Analyse fonctionnelle

Bertrand RÉMY
(d'après un cours de Pierre EYMARD)

26 mars 2018

Première partie

Géométrie dans les espaces vectoriels normés

Dans ce chapitre, la notion centrale est celle d'ensemble *convexe* ; elle donne prise à l'intuition géométrique. Après les propriétés générales de ces ensembles, nous étudierons les relations entre convexité et dimension (théorèmes de Carathéodory, de Helly). L'énoncé le plus important du chapitre est le théorème de Hahn-Banach ; dans sa version analytique, il montre la possibilité de prolonger les formes linéaires continues sans augmenter leur norme, ce qui a des conséquences intéressantes pour la dualité des espaces normés et en théorie de l'approximation. Dans une version géométrique le théorème de Hahn-Banach conduit à l'étude de la séparation des ensembles convexes par des hyperplans. Enfin, un théorème de Krein et Milman permet de reconstituer les ensembles convexes à partir de leurs points extrémaux.

Il s'agit de résultats relativement récents, dont le début remonte aux années vingt ; le point culminant de leur histoire reste l'admirable livre du mathématicien polonais Stefan Banach, publié à Varsovie en farnçais en 1932 sous le titre « Théorie des opérateurs linéaires ».

I.1 Propriétés générales des ensembles convexes

Soit E un espace vectoriel sur $\mathbf{R}$ et soit A une partie de E. On dit que A est *convexe* si, chaque fois que A contient deux point x et y, elle contient le segment $[x; y]$; autrement dit, si les hypothèses $x \in A$, $y \in A$, λ réel, $0 \leqslant \lambda \leqslant 1$ impliquent la conclusion : $\lambda x + (1-\lambda)y \in A$.

Cette définition s'améliore aussitôt en une propriété plus forte :

Proposition I .1 *Soit A convexe. Alors les hypothèses $x_1, \dots x_p \in A$, $\lambda_1, \dots \lambda_p \in \mathbf{R}$, $\lambda_1 \geqslant 0, \dots \lambda_p \geqslant 0$ et $\lambda_1 + \dots \lambda_p = 1$ impliquent la conclusion*

$$\lambda_1 x_1 + \dots \lambda_p x_p \in A.$$

Autrement dit, un ensemble convexe contient tous les barycentres à coefficients positifs de ses sous-ensembles de points. La démonstration est facile, par récurrence sur p ; essayez de l'écrire avant de lire ce qui suit !

Preuve. On fait une récurrence sur p. Si $p = 2$, la proposition n'est autre que la définition de la convexité de A. Soit $p > 2$, et supposons la proposition vraie pour moins de p points. Si l'un des λ_i était nul, on pourrait supprimer x_i et l'on n'aurait affaire qu'à moins de p points ; on peut donc supposer tous les $\lambda_i > 0$. Puisque $\dfrac{1-\lambda_1}{\lambda_2 + \dots \lambda_p} = 1$, on a :

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p = \lambda_1 x_1 + \frac{1-\lambda_1}{\lambda_2 + \dots \lambda_p}(\lambda_2 x_2 + \cdots + \lambda_p x_p)$$

$$= \lambda_1 x_1 + (1-\lambda_1)\Big[\frac{\lambda_2}{\lambda_2 + \dots \lambda_p} x_2 + \cdots + \frac{\lambda_p}{\lambda_2 + \dots \lambda_p} x_p\Big].$$

Par l'hyptohèse de récurrence, nous avons entre crochets un point, disons y, de A, donc

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p = \lambda_1 x_1 + (1 - \lambda_1) y$$

est dans A, par définition de la convexité. □

Si E est normé, toute partie convexe A de E est connexe, et même connexe par arcs puisque deux points peuvent toujours être joints par un segment contenu dans A. En particulier, les parties convexes de $\mathbf{R}$ sont les intervalles (de tout genre).

Dans E normé, toute boule ouverte

$$B = \{x \in E : \|x - a\| < r\}$$

où $a \in E$, r réel > 0, est convexe, car si $\|x - a\| < r$, $\|y - a\| < r$ et $0 \leqslant \lambda \leqslant 1$, on a :

$$\|\lambda x + (1 - \lambda) y - a\| \; = \; \|\lambda(x - a) + (1 - \lambda)(y - a)\|$$

$$\leqslant \lambda \|(x - a)\| + (1 - \lambda)\|(y - a)\| < \lambda r + (1 - \lambda) r = r.$$

De même, toute boule fermée est convexe.

Tout sous-espace vectoriel V de E est convexe, ainsi que, plus généralement, tout *sous-espace affine*

$$V + a = \{x = y + a : y \in V\},$$

où V est un sous-espace vectoriel de E et a est un vecteur donné de E.

Sont aussi convexes les *demi-espaces* affines ouverts (resp. fermés)

$$\{x \in E : f(x) > \alpha \, (\text{resp. } \geqslant \alpha)\},$$

où $\alpha \in \mathbf{R}$, et f une forme linéaire sur E non nulle, sont donnés, car par exemple si $f(x) > \alpha$, $f(x) > \alpha$ et $0 \leqslant \lambda \leqslant 1$, on a :

$$f\left(\lambda x + (1 - \lambda) y\right) = \lambda f(x) + (1 - \lambda) f(y) > \lambda \alpha + (1 - \lambda)\alpha = \alpha.$$

À partir d'exemples connus comme convexes, on en fabrique d'autres grâce à l'une ou l'autre des propriétés de permanence énoncées dans la

Proposition I .2 1° *L'intersection, le produit cartésien l'image directe ou réciproque par application linéaire affine, et (si E est normé) la fermeture, l'intérieur de parties convexes sont encore convexes.*

2° *Si A et B sont des parties convexes dans E et si μ et ν sont des nombres réels, l'ensemble*

$$\mu A + \nu B = \{\mu a + \nu b : a \in A, b \in B\}$$

est convexe dans E ; en particulier $A + B$ et $A - B$ sont convexes.

Les démonstrations sont faciles, sauf peut-être celle qui concerne l'intérieur. Essayez de les écrire avant de lire ce qui suit !

Preuve. Intersections. Il est évident sur la définition que, si $(A_\lambda)_{\lambda\in\Lambda}$ est une famille finie ou infinie d'ensembles convexes dans E, alors l'ensemble $\bigcap_{\lambda\in\Lambda} A_\lambda$ est convexe. Par exemple, dans E normé, toute intersection de demi-espaces fermés est un convexe fermé. Nous verrons plus tard que la réciproque est vraie : tout convexe fermé est l'intersection des demi-espaces fermés qui le contiennent.

Exercice I .3 *Convainquez-vous d'ores et déjà que cet énoncé est vrai dans le cas d'un disque fermé du plan euclidien.*

Produits cartésiens. Si E_1 et E_2 sont deux espaces vectoriels sur $\mathbf{R}$, et si A_1 et A_2 en sont des parties convexes respectives, alors l'ensemble

$$A = A_1 \times A_2 = \{(x_1, x_2) : x_1 \in A_1, x_2 \in A_2\}$$

est convexe dans $E_1 \times E_2$, car si (x_1, x_2) et (y_1, y_2) sont dans A et si on a $0 \leqslant \lambda \leqslant 1$, alors

$$\lambda(x_1, x_2) + (1-\lambda)(y_1, y_2) = (\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2)$$

est encore dans $A_1 \times A_2 = A$. Cet énoncé s'étend immédiatement aux produits cartésiens de $n \geqslant 2$ ensembles convexes. Par exemple dans $\mathbf{R}^n$ tout pavé (i.e. tout produit cartésien d'intervalles) est convexe.

Application linéaire affine. Rappelons qu'on appelle *application linéaire affine* de E dans F toute application

$$x \mapsto \psi(x) = \varphi(x) + b,$$

où φ est une application linéaire de E dans F et b est un vecteur fixé dans F. On vérifie aisément que l'image par ψ de tout segment $[x; y]$ dans E est le segment $[\psi(x); \psi(y)]$ dans F. En effet, pour $0 \leqslant \lambda \leqslant 1$, on a :

$$\lambda\psi(x)+(1-\lambda)\psi(y) = \lambda[\varphi(x)+b]+(1-\lambda)[\varphi(y)+b] = \lambda\varphi(x)+(1-\lambda)\varphi(y)+b$$

$$= \varphi\left(\lambda x + (1-\lambda)y\right) + b = \psi\left(\lambda x + (1-\lambda)y\right).$$

De ce fait, il résulte aussitôt que, si A est convexe dans E, alors $\psi(A)$ est convexe dans F. Enfin, soit B convexe dans F ; si $x \in \psi^{-1}(B)$ et $y \in \psi^{-1}(B)$, on a : $\psi([x;y]) = [\psi(x);\psi(y)] \subset B$, donc $[x;y] \subset \psi^{-1}(B)$, ce qui prouve que l'ensemble $\psi^{-1}(B)$ est convexe.

En particulier, l'image d'un convexe par une homothétie ou une translation est un convexe.

Le second point de la proposition est une conséquence de ce que nous venons de dire, car $\mu A + \nu B$ est l'image de $A \times B$ par l'application linéaire $(x,y) \mapsto \lambda x + \mu y$ de $E \times E$ dans E.

Fermeture. Soit A un convexe dans E et notons $\overline{A}$ sa fermeture dans E. Si $x \in \overline{A}$, $y \in \overline{A}$ et $0 \leqslant \lambda \leqslant 1$, il existe une suite $(x_k)_{k\in\mathbf{N}}$ (resp. $(y_k)_{k\in\mathbf{N}}$) dans A telle que $\lim\limits_{k\to\infty} x_k = x$ (resp. $\lim\limits_{k\to\infty} y_k = y$). Alors :

$$\lambda x + (1-\lambda)y = \lambda \lim_{k\to\infty} x_k + (1-\lambda)\lim_{k\to\infty} y_k = \lim_{k\to\infty}(\lambda x_k + (1-\lambda)y_k)$$

est limite d'éléments de A, donc appartient à $\overline{A}$. Ainsi $\overline{A}$ est convexe.

Intérieur. Notons A° l'intérieur de la partie convexe A : c'est l'ensemble des centres des boules ouvertes contenues dans A. Soient $x \in A^\circ$ et $y \in A^\circ$. On se donne un point $m \in]x;y[$. Puisque x est dans A°, il est le centre d'une boule ouverte B_x contenue dans A. L'homothétie de centre y qui transforme x en m, transforme la boule ouverte B_x en une boule ouverte B_m centrée en m, et contenue dans A car, par raison d'homothétie, tout point de B_m est sur un segment $[y;z]$ où $z \in B_x$ et donc $z \in A$. Ainsi $m \in A^\circ$. □

Remarquons que dans cette démonstration, seule l'hypothèse $y \in A$ (et non $y \in A^\circ$) a servi pour obtenir $m \in A^\circ$. On a en fait prouvé le résultat plus fort :

Proposition I .4 *Soit A une partie convexe d'un espace vectoriel normé. Soient $x \in A^\circ$ et $y \in A$. Alors tout point $m \in [x;y[$ est dans A°.* □

Exercice I .5 1° *Améliorer la proposition précédente en supposant seulement $y \in \overline{A}$.*

2° *Soit A convexe avec $A^\circ \neq \varnothing$. Montrer que $\overline{A} = \overline{A^\circ}$ et $A^\circ = (\overline{A})^\circ$.*

Exercice I .6 *Soit A un sous-ensemble convexe compact de $\mathbf{R}^n$ tel que le point 0 soit dans l'intérieur de A. Soit la sphère unité $S = \{x \in \mathbf{R}^n : \|x\| = 1\}$.*

1° *Montrer que, pour tout $y \in S$, la demi-droite d'origine 0 passant par y coupe la frontière $\mathrm{Fr}(A) = \overline{A}\backslash A^\circ$ de A en un point, disons $\delta(y)$, et un seul.*

2° *Montrer que l'application $y \mapsto \delta(y)$ est une application continue de S sur $\mathrm{Fr}(A)$.*

Exercice I .7 *Un convexe de* $\mathbf{R}^n$ *est dit de dimension d si le plus petit sous-espace affine le contenant est de dimension d, i.e. est le translaté d'un sous-espace vectoriel de dimension d.*

1° *Prouver qu'une partie convexe A est de dimension n si et seulement si son intérieur* A° *est non vide.*

2° *Déduire de l'exercice I .6 que tout convexe compact de dimension d de* $\mathbf{R}^n$ *est homéomorphe à la boule-unité fermé de* $\mathbf{R}^d$*.*

I.2 Enveloppes convexes

Soit E un espace vectoriel sur $\mathbf{R}$, et soit A une partie de E. Il y a au moins une partie convexe de E qui contient A, à savoir E lui-même. Donc il y a une plus petite partie convexe de E qui contient A, à savoir l'intersection de toutes les parties convexes de E qui contiennent A. On la nomme l'*enveloppe convexe* de A, et on la note $\mathrm{conv}(A)$. Ainsi définie « de l'extérieur », on peut aussi la définir « de l'intérieur » :

Proposition I .8 *La partie* $\mathrm{conv}(A)$ *n'est autre que l'ensemble des combinaisons linéaires convexes des points de A :*

$$\mathrm{conv}(A) = \{\sum_{i=1}^{p} \lambda_i x_i : x_i \in A, \lambda_i \in \mathbf{R}_+, \sum_{i=1}^{p} \lambda_i = 1, p \textit{ entier quelconque}\}.$$

Preuve. Notons provisoirement $\widehat{A}$ l'ensemble décrit dans l'énoncé. Déjà $\widehat{A}$ est contenu dans tout ensemble convexe contenant A, ce qui implique que $\widehat{A} \subset \mathrm{conv}(A)$. Pour l'inclusion réciproque, il suffit de voir que $\widehat{A}$ est convexe. On se donne donc $x = \sum_i \lambda_i x_i$ et $y = \sum_j \lambda'_j y_j$ dans $\widehat{A}$, ainsi que μ tel que $0 < \mu < 1$. On a :

$$\mu x + (1-\mu)y = \mu \sum_i \lambda_i x_i + (1-\mu) \sum_j \lambda'_j y_j = \sum_i (\lambda_i \mu) x_i + \sum_j \lambda'_j (1-\mu) y_j,$$

avec $x_i \in A$, $y_j \in A$, $\lambda_i \mu > 0$, $\lambda'_j(1-\mu) > 0$ et :

$$\sum_i (\lambda_i \mu) + \sum_j \lambda'_j (1-\mu) = \mu \sum_i \lambda_i + (1-\mu) \sum_j \lambda_j = \mu + (1-\mu) = 1,$$

donc $\mu x + (1-\mu)y$ est dans $\widehat{A}$, et finalement $\widehat{A}$ est convexe. □

Exercice I .9 *Soit* $P(z)$ *un polynôme à coefficients complexes et soit* $P'(z)$ *son polynôme dérivé. Montrer que dans* $\mathbf{C} \simeq \mathbf{R}^2$ *toute racine de* P' *est dans l'enveloppe convexe de l'ensemble des racines de* P*.*

Supposons maintenant E normé. Soit A une partie de E. Il y au moins une partie convexe fermée de E qui contient A, à savoir E lui-même. Donc il y a une plus petite partie convexe fermée de E qui contient A, à savoir l'intersection de toutes les parties convexes fermées de E qui contiennent A. On la nomme l'*enveloppe convexe fermée* de A.

Proposition I .10 *L'enveloppe convexe fermée de A est l'adhérence dans E de l'enveloppe convexe de A.*

Preuve. Cette adhérence est fermée (par définition), convexe (par la proposition I .2), et elle contient A. Si B est un fermé convexe qui contient A, alors B contient $\mathrm{conv}(A)$, et donc $\overline{\mathrm{conv}(A)}$ puisque B est fermé. □

Sans ambiguïté, nous pourrons donc noter $\overline{\mathrm{conv}(A)}$ l'enveloppe convexe fermée de A.

Exercice I .11 *Soit A une partie de E espace vectoriel réel et normé.*
1° *Si A est fermée, est-il toujours vrai que* $\mathrm{conv}(A)$ *soit fermée ?*
2° *Si A est ouverte, est-il toujours vrai que* $\mathrm{conv}(A)$ *soit ouverte ?*

Si E est de dimension finie, disons n, on peut considérablement améliorer la proposition I .8, en y remplaçant « p quelconque » par « $p \leqslant n+1$ ». C'est le théorème de Carathéodory :

Théorème I .12 *Soit A une partie de $\mathbf{R}^n$.*

$$\mathrm{conv}(A) = \{ \sum_{i=1}^{p} \lambda_i x_i : x_i \in A, \lambda_i \in \mathbf{R}_+, \sum_{i=1}^{p} \lambda_i = 1, p \leqslant n+1 \}.$$

Preuve. Soit x un point fixé dans $\mathrm{conv}(A)$. Il s'écrit de plusieurs façons $x = \sum_{i=1}^{p} \lambda_i x_i$ avec $x_i \in A, \lambda_i \in \mathbf{R}_+, \sum_{i=1}^{p} \lambda_i = 1$ et p entier quelconque. On choisit une telle écriture avec p minimal, et on suppose que $p \geqslant n+2$ afin d'aboutir à une contradiction.

Les vecteurs $x_1 - x, x_2 - x, \dots x_{p-1} - x$ sont en nombre $p-1 \geqslant n+1$ donc sont linéairement dépendants : il existe des constantes réelles non toutes nulles $\mu_1, \mu_2, \dots \mu_{p-1}$ telles que

$$\mu_1(x_1 - x) + \mu_2(x_2 - x) + \cdots + \mu_{p-1}(x_{p-1} - x) = 0.$$

Posons $\mu_p = 0$. Pour tout nombre réel t, on a :

$$\sum_{i=1}^{p} (\lambda_i + t\mu_i)(x_i - x) = 0,$$

car le premier membre n'est autre que

$$\Big[\left(\sum_{i=1}^{p} \lambda_i x_i\right) - x\Big] + t \cdot \sum_{i=1}^{p-1} \mu_i(x_i - x) = 0 + t \cdot 0 = 0.$$

Lemme I .13 *On peut choisir une valeur t_0 de t de telle sorte que*

$\lambda_i + t_0\mu_i \geqslant 0$ *pour tout* $i \in \{1; 2; \ldots p\}$ *;*
$\lambda_k + t_0\mu_k = 0$ *pour un indice* k *au moins dans* $\{1; 2; \ldots p\}$.

Si on admet le lemme, la démonstration s'achève comme suit. Pour un tel choix de t_0, on aurait puisque $\mu_p = 0$:

$$\sum_{i=1}^{p}(\lambda_i + t_0\mu_i) \geqslant \lambda_p > 0$$

puis, d'après $\displaystyle\sum_{i=1}^{p}(\lambda_i + t\mu_i)(x_i - x) = 0$:

$$x = \sum_{i=1, i\neq k}^{p} \frac{\lambda_i + t_0\mu_i}{\sum_{i=1}^{p}(\lambda_i + t_0\mu_i)} x_i,$$

ce qui fournirait pour x une écriture de longueur $< p$, contredisant la minimalité de p.

Il reste à prouver le lemme. Comme les λ_i sont tous > 0 par minimalité de p, on peut supprimer les indices i tels que $\mu_i = 0$, et donc supposer tous les $\mu_i \neq 0$. On veut choisir $t = -\frac{\lambda_k}{\mu_k}$ pour un indice k au moins, tel que de plus

$$t \geqslant -\frac{\lambda_i}{\mu_i} \text{ si } \mu_i > 0 \quad \text{et} \quad t \leqslant -\frac{\lambda_i}{\mu_i} \text{ si } \mu_i < 0.$$

Pour cela, il suffit de choisir pour k l'indice (ou l'un des indices) i tel que $|\frac{\lambda_i}{\mu_i}|$ soit minimal. □

Ce qui est profond dans le théorème de Carathéodory, c'est que l'on passe d'un nombre de termes p indéterminé, arbitrairement grand, à une taille de combinaison convexe uniformément bornée. Ce principe de finitude permet d'obtenir le

Corollaire I .14 *Dans un espace vectoriel normé de dimension finie, l'enveloppe convexe de toute partie compacte est compacte.*

Preuve. Soit A un compact de $\mathbf{R}^n$. Dans $\mathbf{R}^{n+1}$, l'ensemble

$$\Lambda = \{(\lambda_1, \lambda_2, \ldots \lambda_{n+1}) \in \mathbf{R}^{n+1} : \lambda_i \geqslant 0, \sum_{i=1}^{n+1} \lambda_i = 1\}$$

est compact, car fermé et borné (Borel-Lebesgue). Or d'après le théorème de Carathéodory, conv(A) est l'image du compact $\Lambda \times A^{n+1}$ par l'application continue

$$(\lambda_1, \lambda_2, \dots \lambda_{n+1}, x_1, x_2, \dots x_{n+1}) \mapsto \sum_{i=1}^{n+1} \lambda_i x_i$$

de $\mathbf{R}^{n+1} \times \mathbf{R}^{n+1}$ dans $\mathbf{R}^{n+1}$. Donc conv(A) est compacte. □

On remarquera que le corollaire traite de l'enveloppe convexe de A et non pas de l'enveloppe convexe fermée. Le résultat n'est pas valabe en dimension infinie, comme le montre l'exercice qui suit.

Exercice I .15 *Dans l'espace de suites $\ell^\infty(\mathbf{N})$, l'ensemble $\{\frac{e_n}{n}\}_{n \geqslant 1} \cup \{0\}$ est compact mais son enveloppe convexe n'est pas fermée.*

Cependant, en dimension infinie subsiste le résultat moins fort que voici.

Théorème I .16 *Dans un espace de Banach, l'enveloppe convexe fermée de toute partie compacte est compacte.*

Preuve. Notons E l'espace de Banach ambiant. Dans cette démonstration, la notation $B(a, \varepsilon)$ désigne la boule de centre a et de rayon ε. Par ailleurs, par l'inégalité triangulaire, on a la formule :

$$B(a, \varepsilon) = B(a, \frac{\varepsilon}{2}) + B(0, \frac{\varepsilon}{2}).$$

Soit A un compact dans E. On sait déjà que $\overline{\text{conv}(A)}$ est fermée, donc ici complète puisque E l'est par hypothèse. Il suffit donc de voir que $\overline{\text{conv}(A)}$ est précompacte, i.e. qu'elle satisfait à la condition du ε-recouvrement (pour $\varepsilon > 0$ arbitrairement petit), qui est que : pour tout $\varepsilon > 0$, la partie $\overline{\text{conv}(A)}$ admet un recouvrement par une collection finie de boules de rayon ε. Or, puisque A est compact, il existe une collection finie de boules

$$B_i = B(b_i, \frac{\varepsilon}{2}) = \{b_i\} + B(0, \frac{\varepsilon}{2})$$

avec $1 \leqslant i \leqslant N$, telle que $A \subset B_1 \cup B_2 \cup \dots B_N$. On note C l'ensemble conv$(\{b_1; b_2; \dots b_N\})$: il est compact d'après le corollaire précédent appliqué dans l'espace réel de dimension finie engendré par les b_i. Ainsi, il existe $a_1, a_2, \dots a_m \in E$ tels que

$$C \subset \bigcup_{i=1}^{m} B(a_i, \frac{\varepsilon}{2}).$$

Finalement, on obtient la chaîne d'inclusions suivante :

$$\operatorname{conv}(A) \subset \operatorname{conv}\left(\bigcup_i B_i\right) = \operatorname{conv}\left(\{b_1; b_2; \dots b_N\} + B(0, \frac{\varepsilon}{2})\right)$$

$$= C + B(0, \frac{\varepsilon}{2}) \subset [\bigcup_{i=1}^{m} B(a_i, \frac{\varepsilon}{2})] + B(0, \frac{\varepsilon}{2}) = \bigcup_{i=1}^{m} B(a_i, \varepsilon),$$

qui permet de conclure en passant à l'adhérence. □

I.3 Théorème de Helly

Les rapports de la notion de convexité avec celle de dimension, quand elle est finie, apparaissent déjà dans le théorème de Carathéodory : la dimension de E est le plus petit entier n tel qu'on est assuré d'obtenir, par combinaisons linéaires convexes de $n + 1$ éléments de A, l'enveloppe convexe de toute partie A de E. Voici une autre manière de caractériser la dimension par la convexité.

Théorème I .17 *Dans $\mathbf{R}^n$ soient $r \geqslant n+1$ ensembles convexes $A_1, A_2, \dots A_r$. On suppose que, chaque fois qu'on choisit $n + 1$ de ces ensembles, leur intersection est non vide. Alors $A_1 \cap A_2 \cap \dots A_r \neq \varnothing$.*

Songer au cas où $n = 2$ et où r vaut un milliard !

Remarquons qu'on ne peut abaisser $n + 1$ en n dans l'énoncé : dans $\mathbf{R}^2$ prenons les $r = 3$ ensembles convexes que sont les côtés d'un triangle ; deux à deux ils ont des points en commun mais leur intersection globale est vide.

Remarquons aussi que l'hypothèse de convexité sur les A_i est nécessaire également (prendre $n = 1$, $A_1 = \{0; 1\}$, $A_1 = \{1; 2\}$ et $A_1 = \{0; 2\}$).

Preuve. La démonstration se fait par récurrence sur r. Si $r = n + 1$, la conclusion coïncide avec l'hypothèse. Supposons désormais que $r \geqslant n + 2$, et que le théorème vrai pour les familles de $r - 1$ de parties convexes. Par hypothèse de récurrence, pour tout $i = 1, 2, \dots r$ l'ensemble $\bigcap_{j=1, j\neq i}^{r} A_j$ est non vide ; soit $\underline{x}_i \in \bigcap_{j=1, j\neq i}^{r} A_j$. Considérons le système linéaire homogène en les inconnues réelles $\lambda_1, \lambda_2, \dots \lambda_r$:

$$\begin{aligned} &\lambda_1 \underline{x}_1 + \lambda_2 \underline{x}_2 + \dots \lambda_r \underline{x}_r = 0 \\ &\lambda_1 + \lambda_2 + \dots \lambda_r = 0 \end{aligned}$$

c'est-à-dire, scalairement, et en posant $\underline{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$, le système de $n + 1$ équations à r inconnues :

$$x_{11}\lambda_1 + x_{21}\lambda_2 + \dots x_{r1}\lambda_r = 0$$

$$
\begin{aligned}
&x_{12}\lambda_1 + x_{22}\lambda_2 + \dots x_{r2}\lambda_r = 0\\
&\dots\\
&x_{1r}\lambda_1 + x_{2r}\lambda_2 + \dots x_{rr}\lambda_r = 0\\
&\lambda_1 + \lambda_2 + \dots \lambda_r = 0
\end{aligned}
$$

Il y a $n+1$ équations pour $r \geqslant n+2$ inconnues, donc le théorème de Cramer assure qu'il existe des λ_i non tous nuls qui sont solutions du système ci-dessus. Renumérotons ces solutions de sorte que pour un indice p les λ_i avec $i \leqslant p$ soient $\geqslant 0$ et que $\lambda_{p+1}, \lambda_{p+2}, \dots \lambda_r < 0$. Aucun des deux groupes de λ_i n'est vide car la somme de tous les λ_i vaut 0 (dernière équation du système). En outre les λ_i avec $i \leqslant p$ sont non tous nuls.

Le système se réécrit alors

$$
\begin{aligned}
&\sum_{i=1}^{p} \lambda_i \underline{x}_i = \sum_{j=p+1}^{r} (-\lambda_j)\underline{x}_j\\
&\sum_{i=1}^{p} \lambda_i = \sum_{j=p+1}^{r} (-\lambda_j)
\end{aligned}
$$

On définit alors un vecteur par combinaison linéaire de deux façons :

$$
\underline{x} = \frac{\sum_{i=1}^{p} \lambda_i \underline{x}_i}{\sum_{i=1}^{p} \lambda_i} = \frac{\sum_{j=p+1}^{r} (-\lambda_j)\underline{x}_j}{\sum_{j=p+1}^{r} (-\lambda_j)}.
$$

Par sa première écriture, on voit que $\underline{x}$ appartient à chaque A_j pour $j > p$; et par la seconde écriture, on voit que $\underline{x}$ appartient à chaque A_i pour $i \leqslant p$. Ceci achève la démonstration. □

Le théorème de Helly admet une formulation topologique. Rappelons que si $\mathcal{F}$ est une famille (finie ou infinie) de parties compactes dans un espace métrique, telle que l'intersection de toute sous-famille finie soit non vide, alors l'intersection de la famille $\mathcal{F}$ tout entière est non vide. On améliore beaucoup cet énoncé si l'on suppose de plus que les parties compactes sont convexes. Le théorème de Helly fournit alors aussitôt le

Corollaire I .18 *Soit $\mathcal{F}$ une famille finie ou infinie de parties convexes et compactes de $\mathbf{R}^n$. Pour que l'intersection de toutes les parties de $\mathcal{F}$ soit non vide, il suffit que l'intersection de toute sous-famille de $n+1$ parties soit non vide.*

Exercice I .19 *Dans le plan, soit une famille de segments parallèles deux à deux non alignés (supports deux à deux distincts) et admettant trois à trois une droite sécante commune. Alors la famille tout entière admet une droite sécante commune. Indication : on commencera par le cas d'une famille finie ;*

en prenant des axes tels que (Oy) soit parallèle aux segments donnés, pour tout segment S de la famille, soit A_S l'ensemble des $(\alpha, \beta) \in \mathbf{R}^2$ tels que la droite $\{y = \alpha x + \beta\}$ coupe S. Dans le plan auxiliaire des (α, β) représenter graphiquement A_S et appliquer à la famille des A_S le théorème de Helly.

Exercice I .20 *Dans $\mathbf{R}^n$, soit A un ensemble convexe compact et d'intérieur A° non vide.*

1° *Montrer qu'il existe au moins un point $z \in A^\circ$ tel que toute « corde » $[u; v]$ de A passant par z vérifie l'inégalité*

$$\frac{1}{n} \leqslant \frac{uz}{zv} \leqslant n.$$

Indication : remarquer que, si de tels z existent, ce sont les points d'intersection des homothétiques A_u de A par les homothéties de rapport $\frac{n}{n+1}$ et de centre u quand le paramètre u parcourt la frontière de A ; prouver que $A_u = \frac{1}{n+1}u + \frac{n}{n+1}A$, et appliquer aux A_u le corollaire du théorème de Helly.

2° *Soit A un simplexe de $\mathbf{R}^n$, i.e. l'enveloppe convexe de $\{0; x_1; x_2; \ldots x_n\}$ où les x_i forment une base de $\mathbf{R}^n$. Montrer qu'alors z est unique et qu'il est l'isobarycentre des sommets de A, et que dans ce cas les bornes n et $\frac{1}{n}$ de la double inégalité sont atteintes (on traitera d'abord à la main le cas $n = 2$).*

I.4 Théorème de Hahn-Banach

Soit E un espace vectoriel sur $\mathbf{R}$. Par définition une *jauge* sur E est une fonction p définie dans E, à valeurs réelles, telle que pour tous $x, y \in E$ et $\lambda \in \mathbf{R}_{\geqslant 0}$, on ait :

$$p(\lambda x) = \lambda p(x) \quad \text{et} \quad p(x + x) \leqslant p(x) + p(y).$$

On dit que p est une fonction positivement homogène et sous-additive, respectivement.

Les formes linéaires et les normes sur E sont des exemples de jauge, mais voici des cas plus intéressants.

Proposition I .21 *Soit E un espace normé sur $\mathbf{R}$, et soit A un ensemble ouvert convexe dans E, contenant l'origine 0. Pour tout $x \in E$, on pose :*

$$p_A(x) = \inf\{t > 0 : \frac{x}{t} \in A\}.$$

Alors :

(i) *le point x est dans A si et seulement si $p_A(x) < 1$;*

(ii) *l'application p_A est une jauge.*

On appelle p_A la *jauge de* A.

Preuve. Prouvons (i). Soit $x \in A$. Si x est l'origine, on a $p_A(x) = 0 < 1$; supposons désormais que $x \neq 0$. L'intersection de A avec la demi-droite ouverte issue de 0 passant par x est un intervalle ouvert contenant x, donc il existe $t \in]0;1[$ tel que $\frac{x}{t} \in A$ et par suite $p_A(x) < 1$. Réciproquement, soit $x \in E$ tel que $p_A(x) < 1$. Par définition de la borne inférieure, il existe $t_0 < 1$ tel que $\frac{x}{t_0} \in A$; par convexité, le segment $[0; \frac{x}{t_0}]$ est contenu dans A. Comme x appartient à ce segment (puisque $t_0 < 1$), on a $x \in A$.

Prouvons (ii). Soient $x \in E$ et $\lambda > 0$. On a, en posant $\frac{t}{\lambda} = t'$:

$$p_A(\lambda x) = \inf\{t > 0 : \frac{\lambda x}{t} \in A\} = \inf\{t > 0 : \frac{x}{(t/\lambda)} \in A\}$$

$$= \inf\{\lambda t' : t' > 0, \frac{x}{t'} \in A\} = \lambda \inf\{t' > 0 : \frac{x}{t'} \in A\} = \lambda p_A(x),$$

ce qui prouve l'homogénéité positive de p_A.

Soient $x, y \in E$ et $\varepsilon > 0$. Par définition de la borne inférieure, il existe $t > 0$ tel que $\frac{x}{t} \in A$ et $t \leqslant p_A(x) + \varepsilon$, et il existe $s > 0$ tel que $\frac{y}{s} \in A$ et $s \leqslant p_A(y) + \varepsilon$. Par convexité de A, on a :

$$\frac{1}{s+t}(x+y) = \frac{t\frac{x}{t} + s\frac{y}{s}}{s+t}$$

est dans A, et $s + t > 0$ donc :

$$p_A(x+y) \leqslant t + s \leqslant p_A(x) + p_A(y) + 2\varepsilon,$$

d'où, vu que ε est arbitrairement petit, l'inégalité :

$$p_A(x+y) \leqslant p_A(x) + p_A(y),$$

qui est la sous-additivité recherchée. □

Exercice I .22 *Prouver que la jauge p_A est une fonction uniformément continue dans E.*

La démonstration du théorème de Hahn-Banach repose sur une récurrence transfinie, dont nous commençons par extraire l'analogue du passage de n à $n+1$ dans une récurrence ordinaire, en énonçant le

Lemme I .23 *Soit V_1 un espace vectoriel réel. Soit V_0 un sous-espace vectoriel de codimension 1 dans V_1, ce qui signifie qu'il existe $x_1 \in V_1$ tel que $V_1 = V_0 \oplus \mathbf{R}x_1$. Soit p une jauge sur V_1. Soit f_0 une forme linéaire sur V_0 telle que pour tout $x \in V_0$ on a : $f_0(x) \leqslant p(x)$. Alors il existe une forme linéaire f_1 sur V_1 telle que :*

(i) *la forme linéaire f_1 prolonge f_0, i.e. $f_1(x) = f_0(x)$ pour tout $x \in V_0$;*
(ii) *pour tout $x \in V_1$, on a : $f_1(x) \leqslant p(x)$.*

Réflexion préliminaire : il est très facile de prolonger f_0 en une forme linéaire f_1 sur V_1 ; il suffit de choisir une constante réelle α et de poser, pour tout $\lambda \in \mathbf{R}$ et tout $y \in V_0$:

$$f_1(y + \lambda x_1) = f_0(y) + \lambda\alpha.$$

La subtilité est de choisir la constante α pour conserver le contrôle de la forme par la jauge.

Preuve. Quels que soient $y \in V_0$ et $z \in V_0$, on a :

$$f_0(z) + f_0(y) = f_0(z + y) \leqslant p(z + y) \leqslant p(z + x_1) + p(y - x_1),$$

donc

$$f_0(y) - p(y - x_1) \leqslant -f_0(z) + p(z + x_1)$$

pour y et z arbitraires dans V_0 ; par conséquent

$$\sup_{y \in V_0} \left(f_0(y) - p(y - x_1)\right) \leqslant \inf_{z \in V_0} \left(p(z + x_1) + p(y - x_1)\right).$$

Il y a donc des nombres réels entre les bornes supérieure et inférieure ci-dessus ; autrement dit, il existe $\alpha \in \mathbf{R}$ tel que

pour tout $y \in V_0$, on a : $f_0(y) - \alpha \leqslant p(y - x_1)$;
pour tout $y \in V_0$, on a : $f_0(y) + \alpha \leqslant p(y + x_1)$.

Pour tout $y + \lambda x_1 \in V_1$, où $y \in V_0$ et $\lambda \in \mathbf{R}$, décidons de poser :

$$f_1(y + \lambda x_1) = f_0(x) + \lambda\alpha,$$

définition non ambigüe puisqu'il y a une seule manière d'écrire un élément de V_1 sous la forme $y + \lambda x_1$.

Il est clair que f_1 est linéaire sur V_1 et que la restriction de f_1 à V_0 est f_0. Il reste à voir que, pour $\lambda \neq 0$, on a $f_1(y + \lambda x_1) \leqslant p(y + \lambda x_1)$. Distinguons deux cas :

si $\lambda > 0$, on applique la seconde inégalité satisfaite par α avec $\frac{y}{\lambda}$ au lieu de y, pour obtenir :

$$f_1(y+\lambda x_1) = f_0(y)+\lambda\alpha = \lambda\left(f(\frac{y}{\lambda}) + \alpha\right) \leqslant \lambda p\left(\frac{y}{\lambda} + x_1\right) = p(y+\lambda x_1);$$

si $\lambda < 0$, on applique la première inégalité satisfaite par α avec $\frac{y}{-\lambda}$ au lieu de y, pour obtenir :

$$f_1(y+\lambda x_1) = f_0(y)+\lambda\alpha = -\lambda\left(f(\frac{y}{-\lambda}) - \alpha\right) \leqslant (-\lambda)p\left(\frac{y}{-\lambda} - x_1\right) = p(y+\lambda x_1).$$

Ceci prouve la majoration cherchée de la forme linéaire prolongée par la jauge imposée. □

Théorème I .24 (Hahn-Banach, version 1) *Soit E un espace vectoriel sur $\mathbf{R}$ et soit p une jauge sur E. Soit V_0 un sous-espace vectoriel de E et soit f_0 une forme linéaire sur V_0 telle que pour tout $x \in V_0$ on ait : $f_0(x) \leqslant p(x)$. Alors il existe une forme linéaire F sur E telle que*

(i) *la forme linéaire F prolonge f_0, i.e. $F(x) = f_0(x)$ pour tout $x \in V_0$;*
(ii) *pour tout $x \in E$, on a : $F(x) \leqslant p(x)$.*

Commençons par indiquer une démonstration dans un cas particulier instructif : supposons en effet qu'il existe une suite, finie ou infinie, de vecteurs $v_1, v_2, \dots v_n, \dots$ dans E telle qu'en posant $V_n = V_{n-1} + \mathbf{R}v_n$ pour tout $n \geqslant 1$, le vecteur v_n ne soit pas dans V_{n-1} (et donc que $V_n = V_{n-1} \oplus \mathbf{R}v_n$) et qu'on ait $E = \bigcup_n V_n$. Par récurrence (ordinaire) sur n, en appliquant le lemme précédent, on obtient pour tout n une forme linéaire f_n sur V_n telle que la restriction de f_n à V_{n-1} soit f_{n-1} et telle que pour tout $x \in V_n$ on ait $f_n(x) \leqslant p(x)$. En particulier, $f_n|_{V_0} = f_0$. Pour x fixé dans E, posons $F(x) = f_n(x)$ dès lors que $x \in V_n$; cette définition ne dépend pas du n en question car si $m \leqslant n$ et si $x \in V_m \subset V_n$, on a $f_m(x) = f_n(x)$. En outre il est clair que F est une forme linéaire sur E, qu'elle prolonge f_0 et qu'elle est dominée par la jauge p.

Preuve. Dans le cas général, la démonstration du théorème de Hahn-Banach se fait par récurrence transfinie, c'est-à-dire s'appuie sur l'axiome de Zorn. Soit Λ l'ensemble des couples (V, f) où V est un sous-espace vectoriel de E contenant V_0 et f est une forme linéaire sur V prolongeant f_0 et telle que $f(x) \leqslant p(x)$ pour tout $x \in V$. On définit sur Λ une relation d'ordre en posant $(V, f) \preccurlyeq (V', f')$ si, et seulement si, on a :

$$V' \supset V \quad \text{et} \quad f'|_V = f.$$

L'ensemble Λ est non vide car il contient (V_0, f_0). Dans Λ une famille non vide et totalement ordonnée pour $\preccurlyeq$, disons $\mathcal{T}$, est une famille $(V_i, f_i)_{i \in I}$ non vide d'éléments de Λ telle que si $i \neq j$ sont dans I, ou bien $(V_i, f_i) \preccurlyeq (V_j, f_j)$ ou bien $(V_j, f_j) \preccurlyeq (V_i, f_i)$. Pour un telle famille $\mathcal{T}$, posons $V_\mathcal{T} = \bigcup_i V_i$: c'est un sous-espace vectoriel de E qui contient V_0, et pour $x \in V_\mathcal{T}$ posons $f_\mathcal{T}(x) = f_i(x)$ pour un indice $i \in I$ (disponible) choisi de telle sorte que $x \in V_i$; cette définition est indépendante du choix de i en question. Alors $f_\mathcal{T}$ est une forme linéaire sur $V_\mathcal{T}$, elle prolonge f_0 et est majorée par p sur $V_\mathcal{T}$. Autrement dit, on vient de vérifier que $(V_\mathcal{T}, f_\mathcal{T})$ appartient à Λ ; en fait, dans l'ensemble partiellement ordonné $(\Lambda, \preccurlyeq)$ l'élément $(V_\mathcal{T}, f_\mathcal{T})$ est une borne supérieure de l'ensemble totalement ordonné $\mathcal{T}$.

Le paragraphe précédent justifie que l'on peut appliquer le lemme de Zorn dans $(\Lambda, \leqslant)$ afin d'obtenir un élément maximal, disons (W, F), et il reste à voir que $W = E$. Sinon, par l'absurde il existerait dans E un vecteur $v \notin W$ et le lemme précédent assurerait l'existence, dans $W' = W \oplus \mathbf{R}v$, d'une forme linéaire F' prolongeant F, donc f_0 et telle que $F'(x) \leqslant p(x)$ pour tout $x \in W'$. Ainsi on aurait $(W, F) \leqslant (W', F')$ avec $(W, F) \neq (W', F')$, ce qui contredirait la maximalité de (W, F). □

Corollaire I .25 *Soit E un espace vectoriel sur $\mathbf{R}$ et soit p une jauge non nulle sur E. Alors il existe sur E une forme linéaire non nulle F sur E telle que $F(x) \leqslant p(x)$ pour tout $x \in E$.*

Preuve. Choisissons $x_0 \in E$ tel que $p(x_0) \neq 0$, posons $V_0 = \mathbf{R}x_0$ et définissons f_0 sur V_0 par $f_0(\lambda x_0) = \lambda p(x_0)$. Alors f_0 est une forme linéaire non nulle sur V_0, et il suffit de voir que $f_0 \leqslant p|_{V_0}$ pour conclure grâce au théorème de Hahn-Banach qui précède. Faisons donc cette vérification. Pour $y = \lambda x_0$, avec $\lambda \geqslant 0$, on a :

$$f_0(y) = f_0(\lambda x_0) = \lambda p(x_0) = p(\lambda x_0) = p(y) \leqslant p(y),$$

et si $y = \lambda x_0$, cette fois avec $\lambda < 0$, on a : $0 = p(0) \leqslant p(x_0) + p(-x_0)$ donc $-p(x_0) \leqslant p(-x_0)$ et, par conséquent :

$$f(y) = f_0(\lambda x_0) = \lambda p(x_0) = (-\lambda)\left(-p(x_0)\right) \leqslant (-\lambda) p(-x_0) = p\left((-\lambda)(-x_0)\right) = p(y),$$

ce qui conclut la preuve. □

Exercice I .26 *Soit $E = \mathcal{B}(\mathbf{R})$ l'espace vectoriel sur $\mathbf{R}$ des fonctions $x = x(t)$ définies et bornées sur $\mathbf{R}$, à valeurs réelles. Si $\alpha_1, \alpha_2, \dots \alpha_n$ est un ensemble fini de nombres réels, on pose :*

$$M(x; \alpha_1, \alpha_2, \dots \alpha_n) = \sup_{t \in \mathbf{R}} \frac{1}{n} \sum_{k=1}^{n} x(t + \alpha_k),$$

et

$$p(x) = \inf_{\alpha_1, \alpha_2, \dots \alpha_n} M(x; \alpha_1, \alpha_2, \dots \alpha_n),$$

où l'infimum est pris sur tous les systèmes finis $\alpha_1, \alpha_2, \dots \alpha_n$ de nombres réels.

1° *L'application p est une jauge.*

2° *Il existe sur E une forme linéaire F telle que : $F(1) = 1$, $F(x) \geqslant 0$ si $x(t) \geqslant 0$ pour tout $t \in \mathbf{R}$ et F est invariante par translations, i.e. $F(x_\tau) = F(x)$ pour toute fonction x, où x_τ est la fonction « translatée par τ » définie par $t \mapsto x(t - \tau)$.*

3° *En déduire qu'à toute partie A de* $\mathbf{R}$ *on peut attribuer un nombre réel* $\mu(A)$ *de sorte que :* $\mu(A \cup B) = \mu(A) + \mu(B)$ *si* $A \cap B = \varnothing$, $\mu(A) \geqslant 0$, $\mu(\mathbf{R}) = 1$ *et* $\mu(A) = \mu(B)$ *si A et B se déduisent l'un de l'autre par translation.*

On remarquera l'analogie avec la théorie de la mesure des ensembles.

Soit maintenant E un espace vectoriel normé sur le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$. Rappelons que le dual (topologique) de E est l'espace vectoriel E' des formes linéaires f continues sur E, normé par

$$\|f\|_{E'} = \sup_{x \in E, \|x\| \leqslant 1} |f(x)|.$$

L'espace E' est un espace de Banach (i.e. normé complet).

Théorème I .27 (Hahn-Banach, version 2) *Soit E un espace vectoriel normé sur le corps* $\mathbf{K} = \mathbf{R}$ *ou* $\mathbf{C}$. *Soit V un sous-espace vectoriel de E et soit f une forme linéaire continue sur V. Alors il existe une forme linéaire F continue sur E telle que*

(i) *la forme linéaire F prolonge f, i.e. $F(x) = f(x)$ pour tout $x \in V$;*

(ii) *on a :* $\|f\|_{E'} = \|f\|_{V'}$, *c'est-à-dire*

$$\sup_{x \in E, \|x\| \leqslant 1} |F(x)| = \sup_{x \in V, \|x\| \leqslant 1} |f(x)|.$$

Le point essentiel est le (ii) : on peut prolonger la forme linéaire f sans augmenter sa norme.

Preuve. Pour $\mathbf{K} = \mathbf{R}$, c'est un corollaire immédiat de la première version : il suffit d'y prendre $p(x) = \|f\|_{V'} \|x\|_E$, $V_0 = V$ et $f_0 = f$; on obtient alors une forme linéaire F sur E qui prolonge f et vérifie :

$$F(x) \leqslant \|f\|_{V'} \cdot \|x\|_E$$

pour tout $x \in E$. Changeant x en $-x$, on voit qu'elle vérifie même

$$|F(x)| \leqslant \|f\|_{V'} \cdot \|x\|_E$$

pour tout $x \in E$. Donc la forme linéaire F est continue dans E, et

$$\|f\|_{E'} \leqslant \|f\|_{V'},$$

d'où $\|f\|_{E'} = \|f\|_{V'}$.

Pour $\mathbf{K} = \mathbf{C}$, on se ramène au cas $\mathbf{K} = \mathbf{R}$ en considérant l'espace vectoriel réel $E_{\mathbf{R}}$ sous-jacent à E. Soit $V_{\mathbf{R}}$ l'espace vectoriel réel sous-jacent à V. Soit $u = \mathrm{Re}(f)$; c'est une forme linéaire continue sur $V_{\mathbf{R}}$, de norme $\|f\|_{V'}$. On vient de voir qu'elle se prolonge en une forme linéaire réelle U continue dans $E_{\mathbf{R}}$, de norme $\|f\|_{V'}$. Posons, pour tout $x \in E$:

$$F(x) = U(x) - iU(ix).$$

Alors F est une forme linéaire complexe sur E, de norme $\|U\|_{E_{\mathbf{R}}}$, c'est-à-dire égale à $\|f\|_{V'}$, et puisque $f(x) = u(x) - iu(ix)$, il est clair que F prolonge f. □

Exercice I .28 *On suppose que E est un espace de Hilbert dont on note $(\cdot|\cdot)$ le produit scalaire.*

1° *Rappeler la description des formes linéaires continues sur E.*

2° *En utilisant les théorèmes de projections, retrouver une démonstration du théorème de Hahn-Banach dans les espaces de Hilbert.*

3° *Constater que, dans ce cas, le prolongement de f par F avec conservation de la norme est unique et s'annule sur l'orthogonal de V.*

Nous passons désormais à une troisième version, dite géométrique, du théorème de Hahn-Banach.

Théorème I .29 (Hahn-Banach, version 3) *Soit E un espace vectoriel normé sur le corps $\mathbf{R}$. Soit A un ensemble ouvert convexe non vide de E. Soit L un sous-espace affine de E tel que $A \cap L$ soit vide. Alors il existe un hyperplan (affine) fermé H de E tel que $A \cap H$ soit encore vide et contenant L.*

Rappelons qu'un hyperplan fermé est le translaté d'un sous-espace vectoriel fermé de codimension 1 ; c'est-à-dire qu'il s'agit d'un ensemble de la forme $\{x \in E : F(x) = \alpha\}$ où F est une forme linéaire continue non identiquement nulle sur E et α est un nombre réel.

Preuve. Commençons par le cas particulier instructif où L est réduit à un point, disons m (par exemple si E est de dimension 2 et si m est dans la frontière $\mathrm{Fr}(A) = \overline{A} \backslash A^\circ$ de A, il s'agit de voir qu'on peut mener une sorte de « tangente » à A en m).

Par translation, on se ramène au cas où A contient l'origine 0. Soit m un point de E tel que $m \notin A$ (ce qui n'empêche pas m d'être sur la frontière de A). On va prouver qu'il existe un hyperplan fermé H de E contenant m mais ne coupant pas A. Introduisons la jauge p_A de A définie par :

$$p_A(x) = \inf\{t > 0 : \frac{x}{t} \in A\}$$

pour $x \in E$. Par la proposition I .21, on a $p_A(m) \geqslant 1$ car $m \notin A$. Soit $V = \mathbf{R}m$ la droite vectorielle engendrée par m et soit sur V la forme linéaire $f : \lambda \mapsto \lambda p_A(m)$. On a $f(\lambda m) \leqslant p_A(\lambda m)$ car :

— si $\lambda \geqslant 0$, alors $f(\lambda m) = \lambda p_A(m) = p_A(\lambda m)$;
— si $\lambda < 0$, alors $f(\lambda m) = \lambda p_A(m) \leqslant 0 \leqslant p_A(\lambda m)$, car tout $p_A(x)$ est $\geqslant 0$.

D'après le théorème de Hahn-Banach (première version), il existe une forme linéaire F sur E qui prolonge f (et donc en particulier $F(m) = f(m) = p_A(m) \geqslant 1$) et telle que, pour tout $x \in E$, on a $F(x) \leqslant p_A(x)$ (et en particulier $F(x) < 1 < p_A(m)$ pour tout $x \in A$ par la proposition I .21). Posons $\alpha = p_A(m)$. L'hyperplan

$$H = \{x \in E : F(x) = \alpha\}$$

passe par m et, pour tout $x \in A$, on a $F(x) < \alpha$, donc $H \cap A = \varnothing$. Ceci implique que l'hyperplan H n'est pas dense, et donc qu'il est fermé dans E. Ceci achève la démonstration dans le cas particulier où L est réduit à un point.

Passons maintenant à la démonstration dans le cas général. Cette fois, on utilise une translation judicieuse pour se ramener au cas où L contient l'origine 0, autrement dit au cas où L est un sous-espace vectoriel de E. On suppose aussi que L est fermé car, A étant ouvert, l'intersection $L \cap \overline{A}$ est encore vide. Passons à l'espace vectoriel normé quotient $\overline{E} = E/L$, muni de la norme

$$\|\bar{x}\| = \inf_{x \in \bar{x}} \|x\|_E,$$

où $\bar{x} = x + L$. Dans $\overline{E}$ l'ensemble $\overline{A} = \{\bar{x} : x \in A\}$ est convexe et ouvert, car la projection $x \mapsto \bar{x}$ est une application linéaire de E sur $\overline{E}$ qui transforme tout ouvert en un ouvert (on parle d'application *ouverte*). De plus, on a $\bar{0} = L \notin \overline{A}$. D'après le cas particulier traité ci-dessus, il existe dans $\overline{E}$ un hyperplan vectoriel fermé $\mathcal{H}$ tel que $\mathcal{H} \cap \overline{A}$ soit vide. Soit H l'image réciproque de $\mathcal{H}$ dans E ; c'est un hyperplan fermé de E tel que $H \supset L$ et $H \cap A = \varnothing$, ce qui achève la démonstration. □

Proposition I .30 *Si A est ouvert convexe non vide et ne contient pas 0, il existe une forme linéaire continue $f \in E'$ telle que $f(x) > 0$ pour tout $x \in A$.*

Preuve. En effet, d'après le cas particulier traité dans la preuve précédente, il existe un hyperplan vectoriel fermé $\{f = 0\} = \{x \in E : f(x) = 0\}$ qui ne coupe pas A. Puisque A est convexe, il est tout entier dans l'un des deux demi-espaces ouverts délimités par cet hyperplan ; quitte à remplacer f par $-f$, il est dans le demi-espace $\{f > 0\}$. □

Exercice I .31 *Dans le cas $E = \mathbf{R}^2$, donner de la troisième version du théorème de Hahn-Banach une démonstration élémentaire basée sur l'idée suivante. Soit A un ouvert convexe de $\mathbf{R}^2$ ne contenant pas 0 ; il faut montrer* *qu'il existe une droite issue de* 0 *et ne coupant pas* A. *Soit* Γ *le cercle unité et, pour tout* $x \in A$, *soit* $\varphi(x)$ *l'intersection avec* Γ *de la demi-droite issue de* 0 *passant par* x. *On montrera que* $\varphi(A)$ *est un intervalle curviligne ouvert dans* Γ *et de longueur* $\leqslant \pi$.

I.5 Applications à la dualité des espaces normés

Relisez attentivement l'énoncé de la deuxième version du théorème de Hahn-Banach. En voici quelques corollaires.

Corollaire I .32 *Soit, dans un espace vectoriel normé sur le corps* $\mathbf{K} = \mathbf{R}$ *ou* $\mathbf{C}$, *un vecteur* a *non nul. Alors il existe* $f \in E'$ *telle que*

$$\|f\|_{E'} = 1 \quad \text{et} \quad f(a) = \|a\|.$$

En particulier, si $E \neq \{0\}$ *alors* $E' \neq \{0\}$ *: un espace vectoriel normé sur* $\mathbf{K}$ *a toujours des formes linéaires continues non nulles.*

Preuve. Soit la droite vectorielle $V = \mathbf{K}a$. Sur V la forme linéaire $f_0 : \lambda a \mapsto \lambda\|a\|$ est continue de norme 1 ; il suffit de la prolonger à E sans changement de norme. □

Corollaire I .33 *Le dual topologique* E' *sépare les points : si* x *et* y *sont deux vecteurs distincts de* E, *il existe* $f \in E'$ *telle que* $f(x) \neq f(y)$.

Preuve. En effet, si $x \neq y$ il suffit de prendre $a = x - y$ et d'appliquer le précédent corollaire. □

Corollaire I .34 *On a, pour tout* $x \in E$, *la formule :*

$$\|x\|_E = \sup_{f \in E', \|f\|_{E'} \leqslant 1} |f(x)|.$$

Preuve. Si $x = 0$ c'est évident ; supposons désormais que x est non nul. Si $\|f\|_{E'} \leqslant 1$, on a $|f(x)| \leqslant \|x\|_E$; ainsi on sait déjà que

$$\sup_{f \in E', \|f\|_{E'} \leqslant 1} |f(x)| \leqslant \|x\|_E.$$

Réciproquement, puisque $x \neq 0$ le premier corollaire de la section assure l'existence de $f \in E'$ de norme 1 telle que $|f(x)| = \|x\|$; et donc

$$\sup_{f \in E', \|f\|_{E'} \leqslant 1} |f(x)| \geqslant \|x\|_E,$$

ce qui prouve l'énoncé. □

On appelle *bidual* de l'espace vectoriel E normé sur $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$ le dual (topologique) $E'' = (E')'$ de l'espace de Banach E'. D'après le corollaire qui précède, l'application

$$x \mapsto (e_x : f \mapsto f(x))$$

établit une isométrie de E dans son bidual E'', qui permet d'identifier E à un sous-espace vectoriel normé de E''.

Exercice I .35 1° *Quel est le bidual de l'espace c_0 des suites réelles qui tendent vers 0, muni de la norme sup ?*

2° *Citer des espaces vectoriels normés égaux à leur bidual (ces espaces vectoriels normés sont dits réflexifs).*

3° *En considérant le sous-espace vectoriel V de $\ell^\infty(\mathbf{N})$ formé des suites réelles convergentes, et sur V la forme linéaire*

$$x = (x_n)_{n \in \mathbf{N}} \mapsto \lim_{n \to \infty} x_n,$$

montrer qu'il existe sur $\ell^\infty(\mathbf{N})$ une forme linéaire continue non nulle mais identiquement nulle sur c_0. En déduire que l'espace $\ell^1(\mathbf{N})$ n'est pas réflexif.

Corollaire I .36 *Soit A une partie convexe non vide de E. Soit $x \in E$. Pour que x soit dans E limite de combinaisons linéaires de vecteurs de A, il faut et il suffit que toute $f \in E'$ identiquement nulle sur A soit nulle sur x.*

Preuve. Notons $\langle A \rangle$ le sous-espace vectoriel de E engendré par A (i.e. l'ensemble des combinaisons linéaires de vecteurs de A) et notons $\overline{\langle A \rangle}$ son adhérence dans E (i.e. l'ensemble des limites de suites convergentes de telles combinaisons linéaires). La condition proposée est nécessaire car si $x \in \overline{\langle A \rangle}$ et si f est identiquement nulle sur A, alors f est identiquement nulle sur $\langle A \rangle$ par linéarité, et finalement identiquement nulle sur $\overline{\langle A \rangle}$ par continuité et passage à la limite.

Le point intéressant du corollaire, qui demande le plus de travail pour la preuve, est de démontrer que la condition est suffisante. Donnons-nous $x \notin \overline{\langle A \rangle}$ afin de montrer qu'il existe $f \in E'$ telle que f soit nulle sur A mais que $f(x) \neq 0$. Soit $\delta = \inf_{y \in \overline{\langle A \rangle}} \|x - y\|$: c'est un nombre réel > 0, sinon x serait limite d'une suite d'éléments de $\overline{\langle A \rangle}$ donc serait lui-même dans $\overline{\langle A \rangle}$. Posons

$$V = \overline{\langle A \rangle} + \mathbf{K}x.$$

Soit g la forme linéaire sur V définie, pour $y \in \overline{\langle A \rangle}$ et $\lambda \in \mathbf{K}$ par

$$g(y + \lambda x) = \lambda.$$

Puisque, pour $\lambda \neq 0$, on a :

$$\|y + \lambda x\| = |\lambda| \|x + \frac{y}{\lambda}\| \geqslant |\lambda| \delta = \delta |g(y + \lambda x)|,$$

la forme linéaire g est continue sur V et de norme $\|g\|_{V'} \leqslant \frac{1}{\delta}$. D'après le théorème de Hahn-Banach, soit $f \in E'$ une forme linéaire continue qui prolonge g. On a $f(x) = g(x) = 1$, et $f(y) = g(y) = 0$ pour tout $y \in A$. □

Corollaire I .37 *Soit A une partie de E telle que la seule $f \in E'$ identiquement nulle sur A soit la forme nulle. Alors A est totale dans E, c'est-à-dire : tout vecteur de E est limite de combinaisons linéaires de vecteurs de A.*

Preuve. C'est un cas particulier du corollaire précédent. □

Ces deux derniers corollaires jouent un rôle important en théorie de l'approximation.

Exercice I .38 *De la démonstration de l'avant-dernier corollaire, dégager le résultat suivant : soit V un sous-espace vectoriel de E et soit $x \in E$ tel que $\delta = \inf_{y \in V} \|x - y\| > 0$; alors il existe $f \in E'$ telle que :*

$$f(x) = 1, \quad f(y) = 0 \quad \text{pour tout } y \in V \quad \text{et} \quad \|f\|_{E'} \leqslant \frac{1}{\delta}.$$

I.6 Séparation des ensembles convexes

Dans sa troisième version, le théorème de Hahn-Banach a des applications intéressantes à la géométrie des ensembles convexes. Soit E un espace vectoriel normé sur $\mathbf{R}$. Dans E un *hyperplan (affine) fermé*, de la forme $\{x \in E : f(x) = \alpha\}$ où $f \in E' \backslash \{0\}$ et $\alpha \in \mathbf{R}$, délimite deux *demi-espaces ouverts* :

$$\{x \in E : f(x) < \alpha\} \quad \text{et} \quad \{x \in E : f(x) > \alpha\},$$

ainsi que deux *demi-espaces fermés* :

$$\{x \in E : f(x) \leqslant \alpha\} \quad \text{et} \quad \{x \in E : f(x) \geqslant \alpha\}.$$

Soient A et B deux parties de E. Soit H un hyperplan fermé de E. On dit que H *sépare* (resp. *sépare strictement*) A et B, si A est contenu dans l'un et B est contenu dans l'autre des demi-espaces fermés (resp. ouverts) délimités par H.

Dans sa troisième version, le théorème de Hahn-Banach énonce en somme que, si A est ouvert convexe non vide et L un sous-espace affine disjoint de A, alors il existe un hyperplan fermé qui sépare A de L. Cet énoncé peut en fait être très notablement amélioré.

Corollaire I .39 *Dans un espace vectoriel réel normé E, soient A et B deux ensembles convexes non vides disjoints.*

(i) *Si A est ouvert, il existe un hyperplan fermé qui sépare A et B.*

(ii) *Si A et B sont tous deux ouverts, ou si A est fermé et B compact, il existe un hyperplan fermé qui sépare strictement A et B.*

(iii) *Si A et B sont tous les deux fermés, et si E est de dimension finie, il existe un hyperplan fermé qui sépare A et B.*

Preuve. Preuve de (i). L'ensemble $C = A - B$ est ouvert (comme réunion de translatés de l'ouvert A), convexe (d'après la proposition I .30), non vide car $A \neq \varnothing$ et ne contient pas l'origine (car A et B sont disjoints). D'après la proposition I .30 il existe $f \in E'$ telle que $f(x)$ soit > 0 pour tout $x \in C$, donc telles que $f(a) > f(b)$ quels que soient $a \in A$ et $b \in B$. Ainsi $\alpha = \inf_{a \in A} f(a) \geqslant \sup_{b \in B} f(b) = \beta$. Pour γ un nombre réel dans l'intervalle $[\alpha; \beta]$, l'hyperplan $\{x \in E : f(x) = \gamma\}$ sépare A et B.

Preuve de (ii). Si de plus B est ouvert, l'hyperplan H du (i) sépare strictement A et B, car si par exemple il existait un point b dans $H \cap B$, il existerait une boule ouverte $B(b, r)$ avec $r > 0$, contenue dans B (car B est ouvert) ; donc B aurait des points strictement de part et d'autre de H, ce qui contredit le fait que H sépare A et B.

Maintenant, supposons A fermé et B compact. Alors $\delta = \inf_{x \in B} d(x, A)$ est > 0 car la fonction

$$x \mapsto d(x, A) = \inf_{y \in A} \|x - y\|$$

est continue (en fait 1-lipschitzienne) et > 0 sur le compact B, donc atteint sa borne inférieure (laquelle de ce fait ne peut être nulle). Dès lors les ensembles $A + B(0, \frac{\delta}{3})$ et $B + B(0, \frac{\delta}{3})$ sont deux ouverts convexes disjoints, auxquels il suffit d'appliquer l'énoncé (ii) dans le cas précédent : il existe un hyperplan fermé qui sépare strictement ces deux ouverts épaississant A et B, donc qui sépare strictement A et B.

Preuve de (iii). Ici $E = \mathbf{R}^d$, qu'on munit du produit scalaire $(\cdot|\cdot)$ et de la norme euclidienne (équivalente à toute autre norme). Par translation, on suppose que A contient l'origine ; par rotation et homothétie on suppose que B contient le point $e_1 = (1, 0, 0, \ldots, 0)$. Soit S la sphère unité de centre 0. L'ensemble $\Lambda = S \times [0; e_1]$ est compact ; tout point $\lambda = (u, \xi)$ de cet ensemble, où $u \in S$ et $\xi \in [0; e_1]$, définit un hyperplan H_λ issu de ξ et orthogonal à u : c'est l'hyperplan d'équation $(x|u) = (\xi|u)$. C'est avec cet ensemble d'hyperplans H_λ que nous allons travailler.

Pour tout entier $n \geqslant 1$, soit $\overline{B}(0,n)$ la boule fermée de centre 0 et de rayon n ; elle est compacte puisque E est de dimension finie. D'après le point (ii) il existe un hyperplan H_λ séparant le convexe compact $A \cap \overline{B}(0,n)$ du convexe fermé B. Nécessairement cet hyperplan coupe l'axe e_1 entre les points 0 et e_1, donc en un point ξ_n du segment $[0; e_1]$. Il est donc du type $H_n = H_{\lambda_n}$, avec $\lambda_n = (u_n, \xi_n) \in \Lambda$. La séparation par H_n de $A \cap \overline{B}(0,n)$ et B se traduit, en changeant au besoin u_n en $-u_n$, par les inégalités :

(1) $(x|u_n) \geqslant (\xi_n|u_n)$ pour tout $x \in B$;
(2) $(x|u_n) \leqslant (\xi_n|u_n)$ pour tout $x \in A \cap \overline{B}(0,n)$.

Puisque Λ est compact, il existe une suite extraite de $(\lambda_n)_{n\geqslant 0}$, par abus encore notée $(\lambda_n)_{n\geqslant 0}$, qui converge vers un $\lambda = (u,\xi) \in \Lambda$. Alors l'hyperplan $H_\lambda = \{x \in E : (x|u) = (\xi_n|u)\}$ sépare A et B. En effet, si $x \in B$, en faisant tendre n vers $+\infty$ dans les inégalités (1) on obtient :

$$(x|u) \geqslant (\xi|u) \quad \text{pour tout}\, x \in B.$$

Si $x \in A$, pour n assez grand, x finit par être dans tous les ensembles $A \cap \overline{B}\,(0, \varphi(n))$ et, en faisant dès lors tendre n vers $+\infty$ dans (2), on obtient :

$$(x|u) \leqslant (\xi|u) \quad \text{pour tout}\, x \in A,$$

qui est la séparation souhaitée. □

Corollaire I .40 *Tout ensemble convexe fermé est l'intersection des demi-espaces fermés qui le contiennent.*

Preuve. Soit A un convexe fermé non vide, et soit $\widetilde{A}$ l'intersection des demi-espaces fermés qui contiennent A. Évidemment $A \subset \widetilde{A}$. Soit $x \in E$ tel que $x \notin A$, il s'agit de voir que $x \notin \widetilde{A}$. D'après le corollaire qui précède, il existe un hyperplan fermé H séparant strictement x de A. Ainsi x n'appartient pas au demi-espace fermé délimité par H qui contient A, et on a donc $x \notin \widetilde{A}$. □

Exercice I .41 *Tout sous-espace affine fermé est l'intersection des hyperplans affines qui le contiennent.*

Exercice I .42 *Soit C un convexe fermé et A une partie de E. Pour que C contienne A, il faut et il suffit que pour toute $f \in E'$, on ait $f(A) \subset f(C)$.*

On souhaite maintenant préciser le corollaire précédent en réduisant le nombre de demi-espaces fermés mis en jeu.

Definition I .43 *Soit A un ensemble convexe fermé non vide et soit x un point de la frontière de A. Un* hyperplan d'appui *de A au point x est un hyperplan fermé passant par x et tel que A soit tout entier contenu dans l'un des deux demi-espace fermés délimités par H.*

Proposition I .44 *Soit A un ensemble fermé convexe d'intérieur non vide. Alors tout point frontière de A appartient à au moins un hyperplan d'appui de A.*

Preuve. Soit $x \in \mathrm{Fr}(A)$. Alors $\{x\} \cap A^\circ = \varnothing$. Pour obtenir un hyperplan d'appui en x il suffit d'appliquer la troisième version du théorème de Hahn-Banach au sous-espace affine réduit au point x et à l'ouvert convexe A°. □

Corollaire I .45 *Tout ensemble fermé convexe A qui est, soit d'intérieur non vide, soit compact non vide, est l'intersection des demi-espaces fermés qui le contiennent et qui sont bordés par un hyperplan d'appui de A.*

Preuve. Soit $\widehat{A}$ cette intersection. Évidemment $A \subset \widehat{A}$. Soit $x \in E$ tel que $x \notin A$; il s'agit de voir que $x \notin \widehat{A}$.

1°) Supposons A° non vide. Choisissons $y \in A^\circ$. L'intersection de A et du segment $[x; y]$ est un segment $[z; y]$ où $z \in \mathrm{Fr}(A)$ est distinct de x. Par la proposition qui précède, on dispose d'un hyperplan d'appui de A en z. Le demi-espace fermé délimité par H qui contient A ne contient pas x, sinon la droite joignant x, y, z serait tout entière dans H et comme le point y est intérieur à A, s'il était dans H il y aurait des points de A strictement de part et d'autre de H. Ainsi $x \notin \widehat{A}$.

2°) Supposons A compact non vide. D'après le précédent corollaire, il existe un hyperplan fermé H séparant strictement A de $\{x\}$, donc il existe $f \in E'$ et $\alpha \in \mathbf{R}$ tels que : $f(x) < \alpha$, et $f(y) > \alpha$ pour tout $y \in A$. Soit $\beta = \inf_{y \in A} f(y)$. Le demi-espace fermé $\{y \in E : f(y) \geqslant \beta\}$ contient A. De plus l'hyperplan $H_1 = \{z \in E : f(z) = \beta\}$ est d'appui pour A car sur A compact la borne inférieure β est atteinte. Mais x n'appartient pas au demi-espace $\{y \in E : f(y) \geqslant \beta\}$ car $f(x) < \alpha \leqslant \beta$. Ainsi $x \notin \widehat{A}$. □

Exercice I .46 *Dans $\ell^2(\mathbf{N})$ soit A l'enveloppe convexe fermée de l'ensemble $\{\pm\frac{e_n}{n}\}_{n \geqslant 1}$. Montrer que A est compact, que $0 \in \mathrm{Fr}(A)$ mais que par 0 ne passe aucun hyperplan d'appui de A.*

I.7 Points extrémaux

Soit A un ensemble convexe fermé dans un espace vectoriel normé sur **R**. Soit $a \in A$ On dit que a est un *point extrémal* de A s'il n'existe aucun intervalle ouvert contenant a et contenu dans A ; en d'autres termes si les relations

$$\alpha = \lambda x + (1-\lambda)y, \quad x, y \in A, \quad 0 \leqslant \lambda \leqslant 1$$

entraînent $x = y = a$. Il revient au même de dire que a n'est pas le milieu d'un segment de longueur > 0 contenu dans A ; en d'autres termes les relations

$$a = \frac{x+y}{2} \quad x, y \in A$$

entraînent $x = y = a$.

Par exemple, dans le plan les points extrémaux d'un domaine polygonal convexe sont les sommets du polygone frontière.

Toujours si $n = 2$, l'ensemble A ci-dessus a ses points extrémaux figurés à droite.

On notera $\mathrm{Extr}(A)$ l'ensemble des points extrémaux de A. Il est clair que cet ensemble est contenu dans la frontière de A.

Exercice I .47 *Dans* $\mathbf{R}^n$ *muni de sa base canonique* $(e_i)_{1 \leqslant i \leqslant n}$*, déterminer les points extrémaux de la boule unité fermée*

1° *pour la norme sup :* $x = \sum_i x_i e_i \mapsto \max_i |x_i|$ *;*

2° *pour la norme euclidienne :* $x = \sum_i x_i e_i \mapsto \sqrt{\sum_i |x_i|^2}$.

3° *Plus généralement, montrer que si* E *est un espace vectoriel muni d'un produit scalaire et de la norme associée, tout point de la sphère unité est extrémal pour la boule unité fermée de* E.

4° *Il en est de même si* $E = \ell^p(\mathbf{N})$ *avec* $1 < p < +\infty$.

Exercice I .48 1° *Les points extrémaux de la boule unité fermée de* $\ell^\infty(\mathbf{N})$ *sont les suites* $(x_n)_{n\in\mathbf{N}}$ *tels que* $|x_n| = 1$ *pour tout* $n \in \mathbf{N}$.

2° *La boule unité fermée de l'espace* c_0 *des suites qui tendent vers* 0 *n'a pas de point extrémal.*

Exercice I .49 *Dans l'espace euclidien* $\mathbf{R}^3$*, soit* A *l'enveloppe convexe de l'ensemble réunion du cercle* $\{x_3 = 0; x_1^2 + x_2^2 - 2x_2 = 0\}$ *et des deux points* $\{x_1 = x_2 = 0; x_3 = \pm 1\}$*. Montrer que* A *est convexe compact. Déterminer* $\mathrm{Extr}(A)$ *et constater que cet ensemble n'est pas fermé.*

Lemme I .50 *Soit* A *un ensemble convexe fermé dans* E*, et soit* H *un hyperplan d'appui de* A*. Alors :*

$$\mathrm{Extr}(A \cap H) = \mathrm{Extr}(A) \cap H.$$

Preuve. Par définition des points extrémaux, on a déjà :

$$\mathrm{Extr}(A \cap H) \supset \mathrm{Extr}(A) \cap H.$$

Réciproquement, soit $x \in \mathrm{Extr}(A \cap H)$. Alors déjà $x \in H$. Par l'absurde, supposons que $x \notin \mathrm{Extr}(A)$; alors x serait dans un intervalle ouvert I contenu dans A. Cet I serait contenu dans H, sinon il irait de part et d'autre de H, en contradiction avec le fait que H est un hyperplan d'appui de A. Ainsi $x \notin \mathrm{Extr}(A \cap H)$. □

Un théorème, dû à Krein et Milman, affirme que dans un espace vectoriel normé, tout ensemble convexe compact est l'enveloppe convexe fermée de l'ensemble de ses points extrémaux. Nous nous contenterons ici du cas de la dimension finie mais nous y énonçons un résultat plus fort (remplaçant l'enveloppe convexe fermée par la simple enveloppe convexe).

Théorème I .51 (Krein-Milman) *Soit A un ensemble convexe compact non vide dans un espace vectoriel de dimension finie sur* $\mathbf{R}$. *Alors A est l'enveloppe convexe de l'ensemble de ses points extrémaux.*

Preuve. Il est clair que, si on note $\mathrm{Fr}(A)$ la frontière de A, on a

$$A = \mathrm{conv}\,(\mathrm{Fr}(A)) \,.$$

Il suffit donc de prouver que tout $x \in \mathrm{Fr}(A)$ est dans $\mathrm{conv}\,(\mathrm{Extr}(A))$. On raisonne par récurrence sur la dimension de A (i.e. la dimension du plus petit sous-espace affine contenant A). Si $d = 0$ ou 1, l'ensemble A est un segment $[a; b]$ et $\mathrm{Extr}(A) = \{a; b\}$; le théorème est alors évident.

On suppose désormais que $d > 1$ et que l'énoncé est vrai en dimensions inférieures. Plongeons A dans l'espace affine $\mathbf{R}^d$ qu'il engendre. Dans cet espace l'intérieur de A est non vide donc, d'après la proposition I .44, si $x \in \mathrm{Fr}(A)$ est donné, il passe par x un hyperplan d'appui, disons H, de A. L'ensemble $A \cap H$ est convexe compact non vide et de dimension $\leqslant d-1$; par hypothèse de récurrence, on a donc

$$A \cap H = \mathrm{conv}\,(\mathrm{Extr}(A \cap H)) \,.$$

En appliquant le lemme qui précède, on voit que pour le point $x \in \mathrm{Fr}(A)$, on a :

$$x \in A \cap H = \mathrm{conv}\,(\mathrm{Extr}(A \cap H)) = \mathrm{conv}\,(\mathrm{Extr}(A) \cap H) \subset \mathrm{conv}\,(\mathrm{Extr}(A)) \,,$$

ce qui achève la démonstration. □

Exercice I .52 *Soit A un ensemble convexe compact non vide dans $\mathbf{R}^n$. Soit $f : A \to \mathbf{R}$ une fonction continue et convexe sur A, ce qui signifie que*

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda) f(y)$$

pour tous $x, y \in A$ et $\lambda \in [0;1]$. Alors le maximum de A est atteint sur $\mathrm{Extr}(A)$ *; en particulier, si A est un polyèdre il suffit de tester un nombre fini de points.*

Exercice I .53 *Soit E un espace vectoriel de dimension finie p sur $\mathbf{R}$. Dans E soit K un polyèdre convexe plein défini par le système d'équations*

$$f_1(x) \geqslant 0; \quad f_2(x) \geqslant 0; \quad \ldots f_N(x) \geqslant 0,$$

où $N \geqslant p$ et où les f_i sont des fonctions linéaires affines sur E (i.e. de la forme $f_i = \varphi_i + \alpha_i$ avec $\varphi_i \in E'$ et $\alpha_i \in \mathbf{R}$). Soit a un point extrémal de K. Montrer que le nombre des indices $i \in \{1; 2; \ldots N\}$ tels que $f_i(a) = 0$ est $\geqslant p$; autrement dit, tout sommet d'un polyèdre convexe dans $\mathbf{R}^p$ est dans l'intersection d'au moins p faces dudit polyèdre.

Exercice I .54 *Dans l'espace E de dimension n^2 des matrices réelles $[a_{ij}]_{1 \leqslant i,j \leqslant n}$, soit l'ensemble convexe :*

$$A = \{a = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in E : a_{ij} \geqslant 0 \text{ pour tous } i, j \in \{1; 2; \ldots n\}$$

$$\text{et} \quad \sum_{i=1}^{n} a_{ik} = \sum_{j=1}^{n} a_{kj} = 1 \text{ pour tout } k \in \{1; 2; \ldots n\}\}$$

des matrices dites bistochastiques. Soit P l'ensemble des matrices de permutation, i.e. qui ont un coefficient 1 exactement sur chaque ligne et sur chaque colonne, les autres coefficients étant nuls. Montrer que $P = \mathrm{Extr}(A)$ et en déduire un théorème de Birkhoff : toute matrice bistochastique est combinaison linéaire convexes de matrices de permutations, et même d'au plus $(n-1)^2 + 1$ d'entre elles.

Indication : A est un polyèdre convexe dans l'espace affine

$$\left\{a = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in E : \quad \sum_{i=1}^{n} a_{ik} = \sum_{j=1}^{n} a_{kj} = 1 \text{ pour tout } k \in \{1; 2; \ldots n\}\right\}$$

On montrera que cet espace est de dimension $(n-1)^2$ et on appliquera l'exercice précédent.

Deuxième partie

L'espace des fonctions continues sur un compact

Dans ce chapitre sont exposés quelques aspects de l'Analyse dans l'espace de Banach $\mathcal{C}(X)$ dans fonctions continues sur un espace compact X. Une notion de base est celle de sous-ensemble *équicontinu* de $\mathcal{C}(X)$. Cette propriété a la vertu de rendre uniforme la convergence simple. De plus l'équicontinuité est la seule condition qu'il faut ajouter à celles d'être fermée et bornée pour qu'une partie de $\mathcal{C}(X)$ soit compacte (théorème d'Ascoli). En liaison avec l'équicontinuité nous étudions aussi la compacité dans l'espace $\mathcal{O}(U)$ des fonctions holomorphes dans un ouvert U de $\mathbf{C}$ (théorèmes de Montel, de Vitali) : contrairement au cas de $\mathcal{C}(X)$ et, curieusement, de façon analogue au cas des espaces de dimension finie, le seul fait pour une partie de $\mathcal{O}(U)$ d'être borné suffit pour qu'on ait dans cette partie la propriété de Bolzano-Weierstrass.

La seconde moitié du chapitre est consacrée au théorème de Weierstrass d'approximation uniforme des fonctions continues sur $[0;1]$ par des polynômes, et à la superbe axiomatisation de généralisation à $\mathcal{C}(X)$ qu'en a donnée en 1948 le mathématicien américain Marshall Stone, en mettant en évidence le rôle joué dans la question par la structure d'algèbre de $\mathcal{C}(X)$ et par la relation d'ordre dans cet espace. Le théorème de Weierstrass-Stone a d'importantes conséquences en Analyse fonctionnelle.

Enfin, nous terminons le chapitre par l'étude du prolongement à X d'une fonction définie et continue dans une partie fermée de X (théorème d'Urysohn).

II.8 Équicontinuité et convergence dans les espaces d'applications

Dans tout ce chapitre, X et F désignent des espaces métriques (en fait, pour ceux qui connaissent la Topologie générale : X pourrait être un espace topologique – les modifications des énoncés et des démonstrations étant attendues et mineures). On pensera notamment au cas où $F = \mathbf{R}$ ou $\mathbf{C}$. Soit $\mathcal{H}$ un ensemble d'applications $f : X \to F$. On pensera notamment au cas où $\mathcal{H}$ est un ensemble de fonctions définies dans X, à valeurs réelles ou complexes. De plus, l'ensemble favori sera celui où X est le segment $[0;1]$ de l'axe réel.

Soit a un point de X. On dit que $\mathcal{H}$ est *équicontinu en a* si quel que soit $\varepsilon > 0$ il existe $\eta > 0$ tel que, quelle que soit $f \in \mathcal{H}$, on ait

$$d_X(x, a) < \eta \Rightarrow d_F(f(x), f(a)) < \varepsilon.$$

On dit que l'ensemble $\mathcal{H}$ est *équicontinu dans X* si $\mathcal{H}$ est équicontinu en a pour tout $a \in X$.

Ainsi, l'équicontinuité en a requiert une cohérence dans la façon dont les diverses applications $f \in \mathcal{H}$ sont continues en a : à un $\varepsilon > 0$ donné on peut, dans la définition de la continuité, faire correspondre un unique voisinage de a dans X (donné par la boule de centre a et de rayon η dans le cas métrique que l'on considère) valable d'un seul coup pour toutes les fonctions $f \in \mathcal{H}$.

Tout ensemble fini $\mathcal{H} = \{f_1; f_2; \dots f_N\}$ d'applications continues au point $a \in X$ est équicontinu en a car, si à $\varepsilon > 0$ correspond, pour chaque $i \in \{1; 2; \dots N\}$, un η_i tel que « $d_X(x, a) < \eta_i \Rightarrow d_F(f_i(x), f_i(a)) < \varepsilon$ » il suffit de prendre $\eta = \min_{\in\{1;2;\dots N\}}$. Mais voici des exemples plus intéressants.

Exercice II .1 *Soit* $X = [0; 1]$ *et* $F = \mathbf{C}$ *et soit* k *réel* > 0 *donné. On note* $\mathcal{H}_k$ *l'ensemble des fonctions* $f : X \to \mathbf{C}$ *qui sont* k*-lipschitziennes de rapport* k*, c'est-à-dire telles que quels que soient* $x, y \in X$ *on ait :*

$$|f(x) - f(y)| \leqslant k|x - y|.$$

Montrer que $\mathcal{H}_k$ *est équicontinu dans* X*. En déduire que l'ensemble* $\mathcal{H}'_k$ *des* f *qui admettent sur* $[0; 1]$ *une dérivée* f' *telle que* $|f'(x)| \leqslant k$ *pour tout* $x \in X$*, est équicontinu sur* X*.*

Exercice II .2 *Soit* U *un ouvert de* $\mathbf{C}$*. Soit* $\mathcal{H}$ *un ensemble de fonctions holomorphes dans* U *tel que, pour tout disque* D *fermé de rayon* > 0 *et contenu dans* U*, il existe une constante* M_D *telle que pour toute* $f \in \mathcal{H}$ *et tout* $z \in D$ *on ait* $|f(z)| \leqslant M_D$*.*

1° *Montrer que pour tout disque* D *comme ci-dessus, il existe une constante* M'_D *telle que pour toute* $f \in \mathcal{H}$ *on ait* $|f'(z)| \leqslant M'_D$*.*

2° *Montrer que l'ensemble* $\mathcal{H}$ *est équicontinu sur* U*.*

Rappelons quelques notions de base sur les convergences de suites de fonctions. Soit $(f_n)_{n \geqslant 0}$ une suite d'applications de X dans F, et soit g une application de X dans F. On dit que $(f_n)_{n \geqslant 0}$ converge *simplement* (ou *ponctuellement*) vers g dans X si

pour tout $x \in X$ et tout $\varepsilon > 0$ il existe un rang N tel que
pour tout $n \geqslant N$ on a $d_F(g(x), f_n(x)) \leqslant \varepsilon$,

alors qu'on dit que $(f_n)_{n \geqslant 0}$ converge *uniformément* vers g dans X si

pour tout $\varepsilon > 0$ il existe un rang N tel que
pour tout $x \in X$ et pour tout $n \geqslant N$ on a $d_F(g(x), f_n(x)) \leqslant \varepsilon$.

Exercice II .3 *Soit $X = [0;1]$ et $F = \mathbf{R}$. Donner un exemple simple d'une suite $(f_n)_{n\geqslant 0}$ de fonctions continues qui converge simplement sur X vers une fonction discontinue. Donner un exemple simple d'une suite $(f_n)_{n\geqslant 0}$ de fonctions continues qui converge simplement mais pas uniformément sur X vers la fonction nulle.*

Nous allons voir que ces exemples n'auraient pas été possibles sous l'hypothèse supplémentaire que les f_n forment un ensemble équicontinu.

Soient à nouveau X et F des espaces métriques.

Proposition II .4 *Soit $(f_n)_{n\geqslant 0}$ une suite d'applications de X dans F. Soit $a \in X$. On suppose que $\{f_n\}_{n\geqslant 0}$ est équicontinue en a et que $(f_n)_{n\geqslant 0}$ converge simplement vers une fonction g. Alors g est continue en a.*

Preuve. On se donne $\varepsilon > 0$. On écrit l'équicontinuité en a de la famille $\{f_n\}_{n\geqslant 0}$: il existe $\delta > 0$ tel que pour tout $n \geqslant 0$ et tout $x \in B(a,\delta)$ on a $d(f_n(x), f_n(a)) < \varepsilon$. Par inégalité triangulaire, on a :

$$d(g(x), g(a)) \leqslant d(g(x), f_n(x)) + d(f_n(x), f_n(a)) + d(f_n(a), g(a)),$$

pour tout $n \geqslant 0$. Par convergence simple, pour n assez grand on a :

$$d(g(x), f_n(x)) < \varepsilon \quad \text{et} \quad d(f_n(a), g(a)) < \varepsilon.$$

Comme $\varepsilon > 0$ est quelconque, on en déduit la continuité de g en a. □

Dans cette preuve, on ne fait rien d'autre que passer à la limite quand $n \to \infty$ dans $d(f_n(x), f_n(a)) < \varepsilon$; pour $F = \mathbf{R}$ ou $\mathbf{C}$, cela donne $|g(x) - g(a)| \leqslant \varepsilon$ à partir de $|f_n(x) - f_n(a)| < \varepsilon$.

Exercice II .5 *Soit $X = [0;1]$ et $F = \mathbf{R}$. On considère la suite des fonctions $f_n : x \mapsto x^n$. En quels points cette suite de fonctions est-elle équicontinue ?*

Proposition II .6 *On suppose que F est complet et on se donne D une partie dense de X. On se donne aussi une suite d'applications $f_n : X \to F$. On suppose que l'ensemble $\{f_n\}_{n\geqslant 0}$ est équicontinu sur X et que $(f_n)_{n\geqslant 0}$ converge simplement sur D. Alors $(f_n)_{n\geqslant 0}$ converge simplement sur X.*

Preuve. Par complétude de F, il suffit de montrer que pour tout $x \in X$ la suite $(f_n(x))_{n\geqslant 0}$ est de Cauchy. On fixe donc $x \in X$. Soit $\varepsilon > 0$. On écrit l'équicontinuité en x de la famille $\{f_n\}_{n\geqslant 0}$: il existe $\delta > 0$ tel que pour tout $n \geqslant 0$ et tout $y \in B(x,\delta)$ on a $d(f_n(x), f_n(y)) < \varepsilon$. Par densité il existe $d \in D \cap B(x,\delta)$, pour lequel en particulier $d(f_n(x), f_n(d)) < \varepsilon$ pour tout $n \geqslant 0$. Par convergence simple, la suite $(f_n(d))_{n\geqslant 0}$ est convergente, donc de Cauchy, et il existe $n \geqslant 1$ tel que pour tous $m, n \geqslant N$ on ait $d(f_n(d), f_m(d)) < \varepsilon$. Par inégalité triangulaire, pour tous $m, n \geqslant N$ on a :

$$d(f_n(x), f_m(x)) \leqslant d(f_n(x), f_n(d)) + d(f_n(d), f_m(d)) + d(f_m(d), f_m(x)),$$

ce qui prouve bien que $(f_n(x))_{n\geqslant 0}$ est de Cauchy. □

Proposition II .7 *On suppose X compact. On se donne $(f_n)_{n\geqslant 0}$ une suite d'applications $X \to F$. On suppose que $\{f_n\}_{n\geqslant 0}$ est équicontinu sur X et que $(f_n)_{n\geqslant 0}$ converge simplement vers une fonction g. Alors la suite $(f_n)_{n\geqslant 0}$ converge vers g uniformément sur X.*

Preuve. Déjà, la proposition II .4 dit que g est continue sur X. Soit $\varepsilon > 0$. Pour tout $x \in X$, par équicontinuité en x de la famille $\{f_n\}_{n\geqslant 0}$ et continuité de g en x, il existe $\delta_x > 0$ tel que pour tout $n \geqslant 0$ et tout $y \in B(x, \delta_x)$ on ait :

$$d(f_n(x), f_n(y)) \leqslant \varepsilon \quad \text{et} \quad d(g(x), g(y)) \leqslant \varepsilon.$$

Par compacité de X, on extrait du recouvrement $X = \bigcup_{x\in X} B(x, \delta_x)$ un sous-recouvrement fini, disons $X = \bigcup_{i=1}^{k} B(x_i, \delta_i)$ avec $\delta_i = \delta_{x_i}$. Par convergence simple, il existe $n \geqslant 0$ tel que pour tout $i \in \{1, 2, \dots k\}$ et tout $n \geqslant N$, on ait : $d(g(x_i), f_n(x_i)) \leqslant \varepsilon$. Soit maintenant $x \in X$ quelconque : ce point est contenu dans une boule $B(x_i, \delta_i)$ et pour tout $n \geqslant N$ on peut écrire :

$$d(g(x), f_n(x)) \leqslant d(g(x), g(x_i)) + d(g(x_i), f_n(x_i)) + d(f_n(x_i), f_n(x)) \leqslant 3\varepsilon.$$

Bref : $\|g - f_n\|_\infty \leqslant 3\varepsilon$ pour $n \geqslant N$. □

On peut synthétiser les propositions qui précèdent dans le résultat suivant.

Théorème II .8 *Soit X un espace métrique compact et soit F un espace métrique complet. Soit D une partie dense de X. On suppose que $\{f_n\}_{n\geqslant 0}$ est équicontinu sur X et que $(f_n)_{n\geqslant 0}$ converge simplement sur D. Alors la suite $(f_n)_{n\geqslant 0}$ converge uniformément sur X.* □

Exercice II .9 *Soit $\mathcal{H}$ l'ensemble des fonctions polynomiales P à coefficients réels dont le maximum des valeurs absolues des coefficients est $\leqslant 1$.*

1° *On se donne un nombre réel $b \in]0; 1[$. Soient x_0 et x réels tels que $|x| \leqslant b$ et $|x_0| \leqslant b$. Soit N un nombre entier > 0. Montrer que*

$$|P(x) - P(x_0)| \quad \leqslant \quad \sum_{n=0}^{N} |x^n - x_0^n| + 2 \sum_{n=?+1}^{\infty} b^n.$$

2° *En déduire que $\mathcal{H}$ est équicontinu sur $]-1; 1[$. Est-ce vrai sur $[-1; 1]$?*

II.9 Théorème d'Ascoli

Dans un espace métrique, toute partie compacte est fermée et bornée. Réciproquement, dans $\mathbf{R}^n$ il suffit qu'un ensemble soit fermé et borné pour qu'il soit compact : c'est le théorème de Borel-Lebesgue. En revanche, dans un espace normé de dimension infinie, on sait que la boule unité fermée, quoiqu'évidemment bornée, n'est jamais compacte : c'est le théorème de F. Riesz ; dans de tels espaces, il faut donc donner des conditions supplémentaires pour exprimer la compacité.

Soit X un espace compact et soit F un espace métrique complet. On notera $\mathcal{C}(X,F)$ l'espace de toutes les applications continues de X dans F, muni de la distance d_∞ de la convergence uniforme :

$$d_\infty(f,g) = \sup_{x\in X} d\left(f(x), g(x)\right).$$

On sait que l'espace métrique $(\mathcal{C}(X,F), d_\infty)$ est complet. En particulier, si $F = \mathbf{R}$ ou $\mathbf{C}$, l'espace $\mathcal{C}(X,F)$ est un espace vectoriel ; c'est un espace de Banach pour la norme :

$$\|f\|_\infty = \sup_{x\in X} |f(x)|,$$

et d_∞ est la distance associée à cette norme.

Nous allons caractériser les parties compactes de l'espace métrique $(\mathcal{C}(X,F), d_\infty)$. Il est remarquable que l'équicontinuité soit la condition essentielle dans cette caractérisation.

Théorème II .10 (Arzela-Ascoli) *Soit $\mathcal{H}$ une partie de $\mathcal{C}(X,F)$, muni de la distance d_∞ de la convergence uniforme. Si $x \in X$ désignons par $\mathcal{H}(x)$ la partie de F formée des valeurs $f(x)$ quand f parcourt $\mathcal{H}$. Pour que $\mathcal{H}$ soit une partie compacte de $(\mathcal{C}(X,F), d_\infty)$, il faut et il suffit que les trois conditions suivantes soient simultanément satisfaites :*
(i) *la partie $\mathcal{H}$ est fermée dans $\mathcal{C}(X,F)$;*
(ii) *quel que soit $x \in X$ l'adhérence de $\mathcal{H}(x)$ dans F est compacte ;*
(iii) *la partie $\mathcal{H}$ est un ensemble équicontinu d'applications $X \to F$.*

Remarque II .11 *Si $F = \mathbf{R}$ ou $\mathbf{C}$, la condition* (ii) *signifie simplement que, quel que soit $x \in X$, l'ensemble $\mathcal{H}(x)$ est borné dans $\mathbf{R}$ ou $\mathbf{C}$, ce qui, en présence de l'hypothèse* (iii)*, équivaut à l'hypothèse* (ii′)*, en apparence plus forte que* (ii)*, suivante :*

(ii′) *la partie $\mathcal{H}$ est bornée dans $\mathcal{C}(X,F)$, i.e. il existe une constante M telle que, quels que soient $f \in \mathcal{H}$ et $x \in X$, on ait : $|f(x)| \leqslant M$.*

Exercice II .12 *Si $F = \mathbf{R}$ ou $\mathbf{C}$, justifier que* « (ii) *et* (iii) » *équivaut à* « (ii′) *et* (iii) ».

Preuve. Prouvons d'abord la partie banale du théorème, à savoir que les conditions mentionnées sont nécessaires. Soit $\mathcal{H}$ une partie compacte de $(\mathcal{C}(X,F), d_\infty)$. En particulier, elle est fermée, d'où (i). Puis, à $x \in X$ fixé, on utilise que l'application d'évaluation $f \mapsto f(x)$ est continue de $\mathcal{C}(X,F)$ dans F, car on a : $d\,(f(x), g(x)) \leqslant d_\infty(f,g)$, et donc l'image $\mathcal{H}(x)$ du compact $\mathcal{H}$ par cette application est compacte elle aussi. On vient de prouver (ii). Enfin, prouvons (iii). Soit $\varepsilon > 0$. D'après la condition du ε-recouvrement pour le compact $\mathcal{H}$, il existe $f_1, f_2, \ldots, f_p \in \mathcal{H}$, en nombre fini, telles que pour toute $f \in \mathcal{H}$ on ait $d_\infty(f, f_i) \leqslant \frac{\varepsilon}{3}$ pour un indice $i \in \{1; 2; \ldots; p\}$ convenable. Soit $x_0 \in X$. Par équicontinuité de $\{f_1; f_2; \ldots; f_p\}$ en ce point, il existe $\eta > 0$ tel que $d_X(x, x_0) < \eta$ entraîne que $d\,(f_i(x), f_i(x_0)) \leqslant \frac{\varepsilon}{3}$ quel que soit l'indice $i \in \{1; 2; \ldots; p\}$. Dès lors, l'inégalité $d_X(x, x_0) < \eta$ entraîne que quelle soit $f \in \mathcal{H}$ on ait :

$$d\,(f(x), f(x_0)) \leqslant d\,(f(x), f_i(x)) + d\,(f_i(x), f_i(x_0)) + d\,(f_i(x_0), f(x_0)) \leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

où l'indice i est l'indice choisi comme ci-dessus par rapport à f. On a donc prouvé (iii) et le finalement le fait que les conditions mentionnées sont nécessaires à la compacité.

On va prouver maintenant que les conditions en question sont suffisantes pour la compacité d'une partie $\mathcal{H}$ de $(\mathcal{C}(X,F), d_\infty)$, en commençant par le cas particulier où $X = [0;1]$ et $F = \mathbf{C}$. C'est une démonstration instructive car elle utilise le procédé diagonal, et donc toutes les ressources de la dénombrabilité.

Soit donc pour le moment $X = [0;1]$ et $F = \mathbf{C}$; soit $D = \mathbf{Q} \cap [0;1]$ l'ensemble des nombres rationnels r tels que $0 \leqslant r \leqslant 1$: c'est un ensemble dénombrable et dense dans $[0;1]$. Ordonnons-le : $D = \{r_1; r_2; \ldots r_n; \ldots\}$ en une suite (peu importe l'ordre). Soit $\mathcal{H}$ une partie de $\mathcal{C} = \mathcal{C}([0;1], \mathbf{C})$ satisfaisant les conditions (i), (ii) et (iii). Puisque déjà $\mathcal{H}$ est fermée dans $\mathcal{C}$, pour montrer qu'elle est compacte il suffit de prouver la propriété de Bolzano-Weierstrass, à savoir que si $(f_n)_{n\geqslant 0}$ est une suite dans $\mathcal{H}$ on peut en extraire une sous-suite qui converge uniformément sur $[0;1]$.

Or, d'après l'hypothèse (ii), la suite de nombres complexes $(f_n(r_1))_{n\geqslant 0}$ est bornée ; donc d'après le théorème de Bolzano-Weierstrass dans $\mathbf{C}$, il existe une suite $(f_n^1)_{n\geqslant 0}$ extraite de $(f_n)_{n\geqslant 0}$ telle que la suite numérique $\left(f_n^1(r_1)\right)_{n\geqslant 0}$ converge dans $\mathbf{C}$. Puis, d'après l'hypothèse (ii) appliquée au point $x = r_2$, il existe une suite $(f_n^2)_{n\geqslant 0}$ extraite de $(f_n^1)_{n\geqslant 0}$ telle que la suite numérique $\left(f_n^2(r_2)\right)_{n\geqslant 0}$ converge dans $\mathbf{C}$ (alors que la suite $\left(f_n^2(r_1)\right)_{n\geqslant 0}$ reste convergente). On itère ce raisonnement, et l'on prouve ainsi que pour tout entier $p > 1$ il existe une suite $(f_n^p)_{n\geqslant 0}$ extraite de $(f_n^{p-1})_{n\geqslant 0}$, donc de $(f_n)_{n\geqslant 0}$, telle que les suites numériques $(f_n^p(r_1))_{n\geqslant 0}, (f_n^p(r_2))_{n\geqslant 0}, \ldots (f_n^p(r_n))_{n\geqslant 0}$ convergent.

La suite « diagonale » $(g_n)_{n\geqslant 0} = (f_n^n)_{n\geqslant 0}$ est extraite de $(f_n)_{n\geqslant 0}$ et converge simplement sur D car, quel que soit p, dès que $n > p$ la suite $(f_n^n(r_p))_{n\geqslant 0}$ est extraite de la suite convergente $(f_n^p(r_p))_{n\geqslant 0}$. Or l'ensemble des g_n est contenu dans $\mathcal{H}$, donc équicontinu dans X. Le théorème de la section précédente assure donc que la suite $(g_n)_{n\geqslant 0}$ converge uniformément sur X, donc au sens de d_∞ dans $\mathcal{C}$.

Cette démonstration par le procédé diagonal vaut plus généralement chaque fois que l'espace compact X est *séparable*, i.e. contient une partie dénombrable et dense, donc pour la plupart des espaces compacts usuels en Analyse, et ceci même pour F métrique complet quelconque au lieu de $\mathbf{C}$.

Remarque II .13 *D'autres utilisations bien connues du procédé diagonal sont : la démonstration par Cantor de la non dénombrabilité de* $\mathbf{R}$*, le fait que toute limite en norme d'opérateurs compacts est encore un opérateur compact, et le théorème de Montel que nous verrons à la section suivante.*

Par scrupule donnons maintenant la preuve du théorème d'Ascoli, valable dans le cas général (et même si X est un espace topologique compact non nécessairement métrique). Soit $\mathcal{H}$ une partie de $\mathcal{C} = \mathcal{C}(X, F)$ satisfaisant les conditions (i), (ii) et (iii). D'après (i), l'espace $\mathcal{H}$ est complet donc il suffit, pour vérifier sa compacité, de prouver sa précompacité, i.e. de vérifier la condition d'ε-recouvrement fini pour $\varepsilon > 0$ arbitrairement petit. Soit donc $\varepsilon > 0$. D'après (iii) – condition d'équicontinuité, pour tout $y \in X$ il existe un ouvert U_y de X, contenant y et tel que pour toute $f \in \mathcal{H}$ on a : $d\,(f(x), f(y)) \leqslant \frac{\varepsilon}{4}$. Par compacité de X, il existe $y_1, y_2, \dots y_p$ en nombre fini dans X tel que X soit recouvert par les ouverts U_{y_i}. D'après (ii) – condition de compacité des ensembles de valeurs, la réunion $K = \overline{\bigcup_{i=1}^p \mathcal{H}(y_i)}$ est une partie compacte de F, donc il existe des $c_1, c_2, \dots c_m \in K$ tels que les boules fermées $\overline{B}(c_j, \frac{\varepsilon}{4})$ $(1 \leqslant j \leqslant m)$ recouvrent K. Soit Φ l'ensemble des applications $i \mapsto \varphi(i)$ de l'ensemble fini $\{1; 2; \dots p\}$ dans l'ensemble fini $\{1; 2; \dots m\}$: Φ est fini. Pour toute $\varphi \in \Phi$, notons $\mathcal{H}_\varphi$ l'ensemble des $f \in \mathcal{H}$ telles que pour tout $i \in \{1; 2; \dots p\}$ on a : $d\left(f(y_i), c_{\varphi(i)}\right) \leqslant \frac{\varepsilon}{4}$. Il est évident qu'on a : $\mathcal{H} = \bigcup_{\varphi\in\Phi} \mathcal{H}_\varphi$. Pour achever la démonstration, nous allons voir que si f_0 est choisie dans $\mathcal{H}_\varphi$ et si f est quelconque dans $\mathcal{H}_\varphi$, on a : $d_\infty(f, f_0) \leqslant \varepsilon$. En effet, soit $x \in X$, et soit i l'un des indices dans $\{1; 2; \dots p\}$ tels que $x \in U_{y_i}$. On a :

$$
\begin{aligned}
d\,(f_0(x), f(x)) &\leqslant d\,(f_0(x), f_0(y_i)) + d\left(f_0(y_i), c_{\varphi(i)}\right) + d\left(c_{\varphi(i)}, f(y_i)\right) + d\,(f(y_i), f(x)) \\
&\leqslant \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon,
\end{aligned}
$$

par définition des ouverts U_y et des ensembles de fonctions $\mathcal{H}_\varphi$. □

Exercice II .14 *L'adhérence dans* $\mathcal{C}(X, F)$ *d'une partie équicontinue est une partie équicontinue. Donc, dans l'énoncé du théorème d'Ascoli* « (ii) &

(iii) » *est une condition nécessaire et suffisante pour que l'adhérence de* $\mathcal{H}$ *soit compacte.*

Exercice II .15 *Soit* $\mathcal{C} = \mathcal{C}([0;1], \mathbf{C})$. *On se donne une fonction continue* $K(\cdot,\cdot)$ *sur le carré* $[0;1] \times [0;1]$ *de* $\mathbf{R}^2$ *à valeurs complexes. À toute* $f \in \mathcal{C}$ *on fait correspondre la fonction* Tf *définie sur* $[0;1]$ *par :*

$$Tf(x) = \int_0^1 K(x,y) f(y)\, \mathrm{d}y.$$

L'application $f \mapsto Tf$ *s'appelle l'opérateur intégral de noyau* K. *Soit* $\mathcal{H}$ *l'image par* T *de la boule unité de* $\mathcal{C}$. *Montrer que l'adhérence de* $\mathcal{H}$ *dans* $\mathcal{C}$ *est compacte. Cet exercice annonce la théorie de Fredholm, qui sera développée plus loin.*

Exercice II .16 *Si* k *est un nombre réel* > 0, *on note* $\mathcal{H}_k$ *l'ensemble des fonctions définies sur* $[0;1]$, *à valeurs réelles, telles que* $f(0) = 0$ *et satisfaisant*

$$|f(x) - f(y)| \quad \leqslant \quad k|x-y| \qquad \textit{pour tous } x, y \in [0;1].$$

On pose $\mathcal{E} = \bigcup_{k>0} \mathcal{H}_k$ *et pour* $f \in \mathcal{E}$, *on note* $N(f) = \inf\{k > 0 : f \in \mathcal{H}_k\}$.

1° *Justifier que* $\mathcal{E}$ *est un* $\mathbf{R}$*-espace vectoriel et que* N *est une norme sur* $\mathcal{E}$.

2° *Prouver que l'on a* $\|f\|_\infty \leqslant N(f)$ *pour toute* $f \in \mathcal{E}$ *mais que* $\|\cdot\|_\infty$ *et* N *ne sont pas des normes équivalentes sur* $\mathcal{H}_k$.

3° *On fixe* $k > 0$. *Discuter la compacité de* $\mathcal{H}_k$ *dans* $(\mathcal{E}, \|\cdot\|_\infty)$ *et dans* $(\mathcal{E}, N(\cdot))$.

Exercice II .17 *Soit* $\mathcal{B} = \mathcal{B}(\mathbf{R}, \mathbf{C})$ *l'espace vectoriel des fonctions* f *bornées dans* $\mathbf{R}$, *à valeurs complexes, avec la norme* $\|f\|_\infty = \sup_{x\in\mathbf{R}} |f(x)|$. *Soit* E *le sous-espace vectoriel fermé de* $\mathcal{B}$ *engendré par les fonctions exponentielles* $x \mapsto \exp(i\lambda x)$, *où* $\lambda \in \mathbf{R}$. *Les fonctions* $f \in E$ *sont appelées fonctions presque périodiques sur* $\mathbf{R}$. *Si* $f \in E$ *et* $\tau \in \mathbf{R}$, *on note* f_τ *la fonction telle que* $f_\tau(x) = f(x-\tau)$ *pour tout* $x \in \mathbf{R}$. *Montrer que les* f_τ *sont dans* E *et que, si* $f \in E$ *est donnée, l'ensemble* $\mathcal{H}_f$ *des* f_τ, *où* τ *parcourt* $\mathbf{R}$, *est d'adhérence compacte dans* E. *Indication : on commencera par traiter le cas où* f *est une exponentielle* $x \mapsto \exp(i\lambda x)$, *puis celui où* f *est un combinaison linéaire* g *de telles exponentielles ; dans le cas général, on prouvera la précompacité pour* $\mathcal{H}_f$ *à partir de celle de* $\mathcal{H}_g$ *pour* g *approchant convenablement* f.

II.10 Analyse dans l'espace des fonctions holomorphes

Soit U un ensemble ouvert non vide fixé du plan complexe. Notons $\mathcal{O}(U)$ l'espace vectoriel sur $\mathbf{C}$ des fonctions $f = f(z)$ définies et holomorphes dans U. On ne fait pas de $\mathcal{O}(U)$ un espace normé, mais pour y pratiquer néanmoins l'Analyse, on utilise la *convergence uniforme sur tout compact* de U ; une suite $(f_n)_{n\geqslant 0}$ dans $\mathcal{O}(U)$ converge uniformément sur tout compact de U vers une fonction g si :

> Pour tout compact X dans U et pour tout $\varepsilon > 0$, il existe $N = N_{\varepsilon,X}$ tel que pour tout $n \geqslant N$ et pour tout $z \in X$ on a :
>
> $$|g(z) - f_n(z)| \leqslant \varepsilon.$$

Nous dirons plus brièvement que la suite $(f_n)_{n\geqslant 0}$ converge vers g dans $\mathcal{O}(U)$.

Remarque II .18 1. *Si c'est le cas, la fonction limite est automatiquement holomorphe dans U, donc appartient à $\mathcal{O}(U)$. En effet, soit $a \in U$; montrons que g est dérivable au sens complexe au voisinage de a. Soit $r > 0$ assez petit pour que le disque $D = \{z \in \mathbf{C} : |z-a| \leqslant r\}$ soit contenu dans U. Pour tout $z \in D^\circ$, on a la formule intégrale de Cauchy :*

$$f_n(z) = \frac{1}{2i\pi}\int_{|z-a|=r}\frac{f_n(\xi)}{\xi - z}\,\mathrm{d}\xi,$$

qui évidemment converge quand $n \to \infty$ vers

$$g(z) = \frac{1}{2i\pi}\int_{|z-a|=r}\frac{g(\xi)}{\xi - z}\,\mathrm{d}\xi,$$

et cette fonction est bien dérivable au sens complexe par rapport à z, sous le signe somme dans le membre de gauche : pour $|z-a| < r$, on a :

$$g'(z) = \frac{1}{2i\pi}\int_{|z-a|=r}\frac{g(\xi)}{(\xi - z)^2}\,\mathrm{d}\xi.$$

2. *Pour vérifier la convergence uniforme sur tout compact, on peut en fait se contenter de la tester sur une famille particulière (mais bien choisie) de compacts dans X, par exemple sur la famille des disques fermés. Supposons qu'une suite $(f_n)_{n\geqslant 0}$ dans $\mathcal{O}(U)$ satisfasse la condition :*

> *Pour tout disque fermé D contenu dans U et pour tout $\varepsilon > 0$, il existe $N = N_{\varepsilon,D}$ tel que pour tout $n \geqslant N$ et pour tout $z \in D$ on a :*
>
> $$|g(z) - f_n(z)| \leqslant \varepsilon.$$

Il s'agit de voir que $(f_n)_{n\geqslant 0}$ converge uniformément sur tout compact de U vers g. Soit X un compact contenu dans U. Tout point

$a \in X$ est le centre d'un disque fermé D_A de rayon > 0 et contenu dans U. Puisque X est compact, il existe un ensemble fini de points $a_1, a_2, \ldots, a_p$ de X tels que

$$X \subset D_{a_1} \cup D_{a_2} \cup \cdots \cup D_{a_p}.$$

Soit $\varepsilon > 0$, et soient $N_{\varepsilon,D_{a_1}}, N_{\varepsilon,D_{a_2}}, \ldots, N_{\varepsilon,D_{a_p}}$ les entiers donnés par la condition de convergence uniforme sur les disques citée ci-dessus. Posons

$$N_{\varepsilon,X} = \max_{1 \leqslant i \leqslant p} N_{\varepsilon,D_{a_i}}.$$

Si $z \in X$, on a $z \in D_{a_i}$ pour au moins un indice i ; donc, puisque $N_{\varepsilon,X} \geqslant N_{\varepsilon,D_{a_i}}$:

$$n \geqslant N_{\varepsilon,X} \Longrightarrow |g(z) - f_n(z)| \leqslant \varepsilon,$$

ce qui prouve la convergence de $(f_n)_{n \geqslant 0}$ vers g dans $\mathcal{O}(U)$.

Exercice II .19 *Il s'agit d'un rappel de théorie élémentaire des séries entières. Soit $0 < R \leqslant +\infty$, et soit $U = \{z \in \mathbf{C} : |z| < R\}$. Soit $f \in \mathcal{O}(U)$. On a : $f(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n + \ldots$ où la série converge pour tout $z \in U$. Montrer que la suite des sommes partielles f_n définie par $f_n(z) = \sum_{k=0}^{n} a_k z^k$ converge vers $f(z)$ dans $\mathcal{O}(U)$.*

Remarque II .20 *3. Soit $\mathcal{H}$ une partie de $\mathcal{O}(U)$. On dit que $\mathcal{H}$ est bornée dans $\mathcal{O}(U)$ si elle est bornée sur tout compact de U, c'est-à-dire si :*

> *Pour tout compact X de U il existe une constante $M_X > 0$ telle que pour tout $f \in \mathcal{H}$ et pour tout $z \in D$ on ait $|f(z)| \leqslant M_X$.*

Cette condition est équivalente à la suivante :

> *Pour tout disque fermé D de U il existe une constante $M_D > 0$ telle que pour tout $f \in \mathcal{H}$ et pour tout $z \in D$ on ait $|f(z)| \leqslant M_D$.*

En effet, sous la seconde condition, en s'étant donné un compact X on peut trouver un recouvrement comme précédemment

$$X \subset D_{a_1} \cup D_{a_2} \cup \cdots \cup D_{a_p},$$

et il suffit alors de poser

$$M_X = \max_{1 \leqslant i \leqslant p} M_{\varepsilon,D_{a_i}}$$

pour vérifier la condition pour X.

Dans les deux propositions ci-après nous allons voir que, du point de vue de l'Analyse dans l'espace $\mathcal{O}(U)$, la dérivation est une opération très régulière.

Proposition II .21 *Soit $(f_n)_{n\geqslant 0}$ une suite qui converge vers g dans $\mathcal{O}(U)$. Alors la suite des dérivées $(f'_n)_{n\geqslant 0}$ converge vers la dérivée g' dans $\mathcal{O}(U)$.*

Preuve. On va prouver la condition de convergence uniforme sur les disques pour les dérivées. Soit $D_a = \{z \in \mathbf{C} : |z-a| \leqslant r\}$ un disque fermé contenu dans U. Il existe $r_1 > r$ tel que le disque $D_a^1 = \{z \in \mathbf{C} : |z-a| \leqslant r_1\}$ soit encore contenu dans U. Soit $\varepsilon > 0$. Par l'hypothèse que $(f_n)_{n\geqslant 0}$ converge vers g dans $\mathcal{O}(U)$, il existe N_ε tel que pour tout $z \in D_a^1$ on ait :

$$|g(z) - f_n(z)| \leqslant \frac{(r_1 - r)^2}{r_1}\varepsilon.$$

Mais alors, pour tout $z \in D_a$, d'après la formule intégrale de Cauchy pour la dérivée, on a :

$$|g'(z)-f'_n(z)| = |\frac{1}{2i\pi}\int_{|\xi-a|=r_1} \frac{g(\xi)-f_n(\xi)}{(\xi-z)^2}\,\mathrm{d}\xi| \leqslant \frac{2\pi r_1}{2\pi}\frac{1}{(r_1-r)^2}\sup_{|\xi-a|=r_1}|g(\xi)-f_n(\xi)|$$

qui est $\leqslant \varepsilon$ dès que $n \geqslant N_\varepsilon$. Ceci prouve la proposition, par la réduction initiale. □

Exercice II .22 *Constater combien cette situation diffère du cas de la variable réelle, en exhibant une suite $(f_n)_{n\geqslant 0}$ de fonctions dérivables sur $[0;1]$ qui converge uniformément vers une fonction g dérivable, mais telle que les suites numériques $(f'_n(x))_{n\geqslant 0}$ n'aient de limite pour aucun $x \in [0;1]$.*

Proposition II .23 *Soit $\mathcal{H}$ une partie de $\mathcal{O}(U)$ et soit $\mathcal{H}' = \{f' : f \in \mathcal{H}\}$. Si $\mathcal{H}$ est bornée dans $\mathcal{O}(U)$, alors $\mathcal{H}'$ l'est aussi.*

Preuve. On utilise là aussi la réduction à la famille des disques fermés. Soit $D_a = \{z \in \mathbf{C} : |z-a| \leqslant r\}$ un disque fermé contenu dans U. On cherche une constante M'_{D_a} telle que pour toute $f \in \mathcal{H}$ et tout $z \in D_a$ on ait : $|f'(z)| \leqslant M'_{D_a}$. Il existe $r_1 > r$ tel que le disque $D_a^1 = \{z \in \mathbf{C} : |z-a| \leqslant r_1\}$ soit encore contenu dans U, et pour tout $z \in D_a$ on a :

$$|f'(z)| = |\frac{1}{2i\pi}\int_{|\xi-a|=r_1}\frac{f(\xi)}{(\xi-z)^2}\,\mathrm{d}\xi| \leqslant \frac{2\pi r_1}{2\pi}\frac{1}{(r_1-r)^2}M_{D_a^1},$$

où $M_{D_a^1}$ est une constante provenant de la condition que $\mathcal{H}$ est bornée dans $\mathcal{O}(U)$. Ceci prouve la proposition, par la réduction initiale. □

Corollaire II .24 *Toute partie bornée dans $\mathcal{O}(U)$ est équicontinue.*

Preuve. Soit $\mathcal{H}$ une partie bornée dans $\mathcal{O}(U)$. Soit $a \in U$. Pour démontrer l'équicontinuité de $\mathcal{H}$ en a, on se place dans un disque D_a centré en a, de rayon > 0, contenu dans U. D'après la proposition II .23, il existe une constante M'_{D_a} telle que pour toute $f \in \mathcal{H}$ et pour tout $z \in D_a$ on a : $|f'(z)| \leqslant M'_{D_a}$. En intégrant le long du segment $[a; z]$, on obtient que, pour toute $f \in \mathcal{H}$ et pour tout $z \in D_a$, on a :

$$|f(z) - f(a)| = |\int_a^z f'(\xi)\,\mathrm{d}\xi \leqslant |z-a| \sup_{\xi \in [a;z]} |f'(z)| \leqslant M'_{D_a}|z-a| \leqslant \varepsilon$$

dès que $|z - a| \leqslant \dfrac{\varepsilon}{M'_{D_a}}$. □

Corollaire II .25 *Soit $(f_n)_{n \geqslant 0}$ une suite dans $\mathcal{O}(U)$. Supposons que l'ensemble des f_n est borné dans $\mathcal{O}(U)$, ce qui signifie que*

$$\sup_{n \geqslant 0} \sup_{z \in X} |f_n(z)| = M_X < +\infty$$

pour tout compact X contenu dans U. Supposons de plus que la suite $(f_n)_{n \geqslant 0}$ converge simplement dans une partie dense de U. Alors la suite $(f_n)_{n \geqslant 0}$ converge uniformément sur tout compact de U, c'est-à-dire dans $\mathcal{O}(U)$.

Preuve. Vu le corollaire II .24, c'est une conséquence immédiate des propositions II .6 et II .7 appliquées aux parties compactes X de U. □

Théorème II .26 (Montel) *Soit $(f_n)_{n \geqslant 0}$ une suite dans $\mathcal{O}(U)$. Supposons que l'ensemble des f_n est borné dans $\mathcal{O}(U)$. Alors on peut extraire de la suite $(f_n)_{n \geqslant 0}$ une suite convergente dans $\mathcal{O}(U)$.*

Preuve. Les points $r = x + iy$ de U avec $x, y \in \mathbf{Q}$ forment un ensemble $D = \{r_1; r_2; \ldots r_n; \ldots\}$ dénombrable et dense dans U. Pour tout $r \in D$ l'ensemble de nombres complexes $\{f_n(r)\}_{n \geqslant 0}$ est borné car $\{r\}$ est compact. Le procédé diagonal utilisé dans la preuve du théorème d'Ascoli fournit une suite $(g_n)_{n \geqslant 0}$ extraite de la suite $(f_n)_{n \geqslant 0}$, qui converge simplement en tout point de D, donc dans $\mathcal{O}(U)$ d'après le corollaire II .25. □

Cet énoncé est remarquable et même étonnant : l'espace $\mathcal{O}(U)$, de même que l'espace de dimension finie $\mathbf{R}^n$, jouit de la propriété de Bolzano-Weierstrass, à savoir que de toute suite bornée on peut extraire une suite convergente, alors que ceci n'est jamais vrai dans un espace vectoriel normé de dimension infinie (théorème de F. Riesz).

Voici un complément au théorème de Montel.

Théorème II .27 (Vitali) *Soit U un ouvert connexe de $\mathbf{C}$ et soit D une partie de U ayant au moins un point d'accumulation dans U, i.e. D contient* *au moins une suite injective* $(r_p)_{p\geqslant 0}$ *qui converge vers un point de* U*. Soit* $(f_n)_{n\geqslant 0}$ *une suite dans* $\mathcal{O}(U)$*. Supposons que l'ensemble des* f_n *soit borné dans* $\mathcal{O}(U)$ *et que la suite* $(f_n)_{n\geqslant 0}$ *converge simplement dans* D*. Alors la suite* $(f_n)_{n\geqslant 0}$ *converge uniformément sur tout compact de* U*, c'est-à-dire dans* $\mathcal{O}(U)$*.*

Preuve. Remarque préliminaire : puisque U est connexe, si g_1 et g_2 sont holomorphes sur U et si $g_1(r_p) = g_2(r_p)$ pour tout $p \geqslant 0$, alors $g_1 = g_2$ par le principe des zéros isolés appliqué à $g_1 - g_2$.

D'après le théorème II .26, il existe des suites extraites de $(f_n)_{n\geqslant 0}$ qui convergent dans $\mathcal{O}(U)$. Par la remarque préliminaire et l'hyptohèse que la suite $(f_n)_{n\geqslant 0}$ converge simplement dans D, donc en chaque r_p, pour toute telle sous-suite convergente, la limite $g(z)$ est toujours la même (car elle est déterminée par ses valeurs en les r_p). Par l'absurde, supposons que $(f_n)_{n\geqslant 0}$ ne converge pas vers g dans $\mathcal{O}(U)$. Il existe alors un compact X de U et un $\alpha > 0$ tels que, pour tout entier $k \geqslant 1$, il existe un entier n_k et un $z_k \in X$, avec $n_k > n_{k-1}$ et $|g(z_k) - f_{n_k}(z_k)| > \alpha$. De la suite bornée $(f_{n_k})_{k\geqslant 0}$ on ne peut extraire aucune sous-suite convergente (ce qui est en contradiction avec le théorème II .26), car cette dernière suite devrait alors converger vers g tout en satisfaisant les inégalités $|g(z_k) - f_{n_k}(z_k)| > \alpha$. □

Exercice II .28 *On pose, pour tout* $z \in \mathbf{C}$ *:*

$$\exp(z) = 1 + z + \frac{z}{2} + \cdots + \frac{z^n}{n!} + \ldots$$

Montrer que $\exp(z) = \lim_{n\to\infty}\left(1 + \frac{z}{n}\right)^n$ *en procédant comme suit.*

1° *Vérifier que c'est vrai pour tous les* z *réels* > 0 *en passant au logarithme.*

2° *Montrer que la suite* $(f_n)_{n\geqslant 0}$*, où* $f_n(z) = (1 + \frac{z}{n})^n$*, est bornée dans* $\mathcal{O}(\mathbf{C})$*.*

3° *En déduire, par le théorème de Vitali, que* $(f_n)_{n\geqslant 0}$ *converge dans* $\mathcal{O}(\mathbf{C})$*, disons vers* g*.*

4° *Constater que* $g(z) = \exp(z)$ *car c'est vrai pour* z *réel* > 0*.*

Exercice II .29 *Soit* U *le disque* $\{z \in \mathbf{C} : |z - 1| < 1\}$*. Si* $z \in U$ *on pose :*

$$\log(z) = \sum_{m\geqslant 1} -\frac{(1-z)^m}{m},$$

somme d'une série absolument convergente, et

$$f_n(z) = \sum_{m=1}^{n} -\frac{(1-z)^m}{m}.$$

1° *Si z est réel et dans U, constater que* $\log(z)$ *est la fonction bien connue telle que* $\exp(\log(z)) = z$.
2° *Montrer que, quand $n \to \infty$, la suite* $(\exp \circ f_n)_{n \geqslant 1}$ *converge dans* $\mathcal{O}(U)$.
3° *En déduire que, pour tout $z \in U$, on a :* $\exp(\log(z)) = z$.

Exercice II .30 *Soit $\mathcal{C} = \mathcal{C}([-\frac{1}{2}; \frac{1}{2}], \mathbf{R})$, et soit $\mathcal{H}$ le sous-ensemble de $\mathcal{C}$ formé des fonctions polynomiales à coefficients réels tous $\leqslant 1$ en valeur absolue. Soit $\overline{\mathcal{H}}$ l'adhérence de $\mathcal{H}$ dans $\mathcal{C}$.*

1° *Montrer que, pour toute $f \in \mathcal{H}$, la dérivé $\frac{\mathrm{d}f}{\mathrm{d}x}$ est bornée en valeur absolue par 4 sur $[-\frac{1}{2}; \frac{1}{2}]$.*
2° *En déduire que $\mathcal{H}$ est équicontinu dans $[-\frac{1}{2}; \frac{1}{2}]$.*
3° *Montrer que $\overline{\mathcal{H}}$ est un sous-ensemble convexe et compact dans $\mathcal{C}$.*
4° *En appliquant le théorème de Vitali, montrer que toute $f \in \overline{\mathcal{H}}$ est la restriction au segment $[-\frac{1}{2}; \frac{1}{2}]$ d'une fonction holomorphe dans le disque unité ouvert de $\mathbf{C}$, notée $\tilde{f}$.*
5° *Caractériser les fonctions $\tilde{f}$ où $f \in \overline{\mathcal{H}}$, par une propriété de leurs coefficients de Taylor à l'origine. Indication : on pourra utiliser la formule, pour $n \geqslant 0$ et $0 \leqslant r < 1$, donnant ces coefficients a_n, à savoir : $a_n r^n = \dfrac{1}{2\pi} \displaystyle\int_0^{2\pi} \tilde{f}(re^{i\theta}) e^{-in\theta} \,\mathrm{d}\theta$.*
6° *Déterminer l'ensemble des points extrémaux de $\overline{\mathcal{H}}$.*

Exercice II .31 *Soit U la bande $\{z = x + iy : 0 < x < 1; y > 0\}$. Soit f une fonction holomorphe et bornée dans U. Supposons que $l = \lim_{y \to +\infty} f(\frac{1}{2} + iy)$ existe. Montrer qu'alors, pour tout $x \in]0;1[$, on a :*

$$\lim_{y \to +\infty} f(x + iy)$$

et ceci uniformément pour $x \in [a;b]$ quels que soient $a < b$ dans $]0;1[$. Indication : appliquer le théorème de Vitali à la suite $f_n(z) = f(z + in)$ dans le rectangle défini par $0 < x < 1$ et $0 < y < 2$.

II.11 Approximation uniforme des fonctions continues sur un compact

Dans tout ce paragraphe, X désigne un espace compact, et $\mathcal{C} = \mathcal{C}(X, \mathbf{R})$ est l'espace de Banach sur $\mathbf{R}$ des fonctions continues sur X à valeurs réelles, muni de la norme de la convergence uniforme :

$$\|f\|_\infty = \sup_{x\in X} |f(x)|.$$

Outre les opérations vectorielles, on a dans $\mathcal{C}$:

- la multiplication des fonctions : pour $f, g \in \mathcal{C}$, la fonction fg est définie par $(fg)(x) = f(x)g(x)$ pour tout $x \in X$; la compatibilité avec la norme est l'inégalité de *sous-multiplicativité* :

  $$\|fg\|_\infty \leqslant \|f\|_\infty \cdot \|g\|_\infty,$$

  ce qui fait de $\mathcal{C}$ une *algèbre de Banach* dont la théorie générale sera développée plus loin dans ce cours.
- la relation d'ordre $\leqslant$, où $f \leqslant g$ signifie que $f(x) \leqslant g(x)$ pour tout $x \in X$. Si f et g sont dans $\mathcal{C}$, on définit leurs *enveloppe supérieure* $\sup(f, g)$ et *enveloppe inférieure* $\inf(f, g)$ par les formules pour tout $x \in X$:

  $$\sup(f, g)(x) = \max\{f(x); g(x)\} \quad \text{et} \quad \inf(f, g)(x) = \min\{f(x); g(x)\}.$$

Exercice II .32 *Justifier que les fonctions* $\sup(f, g)$ *et* $\inf(f, g)$ *sont dans* $\mathcal{C}$ *puis, en définissant* $|f|$ *par* $|f|(x) = |f(x)|$ *pour tout* $x \in X$*, démontrer les formules :*

$$|f| = \sup(f, -f), \quad \sup(f, g) = \frac{1}{2}\left(f + g + |f - g|\right) \quad \text{et} \quad \inf(f, g) = \frac{1}{2}\left(f + g - |f - g|\right).$$

Montrer que $f \mapsto |f|$ *est une application continue de* $\mathcal{C}$ *dans lui-même.*

Definition II .33 *Soit* $\mathcal{H}$ *une partie de* $\mathcal{C}$ *et soit* $\mathcal{H}$ *son adhérence dans* $\mathcal{C}$*. Soit* $f \in \mathcal{C}$*. On dit que* f *peut être approchée uniformément dans* X *par des fonctions appartenant à* $\mathcal{H}$ *si* f *appartient à* $\mathcal{H}$*, autrement dit si :*

> *quel que soit* $\varepsilon > 0$ *il existe une fonction* $u_\varepsilon \in \mathcal{H}$ *telle que, quel que soit* $x \in X$*, on ait :* $|f(x) - u_\varepsilon(x)| \leqslant \varepsilon$*.*

Dire que toute fonction continue sur X peut être approchée uniformément dans X par des fonctions appartenant à $\mathcal{H}$, c'est donc dire que $\overline{\mathcal{H}} = \mathcal{C}$, ou encore que $\mathcal{H}$ est dense dans l'espace de Banach $\mathcal{C}$. L'exemple historique est le théorème de Weierstrass, dont l'énoncé est que toute fonction continue $[0; 1]$ est limite uniforme sur $[0; 1]$ de fonctions polynomiales. C'est cet énoncé que nous allons retrouver après l'avoir généralisé considérablement, suivant les idées de M. Stone.

Théorème II .34 (Stone-Weierstrass) *Soit X un espace compact. On se donne $\mathcal{H}$ une partie de $\mathcal{C} = \mathcal{C}(X, \mathbf{R})$ telle que :*

(i) *$\mathcal{H}$ est un sous-espace vectoriel de $\mathcal{C}$;*
(ii) *$\mathcal{H}$ est stable par multiplication : si $u, v \in \mathcal{H}$, alors $uv \in \mathcal{H}$;*
(iii) *$\mathcal{H}$ contient la fonction constante égale à 1 ;*
(iv) *$\mathcal{H}$ sépare X : pour tous $x \neq x'$ dans X, il existe $u \in \mathcal{H}$ telle que $u(x) \neq u(y)$.*

Alors $\overline{\mathcal{H}} = \mathcal{C}$.

On retrouve bien le théorème historique en prenant $X = [0; 1]$ et $\mathcal{H}$ l'ensemble des fonctions polynomiales sur $[0; 1]$; le seul polynôme $x \mapsto x$ sert ici à séparer les points de X.

En fait, Stone a su dégager l'importance, dans la question, de la relation d'ordre sur $\mathcal{C}$; le théorème de Stone-Weierstrass sera obtenu comme conséquence du

Théorème II .35 (Stone) *Soit X un espace compact. On se donne $\mathcal{H}$ une partie de $\mathcal{C} = \mathcal{C}(X, \mathbf{R})$ telle que :*

(i) *$\mathcal{H}$ est un sous-espace vectoriel de $\mathcal{C}$;*
(ii′) *$\mathcal{H}$ est stable par prise de valeur absolue : si $u \in \mathcal{H}$, alors $|u| \in \mathcal{H}$;*
(iii) *$\mathcal{H}$ contient la fonction constante égale à 1 ;*
(iv) *$\mathcal{H}$ sépare X.*

Alors $\overline{\mathcal{H}} = \mathcal{C}$.

Nous commençons par établir un lemme dont la démonstration est un bel argument de compacité.

Lemme II .36 *Soit $\mathcal{H}$ une partie de $\mathcal{C}$ telle que si u_1 et u_2 sont dans $\mathcal{H}$, alors $\max(u_1, u_2) \in \mathcal{H}$. Soit $f \in \mathcal{C}(X; \mathbf{R})$. On suppose que pour tout $y \in X$ il existe $u_y \in \mathcal{H}$ telle que $u_y(y) > f(y)$. Alors il existe $v \in \mathcal{H}$ telle que $v(x) > f(x)$ pour tout $x \in X$.*

Preuve. Pour chaque $y \in X$, note $U_y = \{x \in X : u_y(x) > f(x)\}$: c'est un ouvert car c'est l'image réciproque de $\mathbf{R}_+^\times$ par la fonction continue $u_y - f$; en outre $y \in U_y$. Par compacité de X, qui est réunion de tous les U_y, on peut écrire $X = U_{y_1} \cup U_{y_2} \cup \cdots \cup U_{y_p}$ pour des y_i convenables et en nombre fini.

Par récurrence finie, la fonction $v : x \mapsto \max\{u_{y_1}(x); u_{y_2}(x); \ldots u_{y_p}(x)\}$ est dans $\mathcal{H}$. Elle convient car un élément quelconque $x \in X$ est dans un U_{y_i} au moins et on peut écrire :

$$v(x) \geqslant u_{y_i}(x) > f(x). \qquad \square$$

De ce lemme, on déduit déjà le remarquable :

Théorème II .37 (Dini) *Soit X un espace compact. Soit $(f_n)_{n \geqslant 0}$ une suite de fonctions continues de X dans $\mathbf{R}$. On fait les hypothèses suivantes.*

(i) *La suite $(f_n)_{n \geqslant 0}$ converge simplement vers f dans $\mathcal{C}(X, \mathbf{R})$.*
(ii) *La suite $(f_n)_{n \geqslant 0}$ est monotone, c'est-à-dire ou bien décroissante ou bien croissante.*

Alors $(f_n)_{n \geqslant 0}$ converge uniformément vers f, c'est-à-dire converge vers f pour la norme $\| \cdot \|_\infty$.

Preuve. On travaille dans le cas où la suite $(f_n)_{n \geqslant 0}$ est croissante (cas auquel on peut toujours se ramener, quitte à opposer les fonctions). Alors, par monotonie la famille $\mathcal{H} = \{f_n\}_{n \geqslant 0}$ est stable par passage au maximum car pour $m \geqslant n$, on a : $\max\{f_n; f_m\} = f_m$.

Maintenant on se donne $\varepsilon > 0$. Par convergence simple, pour tout $y \in X$ il existe $u \in \mathcal{H}$ telle que $u(y) > f(y) - \varepsilon$. On applique le lemme précédent à la fonction $f - \varepsilon$ (toujours avec $\mathcal{H} = \{f_n\}_{n \geqslant 0}$) : il existe N tel que pour tout $x \in X$ on ait :

$$f_N(x) > f(x) - \varepsilon.$$

Finalement, pour tout indice $n \geqslant N$ et pour tout $x \in X$, on a :

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

soit $\|f - f_n\|_\infty \leqslant \varepsilon$ pour tout $n \geqslant N$. □

Le théorème de Dini a pour conséquence ce premier énoncé, d'apparence anecdotique, mais utile pour prouver le théorème de Weierstrass historique (approximation uniforme des fonctions continues sur un segment compact par des fonctions polynomiales).

Lemme II .38 *On définit une suite $(p_n)_{n \geqslant 0}$ de fonctions $[0; 1] \to \mathbf{R}$ par $p_1(x) = 0$ pour tout $x \in [0; 1]$ et*

$$p_{n+1}(x) = p_n(x) + \frac{1}{2}\left(x - p_n(x)^2\right).$$

Alors les fonctions p_n sont polynomiales et convergent uniformémement vers $x \mapsto \sqrt{x}$ sur $[0; 1]$.

Remarqu'ons qu'en remplaçant $p_n(x)$ par $p_n(x^2)$, on voit que la fonction valeur absolue $x \mapsto |x|$ est limite uniforme de fonctions polynomiales sur $[0; 1]$.

Preuve. Pour tout $n \geqslant 1$, un calcul montre que :

$$\sqrt{x} - p_{n+1}(x) = \left(\sqrt{x} - p_n(x)\right)\left(1 - \frac{1}{2}\left(\sqrt{x} + p_n(x)\right)\right),$$

ce qui permet de voir par récurrence que

$$p_n(x) \geqslant 0 \quad \text{et} \quad p_n(x) \leqslant p_{n+1}(x) \leqslant \sqrt{x} \quad \text{pour tout } x \in [0;1].$$

À x fixé la suite numérique $(p_n(x))_{n\geqslant 0}$ est croissante et majorée. Elle converge donc vers $g(x)$ qui satisfait $g(x) = g(x) + \frac{1}{2}\left(x - g(x)^2\right)$, autrement dit vers $g(x) = \sqrt{x}$. Ainsi on vient de voir que $\sqrt{\cdot}$ est limite simple des p_n sur $[0;1]$, et le théorème de Dini assure que la convergence est uniforme. □

La proposition suivante est au cœur de la démonstration de Stone ; sa démonstration n'est pas vraiment délicate si l'on a compris la preuve du lemme II .36.

Proposition II .39 *Soit $\mathcal{H}$ une partie de $\mathcal{C}$ stable par prise d'enveloppe supérieure et inférieure, i.e. telle que si u_1 et u_2 sont dans $\mathcal{H}$, alors $\sup(u_1, u_2) \in \mathcal{H}$ et $\inf(u_1, u_2) \in \mathcal{H}$. Soit $f \in \mathcal{C}$. On suppose que pour tout $\varepsilon > 0$ et quels que soient $x, y \in X$ il existe une fonction $u_{x,y}$ telle que*

$$|f(x) - u_{x,y}(x)| < \varepsilon \quad \text{et} \quad |f(y) - u_{x,y}(y)| < \varepsilon.$$

Alors $\overline{\mathcal{H}} = \mathcal{C}$.

Preuve. Soit $\varepsilon > 0$. Si $x \in X$, notons $\mathcal{H}_x$ l'ensemble des fonctions $u \in \mathcal{H}$ telles que $u(x) < f(x) + \varepsilon$. Par hypothèse, pour tout $y \in X$, on a :

$$u_{x,y} \in \mathcal{H}_x \quad \text{et} \quad u_{x,y}(y) > f(y) - \varepsilon.$$

De plus $u_1, u_2 \in \mathcal{H}_x$ entraîne que $\sup(u_1, u_2) \in \mathcal{H}_x$. Donc, d'après le lemme II .36, il existe une fonction $u_x \in \mathcal{H}_x$ telle que pour tout $z \in X$ on a :

$$u_x(z) > f(z) - \varepsilon.$$

Puisque $u_x \in \mathcal{H}_x$, on a de plus : $u_x(x) < f(x) + \varepsilon$. Soit $\mathcal{H}'$ l'ensemble des $u \in \mathcal{H}$ telles que l'on ait $u(z) > f(z) - \varepsilon$ pour tout $z \in X$. Si u_1 et u_2 sont dans $\mathcal{H}'$, il est clair que $\inf(u_1, u_2) \in \mathcal{H}'$. Comme les u_x appartiennent à $\mathcal{H}'$ et vérifient $u_x(x) < f(x) + \varepsilon$, d'après le lemme II .36 il existe $u \in \mathcal{H}'$ telle que quel que soit $z \in X$ on a : $u(z) < f(z) + \varepsilon$. Finalement, pour cette fonction u, qui est évidemment dans $\mathcal{H}$, on a : $\|f - u\|_\infty \leqslant \varepsilon$. □

Corollaire II .40 *Soit X un espace compact ayant au moins deux éléments. Soit $\mathcal{H}$ une partie de $\mathcal{C}$ telle que :*

(i) *$\mathcal{H}$ est stable par prise d'enveloppe supérieure et inférieure ;*
(ii) *quels que soient $x \neq y$ dans X et $\alpha, \beta \in \mathbf{R}$, il existe $g \in \mathcal{H}$ telle que $g(x) = \alpha$ et $g(y) = \beta$.*

Alors $\overline{\mathcal{H}} = \mathcal{C}$.

Preuve. Soit $f \in \mathcal{C}$. Montrons que f et $\mathcal{H}$ satisfont aux hypothèse de la proposition précédente. Si $x \neq y$, on y satisfait (pour tout $\varepsilon > 0$) en prenant pour $u_{x,y}$ la fonction g de la condition (ii) ci-dessus pour le choix de $\alpha = f(x)$ et $\beta = f(y)$. Si $x = y$, prenons un $z \neq x$; il suffit alors de choisir pour $u_{x,y}$ la fonction g de la condition (ii) ci-dessus pour le choix de $\alpha = f(x) = f(y)$ et $\beta = f(z)$. ▫

Exercice II .41 *Par ce corollaire, démontrer que dans* $\mathcal{C}([0;1], \mathbf{R})$ *est uniformément dense l'ensemble des fonctions dont le graphe est une ligne polygonales à nombre fini de côtés.*

Maintenant, tout est prêt pour obtenir sans effort la preuve du théorème de Stone.

Preuve. Les hypothèses (i) et (ii′) entraînent que $\mathcal{H}$ est stable par prise d'enveloppe supérieure et inférieure, car :

$$\sup(f,g) = \frac{1}{2}\left(f + g + |f - g|\right) \quad \text{et} \quad \inf(f,g) = \frac{1}{2}\left(f + g - |f - g|\right).$$

Si X est réduit à un point, le théorème de Stone est évident (et sans intérêt). Désormais, on se place dans le cas où X a au moins deux éléments. Si $x \neq y$ sont dans X et si $\alpha, \beta \in \mathbf{R}$, donnons-nous grâce à l'hypothèse (iv) une fonction $h \in \mathcal{H}$ telle que $h(x) \neq h(y)$. On considère alors la fonction g définie par :

$$z \mapsto g(z) = \alpha 1 + (\beta - \alpha)\frac{h(z) - h(x)}{h(y) - h(x)}.$$

Cette fonction g est dans $\mathcal{H}$ par les hypothèses (i) et (iii), et tout a été fait pour que $g(x) = \alpha$ et $g(y) = \beta$. Donc $\overline{\mathcal{H}} = \mathcal{C}$ par le corollaire qui précède. ▫

Exercice II .42 *Soit* (X, d) *un espace métrique compact. Soit* $\mathcal{H}$ *l'ensemble des fonctions lipschitziennes sur* X*, i.e. telles qu'il existe une constante* k *(dépendant de* f*) pour laquelle on ait* $|f(x) - f(y)| \leqslant kd(x,y)$ *quels que soient* $x, y \in X$*. Montrer que* $\mathcal{H}$ *est uniformément dense dans* $\mathcal{C}(X, \mathbf{R})$.

Passons maintenant à la démonstration du théorème de Stone-Weierstrass.

Preuve. On va le déduire du théorème de Stone ; il suffit de voir que les hypothèses (i) à (iv) du théorème de Stone-Weierstrass entraînent l'hypothèse (ii′) du théorème de Stone. En fait, il suffit de voir que ces quatre hypothèse entraînent

$$\text{(ii}''\text{)} \quad u \in \mathcal{H} \Rightarrow |u| \in \overline{\mathcal{H}}.$$

C'est une condition plus faible en apparence que (ii′), mais l'astuce est la suivante : vu la continuité de l'application $u \mapsto |u|$ de $\mathcal{C}$ dans lui-même, on saura alors que $u \in \overline{\mathcal{H}}$ implique $|u| \in \overline{\mathcal{H}}$, et puisque que $\overline{\mathcal{H}}$ est un sous-espace vectoriel, on lui appliquera le théorème de Stone, ce qui donnera $\overline{\overline{\mathcal{H}}} = \mathcal{C}$, soit $\overline{\mathcal{H}} = \mathcal{C}$ puisque $\overline{\overline{\mathcal{H}}} = \overline{\mathcal{H}}$.

Pour vérifier (ii$''$), on se ramène au cas $u \neq 0$. Posons alors $a = \|u\|_\infty$. Soient les polynômes p_n du lemme II .38, qui approchent uniformément la fonction racine carrée sur $[0;1]$. La suite de fonctions $u_n = ap_n(\frac{u^2}{a^2})$ est dans $\mathcal{H}$ d'après les hypothèses (i) à (iii). La suite des fonctions u_n, quand $n \to \infty$, converge uniformément dans X vers la fonction $a\sqrt{\frac{u^2}{a^2}} = |u|$. Donc $u \in \overline{\mathcal{H}}$, ce qui montre que la condition (ii$''$) est satisfaite. □

Relisez l'énoncé du théorème de Stone-Weierstrass et constatez que, plus brièvement, il peut se lire : *toute sous-algèbre de $\mathcal{C}(X, \mathbf{R})$, qui contient* 1 *et sépare X, est dense dans $\mathcal{C}(X, \mathbf{R})$.*

Nous pouvons maintenant dérouler une série de corollaires importants.

Corollaire II .43 (Weierstrass classique) *Soit X une partie fermée et bornée de $\mathbf{R}^n$. Alors toute fonction $f \in \mathcal{C}(X, \mathbf{R})$ est limite uniforme sur X d'une suite de fonctions $(p_k(x_1, x_2, \ldots, x_n))_{n \geqslant 0}$ polynomiales en n variables et à coefficients réels.*

Preuve. En effet, un tel X est compact dans $\mathbf{R}^n$ par Borel-Lebesgue ; l'ensemble $\mathcal{H}$ des fonctions polynomiales satisfait les conditions (i), (ii) et (iii), ainsi que (iv) car si x et y sont distincts dans $\mathbf{R}^n$, l'une au moins de leurs coordonnées diffèrent (les fonctions coordonnées sont bien sûr polynomiales, de degré 1). □

Corollaire II .44 (Stone-Weierstrass complexe) *Soit X un espace compact. On se donne $\mathcal{H}$ une partie de $\mathcal{C} = \mathcal{C}(X, \mathbf{C})$ telle que :*

(i) *$\mathcal{H}$ est un sous-espace vectoriel de $\mathcal{C}$;*
(ii) *$\mathcal{H}$ est stable par produit : si $u, v \in \mathcal{H}$, alors $uv \in \mathcal{H}$;*
(iii) *$\mathcal{H}$ contient la fonction constante égale à* 1 *;*
(iv) *$\mathcal{H}$ sépare X ; et*
(v) *$\mathcal{H}$ est stable par conjugaison complexe : $u \in \mathcal{H}$ entraîne $\bar{u} \in \mathcal{H}$, où $\bar{u}$ est la fonction sur X définie par $\bar{u}(x) = \overline{u(x)}$ pour tout $x \in X$.*

Alors $\overline{\mathcal{H}} = \mathcal{C}$, i.e. toute fonction continue à valeurs complexes sur X peut être approchée uniformément sur X par des fonctions dans $\mathcal{H}$.

Plus brièvement : toute sous-algèbre de $\mathcal{C}(X, \mathbf{C})$, contenant 1, séparante et auto-adjointe, est dense dans l'espace de Banach $\mathcal{C}(X, \mathbf{C})$.

Preuve. Soit $\mathcal{H}_{\mathbf{R}}$ l'ensemble des $u \in \mathcal{H}$ telles que $u(X) \subset \mathbf{R}$. Il est clair que $\mathcal{H}_{\mathbf{R}} \subset \mathcal{C}(X, \mathbf{R})$ satisfait les conditions (i), (ii), (iii) du théorème de Stone-Weierstrass. De plus, $\mathcal{H}_{\mathbf{R}}$ satisfait (iv), c'est-à-dire sépare X : en effet, soient $x, y \in X$ avec $x \neq y$; il existe $u \in \mathcal{H}$ telle que $u(x) \neq u(y)$; les fonctions $R(u) = \frac{u+\bar{u}}{2}$ et $I(u) = \frac{u-\bar{u}}{2i}$ sont dans $\mathcal{H}_{\mathbf{R}}$ et l'une au moins prend des valeurs différentes en x et en y car $u = R(u) + iI(u)$. Ainsi, par le théorème de Stone-Weierstrass toute fonction $f \in \mathcal{C}(X, \mathbf{R})$ est approchée uniformément sur X par des $u \in \mathcal{H}_{\mathbf{R}}$. □

Exercice II .45 *Soit X un sous-ensemble compact de $\mathbf{C}$.*

1° *Montrer que toute fonction $f \in \mathcal{C}(X, \mathbf{C})$ peut être approchée uniformément sur X par des fonctions polynomiales en les variables z et $\bar{z}$, à coefficients complexes.*

2° *Soit X le disque unité $\{z \in \mathbf{C} : |z| \leqslant 1\}$. Soit $\overline{\mathcal{H}}$ l'adhérence dans $\mathcal{C}(X, \mathbf{C})$ de l'ensemble des polynômes à coefficients complexes en la seule variable z. Montrer que pour toute $f \in \overline{\mathcal{H}}$ on a la formule :*

$$f(0) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) \, \mathrm{d}\theta,$$

et que, en revanche, la fonction continue $f(z) = |z|$ ne possède pas cette propriété.

3° *En déduire que $\overline{\mathcal{H}}$ est différent de $\mathcal{C}(X, \mathbf{C})$.*

Exercice II .46 *Soit $a \leqslant b$ des nombres réels. Montrer que l'ensemble des fonctions de la forme $x \mapsto \sum_{k=0}^{n} \lambda_k \exp(n_n x)$ où les λ_k appartiennent à $\mathbf{R}$, les n_k à $\mathbf{N}$ et n est un indice entier, est uniformément dense dans $\mathcal{C}([a; b], \mathbf{R})$.*

Corollaire II .47 *Soient X et Y des espaces compacts. Dans l'espace de Banach $\mathcal{C}(X \times Y, \mathbf{R})$ est dense l'ensemble $\mathcal{H}$ des sommes finies de fonctions $(x, y) \mapsto f(x)g(y)$ où $f \in \mathcal{C}(X, \mathbf{R})$ et $g \in \mathcal{C}(Y, \mathbf{R})$.*

Preuve. En effet, $\mathcal{H}$ est une sous-algèbre de $\mathcal{C}(X \times Y, \mathbf{R})$ et deux points $(x_1, y_1) \neq (x_2, y_2)$ de $X \times Y$, où par exemple $x_1 \neq x_2$, sont séparés par $d_X(x_1, \cdot)$ qui est dans $\mathcal{H}$. □

Ce corollaire, très utile, permet de travailler à variables séparées pour des problèmes à deux variables.

Corollaire II .48 *Toute fonction $f = f(x)$ continue sur $\mathbf{R}$, à valeurs complexes, et de période 2π, est limite uniforme sur $\mathbf{R}$ de polynômes trigonométriques $\sum_{k=-n}^{n} a_k e^{ikx}$ où $n \in \mathbf{N}$ et les a_k appartiennent à $\mathbf{C}$.*

Preuve. Soit X le cercle $\{e \in \mathbf{C} : |z| = 1\}$; il est compact. À toute $f \in \mathcal{C}(\mathbf{R}, \mathbf{C})$ de période 2π, associons la fonction $\tilde{f}$ définie sur X par la formule $\tilde{f}(e^{ix}) = f(x)$ où $0 \leqslant x \leqslant 2\pi$. L'ensemble $\mathcal{H}$ des $\tilde{g}$, où g est un polynôme trigonométrique sur $\mathbf{R}$, n'est autre que l'ensemble $\mathcal{H}$ des fonctions $z \mapsto \sum_{k=-n}^{n} a_k z^k$, où $n \in \mathbf{N}$ et les a_k appartiennent à $\mathbf{C}$. Il est clair que $\mathcal{H}$ remplit les hypothèses (i) à (v) du corollaire II .44, ce qui achève la démonstration puisqu'évidemment :

$$\sup_{x \in \mathbf{R}} |f(x) - g_n(x)| = \sup_{z \in X} |\tilde{f}(z) - \tilde{g}_n(z)|. \qquad \square$$

Exercice II .49 1° *Utilisez vos connaissances classiques sur les séries de Fourier (théorème de Fejer) pour préciser l'énoncé du corollaire précédent.*

2° *Plus généralement, donnez un énoncé du type du corollaire précédent pour les fonctions périodiques de plusieurs variables.*

Exercice II .50 *On suppose connu (par exemple par le théorème de Fejer) le fait que toute fonction continue de période* 2π *est limite uniforme sur* $\mathbf{R}$ *de polynômes trigonométriques. En déduire une démonstration du théorème de Weierstrass historique : on se ramène à démontrer que les fonctions* $x \mapsto \sin(nx)$ *et* $x \mapsto \cos(nx)$ *sont sur* $[0; 2\pi]$ *limites uniformes de fonctions polynomiales, ce qui résulte de leur développement en série de Taylor.*

Revenons au théorème de Weierstrass historique : toute $f \in \mathcal{C}([0; 1], \mathbf{R})$ est limite uniforme sur $[0; 1]$ d'une suite $(P_n)_{n \geqslant 0}$ de fonctions polynomiales à coefficients réels. Si f est donnée, peut-on préciser explicitement de tels P_n qui conviennent pour approximer f ? Réponse : oui.

Théorème II .51 *Si* $f \in \mathcal{C}([0; 1], \mathbf{R})$*, posons pour tout* $n \geqslant 1$ *:*

$$P_n(x) = \sum_{k=0}^{n} C_n^k f(\frac{k}{n}) x^k (1-x)^{n-k} \quad x \in [0; 1].$$

Alors :

$$\lim_{n \to \infty} \|f - P_n\| = 0.$$

La démonstration commence, de façon élémentaire, par le

Lemme II .52 *Pour* $1 \leqslant k \leqslant n$*, posons :*

$$r_k^n(x) = r_k(x) = C_n^k x^k (1-x)^{n-k}.$$

Alors :

1. $\sum_{k=0}^{n} r_k(x) = 1$;
2. $\sum_{k=0}^{n} kr_k(x) = nx$;
3. $\sum_{k=0}^{n} k(k-1)r_k(x) = n(n-1)x^2$;
4. $\sum_{k=0}^{n}(k-nx)^2 r_k(x) = nx(1-x)$.

Preuve. La formule 1. est la formule du binôme. Pour 2, on calcule :

$$\sum_{k=0}^{n} kr_k(x) = nx\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!\,((n-1)-(k-1))!}x^{k-1}(1-x)^{n-1-(k-1)}$$

$$= nx\sum_{k=0}^{n-1}\frac{(n-1)!}{k!(n-1-k)!}x^k(1-x)^{n-1-k} = nx.$$

Pour 3, on calcule également :

$$\sum_{k=0}^{n} k(k-1)C_n^k x^k(1-x)^{n-k} = n(n-1)x^2\sum_{k=2}^{n}\frac{(n-2)!}{(k-2)!\,((n-2)-(k-2))!}x^{k-2}(1-x)^{n-2-(k-2)}$$

$$= n(n-1)x^2\sum_{k=0}^{n-2}\frac{(n-2)!}{k!(n-2-k)!}x^k(1-x)^{n-2-k} = n(n-1)x^2.$$

Enfin, le premier membre de 4 vaut :

$$\sum_k k^2 r_k(x) - 2nx\sum_k kr_k(x) + n^2x^2\sum_k r_k(x) = \sum_k k(k-1)r_k(x) - (2nx-1)\sum_k kr_k(x) + n^2x^2 = nx(1-x),$$

d'après 2 et 3 déjà démontrées. □

Passons à la démonstration du théorème de Bernstein.

Preuve. Soit $f \in \mathcal{C}$ et soit $\varepsilon > 0$. La fonction f est uniformément continue sur $[0;1]$ par le théorème de Heine, donc on peut choisir $\delta > 0$ tel que

$$|x - x'| \leqslant \delta \Rightarrow |f(x) - f(x')| \leqslant \frac{\varepsilon}{2}.$$

Fixons désormais un tel δ (qui ne dépend que de ε).

Pour tout n entier > 0 et tout $x \in [0;1]$, on a :

$$|f(x) - P_n(x)| = |f(x) - \sum_{k=0}^{n} f(\frac{k}{n})r_k(x)| = |\sum_{k=0}^{n}\left(f(x) - f(\frac{k}{n})\right)r_k(x)|$$

$$\leqslant \sum_{k:|k-nx|\leqslant \delta n}|f(x) - f(\frac{k}{n})|r_k(x) \;+\; |\sum_{k:|k-nx|>\delta n}\left(f(x) - f(\frac{k}{n})\right)r_k(x)|,$$

et donc par 1 et choix de δ :

$$|f(x) - P_n(x)| \leqslant \frac{\varepsilon}{2} + |\sum_{k:|k-nx|>\delta n} \left(f(x) - f(\frac{k}{n})\right) r_k(x)|.$$

Mais, d'autre part :

$$|\sum_{k:|k-nx|>\delta n} \left(f(x) - f(\frac{k}{n})\right) r_k(x)| \leqslant 2\|f\|_\infty \sum_{k:|k-nx|>\delta n} r_k(x)$$

$$\leqslant \frac{2\|f\|_\infty}{\delta^2 n^2} \sum_{k:|k-nx|>\delta n} (k-nx)^2 r_k(x) \leqslant \frac{2\|f\|_\infty}{\delta^2 n^2} \sum_{k=0}^{n} (k-nx)^2 r_k(x)$$

$$\frac{2\|f\|_\infty}{\delta^2 n^2} nx(1-x) \leqslant \frac{2\|f\|_\infty}{\delta^2} \frac{1}{n},$$

qui est aussi $\leqslant \frac{\varepsilon}{2}$ dès que n est assez grand. □

Notons qu'à degré n fixé, le polynôme de Bernstein P_n de f n'est pas le meilleur polynôme d'approximation uniforme de f sur $[0;1]$.

II.12 Théorème d'Urysohn

Dans un espace compact X, peut-on prolonger toute fonction définie et continue sur un sous-ensemble fermé de X en une fonction continue sur X tout entier ? La question est naturelle et la réponse affirmative.

Théorème II .53 (Urysohn) *Soit X un espace compact et soit Y un sous-ensemble fermé de X. Alors toute $f \in \mathcal{C}(Y, \mathbf{R})$ est la restriction à Y d'une fonction $F \in \mathcal{C}(X, \mathbf{R})$.*

Preuve. Soit $\mathcal{H}$ l'ensemble des restrictions à Y des fonctions continues sur X. Pour le compact Y, l'ensemble $\mathcal{H}$ satisfait aux hypothèses (i), (ii) et (iii) du théorème de Stone-Weierstrass. De plus, $\mathcal{H}$ sépare Y, car si $y_1 \in Y$, $y_2 \in Y$ et $y_1 \neq y_2$, la fonction $x \mapsto d(y,x)$ est définie et continue sur X et sa restriction à Y sépare y_1 et y_2. En appliquant le théorème de Stone-Weierstrass, on trouve déjà que l'ensemble des restrictions à Y des $F \in \mathcal{C}(X, \mathbf{R})$ est uniformément dense dans $\mathcal{C}(Y, \mathbf{R})$; c'est un pas en direction du théorème mais il faut faire mieux. Donc on sait déjà que :

> pour toute $g \in \mathcal{C}(Y, \mathbf{R})$ et pour tout $\varepsilon > 0$ il existe $g_\varepsilon \in \mathcal{C}(X, \mathbf{R})$ telle que : $|g(y) - g_\varepsilon(y)| \leqslant \varepsilon$ pour tout $y \in Y$.

Dans cette notation g_ε, on peut même supposer que

$$\|g_\varepsilon\|_{\infty, \mathcal{C}(X,\mathbf{R})} = \|g_\varepsilon\|_{\infty, \mathcal{C}(Y,\mathbf{R})}$$

quitte à tronquer g_ε en la remplaçant par la fonction $\sup(\alpha, \inf(\beta, g_\varepsilon))$ où $\alpha = \inf_{y \in Y} g(y)$ et $\beta = \sup_{y \in Y} g(y)$.

Soit $f \in \mathcal{C}(Y, \mathbf{R})$. Par récurrence, définissons une suite $g^{(n)}$ d'éléments de $\mathcal{C}(X, \mathbf{R})$ par les formules :

$$g^{(1)} = f_{\frac{1}{2}} \quad \text{et} \quad g^{(n)} = (f - \sum_{i=1}^{n-1} g^{(i)})_{\frac{1}{2^n}} \quad \text{pour } n > 1.$$

Par définition du symbole g_ε, employé ici pour ε valant les $\frac{1}{2^n}$, on a donc :

$$|f(y) - \sum_{i=1}^{k} g^{(i)}(y)| \leqslant \frac{1}{2^n} \quad \text{pour tout } y \in Y$$

et

$$|g^{(n)}(x)| \leqslant \frac{1}{2^{n-1}} \quad \text{pour tout } x \in X.$$

Ainsi la série de fonctions $\sum_{n=1}^{\infty} g^{(n)}(x)$ converge uniformément sur X vers une fonction F continue et, pour tout $y \in Y$, on a : $F(y) = f(y)$. □

Corollaire II .54 *Soient Y_1 et Y_2 deux fermés disjoints dans un espace compact X. Alors il existe une $f \in \mathcal{C}(X, \mathbf{R})$ telle que*

$$0 \leqslant f \leqslant 1, \quad f(y) = 1 \text{ si } y \in Y_1 \quad \text{et} \quad f(y) = 0 \text{ si } y \in Y_2.$$

Preuve. La fonction constante égale à 1 sur Y_1 et constante égale à 0 sur Y_2 est continue sur $Y_1 \cup Y_2$. Si F est un prolongement de cette fonction en une fonction continue sur X , il suffit de poser $f = \sup(0, \inf(1, F))$. □

Exercice II .55 *Montrer que le corollaire est satisfait dans tout espace métrique X, même non compact, en posant*

$$f(x) = \frac{d(x, Y_2)}{d(x, Y_1) + d(x, Y_2)},$$

où, si Y est fermé dans X, on pose $d(x, Y) = \inf_{y \in Y} d(x, y)$.

Troisième partie

Le théorème de Baire et ses applications

L'Analyse fournit des exemples divers et importants d'espaces métriques complets : les espaces $\mathbf{R}^n$, les espaces L^p de la théorie de la mesure, notamment L^1 et L^2, l'espace $(\mathcal{C}(X), \|\cdot\|_\infty)$ des fonctions continues sur X compact, l'espace de Hilbert des fonctions f holomorphes sur un ouvert U de $\mathbf{C}$ telles que $\|f\|^2 = \int_U |f(z)|^2 \, \mathrm{d}z < +\infty$, l'espace $\mathcal{L}(E, F)$ des applications linéaires continues d'un espace normé vers un espace de Banach F, avec la norme :

$$\|\varphi\| = \sup_{\|x\|_E \leqslant 1} \|\varphi(x)\|_F = \sup_{\|x\|_E = 1} \|\varphi(x)\|_F$$

si $E \neq \{0\}$, sont des espaces complets, i.e. dans lesquels toute suite de Cauchy est convergente. De plus, si E est complet, tout sous-ensemble fermé de E est lui-même complet, en tant que sous-espace métrique de E, par exemple le segment $[0; 1]$ de $\mathbf{R}$.

Influencé par les idées de Cantor sur le dénombrable et le transfini, René Baire (1874-1932) a découvert en 1899 dans le cas de $\mathbf{R}^n$ un théorème, qui en fait est valable dans tout espace métrique complet (et aussi dans tout sous-espace ouvert non vide d'un tel espace), et qui dit que si un tel espace n'est pas vide, il ne peut être réunion d'une suite (i.e. d'une famille dénombrable) de fermés tous d'intérieur vide, alors qu'en revanche, il est par exemple réunion de la famille de tous ses points, lesquels sont fermés d'intérieur vide en général. On voit donc combien les hypothèses de dénombrabilité sont essentielles dans le théorie de Baire.

Ce théorème a d'abord des applications à la théorie des fonctions discontinues ; si une telle fonction est limite simple de fonctions continues, alors elle n'est pas trop discontinue : l'ensemble des points où elle est discontinue est *maigre* en un sens que nous préciserons.

D'autres applications du théorème de Baire s'exercent en théorie des espaces de Banach. Dans cette théorie, il y a trois résultats fondamentaux. L'un d'eux, le théorème de Hahn-Banach, a été étudié au chapitre I. Les deux autres, le théorème de Banach-Steinhaus et le théorème de l'application ouverte (ou théorème de Banach) sont des conséquences du théorème de Baire ; ils sont établis dans le présent chapitre, avec quelques applications à la théorie de l'approximation, par exemple la non convergence uniforme du procédé d'interpolation de Lagrange, la non convergence en général de la série de Fourier d'une application continue etc.

III.13 Théorème de Baire

Commençons par formuler un lemme sur les espaces métriques complet.

Lemme III .1 (Cantor) *Soit (E, d) un espace métrique complet. Soit $(F_n)_{n \geqslant 0}$ une suite décroissante pour l'inclusion de fermés non vides de E dont les* *diamètres*

$$\delta(F_n) = \sup_{x,x' \in F_n} d(x,x')$$

tendent vers 0 *quand* $n \to \infty$. *Alors l'intersection des* F_n *n'est pas vide ; elle est en fait réduite à un point.*

Preuve. Pour tout entier $n \geqslant 0$, choisissons un point x_n dans l'ensemble non vide F_n. Si $m \geqslant n$, les points x_m et x_n sont dans F_n par décroissance de $(F_n)_{n\geqslant 0}$, donc

$$m \geqslant n \Rightarrow d(x_m, x_n) \leqslant \delta(F_n) \leqslant \varepsilon$$

dès que n est assez grand puisque $\lim_{n\to\infty} \delta(F_n) = 0$. Ceci prouve que la suite $(x_n)_{n\geqslant 0}$ est de Cauchy, et donc qu'elle a une limite, disons x, puisque E est complet.

Pour p entier fixé, montrons que $x \in F_p$. Pour tout $k \geqslant 0$, le point x_{p+k} est dans F_p, donc puisque F_p est fermé :

$$x = \lim_{k\to\infty} x_{p+k} \in F_p.$$

En fait, $\bigcap_{n\geqslant 0} F_n = \{x\}$ car si y est dans cette intersection on a pour tout n : $d(x,y) \leqslant \delta(F_n)$, et donc en faisant $n \to \infty$ on obtient : $d(x,y) = 0$. □

Exercice III .2 *Soit* (E,d) *un espace métrique complet. Soit* $0 \leqslant k < 1$. *Soit* f *une application de* E *dans* E *contractante de rapport* k *: pour tous* $x, y \in E$, *on a :*

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

Pour tout entier $n \geqslant 1$, *soit* $F_n = \{x \in E : d(f(x),x) \leqslant \frac{1}{n}\}$. *En appliquant aux parties* F_n *le lemme de Cantor, retrouver un théorème de Banach.*

Rappelons maintenant quelques notions de topologie dans un espace métrique (E,d). Dire qu'une partie X est *dense* dans E, c'est dire que :

> pour tout $x \in E$ on a une suite $(x_n)_{n\geqslant 0}$ dans X telle que $\lim_{n\to\infty} x_n = x$;
> ou encore : l'adhérence de X est E tout entier ;
> ou encore : pour tout $y \in E$ et tout $\varepsilon > 0$, on a $B(y,\varepsilon) \cap X \neq \varnothing$;
> ou encore : pour tout ouvert non vide U de E, a $U \cap X \neq \varnothing$.

Soit X une partie de E. L'*intérieur* X° de X est le plus grand ouvert de E contenu dans X ; autrement dit, c'est la réunion des ouverts de E contenus dans X. Un élément $y \in E$ appartient à X° s'il existe $r > 0$ tel que $B(y,r) \subset X$. Souvent, X° est vide (penser à sous-espace vectoriel X

strictement inclus dans un espace vectoriel normé E) ; dans ce cas, on dit que X est *sans point intérieur*, ou encore d'*intérieur vide*.

Pour relier les deux précédentes notions, on peut remarquer qu'une partie X est dense dans E si, et seulement si, son complémentaire $E\backslash X$ est d'intérieur vide : $\overline{X} = E \Leftrightarrow (E\backslash X)^\circ = \varnothing$, et donc $X^\circ = \varnothing \Leftrightarrow \overline{E\backslash X} = X$.

Rappelons aussi la structure des ouverts de $\mathbf{R}$: tout ouvert de $\mathbf{R}$ est réunion au plus dénombrable d'intervalles ouverts deux à deux disjoints (qui sont ses composantes connexes).

Exercice III .3 *Redémontrez ce fait.*

Dans l'exemple le plus simple, le théorème de Baire dit que $\mathbf{R}$ n'est pas réunion dénombrable de fermés d'intérieur vide. Plus généralement :

Théorème III .4 *Dans un espace métrique complet (E, d) non vide, les assertions suivantes sont satisfaites :*

(i) *l'intersection de toute suite, i.e. de toute famille dénombrable, d'ouverts denses est non vide, et en fait est même encore dense ;*
(ii) *l'intersection de toute suite, i.e. de toute famille dénombrable, de fermés d'intérieur vide est encore d'intérieur vide, et en particulier n'est pas l'espace E tout entier.*

Ainsi, un espace métrique complet n'est pas réunion dénombrable de fermés sans point intérieur.

Preuve. Le point (ii) équivaut à (i) par passage au complémentaire. Prouvons (i). Soit $(U_n)_{n\geqslant 0}$ une suite d'ouverts tous denses dans E. Soit V un ouvert non vide de E. On va prouver que $V \cap \bigcap_{n\geqslant 0} U_n$ n'est pas vide. Par récurrence sur n, construisons une suite de boules ouvertes $B(x_n, r_n)$ telles que

$$\begin{aligned}
&\overline{B(x_1, r_1)} \subset U_1 \cap V \text{ (possible car } U_1 \cap V \neq \varnothing\text{, vu que } U_1 \text{ est dense)}\,;\\
&\overline{B(x_2, r_2)} \subset U_2 \cap B(x_1, r_1) \text{ (possible car } U_2 \text{ est ouvert dense)}\,;\\
&\dots\\
&\overline{B(x_n, r_n)} \subset U_n \cap B(x_{n-1}, r_{n-1}) \text{ et } r_n < \tfrac{1}{2}\cdot r_{n-1} \text{ pour tout } n \geqslant 2.
\end{aligned}$$

Les $F_n = \overline{B(x_n, r_n)}$ sont une suite de décroissante de fermés non vides dont les diamètres $\delta(F_n) = 2r_n$ tendent vers 0. D'après le lemme d'intersection de Cantor, il y a au moins un point x dans $\bigcap_{n\geqslant 0} \overline{B(x_n, r_n)}$ et donc dans $\left(\bigcap_{n\geqslant 0} U_n\right) \cap V$. □

Corollaire III .5 *Soit Ω un ouvert non vide dans un espace métrique complet (E,d). Les assertions* (i) *et* (ii) *du théorème de Baire restent vérifiées en remplaçant (E,d) par Ω muni de la distance induite.*

Preuve. En effet, soit $(U_n)_{n\geqslant 0}$ une suite d'ouverts de Ω tous denses dans Ω ; les ouverts U_n sont encore des ouverts de $\overline{\Omega}$ tous denses dans $\overline{\Omega}$. Mais l'adhérence $\overline{\Omega}$ de Ω dans E est un espace complet, auquel on peut appliquer le théorème de Baire : l'intersection des U_n est dense dans $\overline{\Omega}$, donc a fortiori dans Ω. □

Ainsi, le théorème de Baire est valable dans l'intervalle de $]-1;1[$ de $\mathbf{R}$, bien que ce ne soit pas un espace complet.

III.14 Applications à la théorie des fonctions discontinues

Une fonction $f : \mathbf{R} \to \mathbf{R}$ est dite *de la première classe de Baire* si elle est limite simple sur $\mathbf{R}$ d'une suite $(f_n)_{n\geqslant 0}$ de fonctions continues sur $\mathbf{R}$. Bien entendu, une telle f peut être discontinue, mais nous allons voir qu'elle ne peut pas l'être trop.

Pour étudier cette question, rappelons – ou introduisons – la notion d'*oscillation* d'une fonction $f : \mathbf{R} \to \mathbf{R}$ au point $x \in \mathbf{R}$; c'est par définition le nombre (fini ou infini) :

$$\omega_f(x) = \inf_{\varepsilon>0} \sup_{y,y'\in B(x,\varepsilon)} |f(y') - f(y)|.$$

Proposition III .6 *La fonction f est continue en x si, et seulement si, $\omega_f(x) = 0$.*

Preuve. En effet, d'après le critère de Cauchy, valable parce que $\mathbf{R}$ est complet par construction, dire que $\lim\limits_{y\to x} f(y) = f(x)$ équivaut à dire que : pour tout $\eta > 0$ il existe $\varepsilon > 0$ tel que $y, y' \in B(x,\varepsilon) \Rightarrow |f(y) - f(y')| \leqslant \varepsilon$. Autrement dit :

$$\text{pour tout } \eta > 0 \text{ il existe } \varepsilon > 0 \text{ tel que } \sup_{y,y'\in B(x,\varepsilon)} |f(y') - f(y)| \leqslant \varepsilon.$$

Ou encore : $\inf\limits_{\varepsilon>0} \sup\limits_{y,y'\in B(x,\varepsilon)} |f(y') - f(y)|$. □

Si $f : \mathbf{R} \to \mathbf{R}$ est quelconque, on prendra garde au fait que la fonction $x \mapsto \omega_f(x)$ n'est pas continue en général ; cependant, on a :

Lemme III .7 *Soit $f : \mathbf{R} \to \mathbf{R}$ et soit α un nombre réel > 0. Alors :*

$$U_\alpha = \{x \in \mathbf{R} : \omega_f(x) < \alpha\}$$

est un ensemble ouvert dans $\mathbf{R}$.

Preuve. On remarque d'abord que la fonction $\varepsilon \mapsto \sup_{y,y' \in B(x,\varepsilon)} |f(y') - f(y)|$ est croisssante et $\geqslant 0$. Soit $x \in U_\alpha$. Alors il existe $\eta > 0$ et $\varepsilon > 0$ tels que

$$y, y' \in B(x, \varepsilon) \Rightarrow |f(y') - f(y)| \leqslant \alpha - \eta.$$

Soit $x' \in B(x, \frac{\varepsilon}{2})$. On se donne $y, y' \in B(x', \frac{\varepsilon}{2})$. Par inégalité triangulaire, on a $y, y' \in B(x, \varepsilon)$ et donc $|f(y') - f(y)| \leqslant \alpha - \eta$. En passant à la borne supérieure, on obtient

$$\sup_{y,y' \in B(x', \frac{\varepsilon}{2})} |f(y') - f(y)| \leqslant \alpha - \eta,$$

et donc $\omega_f(x') \leqslant \alpha - \eta$. Ceci prouve donc que $B(x, \frac{\varepsilon}{2}) \subset U_\alpha$, et donc que U_α est ouvert puisque c'est un voisinage de tous ses points. □

Lemme III .8 *Soit* $f : \mathbf{R} \to \mathbf{R}$ *dans la première classe de Baire et soit* α *un nombre réel* > 0. *Alors l'ensemble ouvert :*

$$U_\alpha = \{x \in \mathbf{R} : \omega_f(x) < \alpha\}$$

est dense dans $\mathbf{R}$.

Preuve. Soit $(f_n)_{n \geqslant 0}$ une suite de fonctions continues $\mathbf{R} \to \mathbf{R}$ qui converge simplement sur $\mathbf{R}$ vers f. Soit Ω un intervalle ouvert non vide de $\mathbf{R}$. On va prouver que $\Omega \cap U_\alpha \neq \varnothing$. Pour tout entier $p \geqslant 1$, l'ensemble

$$F_p = \{x \in \Omega : \text{ quel que soit } q \geqslant p, \quad |f_q(x) - f_p(x)| \leqslant \frac{\alpha}{3}\}$$

est fermé dans l'espace métrique Ω, car les fonctions différence $f_q - f_p$ sont continues. De plus, on a $\Omega = \bigcup_{p \geqslant 1} F_p$ car, pour tout $x \in \Omega$ fixé, la suite $(f_p(x))_{p \geqslant 1}$ est de Cauchy dans $\mathbf{R}$. Appliquons à l'espace Ω le corollaire III .5 du théorème de Baire : il existe un indice p_0 tel que F_{p_0} soit d'intérieur non vide dans Ω. On va montrer que :

$$(*) \quad (F_{p_0})^\circ \subset U_\alpha,$$

ce qui achèvera la démonstration car alors $\Omega \cap U_\alpha \supset (F_{p_0})^\circ \neq \varnothing$.

Soit $x \in (F_{p_0})^\circ$. En faisant $q \to \infty$ dans la définition de F_{p_0}, on voit que

$$|f(y) - f_{p_0}(y)| \leqslant \frac{\alpha}{3}$$

pour tout $y \in F_{p_0}$; d'où, par inégalité triangulaire :

$$|f(y) - f(y')| \leqslant |f(y) - f_{p_0}(y)| + |f_{p_0}(y) - f_{p_0}(y')| + |f_{p_0}(y') - f(y')|$$

$$\leqslant \frac{\alpha}{3} + |f_{p_0}(y) - f_{p_0}(y')| + \frac{\alpha}{3} < \alpha$$

dès que y et y' sont assez proches de x tout en étant dans F_{p_0}, ce qui est loisible car $(F_{p_0})^\circ$ est un ouvert non vide contenant x, et vu la continuité de f_{p_0} au point x, qui permet de rendre le terme médian $< \frac{\alpha}{3}$. Ceci prouve que $\omega_f(x) < \alpha$ et donc $(*)$. □

On en déduit le

Théorème III .9 *Soit $f : \mathbf{R} \to \mathbf{R}$ dans la première classe de Baire. Alors l'ensemble X des points de $\mathbf{R}$ où f est continue est intersection dénombrable d'ouverts denses. En particulier, par le théorème de Baire, X est dense dans $\mathbf{R}$.*

Preuve. En effet, on a :

$$X = \{x \in \mathbf{R} : \omega_f(x) = 0\} = \bigcap_{n \geqslant 1} U_{\frac{1}{n}},$$

où les $U_{\frac{1}{n}}$ sont ouverts et denses dans $\mathbf{R}$ par les lemmes qui précèdent. □

Bien entendu, le même énoncé reste valable pour tout espace métrique complet (E, d) à la place de la droite réelle.

Dans E, disons qu'un ensemble est *maigre* s'il est réunion au plus dénombrable d'ensembles dont l'adhérence est d'intérieur vide. Le théorème III .9 se lit encore : si f est limite simple de fonctions continues sur E complet, l'ensemble de ses points de discontinuité est maigre dans E.

Exercice III .10 *Soit $f : \mathbf{R} \to \mathbf{R}$ dans la première classe de Baire. Montrer que, pour tout intervalle ouvert non vide I, il existe un sous-intervalle ouvert non vide J de I sur lequel f est bornée.*

Exercice III .11 *Soit l'espace métrique $\mathbf{Q}$ des nombres rationnels, muni de la distance de la valeur absolue des nombres réels. Soit une fonction $f : \mathbf{Q} \to \mathbf{R}$ continue. Alors l'ensemble*

$$D = \{x \in \mathbf{R} : f(r) \text{ n'a pas de limite quand } r \to x, x \in \mathbf{Q}\}$$

est maigre dans $\mathbf{R}$. Indication : introduire

$$\tilde{\omega}_f(x) = \inf_{\varepsilon > 0} \sup_{r, r' \in B(x, \varepsilon) \cap \mathbf{Q}} |f(r') - f(r)|;$$

remarquer que $D = \{x \in \mathbf{R} : \tilde{\omega}_f(x) \neq 0\}$, puis montrer que les ensembles $U_\alpha = \{x \in \mathbf{R} : \tilde{\omega}_f(x) < \alpha\}$ sont ouverts et denses dans $\mathbf{R}$, la densité n'étant pas difficile à voir ici.

III.15 Théorème de Banach-Steinhaus

Rappelons que si E et F sont deux espaces normés sur le même corps des scalaires $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$, l'espace vectoriel $\mathcal{L}(E, F)$ sur $\mathbf{K}$ de toutes les applications linéaires et continues φ de E dans F est normé par :

$$|||\varphi||| = \sup_{\|x\|_E \leqslant 1} \|\varphi(x)\|_F.$$

Si $E \neq \{0\}$ on peut calculer cette borne supérieure sur la sphère unité : $|||\varphi||| = \sup_{\|x\|_E=1} \|\varphi(x)\|_F$. De plus, pour tout $x \in E$, on a par définition l'inégalité :

$$\|\varphi(x)\|_F \leqslant |||\varphi||| \|x\|_E.$$

Un cas important est celui où $F = \mathbf{K}$ est de dimension 1. Alors $\mathcal{L}(E, \mathbf{K})$ n'est autre que le dual (topologique) E' de E.

Soit Φ une partie de $\mathcal{L}(E, F)$. Supposons Φ bornée dans l'espace normé $\mathcal{L}(E, F)$, c'est-à-dire qu'il existe une constante M telle que pour toute $\varphi \in \Phi$ on ait : $|||\varphi||| \leqslant M$. Alors il est clair que Φ est bornée en chaque point x de E, c'est-à-dire qu'il existe une constante M_x telle que pour toute $\varphi \in \Phi$ on ait : $\|\varphi(x)\|_F \leqslant M_x$. En effet, il suffit de prendre $M_x = |||\varphi||| \|x\|_E$. Cette conclusion semble être un affaiblissement considérable de l'hypothèse. C'est un résultat très fort que la réciproque soit vraie sous l'hypothèse que E est complet.

Théorème III .12 (Banach-Steinhaus) *Soit E un espace de Banach et soit F un espace normé sur le même corps des scalaires $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$. Soit $\Phi \subset \mathcal{L}(E, F)$ un ensemble d'applications linéaires et continues de E dans F. On suppose que pour tout $x \in E$, on a :*

$$\sup_{\varphi \in \Phi} \|\varphi(x)\|_F < +\infty,$$

alors on a :

$$\sup_{\varphi \in \Phi} |||\varphi||| < +\infty.$$

Autrement dit, si la famille Φ est bornée en chaque point x de E, alors elle est uniformément bornée sur tous les points de la boule unité de E ; on parle en anglais de *uniform boundedness principle.*

Preuve. Pour chaque entier $n \geqslant 1$, notons :

$$V_n = \{x \in E : \sup_{\varphi \in \Phi} \|\varphi(x)\|_F > n\}.$$

D'abord, on remarque que chaque V_n est ouvert dans E. Fixons en effet $n \geqslant 1$. Pour $x_0 \in V_n$, il suffit de trouver un ouvert U_0 tel que $x_0 \in U_0 \subset V_n$. Puisque

$$\sup_{\varphi \in \Phi} \|\varphi(x_0)\|_F > n,$$

il existe $\varphi_0 \in \Phi$ telle que $\|\varphi_0(x_0)\|_F > n$; mais la fonction $x \mapsto \|\varphi_0(x)\|_F$ est continue donc l'ensemble

$$U_0 = \{x \in E : \|\varphi_0(x)\|_F > n$$

est ouvert dans E, et visiblement $x_0 \in U_0 \subset V_n$.

Ensuite, on remarque que l'un au moins des V_n n'est pas dense dans E : sinon, par le théorème de Baire III .4, l'intersection $\bigcap_{n \geqslant 1} V_n$ serait non vide : il existerait donc $x \in E$ tel que $\sup_{\varphi \in \Phi} \|\varphi(x)\|_F = +\infty$, contrairement à l'hypothèse.

On prend donc N tel que V_N n'est pas dense dans E. Le complémentaire de V_N est donc d'intérieur non vide : il contient une boule ouverte. Il existe donc $a \in E$ et $r > 0$ tel que : $\|x\| \leqslant r \Rightarrow a + x \notin V_N$. Ainsi, pour toute $\varphi \in \Phi$, on a $\|\varphi(a+x)\| \leqslant N$ dès que $\|x\| \leqslant r$; ce qui entraîne que pour toute $\varphi \in \Phi$:

$$\|\varphi(x)\| \leqslant \|\varphi(a+x)\| + \|\varphi(a)\| \leqslant 2N$$

dès que dès que $\|x\| \leqslant r$; autrement dit $\sup_{\|x\| \leqslant 1} \|\varphi(x)\| \leqslant \frac{2N}{r}$. Ainsi, pour toute $\varphi \in \Phi$, on a :

$$|||\varphi||| \leqslant \frac{2N}{r},$$

et donc $\sup_{\varphi \in \Phi} |||\varphi||| < +\infty$. □

Exercice III .13 *Soit E un espace de Banach et F un espace normé sur le même corps des scalaires $\mathbf{K} = \mathbf{R}$ ou $\mathbf{C}$. Soit $(\varphi)_{n \geqslant 1}$ une suite d'applications linéaires continues qui converge simplement sur E vers une application $\varphi : E \to F$. Montrer que $\Phi = \{\varphi\}_{n \geqslant 1}$ vérifie l'hypothèse du théorème de Banach-Steinhaus ; en déduire que Φ est équicontinue, et finalement que φ est continue.*

III.16 Application à la non-convergence simple des séries de Fourier

Faisons tout d'abord quelques rappels sur les séries de Fourier. À toute fonction $f : \mathbf{R} \to \mathbf{R}$ continue de période 2π, on associe ses *coefficients de Fourier* :

$$c_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, \mathrm{d}x \quad (n \in \mathbf{Z})$$

et sa *série de Fourier* : $\displaystyle\sum_{n\in\mathbf{Z}} c_n(f)e^{inx}$.

Les sommes de Fourier d'ordre n de f sont les

$$S_n(f)(x) = \sum_{k=-n}^{n} c_k(f)e^{ikx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_n(x-t)\ \mathrm{d}t,$$

où D_n est le *noyau de Dirichlet* :

$$D_n(t) = \sum_{k=-n}^{n} e^{ikt} = \frac{\sin\left((n+\frac{1}{2})t\right)}{\sin(\frac{t}{2})}.$$

On sait que :

$$\lim_{n\to\infty} \|S_n(f) - f\|_{\mathrm{L}^2([-\pi;\pi])} = 0$$

et que, par conséquent, il existe une suite extraite de $(S_n(f))_{n\geqslant 1}$ qui converge vers f presque partout.

Il est donc naturel de se demander s'il y a convergence simple de $(S_n(f))_{n\geqslant 1}$ sur $[-\pi;\pi]$ pour toute fonction $f : \mathbf{R} \to \mathbf{R}$ continue de période 2π.

La réponse est négative : grâce au théorème de Banach-Steinhaus, on va prouver que pour tout $x \in \mathbf{R}$ il existe une telle fonction f telle que

$$\sup_{n\geqslant 1} |S_n(f)(x)| = +\infty.$$

La suite des $S_n(f)(x)$ étant non bornée, elle n'a évidemment pas de limite finie et en particulier ne converge pas vers $f(x)$.

Nous utilisons pour cela une estimation du noyau de Dirichlet.

Lemme III .14 *On a :*

$$\|D_n\|_{\mathrm{L}^1([-\pi;\pi])} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\ \mathrm{d}t \ \geqslant\ \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k},$$

et donc :

$$\lim_{n\to\infty} \|D_n\|_{\mathrm{L}^1([-\pi;\pi])} = +\infty.$$

Preuve. On utilise un argument de parité et l'inégalité $\sin(\frac{t}{2}) \leqslant \frac{t}{2}$ pour $0 \leqslant t \leqslant \pi$ pour le calcul qui suit.

$$\|D_n\|_{\mathrm{L}^1([-\pi;\pi])} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\ \mathrm{d}t = \frac{1}{\pi}\int_{0}^{\pi} |D_n(t)|\ \mathrm{d}t \ \geqslant\ \frac{2}{\pi}\int_{0}^{\pi} \left|\sin\left((n+\frac{1}{2})t\right)\right| \frac{\mathrm{d}t}{t},$$

et un changement de variable fournit :

$$\|D_n\|_{\mathrm{L}^1([-\pi;\pi])} \ \geqslant\ \frac{2}{\pi}\int_{0}^{(n+\frac{1}{2})\pi} |\sin(u)|\,\frac{\mathrm{d}u}{u} \ \geqslant\ \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k\pi}\int_{(k-1)\pi}^{k\pi} |\sin(u)|\,\mathrm{d}u = \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k},$$

ce qui est le résultat annoncé. □

Reprenons notre problème de convergence simple de (sommes partielles de) séries de Fourier $(S_n(f))_{n\geqslant 1}$. Soit E l'espace de Banach des fonctions $f : \mathbf{R} \to \mathbf{C}$ continues, à valeurs complexes, de période 2π, qu'on munit de la norme de la convergence uniforme $\|\cdot\|_\infty$:

$$\|f\|_\infty = \sup_{x\in[-\pi;\pi]} |f(x)| = \sup_{x\in\mathbf{R}} |f(x)|.$$

Soit $F = \mathbf{C}$. Par translation, on se ramène à se placer au point $x = 0$. On travaille avec la suite $(\varphi_n)_{n\geqslant 0}$ de formes linéaires dans $E' = \mathcal{L}(E, \mathbf{C})$ définies par évaluation en 0 :

$$\varphi_n : f \mapsto S_n(f)(0) = \sum_{k=n}^{n} c_k(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_n(t)\ \mathrm{d}t.$$

D'après la théorie de l'intégration, on sait que

$$|||\varphi_n||| = \sup_{f\in E, \|f\|_\infty=1} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_n(t)\ \mathrm{d}t = \|D_n\|_{\mathrm{L}^1([-\pi;\pi])}.$$

Ainsi, d'après le lemme III .14, on a : $\sup_{n\geqslant 1} |||\varphi_n||| = +\infty$. En lisant le théorème de Banach-Steinhaus à l'envers, on voit qu'il existe une fonction $f \in E$ telle que

$$\sup_{n\geqslant 1} |\varphi_n(f)| = |S_n(f)(0)| = +\infty,$$

ce qui est ce qu'on avait annoncé. □

Exercice III .15 *Il existe une notion de série de Fourier dans le cadre* L^1 *: à toute fonction intégrable* $f : [-\pi;\pi] \to \mathbf{C}$ *on associe une série de Fourier. Montrer qu'il existe* $f \in \mathrm{L}^1([-\pi;\pi])$ *telle que la suite des* $S_n(f)$ *ne tende pas vers* f *en norme* L^1 *quand* $n \to \infty$*. Indication : on pourra montrer que l'application linéaire* $S_n : f \mapsto S_n(f)$ *a pour norme* $\|D_n\|_{\mathrm{L}^1([-\pi;\pi])}$ *en tant qu'application de* $\mathrm{L}^1([-\pi;\pi])$ *dans lui-même.*

Ainsi, la situation dans le cadre L^1 est bien différente de celle dans le cadre L^2. Information sans démonstration : pour tout $p \in]1; +\infty[$ et toute $f \in \mathrm{L}^p([-\pi;\pi])$, on a : $\lim_{n\to\infty} \|S_n(f) - f\|_{\mathrm{L}^p([-\pi;\pi])} = 0$.

Rappelons aussi qu'en ce qui concerne les fonctions continues 2π-périodiques, la situation est meilleure si l'on s'intéresse à la convergence en moyenne de Cesaro, i.e. si l'on remplace les sommes partielles de Fourier $S_n(f)$ par les sommes partielles de Fejer :

$$\sigma_n(f) = \frac{1}{n+1}\sum_{k=0}^{n} S_k(f),$$

ce qui revient à remplacer le noyau de Dirchlet, par une autre fonction (*noyau de Fejer*), positive ou nulle cette fois. On sait en effet que la suite $(\sigma_n(f))_{n\geqslant 1}$ converge uniformément vers f sur $[-\pi;\pi]$ (et sur $\mathbf{R}$).

Exercice III .16 *En supposant connu le théorème de Fejer, prouver que pour toute fonction $f \in \mathrm{L}^1([-\pi;\pi])$ on a :* $\lim\limits_{n\to\infty} \|f-\sigma_n(f)\|_{\mathrm{L}^1([-\pi;\pi])} = 0$.

III.17 Application à la non-convergence uniforme de l'interpolation de Lagrange

Pour tout entier $n \geqslant 1$ et tout entier k tel que $0 \leqslant k \leqslant n$, le polynôme :

$$l_k^n(x) = \prod_{j=0, j\neq k}^{n} \frac{nx-j}{k-j}$$

est l'unique polynôme de degré n qui vaut 1 en $\frac{k}{n}$ et 0 en les points $\frac{j}{n}$ avec $0 \leqslant j \leqslant n$ et $j \neq k$. À toute fonction $f \in \mathcal{C}([0;1],\mathbf{C})$ on associe ses *polynômes d'interpolation de Lagrange* $p_n(f)$ définis par

$$p_n(f) : x \mapsto \sum_{k=0}^{n} f(\frac{k}{n}) l_k^n(x)$$

pour tout $n \geqslant 1$. Ainsi le polynôme $p_n(f)$ est l'unique polynôme de degré n qui prend les mêmes valeurs que f aux points $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$.

Théorème III .17 *Il existe une fonction continue $f : [0;1] \to \mathbf{C}$ telle que la suite des polynômes d'interpolation de Lagrange $(p_n(f))_{n\geqslant 1}$ ne converge pas uniformément vers f sur $[0;1]$ quand $n \to \infty$.*

Commençons par deux lemmes.

Lemme III .18 *Munissons $E = \mathcal{C}([0;1],\mathbf{C})$ de la norme $\|\cdot\|_\infty$ de la convergence uniforme. Alors l'application $f \mapsto p_n(f)$ de E dans lui-même est de norme* $|||p_n||| = \sup\limits_{x\in[0;1]} \sum\limits_{k=0}^{n} |l_k^n(x)|$.

Preuve. Pour toute $f \in E = \mathcal{C}([0;1],\mathbf{C})$ et tout $x \in [0;1]$, on a :

$$|p_n(f)(x)| = |\sum_{k=0}^{n} f(\frac{k}{n}) l_k^n(x)| \leqslant \sum_{k=0}^{n} |f(\frac{k}{n})||l_k^n(x)| \leqslant \|f\|_\infty \sum_{k=0}^{n} |f(\frac{k}{n})|,$$

donc

$$\|p_n(f)\|_\infty = \sup_{x\in[0;1]} |p_n(f)(x)| \leqslant \|f\|_\infty \sup_{x\in[0;1]} \sum_{k=0}^{n} |f(\frac{k}{n})|.$$

Ceci fournit déjà : $\|\|p_n\|\| \leqslant \sup_{x\in[0;1]} \sum_{k=0}^{n} |l_k^n(x)|$, et il s'agit de prouver l'inégalité inverse.

Soit x_0 un point de $[0;1]$ où la fonction $x \mapsto \sum_{k=0}^{n} |l_k^n(x)|$ atteint son maximum sur $[0;1]$. Choisissons f_0 une fonction continue sur $[0;1]$ telle que $\|f\|_\infty = 1$, et en outre telle que $f_0(\frac{k}{n}) = \text{signe}\,(l_k^n(x_0))$ c'est-à-dire que $f_0(x)$ vaut 1, -1 ou 0 suivant que $l_k^n(x_0)$ est > 0, < 0 ou $= 0$. Alors :

$$\|\|p_n\|\| \geqslant \|p_n(f_0)\|_\infty \geqslant |p_n(f_0)(x_0)| = \sum_{k=0}^{n} |f(x_0)| = \sup_{x\in[0;1]} \sum_{k=0}^{n} |f(x)|,$$

qui est la seconde inégalité recherchée. □

Lemme III .19 *On a :* $\lim_{n\to\infty} \|\|p_n\|\| = +\infty$.

Preuve. En posant $M_n(x) = \sum_{k=0}^{n} |f(x)|$, nous allons estimer $M(\frac{1}{2n})$ et prouver que cette suite tend vers $+\infty$; il en résultera bien que

$$\|\|p_n\|\| = \sup_{x\in[0;1]} M_n(x) \geqslant M(\frac{1}{2n})$$

tend vers $+\infty$. On a :

$$|l_k^n(\frac{1}{2n})| = \prod_{j=0, j\neq k}^{n} \frac{|\frac{n}{2n} - j|}{|k-j|} = \frac{1}{2^n} \prod_{j=0, j\neq k}^{n} \frac{1-2j}{k-j} = \frac{1}{2^n} \frac{1}{|2k-1|} \frac{\prod_{j=0}^{n} |1-2j|}{\prod_{j=0, j\neq k}^{n} |k-j|}$$

$$\frac{1}{2^n} \frac{1}{|2k-1|} \frac{1\cdot 3\cdot \cdots \cdot (2n-1)}{k!(n-k)!} = \frac{1}{2^{2n}} \frac{1}{|2k-1|} \frac{(2n)!}{(n!)^2} \frac{n!}{k!(n-k)!},$$

donc

$$M_n(\frac{1}{2^n}) = \frac{(2n)!}{2^{2n}(n!)^2} \sum_{k=0}^{n} C_n^k \frac{1}{|2k-1|}.$$

Or :

$$\sum_{k=0}^{n} C_n^k \frac{1}{|2k-1|} \geqslant \sum_{k=0}^{n} C_n^k \frac{1}{2k+1} = \int_0^1 \left(\sum_{k=0}^{n} C_n^k x^{2k} \right) \mathrm{d}x = \int_0^1 (1+x^2)\ \mathrm{d}x \geqslant \int_0^1 (2x)^n\ \mathrm{d}x = \frac{2^n}{n+1}$$

D'après la formule de Stirling, on sait que $n! \sim_{n\to\infty} \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$, donc cela fournit :

$$\frac{(2n)!}{2^{2n}(n!)^2} \sim_{n\to\infty} \frac{1}{\sqrt{\pi}} \frac{(2n)^{2n+\frac{1}{2}}}{2^{2n}\cdot n^{2n+1}} = \frac{1}{\sqrt{\pi n}}.$$

Ainsi : $M_n(\frac{1}{2^n}) \geqslant \frac{(2n)!}{2^{2n}(n!)^2} \cdot \frac{2^n}{n+1}$, qui équivaut quand $n \to \infty$ à $\frac{1}{\sqrt{\pi}} \frac{2^n}{n^{\frac{3}{2}}}$, lequel tend bien vers $+\infty$ quand $n \to \infty$. □

Nous pouvons maintenant passer à la démonstration du théorème III .17.

Preuve. On applique le théorème de Bananch-Steinhaus III .12 avec $E = F = \mathcal{C} = \mathcal{C}([0;1], \mathbf{C})$ muni de la norme de la convergence uniforme, et Φ l'ensemble des application p_n. Puisque $\sup_{n \geqslant 1} |||p_n||| = +\infty$, il existe une fonction $f \in \mathcal{C}([0;1], \mathbf{C})$ telle que $\sup_{n \geqslant 1} \|p_n(f)\|_\infty = +\infty$. N'étant pas bornée dans $\mathcal{C}$, la suite $(p_n(f))_{n \geqslant 1}$ ne peut converger dans $\mathcal{C}$. □

Exercice III .20 *Soit $(a_n)_{n \geqslant 0}$ une suite de nombres réels telle que quelle que soit la suite $x = (x_n)_{n \geqslant 0} \in \ell^1(\mathbf{N})$, la série $\sum_{n \geqslant 0} a_n x_n$ converge. Montrer qu'alors la suite $(a_n)_{n \geqslant 1}$ appartient à $\ell^\infty(\mathbf{N})$. Indication : considérer la suite $(\varphi_n)_{n \geqslant 0}$ de formes linéaires sur $\ell^1(\mathbf{N})$ définie par $\varphi_n(x) = \sum_{k=1}^n a_k x_k$ et lui appliquer le théorème de Bananch-Steinhaus. Dans cet énoncé, peut-on remplacer ℓ^1 et ℓ^∞ par ℓ^p et ℓ^q où p et q sont réels et tels que $\frac{1}{p} + \frac{1}{q} = 1$?*

III.18 Théorème de Banach de l'application ouverte

C'est de ce chapitre le théorème dont la démonstration est la plus délicate, même si l'énoncé est facile à comprendre :

Théorème III .21 (théorème de Banach de l'application ouverte) *Soient E et F des espaces de Banach. Soit $\varphi : E \to F$ une application linéaire continue et surjective. Alors φ est ouverte, c'est-à-dire que l'image par φ de tout ouvert de E est un ouvert de F.*

Ce théorème heurte nos habitudes qui sont que, par une application continue, l'image réciproque de tout ouvert est un ouvert ; il s'agit ici de l'image directe.

La linéarité de φ est une condition forte mais essentielle – s'en convaincre en dessinant un graphe judicieux d'une fonction réelle d'une variable réelle. Il est essentiel aussi que non seulement E, mais aussi F, soit complet. Par exemple, si $E = \ell^1(\mathbf{N})$ muni de sa norme $\|x\|_1 = \sum_{n=0}^\infty |x_n|$ et $F = \ell^1(\mathbf{N})$ muni de la norme $\|x\|_\infty = \max_{n \geqslant 0} |x_n|$ pour $x = (x_n)_{n \geqslant 0}$, alors l'application identique de E sur F est continue, surjective, et même bijective, mais sa réciproque φ^{-1} n'est pas continue : sinon, F serait complet comme E, ce qui n'est pas vu que F est dense et strictement inclus dans $\ell^\infty(\mathbf{N})$.

Avant de passer à la démonstration, voyons l'intérêt de ce théorème à travers ses corollaires.

Corollaire III .22 *Soient E et F des espaces de Banach. Soit $\varphi : E \to F$ une application linéaire continue et bijective. Alors φ^{-1} est une application linéaire continue de F sur E.*

Preuve. En effet, si U est un ouvert de E, son image réciproque par φ^{-1} n'est autre que $\varphi(U)$, qui est un ouvert de F par le théorème de Banach III .21 qui précède. ▫

Corollaire III .23 (théorème du graphe fermé) *Soient E et F des espaces de Banach. Soit $\varphi : E \to F$ une application linéaire. Supposons que, chaque fois qu'une suite $(x_n)_{n\geqslant 0}$ tend vers 0 dans E et que la suite image $(\varphi(x_n))_{n\geqslant 0}$ converge dans F, disons vers y, alors la limite y vaut 0. Alors l'application φ est continue.*

En appliquant l'hypothèse à la suite décalée des $x_n - x$ (pour $x \in E$), on voit qu'elle équivaut à la formulation en apparence plus générale suivante :

> chaque fois qu'une suite $(x_n)_{n\geqslant 0}$ converge dans E, disons vers x, et que la suite image $(\varphi(x_n))_{n\geqslant 0}$ converge dans F, disons vers y, alors la limite $y = \varphi(x)$,

qui équivaut encore au fait que

> le *graphe* $\mathrm{gr}(\varphi)$ de φ, c'est-à-dire le sous-ensemble de $E \times F$:
>
> $$\mathrm{gr}(\varphi) = \{\,(x, \varphi(x)) : x \in E\},$$
>
> est fermé dans $E \times F$.

Ainsi le corollaire III .23 énonce que pour qu'une application linéaire entre espaces de Banach soit continue, il faut et il suffit que son graphe soit fermé (la partie « condition nécessaire » est facile, par exemple par le critère séquentiel de continuité).

Bien entendu, ceci ne serait pas vrai si φ n'était pas linéaire. Par exemple, pour $E = F = \mathbf{R}$ et $\varphi(x) = \frac{1}{x}$ pour $x \neq 0$ et $\varphi(0) = 0$, le graphe de φ est fermé sans que la fonction soit continue.

Passons à la preuve du corollaire.

Preuve. L'espace $E \times F$, normé par

$$(x, y) \mapsto \|(x, y)\| = \max\{\|x\|_E ; \|x\|_F\},$$

est un espace de Banach (i.e., il est complet). Puisque φ est linéaire, on voit immédiatement que son graphe

$$\mathrm{gr}(\varphi) = \{\,(x, \varphi(x)) : x \in E\}$$

est un sous-espace vectoriel de $E \times F$. Par hypothèse, il est fermé donc complet : c'est un espace de Banach pour la norme induite. L'application de projection

$$\theta : (x, \varphi(x)) \mapsto x$$

est une application linéaire, continue et bijective de $\mathrm{gr}(\varphi)$ sur E. Par le corollaire III .22 son application réciproque θ^{-1} est continue ; autrement dit, $x \mapsto (x, \varphi(x))$ est continue. En composant cette application avec la seconde projection $\mathrm{pr}_F : E \times F \to F$, on voit que $\varphi = \mathrm{pr}_F \circ \theta^{-1}$ est continue, ce qu'il fallait démontrer. □

Exercice III .24 *Soit X un ensemble et soit $(E, \|\cdot\|)$ un espace de Banach de fonctions définies sur X. On suppose que la norme $\|\cdot\|$ vérifie :*

$$\lim_{n\to\infty} \|f_n\| = 0 \quad \Longrightarrow \quad \text{quel que soit } x \in X, \text{ on a : } \lim_{n\to\infty} f_n(x) = 0.$$

Soit φ une fonction sur X telle que $f \in E \Rightarrow \varphi f \in E$ (on dit que φ est un multiplicateur de E). Montrer que l'application $f \mapsto \varphi f$ est continue pour la norme de E.

Le corollaire III .22 est un moyen puissant pour démontrer que des espaces sont distincts ; faisons-le comprendre à nouveau à propos d'un problème concernant les séries de Fourier. Si $f \in \mathrm{L}^1([-\pi; \pi], \mathbf{C})$, soit $(c_n(f))_{n\in\mathbf{Z}}$ la suite de ses coefficients de Fourier :

$$c_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, \mathrm{d}x \quad (n \in \mathbf{Z})$$

Rappelons que $\lim_{n\to\pm\infty} c_n(f) = 0$; c'est le *lemme de Riemann-Lebesgue.* Le problème est le suivant :

Problème : Étant donnée une suite $(a_n)_{n\in\mathbf{Z}}$ telle que $\lim_{n\to\pm\infty} a_n = 0$, existe-t-il toujours une fonction $f \in \mathrm{L}^1([-\pi; \pi], \mathbf{C})$ telle qu'on ait $c_n(f) = a_n$ pour tout $n \in \mathbf{Z}$.

La réponse est négative. Pour le voir, notons $c_0(\mathbf{Z})$ l'espace de Banach des suites $(a_n)_{n\in\mathbf{Z}}$ telles que $\lim_{n\to\pm\infty} a_n = 0$, muni de la norme sup :

$$\|(a_n)_{n\in\mathbf{Z}}\| = \sup_{n\in\mathbf{Z}} |a_n|.$$

Appliquons le corollaire III .22 avec $E = \mathrm{L}^1([-\pi; \pi], \mathbf{C})$, $F = c_0(\mathbf{Z})$ et $\varphi : f \mapsto (c_n(f))_{n\in\mathbf{Z}}$. Elle est linéaire, continue (car $|c_n(f)| \leqslant \|f\|_1$) et injective (car une fonction dont les coefficients de Fourier sont tous nuls est nulle presque partout pour la mesure de Lebesgue, par la formule d'inversion de Fourier). Le problème demande si φ est surjective. Si elle l'était, le corollaire assurerait la continuité de $\varphi^{-1} : c_0(\mathbf{Z}) \to \mathrm{L}^1([-\pi;\pi], \mathbf{C})$, i.e. l'existence d'une constante $M > 0$ telle que

$$(*) \quad \|f\|_1 \leqslant M\|(c_n(f))_{n\in\mathbf{Z}}\|_\infty$$

pour toute $f \in \mathrm{L}^1([-\pi;\pi], \mathbf{C})$. Dans cette inégalité prenons pour f les noyaux de Dirichlet

$$D_N(x) = \sum_{k=-N}^{N} e^{ikx}.$$

Les $c_n(D_N)$ valent tous 0 ou 1, donc $\|(c_n(D_N))_{n\in\mathbf{Z}}\|_\infty = 1$ pour tout $n \in \mathbf{Z}$, alors qu'on a vu au lemme III .14 que

$$\lim_{N\to\infty} \|D_N\|_{\mathrm{L}^1([-\pi;\pi])} = +\infty.$$

C'est évidemment en contradiction avec les inégalités $(*)$ pour les fonctions $f = D_N$ avec $n \geqslant 1$.

Exercice III .25 *Si $f \in \mathrm{L}^1(\mathbf{R}, \mathrm{d}x)$, sa transformée de Fourier*

$$\hat{f}(t) = \int_{\mathbf{R}} f(x)e^{-itx}\,\mathrm{d}x$$

appartient à l'espace $\mathcal{C}_0(\mathbf{R})$ des fonctions continues sur $\mathbf{R}$ de limite 0 en $\pm\infty$. Montrer qu'il existe des $g \in \mathcal{C}_0(\mathbf{R})$ qui ne sont pas des transformées de fonctions dans $\mathrm{L}^1(\mathbf{R}, \mathrm{d}x)$.

Nous allons maintenant passer à la preuve (difficile) du théorème de Banach de l'application ouverte III .21.

Preuve. Soit U (resp. V) l'ensemble des $x \in E$ (resp. $y \in F$) tels que $\|x\| < 1$ (resp. $\|y\| < 1$). Si $A \subset E$ (resp. $A \subset F$) et si λ est un scalaire, on note λA l'ensemble des λx pour $x \in A$.

– Première étape : il existe $\delta > 0$ tel que pour tout $y \in F$ et tout $\varepsilon > 0$, il existe $x \in E$ tel que $\|x\| \leqslant \frac{1}{\delta}\|y\|$ et $\|y - \varphi(x)\| \leqslant \varepsilon$.

L'application φ étant surjective, l'espace complet F est la réunion de la suite de fermés $\overline{\varphi(kU)}$ où k parcourt les entiers $\geqslant 1$. D'après le théorème de Baire III .4 l'un au moins de ces fermés est d'intérieur non vide. Donc il existe un entier $k \geqslant 1$, un point y_0 dans F et un nombre réel $\eta > 0$ tels que pour tout $y \in F$ vérifiant $\|y\| < \eta$, on ait $y_0 + y \in \overline{\varphi(kU)}$. Posons $\delta = \frac{\eta}{2k}$. Par homothétie, on se ramène à vérifier l'assertion de la première étape pour les y d'une sphère, par exemple ceux vérifiant $\|y\| = \eta$. Pour un tel y, il existe des suites $(x'_i)_{i\geqslant 1}$ et $(x''_i)_{i\geqslant 1}$ telles que $\lim_{i\to\infty} \varphi(x'_i) = y_0$ et $\lim_{i\to\infty} \varphi(x''_i) = y_0 + y$. Posons $x_i = x'_i - x''_i$. Alors $\|x_i\| < 2k = \delta^{-1}\|y\|$ et $\lim_{i\to\infty} \varphi(x_i) = y$.

– Deuxième étape : pour tout $\alpha > 0$, on a : $\delta V \subset \varphi\left((1+\alpha)U\right)$.

Soit $y \in \delta V$. D'après la première étape, il existe $x_1 \in E$ tel que $\|x_1\| < 1$ et $\|y - \varphi(x_1)\| \leqslant \frac{\delta\alpha}{2}$. Par hypothèse de récurrence, supposons que soient déjà choisis $x_1, x_2, \dots x_n$ tels que

$$\|x_j\| < \frac{\alpha}{2^n} \quad \text{et} \quad \|y - \varphi(x_1) - \cdots - \varphi(x_j)\| \leqslant \frac{\delta\alpha}{2^j}.$$

Alors, en utilisant la première étape, où y est remplacé par $y - \varphi(x_1) - \cdots - \varphi(x_n)$, on obtient un $x_{n+1} \in E$ tel que

$$\|x_{n+1}\| < \frac{\alpha}{2^{j-1}} \quad \text{et} \quad \|y - \varphi(x_1) - \cdots - \varphi(x_{n+1})\| \leqslant \frac{\delta\alpha}{2^{n+1}}.$$

On obtient ainsi une suite $(x_n)_{n\geqslant 1}$ dans E et, en posant $s_n = \sum_{k=1}^{n} x_k$, on voit que la suite $(s_n)_{n\geqslant 1}$ est de Cauchy dans E, donc converge vers un $x \in E$. Puisque $\|x_n\| < \frac{\alpha}{2^{n-1}}$, on a $\|x\| < 1 + \alpha$. Enfin, par construction on a : $\lim\limits_{n\to\infty} \varphi(s_n) = y$, et donc $y = \varphi(x)$. Ainsi : $y \in \varphi\left((1+\alpha)U\right)$.

– Troisième étape : pour tout $\alpha > 0$, on a : $\delta V \subset \varphi(U)$.

En effet, d'après la deuxième étape, on a : $\varphi(U) \supset \displaystyle\bigcup_{\alpha>0} \frac{\delta}{1+\alpha} V = \delta V$.

– Quatrième étape : pour tout ouvert O de E, l'image $\varphi(O)$ est ouverte dans F.

Soit $y_0 \in \varphi(O)$ et soit $x_0 \in E$ tel que $\varphi(x_0) = y_0$. Puisque O est ouvert, il existe $r > 0$ tel que la boule $x_0 + rU$ soit contenue dans O. Alors, d'après la linéarité de φ, l'ensemble $\varphi(O)$ contient $\varphi(x_0 + rU) = y_0 + r\varphi(U)$. Donc, d'après la troisième étape, $\varphi(O)$ contient la boule de centre y_0 et de rayon $r\delta$. Comme y était quelconque, ceci prouve que $\varphi(O)$ est un voisinage de tous ses points, donc est ouvert dans F. □

Quatrième partie

Algèbres de Banach

La théorie que nous allons présenter est récente (1941) et presque toute entière due au mathématicien I. M. Gelfand. Dans un même moule, celui des algèbres de Banach, elle regroupe des algèbres aussi différentes que :

- l'algèbre $\mathcal{C}(X)$ des fonctions continues sur un compact X ;
- des algèbres de convolution ;
- des algb̀res de fonctions holomorphes ;
- des algèbres (non commutatives) d'opérateurs dans un espace de Banach.

Les étapes principales de la théorie sont (pour une algèbre de Banach A) :

- l'étude du spectre d'un élément $x \in A$, le fait fondamental étant qu'il n'est jamais vide ;
- le théorème de Gelfand-Mazur ;
- l'holomorphie, sur l'ouvert complémentaire du spectre, de la fonction résolvante $\lambda \mapsto (\lambda e - x)^{-1}$;
- l'étude de l'exponentiation dans A.

Si plus particulièrement, l'algèbre de Banach A est commutative, la théorie se poursuit avec :

- l'identification de l'espace des idéaux maximaux de A avec l'ensemble $\widehat{A}$ des formes linéaires non nulles multiplicatives (ou *caractères*) de A ;
- la transformation de Gelfand de A dans $\mathcal{C}(\widehat{A})$, dont un cas particulier est la transformation de Fourier ;
- la re-démonstration, étonnamment simple, d'un théorème de Wiener (1932).

IV.19 Notion d'algèbre de Banach

On appelle *algèbre de Banach* un espace de Banach $A \neq \{0\}$ sur $\mathbf{C}$ muni d'une multiplication $(x, y) \mapsto xy$ de $A \times A$ dans A telles que, pour cette multiplication et l'addition vectorielle, A soit un anneau, et telle que, quels que soient $x \in A$, $y \in A$ et $\lambda \in \mathbf{C}$, on ait :

$$\lambda(xy) = (\lambda x)y = x(\lambda y) \quad \text{et} \quad \|xy\| \leqslant \|x\|\|y\|.$$

Si l'anneau A a un élément neutre e pour la multiplication, on exige de plus que $\|e\| = 1$ et on dit alors que A est une *algèbre de Banach avec unité*. On dit que A est *commutative* si $xy = yx$ pour tous x et y dans A.

Exercice IV .1 *Montrer que l'application multiplication d'une algèbre de Banach est continue.*

La théorie axiomatise dans un seul moule structurel des exemples concrets très divers. En voici quelques-uns.

Exemple IV .2 *L'algèbre $A = \mathbf{C}$ des nombres complexes.*

Exemple IV .3 *Soit X un espace compact non vide et $A = \mathcal{C}(X) = \mathcal{C}(X, \mathbf{C})$ l'algèbre des fonctions continues à valeurs complexes sur X. Pour la norme uniforme $\|\cdot\|_\infty$ donnée par $\|f\|_\infty = \sup_{x\in X} |f(x)|$, et pour le produit (point par point) ordinaire donné par $(fg)(x) = f(x)g(x)$ pour tout $x \in X$, l'algèbre A est une algèbre de Banach commutative, qui a une unité, à savoir la fonction constante égale à 1.*

Exemple IV .4 *Soit $U = \{z \in \mathbf{C} : |z| < 1\}$ et $\overline{U} = \{z \in \mathbf{C} : |z| \leqslant 1\}$. On prend $A = \mathcal{OC}(\overline{U})$ l'algèbre des fonctions $f = f(z)$ continues sur $\overline{U}$ et holomorphes dans U. Pour la norme uniforme $\|\cdot\|_\infty$ donnée par $\|f\|_\infty = \sup_{z\in\overline{U}} |f(z)|$, et pour le produit (point par point) ordinaire donné par $(fg)(z) = f(z)g(z)$ pour tout $z \in \overline{U}$, l'algèbre A est une algèbre de Banach commutative, qui a une unité, à savoir la fonction constante égale à 1.*

Exemple IV .5 *Soit $\mathcal{D}_m([0;1])$ la sous-algèbre de $\mathcal{C}([0;1])$ constituée des fonctions f qui sont m fois dérivables. Pour la norme $\|f\| = \displaystyle\sum_{j=0}^{m} \frac{1}{j!}\|f^{(j)}\|$, l'algèbre A est une algèbre de Banach commutative avec unité.*

Exemple IV .6 *Soit $A = L^1(\mathbf{R})$ l'algèbre des (classes de) fonctions intégrables au sens de Lebesgue sur la droite réelle. Pour la norme $\|f\|_1 = \displaystyle\int_{\mathbf{R}} |f(x)|\ \mathrm{d}x$, et pour le produit de convolution $*$ donné par :*

$$(f * g)(x) = \int_{\mathbf{R}} f(x-t)g(t)\ \mathrm{d}t$$

pour toutes $f, g \in A$ et tout $x \in \mathbf{R}$, l'algèbre A est une algèbre de Banach commutative.

Exemple IV .7 *Soit $A = \ell^1(\mathbf{Z})$ l'algèbre des suites $x = (x_n)_{n\in\mathbf{Z}}$ de nombres complexes indexées par les entiers relatifs. Pour la norme $\|x\|_1 = \displaystyle\sum_{n\in\mathbf{Z}} |x_n|$, et pour le produit de convolution $*$ donné par :*

$$(x * y)_n = \sum_{p\in\mathbf{Z}} x_{n-p}y_p,$$

pour toutes suites $x = (x_n)_{n\in\mathbf{Z}}$ et $y = (y_n)_{n\in\mathbf{Z}}$ dans A et tout indice $n \in \mathbf{Z}$, l'algèbre A est une algèbre de Banach commutative. Elle admet une unité : la suite e_0 qui vaut 1 en 0 et 0 ailleurs.

Exemple IV .8 *Soit E un espace de Banach non nul. On prend pour A l'algèbre $\mathcal{L}(E,E)$ des applications linéaires continues (ou opérateurs bornés) φ de E dans E. Pour la norme d'opérateur donnée par*

$$|||\varphi||| = \sup_{\|x\|_E \leqslant 1} \|\varphi(x)\|_E,$$

et pour la multiplication donnée par la composition $\circ$, l'algèbre A est une algèbre de Banach, avec pour unité l'application identique id_E de E.

Exemple IV .9 *Munissons $\mathbf{C}^n$ d'une norme vectorielle $\|\cdot\|$. Prenons $A = M_n(\mathbf{C})$ l'algèbre des matrices $M = [m_{ij}]_{1 \leqslant i,j \leqslant n}$ de taille $n \times n$ à coefficients complexes. Pour la norme*

$$|||M||| = \sup_{x \in \mathbf{C}^n, \|x\|=1} \|Mx\|,$$

et pour le produit usuel des matrices, l'algèbre A est une algèbre de Banach, avec pour unité la matrice identité I_n de taille $n \times n$.

Toutes ces algèbres de Banach sont commutatives, sauf les deux derniers exemples. Elles sont toutes avec unité sauf $A = \mathrm{L}^1(\mathbf{R})$.

Exemple IV .10 *Soit A une algèbre de Banach sans unité. On pose $\widetilde{A} = A \times \mathbf{C}$ avec la norme $\|(x,\lambda)\| = \|x\| + |\lambda|$ et le produit*

$$(x,\lambda)(y,\mu) = (xy + \lambda x + \mu y, \lambda\mu).$$

Alors $\widetilde{A}$ est une algèbre de Banach, avec unité égale à $(0_A, 1)$ et dite déduite de A par adjonction d'un élément neutre.

Exercice IV .11 *Vérifier que tous les exemples ci-dessus fournissent bien des algèbres de Banach.*

IV.20 Étude du groupe des inversibles

Notons dans toute la suite G le groupe des éléments inversibles de l'algèbre de Banach A, supposée avec élément unité e dans toute la suite, sauf mention expresse du contraire. Autrement dit, G est l'ensemble des $x \in A$ tels qu'il existe $y \in A$ pour lequel $xy = yx = e$ (auquel cas on notera $y = x^{-1}$ par unicité de l'inverse). L'ensemble G est un groupe pour la multiplication dans A, car si $x \in G$ et $y \in G$ alors $(xy)^{-1} = y^{-1}x^{-1}$ et en outre : $(x^{-1})^{-1} = x$ et $e^{-1} = e$.

Théorème IV .12 *Soit $y \in A$ tel que $\|y\| < 1$. Alors $e + y$ est inversible dans A et $(e+y)^{-1} = \sum_{n=0}^{\infty} (-1)^n y^n$: de plus :*

$$\|(e+y)^{-1} - e + y\| \leqslant \frac{\|y\|^2}{1 - \|y\|}.$$

Plus généralement, soit x inversible dans A. On pose $\alpha = \frac{1}{\|x^{-1}\|}$. Alors pour tout $h \in A$ tel que $\|h\| = \beta < \alpha$, l'élément $x + h$ est inversible dans A et on a :

$$\|(x+h)^{-1} - x^{-1} + x^{-1}hx^{-1}\| \leqslant \frac{\beta^2}{\alpha^2(\alpha - \beta)}.$$

Preuve. La série $\sum_{n=0}^{\infty} (-1)^n y^n$ converge absolument dans l'espace de Banach A car

$$\sum_{n=0}^{\infty} \|(-1)^n y^n\| \leqslant \sum_{n=0}^{\infty} \|y\|^n \quad \text{et} \quad \|y\| < 1.$$

En effet, l'axiome de sous-multiplicativité de la norme implique $\|y^n\| \leqslant \|y\|^n$ par récurrence sur n. Posons

$$s = \sum_{n=0}^{\infty} (-1)^n y^n \quad \text{et} \quad s_N = \sum_{n=0}^{N} (-1)^n y^n$$

pour chaque entier $n \geqslant 0$. Alors

$$\begin{array}{rcl} (e+y)s_N & = & (e+y)(e - y + y^2 - y^3 + \cdots + (-1)^N y^N) \\ & = & e + y - y + y^2 - y^2 + y^3 - y^3 + \cdots + (-1)^N y^N + (-1)^N y^{N+1} \\ & = & e + (-1)^N y^{N+1} \\ & = & s_N(e+y) \end{array}$$

tend vers $(e+y)s = e = s(e+y)$ quand $N \to \infty$, et donc $s = (e+y)^{-1}$. De plus

$$(e+y)^{-1} - e + y = \sum_{n \geqslant 2} (-1)^n y^n = y^2 \sum_{k=0}^{\infty} (-1)^k y^k,$$

donc

$$\|(e+y)^{-1} - e + y\| \leqslant \|y^2\| \|\sum_{k=0}^{\infty} (-1)^k y^k\| \leqslant \|y\|^2 \sum_{k=0}^{\infty} \|y\|^k = \frac{\|y\|^2}{1 - \|y\|}.$$

Passons maintenant à l'énoncé général. On a :

$$\|x^{-1}h\| \leqslant \|x^{-1}\| \|h\| = \frac{\beta}{\alpha},$$

donc d'après ce qu'on vient de faire $e+x^{-1}h$ est inversible. Par suite $x+h = x(e+x^{-1}h)$ est inversible comme produit de deux inversibles et $(x+h)^{-1} = (e+x^{-1}h)^{-1}x^{-1}$. Faisons $y = x^{-1}h$ dans l'inégalité qui précède :

$$\|(e+x^{-1}h)^{-1} - e + x^{-1}h\| \leqslant \frac{\|x^{-1}h\|^2}{1-\|x^{-1}h\|}.$$

En multipliant par x^{-1} l'argument de la norme du membre de gauche, il vient par sous-multiplicativité :

$$\begin{aligned} \|(x+h)^{-1} - x^{-1} + x^{-1}hx^{-1}\| &\leqslant \tfrac{\|x^{-1}\|\|x^{-1}h\|^2}{1-\|x^{-1}h\|} \\ &\leqslant \tfrac{\|x^{-1}\|^3\|h\|^2}{1-\|x^{-1}\|\|h\|} \\ &= \tfrac{\beta^2}{\alpha^3(1-\frac{\beta}{\alpha})} \\ &= \tfrac{\beta^2}{\alpha^2(\alpha-\beta)}, \end{aligned}$$

ce qui conclut la démonstration. □

Corollaire IV .13 *L'ensemble G des éléments inversibles de l'algèbre de Banach avec unité A est ouvert dans A et l'application inverse $x \mapsto x^{-1}$ est une application continue. C'est même une application différentiable dans G, sa différentielle au point $x \in G$ étant l'application linéaire de A dans A définie par $h \mapsto -x^{-1}hx^{-1}$.*

Preuve. L'énoncé général du théorème précédent dit que si $x^{-1} \in G$, et si $\alpha = \|x^{-1}\|^{-1}$, la boule ouverte de centre x^{-1} et de rayon α est incluse dans G. De plus, l'inégalité de ce même énoncé précise que dès que $\|h\| = \beta < \frac{\alpha}{2}$, on a :

$$\|(x+h)^{-1} - x^{-1} + x^{-1}hx^{-1}\| \leqslant \frac{2}{\alpha^3}\|h\|^2,$$

ce qui vérifie l'assertion sur la différentielle. □

On notera que l'assertion sur la différentielle est une géralisation du fait que la dérivée dans $\mathbf{C}^\times$ de la fonction $z \mapsto \frac{1}{z}$ est $z \mapsto \frac{-1}{z^2}$.

Exercice IV .14 *1. Montrer que, dans $M_n(\mathbf{C})$, toute matrice est limite d'une suite de matrices inversibles; autrement dit, que $G = \mathrm{GL}_n(\mathbf{C})$ est dense dans $A = M_n(\mathbf{C})$.*

On va voir que ce fait ne subsiste pas en dimension infinie.

2. Soit en effet $\ell^2(\mathbf{N})$ l'espace des suites complexes $x = (x_n)_{n\geqslant 0}$ de carré intégrable, i.e. telles que $\|x\| = \sum_{n=0}^{\infty} |x_n|^2 < +\infty$. On note $A = \mathcal{L}\left(\ell^2(\mathbf{N}), \ell^2(\mathbf{N})\right)$. Soit S l'opérateur de décalage, « shift » en bon français, défini par :

$$S(x_0, x_1, s_2, \ldots, x_n, \ldots) = (0, x_0, x_1, s_2, \ldots, x_n, \ldots).$$

— *Remarquer que S est isométrique, mais non surjectif, donc n'est pas inversible dans l'algèbre de Banach A.*
— *Par l'absurde, montrer que S n'est pas limite dans A d'une suite $(T_n)_{n\geqslant 0}$ d'opérateurs bornés inversibles. Indication : montrer que la suite $(|||T|||_n)_{n\geqslant 0}$ n'est pas bornée et en déduire qu'il existerait une suite $(U_n)_{n\geqslant 0}$ dans A telle que $|||U|||_n = 1$ pour tout $n \geqslant 0$ et $\lim\limits_{n\to\infty} |||SU_n||| = 0$.*

IV.21 Spectre d'un élément

Soit $x \in A$. On appelle *spectre* de x (dans A), et l'on note $\mathrm{sp}(x)$ ou $\mathrm{sp}_A(x)$, l'ensemble des $\lambda \in \mathbf{C}$ tels que $x - \lambda e$ n'est pas inversible dans A.

Ainsi $\mathrm{sp}(x)$ est une partie de $\mathbf{C}$. Par exemple, le spectre d'une matrice M dans $M_n(\mathbf{C})$ est l'ensemble (fini) de ses valeurs propres complexes. Le spectre d'une fonction $f \in \mathcal{C}(X)$ est l'ensemble $f(X)$ des valeurs prises par f.

Proposition IV .15 *Pour tout $x \in A$, l'ensemble $\mathrm{sp}(x)$ est un compact de $\mathbf{C}$, contenu dans le disque centré à l'origine et de rayon $\|x\|$.*

Preuve. Supposons $|\lambda| > \|x\|$. Alors $\lambda \neq 0$ et $x - \lambda e = -\lambda(e - \frac{x}{\lambda})$ est inversible dans A car $\|\frac{x}{\lambda}\| < 1$. Donc $\lambda \notin \mathrm{sp}(x)$. Ceci prouve que $\mathrm{sp}(x)$ est contenu dans le disque $\{\lambda \in \mathbf{C} : |\lambda| \leqslant \|x\|\}$: en particulier $\mathrm{sp}(x)$ est une partie bornée de $\mathbf{C}$. Par Borel-Lebesgue, il reste à voir que c'est une partie fermée de $\mathbf{C}$, c'est-à-dire que son complémentaire est ouvert. Soit $\lambda_0 \notin \mathrm{sp}(x)$. Alors $x - \lambda_0 e$ est inversible dans A, donc $x - \lambda e$ aussi dès que $|\lambda - \lambda_0|$ est assez petit, car G est ouvert (Théorème IV .12). $\square$

Exercice IV .16 *Soit S l'opérateur de décalage défini dans l'Exercice IV .14. Montrer que le spectre de S dans $A = \mathcal{L}\left(\ell^2(\mathbf{N}), \ell^2(\mathbf{N})\right)$ est exactement le disque unité fermé de $\mathbf{C}$.*

Exercice IV .17 *Soit $A = \ell^1(\mathbf{Z})$ l'algèbre de Banach de l'Exemple IV .7. Pour tout $n \in \mathbf{Z}$, soit e_n l'élément de A qu'est la suite dont tous les termes sont nuls sauf le n-ième, qui vaut 1.*

1. *Vérifier la formule de covolution $e_m * e_n = e_{m+n}$ pour tous m et n dans $\mathbf{Z}$, puis que $(e_n)^{-1} = e_{-n}$ et $e_n = (e_1)^n$ pour tout n dans $\mathbf{Z}$.*
2. *Dans $A = \ell^1(\mathbf{Z})$, tout élément $x = (x_n)_{n\in\mathbf{Z}}$ vérifie :*

$$x = \lim_{N\to\infty} \sum_{n=-N}^{N} x_n e_n = \lim_{N\to\infty} \sum_{n=-N}^{N} x_n (e_1)^n.$$

3. *Montrer que le spectre de e_1 dans A est le cercle unité de $\mathbf{C}$.*

Théorème IV .18 *Soit $x \in A$. Soit φ une forme linéaire continue de l'espace de Banach A. Alors la fonction f d'une variable complexe*

$$\lambda \mapsto f(\lambda) = \varphi\left((x-\lambda e)^{-1}\right)$$

est holomorphe dans l'ouvert $\mathbf{C} - \mathrm{sp}_A(x)$. De plus, la fonction $\lambda \mapsto \lambda f(\lambda)$ est bornée quand $|\lambda| \to \infty$.

Preuve. Soit Ω l'ouvert $\mathbf{C} - \mathrm{sp}_A(x)$ et soit $\lambda \in \Omega$ fixé. Pour $\mu \in \mathbf{C}$, assez proche de λ pour qu'on ait :

$$\mu \in \Omega \quad \text{et} \quad |\lambda - \mu| < \frac{1}{2}\|(x-\lambda e)^{-1}\|,$$

écrivons l'inégalité du cas général du Théorème IV .12, en y remplaçant x par $x - \lambda e$ et h par $(\lambda-\mu)e$, et donc $x+h$ par $x-\mu e$. On obtient :

$$\|(x-\mu e)^{-1} - (x-\lambda e)^{-1} + (\lambda-\mu)(x-\lambda e)^{-2}\| \leqslant C|\lambda-\mu|^2,$$

où C est une constante (pour x et λ fixés). Par conséquent :

$$\lim_{\mu\to\lambda, \mu\neq\lambda} \frac{(x-\mu e)^{-1} - (x-\lambda e)^{-1}}{\mu-\lambda} = (x-\lambda e)^{-2}$$

dans l'espace de Banach A. En appliquant la forme linéaire continue φ à ce calcul de limite, on obtient :

$$\lim_{\mu\to\lambda, \mu\neq\lambda} \frac{f(\mu)-f(\lambda)}{\mu-\lambda} = \varphi\left((x-\lambda e)^{-2}\right).$$

Ainsi f a une dérivée première continue dans Ω : elle est holomorphe dans Ω. En outre :

$$\lambda f(\lambda) = \varphi\left(\lambda(x-\lambda e)^{-1}\right) = \varphi\left(\left(\frac{x}{\lambda} - e\right)^{-1}\right)$$

tend vers $\varphi(-e)$ quand $|\lambda| \to \infty$, donc est bornée. □

On peut en déduire le résultat fondamental suivant.

Corollaire IV .19 *Soit $x \in A$. Alors le spectre de x dans A n'est pas vide.*

Preuve. On procède par l'absurde, en supposant $\mathrm{sp}_A(x)$ vide. Alors f serait holomorphe sur $\mathbf{C}$ tout entier (autrement dit, f serait une fonction *entière*) et bornée (car $\lambda \mapsto \lambda f(\lambda)$ est bornée à l'infini). D'après le théorème de Liouville, ces deux propriétés imposent à f d'être constante sur $\mathbf{C}$, et même nulle (à nouveau car $\lambda \mapsto \lambda f(\lambda)$ est bornée à l'infini). Or, soit λ_0

fixé dans $\mathbf{C}$. Comme $\mathrm{sp}_A(x) = \varnothing$, on dispose de l'inverse $(x - \lambda_0 e)^{-1}$, qui est un vecteur non nul de A. Par Hahn-Banch, il existe $\varphi \in A'$ telle que $\varphi\left(\lambda(x - \lambda_0 e)^{-1}\right) \neq 0$. Pour une telle forme linéaire continue φ, la fonction holomorphe associée comme ci-dessus serait non nulle, ce qui est une contradiction. □

Corollaire IV .20 (Gelfand-Mazur) *Si l'anneau A est un corps, alors $A = \mathbf{C}e$; autrement dit A est isomorphe au corps des nombres complexes.*

On notera que A n'est pas supposée commutative au départ.

Preuve. Il s'agit de voir que pour tout $x \in A$, il existe $\lambda \in \mathbf{C}$ tel que $x = \lambda e$. Sinon, on aurait $(x - \lambda e) \neq 0$ pour tout $\lambda \in \mathbf{C}$ et donc $x - \lambda e$ serait inversible puisque A est supposé être un corps : contradiction avec le fait que $\mathrm{sp}_A(x) \neq \varnothing$. □

Nous passons maintenant à quelques énoncés qui facilitent les calculs de spectre.

Déjà, pour tout $x \in A$ et tout $\lambda \in \mathbf{C}$, il est clair que :

$$\mathrm{sp}_A(\lambda e) = \{\lambda\} \quad \text{et} \quad \mathrm{sp}_A(x + \lambda e) = \mathrm{sp}_A(x) + \lambda.$$

Théorème IV .21 *Soit P un polynôme à coefficients complexes. Alors*

$$\mathrm{sp}_A\left(P(x)\right) = P\left(\mathrm{sp}_A(x)\right).$$

Plus généralement, si Q est un second polynôme à coefficients complexes, tel que $Q(x)$ est inversible dans A, alors en notant R la fraction rationnelle $\frac{P}{Q}$ et $R(x) = P(x)Q(x)^{-1} = Q(x)^{-1}P(x)$, on a :

$$0 \notin Q\left(\mathrm{sp}_A(x)\right) \quad \text{et} \quad \mathrm{sp}_A\left(R(x)\right) = R\left(\mathrm{sp}_A(x)\right).$$

Preuve. Si P est constant c'est vrai car $\mathrm{sp}_A(\lambda e) = \{\lambda\}$. On suppose désormais que P est non constant. Soit $\mu \in \mathbf{C}$. On écrit :

$$P(X) - \mu = \alpha_n(X - \lambda_1)(X - \lambda_2)\ldots(X - \lambda_n),$$

où $\alpha_n \in \mathbf{C}^\times$ et où les λ_i dépendent de μ. L'image réciproque par P de μ est $P^{-1}(\{\mu\}) = \{\lambda_1; \lambda_2; \ldots \lambda_n\}$, et l'on a : $P(x) - \mu = \alpha_n(x - \lambda_1 e)(x - \lambda_2 e)\ldots(x - \lambda_n e)$. Cette dernière égalité, et le fait que les facteurs commutent, permet de voir que $P(x) - \mu$ est inversible si, et seulement si, chaque facteur $x - \lambda_i e$ l'est. Ainsi :

$$\begin{aligned}
\mu \in \mathrm{sp}_A\left(P(x)\right) &\iff \lambda_i \in \mathrm{sp}_A(x) \text{ pour au moins un indice } i \in \{1; 2; \ldots n\} \\
&\iff \mathrm{sp}_A(x) \cap P^{-1}(\{\mu\}) \neq \varnothing \\
&\iff \mu \in P\left(\mathrm{sp}_A(x)\right),
\end{aligned}$$

ce qui prouve la première assertion.

Passons maintenant au cas général. Remarquons d'abord que $PQ = QP$, et donc que $P(x)Q(x) = Q(x)P(x)$, ce qui permet d'écrire aussi $P(x)Q(x)^{-1} = Q(x)^{-1}P(x)$ en multipliant à gauche et à droite par $Q(x)^{-1}$ l'égalité précédente. On note $R(x) = P(x)Q(x)^{-1} = Q(x)^{-1}P(x)$. Puisque $Q(x)$ est inversible, on a bien sûr $0 \notin \mathrm{sp}_A\,(Q(x)) = Q\,(\mathrm{sp}_A(x))$. Il reste donc à prouver la dernière formule $\mathrm{sp}_A\,(R(x)) = R\,(\mathrm{sp}_A(x))$. Soit $\mu \in \mathbf{C}$. Posons $R - \mu = S = \frac{P_1}{Q}$ où $P_1 = P - \mu Q$. Alors on a les équivalences logiques suivantes :

$$\begin{array}{lclcl}
\mu \notin \mathrm{sp}_A\,(R(x)) & \Longleftrightarrow & R(x) - \mu e \,\text{est inversible} & \Longleftrightarrow & S(x)\,\text{est inversible} \\
& \Longleftrightarrow & P_1(x)\,\text{est inversible} & \Longleftrightarrow & 0 \notin \mathrm{sp}_A\,(P_1(x)) \\
& \Longleftrightarrow & 0 \notin P_1\,(\mathrm{sp}_A(x)) & \Longleftrightarrow & 0 \notin S\,(\mathrm{sp}_A(x)) \\
& \Longleftrightarrow & 0 \notin R\,(\mathrm{sp}_A(x)) - \mu & \Longleftrightarrow & \mu \notin R\,(\mathrm{sp}_A(x))
\end{array}$$

ce qui prouve la dernière assertion. □

Corollaire IV .22 *Supposons x inversible et $\|x\| = \|x^{-1}\| = 1$. Alors tout $\lambda \in \mathrm{sp}_A(x)$ est de module 1.*

Preuve. En effet, en prenant $R(X) = \frac{1}{X}$ dans le théorème qui précède, on obtient que $\mathrm{sp}_A(x^{-1}) = \mathrm{sp}_A(x)^{-1}$. Mais $\mathrm{sp}_A(x)$ est contenu dans le disque $\{|\lambda| \leqslant 1\}$ car $\|x\| = 1$, et $\frac{1}{\mathrm{sp}_A(x)} = \mathrm{sp}_A(x^{-1})$ est aussi contenu dans ce disque car $\|x^{-1}\| = 1$. □

Exercice IV .23 *(suite de l'Exercice IV .17). Soient $c_0, c_1, \dots c_k$ des nombres complexes tels que $P(\lambda) = c_0 + c_1\lambda + c_2\lambda^2 + \dots c_k\lambda^k$ soit non nul pour λ sur le cercle unité. Montrer que l'élément $x = (x_n)_{n\in\mathbf{Z}} \in \ell^1(\mathbf{Z})$ tel que $x_i = c_i$ pour $i \in \{0; 1; 2; \dots k\}$ et $x_i = 0$ ailleurs, est inversible dans $\ell^1(\mathbf{Z})$. Calculer l'inverse de x pour les choix $c_0 = -24$, $c_1 = 26$, $c_2 = -9$ et $c_3 = 1$ (avec $k = 3$). Indication : on remarquera que le polynôme $P(\lambda) = -24 + 26\lambda - 9\lambda^2 + \lambda^3$ a pour racine évidente 2, et on décomposera la fraction rationnelle $\frac{1}{P(X)}$ en éléments simples.*

Voyons maintenant le comportement du spectre par *changement d'algèbre.*

Proposition IV .24 *Sous B une sous-algèbre fermé de A telle que $e \in B$. Alors :*

(i) *on a $\mathrm{sp}_B(x) \supset \mathrm{sp}_A(x)$, mais :*
(ii) *la frontière de $\mathrm{sp}_B(x)$ est contenue dans $\mathrm{sp}_A(x)$,*
(iii) *et si $\mathrm{sp}_B(x)$ est d'intérieur vide (par exemple parce qu'il est réel), alors on a : $\mathrm{sp}_B(x) = \mathrm{sp}_A(x)$.*

Preuve. Le point (iii) est une conséquence immédiate de (i) et (ii), car alors l'ensemble $\mathrm{sp}_B(x)$ est égal à sa frontière. Le point (i) est évident car si $x - \lambda e$ est inversible dans B, il l'est dans A.

Prouvons le point (ii). Soit $\lambda_0 \in \mathrm{Fr}\,(\mathrm{sp}_B(x))$. On a en particulier $\lambda_0 \in \mathrm{sp}_B(x)$ puisque les spectres sont des ensembles fermés. Il existe dans $\mathbf{C}$ une suite $(\lambda_n)_{n\geqslant 0}$ de limite λ_0 et telle qu'aucun λ_n ne soit dans $\mathrm{sp}_B(x)$. Pour tout $n \geqslant 1$, l'inverse $(x-\lambda e)^{-1}$ existe dans B, donc a fortiori dans A. Supposons que $x-\lambda_0 e$ ait un inverse dans A ; alors par continuité des opérations dans A, la suite des éléments $(x-\lambda_n e)^{-1}$ de B convergerait dans A vers $(x-\lambda_0 e)^{-1}$. Puisque B est supposée fermée, cela impliquerait que $(x-\lambda_0 e)^{-1} \in B$, en contradiction avec $\lambda_0 \in \mathrm{sp}_B(x)$. □

Exercice IV .25 *Soit* $A = \ell^1(\mathbf{Z})$ *de l'Exemple IV .7, et soit*

$$B = \{x = (x_n)_{n\geqslant 0} \in A : x_n = 0 \text{ pour } n < 0\}.$$

Montrer que B *est une sous-algèbre fermée de* A*, contenant l'élément neutre* e_0*. On a déterminé* $\mathrm{sp}_A(e_1)$ *dans l'Exercice IV .17; déterminer* $\mathrm{sp}_B(e_1)$*.*

Exercice IV .26 *Soient* $U = \{z \in \mathbf{C} : |z| < 1\}$, $\overline{U} = \{z \in \mathbf{C} : |z| \leqslant 1\}$ *et* $\mathbb{T} = \{z \in \mathbf{C} : |z| = 1\}$*. On note* $A = \mathcal{C}(\mathbb{T})$ *et* B *la sous-algèbre des fonctions de* A *qui se prolongent à* $\overline{U}$ *et qui sont holomorphes sur* U*. Soit* f *la fonction identique sur* $\mathbb{T}$*, i.e.* $f(z) = z$*. Déterminer* $\mathrm{sp}_A(f)$ *et* $\mathrm{sp}_B(f)$*.*

Exercice IV .27 *Soit* λ_0 *un point frontière de* $\mathrm{sp}_A(x)$*. Montrer que quand* $\lambda \notin \mathrm{sp}_A(x)$ *tend vers* λ_0*, on a* $\lim \|(x-\lambda e)^{-1}\| = +\infty$*.*

IV.22 Formule du rayon spectral

Rappel technique sur les limites de suites.

Soit $\overline{\mathbf{R}} = \mathbf{R} \sqcup \{+\infty\} \sqcup \{-\infty\}$ la droite achevée. Soit $(u_n)_{n\geqslant 0}$ une suite de $\overline{\mathbf{R}}$. On pose :

$$\limsup_{n\to\infty} u_n = \lim_{p\to\infty} (\sup_{n\geqslant p} u_p) = \inf_{p\geqslant 0} \sup_{n\geqslant p} u_p$$

et

$$\liminf_{n\to\infty} u_n = \lim_{p\to\infty} (\inf_{n\geqslant p} u_p) = \sup_{p\geqslant 0} \inf_{n\geqslant p} u_p.$$

Soit $l \in \overline{\mathbf{R}}$. Les assertions suivantes sont alors équivalentes.

(i) La suite $(u_n)_{n\geqslant 0}$ a une limite et cette limite est l.
(ii) On a : $\limsup\limits_{n\to\infty} u_n \leqslant l \leqslant \liminf\limits_{n\to\infty} u_n$.

Revenons maintenant aux algèbres de Banach. Soit A une telle algèbre avec unité e. Si $x \in A$, on pose :

$$\rho(x) = \sup_{\lambda\in \mathrm{sp}_A(x)} |\lambda|,$$

qu'on appelle le *rayon spectral* de x (dans A). C'est le rayon du plus petit disque fermé de $\mathbf{C}$ centré à l'origine et contenant le compact $\mathrm{sp}_A(x)$. On a vu à la Proposition IV .15 que

$$\rho(x) \leqslant \|x\|,$$

et nous allons prouver la *formule du rayon spectral* :

Proposition IV .28 *On a :*

$$\rho(x) = \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}.$$

On peut voir sur les Exemples IV .7, IV .8 et IV .9 que la formule n'est nullement évidente. La démonstration ci-après, dans le cas général, est assez délicate. Cependant dans le cas de l'algèbre des matrices $M_n(\mathbf{C})$ une démonstration élémentaire est proposée plus loin en exercice.

Preuve. D'après le théorème de Hilbert-Dirac, pour tout $\lambda \in \mathbf{C}$, on a $\lambda^n \in \mathrm{sp}_A(x^n)$, donc $|\lambda^n| \leqslant \|x^n\|$. Ceci implique $|\lambda| \leqslant \|x^n\|^{\frac{1}{n}}$, et donc :

$$\|\lambda\| \leqslant \liminf_{n\to\infty} \|x^n\|^{\frac{1}{n}}.$$

En passant au $\sup_{\lambda\in\mathrm{sp}_A(x)}$, on obtient : $\rho(x) \leqslant \liminf_{n\to\infty} \|x^n\|^{\frac{1}{n}}$.

Désormais, on se donne $\lambda > \|x\|$. Puisque $\|\frac{x}{\lambda}\| < 1$, le Théorème IV .12donne :

$$(*) \qquad -(x-\lambda e)^{-1} = \sum_{n=0}^{\infty} \frac{x^n}{\lambda^{n+1}}.$$

Soit $\varphi \in A'$. Posons $f(\lambda) = \varphi\left((x-\lambda e)^{-1}\right)$; cette fonction f, on l'a vu au Théorème IV .18, est holomorphe dans le complémentaire de $\mathrm{sp}_A(x)$, donc a fortiori dans l'ensemble $\{|\lambda| > \|x\|\}$. Pour $|\lambda| > \|x\|$, en appliquant φ aux deux membres de l'égalité $(*)$, on obtient le développement en série de Laurent :

$$f(\lambda) = -\sum_{n=0}^{\infty} \frac{\varphi(x^n)}{\lambda^{n+1}}.$$

D'après la théorie des fonctions holomorphes, ce développement converge *même* pour $|\lambda| > \rho(x)$, mieux que $|\lambda| > \|x\|$, puisque f est holomorphe dans $\{|\lambda| > \rho(x)\}$. En particulier, si $|\lambda| > \rho(x)$, alors pour toute $\varphi \in A'$

$$\sup_{n \geqslant 0} |\varphi(\frac{x^n}{\lambda^n})| < +\infty,$$

donc d'après le théorème de Banach-Steinhaus appliqué à l'espace de Banach A' et à la famille de formes sur A'

$$\varphi \mapsto \varphi(\frac{x^n}{\lambda^n}) \quad (n \geqslant 0),$$

on a :

$$\sup_{n \geqslant 0} \|\frac{x^n}{\lambda^n}\| < +\infty.$$

Ainsi : $|\lambda| > \rho(x)$ implique $\|\frac{x^n}{\lambda^n}\| \leqslant C(\lambda)$ pour tout $n \geqslant 0$, où la constante $C(\lambda)$ ne dépend que de λ. Par conséquent, pour tout λ tel que $|\lambda| > \rho(x)$, on a :

$$\|x^n\|^{\frac{1}{n}} \leqslant C(\lambda)^{\frac{1}{n}} |\lambda|$$

et donc

$$\limsup_{n \to \infty} \|x^n\|^{\frac{1}{n}} \leqslant |\lambda|.$$

En passant au $\inf_{|\lambda| > \rho(x)}$, on obtient :

$$\limsup_{n \to \infty} \|x^n\|^{\frac{1}{n}} \leqslant \rho(x),$$

ce qui, vu le rappel technique sur les suites, achève la démonstration. □

Exercice IV .29 *On propose ici une preuve élémentaire de la formule du rayon spectral dans le cas de dimension finie, c'est-à-dire :*

$$\lim_{n \to \infty} \|M^n\|^{\frac{1}{n}} = \text{le maximum des modules des valeurs propres de } M,$$

pour tout matrice $M \in M_n(\mathbf{C})$.

1. *Remarquer qu'il suffit de prouver le résultat pour une norme particulière.*
2. *Montrer que, si* $M \in M_n(\mathbf{C})$ *et* $\varepsilon > 0$ *sont fixés, il existe une norme matricielle* $|||\cdot|||$ *telle que*

$$|||M||| \leqslant \rho(M) + \varepsilon.$$

3. *Conclure.*

IV.23 Calcul fonctionnel holomorphe

Soit $x \in A$ et soit $\lambda \mapsto f(\lambda)$ une fonction holomorphe au voisinage de l'ensemble $\mathrm{sp}(x) = \mathrm{sp}_A(x)$. Dans ce paragraphe, on va construire un élément de A, noté $f(x)$, et justifier ce symbolisme. Mais ceci nécessite quelques préliminaires.

• Préliminaires sur l'intégrale vectorielle.

Soit A un espace de Banach sur $\mathbf{C}$, et soient A' et A'' le dual et le bidual de cet espace, respectivement. Soit γ un arc de courbe de classe C^1 par morceaux dans $\mathbf{C}$, et $f : \gamma \to A$ une fonction définie et continue sur γ, à valeurs (vectorielles) dans A. Si $\varphi \in A'$, l'application $\lambda \mapsto \langle f(\lambda)|\varphi\rangle$ est une fonction numérique définie et continue sur l'arc γ, donc son intégrale

$$\int_\gamma \langle f(\lambda)|\varphi\rangle \,\mathrm{d}\lambda$$

est une notion bien connue. L'application

$$\varphi \mapsto \int_\gamma \langle f(\lambda)|\varphi\rangle \,\mathrm{d}\lambda$$

est une forme linéaire en φ, continue sur A'. Donc il existe un élément et un seul $\mathcal{I} \in A''$ tel que

$$\mathcal{I}(\varphi) = \int_\gamma \langle f(\lambda)|\varphi\rangle \,\mathrm{d}\lambda$$

pour tout $\varphi \in A'$. En fait, cet élément est dans A, ce qui est mieux que A'' ; *nous admettrons ce fait*, qui est une conséquence du théorème selon lequel l'enveloppe convexe fermée d'un compact est compacte. On pose :

$$\mathcal{I} = \vec{\int_\gamma} f(\lambda)\,\mathrm{d}\lambda.$$

Ainsi l'intégrale vectorielle $\vec{\int_\gamma} f(\lambda)\,\mathrm{d}\lambda$ est l'unique élément de A tel que

$$\langle \vec{\int_\gamma} f(\lambda)\,\mathrm{d}\lambda|\varphi\rangle = \int_\gamma \langle f(\lambda)|\varphi\rangle \,\mathrm{d}\lambda$$

pour toute $\varphi \in A'$. On a l'inégalité :

$$\|\vec{\int_\gamma} f(\lambda)\,\mathrm{d}\lambda\| \leqslant \int_\gamma \|f(\lambda)\|_A \,\mathrm{d}\lambda.$$

• Préliminaires sur l'espace des fonctions holomorphes au voisinage d'un fermé de $\mathbf{C}$

Soit F un ensemble fermé dans $\mathbf{C}$. Nous noterons $\mathcal{H}(F)$ l'ensemble des fonctions f, à valeurs complexes, définies et holomorphes dans un ouvert U_f (dépendant de f) qui contient F.

On vérifie facilement que si f et g sont dans $\mathcal{H}(F)$ et $\lambda \in \mathbf{C}$, alors λf, $f + g$, fg sont dans $\mathcal{H}(F)$ – les deux affirmations au moyen de l'ouvert $U_f \cap U_g$ – donc $\mathcal{H}(F)$ est un espace vectoriel, et même une algèbre sur $\mathbf{C}$. Pour faire de l'analyse dans cet espace, on définit une notion de limite : soit $(f_n)_{n\geqslant 0}$ une suite dans $\mathcal{H}(F)$ et soit $f \in \mathcal{H}(F)$; alors on dit que $(f_n)_{n\geqslant 0}$ tend vers f dans $\mathcal{H}(F)$ s'il existe un ouvert fixe U indépendant de n qui contient F et tel que les fonctions f_n et f sont définies et holomorphes dans U et la suite f_n converge uniformément vers f sur U.

- Préliminaires sur l'idée du calcul fonctionnel holomorphe

Soit A une algèbre de Banach avec élément neutre e. Soit $x \in A$. On sait déjà définir $P(x)$, élément de A, pour tout polynôme P à coefficients complexes ; si

$$P(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n,$$

on pose :

$$P(x) = a_0 e + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

On voit aussi comment définir $f(x)$, élément de A, pour certaines fonctions de variable complexe. Par exemple, pour

$$f(\lambda) = \exp(\lambda) = 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^n}{n!} + \dots,$$

on pose

$$f(x) = \exp(x) = e + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \dots,$$

en remarquant que la série converge absolument, donc converge, dans l'espace de Banach A, vu que : $\displaystyle\sum_{n=0}^{\infty} \frac{\|x^n\|}{n!} \leqslant \sum_{n=0}^{\infty} \frac{\|x\|^n}{n!} = e^{\|x\|} < +\infty.$

Nous étudierons systématiquement $\exp(x)$ à la section IV.25 .

Plus généralement, si $f(\lambda)$ est une fonction *entière* (i.e. holomorphe sur $\mathbf{C}$ tout entier), de la variable complexe λ ;

$$f(\lambda) = f(0) + \frac{f'(0)}{1!}\lambda + \frac{f^{(2)}(0)}{2!}\lambda^2 + \cdots + \frac{f^{(n)}(0)}{n!}\lambda^n + \dots$$

on peut poser

$$f(x) = f(0)e + \frac{f'(0)}{1!}x + \frac{f^{(2)}(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \dots$$

car cette série converge absolument, donc converge, dans A.

Soit $f \in \mathcal{H}(\mathrm{sp}(x))$ une fonction holomorphe au voisinage de l'ensemble compact $\mathrm{sp}(x)$. On va définir un élément de A qu'on notera $f(x)$ et qui englobera les exemples précédents. L'idée est de partir de la *formule intégrale de Cauchy* :

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(\lambda)}{\lambda - z}\,\mathrm{d}z$$

et de poser

$$f(x) = \frac{1}{2\pi i}\vec{\int_\gamma}(\lambda e - x)^{-1} f(\lambda)\,\mathrm{d}\lambda.$$

Théorème IV .30 *Soit A une algèbre de Banach avec unité e. Soit $x \in A$ un élément donné de cette algèbre, et $\mathrm{sp}(x)$ son spectre dans A.*

Pour toute fonction $f \in \mathcal{H}(\mathrm{sp}(x))$, donc holomorphe dans un ouvert U_f contenant $\mathrm{sp}(x)$, et pour toute courbe fermée γ de classe C^1 par morceaux, contenue dans U_f, entourant l'ensemble $\mathrm{sp}(x)$ (i.e. d'indice 1 par rapport à chaque point de $\mathrm{sp}(x)$), définissons un élément $f(x)$ de A par l'intégrale vectorielle :

$$f(x) = \frac{1}{2\pi i}\vec{\int_\gamma}(\lambda e - x)^{-1} f(\lambda)\,\mathrm{d}\lambda.$$

Alors, les assertions suivantes sont satisfaites.

(i) *L'élément $f(x)$ de A ainsi défini ne dépend pas du choix de lacet γ comme ci-dessus.*

(ii) *L'application $f \mapsto f(x)$ est un homomorphisme continu de l'algèbre $\mathcal{H}(\mathrm{sp}(x))$ dans l'algèbre de Banach A.*

(iii) *On a : $\mathrm{sp}(f(x)) = f(\mathrm{sp}(x))$, ce qui généralise le Théorème de Hilbert-Dirac.*

Preuve. Prouvons le point (i). À deux courbes γ_1 et γ_2 satisfaisant les hypothèses portant sur γ ci-dessus, correspondent a priori deux éléments de A, que nous noterons $f_{\gamma_1}(x)$ et $f_{\gamma_2}(x)$, respectivement. D'après le Théorème IV .18, pour toute $\varphi \in A'$, la fonction

$$\lambda \mapsto \langle (x - \lambda e)^{-1}, \varphi\rangle f(\lambda)$$

est holomorphe dans l'ouvert $V_f = U_f \cap (\mathbf{C}\backslash \mathrm{sp}(x))$. D'après le Théorème de Cauchy, son intégrale le long de γ_1 est égale à son intégrale le long de γ_2 car on peut déformer continûment γ_1 en γ_2 en restant dans V_f. Donc, pour toute $\varphi \in A'$, on a :

$$\langle f_{\gamma_1}(x)|\varphi\rangle = \frac{1}{2\pi i}\int_{\gamma_1}\langle (\lambda e - x)^{-1}|\varphi\rangle f(\lambda)\,\mathrm{d}\lambda = \frac{1}{2\pi i}\int_{\gamma_2}\langle (\lambda e - x)^{-1}|\varphi\rangle f(\lambda)\,\mathrm{d}\lambda = \langle f_{\gamma_2}(x)|\varphi\rangle,$$

ce qui implique que $f_{\gamma_1}(x) = f_{\gamma_2}(x)$ par Hahn-Banach.

Prouvons le point (ii), c'est-à-dire la continuité de la flèche proposée, ainsi que sa compatibilité aux lois multiplicatives et aux combinaisons linéaires.

Tout d'abord 1 et λ sont holomorphes dans $\mathbf{C}$ tout entier. On peut prendre pour γ un cercle de rayon assez grand pour que, si $\lambda \in \gamma$, on ait :

$$(\lambda e - x)^{-1} = \lambda^{-1}e + \lambda^{-2}x + \cdots + \lambda^{-n-1}x^n + \ldots$$

où la série converge absolument dans A. Alors, en utilisant les formules

$$\frac{1}{2\pi i}\int_\gamma \frac{\mathrm{d}\lambda}{\lambda^k} = 0 \quad \text{si}\, k \neq 0 \quad \text{et vaut 1 si}\, k = 1,$$

on voit que pour $f(\lambda) = 1$, on a :

$$f(x) = \frac{1}{2\pi i}\vec{\int_\gamma} (\lambda^{-1}e + \lambda^{-2}x + \cdots + \lambda^{-n-1}x^n + \ldots)\,\mathrm{d}\lambda = \left(\frac{1}{2\pi i}\int_\gamma \frac{\mathrm{d}\lambda}{\lambda}\right) e = e,$$

et pour $f(\lambda) = \lambda$, on a :

$$f(x) = \frac{1}{2\pi i}\vec{\int_\gamma} (\lambda^{-1}e + \lambda^{-2}x + \cdots + \lambda^{-n-1}x^n + \ldots)\,\mathrm{d}\lambda = \left(\frac{1}{2\pi i}\int_\gamma \frac{\mathrm{d}\lambda}{\lambda}\right) x = x,$$

La linéarité de l'application est évidente ; travaillons sur la multiplicativité et donnons-nous f_1 et f_2 dans $\mathcal{H}(\mathrm{sp}(x))$. On a :

$$f_1(x) = \frac{1}{2\pi i}\vec{\int_\gamma} (\lambda e - x)^{-1} f_1(\lambda)\,\mathrm{d}\lambda \quad \text{et} \quad f_2(x) = \frac{1}{2\pi i}\vec{\int_\gamma} (\lambda e - x)^{-1} f_2(\lambda)\,\mathrm{d}\lambda,$$

où l'on peut supposer que γ_2 contient γ_1 en son intérieur. Alors

$$\begin{array}{rcl}
f_1(x)f_2(x) & = & \frac{-1}{4\pi^2}\vec{\int}_{\gamma_1}\vec{\int}_{\gamma_2} (\lambda e - x)^{-1}(\mu e - x)^{-1} f_1(\lambda) f_2(\mu)\,\mathrm{d}\lambda\mathrm{d}\mu \\
 & = & \frac{-1}{4\pi^2}\vec{\int}_{\gamma_1}\vec{\int}_{\gamma_2} \frac{f_1(\lambda)f_2(\mu)}{\mu-\lambda}\left((\lambda e - x)^{-1} - (\mu e - x)^{-1}\right)\,\mathrm{d}\lambda\mathrm{d}\mu \\
 & = & \frac{1}{2\pi i}\vec{\int}_{\gamma_1} (\lambda e - x)^{-1} f_1(\lambda)\left(\frac{1}{2\pi i}\int_{\gamma_2}\frac{f_2(\mu)}{\mu-\lambda}\mathrm{d}\mu\right)\,\mathrm{d}\lambda \\
 & + & \frac{1}{4\pi^2}\vec{\int}_{\gamma_2} (\mu e - x)^{-1} f_2(\mu)\left(\int_{\gamma_1}\frac{f_1(\lambda)}{\mu-\lambda}\mathrm{d}\lambda\right)\,\mathrm{d}\mu.
\end{array}$$

Or, d'après la formule intégrale de Cauchy, puisque λ est à l'intérieur de γ_2 et μ à l'intérieur de γ_1, on a :

$$\frac{1}{2\pi i}\int_{\gamma_2}\frac{f_2(\mu)}{\mu-\lambda}\mathrm{d}\mu = f_2(\lambda) \quad \text{et} \quad \int_{\gamma_1}\frac{f_1(\lambda)}{\mu-\lambda}\mathrm{d}\lambda = 0,$$

donc

$$f_1(x)f_2(x) = \frac{1}{2\pi i}\vec{\int_{\gamma_1}} (\lambda e - x)^{-1}(f_1f_2)(\lambda)\,\mathrm{d}\lambda = (f_1f_2)(x).$$

Pour la continuité, on vérifie le critère séquentiel pour la notion de convergence dans $\mathcal{H}(\mathrm{sp}(x))$ décrite ci-dessus. On a :

$$\begin{array}{rcl} \|f(x)-f_n(x)\|_A & \leqslant & \frac{1}{2\pi}\int_\gamma \|(\lambda e-x)^{-1}\|\cdot|f(\lambda)-f_n(\lambda)|\,\mathrm{d}\lambda \\ & \leqslant & \sup_{\lambda\in U}|f(\lambda)-f_n(\lambda)|\cdot\frac{1}{2\pi}\int_\gamma\|(\lambda e-x)^{-1}\|\,\mathrm{d}\lambda \\ & = & C\sup_{\lambda\in U}|f(\lambda)-f_n(\lambda)|, \end{array}$$

qui tend bien vers 0 quand $n\to\infty$.

Prouvons le point (iii).

Montrons que $\mu\notin f(\mathrm{sp}(x))$ implique que $\mu\notin\mathrm{sp}(f(x))$. En effet, si $\mu\notin f(\mathrm{sp}(x))$, la fonction $\lambda\mapsto f(\lambda)-\mu$ ne s'annule pas sur $\mathrm{sp}(x)$; par compacité, il existe un ouvert $U\supset\mathrm{sp}(x)$ sur lequel cette fonction ne s'annule pas. Par conséquent, la fonction

$$\lambda\mapsto g(\lambda)=\frac{1}{f(\lambda)-\mu}$$

est définie et holomorphe dans U, donc appartient à $\mathcal{H}(\mathrm{sp}(x))$. Par calcul fonctionnel holomorphe, i.e. par le point (ii), on a :

$$g(x)\cdot(f(x)-\mu e)=e,$$

ce qui prouve que $f(x)-\mu e$ est inversible, donc que $\mu\notin\mathrm{sp}(f(x))$.

Montrons que $\mu\in\mathrm{sp}(x)$ implique que $f(\mu)\in\mathrm{sp}(f(x))$. On a :

$$\begin{array}{rcl} f(\mu)e-f(x) & = & \frac{1}{2\pi i}\vec{\int_\gamma}(\lambda e-x)^{-1}(f(\mu)-f(\lambda))\,\mathrm{d}\lambda \\ & = & \frac{1}{2\pi i}\vec{\int_\gamma}(\lambda e-x)^{-1}(\mu-\lambda)\frac{f(\mu)-f(\lambda)}{\mu-\lambda}\mathrm{d}\lambda \\ & = & \frac{1}{2\pi i}\vec{\int_\gamma}(\lambda e-x)^{-1}(\mu-\lambda)h(\lambda)\mathrm{d}\lambda \\ & = & (\mu e-x)h(x) \end{array}$$

par calcul fonctionnel holomorphe, avec la fonction h définie par

$$\lambda\mapsto h(\lambda)=\frac{f(\mu)-f(\lambda)}{\mu-\lambda}\quad\text{si }\lambda\neq\mu\quad\text{et}\quad h(\mu)=f'(\mu),$$

qui est bien une fonction de $\mathcal{H}(\mathrm{sp}(x))$. Vu que

$$f(\mu)e-f(x)=(\mu e-x)h(x),$$

l'élément $f(\mu)e-f(x)$ n'est pas inversible puisque $(\mu e-x)$ ne l'est pas. ▫

Remarque IV .31 *1. Soient $x\in A$ et $f_1,f_2\in\mathcal{H}(\mathrm{sp}(x))$. Alors les éléments $f_1(x)$ et $f_2(x)$ commutent dans A puisque $f_1f_2=f_2f_1$.*

2. D'après le point (iii) du Théorème, pour que $f(x)$ soit inversible dans A, il faut et il suffit que la fonction $\lambda\mapsto f(\lambda)$ ne soit jamais nulle sur $\mathrm{sp}(x)$.

3. *D'après le point (ii) du Théorème, si on prend pour f un polynôme $P(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n$, alors le $f(x) \in A$ que lui fait correpondre le calcul fonctionnel holomorphe :*

$$P(x) = \frac{1}{2\pi i}\vec{\int_\gamma}(\lambda e - x)^{-1}P(\lambda)\,\mathrm{d}\lambda$$

n'est autre que $P(x) = a_0e + a_1x + a_2x^2 + \cdots + a_nx^n$. Plus généralement, soit $x \in A$, et soit $f(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n + \ldots$ une série entière de rayon de convergence $R > \|x\|$. Alors $f \in \mathcal{H}(\mathrm{sp}(x))$ et si $f(x) = \frac{1}{2\pi i}\vec{\int_\gamma}(\lambda e - x)^{-1}f(\lambda)\,\mathrm{d}\lambda$ est l'élément de A que lui fait correspondre le calcul fonctionnel holomorphe, on a la formule :

$$(*) \quad f(x) = a_0e + a_1x + a_2x^2 + \cdots + a_nx^n + \ldots$$

En effet, la série converge absolument dans $\{\|x\| < R\}$, donc le second membre a un sens ; si on pose

$$f_n(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n,$$

les f_n convergent vers $f(\lambda)$ uniformément dans $U = \{|\lambda| < r$ avec $r = \dfrac{R + \|x\|}{2}$, donc on a $U \supset \mathrm{sp}(x)$. Par conséquent les

$$f_n(x) = a_0e + a_1x + a_2x^2 + \cdots + a_nx^n$$

converge dans A vers $f(x)$, d'où $(*)$ d'après le (ii) du théorème qui précède. Par exemple

$$\exp(x) = \sum_{n=0}^{\infty}\frac{x^n}{n!} = \frac{1}{2i\pi}\vec{\int_\gamma}(\lambda e - x)^{-1}e^\lambda\,\mathrm{d}\lambda$$

si γ est un cercle de rayon assez grand pour englober $\mathrm{sp}(x)$ en son intérieur : les deux définitions de l'exponentielle coïncident.

Exercice IV .32 *Commencer par relire l'exercice II .28. En déduire que si x est un élément d'une algèbre de Banach avec neutre e, on a :*

$$\lim_{n\to\infty}\left(e + \frac{x}{n}\right)^n = \exp(x).$$

IV.24 Théorème de Shilov sur la décomposition des opérateurs

Le calcul fonctionnel holomorphe est particulièrement utile en théorie des opérateurs ; si T est un opérateur, il permet de construire de nouveaux opérateurs $f(T)$ pour toute fonction f holomorphe au voisinage du spectre de T. Voici un exemple frappant d'application, dû à G. Shilov : si le spectre de T se casse en deux morceaux, alors l'espace est somme directe de deux sous-espaces fermés stables par T ; c'est l'idée de base qui permet ensuite de réduire, et si possible de diagonaliser, l'opérateur T.

Théorème IV .33 *Soit X un espace de Banach sur $\mathbf{C}$. Notons $\mathcal{L}(X) = \mathcal{L}(X, X)$ l'algèbre de Banach des opérateurs dans X (i.e. des opérateurs bornés de X dans X). Soit $T \in \mathcal{L}(X)$ et*

$$\mathrm{sp}(T) = \mathrm{sp}_{\mathcal{L}(X)}(T) = \{\lambda \in \mathbf{C} : T - \lambda \mathrm{id}_X \text{ n'est pas inversible dans} \, \mathcal{L}(X)\}.$$

Supposons que $\mathrm{sp}(T) = S_1 \sqcup S_2$, où S_1 et S_2 sont des compacts disjoints dans $\mathbf{C}$. Alors il existe deux sous-espaces vectoriels fermés X_1 et X_2 de X, stables par T, tels que $X = X_1 \oplus X_2$, et tels que

$$\mathrm{sp}_{\mathcal{L}(X_1)}(T|_{X_1}) = S_1 \quad \text{et} \quad \mathrm{sp}_{\mathcal{L}(X_2)}(T|_{X_2}) = S_2,$$

où $T|_{X_i}$ désigne la restriction de T à X_i pour $i \in \{1; 2\}$.

Preuve. L'idée est de définir X_1 et X_2 par leurs projecteurs E_1 et E_2, eux-mêmes obtenus par calcul fonctionnel holomorphe à partir de fonctions e_1 et e_2 indicatrices d'ensembles, car les fonctions indicatrices partagent avec les projecteurs la propriété algébrique d'être idempotent (i.e. d'être égal à son carré). Agrandissons S_1 et S_2 en deux ouverts U_1 et U_2, respectivement, tels que $U_1 \cap U_2$ soit encore vide ; c'est possible car S_1 et S_2 étant compacts, on a :

$$d(S_1, S_2) = \inf\{|\lambda_1 - \lambda_2| : \lambda_1 \in S_1, \lambda_2 \in S_2\} > 0.$$

Soit $U = U_1 \sqcup U_2$. La fonction e_1 définie par $e_1(\lambda) = 1$ sur U_1 et 0 sur U_2 est holomorphe dans U donc sur $\mathrm{sp}(T)$. Le calcul fonctionnel holomoprhe lui fait correspondre un opérateur $E_1 = e_1(T) \in \mathcal{L}(X)$. De même, la fonction e_2 définie par $e_2(\lambda) = 1$ sur U_2 et 0 sur U_1 est holomorphe dans U donc sur $\mathrm{sp}(T)$, et le calcul fonctionnel holomorphe lui fait correspondre un opérateur $E_2 = e_2(T) \in \mathcal{L}(X)$. Il est clair que les fonctions e_1 et e_2 satisfont les identités

$$e_1 + e_2 = 1; \quad e_1 e_2 = e_2 e_1 = 0; \quad e_1^2 = e_1; \quad e_2^2 = e_2.$$

Par l'homomorphisme du calcul fonctionnel holomorphe, on a donc dans $\mathcal{L}(X)$ les relations :

$$E_1 + E_2 = \mathrm{id}_X; \quad E_1 E_2 = E_2 E_1 = 0; \quad E_1^2 = E_1; \quad E_2^2 = E_2.$$

Soient $X_1 = \{x \in X : E_1 x = x\}$ et $X_2 = \{x \in X : E_2 x = x\}$ les sous-espaces propres de valeur propre 1 associés aux opérateurs E_1 et E_2 : ce sont évidemment des sous-espaces fermés de X. On a $X_1 \cap X_2 = \{0\}$ car si x est dans cette intersection, on peut écrire $x = E_1 x = E_1(E_2 x) = (E_1 E_2)x = 0$.

On a $X = X_1 + X_2$ car pour tout $x \in X$ on peut écrire : $x = \mathrm{id}_X x = (E_1 + E_2)x = E_1 x + E_2 x$. Ainsi $X = X_1 \oplus X_2$. Enfin, montrons que X_1 et X_2 sont stables par T. Si $x \in X_i$, alors $x = E_i x$, donc $Tx = (TE_i)x = (E_i T)x = E_i(Tx)$ et donc $Tx \in E_1$ (on rappelle que les éléments $f(T)$ commutent entre eux). On peut donc parler des restrictions $T|_{X_i}$. Introduisons aussi

$$T_1 = TE_1 = E_1 T \quad \text{et} \quad T_2 = TE_2 = E_2 T \ ;$$

ce sont des éléments de $\mathcal{L}(X)$. L'opérateur T_1 coïncide avec avec $T|_{X_1}$ sur X_1 et vaut 0 sur X_2. D'autre part

$$T_1 = \frac{1}{2\pi i} \vec{\int}_{\gamma = \gamma_1 \sqcup \gamma_2} \lambda e_1(\lambda)(\lambda \mathrm{id}_X - T)^{-1} \, \mathrm{d}\lambda,$$

donc, vu que $e_1(\lambda) = 0$ sur γ_2 et vaut 1 sur γ_1,

$$T_1 = \frac{1}{2\pi i} \vec{\int}_{\gamma_1} \lambda(\lambda \mathrm{id}_X - T)^{-1} \, \mathrm{d}\lambda,$$

puis, quel que soit μ fixé dans $\mathbf{C}$,

$$T_1 - \mu E_1 = \frac{1}{2\pi i} \vec{\int}_{\gamma_1} (\lambda - \mu)(\lambda \mathrm{id}_X - T)^{-1} \, \mathrm{d}\lambda.$$

Supposons μ *extérieur* à la courbe γ_1 ; par calcul fonctionnel holomorphe définissons l'opérateur dans X

$$Q_\mu = \frac{1}{2\pi i} \vec{\int}_{\gamma_1} \frac{1}{(\lambda - \mu)} (\lambda \mathrm{id}_X - T)^{-1} \, \mathrm{d}\lambda.$$

On a clairement $Q_\mu = E_1 Q_\mu = Q_\mu E_1$, donc X_1 est stable par Q_μ. De plus

$$(T_1 - \mu E_1) Q_\mu = \frac{1}{2\pi i} \vec{\int}_{\gamma_1} (\lambda - \mu) \frac{1}{(\lambda - \mu)} (\lambda \mathrm{id}_X - T)^{-1} \, \mathrm{d}\lambda = \frac{1}{2\pi i} \vec{\int}_{\gamma_1} e_1(\lambda)(\lambda \mathrm{id}_X - T)^{-1} \, \mathrm{d}\lambda = E_1.$$

En restreignant l'égalité $(T_1 - \mu E_1) Q_\mu = E_1$ aux vecteurs de X_1, on voit que l'on a prouvé que $(T - \mu \mathrm{id}_X)|_{X_1}$ est inversible dans $\mathcal{L}(X_1)$, ayant pour inverse $Q_\mu|_{X_1}$; ceci est valable quel que soit $\mu \notin S_1$, car on peut choisir γ_1 laissant μ à l'extérieur. Ainsi est prouvé que

$$\mathrm{sp}_{\mathcal{L}(X_1)}(T|_{X_1}) \subset S_1 \quad \text{et} \quad \mathrm{sp}_{\mathcal{L}(X_2)}(T|_{X_2}) \subset S_2.$$

Reste à voir que, par exemple, tout $\lambda_0 \in S_1$ est un point du spectre dans $\mathcal{L}(X_1)$ de $T|_{X_1}$. On vient de voir que $(T - \lambda_0 \mathrm{id}_X)|_{X_2}$ est inversible dans $\mathcal{L}(X_2)$, car λ_0 est extérieur à γ_2. Donc il existe $Q_2 \in \mathcal{L}(X_2)$ tel que

$$(T - \lambda_0 \mathrm{id}_X) Q_2 y = y \quad \text{pour tout} \quad y \in X_2.$$

Si, par l'absurde, $(T - \lambda_0 \mathrm{id}_X)|_{X_1}$ était inversible, il existerait un opérateur $Q_1 \in \mathcal{L}(X_1)$ tel que

$$(T - \lambda_0 \mathrm{id}_X) Q_1 y = y \quad \text{pour tout} \quad y \in X_1.$$

En posant $Q = Q_1$ sur X_1 et $Q = Q_2$ sur X_2, on obtiendrait un opérateur $Q \in \mathcal{L}(X)$ tel que $(T - \lambda_0 \mathrm{id}_X) Q = \mathrm{id}_X$ dans X tout entier. C'est absurde, car $\lambda_0 \in S_1$, donc $\lambda_0 \in \mathrm{sp}_{\mathcal{L}(X)}(T)$. □

Exercice IV .34 *Soit A une algèbre de Banach avec unité. Soit U un ouvert de $\mathbf{C}$. Soit $\Omega = \{x \in A : \mathrm{sp}_A(x) \subset U\}$. Montrer que Ω est ouvert dans A, et que si f est une fonction holomorphe dans U, alors l'application $x \mapsto f(x)$ de Ω dans l'espace de Banach A est différentiable.*

IV.25 Fonction exponentielle

C'est l'occasion de reprendre ce thème très classique dans un contexte nouveau. Plus qu'une partie de cours ce pourrait être un travail dirigé. Aussi le lecteur est invité à essayer de trouver et de rédiger par lui-même, avant de les lire dans le texte, les démonstrations des énoncés ci-après.

Soit A une algèbre de Banach avec unité e. Pour tout $x \in A$, on pose :

$$\exp(x) = e + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \ldots$$

Nous savons déjà que la série converge (absolument) dans A, que dans A on a la formule

$$\exp(x) = \lim_{n\to\infty} (e + \frac{x}{n})^n,$$

et que, si γ est un cercle de rayon $> \|x\|$ (ou même seulement de rayon assez grand pour englober l'ensemble $\mathrm{sp}(x)$ dans son intérieur), on a :

$$\exp(x) = \frac{1}{2\pi i} \vec{\int_\gamma} (\lambda e - x)^{-1} e^\lambda \, \mathrm{d}\lambda.$$

Proposition IV .35 *L'application $x \mapsto \exp(x)$ de A dans A est continue.*

Preuve. Si $\|x - x_0\| \leqslant 1$ et si N est un entier > 1, on a :

$$\begin{aligned}
\|\exp(x) - \exp(x_0)\| &= \|\textstyle\sum_{n\geqslant 1} \frac{1}{n!}(x^n - x_0^n)\| \\
&= \|\textstyle\sum_{n=1}^N \frac{1}{n!}(x^n - x_0^n) + \sum_{n\geqslant N+1} \frac{1}{n!}(x^n - x_0^n)\| \\
&\leqslant \|\textstyle\sum_{n=1}^N \frac{1}{n!}(x^n - x_0^n)\| + 2\sum_{n=N+1}^\infty \frac{1}{n!}(\|x_0\| + 1)^n
\end{aligned}$$

On fixe N de sorte que $\sum_{n=N+1}^\infty \frac{1}{n!}(\|x_0\| + 1)^n \leqslant \frac{\varepsilon}{4}$, ce qui est possible car c'est le reste d'une série numérique convergente. Pour finir, par continuité des fonctions polynômes on utilise que

$$\|\sum_{n=1}^N \frac{1}{n!}(x^n - x_0^n)\| \leqslant \frac{\varepsilon}{2}$$

dès que $\|x - x_0\|$ est assez petit. $\square$

Proposition IV .36 *Si $x \in A$ et $y \in A$ commutent (i.e. $xy = yx$), alors on a :* $\exp(x+y) = \exp(x)\exp(y)$.

Preuve. Puisque que x et y commutent, on a, pour tout entier $n \geqslant 0$, la formule du binôme de Newton :

$$(x+y)^n = \sum p + q = n \frac{n!}{p!q!} x^p y^q.$$

Le théorème sur la multiplication de Cauchy de deux séries absolument convergentes est valable dans A (même démonstration que pour les séries dans $\mathbf{C}$, en remplaçant les modules par des normes) ; d'après ce théorème, si on pose :

$$\exp(x) = \sum_{p \geqslant 0} \frac{x^p}{p!} = \sum_{p \geqslant 0} u_p \quad \text{et} \quad \exp(y) = \sum_{q \geqslant 0} \frac{y^q}{q!} = \sum_{q \geqslant 0} v_q,$$

alors

$$\exp(x)\exp(y) = \sum_{n=0}^{\infty} w_n$$

où l'on a :

$$w_n = \sum_{p+q=n} u_p v_q = \frac{1}{n!} \sum_{p+q=n} \frac{n!}{p!q!} x^p y^q = \frac{1}{n!}(x+y)^n$$

pour tout $n \geqslant 0$. $\square$

Exercice IV .37 *Soient $A \in M_2(\mathbf{C})$, $x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Calculer* $\exp(x)$, $\exp(y)$, $\exp(x+y)$ *et constater que* $\exp(x)\exp(y) \neq \exp(x+y)$.

Un substitut à la Proposition IV .36, quand x et y ne commutent pas, est la :

Proposition IV .38 *Pour tous x et y dans A, on a la formule :*

$$\exp(x+y) = \lim_{n \to \infty} \left(\exp(\frac{x}{n}) \exp(\frac{y}{n}) \right)^n.$$

Preuve. Pour $n > k$, posons :

$$u_n = (e + \frac{x+y}{n})^n = \sum_{p=0}^{n} \frac{n(n-1)\cdots(n-p+1)}{n^p} \frac{(x+y)^p}{p!},$$

$$s_{n,k} = \sum_{p=0}^{k} \frac{n(n-1)\cdots(n-p+1)}{n^p} \frac{(x+y)^p}{p!}$$

et

$$r_{n,k} = \sum_{p=k+1}^{n} \frac{n(n-1)\cdots(n-p+1)}{n^p}\frac{(x+y)^p}{p!}.$$

D'autre part,

$$z_n = \exp(\frac{x}{n})\exp(\frac{y}{n}) = e + \frac{1}{n}(x+y) + \frac{1}{n^2}w_n,$$

où w_n reste borné quand n varie. Donc, pour $n > k$, on a :

$$(z_n)^n = \left(e + \frac{1}{n}(x+y+\frac{w_n}{n})\right)^n = \sum_{p=0}^{n} \frac{n(n-1)\cdots(n-p+1)}{n^p}\frac{\left(x+y+\frac{w_n}{n}\right)^p}{p!},$$

$$\sigma_{n,k} = \sum_{p=0}^{k} \frac{n(n-1)\cdots(n-p+1)}{n^p}\frac{\left(x+y+\frac{w_n}{n}\right)^p}{p!}$$

et

$$\rho_{n,k} = \sum_{p=k+1}^{n} \frac{n(n-1)\cdots(n-p+1)}{n^p}\frac{\left(x+y+\frac{w_n}{n}\right)^p}{p!}.$$

On peut fixer un rang k assez grand pour que pour tout $n > k$ on ait :

$$(1) \quad \|r_{n,k}\| \leqslant \frac{\varepsilon}{3}$$

et

$$(2) \quad \|\rho_{n,k}\| \leqslant \frac{\varepsilon}{3}$$

car d'une part, quel que soit p, on a $\frac{n(n-1)\cdots(n-p+1)}{n^p} \leqslant 1$ et, d'autre part, $\|x+y+\frac{w_n}{n}\|$ est borné par une constante C indépendante de n, vu que $(w_n)_{n\geqslant 0}$ est bornée ; par conséquent,

$$\|r_{n,k}\| \leqslant \sum_{p>k} \frac{(x+y)^p}{p!} \quad \text{et} \quad \|\rho_{n,k}\| \leqslant \sum_{p>k} \frac{C^p}{p!}$$

sont $\leqslant \frac{\varepsilon}{3}$ pour k assez grand, au titre de restes de séries abslolument convergentes.

L'indice k étant désormais fixé, choisissons $N > k$ tel que pour tout $n > N$ on ait :

$$(3) \quad \|s_{n,k} - \sigma_{n,k}\| \leqslant \frac{\varepsilon}{3}.$$

C'est possible car, à k fixé pour les p en nombre fini tels que $0 \leqslant p \leqslant k$, on a :

$$\lim_{n\to\infty} \frac{n(n-1)\cdots(n-p+1)}{n^p} = 0 \quad \text{et} \quad \lim_{n\to\infty} \frac{1}{n}w_n = 0,$$

donc

$$\lim_{n\to\infty} \frac{n(n-1)\cdots(n-p+1)}{n^p} \frac{(x+y)^p - (x+y+\frac{1}{n}w_n)^p}{p!} = 0.$$

La combinaison de (1), (2) et (3) montre que $\lim_{n\to\infty} (u_n - (z_n)^n) = 0$, et comme on sait depuis l'Exercice IV .32 que $\lim_{n\to\infty} u_n = \exp(x+y)$, on obtient que les deux membres de l'égalité dans l'énoncé valent tous les deux $\lim_{n\to\infty} (z_n)^n$. □

Il est bon de relire pour la suite l'Exercice 15 du Chapitre II. Si $x \in A$ vérifie $\|x - e\| < 1$, on pose

$$\log(x) = \sum_{m\geqslant 1} -\frac{(e-x)^m}{m},$$

ce qui a un sens et définit un élément de A, car la série est absolument convergente, vu que

$$\sum_{m\geqslant 1} \|-\frac{(e-x)^m}{m}\| \leqslant \sum_{m\geqslant 1} \frac{\|e-x\|^m}{m} = -\log(1-\|e-x\|) < +\infty.$$

D'après le calcul fonctionnel holomorphe on a aussi, pour α assez petit :

$$\log(x) = \frac{1}{2\pi i} \vec{\int}_{|1-\lambda|=\alpha} (\lambda e - x)^{-1} \log(\lambda)\, \mathrm{d}\lambda$$

où log est la détermination principale du logarithme.

Proposition IV .39 *Soit $x \in A$ tel que $\|x - e\| < 1$. Alors* $\exp(\log(x)) = x$.

Preuve. Soit U le disque $\{|z-1| < 1\}$ dans $\mathbf{C}$. Pour $z \in U$ et $n > 1$ entier posons :

$$f_n(z) = \sum_{m=1}^{n} -\frac{(1-z)^m}{m}.$$

D'après la théorie des séries entières, $(f_n)_{n\geqslant 1}$ converge vers $\log(z)$ uniformément sur tout compact de U. Par continuité de l'exponentielle, $(e^{f_n})_{n\geqslant 1}$ converge vers $e^{\log(z)} = z$ uniformément sur tout compact de U. Posons

$$f_n(x) = \sum_{m=1}^{n} -\frac{(e-x)^m}{m}.$$

Par définition de $\log(x)$, la suite $f_n(x)$ tend vers $\log(x)$ dans A quand $n \to \infty$, donc par continuité la suite des $e^{f_n(x)}$ tend vers $e^{\log(x)}$ dans A quand $n \to \infty$.

Mais, d'autre part, $\exp(f_n(x))$ n'est autre que l'image dans A, par le calcul fonctionnel holomorphe, de la fonction $e^{f_n(z)}$: la vérification rigoureuse de ceci consiste à se convaincre d'abord que, à k fixé, la fonction e^{z^k} donne $\exp(x^k)$ dans le calcul fonctionnel holomorphe, ce qui est clair sur la série exponentielle. Mais, par ailleurs, $e^{f_n(z)}$ tend vers z dans l'espace $\mathcal{H}(\mathrm{sp}(x))$. Donc $\exp(f_n(x))$ tend dans A d'une part vers x et, d'autre part comme on l'a vu, vers $\exp(\log(x))$. Ainsi $\exp(\log(x)) = x$. $\square$

Soit $\exp A = \{x \in A : \text{il existe } y \in A \text{ tel que } x = \exp(y)\}$ l'image de A dans A par l'application exponentielle. Nous allons comparer $\exp A$ au groupe des inversibles $G = A^\times$. D'abord une évidence : $\exp A \subset G$, car $e = \exp(0)$ et, d'après la Proposition IV .36, pour tout $x \in A$:

$$\exp(x)\exp(-x) = \exp(-x)\exp(x) = \exp(0) = e,$$

prouvant que $\exp(x)$ est inversible dans A avec $\exp(x)^{-1} = \exp(-x)$.

Proposition IV .40 *Le sous-ensemble* $\exp A$ *est une partie connexe de* G*.*

Preuve. Le sous-ensemble $\exp A$ est même connexe par arcs, car pour tout $x = \exp(y)$, le chemin continu $t \mapsto \exp(ty)$ $(0 \leqslant t \leqslant 1)$ relie e à x. $\square$

Ici, ouvrons une parenthèse pour un :

• Rappel sur la connexité.

Un espace métrique est dit *connexe* s'il n'est pas réunion de deux ouverts disjoints non vides. Une partie C d'un espace métrique E est dite *connexe* si le sous-espace métrique C de E est connexe, c'est-à-dire si les hypothèses :

$$U_1 \text{ et } U_2 \text{ ouverts de } E, \quad U_1 \cap C \neq \varnothing, \quad U_2 \cap C \neq \varnothing \quad \text{et} \quad C \subset U_1 \cup U_2$$

entraînent que $C \cap U_1 \cap U_2 \neq \varnothing$.

Si C est connexe, alors $\overline{C}$ est connexe, et toute image continue de C est connexe. Si $(C_\lambda)_{\lambda\in\Lambda}$ est une famille de parties connexes de E telle que $\bigcap_{\lambda\in\Lambda} C_\lambda \neq \varnothing$, alors $\bigcup_{\lambda\in\Lambda} C_\lambda \neq \varnothing$ est connexe.

Dans E normé, toute partie convexe est connexe.

Soit $x \in E$. La *composante connexe de x dans E*, notée C_x, est la réunion des parties connexes de E qui contiennent x ; c'est donc la plus grande partie connexe de E qui contienne x ; elle est fermée dans E. Si x et y sont dans E, ou bien $C_x = C_y$ ou bien $C_x \cap C_y = \varnothing$. La relation $x \sim y$ si et seulement si $C_x = C_y$ est une relation d'équivalence dont les classes s'appellent *composantes connexes* de E.

Suposons maintenant que E (qu'on notera plutôt G) soit un *groupe topologique*, ce qui signifie que sont continues les applications $(x, y) \mapsto xy$ de

$G \times G$ dans G et $x \mapsto x^{-1}$ de G dans G. C'est le cas pour le groupe G des inversibles d'une algèbre de Banach A.

Proposition IV .41 *La composante connexe $G_\circ$ de l'élément neutre e dans un groupe topologique G est un sous-groupe distingué fermé de G.*

Preuve. Si $x \in G_\circ$, alors $x^{-1}G_\circ$ est connexe, comme image continue (par la multiplication à droite par x^{-1}) du connexe $G_\circ$, et contient e, donc $x^{-1}G_\circ \subset G_\circ$ pour tout $x \in G_\circ$. Par conséquent, $G_\circ$ est un sous-groupe de G ; il est fermé comme toute composante connexe. Si $a \in G$, alors $a^{-1}G_\circ a$ est connexe, comme image continue (par la conjugaison par a^{-1}) du connexe $G_\circ$, et contient e, donc $a^{-1}G_\circ a \subset G_\circ$ pour tout $a \in G$. Ainsi $G_\circ$ est-il distingué dans G. □

Théorème IV .42 *Soit A une algèbre de Banach avec unité, notée e. Soit G le groupe de ses éléments inversibles. Soit $G_\circ$ la composante connexe de e dans G. Alors les assertions suivantes sont satisfaites.*

(i) *Le sous-groupe $G_\circ$ est engendré par* $\exp A$ *;*

(ii) *Si A est commutative, alors $G_\circ = \exp A$.*

Preuve. Posons $V = \exp A$. D'après la Proposition IV .40, V est contenue dans $G_\circ$. Le sous-groupe de $G_\circ$ engendré par V n'est autre que

$$V^\infty = \bigcup_{n \geqslant 0} V^n$$

où V^n est l'ensemble des produits de n éléments de V ; ceci découle du fait que V est symétrique (i.e. $V^{-1} = V$). Bref, $\langle V \rangle$ est l'ensemble des produits finis d'éléments de V.

On a vu à la Proposition IV .39 que V contient la boule ouverte de centre e et de rayon 1 (tout élément de cette boule est l'exponentielle de son logarithme), donc e est un point intérieur à V, donc intérieur à V^∞. Par translation, chaque $x \in V^\infty$ est intérieur à V^∞ ; par conséquent le sous-groupe V^∞ est ouvert dans $G_\circ$; du coup il est aussi fermé dans $G_\circ$ (car c'est le complémentaire dans $G_\circ$ des classes à gauche de $G_\circ$ modulo V^∞, autres que V^∞ elle-même). Puisque $G_\circ$ est connexe et que $V^\infty \neq \varnothing$, on a : $V^\infty = G_\circ$. Ceci prouve le point (i). En ce qui concerne (ii), si A est commutative $\exp A$ est déjà un groupe, d'après la Proposition IV .36, et donc $G_\circ = \exp A$. □

Exemple IV .43 *Sachant que le groupe $\mathrm{GL}_n(\mathbf{C})$ des matrices $n \times n$ complexes inversibles est connexe, on obtient que toute matrice $n \times n$ complexe de déterminant $\neq 0$ est produit d'un nombre fini de matrices exponentielles.*

Exercice IV .44 *1. Soit $x \in A$ tel que $x^n = e$ pour un entier $n > 0$ au moins. Montrer que* $\mathrm{sp}(x)$ *est fini, et que $x \in G_\circ$*

2. Supposons A commutative. Soit $x \in G$ tel que $x^n \in G_\circ$ pour un entier $n > 0$ au moins. Montrer qu'alors $x \in G_\circ$. En déduire que si A est commutative, ou bien $G = G_\circ$ ou bien $G/G_\circ$ est sans torsion (et c'est alors un groupe infini).

IV.26 Transformation de Gelfand

Dans tout ce paragraphe, A désigne une algèbre de Banach commutative avec unité, notée e. En particulier, A est un anneau commutatif ; à ce titre, A a des idéaux, et en particulier des idéaux maximaux. Nous allons étudier l'ensemble $\mathcal{M}_A$ des idéaux maximaux de A.

Mais d'autre part A est un espace de Banach, pour lequel on connaît déjà l'importance du dual A', ensemble des formes linéaires continues sur A ; là encore, si on veut tenir compte de l'existence d'un produit sur A, on est amené à faire jouer un rôle privilégié aux $f \in A'$, non identiquement nulles, qui sont de plus *multiplicatives*, c'est-à-dire telles que $f(xy) = f(x)f(y)$ pour tous x et y dans A. Ces formes multiplicatives continues sont appelés les *caractères* de A . Nous allons étudier l'ensemble $\widehat{A}$ des caractères de A.

Il y a une bijection naturelle entre $\mathcal{M}_A$ et $\widehat{A}$, chaque idéal maximal étant le noyau d'un caractère, et réciproquement.

Rappelons qu'un *idéal* I de A est un sous-espace vectoriel de A tel que $x \in I$ et $y \in A$ entraîne $xy \in A$. Un idéal I est dit *propre* si I est différent de A tout entier. Un idéal M est dit *maximal* s'il est propre et si tout idéal propre qui le contient est égal à M. Il revient au même de dire que A/M est un corps (par définition un corps n'est pas l'anneau nul). Tout idéal propre I est contenu dans au moins un idéal maximal (théorème de Krull, ou encore lemme de Zorn). Soit I un idéal propre fermé de l'algèbre de Banach A. Alors l'anneau et espace vectoriel A/I, ensemble des classes $\overline{x} = x + I$, où $x \in A$, avec les opérations évidentes, et la norme

$$\|\overline{x}\| = \inf_{y \in \overline{x} = x+I} \|y\|,$$

est une algèbre de Banach, avec $\overline{e}$ comme unité.

Soit G l'ensemble des inversibles de A. Si $x \in G$, alors x n'est dans aucun idéal propre car $xx^{-1} = e$ appartiendrait à cet idéal, et donc on aurait $y = ye \in I$ pour tout $y \in A$. D'autre part, si $x \notin G$, alors l'idéal xA engendré par x est propre car $e \notin xA$. En conséquence, la réunion des idéaux maximaux de A coïncide avec le complémentaire de l'ensemble G des inversibles de A.

Proposition IV .45 *Soit M un idéal maximal de A. Alors M est fermé dans A et le corps M/A est le corps des nombres complexes.*

Preuve. On vient de voir que $M \cap G$ est vide. Puisque G est ouvert, $\overline{M} \cap G$ est encore vide, donc $\overline{M}$ est un idéal propre qui contient M : on a donc $\overline{M} = M$, ce qui prouve que M est fermé. Le quotient A/M est une algèbre de Banach dont l'anneau est un corps puisque M est maximal ; d'après le théorème de Gelfand-Mazur, ce corps est le corps des nombres complexes. □

Definition IV .46 *Un caractère de A est une fonction $\chi : A \to \mathbf{C}$ qui possède les propriétés suivantes.*

(i) *La fonction χ n'est pas identiquement nulle.*
(ii) *La fonction χ est une forme linéaire sur A.*
(iii) *La fonction χ est multiplicative.*

De cette définition, il résulte que :
(iv) On a $\chi(e) = 1$.
En effet, si $\chi(a) \neq 0$, alors $\chi(a) = \chi(ae) = \chi(a)\chi(e)$.

On appelle *spectre de Gelfand* de A, et l'on note $\widehat{A}$, l'ensemble des caractères de A. On note $\mathcal{M}_A$ l'ensemble des idéaux maximaux de A.

Proposition IV .47 *L'application $\chi \mapsto \mathrm{Ker}(\chi)$ établit une bijection de $\widehat{A}$ sur $\mathcal{M}_A$.*

Preuve. Soit $\chi \in \widehat{A}$. D'après (ii) et (iii) c'est un homomorphisme d'algèbres, qui est surjectif car $\chi(\lambda e) = \lambda\chi(e) = \lambda$ pour tout nombre complexe λ. Ainsi, $A/\mathrm{Ker}(\chi)$ est isomorphe à $\mathbf{C}$, qui est un corps ; donc l'idéal $\mathrm{Ker}(\chi)$ est maximal, et $\chi \mapsto \mathrm{Ker}(\chi)$ est bien une application de $\widehat{A}$ sur $\mathcal{M}_A$. Cette application est surjective, car si $M \in \mathcal{M}_A$, alors l'homomorphisme quotient associé $\chi : A \to A/M \simeq \mathbf{C}$ appartient à $\widehat{A}$, et $\mathrm{Ker}(\chi) = M$. Cette application est injective, car pour $\chi, \chi' \in A$ distinctes, il existe $a \in A$ tel que $\chi(a) \neq \chi'(a)$. Alors $y = a - \chi(a)e$ appartient à $\mathrm{Ker}(\chi)$ mais pas à $\mathrm{Ker}(\chi')$. Donc $\chi \neq \chi'$ entraîne $\mathrm{Ker}(\chi) \neq \mathrm{Ker}(\chi')$. □

Théorème IV .48 *Soit A une algèbre de Banach commutative avec élément unité, noté e. Soit $\widehat{A}$ son spectre de Gelfand, ensemble des caractères de A.*

(i) *Soit $\chi \in \widehat{A}$. Alors la forme linéaire χ sur A est continue et de norme 1. Ainsi, $\widehat{A}$ est-il une partie de la sphère unité du dual A' de A.*
(ii) *Soit $x \in A$. Alors*

$$\mathrm{sp}_A(x) = \{\chi(x) : \chi \in \widehat{A}\}.$$

(iii) *Soit $x \in A$. Pour que x soit inversible, il faut et il suffit qu'on ait $\chi(x) \neq 0$ pour tout $\chi \in \widehat{A}$.*

Preuve. On a : $\chi\,(x - \chi(x)e) = \chi(x) - \chi(x)1 = 0$, donc $x - \chi(x)e$ ne peut être inversible puisqu'il est dans un idéal maximal $M = \mathrm{Ker}(\chi)$ de A, donc dans le complémentaire de G. Ceci prouve que $\mathrm{sp}_A(x) \supset \{\chi(x) : \chi \in \widehat{A}\}$. Il en résulte que, pour $x \in A$, on a $|\chi(x)| \leqslant \|x\|$; autrement dit chaque caractère χ est une forme linéaire continue de norme $\leqslant 1$, et en fait de norme exactement 1 puisque $\chi(e) = 1$. On vient donc de prouver (i) et une moitié de (ii).

Montrons (iii) par la suite d'équivalences logiques :

$$\begin{aligned} x \in G \quad &\Leftrightarrow \quad x \notin \textstyle\bigcup_{M \in \mathcal{M}_A} M \\ &\Leftrightarrow \quad \text{pour tout } \chi \in \widehat{A} \text{ on a } x \notin \mathrm{Ker}(\chi) \\ &\Leftrightarrow \quad \text{pour tout } \chi \in \widehat{A} \text{ on a } \chi(x) \neq 0. \end{aligned}$$

Il reste à prouver que

$$\mathrm{sp}_A(x) \subset \{\chi(x) : \chi \in \widehat{A}\}.$$

Si $\lambda \in \mathrm{sp}_A(x)$, alors $x - \lambda e$ n'est pas inversible, donc, d'après le point (iii) qu'on vient de prouver il existe $\chi \in \widehat{A}$ tel que $\chi(x - \lambda e) = 0$, soit tel que $\lambda = \chi(x)$. □

Arrêtons-nous un peu sur le cas particulier des algèbres de fonctions complexes continues sur un compact X.

Soit $\mathcal{C} = \mathcal{C}(X)$ l'algèbre de Banach des fonctions continues à valeurs complexes sur un espace compact X, munie de la norme $\|\cdot\|_\infty$ de la convergence uniforme. Pour tout $x_0 \in X$ fixé, l'évaluation en x_0 :

$$\chi_{x_0} : f \mapsto f(x_0)$$

est évidemment un caractère de $\mathcal{C}$; l'idéal maximal correspondant est

$$M_{x_0} = \mathrm{Ker}(\chi_{x_0}) = \{f \in \mathcal{C}(X) : f(x_0) = 0\}.$$

Montrons qu'il n'y a pas d'autres idéaux maximaux dans $\mathcal{C}(X)$ que les M_x quand x parcourt X. Autrement dit que X s'identifie au spectre de Gelfand de $\mathcal{C}(X)$. Ceci résulte du

Lemme IV .49 *Soit I un idéal propre de $\mathcal{C}(X)$. Alors il existe au moins un point $x_0 \in X$ tel que toutes les fonctions $f \in I$ s'annulent en ce point x_0.*

Preuve. Supposons le contraire afin d'aboutir à une contradiction. Alors pour tout $y \in X$ il existe $f_y \in I$ non nulle en y et donc par continuité non nulle sur un ouvert U_y autour de y. On pose $g_y = |f_y| = f_y\overline{f}_y$: c'est une fonction de I car $f_y \in I$ et I est un idéal. Ainsi, pour tout $y \in X$ il existe une fonction de I qui est $\geqslant 0$ sur X tout entier et > 0 sur un ouvert U_y de X contenant y. Par compacité on en déduit une suite finie $y_1, y_2, \ldots y_p \in X$ telle que $X = U_{y_1} \cup U_{y_2} \cup \cdots \cup U_{y_p}$. La fonction $g = g_{y_1} + g_{y_2} + \cdots + g_{y_p}$ serait dans I et > 0 sur X tout entier, donc inversible : c'est exclu car I est propre dans $\mathcal{C}(X)$. $\square$

Si M est un idéal maximal de $\mathcal{C}(X)$, donc propre, pour le point x_0 venant du lemme appliqué à $I = M$ on aura $M \subset M_{x_0}$, et donc l'égalité par maximalité de M.

Exercice IV .50 *On va déterminer tous les idéaux fermés de $\mathcal{C}(X)$. Soit E un fermé de X. On introduit les idéaux :*

$$I_E = \{f \in \mathcal{C}(X) : f(x) = 0 \text{ pour tout } x \in E\}$$

et

$$J_E = \{f \in \mathcal{C}(X) : \text{il existe un ouvert } U_f \supset E \text{ tel que } f(x) = 0 \text{ pour tout } x \in U_f\}.$$

1. *Montrer que J_E est l'adhérence de I_E dans $\mathcal{C}(X)$. Indication : penser au Théorème d'Urysohn en considérant les fermés $Y_1 = \{x : |f(x)| \geqslant \frac{2}{n}\}$ et $Y_2 = \{x : |f(x)| < \frac{1}{n}\}$.*
2. *Soit I un idéal propre de $\mathcal{C}(X)$ et soit*

$$E = \{x \in X : f(x) = 0 \text{ pour toute } f \in I\}.$$

Montrer que

$$J_E \subset I \subset I_E.$$

Indication : si $f \in J_E$, soit Y le complémentaire de U_f dans X. Montrer qu'il existe une $g \in I$ qui est strictement positive sur Y, puis en prolongeant $\frac{1}{g}$ de Y à X par le Théorème d'Urysohn, montrer qu'il existe une $u \in I$ qui est identiquement égale à 1 sur Y.

3. *En déduire que les idéaux fermés de $\mathcal{C}(X)$ sont exactement les idéaux I_E comme ci-dessus, quand E parcourt l'ensemble des fermés de X.*
4. *Tout idéal fermé de $\mathcal{C}(X)$ est l'intersection des idéaux maximaux qui le contiennent.*

Exercice IV .51 *Soit $A = \mathcal{D}_m([0;1])$ l'algèbre de Banach définie dans l'Exercice IV .5. Montrer que les idéaux maximaux de A sont, ici encore, les noyaux des évaluations aux points de $[0;1]$, mais qu'il y a plusieurs idéaux fermés distincts à être contenus dans un seul idéal maximal donné.*

Reprenons la théorie générale des algèbres de Banach commutatives. On ne fait pas de l'ensemble $\widehat{A}$, spectre de Gelfand de A, un espace métrique, mais constatant que c'est une partie de la boule unité du dual A', on munit A' de la topologie faible de dualité avec A : par définition, on dit que les $\chi_i \in \widehat{A}$ tendent vers χ quand $i \to \infty$ si $\lim_{i\to\infty} \chi_i(a) = \chi(a)$ pour tout $a \in A$. L'ensemble $\widehat{A}$ devient alors, pour cette notion de limite, un espace topologique (non métrique en général) compact.

À tout $x \in A$, on va attacher une fonction continue sur le compact $\widehat{A}$, la transformée de Gelfand de x. L'extrême simplicité apparente de la définition pourrait en masquer la profondeur :

Definition IV .52 *Pour tout $x \in A$ donné, la fonction $\mathcal{G}x$ définie sur $\widehat{A}$, à valeurs complexes, par la formule :*

$$\mathcal{G}x(\chi) = \chi(x),$$

pour tout $\chi \in \widehat{A}$, s'appelle la transformée de Gelfand de x.

Chaque fonction $\mathcal{G}x$ est continue, car si des $\chi_i \in \widehat{A}$ tendent vers χ dans $\widehat{A}$, alors pour $x \in A$ fixé, les nombres $\chi_i(x)$ tendent vers le nombre $\chi(x)$, c'est-à-dire $\lim_{i\to\infty} \mathcal{G}x(\chi_i) = \mathcal{G}x(\chi)$.

De plus, d'après le Théorème IV .48, on a les relations :

$$\|\mathcal{G}x\|_\infty = \sup_{\chi\in\widehat{A}} |\mathcal{G}x(\chi)| = \sup_{\chi\in\widehat{A}} |\chi(x)| = \rho(x) \leqslant \|x\|$$

et

$$\mathrm{sp}(x) = \{\mathcal{G}x(\chi) : \chi \in \widehat{A}\}.$$

Le spectre de x est donc l'ensemble des valeurs prises par la transformée de Gelfand $\mathcal{G}x$. Il en résulte le

Théorème IV .53 *Pour que x soit inversible dans A, il faut et il suffit que sa transformée de Gelfand $\mathcal{G}x$ ne s'annule pas dans $\widehat{A}$.*

Tout ce que nous avons vu montre aussi que la transformée de Gelfand $x \mapsto \mathcal{G}x$ est un homomorphisme continu de l'algèbre de Banach A dans l'algèbre de Banach $\mathcal{C}(\widehat{A})$ normée par $\|\cdot\|_\infty$. En particulier, on a la formule fondamentale :

$$\mathcal{G}(xy) = (\mathcal{G}x)(\mathcal{G}y),$$

qui généralise le fait que la transformation de Fourier transforme le produit de convolution en le produit ponctuel ordinaire. En effet, on a :

$$\mathcal{G}(xy) = \chi(xy) = \chi(x)\chi(y) = \mathcal{G}x\mathcal{G}y.$$

Si x est inversible dans A, on a $\mathcal{G}(x^{-1}) = \dfrac{1}{\mathcal{G}x}$.

On remarquera que $\mathcal{G}$ n'est pas toujours injective :

Exercice IV .54 *Les quatre ensembles suivants sont égaux.*

(i) *Le noyau de $\mathcal{G}$ dans A.*

(ii) *L'intersection des idéaux maximaux de A.*

(iii) *L'ensemble des $x \in A$ tels que* $\mathrm{sp}(x) = \{0\}$.

(iv) *L'ensemble des $x \in A$ tels que* $\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}} = 0$.

Cette partie s'appelle le radical de A. Donner un exemple d'algèbre de Banach A avec un radical non trivial.

Exercice IV .55 *Montrer que $\mathcal{G}(A)$ est en général distincte de $\mathcal{C}(\widehat{A})$, mais qu'elle sépare toujours $\widehat{A}$.*

Exercice IV .56 *Montrer que $\mathcal{G}$ est une isométrie de A dans $\mathcal{C}(\widehat{A})$ si et seulement si $\|x^2\| = \|x\|^2$ pour tout $x \in A$.*

Dans le cas de l'algèbre de Banach $A = \mathcal{C}(X)$, où $\widehat{A} \simeq X$, la transformée de Gelfand fait correspondre à l'élément $f \in \mathcal{C}(X)$ la fonction f elle-même, ce qui ne présente aucun intérêt. Vu l'Exercice IV .51, il en est de même pour l'algèbre $A = \mathcal{D}_m([0;1])$. En revanche, la profondeur et l'efficacité de la théorie vont apparaître dans son application à l'algèbre de convolution $\ell^1(\mathbf{Z})$.

IV.27 Théorème de Wiener

Reprenons l'algèbre $\ell^1(\mathbf{Z})$ munie du produit de convolution. C'est l'algèbre des suites $x = (x_n)_{n\in\mathbf{Z}}$ de nombres complexes, telles que $\sum_{n\in\mathbf{Z}} |x_n| < +\infty$, munie du produit de convolution :

$$x * y = (x * y)_n \quad \text{où} \quad (x * y)_n = \sum_{p\in\mathbf{Z}} x_{n-p} y_p \text{ pour tout } n \in \mathbf{Z}$$

Dans des exercices précédents, nous avons commencé à étudier cette algèbre de Banach commutative ; résumons ce que nous en savons déjà.

Pour tout $k \in \mathbf{Z}$, on définit une suite e_k par $(e_k)_n = 1$ si $n = k$ et 0 sinon. Alors on a $e_h * e_k = e_{h+k}$ pour tous $h, k \in \mathbf{Z}$, donc $(e_1)^n = e_n$ pour tout $n \in \mathbf{Z}$ et e_0 est l'élément neutre de A.

Si $x = (x_n)_{n\in\mathbf{Z}} \in \ell^1(\mathbf{Z})$, la série $\sum_{n\in\mathbf{Z}} x_n e_n = \sum_{n\in\mathbf{Z}} x_n (e_1)^n$ converge (absolument) vers x dans l'espace de Banach $\ell^1(\mathbf{Z})$, c'est-à-dire :

$$x = \lim_{N\to\infty} \sum_{n=-N}^{N} x_n (e_1)^n$$

au sens de la norme ℓ^1 ci-dessus. De plus, nous savons que le spectre de l'élément e_1 dans $\ell^1(\mathbf{Z})$ est le cercle unité du plan complexe :

$$\mathrm{sp}(e_1) = \mathbb{T} = \{\lambda \in \mathbf{C} : |\lambda| = 1\} = \{e^{i\theta} : \theta \in \mathbf{R}\}.$$

On va voir que le spectre de Gelfand de $\ell^1(\mathbf{Z})$ s'identifie à $\mathbb{T}$.

Pour tout $e^{i\theta} \in \mathbb{T}$ et toute $x = (x_n)_{n\in\mathbf{Z}} \in \ell^1(\mathbf{Z})$, posons :

$$\chi_{e^{i\theta}}(x) = \sum_{n=-\infty}^{\infty} x_n e^{in\theta},$$

ce qui a un sens, car la série converge (absolument) dans $\mathbf{C}$. Il est clair que $x \mapsto \chi_{e^{i\theta}}(x)$ est une forme linéaire non nulle sur $\ell^1(\mathbf{Z})$, car $\chi_{e^{i\theta}}(e_0) = 1$. Vérifions que c'est un caractère, c'est-à-dire qu'elle est multiplicative sur A. Si $x \in A$ et $y \in A$, on a :

$$\begin{aligned}\chi_{e^{i\theta}}(x)\chi_{e^{i\theta}}(y) &= \left(\textstyle\sum_{p=-\infty}^{\infty} x_p e^{ip\theta}\right)\left(\textstyle\sum_{q=-\infty}^{\infty} y_q e^{iq\theta}\right) &= \textstyle\sum_{p,q\in\mathbf{Z}} x_p y_q e^{i(p+q)\theta} \\ &= \textstyle\sum_{n\in\mathbf{Z}}\left(\textstyle\sum_{p+q=n} x_p y_q\right) e^{in\theta} &= \textstyle\sum_{n\in\mathbf{Z}} (x * y)_n e^{in\theta},\end{aligned}$$

ce qui vaut encore $\chi_{e^{i\theta}}(x * y)$; dans les calculs qui précèdent, on a pu regrouper les termes selon les $n = p + q$, vu le théorème sur la multiplication de Cauchy des séries absolument convergentes.

Ainsi tout $\chi_{e^{i\theta}}$ est un caractère de $\ell^1(\mathbf{Z})$; montrons qu'il n'y en a pas d'autres. Soit $\chi \in \widehat{A}$. Cherchons $e^{i\theta}$ tel que $\chi = \chi_{e^{i\theta}}$. La valeur candidate de $e^{i\theta}$ se trouve au moyen de e_1. Plus précisément, puisque $e_{-1} = (e_1)^{-1}$ on a $\chi(e_1)\chi(e_{-1}) = 1$ avec $|\chi(e_1)| \leqslant 1$ et $|\chi(e_{-1})| \leqslant 1$ car $\|e_1\| = \|e_{-1}\| = 1$, et donc $|\chi(e_1)| = 1$ (ce qui résulte aussi du fait que $\mathrm{sp}(e_1) = \mathbb{T}$). On peut donc écrire $\chi(e_1) = e^{i\theta}$ puis, par multiplicativité : $\chi(e_{-1}) = e^{-i\theta}$ et $\chi(e_n) = e^{in\theta}$ pour tout $n \in \mathbf{Z}$. On peut maintenant comparer χ et $\chi_{e^{i\theta}}$: ces deux caractères coïncident sur chaque e_n ($n \in \mathbf{Z}$), donc par linéarité sur toutes les combinaisons linéaires finies $\sum_{n=-N}^{N} x_n (e_1)^n$; donc par passage à la limite quand $N \to \infty$, on a : $\chi(x) = \chi_{e^{i\theta}}(x)$ pour tout $x \in \ell^1(\mathbf{Z})$. En résumé :

Proposition IV .57 *Les caractères de l'algèbre de Banach $A = \ell^1(\mathbf{Z})$ sont exactement les :*

$$x = (x_n)_{n\in\mathbf{Z}} \mapsto \chi_{e^{i\theta}}(x) = \sum_{n=-\infty}^{\infty} x_n e^{in\theta}.$$

On peut donc identifier le spectre de Gelfand $\widehat{A}$ avec le cercle unité $\mathbb{T} = \{e^{i\theta}\}_{\theta\in\mathbf{R}}$ de $\mathbf{C}$, par $\chi_{e^{i\theta}} \leftrightarrow e^{i\theta}$. Après cette identification, la transformée de Gelfand $\mathcal{G}x$ d'un élément $x = (x_n)_{n\in\mathbf{Z}} \in \ell^1(\mathbf{Z})$ est la fonction continue sur $\mathbb{T}$ définie par :

$$\mathcal{G}x(e^{i\theta}) = \sum_{n=-\infty}^{\infty} x_n e^{in\theta}.$$

Corollaire IV .58 *Pour que* $x = (x_n)_{n\in\mathbf{Z}} \in \ell^1(\mathbf{Z})$ *soit inversible dans* $\ell^1(\mathbf{Z})$ *il faut et il suffit que pour tout* $\theta \in \mathbf{R}$*, on ait :* $\sum_{n=-\infty}^{\infty} x_n e^{in\theta} \neq 0$*.*

Ce corollaire peut être reformulé plus classiquement, sous la forme d'un théorème initialement dû à N. Wiener.

Théorème IV .59 *Soit* $t \mapsto f(t)$ *une fonction continue de variable réelle, de période* 2π *et à valeurs complexes. Soient*

$$c_n(f) = \frac{1}{2\pi i}\int_{-\pi}^{\pi} f(t)e^{-int}\,\mathrm{d}t$$

ses coefficients de Fourier ($n \in \mathbf{N}$)*. On fait les deux hypothèses suivantes.*

(i) *La fonction* f *ne s'annule pas.*

(ii) *La série de Fourier de* f *est absolument convergente :*

$$\sum_{n\in\mathbf{Z}} |c_n(f)| < +\infty.$$

Alors, en posant $g(t) = \dfrac{1}{f(t)}$*, on a :*

$$\sum_{n\in\mathbf{Z}} |c_n(g)| < +\infty.$$

Autrement dit, si une fonction continue f sur $\mathbb{T}$ ne s'annule pas – ce qui permet de considérer la fonction inverse $\frac{1}{f}$ – et si f est dans la classe (très particulière) des fonctions à série de Fourier absolument convergente, alors $\frac{1}{f}$ est elle aussi dans cette classe.

Cet énoncé, découvert par Norbert Wiener en 1932, et démontré par de longs calculs de convolution, fut redémontré comme ci-dessus en quelques lignes par Israël Gelfand comme application de sa théorie des algèbres de Banach.

Preuve. Soit $x \in \ell^1(\mathbf{Z})$ telle que $x_n = c_n(f)$ pour tout $n \in \mathbf{Z}$. Alors x satisfait à l'hypothèse du Corollaire, car $f(\theta) = \sum_{n\in\mathbf{Z}} x_n e^{in\theta}$ n'est jamais nulle. Donc x est inversible dans $\ell^1(\mathbf{Z})$; soit $y = (y_n)_{n\in\mathbf{Z}}$ son inverse. Posons $g(\theta) = \sum_{n\in\mathbf{Z}} y_n e^{in\theta}$. Puisque $x * y = e_0$, par transformation de Gelfand on a : $g(\theta)f(\theta) = 1$ pour tout θ, donc $g = \frac{1}{f}$. Mais les coefficients de Fourier de g sont les y_n ; donc $\sum_{n\in\mathbf{Z}} |c_n(g)| < +\infty$. □

Exercice IV .60 *Soit* A *l'algèbre des sommes de séries de Taylor absolument convergentes* $f = f(z) = \sum_{n\geqslant 0} a_n z^n$*, munie de la norme* $\|f\| = \sum_{n=0}^{\infty} |a_n| < +\infty$*.*

1. *Montrer que les caractères de A sont exactement les $\chi_{z_0} : f \mapsto f(z_0)$ où $z_0 \in \mathbf{C}$ avec $|z_0| \leqslant 1$. Ceci permet d'identifier $\widehat{A}$ au disque unité fermé de rayon 1 de $\mathbf{C}$.*
2. *Montrer que si $f \in A$ n'est jamais nulle sur ce disque, alors il existe des nombres complexes b_n tels que*

$$\frac{1}{f(z)} = \sum_{n \geqslant 0} b_n z^n \quad \text{et} \quad \sum_{n=0}^{\infty} |b_n| < +\infty.$$

3. *Soient $f_1, f_2, \dots f_p$ des fonctions appartenant à A et n'ayant aucun zéro commun dans $\{z \in \mathbf{C} : |z| \leqslant 1\}$. Montrer qu'il existe $g_1, g_2, \dots g_p$ telles que $\sum_{i=1}^{p} f_i g_i = 1$.*

Exercice IV .61 *Soit $A = \mathcal{OC}(\overline{U})$ l'algèbre de l'Exemple IV .4. On pose $e_1(z) = z$.*

1. *Montrer que tout $f \in A$ est limite en norme de A de polynômes en e_1.*
2. *Montrer que les seuls caractères de A sont les évaluations aux points de $\overline{U}$.*
3. *Soient $f_1, f_2, \dots f_p$ des fonctions appartenant à A et n'ayant aucun zéro commun dans $\overline{U}$. Montrer qu'il existe $g_1, g_2, \dots g_p$ telles que $\sum_{i=1}^{p} f_i g_i = 1$.*
4. *Montrer que cette algèbre est strictement plus grande que celle de l'Exercice IV .60.*

Exercice IV .62 *Soit A une algèbre de Banach possédant un élément a (appelé générateur) tel que l'ensemble des $P(a)$, où $P \in \mathbf{C}[X]$, soit dense dans A. Montrer que dans ce cas $\chi \mapsto \chi(a)$ établit une bijection de $\widehat{A}$ sur $\mathrm{sp}_A(a)$. Reconsidérer les Exercices IV .60 et IV .61 dans cette perspective.*

Exercice IV .63 *Soit X un espace compact. Soit A une sous-algèbre de $\mathcal{C}(X)$. On fait les hypothèses suivantes*

1. *A sépare X ;*
2. *$1 \in A$;*
3. *A est auto-adjointe, i.e. $f \in A \Rightarrow \overline{f} \in A$;*
4. *si f ne s'annule jamais sur A, alors $\frac{1}{f} \in A$.*

Montrer qu'alors les seuls homomorphismes non identiquement nuls de l'algèbre A dans le corps $\mathbf{C}$ sont les homomorphismes d'évaluation $\chi_x : f \mapsto f(x)$, où x parcourt X. En particulier, si A est une algèbre de Banach, son spectre de Gelfand s'identifie à X.

Vis-à-vis du calcul fonctionnel holomorphe, la transformation de Gelfand se comporte comme on peut l'espérer.

Proposition IV .64 *Soit $x \in A$. Soit F une fonction holomorphe au voisinage de* $\mathrm{sp}(x)$ *et soit $F(x)$ l'élément qui correspond à F dans le calcul fonctionnel holomorphe. Alors*

$$\mathcal{G}F(x) = F \circ \mathcal{G}x.$$

Preuve. C'est un calcul. Pour tout $\chi \in \widehat{A}$, on a :

$$\begin{array}{rclcl} \mathcal{G}F(x)(\chi) & = & \chi\,(F(x)) & = & \langle \frac{1}{2\pi i}\vec{\int_\gamma}(\lambda e - x)^{-1}F(\lambda)\,\mathrm{d}\lambda|\chi\rangle \\ & = & \frac{1}{2\pi i}\int_\gamma \langle(\lambda e - x)^{-1}|\chi\rangle F(\lambda)\,\mathrm{d}\lambda & = & \frac{1}{2\pi i}\int_\gamma (\lambda z - \chi(x))^{-1} F(\lambda)\,\mathrm{d}\lambda \\ & = & F\,(\chi(x)) & = & F\,(\mathcal{G}x(\chi))\,, \end{array}$$

qui est bien $(F \circ \mathcal{G}x)(\chi)$. □

On en déduit une généralisation du Théorème de Wiener :

Théorème IV .65 *Soit $f = f(t)$ une fonction continue de variable réelle et de période 2π, dont la série de Fourier de f est absolument convergente :*

$$\sum_{n\in\mathbf{Z}} |c_n(f)| < +\infty.$$

Soit F une fonction holomorphe au voisinage de l'ensemble $f(\mathbb{T})$ des valeurs prises par f. Alors la fonction

$$(F \circ f) : t \mapsto F\,(f(t))$$

est encore une fonction 2π-périodique à série de Fourier de f absolument convergente.

Évidemment, pour $F(z) = \frac{1}{z}$ on retrouve le Théorème de Wiener.

Preuve. Sur $\mathbb{T} \simeq \widehat{\ell^1(\mathbf{Z})}$, on a :

$$f(\theta) = \sum_{n\in\mathbf{Z}} c_n(f)e^{in\theta} = \mathcal{G}x(\chi_{e^{i\theta}})$$

où $x = (c_n(f))_{n\in\mathbf{Z}}$ appartient à $\ell^1(\mathbf{Z})$ par hypothèse. Le spectre de x est l'ensemble $f(\mathbb{T})$ des valeurs prises par $\mathcal{G}x = f$, donc F est holomorphe au voisinage de $\mathrm{sp}(x)$; on peut donc appliquer la Proposition IV .64 :

$$(*) \quad (F \circ f)(\theta) = \mathcal{G}F(x)(\chi_{e^{i\theta}}) = \sum_{n\in\mathbf{Z}} d_n e^{in\theta}$$

avec $F(x) = (d_n)_{n\in\mathbf{Z}} \in \ell^1(\mathbf{Z})$, donc $\sum_{n\in\mathbf{Z}} |c_n(f)| < +\infty$. Mais, d'après $(*)$, on a $d_n = c_n(F \circ f)$ pour tout $n \in \mathbf{Z}$. □

Exercice IV .66 *Énoncer et démontrer un théorème de type Wiener-Lévy pour l'algèbre de l'Exercice IV .60.*

• Transformation de Fourier comme transformation de Gelfand

On cite ici des résultats sans les démontrer.

L'algèbre de Banach commutative la plus importante est peut-être l'algèbre de convolution $L^1(\mathbf{R})$ pour la norme $\|\cdot\|_1$ et le produit

$$(f_1 * f_2)(x) = \int_{\mathbf{R}} f_1(x-t) f_2(t)\, \mathrm{d}t.$$

Or cette algèbre n'a pas d'élément neutre (un tel élément neutre ne pourrait être que la mesure de Dirac, qui n'est pas une fonction), donc ce qui précède ne s'applique pas directement à $L^1(\mathbf{R})$.

Soit $\widetilde{A}$ l'algbèbre déduite de $L^1(\mathbf{R})$ par adjonction d'un élément neutre, définie dans l'Exemple IV .10. Pour tout $t \in \mathbf{R}$ et tout $(f, \lambda) \in \widetilde{A} = L^1(\mathbf{R}) \times \mathbf{C}$, la formule :

$$\chi_t(f, \lambda) = \int_{\mathbf{R}} f(x) e^{-ixt}\, \mathrm{d}x = \mathcal{F}f(t),$$

qui donne la transformée de Fourier de $f \in L^1(\mathbf{R})$, définit évidemment un caractère de $\widetilde{A}$, car $\mathcal{F}(f_1 * f_2) = \mathcal{F}(f_1)\mathcal{F}(f_2)$. On montre qu'il n'y a pas d'autre caractère de $\widetilde{A}$, excepté χ_∞ défini par $\chi_\infty(f, \lambda) = \lambda$.

Ainsi le spectre de Gelfand de $\widetilde{A}$ s'identifie à $\mathbf{R} \sqcup \{\infty\}$. Dans cette identification, la restriction à $\mathbf{R}$ de la transformée de Gelfand de (f, λ) n'est autre que la transformée de Fourier de la fonction f. Un théorème de type Wiener est le suivant : soient $a < b$ des nombres réels et soit $f \in L^1(\mathbf{R})$ telle que $\mathcal{F}f(t)$ ne soit nulle pour aucun $t \in [a; b]$. Alors il existe $g \in L^1(\mathbf{R})$ telle que $\mathcal{F}g(t) = \frac{1}{\mathcal{F}f(t)}$ pour tout $t \in [a; b]$.

On retrouvera toutes ces idées dans un contexte un peu moins difficile en résolvant l'exercice qui suit.

Exemple IV .67 *L'algèbre de Banach $L^1(\mathbb{T})$ des fonctions intégrables de période 2π, avec le produit de convolution*

$$(f_1 * f_2)(\theta) = \int_{-\pi}^{\pi} f_1(\theta - \varphi) f_2(\varphi)\, \mathrm{d}\varphi.$$

est une algèbre de Banach sans élément unité. Soit $\widetilde{A}$ l'alg ebre obtenue en lui adjoignant un élément unité. Montrer que, pour tout $n \in \mathbf{Z}$

$$(f, \lambda) \mapsto c_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta}\, \mathrm{d}\theta$$

est un caractère de $\widetilde{A}$, et qu'il n'y a pas d'autre caractère de $\widetilde{A}$ hormis c_∞ défini par $c_\infty(f, \lambda) = \lambda$. Ainsi le spectre de Gelfand de $\widetilde{A}$ s'identifie-t-il à $\mathbf{Z} \sqcup \{\infty\}$. En utilisant la mesure de Dirac, pouvez-vous énoncer un théorème de type Wiener dans cette situation ?

Cinquième partie

Théorie spectrale des opérateurs compacts

Dans le chapitre sur les algèbres de Banach, nous avons étudié le spectre en général. Nous allons approfondir cette notion dans le cas d'un opérateur T dans un espace de Banach. Une partie de $\mathrm{sp}(T)$ est formée de *valeurs propres* de T. Si E est de dimension finie, $\mathrm{sp}(T)$ est tout entier formé de valeurs propres de T : c'est l'ensemble (fini) des racines du polynôme caractéristique de l'endomorphisme T.

Si E est de dimension quelconque, finie ou infinie, l'opérateur T est dit *compact* si, pour toute suite $(f_n)_{n\geqslant 0}$ bornée dans E, on peut extraire de la suite $(Tf_n)_{n\geqslant 0}$ une suite de Cauchy dans E. Alors T a des propriétés spectrales remarquables : tout $\lambda \in \mathrm{sp}(T)$ non nul est en fait une valeur propre de T ; ces valeurs propres sont en nombre fini ou forment un ensemble dénombrable $(\lambda_n)_{n\geqslant 0}$ avec $\lim_{n\to\infty} \lambda_n = 0$; leurs sous-espaces propres sont de dimension finie.

Les opérateurs intégraux $f \mapsto Tf$ dans $\mathcal{C}([a;b])$, dits *de Fredholm* et définis par :

$$Tf(x) = \int_a^b K(x,y)f(y)\,\mathrm{d}y$$

sont les exemples les plus intéressants d'opérateurs compacts. Leur théorie spectrale permet de résoudre les équations intégrales associées (alternative de Fredholm, équations de Volterra, d'Abel). À cette théorie se rattache naturellement celle de Sturm-Liouville.

V.28 Valeurs spectrales et valeurs propres

Dans ce chapitre, E désigne un espace de Banach sur $\mathbf{C}$. On note simplement $\mathcal{L}(E)$ l'algèbre de Banach $\mathcal{L}(E,E)$ des opérateurs bornés T dans E, c'est-à-dire des applications linéaires continues $E \to E$, avec la norme :

$$|||T||| = \sup_{f\in E, \|f\|\leqslant 1} \|Tf\|.$$

Cette algèbre, non commutative dès que $\dim E \geqslant 2$, a un élément neutre, l'opérateur identique id_E.

Definition V .1 *Soit T un opérateur dans E et soit $\lambda \in \mathbf{C}$.*

(i) *On dit que λ est une valeur spectrale de T, ou que T appartient au spectre $\mathrm{sp}(T)$ de T, si – conformément à la théorie des algèbres de Banach – l'opérateur $T - \lambda\mathrm{id}_E$ n'est pas inversible dans $\mathcal{L}(E)$; autrement dit, s'il n'existe pas d'opérateur borné $S \in \mathcal{L}(E)$ tel que $(T-\lambda\mathrm{id}_E)\circ S = S\circ(T-\lambda\mathrm{id}_E) = \mathrm{id}_E$.*

(ii) *On dit que λ est valeur propre de T si $T - \lambda\mathrm{id}_E$ n'est pas injectif dans E ; autrement dit, s'il existe $f \in E\backslash\{0\}$ tel que $Tf = \lambda f$. De tels vecteurs f sont appelés des vecteurs propres de T pour la valeur propre λ.*

(iii) *On dit que λ est une valeur propre généralisée de T s'il existe une constante $c > 0$ et une suite $(f_n)_{n\geqslant 0}$ dans E telle que $\|f_n\| \geqslant c$ pour tout $n \geqslant 0$, mais avec $\lim\limits_{n\to\infty} Tf_n - \lambda f_n = 0$.*

Si E est un espace de fonctions, on parle de *fonction propre* plutôt que de vecteur propre. Dans tous les cas, on omet souvent de préciser la valeur propre pour un vecteur propre.

Rappel : on a vu dans un contexte plus général :

1. que $\mathrm{sp}(T)$ est un sous-ensemble compact et non vide de $\mathbf{C}$, contenu dans le disque $\{\lambda \in \mathbf{C} : |\lambda| \leqslant \|T\|\}$;
2. que la fonction $\lambda \mapsto \langle (T - \lambda \mathrm{id}_E)^{-1}, \varphi\rangle$ est, pour toute forme linéaire continue φ sur $\mathcal{L}(E)$, holomorphe dans l'ouvert de $\mathbf{C}$ complémentaire de $\mathrm{sp}(T)$;
3. qu'on a la formule du rayon spectral :
$$\sup_{\lambda\in\mathrm{sp}(T)} |\lambda| = \lim_{n\to\infty} |||T|||^{n\,\frac{1}{n}};$$
4. que, si pour F holomorphe au voisinage de $\mathrm{sp}(T)$, on définit l'opérateur $F(T)$ dans E par :
$$F(T) = \frac{1}{2\pi i}\vec{\int_\gamma} (T - \lambda \mathrm{id}_E)^{-1} F(\lambda)\,\mathrm{d}\lambda,$$
alors $F \mapsto F(T)$ est un homomorphisme d'algèbres de $\mathcal{H}\left(\mathrm{sp}(T)\right)$ dans $\mathcal{L}(E)$ (calcul fonctionnel holomorphe) ;
5. et que $\mathrm{sp}\left(F(T)\right) = F\left(\mathrm{sp}(T)\right)$ (théorème de Hilbert-Dirac généralisé).

Ce qui apparaît en plus dans $\mathcal{L}(E)$, c'est la notion de valeur propre.

Proposition V .2 *Soit T un opérateur dans E et soit $\lambda \in \mathbf{C}$.*
(i) *Si λ est une valeur propre de T, alors λ est une valeur propre généralisée de T.*
(ii) *Si λ est une valeur propre généralisée de T, alors λ est une valeur spectrale de T.*
(iii) *Si λ appartient à la frontière de* $\mathrm{sp}(T)$*, alors λ est une valeur propre généralisée de T.*

Preuve. Les points (i) et (ii) sont évidents, mais ça va mieux en l'écrivant.

Prouvons le point (iii). Soit λ un point frontière du compact $\mathrm{sp}(T)$. Soit $(\lambda_n)_{n\geqslant 0}$ une suite qui tend vers λ mais avec $\lambda_n \notin \mathrm{sp}(T)$ pour tout $n \geqslant 0$.

Ainsi chaque opérateur $T - \lambda_n \mathrm{id}_E$ est inversible dans E, alors que $T - \lambda \mathrm{id}_E$ ne l'est pas. Puisque $T - \lambda \mathrm{id}_E = T - \lambda_n \mathrm{id}_E + (\lambda_n - \lambda)\mathrm{id}_E$, on en déduit que

$$|\lambda - \lambda_n| \geqslant \frac{1}{|||(T - \lambda_n \mathrm{id}_E)^{-1}|||},$$

sinon $T - \lambda \mathrm{id}_E$ serait inversible, au titre d'élément suffisamment proche de l'inversible $T - \lambda_n \mathrm{id}_E$ (cf Théorème IV .12). Ceci prouve que $|||(T - \lambda_n \mathrm{id}_E)^{-1}|||$ tend vers $+\infty$ quand $n \to \infty$. Posons

$$S_n = \frac{(T - \lambda_n \mathrm{id}_E)^{-1}}{|||(T - \lambda_n \mathrm{id}_E)^{-1}|||}.$$

On a $|||S_n||| = 1$, donc il existe des $g_n \in E$ tels que $\|g_n\| = 1$ et $\frac{1}{2} \leqslant \|S_n g_n\| \leqslant 1$ pour tout $n \geqslant 0$. Posons $f_n = S_n g_n$. Alors $\|f_n\| \geqslant \frac{1}{2} = c > 0$ et

$$\begin{aligned}
(T - \lambda \mathrm{id}_E) f_n &= (T - \lambda_n \mathrm{id}_E) f_n + (\lambda_n - \lambda) f_n \\
&= (T - \lambda_n \mathrm{id}_E) S_n g_n + (\lambda_n - \lambda) f_n \\
&= \frac{1}{|||(T - \lambda_n \mathrm{id}_E)^{-1}|||} g_n + (\lambda_n - \lambda) f_n
\end{aligned}$$

tend vers 0 car $\|f_n\| \leqslant \|g_n\| \leqslant 1$. □

Exercice V .3 *Soit S l'opérateur de décalage défini et étudié dans l'exercice IV .14. On sait que le spectre de S est le disque unité fermé de S. Montrer que S n'a pas de valeur propre et que le cercle unité est l'ensemble des valeurs propres généralisées de S.*

V.29 Notion d'opérateur compact

Soit E un espace de Banach sur $\mathbf{C}$ et soit $T \in \mathcal{L}(E)$. On dit que T est un *opérateur compact* (ou *complètement continu*) dans E si :

- quelle que soit la suite $(f_n)_{n \geqslant 0}$ bornée dans E on peut extraire de $(T f_n)_{n \geqslant 0}$ une sous-suite convergente ;

autrement dit, si :

- l'image par T de la boule unité de E est d'adhérence compacte.

On notera $\mathcal{LK}(E)$ l'ensembles des opérateurs compacts dans E. Tout opérateur de rang fini, c'est-à-dire de la forme :

$$f \mapsto T(f) = \alpha_1(f) g_1 + \alpha_2(f) g_2 + \cdots + \alpha_n(f) g_n,$$

où $g_1, g_2, \cdots g_n$ sont des vecteurs donnés dans E et $\alpha_1, \alpha_2, \cdots \alpha_n$ sont des formes linéaire continues données sur E, est compact.

Ceci provient du fait que $T(E)$ est un espace vectoriel de dimension finie, donc l'image par T de la boule unité de E est une partie bornée d'un espace normé de dimension finie : cette image a une adhérence compacte par le théorème de Borel-Lebesgue.

À partir des opérateurs de rang fini, on peut fabriquer beaucoup d'opérateurs compacts au moyen du résultat suivant.

Théorème V .4 *Soit* $T \in \mathcal{L}(E)$ *et soit* $(T_n)_{n\geqslant 0}$ *une suite d'opérateurs compacts dans* E*, telle que* $\lim_{n\to\infty} |||T - T_n||| = 0$*. Alors* T *est un opérateur compact.*

Preuve. On va utiliser le procédé diagonal. Soit $(f_n)_{n\geqslant 0}$ une suite dans E telle que $\|f_n\| \leqslant 1$ pour tout n. On va montrer que la suite $(Tf_n)_{n\geqslant 0}$ admet une sous-suite de Cauchy dans E.

Puisque T_1 est un opérateur compact, on peut extraire de $(f_n)_{n\geqslant 0}$ une sous-suite $(f_n^1)_{n\geqslant 0}$ telle que la suite $(T_1 f_n^1)_{n\geqslant 0}$ soit de Cauchy. Puis, T_2 étant compact, on peut extraire de la suite $(f_n^1)_{n\geqslant 0}$ une sous-suite $(f_n^2)_{n\geqslant 0}$ telle que $(T_2 f_n^2)_{n\geqslant 0}$ soit de Cauchy, etc. Par récurrence, on construit pour tout entier $m \geqslant 1$ une suite $(f_n^m)_{n\geqslant 0}$ extraite de $(f_n^{m-1})_{n\geqslant 0}$ et telle que la suite $(T_m f_n^m)_{n\geqslant 0}$ soit de Cauchy.

Considérons la suite diagonale $(f_n^n)_{n\geqslant 0}$: elle est extraite de la suite initiale $(f_n)_{n\geqslant 0}$, et pour tout $m \geqslant 1$ fixé, la suite $(T_m f_n^n)_{n\geqslant 0}$ est de Cauchy. Pour $\varepsilon > 0$ donné, fixons un indice m tel que $|||T - T_m||| \leqslant \frac{\varepsilon}{4}$. Quels que soient les indices $q \geqslant p$, on a :

$$T(f_q^q) - T(f_p^p) = (T - T_m)(f_q^q - f_p^p) + T_m(f_q^q) - T_m(f_p^p),$$

donc

$$\|T(f_q^q) - T(f_p^p)\| \leqslant 2 \cdot \frac{\varepsilon}{4} + \|T_m(f_q^q) - T_m(f_p^p)\| \leqslant \varepsilon$$

dès que $q \geqslant p \geqslant N_\varepsilon$, vu que la suite $(T_m f_n^n)_{n\geqslant 0}$ est de Cauchy. □

Remarquons que id_E est compact si, et seulement si, E est de dimension finie, car d'après le théorème de F. Riesz, la finitude de la dimension est la condition nécessaire et suffisante pour que la boule unité de E soit compacte.

Remarquons aussi que $\mathcal{LK}(E)$ est un *idéal bilatère* de $\mathcal{L}(E)$: si T est compact et si $S \in \mathcal{L}(E)$, alors ST et TS sont des opérateurs compacts ; car l'image par S de la boule unité de E est borné, et l'image par S d'un compact est un compact.

Des deux remarques précédentes, il résulte que dans un espace E de dimension infinie, un opérateur compact n'est jamais inversible ; dans ce cas, 0 est toujours dans le spectre de l'opérateur.

Mais les opérateurs compacts les plus importants en Analyse sont les opérateurs intégraux que nous allons définir maintenant.

Soit $K = K(x,y)$ une fonction continue donnée, à valeurs complexes, sur le carré $[a;b] \times [a;b]$ de $\mathbf{R}^2$. Dans l'espace de Banach $E = \mathcal{C}([a;b],\mathbf{C})$ des fonctions continues sur $[a;b]$, à valeurs complexes, muni de la norme uniforme $\|\cdot\|_\infty$, l'*opérateur de Fredholm* de noyau K, à savoir $f \mapsto Tf$, où

$$Tf(x) = \int_a^b K(x,y)f(y)\,\mathrm{d}y,$$

est compact. En effet, l'ensemble des fonctions Tf, où $\|f\|_\infty \leqslant 1$, est équicontinu, car :

$$|Tf(x) - TF(y)| \leqslant (b-a) \sup_{y\in[a;b]} |K(x,y) - K(x',y)| \leqslant \varepsilon$$

dès que $|x-x'| \leqslant \eta$, vu l'uniforme continuité de K dans le carré $[a;b]\times[a;b]$ Donc l'image par T de la boule unité $\{\|f\|_\infty \leqslant 1\}$ est d'adhérence compacte dans $E = \mathcal{C}([a;b],\mathbf{C})$, par le Théorème d'Ascoli.

Dans le même espace $E = \mathcal{C}([a;b],\mathbf{C})$ est aussi compact l'opérateur de Volterra, à savoir $f \mapsto Sf$ défini par :

$$Sf(x) = \int_a^x K(x,y)f(y)\,\mathrm{d}y \quad (a \leqslant x \leqslant b)$$

comme on le voit aussi par le Théorème d'Ascoli.

Exercice V .5 *Soit $1 \leqslant p \leqslant +\infty$. On prend pour E l'espace $\ell^p(\mathbf{Z})$. Soit $(a_n)_{n\geqslant 1}$ une suite bornée de nombres complexes donnée. Dans l'espace de Banach ℓ^p soit $x \mapsto y = Tx$ l'opérateur défini par $y_n = a_n x_n$ pour tout $n \geqslant 1$. Montrer que T est compact si et seulement si $\lim_{n\to\infty} a_n = 0$. Indications : on prouvera d'abord que $|||T||| = \sup_n |a_n|$; si $\lim_{n\to\infty} a_n = 0$, on fera apparaître T comme limite d'opérateurs de rang fini.*

V.30 Propriétés spectrales des opérateurs compacts

Dans un espace E de dimension finie, l'Algèbre linéaire nous a appris qu'un opérateur injectif est automatiquement bijectif, donc n'a pas d'autres valeurs spectrales que le valeurs propres, et que ces valeurs propres sont en nombre fini, étant les racines du polynôme caractéristique.

Dans un espace E de dimension quelconque, nous allons voir que les opérateurs compacts ont des propriétés spectrales très particulières, qui généralisent celles des opérateurs en dimension finie.

Lemme V .6 *Soit $T \in \mathcal{LK}(E)$. Soit λ une valeur propre généralisée non nulle de T. Alors λ est une valeur propre de T.*

Preuve. Par hypothèse il existe une suite $(f_n)_{n\geqslant 0}$ telle que $\|f_n\| \geqslant c > 0$ et $\lim_{n\to\infty} (Tf_n - \lambda f_n) = 0$. En divisant f_n par $\|f_n\|$ on se ramène à $\|f_n\| = 1$ pour tout n. De la suite $(Tf_n)_{n\geqslant 0}$ on peut extraire une suite convergente : $\lim_{k\to\infty} Tf_{n_k} = g$. Évidemment $(\lambda f_{n_k})_{k\geqslant 0}$ tend vers λg. On a :

$$\|g\| = |\lambda| \lim_{k\to\infty} \|f_{n_k}\| = |\lambda| \neq 0,$$

donc $g \neq 0$. Enfin, par continuité de T on a :

$$Tg = \lim_{k\to\infty} T(\lambda f_{n_k}) = \lambda \lim_{k\to\infty} T(f_{n_k}) = \lambda g,$$

ce qui prouve que λ est une valeur propre de T. $\square$

Lemme V .7 *Soit $T \in \mathcal{LK}(E)$. Alors pour tout nombre réel $c > 0$, l'opérateur T n'a qu'un nombre fini de valeurs propres distinctes de module $> c$.*

Preuve. On raisonne par l'absurde, en se donnant une suite $(\lambda_n)_{n\geqslant 0}$ de valeurs propres distinctes telles que $|\lambda_n| > c > 0$ pour tout n et où c ne dépend pas de n. Soit $(e_n)_{n\geqslant 0}$ une suite de vecteurs propres associés : $Te_n = \lambda_n e_n$ et $e_n \neq 0$ pour tout $n \geqslant 0$. Les vecteurs e_n forment une famille libre car les valeurs propres sont deux à deux distinctes. Soit L_n le sous-espace vectoriel engendré par les vecteurs e_1 à e_n. Le vecteur e_n n'appartient pas à L_{n-1} et par ailleurs, par le lemme de Riesz, il existe $h_n \in L_n$ tel que $\|h_n\|= 1$ et $\|h_n - f\| \geqslant \frac{1}{2}$ pour tout $f \in L_{n-1}$. Il existe $\alpha \in \mathbf{C}$ et $f_0 \in L_{n-1}$ tels que $h_n = f_0 + \alpha e_n$. On a :

$$Th_n = Tf_0+\alpha Te_n = Tf_0+\lambda_n\alpha e_n = Tf_0+\lambda_n(h_n-f_0) = (T-\lambda_n)f_0 = \lambda_n(h_n-f_n),$$

où $f_n \in L_{n-1}$. Donc, pour $m < n$, on a :

$$Th_n - Th_m = \lambda_n(h_n - f),$$

où $f \in L_{n-1}$. Par suite $\|Th_n - Th_m\| = |\lambda_n| \cdot \|h_n - f\| \geqslant \frac{c}{2}$ pour tout $n > m$, ce qui évidemment interdit à toute suite extraite de $(Th_n)_{n\geqslant 0}$ d'être de Cauchy, contrairement à l'hypothèse suivant laquelle l'opérateur T est compact. $\square$

Lemme V .8 *Soit $T \in \mathcal{LK}(E)$. Soit $\lambda \in \mathrm{sp}(T)\backslash\{0\}$. Alors λ est valeur propre de T.*

Preuve. D'après la Proposition V .2 (iii) et le Lemme V .6, c'est déjà vrai pour tout $\lambda \neq 0$ point frontière de $\mathrm{sp}(T)$. De plus, la Proposition V .2 (iii) et le Lemme V .7 impliquent qu'étant donné $c > 0$, il n'y a qu'un nombre fini de points frontière de $\mathrm{sp}(T)$ de module $> c$. Raisonnons par l'absurde, et supposons que $\mathrm{sp}(T)$ contienne un point intérieur $\lambda \neq 0$. Posons $c = |\lambda|$, et soient $\lambda_1, \lambda_2, \ldots \lambda_n$ les points frontière de $\mathrm{sp}(T)$ de module $> c$. Par λ menons une demi-droite extérieure au disque centré en 0 de rayon c, et qui évite les λ_i. Puisque λ est intérieur à $\mathrm{sp}(T)$, qui est compact, cette demi-droite devrait couper la frontière de $\mathrm{sp}(T)$, ce qui est impossible car cela fournirait un point frontière de module $> c$ supplémentaire. □

Les lemmes ont déjà prouvé l'essentiel du résultat qui suit.

Théorème V .9 *Soit T un opérateur compact dans un espace de Banach. Alors toute valeur spectrale non nulle, i.e. tout $\lambda \neq 0$ tel que $T - \lambda \mathrm{id}_E$ est non inversible, est en fait une valeur propre de T, et le sous-espace propre associé*

$$V_\lambda = \mathrm{Ker}(T - \lambda \mathrm{id}_E)$$

est de dimension finie. De plus, l'ensemble de ces valeurs propres est au plus dénombrable et 0 en est le seul point d'accumulation possible.

Preuve. Il reste seulement à prouver que $\dim V_\lambda < +\infty$. Or, V_λ est stable par T, et donc $T|_{V_\lambda}$ est un opérateur à la fois compact et inversible (car homothétie de rapport $\lambda \neq 0$) de V_λ. On a vu, grâce au théorème de Riesz, que ceci impose $\dim V_\lambda < +\infty$. □

Exercice V .10 *Dans l'exercice V .5, quelles sont les valeurs propres des T ? Quels sont les vecteurs propres correspondants ?*

Exercice V .11 *Montrer que dans $\mathcal{C}([0; \frac{\pi}{2}], \mathbf{C})$, l'opérateur de Fredholm $f \mapsto Tf$ où*

$$Tf(x) = \int_0^{\frac{\pi}{2}} \cos(x - y) f(y) \, \mathrm{d}y,$$

est de rang 2. Quelles sont les valeurs propres non nulles de T et les vecteurs propres associés ?

V.31 Alternative de Fredholm

C'est l'analogue de ce qui se passe en Algèbre linéaire de dimension finie. Rappelons ce qu'on appelle dans ce cas l'*alternative de Cramer*. Soit A une matrice complexe $n \times n$. Alors :

- ou bien : quel que soit $\vec{b} \in \mathbf{C}^n$ l'équation $A\vec{x} = \vec{b}$ a une solution et une seule ;
- ou bien : l'équation $A\vec{x} = \vec{0}$ a une solution non nulle.

En effet, on est dans le premier cas si et seulement si $\det A \neq 0$; dans le second si et seulement $\det A = 0$. Et il faut bien que l'un des deux ait lieu.

Proposition V .12 *Soit T un opérateur compact dans un espace de Banach E sur $\mathbf{C}$. Alors :*

- *ou bien : quel que soit $g \in E$ l'équation $(\mathrm{id}_E - \mu T)f = g$ a une solution f et une seule dans E ;*
- *ou bien : l'équation $(\mathrm{id}_E - \mu T)f = 0$ a une solution f non nulle.*

L'intérêt de l'énoncé est que le second terme de l'alternative de Fredholm est plus facile à discuter que le premier.

Preuve. Si $\mu = 0$, on est trivialement dans le premier cas, et pas dans le second. Désormais on suppose que $\mu \neq 0$, on pose $\lambda = \frac{1}{\mu}$ et on étudie l'équation

$$(T - \lambda \mathrm{id}_E)f = -\lambda g$$

d'inconnue f, qui est équivalente à la précédente. On est nécessairement dans l'un des deux cas suivants (qui s'excluent mutuellement) :

— ou bien $\lambda \notin \mathrm{sp}(T)$, donc $T - \lambda \mathrm{id}_E$ est inversible, et on est dans le premier cas de l'alternative ;

— ou bien $\lambda \in \mathrm{sp}(T)$, donc par le Théorème V .9 λ est valeur propre de T, et on est dans le second cas de l'alternative.

Ceci prouve l'énoncé cherché. □

Reprenons les exemples des opérateurs intégraux.

Proposition V .13 *Soit $K = K(x, y)$ un noyau continu sur $[a; b] \times [a; b]$ et soit $\mu \in \mathbf{C}$. Alors :*

— *ou bien l'équation intégrale*

$$f(x) = \mu \int_a^b K(x, y)f(y)\,\mathrm{d}y + g(x)$$

a, quel que soit $g \in \mathcal{C}([a; b], \mathbf{C})$, une solution $f \in \mathcal{C}([a; b], \mathbf{C})$ et une seule ;

— *ou bien l'équation homogène associée*

$$f(x) = \mu \int_a^b K(x, y)f(y)\,\mathrm{d}y$$

a une solution non identiquement nulle. □

Le second cas ne peut se produire que pour un ensemble fini, ou au plus dénombrable, de valeurs $\mu \in \mathbf{C}$ (qui dans ce dernier cas tendent vers l'infini). Ces valeurs exceptionnelles de μ sont souvent des valeurs physiquement intéressantes (phénomènes de résonance).

Exercice V .14 *Pour toute fonction* $g\mathcal{C}([0;\pi],\mathbf{C})$*, l'équation intégrale*

$$f(x)-\int_0^\pi \sin(x+y)f(y)\,\mathrm{d}y=g(x)$$

a une solution et une seule.

Pour l'équation de Volterra, la dichotomie devient drastique : le second cas n'a jamais lieu ! On a d'abord le lemme suivant :

Lemme V .15 *Soit* $K=K(x,y)$ *un noyau continu sur* $[a;b]\times[a;b]$ *et soit* $f\mapsto Sf$ *l'opérateur de Volterra* $S:\mathcal{C}([a;b],\mathbf{C})\to\mathcal{C}([a;b],\mathbf{C})$ *défini par*

$$Sf(x)=\int_a^x K(x,y)f(y)\,\mathrm{d}y.$$

Alors $\lambda=0$ *est la seule valeur propre possible de* S*. On a donc :* $\mathrm{sp}(S)=\{0\}$*.*

Proposition V .16 *Soit* $K=K(x,y)$ *un noyau continu sur* $[a;b]\times[a;b]$*. Alors, quelle que soit* $g\in\mathcal{C}([a;b],\mathbf{C})$*, l'équation de Volterra :*

$$f(x)-\int_a^x K(x,y)f(y)\,\mathrm{d}y=g(x)$$

a une solution $f\in\mathcal{C}([a;b],\mathbf{C})$ *et une seule.*

En effet, le Lemme ci-dessus indique que, dans la Proposition V .13 avec $\mu=1$, seul le premier volet de l'alternative se présente. □

Preuve du Lemme. Soient $\lambda\neq 0$ et $f\in\mathcal{C}([a;b],\mathbf{C})$ tels que $Sf(x)=\lambda f(x)$. Raisonnons par l'absurde et supposons que f n'est pas la fonction nulle. Puisque $Sf(a)=\int_a^a K(x,y)f(y)\,\mathrm{d}y=0$, on a $f(a)=0$. On se ramène au cas où f n'est pas identiquement nulle au voisinage de a, quitte à changer a en le repoussant à droite jusqu'à cela se produise. Posons :

$$m(\delta)=\sup_{a\leqslant x\leqslant a+\delta}|f(x)|.$$

Le nombre $m(\delta)$ est non nul pour tout $\delta>0$, et tend vers 0 quand $\delta>0$ tend vers 0. De plus, il existe $x_\delta\in[a;a+\delta]$ tel que $|f(x_\delta)|=m(\delta)$. Alors

$$\begin{array}{rclcl}|\lambda|m(\delta)=|\lambda f(x_\delta)| & = & |Sf(x_\delta)| & = & |\int_a^{x_\delta}K(x,y)f(y)\,\mathrm{d}y|\\ & \leqslant & \int_a^{x_\delta}|K(x,y)||f(y)|\,\mathrm{d}y & \leqslant & C\delta m(\delta),\end{array}$$

où $C=\sup_{a\leqslant x,y\leqslant b}|K(x,y)|$. Mais cette inégalité $|\lambda|m(\delta)\leqslant C\delta m(\delta)$, c'est-à-dire $|\lambda|\leqslant C\delta$, est absurde dès que $\delta>0$ est assez petit, vu que $\lambda\neq 0$. □

Exercice V .17 *Exprimer par une quadrature l'unique solution f de l'équation de Volterra :*

$$f(x) - \int_0^x f(y)\,\mathrm{d}y = g(x).$$

En déduire explicitement f quand $g(x) = 1$, quand $g(x) = e^x$.

Exercice V .18 *Soit $E = \mathcal{C}([0;1], \mathbf{C})$ et*

$$K(x,y) = \min\{x;y\}.$$

Si $f \in E$, on pose :

$$g(x) = Tf(x) = \int_0^1 K(x,y)f(y)\,\mathrm{d}y = \int_0^x f(y)\,\mathrm{d}y + x\int_x^1 f(y)\,\mathrm{d}y.$$

1. *Montrer que g est de classe C^2 sur $[0;1]$, que $g''(x) = -f(x)$, et que $g(0) = 0$ et $g'(1) = 0$.*
2. *Soit F le sous-espace vectoriel de E des $g \in E$ de classe C^2 telles que $g(0) = 0$ et $g'(1) = 0$. Montrer que T établit une bijection de E sur F. Préciser l'application $T^{-1} : F \to E$.*
3. *Dans E, soit le produit scalaire :*

   $$(f|g) = \int_0^1 f(x)\overline{g(x)}\,\mathrm{d}x.$$

   Montrer que $(Tf|g) = (f|Tg)$. En déduire que les valeurs propres de T sont réelles.
4. *Vérifier que T n'a pas de valeur propre $\leqslant 0$. Montrer que les valeurs propres de T sont les nombres $\lambda_n = (\frac{\pi}{2} + n\pi)^{-2}$ pour n entier $\geqslant 0$. Déterminer le sous-espace vectoriel $V_{\lambda_n} = \mathrm{Ker}(T - \lambda_n \mathrm{id}_E)$.*

V.32 Cas d'un opérateur compact auto-adjoint

Dans ce paragraphe, E est un espace de Hilbert sur $\mathbf{C}$ dont le produit scalaire est noté $(\cdot|\cdot)$ et la norme associée $\|\cdot\|$; on a $\|f\| = \sqrt{(f|f)}$ pour tout $f \in E$.

Rappelons qu'un opérateur borné $T \in \mathcal{L}(E)$ est dit *auto-adjoint* , ou *hermitien*, si quels que soient f_1 et f_2 dans E on a :

$$(Tf_1|f_2) = (f_1|Tf_2).$$

Les faits suivants sur les opérateurs auto-adjoints sont standard. Soit T un opérateur auto-adjoint.

(i) Toutes les valeurs propres de T sont réelles.
(ii) Deux vecteurs propres pour des valeurs propres distinctes sont orthogonaux.
(iii) Si V est un sous-espace vectoriel T-stable, alors son orthogonal $T^{\perp}$ l'est aussi.

Exercice V .19 *Redémontrer ces faits.*

Nous y ajouterons les propriétés suivantes.

Proposition V .20 *Soit T un opérateur auto-adjoint dans un espace de Hilbert E. Alors :*

$$|||T^2||| = |||T|||^2 \quad \text{et} \quad |||T||| = \sup_{\lambda \in \mathrm{sp}(T)} |\lambda|.$$

Donc si $T \neq 0$, alors $\mathrm{sp}(T) \neq \{0\}$. En particulier, si T est auto-adjoint compact et non nul, alors T a des valeurs propres $\neq 0$.

Preuve. Pour tout f avec $\|f\| \leqslant 1$, on a par l'inégalité de Cauchy-Schwarz :

$$\|Tf\|^2 = (Tf|Tf) = (T^2f|f) \leqslant \|T^2f\| \cdot \|f\| \leqslant \|T^2f\|,$$

donc

$$|||T|||^2 = \sup_{\|f\| \leqslant 1} \|Tf\|^2 \leqslant \sup_{\|f\| \leqslant 1} \|T^2f\| = |||T^2|||.$$

L'inégalité en sens contraire est banale.

Ainsi, on a :

$$|||T||| = |||T^2|||^{\frac{1}{2}} = |||T^4|||^{\frac{1}{4}} = \cdots = |||T^{2^n}|||^{\frac{1}{2^n}}$$

pour tout n ; or, quand $n \to \infty$, par la formule du rayon spectrale on a : $\lim_{n\to\infty} |||T^{2^n}|||^{\frac{1}{2^n}} = \sup_{\lambda \in \mathrm{sp}(T)} |\lambda|$. Ceci fournit bien $|||T||| = \sup_{\lambda \in \mathrm{sp}(T)} |\lambda|$, ce qu'il fallait démontrer. □

Théorème V .21 *Soit T un opérateur non nul, auto-adjoint et compact dans un espace de Hilbert E. Soit $(\lambda_n)_{1 \leqslant n < p}$ la suite finie ou infinie (avec alors $p = \infty$) des valeurs propres (nécessairement réelles) non nulles de T, chacune comptée autant de fois que sa multiplicité, i.e. $\dim V_\lambda$ avec $V_\lambda = \mathrm{Ker}(T - \lambda \mathrm{id}_E)$. Soit $(e_n)_{1 \leqslant n < p}$ une famille orthonormée de vecteurs de E, avec $Te_n = \lambda_n e_n$ pour tout n. Alors :*

(i) *Pour tout $f \in E$, on a la formule :*

$$Tf = \sum_{1 \leqslant n < p} \lambda_n (f|e_n) e_n$$

où, si $p = \infty$, la série converge en norme dans E.

(ii) *Soit μ un nombre complexe distinct de tous les inverses $\frac{1}{\lambda_n}$. Alors, pour tout $g \in E$, l'équation*

$$f - \mu \cdot T(f) = g$$

a une solution $f \in E$ et une seule, qui est donnée par la formule :

$$f = g + \mu \sum_{1 \leqslant n < p} \frac{\lambda_n}{1 - \mu\lambda_n} (g|e_n) e_n$$

où, si $p = \infty$, la série converge en norme dans E.

Preuve. Soit V le sous-espace de Hilbert de E engendré par les e_n : il est T-stable, donc $V^\perp$ l'est également. On a : $E = V \oplus V^\perp$ et la restriction $T|_{V^\perp}$ est un opérateur auto-adjoint compact de $V^\perp$, sans valeur propre $\neq 0$ par construction de V. La Proposition V .20 impose donc $T|_{V^\perp} = 0$.

Soit $f \in E$. Les e_n formant une base orthonormée de V, on a une décomposition

$$f = \sum_{1 \leqslant n < p} (f|e_n) e_n + g$$

avec $g \in V^\perp$. En appliquant T aux deux membres, et en remarquant que $Tg = 0$, on obtient la formule annoncée, ce qui finit de prouver (i).

Le début de (ii) n'est autre que l'alternative de Fredholm. Pour prouver la formule donnant f, remarquons d'abord que si $p = \infty$, la série du second membre converge dans E car la suite $(\lambda_n)_{1 \leqslant n < p}$ tend vers 0 ; ainsi pour $n \geqslant n_0$, on a :

$$\sum_{n \geqslant n_0} \left| \frac{\lambda_n}{1 - \mu\lambda_n} \right|^2 |(g|e_n)|^2 \leqslant \sum_{n \geqslant n_0} |(g|e_n)|^2 \leqslant \|g\|^2.$$

Posons

$$f = g + \mu \sum_{1 \leqslant n < p} \frac{\lambda_n}{1 - \mu\lambda_n} (g|e_n) e_n.$$

Alors

$$Tf = Tg + \mu \sum_{1 \leqslant n < p} \frac{\lambda_n^2}{1 - \mu\lambda_n} (g|e_n) e_n$$

Mais d'après (i), on a :

$$Tg = \sum_{1 \leqslant n < p} \lambda_n (g|e_n) e_n,$$

donc

$$Tf = \sum_{1 \leqslant n < p} \left(\lambda_n + \frac{\mu \lambda_n^2}{1 - \mu \lambda_n} \right) (g|e_n) e_n,$$

et finalement

$$f - \mu \cdot T(f) = g + \sum_{1 \leqslant n < p} \left(\frac{\lambda_n \mu}{1 - \mu \lambda_n} - \mu \lambda_n - \frac{\mu^2 \lambda_n^2}{1 - \mu \lambda_n} \right) (g|e_n) e_n,$$

qui vaut bien g. □

Exercice V .22 *Réciproquement, soit E un espace de Hilbert ; soit $(e_n)_{n \geqslant 1}$ une suite orthonormée dans E et soit $(\lambda_n)_{n \geqslant 1}$ une suite de nombres réels qui tend vers 0. Pour tout $f \in E$ on pose :*

$$Tf = \sum_{n \geqslant 1} \lambda_n (f|e_n) e_n.$$

Montrer que cette série converge dans E, et que l'opérateur T dans E ainsi défini est auto-adjoint compact.

Une classe d'opérateurs auto-adjoints compacts est donnée par les opérateurs à noyau symétrique.

Exemple V .23 *Soient $a < b$ des nombres réels. Soit $E = L^2([a;b], \mathbf{C})$ l'espace de Hilbert des (classes de) fonctions de carré sommable sur $[a;b]$ avec le produit scalaire*

$$(f_1|f_2) = \int_a^b f_1(x) \overline{f_2(x)} \, \mathrm{d}x$$

et la norme

$$\|f\|_2 = \sqrt{(f|f)}.$$

On sait que $\mathcal{C}([a;b], \mathbf{C})$ est dense dans E et que, si $f \in \mathcal{C}([a;b], \mathbf{C})$, on a :

$$\|f\|_2 \leqslant \sqrt{b-a} \cdot \|f\|_\infty.$$

Soit $K = K(x,y)$ un noyau continu sur $[a;b] \times [a;b]$. On pose

$$Tf(x) = \int_a^b K(x,y) f(y) \, \mathrm{d}y$$

pour toute $f \in L^2([a;b], \mathbf{C})$*. Il est clair que* $Tf \in L^2([a;b], \mathbf{C})$*, et même* $Tf \in \mathcal{C}([a;b], \mathbf{C})$ *par continuité sous le signe somme. Nous savons déjà que* $T : \mathcal{C}([a;b], \mathbf{C}) \to \mathcal{C}([a;b], \mathbf{C})$ *est un opérateur compact pour* $\mathcal{C}([a;b], \mathbf{C})$ *muni de la norme* $\|\cdot\|_\infty$ *de la convergence uniforme.*

Voyons que $T : L^2([a;b], \mathbf{C}) \to L^2([a;b], \mathbf{C})$ *est un opérateur compact pour la norme* $\|\cdot\|_2$*. L'ensemble* $\mathcal{H} = \{Tf : f \in L^2([a;b], \mathbf{C}), \|f\|_2 \leqslant 1\}$ *est contenu dans* $\mathcal{C}([a;b], \mathbf{C})$ *; de plus,* $\mathcal{H}$ *est :*

— *équiborné, car*

$$|Tf(x)| = \left|\int_a^b K(x,y)f(y)\,\mathrm{d}y\right| \leqslant \sup_x \left(\int_a^b |K(x,y)^2|\,\mathrm{d}y\right)^{\frac{1}{2}} = M,$$

par Cauchy-Schwarz.

— *équicontinu, car par uniforme continuité de* K*, on a :*

$$|Tf(x) - Tf(x')| \leqslant \int_a^b |K(x,y) - K(x',y)| \cdot |f(y)|\,\mathrm{d}y \leqslant \varepsilon \int_a^b |f(y)|\,\mathrm{d}y \leqslant \varepsilon\sqrt{b-a}$$

dès que $|x - x'| \leqslant \eta_\varepsilon$*. D'après le Théorème d'Ascoli, si* $(g_n)_{n\geqslant 1}$ *est une suite dans* $\mathcal{H}$*, on peut en extraire une suite de Cauchy pour la norme* $\|\cdot\|_\infty$*, donc a fortiori de Cauchy pour la norme* $\|\cdot\|_2$ *car on a vu que* $\|\cdot\|_2 \leqslant \sqrt{b-a} \cdot \|\cdot\|_\infty$*.*

Remarque V .24 *Les fonctions propres de* T *dans* $L^2([a;b], \mathbf{C})$*, pour une valeur propre* $\neq 0$*, sont déjà dans* $\mathcal{C}([a;b], \mathbf{C})$*, car on peut écrire* $f = \frac{1}{\lambda}Tf$ *et* $Tf \in \mathcal{C}([a;b], \mathbf{C})$*. Ainsi, il n'apparaît pas de nouvelle valeur propre quand on étend* T *de* $\mathcal{C}([a;b], \mathbf{C})$ *à* $L^2([a;b], \mathbf{C})$*.*

Revenons maintenant aux opérateurs auto-adjoints.

Proposition V .25 *Soit* $K = K(x,y)$ *un noyau continu sur le carré* $[a;b] \times [a;b]$*, et soit* T *l'opérateur compact dans l'espace sde Hilbert* $L^2([a;b], \mathbf{C})$ *donné par la formule :*

$$Tf(x) = \int_a^b K(x,y)f(y)\,\mathrm{d}y$$

pour toute $f \in L^2([a;b], \mathbf{C})$*. Alors, pour que* T *soit auto-adjoint dans* $L^2([a;b], \mathbf{C})$*, il faut et suffit que* $K(x,y) = \overline{K(y,x)}$ *pour tous* $x, y \in [a;b]$*.*

Dans ce cas, on dit que le noyau K est *symétrique*.

Preuve. En effet, pour f et g dans $L^2([a;b],\mathbf{C})$, par Fubini on a :

$$\begin{aligned}(Tf|g)-(f|Tg) &= \textstyle\int_a^b\int_a^b K(x,y)f(y)\overline{g(x)}\,\mathrm{d}x\mathrm{d}y - \int_a^b\int_a^b f(y)\overline{K(y,x)g(x)}\,\mathrm{d}x\mathrm{d}y \\ &= \textstyle\int_{[a;b]\times[a;b]}\Big(K(x,y)-\overline{K(y,x)}\Big)\,f(y)\overline{g(x)}\,\mathrm{d}x\mathrm{d}y\end{aligned}$$

Mais il faut se rappeler qu'une des conséquences du théorèmes de Stone-Weierstrass est que les fonctions de la forme $(x,y)\mapsto f(x)\overline{g(y)}$, avec f et g continues sur $[a;b]$, sont uniformément denses dans $\mathcal{C}([a;b]\times[a;b],\mathbf{C})$, donc par Cauchy-Schwarz denses aussi pour $\|\cdot\|_2$ dans $\mathcal{C}([a;b]\times[a;b],\mathbf{C})$, lequel est dense pour la norme L^2 dans $L^2([a;b]\times[a;b],\mathbf{C})$. Ceci implique l'équivalence cherchée. □

Exercice V .26 *Sur les noyaux de Hilbert-Schmidt. Soit $K=K(x,y)$ une fonction mesurable de $\mathbf{R}\times\mathbf{R}$ telle que :*

$$\|K\|_2=\left(\iint_{\mathbf{R}^2}|K(x,y)|^2\,\mathrm{d}x\mathrm{d}y\right)^{\frac12}<+\infty.$$

Montrer que la formule :

$$Tf(x)=\int_{\mathbf{R}}K(x,y)f(y)\,\mathrm{d}y$$

définit dans l'espace de Hilbert $L^2(\mathbf{R},\mathrm{d}x)$ un opérateur borné de norme $|||T|||\leqslant\|K\|_2$, que cet opérateur est compact, et qu'il est auto-adjoint si et seulement si $K(x,y)=\overline{K(y,x)}$ presque partout.

Indication : d'après la théorie de la mesure, il existe une suite $(K_n(x,y))_{n\geqslant1}$ de fonctions continues à support compact dans $\mathbf{R}^2$, telle que $\|K-K_n\|_{L^2(\mathbf{R}^2,\,\mathrm{d}x)}$ tende vers 0 quand $n\to\infty$. Chaque fonction K_n est nulle en dehors d'un carré $[a_n;b_n]\times[a_n;b_n]$. Pour $f\in L^2(\mathbf{R},\mathrm{d}x)$ poser

$$T_nf(x)=\int_{a_n}^{b_n}K_n(x,y)f(y\,\mathrm{d}y,$$

et montrer que chaque T_n est un opérateur compact.

Exercice V .27 *Déterminer les valeurs propres de l'opérateur de Fredholm T dans $\mathcal{C}([0;1],\mathbf{C})$ défini par :*

$$Tf(x)=\int_0^1 e^{|x-y|}f(y)\,\mathrm{d}y.$$

Montrer que $\sup\limits_{\|f\|_2\leqslant1}\|Tf\|_2\leqslant1,426.$

Exercice V .28 *On reprend l'exercice V .26 en supposant en outre que le noyau K est symétrique. Soit $(\lambda_n)_{n\geqslant 1}$ la suite finie ou infinie des valeurs propres non nulles de T, comptées autant de fois que leurs multiplicités. Soit $(\varphi_n)_{n\geqslant 1}$ une famille orthonormée de fonctions propres de T, avec $T\varphi_n = \lambda_n\varphi_n$ pour tout $n \geqslant 1$. Prouver la formule :*

$$K(x,y) = \sum_{n\geqslant 1} \lambda_n \varphi_n(x)\overline{\varphi_n(y)},$$

où la série converge dans l'espace de Hilbert $L^2(\mathbf{R}^2, \mathrm{d}x\mathrm{d}y)$.

V.33 Noyaux de Fredholm itérés et noyau résolvant

On peut dans certains cas résoudre l'équation de Fredholm par approximations successives. Soit $K = K(x,y)$ un noyau continu sur le carré $[a;b] \times [a;b]$, et soit T l'opérateur de Fredholm de noyau K sur $\mathcal{C}([a;b], \mathbf{C})$ donné par la formule :

$$Tf(x) = \int_a^b K(x,y)f(y)\,\mathrm{d}y$$

pour toute $f \in \mathcal{C}([a;b], \mathbf{C})$. Par Fubini, son carré T^2 est encore un opérateur de Fredholm, de noyau : $K_2(x,y) = \int_a^b K(x,z)K(z,y)\,\mathrm{d}z$. Et plus généralement, les puissances T^n sont des opérateurs de Fredholm, donnés par les noyaux itérés K_n définis par $K_1 = K$ et

$$K_n(x,y) = \int_a^b K(x,z)K_{n-1}(z,y)\,\mathrm{d}z.$$

Exercice V .29 *Prouver les formules :*

$$K_{m+n}(x,y) = \int_a^b K_m(x,z)K_n(z,y)\,\mathrm{d}z \quad \text{et} \quad K_n(x,y) = \int_a^b\int_a^b K(x,z)K_{n-2}(z,t)K(t,y)\,\mathrm{d}z\mathrm{d}t.$$

Théorème V .30 *Faisons l'hypothèse que :*

$$\int_a^b\int_a^b |K(x,y)|^2\,\mathrm{d}x\mathrm{d}y < 1.$$

Alors la série, dite de Neumann, associée au noyau K :

$$R(x,y) = \sum_{n=1}^{\infty} K_n(x,y)$$

converge uniformément en (x,y) sur le carré $[a;b] \times [a;b]$ et, pour toute $g \in \mathcal{C}([a;b], \mathbf{C})$, la formule :

$$f(x) = g(x) + \int_a^b R(x,y)g(y)\,\mathrm{d}y$$

fournit l'unique solution $f \in \mathcal{C}([a;b], \mathbf{C})$ *de l'équation de Fredholm*

$$f(x) = \int_a^b K(x,y) f(y)\,\mathrm{d}y + g(x).$$

Preuve. Si $N(x,y)$ est un noyau, on notera pour abréger :

$$\|N\|_2 = \left(\int_a^b \int_a^b |N(x,y)|^2\,\mathrm{d}x\mathrm{d}y \right)^{\frac{1}{2}}.$$

Ainsi, par hypothèse $\|K\|_2 < 1$.

Montrons d'abord que l'équation de Fredholm ci-dessus a une solution et une seule. Grâce à l'alternative de Fredholm (cf Proposition V .13), il suffit de s'assurer que $f(x) = \int_a^b K(x,y) f(y)\,\mathrm{d}y$ implique que f est identiquement nulle. Or, d'après Cauchy-Schwarz, on a sous cette hypothèse :

$$|f(x)|^2 \leqslant \int_a^b |K(x,y)|^2\,\mathrm{d}y \cdot \int_a^b |f(y)|^2\,\mathrm{d}y,$$

et en intégrant en x, il vient :

$$\int_a^b |f(x)|^2\,\mathrm{d}x \cdot \left(1 - (\|K\|_2)^2\right) \leqslant 0,$$

ce qui exige que $\int_a^b |f(x)|^2\,\mathrm{d}x = 0$, et finalement que f soit la fonction nulle.

En appliquant l'inégalité de Cauchy-Schwarz à la formule de récurrence définissant les K_n, il vient :

$$|K_n(x,y)|^2 \leqslant \int_a^b |K(x,z)|^2\,\mathrm{d}z \cdot \int_a^b |K_{n-1}(z,y)|^2\,\mathrm{d}z,$$

puis en intégrant les deux membres en x et en y, on obtient :

$$(\|K_n\|_2)^2 \leqslant (\|K\|_2)^2 \cdot (\|K_{n-1}\|_2)^2,$$

d'où l'inégalité

$$\|K_n\|_2 \leqslant (\|K\|_2)^n.$$

De même, en appliquant l'inégalité de Cauchy-Schwarz à la seconde formule de l'Exercice V .29, on a :

$$|K_n(x,y)|^2 \leqslant \left(\int_a^b \int_a^b |K_{n-2}(z,t)|^2\,\mathrm{d}z\mathrm{d}t \right) \cdot \left(\int_a^b \int_a^b |K(x,z) \cdot K(t,y)|^2\,\mathrm{d}z\mathrm{d}t \right).$$

La deuxième intégrale à droite est continue en (x,y), donc bornée. En utilisant $\|K_n\|_2 \leqslant (\|K\|_2)^n$, on voit donc que pour tous x et y dans $[a;b]$, on a :

$$|K_n(x,y)| \leqslant C(\|K\|_2)^n.$$

Par suite la série de Neumann associée à $K(x,y)$:

$$R(x,y) = \sum_{n=1}^{\infty} K_n(x,y),$$

vu l'hypothèse du théorème, converge normalement, donc absolument, sur $[a;b] \times [a;b]$. Par la définition de la série de Neumann et intégration terme à terme, on voit qu'on a aussi :

$$R(x,y) = K(x,y) + \int_a^b K(x,z)R(z,y)\,\mathrm{d}z.$$

Montrons enfin que la solution proposée dans l'énoncé résout bien le problème de Fredholm considéré. On a :

$$\begin{aligned}
f(x) - \int_a^b K(x,y)f(y)\,\mathrm{d}y &= g(x) + \int_a^b R(x,y)g(y)\mathrm{d}y - \int_a^b K(x,y)\left(g(y) + \int_a^b R(y,z)g(z)\mathrm{d}z\right)\mathrm{d}y \\
&= g(x) + \int_a^b \left(R(x,y) - K(x,y) - \int_a^b K(x,z)R(z,y)\,\mathrm{d}z\right) g(y)\,\mathrm{d}y \\
&= g(x),
\end{aligned}$$

par la relation entre R et K obtenue juste auparavant. □

Exercice V .31 *Pour $0 \leqslant x \leqslant 2\pi$ et $0 \leqslant y \leqslant 2\pi$, on pose :*

$$K(x,y) = \sum_{p=1}^{\infty} \frac{\sin(px)\sin\left((p+1)y\right)}{p^2}.$$

1. *Montrer que, pour tout $n \geqslant 1$, on a :*

$$K_n(x,y) = \sum_{p=1}^{\infty} \pi^{n-1} \frac{\sin(px)\sin\left((p+1)y\right)}{p^2(p+1)^2\ldots(p+n-1)^2}.$$

2. *En déduire que la série de Neumann associée $\frac{1}{\lambda}K$:*

$$R_\lambda(x,y) = \sum_{n=1}^{\infty} \frac{1}{\lambda^n} K_n(x,y)$$

converge uniformément quel que soit $\lambda \neq 0$, et que l'opérateur de Fredholm de noyau K n'a aucune valeur propre non nulle.

Appliquons une méthode analogue pour intégrer l'équation de Volterra. Soit S l'opérateur de Volterra de noyau K, défini par :

$$Sf(x) = \int_a^x K(x,y)f(y)\,\mathrm{d}y.$$

Les itérés S^n sont des opérateurs de Volterra dont les noyaux sont cette fois les $K^{(n)}(x,y)$ où $K^{(1)} = K$ et pour $n \geqslant 2$:

$$K^{(n)}(x,y) = \int_y^x K(x,z)K^{(n-1)}(z,y)\,\mathrm{d}z.$$

En posant

$$M = \sup_{a \leqslant x,y \leqslant b} |K(x,y)|,$$

on vérifie par récurrence que

$$|K^{(n)}(x,y)| \leqslant M^n \frac{|x-y|^{n-1}}{(n-1)!}.$$

En effet, c'est vrai pour $n = 1$, et d'après la relation de récurrence qui précède, on a :

$$\begin{array}{rcl} |K^{(n)}(x,y)| & \leqslant & |x-y| M \int_y^x |K^{(n-1)}(z,y)|\,\mathrm{d}z \\ & \leqslant & |x-y| M^n \int_y^x \frac{|z-y|^{n-2}}{(n-2)!}\,\mathrm{d}z \\ & \leqslant & M^n \frac{|x-y|^{n-1}}{(n-1)!}. \end{array}$$

Par conséquent,

$$|K^{(n)}(x,y)| \leqslant \frac{M^n (b-a)^{n-1}}{(n-1)!},$$

et la série

$$R(x,y) = \sum_{n=1}^{\infty} K^{(n)}(x,y)$$

converge normalement, donc uniformément, en (x,y). En terminant comme dans la preuve précédente, on a prouvé *cette fois sans hypothèse sur la norme* $\|\cdot\|_2$ *de* K le résultat suivant.

Théorème V .32 *L'unique solution de l'équation de Volterra*

$$f(x) = \int_a^x K(x,y) f(y)\,\mathrm{d}y + g(x)$$

est donnée par la formule :

$$f(x) = g(x) + \int_a^x R(x,y) g(y)\,\mathrm{d}y,$$

où le noyau résolvant R *est donné par la somme de série uniformément convergente :*

$$R(x,y) = \sum_{n=1}^{\infty} K^{(n)}(x,y)$$

dans laquelle $K^{(1)} = K$ *et*

$$K^{(n)}(x,y) = \int_y^x K(x,z)K^{(n-1)}(z,y)\,\mathrm{d}z$$

pour $n \geqslant 2$. □

On notera la réciprocité entre les noyaux K et R : le noyau résolvant $R(x,y)$ du noyau de Volterra n'est autre que $-K(x,y)$.

Exercice V .33 *Appliquer le théorème qui précède pour résoudre l'équation*

$$f(x) = \int_0^x (y-x)f(y)\,\mathrm{d}y + x,$$

puis retrouver la solution par une méthode élémentaire directe.

V.34 Équation intégrale d'Abel

Soit $E = \mathcal{C}([0;1], \mathbf{C})$ l'espace de Banach des fonctions f à valeurs complexes définies et continues sur $[0;1]$, muni de la norme $\|\cdot\|_\infty$ de la convergence uniforme :

$$\|f\|_\infty = \sup_{0 \leqslant x \leqslant 1} |f(x)|.$$

Soit α un paramètre réel donné tel que $0 < \alpha < 1$. Pour toute fonction $f \in E$ et tout $x \in [0;1]$, on pose :

$$Tf(x) = \int_0^x \frac{f(y)}{(x-y)^\alpha}\,\mathrm{d}y = \int_0^x \frac{f(x-y)}{y^\alpha}\,\mathrm{d}y.$$

L'opérateur T s'appelle l'*opérateur intégral d'Abel* de paramètre α. L'intégrale ci-dessus a un sens ; elle est même absolument convergente car :

$$\int_0^x \frac{|f(y)|}{(x-y)^\alpha}\,\mathrm{d}y = \int_0^x \frac{|f(x-y)|}{y^\alpha}\,\mathrm{d}y \leqslant \frac{x^{1-\alpha}}{1-\alpha}\|f\|_\infty < +\infty.$$

De plus, la fonction $x \mapsto Tf(x)$ est continue sur $[0;1]$ car pour $\varepsilon > 0$ donné, par continuité uniforme de f et par continuité de $x \mapsto x^{1-\alpha}$, on a :

$$|f(x'-y) - f(x-y)| \leqslant \varepsilon \quad \text{et} \quad |(x')^{1-\alpha} - x^{1-\alpha}| \leqslant \varepsilon$$

dès que $|x - x'| \leqslant \eta_\varepsilon$, et donc, si par exemple $x' \geqslant x$, on a :

$$\begin{aligned} |Tf(x') - Tf(x)| &\leqslant \textstyle\int_0^x \frac{|f(x'-y)-f(x-y)|}{y^\alpha}\,\mathrm{d}y + \int_x^{x'} \frac{|f(x'-y)|}{y^\alpha}\,\mathrm{d}y \\ &\leqslant \textstyle\varepsilon \int_0^1 \frac{\mathrm{d}y}{y^\alpha} + \frac{\|f\|_\infty}{1-\alpha}((x')^{1-\alpha} - x^{1-\alpha}) \\ &\leqslant \textstyle\frac{1+\|f\|_\infty}{1-\alpha}\varepsilon \end{aligned}$$

Ainsi, T envoie bien E dans E. De plus, T est un opérateur borné de norme $\leqslant \frac{1}{1-\alpha}$ d'après l'inégalité utilisée pour vérifier la convergence absolue de l'intégrale définissant $Tf(x)$. En fait, on a exactement :

$$|||T||| = \frac{1}{1-\alpha}$$

car cette valeur est atteinte pour la fonction constante égale à 1. L'opérateur T est presque du type Volterra au sens précédent, mais le noyau est ici singulier. Ceci nous oblige à faire une démonstration pour prouver que T est un opérateur compact.

Pour cela considérons, pour $0 < a < 1$, l'opérateur T_a dans E défini par :

$$T_a f(x) = \int_a^x \frac{f(x-y)}{y^\alpha}\,\mathrm{d}y$$

si $x \geqslant a$, et $T_a f(x) = 0$ sinon. On a :

$$(T - T_a)f(x) = \int_0^a \frac{f(x-y)}{y^\alpha}\,\mathrm{d}y \quad \text{si } x \geqslant a$$

et $(T - T_a)f(x) = 0$ si $0 \leqslant x \leqslant a$. Ceci implique que

$$|(T - T_a)f(x)| \leqslant \|f\|_\infty \frac{a^{1-\alpha}}{1-\alpha},$$

d'où

$$\lim_{a>0, a\to 0} |||T - T_a||| = 0.$$

Par le Théorème V .4, il suffit donc de prouver que les opérateurs T_a sont compacts. Pour cela, on regarde les images de la boule unité $\{\|\cdot\|_\infty \leqslant 1\}$ par ces opérateurs, précisément :

$$\mathcal{H}_a = \{T_a f : f \in E \text{ avec } \|f\|_\infty \leqslant 1\}.$$

Les valeurs de toutes les fonctions de tous les $\mathcal{H}_a$ sont uniformément bornées par ce qui précède. Pour appliquer Ascoli et conclure, il suffit donc de vérifier l'équicontinuité de chaque famille $\mathcal{H}_a$. Pour a donné, l'équicontinuité en $x \in [0; a[$ est évidente car toutes les fonctions de $\mathcal{H}_a$ sont identiquement nulles au voisinage de x ; idem pour l'équicontinuité à gauche en a. Pour vérifier l'équicontinuité à droite en a, on calcule :

$$|T_a f(x) - T_a f(a)| = |T_a f(x)| = |\int_a^x \frac{f(x-y)}{y^\alpha}\,\mathrm{d}y| \leqslant \frac{|x-a|}{a^\alpha}.$$

Si maintenant $a < x \leqslant x' \leqslant 1$, on a puisque $\|f\|_\infty \leqslant 1$:

$$\begin{array}{rcl}
|T_af(x) - T_af(a)| & = & \left|\int_a^{x'} \frac{f(x'-y)}{y^\alpha}\,\mathrm{d}y - \int_a^x \frac{f(x-y)}{y^\alpha}\,\mathrm{d}y\right| \\
& = & \left|\int_0^{x'-a} \frac{f(y)}{(x'-y)^\alpha}\,\mathrm{d}y - \int_0^{x-a} \frac{f(y)}{(x-y)^\alpha}\,\mathrm{d}y\right| \\
& \leqslant & \left|\int_{x-a}^{x'-a} \frac{f(y)}{(x'-y)^\alpha}\,\mathrm{d}y\right| + \left|\int_0^{x-a} \left(\frac{1}{(x'-y)^\alpha} - \frac{1}{(x-y)^\alpha}\right) f(y)\,\mathrm{d}y\right| \\
& \leqslant & \int_{x-a}^{x'-a} \frac{f(y)}{(x'-y)^\alpha}\,\mathrm{d}y + \int_0^{x-a} \left(\frac{1}{(x'-y)^\alpha} - \frac{1}{(x-y)^\alpha}\right) f(y)\,\mathrm{d}y \\
& = & \frac{1}{1-\alpha}\big((x'-x+a)^{1-\alpha} - a^{1-\alpha} + (x')^{1-\alpha} - a^{1-\alpha} - (x)^{1-\alpha} + (x'-x+a) \\
& \leqslant & \big(2(x'-x+a)^{1-\alpha} - 2a^{1-\alpha} + (x')^{1-\alpha} - (x)^{1-\alpha}\big) \\
& \leqslant & C(x'-x),
\end{array}$$

où la constante C dépend de a et de α mais pas de x. Ceci prouve l'équicontinuité de $\mathcal{H}_a$.

Lemme V .34 *Pour tout entier $n \geqslant 1$, toute $f \in E$ et tout $x \in [0;1]$, on a :*

$$|T^n f(x)| \leqslant \frac{x^{n\beta}\,(\Gamma(\beta))^n}{\Gamma(1+n\beta)} \cdot \|f\|_\infty,$$

où l'on a posé $\beta = 1 - \alpha$.

Preuve. On fait une récurrence sur $n \geqslant 1$. Pour $n = 1$, cela résulte immédiatement de

$$|Tf(x)| \leqslant \int_0^x \frac{|f(y)|}{(x-y)^\alpha}\,\mathrm{d}y = \int_0^x \frac{|f(x-y)|}{y^\alpha}\,\mathrm{d}y \leqslant \frac{x^{1-\alpha}}{1-\alpha}\|f\|_\infty.$$

Pour l'hérédité, on calcule :

$$|T^{n+1}f(x)| = |(T\circ T^n)f(x)| \leqslant \int_0^x \frac{|T^n f(y)|}{(x-y)^\alpha}\,\mathrm{d}y \leqslant \frac{(\Gamma(\beta))^n}{\Gamma(1+n\beta)}\cdot\|f\|_\infty \int_0^x \frac{y^{n\beta}}{(x-y)^\alpha}\,\mathrm{d}y,$$

donc, par le changement de variable $y = sx$, on obtient :

$$\begin{array}{rcl}
|T^{n+1}f(x)| & \leqslant & \frac{(\Gamma(\beta))^n}{\Gamma(1+n\beta)} \cdot \|f\|_\infty \frac{x^{n\beta}x}{x^\alpha} \int_0^1 s^{n\beta}(1-s)^{-\alpha}\,\mathrm{d}s \\
& = & x^{(n+1)\beta} \frac{(\Gamma(\beta))^n}{\Gamma(1+n\beta)} B(n\beta+1,\beta)\|f\|_\infty \\
& = & x^{(n+1)\beta} \frac{(\Gamma(\beta))^n}{\Gamma(1+n\beta)} \frac{\Gamma(1+n\beta)\Gamma(\beta)}{\Gamma(1+(n+1)\beta)}\|f\|_\infty,
\end{array}$$

ce qui est l'inégalité cherchée pour prouver la récurrence. □

Il résulte du Lemme que :

$$|||T^n||| \leqslant \frac{(\Gamma(\beta))^n}{\Gamma(1+n\beta)}.$$

D'après la formule de Stirling, $\Gamma(x)$ tend vers $+\infty$ quand $x \to +\infty$ plus vite que toute k^x où $k > 0$ est donné. Donc pour tout $\mu \in \mathbf{C}$ donné, la série d'opérateurs $\sum_{n\geqslant 1}\mu^n T^n$ converge absolument dans l'espace de Banach $\mathcal{L}(E)$. Il en résulte immédiatement que *pour tout* $\mu \in \mathbf{C}$*, l'opérateur* $\mathrm{id}_E - \mu T$ *est inversible dans* $\mathcal{L}(E)$*, et*

$$(\mathrm{id}_E - \mu T)^{-1} = \sum_{n=1}^{\infty} \mu^n T^n.$$

Ainsi, comme les opérateurs de Volterra classiques, l'opérateur d'Abel T n'a d'après l'alternative de Fredholm, aucune valeur propre $\neq 0$.

Nous allons enfin prouver la *formule d'inversion d'Abel.*

Théorème V .35 *Si* $f \in E$ *et si*

$$g(x) = \int_0^x \frac{f(y)}{(x-y)^\alpha}\,\mathrm{d}y,$$

alors

$$f(y) = \frac{\sin(\pi\alpha)}{\pi}\frac{\mathrm{d}}{\mathrm{d}y}\int_0^y \frac{g(x)}{(y-x)^{1-\alpha}}\,\mathrm{d}x.$$

Preuve. L'idée est d'utiliser le Théorème d'approximation de Weierstrass afin de réduire la preuve de la formule à une vérification sur les monômes.

Prouvons donc la formule pour $f(x) = x^k$ où k est un entier $\geqslant 0$. Dans ce cas :

$$g(x) = \int_0^x (x-y)^{-\alpha} y^k\,\mathrm{d}y = x^{1-\alpha+k}\int_0^1 (1-s)^{-\alpha} s^k\,\mathrm{d}s = \frac{\Gamma(1-\alpha)\Gamma(k+1)}{\Gamma(2-\alpha+k)} x^{1-\alpha+k}$$

et

$$\begin{aligned}
\textstyle\int_0^y \frac{g(x)}{(y-x)^{1-\alpha}}\,\mathrm{d}x &= \textstyle\frac{\Gamma(1-\alpha)\Gamma(k+1)}{\Gamma(2-\alpha+k)}\int_0^y (y-x)^{\alpha-1} x^{1-\alpha+k}\,\mathrm{d}x \\
&= \textstyle\frac{\Gamma(1-\alpha)\Gamma(k+1)}{\Gamma(2-\alpha+k)} y^{\alpha-1+1-\alpha+k+1}\int_0^1 (1-s)^{\alpha-1} s^{1-\alpha+k}\,\mathrm{d}s \\
&= \textstyle\frac{\Gamma(1-\alpha)\Gamma(k+1)}{\Gamma(2-\alpha+k)}\frac{\Gamma(\alpha)\Gamma(2-\alpha+k)}{\Gamma(k+2)} y^{k+1} \\
&= \textstyle\frac{\pi}{\sin(\pi\alpha)}\frac{y^{k+1}}{k+1},
\end{aligned}$$

donc

$$\frac{\sin(\pi\alpha)}{\pi}\frac{\mathrm{d}}{\mathrm{d}y}\int_0^y \frac{g(x)}{(y-x)^{1-\alpha}}\,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{y^{k+1}}{k+1}\right) = y^k = f(y).$$

Ceci prouve la formule d'inversion dAbel pour les monômes et donc, par combinaison linéaire, pour les fonctions polynomiales.

Faisons maintenant usage du Théorème d'approximation de Weierstrass pour conclure en toute généralité. Soit $f \in E$ une fonction continue quelconque. On dispose d'une suite $(p_n)_{n\geqslant 1}$ de fonctions polynomiales telle que :

$$\lim_{n\to\infty} \|f - p_n\|_\infty = 0.$$

Soit F (resp. P_n) la primitive de f (resp. p_n) qui s'annule en 0. On a aussi

$$\lim_{n\to\infty} \|F - P_n\|_\infty = 0.$$

Soit $g = Tf$ et $g_n = Tp_n$. Puisque l'opérateur T est borné, on a :

$$\lim_{n\to\infty} \|g - g_n\|_\infty = 0.$$

La transformation d'Abel de paramètre $1 - \alpha$ étant elle aussi continue, on voit que :

$$\int_0^y \frac{g_n(x)}{(y-x)^{1-\alpha}}\,\mathrm{d}x \text{ converge vers } \int_0^y \frac{g(x)}{(y-x)^{1-\alpha}}\,\mathrm{d}x$$

uniformément en $y \in [0;1]$. Mais d'après la formule déjà prouvée pour les fonctions polynomiales, on a :

$$\int_0^y \frac{g_n(x)}{(y-x)^{1-\alpha}}\,\mathrm{d}x = P_n(y)\frac{\pi}{\sin(\pi\alpha)}$$

pour tout $n \geqslant 1$. Les fonctions données par le second membre forment une suite qui tend vers $F(y)\frac{\pi}{\sin(\pi\alpha)}$ uniformément en y, ce qui prouve la formule d'inversion d'Abel en toute généralité. □

Noter que la formule d'inversion d'Abel implique que $\mathrm{Ker}(T) = \{0\}$, et donc que T n'a pas de valeur propre du tout.

V.35 Équations de Sturm-Liouville

On sait qu'une équation différentielle du deuxième ordre

$$y'' + p(x)y' + q(x)y = f(x),$$

où les fonctions p, q et f sont données, continues sur $[a;b]$, a une solution et une seule pour des conditions initiales du type Cauchy donnant les valeurs $y(x_0)$ et $y'(x_0)$, en un point fixé $x_0 \in [a;b]$, de la fonction inconnue y et de sa dérivée.

Mais la Physique conduit à chercher des solutions $y = y(x)$ satisfaisant à des conditions aux limites, par example :

$$y(a) = y(b) = 0,$$

ou plus généralement

$$Ay'(a) + By(a) = 0 \quad \text{et} \quad Cy'(b) + Dy(b) = 0,$$

où A, B, C et D sont des constantes données. Pensez par exemple au problème des cordes vibrantes.

Un exemple très simple, lié aux séries de Fourier, montre que la recherche de solutions avec conditions aux limites fournit des résultats bien différents de ceux mettant en jeu des conditions initiales. Si λ est une constante, le problème

$$y'' + \lambda y = 0 \quad \text{avec} \quad y(0) = y(\pi) = 0$$

n'a de solution non identiquement nulle que si $\lambda = n^2$ où n est un nombre entier > 0, auquel cas il y a alors une infinité de solutions.

Sur des exemples, nous allons voir que ce type de problème peut se ramener à la résolution d'équations de Fredholm. Nous avons déjà constaté les liens entre les deux théories dans des exercices ci-dessus. Ils apparaissent de façon élémentaire dans l'exercice qui suit, qui d'ailleurs est à la base des considérations qui vont suivre.

Exercice V .36 *Soit sur* $[a;b] \times [a;b]$ *le noyau symétrique*

$$N(x,y) = \frac{(x-b)(y-a)}{b-a} \text{ pour } y \leqslant x \quad \text{et} \quad N(x,y) = \frac{(x-a)(y-b)}{b-a} \text{ pour } x \leqslant y.$$

Soit $f \in \mathcal{C}([a;b], \mathbf{C})$ *donnée. Montrer que*

$$F(x) = \int_a^b N(x,y) f(y)\, \mathrm{d}y$$

est la fonction de classe C^2 *telle que* $F'' = f$ *et* $F(a) = F(b) = 0$.

Ainsi le problème de Sturm-Liouville élémentaire

$$y'' = f \quad \text{et} \quad y(a) = y(b) = 0$$

est résolu en appliquant à f l'opérateur de Fredholm de noyau N.

On va commencer par étudier un cas simple, celui d'une équation

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \lambda A(x) y + f(x)$$

pour les conditions aux limites

$$y(a) = y(b) = 0,$$

où A et f sont des fonctions continues données sur $[a;b]$, à valeurs complexes, et λ est un paramètre complexe.

En adoptant d'abord une démarche heuristique, cherchons à satisfaire l'équation différentielle, formellement et par un développement ordonné suivant les puissances de λ :

$$y(x) = y_0(x) + \lambda y_1(x) + \lambda^2 y_2(x) + \cdots + \lambda^n y_n(x) + \ldots$$

et où chaque y_n satisfait les conditions aux limites, i.e. $y_n(a) = y_n(b) = 0$.

Le premier terme y_0 s'obtient en résolvant l'équation différentielle avec $\lambda = 0$, c'est-à-dire :

$$y''(x) = f(x) \quad \text{et} \quad y(a) = y(b) = 0,$$

dont la solution, on vient de le voir dans l'Exercice précédent, est :

$$y_0(x) = \int_a^b N(x,s) f(s)\,\mathrm{d}s$$

en posant

$$N(x,y) = \frac{(x-b)(s-a)}{b-a} \text{ pour } s \leqslant x \quad \text{et} \quad N(x,y) = \frac{(x-a)(s-b)}{b-a} \text{ pour } x \leqslant s.$$

Ce noyau est continu et symétrique dans le carré $[a;b] \times [a;b]$. Les autre termes du développement se calculent (toujours formellement) de proche en proche ; on a :

$$\frac{\mathrm{d}^2 y_n}{\mathrm{d}x^2} = \lambda A(x) y_{n-1} \quad \text{et} \quad y_n(a) = y_n(b) = 0,$$

d'où en remplaçant dans ce qui précède f par Ay_{n-1} :

$$y_n(x) = \int_a^b N(x,s) A(s) y_{n-1}(s)\,\mathrm{d}s.$$

Comme on l'a vu, ce sont précisément les calculs qu'il faudrait effectuer pour résoudre par approximations successives l'équation de Fredholm :

$$(*) \quad y(x) = \lambda \int_a^b N(x,s) A(s) y(s)\,\mathrm{d}s + \int_a^b N(x,s) f(s)\,\mathrm{d}s,$$

ce qui rend extrêmement plausible l'énoncé suivant.

Proposition V .37 *Étant donnée l'équation :*

$$(*) \quad \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \lambda A(x) y + f(x),$$

où A et f sont des fonctions continues données sur $[a;b]$, à valeurs complexes, et λ est un paramètre complexe. Introduisons le noyau $N(x,s)$ défini par :

$$N(x,y) = \frac{(x-b)(s-a)}{b-a} \text{ pour } s \leqslant x \quad \text{et} \quad N(x,y) = \frac{(x-a)(s-b)}{b-a} \text{ pour } x \leqslant s,$$

puis le noyau non symétrique :

$$K(x,s) = N(x,s) A(s),$$

et posons :

$$g(x) = \int_a^b N(x,s)f(s)\,\mathrm{d}s.$$

Alors, pour une fonction $y = y(x)$ les deux assertions suivantes sont équivalentes :

(i) *la fonction y est solution de $(*)$ et $y(a) = y(b) = 0$;*

(ii) *la fonction y est solution de l'équation de Fredholm :*

$$y(x) = \lambda \int_a^b K(x,s)y(s)\,\mathrm{d}s + g(x).$$

Preuve. Supposons (i), autrement dit qu'on a : $y''(x) = \lambda A(x)y(x) + f(x)$ et $y(a) = y(b) = 0$. Il suffit d'appliquer l'Exercice d'introduction avec $f(x)$ remplacé par $\lambda A(x)y(x) + f(x)$ pour obtenir (ii).

Réciproquement, supposons (ii). Dans

$$y(x) = \lambda \int_a^b K(x,s)y(s)\,\mathrm{d}s + g(x),$$

remplaçons $K(x,s)$ et $g(x)$ par leurs expressions en fonction de $N(x,s)$, puis revenons à la définition de $N(x,s)$. On constate alors que y satisfaisant $(**)$ satisfait automatiquement (i). □

Corollaire V .38 *Soit λ un nombre complexe. Alors :*

(i) *ou bien l'équation*

$$(*) \qquad \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \lambda A(x)y + f(x),$$

a, pour tout fonction f, une solution $y = y(x)$ et une seule telle que $y(a) = y(b) = 0$;

(ii) *ou bien l'équation homogène associée*

$$(**) \qquad \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \lambda A(x)y$$

a des solution non identiquement nulles telles que $y(a) = y(b) = 0$.

Dans le second cas, nous dirons que λ est une *valeur singulière* pour l'équation $(**)$ ou l'équation $(*)$; cela signifie que $\frac{1}{\lambda}$ est valeur propre de l'opérateur intégral de noyau $N(x,s)A(s)$.

Supposons désormais que A n'est pas identiquement nul.

La théorie spectrale des opérateurs compacts (cf le Théorème V .9) montre qu'il y a au plus une infinité dénombrable de valeurs singulières de λ pour l'équation $(**)$, et au plus un nombre fini dans tout disque $\{|\lambda| \leqslant R\}$. Nous verrons bientôt que si $A(x)$ est une fonction réelle, il y a en effet une infinité de valeurs singulières.

Proposition V .39 *Chaque valeur singulière de* $(**)$ *est de multiplicité* 1 *seulement : pour* λ *singulière, il y a, à multiplication par un facteur constant près, une seule fonction propre, donc non identiquement nulle, solution de* $(**)$ *avec* $y(a) = y(b) = 0$.

Preuve. S'il existait deux telles solutions linéairement indépendantes, elles engendreraient l'espace vectoriel (de dimension 2) des solutions de $(**)$. Ainsi toute solution de ladite équation vérifierait automatiquement les conditions aux limites $y(a) = y(b) = 0$, ce qui serait en contradiction avec le Théorème de Cauchy suivant lequel $(**)$ avec les conditions $y(a) = 1$ et $y'(a) = 0$ admet une solution. □

Proposition V .40 *Supposons que* $A(x)$ *soit continue réelle et non identiquement nulle sur* $[a; b]$. *Alors les valeurs singulières de* $(**)$ *sont toutes réelles.*

Preuve. Le noyau $K(x, s) = N(x, s)A(s)$ de l'équation de Fredholm

$$y(x) = \lambda \int_a^b K(x, s)y(s)\,\mathrm{d}s + g(x)$$

n'est pas symétrique, donc le résultat annoncé ne peut se déduire de considérations sur les opérateurs auto-adjoints. On va travailler directement sur l'équation différentielle.

Soit $\lambda = \alpha + i\beta$ une valeur singulière complexe, et $y(x) = u(x) + iv(x)$ une fonction propre (donc non identiquement nulle). L'équation

$$(**) \quad \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \lambda A(x)y$$

donne, en passant aux parties réelle et imaginaire :

$$u'' = (\alpha u - \beta v)A \quad \text{et} \quad v'' = (\beta u + \alpha v)A,$$

d'où

$$vu'' - uv'' = \beta(u^2 + v^2)A \quad \text{et} \quad uu'' + vv'' = \alpha(u^2 + v^2)A,$$

et donc

$$0 = [vu' - uv']_a^b = -\beta \int_a^b (u^2 + v^2)A\,\mathrm{d}x,$$

et

$$0 = [uu' + vv']_a^b = \int_a^b \left((u')^2 + (v')^2\right)\,\mathrm{d}x + \alpha \int_a^b \int_a^b (u^2 + v^2)A\,\mathrm{d}x.$$

Ainsi, en supposant $\beta \neq 0$ on obtient

$$\int_a^b (u^2 + v^2)A\,\mathrm{d}x = \int_a^b \left((u')^2 + (v')^2\right)\,\mathrm{d}x = 0,$$

ce qui implique $u' = v' = 0$, donc que y' est constante, autrement dit nulle par les conditions aux limites satisfaites par y : c'est exclu, et on a donc nécessairement $\beta = 0$. □

Nous abondonnons provisoirement l'étude des valeurs singulières de $(**)$ pour exposer la *théorie de Sturm* pour les zéros des solutions (réelles) de l'équation différentielle :

$$y''(x) + p(x)y'(x) + q(x)y(x) = 0.$$

Soient $a < b$ deux nombres réels. Soient P et Q deux fonctions réelles continues sur $[a; b]$. On suppose de plus que $P(x)$ est > 0 et admet une dérivée première continue sur $[a; b]$. On va en fait étudier l'équation différentielle :

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(P(x)\frac{\mathrm{d}y}{\mathrm{d}x}\right) + Q(x)y = 0,$$

qui équivaut à la précédente si on pose :

$$p(x) = \frac{P'(x)}{P(x)} \quad \text{et} \quad q(x) = \frac{Q(x)}{P(x)},$$

ou, réciproquement :

$$P(x) = \exp\int_a^x p(t)\,\mathrm{d}t \quad \text{et} \quad Q(x) = q(x)P(x).$$

Proposition V .41 *Une solution $y(x)$ non identiquement nulle de*

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(P(x)\frac{\mathrm{d}y}{\mathrm{d}x}\right) + Q(x)y = 0$$

n'a qu'un nombre fini de zéros dans $[a; b]$.

Preuve. Sinon, il existerait dans $[a; b]$ un point d'accumulation, disons c, des zéros de $y(x)$ qui, d'après le théorème de Rolle serait aussi un point d'accumulation des zéros de $y'(x)$. Par continuité, on aurait $y(c) = y'(c) = 0$, forçant y à être identiquement nulle par unicité de la solution sous conditions de Cauchy. □

Proposition V .42 *Soient $c < d$ deux zéros consécutifs d'une solution y_1 non identiquement nulle de*

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(P(x)\frac{\mathrm{d}y}{\mathrm{d}x}\right) + Q(x)y = 0$$

et soit y_2 une solution de la même équation, non proportionnelle à y_1. Alors y_2 a un zéro et un seul dans $]c; d[$.

Preuve. Prouvons l'existence par l'absurde. Si, pour tout $x \in]c;d[$, on a $y_2(x) \neq 0$, on a aussi $y_2(c) \neq 0$ et $y_2(d) \neq 0$ (car sinon y_2 serait proportionnelle à y_1 comme ayant des conditions initiales proportionnelles). Ainsi $\frac{y_1}{y_2}$ est une fonction continue sur $[c;d]$ nulle en c et en d ; sa dérivé $\frac{y_1'y_2-y_1y_2'}{(y_2)^2}$ s'annule quelque part entre c et d, ce qui est impossible car le wronskien det $\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}$ n'est jamais nul pour deux solutions linéairement indépendantes.

Prouvons l'unicité par l'absurde. Supposons qu'il existe $\alpha < \beta$ dans $]c;d[$ tels que $y_2(\alpha) = y_2(\beta) = 0$. D'après la partie existence qui vient d'être démontrée, y_1 aurait un zéro entre α et β, ce qui est exclu par le fait que c et d sont consécutifs. □

On va maintenant comparer les solutions de deux équations différentielles du même type que ci-dessus, mais différentes.

Proposition V .43 *Soient les deux équations :*

$$(E) \quad \frac{\mathrm{d}}{\mathrm{d}x}\left(P(x)\frac{\mathrm{d}y}{\mathrm{d}x}\right) + Q(x)y = 0$$

et

$$(E') \quad \frac{\mathrm{d}}{\mathrm{d}x}\left(P(x)\frac{\mathrm{d}y}{\mathrm{d}x}\right) + Q_1(x)y = 0$$

où $P(x)$, $Q(x)$ et $Q_1(x)$ sont continues réelles, $P(x)$ est > 0 et où

$$Q(x) < Q_1(x)$$

pour tout $x \in [a;b]$. Soient $c < d$ deux zéros consécutifs d'une solution y non nulle de (E). Alors toute solution z de (E_1) s'annule au moins une fois dans $]c;d[$.

Autrement dit, à $P(x)$ donnée, on augmente le nombre des oscillations des solutions de (E) quand on augmente $Q(x)$.

Preuve. Multipliant (E) par z et (E_1) par y, puis soustrayant, on obtient :

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(P(zy' - yz')\right) = (Q_1 - Q)yz,$$

d'où, en intégrant entre c et d :

$$\int_c^d (Q_1 - Q)yz \,\mathrm{d}x = P(d)z(d)y'(d) - P(c)z(c)y'(c).$$

Par l'absurde, si $z(x)$ n'était jamais nulle dans $]c;d[$, alors $y(x)$ et $z(x)$ garderaient chacune un signe constant sur cet intervalle. Si, par exemple, les fonctions y et z restent > 0 sur $]c; d[$, alors le premier membre est > 0 alors que le second est $\leqslant 0$ car $y'(d) \leqslant 0$ et $y'(c) \geqslant 0$ (vu le signe de y sur $]c; d[$). Les trois autres possibilités de signes conduisent elles aussi à une contradiction à chaque fois. $\square$

Par exemple, soit l'équation :

$$y'' + Q(x)y = 0,$$

où $Q(x)$ est réelle continue, et vérifie l'inégalité

$$0 < \omega^2 < Q(x) < \Omega^2$$

pour tout $x \in [a; b]$, avec ω et Ω constantes. Comparons une solution de cette équation aux solutions de l'équation

$$y'' + \alpha^2 y = 0,$$

où successivement $\alpha = \omega$ et $\alpha = \Omega$. Les solutions sont de la forme $C \sin(\alpha(x - x_0))$ et ont pour zéros les nombres $x_0 + k\frac{\pi}{\alpha}$ où k est un entier. L'écart entre deux zéros consécutifs est constant et égal à $\frac{\pi}{\alpha}$. Il en résulte que, si $\delta = d - c$ est l'écart variable entre deux zéros consécutifs de la solution non identiquement nulle de $y'' + Q(x)y = 0$, on a les inégalités :

$$\frac{\pi}{\Omega} \leqslant \delta \leqslant \frac{\pi}{\omega}.$$

Car si δ était $< \frac{\pi}{\Omega}$, on trouverait un x_0 (par exemple $x_0 = c$) tel que pour aucun entier k, le nombre ne serait dans $]c; d[$. Et si δ était $> \frac{\pi}{\omega}$ il existerait x_0 tel que $[x_0; x_0 + \frac{\pi}{\Omega}]$ soit contenu dans $]c; d[$, donc ne contiendrait aucun zéro de y. Les inégalités $\frac{\pi}{\Omega} \leqslant \delta \leqslant \frac{\pi}{\omega}$ permettent d'assigner des bornes au nombre de zéros de $y(x)$ contenus dans un segment donné.

Exercice V .44 *Soit pour ν réel l'équation de :*

$$(E_\nu) \quad y'' + \frac{1}{x}y' + (1 - \frac{\nu^2}{x^2})y = 0 \quad (x > 0)$$

et sa solution la fonction de Bessel :

$$J_\nu(x) = \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty} \frac{(-\frac{x^2}{4})^k}{\Gamma(k+1)\Gamma(\nu + k + 1)}.$$

1. *En posant $Y = \sqrt{x}y$, transformer (E_ν) en une équation du type considéré dans la Proposition précédente avec $Q(x) = 1 - \frac{4\nu^2 - 1}{4x^2}$.*
2. *Soit $n(X)$ le nombre des zéros positifs de J_ν qui sont $\leqslant X$. Montrer que $\lim_{X\to\infty} \frac{n(X)}{X} = \frac{1}{\pi}$.*

Revenons à l'étude des valeurs singulières λ de l'équation

$$y'' = \lambda A(x) y$$

sous les conditions aux limites

$$y(a) = y(b) = 0$$

où l'on suppose $A(x)$ continue réelle et non identiquement nulle sur $[a;b]$. Si $x_0 \in [a;b]$, y_0 et m_0 sont donnés, notons $y(x,\lambda)$ la solution de $y'' = \lambda A(x)y$ qui satisfait aux conditions de Cauchy, indépendante de λ, suivantes :

$$y(x_0, \lambda) = y_0 \quad \text{et} \quad \frac{\mathrm{d}}{\mathrm{d}x} y(x_0, \lambda) = m_0.$$

Il résulte des théorèmes généraux sur les équations différentielles linéaires à coefficients dépendant d'un paramètre que, pour pour tout $x \in [a;b]$ fixé, la fonction

$$\lambda \mapsto y(x, \lambda)$$

est holomorphe dans $\mathbf{C}$: c'est une fonction entière de la variable complexe λ.

Remarque V .45 *On en déduit une nouvelle démonstration, indépendante de la théorie spectrale, du fait que le nombre de valeurs singulières dans tout disque fermé de rayon fini, est fini.*

Preuve. En effet, soient $y_1(x,\lambda)$ et $y_2(x,\lambda)$ deux solutions de $y'' = \lambda A(x)y$ correspondant à des conditions initiales linéairement indépendantes. La solution générale de cette équation s'écrit alors comme une combinaison linéaire :

$$y(x) = C_1 y_1(x, \lambda) + C_2 y_2(x, \lambda).$$

Alors λ est valeur singulière si et seulement si il existe C_1 et C_2 non toutes deux nulles et telles que

$$y(x) = C_1 y_1(a, \lambda) + C_2 y_2(a, \lambda) = 0$$

et

$$y(x) = C_1 y_1(b, \lambda) + C_2 y_2(b, \lambda) = 0,$$

donc si et seulement si la fonction entière

$$\lambda \mapsto y_1(a, \lambda) y_2(b, \lambda) - y_2(a, \lambda) y_1(b, \lambda)$$

est nulle au point λ. Or cette fonction entière n'est pas identiquement nulle car 0 n'est pas valeur singulière. Donc ses racines sont isolées dans $\mathbf{C}$. □

Lemme V .46 *Soit $y(x,\lambda)$ la solution de*

$$y'' = \lambda A(x) y$$

correspondant à des conditions initiales x_0, y_0 et m_0 indépendantes de λ. Alors il existe un signe $\epsilon = \pm$ tel que quand λ réel tend vers $\epsilon\infty$, le nombre de zéros de cette solution tend vers l'infini.

Preuve. Puisque A n'est pas identiquement nulle, on peut trouver $c < d$ dans $[a;b]$ tels que $|A(x)| > \omega^2 > 0$ pour tout $x \in [c;d]$. Prenons λ du signe de $A(x)$ sur $[c;d]$. On a : $\lambda A(x) > |\lambda|\omega^2$. D'après le Théorème de comparaison de Sturm, $y(x,\lambda)$ s'annule au moins $(d-c)\frac{\omega}{\pi}\sqrt{|\lambda|} - 1$ fois sur $[c;d]$. □

Proposition V .47 *Si $A(x)$ est réelle continue non identiquement nulle sur $[a;b]$, alors l'équation*

$$y'' = \lambda A(x) y$$

a une infinité de valeurs singulières (nécessairement réelles) pour le problème aux limites $y(a) = y(b) = 0$.

Preuve. Soit $y(x,\lambda)$ la solution de $y'' = \lambda A(x)y$ telle que $y(x,\lambda) = 0$ et $y'(x,\lambda) = 1$. Pour fixer les idées, supposons que c'est quand $\lambda \to +\infty$ que le nombre de zéros de $y(x,\lambda)$ augmente indéfiniment (suivant le lemme qui précède). Puisque $y(x,\lambda)$ dépend continûment de λ et puisque tous ses zéros sont simples (par unicité de Cauchy), le nombre de ces zéros ne peut augmenter que lorsqu'on passe par une valeur de λ telle que $y(b,\lambda) = 0$, c'est-à-dire une valeur singulière. Quand $\lambda \to +\infty$ en croissant, le lemme impose donc qu'on passe une infinité de fois par des valeurs singulières. □

Étudions donc pour terminer les cas particulier où $A(x)$ est < 0 pour tout $x \in [a;b]$. En posant alors :

$$Q(x) = -A(x)$$

on étudie le problème :

$$y'' = \lambda Q(x) y \quad \text{avec} \quad y(a) = y(b) = 0,$$

où $Q(x)$ est > 0 pour tout $x \in [a;b]$. À tout ce qui a été déjà écrit, on peut ajouter les compléments suivants.

Les valeurs singulières sont toutes réelles strictement positives.

En effet, puisque $\int_a^b y(y'' + \lambda Q y)\,\mathrm{d}x = 0$, on a :

$$0 = [yy']_a^b = \int_a^b (y')^2\,\mathrm{d}x - \lambda \int_a^b Q y^2\,\mathrm{d}x,$$

donc $\lambda \int_a^b Qy^2 \,\mathrm{d}x \geqslant 0$, puis $\lambda \geqslant 0$, donc $\lambda > 0$ puisque 0 n'est pas valeur singulière.

Rangeons les valeurs singulières par ordre croissant : $\lambda_1 < \lambda_2 < \cdots < \lambda_n \ldots$, *et soit* y_n *une fonction propre de l'équation ci-dessus associée à* λ_n*. Alors* y_n *a exactement* $(n-1)$ *racines dans* $]a;b[$.

On le voit en approfondissant un peu la méthode de démonstration de la Proposition V .47 ; et en vérifiant que, chaque nouvelle fois que $y(b,\lambda)$ s'annule, il s'introduit effectivement un nouveau zéro de $y(x,\lambda)$. Autrement dit, il suffit de vérifier que

$$\frac{\partial y}{\partial \lambda}(b,\lambda) \cdot \frac{\partial y}{\partial x}(b,\lambda) > 0$$

au moment où $y(b,\lambda) = 0$. Or, en dérivant par rapport à λ l'équation différentielle $y'' = \lambda Q(x) y$, on obtient :

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial y}{\partial \lambda}\right) + Qy + \lambda Q \frac{\partial y}{\partial x} = 0,$$

d'où :

$$\frac{\partial y}{\partial \lambda}\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - y\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial y}{\partial \lambda}\right) = y^2 Q,$$

et en intégrant de a à b :

$$\left[\frac{\partial y}{\partial \lambda}\frac{\mathrm{d}y}{\mathrm{d}x} - y\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial y}{\partial \lambda}\right)\right]_a^b > 0.$$

Puisque $y(a) = y(b) = 0$ et $\frac{\partial y}{\partial \lambda}(a,\lambda) = 0$, on a bien l'inégalité stricte recherchée.

Avec les notations précédentes, on a les relations d'orthogonalité :

$$\int_a^b y_n(x) y_m(x) \,\mathrm{d}x = 0$$

dès que $m \neq n$.

En effet, de :

$$y_n'' + \lambda_n Q y_n = 0 \quad \text{et} \quad y_m'' + \lambda_m Q y_m = 0,$$

avec $y_n(a) = y_n(b) = y_m(a) = y_m(b) = 0$, on déduit que :

$$0 = [y_m y_n' - y_n y_m']_a^b = (\lambda_m - \lambda_n) \int_a^b Q y_n y_m \,\mathrm{d}x.$$

À partir de ces remarques, on pourrait développer pour les fonctions f sur $[a; b]$ une analyse harmonique en série de fonctions orthogonales :

$$f(x) \sim \sum_{n=1}^{\infty} c_n(f) y_n(x),$$

avec $c_n(f) = \int_a^b f(x) y_n(x) Q(x) \, \mathrm{d}x$, qui généraliserait la théorie classique des séries Fourier.

Table des matières