Analyse fonctionnelle

Master de Mathématiques

Table des matières

Chapitre 1

Opérateurs entre espaces de Hilbert

1. Micro-rappels et notations

1.1. Opérateurs.

• Si X et Y sont des espaces vectoriels normés sur $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$, on note $\mathcal{L}(X,Y)$ l'espace des applications linéaires continues $T:X\to Y$, muni de sa norme naturelle :

$$\|T\|=\sup\left\{\frac{\|T(x)\|}{\|x\|};\ x\neq 0\right\}=\sup\left\{\|T(x)\|;\ \|x\|=1\right\}=\sup\left\{\|T(x)\|;\ \|x\|\leqslant 1\right\};$$

et on pose $\mathcal{L}(X):=\mathcal{L}(X,X)$. Un élément de $\mathcal{L}(X,Y)$ s'appelle un **opérateur borné** de X dans Y.

• Mode d'emploi : si $T:X\to Y$ est linéaire et $C\in\mathbb{R}^+$, alors

$$T \text{ est continue et } \|T\|\leqslant C \iff \forall x\in X\ :\ \|T(x)\|\leqslant C\,\|x\|.$$

• On écrit en général Tx au lieu de $T(x)$, et AB au lieu de $A\circ B$.
• On a toujours $\|AB\|\leqslant\|A\|\,\|B\|$.
• l'opérateur“identité” de X dans X se note I_X, ou simplement I.
• Un opérateur $T\in\mathcal{L}(X,Y)$ est **inversible** si T est bijectif et si $T^{-1}:Y\to X$ est continue. Si X et Y sont des espaces de Banach, tout opérateur bijectif $T\in\mathcal{L}(X,Y)$ est automatiquement inversible : c'est le *Théorème d'isomorphisme de Banach*, qu'on verra au Chapitre 3.

• Un opérateur $T\in\mathcal{L}(X,Y)$ est un **plongement** s'il existe une constante $c>0$ telle que

$$\forall x\in X\ :\ \|Tx\|\geqslant c\,\|x\|.$$

Autrement dit : T est un plongement si et seulement si T est injectif et “T^{-1}” : $\mathrm{Im}(T)\to X$ est continu.

LEMME 1.1. *Soient X et Y des espces de Banach, et soit $T\in\mathcal{L}(X,Y)$. Alors T est un plongement si et seulement si T est injectif et* $\mathrm{Im}(T)$ *est fermé dans Y.*

Démonstration. Supposons que T soit un plongement avec “témoin” $c>0$. Alors T est injectif. Soit (y_n) une suite dans $\mathrm{Im}(T)$ convergeant vers $y\in Y$, et écrivons $y_n=Tx_n$. On a $\|y_q-y_p\|=\|T(x_q-x_p)\|\geqslant\|x_q-x_p\|$ pour tous $p,q\in\mathbb{N}$; donc la suite (x_n) est de Cauchy dans X puisque (y_n) est de Cauchy dans Y. Comme X est complet, la suite (x_n) converge, $x_n\to x\in X$. Alors $y=Tx$ par continuité de T ; donc $y\in\mathrm{Im}(T)$. Ainsi $\mathrm{Im}(T)$ est fermé dans Y.
Inversement, supposons que T soit injectif et à image fermée. Alors $\mathrm{Im}(T)$ est un espace de Banach puisque Y en est un ; et $T:X\to\mathrm{Im}(T)$ est bijectif. D'après le Théorème d'isomorphisme de Banach, l'application linéaire $T^{-1}:\mathrm{Im}(T)\to X$ est continue ; et donc T est un plongement. $\square$

1.2. Espaces de Hilbert. On suppose connus les faits de base concernant les espaces de Hilbert, c'est-à-dire le théorème de projection et ses conséquences habituelles. Une précision importante :

Convention. Si H est un espace de Hilbert complexe, le produit scalaire $\langle x, y\rangle$ est linéaire par rapport à x et antilinéaire par rapport à y.

1.2.1. *Espaces ℓ^2 et L^2.*

• Si I est un ensemble quelconque (non vide), on pose

$$\ell^2(I) := \left\{ x = (x_i)_{i\in I} \in \mathbb{C}^I;\ \|x\|_2^2 := \sum_{i\in I} |x_i|^2 < \infty \right\}.$$

• En raison de l'existence de bases hilbertiennes, on sait que tout espace de Hilbert est isométrique à un espace $\ell^2(I)$; et que tout espace de Hilbert séparable de dimension infinie est isométrique à $\ell^2(\mathbb{N})$.

• Si $(\Omega, \mathfrak{B}, m)$ est un espace mesuré, on note $L^2(\Omega, m)$ l'espace L^2 associé. Le produit scalaire sur $L^2(\Omega, m)$ est donné par

$$\langle f, g\rangle_{L^2(\Omega,m)} = \int_\Omega f(x)\,\overline{g(x)}\,dm(x).$$

• On a $\ell^2(I) = L^2(I, m)$, où m est la mesure de comptage sur I.

• Si I est un intervalle de $\mathbb{R}$, on pose $L^2(I) := L^2(I, dx)$.

• Il est bien connu, et non trivial, que l'espace $L^2(I)$ est *séparable.*

• On pose $L^2(\mathbb{T}) := L^2(\mathbb{T}, m)$, où m est la mesure de Lebesgue normalisée sur le cercle $\mathbb{T}$. La mesure m est définie par

$$m(A) := \frac{1}{2\pi}\lambda_1\big(\{\theta \in [0, 2\pi[;\ e^{i\theta} \in A\}\big),$$

où λ_1 est la mesure de Lebesgue sur $\mathbb{R}$. Si f est une fonction borélienne sur $\mathbb{T}$, positive ou intégrable par rapport à m, alors

$$\int_{\mathbb{T}} f\,dm = \int_0^{2\pi} f(e^{i\theta})\,\frac{d\theta}{2\pi}\cdot$$

• Si I est un ensemble quelconque et si H est un espace de Hilbert, on pose

$$\ell^2(I, H) := \left\{ x = (x_i)_{i\in I} \in H^I;\ \|x\|^2 := \sum_{i\in I} \|x_i\|^2 < \infty \right\}.$$

1.2.2. *Produit scalaire et norme.* Le lemme suivant sera constamment utilisé.

Lemme 1.2. *Soit H un espace de Hilbert. Pour tout $z \in H$, on a*

$$\|z\| = \sup\big\{|\langle z, y\rangle|;\ \|y\| = 1\big\} = \max\big\{|\langle z, y\rangle|;\ \|y\| = 1\big\}.$$

Démonstration. Le sup en question est $\leqslant \|z\|$ par l'inégalité de Cauchy-Schwarz ; puis on prend $y := \frac{z}{\|z\|}$ si $z \neq 0$. □

Corollaire 1.3. *Si H_1, H_2 sont des espaces de Hilbert (sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) et si $T \in \mathcal{L}(H_1, H_2)$, alors*

$$\|T\| = \sup\Big\{|\langle Tx, y\rangle|;\ \|x\| = 1 = \|y\|\Big\}.$$

Démonstration. C'est évident par le lemme :

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\|=1} \sup_{\|y\|=1} |\langle Tx, y\rangle|.$$

□

Corollaire 1.4. *Si $T : H_1 \to H_2$ est linéaire et $C \in \mathbb{R}^+$, alors*

$$T \textit{ continue et } \|T\| \leqslant C \iff \forall (x,y) \in H_1 \times H_2 \ : \ |\langle Tx, y\rangle| \leqslant C\,\|x\|\,\|y\|.$$

Le lemme montre en particulier que si H est un espace de Hilbert réel ou complexe et si $T \in \mathcal{L}(H)$, alors

$$T = 0 \iff \forall x, y \in H \ : \ \langle Tx, y\rangle = 0.$$

Dans le cas d'un espace de Hilbert complexe, ceci peut être amélioré :

Lemme 1.5. *Soit H un espace de Hilbert* complexe*, et soit $T : H \to H$ une application linéaire. Si on a $\langle Tu, u\rangle = 0$ pour tout $u \in H$, alors $T = 0$.*

Démonstration. Si $x, y \in H$, alors

$$\begin{aligned} 0 = \langle T(x+y), x+y\rangle &= \langle Tx, x\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Ty, y\rangle \\ &= \langle Tx, y\rangle + \langle Ty, x\rangle; \end{aligned}$$

donc $\langle Tx, y\rangle = -\langle Ty, x\rangle$. En remplaçant y par iy, on obtient $-i\langle Tx, y\rangle = -i\langle Ty, x\rangle$, *i.e.* $\langle Tx, y\rangle = \langle Ty, x\rangle$. Ainsi, on a en fait $\langle Tx, y\rangle = 0$ pour tous $x, y \in H$, et donc $T = 0$. □

Remarque. Le résultat est faux sur un espace de Hilbert réel : il suffit de considérer $H := \mathbb{R}^2$ et de prendre pour T une rotation d'angle $\pm\frac{\pi}{2}$.

1.3. Adjoint d'un opérateur hilbertien.

Rappel. Soient H_1, H_2 des espaces de Hilbert. Si $T \in \mathcal{L}(H_1, H_2)$, il existe un unique opérateur $T^* \in \mathcal{L}(H_2, H_1)$ tel que

$$\forall (x,y) \in H_1 \times H_2 \ : \ \langle Tx, y\rangle = \langle x, T^*y\rangle.$$

On dit que T^* est l'**adjoint** (hilbertien) de l'opérateur T.

Exemple 1. Dans le cas "trivial" $H_1 = \mathbb{C} = H_2$, tout opérateur $T \in \mathcal{L}(\mathbb{C})$ est de la forme $T = \lambda I$ pour un certain $\lambda \in \mathbb{C}$, et on a alors $T^* = \bar{\lambda} I$. Autrement dit, dans ce cas trivial l'opération de passage à l'adjoint s'identifie à la conjugaison complexe. Dans le cas d'un espace de Hilbert complexe H général, il est utile de garder à l'esprit que l'opération de passage à l'adjoint sur $\mathcal{L}(H)$ est l'analogue de la conjugaison complexe.

Exemple 2. Si $T \in \mathcal{L}(\mathbb{K}^d)$ a pour matrice $A = \big(a_{i,j}\big)_{1\leqslant i,j\leqslant d}$, alors $T^* \in \mathcal{L}(\mathbb{K}^d)$ a pour matrice $A^* := \big(\overline{a_{j,i}}\big)_{1\leqslant i,j\leqslant d}$.

Propriétés formelles. *L'opération de "passage à l'adjoint" possède les propriétés suivantes.*

(i) $(A+B)^* = A^* + B^*$.

(ii) $(\lambda A)^* = \bar{\lambda} A^*$ *si* $\lambda \in \mathbb{K}$.

(iii) $(AB)^* = B^* A^*$.

(iv) $I^* = I$.

(v) $(T^*)^* = T$.

(vi) $\|T^*\| = \|T\|$.

Démonstration. Les points (i), ... , (v) sont des exos très faciles. Pour (vi), on écrit

$$\|T^*\| = \sup\big\{|\langle T^*y, x\rangle|;\ \|y\| = 1 = \|x\|\big\} = \sup\big\{|\langle y, Tx\rangle|;\ \|y\| = 1 = \|x\|\big\} = \|T\|.$$

□

REMARQUE. De (iii), (iv) et (v) on déduit qu'un opérateur $T \in \mathcal{L}(H_1, H_2)$ est inversible si et seulement si T^* est inversible, et qu'on a alors $(T^{-1})^* = (T^*)^{-1}$.

LEMME 1.6. *Si $T \in \mathcal{L}(H_1, H_2)$, alors $\|T^*T\| = \|T\|^2$.*

Démonstration. D'abord, $\|T^*T\| \leqslant \|T\|\,\|T^*\| = \|T\|^2$. Ensuite, $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leqslant \|T^*Tx\|\,\|x\| \leqslant \|T^*T\|\,\|x\|^2$ pour tout $x \in H$, donc $\|T\|^2 \leqslant \|T^*T\|$. □

PROPOSITION 1.7. *Si $T \in \mathcal{L}(H_1, H_2)$, alors $\ker(T^*) = \mathrm{Im}(T)^\perp$ et $\overline{\mathrm{Im}(T^*)} = \ker(T)^\perp$.*

Démonstration. Par définition, si $y \in H_2$, alors

$$\begin{aligned} y \in \ker(T^*) &\iff \forall x \in H_1\ :\ \langle T^*y, x\rangle = 0 \\ &\iff \forall x \in H_1\ :\ \langle y, Tx\rangle = 0 \iff y \in \mathrm{Im}(T)^\perp. \end{aligned}$$

Donc $\ker(T^*) = \mathrm{Im}(T)^\perp$; et donc $\ker(T)^\perp = \big(\mathrm{Im}(T^*)^\perp\big)^\perp = \overline{\mathrm{Im}(T^*)}$ en appliquant ceci à T^*. □

COROLLAIRE 1.8. *Pour tout opérateur $T \in \mathcal{L}(H_1, H_2)$, on a $H_1 = \ker(T) \overset{\perp}{\oplus} \overline{\mathrm{Im}(T^*)}$ et $H_2 = \ker(T^*) \overset{\perp}{\oplus} \overline{\mathrm{Im}(T)}$.*

COROLLAIRE 1.9. *Si $T \in \mathcal{L}(H_1, H_2)$ alors (T^* injectif) $\iff$ (T à image dense), et (T injectif) $\iff$ (T^* à image dense).*

COROLLAIRE 1.10. *Pour $T \in \mathcal{L}(H_1, H_2)$, les propriétés suivantes sont équivalentes :*

(i) *T est inversible ;*

(ii) *T et T^* sont des plongements.*

Démonstration. Si T est inversible alors T est certainement un plongement, et donc T^* est lui aussi un plongement car T^* est inversible. Inversement, supposons que T et T^* soient des plongements. Alors T est injectif, $\mathrm{Im}(T)$ est fermé dans H_2 car H_1 est complet, et $\mathrm{Im}(T)$ est dense dans H_2 car T^* est injectif ; donc T est en fait bijectif. Comme T est un plongement, T^{-1} est continu ; donc T est inversible. □

1.4. Matrices infinies. Dans cette section, H est l'espace de Hilbert $\ell^2 := \ell^2(\mathbb{N})$, et on note $(e_i)_{i\in\mathbb{N}}$ la base canonique de ℓ^2. Si $x = (x_0, x_1, \dots) \in \ell^2$, on a donc $x_i = \langle x, e_i\rangle$ pour tout $i \in \mathbb{N}$; et $x = \sum_{i=0}^\infty x_i e_i$, où la série converge dans ℓ^2.

DÉFINITION 1.11. *Si $T \in \mathcal{L}(\ell^2)$, la **matrice de T relativement à** (e_i) est la matrice infinie $A := \big(\langle Te_j, e_i\rangle\big)_{i,j\in\mathbb{N}}$.*

REMARQUE 1.12. Un opérateur $T \in \mathcal{L}(\ell^2)$ est entièrement déterminé par sa matrice.

Démonstration. Si T_1, T_2 ont la même matrice, alors $\langle T_1 u, v\rangle = \langle T_2 u, v\rangle$ pour tous $u, v \in c_{00} := \operatorname{vect}\{e_i;\ i \in \mathbb{N}\}$ par linéarité, donc $\langle T_1 x, y\rangle = \langle T_2 x, y\rangle$ pour tous $x, y \in \ell^2$ par continuité car c_{00} est dense dans ℓ^2, et donc $T_1 = T_2$. □

Remarque 1.13. Si $T \in \mathcal{L}(\ell^2)$ a pour matrice $A = (a_{i,j})_{i,j\in\mathbb{N}}$ relativement à (e_i), alors T^* a pour matrice $A^* := \big(\overline{a_{j,i}}\big)_{i,j\in\mathbb{N}}$. En particulier, on a

$$\forall j \in \mathbb{N}\ :\ \sum_{i=0}^{\infty} |a_{i,j}|^2 = \|Te_j\|^2 < \infty \qquad \text{et} \qquad \forall i \in \mathbb{N}\ :\ \sum_{j=0}^{\infty} |a_{i,j}|^2 = \|T^* e_i\|^2 < \infty.$$

Démonstration. Exo. □

Remarque 1.14. Si $T \in \mathcal{L}(\ell^2)$ a pour matrice $A = \big(a_{i,j}\big)$ relativement à (e_i), alors

$$\forall x \in \ell^2\ \ \forall i \in \mathbb{N}\ :\ (Tx)_i = \sum_{j=0}^{\infty} a_{i,j} x_j, \qquad \text{où la série converge } \textit{absolument.}$$

Démonstration. On a $x = \sum_{j=0}^{\infty} x_j\, e_j$ où la série converge dans ℓ^2 ; donc, par continuité de T et du produit scalaire, on peut écrire

$$(Tx)_i = \langle Tx, e_i\rangle = \sum_{j=0}^{\infty} x_j \,\langle Te_j, e_i\rangle = \sum_{j=0}^{\infty} a_{i,j} x_j .$$

La série converge absolument car $\sum_{j=0}^{\infty} |a_{i,j}|^2 < \infty$ et $\sum_{j=0}^{\infty} |x_j|^2 < \infty$. □

2. Quelques exemples

2.1. Projections orthogonales.

Rappel. Soit H un espace de Hilbert, et soit E un sous-espace fermé de H. Pour tout $x \in H$, il existe un unique $z \in E$ tel que $(x - z) \perp E$; et ce z est caractérisé par le fait que $z \in E$ et $\|x - z\| = \operatorname{dist}(x, E)$. Notation : $z = p_E(x)$. On dit que $p_E(x)$ est le **projeté orthogonal** de x sur E.

Propriétés. *Soit $p := p_E : H \to H$. Alors p possède les propriétés suivantes.*

- *p est linéaire et $p^2 = p$.*
- *p est continue et $\|p\| = 1$.*
- *$\forall x, y \in H\ :\ \langle px, y\rangle = \langle px, py\rangle = \langle x, py\rangle$; donc $p^* = p$.*
- *$\forall x \in H\ :\ \langle px, x\rangle = \|px\|^2 \geqslant 0$.*

Démonstration. Exo. □

Lemme 2.1. *Soit $p \in \mathcal{L}(H)$. Alors p est une projection orthogonale* si et seulement si *$p^2 = p$ et $p^* = p$.*

Démonstration. Une implication a déjà été vue. Inversement, supposons que $p^* = p = p^2$. Comme $p^2 = p$, on a $\operatorname{Im}(p) = \ker(I - p)$ et $H = \ker(p) \oplus \operatorname{Im}(p)$; et comme $p^* = p$, on a $\ker(p) = \ker(p^*) = \operatorname{Im}(p)^{\perp}$. Donc p est une projection orthogonale. □

Remarque. Soit H un espace de Hilbert, et soit E un sous-espace fermé de H. Soit $\pi : H \to E$ la projection orthogonale de H sur E, *considérée comme un opérateur de H dans E*. Alors π^* est l'injection canonique $E \hookrightarrow H$.

Démonstration. Exo. □

2.2. Opérateurs diagonaux, opérateurs de multiplication.

LEMME 2.2. *Si $a = (a_i)_{i\in\mathbb{N}} \in \ell^\infty(\mathbb{N})$, on définit un opérateur borné $\Delta_a : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ en posant*

$$\Delta_a x := (a_0x_0, a_1x_1, a_2x_2 \dots) \quad \textit{pour } x = (x_0, x_1, x_2, \dots) \in \ell^2(\mathbb{N}).$$

On a $\|\Delta_a\| = \|a\|_\infty = \sup_{i\geqslant 0} |a_i|$. La matrice de Δ_a dans la base canonique de ℓ^2 est la matrice diagonale

$$D_a := \begin{pmatrix} a_0 & & & \\ & a_1 & & \\ & & a_2 & \\ & & & \ddots \end{pmatrix}$$

Démonstration. Exo. □

LEMME 2.3. *Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré. Si $\phi \in L^\infty(\Omega, m)$, on définit un opérateur borné $M_\phi : L^2(\Omega, m) \to L^2(\Omega, m)$ en posant $M_\phi f := \phi f$ pour toute $f \in L^2(\Omega, m)$. On a $\|M_\phi\| \leqslant \|\phi\|_\infty$, et $\|M_\phi\| = \|\phi\|_\infty$ si la mesure m est* sigma-finie.

Démonstration. Le fait que M_ϕ soit un opérateur borné avec $\|M_\phi\| \leqslant \|\phi\|_\infty$ est laissé en exo.
Montrons que $\|M_\phi\| \geqslant \|\phi\|_\infty$ si m est sigma-finie. Soit $\alpha < \|\phi\|_\infty$ quelconque. Par définition de la norme $\|\cdot\|_\infty$, l'ensemble $A := \{|\phi| \geqslant \alpha\}$ vérifie $m(A) > 0$. Comme la mesure m est sigma-finie, on peut trouver $E \subseteq A$ tel que $0 < m(E) < \infty$. Alors $f := \mathbf{1}_E \in L^2\backslash\{0\}$, et $\|M_\phi f\|_2^2 = \int_E |\phi|^2\, dm \geqslant \alpha^2 m(E) = \alpha^2 \|f\|_2^2$. Ainsi, on a $\|M_\phi\| \geqslant \alpha$ pour tout $\alpha < \|\phi\|_\infty$. □

REMARQUE. Les opérateurs diagonaux Δ_a sont des cas particuliers d'opérateurs de multiplication : si on prend $\Omega = \mathbb{N}$ avec pour m la mesure de comptage, alors $L^2(\Omega, m) = \ell^2(\mathbb{N})$ et $M_\phi = \Delta_a$ où $a := (\phi(n))_{n\in\mathbb{N}}$.

EXERCICE. On a $M_\phi M_\psi = M_{\phi\psi} = M_\psi M_\phi$ et $(M_\phi)^* = M_{\overline{\phi}}$ pour toutes $\phi, \psi \in L^\infty$.

2.3. Shifts.

DÉFINITION 2.4. *Le "forward shift" sur $\ell^2(\mathbb{N})$ est l'opérateur $S : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ défini par*

$$S(x_0, x_1, x_2, \dots) := (0, x_0, x_1, x_2, \dots)$$

Le "backward shift" B est défini par

$$B(x_0, x_1, x_2, \dots) := (x_1, x_2, \dots)$$

On a ainsi

$$Se_i = e_{i+1} \qquad \textit{pour tout } i \in \mathbb{N},$$
$$Be_0 = 0 \qquad \text{et} \qquad Be_i = e_{i-1} \quad \textit{pour } i \geqslant 1.$$

Matriciellement :

$$S \sim \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & 1 & \ddots & \\ & & & \ddots & \ddots \end{pmatrix} \qquad \text{et} \qquad B \sim \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & \ddots & \ddots \\ & & & & \ddots \end{pmatrix}$$

LEMME 2.5. *Les opérateurs S et B possèdent les propriétés suivantes.*

- *S est une isométrie, i.e. $\|Sx\| = \|x\|$ pour tout $x \in \ell^2$.*
- *$\|B\| = 1$.*
- *$S^* = B$.*
- *$BS = I$.*
- *B est surjectif et $\ker(B) = [e_0]$.*
- *S est injectif et $\mathrm{Im}(S) = [e_n;\ n \geqslant 1] = [e_0]^\perp$.*

Démonstration. Exo. □

REMARQUE. Précisons une notation qu'on aura l'occasion de ré-utiliser : si $(x_i)_{i\in I}$ est une famille d'éléments d'un espace vectoriel normé X, on note $[x_i;\ i \in I]$ le sous-espace vectoriel *fermé* de X engendré par les x_i :

$$[x_i;\ i \in I] := \overline{\mathrm{vect}\,\{x_i;\ i \in I\}}.$$

Pour une famille *finie* $(x_1, \dots, x_N)$, on a $[x_1, \dots, x_N] = \mathrm{vect}\,(x_1, \dots, x_N)$ car tout sous-espace vectoriel de dimension finie est fermé dans X.

DÉFINITION 2.6. *Le "forward shift" S et le "backward shift" B sur $\ell^2(\mathbb{Z})$ sont les opérateurs sur $\ell^2(\mathbb{Z})$ définis par*

$$Se_i := e_{i+1} \qquad \text{et} \qquad Be_i := e_{i-1} \qquad \textit{pour tout } i \in \mathbb{Z}.$$

Autrement dit, si $x = (x_i)_{i\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$, alors

$$Sx = (x_{i-1})_{i\in\mathbb{Z}} \qquad \text{et} \qquad Bx = (x_{i+1})_{i\in\mathbb{Z}}.$$

LEMME 2.7. *Les shifts B et S sur $\ell^2(\mathbb{Z})$ sont des isométries bijectives inverses l'une de l'autre.*

Démonstration. Exo. □

REMARQUE. Plus généralement, si H est un espace de Hilbert quelconque, on peut définir les shifts canoniques B et S sur $\ell^2(\mathbb{N}, H)$ et $\ell^2(\mathbb{Z}, H)$.

2.4. Opérateurs à noyaux. Dans cette section $(\Omega, \mathfrak{B}, m)$ est un espace mesuré, et on suppose que la mesure m est *sigma-finie* pour pouvoir appliquer le Théorème de Fubini. .

DÉFINITION 2.8. *Un **noyau sur** Ω est une fonction mesurable $K : \Omega \times \Omega \to \mathbb{C}$.*

NOTATION. Soit $K : \Omega \times \Omega \to \mathbb{C}$ un noyau sur Ω. Si $f : \Omega \to \mathbb{C}$ est une fonction mesurable et si $x \in \Omega$ est tel que la fonction $y \mapsto K(x, y)f(y)$ est intégrable sur Ω, on pose

$$T_K f(x) := \int_\Omega K(x, y)f(y)\, dm(y).$$

DÉFINITION 2.9. *On dit qu'un noyau $K : \Omega\times\Omega \to \mathbb{C}$ **définit un opérateur borné sur** $L^2(\Omega, m)$ si*

- *pour toute $f \in L^2(\Omega, m)$, la fonction $T_K f$ est bien définie m-presque partout ;*
- *$T_K f \in L^2(\Omega, m)$ pour toute $f \in L^2(\Omega, m)$;*

- *l'application linéaire $T_K : L^2(\Omega, m) \to L^2(\Omega, m)$ est continue.*

Remarque 2.10. Si $\Omega = \mathbb{N}$ et si m est la mesure de comptage, alors $L^2(\Omega, m) = \ell^2(\mathbb{N})$ et “presque partout” veut dire “partout”. Donc, un noyau $K : \mathbb{N} \times \mathbb{N} \to \mathbb{C}$ définit un opérateur borné sur $L^2(\mathbb{N}, m)$ exactement quand la matrice $A := \big(K(i,j)\big)_{i,j \in \mathbb{N}}$ est la matrice d'un opérateur borné sur ℓ^2 (*cf* la Remarque 1.14.)

Définition 2.11. *On dit qu'un noyau $K : \Omega \times \Omega \to \mathbb{C}$ est **absolument L^2 - borné** s'il existe une constante $M < \infty$ telle que*

$$\forall f \in L^2(\Omega, m)\ :\ \int_\Omega \left(\int_\Omega |K(x,y) f(y)|\, dm(y) \right)^2 dm(x) \leqslant M \int_\Omega |f(y)|^2 dm(y).$$

Exemple 2.12. Si $K \in L^2(\Omega \times \Omega, m \otimes m)$, alors K est absolument L^2 - borné avec constante $M := \|K\|_{L^2}^2$.

Démonstration. Application “mécanique” de l'inégalité de Cauchy-Schwarz et du Théorème de Fubini (exo). □

Lemme 2.13. *Si $K : \Omega \times \Omega \to \mathbb{C}$ est un noyau absolument L^2 - borné, avec constante M, alors K définit un opérateur borné sur $L^2(\Omega, m)$ avec $\|T_K\| \leqslant \sqrt{M}$.*

Démonstration. Soit $f \in L^2(\Omega, m)$. Pour $x \in \Omega$, posons

$$\Phi(x) := \int_\Omega |K(x,y) f(y)|\, dm(y).$$

Comme la mesure m est sigma-finie, Φ est une fonction *mesurable* à valeurs dans $[0, \infty]$ (c'est un petit bout de l'énoncé du Théorème de Fubini). Par hypothèse, on a $\int_\Omega \Phi(x)^2\, dm(x) \leqslant M\, \|f\|_2^2 < \infty$. Donc $\Phi(x) < \infty$ pp, et donc $T_K f(x)$ est bien défini presque partout. Par définition de Φ, on a $|T_K f(x)| \leqslant \Phi(x)$ pp, donc $T_K f \in L^2$ et $\|T_K f\|_2^2 \leqslant \|\Phi\|_2^2 \leqslant M\, \|f\|_2^2$. □

Corollaire 2.14. *Si $K \in L^2(\Omega \times \Omega, m \otimes m)$, alors K définit un opérateur borné sur $L^2(\Omega, m)$ avec $\|T_K\| \leqslant \|K\|_{L^2}$. En particulier, si $A = \big(a_{i,j}\big)_{i,j \in \mathbb{N}}$ est une matrice infinie telle que $\sum_{i,j=0}^\infty |a_{i,j}|^2 < \infty$, alors A est la matrice d'un opérateur borné sur $\ell^2(\mathbb{N})$.*

Remarque 2.15. Si $K : \Omega \to \Omega \to \mathbb{C}$ est un noyau absolument L^2 - borné, et si le “noyau adjoint” $K^*(x,y) := \overline{K(y,x)}$ est lui aussi absolument L^2 - borné, alors $(T_K)^* = T_{K^*}$.

Démonstration. Exo. □

Théorème 2.16. (test de Schur)
Soit $K : \Omega \times \Omega \to \mathbb{C}$ un noyau. On suppose qu'il existe deux fonctions mesurables strictement positives $w_1, w_2 : \Omega \to \mathbb{R}$ et deux constantes $C_1, C_2 < \infty$ telles que

- $\forall x \in \Omega\ :\ \int_\Omega |K(x,y)| w_1(y)\, dm(y) \leqslant C_1\, w_2(y)$ *;*
- $\forall y \in \Omega\ :\ \int_\Omega |K(x,y)|\ w_2(x)\, dm(x) \leqslant C_2\, w_1(y).$

Alors K est absolument L^2-borné avec constante $M := C_1C_2$, et donc K définit un opérateur borné sur $L^2(\Omega, m)$ avec $\|T_K\| \leqslant \sqrt{C_1C_2}$.

Démonstration. Soit $f \in L^2(\mathbb{R})$. Pour tout $x \in \Omega$, on a

$$\begin{aligned}\int_\Omega |K(x,y)f(y)|\, dm(y) &= \int_\Omega \big(|K(x,y)|w_1(y)\big)^{1/2} \times \left(\frac{|K(x,y)|}{w_1(y)}\right)^{1/2} |f(y)|\, dm(y)\\ &\leqslant \left(\int_\Omega |K(x,y)|w_1(y)\, dm(y)\right)^{1/2} \left(\int_\Omega \frac{|K(x,y)|}{w_1(y)}\, |f(y)|^2 dm(y)\right)^{1/2}\\ &\leqslant \sqrt{C_1\, w_2(x)} \left(\int_\Omega \frac{|K(x,y)|}{w_1(y)}\, |f(y)|^2 dm(y)\right)^{1/2}.\end{aligned}$$

Donc (par Fubini)

$$\begin{aligned}&\int_\Omega \left(\int_\Omega |K(x,y)f(y)|\, dm(y)\right)^2 dm(x)\\ &\qquad\leqslant C_1 \int_\Omega w_2(x) \left(\int_\Omega \frac{|K(x,y)|}{w_1(y)}\, |f(y)|^2 dm(y)\right) dm(x)\\ &\qquad= C_1 \int_\Omega |f(y)|^2 \times \frac{1}{w_1(y)} \left(\int_\Omega |K(x,y)|w_2(x)\, dm(x)\right) dm(y)\\ &\qquad\leqslant C_1C_2 \int_\Omega |f(y)|^2 dm(y).\end{aligned}$$

□

Corollaire 2.17. *S'il existe deux constantes C_1 et C_2 telles que*

$$\forall x \in \Omega \ : \ \int_\Omega |K(x,y)|\, dm(y) \leqslant C_1 \qquad \text{et} \qquad \forall y \in \Omega \ : \ \int_\Omega |K(x,y)|\, dm(x) \leqslant C_2,$$

alors K est absolument L^2-borné avec constante C_1C_2, et donc définit un opérateur borné sur $L^2(\Omega, m)$ avec $\|T_K\| \leqslant \sqrt{C_1C_2}$.

Démonstration. On applique le test de Schur avec $w_1(t) := 1 = w_2(t)$. □

Exemple 2.18. Soit $k \in L^1(\mathbb{R})$ et soit $K : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ définie par $K(x,y) := k(x-y)$. Si $f : \Omega \to \mathbb{C}$ est mesurable, alors

$$T_K f(x) = \int_{\mathbb{R}} k(x-y)f(y)\, dy = k * f(x)$$

en tout point $x \in \mathbb{R}$ où ceci a un sens. On a $\forall x \in \mathbb{R} \ : \ \int_{\mathbb{R}} |k(x-y)|dy = \|k\|_1$ et $\forall y \in \mathbb{R} \ : \ \int_{\mathbb{R}} |k(y-x)|\, dx = \|k\|_1$. Donc K définit un opérateur borné sur L^2 avec $\|T_K\| \leqslant \sqrt{\|k\|_1 \times \|k\|_1} = \|k\|_1$. Autrement dit : si $f \in L^2(\mathbb{R})$ alors $k * f$ est bien définie presque partout et appartient à $L^2(\mathbb{R})$, avec

$$\|k * f\|_2 \leqslant \|k\|_1\, \|f\|_2.$$

Remarque. Le noyau $K(x,y) = k(x-y)$ n'appartient pas à $L^2(\mathbb{R} \times \mathbb{R})$.

Exercice. Imaginer un “test de Schur L^p” pour $1 < p < \infty$, et vérifier qu'il s'applique pour la convolution L^1-L^p.

EXEMPLE 2.19. Soit $A = \big(a_{i,j}\big)_{i,j\in\mathbb{N}}$ une matrice infinie. S'il existe $r < 1$ tel que

$$\forall i, j \in \mathbb{N} \ : \ |a_{i,j}| \leqslant r^{|j-i|},$$

alors A définit un opérateur borné sur $\ell^2(\mathbb{N})$ avec $\|A\| \leqslant \frac{1+r}{1-r}\cdot$

Démonstration. Il suffit de montrer que la matrice $A_r := \big(r^{|j-i|}\big)_{i,j\in\mathbb{N}}$ est absolument L^2-bornée avec constante $\big(\frac{1+r}{1-r}\big)^2$. Comme A_r est symétrique, il suffit donc de vérifier que

$$\forall i \in \mathbb{N} \ : \ \sum_{j=0}^{\infty} r^{|j-i|} \leqslant \frac{1+r}{1-r}\cdot$$

Il faut juste accepter de faire le calcul :

$$\begin{aligned}\sum_{j=0}^{\infty} r^{|j-i|} &= \sum_{0\leqslant j<i} r^{i-j} + \sum_{j=i}^{\infty} r^{j-i}\\ &= \sum_{k=1}^{i} r^k + \sum_{k=0}^{\infty} r^k\\ &= r\,\frac{1-r^i}{1-r} + \frac{1}{1-r}\\ &\leqslant \frac{1+r}{1-r}\cdot\end{aligned}$$

□

3. Classes importantes d'opérateurs

3.1. Opérateurs auto-adjoints.

DÉFINITION 3.1. *Un opérateur $T \in \mathcal{L}(H)$ est dit* ***auto-adjoint*** *si $T^* = T$; autrement dit, si*

$$\forall x, y \in H \ : \ \langle Tx, y\rangle = \langle x, Ty\rangle.$$

EXEMPLES.

(1) Soit M_ϕ un opérateur de multiplication sur $L^2(\Omega, m)$. Alors M_ϕ est auto-adjoint si et seulement si la fonction ϕ est pp à valeurs réelles.
(2) Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré, et soit $K : \Omega \times \Omega$ un noyau absolument L^2-borné tel que le noyau adjoint $K^*(x, y) := \overline{K(y, x)}$ est lui aussi absolument L^2-borné. Alors T_K est autoadjoint si et seulement si $K(y, x) = \overline{K(x, y)}$ pp sur $\Omega \times \Omega$.

Remarque. Si $T \in \mathcal{L}(H)$ est auto-adjoint, alors $\langle Tx, x\rangle \in \mathbb{R}$ pour tout $x \in H$.

Exercice. Soit H un espace de Hilbert complexe. Montrer que si $T \in \mathcal{L}(H)$ est auto-adjoint, alors les valeurs propres éventuelles de T sont réelles.

Comme l'opération de passage à l'adjoint est l'analogue de la conjugaison complexe, les opérateurs auto-adjoints sont aux opérateurs généraux ce que les nombres réels sont aux nombres complexes. Le lemme suivant n'est donc pas surprenant.

LEMME 3.2. *Soit H un espace de Hilbert* complexe. *Si $T \in \mathcal{L}(H)$, alors T peut s'écrire de manière unique sous la forme $T = A + iB$, où A, B sont auto-adjoints. Si de plus $AB = BA$, alors $T^*T = A^2 + B^2$ et $\forall x \in H\ :\ \|Tx\|^2 = \|Ax\|^2 + \|Bx\|^2$.*

Démonstration. Si $T = A + iB$ avec A et B auto-adjoints, alors $T^* = A - iB$; donc on doit avoir $A = \frac{T+T^*}{2}$ et $B = \frac{T-T^*}{2i}\cdot$ Inversement A et B donnés par ces formules conviennent.
Supposons que $AB = BA$. Comme A et B sont auto-adjoints, on a alors $T^*T = (A - iB)(A + iB) = A^2 + B^2 = A^*A + B^*B$, donc $\forall x \in H\ :\ \|Tx\|^2 = \langle T^*Tx, x\rangle = \langle A^*Ax, x\rangle + \langle B^*Bx, x\rangle = \|Ax\|^2 + \|Bx\|^2$. □

COROLLAIRE 3.3. *Soit H un espace de Hilbert* complexe. *Un opérateur $T \in \mathcal{L}(H)$ est auto-adjoint* si et seulement si $\langle Tx, x\rangle \in \mathbb{R}$ *pour tout $x \in H$.*

Démonstration. Supposons que $\langle Tx, x\rangle \in \mathbb{R}$ pour tout $x \in H$, et écrivons $T = A + iB$ avec A et B auto-adjoints. Alors $\langle Tx, x\rangle = \langle Ax, x\rangle + i\langle Bx, x\rangle$ pour tout $x \in H$. Comme A et B sont auto-adjoint, on a $\langle Ax, x\rangle \in \mathbb{R}$ et $\langle Bx, x\rangle \in \mathbb{R}$. On a donc nécessairement $\langle Bx, x\rangle = 0$ pour tout $x \in H$. Donc $B = 0$ d'après le Lemme 1.5 ; et donc $T = A$ est auto-adjoint. □

PROPOSITION 3.4. *Si $T \in \mathcal{L}(H)$ est auto-adjoint, alors*

$$\|T\| = \sup \left\{ |\langle Tx, x\rangle|;\ \|x\| = 1 \right\}.$$

Démonstration. Posons $M := \sup\left\{|\langle Tx, x\rangle;\ \|x\| = 1\right\}$. On a évidemment $M \leqslant \|T\|$. Pour l'inégalité inverse, il suffit de montrer qu'on a $|\langle Tx, y\rangle| \leqslant M$ pour tous $x, y \in H$ vérifiant $\|x\| = 1 = \|y\|$.
Comme T est auto-adjoint, on a (exo)

$$\langle T(x + y), x + y\rangle = \langle Tx, x\rangle + 2\mathrm{Re}\,\langle Tx, y\rangle + \langle Ty, y\rangle;$$

et de même

$$\langle T(x - y), x - y\rangle = \langle Tx, x\rangle - 2\mathrm{Re}\,\langle Tx, y\rangle + \langle Ty, y\rangle.$$

En retranchant ces deux identités, on obtient

$$4\mathrm{Re}\,\langle Tx, y\rangle = \langle T(x + y), x + y\rangle - \langle T(x - y), x - y\rangle.$$

Donc

$$\mathrm{Re}\,\langle Tx, y\rangle \leqslant \frac{1}{4}\left(M\,\|x + y\|^2 + M\,\|x - y\|^2\right) = \frac{M}{2}\left(\|x\|^2 + \|y\|^2\right) = M.$$

Maintenant, soit $\omega \in \mathbb{K}$ tel que $|\omega| = 1$ et $|\langle Tx, y\rangle| = \omega\,\langle Tx, y\rangle$. Comme $\|\omega x\| = 1 = \|y\|$, on a

$$|\langle Tx, y\rangle| = \langle T(\omega x), y\rangle = \mathrm{Re}\,\langle T(\omega x), y\rangle \leqslant M.$$

□

COROLLAIRE 3.5. *Si $A, B \in \mathcal{L}(H)$ sont auto-adjoints et si on a $\langle Ax, x\rangle = \langle Bx, x\rangle$ pour tout $x \in H$, alors $A = B$.*

Démonstration. L'opérateur $T := B - A$ est auto-adjoint et $\langle Tx, x\rangle = 0$ pour tout $x \in H$, donc $\|T\| = 0$. □

3.2. Opérateurs positifs.

DÉFINITION 3.6. *Soit H un espace de Hilbert. On dit qu'un opérateur $T \in \mathcal{L}(H)$ est **positif** si T est auto-adjoint et $\forall x \in H\;:\;\langle Tx, x\rangle \geqslant 0$. On écrit alors $T \geqslant 0$.*

EXEMPLES.

(1) Toute projection orthogonale est un opérateur positif.

(2) Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré sigma-finie., et soit $L^2 := L^2(\Omega, \mathfrak{B}, m)$. Un opérateur de multiplication $M_\phi : L^2 \to L^2$ est positif si et seulement si la fonction ϕ est pp à valeurs dans $\mathbb{R}^+$.

Démonstration. Exo. □

Exercice. Montrer que si $T \in \mathcal{L}(H)$ est un opérateur positif, alors les valeurs propres éventuelles de T sont $\geqslant 0$.

LEMME 3.7. *Si $T \in \mathcal{L}(H_1, H_2)$, alors $T^*T \geqslant 0$.*

Démonstration. L'opérateur $T^*T \in \mathcal{L}(H_1)$ est bien défini puisque T va de H_1 dans H_2 et T^* va de H_2 dans H_1 ; il est auto-adjoint car $(T^*T)^* = T^*(T^*)^* = T^*T$; et $\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 \geqslant 0$ pour tout $x \in H_1$. □

LEMME 3.8. *Pour tout $T \in \mathcal{L}(H_1, H_2)$, on a l'équivalence*

$$\|T\| \leqslant 1 \iff I - T^*T \geqslant 0.$$

Démonstration. Comme $I - T^*T$ est auto-adjoint, il s'agit de montrer que $\|T\| \leqslant 1$ si et seulement si $\langle (I - T^*T)x, x\rangle \geqslant 0$ pour tout $x \in H_1$. Mais $\langle (I - T^*T)x, x\rangle = \langle x, x\rangle - \langle T^*Tx, x\rangle = \|x\|^2 - \|Tx\|^2$, donc tout est clair. □

PROPOSITION 3.9. *Si $T \in \mathcal{L}(H)$ est un opérateur positif, alors*

$$\forall x, y \in H\;:\;|\langle Tx, y\rangle| \leqslant \langle Tx, x\rangle^{1/2} \langle Ty, y\rangle^{1/2}.$$

Démonstration. C'est l'inégalité de Cauchy-Schwarz pour la forme hermitienne $\geqslant 0$ définie par $B(x, y) := \langle Tx, y\rangle$. □

COROLLAIRE 3.10. *Si $T \in \mathcal{L}(H)$ est $\geqslant 0$, alors*

$$\forall x \in H\;:\;\|Tx\| \leqslant \|T\|^{1/2} \langle Tx, x\rangle^{1/2}.$$

Démonstration. Par la proposition, on a $|\langle Tx, y\rangle| \leqslant \langle Tx, x\rangle^{1/2} \times \big(\|T\|\,\|y\|^2\big)^{1/2} = \|T\|^{1/2} \langle Tx, x\rangle^{1/2} \|y\|$ pour tout $y \in H$. □

COROLLAIRE 3.11. *Si $T \in \mathcal{L}(H)$ est positif, alors*

$$\ker(T) = \big\{x \in H; \langle Tx, x\rangle = 0\big\}.$$

Démonstration. Micro-exo. □

COROLLAIRE 3.12. *Si $T \in \mathcal{L}(H_1, H_2)$ vérifie $\|T\| \leqslant 1$, alors*

$$\ker(I - T^*T) = \big\{x;\ \|Tx\| = \|x\|\big\}.$$

Démonstration. Exo. □

3.3. Isométries, opérateurs unitaires.

DÉFINITION 3.13. *Soit $T \in \mathcal{L}(H_1, H_2)$. On dit que T est une* ***isométrie*** *si on a $\|Tx\| = \|x\|$ pour tout $x \in H$; et on dit que T est* ***unitaire*** *si T est une isométrie bijective.*

Remarque 1. Si T est unitaire, alors T^{-1} est aussi unitaire ; et si T_1 et T_2 sont unitaires, alors $T_1 T_2$ est unitaire. Donc les opérateurs unitaires forment un *groupe*, qu'on appelle assez logiquement le **groupe unitaire** de H.

Remarque 2. Si $T : H_1 \to H_2$ est unitaire, alors T change toute base hilbertienne de H_1 en une base hilbertienne de H_2. Inversement, si T change *une* base hilbertienne de H_1 en une base hilbertienne de H_2, alors T est unitaire.

Démonstration. Exo. (Pour la 1ère partie, on a besoin du Lemme 3.14.) □

EXEMPLES.

(1) Si $H_1 = \mathbb{K}^d = H_2$, alors "isométrie" $\iff$ "unitaire".

(2) Le forward shift S sur $\ell^2(\mathbb{N})$ est une isométrie, non unitaire.

(3) Les shifts B et S sur $\ell^2(\mathbb{Z})$ sont des opérateurs unitaires.

(4) Un opérateur de multiplication $M_\phi : L^2 \to L^2$ est unitaire si et seulement si $|\phi(x)| = 1$ pp.

(5) Pour $f \in L^2(\mathbb{T})$, notons $\widehat{f}(n)$ les coefficients de Fourier de f :
$$\forall n \in \mathbb{Z} \; : \; \widehat{f}(n) = \int_{\mathbb{T}} f(z) z^{-n} dm(z) = \int_0^{2\pi} f(e^{i\theta})\, e^{-in\theta}\, \frac{d\theta}{2\pi}\cdot$$
La transformation de Fourier $L^2(\mathbb{T}) \ni f \mapsto (\widehat{f}(n))_{n\in\mathbb{Z}}$ est un opérateur unitaire de $L^2(\mathbb{T})$ sur $\ell^2(\mathbb{Z})$.

(6) La transformation de Fourier-Plancherel $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ est un opérateur unitaire.

Démonstration. Exo. □

LEMME 3.14. *Si $T \in \mathcal{L}(H_1, H_2)$ est une isométrie, alors $\langle Tx, Ty\rangle = \langle x, y\rangle$ pour tous $x, y \in H_1$.*

Démonstration. Pour tous $u, v \in H$, on a
$$\|u+v\|^2 - \|u-v\|^2 = 4\mathrm{Re}\,\langle u, v\rangle.$$
Donc, par linéarité,
$$4\mathrm{Re}\,\langle Tx, Ty\rangle = \|T(x+y)\|^2 - \|T(x-y)\|^2 = \|x+y\|^2 - \|x-y\|^2 = 4\mathrm{Re}\,\langle x, y\rangle.$$
Ainsi, on a $\mathrm{Re}\,\langle Tx, Ty\rangle = \mathrm{Re}\,\langle x, y\rangle$ pour tous $x, y \in H$. En changeant y en iy si $\mathbb{K} = \mathbb{C}$, on obtient aussi $\mathrm{Im}\,\langle Tx, Ty\rangle = \mathrm{Im}\,\langle x, y\rangle$, d'où finalement $\langle Tx, Ty\rangle = \langle x, y\rangle$. □

LEMME 3.15. *Soit $T \in \mathcal{L}(H_1, H_2)$.*

(a) *T est une isométrie si et seulement si $T^*T = I$.*

(b) *T est unitaire si et seulement si T est inversible et $T^{-1} = T^*$, i.e. $T^*T = I = TT^*$.*

Démonstration. (a) Si T est une isométrie, on a en particulier $\|T\| = 1$) ; donc $\ker(I - T^*T) = \{x \in H_1;\ \|Tx\| = \|x\|\} = H_1$, et donc $T^*T = I$. Inversement, si $T^*T = I$, alors $\|Tx\|^2 = \langle T^*Tx, x\rangle = \|x\|^2$ pour tout $x \in H_1$.
(b) Si T est unitaire alors T^{-1} est une isométrie, donc $\langle T^{-1}u, T^{-1}v\rangle = \langle u, v\rangle$ pour tous $u, v \in H_2$. Donc, si $x \in H_1$ et $y \in H_2$, alors $\langle Tx, y\rangle = \langle T^{-1}Tx, T^{-1}y\rangle = \langle x, T^{-1}y\rangle$; et donc $T^{-1} = T^*$. Inversement, si $T^*T = I = TT^*$, alors T est une isométrie par (a), et T est inversible. □

COROLLAIRE 3.16. *T est unitaire si et seulement si T et T^* sont des isométries.*

DÉFINITION 3.17. *Soient $T_1 \in \mathcal{L}(H_2)$ et $T_2 \in \mathcal{L}(H_2)$. On dit que T_1 et T_2 sont **unitairement équivalents** s'il existe un opérateur unitaire $U : H_1 \to H_2$ tel que $T_2U = UT_1$, i.e. $T_1 = U^{-1}T_2U = U^*T_2U$.*

EXEMPLE 1. Un opérateur $T \in \mathcal{L}(\mathbb{K}^d)$ est diagonalisable en base orthononormée si et seulement si T est unitairement équivalent à un opérateur diagonal. Si H est un Hilbert séparable de dimension infinie, un opérateur $T \in \mathcal{L}(H)$ est unitairement équivalent à un opérateur diagonal $D \in \mathcal{L}(\ell^2)$ si et seulement si T est diagonalisable en base orthonormée, *i.e.* H possède une base hilbertienne formée de vecteurs propres pour T.

EXEMPLE 2. Soit $\mathbf{z} \in L^\infty(\mathbb{T})$ la fonction $z \mapsto z$. L'opérateur $M_{\mathbf{z}} : L^2(\mathbb{T}) \to L^2(\mathbb{T})$ est unitairement équivalent au forward shift $S : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$.

Démonstration. Si $f \in L^2(\mathbb{T})$, alors

$$\forall n \in \mathbb{Z}\ :\ \widehat{\mathbf{z}f}(n) = \int_{\mathbb{T}} zf(z)z^{-n}dm(z) = \int_{\mathbb{T}} f(z)z^{-(n-1)}dm(z) = \widehat{f}(n-1).$$

Autrement dit, en notant $\mathcal{F} : L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ la transformation de Fourier, on a $\mathcal{F}(M_{\mathbf{z}}f) = S(\mathcal{F}f)$ pour toute $f \in L^2(\mathbb{T})$; et donc $M_{\mathbf{z}} = \mathcal{F}^{-1}S\mathcal{F}$. □

3.4. Opérateurs normaux.

DÉFINITION 3.18. *Un opérateur $T \in \mathcal{L}(H)$ est dit **normal** si on a $T^*T = TT^*$.*

EXEMPLES.

(1) Tout opérateur auto-adjoint est normal.
(2) Tout opérateur unitaire est normal.
(3) Tout opérateur de multiplication $M_\phi : L^2 \to L^2$ est normal.

LEMME 3.19. *Un opérateur $T \in \mathcal{L}(H)$ est normal si seulement si il s'écrit $T = A + iB$ avec A et B auto-adjoints et $AB = BA$.*

Démonstration. Si $T = A+iB$ avec A et B auto-adjoints, alors $T^*T = (A-iB)(A+iB) = (A^2 + B^2) + i(AB - BA)$ et $TT^* = (A + iB)(A - iB) = A^2 + B^2 - i(AB - BA)$. Donc T est normal si eut seulement si $(AB - BA) = -(AB - BA)$, *i.e.* $AB = BA$. □

LEMME 3.20. *Un opérateur $T \in \mathcal{L}(H)$ est normal si seulement si $\|Tx\| = \|T^*x\|$ pour tout $x \in H$.*

Démonstration. Comme T^*T et TT^* sont auto-adjoints, on a $T^*T = TT^*$ si et seulement si $\langle T^*Tx, x\rangle = \langle TT^*x, x\rangle$ pour tout $x \in H$, *i.e.* $\|Tx\|^2 = \|T^*x\|^2$. □

Corollaire 3.21. *Si* $T \in \mathcal{L}(H)$ *est normal, alors* $\ker(T^*) = \ker(T)$.

Démonstration. C'est évident. □

Corollaire 3.22. *Si* $T \in \mathcal{L}(H)$ *est normal, alors* $H = \ker(T) \overset{\perp}{\oplus} \overline{\mathrm{Im}(T)}$.

Démonstration. On a $H = \ker(T^*) \overset{\perp}{\oplus} \overline{\mathrm{Im}(T)}$ pour tout opérateur $T \in \mathcal{L}(H)$, et $\ker(T^*) = \ker(T)$ si T est normal. □

4. Théorème ergodique de von Neumann

On sait bien que si $\lambda \in \mathbb{C}$ vérifie $|\lambda| \leqslant 1$, alors

$$\frac{1}{n+1}\sum_{k=0}^{n} \lambda^k \xrightarrow{n\to\infty} \begin{cases} 0 & \text{si } \lambda \neq 1 \\ 1 & \text{si } \lambda = 1 \end{cases}$$

Le théorème suivante est une généralisation "opératorielle" qui a de très nombreuses applications.

Théorème 4.1. *Soit* H *un espace de Hilbert, et soit* $T \in \mathcal{L}(H)$ *vérifiant* $\|T\| \leqslant 1$. *Pour* $n \in \mathbb{N}$, *posons*

$$A_n := \frac{1}{n+1}\sum_{k=0}^{n} T^k.$$

Pour tout $x \in H$, *la suite* $(A_n x)$ *converge vers* px, *où* p *est la projection orhogonale sur* $\ker(T-I)$.

Démonstration.

Fait 1. On a $H = \ker(T-I) \overset{\perp}{\oplus} \overline{\mathrm{Im}(T-I)}$.

Preuve du Fait 1. Il suffit de montrer que $\ker(T-I) = \mathrm{Im}(T-I)^\perp$; autrement dit que $\ker(T-I) = \ker((T-I)^*)$, *i.e.* $\ker(T-I) = \ker(T^*-I)$.
Si $x \in H$ vérifie $Tx = x$, alors $\|Tx\|^2 = \|x\|^2$, autrement dit $\langle (I - T^*T)x, x\rangle = 0$. Comme $I - T^*T \geqslant 0$ puisque $\|T\| \leqslant 1$, on a donc $(I - T^*T)x = 0$, *i.e.* $T^*Tx = x$; et donc $T^*x = x$ puisque $Tx = x$. Ainsi, $\ker(I-T) \subseteq \ker(I-T^*)$; et en appliquant ceci à T^* (qui vérifie $\|T^*\| = \|T\| \leqslant 1$), on obtient l'autre inclusion.
Si on ne veut pas utiliser la notion d'opérateur positif, voici un preuve directe du fait que $\ker(T-I) = \ker(T^*-I)$. Si $x \in H$, alors

$$\begin{aligned}\|(I-T^*)x\|^2 &= \|x\|^2 - 2\mathrm{Re}\,\langle x, T^*x\rangle + \|T^*x\|^2 \\ &\leqslant 2\big(\|x\|^2 - \mathrm{Re}\,\langle x, T^*x\rangle\big) \qquad \text{car } \|T^*\| = \|T\| \leqslant 1 \\ &= 2\big(\|x\|^2 - \mathrm{Re}\,\langle Tx, x\rangle\big).\end{aligned}$$

Si $Tx = x$, on obtient donc $\|(I-T^*)x\|^2 \leqslant 0$, *i.e.* $T^*x = x$. Donc $\ker(T-I) \subseteq \ker(T^*-I)$; et l'inclusion inverse en appliquant ceci à T^*. □

Fait 2. $\|A_n(T-I)\| \to 0$ quand $n \to \infty$, et donc $A_n v \to 0$ pour tout $v \in \mathrm{Im}(T-I)$.

Preuve du Fait 2. On a $A_n(T-I) = \frac{1}{n+1}\left(\sum_{k=0}^n T^{k+1} - \sum_{k=0}^n T^k\right) = \frac{1}{n+1}(T^{n+1} - I)$, donc $\|A_n(T-I)\| \leqslant \frac{2}{n+1}$ car $\|T^{n+1}\| \leqslant \|T\|^{n+1} \leqslant 1$. □

Par le Fait 2 et comme $\|A_n\| \leqslant \frac{1}{n+1}\sum_{k=0}^n \|T^k\| \leqslant 1$ pour tout $n \in \mathbb{N}$, on voit (exo) que $A_n v \to 0$ pour tout $v \in \overline{\mathrm{Im}(T-I)}$. Comme évidemment $A_n u = u$ pour tout $u \in \ker(I-T)$, on en déduit que $A_n x \to px$ pour tout $x \in H$ d'après le Fait 1. □

Corollaire 4.2. *Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré, et soit $\phi : \Omega \to \Omega$ une application mesurable* préservant la mesure m, *i.e. telle que*

$$m\big(\phi^{-1}(E)\big) = m(E) \qquad \text{pour tout ensemble mesurable } E \subseteq \Omega.$$

Pour $k \in \mathbb{N}$, posons $\phi^k := \phi\circ\cdots\circ\phi$. Alors, pour toute $f \in L^2 := L^2(\Omega, m)$, les fonctions $f\circ\phi^k$ sont dans L^2 et la suite

$$f_n := \frac{1}{n+1}\sum_{k:=0}^n f\circ\phi^k$$

converge en norme L^2 vers une fonction $\widetilde{f}$ telle que $\widetilde{f}\circ\phi = \widetilde{f}$ pp.

Démonstration. Comme ϕ préserve la mesure m, on voit que si f_1 et f_2 sont deux fonctions mesurables égales m-pp, alors $f_1\circ\phi$ et $f_2\circ\phi$ sont également égale m-pp (exo). Donc, si $f \in L^2(\Omega, m)$, alors $Vf := f\circ\phi$ est bien définie pp.
Notons $\mathcal{E}\cap L^2$ l'ensemble des fonctions étagées appartenant à L^2. On sait que $\mathcal{E}\cap L^2$ est dense dans L^2. Si $f \in \mathcal{E}\cap L^2$ et si on écrit

$$f = \sum_{i=1}^N c_i \mathbf{1}_{E_i} \qquad \text{avec des } E_i \text{ disjoints,}$$

alors

$$Vf = \sum_{i=1}^N c_i \mathbf{1}_{\phi^{-1}(E_i)} \qquad \text{et les } \phi^{-1}(E_i) \text{ sont disjoints.}$$

Donc

$$\|Vf\|_2^2 = \sum_{i=1}^N |c_i|^2 m\big(\phi^{-1}(E_i)\big) = \sum_{i=1}^N |c_i|^2 m(E_i) = \|f\|_2^2.$$

Comme $\mathcal{E}\cap L^2$ est dense dans L^2, on en déduit que V envoie L^2 dans L^2 et est même une *isométrie*. En particulier, $\|V\| \leqslant 1$.
Si $f \in L^2$ est quelconque, alors $f_n = \frac{1}{n+1}\sum_{k=0}^n V^k f$. D'après le théorème appliqué avec $H := L^2$ et $T := V$, on déduit que (f_n) converge en norme L^2 vers $\widetilde{f}$, projeté orthogonal de f sur $\ker(V-I)$; ce qui donne le résultat souhaité. □

5. Racine carrée d'un opérateur positif

Théorème 5.1. *Si $A \in \mathcal{L}(H)$ est un opérateur positif, alors il existe un unique opérateur positif S tel que $S^2 = A$. On dit que S est la* ***racine carré*** *de A, et on écrit $S = \sqrt{A}$.*

Démonstration. On peut supposer que $\|A\| \leqslant 1$. Alors $B := I - A$ vérifie $0 \leqslant B \leqslant I$, et en particulier $\|I - A\| \leqslant 1$.

Fait 1. Soit $\mathcal{A}$ l'ensemble des fonctions $f : [-1, 1] \to \mathbb{R}$ de la forme $f(t) = \sum_{n=0}^\infty a_n t^n$ avec $\sum_{n=0}^\infty |a_n| < \infty$. Pour tout $T \in \mathcal{L}(H)$ auto-adjoint tel que $\|T\| \leqslant 1$ et pour toute $f \in \mathcal{A}$, $f(t) = \sum_0^\infty a_n t^n$, la série $\sum a_n T^n$ converge dans $\mathcal{L}(H)$ et on peut donc poser $f(T) := \sum_{n=0}^\infty a_n T^n$. Si $f, g \in \mathcal{A}$, alors $fg \in \mathcal{A}$ et $(fg)(T) = f(T)g(T)$.

Preuve du Fait 1. Exo. $\square$

FAIT 2. Il existe une suite $(c_n)_{n\geqslant 1}$ de nombres réels $\geqslant 0$ telle que $\sum_{n=1}^{\infty} c_n = 1$ et $\forall t \in [-1,1]\ :\ \sqrt{1-t} = 1 - \sum_1^{\infty} c_n t^n$.

Preuve du Fait 2. On sait que le développement en série entière de $\sqrt{1+u}$ sur l'intervalle $]-1,1[$ est

$$\sqrt{1+u} = \sum_{n=0}^{\infty} \binom{1/2}{n} u^n = 1 + \sum_{n=1}^{\infty} \binom{1/2}{n} u^n.$$

De plus $\binom{1/2}{n} = \frac{1/2(1/2-1)\cdots(1/2-n+1)}{n!}$ est $\geqslant 0$ si $n \geqslant 1$ est impair et < 0 si n est pair. Donc $c_n := (-1)^{n+1}\binom{1/2}{n}$ est $\geqslant 0$ pour tout $n \geqslant 1$, et

$$\forall t \in \,]-1,1[\ :\ \sqrt{1-t} = 1 - \sum_{n=1}^{\infty} c_n t^n.$$

Comme les c_n sont $\geqslant 0$, on a $\lim_{t\to 1^-} \sum_{n=1}^{\infty} c_n t^n = \sum_{n=1}^{\infty} c_n$ par convergence monotone. Donc $\sum_{n=1}^{\infty} c_n = 1$ par continuité de $\sqrt{1-t}$ en 1. En particulier $\sum_{n=1}^{\infty} c_n < \infty$, donc la série $\sum c_n t^n$ converge normalement sur $[-1,1]$, et donc la formule précédente est aussi valable pour $t = -1$ par continuité. $\square$

Soit $f : [-1,1] \to \mathbb{R}$ définie par $f(t) := \sqrt{1-t}$. Par les Faits 1 et 2 et comme $\|B\| \leqslant 1$, on peut poser $S := f(B) = I - \sum_{n=1}^{\infty} c_n B^n$. Comme $B^* = B$ et $c_n \in \mathbb{R}$, on a $S^* = S$. Si $x \in H$, alors $\langle Sx, x\rangle = \|x\|^2 - \sum_{n=1}^{\infty} c_n \langle B^n x, x\rangle \geqslant \|x\|^2 - \sum_{n=1}^{\infty} c_n \|B^n\|\,\|x\|^2$ car $c_n \geqslant 0$, et donc $\langle Sx, x\rangle \geqslant (1 - \sum_{n=1}^{\infty} c_n)\,\|x\|^2 = 0$ car $\|B\| \leqslant 1$. Donc l'opérateur S est positif. Enfin, par le Fait 1 et comme $f(t)^2 \equiv 1 - t$, on a $S^2 = f^2(B) = I - B = A$.

On a donc montré l'existence d'un opérateur $S \geqslant 0$ tel que $S^2 = A$. De plus, S est limite d'une suite de polynômes en A, donc S commute avec tout opérateur $T \in \mathcal{L}(H)$ tel que $TA = AT$ (exo).

Soit maintenant T un autre opérateur $\geqslant 0$ tel que $T^2 = A$. Alors $TA = T^3 = AT$, donc $TS = ST$. Comme $T^2 = S^2$, on en déduit $(T-S)(T+S) = 0$. On a donc $\mathrm{Im}(T+S) \subseteq \ker(T-S)$. Mais comme S et T sont $\geqslant 0$, on a aussi $\mathrm{Im}(T+S)^{\perp} = \ker(T+S) = \ker(T) \cap \ker(S)$, car si $(S+T)x = 0$, alors $0 = \langle Tx + Sx, x\rangle = \langle Tx, x\rangle + \langle Sx, x\rangle$, donc $\langle Tx, x\rangle = 0 = \langle Sx, x\rangle$ car les deux termes sont $\geqslant 0$, et donc $Tx = 0 = Sx$; donc certainement $\mathrm{Im}(T+S)^{\perp} \subseteq \ker(T-S)$. Au total, $T - S \equiv 0$ sur $\mathrm{Im}(T+S)$ et sur $\mathrm{Im}(T+S)^{\perp}$, donc $T - S = 0$. $\square$

REMARQUE 5.2. La preuve du théorème a montré que $\sqrt{A}$ est limite d'une suite de polynômes en A, et donc $\sqrt{A}$ commute avec tout opérateur $T \in \mathcal{L}(H)$ tel que $TA = AT$.

NOTATION. Si $T \in \mathcal{L}(H)$, on pose $|T| := \sqrt{T^*T}$.

LEMME 5.3. *Si $T \in \mathcal{L}(H)$, alors $\forall x \in H\ :\ \|\,|T|x\| = \|Tx\|$. En particulier,* $\ker(|T|) = \ker(T)$.

Démonstration. On a $\||T|x\|^2 = \langle |T|^*|T|x, x\rangle = \langle |T|^2 x, x\rangle = \langle T^*Tx, x\rangle = \|Tx\|^2$. $\square$

6. Décomposition polaire

On sait bien que tout nombre complexe z peut s'écrire sous la forme $z = re^{i\theta}$. Dans cette section, on va généraliser cela aux opérateurs hilbertiens.

DÉFINITION 6.1. *Soit $W \in \mathcal{L}(H)$. On dit que W est une **isométrie partielle** si la restriction de W à $\ker(W)^{\perp}$ est une isométrie.*

Remarque. Évidemment, toute isométrie est une isométrie partielle.

Exemple. Le backward shift B sur $\ell^2(\mathbb{N})$ est une isométrie partielle : on a $\ker(B) = [e_0]$, donc $\ker(B)^{\perp} = [e_n;\ n \geqslant 1]$; et si $x = \sum_{n\geqslant 1} x_n e_n \in \ker(B)^{\perp}$, alors $Bx = \sum_{n\geqslant 1} x_n e_{n-1}$ et donc $\|Bx\|^2 = \sum_{n\geqslant 1} |x_n|^2 = \|x\|^2$.

LEMME 6.2. *Soit $W \in \mathcal{L}(H)$. Alors W est une isométrie partielle si et seulement si W^*W est une projection ; et dans ce cas, W^*W est la projection orthogonale sur $\ker(W)^{\perp}$.*

Démonstration. Supposons que W soit une isométrie partielle, et posons $E := \ker(W)^{\perp}$. D'abord, on a $\|W\| \leqslant 1$, car si $x \in H$, on peut écrire $x = e + f$ avec $e \in E$ et $f \in \ker(W)$, donc $Wx = We$ et $\|Wx\| = \|We\| = \|e\| \leqslant \|x\|$. Donc $I - W^*W \geqslant 0$. Ensuite, si $x \in E$, alors $\langle I - W^*Wx, x\rangle = \|x\|^2 - \langle W^*Wx, x\rangle = \|x\|^2 - \|x\|^2 = 0$, donc $(I - W^*W)x = 0$, *i.e.* $W^*Wx = x$; et si $x \in E^{\perp} = \ker(W)$, alors évidemment $W^*Wx = 0$. Donc W^*W est en effet la projection orthogonale sur $E = \ker(W)^{\perp}$.
Inversement, supposons que W^*W soit une projection, et posons $E := \mathrm{Im}(W^*W)$. Si $x \in E$, alors $W^*Wx = x$, donc $\|Wx\|^2 = \langle W^*Wx, x\rangle = \|x\|^2$. De plus, comme W^*W est auto-adjoint, on a $E^{\perp} = \ker(W^*W) = \ker(W)$, et donc $E = \ker(W)^{\perp}$ car E est fermé (W^*W est une projection, donc $E = \ker(I - W^*W)$). □

THÉORÈME 6.3. *Si $T \in \mathcal{L}(H)$ alors T peut s'écrire sous la forme $T = WS$, où S est un opérateur positif et W est une isométrie partielle avec $\ker(W)^{\perp} = \overline{\mathrm{Im}(S)}$. De plus, cette écriture est unique, et on a $S = |T|$ et $\ker(W) = \ker(T)$. L'écriture $T = W\,|T|$ s'appelle la **décomposition polaire** de T.*

Démonstration. Montrons d'abord l'existence d'une isométrie partielle W telle que $W|T| = T$ et $\ker(W)^{\perp} = \overline{\mathrm{Im}(|T|)}$. Comme $\ker(T) = \ker(|T|)$, il existe une unique application linéaire $L : \mathrm{Im}(|T|) \to H$ telle que $Tx = L(|T|x)$ pour tout $x \in H$. De plus, si $y = |T|x \in \mathrm{Im}(|T|)$, alors $\|L(y)\| = \|Tx\| = \||T|x\| = \|y\|$; donc L est une isométrie de $\mathrm{Im}(|T|)$ dans H. Par le théorème de prolongement des applications linéaires continues (H est complet...), on peut prolonger L en une isométrie linéaire $\widetilde{L} : \overline{\mathrm{Im}(|T|)} \to H$. On pose alors $W := \widetilde{L}p$, où $p : H \to \overline{\mathrm{Im}(|T|)}$ est la projection orthogonale de H sur $\overline{\mathrm{Im}(|T|)}$. On a $W|T| = T$ par définition : si $x \in H$, alors $p|T|x = |T|x$ et donc $W|T|x = \widetilde{L}|T|x = Tx$. De plus, la restriction de W à $\overline{\mathrm{Im}(|T|)}$ est égale à $\widetilde{L}$, donc c'est une isométrie ; et comme $\widetilde{L}$ est injective, on a $\ker(W) = \ker(p) = \overline{\mathrm{Im}(|T|)}^{\perp}$, donc $\ker(W)^{\perp} = \overline{\mathrm{Im}(|T|)}$. Ainsi, W est une isométrie partielle qui convient. Enfin, comme $\ker(W)$ est un sous-espace fermé de H et que $|T|$ est auto-adjoint, on a $\ker(W) = \overline{\mathrm{Im}(|T|)}^{\perp} = \ker(|T|) = \ker(T)$.
Montrons maintenant l'unicité. Soient S un opérateur positif et W_1 une isométrie partielle tels que $T = W_1 S$ et $\ker(W_1)^{\perp} = \overline{\mathrm{Im}(S)}$. Alors $T^*T = SW_1^*W_1S$. Mais comme W_1 est une isométrie partielle, $W_1^*W_1$ est la projection orthogonale sur $\ker(W_1)^{\perp} =$

$\overline{\mathrm{Im}(S)}$. Donc $W_1^*W_1S = S$, et donc $T^*T = S^2$. Comme S est positif, on a donc nécessairement $S = \sqrt{T^*T} = |T|$. Enfin, l'équation $T = W_1S$ détermine entièrement W_1 sur $\mathrm{Im}(S)$, donc sur $\overline{\mathrm{Im}(S)} = \ker(S)^\perp$; et comme $W_1 \equiv 0$ sur $\ker(S)$ puisque $\ker(S) = \overline{\mathrm{Im}(S)}^\perp = \ker(W_1)$, on voit que W_1 est déterminé de manière unique. □

COROLLAIRE 6.4. *Si $T \in \mathcal{L}(H)$ est injectif et à image dense, il existe un unique opérateur unitaire U tel que $T = U|T|$.*

Démonstration. Si $T = W|T|$ est la décomposition polaire de T, alors $\ker(W) = \ker(T) = \{0\}$, donc W est une isométrie ; et $\mathrm{Im}(W)$ contient $\mathrm{Im}(T)$, donc $\mathrm{Im}(W)$ est dense dans H, et donc W est surjective puisque $\mathrm{Im}(W)$ est aussi fermé (W est une isométrie !). Ainsi, W est unitaire et on peut prendre $U := W$. Pour l'unicité, on observe que l'opérateur $|T|$ est à image dense car il est auto-adjoint et injectif ($\ker(|T|) = \ker(T) = \{0\}$). L'équation $T = U|T|$ détermine donc U de manière unique par continuité. □

COROLLAIRE 6.5. *Si H est* de dimension finie*, alors tout opérateur $T \in \mathcal{L}(H)$ peut s'écrire sous la forme $T = US$ avec U unitaire et $S \geqslant 0$.*

Démonstration. Soit $T = WS$ la décomposition polaire de T. Alors $W_{|\mathrm{Im}(S)}$ est une isométrie de $\mathrm{Im}(S)$ sur $\mathrm{Im}(T)$. Comme on est en dimension finie, on peut écrire $\dim \mathrm{Im}(S)^\perp = \dim(H) - \dim \mathrm{Im}(S) = \dim(H) - \dim \mathrm{Im}(T) = \dim \mathrm{Im}(T)^\perp$. Donc on peut choisir un opérateur unitaire V de $\mathrm{Im}(S)^\perp$ sur $\mathrm{Im}(T)^\perp$. Alors $U := W_{|\mathrm{Im}(S)} \oplus V \in \mathcal{L}(H)$ est un opérateur unitaire, et on a $T = US$ puisque $U_{|\mathrm{Im}(S)} = W_{|\mathrm{Im}(S)}$. □

7. Inégalité de von Neumann

THÉORÈME 7.1. *Soit H un espace de Hilbert complexe, et soit $T \in \mathcal{L}(H)$ vérifiant $\|T\| \leqslant 1$. Pour tout polynôme $P \in \mathbb{C}[\mathbf{z}]$, on a*

$$\|P(T)\| \leqslant \|P\|_\infty := \sup\big\{|P(z)|;\ |z| = 1\big\}.$$

Démonstration. Si $P = c_0\mathbf{1} + c_1\mathbf{z} + \cdots + c_d\mathbf{z}^d$, alors l'opérateur $P(T)$ est bien entendu défini par $P(T) := c_0I + c_1T + \cdots + c_dT^d$.

CAS 1. $\dim H < \infty$ et l'opérateur T est unitaire.

On identifie H à $\mathbb{C}^d$, où $d := \dim(H)$, et $\mathcal{L}(\mathbb{C}^d)$ à $M_d(\mathbb{C})$. Comme T est unitaire, il est diagonalisable en base orthonormée et ses valeurs propres sont de module 1. On peut donc trouver une matrice unitaire U et $\lambda_1, \dots, \lambda_d \in \mathbb{C}$ avec $|\lambda_i| = 1$ tels que

$$T = U^* \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_d \end{pmatrix} U.$$

En notant D la matrice diagonale impliquée, on a $T^k = U^*D^kU$ pour tout $k \in \mathbb{N}$, et donc $P(T) = U^*P(D)U$, *i.e.*

$$P(T) = U^* \begin{pmatrix} P(\lambda_1) & & & \\ & P(\lambda_2) & & \\ & & \ddots & \\ & & & P(\lambda_d) \end{pmatrix} U.$$

Donc
$$\|P(T)\| = \|P(D)\| = \max\big(|P(\lambda_1)|, \dots, |P(\lambda_d)|\big) \leqslant \|P\|_\infty.$$

Cas 2. $\dim(H) < \infty$ et T est quelconque.

D'après la décomposition polaire, on peut écrire $T = US$, où U est unitaire et $S \geqslant 0$. On a $\|S\| = \|U^*T\| = \|T\| \leqslant 1$. Comme S est auto-adjoint, il est diagonalisable en base orthonormée, et comme $S \geqslant 0$ et $\|S\| \leqslant 1$, on peut donc trouver une matrice unitaire V et $s_1, \dots, s_d \in \mathbb{R}$ avec $0 \leqslant s_i \leqslant 1$ tels que
$$S = V^* \begin{pmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_d \end{pmatrix} V.$$
On a ainsi
$$T = UV^* \begin{pmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_d \end{pmatrix} V.$$
Introduisons alors l'application $A : \mathbb{C}^d \to M_d(\mathbb{C})$ définie par
$$A(z_1, \dots, z_d) := UV^* \begin{pmatrix} z_1 & & & \\ & z_2 & & \\ & & \ddots & \\ & & & z_d \end{pmatrix} V,$$
de sorte que
$$P(T) = P\big(A(s_1, \dots, s_d)\big).$$
Pour tous $x, y \in \mathbb{C}^d$, l'application $(z_1, \dots, z_d) \mapsto \langle P\big(A(z_1, \dots, z_d)\big)x, y\rangle$ est polynomiale, donc holomorphe sur $\mathbb{C}$ par rapport à chaque variable z_i. En appliquant le *principe du maximum* variable par variable dans le disque $\mathbb{D} = \{z \in \mathbb{C};\ |z| < 1\}$ et comme $s_1, \dots, s_d \in \overline{\mathbb{D}}$, on en déduit
$$\begin{aligned} |\langle P(T)x, y\rangle| &= \big|\langle P\big(A(s_1, \dots, s_d)\big)x, y\rangle\big| \\ &\leqslant \sup\Big\{\big|\langle P\big(A(z_1, \dots, z_d)\big)x, y\rangle\big|;\ |z_1| = \dots = |z_d| = 1\Big\}. \end{aligned}$$
Ceci étant vrai pour tous $x, y \in \mathbb{C}^d$, on obtient ainsi
$$\|P(T)\| \leqslant \sup\Big\{\big\|P\big(A(z_1, \dots, z_d)\big)\big\|;\ |z_1| = \dots = |z_d| = 1\Big\}.$$
Mais si $|z_1| = \dots = |z_d| = 1$, alors l'opérateur $A(z_1, \dots, z_d)$ est unitaire (micro-exo) ; et donc $\big\|P\big(A(z_1, \dots, z_d)\big)\big\| \leqslant \|P\|_\infty$ d'après le Cas 1. Donc $\|P(T)\| \leqslant \|P\|_\infty$.

Cas 3. Cas général.

Il suffit de montrer qu'on a $\|P(T)x\| \leqslant \|P\|_\infty$ pour tout $x \in H$ vérifiant $\|x\| \leqslant 1$. Fixons un tel x.
Si on pose $E := [T^k x;\ k \in \mathbb{N}]$, alors E est un sous-espace fermé de H invariant par T, et E est séparable. Comme $P(T_{|E}) = P(T)_{|E}$, on peut donc supposer que H est séparable, quitte à remplacer T par $T_{|E}$.

Soit $(e_n)_{n\geqslant 0}$ une base orthonormée de H. Pour $n \in \mathbb{N}$, notons p_n la projection orthogonale sur $H_n := [e_0, \dots, e_n]$. Alors $\|p_n\| \leqslant 1$ et $p_n z \to z$ pour tout $z \in H$ quand $n \to \infty$. On en déduit que si on pose $T_n := p_n T p_n$, alors $\|T_n\| \leqslant \|T\| \leqslant 1$ et $T_n x \to Tx$. En effet : $\|T_n x - Tx\| \leqslant \|p_n T(p_n x - x)\| + \|p_n Tx - Tx\| \leqslant \|T\| \, \|p_n x - x\| + \|p_n Tx - Tx\|$. De même, $T_n^2 x \to T^2 x$ car $\|T_n^2 x - T^2 x\| \leqslant \|T_n(T_n x - Tx)\| + \|T_n(Tx) - T(Tx)\| \leqslant \|T_n x - Tx\| + \|T_n(Tx) - T(Tx)\|$. Par récurrence, on montre que $T_n^k x \to T^k x$ pour tout $k \in \mathbb{N}$ quand $n \to \infty$, et donc que $P(T_n)x \to P(T)x$ pour tout $x \in H$. Mais les T_n “vivent” sur des espaces de dimension finie et vérifient $\|T_n\| \leqslant 1$. Donc $\|P(T_n)\| \leqslant \|P\|_\infty$ pour tout $n \in \mathbb{N}$, et donc $\|P(T)x\| = \lim_{n\to\infty} \|P(T_n)x\| \leqslant \|P\|_\infty$. □

8. Opérateurs compacts

8.1. Généralités minimales.

Notation. Si X est un espace vectoriel normé, on note B_X la boule unité fermée de X.

Définition 8.1. *Soient X et Y des evn. On dit qu'un opérateur $T \in \mathcal{L}(X, Y)$ est* ***compact*** *si $\overline{T(B_X)}$ est une partie compacte de Y. On note $\mathcal{K}(X, Y)$ l'ensemble des opérateurs compacts de X dans Y, et on pose $\mathcal{K}(X) := \mathcal{K}(X, X)$.*

Remarque 1. Si $T \in \mathcal{L}(X, Y)$ est compact, alors $\overline{T(B)}$ est compact pour tout ensemble borné $B \subseteq X$.

Remarque 2. Un opérateur $T \in \mathcal{L}(X, Y)$ est compact si et seulement si : pour toute suite bornée $(x_n) \subseteq X$, la suite (Tx_n) possède une sous-suite convergente.

Exercice. Montrer que dans la définition d'un opérateur compact, il n'est pas nécessaire de supposer *a priori* que T est un opérateur borné ; autrement dit : si $T : X \to Y$ est une application linéaire telle que $\overline{T(B_X)}$ est un compact de Y, alors T est nécessairement continue.

Exemple 1. Tout opérateur **de rang fini** est compact.

Démonstration. Exo. □

Exemple 2. Soit $[a, b]$ un intervalle compact de $\mathbb{R}$. Notons $\mathcal{C}^1([a, b])$ l'espace de toutes les fonctions $f : [a, b] \to \mathbb{C}$ de classe $\mathcal{C}^1$, muni de la norme $\|f\|_{\mathcal{C}^1} := \|f\|_\infty + \|f'\|_\infty$. Alors l'injection canonique $J : \mathcal{C}^1([a, b]) \hookrightarrow \mathcal{C}([a, b]$ est un opérateur compact.

Démonstration. On a $\|f\|_\infty \leqslant \|f\|_{\mathcal{C}^1}$ pour toute $f \in \mathcal{C}^1([a, b])$, donc J est un opérateur borné (et $\|J\| \leqslant 1$).
Si (f_n) est une suite bornée dans l'espace $\mathcal{C}^1([a, b])$ et si on pose $M := \sup_{n\in\mathbb{N}} \|f_n\|_{\mathcal{C}^1}$, alors (f_n) est uniformément bornée, et les f_n sont toutes M-lipschitziennes puisque $\|f_n'\|_\infty \leqslant M$. Par le *Théorème d'Ascoli*, on en déduit que (f_n) possède une sous-suite (f_{n_k}) qui converge uniformément sur $[a, b]$, *i.e.* (Jf_{n_k}) converge dans $\mathcal{C}([a, b])$. Donc J est compact. □

Remarque 8.2. Soient $A \in \mathcal{L}(Y, Z)$ et $B \in \mathcal{L}(X, Y)$. Si A *ou* B est compact, alors AB est compact.

Démonstration. Exo. □

Proposition 8.3. *Si Y est un espace de Banach, alors $\mathcal{K}(X,Y)$ est un sous-espace vectoriel fermé de $\mathcal{L}(X,Y)$.*

Démonstration. Le fait que $\mathcal{K}(X,Y)$ soit un sous-ev de $\mathcal{L}(X,Y)$ est un exo facile. Soit $(T_n) \subseteq \mathcal{K}(X,Y)$ convergeant vers $T \in \mathcal{L}(X,Y)$. On veut montrer que T est compact.
Comme Y est complet, il suffit de montrer que $T(B_X)$ est *précompact* ; autrement dit, que pour tout $\varepsilon > 0$, on peut trouver un ε-réseau fini pour $T(B_X)$. On fixe donc $\varepsilon > 0$. Soit $n \in \mathbb{N}$ tel que $\|T_n - T\| < \varepsilon/2$. Alors $T_n(B_X)$ est un $(\varepsilon/2)$-réseau pour $T(B_X)$ (exo). Comme T_n est compact, $T_n(B_X)$ est précompact, dont admet un $(\varepsilon/2)$-réseau fini F. Alors F est un ε-réseau pour $T(B_X)$ (autre exo). □

Corollaire 8.4. *Si $T \in \mathcal{L}(X,Y)$ et s'il existe une suite d'opérateurs de rangs finis (R_n) telle que $\|R_n - T\| \to 0$, alors T est compact.*

Lemme 8.5. *Soit $T \in \mathcal{L}(X,Y)$ un opérateur compact, et soit $A \in \mathcal{L}(Y,Z)$. Si $(A_n) \subseteq \mathcal{L}(Y,Z)$ est une suite bornée telle que $A_n y \to Ay$ pour tout $y \in Y$ quand $n \to \infty$, alors $\|A_nT - AT\| \to 0$.*

Démonstration. Comme la suite (A_n) est bornée dans $\mathcal{L}(Y,Z)$, toutes les applications A_n sont C-lipschitziennes pour une même constante C, et en particulier la suite (A_n) est *équicontinue*. Donc (exo) *$A_n y \to Ay$ uniformément sur tout compact $K \subseteq Y$*. Comme $K := \overline{T(B_X)} \subseteq Y$ est compact, on en déduit que $A_nTx \to ATx$ uniformément sur B_X, et donc $\|A_nT - AT\| = \sup\{\|T_nx - Tx\|;\ x \in B_X\} \xrightarrow{n\to\infty} 0$. □

Corollaire 8.6. *Soit H un espace de Hilbert séparable (de dimension infinie), et soit $(e_i)_{i\in\mathbb{N}}$ une base hilbertienne de H. Pour $n \in \mathbb{N}$, soit $p_n \in \mathcal{L}(H)$ la projection orthogonale sur $E_n := [e_0, \dots, e_n]$. Si $T \in \mathcal{L}(X,H)$ est un opérateur compact, alors $\|p_nT - T\| \to 0$.*

Démonstration. On a $\|p_n\| \leqslant 1$ pour tout n, et $p_n y \to y$ pour tout $y \in H$. □

Corollaire 8.7. *Si H est un espace de Hilbert, alors tout opérateur compact $T \in \mathcal{L}(X,H)$ est limite d'une suite d'opérateurs de rang fini.*

Démonstration. Si H est séparable, cela découle du corollaire précédent car les p_n sont de rangs finis. Dans le cas général, on observe que $K := \overline{T(B_X)}$ est séparable car c'est un compact de l'espace métrique H. Donc $Y_0 := \overline{\mathrm{Im} T} = \overline{\mathrm{vect}(K)}$ est un sous-espace séparable de H ; et en considérant T comme un opérateur de X dans H_0, on peut appliquer le 1er cas. □

Exemple 8.8. Soit $a = (a_n)_{n\in\mathbb{N}} \in \ell^\infty(\mathbb{N})$, et soit $\Delta_a : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ l'opérateur diagonal associé. Alors Δ_a est compact si et seulement si $a \in c_0$, *i.e.* $a_n \to 0$ quand $n \to \infty$.

Démonstration. Pour $n \in \mathbb{N}$, notons $p_n \in \mathcal{L}(\ell^2)$ la projection orthogonale sur $[e_0, \dots, e_n]$. D'après 8.4 et 8.6, Δ_a est compact si et seulement si $\|p_n\Delta_a - \Delta_a\| \to 0$ quand $n \to \infty$. Mais $\Delta_a - p_n\Delta_a = \Delta_{b^{(n)}}$, où $b^{(n)} := (0, \dots, 0, a_{n+1}, a_{n+2}, \dots)$. Donc $\|\Delta_a - p_n\Delta_a\| = \|b^{(n)}\|_\infty = \sup_{i>n} |a_i|$, et donc $\|\Delta_a - p_n\Delta_a\| \to 0$ si et seulement si $a \in c_0$. □

REMARQUE 8.9. Soit $T \in \mathcal{L}(H_1, H_2)$. Si T est compact alors, pour toute suite orthonormale $(f_n)_{n\in\mathbb{N}} \subseteq H_1$, on a que $\|Tf_n\| \to 0$.

Démonstration. Le point clé est que $\langle f_n, x\rangle \to 0$ pour tout $x \in H_1$, d'après l'inégalité de Bessel. (Plus tard, on dira que $f_n \to 0$ *faiblement.*)
Supposons que $\|Tf_n\|$ ne tende pas vers 0. Alors, quitte à extraire une sous-suite de (f_n), on peut supposer qu'il existe $\varepsilon > 0$ tel que $\forall n \in \mathbb{N}\ :\ \|Tf_n\| \geqslant \varepsilon$; et comme T est compact, on peut également supposer que la suite (Tf_n) converge vers un certain $z \in H_2$. Alors $\|z\| \geqslant \varepsilon > 0$, et $\langle Tf_n, z\rangle \to \|z\|^2$; mais $\langle Tf_n, z\rangle = \langle f_n, T^*z\rangle$, donc $\langle Tf_n, z\rangle$ doit tendre vers 0, ce qui est une contradiction. □

Exercice. Redémontrer que l'opérateur diagonal Δ_a n'est pas compact si $a \notin c_0$.

8.2. Opérateurs de Hilbert-Schmidt.

LEMME 8.10. *Soient H_1, H_2 deux espaces de Hilbert. Si $T \in \mathcal{L}(H_1, H_2)$ et si $(e_i)_{i\in I}$ et $(f_j)_{j\in J}$ sont des bases hilbertiennes de H_1 et H_2, alors*

$$\sum_{i\in I} \|Te_i\|^2 = \sum_{j\in J} \|T^*f_j\|^2.$$

En particulier, la quantité

$$\|T\|_{\mathrm{HS}}^2 := \sum_{i\in I} \|Te_i\|^2$$

ne dépend pas de la base hilbertienne $(e_i)_{i\in I}$ de H_1 ; et on a de plus $\|T\|_{\mathrm{HS}}^2 = \|T^*\|_{\mathrm{HS}}^2$.

Démonstration. Par "Fubini positif" pour les sommes,

$$\sum_{i\in I} \|Te_i\|^2 = \sum_{i\in I}\sum_{j\in J} |\langle Te_i, f_j\rangle|^2 = \sum_{j\in J}\sum_{i\in I} |\langle T^*f_j, e_i\rangle|^2 = \sum_{j\in J} \|T^*f_j\|^2.$$

□

DÉFINITION 8.11. *Si $T \in \mathcal{L}(H_1, H_2)$, on pose*

$$\|T\|_{\mathrm{HS}}^2 := \sum_{i\in I} \|Te_i\|^2,$$

où $(e_i)_{i\in I}$ est une base hilbertienne quelconque de H_1. On dit que T est un opérateur ***de Hilbert-Schmidt*** *si on a $\|T\|_{\mathrm{HS}} < \infty$.*

REMARQUE 1. Par définition, T est de Hilbert-Schmidt si et seulement si T^* est de Hilbert-Schmidt.

REMARQUE 2. On a $\|T\| \leqslant \|T\|_{\mathrm{HS}}$ pour tout $T \in \mathcal{L}(H_1, H_2)$.

Démonstration. Soit $(f_j)_{j\in J}$ une base hilbertienne de H_2. Si $x \in H_1$, alors

$$\|Tx\|^2 = \sum_{j\in J} |\langle Tx, f_j\rangle|^2 = \sum_{j\in J} |\langle x, T^*f_j\rangle|^2 \leqslant \sum_{j\in J} \|x\|^2\, \|T^*f_j\|^2 = \|T^*\|_{\mathrm{HS}}^2\, \|x\|^2.$$

Donc $\|T\| \leqslant \|T^*\|_{\mathrm{HS}} = \|T\|_{\mathrm{HS}}$. □

REMARQUE 3. Si $B \in \mathcal{L}(H_1, H_2)$ et $A \in \mathcal{L}(H_2, H_3)$, alors $\|AB\|_{\mathrm{HS}} \leqslant \|A\|_{\mathrm{HS}}\|B\|$ et $\|AB\|_{\mathrm{HS}} \leqslant \|A\|\, \|B\|_{\mathrm{HS}}$. En particulier, si A *ou* B est de Hilbert-Schmidt, alors AB est de Hilbert-Schmidt.

Démonstration. Soit $(e_i)_{i\in I}$ une base hilbertienne de H_1. On a

$$\|AB\|_{\mathrm{HS}}^2 = \sum_{i\in I} \|ABe_i\|^2 \leqslant \sum_{i\in I} \|A\|^2\|Be_i\|^2 = \|A\|^2\|B\|_{\mathrm{HS}}^2,$$

donc $\|AB\|_{\mathrm{HS}} \leqslant \|A\|_{\mathrm{HS}}\|B\|$. On en déduit

$$\|AB\|_{\mathrm{HS}} = \|(AB)^*\|_{\mathrm{HS}} = \|B^*A^*\|_{\mathrm{HS}} \leqslant \|B^*\|\,\|A^*\|_{\mathrm{HS}} = \|A\|\,\|B\|_{\mathrm{HS}}.$$

□

Proposition 8.12. *Tout opérateur de Hilbert-Schmidt est compact.*

Démonstration. Soit $T \in \mathcal{L}(H_1, H_2)$ de Hilbert-Schmidt. Si $\dim(H_1) < \infty$, alors T est de rang fini, donc compact ; donc on suppose que $\dim(H_1) = \infty$. Pour simplifier, on va aussi supposer que H_1 est séparable.
Soit $(e_i)_{i\in\mathbb{N}}$ une base hilbertienne de H_1. Pour $n \in \mathbb{N}$, soit $p_n \in \mathcal{L}(H_1)$ la projection orthogonale sur $E_n := [e_0, \dots, e_n]$. On a $Tp_ne_i = Te_i$ pour tout $i \leqslant n$, et $Tp_ne_i = 0$ si $i > n$; donc

$$\|Tp_n - T\|_{\mathrm{HS}}^2 = \sum_{i>n} \|Te_i\|^2 \xrightarrow{n\to\infty} 0.$$

A fortiori $\|Tp_n - T\| \to 0$, et donc T est compact car les Tp_n sont de rang fini. □

Exemple 1. Soit $a \in \ell^\infty(\mathbb{N})$. L'opérateur diagonal $\Delta_a : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ est de Hilbert-Schmidt si et seulement si $a \in \ell^2$, *i.e.* $\sum_{i=0}^\infty |a_i|^2 < \infty$, et on a $\|\Delta_a\|_{\mathrm{HS}} = \|a\|_2$.

Démonstration. Exo. □

Exemple 2. Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré sigma-fini tel que l'espace $L^2(\Omega, m)$ soit séparable (par exemple, Ω pourrait être un intervalle de $\mathbb{R}$). Si $K : \Omega\times\Omega \to \mathbb{C}$ est un noyau appartenant à $L^2(\Omega\times\Omega, m\otimes m)$, alors l'opérateur associé $T_K : L^2(\Omega, m) \to L^2(\Omega, m)$ est de Hilbert-Schmidt, et $\|T_K\|_{\mathrm{HS}} = \|K\|_2$.

Démonstration. Pour $x \in \Omega$, soit $K_x : \Omega \to \mathbb{C}$ la fonction définie par $K_x(y) := K(x, y)$. Par "Fubini positif", $K_x \in L^2(\Omega, m)$ pour presque tout $x \in \Omega$ et

$$\|K\|_{L^2(\Omega\times\Omega)}^2 = \int_\Omega \left(\int_\Omega |K(x,y)|^2 dm(y)\right) dm(x) = \int_\Omega \|K_x\|_{L^2(\Omega)}^2.$$

Soit $(e_i)_{i\in I}$ une base hilbertienne de $L^2(\Omega, m)$. Comme $L^2(\Omega, m)$ est supposé séparable, l'ensemble d'indices I est dénombrable. Pour presque tout $x \in \Omega$, on a

$$T_Ke_i(x) = \int_\Omega K(x,y)e_i(y)\,dm(y) = \langle K_x, \overline{e_i}\rangle_{L^2(\Omega)};$$

et donc

$$\|T_Ke_i\|^2 = \int_\Omega |\langle K_x, \overline{e_i}\rangle|^2 dm(x) = \int_\Omega |\langle \overline{K_x}, e_i\rangle|^2 dm(x).$$

Donc, par "Fubini positif" à nouveau (où on a réellement besoin que l'ensemble I soit dénombrable) :

$$\begin{aligned}\|T_K\|_{\mathrm{HS}}^2 &= \sum_{i\in I}\int_\Omega |\langle \overline{K_x}, e_i\rangle|^2 dm(x)\\ &= \int_\Omega \left(\sum_{i\in I}|\langle \overline{K_x}, e_i\rangle|^2\right) dm(x)\\ &= \int_\Omega \|K_x\|_2^2\, dm(x) \qquad \text{car } \|\overline{K_x}\|_2 = \|K_x\|_2\\ &= \|K\|^2_{L^2(\Omega\times\Omega)}.\end{aligned}$$

□

8.3. Diagonalisation des opérateurs normaux compacts.

8.3.1. *Objectif.* On sait que si $T \in \mathcal{L}(\mathbb{R}^n)$ est auto-adjoint, alors T est diagonalisable en base orthonormée. De même, si $T \in \mathcal{L}(\mathbb{C}^n)$ est normal, alors T est diagonalisable en base orthonormée. On va retrouver et grandement généraliser ces résultats en démontrant le théorème suivant.

THÉORÈME 8.13. *Si H est un espace de Hilbert sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, alors tout opérateur auto-adjoint* compact *$T \in \mathcal{L}(H)$ est diagonalisable en base orthonormée, i.e. H possède une base hilbertienne formée de vecteurs propres pour T. Si $\mathbb{K} = \mathbb{C}$, alors tout opérateur normal compact $T \in \mathcal{L}(H)$ est diagonalisable en base orthonormée.*

8.3.2. *Préliminaires.*

LEMME 8.14. *Soit X un espace vectoriel normé. Si $T \in \mathcal{L}(X)$ est compact et si $\lambda \in \mathbb{K}\backslash\{0\}$, alors* $\dim\ker(T-\lambda I) < \infty$.

Démonstration. Posons $E_\lambda := \ker(T-\lambda I)$. D'après le *Théorème de Riesz*, il suffit de montrer que B_{E_λ} est compacte ; ce qui revient à montrer que toute suite bornée $(x_n) \subseteq E_\lambda$ possède une sous-suite qui converge dans E_λ. Comme $\lambda \neq 0$, on peut écrire $x_n = \frac{1}{\lambda}Tx_n$. Comme T est compact, la suite (x_n) possède une sous-suite (x_{n_k}) telle que $Tx_{n_k} \to z \in X$. Alors $x_{n_k} \to x := \frac{1}{\lambda}z$, et $x \in E_\lambda$ car $E_\lambda = \ker(T-\lambda I)$ est *fermé* dans X. □

LEMME 8.15. *Soit $T \in \mathcal{L}(H)$. Si $E \subseteq H$ est un sous-espace vectoriel* invariant par *T, i.e. $T(E) \subseteq E$, alors $E^\perp$ est invariant par T^*. Si E est un* sous-espace propre *de T, alors E est invariant par tout opérateur $A \in \mathcal{L}(H)$ tel que $AT = TA$.*

Démonstration. Si $x \in E^\perp$ et $z \in E$, alors $\langle T^*x, z\rangle = \langle x, Tz\rangle = 0$ car $Tz \in E$; donc $T^*x \in E^\perp$. Si $E = \ker(T-\lambda I)$ pour un certain $\lambda \in \mathbb{K}$ et si $AT = TA$, alors $\forall x \in E\ :\ TAx = A(Tx) = A(\lambda x) = \lambda Ax$, donc $A(E) \subseteq E$. □

LEMME 8.16. *Si $T \in \mathcal{L}(H)$ est normal et si $\lambda \in \mathbb{K}$, alors* $\ker(T-\lambda I) = \ker(T^*-\bar\lambda I)$.

Démonstration. L'opérateur $T-\lambda I$ est normal, donc

$$\ker(T-\lambda I) = \ker\big((T-\lambda I)^*\big) = \ker(T^*-\bar\lambda I).$$

□

Corollaire 8.17. *Si $T \in \mathcal{L}(H)$ est normal, alors les sous-espaces propres de T sont deux à deux orthogonaux.*

Démonstration. Soient $x, y \in H$ vérifiant $Tx = \lambda x$ et $Ty = \mu y$ avec $\lambda \neq \mu$. On a d'une part $\langle Tx, y\rangle = \langle \lambda x, y\rangle = \lambda \langle x, y\rangle$; et d'autre part, d'après le lemme : $\langle Tx, y\rangle = \langle x, T^*y\rangle = \langle x, \bar{\mu} y\rangle = \mu \langle x, y\rangle$. Donc $\lambda \langle x, y\rangle = \mu \langle x, y\rangle$, et donc $\langle x, y\rangle = 0$. □

Exercice. Montrer que si $V \in \mathcal{L}(H)$ est une *isométrie*, alors les sous-espaces propres de V sont deux à deux orthogonaux.

Lemme 8.18. *Soit $T \in \mathcal{L}(H)$ un opérateur compact, et supposons $H \neq \{0\}$. Si T est auto-adjoint, alors $\|T\|$ ou $-\|T\|$ est valeur propre de T. Si $\mathbb{K} = \mathbb{C}$ et si T est normal, alors T admet une valeur propre λ telle que $|\lambda| = \|T\|$.*

Démonstration. Si $T = 0$, il n'y a rien à démontrer ; donc on supposera que $T \neq 0$. Dans la suite, on posera $r := \|T\|$.

Supposons d'abord que T soit auto-adjoint. Dans ce cas, on sait que

$$r = \sup\big\{|\langle Tx, x\rangle|;\ \|x\| = 1\big\}.$$

Comme $\langle Tx, x\rangle \in \mathbb{R}$ pour tout $x \in H$, on peut donc trouver une suite $(x_n)_{n\in\mathbb{N}} \subseteq H$ telle que $\|x_n\| = 1$ pour tout n et $\langle Tx_n, x_n\rangle \to \pm r$ quand $n \to \infty$. Supposons par exemple que $\langle Tx_n, x_n\rangle \to r$, et montrons que dans ce cas r est valeur propre de T.

Le point clé est que l'opérateur auto-adjoint $rI - T$ est *positif* : en effet, comme $r = \|T\|$, on a $\langle (rI - T)x, x\rangle = r\,\|x\|^2 - \langle Tx, x\rangle \geqslant 0$ pour tout $x \in H$.
Par définition de (x_n), on voit que

$$\langle (rI - T)x_n, x_n\rangle = r\,\|x_n\|^2 - \langle Tx_n, x_n\rangle = r - \langle Tx_n, x_n\rangle \xrightarrow{n\to\infty} 0.$$

Mais comme $rI - T$ est positif, on a

$$\|(rI - T)x_n\| \leqslant \|rI - T\|^{1/2} \langle (rI - T)x_n, x_n\rangle^{1/2},$$

d'après le Corollaire 3.10 ; donc en fait $\|(rI - T)x_n\| \to 0$.
Pour conclure, on utilise (enfin) le fait que T est compact : comme la suite (x_n) est bornée, on peut supposer, quitte à extraire une sous-suite, que la suite (Tx_n) est convergente, $Tx_n \to y \in H$. Comme $\|(rI - T)x_n\| \to 0$ et comme $r \neq 0$, on voit alors que la suite (x_n) converge vers $x := \frac{1}{r}\,y$. On a $\|x\| = 1$ car $\|x_n\| = 1$ pour tout n, et $Tx - rx = \lim_{n\to\infty}(T - rI)x_n = 0$; donc r est bien valeur propre de T.

Supposons maintenant que $\mathbb{K} = \mathbb{C}$ et que T soit un opérateur normal. L'opérateur $S := T^*T$ est positif, et il est aussi compact car T est compact ; donc $\|T^*T\| = r^2$ est valeur propre de T^*T d'après le "cas auto-adjoint" (on utilise le fait que les valeurs propres d'un opérateur positif sont nécessairement $\geqslant 0$). Posons $E := \ker(T^*T - r^2 I)$. Alors $E \neq \{0\}$, et $\dim(E) < \infty$ car T^*T est compact et $r^2 \neq 0$. De plus E est invariant par T car l'opérateur normal T commute avec T^*T. On peut donc définir $T_{|E} : E \to E$, et comme $\dim(E) < \infty$ et que $\mathbb{K} = \mathbb{C}$, on voit que $T_{|E}$ possède au moins une valeur propre λ. Il existe ainsi $x \in E\backslash\{0\}$ tels que $Tx = \lambda x$. Alors $T^*x = \bar{\lambda} x$ car T est normal, donc $r^2 x = T^*Tx = |\lambda|^2 x$, et donc $|\lambda| = r = \|T\|$. □

8.3.3. *Preuve du théorème de diagonalisation.* Soit $T \in \mathcal{L}(H)$ ou bien auto-adjoint, ou bien normal avec $\mathbb{K} = \mathbb{C}$. Posons $E := \operatorname{vect}\left(\bigcup_{\lambda\in\mathbb{K}} \ker(T-\lambda I)\right)$. Alors E est un sous-espace vectoriel de H invariant par T ; et E est également invariant par T^* car les $\ker(T-\lambda I)$ sont aussi des sous-espaces propres de T^*. Donc $F := E^\perp$ est un sous-espace fermé de H invariant par T et par T^*. On a $(T_{|F})^* = T^*_{|F}$ (exo), donc $T_{|F}$ est auto-adjoint ou normal. Donc, si on avait $F \neq \{0\}$, alors $T_{|F}$ possèderait une valeur propre, ce qui est exclu par définition de E. Ainsi, on a $E^\perp = \{0\}$; autrement dit :

$$\overline{\operatorname{vect}}\left(\bigcup_{\lambda\in\mathbb{K}} \ker(T-\lambda I)\right) = H.$$

Comme les $\ker(T-\lambda)$ sont deux à deux orthogonaux, on obtient donc une base hilbertienne de H formée de vecteurs propres pour T en choisissant, pour chaque $\lambda \in \mathbb{K}$ tel que $\ker(T-\lambda I) \neq \{0\}$, une base hilbertienne $(e_{i,\lambda})_{i\in I_\lambda}$ de $\ker(T-\lambda I)$.

Corollaire 8.19. *Si H est un espace de Hilbert complexe, séparable et de dimension infinie, alors tout opérateur compact normal $T \in \mathcal{L}(H)$ est unitairement équivalent à un opérateur diagonal $\Delta_a : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$, où $a \in c_0$.*

Démonstration. C'est évident par le théorème. □

Corollaire 8.20. *Si $\mathbb{K} = \mathbb{C}$, un opérateur compact normal $T \in \mathcal{L}(H)$ est auto-adjoint si et seulement si toutes ses valeurs propres sont réelles. Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, un opérateur compact auto-adjoint $T \in \mathcal{L}(H)$ est positif si et seulement si toutes ses valeurs propres sont $\geqslant 0$.*

Démonstration. Exo. □

8.3.4. *Cas des opérateurs à noyaux.* Dans cette section, la seule mesure considérée est la mesure de Lebesgue (en dimensions 1 et 2).

Proposition 8.21. *Soit I un intervalle de $\mathbb{R}$, et soit $K : I \times I \to \mathbb{C}$ un noyau appartenant à $L^2(I\times I)$ et vérifiant $K(y,x) = \overline{K(x,y)}$ pour tous $x, y \in I$.*

(1) *L'espace $L^2(I)$ possède une base hilbertienne $(\varphi_n)_{n\in\mathbb{N}}$ formée de vecteurs propres pour l'opérateur T_K.*
(2) *En notant $(\lambda_n)_{n\in\mathbb{N}}$ la suite de valeurs propres associées et en posant*
$$\psi_n(x,y) := \varphi_n(x)\overline{\varphi_n(y)},$$
on a $K = \sum_{n=0}^{\infty} \lambda_n \psi_n$, où la série converge dans $L^2(I\times I)$.
(3) *Avec les mêmes notations, on a*
$$\sum_{n=0}^{\infty} |\lambda_n|^2 = \|K\|^2_{L^2(I\times I)} = \int_{I\times I} |K(x,y)|^2 dxdy.$$
(4) *Supposons que l'intervalle I soit compact et que $K : I\times I \to \mathbb{C}$ soit continue.Alors, pour tout $n \in \mathbb{N}$ tel que $\lambda_n \neq 0$, la fonction φ_n "est" continue.*

(5) *Dans la situation de* (4), *on suppose de plus que la série* $\sum \lambda_n\, \varphi_n(x)\overline{\varphi_n(y)}$ *converge uniformément sur* $I \times I$. *Alors*

$$\forall (x,y) \in I \times I \;:\; K(x,y) = \sum_{n=0}^{\infty} \lambda_n\, \varphi_n(x)\overline{\varphi_n(y)}.$$

(6) *Sous les mêmes hypothèses que* (5), *la série* $\sum \lambda_n$ *converge et on a la* ***formule de trace***

$$\sum_{n=0}^{\infty} \lambda_n = \int_I K(x,x)\, dx.$$

Démonstration. (1) Comme $L^2(I)$ est séparable et $K \in L^2(I \times I)$, l'opérateur T_K est de Hilbert-Schmidt (donc compact) et auto-adjoint.
Pour (2) et (3), on a besoin du fait suivant.

Fait. Pour $(n,m) \in \mathbb{N}\times\mathbb{N}$, soit $e_{n,m} : I\times I \to \mathbb{C}$ la fonction définie par $e_{n,m}(x,y) := \varphi_n(x)\overline{\varphi_m}(y)$. La famille $\big(e_{n,m}\big)_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ est une base hilbertienne de $L^2(I \times I)$.

Preuve du Fait. On vérifie sans difficulté que la famille $(e_{n,m})$ est orthonormale (exo) ; donc il suffit de montrer que si $F \in L^2(I \times I)$ est orthogonale à toutes les $e_{n,m}$, alors $F = 0$. Mais

$$\langle F, e_{n,m}\rangle_{L^2(I\times I)} = \int_I \left(\int_I F(x,y)\varphi_m(y)\, dy\right) \overline{\varphi_n(x)}\, dx = \langle T_F\varphi_m, \varphi_n\rangle_{L^2(I)};$$

donc, si F est orthogonale à toutes les $e_{n,m}$, alors $T_F = 0$ puisque les φ_i forment une base hilbertienne de $L^2(I)$; et donc $F = 0$ car $\|F\|_{L^2(I\times I)} = \|T_F\|_{\mathrm{HS}}$. □

Si $(n,m) \in \mathbb{N} \times \mathbb{N}$, alors (*cf* plus haut)

$$\langle K, e_{n,m}\rangle_{L^2(I\times I)} = \langle T_K\varphi_m, \varphi_n\rangle_{L^2(I)} = \lambda_m\, \langle \varphi_m, \varphi_n\rangle_{L^2(I)};$$

autrement dit, $\langle K, e_{n,m}\rangle = 0$ si $m \neq n$ et $\langle K, e_{n,n}\rangle = \lambda_n$. Par le Fait et comme $e_{n,n} = \psi_n$, on en déduit (2) et (3) en même temps.
(4) Comme $\lambda_n \neq 0$, on peut écrire $\varphi_n = \frac{1}{\lambda_n} T_K\varphi_n$. Mais comme K est continue et I compact, la formule $T_K f(x) = \int_I K(x,y) f(y)\, dy$ définit une fonction continue pour toute $f \in L^1(I)$ (exo), donc en particulier pour toute $f \in L^2(I)$; et donc, φ_n "est" continue.
(5) Comme I est borné, la convergence uniforme entraîne la convergence dans $L^2(I)$. Donc, par (3), la fonction $F(x,y) := \sum_{n=0}^{\infty} \lambda_n\, \varphi_n(x)\overline{\varphi_n(y)}$ est presque partour égale à K. De plus, F est continue, par (4) et par la convergence uniforme de la série. Comme K est également continue, on a $K(x,y) = F(x,y)$ pour tout $(x,y) \in I \times I$.
(6) Par (5), on a $K(x,x) = \sum_{n=0}^{\infty} \lambda_n\, |\varphi_n(x)|^2$ pour tout $x \in I$, avec convergence uniforme de la série. Comme I est borné, on peut intégrer terme à terme, ce qui donne (6) car $\int_I |\varphi_n(x)|^2 dx = 1$ pour tout $n \in \mathbb{N}$. □

Remarque. Dans la situation de (4), on peut montrer que si l'opérateur T_K est *positif*, alors la série $\sum \lambda_n \phi_n(x)\overline{\phi_n(y)}$ converge uniformément sur $I \times I$. Ce résultat s'appelle d'habitude le *Théorème de Mercer.*

Voici une illustration de la Proposition 8.21, donnée sous forme d'exercice.

Exercice 1.1. Soit $K : [0,1] \times [0,1] \to \mathbb{R}$ le noyau (continu) défini par $K(x,y) := \min(x,y) - xy$.

(1) Montrer que si $f \in L^2([0,1])$ est continue, alors il existe unique fonction $g \in \mathcal{C}^2([0,1])$ vérifiant $g'' = -f$ et $g(0) = 0 = g(1)$, et que cette fonction g est égale à $T_K f$.

(2) En utilisant (1), établir les faits suivants.

   (i) L'opérateur T_K est à image dense et injectif.

   (ii) Si $\lambda \neq 0$ alors une fonction $f \in L^2([0,1])$ vérifie $T_K f = \lambda f$ si et seulement si elle est de classe $\mathcal{C}^2$ et solution de l'équation différentielle $f'' = -\frac{1}{\lambda} f$ avec "conditions aux limites" $f(0) = 0 = f(1)$.

(3) *Déduire de* (2) que les fonctions $e_n(t) := \sqrt{2}\sin(\pi n t)$, $n \geqslant 1$ forment une base hilbertienne de $L^2([0,1])$, et démontrer les formules suivantes :

$$\forall x, y \in [0,1] \ : \ \sum_{n=1}^{\infty} \frac{\sin(n\pi x)\sin(n\pi y)}{n^2} = \frac{\pi^2}{2}\big(\min(x,y) - xy\big);$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \qquad \text{et} \qquad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

9. Opérateurs sur l'espace de Hardy H^2

9.1. L'espace H^2.

NOTATION. On pose $\mathbb{D} := \{z \in \mathbb{C};\ |z| < 1\}$, et on note $H(\mathbb{D})$ l'ensemble de toutes les fonctions holomorphes sur $\mathbb{D}$. Si $f \in H(\mathbb{D})$, on écrit

$$f(z) = \sum_{n=0}^{\infty} c_n(f) z^n.$$

DÉFINITION 9.1. *Pour toute fonction $f \in H(\mathbb{D})$, on pose*

$$\|f\|^2_{H^2(\mathbb{D})} := \sum_{n=0}^{\infty} |c_n(f)|^2;$$

et on note $H^2(\mathbb{D})$ l'ensemble des $f \in H(\mathbb{D})$ vérifiant $\|f\|^2_{H^2(\mathbb{D})} < \infty$.

Exercice. Montrer que $H^2(\mathbb{D})$ est un espace vectoriel et que $\|\cdot\|_{H^2(\mathbb{D})}$ est une norme.

NOTATION. On note $\mathbf{z}$ la fonction $\mathbb{D} \ni z \mapsto z$.

LEMME 9.2. *Soit Soit $U : H^2(\mathbb{D}) \to \ell^2(\mathbb{N})$ l'application (linéaire) définie par*

$$Uf := \big(c_n(f)\big)_{n\geqslant 0}.$$

Alors U est une isométrie bijective, et $U\mathbf{z}^n = e_n$ pour tout $n \in \mathbb{N}$, où (e_n) est la base canonique de $\ell^2(\mathbb{N})$.

Démonstration. Par définition, U est une isométrie et $U\mathbf{z}^n = e_n$ pour tout $n \in \mathbb{N}$. Pour montrer la surjectivité, on observe que si $x = (c_n)_{n\geqslant 0} \in \ell^2(\mathbb{N})$, alors le rayon de convergence de la série entière $\sum c_n z^n$ est au moins égal à 1 car la suite (c_n) est bornée ; donc $f(z) := \sum_{n=0}^{\infty} c_n z^n$ est une fonction holomorphe bien définie sur $\mathbb{D}$, et par définition $f \in H^2(\mathbb{D})$ et $Uf = x$. □

Corollaire 9.3. *$H^2(\mathbb{D})$ est un espace de Hilbert et les fonction $\mathbf{z}^n$, $n \in \mathbb{N}$ forment une base hilbertienne de $H^2(\mathbb{D})$. Si $f \in H^2(\mathbb{D})$, alors*

$$f = \sum_{n=0}^{\infty} c_n(f)\mathbf{z}^n \quad \textit{où la série converge dans } H^2(\mathbb{D}).$$

Si $f, g \in H^2(\mathbb{D})$, alors

$$\langle f, g\rangle_{H^2(\mathbb{D})} = \sum_{n=0}^{\infty} c_n(f)\overline{c_n(g)}.$$

Proposition 9.4. *La convergence dans $H^2(\mathbb{D})$ entraine la convergence uniforme sur tout compact. En particulier, pour tout $a \in \mathbb{D}$, la forme linéaire $\delta_a : f \mapsto f(a)$ est continue sur $H^2(\mathbb{D})$*

Démonstration. Il suffit (exo) de montrer que pour tout compact $K \subseteq \mathbb{D}$, il existe une constante C_K telle que

$$\forall f \in H^2(\mathbb{D}) \ : \ \sup_{z\in K} |f(z)| \leqslant C_K \, \|f\|_{H^2(\mathbb{D})}.$$

Si $z \in K$, alors

$$|f(z)|^2 = \left|\sum_{n=0}^{\infty} c_n(f)z^n\right|^2 \leqslant \left(\sum_{n=0}^{\infty} |c_n(f)|^2\right) \times \left(\sum_{n=0}^{\infty} |z|^{2n}\right) = \frac{1}{1-|z|^2}\,\|f\|_{H^2};$$

d'où le résultat avec $C_K := \sup_{z\in K} \frac{1}{1-|z|^2}$. □

Corollaire 9.5. *Pour tout $a \in \mathbb{D}$, il existe une unique fonction $k_a \in H^2(\mathbb{D})$ telle que*

$$\forall f \in H^2(\mathbb{D}) \ : \ f(a) = \langle f, k_a\rangle_{H^2(\mathbb{D})}.$$

On dit que k_a est le **noyau reproduisant au point** *a pour l'espace $H^2(\mathbb{D})$.*

Démonstration. C'est le "Théorème de représentation de Riesz" pour la forme linéaire continue δ_a. □

Lemme 9.6. *Pour tout $a \in \mathbb{D}$, on a* $k_a(z) = \dfrac{1}{1-\bar{a}z}$.

Démonstration. Comme $(\mathbf{z}^n)_{n\geqslant 0}$ est une base hilbertienne de $H^2(\mathbb{D})$, on peut écrire $k_a = \sum_{n=0}^{\infty} c_n(k_a)\mathbf{z}^n$, où la série converge dans $H^2(\mathbb{D})$. Mais

$$c_n(k_a) = \langle k_a, \mathbf{z}^n\rangle_{H^2(\mathbb{D})} = \overline{\langle \mathbf{z}^n, k_a\rangle_{H^2(\mathbb{D})}} = \overline{\mathbf{z}^n(a)} = \bar{a}^n.$$

Comme la convergence dans $H^2(\mathbb{D})$ entraine la convergence simple, on a donc

$$\forall z \in \mathbb{D} \ : \ k_a(z) = \sum_{n=0}^{\infty} \bar{a}^n z^n = \frac{1}{1-\bar{a}z}.$$

□

Exercice. Calculer $\|k_a\|_{H^2}$.

PROPOSITION 9.7. *Pour toute fonction $f \in H(\mathbb{D})$, on a*

$$\|f\|^2_{H^2(\mathbb{D})} = \sup_{0\leqslant r<1} \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi} = \lim_{r\to 1^-} \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi}.$$

En particulier : $f \in H^2(\mathbb{D}) \iff \sup_{0\leqslant r<1} \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi} < \infty.$

Démonstration. Pour $0 \leqslant r < 1$, soit $u_r : \mathbb{R} \to \mathbb{C}$ la fonction définie par

$$u_r(\theta) := f(re^{i\theta}) = \sum_{n=0}^{\infty} c_n(f) r^n e^{in\theta}.$$

Par convergence normale de la série, la fonction u_r est continue, 2π-périodique, et ses coefficients de Fourier sont donnés par $\widehat{u_r}(n) = 0$ si $n < 0$ et $\widehat{u_r}(n) = c_n(f) r^n$ si $n \neq 0$. Par la formule de Parseval, on a donc

$$\int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi} = \sum_{n=0}^{\infty} |c_n(f)|^2 r^{2n} \qquad \text{pour tout } 0 \leqslant r < 1.$$

D'où le résultat par le Théorème de convergence monotone pour les sommes, car le membre de droite est une fonction croissante de $r \in [0, 1[$. □

COROLLAIRE 9.8. *Si $f : \mathbb{D} \to \mathbb{C}$ est une fonction holomorphe* bornée, *alors $f \in H^2(\mathbb{D})$ et $\|f\|_{H^2} \leqslant \|f\|_\infty$.*

Démonstration. Exo. □

NOTATION. Si $u \in L^2(\mathbb{T})$, on note $\widehat{u}(n)$ les coefficients de Fourier de u :

$$\forall n \in \mathbb{Z} \; : \; \widehat{u}(n) = \int_{\mathbb{T}} u(z) z^{-n} dm(z) = \int_0^{2\pi} u(e^{i\theta})\, e^{-in\theta} \frac{d\theta}{2\pi}.$$

DÉFINITION 9.9. *On pose*

$$H^2(\mathbb{T}) := \big\{u \in L^2(\mathbb{T});\ \widehat{u}(n) = 0 \quad \textit{pour tout } n < 0\big\}.$$

Remarque. Par définition, $H^2(\mathbb{T})$ est un sous-espace fermé de $L^2(\mathbb{T})$, et donc un espace de Hilbert.

LEMME 9.10. *Si $u \in H^2(\mathbb{T})$, alors la fonction f définie par $f(z) := \sum_0^\infty \widehat{u}(n) z^n$ appartient à $H^2(\mathbb{D})$ et $\|f\|_{H^2(\mathbb{D})} = \|u\|_{L^2(\mathbb{T})}$. Inversement, si $f \in H^2(\mathbb{D})$, alors il existe une unique $u \in H^2(\mathbb{T})$ telle que $\widehat{u}(n) = c_n(f)$ pour tout $n \geqslant 0$. On dit que u est la* ***valeur au bord de*** *f, et on écrit $u = bf$.*

Démonstration. La première partie est simplement la formule de Parseval pour u. La deuxième partie vient de la complétude de $L^2(\mathbb{T})$: si $f \in H^2(\mathbb{D})$, les sommes partielles de la série $\sum_{n\geqslant 0} c_n(f) e^{in\theta}$ forment une suite de Cauchy dans $L^2(\mathbb{T})$, donc cette série converge dans $L^2(\mathbb{T})$ vers une fonction u qui convient. □

COROLLAIRE 9.11. *Les deux espaces de Hilbert $H^2(\mathbb{T})$ et $H^2(\mathbb{D})$ sont "canoniquement" isométriques via l'application $H^2(\mathbb{T}) \ni u \mapsto f(z) = \sum_{n=0}^\infty \widehat{u}(n) z^n \in H^2(\mathbb{D})$ et son inverse $H^2(\mathbb{D}) \ni f \mapsto bf \in H^2(\mathbb{T})$.*

COROLLAIRE 9.12. *Si $f, g \in H^2(\mathbb{D})$, alors*

$$\langle f, g\rangle_{H^2(\mathbb{D})} = \langle bf, bg\rangle_{L^2(\mathbb{T})} = \int_0^{2\pi} bf(e^{i\theta})\, \overline{bg(e^{i\theta})} \frac{d\theta}{2\pi}.$$

ATTENTION. Si $f \in H^2(\mathbb{D})$, alors $bf(e^{i\theta})$ n'est définie que presque partout.

NOTATION. Si $f \in H^2(\mathbb{D})$ et $0 \leqslant r < 1$, on note $f_r : \mathbb{T} \to \mathbb{C}$ la fonction définie par

$$f_r(e^{i\theta}) := f(re^{i\theta}) = \sum_{n=0}^{\infty} c_n(f) r^n e^{in\theta}.$$

On a $f_r \in \mathcal{C}(\mathbb{T}) \cap H^2(\mathbb{T})$ et $\widehat{f_r}(n) = c_n(f) r^n$ pour tout $n \geqslant 0$.

PROPOSITION 9.13. *Si $f \in H^2(\mathbb{D})$, alors $f_r \to bf$ en norme $L^2(\mathbb{T})$ quand $r \to 1^-$.*

Démonstration. Par la formule de Parseval, on a

$$\|bf - f_r\|_{L^2}^2 = \sum_{n=0}^{\infty} (1 - r^n)^2 \, |c_n(f)|^2;$$

d'où le résultat par convergence dominée (pour les séries). □

COROLLAIRE 9.14. *Si $F : \overline{\mathbb{D}} \to \mathbb{C}$ est une fonction continue sur $\overline{\mathbb{D}}$, alors $f := F_{|\mathbb{D}} \in H^2(\mathbb{D})$ et $bf = F_{|\mathbb{T}}$.*

Démonstration. La fonction f est holomorphe et *bornéee* car F est bornée sur le compact $\overline{\mathbb{D}}$; donc $f \in H^2(\mathbb{D})$. Ensuite, F est uniformément continue sur $\overline{\mathbb{D}}$, donc $f_r(e^{i\theta}) = F(re^{i\theta}) \to F(e^{i\theta})$ uniformément quand $r \to 1$, et donc $f_r \to F_{|\mathbb{T}}$ en norme $L^2(\mathbb{T})$; ce qui prouve que $bf = F_{|\mathbb{T}}$ par la proposition. □

COROLLAIRE 9.15. *Si $a \in \mathbb{D}$ alors bk_a "est" continue sur $\mathbb{T}$ et $bk_a(e^{i\theta}) \equiv \frac{1}{1-\bar{a}e^{i\theta}}$.*

COROLLAIRE 9.16. *Si $f \in H^2(\mathbb{D})$, alors f vérifie la la "formule de Cauchy au bord" :*

$$\forall z \in \mathbb{D} \; : \; f(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{bf(e^{i\theta})}{e^{i\theta} - z} \, e^{i\theta} d\theta.$$

Démonstration. On a $f(z) = \langle f, k_z \rangle_{H^2(\mathbb{D})} = \langle bf, bk_z \rangle_{L^2(\mathbb{T})}$; ce qui donne le résultat en écrivant explicitement $\langle bf, bk_z \rangle_{L^2(\mathbb{T})}$. □

COROLLAIRE 9.17. *Si $f \in H^2(\mathbb{D})$ alors, pour toute suite (r_n) tendant vers 1^-, on peut trouver une sous-suite (r_{n_k}) telle que $f(r_{n_k} e^{i\theta}) \to bf(e^{i\theta})$ pp.*

Démonstration. La convergence L^2 entraine la convergence pp d'une sous-suite. □

REMARQUE. En fait, on a un résultat "bien meilleur" que le Corollaire 9.17, qu'on appelle le **Théorème de Fatou** : *Si $f \in H^2(\mathbb{D})$, alors $f(re^{i\theta}) \to bf(e^{i\theta})$ presque partout quand $r \to 1^-$.* (Inutile d'extraire une sous-suite.) La preuve est plus délicate que tout ce qui a été vu jusqu'à maintenant, et on ne la fera pas.

9.2. Espace H^∞ et opérateurs de multiplication sur H^2.

DÉFINITION 9.18. *On note $H^\infty(\mathbb{D})$ l'ensemble des fonctions holomorphes bornées sur $\mathbb{D}$.*

Remarque 1. On a vu que $H^\infty(\mathbb{D}) \subseteq H^2(\mathbb{D})$.

Remarque 2. Si $\phi \in H^\infty(\mathbb{D})$, alors $\phi f \in H^2(\mathbb{D})$ pour toute $f \in H^2(\mathbb{D})$ et $\|\phi f\|_{H^2} \leqslant \|\phi\|_\infty \|f\|_{H^2}$. Autrement dit, on définit un opérateur borné $M_\phi : H^2(\mathbb{D}) \to H^2(\mathbb{D})$ en posant $M_\phi f := \phi f$ pour toute $f \in H^2(\mathbb{D})$, et on a $\|M_\phi\| \leqslant \|\phi\|_\infty$.

Démonstration. Pour tout $0 \leqslant r < 1$, on a

$$\int_0^{2\pi} |(\phi f)(re^{i\theta})|^2 \,\frac{d\theta}{2\pi} \leqslant \|\phi\|_\infty^2 \int_0^{2\pi} |f(re^{i\theta})|^2 \,\frac{d\theta}{2\pi};$$

donc $\|\phi f\|_{H^2} \leqslant \|\phi\|_\infty \, \|f\|_{H^2}$ en faisant $r \to 1^-$.

□

PROPOSITION 9.19. *Si $\phi \in H^\infty(\mathbb{D})$, alors $\|M_\phi\| = \|\phi\|_\infty$.*

Démonstration. On a vu que $\|M_\phi\| \leqslant \|\phi\|_\infty$. L'inégalité inverse va se déduire du fait suivant.

FAIT. Pour tout $a \in \mathbb{D}$, on a $M_\phi^* k_a = \overline{\phi(a)}\, k_a$.

Preuve du Fait. Si $f \in H^2(\mathbb{D})$, alors

$$\langle M_\phi^* k_a, f\rangle = \langle k_a, M_\phi f\rangle = \overline{\langle \phi f, k_a\rangle} = \overline{\phi(a) f(a)} = \overline{\phi(a)}\,\langle k_a, f\rangle = \langle \overline{\phi(a)}\, k_a, f\rangle.$$

□

Par le Fait et comme $k_a \neq 0$, on voit que $\|M_\phi^*\| \geqslant |\overline{\phi(a)}| = |\phi(a)|$ pour tout $a \in \mathbb{D}$, et donc $\|M_\phi\| = \|M_\phi^*\| \geqslant \|\phi\|_\infty$. □

COROLLAIRE 9.20. *Si $\phi \in H^\infty(\mathbb{D})$, alors $\|\phi\|_\infty = \|b\phi\|_{L^\infty(\mathbb{T})}$.*

Démonstration. On sait qu'on peut trouver une suite (r_n) tendant vers 1^- telle que $\phi(r_n e^{i\theta}) \to b\phi(e^{i\theta})$ pp. Comme $|\phi(r_n e^{i\theta})| \leqslant \|\phi\|_\infty$ pour tout n et pour tout θ, on en déduit $\|b\phi(e^{i\theta})\|_{L^\infty(\mathbb{T})} \leqslant \|\phi\|_\infty$.
Pour l'inégalité inverse, il suffit de montrer que $\|M_\phi\| \leqslant \|b\phi\|_{L^\infty(\mathbb{T})}$. On utilise pour cela le fait suivant.

FAIT. Pour toute $f \in H^2(\mathbb{D})$, on a $b(\phi f) = (b\phi)\,(bf)$.

Preuve du Fait. On peut trouver une suite (r_n) tendant vers 1^- telle que $\phi(r_n e^{i\theta}) \to b\phi(e^{i\theta})$ pp, $f(r_n e^{i\theta}) \to bf(e^{i\theta})$ pp et $(\phi f)(r_n e^{i\theta}) \to b(\phi f)(e^{i\theta})$ pp ; d'où le résultat puisque $(\phi f)(r_n e^{i\theta}) = \phi(r_n e^{i\theta})\, f(r_n e^{i\theta})$. □

Par le Fait, on a

$$\begin{aligned}
\forall f \in H^2(\mathbb{D})\ :\ \|M_\phi f\|_{H^2(\mathbb{D})} &= \|\phi f\|_{H^2(\mathbb{D})}\\
&= \|b(\phi f)\|_{L^2(\mathbb{T})}\\
&= \|(b\phi)\,(bf)\|_{L^2(\mathbb{T})}\\
&\leqslant \|b\phi\|_{L^\infty(\mathbb{T})}\, \|f\|_{L^2(\mathbb{T})} = \|b\phi\|_{L^\infty(\mathbb{T})}\, \|bf\|_{H^2(\mathbb{D})};
\end{aligned}$$

donc en effet $\|M_\phi\| \leqslant \|b\phi\|_{L^\infty(\mathbb{T})}$. □

DÉFINITION 9.21. *On dit qu'une fonction $\varphi \in H^\infty(\mathbb{D})$ est **intérieure** si on a $|b\varphi(e^{i\theta})| = 1$ pp.*

Remarque. Si ϕ est une fonction intérieure non constante, alors $|\phi(z)| \leqslant 1$ pour tout $z \in \mathbb{D}$ d'après 9.20 ; et en fait $|\phi(z)| < 1$ pour tout $z \in \mathbb{D}$ si ϕ n'est pas constante, d'après le principe du maximum (exo).

EXEMPLE 9.22. Pour $a \in \mathbb{D}$, la fonction φ_a définie par $\varphi_a(z) := \frac{a-z}{1-\bar{a}z}$ est intérieure. De plus, φ_a est un **automorphisme de** $\mathbb{D}$, *i.e.* une bijection holomorphe de $\mathbb{D}$ sur $\mathbb{D}$, et $\varphi_a^{-1} = \varphi_a$.

Démonstration. La formule définissant φ_a a un sens pour tout $z \in \mathbb{C}\backslash\{1/\bar{a}\}$, donc définit en particulier une fonction continue sur $\overline{\mathbb{D}}$. Donc $\varphi_a \in H^\infty(\mathbb{D})$ et $b\varphi_a(e^{i\theta}) = \frac{a-e^{i\theta}}{1-\bar{a}e^{i\theta}}$ pour tout $e^{i\theta} \in \mathbb{T}$. En factorisant par $e^{i\theta}$ au dénominateur, on en déduit que $|b\varphi_a(e^{i\theta})| = \left|\frac{e^{i\theta}-a}{e^{-i\theta}-\bar{a}}\right| \equiv 1$ sur $\mathbb{T}$; donc φ_a est intérieure. Comme φ_a n'est pas constante, on a $|\varphi_a(z)| < 1$ pour tout $z \in \mathbb{D}$, *i.e.* $\varphi_a(\mathbb{D}) \subseteq \mathbb{D}$. En résolvant l'équation $\varphi_a(z) = w$, on voit que φ_a est une bijection de $\mathbb{D}$ sur $\mathbb{D}$ avec $\varphi_a^{-1} = \varphi_a$ (exo). □

REMARQUE. On "sait bien" que tout automorphisme de $\mathbb{D}$ est de la forme $c\,\varphi_a(z)$, où $|c| = 1$.

Exercice. Montrer que si $\phi \in H^\infty(\mathbb{D})$ est une fonction intérieure, alors l'opérateur $M_\phi : H^2(\mathbb{D}) \to H^2(\mathbb{D})$ est une isométrie.

9.3. Opérateurs de composition.

NOTATION. On note $H(\mathbb{D}, \mathbb{D})$ l'ensemble de toutes les fonctions $\varphi \in H(\mathbb{D})$ telles que $\varphi(\mathbb{D}) \subseteq \mathbb{D}$.

Remarque. Si $\varphi \in H(\mathbb{D}, \mathbb{D})$ et $f \in H(\mathbb{D})$, alors $C_\varphi := f \circ \varphi \in H(\mathbb{D})$. On dit que l'application linéaire $C_\varphi : H(\mathbb{D}) \to H(\mathbb{D})$ est l'**opérateur de composition** associé à φ.

THÉORÈME 9.23. *Si $\varphi \in H(\mathbb{D}, \mathbb{D})$, alors $C_\varphi f = f \circ \varphi \in H^2$ pour toute $f \in H^2$ et $C_\varphi : H^2 \to H^2$ est continu, avec $\|C_\varphi\| \leqslant \sqrt{\frac{1+|\varphi(0)|}{1-|\varphi(0)|}}$.*

Démonstration. Remarquons d'abord que $\varphi \in H^\infty$, donc $C_\varphi \mathbf{z}^n = \varphi^n \in H^\infty \subseteq H^2$ pour tout $n \in \mathbb{N}$, et donc $C_\varphi f \in H^2$ pour toute fonction *polynomiale* $f \in H^2$. Comme les fonctions polynomiales sont denses dans H^2, il suffit de montrer qu'on a

$$\|C\varphi f\|^2_{H^2(\mathbb{D})} \leqslant \frac{1+|\varphi(0)|}{1-|\varphi(0)|}\,\|f\|^2_{H^2(\mathbb{D})} \qquad \text{pour toute } f \in H^2 \text{ polynomiale.}$$

CAS 1. $\varphi(0) = 0$.

On veut dans ce cas montrer que

$$\|C_\varphi f\|_{H^2} \leqslant \|f\|_{H^2} \qquad \text{pour toute } f \in H^2 \text{ polynomiale.}$$

Notons $B : H^2 \to H^2$ le "backward shift" sur H^2 :

$$B\Big(\sum_{n=0}^\infty c_n \mathbf{z}^n\Big) := \sum_{n=1}^\infty c_n \mathbf{z}^{n-1}.$$

Si $f \in H^2$ est polynomiale, $f = \sum\limits_{n=0}^{d} c_n \mathbf{z}^n$, alors

$$\begin{aligned} C_\varphi f &= \sum_{n=0}^{d} c_n \varphi^n \\ &= c_0 \mathbf{1} + \varphi \sum_{n=1}^{d} c_n \varphi^{n-1} \\ &= c_0 \mathbf{1} + \varphi \times C_\varphi(Bf). \end{aligned}$$

Comme $\varphi(0) = 0$, la fonction $g := \varphi \times C_\varphi(Bf) \in H^2$ vérifie également $g(0) = 0$, donc $\langle g, \mathbf{1}\rangle_{H^2} = 0$. Par conséquent

$$\begin{aligned} \|C_\varphi f\|_{H^2}^2 &= |c_0|^2 + \|\varphi \times C_\varphi(Bf)\|_{H^2}^2 \\ &\leqslant |c_0|^2 + \|C_\varphi(Bf)\|_{H^2}^2 \qquad \text{car } \|\varphi\|_\infty \leqslant 1. \end{aligned}$$

Comme Bf est une fonction polynomiale de terme constant c_1, on peut appliquer ceci à Bf, et on obtient $\|C_\varphi f\|_{H^2}^2 \leqslant |c_0|^2 + |c_1|^2 + \|C_\varphi(B^2 f)\|_{H^2}^2$. Plus généralement, comme $B^k f$ est une fonction polynomiale de terme constant c_k pour tout $k \in [\![0, d]\!]$ et comme $B^{d+1} f = 0$, on peut répéter d fois ce raisonnement, ce qui donne

$$\|C_\varphi f\|_{H^2}^2 \leqslant |c_0|^2 + \cdots + |c_d|^2 + 0 = \|f\|_{H^2}^2.$$

Cas 2. La fonction φ est intérieure.

Posons $r := |\varphi(0)|$. Il suffit de montrer que la matrice $A := \big(\langle C_\varphi \mathbf{z}^i, C_\varphi \mathbf{z}^j\rangle\big)_{i,j\in\mathbb{N}}$ définit un opérateur borné sur $\ell^2(\mathbb{N})$, avec $\|A\| \leqslant C := \frac{1+r}{1-r}\cdot$ En effet, si ce point est acquis, on aura pour toute fonction polynomiale $f(z) = \sum\limits_{n=0}^{d} c_n z^n$:

$$\begin{aligned} \|C_\varphi f\|_{H^2}^2 &= \sum_{i,j=0}^{d} c_i \overline{c_j} \, \langle C_\varphi \mathbf{z}^i, C_\varphi \mathbf{z}^j\rangle \\ &= \Big\langle A\Big(\sum_{n=0}^{d} c_n e_n\Big), \sum_{n=0}^{d} c_n e_n \Big\rangle_{\ell^2} \\ &\leqslant C \,\Big\| \sum_{n=0}^{d} c_n e_n \Big\|_{\ell^2}^2 = C\, \|f\|_{H^2}^2. \end{aligned}$$

Si $i, j \in \mathbb{N}$, alors

$$\begin{aligned} \langle C_\varphi \mathbf{z}^i, C_\varphi \mathbf{z}^j\rangle &= \langle \varphi^i, \varphi^j\rangle \\ &= \int_{\mathbb{T}} (b\varphi)^i \overline{(b\varphi)^j}\, dm \\ &= \int_{\mathbb{T}} (b\varphi)^{i-j}\, dm \quad \text{car } |b\varphi| = 1 \text{ pp.} \end{aligned}$$

D'après la formule de Cauchy, on en déduit que

$$\langle C_\varphi \mathbf{z}^i, C_\varphi \mathbf{z}^j\rangle = \varphi(0)^{i-j} \qquad \text{si } i > j,$$

et (en utilisant à nouveau que $|b\varphi| = 1$ pp)

$$\langle C_\varphi \mathbf{z}^i, C_\varphi \mathbf{z}^j \rangle = \int_{\mathbb{T}} \overline{(b\varphi)^{j-i}} dm = \overline{\int_{\mathbb{T}} (b\varphi)^{j-i} dm} = \overline{\varphi(0)}^{\,j-i} \quad \text{si } j \geqslant i.$$

Comme on a posé $r := |\varphi(0)|$, on a donc

$$|\langle C_\varphi \mathbf{z}^i, C_\varphi \mathbf{z}^j \rangle| \leqslant r^{|j-i|} \qquad \text{pour tous } i, j \in \mathbb{N}.$$

D'après l'Exemple 2.19, on en déduit que la matrice $A = \big(\langle C_\varphi \mathbf{z}^i, C_\varphi \mathbf{z}^j \rangle\big)_{i,j \in \mathbb{N}}$ définit en effet un opérateur borné sur $\ell^2(\mathbb{N})$ avec $\|A\| \leqslant \frac{1+r}{1-r}\cdot$

Cas 3. Cas général.

Posons $a := \varphi(0)$ et $\psi := \varphi_a \circ \varphi$, où $\varphi_a(z) = \frac{a-z}{1-\bar{z}}\cdot$ Alors $\varphi = \varphi_a^{-1} \circ \psi = \varphi_a \circ \psi$, et donc $C_\varphi = C_{\varphi_a} \circ C_\psi$. Comme $\psi(0) = \varphi_a(a) = 0$ et comme φ_a est intérieure avec $\varphi_a(0) = a = \varphi(0)$, on en déduit le résultat souhaité d'après les 2 premiers cas. $\square$

Chapitre 2

Un peu de topologie “générale”

1. Vocabulaire

1.1. Espaces topologiques.

RAPPEL. Une **topologie** sur un ensemble X est une famille τ de parties de X contenant $\emptyset$ et X et qui est stable *réunions quelconques* et par *intersections finies*. Les éléments de τ s'appellent les **ouverts** de la topologie τ. Un **espace topologique** est un ensemble X muni d'une topologie.

EXEMPLE 1. Toute distance d sur X définit une topologie τ_d : un ensemble $O \subseteq X$ est τ_d - ouvert ssi, pour tout $x \in O$, on peut trouver $r > 0$ tel que $B(x, r) \subseteq O$.

EXEMPLE 2. Si X est un ensemble quelconque, alors $\tau_{\min} := \{\emptyset, X\}$ et $\tau_{\max} := \mathcal{P}(X)$ sont des topologies, qui sont respectivement la plus petite et la plus grande topologie sur X. La topologie $\tau_{\min}$ s'appelle la **topologie grossière**, et $\tau_{\max}$ s'appelle la **topologie discrète**. (La topologie discrète est beaucoup plus “importante” que la topologie grossière.)

REMARQUE. Notions topologiques les plus “primitives”.

- Ouverts, fermés.
- Voisinages d'un point $a \in X$. *Notation* : $\mathcal{V}(a) := \big\{V \subseteq X;\ V \text{ voisinage de } a\big\}$.
- Intérieur, adhérence d'un ensemble $A \subseteq X$. *Notations* : $\mathring{A}$ et $\overline{A}$.
- Continuité d'une application $f : X \to Y$.
  - f est continue sur X ssi $f^{-1}(V)$ est ouvert dans X pour tout ouvert $V \subseteq Y$.
  - f est continue *en un point* $a \in X$ ssi : pour tout ouvert $V \ni f(a)$, on peut trouver un ouvert $U \ni a$ tel que $f(U) \subseteq V$; autrement dit : pour tout voisinage V de $f(a)$, $f^{-1}(V)$ est un voisinage de a.

DÉFINITION 1.1. *Soit (X, τ) un espace topologique, et soit $E \subseteq X$. La famille de tous les ensembles $O \subseteq E$ de la forme $O = \widetilde{O} \cap E$ où $\widetilde{O}$ est un ouvert de X, est une topologie sur E, qu'on appelle la topologie* ***induite*** *par la topologie τ. On la note τ_E.*

Exercice. Vérifier que τ_E est bien une topologie, et que si $\tau = \tau_d$ pour une certaine distance d, alors τ_E est la topologie définie par $d_{|E\times E}$.

1.2. Bases.

DÉFINITION 1.2. *Soit (X, τ) un espace topologique.*

(1) *Étant donné $a \in X$, on dit qu'une famille $\mathcal{V}$ de voisinages de a est une* ***base de voisinages*** *pour a si, pour tout voisinage V de a, on peut trouver $V' \in \mathcal{V}$ tel que $V' \subseteq V$.*

(2) *On dit qu'une famille d'ouverts $\mathcal{B}$ est une* ***base pour la topologie*** *τ si, pour tout ouvert $U \subseteq X$ et pour tout $a \in U$, on peut trouver $B \in \mathcal{B}$ tel que $a \in B$ et $B \subseteq U$; autrement dit : si tout ouvert non-vide $U \subseteq X$ est réunion d'éléments de $\mathcal{B}$.*

Exemples. Soit (X,d) un espace métrique.

(1) si $a \in X$, alors $\mathcal{V} := \{B(a,1/n);\ n \in \mathbb{N}^*\}$ est une base de voisinages pour a. En particulier, tout point $a \in X$ possède une base *dénombrable* de voisinages.

(2) La famille $\mathcal{B} := \{B(a,1/n);\ a \in X, n \in \mathbb{N}^*\}$ est une base pour la topologie de X. Plus généralement, si $D \subseteq X$ est un ensemble *dense* dans X, alors $\mathcal{B} := \{B(z,1/n);\ z \in D, n \in \mathbb{N}^*\}$ est une base pour la topologie de X.

Démonstration. Exo. □

1.3. Métrisabilité.

Définition 1.3. *Un espace topologique (X,τ) est dit* ***métrisable*** *si sa topologie peut-être définie par une distance, i.e. il existe une distance d sur X telle que $\tau = \tau_d$; et (X,τ) est dit* ***complètement métrisable*** *si sa topologie peut être définie par une distance d telle que (X,d) est complet.*

Exercice. Montrer que si X est un ensemble quelconque, alors la topologie discrète sur X est complètement métrisable.

Remarque 1.4. Si (X,τ) est métrisable, alors τ peut être définie par une distance δ telle que $\delta(x,y) \leqslant 1$ pour tous $x,y \in X$. Si (X,τ) est complètement métrisable, alors τ peut être définie par un distance δ telle que (X,δ) est complet et $\delta(x,y) \leqslant 1$ pour tous $x,y \in X$.

Démonstration. Exo. (*Si d est une distance sur X telle que $\tau = \tau_d$, poser $\delta(x,y) := \min\big(1, d(x,y)\big)$. Vérifier que (X,δ) est complet si (X,d) est complet.*) □

Remarque 1.5. Si (X,τ) est un espace topologique complètement métrisable, alors tout fermé $F \subseteq X$ est complètement métrisable, et tout ouvert $V \subseteq X$ est complètement métrisable.

Démonstration. Soit d une distance sur X telle que $\tau = \tau_d$ et (X,d) est complet. Alors (F,d) est aussi complet, donc F est complètement métrisable. Pour montrer que V est complètement métrisable, on définit une distance δ sur V en posant

$$\forall x,y \in V\ :\ \delta(x,y) := d(x,y) + \left| \frac{1}{\mathrm{dist}(y, X\backslash V)} - \frac{1}{\mathrm{dist}(x, X\backslash V)} \right|.$$

On vérifie que $\tau_\delta = (\tau_d)_V$, et que (V,δ) est complet (exo). □

1.4. Espaces séparés.

Définition 1.6. *Un espace topologique X est dit* ***séparé*** *si, pour tous $a,b \in X$ tels que $a \neq b$, on peut trouver des ouverts $U \ni a$ et $V \ni b$ tels que $U \cap V = \varnothing$. (En anglais, on dit* ***de Hausdorff*** *au lieu de “séparé”.)*

Exemple 1. Tout espace topologique métrisable est séparé.

Exemple 2. Soit X un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et soit $\| \cdot \|$ une *semi-norme* sur X. La topologie définie par $\| \cdot \|$ est séparée si et seulement si $\| \cdot \|$ est une norme.

Exercice 1. A quelle condition sur un ensemble X la topologie grossière sur X est-elle séparée ?

Exercice 2. Montrer que dans un espace topologique séparé, tous les singletons sont des ensembles fermés.

1.5. Compacité.

Définition 1.7. *Un espace topologique X est dit **compact** si X est séparé et si tout* recouvrement ouvert *de X possède un* sous-recouvrement fini *; autrement dit : pour toute famille d'ouverts $(O_i)_{i\in I}$ telle que $\bigcup_{i\in I} O_i = X$, on peut trouver un ensemble fini $J \subseteq I$ tel que $\bigcup_{i\in J} O_i = X$ (**propriété de Borel-Lebesgue**).*

Exercice. Montrer qu'on définit une topologie sur $\mathbb{R}$ en décrétant que les ensembles fermés sont $\mathbb{R}$ et les ensembles finis ; et montrer que cette topologie possède la propriété de Borel-Lebesgue. Cette topologie est-elle séparée ?

Rappel. Un espace topologique *métrisable* X est compact si et seulement si : toute suite $(x_n)_{n\in\mathbb{N}}$ de points de X possède une sous-suite (x_{n_k}) qui converge dans X (**propriété de Bolzano-Weierstrass**).

Remarque 1. Soit X un espace topologique séparé, et soit $K \subseteq X$. Alors K est compact si et seulement si : pour toute famille $(O_i)_{i\in I}$ d'ouverts *de* X telle que $K \subseteq \bigcup_{i\in I} O_i$, on peut trouver un ensemble fini $J \subseteq I$ tel que $K \subseteq \bigcup_{i\in J} O_i$.

Remarque 2. Soit X un espace topologique séparé. On dit qu'une famille $(F_i)_{i\in I}$ de fermés de X possède la *Propriété d'Intersection Finie* si on a $\bigcap_{i\in J} F_i \neq \varnothing$ pour tout ensemble fini $J \subseteq I$. Alors X est compact si et seulement si : toute famille $(F_i)_{i\in I}$ de fermés de X possédant la PIF est telle que $\bigcap_{i\in I} F_i \neq \varnothing$.

Exercice important. Soient X et Y des espaces topologiques séparés.

(1) Si $f : X \to Y$ est continue et si $K \subseteq X$ est compact, alors $f(K)$ est un compact de Y.
(2) Si X est compact, alors tout fermé $F \subseteq X$ est compact.
(3) Si $K \subseteq X$ est compact, alors K est fermé dans X.
(4) Si X est compact et si $f : X \to Y$ est continue et bijective, alors f est un homéomorphisme, *i.e.* $f^{-1} : Y \to X$ est continue.

1.6. Séparabilité.

Définition 1.8. *Un espace topologique X est dit **séparable** s'il existe un ensemble $D \subseteq X$ à la fois dénombrable et dense dans X.*

Exemples.

(1) $\mathbb{R}$ et $\mathbb{C}$ sont séparables.
(2) Tout evn de dimension finie est séparable.

(3) Tout espace compact métrisable est séparable.

(4) Si K est un espace topologique compact métrisable, alors l'espace de Banach $\mathcal{C}(K)$ est séparable.

(5) Si I est un intervalle de $\mathbb{R}$, alors $L^p(I)$ est séparable pour tout $1 \leqslant p < \infty$.

Démonstration. (1) et (2) sont laissés en exo.

(3) Soit K un espace compact métrisable, et soit d une distance définissant la topologie de K. Pour tout $n \in \mathbb{N}^*$, on peut trouver un ensemble fini $F_n \subseteq K$ tel que $K = \bigcup_{z\in F_n} B(z, \frac{1}{n})$. Alors $D := \bigcup_{n\geqslant 1} F_n$ est dénombrable et dense dans K (exo).

(4) Soit d une distance défiinissant la topologie de K, et soit $D = \{z_n;\ n \in \mathbb{N}\}$ un ensemble dénombrable dense dans K (un tel D existe par (3)). Pour $n \in \mathbb{N}$, notons $f_n \in \mathcal{C}(K)$ la fonction définie par $f_n(x) := d(x, z_n)$. Comme D est dense dans K, les fonctions f_n séparent les points de K (exo). Par le *Théorème de Stone-Weierstrass*, on en déduit que la sous-algèbre $\mathcal{A} \subseteq \mathcal{C}(K)$ engendrée par les f_n est dense dans $\mathcal{C}(K)$. Soit alors $\mathcal{D}$ l'ensemble des fonctions $f \in \mathcal{C}(K)$ qui sont combinaisons linéaires à coefficients dans $\mathbb{Q} + i\mathbb{Q}$ de fonctions de la forme $f_I := \prod_{n\in I} f_n$, où $I \subseteq \mathbb{N}$ est fini. L'ensemble $\mathcal{D}$ est dénombrable, et il est dense dans $\mathcal{C}(K)$ par densité de $\mathcal{A}$ (exo).

(5) Supposons d'abord que l'intervalle I soit *borné*, et donc de mesure finie. Soit $\mathcal{J}$ la famille de tous les intervalles $J \subseteq I$ à extrémités rationnelles, et soit $\mathcal{D} \subseteq L^p(I)$ l'ensemble des fonctions f qui sont combinaisons linéaires à coefficients dans $\mathbb{Q} + i\mathbb{Q}$ de fonctions de la forme $\mathbf{1}_J$ où $J \in \mathcal{J}$. L'ensemble $\mathcal{D}$ est dénombrable. En utilisant le *Théorème des classes monotones*, il n'est pas trop difficile de montrer que $\mathbf{1}_A \in \overline{\mathcal{D}}^{L^p}$ pour tout borélien $A \subseteq I$ (exo). Comme les fonctions étagées sont denses dans L^p, on en déduit que $\mathcal{D}$ est dense dans $L^p(I)$.
Pour un intervalle I quelconque, on écrit $I = \bigcup_{n\in\mathbb{N}} I_n$ où (I_n) est une suite croissante d'intervalles bornés. Si on pose $\mathcal{E}_n := \big\{f \in L^p(I);\ f \equiv 0 \text{ en dehors de } I_n\big\} \subseteq L^p(I)$, alors $\bigcup_{n\in\mathbb{N}} \mathcal{E}_n$ est dense dans $L^p(I)$ (exo). Mais $\mathcal{E}_n$ s'identifie de manière évidente à $L^p(I_n)$, donc $\mathcal{E}_n$ est séparable d'après le 1er cas. Si on choisit pour tout n un ensemble $\mathcal{D}_n \subseteq \mathcal{E}_n$ dénombrable et dense dans $\mathcal{E}_n$, alors $\mathcal{D} := \bigcup_{n\in\mathbb{N}} \mathcal{D}_n$ est dénombrable et dense dans $L^p(I)$. □

Lemme 1.9. *Un* espace vectoriel normé X *est séparable si et seulement si il existe une suite croissante* $(E_n)_{n\in\mathbb{N}}$ *de sous-espaces vectoriels de dimension finie telle que* $\bigcup_{n\in\mathbb{N}} E_n$ *est dense dans* X.

Démonstration. Si X séparable et si $D = \{z_0, z_1, z_2, \dots\}$ est dénombrable et dense dan X, il suffit de poser $E_n := [z_0, \dots, z_n]$. Inversement, on prend $D := \bigcup_{n\in\mathbb{N}} D_n$, où $D_n \subseteq E_n$ est dénombrable et dense dans E_n. □

Exemple. $c_0(\mathbb{N})$ est séparable ; $\ell^p(\mathbb{N})$ est séparable pour tout $1 \leqslant p < \infty$.

Démonstration. Posons $E_n := [e_0, \dots, e_n]$, où $(e_n)_{n\in\mathbb{N}}$ est la "base canonique" de $X = c_0$ ou ℓ^p. Alors $\bigcup_{n\in\mathbb{N}} E_n = c_{00} := \text{vect}\,\big\{e_n;\ n \in \mathbb{N}\big\}$ est dense dans X car si $x \in X$ est quelconque, alors $x = \sum_{n=0}^{\infty} x_n e_n$, où la série converge pour la norme de X (exo) et les sommes partielles de la série appartiennent à c_{00}. □

Proposition 1.10. *Soit X un espace topologique* métrisable. *Alors X est séparable si et seulement si sa topologie possède une base dénombrable.*

Démonstration. Supposons que X soit séparable. Si $D \subseteq X$ est dénombrable et dense dans X et si d une distance définissant la topologie de X, alors on vu que $\mathcal{B} := \{B(z, 1/n);\ z \in D, n \in \mathbb{N}^*\}$ est une base pour la topologie de X ; et $\mathcal{B}$ est dénombrable. Inversement, si $\mathcal{B} = (B_n)_{n\in\mathbb{N}}$ est une base dénombrable pour la topologie de X avec $B_n \neq \varnothing$ et si on choisit pour tout $n \in \mathbb{N}$ un point $z_n \in B_n$, alors $D := \{z_n;\ n \in \mathbb{N}\}$ est dense dans X (exo) ; donc X est séparable. □

Corollaire 1.11. *Si X est métrisable et séparable, alors tout ensemble $E \subseteq X$ est séparable (pour la topologie induite).*

Démonstration. Si $\mathcal{B}$ est une base dénombrable pour la topologie de X, alors $\mathcal{B}_E := \{O \cap E;\ O \in \mathcal{B}\}$ est une base dénombrable pour la topologie de E. □

Corollaire 1.12. (propriété de Lindelöf)
Soit X un espace métrisable séparable. Si $(O_i)_{i\in I}$ est une famille quelconque d'ouverts de X, alors il existe un ensemble dénombrable $J \subseteq I$ tel que $\bigcup_{i\in J} O_i = \bigcup_{i\in I} O_i$.

Démonstration. Soit $\mathcal{B}$ une base dénombrable pour la topologie de X, et posons $\widetilde{\mathcal{B}} := \{B \in \mathcal{B};\ \exists i \in I\ :\ B \subseteq O_i\}$. Pour tout $B \in \widetilde{\mathcal{B}}$, choisissons un indice $i(B) \in I$ tel que $B \subseteq O_{i(B)}$. Alors $J := \{i(B);\ B \in \widetilde{\mathcal{B}}\}$ convient (exo). □

Lemme 1.13. *Soit X un espace topologique séparable. Si $\mathcal{O} = (O_i)_{i\in I}$ est une famille d'ouverts de X non vides et deux à deux disjoints, alors $\mathcal{O}$ est nécessairement dénombrable.*

Démonstration. Soit $D \subseteq X$ un ensemble dénombrable dense. Pour tout $i \in I$, on peut choisir un point $z_i \in D \cap O_i$ puisque D est dense dans X. Comme les O_i sont disjoints, l'application $i \mapsto z_i$ est *injective*. Donc I s'injecte dans l'ensemble dénombrable D ; et donc I est dénombrable. □

Corollaire 1.14. *L'espace $\ell^\infty(\mathbb{N})$* n'est pas *séparable. Si I est un intervalle de $\mathbb{R}$ non trivial, alors $L^\infty(I)$ n'est pas séparable.*

Démonstration. Si $A, B \subseteq \mathbb{N}$ et $A \neq B$, alors $\|\mathbf{1}_B - \mathbf{1}_A\|_\infty = 1$. Donc la famille $\{B(\mathbf{1}_A, 1/2);\ A \in \mathcal{P}(\mathbb{N})\}$ est une famille non-dénombrable de boules ouvertes deux à deux disjointes dans $\ell^\infty(\mathbb{N})$; et donc $\ell^\infty(\mathbb{N})$ n'est pas séparable. Le cas de $L^\infty(I)$ est laissé en exo. □

Définition 1.15. *Un **espace polonais** est un espace topologique séparable et complètement métrisable.*

2. Suites généralisées

2.1. Définitions et exemples.

Définition 2.1. *Un **ensemble préordonné** $(P, \preceq)$ est un ensemble P muni d'une relation binaire $\preceq$ qui est* transitive *: $(p \preceq q$ et $q \preceq r) \implies p \preceq r$. On dit qu'un ensemble préordonné $(P, \preceq)$ est **filtrant** si*

$$\forall p_1, p_2 \in P\ \exists p \in P \quad \text{tel que} \quad p \succeq p_1 \text{ et } p \succeq p_2.$$

Exemple. Les ensembles suivants sont préordonnés filtrants.

(1) $(\mathbb{N}, \leqslant)$, $(\mathbb{R}^+, \leqslant)$.

(2) $(\mathbb{R}, \preceq)$, où $x \preceq y$ ssi $|x| \leqslant |y|$.
(3) $(]0,1], \geqslant)$.
(4) $(\mathtt{FIN}(I), \subseteq)$, où I est un ensemble quelconque et $\mathtt{FIN}(I)$ est l'ensemble des partis finies de I.
(5) $(\mathcal{V}(a), \supseteq)$, où $a \in X$ espace topologique.

Définition 2.2. *Soit X un ensemble quelconque. Une **suite généralisée** dans X est une famille $(x_p)_{p\in P}$ d'éléments de X indexée par un ensemble préordonné filtrant $(P, \preceq)$.*

Remarque 1. Le plus souvent, on écrira "s.g." au lieu de "suite généralisée".

Remarque 2. Évidemment, toute *suite* est une s.g.

2.2. Convergence.

Définition 2.3. *Soit X un espace topologique, $(x_p)_{p\in P}$ une s.g. dans X. Étant donné $a \in X$, on dit que (x_p) **tend vers** a, et on écrit $x_p \to a$ si*

$$\forall V \text{ voisinage de } a,\ \exists p_0 \in P \text{ tel que } \forall p \succeq p_0\ :\ x_p \in V.$$

*On dit que la s.g. (x_p) **converge dans** X s'il existe $a \in X$ telle que $x_p \to a$.*

Remarque. Si $(x_p)_{p\in\mathbb{N}}$ est une vraie suite, on retrouve la définition habituelle.

Exercice. Si l'espace topologique X est *séparé*, alors on a "unicité de la limite" : on ne peut pas avoir $x_p \to a$ et $x_p \to b$ avec $a \neq b$.

Exemple. Soit $a \in X$. Pour tout V voisinage de a, choisissons un point $x_V \in V$. Alors la s.g. $(x_V)_{V\in\mathcal{V}(a)}$ tend vers a.

Démonstration. C'est évident par définition. □

Proposition 2.4. *Soit X un espace topologique et soit $A \subseteq X$. Soit aussi $a \in X$. Alors*

$$a \in \overline{A} \iff \text{il existe une s.g. } (x_p) \subseteq A \text{ telle que } x_p \to a.$$

Démonstration. Supposons que $a \in \overline{A}$. Alors $V \cap A \neq \varnothing$ pour tout $V \in \mathcal{V}(a)$. Choisissons un point $x_V \in V \cap A$ pour tout $V \in \mathcal{V}(a)$. Alors la s.g. $(x_V)_{V\in\mathcal{V}(a)}$ est une s.g. de points de A et $x_V \to a$.
Variante sans choix : on pose $P := \{(V, z);\ V \in \mathcal{V}(a),\ z \in V \cap A\}$ et on décrète que $(V, z) \preceq (V', z')$ ssi $V \supseteq V'$. Alors $(P, \preceq)$ est filtrant car $a \in \overline{A}$ (exo). Pour $p = (V, z) \in P$, on pose $x_{V,z} := z$. Alors $x_p \to a$.
Inversement, supposons qu'il existe une s.g. $(x_p)_{p\in P} \subseteq A$ telle que $x_p \to a$. Alors $\forall V \in \mathcal{V}(a)\ \exists p \in P\ :\ x_p \in V$, et donc certainement $V \cap A \neq \varnothing$ pour tout $V \in \mathcal{V}(a)$; donc $a \in \overline{A}$. □

Corollaire 2.5. *Un ensemble $F \subseteq X$ est fermé dans X si et seulement si il contient tous ses "points limites", autrement dit : chaque fois qu'une s.g. $(x_p) \subseteq F$ tend vers un point $a \in X$, on a que $a \in F$.*

Remarque. En général, on ne peut pas remplacer "s.g." par "suite" dans la proposition 2.4 ; mais on sait bien qu'on peut si X est métrisable !

Exercice. Montrer qu'on peut remplacer "s.g." par "suite" si le point a possède une base dénombrable de voisinages.

PROPOSITION 2.6. *Soient X et Y des espaces topologiques, et soit $f : X \to Y$. Soit également $a \in X$. Alors*

$$(f \text{ continue en } a) \iff f(x_p) \to f(a) \text{ pour toute s.g. } (x_p) \subseteq X \text{ telle que } x_p \to a.$$

Démonstration. On laisse $\Longrightarrow$ en exo.
Inversement, supposons f *non* continue en a. Alors on peut trouver un voisinage W_0 de $f(a)$ tel qu'on n'ait $f(V) \subseteq W_0$ pour *aucun* voisinage V de a. Donc, pour tout $V \in \mathcal{V}(a)$, on peut choisir un point $x_V \in V$ tel que $f(x_V) \notin W_0$. Alors $x_V \to a$, mais $f(x_V) \not\to f(a)$ puisque $\forall V \ : \ f(x_V) \notin W_0$. □

Exercice. Faire une preuve "sans choix".

COROLLAIRE 2.7. *Soient τ et τ' deux topologies sur X. Alors $\tau \subseteq \tau'$ si et seulement si la convergence d'une s.g. pour τ' entraine sa convergence pour τ (vers la même limite). Et donc, $\tau = \tau'$ si et seulement si τ et τ' ont les mêmes s.g. convergentes (avec les mêmes limites).*

Démonstration. Dire que $\tau \subseteq \tau'$ signifie que l'application $id : (X, \tau') \to (X, \tau)$ est continue ; donc le résultat est évident. □

2.3. Sous-s.g. et valeurs d'adhérences.

DÉFINITION 2.8. *Soit $(x_p)_{p \in P}$ une s.g. dans un ensemble X. Une* ***sous-s.g.*** *de (x_p) est une s.g. $(z_{p'})_{p' \in P'}$ de la forme $z_{p'} = x_{\phi(p')}$, où $\phi : P' \to P$ est une application* ***cofinale****, i.e. telle que $\forall p_0 \in P \ \exists p'_0 \in P' \ \forall p' \geq p'_0 \ : \ \phi(p') \geq p_0$.*

Exemple. Si $(x_n)_{n \in \mathbb{N}}$ est une suite dans X et si $(n_k)_{k \in \mathbb{N}}$ est une suite d'entiers telle que $n_k \to \infty$, alors $(x_{n_k})_{k \in \mathbb{N}}$ est une sous-s.g. de (x_n).

Exercice. Une sous-s.g. d'une sous-s.g. est une sous-s.g.

DÉFINITION 2.9. *Soit $(x_p)_{p \in P}$ une s.g. dans un espace topologique X, et soit $a \in X$. On dit que a est une* ***valeur d'adhérence*** *de (x_p) si*

$$\forall V \in \mathcal{V}(a) \ \forall p_0 \in P \ \exists p \geq p_0 \ : \ x_p \in V.$$

LEMME 2.10. *Un point $a \in X$ est une valeur d'adhérence de (x_p) si et seulement si (x_p) admet une sous-s.g. $(z_{p'})$ telle que $z_{p'} \to a$.*

Démonstration. Supposons que a soit valeur d'adhérence de (x_p). Posons

$$P' := \{(V, p); \ V \in \mathcal{V}(a) \text{ et } x_p \in V\};$$

et décrétons que

$$(V_1, p_1) \preceq (V_2, p_2) \qquad \text{ssi} \qquad V_1 \supseteq V_2 \quad \text{et} \quad p_1 \preceq p_2.$$

Alors $(P', \preceq)$ est filtrant car a est une valeur d'adhérence de (x_p). Détail : étant donné $p'_1 = (V_1, p_1)$ et $p'_2 = (V_2, p_2)$, on peut trouver $p_0 \in P$ tel que $p_0 \succeq p_1$ et $p_0 \succeq p_2$ car P est filtrant, puis $p \succeq p_0$ tel que $x_p \in V_1 \cap V_2$ car a est valeur d'adhérence de (x_p) et $V_1 \cap V_2$ est un voisinage de a ; et alors $p' := (V_1 \cap V_2, p)$ est tel que $p' \succeq p'_1$ et $p' \succeq p'_2$. Pour $p' = (V, p) \in P'$, posons $z_{p'} := x_p$. Alors $(z_{p'})_{p' \in P'}$ est une sous-s.g. de (x_p) et $z_{p'} \to a$. "Détail" : l'application $\phi : P' \to P$ définie par $\phi(V, p) := p$ est cofinale ar

$(X,p) \in P'$ pour tout $p \in P$; et $z_{p'} \to a$ par définition de la relation $\preceq$ car pour tout $V_0 \in \mathcal{V}(a)$, on peut trouver $p_0 \in P$ tel que $x_{p_0} \in V_0$.

La réciproque est laissée en exo. □

Proposition 2.11. *Soit X un espace topologique séparé. Alors X est compact si et seulement si toute s.g. $(x_p)_{p\in P}$ dans X possède une sous-s.g. qui converge dans X, si et seulement si toute s.g. $(x_p) \subseteq X$ possède une valeur d'adhérence dans X.*

Démonstration. Supposons X compact, et soit $(x_p)_{p\in P}$ une s.g. dans X. Supposons que (x_p) ne possède aucune valeur d'adhérence dans X. Alors, pour tout $x \in X$, on peut trouver un voisinage ouvert V_x de x dans X et un indice $p_x \in P$ tels que $\forall p \succeq p_x \;:\; x_p \notin V_x$. La famille $(V_x)_{x\in X}$ est un recouvrement ouvert de X, donc on peut trouver $x_1, \dots, x_r \in X$ tels que $V_{x_1} \cup \dots \cup V_{x_r} = X$. Comme P est filtrant, on peut ensuite trouver $p \in P$ tel que $p \succeq p_{x_1}, \dots, p_{x_r}$. Alors $x_p \notin V_{x_i}$ pour $i = 1, \dots, r$, ce qui est absurde puisque quand même $x_p \in X$.

Inversement, supposons que X ne soit pas compact : il faut trouver une s.g. $(x_p)_{p\in P}$ sans valeur d'adhérence. Soit $(O_i)_{i\in I}$ un recouvrement ouvert de X n'admettant aucun sous-recouvrement fini. Posons $(P, \preceq) := (\mathtt{FIN}(I), \subseteq)$. Pour tout $J \in P = \mathtt{FIN}(I)$, choisissons un point $x_J \in X$ tel que $x_J \notin \bigcup_{i\in J} O_i$. Alors $(x_J)_{J\in\mathcal{P}_f(I)}$ est une s.g. dans X. Supposons que (x_J) possède une valeur d'adhérence $a \in X$. Comme $X = \bigcup_{i\in I} O_i$, on peut trouver $i_0 \in I$ tel que $a \in O_{i_0}$. Comme a est censé être une valeur d'adhérence de (x_J) et que O_{i_0} est un voisinage de a, on peut trouver $J \in \mathcal{P}_f(I)$ tel que $\{i_0\} \subseteq J$, *i.e.* $i_0 \in J$, et $x_J \in O_{i_0}$; mais c'est absurde car on doit avoir $x_J \notin \bigcup_{i\in J} O_i$ et $i_0 \in J$. □

Exercice. Refaire l'“Exercice important” de 1.5 en utilisant des s.g.

2.4. Complétude et s.g. de Cauchy.

Définition 2.12. *Soit (X,d) un espace métrique. On dit qu'une suite généralisée $(x_p)_{p\in P} \subseteq X$ est **de Cauchy** si*

$$\forall \varepsilon > 0 \; \exists p_\varepsilon \in P \; \forall p, q \succeq p_\varepsilon \;:\; d(x_p, x_q) \leqslant \varepsilon.$$

Exercice. Montrer que toute s.g. convergente est de Cauchy.

Lemme 2.13. *Si (X,d) est un espace métrique complet, alors toute s.g. de Cauchy $(x_p)_{p\in P} \subseteq X$ est convergente.*

Démonstration. Comme (x_p) est de Cauchy, on peut pour tout $n \in \mathbb{N}$ choisir $p_n \in P$ tel que

$$\forall p, q \succeq p_n \;:\; d(x_p, x_q) \leqslant 2^{-n}.$$

De plus, comme l'ensemble préordonné P est filtrant, on peut supposer que la suite $(p_n)_{n\in\mathbb{N}}$ est croissante : si $p_0, \dots, p_n$ ont été choisis, on peut toujours prendre $p_{n+1} \succeq p_0, \dots, p_n$.

Par définition, la *suite* $(x_{p_n})_{n\in\mathbb{N}}$ est de Cauchy dans (X,d) ; donc elle converge dans X puisque (X,d) est complet, $x_{p_n} \to x \in X$. Si maintenant $\varepsilon > 0$ est quelconque, on peut choisir $n_0 \in \mathbb{N}$ tel que $2^{-n_0} \leqslant \varepsilon$, et on a alors $d(x_p, x_{p_n}) \leqslant \varepsilon$ pour tout $p \succeq p_{n_0}$ et pour tout $n \geqslant n_0$ puisque $p_n \succeq p_{n_0}$. En faisant $n \to \infty$, on en déduit que $d(x_p, x) \leqslant \varepsilon$ pour tout $p \succeq p_{n_0}$. Donc $x_p \to x$. □

2.5. Un exemple : familles sommables.

DÉFINITION 2.14. *Soit X un espace vectoriel normé, et soit $(x_i)_{i\in I}$ une famille d'éléments de X. On dit que la famille (x_i) est* ***sommable*** *si la suite généralisée $\left(\sum_{i\in F} x_i\right)_{F\in \mathtt{FIN}(I)}$ est convergente. Autrement dit, (x_i) est sommable s'il existe $x \in X$ (uniquement déterminé) tel que la chose suivante ait lieu :*

$$\forall \varepsilon > 0\ \exists F_\varepsilon \subseteq I \text{ fini} \quad \textit{tel que} \quad \forall F \supseteq F_\varepsilon \textit{ fini} \ : \ \left\| \sum_{i\in F} x_i - x \right\| \leqslant \varepsilon.$$

Si la famille (x_i) est sommable, on pose $\displaystyle\sum_{i\in I} x_i := \lim_F \sum_{i\in F} x_i$.

Exercice. Montrer que si une *suite* $(x_n)_{n\in\mathbb{N}} \subseteq X$ est sommable, alors la série $\sum x_n$ converge dans X et $\sum_{n=0}^{\infty} x_n = \sum_{n\in\mathbb{N}} x_n$. Montrer de même que si une suite $(x_n)_{n\in\mathbb{Z}}$ indexée par $\mathbb{Z}$ est sommable, alors $\sum_{n=-\infty}^{\infty} x_n := \lim_{N\to\infty} \sum_{n=-N}^{N} x_n$ existe et est égale à $\sum_{n\in\mathbb{Z}} x_n$.

REMARQUE 2.15. (Critère de Cauchy)
Supposons que X soit un espace de Banach. Alors une famille $(x_i)_{i\in I} \subseteq X$ est sommable si et seulement si la chose suivante a lieu : pour tout $\varepsilon > 0$, on peut trouver un ensemble fini $F_\varepsilon \subseteq I$ tel que

$$\left\| \sum_{i\in J} x_i \right\| \leqslant \varepsilon \qquad \text{pour tout } J \subseteq I \text{ fini tel que } J \cap F_\varepsilon = \varnothing.$$

Démonstration. C'est une conséquence du Lemme 2.13. Les détails sont laissés en exo. □

EXEMPLE 1. Supposons que X soit un espace de Banach. Si $(x_i)_{i\in I}$ et une famille d'éléments de X telle que $\sum_{i\in I} \|x_i\| < \infty$, alors la famille (x_i) est sommable.

Démonstration. Rappelons que si $(a_i)_{i\in I}$ est une famille de nombres positifs, alors, par définition,

$$\sum_{i\in} a_i = \sup_{F\in \mathtt{FIN}(I)} \sum_{i\in F} a_i \ \in [0,\infty].$$

Soit $\varepsilon > 0$. Comme $\sum_{i\in I} \|x_i\| < \infty$, on peut trouver un ensemble fini $F_\varepsilon \subseteq I$ tel que $\sum_{i\in I} \|x_i\| \leqslant \sum_{i\in F} \|x_i\| + \varepsilon$. Pour tout ensemble fini $J \subseteq I$ tel que $J \cap F_\varepsilon = \varnothing$, on a alors $\sum_{i\in J} \|x_i\| + \sum_{i\in F_\varepsilon} \|x_i\| = \sum_{i\in J\cup F_\varepsilon} \|x_i\| \leqslant \sum_{i\in I} \|x_i\| \leqslant \sum_{i\in F} \|x_i\| + \varepsilon$, donc $\sum_{i\in J} \|x_i\| \leqslant \varepsilon$, et donc *a fortiori* $\|\sum_{i\in J} x_i\| \leqslant \varepsilon$. Donc la famille (x_i) est sommable par le critère de Cauchy (Remarque 2.15). □

EXEMPLE 2. Soit H un espace de Hilbert. Si $(x_i)_{i\in I}$ est une famille d'éléments de H deux à deux orthogonaux telle que $\sum_{i\in I} \|x_i\|^2 < \infty$, alors la famille (x_i) est sommable.

Démonstration. La preuve est la même que pour l'Exemple 1, en mettant des carrés partout et en observant qu'on a $\left\|\sum_{i\in J} x_i\right\|^2 = \sum_{i\in J} \|x_i\|^2$ pour tout ensemble fini $J \subseteq I$, par orthogonalité des x_i. □

Remarque. L'Exemple 2 montre en particulier que si $(e_i)_{i\in I}$ est une base hilbertienne de H alors, pour tout $x \in H$, la famille $\left(\langle x, e_i\rangle e_i\right)_{i\in I}$ est sommable et $\sum_{i\in I} \langle x, e_i\rangle e_i = x$.

Théorème 2.16. *Soit X un espace de Banach, et soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X. Les propriétés suivantes sont équivalentes.*

(1) *La suite (x_n) est sommable.*
(2) *Pour toute permutation $\sigma : \mathbb{N} \to \mathbb{N}$, la série $\sum x_{\sigma(n)}$ est convergente.*
(3) *Pour toute application injective $\sigma : \mathbb{N} \to \mathbb{N}$, la série $\sum x_{\sigma(n)}$ est convergente.*
(4) *Pour toute suite strictement croissante d'entiers $(n_k)_{k\in\mathbb{N}}$, la série $\sum x_{n_k}$ est convergente.*
(5) *Pour toute suite $(\varepsilon_n) \in \{0,1\}^{\mathbb{N}}$, la série $\sum \varepsilon_n x_n$ est convergente.*

De plus, si la suite (x_n) est sommable alors $\sum_{n\in\mathbb{N}} x_n = \sum_0^\infty x_{\sigma(n)}$ pour toute permutation $\sigma : \mathbb{N} \to \mathbb{N}$.

Démonstration. Les implications (3) $\implies$ (2) et (3) $\implies$ (4) sont évidentes ; et l'équivalence de (4) et (5) est un micro-exo.

Supposons (1) vérifiée, et montrons (3). On fixe une application injective $\sigma : \mathbb{N} \to \mathbb{N}$. Soit $\varepsilon > 0$. Par le critère de Cauchy, on peut trouver un ensemble fini $F \subseteq \mathbb{N}$ tel que $\left\|\sum_{n\in J} x_n\right\| \leqslant \varepsilon$ pour tout $J \subseteq \mathbb{N}$ fini vérifiant $J \cap F = \varnothing$. Posons alors $F' := \sigma^{-1}(F)$. Comme σ est injective, F' est un ensemble fini. Si $J' \subseteq \mathbb{N}$ est fini et $J' \cap F' = \varnothing$, alors $\sigma(J') \cap F = \varnothing$; donc, par injectivité de σ, on a $\left\|\sum_{n\in J'} x_{\sigma(n)}\right\| = \left\|\sum_{k\in\sigma(J')} x_k\right\| \leqslant \varepsilon$. Donc la suite $(x_{\sigma(n)})$ est sommable par le critère de Cauchy, et donc la série $\sum x_{\sigma(n)}$ converge.

Montrons (2) $\implies$ (1) par contraposée. Supposons que la suite (x_n) ne soit pas sommable. Alors le critère de Cauchy n'est pas vérifié ; donc il existe un $\varepsilon > 0$ tel que : pour tout ensemble fini $F \subseteq \mathbb{N}$, on peut trouver une ensemble fini F' tel que $F' \cap F = \varnothing$ et $\left\|\sum_{m\in F'} x_m\right\| \geqslant \varepsilon$. On peut donc construire par récurrence une suite de parties finies de $\mathbb{N}$ deux à deux disjointes $(F_r)_{r\geqslant 0}$ telle que $\left\|\sum_{m\in F_r} x_m\right\| \geqslant \varepsilon$ pour tout $r \geqslant 0$. Soit $(I_r)_{r\geqslant 0}$ une suite d'intervalles de $\mathbb{N}$ consécutifs ($\min I_{r+1} > \max I_r$ pour tout k) telle que $|I_r| = |F_r|$ pour tout r et telle que $\mathbb{N}\setminus\bigcup I_r$ ait la même cardinalité (finie ou infinie) que $\mathbb{N}\setminus\bigcup_{r\geqslant 0} F_r$. Par le choix des I_r, il existe une permutation $\sigma : \mathbb{N} \to \mathbb{N}$ telle que $\sigma(I_r) = F_r$ pour tout $r \geqslant 0$. On a alors $\left\|\sum_{n\in I_r} x_{\sigma(n)}\right\| = \left\|\sum_{m\in F_r} x_m\right\| \geqslant \varepsilon$ pour tout r ; donc les sommes partielles de la série $\sum x_{\sigma(n)}$ ne vérifient pas le critère de Cauchy (micro-exo), et donc cette série ne converge pas. Ainsi, (2) n'est pas vérifiée.

La preuve de (4) $\implies$ (1) est très semblable. Si (1) n'est pas vérifiée, il existe un $\varepsilon > 0$ témoignant que le critère de Cauchy n'est pas vérifié. On peut construire par récurrence une suite $(F_r)_{r\geqslant 0}$ de parties finies de $\mathbb{N}$ telle que $\min F_{r+1} > \max F_r$ et $\left\|\sum_{n\in F_r} x_n\right\| \geqslant \varepsilon$ pour tout $r \geqslant 0$: si $F_0, \dots, F_r$ sont construits, on pose $N := \max(F_0 \cup \cdots \cup F_r)$, et on choisit F_{r+1} tel que $F_{r+1} \cap [\![0, N]\!] = \varnothing$ et $\left\|\sum_{n\in F_{r+1}} x_n\right\| \geqslant \varepsilon$. Si on note $(n_k)_{k\geqslant 0}$ l'énumération croissante de l'ensemble $\bigcup_{r\geqslant 0} F_r$, alors la série $\sum x_{n_k}$ diverge (exo) ; donc (5) n'est pas vérifiée.

Montrons pour finir que si la suite (x_n) est sommable, alors $\sum_{n\in\mathbb{N}} x_n = \sum_{n=0}^\infty x_{\sigma(n)}$ pour toute permutation $\sigma : \mathbb{N} \to \mathbb{N}$. Posons $x := \sum_{n\in\mathbb{N}} x_n$. Soit $\varepsilon > 0$, et soit $F_\varepsilon \subseteq \mathbb{N}$ un ensemble fini tel que $\left\|\sum_{m\in F} x_m - x\right\| \leqslant \varepsilon$ pour tout ensemble fini $F \supseteq F_\varepsilon$. Comme $\sigma : \mathbb{N} \to \mathbb{N}$ est surjective, on peut trouver un entier N_0 tel que $\sigma\big([\![0, N_0]\!]\big) \supseteq F_\varepsilon$. Alors $\sigma\big([\![0, N]\!]\big) \supseteq F_\varepsilon$ pour tout $N \geqslant N_0$, donc $\left\|\sum_{m\in\sigma([\![0,N]\!])} x_m - x\right\| \leqslant \varepsilon$. Mais comme σ est injective, on a $\sum_{m\in\sigma([\![0,N]\!])} x_m = \sum_{n=0}^N x_{\sigma(n)}$. Donc $\left\|\sum_{n=0}^N x_{\sigma(n)} - x\right\| \leqslant \varepsilon$ pour tout $N \geqslant N_0$, ce qui prouve que $\sum_{n=0}^\infty x_{\sigma(n)} = x$. $\square$

Corollaire 2.17. *soit X un evn* de dimension finie*, et soit $(x_n)_{n\in\mathbb{N}} \subseteq X$. Alors (x_n) est sommable* si et seulement si $\sum_0^\infty \|x_n\| < \infty$.

Démonstration. Par équivalence des normes, on se ramène au cas où $X = \mathbb{K}$ (exo). Et dans ce cas, le résultat est "bien connu" : si (x_n) est une suite de scalaires telle que $\sum x_{\sigma(n)}$ converge pour toute permutation $\sigma : \mathbb{N} \to \mathbb{N}$, alors "on sait bien" que la série $\sum x_n$ est absolument convergente. □

Définition 2.18. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite dans un espace de Banach X. On dit que* ***la série $\sum x_n$ est inconditionnellement convergente*** *si la suite (x_n) est sommable.*

Proposition 2.19. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite dans un espace de Banach X. Si la série $\sum x_n$ est inconditionnellement convergente alors, pour toute suite bornée $(a_n) \subseteq \mathbb{K}$, la série $\sum a_n x_n$ est convergente.*

Démonstration. En considérant parties réelles et imaginaires séparément si $\mathbb{K} = \mathbb{C}$, on se ramène au cas où les a_n sont réels. De plus, quitte à remplacer a_n par $\frac{M-a_n}{2M}$ où $M = \sup_k |a_k|$, on peut supposer qu'on a $0 \leqslant a_n \leqslant 1$ pour tout $n \in \mathbb{N}$.
Soit $\varepsilon > 0$. Comme la suite (x_n) est sommable, on peut trouver un entier N tel que $\|\sum_{n\in J} x_n\| \leqslant \varepsilon$ pour tout ensemble fini J tel que $\min J > N$. On va montrer qu'on a $\left\|\sum_{n=p+1}^q a_n x_n\right\| \leqslant \varepsilon$ pour tous $N \leqslant p < q$, ce qui prouvera que les sommes partielles de la série $\sum a_n x_n$ vérifient le critère de Cauchy.
Fixons p et q tels que $N \leqslant p < q$, et posons $I := [\![p+1, q]\!]$. Alors la suite finie $(a_n)_{n\in I}$ appartient à $[0,1]^I$, donc elle est combinaison convexe de suites $\alpha \in \{0,1\}^I$ (exo). Pour une telle suite α, on a $\sum_{n\in I} \alpha_n x_n = \sum_{n\in J} x_n$ où $J := \{n \in I;\ \alpha_n = 1\}$ vérifie $\min J > N$; donc $\left\|\sum_{n\in I} \alpha_n x_n\right\| \leqslant \varepsilon$. Par convexité de la norme, on en déduit $\left\|\sum_{n\in I} a_n x_n\right\| \leqslant \varepsilon$. □

3. Topologie engendrée par une famille d'applications

3.1. Définition générale.

Lemme-Définition. Soit X un ensemble, et soit $(\pi_i)_{i\in I}$ une famille d'applications, $\pi_i : X \to E_i$ où E_i est un espace topologique. Il existe une *plus petite* topologie τ sur X rendant continues toutes les applications $\pi_i : X \to E_i$. On dit que τ est la topologie **engendrée par les applications π_i.**

Démonstration. Il y a certainement au moins une topologie rendant continues les π_i, à savoir la topologie discrète. De plus l'intersection d'une famille quelconque de topologies sur X rendant continues les π_i est encore une topologie rendant continues les π_i (exo). Donc la topologie cherchée est l'intersection de *toutes* les topologies sur X rendant continues les π_i. □

Dans la suite, X est un espace topologique dont la topologie τ est engendrée par une famille d'applications $\pi_i : X \to E_i$, $i \in I$.

Lemme 3.1. *Un ensemble $O \subseteq X$ est τ-ouvert si et seulement s'il est réunion d'ensembles de la forme*

$$[V_{i_1}, \dots, V_{i_r}] := \left\{x \in X;\ \pi_{i_1}(x) \in V_{i_1}, \dots, \pi_{i_r}(x) \in V_{i_r}\right\},$$

où $i_1, \dots, i_r \in I$ et V_i est un ouvert de E_i pour $i = i_1, \dots, i_r$.

Démonstration. Notons τ' la famille de tous les ensembles $O \subseteq X$ qui sont réunion d'ensembles de la forme $[V_{i_1}, \dots, V_{i_r}]$. Comme $[V_{i_1}, \dots, V_{i_r}] = \bigcap_{s=1}^r \pi_{i_s}^{-1}(V_{i_s})$, on voit que $[V_{i_1}, \dots, V_{i_r}]$ est τ-ouvert par continuité des π_i, et donc tout ensemble $O \in \tau'$ est τ-ouvert. Ainsi, $\tau' \subseteq \tau$. Inversement, on vérifie que τ' est une topologie sur X (exo), et que τ' rend les π_i continues (autre exo). Donc $\tau' \supseteq \tau$. $\square$

Remarque. On dira que les ensembles de la forme $[V_{i_1}, \dots, V_{i_r}]$ sont les **ouverts élémentaires** pour la topologie τ.

LEMME 3.2. *Si $(x_p)_{p\in P}$ est une suite généralisée dans X et si $a \in X$, alors*
$$x_p \to a \iff \pi_i(x_p) \to \pi_i(a) \text{ pour tout } i \in I.$$

Démonstration. Si $x_p \to a$, alors certainement $\pi_i(x_p) \to \pi_i(a)$ pour tout $i \in I$ par continuité de π_i. Inversement, supposons que $\pi_i(x_p) \to \pi_i(a)$ pour tout $i \in I$. Soit V un voisinage quelconque de a dans X. Par le Lemme 3.1, on peut trouver un ouvert élémentaire $[V_{i_1}, \dots, V_{i_r}]$ tel que $a \in [V_{i_1}, \dots, V_{i_r}] \subseteq V$. Comme $\pi_i(x_p) \to \pi_i(a)$ pour tout i, on peut trouver $p_1, \dots, p_r \in P$ tels que pour $s = 1, \dots, r$, on ait que $\forall p \succeq p_s \ : \ \pi_{i_s}(x_p) \in V_{i_s}$. Comme P est filtrant, on peut trouver $p_0 \in P$ tel que $p_0 \succeq p_1, \dots, p_r$. Alors $\forall p \succeq p_0 \ : \ x_p \in [V_{i_1}, \dots, V_{i_r}] \subseteq V$. Donc $x_p \to a$. $\square$

COROLLAIRE 3.3. *Si Z est un autre espace topologique, une application $f : Z \to X$ est continue si et seulement si $\pi_i \circ f : Z \to E_i$ est continue pour tout $i \in I$.*

Démonstration. Exo. $\square$

3.2. Topologie produit.

DÉFINITION 3.4. *Soit $(X_i)_{i\in I}$ une famille d'espaces topologiques, et soit $X := \prod_{i\in I} X_i$. Un “point” de X est donc de la forme $x = (x_i)_{i\in I}$ où $x_i \in X_i$ pour tout $i \in I$; on pourra aussi écrire $x = (x(i))_{i\in I}$. La **topologie produit** sur X est la topologie engendrée par les projections canoniques $\pi_i : X \to X_i$.*

MODE D'EMPLOI. Soit $X = \prod_{i\in I} X_i$ muni de la topologie produit.

- Une s.g. $(x_p)_{p\in P}$ dans X converge vers $a \in X$ ssi $x_p(i) \to a(i)$ dans X_i pour tout $i \in I$.
- Soit Z un espace topologique, et soit $f = Z \to X$. Écrivons $f(z) = \big(f_i(z)\big)_{i\in I}$. Alors f est continue ssi chaque application $f_i : Z \to X_i$ est continue.
- Une base pour la topologie produit est donnée par les ouverts élémentaires
$$\text{“}V_{i_1} \times \cdots \times V_{i_r}\text{”} := \big\{x \in X;\ x_{i_1} \in V_{i_1}, \dots, x_{i_r} \in V_{i_r}\big\},$$
où V_i est ouvert dans X_i pour $i = 1, \dots, r$.

EXEMPLE 1. Soient E et F deux evn sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. La topologie produit sur $E \times F$ est la topologie définie par la **norme produit** $\|(x, y)\| := \max\big(\|x\|_E, \|y\|_F\big)$.

EXEMPLE 2. Prenons tous les X_i égaux à un même espace topologique E. Alors $\prod_{i\in I} X_i = E^I$ est l'ensemble de toutes les applications $f : I \to E$, qu'on notera souvent $\mathcal{F}(I, E)$. La topologie produit sur $\mathcal{F}(I, E)$ s'appelle la **topologie de la convergence simple**.

Remarque. Il n'est pas clair que si $X_i \neq \varnothing$ pour tout $i \in I$, alors $\prod_{i\in I} X_i \neq \varnothing$. En fait, c'est exactement ce que dit l'**axiome du choix**. (À titre d'exercice, on pourra montrer que $\prod_{i\in I} X_i \neq \varnothing$ en utilisant le Lemme de Zorn.)

Exercice. Montrer que si tous les X_i sont séparés, alors $X = \prod_{i\in I} X_i$ est séparé.

3.3. Topologie définie par une famille de semi-normes.

Rappel. Soit X un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une *semi-norme* sur X est une application $\|\cdot\| : X \to \mathbb{R}^+$ satisfaisant toutes les propriétés d'une norme sauf peut-être l'implication $\|0\| = 0 \implies x = 0$. Toute semi-norme $\|\cdot\|$ définit une topologie $\tau_{\|\cdot\|}$ sur X : un ensemble $O \subseteq X$ est $\tau_{\|\cdot\|}$-ouvert ssi, pour tout $a \in O$, on peut trouver $r > 0$ tel que $B_{\|\cdot\|}(a, r) \subseteq O$. (On a vu en exo que cette topologie est séparée si et seulement si la semi-norme $\|\cdot\|$ est une norme.)

Définition 3.5. *Soit X un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et soit $(\|\cdot\|_i)_{i\in I}$ une famille de semi-normes sur X. La **topologie définie par les** $\|\cdot\|_i$ est la topologie sur X engendrée par les "applications identité" $X \hookrightarrow (X, \|\cdot\|_i)$.*

Traduction. On garde les notations de la définition.

- Si (x_p) est une s.g. dans X et $a \in X$, alors $x_p \to a \iff \|x_p - a\|_i \to 0$ pour toute semi-norme $\|\cdot\|_i$.
- Soit $a \in X$. Pour $i_1, \dots, i_r \in I$ et $\varepsilon > 0$, posons
$$B_{i_1,\dots,i_r}(a, \varepsilon) := \big\{x \in X;\ \|x - a\|_{i_1} < \varepsilon, \dots, \|x - a\|_{i_r} < \varepsilon\big\}.$$
Alors les $B_{i_1,\dots,i_r}(a, \varepsilon)$ forment une base de voisinages pour a.

Exemples.

(1) Soit I un ensemble quelconque, et soit Y un evn. La topologie de la convergence simple sur $\mathcal{F}(I, Y)$ est engendrée par les semi-normes $\|\cdot\|_i$, $i \in I$ définies par $\|f\|_i := \|f(i)\|$.

(2) Soient X et Y des evn. La **topologie opératorielle forte** sur $\mathcal{L}(X, Y)$ est la topologie engendrée par les semi-normes $\|\cdot\|_x$, $x \in X$ définies par $\|T\|_x := \|Tx\|$. Cette topologie se note `SOT`. Autrement dit : `SOT` est la topologie sur $\mathcal{L}(X, Y)$ induite par la topologie de la convergence simple sur $\mathcal{F}(X, Y)$.

(3) Soit Ω un espace topologique, et soit $\mathcal{C}(\Omega) := \{f : \Omega \to \mathbb{C} \text{ continues}\}$. Pour tout compact $K \subseteq \Omega$, soit $\|\cdot\|_K$ la semi-norme sur $\mathcal{C}(\Omega)$ définie par $\|f\|_K := \sup\{|f(x)|;\ x \in K\}$. La **topologie de la convergence uniforme sur tout compact** sur $\mathcal{C}(\Omega)$ est la topologie engendrée par les semi-normes $\|\cdot\|_K$.

(4) Soit Ω un ouvert de $\mathbb{R}^d$, et notons $\mathcal{C}^\infty(\Omega)$ l'espace vectoriel constitué par toutes les fonctions $f : \Omega \to \mathbb{C}$ de classe $\mathcal{C}^\infty$. Pour tout compact $K \subseteq \Omega$ et pour tout $n \in \mathbb{N}$, notons $\|\cdot\|_{K,n}$ la semi-norme sur $\mathcal{C}^\infty(\Omega)$ définie par
$$\|f\|_{K,n} := \sup\big\{|\partial^\alpha f(x)|;\ |\alpha| \leqslant n\,,\ x \in K\big\}.$$
La **topologie** $\mathcal{C}^\infty$ sur $\mathcal{C}^\infty(\Omega)$ est la topologie engendrée par les semi-normes $\|\cdot\|_{K,n}$. Ainsi, une sg $(f_p) \subseteq \mathcal{C}^\infty(\Omega)$ converge vers $f \in \mathcal{C}^\infty(\Omega)$ en topologie $\mathcal{C}^\infty$ si et seulement si $\partial^\alpha f_p \to \partial^\alpha f$ uniformément sur tout compact pour tout $\alpha \in \mathbb{N}^d$ (micro-exo).

4. Plus de choses sur les produits

4.1. Produits dénombrables.

PROPOSITION 4.1. *Soit* $(X_i)_{i\in I}$ *une famille* dénombrable *d'espaces topologiques, et soit* $X := \prod_{i\in I} X_i$.

(1) *Si les* X_i *sont métrisables, alors* X *est métrisable.*

(2) *Si les* X_i *sont complètement métrisables, alors* X *aussi.*

(3) *Si les* X_i *sont séparables, alors* X *aussi.*

(4) *Si les* X_i *sont polonais, alors* X *aussi.*

Démonstration. On va faire la preuve dans le cas où I est infini, ce qui revient à supposer que $I = \mathbb{N}$. (Le cas où I fini est plus simple.)
(1) Pour tout $i \in \mathbb{N}$, soit δ_i une distance sur X_i définissant la topologie de X_i, avec de plus $\delta_i(u,v) \leqslant 1$ pour tous $u, v \in X_i$. Pour $x, y \in X$, on pose

$$\delta(x,y) := \sum_{i=0}^{\infty} 2^{-i}\delta_i\big(x(i), y(i)\big).$$

On vérifie (exo) que δ est une distance sur X. Il faut montrer que δ définit la topologie produit, *autrement dit* qu'une s.g. $(x_p)_{p\in P} \subseteq X$ converge vers $a \in X$ pour la topologie produit si et seulement si elle converge pour δ.
Supposons que $x_p \to a$ pour la topologie produit, *i.e.* $x_p(i) \to a(i)$ pour tout $i \in \mathbb{N}$. Soit $\varepsilon > 0$. Pour tout $N \in \mathbb{N}$ et pour tout p, on a

$$\delta(x_p, a) \leqslant \sum_{i=0}^{N} 2^{-i}\delta_i\big(x_p(i), a(i)\big) + \sum_{i>N} 2^{-i}.$$

Donc, si on fixe N assez grand, on a

$$\forall p \; : \; \delta(x_p, a) \leqslant \sum_{i=0}^{N} 2^{-i}\delta_i\big(x_p(i), a(i)\big) + \varepsilon/2.$$

Comme $\delta_i\big(x_p(i), a(i)\big) \to 0$ pour $i = 0, \dots, N$, on en déduit (exo) que $\delta(x_p, a) \to 0$.
Inversement, si $\delta(x_p, a) \to 0$, alors $x_p(i) \to a(i)$ pour tout $i \in \mathbb{N}$ car $\delta_i\big(x_p(i), a(i)\big) \leqslant 2^i\delta(x_p, a)$; donc $x_p \to a$ pour la topologie produit.
(2) Avec les notations de (1), il faut juste montrer que si (X_i, δ_i) est complet pour tout $i \in I$, alors (X, δ) est complet ; autrement dit, *d'après* (1), que toute suite $(x_n) \subseteq X$ qui est de Cauchy pour δ converge coordonnée par coordonnée. On laisse cela en exo.
(3) Pour $i \in \mathbb{N}$, soit D_i un ensemble dénombrable dense dans X_i. Soit aussi $a \in X$ fixé. Pour tout ensemble fini $F \subseteq \mathbb{N}$, soit

$$\mathcal{D}_F := \big\{x \in X;\; x_i \in D_i \text{ pour tout } i \in F \text{ et } x_i = a_i \text{ pour tout } i \notin F\big\}.$$

Alors chaque $\mathcal{D}_F$ est dénombrable (exo), donc $\mathcal{D} := \bigcup_{F\in\mathtt{FIN}(\mathbb{N})} \mathcal{D}_F$ aussi ; et on vérifie que $\mathcal{D}$ est dense dans X (autre exo).
(4) est immédiat par (2) et (3). □

REMARQUE. Si I est un ensemble non dénombrable, alors $\mathcal{F}(I, \mathbb{R}) = \mathbb{R}^I$ n'est *pas* métrisable pour la topologie de la convergence simple.

Démonstration. Soit $\mathcal{A} := \{\mathbf{1}_F;\ F \subseteq I \text{ fini}\} \subseteq \mathcal{F}(I, \mathbb{R})$. Alors la fonction $\mathbf{1}$ appartient à $\overline{\mathcal{A}}$ car la suite généralisée $(\mathbf{1}_F)_{F\in \mathtt{FIN}(I)}$ tend vers $\mathbf{1}$ (exo). Mais aucune *suite* $(\mathbf{1}_{F_n})_{n\in\mathbb{N}}$ de points de $\mathcal{A}$ ne peut tendre vers $\mathbf{1}$. En effet : l'ensemble $\bigcup_{n\in\mathbb{N}} F_n$ est dénombrable et I ne l'est pas, donc on peut trouver $i_0 \in I\backslash \bigcup_{n\in\mathbb{N}} F_n$, et alors $\mathbf{1}_{F_n}(i_0) = 0 \not\to 1$. □

Corollaire 4.2. *Soit E un espace topologique complètement métrisable. Si $G \subseteq E$ est un ensemble G_δ, i.e. une intersection dénombrable d'ouverts, alors G – muni de la topologie induite par celle de E – est complètement métrisable.*

Démonstration. Écrivons $G = \bigcap_{i\in\mathbb{N}} V_i$, où les V_i sont des ouverts de E. Les V_i sont complètement métrisables (vu en exo). Soit

$$\widetilde{G} := \left\{ x \in \prod_{i\in\mathbb{N}} V_i;\ x_i = x_j \text{ pour tous } i, j \in \mathbb{N} \right\}.$$

Alors $\widetilde{G}$ est fermé dans $\prod_{i\in I} V_i$, donc complètement métrisable. De plus, si on définit $\phi : G \to \widetilde{G}$ par $\phi(z) := (z, z, z \dots)$, alors on vérifie que ϕ est un homéomorphisme de G sur $\widetilde{G}$ (de réciproque donnée par $\phi^{-1}(x) := x_0$). Donc G est complètement métrisable puisque $\widetilde{G}$ l'est. □

Exercice. Montrer que si E est un espace topologique métrisable, alors tout fermé de E est G_δ.

Remarque 4.3. Il se trouve que la réciproque du Corollaire 4.2est vraie : si $G \subseteq E$ est complètement métrisable, alors G est un G_δ de E. Ainsi, un ensemble $G \subseteq E$ est complètement métrisable *si et seulement si* G est un G_δ de E. Ce résultat s'appelle le **Théorème d'Alexandrov**.

Démonstration. Soit d une distance définissant la topologie de E, et soit δ une distance sur G topologiquement équivalente à $d_{|G\times G}$ telle que (G, δ) soit complet. On notera B_δ les boules ouvertes de G relativement à la distance δ, et $\delta\text{-}\mathrm{diam}(A)$ le diamètre d'un ensemble $A \subseteq G$ relativement δ. De même, on notera B_d les boules ouvertes de E relativement à la distance d.

Observons d'abord que $\overline{G}$ est un G_δ de E car c'est un fermé de E et E est métrisable. Donc, pour montrer que G est un G_δ de E, il suffit de montrer que c'est un G_δ *de* $\overline{G}$ (micro-exo). On peut donc en fait supposer que $\overline{G} = E$, *i.e.* que *G est dense dans E.*

Pour $n \in \mathbb{N}^*$, posons

$$O_n := \left\{ x \in E;\ \exists V \text{ voisinage ouvert de } x \quad \text{tel que} \quad \delta\text{-}\mathrm{diam}(V \cap G) < \frac{1}{n} \right\}.$$

Par définition, les O_n sont des ouverts de E. On va montrer que $G = \bigcap_{n\geqslant 1} O_n$.
Soit $x \in G$, et soit $n \in \mathbb{N}^*$ fixé. Si on pose $B := B_\delta\big(x, \frac{1}{3n}\big)$, alors B est un ouvert de G tel que $\delta\text{-}\mathrm{diam}(B) \leqslant \frac{2}{3n} < \frac{1}{n}\cdot$ Donc, si on choisit un ouvert V de E tel que $B = V \cap G$, alors V témoigne que $x \in O_n$. Ceci étant vrai pour tout $n \in \mathbb{N}^*$, on a donc montré l'inclusion $G \subseteq \bigcap_{n\geqslant 1} O_n$.
Inversement, soit $x \in \bigcap_{n\geqslant 1} O_n$ quelconque, et montrons que $x \in G$. Pour tout $n \geqslant 1$, choisissons un ouvert $V_n \subseteq E$ tel que $x \in V_n$ et $\delta\text{-}\mathrm{diam}(V_n \cap G) < \frac{1}{n}\cdot$ Quitte à remplacer V_n par $V_n \cap B_d\big(x, \frac{1}{n}\big)$, on peut supposer que $V_n \subseteq B_d\big(x, \frac{1}{n}\big)$; et quitte à

remplacer ensuite V_n par $V_n \cap V_{n-1}$ pour $n \geqslant 2$, on peut également supposer que la suite $(V_n)_{n\geqslant 1}$ est décroissante.
Comme G est dense dans E, on a $V_n \cap G \neq \varnothing$ pour tout $n \in \mathbb{N}$, donc on peut choisir un point $x_n \in V_n \cap G$. Alors $x_n \to x$ car $x_n \in V_n \subseteq B_d\big(x, \frac{1}{n}\big)$ De plus, si $p, q \in \mathbb{N}$, alors x_p et x_q sont tous les deux dans $V_{\min(p,q)} \cap G$ car la suite (V_n) est décroissante, donc

$$\delta(x_p, x_q) \leqslant \delta\text{-}\operatorname{diam}(V_{\min(p,q)} \cap G) \xrightarrow{p,q\to\infty} 0.$$

Par conséquent, la suite (x_n) est de Cauchy dans (G, δ), qui est complet ; donc (x_n) converge pour δ vers un certain point $x_\infty \in G$. Mais δ est topologiquement équivalente à $d_{|G\times G}$, donc $x_n \to x_\infty$ pour d ; et donc $x_\infty = x$ car $x_n \to x$ pour d. Donc $x \in G$! $\square$

Corollaire 4.4. *Si Ω est un ouvert de $\mathbb{R}^d$, alors $\mathcal{C}(\Omega)$ est polonais (pour la topologie de la convergence uniforme sur tout compact).*

Démonstration. Soit $(K_i)_{i\in\mathbb{N}}$ une **suite exhaustive de compacts** pour Ω, *i.e.* (K_i) est une suite croissante de compacts de Ω telle que tout compact $K \subseteq \Omega$ est contenu dans un K_i. Alors chaque $\mathcal{C}(K_i)$ est un espace de Banach *séparable*, donc un espace polonais ; donc $X := \prod_{i\in\mathbb{N}} \mathcal{C}(K_i)$ est un espace polonais. Soit $\Phi : \mathcal{C}(\Omega) \to X$ l'application définie par $\Phi(f) := \big(f_{|K_i}\big)_{i\in\mathbb{N}}$. On a

$$\Phi\big(\mathcal{C}(\Omega)\big) = \big\{\bar{f} = (f_i)_{i\in\mathbb{N}} \in X;\ f_j \equiv f_i \text{ sur } K_i \text{ pour tous } i \leqslant j\big\};$$

donc $\Phi\big(\mathcal{C}(\Omega)\big)$ est fermé dans X, et donc $\Phi\big(\mathcal{C}(\Omega)\big)$ est un espace polonais. Par définition de la topologie de $\mathcal{C}(\Omega)$ et de la topologie produit sur X, on voit (exo) que Φ est un homéomorphisme de $\mathcal{C}(\Omega)$ sur $\Phi\big(\mathcal{C}(\Omega)\big)$; donc $\mathcal{C}(\Omega)$ est polonais. $\square$

Corollaire 4.5. *Soit K un espace topologique compact. S'il existe une famille dénombrable $(f_i)_{i\in I}$ de fonctions continues sur K qui **sépare les points de** K, alors K est métrisable.*

Démonstration. Soit $\Phi : K \to \mathbb{C}^I$ l'application définie par $\Phi(x) := (f_i(x))_{i\in I}$. Alors Φ est continue, et elle est *injective* car les f_i séparent les points de K. Comme K est compact, on en déduit que K est un homéomorphisme de K sur $\Phi(K)$. Ainsi, K est homéomorphe à une partie de $\mathbb{C}^I$ et est donc métrisable. $\square$

Proposition 4.6. *Tout espace topologique* métrisable séparable *est homéomorphe à une partie de* $[0,1]^{\mathbb{N}}$.

Démonstration. Soit E un espace topologique métrisable séparable. Soit δ une distance définissant la topologie de E telle que $\delta(x, y) \leqslant 1$ pour tous $x, y \in E$, et soit $D = \{z_i;\ i \in \mathbb{N}\}$ un ensemble dénombrable et dense dans E. On définit $\Phi : E \to [0,1]^{\mathbb{N}}$ par

$$\Phi(x) := \big(\delta(x, z_i)\big)_{i\in\mathbb{N}}.$$

Alors Φ est continue, et elle est aussi injective : en effet, si $x \neq y$, on peut trouver $i \in \mathbb{N}$ tel que $\delta(z_i, x) < \delta(x, y)/2$, et alors $\delta(z_i, x) < \delta(z_i, y)$ car $\delta(z_i, y) \geqslant \delta(x, y) - \delta(x, z_i) > \delta(x, y)/2$; donc $\Phi(x) \neq \Phi(y)$. Enfin, $\Phi^{-1} : \Phi(E) \to E$ est continue. En effet, supposons que $(x_n)_{n\in\mathbb{N}}$ soit une suite de points de E telle que $\Phi(x_n) \to \Phi(x)$ où $x \in E$; autrement dit, $\delta(z_i, x_n) \to \delta(z_i, x)$ pour tout $i \in \mathbb{N}$. Si (x_n) ne tendait pas vers x, on pourrait trouver une sous-suite (x_{n_k}) telle que $\delta(x_{n_k}, x) \geqslant \varepsilon > 0$ pour tout k. En choisissant $i \in \mathbb{N}$ tel que $\delta(z_i, x) < \varepsilon/2$, on aurait alors $\delta(x_{n_k}, z_i) \geqslant \delta(x_{n_k}, x) - \delta(z_i, x) > \varepsilon/2$ pour tout k, ce qui est absurde puisque $\delta(z_i, x_n) \to \delta(z_i, x)$ quand $n \to \infty$. $\square$

4.2. Produits de compacts. Le théorème suivant est sans doute le résultat le plus important concernant la compacité.

THÉORÈME 4.7. (Tychonoff)
Tout produit de compacts est compact : si $(X_i)_{i\in I}$ est une famille quelconque d'espaces topologiques compacts, alors $X := \prod_{i\in I} X_i$ est compact.

Démonstration. On a vu en exo que X est séparé. Soit $(x_p)_{p\in P}$ une s.g. dans X : on veut montrer que (x_p) possède une valeur d'adhérence.
Pour tout $\varnothing \neq J \subseteq I$, on posera $X_J := \prod_{i\in J} X_i$. En particulier, on a $X_I = X$ et $X_{\{i\}} = X_i$ pour tout $i \in I$. Si $J \subseteq J' \subseteq I$ et $z = (z(i))_{i\in J'} \in X_{J'}$, on pose $z_{|J} := (z(i))_{i\in J}$.
Soit

$$\mathcal{E} := \Big\{(J,z);\ \varnothing \neq J \subseteq I\ , z \in X_J \text{ et } z \text{ est une valeur d'adhérence de } \big((x_p)_{|J}\big)_{p\in P}\Big\}.$$

FAIT 0. L'ensemble $\mathcal{E}$ est non vide.

Preuve du Fait 0. Soit $i_0 \in I$. Comme X_{i_0} est compact, la s.g. $(x_p(i_0))_{p\in P}$ possède une valeur d'adhérence $z \in X_{i_0}$; et alors $(\{i_0\}, z) \in \mathcal{E}$. □

On définit une relation d'ordre $\leqslant$ sur $\mathcal{E}$ de la façon suivante :

$$(J,z) \leqslant (J',z') \qquad \text{ssi} \qquad J \subseteq J' \quad \text{et} \quad z'_{|J} = z.$$

FAIT 1. $\mathcal{E}$ possède un élément maximal pour $\leqslant$.

Preuve du Fait 1. D'après le *Lemme de Zorn*, il suffit de montrer que l'ordre $\leqslant$ est inductif. Soit $\mathcal{C} = \{(J_s, z_s);\ s \in S\}$ une partie de $\mathcal{E}$ totalement ordonnée pour $\leqslant$. Posons $J := \bigcup_{s\in S} J_s$. Comme $\mathcal{C}$ est totalement ordonnée, il existe un $z \in X_J$ bien défini tel que $z_{|J_s} = z_s$ pour tout $s \in S$ (exo). Montrons que z est une valeur d'adhérence de $\big((x_p)_{|J}\big)_{p\in P}$. Soit V un voisinage de z dans $X_J = \prod_{i\in J} X_i$, et soit $p_0 \in P$: on cherche $p \geq p_0$ tel que $(x_p)_{|J} \in V$. Soient $i_1, \dots, i_r \in J$ et $V_{i_1}, \dots, V_{i_r}$ ouverts dans $X_{i_1}, \dots, X_{i_r}$ tels que $z \in$ "$V_{i_1} \times \cdots \times V_{i_r}$" $\subseteq V$. Comme $\mathcal{C}$ est totalement ordonné, on peut trouver $s \in S$ tel que $\{i_1, \dots, i_r\} \subseteq J_s$. Comme $z_{|J_s} = z_s$, on peut considérer "$V_{i_1} \times \cdots \times V_{i_r}$" comme un voisinage de z_s dans $X_{J_s} = \prod_{i\in J_s} X_i$; et comme z_s est une valeur d'adhérence de $\big((x_p)_{|J_s}\big)$, on peut trouver $p \geq p_0$ tel que $(x_p)_{|J_s} \in$ "$V_{i_1} \times \cdots \times V_{i_r}$". En considérant maintenant à nouveau "$V_{i_1} \times \cdots \times V_{i_r}$" comme une partie de X_J, on a alors $(x_p)_{|J} \in$ "$V_{i_1} \times \cdots \times V_{i_r}$" $\subseteq V$. Ainsi z est bien une valeur d'adhérence de $\big((x_p)_{|J}\big)$, donc $(J,z) \in \mathcal{E}$; et par définition, (J,z) majore tous les éléments de $\mathcal{C}$. □

FAIT 2. Si $(J,z) \in \mathcal{E}$ et $J \neq I$, alors (J,z) n'est pas maximal pour $\leqslant$.

Preuve du Fait 2. Comme z est une valeur d'adhérence de $\big((x_p)_{|J}\big)$, on peut trouver une sous-s.g. $(u_{p'})_{p'\in P'}$ de (x_p) telle que $(u_{p'})_{|J} \to z$. Maintenant, soit $i_0 \in I\backslash J$. Comme X_{i_0} est compact, on peut trouver une sous-s.g. $(v_{p''})_{p''\in P''}$ de $(u_{p'})$ et un point $a_0 \in X_{i_0}$ tels que $v_{p''}(i_0) \to a_0$. Si on définit $\widetilde{z} \in X_{J\cup\{i_0\}}$ par $\widetilde{z}_{|J} := z$ et $\widetilde{z}(i_0) := a_0$, alors $(v_{p''})_{|J\cup\{i_0\}} \to \widetilde{z}$. Donc $\widetilde{z}$ est une valeur d'adhérence de $\big((x_p)_{|J\cup\{i_0\}}\big)$ car $(v_{p''})$ est une sous-s.g. de (x_p) ; et donc, $(J\cup\{i_0\}, \widetilde{z})$ appartient à $\mathcal{E}$ et majore strictement (J,z). □

On peut maintenant terminer la preuve du Théorème de Tychonoff. Par les Faits 1 et 2, on peut trouver un élément de $\mathcal{E}$ de la forme (I,z). Cela signifie que $z \in X_I = X$ et que z est une valeur d'adhérence de $\big((x_p)_{|I}\big) = (x_p)$. □

Corollaire 4.8. *Soit I un ensemble quelconque, et soit $(f_p)_{p\in P}$ une s.g. de fonctions, $f_p : I \to \mathbb{C}$. On suppose que (f_p) est* ***simplement bornée****, i.e. que pour tout $i \in I$, la s.g. $(f_p(i))_{p\in P} \subseteq \mathbb{C}$ est bornée. Alors (f_p) possède une sous-s.g. $(g_{p'})_{p'\in P'}$ qui converge simplement vers une fonction $f : I \to \mathbb{C}$. Si de plus I est* dénombrable *et si $(f_n)_{n\in\mathbb{N}}$ est une* suite *de fonctions simplement bornée, alors (f_n) possède une* sous-suite *(f_{n_k}) qui converge simplement sur I.*

Démonstration. Pour $i \in I$, soit $M_i := \sup\{|f_p(i)|;\ p \in P\}$ et $K_i := \overline{D}(0, M_i) \subseteq \mathbb{C}$. Alors on peut considérer les f_p comme des éléments de $X := \prod_{i\in I} K_i$, qui est compact ; d'où la 1ère partie. Si de plus I est dénombrable, alors X est compact et *métrisable*, ce qui donne la 2ème partie. □

Exercice. Démontrer directement la 2ème partie du Corollaire 4.8, sans utiliser la topologie produit. (La formule magique est : “procédé diagonal”.)

Exemple 1. (Théorème d'Ascoli)
Soit K un espace topologique compact, et soit $\mathcal{F} \subseteq \mathcal{C}(K)$. On suppose que la famille $\mathcal{F}$ est simplement bornée, et que de plus $\mathcal{F}$ est *équicontinue* en tout point, autrement dit que pour tout $z \in K$ et pour tout $\varepsilon > 0$, on peut trouver un voisinage ouvert V_z de z dans K tel que $\forall x \in V_z\ \forall f \in \mathcal{F}\ :\ |f(x) - f(z)| < \varepsilon$. Alors toute suite $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{F}$ possède une sous-suite uniformément convergente.

Démonstration. L'adhérence de $\mathcal{F}$ dans $(\mathcal{C}(K), \|\cdot\|_\infty)$ vérifie les mêmes hypothèses que $\mathcal{F}$ (exo) ; donc on peut supposer que $\mathcal{F}$ est une partie fermée de $\mathcal{C}(K)$. Il s'agit alors de montrer que $\mathcal{F}$ est un compact de $\mathcal{C}(K)$. On va le faire en montrant que toute suite généralisée $(f_p) \subseteq \mathcal{F}$ possède une sous-s.g. qui converge dans $\mathcal{C}(K)$.

Soit $(f_p)_{p\in P}$ une s.g. d'éléments de $\mathcal{F}$. Alors (f_p) est simplement bornée, donc elle possède une sous-s.g. $(g_{p'})_{p'\in P'}$ qui converge simplement vers une fonction $f : K \to \mathbb{C}$. On va montrer qu'en fait $g_{p'} \to f$ *uniformément*, ce qui terminera la preuve.
Soit $\varepsilon > 0$. Par équicontinuité de $\mathcal{F}$ et par compacité de K, on peut trouver $z_1, \dots, z_N \in K$ et des ouverts $V_i \ni z_i$ tels que $V_1 \cup \cdots \cup V_N = K$ et $\forall i\ \forall x \in V_i\ \forall p' \in P'\ :\ |g_{p'}(x) - g_{p'}(z_i)| \leqslant \varepsilon$. Comme $g_{p'} \to f$ simplement, ceci entraine que $\forall i\ \forall x \in V_i\ :\ |f(z_i) - f(x)| \leqslant \varepsilon$. Ensuite, comme $g_{p'}(z_i) \to f(x_i)$ pour $i = 1, \dots, N$ et comme P' est filtrant, on peut trouver $p'_0 \in P'$ tel que $\forall p' \succeq p'_0\ \forall i\ :\ |g_{p'}(z_i) - f(z_i)| \leqslant \varepsilon$. Un coup d'inégalité triangulaire montre alors que $\forall p' \succeq p'_0\ \forall i\ \forall x \in V_i\ :\ |g_{p'}(x) - f(x)| \leqslant 3\varepsilon$; autrement dit $\forall p' \succeq p'_0\ \forall x \in K\ :\ |g_{p'}(x) - f(x)| \leqslant 3\varepsilon$ puisque $V_1 \cup \cdots \cup V_N = K$. Donc, en effet, $g_{p'} \to f$ uniformément. □

Exemple 2. Pour $n \in \mathbb{N}$, soit $f_n : [0, 2\pi] \to \mathbb{C}$ définie par $f_n(t) := e^{int}$. Alors la suite (f_n) est simplement bornée, mais ne possède pas de *sous-suite* simplement convergente.

Démonstration. Supposons qu'il existe une sous-suite (f_{n_k}) qui converge simplement vers une fonction $f : [0, 2\pi] \to \mathbb{C}$, *i.e.* $e^{in_k t} \to f(t)$ pour tout $t \in [0, 2\pi]$. Alors f est borélienne et $|f(t)| \equiv 1$, donc f est intégrable sur $[0, 2\pi]$. Par le Lemme de Riemann-Lebesgue, on en déduit que $\int_0^{2\pi} f(t)e^{-in_k t}dt \to 0$ quand $k \to \infty$. Mais par ailleurs, $f(t)e^{-in_k t} \to |f(t)|^2 = 1$ pour tout $t \in [0, 2\pi]$ et $|f(t)e^{-in_k t}| \equiv 1$; donc, par convergence dominée, $\int_0^{2\pi} f(t)e^{-in_k t}dt \to \int_0^{2\pi} 1\, dt = 2\pi$. On obtient donc une contradiction. □

5. L'espace de Cantor

Définition 5.1. *L'**espace de Cantor** est l'espace $\Delta := \{0,1\}^{\mathbb{N}}$ de toutes les suites infinies de* 0 *et de* 1*, muni de la topologie produit.*

Fait 5.2. *L'espace Δ est compact et métrisable.*

Démonstration. C'est évident puisque $\{0,1\}$ est certainement compact métrisable et que $\mathbb{N}$ est dénombrable. $\square$

Lemme 5.3. *Notons $\mathcal{S}$ l'ensemble de toutes les suites* finies *de* 0 *et de* 1*. Pour $s = (s_0, \dots, s_n)$, posons $W_s := \big\{\alpha \in \Delta;\ \alpha$ commence par $s\big\}$. Alors les W_s sont ouverts et fermés dans Δ, et forment une base pour la topologie de Δ.*

Démonstration. On a $W_s = \big\{\alpha \in \Delta;\ \alpha(i) = s_i$ pour $i = 0, \dots, n\big\}$, donc W_s est ouvert et fermé car les projections canoniques $\alpha \mapsto \alpha(i)$ sont continues de Δ dans $\{0,1\}$ et que les singletons $\{s_i\}$ sont ouverts et fermés dans $\{0,1\}$. Si $\alpha \in \Delta$ et si V est un voisinage ouvert de α dans Δ, alors on peut trouver un ouvert élémentaire "$V_0 \times \cdots \times V_n$" tel que $\alpha \in$ "$V_0 \times \cdots \times V_n$" $\subseteq V$ (les V_i sont des ouverts de $\{0,1\}$). Alors $\alpha(i) \in V_i$ pour $i = 0, \dots, n$; donc, si on pose $s := (\alpha(0), \dots, \alpha(n))$, on a $\alpha \in W_s \subseteq$ "$V_0 \times \cdots \times V_n$" $\subseteq V$. Ainsi, les W_s forment une base pour la topologie de Δ. $\square$

Corollaire 5.4. *L'espace Δ est **totalement discontinu** : les seules parties connexes de Δ sont $\varnothing$ et les singletons.*

Démonstration. Il s'agit de montrer que si $A \subseteq \Delta$ contient au moins 2 points $\alpha \neq \beta$, alors A n'est pas connexe. Par le lemme, on peut trouver $s \in \mathcal{S}$ tel que $\alpha \in W_s$ et $\beta \notin W_s$. Alors $W_s \cap A$ est ouvert et fermé dans A, non-vide et différent de A ; donc A n'est pas connexe. $\square$

Proposition 5.5. *Tout espace topologique X complètement métrisable et sans point isolé contient une "copie" de Δ, i.e. il existe un compact $K \subseteq X$ homéomorphe à Δ.*

Démonstration. Soit d une distance définissant la topologie de X et telle que (X,d) soit complet. Comme X n'a pas de point isolé, il contient au moins 2 points ; et donc, on peut trouver deux ouverts non vides $V_0, V_1 \subseteq X$ tels que $\overline{V_0} \cap \overline{V_1} = \varnothing$, avec de plus $\mathrm{diam}(\overline{V_0}) \leqslant \frac{1}{2}$ et $\mathrm{diam}(\overline{V_1}) \leqslant \frac{1}{2}\cdot$ On peut faire de même dans V_0 et V_1, et ainsi de suite. De façon précise, en notant $\mathcal{S}$ l'ensemble de toutes les suites finies de 0 et de 1, on construit une famille $(V_s)_{s \in \mathcal{S}}$ d'ouverts non vides de X telle que pour tout $s \in \mathcal{S}$, les choses suivantes aient lieu :

- $\overline{V_{s0}} \cup \overline{V_{s1}} \subseteq V_s$ et $\overline{V_{s0}} \cap \overline{V_{s1}} = \varnothing$;
- $\mathrm{diam}(\overline{V_s}) \leqslant 2^{-|s|}$, où $|s|$ est la longueur de s.

Si $\alpha \in \Delta$ alors, par le théorème des fermés emboités, l'intersection $\bigcap_{n=0}^{\infty} \overline{V_{(\alpha_0, \dots, \alpha_n)}}$ est non vide et réduite à un point x_α. On peut donc poser $j(\alpha) := x_\alpha$, ce qui définit un application $j : \Delta \to X$.

L'application $j : \Delta \to X$ est *continue.* En effet, soit $\alpha \in \Delta$ quelconque. Si V est un voisinage ouvert de $j(\alpha) = x_\alpha$, alors on peut trouver $n \in \mathbb{N}$ tel que $V_{(\alpha_0, \dots, \alpha_n)} \subseteq V$ car $\mathrm{diam}(V_{(\alpha_0, \dots, \alpha_n)}) \to 0$ quand $n \to \infty$. Si on pose $s := (\alpha_0, \dots, \alpha_n)$, alors toute suite $\beta \in \Delta$ qui commence par s est telle que $j(\beta) = x_\beta \in V_s \subseteq V$; donc W_s est un voisinage ouvert de α tel que $j(W_s) \subseteq V$.

L'application j est de plus *injective*. En effet, si $\alpha \neq \beta$, soit n le plus petit indice tel que $\alpha_n \neq \beta_n$. Alors $V_{(\alpha_0,\dots,\alpha_n)} \cap V_{(\beta_0,\dots,\beta_n)} = \emptyset$, et donc $j(\alpha) = x_\alpha \neq x_\beta = j(\beta)$ puisque $x_\alpha \in V_{(\alpha_0,\dots,\alpha_n)}$ et $x_\beta \in V_{(\beta_0,\dots,\beta_n)}$.
Comme Δ est compact, on peut maintenant conclure que $K := j(\Delta) \subseteq X$ est un compact homéomorphe à Δ. □

COROLLAIRE 5.6. *Tout espace polonais non dénombrable contient une copie de Δ.*

Démonstration. Soit X un espace polonais non dénombrable. Il suffit de montrer que X contient un fermé non vide F sans point isolé. En effet, F est alors un espace complètement métrisable et sans point isolé, donc contient une copie de Δ par la proposition.

FAIT. Notons O l'ensemble de tous les points $x \in X$ possèdant un voisinage ouvert *dénombrable*. Alors O est un ouvert de X, et O est dénombrable.

Preuve du Fait. Pour tout $x \in X$, choisissons un voisinage ouvert V_x de x qui soit dénombrable. Alors $V_x \subseteq O$ pour tout $x \in O$, puisque V_x est dénombrable et est un voisinage de chacun de ses points. On a donc $O = \bigcup_{x \in O} V_x$, ce qui prouve que O est un ouvert de X. Enfin, O est dénombrable par la propriété de Lindelöf (exo). □

Soit $F := X \backslash O$. Par le Fait, F est un fermé de X et $F \neq \emptyset$ car X est non dénombrable. De plus, si $x \in F$, alors tout voisinage ouvert V de x dans X est non dénombrable, par définition de F ; en particulier, $V \cap F$ n'est pas réduit à $\{x\}$ car sinon $V = (V \cap F) \cup (V \cap O)$ serait dénombrable. Donc F n'a pas de point isolé. □

REMARQUE. Soit $j : \Delta \to \mathbb{R}$ l'application définie par

$$j(\alpha) := \sum_{i=0}^{\infty} \frac{2\alpha_i}{3^{i+1}}.$$

Alors j est injective et continue, donc $K_3 := j(\Delta)$ est homéomorphe à Δ. L'ensemble K_3 est l'**ensemble triadique de Cantor**.

Démonstration. Exo. □

THÉORÈME 5.7. *Tout espace compact métrisable est* image continue *de Δ. Autrement dit : si K est un espace compact métrisable, alors il existe une surjection continue $\phi : \Delta \twoheadrightarrow K$.*

Démonstration.

FAIT 1. Il existe une surjection continue de Δ sur $[0,1]$.

Preuve du Fait 1. Soit $s : \Delta \to \mathbb{R}$ l'application définie par

$$s(\alpha) := \sum_{i=0}^{\infty} \frac{\alpha(i)}{2^{i+1}}.$$

L'application s est continue car les sommes partielles de la série sont des fonctions continues de $\alpha \in \Delta$ et la série converge normalement sur Δ. De plus, s est à valeurs dans $[0,1]$ car $\sum_{i=0}^{\infty} 2^{-i-1} = 1$. Enfin, $s(\Delta) = [0,1]$ car tout nombre $x \in [0,1]$ possède un développement en base 2. □

Fait 2. Δ est homéomorphe à $\Delta^{\mathbb{N}}$.

Preuve du Fait 2. Soit $(i,j) \mapsto n_{i,j}$ une bijection de $\mathbb{N}\times\mathbb{N}$ sur $\mathbb{N}$, de réciproque $n \mapsto (i_n, j_n)$. Définissons $\Phi : \Delta^{\mathbb{N}} \to \Delta$ comme suit : si $\bar{x} = (\bar{x}_0, \bar{x}_1, \dots) \in \Delta^{\mathbb{N}}$, alors

$$\Phi(\bar{x}) := \big(\bar{x}_{i_n}(j_n)\big)_{n\in\mathbb{N}}$$

Alors Φ est une bijection de $\Delta^{\mathbb{N}}$ sur Δ, de réciproque $\Delta \ni \alpha \mapsto (\bar{x}_0, \bar{x}_1, \dots)$, où $\bar{x}_i(j) := \alpha(n_{i,j})$. De plus, Φ est continue car chaque application $\bar{x} \mapsto \bar{x}_i(j)$ est continue, et Φ^{-1} est continue car chaque application $\alpha \mapsto \alpha(n)$ est continue. Donc Φ est un homéomorphisme de $\Delta^{\mathbb{N}}$ sur Δ. □

Fait 3. Il existe une surjection continue de Δ sur $[0,1]^{\mathbb{N}}$.

Preuve du Fait 3. Fait 1+Fait 2 (exo). □

Fait 4. Si F est un fermé de Δ, il existe une **rétraction continue** de Δ sur F, *i.e.* une application continue $r : \Delta \to F$ telle que $r(\beta) = \beta$ pour tout $\beta \in F$.

Preuve du Fait 4. Soit d la distance sur Δ définie par

$$d(\alpha,\beta) := \sum_{i=0}^{\infty} \frac{|\beta(i) - \alpha(i)|}{3^i}.$$

La distance d définit la topologie de Δ (exo) ; en particulier, si $\alpha \in \Delta$ est fixé, la fonction $\beta \mapsto d(\alpha,\beta)$ est continue sur le compact F, donc il existe au moins 1 point $\beta \in F$ tel que $d(\alpha,\beta) = \mathrm{dist}(\alpha, F)$. De plus, comme l'application $\{0,1\}^{\mathbb{N}} \ni \varepsilon \mapsto \sum_{i=0}^{\infty} \frac{\varepsilon_i}{3^i}$ est *injective* (exo), on voit que si α est fixé, alors

$$d(\alpha,\beta) = d(\alpha,\beta') \implies |\beta(i) - \alpha(i)| = |\beta'(i) - \alpha(i)| \quad \text{pour tout } i \implies \beta = \beta'.$$

Donc, pour tout $\alpha \in \Delta$, il existe *exactement 1* point $\beta \in F$ tel que $d(\alpha,\beta) = \mathrm{dist}(\alpha,F)$. On note ce point $r(\alpha)$. Par définition, on a $r(\beta) = \beta$ pour tout $\beta \in F$. De plus, le graphe de l'application $r : \Delta \to F$ est fermé dans $\Delta \times F$, car

$$r(\alpha) = \beta \iff \beta \in F \quad \text{et} \quad \forall \beta' \in F \;:\; d(\alpha,\beta) \leqslant d(\alpha,\beta').$$

Donc le graphe de r est *compact*, et donc r est continue (exo classique). □

On peut maintenant finir la preuve du théorème. Soit K un espace compact métrisable quelconque. Alors K est métrisable séparable, donc est homéomorphe à une partie de $[0,1]^{\mathbb{N}}$. On peut donc supposer que $K \subseteq [0,1]^{\mathbb{N}}$. Soit $s : \Delta \twoheadrightarrow [0,1]^{\mathbb{N}}$ une surjection continue donnée par le Fait 3. Alors $F := s^{-1}(K)$ est un fermé de Δ ; donc le Fait 4 fournit un rétraction continue $r : \Delta \twoheadrightarrow F$. Alors $\phi := s \circ r : \Delta \to K$ est une surjection continue de Δ sur K (micro-exo). □

Exercice. Montrer que l'espace vectoriel engendré par les fonctions $\mathbf{1}_{W_s}$, $s \in \mathcal{S}$ est dense dans $\mathcal{C}(\Delta)$; et en déduire que si K est un espace métrique compact quelconque, alors l'espace de Banach $\mathcal{C}(K)$ est séparable.

Chapitre 3

Espaces vectoriels topologiques

1. Généralités

1.1. Groupes topologiques, espaces vectoriels topologiques.

Définition 1.1. *Un **groupe topologique** est un groupe $(G, \cdot)$ muni d'une topologie τ séparée telle que les applications $(g,h) \mapsto g \cdot h$ et $g \mapsto g^{-1}$ soient continues respectivement de $G \times G$ dans G et de G dans G. Un **espace vectoriel topologique** est un espace vectoriel X sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ muni d'une topologie τ séparée telle que les applications $(x,y) \mapsto x+y$ et $(\lambda, x) \mapsto \lambda x$ soient continues respectivement de $X \times X$ dans X et de $\mathbb{K} \times X$ dans X.*

Remarque. Évidemment, si X est un evt, alors $(X,+)$ est un groupe topologique, commutatif.

Exemple 1. Tout evn est un evt.

Démonstration. Exo. □

Exemple 2. Soit X un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et soit $(\| \cdot \|_i)_{i \in I}$ une famille de semi-normes sur X. On note τ la topologie définie par les $\| \cdot \|_i$. Alors (X, τ) est un evt si et seulement si la famille $(\| \cdot \|_i)$ est **séparante**, *i.e.* telle que $\forall x \in X \backslash \{0\}\ \exists i\ :\ \|x\|_i \neq 0$.

Démonstration. La continuité des opérations est laissée en exo. Supposons que la famille $(\| \cdot \|_i)$ soit séparante, et montrons que (X,τ) est séparé. Soient $x, y \in X$ avec $x \neq y$. Comme $(\| \cdot \|_i)$ est séparante, on peut trouver $i \in I$ tel que $\delta := \|y - x\|_i$ est > 0. Alors $V_x := B_i(x, \delta/2)$ et $V_y := B_i(y, \delta/2)$ sont des voisinages ouverts de x et y tels que $V_x \cap V_y = \varnothing$. Inversement, supposons que (X, τ) soit séparé. Soit $x \in X \backslash \{0\}$ quelconque. Comme τ est séparée, on peut trouver $i_1, \dots, i_r \in I$ et $\varepsilon > 0$ tels que $x \notin B_{i_1, \dots, i_r}(0, \varepsilon)$. Alors $\|x\|_{i_k} \geqslant \varepsilon > 0$ pour au moins 1 indice k ; ce qui montre que la famille $(\| \cdot \|_i)$ est séparante. □

Exemple 3. Le cercle $\mathbb{T} := \{z \in \mathbb{C};\ |z| = 1\}$ est un groupe topologique (pour la multiplication).

Exemple 4. Soit X un espace de Banach, et soit $GL(X) := \{T \in \mathcal{L}(X) \text{ inversibles}\}$. Alors $\big(GL(X), \| \cdot \|\big)$ est un groupe topologique (pour la composition).

Démonstration. C'est "bien connu". □

Exercice. Soit X un evt. Montrer que pour tout $a \in X$, l'application $x \mapsto x + a$ est un homéomorphisme de X sur X, et que pour tout $\lambda \in \mathbb{K} \backslash \{0\}$, l'application $x \mapsto \lambda x$ est un homéomorphisme de X sur X.

1.2. Petites choses utiles. Dans ce qui suit X est un evt. Pour $A, B \subseteq X$, $u \in X$ et $\lambda \in \mathbb{K}$, on pose

$$\begin{aligned} A + B &:= \{x + y;\ x \in A, y \in B\}, \\ u + A &:= \{u + x;\ x \in A\}, \\ \lambda A &:= \{\lambda x;\ x \in A\}. \end{aligned}$$

Exercice. Montrer que si $O \subseteq X$ est ouvert, alors $O + B$ est ouvert pour tout $B \subseteq X$.

Fait 1. Soit $a \in X$. Un ensemble $V \subseteq X$ est un voisinage (ouvert) de a si et seulement si $V = a + W$ où W est un voisinage (ouvert) de 0.

Démonstration. Par définition d'un evt, l'application $\tau_a : x \mapsto x + a$ est un homéomorphisme de X sur X. Donc, si V est un voisinage ouvert de a, alors $W := V - a = \tau_a^{-1}(V)$ est un voisinage ouvert de 0 ; et si W est un voisinage ouvert de 0, alors $a + W = \tau_a(V)$ est un voisinage ouvert de a. □

Fait 2. Si $W \subseteq X$ est un voisinage de 0 alors W est **absorbant** : pour tout $x \in X$, on peut trouver $A < \infty$ tel que $\forall \alpha \geqslant A\ :\ x \in \alpha W$.

Démonstration. Comme $rx \to 0$ quand $r \to 0^+$, on peut trouver $\delta > 0$ tel que $rx \in W$ pour tout $0 < r \leqslant \delta$, *i.e.* $x \in (1/r)W$. Alors $A := 1/\delta$ convient. □

Fait 3. Si W est un voisinage de 0 dans X, alors on peut trouver un voisinage W' de 0 tel que $W' + W' \subseteq W$ et $W' - W' \subseteq W$.

Démonstration. Par continuité des applications $(x, y) \mapsto x + y$ et $(x, y) \mapsto x - y$ en $(0, 0)$ et par définition de la topologie produit, on peut trouver U_1, V_1, U_2, V_2 voisinages ouverts de 0 dans X tels que $U_1 + V_1 \subseteq W$ et $U_2 - V_2 \subseteq W$. Alors $W' := U_1 \cap V_1 \cap U_2 \cap V_2$ convient. □

Fait 4. Si W est un voisinage de 0 dans X, alors on peut trouver un voisinage W' de 0 tel que $\overline{W'} \subseteq W$.

Démonstration. Soit W' un voisinage ouvert de 0 tel que $W' - W' \subseteq W$; montrons que W' convient. Soit $x \in \overline{W'}$ quelconque. Comme $x + W'$ est un voisinage ouvert de x, on a $(x + W') \cap W' \neq \varnothing$. Donc $x \in W' - W'$; et donc $x \in W$. □

Fait 5. Si W est un voisinage de 0 dans X, alors on peut trouver un voisinage W' de 0 tel que $W' \subseteq W$ et qui est **équilibré**, *i.e.* tel que $\lambda W' \subseteq W'$ pour tout $\lambda \in \mathbb{K}$ vérifiant $|\lambda| \leqslant 1$. Si de plus W est *convexe*, on peut imposer que W' soit lui aussi convexe.

Démonstration. Posons $\Lambda := \{\lambda \in \mathbb{K};\ |\lambda| \leqslant 1\}$, qui est un compact de $\mathbb{K}$. L'ensemble $\{(\lambda, x) \in X \times \Lambda;\ \lambda x \notin W\}$ est un fermé de $\Lambda \times X$ car W est ouvert dans X et l'aplication $(\lambda, x) \mapsto \lambda x$ est continue. Comme Λ est compact, on en déduit (exo) que l'ensemble $F := \{x \in X;\ \exists \lambda \in \Lambda\ :\ \lambda x \notin W\}$ est un fermé de X. Donc $W' := \{x \in X;\ \forall \lambda \in \Lambda\ :\ \lambda x \in W\} = X \backslash F$ est ouvert ; et on vérifie sans peine que W' convient (exo). □

Remarque. Les Faits 1, 3, 4 sont valables dans n'importe quel groupe topologique commutatif $(G, +, \tau)$.

1.3. Métrisabilité, complète métrisabilité.

PROPOSITION 1.2. *Soit X un evt. On suppose que X est métrisable. Alors il existe une distance d sur X qui définit la topologie de X et qui est de plus* ***invariante par translations*** *: pour tous $x, y, a \in X$, on a $d(x+a, y+a) = d(x,y)$.*

Démonstration. Fixons une distance ρ définissant la topologie de X. Soit également W un voisinage de 0 équilibré. On posera $W(0) := \{0\}$, $W(1) := W$ et, pour $k \in \mathbb{N}^*$,

$$W(k) := \underbrace{W + \cdots + W}_{k \text{ fois}}.$$

Notons que $W(k)$ contient kW. Comme W est un voisinage de 0 (donc un ensemble absorbant...) on en déduit que pour tout $u \in X$, on peut trouver $k \in \mathbb{N}^*$ tel que $u \in W(k)$.
En utilisant les "petites choses utiles", on peut construire par récurrence une suite de voisinages de 0 équilibrés notée $\big(W(2^{-n})\big)_{n\geqslant 1}$ telle que, pour tout $n \in \mathbb{N}^*$, on ait

$$W(2^{-n}) + W(2^{-n}) \subseteq W(2^{-(n-1)}) \qquad \text{et} \qquad W(2^{-n}) \subseteq B\big(0, 1/n\big).$$

Notons $\mathbf{D}$ l'ensemble des *nombres réels dyadiques*, *i.e.* l'ensemble des nombres réels positifs de la forme

$$s = k + \sum_{i=1}^{n} \alpha_i \, 2^{-i},$$

où $k \in \mathbb{N}$ et $\alpha_i = 0$ ou 1. Pour $s = k + \sum_{i=1}^n \alpha_i \, 2^{-i} \in \mathbf{D}$, on pose

$$W(s) := W(k) + \alpha_1 W(2^{-1}) + \cdots + \alpha_n W(2^{-n}).$$

Par le choix des $W(2^{-n})$ et la définition des $W(k)$, on voit (exo) que

$$\forall s, s' \in \mathbf{D} \ : \ W(s) + W(s') \subseteq W(s+s').$$

On en déduit que si $s, t \in \mathbf{D}$ et $s \leqslant t$, alors $W(s) \subseteq W(t)$: en effet, on peut écrire $t = s + s'$ avec $s' \in \mathbf{D}$, donc $W(t) \supseteq W(s) + W(s') \supseteq W(s)$ puisque $0 \in W(s')$.

FAIT. Soit $\phi : G \to \mathbb{R}^+$ la fonction définie par

$$\phi(u) := \inf \big\{ s \in \mathbf{D};\ u \in W(s) \big\}.$$

Alors ϕ possède les propriétés suivantes.

(i) $\phi(u+u') \leqslant \phi(u) + \phi(u')$ pour tous $u, u' \in X$.

(ii) $\phi(0) = 0$ et $\phi(u) > 0$ si $u \neq 0$.

(iii) $\phi(-u) = \phi(u)$ pour tout $u \in X$.

(iv) Pour tout $n \in \mathbb{N}^*$ et pour tout $u \in X$, on a les implications

$$\phi(u) < 2^{-n} \implies u \in W\big(2^{-n}\big) \implies \phi(u) \leqslant 2^{-n}.$$

Preuve du Fait. La fonction ϕ est bien définie car on a observé plus haut que pour tout $u \in X$, on peut trouver $k \in \mathbb{N}^*$ tel que $u \in W(k)$.
(i) Soient $u, u' \in X$. Pour tout $\varepsilon > 0$, on peut trouver $s, s' \in \mathbf{D}$ tels que : $s < \phi(u) + \varepsilon$ et $u \in W(s)$; $s' < \phi(u') + \varepsilon$ et $u' \in W(s')$. Alors $u + u' \in W(s) + W(s') \subseteq W(s+s')$, et donc $\phi(u+u') \leqslant s + s' \leqslant \phi(u) + \phi(u') + 2\varepsilon$. Ceci étant vrai pour tout $\varepsilon > 0$, on a donc $\phi(u+u') \leqslant \phi(u) + \phi(u')$.

(ii) Il est clair que $\phi(0) = 0$. Si $u \neq 0$, on peut trouver $n \in \mathbb{N}^*$ tel que $\rho(u, 0) \geqslant 1/n$. Alors $u \notin W(2^{-n})$ puisque $W(2^{-n}) \subseteq B(0, 1/n)$; donc $u \notin W(s)$ pour $s \leqslant 2^{-n}$ puisque $W(s) \subseteq W(2^{-n})$, et donc $\phi(u) \geqslant 2^{-n} > 0$.

(iii) On a $\phi(u) = \phi(-u)$ car tous les ensembles $W(s)$ sont équilibrés, donc symétriques ($W(s) = -W(s)$).

(iv) Si $u \in W(2^{-n})$, alors $\phi(u) \leqslant 2^{-n}$ par définition de ϕ. Si $\phi(u) < 2^{-n}$, alors on peut trouver $s \in \mathbf{D}$ tel que $s < 2^{-n}$ et $u \in W(s)$; donc $u \in W(2^{-n})$ puisque $W(s) \subseteq W(2^{-n})$. □

On peut maintenant terminer la preuve de la proposition. Pour $x, y \in X$, posons

$$d(x, y) := \phi(x - y).$$

En utilisant les propriétés (i), (i), (iii) du Fait, on montre sans difficulté que d est une distance sur X (exo). De plus, il est clair par définition que d est invariante par translations : si $x, y, a \in X$, alors

$$d(x + a, y + a) = \phi\big((x + a) - (y + a)\big) = \phi(x - y) = d(x, y).$$

Pour montrer que d définit la topologie de X, il suffit de montrer qu'une suite $(x_k) \subseteq X$ converge vers 0 pour d si et seulement si elle converge vers 0 pour la distance initiale ρ. Supposons que $d(x_k, 0) \to 0$. Pour $\varepsilon > 0$ donné, on peut trouver $n \in \mathbb{N}^*$ tel que $1/n < \varepsilon$, puis K tel $\forall k \geqslant K\ :\ d(x_k, 0) < 2^{-n}$. Alors $\phi(x_k) < 2^{-n}$ pour tout $k \geqslant K$, donc $x_k \in W(2^{-n})$ par (iv), et donc $\rho(x_k, 0) < 1/n < \varepsilon$ puisque $W(2^{-n}) \subseteq B(0, 1/n)$; donc $\rho(x_k, 0) \to 0$. Inversement, supposons que $\rho(x_k, 0) \to 0$, *i.e.* $x_k \to 0$ pour la topologie de X. Soit $\varepsilon > 0$, et soit $n \in \mathbb{N}$ tel que $2^{-n} < \varepsilon$. Comme $W(2^{-n})$ est un voisinage de 0, on peut trouver K tel que $\forall k \geqslant K\ :\ x_k \in W(2^{-n})$. Alors $\phi(x_k) \leqslant 2^{-n}$ pour tout $k \geqslant K$, *i.e.* $d(x_k, 0) \leqslant 2^{-n} < \varepsilon$; donc $d(x_k, 0) \to 0$. □

Remarque. En fait, la proposition est valable pour tout *groupe topologique commutatif* $(G, +, \tau)$; mais la preuve donnée ne “passe” pas telle quelle : on n'est pas certain que pour tout $x \in G$, on puisse trouver $k \in \mathbb{N}^*$ tel que $x \in W(k)$, et donc la fonction ϕ n'est peut-être pas bien définie.

Remarque 1.3. Si X est un evt dont la topologie est définie par une *suite* de semi-normes $(\|\cdot\|)_{i \in \mathbb{N}}$, alors X est métrisable et on définit une distance compatible avec la topologie de X et invariante par translations en posant

$$d(x, y) := \sum_{i=0}^{\infty} 2^{-i} \min\left(1, \|y - x\|_i\right).$$

Démonstration. Exo. □

Remarque 1.4. Soit $(G, +)$ un groupe commutatif. Si d est une distance invariante par translations sur G, alors

$$\forall x, y, x', y' \in G\ :\ d(x + y, x' + y') \leqslant d(x, x') + d(y, y').$$

En particulier, on a $d(x + y, 0) \leqslant d(x, 0) + d(y, 0)$ pour tous $x, y \in G$.

Démonstration. On a $d(x+y, x'+y') = d(x-x', y'-y) \leqslant d(x-x', 0) + d(0, y'-y) = d(x, x') + d(y, y')$. □

Proposition 1.5. *Si* $(G,+,\tau)$ *est un groupe topologique commutatif complètement métrisable, alors* G *est complet pour* toute *distance* d *compatible avec sa topologie et invariante par translations.*

Démonstration. Fixons une distance d compatible avec la topologie de G et invariante par translations. Dans la suite, on notera $(\widehat{G},\widehat{d}\,)$ le *complété* de l'espace métrique (G,d) ; autrement dit, $(\widehat{G},\widehat{d}\,)$ est un espace métrique complet qui contient G avec $\widehat{d}_{G\times G}=d$, et G est dense dans $\widehat{G}$. On va montrer que $G=\widehat{G}$, ce qui donnera le résultat.

Fait. On peut munir $(\widehat{G},\widehat{d}\,)$ d'une structure de groupe topologique (commutatif) qui prolonge celle de G.

Démonstration. Munissons $G\times G$ de la distance produit associée à d, que l'on notera ρ :

$$\rho\big((x,x'),(y,y')\big)=\max\big(d(x,y),d(x',y')\big).$$

On vérifie sans difficulté (exo) que le complété de $(G\times G,\rho)$ est $(\widehat{G}\times\widehat{G},\widehat{\rho})$, où $\widehat{\rho}$ est la distance produit associée à $\widehat{d}$. De plus, l'application $(x,y)\mapsto x+y$ est *uniformément continue* de $(G\times G,\rho)$ dans $(G,d)\subseteq(\widehat{G},\widehat{d}\,)$. En effet, si $(x,x'),(y,y')\in X\times X$, alors, par la Remarque 1.4, on a

$$d(x+y,x'+y')\leqslant d(x,x')+d(y,y')\leqslant 2\rho\big((x,x'),(y,y')\big);$$

donc l'application $(x,y)\mapsto x+y$ est en fait 2-lipschitzienne. Comme $(\widehat{G},\widehat{d}\,)$ est complet et que $G\times G$ est dense dans $\widehat{G}\times\widehat{G}$, on en déduit que l'application $(x,y)\mapsto x+y$ se prolonge de manière unique en une application continue de $\widehat{G}\times\widehat{G}$ dans $\widehat{G}$, qu'on notera encore $(x,y)\mapsto x+y$. De même, l'application $x\mapsto -x$ se prolonge continûment à $\widehat{G}$. On vérifie alors que $\widehat{G}$ muni de la loi $+$ et de la topologie définie par $\widehat{d}$ est un groupe topologique (exo) □

On sait maintenant que $\widehat{G}$ est un groupe topologique complètement métrisable, et que $G\subseteq\widehat{G}$ est complètement métrisable pour la topologie induite (qui est définie par d). D'après le Théorème d'Alexandrov (Remarque 4.3 du Chapitre 2), G est un G_δ de $\widehat{G}$; et bien sûr G est toujours dense dans $\widehat{G}$. On en déduit que si a est un point quelconque de $\widehat{G}$, alors $a+G\subseteq\widehat{G}$ est également un G_δ dense de $\widehat{G}$, car l'application $x\mapsto a+x$ est un homéomorphisme de $\widehat{G}$ sur $\widehat{G}$. D'après le *Théorème de Baire*, on a donc $G\cap(a+G)\neq\varnothing$. Ainsi $a\in G-G$, et donc $a\in G$ puisque G est un sous-groupe de $\widehat{G}$. Ceci étant vrai pour tout $a\in\widehat{G}$, on a donc bien $\widehat{G}=G$. □

Corollaire 1.6. *Soit* X *un evt métrisable. Les choses suivantes sont équivalentes.*

(1) X *est complètement métrisable.*

(2) *Toute suite* $(x_n)_{n\in\mathbb{N}}\subseteq X$ *telle que* $x_q-x_p\to 0$ *quand* $p,q\to\infty$ *converge dans* X.

Démonstration. Soit d une distance invariante par translations définissant la topologie de X. Comme $d(x,y)=d(x-y,0)$ pour tous $x,y\in X$, on voit que les suites $(x_n)\subseteq X$ vérifiant l'hypothèse de (2) sont exactement les suites de Cauchy pour la distance d. Donc (2) est vérifiée si et seulement si X est complet pour d ; ce qui est équivalent à (1) par la Proposition. □

Corollaire 1.7. *Soit X un evt complètement métrisable, et soit d une distance invariante par translations définissant la topologie de X. Soit aussi $(x_k)_{k\in\mathbb{N}}$ une suite de points de X. Si la série $\sum d(x_k,0)$ est convergente, alors la série $\sum x_k$ converge dans X, i.e. les sommes partielles $S_n := \sum_{k=0}^n x_k$ convergent dans X.*

Démonstration. Comme (X,d) est complet, il suffit de montrer que la suite (S_n) est de Cauchy dans (X,d) ; ce qui découle de l'invariance par translations : si $p<q$, alors

$$\begin{aligned} d(S_q,S_p) = d(S_q-S_p,0) &= d(x_{p+1}+\cdots+x_q,0)\\ &\leqslant d(x_{p+1},0)+\cdots+d(x_q,0) \xrightarrow{p,q\to\infty} 0. \end{aligned}$$

□

1.4. Espaces localement convexes, espaces de Fréchet.

Définition 1.8. *Soit X un evt.*

(1) *On dit que X est **localement convexe** si tout point $a\in X$ possède une base de voisinage formée d'ensembles ouverts convexes.*
(2) *On dit que X est un **espace de Fréchet** si X est localement convexe et complètement métrisable.*

Remarque. Tout sous-ev d'un espace localement convexe est localement convexe. Tout sous-ev fermé d'un espace de Fréchet est un espace de Fréchet.

Démonstration. Exo. □

Exemple 1. Tout evt X dont la topologie est définie par une famille de semi-normes $(\|\cdot\|_i)_{i\in I}$ est localement convexe.

Démonstration. Si $a\in X$, toutes les "boules" $B_{i_1,\dots,i_r}(a,\varepsilon)$ sont des ensembles convexes. □

Exemple 2. Si Ω est un ouvert de $\mathbb{R}^d$, alors $\mathcal{C}(\Omega)$ est un espaces de Fréchet.

Démonstration. On a vu au Chapitre 2 que $\mathcal{C}(\Omega)$ est complètement métrisable. De plus, $\mathcal{C}(\Omega)$ est localement convexe puisque sa topologie est engendrée par une famille de semi-normes. □

Exemple 3. Si Ω est un ouvert de $\mathbb{R}^d$, alors $\mathcal{C}^\infty(\Omega)$ est un espace de Fréchet.

Démonstration. L'espace $\mathcal{C}^\infty(\Omega)$ est localement convexe puisque sa topologie est engendrée par une famille de semi-normes.
Soit $(K_i)_{i\in\mathbb{N}}$ une suite exhaustive de compacts pour Ω, *i.e.* une suite croissante de compacts de Ω telle que tout compact $K\subseteq\Omega$ est contenu dans l'un des K_i. Pour $f\in\mathcal{C}^\infty(\Omega)$, posons

$$\|f\|_i := \sup\big\{|\partial^\alpha f(x)|;\, x\in K_i\,,\, |\alpha|\leqslant i\big\};$$

autrement dit, avec les notations du Chapitre 2, $\|\cdot\|_i=\|\cdot\|_{K_i,i}$. Si K est un compact quelconque de Ω et si $n\in\mathbb{N}$, alors on peut trouver i tel que $K\subseteq K_i$ et $n\leqslant i$, et donc tel que $\|\cdot\|_{K,n}\leqslant\|\cdot\|_i$. On voit ainsi que la convergence pour toutes les semi-normes $\|\cdot\|_i$ suffit à assurer la convergence pour toutes les semi-normes $\|\cdot\|_{K,n}$. Par conséquent, la topologie de $\mathcal{C}^\infty(\Omega)$ est engendrée par les $\|\cdot\|_i$, $i\in\mathbb{N}$; et par la Remarque 1.3, cela montre que $\mathcal{C}^\infty(\Omega)$ est métrisable.

Soit (f_n) une suite dans $\mathcal{C}^\infty(\Omega)$ telle que $f_q - f_p \to 0$ quand $p, q \to \infty$. Alors, pour tout $\alpha \in \mathbb{N}^d$, la suite $(\partial^\alpha f_n)$ vérifie le critère de Cauchy uniforme sur tout compact $K \subseteq \Omega$, donc converge uniformément sur tout compact. Par un théorème bien connu de calcul différentiel, on en déduit qu'il existe une fonction $f \in \mathcal{C}^\infty(\Omega)$ telle que $\partial^\alpha f_n \to \partial^\alpha f$ uniformément sur tout compact. Autrement dit, (f_n) converge dans $\mathcal{C}^\infty(\Omega)$. On a ainsi montré que $\mathcal{C}^\infty(\Omega)$ est complètement métrisable. □

Exemple 4. Soit Ω un ouvert de $\mathbb{R}^d$. Si K est un compact de Ω, alors l'espace $\mathcal{D}_K(\Omega) := \{\varphi \in \mathcal{C}^\infty(\Omega);\ \mathrm{supp}(\varphi) \subseteq K\}$ est un espace de Fréchet quand on le munit de la topologie induite par $\mathcal{C}^\infty(\Omega)$.

Démonstration. Il suffit d'observer que $\mathcal{D}_K(\Omega)$ est un sous-espace fermé de $\mathcal{C}^\infty(\Omega)$ (micro-exo). □

Exemple 5. Si Ω est un ouvert de $\mathbb{C}$, alors $H(\Omega)$, l'espace des fonctions holomorphes sur Ω, est un espace de Fréchet quand on le munit de la topologie de la convergence uniforme sur tout compact.

Démonstration. Daprès le *Théorème de convergence de Weierstrass*, $H(\Omega)$ est un sous-espace fermé de $\mathcal{C}(\Omega)$. □

1.5. Applications linéaires continues.

Lemme 1.9. *Soient X et Y des evt, et soit $T : X \to Y$ une application linéaire. Les choses suivantes sont équivalentes.*

(i) *T est continue ;*

(ii) *T est continue en 0 ;*

(ii') *Pour tout voisinage V de 0 dans Y, il existe un voisinage W de 0 dans X tel que $T(W) \subseteq V$.*

Démonstration. Il est clair que (i) $\Longrightarrow$ (ii) et que (ii) $\Longleftrightarrow$ (ii'). Inversement, supposons T continue en 0. Soit $a \in X$ quelconque, et soit (x_p) une suite généralisée convergeant vers a. Alors $x_p - a \to 0$, donc $T(x_p - a) \to 0$ par continuité de T en 0, et donc $T(x_p) \to T(a)$ puisque $T(x_p - a) = T(x_p) - T(a)$. □

Corollaire 1.10. *Soit X un evt et soit Y un* espace vectoriel normé. *Une application linéaire $T : X \to Y$ est continue si et seulement si elle est* bornée sur un voisinage de 0, *i.e. il existe un voisinage W de 0 dans X et une constante C tels que $\forall x \in W\ :\ \|Tx\| \leqslant C$.*

Démonstration. L'implication (T continue) $\Longrightarrow$ (T bornée sur un voisinage de 0) est laissée en exo. Inversement, supposons qu'il existe un voisinage W de 0 et une constante C tels que $\forall x \in W\ :\ \|Tx\| \leqslant C$. Si $\varepsilon > 0$ est donné, alors $W_\varepsilon := \frac{\varepsilon}{C} W$ est un voisinage de 0 tel que $\|T(x)\| \leqslant \varepsilon$ pour tout $x \in W_\varepsilon$. Donc T est continue en 0, et donc continue. □

Corollaire 1.11. *Soit X un evt, et soit $\Phi : X \to \mathbb{K}$ une* forme linéaire. *Alors Φ est continue si et seulement* $\ker(\Phi)$ *est fermé dans X.*

Démonstration. Si Φ est continue, alors $\ker(\Phi) = \Phi^{-1}(\{0\})$ est évidemment fermé. Inversement, supposons que Φ ne soit pas continue (donc en particulier $\Phi \neq 0$), et montrons que $\ker(\Phi)$ n'est pas fermé. On va en fait montrer que $\ker(\Phi)$ est dense dans X, ce qui donnera le résultat puisque $\ker(\Phi) \neq X$

Fait. Pour tout voisinage $W >$ de 0 dans X, on a $\Phi(W) = \mathbb{K}$.

Preuve du Fait. On peut supposer que W est équilibré. Comme Φ n'est pas continue, elle n'est pas bornée sur W, *i.e.* $\Phi(W)$ est une partie non bornée de $\mathbb{K}$. Mais $\Phi(W)$ est aussi un ensemble équilibré par linéarité de Φ (micro-exo) ; donc $\Phi(W) = \mathbb{K}$ (autre exo). $\square$

Montrons maintenant que $\ker(\Phi)$ est dense dans X. Soit $x \in X$ quelconque. Par le Fait, pour tout voisinage W de 0, on peut trouver $w \in W$ tel que $\Phi(w) = -\Phi(x)$. Alors $z := x + w \in \ker(\Phi)$ par linéarité de Φ ; donc $(x + W) \cap \ker(\Phi) \neq \varnothing$ pour tout voisinage W de 0, et donc $x \in \overline{\ker(\Phi)}$. $\square$

Exercice. Soit X un evt, et soit p une semi-norme sur X. Montrer que p est continue si et seulement si il existe un voisinage W de 0 tel que p est bornée sur W.

Proposition 1.12. *Soit X un evt dont la topologie est définie par une famille de semi-normes $(\|\cdot\|)_{i\in I}$, soit Y un evn, et soit $T : X \to Y$ une application linéaire. Les choses suivantes sont équivalentes.*

(1) *T est continue.*

(2) *Il existe $i_1, \dots, i_r \in I$ et une constante C tels que*

$$\forall x \in X \ : \ \|T(x)\| \leqslant C \max\big(\|x\|_{i_1}, \dots, \|x\|_{i_r}\big).$$

Démonstration. Supposons T continue, et montrons que (2) est vérifiée. Soit $W := \{y \in Y;\ \|y\| \leqslant 1\}$. C'est un voisinage de 0 dans Y, donc on peut trouver un voisinage W' de 0 dans X tel que $T(W') \subseteq W$. Soient $i_1, \dots, i_r \in I$ et $\varepsilon > 0$ tels que $\overline{B}_{i_1,\dots,i_r}(0, \varepsilon) \subseteq W'$. On a ainsi l'implication suivante pour tout $u \in X$:

$$\max\big(\|u\|_{i_1}, \dots, \|u\|_{i_r}\big) \leqslant \varepsilon \implies \|T(u)\| \leqslant 1.$$

Posons alors $C := 1/\varepsilon$ et montrons que $\|T(x)\| \leqslant C \max\big(\|x\|_{i_1}, \dots, \|x\|_{i_r}\big)$ pour tout $x \in X$. Posons $p(x) := \max\big(\|x\|_{i_1}, \dots, \|x\|_{i_r}\big)$. Si $p(x) = 0$, alors $p(\lambda x) = \lambda p(x) = 0$ pour tout $\lambda > 0$, donc $\|T(\lambda x)\| \leqslant 1$, *i.e.* $\|T(x)\| \leqslant 1/\lambda$; donc $\|T(x)\| = 0$ en faisant $\lambda \to \infty$, et donc l'inégalité souhaitée est certainement vraie. Si $p(x) \neq 0$, on peut poser $u := \varepsilon \frac{x}{p(x)}$, et on obtient l'inégalité souhaitée "par homogénéité" (exo).

Inversement, si (2) est vérifiée, alors T est certainement continue en 0 (micro-exo), donc continue. $\square$

Remarque. Si la famille de semi-normes $(\|\cdot\|)_{i\in I}$ est **filtrante**, *i.e.* si pour tous $i_1, \dots, i_r \in I$ on peut trouver $i \in I$ tel que $\|\cdot\|_i \geqslant \|\cdot\|_{i_k}$ pour $k = 1, \dots, r$, il revient au même de dire qu'il existe $i \in \mathbb{N}$ et une constante C tels que $\forall x \in X \ : \ \|T(x)\| \leqslant C\,\|x\|_i$.

Démonstration. Micro-exo. $\square$

Exemple 1. Soit un ouvert de $\mathbb{R}^d$ et soit $T : \mathcal{D}(\Omega) \to \mathbb{C}$ une forme linéaire. Alors T est une distribution sur Ω si et seulement si la restriction de T à chaque sous-espace $\mathcal{D}_K(\Omega)$ est continue.

Démonstration. Dire que T est une distribution signifie, par définition, que pour tout compact $K \subseteq \Omega$, il existe une constante C et un entier n tels que $\forall \varphi \in \mathcal{D}_K(\Omega)$: $|T(\varphi)| \leqslant C \sup\{|\partial^\alpha \varphi(x)|;\ |\alpha| \leqslant n\,,\ x \in \Omega\}$. Mais si $\varphi \in \mathcal{D}_K(\Omega)$, alors le "sup" précédent est égal au "sup sur K", autrement dit à $\|\varphi\|_{K,n}$. Donc le résultat est clair par le Corollaire 1.12 (exo). $\square$

EXEMPLE 2. Si $T : \mathcal{C}^\infty(\Omega) \to \mathbb{C}$ est une forme linéaire continue sur $\mathcal{C}^\infty(\Omega)$, alors sa restriction à $\mathcal{D}(\Omega)$ est une distribution *à support compact*. Inversement, toute distribution à support compact se prolonge de manière unique en une forme linéaire continue sur $\mathcal{C}^\infty(\Omega)$.

Démonstration. Par l'Exemple 1, $T_{|\mathcal{D}(\Omega)}$ est une distribution. Comme T est continue sur $\mathcal{C}^\infty(\Omega)$, on peut trouver une constante C, un compact $K \subseteq \Omega$ et un entier n tels que $\forall f \in \mathcal{C}^\infty(\Omega)\ :\ |T(f)| \leqslant C \sup\big\{|\partial^\alpha f(x)|;\ |\alpha| \leqslant n\,,\, x \in K\big\}$. En particulier, si $\varphi \in \mathcal{D}(\Omega)$ est identiquement nulle au voisinage de K, alors $T(\varphi) = 0$ puisque $\partial^\alpha\varphi(x) \equiv 0$ sur K pour tout $\alpha \in \mathbb{N}^d$; donc le support de T en tant que distribution est contenu dans le compact K.
Inversement, soit $S : \mathcal{D}(\Omega) \to \mathbb{C}$ une distribution à support compact, et notons $K \subseteq \Omega$ le support de S. Choisissons une fonction $\chi \in \mathcal{D}(\Omega)$ telle que $\chi(x) \equiv 1$ au voisinage de K. Alors $\langle S, \varphi\rangle = \langle S, \chi\varphi\rangle$ pour toute $\varphi \in \mathcal{D}(\Omega)$ par le choix de χ. Notons L le support de χ, qui est donc un compact de Ω. Si f est une fonction quelconque de $\mathcal{C}^\infty(\Omega)$, alors la fonction χf appartient à $\mathcal{D}_L(\Omega)$. De plus, on vérifie sans trop de peine que l'application $f \mapsto \chi f$ est continue de $\mathcal{C}^\infty(\Omega)$ dans $\mathcal{D}_L(\Omega)$: il faut utiliser le fait que si $\alpha \in \mathbb{N}^d$, alors $\partial^\alpha(\chi f)$ est combinaison linéaire de fonctions de la forme $\partial^\beta\chi\,\partial^\gamma f$ avec des coefficients indépendants de χ et f (exo). Comme S est une distribution, on en déduit que la formule $T(f) := \langle S, \chi f\rangle$ définit une forme linéaire continue sur $\mathcal{C}^\infty(\Omega)$, qui prolonge S. Le prolongement est unique car $\mathcal{D}(\Omega)$ est dense dans $\mathcal{C}^\infty(\Omega)$. □

NOTATION. Si X et Y sont des evt, on note $\mathcal{L}(X, Y)$ l'ensemble des applications linéaires continues $T : X \to Y$, et on pose $\mathcal{L}(X) := \mathcal{L}(X, X)$. On rappelle que si X et Y sont des evn, alors $\mathcal{L}(X, Y)$ est toujours muni de la norme définie par

$$\|T\| := \sup\left\{\frac{\|T(x)\|}{\|x\|};\ x \neq 0\right\} = \sup\left\{\|T(x)\|;\ \|x\| = 1\right\} = \sup\big\{\|T(x)\|;\ \|x\| \leqslant 1\big\}.$$

1.6. Quotients. On sait bien qu'il est parfois utile de quotienter... Les lemmes suivants résument ce qu'il faut savoir sur les evt quotients.

RAPPEL. Soit X un espace topologique, soit Z un ensemble abstrait, et soit $\pi : X \to Z$. On munit Z d'une topologie en décrétant qu'un ensemble $O \subseteq Z$ est ouvert si et seulement si $\pi^{-1}(O)$ est un ouvert de X. Cette topologie s'appelle la **topologie quotient** associée à π. Par définition, l'application $\pi : X \to Z$ est continue. Si Y est un espace topologique, alors une application $f : Z \to Y$ est continue si et seulement si $f \circ \pi : X \to Y$ est continue.

Démonstration. Exo. □

EXEMPLE. La topologie du cercle $\mathbb{T} := \{z \in \mathbb{C};\ |z| = 1\}$ est la topologie quotient définie par l'application $\mathbb{R} \ni t \mapsto e^{it}$.

Démonstration. Exo. □

LEMME 1.13. *Soit X un evt. Si $E \subseteq X$ un sous-espace vectoriel* fermé *de X, on note X/E l'espace vectoriel quotient, et $\pi : X \to X/E$ la surjection canonique.*

(1) *Muni de la topologie quotient, E/X est un evt.*
(2) *La surjection canonique $\pi : X \to X/E$ est une application (linéaire continue et)* ouverte *: pour tout ouvert $V \subseteq X$, l'ensemble $\pi(V)$ est ouvert dans X/E.*

(3) *Si Y est un autre evt et si $T : X \to Y$ est une application linéaire continue, alors il existe une unique application linéaire $\widetilde{T} : X/\ker(T) \to Y$ telle que $\widetilde{T} \circ \pi = T$; et cette application linéaire $\widetilde{T}$ est injective et continue.*

Démonstration. (2) Soit V un ouvert de X. Par définition de π, on a $\pi^{-1}(\pi(V)) = E + V$ (micro-exo) ; autrement dit $\pi^{-1}(\pi(V)) = \bigcup_{e \in E}(e + V)$. Donc $\pi^{-1}(\pi(V))$ est un ouvert de X (réunion d'ouverts), et donc $\pi(V)$ est ouvert dans X/E.

(1) La continuité des opérations est laissée en exo (un peu fastidieux, et pas très palpitant ; on a besoin de (1)). Montrons que X/E est *séparé*. Soient $z, z' \in E/X$ tels que $z \neq z'$, et choisissons $x, x' \in X$ tels que $\pi(x) = z$ et $\pi(x') = z'$. Alors $x - x' \notin E$ puisque $\pi(x) \neq \pi(x')$. Comme E est fermé dans X, on peut donc trouver un voisinage V de 0 dans X tel que $\big(x - x' + V - V\big) \cap E = \varnothing$. Alors $W := \pi(V)$ est un voisinage ouvert de 0 dans X/E par (1), et $0 \notin z - z' + W - W$ par linéarité de π. Donc $U := z + W$ et $U' := z' + W$ sont des voisinages de u et u' tels que $U \cap U' = \varnothing$, ce qui prouve que X/E est séparé.

(3) est évident. □

Lemme 1.14. *Soit X un* espace vectoriel normé, *soit $E \subseteq X$ un sous-espace vectoriel fermé et soit $\pi : X \to X/E$ la surjection canonique.*

(1) *La topologie de X/E est engendrée par la **norme quotient**, qui est définie comme suit : si $z \in X/E$, alors*
$$\|z\|_{X/E} := \inf \big\{\|x\|;\ \pi(x) = z\big\}.$$
Autrement dit :
$$\forall x \in X\ :\ \|\pi(x)\|_{X/E} = \inf \big\{\|x + u\|;\ u \in E\big\} = \mathrm{dist}(x, E).$$

(2) *On a $\pi\big(B_X(a, r)\big) = B_{X/E}\big(\pi(a), r\big)$ pour tout $a \in X$ et pour tout $r > 0$. En particulier, $\|\pi\| = 1$.*

(3) *Si X est un espace de Banach, alors X/E est un espace de Banach.*

(4) *Si Y est un autre evn et si $T \in \mathcal{L}(X, Y)$, alors l'application linéaire $\widetilde{T} : X/\ker(T) \to Y$ associée vérifie $\|\widetilde{T}\| = \|T\|$.*

Démonstration. (1) Le fait que la norme quotient soit effectivement une norme sur X/E est laissé en exo (il est important que E soit fermé) ; montrons qu'elle définit la topologie quotient. Par définition de la norme quotient, on a $\|\pi(x)\|_{X/E} \leqslant \|x\|$ pour tout $x \in X$, donc l'application linéaire $\pi : X \to X/E$ est continue de X dans $\big(X/E, \|\cdot\|_{X/E}\big)$ avec $\|\pi\| \leqslant 1$. Donc, tout ouvert pour $\|\cdot\|_{X/E}$ est ouvert pour la topologie quotient, par définition de la topologie quotient. Par ailleurs, la définition de la norme quotient montre que l'image de toute boule ouvert $B(a, r) \subseteq X$ est la boule ouverte $B(\pi(a), r)$ pour la norme quotient : en effet, dire que $z \in X/E$ appartient à $B(\pi(a), r)$, *i.e.* que $\|z - \pi(a)\|_{X/E} < r$, signifie qu'on peut trouver $u \in X$ tel que $\pi(u) = z - \pi(a)$ et $\|u\| < r$, autrement dit $\pi(a + u) = z$ et $\|u\| < r$, ce qui revient à dire que $z \in \pi\big(B(a, r)\big)$. Comme tout ouvert de X est réunion de boules ouvertes, on en déduit que $\pi : X \to \big(X/E, \|\cdot\|_{X/E}\big)$ est une application ouverte. Si maintenant $O \subseteq X$ est ouvert pour la topologie quotient, alors $O = \pi\big(\pi^{-1}(O)\big)$ car π est surjective, donc O est ouvert pour $\|\cdot\|_{X/E}$ puisque $\pi^{-1}(O)$ est un ouvert de X.

(2) La 1ère partie a été démontrée dans la preuve de (1). En particulier $\pi\big(B(0,1)\big) = B(0,1)$, ce qui entraine que $\|\pi\| = 1$ (exo).

(3) C'est un exo classique : on montre que si $(z_k)_{k\in\mathbb{N}}$ est une suite d'éléments de X/E vérifiant $\sum_{k=0}^{\infty} \|z_k\| < \infty$, alors la série $\sum z_k$ converge dans X/E.

(4) Si $z \in X/E$ alors, pour tout $\varepsilon > 0$, on peut trouver $x = x_\varepsilon \in X$ tel que $\|x\| \leqslant (1+\varepsilon)\|z\|$ et $\pi(x) = z$. Alors $\widetilde{T}z = Tx$, donc $\|\widetilde{T}z\| \leqslant \|T\|\,\|x\| \leqslant (1+\varepsilon)\|z\|$. C'est vrai pour tout $\varepsilon > 0$, donc $\|\widetilde{T}z\| \leqslant \|T\|\,\|z\|$. Ainsi, $\widetilde{T}$ est continue et $\|\widetilde{T}\| \leqslant \|T\|$. Mais comme $T = \widetilde{T}\pi$, on a aussi $\|T\| \leqslant \|\widetilde{T}\|\,\|\pi\| \leqslant \|\widetilde{T}\|$. □

Exemple. La preuve "la plus courte" du Théorème de Riesz.

Soit X un evn dont la boule unité fermée B_X est compacte. Alors la boule ouverte $B_X(0,1)$ est précompacte, donc on peut trouver $x_1, \dots, x_N \in X$ tels que $B_X(0,1) \subseteq \bigcup_{i=1}^N B_X(x_i, 1/2)$. Soit $E := \mathrm{Vect}(x_1, \dots, x_N)$, qui est un sous-espace vectoriel *fermé* de X car de dimension finie, et soit $\pi : X \to X/E$ la surjection canonique. Comme $x_i \in E$, on a $\pi(x_i) = 0$ et donc $\pi\big(B_X(x_i, 1/2)\big) = B_{X/E}\big(\pi(x_i), 1/2\big) = B_{X/E}(0, 1/2)$ pour $i = 1, \dots, N$. Donc $B_{X/E}(0,1) = \pi\big(B_X(0,1)\big) \subseteq \bigcup_{i=1}^N \pi\big(B_X(x_i, 1/2)\big) = B_{X/E}(0, 1/2)$. Cela impose $X/E = \{0\}$ (micro-exo) ; donc $X = E$ et donc $\dim(X) < \infty$.

2. Théorèmes de Hahn-Banach

2.1. Fonctionnelles sous-linéaires.

Définition 2.1. *Soit X un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une* **fonctionnelle sous-linéaire** *sur X est une application $p : X \to \mathbb{R}$ vérifiant les propriétés suivantes.*

(i) *$p(\lambda x) = \lambda p(x)$ pour tout $x \in X$ et tout $\lambda \geqslant 0$; en particulier, $p(0) = 0$.*

(ii) *$p(x+y) \leqslant p(x) + p(y)$ pour tous $x, y \in X$.*

Exemple 1. Une forme $\mathbb{R}$-linéaire $\varphi : X \to \mathbb{R}$ est une fonctionnelle sous-linéaire.

Exemple 2. Une *semi-norme* est une fonctionnelle sous-linéaire. En particulier, une norme est une fonctionnelle sous-linéaire.

Lemme 2.2. *Soit X un evt, et soit $C \subseteq X$ une partie convexe de X telle que $0 \in \mathring{C}$. On définit une fonctionnelle sous-linéaire $p_C : X \to \mathbb{R}^+$ en posant*

$$p_C(x) := \inf\left\{\alpha > 0;\ \frac{x}{\alpha} \in C\right\}.$$

On dit que p_C est la **fonctionnelle de Minkowski** *du convexe C. On a $p_C(x) \leqslant 1$ pour tout $x \in C$, $p_C(x) \geqslant 1$ pour tout $x \in X \backslash C$, et $p_C(x) > 1$ pour tout $x \in X \backslash \overline{C}$. Si de plus C est* équilibré*, alors p_C est une semi-norme continue. Enfin, si C est ouvert, alors $C = \{x \in X;\ p(x) < 1\}$.*

Démonstration. Si $x \in X$ est quelconque, alors l'ensemble $\{\alpha > 0;\ x/\alpha \in C\}$ est non vide car C est absorbant (c'est un voisinage de 0) ; donc $p_C(x)$ est *bien défini* pour tout $x \in C$.

Il est facile de voir que $p_C(\lambda x) = \lambda p_C(x)$ pour tout $x \in X$ et pour tout $\lambda > 0$. On a de plus $p_C(0) = 0$ car $\frac{0}{\alpha} \in C$ pour tout $\alpha > 0$. Donc $p_C(\lambda x) = \lambda p_C(x)$ pour tout $\lambda \geqslant 0$.

Soient $x, y \in X$, et soient $\alpha, \beta > 0$ tels que $\frac{x}{\alpha} \in C$ et $\frac{y}{\beta} \in C$. En écrivant

$$\frac{x+y}{\alpha+\beta} = \frac{\alpha}{\alpha+\beta}\,\frac{x}{\alpha} + \frac{\beta}{\alpha+\beta}\,\frac{y}{\beta} := \lambda\,\frac{x}{\alpha} + (1-\lambda)\,\frac{y}{\beta}\,,$$

on voit que $\frac{x+y}{\alpha+\beta}$ est combinaison convexe de deux éléments de C, et par conséquent $\frac{x+y}{\alpha+\beta} \in C$. Par définition de p_C, on en déduit

$$p_C(x+y) \leqslant \alpha + \beta,$$

pour tous $\alpha, \beta > 0$ tels que $\frac{x}{\alpha} \in C$ et $\frac{y}{\beta} \in C$. En passant indépendemment à la borne supérieure en α et en β, on obtient donc $p_C(x+y) \leqslant p_C(x) + p_C(y)$. Ainsi, p_C est une fonctionnelle sous-linéaire.

Si $x = \frac{x}{1} \in C$, alors certainement $p_C(x) \leqslant 1$. Si $p_C(x) < 1$, alors on peut trouver $\alpha < 1$ et $y \in C$ tels que $x = \alpha y$; et donc $x \in C$ par convexité puisque $0 \in C$. Si $x \notin \overline{C}$, alors on peut trouver $r < 1$ tel que $rx \notin \overline{C}$ car $X \backslash \overline{C}$ est un ouvert de X et $rx \to x$ quand $r \to 1^-$. Alors, comme $0 \in C$ et comme C est convexe, on a $sx \notin C$ pour tout $s \geqslant r$; autrement dit $x/\alpha \notin C$ pour tout $\alpha \leqslant 1/r$. Donc $p_C(x) \geqslant 1/r > 1$.

Si C est équilibré, alors $p_C(\omega x) = p_C(x)$ pour tout $x \in X$ et pour tout $\omega \in \mathbb{K}$ vérifiant $|\omega| = 1$ (micro-exo). Donc $p_C(\lambda x) = |\lambda|\, p_C(x)$ pour tout $x \in X$ et pour tout $\lambda \in \mathbb{K}$ puisqu'on peut écrire $\lambda x = \omega t x$ avec $|\omega| = 1$ et $t = |\lambda| \geqslant 0$; et donc p_C est une semi-norme. Comme la semi-norme p_C est bornée sur C (qui est un voisinage de 0), elle est continue (exo déjà posé).

On sait déjà que $\{x \in X;\ p_C(x) < 1\} \subseteq C$. Si maintenant C est ouvert et si $x \in C$, alors on peut trouver $r > 1$ tel que $rx \in C$ car $rx \to x$ quand $r \to 1^+$; donc $p_C(x) \leqslant 1/r < 1$. □

Remarque. Si X est un evn et si $C = B(0,1)$, alors $p_C = \|\cdot\|$.

COROLLAIRE 2.3. *Un evt X est localement convexe si et seulement si sa topologie est engendrée par une famille de semi-normes.*

Démonstration. Une implication a déjà été vue. Inversement, supposons que X soit localement convexe. Notons $\mathcal{W}$ la famille de tous les voisinages ouverts de 0 convexes et équilibrés. Par le Lemme 2.2, on sait que si $W \in \mathcal{W}$, alors p_W est une semi-norme continue sur X. On va montrer que la topologie de X est engendrée par la famille $(p_W)_{W \in \mathcal{W}}$. Pour cela, il suffit de montrer qu'une suite généralisée $(x_d)_{d \in D} \subseteq X$ converge vers 0 si et seulement si $p_W(x_d) \to 0$ pour tout $W \in \mathcal{W}$ (exo). Si $x_d \to 0$, alors $p_W(x_d) \to p_W(0) = 0$ pour tout $W \in \mathcal{W}$ car p_W est continue. Inversement, supposons que $p_W(x_d) \to 0$ pour tout $W \in \mathcal{W}$. Soit V un voisinage de 0 quelconque dans X. Comme X est localement convexe, on peut trouver $W \in \mathcal{W}$ tel que $W \subseteq V$ (micro-exo ; retourner voir les “petites choses utiles”) ; et comme $p_W(x_d) \to 0$, on peut trouver $d_0 \in D$ tel que $p_W(x_d) < 1$ pour tout $d \succeq d_0$. Alors $x_d \in W \subseteq V$ pour tout $d \succeq d_0$, ce qui prouve que $x_d \to 0$. □

2.2. Hahn-Banach “abstrait”.

THÉORÈME 2.4. *Soit X un espace vectoriel sur $\mathbb{R}$, et soit $p : X \to \mathbb{R}$ une fonctionnelle sous-linéaire. Soit également E un sous-espace vectoriel de X, et soit $\varphi : E \to \mathbb{R}$ une forme linéaire* majorée par p, *c'est-à-dire vérifiant $\varphi(x) \leqslant p(x)$ pour tout $x \in E$. Alors il existe une forme linéaire $\Phi : X \to \mathbb{R}$ définie sur X tout entier, qui prolonge φ et reste majorée par p.*

Démonstration. On dira qu'une forme linéaire $\Phi : Z \to \mathbb{R}$ définie sur un sous-espace vectoriel $Z \subseteq X$ tel que $Z \subseteq E$ est un *prolongement admissible* de φ si $\Phi_{|E} = \varphi$ et $\Phi(z) \leqslant p(z)$ pour tout $z \in Z$. Le sous-ev Z s'appelle le *domaine* de Φ et se note $\mathrm{dom}(\Phi)$. Par exemple, φ est un prolongement admissible de φ avec $\mathrm{dom}(\varphi) = E$; et ce qu'on cherche est un prolongement admissible Φ de φ tel que $\mathrm{dom}(\Phi) = X$.

Notons $\mathcal{E}$ l'ensemble des prolongement admissibles de φ. On définit une relation d'ordre $\leqslant$ sur $\mathcal{E}$ comme suit :

$$\Phi \leqslant \Phi' \iff \mathrm{dom}(\Phi) \subseteq \mathrm{dom}(\Phi') \quad \text{et} \quad \Phi' \text{ prolonge } \Phi.$$

Fait 1. $\mathcal{E}$ possède un élément maximal pour $\leqslant$.

Preuve du Fait 1. C'est une "application directe" du Lemme de Zorn, qu'on laisse en exo. □

Fait 2. Si $\Phi \in \mathcal{E}$ et $Z := \mathrm{dom}(\Phi) \neq X$, alors Φ n'est pas maximal pour $\leqslant$.

Preuve du Fait 2. Soit $e \in X \backslash Z$, et soit $Z' := Z \oplus \mathbb{R}e$. Soit également $c \in \mathbb{R}$, et soit $\Phi' : Z' \to \mathbb{R}$ la forme linéaire définie par

$$\Phi'(z + te) := \Phi(z) + ct \qquad \text{pour tout } z' = z + te \in Z'.$$

Par définition, Φ' prolonge Φ et $\mathrm{dom}(\Phi')$ est strictement plus grand que $\mathrm{dom}(\Phi)$. Donc il suffit de montrer qu'on peut choisir c de façon à avoir $\Phi' \in \mathcal{E}$, *i.e.*

$$\forall z \in Z\ \forall t \in \mathbb{R}\ :\ \Phi(z) + ct \leqslant p(z + te). \tag{2.1}$$

L'inégalité (2.1) est satisfaite pour $t = 0$ car $\Phi \leqslant p$ sur Z. Donc il suffit de montrer qu'il existe un nombre réel c tel que

$$\forall z \in Z\ \forall \lambda > 0\ :\ \Phi(z) \pm c\lambda \leqslant p(z \pm \lambda e).$$

En divisant par λ et en remplaçant z par z/λ, cela revient à montrer qu'il existe $c \in \mathbb{R}$ tel que

$$\forall z \in Z\ :\ \Phi(z) \pm c \leqslant p(z \pm e),$$

autrement dit

$$\forall z \in Z\ :\ \Phi(z) - p(z - e) \leqslant c \leqslant p(z + e) - \Phi(z).$$

Il s'agit donc de voir si on a

$$\sup\big\{\Phi(u) - p(u - e);\ u \in Z\big\} \leqslant \inf\big\{p(v + e) - \Phi(v);\ v \in Z\big\};$$

ou encore : $\forall u, v \in Z\ :\ \Phi(u) - p(u - e) \leqslant p(v + e) - \Phi(v)$. C'est effectivement le cas car si $u, v \in Z$, alors

$$\Phi(u) + \Phi(v) = \Phi(u + v) \leqslant p(u + v) \leqslant p(u - e) + p(v + e).$$

□

La preuve du théorème est maintenant terminée : Par les Faits 1 et 2, il existe un prolongement admissible Φ de φ tel que $\mathrm{dom}(\Phi) = X$. □

2.3. Prolongement des formes linéaires continues.

THÉORÈME 2.5. *Soit X un evt sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit également E un sous-espace vectoriel de X, et soit $\varphi : E \to \mathbb{K}$ une forme linéaire continue. Alors il existe une forme linéaire continue $\Phi : X \to \mathbb{K}$ telle que $\Phi_{|E} = \varphi$. De plus, si X est un evn, on peut imposer que $\|\Phi\| = \|\varphi\|$; et on dit alors que Φ est un* ***prolongement de Hahn-Banach*** *de φ.*

Démonstration. Commençons par le cas plus familier où X est un evn. On traite séparément le cas réel et le cas complexe.

CAS 1. $\mathbb{K} = \mathbb{R}$.

Posons $C := \|\Phi\|$. En appliquant le Théorème 2.4 avec $p(x) := C\,\|x\|$, on obtient une forme linéaire $\Phi : X \to \mathbb{R}$ qui prolonge φ et vérifie $\Phi(x) \leqslant C\,\|x\|$ pour tout $x \in X$. On a alors aussi $-\Phi(x) = \Phi(-x) \leqslant C\,\|-x\| = C\,\|x\|$, d'où $|\Phi(x)| \leqslant C\,\|x\|$ pour tout $x \in X$. Donc Φ est continue et $\|\Phi\| \leqslant C = \|\varphi\|$. Mais comme Φ prolonge φ, on a aussi $\|\Phi\| \geqslant \|\varphi\|$, d'où $\|\Phi\| = \|\varphi\|$. ▫

CAS 2. $\mathbb{K} = \mathbb{C}$.

On utilise le fait suivant.

FAIT 2.6. *Soit Z un espace vectoriel sur $\mathbb{C}$. Si $\theta : Z \to \mathbb{R}$ est une forme $\mathbb{R}$-linéaire, il existe une unique forme $\mathbb{C}$-linéaire $\theta_{\mathbb{C}} : Z \to \mathbb{C}$ telle que $\mathrm{Re}(\theta_{\mathbb{C}}) = \theta$; elle est donnée par la formule*

$$\theta_{\mathbb{C}}(x) := \theta(x) - i\theta(ix).$$

Démonstration. Exo d'algèbre. □

Posons $\varphi_{\mathbb{R}} = \mathrm{Re}(\varphi)$. Alors $\varphi_{\mathbb{R}} : E \to \mathbb{R}$ est une forme $\mathbb{R}$-linéaire continue, et on a $\|\varphi_{\mathbb{R}}\| \leqslant \|\varphi\|$ car $|\varphi_{\mathbb{R}}(x)| \leqslant |\varphi(x)|$ pour tout $x \in E$. D'après le cas 1, il existe donc une forme $\mathbb{R}$-linéaire continue $\Phi_{\mathbb{R}} : X \to \mathbb{R}$ qui prolonge $\varphi_{\mathbb{R}}$ et vérifie $\|\Phi_{\mathbb{R}}\| \leqslant \|\varphi\|$. Soit $\Phi : X \to \mathbb{C}$ l'unique forme $\mathbb{C}$-linéaire telle que $\mathrm{Re}(\Phi) = \Phi_{\mathbb{R}}$. On a $\mathrm{Re}\big(\Phi_{|E}\big) = (\Phi_{\mathbb{R}})_{|E} = \varphi_{\mathbb{R}} = \mathrm{Re}(\varphi)$, donc $\Phi_{|E} = \varphi$.
Il reste à voir que Φ est continue et $\|\Phi\| = \|\varphi\|$. Si $x \in X$, alors il existe un nombre complexe ω tel que $|\omega| = 1$ et $|\Phi(x)| = \omega\Phi(x)$. On a alors $|\Phi(x)| = \Phi(\omega x) = \mathrm{Re}(\Phi(\omega x))$ puisque $|\Phi(x)| \in \mathbb{R}$; autrement dit $|\Phi(x)| = \Phi_{\mathbb{R}}(\omega x)$. On en déduit $|\Phi(x)| \leqslant \|\Phi_{\mathbb{R}}\|\,\|\omega x\| = \|\varphi\|\,\|x\|$ pour tout $x \in X$, donc Φ est continue $\|\Phi\| \leqslant \|\varphi\|$. L'inégalité inverse est évidente puisque Φ prolonge φ.
Supposons maintenant que X soit un evt localement convexe quelconque. Par ce qui précède, on voit qu'on peut supposer que X est réel. Comme la forme linéaire $\varphi : E \to \mathbb{R}$ est continue, on peut trouver un voisinage ouvert W de 0 dans E tel que $|\varphi(w)| < 1$ pour tout $w \in W$; puis un voisinage ouvert $\widetilde{W}$ de 0 dans X tel que $W = \widetilde{W} \cap E$. De plus, comme X est localement convexe, on peut suposer (quitte à diminuer W) que $\widetilde{W}$ est convexe et équilibré. Soit alors $p := p_{\widetilde{W}}$, la fonctionnelle de Minkowski de $\widetilde{W}$. Par le Lemme 2.2, on sait que p est une semi-norme (continue) sur X et que $\widetilde{W} = \{x \in X;\ p(x) < 1\}$. De plus on a $|\varphi(x)| \leqslant p(x)$ pour tout $x \in E$. En effet, si on avait $|\varphi(x)| > p(x)$, on pourrait trouver $c \in \mathbb{R}$ tel que $p(x) < c < |\varphi(x)|$; alors $w := x/c$ vérifierait d'une part $p(w) < 1$, donc $w \in \widetilde{W} \cap E = W$, et d'autre part $|\varphi(w)| > 1$, ce qui est absurde. Par le Théorème 2.4, on peut donc trouver une forme

linéaire $\Phi : X \to \mathbb{R}$ prolongeant φ telle que $\Phi(x) \leqslant p(x)$ pour tout $x \in X$; et comme p est une semi-norme, on a en fait $|\Phi(x)| \leqslant p(x)$ pour tout $x \in X$ (exo). En particulier $|\Phi(x)| \leqslant 1$ pour tout $x \in \widetilde{W}$; donc Φ est bornée sur un voisinage de 0, et donc Φ est continue. □

Notation. Si X est un evt sur $\mathbb{K}$, on note X^* l'ensemble des formes linéaires continues $\Phi : X \to \mathbb{K}$; autrement dit, $X^* = \mathcal{L}(X, \mathbb{K})$. Les éléments de X^* se noteront souvent avec des étoiles : $x^*, y^*, z^*, \dots$. Si $x^* \in X^*$ et $x \in X$, on écrit $\langle x^*, x\rangle$ au lieu de $x^*(x)$.

Corollaire 2.7. *Soit X un evt localement convexe. Pour tout $x \in X\backslash\{0\}$, on peut trouver $x^* \in X^*$ telle que $\langle x^*, x\rangle \neq 0$. Si de plus X est un evn, alors on peut trouver $x^* \in X^*$ telle que $\|x^*\| = 1$ et $\langle x^*, x\rangle = \|x\|$.*

Démonstration. Fixons $x_0 \in X\backslash\{0\}$, et posons $E := \mathbb{K}x_0$. Soit également $c > 0$, et soit $\varphi : \mathbb{K}x_0 \to \mathbb{K}$ l'unique forme linéaire telle que $\varphi(x_0) = c$. On a $\varphi(\lambda x_0) = c\,\lambda$ pour tout $\lambda x_0 \in \mathbb{K}x_0$, donc $\ker(\varphi) = \{0\}$. En particulier $\ker(\varphi)$ est fermé dans $\mathbb{K}x_0$, et donc φ est continue d'après un exo déjà posé. Si X est un evn, on prend $c := \|x_0\|$: alors $|\varphi(\lambda x_0)| = \big|\lambda\,\|x_0\|\big| = \|\lambda x_0\|$ pour tout $\lambda x_0 \in E$; donc $\|\varphi\| = 1$. D'où le résultat en prolongeant φ par Hahn-Banach. □

Corollaire 2.8. *Si X est un evt localement convexe, alors X^** sépare les points *de X : si $u, v \in X$ et $u \neq v$, on peut trouver $x^* \in X^*$ telle que $\langle x^*, u\rangle \neq \langle x^*, v\rangle$.*

Démonstration. Micro-exo. □

Corollaire 2.9. *Soit X un evt localement convexe, et soit E un sous-espace vectoriel de X. Posons*

$$E^\perp := \big\{\Phi \in X^*;\ \Phi(u) = 0 \textit{ pour tout } u \in E\big\}.$$

Pour tout $a \in X$ tel que $a \notin \overline{E}$, on peut trouver $\Phi \in X^*$ telle que $\Phi \in E^\perp$ et $\Phi(a) \neq 0$. En particulier, E est dense *dans X si et seulement si $E^\perp = \{0\}$. (**Critère dual de densité**.)*

Démonstration. Comme $\overline{E}$ est un sous-espace fermé de X, on peut considérer l'evt quotient $X/\overline{E}$. Soit $\pi : X \to X \to \overline{E}$ la surjection canonique. Comme $a \notin \overline{E}$, on a $\pi(a) \neq 0$. De plus, comme X est localement convexe, on vérifie sans peine que $X/\overline{E}$ est localement convexe (exo). Donc on peut trouver $\Psi \in \big(X/\overline{E}\big)^*$ telle que $\Psi(\pi(a)) \neq 0$. Alors $\Phi := \Psi \circ \pi \in X^*$ est telle que $\Phi \in E^\perp$ et $\Phi(a) \neq 0$. □

Preuve sans quotient. Soit $F := \mathbb{K}a \oplus \overline{E}$, et soit $\varphi : F \to \mathbb{K}$ l'unique forme linéaire telle que $\varphi(a) = 1$ et $\varphi \equiv 0$ sur $\overline{E}$. On a $\ker(\varphi) = \overline{E}$ (micro-exo) ; donc φ est *continue.* Comme X est localement convexe, on peut prolonger φ en une forme linéaire continue $\Phi \in X^*$; qui convient. □

Exemple 1. Soit X un evn. Si $E \subseteq X$ est un sous-espace de dimension finie, alors E admet un *supplémentaire fermé* : il existe un sous-espace fermé $F \subseteq X$ tel que $X = E \oplus F$.

Démonstration. Soit $(e_1, \dots, e_d)$ une base de E, et soient $\varphi_1, \dots, \varphi_d = E \to \mathbb{K}$ les formes linéaires définies par les relations $\varphi_i(e_j) := \delta_{i,j}$. Comme $\dim(E) < \infty$, les φ_i sont continues ; donc on peut les prolonger en $\Phi_1, \dots, \Phi_d \in X^*$. Si on définit $\pi : X \to E$ par

$$\pi(x) := \sum_{i=1}^{d} \Phi_i(x)\, e_i,$$

alors π est linéaire continue, et π est une *projection* de X sur E, *i.e.* $\pi(x) = x$ pour tout $x \in E$ (exo). On a donc $X = E \oplus \ker(\pi)$; et donc $F := \ker(\pi)$ convient. □

Exemple 2. Soit $\Omega := \{\lambda \in \mathbb{C};\ |\lambda| > 1\}$. Pour tout $\lambda \in \Omega$, soit $f_\lambda : [-1,1] \to \mathbb{C}$ la fonction continue définie par $f_\lambda(t) := \frac{1}{t-\lambda}\cdot$ Alors $E := \mathrm{Vect}\,\{f_\lambda;\ \lambda \in \Omega\}$ est dense dans $\mathcal{C}([-1,1])$. Plus généralement, si $\Lambda \subseteq \Omega$ est un ensemble ou bien non borné, ou bien possèdant un point d'accumulation dans Ω, alors $E_\Lambda := \mathrm{Vect}\,\{f_\lambda;\ \lambda \in \Lambda\}$ est dense dans $\mathcal{C}([-1,1])$.

Démonstration. Il suffit de montrer que $E_\Lambda^\perp = \{0\}$. Soit donc $\Phi \in \big(\mathcal{C}([-1;1])\big)^*$ vérifiant $\Phi(f_\lambda) = 0$ pour tout $\lambda \in \Lambda$: on veut montrer que $\Phi = 0$.
Si $\lambda \in \Omega$, on peut écrire

$$f_\lambda(t) = \frac{1}{\lambda}\,\frac{1}{1 - \frac{t}{\lambda}} = \sum_{k=0}^{\infty} \frac{t^k}{\lambda^{k+1}}\cdot$$

où la série converge normalement sur $[-1,1]$. Donc, en notant $\mathbf{t}^k$ la fonction $[-1,1] \ni t \mapsto t^k$, on a

$$f_\lambda = \sum_{k=0}^{\infty} \frac{\mathbf{t}^k}{\lambda^{k+1}},$$

où la série converge dans $\mathcal{C}([-1,1])$. Comme Φ est une forme linéaire continue sur $\mathcal{C}([-1,1])$, on en déduit

$$\Phi(f_\lambda) = \sum_{k=0}^{\infty} c_k \times \left(\frac{1}{\lambda}\right)^{k+1} = F\left(\frac{1}{\lambda}\right),$$

où $c_k := \Phi(\mathbf{t}^k)$ et $F(z) := \sum_0^\infty c_k z^{k+1}$. Comme la suite (c_k) est bornée ($|c_k| \leqslant \|\Phi\|\,\|\mathbf{t}^k\|_\infty = \|\Phi\|$), la fonction F est holomorphe dans le disque unité $\mathbb{D} = \{|z| < 1\}$. Par hypothèse sur Φ, on a $F\left(\frac{1}{\lambda}\right) = 0$ pour tout $\lambda \in \Lambda$. De plus, l'ensemble $\{1/\lambda;\ \lambda \in \Lambda\}$ possède un point d'accumulation dans $\mathbb{D}$ par hypothèse sur Λ. D'après le principe des zéros isolés, on a donc $F = 0$, d'où $c_k = 0$ pour tout $k \in \mathbb{N}$. Comme $c_k = \Phi(\mathbf{t}^k)$, on a donc $\Phi(P) = 0$ pour toute fonction polynomiale P par linéarité, d'où $\Phi = 0$ car les fonctions polynomiales sont denses dans $\mathcal{C}([-1,1])$. □

Exemple 3. (fonctions holomorphes à valeurs vectorielles)
Soit X un espace de Banach complexe. Si Ω est un ouvert de $\mathbb{C}$, une fonction $F = \Omega \to X$ est dite *holomorphe* si elle est $\mathbb{C}$-dérivable en tout point de Ω.

(1) La formule de Cauchy est valable pour les fonctions holomorphes à valeurs vectorielles : si $F : \Omega \to X$ est holomorphe et si par exemple D est un disque ouvert tel que $\overline{D} \subseteq \Omega$, alors

$$\forall z \in D\ :\ F(z) = \frac{1}{2i\pi}\int_{\partial D} \frac{F(\xi)}{\xi - z}\, d\xi.$$

(2) De même, le Théorème de Liouville est valable : si $F : \mathbb{C} \to X$ est une fonction holomorphe et bornée, alors F est constante.

Démonstration. Le point clé et que si $F : \Omega \to X$ est holomorphe alors, pour toute forme linéaire continue $x^* \in X^*$, la fonction $x^* \circ F : \Omega \to \mathbb{C}$ est holomorphe, avec $(x^* \circ F)'(z) = \langle x^*, F'(z)\rangle$. Cela permet d'obtenir "par scalarisation" des résultats sur les fonctions holomorphes à valeurs vectorielles en utilisant les résultats correspondants sur les fonctions à valeurs complexes.

(1) Soit $z \in D = D(z_0, r)$. L'intégrale curviligne vectorielle $\int_{\partial D} \frac{F(\xi)}{\xi - z}\, d\xi$ a un sens car la fonction $\xi \mapsto \frac{F(\xi)}{\xi - z}\, d\xi$ est continue sur ∂D à valeurs dans l'espace de Banach X. Par la formule de Cauchy scalaire, on a

$$\forall x^* \in X^* \ : \ \left\langle x^*, \int_{\partial D} \frac{F(\xi)}{\xi - z}\, d\xi \right\rangle = \int_{\partial D} \frac{\langle x^*, F(\xi)\rangle}{\xi - z}\, d\xi = \langle x^*, F(z)\rangle;$$

donc , $\int_{\partial D} \frac{F(\xi)}{\xi - z}\, d\xi = F(z)$ d'après le Théorème de Hahn-Banach.

(2) Pour toute $x^* \in X^*$, la fonction $z \mapsto \langle x^*, F(z)\rangle$ est holomorphe sur $\mathbb{C}$ et bornée (exo) ; donc cette fonction est constante par le Théorème de Liouville classique. Ainsi, on a

$$\forall x^* \in X^* \ : \ \langle x^*, F(z)\rangle \equiv \langle x^*, F(0)\rangle,$$

et donc $F(z) \equiv F(0)$ par Hahn-Banach. □

2.4. Séparation des ensembles convexes.

Rappel. Si X est un evt, les éléments de X^* se noteront $x^*, y^*, z^*, \dots$ Si $x^* \in X^*$ et $x \in X$, on écrit $\langle x^*, x\rangle$ au lieu de $x^*(x)$.

Définition 2.10. *Soit X un evt réel, soient $A, B \subseteq X$ vérifiant $A \cap B = \varnothing$, et soit $x^* \in X^* \backslash \{0\}$.*

(i) *On dit que x^* **sépare strictement** A **de** B si on a*

$$\inf \big\{\langle x^*, u\rangle;\ u \in A\big\} > \sup \big\{\langle x^*, v\rangle;\ v \in B\big\};$$

autrement dit, s'il existe $\alpha, \beta \in \mathbb{R}$ tels que

$$\langle x^*, u\rangle \geqslant \beta > \alpha \geqslant \langle x^*, v\rangle \qquad \textit{pour tous } u \in A,\ v \in B.$$

(ii) *On dit que x^* **sépare** A **de** B **au sens large** si on a*

$$\inf \big\{\langle x^*, u\rangle;\ u \in A\big\} \geqslant \sup \big\{\langle x^*, v;\ v \in B\big\}.$$

Remarque. Le fait que x^* sépare strictement A et B s'interpréte géométriquement de façon très simple : si on choisit $\lambda \in \mathbb{R}$ tel que $\sup_B x^* < \lambda < \inf_A x^*$, alors A et B sont situés "de part et d'autre" de l'hyperplan fermé $H_\lambda := \big\{x \in X;\ \langle x^*, x\rangle = \lambda\big\}$.

Théorème 2.11. *Soit X un evt localement convexe réel, et soient A, B deux parties* convexes *de X telles que $A \cap B = \varnothing$.*

(1) *Si A est* compact *et si B est* fermé, *alors il existe une forme linéaire continue $x^* \in X^* \backslash \{0\}$ qui sépare strictement A de B.*

(2) *Si B est* ouvert, *alors il existe $x^* \in X^* \backslash \{0\}$ qui sépare A de B au sens large.*

Démonstration. (1) On distingue deux cas.

Cas 1. A est réduit à 1 point $\{a\}$.

Dans ce cas, on cherche $x^* \in X^*$ telle que $\langle x^*, a\rangle > \sup\big\{\langle x^*, x\rangle;\ x \in B\big\}$. Par translation, on peut supposer que $0 \in B$.

Fait. On peut trouver un ouvert convexe $\Omega \subseteq X$ tel que $B \subseteq \Omega$ et $a \notin \overline{\Omega}$.

Preuve du Fait. Comme $a \notin B$ et que B est fermé, $X\backslash B$ est un voisinage ouvert de a. Par continuité de l'application $(x, y) \mapsto x - y$ et comme X est localement convexe, on peut donc trouver un voisinage ouvert convexe V de 0 tel que $a + V - V \subseteq X\backslash B$, *i.e.* $(a + V - V) \cap B = \varnothing$. (Si X est un evn, il suffit de prendre pour V une boule $B(0, \varepsilon)$ avec $\varepsilon > 0$ assez petit.) Alors $\Omega := B + V$ convient : il est ouvert et convexe (exo) ; on a $B \subseteq \Omega$ car $0 \in V$; et $a \notin \overline{\Omega}$ car $(a + V) \cap \Omega = \varnothing$ et $a + V$ est un voisinage de a. $\square$

Soit p_Ω la fonctionnelle de Minkowski de Ω,
$$p_\Omega(x) = \inf\left\{\alpha > 0;\ \frac{x}{\alpha} \in \Omega\right\}.$$
On sait que p_Ω est une fonctionnelle sous-linéaire. On a $p_\Omega(x) \leqslant 1$ pour tout $x \in \Omega$, et $p_\Omega(a) > 1$ car $a \notin \overline{\Omega}$. Soit $\varphi : \mathbb{R}a \to \mathbb{R}$ l'unique forme linéaire telle que $\varphi(a) = p_\Omega(a)$. Si $x = \lambda a \in \mathbb{R}a$, on a $\varphi(x) = \lambda p_\Omega(a)$; donc $\varphi(x) \leqslant 0 \leqslant p_\Omega(x)$ si $\lambda \leqslant 0$, et $\varphi(x) = p_\Omega(\lambda a) = p_\Omega(x)$ si $\lambda \geqslant 0$. La forme linéaire φ est donc majorée par p_Ω. D'après le théorème de Hahn-Banach, on peut trouver une forme linéaire $\Phi : X \to \mathbb{R}$ prolongeant φ et majorée par p_Ω. On a alors $\Phi(a) = p_\Omega(a) > 1$, et $\Phi(x) \leqslant p_\Omega(x) \leqslant 1$ si $x \in B \subseteq \Omega$. Enfin, $|\Phi(x)| = \Phi(\pm x) \leqslant 1$ pour tout $x \in \Omega \cap (-\Omega)$. Comme $W := \Omega \cap (-\Omega)$ est un voisinage de 0, on en déduit que Φ est continue. Ainsi, $x^* := \Phi$ convient.

Cas 2. Cas général.

On applique le premier cas à $\widetilde{A} := \{0\}$ et
$$\widetilde{B} := B - A = \big\{v - u;\ (u, v) \in A \times B\big\}.$$
L'ensemble $\widetilde{B}$ est convexe car A et B sont convexes. Comme A est compact et B fermé, on vérifie (exo) que $\widetilde{B}$ est également *fermé* dans X. Enfin, on a $0 \notin \widetilde{B}$ puisque $A \cap B = \varnothing$. Par le Cas 1, on peut trouver $x^* \in X^*$ telle que
$$0 = \langle x^*, 0\rangle > \sup\big\{\langle x^*, v - u\rangle;\ (u, v) \in A \times B\big\}.$$
Visiblement, cette x^* convient.

(2) Par la preuve de (1), on voit qu'il suffit de traiter le cas où A est réduit à un point $\{a\}$, et qu'on peut de plus supposer que $0 \in B$. Comme plus haut, on peut trouver une forme linéaire continue $\Phi : X \to \mathbb{R}$ telle que $\Phi(a) = p_B(a) \geqslant 1$ et $\Phi(x) \leqslant p_B(x)$ pour tout $x \in X$. Alors $\Phi \neq 0$, et $\Phi(a) \geqslant \Phi(v)$ pour tout $v \in B$ puisque $p_B(v) \leqslant 1$; donc $x^* := \Phi$ convient. $\square$

Remarque. Pour (2), il n'est en fait pas nécessaire de supposer que X est localement convexe : la preuve fonctionne dans un evt X quelconque.

Corollaire 2.12. *Soit X un evt localement convexe réel.*

1. *Si $C \subseteq X$ est un ensemble convexe* fermé *et si $a \in X$ est tel que $a \notin C$, alors on peut trouver $x^* \in X^*$ telle que $\langle x^*, a\rangle > \sup\big\{\langle x^*, v\rangle;\ v \in C\big\}$.*
2. *Si $\Omega \subseteq X$ est un* ouvert *convexe et si $a \in X$ est tel que $a \notin \Omega$, alors on peut trouver $x^* \in X^*\backslash\{0\}$ telle que $\langle x^*, a\rangle \geqslant \sup\big\{\langle x^*, z\rangle;\ z \in \Omega\big\}$.*

Démonstration. On applique le théorème avec $A := \{a\}$ et $B := C$ ou Ω. $\square$

Corollaire 2.13. *Soit X un* espace vectoriel normé *réel, et soient $A, B \subseteq X$ deux ensembles convexes. Si A et B sont* à distance strictement positive, *autrement dit s'il existe $\varepsilon > 0$ tel que $\forall (u,v) \in A \times B\ :\ \|u - v\| \geqslant \varepsilon$, alors il existe $x^* \in X^*$ qui sépare strictement A de B.*

Démonstration. Par hypothèse, l'ensemble $B - A$ ne rencontre pas la boule ouverte $B(0,\varepsilon)$; donc $0 \notin \Omega := (B - A) + B(0,\varepsilon)$. Comme Ω est un ouvert de X (exo), on peut trouver $x^* \in X^* \backslash \{0\}$ telle que $0 = \langle x^*, 0\rangle \geqslant \langle x^*, z\rangle$ pour tout $z \in \Omega$. Comme $x^* \neq 0$, on peut trouver $x \in B(0,\varepsilon)$ tel que $\alpha := \langle x^*, x\rangle > 0$. Alors $\langle x^*, v\rangle - \langle x^*, u\rangle + \alpha = \langle x^*, v - u + x\rangle \leqslant 0$ pour tout $u \in A$ et pour tout $v \in B$, puisque $z := v - u + x \in \Omega$. Donc $\inf\{\langle x^*, u\rangle;\ u \in A\} \geqslant \sup\{\langle x^*, v\rangle;\ v \in B\} + \alpha$, et donc x^* sépare strictement A de B. $\square$

Corollaire 2.14. *Soit X un evt localement convexe réel, et soit $E \subseteq X$. Notons* $\mathrm{conv}(E)$ *l'*enveloppe convexe *de E, i.e. le plus petit convexe $C \subseteq X$ contenant E. Pour $a \in X$, on a l'équivalence suivante :*

$$a \in \overline{\mathrm{conv}(E)} \iff \forall x^* \in X^*\ :\ \langle x^*, a\rangle \leqslant \sup\big\{\langle x^*, v\rangle;\ v \in E\big\}.$$

Démonstration. L'implication $\Rightarrow$ est laissée en exo ; et $\Leftarrow$ est une conséquence de 2.12 : prendre $C := \overline{\mathrm{conv}(E)}$. $\square$

Corollaire 2.15. *Soit X un evt localement convexe réel. Si $C \subseteq X$ est un ensemble convexe fermé, alors C est* intersection de demi-espaces fermés.

Démonstration. Par le Corollaire 2.12, pour tout point $a \in X\backslash C$, on peut trouver un demi-espace fermé H tel que $C \subseteq H$ et $a \notin H$. Donc, C est l'intersection de tous les demi-espaces fermés qui le contiennent. $\square$

Corollaire 2.16. *Soit X un evt localement convexe* complexe, *et soit $C \subseteq X$ un ensemble convexe fermé. Si $a \in X$ et $a \notin C$, alors on peut trouver $x^* \in X^*$ telle que* $\mathrm{Re}\,\langle x^*, a\rangle > \sup\big\{\mathrm{Re}\,\langle x^*, v\rangle;\ v \in C\big\}$.

Démonstration. D'après le Corollaire 2.12, on peut trouver une forme $\mathbb{R}$-linéaire continue $\varphi : X \to \mathbb{R}$ telle que $\varphi(a) > \sup\big\{\varphi(v);\ v \in C\big\}$; et on a vu qu'il existe une unique $x^* \in X^*$ telle que $\mathrm{Re}\, x^* = \varphi$. $\square$

Exemple 1. Convexes et intégrales.

Lemme 2.17. *Soit $(\Omega, \mathfrak{B}, \mathbb{P})$ un espace de probabilité, et soit $f : \Omega \to \mathbb{R}^d$ une fonction intégrable. Soit également $C \subseteq \mathbb{R}^d$ un ensemble convexe fermé. Si on a $f(t) \in C$ pour presque tout $t \in \Omega$, alors $\int_\Omega f\, d\mathbb{P} \in C$.*

Démonstration. Par le Corollaire 2.15, il suffit de prouver le résultat dans le cas où C est un demi-espace fermé, *i.e.* $C = \{x \in \mathbb{R}^d;\ \langle x^*, x\rangle \leqslant \alpha\}$ où $x^* \in (\mathbb{R}^d)^*$ et $\alpha \in \mathbb{R}$. Et dans ce cas, c'est évident :

$$\left\langle x^*, \int_\Omega f\, d\mathbb{P} \right\rangle = \int_\Omega \langle x^*, f(t)\rangle\, d\mathbb{P}(t) \leqslant \int_\Omega \alpha\, d\mathbb{P}(t) = \alpha.$$

$\square$

Exemple 2. Fonctions convexes et fonctions affines.

Proposition 2.18. *Soit E un evt localement convexe réel, et soit $f : E \to \mathbb{R}$ une fonction* convexe *continue. Pour tout point $x_0 \in E$, il existe une fonction* affine *continue $\varphi : E \to \mathbb{R}$ telle que $\varphi(x_0) = f(x_0)$ et $\forall x \in E\ :\ \varphi(x) \leqslant f(x)$.*

Démonstration. Fixons $x_0 \in E$, et posons

$$\Omega := \big\{(x,\lambda) \in E \times \mathbb{R};\ \lambda > f(x)\big\}.$$

Comme f est convexe continue, Ω est un ouvert convexe de $Z := E \times \mathbb{R}$. De plus, le point $a := \big(x_0, f(x_0)\big)$ n'appartient pas à Ω. D'après 2.12, on peut donc trouver une forme linéaire continue non nulle $\Phi : E \times \mathbb{R} \to \mathbb{R}$ telle que $\Phi(a) \geqslant \Phi(z)$ pour tout $z \in \Omega$. La forme linéaire Φ est donnée par

$$\Phi(x,\lambda) = \langle x^*, x\rangle + c\lambda\,,$$

où $x^* \in E^*$ et c est une constante. Par définition de Φ, on a ainsi

$$\langle x^*, x_0\rangle + cf(x_0) \geqslant \langle x^*, x\rangle + c\lambda \tag{2.2}$$

pour tout $x \in E$ et pour tout $\lambda > f(x)$. Si on avait $c = 0$, on en déduirait $\langle x^*, x_0\rangle \geqslant \langle x^*, x\rangle$ pour tout $x \in E$, d'où $x^* = 0$, ce qui est exclu puisque $\Phi \neq 0$. Si on avait $c > 0$, alors (2.2) conduirait à une contradiction en faisant tendre λ vers $+\infty$. On a donc $c < 0$; et quitte à diviser par $-c$, on peut supposer $c = -1$. Ainsi, on a $\langle x^*, x_0\rangle - f(x_0) \geqslant \langle x^*, x\rangle - \lambda$ pour tout $x \in E$ et pour tout $\lambda > f(x)$. En fixant x et en faisant tendre λ vers $f(x)$, on en déduit

$$\langle x^*, x_0\rangle - f(x_0) \geqslant \langle x^*, x\rangle - f(x)$$

pour tout $x \in E$. Par conséquent, il suffit de poser $\varphi(x) := f(x_0) + \langle x^*, x - x_0\rangle$. □

Remarque. Si on note G_f le graphe de f et H le graphe de la fonction affine φ, alors H est un hyperplan affine de $E \times \mathbb{R}$, qui "touche" G_f au point $\big(x_0, f(x_0)\big)$, et qui est situé "en dessous" de G_f. On dit que H est un **hyperplan d'appui** à G_f au point $\big(x_0, f(x_0)\big)$.

Exemple 3. Enveloppe convexe des matrices orthogonales.

Proposition 2.19. *Toute matrice $A \in M_d(\mathbb{R}) = \mathcal{L}(\mathbb{R}^d)$ telle que $\|A\| \leqslant 1$ est combinaison convexe de matrices orthogonales.*

Démonstration. Fixons $A \in M_d(\mathbb{R})$ avec $\|A\| \leqslant 1$. On veut montrer que $A \in \mathrm{conv}\big(O_d(\mathbb{R})\big)$. On sait que $O_d(\mathbb{R})$ est un compact de $M_d(\mathbb{R})$. *Donc*, $\mathrm{conv}\big(O_d(\mathbb{R})\big)$ est également un compact de $M_d(\mathbb{R})$: c'est une conséquence du *Théorème de Carathéodory.* En particulier $\mathrm{conv}\big(O_d(\mathbb{R})\big) = \overline{\mathrm{conv}\big(O_d(\mathbb{R})\big)}$. D'après le Théorème de séparation des convexes, il suffit donc de montrer que pour toute forme linéaire $\Phi : M_d(\mathbb{R}) \to \mathbb{R}$, on a

$$\Phi(A) \leqslant \sup\big\{\Phi(V);\ V \in O_d(\mathbb{R})\big\}. \tag{2.3}$$

On sait (ou pas) qu'on peut trouver une matrice $B \in M_d(\mathbb{R})$ telle que

$$\forall M \in M_d(\mathbb{R})\ :\ \Phi(M) = \mathrm{trace}(MB).$$

D'après le Théorème de *décomposition polaire*, on peut trouver $U \in O_d(\mathbb{R})$ et $P \geqslant 0$ telles que $B = UP$. Comme la matrice P est $\geqslant 0$, elle est diagonalisable en base orthonormée et ses valeurs propres sont $\geqslant 0$; on peut donc trouver une b.o.n. $(f_1, \dots, f_d)$ de $\mathbb{R}^d$ et $\lambda_1, \dots, \lambda_d \geqslant 0$ tels que $Pf_i = \lambda_i f_i$ pour $i = 1, \dots, d$. On a alors

$$\Phi(A) = \mathrm{trace}(AUP) = \sum_{i=1}^{d} \langle AUPf_i, f_i\rangle = \sum_{i=1}^{d} \lambda_i \langle AUf_i, f_i\rangle.$$

Comme $\|A\| \leqslant 1$, on a $\langle AUf_i, f_i\rangle \leqslant \|AUf_i\|\,\|f_i\| \leqslant \|Uf_i\|\,\|f_i\| = 1$ pour $i = 1, \dots, d$; donc, comme les λ_i sont $\geqslant 0$:

$$\Phi(A) \leqslant \sum_{i=1}^{d} \lambda_i = \mathrm{trace}(P).$$

Mais $\mathrm{trace}(P) = \mathrm{trace}(U^{-1}UP) = \Phi(U^{-1})$; donc (2.3) est bien vérifiée. $\square$

3. Théorème de Banach-Steinhaus

Définition 3.1. *Soit X un evt, et soit $B \subseteq X$. On dit que B est **borné** s'il est "absorbé par tout voisinage de 0" : pour tout voisinage W de 0 dans X, on peut trouver $\alpha \in \mathbb{R}^+$ tel que $B \subseteq \alpha W$.*

Exemple 1. Supposons que X soit localement convexe, et soit $(\|\cdot\|)_{i\in I}$ une famille de semi-normes engendrant la topologie de X. Alors un ensemble $B \subseteq X$ est borné si et seulement si toutes les semi-normes $\|\cdot\|_i$ sont bornées sur B, autrement dit : pour tout $i \in I$, on peut trouver une constante M_i telle que $\forall x \in B \,:\, \|x\|_i \leqslant M_i$. En particulier, si X est un *espace vectoriel normé*, alors $B \subseteq X$ est borné si et seulement si il est borné au sens usuel, *i.e.* il existe une constante M telle que $\forall x \in B \,:\, \|x\| \leqslant M$.

Démonstration. Supposons B borné. Pour tout $i \in I$, l'ensemble $B_i(0,1)$ est un voisinage de 0, donc on peut trouver $\alpha \in \mathbb{R}^+$ tel que $B \subseteq \alpha\, B_i(0,1) = B_i(0,\alpha)$. Alors $\|x\|_i \leqslant \alpha$ pour tout $x \in B$, donc la semi-norme $\|\cdot\|_i$ est bornée sur B. Inversement, supposons que toutes les semi-normes $\|\cdot\|_i$ soient bornées sur B, avec "témoins" M_i. Soit $W \subseteq X$ un voisinage de 0 quelconque. On peut trouver $i_1, \dots, i_r \in I$ et $\varepsilon > 0$ tels que $\overline{B}_{i_1,\dots,i_r}(0,\varepsilon) \subseteq W$. Posons $M := \max(M_{i_1}, \dots, M_{i_r})$. Si $x \in B$, alors $w := \frac{\varepsilon}{M}\, x$ vérifie $\|w\|_{i_k} = \frac{\varepsilon}{M}\, \|x\|_{i_k} \leqslant \varepsilon$ pour $k = 1, \dots, r$, *i.e.* $w \in \overline{B}_{i_1,\dots,i_r}(0,\varepsilon) \subseteq W$. Donc, si on pose $\alpha := \frac{M}{\varepsilon}$, on voit que $B \subseteq \alpha W$. $\square$

Exemple 2. Tout ensemble *fini* $B \subseteq X$ est borné.

Démonstration. Soit $B = \{x_1, \dots, x_N\} \subseteq X$ fini, et soit W un voisinage de 0. Pour $k = 1, \dots, N$, on peut trouver $r_k > 0$ tel que $rx_k \in W$ pour tout $0 \leqslant r \leqslant r_k$, car $rx_k \to 0$ quand $r \to 0^+$. Si on pose $r := \min(r_1, \dots, r_k)$, alors $rB \subseteq W$; donc $B \subseteq \alpha W$ avec $\alpha := 1/r$. $\square$

Exemple 3. Si $(x_n)_{n\in\mathbb{N}}$ est une *suite* d'éléments de X qui converge dans X, alors l'ensemble $B := \{x_n;\ n \in \mathbb{N}\}$ est borné.

Démonstration. Par translation, on peut supposer que $x_n \to 0$ (exo). Soit W un voisinage de 0 dans X, qu'on peut supposer équilibré. Comme $x_n \to 0$, on peut trouver $N \in \mathbb{N}$ tel que $\forall n > N \,:\, x_n \in W$; puis, comme $\{x_1, \dots, x_N\}$ est fini, on peut trouver $\gamma \in \mathbb{R}^+$ tel que $x_k \in \gamma W$ pour $k = 1, \dots, N$. Si on pose $\alpha := \max(\gamma, 1)$, alors $B \subseteq \alpha W$ car W est équilibré (micro-exo). $\square$

Exercice 1. Toute réunion fini d'ensembles bornés est un ensemble borné.

Exercice 2. Si $B \subseteq X$ est borné, alors $a + B$ est borné pour tout $a \in X$.

Exercice 3. Montrer qu'un ensemble $B \subseteq X$ est borné si et seulement si la chose suivante a lieu : pour toute suite $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{K}$ tendant vers 0 et pour toute suite $(x_n) \subseteq B$, on a que $\lambda_n x_n \to 0$.

THÉORÈME 3.2. *Soient X un evt* complètement métrisable*, Y un evt, et $(T_i)_{i\in I}$ une famille d'applications linéaires continues, $T_i : X \to Y$. On suppose que pour tout $x \in X$, l'ensemble $\{T_i(x);\ i \in I\}$ est borné dans Y. Alors les choses suivantes ont lieu.*

(1) *La famille $(T_i)_{i\in I}$ est **équicontinue** : pour tout voisinage W de 0 dans Y, il existe un voisinage V de 0 dans X tel que $\forall i \in I\ :\ T_i(V) \subseteq W$.*

(2) *Pour tout ensemble borné $B \subseteq X$, l'ensemble $\{T_i(x);\ i \in I, x \in B\}$ est borné dans Y.*

Démonstration. (1) Soit W un voisinage de 0 dans Y, et soit W' un voisinage de 0 équilibré tel que $\overline{W'} - \overline{W'} \subseteq W$ (micro-exo : ça existe). Si $x \in X$ alors, comme l'ensemble $B_x := \{T_i(x);\ i \in I\}$ est borné et comme W' est un voisinage de 0 équilibré, on peut trouver $n_x \in \mathbb{N}^*$ tel que $B_x \subseteq n_x W'$, et donc *a fortiori* $B_x \subseteq n_x\overline{W'}$. Donc, si on pose pour $n \in \mathbb{N}^*$:

$$F_n := \{x \in X;\ \forall i \in I\ :\ T_i(x) \in n\overline{W'}\},$$

alors $\bigcup_{n\in\mathbb{N}^*} F_n = X$. De plus, les F_n sont des fermés de X car les T_i sont continues. Comme X est complètement métrisable, on peut appliquer le *Théorème de Baire* : il existe un entier N tel que F_N est d'intérieur non-vide. On peut donc trouver $a \in X$ et un voisinage ouvert V' de 0 dans X tels que

$$\forall x \in a + V'\ \forall i \in I\ :\ T_i(x) \in N\overline{W'}.$$

Comme W' est équilibré et donc symétrique ($-W' = W'$), la même chose vaut pour $x \in -(a + V')$ par linéarité des T_i. Toujours par linéarité des T_i et comme $V' - V' = (a + V') - (a + V')$, on en déduit que

$$\forall x \in V' - V'\ \forall i \in I\ :\ T_i(x) \in N \cdot \left(\overline{W'} - \overline{W'}\right).$$

Donc, si on pose $V := \frac{1}{N} \cdot (V' - V')$, qui est bien un voisinage ouvert de 0 dans X (micro-exo), alors V convient : $T_i(V) \subseteq W$ pour tout $i \in I$.

(2) Soit $B \subseteq X$ un ensemble borné, et posons $\widetilde{B} := \{T_i(x);\ x \in B, i \in I\}$. Soit W un voisinage de 0 quelconque dans Y. Par (1), on peut trouver un voisinage V de 0 dans X tel que $T_i(V) \subseteq W$ pour tout $i \in I$. Comme B est borné, on peut ensuite trouver $\alpha \in \mathbb{R}^+$ tel que $B \subseteq \alpha V$. Alors $\widetilde{B} \subseteq \alpha W$ par le choix de V et par linéarité des T_i (micro-exo) ; donc $\widetilde{B}$ est borné. □

COROLLAIRE 3.3. *Soient X un espace de Banach, Y un espace vectoriel normé, et $(T_i)_{i\in I}$ une famille d'applications linéaires continues, $T_i : X \to Y$. Si la famille (T_i) est* simplement bornée*, alors elle est* bornée en norme*. Autrement dit : si $\forall x \in I\ :\ \sup_{i\in I} \|T_i x\| < \infty$, alors $\sup_{i\in I} \|T_i\| < \infty$.*

Démonstration. On applique (2) en prenant pour B la boule unité fermée de X. □

COROLLAIRE 3.4. *Soit* $(T_n)_{n\in\mathbb{N}}$ *une* suite *d'applications linéaires continues,* $T_n : X \to Y$*, où* X *est un evt complètement métrisable et* Y *est un evt. On suppose que la suite* (T_n) *est simplement convergente, autrement dit que* $T_n(x)$ *converge dans* Y *pour tout* $x \in X$*. Alors* $T := \lim_n T_n$ *est une application linéaire continue.*

Démonstration. Il est clair que T est linéaire. Soit W un voisinage de 0 dans Y : on cherche un voisinage V de 0 dans X tel que $T(V) \subseteq W$. Dans ce qui suit, on fixe un voisinage W' de 0 dans Y tel que $\overline{W'} \subseteq W$.
Pour tout $x \in X$, l'ensemble $\{T_n(x);\ n \in \mathbb{N}\}$ est borné dans Y puisque la suite $(T_n(x))$ est convergente. Par (1), on peut donc trouver un voisinage V de 0 dans X tel que $\forall n \in \mathbb{N}\ :\ T_n(V) \subseteq W'$. Comme $T(x) = \lim T_n(x) \in \overline{T(V)}$ pour tout $x \in V$, on en déduit $T(V) \subseteq \overline{W'} \subseteq W$. □

EXEMPLE 1. Soit Ω un ouvert de $\mathbb{R}^d$, et soit $(T_n)_{n\in\mathbb{N}}$ une suite de distributions sur Ω. On suppose que $\lim_{n\to\infty} T_n(\varphi)$ existe pour toute $\varphi \in \mathcal{D}(\Omega)$. Alors la forme linéaire $T := \lim T_n$ est une distribution.

Démonstration. Il faut montrer que pour tout compact $K \subseteq \Omega$, la restriction de T à $\mathcal{D}_K(\Omega)$ est continue ; et ceci est clair puisque la restriction de chaque T_n à $\mathcal{D}_K(\Omega)$ est continue et que l'espace $\mathcal{D}_K(\Omega)$ est complètement métrisable. □

EXEMPLE 2. Divergence des séries de Fourier.

Soit $\mathcal{C}_{2\pi}$ l'espace des fonctions $f : \mathbb{R} \to \mathbb{C}$ continues et 2π-périodiques. Si $f \in \mathcal{C}_{2\pi}$, on note $c_k(f)$, $k \in \mathbb{Z}$ ses coefficients de Fourier :

$$c_k(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\,dt.$$

Pour $n \in \mathbb{N}$, on pose

$$S_n f(t) := \sum_{k=-n}^{n} c_k(f)\,e^{ikt}.$$

PROPOSITION 3.5. *Il existe une fonction* $f \in \mathcal{C}_{2\pi}$ *telle que la suite* $\big(S_n f(0)\big)_{n\in\mathbb{N}}$ *n'est pas bornée. En particulier,* $S_n f(0) \not\to f(0)$ *quand* $n \to \infty$*.*

Démonstration. On rappelle que si $f \in \mathcal{C}_{2\pi}$, alors

$$\forall n \in \mathbb{N}\ \forall x \in \mathbb{R}\ :\ S_n f(x) = D_n * f(x) := \int_{-\pi}^{\pi} f(t)D_n(x-t)\,\frac{dt}{2\pi},$$

où D_n est le **noyau de Dirichlet** :

$$D_n(t) := \sum_{k=-n}^{n} e^{ikt} = \frac{\sin(n+\frac{1}{2})t}{\sin\frac{t}{2}}\cdot$$

En particulier :

$$S_n f(0) = \int_{-\pi}^{\pi} f(t)D_n(t)\,\frac{dt}{2\pi}\cdot$$

Dans la suite, on munit $\mathcal{C}_{2\pi}$ de la norme $\|\cdot\|_\infty$, qui en fait un espace de Banach car c'est un sous-espace fermé de $(\ell^\infty(\mathbb{R}), \|\cdot\|_\infty)$.

Fait 1. Pour $n \in \mathbb{N}$, soit $\Phi_n : \mathcal{C}_{2\pi} \to \mathbb{C}$ la forme linéaire définie par $\Phi_n(f) := S_n f(0)$. Alors (Φ_n est continue et)

$$\|\Phi_n\| = \|D_n\|_1 := \int_{-\pi}^{\pi} |D_n(t)| \, \frac{dt}{2\pi}.$$

Preuve du Fait 1. On a $|\Phi_n(f)| \leqslant \int_{-\pi}^{\pi} |f(t)||D_n(t)| \frac{dt}{2\pi} \leqslant \|f\|_\infty \times \|D_n\|_1$ pour toute $f \in \mathcal{C}_{2\pi}$, donc Φ_n est continue et $\|\Phi_n\| \leqslant \|D_n\|_1$. Inversement, on peut (exo) construire une suite $(f_k)_{k\in\mathbb{N}} \subseteq \mathcal{C}_{2\pi}$ telle que $\|f_k\|_\infty = 1$ pour tout k et $\lim_{k\to\infty} f_k(t) = \mathrm{signe}(D_n(t))$ pour tout $t \in [-\pi, \pi]$. Alors $\Phi_n(f_k) = \int_{-\pi}^{\pi} f_k(t) D_n(t) \frac{dt}{2\pi} \to \int_{-\pi}^{\pi} |D_n(t)| \frac{dt}{2\pi} = \|D_n\|_1$ quand $k \to \infty$, par convergence dominée. Comme $\|f_k\|_\infty = 1$ pour tout k, on en déduit $\|\Phi_n\| \geqslant \|D_n\|_1$. □

Fait 2. On a $\lim_{n\to\infty} \|D_n\|_1 = \infty$.

Preuve du Fait 2. Comme $|\sin(u)| \leqslant u$ pour tout $u \geqslant 0$, on a

$$\int_0^{\pi} |D_n(t)| \, \frac{dt}{2\pi} = \int_0^{\pi} \left| \frac{\sin(n+\frac{1}{2})t}{\sin\frac{t}{2}} \right| \frac{dt}{2\pi} \geqslant \frac{1}{\pi} \int_0^{\pi} \frac{|\sin(n+\frac{1}{2})t|}{t} \, dt = \frac{1}{\pi} \int_0^{(n+\frac{1}{2})\pi} \frac{|\sin u|}{u} \, du,$$

d'où le résultat car $\int_0^\infty \frac{|\sin u|}{u} \, du = \infty$. □

La preuve de la proposition est maintenant terminée : par les Faits 1 et 2, la suite de formes linéaires $(\Phi_n)_{n\in\mathbb{N}}$ n'est pas bornée en norme ; donc elle n'est pas simplement bornée, par Banach-Steinhaus. C'est exactement la conclusion souhaitée. □

4. Théorèmes de l'application ouverte et du graphe fermé

Théorème 4.1. *Soient X et Y deux evt* complètement métrisables*, et soit $T \in \mathcal{L}(X, Y)$. On suppose que T est* surjective*. Alors, pour tout voisinage W de 0 dans X, l'ensemble $T(W)$ est un voisinage de 0 dans Y.*

Démonstration. La preuve repose sur le fait suivant.

Fait. Pour tout voisinage U de 0 dans X, l'ensemble $\overline{T(U)}$ est un voisinage de 0 dans Y.

Preuve du Fait. Soit U' un voisinage de 0 équilibré tel que $U' - U' \subseteq U$. On a $X = \bigcup_{n\in\mathbb{N}^*} n\, U'$, donc $Y = T(X) = \bigcup_{n\in\mathbb{N}^*} n\, T(U')$, et *a fortiori* $Y = \bigcup_{n\in\mathbb{N}^*} n\, \overline{T(U')}$. D'après le Théorème de Baire (applicable car Y est complètement métrisable), l'un des ensembles fermés $n\, \overline{T(U')}$ est d'intérieur non-vide. Donc $\overline{T(U')}$ est d'intérieur non-vide (micro-exo). Notons O l'intérieur de $\overline{T(U')}$. Alors $V := O - O$ est un voisinage ouvert de 0, et $V \subseteq \overline{T(U')} - \overline{T(U')}$. De plus, on a $\overline{T(U')} - \overline{T(U')} \subseteq \overline{T(U') - T(U')}$ (exo), et $T(U') - T(U') = T(U' - U')$ par linéarité de T ; donc $V \subseteq \overline{T(U)}$, ce qui prouve le Fait. □

Soit maintenant W un voisinage de 0 quelconque dans X. Soit également d une distance définissant la topologie de X et invariante par translations. Choisissons $r > 0$ tel que $B_X(0, 2r) \subseteq W$, et posons $B_0 := B_X(0, r)$. Par le Fait, on sait que $\overline{T(B_0)}$ est un voisinage de 0 dans Y. Pour démontrer le théorème, il suffit donc de montrer que $\overline{T(B_0)} \subseteq T(W)$. Dans la suite, on posera $B_n := B_X(0, 2^{-n} r)$ pour tout $n \in \mathbb{N}$. On fixe également une distance d_Y définissant la topologie de Y. Enfin, on se donne une suite $(\varepsilon_n)_{n\geqslant 1} \subseteq \,]0, \infty[$ tendant vers 0.

Soit $y \in \overline{T(B_0)}$ quelconque : il s'agit de trouver $x \in W$ tel que $T(x) = y$.
Comme $V_1 := \overline{T(B_1)} \cap B_Y(0, \varepsilon_1)$ est un voisinage de 0 dans Y par le Fait, l'ensemble $y - V_1$ est un voisinage de y, et donc $(y - V_1) \cap T(B_0) \neq \varnothing$. Ainsi, on peut trouver $x_0 \in B_0$ tel que $y - T(x_0) \in V_1$; autrement dit

$$y - T(x_0) \in \overline{T(B_1)} \qquad \text{et} \qquad d_Y(y - T(x_0), 0) < \varepsilon_1.$$

En répétant ce raisonnement avec $y - T(x_0)$ au lieu de y et B_1 au lieu de B_0, on obtient un point $x_1 \in B_1$ tel que $y - T(x_0) - T(x_1) \in \overline{T(B_2)}$ et $d_Y(y - T(x_0) - T(x_1), 0) < \varepsilon_2$; autrement dit, par linéarité de T :

$$y - T(x_0 + x_1) \in \overline{T(B_2)} \qquad \text{et} \qquad d_Y\big(y - T(x_0 + x_1), 0\big) < \varepsilon_2.$$

Et ainsi de suite : par récurrence, on construit une suite $(x_n)_{n \in \mathbb{N}} \subseteq X$ telle que pour tout $n \in \mathbb{N}$, les choses suivantes aient lieu :

(i) $x_n \in B_n$, *i.e.* $d(x_n, 0) < 2^{-n} r$;

(ii) $d_Y\big(y - T(x_0 + x_1 + \cdots + x_n), 0\big) < \varepsilon_{n+1}$;

(iii) $y - T(x_0 + x_1 + \cdots + x_n) \in \overline{T(B_{n+1})}$.

Comme la série $\sum d(x_k, 0)$ est convergente, que la distance d est invariante par translations et que (X, d) est complet, la série $\sum x_k$ converge dans X d'après le Corollaire 1.7. Posons $x := \sum_{k=0}^{\infty} x_k$. Alors $T(x) = y$ par (ii) et par continuité de T. De plus, en utilisant (i) on vérifie (exo) qu'on a $d(x, 0) < \sum_{k=0}^{\infty} 2^{-k} r = 2r$; donc $x \in W$ puisque $B_X(0, 2r) \subseteq W$. □

Corollaire 4.2. (Théorème de majoration *a priori*)
Soient X et Y deux espaces de Banach, et soit $T \in \mathcal{L}(X, Y)$. On suppose que T est surjective. Alors, il existe une constante $C < \infty$ telle que :

$$\forall y \in Y\ \exists x \in X \quad \textit{vérifiant} \quad Tx = y \quad \text{et} \quad \|x\| \leqslant C\,\|y\|.$$

Démonstration. Notons B_X la boule unité fermée de X. Par le Théorème, $T(B_X)$ est un voisinage de 0 dans Y, donc contient une boule fermée $\overline{B}(0, c)$. Ainsi, pour tout $y \in Y$ vérifiant $\|y\| \leqslant c$, on peut trouver $x \in X$ tel que $\|x\| \leqslant 1$ et $T(x) = y$. Par homogénéité, on en déduit le résultat souhaité avc $C := 1/c$ (exo). □

Corollaire 4.3. (Théorème de l'application ouverte)
Soient X et Y des evt complètement métrisables. Si $T \in \mathcal{L}(X, Y)$ est surjective, alors T est une application ouverte : pour tout ouvert $O \subseteq X$, l'ensemble $T(O)$ est un ouvert de Y.

Démonstration. Soit $y \in T(O)$ quelconque, et soit $x \in O$ tel que $T(x) = y$. Comme O est un voisinage de x, on peut écrire $O = x + W$ où W est un voisinage de 0. Alors $T(O) = y + T(W)$ par linéarité de T ; donc $T(O)$ est un voisinage de y puisque $T(W)$ est un voisinage de 0. Ceci étant vrai pour tout $y \in T(O)$, on en déduit que $T(O)$ est un ouvert de Y. □

Corollaire 4.4. (Théorème d'isomorphisme de Banach)
Soient X et Y des evt complètement métrisables. Si $T \in \mathcal{L}(X, Y)$ est bijective, alors l'application linéaire $T^{-1} : Y \to X$ est continue.

Démonstration. Conséquence "immédiate" du Théorème de l'application ouverte (micro-exo). □

Corollaire 4.5. (Théorème du graphe fermé)
Soient X et Y des evt complètement métrisables, et soit $T : X \to Y$ une application linéaire. Soit $G_T := \big\{(x, Tx);\ x \in X\big\}$ le graphe de T. Si G_T est fermé dans $X \times Y$, alors T est continue.

Démonstration. L'espace $X \times Y$ muni de la topologie produit est un evt (exo), et il est complètement métrisable car X et Y le sont. Donc G_T est un evt complètement métrisable puisque c'est un sous-espace vectoriel de $X \times Y$ et qu'on suppose qu'il est *fermé* dans $X \times Y$.
Soit $J : X \to G_T$ l'application linéaire définie par $J(x) := (x, Tx)$. Par définition, J est *bijective*, avec $J^{-1}(x, Tx) = x$ pour tout $(x, Tx) \in G_T$. Comme la projection canonique $\pi_X : X \times Y \to X$ est continue, on voit que J^{-1} est continue. Comme G_T et X sont des evt complètement métrisables, on en déduit que $J = (J^{-1})^{-1}$ est continue. Et donc T est continue car $T = \pi_Y \circ J$. □

Remarque 1. La réciproque du théorème du graphe fermé est vraie et à peu près évidente : le graphe de n'importe quelle application continue $f : X \to Y$ est toujours fermé dans $X \times Y$.

Démonstration. Exo. □

Remarque 2. Soit $T : X \to Y$ linéaire. Pour montrer que le graphe de T est fermé, il suffit de vérifier la chose suivante : pour toute suite $(x_n) \subseteq X$ telle que $x_n \to 0$ et telle que la suite (Tx_n) converge dans Y, on a que $Tx_n \to 0$.

Démonstration. Exo. □

Exemple 1. Projections continues.

Lemme 4.6. *Soit X un evt complètement métrisable. Si E et F sont des sous-espaces fermés de X tels que $X = E \bigoplus F$, alors les projections $p : X \to E$ et $q : X \to F$ associées à cette décomposition sont continues.*

Démonstration. Pour montrer que $p : X \to E$ est continue, on applique le Théorème du graphe fermé, ce qui est possible car X et E sont complètement métrisables. Soit $(x_n) \subseteq X$ une suite telle que $x_n \to x \in X$ et $px_n \to z \in E$. Alors $pz = z$, et $x_n - px_n \to x - z$. Comme $x_n - px_n \in \ker(p) = F$ et comme F est fermé dans X, on a donc $x - z \in F = \ker(p)$, et donc $z = pz = px$. □

Exemple 2. Non-surjectivité de la transformation de Fourier.

Si $f \in L^1(\mathbb{R})$, sa transformée de Fourier est la fonction $\widehat{f} : \mathbb{R} \to \mathbb{C}$ définie par
$$\widehat{f}(x) := \int_{\mathbb{R}} f(t) e^{-ixt} dt.$$
On sait que $\widehat{f}$ est continue bornée, avec $\|\widehat{f}\|_\infty \leqslant \|f\|_1$. De plus, le *Lemme de Riemann-Lebesgue* dit que $\widehat{f}(x) \to 0$ quand $x \to \pm\infty$. Ainsi, en notant $\mathcal{C}_0(\mathbb{R})$ l'ensemble des fonctions continues $g : \mathbb{R} \to \mathbb{C}$ tendant vers 0 à l'infini, la transformation de Fourier est un opérateur borné de $L^1(\mathbb{R})$ dans $(\mathcal{C}_0(\mathbb{R}), \|\cdot\|_\infty)$.

Proposition 4.7. *La transformation de Fourier $\mathcal{F} : L^1(\mathbb{R}) \to \mathcal{C}_0(\mathbb{R})$ n'est pas surjective. Autrement dit : il existe une fonction $g \in \mathcal{C}_0(\mathbb{R})$ qui n'est la transformée de Fourier d'aucune $f \in L^1(\mathbb{R})$.*

Démonstration. Par la formule d'inversion de Fourier, on sait que $\mathcal{F} : L^1(\mathbb{R}) \to \mathcal{C}_0(\mathbb{R})$ est injective. Supposons que $\mathcal{F}$ soit surjective. Alors, comme $L^1(\mathbb{R})$ et $\mathcal{C}_0(\mathbb{R})$ sont des espaces de Banach, $\mathcal{F}$ est inversible ; donc on peut trouver une constante $c > 0$ telle que $\forall f \in L^1(\mathbb{R}) \ : \ \|\widehat{f}\|_\infty \geqslant c\,\|f\|_1$. Montrons que cela mène à une contradiction.
Pour $0 < \varepsilon < 1$, soit $g_\varepsilon : \mathbb{R} \to \mathbb{R}$ la fonction continue valant 1 sur $[-1+\varepsilon, 1-\varepsilon]$, nulle en dehors de $[-1-\varepsilon, 1+\varepsilon]$, et affine sur les intervalles $[-1-\varepsilon, -1+\varepsilon]$ et $[1-\varepsilon, 1+\varepsilon]$. On a (exo)

$$g_\varepsilon = \frac{1}{2\varepsilon}\,\mathbf{1}_{[-\varepsilon,\varepsilon]} * \mathbf{1}_{[-1,1]}.$$

Un calcul direct donne alors

$$\widehat{g_\varepsilon}(t) = \frac{2}{\varepsilon}\,\frac{\sin t\,\sin(\varepsilon t)}{t^2} := f_\varepsilon(t).$$

On en déduit (exo) que $f_\varepsilon = \widehat{g_\varepsilon} \in L^1(\mathbb{R})$; et donc, par la formule d'inversion de Fourier et comme g_ε est une fonction paire, que $g_\varepsilon = \frac{1}{2\pi}\,\widehat{f_\varepsilon}$. Donc, en posant $\kappa := c/2\pi$, on doit avoir

$$\|g_\varepsilon\|_\infty \geqslant \kappa\,\|f_\varepsilon\|_1 \qquad \text{pour tout } \varepsilon > 0.$$

Cependant, on a $\|g_\varepsilon\|_\infty = 1$; et comme $\sin u \geqslant \frac{2}{\pi}\,u$ si $u \in [0, \frac{\pi}{2}]$,

$$\|f_\varepsilon\|_1 \geqslant \frac{1}{\varepsilon}\int_0^{\frac{\pi}{2\varepsilon}} \frac{|\sin t|}{t} \times \frac{|\sin(\varepsilon t)|}{t}\,dt \geqslant \frac{2}{\pi}\int_0^{\frac{\pi}{2\varepsilon}} \frac{|\sin t|}{t}\,dt \xrightarrow{\varepsilon\to 0^+} \infty.$$

D'où la contradiction cherchée. ☐

Chapitre 4

Mesures

1. Mesures complexes

1.1. Définition ; variation totale.

DÉFINITION 1.1. *Soit* $(\Omega, \mathfrak{B})$ *un espace mesurable. Une* ***mesure complexe*** *sur* $(\Omega, \mathfrak{B})$ *est une fonction* $\mu : \mathfrak{B} \to \mathbb{C}$ dénombrablement additive *: pour toute suite* $(A_k)_{k\in\mathbb{N}}$ *d'ensembles mesurables deux à deux disjoints, la série* $\sum \mu(A_k)$ *converge et on a* $\mu\big(\bigcup_{k=0}^\infty A_k\big) = \sum_{k=0}^\infty \mu(A_k)$. *Une* ***mesure réelle*** *est une mesure complexe prenant ses valeurs dans* $\mathbb{R}$.

Remarque 0. Une *mesure positive* μ sur $(\Omega, \mathfrak{B})$ est une mesure complexe au sens de la définition si et seulement si c'est une mesure *finie*.

Remarque 1. Si μ est une mesure complexe, alors $\mathrm{Re}(\mu)$ et $\mathrm{Im}(\mu)$ sont des mesures réelles. Donc, toute mesure complexe μ s'écrit $\mu = \mu_1 + i\mu_2$, où μ_1 et μ_2 sont des mesures réelles.

Remarque 2. Si μ est une mesure complexe et si $(A_k)_{k\in\mathbb{N}}$ est une suite d'ensembles mesurables deux à deux disjoints, alors la série $\sum \mu(A_k)$ converge *absolument.*

Démonstration. Il découle de la définition que pour toute permutation $\sigma : \mathbb{N} \to \mathbb{N}$, la série $\sum \mu\big(A_{\sigma(k)}\big)$ doit converger ; et ceci entraine la convergence absolue de la série $\sum \mu(A_k)$ $\square$

Remarque 3. Si μ est une mesure complexe, on a toujours $\mu(\varnothing) = 0$.

Démonstration. Si on prend $A_k := \varnothing$ pour tout $k \in \mathbb{N}$, on voit que la série $\sum \mu(\varnothing)$ doit converger. $\square$

Remarque 4. L'ensemble de toutes les mesures complexes sur $(\Omega, \mathfrak{B})$ est un $\mathbb{C}$-espace vectoriel, qu'on notera $M(\Omega, \mathfrak{B})$, ou plus simplement $M(\Omega)$ en “oubliant” la tribu $\mathfrak{B}$. On note aussi $M_\mathbb{R}(\Omega)$ l'ensemble des mesures réelles (qui est un $\mathbb{R}$-espace vectoriel), et $M_+(\Omega)$ l'ensemble des mesures positives finies.

EXEMPLE. Soit m une mesure positive sur $(\Omega, \mathfrak{B})$. Pour toute $h \in L^1(\Omega, m)$, on définit une mesure complexe sur $(\Omega, \mathfrak{B})$ en posant

$$\mu(A) := \int_A h\, dm \qquad \text{pour tout } A \in \mathfrak{B}.$$

Cette mesure se note $\mu = hm$.

Démonstration. Si $(A_k)_{k\in\mathbb{N}}$ est une suite d'ensembles mesurables deux à deux disjoints et si $A := \bigcup_{k=0}^\infty A_k$, alors

$$\mu(A) = \int_\Omega \mathbf{1}_A f\, dm = \int_\Omega \sum_{k=0}^\infty \mathbf{1}_{A_k} f\, dm = \sum_{k=0}^\infty \int_\Omega \mathbf{1}_{A_k} f\, dm = \sum_{k=0}^\infty \mu(A_k),$$

où on a utilisé le Théorème de convergence dominée pour intervertir la somme et l'intégrale. $\square$

EXERCICE. Montrer que si $h_1, h_2 \in L^1(\Omega, m)$ et si $h_1 m = h_2 m$ alors $h_1 = h_2$. En déduire que si $h \in L^1(\Omega, m)$, alors la mesure $\mu := hm$ est réelle si et seulement si h est m-pp à valeurs réelles, et que μ est positive si et seulement si $h(x) \geqslant 0$ pp.

TERMINOLOGIE. On appellera *partition mesurable* d'un ensemble mesurable $A \subseteq \Omega$ toute famille $(E_i)_{i\in I}$ d'ensembles mesurables deux à deux disjoints telle que $\bigcup_{i\in I} E_i = A$. (On ne demande pas que les E_i soient $\neq \varnothing$.)

THÉORÈME 1.2. *Soit μ une mesure complexe sur $(\Omega, \mathfrak{B})$. Pour tout $A \in \mathfrak{B}$, posons*

$$|\mu|(A) := \sup \left\{ \sum_{n=0}^{\infty} |\mu(E_n)|;\ (E_n)_{n\in\mathbb{N}} \textit{ partition mesurable de } A \right\}.$$

Alors $|\mu|$ est une mesure positive finie *sur $(\Omega, \mathfrak{B})$, qu'on appelle la* ***variation totale*** *de μ. On a $\forall A \in \mathfrak{B}\ :\ |\mu|(A) \geqslant |\mu(A)|$, et $|\mu|$ est la* plus petite *mesure positive possèdant cette propriété.*

Démonstration. (i) Montrons que $|\mu|$ est une mesure. Comme $\mu(\varnothing) = 0$, on a visiblement $|\mu|(\varnothing) = 0$. Soit $(A_k)_{k\in\mathbb{N}}$ une suite d'ensembles mesurables deux à deux disjoints, et soit $A := \bigcup_{k\in\mathbb{N}} A_k$. Si $(E_n)_{n\in\mathbb{N}}$ est une partition mesurable de A alors $(E_n \cap A_k)_{n\in\mathbb{N}}$ est une partition mesurable de A_k pour tout $k \in \mathbb{N}$. Donc

$$\forall k \in \mathbb{N}\ :\ \sum_{n=0}^{\infty} |\mu(E_n \cap A_k)| \leqslant |\mu|(A_k).$$

De plus, comme $E_n \subseteq A$ on a $E_n = \bigcup_{k=0}^{\infty} (E_n \cap A_k)$, et la réunion est disjointe ; donc

$$\mu(E_n) = \sum_{k=0}^{\infty} \mu(E_n \cap A_k) \qquad \text{pour tout } n \in \mathbb{N}.$$

On en déduit

$$\begin{aligned} \sum_{n=0}^{\infty} |\mu(E_n)| &= \sum_{n=0}^{\infty} \left| \sum_{k=0}^{\infty} \mu(E_n \cap A_k) \right| \\ &\leqslant \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |\mu(E_n \cap A_k)| \\ &= \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} |\mu(E_n \cap A_k)| \leqslant \sum_{k=0}^{\infty} |\mu|(A_k). \end{aligned}$$

Ceci étant vrai pour toute partition mesurable de A, on a donc

$$|\mu|(A) \leqslant \sum_{k=0}^{\infty} |\mu|(A_k).$$

Inversement, montrons qu'on a $\sum_{k=0}^{\infty} |\mu|(A_k) \leqslant |\mu|(A)$. Soit $\mathbf{K} := \{k \in \mathbb{N};\ |\mu|(A_k) > 0\}$, et soit $\widetilde{A} := \bigcup_{k\in\mathbf{K}} A_k$. Comme $\widetilde{A} \subseteq A$, on a $|\mu|(\widetilde{A}) \leqslant |\mu|(A)$ car la définition de $|\mu|$ entraîne que $|\mu|$ est une fonction croissante (exo). Donc, quitte à remplacer A par $\widetilde{A}$, on peut supposer que $|\mu|(A_k) > 0$ pour tout $k \in \mathbb{N}$. Soit $\varepsilon > 0$, et soit $(t_k)_{k\in\mathbb{N}}$ une suite de nombres réels telle que $0 \leqslant t_k < |\mu|(A_k)$ pour tout k et $\sum_{k=0}^{\infty} t_k \geqslant \sum_{k=0}^{\infty} |\mu|(A_k) - \varepsilon$.

Pour tout $k \in \mathbb{N}$, on peut choisir une une partition mesurable $(E_{k,n})_{n\in\mathbb{N}}$ de A_k telle que $\sum_{n=0}^{\infty} |\mu(E_{k,n})| > t_k$. Alors $(E_{k,n})_{k,n\in\mathbb{N}}$ est une partition mesurable de A ; donc

$$|\mu|(A) \geqslant \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} |\mu(E_{k,n})| \geqslant \sum_{k=0}^{\infty} t_k \geqslant \sum_{k=0}^{\infty} |\mu|(A_k) - \varepsilon.$$

Ceci étant vrai pour tout $\varepsilon > 0$, on en déduit $|\mu|(A) \geqslant \sum_{k=0}^{\infty} |\mu|(A_k)$.

(ii) Montrons maintenant que la mesure $|\mu|$ est finie. Supposons que $|\mu|(\Omega) = \infty$, et cherchons à obtenir une contradiction.

Fait. Soit $H \in \mathfrak{B}$ et supposons que $|\mu|(H) = \infty$. Alors, on peut trouver deux ensembles mesurables $A, H' \subseteq H$ tels que $A \cap H' = \varnothing$, $|\mu|(A) > 1$ et $|\mu|(H') = \infty$.

Preuve du Fait. Soit $R \in \mathbb{R}^+$ à fixer ultérieurement. Comme $|\mu|(H) = \infty$, on peut choisir $(E_n)_{n\in\mathbb{N}}$, partition mesurable de H telle que $\sum_{n=0}^{\infty} |\mu(E_n)| > R$, puis $N \in \mathbb{N}$ tel que $\sum_{n=0}^{N} |\mu(E_n)| > R$. On a alors par exemple $\sum_{n=0}^{N} |\mathrm{Re}\,\mu(E_n)| > R/2$. Donc, en posant $I := \big\{n \in [\![0, N]\!];\ \mathrm{Re}\,\mu(E_n) > 0\big\}$ et $J := \big\{n \in [\![0, N]\!];\ \mathrm{Re}\,\mu(E_n) \leqslant 0\big\}$, on a

$$\mathrm{Re}\left(\sum_{n\in I} \mu(E_n)\right) - \mathrm{Re}\left(\sum_{n\in J} \mu(E_n)\right) > R/2.$$

On peut donc par exemple supposer que

$$\left|\sum_{n\in I} \mu(E_n)\right| > R/4.$$

Autrement dit, en posant $E := \bigcup_{n\in I} E_n$, on a

$$E \subseteq H \qquad \text{et} \qquad |\mu(E)| > R/4.$$

On en déduit

$$|\mu(H\backslash E)| = |\mu(H) - \mu(E)| \geqslant |\mu(E)| - |\mu(H)| > R/4 - |\mu(H)|.$$

Si on a choisi au départ R tel que $R/4 > 1 + |\mu(H)|$, on aura ainsi

$$|\mu(E)| > 1 \qquad \text{et} \qquad |\mu(H\backslash E)| > 1.$$

Enfin, comme $|\mu|$ est une mesure, on a $\infty = |\mu|(H) = |\mu|(E) + |\mu|(H\backslash E)$, donc $|\mu|(E) = \infty$ ou $|\mu|(H\backslash E) = \infty$. Il suffit donc de poser $A := E$ et $H' := H\backslash E$, ou le contraire. ☐

En posant $H_0 := \Omega$ et en appliquant le Fait de manière répétée, on construit par récurrence deux suites d'ensembles mesurables $(A_k)_{k\geqslant 0}$ et $(H_k)_{k\geqslant 0}$ telles que pour tout $k \in \mathbb{N}$, les propriétés suivantes soient satisfaites :

- $A_k \cup H_{k+1} \subseteq H_k$;
- $A_k \cap H_{k+1} = \varnothing$;
- $|\mu(A_k)| > 1$ et $|\mu|(H_{k+1}) = \infty$.

Par construction, les A_k sont deux à deux disjoints : si $k < l$ alors $A_l \subseteq H_l \subseteq H_{k+1}$ et $H_{k+1} \cap A_k = \varnothing$, donc $A_k \cap A_l = \varnothing$; mais la série $\sum \mu(A_k)$ ne converge pas puisque $|\mu(A_k)| > 1$ pour tout k. Comme μ est supposée être une mesure complexe, on aboutit ainsi à une contradiction.

(iii) Par définition, on a $|\mu|(A) \geqslant |\mu(A)|$ pour tout $A \in \mathfrak{B}$ car $(A, \varnothing, \varnothing, \dots)$ est une partition mesurable de A (!) Soit m une mesure positive telle que $\forall A \in \mathfrak{B}$: $m(A) \geqslant |\mu(A)|$. Si $A \in \mathfrak{B}$ et si $(E_n)_{n\in\mathbb{N}}$ est une partition mesurable quelconque de

A, alors $m(A) = \sum_{n=0}^{\infty} m(E_n) \geqslant \sum_{n=0}^{\infty} |\mu(E_n)|$; donc $m(A) \geqslant |\mu|(A)$ par définition de $|\mu|(A)$. □

REMARQUE. La mesure μ est $\geqslant 0$ si et seulement si $|\mu| = \mu$.

Démonstration. Si $|\mu| = \mu$, alors certainement $\mu \geqslant 0$. Inversement, si $\mu \geqslant 0$ alors $\mu = |\mu|$ par la propriété de minimalité de $|\mu|$. □

COROLLAIRE 1.3. *Soit μ une mesure* réelle *sur $(\Omega, \mathfrak{B})$. Si on pose $\mu^+ := \frac{1}{2}(|\mu| + \mu)$ et $\mu^- := \frac{1}{2}(|\mu| - \mu)$, alors μ^+ et μ^- sont des mesures positives finies telles que $\mu = \mu^+ - \mu^-$ et $\mu^+ + \mu^- = |\mu|$. On dit que μ^+ est la **partie positive** de μ, et que μ^- est la **partie négative** de μ.*

Démonstration. Par définition, μ^+ et μ^- sont des mesures réelles telles que $\mu = \mu^+ - \mu^-$ et $\mu^+ + \mu^- = |\mu|$. De plus, on a $\mu^+(A) \geqslant \frac{1}{2}(|\mu|(A) - |\mu(A)|) \geqslant 0$ pour tout $A \in \mathfrak{B}$, donc μ^+ est une mesure positive ; et de même pour μ^-. □

COROLLAIRE 1.4. *Toute mesure complexe est combinaison linéaire de 4 mesures positives finies.*

EXEMPLE. Soit m une mesure positive sur $(\Omega, \mathfrak{B})$, soit $h \in L^1(\Omega, m)$, et soit $\mu := hm$. Alors $|\mu| = |h|m$. Si h est une fonction réelle, alors $\mu^+ = h^+m$ et $\mu^- = h^-m$.

Démonstration. Il est assez clair que $\nu := |h|m$ est une mesure positive telle que $\forall A \in \mathfrak{B} \ : \ \nu(A) \geqslant |\mu(A)|$; donc $|\mu| \leqslant |h|m$. Inversement, montrons que pour tout $A \in \mathfrak{B}$, on a $\int_A |h|\, dm \leqslant |\mu|(A)$. Soit θ une fonction mesurable telle que $|\theta(x)| \equiv 1$ et $|h| = \theta h$. On peut trouver une suite $(\varphi_n)_{n\in\mathbb{N}}$ de fonctions étagées à valeurs complexes telle que $|\varphi_n| \leqslant 1$ et $\varphi_n(x) \to \theta(x)$ pour tout $x \in \Omega$. Alors $\int_A \varphi_n h\, dm \to \int_A |h|\, dm$ par convergence dominée ; donc il suffit de montrer qu'on a $\left|\int_A \varphi_n h\, dm\right| \leqslant |\mu|(A)$ pour tout $n \in \mathbb{N}$. Si on écrit $\varphi_n = \sum_{i=1}^K \alpha_i \mathbf{1}_{E_i}$ où les E_i forment une partition de Ω et $|\alpha_i| \leqslant 1$, on obtient $\left|\int_A \varphi_n h\, dm\right| \leqslant \sum_{i=1}^K |\alpha_i| \left|\int_{A\cap E_i} h\, dm\right| \leqslant \sum_{i=1}^K |\mu(A \cap E_i)| \leqslant |\mu|(A)$ par définition de $|\mu|$. Ainsi, on a bien $|\mu| = |h|m$. Si h est réelle, on a donc $\mu^+ = \frac{1}{2}(|h| + h)m$ et $\mu^- = \frac{1}{2}(|h| - h)m$, autrement dit $\mu^+ = h^+m$ et $\mu^- = h^-m$. □

1.2. Intégration par rapport à une mesure complexe.

DÉFINITION 1.5. *Soit $f : \Omega \to \mathbb{C}$ une fonction* mesurable bornée. *Pour toute mesure réelle μ sur $(\Omega, \mathfrak{B})$, on pose $\int_\Omega f\, d\mu := \int_\Omega f\, d\mu^+ - \int_\Omega f\, d\mu^-$. Pour toute mesure complexe μ s'écrivant $\mu = \mu_1 + i\mu_2$ avec μ_1, μ_2 réelles, on pose $\int_\Omega f\, d\mu := \int_\Omega f\, d\mu_1 + i \int_\Omega f\, d\mu_2$.*

EXEMPLE 1. Si f est une fonction *étagée* et si on écrit $f = \sum_{k=1}^N \alpha_k \mathbf{1}_{E_k}$, alors $\int_\Omega f\, d\mu = \sum_{k=1}^N \alpha_k \mu(E_k)$.

Démonstration. Exo. □

EXEMPLE 2. Soit m une mesure positive sur $(\Omega, \mathfrak{B})$, et soit $\mu := hm$, où $h \in L^1(\Omega, m)$. Pour toute fonction mesurable bornée $f : \Omega \to \mathbb{C}$, on a $\int_\Omega f\, d\mu = \int_\Omega fh\, dm$.

Démonstration. Par l'Exemple 1, le résultat est vrai pour une fonction f étagée : si $f = \sum_{k=1}^{N} \alpha_k \mathbf{1}_{E_k}$, alors $\int_\Omega f\,d\mu = \sum_{k=1}^{N} \alpha_k \mu(E_k) = \sum_{k=1}^{N} \alpha_k \int_{E_k} h\,dm = \int_\Omega fh\,dm$. Pour une fonction f mesurable bornée quelconque, on peut trouver une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions étagées telle que $f_n \to f$ *uniformément* sur Ω. Alors $\int_\Omega f_n\,d\mu \to \int_\Omega f\,d\mu$ (exo), d'où le résultat. □

REMARQUE. L'application $(\mu, f) \mapsto \int_\Omega f\,d\mu$ est bilinéaire.

Démonstration. Exo. □

1.3. Norme d'une mesure complexe.

LEMME 1.6. *Pour toute mesure $\mu \in M(\Omega)$, posons $\|\mu\|_M := |\mu|(\Omega)$. Alors $\|\cdot\|_M$ est une norme sur $M(\Omega)$.*

Démonstration. Pour toute partition mesurable $(E_n)_{n\in\mathbb{N}}$ de Ω, l'application $\mu \mapsto \sum_{n=0}^{\infty} |\mu(E_n)|$ est une semi-norme sur $M(\Omega)$. Comme un sup de semi-normes est encore une semi-norme (exo), on en déduit que $\|\cdot\|_M$ est une semi-norme. Enfin, si $\|\mu\|_M = 0$, *i.e.* $|\mu|(\Omega) = 0$, alors $|\mu| = 0$ car $|\mu|$ est une mesure positive, donc $\mu = 0$ car $|\mu(A)| \leqslant |\mu|(A)$ pour tout $A \in \mathfrak{B}$. □

EXEMPLE. Si $\mu := hm$, où m est une mesure positive et $h \in L^1(\Omega, m)$, alors $\|\mu\|_M = \|h\|_{L^1(\Omega,m)}$.

Démonstration. On a vu que $|\mu| = |h|m$, donc $|\mu|(\Omega) = \int_\Omega |h|\,dm$. □

2. Théorème de Lebesgue-Radon-Nikodym

DÉFINITION 2.1. *Soit m une mesure $\geqslant 0$ sur $(\Omega, \mathfrak{B})$, et soit μ une mesure complexe.*

(a) *On dit que μ est **singulière par rapport à** m, et on écrit $\mu \perp m$, s'il existe un ensemble mesurable $D \subseteq \Omega$ tel que $m(D) = 0$ et μ est* portée par *D, i.e. $\mu(A) = \mu(A \cap D)$ pour tout $A \in \mathfrak{B}$ (ou encore : $\mu(B) = 0$ pour tout $B \in \mathfrak{B}$ tel que $B \cap D = \varnothing$).*

(b) *On dit que μ est **absolument continue par rapport à** m, et on écrit $\mu \ll m$, si l'implication suivante a lieu pour tout ensemble mesurable $A \subseteq \Omega$: $m(A) = 0 \implies \mu(A) = 0$.*

EXEMPLES. Si $h \in L^1(\Omega, m)$, alors $\mu := hm \ll m$. Si $m = \lambda_1$, la mesure de Lebesgue sur $\mathbb{R}$, alors $\delta_a \perp m$ pour tout $a \in \mathbb{R}$.

REMARQUE. Si m est une mesure $\geqslant 0$ sur $(\Omega, \mathfrak{B})$, la seule mesure complexe μ telle que $\mu \perp m$ et $\mu \ll m$ est $\mu = 0$.

Démonstration. Exo. □

THÉORÈME 2.2. *Soit m une mesure $\geqslant 0$ et sigma-finie sur $(\Omega, \mathfrak{B})$, et soit μ une mesure complexe.*

(1) *On peut écrire de manière unique μ sous la forme $\mu = \lambda + \nu$, où λ et ν sont des mesures complexes telles que $\lambda \ll m$ et $\nu \perp m$. De plus, si $\mu \geqslant 0$ alors $\lambda \geqslant 0$ et $\nu \geqslant 0$. L'écriture $\mu = \lambda + \nu$ est la **décomposition de Lebesgue** de la mesure μ.*

(2) *La mesure λ est de la forme $\lambda = hm$, pour une certaine $h \in L^1(\Omega, m)$.*

Démonstration. Comme μ est combinaison linéaire de mesures positives finies, on peut se contenter de démontrer le théorème pour $\mu \geqslant 0$ (et finie).
Par ailleurs, l'unicité dans (1) est facile à démontrer : si $\mu = \lambda_1 + \nu_1 = \lambda_2 + \nu_2$ avec $\lambda_i \ll m$ et $\nu_i \perp m$, alors la mesure $\lambda_2 - \lambda_1 = \nu_1 - \nu_2$ est à la fois $\ll m$ et $\perp m$ (exo), donc $\lambda_2 - \lambda_1 = 0 = \nu_1 - \nu_2$. Il suffit donc de prouver la partie "existence" de (1), et (2). On va en fait démontrer les deux choses en même temps.

Cas 1. On suppose que la mesure m est finie.

L'idée est de chercher λ sous la forme $\lambda = hm$, où h est la "plus grande fonction possible". Posons

$$\begin{aligned}\mathcal{H} : &= \big\{h : \Omega \to [0,\infty] \text{ mesurable; } hm \leqslant \mu\big\}\\ &= \Big\{h \geqslant 0;\ \forall A \in \mathfrak{B}\ :\ \int_A h\, dm \leqslant \mu(A)\Big\}.\end{aligned}$$

Comme $\mu \geqslant 0$, l'ensemble $\mathcal{H}$ contient $h := 0$, et donc $\mathcal{H} \neq \emptyset$. On peut donc poser

$$\alpha := \sup\Big\{\int_\Omega h\, dm;\ h \in \mathcal{H}\Big\};$$

et on a $\alpha \leqslant \mu(\Omega) < \infty$.

Fait 1. Il existe une fonction $h \in \mathcal{H}$ telle que $\int_\Omega h\, dm = \alpha$.

Preuve du Fait 1. On remarque d'abord que si $h_1, h_2 \in \mathcal{H}$, alors $h := \max(h_1, h_2) \in \mathcal{H}$. En effet, on a $h = \mathbf{1}_E h_1 + \mathbf{1}_{E^c} h_2$, où $E := \{h_1 \geqslant h_2\}$, donc $\forall A \in \mathfrak{B}\ :\ \int_A h\, dm = \int_{A\cap E} h_1\, dm + \int_{A\cap E^c} h_2\, dm \leqslant \mu(A \cap E) + \mu(A \cap E^c) = \mu(A)$.
On en déduit qu'il existe une suite *croissante* $(h_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$ telle que $\int_\Omega h_n\, dm \to \alpha$: si (u_n) est une suite d'éléments de $\mathcal{H}$ telle que $\int_\Omega u_n\, dm \to \alpha$, on peut prendre $h_n := \max(u_0, \dots, u_n)$. Alors la fonction $h := \sup_{n\in\mathbb{N}} h_n$ convient (exo). $\square$

Comme $h \in \mathcal{H}$, la mesure hm est positive, finie, et $hm \leqslant \mu$. Donc $\nu := \mu - hm$ est une mesure positive bien définie.Pour terminer la preuve, il suffit de montrer que $\nu \perp m$.
Pour $\varepsilon > 0$, posons

$$\mathfrak{E}_\varepsilon := \big\{E \in \mathfrak{B};\ \nu(E) < \varepsilon\, m(E)\big\}.$$

Notons que si $E \in \mathfrak{E}_\varepsilon$, alors $m(E) > 0$. Cependant, il n'est pas immédiatement apparent que $\mathfrak{E}_\varepsilon \neq \emptyset$.

Fait 2. Soit $\varepsilon > 0$. Si $A \in \mathfrak{B}$ vérifie $m(A) > 0$, alors on peut trouver $E \in \mathfrak{E}_\varepsilon$ tel que $E \subseteq A$.

Preuve du Fait 2. Comme $m(A) > 0$, on a $\int_\Omega (h + \varepsilon\mathbf{1}_A)\, dm = \alpha + \varepsilon m(A) > \alpha$; donc la fonction $h + \varepsilon 1_A$ n'appartient pas à $\mathcal{H}$. Par définition de $\mathcal{H}$, on peut donc trouver $B \in \mathfrak{B}$ tel que $\int_B (h + \varepsilon\mathbf{1}_A)\, dm > \mu(B)$. On a ainsi $\mu(B) - \int_B h\, dm < \varepsilon\, m(A \cap B)$, *i.e.* $\nu(B) < \varepsilon\, m(A \cap B)$. Comme ν est une mesure $\geqslant 0$, on a donc *a fortiori* $\nu(A \cap B) < \varepsilon\, m(A \cap B)$, donc $E := A \cap B$ convient. $\square$

Fait 3. Pour tout $\varepsilon > 0$, on peut trouver $E \in \mathfrak{E}_\varepsilon$ tel que $m(\Omega\backslash E) = 0$.

Preuve du Fait 3. Par le Lemme de Zorn, il existe une partie $\mathfrak{F}$ de $\mathfrak{E}_\varepsilon$ qui est formée d'ensembles deux à deux disjoints et *maximale* relativement à cette propriété. Comme $m(F) > 0$ pour tout $F \in \mathfrak{F}$ puisque $\mathfrak{F} \subseteq \mathfrak{E}_\varepsilon$ et comme la mesure m est (sigma-)finie, $\mathfrak{F}$ est nécessairement dénombrable (exo) ; donc on peut écrire $\mathfrak{F} = \{F_k;\ k \in \mathbf{K}\}$, où

$\mathbf{K} \subseteq \mathbb{N}$. Posons $E := \bigcup_{k\in\mathbf{K}} F_k$. Comme les F_k sont disjoints, on a $\nu(E) = \sum_{k\in\mathbf{K}} \nu(F_k) < \sum_{k\in\mathbf{K}} \varepsilon\, m(F_k) = \varepsilon\, m(E)$; donc $E \in \mathfrak{F}_\varepsilon$. Si on avait $m(\Omega\backslash E) > 0$, on pourrait trouver graâce au Fait 2 un ensemble $E \in \mathfrak{E}_\varepsilon$ tel que $E \subseteq \Omega\backslash \bigcup_{k\in\mathbb{N}} F_k$. Alors $\mathfrak{F} \cup \{E\}$ serait une partie de $\mathfrak{E}_\varepsilon$ contenant strictement $\mathfrak{F}$ et formée d'ensembles deux à deux disjoints, ce qui contredirait la maximalité de $\mathfrak{F}$. $\square$

Par le Fait 3, on peut choisir pour tout $n \in \mathbb{N}^*$ un ensemble $E_n \in \mathfrak{E}_{1/n}$ tel que $m(\Omega\backslash E_n) = 0$. Posons alors

$$D := \bigcup_{n=1}^{\infty} (\Omega\backslash E_n) = \Omega\backslash \bigcap_{n=1}^{\infty} E_n.$$

Par définition de D, on a $m(D) = 0$. De plus,

$$\nu(\Omega\backslash D) \leqslant \nu(E_n) < \frac{1}{n}\, m(E_n) \leqslant \frac{1}{n}\, m(\Omega) \quad \text{pour tout } n \geqslant 1,$$

et donc $\nu(\Omega\backslash D) = 0$ car la mesure m est *finie*. Comme la mesure ν est $\geqslant 0$, on a donc $\nu(B) = 0$ pour tout $B \subseteq \Omega\backslash D$, et donc $\nu \perp m$ puisque $m(D) = 0$.

CAS 2. La mesure m est seulement sigma-finie.

On se ramène au Cas 1 grâce au fait suivant.

FAIT 2.3. *Il existe une mesure positive* finie m' *telle que* $m' \ll m$ *et* $m \ll m'$.

Démonstration. Comme la mesure m est sigma-finie, on peut trouver (exo) une suite $(E_k)_{k\geqslant 1}$ d'ensembles mesurables deux à deux disjoints tels que $m(E_k) < \infty$ pour tout k et $\Omega = \bigcup_{k=1}^{\infty} E_k$. En supposant $m(E_k) > 0$ pour tout k, posons alors

$$w := \sum_{k=1}^{\infty} \frac{2^{-k}}{m(E_k)} \mathbf{1}_{E_k}.$$

La fonction w est une fonction mesurable positive bien définie, et $w(x) > 0$ pour tout $x \in \Omega$. On a $\int_\Omega w\, dm = \sum_{k=1}^{\infty} 2^{-k} = 1$, donc la mesure $m' := wm$ est finie ; et $m' \ll m$ par définition. Inversement, si $A \in \mathfrak{B}$ vérifie $m'(A) = \int_A w\, dm = 0$, alors $m(A) = 0$ car $w(x) > 0$ pour tout $x \in A$; donc $m \ll m'$. $\square$

Si on applique le Cas 1 à la mesure m' donnée par le Fait 2.3, on voit qu'on peut écrire $\mu = \lambda + \nu$ où $\lambda \ll m'$ et $\nu \perp m'$. Mais alors $\lambda \ll m$ car $m' \ll m$; et $\nu \perp m$ car $m \ll m'$. $\square$

COROLLAIRE 2.4. (Théorème de Radon-Nikodym)
Soit m une mesure positive sigma-finie sur $(\Omega, \mathfrak{B})$, et soit μ une mesure complexe. Si $\mu \ll m$, alors $\mu = hm$ pour une certaine $h \in L^1(\Omega, m)$.

Démonstration. C'est évident par le théorème. $\square$

COROLLAIRE 2.5. *Une variable aléatoire réelle X est "à densité" si et seulement si sa loi $\mathbb{P}_X$ est absolument continue par rapport à la mesure de Lebesgue.*

COROLLAIRE 2.6. *Si μ est une mesure complexe sur $(\Omega, \mathfrak{B})$, alors $\mu = h|\mu|$ pour une certaine fonction mesurable $h : \Omega \to \mathbb{C}$ telle que $|h(x)| \equiv 1$.*

Démonstration. Comme $|\mu(A)| \leqslant |\mu|(A)$ pour tout $A \in \mathfrak{B}$, on a certainement $\mu \ll |\mu|$. Donc $\mu = h|\mu|$ pour une certaine $h \in L^1(\Omega, |\mu|)$. Alors $|\mu| = |h||\mu|$; donc $|h| = 1$ au sens de $L^1(|\mu|)$; et quitte à redéfinir h sur un ensemble $|\mu|$-négligeable, on peut supposer que $|h(x)| = 1$ pour tout $x \in \Omega$. $\square$

Corollaire 2.7. *Si μ est une mesure réelle sur $(\Omega, \mathfrak{B})$, alors il existe une partition mesurable (E^+, E^-) de Ω telle que $\forall A \in \mathfrak{B}\ :\ \mu^+(A) = \mu(A \cap E^+) = |\mu|(A \cap E^+)$ et $\mu^-(A) = -\mu(A \cap E^-) = |\mu|(A \cap E^-)$. En particulier, μ^+ est portée par E^+ et μ^- est portée par E^-, et donc $\mu^+ \perp \mu^-$.*

Démonstration. Comme μ est réelle, on peut écrire $\mu = h|\mu|$ pour une certaine fonction mesurable $h : \Omega \to \mathbb{R}$ telle que $\forall x \in \Omega\ :\ h(x) = \pm 1$. Posons $E^+ := \{h = 1\}$ et $E^- := \{h = -1\}$. Alors (E^+, E^-) est une partition mesurable de Ω. On a $\mu^+ = \frac{1}{2}(|\mu| + \mu) = \frac{1}{2}(1 + h)|\mu| = \mathbf{1}_{E^+}|\mu|$, donc $\forall A \in \mathfrak{B}\ :\ \mu^+(A) = |\mu|(A \cap E^+)$; et comme $h \equiv 1$ sur E^+, on a aussi $\mu(A \cap E^+) = \int_{A \cap E^+} h\, d|\mu| = |\mu|(A \cap E^+)$. Même type de raisonnement pour μ^-. □

Corollaire 2.8. *L'espace $M(\Omega)$ est complet pour la norme $\| \cdot \|_M$.*

Démonstration. La preuve repose sur le fait suivant.

Fait. Si $(\mu_n)_{n \in \mathbb{N}}$ est une suite de mesures complexes sur Ω, alors il existe une mesure positive finie m telle que $\forall n \in \mathbb{N}\ :\ \mu_n \ll m$.

Preuve du Fait. On peut supposer que $\mu_n \neq 0$ pour tout $n \in \mathbb{N}$. On peut donc poser $m := \sum_{n=0}^{\infty} \frac{2^{-n}}{|\mu_n|(\Omega)}\, |\mu_n|$, et on vérifie que m convient. □

Remarque. Si on veut juste une mesure m sigma-finie (ce qui suffit pour la suite) il suffit de poser $m := \sum_{n=0}^{\infty} |\mu_n|$.

Soit maintenant $(\mu_n)_{n \in \mathbb{N}}$ une suite de Cauchy dans $M(\Omega)$. Par le Fait et le Théorème de Radon-Nikodym, on peut trouver une mesure positive finie m et une suite $(h_n) \subseteq L^1(\Omega, m)$ telles que $\forall n \in \mathbb{N}\ :\ \mu_n = h_n m$. Alors $\|h_q - h_p\|_1 = \|\mu_q - \mu_p\|_M$ pour tous $p, q \in \mathbb{N}$; donc la suite (h_n) est de Cauchy dans $L^1(\Omega, m)$, et donc elle converge dans $L^1(\Omega, m)$ vers une fonction h puisque $L^1(\Omega, m)$ est complet. Donc $\mu_n \to \mu := hm$ pour la norme de $M(\Omega)$. □

3. Formes linéaires continues sur L^p

Dans cette section $(\Omega, \mathfrak{B}, m)$ est un espace mesuré fixé une fois pour toutes. Bien que ce ne soit pas strictement indispensable, on suppose que

la mesure m est sigma-finie.

Pour $1 \leqslant p \leqslant \infty$, on pose $L^p := L^p(\Omega, m)$. Si $p < \infty$, on pose $\|f\|_p := \left(\int_\Omega |f|^p dm\right)^{1/p}$ pour toute fonction mesurable $f : \Omega \to \mathbb{C}$ (même si $f \notin L^p$).

Rappel. Soit $p, q \in [1, \infty]$ des exposants conjugués ($\frac{1}{p} + \frac{1}{q} = 1$). Pour toutes fonctions mesurables $f, g : \Omega \to \mathbb{C}$, on a $\|fg\|_1 \leqslant \|f\|_p\, \|g\|_q$. En particulier : si $f \in L^p$ et $g \in L^q$, alors $fg \in L^1$.

Démonstration. Hölder. □

Lemme 3.1. *Soit $p \in [1, \infty]$, et soit q l'exposant conjugué. Pour $g \in L^q$, on note $\Phi_g : L^p \to \mathbb{C}$ la forme linéaire définie par $\Phi_g(f) := \int_\Omega fg\, dm$. Alors Φ_g est continue et $\|\Phi_g\| = \|g\|_q$.*

Démonstration. Par Hölder, Φ_g est continue et $\|\Phi_g\| \leqslant \|g\|_q$. Il reste à montrer l'inégalité inverse $\|\Phi_g\| \geqslant \|g\|_q$. Bien entendu, on peut supposer que $g \neq 0$.

Dans ce qui suit, on fixe une fonction mesurable $\theta : \Omega \to \mathbb{C}$ telle que

$$|\theta(x)| \equiv 1 \qquad \text{et} \qquad \theta g = |g|.$$

Supposons que $p = 1$, et donc $q = \infty$. Soit $\alpha < \|g\|_\infty$. Par définition de la norme $\|\cdot\|_\infty$, l'ensemble $E_\alpha := \{|g| > \alpha\}$ vérifie $m(E_\alpha) > 0$. Comme la mesure m est sigma-finie, on peut donc trouver un ensemble mesurable $E \subseteq E_\alpha$ tel que $0 < m(E) < \infty$. Alors $f := \mathbf{1}_E\theta \in L^1$ et $f \neq 0$. On a $\|f\|_1 = \int_E |\theta|\, dm = m(E)$, et $\Phi_g(f) = \int_E \theta g\, dm = \int_E |g|\, dm \geqslant \alpha m(E) = \alpha\|f\|_1$. Donc $\|\Phi\| \geqslant \alpha$; et comme $\alpha < \|g\|_\infty$ est quelconque, on en déduit $\|\Phi\| \geqslant \|g\|_\infty$.

Supposons maintenant que $p = \infty$, et donc $q = 1$. Alors $f := \theta \in L^\infty$ avec $\|f\|_\infty = 1$, et $\Phi_g(f) = \int_\Omega \theta g\, dm = \int_\Omega |g|\, dm = \|g\|_1$; donc $\|\Phi_g\| \geqslant \|g\|_1$.

Supposons enfin que $1 < p, q < \infty$. Posons $f := \theta|g|^{q-1}$. Comme $p + q = pq$ et $|\theta| \equiv 1$, on a $|f|^p = |g|^{p(q-1)} = |g|^q$; donc $f \in L^p$ et $\|f\|_p = \|g\|_q^{q/p}$. Par ailleurs, comme $\theta g = |g|$, on a aussi $fg = g\theta|g|^{q-1} = |g|^q$; donc $\Phi_g(f) = \int_\Omega |g|^q dm = \|g\|_q^q$. Par conséquent, $\|\Phi_g\| \geqslant \frac{\Phi_g(f)}{\|f\|_p} = \|g\|_q^{q-\frac{q}{p}} = \|g\|_q$. $\square$

REMARQUE. Le fait que la mesure m soit sigma-finie a servi uniquement pour le cas $p = 1$.

COROLLAIRE 3.2. *Si $g_1, g_2 \in L^q$ et si $\Phi_{g_1} = \Phi_{g_2}$, alors $g_1 = g_2$.*

Démonstration. L'application $g \mapsto \Phi_g$ étant linéaire, on a $\Phi_{g_2 - g_1} = 0$, et donc $\|g_2 - g_1\|_q = 0$. $\square$

COROLLAIRE 3.3. ("principe de dualité" dans L^p)
Soit $1 \leqslant p \leqslant \infty$ d'exposant conjugué q, et soit $f : \Omega \to \mathbb{C}$ une fonction mesurable. Soit également $C \in \mathbb{R}^+$. On suppose que

$$\forall g \in L^q \text{ étagée } : \int_\Omega |fg|\, dm \leqslant C\, \|g\|_q.$$

Alors on peut conclure que $f \in L^p$ et $\|f\|_p \leqslant C$.

Démonstration. Supposons d'abord que $p < \infty$. Comme la mesure m est sigma-finie, on peut trouver une suite croissante d'ensembles mesurables $(\Omega_n)_{n \geqslant 0}$ telle que $m(\Omega_n) < \infty$ pour tout n et $\bigcup_{n=0}^\infty \Omega_n = \Omega$. Pour $n \in \mathbb{N}$, posons $f_n := |f|\mathbf{1}_{E_n}$, où $E_n := \Omega_n \cap \{|f| \leqslant n\}$. Alors $f_n \geqslant 0$, la suite (f_n) est croissante et $f_n(x) \to |f(x)|$ pour tout $x \in \Omega$. Donc $\|f_n\|_p \to \|f\|_p$ par convergence monotone. Par ailleurs, on a $f_n \in L^p$ pour tout $n \in \mathbb{N}$, car $|f_n| \leqslant n\mathbf{1}_{\Omega_n}$ et $m(\Omega_n) < \infty$. De plus, on a $\int_\Omega |f_n g|\, dm \leqslant \int_\Omega |fg|\, dm \leqslant C\, \|g\|_q$ pour toute $g \in L^q$ étagée. Comme les fonctions étagées sont denses dans L^q, on en déduit que la forme linéaire continue $\Phi_{f_n} : L^q \to \mathbb{C}$ vérifie $\|\Phi_{f_n}\| \leqslant C$. Par le lemme, on a donc $\|f_n\|_p \leqslant C$, pour tout $n \in \mathbb{N}$. Donc $\|f\|_p = \lim_{n\to\infty} \|f_n\|_p \leqslant C$.
Si $p = \infty$, la preuve est la même en considérant les fonctions $f_n := |f|\, \mathbf{1}_{\{|f| \leqslant n\}}$ (on n'a pas besoin des Ω_n). On a $f_n \in L^\infty$ pour tout $n \in \mathbb{N}$, et $\|f_n\|_\infty \leqslant C$ par le lemme. Comme $f_n(x) \to |f(x)|$ pour tout $x \in \Omega$, on en déduit que $|f(x)| \leqslant C$ pp, *i.e.* $f \in L^\infty$ et $\|f\|_\infty \leqslant C$. $\square$

Corollaire 3.4. (Inégalité de Minkowski intégrale)
Soit $f:\Omega\to\mathbb{C}$ une fonction mesurable. On suppose que f est de la forme $f(x)=\int_\Lambda F_t(x)\,d\nu(t)$, où $(\Lambda,\mathfrak{T},\nu)$ est un autre espace mesuré sigma-fini et la fonction $(x,t)\mapsto F_t(x)$ est mesurable. Pour tout $1\leqslant p<\infty$, on a $\|f\|_p\leqslant\int_\Lambda\|F_t\|_p\,d\nu(t)$.

Démonstration. On peut évidemment supposer que $C:=\int_\Lambda\|F_t\|_p\,d\nu(t)<\infty$. D'après le principe de dualité, il suffit de montrer que

$$\forall g\in L^q\ :\ \int_\Omega|fg|\,dm\leqslant C\,\|g\|_q.$$

C'est une majoration "automatique" : par Fubini+Hölder,

$$\begin{aligned}\int_\Omega|fg|\,dm&=\int_\Omega|g(x)|\left|\int_\Lambda F_t(x)\,d\nu(t)\right|dm(x)\\&\leqslant\int_\Lambda\left(\int_\Omega|F_t(x)g(x)|\,dm(x)\right)d\nu(t)\\&\leqslant\int_\Lambda\|F_t\|_p\|g\|_q\,d\nu(t)=C\,\|g\|_q.\end{aligned}$$

□

Théorème 3.5. *Si $1\leqslant p<\infty$, alors toute forme linéaire continue sur L^p est du type Φ_g. Autrement dit, si $\Phi:L^p\to\mathbb{C}$ est une forme linéaire continue, alors il existe une unique $g\in L^q$ telle que*

$$\forall f\in L^p\ :\ \Phi(f)=\int_\Omega fg\,dm.$$

Démonstration. L'unicité découle du Lemme 3.1. Pour l'existence, on va distinguer 2 cas.

Cas 1. La mesure m est finie.

Dans ce cas, on a $\mathbf{1}_A\in L^p$ pour tout $A\in\mathfrak{B}$. On peut donc poser

$$\mu(A):=\Phi(\mathbf{1}_A).$$

La fonction μ est une mesure complexe sur $(\Omega,\mathfrak{B})$. En effet, si $(A_k)_{k\in\mathbb{N}}$ est une suite d'ensembles mesurables deux à deux disjoints, alors

$$\mathbf{1}_{\bigcup_{k=0}^\infty A_k}=\sum_{k=0}^\infty\mathbf{1}_{A_k},$$

où la série converge dans L^p car $p<\infty$ (exo). Comme $\Phi\in(L^p)^*$, on a donc

$$\mu\left(\bigcup_{k=0}^\infty A_k\right)=\Phi\left(\mathbf{1}_{\bigcup_{k=0}^\infty A_k}\right)=\sum_{k=0}^\infty\Phi(\mathbf{1}_{A_k})=\sum_{k=0}^\infty\mu(A_k).$$

De plus, la mesure μ est *absolument continue* par rapport à m : si $m(A)=0$, alors $\mathbf{1}_A=0$ dans $L^p=L^p(\Omega,m)$, donc $\mu(A)=\Phi(\mathbf{1}_A)=0$.
Par le Théorème de Radon-Nikodym, on a donc $\mu=gm$ pour une certaine $g\in L^1(\Omega,m)$. Ainsi,

$$\Phi(\mathbf{1}_A)=\int_A g\,dm=\int_\Omega\mathbf{1}_Ag\,dm\qquad\text{pour tout }A\in\mathfrak{B};$$

et donc, par linéarité :

$$\Phi(f) = \int_\Omega fg\,dm \qquad \text{pour toute fonction } f \textit{ étagée}. \tag{3.1}$$

On en déduit en particulier qu'on a $\left|\int_\Omega fg\,dm\right| \leqslant \|\Phi\|\,\|f\|_p$ pour toute fonction f étagée ; donc, d'après le "principe de dualité" (Corollaire 3.3 où on a échangé p et q), on peut affirmer que $g \in L^q$. La forme linéaire Φ_g est donc bien définie et continue sur L^p. Par (3.1), on a $\Phi(f) = \Phi_g(f)$ pour toute fonction f étagée ; donc $\Phi = \Phi_g$ car les fonctions étagées sont denses dans L^p. Ceci termine la preuve dans le cas où la mesure m est finie.

Cas 2. La mesure m est seulement sigma-finie.

Soit $(\Omega_n)_{n\in\mathbb{N}}$ une suite croissante d'ensembles mesurables telle que $\bigcup_{n=0}^\infty \Omega_n = \Omega$. Pour $n \in \mathbb{N}$, notons $L^p(\Omega_n)$ l'espace $L^p\big(\Omega_n, m_{|\Omega_n}\big)$. On considère $L^p(\Omega_n)$ comme un sous-espace de $L^p = L^p(\Omega, m)$ en prolongeant par 0 les éléments de $L^p(\Omega_n)$; autrement dit, on identifie $L^p(\Omega_n)$ à $\{f \in L^p;\ f = 0 \text{ pp sur } \Omega\backslash\Omega_n\}$. Pour tout $n \in \mathbb{N}$, la restriction de Φ à $L^p(\Omega_n)$ est une forme linéaire continue de norme $\leqslant \|\Phi\|$. D'après la Cas 1, il existe donc une unique $g_n \in L^q(\Omega_n)$ telle que $\forall f \in L^p(\Omega_n)\ :\ \Phi(f) = \int_{\Omega_n} fg\,dm = \int_\Omega fg\,dm$, et on a $\|g_n\|_{L^q(\Omega_n)} \leqslant \|\Phi\|$. Par unicité, on a $g_m = g_n$ pp sur Ω_n si $n \leqslant m$. Donc il existe une fonction mesurable pp définie $g : \Omega \to \mathbb{C}$ telle que $\forall n \in \mathbb{N}\ :\ g = g_n$ pp sur Ω_n. On a alors $\Phi(f) = \int_\Omega fg\,dm$ pour toute $f \in \bigcup_{n\in\mathbb{N}} L^p(\Omega_n)$. De plus, on a $\|\mathbf{1}_{\Omega_n} g\|_q = \|g_n\|_{L^q(\Omega_n)} \leqslant \|\Phi\|$ pour tout $n \in \mathbb{N}$, donc $g \in L^q$ et $\|g\|_q \leqslant \|\Phi\|$ (exo). Comme $\bigcup_{n\in\mathbb{N}} L^p(\Omega_n)$ est dense dans L^p car $p < \infty$ (autre exo), on en déduit que $\Phi = \Phi_g$. $\square$

Remarque. On peut montrer que si $1 < p < \infty$, alors on n'a en fait pas besoin de supposer que la mesure m est sigma-finie.

Corollaire 3.6. *Si $1 \leqslant p < \infty$, alors $(L^p)^*$ s'identifie isométriquement à L^q, où q est l'exposant conjugué de p.*

Remarque. "En général", le résultat est faux pour $p = \infty$. Plus précisément, s'il existe une suite $(E_n)_{n\in\mathbb{N}}$ d'ensembles mesurables deux à deux disjoints tels que $m(E_n) > 0$ pour tout $n \in \mathbb{N}$, alors L^1 n'est pas le dual de L^∞, *i.e.* il existe des formes linéaires continues sur L^∞ qui ne sont pas de la forme Φ_g avec $g \in L^1$.

Démonstration. Pour toute suite $a \in \ell^\infty(\mathbb{N})$, posons

$$f_a := \sum_{n=0}^\infty a_n \mathbf{1}_{E_n} \in L^\infty.$$

L'application $a \mapsto f_a$ est linéaire, et comme $m(E_n) > 0$ pour tout $n \in \mathbb{N}$, on a $\|f_a\|_\infty = \|a\|_{\ell^\infty}$. Donc l'application $a \mapsto f_a$ est une isométrie, et en particulier cette application est injective.

Soit maintenant $\mathsf{c} \subseteq \ell^\infty(\mathbb{N})$ le sous-espace vectoriel constitué par toutes les suites a *admettant une limite dans* $\mathbb{C}$, et soit $Z := \{f_a;\ a \in \mathsf{c}\} \subseteq L^\infty$; autrement dit, Z est le sous-espace vectoriel de L^∞ constitué par toutes les fonctions $f \in L^\infty$ s'écrivant sous la forme

$$f = \sum_{n=0}^\infty a_n \mathbf{1}_{E_n}\,, \qquad \text{avec } a = (a_n) \in \mathsf{c}.$$

Fait. Il existe une forme linéaire continue $\Phi : L^\infty \to \mathbb{C}$ telle que

$$\forall f = \sum_{n=0}^{\infty} a_n \mathbf{1}_{E_n} \in Z \ : \ \Phi(f) = \lim_{n\to\infty} a_n.$$

Preuve du Fait. Si $f = f_a \in Z$, on peut poser $\varphi(f) := \lim a_n$ car a est uniquement déterminée par f. Alors $|\varphi(f)| \leqslant \sup_{n\in\mathbb{N}} |a_n| = \|f\|_\infty$; donc φ est une forme linéaire continue sur $(Z, \|\cdot\|_\infty)$, d'où le résultat par le Théorème de Hahn-Banach. □

Montrons que la forme linéaire Φ donnée par le Fait n'est pas du type Φ_g. Supposons qu'il existe $g \in L^1$ telle que

$$\forall f = \sum_{n=0}^{\infty} a_n \mathbf{1}_{E_n} \in Z \ : \ \int_\Omega fg\, dm = \lim_{n\to\infty} a_n.$$

Alors

$$\int_{E_n} g\, dm = \Phi(\mathbf{1}_{E_n}) = 0 \qquad \text{pour tout } n \in \mathbb{N}.$$

Mais par ailleurs, comme $g \in L^1$, on a par convergence dominée :

$$\int_\Omega fg\, dm = \sum_{n=0}^{\infty} a_n \int_{E_n} g\, dm \qquad \text{pour toute } f = \sum_{n=0}^{\infty} a_n \mathbf{1}_{E_n} \in Z.$$

Donc $\int_\Omega fg\, dm = 0$ pour toute $f \in Z$, autrement dit $\forall f \in Z \ : \ \Phi(f) = 0$; mais ceci est faux car

$$\Phi\left(\sum_{n=0}^{\infty} \mathbf{1}_{E_n}\right) = \lim_{n\to\infty} 1 = 1.$$

□

4. Formes linéaires continues sur $\mathcal{C}(\Omega)$

Dans cette section, Ω est un espace topologique *compact métrisable.* On note $\mathcal{C}(\Omega)$ l'espace des fonctions continues $f : \Omega \to \mathbb{C}$ muni de la norme $\|\cdot\|_\infty$. De plus, on note $\mathcal{C}_\mathbb{R}(\Omega)$ l'ensemble des $f \in \mathcal{C}(\Omega)$ à valeurs réelles, et $\mathcal{C}_+(\Omega)$ l'ensemble des $f \in \mathcal{C}(\Omega)$ à valeurs réelles $\geqslant 0$. On note aussi $M(\Omega)$ l'ensemble des mesures boréliennes complexes sur Ω, $M_\mathbb{R}(\Omega)$ l'ensemble des mesures réelles, et $M_+(\Omega)$ l'ensemble des mesures positives finies.

Lemme 4.1. *Soit $\mu \in M(\Omega)$. Notons $L_\mu : \mathcal{C}(\Omega) \to \mathbb{C}$ la forme linéaire définie par*

$$L_\mu(f) := \int_\Omega f\, d\mu.$$

Alors L_μ est continue et $\|L_\mu\| = \|\mu\|_M = |\mu|(\Omega)$.

Démonstration. Écrivons $\mu = h|\mu|$, où $|h(x)| \equiv 1$. Si $f \in \mathcal{C}(\Omega)$, alors $L_\mu(f) = \int_\Omega fh\, d|\mu|$, donc $|L_\mu(f)| \leqslant \int_\Omega |fh|\, d|\mu| = \int_\Omega |f|\, d|\mu| \leqslant \|f\|_\infty\, |\mu|(\Omega)$. Donc L_μ est continue et $\|L_\mu\| \leqslant |\mu|(\Omega)$.
L'inégalité inverse $\|L_\mu\| \geqslant |\mu|(\Omega)$ est plus délicate. Le point clé est le fait suivant.

Fait 4.2. *Soit $m \in M_+(\Omega)$. Si $\theta : \Omega \to \mathbb{C}$ est une fonction mesurable bornée, alors, pour tout $\varepsilon > 0$, on peut trouver une fonction $f \in \mathcal{C}(\Omega)$ telle que $\|f\|_\infty \leqslant \|\theta\|_\infty$ et $m\big(\{|f-\theta| \geqslant \varepsilon\}\big) < \varepsilon$.*

Preuve du Fait. Comme θ est (mesurable et) bornée on peut trouver une fonction étagée $\widetilde{\theta}$ telle que $\|\widetilde{\theta}-\theta\|_\infty < \varepsilon$ et $|\widetilde{\theta}(x)| \leqslant |\theta(x)|$ pour tout $x \in \Omega$. Écrivons $\widetilde{\theta} = \sum_{i=1}^N c_i \mathbf{1}_{A_i}$, où les A_i forment une partition mesurable de Ω. Soit également $\eta > 0$ à choisir. Comme Ω est métrisable, il est “bien connu” que la mesure m est **régulière** : pour tout borélien $A \subseteq \Omega$, on peut trouver un fermé E tel que $E \subseteq A$ et $m(A\backslash E) < \eta$. On peut donc trouver, pour $i = 1, \dots, N$, un compact $E_i \subseteq A_i$ tel que $m(A_i\backslash E_i) < \eta$. Soit $E := \bigcup_{i=1}^N E_i$. Comme les E_i sont des fermés de Ω deux à deux disjoints, la fonction $\phi := \widetilde{\theta}_{|E}$ (qui est constante sur chaque E_i) est continue sur E. D'après le *Théorème de Tietze-Urysohn*, on peut prolonger ϕ en une fonction f continue sur Ω telle que $\|f\|_\infty = \|\phi\|_\infty$. Autrement dit, on peut trouver une fonction $f \in \mathcal{C}(\Omega)$ telle que $f(x) \equiv c_i$ sur E_i pour $i = 1, \dots, N$ et $\|f\|_\infty \leqslant \max_{1\leqslant i\leqslant N} |c_i| = \|\widetilde{\theta}\|_\infty$. Alors $\|f\|_\infty \leqslant \|\theta\|_\infty$, et comme $f \equiv \widetilde{\theta}$ sur $E = \bigcup_{i=1}^N E_i$, on a $|f-\theta| < \varepsilon$ sur E. Donc $m\big(\{|f-\theta| \geqslant \varepsilon\}\big) \leqslant m(\Omega\backslash E) = \sum_{i=1}^N m(A_i\backslash E_i) < N\eta < \varepsilon$ si on prend $\eta < \varepsilon/N$. □

Soit $\varepsilon > 0$. On applique le Fait avec $m := |\mu|$ et $\theta := \bar{h} = \frac{1}{h}$, où $h : \Omega \to \mathbb{C}$ est telle que $|h(x)| \equiv 1$ et $\mu = h|\mu|$: cela fournit $f \in \mathcal{C}(\Omega)$ telle que $\|f\|_\infty \leqslant 1$ et $|\mu|\big(\{|f-\theta| \geqslant \varepsilon\}\big) < \varepsilon$. Alors

$$\begin{aligned}
|L_\mu(f)| = \left|\int_\Omega fh\, d|\mu|\right| &\geqslant \left|\int_\Omega \theta h\, d|\mu|\right| - \left|\int_\Omega (f-\theta)\, d|\mu|\right| \\
&= |\mu|(\Omega) - \left|\int_\Omega (f-\theta)\, d|\mu|\right| \\
&\geqslant |\mu|(\Omega) - \int_\Omega |f-\theta|\, d|\mu| \\
&\geqslant |\mu|(\Omega) - \big(\varepsilon|\mu|(\Omega) + 2\varepsilon\big).
\end{aligned}$$

Comme $\varepsilon > 0$ est arbitraire, on en déduit $\|L_\mu\| \geqslant |\mu|(\Omega)$. □

COROLLAIRE 4.3. *Si $\mu_1, \mu_2 \in M(\Omega)$ et si $L_{\mu_1} = L_{\mu_2}$, alors $\mu_1 = \mu_2$.*

DÉFINITION 4.4. *Soit $L : \mathcal{C}(\Omega) \to \mathbb{C}$ une forme linéaire. On dit que L est **réelle** si on a $L(f) \in \mathbb{R}$ pour toute $f \in \mathcal{C}_\mathbb{R}(\Omega)$; et on dit que L est **positive** si on a $L(f) \geqslant 0$ pour toute $f \in \mathcal{C}_+(\Omega)$.*

Remarque 1. Toute forme linéaire positive est réelle.

Démonstration. C'est clair puisque $\mathcal{C}_\mathbb{R}(\Omega) = \mathcal{C}_+(\Omega) - \mathcal{C}_+(\Omega)$. □

Remarque 2. Une forme linéaire $L : \mathcal{C}(\Omega) \to \mathbb{C}$ est réelle si et seulement si $L(\bar{f}) = \overline{L(f)}$ pour toute $f \in \mathcal{C}(\Omega)$.

Démonstration. Exo. □

EXEMPLE. Soit $\mu \in M(\Omega)$. Alors μ est réelle si et seulement si la forme linéaire L_μ est réelle ; et μ est $\geqslant 0$ si et seulement si L_μ est positive.

Démonstration. Il est clair que si μ est réelle alors L_μ est réelle, et que si μ est $\geqslant 0$ alors L_μ est positive.
Écrivons $\mu = \mu_1 + i\mu_2$, avec μ_1 et μ_2 réelles. Si L_μ est réelle, alors $\forall f \in \mathcal{C}_\mathbb{R}(\Omega)$: $L_{\mu_2}(f) = \mathrm{Im}\big(L_\mu(f)\big) = 0$; donc $L_{\mu_2} = 0$ par linéarité de L_{μ_2}, *et donc* $\mu_2 = 0$ par le Corollaire 4.3. On montre de même que si L_μ est positive, alors $\mu \geqslant 0$. □

Lemme 4.5. *Soit $L : \mathcal{C}(\Omega) \to \mathbb{C}$ une forme linéaire. Alors L est positive si et seulement si elle est continue avec $\|L\| = L(\mathbf{1})$.*

Démonstration. Supposons L positive. Alors L est en particulier réelle, donc $L(\bar f) = \overline{L(f)}$ pour toute $f \in \mathcal{C}(\Omega)$, ce qui entraine que $\forall f \in \mathcal{C}(\Omega)\ :\ L(\mathrm{Re}(f)) = \mathrm{Re}\big(L(f)\big)$. Soit $f \in \mathcal{C}(\Omega)$ quelconque, et choisissons $\omega \in \mathbb{C}$ tel que $|\omega| = 1$ et $|L(f)| = \omega L(f)$. Alors $|L(f)| = \mathrm{Re}\big(L(\omega f)\big)$ car $|L(f)| \in \mathbb{R}$, et donc $|L(f)| = L\big(\mathrm{Re}(\omega f)\big)$ puisque L est réelle. Maintenant, la fonction $g := \|f\|_\infty \mathbf{1} - \mathrm{Re}(\omega f)$ est réelle $\geqslant 0$; donc $L(g) \geqslant 0$, *i.e.* $\|f\|_\infty L(\mathbf{1}) \geqslant L\big(\mathrm{Re}(\omega f)\big) = |L(f)|$. Donc L est continue et $\|L\| \leqslant L(\mathbf{1})$. Mais $|L(\mathbf{1})| \leqslant \|L\|\,\|\mathbf{1}\|_\infty = \|L\|$, donc $\|L\| = L(\mathbf{1})$.

Inversement, supposons que L soit continue avec $\|L\| = L(\mathbf{1})$, et montrons que L est positive. On peut supposer que $\|L\| = 1 = L(\mathbf{1})$. Il suffit de montrer que si $f \in \mathcal{C}(\Omega)$ vérifie $0 \leqslant f \leqslant 1$, alors $L(f) \in [0,1]$. Supposons que ce ne soit pas le cas. Alors on peut trouver un disque ouvert $D = D(a,r) \subseteq \mathbb{C}$ tel que $[0,1] \subseteq D$ et $L(f) \notin D$ (faire un dessin). Comme $f(\Omega) \subseteq [0,1] \subseteq D$, la fonction $f - a\mathbf{1} \in \mathcal{C}(\Omega)$ vérifie $\|f - a\mathbf{1}\|_\infty < r$. Donc $|L(f - a\mathbf{1})| < r$ puisque $\|L\| = 1$. Mais comme $L(\mathbf{1}) = 1$, on a $L(f - a\mathbf{1}) = L(f) - a$, et donc $|L(f - a\mathbf{1})| \geqslant r$ puisque $L(f) \notin D$; d'où une contradiction.

Autre preuve. Supposons comme plus haut que $\|L\| = 1 = L(\mathbf{1})$. Montrons d'abord que L est réelle, *i.e.* que $L(f) \in \mathbb{R}$ pour toute $f \in \mathcal{C}_\mathbb{R}(\Omega)$. Écrivons $L(f) = a + ib$, et supposons par exemple $b \geqslant 0$. Supposons également que $\|f\|_\infty = 1$. Comme f est réelle, on a $\|f + in\|_\infty^2 = \|f\|_\infty^2 + n^2 = 1 + n^2$ pour tout $n \in \mathbb{N}$. Donc $|L(f + in)|^2 \leqslant 1 + n^2$ puisque $\|L\| = 1$. Mais $L(f + in) = a + i(b + n)$, donc $|L(f + in)|^2 = a^2 + (b + n)^2 = a^2 + b^2 + 2nb + n^2$. On obtient ainsi $2nb \leqslant 1 - a^2 - b^2$ pour tout $n \in \mathbb{N}$. Comme le second membre ne dépend pas de n, on en déduit $b \leqslant 0$, et donc $b = 0$ puisqu'on suppose que $b \geqslant 0$. Montrons maintenant que L est positive. Soit $f \in \mathcal{C}_+(\Omega)$, et supposons $\|f\|_\infty \leqslant 1$. Alors $0 \leqslant f \leqslant 1$, donc $\|\mathbf{1} - f\|_\infty \leqslant 1$. Comme $\|L\| = 1 = L(\mathbf{1})$, on a donc $|L(f)| \leqslant 1$ et $|1 - L(f)| = |L(\mathbf{1} - f)| \leqslant 1$. Mais $t := L(f)$ est un nombre réel puisque L est réelle ; donc on a nécessairement $0 \leqslant t \leqslant 1$. Ainsi $L(f) \geqslant 0$ pour toute $f \in \mathcal{C}_+(\Omega)$ telle que $\|f\|_\infty \leqslant 1$; donc L est positive.

□

Lemme 4.6. *Soit $L : \mathcal{C}(\Omega) \to \mathbb{C}$ une forme linéaire.*

(1) *L s'écrit de manière unique sous la forme $L = L_1 + iL_2$, où L_1 et L_2 sont des formes linéaires réelles, qui sont continues si L est continues.*

(2) *Si L est réelle et continue, alors on peut écrire $L = L^+ - L^-$, où L^+ et L^- sont des formes linéaires positives (**décomposition de Riesz**).*

Démonstration. (1) On se convainc assez vite qu'on doit nécessairement poser

$$L_1(f) := \frac{1}{2}\Big(L(f) + \overline{L(\bar f)}\Big) \qquad \text{et} \qquad L_2(f) := \frac{1}{2i}\Big(L(f) - \overline{L(\bar f)}\Big);$$

et on vérifie que L_1 et L_2 ainsi définies conviennent.

(2) On peut supposer que $\|L\| = 1$. Pour toute $f \in \mathcal{C}_\mathbb{R}(\Omega)$, on a ainsi

$$|L(f)| \leqslant \|f\|_\infty = \|f^+\|_\infty + \|f^-\|_\infty. \tag{4.1}$$

Posons alors

$$E := \{(f,f);\ f \in \mathcal{C}_\mathbb{R}(\Omega)\} \subseteq \mathcal{C}_\mathbb{R}(\Omega) \times \mathcal{C}_\mathbb{R}(\Omega),$$

qui est un sous-espace vectoriel de $\mathcal{C}_\mathbb{R}(\Omega)\times\mathcal{C}_\mathbb{R}(\Omega)$, et soit $p:\mathcal{C}_\mathbb{R}(\Omega)\times\mathcal{C}_\mathbb{R}(\Omega)\to\mathbb{R}$ la fonction définie par

$$p(u,v) := \|u^+\|_\infty + \|v^-\|_\infty.$$

On vérifie sans difficulté que p est une fonctionnelle sous-linéaire (exo). Si on définit $\varphi : E\to\mathbb{R}$ par $\varphi(f,f) := L(f)$, alors φ est une forme $\mathbb{R}$-linéaire, et (4.1) dit que φ est majorée par p. D'après le Théorème de Hahn-Banach, on peut donc trouver une forme $\mathbb{R}$-linéaire $\Phi : \mathcal{C}_\mathbb{R}(\Omega)\times\mathcal{C}_\mathbb{R}(\Omega)\to\mathbb{R}$ telle que $\Phi_{|E}=\varphi$ et $\Phi\leqslant p$. Autrement dit, on peut trouver deux formes $\mathbb{R}$-linéaires $\Phi_1,\Phi_2 : \mathcal{C}_\mathbb{R}(\Omega)\to\mathbb{R}$ telles que

$$\forall f\in\mathcal{C}_\mathbb{R}(\Omega)\;:\;\Phi_1(f)+\Phi_2(f)=L(f)\qquad\text{et}$$

$$\forall (u,v)\in\mathcal{C}_\mathbb{R}(\Omega)\times\mathcal{C}_\mathbb{R}(\Omega)\;:\;\Phi_1(u)+\Phi_2(v)\leqslant\|u^+\|_\infty+\|v^-\|_\infty.$$

En prenant $v:=0$, on voit que $\Phi(u)\leqslant\|u^+\|_\infty$ pour toute $u\in\mathcal{C}_\mathbb{R}(\Omega)$. Donc $\Phi_1(u)\leqslant 0$ pour toute $u\leqslant 0$; autrement dit, $\Phi_1(u)\geqslant 0$ pour toute $u\in\mathcal{C}_+(\Omega)$. De même, $\Phi_2(v)\leqslant 0$ pour toute $v\in\mathcal{C}_+(\Omega)$. Ainsi, Φ_1 et $-\Phi_2$ sont des formes $\mathbb{R}$-linéaire positives sur $\mathcal{C}_\mathbb{R}(\Omega)$. On peut donc prolonger Φ_1 et $-\Phi_2$ en des formes $\mathbb{C}$-linéaires positives L^+ et L^- sur $\mathcal{C}(\Omega)$ en posant $L^+(f+ig):=L_1(f)+iL_1(g)$ et $L^-(f+ig):=-\Phi_2(f)-i\Phi_2(g)$. Alors $L(f)=\Phi_1(f)+\Phi_2(f)=L^+(f)-L^-(f)$ pour toute $f\in\mathcal{C}_\mathbb{R}(\Omega)$, et donc $L=L^+-L^-$. □

Remarque. Voici une autre façon de démontrer (2). On commence par définir L^+ sur $\mathcal{C}_+(\Omega)$ en posant

$$L^+(f) := \sup\big\{L(u);\ 0\leqslant u\leqslant f\big\}.$$

Comme L est continue, le “sup” est bien un nombre réel, donc la définition a un sens ; et on a $L^+(f)\geqslant 0$ et $L^+(f)\geqslant L(f)$ pour toute $f\in\mathcal{C}_+(\Omega)$. Il est clair également qu'on a $L^+(\lambda f)=\lambda L^+(f)$ pour toute constante $\lambda\geqslant 0$. Le point clé est de montrer que L^+ est additive, *i.e.* $L^+(f+g)=L^+(f)+L^+(g)$ pour toutes $f,g\in\mathcal{C}_+(\Omega)$. L'inégalité $L^+(f)+L^+(g)\leqslant L^+(f+g)$ découle directement de la définition (exo). L'inégalité inverse se prouve en montrant que si $u\in\mathcal{C}(\Omega)$ vérifie $0\leqslant u\leqslant f+g$, alors on peut écrire $u=u_1+u_2$ où $u_1,u_2\in\mathcal{C}(\Omega)$ et $0\leqslant u_1\leqslant f$, $0\leqslant u_2\leqslant g$: il suffit d'écrire $u=f-(f-u)=\big(f-(f-u)^+\big)+(f-u)^-$ et de vérifier que cela convient.
Le reste est ensuite de la routine. On prolonge L^+ en une forme $\mathbb{R}$-linéaire sur $\mathcal{C}_\mathbb{R}(\Omega)$ en posant $L^+(f):=L^+(f^+)-L^+(f^-)$, puis en une forme $\mathbb{C}$-linéaire sur $\mathcal{C}(\Omega)$ en posant $L^+(f_1+if_2):=L^+(f_1)+iL^+(f_2)$. Alors L^+ est une forme linéaire positive, et $L^-:=L^+-L$ aussi puisque $L^+(f)\geqslant L(f)$ pour toute $f\in\mathcal{C}_+(\Omega)$ par définition.

Corollaire 4.7. *Si $L:\mathcal{C}(\Omega)\to\mathbb{C}$ est une forme linéaire continue, alors L est combinaison linéaire de 4 formes linéaires positives.*

On peut maintenant énoncer et démontrer le résultat principal de cette section, qu'on appelle souvent le **Théorème de représentation de Riesz**

Théorème 4.8. *Toute forme linéaire continue sur $\mathcal{C}(\Omega)$ est du type L_μ. Plus précisément : si $L:\mathcal{C}(\Omega)\to\mathbb{C}$ est une forme linéaire continue, alors il existe une unique $\mu\in M(\Omega)$ telle que*

$$\forall f\in\mathcal{C}(\Omega)\;:\;L(f)=\int_\Omega f\,d\mu.$$

De plus, μ est réelle si L est réelle, et μ est $\geqslant 0$ si L est positive.

Démonstration. L'unicité découle du Lemme 4.1. Pour l'existence, on va distinguer 2 cas.

Cas 1. Ω est l'espace de Cantor $\Delta := \{0,1\}^{\mathbb{N}}$.

Soit $L : \mathcal{C}(\Delta) \to \mathbb{C}$ une forme linéaire continue sur $\mathcal{C}(\Delta)$: on cherche une mesure complexe $\mu \in M(\Delta)$ telle que $L = L_\mu$. Par la Lemme 4.6, on peut supposer que la forme linéaire L est *positive* ; et on cherche alors on cherche une mesure μ positive et finie.
Notons $\mathfrak{A}$ la famille de tous les ensembles $W \subseteq \Delta$ à la fois *ouverts et fermés*. Il est clair que $\mathfrak{A}$ est une *algèbre* de parties de Δ, *i.e.* $\mathfrak{A}$ est stable par réunions finies, intersections finies et complémentation.

Fait 1. L'algèbre $\mathfrak{A}$ engendre la tribu borélienne de Δ.

Preuve du Fait 1. On a vu que $\mathfrak{A}$ est une base pour la topologie de Δ. Donc, tout ouvert $O \subseteq \Delta$ est réunion d'éléments de $\mathfrak{A}$. Comme Δ est métrisable séparable, on en déduit que tout ouvert de Δ est réunion *dénombrable* d'éléments de $\mathfrak{A}$. (En fait, il n'est pas très difficile de montrer que $\mathfrak{A}$ est dénombrable.) Donc la tribu $\sigma(\mathfrak{A})$ contient tous les ouverts. $\square$

Si $W \in \mathfrak{A}$, alors la fonction $\mathbf{1}_W$ est *continue* sur Δ, *i.e.* $\mathbf{1}_W \in \mathcal{C}(\Omega)$. On peut donc poser
$$\alpha(W) := L\big(\mathbf{1}_W\big) \qquad \text{pour tout } W \in \mathfrak{A};$$
et on a $\alpha(W) \geqslant 0$ car L est une forme linéaire positive et $\mathbf{1}_W \geqslant 0$.

Fait 2. La fonction $\alpha : \mathfrak{A} \to \mathbb{R}^+$ est une *mesure sur* $\mathfrak{A}$: si $(W_k)_{k\in\mathbb{N}}$ est une suite d'éléments de $\mathfrak{A}$ deux à deux disjoints *telle que* $\bigcup_{k=0}^\infty W_k \in \mathfrak{A}$, alors $\alpha\Big(\bigcup_{k=0}^\infty W_k\Big) = \sum_{k=0}^\infty \alpha(W_k)$.

Preuve du Fait 2. Comme $W := \bigcup_{k=0}^\infty W_k$ est fermé dans Δ, donc compact, et que les W_k sont ouverts, on peut trouver $K \in \mathbb{N}$ tel que $W = \bigcup_{k=0}^K W_k$; autrement dit, la réunion dénombrable est en fait une réunion finie. Alors $W_k = \varnothing$ pour tout $k > K$ car les W_k sont disjoints ; et donc $\alpha(W_k) = 0$ pour tout $k > K$. De plus, toujours par disjointude, on a $\mathbf{1}_{\bigcup_{k=0}^K W_k} = \sum_{k=0}^K \mathbf{1}_{W_k}$. Donc
$$\alpha(W) = L\left(\sum_{k=0}^K \mathbf{1}_{W_k}\right) = \sum_{k=0}^K L\big(\mathbf{1}_{W_k}\big) = \sum_{k=0}^K \alpha(W_k) = \sum_{k=0}^\infty \alpha(W_k).$$
$\square$

Fait 3. $\operatorname{Vect}\big\{\mathbf{1}_W;\ W \in \mathfrak{A}\big\}$ est dense dans $\mathcal{C}(\Delta)$.

Preuve du Fait 3. Posons $\mathcal{A} := \operatorname{Vect}\big\{\mathbf{1}_W;\ W \in \mathfrak{A}\big\}$. Comme $\mathfrak{A}$ est stable par intersections finies, $\mathcal{A}$ est stable par produit. Donc $\mathcal{A}$ est une sous-algèbre de $\mathcal{C}(\Delta)$, qui est clairement auto-conjuguée. De plus, l'ensemble $\{\mathbf{1}_W;\ W \in \mathfrak{A}\}$ sépare les points de Δ car $\mathfrak{A}$ est une base pour la topologie de Δ. *A fortiori*, l'algèbre $\mathcal{A}$ sépare les points de Δ. Donc on peut appliquer le Théorème de Stone-Weierstrass. $\square$

Preuve directe, sans Stone-Weierstrass. Soit $f \in \mathcal{C}(\Delta)$, et soit $\varepsilon > 0$. On cherche une fonction $\varphi \in \operatorname{Vect}\big\{\mathbf{1}_W;\ W \in \mathfrak{A}\big\}$ telle que $\|\varphi - f\|_\infty \leqslant \varepsilon$.
Comme f est continue, on peut trouver pour tout $x \in \Delta$ un voisinage V_x de x tel que $\operatorname{diam}\big(f(W_x)\big) \leqslant \varepsilon$. De plus, on peut supposer que $V_x \in \mathfrak{A}$ car $\mathfrak{A}$ est une base pour la topologie de Δ. Alors $\Delta = \bigcup_{x\in\Delta} V_x$; donc, par compacité, on peut trouver $x_1, \dots, x_N \in \Delta$ tels que $\Delta = V_{x_1} \cup \dots \cup V_{x_N}$. Si on pose $W_1 := V_{x_1}$ et $W_i := V_{x_i} \backslash (V_{x_1} \cup \dots \cup V_{x_{i-1}})$

pour $1 < i \leqslant N$, alors les W_i sont dans $\mathfrak{A}$, forment une partition de Δ (en supposant $W_i \neq \varnothing$), et $\operatorname{diam}\big(f(W_i)\big) \leqslant \varepsilon$. Si on choisit un point $w_i \in W_i$ pour $i = 1, \dots, N$, alors la fonction $\varphi := \sum_{i=1}^N f(w_i)\mathbf{1}_{W_i}$ convient : si $x \in \Delta$ est quelconque, alors $x \in W_i$ pour un unique $i \in [\![1, N]\!]$, et $|f(x) - \varphi(x)| = |f(x) - f(w_i)| \leqslant \operatorname{diam}\big(f(W_i)\big) \leqslant \varepsilon$. □

On peut maintenant conclure la preuve du Cas 1. Par les Faits 1 et 2 et le théorème “standard” de prolongement de mesures, la mesure $\alpha : \mathfrak{A} \to \mathbb{R}^+$ se prolonge en une mesure borélienne positive finie μ. Par définition de μ, on a $L(\mathbf{1}_W) = \mu(W)$ pour tout $W \in \mathfrak{A}$. Donc $L(f) = \int_\Delta f\,d\mu = L_\mu(f)$ pour toute $f \in \operatorname{Vect}\big\{\mathbf{1}_W;\ W \in \mathfrak{A}\big\}$ par linéarité ; et donc $L = L_\mu$ d'après le Fait 3 puisque L et L_μ sont continues.

Cas 2. Ω compact métrisable quelconque.

On sait qu'il existe une surjection continue $s : \Delta \twoheadrightarrow \Omega$. Alors l'appplication $J : \mathcal{C}(\Omega) \to \mathcal{C}(\Delta)$ définie par $Jf := f \circ s$ est une *isométrie* linéaire de $\mathcal{C}(\Omega)$ dans $\mathcal{C}(\Delta)$, par surjectivité de s (exo).
Soit maintenant $L : \mathcal{C}(\Omega) \to \mathbb{C}$ une forme linéaire continue. D'après ce qui précède, on définit une forme linéaire continue $\varphi : J\big(\mathcal{C}(\Omega)\big) \to \mathbb{C}$ en posant $\varphi(Jf) := L(f)$ pour toute $f \in \mathcal{C}(\Omega)$. On a donc

$$\forall f \in \mathcal{C}(\Omega)\ :\ L(f) = \varphi(f \circ s).$$

D'après le Théorème de Hahn-Banach, φ se prolonge en une forme linéaire continue $\Phi : \mathcal{C}(\Delta) \to \mathbb{C}$; et par le Cas 1, Φ est de la forme L_ν pour une certaine mesure $\nu \in M(\Delta)$. Ainsi, on a

$$\forall f \in \mathcal{C}(\Omega)\ :\ L(f) = \int_\Delta (f \circ s)\,d\nu.$$

Soit alors $\mu \in M(\Omega)$ la *mesure image* de ν par s :

$$\mu(A) := \nu\big(s^{-1}(A)\big) \qquad \text{pour tout borélien } A \subseteq \Omega.$$

Par définition, on a $\int_\Omega f\,d\mu = \int_\Delta (f \circ s)\,d\nu$ pour toute fonction borélienne étagée $f : \Omega \to \mathbb{C}$. Par approximation, ceci est encore vrai pour toute fonction borélienne bornée $f : \Omega \to \mathbb{C}$; et donc en particulier pour toute $f \in \mathcal{C}(\Omega)$. On a donc montré que $L(f) = \int_\Omega f\,d\mu$ pour toute $f \in \mathcal{C}(\Omega)$, ce qui termine la preuve du théorème. □

Corollaire 4.9. *L'espace $\big(\mathcal{C}(\Omega)\big)^*$ s'identifie isométriquement à $M(\Omega)$.*

Corollaire 4.10. *Toute forme $\mathbb{R}$-linéaire continue $L : \mathcal{C}_\mathbb{R}(\Omega) \to \mathbb{R}$ est de la forme L_μ pour une unique mesure $\mu \in M_\mathbb{R}(\Omega)$.*

Démonstration. La forme linéaire L se prolonge de manière unique en une forme $\mathbb{C}$-linéaire continue $\widetilde{L} : \mathcal{C}(\Omega) \to \mathbb{C}$ par la formule $\widetilde{L}(f_1 + if_2) := L(f_1) + iL(f_2)$. Cette forme linéaire $\widetilde{L}$ est réelle puisqu'elle prolonge L, donc elle est donnée par une mesure réelle ; d'où le résultat. □

Remarque 4.11. Le Théorème de représentation de Riesz reste vrai dans le cas d'un espace topologique compact Ω pas forcément métrisable, avec exactement le même énoncé, à un petit point de “détail” près : ce qu'on doit noter $M(\Omega)$ est l'ensemble des mesures boréliennes complexes *régulières* sur Ω. Par définition, une mesure complexe μ est dite régulière si sa variation totale $m := |\mu|$ est régulière au sens habituel, *i.e.*

pour tout borélien $A \subseteq \Omega$ et pour tout $\varepsilon > 0$, on peut trouver un fermé E et un ouvert V tels que $E \subseteq A \subseteq V$ et $m(V \backslash E) < \varepsilon$. (Il est "bien connu" que si Ω est un espace topologique métrisable, alors toute mesure borélienne positive finie m sur Ω est régulière, donc il n'y avait pas lieu d'introduire cette terminologie pour ce qui nous a occupé ; mais on a fortement utilisé la régularité quand on a montré qu'une mesure $\mu \in M(\Omega)$ est entièrement déterminée par la forme linéaire L_μ.) Il est possible de donner une preuve du Théorème de représentation de Riesz suivant le même scénario que celle faite dans le cas métrisable ; mais les "détails techniques" sont plus délicats.

5. Intégrale vectorielle

Dans cette section $(\Omega, \mathfrak{B}, m)$ est un espace mesuré et X est un espace de Banach. Une fonction $f : \Omega \to X$ est dite *mesurable* si elle est mesurable de $(\Omega, \mathfrak{B})$ dans $(X, \mathfrak{B}(X))$, où $\mathfrak{B}(X)$ est la tribu borélienne de X.

REMARQUE 5.1. Si $f : \Omega \to X$ est mesurable, alors la fonction positive $\|f\|$ est mesurable.

Démonstration. Par composition. □

NOTATIONS. On note $\mathcal{E}(\Omega, X)$ l'ensemble de toutes les fonctions mesurables $\varphi : \Omega \to X$ s'écrivant

$$\varphi = \sum_{i=1}^{N} x_i \, \mathbf{1}_{E_i},$$

où $x_1, \dots, x_N \in X$ et $m(E_i) < \infty$ pour $i = 1, \dots, N$. Pour une telle fonction φ, on note $\int_\Omega \varphi \, dm$ l'élément de X défini par

$$\int_\Omega \varphi \, dm := \sum_{i=1}^{N} m(E_i) x_i.$$

(On vérifie que ceci a bien un sens, *i.e.* que $\sum_{i=1}^N m(E_i) x_i$ ne dépend que de φ et pas de la façon d'écrire φ sous la forme $\varphi = \sum_{i=1}^N x_i \, \mathbf{1}_{E_i}$.)

FAIT 5.2. *$\mathcal{E}(\Omega, X)$ est un espace vectoriel, l'application $\varphi \mapsto \int_\Omega \varphi \, dm$ est linéaire de $\mathcal{E}(\Omega, X)$ dans X, et on a $\left\| \int_\Omega \varphi \, dm \right\| \leqslant \int_\Omega \|\varphi\| \, dm$ pour toute $\varphi \in \mathcal{E}(\Omega, X)$.*

Démonstration. Les 2 premiers points sont évidents. Pour la majoration, on écrit $\varphi = \sum_{i=1}^N u_i \, \mathbf{1}_{E_i}$ avec des E_i deux à deux disjoints : alors $\|\varphi\| = \sum_{i=1}^N \|x_i\| \, \mathbf{1}_{E_i}$, donc $\int_\Omega \|\varphi\| \, dm = \sum_{i=1}^N \|x_i\| \, m(E_i) \geqslant \left\| \int_\Omega \varphi \, dm \right\|$. □

FAIT 5.3. *Si $f : \Omega \to X$ est mesurable et si $\varphi \in \mathcal{E}(\Omega, X)$, alors $f + \varphi$ est mesurable.*

Démonstration. Il suffit de le prouver pour $\varphi = x\mathbf{1}_E$, où $x \in X$ et $E \in \mathfrak{B}$. Si B est un borélien de X et si $t \in \Omega$, alors

$$(f + \varphi)(t) \in B \iff (t \in E \text{ et } f(t) \in B - x) \quad \text{ou} \quad (t \notin E \text{ et } f(t) \in B)\,.$$

Donc $(f + \varphi)^{-1}(B) = \Big(E \cap f^{-1}(B - x) \Big) \cup \Big(E^c \cap f^{-1}(B) \Big)$, et donc $(f + \varphi)^{-1}(B)$ est mesurable. □

DÉFINITION 5.4. *On dit qu'une fonction $f : \Omega \to X$ est m-**mesurable** si f est pp égale à une fonction mesurable $\widetilde{f} : \Omega \to X$.*

Remarque. Si $f : \Omega \to X$ est m-mesurable, on pose $\int_\Omega \|f\|\, dm := \int_\Omega \|\widetilde{f}\|\, dm$, où $\widetilde{f}$ est n'importe quelle fonction mesurable telle que $f = \widetilde{f}$ pp.

DÉFINITION 5.5. *Soit $f : \Omega \to X$. On dit que f est* ***à image essentiellement séparable*** *s'il existe un sous-ensemble séparable $Z \subseteq X$ tel que $f(t) \in Z$ pp.*

Remarque. Quitte à remplacer Z par $\overline{\mathrm{vect}}\,(Z)$, on peut toujours supposer que Z est un sous-espace vectoriel fermé de X.

EXEMPLE. Si Ω est un espace topologique séparable, alors toute fonction continue $f : \Omega \to X$ est à image essentiellement séparable.

Démonstration. Soit $D \subseteq X$ un ensemble dénombrable dense. Alors $Z := \overline{f(D)}$ est un sous-ensemble séparable de X, et on a $f(\Omega) = f(\overline{D}) \subseteq Z$ par continuité de f. $\square$

LEMME 5.6. *Soit $f : \Omega \to X$ une fonction m-mesurable. On suppose que f est à image essentiellement séparable, et que $\int_\Omega \|f\|\, dm < \infty$. Alors les choses suivantes ont lieu.*

(1) *Il existe une suite $(\varphi_n)_{n\in\mathbb{N}} \subseteq \mathcal{E}(\Omega, X)$ telle que $\int_\Omega \|\varphi_n - f\|\, dm \to 0$. On dit qu'une telle suite (φ_n) est une* ***suite approximante*** *pour f.*

(2) *La limite $\lim_{n\to\infty} \int_\Omega \varphi_n\, dm$ existe dans X pour toute suite approximante (φ_n) pour f, et elle est la même pour toutes les suites approximantes.*

Démonstration. (1) Il suffit de montrer que pour tout $\varepsilon > 0$, on peut trouver une fonction $\varphi \in \mathcal{E}(\Omega, X)$ telle que $\int_\Omega \|\varphi - f\|\, dm < \varepsilon$. On fixe donc $\varepsilon > 0$.

Quitte à remplacer f par une fonction mesurable égale à f pp, on peut supposer que f est mesurable. Soit $Z \subseteq X$ un sous-espace fermé séparable tel que $f(t) \in Z$ pp. Quitte à redéfinir f en posant $f(t) := 0$ pour $t \in f^{-1}(X\backslash Z)$, on peut supposer que $f(t) \in Z$ pour tout $t \in \Omega$. Donc, quitte à remplacer X par Z, on peut supposer d'emblée que l'espace de Banach X est séparable.

Pour $k \in \mathbb{N}$, posons $\Omega_k := \{x \in \Omega;\ 2^{-k} \leqslant \|f(t)\| \leqslant k\}$. Les Ω_k sont mesurables, la suite (Ω_k) est croissante, et $\bigcup_{k=0}^\infty \Omega_k = \{t;\ f(t) \neq 0\}$. Comme la fonction $\|f\|$ est intégrable sur Ω, on en déduit que $\int_{\Omega\backslash\Omega_k} \|f\|\, dm \to 0$ quand $k \to \infty$. Donc on peut choisir $k \in \mathbb{N}$ tel que $\int_{\Omega\backslash\Omega_k} \|f\|\, dm < \varepsilon/3$. Notons que $m(\Omega_k) < \infty$ d'après l'inégalité de Markov.
Soit $\eta > 0$ à choisir. Comme X est séparable, on peut trouver une suite $(B_i)_{i\in\mathbb{N}}$ de boréliens de X telle que les B_i forment une partition de X et $\mathrm{diam}(B_i) < \eta$ pour tout $i \in \mathbb{N}$. Pour $i \in \mathbb{N}$, posons $E_i := \Omega_k \cap f^{-1}(B_i)$. Les E_i sont mesurables, deux à eux disjoints, et $\bigcup_{i\in\mathbb{N}} E_i = \Omega_k$. Comme $\int_{\Omega_k} \|f\|\, dm < \infty$, on peut donc trouver $N \in \mathbb{N}$ tel que $\int_{\Omega_k\backslash\bigcup_{i=0}^N E_i} \|f\|\, dm < \varepsilon/3$. De plus, on peut évidemment supposer que les E_i sont non vides. Choisissons un point $t_i \in E_i$ pour $i = 0, \dots, N$, et posons

$$\varphi := \sum_{i=0}^N \mathbf{1}_{E_i} f(t_i).$$

Alors $\varphi \in \mathcal{E}(\Omega, X)$; et par définition, on a $\|\varphi(t) - f(t)\| < \eta$ sur $\bigcup_{i=0}^N E_i$ et $\varphi(t) = 0$ sur $\Omega \backslash \bigcup_{i=0}^N E_i$. Donc

$$\begin{aligned}
\int_\Omega \|f - \varphi\|\, dm &= \int_{\Omega \backslash \bigcup_{i=0}^N E_i} \|f\|\, dm + \int_{\bigcup_{i=0}^N E_i} \|f - \varphi\|\, dm \\
&= \int_{\Omega\backslash\Omega_k} \|f\|\, dm + \int_{\Omega_k \backslash \bigcup_{i=0}^N E_i} \|f\|\, dm + \int_{\bigcup_{i=0}^N E_i} \|f - \varphi\|\, dm \\
&< 2\varepsilon/3 + \eta\, m(\Omega_k).
\end{aligned}$$

Donc, si on choisit η tel que $\eta\, m(\Omega_k) < \varepsilon/3$, la fonction φ convient.

(2) Si (φ_n) est une suite approximante pour f, alors la suite $\left(\int_\Omega \varphi_n\, dm\right)_{n\in\mathbb{N}}$ est de Cauchy dans X car $\left\|\int_\Omega \varphi_q\, dm - \int_\Omega \varphi_p\, dm\right\| \leqslant \int_\Omega \|\varphi_q - \varphi_p\|\, dm \leqslant \int_\Omega \|\varphi_q - f\|\, dm + \int_\Omega \|f - \varphi_p\|\, dm$ pour tous $p, q \in \mathbb{N}$. Donc $\lim_{n\to\infty} \varphi_n$ existe dans X puisque X est supposé complet. Si (ψ_n) est une autre suite approximante pour f, alors $\int_\Omega \|\varphi_n - \psi_n\|\, dm \to 0$ car $\|\varphi_n - \psi_n\| \leqslant \|\varphi_n - f\| + \|f - \psi_n\|$; donc on doit avoir $\lim \int_\Omega \varphi_n = \lim \int_\Omega \psi_n$. □

Exercice. Soit $f : \Omega \to X$ une fonction m-mesurable telle que $\int_\Omega \|f\|\, dm < \infty$. Montrer que f est à image essentiellement séparable si et seulement si on peut trouver une suite $(\varphi_n)_{n\in\mathbb{N}} \subseteq \mathcal{E}(\Omega, X)$ telle que $\varphi_n(t) \to f(t)$ pp.

Définition 5.7. *Soit $f : \Omega \to X$. On dit que f est **m-intégrable au sens de Bochner** si f est m-mesurable, à image essentiellement séparable, et si on a $\int_\Omega \|f\|\, dm < \infty$. On pose alors $\int_\Omega f\, dm = \lim \int_\Omega \varphi_n\, dm$, où (φ_n) est n'importe quelle suite approximante pour f.*

Exemple. Supposons que Ω soit un espace topologique et que m soit une mesure borélienne. Si Ω est *séparable* et si la mesure m est *finie*, alors toute fonction continue bornée $f : \Omega \to X$ est m-intégrable au sens de Bochner.

Proposition 5.8. *Notons $L^1(\Omega, m, X)$ l'ensemble des fonctions $f : \Omega \to X$ m-intégrables au sens de Bochner.*

(1) *$L^1(\Omega, m, X)$ est un espace vectoriel et l'application $f \mapsto \int_\Omega f\, dm$ est linéaire de $L^1(\Omega, m, X)$ dans X.*
(2) *Si $f \in L^1(\Omega, m, X)$, alors $\left\|\int_\Omega f\, dm\right\| \leqslant \int_\Omega \|f\|\, dm$.*
(3) *Si Y est un autre espace de Banach et si $L \in \mathcal{L}(X, Y)$, alors $L \circ f \in L^1(\Omega, m, Y)$ pour toute $f \in L^1(\Omega, m, X)$, et*
$$L\left(\int_\Omega f\, dm\right) = \int_\Omega L \circ f\, dm.$$

Démonstration. (1) On se contente de montrer que si $f, g \in L^1(\Omega, m, X)$, alors $f + g \in L^1(\Omega, m, X)$. Bizarrement, la difficulté est en fait de montrer que $f + g$ est m-mesurable.

Quitte à modifier f et g sur des ensembles de mesure 0, on peut supposer que f et g sont mesurables. Par hypothèse, on peut trouver des sous-espaces séparables $Z_f, Z_g \subseteq X$ tels que $f(x) \in E_f$ pp et $g(x) \in E_g$ pp. Alors $Z := \overline{Z_f + Z_g}$ est un sous-espace séparable de X, et on a $(f(x) \in Z$ et $g(x) \in Z)$ pp. De plus, Z est fermé dans X, donc l'ensemble $E := \{x \in \Omega;\ f(x) \in Z \text{ et } g(x) \in Z\}$ est mesurable. Donc, quitte à re-modifier f et g en dehors de E, on peut supposer que $\forall x \in \Omega\ :\ f(x) \in Z$ et $g(x) \in Z$. Le point clé est maintenant le fait suivant.

FAIT. L'application $P : \Omega \to Z \times Z$ définie par $P(x) := (f(x), g(x))$ est mesurable.

Preuve du Fait. Comme Z est un espace métrique séparable, la tribu borélienne $\mathfrak{B}(Z \times Z)$ est la tribu produit $\mathfrak{B}(Z) \otimes \mathfrak{B}(Z)$: en effet, tout ouvert de $Z \times Z$ est réunion de produits d'ouverts, donc réunion *dénombrable* de produits d'ouverts par la propriété de Lindelöf ; donc la tribu produit $\mathfrak{B}(Z) \otimes \mathfrak{B}(Z)$ contient tous les ouverts, et donc $\mathfrak{B}(Z \times Z) \subseteq \mathfrak{B}(Z) \otimes \mathfrak{B}(Z)$ (l'inclusion inverse est vraie pour tout espace topologique Z). □

Par le Fait, la fonction $f + g$ est mesurable car $f + g = S \circ P$ où $S : Z \times Z \to X$ est l'application continue définie par $S(u, v) := u + v$. De plus, $f + g$ est à valeur dans le sous-espace séparable Z, et $\int_\Omega \|f + g\|\, dm \leqslant \int_\Omega \|f\|\, dm + \int_\Omega \|g\|\, dm < \infty$. Donc $f + g \in L^1(\Omega, m, X)$. La linéarité de l'application $f \mapsto \int_\Omega f\, dm$ est évidente par définition et d'après le Fait 5.2.

(2) est clair par le Fait 5.2 ; et (3) est un exo facile. □

COROLLAIRE 5.9. *Si $f \in L^1(\Omega, m, X)$, alors*

$$\forall x^* \in X^* \ : \ \left\langle x^*, \int_\Omega f\, dm \right\rangle = \int_\Omega \langle x^*, f(t)\rangle\, dm(t).$$

REMARQUE 5.10. Le Théorème de convergence dominée est valable pour les fonctions à valeurs vectorielles : si $(f_n)_{n \in \mathbb{N}}$ est une suite dans $L^1(\Omega, m, X)$ qui converge presque partout vers une fonction $f : \Omega \to X$ et s'il existe une fonction $g \in L^1(\Omega, m)$ telle que $\forall n \in \mathbb{N} \ : \ \|f(t)\| \leqslant g(t)$ pp, alors $f \in L^1(\Omega, m, X)$ et $\int_\Omega \|f_n - f\|\, dm \to 0$; en particulier $\int_\Omega f_n\, dm \to \int_\Omega f\, dm$.

Démonstration. Il suffit d'appliquer le Théorème de convergence dominée habituel aux fonctions $\|f_n - f\|$. □

Chapitre 5

Dualité

1. Le dual d'un espace vectoriel normé

1.1. Rappels et exemples.

RAPPEL. Si X est un espace vectoriel topologique sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, on note X^* l'ensemble des formes linéaires continues de X dans $\mathbb{K}$. On dit que X^* est le **dual** de X. Si X est un espace vectoriel normé, alors $X^* = \mathcal{L}(X, \mathbb{K})$ est un espace de Banach (même si X n'est pas complet).

NOTATIONS. On note en général x^*, y^*, z^* les éléments de X^*. Si $x^* \in X^*$ et $x \in X$, on écrit en général $\langle x^*, x\rangle$ au lieu de $x^*(x)$.

REMARQUE. Soit X un espace vectoriel normé, et notons B_X la boule unité fermée de X. Si $x^* \in X^*$, alors

$$\|x^*\| = \sup \big\{|\langle x^*, x\rangle|;\ x \in B_X\big\} = \sup \big\{\mathrm{Re}\,\langle x^*, x\rangle;\ x \in B_X\big\}.$$

Démonstration. La 1ère égalité est la définition de $\|x^*\|$. Notons A le 1er "sup" et B le 2ème. On a évidemment $A \geqslant B$. Inversement, si $x \in B_X$, on peut trouver $\omega \in \mathbb{K}$ tel que $|\omega| = 1$ et $|\langle x^*, x\rangle| = \omega\,\langle x^*, x\rangle$; donc $|\langle x^*, x\rangle| = \mathrm{Re}\,|\langle x^*, x\rangle| = \mathrm{Re}\,\langle x^*, \omega x\rangle \leqslant B$ car $\omega x \in B_X$. $\square$

EXEMPLES. On sait décrire complètement les duaux de 3 classes très importantes d'espaces.

(1) Si H est un espace de Hilbert, alors "$H^* = H$".

(2) Si $(\Omega, \mathfrak{B}, m)$ est un espace mesuré avec une mesure m sigma-finie et si $1 \leqslant p < \infty$, alors "$(L^p)^* = L^q$", où $\frac{1}{p} + \frac{1}{q} = 1$.

(3) Si Ω est un espace topologique compact, alors "$\big(\mathcal{C}(\Omega)\big)^* = M(\Omega)$".

Remarque. Dans le cas de l'espace de Hilbert $H := L^2 = L^2(\Omega, m)$, il y a une subtilité : si on identifie $(L^2)^*$ à L^2 par la dualité "banachique" L^p-L^q avec $p = 2 = q$, cela veut dire qu'on identifie une $f \in L^2$ avec la forme linéaire $\Phi_f : g \mapsto \int_\Omega gf\, dm$; mais si on identifie $(L^2)^*$ à L^2 par la dualité "hilbertienne", cela veut dire qu'on identifie une $f \in L^2$ avec la forme linéaire $\Psi_f : g \mapsto \langle g, f\rangle_{L^2} = \int_\Omega g\bar{f}\, dm$, *i.e.* $\Psi_f = \Phi_{\bar{f}}$. Dans le cas "banachique", l'identification canonique entre $(L^2)^*$ et L^2 est linéaire ; et dans le cas "hilbertien", elle est antilinéaire.

LEMME 1.1. *Soit X un espace vectoriel normé. Pour tout $x \in X$, on peut trouver $x^* \in X^*$ telle que $\|x^*\| = 1$ et $\langle x^*, x\rangle = \|x\|$. En particulier :*

$$\forall x \in X\ :\ \|x\| = \sup \big\{|\langle x^*, x\rangle|;\ x^* \in B_{X^*}\big\} = \sup \big\{\mathrm{Re}\,\langle x^*, x\rangle;\ x^* \in B_{X^*}\big\}.$$

Démonstration. Conséquence du Théorème de Hahn-Banach. $\square$

Corollaire 1.2. *Si X est un espace vectoriel normé quelconque, alors X se plonge isométriquement dans son* ***bidual*** *$X^{**} := (X^*)^*$. Plus précisément, on définit une isométrie linéaire $j_X : X \to X^{**}$ en posant*

$$\langle j_X(x), x^* \rangle := \langle x^*, x \rangle \qquad \textit{pour tous } x \in X \textit{ et } x^* \in X^*.$$

On dit que $j_X : X \to X^{**}$ est le ***plongement canonique*** *de X dans X^{**}.*

Démonstration. C'est évident. $\square$

Remarque. Si $x \in X$, on écrira parfois δ_x au lieu de $j_X(x)$.

Corollaire 1.3. *Tout espace vectoriel normé X admet un* complété *: il existe un espace de Banach $\widehat{X}$ tel que "$X \subseteq \widehat{X}$" et X est dense dans $\widehat{X}$.*

Démonstration. Il suffit de poser $\widehat{X} := \overline{j_X(X)} \subseteq X^{**}$, et de se souvenir que $X^{**} = (X^*)^*$ est complet. $\square$

Exercice. Montrer que le complété est unique à isomorphisme isométrique près.

Corollaire 1.4. *Si X est un evn, alors* $\dim(X^*) = \dim(X)$.

Démonstration. Si $\dim(X) < \infty$, alors X^* est l'ensemble de toutes les formes linéaires sur X, donc l'algèbre linéaire nous dit que $\dim(X^*) = \dim(X) < \infty$. On en déduit que si $\dim(X^*) < \infty$, alors $\dim(X^{**}) < \infty$, et donc $\dim(X) < \infty$ puisque $X \subseteq X^{**}$; autrement dit : si $\dim(X) = \infty$, alors $\dim(X^*) = \infty = \dim(X)$. $\square$

Notations. Soit X un espace vectoriel topologique.

- Si $A \subseteq X$, on pose $A^\perp := \big\{x^* \in X^*;\ \langle x^*, u\rangle = 0 \text{ pour tout } u \in A\big\}$.
- Si $B \subseteq X^*$, on pose $B_\perp := \big\{x \in X;\ \langle u^*, x\rangle = 0 \text{ pour tout } u^* \in B\big\}$.

Lemme 1.5. *Soit X un evt localement convexe. Si $A \subseteq X$, alors*

$$\big(A^\perp\big)_\perp = \overline{\mathrm{vect}\,(A)}.$$

Démonstration. Soit $E := \mathrm{vect}(A)$. Par linéarité, on a $A^\perp = E^\perp$, donc $\big(A^\perp\big)_\perp = \big(E^\perp\big)_\perp$. Il est clair que $\big(E^\perp\big)_\perp$ est un sous-espace fermé de X et que $E \subseteq \big(E^\perp\big)_\perp$, donc $\big(E^\perp\big)_\perp \supseteq \overline{E}$. Inversement, si $x \in X$ et $x \notin \overline{E}$, alors on peut trouver $x^* \in E^\perp$ tel que $\langle x^*, x\rangle \neq 0$ (conséquence du Théorème de Hahn-Banach), donc $x \notin \big(E^\perp\big)_\perp$. Ainsi $X \backslash \overline{E} \subseteq X \backslash \big(E^\perp\big)_\perp$, et au total $\overline{E} = \big(E^\perp\big)_\perp$. $\square$

1.2. Sous-espaces et quotients. La proposition suivante peut sembler totalement formelle (et d'une certaine façon, elle l'est) ; mais elle est en fait très utile.

Proposition 1.6. *Soit X un espace vectoriel normé, et soit E un sous-espace vectoriel de X.*

(1) *On a "$E^* = X^*/E^\perp$". Plus précisément : l'****opérateur de restriction*** *$R : X^* \to E^*$ défini par $Rx^* := (x^*)_{|E}$ induit par passage au quotient une isométrie bijective de $X^*/E^\perp$ sur E^*, d'inverse $E^* \ni \varphi \mapsto \pi_{X^*/E^\perp}\Phi$, où $\Phi \in X^*$ est n'importe quel prolongement de Hahn-Banach de φ.*

(2) *Si E est un sous-espace fermé, alors "$\big(X/E\big)^* = E^\perp$". Plus précisément, l'application $\Phi \mapsto \Phi \circ \pi_{X/E}$ est une isométrie bijective de $(X/E)^*$ sur $E^\perp$.*

Démonstration. (1) Comme X^* est un espace de Banach et que $E^\perp$ est un sous-espace fermé de X^*, le quotient $X^*/E^\perp$ est bien un espace de Banach.
D'après le Théorème de Hahn-Banach, l'opérateur R est surjectif; plus précisément, on a $R\big(\mathring{B}_{X^*}\big) = \mathring{B}_{E^*}$. De plus, $\ker(R) = E^\perp$; donc R "passe au quotient", et l'application quotient $\widetilde{R} : X^*/E^\perp \to E^*$ est bijective puisque R est surjectif. Comme $\pi_{X^*/E^\perp}\big(\mathring{B}_{X^*}\big) = \mathring{B}_{X^*/E^\perp}$, on a $\widetilde{R}\big(\mathring{B}_{X^*/E^\perp}\big) = R\big(\mathring{B}_{X^*}\big) = \mathring{B}_{E^*}$; et comme $\widetilde{R}$ est bijective, on en déduit que $\widetilde{R}$ est une isométrie (exo). Soit $J := \widetilde{R}^{-1} : E^* \to X^*/E^\perp$. Si $\varphi \in E^*$ et si si Φ est un prolongement de Hahn-Banach de φ, alors $R(\Phi) = \varphi$, autrement dit $\widetilde{R}(\pi_{X^*/E^\perp}\Phi) = \varphi$; et donc $J\varphi = \pi_{X^*/E^\perp}\Phi$.
(2) Si $\Phi \in (X/E)^*$, alors $J\Phi := \Phi \circ \pi_{X/E} \in X^*$ et $J\Phi \equiv 0$ sur E par définition ; autrement dit $J\Phi \in E^\perp$. Inversement, si $x^* \in E^\perp$, alors x^* "passe au quotient" puisque $\ker(x^*) \supseteq E$: il existe une unique $\Phi \in \mathcal{L}(X/E, \mathbb{K}) = (X/E)^*$ telle que $\Phi \circ \pi_{X/E} = x^*$, *i.e.* $x^* = J\Phi$, et on a $\|\Phi\| = \|x^*\|$. Donc l'application linéaire $J : (X/E)^* \to E^\perp$ est une isométrie bijective. □

1.3. Séparabilité.

Le résultat suivant rend parfois service.

PROPOSITION 1.7. *Soit X un espace vectoriel normé. Si X^* est séparable, alors X est séparable.*

Démonstration. Soit $D = \{z_n^*;\ n \in \mathbb{N}\}$ un ensemble dénombrable dense dans $S_{X^*} := \{x^* \in X^*;\ \|x^*\| = 1\}$. Pour tout $n \in \mathbb{N}$, on peut choisir $z_n \in B_X$ tel que $|\langle z_n^*, z_n\rangle| \geqslant 1/2$. Montrons que $E := \mathrm{vect}\{z_n;\ n \in \mathbb{N}\}$ est dense dans X. Par le critère dual de densité, il suffit de vérifier que $E^\perp = \{0\}$. Par l'absurde, supposons que $E^\perp \neq \{0\}$. Soit $x^* \in E^\perp$ avec $\|x^*\| = 1$. Comme D est dense dans S_{X^*}, on peut trouver $n \in \mathbb{N}$ tel que $\|z_n^* - x^*\| < 1/2$. Alors $|\langle x^*, z_n\rangle| \geqslant |\langle z_n^*, z_n\rangle| - |\langle x^* - z_n^*, z_n\rangle| \geqslant \frac{1}{2} - \|z_n^* - x^*\|\,\|z_n\| > \frac{1}{2} - \frac{1}{2} = 0$, ce qui est absurde puisqu'on doit avoir $\langle x^*, z_n\rangle = 0$. □

Remarque. La réciproque est fausse : si X est séparable, alors X^* n'a pas de raison de l'être. Par exemple, $X := \ell^1(\mathbb{N})$ est séparable, mais $X^* = \ell^\infty(\mathbb{N})$ ne l'est pas.

2. Adjoint d'un opérateur

2.1. Généralités.

DÉFINITION 2.1. *Soient X et Y deux espaces de Banach, et soit $T \in \mathcal{L}(X, Y)$. L'**adjoint banachique** de T est l'opérateur $T^* = Y^* \to X^*$ défini par $T^*y^* := y^* \circ T$; autrement dit :*
$$\forall y^* \in Y^*\ \forall x \in X\ :\ \langle T^*y^*, x\rangle = \langle y^*, Tx\rangle.$$

Remarque 1. On a bien $T^* \in \mathcal{L}(Y^*, X^*)$ car T^* est linéaire et $\|T^*y^*\| \leqslant \|y^*\|\,\|T\|$ pour toute $y^* \in Y^*$.

Remarque 2. On écrit parfois ${}^tT$ au lieu de T^*.

EXEMPLE 1. Soient H_1, H_2 deux espaces de Hilbert. Si $T \in \mathcal{L}(H_1, H_2)$, alors l'adjoint Banachique ${}^tT : (H_2)^* \to (H_1)^*$ s'identifie à l'adjoint Hilbertien $T^* : H_2 \to H_1$ *via* les identifications canoniques $(H_i)^* \equiv H_i$.

Démonstration. Exo. □

Remarque. Il faut cependant faire attention : l'application $T \mapsto {}^tT$ est linéaire, mais l'application $T \mapsto T^*$ est *antilinéaire* si H_1 et H_2 sont des Hilbert complexes.

Exemple 2. Soit X un espace de Banach, et soit E un sous-espace fermé de X.

(i) Si on note $\pi : X \to X/E$ l'application quotient canonique et si on identifie $(X/E)^*$ et $E^\perp$, alors π^* est l'injection canonique $E^\perp \hookrightarrow X^*$.

(ii) Si on note $j : E \hookrightarrow X$ l'injection canonique, alors j^* est l'opérateur de restriction $R : X^* \to E^*$; et si on identifie E^* et $X^*/E^\perp$, alors j^* est l'application quotient canonique $X^* \to X^*/E^\perp$.

Démonstration. Par définition, on a $\pi^*\Phi = \Phi \circ \pi$ pour toute $\Phi \in (X/E)^*$; autrement dit $\pi^* x^* = x^*$ pour toute $x^* \in E^\perp$. De même, on a $j^*x^* = x^* \circ j = (x^*)_{|E}$ pour toute $x^* \in X^*$. □

Propriétés formelles. Soient X, Y, Z des espaces de Banach.

(1) On a $\|T^*\| = \|T\|$ pour tout $T \in \mathcal{L}(X,Y)$. Donc, l'application $T \mapsto T^*$ est une isométrie linéaire de $\mathcal{L}(X,Y)$ dans $\mathcal{L}(Y^*,X^*)$.

(2) Si $T \in \mathcal{L}(X,Y)$ et si on pose $T^{**} := (T^*)^* \in \mathcal{L}(X^{**},Y^{**})$, alors "$(T^{**})_{|X} = T$". De façon précise : $T^{**}j_X = j_Y T$.

(3) On a $(I_X)^* = I_{X^*}$.

(4) Si $B \in \mathcal{L}(X,Y)$ et $A \in \mathcal{L}(Y,Z)$, alors $(AB)^* = B^*A^*$.

(5) Si $T \in \mathcal{L}(X,Y)$ est inversible, alors T^* est inversible et $(T^*)^{-1} = (T^{-1})^*$.

Démonstration. (1) C'est un calcul "automatique" utilisant le Lemme 1.1 :

$$\begin{aligned}\|T^*\| &= \sup_{\|y^*\|\leqslant 1} \|T^*y^*\| \\ &= \sup_{\|y^*\|\leqslant 1}\ \sup_{\|x\|\leqslant 1} |\langle T^*y^*, x\rangle| \\ &= \sup_{\|x\|\leqslant 1}\ \sup_{\|y^*\|\leqslant 1} |\langle y^*, Tx\rangle| \\ &= \sup_{\|x\|\leqslant 1} \|Tx\| = \|T\|.\end{aligned}$$

(2) Par définition, on a $\forall x \in X\, \forall y^* \in Y^*$: $\langle T^{**}j_X(x), y^*\rangle = \langle j_X(x), T^*y^*\rangle = \langle T^*y^*, x\rangle = \langle y^*, Tx\rangle = \langle j_Y Tx, y^*\rangle$.
(3), (4) et (5) sont laissés en exo. □

2.2. Noyaux et images.

Proposition 2.2. *Soient X et Y des espaces de Banach. Si $T \in \mathcal{L}(X,Y)$, alors* $\ker(T^*) = \mathrm{Im}(T)^\perp$ *et* $\ker(T) = \mathrm{Im}(T^*)_\perp$.

Démonstration. Par définition, $y^* \in \ker(T^*) \iff \forall x \in X\ : \langle T^*y^*, x\rangle = 0 \iff \forall x \in X\ :\ \langle y^*, Tx\rangle = 0$; donc $\ker(T^*) = \mathrm{Im}(T)^\perp$. *Idem* pour $\ker(T)$. □

Corollaire 2.3. *Si $T \in \mathcal{L}(X,Y)$, alors* $\ker(T^*)_\perp = \overline{\mathrm{Im}(T)}$. *En particulier, T^* est injectif si et seulement si* $\mathrm{Im}(T)$ *est dense dans Y.*

Démonstration. On a $(E^\perp)_\perp = \overline{E}$ pour tout sous-espace vectoriel $E \subseteq Y$. □

COROLLAIRE 2.4. *Un opérateur $T \in \mathcal{L}(X,Y)$ est inversible* si et seulement si T^* *est inversible.*

Démonstration. On sait déjà que si T est inversible, alors T^* est inversible. Inversement, supposons que T^* soit inversible. Alors $T^{**} : X^{**} \to X^{**}$ est inversible, et $A := (T^{**})^{-1} = \big((T^*)^{-1}\big)^*$. Comme $AT^{**} = I_{X^{**}}$ et $(T^{**})_{|X} = T$, on voit que $ATx = x$ pour tout $x \in X$. Donc $\|x\| \leqslant \|A\|\,\|Tx\|$ pour tout $x \in X$, et donc T est un plongement. Comme X est un epace de Banach, on en déduit que T est injectif et à image fermée. Mais T est aussi à image dense car T^* est injectif ; donc T est en fait bijectif, et par suite inversible car on sait déjà que T est un plongement. (On n'a pas besoin d'appliquer le Théorème d'isomorphisme de Banach). □

2.3. Adjoint et surjectivité.

THÉORÈME 2.5. *Soient X, Y des espaces de Banach, et soit $T \in \mathcal{L}(X,Y)$. Alors T est surjectif si et seulement si T^* est un plongement.*

Pour la preuve, on a besoin du lemme suivant, qui est par ailleurs intéressant et utile en d'autres circonstances.

LEMME 2.6. *Soient X, Y des espaces de Banach, et soit $T \in \mathcal{L}(X,Y)$. Soient également $C \in \mathbb{R}^+$ et α tel que $0 < \alpha < 1$. On suppose que*

$$\forall y \in B_Y(0,1)\ \exists x \in B_X(0,C)\ :\ \|Tx - y\| \leqslant \alpha. \tag{2.1}$$

Alors $T\big(B_X(0,C)\big)$ contient la boule $B_Y(0,1-\alpha)$.

Démonstration. On va montrer que $T\big(B_X(0,\frac{C}{1-\alpha})\big)$ contient $B_Y(0,1)$, ce qui revient au même par linéarité de T.

Remarquons d'abord que (2.1) est en fait vraie avec une inégalité *stricte*. En effet, si $y \in B_Y(0,1)$, on peut trouver $\lambda > 1$ tel que $\widetilde{y} := \lambda y$ vérifie encore $\widetilde{y} \in B_Y(0,1)$, puis $\widetilde{x} \in B_X(0,C)$ tel que $\|T\widetilde{x} - \widetilde{y}\| \leqslant \alpha$. Alors $x := \frac{\widetilde{x}}{\lambda} \in B_X(0,1)$ et $\|Tx - y\| = \frac{1}{\lambda}\|T\widetilde{x} - \widetilde{y}\| \leqslant \alpha/\lambda < \alpha$.

Soit $y \in B_Y(0,1)$ quelconque : on cherche $x \in X$ tel que $\|x\| < \frac{C}{1-\alpha}$ et $Tx = y$. Par (2.1), on peut trouver $x_0 \in B_X(0,C)$ tel que $\|y - Tx_0\| < \alpha$. Alors $y_1 := \frac{1}{\alpha}(y - Tx_0) \in B_Y(0,1)$, donc on peut trouver $x_1 \in B_X(0,C)$ tel que $\|Tx_1 - \frac{1}{\alpha}(y - Tx_0)\| < \alpha$, *i.e.*

$$\big\|y - T(x_0 + \alpha x_1)\big\| < \alpha^2.$$

Alors $y_2 := \frac{1}{\alpha^2}\big(y - T(x_0 + \alpha x_1)\big) \in B_Y(0,1)$, donc on peut trouver $x_2 \in B_X(0,C)$ tel que $\|y_2 - Tx_2\| < \alpha$, et ainsi de suite : par récurrence, on construit une suite $(x_k)_{k \geqslant 0} \subseteq B_X(0,C)$ telle que

$$\forall n \in \mathbb{N}\ :\ \big\|y - T(x_0 + \alpha x_1 + \cdots + \alpha^n x_n)\big\| < \alpha^{n+1}. \tag{2.2}$$

Comme $\alpha < 1$ et $\|x_k\| < C$ pour tout k, on a $\sum_{k=0}^\infty \|\alpha^k x_k\| < C\sum_{k=0}^\infty \alpha^k = \frac{C}{1-\alpha} < \infty$. En particulier, la série $\sum \alpha^k x_k$ converge dans X car X est un espace de Banach. Si on pose $x := \sum_{k=0}^\infty \alpha^k x_k$, alors $\|x\| \leqslant \sum_{k=0}^\infty \|\alpha^k x_k\| < \frac{C}{1-\alpha}$; et par (2.2), on a $Tx = \lim_{n\to\infty} T(x_0 + \alpha x_1 + \cdots + \alpha^n x_n) = y$. □

COROLLAIRE 2.7. *Si $\overline{T\big(B_X(0,C)\big)}$ contient $B_Y(0,1)$, alors en fait $T\big(B_X(0,C)\big)$ contient déjà $B_Y(0,1)$.*

Démonstration. L'hypothèse dit que la propriété (2.1) est vérifiée pour *tout* $0 < \alpha < 1$. Donc $T\big(B_X(0,C)\big)$ contient $\bigcup_{n \in \mathbb{N}^*} B_Y(0, 1 - \frac{1}{n}) = B_Y(0,1)$. □

Preuve du Théorème 2.5. (i) Supposons que T soit surjectif. Comme X et Y sont des espaces de Banach, il existe alors une constante $c > 0$ telle que $T(B_X) \supseteq c\,B_Y$. Pour tout $y^* \in Y^*$, on a donc

$$\|T^*y^*\| = \sup_{x\in B_X} |\langle T^*y^*, x\rangle = \sup_{x\in B_X} |\langle y^*, Tx\rangle| \geqslant \sup_{y\in cB_Y} |\langle y^*, y\rangle| = c\,\|y^*\|;$$

donc T^* est un plongement.
Voici une autre façon de dire la même chose. Notons $\pi : X \to \ker(T)$ l'application quotient canonique et $\widetilde{T} : X/\ker(T) \to Y$ l'opérateur induit. Alors $\widetilde{T}$ est bijectif car T est surjectif; donc $\widetilde{T}$ est inversible car $X/\ker(T)$ et Y sont des espaces de Banach. On a $T = \widetilde{T}\pi$, donc $T^* = \pi^*\widetilde{T}^*$; et donc T^* est un plongement car $\widetilde{T}^*$ est inversible et π^* est un plongement (isométrique).

(ii) Inversement, supposons que T^* soit un plongement. On a donc une constante $c > 0$ telle que $\forall y^* \in Y^*\ :\ \|T^*y^*\| \geqslant c\,\|y^*\|$.

Fait. On a $B_Y \subseteq \overline{T\big(\frac{1}{c}\,B_X\big)}$.

Preuve du Fait. Il s'agit de montrer que si $y \in Y$ est tel que $y \notin \overline{T\big(\frac{1}{c}\,B_X\big)}$, alors $\|y\| > 1$. Par le Théorème de séparation des convexes, on peut trouver $y^* \in Y^*$ tel que

$$\mathrm{Re}\,\langle y^*, y\rangle > \sup\big\{\mathrm{Re}\,\langle y^*, Tx\rangle;\ x \in (1/c)\,B_X\big\}.$$

Autrement dit :

$$\mathrm{Re}\,\langle y^*, y\rangle > \frac{1}{c}\,\sup\big\{\mathrm{Re}\,\langle T^*y^*, x\rangle;\ x \in B_X\big\} = \frac{1}{c}\,\|T^*y^*\|.$$

Comme $\|T^*y^*\| \geqslant c\,\|y^*\|$, on en déduit $|\langle y^*, y\rangle| > \|y^*\|$, et donc $\|y\| > 1$. □

Par le Fait, on a (exo) $\mathring{B}_Y \subseteq \overline{T\big(\frac{1}{c}\,\mathring{B}_X\big)}$. Comme X est un espace de Banach, on peut en fait enlever l'adhérence grâce au Lemme 2.6. On obtient ainsi que $T(\mathring{B}_X)$ contient $c\,\mathring{B}_Y$, et donc T est surjectif. □

Remarque. La preuve du théorème a montré que si $T \in \mathcal{L}(X,Y)$ et $c > 0$, alors on a l'équivalence suivante :

$$T \text{ est un plongement avec constante } c \iff T(B_X) \supseteq c\,B_Y.$$

Corollaire 2.8. *Soient X, Y des espaces de Banach. Un opérateur $T \in \mathcal{L}(X,Y)$ est inversible si et seulement si T est injectif et T^* est un plongement.*

Démonstration. Par le Théorème d'isomorphisme de Banach, T est inversible si et seulement si il est bijectif, *i.e.* injectif et surjectif. □

Corollaire 2.9. *Soient X, Y des espaces de Banach. Pour $T \in \mathcal{L}(X,Y)$, les propriétés suivantes sont équivalentes.*

(1) *T est à image fermée ;*

(2) $\mathrm{Im}(T^*) = \ker(T)^\perp$ *;*

(3) *T^* est à image fermée.*

Démonstration. (1) $\implies$ (2). Supposons que T soit à image fermée. Posons $Y_0 := \operatorname{Im}(T)$, qui est donc un espace de Banach, et notons $T_0 : X \to Y_0$ l'opérateur T considéré comme opérateur de X dans Y_0. Alors T_0 est surjectif et $\ker(T_0) = \ker(T)$. Donc T_0 induit un opérateur inversible $\widetilde{T_0} : X/\ker(T) \to Y_0$; et on a $T = i_0\widetilde{T_0}\pi$, où $\pi : X \to X/\ker(T)$ est l'application quotient et $i_0 : Y_0 \to Y$ est l'injection canonique. Donc $\operatorname{Im}(T^*) = \pi^*(\widetilde{T_0})^* i_0^*(Y^*)$. Mais $i_0^* : Y^* \to Y_0^*$ est surjectif (c'est l'opérateur de restriction), et $(\widetilde{T_0})^* : Y_0^* \to (X/\ker(T))^*$ est inversible. Donc $(\widetilde{T_0})^* i_0^*(Y^*) = (X/\ker(T))^*$, et donc $\operatorname{Im}(T^*) = \pi^*\big((X/\ker(T))^*\big) = \ker(T)^\perp$.

Preuve sans quotient. On suppose que T est à image fermée et on veut montrer que $\operatorname{Im}(T^*) = \ker(T)^\perp$. L'inclusion $\operatorname{Im}(T^*) \subseteq \ker(T)^\perp$ est claire (micro-exo). Inversement, soit $x^* \in \ker(T)^\perp$: on cherche $y^* \in Y^*$ telle que $x^* = T^*y^*$. Comme $x^* \in \ker(T)^\perp$, il existe une unique forme linéaire $\varphi : \operatorname{Im}(T) \to \mathbb{K}$ telle que

$$\forall x \in X \;:\; \varphi(Tx) = \langle x^*, x\rangle.$$

De plus, φ est *continue*. En effet, comme $\operatorname{Im}(T)$ est fermé dans Y, c'est un espace de Banach, et $T : X \to \operatorname{Im}(T)$ est surjectif. Par le Théorème de majoration *a priori*, il existe une constante C telle que, pour tout $y \in \mathrm{Im(T)}$, on peut trouver $x \in X$ vérifiant $Tx = y$ et $\|x\| \leqslant C\,\|y\|$. On a alors $|\varphi(y)| = |\langle x^*, x\rangle| \leqslant \|x^*\|\,\|x\| \leqslant C\|x^*\|\,\|y\|$; donc φ est continue avec $\|\varphi\| \leqslant C\|x^*\|$. En notant y^* une extension de Hahn-Banach de φ, on a $T^*y^* = x^*$ par définition de φ.

(2) $\implies$ (3) est évident : $E^\perp$ est fermé dans X^* pour tout $E \subseteq X$.

(3) $\implies$ (1). Supposons que $\operatorname{Im}(T^*)$ soit fermé dans X^*. Posons $Y_0 := \overline{\operatorname{Im}(T)} \subseteq Y$, et soit $T_0 : X \to Y_0$ l'opérateur T considéré comme opérateur de X dans Y_0. Alors $T = i_0T_0$ où $i_0 : Y_0 \to Y$ est l'injection canonique, donc $T^* = T_0^* i_0^*$, et donc $\operatorname{Im}(T_0^*) = \operatorname{Im}(T^*)$ car $i_0^* : Y^* \to Y_0^*$ est surjectif ; en particulier T_0^* est à image fermée. De plus T_0^* est aussi injectif car T_0 est par définition à image dense. Donc T_0^* est un plongement. Par le théorème, on en déduit que T_0 est surjectif. Donc $\operatorname{Im}(T) = \operatorname{Im}(T_0) = Y_0 = \overline{\operatorname{Im}(T)}$; autrement dit T est à image fermée.

Preuve directe de (1) $\iff$ (3). Soit $Y_0 := \overline{\operatorname{Im}(T)}$, et soit $T_0 : X \to Y_0$ l'opérateur T considéré comme opérateur de X dans Y_0. Alors T est à image fermée si et seulement si T_0 est surjectif ; ce qui revient à dire que T_0^* est un plongement puisque Y_0 est un espace de Banach. De plus, T_0^* est injectif puisque T_0 est à image dense. Donc $\operatorname{Im}(T)$ est fermée si et seulement $\operatorname{Im}(T_0^*)$ est fermée. Mais $T = i_0T_0$, où $i_0 : Y_0 \to Y$ est l'injection canonique ; donc $T^* = T_0^* i_0^* = T_0^* R$, où $R : Y^* \to Y_0^*$ est l'opérateur de restriction ; et donc $\operatorname{Im}(T^*) = \operatorname{Im}(T_0^*)$ car R est surjectif par Hahn-Banach. On a donc bien montré que $\operatorname{Im}(T)$ est fermée si et seulement si $\operatorname{Im}(T^*)$ est fermée. $\square$

Corollaire 2.10. *Soient X et Y des espaces de Banach, et soit $T \in \mathcal{L}(X,Y)$. Alors T est un plongement si et seulement si T^* est surjectif.*

Démonstration. Si T est un plongement, alors T est injectif et à image fermée, donc $\operatorname{Im}(T^*) = \{0\}^\perp = X^*$ par le corollaire précédent. Inversement, si T^* est surjectif, alors T est injectif puisque $\ker(T) = \operatorname{Im}(T^*)_\perp$, et T est à image fermée car T^* est à image fermée ; donc T est un plongement. $\square$

2.4. Compacité.

Théorème 2.11. *Soient X, Y des espaces de Banach, et soit $T \in \mathcal{L}(X,Y)$. Alors T est compact si et seulement si T^* est compact.*

Démonstration. Supposons que T soit compact, *i.e.* que $K := \overline{T(B_X)}$ soit un compact de Y. Soit $(y_n^*)_{n\in\mathbb{N}}$ une suite bornée dans Y^* : on veut montrer que la suite $(T^*y_n^*)$ possède une sous-suite qui converge dans X^*. Pour $n \in \mathbb{N}$, posons $f_n := (y_n^*)_{|K} \in \mathcal{C}(K)$. La suite (f_n) est bornée dans $\mathcal{C}(K)$ car (y_n^*) est bornée et K est borné. De plus, si on pose $M := \sup_{n\in\mathbb{N}} |y_n^*|$, alors toutes les fonctions f_n sont M-lipschitziennes. D'après le Théorème d'Ascoli, on en déduit que (f_n) possède une sous-suite (f_{n_k}) qui converge uniformément sur K. Alors (f_{n_k}) vérifie le critère de Cauchy uniforme :

$$\|f_{n_q} - f_{n_p}\|_\infty \to 0 \qquad \text{quand } p, q \to \infty.$$

Mais

$$\begin{aligned}
\|f_{n_q} - f_{n_p}\|_\infty &= \sup \left\{ |\langle y_{n_q}^* - y_{n_p}^*, y\rangle|;\ y \in \overline{T(B_X)} \right\} \\
&\geqslant \sup \left\{ |\langle y_{n_q}^* - y_{n_p}^*, Tx\rangle|;\ x \in B_X \right\} \\
&= \sup \left\{ |\langle T^*y_{n_q}^* - T^*y_{n_p}^*, x\rangle|;\ x \in B_X \right\} \\
&= \|T^*y_{n_q}^* - T^*y_{n_p}^*\|.
\end{aligned}$$

Donc la suite $(T^*y_{n_k}^*)$ est de Cauchy dans X^*, et donc converge dans X^* car X^* est complet.

Inversement, si T^* est compact, alors T^{**} est compact par ce qui précède ; et donc T est compact puisque $T^{**}_{|X} = T$ (micro-exo ; on a besoin que Y soit complet). ☐

3. Topologies faible et préfaible

3.1. Généralités.

Définition 3.1. *Soit X un espace vectoriel normé..*

(i) *La **topologie faible de** X est la topologie sur X définie par les semi-normes $x \mapsto |\langle x^*, x\rangle|$, où $x^* \in X^*$. Cette topologie se note $\sigma(X, X^*)$, ou w.*

(ii) *La **topologie préfaible de** X^* est la topologie sur X^* définie par les semi-normes $x^* \mapsto |\langle x^*, x\rangle|$, où $x \in X$. Cette topologie se note $\sigma(X^*, X)$, ou w^*.*

Reformulation. Soit X un espace vectoriel normé.

(ii) La topologie w^* est la topologie sur X^* induite par la topologie de la convergence simple sur $\mathcal{F}(X, \mathbb{K}) = \mathbb{K}^X$.

(i) La topologie w est la topologie sur X induite par la topologie de la convergence simple sur $\mathcal{F}(X^*, \mathbb{K}) = \mathbb{K}^{X^*}$ quand on considère X comme une partie de $\mathcal{F}(X^*, \mathbb{K})$ *via* le plongement canonique $j_X : X \to X^{**}$.

Mode d'emploi. Soit X un espace vectoriel normé.

(1a) Si $x \in X$, une base de voisinages pour x dans (X, w) est fournie par les ensembles de la forme

$$B_{x_1^*,\dots,x_N^*}(x, \varepsilon) := \left\{ u \in X;\ |\langle x_k^*, u\rangle - \langle x_k^*, x\rangle| < \varepsilon \quad \text{pour } k = 1, \dots, N \right\},$$

où $x_1^*, \dots, x_N^* \in X^*$ et $\varepsilon > 0$.

(1b) Si $(x_p)_{p\in P}$ est une suite généralisée dans X et si $x \in X$, alors

$$x_p \xrightarrow{w} x \iff \langle x^*, x_p\rangle \to \langle x^*, x\rangle \quad \text{pour toute } x^* \in X^*.$$

(2a) Si $x^* \in X^*$, une base de voisinages pour x^* dans (X^*, w^*) est fournie par les ensembles de la forme

$$B_{x_1,\dots,x_N}(x^*,\varepsilon) := \big\{u^* \in X^*;\ |\langle u^*, x_k\rangle - \langle x^*, x_k\rangle| < \varepsilon \quad \text{pour } k = 1,\dots,N\big\},$$

où $x_1, \dots, x_N \in X$ et $\varepsilon > 0$.

(2b) Si $(x_p^*)_{p\in P}$ est une suite généralisée dans X^* et si $x^* \in X^*$, alors

$$x_p^* \xrightarrow{w^*} x^* \iff \langle x_p^*, x\rangle \to \langle x^*, x\rangle \quad \text{pour tout } x \in X.$$

Remarque 1. Les topologie w et w^* sont plus faibles que les topologies définies par les normes de X et de X^* : tout ensemble w-ouvert ou w^*-ouvert est ouvert en norme. Cependant :

- si $\dim(X) < \infty$, alors w et w^* coïncident avec les topologies définies par les normes de X et de X^* ;
- toute boule fermée $B \subseteq X$ est w-fermée et toute boule fermée $B \subseteq X^*$ est w^*-fermée.

Démonstration. La 1ère partie est évidente si on regarde par exemple la convergence des suites généralisées.
Si $\dim(X) < \infty$, on peut supposer que $X = \mathbb{K}^d$ en choisissant une base. Alors la convergence faible entraine la convergence coordonnée par coordonnée, laquelle est équivalente à la convergence pour $\|\cdot\|$ par équivalence des normes en dimension finie, donc $w = \tau_{\|\cdot\|}$; et *idem* pour w^* sur X^* car $\dim(X^*) < \infty$.
Soit B une boule fermée de X. Par translation et dilatation, on peut supposer que B est la boule unité B_X. D'après le Théorème de Hahn-Banach, $x \in B_X \iff \forall x^* \in B_{X^*}\ :\ |\langle x^*, x\rangle| \leqslant 1$; autrement dit $B_X = \bigcap_{x^*\in B_{X^*}} \big\{x \in X\ :\ |\langle x^*, x\rangle| \leqslant 1\big\}$. Ainsi, B_X est w-fermée en tant qu'intersection d'ensemble w-fermés. Même preuve pour une boule fermée $B \subseteq X^*$, en plus simple car on n'a pas besoin de Hahn-Banach. □

Remarque 2. Si on considère X comme contenu dans X^{**}, la topologie faible de X est la topologie induite sur X par la topologie préfaible $w^{**} = \sigma(X^{**}, X^*)$ de X^{**}.

Remarque 3. Si E est un sous-espace vectoriel de X, alors la topologie faible de E est la topologie induite sur E par la topologie faible de X.

Démonstration. Exo utilisant le Théorème de Hahn-Banach. □

LEMME 3.2. *Si X est un espace de Banach, alors les topologies w et w^* sont* séparées *; et (X, w) et (X^*, w^*) sont des espaces vectoriels topologiques localement convexes.*

Démonstration. Il faut juste vérifier que les topologie w et w^* sont séparées.
Si $a^*, b^* \in X^*$ et $a^* \neq b^*$, on peut trouver $x \in X$ tel que $\langle a^*, x\rangle \neq \langle b^*, x\rangle$; et si on pose $\varepsilon := |\langle a^*, x\rangle - \langle b^*, x\rangle|$, alors $U := B_x(a^*, \varepsilon/2)$ et $V := B_x(b^*, \varepsilon/2)$ sont des ouverts de (X^*, w^*) tels que $a^* \in U$, $b^* \in V$ et $U \cap V = \varnothing$. Donc la topologie w^* est séparée. On montre de même que la topologie w est séparée, en utilisant le fait que X^* *sépare les points de X* ; ce qui, évidemment, est tout à fait non-trivial (Hahn-Banach). □

3.2. Métrisabilité.

Proposition 3.3. *Soit X un espace de Banach. Si $\dim(X) = \infty$, alors les topologies w et w^* ne sont pas métrisables.*

La preuve repose sur le lemme suivant, dont on se re-servira.

Lemme 3.4. *Soit Z un espace vectoriel, et soient $\Phi_1, \dots, \Phi_N, \Phi$ des formes linéaires sur Z. Si on a $\bigcap_{k=1}^N \ker(\Phi_k) \subseteq \ker(\Phi)$, alors Φ est combinaison linéaire des Φ_k.*

Démonstration. C'est un exo bien connu d'algèbre linéaire. Soit $\Theta : Z \to \mathbb{K}^N$ l'application linéaire définie par $\Theta(x) := (\Phi_1(x), \dots, \Phi_N(x))$. On a $\ker(\Theta) = \bigcap_{k=1}^N \ker(\Phi_k)$, donc $\ker(\Theta) \subseteq \ker \Phi$. Il existe donc une application linéaire $L : \mathbb{K}^N \to \mathbb{K}$ telle que $\Phi : L \circ \Theta$, ce qui donne le résultat. □

Preuve de la Proposition 3.3. Supposons que l'espace (X, w) soit métrisable. Alors 0 possède une base dénombrable de voisinages dans (X, w) ; donc, on peut trouver un ensemble dénombrable $D = \{z_k^*;\ k \in \mathbb{N}\} \subseteq X^*$ vérifiant la propriété suivante : pour tout voisinage V de 0 dans (X, w), on peut trouver $k_1, \dots k_N \in \mathbb{N}$ et $\varepsilon > 0$ tels que $B_{z_{k_1}^*, \dots, z_{k_N}^*}(0, \varepsilon) \subseteq V$.

Fait. Pour toute $x^* \in X^*$, on peut trouver $k_1, \dots, k_N \in \mathbb{N}$ tels que $\bigcap_{i=1}^N \ker(z_{k_i}^*) \subseteq \ker(x^*)$.

Preuve du Fait 1. L'ensemble $V := \{x \in X;\ |\langle x^*, x\rangle| < 1\}$ est un voisinage de 0 dans (X, w), donc on peut trouver $k_1, \dots, k_N \in \mathbb{N}$ et $\varepsilon > 0$ tels que $B_{z_{k_1}^*, \dots, z_{k_N}^*}(0, \varepsilon) \subseteq V$. Si $x \in \bigcap_{i=1}^N \ker(x_{k_i}^*)$, alors $Rx \in \bigcap_{i=1}^N \ker(z_{k_i}^*)$ pour tout $R > 0$; donc $Rx \in B_{z_{k_1}^*, \dots, z_{k_N}^*}(0, \varepsilon) \subseteq V$, et donc $|\langle x^*, x\rangle| < 1/R$ pour tout $R > 0$. On a donc bien $\bigcap_{i=1}^N \ker(z_{k_i}^*) \subseteq \ker(x^*)$. □

Par le Fait et le Lemme 3.4, on voit que $X^* = \operatorname{vect}\{z_k^*;\ k \in \mathbb{N}\}$; donc X^* est de dimension algébrique dénombrable. Comme X^* est un espace de Banach, ceci n'est possible que si $\dim(X^*) < \infty$ (conséquence du Théorème de Baire). Donc $\dim(X) < \infty$ par le Corollaire 1.4.

On montre de la même façon que si la topologie w^* est métrisable, alors $\dim(X) < \infty$. (On a besoin ici que X soit un espace de Banach.) □

Corollaire 3.5. *Si X est un espace de Banach de dimension infinie, alors les topologies w et w^* sont* strictement plus faibles *que les topologies définies par les normes de X et de X^*.*

Exercice. Donner une preuve directe de ce denier résultat.

Exemple. Soit H un espace de Hilbert. Si $(e_n)_{n \in \mathbb{N}}$ est une suite orthonormale dans H, alors $e_n \xrightarrow{w} 0$ mais $e_n \not\to 0$ en norme.

Démonstration. Si $y \in H$, alors $\sum_{n=0}^\infty |\langle e_n, y\rangle|^2 \leqslant \|y\|^2 < \infty$ d'après l'inégalité de Bessel. Donc $\langle e_n, y\rangle \to 0$ pour tout $y \in H$, et donc $e_n \xrightarrow{w} 0$ puisque "$H^* = H$". □

Proposition 3.6. *Soit X un espace de Banach. Si X est séparable, alors les parties bornées de X^* sont w^*-métrisables. Si X^* est séparable, alors les parties bornées de X sont w-métrisables.*

Démonstration. Supposons que X soit séparable. Il suffit de montrer que pour tout $M \in \mathbb{R}^+$, la boule $B_M := \overline{B(0,M)} \subseteq X^*$ est w^*-métrisable. Soit $D = \{x_n;\ n \in \mathbb{N}\}$ un ensemble dénombrable dense dans X, et soit $J : B_M \to \mathbb{K}^{\mathbb{N}}$ l'application définie par $J(x^*) := \big(\langle x^*, x_n\rangle\big)_{n\in\mathbb{N}}$. Cette application J est continue, et elle est injective car D est dense dans X. Montrons que $J^{-1} : J(B_M) \to (X^*, w^*)$ est continue. Soit (x_p^*) une suite généralisée dans B_M et soit $x^* \in B_M$ tels que $J(x_p^*) \to J(x^*)$: on veut montrer que $x_p^* \xrightarrow{w^*} x^*$, *i.e.* que $\langle x_p^*, x\rangle \to \langle x^*, x\rangle$ pour tout $x \in X$. Fixons $x \in X$. Pour $\varepsilon > 0$ donné, on peut trouver $n \in \mathbb{N}$ tel que $\|x_n - x\| < \varepsilon$. Comme $J(x_p) \to J(x)$, on peut ensuite trouver $p_0 \in P$ tel que $\forall p \geq p_0\ :\ |\langle x_p^*, x_n\rangle - \langle x^*, x_n\rangle| < \varepsilon$. Pour $p \geq p_0$, on a alors

$$\begin{aligned}|\langle x_p^*, x\rangle - \langle x^*, x\rangle| &\leqslant |\langle x_p^*, x - x_n\rangle| + |\langle x_p^*, x_n\rangle - \langle x^*, x_n\rangle| + |\langle x^*, x_n - x\rangle| \\ &\leqslant \|x_p^*\|\,\|x_n - x\| + |\langle x_p^*, x_n\rangle - \langle x^*, x_n\rangle| + \|x^*\|\,\|x_n - x\| \\ &< (2M+1)\varepsilon.\end{aligned}$$

Donc en effet $\langle x_p^*, x\rangle \to \langle x^*, x\rangle$. On a ainsi montré que B_M est homéomorphe à $J(B_M) \subseteq \mathbb{K}^{\mathbb{N}}$, ce qui prouve que B_M est métrisable puisque $\mathbb{K}^{\mathbb{N}}$ l'est.

Si X^* est séparable, alors les parties bornées de X^{**} sont w^{**}-métrisables, donc les parties bornées de X sont w-métrisables puisque $X \subseteq X^{**}$ et que w est la topologie induite par w^{**} sur X. □

3.3. Ensembles bornés. Soit X un espace vectoriel normé. Convenons de dire qu'un ensemble $A \subseteq X$ est **w-borné** s'il est borné dans (X, w) ; autrement dit, si $\forall x^* \in X^*\ :\ \sup_{x\in A} |\langle x^*, x\rangle| < \infty$. De même, un ensemble $B \subseteq X^*$ sera dit **w^*-borné** si B est borné dans (X^*, w^*), it i.e. $\forall x \in X\ :\ \sup_{x^*\in B} |\langle x^*, x\rangle| < \infty$.

Proposition 3.7. *Soit X un espace vectoriel normé.*

(1) *Tout ensemble w-borné $A \subseteq X$ est borné*

(2) *Si X est complet, alors tout ensemble w^*-borné $B \subseteq X^*$ est borné.*

Démonstration. (2) est "simplement" le Théorème de Banach-Steinhaus. Pour (1), on applique (2) à X^*, qui est complet : on en déduit que tout ensemble w^{**}-borné $A \subseteq X^{**}$ est borné, ce qui donne en particulier (1). □

Corollaire 3.8. *Si X est un evn, alors toute* suite *w-convergente $(x_n)_{n\in\mathbb{N}} \subseteq X$ est bornée. Si X est un espace de Banach, alors toute suite w^*-convergente $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ est bornée.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}} \subseteq X$ une suite w-convergente. Pour tout $x^* \in X^*$, la suite $(\langle x^*, x_n\rangle)_{n\in\mathbb{N}}$ converge dans $\mathbb{K}$, donc elle est bornée. Autrement dit, l'ensemble $\{x_n;\ n \in \mathbb{N}\}$ est w-borné ; et donc cet ensemble est borné. Même preuve pour une suite w^*-convergente dans X^*. □

Remarque. Le résultat est faux pour des suites généralisées.

3.4. Applications linéaires continues.

Lemme 3.9. *Soit X un espace vectoriel normé sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.*

(1) *Une forme linéaire $\Phi : X \to \mathbb{K}$ est continue sur (X, w) si et seulement si elle est continue sur $(X, \|\cdot\|)$, i.e. $\Phi \in X^*$. Autrement dit : $(X, w)^* = X^*$.*

(2) *Une forme linéaire $\Phi : X^* \to \mathbb{K}$ est continue sur (X^*, w^*) si et seulement si elle est de la forme $\Phi = \delta_x$ pour un certain $x \in X$. Autrement dit : $(X^*, w^*)^* = X$.*

Démonstration. (1) Si Φ est continue sur (X, w), alors elle est continue sur $(X, \|\cdot\|)$ puisque $w \subseteq \tau_{\|\cdot\|}$. Inversement, supposons que Φ soit continue sur $(X, \|\cdot\|)$, *i.e.* $\Phi \in X^*$. Si $(x_p)_{p \in P}$ est une suite généralisée telle que $x_p \xrightarrow{w} x \in X$, alors $\Phi(x_p) \to \Phi(x)$ par définition de la topologie w ; donc Φ est continue sur (X, w).
(2) Rappelons que si $x \in X$, alors $\delta_x = j_X(x)$, *i.e.* $\delta_x : X^* \to \mathbb{K}$ est la forme linéaire définie par $\delta_x(x^*) := \langle x^*, x\rangle$. Si $\Phi = X^* \to \mathbb{K}$ est de la forme δ_x, alors Φ est continue sur (X^*, w^*) par définition de la topologie w^* (exo). Inversement, supposons que Φ soit continue sur (X^*, w^*). Alors on peut trouver un voisinage V de 0 dans (X^*, w^*) tel que $|\Phi(x^*)| < 1$ pour tout $x^* \in V$. Par définition de w^*, il existe donc $x_1, \dots, x_N \in X$ et $\varepsilon > 0$ tels que

$$x^* \in B_{x_1,\dots,x_N}(0, \varepsilon) \implies |\Phi(x^*)| < 1;$$

autrement dit :

$$\big(|\delta_{x_k}(x^*)| < \varepsilon \ \text{ pour } k = 1, \dots, N\big) \implies |\Phi(x^*)| < 1.$$

On en déduit, comme dans la preuve de la Proposition 3.3, qu'on a

$$\bigcap_{k=1}^{N} \ker(\delta_{x_k}) \subseteq \ker(\Phi);$$

et donc, d'après le Lemme 3.4, que Φ est combinaison linéaire des δ_{x_k}. Ainsi, on peut trouver $a_1, \dots, a_n \in \mathbb{K}$ tels que $\Phi = \sum_{k=1}^N a_k \delta_{x_k}$; autrement dit $\Phi = \delta_x$ où $x := \sum_{k=1}^N a_k x_k$. $\square$

Corollaire 3.10. *Si E est un sous-espace vectoriel de X^*, alors $\big(E_\perp\big)^\perp = \overline{E}^{w^*}$.*

Démonstration. Si on pose $Z := (X^*, w^*)$, alors Z est un evt localement convexe et "$Z^* = X$". Donc $E_\perp = E^\perp \subseteq Z^*$, et $(E_\perp)^\perp = (E^\perp)_\perp \subseteq Z$. D'après le Lemme 1.5, on a donc $(E_\perp)^\perp = \overline{E}^Z = \overline{E}^{w^*}$. $\square$

Proposition 3.11. *Soient X et Y des espaces vectoriels normés. Pour une application linéaire $T : X \to Y$, les propriétés suivantes sont équivalentes :*

(i) *T est continue ;*

(ii) *T est continue de $(X, \|\cdot\|)$ dans (Y, w) ;*

(iii) *T est continue de (X, w) dans (Y, w).*

Démonstration. Il est clair que (i) $\implies$ (ii) et que (iii) $\implies$ (ii).
Supposons que (ii) soit vérifiée, et montrons (iii). Par définition de la topologie w sur Y, il suffit de montrer que pour tout $y^* \in Y^*$, l'application $x \mapsto \langle y^*, Tx\rangle$ est continue sur (X, w). Par (ii), cette application est continue sur $(X, \|\cdot\|)$; et comme il s'agit d'une forme linéaire, elle est continue sur (X, w) par le Lemme 3.9.

Supposons maintenant que (ii) soit vérifiée, et montrons (i). Par hypothèse, pour tout $y^* \in Y^*$, la forme linéaire $x \mapsto \langle y^*, Tx\rangle$ est continue ; donc, on a $\forall y^* \in Y^*\ :\ \sup\{|\langle y^*, Tx\rangle|;\ x \in B_X\} < \infty$. Donc l'ensemble $T(B_X)$ est faiblement borné dans Y, et donc $T(B_X)$ est borné d'après la Proposition 3.7. Autrement dit : T est bornée sur B_X, et donc continue. □

Corollaire 3.12. *Soient X et Y des espaces de Banach. Si $T \in \mathcal{L}(X,Y)$ est un opérateur compact, alors T change toute suite w-convergente en une suite $\|\cdot\|$-convergente : si $(x_n)_{n\in\mathbb{N}}$ est une suite d'éléments de X telle que $x_n \xrightarrow{w} x \in X$, alors $\|Tx_n - Tx\| \to 0$.*

Démonstration. Comme T est continue de (X,w) dans (X,w), on sait déjà que $Tx_n \xrightarrow{w} Tx$. De plus, (x_n) est une *suite* w-convergente, donc elle est bornée. Comme T est compact, la suite (Tx_n) vit donc dans un compact de $(Y, \|\cdot\|)$. Mais comme $Tx_n \xrightarrow{w} Tx$, la seule valeur d'adhérence possible (en norme) pour la suite (Tx_n) est $y = Tx$; donc $Tx_n \xrightarrow{\|\cdot\|} Tx$. □

Proposition 3.13. *Soient X et Y des evn. Pour un opérateur $A \in \mathcal{L}(Y^*, X^*)$, les propriétés suivantes sont équivalentes.*

(i) *A est continu de (Y^*, w^*) dans (X^*, w^*).*

(ii) *A est un adjoint, i.e. existe $T \in \mathcal{L}(X,Y)$ tel que $A = T^*$;*

Démonstration. Supposons (ii) vérifiée. Alors, pour tout $x \in X$, la forme linéaire $y^* \mapsto \langle Ay^*, x\rangle$ est continue sur (Y^*, w^*) car $\langle Ay^*, x\rangle = \langle y^*, Tx\rangle$; donc (i) est vérifiée. Inversement, supposons (i) vérifiée. Alors, pour tout $x \in X$, la forme linéaire Φ_x : $y^* \mapsto \langle Ay^*, x\rangle$ est continue sur (Y^*, w^*). Donc, par le Lemme 3.9, il existe un unique $y = y_x \in Y$ tel que $\Phi_x = \delta_{y_x}$. Posons $Tx := y_x$. L'application $T : X \to Y$ est visiblement linéaire, et on a par définition

$$\forall y^* \in Y^*\ :\ \langle y^*, Tx\rangle = \langle Ay^*, x\rangle.$$

Cette identité montre que T est continue de (X,w) dans (X,w). Donc T est continue par la Proposition 3.11 ; et la même identité montre que $A = T^*$. □

3.5. Adhérence d'un convexe. On a vu que dans un evn, toute boule fermée est w-fermée. Cela se généralise à n'importe quel ensemble convexe :

Proposition 3.14. (Théorème de Mazur)
Soit X un espace vectoriel normé. Si $C \subseteq X$ est un ensemble convexe, *alors $\overline{C}^{w} = \overline{C}^{\|\cdot\|}$.*

Démonstration. Comme $w \subseteq \tau_{\|\cdot\|}$, on a $\overline{C}^{\|\cdot\|} \subseteq \overline{C}^{w}$. Pour l'inclusion inverse, on montre que $X \backslash \overline{C}^{\|\cdot\|} \subseteq X \backslash \overline{C}^{w}$. Soit $x \in X$ tel que $x \notin \overline{C}^{\|\cdot\|}$. D'après le Théorème de Hahn-Banach, on peut trouver $x^* \in X^*$ et $\alpha < \beta$ tels que

$$\mathrm{Re}\,\langle x^*, x\rangle \geqslant \alpha > \beta \geqslant \sup\left\{\mathrm{Re}\,\langle x^*, u\rangle;\ u \in C\right\}. \tag{3.1}$$

Si on avait $x \in \overline{C}^{w}$, on pourrait trouver une suite généralisée $(u_p) \subseteq C$ telle que $u_p \xrightarrow{w} x$. Alors $\langle x^*, u_p\rangle \to \langle x^*, x\rangle$, donc $\mathrm{Re}\,\langle x^*, u_p\rangle \to \mathrm{Re}\,\langle x^*, x\rangle$, ce qui est impossible par (3.1). □

Corollaire 3.15. *Soit X un espace vectoriel normé. Si $C \subseteq X$ est un ensemble convexe, alors C est fermé dans X si et seulement si C est* faiblement fermé*, i.e. fermé dans (X, w).*

Démonstration. C'est évident. $\square$

Corollaire 3.16. *Soit X un espace vectoriel normé. Pour tout ensemble $A \subseteq X$, on a $\overline{A}^{w} \subseteq \overline{\mathrm{conv}(A)}^{\|\cdot\|}$. Autrement dit : si $x \in \overline{A}^{w}$, alors il existe une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de $\mathrm{conv}(A)$ telle que $\|x_n - x\| \to 0$.*

Démonstration. C'est à nouveau évident : $\overline{A}^{w} \subseteq \overline{\mathrm{conv}(A)}^{w} = \overline{\mathrm{conv}(A)}^{\|\cdot\|}$. $\square$

Corollaire 3.17. *Soit X un evn, et soit $(x_k)_{k\in\mathbb{N}}$ une suite d'éléments de X. Si $x_k \xrightarrow{w} 0$, alors il existe une suite (z_n) formée de combinaisons convexes des x_k telle que $z_n \xrightarrow{\|\cdot\|} 0$.*

Démonstration. On applique 3.16 à $A := \{x_k;\ k \in \mathbb{N}\}$ et $x := 0$. $\square$

Remarque. Le Théorème de Mazur est un résultat spécifique à la topologie *faible* w : il n'est pas valable pour la topologie préfaible w^*.

Exemple. Soit K un espace topologique compact, et soit $(f_k)_{k\in\mathbb{N}}$ une suite de fonctions continues sur K. On suppose que la suite (f_k) est uniformément bornée, et que $f_k(t) \to 0$ pour tout $t \in K$. Alors il existe une suite (g_n) de combinaisons convexes des f_k telle que $g_n \to 0$ *uniformément* sur K.

Démonstration. Il suffit de montrer que $f_k \xrightarrow{w} 0$ dans l'espace de Banach $\mathcal{C}(K)$. Comme $\big(\mathcal{C}(K)\big)^* = M(K)$, cela revient à montrer que $\int_K f_k\, d\mu \to 0$ pour toute mesure borélienne complexe μ sur K ; et ceci est clair par convergence dominée puisque (f_k) est uniformément bornée et que $f_k \to 0$ simplement. $\square$

Exercice. Essayer de démontrer directement ce résultat.

3.6. Compacité préfaible. Vu l'utilité de la compacité, on conçoit que le résultat suivant soit particulièrement important.

Théorème 3.18. (Théorème de Banach-Alaoglu)
Si X est un espace vectoriel normé, alors toute boule fermée $B \subseteq X^*$ est w^*- compacte.

Démonstration. Par translation et dilatation, on peut supposer que B est la boule unité fermée B_{X^*}. Comme la topologie w^* est induite par la topologie de la convergence simple sur $\mathcal{F}(X, \mathbb{K}) = \mathbb{K}^X$, il s'agit de montrer que toute suite généralisée $(x_p^*) \subseteq B_{X^*}$ possède une sous-s.g. qui converge simplement vers un élément de B_{X^*}.
Si $x \in X$, alors $\forall p \in P\ :\ |\langle x_p^*, x\rangle| \leqslant \|x\|$. Donc la s.g. (x_p^*) est *simplement bornée*. Par le Théorème de Tychonoff, (x_p^*) possède une sous-s.g. $(z_{p'}^*)_{p'\in P'}$ qui converge simplement sur X vers une fonction $\Phi : X \to \mathbb{K}$. Comme les $z_{p'}^*$ sont des formes linéaires, Φ est une forme linéaire ; et comme $|\langle z_{p'}^*, x\rangle| \leqslant \|x\|$ pour tout p' et pour tout $x \in X$, on a $|\Phi(x)| \leqslant \|x\|$ pour tout $x \in X$, donc Φ est continue et $\|\Phi\| \leqslant 1$, *i.e.* $\Phi \in B_{X^*}$. Ceci termine la preuve du théorème. $\square$

Corollaire 3.19. *Si X est un espace vectoriel normé, alors toute suite bornée $(x_n^*)_{n\in\mathbb{N}} \subseteq X^*$ possède une* sous-s.g. *w^*- convergente. Si de plus X est* séparable*, alors toute suite bornée $(x_n^*) \subseteq X^*$ possède une* sous-suite *w^*- convergente.*

Démonstration. La 1ère partie est évidente ; et la 2ème partie vient du fait que les parties bornées de X^* sont w^*- métrisables si X est séparable. $\square$

EXEMPLE 1. Soit K un espace topologique compact métrisable. Si $(\mu_n)_{n\in\mathbb{N}}$ est une suite de mesures de probabilité boréliennes sur K, alors il existe une mesure de probabilité μ et une sous-suite (μ_{n_k}) de (μ_n) telles que

$$\forall f \in \mathcal{C}(K) \ : \ \int_K f\, d\mu_{n_k} \to \int_K f\, d\mu.$$

Démonstration. Les μ_n habitent dans $M(K) = \mathcal{C}(K)^*$, et comme elles sont $\geqslant 0$, on a $\|\mu_n\| = \mu_n(K) = 1$ pour tout $n \in \mathbb{N}$. D'après le Théorème de Banach-Alaoglu et comme $\mathcal{C}(K)$ est séparable, (μ_n) a une sous-suite (μ_{n_k}) qui converge w^* vers une mesure $\mu \in M(K)$; autrement dit $\int_K f\, d\mu_{n_k} \to \int_K f\, d\mu$ pour toute $f \in \mathcal{C}(K)$. La mesure μ est positive car $\int_K f\, d\mu = \lim \int_K f\, d\mu_{n_k} \geqslant 0$ pour toute $f \in \mathcal{C}_+(K)$; et on a $\mu(K) = \int_K \mathbf{1}\, d\mu = \lim \int_K \mathbf{1}\, d\mu_{n_k} = 1$, donc μ est une mesure de probabilité. $\square$

EXEMPLE 2. (Théorème de Banach-Mazur).
Tout espace de Banach séparable sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est isométrique à un sous-espace fermé de $\mathcal{C}([0,1], \mathbb{K})$.

Démonstration. Soit X un espace de Banach séparable. Alors $K := (B_{X^*}, w^*)$ est compact et métrisable. De plus, si $x \in X$, alors la fonction $Jx : K \to \mathbb{K}$ définie par $Jx(x^*) := \langle x^*, x\rangle$ est continue sur K, et on a $\|Jx\|_\infty = \sup_{x^*\in B_{X^*}} |\langle x^*, x\rangle| = \|x\|$. On a donc une isométrie linéaire $J : X \to \mathcal{C}(K)$; autrement dit, X est isométrique à un sous-espace de $\mathcal{C}(K)$, nécessairement fermé car X est complet. Il suffit donc de montrer que $\mathcal{C}(K)$ est isométrique à un sous-espace fermé de $\mathcal{C}([0,1])$. On va le faire en 2 étapes.

FAIT 1. $\mathcal{C}(K)$ se plonge isométriquement dans $\mathcal{C}(\Delta)$, où $\Delta = \{0,1\}^{\mathbb{N}}$.

Preuve du Fait 1. Comme K est métrisable, on sait qu'il existe une surjection continue $s : \Delta \twoheadrightarrow K$. On définit donc une isométrie linéaire $J : \mathcal{C}(K) \to \mathcal{C}(\Delta)$ en posant $Jf := f \circ s$. $\square$

FAIT 2. $\mathcal{C}(\Delta)$ se plonge isométriquement dans $\mathcal{C}([0,1])$.

Preuve du Fait 2. On sait que Δ est homéomorphe à un compact $E \subseteq \mathbb{R}$. Alors $\mathcal{C}(\Delta)$ et $\mathcal{C}(E)$ sont isométriques (exo), donc il suffit de montrer que $\mathcal{C}(E)$ se plonge isométriquement dans $\mathcal{C}([0,1])$. On peut supposer que $E \subseteq [0,1]$ avec $0 \in E$ et $1 \in E$ (exo). Alors $[0,1]\backslash E$ est un ouvert de $\mathbb{R}$, donc il est réunion d'une famille (dénombrable) d'intervalles ouverts deux à deux disjoints $J_i =]a_i, b_i[$, où $a_i, b_i \in E$. Pour toute fonction continue $f : E \to \mathbb{K}$, notons $Jf : [0,1] \to \mathbb{K}$ la fonction continue égale à f sur E et affine sur chaque intervalle $[a_i, b_i]$. L'application $f \mapsto Jf$ est visiblement linéaire. On a $\|Jf\|_\infty \geqslant \|f\|_\infty$ car Jf prolonge f. Inversement, si $t \in [0,1]\backslash E$, alors t est dans un intervalle $]a_i, b_i[$, donc $|Jf(t)| \leqslant \max\big(|Jf(a_i)|, |Jf(b_i)|\big) = \max\big(|f(a_i)|, |f(b_i)|\big) \leqslant \|f\|_\infty$ puisque Jf est affine sur $[a_i, b_i]$. Donc $\|Jf\|_\infty = \|f\|_\infty$ pour toute $f \in \mathcal{C}(E)$. On a ainsi défini une isométrie linéaire $J : \mathcal{C}(E) \to \mathcal{C}([0,1])$. $\square$

Par les Faits 1 et 2, la preuve est maintenant terminée. $\square$

4. Bidual, réflexivité

Rappel. Si X est un espace vectoriel normé, on considère que $X \subseteq X^{**}$ *via* le plongement canonique $j_X : X \hookrightarrow X^{**}$.

4.1. Densité d'un espace dans son bidual.

Lemme 4.1. (Théorème de Goldstine)
Si X est un espace vectoriel normé, alors B_X est dense dans $(B_{X^{**}}, w^{**})$.

Démonstration. Supposons que B_X ne soit pas dense dans $(B_{X^{**}}, w^{**})$, *i.e.* que $C := \overline{B_X}^{w^{**}} \neq B_{X^{**}}$. Soit $x^{**} \in B_{X^{**}}$ tel que $x^{**} \notin C$. Comme C est convexe et w^{**}-fermé et comme (X^{**}, w^{**}) est localement convexe, on peut trouver une forme linéaire continue $\Phi : (X^{**}, w^{**}) \to \mathbb{K}$ et $\alpha, \beta \in \mathbb{R}$ tels que

$$\operatorname{Re}\Phi(x^{**}) > \sup\left\{\operatorname{Re}\Phi(u^{**});\ u^{**} \in C\right\} \geqslant \sup\left\{\operatorname{Re}\Phi(x);\ x \in B_X\right\}.$$

D'après le Lemme 3.9, Φ est de la forme δ_{x^*} pour une certaine $x^* \in X^*$. On a donc

$$\operatorname{Re}\langle x^{**}, x^*\rangle > \sup\left\{\operatorname{Re}\langle x^*, x\rangle;\ x \in B_X\right\} = \|x^*\|;$$

ce qui est impossible car $\|x^{**}\| \leqslant 1$. □

4.2. Espaces réflexifs.

Définition 4.2. *Soit X un espace de Banach. On dit que X est **réflexif** si "$X^{**} = X$"; de façon précise : si le plongement canonique $j_X : X \hookrightarrow X^{**}$ est surjectif. Autrement dit, X est réflexif si et seulement si toute forme linéaire continue $\Phi : X^* \to \mathbb{K}$ est de la forme $\Phi = \delta_x$ pour un certain $x \in X$.*

Exemple 0. Tout espace de dimension finie est réflexif.

Démonstration. Si $\dim(X) < \infty$, alors $\dim(X^*) = \dim(X) < \infty$, donc $\dim(X^{**}) = \dim(X^*) = \dim(X)$, et donc $X^{**} = X$ puisque $X \subseteq X^{**}$. □

Exemple 1. Tout espace de Hilbert H est réflexif, car "$H^* = H$" et donc "$H^{**} = H^* = H$".

Exemple 2. Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré sigma-fini. Si $1 < p < \infty$, alors $L^p = L^p(\Omega, m)$ est réflexif, car "$(L^p)^* = L^q$" et $1 < q < \infty$, donc "$(L^p)^{**} = (L^q)^* = L^p$.

Exemple 3. L'espace $X := c_0(\mathbb{N})$ n'est pas réflexif car "$c_0^* = \ell^1$" et donc "$c_0^{**} = (\ell^1)^* = \ell^\infty \neq c_0$".

Exemple 4. Si K est un espace compact métrisable infini, alors $X := \mathcal{C}(K)$ n'est pas réflexif.

Démonstration. On sait que $\mathcal{C}(K)^* = M(K)$. Par ailleurs, comme K est infini, il existe une fonction borélienne bornée $g : K \to \mathbb{C}$ non continue (exo). Alors la forme linéaire $\Phi : M(K) \to \mathbb{C}$ définie par $\Phi(\mu) := \int_K g\, d\mu$ est continue, mais Φ n'est pas de la forme δ_f pour une certaine $f \in \mathcal{C}(K)$ car g n'est pas continue (exo). □

Remarque 4.3. Si X est un espace de Banach réflexif, alors toute forme linéaire continue $x^* \in X^*$ **atteint sa norme** : il existe $x \in X$ tel que $\|x\| = 1$ et $\langle x^*, x\rangle = \|x^*\|$. Inversement, on peut montrer que si toute forme linéaire continue $x^* \in X^*$ atteint sa norme, alors X est réflexif (*Théorème de James*).

Démonstration. D'après le Théorème de Hahn-Banach, on peut trouver $x^{**} \in X^{**}$ telle que $\|x^{**}\| = 1$ et $\langle x^{**}, x^* \rangle = \|x^*\|$; et $x^{**} \in X$ puisque X est réflexif. □

PROPOSITION 4.4. *Si X est un espace de Banach réflexif, alors tout sous-espace fermé $E \subseteq X$ est réflexif.*

Démonstration. On a $E^* = X^*/E^\perp$, donc $E^{**} = \big(E^\perp\big)^\perp$. Mais comme X est réflexif, $(E^\perp)^\perp$ est contenu dans X ; donc $(E^\perp)^\perp = (E^\perp)_\perp = \overline{E} = E$.

Preuve sans quotient. Soit $\Phi \in E^{**}$: on veut montrer que Φ est de la forme $\Phi = \delta_u$ pour un certain $u \in E$, autrement dit qu'il existe $u \in E$ tel que $\forall u^* \in E^* \ : \ \Phi(u^*) = \langle u^*, u \rangle$. Soit $\widetilde{\Phi} : X^* \to \mathbb{K}$ la forme linéaire définie par $\widetilde{\Phi}(x^*) := \Phi\big(x^*_{|E}\big)$, $x^* \in X^*$. La forme linéaire $\widetilde{\Phi}$ est continue (micro-exo), *i.e.* $\widetilde{\Phi} \in X^{**}$. Comme X est réflexif, il existe donc $u \in X$ tel que $\forall x^* \in X^* \ : \ \Phi\big(x^*_{|E}\big) = \langle x^*, u \rangle$. Il est *évident* par définition qu'on a $u \in (E^\perp)_\perp$. Donc $u \in \overline{E} = E$. Si $u^* \in E^*$ et si on note $x^* \in X^*$ un prolongement de Hahn-Banach de u^*, alors $\Phi(u) = \Phi(x^*_{|E}) = \langle x^*, u \rangle = \langle u^*, u \rangle$ puisque $u \in E$. Donc $\Phi = \delta_u$. □

COROLLAIRE 4.5. *Un espace de Banach X est réflexif si et seulement si X^* est réflexif.*

Démonstration. Si X est réflexif, alors $X^{**} = X$, donc $(X^*)^{**} = \big(X^{**}\big)^* = X^*$, et donc X^* est réflexif. Inversement, si X^* est réflexif, alors $X^{**} = (X^*)^*$ est réflexif par ce qui précède, et donc X est réflexif car X est un sous-espace fermé de X^{**}. □

PROPOSITION 4.6. *Un espace de Banach X est réflexif si et seulement si sa boule unité fermée B_X est w - compacte.*

Démonstration. Si X est réflexif, alors $B_X = B_{X^{**}}$, donc B_X est w^{**}- compacte d'après le Théorème de Banach-Alaoglu, et donc B_X est w - compacte car $w = w^{**}$ (puisque $X = X^{**}$). Inversement, supposons que B_X soit w - compacte. Alors B_X est w^{**}- compacte dans X^{**} puisque w est la topologie induite par w^{**} sur X ; donc B_X est w^{**}- *fermée* dans X^{**}. Donc $B_X = B_{X^{**}}$ d'après le Théorème de Goldstine ; et donc $X = X^{**}$. □

COROLLAIRE 4.7. *Si X est un espace de Banach réflexif et si $A \subseteq X$ est borné, alors $\overline{A}^w$ est w - compact.*

Démonstration. L'ensemble A est contenu dans une boule fermée B, qui est w - compacte car X est réflexif. Alors B est w - fermée, donc $\overline{A}^w \subseteq B$; donc $\overline{A}^w$ est une partie w - fermée de B, et donc $\overline{A}^w$ est w - compact puisque B est w-compacte. □

COROLLAIRE 4.8. *Si X est un espace de Banach réflexif, alors tout ensemble convexe fermé et borné $C \subseteq X$ est w - compact.*

Démonstration. L'ensemble C est convexe et fermé, donc il est w - fermé par le Théorème de Mazur ; donc $C = \overline{C}^w$ est w - compact par le corollaire précédent. □

COROLLAIRE 4.9. *Soit X un espace de Banach réflexif, et soit $C \subseteq X$ un ensemble convexe fermé. Si $\varphi : C \to \mathbb{R}$ est une fonction convexe continue telle que $\varphi(x) \to +\infty$ quand $\|x\| \to \infty$, alors φ possède un minimum sur C. En particulier, pour tout $a \in X$, on peut trouver un point $x \in C$ tel que $\|x - a\| = \mathrm{dist}(a, C)$.*

Démonstration. Posons $m := \inf\{\varphi(x);\ x \in C\}$, qui est bien défini mais peut-être égal à $-\infty$. Il suffit de montrer qu'il existe un $x \in C$ tel que $\varphi(x) \leqslant m$, ce qui prouvera en même temps que $m > -\infty$ et que m est atteint. Soit $(\alpha_n)_{n\in\mathbb{N}}$ une suite décroissante de nombres réels telle que $\alpha_n \to m$ et $\alpha_n > m$ pour tout $n \in \mathbb{N}$. Pour $n \in \mathbb{N}$, posons

$$C_n := \{x \in C;\ \varphi(x) \leqslant \alpha_n\}.$$

Par définition, les C_n sont tous non-vides, et ils sont fermés dans X car C est fermé et φ est continue. De plus, chaque C_n est borné car $\varphi(x) \to +\infty$ quand $\|x\| \to \infty$ (exo). Enfin, les C_n sont convexes car φ est convexe (autre exo). Donc les C_n sont *w - compacts.* Comme ils sont $\neq \varnothing$ et que la suite (C_n) est décroissante, on en déduit que $\bigcap_{n\in\mathbb{N}} C_n \neq \varnothing$; ce qui est exactement la conclusion souhaitée.
Si $a \in X$ est fixé et si on applique le résultat à la fonction $\varphi : C \to \mathbb{R}$ définie par $\varphi(x) := \|x - a\|$, on obtient la 2ème partie du corollaire. □

Corollaire 4.10. *Si X est un espace de Banach réflexif, alors toute suite bornée $(x_n)_{n\in\mathbb{N}}$ possède une sous-suite w - convergente.*

Démonstration. On peut supposer que $\|x_n\| \leqslant 1$ pour tout $n \in \mathbb{N}$, de sorte que $x_n \in B_X$.
Suppososons d'abord que X soit *séparable.* Alors X^* est séparable car $(X^*)^* = X$ est séparable ; donc B_X est w - métrisable. Comme B_X est w - compacte puisque X est réflexif, on a donc le résultat dans ce cas.
Dans le cas général, posons $E := \overline{\mathrm{vect}}\,\{x_n;\ n \in \mathbb{N}\}$. Alors E est un sous-espace fermé de X, donc E est réflexif ; et E est séparable par définition. D'après le 1er cas, (x_n) possède donc une sous-suite qui converge dans (E, w) ; et évidemment, cette sous-suite converge dans (X, w). □

Remarque. On peut montrer que la réciproque est vraie : si toute suite bornée $(x_n) \subseteq X$ possède une sous-suite w - convergente, alors X est réflexif. (*Théorème d'Eberlein-Smulian.*)

Corollaire 4.11. *Soient X et Y deux espaces de Banach, avec X réflexif. Pour un opérateur $T \in \mathcal{L}(X, Y)$, les propriétés suivantes sont équivalentes :*

(i) *T est compact ;*

(ii) *pour toute suite $(z_k)_{n\in\mathbb{N}} \subseteq X$ telle que $z_k \xrightarrow{w} 0$, on a que $\|Tz_k\| \to 0$.*

Démonstration. On a déjà vu que (i) $\Longrightarrow$ (ii), sans hypothèse sur X. Inversement, supposons (ii) vérifiée. Soit $(x_n)_{n\in\mathbb{N}}$ une suite bornée d'éléments de X : on veut montrer que la suite (Tx_n) possède une sous-suite convergente. Comme X est réflexif, (x_n) possède une sous-suite (x_{n_k}) qui converge faiblement, $x_{n_k} \xrightarrow{w} x \in X$. Alors $Tx_{n_k} \to y := Tx$ par (ii) appliqué à $z_k := x_{n_k} - x$. □

Exercice. Soient X et Y des espaces de Banach. On note `WOT` la topologie sur $\mathcal{L}(X, Y)$ engendrée par les semi-normes $p_{x,y^*}(T) := |\langle y^*, Tx\rangle|$, où $x \in X$ et $y^* \in Y^*$. (Cette topologie s'appelle la **topologie opératorielle faible**.) En termes de suites généralisées, $T_p \xrightarrow{\texttt{WOT}} T$ signifie donc que $T_p x \xrightarrow{w} Tx$ pour tout $x \in X$. Montrer que si Y est réflexif, alors la boule unité fermée de $\mathcal{L}(X, Y)$ est `WOT`-compacte.

4.3. Uniforme convexité.

Définition 4.12. *Soit X un espace de Banach. On dit que X est* ***uniformément convexe*** *si "la boule unité de X est bien ronde". De façon précise, X est uniformément convexe si, pour tout $\varepsilon > 0$, on peut trouver $\delta = \delta(\varepsilon) > 0$ tel que*

$$\forall x, y \in B_X \quad \textit{vérifiant } \|x - y\| \geqslant \varepsilon \ : \ \textit{on a} \quad \left\| \frac{x+y}{2} \right\| \leqslant 1 - \delta.$$

Remarque. Un espace X uniformément convexe est en particulier **strictement convexe** : si $x, y \in X$ vérifient $\|x\|, \|y\| \leqslant 1$ et $x \neq y$, alors $\|(1-\lambda)x + \lambda y\| < 1$ pour tout $\lambda \in]0,1[$.

Démonstration. Soient $x, y \in B_X$ avec $x \neq y$, et supposons que $\|(1-\lambda)x + \lambda y\| = 1$ pour un certain $\lambda \in]0,1[$. Comme $0 < \lambda < 1$ et $x \neq y$, on peut trouver $u, v \in [x, y]$ avec $u \neq y$ tels que $(1-\lambda)x + \lambda y = \frac{u+v}{2}$. Mais $u, v \in B_X$ car B_X est un ensemble convexe, donc $\left\| \frac{u+v}{2} \right\| \leqslant 1 - \delta(\|v - u\|) < 1$. On obtient donc une contradiction $\square$

Exercice. Montrer qu'un espace de Banach X est strictement convexe si et seulement si sa sphère unité S_X ne contient pas de segment non. trivial.

Exemple 1. Tout espace de Hilbert est uniformément convexe.

Démonstration. Soit H un espace de Hilbert. Si $x, y \in B_H$, alors

$$\|x - y\|^2 + \|x + y\|^2 = 2(\|x\|^2 + \|y\|^2) \leqslant 4.$$

Donc, si $x, y \in B_H$ et $\|x - y\| \geqslant \varepsilon$, alors

$$\left\| \frac{x+y}{2} \right\|^2 = \frac{1}{4} \|x + y\|^2 \leqslant 1 - \left\| \frac{x-y}{2} \right\|^2 \leqslant 1 - \frac{\varepsilon^2}{4}.$$

Par conséquent, on peut prendre $\delta(\varepsilon) := 1 - \sqrt{1 - \varepsilon^2/4}$. $\square$

Exemple 2. L'espace $X = (\mathbb{R}^2, \| \cdot \|_\infty)$ n'est pas uniformément convexe.

Démonstration. Si on pose $x := (1,0)$ et $y := (1,1)$, alors le segment $[x, y]$ est contenu dans S_X. Donc X n'est pas strictement convexe. $\square$

Remarque 4.13. Pour un espace de Banach X, les propriétés suivantes sont équivalentes.

(1) X est uniformément convexe.

(2) Pour toute suite $(x_n, y_n)_{n \in \mathbb{N}} \subseteq B_X \times B_X$ telle que $\left\| \frac{x_n + y_n}{2} \right\| \to 1$, on a que $\|x_n - y_n\| \to 0$.

Démonstration. Exo. $\square$

Proposition 4.14. *Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré. Si $1 < p < \infty$, alors $L^p = L^p(\Omega, m)$ est uniformément convexe.*

Démonstration. La preuve repose sur le fait suivant.

Fait 1. Soit $\alpha : \mathbb{C}^2 \to \mathbb{R}$ la fonction définie par

$$\alpha(u, v) := \frac{|u|^p + |v|^p}{2} - \left| \frac{u+v}{2} \right|^p.$$

(a) On a $\alpha(u,v) \geqslant 0$ pour tout $(u,v) \in \mathbb{C}^2$, et $\alpha(u,v) > 0$ si $u \neq v$.

(b) Pour tout $\eta > 0$, il existe une constante $C_\eta < \infty$ telle que l'implication suivante ait lieu : $|u-v|^p \geqslant \eta\left(|u|^p + |v|^p\right) \implies |u|^p + |v|^p \leqslant C_\eta\, \alpha(u,v)$.

Preuve du Fait 1. (a) Si $|u| \neq |v|$, alors $\alpha(u,v) \geqslant \frac{|u|^p+|v|^p}{2} - \left(\frac{|u|+|v|}{2}\right)^p > 0$ car la fonction $t \mapsto t^p$ est *strictement convexe* sur $\mathbb{R}^+$. Si $|u| = |v| := t$ et $u \neq v$, alors $|u+v| < |u| + |v| = 2t$, donc $\alpha(u,v) > \frac{2t^p}{2} - t^p = 0$. Et si $u = v$, alors $\alpha(u,v) = |u|^p - |u|^p = 0$.
(b) L'ensemble $K_\eta := \left\{(x,y) \in \mathbb{C}^2;\ |x|^p + |y|^p = 1\ ,\ |x-y|^p \geqslant \eta\right\}$ est compact car fermé borné dans $\mathbb{C}^2$, et $\alpha(u,v) > 0$ sur K par (a). Donc il existe une constante $\theta = \theta_\eta > 0$ telle que $\forall (x,y) \in K_\eta\ :\ \alpha(x,y) \geqslant \theta$. Si maintenant $(u,v) \in \mathbb{C}^2$ vérifie $|u-v|^p \geqslant \eta(|u|^p + |v|^p)$ et $(u,v) \neq (0,0)$ et si on pose $M := \|(u,v)\|_p = (|u|^p + |v|^p)^{1/p}$, alors $(x,y) := \left(\frac{u}{M}, \frac{v}{M}\right)$ appartient à K. Comme $\alpha(x,y) = \frac{1}{M^p}\,\alpha(u,v)$, on en déduit $\alpha(u,v) \geqslant \theta\,(|u|^p + |v|^p)$; donc $C_\eta := 1/\theta$ convient. □

Remarque. Il découle en particulier du Fait 1 que

$$\forall u, v \in \mathbb{C}\ :\ |u+v|^p \leqslant 2^{p-1}\big(|u|^p + |v|^p\big).$$

Démonstration. On a $|u+v|^p = 2^p\left(\frac{|u|^p+|v|^p}{2} - \alpha(u,v)\right)$ et $\alpha(u,v) \geqslant 0$. □

FAIT 2. Si $f, g \in B_{L^p}$, alors $\int_\Omega \alpha\big(f(t), g(t)\big)\, dm(t) \leqslant 1 - \left\|\frac{f+g}{2}\right\|_p^p$.

Preuve du Fait 2. On a $\int_\Omega \alpha\big(f(t), g(t)\big)\, dm(t) = \frac{\|f\|_p^p + \|g\|_p^p}{2} - \left\|\frac{f+g}{2}\right\|_p^p$. □

Soit maintenant (f_n, g_n) une suite dans $B_{L^p} \times B_{L^p}$ telle que $\left\|\frac{f_n+g_n}{2}\right\|_p \to 1$. Il s'agit de montrer que $\|f_n - g_n\|_p \to 0$.
Soit $\eta > 0$ à choisir. Fixons également $n \in \mathbb{N}$, et posons

$$E := \left\{t \in \Omega;\ |f_n(t) - g_n(t)|^p \geqslant \eta\,(|f_n(t)|^p + |g_n(t)|^p)\right\}.$$

Par définition de E, on a

$$\begin{aligned}
\|f_n - g_n\|_p^p &= \int_E |f_n - g_n|^p dm + \int_{\Omega\setminus E} |f_n - g_n|^p dm \\
&\leqslant \int_E |f_n - g_n|^p dm + \eta \int_{\Omega\setminus E} (|f_n|^p + |g_n|^p)\, dm \\
&\leqslant \int_E |f_n - g_n|^p dm + 2\eta \qquad \text{car } \|f_n\|_p, \|g_n\|_p \leqslant 1.
\end{aligned}$$

Par ailleurs, on a aussi (d'après le Fait 1)

$$\begin{aligned}
\forall t \in E\ :\ |f_n(t) - g_n(t)|^p &\leqslant 2^{p-1}\left(|f_n(t)|^p + |g_n(t)|^p\right) \\
&\leqslant 2^{p-1} C_\eta\, \alpha\big(f_n(t), g_n(t)\big).
\end{aligned}$$

Comme la fonction α est $\geqslant 0$, on en déduit

$$\begin{aligned}
\int_E |f_n - g_n|^p dm &\leqslant 2^{p-1} C_\eta \int_\Omega \alpha\big(f_n(t), g_n(t)\big)\, dm(t) \\
&\leqslant 2^{p-1} C_\eta \times \left(1 - \left\|\frac{f_n + g_n}{2}\right\|_p^p\right) \qquad \text{d'après le Fait 2.}
\end{aligned}$$

On a donc montré que

$$\forall n \in \mathbb{N} \;:\; \|f_n - g_n\|_p^p \leqslant 2\eta + 2^{p-1} C_\eta \times \left(1 - \left\|\frac{f_n + g_n}{2}\right\|_p^p\right).$$

Comme $\left\|\frac{f_n+g_n}{2}\right\|_p \to 1$, on en déduit $\overline{\lim}\,\|f_n - g_n\|_p^p \leqslant 2\eta$, pour tout $\eta > 0$. Donc $\|f_n - g_n\|_p \to 0$. □

THÉORÈME 4.15. *Tout espace de Banach uniformément convexe est réflexif.*

Démonstration. Soit X un espace de Banach uniformément convexe. Pour montrer que X est réflexif, il suffit de montrer que si $x^{**} \in X^{**}$ et $\|x^{**}\| = 1$, alors en fait $x^{**} \in X$.
Par le Théorème de Goldstine, on peut trouver une suite généralisée $(x_p)_{p\in P} \subseteq B_X$ telle que $x_p \xrightarrow{w^{**}} x^{**}$. On va montrer qu'en fait $\|x_p - x^{**}\| \to 0$, ce qui prouvera que $x^{**} \in X$ puisque X, étant complet, est un sous-espace fermé de X^{**}.
Soit $\varepsilon > 0$. Comme $\|x^{**}\| = 1$, on peut trouver $x^* \in B_X^*$ tel que $\mathrm{Re}\langle x^{**}, x^*\rangle > 1 - \delta(\varepsilon)$, où $\delta(\varepsilon)$ est donné par la définition de l'uniforme convexité ; et comme $x_p \xrightarrow{w^{**}} x^{**}$, on peut ensuite trouver $p_0 \in P$ tel que $\forall p \succeq p_0 \;:\; \mathrm{Re}\langle x^*, x_p\rangle > 1 - \delta(\varepsilon)$. Comme $\|x^*\| \leqslant 1$, on a alors $\forall p, p' \succeq p_0 \;:\; \left\|\frac{x_p + x_{p'}}{2}\right\| \geqslant \mathrm{Re}\left\langle x^*, \frac{x_p+x_{p'}}{2}\right\rangle > 1 - \delta(\varepsilon)$; et donc, par définition de $\delta(\varepsilon)$:

$$\forall p, p' \succeq p_0 \;:\; \|x_p - x_{p'}\| \leqslant \varepsilon.$$

Comme $x_{p'} \xrightarrow{w^{**}} x^{**}$ et comme la boule fermée $\overline{B}(x_p, \varepsilon) \subseteq X^{**}$ est w^{**}- fermée dans X^{**}, on en déduit

$$\forall p \succeq p_0 \;:\; \|x_p - x^{**}\| \leqslant \varepsilon;$$

ce qui termine la preuve. □

Remarque. Voici une preuve presque identique mais un peu plus “habillée”. On considère l'ensemble préordonné filtrant $P \times P$ (avec l'ordre produit). On a

$$\forall x^* \in B_{X^*} \;:\; 1 \geqslant \overline{\lim} \left\|\frac{x_p + x_{p'}}{2}\right\| \geqslant \mathrm{Re}\left\langle x^*, \frac{x_p + x_{p'}}{2}\right\rangle = \mathrm{Re}\langle x^{**}, x^*\rangle.$$

En prenant le “sup en x^*”, on en déduit que $\left\|\frac{x_p+x_{p'}}{2}\right\| \to 1$. Donc $\|x_p - x_{p'}\| \to 0$ par uniforme convexité. Par conséquent, la s.g. $(x_p)_{p\in P}$ est *de Cauchy*. Donc (x_p) converge dans X car X est complet, nécessairement vers x^{**} puisque $x_p \xrightarrow{w^{**}} x^*$; et donc $x^{**} \in X$.

EXEMPLE. Retour sur le dual de L^p.

Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré, et soit $1 < p < \infty$. Montrons à nouveau que “le dual de L^p est L^q”.
Soit $J : L^q \to (L^p)^*$ le plongement canonique de L^q dans $(L^p)^*$: si $g \in L^q$, alors Jg est la forme linéaire $\Phi_g : L^p \to \mathbb{C}$ définie par $\Phi_g(f) := \int_\Omega fg\, dm$. On sait que J est une isométrie ; donc $J(L^q)$ est un sous-espace fermé de $(L^p)^*$. On voudrait montrer que

$J(L^q) = (L^p)^*$. D'après le critère dual de densité, il suffit de montrer que $J(L^q)^\perp = \{0\}$. Mais L^p est réflexif car uniformément convexe ; donc

$$J(L^q)^\perp = J(L^q)_\perp = \Big\{ f \in L^p \ : \ \forall g \in L^q \ : \ \int_\Omega fg\, dm = 0 \Big\}.$$

Autrement dit, $J(L^q)^\perp = \{f \in L^p;\ \Phi_f = 0\}$, où $\Phi_f : L^q \to \mathbb{C}$ est la forme linéaire définie par $\Phi_f(g) := \int_\Omega gf\, dm$. Donc $J(L^q)^\perp = \{0\}$ puisque l'application $f \mapsto \Phi_f$ est une isométrie.

5. Points extrémaux

5.1. Définition et exemples.

DÉFINITION 5.1. *Soit X un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et soit $C \subseteq X$. On dit qu'un point $e \in C$ est un* ***point extrémal*** *de C s'il n'est pas possible d'écrire e comme combinaison convexe de deux points de C différents de e ; autrement dit : e extrémal si et seulement si, pour tout segment $[u, v]$ a extrémités dans C tel que $e \in [u, v]$, on a $u = e$ ou $v = e$. On note $\operatorname{ext}(C)$ l'ensemble des points extrémaux de C.*

Remarque. En général, on parle de points extrémaux uniquement pour les ensembles convexes ; mais la définition a un sens pour un ensemble C quelconque.

REFORMULATION. Soit $C \subseteq X$, et soit $e \in C$. Alors $e \in \operatorname{ext}(C)$ si et seulement si l'implication suivante a lieu :

$$\Big(e = (1-\lambda)u + \lambda v \text{ avec } u, v \in C \text{ et } 0 < \lambda < 1 \Big) \implies u = v = e.$$

Si on suppose de plus que C est convexe, alors $e \in \operatorname{ext}(C)$ si et seulement si l'implication suivante a lieu :

$$e = \frac{u+v}{2} \text{ avec } u, v \in C \implies u = e = v.$$

Démonstration. Exo. □

REMARQUE. Si X est un espace vectoriel normé, alors $\operatorname{ext}(C) \subseteq \partial C$ pour tout ensemble $C \subseteq X$. En particulier, $\operatorname{ext}(B_X) \subseteq S_X$.

Démonstration. Il est évident géométriquement qu'un point extrémal de C ne peut pas appartenir à $\mathring{C}$. Écrire les détails. □

EXEMPLE 1. Prenons $X := \mathbb{R}^2$.

- Si $C \subseteq \mathbb{R}^2$ est un triangle ou un rectangle, alors les points extrémaux de C sont ses sommets.
- Si $C \subseteq \mathbb{R}^2$ est un disque, alors $\operatorname{ext}(C) = \partial C$.

Démonstration. On le voit bien. □

EXEMPLE 2. Si X est un espace de Banach strictement convexe, alors $\operatorname{ext}(B_X) = S_X$.

Démonstration. C'est évident. □

EXEMPLE 3. Les points extrémaux de $B_{M_d(\mathbb{R})} = B_{\mathcal{L}(\mathbb{R}^d)}$ sont les matrices orthogonales.

Démonstration. On identifie $M_d(\mathbb{R})$ et $\mathcal{L}(\mathbb{R}^d)$; donc $O_d(\mathbb{R}) \subseteq \mathcal{L}(\mathbb{R}^d)$ est l'ensemble des isométries linéaires de $\mathbb{R}^d$.
Montrons d'abord que toute isométrie J est un point extrémal de $B_{\mathcal{L}(\mathbb{R}^d)}$. Soient $A, B \in B_{\mathcal{L}(\mathbb{R}^d)}$ tels que $J = (1-\lambda)A + \lambda B$ avec $0 < \lambda < 1$. Si $x \in \mathbb{R}^d$ vérifie $\|x\| = 1$, alors

$$1 = \|x\|^2 = \|Jx\|^2 = \|(1-\lambda)Ax + \lambda Bx\|^2 \qquad \text{et} \qquad \|Ax\|, \|Bx\| \leqslant 1.$$

Comme la fonction $u \mapsto \|u\|^2$ est strictement convexe sur $\mathbb{R}^d$, cela n'est possible que si $Ax = Bx$, pour tout $x \in \mathbb{R}^d$ de norme 1. Donc $A = B$, et donc $A = J = B$.
Inversement, soit A un point extrémal de $B_{\mathcal{L}(\mathbb{R}^d)}$ et montrons que A est une isométrie. D'après la décomposition polaire, on peut écrire $A = UP$, où U est une isométrie et P est un opérateur positif. Il suffit de montrer que $P = I$.
Comme A est un point extrémal de $B_{\mathcal{L}(\mathbb{R}^d)}$, il en va de même pour P car U est une isométrie (exo). De plus, P est diagonalisable en base orthonormée et ses valeurs propres sont comprises entre 0 et 1 puisque P est un opérateur positif et $\|P\| \leqslant 1$. On peut donc trouver une isométrie J telle que $D := J^{-1}PJ$ soit un opérateur diagonal dans la base canonique de $\mathbb{R}^d$, avec des coefficient diagonaux $\lambda_i \in [0,1]$. Comme P est extrémal et comme J est une isométrie, l'opérateur diagonal D est un point extrémal de $B_{\mathcal{L}(\mathbb{R}^d)}$. Ceci n'est possible que si $\lambda_i = 1$ pour tout $i \in [\![1, d]\!]$. En effet, si on avait $\lambda_i < 1$ pour un certain i, on pourrait écrire $\lambda_i = \frac{\alpha_1+\alpha_2}{2}$ avec $-1 \leqslant \alpha_1 < \lambda_i < \alpha_2 \leqslant 1$. On aurait alors $D = \frac{D_1+D_2}{2}$, où D_1 et D_2 sont les opérateurs diagonaux obtenus à partir de D en remplaçant λ_i par α_1 et α_2 ; donc D ne serait pas un point extrémal de $B_{\mathcal{L}(\mathbb{R}^d)}$. Ainsi, on a $\lambda_i = 1$ pour tout $i \in [\![1, d]\!]$, donc $D = I$, et donc $P = I$ puisque $P = JDJ^{-1}$. □

EXEMPLE 4. Soit Ω un espace topologique ou bien métrisable séparable, ou bien compact, et soit $M(\Omega)$ l'espace des mesures boréliennes complexes sur Ω (supposées régulières si Ω est compact). Les points extrémaux de $B_{M(\Omega)}$ sont les mesures de la forme $\mu = c\,\delta_a$, où $a \in \Omega$ et $|c| = 1$.

Démonstration. On a besoin du fait général suivant.

FAIT 5.2. *Si $\mu \in M(\Omega)$, il existe un* plus grand *ouvert O tel que $|\mu|(O) = 0$. Le complémentaire de cet ouvert O s'appelle le* ***support*** *de la mesure μ, et se note* $\mathrm{supp}(\mu)$.

Preuve du Fait. Soit $(V_i)_{i\in I}$ la famille de tous les ouverts $V \subseteq \Omega$ tels que $|\mu|(V) = 0$, et soit $O := \bigcup_{i\in I} V_i$. Alors O est certainement un ouvert de Ω, et il suffit visiblement de montrer que $|\mu|(O) = 0$.
Si Ω est métrisable séparable, on peut trouver un ensemble *dénombrable* $J \subseteq I$ tel que $O = \bigcup_{i\in J} V_i$ (propriété de Lindelöf). Donc $|\mu|(O) = 0$ puisque $|\mu|$ est une mesure.
Si Ω est compact, on utilise la *régularité* de la mesure $|\mu|$: on a $|\mu|(K) = 0$ pour tout compact $K \subseteq O$ car K est contenu dans une réunion *finie* d'ouverts V_i ; donc $|\mu|(O) = 0$ par régularité. □

Remarque 1. Soit $\mu \in M(\Omega)$. Un point $a \in \Omega$ appartient au support de μ si et seulement si $|\mu|(V) > 0$ pour tout voisinage V de a dans Ω.

Démonstration. C'est évident par définition du support. □

Remarque 2. Une mesure $\mu \in M(\Omega)$ non nulle est de la forme $\mu = c\,\delta_a$ si et seulement si $\mathrm{supp}(\mu)$ est réduit à 1 point.

Démonstration. Exo. □

Soit μ un point extrémal de $B_{M(\Omega)}$. On a nécessairement $\|\mu\| = 1$. Donc, pour montrer que μ et de la forme $c\,\delta_a$ avec $|c| = 1$, il suffit de montrer que $\mathrm{supp}(\mu)$ est réduit à 1 point.
Supposons que $\mathrm{supp}(\mu)$ contiennent au moins 2 points a_1 et a_2. Soient V_1 et V_2 deux ouverts de Ω tels que $a_i \in V_i$ et $V_1 \cap V_2 = \varnothing$. Comme a_1 et a_2 appartiennent au support de μ, on a $|\mu|(V_1) > 0$ et $|\mu|(V_2) > 0$. Donc, si on pose $E := V_1$, alors $\alpha := |\mu|(E) > 0$ et $\beta := |\mu|(E^c) > 0$. Comme de plus $|\mu|(\Omega) = \|\mu\| = 1$, on a $\alpha + \beta = 1$; donc $0 < \alpha < 1$ et $\beta = 1 - \alpha$. Si on pose $\mu_1 := \frac{1}{\alpha}\,\mathbf{1}_E\,\mu$ et $\mu_2 := \frac{1}{1-\alpha}\mathbf{1}_{E^c}\,\mu$, alors $\|\mu_1\| = 1 = \|\mu_2\|$ et $\mu = \alpha\,\mu_1 + (1-\alpha)\,\mu_2$; ce qui contredit l'extrémalité de μ puisque $\mu_1 \neq \mu_2$.
Le fait que toute mesure de la forme $c\,\delta_a$ avec $|c| = 1$ soit un point extrémal de $B_{M(\Omega)}$ est laissé en exo. □

Remarque. Notons $\mathbf{P}(\Omega)$ l'ensemble des mesures *de probabilité* sur Ω. On montre essentiellement de la même façon que les points extrémaux de $\mathbf{P}(\Omega)$ sont les masses de Dirac δ_a, $a \in \Omega$.

Exemple 5. La boule unité de $c_0(\mathbb{N})$ et la boule unité de $L^1([0,1])$ n'ont pas de points extrémaux.

Démonstration. Soit $x = (x_i)_{i\in\mathbb{N}} \in B_{c_0}$ quelconque. Comme $x \in c_0$, on peut choisir $i \in \mathbb{N}$ tel que $|x_i| < 1$, puis $\alpha, \beta \in \mathbb{C}$ de module < 1 tels que $\alpha \neq \beta$ et $x_i = \frac{\alpha+\beta}{2}$. Alors $x = \frac{u+v}{2}$, où u et v sont obtenus à partir de x en remplaçant la coordonnée x_i respectivement par α et par β. Comme u et v sont dans B_{c_0} et $u \neq v$, cela montre que x n'est pas un point extrémal de B_{c_0}. Ainsi, $\mathrm{ext}(B_{c_0}) = \varnothing$.
Soit maintenant $f \in B_{L^1}$ quelconque, et supposons que $\|f\|_1 = 1$. Comme la mesure de Lebesgue m ne charge pas les points, le support de la mesure $|f|m$ n'est pas réduit à 1 point. Comme dans la preuve de l'Exemple 3, on en déduit qu'on peut trouver un ensemble mesurable $E \subseteq [0,1]$ tel que $\alpha := \int_E |f| > 0$ et $1 - \alpha = \int_{E^c} |f| > 0$. Si on pose $u := \frac{1}{\alpha}\,\mathbf{1}_E f$ et $v := \frac{1}{1-\alpha}\,\mathbf{1}_{E^c} f$, alors $\|u\|_1 = 1 = \|v\|_1$ et $\alpha u + (1-\alpha)v = f$. Comme $u \neq v$, cela montrer que f n'est pas un point extrémal de B_{L^1}. Donc $\mathrm{ext}(B_{L^1}) = \varnothing$. □

Exercice. Soient X et Y des espaces vectoriels normés. Montrer que si $J : X \to Y$ est une isométrie bijective (linéaire), alors $J\big(\mathrm{ext}(B_X)\big) = \mathrm{ext}(B_Y)$.

5.2. Isométries entre espaces $\mathcal{C}(K)$. Le résultat suivant montre que si K est un espace topologique compact, alors l'espace de Banach $\mathcal{C}(K)$ "détermine entièrement K". Le lien avec les points extrémaux n'est pas immédiatement apparent. En fait, c'est la *preuve* du théorème qui utilise la notion de point extrémal.

Théorème 5.3. (Banach-Stone)
Soient K et L deux espaces topologiques compacts. Si les espaces de Banach $\mathcal{C}(K)$ et $\mathcal{C}(L)$ sont isométriques, alors K et L sont homéomorphes. De plus, si $J : \mathcal{C}(K) \to \mathcal{C}(L)$ est une isométrie bijective, alors J est nécessairement de la forme

$$Jf(y) = c(y)\,f(\phi(y)),$$

où $\phi : L \to K$ est un homéomorphisme et $c : L \to \mathbb{C}$ est une fonction continue telle que $|c(y)| \equiv 1$.

Démonstration. Il suffit évidemment de démontrer la partie "De plus".
On aura besoin du fait suivant. Rappelons que si Ω est un espace topologique compact, alors on identifie $\mathcal{C}(\Omega)^*$ et $M(\Omega)$ *via* le Théorème de représentation de Riesz.

FAIT. Si Ω est un espace topologique compact, alors l'application $z \mapsto \delta_z$ est un homéomorphisme de Ω sur $\{\delta_z;\ z \in \Omega\} \subseteq (M(\Omega), w^*)$.

Preuve du Fait. Si $z_p \to z$, alors $\langle \delta_{z_p}, f\rangle = f(z_p) \to f(z) = \langle \delta_z, f\rangle$ pour toute $f \in \mathcal{C}(\Omega)$; donc l'application $z \mapsto \delta_z$ est continue de Ω dans $(M(\Omega), w^*)$. Cette application est de plus injective (si $z_1 \neq z_2$, alors $\mathrm{supp}(\delta_{z_1}) = \{z_1\} \neq \{z_2\} = \mathrm{supp}(\delta_{z_2})$) ; donc c'est un homéomorphisme de Ω sur son image puisque Ω est compact. $\square$

Soit $J : \mathcal{C}(K) \to \mathcal{C}(L)$ une isométrie bijective. Pour montrer que J est de la forme annoncée, l'idée de départ est d'écrire que si $f \in \mathcal{C}(K)$, alors

$$\forall y \in L\ :\ Jf(y) = \langle \delta_y, f\rangle = \langle J^*\delta_y, f\rangle.$$

On va donc s'intéresser de plus près à l'opérateur J^*.
Comme J est une isométrie bijective, $J^* : M(K) \to M(L)$ est inversible avec $\|J^*\| = \|J\| = 1$ et $\|(J^*)^{-1}\| = \|(J^{-1})^*\| = \|J^{-1}\| = 1$. Donc J^* est une isométrie bijective. Par conséquent, on a $J^*\big[\mathrm{ext}\big(B_{M(L)}\big)\big] = \mathrm{ext}\big(B_{M(K)}\big)$. D'après la description des points extrémaux de $B_{M(\Omega)}$, on en déduit que pour tout $y \in L$, il existe un point $\phi(y) \in X$ et un nombre $c(y) \in \mathbb{C}$ de module 1 tels que

$$J^*\delta_y = c(y)\,\delta_{\phi(y)}.$$

Comme $c(y) = \langle c(y)\delta_{\phi(y)}, \mathbf{1}\rangle = \langle J^*\delta_y, \mathbf{1}\rangle$ et comme J^* est w^*-w^* continu, on voit que la fonction $y \mapsto c(y)$ est continue sur L. Donc l'application $y \mapsto \delta_{\phi(y)}$ est continue de L dans $(M(K), w^*)$ puisque $\delta_{\phi(y)} = \frac{1}{c(y)}\,J^*\delta_y$; et donc l'application $y \mapsto \phi(y)$ est continue par le Fait.
De même, il existe une fonction continue $d : K \to \mathbb{C}$ et une application continue $\psi : K \to (M(L), w^*)$ telles que $\forall x \in K\ :\ (J^*)^{-1}(\delta_x) = d(x)\,\delta_{\psi(x)}$. Donc $\delta_y = (J^*)^{-1}J^*\delta_y = c(y)d(\phi(y))\,\delta_{\psi(\phi(y))}$ pour tout $y \in L$, ce qui entraine que $\psi(\phi(y)) = y$. Ainsi, on a $\psi \circ \phi = id_Y$; et de même $\phi \circ \psi = id_X$. Donc ϕ est un homéomorphisme de L sur K, avec $\phi^{-1} = \psi$.
Si maintenant $f \in \mathcal{C}(K)$, alors $Jf(y) = \langle J^*\delta_y, f\rangle = c(y)\langle \delta_{\phi(y)}, f\rangle = c(y)\,f(\phi(y))$ pour tout $y \in L$, ce qui termine la preuve. $\square$

5.3. Convexes compacts. Le résultat suivant est spectaculaire et a de nombreuses applications.

THÉORÈME 5.4. (Krein-Milman)
Soit X un espace vectoriel topologique localement convexe. Si $K \subseteq X$ est un ensemble compact non vide, alors $K \subseteq \overline{\mathrm{conv}}\,\big(\mathrm{ext}(K)\big)$, et donc $K = \overline{\mathrm{conv}}\,\big(\mathrm{ext}(K)\big)$ si K est de plus convexe. En particulier, $\mathrm{ext}(K) \neq \varnothing$.

Remarque. La notation $\overline{\mathrm{conv}}(A)$ est synonyme de $\overline{\mathrm{conv}(A)}$.

Ce théorème va découler presque immédiatement du lemme suivant.

LEMME 5.5. *Soit X un evt localement convexe, et soit $K \subseteq X$ un compact non vide. Si $\varphi : X \to \mathbb{R}$ est une fonction convexe continue, alors $\varphi_{|K}$ atteint son maximum en un point extrémal de K, i.e. il existe un point $e \in \mathrm{ext}(K)$ tel que $\varphi(e) = \max_K \varphi$.*

Démonstration. On utilisera la terminologie suivante. Si $C \subseteq X$, on dira qu'un ensemble $F \subseteq C$ est une **partie extrémale de** C si l'implication suivante a lieu :

$$\Big(u, v \in C \text{ et } (1-\lambda)u + \lambda v \in F \text{ pour un certain } \lambda \in]0,1[\Big) \implies \big(u \in F \text{ et } v \in F\big).$$

Avec cette terminologie, on voit qu'un point $e \in C$ est un point extrémal de C si et seulement si $\{e\}$ est une partie extrémale de C.

Fait 1. Soit $C \subseteq X$ un ensemble compact non vide. Si $\varphi : X \to \mathbb{R}$ est une fonction convexe continue, alors l'ensemble $F(\varphi, C) := \{x \in C;\ \varphi(x) = \max_C \varphi\}$ est une partie extrémale de C, et $F(\varphi, C) \neq \varnothing$.

Preuve du Fait 1. La définition de $F(\varphi, C)$ a un sens et $F(\varphi, C) \neq \varnothing$, car φ est continue et C est compact. Posons $M := \max_C \varphi$. Soient $u, v \in C$ tels que $(1-\lambda)u+\lambda v \in F(\varphi, C)$ pour un certain $\lambda \in]0,1[$. Comme φ est convexe, on a $M = \varphi\big((1-\lambda)u+\lambda v\big) \leqslant (1-\lambda)\varphi(u) + \lambda\varphi(v) \leqslant (1-\lambda)M + \lambda M = M$. Donc $\varphi(u) = M = \varphi(v)$ car $0 < \lambda < 1$, *i.e.* $u, v \in F(\varphi, C)$. □

Fait 2. Soit $C \subseteq X$, et soit $F \subseteq C$ une partie extrémale de C. Si $F' \subseteq F$ est une partie extrémale de F, alors F' est une partie extrémale de C.

Preuve du Fait 2. C'est un exo facile. □

Fait 3. Si $M \subseteq X$ est compact et non vide, alors $\mathrm{ext}(M) \neq \varnothing$.

Preuve du Fait 3. Notons $\mathcal{F}$ la famille de toutes les parties extrémales de M compactes et non vides. Une application du Lemme de Zorn montre qu'il existe une $F_0 \in \mathcal{F}$ qui est minimale pour l'inclusion (exo ; il faut observer que si $(F_i)_{i\in I}$ est une famille totalement ordonnée de compacts non vides, alors $\bigcap_{i\in I} F_i \neq \varnothing$).
Supposons que F_0 contienne au moins 2 points distincts a et b. Comme X est localement convexe, on peut trouver une forme $\mathbb{R}$-linéaire continue $\varphi : X \to \mathbb{R}$ telle que $\varphi(a) \neq \varphi(b)$. Alors $F_1 := F(\varphi, F_0)$ ne peut pas contenir a et b, donc F_1 est strictement contenue dans F_0. De plus, $F_1 \neq \varnothing$ et F_1 est une partie extrémale de F_0 par le Fait 1. Donc F_1 est une partie extrémale de M par le Fait 2 ; ce qui contredit la minimalité de F_0. Ainsi F_0 est réduit à 1 point $\{e\}$, et $e \in \mathrm{ext}(M)$ puisque F_0 est une partie extrémale de M. □

Soit maintenant $K \subseteq X$ un ensemble compact non vide, et soit $\varphi : X \to \mathbb{R}$ convexe continue. Par le Fait 3, l'ensemble $M := F(\varphi, K)$ possède au moins 1 point extrémal e. Alors $\{e\}$ est une partie extrémale de M, donc une partie extrémale de K par les Faits 1 et 2 ; donc $e \in \mathrm{ext}(K)$. Ceci termine la preuve du lemme. □

Preuve du Théorème de Krein-Milman. Soit $K \subseteq X$ un ensemble compact non vide, et soit $x \in K$ quelconque. Supposons que $x \notin \overline{\mathrm{conv}}\big(\mathrm{ext}(K)\big)$. Alors, par le Théorème de séparation des convexes, on peut trouver une forme $\mathbb{R}$-linéaire continue $\varphi : X \to \mathbb{R}$ telle que $\varphi(x) > \sup\big\{\varphi(z);\ z \in \overline{\mathrm{conv}}\big(\mathrm{ext}(K)\big)\big\}$. En particulier, $\varphi(x) > \varphi(e)$ pour tout $e \in \mathrm{ext}(K)$. Mais par le Lemme 5.5, il existe un point $e \in \mathrm{ext}(K)$ tel que $\varphi(e) = \max_K \varphi$. On a donc $\varphi(e) \geqslant \varphi(x) > \varphi(e)$, ce qui pose problème. □

Corollaire 5.6. *Toute matrice $M \in M_d(\mathbb{R})$ vérifiant $\|M\| \leqslant 1$ est combinaison convexe de matrices orthogonales.*

Démonstration. On a vu que $\mathrm{ext}\big(B_{M_d(\mathbb{R})}\big) = O_d(\mathbb{R})$. Donc $M \in \overline{\mathrm{conv}}(O_d(\mathbb{R}))$ par Krein-Milman. Mais $O_d(\mathbb{R})$ est un compact de $M_d(\mathbb{R})$, *et donc* $\mathrm{conv}(O_d(\mathbb{R}))$ est compact car on est en dimension finie. En particulier $\mathrm{conv}(O_d(\mathbb{R}))$ est un fermé de $M_d(\mathbb{R})$; donc $M \in \mathrm{conv}(O_d(\mathbb{R}))$, ce qui est le résultat annoncé. □

COROLLAIRE 5.7. *Si X est un espace vectoriel normé, alors B_{X^*} est l'enveloppe convexe w^*-fermée de ses points extrémaux. En particulier,* $\mathrm{ext}\big(B_{X^*}\big) \neq \varnothing$.

Démonstration. Par le Théorème de Banach-Alaoglu, B_{X^*} est un (convexe) compact de l'espace localement convexe (X^*, w^*). □

Remarque. Ce résultat montre en particulier qu'il n'existe pas d'espace vectoriel normé X tel que X^* soit isométrique à $L^1([0,1])$ ou à $c_0(\mathbb{N})$.

COROLLAIRE 5.8. *Soit X un espace de Banach* réflexif. *Si $C \subseteq X$ est un ensemble convexe, fermé et borné, alors $C = \overline{\mathrm{conv}}\big(\mathrm{ext}(C)\big)$.*

Démonstration. Comme X est réflexif, C est un convexe compact de l'espace localement convexe (X, w). Donc $C = \overline{\mathrm{conv}\big(\mathrm{ext(C)}\big)}^{w}$ par Krein-Milman ; et donc $C = \overline{\mathrm{conv}}\big(\mathrm{ext}(C)\big)$ car l'adhérence faible d'un convexe est égale à son adhérence en norme. □

COROLLAIRE 5.9. *Soit X un evt localement convexe, et soit $K \subseteq X$ un ensemble convexe compact. Notons $A(K)$ l'ensemble des fonctions affines continues $f : K \to \mathbb{C}$. Pour tout point $x \in K$, on peut trouver une mesure de probabilité borélienne μ sur $\overline{\mathrm{ext}(K)}$ telle que*

$$\forall f \in A(K) \ : \ f(x) = \int_{\overline{\mathrm{ext}(K)}} f(z)\, d\mu(z).$$

Démonstration. On posera $E := \overline{\mathrm{ext}(K)}$. Comme $\mathrm{ext}(K) \subseteq K$ et comme K est compact, E est un compact de X.

D'après le Théorème de Krein-Milman, on peut trouver une suite généralisée $(z_p)_{p\in P} \subseteq \mathrm{conv}\big(\mathrm{ext}(K)\big)$ telle que $z_p \to x$.

Pour tout $p \in P$, écrivons

$$z_p = \sum_{i=1}^{N_p} c_{i,p} e_{i,p},$$

où $e_{i,p} \in \mathrm{ext}(K)$ et $c_{i,p} \geqslant 0$ avec $\sum_i c_{i,p} = 1$. Alors

$$\mu_p := \sum_{i=1}^{N_p} c_{i,p} \delta_{e_{i,p}}$$

est une mesure de probabilité borélienne sur E. En particulier $\mu_p \in B_{M(E)}$.

Comme E est compact, on a $M(E) = \mathcal{C}(E)^*$ d'après le Théorème de représentation de Riesz. Donc (μ_p) est une suite généralisée dans $B_{\mathcal{C}(E)^*}$, et donc (μ_p) possède une sous-s.g. qui converge w^* dans $\mathcal{C}(E)^* = M(E)$. On peut donc supposer qu'il existe une mesure $\mu \in M(E)$ telle que $\mu_p \xrightarrow{w^*} \mu$, *i.e.*

$$\forall f \in \mathcal{C}(E) \ : \ \int_E f\, d\mu_p \to \int_E f\, d\mu.$$

La mesure μ est positive car les μ_p le sont, et $\mu(E) = \int_E \mathbf{1}\, d\mu = \lim \int_E \mathbf{1}\, d\mu_p = 1$ car $\mu_p(E) = 1$ pour tout p ; donc μ est une mesure de probabilité.

Si $f : K \to \mathbb{C}$ est affine continue, alors

$$\forall p \in P \ : \ \int_E f\, d\mu_p = \sum_{i=1}^{N_p} c_{i,p} f(e_{i,p}) = f\left(\sum_{i=1}^{N_p} c_{i,p} e_{i,p}\right) = f(z_p).$$

Comme $f(z_p) \to f(x)$ par continuité de f, on obtient donc le résultat souhaité :

$$f(x) = \int_E f\, d\mu \qquad \text{pour toute } f \in A(K).$$

$\square$

5.4. Stone-Weierstrass. Dans cette section, on utilise le Théorème de Krein-Milman pour donner une preuve du *Théorème de Stone-Weierstrass*, dont on rappelle l'énoncé.

THÉORÈME 5.10. *Soit K un espace topologique compact. Si $\mathcal{A} \subseteq \mathcal{C}(K)$ est une sous-algèbre stable par conjugaison, contenant $\mathbf{1}$ et qui sépare les points de K, alors $\mathcal{A}$ est dense dans $\mathcal{C}(K)$.*

Démonstration. Il suffit de montrer que l'ensemble

$$\mathcal{A}^\perp := \left\{\mu \in M(K);\ \int_K f\, d\mu = 0 \text{ pour toute } f \in \mathcal{A}\right\} \subseteq M(K) = \mathcal{C}(K)^*$$

est réduit à $\{0\}$. Dans la suite, on posera

$$C := \{\mu \in \mathcal{A}^\perp;\ \|\mu\| \leqslant 1\} = B_{\mathcal{A}^\perp}.$$

Comme $\mathcal{A}^\perp$ est w^*- fermé dans $M(K)$, l'ensemble C est w^*- compact. On va raisonner par l'absurde : on suppose que $C \neq \{0\}$ et on cherche à obtenir une contradiction.

FAIT 1. Si $\mu \in \mathcal{A}^\perp$, alors $h\mu \in \mathcal{A}^\perp$ pour toute $h \in \mathcal{A}$.

Preuve du Fait 1. Si $f \in \mathcal{A}$, alors $fh \in \mathcal{A}$ et donc $\int_K f\, d(h\mu) = \int_K fh\, d\mu = 0$. $\square$

FAIT 2. Si $a, b \in K$ et $a \neq b$, on peut trouver une fonction $g \in \mathcal{A}$ à valeurs réelles telle que $0 \leqslant g \leqslant \mathbf{1}$ et $g(a) \neq g(b)$.

Preuve du Fait 2. Comme $\mathcal{A}$ est un espace vectoriel stable par conjugaison qui sépare les points de K, on peut trouver $f \in \mathcal{A}$ à valeurs réelles telle que $f(a) \neq f(b)$ et $\|f\|_\infty = 1$. Alors $g := \frac{\mathbf{1}+f}{2}$ convient ; on a bien $g \in \mathcal{A}$ car $\mathbf{1} \in \mathcal{A}$. $\square$

FAIT 3. Soit $\mu \in \mathrm{ext}(C)$. Si $\mu_1, \mu_2 \in \mathcal{A}^\perp$ sont telles que $\mu_1 + \mu_2 = \mu$ et $\|\mu_1\| + \|\mu_2\| = \|\mu\|$, alors μ_1 et μ_2 sont proportionnelles à μ.

Preuve du Fait 3. Il n'y a rien à démontrer si $\mu_1 = 0$ ou $\mu_2 = 0$; donc on suppose que $\mu_1 \neq 0$ et $\mu_2 \neq 0$. Alors $\nu_1 := \frac{\mu_1}{\|\mu_1\|}$ et $\nu_2 := \frac{\mu_2}{\|\mu_2\|}$ sont bien définies et appartiennent à C. De plus, comme $\mu \in \mathrm{ext}(C)$ et $C \neq \{0\}$, on a $\|\mu\| = 1$, donc $\|\mu_1\| + \|\mu_2\| = 1$. Si on pose $\alpha := \|\mu_1\|$, on voit ainsi que $0 < \alpha < 1$ et $\mu = \alpha\, \nu_1 + (1-\alpha)\, \nu_2$. Donc $\nu_1 = \mu = \nu_2$ puisque $\mu \in \mathrm{ext}(C)$, ce qui prouve le Fait. $\square$

FAIT 4. Si $\mu \in \mathrm{ext}(C)$, alors $\mathrm{supp}(\mu)$ est réduit à 1 point.

Preuve du Fait 4. Comme on suppose que $C \neq \{0\}$, on a nécessairement $\|\mu\| = 1$ et donc $\mu \neq 0$. Supposons que $\text{supp}(\mu)$ contienne au moins 2 points distincts a et b. Par le Fait 2, on peut trouver $g \in \mathcal{A}$ à valeurs réelles telle que $0 \leqslant g \leqslant \mathbf{1}$ et $g(a) \neq g(b)$. Alors $\mu_1 := g\mu$ et $\mu_2 := (\mathbf{1} - g)\mu$ sont dans $\mathcal{A}^\perp$ par le Fait 1, on a $\mu_1 + \mu_2 = \mu$ par définition, et $\|\mu_1\| + \|\mu_2\| = \int_K g\, d|\mu| + \int_K (\mathbf{1} - g)\, d|\mu| = \int_K d|\mu| = \|\mu\|$. Par le Fait 3, μ_1 et μ_2 sont proportionnelles à μ. On a donc une constante c telle que $\mu_1 = c\mu$. Alors $|c||\mu| = |\mu_1| = g|\mu|$ car $g \geqslant 0$; donc la fonction g est $|\mu|$-pp égale à la constante $|c|$. Comme g est continue, elle doit donc être constante sur $\text{supp}(\mu)$ (exo) ; mais ceci n'est pas vrai puisque $g(a) \neq g(b)$. □

On peut maintenant conclure la preuve du théorème. Par le Fait 4 et le Théorème de Krein-Milman, on peut trouver une mesure $\mu \in \mathcal{A}^\perp$ telle que $\text{supp}(\mu)$ est un singleton $\{a\}$. Alors $\mu = c\,\delta_a$ pour une certaine constante $c \neq 0$; mais comme $\mathbf{1} \in \mathcal{A}$, on a $c = \int_K \mathbf{1}\, d\mu = 0$ puisque $\mu \in \mathcal{A}^\perp$. On a donc obtenu la contradiction souhaitée. □

5.5. Fonctions de type positif. Dans cette cette section, on donne une application assez élaborée du Théorème de Krein-Milman, qu'on aura l'occasion d'utiliser au Chapitre 6.

Définition 5.11. *Un* ***semi-groupe commutatif*** *est un ensemble S muni d'un loi interne $(s,t) \mapsto s+t$ qui est associative, commutative, et possède un élément neutre noté* 0.

Exemples.

(1) Tout groupe commutatif est un semi-groupe commutatif.

(2) $\mathbb{R}_+$ et $\mathbb{N}$ sont des semi-groupes commutatifs. Plus généralement, $\mathbb{R}_+^d$ et $\mathbb{N}^d$ sont des semi-groupes commutatifs pour tout $d \geqslant 1$.

Définition 5.12. *Soit $(S, +)$ un semi-groupe commutatif. Une* ***involution*** *sur S est une application $s \mapsto s^*$ de S dans S qui vérifie $(s+t)^* = s^* + t^*$ et $(s^*)^* = s$ pour tous $s, t \in S$.*

Remarque. On a nécessairement $0^* = 0$, car $0^* = 0^* + 0 = (0 + 0^*)^* = (0^*)^*$.

Exemples.

(1) L'application $s \mapsto s$ est une involution !

(2) Si S est un groupe commutatif, alors l'application $s \mapsto -s$ est une involution

(3) Si $S = \mathbb{N} \times \mathbb{N}$, l'application $(k,l) \mapsto (l,k)$ est une involution.

Définition 5.13. *Soit $S = (S, +, *)$ un semi-groupe commutatif avec involution. Un* ***caractère*** *de S est une application $\gamma : S \to \mathbb{C}$ qui possède les propriétés suivantes :*

(i) *$\gamma(s+t) = \gamma(s)\gamma(t)$ pour tous $s, t \in S$, et $\gamma(0) = 1$;*

(ii) *$\gamma(s^*) = \overline{\gamma(s)}$ pour tout $s \in S$.*

Exemples.

(1) Si S est un groupe commutatif muni de l'involution $s \mapsto -s$, alors les caractères de S sont exactement les homomorphismes de S dans le groupe multiplicatif $\mathbb{T}$.

(2) Soit $S := \mathbb{N} \times \mathbb{N}$ muni de l'involution $(k,l)^* = (l,k)$. Les caractères de S sont les applications de la forme $\gamma_z(k,l) := z^k \bar{z}^l$, où $z \in \mathbb{C}$.

Démonstration. (1) Si $\gamma : S \to \mathbb{C}$ est un caractère alors, par (i), on a $\gamma(s) \neq 0$ pour tout $s \in S$ et γ est un homomorphisme de S dans le groupe multiplicatif $\mathbb{C}^*$. De plus, par (ii) on a $\overline{\gamma(s)} = \gamma(-s) = 1/\gamma(s)$, autrement dit $|\gamma(s)| = 1$ pour tout $s \in S$. Inversement, tout homomorphisme $\gamma : S \to \mathbb{T}$ est un caractère de S, par le même argument.
(2) Soit $\gamma : \mathbb{N} \times \mathbb{N} \to \mathbb{C}$ un caractère. Comme $(k, l) = (k, 0) + (0, l)$, on a $\gamma\big((k, l)\big) = \gamma\big((k, 0)\big)\gamma\big((0, l)\big) = \gamma\big(1, 0)\big)^k \gamma\big((0, 1)\big)^l$ pour tout $(k, l) \in \mathbb{N} \times \mathbb{N}$ par (i). Mais $\gamma\big((0, 1)\big) = \overline{\gamma\big((1, 0)\big)}$ par (ii), donc γ est de la forme γ_z avec $z := \gamma\big((1, 0)\big)$. La réciproque est laissée en exo. □

Définition 5.14. *Soit $S = (S, +, *)$ un semi-groupe commutatif avec involution, et soit $\phi : S \to \mathbb{C}$. On dit que ϕ est une* ***fonction de type positif*** *si, pour tous $s_1, \dots, s_d \in S$ et pour tous $c_1, \dots, c_d \in \mathbb{C}$, on a*

$$\sum_{i,j=1}^{d} c_i \overline{c_j}\, \phi(s_i + s_j^*) \geqslant 0.$$

(Si $u \in \mathbb{C}$, on écrit évidemment "$u \geqslant 0$" pour signifier que u est réel et positif !)
Autrement dit, ϕ est de type positif si et seulement si, pour tous $s_1, \dots, s_d \in S$, la matrice $\Big(\phi(s_i + s_j^*)\Big)_{1 \leqslant i,j \leqslant d} \in M_d(\mathbb{C})$ est positive.

Exemple. Tout caractère de S est une fonction de type positif.

Démonstration. Soit $\gamma : S \to \mathbb{C}$ un caractère. Si $s_1, \dots, s_d \in S$ et $c_1, \dots, c_d \in \mathbb{C}$, alors

$$\sum_{i,j=1}^{d} c_i \overline{c_j}\, \gamma(s_i + s_j^*) = \sum_{i,j=1}^{d} c_i \overline{c_j}\, \gamma(s_i)\overline{\gamma(s_j)} = \left|\sum_{n=1}^{d} c_n \gamma(s_n)\right|^2 \geqslant 0.$$

□

Remarque 5.15. Si $\phi : S \to \mathbb{C}$ est de type positif, alors $\phi(0) \geqslant 0$, $\phi(s^*) = \overline{\phi(s)}$ pour tout $s \in S$, et $|\phi(s)|^2 \leqslant \phi(0)\phi(s + s^*)$. Si de plus ϕ est bornée, alors $|\phi(s)| \leqslant \phi(0)$ pour tout $s \in S$.

Démonstration. Pour tout $s \in S$, la matrice

$$M_s := \begin{pmatrix} \phi(0) & \phi(s) \\ \phi(s^*) & \phi(s + s^*) \end{pmatrix}$$

est positive. On a donc $\phi(s^*) = \overline{\phi(s)}$, $\phi(0) \geqslant 0$ et $\det(M_s) = \phi(0)\phi(s+s^*) - \phi(s)\phi(s^*) \geqslant 0$; donc

$$|\phi(s)|^2 \leqslant \phi(0)\phi(s + s^*) \qquad \text{pour tout } s \in S.$$

Comme de plus ϕ est bornée, on peut poser $C := \|\phi\|_\infty$. L'inégalité précédente donne alors $|\phi(s)|^2 \leqslant C\phi(0)$ pour tout $s \in S$, donc $C^2 \leqslant C\phi(0)$, et donc $\|\phi\|_\infty = C \leqslant \phi(0)$. □

Notations. Si $S = (S, +, *)$ est un semi-groupe commutatif avec involution, on note $\mathcal{P}_1(S)$ l'ensemble des fonctions de type positif $\phi : S \to \mathbb{C}$ bornées et vérifiant $\phi(0) = 1$, et $\widehat{S}$ l'ensemble des caractères bornés $\gamma : S \to \mathbb{C}$. Enfin, on note $\mathcal{F}(S)$ l'ensemble de toutes les fonctions $f : S \to \mathbb{C}$, *i.e.* $\mathcal{F}(S) = \mathbb{C}^S$, et on munit $\mathcal{F}(S)$ de la topologie de la convergence simple.

LEMME 5.16. *Si $S = (S, +, *)$ est un semi-groupe commutatif avec involution, alors $\mathcal{P}_1(S) \subseteq \mathcal{F}(S)$ est convexe compact, $\widehat{S}$ est compact, et $\operatorname{ext}(\mathcal{P}_1(S)) = \widehat{S}$.*

Démonstration. Il est assez clair sur la définition que $\mathcal{P}_1(S)$ est une partie convexe et fermée de $\mathcal{F}(S) = \mathbb{C}^S$. De plus, on a $\mathcal{P}_1(S) \subseteq \overline{\mathbb{D}}^S$ d'après la Remarque 5.15 ; donc $\mathcal{P}_1(S)$ est compact d'après le Théorème de Tychonoff. Comme $\widehat{S} \subseteq \mathcal{P}_1(S)$ et que $\widehat{S}$ est clairement fermé dans $\mathcal{F}(S)$, on en déduit que $\widehat{S}$ est compact également.
Montrons que $\operatorname{ext}(\mathcal{P}_1(S)) \subseteq \widehat{S}$. Soit ϕ un point extrémal de $\mathcal{P}_1(S)$. Comme $\phi \in \mathcal{P}_1(S)$, on sait déjà que $\phi(s^*) = \overline{\phi(s)}$ pour tout $s \in S$ et $\phi(0) = 1$. Donc il reste seulement à prouver qu'on a $\phi(t+s) = \phi(t)\phi(s)$ pour tous $s, t \in S$. Dans ce qui suit, on fixe $t \in S$.

FAIT 1. Si ψ_1 et ψ_2 sont des fonctions bornées de type positif sur S telles que $\psi_1 + \psi_2 = \phi$, alors ψ_1 et ψ_2 sont proportionnelles à ϕ.

Preuve du Fait 1. On a $\psi_i(0) \geqslant 0$ et $\psi_1(0) + \psi_2(0) = \phi(0) = 1$; donc $\alpha := \psi_1(0) \in [0,1]$ et $\psi_2(0) = 1 - \alpha$. Si $\alpha = 0$ ou $\alpha = 1$, alors $\psi_1 = 0$ ou $\psi_2 = 0$ par la Remarque 5.15, et il n'y a rien à démontrer ; donc on suppose que $0 < \alpha < 1$. Si on pose $\phi_1 := \frac{1}{\alpha}\,\psi_1$ et $\phi_2 := \frac{1}{1-\alpha}\,\psi_2$, alors ϕ_1 et ϕ_2 appartiennent à $\mathcal{P}_1(S)$ puisque $\phi_i(0) = 1$, et $\alpha\phi_1 + (1-\alpha)\phi_2 = \phi$. Donc $\phi_1 = \phi = \phi_2$ car $\phi \in \operatorname{ext}(\mathcal{P}_1(S))$; ce qui prouve le fait. $\square$

FAIT 2. Pour tout $w \in \mathbb{C}$ vérifiant $|w| \leqslant 1$, la fonction $\phi_w : S \to \mathbb{C}$ définie par
$$\phi_w(s) := 2\,\phi(s) + w\,\phi(s+t) + \overline{w}\,\phi(s+t^*)$$
est de type positif

Preuve du Fait 2. Si $s_1, \dots, s_d \in S$ et $c_1, \dots, c_d \in \mathbb{C}$, alors
$$\sum_{i,j=1}^{d} c_i\overline{c_j}\phi_w(s_i + s_j^*) = 2\psi(0) + 2\mathrm{Re}\big(w\,\psi(t)\big),$$
où ψ est la fonction définie par
$$\psi(s) := \sum_{i,j=1}^{d} c_i\overline{c_j}\,\phi(s_i + s_j^* + s).$$
On vérifie sans difficulté majeure que la fonction ψ est de type positif (exo). De plus, ψ est bornée car ϕ est bornée. Donc $|\psi(s)| \leqslant \psi(0)$ pour tout $s \in S$ d'après la Remarque 5.15. Comme $|w| \leqslant 1$, on en déduit que $\psi(0) + \mathrm{Re}\big(w\,\psi(t)\big) \geqslant 0$, ce qui termine la preuve du fait. $\square$

Avec les notations du Fait 5.5, on a $\phi_1 + \phi_{-1} = 4\phi = \phi_i + \phi_{-i}$. Donc, les 4 fonctions $\phi_1, \phi_{-1}, \phi_i, \phi_{-i}$ sont proportionneles à ϕ d'après le Fait 5.5. En regardant ϕ_1 et ϕ_i, on en déduit que les fonctions $s \mapsto \phi(s+t) + \phi(s+t^*)$ et $s \mapsto \phi(s+t) - \phi(s+t^*)$ sont proportionnelles à ϕ. Par conséquent, la fonction $s \mapsto \phi(s+t)$ est proportionnelle à ϕ ; autrement dit, il existe une constante $k(t)$ telle que
$$\forall s \in S \;:\; \phi(s+t) = k(t)\phi(s).$$
En prenant $s := 0$, on voit que $k(t) = \phi(t)$; ce qui termine la preuve du fait que ϕ est un caractère. Ainsi, on a montré que $\operatorname{ext}(\mathcal{P}_1(S)) \subseteq \widehat{S}$.
Montrons maintenant que tout caractère $\gamma \in \widehat{S}$ est un point extrémal de $\mathcal{P}_1(S)$. Soient $\phi, \psi \in \mathcal{P}_1(S)$ telles que $\gamma = \frac{\phi+\psi}{2}\cdot$. D'après la Remarque 5.15, on a $|\phi(s)|^2 \leqslant \phi(s+s^*)$ et

$|\psi(s)|^2 \leqslant \psi(s+s^*)$ pour tout $s \in S$, tandis que $|\gamma(s)|^2 = \gamma(s)\overline{\gamma(s)} = \gamma(s+s^*)$ puisque γ est un caractère. Donc

$$\left|\frac{\phi(s)+\psi(s)}{2}\right|^2 = \gamma(s+s^*) = \frac{1}{2}\big(\phi(s+s^*)+\psi(s+s^*)\big) \geqslant \frac{1}{2}\left(|\phi(s)|^2+|\psi(s)|^2\right);$$

ce qui force $\phi(s) = \psi(s)$ car la fonction $z \mapsto |z|^2$ est strictement convexe sur $\mathbb{C}$. Ceci étant vrai pour tout $s \in S$, on a donc $\phi = \gamma = \psi$. □

Théorème 5.17. (Herglotz)
Soit $S = (S, +, *)$ un semi-groupe commutatif dénombrable avec involution. Pour une fonction $\phi : S \to \mathbb{C}$, les propriétés suivantes sont équivalentes.

(1) *ϕ est bornée et de type positif, avec $\phi(0) = 1$.*
(2) *Il existe une mesure de probabilité borélienne m sur $\widehat{S}$ telle que*

$$\forall s \in S \;:\; \phi(s) = \int_{\widehat{S}} \gamma(s)\, dm(\gamma).$$

Démonstration. Comme tout $\gamma \in \widehat{S}$ est une fonction de type positif, l'implication (2) $\Longrightarrow$ (1) est un exo facile.
Inversement, supposons que ϕ soit bornée et de type positif, avec $\phi(0) = 1$. Alors $\phi \in \mathcal{P}_1(S)$. D'après le Lemme 5.16 et le Théorème de Krein-Milman (Corollaire 5.9), et comme de plus $\mathrm{ext}\big(\mathcal{P}_1(S)\big) = \widehat{S}$ est compact, on peut donc trouver une mesure de probabilité borélienne m sur $\widehat{S}$ telle que

$$F(\phi) = \int_{\widehat{S}} F(\gamma)\, dm(\gamma) \qquad \text{pour toute fonction affine continue } F : \mathcal{P}_1(S) \to \mathbb{C}.$$

Comme la fonction $\psi \mapsto \psi(s)$ est affine et continue sur $\mathcal{P}_1(S)$ pour tout $s \in S$ (puisque $\mathcal{P}_1(S)$ est muni de la topologie de la convergence simple), on obtient donc le résultat souhaité :

$$\forall s \in S \;:\; \phi(s) = \int_{\widehat{S}} \gamma(s)\, dm(\gamma).$$

□

Exemple. Si $\phi = \mathbb{Z} \to \mathbb{C}$ est une fonction bornée de type positif telle que $\phi(0) = 1$, alors il existe une mesure de probabilité borélienne μ sur $\mathbb{T}$ telle que

$$\forall n \in \mathbb{Z} \;:\; \phi(n) = \widehat{\mu}(n) := \int_{\mathbb{T}} z^n d\mu(z).$$

Démonstration. On prend pour S le groupe $\mathbb{Z}$ avec l'involution $n \mapsto -n$. Les caractères de $\mathbb{Z}$ sont les homomorphismes $\gamma : \mathbb{Z} \to \mathbb{T}$. Un tel homomorphisme est de la forme $\gamma(n) = z^n$ pour un certain $z \in \mathbb{T}$, à savoir $z := \gamma(1)$. L'application $\gamma \mapsto \gamma(1)$ est une bijection de $\widehat{\mathbb{Z}}$ sur $\mathbb{T}$, et elle est continue car la topologie de $\widehat{\mathbb{Z}} \subseteq \mathcal{F}(\mathbb{Z})$ est la topologie de la convergence simple. Donc cette application est un homéomorphisme de $\widehat{\mathbb{Z}}$ sur $\mathbb{T}$ car $\widehat{\mathbb{Z}}$ est compact. (Si on préfère, il est très facile de voir directement que la réciproque est continue.) On peut ainsi identifier complètement $\widehat{\mathbb{Z}}$ et $\mathbb{T}$, de sorte que le résultat découle immédiatement du théorème. □

Chapitre 6

Rudiments de théorie spectrale

1. Algèbres de Banach

1.1. Définition et exemples. Dans cette section, on prend $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

DÉFINITION 1.1. *Une* $\mathbb{K}$-***algèbre*** *est un* $\mathbb{K}$*-espace vectoriel* A *muni d'une "multiplication interne", i.e. une application* $(a, b) \mapsto a \cdot b$ *de* $A \times A$ *dans* A *qui est* associative $(a \cdot (b \cdot c) = (a \cdot b) \cdot c)$ *et* distributive par rapport à l'addition $(a \cdot (b + c) = a \cdot b + a \cdot c$ *et* $(b + c) \cdot a = b \cdot a + c \cdot a)$. *On dit que* A *est* ***commutative*** *si la multiplication est commutative* $(a \cdot b = b \cdot a$ *pour tous* $a, b \in A)$. *On dit que* A *est* ***unitaire*** *si* $A \neq \{0\}$ *et s'il existe un élément* $e \in A$ *– nécessairement unique – tel que* $\forall a \in A \;:\; a \cdot e = a = e \cdot a$.

EXEMPLES. L'ensemble des polynômes à coefficients dans $\mathbb{K}$ est une $\mathbb{K}$-algèbre commutative et unitaire, qu'on notera $\mathbb{K}[\mathbf{z}]$ ou $\mathbb{K}[\mathbf{x}]$. Si $d \in \mathbb{N}^*$, alors $M_d(\mathbb{K})$ est une $\mathbb{K}$-algèbre unitaire, non commutative dès que $d \geqslant 2$.

Remarque 1. On écrira ab au lieu de $a \cdot b$.

Remarque 2. On notera en général $\mathbf{1}$ l'élément unité pour la multiplication (si A est unitaire).

Remarque 3. Si A est une $\mathbb{K}$-algèbre, alors $(A, +, \cdot)$ est un anneau. Donc on peut se demander si A est *intègre* $(a \cdot b = 0 \implies a = 0$ ou $b = 0)$, ou si A est un *corps* quand A est commutative et unitaire.

DÉFINITION 1.2. *Une* ***algèbre normée*** *sur* $\mathbb{K}$ *est une* $\mathbb{K}$*-algèbre* $A \neq \{0\}$ *munie d'une norme telle que*

$$\forall a, b \in A \;:\; \|ab\| \leqslant \|a\|\,\|b\|.$$

Si A *est unitaire, on impose de plus que* $\|\mathbf{1}\| = 1$. *Une* ***algèbre de Banach*** *est une algèbre normée complète.*

REMARQUE. Si A est une algèbre normée, alors la multiplication est continue de $A \times A$ dans A.

Démonstration. Exo. □

EXEMPLES.

(1) $A := \mathbb{K}$ est une algèbre de Banach !
(2) Si K est un espace topologique compact, alors $A := \mathcal{C}(K)$ est une algèbre de Banach, commutative et unitaire.
(3) Si $(\Omega, \mathfrak{B}, m)$ est un espace mesuré, alors $A := L^\infty(\Omega, m)$ est une algèbre de Banach, commutative et unitaire.

(4) Si X est un espace de Banach , alors $A := \mathcal{L}(X)$ est une algèbre de Banach, unitaire ($\mathbf{1} = I_X$) mais non commutative si $\dim(X) \geqslant 2$.

(5) $A := L^1(\mathbb{R})$ muni du *produit de convolution* $*$ est une algèbre de Banach commutative, non unitaire.

Définition 1.3. *Soient A et B deux $\mathbb{K}$-algèbres. Un* ***homomorphisme*** *de A dans B est une application $\Phi : A \to B$ qui est linéaire et* multiplicative *: $\Phi(uv) = \Phi(u)\Phi(v)$ pour tous $u, v \in A$. Si A et B sont unitaires, on demande de plus que $\Phi(\mathbf{1}) = \mathbf{1}$.*

Exemple. Soit A une $\mathbb{K}$-algèbre unitaire. Pour tout polynôme

$$P = c_0\mathbf{1} + c_1\mathbf{z} + \cdots + c_n\mathbf{z}^n \in \mathbb{K}[\mathbf{z}],$$

posons

$$P(a) := c_0\mathbf{1} + c_1 a + \cdots c_n a^n.$$

Alors l'application $P \mapsto P(a)$ est un homomorphisme de $\mathbb{K}[\mathbf{z}]$ dans A ; plus précisément, c'est l'unique homomorphisme $\Phi : \mathbb{K}[\mathbf{z}] \to A$ tel que $\Phi(\mathbf{z}) = a$.

1.2. Éléments inversibles. On prend toujours $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Définition 1.4. *Soit A une $\mathbb{K}$-algèbre unitaire. On dit qu'un élément a de A est* ***inversible*** *(dans A) s'il existe $b \in A$ – nécessairement unique – tel que $ab = \mathbf{1} = ba$; et on écrit alors $b = a^{-1}$. On note* $\mathrm{G}(A)$ *l'ensemble des éléments inversibles de A.*

Remarque. $\mathrm{G}(A)$ est un *groupe* pour la multiplication ; et comme dans tout groupe : si $a, b \in \mathrm{G}(A)$, alors $(ab)^{-1} = b^{-1}a^{-1}$.

Exercice. Montrer qu'un élément a de A est inversible si et seulement si il est inversible à gauche et à droite, *i.e.* il existe $b_1, b_2 \in A$ tels que $b_1 a = \mathbf{1} = ab_2$.

Exemple 1. Soit $A := \mathcal{L}(X)$, où X est un espace de Banach. D'après le Théorème d'isomorphisme de Banach, un opérateur $T \in \mathcal{L}(X)$ est inversible si et seulement si T est bijectif, *i.e.* injectif et surjectif. Si $\dim(X) < \infty$, alors T inversible $\iff T$ injectif $\iff T$ surjectif.

Exemple 2. On prend $A := \mathcal{L}(\ell^2(\mathbb{N}))$. Soit $a \in \ell^\infty(\mathbb{N})$, et soit $\Delta_a : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ l'opérateur diagonal associé. Si $a_n \neq 0$ pour tout $n \in \mathbb{N}$, alors Δ_a est injectif ; mais si $\inf_{n\in\mathbb{N}} |a_n| = 0$, par exemple si $a_n \to 0$, alors Δ_a n'est pas un plongement, et n'est donc pas inversible.

Démonstration. Que Δ_a soit injectif si $a_n \neq 0$ pour tout n est un micro-exo. En notant (e_n) la base canonique de $\ell^2(\mathbb{N})$, on a $\Delta_a e_n = a_n e_n$ pour tout $n \in \mathbb{N}$; donc il est à peu près clair que si $\inf_n |a_n| = 0$, alors Δ_a n'est pas un plongement. □

Exemple 3. On prend $A := \mathcal{C}(K)$, où K est un espace topologique compact. Une fonction $f \in \mathcal{C}(K)$ est inversible dans $\mathcal{C}(K)$ si et seulement si $f(t) \neq 0$ pour tout $t \in K$.

Démonstration. Si f est inversible, il existe une fonction $g : K \to \mathbb{C}$ telle que $f(t)g(t) \equiv 1$ sur K, et donc certainement $f(t) \neq 0$ pour tout $t \in K$. Inversement, si f ne s'annule pas, alors $g := 1/f$ est bien définie et *continue* sur K, donc f est inversible *dans* $\mathcal{C}(K)$. □

Exemple 4. On prend $A := L^\infty = L^\infty(\Omega, m)$, où $(\Omega, \mathfrak{B}, m)$ est un espace mesuré. Une $f \in L^\infty$ est inversible dans L^∞ si et seulement si il existe une constante $\varepsilon > 0$ telle que $|f(t)| \geqslant \varepsilon$ pp.

Démonstration. Supposons f inversible dans L^∞, et soit $g := f^{-1}$. On a $f(t)g(t) = 1$ pp, donc $g(t) \neq 0$ pp et $|f(t)| = 1/|g(t)| \geqslant \varepsilon := 1/\|g\|_\infty$ pp. Inversement, supposons qu'on ait $|f(t)| \geqslant \varepsilon > 0$ pp. Alors $g(t) := 1/f(t)$ est bien défini pp et $|g(t)| \leqslant 1/\varepsilon$ pp, donc $g \in L^\infty$; et $fg = \mathbf{1} = gf$ par définition. $\square$

Lemme 1.5. *Soit A une algèbre de Banach unitaire. Si $u \in A$ vérifie $\|u\| < 1$, alors $\mathbf{1} - u$ est inversible et*

$$(\mathbf{1} - u)^{-1} = \sum_{k=0}^{\infty} u^k.$$

En particulier, on a $\left\|(\mathbf{1} - u)^{-1}\right\| \leqslant \frac{1}{1-\|u\|}$ *et* $\|(\mathbf{1} - u)^{-1} - \mathbf{1}\| \leqslant \frac{\|u\|}{1-\|u\|}$.

Démonstration. Comme $\|u\| < 1$, on a $\sum_{k=0}^\infty \|u^k\| \leqslant \sum_{k=0}^\infty \|u\|^k < \infty$. Donc la série $\sum u^k$ converge dans A car A est un espace de Banach. Si on pose $b := \sum_{k=0}^\infty u^k$, on vérifie que $b(\mathbf{1} - u) = \mathbf{1} = (\mathbf{1} - u)b$ en utilisant la continuité du produit ; donc $\mathbf{1} - u$ est inversible et $(\mathbf{1} - u)^{-1} = b$.
On a $\|(\mathbf{1} - u)^{-1}\| = \left\|\sum_{k=0}^\infty u^k\right\| \leqslant \sum_{k=0}^\infty \|u^k\| \leqslant \sum_{k=0}^\infty \|u\|^k = \frac{1}{1-\|u\|}$. De même, comme $(\mathbf{1} - u)^{-1} - \mathbf{1} = \sum_{k=1}^\infty u^k$, on a $\|(\mathbf{1} - u)^{-1} - \mathbf{1}\| \leqslant \sum_{k=1}^\infty \|u\|^k = \frac{\|u\|}{1-\|u\|}$. $\square$

Exercice. Soit A une algèbre normée unitaire (pas forcément complète), et soit $u \in A$. On suppose que $\mathbf{1} - u$ est inversible. Montrer que si $\|u\| < 1$, alors $\|(\mathbf{1} - u)^{-1}\| \leqslant \frac{1}{1-\|u\|}$, et que si $\|u\| > 1$, alors $\|(\mathbf{1} - u)^{-1}\| \leqslant \frac{1}{\|u\|-1}$.

Corollaire 1.6. *Soit A une algèbre de Banach unitaire. Si $a \in \mathrm{G}(A)$, alors $B\big(a, 1/\|a^{-1}\|\big) \subseteq \mathrm{G}(A)$. De plus, on a la majoration suivante :*

$$\forall x \in \overline{B}\big(a, 1/2\|a^{-1}\|\big) \ : \ \|x^{-1} - a^{-1}\| \leqslant 2\|a^{-1}\|^2 \, \|x - a\|.$$

Démonstration. Si $x \in B\big(a, 1/\|a^{-1}\|\big)$, alors $x = a + (x - a) = a(\mathbf{1} - u)$ où $u := a^{-1}(a - x)$ vérifie $\|u\| \leqslant \|a^{-1}\| \, \|a - x\| < 1$; donc $x \in \mathrm{G}(A)$ et $x^{-1} = (\mathbf{1} - u)^{-1}a^{-1}$. Si de plus $x \in \overline{B}\big(a, 1/2\|a^{-1}\|\big)$, alors $\|u\| \leqslant 1/2$; donc

$$\begin{aligned}\|x^{-1} - a^{-1}\| &= \left\|\big((\mathbf{1} - u)^{-1} - \mathbf{1}\big)a^{-1}\right\| \\ &\leqslant \frac{\|u\|}{1 - \|u\|} \, \|a^{-1}\| \\ &\leqslant 2\|u\| \, \|a^{-1}\| \\ &\leqslant 2\|a^{-1}\|^2 \, \|x - a\|.\end{aligned}$$

$\square$

Remarque. Si $a \in A$ est inversible, alors $\|a^{-1}\| \geqslant \frac{1}{\|a\|}$, mais en en général on n'a pas égalité. Donc $1/\|a^{-1}\| \leqslant \|a\|$, mais on ne peut en général pas remplacer $B\big(a, 1/\|a^{-1}\|\big)$ par $B(a, \|a\|)$ dans (1.6).

Démonstration. $\|a\| \, \|a^{-1}\| \geqslant \|aa^{-1}\| = \|\mathbf{1}\| = 1$. $\square$

COROLLAIRE 1.7. *Si A est une algèbre de Banach unitaire, alors* $\mathrm{G}(A)$ *est un ouvert de A, et l'application $x \mapsto x^{-1}$ est continue sur* $\mathrm{G}(A)$.

Démonstration. Évident par le corollaire précédent. □

COROLLAIRE 1.8. *Si A est une algèbre de Banach unitaire, alors* $\mathrm{G}(A)$ *est un groupe topologique.*

LEMME 1.9. *Soit A une algèbre de Banach unitaire, et soit $(a_n)_{n\in\mathbb{N}}$ une suite d'éléments de* $\mathrm{G}(A)$. *Si $a_n \to a \in A$ et si $a \notin \mathrm{G}(A)$, alors $\|a_n^{-1}\| \to \infty$.*

Démonstration. Supposons que $\|a_n^{-1}\| \not\to \infty$. Alors, quitte à extraire une sous-suite, on peut supposer qu'il existe une constante $C < \infty$ telle que $\forall n \in \mathbb{N} \ : \ \|a_n^{-1}\| \leqslant C$. Comme $a_n \to a$, on peut trouver n tel que $\|a - a_n\| < 1/C$. Alors $\|a - a_n\| < 1/\|a_n^{-1}\|$, donc $a \in \mathrm{G}(A)$ par le Corollaire 1.6 ; ce qui est faux par hypothèse. □

COROLLAIRE 1.10. *Si $a \in A$ est tel que $a \in \partial\mathrm{G}(A)$, alors a est un* ***diviseur de 0 topologique*** *: on peut trouver une suite $(b_n)_{n\in\mathbb{N}} \subseteq A$ telle que $\|b_n\| = 1$ pour tout $n \in \mathbb{N}$ et cependant $b_n a \to 0$ et $ab_n \to 0$.*

Démonstration. Par hypothèse, on a $a \notin \mathrm{G}(A)$ mais on peut trouver une suite $(a_n) \subseteq \mathrm{G}(A)$ telle que $a_n \to a$. Alors $\|a_n^{-1}\| \to \infty$ par le lemme ; donc, si on pose $b_n := \frac{a_n^{-1}}{\|a_n^{-1}\|}$, on a $\|b_n\| = 1$ et $\|b_n a\| \leqslant \|b_n a_n\| + \|b_n(a - a_n)\| = \frac{1}{\|a_n^{-1}\|} + \|b_n(a - a_n)\| \leqslant \frac{1}{\|a_n^{-1}\|} + \|a - a_n\| \xrightarrow{n\to\infty} 0$; et de même $ab_n \to 0$. □

REMARQUE 1.11. Soient A et B deux $\mathbb{K}$-algèbres unitaires. Si $\Phi : A \to B$ est un homomorphisme d'algèbres, alors $\Phi\big(\mathrm{G}(A)\big) \subseteq \mathrm{G}(B)$ et $\Phi(x)^{-1} = \Phi(x^{-1})$ pour tout $x \in \mathrm{G}(A)$.

Démonstration. Exo. □

1.3. Spectre, rayon spectral. On prend toujours $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

DÉFINITION 1.12. *Soit A une $\mathbb{K}$-algèbre unitaire, et soit $a \in A$. Le* ***spectre de*** *a dans l'algèbre A est l'ensemble de tous les scalaires λ tels que $a - \lambda\mathbf{1}$ n'est* pas *inversible dans A. Cet ensemble se note $\sigma_A(a)$, ou simplement $\sigma(a)$:*

$$\sigma(a) = \big\{\lambda \in \mathbb{K};\ a - \lambda\mathbf{1} \notin \mathrm{G}(A)\big\}.$$

REMARQUE 0. Dire que $0 \in \sigma(a)$ signifie que a n'est pas inversible dans A.

REMARQUE 1. On a $\sigma(\mathbf{1}) = \{1\}$.

Démonstration. Exo. □

EXERCICE. Soient A et B deux $\mathbb{K}$-algèbres unitaires. Si $\Phi = A \to B$ est un homomorphisme, alors $\forall a \in A \ : \ \sigma(\Phi(a)) \subseteq \sigma(a)$; et si Φ est un isomorphisme, alors $\forall a \in A \ : \ \sigma(\Phi(a)) = \sigma(a)$.

EXEMPLE 1. Soit $A := M_d(\mathbb{K}) = \mathcal{L}(\mathbb{K}^d)$. Si $A \in M_d(\mathbb{K})$, alors $\sigma(A)$ est l'ensemble des valeurs propres de A.

Exemple 2. On prend $A := \mathcal{L}(\ell^2(\mathbb{N}))$. Soit $a \in \ell^\infty(\mathbb{N})$, et soit $\Delta_a : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ l'opérateur diagonal associé. Si $a_n \neq 0$ pour tout $n \in \mathbb{N}$ et $\inf_{n\in\mathbb{N}} |a_n| = 0$, alors $0 \in \sigma(\Delta_a)$ mais 0 n'est pas valeur propre de Δ_a.

Démonstration. On a vu que sous ces hypothèses, Δ_a est injectif mais n'est pas un plongement. □

Exercice. Montrer que si $a \in \ell^\infty(\mathbb{N})$, alors $\sigma(\Delta_a) = \overline{\{a_n;\ n \in \mathbb{N}\}}$.

Exemple 3. On prend $A := \mathcal{C}(K)$, où K est un espace topologique compact. Si $f \in \mathcal{C}(K)$, alors $\sigma(f) = f(K)$.

Démonstration. Par définition, $\lambda \notin \sigma(f) \iff f - \lambda\mathbf{1}$ inversible dans $\mathcal{C}(K) \iff f(t) - \lambda \neq 0$ pour tout $t \in K \iff \lambda \notin f(K)$. □

Exemple 4. On prend $A := L^\infty = L^\infty(\Omega, m)$.

(1) Si $\phi \in L^\infty$ et $\lambda \in \mathbb{C}$, alors
$$\lambda \in \sigma_{L^\infty}(\phi) \iff \forall \varepsilon > 0\ :\ m\big(\{t \in \Omega;\ |\phi(t) - \lambda| < \varepsilon\}\big) > 0.$$

(2) Si $\phi \in L^\infty$, alors $\sigma_{L^\infty}(\phi)$ est le plus petit fermé $K \subseteq \mathbb{C}$ tel que $\phi(t) \in K$ pp. Pour cette raison, on dit que $\sigma_{L^\infty}(\phi)$ est l'**image essentielle** de ϕ.

Démonstration. (1) Par définition $\lambda \notin \sigma_{L^\infty}(\phi) \iff \phi - \lambda\mathbf{1}$ inversible dans $L^\infty \iff \exists \varepsilon > 0\ :\ |\phi(t) - \lambda| \geqslant \varepsilon$ pp.
(2) Par (1), un nombre complexe λ n'appartient pas à $\sigma_{L^\infty}(\phi)$ si et seulement si il existe un voisinage ouvert V de λ tel que $m\big(\phi^{-1}(V)\big) = 0$. Autrement dit, $\Omega := \mathbb{C}\backslash\sigma_{L^\infty}(\phi)$ est la réunion de tous les ouverts $V \subseteq \mathbb{C}$ tels que $m\big(\phi^{-1}(V)\big) = 0$. On en déduit d'une part que Ω est un ouvert de $\mathbb{C}$; et d'autre part, en utilisant la propriété de Lindelöf, qu'on a $m\big(\phi^{-1}(\Omega)\big) = 0$. Ainsi, $\mathbb{C}\backslash\sigma_{L^\infty}(\phi)$ est *le plus grand* ouvert $V \subseteq \mathbb{C}$ tel que $m\big(\phi^{-1}(V)\big) = 0$; ce qui donne (2). □

Proposition 1.13. *Soit A une $\mathbb{K}$-algèbre unitaire, et soit $P \in \mathbb{K}[\mathbf{z}]$. Pour tout $\lambda \in \sigma(a)$, on a $P(\lambda) \in \sigma\big(P(a)\big)$; autrement dit, $\sigma(P(a)) \supseteq P(\sigma(a)) := \{P(\lambda);\ \lambda \in \sigma(a)\}$. Si de plus $\mathbb{K} = \mathbb{C}$ alors $\sigma\big(P(a)\big) = P(\sigma(a))$.*

Démonstration. Soit $\lambda \in \sigma(a)$. Le polynôme $P - P(\lambda)$ s'annule en λ, donc on peut écrire $P(\mathbf{z}) - P(\lambda) = (\mathbf{z} - \lambda)Q(\mathbf{z}) = Q(\mathbf{z})(\mathbf{z} - \lambda)$ pour un certain polynôme $Q \in \mathbb{K}[\mathbf{z}]$. Donc $P(a) - P(\lambda)\mathbf{1} = (a - \lambda\mathbf{1})Q(a) = Q(a)(a - \lambda\mathbf{1})$. Comme $a - \lambda\mathbf{1}$ n'est pas inversible, on en déduit que $P(a) - P(\lambda)\mathbf{1}$ n'est pas inversible : s'il l'était, alors $a - \lambda\mathbf{1}$ serait inversible à gauche et à droite (exo), donc inversible. Ainsi, on a $P(\lambda) \in \sigma\big(P(a)\big)$ pour tout $\lambda \in \sigma(a)$.
Inversement, supposons que $\mathbb{K} = \mathbb{C}$, et soit $\mu \in \sigma\big(P(a)\big)$: on veut montrer qu'il existe $\lambda \in \sigma(a)$ tel que $\mu = P(\lambda)$. C'est évident si P est constant ; donc on suppose que $\deg(P) \geqslant 1$. Comme on est sur $\mathbb{C}$, on peut écrire $P(\mathbf{z}) - \mu = c\,(\mathbf{z} - \lambda_1)\cdots(\mathbf{z} - \lambda_d)$, où $c \neq 0$. Alors $P(a) - \mu\mathbf{1} = c(a - \lambda_1\mathbf{1})\cdots(a - \lambda_d\,\mathbf{1})$. Comme $P(a) - \mu\mathbf{1}$ n'est pas inversible, on en déduit qu'il existe $k \in [\![1, d]\!]$ tel que $a - \lambda_k\mathbf{1}$ n'est pas inversible. On a ainsi $\lambda_k \in \sigma(a)$, et $P(\lambda_k) = \mu$ puisque λ_k est racine de $P(\mathbf{z}) - \mu$. □

Lemme 1.14. *Soit A une algèbre de Banach unitaire. Si $a \in A$, alors $\sigma(a)$ est un* compact *de $\mathbb{K}$, et $\sigma(a) \subseteq \overline{D}(0, \|a\|)$.*

Démonstration. Par définition, $\sigma(a)$ est un fermé de $\mathbb{K}$ car $\lambda \in \sigma(a) \iff a - \lambda\mathbf{1} \in A\backslash \mathrm{G}(A)$ et l'application $\lambda \mapsto a - \lambda\mathbf{1}$ est continue de $\mathbb{K}\backslash\sigma(a)$ dans A. Si $\lambda \in \mathbb{K}$ vérifie $|\lambda| > \|a\|$, alors $a - \lambda\mathbf{1} = -\lambda(\mathbf{1} - u)$ où $u := \frac{1}{\lambda}a$ vérifie $\|u\| < 1$; donc $a - \lambda\mathbf{1}$ est inversible, *i.e.* $\lambda \notin \sigma(a)$. On a donc $\sigma(a) \subseteq \overline{D}(0, \|a\|)$; et en particulier $\sigma(a)$ est compact car fermé et borné dans $\mathbb{K}$. $\square$

Définition 1.15. *Soit A une algèbre de Banach unitaire. Si $a \in A$, on pose $r(a) := \sup\big\{|\lambda|;\ \lambda \in \sigma(a)\big\}$, avec la convention $\sup\varnothing = -\infty$. On dit que $r(a)$ est le* ***rayon spectral*** *de a.*

Remarque 1. Comme $\sigma(\mathbf{1}) = \{1\}$, on a $r(\mathbf{1}) = 1$.

Remarque 2. On a toujours $r(a) \leqslant \|a\|$.

Remarque 3. Pour tout $n \in \mathbb{N}$, on a $r(a^n) \geqslant r(a)^n$; et $r(a^n) = r(a)^n$ si $\mathbb{K} = \mathbb{C}$.

Démonstration. Par la proposition 1.13, on a $\lambda^n \in \sigma(a^n)$ pour tout $\lambda \in \sigma(a)$; donc $r(a)^n = \sup\big\{|\lambda^n|;\ \lambda \in \sigma(a)\big\} \leqslant r(a^n)$. Si $\mathbb{K} = \mathbb{C}$, on a $\sigma(a^n) = \{\lambda^n;\ \lambda \in \sigma(a)\}$ toujours par la Proposition 1.13, donc $r(a^n) = r(a)^n$. $\square$

Exercice. On prend $A := L^\infty = L^\infty(\Omega, m)$. Montrer que si $\phi \in L^\infty$, alors $r(\phi) = \|\phi\|_\infty$.

Désormais, on prend $\mathbb{K} = \mathbb{C}$.

Notation. Soit A une algèbre de Banach unitaire, et soit $a \in A$. Pour tout $\lambda \in \mathbb{C}\backslash\sigma(a)$, on pose $R_a(\lambda) := (a - \lambda\mathbf{1})^{-1}$. On dit que l'application $R_a : \mathbb{C}\backslash\sigma(a) \to A$ est la **résolvante** de l'élément a.

Fait 1.16. *Soit $a \in A$. Pour tous $\lambda, \mu \in \mathbb{C}\backslash\sigma(a)$, on a $R_a(\lambda)R_a(\mu) = R_a(\mu)R_a(\lambda)$ et*

$$R_a(\mu) - R_a(\lambda) = (\mu - \lambda)R_a(\mu)R_a(\lambda) \qquad (\textit{\textbf{équation résolvante}}).$$

Démonstration. Comme $(a - \mu\mathbf{1})(a - \lambda\mathbf{1}) = (a - \lambda\mathbf{1})(a - \mu\mathbf{1})$, on a $R_a(\lambda)R_a(\mu) = R_a(\mu)R_a(\lambda)$. De plus,

$$\begin{aligned}
R_a(\mu) - R_a(\lambda) &= (a - \mu\mathbf{1})^{-1} - (a - \lambda\mathbf{1})^{-1} \\
&= (a - \mu\mathbf{1})^{-1}\Big(\mathbf{1} - (a - \mu\mathbf{1})(a - \lambda\mathbf{1})^{-1}\Big) \\
&= (a - \mu\mathbf{1})^{-1}\Big((a - \lambda\mathbf{1}) - (a - \mu\mathbf{1})\Big)(a - \lambda\mathbf{1})^{-1} \\
&= R_a(\mu)\,(\mu - \lambda)\mathbf{1}\,R_a(\lambda).
\end{aligned}$$

$\square$

Lemme 1.17. *Soit A une $\mathbb{C}$-algèbre de Banach unitaire, et soit $a \in A$.*

(1) *La fonction $R_a : \mathbb{C}\backslash\sigma(a) \to A$ est holomorphe, avec $R_a'(\lambda) = R_a(\lambda)^2$. De plus, $R_a(\lambda) \to 0$ quand $|\lambda| \to \infty$.*

(2) *Pour tout* λ *vérifiant* $|\lambda| > r(a)$*, on a*

$$R_a(\lambda) = -\sum_{k=0}^{\infty} \frac{a^k}{\lambda^{k+1}},$$

où la série converge dans A*.*

Démonstration. (1) Soit $\lambda \in \mathbb{C}\backslash\sigma(a)$ quelconque. Si $\mu \in \mathbb{C}\backslash\sigma(a)$ et $\mu \neq \lambda$, alors

$$\frac{R_a(\mu) - R_a(\lambda)}{\mu - \lambda} = R_a(\mu)R_a(\lambda) = (a - \mu\mathbf{1})^{-1}(a - \lambda\mathbf{1})^{-1}.$$

Par continuité de l'application $x \mapsto x^{-1}$ sur $\mathrm{G}(A)$ et de la multiplication de A, on en déduit que

$$\frac{R_a(\mu) - R_a(\lambda)}{\mu - \lambda} \xrightarrow{\mu\to\lambda} (a - \lambda\mathbf{1})^{-1}(a - \lambda\mathbf{1})^{-1} = R_a(\lambda)^2.$$

Donc R_a est $\mathbb{C}$-dérivable en tout point $\lambda \in \mathbb{C}\backslash\sigma(a)$, avec $R_a'(\lambda) = R_a(\lambda)^2$.
Si $|\lambda| \geqslant 2\|a\|$, alors $a - \lambda\mathbf{1} = -\lambda(\mathbf{1} - u)$, où $u := \frac{a}{\lambda}$ vérifie $\|u\| \leqslant 1/2$; donc $R_a(\lambda) = \frac{1}{\lambda}\,(\mathbf{1} - u)^{-1}$ et donc $\|R_a(\lambda)\| = \frac{1}{|\lambda|}\,\|(\mathbf{1} - u)^{-1}\| \leqslant \frac{2}{|\lambda|} \xrightarrow{|\lambda|\to\infty} 0$.
(2) Si λ vérifie $|\lambda| > \|a\|$, alors $\left\|\frac{a}{\lambda}\right\| < 1$, donc on peut écrire

$$R_a(\lambda) = -\frac{1}{\lambda}\left(\mathbf{1} - \frac{a}{\lambda}\right)^{-1} = -\frac{1}{\lambda}\sum_{k=0}^{\infty}\left(\frac{a}{\lambda}\right)^k = -\sum_{k=0}^{\infty}\frac{a^k}{\lambda^{k+1}}.$$

Il faut voir que cette identité est encore valable si λ vérifie seulement $|\lambda| > r(a)$. En particulier, il faut montrer que la série converge, ce qui n'est pas évident.
Soit $F : D\big(0, 1/r(a)\big) \to A$ la fonction définie par $F(0) := 0$ et $F(z) := R_a(1/z)$ si $0 < |z| < 1/r(a)$. Par (1), la fonction F est bien définie, holomorphe sur $D\big(0, 1/r(a)\big)\backslash\{0\}$ avec $F'(z) = -\frac{1}{z^2}R_a(1/z)^2$, et continue en 0. De plus,

$$\frac{F(z) - F(0)}{z - 0} = \frac{1}{z}\left(a - \frac{1}{z}\mathbf{1}\right)^{-1} = (za - \mathbf{1})^{-1} \xrightarrow{z\to 0} -\mathbf{1},$$

par continuité de l'application $x \mapsto x^{-1}$; donc F est $\mathbb{C}$-dérivable en 0 avec $F'(0) = -\mathbf{1}$. Ainsi, la fonction F est holomorphe sur tout le disque $D\big(0, 1/r(a)\big)$.
Exactement comme dans le cas des fonctions holomorphes à valeurs scalaires, on en déduit (en utilisant la formule de Cauchy) que F se développe en série entière dans le disque $D\big(0, 1/r(a)\big)$. Il existe donc une suite de "coefficients" $(c_n)_{n\geqslant 0} \subseteq A$ uniquement déterminée telle que

$$\forall z \in D\big(0, 1/r(a)\big)\ :\ \sum_{n=0}^{\infty} c_n z^n,$$

où la série converge dans A. Comme de plus $F(0) = 0$, on a $c_0 = 0$; donc la somme démarre à $n = 1$, et ainsi

$$F(z) = \sum_{k=0}^{\infty} c_{k+1} z^{k+1}.$$

Par ailleurs, si z vérifie $|z| < 1/\|a\|$, alors $|1/z| > \|a\|$, donc

$$F(z) = R_a(1/z) = -\sum_{k=0}^{\infty}\frac{a^k}{(1/z)^{k+1}} = -\sum_{k=0}^{\infty} a^k z^{k+1}.$$

Par unicité des coefficients du développement en série entière de F dans le disque $D(0,1/\|a\|)$, on en déduit qu'on a $c_{k+1}=-a_k$ pour tout $k\geqslant 0$.
Si maintenant $\lambda\in\mathbb{C}$ vérifie $|\lambda|>r(a)$, alors $z:=1/\lambda\in D\big(0,1/r(a)\big)$; donc on peut écrire

$$R_a(\lambda)=F(1/\lambda)=\sum_{k=0}^{\infty}\frac{1}{\lambda^{k+1}}\,c_{k+1}=-\sum_{k=0}^{\infty}\frac{a^k}{\lambda^{k+1}},$$

où la série converge dans A. □

Théorème 1.18. *Soit A une algèbre de Banach unitaire sur $\mathbb{C}$. Si $a\in A$, alors $\sigma(a)\neq\varnothing$. De plus, on a la* ***formule du rayon spectral*** *:*

$$r(a)=\lim_{n\to\infty}\|a^n\|^{1/n}=\inf_{n\geqslant 1}\|a^n\|^{1/n}.$$

Démonstration. (i) Supposons que $\sigma(a)=\varnothing$. Alors la résolvante R_a est holomorphe sur $\mathbb{C}$. Comme $R_a(\lambda)\to 0$ quand $|\lambda|\to\infty$, la fonction R_a est bornée sur $\mathbb{C}$. Donc R_a est *constante* d'après le Théorème de Liouville, et donc $R_a(\lambda)\equiv 0$ puisque $R_a(\lambda)\to 0$ quand $|\lambda|\to\infty$; mais ceci est absurde puisque $R_a(\lambda)$ est toujours un élément inversible de A.
(ii) Pour montrer la formule du rayon spectral, on a besoin du fait suivant.

Fait. Si $(u_n)_{n\geqslant 1}$ est une suite de nombres réels $\geqslant 0$ **sous-multiplicative**, *i.e.* vérifiant

$$\forall n,m\in\mathbb{N}^*\ :\ u_{n+m}\leqslant u_nu_m,$$

alors la suite $\big(u_n^{1/n}\big)$ est convergente et $\lim_{n\to\infty}u_n^{1/n}=\inf_{n\geqslant 1}u_n^{1/n}$.

Preuve du Fait. Exo bien connu. □

Si on pose $u_n:=\|a^n\|$, alors la suite (u_n) est sous-multiplicative : $\|a^{n+m}\|=\|a^na^m\|\leqslant\|a^n\|\,\|a^m\|$. Donc $l:=\lim_{n\to\infty}\|a^n\|^{1/n}$ existe, et $l=\inf_{n\geqslant 1}\|a^n\|^{1/n}$. Il s'agit de voir que $l=r(a)$.
On a vu que pour tout $n\in\mathbb{N}$, on a $r(a)^n\leqslant r(a^n)\leqslant\|a^n\|$; donc $\forall n\geqslant 1\ :\ r(a)\leqslant\|a^n\|^{1/n}$, et donc $r(a)\leqslant l$.
Pour l'inégalité inverse, il suffit de montrer que pour tout $r>r(a)$, on a $r\geqslant l$. Comme $r>r(a)$, la série $\sum\frac{a^k}{r^{k+1}}$ converge dans A, donc la série $\sum\frac{a^k}{r^k}$ converge également. En particulier, la suite $\big(\frac{a^k}{r^k}\big)$ est bornée dans A. Si on pose $M:=\sup_{k\geqslant 0}\big\|\frac{a^k}{r^k}\big\|$, alors $\|a^k\|^{1/k}\leqslant M^{1/k}\,r$ pour tout $k\geqslant 1$; donc $l=\lim\|a^k\|^{1/k}\leqslant r$ car $M^{1/k}\to 1$.

□

1.4. Spectre dans une sous-algèbre.

Proposition 1.19. *Soit A une $\mathbb{C}$-algèbre de Banach unitaire, et soit B une sous-algèbre fermée de A contenant $\mathbf{1}$. Si $a\in B$, alors $\sigma_A(a)\subseteq\sigma_B(a)$ et $\partial\sigma_B(a)\subseteq\sigma_A(a)$.*

Démonstration. L'inclusion $\sigma_A(a)\subseteq\sigma_B(a)$ est évidente : si $a-\lambda\mathbf{1}$ n'est pas inversible dans A, il n'est *a fortiori* pas inversible dans B.
Soit $\lambda\in\partial\sigma_B(a)$. Alors $a-\lambda\mathbf{1}\notin\mathrm{G}(B)$, mais on peut trouver une suite $(\lambda_n)\subseteq\mathbb{C}$ tel que $a-\lambda_n\mathbf{1}\in\mathrm{G}(B)$ pour tout $n\in\mathbb{N}$ et $\lambda_n\to\lambda$. Comme $a-\lambda_n\mathbf{1}\to a-\lambda\mathbf{1}$, on en déduit que $a-\lambda\mathbf{1}\in\partial G(B)$. Donc $a-\lambda\mathbf{1}$ est un diviseur de 0 topologique dans B d'après le Lemme 1.9 : on peut trouver une suite $(b_n)\subseteq B$ telle que $\|b_n\|=1$ pour tout $n\in\mathbb{N}$ et $(a-\lambda\mathbf{1})b_n\to 0$. Si $a-\lambda\mathbf{1}$ était inversible dans A, on pourrait écrire

$b_n = (a-\lambda\mathbf{1})^{-1}(a-\lambda\mathbf{1})b_n$, donc $1 = \|b_n\| \leqslant \|(a-\lambda\mathbf{1})^{-1}\|\,\|(a-\lambda\mathbf{1})b_n\| \xrightarrow{n\to\infty} 0$, ce qui est absurde. Donc $\lambda \in \sigma_A(a)$. □

Corollaire 1.20. *Soit A une $\mathbb{C}$-algèbre de Banach unitaire, et soit B une sous-algèbre fermée de A contenant $\mathbf{1}$. Si $a \in B$, alors $\sigma_B(a)$ est la réunion de $\sigma_A(a)$ et de certaines composantes connexes bornées de $\mathbb{C}\backslash\sigma_A(a)$. En particulier, si $\mathbb{C}\backslash\sigma_A(a)$ est connexe, alors $\sigma_B(a) = \sigma_A(a)$.*

Démonstration. Notons $(O_i)_{i\in I}$ la famille des composantes connexes de $\mathbb{C}\backslash\sigma_A(a)$. Comme $\mathbb{C}\backslash\sigma_A(a)$ est un ouvert de $\mathbb{C}$, les O_i sont des *ouverts* connexes de $\mathbb{C}$, et $\mathbb{C}\backslash\sigma(a) = \bigcup_{i\in I} O_i$. Soit $I_0 := \{i \in I;\ O_i \cap \sigma_B(a) \neq \varnothing\}$. Alors $\sigma_B(a) = \sigma_A(a) \cup \bigcup_{i\in I_0} \big(O_i \cap \sigma_B(a)\big)$. Il suffit de montrer qu'on a $O_i \subseteq \sigma_B(a)$ pour tout $i \in I_0$: en effet, on en déduira que $\sigma_B(a) = \sigma_A(a) \cup \bigcup_{i\in I_0} O_i$, donc que $\sigma_B(a)$ la réunion de $\sigma_A(a)$ et de certaines composantes connexes de $\mathbb{C}\backslash\sigma_A(a)$, nécessairement bornées puisque $\sigma_B(a)$ est borné. Soit $i \in I_0$. Alors $E := \sigma_B(a) \cap O_i$ est un fermé non vide de O_i. De plus, on a en fait $E \subseteq \overset{\circ}{\widehat{\sigma_B(a)}}$, car $E\backslash\overset{\circ}{\widehat{\sigma_B(a)}} = \Big(\sigma_B(a)\backslash\overset{\circ}{\widehat{\sigma_B(a)}}\Big) \cap O_i = \partial\sigma_B(a) \cap O_i \subseteq \sigma_A(a) \cap O_i$ par la proposition et $\sigma_A(a) \cap O_i = \varnothing$. Donc $E = \overset{\circ}{\widehat{\sigma_B(a)}} \cap O_i$, et donc E est un *ouvert* de O_i. Comme O_i est connexe et comme E est aussi fermé dans O_i, on en déduit que $E = O_i$, ce qu'on voulait montrer. □

Exemple. Soit $A := \mathcal{C}(\mathbb{T})$, et soit B l'adhérence des fonctions polynomiales dans $\mathcal{C}(\mathbb{T})$. En notant $\mathbf{z}$ la fonction $z \mapsto z$, on a $\sigma_A(\mathbf{z}) = \mathbb{T}$ et $\sigma_B(\mathbf{z}) = \overline{\mathbb{D}}$.

Démonstration. On a $\mathbf{z}(\mathbb{T}) = \mathbb{T}$, donc $\sigma_{\mathcal{C}(\mathbb{T})}(\mathbf{z}) = \mathbb{T}$.
Les composantes connexes de $\mathbb{C}\backslash\mathbb{T}$ sont $\mathbb{D}$ et $\mathbb{C}\backslash\overline{\mathbb{D}}$; donc on a $\sigma_B(\mathbf{z}) = \mathbb{T}$ ou $\sigma_B(\mathbf{z}) = \mathbb{T} \cup \mathbb{D} = \overline{\mathbb{D}}$ d'après le Corollaire 1.20. Il suffit donc de montrer que $\sigma_B(\mathbf{z})$ contient $\overline{\mathbb{D}}$. Soit $\lambda \in \overline{\mathbb{D}}$, et supposons que $\lambda \notin \sigma_B(\mathbf{z})$. Alors on peut trouver une fonction $g \in B$ telle que $(\mathbf{z} - \lambda\mathbf{1})g = \mathbf{1}$, *i.e.* $(z-\lambda)g(z) = 1$ pour tout $z \in \mathbb{T}$. Par définition de B, on peut trouver une suite de polynômes $(P_n)_{n\in\mathbb{N}}$ telle que $P_n \to g$ dans $\mathcal{C}(\mathbb{T})$, *i.e.* $P_n(z) \to g(z)$ uniformément sur $\mathbb{T}$. Alors $Q_n(z) := 1 - (z-\lambda)P_n(z) \to 0$ uniformément sur $\mathbb{T} = \partial\mathbb{D}$. Mais les Q_n sont des fonctions polynomiales, donc des fonctions holomorphes sur $\mathbb{C}$. D'après le *Principe du maximum*, on en déduit que $Q_n(z) \to 0$ uniformément sur $\overline{\mathbb{D}}$. En particulier $Q_n(\lambda) \to 0$, ce qui est absurde puisque $Q_n(\lambda) \equiv 1$. □

1.5. Cas des opérateurs bornés. Dans cette section, X est un espace de Banach, et $A = \mathcal{L}(X)$. Donc si $T \in \mathcal{L}(X)$ et $\lambda \in \mathbb{K}$, alors

$$\lambda \in \sigma(T) \iff T - \lambda I \text{ n'est pas bijectif.}$$

1.5.1. *Adjoints.*

Lemme 1.21. *Soit $T \in \mathcal{L}(X)$.*

(a) *En notant $T^* : X^* \to X^*$ l'adjoint banachique de T, on a $\sigma(T^*) = \sigma(T)$.*

(b) *Si X est un espace* de Hilbert *H et si on note $T^* \in \mathcal{L}(H)$ l'adjoint* hilbertien *de T, alors $\sigma(T^*) = \big\{\bar{\lambda};\ \lambda \in \sigma(T)\big\}$.*

Démonstration. Si $\lambda \in \mathbb{C}$, alors

$$\lambda \notin \sigma(T) \iff T - \lambda I \text{ inversible} \iff (T-\lambda I)^* \text{ inversible;}$$

d'où le résultat car $(T-\lambda I)^* = T^* - \lambda I$ dans le cas banachique et $(T-\lambda I)^* = T^* - \bar{\lambda} I$ dans le cas hilbertien. □

Remarque. On n'a rien de semblable pour les valeurs propres. Par exemple, si B est le "backward shift" sur $\ell^2(\mathbb{N})$ et si S est le "forward shift", alors $B = S^*$, et 0 est valeur propre de B mais pas de S.

1.5.2. *Particularités hilbertiennes.* Dans ce qui suit, H est un espace de Hilbert complexe.

Proposition 1.22. *Soit $T \in \mathcal{L}(H)$.*

(1) *Si T est auto-adjoint, alors $\sigma(T) \subseteq \mathbb{R}$.*
(2) *Si T est positif, alors $\sigma(T) \subseteq \mathbb{R}^+$.*
(3) *Si T est unitaire, alors $\sigma(T) \subseteq \mathbb{T}$.*

Démonstration. (1) Supposons que $T^* = T$. On veut montrer que $T - \lambda I$ est inversible pour tout $\lambda \in \mathbb{C}\backslash\mathbb{R}$; et pour cela, il suffit de vérifier que $T - \lambda I$ et $(T - \lambda I)^*$ sont des plongements. Écrivons $\lambda = a + ib$, avec $a, b \in \mathbb{R}$ et donc $b \neq 0$. On a $T - \lambda I = A + iB$, où $A := T - aI$ et $B := bI$ sont auto-adjoints avec $AB = BA$. Donc $\forall x \in H\ :\ \|(T - \lambda I)x\|^2 = \|Ax\|^2 + \|Bx\|^2 \geqslant \|Bx\|^2 = b^2\|x\|^2$, et donc $T - \lambda I$ est un plongement. De même, $(T - \lambda I)^* = T - \bar{\lambda} I$ est un plongement car $\bar{\lambda} \notin \mathbb{R}$.
(2) Comme T est auto-adjoint, on sait déjà que $\sigma(T) \subseteq \mathbb{R}$; donc il suffit de montrer que $\sigma(T) \cap]-\infty, 0[= \varnothing$, *i.e.* que $T - \lambda I$ est inversible pour tout $\lambda < 0$. Si on écrit $\lambda = -c$ où $c > 0$, on a $\forall x \in H\ :\ \langle (T - \lambda I)x, x\rangle = \langle Tx, x\rangle + c\,\|x\|^2 \geqslant c\,\|x\|^2$; donc, par Cauchy-Schwarz : $\|(T - \lambda I)x\| \geqslant c\,\|x\|$ pour tout $x \in H$. Donc $T - \lambda I$ est un plongement, et donc $T - \lambda I$ est inversible puisque $T - \lambda I$ est auto-adjoint.
(3) Il faut montrer que $T - \lambda I$ est inversible pour tout $\lambda \in \mathbb{C}\backslash\mathbb{T}$. Si $|\lambda| > 1$, alors $\|T\| = 1 < \frac{1}{1/|\lambda|} = \frac{1}{\|(\lambda I)^{-1}\|}$; donc $T - \lambda I$ est inversible par le Lemme 1.5. De même, si $|\lambda| < 1$, alors $\|\lambda I\| = |\lambda| < 1 = \frac{1}{\|T^{-1}\|}$, donc $T - \lambda I$ est inversible. □

Proposition 1.23. *Si $T \in \mathcal{L}(H)$ est un opérateur* normal*, alors $r(T) = \|T\|$.*

Démonstration. Par la formule du rayon spectral, on a $r(T) = \lim_{n\to\infty} \|T^{2^n}\|^{1/2^n}$; donc il suffit de montrer que $\|T^{2^n}\| = \|T\|^{2^n}$ pour tout $n \geqslant 0$. C'est évident si $n = 0$; donc, par une récurrence immédiate, il suffit de montrer qu'on a $\|T^2\| = \|T\|^2$ pour tout opérateur normal T.
L'idée est de regarder $\|T\|^4$. On a $\|T\|^4 = \big(\|T\|^2\big)^2 = \|T^*T\|^2 = \|(T^*T)^*(T^*T)\| = \|T^*TT^*T\|$. Mais T est normal, donc $T^*TT^*T = (T^*)^2T^2 = (T^2)^*T^2$. On obtient donc $\|T\|^4 = \|(T^2)^*T^2\| = \|T^2\|^2$, *i.e.* $\|T\|^2 = \|T^2\|$. □

Corollaire 1.24. *Pour tout opérateur $T \in \mathcal{L}(H)$, on a $\|T\| = \sqrt{r(T^*T)}$.*

Démonstration. Comme T^*T est auto-adjoint, on a $r(T^*T) = \|T^*T\| = \|T\|^2$. □

Théorème 1.25. *Soit $T \in \mathcal{L}(H)$* auto-adjoint*. Posons $m := \inf\big\{\langle Tx, x\rangle;\ \|x\| = 1\big\}$ et $M := \sup\big\{\langle Tx, x\rangle;\ \|x\| = 1\big\}$. Alors m et M appartiennent à $\sigma(T)$ et $\sigma(T) \subseteq [m, M]$.*

Démonstration. Comme $\sigma(T)$ est un compact de $\mathbb{R}$, il possède un plus petit élément et un plus grand élément. Il s'agit de vérifier que $\min\big(\sigma(T)\big) = m$ et $\max\big(\sigma(T)\big) = M$. On va se contenter de $\min\big(\sigma(T)\big)$.

Par définition de m, l'opérateur $T - mI$ est positif. Donc $\sigma(T - mI) \subseteq \mathbb{R}^+$; et comme $\sigma(T - mI) = \sigma(T) - m$, on en déduit que $\min\bigl(\sigma(t)\bigr) \geqslant m$. Pour l'inégalité inverse, il suffit de vérifier que $m \in \sigma(T)$. Par définition de m, on peut trouver une suite $(x_n) \subseteq H$ telle que $\|x_n\| = 1$ pour tout $n \in \mathbb{N}$ et $\langle Tx_n, x_n\rangle \to m$, autrement dit $\langle (T - mI)x_n, x_n\rangle \to 0$. De plus, comme l'opérateur $T - mI$ est positif, on a aussi $\|(T - mI)x_n\|^2 \leqslant \|T - mI\|\,\langle (T - mI)x_n, x_n\rangle$ pour tout $n \in \mathbb{N}$. Donc $(T - mI)x_n \to 0$. Comme $\|x_n\| = 1$ pour tout $n \in \mathbb{N}$, on en déduit que $T - mI$ n'est pas un plongement ; en particulier, $T - mI$ n'est pas inversible, *i.e.* $m \in \sigma(T)$. □

Corollaire 1.26. *Si $T \in \mathcal{L}(H)$ est auto-adjoint, alors $\|T\| \in \sigma(T)$ ou $-\|T\| \in \sigma(T)$.*

Démonstration. Comme T est auto-adjoint, on a $\|T\| = \sup\bigl\{|\langle Tx, x\rangle|;\ \|x\| = 1\bigr\} = M$ ou $-m$. □

Corollaire 1.27. *Soit $T \in \mathcal{L}(H)$, et supposons que T soit auto-adjoint. Alors T est positif* si et seulement si *$\sigma(T) \subseteq \mathbb{R}^+$.*

Démonstration. On sait déjà que si T est positif, alors $\sigma(T) \subseteq \mathbb{R}^+$. Inversement, si $\sigma(T) \subseteq \mathbb{R}^+$, alors $m := \inf\bigl\{\langle Tx, x\rangle;\ \|x\| = 1\bigr\}$ est $\geqslant 0$ par le théorème, donc T est positif. □

Remarque 1.28. La preuve du Théorème 1.25 est valable sans aucune modification pour un opérateur auto-adjoint T sur un espace de Hilbert *réel*. On en déduit en particulier que si H est un espace de Hilbert réel et si $T \in \mathcal{L}(H)$ est auto-adjoint, alors $\sigma(T) \neq \varnothing$ et $r(T) = \|T\|$.

1.5.3. *Valeurs propres, valeurs propres approchées.* On revient au cas général d'un espace de Banach X (sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

Définition 1.29. *Soit $T \in \mathcal{L}(X)$.*

(1) *Le **spectre ponctuel** de T est l'ensemble des valeurs propres de T. On le note $\sigma_p(T)$. Donc : $\lambda \in \sigma_p(T)$ si et seulement si existe $x \in X$ tel que $\|x\| = 1$ et $Tx - \lambda x = 0$.*

(2) *On dit que $\lambda \in \mathbb{K}$ est une **valeur propre approchée** de T si $T - \lambda I$ n'est pas un plongement ; autrement dit, s'il existe une suite $(x_n)_{n\in\mathbb{N}} \subseteq X$ telle que $\|x_n\| = 1$ pour tout $n \in \mathbb{N}$ et $Tx_n - \lambda x_n \to 0$. On note $\sigma_{ap}(T)$ l'ensemble des valeurs propres approchées de T.*

Remarque. Évidemment, $\sigma_p(T) \subseteq \sigma_{ap}(T) \subseteq \sigma(T)$.

Exercice. Montrer que $\sigma_{ap}(T)$ est un compact de $\mathbb{C}$.

Exemple. Soit $a \in \ell^\infty(\mathbb{N})$ et soit $\Delta_a \in \mathcal{L}\bigl(\ell^2(\mathbb{N})\bigr)$ l'opérateur diagonal associé. Si $a_n \to 0$ et $a_n \neq 0$ pour tout $n \in \mathbb{N}$, alors $0 \in \sigma_{ap}(\Delta_a)$ mais $0 \notin \sigma_p(\Delta_a)$.

Démonstration. Exo. □

Exercice. Montrer que si $T \in \mathcal{L}(X)$ est compact, alors $\sigma_{ap}(T)\backslash\{0\} \subseteq \sigma_p(T)$. (On verra plus loin qu'en fait $\sigma(T)\backslash\{0\} \subseteq \sigma_p(T)$.)

Proposition 1.30. *Soit $T \in \mathcal{L}(X)$.*

(a) *On a $\sigma(T) = \sigma_{ap}(T) \cup \sigma_p(T^*) = \sigma_p(T) \cup \sigma_{ap}(T^*)$.*

(b) *Si X est un espace de Hilbert H et si on note T^* l'adjoint hilbertien de T, alors $\sigma(T) = \sigma_{ap}(T) \cup \big\{\bar\lambda;\ \lambda \in \sigma_p(T^*)\big\} = \sigma_p(T) \cup \big\{\bar\lambda;\ \lambda \in \sigma_{ap}(T^*)\big\}$.*

Démonstration. (a) Soit $\lambda \in \sigma(T)$. Si $\lambda \notin \sigma_{ap}(T)$, alors $T-\lambda I$ est injectif et à image fermée, mais n'est pas surjectif ; donc $T-\lambda I$ n'est pas à image dense, et donc $(T-\lambda I)^* = T^* - \lambda I$ n'est pas injectif, *i.e.* $\lambda \in \sigma_p(T^*)$. On a donc $\sigma(T) \subseteq \sigma_{ap}(T) \cup \sigma_p(T^*)$; et l'inclusion inverse est évidente car $\sigma_{ap}(T) \subseteq \sigma(T)$ et $\sigma_p(T^*) \subseteq \sigma(T^*) = \sigma(T)$. De même, si $\lambda \in \sigma(T)$ mais $\lambda \notin \sigma_p(T)$, alors $T - \lambda I$ est injectif mais pas surjectif ; donc $T^* - \lambda I = (T - \lambda I)^*$ n'est pas un plongement, *i.e.* $\lambda \in \sigma_{ap}(T^*)$. On a donc $\sigma(T) \subseteq \sigma(T) \cup \sigma_{ap}(T^*)$, et en fait égalité.
La preuve de (b) est identique. On peut aussi dire que (b) découle de (a) *via* l'identification antilinéaire entre l'adjoint banachique et l'adjoint hilbertien. □

Corollaire 1.31. *Soit H un espace de Hilbert. Si $T \in \mathcal{L}(H)$ est normal, alors $\sigma(T) = \sigma_{ap}(T)$.*

Démonstration. Il suffit de montrer que $\big\{\bar\lambda;\ \lambda \in \sigma_p(T^*)\big\} \subseteq \sigma_{ap}(T)$. Comme T est normal, $T - \lambda I$ est normal pour tout $\lambda \in \mathbb{K}$, et donc $\ker(T - \lambda I) = \ker\big((T - \lambda I)^*\big) = \ker(T^* - \bar\lambda I)$. En particulier $\ker(T - \lambda I) \neq \{0\} \iff \ker(T^* - \bar\lambda I) \neq \{0\}$; donc $\big\{\bar\lambda;\ \lambda \in \sigma_p(T^*)\big\} = \sigma_p(T)$, ce qui est plus que ce qu'on voulait. □

Proposition 1.32. *Pour tout $T \in \mathcal{L}(X)$, on a $\partial\sigma(T) \subseteq \sigma_{ap}(T)$.*

Démonstration. Soit $\lambda \in \partial\sigma(T)$. Alors $T - \lambda I$ n'est pas inversible, mais on peut trouver une suite $(\lambda_n) \subseteq \mathbb{C}$ telle que $\lambda_n \to \lambda$ et $T - \lambda_n I$ est inversible pour tout $n \in \mathbb{N}$. Par le Lemme 1.9, on sait que $\|(T - \lambda_n I)^{-1}\| \to \infty$. Donc on peut trouver une suite $(y_n) \subseteq X$ telle que $\|y_n\| = 1$ pour tout $n \in \mathbb{N}$ et $\|(T - \lambda_n I)^{-1} y_n\| \to \infty$. Si on pose $x_n := \frac{(T-\lambda_n I)^{-1} y_n}{\|(T-\lambda_n I)^{-1} y_n\|}$, alors $\|x_n\| = 1$ et $(T - \lambda_n I)x_n = \frac{y_n}{\|(T-\lambda_n I)^{-1} y_n\|} \to 0$ car $\|y_n\| = 1$. Comme $\lambda_n \to \lambda$, on en déduit que $(T - \lambda I)x_n \to 0$ (exo) ; et donc $\lambda \in \sigma_{ap}(T)$ puisque $\|x_n\| = 1$ pour tout $n \in \mathbb{N}$. □

Exemple 1.33. Soient S et B le forward shift et le backward shift sur $\ell^2(\mathbb{N})$.

(i) On a $\sigma_p(B) = \mathbb{D}$ et $\sigma(B) = \sigma_{ap}(B) = \overline{\mathbb{D}}$. De plus, si $\lambda \in \mathbb{D}$, alors $\ker(B - \lambda I)$ est de dimension 1.

(ii) On a $\sigma(S) = \overline{\mathbb{D}}$, $\sigma_{ap}(S) = \mathbb{T}$ et $\sigma_p(S) = \varnothing$.

Démonstration. (i) Si on résout formellement l'équation $Bx = \lambda x$ avec $x \in \mathbb{C}^{\mathbb{N}}$, on trouve que $x_{i+1} = \lambda x_i$ pour tout $i \in \mathbb{N}$, et donc $x_n = \lambda^n x_0$ pour tout $n \in \mathbb{N}$. Donc $x \neq 0$ si et seulement si $x_0 \neq 0$; et sous cette hypothèse, x appartient à $\ell^2(\mathbb{N})$ si et seulement si $|\lambda| < 1$. Cela prouve que $\sigma_p(B) = \mathbb{D}$; et aussi que si $\lambda \in \mathbb{D}$, alors $\ker(B - \lambda I)$ est de dimension 1, engendré par $e_\lambda := (1, \lambda, \lambda^2, \dots)$.
Comme $\|B\| = 1$, on a $r(B) \leqslant 1$, *i.e.* $\sigma(B) \subseteq \overline{\mathbb{D}}$. Comme $\sigma(B)$ est compact et contient $\mathbb{D}$ par ce qui précède, on en déduit $\sigma(B) = \overline{\mathbb{D}}$. De plus, $\mathbb{D} = \sigma_p(B) \subseteq \sigma_{ap}(B)$ et $\partial\mathbb{D} \subseteq \partial\sigma(B) \subseteq \sigma_{ap}(B)$; donc $\sigma_{ap}(B) = \overline{\mathbb{D}}$ également. (Plus simplement, on peut aussi dire que $\sigma_{ap}(B)$ est compact avec $\mathbb{D} \subseteq \sigma_{ap}(B) \subseteq \overline{\mathbb{D}}$, donc $\sigma_{ap}(B) = \overline{\mathbb{D}}$.)
(ii) Comme $S = B^*$, on a $\sigma(S) = \{\bar\lambda;\ \lambda \in \sigma(B)\} = \overline{\mathbb{D}}$. Par la Proposition 1.30 et comme $S^* = B$, on en déduit que $\overline{\mathbb{D}} = \sigma_{ap}(S) \cup \big\{\bar\lambda;\ \lambda \in \sigma_p(B)\big\} = \sigma_{ap}(S) \cup \mathbb{D}$, et donc $\sigma_{ap}(S) = \mathbb{T}$.

Montrons pour finir que $\sigma_p(S) = \varnothing$. Si $\lambda \in \mathbb{C}$, l'équation $Sx = \lambda x$ donne $\lambda x_0 = 0$ et $\lambda x_i = x_{i-1}$ pour tout $i \geqslant 1$. Donc l'équation n'a pas de solution $x \neq 0$ si $\lambda \neq 0$; et elle n'en a pas non plus si $\lambda = 0$! Ainsi, aucun $\lambda \in \mathbb{C}$ n'est valeur propre de S. □

2. Propriétés spectrales des opérateurs compacts

Dans cette section, X est un espace de Banach sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On voudrait démontrer le théorème suivant.

THÉORÈME 2.1. *Soit $K \in \mathcal{L}(X)$ un opérateur compact.*

(1) *Si* $\dim(X) = \infty$, *alors* $0 \in \sigma(K)$.

(2) *Si* $\lambda \in \mathbb{K}\backslash\{0\}$, *alors*

$$\lambda \in \sigma(K) \iff K - \lambda I \textit{ non injectif} \iff K - \lambda I \textit{ non surjectif.}$$

En particulier, $\sigma(K)\backslash\{0\}$ *est constitué uniquement de* valeurs propres *de* K.

(3) *Si* $\lambda \in \mathbb{K}\backslash\{0\}$, *alors* $\lambda \in \sigma_p(K) \iff \lambda \in \sigma_p(K^*)$, *où* K^* *est l'adjoint banachique de* K.

(4) *Si* $\lambda \in \mathbb{K}\backslash\{0\}$, *alors* $\dim\ker(K - \lambda I) < \infty$, *et* $\operatorname{Im}(K - \lambda I)$ *est* fermé *dans* X *avec* $\dim\big(X/\operatorname{Im}(K - \lambda I)\big) < \infty$. *Les mêmes propriétés sont satisfaites par* $(K - \lambda I)^n$ *pour tout* $n \in \mathbb{N}$.

(5) *Si* $\lambda \in \mathbb{K}\backslash\{0\}$, *alors il existe un entier* n *tel que*

$$X = \ker(K - \lambda I)^n \oplus \operatorname{Im}(K - \lambda I)^n.$$

De plus, la restriction de $(K - \lambda I)^n$ *à* $\operatorname{Im}(K - \lambda I)^n$ *est un isomorphisme de* $\operatorname{Im}(K - \lambda I)^n$ *sur* $\operatorname{Im}(K - \lambda I)^n$.

(6) *Si* $\lambda \in \mathbb{K}\backslash\{0\}$, *alors*

$$\dim\,\ker(K - \lambda I) = \dim\big(X/\operatorname{Im}(K - \lambda I)\big) = \dim\ker(K^* - \lambda I).$$

(7) $\sigma(K)$ *est ou bien fini, ou bien de la forme* $\{0\} \cup \{\lambda_n;\ n \in \mathbb{N}\}$, *où les* λ_n *sont deux à deux distincts et* $\lambda_n \to 0$.

Pour la preuve, on a besoin de plusieurs lemmes dont certains sont intéressantes pour eux même.

LEMME 2.2. *Soient* Y *et* Z *deux espaces de Banach. S'il existe un plongement compact* $J : Y \to Z$, *alors* $\dim(Y) < \infty$.

Démonstration. Comme J est injectif, il suffit de montrer que $\dim\big(J(Y)\big) < \infty$. D'après le Théorème de Riesz, cela revient montrer que toute suite bornée $(z_n)_{n\in\mathbb{N}} \subseteq J(Y)$ admet une sous-suite qui converge dans $J(Y)$. Si on écrit $z_n = Jy_n$, alors la suite (y_n) est bornée dans Y car J est un plongement ; donc la suite $(z_n) = (Jy_n)$ admet une sous-suite convergente car J est compact, $z_{n_k} \to z \in Z$; et $z \in J(Y)$ car $J(Y)$ est fermé dans Z. □

Exercice. Montrer que s'il existe un opérateur $R \in \mathcal{L}(Y, Z)$ compact et surjectif, alors $\dim(Z) < \infty$.

COROLLAIRE 2.3. *Si* $T \in \mathcal{L}(X)$ *est compact et si* $J \in \mathcal{L}(X)$ *est un plongement, alors* $\dim\,\ker(T - J) < \infty$.

Démonstration. On applique le lemme avec $Y := \ker(T - J)$ et $Z := X$. L'opérateur $J_{|Y} : Y \to X$ est un plongement, et il est compact car $J_{|Y} = T_{|Y}$ par définition de Y. □

Lemme 2.4. *Soient Y et Z deux espaces de Banach, $K \in \mathcal{L}(Y,Z)$ un opérateur compact et $J \in \mathcal{L}(Y,Z)$ un plongement. Si $K - J$ est injectif, alors $K - J$ est un plongement.*

Démonstration. Supposons que $K - J$ ne soit pas un plongement. Alors, on peut trouver une suite $(y_n) \subseteq Y$ telle que $\|y_n\| = 1$ pour tout $n \in \mathbb{N}$ et $(K-J)y_n \to 0$. Comme K est un opérateur compact, on peut trouver une sous-suite (y_{n_k}) de (y_n) telle que $Ky_{n_k} \to z \in X$. Alors $Jy_{n_k} = Ky_{n_k} - (K-J)y_{n_k} \to z$. Comme J est un plongement, on en déduit que la suite (x_{n_k}) est de Cauchy dans X ; donc $y_{n_k} \to y \in Y$. On a $y \neq 0$ car $\|y\| = \lim \|y_{n_k}\| = 1$, et $(K-J)x = \lim (K-J)x_{n_k} = 0$; donc $K - J$ n'est pas injectif. □

Corollaire 2.5. *Si $K \in \mathcal{L}(X)$ est compact et si $J \in \mathcal{L}(X)$ est un plongement, alors $\mathrm{Im}(K-J)$ est un sous-espace fermé de X.*

Démonstration. Comme K est compact et que J est un plongement, $\ker(K-J)$ est de dimension finie. Donc, on peut trouver un sous-espace *fermé* $Y \subseteq X$ tel que $X = \ker(K-J) \oplus Y$. Alors Y est un espace de Banach, et $(K-J)_{|Y} : Y \to X$ est injectif. Donc $(K-J)_{|Y}$ est un plongement d'après le lemme puisque $K_{|Y}$ est compact et $J_{|Y}$ est un plongement ; et donc $(K-J)(Y)$ est fermé dans X. Mais $(K-J)(Y) = \mathrm{Im}(K-J)$ par le choix de Y, donc $\mathrm{Im}(K-J)$ est fermé. □

Lemme 2.6. *Soit Z un espace de Banach. Si $E \subseteq Z$ est un sous-espace fermé tel que $E \neq Z$, alors on peut trouver $x \in Z$ tel que $\|x\| = 1$ et $\mathrm{dist}(x,E) \geqslant 1/2$.*

Démonstration. Soit $z \in Z \backslash E$ quelconque. On a $\mathrm{dist}(z,E) > 0$ car E est fermé, donc on peut trouver $e \in E$ tel que $\|z - e\| \leqslant 2\,\mathrm{dist}(z,E)$. Alors $x := \frac{z-e}{\|z-e\|}$ convient, car on a $\mathrm{dist}(x,E) = \frac{1}{\|z-e\|}\,\mathrm{dist}\big(z-e,E\big) = \frac{1}{\|z-e\|}\,\mathrm{dist}(z,E)$. □

Lemme 2.7. *Soit $K \in \mathcal{L}(X)$ un opérateur compact, et soit $J \in \mathcal{L}(X)$ un plongement. Soit également $(E_n)_{n\geqslant 0}$ une suite de sous-espaces fermés de X. On suppose qu'on a $J(E_n) = E_n$ pour tout $n \in \mathbb{N}$, et on fait de plus l'une des deux hypothèses suivantes :*

(a) *la suite (E_n) est croissante et $(K-J)(E_{n+1}) \subseteq E_n$ pour tout $n \in \mathbb{N}$;*

(b) *la suite (E_n) est décroissante et $(K-J)(E_n) \subseteq E_{n+1}$ pour tout $n \in \mathbb{N}$.*

Alors la suite (E_n) est stationnaire *: il existe $n_0 \in \mathbb{N}$ tel que $\forall n \geqslant n_0 \ : \ E_n = E_{n_0}$.*

Démonstration. On traite seulement le cas (a) ; la preuve pour (b) est identique. Comme J est un plongement, on peut fixer une constante $c > 0$ telle que $\|Ju\| \geqslant c\,\|u\|$ pour tout $u \in X$.
Supposons que la suite (E_n) ne soit pas stationnaire. Alors, quitte à extraire une sous-suite, on peut supposer que (E_n) est strictement croissante (exo). D'après le Lemme 2.6, on peut trouver pour tout $n \geqslant 1$ un point $x_n \in E_n$ tel que $\|x_n\| = 1$ et $\mathrm{dist}(x_n, E_{n-1}) \geqslant 1/2$. Si $1 \leqslant p < q$, on a $Kx_q - Kx_p = (K-J)(x_q - x_p) + J(x_q - x_p) = Jx_q - z$, où $z := Jx_p + (K-J)x_p - (K-J)x_q \in E_p + E_{p-1} - E_{q-1} \subseteq E_{q-1} = J(E_{q-1})$. On a ainsi $z = Jy$ pour un certain $y \in E_{q-1}$, et donc $\|Kx_q - Kx_p\| = \|J(x_q - y)\| \geqslant c\,\|x_q - y\| \geqslant c/2$. Ceci étant vrai pour tous $1 \leqslant p < q$, on en déduit que la suite (Kx_n) ne possède aucune sous suite convergente ; ce qui contredit la compacité de K. □

Corollaire 2.8. *Si $K \in \mathcal{L}(X)$ est compact et si $\lambda \in \mathbb{K}\backslash\{0\}$, alors la suite croissante des* $\ker(K-\lambda I)^n$ *est stationnaire, et la suite décroissante des* $\mathrm{Im}(K-\lambda)^n$ *est stationnaire.*

Démonstration. On applique le lemme avec $J := \lambda I$ et $E_n := \ker(K-\lambda I)^n$ ou $E_n := \mathrm{Im}(K-\lambda I)^n$. On a bien $J(E_n) = E_n$ pour tout $n \in \mathbb{N}$; et (a) ou (b) est visiblement vérifiée. Il faut cependant vérifier que $\mathrm{Im}(K-\lambda I)^n$ est bien fermé pour tout $n \in \mathbb{N}$. C'est évident pour $n = 0$. Si $n \geqslant 1$, on voit en utilisant la formule du binôme que $(K-\lambda I)^n = R + \mu I$, où $\mu := (-\lambda)^n \neq 0$ et R est un opérateur compact ; donc $\mathrm{Im}(K-\lambda I)^n$ est fermé d'après le Corollaire 2.5. □

Preuve du Théorème 2.1. (1) Si $0 \notin \sigma(K)$, alors $K : X \to X$ est inversible, donc $\dim(X) < \infty$ par le Lemme 2.2 avec $J := K$.

(2) Montrons par exemple que $K-\lambda I$ injectif $\implies$ $K-\lambda I$ inversible. Suposons que $K-\lambda I$ soit injectif. Si $K-\lambda I$ n'est pas inversible, alors il n'est pas surjectif. Si on choisit $z \in X\backslash\mathrm{Im}(K-\lambda I)$, alors $\forall n \in \mathbb{N} \ : \ (K-\lambda I)^n z \in \mathrm{Im}(K-\lambda I)^n \backslash \mathrm{Im}(K-\lambda I)^{n+1}$ par injectivité de $K-\lambda I$; donc la suite des $\mathrm{Im}(K-\lambda I)^n$ est *strictement* décroissante, ce qui contredit le Corollaire 2.8.

(3) Par (2) et comme K et K^* sont compacts, on a $\sigma_p(K)\backslash\{0\} = \sigma(K)\backslash\{0\} = \sigma(K^*)\backslash\{0\} = \sigma_p(K^*)\backslash\{0\}$.

(4) Par les Lemmes 2.4 et 2.7 avec $J := \lambda I$, on voit que $\ker(K-\lambda I)$ est de dimension finie et que $\mathrm{Im}(K-\lambda I)$ est fermé.

Comme $\mathrm{Im}(K-\lambda I)$ est fermé, on sait que $X/\mathrm{Im}(K-\lambda I)$ est un espace de Banach et que $\Big(X/\mathrm{Im}(K-\lambda I)\Big)^*$ s'identifie à $\mathrm{Im}(K-\lambda I)^\perp = \ker(K^*-\lambda I)$. Donc $\dim\Big(X/\mathrm{Im}(K-\lambda I)\Big)^* < \infty$ car K^* est compact ; et donc, par le Corollaire 1.4 du Chapitre 5, $\dim\Big(X/\mathrm{Im}(K-\lambda I)\Big) < \infty$ avec plus précisément $\dim\Big(X/\mathrm{Im}(K-\lambda I)\Big) = \dim\Big(X/\mathrm{Im}(K-\lambda I)\Big)^* = \dim\ker(K^*-\lambda I)$. Si on veut absolument éviter les quotients, on peut écrire les choses ainsi : comme $\mathrm{Im}(K-\lambda I)$ est fermé, on a $\mathrm{Im}(K-\lambda I) = \big(\mathrm{Im}(K-\lambda I)^\perp\big)_\perp = \ker(K^*-\lambda I)_\perp$; et donc, puisque K^* est compact, $\mathrm{Im}(K-\lambda I)$ est de codimension finie, égale à $\dim\ker(K^*-\lambda I)$.

Par la formule du binôme (applicable car K et λI commutent), on a $(K-\lambda I)^n = \sum_{j=0}^n (-1)^{n-j}\binom{n}{j}\lambda^{n-j}K^j$. Comme K^j est compact pour tout $j \geqslant 1$, on en déduit que $(K-\lambda I)^n = R + (-1)^n\lambda^n I := R - \mu I$, où R est un opérateur compact. Donc on peut appliquer ce qui précède.

(5) Par le Corollaire 2.8, la suite des $\ker(K-\lambda I)^n$ et la suite des $\mathrm{Im}(K-\lambda I)^n$ sont stationnaires. Donc on peut trouver un entier n tel que $\ker(K-\lambda I)^{n+1} = \ker(K-\lambda I)^n$ et $\mathrm{Im}(K-\lambda I)^{n+1} = \mathrm{Im}(K-\lambda I)^n$.

Notons T la restriction de $K-\lambda I$ à $\mathrm{Im}(K-\lambda I)^n$. On a $\mathrm{Im}(T) = \mathrm{Im}(K-\lambda I)^{n+1} = \mathrm{Im}(K-\lambda I)^n$, donc $T : \mathrm{Im}(K-\lambda I)^n \to \mathrm{Im}(K-\lambda I)^n$ est un opérateur surjectif. De plus, T est aussi injectif. En effet : si $x \in \ker(T)$, alors $x \in \mathrm{Im}(K-\lambda I)^n$ et $(K-\lambda I)x = 0$; on a donc $x = (K-\lambda I)^n z$ pour un certain $z \in X$ tel que $(K-\lambda I)^{n+1}z = (K-\lambda I)x = 0$; donc $z \in \ker(K-\lambda I)^{n+1} = \ker(K-\lambda I)^n$, et donc $x = (K-\lambda I)^n z = 0$. Ainsi, $T : \mathrm{Im}(K-\lambda I)^n \to \mathrm{Im}(K-\lambda I)^n$ est inversible.

Soit $S := T^{-1} : \mathrm{Im}(K-\lambda I)^n \to \mathrm{Im}(K-\lambda I)^n$, et soit $\pi : X \to X$ l'opérateur défini par $\pi(x) := S^n(K-\lambda I)^n x$. Comme $S^n : \mathrm{Im}(K-\lambda I)^n \to \mathrm{Im}(K-\lambda I)^n$ est inversible, on a $\mathrm{Im}(\pi) = \mathrm{Im}(K-\lambda I)^n$. Si $v \in \mathrm{Im}(\pi) = \mathrm{Im}(K-\lambda I)^n$, on peut écrire $\pi(v) = S^n T^n v$, donc $\pi(v) = v$. Donc π est une projection d'image $\mathrm{Im}(K - \lambda I)^n$. De plus, comme S^n est injectif, on a $\ker(\pi) = \ker(K - \lambda I)^n$. On a ainsi trouvé une projection d'image $\mathrm{Im}(K - \lambda I)^n$ et de noyau $\ker(T - \lambda I)^n$, donc $X = \ker(K - \lambda I)^n \oplus \mathrm{Im}(K - \lambda I)^n$.

(6) Avec les notations de (5), posons $E := \ker(K - \lambda I)^n$ et $F := \mathrm{Im}(K - \lambda I)^n$. Alors E et F sont invariants par $K - \lambda I$, on a $X = E \oplus F$, et $(K - \lambda I)_{|F}$ est un isomorphisme de F sur F. On en déduit que $\mathrm{Im}(K - \lambda I) = (K - \lambda I)(E) \oplus F = \mathrm{Im}\big((K - \lambda I)_{|E}\big) \oplus F$, et que $\ker(K - \lambda I) = \ker(K - \lambda I) \cap E = \ker\big((K - \lambda I)_{|E}\big)$. Comme $X = E \oplus F$ et $\mathrm{Im}(K - \lambda I) = \mathrm{Im}\big((K - \lambda I)_{|E}\big) \oplus F$, la codimension de $\mathrm{Im}(K - \lambda I)$ dans X est égale à la codimension de $\mathrm{Im}\big((K - \lambda I)_{|E}\big)$ dans E ; et comme $\dim(E) < \infty$, on a donc $\dim\big(X/\mathrm{Im}(K - \lambda I)\big) = \dim(E) - \dim \mathrm{Im}\big((K - \lambda I)_{|E}\big) = \dim \ker\big((K - \lambda I)_{|E}\big) = \dim\ker(K - \lambda I)$. Enfin, on a vu dans la preuve de (4) que $\dim\Big(X/\mathrm{Im}(K - \lambda I)\Big) = \dim\Big(X/\mathrm{Im}(K - \lambda I)\Big)^* = \dim \ker(K^* - \lambda I)$.

(7) La preuve repose sur le fait suivant.

Fait. Si $\lambda \in \sigma(K)$ et $\lambda \neq 0$, alors λ est un point isolé de $\sigma(K)$.

Preuve du Fait. Il s'agit de montrer qu'il existe un voisinage V de λ dans $\mathbb{C}$ tel que $V \cap \sigma(K)$ ne contienne pas d'autre point que λ. Autrement dit, en posant $T := K - \lambda I$, on cherche $\varepsilon > 0$ tel que $T - \mu I$ est inversible pour tout $\mu \in D(0, \varepsilon)$ avec $\mu \neq 0$.
Par (5), on sait qu'il existe un entier n et deux sous-espaces fermés $E, F \subseteq X$ invariants par T tels que $X = E \oplus F$, $\dim(E) < \infty$, $(T_{|E})^n = 0$ et $T_{|F} : F \to F$ est inversible. Comme l'ensemble des opérateurs inversibles est un ouvert de $\mathcal{L}(F)$, on peut trouver $\varepsilon > 0$ tel que $(T - \mu I)_{|F} : F \to F$ est inversible pour tout $\mu \in D(0, \varepsilon)$. De plus, si $\mu \neq 0$, alors $(T - \mu I)_{|E} : E \to E$ est inversible car $\dim(E) < \infty$ et la seule valeur propre de $T_{|E}$ est 0 (puisque $(T_{|E})^n = 0$). Donc, pour $\mu \in D(0, \varepsilon)$ et $\mu \neq 0$, on a bien que $T - \mu I$ est inversible. □

Par le Fait, on voit que pour tout $\varepsilon > 0$, l'ensemble $E_\varepsilon := \{\lambda \in \sigma(K);\ |\lambda| \geqslant \varepsilon\}$ est *fini*. En effet, comme $E_\varepsilon \subseteq \sigma(K)\backslash\{0\}$, le Fait dit que tout point $\lambda \in E_\varepsilon$ possède un voisinage ouvert V_λ tel que $V_\lambda \cap \sigma(K) = \{\lambda\}$. Comme de plus E_ε est compact, on peut trouver $\lambda_1, \dots, \lambda_N \in E_\varepsilon$ tels que $E_\varepsilon \subseteq V_{\lambda_1} \cup \dots \cup V_{\lambda_N}$; et on a ainsi $E_\varepsilon = \{\lambda_1, \dots, \lambda_N\}$.
Supposons que $\sigma(K)$ soit infini Soit $(\varepsilon_k)_{k \in \mathbb{N}}$ une suite décroissante de nombres réels > 0 telle que $\varepsilon_k \to 0$ et $\sigma(K) \subseteq D(0, \varepsilon_0)$. On a $\sigma(K) = \{0\} \cup \bigcup_{n \in \mathbb{N}} \Lambda_k$, où $\Lambda_k := \{\lambda \in \sigma(K);\ \varepsilon_{k+1} \leqslant |\lambda| < \varepsilon_k\}$. Tous les ensembles Λ_k sont finis puisque $\Lambda_k \subseteq E_{\varepsilon_{k+1}}$. Donc, en énumérant les Λ_k non vides, on obtient une suite (λ_n) telle que $\lambda_n \to 0$ et $\sigma(K) = \{0\} \cup \{\lambda_n;\ n \in \mathbb{N}\}$.

□

3. Calcul fonctionnel

Rappel. Soit A une algèbre de Banach unitaire sur $\mathbb{C}$, et soit $a \in A$. Si $P = c_0\mathbf{1} + c_1\mathbf{z} + \dots + c_d\mathbf{z}^d \in \mathbb{C}[\mathbf{z}]$, on peut définir

$$P(a) := c_0\mathbf{1} + c_1 a + \dots + c_d a^d.$$

L'application $P \mapsto P(a)$ est un homomorphisme de $\mathbb{C}[\mathbf{z}]$ dans A, et on a vu que pour tout $P \in \mathbb{C}[\mathbf{z}]$, on a

$$\sigma\big(P(a)\big) = P(\sigma(a)) = \{P(\lambda);\ \lambda \in \sigma(a)\}.$$

On aimerait maintenant définir $f(a)$ pour des fonctions f plus générales que les fonctions polynomiales, et que l'application $f \mapsto f(a)$ conserve de bonnes propriétés.

3.1. Calcul fonctionnel holomorphe. Dans cette section, on va voir qu'on peut définir $f(a)$ pour toute fonction f *holomorphe au voisinage du spectre de* a. On va commencer par le cas des fonctions rationnelles, qui est purement “algébrique” et est indispensable pour traiter ensuite le cas général. Puis on considérera le cas d'une fonction f holomorphe *dans un disque* contenant $\sigma(a)$. Ce cas n'est pas indispensable pour la suite ; mais comme il est simple et naturel, il aurait été dommage de ne pas en parler séparément. Enfin, on traitera le cas général, qui demande un peu de travail.

3.1.1. *Fonctions rationnelles.*

Notation. On notera $\mathcal{R}(a)$ l'ensemble des fractions rationnelles $R \in \mathbb{C}(\mathbf{z})$ ne possèdant pas de pôles dans $\sigma(a)$. Autrement dit, $R \in \mathcal{R}(a)$ si et seulement si $R = \frac{P}{Q}$ où P et Q sont des polynômes et Q n'a pas de racines dans $\sigma(a)$.

Remarque. $\mathcal{R}(a)$ est une sous-algèbre de $\mathbb{C}(\mathbf{z})$.

Fait 3.1. *Si $R = \frac{P}{Q} \in \mathcal{R}(a)$, alors $Q(a)$ est inversible dans A et $P(a)Q(a)^{-1} = Q(a)^{-1}P(a)$. De plus, si $R = \frac{P_1}{Q_1} = \frac{P_2}{Q_2}$, alors $P_1(a)Q_1(a)^{-1} = P_2(a)Q_2(a)^{-1}$.*

Démonstration. Comme Q n'a pas de racines dans $\sigma(a)$, on a $0 \notin Q(\sigma(a))$. Donc $0 \notin \sigma\big(Q(a)\big)$, autrement dit $Q(a)$ est inversible dans A. Comme $P(a)Q(a) = (PQ)(a) = (QP)(a) = Q(a)P(a)$, on a $Q(a)^{-1}P(a) = P(a)Q(a)^{-1}$.
Si $R = \frac{P_1}{Q_1} = \frac{P_2}{Q_2}$, alors $P_1Q_2 = P_2Q_1$, donc $P_1(a)Q_2(a) = P_2(a)Q_1(a) = Q_1(a)P_2(a)$, et donc $Q_1(a)^{-1}P_1(a) = P_2(a)Q_2^{-1}(a)$. □

Définition 3.2. *Si $R = \frac{P}{Q} \in \mathcal{R}(a)$, on pose $R(a) := P(a)Q(a)^{-1} = Q(a)^{-1}P(a)$.*

Proposition 3.3. *L'application $R \mapsto R(a)$ est un homomorphisme de $\mathcal{R}(a)$ dans A ; plus précisément, c'est l'unique homomorphisme $\Phi : \mathcal{R}(a) \to A$ dans A tel que $\Phi(\mathbf{z}) = a$. De plus, si $R \in \mathcal{R}(a)$, alors $\sigma(R(a)) = R\big(\sigma(a)\big)$.*

Démonstration. (i) Si $R_1 = \frac{P_1}{Q_1}$ et $R_2 = \frac{P_2}{Q_2}$, alors on doit *a priori* écrire

$$(R_1R_2)(a) = (P_1P_2)(a)\big(Q_1Q_2(a)\big)^{-1} = P_1(a)P_2(a)Q_2(a)^{-1}Q_1(a)^{-1}.$$

Mais tout commute par le Fait 3.1 ; donc $(R_1R_2)(a) = P_1(a)Q_1(a)^{-1}P_2(a)Q_2(a)^{-1} = R_1(a)R_2(a)$. On montre de même que $(R_1 + R_2)(a) = R_1(a) + R_2(a)$ (exo). Comme $\mathbf{1}(a) = \mathbf{1}$, l'aplication $R \mapsto R(a)$ est donc un homomorphisme de $\mathcal{R}(a)$ dans A, et bien sûr $\mathbf{z}(a) = a$. Si $\Phi : \mathcal{R}(a) \to A$ est un autre homomorphisme tel que $\Phi(\mathbf{z}) = a$, alors $\Phi(P) = P(a)$ pour tout polynôme P. Donc, si $R = \frac{P}{Q} = PQ^{-1}$, alors $\Phi(R) = \Phi(P)\Phi(Q)^{-1} = P(a)Q(a)^{-1} = R(a)$.
(ii) Soit $R = \frac{P}{Q} \in \mathcal{R}(a)$. Si $\mu \in \mathbb{C}$, alors

$$R(a) - \mu\mathbf{1} = P(a)Q(a)^{-1} - \mu\mathbf{1} = \big(P(a) - \mu Q(a)\big)Q(a)^{-1}.$$

Comme $Q(a)^{-1}$ est inversible on en déduit que $R(a) - \mu\mathbf{1}$ est inversible si et seulement si $P(a) - \mu Q(a)$ est inversible. Autrement dit :

$$\mu \in \sigma(a) \iff 0 \in \sigma\big(P(a) - \mu Q(a)\big).$$

Mais $P(a) - \mu Q(a) = (P - \mu Q)(a)$; donc $\sigma\big(P(a) - \mu Q(a)\big) = (P - \mu Q)(\sigma(a)) = \big\{P(\lambda) - \mu\, Q(\lambda);\ \lambda \in \sigma(a)\big\}$. Donc

$$\mu \in \sigma(R(a)) \iff \exists\lambda \in \sigma(a)\ :\ P(\lambda) - \mu Q(\lambda) = 0 \iff \exists\lambda \in \sigma(a)\ :\ \mu = R(\lambda).$$

□

3.1.2. *Séries entières.* Si Ω est un ouvert de $\mathbb{C}$, on note $H(\Omega)$ l'ensemble de toutes les fonctions holomorphes sur Ω. Évidemment, $H(\Omega)$ est une $\mathbb{C}$-algèbre.

Lemme 3.4. *Soit $R > r(a)$. Si $f \in H\big(D(0,R)\big)$ et si on écrit $f(z) = \sum_{k=0}^{\infty} c_k z^k$, alors la série $\sum c_k a^k$ converge dans A. On peut donc poser*

$$f(a) := \sum_{k=0}^{\infty} c_k a^k.$$

Démonstration. Choisissons r tel que $r(a) < r < R$. Comme $r(a) = \lim_{k\to\infty} \|a^k\|^{1/k}$, on peut trouver k_0 tel que $\forall k \geqslant k_0\ :\ \|a^k\|^{1/k} \leqslant r$. Alors $\|c_k a^k\| \leqslant |c_k| r^k$ pour tout $k \geqslant k_0$, donc la série $\sum \|c_k a^k\|$ est convergente, et donc la série $\sum c_k a^k$ converge dans A puisque A est un espace de Banach. □

Exemple. Pour tout $a \in A$, on a $e^a := \displaystyle\sum_{k=0}^{\infty} \frac{a^k}{k!}\cdot$

Exercice. Montrer que si $a, b \in A$ et si $ab = ba$, alors $e^{a+b} = e^a e^b$. En déduire que si $a \in A$, alors e^a est inversible et $\big(e^a\big)^{-1} = e^{-a}$

Lemme 3.5. (Formule de Cauchy)
Soit $R > r(a)$, et soit r tel que $r(a) < r < R$. Pour toute fonction fonction $f \in H(D(0,R))$, on a

$$f(a) = \int_0^{2\pi} f(re^{i\theta})(re^{i\theta}\mathbf{1} - a)^{-1}\, re^{i\theta}\, \frac{d\theta}{2\pi}\cdot$$

Démonstration. Il faut d'abord expliquer pourquoi l'énoncé a un sens. Comme $|re^{i\theta}| = r > r(a)$, on voit que $re^{i\theta} - a$ est inversible pour tout $\theta \in [0, 2\pi]$. Donc l'application $\theta \mapsto f(re^{i\theta})(re^{i\theta}\mathbf{1} - a)^{-1}\, re^{i\theta}$ est bien définie et continue sur $[0, 2\pi]$, à valeurs dans l'espace de Banach A, et donc on peut l'intégrer sur $[0, 2\pi]$.
Écrivons $f(z) = \sum_{k=0}^{\infty} c_k z^k$. Alors

$$f(re^{i\theta})(re^{i\theta}\mathbf{1} - a)^{-1}\, re^{i\theta} = \sum_{k=0}^{\infty} c_k (re^{i\theta})^{k+1}(re^{i\theta}\mathbf{1} - a)^{-1},$$

et la série converge normalement sur $[0, 2\pi]$. Donc on peut écrire

$$\int_0^{2\pi} f(re^{i\theta})(re^{i\theta}\mathbf{1} - a)^{-1}\, re^{i\theta}\, \frac{d\theta}{2\pi} = \sum_{k=0}^{\infty} c_k \int_0^{2\pi} (re^{i\theta})^{k+1}(re^{i\theta}\mathbf{1} - a)^{-1}\cdot$$

Il suffit donc de montrer que

$$\forall k \in \mathbb{N}\ :\ a^k = \int_0^{2\pi} r^{k+1} e^{i(k+1)\theta}(re^{i\theta}\mathbf{1} - a)^{-1}\, \frac{d\theta}{2\pi}\cdot$$

Comme $|re^{i\theta}| = r > r(a)$, on a

$$(re^{i\theta}\mathbf{1} - a)^{-1} = \sum_{n=0}^{\infty} \frac{a^n}{(re^{i\theta})^{n+1}},$$

où la série converge dans A. La série converge en fait normalement sur $[0, 2\pi]$, donc on peut écrire

$$\int_0^{2\pi} r^{k+1} e^{i(k=1)\theta} (re^{i\theta}\mathbf{1} - a)^{-1} \frac{d\theta}{2\pi} = \sum_{n=0}^{\infty} a^n r^{k-n} \int_0^{2\pi} e^{i(k-n)\theta} \frac{d\theta}{2\pi} = a^k,$$

car toutes les intégrales $\int_0^{2\pi} e^{i(k-n)\theta} \frac{d\theta}{2\pi}$ valent 0, sauf celle pour $n := k$ qui vaut 1. □

PROPOSITION 3.6. *Soit $R > r(a)$. L'application $f \mapsto f(a)$ est un homomorphisme de $H\big(D(0,R)\big)$ dans A ; plus précisément, c'est l'unique homomorphisme continu Φ : $H\big(D(0,R)\big) \to A$ tel que $\Phi(\mathbf{z}) = a$.*

Démonstration. On a besoin des deux faits suivants.

FAIT 1. Soit r tel que $r(a) < r < R$, et soit $K_r := \overline{D}(0,r)$. Il existe une constante $C = C_r$ telle que

$$\forall f \in H\big(D(0,R)\big) \; : \; \|f(a)\| \leqslant C \, \sup\big\{|f(z)|;\; z \in K_r\big\}.$$

Preuve du Fait 1. D'après la Formule de Cauchy, on a

$$\begin{aligned} \|f(a)\| &= \left\| \int_0^{2\pi} f(re^{i\theta})(re^{i\theta}\mathbf{1} - a)^{-1}\, re^{i\theta} \frac{d\theta}{2\pi} \right\| \\ &\leqslant \int_0^{2\pi} |f(re^{i\theta})| \, \|(re^{i\theta}\mathbf{1} - a)^{-1}\| \, \frac{d\theta}{2\pi} \\ &\leqslant C \, \sup\big\{|f(z)|;\; z \in K_r\big\}, \end{aligned}$$

où $C = C_r := \int_0^{2\pi} \|(re^{i\theta}\mathbf{1} - a)^{-1}\| \, \frac{d\theta}{2\pi}$. □

FAIT 2. Les fonctions polynomiales sont denses dans $H\big(D(0,R)\big)$.

Preuve du Fait 2. C'est évident : si $f \in H\big(D(0,R)\big)$ s'écrit $f(z) = \sum_{k=0}^{\infty} c_k z^k$ et si on pose $P_n(z) := \sum_{k=0}^{n} c_k z^k$, alors $P_n(z) \to f(z)$ uniformément sur tout compact de $D(0,R)$, *i.e.* $P_n \to f$ pour la topologie de $H\big(D(0,R)\big)$. □

Par le Fait 1 et par définition de la topologie de $H\big(D(0,R)\big)$, l'application linéaire $f \mapsto f(a)$ est continue de $H\big(D(0,R)\big)$ dans A. Comme $(PQ)(a) = P(a)Q(a)$ pour tous polynômes P, Q et comme les fonctions polynomiales sont denses dans $H\big(D(0,R)\big)$ par le Fait 2, on en déduit qu'on a $(fg)(a) = f(a)g(a)$ pour toutes $f, g \in H\big(D(0,R)\big)$. Donc l'application $f \mapsto f(a)$ est un homomorphisme de $H\big(D(0,R)\big)$ dans A.
Si $\Phi : H\big(D(0,R)\big) \to A$ est un autre homomorphisme continu tel que $\Phi(\mathbf{z}) = a$, alors $\Phi(P) = P(a)$ pour toute fonction polynomiale P, et donc $\Phi(a) = f(a)$ pour toute $f \in H\big(D(0,R)\big)$ par densité des fonctions polynomiales dans $H\big(D(0,R)\big)$. □

Remarque. Il est également vrai qu'on a $\sigma(f(a)) = f(\sigma(a))$ pour toute $f \in H\big(D(0,R)\big)$; mais à ce stade, ce n'est pas du tout évident ! Comme dans le cas des polynômes, on montre que $f(\sigma(a)) \subseteq \sigma(f(a))$ en factorisant. Pour l'inclusion inverse, il suffirait de savoir que si une fonction $f \in H\big(D(0,R)\big)$ ne s'annule pas sur $\sigma(a)$, alors f est inversible dans $H\big(D(0,R)\big)$, autrement dit ne s'annule pas sur le disque $D(0,R)$; mais ceci est visiblement faux.

3.1.3. *Cas général.*

Définition 3.7. *On note $\mathcal{H}(a)$ l'ensemble de toutes les fonctions f définies et holomorphes sur un ouvert $\Omega = \Omega_f \subseteq \mathbb{C}$ contenant $\sigma(a)$. On identifie deux fonctions $f, g \in \mathcal{H}(a)$ s'il existe un ouvert $U \supseteq \sigma(a)$ tel que $f \equiv g$ sur U.*

Remarque 1. $\mathcal{H}(a)$ est une $\mathbb{C}$-algèbre commutative unitaire.

Remarque 2. Si $R \in \mathcal{R}(a)$, alors la *fonction* rationnelle associée $\widetilde{R}$ appartient à $\mathcal{H}(a)$, et l'application $R \mapsto \widetilde{R}$ est un homomorphisme injectif. Donc, on peut considérer que $\mathcal{R}(a) \subseteq \mathcal{H}(a)$.

Démonstration. Tout est évident sauf l'injectivité : il faut voir que si $R_1, R_2 \in \mathcal{R}(a)$ et si $R_1(z) = R_2(z)$ sur un ouvert $U \supseteq \sigma(a)$, alors $R_1 = R_2$. Si on écrit $R_i = \frac{P_i}{Q_i}$, alors $P_1(z)Q_2(z) \equiv P_2(z)Q_1(z)$ sur U. Mais P_1Q_2 et P_2Q_1 sont des polynômes et U est infini ; donc $P_2Q_1 = P_1Q_2$, et donc $R_1 = R_2$. $\square$

Remarque 3. Une $f \in \mathcal{H}(a)$ est inversible dans $\mathcal{H}(a)$ si et seulement si $f(z) \neq 0$ pour tout $z \in \sigma(a)$.

Démonstration. La fonction f est définie et holomorphe sur un ouvert $\Omega \supseteq \sigma(a)$. Si f est inversible dans $\mathcal{H}(a)$, on peut trouver une fonction g holomorphe au voisinage de $\sigma(a)$ tellle que $f(z)g(z) \equiv 1$ sur un ouvert $U \supseteq \sigma(a)$; en particulier $f(z)g(z) = 1$ pour tout $z \in \sigma(a)$, et donc $f(z) \neq 0$ sur $\sigma(a)$. Inversement, supposons que $f(z) \neq 0$ pour tout $z \in \sigma(a)$. Comme f est continue sur Ω, l'ensemble $\Omega' := \{z \in \Omega;\ f(z) \neq 0\}$ est un ouvert de $\mathbb{C}$, et $\Omega' \supseteq \sigma(a)$. La fonction $g := 1/f$ est bien définie et holomorphe sur Ω', donc $g \in \mathcal{H}(a)$; et $gf = \mathbf{1} = fg$ dans $\mathcal{H}(a)$. $\square$

Définition 3.8. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{H}(a)$, et soit $f \in \mathcal{H}(a)$. On dit que "f_n tend vers f dans $\mathcal{H}(a)$" s'il existe un ouvert $U \supseteq \sigma(a)$ tel que*

- *f et les f_n sont définies et holomorphes sur U ;*
- *$f_n \to f$ dans $H(U)$, i.e. $f_n(z) \to f(z)$ uniformément sur tout compact de U.*

Remarque. On dira qu'une application $\Phi = \mathcal{H}(a) \to A$ est *séquentiellement continue* si, pour toute suite $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}(a)$ telle que $f_n \to f \in \mathcal{H}(a)$, on a que $\Phi(f_n) \to \Phi(f)$.

Théorème 3.9. *Il existe un unique homomorphisme séquentiellement continu $\Phi_a : \mathcal{H}(a) \to A$ tel que $\Phi_a(\mathbf{z}) = a$. Si on pose $f(a) := \Phi_a(f)$, alors les choses suivantes ont lieu pour toute $f \in \mathcal{H}(a)$:*

- *$f(a)$ est limite d'éléments de A de la forme $R(a)$, où $R \in \mathcal{R}(a)$;*
- *$f(a)b = bf(a)$ pour tout $b \in A$ tel que $ab = ba$;*
- *$\sigma(f(a)) = f(\sigma(a))$.*

On dit que Φ_a est le ***calcul fonctionnel holomorphe*** *associé à a.*

L'idée de la preuve de ce théorème est très simple : on va montrer que le "calcul fonctionnel rationnel" $\mathcal{R}(a) \ni R \mapsto R(a)$ possède une bonne propriété de continuité, qui permet de l'étendre à $\mathcal{H}(a)$ "par densité". Les détails techniques sont cependant non triviaux. On va avoir besoin de deux lemmes.

LEMME 3.10. *Si V est un ouvert borné tel que $\sigma(a) \subseteq V$, il existe une constante $C = C_V < \infty$ telle que : pour toute fraction rationnelle R n'ayant pas de pôles dans $\overline{V}$, on a*

$$\|R(a)\| \leqslant C_V \sup\left\{|R(z)|;\ z \in \overline{V}\right\}.$$

Démonstration. Si V était un disque $D(0,r)$, il suffirait d'appliquer la formule de Cauchy comme dans la preuve de la Proposition 3.6 (Fait 1). Pour un ouvert V quelconque, l'idée est la même mais la mise en oeuvre est plus délicate.
Dans ce qui suit, on note m la mesure de Lebesgue sur $\mathbb{C} = \mathbb{R}^2$. D'autre part, si φ est une fonction de classe $\mathcal{C}^1$ sur un ouvert $\Omega \subseteq \mathbb{C} = \mathbb{R}^2$, on pose

$$\overline{\partial}\varphi := \frac{1}{2}\left(\partial_x\varphi + \partial_y\varphi\right).$$

Avec cette notation, on voit en particulier qu'une fonction $\varphi \in \mathcal{C}^1(\Omega)$ est holomorphe sur Ω si et seulement si $\overline{\partial}\varphi = 0$ (*équation de Cauchy-Riemann*). Enfin, on note $\mathcal{C}^1_{00}(\mathbb{C})$ l'ensemble des fonctions de classe $\mathcal{C}^1$ à support compact sur $\mathbb{C}$.

FAIT 1. Pour toute fonction $\varphi \in \mathcal{C}^1_{00}(\mathbb{C})$ et pour tout $z \in \mathbb{C}$, la fonction $u \mapsto \frac{\overline{\partial}\varphi(u)}{u-z}$ est intégrable sur $\mathbb{C} = \mathbb{R}^2$, et on a

$$\varphi(z) = -\frac{1}{\pi}\int_{\mathbb{C}} \frac{\overline{\partial}\varphi(u)}{u-z}\, dm(u).$$

Preuve du Fait 1. Par translation, il suffit de démontrer le résultat pour $z := 0$. La fonction $u \mapsto 1/u$ est définie sur $\mathbb{C}\backslash\{0\}$, donc m-presque partout, et elle est intégrable sur tout disque $\overline{D}(0,R) \subseteq \mathbb{C}$, car

$$\int_{\overline{D}(0,R)} \frac{1}{|z|}\, dm(u) = \int_0^R\int_0^{2\pi} \frac{1}{r}\, r dr d\theta = 2\pi R < \infty.$$

Comme $\overline{\partial}\phi$ est continue à support compact, on en déduit que la fonction $u \mapsto \frac{\overline{\partial}\phi(u)}{u}$ est intégrable sur $\mathbb{C}$.
Si on passe en coordonnées polaires, $u = (x,y) = re^{i\theta}$, on trouve (exo classique)

$$\overline{\partial}\varphi = \frac{e^{i\theta}}{2}\left(\frac{\partial\varphi}{\partial r} + \frac{i}{r}\frac{\partial\varphi}{\partial\theta}\right).$$

Donc

$$\begin{aligned}
\int_{\mathbb{C}} \frac{\overline{\partial}\varphi(u)}{u}\, dm(u) &= \int_0^\infty\int_0^{2\pi} \frac{e^{i\theta}}{2}\left(\frac{\partial\varphi}{\partial r} + \frac{i}{r}\frac{\partial\varphi}{\partial\theta}\right) \times \frac{1}{re^{i\theta}}\, r dr d\theta\\
&= \frac{1}{2}\int_0^{2\pi}\left(\int_0^\infty \frac{\partial\varphi}{\partial r}\, dr\right) d\theta + \frac{1}{2}\int_0^\infty \frac{i}{r}\left(\int_0^{2\pi}\frac{\partial\varphi}{\partial\theta}\, d\theta\right) dr\\
&= \frac{1}{2}\int_0^{2\pi}\Big[\varphi(re^{i\theta})\Big]_{r=0}^{r=\infty} d\theta + \frac{1}{2}\int_0^\infty \frac{i}{r}\Big[\varphi(re^{i\theta})\Big]_{\theta=0}^{\theta=2\pi} dr\\
&= -\pi\,\varphi(0) + 0.
\end{aligned}$$

□

Fait 2. Si $\chi \in \mathcal{C}^1_{00}(\mathbb{C})$ est telle que $\chi(u) \equiv 0$ sur un ouvert $U \subseteq \mathbb{C}$, alors

$$\forall \alpha \in U \ \forall n \in \mathbb{Z} \ : \ \int_{\mathbb{C}} \overline{\partial}\chi(u)(u-\alpha)^n dm(u) = 0.$$

Preuve du Fait 2. Comme $\overline{\partial}\chi :\equiv 0$ sur U, les fonctions $u \mapsto \overline{\partial}\chi(u)(u-\alpha)^n$ sont bien définies sur $\mathbb{C}$ (même pour $n < 0$), et elles sont continues à support compact. Donc les intégrales ont un sens.
Par le Fait 1, on a $\int_{\mathbb{C}} \overline{\partial}\chi(u)(u-\alpha)^{-1}dm(u) = -\pi\chi(\alpha) = 0$ pour tout $\alpha \in U$. En dérivant par rapport à α, on en déduit $\int_{\mathbb{C}} \overline{\partial}\chi(u)(u-\alpha)^n dm(u) = 0$ pour tout $\alpha \in U$ et pour tout $n \leqslant -1$.
Si $n \geqslant 0$, la fonction $f_n(u) := (u-\alpha)^n$ est holomorphe sur $\mathbb{C}$, donc $\overline{\partial} f_n \equiv 0$ et donc $\overline{\partial}\chi(u)(u-\alpha)^n = \overline{\partial}(\chi f_n)(u)$ puisque $\overline{\partial}(\chi f_n) = (\overline{\partial}\chi) f_n + \chi(\overline{\partial} f_n)$. La fonction χf_n est de classe $\mathcal{C}^1$ à support compact, donc $\int_{\mathbb{C}} \overline{\partial}(\chi f_n)\, dm = 0$, et donc $\int_{\mathbb{C}} \overline{\partial}\chi(u)(u-\alpha)^n dm(u) = 0$. □

Le fait suivant joue le rôle de la formule de Cauchy.

Fait 3. Soit W un ouvert de $\mathbb{C}$ tel que $\sigma(a) \subseteq W$ et $\overline{W} \subseteq V$, et soit $\chi \in \mathcal{C}^1_{00}(\mathbb{C})$ telle que $\chi \equiv 1$ sur W et $\chi \equiv 0$ en dehors de V. Pour toute fraction rationnelle R n'ayant pas de pôles dans $\overline{V}$, on a

$$R(a) = -\frac{1}{\pi}\int_{\overline{V}\backslash W} \overline{\partial}\chi(u)R(u)(a-u\mathbf{1})^{-1}\, dm(u).$$

Preuve du Fait 3. Comme $\sigma(a) \subseteq W$ et que R n'a pas de pôles dans $\overline{V}$, la fonction $u \mapsto \overline{\partial}\chi(u)R(u)(a-u\mathbf{1})^{-1}$ est bien définie et continue sur le compact $\overline{V}\backslash W$, à valeurs dans l'espace de Banach A ; donc on peut l'intégrer sur $\overline{V}\backslash W$. Par ailleurs, la fonction $\overline{\partial}\chi$ est identiquement nulle sur W et en dehors de $\mathbb{C}\backslash\overline{V}$. Donc, au lieu de $\int_{\overline{V}\backslash W}$, on peut écrire $\int_{\mathbb{C}}$; et donc il s'agit de montrer que

$$R(a) = -\frac{1}{\pi}\int_{\mathbb{C}} \overline{\partial}\chi(u)R(u)(a-u\mathbf{1})^{-1}\, dm(u).$$

La fraction rationelle R peut se décomposer en éléments simples : elle est donc somme d'un polynôme P et de termes de la forme $\frac{1}{(\mathbf{z}-\alpha)^k}$, avec $k \in \mathbb{N}$ et $\alpha \in \mathbb{C}\backslash\overline{V}$. De plus, si on fixe $\alpha \in \mathbb{C}\backslash\overline{V}$, le polynôme P est somme de termes de la forme $(\mathbf{z}-\alpha)^k$, $k \in \mathbb{N}$. Donc, pour démontrer la formule souhaitée, il suffit de prouver que

$$\forall \alpha \in \mathbb{C}\backslash\overline{V} \ \forall n \in \mathbb{Z} \ : \ (a-\alpha\mathbf{1})^n = -\frac{1}{\pi}\int_{\mathbb{C}} \overline{\partial}\chi(u)(u-\alpha)^n(a-u\mathbf{1})^{-1}\, dm(u).$$

Dans la suite, on fixe $\alpha \in \mathbb{C}\backslash\overline{V}$, et pour $n \in \mathbb{Z}$ on pose

$$b_n := \int_{\mathbb{C}} \overline{\partial}\chi(u)(u-\alpha)^n(a-u\mathbf{1})^{-1}\, dm(u).$$

D'après l'équation résolvante, on a

$$\forall u \in \mathbb{C}\backslash\sigma(a) \ : \ (a-u\mathbf{1})^{-1} = (a-\alpha\mathbf{1})^{-1} + (a-\alpha\mathbf{1})^{-1}(u-\alpha)(a-u\mathbf{1})^{-1}.$$

Donc on peut écrire

$$\begin{aligned} b_n = (a-\alpha\mathbf{1})^{-1} &\int_{\overline{V}\backslash W} \overline{\partial}\chi(u)(u-\alpha)^n dm(u) \\ &+ (a-\alpha\mathbf{1})^{-1} \int_{\overline{V}\backslash W} \overline{\partial}\chi(u)(u-\alpha)^{n+1}(a-u\mathbf{1})^{-1}\, dm(u); \end{aligned}$$

autrement dit

$$b_n = (a-\alpha\mathbf{1})^{-1} \int_{\overline{V}\backslash W} \overline{\partial}\chi(u)(u-\alpha)^n dm(u) + (a-\alpha\mathbf{1})^{-1} b_{n+1}.$$

De plus, comme $\chi \equiv 0$ sur $U := \mathbb{C}\backslash\overline{V}$, on a

$$\int_{\mathbb{C}} \overline{\partial}\chi(u)(u-\alpha)^n dm(u) = 0 \quad \text{d'après le Fait 2.}$$

On obtient ainsi

$$b_{n+1} = (a-\alpha\mathbf{1})\, b_n \qquad \text{pour tout } n \in \mathbb{Z}.$$

Pour conclure, il suffit donc de montrer que $\frac{1}{\pi}\, b_0 = -\mathbf{1}$, autrement dit

$$\frac{1}{\pi}\int_{\mathbb{C}} \overline{\partial}\chi(u)(a-u\mathbf{1})^{-1} dm(u) = -\mathbf{1}.$$

Pour cela, l'idée est de *changer de fonction* χ. Soit $R > \|a\|$, et soit $\eta \in \mathcal{C}^1_{00}(\mathbb{C})$ telle que $\eta(u) \equiv 1$ sur $D(0,R) \supseteq \sigma(a)$. Alors $\chi(u) - \eta(u) \equiv 0$ au voisinage de $\sigma(a)$. Donc si on pose $f(u) := (a-u\mathbf{1})^{-1}$, la fonction $(\chi-\eta)f$ est bien définie sur $\mathbb{C}$ et de classe $\mathcal{C}^1$ à support compact (à valeurs dans A) ; et comme f est holomorphe sur $\mathbb{C}\backslash\sigma(a)$, on peut écrire $\overline{\partial}\chi(u)(a-u\mathbf{1})^{-1} - \overline{\partial}\eta(u)(a-u\mathbf{1})^{-1} = \overline{\partial}\big((\chi-\eta)f\big)(u)$. Comme $\int_{\mathbb{C}} \overline{\partial}\big((\chi-\eta)f\big)\, dm = 0$ puisque $\overline{\partial}\big((\chi-\eta)f\big)$ est de classe $\mathcal{C}^1$ à suport compact, on a donc

$$\begin{aligned} \int_{\mathbb{C}} \overline{\partial}\chi(u)(a-u\mathbf{1})^{-1} dm(u) &= \int_{\mathbb{C}} \overline{\partial}\eta(u)(a-u\mathbf{1})^{-1} dm(u) \\ &= \int_{\mathbb{C}\backslash\overline{D}(0,R)} \overline{\partial}\eta(u)(a-u\mathbf{1})^{-1} dm(u). \end{aligned}$$

Maintenant, si $u \in \mathbb{C}\backslash\overline{D}(0,R)$ alors $|u| > \|a\|$, donc on peut écrire

$$(a-u\mathbf{1})^{-1} = -\sum_{k=0}^{\infty} \frac{a^k}{u^{k+1}};$$

et la série converge normalement sur tout compact de $\mathbb{C}\backslash\overline{D}(0,R)$. Comme $\overline{\partial}\eta$ est continue à support compact, on a donc

$$\begin{aligned} \int_{\mathbb{C}\backslash\overline{D}(0,R)} \overline{\partial}\eta(u)(a-u\mathbf{1})^{-1} dm(u) &= -\sum_{k=0}^{\infty} \int_{\mathbb{C}\backslash\overline{D}(0,R)} \overline{\partial}\eta(u)\, \frac{a^k}{u^{k+1}}\, dm(u) \\ &= -\sum_{k=0}^{\infty} a^k \int_{\mathbb{C}} \frac{\overline{\partial}\eta(u)}{u^{k+1}}\, dm(u). \end{aligned}$$

Mais comme $\eta \equiv 1$ sur le disque $D(0,R)$, on a $\int_{\mathbb{C}} \frac{\overline{\partial}\eta(u)}{u-z}\, dm(u) \equiv \pi$ sur $D(0,R)$ d'après la Fait 1, et donc $\int_{\mathbb{C}} \frac{\overline{\partial}\eta(u)}{(u-z)^n}\, dm(u) \equiv 0$ sur $D(0,R)$ pour tout $n \geqslant 2$ en dérivant par rapport à z. En particulier, en prenant $z := 0$, on voit que $\int_{\mathbb{C}} \frac{\overline{\partial}\eta(u)}{u^{k+1}}\, dm(u)$ vaut π pour $k := 0$ et 0 pour tout $k \geqslant 1$. Donc on obtient

$$\frac{1}{\pi}\int_{\mathbb{C}} \overline{\partial}\chi(u)(a-u\mathbf{1})^{-1}dm(u) = -a^0 = -\mathbf{1}.$$

□

Il est maintenant très facile de terminer la preuve du lemme. Si χ est comme dans le Fait 3, alors on a pour toute fraction rationnelle R sans pôles dans $\overline{V}$:

$$\begin{aligned}\|R(a)\| &\leqslant \frac{1}{\pi}\int_{\overline{V}\setminus W} \|\overline{\partial}\chi(u)R(u)(a-u\mathbf{1})^{-1}\|\, dm(u)\\ &\leqslant C\, \sup\{R(u);\ u\in\overline{V}\},\end{aligned}$$

où $C := \frac{1}{\pi}\int_{\overline{V}\setminus W} \|\overline{\partial}\chi(u)(a-u\mathbf{1})^{-1}\|\, dm(u)$. □

Lemme 3.11. *Si Ω est un ouvert de $\mathbb{C}$, alors les fonctions rationnelles sans pôles dans Ω sont denses dans $H(\Omega)$ pour la topologie de la convergence uniforme sur tout compact.*

Démonstration. Il s'agit d'un cas particulier du **Théorème de Runge**, un résultat “bien connu” d'analyse complexe. Même si on a tous les outils pour le démontrer, on *admettra* le résultat. □

Corollaire 3.12. *Pour toute $f \in \mathcal{H}(a)$, il existe une suite $(R_n)_{n\in\mathbb{N}} \subseteq \mathcal{R}(a)$ telle que $R_n \to f$ dans $\mathcal{H}(a)$.*

Démonstration. C'est évident par le lemme. □

Preuve du Théorème 3.9. (i) Pour l'existence du calcul fonctionnel Φ_a, tout va découler du fait suivant.

Fait. Soit $f \in \mathcal{H}(a)$. Si $(R_n)_{n\in\mathbb{N}} \subseteq \mathcal{R}(a)$ est telle que $R_n \to f$ dans $\mathcal{H}(a)$, alors la suite $\big(R_n(a)\big)$ converge dans A. De plus, $\lim R_n(a)$ est indépendante de la suite (R_n) telle que $R_n \to f$.

Preuve du Fait. Par définition de la convergence dans $\mathcal{H}(a)$, il existe un ouvert $U \subseteq \sigma(a)$ tel que f est holomorphe sur U, les R_n n'ont pas de pôle dans U et $R_n(z) \to f(z)$ uniformément sur tout compact de U. Comme $\sigma(a)$ est compact, on peut trouver un ouvert borné V tel que $\sigma(a) \subseteq V$ et $\overline{V} \subseteq U$. Alors la suite (R_n) converge uniformément sur $\overline{V}$, donc $|R_q(z) - R_p(z)| \to 0$ uniformément sur $\overline{V}$ quand $p, q \to \infty$. D'après le Lemme 3.10, on en déduit que $\|R_q(a) - R_p(a)\| \to 0$ quand $p, q \to \infty$. Donc la suite $(R_n(a))$ est de Cauchy dans A, et donc elle converge dans A. Si $(S_n) \subseteq \mathcal{R}(a)$ est une autre suite telle que $S_n \to f$, alors $R_n - S_n \to 0$ dans $\mathcal{H}(a)$, donc $\|R_n(a) - S_n(a)\| \to 0$ grâce au Lemme 3.10, et donc $\lim S_n(a) = \lim R_n(a)$. □

Par le Fait, on peut poser pour toute $f \in \mathcal{H}(a)$:

$$\Phi_a(f) := \lim_{n\to\infty} R_n(a),$$

où (R_n) est n'importe quelle suite d'éléments de $\mathcal{R}(a)$ telle que $R_n \to f$ dans $\mathcal{H}(a)$. Comme l'application $R \mapsto R(a)$ est un homomorphisme de $\mathcal{R}(a)$ dans A, on vérifie très facilement que Φ_a est un homomorphisme de $\mathcal{H}(a)$ dans A. De plus, on a $\Phi_a(R) = R(a)$ pour toute $R \in \mathcal{R}(a)$ (on peut prendre $R_n := R$), donc certainement $\Phi_a(\mathbf{z}) = a$.

Montrons que l'application $\Phi_a : \mathcal{H}(a) \to A$ est séquentiellement continue. Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{H}(a)$ telle que $f_n \to f \in \mathcal{H}(a)$. Par définition de la convergence dans $\mathcal{H}(a)$, on peut trouver un ouvert $U \supseteq \sigma(a)$ tel que f et les f_n sont holomorphes sur U et $f_n(z) \to f(z)$ uniformément sur tout compact de U. Soit V un ouvert borné tel que $\sigma(a) \subseteq V$ et $\overline{V} \subseteq U$. Par le Lemme 3.11 appliqué avec $\Omega := U$, on peut choisir des suites $(R_k)_{k\in\mathbb{N}}$ et $(R_{k,n})_{k\in\mathbb{N}}$ de fractions rationnelles sans pôles dans U telles que $R_k \to f$ dans $H(U)$ et $\forall n \in \mathbb{N}\ :\ R_{k,n} \to f_n$ dans $H(U)$ quand $k \to \infty$. Avec les notations du Lemme 3.10, on a $\|R_{k,n}(a) - R_k(a)\| \leqslant C_V \sup\{|R_{k,n}(z) - R_k(z)|;\ z \in \overline{V}\}$ pour tous $k, n \in \mathbb{N}$. En faisant $k \to \infty$, on en déduit que

$$\forall n \in \mathbb{N}\ :\ \|\Phi_a(f_n) - \Phi_a(f)\| \leqslant C_V \sup\{|f_n(z) - f(z)|;\ z \in \overline{V}\}.$$

Comme $f_n(z) \to f(z)$ uniformément sur $\overline{V}$ puisque $\overline{V}$ est un compact de U, cela montre que $\Phi_a(f_n) \to \Phi_a(f)$.

(ii) L'unicité est facile à démontrer : si Φ est un homomorphisme de $\mathcal{H}(a)$ dans A tel que $\Phi(\mathbf{z}) = a$, on a nécessairement $\Phi(R) = R(a)$ pour toute $R \in \mathcal{R}(a)$; et si Φ est de plus séquentiellement continu, on en déduit que $\Phi = \Phi_a$ par "densité séquentielle" de $\mathcal{R}(a)$ dans $\mathcal{H}(a)$.

(iii) Désormais, on écrit $f(a)$ au lieu de $\Phi_a(f)$. Soit $f \in \mathcal{H}(a)$.

Par définition, $f(a)$ est limite d'éléments de A de la forme $R(a)$, où $R \in \mathcal{R}(a)$; donc $f(a)b = bf(a)$ pour tout $b \in A$ tel que $ab = ba$ (exo). Il reste à montrer que $\sigma(f(a)) = f(\sigma(a))$.

Si $\lambda \in \sigma(a)$, alors la fonction $z \mapsto f(z) - f(\lambda)$ appartient à $\mathcal{H}(a)$ et s'annule en λ. Donc on peut trouver $g \in \mathcal{H}(a)$ telle que $f(z) - f(\lambda) = (z - \lambda)g(z)$ au voisinage de $\sigma(a)$. On a donc $f - f(\lambda)\mathbf{1} = (\mathbf{z} - \lambda\mathbf{1})g = g(\mathbf{z} - \lambda\mathbf{1})$ dans $\mathcal{H}(a)$, et donc $f(a) - f(\lambda)\mathbf{1} = (a - \lambda\mathbf{1})g(a) = g(a)(a - \lambda\mathbf{1})$. Comme $a - \lambda\mathbf{1}$ n'est pas inversible dans A, on en déduit que $f(a) - f(\lambda)\mathbf{1}$ n'est pas inversible dans A, *i.e.* $f(\lambda) \in \sigma(f(a)$. Ainsi, on a montré que $f(\sigma(a)) \subseteq \sigma(f(a))$.

Inversement, soit $\mu \in \sigma(f(a))$. Alors $(f - \mu\mathbf{1})(a) = f(a) - \mu\mathbf{1}$ n'est pas inversible dans A, donc $f - \mu\mathbf{1}$ n'est pas inversible dans $\mathcal{H}(a)$. Donc $f - \mu\mathbf{1}$ doit s'annuler en au moins 1 point de $\sigma(a)$; autrement dit $\mu \in f(\sigma(a))$. □

Remarque 1. Comme l'algèbre $\mathcal{H}(a)$ est commutative, on a $f(a)g(a) = g(a)f(a)$ pour toutes $f, g \in \mathcal{H}(a)$.

Remarque 2. Si $f \in \mathcal{H}(a)$ est holomorphe dans un disque $D(0, R)$ avec $R > r(a)$ et si on écrit $f(z) = \sum_{k=0}^{\infty} c_k a^k$, alors $f(a) = \sum_{k=0}^{\infty} c_k a^k$.

Démonstration. Si on pose $P_n(z) := \sum_{k=0}^{n} c_k z^k$, alors $P_n(z) \to f(z)$ uniformément sur tout compact de $D(0, R)$, donc $P_n \to f$ dans $\mathcal{H}(a)$, et donc $P_n(a) \to f(a)$. Comme $P_n(a) = \sum_{k=0}^{\infty} c_k a^k$ pour tout $n \in \mathbb{N}$, on a donc $f(a) = \sum_{k=0}^{\infty} c_k a^k$. □

Remarque 3. Si $f \in \mathcal{H}(a)$, on peut donner une formule pour $f(a)$: on a

$$f(a) = -\frac{1}{\pi} \int_{\mathbb{C}} \overline{\partial}\chi(u) f(u)(a - u\mathbf{1})^{-1}\, dm(u),$$

où χ est n'importe quelle fonction de classe $\mathcal{C}^1$ à support compact telle que $\chi(z) \equiv 1$ au voisinage de $\sigma(a)$.

Démonstration. Cela a été démontré dans la preuve du Lemme 3.10 lorsque f est une fonction rationnelle (Fait 3) ; et le cas général s'en déduit par approximation. □

Corollaire 3.13. (Théorème de décomposition de Riesz)
Soit X un espace de Banach complexe, et soit $T \in \mathcal{L}(X)$. On suppose que $\sigma(T)$ peut se décomposer sous la forme $\sigma(T) = \sigma_1 \cup \sigma_2$, où σ_1 et σ_2 sont des compacts non vides tels que $\sigma_1 \cap \sigma_2 = \varnothing$. Alors on peut décomposer l'espace X sous la forme $X = X_1 \oplus X_2$, où X_1 et X_2 sont des sous-espaces fermés invariants par T tels que $\sigma\big(T_{|X_1}\big) = \sigma_1$ et $\sigma\big(T_{|X_2}\big) = \sigma_2$. De plus, cette décomposition est unique.

Démonstration. Soient V_1 et V_2 des ouverts de $\mathbb{C}$ tels que $\sigma_i \subseteq V_i$ et $V_1 \cap V_2 = \varnothing$, et posons $\Omega := V_1 \cup V_2$. Alors les fonctions $\mathbf{1}_{V_1}$ et $\mathbf{1}_{V_2}$ sont holomorphes sur Ω (!) donc appartiennent à $\mathcal{H}(T)$ puisque $\Omega \supseteq \sigma(T)$. Posons $p_i := \mathbf{1}_{V_i}(T)$ pour $i = 1, 2$. Comme $\mathbf{1}_{V_i}^2 = \mathbf{1}_{V_i}$, on a $p_i^2 = p_i$, autrement dit p_1 et p_2 sont des projections. De plus, on a $p_1 + p_2 = I$ car $\mathbf{1}_{V_1} + \mathbf{1}_{V_2} = \mathbf{1}$, et $p_1p_2 = 0 = p_2p_1$ car $\mathbf{1}_{V_1}\mathbf{1}_{V_2} = 0$. Donc, si on pose $X_i := \mathrm{Im}(p_i) = \ker(I - p_i)$, alors $X = X_1 \oplus X_2$; et les sous-espaces X_i sont invariants par T car $p_iT = Tp_i$.
Posons $T_i := T_{|X_i} : X_i \to X_i$, et montrons que $\sigma(T_i) = \sigma_i$ pour $i = 1, 2$. Comme $T = T_1 \oplus T_2$, on a $\sigma(T) = \sigma(T_1) \cup \sigma(T_2)$ (exo). Comme par ailleurs $\sigma(T) = \sigma_1 \cup \sigma_2$ et $\sigma_1 \cap \sigma_2 = \varnothing$, il suffit donc de montrer qu'on a $\sigma(T_i) \subseteq \sigma_i$ pour $i = 1, 2$. On peut évidemment se contenter de montrer que $\sigma(T_1) \subseteq \sigma_1$. Soit $\lambda \in \mathbb{C}\backslash\sigma_1$. Comme $z - \lambda$ ne s'annule pas au voisinage de σ_1, il existe une fonction g holomorphe au voisinage de $\sigma(T)$ telle que $(z - \lambda)g(z) \equiv 1$ au voisinage de σ_1 et $g(z) \equiv 0$ sur V_2. Alors $(\mathbf{z} - \lambda\mathbf{1})g = \mathbf{1}_{V_1}$ dans $\mathcal{H}(T)$, donc $(T - \lambda I)g(T) = p_1 = g(T)(T - \lambda I)$. Comme $(p_1)_{|X_1} = I_{X_1}$, on en déduit que si on pose $R := g(T)$, alors $T_1 - \lambda I_{X_1}$ est inversible d'inverse $R_{|X_1}$. Donc $\lambda \notin \sigma(T_1)$; et ainsi, on a montré que $\sigma(T_1) \subseteq \sigma_1$.
Supposons qu'on ait une autre décomposition $X = X_1' \oplus X_2'$ avec $\sigma(T_{X_i'}) = \sigma_i$. Notons p_1' et p_2' les projections associées à cette décomposition. Comme X_1' et X_2' sont invariants par T, les projections p_i' commutent avec T. Donc $p_i'p_j = p_jp_i'$ pour tous $i, j \in \{1, 2\}$ car p_1 et p_2 ont été obtenues par calcul fonctionnel holomorphe. On en déduit (exo) qu'on a $X_1 = (X_1 \cap X_1') \oplus (X_1 \cap X_2')$ et $X_2' = (X_1 \cap X_2') \oplus (X_2 \cap X_2')$, tous les sous-espaces considérés étant invariants par T. Par conséquent, si on suppose que $X_1 \cap X_2' \neq \{0\}$, on a d'une part $\sigma(T_{|X_1 \cap X_2'}) \subseteq \sigma(T_{|X_1}) = \sigma_1$, et d'autre part $\sigma(T_{|X_1 \cap X_2'}) \subseteq \sigma(T_{|X_2'}) = \sigma_2$; donc $\sigma(T_{|X_1 \cap X_2'}) = \varnothing$, ce qui est absurde. Ainsi, on a $X_1 \cap X_2' = \{0\}$; donc $X_1 = X_1 \cap X_1'$, *i.e.* $X_1' \subseteq X_1$. L'inclusion inverse se démontre de la même façon ; donc $X_1 = X_1'$, et de même $X_2 = X_2'$. □

Théorème 3.14. *Si $f \in \mathcal{H}(a)$ et $g \in \mathcal{H}(f(a))$, alors $g \circ f \in \mathcal{H}(a)$ et $(g \circ f)(a) = g\big(f(a)\big)$.*

Démonstration. La fonction f est holomorphe sur un ouvert $U \supseteq \sigma(a)$, et g est holomorphe sur un ouvert $V \supseteq \sigma(f(a))$. Comme $\sigma(f(a)) = f(\sigma(a))$, on a $f(\sigma(a)) \subseteq V$. Donc $\Omega := f^{-1}(V)$ est un ouvert contenant $\sigma(a)$, sur lequel $g \circ f$ est bien définie et holomorphe ; et donc $g \circ f \in \mathcal{H}(a)$. Si on pose $\Phi(g) := (g \circ f)(a)$, alors Φ est visiblement un homomorphisme de $\mathcal{H}(f(a))$ dans A tel que $\Phi(\mathbf{z}) = f(a)$. De plus, on vérifie sans difficulté que Φ est séquentiellement continu. Donc $\Phi(a) = g(f(a))$ pour toute $g \in \mathcal{H}(f(a))$. □

Corollaire 3.15. *Si $a \in A$ vérifie $r(a - \mathbf{1}) < 1$, en particulier si $\|a - \mathbf{1}\| < 1$, alors a admet un logarithme : il existe $u \in A$ tel que $e^u = a$.*

Démonstration. Par hypothèse, on a $\sigma(a - \mathbf{1}) \subseteq D(0, 1)$. Comme $\sigma(a - \mathbf{1}) = \sigma(a) - 1$, on en déduit $\sigma(a) \subseteq D(1, 1) \subseteq \mathbb{C}\backslash\mathbb{R}^-$. Soit log la détermination principale du logarithme complexe. La fonction log est holomorphe sur $\mathbb{C}\backslash\mathbb{R}^-$, donc $\log \in \mathcal{H}(a)$. On peut donc définir $u := \log(a)$. Comme $e^{\log(z)} = z$ pour tout $z \in \mathbb{C}\backslash\mathbb{R}^-$, on a $\exp \circ \log = \mathbf{z}$ dans $\mathcal{H}(a)$, donc $e^u = e^{\log(a)} = \exp \circ \log(a) = a$ par le théorème. $\square$

3.2. Calcul fonctionnel continu. Dans cette section, H est un espace de Hilbert complexe.

3.2.1. *Opérateurs auto-adjoints.*

Théorème 3.16. *Soit $T \in \mathcal{L}(H)$ un opérateur* auto-adjoint. *Il existe un unique homomorphisme continu $\Psi_T : \mathcal{C}(\sigma(T)) \to \mathcal{L}(H)$ tel que $\Psi_T(\mathbf{x}) = T$, où $\mathbf{x} \in \mathcal{C}(\sigma(T))$ est la fonction $x \mapsto x$. Si on pose $\Psi_T(f) := f(T)$, alors les propriétés suivantes ont lieu pour toute $f \in \mathcal{C}(\sigma(T))$.*

(i) *$f(T)$ est limite d'opérateurs de la forme $P(T)$ où P est un polynôme.*
(ii) *$\|f(T)\| = \|f\|_\infty$.*
(iii) *$f(T)^* = \bar{f}(T)$.*
(iv) *$f(T)A = Af(T)$ pour tout $A \in \mathcal{L}(H)$ tel que $TA = AT$.*
(v) *$\sigma(f(T)) = f(\sigma(T))$.*

On dit que Ψ_T est le ***calcul fonctionnel continu*** *associé à l'opérateur T.*

Démonstration. Le point clé est le fait suivant.

Fait. Pour tout polynôme $P \in \mathbb{C}[\mathbf{x}]$, on a $\sigma(P(T)) = P(\sigma(T))$ et $\|P(T)\| = \sup\big\{|P(x)|;\ x \in \sigma(T)\big\}$.

Preuve du Fait 1. On sait depuis longtemps que $\sigma(P(T)) = P(\sigma(T)$.
Comme T est auto-adjoint (il suffirait que T soit normal), $P(T)$ est un opérateur normal ; donc $\|P(T)\| = r(P(T))$, autrement dit $\|P(T)\| = \sup\big\{|\lambda|;\ \lambda \in \sigma(P(T))\big\}$; d'où le résultat puisque $\sigma(P(T)) = P(\sigma(T)) = \big\{P(x);\ x \in \sigma(T)\big\}$. $\square$

Notons $\mathcal{P} \subseteq \mathcal{C}(\sigma(T))$ l'ensemble des fonctions polynomiales. Par le Fait, si $p \in \mathcal{P}$ et si P_1 et P_2 sont deux polynômes tels que $P_1(x) \equiv p(x) \equiv P_2(x)$ sur $\sigma(T)$, alors $P_1(T) - P_2(T) = (P_1 - P_2)(T) = 0$ puisque $(P_1 - P_2)(x) \equiv 0$ sur $\sigma(T)$; donc $P_1(T) = P_2(T)$. On peut donc poser $\psi_T(p) := P(T)$, où P est n'importe quel polynôme tel que $P(x) \equiv p(x)$ sur $\sigma(T)$. L'application $\psi_T : \mathcal{P} \to \mathcal{L}(H)$ ainsi définie est un homomorphisme de $\mathcal{P}$ dans $\mathcal{L}(H)$, et ψ_T est de plus une *isométrie* par le Fait : on a $\|\psi_T(p)\| = \|p\|_\infty$ pour toute $p \in \mathcal{P}$.
Comme T est auto-adjoint, $\sigma(T)$ est un compact de $\mathbb{R}$; donc $\mathcal{P}$ est dense dans $\mathcal{C}(\sigma(T))$ d'après le Théorème de Weierstrass. On peut donc prolonger l'application linéaire continue ψ_T en une application linéaire continue $\Psi_T : \mathcal{C}(\sigma(T)) \to \mathcal{L}(H)$; et Ψ_T est un homomorphisme d'algèbres isométrique car ψ_T en est un (exo). Ainsi, en posant $f(T) := \Psi_T(f)$, on a un homomorphisme continu $f \mapsto f(T)$ tel que $\|f(T)\| = \|f\|_\infty$ pour toute $f \in \mathcal{C}(\sigma(T))$.
Si $\Psi : \mathcal{C}(\sigma(T)) \to \mathcal{L}(H)$ est un autre homomorphisme tel que $\Psi(\mathbf{x}) = T$, alors $\Psi(p) = p(T) = \Psi_T(p)$ pour toute fonction polynomiale p ; et si Ψ est de plus continu, alors $\Psi = \Psi_T$ par densité de $\mathcal{P}$ dans $\mathcal{C}(\sigma(T))$.
La propriété (i) est vraie par définition de Ψ_T : si $f \in \mathcal{C}(\sigma(T))$, on peut trouver une suite de polynômes (P_n) telle que $P_n(x) \to f(x)$ uniformément sur $\sigma(T)$, et alors $P_n(T) = \Psi_T(P_n) \to \Psi_T(f) = f(T)$ par continuité de Ψ_T. La propriété (ii) a déjà été vue. Pour (iii), on observe que si $p = a_0\mathbf{1} + a_1\mathbf{x} + \cdots + a_d\mathbf{x}^d \in \mathcal{P}$, alors $p(T)^* =$

$\overline{a_0}\, I + \overline{a_1}\, T + \cdots + \overline{a_d}\, T^d = \overline{p}(T)$ car $T = T^*$ et (donc) $\mathbf{x}$ est à valeurs réelles sur $\sigma(T)$; donc (iii) découle de (i) par continuité de Ψ_T. La propriété (iv) découle également de (i), car si $AT = TA$, alors $A^k T = TA^k$ pour tout $k \in \mathbb{N}$ et donc $AP(T) = P(T)A$ pour tout polynôme P.

Il reste à montrer qu'on a $\sigma\big(f(T)\big) = f(\sigma(T))$ pour toute $f \in \mathcal{C}(\sigma(T))$.

Comme Ψ_T est un homomorphisme de $\mathcal{C}(\mathbb{T})$ dans $\mathcal{L}(H)$, on a

$$\sigma(f(T)) = \sigma_{\mathcal{L}(H)}(\Psi_T(f)) \subseteq \sigma_{\mathcal{C}(\sigma(T))}(f) = f(\sigma(T)).$$

Inversement, soit $\lambda \in f(\sigma(T))$, et choisissons $x_0 \in \sigma(T)$ tel que $f(x_0) = \lambda$. Soit (p_n) une suite de fonctions polynomiales telle que $p_n \to f$ dans $\mathcal{C}(\sigma(T))$. Alors $p_n(x_0) \in p_n(\sigma(T)) = \sigma(p_n(T))$ pour tout $n \in \mathbb{N}$, autrement dit $p_n(T) - p_n(x_0)\mathbf{1}$ n'est pas inversible dans $\mathcal{L}(H)$. Comme $p_n(T) - p_n(x_0)\mathbf{1} \to f(T) - \lambda\mathbf{1}$ par continuité du calcul fonctionnel et comme l'ensemble des opérateurs non-inversibles est fermé dans $\mathcal{L}(H)$, on en déduit que $f(T) - \lambda I$ est non-inversible, *i.e.* $\lambda \in \sigma(f(T))$. □

Remarque. Le calcul fonctionnel continu est compatible avec le calcul fonctionnel holomorphe : si $f \in \mathcal{H}(T)$, alors $f(T) = \big(f_{|\sigma(T)}\big)(T)$.

Démonstration. L'application $\Phi : f \mapsto \big(f_{|\sigma(T)}\big)(T) = \Psi_T(f_{|\sigma(T)})$ est un homomorphisme de $\mathcal{H}(T)$ dans $\mathcal{L}(H)$ tel que $\Phi(\mathbf{z}) = T$; et Φ est séquentiellement continu car l'aplication $f \mapsto f_{|\sigma(T)}$ est séquentiellement continue et Ψ_T est continu. Donc Φ est le calcul fonctionnel holomorphe Φ_T. □

Théorème 3.17. *Soit $T \in \mathcal{L}(H)$ auto-adjoint. Si $f \in \mathcal{C}(\sigma(T))$ est à valeurs réelles et si $g \in \mathcal{C}(f(\sigma(T)))$, alors $(g \circ f)(T) = g(f(T))$.*

Démonstration. La fonction $g \circ f$ est bien définie et continue sur $\sigma(T)$, donc l'opérateur $(g \circ f)(T)$ est bien défini. De même, $g(f(T))$ a un sens car l'opérateur $f(T)$ est auto-adjoint (f est à valeurs réelles) et g est continue sur $\sigma(f(T)) = f(\sigma(T))$. Comme l'application $g \mapsto g(f(T))$ est un homomorphisme de $\mathcal{C}(\sigma(f(T))$ dans $\mathcal{L}(H)$, on a $P(f(T)) = (P \circ f)(T)$ pour tout polynôme P ; donc $g(f(T)) = (g \circ f)(T)$ pour toute $g \in \mathcal{C}(\sigma(f(T))$ par approximation. □

Comme illustration du calcul fonctionnel, on peut maintenant redémontrer très rapidement l'existence et l'unicité de la racine carrée d'un opérateur positif, qu'on avait établie "à la main" au Chapitre 1.

Corollaire 3.18. *Si $T \in \mathcal{L}(H)$ est un opérateur positif, alors il existe un unique opérateur positif S tel que $S^2 = T$. De plus, on a $SA = AS$ pour tout $A \in \mathcal{L}(H)$ tel que $TA = AT$. On dit que S est la* ***racine carrée*** *de T, et on écrit $S = \sqrt{T}$ ou $S = T^{1/2}$.*

Démonstration. Comme T est positif, il est auto-adjoint avec $\sigma(T) \subseteq \mathbb{R}^+$. Donc la fonction $f(x) := \sqrt{x}$ est bien définie et continue sur $\sigma(T)$, et on peut poser $S := f(T)$. Alors $S^* = \overline{f}(T) = f(T) = S$ puisque f est à valeurs réelles, S est positif car $\sigma(S) = f(\sigma(T)) \subseteq \mathbb{R}^+$, et $S^2 = f^2(T) = \mathbf{x}(T) = T$. Donc S convient. De plus, on a $AS = AS$ pour tout $A \in \mathcal{L}(H)$ tel que $AT = TA$ puisque S est de la forme $f(T)$.

Pour l'unicité, on utilise le théorème de composition : si S' est un opérateur positif tel que $S'^2 = T$ alors $S' = \mathbf{x}(S') = \sqrt{\mathbf{x}^2}(S') = \sqrt{\mathbf{x}^2(S')} = \sqrt{S'^2} = \sqrt{T} = S$. □

Voici une autre illustration du calcul fonctionnel.

Corollaire 3.19. *Un opérateur* $T \in \mathcal{L}(H)$ *est* positif et inversible *si et seulement si il existe un opérateur auto-adjoint* L *tel que* $T = e^L$*. Dans ce cas, l'opérateur* L *est unique.*

Démonstration. Si $T = e^L$ avec $L^* = L$, alors T est auto-adjoint car la fonction $\exp_{|\sigma(L)} \in \mathcal{C}(\sigma(L))$ est à valeurs réelles, et $\sigma(T) = \exp(\sigma(L)) \subseteq\,]0, \infty[$. Donc T est positif et inversible.
Inversement, si T est positif et inversible, alors $\sigma(T) \subseteq]0, \infty[$, donc la fonction ln est (bien définie et) continue sur $\sigma(T)$. On peut donc poser $L := \ln(T)$. Alors L est auto-adjoint car ln est à valeurs réelles, et $e^L = \exp(\ln(T)) = (\exp \circ \ln)(T) = \mathbf{x}(T) = T$. Si L' est un autre opérateur auto-adjoint tel que $e^{L'} = T$, alors $L' = (\ln \circ \exp)(L') = \ln(e^{L'}) = \ln(T) = L$. □

Exercice. Soit K un compact de $\mathbb{R}$, et soit $f : K \to \mathbb{R}$ une fonction continue et *injective*. Montrer que si $T_1, T_2 \in \mathcal{L}(H)$ sont des opérateurs auto-adjoints tels que $\sigma(T_1) = K = \sigma(T_2)$ et $f(T_1) = f(T_2)$, alors $T_1 = T_2$.

Remarque 3.20. Si T est un opérateur auto-adjoint sur un espace de Hilbert *réel*, alors on peut définir $f(T)$ pour toute fonction $f \in \mathcal{C}(\sigma(T))$ à valeurs réelles, et les propriétés du calcul fonctionnel restent les mêmes.

Démonstration. Il suffit de montrer que pour tout polynôme P à coefficients réels, on a $\sigma(P(T)) = P(\sigma(T))$: si ce point est acquis, on peut recopier la preuve du Théorème 3.16 en utilisant la Remarque 1.28 et la densité des fonctions polynomiales à coefficients réels dans $\mathcal{C}_{\mathbb{R}}(\sigma(T))$.
On a $P(\sigma(T)) \subseteq \sigma(P(T))$ pour tout $P \in \mathbb{R}[\mathbf{x}]$ d'après la Proposition 1.13.
Pour la réciproque, il suffit de montrer que si $P \in \mathbb{R}[\mathbf{x}]$ ne s'annule pas sur $\sigma(T)$, alors $P(T)$ est inversible dans $\mathcal{L}(H)$.
Le polynôme P se décompose sous la forme $P = (\mathbf{x} - \alpha_1) \cdots (\mathbf{x} - \alpha_r)Q$, où $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ et $Q \in \mathbb{R}[\mathbf{x}]$ n'a pas de racines réelles ; donc $P(T) = Q(T)(T - \alpha_1 I) \cdots (T - \alpha_r I)$. Comme P ne s'annule pas sur $\sigma(T)$, les α_i n'appartiennent pas à $\sigma(T)$ et donc l'opérateur $(T - \alpha_1 I) \cdots (T - \alpha_r I)$ est inversible. Il suffit donc de montrer que l'opérateur $Q(T)$ est inversible.
Comme Q n'a pas de racines réelles, il est par exemple > 0 sur $\mathbb{R}$; et comme c'est un polynôme, $Q(x)$ tend vers $+\infty$ quand $x \to \pm\infty$. Donc on peut trouver $\varepsilon > 0$ tel que $Q(x) > \varepsilon$ pour tout $x \in \mathbb{R}$. Si on pose $R := Q - \varepsilon \mathbf{1}$, alors $R \in \mathbb{R}[\mathbf{x}]$ et $R(x) > 0$ pour tout $x \in \mathbb{R}$. En décomposant R dans $\mathbb{C}[\mathbf{z}]$, on voit qu'on peut trouver des polynômes A et B à coefficients réels tels que $R = A^2 + B^2$. Alors $R(T) = A(T)^2 + B(T)^2 = A(T)^* A(T) + B(T)^* B(T)$ car T est auto-adjoint ; donc $R(T)$ est un opérateur positif. Comme $R(T) = Q(T) - \varepsilon I$, on a donc $\langle Q(T)u, u \rangle \geqslant \varepsilon \|u\|^2$ pour tout $u \in H$. Donc $Q(T)$ est un plongement, et donc $Q(T)$ est inversible car $Q(T)$ est auto-adjoint. □

3.2.2. *Opérateurs normaux.* On voudrait maintenant montrer l'existence d'un calcul fonctionnel continu pour tout opérateur *normal* $T \in \mathcal{L}(H)$. Les idées sont les mêmes que dans le cas auto-adjoint, mais la preuve est plus délicate.

Notation. On note $\mathbb{C}[X, Y]$ l'algèbre des polynômes à 2 indéterminées à coefficients complexes. Si $P = \sum c_{k,l} X^k Y^k \in \mathbb{C}[X, Y]$, si A est une algèbre de Banach avec unité et si $a, b \in A$, on pose

$$P(a, b) := \sum c_{k,l} a^k b^l.$$

Remarque. Si $ab = ba$, alors l'application $P \mapsto P(a,b)$ est un homomorphisme de $\mathbb{C}[X,Y]$ dans A.

Théorème 3.21. *Soit $T \in \mathcal{L}(H)$ un opérateur* normal. *Il existe un unique homomorphisme continu $\Psi_T : \mathcal{C}(\sigma(T)) \to \mathcal{L}(H)$ tel que $\Psi_T(\mathbf{z}) = T$ et $\Psi_T(\bar{\mathbf{z}}) = T^*$. Si on pose $f(T) := \Psi_T(f)$, alors les propriétés suivantes ont lieu pour toute $f \in \mathcal{C}(\sigma(T))$.*

(i) *$f(T)$ est limite d'opérateurs de la forme $P(T,T^*)$ où $P \in \mathbb{C}[X,Y]$.*

(ii) *$\|f(T)\| = \|f\|_\infty$.*

(iii) *$f(T)^* = \bar{f}(T)$.*

(iv) *$f(T)A = Af(T)$ pour tout $A \in \mathcal{L}(H)$ tel que $TA = AT$ et $T^*A = AT^*$.*

(v) *$\sigma(f(T)) = f(\sigma(T))$.*

Remarque. L'hypothèse faite sur A dans (iv) est en fait redondante : on peut montrer que si $AT = TA$, alors $AT^* = T^*A$ (*Théorème de Fuglede-Putnam*).

Pour la preuve du Théorème 3.21, le point clé est le fait suivant.

Fait 3.22. *Soit $T \in \mathcal{L}(H)$ un opérateur normal. Pour tout polynôme $P \in \mathbb{C}[X,Y]$, on a $\sigma\big(P(T,T^*)\big) = \big\{P(z,\bar{z});\ z \in \sigma(T)\big\}$ et $\|P(T,T^*)\| = \sup\big\{|P(z,\bar{z})|;\ z \in \sigma(T)\big\}$.*

Ce résultat n'est absolument pas évident. Pour le montrer, on a besoin de 4 lemmes intéressants pour eux même. Dans ce qui suit, on dira qu'une sous-algèbre $\mathcal{B} \subseteq \mathcal{L}(H)$ est *auto-adjointe* si elle est stable par passage à l'adjoint : $T \in \mathcal{B} \implies T^* \in \mathcal{B}$.

Lemme 3.23. *Si $\mathcal{B} \subseteq \mathcal{L}(H)$ est une sous-algèbre fermée et auto-adjointe contenant I, alors $\forall S \in \mathcal{B}\ :\ \sigma_{\mathcal{L}(H)}(S) = \sigma_{\mathcal{B}}(S)$.*

Démonstration. Il faut seulement prouver que $\forall S \in \mathcal{B}\ :\ \sigma_{\mathcal{B}}(S) \subseteq \sigma_{\mathcal{L}(H)}(S)$; ce qui revient à montrer que si $R \in \mathcal{B}$ est inversible dans $\mathcal{L}(H)$, alors R est inversible dans $\mathcal{B}$, *i.e.* $R^{-1} \in \mathcal{B}$.
Comme R est inversible dans $\mathcal{L}(H)$, l'opérateur R^*R est également inversible dans $\mathcal{L}(H)$, donc $0 \notin \sigma_{\mathcal{L}(H)}(R^*R)$. Mais R^*R est auto-adjoint, donc $\sigma_{\mathcal{L}(H)} \subseteq \mathbb{R}$. En particulier, $\mathbb{C}\backslash\sigma_{\mathcal{L}(H)}(R^*R)$ est *connexe* car $\mathbb{C}\backslash K$ est connexe pour tout compact $K \subseteq \mathbb{R}$. D'après la Proposition 1.19, on en déduit que $\sigma_{\mathcal{L}(H)}(R^*R) = \sigma_{\mathcal{B}}(R^*R)$; et donc $0 \notin \sigma_{\mathcal{B}}(R^*R)$. Ainsi R^*R est inversible dans $\mathcal{B}$, *i.e.* $(R^*R)^{-1} \in \mathcal{B}$. Mais $R^{-1} = R^{-1}(R^*)^{-1}R^* = (R^*R)^{-1}R^*$, donc $R^{-1} \in \mathcal{B}$ puisque $\mathcal{B}$ est une algèbre auto-adjointe. □

Notation. Si A est une $\mathbb{C}$-algèbre avec unité, on note $\mathrm{Hom}(A,\mathbb{C})$ l'ensemble de tous les homorphismes d'algèbres $\gamma : A \to \mathbb{C}$.

Lemme 3.24. *Soit A une algèbre de Banach commutative avec unité. Si $a \in A$, on a l'équivalence suivante :*

$$a \text{ inversible dans } A \iff \gamma(a) \neq 0 \text{ pour tout } \gamma \in \mathrm{Hom}(A,\mathbb{C}).$$

Démonstration. Si a est inversible, alors $1 = \gamma(\mathbf{1}) = \gamma(a)\gamma(a^{-1})$ pour tout homomorphisme $\gamma : A \to \mathbb{C}$, et donc $\gamma(a) \neq 0$. (La commutativité n'est pas nécessaire ici.)
Inversement, supposons a non inversible dans A. Alors $I_a := \{ax;\ x \in A\}$ est un *idéal propre* de l'anneau commutatif A, *i.e.* I_a est un idéal et $I_a \neq A$. Donc I_a est contenu

dans un idéal propre maximal I. L'idéal I est nécessairement *fermé* dans A : en effet, si I n'était pas fermé, on aurait $\overline{I} = A$ par maximalité de I car $\overline{I}$ est un idéal de A ; donc on pourrait trouver $x \in I$ tel que $\|x - \mathbf{1}\| < 1$, ce qui est impossible car un tel x serait inversible dans A et on aurait alors $\mathbf{1} = xx^{-1} \in I$.
Comme I est un idéal propre de A et que A est une algèbre unitaire, I est un sous-espace vectoriel de A et l'espace vectoriel quotient A/I est une algèbre unitaire ; et comme I est fermé, on peut munir A/I de la norme quotient. On vérifie sans difficulté majeure (exo) que A/I est une algèbre de Banach. De plus, A/I est un *corps* car I est un idéal maximal ; autrement dit, tout $x \in A/I$ non nul est inversible dans A/I. Mais si $x \in A/I$, on sait que $\sigma_{A/I}(x) \neq \emptyset$; donc on peut trouver $\lambda \in \mathbb{C}$ tel que $x - \lambda\mathbf{1}_{A/I}$ n'est pas inversible dans A/I, et on a alors $x - \lambda\mathbf{1}_{A/I} = 0$ puisque A/I est un corps. Ainsi, on voit que $A/I = \mathbb{C}\mathbf{1}_{A/I}$, et donc l'algèbre A/I isomorphe à $\mathbb{C}$. On peut donc considérer l'application quotient $\pi : A \to A/I$ comme un homomorphisme $\gamma : A \to \mathbb{C}$; et on a $\gamma(a) = 0$ puisque $a \in I$. □

Corollaire 3.25. *Soit A une algèbre de Banach commutative avec unité, et soit $a \in A$. Pour tout $a \in A$, on a*

$$\sigma(a) = \big\{\gamma(a);\ \gamma \in \mathrm{Hom}(A, \mathbb{C})\big\}.$$

Démonstration. Si $\gamma \in \mathrm{Hom}(A,\mathbb{C})$ et $\lambda \in \mathbb{C}$, alors $\gamma(\mathbf{1}) = 1$ et donc $\gamma(a) - \lambda = \gamma(a - \lambda\mathbf{1})$. Donc, par le lemme, un nombre complexe λ appartient à $\sigma(a)$ si et seulement si il existe un $\gamma \in \mathrm{Hom}(A,\mathbb{C})$ tel que $\gamma(a) - \lambda = 0$. □

Exercice. Soit A une algèbre de Banach commutative avec unité. Montrer que pour tous $a, b \in A$, on a $\sigma(a + b) \subseteq \sigma(a) + \sigma(b)$ et $\sigma(ab) \subseteq \sigma(a)\sigma(b)$.

Lemme 3.26. *Soit A une algèbre de Banach avec unité. Si $\gamma \in \mathrm{Hom}(A,\mathbb{C})$, alors γ est (une forme linéaire) continue et $\|\gamma\| \leqslant 1$.*

Démonstration. Si $x \in A$, alors $\gamma(x) \in \sigma(x)$ car γ est un homomorphisme, donc $|\gamma(x)| \leqslant r(x) \leqslant \|x\|$. □

Lemme 3.27. *Soit $\mathcal{B} \subseteq \mathcal{L}(H)$ une sous-algèbre fermée auto-adjointe et commutative contenant I. Si $\gamma \in \mathrm{Hom}(A,\mathbb{C})$, alors*

$$\forall S \in \mathcal{B}\ :\ \gamma(S^*) = \overline{\gamma(S)}.$$

Démonstration. Si $S \in \mathcal{B}$, on peut écrire $S = A + iB$ où $A := \frac{S+S^*}{2}$ et $B := \frac{S-S^*}{2i}$ sont auto-adjoints et appartiennent tous les deux à $\mathcal{B}$. Alors $\gamma(S) = \gamma(A) + i\gamma(B)$ et $\gamma(S^*) = \gamma(A - iB) = \gamma(A) - i\gamma(B)$; donc il suffit de montrer que $\gamma(A) \in \mathbb{R}$ et $\gamma(B) \in \mathbb{R}$. Ainsi, on s'est ramené à montrer que $\gamma(S) \in \mathbb{R}$ pour tout $S \in \mathcal{B}$ *auto-adjoint*.
Comme S est auto-ajoint, l'opérateur $e^{iS} = \sum_{k=0}^{\infty} \frac{(iS)^k}{k!}$ est *unitaire*, car e^{iS} est inversible et $(e^{iS})^* = \sum_{k=0}^{\infty} \frac{(-iS)^k}{k!} = e^{-iS} = (e^{iS})^{-1}$. De plus, $e^{iS} \in \mathcal{B}$ car $\mathcal{B}$ est une sous-algèbre fermée de $\mathcal{L}(H)$; donc on peut considérer $\gamma(e^{iS})$. Par le Lemme 3.26, on a $|\gamma(e^{iS})| \leqslant \|e^{iS}\| = 1$. Mais γ est un homomorphisme continu, donc $\gamma(e^{iS}) = \sum_{k=0}^{\infty} \frac{i^k\gamma(S)^k}{k!} = e^{i\gamma(S)}$. On a donc $|e^{i\gamma(S)}| \leqslant 1$; et de même $|e^{-i\gamma(S)}| \leqslant 1$, *i.e.* $|e^{i\gamma(S)}| \geqslant 1$. On voit ainsi que $|e^{i\gamma(S)}| = 1$, ce qui signifie que $\gamma(S) \in \mathbb{R}$. □

Preuve du Fait 3.22. Soit $T \in \mathcal{L}(H)$ un opérateur normal. Notons $\mathcal{C}^*(T) \subseteq \mathcal{L}(H)$ la sous-algèbre fermée de $\mathcal{L}(H)$ engendrée par I, T et T^*. Comme $T^*T = TT^*$, on a

$$\mathcal{C}^*(T) = \overline{\{P(T,T^*);\; P \in \mathbb{C}[X,Y]\}}.$$

En particulier, $\mathcal{C}^*(T)$ est commutative et auto-adjointe. Par les Lemmes 3.23 et 3.24, on en déduit que

$$\forall S \in \mathcal{C}^*(T) \;:\; \sigma(S) = \sigma_{C^*(T)}(S) = \Big\{\gamma(S);\; \gamma \in \mathrm{Hom}\big(\mathcal{C}^*(T), \mathbb{C}\big)\Big\}.$$

Maintenant, soit $P = \sum c_{k,l} X^k Y^l \in \mathbb{C}[X,Y]$. Si $\gamma \in \mathrm{Hom}\big(\mathcal{C}^*(T), \mathbb{C}\big)$ alors, par le Lemme 3.27 :

$$\begin{aligned}
\gamma\big(P(T,T^*)\big) &= \gamma\left(\sum c_{k,l} T^k T^{*l}\right)\\
&= \sum c_{k,l}\gamma(T)^k \gamma(T^*)^l\\
&= \sum c_{k,l}\gamma(T)^k \overline{\gamma(T)}^l\\
&= P\big(\gamma(T), \overline{\gamma(T)}\big).
\end{aligned}$$

En utilisant à nouveau les Lemme 3.23 et 3.24, on obtient donc

$$\begin{aligned}
\sigma\big(P(T,T^*)\big) &= \Big\{P\big(\gamma(T), \overline{\gamma(T)}\,\big);\; \gamma \in \mathrm{Hom}\big(\mathcal{C}^*(T), \mathbb{C}\big)\Big\}\\
&= \big\{P(z,\bar z);\; z \in \sigma_{\mathcal{C}^*(T)}(T)\big\}\\
&= \big\{P(z,\bar z);\; z \in \sigma(T)\big\}.
\end{aligned}$$

Enfin, l'opérateur $P(T,T^*)$ est normal car T est normal (micro-exo), donc $\|P(T,T^*)\| = r\big(P(T,T^*)\big)$; autrement dit

$$\|P(T,T^*)\| = \sup\big\{|\lambda|;\; \lambda \in \sigma\big(P(T,T^*)\big)\big\} = \sup\big\{|P(z,\bar z)|;\; z \in \sigma(T)\big\}.$$

□

Preuve du Théorème 3.21. Notons $\mathcal{P} \subseteq \mathcal{C}(\sigma(T))$ l'ensemble des fonctions f de la forme $f(z) = P(z,\bar z)$, où $P \in \mathbb{C}[X,Y]$. Par le Fait 3.22, si $f \in \mathcal{P}$ et si $P_1, P_2 \in \mathbb{C}[X,Y]$ sont tels que $P_1(z,\bar z) \equiv f(z) \equiv P_2(z,\bar z)$ sur $\sigma(T)$, alors $P_1(T,T^*) - P_2(T,T^*) = (P_1-P_2)(T,T^*) = 0$ puisque $(P_1-P_2)(z,\bar z) \equiv 0$ sur $\sigma(T)$; donc $P_1(T,T^*) = P_2(T,T^*)$. On peut donc poser $\psi_T(f) := P(T,T^*)$, où P est n'importe quel polynôme de $\mathbb{C}[X,Y]$ tel que $P(z,\bar z) \equiv f(z)$ sur $\sigma(T)$. L'application $\psi_T : \mathcal{P} \to \mathcal{L}(H)$ ainsi définie est un homomorphisme de $\mathcal{P}$ dans $\mathcal{L}(H)$, et ψ_T est une isométrie par le Fait 3.22.

Par définition, $\mathcal{P}$ est une sous-algèbre auto-conjuguée de $\mathcal{C}(\sigma(T))$ qui contient les fonctions constantes, et $\mathcal{P}$ sépare les points de $\sigma(T)$ car la fonction $\mathbf{z}$ sépare déjà les points à elle toute seule. Donc $\mathcal{P}$ est dense dans $\mathcal{C}(\sigma(T))$ d'après le Théorème de Stone-Weierstrass.

On peut maintenant laisser sans inquiétude la fin de la preuve du théorème en exo : il suffit de recopier ce qui a été fait dans le cas auto-adjoint. □

Remarque. Comme dans le cas auto-adjoint, le calcul fonctionnel continu pour un opérateur normal est compatible avec le calcul fonctionnel holomorphe.

Corollaire 3.28. *Soit $T \in \mathcal{L}(H)$ un opérateur normal. Alors T est auto-adjoint si et seulement si $\sigma(T) \subseteq \mathbb{R}$; et donc T est positif si et seulement si $\sigma(T) \subseteq \mathbb{R}^+$.*

Démonstration. Si $\sigma(T) \subseteq \mathbb{R}$, alors la fonction $\mathbf{z} \in \mathcal{C}(\sigma(T))$ est à valeurs réelles, donc $T = \mathbf{z}(T) = \bar{\mathbf{z}}(T) = T^*$. $\square$

COROLLAIRE 3.29. *Soit $T \in \mathcal{L}(H)$ un opérateur normal. Si $\lambda \in \sigma(T)$ est un point isolé de $\sigma(T)$, alors la décomposition de Riesz $H = E_\lambda \oplus F$ associée à la décomposition $\sigma(T) = \{\lambda\} \cup \sigma(T)\backslash\{\lambda\}$ est orthogonale, et on a $E_\lambda = \ker(T - \lambda I)$. En particulier, tout point isolé de $\sigma(T)$ est une valeur propre de T.*

Démonstration. La projection $P_\lambda : H \to E_\lambda$ associée à la décomposition $H = E_\lambda \oplus F$ est obtenue par calcul fonctionnel holomorphe ; plus précisément, $P_\lambda = f(T)$ où $f \in \mathcal{H}(T)$ vaut 1 au voisinage de λ et 0 au voisinage de $\sigma(T)\backslash\{\lambda\}$. Mais comme T est normal, on peut aussi utiliser le calcul fonctionnel continu : on a $f \equiv \mathbf{1}_{\{\lambda\}}$ sur $\sigma(T)$, et la fonction $\mathbf{1}_{\{\lambda\}}$ est continue sur $\sigma(T)$, donc $P_\lambda = \mathbf{1}_{\{\lambda\}}(T)$. Comme $\mathbf{1}_{\{\lambda\}}$ est à valeurs réelles, la projection P_λ est un opérateur auto-adjoint, et donc c'est une projection *orthogonale*. Ainsi, la décomposition $H = E_\lambda \oplus F$ est orthogonale. On en déduit que les sous-espaces E_λ et F sont invariants par T^*. Par conséquent, l'opérateur $T_\lambda := T_{|X_\lambda}$ vérifie $T_\lambda^* = T^*_{|X_\lambda}$, et donc T_λ est normal. Donc $T_\lambda - \lambda I_{E_\lambda}$ est normal également, avec $\sigma(T_\lambda - \lambda I_{E_\lambda}) = \{0\}$. Comme $r(A) = \|A\|$ pour tout opérateur normal A, on a donc $T_\lambda = \lambda I_{E_\lambda}$. Ainsi $E_\lambda \subseteq \ker(T - \lambda I)$, et en particulier λ est valeur propre de T. Enfin, $\ker(T - \lambda I)$ est exactement égal à E_λ car $(T - \lambda I)_{|F}$ est inversible (exo). $\square$

3.3. C^*- algèbres commutatives.

DÉFINITION 3.30. *Soit A une $\mathbb{C}$-algèbre. Une **involution** sur A est une application $a \mapsto a^*$ de A dans A qui est antilinéaire, et vérife*

- *$(ab)^* = b^*a^*$ pour tous $a, b \in A$;*
- *$(a^*)^* = a$ pour tout $a \in A$.*

REMARQUE. Si A est unitaire, on a nécessairement $\mathbf{1}^* = \mathbf{1}$.

Démonstration. Exo. $\square$

DÉFINITION 3.31. *Une **C^*- algèbre** est une $\mathbb{C}$-algèbre de Banach A munie d'une involution $a \mapsto a^*$ telle que*

$$\forall a \in A \ : \ \|a^*a\| = \|a\|^2.$$

EXEMPLES.

(1) $A := \mathbb{C}$ est une C^*- algèbre (avec l'involution $z^* := \bar{z}$).

(2) Si H est un espace de Hilbert complexe, alors $\mathcal{L}(H)$ est une C^*- algèbre. Plus généralement, toute sous-algèbre auto-adjointe et fermée $\mathcal{B} \subseteq \mathcal{L}(H)$ est une C^*- algèbre.

(3) Si K est un espace topologique compact, alors $\mathcal{C}(K)$ est une C^*- algèbre (avec l'involution $f^* := \bar{f}$).

(4) Si $(\Omega, \mathfrak{B}, m)$ est un espace mesuré, alors $L^\infty := L^\infty(\Omega, m)$ est une C^*- algèbre ($f^* := \bar{f}$).

DÉFINITION 3.32. *Si A et B sont des C^*- algèbres, un **$*$-homomorphisme** de A dans B est un homomorphisme $\Phi : A \to B$ tel que $\forall a \in A \ : \ \Phi(a^*) = \Phi(a)^*$.*

Exemple. Soit Hun espace de Hilbert complexe. Si $T \in \mathcal{L}(H)$ est un opérateur normal, alors le calcul fonctionnel continu $\Psi_T = \mathcal{C}(\sigma(T)) \to \mathcal{L}(H)$ est un $*$-isomorphisme isométrique.

Théorème 3.33. (Gelfand-Naimark)
Si A est une C^*- algèbre commutative *avec unité, alors A est isométriquement $*$-isomorphe à $\mathcal{C}(\Gamma)$, pour un certain espace topologique compact Γ.*

Remarque. D'après le Théorème de Banach-Stone (Théorème 5.3 du Chapitre 5), l'espace Γ est uniquement déterminé à homéomorphisme près.

Preuve du Théorème de Gelfand-Naimark. Soit $\Gamma_A := \mathrm{Hom}(A, \mathbb{C})$, l'ensemble de tous les homomorphismes $\gamma : A \to \mathbb{C}$. Comme A est une $\mathbb{C}$-algèbre de Banach commutative, Γ_A est non-vide d'après le Lemme 3.24, et on a

$$\forall x \in A \ : \ r(x) = \sup \big\{ |\gamma(x)|;\ \gamma \in \Gamma_A \big\}. \tag{3.1}$$

D'après le Lemme 3.26, tout homomorphisme $\gamma : A \to \mathbb{C}$ est une forme linéaire continue avec $\|\gamma\| \leqslant 1$. Donc, Γ_A est une partie de B_{A^*}. De plus, il est facile de vérifier que Γ_A est w^*- fermé dans A^* (exo). Donc, (Γ_A, w^*) est un espace topologique compact.
Pour tout $x \in A$, notons $\widehat{x} : \Gamma_A \to \mathbb{C}$ la fonction définie par

$$\widehat{x}(\gamma) := \gamma(x) = \langle \gamma, x\rangle \qquad \text{pour tout } \gamma \in \Gamma_A.$$

Par définition de la topologie w^*, on voit que $\widehat{x}$ est une fonction *continue* sur (Γ_A, w^*) ; autrement dit $\widehat{x} \in \mathcal{C}(\Gamma_A)$.
L'application $\mathcal{G}_A : x \mapsto \widehat{x}$ est un *homomorphisme* de A dans $\mathcal{C}(\Gamma_A)$, car si $x, y \in A$ et $\gamma \in \Gamma_A$, alors $\widehat{xy}(\gamma) = \gamma(xy) = \gamma(x)\gamma(y) = \widehat{x}(\gamma)\widehat{y}(\gamma)$, et $\widehat{\mathbf{1}}(\gamma) = \gamma(\mathbf{1}) = 1$ pour tout $\gamma \in \Gamma_A$. De plus, (3.1) dit exactement que

$$\forall x \in A \ : \ \|\widehat{x}\|_\infty = r(x).$$

On dit que $\mathcal{G}_A$ est la **transformation de Gelfand** pour l'algèbre de Banach A.
On montre en recopiant la preuve du Lemme 3.27 que si $x \in A$ et $\gamma \in \Gamma_A$, alors $\gamma(x^*) = \overline{\gamma(x)}$. Ainsi, on a $\widehat{x^*} = \overline{\widehat{x}}$ pour tout $x \in A$; autrement dit $\mathcal{G}_A$ est un $*$-homomorphisme.
Comme A est une C^*- algèbre commutative, on montre en recopiant la preuve de la Proposition 1.23 que si $x \in A$, alors $r(x) = \|x\|$. Ainsi, on a

$$\forall x \in A \ : \ \|\widehat{x}\|_\infty = \|x\|.$$

Autrement dit, $\mathcal{G}_A$ est une isométrie de A dans $\mathcal{C}(\Gamma_A)$.
Pour conclure, il suffit de montrer que $\mathcal{G}_A$ est bijective. Soit $\widehat{A} := \mathcal{G}_A(A) \subseteq \mathcal{C}(\Gamma_A)$. Comme $\mathcal{G}_A$ est un $*$-homomorphisme, $\widehat{A}$ est une sous-algèbre auto-conjuguée de $\mathcal{C}(\Gamma_A)$ contenant $\mathbf{1}$. De plus, $\widehat{A}$ sépare les points de Γ_A : si $\gamma \neq \gamma'$, alors on a un $x \in A$ tel que $\gamma(x) \neq \gamma'(x)$, *i.e.* $\widehat{x}(\gamma) \neq \widehat{x}(\gamma')$. Par le Théorème de Stone-Weierstrass, on en déduit que $\widehat{A}$ est dense dans $\mathcal{C}(\Gamma_A)$. Mais $\widehat{A}$ est aussi fermée dans $\mathcal{C}(\Gamma_A)$ car $\mathcal{G}_A$ est une isométrie. Donc $\widehat{A} = \mathcal{C}(\Gamma_A)$, ce qui termine la preuve. □

Le Théorème de Gelfand-Naimark montre en particulier que tout espace du type $L^\infty = L^\infty(\Omega, m)$ "est" en fait un espace $\mathcal{C}(\Gamma)$. La preuve du théorème a montré que le compact Γ est l'espace des homomorphismes de L^∞ dans $\mathbb{C}$, ce qui reste assez formel et n'est peut-être pas extraordinairement parlant. Dans le cas de $\ell^\infty(\mathbb{N})$, on peut donner une description plus “concrète” de Γ. C'est le contenu du très long exercice suivant.

EXERCICE. On dit qu'une famille $\mathcal{F}$ de parties de $\mathbb{N}$ est un **filtre** si : $\mathbb{N}\in\mathcal{F}$; tous les éléments de $\mathcal{F}$ sont non vides ; $\mathcal{F}$ est stable par intersections finies ; $\mathcal{F}$ est co-héréditaire, *i.e.* $(I\in\mathcal{F}$ et $I'\supseteq I)\implies (I'\in\mathcal{F})$. Un **ultrafiltre** est un filtre $\mathcal{U}$ qui est maximal pour l'inclusion. On note $\beta\mathbb{N}$ l'ensemble de tous les ultrafiltres $\mathcal{U}\subseteq\mathcal{P}(\mathbb{N})$.

(1) Montrer que tout filtre $\mathcal{F}\subseteq\mathcal{P}(\mathbb{N})$ est contenu dans un ultrafiltre.
(2) Montrer que si $n\in\mathbb{N}$, alors $\mathcal{U}_n:=\{A\subseteq\mathbb{N};\ A\ni n\}$ est un ultrafiltre. Un ultrafiltre de ce type est dit *trivial*.
(3) Montrer que pour un filtre $\mathcal{U}\subseteq\mathcal{P}(\mathbb{N})$, les choses suivantes sont équivalentes :
    (i) $\mathcal{U}$ est un ultrafiltre ;
    (ii) pour tout $A\subseteq\mathbb{N}$, on a $A\in\mathcal{U}$ ou $A^c\in\mathcal{U}$;
    (iii) $\forall A_1,\dots,A_n\subseteq\mathbb{N}\ :\ (A_1\cup\cdots\cup A_n\in\mathcal{U})\implies (A_1\in\mathcal{U}\text{ ou}\cdots\text{ou }A_n\in\mathcal{U})$.
(4) Montrer qu'un ultrafiltre $\mathcal{U}$ est non-trivial si et seulement si il contient le *filtre de Fréchet* $\mathcal{F}_\infty:=\big\{A\subseteq\mathbb{N};\ A\text{ cofini}\big\}$.
(5) Étant donné un filtre $\mathcal{F}\subseteq\mathcal{P}(\mathbb{N})$, on dit qu'une suite de nombre complexes $(a_n)_{n\in\mathbb{N}}$ **converge le long de** $\mathcal{F}$ s'il existe $l\in\mathbb{C}$ tel que
$$\forall\varepsilon>0\ :\ \big\{n\in\mathbb{N};\ |a_n-l|<\varepsilon\big\}\in\mathcal{F}.$$
On écrit alors $a_n\underset{\mathcal{F}}{\longrightarrow} l$. Montrer qu'on a “unicité de la limite” : si $a_n\underset{\mathcal{F}}{\longrightarrow} l$ et $a_n\underset{\mathcal{F}}{\longrightarrow} l'$, alors $l=l'$. On peut donc écrire $l=\mathcal{F}\text{-}\lim a_n$. Vérifier également qu'une suite (a_n) converge le long du filtre de Fréchet $\mathcal{F}_\infty$ si et seulement si elle converge au sens usuel, et qu'elle converge le long d'un ultrafiltre trivial $\mathcal{U}_{n_0}$ si et seulement si $a_{n_0}=l$.
(6) Montrer que si $(a_n)_{n\in\mathbb{N}}\subseteq\mathbb{C}$ est une suite *bornée*, alors (a_n) converge le long de tout ultrafiltre $\mathcal{U}$. (*Commencer par montrer que* $\bigcap_{I\in\mathcal{U}}\overline{\{a_n;\ n\in I\}}\neq\varnothing$.)
(7) Pour tout $A\subseteq\mathbb{N}$, on pose $[A]:=\{\mathcal{U}\in\beta\mathbb{N};\ \mathcal{U}\ni A\}$. Établir les propriétés suivantes :
    - $[\varnothing]=\varnothing$ et $[\mathbb{N}]=\beta\mathbb{N}$;
    - $[A^c]=[A]^c$;
    - $[A_1\cup\cdots\cup A_n]=[A_1]\cup\cdots\cup[A_n]$ et $[A_1\cap\cdots\cap A_n]=[A_1]\cap\cdots\cap[A_n]$.
(8) Montrer qu'on définit une topologie sur $\beta\mathbb{N}$ en décrétant qu'un ensemble $\mathbf{O}\subseteq\beta\mathbb{N}$ est ouvert si et seulement si : pour tout $\mathcal{U}\in\mathbf{O}$, on peut trouver $A\subseteq\mathbb{N}$ tel que $\mathcal{U}\in[A]$ et $[A]\subseteq\mathbf{O}$. Montrer également que $\mathfrak{A}:=\big\{[A];\ A\subseteq\mathbb{N}\big\}$ est une base pour cette topologie.
(9) Montrer que $\beta\mathbb{N}$ est compact. (*Commencer par montrer que si* $(A_i)_{i\in I}$ *est une famille de parties de* $\mathbb{N}$ *telle que* $\bigcup_{i\in F}A_i\neq\mathbb{N}$ *pour tout ensemble fini* $F\subseteq I$*, alors il existe un ultrafiltre* $\mathcal{U}$ *tel que* $\forall i\in I\ :\ A_i^c\in\mathcal{U}$.)
(10) On munit $\mathbb{N}$ de la topologie discrète. Montrer que l'application $\mathbf{i}:\mathbb{N}\to\beta\mathbb{N}$ définie par $\mathbf{i}(n):=\mathcal{U}_n$ est un homéomorphisme de $\mathbb{N}$ sur $\mathbf{i}(\mathbb{N})$, et que $\mathbf{i}(\mathbb{N})$ est dense dans $\beta\mathbb{N}$. Montrer également que $\mathbf{i}(\mathbb{N})$ est un ouvert de $\beta\mathbb{N}$. Dans la suite, on pourra donc considérer $\mathbb{N}$ comme un ouvert dense de $\beta\mathbb{N}$.

(11) Pour toute suite $x = \big(x(n)\big)_{n\in\mathbb{N}} \in \ell^\infty(\mathbb{N})$, on définit une fonction $\widehat{x} : \beta\mathbb{N} \to \mathbb{C}$ comme suit :

$$\forall \mathcal{U} \in \beta\mathbb{N} \;:\; \widehat{x}(\mathcal{U}) := \mathcal{U}\text{-}\lim a(n).$$

Montrer que la fonction $\widehat{x}$ est continue. Plus précisément : montrer que $\widehat{x}$ est l'unique fonction continue $f : \beta\mathbb{N} \to \mathbb{C}$ telle que $f_{|\mathbb{N}} = x$.

(12) Montrer que l'application $x \mapsto \widehat{x}$ est un $*$-isomorphisme isométrique de $\ell^\infty(\mathbb{N})$ sur $\mathcal{C}(\beta\mathbb{N})$.

4. Théorème spectral pour les opérateurs normaux

4.1. Opérateurs normaux et opérateurs de multiplication.

LEMME 4.1. *Soit $(\Omega, \mathfrak{B}, m)$ un espace mesuré sigma-fini. Si $\phi \in L^\infty = L^\infty(\Omega, m)$, on note $M_\phi : L^2 \to L^2$ l'opérateur de multiplication associé.*

(i) *L'application $\phi \mapsto M_\phi$ est un homomorphisme isométrique de L^∞ dans $\mathcal{L}(L^2)$.*

(ii) *On a $(M_\phi)^* = M_{\bar\phi}$ pour toute $\phi \in L^\infty$; donc M_ϕ est un opérateur normal.*

(iii) *On a $\sigma(M_\phi) = \sigma_{L^\infty}(\phi)$ pour toute $\phi \in L^\infty$.*

Démonstration. On a déjà montré (i) et (ii). Pour démontrer (iii), notons $J : L^\infty \to \mathcal{L}(L^2)$ l'application $\phi \mapsto M_\phi$, et posons $B := J(L^\infty) \subseteq \mathcal{L}(L^2)$. Par (i) et (ii), B est une sous-algèbre fermée et auto-adjoite de $\mathcal{L}(L^2)$ contenant I, et J est un isomorphisme de L^∞ sur B. Par le Lemme 3.23, on en déduit que si $\phi \in L^\infty$, alors $\sigma_{L^\infty}(\phi) = \sigma_B(J\phi) = \sigma_B(M_\phi) = \sigma_{\mathcal{L}(L^2)}(M_\phi)$. □

REMARQUE. On a vu que si $\phi \in L^\infty = L^\infty(\Omega, m)$, alors $\sigma_{L^\infty}(\phi)$ est le plus petit fermé $K \subseteq \mathbb{C}$ tel que $\phi(t) \in K$ pp.

RAPPEL. Soient H_1 et H_2 deux espaces de Hilbert, et soient $T_1 \in \mathcal{L}(H_1)$ et $T_2 \in \mathcal{L}(H_2)$. On dit que T_1 et T_2 sont *unitairement équivalents* s'il existe un opérateur unitaire $U = H_1 \to H_2$ tel que $T_2U = UT_1$.

THÉORÈME 4.2. *Soit H un espace de Hilbert complexe séparable. Si $T \in \mathcal{L}(H)$ est un opérateur normal, alors il existe un espace mesuré sigma-fini $(\Omega, \mathfrak{B}, m)$ et une fonction $\phi \in L^\infty(\Omega, m)$ tels que T est unitairement équivalent à l'opérateur de multiplication $M_\phi : L^2(\Omega, m) \to L^2(\Omega, m)$. De plus, s'il existe un vecteur $x_0 \in H$ tel que le sous-espace $\operatorname{vect}\{T^kT^{*l}x_0;\; k, l \in \mathbb{N}\}$ soit dense dans H, alors on peut prendre $\Omega := \sigma(T)$ et $\phi := \mathbf{z}$, et on peut prendre pour m une mesure de probabilité.*

EXEMPLE. Soit S le “forward shift” sur $\ell^2(\mathbb{Z})$. L'opérateur S est unitaire, donc normal, et $\operatorname{vect}\{S^kS^{*l}e_0;\; k, l \in \mathbb{N}\} = \operatorname{vect}\{S^ne_0;\; n \in \mathbb{Z}\}$ est dense dans $\ell^2(\mathbb{Z})$. Comme $\sigma(S) = \mathbb{T}$, le théorème affirme que S est unitairement équivalent à $M_\mathbf{z}$ agissant sur $L^2(\mathbb{T}, m)$ pour une certaine mesure de probabilité m sur $\mathbb{T}$; et de fait, on a vu au Chapitre 1 qu'on peut prendre pour m la mesure de Lebesgue sur $\mathbb{T}$ (l'équivalence unitaire étant donnée par la transformation de Fourier).

Preuve du Théorème 4.2. Comme on s'en doute au vu de l'énoncé du théorème, on va distinguer deux cas.

CAS 1. On supose qu'il existe $x_0 \in H$ tel que $\overline{\operatorname{vect}}\{T^kT^{*l}x_0;\; k, l \in \mathbb{N}\} = H$.

On peut évidemment supposer que $\|x_0\| = 1$.

FAIT 1. Il existe une mesure de probabilité borélienne m sur $\sigma(T)$ telle que

$$\forall f \in \mathcal{C}(\sigma(T)) \ : \ \langle f(T)x_0, x_0\rangle = \int_{\sigma(T)} f\, dm.$$

Preuve du Fait 1. Soit $L : \mathcal{C}(\sigma(T)) \to \mathbb{C}$ la forme linéaire définie par

$$L(f) := \langle f(T)x_0, x_0\rangle.$$

Si $f \in \mathcal{C}(\sigma(T))$ est réelle $\geqslant 0$, alors l'opérateur $f(T)$ est positif; donc $L(f) \geqslant 0$. Donc L est une forme linéaire positive. D'après le Théorème de représentation de Riesz, il existe donc une mesure $m \in M_+(\sigma(T))$ telle que $L(f) = \int_{\sigma(T)} f\, dm$ pour toute $f \in \mathcal{C}(\sigma(T))$. On a $m(\sigma(T)) = L(\mathbf{1}) = \langle x_0, x_0\rangle = \|x_0\|^2 = 1$, donc m est une mesure de probabilité. □

FAIT 2. Pour toute $f \in \mathcal{C}(\sigma(T))$, on a

$$\|f(T)x_0\|^2 = \int_{\sigma(T)} |f|^2\, dm.$$

Preuve du Fait 2. C'est "immédiat" :

$$\begin{aligned}\|f(T)x_0\|^2 &= \langle f(T)x_0, f(T)x_0\rangle = \langle f(T)^* f(T)x_0, x_0\rangle \\ &= \langle (\bar{f}f)(T)x_0, x_0\rangle = \langle |f|^2(T)x_0, x_0\rangle = \int_{\sigma(T)} |f|^2\, dm.\end{aligned}$$

□

Par le Fait 2, on définit une isométrie $V : \big(\mathcal{C}(\sigma(T)), \|\cdot\|_{L^2(m)}\big) \to H$ en posant

$$Vf := f(T)x_0 \qquad \text{pour toute } f \in \mathcal{C}(\sigma(T)).$$

De plus, $\mathrm{Im}(V)$ contient $\big\{P(T,T^*)x_0;\ P \in \mathbb{C}[X,Y]\big\} = \mathrm{vect}\,\big\{T^kT^{*l}x_0;\ k,l \in \mathbb{N}\big\}$, donc $\mathrm{Im}(V)$ est dense dans H. Comme $\mathcal{C}(\sigma(T))$ est dense dans $L^2(\sigma(T), m)$, on en déduit que V se prolonge en un opérateur unitaire $U : L^2(\sigma(T), m) \to H$. Si $f \in \mathcal{C}(\sigma(T))$, alors $T(Vf) = Tf(T)x_0 = (\mathbf{z}f)(T)x_0 = V(\mathbf{z}f)$. Donc $T(Uf) = U(\mathbf{z}f)$ pour toute $f \in L^2(\sigma(T), m)$ par densité de $\mathcal{C}(\sigma(T))$ dans L^2 ; autrement dit $TU = UM_{\mathbf{z}}$. On a donc prouvé le résultat souhaité.

CAS 2. Cas général.

Pour $x \in H$, on notera E_x le "sous-espace fermé $\mathcal{C}^*(T)$-cyclique engendré par x" :

$$E_x := \overline{\mathrm{vect}}\,\big\{T^kT^{*l}x;\ k,l \in \mathbb{N}\big\}.$$

FAIT 1. Il existe une famille dénombrable $(x_i)_{i\in I} \subseteq H$ telle que $\|x_i\| = 1$ pour tout $i \in I$, les E_{x_i} sont deux à deux orthogonaux et $H = \overline{\mathrm{vect}}\Big(\bigcup_{i\in I} E_{x_i}\Big)$.

Preuve du Fait. Par le Lemme de Zorn, il existe un ensemble $\mathcal{F} \subseteq S_H$ tel que les E_x, $x \in \mathcal{F}$ sont deux à deux orthogonaux et $\mathcal{F}$ est *maximal* relativement à cette propriété. Écrivons $\mathcal{F} = \{x_i;\ i \in I\}$, où les x_i sont deux à deux distincts. Alors I est nécessairement dénombrable car $(x_i)_{i\in I}$ est une famille orthonormale et H est séparable.

Soit $E := \overline{\mathrm{vect}}\Big(\bigcup_{i\in I} E_{x_i}\Big)$. Si on avait $E \neq H$, on pourrait trouver $x \in S_H$ tel que $x \perp E$. Alors $\langle T^kT^{*l}x, T^{k'}T^{*l'}x_i\rangle = \langle x, T^{l+k'}T^{*(k+l')}x_i\rangle = 0$ pour tout $i \in I$ et pour tous $k, l, k', l' \in \mathbb{N}$; donc $E_x \perp E_{x_i}$ pour tout $i \in I$ par continuité du produit scalaire, ce qui contredit la maximalité de $\mathcal{F}$. Par conséquent, $E = H$. □

Fait 2. Soit $\mathcal{H}$ un espace de Hilbert, et soit $(\mathcal{E}_i)_{i\in I}$ une famille de sous-espaces fermés de $\mathcal{H}$ deux à deux orthogonaux, avec I fini ou $I = \mathbb{N}$. Posons

$$\mathcal{E} := \overline{\mathrm{vect}}\left(\bigcup_{i\in I} \mathcal{E}_i\right).$$

Alors tout vecteur $h \in \mathcal{E}$ s'écrit de manière unique sous la forme

$$h = \sum_{i\in I} h_i,$$

où $h_i \in \mathcal{E}_i$ et la série converge dans $\mathcal{H}$ si $I = \mathbb{N}$, et on a

$$\|h\|^2 = \sum_{i\in I} \|h_i\|^2.$$

En conséquence, on écrit

$$\mathcal{E} = \overset{\perp}{\bigoplus_{i\in I}} \mathcal{E}_i.$$

Preuve du Fait 2. C'est un exo. On a $h_i = p_i h$, où p_i est la projection orthogonale de $\mathcal{H}$ sur $\mathcal{E}_i$. □

Soit $(x_i)_{i\in I} \subseteq H$ donnée par le Fait 1,. Comme I est dénombrable, on peut supposer que I est fini ou bien que $I = \mathbb{N}$.

Posons $H_i := E_{x_i}$ et $T_i := T_{|H_i} \in \mathcal{L}(H_i)$. Comme les H_i sont invariants par T et par T^*, on a $T_i^* = (T^*)_{|H_i}$, donc les T_i sont des opérateurs normaux puisque T est normal ; et par définition de H_i, on peut appliquer le Cas 1 à chaque T_i. Il existe donc un espace de probabilité $(\Omega_i, \mathfrak{B}_i, m_i)$, une fonction $\phi_i \in L^\infty(\Omega_i, m_i)$ et un opérateur unitaire $U_i : H_i \to L^2(\Omega_i, m_i)$ tels que $T_i = U_i^{-1} M_{\phi_i} U_i$.

Soit alors $(\Omega, \mathfrak{B}, m) := \bigsqcup_{i\in I} (\Omega_i, \mathfrak{B}_i, m_i)$ la “réunion disjointe” des espaces mesurés $(\Omega_i, \mathfrak{B}_i, m_i)$. Formellement : Ω est la réunion disjointe des Ω_i, *i.e.* $\Omega = \bigcup_{i\in I}\{i\} \times \Omega_i$; en identifiant Ω_i à $\{i\} \times \Omega_i \subseteq \Omega$, la tribu $\mathfrak{B}$ est l'ensemble des $A \subseteq \Omega$ tels que $A \cap \Omega_i \in \mathfrak{B}_i$ pour tout $i \in I$; et la mesure m est définie par

$$m(A) := \sum_{i\in I} m_i(A \cap \Omega_i) \qquad \text{pour tout } A \in \mathfrak{B}.$$

La mesure m est sigma-finie car I est dénombrable et les m_i sont finies ; et par définition de $(\Omega, \mathfrak{B}, m)$, on a

$$L^2(\Omega, m) = \overset{\perp}{\bigoplus_{i\in I}} L^2(\Omega_i, m_i),$$

où on a identifié $L^2(\Omega_i, m_i)$ avec $\mathcal{E}_i := \{f \in L^2(\Omega, m);\ f \equiv 0 \text{ sur } \Omega\backslash\Omega_i\}$.

Comme $H = \overset{\perp}{\bigoplus}_{i\in I} H_i$, on définit donc un opérateur unitaire $U : H \to L^2(\Omega, m)$ en posant

$$U\big(\oplus_{i\in I} h_i\big) := \oplus_{i\in I} U_i h_i.$$

Et par définition, on a $T = U^{-1}M_\phi U$, où $\phi \in L^\infty(\Omega, m)$ est l'unique fonction telle que $\phi \equiv \phi_i$ sur Ω_i pour tout $i \in I$.

$\square$

Remarque. Dans le Théorème spectral, on peut en fait toujours prendre pour m une mesure *de probabilité*.

Démonstration. Il suffit de remplacer la mesure m donnée par le théorème par $\widetilde{m} := hm$, où h est une fonction mesurable strictement positive telle que $\int_\Omega h\, dm = 1$ (exo : il en existe), et de remplacer l'opérateur $U : H \to L^2(\Omega, m)$ par $\widetilde{U} := JU : H \to L^2(\Omega, \widetilde{m})$, où $J : L^2(\Omega, m) \to L^2(\Omega, \widetilde{m})$ est l'opérateur unitaire défini par $Jf := h^{-1/2}f$. $\square$

4.2. Calcul fonctionnel borélien.

Notation. Dans ce qui suit, la lettre M désigne un espace métrique. On note $\mathbf{B}(M)$ l'ensemble de toutes les fonctions *boréliennes bornées* $f : M \to \mathbb{C}$.

Définition 4.3. *Soit $(f_n)_{,\in\mathbb{N}}$ une suite d'éléments de $\mathbf{B}(M)$, et soit $f \in \mathbf{B}(M)$. On dit que f_n* ***tend vers*** f ***dans*** $\mathbf{B}(M)$ *si (f_n) est uniformément bornée et $f_n(z) \to f(z)$ pour tout $z \in M$.*

Remarque 1. Soit H un espace de Hilbert. Si $\mathcal{E} \subseteq \mathbf{B}(M)$, on dira qu'une application $\Theta : \mathcal{E} \to \mathcal{L}(H)$ est $\mathtt{SOT}$ *séquentiellement continue* (qu'on abrège en "$\mathtt{SOT}$-sc") si, pour toute suite $(f_n) \subseteq \mathcal{E}$ telle que $f_n \to f \in \mathcal{E}$ dans $\mathbf{B}(M)$, on a que $\Theta(f_n) \xrightarrow{\mathtt{SOT}} \Theta(f)$, *i.e.* $\Theta(f_n)x \to \Theta(f)x$ pour tout $x \in H$. De même, on dit que $\Theta : \mathcal{E} \to \mathcal{L}(H)$ est $\mathtt{WOT}$-*séquentiellement continue* ($\mathtt{WOT}$-sc) si, pour toute suite $(f_n) \subseteq \mathcal{E}$ telle que $f_n \to f \in \mathcal{E}$ dans $\mathbf{B}(M)$, on a que $\Theta(f_n)x \xrightarrow{w} \Theta(f)x$ pour tout $x \in H$.

Exemple. Si $T \in \mathcal{L}(H)$ est un opérateur normal, alors le calcul fonctionnel continu $\Psi_T : \mathcal{C}(\sigma(T)) \to \mathcal{L}(H)$ est $\mathtt{SOT}$-sc.

Démonstration. Montrons d'abord que Ψ_T est $\mathtt{WOT}$-sc. Pour cela, il suffit de montrer que si $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}(\sigma(T))$ est une suite bornée telle que $f_n(z) \to 0$ pout tout $z \in \sigma(T)$, alors $f(T_n)x \xrightarrow{w} 0$ pour tout $x \in H$. Mais ceci est clair : comme (f_n) est bornée et tend simplement vers 0, on voit que $f_n \xrightarrow{w} 0$ dans $\mathcal{C}(\sigma(T))$ en utilisant le fait que $\mathcal{C}(\sigma(T))^* = M(\sigma(T))$ et le théorème de convergence dominée ; donc $f(T_n)x \xrightarrow{w} 0$ car l'application linéaire $f \mapsto f(T)x$ est continue de $\mathcal{C}(\sigma(T))$ dans H, et donc continue de $\big(\mathcal{C}(\sigma(T)), w\big)$ dans (H, w).
Montrons maintenant que Ψ_T est en fait $\mathtt{SOT}$-sc. Si $(f_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}(\sigma(T))$ est une suite bornée telle que $f_n(z) \to 0$ pout tout $z \in \sigma(T)$ et si $x \in H$, alors $\|f_n(T)x\|^2 = \langle f_n(T)x, f_n(T)x\rangle = \langle f_n(T)^* f_n(T)x, x\rangle = \langle |f_n|^2(T)x, x\rangle$; donc $\|f_n(T)x\| \to 0$ par $\mathtt{WOT}$ continuité séquentielle, puisque la suite $(|f_n|^2)_{n\in\mathbb{N}} \subseteq \mathcal{C}(\sigma(T))$ est bornée et tend simplement vers 0. $\square$

Remarque 2. On dira qu'un ensemble $\mathcal{E} \subseteq \mathbf{B}(M)$ est *séquentiellement fermé* si, pour toute suite $(f_n) \subseteq \mathcal{E}$ telle que $f_n \to f \in \mathbf{B}(M)$, on a que $f \in \mathcal{E}$.

Lemme 4.4. *Soit $\mathcal{C}_b(M) := \{f : M \to \mathbb{C}$ continues bornées$\}$. Si $\mathcal{E} \subseteq \mathbf{B}(M)$ est un sous-espace vectoriel contenant $\mathcal{C}_b(M)$ et séquentiellement fermé, alors $\mathcal{E} = \mathbf{B}(M)$.*

Démonstration. Comme toute fonction $f \in \mathbf{B}(M)$ est limite uniforme de fonctions boréliennes étagées, il suffit de montrer que $\mathcal{E}$ contient toutes les fonctions étagées ; et comme $\mathcal{E}$ est un espace vectoriel, il suffit de montrer que $\mathbf{1}_A \in \mathcal{E}$ pour tout borélien $A \subseteq M$. Autrement dit, en posant $\mathcal{M} := \{A \subseteq M;\ \mathbf{1}_A \in \mathcal{E}\}$, il s'agit de voir que $\mathcal{M}$ contient tous les boréliens de M.
Il est très facile de vérifier que la famille $\mathcal{M}$ est une *classe monotone* :

- si $A, A' \in \mathcal{M}$ et $A \subseteq A'$, alors $A' \backslash A \in \mathcal{M}$ (car $\mathbf{1}_{A'\backslash A} = \mathbf{1}_{A'} - \mathbf{1}_A$) ;
- si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'éléments de $\mathcal{M}$, alors $A := \bigcup_{n=0}^\infty A_n \in \mathcal{M}$ (car $\mathbf{1}_A = \lim \mathbf{1}_{A_n}$).

Par le Théorème des classes monotones, il suffit donc de montrer que $\mathcal{M}$ contient tous les fermés de M. Mais ceci est clair, car il est bien connu que si $F \subseteq M$ est fermé, alors il existe une suite de fonctions continues $(f_n)_{n\in\mathbb{N}}$ telle que $0 \leqslant f_n \leqslant 1$ et $f_n(x) \to \mathbf{1}_F(x)$ pour tout $x \in M$. Par exemple, on peut prendre $f_n(x) := \max\big(0, 1 - n\,\mathrm{dist}(x, M)\big)$ (exo). □

THÉORÈME 4.5. *Soit H un espace de Hilbert complexe séparable, et soit $T \in \mathcal{L}(H)$ un opérateur normal.*

(1) *Il existe un unique homomorphisme* `WOT`*-sc $\Theta_T : \mathbf{B}(\sigma(T)) \to \mathcal{L}(H)$ tel que $\Theta_T(\mathbf{z}) = T$ et $\Theta_T(\bar{\mathbf{z}}) = T^*$; et cet homomorphisme est en fait* `SOT`*-sc. L'homomorphisme Θ_T prolonge le calcul fonctionnel continu Ψ_T, et on a $\|\Theta_T(f)\| \leqslant \|f\|_\infty := \sup\big\{|f(z)|;\ z \in \sigma(T)\big\}$ pour toute $f \in \mathbf{B}(\sigma(T))$. On dit que Θ_T est le* ***calcul fonctionnel borélien*** *pour l'opérateur T.*

(2) *Si $(\Omega, \mathfrak{B}, m)$ et $\phi \in L^\infty(\Omega, m)$ sont tels que T est unitairement équivalent à $M_\phi : L^2(\Omega, m) \to L^2(\Omega, m)$ avec "témoin" $U : L^2(\Omega, m) \to H$, alors $f(T) := \Theta_T(f)$ est unitairement équivalent à $M_{f\circ\phi}$ pour toute $f \in \mathbf{B}(\sigma(T))$, avec le même témoin U. En particulier, on a*
   - $f(T)^* = \bar{f}(T)$,
   - $\|f(T)\| = \|f\|_{L^\infty(\Omega,m)}$,
   - $\sigma(f(T)) = \sigma_{L^\infty(\Omega,m)}(f \circ \phi)$.

(3) *Si $f \in \mathbf{B}(\sigma(T))$, alors $f(T)A = Af(T)$ pour tout opérateur $A \in \mathcal{L}(H)$ tel que $AT = TA$ et $AT^* = T^*A$.*

Démonstration. (1) D'après le Théorème spectral, il existe un espace mesuré sigma-fini $(\Omega, \mathfrak{B}, m)$, une fonction $\phi \in L^\infty = L^\infty(\Omega, m)$ et un opérateur unitaire $U : H \to L^2(\Omega, m)$ tels que $T = U^{-1} M_\phi U$.
Comme $\sigma(T) = \sigma(M_\phi) = \sigma_{L^\infty}(\phi)$, on sait que $\phi(t) \in \sigma(T)$ m-pp. Donc, si $f \in \mathbf{B}(\sigma(T))$, alors la fonction $f \circ \phi$ est bien définie m-pp, et $f \circ \phi \in L^\infty(\Omega, m)$ car f est bornée. L'opérateur $M_{f\circ\phi}$ est donc bien défini, et on peut poser

$$\Theta(f) := U^{-1} M_{f\circ\phi} U \in \mathcal{L}(H).$$

Par définition, l'application $\Theta : \mathbf{B}(\sigma(T)) \to \mathcal{L}(H)$ est un homomorphisme d'algèbres car l'application $f \mapsto f \circ \phi$ en est un ; et on a $\Theta(\mathbf{z}) = U^{-1} M_\phi U = T$ et $\Theta(\bar{\mathbf{z}}) = U^{-1} M_{\bar{\phi}} U = T^*$.

Montrons que Θ est `SOT`-sc. Soit $(f_n) \subseteq \mathbf{B}(\sigma(T))$ une suite convergeant vers 0 dans $\mathbf{B}(\sigma(T))$. Alors $f_n \circ \phi \to 0$ m-pp et la suite $(f_n \circ \phi)$ est bornée dans L^∞. Par convergence dominée, on en déduit que si $u \in L^2(\Omega, m)$, alors

$$\|M_{f_n\circ\phi}u\|^2_{L^2(\Omega,m)} = \int_\Omega |f_n \circ \phi|^2\, |u|^2\, dm \to 0.$$

Donc, si $x \in H$, alors $\|\Theta(f_n)x\| = \left\|U^{-1}M_{f_n\circ\phi}Ux\right\| = \|M_{f_n\circ\phi}Ux\|_{L^2(\Omega,m)} \to 0$.

Si $f \in \mathbf{B}(\sigma(T))$, alors $\|\Theta(f)\| = \|M_{f\circ\phi}\| = \|f \circ \phi\|_{L^\infty(\Omega,m)} \leqslant \|f\|_\infty$. En particulier, la restriction de Θ à $\mathcal{C}(\sigma(T))$ est continue de $\big(\mathcal{C}(\sigma(T)), \|\cdot\|_\infty\big)$ dans $\mathcal{L}(H)$; et donc Θ prolonge le calcul fonctionnel continu $\Psi_T : \mathcal{C}(\sigma(T)) \to \mathcal{L}(H)$ puisque Θ est un homomorphisme tel que $\Theta(\mathbf{z}) = T$ et $\Theta(\bar{\mathbf{z}}) = T^*$.

Soit $\Theta' : \mathbf{B}(\sigma(T)) \to \mathcal{L}(H)$ un autre homomorphisme `WOT`-sc tel que $\Theta(\mathbf{z})$ et $\Theta(\bar{\mathbf{z}}) = T^*$. En appliquant le Théorème du graphe fermé, on voit que la restriction de Φ à $\mathcal{C}(\sigma(T))$ est continue de $\big(\mathcal{C}(\sigma(T)), \|\cdot\|_\infty\big)$ dans $\mathcal{L}(H)$ (exo). Donc Θ' prolonge le calcul fonctionnel continu Ψ_T, et donc $\Theta'(f) = \Theta(f)$ pour toute $f \in \mathcal{C}(\sigma(T))$. Si on pose $\mathcal{E} := \{f \in \mathbf{B}(\sigma(t));\ \Theta'(f) = \Theta(f)\}$, alors $\mathcal{E}$ est un sous-espace vectoriel de $\mathbf{B}(\sigma(T))$, qui est séquentiellement fermé car Θ et Θ' sont `SOT`-sc (exo) ; et on vient de voir que $\mathcal{E}$ contient $\mathcal{C}(\sigma(T))$. Donc $\mathcal{E} = \mathbf{B}(\sigma(T))$ d'après le Lemme 4.4 ; autrement dit $\Theta' = \Theta$.

(2) est une conséquence de la définition de $\Theta_T = \Theta$ donnée dans la preuve de (1).

(3) Soit $A \in \mathcal{L}(H)$ tel que $AT = TA$ et $AT^* = T^*A$. Posons

$$\mathcal{E} := \{f \in \mathbf{B}(\sigma(T));\ f(T)A = Af(T)\}.$$

Il est clair que que $\mathcal{E}$ est un sous-espace vectoriel de $\mathbf{B}(\sigma(T))$; et $\mathcal{E}$ est séquentiellement fermé car le calcul fonctionnel est `WOT`-sc (exo). De plus, $\mathcal{E}$ contient $\mathcal{C}(\sigma(T))$ d'après le Théorème 3.21. Donc $\mathcal{E} = \mathbf{B}(\sigma(T))$, ce qui donne (3). $\square$

Exercice. Montrer que si $f \in \mathbf{B}(\sigma(T))$, alors $\sigma(f(T)) \subseteq \overline{f(\sigma(T))}$. En déduire que si $f \in \mathbf{B}(\sigma(T))$ et si g est une fonction borélienne bornée sur $\overline{f(\sigma(T))}$, alors on peut écrire $g(f(T)) = (g \circ f)(T)$.

Exemple 4.6. Soit $T \in \mathcal{L}(H)$ un opérateur normal, et soit $\lambda \in \sigma(T)$. Pour $\varepsilon > 0$, l'opérateur $P := \mathbf{1}_{D(\lambda,\varepsilon)}(T)$ est une projection orthogonale non-nulle telle que $TP = PT$ et $\|(T - \lambda I)P\| \leqslant \varepsilon$. Donc, $E := \mathrm{Im}(P)$ est un sous-espace fermé de H non réduit à $\{0\}$ tel que $\|(T - \lambda I)_{|E}\| \leqslant \varepsilon$.

Démonstration. Il y a un léger abus de notation : au lieu de $\mathbf{1}_{D(0,\varepsilon)}$, on devrait écrire $\big(\mathbf{1}_{D(0,\varepsilon)}\big)_{|\sigma(T)}$. Dans ce qui suit, on pose $D := D(\lambda, \varepsilon)$; donc $P = \mathbf{1}_D(T)$.

L'opérateur P est défini par calcul fonctionnel borélien, donc $TP = PT$. On a $P^2 = P$ car $\mathbf{1}_D^2 = \mathbf{1}_D$, et P est auto-adjoint car la fonction $\mathbf{1}_D$ est à valeurs réelles. Donc P est une projection orthogonale.

Supposons que $P = 0$. Alors $T = T\mathbf{1}_{\sigma(T)\backslash D}(T)$ car $(\mathbf{1}_D + \mathbf{1}_{\sigma(T)\backslash D})(T) = \mathbf{1}(T) = I$. Autrement dit, $T = f(T)$ où $f(z) := z\mathbf{1}_{\sigma(T)\backslash D}(z)$. Donc $\sigma(T) \subseteq \overline{f(\sigma(T))}$. Mais $f(\sigma(T)) = \sigma(T)\backslash D$ par définition de f, et $\sigma(T)\backslash D$ est un fermé de $\mathbb{C}$ puisque D est un disque ouvert ; donc $\sigma(T) \subseteq \sigma(T)\backslash D$, ce qui est absurde puisque $\sigma(T)\backslash D$ est strictement contenu dans $\sigma(T)$. Donc $P \neq 0$.

Enfin, comme $(T - \lambda I)P = (\mathbf{z} - \lambda\mathbf{1})\mathbf{1}_D(T)$, on a $\|(T - \lambda I)P\| \leqslant \|(\mathbf{z} - \lambda\mathbf{1})\mathbf{1}_D\|_\infty$, et donc $\|(T - \lambda I)P\| \leqslant \varepsilon$ puisque $|z - \lambda| < \varepsilon$ pour tout $z \in D = D(\lambda, \varepsilon)$. $\square$

4.3. Une autre preuve du Théorème spectral. Dans cette section, on donne les grandes lignes d'une preuve du Théorème spectral qui est "directe" au sens où elle n'utilise pas le calcul fonctionnel continu. Un des intérêts d'avoir une preuve directe est qu'une fois le Théorème spectral démontré, on peut en déduire l'existence du calcul fonctionnel borélien, *et donc* le calcul fonctionnel continu. D'un certain point de vue, il parait plus naturel de procéder ainsi plutôt que de commencer par démontrer l'existence du calcul fonctionnel continu, d'en déduire le Théorème spectral et de démontrer ensuite l'existence du calcul fonctionnel borélien.

Il suffit de traiter le cas d'un opérateur normal $T \in \mathcal{L}(H)$ pour lequel il existe un vecteur $x_0 \in H$ tel que $\|x_0\| = 1$ et $\overline{\mathrm{vect}}\,\big\{T^kT^{*l}x_0;\ k,l \geqslant 0\big\} = H$. On peut également supposer que $\|T\| \leqslant 1$. Le point clé est alors le lemme suivant.

Lemme 4.7. *Soit $T \in \mathcal{L}(H)$ un opérateur normal tel que $\|T\| \leqslant 1$, et soit $x_0 \in H$ vérifiant $\|x_0\| = 1$. Il existe une mesure de probabilité borélienne m sur le disque unité fermé $\overline{\mathbb{D}} = \{z \in \mathbb{C};\ |z| \leqslant 1\}$ telle que*

$$\forall k,l \in \mathbb{N}\ :\ \langle T^kT^{*l}x_0, x_0\rangle = \int_{\overline{\mathbb{D}}} z^k\bar z^l\, dm(z).$$

Si on tient ce lemme pour acquis, la preuve du Théorème spectral n'est plus qu'un "formalité". On vérifie d'abord que si $(c_{k,l})_{k,l\geqslant 0}$ est une suite double de nombre complexes dont tous les termes sont nuls sauf un nombre fini, alors

$$\Big\|\sum_{k,l} c_{k,l}T^kT^{*l}x_0\Big\|^2 = \int_{\overline{\mathbb{D}}}\Big|\sum_{k,l} c_{k,l}z^k\bar z^l\Big|^2\, dm(z).$$

Cela permet de définir sans ambiguité une isométrie $V : \mathrm{vect}\,\big\{T^kT^{*l}x_0;\ k,l \geqslant 0\big\} \to L^2(\overline{\mathbb{D}}, m)$ en posant

$$V(T^kT^{*l}x_0)(z) := z^k\bar z^l \qquad \text{pour tous } k,l \geqslant 0.$$

L'image de V contient toutes les fonctions f de la forme $f(z) = P(z,\bar z)$ où $P \in \mathbb{C}[X,Y]$. Comme ces fonctions sont denses dans $L^2(\overline{\mathbb{D}}, m)$ et comme $\mathrm{vect}\,\big\{T^kT^{*l}x_0;\ k,l \geqslant 0\big\}$ est dense dans H, on peut donc prolonger V en un opérateur unitaire $U : H \to L^2(\overline{\mathbb{D}}, m)$; et on vérifie sans difficulté qu'on a $M_{\mathbf{z}}U = UT$.
Ainsi, T est unitairement équivalent à l'opérateur $M_{\mathbf{z}}$ agissant sur $L^2(\overline{\mathbb{D}}, m)$. Ce n'est pas tout à fait ce qu'on voulait, car il faudrait avoir une mesure m sur $\sigma(T)$.
Cependant, on a $\sigma(T) = \sigma(M_{\mathbf{z}}) = \sigma_{L^\infty(\overline{\mathbb{D}},m)}(\mathbf{z})$. Donc la fonction $\mathbf{z}$ est m-pp à valeurs dans $\sigma(T)$. Autrement dit : $m\big(\overline{\mathbb{D}}\backslash\sigma(T)\big) = 0$. On peut donc considérer que m est en fait une mesure sur $\sigma(T)$ et que $M_{\mathbf{z}}$ agit en fait sur $L^2(\sigma(T), m)$; et cette fois, on a exactement ce qu'on voulait.

Preuve du Lemme 4.7. On va appliquer le Théorème de Herglotz (Théorème 5.17 du Chapitre 5) au semi-groupe $S := \mathbb{N}\times\mathbb{N}$ muni de l'involution $(k,l)^* := (l,k)$. Rappelons qu'on note $\widehat S$ l'ensemble des caractères bornés $\gamma : S \to \mathbb{C}$.
On sait que les caractères de S sont de la forme $\gamma_z(k,l) := z^k\bar z^l$ avec $z \in \mathbb{C}$; donc les éléments de $\widehat S$ sont les γ_z avec $z \in \overline{\mathbb{D}}$. L'application $z \mapsto \gamma_z$ est une bijection de $\overline{\mathbb{D}}$ sur $\widehat S$, et elle est continue car $\widehat S$ est muni de la topologie de la convergence simple. Sa réciproque est l'application $\widehat S \ni \gamma \mapsto \gamma(1,0)$, qui est également continue. Donc l'application $z \mapsto \gamma_z$ est un homéomorphisme, et on peut identifier complètement $\widehat S$ et $\overline{\mathbb{D}}$. D'après le Théorème de Herglotz, il suffit donc de vérifier que la fonction

$\phi : \mathbb{N}\times\mathbb{N}\to\mathbb{C}$ définie par $\phi(k,l) := \langle T^kT^{*l}x_0, x_0\rangle$ est de type positif, bornée et telle que $\phi(0,0)=1$.

On a $\phi(0,0) = \langle x_0, x_0\rangle = \|x_0\|^2 = 1$; et ϕ est bornée car $\|T\|\leqslant 1$. La preuve que ϕ est de type positif “se fait toute seule” : si $(k_1,l_1),\dots,(k_d,l_d)\in\mathbb{N}\times\mathbb{N}$ et $c_1,\dots,c_d\in\mathbb{C}$, alors

$$\begin{aligned}\sum_{i,j=1}^d c_i\overline{c_j}\,\phi\big((k_i,l_i)+(k_l,l_j)^*\big) &= \sum_{i,j=1}^d c_i\overline{c_j}\,\langle T^{k_i+l_j}T^{*(l_i+k_j)}x_0, x_0\rangle\\ &= \sum_{i,j=1}^d \langle c_iT^{k_i}T^{*l_i}x_0, c_jT^{k_j}T^{*l_j}x_0\rangle\\ &= \left\|\sum_{n=1}^d c_nT^{k_n}T^{*l_n}x_0\right\|^2 \geqslant 0.\end{aligned}$$

□

La preuve qu'on vient d'esquisser n'a vraiment rien d'élémentaire : elle repose sur le Théorème de Herglotz, lequel a été démontré en utilisant le Théorème de Krein-Milman et en se fatiguant un peu.

On va voir maintenant que dans le cas d'un opérateur T *auto-adjoint*, on peut donner une preuve du Théorème spectral nettement plus économique en moyens.

La seule chose à établir est l'analogue “auto-adjoint” du Lemme 4.7 ; à savoir le résultat suivant.

Lemme 4.8. *Soit $T\in\mathcal{L}(H)$ un opérateur auto-adjoint tel que $\|T\|\leqslant 1$, et soit $x_0\in H$ vérifiant $\|x_0\|=1$. Il existe une mesure de probabilité borélienne m sur $[-1,1]$ telle que*

$$\forall k\in\mathbb{N}\ :\ \langle T^kx_0, x_0\rangle = \int_{-1}^1 x^k\,dm(x).$$

Démonstration. On distingue deux cas.

Cas 1. $\dim(H)<\infty$.

Dans ce cas, par le Théorème spectral fini-dimensionnel, on sait que T est diagonalisable en base orthonormée. Soit $(e_1,\dots,e_d)$ une b.o.n. de H formée de vecteurs propres pour T, et notons $\lambda_1,\dots,\lambda_d$ les valeurs propres correspondantes. Comme T est auto-adjoint et $\|T\|\leqslant 1$, on a $-1\leqslant\lambda_i\leqslant 1$ pour $i=1,\dots,d$.

Si $k\in\mathbb{N}$, alors $T^ke_i = \lambda_i^ke_i$ pour $i=1\dots,d$. Donc

$$T^kx_0 = T^k\left(\sum_{i=1}^d\langle x_0,e_i\rangle\,e_i\right) = \sum_{i=1}^d\langle x_0,e_i\rangle\lambda_i^ke_i,$$

et donc

$$\langle T^kx_0,x_0\rangle = \sum_{i=1}^d|\langle x_0,e_i\rangle|^2\,\lambda_i^k.$$

Soit alors m la mesure positive sur $[-1,1]$ définie par

$$m := \sum_{i=1}^d|\langle x_0,e_i\rangle|^2\delta_{\lambda_i}.$$

Comme $\sum_{i=1}^{d} |\langle x_0, e_i\rangle|^2 = \|x_0\|^2 = 1$, on voit que m est une mesure de probabilité ; et par définition, on a

$$\int_{-1}^{1} x^k dm(x) = \sum_{i=1}^{d} |\langle x_0, e_i\rangle|^2 \lambda_i^k = \langle T^k x_0, x_0\rangle \quad \text{pour tout } k \in \mathbb{N}.$$

Cas 2. Cas général.

Comme l'espace de Hilbert H est supposé séparable, on peut trouver trouver une suite croissante de sous-espaces de dimension finie $(E_n)_{n\in\mathbb{N}}$ telle que $E := \bigcup_{n\in\mathbb{N}} E_n$ est dense dans H. Si on note $P_n \in \mathcal{L}(H)$ la projection orthogonale de H sur E_n, alors $P_n x \to x$ pour tout $x \in H$: en effet, c'est vrai pour $x \in E$ car la suite (E_n) est croissante, donc c'est vrai pour tout $x \in H$ car E est dense dans H et la suite (P_n) est bornée.
Pour $n \in \mathbb{N}$, posons $T_n := P_n T P_n$. Comme T et P_n sont auto-adjoints, T_n *est auto-adjoint*. De plus $T_n \xrightarrow{\texttt{SOT}} T$ car $P_n \xrightarrow{\texttt{SOT}} I$ (exo). Comme (T_n) est bornée, on en déduit par récurrence sur k que $T_n^k \xrightarrow{\texttt{SOT}} T^k$ pour tout $k \in \mathbb{N}$, puis que $T_n^k P_n \xrightarrow{\texttt{SOT}} T^k$ pour tout $k \in \mathbb{N}$.
Soit n_0 tel que $P_n x_0 \neq 0$ pour tout $n \geqslant n_0$. Si $n \geqslant n_0$ alors, comme T_n est auto-adjoint et "vit" sur l'espace de dimension finie E_n, on peut appliquer le Cas 1 à T_n et $x_n := \frac{1}{\|P_n x_0\|} P_n$: il existe donc une mesure de probabilité μ_n sur $[-1, 1]$ telle que

$$\forall k \in \mathbb{N} \ : \int_{-1}^{1} x^k d\mu_n(x) = \frac{1}{\|P_n x_0\|^2} \langle T_n^k P_n x_0, x_0\rangle.$$

Par le Théorème de Banach-Alaoglu, la suite (μ_n) possède une sous-suite (μ_{n_i}) qui converge w^* vers une mesure de probabilité m :

$$\forall f \in \mathcal{C}([-1, 1]) \ : \int_{-1}^{1} f\, d\mu_{n_i} \to \int_{-1}^{1} f\, dm.$$

Comme $\|P_{n_i} x_0\|^2 \to \|x_0\|^2 = 1$ et $T_{n_i}^k P_{n_i} x_0 \to T^k x_0$ pour tout $k \in \mathbb{N}$, on obtient ainsi le résultat souhaité (en prenant $f(x) := x^k$) :

$$\forall k \in \mathbb{N} \ : \int_{-1}^{1} x^k dm(x) = \langle T^k x_0, x_0\rangle.$$

$\square$

Maintenant que le Lemme 4.8 est démontré, on peut procéder exactement comme plus haut.

Remarque. Il est naturel de se demander pourquoi on n'a pas démontré le Lemme 4.7 de la même manière que le Lemme 4.8, en approchant l'opérateur normal T par des opérateurs fini-dimensionnels $T_n = P_n T P_n$. La raison est que ces opérateurs T_n n'ont *a priori* aucune raison d'être normaux, ce qui empêche de leur appliquer le Théorème spectral fini-dimensionnel.

5. Opérateurs de Fredhom

5.1. Définition et exemples "immédiats".

Notation. Si Z est un espace vectoriel et si $E \subsetneqq Z$ est un sous-espace vectoriel, on note $\mathrm{codim}_Z(E)$ ou simplement $\mathrm{codim}(E)$ la **codimension** de E dans Z, *i.e.* la dimension de l'espace vectoriel quotient Z/E.

DÉFINITION 5.1. *Soient X et Y des espaces de Banach. On dit qu'un opérateur $T \in \mathcal{L}(X,Y)$ est un* ***opérateur de Fredholm*** *si*

(a) $\dim \ker(T) < \infty$ *;*

(b) $\mathrm{Im}(T)$ *est fermé dans* Y *et* $\mathrm{codim}\,\mathrm{Im}(T) < \infty$.

On note $\mathrm{Fred}(X,Y)$ *l'ensemble des opérateurs de Fredholm de X dans Y, et on pose* $\mathrm{Fred}(X) := \mathrm{Fred}(X,X)$. *Si $T \in \mathrm{Fred}(X,Y)$, l'**indice de Fredholm** de T est le nombre entier* $\mathrm{ind}(T)$ *défini par*

$$\mathrm{ind}(T) := \dim \ker(T) - \mathrm{codim}\,\mathrm{Im}(T).$$

EXEMPLE 1. Si X et Y sont de dimension finie, alors tout opérateur $T \in \mathcal{L}(X,Y)$ est de Fredholm, et on a $\mathrm{ind}(T) = \dim(X) - \dim(Y)$ pour tout $T \in \mathcal{L}(X,Y)$.

Démonstration. Il est évident que tout opérateur $T \in \mathcal{L}(X,Y)$ est de Fredholm. De plus, l'algèbre linéaire nous dit que $\dim\ker(T)+\dim \mathrm{Im}(T) = \dim(X)$ et $\mathrm{codim}\,\mathrm{Im}(T) = \dim(Y) - \dim \mathrm{Im}(T)$; donc $\mathrm{ind}(T) = \big(\dim(X) - \dim \mathrm{Im}(T)\big) - \big(\dim(Y) - \dim \mathrm{Im}(T)\big) = \dim(X) - \dim(Y)$. □

EXEMPLE 2. Si $T \in \mathcal{L}(X,Y)$ est *inversible*, alors $T \in \mathrm{Fred}(X,Y)$ et $\mathrm{ind}(T) = 0$.

Démonstration. C'est évident. □

EXEMPLE 3. Le forward shift S et le backward shift B sur $\ell^2(\mathbb{N})$ sont des opérateurs de Fredholm, avec $\mathrm{ind}(S) = -1$ et $\mathrm{ind}(B) = 1$.

Démonstration. En notant $(e_n)_{n\in\mathbb{N}}$ la base canonique de $\ell^2(\mathbb{N})$, on sait que $\ker(B) = [e_0]$ et que B est surjectif ; donc B est de Fredholm et $\mathrm{ind}(B) = 1-0$. De même, S est injectif et $\mathrm{Im}(S) = [e_0]^\perp$; donc S est de Fredholm et $\mathrm{ind}(S) = 0-1$. □

EXEMPLE 4. Si $K \in \mathcal{L}(X)$ est compact et si $\lambda \neq 0$, alors $T := \lambda I + K \in \mathrm{Fred}(X)$ et $\mathrm{ind}(T) = 0$.

Démonstration. C'est une partie du Théorème 2.1. Plus précisément : le fait que T soit de Fredholm découle de la partie (4) de ce théorème ; et on a $\mathrm{ind}(T) = 0$ par la partie (6). □

REMARQUE 1. Dans la définition d'un opérateur de Fredholm, l'hypothèse que $\mathrm{Im}(T)$ est fermée est en fait superflue : si $T \in \mathcal{L}(X,Y)$ est tel que $\mathrm{codim}\,\mathrm{Im}(T) < \infty$, alors $\mathrm{Im}(T)$ est *nécessairement* fermée dans Y.

Démonstration. Quitte à remplacer X par $X/\ker(T)$, on peut supposer que T est injectif. Comme $\mathrm{codim}\,\mathrm{Im}(T) < \infty$, on peut trouver un sous-espace vectoriel $F \subseteq Y$ tel que $\dim(F) < \infty$ et $Y = \mathrm{Im}(T) \oplus F$. Alors F est fermé dans Y, donc c'est un espace de Banach, et donc $X \times F$ est un espace de Banach. Comme T est injectif et $Y = \mathrm{Im}(T) \oplus F$, l'opérateur $S : X \times F \to Y$ défini par $S(x,u) := Tx + u$ est bijectif ; donc S^{-1} est continu d'après le Théorème d'isomorphisme de Banach, et donc $\mathrm{Im}(T) = S(X \times \{0\}) = (S^{-1})^{-1}(X \times \{0\})$ est fermé dans Y puisque $X \times \{0\}$ est fermé dans $X \times F$. □

REMARQUE 2. Si $T \in \mathrm{Fred}(X,Y)$, alors $\dim\ker(T^*) < \infty$ et

$$\mathrm{ind}(T) = \dim\ker(T) - \dim\ker(T^*).$$

Démonstration. Comme $\mathrm{Im}(T)$ est supposé fermé, on peut munir $Y/\mathrm{Im}(T)$ de la norme quotient ; et comme $Y/\mathrm{Im}(T)$ est de dimension finie, on a $\dim\big(Y/\mathrm{Im}(T)\big) = \dim\big(Y/\mathrm{Im}(T)\big)^*$. Mais $\big(Y/\mathrm{Im}(T)\big)^*$ s'identifie à $\mathrm{Im}(T)^\perp = \ker(T^*)$. Donc $\dim\ker(T^*)$ est finie et égale à $\dim\big(Y/\mathrm{Im}(T)\big)$. □

Remarque 3. Un opérateur $T \in \mathcal{L}(X,Y)$ est de Fredholm si et seulement si son adjoint T^* est de Fredholm ; et dans ce cas on a $\mathrm{ind}(T^*) = -\mathrm{ind}(T)$.

Démonstration. Supposons que T soit de Fredholm. Par la Remarque 2, on sait déjà que $\dim\ker(T^*) < \infty$. Comme $\mathrm{Im}(T)$ est fermée dans Y, on peut affirmer que $\mathrm{Im}(T^*)$ est fermée dans Y^*, et plus précisément que $\mathrm{Im}(T^*) = \ker(T)^\perp$. Donc $X^*/\mathrm{Im}(T^*) = X^*/\ker(T)^\perp$, et donc l'espace $\big(X^*/\mathrm{Im}(T^*)\big)^*$ s'identifie à $\big(\ker(T)^\perp\big)^\perp \subseteq X^{**}$. Mais en considérant $\ker(T)$ comme un sous-espace vectoriel de X^{**}, on a $\big(\ker(T)^\perp\big)^\perp = \big(\ker(T)_\perp\big)^\perp = \overline{\ker(T)}^{w^{**}}$, et donc $\big(\ker(T)^\perp\big)^\perp = \ker(T)$ car $\ker(T)$, étant de dimension finie, est w^{**}- fermé dans X^{**} (exo). Ainsi, $\big(X^*/\mathrm{Im}(T^*)\big)^*$ s'identifie à $\ker(T)$. En particulier $\big(X^*/\mathrm{Im}(T^*)\big)^*$ est de dimension finie égale à $\dim\ker(T)$. Donc $X^*/\mathrm{Im}(T^*)$ est lui aussi de dimension finie égale à $\dim\ker(T)$. Ainsi, on a montré que T^* est un opérateur de Fredholm, avec $\mathrm{ind}(T^*) = \dim\ker(T^*) - \dim\ker(T) = -\mathrm{ind}(T)$.
Supposons maintenant que T^* soit de Fredholm. Alors T^{**} est de Fredholm par ce qui précède, donc $\dim\ker(T^{**}) < \infty$. Mais $\ker(T^{**}) \supseteq \ker(T)$ puisque $T^{**}_{|X} = T$, donc $\dim\ker(T) < \infty$. De plus, $\mathrm{Im}(T)$ est fermé dans Y car $\mathrm{Im}(T^*)$ est fermé dans X^* ; et $\dim\ker(T^*) < \infty$, donc $\mathrm{codim}\,\mathrm{Im}(T) < \infty$ puisque $\ker(T^*)$ s'identifie à $\big(Y/\mathrm{Im}(T)\big)^*$. Donc T est de Fredholm. □

Exercice. Montrer que si T est un opérateur de Fredholm, alors $\ker(T^{**}) = \ker(T)$.

5.2. Propriétés de base. Dans ce qui suit, X et Y sont des espaces de Banach.

Lemme 5.2. *Pour $T \in \mathcal{L}(X,Y)$, les propriétés suivantes sont équivalentes.*

(a) *T est de Fredholm.*

(b) *Il existe des sous-espaces fermés $X_0, X_1 \subseteq X$ et $Y_0, Y_1 \subseteq Y$ tels que*
- *$X = X_0 \oplus X_1$ et $Y = Y_0 \oplus Y_1$;*
- *$\dim(X_1) < \infty$ et $\dim(Y_1) < \infty$;*
- *T admet, relativement aux décompositions $X = X_0 \oplus X_1$ et $Y = Y_0 \oplus Y_1$, l'écriture matricielle*
$$T = \begin{pmatrix} T_0 & 0 \\ 0 & 0 \end{pmatrix},$$
où $T_0 \in \mathcal{L}(X_0, Y_0)$ est inversible.

(c) *Il existe des sous-espaces fermés $E \subseteq X$ et $M \subseteq Y$ tels que*
- *$\mathrm{codim}(E) < \infty$ et $\dim(F) < \infty$;*
- *l'opérateur $T_{|E} \oplus I_M : E \times M \to Y$ défini par $T_{|E} \oplus I_M(x,v) := Tx + v$ est un isomorphisme de $E \times M$ sur Y.*

(d) *Il existe des sous-espaces vectoriels de codimension finie $E \subseteq X$ et $F \subseteq Y$ tels que T envoie bijectivement E sur F.*

De plus, si (b) *a lieu avec X_0, X_1, Y_0, Y_1, alors $X_1 = \ker(T)$, $Y_0 = \mathrm{Im}(T)$ et $\mathrm{ind}(T) = \dim(X_1) - \dim(Y_1)$; et si* (d) *a lieu avec E et F, alors $\mathrm{ind}(T) = \mathrm{codim}(E) - \mathrm{codim}(F)$.*

Démonstration. L'implication (a) $\implies$ (b) est claire : si T est de Fredholm, on peut écrire $X = X_0 \oplus \ker(T)$ pour un certain sous-espace fermé X_0 car $\dim\ker(T) < \infty$, et on peut écrire $Y = \operatorname{Im}(T) \oplus Y_1$ pour un certain sous-espace Y_1 tel que $\dim(Y_1) < \infty$ puisque $\operatorname{codim}\operatorname{Im}(T) < \infty$; donc (b) est visiblement satisfaite avec $X_0, X_1 := \ker(T), Y_0 := \operatorname{Im}(T)$ et Y_1.
De plus, si (b) est satisfaite avec X_0, X_1, Y_0, Y_1, alors on a nécessairement $X_1 = \ker(T)$ et $Y_0 = \operatorname{Im}(T)$ puisque T_0 est inversible (exo) ; et donc $\operatorname{ind}(T) = \dim(X_1) - \dim(Y_1)$.
Les implication (b) $\implies$ (c) et $(c) \implies (d)$ sont claires également : si (b) est satisfaite, alors (c) est satisfaite avec $E := Y_0$ et $M := Y_1$; et si (c) est satisfaite avec E et M, alors (d) est satisfaite avec E et $F := T(E)$, qui est bien de codimension finie, égale à $\dim(M)$.
Supposons maintenant (d) vérifiée, et montrons que T est de Fredholm avec $\operatorname{ind}(T) = \operatorname{codim}(E) - \operatorname{codim}(F)$.
Comme $\ker(T) \cap E = \{0\}$ et $\operatorname{Im}(T) \supseteq F$, on a $\dim\ker(T) < \infty$ et $\operatorname{codim}\operatorname{Im}(\mathrm{T}) < \infty$; donc T est un opérateur de Fredholm.
Comme $\ker(T) \cap E = \{0\}$, on peut trouver un sous-espace $M \subseteq X$ tel que $X = E \oplus M \oplus \ker(T)$. Alors $\operatorname{codim}(E) = \dim\ker(T) + \dim(M)$. De plus, $\operatorname{Im}(T) = T(E) \oplus T(M) = F \oplus T(M)$, et $\dim\big(T(M)\big) = \dim(M)$ car $T_{|M}$ est injectif ; donc $\operatorname{codim}(F) = \operatorname{codim}\operatorname{Im}(\mathrm{T}) + \dim(M)$. Ainsi, on voit que $\operatorname{ind}(T) = \operatorname{codim}(E) - \operatorname{codim}(F)$.

□

Remarque. La preuve du lemme montre que dans (d), on peut prendre E et F *fermés* (vérifier). On voit donc qu'un opérateur $T \in \mathcal{L}(X, Y)$ est de Fredholm si et seulement si il existe un sous-espace fermé $E \subseteq X$ de codimension finie tel que $T_{|E} : E \to Y$ est un plongement et $T(E)$ est de codimension finie dans Y.

Exercice. Montrer directement l'implication $(a) \implies (d)$.

Le résultat suivant est habituellement appelé le **Théorème d'Atkinson**.

Théorème 5.3. *Pour $T \in \mathcal{L}(X, Y)$, les propriétés suivantes sont équivalentes.*
(1) *T est de Fredholm.*
(2) *Il existe des opérateurs $S_1, S_2 \in \mathcal{L}(Y, X)$ tels que $TS_1 - I_Y$ et $S_2T - I_X$ sont compacts.*
(2') *Il existe un opérateur $S \in \mathcal{L}(Y, X)$ tel que $TS - I_Y$ et $ST - I_X$ sont compacts.*
(3) *Il existe un opérateur $S \in \mathcal{L}(Y, X)$ tel que $TS - I_Y$ et $ST - I_X$ sont de rang fini.*

Démonstration. (1) $\implies$ (3). Supposons que T soit de Fredholm, et soient $X_0, X_1 \subseteq X$ et $Y_0, Y_1 \subseteq Y$ comme dans le Lemme 5.2. Relativement aux décompositions $X = X_0 \oplus X_1$ et $Y = Y_0 \oplus Y_1$, l'opérateur T s'écrit

$$T = \begin{pmatrix} T_0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{avec } T_0 \in \mathcal{L}(X_0, Y_0) \text{ inversible.}$$

Si on note $S : Y \to X$ l'opérateur défini par

$$S := \begin{pmatrix} T_0^{-1} & 0 \\ 0 & 0 \end{pmatrix},$$

alors

$$ST = \begin{pmatrix} I_{X_0} & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{et} \qquad TS = \begin{pmatrix} I_{Y_0} & 0 \\ 0 & 0 \end{pmatrix}.$$

Donc $ST - I_X$ et $TS - I_Y$ sont de rang fini puisque $\dim(X_1) < \infty$ et $\dim(Y_1) < \infty$.
Les implications (3) $\implies$ (2') $\implies$ (2) sont évidente.
(2) $\implies$ (1). Supposons (2) vérifiée avec $S_1, S_2 \in \mathcal{L}(Y, X)$. Alors $S_1T = I_X - K_1$ et $TS_2 = I_Y - K_2$ avec K_1 et K_2 compacts. Comme $\ker(T) \subseteq \ker(S_1T)$ et $\mathrm{Im}(T) \supseteq \mathrm{Im}(TS_2)$, on a donc

$$\dim\ker(T) \leqslant \dim\ker(I_X - K_1) < \infty \qquad \text{et} \qquad \operatorname{codim}\mathrm{Im}(T) \leqslant \operatorname{codim}\mathrm{Im}(I_Y - K_2) < \infty.$$

Donc T est de Fredholm. □

Corollaire 5.4. *Si $T \in \mathcal{L}(X, Y)$ est de Fredholm et si $K \in \mathcal{L}(X, Y)$ est compact, alors $T + K$ est de Fredholm.*

Démonstration. Si S_1, S_2 vérifient (2) pour T, alors ils le vérifient également pour $T + K$ (exo). □

Le théorème suivant est fondamental.

Théorème 5.5. *L'ensemble des opérateurs de Fredholm est un ouvert de $\mathcal{L}(X, Y)$, et $\mathrm{ind}(T)$ dépend continûment de $T \in \mathrm{Fred}(X, Y)$.*

Démonstration. Il s'agit de voir que si $T \in \mathrm{Fred}(X, Y)$, alors tout opérateur T' suffisamment proche de T est de Fredholm avec $\mathrm{ind}(T') = \mathrm{ind}(T)$. Mais ceci est évident par le Lemme 5.2 : si T vérifie la propriété (c) de ce lemme avec des sous-espaces $E \subseteq X$ et $M \subseteq Y$, alors tout opérateur T' proche de T est tel que $T'_{|E} \oplus I_M$ est un isomorphisme de $E \times M$ sur Y car l'ensemble des opérateurs inversibles est un ouvert de $\mathcal{L}(E \times M, Y)$; et donc T' vérifie (c) *avec les mêmes E et M*. □

Corollaire 5.6. *Si T_0 et T_1 sont deux opérateurs de Fredholm et appartiennent à la même composante connexe de* $\mathrm{Fred}(X, Y)$, *alors* $\mathrm{ind}(T_1) = \mathrm{ind}(T_0)$.

Démonstration. C'est évident puisque ind est à valeurs dans $\mathbb{Z}$. □

Corollaire 5.7. *L'indice de Fredholm est* stable par perturbations compactes *: si $T \in \mathcal{L}(X, Y)$ est de Fredholm et si $K \in \mathcal{L}(X, Y)$ est compact, alors* $\mathrm{ind}(T + K) = \mathrm{ind}(T)$.

Démonstration. On sait que $T + K$ est de Fredholm d'après le théorème d'Atkinson. Si on définit $\gamma : [0, 1] \to \mathcal{L}(X, Y)$ par $\gamma(s) := T + sK$, alors l'application γ est continue, et à valeurs dans $\mathrm{Fred}(X, Y)$ car $T \in \mathrm{Fred}(X, Y)$ et sK est compact pour tout $s \in [0, 1]$. Donc $T = \gamma(0)$ et $T + K = \gamma(1)$ sont dans la même composante connexe de $\mathrm{Fred}(X, Y)$, et donc $\mathrm{ind}(T + K) = \mathrm{ind}(T)$. □

Remarque. On obtient ainsi une nouvelle preuve du fait que si $K \in \mathcal{L}(X)$ est un opérateur compact, alors $\dim\ker(\lambda I + K) = \operatorname{codim}\mathrm{Im}(\lambda I + K)$ pour tout $\lambda \neq 0$: par continuité de l'indice, on a $\mathrm{ind}(\lambda I + K) = \mathrm{ind}(\lambda I) = 0$.

Corollaire 5.8. *Un opérateur $T \in \mathcal{L}(X, Y)$ est de la forme $T = J + K$ avec J inversible et K compact si et seulement si $T \in \mathrm{Fred}(X, Y)$ et $\mathrm{ind}(T) = 0$; et si tel est le cas, alors T est en fait de la forme $T = J + R$ avec J inversible et R de rang fini.*

Démonstration. Si T est de la forme $J + K$, alors T est de Fredholm et $\mathrm{ind}(T) = \mathrm{ind}(J) = 0$ par le corollaire précédent.
Inversement, supposons que $T \in \mathrm{Fred}(X, Y)$ et $\mathrm{ind}(T) = 0$. Alors on peut écrire $X = E \oplus \ker(T)$ et $Y = \mathrm{Im}(T) \oplus F$ avec $\dim(F) = \dim\ker(T) < \infty$. Soit $P : X \to \ker(T)$

la projection associée à la décomposition $X = E \oplus \ker(T)$, et soit $L : \ker(T) \to Y$ un opérateur injectif tel que $\mathrm{Im}(L) = F$. Comme $T_{|E}$ est un isomorphisme de E sur $\mathrm{Im}(T)$ et comme $E = \ker(P)$, l'opérateur $J := T + LP \in \mathcal{L}(X, Y)$ est inversible (exo) ; et $R := T - J = -LP$ est de rang fini. $\square$

Pour terminer cette section, on va démontrer le **Théorème du produit** pour l'indice de Fredholm. C'est un résultat important... dont on ne fera aucun usage.

THÉORÈME 5.9. *Soient X, Y, Z des espaces de Banach. Si $A \in \mathrm{Fred}(X, Y)$ et $B \in \mathrm{Fred}(Y, Z)$, alors $BA \in \mathrm{Fred}(X, Z)$ et $\mathrm{ind}(BA) = \mathrm{ind}(B) + \mathrm{ind}(A)$.*

Démonstration. Par le Lemme 5.2, on peut trouver des sous-espaces de codimension finie $E \subseteq X$, $F, E' \subseteq Y$ et $F' \subseteq Z$ tels que A envoie bijectivement E sur F et B envoie bijectivement E' sur F'. Alors BA envoie bijectivement $\widetilde{E} := (A_{|E})^{-1}(E' \cap F)$ sur $\widetilde{F} := B(E' \cap F)$. De plus, $\widetilde{E}$ est de codimension finie dans E car $E' \cap F$ est de codimension finie dans F et $(A_{|E})^{-1} : F \to E$ est un isomorphisme ; donc $\widetilde{E}$ est de codimension finie dans X puisque $\mathrm{codim}(E) < \infty$. De même, $\widetilde{F}$ est de codimension finie dans Z. Donc, à nouveau par le Lemme 5.2, BA est de Fredholm avec $\mathrm{ind}(BA) = \mathrm{codim}(\widetilde{E}) - \mathrm{codim}(\widetilde{F})$. Pour montrer que $\mathrm{ind}(BA) = \mathrm{ind}(B) + \mathrm{ind}(A)$, introduisons des sous-espaces (de dimension finie) $M, M' \subseteq Y$ tels que

$$(E' \cap F) \oplus M = F \qquad \text{et} \qquad (E' \cap F) \oplus M' = E'.$$

Comme $E = (A_{|E})^{-1}(F)$ et que $(A_{|E})^{-1} : F \to E$ est un isomorphisme, on a alors $E = (A_{|E})^{-1}(E' \cap F) \oplus (A_{|E})^{-1}(M)$, *i.e.* $E = \widetilde{E} \oplus (A_{|E})^{-1}(M)$; et de même, $F' = \widetilde{F} \oplus B(M')$. Donc $\mathrm{codim}(\widetilde{E}) = \mathrm{codim}(E) + \dim\big((A_{|E})^{-1}(M)\big) = \mathrm{codim}(E) + \dim(M)$ et $\mathrm{codim}(\widetilde{F}) = \mathrm{codim}(F') + \dim\big(B(M')\big) = \mathrm{codim}(F') + \dim(M')$, de sorte que

$$\mathrm{ind}(BA) = \mathrm{codim}(E) + \dim(M) - \mathrm{codim}(F') - \dim(M').$$

Comme de plus $\dim(M) = \mathrm{codim}(E' \cap F) - \mathrm{codim}(F)$ et $\dim(M') = \mathrm{codim}(E' \cap F) - \mathrm{codim}(E')$, on obtient ainsi

$$\begin{aligned} \mathrm{ind}(BA) &= \mathrm{codim}(E) - \mathrm{codim}(F) - \mathrm{codim}(F') + \mathrm{codim}(E') \\ &= \mathrm{ind}(A) + \mathrm{ind}(B). \end{aligned}$$

$\square$

Autre preuve de $\mathrm{ind}(BA) = \mathrm{ind}(B) + \mathrm{ind}(A)$. La preuve qui suit est un peu plus longue, mais donne plus d'informations. Pour simplifier les écritures, on posera pour tout opérateur T :

$$\alpha(T) := \dim \ker(T) \qquad \text{et} \qquad \beta(T) := \mathrm{codim}\, \mathrm{Im}(T).$$

Si T est un opérateur de Fredholm, on a donc

$$\mathrm{ind}(T) = \alpha(T) - \beta(T).$$

Une autre notation : si H est un espace vectoriel et si K est un sous-espace vectoriel de H, on note $[H : K]$ la dimension de l'espace quotient H/K :

$$[H : K] = \dim(H/K).$$

On a donc $\dim(H) = \dim(K) + [H : K]$ (que les dimensions soient finies ou non).

Fait 1. On a les identités suivantes :

$$\alpha(BA) = \alpha(A) + \dim\big(\mathrm{Im}(A) \cap \ker(B)\big);$$

$$\beta(BA) = [Y : \mathrm{Im}(A) + \ker(B)] + \beta(B).$$

Preuve du Fait 1. Pour la 1ère identité, on observe qu'on a $\mathrm{Im}(A) \cap \ker(B) = A\big(\ker(BA)\big) = \mathrm{Im}\big(A_{|\ker(BA)}\big)$, et $\ker(A) = \ker\big(A_{|\ker(BA)}\big)$ puisque $\ker(A) \subseteq \ker(BA)$. Donc $\mathrm{Im}(A) \cap \ker(B)$ est isomorphe à $\ker(BA)/\ker(A)$, ce qui donne la formule souhaitée.

Pour la 2ème identité, on introduit un supplémentaire U de $\ker(B)$ dans $\mathrm{Im}(A)+\ker(B)$, et un supplémentaire V de $\mathrm{Im}(A) + \ker(B)$ dans Y. Ainsi :

$$\mathrm{Im}(A) + \ker(B) = U \oplus \ker(B),$$

$$Y = U \oplus \ker(B) \oplus V.$$

Comme $\mathrm{Im}(BA) = B\big(\mathrm{Im}(A) + \ker(B)\big) = B(U)$, on a $\beta(BA) = [Z : B(U)]$. Mais on a aussi $\mathrm{Im}(B) = B(U) \oplus B(V)$, et donc $[Z : B(U)] = \beta(B) + \dim\big(B(V)\big)$. Comme $\dim\big(B(V)\big) = \dim(V) = [Y : \mathrm{Im}(A) + \ker(B)]$ puisque $B_{|V}$ est injectif, on a donc bien montré que $\beta(BA) = [Y : \mathrm{Im}(A) + \ker(B)] + \beta(B)$. □

Fait 2. Si M et N sont des sous-espaces vectoriels de Y, alors

$$\dim(N) = \dim(M \cap N) + [N : M \cap N] \quad \text{et}$$

$$[Y : M] = [N : M \cap N] + [Y : M + N].$$

En particulier, si $\dim(N) < \infty$ et $[Y : M] < \infty$, alors

$$\dim(N) - [Y : M] = \dim(M \cap N) - [Y : M + N].$$

Preuve du Fait 2. La 1ère formule est évidente : c'est l'identité $\dim(H) = \dim(K) + [H : K]$ avec $H := N$ et $K := M \cap N$.

Pour démontrer la 2ème formule, on introduit un supplémentaire U de $M \cap N$ dans M, et un supplémentaire U' de $M \cap N$ dans N :

$$M = U \oplus (M \cap N) \qquad \text{et} \qquad N = (M \cap N) \oplus U'.$$

On a alors (exo)

$$M + N = U \oplus (M \cap N) \oplus U' = M \oplus U';$$

et donc $[Y : M] = [Y : M + N] + \dim(U') = [Y : M + N] + [N : M \cap N]$. □

Même si ce n'est peut-être pas immédiatement apparent, la preuve est maintenant terminée. Si on pose $p := \dim\big(\mathrm{Im(A)} \cap \ker(B)\big)$ et $q := [Y : \mathrm{Im}(A) + \ker(B)]$, et si on applique le Fait 1 puis le Fait 2 avec $M := \mathrm{Im}(A)$ et $N := \ker(B)$, on obtient

$$\begin{cases} \alpha(BA) = \alpha(A) + p, \\ \beta(BA) = q + \beta(B), \\ \alpha(B) - \beta(A) = p - q; \end{cases}$$

d'où on déduit immédiatement que $\alpha(BA) - \beta(BA) = \alpha(A) + \alpha(B) - \beta(A) - \beta(B)$, *i.e.* $\mathrm{ind}(BA) = \mathrm{ind}(B) + \mathrm{ind}(A)$. □

Encore une autre preuve. Pour montrer que $\mathrm{ind}(BA) = \mathrm{ind}(B) + \mathrm{ind}(A)$, on utilise une “astuce” difficilement compréhensible de prime abord, mais assez naturelle *a posteriori* : on introduit les opérateurs $M_0 := A \oplus B : X \oplus Y \to Y \oplus Z$ et $M := (BA) \oplus (-I_Y) : X \oplus Y \to Y \oplus Z$. Matriciellement, on a ainsi

$$M_0 = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \qquad \text{et} \qquad M = \begin{pmatrix} 0 & -I_Y \\ BA & 0 \end{pmatrix}.$$

Il est assez clair que M_0 est de Fredholm et $\mathrm{ind}(M_0) = \mathrm{ind}(A) + \mathrm{ind}(B)$. De même, comme BA est de Fredholm et $-I_Y$ inversible, on voit que M est de Fredholm et $\mathrm{ind}(M) = \mathrm{ind}(BA)$. Il suffit donc de montrer qu'on a $\mathrm{ind}(M) = \mathrm{ind}(M_0)$.
Maintenant, on peut jouer avec des matrices 2×2 et observer qu'on a d'une part

$$M_0 = \begin{pmatrix} I_Y & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I_Y & 0 \\ 0 & I_Y \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I_Y \end{pmatrix},$$

et d'autre part

$$M = \begin{pmatrix} I_Y & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} 0 & -I_Y \\ I_Y & 0 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I_Y \end{pmatrix}.$$

Cela incite à considérer, pour $\theta \in \mathbb{R}$, l'opérateur $M_\theta \in \mathcal{L}(X \oplus Y, Y \oplus Z)$ défini par

$$M_\theta := \begin{pmatrix} I_Y & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} \cos\theta\, I_Y & -\sin\theta\, I_Y \\ \sin\theta\, I_Y & \cos\theta\, I_Y \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I_Y \end{pmatrix}.$$

On a ainsi $M = M_{\pi/2}$ et $M_0 = \dots M_0$.
Pour tout $\theta \in \mathbb{R}$, l'opérateur $R_\theta := \begin{pmatrix} \cos\theta\, I_Y & -\sin\theta\, I_Y \\ \sin\theta\, I_Y & \cos\theta\, I_Y \end{pmatrix} \in \mathcal{L}(Y \oplus Y)$ est inversible, d'inverse $R_{-\theta}$ (exo) ; en particulier, R_θ est de Fredholm. Donc M_θ apparait comme le produit de 3 opérateurs de Fredholm, et par conséquent M_θ est de Fredholm. Comme l'application $\theta \mapsto M_\theta$ est visiblement continue, on en déduit par connexité que $\mathrm{ind}(M_\theta)$ ne dépend pas de θ. Donc $\mathrm{ind}(BA) = \mathrm{ind}(M_{\pi/2}) = \mathrm{ind}(M_0) = \mathrm{ind}(A) + \mathrm{ind}(B)$. $\square$

Exemple. Stabilité et continuité de l'indice de Fredholm.

On peut utiliser le Théorème du produit pour redémontrer la stabilité par perturbations compactes et la continuité de l'indice de Fredholm.

Montrons d'abord la stabilité par perturbations compactes. Soit $T \in \mathrm{Fred}(X, Y)$, et soit $K \in \mathcal{K}(X, Y)$. Par le Théorème d'Atkinson, on peut trouver $S \in \mathcal{L}(Y, X)$ tel que $TS - I_Y$ et $ST - I_X$ sont compacts ; et S est de Fredholm, à nouveau par le Théorème d'Atkinson. Posons $L := ST - I_X$. Comme L est compact, on a $\mathrm{ind}(ST) = \mathrm{ind}(I_X + L) = 0$; et donc $\mathrm{ind}(S) = -\mathrm{ind}(T)$ par le Théorème du produit. De même, $\mathrm{ind}\big((T + K)S\big) = \mathrm{ind}(I_X + L + KS) = 0$ car $L + KS$ est compact ; donc $\mathrm{ind}(T + K) = -\mathrm{ind}(S) = \mathrm{ind}(T)$.

Montrons maintenant la continuité. Soit $T \in \mathrm{Fred}(X, Y)$: on voudrait montrer que pour tout opérateur $T' \in \mathcal{L}(X, Y)$ assez proche de T, on a $\mathrm{ind}(T') = \mathrm{ind}(T)$. Comme plus haut, soit $S \in \mathrm{Fred}(Y, X)$ tel que $L := ST - I_X$ est compact. On a $\mathrm{ind}(ST) = \mathrm{ind}(I_X + L) = 0$, donc $\mathrm{ind}(T) = -\mathrm{ind}(S)$. Si $T' \in \mathcal{L}(X, Y)$, alors $ST' = ST + S(T' - T) = L + I_X + S(T' - T)$. Par ailleurs, si T' est assez proche de T, précisément si $\|S\|\,\|T' - T\| < 1$, alors $J := I_X + S(T' - T)$ est inversible. Comme $ST' = J + L$ et que L est compact, on en déduit (par stabilité de l'indice de Fredholm) que $\mathrm{ind}(ST') = \mathrm{ind}(J) = 0$, et donc $\mathrm{ind}(T') = -\mathrm{ind}(S) = \mathrm{ind}(T)$.

5.3. Deux exemples "non immédiats". Dans cette section, on donne d'une part un exemple "concret" d'utilisation de la notion d'opérateur de Fredholm, et d'autre part un exemple non trivial de calcul d'indice.

Exemple 1. Soit $\phi : [0,1] \to \mathbb{R}$ une fonction continue $\geqslant 0$. Pour toute fonction continue $g : [0,1] \to \mathbb{C}$, il existe une unique fonction $f : [0,1] \to \mathbb{C}$ de classe $\mathcal{C}^2$ telle que $f'' - \phi f = g$ et $f(0) = 0 = f(1)$.

Démonstration. Munissons l'espace $\mathcal{C}^2([0,1])$ de la norme $\|\cdot\|_{\mathcal{C}^2}$ définie par $\|f\|_{\mathcal{C}^2} := \|f\|_\infty + \|f'\|_\infty + \|f''\|_\infty$, qui en fait un espace de Banach. Soit

$$X := \bigl\{f \in \mathcal{C}^2([0,1]);\ f(0) = 0 = f(1)\bigr\}.$$

C'est un sous-espace fermé de $\mathcal{C}^2([0,1])$, donc un espace de Banach. Il s'agit de montrer que l'opérateur $T : X \to \mathcal{C}([0,1])$ défini par $Tf := f'' - \phi f$ est bijectif.
On a $T = \Delta + K$, où $\Delta f := f''$ et $Kf := -\phi f$. Il est assez clair par définition de X que $\Delta_{|X} : X \to \mathcal{C}([0,1])$ est bijectif, *i.e.* inversible (exo). Quant à K, il n'est pas difficile de voir que c'est un opérateur compact : en effet, l'injection canonique $J : \mathcal{C}^2([0,1]) \hookrightarrow \mathcal{C}([0,1])$ est un opérateur compact (exo), et l'opérateur de multiplication $M_\phi : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ est borné, donc $K = -M_\phi J_{|X}$ est compact. Donc $T = \Delta + K$ est un opérateur de Fredholm et $\mathrm{ind}(T) = 0$. Pour conclure, il suffit donc de montrer que T est injectif.
Le point clé est que l'opérateur Δ est "négatif", ce qui se voit en intégrant par parties :

$$\forall f \in X\ :\ \int_0^1 f''\,\bar{f} = \bigl[f'\,\bar{f}\,\bigr]_0^1 - \int_0^1 |f'|^2 = -\int_0^1 |f'|^2 \leqslant 0.$$

On en déduit que si $f \in X$ vérifie $Tf = 0$, *i.e.* $f'' = \phi f$, alors $\int_0^1 \phi|f|^2 \leqslant 0$. Mais la fonction $\phi|f|^2$ est $\geqslant 0$ et continue. Donc $\phi|f|^2 = 0$; autrement dit $\phi f = 0$. Donc $f'' = \phi f = 0$, et donc $f = 0$ puisque $f(0) = 0 = f(1)$. On a donc bien montré que l'opérateur T est injectif. □

Exemple 2. Soient $L^2 := L^2(\mathbb{T})$ et $H^2 := H^2(\mathbb{T})$. Pour toute $\phi \in L^\infty = L^\infty(\mathbb{T})$, on note $T_\phi : H^2 \to H^2$ l'opérateur défini par $T_\phi := P_+ M_\phi$, où $P_+ : L^2 \to H^2$ est la projection orthogonale de L^2 sur H^2 et $M_\phi : H^2 \to L^2$ est la restriction à H^2 de l'opérateur de multiplication par ϕ. (On dit que T_ϕ est l'**opérateur de Toeplitz** associé à ϕ.) Si $\phi : \mathbb{T} \to \mathbb{C}$ est une fonction continue telle que $\phi(z) \neq 0$ pour tout $z \in \mathbb{T}$, alors T_ϕ est un opérateur de Fredholm, et $\mathrm{ind}(T_\phi) = -\deg(\phi)$, où $\deg(\phi)$ est le **degré** de ϕ, *i.e.* l'indice du chemin fermé $\gamma : [0, 2\pi] \to \mathbb{C}^*$ défini par $\gamma(t) := \phi(e^{it})$ par rapport à 0.

Preuve abrégée. Pour $k \in \mathbb{Z}$, on notera $\mathbf{z}^k$ la fonction $\mathbb{T} \ni z \mapsto z^k$, qui appartient à $\mathcal{C}(\mathbb{T}) \subseteq L^\infty$. La fonction $\mathbf{z}^k$ ne s'annule pas sur $\mathbb{T}$, et "on sait bien" que $\deg(\mathbf{z}^k) = k$.

Fait 1. $T_{\mathbf{z}^k}$ est de Fredholm pour tout $k \in \mathbb{Z}$, avec $\mathrm{ind}(T_{\mathbf{z}^k}) = -k$.

Preuve du Fait 1. Si $k \geqslant 0$, alors $M_{\mathbf{z}^k}$ envoie H^2 dans H^2, donc $T_{\mathbf{z}^k} = M_{\mathbf{z}^k}$. De plus, $M_{\mathbf{z}^k}$ est une isométrie de H^2 dans H^2, et $\mathrm{Im}(M_{\mathbf{z}^k}) = \overline{\mathrm{Vect}\,\{\mathbf{z}^n;\ n \geqslant k\}}^{H^2}$ (exo), qui est de codimension k dans H^2. Donc $T_{\mathbf{z}^k}$ est de Fredholm avec $\mathrm{ind}(T_{\mathbf{z}^k}) = 0 - k = -k$. Si $k < 0$, alors $T_{\mathbf{z}^k}$ est surjectif et $\ker(T_{\mathbf{z}^k}) = \mathrm{Vect}\,\{\mathbf{z}^n;\ 0 \leqslant n < -k\}$ (exo) ; donc $T_{\mathbf{z}^k}$ est de Fredholm avec $\mathrm{ind}(T_{\mathbf{z}^k}) = -k - 0 = -k$. □

Fait 2. Si $\phi \in \mathcal{C}(\mathbb{T})$, alors l'opérateur $T_\phi - M_\phi$ est compact.

Preuve du Fait 2. On considère ici T_ϕ comme un opérateur de H^2 dans L^2. Dans ce qui suit, on pose $A_\phi := T_\phi - M_\phi$.
Comme les polynômes trigonométriques sont denses dans $\mathcal{C}(\mathbb{T})$, on peut trouver une suite de polynômes trigonométriques (p_n) telle que $\|p_n - \phi\|_\infty \to 0$. Alors $\|A_{p_n} - A_\phi\| \to 0$ car $\|A_\psi\| \leqslant 2\|\psi\|_\infty$ pour toute $\psi \in L^\infty$ (micro-exo). Pour montrer que A_ϕ est compact, il suffit donc de montrer que les A_{p_n} le sont. Ainsi, on s'est ramené à montrer que $T_p - M_p$ est compact pour tout polynôme trigonométrique p.
Si on écrit $p = \sum_{k=-N}^{N} c_k \mathbf{z}^k$, on voit que $p\mathbf{z}^n \in H^2$ pour tout $n \geqslant N$, et donc $T_p \mathbf{z}^n = p\mathbf{z}^n$. Donc $\ker(T_p - M_p)$ contient le sous-espace fermé de H^2 engendré par les $\mathbf{z}^n$, $n \geqslant N$. Par conséquent, $\ker(T_p - M_p)$ est de codimension finie dans H^2 ; et donc $T_p - M_p$ est *de rang fini*, donc compact. □

Soit maintenant $\phi \in \mathcal{C}(\mathbb{T})$ ne s'annulant pas sur $\mathbb{T}$, et montrons que T_ϕ est de Fredholm. Comme ϕ ne s'annule pas, la fonction $1/\phi$ est bien définie et appartient à $\mathcal{C}(\mathbb{T})$. Par le Fait 2, on peut écrire $T_{1/\phi}T_\phi = T_{1/\phi}(M_\phi + K)$ où $K : H^2 \to L^2$ est un opérateur compact. Comme $T_{1/\phi}M_\phi = P_+ M_{1/\phi}M_\phi = I_{H^2}$, on en déduit que $T_{1/\phi}T_\phi = I_{H^2} + L$, où $L := T_{1/\phi}K$ est compact. On montre de même que $T_\phi T_{1/\phi} = I_{H^2} + L'$ avec L' compact, en appliquant le Fait 2 à $1/\phi$. Donc T_ϕ est de Fredholm d'après le Théorème 5.3.
Montrons pour finir que si $\phi \in \mathcal{C}(\mathbb{T})$ ne s'annule pas sur $\mathbb{T}$, alors $\operatorname{ind}(T_\phi) = -\deg(\phi)$. Posons $k := \deg(\phi)$. D'après les propriétés classiques du degré, la fonction ϕ est "homotope dans $\mathcal{C}(\mathbb{T}, \mathbb{C}^*)$" à la fonction $\mathbf{z}^k$: il existe une application continue $H : [0,1] \times \mathbb{T} \to \mathbb{C}^*$ telle que, en posant $\phi_t(z) := H(t,z)$, on ait $\phi_0 = \phi$ et $\phi_1 = \mathbf{z}^k$. Par ce qui précède, tous les opérateurs T_{ϕ_t} sont de Fredholm puisque les ϕ_t ne s'annulent pas, et l'application $t \mapsto T_{\phi_t}$ est continue de $[0,1]$ dans $\mathcal{L}(H^2)$ par continuité de H (exo). Par continuité de l'indice et d'après le Fait 1, on a donc $\operatorname{ind}(T_\phi) = \operatorname{ind}(T_{\phi_0}) = \operatorname{ind}(T_{\phi_1}) = \operatorname{ind}(T_{\mathbf{z}^k}) = -k$. □

5.4. Spectre essentiel. Dans ce qui suit, X est un espace de Banach complexe *de dimension infinie*. On rappelle qu'on note $\mathcal{K}(X)$ l'ensemble de tous les opérateurs compacts $K : X \to X$. On sait que $\mathcal{K}(X)$ est un sous-espace vectoriel fermé de $\mathcal{L}(X)$; et on sait aussi que $\mathcal{K}(X)$ est une *idéal bilatère* de l'algèbre $\mathcal{L}(X)$: si $K \in \mathcal{K}(X)$, alors $TK \in \mathcal{K}(X)$ et $KT \in \mathcal{L}(X)$ pour tout $T \in \mathcal{L}(X)$. De plus, $\mathcal{K}(X) \neq \mathcal{L}(X)$ car $\dim(X) = \infty$. Donc l'espace vectoriel quotient

$$\mathfrak{C}(X) := \mathcal{L}(X)/\mathcal{K}(X)$$

est muni canoniquement d'une structure d'algèbre de Banach unitaire. (Il faut vérifier que la norme quotient est bien sous-multiplicative et que $\|\mathbf{1}_{\mathfrak{C}(X)}\| = 1$; ce qui n'est pas difficile.) On dit que $\mathfrak{C}(X)$ est l'**algèbre de Calkin** de l'espace de Banach X.

Notation. On note $\pi : \mathcal{L}(X) \to \mathfrak{C}(X)$ l'application quotient canonique ; et si $T \in \mathcal{L}(X)$, on pose $[T] := \pi(T)$.

Avec ces notations, on peut reformuler une partie du Théorème d'Atkinson 5.3 comme suit : si $T \in \mathcal{L}(X)$, alors

$$T \text{ est de Fredholm} \iff [T] \text{ est inversible dans } \mathfrak{C}(X).$$

Définition 5.10. *Si $T \in \mathcal{L}(X)$, le **spectre essentiel** de T est le spectre de $[T] = \pi(T)$ dans l'algèbre $\mathfrak{C}(X)$. On le note $\sigma_e(T)$. Ainsi :*

$$\lambda \in \sigma_e(T) \iff [T - \lambda I] \text{ n'est pas inversible dans } \mathfrak{C}(X).$$

Remarque 1. Par le Théorème d'Atkinson, on voit que

$$\lambda \in \sigma_e(T) \iff T - \lambda I \text{ n'est pas un opérateur de Fredholm.}$$

Remarque 2. Par la théorie générale des algèbres de Banach, $\sigma_e(T)$ est un compact de $\mathbb{C}$ et $\sigma_e(T) \neq \varnothing$. De plus, on a toujours $\sigma_e(T) \subseteq \sigma(T)$.

Démonstration. On a $\sigma_e(T) = \sigma(\pi(T)) \subseteq \sigma(T)$ car $\pi : \mathcal{L}(X) \to \mathfrak{C}(X)$ est un homomorphisme d'algèbres. □

Remarque 3. Le spectre essentiel est "invariant par perturbations compactes" : on a $\sigma_e(T) = \sigma_e(T + K)$ pour tout opérateur compact $K \in \mathcal{L}(X)$.

Démonstration. C'est évident puisque $[T + K] = [T]$. □

Remarque 4. On a $\sigma_e(T^*) = \sigma_e(T)$. Si X est un espace de Hilbert H et si T^* est l'adjoint hilbertien de T, alors $\sigma_e(T^*) = \{\bar{\lambda};\ \lambda \in \sigma_e(T)\}$.

Démonstration. On a vu qu'un opérateur R est de Fredholm si et seulement si R^* est de Fredholm. □

Exemple 1. Si $K \in \mathcal{L}(X)$ est compact, alors $\sigma_e(K) = \{0\}$.

Démonstration. C'est évident puisque $[K] = 0$. □

Exemple 2. Si B et S sont les shifts canoniques sur $\ell^2(\mathbb{N})$, alors $\sigma_e(B) = \mathbb{T} = \sigma_e(S)$.

Démonstration. Comme $B^* = S$, il suffit de montrer que $\sigma_e(S) = \mathbb{T}$.
Par l'Exemple 1.33, on sait que $\sigma(S) = \overline{\mathbb{D}}$; donc $\sigma_e(S) \subseteq \overline{\mathbb{D}}$.
Montrons que $\sigma_e(S) \subseteq \mathbb{T}$; autrement dit, que si $\lambda \in \mathbb{D}$, alors $S - \lambda I$ est un opérateur de Fredholm. Comme S est une isométrie, on a $\|(S-\lambda I)x\| \geqslant \|Sx\| - |\lambda|\,\|x\| = (1 - |\lambda|)\,\|x\|$ pour tout $x \in \ell^2$, donc $S - \lambda I$ est un plongement. De plus, $\ker\big((S-\lambda I)^*\big) = \ker(B^* - \bar{\lambda} I)$ est de dimension 1 par l'Exemple 1.33. Donc $S - \lambda I$ est de Fredholm avec $\mathrm{ind}(S - \lambda I) = 0 - 1 = -1$.
Montrons maintenant que $\sigma_e(S)$ contient $\mathbb{T}$. Soit $\lambda \in \mathbb{T}$ quelconque. L'opérateur $S - \lambda I$ est injectif car $\sigma_p(S) = \varnothing$, et il est aussi à image dense car $(S - \lambda I)^* = B - \bar{\lambda} I$ est injectif ($\lambda \notin \mathbb{D}$). Comme $S - \lambda I$ n'est pas inversible puisque $\lambda \in \sigma(\mathbb{T})$, on en déduit que $S - \lambda I$ n'est pas à image fermée ; et donc $S - \lambda I$ n'est pas de Fredholm. □

Proposition 5.11. *Soit $T \in \mathcal{L}(X)$, et soit $\lambda \in \sigma(T)$ un point isolé de $\sigma(T)$. Soit $X = E_\lambda \oplus F$ la décomposition de Riesz associée à la décomposition $\sigma(T) = \{\lambda\} \cup \sigma(T)\backslash\{\lambda\}$. Alors $\lambda \in \sigma_e(T) \iff \dim(E_\lambda) = \infty$. En particulier, si λ n'est pas valeur propre de T, alors $\lambda \in \sigma_e(T)$.*

Démonstration. Relativement à la décomposition $X = E_\lambda \oplus F$, l'opérateur $T - \lambda I$ s'écrit matriciellement

$$T - \lambda I = \begin{pmatrix} A_\lambda & 0 \\ 0 & B \end{pmatrix}$$

où $A_\lambda \in \mathcal{L}(E_\lambda)$ vérifie $\sigma(A_\lambda) = \{0\}$ et $B \in \mathcal{L}(F)$ est inversible. Comme B est inversible, $T - \lambda I$ est de Fredholm si et seulement si A_λ est de Fredholm (exo). Si $\dim(E_\lambda) < \infty$, cela est certainement vérifié. À l'inverse, si $\dim(E_\lambda) = \infty$, alors $\varnothing \neq \sigma_e(A_\lambda) \subseteq \sigma(A_\lambda) = \{0\}$, donc $\sigma_e(A_\lambda) = \{0\}$, et donc A_λ n'est pas de Fredholm. Ainsi, on voit que $T - \lambda I$ est

de Fredholm si et seulement si $\dim(E_\lambda) < \infty$; autrement dit, $\lambda \in \sigma_e(T)$ si et seulement si $\dim(X_\lambda) = \infty$.
Si $\dim(E_\lambda) < \infty$, alors λ est valeur propre de T car $\lambda \in \sigma(T_{|E_\lambda})$. Donc, par ce qui précède : si λ n'est pas valeur propre de T alors $\lambda \in \sigma_e(T)$. $\square$

PROPOSITION 5.12. *Soit $T \in \mathcal{L}(X)$, et soit $\lambda \in \mathbb{C}$. On suppose qu'il existe une suite $(x_n)_{n\in\mathbb{N}} \subseteq X$ telle que $x_n \xrightarrow{w} 0$, $\|x_n\| = 1$ pour tout n et $\|Tx_n - \lambda x_n\| \to 0$. Alors $\lambda \in \sigma_e(T)$.*

Démonstration. On doit montrer que l'opérateur $R := T - \lambda I$ n'est pas de Fredholm. Si $\dim \ker(R) = \infty$, il n'y a rien à faire ; donc on suppose que $\dim \ker(R) < \infty$. Soit $E \subseteq X$ un sous-espace fermé tel que $X = \ker(R) \oplus E$, et soit $P : X \to \ker(R)$ la projection sur $\ker(R)$ associée à cette décomposition. Alors P est de rang fini, donc c'est un opérateur compact. Comme $x_n \xrightarrow{w} 0$, on en déduit que $\|Px_n\| \to 0$. Donc, si on pose $z_n := (I - P)x_n = x_n - Px_n$, alors $\|z_n\| \geqslant 1/2$ pour n assez grand puisque $\|x_n\| = 1$. De plus, $z_n \in E$ par définition et $\|Rz_n\| = \|Rx_n\| \to 0$. On en déduit que $R_{|E}$ n'est pas un plongement ; et donc $R_{|E}$ n'est pas à image fermée puisque $R_{|E}$ est injectif. Comme $R(E) = \operatorname{Im}(R)$, cela signifie que R n'est pas à image fermée ; donc R n'est pas de Fredholm. $\square$

PROPOSITION 5.13. *Soit $T \in \mathcal{L}(X)$. Si $\lambda \in \sigma(T)\backslash\sigma_e(T)$, alors ou bien λ est un point isolé de $\sigma(T)$, ou bien λ est un point intérieur à $\sigma(T)$, i.e. il existe un voisinage ouvert W de λ dans $\mathbb{C}$ tel que $W \subseteq \sigma(T)$.*

La preuve de cette proposition repose sur le lemme suivant.

LEMME 5.14. *Si $A \in \mathcal{L}(X)$ est un opérateur de Fredholm, alors il existe un disque ouvert $D(0, \varepsilon) \subseteq \mathbb{C}$ tel que :* $\dim \ker(A - \mu I)$ *et* $\operatorname{codim} \operatorname{Im}(A - \mu I)$ *ne dépendent pas de μ si $\mu \in D(0, \varepsilon)$ et $\mu \neq 0$.*

Preuve du Lemme 5.14. Pour $n \in \mathbb{N}$, soit $Z_n := \operatorname{Im}(A^n)$; et soit aussi $Z_\infty := \bigcap_{n\in\mathbb{N}} Z_n$. Comme $A^n \in \operatorname{Fred}(X)$ pour tout $n \in \mathbb{N}$ puisque $A \in \operatorname{Fred}(X)$, on sait que les Z_n sont des sous-espaces fermés de X ; donc Z_∞ est un sous-espace fermé de X. De plus, la suite (Z_n) est décroissante.

FAIT 1. On a $A(Z_\infty) = Z_\infty$.

Preuve du Fait 1. Il est clair que $A(Z_\infty) \subseteq Z_\infty$ (micro-exo). Inversement, soit $z \in Z_\infty$ quelconque. Alors $E_n := \{x \in Z_n ;\ Ax = z\}$ est non vide pour tout $n \in \mathbb{N}$ car $z \in Z_{n+1} = A(Z_n)$. De plus, les E_n sont des sous-espaces affines de X, ils sont de dimension finie car $\{x \in X ;\ Ax = z\}$ est de dimension finie (c'est un translaté de $\ker(A)$), et la suite (E_n) est décroissante. Donc la suite (E_n) est stationnaire (exo) ; et comme tous les E_n sont non vides, on en déduit que $\bigcap_{n\in\mathbb{N}} E_n \neq \varnothing$. Si on choisit $x \in \bigcap_{n\in\mathbb{N}} E_n$, alors $x \in \bigcap_{n\in\mathbb{N}} Z_n = Z_\infty$ et $Ax = z$; donc $z \in A(Z_\infty)$. $\square$

FAIT 2. Si Z est un espace de Banach, alors l'ensemble des opérateurs surjectifs $S : Z \to Z$ est un ouvert de $\mathcal{L}(Z)$.

Preuve du Fait 2. On sait qu'un opérateur $S : Z \to Z$ est surjectif si et seulement si $S^* : Z^* \to Z^*$ est un plongement. De plus, il n'est pas difficile de montrer (exo) que l'ensemble de tous les plongements $R : Z^* \to Z^*$ est un ouvert de $\mathcal{L}(Z^*)$. D'où le résultat par continuité de l'application $S \mapsto S^*$. $\square$

Posons $B := A_{|Z_\infty} : Z_\infty \to Z_\infty$. Par le Fait 1, B est un opérateur surjectif. De plus, on a $\ker(B) \subseteq \ker(A)$ et donc $\dim\ker(B) < \infty$ puisque A est de Fredholm. Donc B est un opérateur de Fredholm.
Par le Fait 2, l'ensemble des opérateurs $S \in \mathcal{L}(Z_\infty)$ qui sont de Fredholm et surjectifs est un ouvert de $\mathcal{L}(Z_\infty)$. Donc on peut trouver $\varepsilon > 0$ tel que $B - \mu I_{Z_\infty}$ est de Fredholm et surjectif pour tout $\mu \in D(0,\varepsilon)$, avec de plus $\operatorname{ind}(B - \mu I_{Z_\infty}) = \operatorname{ind}(B)$. On a donc $\dim\ker(B - \mu I_{Z_\infty}) = \dim\ker(B)$ pour tout $\mu \in D(0,\varepsilon)$ puisque $B - \mu I_{Z_\infty}$ et B sont surjectifs. De plus, quitte à diminuer ε, on peut aussi supposer que $A - \mu I \in \operatorname{Fred}(X)$ pour tout $\mu \in D(0,\varepsilon)$, avec $\operatorname{ind}(A - \mu I) = \operatorname{ind}(A)$.

Fait 3. Si $\mu \neq 0$, alors $\ker(A - \mu I) = \ker(B - \mu I_{Z_\infty})$.

Démonstration. Il s'agit de montrer que si $x \in \ker(A - \mu I)$, alors $x \in Z_\infty$. Comme $Ax = \mu x$, on a $A^n x = \mu^n x$ pour tout $n \in \mathbb{N}$, et donc $x = \mu^{-n} A^n x = A^n(\mu^{-n} x) \in Z_n$. Donc en effet $x \in Z_\infty = \bigcap_{n\in\mathbb{N}} Z_n$. □

On peut maintenant finir la preuve du lemme. Par le Fait 3 et le choix de ε, on a $\dim\ker(A-\mu I) = \dim\ker(B)$ pour tout $\mu \in D(0,\varepsilon)$ tel que $\mu \neq 0$; donc $\dim\ker(A-\mu I)$ ne dépend pas de μ si $\mu \in D(0,\varepsilon)$ et $\mu \neq 0$. De plus $\operatorname{ind}(A - \mu I)$ ne dépend pas non plus de $\mu \in D(0,\varepsilon)$, et donc $\operatorname{codim}\operatorname{Im}(A - \mu I)$ ne dépend pas de μ si $\mu \in D(0,\varepsilon)$ et $\mu \neq 0$.

□

Preuve de la Proposition 5.13. Supposons que λ ne soit pas un point isolé de $\sigma(T)$. Soit $A := T - \lambda I$. Comme on suppose que $\lambda \notin \sigma_e(T)$, l'opérateur A est de Fredholm. Soit $\varepsilon > 0$ donné par le Lemme 5.13 : par définition, $\dim\ker(T - wI)$ et $\operatorname{codim}\operatorname{Im}(T - wI)$ ne dépendent pas de w si $w \in D(\lambda,\varepsilon)$ et $w \neq \lambda$. Posons $W := D(\lambda,\varepsilon)$. Comme λ n'est pas un point isolé de $\sigma(T)$, on peut trouver $w_0 \in W$ tel que $w_0 \neq \lambda$ et $w_0 \in \sigma(T)$. Alors ou bien $\dim\ker(T - w_0 I) > 0$, ou bien $\operatorname{codim}\ker(T - w_0 I) > 0$. Par définition de W et comme $w_0 \neq \lambda$, on a donc ou bien $\dim\ker(T - wI) > 0$ pour tout $w \in W\backslash\{\lambda\}$, ou bien $\operatorname{codim}\ker(T - wI) > 0$ pour tout $w \in W\backslash\{\lambda\}$. En particulier $W\backslash\{\lambda\} \subseteq \sigma(T)$; et donc $W \subseteq \sigma(T)$ puisque $\lambda \in \sigma(T)$. □

Corollaire 5.15. *Soit $T \in \mathcal{L}(X)$, et soit $\lambda \in \mathbb{C}$. Si $\lambda \in \partial\sigma(T)$ et si λ n'est pas un point isolé de $\sigma(T)$, alors $\lambda \in \sigma_e(T)$.*

Démonstration. C'est évident par la proposition. □

Corollaire 5.16. *Soit $T \in \mathcal{L}(X)$. Si Ω est une composante connexe de $\mathbb{C}\backslash\sigma_e(T)$, alors ou bien $\Omega \subseteq \sigma(T)$, ou bien tous les points de $\Omega \cap \sigma(T)$ sont des points isolés de $\sigma(T)$.*

Démonstration. Posons

$$\Omega_1 := \big\{\lambda \in \Omega;\ \exists W \text{ voisinage de } \lambda\ :\ W \subseteq \sigma(T)\big\},$$
$$\Omega_2 := \big\{\lambda \in \Omega;\ \exists W \text{ voisinage de } \lambda\ :\ W \cap \sigma(T) \subseteq \{\lambda\}\big\}.$$

Il est évident que Ω_1 est un ouvert de Ω, et facile de vérifier que Ω_2 est également un ouvert de Ω (exo ; il faut utiliser le fait que $\sigma(T)$ est un fermé de $\mathbb{C}$). De plus, il est évident que $\Omega_1 \cap \Omega_2 = \varnothing$. Enfin, on a $\Omega = \Omega_1 \cup \Omega_2$ par la Proposition 5.13. Comme Ω est connexe, on a donc ou bien $\Omega = \Omega_1$, ou bien $\Omega = \Omega_2$. Ceci démontre le résultat souhaité : si $\Omega = \Omega_1$, alors $\Omega \subseteq \sigma(T)$; et si $\Omega = \Omega_2$, alors tous les points de $\Omega \cap \sigma(T)$ sont des points isolés de $\sigma(T)$. □

Corollaire 5.17. *Si $T \in \mathcal{L}(X)$ et si $\mathbb{C}\backslash\sigma_e(T)$ est connexe, alors tous les points de $\sigma(T)\backslash\sigma_e(T)$ sont des points isolés de $\sigma(T)$ et des valeurs propres de T.*

Démonstration. On ne peut pas avoir $\mathbb{C}\backslash\sigma_e(T) \subseteq \sigma(T)$ car $\sigma(T)$ est un ensemble borné et $\mathbb{C}\backslash\sigma_e(T)$ est non borné ; donc le résultat est clair par le corollaire précédent et la Proposition 5.11. □

Voici pour finir deux résultats propres aux opérateurs normaux.

Proposition 5.18. *Soit H un espace de Hilbert complexe séparable (de dimension infinie), et soit $T \in \mathcal{L}(H)$ un opérateur* normal. *Pour $\lambda \in \sigma(T)$, les choses suivantes sont équivalentes.*

(1) *$\lambda \in \sigma_e(T)$.*
(2) *λ n'est pas un point isolé de $\sigma(T)$, ou bien $\dim\ker(T-\lambda I) = \infty$.*
(3) *Il existe une suite orthonormale $(x_n)_{n\in\mathbb{N}} \subseteq H$ telle que $\|Tx_n - \lambda x_n\| \to 0$.*

Démonstration. (1) $\Longrightarrow$ (2) Supposons que $\lambda \in \sigma_e(T)$, et que λ soit un point isolé de $\sigma(T)$: il faut voir que $\dim\ker(T-\lambda I) = \infty$. Soit $H = E_\lambda \oplus F$ la décomposition de Riesz associée à la décomposition $\sigma(T) = \{\lambda\} \cup \sigma(T)\backslash\{\lambda\}$. Par la Proposition 5.11, on a $\dim E_\lambda = \infty$. Mais comme T est normal, on a $E_\lambda = \ker(T-\lambda I)$ par le Corollaire 3.29. Donc $\dim\ker(T-\lambda I) = \infty$.
(2) $\Longrightarrow$ (3) Supposons (2) vérifiée. Si $\dim\ker(T-\lambda I) = \infty$, alors (3) est certainement vérifiée : il suffit de prendre pour (x_n) une suite orthonormale dans $\ker(T-\lambda I)$. Supposons maintenant que λ ne soit pas un point isolé de $\sigma(T)$. Soit $(\lambda_n) \subseteq \sigma(T)$ une suite de points deux à deux distincts telle que $\lambda_n \to \lambda$. Choisissons une suite (ε_n) de nombres réels > 0 tendant vers 0 telle que les disques $D_n := D(\lambda_n, \varepsilon_n)$ soient deux à deux disjoints, et posons $P_n := \mathbf{1}_{D_n}(T)$. Par l'Exemple 4.6, les P_n sont des projections orthogonales non-nulles et $\|(T-\lambda_n I)P_n\| \leqslant \varepsilon_n$ pour tout $n \in \mathbb{N}$. De plus, on a $P_nP_m = 0$ si $n \neq m$ car $D_n \cap D_m = \varnothing$. Donc les sous-espaces $E_n := \mathrm{Im}(P_n)$ sont deux à deux orthogonaux. Pour tout $n \in \mathbb{N}$, on peut choisir $x_n \in E_n$ tel que $\|x_n\| = 1$. Alors la suite (x_n) est orthonormale, et $\|Tx_n - \lambda_n x_n\| = \|(T-\lambda_n I)P_n x_n\| \leqslant \varepsilon_n$ pour tout $n \in \mathbb{N}$. Donc $\|Tx_n - \lambda_n x_n\| \to 0$, et donc $\|Tx_n - \lambda x_n\| \to 0$ puisque $\lambda_n \to \lambda$. Ainsi, on voit que (3) est vérifiée.
L'implication (3) $\Longrightarrow$ (1) découle de la Proposition 5.12. □

Proposition 5.19. *Soit H un espace de Hilbert complexe (de dimension infinie). Si $T \in \mathcal{L}(H)$ est un opérateur* normal, *alors $\sigma_e(T) = \bigcap_{K\in\mathcal{K}(H)} \sigma(T+K)$.*

Démonstration. On a $\sigma_e(T) = \sigma_e(T+K) \subseteq \sigma(T+K)$ pour tout opérateur compact K ; donc $\sigma_e(T) \subseteq \bigcap_{K\in\mathcal{K}(H)} \sigma(T+K)$, sans hypothèse sur T.
Inversement, soit $\lambda \in \bigcap_{K\in\mathcal{K}(H)} \sigma(T+K)$. Alors $T - \lambda I + K$ n'est inversible pour aucun opérateur compact K ; autrement dit, on ne peut pas écrire $T - \lambda I = J + K$ avec J inversible et K compact. Par le Corollaire 5.8, cela signifie que $T - \lambda I$ n'est pas "de Fredholm avec indice 0". Mais $R := T - \lambda I$ est normal, donc $\ker(R) = \ker(R^*)$. Donc, si R était de Fredholm, on aurait nécessairement $\mathrm{ind}(R) = 0$. Par conséquent, $R = T - \lambda I$ *n'est pas* de Fredholm, *i.e.* $\lambda \in \sigma_e(T)$. □