Prépa. Agrég. écrit d'Analyse, séance n° 12, Mercredi 21 Février 2001.

Compacité

Métrisabilité

Définition. Un espace topologique X est dit *métrisable* s'il existe une distance d sur X telle que la topologie de X coïncide avec la topologie définie sur X par la distance d.

Exercice. Soit φ une fonction sous-additive croissante, continue, nulle en 0 et seulement en 0 ; si d est une distance sur X, montrer que $\varphi \circ d$ est encore une distance sur X, qui définit la même topologie que d. Montrer que toute fonction concave sur $[0, +\infty[$, croissante continue, nulle en zéro, non identiquement nulle, vérifie les conditions précédentes. Vérifier que $t \to \min(t, 1)$ est dans cette classe de fonctions.

Si X topologique est métrisable, on peut donc prendre si on veut une distance *bornée* pour définir la topologie, en remplaçant d par $\min(d, 1)$.

Définition. Soient X un espace topologique et $\mathcal{B}$ une famille d'ouverts de X ; on dit que $\mathcal{B}$ est une *base* de la topologie de X si tout ouvert U de X est réunion de certains ouverts de la famille $\mathcal{B}$. Cela revient exactement à dire que pour tout ouvert U de X

$$\forall x \in \mathrm{U},\ \exists \mathrm{B} \in \mathcal{B},\quad x \in \mathrm{B} \subset \mathrm{U}.$$

La topologie d'un espace métrique (X, d) est un exemple de topologie définie à partir d'une base : on prend pour base de la topologie la famille de toutes les boules ouvertes de la distance d.

Distance et produit dénombrable

Sur le produit $\mathrm{X} = \prod_{i \in \mathrm{I}} \mathrm{X}_i$ d'une famille $(\mathrm{X}_i)_{i \in \mathrm{I}}$ d'espaces topologiques on définit la *topologie produit* de la façon suivante : un ensemble V est un *ouvert élémentaire* s'il existe un sous-ensemble fini $\mathrm{J} \subset \mathrm{I}$ et une famille $(\mathrm{U}_j)_{j \in \mathrm{J}}$ où U_j est ouvert de X_j pour chaque $j \in \mathrm{J}$, de sorte que

$$\mathrm{V} = \{x = (x_i)_{i \in \mathrm{I}} : \forall j \in \mathrm{J},\ x_j \in \mathrm{U}_j\}.$$

On dit ensuite que $\mathrm{U} \subset \mathrm{X}$ est ouvert pour la topologie produit s'il est réunion d'ouverts élémentaires, c'est à dire que pour tout $x \in \mathrm{U}$ il existe V ouvert élémentaire tel que $x \in \mathrm{V} \subset \mathrm{U}$.

La topologie de la convergence simple sur l'espace $\mathcal{F}(\mathrm{Y}, \mathbb{R})$ de toutes les fonctions réelles $f : \mathrm{Y} \to \mathbb{R}$ est identique à la topologie produit de l'ensemble $\mathbb{R}^{\mathrm{Y}}$.

Dans le cas d'un produit dénombrable $\mathrm{X} = \prod_{n=0}^{+\infty} \mathrm{X}_n$ on a le résultat suivant :

Proposition. *Si chaque* X_n *est métrisable, l'espace produit* $\mathrm{X} = \prod_{n \in \mathbb{N}} \mathrm{X}_n$ *est métrisable.*

Démonstration. Pour chaque entier $n \ge 0$, on peut définir la topologie de X_n par une distance $d_n \le 1$. La fonction d sur $\mathrm{X} \times \mathrm{X}$ (où $\mathrm{X} = \prod_{n \in \mathbb{N}} \mathrm{X}_n$) définie par

$$\forall x, y \in \mathrm{X},\quad d(x, y) = \sum_{n=0}^{+\infty} 2^{-n} d_n(x_n, y_n)$$

est finie, continue sur $\mathrm{X} \times \mathrm{X}$ (série normalement convergente de fonctions continues), donc d définit une topologie moins fine que la topologie produit. Inversement il suffit de montrer que tout ouvert élémentaire V est ouvert pour d ; un tel V est donné par un nombre fini d'ouverts $\mathrm{U}_j \subset \mathrm{X}_j$ pour $j = 0, \ldots, \mathrm{N}$ et V est l'ensemble des $y \in \mathrm{X}$ tels que $y_j \in \mathrm{U}_j$ pour tout $j = 0, \ldots, \mathrm{N}$. Si on a $x = (x_n)_{n \in \mathbb{N}} \in \mathrm{V}$, on a $x_j \in \mathrm{U}_j$ pour $j = 0, \ldots, \mathrm{N}$ et on peut trouver $\varepsilon > 0$ tel que $d_j(x_j, y_j) < 2^j \varepsilon$ implique $y_j \in \mathrm{U}_j$ pour tout $j = 0, \ldots, \mathrm{N}$; alors $d(x, y) < \varepsilon$ implique $y \in \mathrm{V}$. On a ainsi montré que V est réunion de d-boules.

Espaces métrisables séparables

Exercice. L'espace métrique (X, d) est **non** séparable si et seulement s'il existe $\varepsilon > 0$ et une famille *non dénombrable* $(x_i)_{i \in \mathrm{I}}$ de points de X à distances mutuelles $\geq \varepsilon$.

Proposition. *Si chaque espace topologique X_n est séparable, l'espace produit $\prod_{n \in \mathbb{N}} \mathrm{X}_n$ est séparable.*

Démonstration. Pour chaque entier $n \geq 0$, soit D_n un ensemble dénombrable dense dans X_n, et soit D l'ensemble des $x = (x_n)$ tels que pour un certain N (dépendant de $x \in \mathrm{D}$) on ait $x_j \in \mathrm{D}_j$ pour $0 \leq j \leq \mathrm{N}$ et $x_n = d_n^{(0)}$ pour tout $n > \mathrm{N}$.

Théorème. *Soit X un espace topologique métrisable ; l'espace X est séparable si et seulement si sa topologie admet une base dénombrable.*

Corollaire. *Tous les sous-espaces sont alors séparables.*

Démonstration. On va montrer que si X est séparable, il existe une suite d'ouverts (V_n) telle que tout ouvert soit réunion de certains de ces V_n. Soient (r_q) une énumération des rationnels positifs, (x_p) une suite dense dans X et $\mathrm{V}_{p,q} = \mathrm{B}_d(x_p, r_q)$; si U est ouvert dans X et $x \in \mathrm{U}$, on trouve d'abord $r > 0$ tel que $\mathrm{B}(x, r) \subset \mathrm{U}$, puis r_q tel que $0 < r_q < r/2$ et x_p tel que $d(x, x_p) < r_q$; on a alors $x \in \mathrm{B}(x_p, r_q) \subset \mathrm{B}(x, r) \subset \mathrm{U}$.

Réciproquement, si (V_n) est une base dénombrable pour la topologie de X et si on prend $x_n \in \mathrm{V}_n$ (pour chaque $\mathrm{V}_n \neq \emptyset$) il est clair que la suite (x_n) sera dense dans X.

Définition : espace polonais. Un espace topologique X est dit *polonais* s'il est séparable et si sa topologie peut être définie par une distance d qui rend l'espace X complet.

Exercice : tout espace polonais est image continue de $\mathbb{N}^{\mathbb{N}}$; tout borélien de $[0, 1]$ est image bijective d'un polonais ; voir exercices 12.1 et 12.2.

Compacité

Définition. On dit que l'espace topologique X est *compact* s'il est séparé et s'il a la propriété de sous-recouvrement ouvert fini : pour tout recouvrement $(\mathrm{U}_i)_{i \in \mathrm{I}}$ de X par des ouverts, il existe un sous-ensemble fini $\mathrm{J} \subset \mathrm{I}$ tel que $\mathrm{X} = \bigcup_{j \in \mathrm{J}} \mathrm{U}_j$.

Ceci équivaut à dire que X est séparé et qu'il a la *propriété d'intersection finie* pour les fermés, c'est à dire que toute famille $\mathcal{F}$ de fermés de X dont toutes les sous-familles finies ont une intersection non vide, a elle-même une intersection non vide. Ceci équivaut encore à dire que X est séparé et qu'on a intersection non vide pour les ordonnés filtrants décroissants de fermés non vides.

Théorème : théorème de Tykhonov. *Tout produit $\mathrm{X} = \prod_{i \in \mathrm{I}} \mathrm{X}_i$ d'une famille $(\mathrm{X}_i)_{i \in \mathrm{I}}$ de compacts est compact.*

Les curieux pourront lire deux démonstrations de ce théorème dans un document en annexe.

Exercice. Contre-exemple à l'extraction de sous-suites convergentes ; désignons par K le disque unité fermé de $\mathbb{C}$; montrer que la suite (f_n) dans le compact $\mathrm{K}^{\mathbb{R}}$ définie par $f_n(t) = \mathrm{e}^{int}$ n'admet aucune sous-suite simplement convergente.

Espace métrique vérifiant Bolzano

Soit (X, d) un espace métrique dans lequel toute suite admet des sous-suites convergentes. On définit le processus suivant dans X (supposé non vide) : on choisit $x_0 \in \mathrm{X}$; ensuite, tant que $\mathrm{U}_n = \bigcup_{j=0}^{n} \mathrm{B}(x_j, \varepsilon)$ est différent de X on choisit $x_{n+1} \notin \mathrm{U}_n$; si ce processus peut continuer indéfiniment on aura trouvé une suite $(x_n) \subset \mathrm{X}$ telle que $d(x_m, x_n) \geq \varepsilon$ pour tous $m \neq n$; une telle suite ne peut pas avoir de sous-suite convergente, ce qui contredit notre hypothèse sur X. Il en résulte que pour tout $\varepsilon > 0$, l'espace X peut être recouvert par une famille finie de boules de rayon ε. Il en résulte en particulier que X est séparable.

Exercice. Packing et recouvrement ; on considère un espace métrique (X, d) et une partie compacte $\mathrm{K} \subset \mathrm{X}$. Pour tout $\varepsilon > 0$ désignons par $\mathrm{N_P}(\mathrm{K}, \varepsilon)$ le max du nombre N de points $x_1, \ldots, x_\mathrm{N}$ de K qui sont à distances $\geq \varepsilon$, c'est à dire $d(x_i, x_j) \geq \varepsilon$ pour tous $i \neq j$. Notons $\mathrm{N_R}(\mathrm{K}, \varepsilon)$ le min du nombre N de points de K tel que K soit recouvert par les boules $\mathrm{B}(x_i, \varepsilon)$, $i = 1, \ldots, \mathrm{N}$. Montrer que

$$\mathrm{N_R}(\mathrm{K}, \varepsilon) \leq \mathrm{N_P}(\mathrm{K}, \varepsilon) \leq \mathrm{N_R}(\mathrm{K}, \varepsilon/2).$$

On suppose maintenant que $\mathrm{X} = \mathbb{R}^p$ est muni d'une norme, et on prend pour d la distance déduite de la norme, et pour K la boule unité B de cet espace normé. Montrer que :

$$\frac{1}{\varepsilon^p} \leq \mathrm{N_R}(\mathrm{B}, \varepsilon); \quad \mathrm{N_P}(\mathrm{B}, 2\varepsilon) \leq \frac{(1+\varepsilon)^p}{\varepsilon^p} \cdot$$

On boucle le cercle Bolzano - compacité

On suppose que (X, d) est un espace métrique qui vérifie le théorème de Bolzano-Weierstrass : toute suite $(x_n)_{n \in \mathbb{N}}$ de points de X admet des sous-suites convergentes dans X. On montre d'abord que les suites décroissantes de fermés non vides de X ont une intersection non vide : soit (F_n) une suite décroissante de fermés non vides, et choisissons $x_n \in \mathrm{F}_n$ pour tout $n \geq 0$. D'après Bolzano, il existe une sous-suite (x_{n_k}) convergente vers un $x \in \mathrm{X}$. On vérifie facilement que x est dans l'intersection des (F_n).

En passant au complémentaire, on déduit que si X est recouvert par une famille dénombrable d'ouverts (W_j), il existe un sous-recouvrement fini (prendre les ensembles fermés $\mathrm{F}_n = (\mathrm{W}_0 \cup \ldots \cup \mathrm{W}_n)^c$, décroissants et dont l'intersection est vide ; cela montre que l'un d'entre eux, F_{n_0}, doit être déjà vide, ce qui signifie que X est recouvert par $\mathrm{W}_0, \ldots, \mathrm{W}_{n_0}$).

Si (X, d) vérifie Bolzano, on a vu qu'il est séparable ; sa topologie admet donc une base dénombrable (V_n). Si les $(\mathrm{U}_i)_{i \in \mathrm{I}}$ ouverts recouvrent (X, d) (où I est un ensemble d'indices quelconque) soit (W_j) la famille de ceux des V_n qui sont contenus dans un U_i ; d'après la propriété de *base* d'une topologie, la famille dénombrable (W_j) forme un recouvrement de X ; d'après le paragraphe précédent, il existe un sous-recouvrement fini $\mathrm{W}_0 \cup \ldots \cup \mathrm{W}_\mathrm{N}$; choisissant pour $j = 0, \ldots, \mathrm{N}$ un ouvert U_{i_j} qui contienne W_j, on forme aussi un sous-recouvrement fini $(\mathrm{U}_{i_j})_{j=0}^{\mathrm{N}}$ du recouvrement ouvert initial $(\mathrm{U}_i)_{i \in \mathrm{I}}$.

Prépa. Agrég. écrit d'Analyse, séance n° 13, Mercredi 28 Février 2001.

Compacité 2

Proposition. *Si* K *est un espace topologique métrisable, il est compact si et seulement s'il a la propriété de Bolzano-Weierstrass : toute suite* (x_n) *de points de* K *admet des sous-suites convergentes.*

Démonstration. On a vu le sens (Bolzano implique compact) la semaine précédente. Inversement, supposons (K, d) métrique compact et soit (x_n) une suite dans K ; pour tout $n \geq 0$ on pose $\mathrm{E}_n = \{x_{n+1}, x_{n+2}, \ldots\}$ et on pose ensuite $\mathrm{F}_n = \overline{\mathrm{E}_n}$; la suite (F_n) est une suite décroissante de fermés non vides ; posons $\mathrm{Z} = \bigcap_n \mathrm{F}_n$ et soit z un point quelconque de l'ensemble fermé non vide Z ; on va construire une sous-suite (x_{n_k}) qui converge vers z. Supposons les indices $n_0 < n_1 < \ldots < n_k$ déjà choisis, avec $d(x_{n_j}, z) < 2^{-j}$ pour $j = 0, \ldots, k$. Puisque z est dans F_{n_k}, qui est l'adhérence de E_{n_k}, la boule ouverte $\mathrm{B}(z, 2^{-k-1})$ rencontre E_{n_k} donc il existe un entier $n > n_k$ tel que l'ouvert $\mathrm{B}(z, 2^{-k-1})$ contienne x_n : on posera $n_{k+1} = n$, et on continuera ainsi.

Produit dénombrable de compacts métriques : la technique de la suite diagonale

Proposition. *Tout produit dénombrable de compacts métrisables vérifie Bolzano, donc est compact.*

Démonstration. Pour exprimer la démonstration, il est utile d'introduire une convention de notation. Si $\mathrm{M} = \{n_0 < \ldots < n_j < \ldots\}$ est un sous-ensemble infini de $\mathbb{N}$, et si (x_n) est une suite d'éléments d'un ensemble X quelconque, convenons de noter la sous-suite (x_{n_j}) par $(x_n)_{n \in \mathrm{M}}$.

Soit donc $(x^{(n)})$ une suite d'éléments du produit dénombrable $\mathrm{X} = \prod_{k \in \mathbb{N}} \mathrm{X}_k$; pour chaque entier $k \geq 0$, on a une suite $(x_k^{(n)})_n$ dans X_k, formée des $k$ièmes coordonnées des $x^{(n)}$. Puisque X_0 est métrique compact, la suite $(x_0^{(n)})$ admet une sous-suite convergente $(x_0^{(n)})_{n \in \mathrm{M}_0}$. La suite $(x_1^{(n)})_{n \in \mathrm{M}_0}$ est dans le compact métrique X_1, donc on peut trouver un nouvel ensemble infini $\mathrm{M}_1 \subset \mathrm{M}_0$ tel que la sous-suite $(x_1^{(n)})_{n \in \mathrm{M}_1}$ soit convergente dans X_1. En continuant ainsi, on construit une suite décroissante $\mathrm{M}_0 \supset \mathrm{M}_1 \supset \ldots \supset \mathrm{M}_j \supset \ldots$ de sous-ensembles infinis de $\mathbb{N}$, telle que la sous-suite $(x_j^{(n)})_{n \in \mathrm{M}_j}$ soit convergente pour tout $j \geq 0$.

C'est ici qu'intervient le procédé de *la suite diagonale*. Construisons un ensemble infini M formé du premier élément n_0 de M_0, puis du premier élément n_1 de M_1 qui soit $> n_0$, etc... On constate que pour tout entier $k \geq 0$, la sous-suite $(x_k^{(n)})_{n \in \mathrm{M}}$ est convergente dans X_k vers un élément y_k : en effet, l'ensemble M est contenu dans M_k à un ensemble fini près, pour tout $k \geq 0$. Cela signifie que la sous-suite $(x^{(n)})_{n \in \mathrm{M}}$ converge dans l'espace produit vers $y = (y_k)$.

L'espace produit $\prod_{k \in \mathbb{N}} \mathrm{X}_k$ est métrisable et vérifie Bolzano, donc il est compact.

Métrique compact et métrique complet

Rappelons qu'une partie A d'un espace topologique séparé X est dite *relativement compacte* dans X si son adhérence $\overline{\mathrm{A}}$ dans X est compacte.

Définition. Un espace métrique X est dit *précompact* si, pour tout $\varepsilon > 0$ il existe un recouvrement fini de X par des parties de diamètre $\leq \varepsilon$.

Une partie A de l'espace métrique X est précompacte si pour tout $\varepsilon > 0$ il existe un recouvrement fini de A par des boules de X de rayon ε (les centres peuvent être dans A ou pas ; si on peut recouvrir A avec N boules de rayon $\varepsilon/2$ centrées en des points de X, on pourra recouvrir A avec N boules de rayon ε centrées en des points de A).

Exemple. Si E est un espace vectoriel normé réel dont la boule unité peut être recouverte par un nombre fini de boules de rayon $r < 1$, alors E est de dimension finie.

En particulier, si la boule unité fermée de E est compacte, l'espace E est de dimension finie (c'est un théorème de F. Riesz, qui sert à dire que certains sous-espaces propres sont de dimension finie ; on y reviendra).

Soit Z un sous-espace de dimension finie d de E, munissons Z d'une mesure μ de Lebesgue en choisissant un isomorphisme de Z avec $\mathbb{R}^d$; les propriétés essentielles de cette mesure μ sont l'invariance par translation (par les vecteurs de Z) et le fait que $\mu(\lambda \mathrm{A}) = \lambda^d \mu(\mathrm{A})$ pour tout $\lambda > 0$ et tout borélien $\mathrm{A} \subset \mathrm{Z}$ (si on réfléchit bien on voit que la deuxième propriété résulte de la première), et la finitude de μ sur les compacts de Z. Posons

$$\mathrm{V} = \sup\{\mu(\mathrm{Z} \cap \mathrm{B}(x, 1)) : x \in \mathrm{E}\}.$$

Il est clair que V est fini (majoré par $\mu(\mathrm{B}(0, 2))$: si on choisit $z \in \mathrm{Z} \cap \mathrm{B}(x, 1)$, alors $\mathrm{Z} \cap \mathrm{B}(x, 1) \subset \mathrm{B}(z, 2)$ et $\mu(\mathrm{Z} \cap \mathrm{B}(z, 2)) = \mu(\mathrm{B}(0, 2))$ par invariance par translation de vecteur $z \in \mathrm{Z}$). Pour toute boule de rayon r on aura alors $\mu(\mathrm{Z} \cap \mathrm{B}(x, r)) \le r^d \mathrm{V}$. Soit $x_0 \in \mathrm{E}$ tel que $\mu(\mathrm{Z} \cap \mathrm{B}(x_0, 1)) > \mathrm{V} - \varepsilon$; puisque la boule unité de E est recouverte par N boules de rayon r, la boule $\mathrm{B}(x_0, 1)$ aussi, donc $\mathrm{Z} \cap \mathrm{B}(x_0, 1)$ est recouvert par N ensembles $\mathrm{Z} \cap \mathrm{B}(x_i, r)$, $i = 1, \ldots, \mathrm{N}$. En passant à la mesure,

$$\mathrm{V} - \varepsilon < \mu(\mathrm{Z} \cap \mathrm{B}(x_0, 1)) \le \sum_{i=1}^{\mathrm{N}} \mu(\mathrm{Z} \cap \mathrm{B}(x_i, r)) \le \mathrm{N}\, r^d \mathrm{V}$$

ce qui implique $\mathrm{N} r^d \ge 1$ et $d \le \mathrm{N}/(\ln r^{-1})$. Comme Z est un sous-espace vectoriel quelconque dans E, il en résulte que $\dim \mathrm{E} \le \mathrm{N}/(\ln r^{-1})$.

Proposition. *Dans un espace métrique complet* (X, d), *une partie* A *a une adhérence compacte si et seulement si elle est précompacte.*

Un espace métrique est donc compact si et seulement s'il est précompact et complet.

Démonstration. Si l'adhérence de A est compacte il est facile de vérifier que A peut être recouvert par un nombre fini de boules de rayon $< \varepsilon$; on va esquisser la démonstration de l'autre direction. Supposons A précompact et soit (x_n) une suite dans $\overline{\mathrm{A}}$; on va montrer d'abord que pour tout $\varepsilon > 0$ on peut extraire une sous-suite telle que $\|x_{n_k} - x_{n_\ell}\| < \varepsilon$ pour tous k, ℓ. Par définition de l'adhérence on trouve d'abord $y_n \in \mathrm{A}$ tels que $d(x_n, y_n) < \varepsilon/4$. Puisque A peut être recouvert par une famille finie de boules de rayon $\varepsilon/4$, on peut trouver une infinité d'indices m, formant un sous-ensemble infini $\mathrm{M} \subset \mathbb{N}$ tels que les y_m pour $m \in \mathrm{M}$ tombent dans la même boule $\mathrm{B}(x, \varepsilon/4)$. Pour $m, n \in \mathrm{M}$ on aura $d(y_m, y_n) < \varepsilon/2$, donc $d(x_m, x_n) < \varepsilon$.

On appliquera ce premier pas successivement avec $\varepsilon = 1/2, 1/4$, etc… en prenant à chaque fois une sous-suite de la sous-suite précédente, puis on prendra une sous-suite diagonale qui sera de Cauchy, donc convergente dans l'espace *complet* E. On conclut que $\overline{\mathrm{A}}$ est compact puisqu'il vérifie Bolzano.

Exercice : régularité des mesures sur la tribu borélienne d'un polonais. Soit μ une probabilité sur la tribu borélienne d'un espace polonais X ; montrer que pour tout $\varepsilon > 0$ il existe un compact $\mathrm{K} \subset \mathrm{X}$ tel que $\mu(\mathrm{K}) > 1 - \varepsilon$.

Soit d une distance qui définit la topologie de X ; on montre d'abord que pour tous $r > 0$ et $\alpha > 0$, il existe un fermé F de X qui est réunion finie de boules de rayon r, et tel que $\mu(\mathrm{F}) > 1 - \alpha$.

Soit (x_n) une suite dense dans X ; posons $\mathrm{A}_n = \bigcup_{i=0}^n \overline{\mathrm{B}(x_i, r)}$. La suite (A_n) est croissante et recouvre X (parce que (x_n) est dense) donc il existe n_0 tel que $\mu(\mathrm{A}_{n_0}) > 1 - \alpha$. Alors $\mathrm{F} = \mathrm{A}_{n_0}$ convient.

Pour conclure on applique ce qui précède avec $r = 2^{-k}$ et $\alpha = \varepsilon/2^{k+1}$, pour tout $k \geq 0$. On obtient ainsi une suite (F_k) de fermés telle que $\mu(\mathrm{F}_k) > 1 - 2^{-k-1}\varepsilon$ pour tout k, ce qui implique que si $\mathrm{K} = \bigcap_{k\in\mathbb{N}} \mathrm{F}_k$, on aura $\mu(\mathrm{K}) > 1 - \varepsilon$. Par ailleurs, on voit que K est fermé et précompact dans l'espace complet X, donc K est compact.

Compacité dans le cas vectoriel normé

Dans le cas d'un sous-ensemble A d'un espace de Banach E, il est agréable de retenir un critère (qui sera noté **C**) qui utilise le caractère vectoriel de l'espace ambiant : pour que l'adhérence de A soit compacte dans l'espace de Banach E, il faut et il suffit que A vérifie les deux conditions suivantes :

C1 – l'ensemble A est borné ;

C2 – pour tout $\varepsilon > 0$, il existe un sous-espace vectoriel $\mathrm{L}_\varepsilon \subset \mathrm{E}$ *de dimension finie* tel que tout point de A soit à une distance $< \varepsilon$ de L_ε :

$$\forall x \in \mathrm{A}, \quad \mathrm{dist}(x, \mathrm{L}_\varepsilon) < \varepsilon.$$

Il est évident que tout ensemble précompact vérifie les deux conditions précédentes. Inversement, supposons que $\mathrm{A} \subset \mathrm{E}$ soit borné par M et vérifie **C2**. Donnons $\varepsilon > 0$ et choisissons L de dimension finie qui vérifie **C2** avec la valeur $\varepsilon/2$. Considérons par ailleurs le compact

$$\mathrm{K} = \{y \in \mathrm{L} : \|y\| \leq \mathrm{M} + \varepsilon\}$$

(fermé borné en dimension finie). Si on recouvre le compact K par un nombre fini de boules $\mathrm{B}(x_i, \varepsilon/2)$, les boules $\mathrm{B}(x_i, \varepsilon)$ recouvriront A, ce qui prouve la précompacité de l'ensemble A.

Définition : opérateurs compacts. On dit qu'une application linéaire continue T d'un espace de Banach E dans un espace de Banach F est *compacte* si l'adhérence de l'image de la boule unité de E est compacte dans F.

On montre que l'ensemble $\mathcal{K}(\mathrm{E}, \mathrm{F})$ des opérateurs compacts de E dans F est un sous-espace vectoriel fermé de $\mathcal{L}(\mathrm{E}, \mathrm{F})$, et qui a de plus la propriété d'idéal.

Proposition : critère d'approximation par des opérateurs bornés de rang fini. *Si* (T_n) *est une suite d'opérateurs de rang fini qui tend vers* T *dans* $\mathcal{L}(\mathrm{E}, \mathrm{F})$*, l'opérateur* T *est compact.*

Démonstration. C'est à peu près celle qui montre que $\mathcal{K}(\mathrm{E}, \mathrm{F})$ est fermé dans $\mathcal{L}(\mathrm{E}, \mathrm{F})$. On va montrer que $\mathrm{A} = \mathrm{T}(\mathrm{B}_\mathrm{E})$ vérifie les deux points du critère **C** précédent. On sait déjà que T est borné, donc A est un ensemble borné. Soit $\varepsilon > 0$ donné ; pour n assez grand, on aura $\|\mathrm{T}_n - \mathrm{T}\| < \varepsilon$, ce qui garantit que tout point de A a une distance $< \varepsilon$ à l'ensemble $\mathrm{A}_n = \mathrm{T}_n(\mathrm{B})$ (en effet, tout point $y \in \mathrm{A}$ est de la forme $y = \mathrm{T}(x)$ pour un $x \in \mathrm{E}$ tel que $\|x\| \leq 1$; si $y_n = \mathrm{T}_n(x)$ on aura $y_n \in \mathrm{T}_n(\mathrm{B})$ et $\|y - y_n\| < \varepsilon)$; mais $\mathrm{L}_n = \mathrm{T}_n(\mathrm{E})$ est un sous-espace vectoriel de dimension finie de F, qui contient A_n, d'où le résultat.

Exercice. Opérateurs diagonaux sur les espaces ℓ_p. Soit $\alpha = (\alpha_n)$ une suite scalaire qui tend vers 0 et définissons un opérateur borné Δ_α sur ℓ_p, $1 \leq p < +\infty$ en posant

$$\Delta_\alpha\big((x_n)_{n\geq 0}\big) = (\alpha_n x_n)_{n\geq 0}.$$

Montrer que Δ_α est compact de ℓ_p dans ℓ_p. Réciproque ?

Partition de l'unité sur un espace topologique X

C'est une famille $(\varphi_i)_{i\in\mathrm{I}}$ de fonctions réelles continues sur X, en général finie, telle que $0 \leq \varphi_i \leq 1$ pour chaque $i \in \mathrm{I}$ et $\sum_i \varphi(x) = 1$ pour tout $x \in \mathrm{X}$.

Sur un métrique compact (K, d) il est très facile de fabriquer une partition de l'unité, subordonnée à un recouvrement ouvert fini $(\mathrm{U}_i)_{i=1}^n$ donné, ce qui signifiera pour nous que l'on a une famille $(\varphi_i)_{i=1}^n$ telle que $\varphi_i = 0$ en dehors de U_i pour chaque indice i (on peut demander plus, à savoir que le *support* de φ_i soit contenu dans l'ouvert U_i). On pose d'abord $\psi_i(x) = \mathrm{dist}(x, \mathrm{U}_i^c)$; la fonction ψ_i est continue ≥ 0, nulle hors de U_i, $\psi_i > 0$ sur U_i et $\psi = \sum_{i=1}^n \psi_i > 0$ parce que (U_i) est un recouvrement de K. On pose finalement $\varphi_i = \psi_i/\psi$.

Proposition. *Soient* (K, d) *un espace métrique compact et* $\varphi_1, \ldots, \varphi_\mathrm{N}$ *une partition de l'unité subordonnée à un recouvrement de* K *par des boules* $\mathrm{B}(x_i, \delta)$, $\delta > 0$ *et* $i = 1, \ldots, \mathrm{N}$ *; pour toute fonction continue* f *sur* K, *on a dans l'espace de Banach* C(K)

$$\mathrm{dist}(f, \mathrm{V_N}) \leq \omega_f(\delta)$$

où $\mathrm{V_N} = \mathrm{Vect}(\varphi_1, \ldots, \varphi_\mathrm{N})$.

Démonstration. On rappelle que

$$\omega_f(\delta) = \sup\{|f(x) - f(y)| : x, y \in \mathrm{K},\ d(x, y) \leq \delta\}.$$

Posons

$$g = \sum_{i=1}^{\mathrm{N}} f(x_i)\, \varphi_i \in \mathrm{V_N}.$$

On a

$$f(x) - g(x) = \sum_{i=0}^{\mathrm{N}} (f(x) - f(x_i))\, \varphi_i(x).$$

Considérons un terme $t_i = (f(x) - f(x_i))\, \varphi_i(x)$ de la somme précédente ; si ce terme t_i est non nul, c'est que $\varphi_i(x) \neq 0$, donc $x \in \mathrm{U}_i = \mathrm{B}(x_i, \delta)$, donc $d(x, x_i) < \delta$ et $|f(x) - f(x_i)| \leq \omega_f(\delta)$; on a donc toujours $|t_i| \leq \omega_f(\delta)\, \varphi_i(x)$, donc

$$|f(x) - g(x)| \leq \sum_{i=0}^{\mathrm{N}} |f(x) - f(x_i)|\, \varphi_i(x) \leq \sum_{i=0}^{\mathrm{N}} \omega_f(\delta)\, \varphi_i(x) = \omega_f(\delta).$$

On a bien obtenu que $\|f - g\| \leq \omega_f(\delta)$, donc $\mathrm{dist}(f, \mathrm{V_N}) \leq \omega_f(\delta)$.

Théorème. *Soit* (K, d) *un espace métrique compact ; on considère un ensemble* A **borné** *dans* C(K) *et on suppose qu'il existe une fonction* $\delta > 0 \to \omega(\delta)$ *nulle et continue en* 0 *telle que*

$$\forall f \in \mathrm{A},\ \forall \delta > 0,\ \forall x, y \in \mathrm{K}, \quad \bigl(d(x,y) \le \delta \Rightarrow |f(x) - f(y)| \le \omega(\delta)\bigr).$$

Alors l'adhérence de A *est compacte dans* C(K).

Démonstration. On va utiliser le critère de précompacité **C**, adapté au cas vectoriel. On sait déjà que A est borné, il reste à montrer la propriété d'approximation par des sous-espaces vectoriels de dimension finie. Soit $\varepsilon > 0$; on peut trouver $\delta > 0$ tel que $\omega(\delta) < \varepsilon$. A l'évidence, la propriété de l'énoncé signifie que

$$\forall f \in \mathrm{A}, \quad \omega_f(\delta) \le \omega(\delta)$$

et la proposition précédente nous donne $\mathrm{dist}(f, \mathrm{V_N}) < \varepsilon$, où $\mathrm{V_N}$ sera l'espace de dimension finie engendré par les fonctions $\varphi_1, \ldots, \varphi_\mathrm{N}$ d'une partition de l'unité subordonnée à un recouvrement de K par des boules de rayon δ, en nombre N.

Remarque : équicontinuité, équicontinuité uniforme. La propriété de l'énoncé précédent s'appelle *équicontinuité uniforme* ; en fait, pour avoir la compacité de l'adhérence de A dans C(K), il suffit que l'ensemble $\mathrm{A} \subset \mathrm{C(K)}$ soit borné et simplement *équicontinu*

$$\forall x \in \mathrm{K},\ \forall \varepsilon > 0,\ \exists \delta > 0,\ \forall f \in \mathrm{A}, \quad \bigl(d(y,x) < \delta \Rightarrow |f(y) - f(x)| < \varepsilon\bigr).$$

On montre que l'équicontinuité implique l'équicontinuité uniforme comme on montre qu'une fonction continue sur le compact K est uniformément continue.

Dans la pratique, on voit en général pour le même prix que l'ensemble à étudier est *uniformément* équicontinu.

Prépa. Agrég. écrit d'Analyse, séance n° 14, Mercredi 7 Mars 2001.

Compacité 3

Résolution d'exercice proposé : si $\mathrm{A} \subset \mathrm{C(K)}$ est équicontinu sur (K, d) métrique compact, il est uniformément équicontinu, c'est à dire :

pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour toute $f \in \mathrm{A}$ et tous $y_1, y_2 \in \mathrm{K}$, la condition $d(y_1, y_2) < \delta$ entraîne $|f(y_1) - f(y_2)| < \varepsilon$.

Solution. D'après la définition de l'équicontinuité de la famille A, on peut trouver pour tout point $x \in \mathrm{K}$ une boule ouverte $\mathrm{B}_x = \mathrm{B}(x, r(x))$ telle que $r(x) > 0$ et

$$\forall y \in \mathrm{B}_x, \forall f \in \mathrm{A}, \quad |f(y) - f(x)| < \varepsilon/2.$$

Puisque l'espace K est compact, on peut extraire un sous-recouvrement fini $\mathrm{U}_1, \ldots, \mathrm{U}_n$ du recouvrement ouvert (B_x), où $\mathrm{U}_j = \mathrm{B}(x_j, r_j)$ pour $j = 1, \ldots, \mathrm{N}$. La fonction ψ définie sur K par $\psi(y) = \max d(y, \mathrm{U}_j^c)$ est continue et > 0 sur K, donc son minimum δ est > 0. Cela signifie que pour tout $y \in \mathrm{K}$, la boule $\mathrm{B}(y, \delta)$ est contenue dans l'un des ouverts U_j. Etant donnés $y_1, y_2 \in \mathrm{K}$ tels que $d(y_1, y_2) < \delta$, on pourra trouver un indice j tel que $\mathrm{B}(y_1, \delta) \subset \mathrm{U}_j$, donc $y_1, y_2 \in \mathrm{U}_j$ et il en résultera que $|f(y_1) - f(y_2)| < \varepsilon$ pour toute $f \in \mathrm{A}$. On a montré l'équicontinuité *uniforme* de la famille A.

Exercice. Si K est un espace topologique compact, K est métrisable si et seulement si C(K) est séparable.

Expliquons le sens qui part de l'hypothèse K métrisable ; pour $\varepsilon = 2^{-k}$, avec $k = 0, 1, \ldots$ on peut trouver une partition de l'unité $(\varphi_j^{(k)})_{j=1,\ldots,\mathrm{N}_k}$ subordonnée à un recouvrement fini de (K, d) par des boules de rayon 2^{-k}. On a vu que pour toute $f \in \mathrm{C(K)}$, on a $\mathrm{dist}(f, \mathrm{V}_k) \le \omega_f(2^{-k})$, où V_k désigne l'espace vectoriel de dimension finie N_k engendré par les $(\varphi_j^{(k)})$, $j = 1, \ldots, \mathrm{N}_k$. L'espace vectoriel V engendré par toutes les fonctions $(\varphi_j^{(k)})$, où j et k varient, est un sous-ensemble séparable de C(K) et de plus, pour toute fonction f continue on sait que $\mathrm{dist}(f, \mathrm{V}) \le \omega_f(2^{-k})$ pour tout k, donc V est dense dans C(K). On pourrait aussi dire que l'ensemble dénombrable des $(\varphi_j^{(k)})$ est *total* dans C(K).

Pour montrer l'autre direction, il faut savoir que si K est *topologique* compact et si x, y sont deux points distincts dans K, on peut trouver une fonction réelle continue f sur K telle que $f(x) \neq f(y)$. En termes pompeux : tout espace compact se plonge dans un espace produit $[0, 1]^{\mathrm{I}}$.

Un espace *topologique* compact K peut être séparable sans être métrisable (voir exercice 13.1).

Exercice : l'espace des fonctions höldériennes sur un métrique compact (K, d). Pour $0 < \alpha \le 1$, on désignera par $\Gamma_\alpha(\mathrm{K}, d)$ l'espace vectoriel des fonctions f réelles et continues sur K telles qu'il existe une constante K telle que $|f(x) - f(y)| \le \mathrm{K}\, d(x, y)^\alpha$ pour tous $x, y \in \mathrm{K}$. Montrer que cet espace est complet pour la norme

$$\|f\|_{\Gamma_\alpha} = \max\bigl(\sup\{d(x, y)^{-\alpha}|f(x) - f(y)| : x \neq y\},\ \|f\|_\infty\bigr).$$

Pour $\alpha = 1$, c'est l'espace de Banach des fonctions lipschitziennes sur $\mathrm{K}, d)$.

Montrer que la boule unité fermée de $\Gamma_\alpha(\mathrm{K}, d)$ est compacte dans $\mathrm{C}(\mathrm{K}, d)$.

Soit A cette boule unité ; la précompacité de A dans C(K) est immédiate par Ascoli, car $\|f\|_\infty \le 1$ pour toute $f \in \mathrm{A}$ (l'ensemble A est borné en norme uniforme) et de plus la condition $|f(x) - f(y)| \le |x-y|^\alpha$ donne l'équicontinuité uniforme de A puisque $\alpha > 0$. Pour terminer il reste à voir que A est uniformément fermé, ce qui est facile.

L'espace $\mathrm{C}^{1,\alpha}$ joue un rôle dans certaines questions d'Analyse : c'est l'espace des fonctions de classe C^1 dont toutes les dérivées partielles sont α-hölderiennes.

Exercice. Pour toute fonction $f \in \mathrm{L}_1(0,1)$ on définit une fonction $\mathrm{P}f$ continue sur $[0,1]$ par la formule

$$\forall x \in [0,1], \quad (\mathrm{P}f)(x) = \int_0^x f(t)\,dt.$$

L'application de primitive P est compacte de tous les L_p sauf L_1, à valeurs dans $\mathrm{C}(0,1)$.

Pour être plus précis, désignons par P_p l'opérateur linéaire borné de $\mathrm{L}_p(0,1)$ dans $\mathrm{C}(0,1)$ qui associe à chaque $f \in \mathrm{L}_p$ la fonction continue $\mathrm{P}f$. Alors P_p est compact pour $1 < p \le +\infty$, mais P_1 n'est pas compact.

Plus généralement, soit K une fonction mesurable bornée sur $[0,1]^2$; on pose

$$\forall x \in [0,1], \quad (\mathrm{T_K}f)(x) = \int_0^1 \mathrm{K}(x,y) f(y)\,dy.$$

Cet opérateur (opérateur à noyau borné) est compact de L_p dans C pour tout $p > 1$. L'opérateur de primitive est de cette forme : il suffit de poser $\mathrm{K}(x,y) = \mathbf{1}_{0 \le y \le x \le 1}$.

Expliquons pourquoi P_1 n'est pas compact de L_1 dans C. Etant donné un intervalle $\mathrm{J} \subset [0,1]$ désignons par $f_\mathrm{J} \in \mathrm{L}_1(0,1)$ la fonction f nulle hors de J, telle que $f(x) = 1/|\mathrm{J}|$ sur la première moitié de J et $f(x) = -1/|\mathrm{J}|$ sur la deuxième moitié. On voit que $\|f_\mathrm{J}\|_1 = 1$, et que $\mathrm{P}f$ est une “fonction triangle”, nulle hors de J, égale à $1/2$ au milieu de l'intervalle J. Il est facile de voir que quand J varie, la famille de toutes les fonctions $(\mathrm{P}f_\mathrm{J})$, qui est contenue dans l'image de la boule unité de $\mathrm{L}_1(0,1)$, n'est pas relativement compacte dans $\mathrm{C}(0,1)$ (considérer une suite (J_n) d'intervalles disjoints).

Remarque. L'opérateur d'intégration est compact de $\mathrm{L}_2(0,1)$ dans $\mathrm{C}(0,1)$, donc il est compact aussi comme endomorphisme de $\mathrm{L}_2(0,1)$ (ou comme endomorphisme de $\mathrm{C}(0,1)$), en composant l'opérateur compact P_2 à gauche (ou à droite) avec l'injection de $\mathrm{C}(0,1)$ dans $\mathrm{L}_2(0,1)$.

Expression d'Ascoli avec les translations dans $\mathbb{R}^d$

Si f est une fonction sur $\mathbb{R}^d$ et $h \in \mathbb{R}^d$ un vecteur, on posera

$$\forall x \in \mathbb{R}^d, \quad (\tau_h f)(x) = f(x-h).$$

Soit $\mathrm{A} \subset \mathrm{C}_{\mathrm{comp}}(\mathbb{R}^d)$ *un ensemble de fonctions continues avec un support contenu dans un borné* $\mathrm{B} \subset \mathbb{R}^d$ *fixé ; l'ensemble* A *est relativement compact en norme uniforme si et seulement s'il est borné en norme du sup et si, pour tout* $\varepsilon > 0$ *donné il existe* $\delta > 0$ *tel que*

$$\forall f \in \mathrm{A}, \quad \|\tau_h f - f\|_\infty < \varepsilon$$

pour toute translation τ_h *dont le vecteur* $h \in \mathbb{R}^d$ *vérifie* $\|h\| < \delta$.

C'est une simple transcription de l'équicontinuité uniforme. On pourra remarquer (ça n'est pas important) que l'hypothèse sur les translations et le fait que les fonctions $f \in \mathrm{A}$ soient à support dans un même borné entraînent automatiquement que A est borné en norme uniforme, mais je préfère garder cette hypothèse, que l'on a tendance à oublier dans Ascoli à des endroits où elle est nécessaire.

ATTENTION ! C'est faux si on supprime la restriction que les fonctions de A doivent avoir des supports contenus dans un borné fixé :

il suffit de prendre $d = 1$, une première fonction φ continue sur $\mathbb{R}$, à support compact, non nulle, et l'ensemble A formé de toutes les translatées $\tau_h\varphi$, $h \in \mathbb{R}$. On voit facilement que cet ensemble A vérifie les deux conditions ci-dessus, mais la suite $(\tau_n\varphi) \subset \mathrm{A}$ n'admet aucune sous-suite uniformément convergente, puisqu'elle converge simplement vers 0 sur $\mathbb{R}$, mais pas uniformément.

Exercice. Montrer qu'on peut garder la conclusion de précompacité dans le critère précédent si on adoucit l'hypothèse sur les supports en la suivante : il existe une fonction continue φ sur $\mathbb{R}^d$, qui tend vers 0 à l'infini, et telle que $|f| \le \varphi$ pour toute $f \in \mathrm{A}$ (on garde bien entendu l'hypothèse sur les translations). Montrer que l'hypothèse de domination précédente est *nécessaire* pour avoir la précompacité uniforme d'un ensemble contenu dans $\mathrm{C}_{\mathrm{comp}}(\mathbb{R}^d)$.

Compacité dans $\mathrm{L}_p(\mathbb{R}^d)$

Pour ce type de résultat et pour l'Analyse Fonctionnelle en général, consulter le livre Hirsch-Lacombe.

Théorème. *Si* A_p *est un ensemble borné de* $\mathrm{L}_p(\mathbb{R}^d)$ *dont toutes les fonctions sont nulles en dehors d'un même ensemble* B *borné, et si pour tout* $\varepsilon > 0$ *donné il existe* $\delta > 0$ *tel que*

$$\forall f \in \mathrm{A}_p,\ \forall h \in \mathbb{R}^d, \quad \bigl(\|h\| < \delta \Rightarrow \|\tau_h f - f\|_p < \varepsilon\bigr)$$

alors A_p *est relativement compact dans* $\mathrm{L}_p(\mathbb{R}^d)$.

ATTENTION ! C'est faux si on enlève la condition que les fonctions $f \in \mathrm{A}_p$ sont nulles en dehors d'un borné fixé, pour la même raison qui a été donnée ci-dessus pour Ascoli ; mais il est facile d'obtenir un énoncé valable sur $\mathbb{R}^d$ tout entier : il suffit de demander **en plus** que pour tout $\varepsilon > 0$, il existe un ensemble borné $\mathrm{B}_\varepsilon \subset \mathbb{R}^d$ tel que $\|f - \mathbf{1}_{\mathrm{B}_\varepsilon} f\|_p < \varepsilon$ pour toute $f \in \mathrm{A}_p$.

Démonstration. On se donne $\varepsilon > 0$ et le $\delta > 0$ qui correspond dans la condition de l'énoncé ci-dessus. On choisit une fonction $\varphi \ge 0$ continue, d'intégrale 1, avec un support compact contenu dans la boule $\mathrm{B}(0, \delta)$ de $\mathbb{R}^d$. On considère l'ensemble

$$\mathrm{A}_\infty = \{f * \varphi : f \in \mathrm{A}_p\}$$

qui est formé de fonctions continues sur $\mathbb{R}^d$, à supports compacts ; on va montrer que A_∞ est relativement compact en norme uniforme en appliquant le critère précédent ; ensuite, on montrera que A_∞ est proche en norme L_p de l'ensemble A_p.

On commence par l'étude de A_∞ : les supports des $f * \varphi$ sont dans un borné fixe, un peu plus grand que B et A_∞ est borné en norme uniforme parce que $\|f * \varphi\|_\infty \le \|f\|_p \, \|\varphi\|_q$ (qui est finie parce que φ est continue à support compact) et parce que A_p est borné dans $\mathrm{L}_p(\mathbb{R}^d)$ par un certain nombre M. Ensuite, pour tout $\varepsilon_1 > 0$ donné, on a si $\|h\| < \delta_1$ convenable, où $0 < \delta_1 < \delta$,

$$\|\tau_h(f * \varphi) - f * \varphi\| = \|f * (\tau_h\varphi - \varphi)\|_\infty \le \|f\|_p \, \|\tau_h\varphi - \varphi\|_q \le \varepsilon_1 \, \mathrm{M}$$

car $h \to \tau_h\varphi$ est continue à valeurs dans $\mathrm{L}_q(\mathbb{R}^d)$ pour tout $q \in [1, +\infty]$ (pour $q = +\infty$, cela vient de la continuité uniforme de φ) ; tout ceci montre que A_∞ est relativement compact en norme uniforme ; comme les fonctions de A_∞ sont uniformément bornées et à support dans un borné fixé, cela entraîne que A_∞ est relativement compact dans $L_p(\mathbb{R}^d)$. De plus, puisque $\varphi(y)\,dy$ est une probabilité sur $\mathbb{R}^d$,

$$\bigl|(f*\varphi)(x) - f(x)\bigr|^p = \Bigl|\int \bigl(f(x-y) - f(x)\bigr)\,\varphi(y)\,dy\Bigr|^p \le \int \bigl|f(x-y) - f(x)\bigr|^p\,\varphi(y)\,dy$$

qui implique par Fubini que

$$\|f*\varphi - f\|_p^p \le \int \|\tau_h f - f\|_p^p\;\varphi(h)\,dh \le \varepsilon^p$$

parce que le support de φ est contenu dans $B(0,\delta)$.

En conclusion, l'ensemble A_p est précompact dans $L_p(\mathbb{R}^d)$ puisque pour tout $\varepsilon > 0$, il peut être approché à ε près (en norme L_p) par un ensemble précompact dans $L_p(\mathbb{R}^d)$.

Compacité de l'injection de $H_0^1(\Omega)$ *dans* L_2*, quand* Ω *est un ouvert borné de* $\mathbb{R}^d$

On considère l'espace $H_0^1(\Omega)$ d'un ouvert borné Ω de $\mathbb{R}^d$ et on prend pour ensemble $A_2 \subset L_2(\mathbb{R}^d)$ les fonctions de la boule unité de $H_0^1(\Omega)$, prolongées par 0 en dehors de Ω. Cet ensemble est compact dans $L_2(\mathbb{R}^d)$. En effet on va estimer $\|\tau_h f - f\|_2$ en utilisant la dérivée dans la direction u, vecteur de norme un colinéaire à h, $h = tu$, en commençant par le cas d'une fonction $f \in \mathcal{D}(\Omega)$, prolongée par 0 hors de Ω (on terminera par la densité de $\mathcal{D}(\Omega)$ dans $H_0^1(\Omega)$, vraie par définition de l'espace). On a donc pour tout $x \in \mathbb{R}^d$

$$f(x-tu) - f(x) = -\int_0^t \nabla f(x-su)\,.\,u\,ds$$

d'où on déduit aisément que

$$|f(x-tu) - f(x)|^2 \le t\int_0^t |\nabla f(x-su)|^2\,ds$$

puis après intégration en $x \in \mathbb{R}^d$ et Fubini

$$\|\tau_h f - f\|_2^2 \le t^2\,\|\nabla f\|_2^2$$

ce qui donne $\|\tau_h f - f\|_2 \le \|h\|$ pour toute f de la boule unité de $H_0^1(\Omega)$. Le critère de précompacité dans $L_p(\mathbb{R}^d)$ s'applique alors sans problème. De plus, on peut montrer que la boule unité fermée de $H_0^1(\Omega)$ est fermée dans $L_2(\mathbb{R}^d)$.

Le résultat de compacité est faux pour l'injection de $H^1(\Omega)$ dans $L_2(\Omega)$. C'est clairement faux si Ω est un ouvert borné constitué d'une infinité de composantes connexes (Ω_n) ; la fonction f_n égale à $|\Omega_n|^{-1/2}$ sur Ω_n et à 0 sur toutes les autres composantes est de norme 1 dans $H^1(\Omega)$ (les dérivées partielles sont nulles !), mais $\|f_n - f_m\|_2 = \sqrt{2}$ quand $m \ne n$, ce qui montre qu'aucune sous-suite ne peut converger dans $L_2(\Omega)$.

On peut aussi réaliser ce phénomène avec un ouvert connexe, en joignant les cellules (Ω_n) par des "couloirs" très étroits dans lesquels les fonctions (f_n) rejoindront la valeur 0 rapidement en sortant de Ω_n.

Prépa. Agrég. écrit d'Analyse, séance n^{o} 15, Mercredi 14 Mars 2001.

Rappels Hilbert

Le corps de base sera $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Pour les produits scalaires complexes, je choisis la notation $\langle x, y\rangle$ et je suppose que l'antilinéarité est du côté droit,

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

Je voudrais rappeler brièvement l'enchaînement des idées dans le début de ce chapitre classique.

On considère donc un $\mathbb{K}$-espace vectoriel E muni d'une application $(x, y) \in \mathrm{E}\times\mathrm{E} \to \langle x, y\rangle \in \mathbb{K}$ qui possède les propriétés suivantes :

– pour tout $y \in \mathrm{E}$ fixé, l'application $x \in \mathrm{E} \to \langle x, y\rangle$ est $\mathbb{K}$-linéaire de E dans $\mathbb{K}$.

– pour tous $x, y \in \mathrm{E}$ on a

$$\langle y, x\rangle = \overline{\langle x, y\rangle}$$

– et enfin $\langle x, x\rangle$, qui est réel d'après la ligne précédente, est > 0 pour tout vecteur $x \in \mathrm{E}$ non nul.

Dans ces conditions l'application $x \to \sqrt{\langle x, x\rangle}$ est une norme sur E (ceci résulte de Cauchy-Schwarz). On dit que E est un *espace de Hilbert* quand il est complet pour cette norme.

Théorème de projection et applications

Théorème. *Soient* H *un espace de Hilbert et* C *un ensemble convexe fermé non vide contenu dans* H *; pour tout point* $x \in \mathrm{H}$, *il existe un unique* $y \in \mathrm{C}$ *qui réalise le minimum de la distance de* x *aux points de* C,

$$\|x - y\| = \min\{\|x - y'\| : y' \in \mathrm{C}\}.$$

Lorsque $\mathrm{C} = \mathrm{E}$ *est un* $\mathbb{K}$*-sous-espace vectoriel fermé, le point* $y = \mathrm{P_E}x$ *est caractérisé par les deux propriétés suivantes :* $y \in \mathrm{E}$ *et* $y - x \perp \mathrm{E}$ *; dans ce cas, on montre que* $\mathrm{P_E}$ *est une application linéaire, de norme* ≤ 1 *(de norme un exactement, sauf si* $\mathrm{E} = \{0\}$*).*

La démonstration du fait que $x - \mathrm{P_E}x$ est orthogonal à tous les vecteurs de E utilise le lemme suivant, qui servira au moins une autre fois plus loin.

Lemme 0. *Si* $x \in \mathrm{H}$ *et si on a pour tout vecteur* z *d'un* $\mathbb{K}$*-sous-espace vectoriel* $\mathrm{E} \subset \mathrm{H}$ *la relation*

$$\mathrm{Re}\,\langle x, z\rangle = 0,$$

alors $\langle x, z\rangle = 0$ *pour tout* $z \in \mathrm{E}$, *c'est à dire* $x \perp \mathrm{E}$.

Démonstration. Soit $z \in \mathrm{E}$; on écrit $\langle x, z\rangle = r\,\mathrm{e}^{i\theta}$ avec $r \ge 0$; on peut appliquer l'hypothèse à $z' = \mathrm{e}^{i\theta}\,z \in \mathrm{E}$, et on obtient

$$0 = \mathrm{Re}\,\langle x, z'\rangle = \mathrm{Re}\Big(\mathrm{e}^{-i\theta}\langle x, z\rangle\Big) = r$$

donc $\langle x, z\rangle = 0$.

Exemple : projection orthogonale sur le sous-espace (de dimension finie) engendré par une suite orthonormée finie. Si $e_1, \ldots, e_n$ est orthonormée et $\mathrm{E} = \mathrm{Vect}(e_1, \ldots, e_n)$, on constate que

$$\forall x \in \mathrm{H}, \quad \mathrm{P_E} x = \sum_{i=1}^{n} \langle x, e_i \rangle \, e_i.$$

En effet, le vecteur candidat $y = \sum_{i=1}^{n} \langle x, e_i \rangle \, e_i$ est dans E, et

$$\langle y, e_j \rangle = \sum_{i=1}^{n} \langle x, e_i \rangle \, \langle e_i, e_j \rangle = \langle x, e_j \rangle$$

pour tout $j = 1, \ldots, n$ ce qui implique que $x - y$ est orthogonal à chaque e_j, donc orthogonal à E par linéarité du produit scalaire.

Conséquence : identification du dual. *Toute forme linéaire continue ℓ sur* H *est de la forme $x \in \mathrm{H} \to \langle x, z \rangle$ pour un certain vecteur $z \in \mathrm{H}$, uniquement déterminé.*

Esquisse. Si $\ell = 0$ on prend $z = 0$; sinon, on considère $\mathrm{E} = \ker \ell \neq \mathrm{H}$, $x \notin \mathrm{E}$ et on prend pour z un multiple convenable du vecteur non nul $x - \mathrm{P_E} x$.

Conséquence 2. *Pour tout sous-espace vectoriel fermé* E *de* H *on a*

$$\mathrm{H} = \mathrm{E} \oplus \mathrm{E}^{\perp}; \quad \mathrm{E}^{\perp\perp} = \mathrm{E}.$$

Esquisse. Pour la première relation il suffit d'écrire $x = \mathrm{P_E} x + (x - \mathrm{P_E} x)$. Pour la seconde on note d'abord que $\mathrm{E} \subset \mathrm{E}^{\perp\perp}$ (évident si on y pense) ; ensuite, si $x \in \mathrm{E}^{\perp\perp}$, le vecteur $x - \mathrm{P_E}(x)$ est dans $\mathrm{E}^{\perp\perp}$, mais il est aussi dans $\mathrm{E}^{\perp}$, donc il est nul.

Définition : système orthonormé, éventuellement autoindexé. On dit qu'une famille $(e_i)_{i \in \mathrm{I}}$ de vecteurs de E est orthonormée si $\langle e_i, e_j \rangle = \delta_{i,j}$ (symbole de Kronecker) pour tous $i, j \in \mathrm{I}$. On dit qu'une partie $\mathrm{O} \subset \mathrm{H}$ est orthonormée si $\|x\| = 1$ pour tout vecteur $x \in \mathrm{O}$ et si $\langle x, y \rangle = 0$ chaque fois que $x, y \in \mathrm{O}$ et $x \neq y$.

Corollaire : inégalité de Bessel. *Pour toute famille $(e_i)_{i \in \mathrm{I}}$ orthonormée dans* H *et pour tout vecteur $x \in \mathrm{H}$, on a l'inégalité*

$$\sum_{i \in \mathrm{I}} |\langle x, e_i \rangle|^2 \leq \|x\|^2.$$

Démonstration. Pour tout sous-ensemble fini $\mathrm{J} \subset \mathrm{I}$ le vecteur $\sum_{i \in \mathrm{J}} \langle x, e_i \rangle \, e_i$ est la projection orthogonale $\mathrm{P_J} x$ de x sur $\mathrm{E_J} = \mathrm{Vect}(e_i, \ i \in \mathrm{J})$. Puisque la norme de la projection $\mathrm{P_J}$ est ≤ 1, on a

$$\sum_{i \in \mathrm{J}} |\langle x, e_i \rangle|^2 = \Big\| \sum_{i \in \mathrm{J}} \langle x, e_i \rangle \, e_i \Big\|^2 = \|\mathrm{P_J} x\|^2 \leq \|x\|^2.$$

(A peu près) par définition, $\sum_{i \in \mathrm{I}}$ est le sup des $\sum_{i \in \mathrm{J}}$ quand J décrit la famille des sous-ensembles finis de I, d'où le résultat.

Définition. On dit qu'une famille orthonormée de vecteurs de H est une *base hilbertienne* de H si elle est totale dans H, c'est à dire que le sous-espace vectoriel engendré par cette famille est dense dans H.

Existence d'une base hilbertienne d'un espace de Hilbert H : on trouve (avec Zorn) un système maximal $\mathrm{O} \subset \mathrm{H}$ de vecteurs de norme un deux à deux orthogonaux. Soit F le sous-espace vectoriel fermé engendré par O ; si on avait $\mathrm{F} \neq \mathrm{H}$, on pourrait trouver un vecteur v de norme un orthogonal à F, donc orthogonal à tous les vecteurs de O, et alors $\mathrm{O} \cup \{v\}$ contredirait la maximalité de O.

Si l'espace de Hilbert H est séparable, ce système orthonormé de vecteurs est au plus dénombrable, puisque les distances entre deux éléments distincts sont égales à $\sqrt{2}$. Si H est à la fois séparable et de dimension infinie, on peut écrire une telle base sous la forme $(e_n)_{n \in \mathbb{N}}$ (si on veut, mais dans certains exemples comme celui du système de Fourier l'ensemble naturel d'indices n'est pas $\mathbb{N}$). Dans ce cas tout vecteur $x \in \mathrm{H}$ se représente par la série convergente dans H

$$(*) \qquad x = \sum_{n=0}^{+\infty} \langle x, e_n \rangle \, e_n.$$

et le produit scalaire s'écrit $\langle x, y \rangle = \sum_{n=0}^{+\infty} \langle x, e_n \rangle \overline{\langle y, e_n \rangle}$, en particulier

$$\|x\|^2 = \sum_{n=0}^{+\infty} |\langle x, e_n \rangle|^2.$$

Exercice : somme hilbertienne de sous-espaces fermés. Soit $(\mathrm{E}_n)_{n \in \mathbb{N}}$ une suite de sous-espaces vectoriels fermés de H, deux à deux orthogonaux ; la somme directe hilbertienne de la famille (E_n) est l'adhérence F du sous-espace vectoriel

$$\mathrm{L} = \bigcup_{n \geq 0} \mathrm{E}_0 \oplus \cdots \oplus \mathrm{E}_n.$$

Montrer que la somme hilbertienne F est égale au sous-espace vectoriel de H formé de toutes les sommes x de séries $\sum_{n=0}^{+\infty} x_n$ où $\sum \|x_n\|^2 < +\infty$ et $x_n \in \mathrm{E}_n$ pour tout $n \geq 0$.

On pourrait convenir de noter $\sum_{n=0}^{+\infty} \mathrm{E}_n = \mathrm{F}$ la somme hilbertienne d'une famille de sous-espaces deux à deux orthogonaux, dans la mesure où une notation de série infinie de sous-espaces vectoriels n'avait pas de sens *a priori* en algèbre.

Systèmes de Haar

On va d'abord décrire le système de Haar sur $\mathbb{R}$. On part de la fonction de base

$$h_{0,0}(x) = \mathbf{1}_{[0,1/2[} - \mathbf{1}_{[1/2,1[}$$

qui est d'intégrale nulle et de norme un dans $\mathrm{L}_2(\mathbb{R})$. Dans ce paragraphe, on appellera *support* d'une fonction f l'ensemble des points x tels que $f(x) \neq 0$; ainsi, le support de $h_{0,0}$ est $[0, 1[$. On définit les fonctions $h_{0,j}$, $j \in \mathbb{Z}$, par translation de $h_{0,0}$, de longueur j ; quand j varie dans $\mathbb{Z}$, ces fonctions $(h_{0,j})$ sont à supports disjoints, donc orthogonales ; ensuite on prend tous les changements d'échelle, et on normalise dans L_2, ce qui donne $h_{k,j}(x) = 2^{k/2} \, h_{0,j}(2^k x)$ pour tout $k \in \mathbb{Z}$, c'est à dire

$$\forall x \in \mathbb{R}, \quad h_{k,j}(x) = 2^{k/2} \, h_{0,0}(2^k x - j).$$

Puisque le support de $h_{0,0}$ est $[0,1[$, il en résulte que $h_{k,j}(x)$ est non nul si et seulement si $0 \le 2^k x - j < 1$ ce qui donne $j\,2^{-k} \le x < (j+1)\,2^{-k}$. Le support de $h_{k,j}$ est donc l'intervalle $[j\,2^{-k}, (j+1)\,2^{-k}[$. Notons pour plus tard que la fonction $h_{k,j}$ est à support dans $[0,1[$ si et seulement si $k \ge 0$ et $j = 0, \ldots, 2^k - 1$.

La famille $(h_{k,j})_{k,j\in\mathbb{Z}}$ est une famille orthonormée dans $\mathrm{L}_2(\mathbb{R})$.

Par changement d'échelle et translation, on se ramène à montrer que $h_{0,0}$ est orthogonale à toute fonction $h_{k,j}$, $k \ge 0$ et $(k,j) \ne (0,0)$; si $k = 0$, alors $j \ne 0$ et les deux fonctions $h_{0,0}$ et $h_{0,j}$ sont à supports disjoints, donc orthogonales. Dans le cas où $k > 0$ le support de $h_{k,j}$ est un intervalle de la génération dyadique $k > 0$; il est ou bien contenu dans $[0,1/2[$ ou bien disjoint de $[0,1/2[$; dans les deux cas

$$\int_0^{1/2} h_{k,j} = 0$$

d'où le résultat en faisant de même avec $[1/2,1[$ et en ajoutant,

$$\int h_{0,0} h_{k,j} = \int_0^{1/2} h_{k,j} - \int_{1/2}^1 h_{k,j} = 0.$$

Pour travailler sur $[0,1]$ on ajoute la fonction $1 = h_\emptyset$. On considère donc la famille formée de 1 et des $(h_{k,j})$ pour $k \ge 0$ et $0 \le j < 2^k$; cette famille est orthonormée dans $\mathrm{L}_2(0,1)$, et on a de plus :

la famille formée de 1 et des fonctions $(h_{k,j})$ pour $k \ge 0$ et $0 \le j < 2^k$ est une base orthonormée de $\mathrm{L}_2(0,1)$.

pour montrer cela il faut montrer que l'espace engendré est dense dans $\mathrm{L}_2(0,1)$. Si on désigne par E_n l'espace de dimension 2^n engendré par les fonctions indicatrices des intervalles $[j\,2^{-n}, (j+1)\,2^{-n}[$ pour $j = 0, \ldots, 2^n - 1$, on constate que toutes les fonctions $h_{k,j}$ pour $0 \le k < n$ et la fonction 1 sont dans E_n. Or on a ici $1 + 1 + 2 + \cdots + 2^{n-1} = 2^n$ fonctions linéairement indépendantes, donc elles engendrent E_n. Maintenant, toute fonction continue sur $[0,1]$ peut être approchée uniformément pour n grand par une fonction de E_n, ce qui montre que le système de Haar est total.

Remarquons qu'une fonction $f \in \mathrm{L}_2(0,1)$, d'intégrale nulle, s'exprime avec les fonctions de Haar pures, sans utiliser la constante 1, puisque f est orthogonale à 1 (écrire la série $(*)$). Cette remarque est utile dans le paragraphe qui suit.

On veut maintenant montrer que le système de Haar entier $(h_{k,j})$ avec $k,j \in \mathbb{Z}$ est une base hilbertienne de $\mathrm{L}_2(\mathbb{R})$; pour cela il suffit d'approcher toute fonction à support borné, disons f nulle hors de $[-a,a]$. On découpe le problème en deux intervalles $[-a,0[$ et $[0,a[$. Soient f à support dans $[0,a[$ et m l'intégrale de f ; alors $f_n = f - m\,2^{-n}\mathbf{1}_{[0,2^n[}$ est d'intégrale nulle, et converge vers f dans $\mathrm{L}_2(\mathbb{R})$ quand $n \to +\infty$, donc il suffit de représenter f_n au moyen du système de Haar ; par changement d'échelle à partir de ce qui a été dit pour $[0,1[$ on voit que f_n, étant d'intégrale nulle, est somme d'une série convergente de multiples de fonctions de Haar à support dans $[0,2^n[$.

Convergence faible des suites

On dit que la suite $(x_n) \subset \mathrm{H}$ *tend faiblement* vers $x \in \mathrm{H}$ si on a

$$\langle x, y\rangle = \lim_n \langle x_n, y\rangle$$

pour tout $y \in \mathrm{H}$. Compte-tenu de l'identification du dual de H, ce qui précède est une façon de dire que pour toute forme linéaire continue ℓ sur H, la suite $(\ell(x_n))$ converge vers $\ell(x)$. Cette deuxième façon de dire est la définition générale de la convergence faible des suites, pour un espace normé quelconque.

Bien entendu, la convergence en norme implique la convergence faible.

Toute suite faiblement convergente est bornée ; c'est une conséquence immédiate du théorème de Banach-Steinhaus, appliqué à la suite d'opérateurs linéaires (T_n) de H dans $\mathbb{C}$ définis par $\mathrm{T}_n(y) = \langle y, x_n\rangle$ pour tout $n \geq 0$: puisqu'il y a convergence simple des (T_n) en tout point de l'espace complet H, la suite des normes $(\|\mathrm{T}_n\|)$ est bornée, et pour finir on voit que $\|\mathrm{T}_n\| = \|x_n\|$ pour tout n.

Lorsque la suite (x_n) est contenue dans un convexe *fermé* C, la limite faible reste dans ce convexe ; c'est un cas d'application d'arguments de *séparation*, obtenus en général par le théorème de Hahn-Banach, mais qu'on obtient ici facilement avec le théorème de projection : si on avait $x \notin \mathrm{C}$, on considérerait sa projection y sur C, puis le vecteur $u = x - y \neq 0$ (puisque $x \notin \mathrm{C}$) et alors

$$0 < \|u\|^2 = \langle x - y, u\rangle = \lim_n \langle x_n - y, u\rangle = \lim_n \mathrm{Re}\,\langle x_n - y, u\rangle \leq 0$$

par les propriétés de la projection ; on a obtenu une contradiction, donc $x \in \mathrm{C}$.

Compacité faible de la boule unité d'un Hilbert

Théorème. *Toute suite bornée (x_n) dans un espace de Hilbert admet des sous-suites faiblement convergentes.*

Démonstration. Soit (x_n) une suite de vecteurs de la boule unité fermée $\mathrm{B_H}$ d'un espace de Hilbert H ; le sous-espace fermé H_0 engendré par cette suite est séparable donc il admet une base hilbertienne dénombrable $(e_k)_{k\in\mathbb{N}}$. On considère l'application φ de $\mathrm{B_H}$ dans le compact $\mathrm{K} = \overline{\mathrm{U}}^{\mathbb{N}}$ donnée par

$$\forall y \in \mathrm{B_H}, \quad \varphi(y) = (\langle y, e_k\rangle)_{k\in\mathbb{N}}$$

où on a noté U le disque unité ouvert dans $\mathbb{C}$. On a une suite $u_n = \varphi(x_n)$ dans le compact métrisable K, dont on peut extraire une sous-suite (u_{n_j}) qui converge pour la topologie produit, c'est à dire que pour tout $k \geq 0$ il existe un nombre $z_k \in \overline{\mathrm{U}}$ tel que

$$z_k = \lim_j \langle x_{n_j}, e_k\rangle\,;$$

cette convergence et le fait que $\|x_n\| \leq 1$ pour tout n impliquent que $\sum_k |z_k|^2 < +\infty$, donc $x = \sum_k z_k e_k$ existe dans H et on vérifie la convergence faible de (x_{n_j}) vers x.

Il en résulte que *l'image de la boule unité fermée par un opérateur compact partant d'un Hilbert* H *est fermée*, donc compacte dans l'espace de Banach d'arrivée F.

En effet si $(x_n) \subset \mathrm{B_H}$ et si $\mathrm{T}x_n \to y \in \mathrm{F}$, on extrait de la suite (x_n) une sous-suite qui tend faiblement vers un $x \in \mathrm{B_H}$; alors $\mathrm{T}x_{n_k}$ tend faiblement vers $\mathrm{T}x$, mais aussi vers y, d'où $y = \mathrm{T}x$ (si on veut le dire avec un Banach F à l'arrivée, on a besoin d'un minimum de Hahn-Banach pour montrer l'unicité de la limite faible d'une suite ; si on se contente du cas où F est un Hilbert, le théorème de projection suffit).

Opérateur adjoint

Si T est une application linéaire continue de H_1 dans H_2 (deux espaces de Hilbert) il existe une unique application linéaire continue $\mathrm{T}^* \in \mathcal{L}(\mathrm{H}_2, \mathrm{H}_1)$ telle que

$$\forall x_1 \in \mathrm{H}_1,\ \forall x_2 \in \mathrm{H}_2, \quad \langle \mathrm{T}x_1, x_2\rangle = \langle x_1, \mathrm{T}^* x_2\rangle.$$

On dit que $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ est *hermitien* quand $\mathrm{T}^* = \mathrm{T}$, ce qui équivaut à dire que

$$\langle \mathrm{T}x, y\rangle = \langle x, \mathrm{T}y\rangle$$

pour tous vecteurs $x, y \in \mathrm{H}$.

Opérateurs hermitiens compacts

Lemme. *Soient* H *un espace de Hilbert et* T *un opérateur hermitien et compact de* H *dans* H *; pour tout sous-espace fermé non nul* $\mathrm{E} \subset \mathrm{H}$ *stable par* T*, il existe un vecteur propre* $u \in \mathrm{E}$ *pour* T *(c'est à dire un vecteur* $u \in \mathrm{E}$ non nul *tel que* $\mathrm{T}u \in \mathbb{K}\,u$*).*

Démonstration. Puisque T est hermitien, la quantité

$$f(x) = \langle x, \mathrm{T}x\rangle = \langle \mathrm{T}x, x\rangle = \overline{\langle x, \mathrm{T}x\rangle}$$

est réelle pour tout $x \in \mathrm{E}$. Si $f(x) = 0$ est nul pour tout $x \in \mathrm{E}$, on en déduit que $\langle \mathrm{T}x, y\rangle = 0$ pour tous $x, y \in \mathrm{E}$, par polarisation (calculer

$$0 = f(x+y) - f(x-y) = 4 \operatorname{Re} \langle \mathrm{T}x, y\rangle$$

et appliquer le lemme 0) ; ceci implique que $\mathrm{T}x = 0$ pour tout $x \in \mathrm{E}$ (appliquer à $y = \mathrm{T}x \in \mathrm{E}$) ; on prend alors $u \in \mathrm{E}$ non nul quelconque (ce qui est possible puisque $\mathrm{E} \neq \{0\}$) et on a $\mathrm{T}u = 0 \in \mathbb{K}u$.

Dans le cas contraire il existe $x_0 \in \mathrm{E}$ tel que $f(x_0) \neq 0$, par exemple tel que $f(x_0) > 0$. Dans ce cas, on maximise la fonction f sur la boule unité fermée $\mathrm{B_E}$ de E ; le maximum est atteint grâce à la compacité de T : si $(x_n) \subset \mathrm{B_E}$ est une suite maximisante pour f sur $\mathrm{B_E}$, on prend une sous-suite (x_{n_j}) telle que $\mathrm{T}x_{n_j} \to y$ pour un certain $y \in \mathrm{E}$ (compacité de l'image de la boule unité), et on peut supposer aussi que (x_{n_j}) converge faiblement vers un $x \in \mathrm{B_E}$ (convexe fermé, donc faiblement fermé) ; on sait alors que $y = \mathrm{T}x$. Alors $|f(x_{n_j}) - \langle y, x_{n_j}\rangle| \le \|\mathrm{T}(x_{n_j}) - y\|$ tend vers 0 et

$$f(x) = \langle \mathrm{T}x, x\rangle = \langle y, x\rangle = \lim_j \langle y, x_{n_j}\rangle = \lim_j f(x_{n_j})$$

donc $f(x) = \lim_n f(x_n) = \sup f(\mathrm{B_E}) \ge f(x_0) > 0$ et on a montré que $f(x)$ est la valeur maximale de f sur $\mathrm{B_E}$.

Pour tout vecteur $z \in \mathrm{E}$ orthogonal à x on pose

$$\forall t \in \mathbb{R}, \quad x(t) = \cos(t)\,x + \sin(t)\,z.$$

On obtient pour tout t un vecteur de la boule unité de E, donc $f(x(t)) \le f(x) = f(x(0))$. En annulant la dérivée en t pour $t = 0$ de la fonction $t \to f(x(t))$ on obtient

$$\operatorname{Re} \langle \mathrm{T}x, z\rangle = 0.$$

D'après le lemme 0 on en déduit que $\mathrm{T}x$ est orthogonal à l'espace des vecteurs z de E qui sont orthogonaux à x ; puisque $\mathrm{T}x \in \mathrm{E}$, ceci implique que $\mathrm{T}x \in \mathbb{K}x$. Par ailleurs $x \neq 0$ puisque $f(x) > 0$.

Théorème. *Soient* H *un espace de Hilbert et* T *un opérateur hermitien et compact de* H *dans* H*; il existe une base orthonormée de* H *formée de vecteurs propres de* T*. Les valeurs propres de* T *sont réelles; les valeurs propres non nulles peuvent être rangées dans une suite qui tend vers* 0 *(ou sinon elles sont en nombre fini).*

Démonstration. Considérons un ensemble $\mathrm{A} \subset \mathrm{H}$ maximal avec les propriétés suivantes :

– si $x \in \mathrm{A}$, alors $\|x\| = 1$ et $\mathrm{T}x \in \mathbb{K}x$;
– si $x, y \in \mathrm{A}$ et $x \neq y$, alors $x \perp y$.

Il est clair que l'ensemble $\mathcal{A}$ des parties A avec ces deux propriétés, ordonné par inclusion, est inductif, donc il existe des éléments maximaux. Montrons qu'un tel A maximal est *total* : soient $\mathrm{F} = \operatorname{Vect} \mathrm{A}$ et E l'orthogonal de F ; on veut montrer que $\mathrm{E} = \{0\}$, ce qui signifie que F est dense.

Tout d'abord, E est stable par T (voir plus loin), ce qui permet d'appliquer le lemme précédent : si on avait $\mathrm{E} \neq \{0\}$ on trouverait d'après ce lemme un vecteur de norme un $u \in \mathrm{E}$ tel que $\mathrm{T}u \in \mathbb{K}u$, et $\mathrm{A} \cup \{u\}$ serait un élément de $\mathcal{A}$ qui contredirait la maximalité de A. On a donc $\mathrm{E} = \{0\}$ et A est total ; c'est donc une base hilbertienne de H, formée de vecteurs propres de T.

Pour montrer que $\mathrm{E} = \mathrm{F}^{\perp}$ est stable par T on observe d'abord que F est stable par T, ce qui est clair puisque F est engendré par des vecteurs propres de T ; ensuite, si $x \perp \mathrm{F}$ on aura pour tout $y \in \mathrm{F}$

$$\langle \mathrm{T}x, y \rangle = \langle x, \mathrm{T}y \rangle = 0$$

puisque $\mathrm{T}y \in \mathrm{F}$; on a bien que $\mathrm{T}x \in \mathrm{F}^{\perp}$, et on a montré que $\mathrm{T}(\mathrm{F}^{\perp}) \subset \mathrm{F}^{\perp}$.

Les valeurs propres sont réelles parce que $\langle \mathrm{T}x, x \rangle$ est réel pour tout x, en particulier pour tout vecteur propre.

On a donc une base hilbertienne $(e_i)_{i \in \mathrm{I}}$ de H telle que $\mathrm{T}e_i = \lambda_i e_i$ pour tout $i \in \mathrm{I}$; si H est séparable on peut garantir que I est fini ou dénombrable ; pour finir on montre que pour tout $\varepsilon > 0$, l'ensemble $\mathrm{J}_\varepsilon \subset \mathrm{I}$ des indices $i \in \mathrm{I}$ tels que $|\lambda_i| \geq \varepsilon$ est *fini* ; en effet, pour tous ces indices $i, j \in \mathrm{J}_\varepsilon$ on a quand $i \neq j$

$$\|\mathrm{T}e_i - \mathrm{T}e_j\| = \sqrt{\lambda_i^2 + \lambda_j^2} \geq \sqrt{2}\,\varepsilon$$

quand $i \neq j$; dans le compact $\mathrm{T}(\mathrm{B_H})$ cette configuration “écartée” n'est possible que pour un nombre fini de points. La finitude de J_ε implique que l'on peut ranger les valeurs propres non nulles dans une suite qui tend vers 0.

Théorème. *Pour tout opérateur compact* T *sur un espace de Banach* E *et pour tout* $\lambda \neq 0$ *le sous-espace propre* $\ker(\mathrm{T} - \lambda \,\mathrm{Id})$ *est de dimension finie.*

Démonstration. C'est pour ça que le théorème de Riesz est fait. Soient $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ compact et $\mathrm{F} = \ker(\mathrm{T} - \lambda \,\mathrm{Id})$, avec $\lambda \neq 0$; puisque $\mathrm{T}y = \lambda y$ pour tout $y \in \mathrm{F}$ on a $\mathrm{T}(\mathrm{B_F}) = \lambda \mathrm{B_F}$, donc puisque $\lambda \neq 0$ on a

$$\mathrm{B_F} = \lambda^{-1}\mathrm{T}(\mathrm{B_F}) \subset \lambda^{-1}\mathrm{T}(\mathrm{B_E})$$

qui est relativement compact, donc $\mathrm{B_F}$ est relativement compacte et il en résulte que $\dim \mathrm{F} < +\infty$ par le théorème de Riesz.

Prépa. Agrég. écrit d'Analyse, séance n$^{\text{o}}$ 16, Mercredi 21 Mars 2001.

Rappels Hilbert 2

Correction d'exercice. Etant donnée une suite numérique $\alpha = (\alpha_n)_{n\ge 0}$ on définit un opérateur Δ_α de $\ell_p = \ell_p(\mathbb{N})$ dans lui-même par

$$\forall \mathbf{x} = (x_n)_{n\ge 0} \in \ell_p, \quad \Delta_\alpha(\mathbf{x}) = (\alpha_n x_n)_{n\ge 0}.$$

Il est clair que $\|\Delta_\alpha\| \le \|\alpha\|_\infty$, puisque (dans le cas $p < +\infty$) on a pour tout $\mathbf{x} \in \ell_p$

$$\|\Delta_\alpha(\mathbf{x})\|_p^p = \sum_{n\ge 0} |\alpha_n x_n|^p \le \big(\sup_{n\ge 0} |\alpha_n|^p\big) \sum_{n\ge 0} |x_n|^p = \|\alpha\|_\infty^p \ \|\mathbf{x}\|_p^p$$

et en regardant l'action sur les vecteurs "de base" $\mathbf{e}_n$ de ℓ_p on se convainc aisément que $\|\Delta_\alpha\| = \|\alpha\|_\infty$ (le vecteur $\mathbf{e}_n$ a toutes ses coordonnées nulles sauf la $n$ième qui vaut 1 ; c'est un vecteur de norme 1, donc $\|\Delta_\alpha\| \ge \|\Delta_\alpha(\mathbf{e}_n)\|_p = |\alpha_n|$ pour tout $n \ge 0$). De plus, l'application qui à $\alpha \in \ell_\infty$ associe $\Delta_\alpha \in \mathcal{L}(\ell_p)$ est linéaire.

Supposons maintenant que $\alpha \in c_0$ c'est à dire $\lim_n \alpha_n = 0$. Pour chaque entier $k > 0$ découpons α en $\alpha = \beta^{(k)} + \gamma^{(k)}$ où $\beta^{(k)}_n = \alpha_n$ si $n \le k$ et $\beta^{(k)}_n = 0$ si $n > k$. La suite $\beta^{(k)}$ est nulle à partir du rang $k+1$, ce qui entraîne que l'opérateur $\Delta_{\beta^{(k)}}$ associé est de rang $\le k+1$ (son image est engendrée par les vecteurs $\mathbf{e}_0, \ldots, \mathbf{e}_k$). Par ailleurs $\|\gamma^{(k)}\|_\infty = \max_{n>k} |\alpha_n|$ tend vers 0 quand $k \to +\infty$, donc

$$\lim_k \|\Delta_\alpha - \Delta_{\beta^{(k)}}\| = \lim_k \|\Delta_{\gamma^{(k)}}\| = \lim_k \|\gamma^{(k)}\|_\infty = 0,$$

donc Δ_α est compact comme limite (en norme d'opérateur) d'une suite d'opérateurs de rang fini.

Inversement, si la suite (α_n) ne tend pas vers 0, l'opérateur Δ_α n'est pas compact : il existe un effet $\delta > 0$ et une sous-suite (α_{n_j}) telle que $|\alpha_{n_j}| \ge \delta$ pour tout indice j ; si $i \ne j$, on voit aisément que $\|\Delta_\alpha \mathbf{e}_{n_i} - \Delta_\alpha \mathbf{e}_{n_j}\|_p \ge \delta$, ce qui montre que l'image par Δ_α de la boule unité de ℓ_p contient une infinité de points à distance mutuelles $\ge \delta > 0$, donc l'image de la boule unité n'est pas relativement compacte dans ℓ_p.

En conclusion, l'opérateur Δ_α est compact de ℓ_p dans ℓ_p si et seulement si $\alpha \in c_0$.

Complément à la diagonalisation des hermitiens compacts

Définition. On dit qu'un endomorphisme $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ est *normal* si $\mathrm{T}^*\mathrm{T} = \mathrm{T}\,\mathrm{T}^*$.

Théorème. *Soient* H *un espace de Hilbert* **complexe** *et* T *un opérateur normal et compact de* H *dans* H *; il existe une base orthonormée de* H *formée de vecteurs propres de* T*. Les valeurs propres non nulles peuvent être rangées dans une suite qui tend vers* 0 *(ou sinon elles sont en nombre fini).*

Esquisse. On peut considérer l'opérateur hermitien compact $\mathrm{A} = \mathrm{T}^*\mathrm{T}$. Il existe une décomposition de H en somme hilbertienne $\mathrm{H} = \sum_{n=0}^{+\infty} \mathrm{E}_n$ de sous-espaces propres de A. Supposons que $\mathrm{E}_0 = \ker \mathrm{A}$; il est facile de vérifier que $\ker \mathrm{A} = \ker \mathrm{T}$. Pour $n > 0$, l'espace E_n est alors le sous-espace propre de A pour une certaine valeur $\lambda \ne 0$; on sait que le sous-espace propre est de dimension finie dans ce cas (théorème de Riesz) et il est stable par T puisque T commute avec A ; on est alors ramené à un problème de diagonalisation en dimension finie, et avec un espace complexe, pour la restriction de T à E_n. Dans le meilleur des cas, on a vu ce théorème en DEUG, sinon on bricolera une preuve ainsi :

soit λ une valeur propre non nulle de $\mathrm{A} = \mathrm{T}^*\mathrm{T}$ et soit $\mathrm{E} = \ker(\mathrm{A} - \lambda \,\mathrm{Id}_\mathrm{H})$ le sous-espace propre correspondant ; puisque $\dim \mathrm{E} < +\infty$, que $\mathrm{T}(\mathrm{E}) \subset \mathrm{E}$ et puisque les scalaires sont complexes, il existe au moins un vecteur non nul $x \in \mathrm{E}$ et un scalaire $\mu \in \mathbb{C}$ tels que $\mathrm{T}(x) = \mu$. On voit que $\ker(\mathrm{T} - \mu \,\mathrm{Id}_\mathrm{H}) = \ker(\mathrm{T}^* - \overline{\mu} \,\mathrm{Id}_\mathrm{H})$ en utilisant le fait que T est normal ; il en résulte que $|\mu|^2 = \lambda$ et que $\mathrm{F} = \ker(\mathrm{T} - \mu \,\mathrm{Id})$ est contenu dans E ; de plus, puisque F est stable par T^*, son orthogonal $\mathrm{F}^\perp$ est stable par T, et on recommence le raisonnement précédent avec le nouveau sous-espace $\mathrm{F}^\perp \cap \mathrm{E}$ stable par T. On peut continuer, jusqu'à ce que E soit décomposé en sous-espaces propres de T, deux à deux orthogonaux par construction.

Finalement on décomposera chaque sous-espace propre de A en sous-espaces propres de T, d'où le résultat.

Remarque. Ce résultat n'est évidemment pas vrai en scalaires réels : considérer une rotation dans $\mathbb{R}^2$, d'angle non multiple de π. C'est un opérateur normal, mais il n'admet aucun vecteur propre.

Opérateurs à noyau de Hilbert-Schmidt

On considère deux espaces mesurés σ-finis $(\mathrm{X}, \mathcal{A}, \mu)$ et $(\mathrm{Y}, \mathcal{B}, \nu)$ et on considère une fonction complexe $\mathrm{K} \in \mathrm{L}_2(\mathrm{X} \times \mathrm{Y}, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$. Pour chaque $f \in \mathrm{L}_2(\mathrm{Y}, \nu)$ considérons l'intégrale positive

$$(*) \qquad \Bigl(\int_\mathrm{Y} |\mathrm{K}(x,y)|\,|f(y)|\,d\nu(y)\Bigr)^2 \le \Bigl(\int_\mathrm{Y} |\mathrm{K}(x,y)|^2\,d\nu(y)\Bigr)\Bigl(\int_\mathrm{Y} |f(y)|^2\,d\nu(y)\Bigr)$$

et en appliquant Fubini on aura

$$\int_\mathrm{X}\Bigl(\int_\mathrm{Y} |\mathrm{K}(x,y)|\,|f(y)|\,d\nu(y)\Bigr)^2 d\mu(x) \le \Bigl(\int_{\mathrm{X}\times\mathrm{Y}} |\mathrm{K}(x,y)|^2\,d\mu(x)\,d\nu(y)\Bigr)\ \|f\|_2^2 < +\infty.$$

Ce calcul préliminaire montre que pour μ-presque tout $x \in \mathrm{X}$, la fonction $y \in \mathrm{Y} \to \mathrm{K}(x,y)\,f(y)$ est ν-intégrable, ce qui permet d'introduire une fonction $\mathrm{T_K}f$ sur X définie pour μ-presque tout $x \in \mathrm{X}$ par la formule

$$(\mathrm{T_K}f)(x) = \int_\mathrm{Y} \mathrm{K}(x,y)f(y)\,d\nu(y).$$

Il est clair que $\mathrm{T_K}$ est linéaire, à valeurs dans l'espace vectoriel des fonctions mesurables sur X, mais de plus le calcul préliminaire montre que

$$\int_\mathrm{X} |(\mathrm{T_K}f)(x)|^2\,d\mu(x) \le \int_\mathrm{X}\Bigl(\int_\mathrm{Y} |\mathrm{K}(x,y)|\,|f(y)|\,d\nu(y)\Bigr)^2 d\mu(x) \le \|\mathrm{K}\|^2_{\mathrm{L}_2(\mathrm{X}\times\mathrm{Y})}\ \|f\|_2^2 < +\infty$$

autrement dit, $\mathrm{T_K}$ définit un opérateur linéaire borné de $\mathrm{L}_2(\mathrm{Y}, \nu)$ dans $\mathrm{L}_2(\mathrm{X}, \nu)$ et

$$\|\mathrm{T_K}\|_{\mathcal{L}(\mathrm{L}_2(\mathrm{Y}),\mathrm{L}_2(\mathrm{X}))} \le \|\mathrm{K}\|_{\mathrm{L}_2(\mathrm{X}\times\mathrm{Y})}.$$

De plus, $\mathrm{T_K}$ est compact ; soit (f_n) une suite dans la boule unité de $\mathrm{L}_2(\mathrm{Y}, \nu)$, et montrons que $(\mathrm{T_K}f_n)$ admet une sous-suite convergente en norme dans $\mathrm{L}_2(\mathrm{X}, \mu)$. Quitte à passer à une première sous-suite, on peut supposer que (f_n) converge faiblement vers une fonction f de la boule unité de $\mathrm{L}_2(\mathrm{Y}, \nu)$. L'inégalité $(*)$ montre que pour tout $x \in \mathrm{X}$,

$$|\mathrm{T_K}(f_n - f)(x)|^2 \le g(x)\,\|f_n - f\|_2^2,$$

où $g(x) = \int_{\mathrm{Y}} |\mathrm{K}(x,y)|^2\, d\nu(y)$ est une fonction μ-intégrable sur X. Par ailleurs, pour μ-presque tout $x \in \mathrm{X}$ fixé, $\mathrm{T_K}(f_n - f)(x)$ peut être considéré comme le produit scalaire dans $\mathrm{L}_2(\mathrm{Y}, \nu)$ de la fonction $f_n - f$ avec la fonction $y \to \overline{\mathrm{K}(x,y)}$, donc ce produit scalaire tend vers 0 par définition de la convergence faible de (f_n) vers f. La suite $|\mathrm{T_K}(f_n - f)|^2$ tend donc simplement vers 0 presque partout, en étant dominée par une fonction intégrable fixe. Par le théorème de convergence dominée, il en résulte que

$$\int_{\mathrm{X}} |\mathrm{T_K}(f_n - f)(x)|^2\, d\mu(x) = \|\mathrm{T_K} f_n - \mathrm{T_K} f\|_2^2 \to 0,$$

ce qu'il fallait démontrer.

On remarquera avant de continuer que la correspondance $\mathrm{K} \in \mathrm{L}_2(\mathrm{X} \times \mathrm{Y}) \to \mathrm{T_K} \in \mathcal{L}(\mathrm{L}_2(\mathrm{Y}), \mathrm{L}_2(\mathrm{X}))$ est injective. En effet, si $\mathrm{T_K} = 0$ on aura, en calculant $\langle \mathbf{1}_{\mathrm{A}}, \mathrm{T_K} \mathbf{1}_{\mathrm{B}} \rangle = 0$ par Fubini

$$\int_{\mathrm{X} \times \mathrm{Y}} \mathbf{1}_{\mathrm{A}}(x) \mathbf{1}_{\mathrm{B}}(y) \mathrm{K}(x,y)\, d\mu(x)\, d\nu(y) = 0,$$

pour tous $\mathrm{A} \in \mathcal{A}$ et $\mathrm{B} \in \mathcal{B}$. Ceci montre que K est orthogonale à toutes les $\mathbf{1}_{\mathrm{A} \times \mathrm{B}}$, dont on sait qu'elles sont totales dans $\mathrm{L}_2(\mathrm{X} \times \mathrm{Y}, \mu \otimes \nu)$. Il en résulte bien que $\mathrm{K} = 0$.

Il est facile de voir avec Fubini que l'opérateur adjoint $(\mathrm{T_K})^*$ est l'opérateur à noyau $\mathrm{T_{K^*}}$ où K^* est la fonction de carré intégrable sur $\mathrm{Y} \times \mathrm{X}$ définie par

$$\forall (y,x) \in \mathrm{Y} \times \mathrm{X}, \quad \mathrm{K}^*(y,x) = \overline{\mathrm{K}(x,y)}.$$

Dans le cas où $\mathrm{X} = \mathrm{Y}$ et $\mu = \nu$, on en déduit que $\mathrm{T_K}$ est hermitien si (et seulement si, par l'injectivité mentionnée ci-dessus) le noyau K possède la *symétrie hermitienne*

$$\forall (x,y) \in \mathrm{X} \times \mathrm{X}, \quad \mathrm{K}(y,x) = \overline{\mathrm{K}(x,y)}.$$

On aura alors un opérateur $\mathrm{T_K}$ hermitien et compact, qui admettra donc une base hilbertienne de vecteurs propres.

Cas particulier : $\mathrm{X} = \mathrm{Y} = \mathbb{N}$, et $\mu = \nu$ est la mesure de comptage $\mathbf{c}$, définie par $\mathbf{c}(\{n\}) = 1$ pour tout $n \in \mathbb{N}$; dans ce cas la “fonction” K sur $\mathbb{N} \times \mathbb{N}$ s'appelle plutôt une suite double $k = (k_{m,n})$, ou *matrice infinie*. Si la matrice k vérifie $\sum_{m,n} |k_{m,n}|^2 < +\infty$, elle permet donc de définir un opérateur linéaire borné T_k de $\ell_2 = \mathrm{L}_2(\mathbb{N}, \mathbf{c})$ dans lui-même, par une formule qui étend naturellement le calcul matriciel,

$$\forall \mathbf{x} = (x_n)_{n \ge 0} \in \ell_2,\ \forall m \ge 0, \quad (\mathrm{T}_k \mathbf{x})_m = \sum_{n=0}^{+\infty} k_{m,n}\, x_n.$$

On expliquera plus loin que le choix de bases hilbertiennes dans L_2 permet de ramener l'étude d'un $\mathrm{T_K}$ quelconque au cas d'un T_k matriciel ; le point de vue matriciel est le point de vue originel de Hilbert, vers 1900 ; le point de vue des noyaux de carré intégrable a été développé par son élève Schmidt ; les opérateurs à noyau (continus, ou au moins continus par morceaux) existaient déjà depuis la fin du 19ème siècle (Volterra).

Correction d'un exercice proposé

On étudie l'opérateur T défini sur $\mathrm{L}_2(0,\pi/2)$ par

$$\forall f \in \mathrm{L}_2(0,\pi/2), \quad (\mathrm{T}f)(x) = \sin(x)\int_0^x \cos(y)f(y)\,dy + \cos(x)\int_x^{\pi/2} \sin(y)f(y)\,dy.$$

On repère tout de suite qu'il s'agit d'un opérateur à noyau T_K, avec

$$\begin{aligned} \mathrm{K}(x,y) &= \sin(x)\cos(y) \ \text{ si } \ 0 \le y \le x \le \pi/2, \\ \mathrm{K}(x,y) &= \sin(y)\cos(x) \ \text{ si } \ 0 \le x \le y \le \pi/2. \end{aligned}$$

Le noyau est réel et symétrique, donc possède la symétrie hermitienne et $\mathrm{T} = \mathrm{T}_\mathrm{K}$ est un opérateur hermitien et compact de l'espace de Hilbert $\mathrm{H} = \mathrm{L}_2(0,\pi/2)$ (pour la mesure de Lebesgue).

Une première remarque est que $\mathrm{T}f$ est une fonction continue, pour toute $f \in \mathrm{H}$; c'est clair sur la formule de définition avec des intégrales fonctions de leur borne (ajouter un zeste de convergence dominée). Mais si f est déjà continue, on peut dire beaucoup plus : on sait depuis le DEUG que $\mathrm{T}f$ est alors dérivable, et

$$\begin{aligned} (\mathrm{T}f)'(x) = \cos(x)&\int_0^x \cos(y)f(y)\,dy + \sin(x)\cos(x)f(x) \\ &- \sin(x)\int_x^{\pi/2} \sin(y)f(y)\,dy - \cos(x)\sin(x)f(x)\,; \end{aligned}$$

il y a un petit miracle, puisque la formule

$$(**) \qquad (\mathrm{T}f)'(x) = \cos(x)\int_0^x \cos(y)f(y)\,dy - \sin(x)\int_x^{\pi/2} \sin(y)f(y)\,dy$$

montre que $(\mathrm{T}f)'$ a une expression similaire à celle de $\mathrm{T}f$, ce qui permet, toujours lorsque f est continue, de dériver une fois de plus,

$$\begin{aligned} (\mathrm{T}f)''(x) = -\sin(x)&\int_0^x \cos(y)f(y)\,dy + \cos(x)^2 f(x) \\ &- \cos(x)\int_x^{\pi/2} \sin(y)f(y)\,dy + \sin^2(x)f(x) \end{aligned}$$

ce qui donne

$$(\mathrm{T}f)''(x) = -(\mathrm{T}f)(x) + f(x).$$

On voit donc que $g = \mathrm{T}f$ est une fonction de classe C^2 qui vérifie l'équation différentielle $g'' + g = f$. De plus l'équation $(**)$ sur $g' = (\mathrm{T}f)'$ montre que $g'(0) = 0$ et $g'(\pi/2) = 0$. Mais il y a une seule solution de l'équation différentielle vérifiant ces conditions aux limites (la différence h de deux solutions g_1 et g_2 vérifierait $h'' + h = 0$ et $h'(0) = 0$, $h'(\pi/2) = 0$. On sait que $h(x) = \mathrm{A}\cos x + \mathrm{B}\sin x$, et les conditions aux limites donnent $\mathrm{A} = \mathrm{B} = 0$). Résumons,

lorsque la fonction f est continue, $\mathrm{T}f$ *est l'unique solution g de l'équation différentielle $g'' + g = f$ qui vérifie les conditions $g'(0) = g'(\pi/2) = 0$.*

On va maintenant identifier les valeurs propres non nulles de T ; on sait déjà qu'elles sont réelles. Si $\lambda \in \mathbb{R}$ est valeur propre non nulle de T, il existe $f \in \mathrm{H}$ de norme un telle que $\mathrm{T}f = \lambda f$. Mais alors la relation $f = \lambda^{-1}\mathrm{T}f$ montre que f est continue, donc $g = \mathrm{T}f$ est de classe C^2 et vérifie l'équation différentielle $g'' + g = f$ et les conditions $g'(0) = g'(\pi/2) = 0$. En revenant à $f = \lambda^{-1}g$ qui est aussi de classe C^2 on aura

$$f'' = (\lambda^{-1} - 1)\, f$$

et $f'(0) = f'(\pi/2) = 0$. Supposons d'abord que $\lambda^{-1} - 1 = \omega^2 > 0$. Les solutions sont de la forme $f(x) = \mathrm{A}\,\mathrm{ch}(\omega x) + \mathrm{B}\,\mathrm{sh}(\omega x)$, mais les conditions aux limites imposent $\mathrm{B} = 0$, puis $\mathrm{A}\,\mathrm{sh}(\omega\pi/2) = 0$ qui entraîne $\mathrm{A} = 0$ et $f = 0$, contrairement à l'hypothèse $\|f\|_2 = 1$. Ce cas est donc impossible.

La deuxième possibilité est $\lambda^{-1} - 1 = 0$, c'est à dire $\lambda = 1$. Dans ce cas l'équation devient $f'' = 0$, qui admet pour solutions les fonctions affines $\mathrm{A} + \mathrm{B}x$; la condition aux limites impose $\mathrm{B} = 0$, et la normalisation donne $f = \sqrt{2/\pi}$.

La dernière possibilité est $\lambda^{-1} - 1 = -\omega^2 < 0$. Dans ce cas les solutions sont de la forme $f(x) = \mathrm{A}\cos(\omega x) + \mathrm{B}\sin(\omega x)$, et les conditions aux limites donnent d'abord $\mathrm{B} = 0$, puis $\sin(\omega\pi/2) = 0$. Ceci n'est possible que si $\omega = 2k$ avec k entier ≥ 1. La valeur correspondante pour λ est $\lambda_k = (1 - 4k^2)^{-1}$. Pour vérifier que $f(x) = \cos(2kx)$ est effectivement vecteur propre de T, on note que $g = f$ est de classe C^2 et vérifie $g'' + g = (-4k^2 + 1)g = \lambda_k^{-1}g$, ce qui prouve que $\mathrm{T}(\lambda_k^{-1}g) = g$, ou bien $\mathrm{T}(f) = \lambda_k f$.

Si on veut normaliser le vecteur propre on posera

$$f_k(x) = \frac{2}{\sqrt{\pi}} \cos(2kx).$$

Pour les valeurs propres, on peut inclure $\lambda = 1$ dans la liste en disant que les valeurs propres non nulles de T sont les $\lambda_k = (1 - 4k^2)^{-1}$ pour $k = 0, 1, 2, \ldots$

On a identifié tous les vecteurs propres correspondant à des valeurs propres non nulles ; il reste à s'occuper du noyau de T. On sait que le noyau de T est l'orthogonal de l'image de T ; en effet la relation

$$\forall y \in \mathrm{H}, \quad \langle x, \mathrm{T}y\rangle = 0$$

qui dit que x est orthogonal à l'image de T équivaut, puisque T est hermitien à

$$\forall y \in \mathrm{H}, \quad \langle \mathrm{T}x, y\rangle = 0$$

qui signifie que $\mathrm{T}x = 0$. On va voir que l'image de T est dense, donc le noyau nul. Si g est une fonction de classe C^2 à support dans $[\varepsilon, \pi/2 - \varepsilon]$ pour un $\varepsilon \in]0, \pi/4[$ et si on pose $f = g'' + g$, la fonction g vérifie bêtement l'équation $g'' + g = f$, ainsi que les conditions aux limites $g'(0) = g'(\pi/2) = 0$, donc $g = \mathrm{T}(f)$ est dans l'image de T. Il est facile de voir, à partir des résultats classiques de densité, que cet ensemble de fonctions g est dense dans $\mathrm{L}_2(0, \pi/2)$.

Opérateurs de Hilbert-Schmidt et bases hilbertiennes

Si $(g_n)_{n\geq 0}$ est une base hilbertienne de $\mathrm{L}_2(\mathrm{Y})$, on note que les fonctions conjuguées $(\overline{g_n})_{n\geq 0}$ forment une autre base hilbertienne. Ensuite, si $(f_n)_{n\geq 0}$ est une base de $\mathrm{L}_2(\mathrm{X})$, la suite double $h_{m,n}(x, y) = f_m(x)\overline{g_n(y)}$, $(m, n) \in \mathbb{N} \times \mathbb{N}$, est une base hilbertienne de $\mathrm{L}_2(\mathrm{X} \times \mathrm{Y})$. On notera $f_m \otimes \overline{g_n}$ la fonction définie par $f_m \otimes \overline{g_n}(x, y) = f_m(x)\,\overline{g_n(y)}$.

Etant donné un noyau K dans $\mathrm{L}_2(\mathrm{X} \times \mathrm{Y})$, il est alors possible de le décomposer sur cette base,

$$\mathrm{K} = \sum_{m,n} k_{m,n}\, f_m \otimes \overline{g_n},$$

et $\sum_{m,n\geq 0}|k_{m,n}|^2 = \|\mathrm{K}\|_2^2$. Pour μ-presque tout $x \in \mathrm{X}$ l'expression

$$\sum_{n=0}^{+\infty}\Bigl(\sum_{m=0}^{+\infty} k_{m,n} f_m(x)\Bigr)\,\overline{g_n}$$

est la décomposition de $y \to \mathrm{K}(x,y)$ dans la base $(\overline{g_n})$, donc par orthogonalité

$$(\mathrm{T_K} g_{n_0})(x) = \sum_{n=0}^{+\infty}\Bigl(\sum_{m=0}^{+\infty} k_{m,n} f_m(x)\Bigr)\int_{\mathrm{Y}} \overline{g_n}(y) g_{n_0}(y)\,d\nu(y) = \sum_{m=0}^{+\infty} k_{m,n_0} f_m(x).$$

Considérons l'isométrie surjective (opérateur unitaire) U de $\mathrm{L}_2(\mathrm{X})$ sur ℓ_2 qui associe à toute $f \in \mathrm{L}_2(\mathrm{X})$ la suite $(\langle f, f_n\rangle)$ de ses coefficients dans la base (f_n), et l'isométrie surjective V de $\mathrm{L}_2(\mathrm{Y})$ sur ℓ_2 qui associe à toute $g \in \mathrm{L}_2(\mathrm{Y})$ la suite $(\langle g, g_n\rangle)$ de ses coefficients dans la base (g_n).

Comparons $\mathrm{T_K}$ et $\mathrm{U^* T}_k \mathrm{V}$; il suffit de tester sur les vecteurs de la base (g_n). L'opérateur V envoie g_{n_0} sur le vecteur $\mathbf{e}_{n_0}$ de la base hilbertienne canonique de ℓ_2. Ensuite T_k envoie $\mathbf{e}_{n_0}$ sur la suite $(k_{m,n_0})_m$, et U^* pour terminer envoie cette suite sur $\sum_m k_{m,n_0} f_m$: c'est bien $\mathrm{T_K}(g_{n_0})$, ce qui montre que $\mathrm{T_K} = \mathrm{U}^*\,\mathrm{T}_k\,\mathrm{V}$.

C'est ce que nous disions précédemment sur le point de vue de Hilbert : on peut toujours ramener $\mathrm{T_K}$ à l'étude d'un opérateur sur ℓ_2 défini par une matrice infinie.

Dans le cas où $\mathrm{X} = \mathrm{Y}$ et $\mu = \nu$, on prendra aussi de préférence $f_n = g_n$, et on aura donc $\mathrm{U} = \mathrm{V}$. On obtiendra une relation d'équivalence unitaire entre $\mathrm{T_K}$ et T_k,

$$\mathrm{T_K} = \mathrm{U}^*\,\mathrm{T}_k\,\mathrm{U}.$$

Une telle relation permet de transposer la diagonalisation de T_k en diagonalisation de $\mathrm{T_K}$, complétant ainsi le tableau.

Supposons que K soit un noyau hermitien ; il existe une base hilbertienne (f_n) de $\mathrm{L}_2(\mathrm{X})$ formée de vecteurs propres de $\mathrm{T_K}$, $(\mathrm{T_K} f_n) = \lambda_n f_n$ pour tout $n \geq 0$. Appliquons ce qui précède avec la base $(f_m \otimes \overline{f_n})$. On a vu que

$$\mathrm{T_K} f_{n_0} = \sum_m k_{m,n_0}\, f_m$$

qui doit être égal ici à $\lambda_{n_0} f_{n_0}$. Il en résulte que la matrice k est diagonale, avec $k_{n,n} = \lambda_n$ pour tout $n \geq 0$ (comme en dimension finie : dans une base de vecteurs propres, la matrice est diagonale !), et on a au sens de $\mathrm{L}_2(\mathrm{X} \times \mathrm{X})$

$$\mathrm{K} = \sum_{n=0}^{+\infty} \lambda_n\, f_n \otimes \overline{f_n}.$$

Si *par hasard* la série $\sum |\lambda_n|$ converge, et si les fonctions (f_n) sont uniformément bornées, on a aussi convergence ponctuelle de la série

$$\sum_{n=0}^{+\infty} \lambda_n\, f_n(x)\overline{f_n(y)},$$

et la somme ne peut être que la fonction K. C'est le cas avec l'exemple qui a été traité précédemment. On obtient ainsi quand $0 \leq y \leq x \leq \pi/2$

$$\sin(x)\cos(y) = \frac{2}{\pi} - \sum_{k=1}^{+\infty} \frac{4}{(4k^2-1)\pi}\,\cos(2kx)\cos(2ky).$$

Prépa. Agrég. écrit d'Analyse, séance n° 17, Mercredi 28 Mars 2001.

Algèbres de Banach

Le corps de base est $\mathbb{C}$ dans tout ce qui suit.

Définition. Une *algèbre de Banach unitaire complexe* A est un espace de Banach complexe, avec une multiplication bilinéaire $(a, b) \in \mathrm{A} \times \mathrm{A} \to ab \in \mathrm{A}$ continue, associative avec élément unité ; l'élément unité sera noté $\mathbf{1}_{\mathrm{A}}$ ou simplement $\mathbf{1}$; on suppose que

$$\|\mathbf{1}\| = 1\,; \quad \|ab\| \le \|a\|\,\|b\|$$

pour tous $a, b \in \mathrm{A}$.

Il en résulte $\|a^n\| \le \|a\|^n$ pour tout $n \ge 0$ (on fait bien sûr la convention $a^0 = \mathbf{1}$). Pour tout $a \in \mathrm{A}$ la limite

$$\rho(a) = \lim_n \|a^n\|^{1/n}$$

existe. Evidemment $\rho(a) \le \|a\|$.

Exercice. Démontrer l'existence de $\lim_n \|a^n\|^{1/n}$ en montrant que la suite $(\|a^n\|^{1/n})$ converge vers $\inf_{n>0} \|a^n\|^{1/n}$ (indication : si $\|a^{m_0}\|^{1/m_0}$ approche l'inf, on écrira la division $n = m_0 q + r$, $0 \le r < m_0$).

Remarque. L'algèbre $\mathrm{A} = \{0\}$ vérifierait toutes les propriétés (il faudrait poser $\mathbf{1} = 0$), sauf $\|\mathbf{1}\| = 1$; pour certains développements il est utile d'admettre l'algèbre nulle. Je ne le ferai pas ici ; le fait que $\mathbf{1} \ne 0_{\mathrm{A}}$ interviendra insidieusement dans le théorème qui dit que le spectre d'un élément est non vide.

Exemples.

1. Pour tout espace normé complexe $\mathrm{E} \ne \{0\}$, l'algèbre $\mathrm{A} = \mathcal{L}(\mathrm{E})$ est une algèbre de Banach unitaire, où le produit est la composition des endomorphismes et la norme est la norme des opérateurs. C'est un exemple éminemment non commutatif.

2. L'espace $\mathrm{C}(\mathrm{K})$ des fonctions complexes continues sur un compact K non vide, avec la norme du sup et la multiplication usuelle des fonctions, est une algèbre de Banach commutative.

3. L'espace $\mathrm{L}_1(\mathbb{R})$ avec sa norme usuelle et le produit de convolution serait un autre exemple commutatif, mais il lui manque une unité ; l'espace $\mathrm{M}_b(\mathbb{R})$ des mesures bornées sur $\mathbb{R}$ avec la convolution possède pour unité δ_0, la mesure de Dirac en 0 ; c'est un exemple d'algèbre de Banach unitaire commutative ; il permet de considérer la sous-algèbre unitaire B des mesures de la forme $f(x)\,dx + \lambda\delta_0$ où $f \in \mathrm{L}_1(\mathbb{R})$ et $\lambda \in \mathbb{C}$: le passage de $\mathrm{L}_1(\mathbb{R})$ à l'algèbre B donne une idée de la procédure d'adjonction d'une unité aux algèbres non unitaires.

4. L'espace $\ell_1(\mathbb{Z})$ avec sa norme usuelle et le produit de convolution des suites est encore un exemple, pour lequel on peut donner une autre représentation : on désigne par W l'ensemble de toutes les fonctions f sur le cercle unité $\mathbb{T}$ qui sont de la forme

$$\forall t \in [0, 2\pi], \quad f(\mathrm{e}^{it}) = \sum_{n\in\mathbb{Z}} c_n\,\mathrm{e}^{int}$$

avec $\sum |c_n| < +\infty$, avec pour norme $\|f\|_{\mathrm{W}} = \sum_{n\in\mathbb{Z}} |c_n|$. Cet exemple est une algèbre de Banach commutative pour le produit ponctuel des fonctions sur $\mathbb{T}$; cet exemple est évidemment isomorphe à l'algèbre de convolution $\ell_1(\mathbb{Z})$.

5. C*-algèbres : on se donne une algèbre de Banach unitaire complexe A, avec en plus une opération involutive antilinéaire $a \to a^*$ de A dans A telle que $(ab)^* = b^* a^*$ et $\|a^* a\| = \|a\|^2$ pour tout $a \in \mathrm{A}$. Il en résulte que $\mathbf{1}^* = \mathbf{1}$ et $\|a^*\| = \|a\|$.

Exemple de base : $\mathcal{L}(\mathrm{H})$ où H est un espace de Hilbert $\neq \{0\}$, l'opération $*$ étant donnée par la prise de l'opérateur adjoint, $\mathrm{T} \to \mathrm{T}^*$; l'espace complexe C(K) avec conjugaison complexe des fonctions comme opération $*$ est un autre exemple de C*-algèbre (on suppose K non vide).

Eléments inversibles

On dit que $a \in \mathrm{A}$ est *inversible* (on ajoute : *dans* A, s'il y a risque de confusion) quand il existe $b \in \mathrm{A}$ tel que $ab = ba = \mathbf{1}$. On sait que quand a est *à la fois* inversible à droite et à gauche, il est inversible.

Proposition. *Si $v \in \mathrm{A}$ vérifie $\rho(v) < 1$, la série $\sum v^n$ converge dans A et sa somme est l'inverse de $\mathbf{1} - v$,*

$$(\mathbf{1} - v)^{-1} = \sum_{n=0}^{+\infty} v^n.$$

Démonstration. Si on choisit r tel que $\rho(v) < r < 1$ on aura $\|v^n\| \leq r^n$ pour n assez grand, d'où la convergence en norme de la série des puissances. Si $\mathrm{S} \in \mathrm{A}$ désigne la somme de la série, la continuité du produit par v donne $v\,\mathrm{S} = \sum_{n \geq 0} v^{n+1} = \mathrm{S} - \mathbf{1}$ d'où $(\mathbf{1} - v)\,\mathrm{S} = \mathbf{1}$, et la même chose de l'autre côté.

Exercice. Si $\|v\| < 1$, montrer qu'on a la majoration $\|(\mathbf{1} - v)^{-1}\| \leq \big(1 - \|v\|\big)^{-1}$.

Conséquence. Si a est inversible et v petit (précisément : $\|v\| < \|a^{-1}\|^{-1}$), $a - v$ est inversible et

$$(a - v)^{-1} = a^{-1} + a^{-1} v a^{-1} + a^{-1} v a^{-1} v a^{-1} + a^{-1} v a^{-1} v a^{-1} v a^{-1} + \cdots$$

En particulier si a est inversible et si $h \in \mathbb{C}$ est petit on a un développement de la forme

$$(*) \qquad (a - hv)^{-1} = \sum_{n=0}^{+\infty} h^n\, a_n$$

avec $(a_n) \subset \mathrm{A}$, développement valable pour tout $h \in \mathbb{C}$ tel que $|h| < h_0 = \|v\|^{-1} \|a^{-1}\|^{-1}$.

Ouvert des éléments inversibles

Considérons l'ensemble

$$\mathrm{U}(\mathrm{A}) = \{a \in \mathrm{A} : a \ \text{ est inversible dans A}\}.$$

L'ensemble U(A) est ouvert dans A : c'est une conséquence immédiate de ce qui précède.

Exercice. Soit $a \in \partial \mathrm{U}(\mathrm{A})$ (la frontière de U(A) ; donc a n'est pas inversible) ; montrer qu'il existe une suite $(b_n) \subset \mathrm{A}$ d'éléments de norme un telle que $\lim_n ab_n = 0_{\mathrm{A}}$ **ou bien** $\lim_n b_n a = 0_{\mathrm{A}}$ (indication : puisque a n'est pas inversible, il n'existe par exemple aucun b tel que $ab = \mathbf{1}$; dans ce cas on montre qu'il n'y a pas de constante $c > 0$ telle que $\|ax\| \geq c\,\|x\|$ pour tout $x \in \mathrm{A}$, ce qui revient à dire qu'il existe une suite (b_n) d'éléments de norme un telle que $\lim_n ab_n = 0_{\mathrm{A}}$).

En conséquence, l'image de a par tout homomorphisme d'algèbres de Banach unitaires de A dans une autre algèbre B restera non inversible.

Rayon spectral et spectre

On utilisera plus bas la forme suivante du théorème de Hahn-Banach : pour tout vecteur x d'un espace normé E, il existe une forme linéaire continue $x^* \in \mathrm{E}^*$ telle que $\|x^*\| \le 1$ et $x^*(x) = \|x\|$.

Définition. Soit $a \in \mathrm{A}$; on appelle *spectre* de a le sous-ensemble de $\mathbb{C}$ défini par

$$\mathrm{Sp}(a) = \{\lambda \in \mathbb{C} : a - \lambda \mathbf{1} \ \text{ non inversible } \}.$$

Il est clair que le spectre de a est fermé dans $\mathbb{C}$.

Exemple. Dans C(K) un élément f est inversible si et seulement si f ne s'annule pas sur le compact K (supposé non vide). Par conséquent, $\mathrm{S}p(f)$ est l'ensemble $f(\mathrm{K}) \subset \mathbb{C}$ des valeurs prises par f sur K. C'est un compact non vide de $\mathbb{C}$, conformément au théorème qui suit.

Théorème. *Le spectre de tout* $a \in \mathrm{A}$ *est compact et non vide.*

Démonstration. On fixe $a \in \mathrm{A}$ et on étudie la fonction $f : z \in \mathrm{V} \to (\mathbf{1} - za)^{-1} \in \mathrm{A}$, où l'ensemble $\mathrm{V} \subset \mathbb{C}$ est défini par

$$\mathrm{V} = \{z \in \mathbb{C} : \mathbf{1} - za \ \text{ inversible } \} = \{0\} \cup \{z \ne 0 : z^{-1} \notin \mathrm{Sp}(a)\}.$$

C'est un ouvert, non vide : si $\rho(za) = |z|\,\rho(a) < 1$, on sait que $z \in \mathrm{V}$, c'est à dire que V contient le disque ouvert de rayon $\rho(a)^{-1}$ (valeur $+\infty$ admise) ; en passant à la variable $1/z$ on obtient que $\mathrm{Sp}(a)$ est borné :

l'ensemble $\mathrm{Sp}(a)$ est contenu dans le disque fermé de rayon $\rho(a)$.

On sait donc que $\mathrm{Sp}(a)$ est compact. On va montrer maintenant qu'on ne peut pas trouver de disque contenant $\mathrm{Sp}(a)$ qui soit plus petit que le disque de rayon $\rho(a)$; précisément on va montrer qu'il y a effectivement un point λ du spectre de a dont le module est $\rho(a)$. On aura ainsi terminé la preuve, et obtenu une information supplémentaire, la *formule du rayon spectral*.

Reprenons la fonction f. Si $z \in \mathrm{V}$ on aura d'après $(*)$ un développement de la forme

$$f(z+h) = \big((\mathbf{1} - za) - ha\big)^{-1} = \sum_{n=0}^{+\infty} h^n\, a_n(z)$$

valable pour h assez petit ; si a^* est une forme linéaire continue sur A on aura donc que $g(z) = a^*(f(z))$ est une fonction holomorphe dans V ; par ailleurs pour z assez petit on sait que $f(z) = \sum_{n=0}^{+\infty} z^n\, a^n$ donc $g(z) = \sum_{n=0}^{+\infty} z^n\, a^*(a^n)$ est le développement de Taylor à l'origine pour la fonction holomorphe g.

Supposons maintenant que $r > 0$ soit tel que $\overline{\mathrm{B}(0,r)}$ soit contenu dans V. La fonction $z \to \|f(z)\|$ est définie et continue sur le cercle de rayon r, donc y admet un maximum $\mathrm{M}(r)$; les inégalités de Cauchy, appliquées à la fonction holomorphe g, donnent que

$$|a^*(a^n)| \le \frac{\max_{|z|=r} |g(z)|}{r^n} \le \|a^*\|\ \frac{\mathrm{M}(r)}{r^n}$$

pour tout $n \ge 0$, donc par le théorème de Hahn-Banach rappelé ci-dessus,

$$\|a^n\| \le \frac{\mathrm{M}(r)}{r^n}$$

pour tout n. Ceci donne $\rho(a) \le 1/r$, c'est à dire $r \le \rho(a)^{-1}$ et finalement on a montré qu'il n'y a pas de disque de rayon plus grand que $\rho(a)^{-1}$ qui puisse être contenu dans V. Dans le cas où $\rho(a) > 0$, cela signifie que le cercle de rayon $\rho(a)^{-1}$ ne peut pas être contenu dans V : il existe donc au moins un z_0 tel que $|z_0| = \rho(a)^{-1}$ et tel que $\mathbf{1} - z_0 a$ ne soit pas inversible. En posant $\lambda = z_0^{-1}$ on trouve $\lambda \in \mathrm{Sp}(a)$ tel que $|\lambda| = \rho(a)$, ce qui prouve en même temps que le spectre est non vide ; le cas où $\rho(a) = 0$ est un peu plus compliqué.

Si on avait $\rho(a) = 0$ et un spectre vide, $f(z)$ serait définie sur $\mathbb{C}$ tout entier et de plus a^{-1} existerait ; mais on voit qu'alors $-zf(z) = (a - z^{-1}\mathbf{1})^{-1}$ convergerait vers a^{-1} lorsque $|z| \to +\infty$, donc $\|f(z)\| = \mathrm{O}(|z|^{-1})$ tendrait vers 0 dans les mêmes conditions, et pour tout a^* la fonction g correspondante définie par $g(z) = a^*(f(z))$ serait holomorphe sur $\mathbb{C}$ et tendant vers 0 à l'infini. En appliquant le théorème de Liouville à la fonction g on trouverait $g = 0$ pour tout a^*, donc $f(z) = 0$, ce qui est impossible puisque les valeurs de f sont des inverses, nécessairement non nuls **puisque** $\mathbf{1}_\mathrm{A} \ne 0_\mathrm{A}$.

On a obtenu dans le cours de la démonstration précédente la formule importante suivante, dite *formule du rayon spectral*,

$$\rho(a) = \max\{|\lambda| : \lambda \in \mathrm{Sp}(a)\}.$$

Remarque. Même pour une matrice complexe de taille $n \times n$, cette formule est intéressante et non triviale, bien qu'on puisse l'obtenir sans parler de fonctions holomorphes. Si on met une norme complexe quelconque sur $\mathbb{C}^n$, on obtient un espace normé E, une algèbre de Banach $\mathrm{A} = \mathcal{L}(\mathrm{E})$ à laquelle les résultats précédents s'appliquent.

Corollaire : Corps de Banach. *Si dans* A *tout élément non nul est inversible, alors tout* $a \in \mathrm{A}$ *est de la forme* $a = \lambda \mathbf{1}_\mathrm{A}$ *pour un certain* $\lambda \in \mathbb{C}$, *donc* A *est isomorphe à* $\mathbb{C}$.

Démonstration. Il existe $\lambda \in \mathbb{C}$ tel que $a - \lambda\mathbf{1}$ soit non inversible, donc $a - \lambda\mathbf{1} = 0_\mathrm{A}$.

Idéaux et idéaux maximaux d'une algèbre de Banach commutative

Dans cette section les algèbres sont commutatives.

Les idéaux I considérés seront tous *propres*, c'est à dire $\mathrm{I} \ne \mathrm{A}$. Si I est propre, il ne peut pas rencontrer les inversibles ; on a donc $\mathrm{I} \cap \mathrm{U(A)} = \emptyset$ et puisque U(A) est ouvert on a aussi $\bar{\mathrm{I}} \cap \mathrm{U(A)} = \emptyset$. Il est facile de voir que l'adhérence $\bar{\mathrm{I}}$ est encore un idéal, et c'est encore un idéal propre. Si I est un idéal maximal, il est donc fermé et on peut considérer l'algèbre de Banach quotient A/I ; mais cette algèbre est un corps, donc directement isomorphe à $\mathbb{C}$ par le corollaire précédent, et l'application $\chi : \mathrm{A} \to \mathrm{A/I} \to \mathbb{C}$ est linéaire et multiplicative ; son noyau est I.

On appelle *caractère* de A toute application linéaire et multiplicative de A dans $\mathbb{C}$. On vient de dire que tout idéal maximal est le noyau d'un caractère. Réciproquement, il est clair que pour tout caractère χ sur A, $\ker\chi$ est un idéal maximal. Puisque le noyau de χ est fermé, χ est une forme linéaire continue (attention, les caractères de **groupes** topologiques peuvent être discontinus, par exemple les caractères de $\mathbb{R}/\mathbb{Z}$: à côté des caractères continus $\theta \to \mathrm{e}^{2\pi i n\theta}$ pour $n \in \mathbb{Z}$, il existe des caractères étranges qu'on obtient avec l'axiome du choix).

Les idéaux maximaux de A *sont les noyaux des caractères de* A.

Tout caractère vérifie $\chi(\mathbf{1}) = 1$, ce qui entraîne $\|\chi\| \ge 1$. Mais si $\chi(a) = z$ avec $\|a\| \le 1$, on aura $|z|^n = |\chi(a^n)| \le \|\chi\|\,\|a\|^n \le \|\chi\|$ pour tout $n > 0$, donc $|z| \le 1$ et finalement $\|\chi\| = 1$.

Exemple. Dans C(K) les fonctions nulles en un point $x_0 \in \mathrm{K}$ forment un idéal maximal ; le caractère correspondant est l'application δ_{x_0} d'évaluation en x_0, $\delta_{x_0}(f) = f(x_0)$; si on préfère le dire ainsi, δ_{x_0} est la mesure de Dirac du point x_0.

Exercice : montrer que tous les caractères de C(K) sont de cette forme.

Le spectre d'une algèbre commutative

Définition. Le spectre d'une algèbre commutative A est l'ensemble Sp(A) formé de tous les caractères χ sur A.

On a vu ci-dessus que Sp(A) est un sous-ensemble de la sphère unité du dual A^*. On le munit de la topologie induite par la topologie $*$-faible de la boule unité (fermée) $\mathrm{B}_{\mathrm{A}^*}$ de A^*, qui est la topologie de la convergence simple sur A. Le théorème de Banach-Alaoglu dit que $\mathrm{B}_{\mathrm{A}^*}$ est compacte pour cette topologie. Le spectre Sp(A) est fermé dans $\mathrm{B}_{\mathrm{A}^*}$: en effet, il est clair que le sous-ensemble des fonctions multiplicatives est fermé pour la convergence simple. Le spectre de A est donc un espace topologique compact, en général non métrisable.

Exemple. Le spectre de C(K) est homéomorphe à K (l'homéomorphisme inverse est l'application $x \in \mathrm{K} \to \delta_x$).

Proposition. *Un élément $a \in \mathrm{A}$ est inversible si et seulement si $\chi(a) \neq 0$ pour tout $\chi \in \mathrm{Sp(A)}$.*

Démonstration. Si a est inversible on a $\chi(a)\chi(a^{-1}) = 1$ donc $\chi(a) \neq 0$; inversement si a est non inversible l'ensemble $a\mathrm{A}$ est un idéal propre, qui sera contenu dans un idéal maximal I, donc il existe un caractère χ qui s'annule au point a, à savoir le caractère χ tel que $\mathrm{I} = \ker \chi$.

On peut voir le résultat d'une autre façon. Désignons par γ l'application de A dans C(Sp(A)) qui associe à chaque $a \in \mathrm{A}$ la fonction continue $\gamma(a)$ sur Sp(A) définie par $\gamma(a)(\chi) = \chi(a)$ pour tout point $\chi \in \mathrm{Sp(A)}$. Il est immédiat de vérifier que γ est un homomorphisme d'algèbres unitaires. De plus $\|\gamma\| \leq 1$.

L'élément $a \in \mathrm{A}$ est inversible dans A si et seulement si son image $\gamma(a)$ est une fonction sur Sp(A) qui ne s'annule en aucun point $\chi \in \mathrm{Sp(A)}$; autrement dit a est inversible dans A si et seulement si son image $\gamma(a)$ est inversible dans C(Sp(A)).

On en déduit encore :

Le spectre de $a \in \mathrm{A}$ est identique au spectre de $\gamma(a)$ dans C(Sp(A)), c'est à dire que

$$\mathrm{Sp}(a) = \{\chi(a) : \chi \in \mathrm{Sp(A)}\}.$$

L'application γ est une isométrie surjective pour les C*-algèbres : *toute C*-algèbre commutative est isomorphe à un espace* C(K). Voir appendice.

Un exemple : le spectre de l'algèbre de Wiener W.

Pour chaque $n \in \mathbb{Z}$ considérons la fonction $e_n \in \mathrm{W}$ définie par $e_n(\mathrm{e}^{it}) = \mathrm{e}^{int}$, pour t variant dans $[0, 2\pi]$; puisque la série de Fourier de e_n comporte un seul terme non nul, avec coefficient 1, on a $\|e_n\|_{\mathrm{W}} = 1$ pour tout $n \in \mathbb{Z}$.

Si χ est un caractère de W posons $\lambda = \chi(e_1)$; on a $|\lambda| = |\chi(e_1)| \le \|\chi\| \, \|e_1\|_{\mathrm{W}} = 1$; mais $e_1 e_{-1} = \mathbf{1}$ donc $\chi(e_{-1}) = \lambda^{-1}$ et $|\lambda^{-1}| \le 1$ pour la même raison. On a donc $|\lambda| = 1$, c'est à dire que $\lambda \in \mathbb{T}$. On en déduit que

$$\forall f \in \mathrm{W}, \quad \chi(f) = f(\lambda)$$

c'est à dire que le spectre Sp(W) est naturellement isomorphe à $\mathbb{T}$, comme pour l'algèbre $\mathrm{C}(\mathbb{T})$; les caractères de W sont les fonctions d'évaluation aux points de $\mathbb{T}$.

Théorème (de Wiener). *Si $f \in \mathrm{W}$ ne s'annule pas sur $\mathbb{T}$, la fonction $z \in \mathbb{T} \to 1/f(z)$ est encore dans* W.

Démonstration. Si f ne s'annule pas sur $\mathbb{T}$, le résultat précédent sur l'identification de Sp(W) nous dit que $\chi(f) \ne 0$ pour tout caractère, ce qui implique que f est inversible dans W ; cet inverse ne peut être que la fonction $1/f$.

Algèbre de Calkin. Spectre essentiel

Soit E un espace de Banach complexe de dimension infinie ; on pose

$$\mathcal{C}(\mathrm{E}) = \mathcal{L}(\mathrm{E})/\mathcal{K}(\mathrm{E}).$$

Le sous-espace $\mathcal{K}(\mathrm{E})$ est fermé, et c'est un idéal bilatère, ce qui permet de munir le quotient d'une structure naturelle d'algèbre de Banach ; de plus Id_{E} n'est pas compacte puisque E est de dimension infinie (théorème de Riesz), donc sa classe au quotient n'est pas nulle. On a donc $\mathbf{1} \ne 0$ dans le quotient, ce qui permet d'avoir une notion de spectre intéressante.

Pour tout $\mathrm{T} \in \mathcal{L}(\mathrm{E})$, on appelle *spectre essentiel* de T le compact non vide de $\mathbb{C}$ qui est le spectre de la classe de T dans l'algèbre de Calkin $\mathcal{C}(\mathrm{E})$. Pour comprendre un peu cette notion, il faut commencer par s'intéresser à l'inversibilité dans $\mathcal{C}(\mathrm{E})$, c'est à dire l'inversibilité de $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *modulo les opérateurs compacts*.

Prenons un exemple. Considérons le shift à droite S sur $\ell_2(\mathbb{N})$ et son adjoint S^* qui est le shift à gauche. On voit que $\mathrm{S}^*\mathrm{S} = \mathrm{Id}$ et que $\mathrm{S}\,\mathrm{S}^* - \mathrm{Id}$ est de rang un, donc compact ; on en déduit que la classe de S dans $\mathcal{C}(\ell_2)$ admet pour inverse la classe de S^* : le shift n'est pas inversible dans $\mathcal{L}(\ell_2)$, mais il est inversible modulo les compacts.

Un opérateur T est dit *de Fredholm* si son noyau est de dimension finie et si son image est fermée et de codimension finie. Comme le noyau N de T est de dimension finie, on peut lui trouver un supplémentaire topologique E_1 et $\mathrm{T}_1 = \mathrm{T}_{|\mathrm{E}_1}$ est alors une bijection algébrique de E_1 sur l'espace de Banach T(E) ; d'après le théorème des isomorphismes, T_1 est un isomorphisme de E_1 sur $\mathrm{F}_1 = \mathrm{T}(\mathrm{E})$.

On peut voir que $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ est Fredholm si et seulement s'il existe deux décompositions $\mathrm{E} = \mathrm{E}_0 \oplus \mathrm{E}_1$ et $\mathrm{E} = \mathrm{F}_0 \oplus \mathrm{F}_1$ en somme directe topologique, avec E_0 et F_0 de dimension finie, et où la restriction T_1 de T à E_1 est un isomorphisme de E_1 sur F_1.

Dans ces conditions, T est inversible modulo les compacts. En effet, désignons par $\mathrm{V} \in \mathcal{L}(\mathrm{E})$ l'opérateur nul sur F_0, égal à l'inverse de T_1 sur F_1. Alors $\mathrm{TV} - \mathrm{Id}$ et $\mathrm{VT} - \mathrm{Id}$ sont de rang fini, donc compacts, et la classe de V dans $\mathcal{C}(\mathrm{E})$ est l'inverse de la classe de l'opérateur T.

Réciproquement, on montre dans tout cours raisonnable d'Analyse Fonctionnelle que tout opérateur de la forme $\mathrm{Id} - \mathrm{K}$, K compact, a un noyau de dimension finie (théorème de Riesz) et une image fermée de codimension finie. S'il existe S tel que $\mathrm{Id} - \mathrm{ST}$ et $\mathrm{TS} - \mathrm{Id}$ soient compacts, on en déduit que ST et TS ont un noyau de dimension finie et une image fermée de codimension finie, donc T a lui aussi un noyau de dimension finie et une image fermée de codimension finie. Il en résulte aussi que :

$\mathrm{T} - \lambda\,\mathrm{Id}$ *est Fredholm si et seulement si* λ *n'est pas dans le spectre essentiel de* T.

Appendice : γ *est isométrique pour les* C^**-algèbres*

Montrons que l'homomorphisme γ qui a été décrit dans la section sur les algèbres commutatives est isométrique pour toute C^*-algèbre commutative A. Lorsque $u \in \mathrm{A}$ est hermitien, c'est à dire $u = u^*$, on voit en considérant la série entière de l'exponentielle que si on pose $a = \mathrm{e}^{iu}$, alors $a^* = \mathrm{e}^{-iu}$, donc $a^*a = aa^* = \mathbf{1}_\mathrm{A}$ (on dit que a est *unitaire*). Alors $a^* = a^{-1}$ et $\|a\| = \|a^{-1}\| = 1$, puisque $1 = \|\mathbf{1}\| = \|a^*a\| = \|a\|^2$; ceci implique que la suite $(a^n)_{n\in\mathbb{Z}}$ est bornée, donc $(\chi(a^n))_{n\in\mathbb{Z}}$ est bornée et par conséquent $1 = |\chi(a)|$; mais $\chi(a) = \chi(\mathrm{e}^{iu}) = \mathrm{e}^{i\chi(u)}$ par la propriété de caractère, donc $\chi(u)$ est réel pour tout hermitien u ; voici une version plus élémentaire : on aura pour t réel petit, puisque $(\mathbf{1} + itu)^* = \mathbf{1} - itu$

$$\|\mathbf{1} + itu\|^2 = \|(\mathbf{1} - itu)(\mathbf{1} + itu)\| = \|\mathbf{1} + t^2u^2\| = 1 + \mathrm{O}(t^2)$$

et $|\chi(\mathbf{1} + itu)| = |1 + it\chi(u)| = 1 + \mathrm{O}(t^2)$ implique que $\chi(u)$ est réel. Ceci implique, en considérant les éléments hermitiens $a + a^*$ et $i(a - a^*)$ que

$$\forall a \in \mathrm{A}, \quad \chi(a^*) = \overline{\chi(a)}.$$

Si u est hermitien, on a $r(u) = \|u\|$ (commencer avec $\|u\|^2 = \|u^*u\| = \|u^2\|$), donc d'après la formule du rayon spectral il existe $\chi \in \mathrm{Sp(A)}$ tel que $|\chi(u)| = \|u\|$. Ceci implique que pour tout $a \in \mathrm{A}$, il existe χ tel que

$$|\chi(a)|^2 = \chi(a^*a) = \|a^*a\| = \|a\|^2,$$

ce qui montre que γ est isométrique de A dans C(Sp(A)) ; l'image $\gamma(\mathrm{A})$ est donc fermée, mais c'est aussi une sous-algèbre de C(Sp(A)), stable par conjugaison complexe et qui sépare les points du compact Sp(A) (évident) ; d'après le théorème de Stone-Weierstrass, $\gamma(\mathrm{A})$ est dense dans C(Sp(A)), donc égale puisqu'aussi fermée.

Exemple. L'espace $\mathrm{L}_\infty(0, 1)$ muni de la norme du sup essentiel et du produit ponctuel, est une C^*-algèbre commutative, si on définit f^* comme étant la complexe conjuguée de la fonction f. Cet espace est donc un C(K) ! Mais le compact K n'est pas très intuitif : c'est le spectre de $\mathrm{L}_\infty(0, 1)$; si vous voyez passer un caractère ou un idéal maximal de $\mathrm{L}_\infty(0, 1)$, prévenez-moi !