Table des matières

Liste des notations et symboles

$\langle \cdot, \cdot \rangle$	produit scalaire.
$\\| \cdot \\|$	norme.
$\oplus$	somme directe (orthogonale).
$A^{\perp}$	le complémentaire orthogonal de A.
$\overline{A}$	l'adhérence de A.
$\overset{o}{A}$	l'intérieur de A.
$B(a, r)$	la boule ouverte de centre a et de rayon r.
$\bar{B}(a, r)$	la boule fermée de centre a et de rayon r.
$\mathcal{L}(X, Y)$	l'espace des opérateurs continus (applications linéaires continues) de X dans Y.
$\mathcal{L}(X)$	$= \mathcal{L}(X, X)$.
c	l'espace des suite convergentes.
c_0	l'espace des suites qui convergent vers 0.
$\mathbb{N}$	l'espace des entiers naturels.
$\mathbb{Z}$	l'ensemble des entiers relatifs.
$\mathbb{Q}$	l'ensemble des nombres rationnels.
$\mathbb{R}$	l'ensemble des nombres réels.
$\mathbb{C}$	l'espace des nombres complexes.
$\mathbb{K}$	désigne $\mathbb{R}$ ou $\mathbb{C}$.
$\mathfrak{Re}(z)$	la partie réelle du nombre complexe z.
$\mathfrak{Im}(z)$	la partie imaginaire du complexe z.
$C(\Omega)$	l'espace des fonctions continues sur Ω.
$C^k(\Omega)$	l'espace des fonctions dont la différentielle d'ordre k est continue sur Ω.
$C^\infty(\Omega)$	l'espace des fonctions indéfiniment différentiables sur Ω.
$C_c(\Omega)$	l'espace des fonctions continues et à support compact sur Ω.
$\mathcal{D}(\Omega)$	l'espace des fonctions indéfiniment différentiables et à support compact sur Ω.
$d(x, y)$	la distance entre x et y.
$d(x, A)$	$= \inf_{y \in A} d(x, y)$, la distance de x à A.
$\emptyset$	l'ensemble vide.
$H^m(\Omega)$	l'espace de Sobolev d'ordre m.
$H_0^m(\Omega)$	l'adhérence de $\mathcal{D}$ dans H^m.
$\operatorname{Im} T$	l'image de l'opérateur T.
$\operatorname{Ker} T$	le noyau de l'opérateur T.
ℓ^∞	l'espace des suites bornées.
ℓ^p	l'espace des suites, p-sommable
$L^\infty(E)$	l'espace des classes de fonctions essentiellement bornées sur E.
$L^p(E)$	l'espace des classes de fonctions p-intégrable sur E.
∂A	$= Fr(A)$ la frontière de A.
$\sigma(T)$	le spectre de l'opérateur T.

$\sigma_p(T)$	le spectre ponctuel de l'opérateur T.
$\text{Vect}(M)$	l'ensemble des combinaisons linéaires finies d'éléments de M.
T^*	l'operateur adjoint de T.
X'	$X^* = \mathcal{L}(X, \mathbb{K})$, l'espace dual topologique de X .
X''	$X^{**} = \mathcal{L}(X', \mathbb{K})$, l'espace bidual topologique de X .

Chapitre 1

Espaces de Banach et applications linéaires

1.1 Préliminaires

Nous allons maintenant rappeler quelques-unes des inégalités simples, mais importantes que nous allons utiliser dans les chapitres suivants.

(T_1) (L'inégalité triangulaire) $\forall\, a, b \in \mathbb{C}$, $|a+b| \leq |a| + |b|$.

Preuve : Comme $|\mathfrak{Re}(z)| \leq |z|$ pour tout $z \in \mathbb{C}$, le résultat s'obtient en écrivant

$$|a+b|^2 = (a+b)(\bar{a}+\bar{b}) \leq |a|^2 + |b|^2 + 2|a||b|. \qquad \blacksquare$$

(T_2)

$$\frac{|a+b|}{1+|a+b|} \leq \frac{|a|}{1+|a|} + \frac{|b|}{1+|b|} \quad \forall\, a, b \in \mathbb{C}. \tag{1.1.1}$$

Preuve : Soit $f(t) = t(1+t)^{-1}$ for $t > -1$. Comme $f'(t) = (1+t)^{-2}$, $f(t)$ est croissante sur $(-1, +\infty)$. D'après l'inégalité précédente on a
$f(|a+b|) \leq f(|a|+|b|)$, d'où

$$\frac{|a+b|}{1+|a+b|} \leq \frac{|a|+|b|}{1+|a|+|b|} = \frac{|a|}{1+|a|+|b|} + \frac{|b|}{1+|a|+|b|} \leq \frac{|a|}{1+|a|} + \frac{|b|}{1+|b|}. \qquad \blacksquare$$

(T_3) (L'inégalité de Young) Soit $p > 1$, $1/p + 1/q = 1$, $a \geq 0$, $b \geq 0$. Alorsn

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

avec égalité si et seulement si $a^p = b^q$.

Démonstration: Soit $f(t) = 1 - \lambda + \lambda t - t^{\lambda}$, où $\lambda = 1/p$ and $t \geq 0$. Alors $f'(t) < 0$ for $0 < t < 1$ and $f'(t) > 0$ for $t > 1$. Comme $f(t) \geq f(1) = 0$, avec égalité si et seulement si $t = 1$. D'où on a

$$t^{\lambda} \leq (1 - \lambda) + \lambda t \tag{1.1.2}$$

pour $t \geq 0$. Si $b = 0$ alors $ab = 0 \leq a^p/p$. Si $b > 0$ on pose $t = a^p b^{-q}$ in (1.1.2) d'où le résultat. ∎

(T_4) (L'inégalité de Hölder) Soit $p > 1$, $1/p + 1/q = 1$, $\xi_k, \eta_k \in \mathbb{K}$ pour $k = 1, 2, \cdots, n$. Alors

$$\sum_{k=1}^{n} |\xi_k \eta_k| \leq \left(\sum_{k=1}^{n} |\xi_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |\eta_k|^q \right)^{\frac{1}{q}}, \tag{1.1.3}$$

on a aussi,

$$\sum_{k=1}^{n} |\xi_k \eta_k| \leq \left(\sum_{k=1}^{n} |\xi_k| \right) \max |\eta_k|. \tag{1.1.4}$$

Démonstration: L'inéquation (1.1.3) est connue sous le nom d'inégalité de Hölder. L'inégalité (1.1.4) est triviale et est un cas particulier de l'inégalité de Hölder pour le cas $p = 1$. Pour montrer (1.1.3) on pose

$$A = (\sum |\xi_k|^p)^{1/p}, \quad B = (\sum_{k=0}^{n} |\eta_k|^q)^{1/q},$$

Si $AB = 0$ on a $A = 0$ ou $B = 0$. Dans les deux cas les deux membres de (1.1.3) sont égaux à 0. On suppose donc que $AB > 0$, d'après l'inégalité de Young,

$$\frac{|\xi_k|}{A} \frac{|\eta_k|}{B} \leq \frac{|\xi_k|^p}{pA^p} + \frac{|\eta_k|^q}{qB^q},$$

d'où $\sum_{k=0}^{n} |\xi_k||\eta_k| \leq (1/p + 1/q)AB = AB$, qui n'est autre que (1.1.3). ∎

1.1.3 Remarque

En utilisant l'argument de la preuve de l'inégalité de Hölder's (1.1.3)*, on a les deux formes suivantes de l'inégalité de Hölder*

$$\sum_{k=1}^{\infty} |\xi_k \eta_k| \leq \left(\sum_{k=1}^{\infty} |\xi_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |\eta_k|^q \right)^{\frac{1}{q}}, \tag{1.1.5}$$

$\xi_k, \eta_k \in \mathbb{K}$ *for* $k \in \mathbb{N}$, $p > 1$, $1/p + 1/q = 1$;

$$\int_E |x(t)y(t)|dt \leq \left(\int_E |x(t)|^p dt \right)^{\frac{1}{p}} \left(\int_E |y(t)|^q dt \right)^{\frac{1}{q}}, \tag{1.1.6}$$

où E *est un ensemble mesurable (au sens de lebesgue) dans* $\mathbb{R}$, $x(t)$ *et* $y(t)$ *des fonctions mesurables sur* E, $p > 1$, $1/p + 1/q = 1$.

1.1.4 Remarque

Lorsque $p = q = 2$, on a (1.1.3) ou (1.1.5) (1.1.6) est l'inégalité de Cauchy-Schwartz .

(T_5) (L'inégalité de Minkowski) Soit $p \geq 1$, $\xi_k, \eta_k \in \mathbb{K}$ for $k = 1, 2, \cdots, n$. Alors

$$\left(\sum_{k=1}^{n} |\xi_k + \eta_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{n} |\xi_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |\eta_k|^p\right)^{\frac{1}{p}}, \qquad (1.1.7)$$

Démonstration: Le cas $p = 1$ est trivial. On suppose que $p > 1$. d'après l'inégalité triangulaire et celle de Hölder (1.1.3),

$$\begin{aligned}
\sum_{k=0}^{n} |\xi_k + \eta_k|^p &\leq \sum_{k=0}^{n} |\xi_k||\xi_k + \eta_k|^{p-1} + \sum_{k=0}^{n} |\eta_k||\xi_k + \eta_k|^{p-1} \\
&\leq \left(\sum_{k=0}^{n} |\xi_k|^p\right)^{1/p} \left(\sum_{k=0}^{n} |\xi_k + \eta_k|^{q(p-1)}\right)^{1/q} \\
&\quad + \left(\sum_{k=0}^{n} |\eta_k|^p\right)^{1/p} \left(\sum_{k=0}^{n} |\xi_k + \eta_k|^{q(p-1)}\right)^{1/q},
\end{aligned}$$

D'où l'inégalité de Minkowski. ∎

1.1.6 Remarque

les deux formes suivantes de l'inégalité de Minkowski ont lieu :

$$\left(\sum_{k=1}^{\infty} |\xi_k + \eta_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |\eta_k|^p\right)^{\frac{1}{p}}, \qquad (1.1.8)$$

où $\xi_k, \eta_k \in \mathbb{K}$, $k \in \mathbb{N}$, $p \geq 1$;

$$\left(\int_E |x(t) + y(t)|^p dt\right)^{\frac{1}{p}} \leq \left(\int_E |x(t)|^p dt\right)^{\frac{1}{p}} + \left(\int_E |y(t)|^p dt\right)^{\frac{1}{p}}, \qquad (1.1.9)$$

où $x(t)$ et $y(t)$ sont des fonctions mesurables de E dans $\mathbb{C}$, $p \geq 1$.

Démonstration: On va montrer (1.1.8). (1.1.9) se démontre de la même manière Le résultat est trivial si le membre de droite de l'inégalité (1.1.8) est infini. D'autre part, $\forall\, a, b \in \mathbb{C}$, on a

$$(|a| + |b|)^p \leq (2\max(|a|, |b|))^p \leq 2^p(|a|^p + |b|^p),$$

ce qui donne

$$\sum_{k=1}^{\infty} |\xi_k + \eta_k|^p \leq 2^p \left(\sum_{k=1}^{\infty} |\xi_k|^p + \sum_{k=1}^{\infty} |\eta_k|^p\right).$$

Si le membre de gauche de l'inégalité (1.1.8) est infini, l'un au moins des membres de droite de l'inégalité (1.1.8) est infini. On peut donc supposer que toutes les sommes sont finies. En utilisant le même argument que dans la preuve de (1.1.7), on obtient (1.1.8). ■

(T_6) Soit $0 < p \leq 1$, $\xi_k, \eta_k \in \mathbb{C}$ pour $k \in \mathbb{N}$. Alors

$$\sum_{k=1}^{m} |\xi_k + \eta_k|^p \leq \sum_{k=1}^{m} |\xi_k|^p + \sum_{k=1}^{m} |\eta_k|^p, \tag{1.1.10}$$

où m est fini ou infini.

Démonstration: Il suffit de montrer l'inégalité $(a+b)^p \leq a^p + b^p$ pour $0 < p \leq 1$, $a \geq 0$, $b \geq 0$. Pour montrer cela, on considère

$$f(t) = 1 + t^p - (1+t)^p$$

pour $t \geq 0$. En considérant la dérivée, on trouve que $f(t) \geq 0$ for $t \geq 0$, i.e. $(1+t)^p \leq 1 + t^p$. Si $b = 0$ alors $(a+b)^p = a^p + b^p$, et si $b > 0$ on pose $t = a/b$ dans $(1+t)^p \leq 1 + t^p$. ■

1.2 Espaces vectoriels et opérateurs linéaires

1.2.1 Des fonctions aux espaces de fonctions

Dans cette section, notre discussion est générale et pas très précis. Pour l'instant, notre objectif est de voir une grande image.

Dans l'Antiquité, les propriétés des *nombres* étaient d'une grande importance. Que $\sqrt{2}$ soit rationnel, par exemple, était une source de grandes discussions. Plus tard, avec le développement du calcul au XVIIe siècle, l'accent passe des nombres aux *fonctions*. Une fonction associe des nombres d'une certaine manière, et c'est cette association qui nous importe maintenant, plus que les nombres. Plus tard, dans la moitié du XIXe, l'intêret s'est déplacé encore plus loin, de l'étude des fonctions individuellement (étudier leur dérivabilité, intégrabilité) aux *espaces de fonctions*. On peut dire qu'un espace fonctionnel, mène fonctions d'intérêt en un objet géométrique. La géométrie de l'espace des fonctions reflète les propriétés importantes des fonctions. Cela a conduit au développement de *l'analyse fonctionnelle*.

En analyse fonctionnelle, nous considérons les fonctions comme des points ou des vecteurs dans un espace de fonctions. Comme, on peut additionner des fonctions défines sur un domaine commun en définissant $(f+g)(x) := f(x) + g(x)$ et les multiplier en défnissant $(af)(x) = af(x)$, on voit ainsi un espace fonctionnel est un *espace vectoriel*.

De plus, on peut envisager une sorte de *distance* sur un espace fonctionnel, qui quantifie les similarités (ou dissimilarités) des fonctions. Le choix d'une telle distance dépend de l'application. Un choix naturel de la distance entre f et g est la *"distance sup"*

$$\|f - g\|_\infty := \sup_{x \in [a,b]} |f(x) - g(x)|.$$

Il s'agit clairement d'une mesure, donc un espace de fonctions devient non seulement un espace vectoriel mais aussi un *espace métrique*. Ces espaces seront appelés *des espaces vectoriels normés*. Un autre choix de distance serait par exemple

$$\|f - g\|_1 := \int_a^b |f(x) - g(x)| \, dx$$

Heuristiquement, une norme $\| f - g\|_\infty$ petite force les valeurs de f et g à rester proches partout sur $[a, b]$, tandis qu'une mesure $\|f - g\|_1$ petite force les valeurs de f et g à rester proche seulement "en moyenne".

1.2.2 Exemples d'espaces vectoriels

Des exemples classiques d'espaces vectoriels sont donnés dans les cours d'algèbre linéaire comme $\mathbb{R}^n$ et $\mathbb{C}^n$. Voici quelques exemples d'espaces de fonctions :

1. $F = \{$toutes les fonctions de $\mathbb{K} \to \mathbb{K}\}$. Cet espace est immense, et n'est pas en général étudié.
2. {toutes les solutions homogènes d'une équation linéaire aux dérivées partielles }
3. $L^1[a, b] = \{$les (classes) fonctions Lebesgue intégrable sur $[a, b]\}$
4. $L^\infty[a, b] = \{$ les (classes) fonctions essentiellement bornées sur $[a, b]\}$
5. $C[a, b] = \{$toutes les fonctions continues sur $[a, b]\}$
6. $C^1[a, b] = \{$toutes les fonctions continûment dérivables sur $[a, b]\}$
7. $C^\infty[a, b] = \{$toutes les fonctions indéfiniment différentiables sur $[a, b]\}$
8. $\mathbb{K}[x] = \{$tous les polynômes à une variable$\}$
9. $\mathbb{K}_n[x] = \{$tous les polynômes à une variable de degré au plus $n\}$

Il y a aussi de nombreux exemples naturels d'espaces de suites (à vérifier !) :

1. $\mathbb{K}^\mathbb{N} = \{ (x_n)_{n\in\mathbb{N}}, |$ les suites à valeurs dans $\mathbb{K}$ $\}$. Cet espace est trop gros. On s'intéressera plutôt à certains de ses sous-espaces :
2. $\ell^p(\mathbb{K}, \mathbb{N}) = \{$ des suites absolument $(x_n)_{n\in\mathbb{N}}$, p-sommables, c'est à dire $\sum_{n\in\mathbb{N}} |x_n|^p < \infty\}$ $p \in \mathbb{R}_+$
3. $\ell^\infty(\mathbb{K}, \mathbb{N}) = \{$des suites bornées à valeurs dans $\mathbb{K}\}$

4. $c(\mathbb{K}, \mathbb{N}) = \{$toutes les suites convergentes d'éléments de $\mathbb{K}\}$
5. $c_0(\mathbb{K}, \mathbb{N}) = \{$des suites d'éléments de $\mathbb{K}$ convergeant vers $0\}$
6. $c_{00}(\mathbb{K}, \mathbb{N}) = \{$toutes les suites d'éléments de $\mathbb{K}$ à support fini$\}$ i.e. $x = (x_n)_{n\in\mathbb{N}} \in c_{00}$ s'il existe $N(x) \in \mathbb{N}$ tel que $x_n = 0$ pour tout $n \geq N(x)$.

1.2.3 Sous-espace vectoriel

Un sous-espace vectoriel d'un espace vectoriel sue un corps $\mathbb{K}$ est un sous-ensemble stable pour les opérations d'addition de vecteurs et de multiplication par des scalaires :

1.2.1 DÉFINITION
Un sous-ensemble E_1 d'un espace vectoriel E est appelé *sous-espace vectoriel* si pour tout $x, y \in E_1$, $a, b \in \mathbb{K}$, on a $ax + by \in E_1$.

1.2.2 EXEMPLE. On peut vérifier que :
$\mathbb{K}_n[x] \subset \mathbb{K}[x] \subset C^\infty[a,b] \subset C^k[a,b] \subset C[a,b] \subset L^\infty[a,b] \subset L^p[a,b] \subset F,$

$c_{00} \subset \ell^p \subset c_0 \subset c \subset \ell^\infty \subset \mathbb{K}^{\mathbb{N}}.$
Tous ces inclusions sont *inclusions de sous-espace*, ie $\mathbb{K}_n[x]$ est un sous-espace de $\mathbb{K}[x]$, etc... (le vérifier !)

1.2.3 EXEMPLE. Soit E un espace vectoriel. Montrer que $\{0\}$ et E sont des sous-espaces de E. Montrer que l'intersection d'une collection arbitraire de sous-espaces de E est encore un sous-espace E.

1.2.4 Base de Hamel

Comme nous le savons, tout espace vectoriel de *dimension finie* E admet une base $\{x_1, \ldots, x_n\}$. Une base est un sous-ensemble linéairement indépendant maximal de vecteurs de E. Le nombre n d'éléments de base est appelé la dimension de E ; ce nombre est indépendant du choix de la base. Chaque vecteur $x \in E$ peut être uniquement exprimé en une combinaison linéaire d'éléments de la base :

$$x = \sum_{k=1}^{n} x_i a_k, \quad \text{pour certains } a_k \in \mathbb{K}. \tag{1.2.1}$$

La notion de base peut être généralisée à tout espace vectoriel.

1.2.4 Définition (base de Hamel)

Un sous-ensemble $\mathcal{B}$ d'un espace vectoriel E est appelé *base de Hamel* de E si chaque vecteur $x \in E$ peut être uniquement exprimé comme une combinaison linéaire finie de certains éléments de $\mathcal{B}$:

$$x = \sum_{k=1}^{n} a_k x_k \tag{1.2.2}$$

pour certains scalaires non nuls $a_k \in \mathbb{K}$ et vecteurs $x_k \in \mathcal{B}$.

1.2.5 Exercice Montrer que chacune des deux affirmations suivantes donne une définition équivalente de la base de Hamel :

1. Une base de Hamel est un sous-ensemble linéairement indépendant maximal [1] $\mathcal{B} \subset E$.
2. Une base de Hamel est un sous-ensemble linéairement indépendant $\mathcal{B}$ de E qui *engendre* E. Ceci signifie que le $\mathrm{Vect}(\mathcal{B})$, défini par

$$\mathrm{Vect}(\mathcal{B}) := \left\{x = \sum_{k=1}^{n} a_k x_k : \ a_k \in \mathbb{K},\ x_k \in \mathcal{B},\, n \in \mathbb{N}\right\},$$

coïncide avec E.

Comme nous n'avons aucune topologie sur E, il faut tenir compte seulement des sommes finies dans (1.2.2). Cette exigence est trop forte, ce qui rend les bases de Hamel peut intéréssante dans la pratique. Nous reviendrons sur une notion plus intéréssante, celle de base de Schauder.

1.2.6 Proposition

Chaque espace vectoriel E admet une base de Hamel.

Pour un espace de dimension finie E, ce résultat est généralement prouvé en algèbre linéaire (du premier cycle) à l'aide d'une récurrence. On rajoute à chaque étape un élément du complémentaire du Vect des vecteurs déjà construits jusqu'à recouvrir l'espace E. Cet argument peut être utilisé en dimension infinie où la récurrence habituelle est remplacée par *la récurrence transfinie*. la récurrence transfinie est mieux appréhendée avec le lemme de Zorn :

1.2.7 Lemme (lemme de Zorn)

Tout ensemble partiellement ordonné inductif non-vide possède au moins un élément maximal.

---

1. l'indépendance linéaire signifie que tout *sous-ensemble fini* de $\mathcal{B}$ est linéairement indépendant dans le sens habituel. En d'autres termes, si $\sum_{k=1}^{n} a_k x_k = 0$ pour certains $a_k \in \mathbb{K}$, $x_k \in \mathcal{B}$, $n \in \mathbb{N}$, alors tout $a_k = 0$ pour tout $k \in \{1, \ldots, n\}$.

voir la section 1.2.9 pour des explications sur le lemme de Zorn et sa relation avec l'axiome du choix et la démonstration de la proposition 1.2.6.

Comme dans le cas de dimension finie, la cardinalité de la base de Hamel E est appelée *dimension* de E ; on peut montrer que la dimension est indépendante du choix d'une base de Hamel.

1.2.8 EXEMPLE. Nous considérons quelques exemples d'espaces vectoriels donnée dans la section 1.2.2.

1. $\dim_{\mathbb{R}}(\mathbb{R}^n) = n$, $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$, $\dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$.
2. $\dim(\mathbb{K}_n[X]) = n+1$, les monômes $\{1, x, x^2, \ldots, x^n\}$ forment une base.
3. $\dim(\mathbb{K}[X]) = \infty$, le système de monômes $\{1, x, x^2, \ldots\}$ forme une base de Hamel.
4. $\dim(c_{00}) = \infty$, les vecteurs de coordonnées $(0, \ldots, 0, 1, 0, \ldots)$ forment une base de Hamel.

1.2.9 REMARQUE

Malheureusement, mise à part des espaces comme $\mathbb{K}[X]$ ou c_{00} (qui sont isomorphes - pourquoi ?) aucune construction explicite de base de Hamel n'est connue en générale pour les espaces vectoriels de dimension infinie. Il serait intéréssant d'avoir une construction d'une base de Hamel de $C[0,1]$, par exemple. D'autre part, les bases Hamel sont généralement non dénombrables (voir conséquence du Théorème de Baire).

1.2.5 Espaces quotients

———————————————-

La notion d'espace quotient permet de négliger certaines directions dans les espaces vectoriels tout en gardant la structure vectorielle,

1.2.10 DÉFINITION (ESPACE QUOTIENT)

Soit E_1 un sous-espace d'un espace vectoriel E. Considérons une relation d'équivalence sur E définie par

$$x \sim y \quad \text{si} \quad x - y \in E_1.$$

L'espace *quotient* E/E_1 est alors l'ensemble des classes d'équivalences $[x]$ où $x \in E$.

L'espace quotient a une structure naturelle d'espace vectoriel, si on le munit des opérations

$$[x] + [y] := [x+y], \quad a[x] := [ax] \quad \text{pour } x, y \in E,\ a \in \mathbb{R}.$$

La dimension de l'espace quotient de E par E_1 est appelée la *codimension* de E_1 dans E, ainsi

$$\operatorname{codim}(E_1) := \dim(E/E_1).$$

1.2.11 EXEMPLE. Montrer que les opérations ci-dessus sont bien définies, et que l'espace quotient est en effet un espace vectoriel.

1.2.12 REMARQUE
1. Observons que $[x]$ est un sous-espace affine :

$$[x] = x + E_1 := \{x + h : h \in E_1\}.$$

2. La définition de la relation d'équivalence $x \sim y$ est censée ignorer les directions E_1, et donc d'identifier les points x, y s'il ne diffèrent que par un vecteur de E_1.

3. De l'algèbre linéaire, nous savons que si E est de dimension finie alors tout sous-espace E_1 vérifie

$$\dim(E_1) + \operatorname{codim}(E_1) = \dim(E).$$

1.2.13 EXEMPLE (L'ESPACE L^1). La notion d'espace quotient est très utile lorsque nous définissons *l'espace de fonctions intégrables* $L^1 = L^1(\Omega, \Sigma, \mu)$ où (Ω, Σ, μ) est un espace de mesure arbitraire. Nous considérons d'abord

$$E := \{\text{toutes les fonctions intégrables } f \text{ sur } (\Omega, \Sigma, \mu)\}.$$

Pour identifier les fonctions qui sont égales μ-presque partout, nous considérons le sous-espace que nous tenons à négliger :

$$E_1 := \{\text{toutes les fonctions } f = 0 \ \ \mu\text{-presque partout}\}$$

Alors, nous définissons

$$L^1 = L^1(\Omega, \Sigma, \mu) := E/E_1.$$

De cette façon, les éléments de L^1 ne sont, à proprement parler, pas des fonctions mais des classes d'équivalences.[2] Mais dans la pratique, on pense à un $f \in L^1$ comme une fonction, en gardant à l'esprit que les fonctions qui coïncider μ -presque partout sont "sont confondues".

1.2.14 EXEMPLE (L'ESPACE L^∞). Une procédure similaire est utilisée pour définir *l'espace des fonctions essentiellement bornées* $L^\infty = L^\infty(\Omega, \Sigma, \mu)$. Une fonction réelle f sur Ω est dite essentiellement bornée s'il existe une fonction bornée g sur Ω tel que $f = g$

2. De plus, les fonctions représentatives de f dans L^1 peut prendre des valeurs infinies, aussi. Cependant, chaque fonction intégrable est finie. Ainsi, chaque fonction est équivalente à une fonction qui est finie partout.

μ-presque partout. Comme dans l'exemple précédent, nous considérons l'espace vectoriel

$$E := \{\text{de toutes les fonctions essentiellement bornées } f \text{ sur } (\Omega, \Sigma, \mu)\}$$

et le sous-espace que nous tenons à négliger :

$$E_1 := \{\text{de toutes les fonctions } f = 0 \quad \mu\text{-presque partout}\}$$

Alors, nous définissons

$$L^\infty = L^\infty(\Omega, \Sigma, \mu) := E/E_1.$$

1.2.15 EXEMPLE. Comme nous le savons, l'espace c_0 des suites qui convergent vers 0 est un sous-espace de l'espace c des suites convergentes. Notons que

$$\text{codim}\,(c_0) = 1.$$

En effet, chaque suite $x \in c$ peut être représentée de manière unique comme

$$x = a\mathbb{1} + z \quad \text{pour un certain } a \in \mathbb{K}, \text{ et un certain } z \in c_0$$

où $\mathbb{1} = (1, 1, \ldots)$. (Comment peut-on choisir la valeur de a?). D'où

$$[x] = a[\mathbb{1}] + [z] = a[\mathbb{1}].$$

Il s'ensuit que tout élément $[x] \in c/c_0$ est un multiple constant de l'élément $[\mathbb{1}]$. Par conséquent, $\dim(c/c_0) = 1$.

Cet exemple montre que l'espace c_0 représente à peut près tout l'espace c, sauf pour une dimension donnée par les suites constantes. Cela explique pourquoi l'espace c est rarement utilisé dans la pratique ; on préfère travailler avec c_0, qui est très semblable à c, avec l'avantage que l'on connaît les limites de toutes ses suites à savoir 0.

1.2.6 Opérateur linéaire (application linéaire)

Il s'agit d'un examen rapide de la notion d'algèbre linéaire classique.

1.2.16 DÉFINITION (OPÉRATEUR LINÉAIRE)
Une application $T : E \to F$ entre deux espaces vectoriels E et F est appelée *opérateur linéaire (application linéaire)* si elle préserve les opérations d'addition de vecteurs et de multiplication par des scalaires,

$$T(ax + by) = aT(x) + bT(y) \quad \text{pour tout } x, y \in E, \ a, b \in \mathbb{K}.$$

Le *noyau* et *l'image* de T sont définis respectivement par [3]

$$\ker(T) = \{x \in E : \ Tx = 0\}, \quad \text{Im}(T) = \{Tx : \ x \in E\}.$$

3. On écrit généralement Tx au lieu de $T(x)$

1.2.17 EXEMPLE (OPÉRATEUR DIFFÉRENTIEL). L'exemple le plus simple d'un opérateur différentiel est obtenu en prenant la dérivée d'une fonction :

$$T(f) = f'.$$

Un tel opérateur est bien défini par exemple sur l'espace des polynômes $T : \mathbb{K}[x] \to \mathbb{K}[x]$. Mais en général, on préfère avoir un opérateur différentiel sur un espace plus grand, par exemple $T : C^1[0,1] \to C[0,1]$ ou il est également bien défini.

1.2.18 EXEMPLE (PLONGEMENT ET APPLICATION QUOTIENT). étant donné un sous-espace E_1 d'un espace vectoriel E, il existe deux opérateurs linéaires canoniques qui lui sont associés :

1. *le plongement ou l'injection canonique* $\iota : E_1 \to E$, qui agit comme l'identité $\iota(x) = x$;
2. l'application quotient $p : E \to E/E_1$, définie par $p(x) = [x]$.

1.2.19 REMARQUE
L'application ι est injective et l'application p est surjective (à vérifier)

1.2.20 Exercice (espaces de suite) Sur un espace de suite telles que c_{00}, c_0 , c, ℓ^∞, ℓ^1, on peut définir les oprérateurs : de décalage à droite (resp de décalage à gauche) par

$$R(x) = (0, x_1, x_2, \ldots); \quad (\text{resp. } L(x) = (x_2, x_3, \ldots,)) \quad \text{pour } x = (x_1, x_2, \ldots).$$

1.2.21 Exercice Calculer les images et les noyaux de l'intégration, l'application quotient, et les opérateurs de décalage dans les exemples ci-dessus.

1.2.7 Exercices supplémentaires

1.2.22 Exercice Montrer que l'intersection d'une collection arbitraire de sous-espaces d'un espace vectoriel E est encore un sous-espace de E.

1.2.23 Exercice Montrer que tout sous-ensemble linéairement indépendant d'un espace vectoriel E peut être prolongé en une base de Hamel de E (le théorème de la base incomplète).

1.2.24 Exercice (Sous-espaces supplémentaires) Soit E_1 un sous-espace d'un espace vectoriel E. Montrer qu'il existe un sous-espace E_2 de E tel que

$$E_1 \cap E_2 = \{0\}, \quad \mathrm{Vect}(E_1 \cup E_2) = E.$$

(Indication : étendre une base de Hamel E_1 à E ; utiliser l'extension pour construire E_2). Les sous-espaces E_1 et E_2 sont appelés *espaces supplémentaires*. On note dans ce cas $E = E_1 \oplus E_2$.
Montrer que E_1 et E_2 sont supplémentaires si et seulement si tout vecteur $x \in E$ peut être représenté de manière unique comme une somme

$$x = x_1 + x_2 \quad \text{avec } x_1 \in E_1,\ x_2 \in E_2.$$

1.2.25 Exercice (Injectivisation) Il s'agit d'une version linéaire du théorème fondamental sur les homomorphismes de groupes. Considérons un opérateur linéaire $T : E \to F$ agissant entre espaces vectoriels E et F. L'opérateur T n'est peut-être pas injectif, nous aimerions le rendre injectif. Pour cela, nous considérons l'application $\tilde{T} : E/\ker T \to F$ qui envoie chaque classe $[x]$ sur le vecteur Tx, c'est à dire $\tilde{T}[x] = Tx$.

(i) Montrer que $\tilde{T}$ est bien définie, c'est à dire $[x_1] = [x_2]$ implique $Tx_1 = Tx_2$.

(ii) Vérifiez que $\tilde{T}$ est un opérateur linéaire et injectif.

(iii) Vérifier que si T est surjective alors $\tilde{T}$ est aussi surjective, et donc $\tilde{T}$ est un isomorphisme linéaire entre $E/\ker T$ et F.

(iv) Montrer que $T = \tilde{T} \circ p$, où $p : E \to E/\ker T$ est l'application quotient. En d'autres termes, chaque opérateur linéaire peut être représenté comme une composition d'un opérateur surjectif et d'un opérateut injectif.

1.2.8 Annexe

1.2.9 Ensembles partiellement ordonnés

1.2.26 Définition
Soient X un ensemble et R une relation binaire sur X : $R \subseteq X \times X$. On dit que R est une *relation d'ordre partiel*, ou tout simplement un *ordre partiel,* si elle vérifie les conditions suivantes :

1. R est réflexive : xRx pour tout $x \in X$.
   (De façon équivalente : $\Delta \subseteq R$, où $\Delta = \{(x,x) \colon x \in X\}$ est la *diagonale* de $X \times X$).
2. R est antisymérique : si xRy et yRx, alors $x = y$.
   (En d'autres termes : $R^{-1} = R$, où $R^{-1} = \{(y,x) \colon (x,y) \in R\}$ est la *relation réciproque* de R).
3. R est transitive : si xRy et yRz, alors xRz.
   (On peut écrire cette propriété comme $R \circ R \subseteq R$, où la *composition des relations*, R et S, est définie par $R \circ S = \{(x,z) \colon \exists y \ \ (x,y) \in R \text{ et } (y,z) \in S\}$).

D'habitude, on note une relation d'ordre partiel par le symbole $\leq$, où bien $\preceq$. Un couple $(X, \leq)$ composé d'un ensemble X et un ordre partiel $\leq$ sur X est appelé *ensemble partiellement ordonné*.

1.2.27 DÉFINITION
Un ordre partiel $\leq$ sur un ensemble X est dit *ordre total* si deux éléments quelconque de X sont comparables :

$$\forall x, y \in X, \quad x \leq y \text{ ou bien } y \leq x.$$

Dans ce cas, on dit aussi que l'ensemble $(X, \leq)$ est *totalement ordonné.*

1.2.28 EXEMPLE. La droite réelle $\mathbb{R}$, munie de l'ordre usuel, est totalement ordonnée.

1.2.29 EXEMPLE. Soit X ensemble quelconque. L'ensemble $\mathcal{P}(X)$ de toutes les parties de X est partiellement ordonné par la relation d'inclusion :

$$A \leq B \iff A \subseteq B.$$

Si $|X| \geq 2$, l'ordre ci-dessus n'est pas total : il y a toujours au moins deux sous-ensembles $A, B \subseteq X$ qui ne sont pas comparables :

$$A \not\subseteq B \text{ et } B \not\subseteq A.$$

1.2.30 EXEMPLE. Un élément x d'un ensemble partiellement ordonné X est dit *maximal* s'il n'y a pas d'éléments strictement plus grands que x :

$$\forall y \in X, \quad x \leq y \Rightarrow x = y.$$

1.2.31 EXEMPLE. L'élément 1 est maximal dans le segment $[0, 1]$ muni de l'ordre usuel.

1.2.32 EXEMPLE. Soit V un espace vectoriel (sur un corps $\mathbb{K}$). Notons $\mathfrak{B}$ la famille de tous les sous-ensembles linéairement indépendents de V :

$$\begin{aligned} X \in \mathfrak{B} \quad \iff \quad & X \subseteq V \text{ et } \forall n \in \mathbb{N}_+,\ \forall \lambda_i \in \mathbb{K},\ \forall x_i \in X, \\ & i = 1, 2, \ldots, n, \quad x_i \neq x_j \text{ pour } i \neq j, \\ & \text{si } \sum_{i=1}^{n} \lambda_i x_i = 0, \text{ alors } \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0. \end{aligned}$$

Alors l'ensemble $\mathfrak{B}$, muni de la relation de l'inclusion, est partiellement ordonné, et un élément $X \in \mathfrak{B}$ est maximal si est seulement si X est une base de V (c.à.d. X n'est contenu dans aucun ensemble linéairement indépendent strictement plus grand).

1.2.33 REMARQUE
La notion d'élément maximal est différente de celle de plus grand d'élément. Dans l'exemple précedent, même si $\dim V = 1$, il n'existe aucun plus grand élément dans $\mathfrak{B}$, mais il existe plusieurs éléments maximaux.

1.2.34 DÉFINITION
Soit $(X, \leq)$ un ensemble partiellement ordonné. Un sous-ensemble C de X est dit *chaîne* si la restriction de l'ordre partiel $\leq$ sur C est un ordre total. En d'autres termes, pour tous $x, y \in C$, ou bien $x \leq y$, ou bien $y \leq x$.

1.2.35 DÉFINITION
Un ensemble partiellement ordonné X est dit *inductif* si toute chaîne $C \subseteq X$ est majorée : il existe $x \in X$ tel que

$$\forall c \in C \;\; c \leq x$$

(un *majorant* pour C).

Le résultat suivant se trouve parmi les outils les plus fréquemment utilisés dans les mathématiques.

1.2.36 THÉORÈME (LEMME DE ZORN)
Tout ensemble partiellement ordonné inductif non-vide possède au moins un élément maximal.

1.2.37 Exercice Montrer le lemme de Zorn dans le cas où X est démonbrable, en utilisant une récurrence .

Voilà une application typique du lemme de Zorn.

1.2.38 THÉORÈME
Tout espace vectoriel possède une base (de Hamel).

Démonstration: Soit V un espace vectoriel quelconque sur un corps $\mathbb{K}$. Définissons l'ensemble partiellement ordonné $\mathscr{B}$ comme dans l'exemple 22.8 : $\mathscr{B}$ consiste de tous les sous-ensembles linéairement indépendents (libres) de V. Montrons que $\mathscr{B}$ est inductif. Soit $\mathscr{C}$ une chaîne dans $\mathscr{B}$, c'est-à-dire $\mathscr{C}$ est une sous-famille de $\mathscr{B}$ qui est totalement ordonné par la relation d'inclusion :

$$\forall A, B \in \mathscr{C}, \quad A \subseteq B \text{ ou } B \subseteq A.$$

Notons que $A \cup B$ appartient à $\mathscr{C}$ pour tous les $A, B \in \mathscr{C}$. Par induction finie, quel que soient $n \in \mathbb{N}$ et $A_1, A_2, \ldots, A_n \in \mathscr{C}$, on a $\cup_{i=1}^n A_i \in \mathscr{C}$.

Posons $X = \cup\mathscr{C} = \cup\{A \colon A \in \mathscr{C}\}$. Montrons que $X \in \mathscr{B}$, c.à.d. que X est libre. Soient $n \in \mathbb{N}$, $\lambda_i \in \mathbb{K}$, $x_i \in X$, $i = 1, 2, \ldots, n$, tels que $x_i \neq x_j$ pour $i \neq j$. Supposons que $\sum_{i=1}^n \lambda_i x_i = 0$. Il existe $A_i \in \mathscr{C}$ tels que $x_i \in A_i$, $i = 1, 2, \ldots, n$. L'ensemble $A = \cup_{i=1}^n A_i$ appartient à $\mathscr{C}$, et par conséquent A est libre. On en déduit : $\lambda_i = 0$ pour tous les $i = 1, 2, \ldots, n$.

On en conclut : la famille $\mathscr{B}$ est inductive. D'après le lemme de Zorn, il possède au moins un éléments maximal, X. En d'autres termes, X est un sous-ensemble libre de V qui n'est pas contenu dans un ensemble libre strictement plus grand. Il est bien connu (et montré aisément) qu'un tel X est une base pour V. ■

1.2.10 Axiome du choix

1.2.40 THÉORÈME (AXIOME DU CHOIX A)

Soit Γ une famille quelconque d'ensembles non-vides deux à deux disjoints. Alors il existe un ensemble $S \subseteq \bigcup_{A\in\Gamma} A$ tel que pour tout $A \in \Gamma$, l'intersection $A \cap S$ consiste en un seul point :

$$|A \cap S| = 1.$$

1.2.41 THÉORÈME (AXIOME DU CHOIX B)

Soit Γ une famille quelconque d'ensembles non-vides. Alors il existe une fonction $f \colon \Gamma \to \bigcup_{A\in\Gamma} A$ (une *fonction de choix*) telle que, pour tout $A \in \Gamma$,

$$f(A) \in A.$$

1.2.42 Exercice Vérifier que les deux formes de l'axiome du choix sont équivalentes.

1.2.43 Exercice Vérifier l'axiome du choix au cas où $\bigcup_{A\in\Gamma} A$ est démonbrable, en utilisant une récurrence .

Le lemme de Zorn est une des conséquences (même, une forme équivalente) de l'axiome de Choix. Il y en a d'autres.

C'est un *théorème,* dont la démonstration est rigoureuse et irréfutable, et elle s'appelle "un paradoxe" uniquement à cause de son caractère contre-intuitif.

En effet, il existe un résultat plus général et même plus étonnant. Notons B_r la boule fermée euclidien du rayon $r > 0$ dans l'espace $\mathbb{R}^3$:

$$\bar{B}_r = \{x \in \mathbb{R}^3 \colon \| x - a \| \leq r\}.$$

1.2.44 THÉORÈME (LE PARADOX DE BANACH ET TARSKI, 1924)

Soient $r > 0$ et $R > 0$ deux nombres réels strictement positifs. Alors il existe des partitions finies des boules B_r et B_R,

$$B_r = \cup_{i=1}^{n} A_i, \qquad B_R = \cup_{i=1}^{n} A'_i$$

en n parties, telles que

pour tout $i \in \{1, \ldots\}$, A'_i est isométrique à A_i.

Il est donc possible de découper une boule petite comme un pois ou comme une goutte d'eau en un nombre fini de morceaux, et de les rassembler de manière à obtenir une boule plus grande comme la terre, ou même comme le soleil.

On peut essayer de réfuter le théorème ci-dessus en utilisant la notion de *volume* dans l'espace $\mathbb{R}^3$. Le volume d'une boule B_r est donné par

$$\mathrm{vol}(B_r) = \frac{4}{3}\pi r^3,$$

et le volume d'une réunion disjointe est égal à la somme des volumes. De plus, le volume est invariant par isométrie. Par conséquent,

$$\begin{aligned}
\frac{4}{3}\pi r^3 &= \mathrm{vol}(B_r) \\
&= \mathrm{vol}\left(\cup_{i=1}^{n} A_i\right) = \cup_{i=1}^{n} \mathrm{vol}(A_i) \\
&= \cup_{i=1}^{n} \mathrm{vol}(A'_i) = \mathrm{vol}\left(\cup_{i=1}^{n} A'_i\right) \\
&= \mathrm{vol}(B_R) = \frac{4}{3}\pi R^3,
\end{aligned}$$

ce qui donne une contradiction si $r \neq R$.

Est-ce une contradiction ?

Indication : les éléments de la partition ne sont certainement pas tous mesurables.

La preuve du théorème de Banach-Tarski utilise l'axiome de choix de manière essentielle. Historiquement, ce paradoxe (qui n'en est pas un) a aidé au développement de la théorie des groupes *moyennables* (*amenable groups,* en anglais).

1.3 Espaces métriques

1.3.1 DÉFINITION
Soit E un ensemble non vide. Une distance (ou métrique) sur E, est la donnée d'une application $d : E \times E \to \mathbb{R}_+$ vérifiants les axiomes suivants :
(M1) $d(x,y) = 0 \Longleftrightarrow x = y$; (propriété de séparation)
(M2) pour tout $x, y \in E$, $d(x,y) = d(y,x)$; (propriété de symétrie)
(M3) pour tout $x, y, z \in E$, $d(x,z) \leq d(x,y) + d(y,z)$. (inégalité triangulaire)

1.3.2 DÉFINITION
On appelle espace métrique tout couple (E, d) où $E \neq \emptyset$ est un ensemble et d est une distance sur E.
(On dira seulement E est un espace métrique lorsque le contexte est clair.)

1.3.3 EXEMPLE. 1. $E = \mathbb{R}$, $d(x,y) = |x - y|$

2. $E = \mathbb{R}$. Soit $f : x \mapsto f(x)$ une fonction concave définie $\forall x \geq 0$, et t.q. $\{f(x) = 0\} \Leftrightarrow \{x = 0\}$. Alors $d(x,y) = f(|x-y|)$ est une distance. En effet, les propriétés D1 et D2 sont évidentes et D3 suit de la condition de concavité.
Une fonction est concave sur un intervalle I si $x_0, x_1, x_2, x_3 \in I$ et $x_0 < x_1 < x_2 < x_3$ alors, $\frac{f(x_1)-f(x_0)}{x_1-x_0} \geq \frac{f(x_3)-f(x_2)}{x_3-x_2}$. (Géométriquement, c'est une remarque sur la relation entre les pentes de deux droites qui lient les points de coordonnées $(x_0, f(x_0)$ et $(x_1, f(x_1))$ et $(x_2, f(x_2)$ et $(x_3, f(x_3))$. Faites un dessin !). Donc si on prend $x_0 = 0, x_1 = a, x_2 = b, x_3 = a+b$ on a $\frac{f(a)-f(0)}{a-0} \geq \frac{f(a+b)-f(b)}{a+b-b}$. Mais $f(0) = 0$ alors, si $0 < a < b$ on a $f(a) \geq f(a+b) - f(b)$ et donc $f(a+b) \leq f(a) + f(b)$.
On a beaucoup d'exemples de distances différentes sur $\mathbb{R}$. En particulier, $d(x,y) = \sqrt{|x-y|}$ ou $d(x,y) = \dfrac{|x-y|}{1+|x-y|}.$ Le dernier exemple définit une distance sur $\mathbb{R}$ qui, pour tout point, est inférieure à 1.

3. Métriques sur $E = \mathbb{R}^p$, soit $X = (x_1, \cdots, x_p) \in \mathbb{R}^p$ et $Y = (y_1, \cdots, y_p) \in \mathbb{R}^p$.
On a $d_2(X,Y) = \left(\sum_{i=1}^{p} |x_i - y_i|^2\right)^{1/2}$ (métrique euclidienne),
ou $d_1(X,Y) = \sum_{i=1}^{p} |x_i - y_i|$,
ou $d_\infty(X,Y) = \sup_{i=[1,\cdots,p]} |x_i - y_i|$

4. Soit E un ensemble quelconque. On définit la topologie discrète par la distance définie par : pour tout $x, y \in E$ on pose $d(x,y) = \begin{cases} 0 & \text{si } x = y, \\ 1 & \text{sinon.} \end{cases}$

1.3.1 Boules ouvertes, fermées. Sphères. Parties bornées

1.3.4 DÉFINITION
Soit a un point de E et $r > 0$ un nombre réel.

1. $\overline{B}(a,r) := \{x \in E \mid d(a,x) \leq r\}$ est appelée boule fermée de centre a et de rayon r.
2. Une boule ouverte de centre a et de rayon r est $B(a,r) := \{x \in E \mid d(a,x) < r\}$
3. Une sphère de centre a et de rayon r est $S(a,r) = \{x \in E \mid d(a,x) = r\}$

On obtient des boules de formes différentes pour des espaces métriques différents. Pour le voir je recommande vivement de dessiner des boules unité dans $\mathbb{R}^2$ pour les distances d_1, d_2 et d_∞.

1.3.5 DÉFINITION
Une partie bornée P de E est une partie de E pour laquelle on peut trouver une boule (ouverte ou fermée) qui contient tous les points de P.

1.3.2 Ouverts et Fermés

1.3.6 DÉFINITION
Une partie ouverte (ou un ouvert) de E est une partie U t.q. $\forall u \in U$, $\exists r > 0$ tel que $B(u,r) \subset U$ *ie* tout point de U est le centre d'une boule ouverte, de rayon non-nul, incluse dans U.

Une partie fermée (ou un fermé) de E est une partie telle que son complémentaire U dans E est un ouvert.

1.3.7 REMARQUE
E et $\emptyset$ sont à la fois ouverts et fermés.

1.3.8 PROPOSITION
Dans un espace métrique (E,d),

1. une boule ouverte est un ouvert
2. une boule fermée est un fermé.

Démonstration: (1) Soit $y \in B(a,r)$. Alors choisissons $\epsilon > 0$ t.q. $d(a,y) < r - \epsilon$ (un tel ϵ existe, car $d(a,y)$ est strictement plus petit que r). Pour tout $z \in B(y,\epsilon)$, montrons que $z \in B(a,r)$, cela veut dire qu'autour de chaque point y de $B(a,r)$ il existe une boule ouverte entièrement contenue dans $B(a,r)$.
Par inégalité triangulaire $d(a,z) \leq d(a,y) + d(z,y) \Rightarrow d(a,z) < r - \epsilon + \epsilon = r$. Donc $z \in B(a,r)$, i.e. chaque point de $B(y,\epsilon)$ appartient à $B(a,r)$ et $B(y,\epsilon) \subset B(a,r)$.

(2) Soit $\complement\overline{B}(a,r)$ le complémentaire de $\overline{B}(a,r)$. Il faut montrer que $\complement\overline{B}(a,r)$ est un ouvert. Soit $y \in \complement\overline{B}(a,r)$. Montrons qu'il existe une boule contenant y entièrement contenue dans $\complement\overline{B}(a,r)$.
Alorsque y est en dehors de $\overline{B}(a,r)$, $d(a,y) > r$. Soit $\epsilon = d(a,y) - r > 0$.
Pour tout $z \in B(y,\epsilon)$ montrons que $z \in \complement\overline{B}(a,r)$. En effet, par inégalité triangulaire $d(a,z) + d(z,y) \geq d(a,y) = r + \epsilon$. Donc $d(a,z) \geq r + \epsilon - d(z,y)$. Alorsque $z \in B(y,\epsilon)$ on a $\epsilon > d(z,y)$ donc $d(a,z) > r + \epsilon - d(z,y) > r + \epsilon - \epsilon = r \Rightarrow z \in \complement\overline{B}(a,r)$. Donc $\overline{B}(a,r)$ est un complément d'un ouvert, c'est donc un fermé. ■

1.3.10 Définition

Soit E un ensemble non-vide et $\mathcal{P}(E)$ l'ensemble de ses parties. On appelle <u>topologie</u> l'ensemble des ouverts $\mathcal{T} \subset \mathcal{P}(E)$ vérifiant les propriétés suivantes :

1. E et $\emptyset$ sont des éléments de $\mathcal{T}$
2. Toute intersection finie d'éléments de $\mathcal{T}$ appartient à $\mathcal{T}$
3. Toute réunion d'éléments de $\mathcal{T}$ appartient à $\mathcal{T}$.

1.3.11 Définition

<u>Position d'un point par rapport à une partie de (E,d).</u>

Soit $A \subset E$.

1. On dit que a est <u>intérieur</u> à A si on peut trouver un ouvert $U \in E$ t.q. $a \in U$ et $U \subset A$. L'<u>intérieur</u> de A, noté $\overset{o}{A}$, est le plus grand ouvert inclus dans A.
2. On dit que a est un <u>point frontière</u> de A si tout ouvert $U \subset E$ contenant a rencontre à la fois A et le complémentaire de A.
3. On dit que a est <u>adhérent</u> à A si tout ouvert $U \subset E$ contenant a rencontre A.
4. L'adhérence de A, notée $\overline{A}$, est le plus petit fermé qui contient A.

1.3.12 Définition

On dit qu'une partie V de (E,d) est un <u>voisinage</u> de $x \in E$ si V contient un ouvert contenant x.

1.3.13 Exercice Démontrer l'équivalence avec la définition suivante : On dit que $V \subset E$ est un voisinage de x ssi $\exists \varepsilon > 0$ tel que $B(x,\varepsilon) \subset V$.

1.3.3 Espaces complets

1.3.14 DÉFINITION
Un espace métrique est dit *complet* si ses suites de Cauchy sont convergentes.

1.3.15 THÉORÈME
Soit Y un sous-espace d'un espace métrique E. Si Y est complet, alors Y est fermé dans E.

Démonstration: Soit x un point adhérent à Y dans X :

$$x \in \bar{Y}.$$

Il existe une suite (y_n) de points de Y convergente vers x. Par conséquent, (y_n) est une suite de Cauchy dans E, mais comme la distance dans Y est induite par celle de E, la suite (y_n) est encore une suite de Cauchy par rapport dans Y. Elle possède une limite, y, dans Y, puisque Y est complet. Bien sûr, $y_n \to y$ dans E aussi. L'unicité de la limite implique que $x = y$, d'où on conclut que $x = y \in Y$, et Y est fermé. ∎

1.3.17 THÉORÈME
Chaque sous-espace fermé d'un espace métrique complet est complet.

Démonstration: Soient E un espace métrique complet et Y une partie de E munie de la fonction de distance induite. Soit (y_n) une suite de Cauchy dans Y. Cette suite-là est également une suite de Cauchy dans E, donc elle possède une limite, x :

$$y_n \to x \in E.$$

Ce point, x, est un point adhérent à Y dans E. Comme Y est fermé, on a $x \in Y$. Évidemment, $x_n \to x$ dans Y. ∎

1.3.19 THÉORÈME
Soit X un espace métrique. Il existe un plongement isométrique $i\colon X \hookrightarrow \hat{X}$ de X dans un espace métrique complet en tant que sous-espace dense. De plus, un tel plongement est unique à une isométrie près : si $j\colon X \hookrightarrow \tilde{X}$ est un autre plongement de X dans un espace métrique complet $\tilde{X}$ en tant que sous-espace dense, alors il existe une isométrie $\varkappa\colon \hat{X} \to \tilde{X}$ telle que $\varkappa \circ i = j$.

1.3.20 DÉFINITION

Le couple (X, ι) qui consiste d'un espace métrique complet $\widehat{X}$ et d'un plongement isométrique ι de X dans $\widehat{X}$ en tant qu'un sous-espace métrique dense s'appelle le *complété* de X. Souvent, par abus de langage, on parle de l'espace $\widehat{X}$ comme le complété de X.

Dans le théorème suivant, on peut penser du symbole $|X|$ tout simplement comme l'ensemble sous-jacent de l'espace métrique donné X.

1.3.21 THÉORÈME

Tout espace métrique X se plonge isométriquement dans l'espace métrique $\ell^\infty(|X|)$. où $\ell^\infty(|X|) = \{x = (x(y))_{y\in X} \,|\, x(y) \in \mathbb{R}, \quad \|x\|_\infty = \sup\limits_{y\in X} |x(y)| < \infty\}$ i.e. l'ensemble des applications bornées de X à valeurs dans $\mathbb{K}$.

Démonstration: Nous allons définir un plongement isométrique

$$i \colon X \hookrightarrow \ell^\infty(|X|),$$

appelé le *plongement de Kuratowski.* (Pour tout x, l'image $i(x)$ sera une fonction $X \to \mathbb{R}$, dont la valeur en un point $y \in X$ sera notée $i(x)_y$:

$$i(x) \colon y \mapsto i(x)_y \in \mathbb{R}.$$

Fixons un point $x_0 \in X$ quelconque, et pour tout $x \in X$ définissons l'image $i(x)$ dans $\ell^\infty(|X|)$ comme suit :

$$i(x)_y = d(x,y) - d(x_0,y).$$

D'abord, étant donné un point $x \in X$, la fonction

$$X \ni y \mapsto i(x)_y = d(x,y) - d(x_0,y)$$

est majorée grâce à l'inégalité triangulaire :

$$|d(x,y) - d(x_0,y)| \leq d(x,x_0),$$

parce que la borne ne depend pas de y, seulement de x.

Étant donné deux points $x, z \in X$ quelconques, on a pour tout $y \in X$:

$$\begin{aligned} |i(x)_y - i(z)_y| &= |d(x,y) - d(x_0,y) - d(z,y) + d(x_0,y)| \\ &= |d(x,y) - d(z,y)| \\ &\leq d(x,y). \end{aligned}$$

D'autre part, en posant $y = x$, on obtient

$$|i(x)_x - i(z)_x| = |d(x,x) - d(z,x)| = d(x,z).$$

On en déduit :

$$\begin{aligned} d_\infty(i(x), i(z)) &= \sup_{y \in X} |i(x)_y - i(z)_y| \\ &= d(x,z). \end{aligned}$$

■

1.3.23 THÉORÈME

L'espace métrique $\ell^\infty(\Gamma) = \{x = (x(\gamma))_{\gamma \in \Gamma} \,|\, x(\gamma) \in \mathbb{K}, \quad \|x\|_\infty = \sup_{\gamma \in \Gamma} |x(\gamma)| < \infty\}$ est complet, quelque soit l'ensemble Γ.

Démonstration: (1) Soit (x_n) une suite de Cauchy dans $\ell^\infty(\Gamma)$. Cela veut dire que pour tout $\varepsilon > 0$ il existe N tel que, quel que soit $n \geq N$, on a

$$x_n \in B_\varepsilon(x_N). \tag{1.3.1}$$

La condtion (1.3.1) est équivalente à la condition suivante :

$$\sup_{\gamma \in \Gamma} |x_n(\gamma) - x_N(\gamma)| < \varepsilon, \tag{1.3.2}$$

qui implique à son tour que pour tout $\gamma \in \Gamma$,

$$|x_n(\gamma) - x_N(\gamma)| < \varepsilon$$

une fois $n \geq N$. On en conclut : pour chaque $\gamma \in \Gamma$,

$$\forall \varepsilon > 0 \;\; \exists N, \;\; \forall n \geq N, \;\; |x_n(\gamma) - x_N(\gamma)| < \varepsilon.$$

Par conséquent, la suite $(x_n(\gamma)$ est une suite de Cauchy dans $\mathbb{R}$, donc elle possède une limite. Notons-là $x(\gamma)$:

$$x_n(\gamma) \to x(\gamma) \text{ quand } n \to \infty.$$

Nous avons construit une fonction

$$x\colon \Gamma \to \mathbb{R},$$

telle que

$$\forall \gamma \in \Gamma, \;\; x_n(\gamma) \to x(\gamma). \tag{1.3.3}$$

(2) Montrons que cette fonction est bornée. Choisissons une valeur de $\varepsilon_0 > 0$ quelconque (par exemple, $\varepsilon_0 = 1$). Il existe N tel que pour tout $n \geq N$ et tout $\gamma \in \Gamma$,

$$|x_n(\gamma) - x_N(\gamma)| < 1 \tag{1.3.4}$$

et par conséquent,

$$|x(\gamma) - x_N(\gamma)| \leq 1. \tag{1.3.5}$$

Cela implique

$$\forall \gamma \in \Gamma, \;\; |x(\gamma)| \leq |x_N(\gamma)| + 1,$$

et par conséquent

$$\sup_{\gamma\Gamma} |x(\gamma)| \leq \sup_{\gamma\in\Gamma} |x_N(\gamma)| + 1 < \infty.$$

On en conclut : $x \in \ell^\infty(\Gamma)$.

(3) Finalement, on va montrer que $x_n \to x$ par rapport à la distance uniforme d_∞ (la condition (1.3.3) seulement implique que (x_n) converge vers x au sens de la convergence simple, voir la remarque 1.3.25). Soit $\varepsilon > 0$. Fixons N tel que quels que soient $n, m \geq N$, on a

$$d_\infty(x_m, x_n) < \varepsilon.$$

On en déduit que pour tout $\gamma \in \Gamma$,

$$|x_m(\gamma) - x_n(\gamma)| < \varepsilon,$$

donc, en prenant la limite $m \to \infty$ dans $\mathbb{R}$,

$$|x(\gamma) - x_n(\gamma)| \leq \varepsilon$$

et, en prenant la borne supérieure sur tous $\gamma \in \Gamma$, on obtient

$$d_\infty(x, x_n) = \sup_{\gamma\in\Gamma} |x_m(\gamma) - x_n(\gamma)| \leq \varepsilon.$$

On a montré que pour tout $\varepsilon > 0$ il existe N tel que quel que soit $n \geq N$, on a

$$d_\infty(x, x_n) \leq \varepsilon,$$

ce qui veut dire

$$x_n \to x \text{ lorsque } n \to \infty,$$

par rapport à la distance d_∞ du type ℓ^∞ (la distance uniforme) sur l'espace ℓ^∞. ■

1.3.25 REMARQUE
Soit X un ensemble quelconque, Y un espace métrique. Le symbole Y^X note la famille des toutes les applications de X dans Y. On dit qu'une suite (f_n) de fonctions de Y^X converge vers une fonction $f \in Y^X$ *simplement* si

$$\forall x \in X, \ \ f_n(x) \to f(x) \text{ dans l'espace } Y.$$

En d'outres termes ,

$$\forall \varepsilon > 0, \ \ \forall x \in X, \ \ \exists N = N(\varepsilon, x), \ \ \forall n \geq N, \ \ d(f_n(x), f(x)) < \varepsilon.$$

On dit que x_n converge vers x *uniformément* si un tel N peut être choisi de façon indépendante de x, c.à.d., uniformément sur X :

$$\forall \varepsilon > 0, \ \ \exists N = N(\varepsilon), \ \ \forall n \geq N, \ \ \forall x \in X, \ \ d(f_n(x), f(x)) < \varepsilon.$$

Évidemment, la convergence uniforme implique la convergence simple. Au même temps, la réciproque n'est pas toujours vraie.

1.3.26 EXEMPLE. Voilà un tel exemple, probablement le plus simple qu'on puisse s'imaginer. Soit $X = \mathbb{N} - \{0\}$, $Y = \{0, 1\}$ muni de la distance usuelle ($d(0, 1) = 1$). Les éléments de $Y^X = \{0, 1\}^{\mathbb{N} - \{0\}}$ sont des suites de zéro et de un. Définissons donc

$$f = (0, 0, 0, \ldots, 0, \ldots,)$$

et pour tout $n \in \mathbb{N} - \{0\}$,

$$f_n = (\underbrace{0, 0, \ldots, 0}_{n \text{ fois}}, 1, 1, 1, \ldots, 1, \ldots)$$

Alors la suite $(f_n)_{n=1}^{\infty}$ converge vers f simplement, mais pas uniformément, parce que, quel que soit n, on a toujours $d_\infty(f, f_n) = 1$.

1.4 Espaces vectoriels normés

1.4.1 Définition et exemples

Une norme est un concept général de *longueur de vecteurs*. Une fois que nous avons une norme, nous pouvons "géométriser l'analyse" dans un certain sens, parce que nous aurions une métrique sur nos espaces vectoriels. Par exemple, cela nous permettrait d'étudier les fonctions grâce à la géométrie des espaces fonctionnels.

Une norme est l'association d'un nombre positif $\|x\|$ à chaque vecteur x dans un espace vectoriel E. Afin de donner un sens à la longueur, ce nombre doit satisfaire certains axiomes naturels :

1.4.1 DÉFINITION (ESPACE VECTORIEL NORMÉ)

Soit E un espace vectoriel. Un *norme* sur E est une fonction $\|\cdot\| : E \to \mathbb{R}$, qui satisfait les axiomes suivants :

(i) $\|x\| \geq 0$ pour tout $x \in E$, $\|x\| = 0$ si et seulement si $x = 0$;

(ii) $\|ax\| = |a|\|x\|$ pour tout $x \in E$, $a \in \mathbb{K}$;

(iii) $\|x+y\| \leq \|x\| + \|y\|$ pour tout $x, y \in E$.

L'espace vectoriel E muni de la norme $\|\cdot\|$ est appelé *espace vectoriel normé*, et noté $(E, \|\cdot\|)$ ou seulement E .

L'axiome (iii) est appelée *inégalité triangulaire* pour la raison suivante. étant donné un triangle arbitraire dans E de sommets $x, y, z \in E$, ses longueurs de ses côtés, vérifient l'inégalité

$$\|x-z\| \leq \|x-y\| + \|y-z\|, \tag{1.4.1}$$

qui découle de l'axiome norme (iii).

L'espace vectoriel normé est naturellement un *espace métrique*, avec la métrique induite définie par

$$d(x,y) := \|x-y\|.$$

Les axiomes de la norme, et en particulier l'inégalité triangulaire (1.4.1), montrer que c'est en effet une métrique (à vérifier !)

1.4.2 PROPOSITION

Soit E un e.v.n. L'application $d : E \times E \to \mathbb{R}_+$ qui au couple (x,y) associe $d(x,y) :=\| x-y \|$ est une distance sur E.

On l'appelle distance induite sur E par la norme. Elle possède les propriétés suivantes :

- $\forall x \in E,\ d(0,x) = \| x \|$
- $\forall \lambda \in \mathbb{R},\ \forall (x,y) \in E^2,\ d(\lambda x, \lambda y) = |\lambda| d(x,y)$
- $\forall (x,y,z) \in E \times E \times E,\ d(x+z, y+z) = d(x,y)$, (invariance par translation)

1.4.3 EXEMPLE. L'ensemble $\mathbb{C}$ des nombres complexes $z = \mathfrak{Re} z + i\mathfrak{Im} z$, $\mathfrak{Re} z, \mathfrak{Im} z \in \mathbb{R}$, est un espace vectoriel complexe de dimension 1. Muni de la norme

$$\| z \| = |z| = \sqrt{\mathfrak{Re} z^2 + \mathfrak{Im} z^2}$$

il est un espace normé complexe.

1.4.4 EXEMPLE. La norme euclidienne standard sur l'espace vectoriel $\mathbb{R}$, $n \in \mathbb{N}$, est donnée par

$$\| x \|_2 = \sqrt{\sum_{i=1}^{n} |x_n|^2}.$$

On a déjà montré que $\| \cdot \|_2$ est une norme : la propriété (3) d'une norme est l'inégalité de Minkowski (Theorem 1.35).

L'espace vectoriel $\mathbb{R}$ muni de la norme $\| \cdot \|_2$ ci-dessus est noté $\ell^2(n)$, ou bien $\ell^2_{\mathbb{R}}(n)$ pour souligner que c'est la version réelle.

1.4.5 EXEMPLE. L'espace vectoriel $\ell^\infty(\Gamma)$, muni de la norme

$$\| x \|_\infty = \sup_{\gamma \in \Gamma} |x(\gamma)|,$$

est un espace normé.

1.4.6 REMARQUE

Soit $E = (E, \| \cdot \|)$ un espace normé. Il est facile à vérifier que la formule

$$d(x,y) = \| x - y \|$$

définit une fonction de distance (métrique) sur E, appelé la fonction de distance induite par la norme, ou associée à la norme.

1.4.7 REMARQUE

La norme $\| x \|$ est la distance entre x et zéro :

$$\| x \| = d_E(x, 0).$$

Pour cette raison, la norme $N \colon E \to \mathbb{R}$ est une fonction continue sur l'espace normé E.

1.4.8 REMARQUE

La norme sur un espace normé non-trivial $E \neq \{0\}$ n'est jamais bornée. Si $x \in E$, $x \neq 0$, alors $\| x \| \neq 0$, et pour $\lambda \to \infty$ on a $\| \lambda x \| = |\lambda| \| x \| \to \infty$ aussi.

1.4.9 EXEMPLE. Il est facile à voir que la fonction de distance d_E induite par une norme de l'espace normé E est *invariant par translations* : quels que soient $x, y, z \in E$, on a

$$d_E(x + z, y + z) = d_E(x, y).$$

Au même temps, il est facile de construire une métrique sur un espace vectoriel qui est invariante par translations mais n'est pas induite par une norme. Telle est la distance

$$d(x,y) = \min\{1, |x - y|\}$$

sur $\mathbb{R}$. Elle est bornée, et on peut utiliser la remarque 1.4.8.

La distance (4) de l'exemple 1.3.3 n'est pas induite par aucune norme.

1.4.10 Remarque

Un espace normé E est compact si est seulement si $E = \{0\}$. Si $E \neq \{0\}$, il existe $x \in E \setminus \{0\}$, et la suite des points $x_n = nx$ ne contient aucune suite extraite de Cauchy, parce que pour tout m, n, $m \neq n$, on a

$$\| x_m - x_n \| = \| mx - nx \| = \| (m-n)x \| = |m-n| \cdot \| x \| \geq \| x \| > 0.$$

Exemple de normes sur $\mathbb{R}$. Soit $x \in \mathbb{R},\ x = (x_1, \cdots, x_n),\ x_i \in \mathbb{R}, \forall i \in [1, \cdots, n]$. Alors

$$\begin{aligned} \| x \|_1 &= \textstyle\sum_1^p |x_i| \\ \| x \|_2 &= \left(\textstyle\sum_1^p |x_i|^2\right)^{1/2} \text{ (norme euclidienne)} \\ \| x \|_p &= \left(\textstyle\sum_1^p |x_i|^p\right)^{1/p} \\ \| x \|_\infty &= \max_{1\leq i \leq p} |x_i| \end{aligned}$$

sont des normes sur $\mathbb{R}$.

$\mathbb{R}$ munit de la norme $\|.\|_p$ est noté parfois ℓ_p^n.

1.4.11 Exemple (d'espaces vectoriels normés : ℓ^∞, c, c_0, ℓ^1, $C(K)$, L^1, L^∞).

Beaucoup d'espaces vectoriels présentés dans la section 1.2.2 et dans l'exemple 1.2.12 sont en fait des espaces normés. Vérifiez les axiomes de la norme.

1. L'espace des suites bornées ℓ^∞ est un espace vectoriel normé, avec la norme définie par
$$\|x\|_\infty := \sup_i \{|x_i|\}. \tag{1.4.2}$$
2. Les espaces c et c_0 sont des espaces normés, avec la même norme du sup, comme dans (1.4.2).
3. L'espace des suites sommables ℓ^1 est un espace vectoriel normé, avec la norme définie par
$$\|x\|_1 := \sum_{i=1}^\infty |x_i|.$$
4. L'espace $C(K)$ des fonctions continues sur un espace topologique compact K est un espace vectoriel normé avec la norme [4]
$$\|f\|_\infty := \max_k |f(x)|.$$

1.4.12 Définition

Normes équivalentes. Deux normes $\|\cdot\|$ et $\|\cdot\|'$ sur E sont équivalentes s'il existe deux constantes $\lambda > 0, \mu > 0$ telles que $\forall x \in E, \lambda\|x\| \leq \|x\|' \leq \mu\|x\|$. On note $\|\cdot\| \sim \|\cdot\|'$.

4. Le maximum est atteint car K est compact.

1.4.13 PROPOSITION
Cette définition induit une relation d'équivalence.

Démonstration: – **reflexivité** : $\|\cdot\| \sim \|\cdot\|$
- **symétrie** : si $\lambda\|x\| \leq \|x\|' \leq \mu\|x\|$ alors $\frac{1}{\mu}\|x\|' \leq \|x\| \leq \frac{1}{\lambda}\|x\|'$.
- **transitivité** : $\lambda\|x\| \leq \|x\|' \leq \mu\|x\|$ et $\beta\|x\|' \leq \|x\|'' \leq \gamma\|x\|'$ implique $\beta\lambda\|x\| \leq \|x\|'' \leq \gamma\mu\|x\|$. ■

1.4.15 EXEMPLE. Les normes $\| x\|_2 = \left(\sum_1^p |x_i|^2\right)^{1/2}$ et $\| x \|_\infty = \max_{1\leq i\leq p}|x_i|$ sont équivalentes. En effet, on a $\|x\|_2 \leq (p \cdot \|x\|_\infty^2)^{1/2} = \sqrt{p}\|x\|_\infty$. Soit $k \in \{1, \cdots, p\}$ tel que $x_k = \max\{x_1, \cdots, x_p\} = \|x\|_\infty$, alors $\|X\|_\infty = (x_k^2)^{1/2} \leq \left(\sum_1^p |x_i|^2\right)^{1/2} = \|x\|_2$. Donc $\frac{1}{\sqrt{p}}\|x\|_2 \leq \|x\|_\infty \leq \|x\|_2$.

1.4.16 Exercice 1. Montrer que toutes les normes $\|\cdot\|_n, n \in [1, +\infty]$ sont équivalentes.
2. Si $\|\cdot\| \sim \|\cdot\|'$ montrer qu'il existe une constante $\lambda > 0$ t.q. $\lambda\|x\| \leq \|x\|' \leq \frac{1}{\lambda}\|x\|$ et $\lambda\|x\|' \leq \|x\| \leq \frac{1}{\lambda}\|x\|'$.

1.4.17 THÉORÈME
Deux normes équivalentes induisent la même topologie.

Ie si les normes sont équivalentes on trouve que deux ensembles

$$\mathcal{T} = \{U \in \mathcal{P}(E), U \text{ ouvert pour la norme } \|\cdot\|\}$$

et $\mathcal{T}' = \{U \in \mathcal{P}(E),\ U$ ouvert pour la norme $\|\cdot\|'\}$, sont égaux : $\mathcal{T} = \mathcal{T}'$.

Démonstration: Soit U un élément de $\mathcal{T}$, il faut montrer que c'est aussi un élément de $\mathcal{T}'$.
Cela se traduit :

Soit U un ouvert pour la norme $\|\cdot\| \Leftrightarrow \forall x \in U,\ \exists \varepsilon > 0$ tel que $B(x, \varepsilon) \subset U$. On va m.q. U est un ouvert pour la norme $\|\cdot\|'$. Pour tout $x \in U$ il faut montrer qu'il existe $\varepsilon' > 0$ tel que $B'(x, \varepsilon')$, une boule pour la norme $\|\cdot\|'$ est un sous-ensemble de U. Pour cela on va trouver ε' tel que tout point y de $B'(x, \varepsilon')$ appartienne aussi à $B(x, \varepsilon)$ et donc à U. Par équivalence des normes $\exists\lambda > 0$ tel que $\forall Z \in E\|z\| \leq \lambda\|z\|'$. Soit $y \in B'(x, \frac{\varepsilon}{\lambda})$ on a $\|x - y\| \leq \lambda\|x - y\|' < \lambda\dfrac{\varepsilon}{\lambda} = \varepsilon$ donc $B'(x, \dfrac{\varepsilon}{\lambda}) \subset U$. Donc si U est un ouvert pour $\|\cdot\|$, alors pour tout $x \in U$, il existe $\varepsilon' = \dfrac{\varepsilon}{\lambda} > 0$ tel que $B'(x, \varepsilon') \subset U$. Donc U est un élément de $\mathcal{T}'$.

De la même manière on montre que si U est un élément de $\mathcal{T}'$, c'est aussi un élément de $\mathcal{T}$.

1.4.19 Exercice (Continuité des normes) Montrer que l'application norme $x \mapsto \|x\|$ est une fonction continue sur l'espace vectoriel normé. Plus précisément, montrer que si $\|x_n - x\| \to 0$, alors $\|x_n\| \to \|x\|$.

1.4.2 La convexité des normes et des boules

La géométrie d'un espace vectoriel normé peut être très différente de celle de la géométrie euclidienne usuelle. Les boules n'ont pas besoin d'êtres rondes. Par exemple, la boule de ℓ^∞ ressemble à un cube (pourquoi ?). Néanmoins, une propriété importante reste valable : les boules sont toujours *d'ensembles convexes*, et la norme est une fonction convexe. Les considérations de convexité sont très utiles lorsque l'on travaille dans des espaces normés.

Rappelons tout d'abord quelques notions issues de la géométrie des espaces métriques.

1.4.20 Définition (Boules, sphères d'espaces vectoriels normés)
Soit X un espace vectoriel normé. Une *boule* ouverte centrée en un point $x_0 \in X$ et de rayon $r > 0$ est la partie définie par

$$B(x_0, r) := \{x \in X : \ \|x - x_0\| < r\}.$$

Une *boule* fermée centrée en un point $x_0 \in X$ et de rayon $r > 0$ est la partie définie par

$$\bar{B}(x_0, r) := \{x \in X : \ \|x - x_0\| < r\}.$$

La boule unité fermée de X est définie par

$$B_X := \bar{B}(0, 1) = \{x \in X : \ \|x\| \leq 1\}.$$

La sphère unité de X est la frontière de la boule unité, c'est-à-dire

$$S_X := \{x \in X : \ \|x\| = 1\}.$$

1.4.21 Définition (Fonctions convexes)
Soit E un espace vectoriel normé.

Une fonction $f : E \to \mathbb{R}$ est *une fonction convexe* si pour tout $x, y \in E$, $\lambda \in [0, 1]$ on a

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Un ensemble $K \subseteq E$ est *convexe* si pour tout $x, y \in K$, $\lambda \in [0, 1]$ on a

$$\lambda x + (1 - \lambda)y \in K.$$

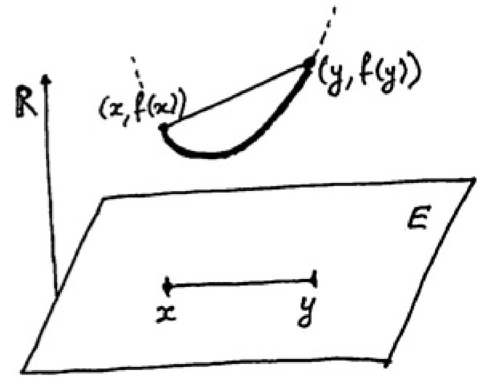


FIGURE 1.1 – Fonction convexe f sur un espace vectoriel E

Un sens géométrique de la convexité est la suivante. Une fonction f est convexe sur E si son graphe sur un intervalle $[x, y] \subset E$ se situe en dessous du segment joignant les points $(x, f(x))$ et $(y, f(y))$ (voir l'image). Un ensemble K est convexe si, pour tous $x, y \in K$, il contient le segment $[x, y]$.

1.4.22 PROPOSITION (LES AXIOMES DE LA NORME ENTRÎNENT LA CONVEXITÉ)
Soit X un espace vectoriel normé. Alors :

1. La fonction $x \mapsto \|x\|$ est convexe sur X.
2. La boule unité B_X est un fermé, symétrique par rapport à l'origine [5], et un sous-ensemble convexe de X.

Démonstration: 1. La convexité de la norme découle des axiomes. En effet, pour $x, y \in E$, $\lambda \in [0, 1]$, nous avons

$$\|\lambda x + (1-\lambda)y)\| \leq \|\lambda x\| + \|(1-\lambda)y)\| = \lambda\|x\| + (1-\lambda)\|y\|.$$

2. l'adhérence de B_X découle de la continuité de la norme (voir l'exercice **??**). La symétrie par rapport à l'origine, est une conséquence de l'axiome (ii) avec λ = -1. Enfin, pour prouver la convexité de B_X, nous choisissons arbitrairement $x, y \in B_X$, $\lambda \in [0, 1]$ et utiliser l'inégalité ci-dessus pour obtenir

$$\|\lambda x + (1-\lambda)y\| \leq \lambda\|x\| + (1-\lambda)\|y\| \leq \lambda + (1-\lambda) = 1.$$

Il en résulte que $\lambda x + (1-\lambda)y \in B_X$ tel que requis. ∎

La réciproque à la proposition 1.4.24 est aussi vraie :

5. symétrique par rapport à l'origine signifie que si $x \in B_X$ alors $-x \in B_X$

1.4.24 PROPOSITION (CONVEXITÉ IMPLIQUE L'INÉGALITÉ TRIANGULAIRE)
Soit $x \mapsto \|x\|$ une fonction à valeurs réelles définie sur un espace vectoriel E. Supposons que cette fonction satisfait les axiomes (i) et (ii) de la norme. Alors :

1. Si la fonction $x \to \|x\|$ est convexe, l'inégalité triangulaire est satisfaite, et $\|\cdot\|$ est une norme sur E.
2. Si l'ensemble de "sous-niveau" $\{x \in X : \ \|x\| \leq 1\}$ est convexe, alors $\|\cdot\|$ est une norme sur E.

Démonstration: 1. Convexité assure que pour chaque $x, y \in E$, $\lambda \in [0,1]$, nous avons

$$\|\lambda x + (1-\lambda)y)\| \leq \lambda\|x\| + (1-\lambda)\|y\|.$$

Inégalité triangulaire en découle en prenant $\lambda = 1/2$.

2. Cette affirmation est moins triviale, et ne peut être obtenue directement à partir de la première. En effet, s'il est vrai que les ensembles de sous-niveaux d'une fonctions convexes sont d'ensembles convexes, la réciproque n'est pas toujours vraie (construire un contre-exemple !)

On veut montrer que, pour $u, v \in E$ on a :

$$\text{si}\|u\| \leq 1, \ \|v\| \leq 1, \ \lambda \in [0,1], \quad \text{alors} \quad \|\lambda u + (1-\lambda)v\| \leq 1. \tag{1.4.3}$$

Soit $x, y \in E$, nous voulons montrer que $\|x+y\| \leq \|x\| + \|y\|$. Ceci est équivalent à

$$\left\|\frac{x}{\|x\|+\|y\|} + \frac{y}{\|x\|+\|y\|}\right\| \leq 1.$$

Nous obtenons cette inégalité de (1.4.3)

$$u = \frac{x}{\|x\|}, \ v = \frac{y}{\|y\|}, \ \lambda = \frac{\|x\|}{\|x\|+\|y\|}.$$

Ceci termine la preuve. ■

1.4.3 Espaces L^p. Inégalité de Minkowski

———————————————————-

Nous avons déjà rencontré les espaces L^1 et ℓ^1. Ce sont des cas particuliers d'une grande famille d'espaces L^p et ℓ^p que nous allons étudier maintenant.

Considérons un espace mesuré (Ω, Σ, μ) et un exposant $p \in [1, \infty)$. Nous définissons *l'espace des fonctions p-intégrable* $L^p = L^p(\Omega, \Sigma, \mu)$ comme l'ensemble des fonctions mesurables $f : \Omega \to \mathbb{R}$ telles que

$$\int_\Omega |f(x)|^p \, d\mu < \infty.$$

1.4.26 PROPOSITION
$L^p(\Omega, \Sigma, \mu)$ est un espace vectoriel pour $p \in [1, \infty)$.

Démonstration: Le seul point non-trivial est que L^p est stable par l'addition, soient $f, g \in L^p$ implique $f+g \in L^p$. Nous obtiendrons ce par la convexité de la fonction $z \mapsto |z|^p$ sur $\mathbb{R}$ pour $p \geq 1$. La convexité implique l'inégalité

$$\Big|\frac{|f(t)+g(t)|}{2}\Big|^p \leq \frac{|f(t)|^p + |g(t)|^p}{2}, \quad t \in \Omega.$$

Alors en intègre deux part et d'autre de cette inégalité. ■

On munit L^p d'une structure d'espace vectoriel normé par la norme

$$\|f\|_p := \Big(\int_\Omega |f(x)|^p \, d\mu\Big)^{1/p} \quad \text{pour} f \in L^p(\Omega, \Sigma, \mu).$$

1.4.28 PROPOSITION
$L^p(\Omega, \Sigma, \mu)$ est un espace vectoriel normé, muni de la norme $\|f\|_p$ pour $p \in [1, \infty)$.

Démonstration: Norm axiomes (i) et (ii) sont simples. L'axiome (iii), inégalité triangulaire, suivront dès la Proposition 1.4.22. pour cela, il suffit de vérifier que l'ensemble sous-niveau

$$B_p := \{f \in L_p : \ \|f\|_p \leq 1\}$$

est un ensemble convexe. Pour le prouver, nous fixons $f, g \in B_p$ et $\lambda \in [0, 1]$. Comme la fonction $z \mapsto |z|^p$ est convexe sur $\mathbb{R}$ pour $p \geq 1$, nous avons l'inégalité

$$\Big|\lambda f(t) + (1-\lambda) g(t)\Big|^p \leq \lambda |f(t)|^p + (1-\lambda)|g(t)|^p.$$

L'intégration des deux côtés de cette inégalité donne

$$\int_\Omega \Big|\lambda f(t) + (1-\lambda) g(t)\Big|^p d\mu \leq \lambda \int_\Omega |f(t)|^p \, d\mu + (1-\lambda) \int_\Omega |g(t)|^p \, d\mu \leq \lambda + (1-\lambda) = 1.$$

Nous avons montré que $\|\lambda f + (1-\lambda) g\|_p \leq 1$, donc $\lambda f + (1-\lambda) g \in B_p$. Par conséquent, le sous-niveau B_p est convexe. La preuve est terminée par la proposition 1.4.24. ■

écrire sur l'inégalité triangulaire $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ sous forme analytique, permet d'obtenir *l'inégalité Minkowski* :

1.4.30 THÉORÈME (INÉGALITÉ DE MINKOWSKI DANS L^p)
Soit $p \in [1, \infty)$. Alors, pour tous $f, g \in L^p(\Omega, \Sigma, \mu)$ on a

$$\Big(\int_\Omega |f(t)+g(t)|^p \, d\mu\Big)^{1/p} \leq \Big(\int_\Omega |f(t)|^p \, d\mu\Big)^{1/p} + \Big(\int_\Omega |g(t)|^p \, d\mu\Big)^{1/p}.$$

1.4.4 Espaces ℓ^p et ℓ_n^p.

Un cas particulier important des espaces $L^p(\Omega, \Sigma, \mu)$ est l'espace ℓ^p obtenue en choisissant $\Omega = \mathbb{N}$ et μ est la mesure de comptage sur $\mathbb{N}$. Pour $p \in [1, \infty)$, *l'espace des suites p-sommables* ℓ^p est l'espace des suites $x = (x_i)_{i=1}^{\infty}$ pour lesquels

$$\sum_{i=1}^{\infty} |x_i|^p < \infty.$$

On munit ℓ^p d'une structure d'espace vectoriel normé avec

$$\|x\|_p := \Big(\sum_{i=1}^{\infty} |x_i|^p\Big)^{1/p}.$$

L'écriture de l'inégalité de Minkowski pour cet espace mesuré, on obtient :

1.4.31 Théorème (inégalité de Minkowski dans ℓ^p)
Soit $p \in [1, \infty)$. Alors, pour toutes suites $x, y \in \ell^p$ on a

$$\Big(\sum_{i=1}^{\infty} |x_i + y_i|^p\Big)^{1/p} \leq \Big(\sum_{i=1}^{\infty} |x_i|^p\Big)^{1/p} + \Big(\sum_{i=1}^{\infty} |y_i|^p\Big)^{1/p}.$$

Une famille remarquable d'espaces $L^p(\Omega, \Sigma, \mu)$ de *dimension finie* est obtenue en considérant un ensemble fini Ω, disons que $\Omega = \{1, \ldots, n\}$ et μ la mesure de comptage sur Ω. L'espace qui en résulte est noté ℓ_n^p. Un élément de ℓ_n^p peut être identifié avec un vecteur de $\mathbb{R}^n$. Ainsi $\ell_n^p = (\mathbb{R}^n, \|\cdot\|_p)$ muni de la norme

$$\|x\|_p := \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p}.$$

Lorsque $p = 2$, c'est l'espace euclidien usuel $\mathbb{R}^n$. Cependant, pour $p \neq 2$, la géométrie de ℓ_n^p est assez différente de la géométrie euclidienne. En effet, en dimension 2, la boule unité de ℓ_2^1 est un "diamant" de sommets $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, -1)$; alors que la boule unité de ℓ_2^∞ est le carré de sommets $(1, 1)$, $(1, -1)$, $(-1, 1)$, $(-1, -1)$.

1.4.32 Exercice (ℓ^∞ est la limite de ℓ^p) Cet exercice explique la présence de l'indice ∞ dans les notations ℓ^∞ et L^∞.

1. Montrer que pour $x \in \ell^{p_0}$ avec $p_0 \geq 1$, alors

$$\|x\|_p \to \|x\|_\infty \quad \text{comme} p_0 \leq p \to \infty.$$

2. Considérer l'espace $L^\infty = L^\infty(\Omega, \Sigma, \mu)$ avec une mesure finie i.e. $\mu(\Omega) < \infty$. Montrer que si $f \in L^\infty$, alors

$$\|f\|_p \to \|f\|_\infty \quad comme p \to \infty.$$

1.4.5 Les sous-espaces d'espaces normés

1.4.33 DÉFINITION (SOUS-ESPACE)
Soit X un espace vectoriel normé. Un sous-espace vectoriel Y de X est un espace vectoriel normé équipé muni de la norme induite par la restriction de celle de X.

Ce concept est familier en topologie, où un sous-ensemble est un sous-espace topologique muni de la induite par la topologie.

1.4.34 EXEMPLE. 1. L'espace des polynômes $\mathbb{R}[x]$ est un sous-espace dense de $C[0,1]$. Il s'agit du théorème d'approximation de Weierstrass.

2. L'ensemble des fonctions continues $C[0,1]$ forme un sous-espace dense de $L^1[0,1]$. (Bien sûr, les deux espaces sont munis de la norme L^1) Cela découle d'un théorème en théorie de la mesure qu'une fonction intégrable peut être approchée par une fonction continue (pourquoi ?)

1.4.35 EXEMPLE. Montrer que l'ensemble de suites convergentes c et l'ensemble de suites qui convergent vers 0, c_0 sont des fermés de ℓ^∞.

Pour tout $p \in [1, \infty[$, montrer que l'ensemble des suites p-sommables ℓ^p est un sous-espace fermé de ℓ^∞, mais c'est un sous-espace dense de c_0.

1.4.6 Espaces quotients d'espaces normés

Dans la section 1.2.5, nous avons défini les espaces quotients d'espaces vectoriels. Si l'espace ambiant est un espace vectoriel normé, on peut également définire une norme (canonique) sur l'espace quotient de la manière suivante.

1.4.36 DÉFINITION (ESPACE QUOTIENT D'UN ESPACE VECTORIEL NORMÉ)

1. Soit X un espace vectoriel normé et Y un sous-espace **fermé** de X. Nous définissons une norme sur X/Y, comme suit. Pour chaque classe $[x] = x + Y$, nous posons

$$\|[x]\| := \inf_{y \in Y} \|x + y\|.$$

Il est facile de comprendre géométriquement la norme dans l'espace quotient comme *la distance* de la classe $[x]$ à l'origine . En effet, soit $\operatorname{dist}(b, A)$ la distance d'un point b de X à un ensemble A :

$$\operatorname{dist}(b, A) = \inf_{a \in A} \|b - a\|.$$

Il est clair que

$$\|[x]\| = \operatorname{dist}(0, [x]). \tag{1.4.4}$$

1.4.37 **Proposition**
L'application $x \mapsto \|[x]\|$ ci-dessus définie une norme sur X/Y.

Démonstration: Premièrement, nous observons que le nombre $\|[x]\|$ est bien défini, c'est à dire qu'il ne dépend pas du choix d'un représentant x dans la classe d'équivalence $[x]$. Cela découle clairement de la définition géométrique (1.4.4).

Alors, il faut vérifier les trois axiomes de la norme.

(I) Supposons que $\|[x]\| = 0$. Alors, à partir de la définition géométrique (1.4.4) nous constatons que 0 est un point limite de $[x]$. Comme Y est fermé, est que $[x] = x + Y$. $0 \in [x]$. Ainsi $[x] = [0]$, vérifie l'axiome (i) de la norme .

(Ii) Soit $x \in X$ et $\lambda \in \mathbb{R}$. Alors

$$\|[\lambda x]\| = \inf_{y \in Y} \|\lambda x + y\| = \inf_{y \in Y} \|\lambda x + \lambda y\| = \lambda \cdot \inf_{y \in Y} \|x + y\| = \lambda \|[x]\|.$$

Ceci vérifie l'axiome (ii) de la norme.

(iii) Soit $x_1, x_2 \in X$, nous voulons montrer que $\|[x_1 + x_2]\| \leq \|[x_1]\| + \|[x_2]\|$. pour cela, fixons un réel arbitraire $\varepsilon > 0$. Par définition de la norme quotient, il existe $y_1, y_2 \in Y$, tels que

$$\|x_1 + x_1\| - \varepsilon \leq \|[x_1]\| \leq \|x_1 + y_1\|, \quad \|x_2 + y_2\| - \varepsilon \leq \|[x_2]\| \leq \|x_2 + y_2\|.$$

En utilisant l'inégalité triangulaire, on obtient

$$\|x_1 + x_2 + y_1 + y_2\| \leq \|x_1 + y_1\| + \|x_2 + y_2\| \leq \|[x_1]\| + \|[x_2]\| + 2\varepsilon.$$

Nous concluons que

$$\|[x_1 + x_2]\| = \inf_{y \in Y} \|x_1 + x_2 + y\| \leq \|x_1 + x_2 + y_1 + y_2\| \leq \|[x_1]\| + \|[x_2]\| + 2\varepsilon.$$

Comme $\varepsilon > 0$ est arbitraire, en faisant tendre ε vers 0 on termine la preuve de l'axiome (iii) de la norme. ■

1.4.39 **Exercice** Considérons le sous-espace Y de $C(K)$ des fonctions constantes. Montrer la formule suivante pour la norme de l'espace quotient $C(K)/Y$:

$$\|[f]\| = \frac{1}{2}\big(\max_{t \in K} f(t) - \min_{t \in K} f(t)\big) \quad \text{pour} f \in C(K).$$

1.4.40 **Exercice** Montrer la formule suivante pour la norme de l'espace quotient ℓ^∞ / c_0 :

$$\|[a]\| = \limsup |a_i| \quad \text{pour } a = (a_i)_{i=1}^\infty \in \ell^\infty.$$

1.4.7 Exercices supplémentaires

1.4.41 Exercice (Produit de deux espaces normés) Soit X et Y deux espaces vectoriels normés. Le produit (cartésien direct) de X et Y est l'espace

$$X \oplus_1 Y = \{(x, y) : \ x \in X, \, y \in Y\}.$$

Montrer que $X \oplus_1 Y$ est un espace vectoriel normé, avec la norme définie par

$$\|(x, y)\| := \|x\| + \|y\|.$$

1.4.42 Exercice (la fonctionnelle de Minkowski) Soit K un sous-ensemble de $\mathbb{R}^n$, fermé, convexe, symétrique par rapport à l'origine et d'intérieur non vide.

La fonctionnelle de Minkowski de K est la fonction définie sur $\mathbb{R}^n$ par

$$\|x\|_K := \inf\left\{t > 0 \ : \ \frac{x}{t} \in K\right\}.$$

Montrer que $\|\cdot\|_K$ est une norme sur $\mathbb{R}^n$, et que la boule unité pour cette norme est K.

1.4.43 Exercice (semi-normes) Une semi-norme sur un espace vectoriel E est une fonction $p : E \to \mathbb{R}$, qui satisfait tous les axiomes de la norme, sauf l'axiome (i). Autrement dit, il peut exister un vecteur non nul x tel que $p(x) = 0$.

Montrer que l'on peut transformer une semi-norme en une norme par l'affacturage sur les directions zéro.

1. Montrer que $\ker(p) := \{x \in E : p(x) = 0\}$ est un sous-espace vectoriel de E.
2. Montrer
$$\|[x]\| := p(x) \quad \text{pour } x \in E.$$
est une norme sur l'espace quotient $E/\ker(p)$.

On illustre cette procédure en construisant les espaces vectoriels normés L^p.

1.4.44 Exemple (**Espaces L^p :**). Soit (X, Σ, μ) un espace mesuré.

Soit $p \in [1, +\infty[$. On définit une semi-norme sur l'espace des fonctions mesurables $f : X \to \bar{\mathbb{C}}$ par

$$\|f\|_p := \left(\int_X |f|^p \, d\mu\right)^{1/p}$$

On note par $\mathcal{L}^p(X)$ l'ensemble des fonctions mesurable à valeurs dans $\bar{\mathbb{C}}$ pour lesquelles $\|f\|_p < +\infty$.

On notera que $\mathcal{L}^p(X)$ est un espace vectoriel ; en effet, pour $f, g \in \mathcal{L}^p(X)$ et $\lambda, \beta \in \mathbb{C}$ on a : $|\lambda f + \beta g|^p \leq 2^p \max(|\lambda||f|, |\beta||g|)^p \leq 2^p(|\lambda| + |\mu|)^p \max(|f|^p, |g|^p) \leq 2^p(|\lambda| + |\mu|)^p(|f|^p + |g|^p)$. D'où $\lambda f + \beta g \in \mathcal{L}^p(X)$.
$\|\|_p$ n'est pas une norme une $\mathcal{L}^p(X)$, en effet

1.4.45 Lemme

Soit $f : X \to \overline{\mathbb{C}}$ une fonction mesurable. $\|f\|_p = 0$ si et seulement si $f(x) = 0$ p.p.

Démonstration: Pour $n \geq 1$ on pose $A_n = \{x \in X|\,|f(x)| \geq \frac{1}{n}\}$. Alors $A_n \subset A_{n+1}$ et $A = \bigcup_{n\geq 1} A_n = \{x \in X|f(x) \neq 0\}$.
On a pour tout $n \geq 1$, $0 = \int_X |f|^p\,d\mu \geq \int_{A_n} |f|^p\,d\mu \geq \left(\frac{1}{n}\right)^p \mu(A_n)$, ainsi $\mu(A_n) = 0$ et par suite $\mu(A) = \lim_n \mu(A_n) = 0$. La réciproque est évidente. ∎

On pose alors $\mathcal{N}(X) = \{f \in \mathcal{L}^p(X)\,|\,f(x) = 0 \text{ p.p. }\}$.

1.4.47 Lemme

$\mathcal{N}(X)$ est sous-espace vectoriel fermé de l'espace des fonctions mesurables.

Il s'en suit que pour obtenir une norme il faudrait identifier les fonctions égales p.p. i.e. $f \sim g$ si et seulement si $\mu(\{x \in X\,|\,|f(x) \neq g(x)\}) = 0$ si et seulement si $f - g \in \mathcal{N}(X)$.

Comme $\mathcal{N}(X)$ est un sous-espace vectoriel fermé de $\mathcal{L}^p(X)$, l'espace vectoriel quotient, est un evn :

1.4.48 Définition

$L^p = L^p(X, \Sigma, \mu) := \mathcal{L}^p(X)/\mathcal{N}(X)$

Les éléments de $L^p(X)$ sont des classes d'équivalences de fonctions $[f]$, telle que un représentant $f : X \to \overline{\mathbb{C}}$: est mesurable et $\left(\int_X |f|^p\,d\mu\right)^{1/p} < \infty$

Notez que ℓ^p est un cas particulier de l'espace $L^p(X, \Sigma, \mu)$ où $X = \mathbb{N}$ et μ est la mesure de comptage sur $\mathbb{N}$.

Espaces L^∞ : Pour le cas $p = +\infty$, on définit le supremum essentiel par :

1.4.49 Définition

Si $f : X \to \overline{\mathbb{C}}$ est mesurable, son supremum essentiel (ou sa borne supérieure essentielle) , notée $\|f\|_\infty$, est définie par

$$\|f\|_\infty = \inf\{\alpha > 0\,|\,\mu(\{x \in X||f(x)| \geq \alpha\}) = 0\}\,.$$

On utilise parfois la notation *esssup* f ou *supess* f pour désigner $\|f\|_\infty$.

On note par $\mathcal{L}^\infty(X)$ l'ensemble des fonctions mesurable à valeurs dans $\mathbb{C}$ pour lesquelles $\|f\|_\infty < +\infty$.

et on définit de même :

1.4.50 DÉFINITION
$L^\infty = L^\infty(X, \Sigma, \mu) := \mathcal{L}^\infty(X)/\mathcal{N}(X)$

Les éléments de $L^\infty(X)$ sont des classes d'équivalences de fonctions $[f]$ (identifié à $f + \mathcal{N}(X)$), telle que un représentant $f : X \to \overline{\mathbb{C}}$: mesurable et $\left(\int_X |f|^\infty \, d\mu\right)^{1/p} < \infty$.

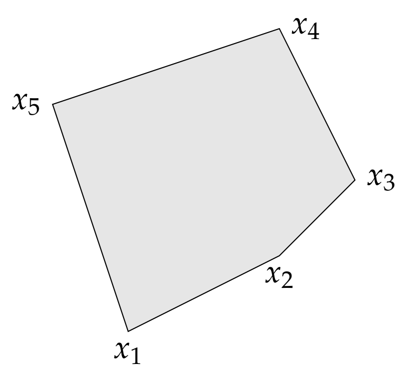
Notez que ℓ^∞ est un cas particulier de l'espace $L^\infty(\Omega, \Sigma, \mu)$ où $\Omega = \mathbb{N}$ et μ est la mesure de comptage sur $\mathbb{N}$.

1.4.51 Exercice (sup essentiel) Montrer que la norme dans $L^\infty(\Omega, \Sigma, \mu)$ peut être aussi définie par

$$\|f\|_\infty = \inf_{\mu(A)=0} \sup_{t \in \Omega \setminus A} |f(t)|$$

où l'infimum porte sur tous les sous-ensembles $A \subset \Omega$ de mesure nulle.

1.4.52 Exercice (Enveloppe convexe) L'enveloppe *convexe* d'un sous-ensemble A d'un espace vectoriel X est définie comme étant le plus petit convexe qui contient A. L'enveloppe convexe de A est noté $conv(A)$. (Voir la figure pour un exemple).



La région hachurée est l'enveloppe convexe $\text{conv}(\{x_1, \ldots, x_5\})$

Une *combinaison convexe* de vecteurs $x_1, \ldots, x_n$ dans un espace vectoriel X est un vecteur de la forme

$$x = \sum_{k=1}^{n} \lambda_k x_k$$

où $\lambda_k \geq 0$ sont des nombres réels tels que $\sum_{k=1}^{n} \lambda_k = 1$. Montrer que $conv(A)$ coïncide avec l'ensemble de toutes les combinaisons convexes d'un nombre fini de vecteurs de A.

1.5 Espaces de Banach

1.5.1 Définition. Complétude de $C(K)$.

Il s'avère que le concept d'espace vectoriel normé est déficient ; de nombreux résultats de l'analyse ne peuvent pas être obtenus juste en se fondant sur les axiomes de norme. Un axiome supplémentaire est nécessaire.

Rappelons que X est un espace métrique est dit **complet** si toute suite de Cauchy dans X converge vers un point dans X. Par exemple, $\mathbb{R}$ est un espace métrique complet alors que $\mathbb{Q}$ ne l'est pas.

Rappelons qu'une suite $(x_i)_{i=1}^{\infty}$ dans un espace vectoriel normé X est de *Cauchy* si

$$\|x_n - x_m\| \to 0 \quad \text{lorsque } n, m \to \infty,$$

c'est à dire pour chaque $\varepsilon > 0$, il existe $N = N(\varepsilon)$ tel que

$$\|x_n - x_m\| < \varepsilon \quad \text{pour tout } \ n, m > N.$$

1.5.1 DÉFINITION (ESPACE DE BANACH)
Un espace vectoriel normé complet est appelé espace de **Banach**.

1.5.2 Exercice (sous-espaces, les quotients d'espaces de Banach) Soit X un espace de Banach et Y un sous-espace (linéaire) de X. Montrer que :

1. Y est un espace de Banach si et seulement si Y est fermé.
2. Si Y est fermé[6] alors X/Y est un espace de Banach pour la norme quotient.

De nombreux espaces sont des espaces de Banach.

1.5.3 THÉORÈME
Pour un espace topologique compact K, $C(K)$ est un espace de Banach.

Démonstration: La plupart des preuves de complétude, réduise le problème à la complétude de $\mathbb{K}$. Soit une suite de Cauchy (f_n) dans $C(K)$, c'est-à-dire

$$\|f_n - f_m\|_\infty \to 0, \quad n, m \to \infty. \tag{1.5.1}$$

6. On rappelle que Y doit être fermée pour que X/Y soit un espace vectoriel normé.

Par conséquent, pour tout $t \in K$, nous avons $|f_n(t) - f_m(t)| \to 0$. En d'autres termes, $(f_n(t))$ est une suite de Cauchy dans $\mathbb{K}$ pour chaque t. Comme $\mathbb{K}$ est complet, cette suite a une limite que nous appelons $f(t)$. Nous avons construit une fonction $f(t)$ tel que $f_n \to f$ simplement.

Nous affirmons que $f_n \to f$ uniformément, ie $\|f_n - f\|_\infty \to 0$. Cela complétera la preuve, car la limite d'une suite uniformément convergente de fonctions continues sur l'espace compact K est une fonction continue. Montrons notre revendication Alors. Par (1.5.1), pour tout ε, il existe $N = N(\varepsilon)$ de telle sorte que

$$|f_n(t) - f_m(t)| < \varepsilon \quad \text{pour tout } n, m > N,\ t \in K.$$

On fait tendre $m \to \infty$ dans cette inégalité (tout en gardant tout le reste fixe), nous concluons que

$$|f_n(t) - f(t)| < \varepsilon \quad \text{pour tout } n, m > N,\ t \in K.$$

Cela signifie que $\|f_n - f\|_\infty \to \infty$, ce que nous voulions montrer. ∎

1.5.5 Exercice (espaces de Banach ℓ^∞, c_0, L^∞)
1. Montrer que ℓ^∞ et L^∞ sont des espaces de Banach. (Indication : modifier la preuve du théorème 1.5.3.)
2. Montrer que c_0 est un espace de Banach (Indication : utiliser l'exercice 1.4.35.) Montrer que l'ensemble des fonctions

$$\{f \in C([0,1]) :\ f(0) = f(1) = 0\}$$

est un espace de Banach par rapport à la norme-sup. (Indication : vérifiez qu'il s'agit d'un sous-espace fermé de $C[0,1]$.)
3. Montrer que l'ensemble des polynômes $\mathbb{K}(x)$ limités à $[0,1]$ n'est pas un espace de Banach pour rapport à la norme-sup. (Indication : rappel que $\mathbb{K}(x)$ est dense dans $C[0,1]$ par le théorème d'approximation de Weierstrass.)
4. Montrer que c_{00} n'est pas un espace de Banach par rapport à toute norme $\|\cdot\|_p$, $1 \le p \le \infty$. (Indication : montrer que c_{00} est dense dans ℓ^p pour $p \in [1, \infty)$, et n'est pas fermé dans ℓ^∞).

1.5.2 Séries dans les espaces de Banach. Complétude de L^p

———————————————-

Nous allons donner un critère utile de complétude des espaces normés en termes de convergence des séries au lieu des suites. Nous allons utiliser ce critère pour prouver la complétude des espaces L^p.

1.5.6 Définition

Série Soit (x_k) une suite de vecteurs dans un espace vectoriel normé X. Si la suite des sommes partielles
$$S_n := \sum_{k=1}^{n} x_k$$
converge vers un vecteur $x \in X$, lorsque $n \to \infty$, nous dirons que la série $\sum_k x_k$ *converge* dans X, et nous écrivons
$$\sum_{k=1}^{\infty} x_k = x.$$
Une série $\sum_k x_k$ est dite *normalement (absolument) convergente* si
$$\sum_{k=1}^{\infty} \|x_k\| < \infty.$$

Rappelons que dans le cas scalaire, où $X = \mathbb{R}$ ou $\mathbb{C}$, la convergence absolue de la série implique sa convergence (la réciproque est en général fausse). Comme le montre le théorème suivant, cela arrive justement grâce à de la complétude de $\mathbb{R}$ et $\mathbb{C}$.

1.5.7 Théorème (critère de complétude)

Un espace vectoriel normé X est un espace de Banach si et seulement si chaque série absolument convergente dans X converge dans X.

Démonstration: 1. Nécessité : Soit X un espace de Banach, et une série absolument convergente
$$\sum_{k=1}^{\infty} \|x_k\| < \infty. \tag{1.5.2}$$
Nous voulons montrer que la série $\sum_k x_k$ converge. Par complétude de X, il suffit de montrer que les sommes partielles de cette série forment une suite de Cauchy, c'est à dire que $\|S_n - S_M\| \to 0$ $n > m \to \infty$. pour cela, nous utilisons l'inégalité triangulaire et notre hypothèse de (1.5.2) pour obtenir
$$\|S_n - S_M\| = \Big\| \sum_{k=n+1}^{m} x_k \Big\| \leq \sum_{k=n+1}^{m} \|x_k\| \to 0.$$
Ceci termine la preuve de la nécessité.

2. Suffisance : Supposons que X ne soit pas complet ; nous voulons construire une série divergente et qui soit absolument convergente. Par la non complétude de X, il existe une suite de Cauchy (v_n) dans X, qui diverge.

Alors, toute sous-suite de (v_n) diverge. Par conséquent, une sous-suite (w_n) de (v_n) telle que

$$\|w_2 - w_1\| \leq \frac{1}{2}, \quad \|w_3 - w_2\| \leq \frac{1}{2^2}, \ldots, \quad \|w_{n+1} - w_n\| \leq \frac{1}{2^n}, \quad \ldots \tag{1.5.3}$$

diverge.

Il s'ensuit que la suite (x_k) définie par

$$x_1 := w_2 - w_1, \quad x_2 := w_3 - w_2, \ldots, \quad x_k := w_{k+1} - w_k, \quad \ldots$$

forme les termes d'une série absolument convergente :

$$\sum_{k=1}^{\infty} \|x_k\| \leq \sum_{k=1}^{\infty} \frac{1}{2^k} = 1.$$

D'autre part, la suite des sommes partielles

$$S_n = \sum_{k=1}^{n} x_k = w_{n+1} - w_1$$

diverge. Donc, nous avons construit une série absolument convergente dans X qui diverge. Ceci termine la preuve. ■

1.5.9 Exercice Valider les deux étapes manquantes dans la preuve du théorème 1.5.7. Soit X un espace vectoriel normé.

1. Soit (v_n) une suite de Cauchy X qui diverge. Montrer que tout sous-suite de (v_n) diverge.
2. Soit (v_n) est une suite de Cauchy dans X. Construire une sous-suite "rapidement de Cauchy" (w_n) de (v_n), c'est à dire celle qui satisfait (1.5.3).

1.5.10 Théorème (de Riesz-Fischer)
Pour chaque $p \in [1, \infty]$, l'espace $L^p = L^p(X, \Sigma, \mu)$ est un espace de Banach.

Démonstration: On va commencer par le cas $\boxed{1 \leq p < +\infty}$. Soit (f_n) une suite de Cauchy dans L^p ; remplaçant au besoin cette suite par une sous-suite on peut supposer que

$$\|f_{n+1} - f_n\|_p \leq \frac{1}{2^n};$$

posant $g_0 = f_0$ et $g_n = f_n - f_{n-1}$, pour $n \geq 1$ et $G = \sum_{n=0}^{+\infty} |g_n|$.

Alors, la suite $G_n = \sum_{k=0}^{n} |g_k|$ est dans $L^p(X)$, en effet pour tout n, $\|G_n\|_p = \| \sum_{k=0}^{n} |g_k| \|_p \leq \sum_{k=0}^{n} \|g_k\|_p \leq \sum_{k=0}^{n} \frac{1}{2^{k-1}} + \|f_0\|_p \leq 1 - \frac{1}{2^{n-1}} + \|f_0\|_p \leq 1 + \|f_0\|_p$.

D'autre part, la suite G_n est positive, croissante et $\lim_n G_n(x) = G(x)$, d'après le théorème de convergence monotone

$$\lim_{n\to+\infty} \int_X G_n^p d\mu = \int_X G^p d\mu$$

i.e. $\|G\|_p = \lim_{n\to+\infty} \|G_n\|_p \leq 1 + \|f_0\|_p$. Ainsi, $G \in L^p(X)$, en particulier, $0 \leq G(x) \leq +\infty$ p.p. par suite la série et $\sum_{n=0}^{+\infty} g_n(x)$ est absolument convergente dans $\mathbb{C}$, p.p. donc convergente p.p. i.e. $\sum_{n=0}^{+\infty} g_n(x) = \lim_{n\to+\infty} f_n(x)$ existe p.p. on note par $f(x)$ cette limite quand elle existe et on pose $f(x) = 0$ si la limite n'existe pas. Alors $f(x) = \lim_{n\to+\infty} f_n(x)$ p.p., donc f est mesurable, il reste à montrer que $f \in L^p$ et que $\lim_{n\to+\infty} \|f_n - f\|_p = 0$. Comme $f(x) = \sum_{n=0}^{+\infty} g_n(x)$ p.p., $|f(x)| \leq \sum_{n=1}^{+\infty} |g_n(x)|$ p.p. d'où $\|f\|_p \leq \sum_{n=0}^{+\infty} \|g_n\|_p \leq +\infty$. i.e. $f \in L^p$. On a $\lim_{n\to+\infty} |f_n(x) - f(x)|^p = 0$, $|f_n(x) - f(x)|^p \leq 2^p G^p(x)$ p.p. et $2^p G^p \in L^1(X)$, et d'après le théorème de convergence dominée :

$$\lim_{n\to+\infty} \int_X |f_n(x) - f(x)|^p d\mu = 0;$$

on voit ainsi que f_n converge vers f dans L^p.

$\boxed{\text{Le cas } p = +\infty}$.

Soit $\sum_n f_n$ une série normalement convergente de $L^\infty(X)$ i.e. $\sum_n \|f_n\|_\infty < \infty$. On doit montrer que la série converge. Posons $A_n = \{x \in X |\, |f_n(x)| > \|f_n\|_\infty\}$. Par définition de $\|.\|_\infty$, A_n est de mesure nulle et par suite $A = \cup_{n=0}^{+\infty} A_n$ est de mesure nulle et pour tout $x \in X \setminus A$; $\sum_{n=0}^{+\infty} |f_n(x)| < \infty$. Comme $\mathbb{C}$ est complet, pour tout $x \in X \setminus A$; $\sum_{n=0}^{+\infty} f_n(x)$ converge, on note par $f(x)$ cette somme et on pose $f(x) = 0$ si $x \in A$. Alors f est mesurable comme somme d'une série de fonctions mesurables convergente p.p. et $|f(x)| \leq \sum_{n=0}^{+\infty} \|f_n\|_\infty < \infty$ d'où $f \in L^\infty(X)$.

D'autre part $\left| f(x) - \sum_{n=0}^{N} f_n(x) \right| \leq \sum_{n=N+1}^{+\infty} \|f_n\|_\infty$ pour tout $x \in X \setminus A$ ainsi

$$\left\| f - \sum_{n=0}^{N} f_n \right\|_\infty \leq \sum_{n=N+1}^{+\infty} \|f_n\|_\infty \to 0 \quad \text{lorsque } N \to +\infty$$

i.e. la série $\sum_n f_n$ converge dans $L^\infty(X)$.

1.6 Applications linéaires continues (ou opérateurs linéaires bornés)

Dans cette section, nous allons étudier les applications linéaires continues $T : X \to Y$ entre espaces vectoriels normés X et Y. Les formes linéaires des cas particuliers d'opérateurs linéaires avec Y, soit le corps des scalaires, $\mathbb{R}$ ou $\mathbb{C}$. Ainsi, plusieurs résultats pour les opérateurs linéaires seront des généralisations de ceux que nous avons pour les formes linéaires

1.6.1 Norme d'opérateur. Continuité et bornitude.

1.6.1 DÉFINITION (NORME D'OPÉRATEUR)
opérateur linéaire ! Norme Un opérateur linéaire $T : X \to Y$ entre espaces vectoriels normés X et Y est dit *borné* s'il existe un nombre $C \geq 0$ tel que

$$\|Tx\| \leq C\|x\| \quad \text{pour tout } x \in X$$

La *norme de T est la plus petite constante C dans cette inégalité, elle est donc définie comme*

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = sup_{\|x\|=1} \|Tx\|.$$

1.6.2 REMARQUE
Nous avons toujours l'inégalité

$$\|Tx\| \leq \|T\| \, \|x\|, \quad \text{pour tout } x \in X.$$

1.6.3 EXEMPLE (COMPOSITION DES OPÉRATEURS). Soit $T : X \to Y$ et $S : Y \to Z$ sont des opérateurs linéaires bornés entre espaces vectoriels normés, alors $S \circ T : X \to Z$ est un opérateur linéaire borné, et

$$\|S \circ T\| \leq \|T\| \, \|S\|.$$

on note souvent ST à la place de $S \circ T$.

1.6.4 PROPOSITION
Un opérateur linéaire est continu si et seulement si il est borné.

1.6.5 Exercice Prouver qu'un opérateur linéaire $T : X \to Y$ est continu si et seulement si l'image de toute suite qui convergent vers zéro est bornée.

1.6.2 Espace d'opérateurs

Soit X et Y des espaces normés. L'espace des applications linéaires continues(appelé aussi espace des opérateurs linéaires bornés) $T : X \to Y$ muni de la norme d'opérateur est noté $L(X, Y)$.

A titre d'exemple, l'espace dual topologique $X' = L(X, \mathbb{K})$.

1.6.6 PROPOSITION

$L(X, Y)$ est un espace normé. De plus, si Y est un espace de Banach alors $L(X, Y)$ est aussi un espace de Banach. En particulier, l'espace dual X' est toujours un espace de Banach.

Démonstration: Les vérifications des axiomes de norme sont simples. Pour prouver la complétude, on prend une suite de Cauchy $T_n \in L(X, Y)$, c'est à dire

$$\|T_n - T_m\| \to 0, | n, m \to \ infty.$$

En appliquant cela à un élément arbitraire $x\ inX$, et notant que

$$\|T_n x - T_m x\| \leq \|T_n - T_m\| \, \|x\| \to 0, \quad n, m \to \infty$$

nous constatons que $(T_n x)$ est une suite de Cauchy dans Y. Par compluétude de Y, elle converge. On définit l'application T par

$$Tx := \lim_n T_n x.$$

Si on montre que T est la limite de T_n dans $L(X, Y)$, ceci achèvera la démonstration.

Il est facile de vérifier que $T : X \to Y$ est une application linéaire.

Pour montrer que T est continu, nous choisissons un $x \in X$, arbitraire, et on utilise la continuité de la norme :

$$\|Tx\| = \lim_n \|T_n x\| \leq \sup_n \|T_n\| \|x\|.$$

Comme toute suite de Cauchy est bornée, $\sup_n \|T_n\| < \ infty$. Il en résulte que T est un opérateur linéaire borné, ie $T \in L(X, Y)$.

Il reste à montrer que T_n tend vers T dans $L(X, Y)$, c'est à dire pour la norme d'opérateur. Comme T_n est de Cauchy, pour tout $\varepsilon > 0$, il existe un tel $N \in \mathbb{N}$, tel que

$$\|T_n - T_m\| < \varepsilon \quad \text{pour} n, m > N.$$

En appliquant cela à un élément arbitraire $x \in X$, on obtient

$$\|T_n x - T_m x\| < \varepsilon \|x\| \quad \text{pour } n, m > N.$$

On fait tendre $m \to \infty$, et on obtient

$$\|T_n x - Tx\| < \varepsilon \|x\| \quad \text{pour } n > N.$$

Comme x est arbitraire, il s'ensuit que

$$\|T_n - T\| < \varepsilon \quad \text{pour } n > N.$$

Cela signifie que T_n converge vers T dans $L(X, Y)$ comme souhaité. ■

1.6.3 Les opérateurs sur l'espace euclidien de dimension finie

Soit $X = Y = \ell_2^n$; rappel que ℓ_2^n est l'espace euclidien $(\mathbb{R}^n, \|\cdot\|_2)$. Comme nous le savons de l'algèbre linéaire, T peuvent être identifiés avec sa matrice $n \times n$ dans la base canonique (T_{ij}), où

$$t_{ij} = \langle Te_j, e_i \rangle, \quad i, j = 1, \ldots, n$$

et où (e_i) désigne la base canonique de ℓ_2^n. De cette manière, la i-ième coordonnée du vecteur Tx peut être calculée par

$$(Tx)_i = \sum_{j=1}^{n} t_{ij} x_j. \tag{1.6.1}$$

1.6.8 PROPOSITION
Tout opérateur linéaire $T : \ell_2^n \to \ell_2^n$ est borné. Plus précisément,

$$\|T\| \leq \|T\|_{\mathrm{HS}}$$

où $\|T\|_{\mathrm{HS}}$ désigne la *norme de Hilbert-Schmidt* de T ! opérateur de Hilbert-Schmidt défini comme

$$\|T\|_{\mathrm{HS}} = \Big(\sum_{i,j=1}^{n} |t_{ij}|^2 \Big)^{1/2}.$$

.

Démonstration: Utiliser (1.6.1) et en appliquant l'inégalité de Cauchy-Schwarz, on obtient

$$\|Tx\|_2^2 = \sum_{i=1}^{n} |(Tx)_i|^2 = \sum_{i=1}^{n} \Big| \sum_{j=1}^{n} t_{ij} x_j \Big|^2 \leq \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} |t_{ij}|^2 \Big) \Big(\sum_{j=1}^{n} |x_j|^2 \Big) = \|T\|_{\mathrm{HS}}^2 \, \|x\|_2^2.$$

D'où le résultat. ■

1.6.4 Complétion

1.6.10 THÉORÈME (COMPLÉTION)
Soit X un espace vectoriel normé. Il existe un espace de Banach $\widehat{X}$, appelé *complétion de X*, ayant les propriétés suivantes : il existe une application linéaire $\iota : X \to \widehat{X}$ telle que :

(i) $\|\iota(x)\| = \|x\|$ pour tout $x \in X$ (c.-à-dire que ι est un plongement isométrique de X dans $\widehat{X}$;

(ii) l'image de ι, $\mathrm{Im}(\iota)$ est dense dans $\widehat{X}$.

Le complété de X est unique à une isométrie près.[7]

Démonstration: On considère l'ensemble $\tilde{X}$ de toutes les suites de Cauchy de X. On définit une relation d'équivalence $\sim$ sur $\tilde{X}$ par deux suites de Cauchy sont équivalentes si leur différence tend vers 0. On note $\widehat{X} = \tilde{X}/\sim$ l'ensemble des classes d'équivalences. Il n'est pas difficile de voir que $\widehat{X}$ et $\tilde{X}$ héritent d'une structure d'espace vectoriel.

1.6.12 LEMME
Si $(x_n)_{n\in\mathbb{N}}$ est une suite de cauchy de E alors la suite $(\|x_n\|)_{n\in\mathbb{N}}$ converge dans $\mathbb{R}$.

Démonstration: Grâce à l'inégalité $|\|x_n\| - \|x_m\|| \leq \|x_n - x_m\|$, $(\|x_n\|)_{n\in\mathbb{N}}$ est une suite de Cauchy et la complétude de $\mathbb{R}$ permet de conclure. ■

On peut alors définir la norme d'une suite de Cauchy $(x_n)_{n\in\mathbb{N}}$, par

$$\|(x_n)\|_0 = \lim_{n\to+\infty} \|x_n\|.$$

Vérifier que $\|.\|_0$ est une norme sur $\widehat{X}$ (mais pas sur $\tilde{X}$ pourquoi ?).

Soit $(x^k) = \left([(x_n^k)_{n\in\mathbb{N}}]\right)$ est une suite de Cauchy de $\widehat{X}$, alors $y = [(x_n^n)_{n\in\mathbb{N}}]$ est sa limite (à vérifier). D'où $\widehat{X}$ est complet.

On identifie $x \in X$ dans $\widehat{X}$ avec la classe des suites qui convergent vers x, donc $\iota : X \to \widehat{X}$ est définie par $\iota(x) = [x]$ d'où le résultat. ■

La notion de complétude suggère une construction alternative de l'espace $L^p[a,b]$, $p \in [1,\infty)$. De la théorie mesure, nous savons que l'ensemble des fonctions continues est un sous-ensemble dense de $L^p[a,b]$ (pourquoi ?). De plus, $L^p[a,b]$ est un espace vectoriel normé complet. Par conséquent (par l'unicité de la complétion), *$L^p[a,b]$ est le complété de l'espace $C[a,b]$ pour la norme $\|\cdot\|_p$.*

Il s'agit d'une définition alternative de $L^p[a,b]$. Cela donne aussi une construction alternative de l'intégrale de Lebesgue. En effet, pour les fonctions continues, les intégrales de Riemann et de Lebesgue coïncident. Par conséquent, l'espace $C[a,b]$ avec $\|\cdot\|_1$ peut être construit en utilisant uniquement intégrale de Riemann et sa complétion donne l'intégrante de Lebesgue.

---

7. Plus précisément, cela signifie que pour toute autre complété $\widehat{X'}$, il existe une bijection linéaire $T : \widehat{X} \to \widehat{X'}$ telle que $\|Tx\| = \|x\|$ pour tout $x \in X$

1.6.5 Exercices supplémentaires

1.6.14 EXEMPLE (ESPACE DE FONCTIONS CONTINUES PÉRIODIQUES). Montrer que l'ensemble de fonctions

$$\{f \in C[0,1] : f(0) = f(1)\}$$

est un espace de Banach par rapport à la norme-sup. (Indication : identifier cet espace à $C(\mathbb{T})$ où $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ est le tore à une dimension.)

1.6.15 EXEMPLE (ESPACE DE PERMANENCE FONCTIONS DIFFÉRENTIABLES). Montrer que l'espace $C^k[0,1]$ des fonctions k-fois continûment différentiables n'est pas un espace de Banach par rapport à la norme-sup.

Montrer que par contre $C^k[0,1]$ est un espace de Banach par rapport à la norme

$$\|f\|_{C^k} = \|f\|_\infty + \|f'\|_\infty + \cdots + \|f^{(k)}\|_\infty.$$

1.6.16 EXEMPLE (COMPLÉTUDE D'UNE SOMME DIRECTE). Soient X et Y des espaces de Banach. Montrer que la somme directe $X \oplus_1 Y$ définie dans l'exercice 1.4.41 est un espace de Banach.

Chapitre 2

Théorèmes fondamentaux de l'analyse fonctionnelle

Dans ce chapitre, nous allons étudier les trois théorèmes qui, avec le théorème de Hahn-Banach, forment les grands principes de l'analyse fonctionnelle. Ce sont le théorème de Banach-Steinhaus (le principe de la borne uniforme), le théorème de l'application ouverte et le théorème du graphe fermé.

2.1 Théorème de Baire

L'étude des espaces métriques complets nous permet de remplacer certains arguments récurrents par des principes généraux. L'un des plus important de ces principes est donnée par le théorème de catégorie de Baire.

2.1.1 DÉFINITION
Soit (X, d) un espace métrique.

1. Un sous-ensemble A de X est dit *rare* (en anglais : *nowhere dense*) lorsque son adhérence $\bar{A}$ est d'intérieur vide.
2. Un sous-ensemble B de X est dit *maigre* (en anglais : *Baire first category*) lorsqu'il est réunion dénombrable de partie rare.

2.1.2 REMARQUE
Une partie d'un ensemble maigre est maigre.

2.1.3 DÉFINITION
On dit que (X, d) un espace espace de Baire lorsqu'il vérifie l'une des assertion équivalentes suivantes :

1. Toute intersection dénombrable d'ouverts dense est dense.
2. Toute réunion dénombrable de fermés d'intérieurs vide est d'intérieur vide.
3. Le complémentaire d'une partie maigre est dense.
4. Toute partie maigre est d'intérieur vide.
5. Tout ouvert non vide n'est pas maigre.

On peut maintenant formuler le celèbre résultat qui se trouve parmi les théorèmes les plus utilisés de toutes les mathématiques.

2.1.4 Théorème (Théorème de Baire)

1. Un espace métrique complet est un espace de Baire.
2. Un espace métrique localement compact est un espace de Baire.
   (localement compact signifie que pour tout $x \in X$ il existe $r_x > 0$ tel que $\bar{B}(x, r_x)$ soit un compact)

Démonstration: 1. Soit $(U_n)_{n\in\mathbb{N}}$ une suite d'ouverts denses de X. Alors, pour tout n, $F_n = X \setminus U_n$ est un fermé rare.

Soit $x_0 \in X$ et $\delta_0 > 0$.

Nous allons montrer qu'il existe un $y \in B(x_0, \delta_0)$ telle que $y \notin \bigcup_{n\in\mathbb{N}} F_n$ ce qui est équivalent $y \in \bigcap_{n\in\mathbb{N}} U_n$ ce qui montre la densité de l'intersection des U_n.

Pour cela, procédons à la construction par récurrece suivante :

on suppose qu'on a déjà construit $x_{n-1} \in X$ et $\delta_{n-1} > 0$, on peut alors trouver $x_n \in X$ tel que $x_n \in B(x_{n-1}, \delta_n/4)$, et $x_n \notin F_n$ (sinon, nous aurions $B(x_{n-1}, \delta_{n-1}/4) \subseteq F_n$ et F_n aurait un intérieur non vide). Comme F_n est fermé et $x_n \notin F_n$ nous pouvons maintenant trouver un δ_n, $\delta_{n-1}/2 > \delta_n > 0$ tel que $B(x_n, \delta_n) \bigcap F_n = \emptyset$.

Maintenant, observons que

$$\delta_m \leq 2^{-1}\delta_{m-1} \leq 2^{-2}\delta_{m-2} \leq \ldots 2^{n-m}\delta_n$$

pour tout $m \geq n \geq 0$. Il s'ensuit que, si $r \geq s$

$$d(x_r, x_s) \leq \sum_{j=r}^{s-1} d(x_{j+1}, x_j) \leq \sum_{j=r}^{s-1} \delta_r/4 \leq 4^{-1}\delta_0 \sum_{j=r}^{s-1} 2^{-r} \leq 2^{-r-1}\delta_0.$$

Ainsi, x_r est une suite de Cauchy et converge donc dans (X, d) vers un certain point y.

Le même genre de calcul donne

$$d(x_r, x_s) \leq \sum_{j=r}^{s-1} d(x_{j+1}, x_j) \leq \sum_{j=r}^{s-1} \delta_r/4 \leq 4^{-1}\delta_r \sum_{j=0}^{sr-1} 2^{-r} \leq \delta_r/2,$$

lorsque $s \geq r$ et ainsi

$$d(x_r, y) \leq d(x_r, x_s) + d(x_s, y) \leq \delta_r/2 + d(x_s, y) \to \delta_{r-1}/2 \qquad \blacksquare$$

lorsque $s \to \infty$. Nous avons donc $d(x_r, y) \leq \delta_r/2$ d' où $y \in B(x_0, \delta_r)$ et $y \notin F_r$ pour chaque $r \geq 1$ ce qu'il fallait démontrer.

2. La cas (X, d) localement compact se traite de la même manière.
   Soit U_n une suite d'ouverts denses de X. Alors pour tout n, $F_n = X \setminus U_n$ est un fermé rare.
   Soit $x_0 \in X$ et $\delta_0 > 0$ tel que $\delta_0 < r_{x_0}$ où $\bar{B}(x_0, r_{x_0})$ est un compact.
   la démonstartion se poursuit alors comme dans le cas complet, car la suite de cauchy x_n construite est dans le compact $\bar{B}(x_0, r_{x_0})$, donc converge vers un $y \in \bar{B}(x_0, r_{x_0})$ car tout compact est complet.

Le théorème de Baire peut être reformulé comme suit.

2.1.6 Corollaire

Soit (X, d) un espace métrique non-vide et complet. Supposons que pour tout $n \in \mathbb{N}$, P_n soit une propriété telle que :

(i) La propriété d'être P_n est *stable* dans le sens où, étant donné $x \in X$, qui a la propriété P_n, nous pouvons trouver un $\epsilon > 0$ telle que chaque fois que $d(x, y) < \epsilon$ le point y a la propriété P_n.

(ii) La propriété de ne pas être P_n est *instable* dans le sens où, étant donné $x \in X$ qui n'a pas la propiété P_n et $\epsilon > 0$, on peut trouver un $y \in X$ avec $d(x, y) < \epsilon$, et qui a la propriété P_n.

Alors, il existe un $x_0 \in X$, qui vérifie toutes les propriétés P_1, $P_2, \ldots$, en fait l'ensemble de tels points est dense dans X.

Démonstration: La preuve s'obtient facilement du théorème précédent en remarquant que : x a la propriété P_n si et seulement si $x \notin F_n$. ■

2.1.8 Exercice Montrer que l'union dénombrable d'ensembles maigres est un ensemble maigre.

Vous avez certainement déjà rencontré le théorème suivant.

2.1.9 THÉORÈME
$\mathbb{R}$ n'est pas dénombrable.

Démonstration: Si nous munissons $\mathbb{R}$ de la métrique usuelle, les singletons $\{x\}$ sont des fermés et d'intérieurs vides. Il s'ensuit que, si E est un sous-ensemble dénombrable de $\mathbb{R}$, alors $E = \bigcup_{e \in E} \{e\}$ est maigre et ainsi $E \neq \mathbb{R}$. ■

La preuve standard de ce résultat (vue au premier cycle) utilise les développements décimaux, mais celle-ci évite de parler de la relation entre nombres réels et nombres décimaux. Elle est également beaucoup plus proche de la preuve originale de Cantor.

2.1.11 Exercice Si (X, d) est un espace métrique. Nous disons qu'un point $x \in X$ est *isolé* si nous pouvons trouver un $\delta > 0$ tel que $B(x, \delta) = \{x\}$.

(i) Montrer que le point $x \in X$ est isolé si et seulement si $\{x\}$ est ouvert.

(ii) Montrer que toute espace métrique complet non vide et sans points isolés n'est pas dénombrable.

(iii) Donner un exemple d'un espace métrique complet infini et dénombrable

(vi) Donner un exemple d'un espace métrique complet non dénombrable tel que tout point est isolé.

2.1.12 Exercice La procédure qui suit est bien connue pour la construction d'ensemble de Cantor. Soit $E_0 = [0, 1]$ et soit $\zeta_1, \zeta_2, \ldots$ une suite des nombres réels avec $0 < \zeta_n < 1$. À l'étape n, E_n est l'union de 2^n intervalles fermés disjoints $I(r, n)$ tous de même longueur. Nous définissons E_{n+1} comme l'union de 2^{n+1} intervalles fermés disjoints obtenus en enlevant un intervalle ouvert $J(r, n)$ de longueur ζ_n fois la longueur de l'origine de $l(r, n)$ au centre de $l(r, n)$. (Ainsi, si $I(r, n) = [c_{r,n} - \delta_n, c_{r,n} + \delta]$ nous prenons $J(r, n) =]c_{r,n} - \zeta_n \delta_n, c_{r,n} + \zeta_n \delta_n[$ et

$$E_{n+1} = \bigcup_{r=1}^{2^n} (I(r, n) \setminus J(r, n)).$$

1. Expliquez pourquoi $\zeta = \prod_{n=1}^{\infty} \zeta_n$ est bien défini. Montrer que ζ peut prendre n'importe quelle valeur dans $[0, 1[$.
2. Montrer que $E = \bigcap_{n=1}^{\infty} E_n$ est un fermé nulle part dense et sans points isolés. Montrer que E est de mesure de Lebesgue ζ.
3. Construire un sous-ensemble maigre $H \subseteq [0, 1]$ mais de mesure de Lebesgue 1.
4. Construire un sous-ensemble $P \subseteq \mathbb{R}$ maigre tel que $\mathbb{R} \setminus P$ soit de mesure de Lebesgue nulle.

2.2 Théorème de Banach-Steinhaus

2.2.1 Énoncé et démonstration

Le principe de la borne uniforme est un résultat dû à Banach et Steinhaus.

2.2.1 THÉORÈME (DE BANACH ET STEINHAUS)
Soit X un de Banach et Y un espace vectoriel normé.
Soit une famille d'opérateurs linéaires continus $\{L_\alpha\}_{\alpha \in I} \subseteq L(X,Y)$.
Alors la famille $\{L_\alpha\}_{\alpha\in I}$ *est uniformément bornée*, i.e. $\sup_{\alpha\in I} \|L_\alpha\| < +\infty$
si et seulement si $F := \{x \in X \mid \sup_{\alpha\in I} \|L_\alpha(x)\| < +\infty\}$ n'est pas maigre.

Démonstration: .

"$\Longrightarrow$" Soit $C = \sup_{\alpha\in I} \|L_\alpha\| < +\infty$.
Comme, pour tout $x \in X$, $\|L_\alpha(x)\| \leq \|L_\alpha\|.\|x\|$, on aura $\|L_\alpha(x)\| \leq C\|x\|$, ainsi $F := \{x \in X \mid \sup_{\alpha\in I} \|L_\alpha(x)\| < +\infty\} = X$, n'est donc pas maigre, car X est complet.

"$\Longleftarrow$" Pour chaque $x \in X$, on pose

$$M(x) = \sup_{\alpha\in I} \|L_\alpha(x)\| \quad \text{et pour tout } n \in \mathbb{N} \quad F_n = \{x \in X : \ M(x) \leq n\}.$$

On remarquera que $F_n = \bigcap_{\alpha\in I} \{x \in X : \ \|L_\alpha(x)\| \leq n\}$ est un fermé de X.
Par hypothèse on a

$$F = \bigcup_{n\in\mathbb{N}} F_n,$$

n'est pas un sous-ensemble maigre et le théorème de Baire entraîne alors que l'un des F_n n'est pas d'intérieur vide. En résumé, nous avons montré qu'il existe $m \in \mathbb{N}^*$, $x_0 \in X$ et $\varepsilon > 0$ tel que

$$x_0 + \varepsilon B_X \subseteq F_m.$$

D'où, pour tout $x \in B_X$ la boule unité fermée, on a

$$\|L_\alpha(x)\| = \frac{1}{\varepsilon}\|L_\alpha(x_0 + \varepsilon x - x_0)\| \leq \frac{1}{\varepsilon}\|L_\alpha(x_0 + \varepsilon x)\| + \frac{1}{\varepsilon}\|L_\alpha(x_0)\| \leq \frac{2m}{\varepsilon}$$

Ainsi,

$$\sup_{\alpha\in I} \|L_\alpha\| \leq \frac{2m}{\varepsilon} < \infty$$

comme souhaité.

2.2.3 Remarque

1. Le théorème de Banach-Steinhaus est équivalent à l'alternative suivante :
   (i) ou bien $\sup_{\alpha \in I} \|L_\alpha\| < +\infty$
   (ii) ou bien $\{x \in X \,|\, \sup_{\alpha \in I} \|L_\alpha(x)\| = +\infty\}$ est dense dans X.
2. le théorème de Banach-Steinhaus s'énonce comme une inversion de quantificateurs :
   $\forall x \in X, \exists M_x > 0$ tel que $\forall \alpha \in I, \quad \|L_\alpha(x)\| < M_x\|x\|$
   $\Longrightarrow \exists M > 0$ tel que $\forall x \in X, \forall \alpha \in I, \quad \|L_\alpha(x)\| < M\|x\|$
3. Dans la preuve du théorème 2.2.1, seulement la complétude X a été utilisée. Ainsi, le résultat est vrai si X est un espace de Banach et Y est un espace bectoriel nomé

2.2.4 Exercice Vérifier que les sous-ensembles F_n dans la preuve ci-dessus sont fermés, convexes et symétriques.

Un cas particulier important du théorème précédent où $F = X$

2.2.5 Corollaire (Le principe de la borne uniforme)

Soit X un de Banach et Y un espace vectoriel normé.
Soit une famille d'opérateurs linéaires continus $\{L_\alpha\}_{\alpha \in I} \subseteq L(X, Y)$.
On suppose que pour tout $x \in X$ on a

$$\sup_{\alpha \in I} \|L_\alpha(x)\| < +\infty.$$

Alors la famille $\{L_\alpha\}_{\alpha \in I}$ *est uniformément bornée*, i.e.

$$\sup_{\alpha \in I} \|L_\alpha\| < +\infty.$$

2.2.6 Corollaire

Soit X un de Banach et Y un espace vectoriel normé. Soit une suite $\{f_n\}_{n \in \mathbb{N}} \in L(X, Y)$ telle que pour tout $x \in X$, la suite $(f_n(x))_{n \in \mathbb{N}}$ converge dans Y.

Alors la limite est dans $L(X, Y)$ i.e. si on pose pour tout $x \in X$, $f(x) := \lim_{n \to +\infty} f_n(x)$, alors $f \in L(X, Y)$.

Démonstration: On montre facilement en utilisant l'unicité de la limite que f est linéaire.

Montrons que f est continue. La suite $(f_n(x))_{n\in\mathbb{N}}$ converge, elle est donc bornée ainsi pour tout $x \in X$ i.e. $\sup_{n\in\mathbb{N}} \|f_n(x)\| < +\infty$. D'après le théorème de Banach-Steinhaus, il existe $C > 0$ tel que $\sup_{n\in\mathbb{N}} \|f_n\| < C$, ainsi pour tout $x \in X$, et pour tout $n \in \mathbb{N}$, $\|f_n(x)\| < C\|x\|$, et par passage à la limite on aura $\|f(x)\| < C\|x\|$ i.e. $f \in L(X,Y)$. ■

2.2.8 REMARQUE
Le corollaire précédent ne garantit pas la convergence en norme. Par exemple, si $X = c_0$ et $f_k(x) = x_k$ pour $x = (x_n)_{n\in\mathbb{N}}$. Alors

1. pour tout $x \in c_0$, $\lim_{k\to+\infty} f_k(x) = \lim_{k\to+\infty} x_k = 0$.
2. on a $|f_k(x)| \leq \|x\|_\infty$ et $|f_k(e_k)| = 1$ pour $e_k = (\delta_{k,n})$ donne $\|f_k\| = 1$.

Donc $\|f_k\|$ ne converge pas vers 0.

2.2.2 Applications

I) Densité des fonctions non différentiables

On considère l'espace $C[0,1]$ muni de la norme de convergence uniforme. On va montrer que la majorité des éléments de $C[0,1]$ ne sont pas différentiables.

Soit $x_0 \in]0,1[$ fixé. Soit $\delta > 0$ tel que $]x_0 - \delta, x_0 + \delta[\subset]0,1[$.

Pour $0 < h < \delta$ on définit la forme linéaire ϕ_h sur $C[0,1]$ par : pour tout $f \in C[0,1]$

$$\phi_h(f) := \frac{f(x+h) - f(x-h)}{2h}.$$

On vérifie facilement que ϕ_h est linéaire et que $\|\phi_h\| = \frac{1}{h}$. D'où $\sup_{0<h<\delta} \|\phi_h\| = +\infty$.

Par conséquent, d'après le théorème de Banach-Steinhaus, l'ensemble des f tel que
$\sup_{0<h<\delta} |\phi_h(f)| = +\infty$, est dense dans $C[0,1]$. En d'autre termes, l'ensemble des fonctions qui ne sont pas différentiables en x_0 est dense dans $C[0,1]$.

II) Application à la convergence des séries de Fourier

Une question fondamentale et classique dans l'analyse de Fourier est : de savoir quand une série de Fourier d'une fonction f sur un intervalle converge vers f ?

Les techniques des espaces de Hilbert fournissent une réponse complète à cette question dans le cas de l'espace L^2 (voir le chapitre sur les espaces de Hilbert).

Dans les espaces de fonctions autres que L^2, la réponse à ce problème est souvent trivial et même négative. Malheureusement, telle est la situation dans l'espace des fonctions continues $C[-\pi,\pi]$. Il existe des fonctions continues f dont la série de Fourier ne converge pas dans $C[-\pi,\pi]$. Cela découle d'un résultat un peu plus fort, qui à son tour est une conséquence du théorème de Banach-Steinhaus

2.2.9 Théorème (Série de Fourier divergente)
Il existe une fonction $f \in C[-\pi,\pi]$ dont les sommes partielles de Fourier

$$(S_n f)(t) = \sum_{k=-n}^{n} \widehat{f}(k) e^{ikt}$$

forme une suite non bornée de nombres complexes en $t = 0$. En particulier, la série de Fourier de f n'est pas bornée (donc diverge) en $t = 0$.

Démonstration: Rappelons que $\widehat{f}(k) = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(s)e^{-ik\pi}\,ds$, d'où $(S_n f)(t) = \displaystyle\sum_{k=-n}^{n} \widehat{f}(k)e^{ikt} =$ $\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(s)\left(\sum_{k=-n}^{n}(k)e^{ik(t-s)}\right) ds.$

Un calcul de somme de série donne pour $\theta \neq 0$

$$D_n(\theta) = \sum_{k=-n}^{n} e^{ik\theta} = \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}$$

et $D_n(0) = 2n+1$. Ainsi les sommes partielles de Fourier peuvent être représentées par le produit de convolution de f avec le noyau de Dirichlet D_n i.e. :

$$(S_n f)(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(t-s)f(s)\,ds,$$

Nous sommes intéressés par le comportement de $(S_n f)(0)$. Ceux-ci sont des formes linéaires évidemment sur $C[-\pi,\pi]$, que nous noterons

$$\Phi_n(f) := (S_n f)(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(s)f(s)\,ds.$$

(on utilise la parité de la fonction D_n).

Comme le noyau de Dirichlet D_n est continu, Φ_n sont des formes linéaires continue sur $C[-\pi,\pi]$, et

$$\|\Phi_n\| = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(s)|\,ds = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})s}{\sin\frac{1}{2}s}\right| ds = \frac{1}{\pi}\int_{0}^{\pi}\frac{\left|\sin(n+\frac{1}{2})s\right|}{\sin\frac{1}{2}s}$$

2.2.11 Lemme

On a $\|\Phi_n\| = \frac{4}{\pi^2}\ln(n) + O(1)$ lorsque $n \to +\infty$

Démonstration: En remarquant que pour tout $t \in [0,\pi]$: $\| \sin(n+\frac{1}{2}) - \sin(nt)| \leq 2\sin(\frac{t}{4})$ on aura que $\|\Phi_n\| = \frac{1}{\pi}\int_0^\pi \frac{|\sin(n+\frac{1}{2})s|}{\sin\frac{1}{2}s} = \frac{1}{\pi}\int_0^\pi \frac{|\sin ns|}{\sin\frac{1}{2}s}\,ds + O(1)$
Alors en remarquant que la fonction $s \mapsto \frac{1}{\sin\frac{1}{2}s} - \frac{2}{s}$ est continue sur $]0,\pi]$ et prolongeable par continuité en $s = 0$, et en écrivant

$$\frac{1}{\pi}\int_0^\pi \frac{|\sin ns|}{\sin\frac{1}{2}s}\,ds = \frac{2}{\pi}\int_0^\pi \frac{|\sin ns|}{s}\,ds + \frac{1}{\pi}\int_0^\pi |\sin ns|\left(\frac{1}{\sin\frac{1}{2}s} - \frac{2}{s}\right)ds + O(1)$$

on obtient que

$$\|\Phi_n\| = \frac{2}{\pi}\int_0^\pi \frac{|\sin ns|}{s}\,ds + O(1)$$

Maintenant,

$$\frac{2}{\pi}\int_0^\pi \frac{|\sin ns|}{s}\,ds = \frac{2}{\pi}\int_0^{n\pi} \frac{|\sin t|}{t}\,dt = \frac{2}{\pi}\sum_{k=1}^{n-1}\int_{k\pi}^{(k+1)\pi} \frac{|\sin t|}{t}\,dt$$

$$= \frac{2}{\pi}\sum_{k=1}^{n-1}\int_{k\pi}^{(k+1)\pi} \frac{|\sin t|}{t}\,dt = \frac{2}{\pi}\sum_{k=1}^{n-1}\int_0^\pi \frac{|\sin t|}{t+k\pi}\,dt = \frac{2}{\pi}\int_0^\pi\left(\sum_{k=1}^{n-1}\frac{1}{t+k\pi}\right)|\sin t|\,dt$$

Comme pour $t \in [o,\pi]$, $k\pi \leq s + k\pi \leq (k+1)\pi$ et que $\int_0^\pi |\sin t|\,dt = 2$ on aura

$$\frac{4}{\pi^2}\sum_{k=2}^{n}\frac{1}{k} \leq \frac{2}{\pi}\int_0^\pi \frac{|\sin ns|}{s}\,ds \leq \frac{4}{\pi^2}\sum_{k=1}^{n-1}\frac{1}{k}$$

Ainsi $\frac{2}{\pi}\int_0^\pi \frac{|\sin ns|}{s}\,ds = \frac{4}{\pi^2} + O(1)$
d'où $\|\Phi_n\| = \frac{4}{\pi^2}\ln(n) + O(1)$ lorsque $n \to +\infty$. Ce qui termine la preuve du lemme ∎

Donc

$$\|\Phi_n\| \to \infty \quad \text{comme} n \to \infty. \tag{2.2.1}$$

Par conséquent, $(\Phi_n)_{n\in\mathbb{N}}$ n'est pas une famille uniformément bornée de formes linéaires.

Par le théorème de Banach-Steinhaus, cette famille n'est pas ponctuellement bornée sur un ensemble dense. Cela signifie qu'il existe une fonction $f \in C[-\pi,\pi]$ tel que l'ensemble des nombres $\{\Phi_n(f) : n \in \mathbb{N}\}$ n'est bornée pas. C'est exactement ce que nous voulions montrer. ∎

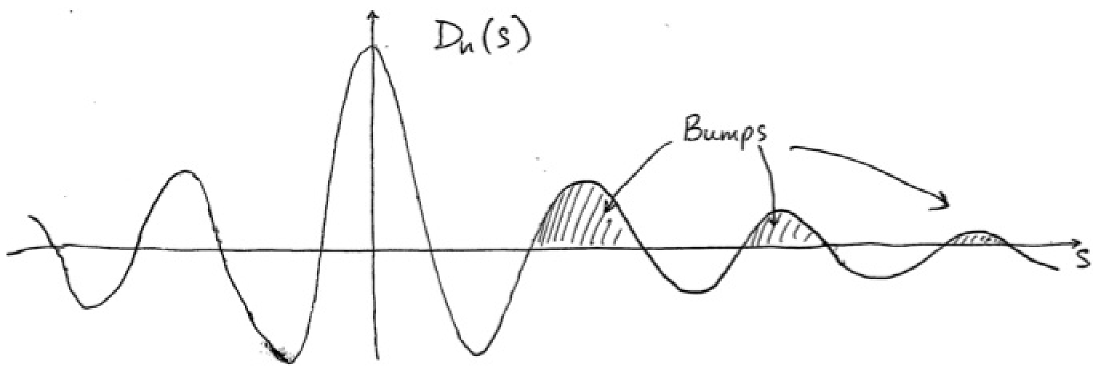


FIGURE 2.1 – noyau de Dirichlet

2.2.13 REMARQUE (CONVERGENCE DES SÉRIES DE FOURIER)

(i) L'analyse de la preuve du théorème 2.2.9 montre que les fonctions continues dont la série de Fourier converge ponctuellement sont rares. Précisément, l'ensemble de ces fonctions est de *maigre* dans $C[-\pi, \pi]$ (C'est à dire qu'il s'agit d'une union dénombrable d'ensembles nulle part denses).

(ii) Néanmoins, pour toute fonction continue f, et même pour $f \in L^p[-\pi, \pi]$, $p > 1$, la série de Fourier converge vers f presque partout. Ainsi, l'ensemble des points de divergence est toujours négligeable. Il s'agit d'un résultat profond de L. Carleson 4.7.9.

(iii) Pour les fonctions de L^1, le résultat de Carleson est généralement faux. Kolmogorov

a construit une fonction $f \in L^1[-\pi, \pi]$ dont la série de Fourier diverge partout.

(iv) A noter que si f est *différentiable* en un point t, alors la série de Fourier de f converge vers $f(t)$. C'est ce qu'on appelle la condition de Dirichlet-Dini.

2.2.14 Exercice 1. Soit f une fonction entière (holomorphe sur $\mathbb{C}$) telle que pour tout $z \in \mathbb{C}$, il existe $n_z \in \mathbb{N}$ avec $f^{(n_z)}(z) = 0$. Montrer que f est polynomiale.

2. Montrer que le résultat est vraie si on remplace fonction entière par fonction réelle C^∞.

3. Soit E un espace de Banach et $T \in \mathcal{L}(E)$ tel que pour tout $x \in E$ il existe $n_x \in \mathbb{N}$ avec $T^{n_x}(x) = 0$. Montrer qu'il existe $n \in \mathbb{N}$ tel que $T^n = 0$.

2.2.3 Bornitude faible et forte

Le principe de la borne uniforme peut être utilisée pour vérifier si un ensemble donné dans un espace de Banach est borné.

2.2.15 COROLLAIRE (BORNÉ FAIBLE ET FORTE)
Soit A un sous-ensemble d'un espace de Banach X. Supposons que A soit *faiblement borné*, c'est à dire

$$\sup_{x \in A} |f(x)| < \infty \quad \text{pour tout } f \in X'.$$

Alors A est (fortement) borné, à savoir

$$\sup_{x \in A} \|x\| < \infty.$$

Ici encore, la proposition réciproque est trivialement vraie - fortement borné implique trivialement faiblement borné.

Démonstration: On identifie X a une partie de X'' en utilisant le plongement canonique. Donc, nous considérons les vecteurs $x \in A$ comme des formes linéaires continues sur X' agissant comme $x(f) := f(x)$, $f \in X'$. La réecriture de l'hypothèse faiblement borné donne que $\sup_{x \in A} |x(f)| < \infty$ pour $f \in X'$, on peut comprendre que cette hypothèse borné ponctuelle de la famille $A \subseteq X'' = L(X', \mathbb{R})$. Le principe de la borne uniforme implique que $\sup_{x \in A} \|x\|_{X''} = \sup_{x \in A} \|x\|_X < \infty$, comme souhaité. ∎

2.2.17 REMARQUE
En utilisant le corollaire 2.2.15, on peut affaiblir l'hypothèse (2.2.5) dans le principe de la borne uniforme par :

$$\sup_{T \in \mathcal{T}} |f(Tx)| < \infty \quad \text{pour tout } x \in X,\ f \in Y'.$$

2.2.18 REMARQUE (PRINCIPE DE CONDENSATION DES SINGULARITÉS)
Banach et Steinhaus ont appelé leur théorème 2.2.1 *le principe de la condensation des singularités* pour la raison suivante. Supposons qu'une famille $\mathcal{T} \subseteq L(X, Y)$ ne soit pas uniformément bornée. Cela signifie que l'ensemble des vecteurs $\{Tx : x \in B_X, T \in \mathcal{T}\}$ n'est pas borné. Le théorème 2.2.1 stipule que $\mathcal{T}$ n'est même pas ponctuellement bornée, donc il existe *un vecteur* $x \in X$ telle que la trajectoire $\{Tx : T \in \mathcal{T}\}$ est sans limite. On peut dire que la non bornitude uniforme de la famille $\mathcal{T}$ est condensée en une "singularité" x.

2.3 Le théorème de l'application ouverte

2.3.1 THÉORÈME (THÉORÈME DE L'APPLICATION OUVERTE)
Soit X, Y des espaces de Banach. Alors, tout opérateur linéaire surjectif $T \in L(X, Y)$ est une *application ouverte*, i.e. T envoie les ouverts dans X sur les ouverts de Y.

La preuve du théorème de l'application ouverte repose sur le théorème de Baire, qui stipule que tout espace métrique complet M n'est par un ensemble maigre i.e. M ne peut pas être représenté par une réunion dénombrable d'ensembles nulle part denses. Rappelons qu'un sous-ensemble $A \subseteq M$ est nulle part dense s'il n'y a pas de boule dans X dans laquelle A est dense. De façon équivalente, A est nulle part dense si l'intérieur de l'adhérence de A est vide.

Le théorème de l'application ouverte dit que pour tout ouvert $U \subseteq X$, tout $y \in TU$ est un point intérieur de TU. Nous affirmons qu'il suffit de montrer cela lorsque U est la boule unité B_X et pour $y = 0$:

2.3.2 LEMME
Pour prouver le théorème de l'application ouverte, il suffit de trouver $\varepsilon > 0$ tel que

$$TB_X \supseteq \varepsilon B_Y. \tag{2.3.1}$$

Démonstration: (Preuve du lemme) Soit $U \subseteq X$ un ouvert et choisissons $y \in TU$. Nous trouvons $x \in U$ tel que $y = Tx$. Comme U est ouvert, il existe $\delta > 0$ tel que

$$U \supseteq x + \delta B_X.$$

On applique T aux deux parties et en utilisant (2.3.1), nous concluons que

$$TU \supseteq T(x + \delta B_X) = y + \delta TB_X \supseteq y + \delta\varepsilon B_Y,$$

donc y est un point intérieur de TU. ■

Nous allons maintenant prouver le lemme. En vue de l'application du théorème de Baire, nous représentons

$$X = \bigcup_{n \in \mathbb{N}} nB_X.$$

Donc

$$Y = TX = \bigcup_{n \in \mathbb{N}} nTB_X.$$

Par le théorème de Baire, il existe $n \in N$ tel que nTB_X n'est pas un ensemble nulle part dense. Ainsi TB_X n'est pas un ensemble nulle part dense, c'est à dire l'adhérence de son intérieur est non vide. Donc il existe $y \in Y$ et $\varepsilon > 0$ tel que

$$\overline{TB_X} \supseteq y + \varepsilon B_Y.$$

Par symétrie, $\overline{TB_X} \supseteq -y + \varepsilon B_Y$. Donc, par convexité (à vérifier !), nous avons

$$\overline{TB_X} \supseteq \varepsilon B_Y.$$

Comme nous le voyons, nous avons presque prouvé l'assertion, mais pour l'adhérence. Malheureusement, en général $\overline{K} \supseteq D$ n'implique pas $K \supseteq D$, même pour d'ensembles convexes et symétriques K, D dans un espace de Banach (donner un contre-exemple !). Toutefois, cela est vrai pour *d'ensembles parfaitement convexes*, définis comme suit :

2.3.4 Définition (Ensemble parfaitement convexe)
Un ensemble K dans un espace de Banach X est *parfaitement convexe* si pour toute suite $(x_k)_{k=1}^{\infty}$ de K et tous nombres $\lambda_k \geq 0$ tels que $\sum_{k=1}^{\infty} \lambda_k = 1$, on a $\sum_{k=1}^{\infty} \lambda_k x_k \in K$.

Un ensemble convexe satisfait cette propriété uniquement pour les *suites finies* (x_k). Par conséquent, chaque ensemble parfaitement convexe est convexe, mais la réciproque est fausse. (Donnez un exemple.)

2.3.5 Lemme
Soit K un ensemble parfaitement convexe dans un espace de Banach Y.
Si $\varepsilon B_Y \subseteq \overline{K}$ pour $\varepsilon > 0$, alors $\frac{\varepsilon}{2} B_Y \subseteq K$.

Démonstration: Supposons $B := \varepsilon B_Y \subseteq \overline{K}$, on veut montrer que $\frac{1}{2}B \subseteq K$. L'hypothèse entraîne que pour tout $y \in B$, $y + \frac{1}{2}B \cap K \neq \emptyset$ donc $y \in K + \frac{1}{2}B$ i.e.

$$B \subseteq K + \frac{1}{2}B,$$

Par Itération, on trouve

$$\begin{aligned} B &\subseteq K + \frac{1}{2}\big(K + \frac{1}{2}B\big) = K + \frac{1}{2}K + \frac{1}{4}B \\ &\subseteq K + \frac{1}{2}K + \frac{1}{4}\big(K + \frac{1}{2}B\big) = K + \frac{1}{2}K + \frac{1}{4}K + \frac{1}{8}\,B \cdots \\ &\subseteq K + \frac{1}{2}K + \frac{1}{4}K + \cdots + \frac{1}{2^n}K + \cdots \end{aligned}$$

Par conséquent [1]

$$\frac{1}{2}B \subseteq \frac{1}{2}K + \frac{1}{4}K + \frac{1}{8}K + \cdots + \frac{1}{2^n}K$$

1. Les sommes et séries en jeux sont dans le sens de Minkowski. La somme d'ensembles est définie par $\sum_k A_k = \{\sum_k a_k : \ a_k \in A_k\}$. De même pour les sommes infinies (série), où nous insistons sur la convergence de $\sum_k a_k$.

Comme K est parfaitement convexe

$$\frac{1}{2}K + \frac{1}{4}K + \frac{1}{8}K + \frac{1}{16}K + \cdots \subseteq K.$$

D'où $\frac{1}{2}B \subseteq K$, ce qui termine la preuve le lemme. ■

2.3.7 EXEMPLE. Les étapes consistant à vérifier la preuve ci-dessus où l'on utilise les sommes de Minkowski et des séries d'ensembles.

Maintenant, nous sommes prêts à achever la preuve du théorème de l'application ouverte. Par le lemme 2.3.5, il suffit de montrer que $K = TB_X$ est parfaitement convexe. C'est facile à vérifier. En effet, considérons une suite $(Tx_k) \subseteq TB_X$ avec $x_k \in B_X$, et le nombre λ_k tel que $\sum_k \lambda_k = 1$. Alors

$$\sum_k \lambda_k T x_k = T\big(\sum_k \lambda_k x_k\big), \tag{2.3.2}$$

à condition que la série $\sum_k \lambda_k x_k$ converge. En fait elle converge absolument :

$$\Big\| \sum_k \lambda_k x_k \Big\| \leq \sum_k \lambda_k \|x_k\| \leq \sum_k \lambda_k = 1.$$

Par la complétude de X, la série $\sum_k \lambda_k x_k$ converge vers un vecteur dans B_X. Il s'ensuit que le membre de droite de (2.3.2) dans l'image appartient à $TB_X = K$, comme on voulait. Ceci termine la démonstration du théorème de l'application ouverte.

2.3.1 Le théorème d'isomorphisme de Banach

Comme conséquence immédiate du théorème de l'application ouverte, on a :

2.3.8 THÉORÈME (LE THÉORÈME D'ISOMORPHISME DE BANACH)
Soit X, Y des espaces de Banach. Alors, tout opérateur linéaire bijectif $T \in L(X, Y)$ est un isomorphisme, donc $T^{-1} \in L(Y, X)$.

Démonstration: Les états ouvert théorème de l'application que les préimages d'ensembles ouverts de moins de T^{-1} sont ouverte, d'où T^{-1} est une application continue. ■

Le théorème d'isomorphisme de Banach est souvent utilisé pour établir la stabilité de solutions des équations linéaires. Soit une équation linéaire en x dans l'espace de Banach X :

$$Tx = b \tag{2.3.3}$$

avec $T \in L(X)$ et $b \in X$.

Supposons que pour chaque $b \in X$ il existe et est unique solution $x \in X$. Alors, par le théorème d'isomorphisme de Banach, la solution $x = x(b)$ est *continue* par rapport à b. En d'autres termes, la solution est stable par des perturbations du membre de droite de (2.3.3).

Dans le cas où T n'est pas injective (mais surjective), on peut encore appliquer le théorème d'isomorphisme de Banach mais à l'injectivisation de T :

2.3.10 COROLLAIRE (LES OPÉRATEURS SURJECTIFS SONT DES APPLICATIONS QUOTIENTS)
Soit X, Y des espaces de Banach. Alors, chaque opérateur linéaire surjective $T \in L(X,Y)$ est la composée d'une application quotient et d'un isomorphisme. Plus précisément,

$$T = \tilde{T} \circ p,$$

où $p : X \to X/\ker T$ est l'application quotient, $\tilde{T} : X/\ker T \to Y$ est un isomorphisme.

Démonstration: Soit $\tilde{T}$ est l'injectivisation de T construit dans l'exemple 1.2.25. Par construction, $T = \tilde{T} \circ p$ et $\tilde{T}$ est injectif. Comme T est surjectif, $\tilde{T}$ est aussi surjectif. Par conséquent, par le théorème d'isomorphisme, $\tilde{T}$ est un isomorphisme, ce qui complète la preuve. ■

2.3.2 Normes équivalentes

Parfois, on veut considérer plusieurs normes sur le même espace, par exemple $\|\cdot\|_\infty$ et $\|\cdot\|_1$ sur $C[0,1]$. Dans ce cas, il faut naturellement les comparer.

2.3.12 DÉFINITION (NORMES ÉQUIVALENTES)
Deux normes $\|\cdot\|$ et $|||\cdot|||$ sur un espace vectoriel E sont dites *équivalentes* s'il existe $C, c > 0$ tel que

$$c\|x\| \leq |||x||| \leq C\|x\| \quad \text{pour tout } x \in E.$$

En d'autres termes, les normes $\|\cdot\|$ et $|||\cdot|||$ sont équivalentes si l'identité

$$\begin{array}{rccc} I : & (E, \|\cdot\|) & \to & (E, |||\cdot|||) \\ & x & \mapsto & x \end{array} \tag{2.3.4}$$

est un isomorphisme.

2.3.13 Exercice Montrer que deux normes sur un espace vectoriel E sont équivalentes si et seulement si elles définissent la même topologie sur E.

2.3.14 PROPOSITION (DOMINATION ET L'ÉQUIVALENCE DES NORMES) Considérons deux normes $\|\cdot\|$ et $|||\cdot|||$ sur un de Banach E par rapport aux deux normes. Supposons que la première norme *domine l'autre, c'est à dire qu'on peut trouver C tel que*

$$|||x||| \leq C\|x\| \quad \text{pour tout } x \in E.$$

Alors, les deux normes sont équivalentes.

Démonstration: L'affirmation découle du théorème d'isomorphisme de Banach appliqué à l'identité (2.3.4). ■

La proposition 2.3.14 indique qu'il est difficile de construire de bonnes normes non équivalentes sur le même espace. Soit l'espace sera non complet ou les normes ne sont pas comparables. Il s'agit d'un moyen simple de prouver la non complétude des espaces.

A titre d'exemple, on considère les normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ sur $C[0,1]$. D'une part, $\|\cdot\|_1 \leq \|\cdot\|_\infty$. D'autre part, les normes ne sont pas équivalentes - on peut facilement construire une suite de fonctions f_n telles que $\|f_n\|_1 = 1$ et $\|f_n\|_\infty \geq n$ est arbitrairement grand. Par la proposition 2.3.14, $C[0,1]$ ne doit pas être complet pour l'une de ces normes. Comme il est complet pour la norme (naturelle) $\|\cdot\|_\infty$, il s'ensuit que $C[0,1]$ n'est pas complet par rapport à $\|\cdot\|_1$.

Cet argument est flexible et s'applique à toute une gamme de normes. Cela implique qu'il n'y a en pratique qu'une seule norme naturelle sur $C[0,1]$, à savoir la somme norme du sup $\|\cdot\|_\infty$.

2.3.16 Exercice (Produits directes d'espaces normés) Soit X et Y des espaces normés et $p \in [1,\infty]$. Définissez $X \oplus_p Y$ comme le produit cartésien $X \times Y$ muni de la norme

$$\|(x,y)\|_p := (\|x\|^p + \|y\|^p)^{1/p} \text{ si } p < \infty, \quad \|(x,y)\|_\infty := \max(\|x\| \; \|y\|) \text{ si } p = \infty.$$

Montrer que $X \oplus_p Y$ est un espace vectoriel normé, et que les normes $\|(x,y)\|_p$, $p \in [1,\infty]$, sont toutes équivalentes.

Pour cette raison, l'indice p est habituellement omis, et l'espace $X \oplus Y$ est appelé la *somme directe* de X et Y.

2.3.3 Plongements isomorphes

Comme nous le savons, le noyau de chaque opérateur $T \in L(X, Y)$ est toujours un sous-espace fermé. Mais, l'image de T n'est pas toujours fermée. Le résultat suivant caractérise les opérateurs dont les images sont fermées.

2.3.17 PROPOSITION (PLONGEMENTS ISOMORPHES)
Soit $T \in L(X, Y)$ un opérateur entre espaces de Banach X et Y. Alors les conditions suivantes sont équivalentes :

(i) T est un *plongement isomorphe, i.e. T est un isomorphisme entre les espaces X et* $\operatorname{Im} T \subseteq Y$ *;*

(ii) T est injective et a une image fermée *;*

(iii) T est minoré, *i.e. il existe $c > 0$ tel que*

$$\|Tx\| \geq c\|x\| \quad \textit{pour tout } x \in X.$$

Démonstration: (i) $\Rightarrow$ (ii). Rappelons qu'un isomorphisme préserve la complétude des espaces. Comme X est complet, il s'ensuit que le sous-espace $\operatorname{Im} T \subseteq Y$ est complet. D'après l'exercice 1.5.2, $\operatorname{Im} T$ est fermé. L'injectivité de T est une conséquence directe de la définition de plongement isomorphe.

(ii) $\Rightarrow$ (iii). Considérant T comme un opérateur de X sur $\operatorname{Im} T$ on voit que T est injectif et surjectif. Par le théorème d'isomorphisme de Banach, T est un isomorphisme.

(iii) $\Rightarrow$ (i). Nous avons

$$c\|x\| \leq \|Tx\| \leq C\|x\| \quad \text{pour tout } x \in X$$

où $C = \|T\|$. Il en résulte que T est un plongement isomorphe. ∎

2.3.4 Espaces normés de dimension finie

Les espaces normés de dimension finie sont les exemples les plus simples d'espaces de Banach. Comme nous allons le voir maintenant, ils sont tous complets, tous isomorphes (mais pas nécessairement isométrique) entre eux l'autre dès qu'ils ont la même dimension, leurs sous-espaces sont fermés et tous les opérateurs linéaires définis sur eux sont bornés.

2.3.19 THÉORÈME
Tout espace normé de dimension n, X est isomorphe à l'espace euclidien ℓ_n^2. Par conséquent, tous les espaces normés de dimension n sont isomorphes.

Démonstration: Nous construisons un isomorphisme entre X et ℓ_n^2 qui identifie une base de X avec la base canonique de ℓ_n^2. Pour cela, considérons une base $u_1, \ldots, u_n$ de X, et on définit l'opérateur $T : \ell_n^2 \to X$ par

$$Tx = \sum_{k=1}^{n} x_k u_k \quad \text{pour } x = (x_1, \ldots, x_n) \in \ell_n^2.$$

La continuité de de T découle des inégalités triangulaires et Cauchy-Schwarz :

$$\|Tx\| \leq \sum_{k=1}^{n} |x_k| \, \|u_k\| \leq \Big(\sum_{k=1}^{n} |x_k|^2 \Big)^{1/2} \Big(\sum_{k=1}^{n} \|u_k\|^2 \Big)^{1/2} = M\|x\|,$$

où $M = \big(\sum_{k=1}^{n} \|u_k\|^2 \big)^{1/2}$.

De plus, T est bijectif, car (u_k) est une base de X. Par le théorème d'isomorphisme de Banach, T est un isomorphisme. ∎

2.3.21 **Exercice** Au lieu d'utiliser le théorème d'isomorphisme de Banach, donner une preuve basée sur la compacité de la sphère unité de ℓ_n^2 .

2.3.22 **Exercice** Deux espaces normés de dimension n, X et Y sont isomorphes mais pas nécessairement *isométriques. Montrer que Y et X sont isométriques si et seulement si $B_Y = TB_X$ pour une certaine application linéaire inversible $T : X \to Y$. Montrer par exemple que ℓ_2^2 n'est pas isométrique à ℓ_2^∞.*

2.3.23 **Remarque (La distance de Banach-Mazur)**

Une notion quantitative de l'isomorphisme est donnée par la *distance de Banach-Mazur $d(X,Y)$ entre deux espaces normés isomorphes X et Y. Elle est définie par*

$$d(X,Y) = \inf\{\|T\|, \|T^{-1}\| : T : X \to Y \textit{est un isomorphisme}\}.$$

Un théorème de F. John donne une forme quantitative du théorème d'isomorphisme 2.3.19. Il précise que chaque espace normé de dimension n, X satisfait

$$d(X, \ell_n^2) \leq \sqrt{n}.$$

Il s'ensuit que, pour deux espaces normés de dimension n, X et Y, on a

$$d(X,Y) \leq n.$$

(Pourquoi ?) E. Gluskin a montré en 1981 que cette borne supérieure est asymptotiquement optimale, i.e. il existe une constante $c > 0$ telle que, pour tout $n \in \mathbb{N}$, on peut construire des espaces normés de dimension n ,X_n et Y_n tels que

$$d(X_n, Y_n) \geq cn.$$

2.3.24 COROLLAIRE

(i) Chaque espace vectoriel normé de dimension finie est un espace de Banach.

(ii) Chaque sous-espace de dimension finie d'un espace vectoriel normé est fermé.

(iii) Chaque opérateur linéaire sur un espace normé de dimension finie est bornée.

(iv) Dans un espace normé de dimension finie les normes sont équivalentes.

Démonstration: (i) Comme un espace vectoriel normé de dimension n X est isomorphe à l'espace complet ℓ_n^2, X est lui-même complet.

(ii) Soit Y un sous-espace de dimension finie d'un espace vectoriel normé X. Alors Y est complet d'après (i). Donc Y est fermé. (En effet, si une suite $(y_n) \subset Y$ converge vers $x \in X$ alors (y_n) est de Cauchy dans Y, d'où sa limite x est dans Y.)

(iii) Chaque opérateur linéaire sur ℓ_n^2 est continu, $\|T\| \leq \|T\|_{\mathrm{HS}}$. Comme un espace vectoriel normé de dim ension n X est isomorphe ℓ_n^2, la même chose est vraie pour les opérateurs linéaires sur X.

(iv) Soient $\|\cdot\|$ et $|||\cdot|||$ deux normes sur X. En partie (iii), l'application d'identité $I : (X, \|\cdot\|) \to (X, |||\cdot|||)$ ainsi que son inverse sont bornés. Ceci termine la preuve. ∎

2.3.26 Exercice (Opérateurs de espaces de dimension finie) Soit X et Y des espaces normés et X est de dimension finie. Montrer que tout opérateur linéaire $T : X \to Y$ est borné. (Indication : identifier X avec ℓ_n^2 par un certain isomorphisme, montrer que T est borné en utilisant un argument similaire à celui du théorème 2.3.19.)

2.4 Le théorème du graphe fermé

Le théorème du graphe fermé est une autre façon de vérifier si un opérateur linéaire donnée est continu. Ce résultat caractérise opérateurs bornés en termes de graphes.

2.4.1 DÉFINITION (GRAPHE D'UN OPÉRATEUR)

Soit $T : X \to Y$ un opérateur linéaire entre espaces vectoriels normés X et Y. Le *graphe* de T est le sous-ensemble suivant de la somme directe[2] $X \oplus_1 Y$:

$$\Gamma(T) = \{(x, Tx) : x \in X\}.$$

2. La somme directe d'espaces normés est étudiée dans l'exercice 2.3.16. Au lieu de $X \oplus_1 Y$, on peut choisir de travailler avec $X \oplus_p Y$ pour tout $p \in [1, \infty]$. Comme nous le savons ceci définit la même topologie.

Il est clair que $\Gamma(T)$ est un sous-espace vectoriel de l'espace vectoriel normé $X \oplus_1 Y$. Le résultat principal de cette section est que $\Gamma(T)$ est fermé si et seulement si T est borné.

Que l'on compare ces deux notions, borné (de façon équivalente, la continuité) de T et celle du graphe fermé. T est continue si et seulement si

$$x_n \to x \in X \quad \text{entraîne} \quad Tx_n \to Tx. \tag{2.4.1}$$

En revanche, $\Gamma(T)$ est fermé si et seulement si

$$x_n \to x \in X \text{ et } Tx_n \to y \in Y \quad \text{alors} \quad y = Tx. \tag{2.4.2}$$

Il ressort clairement de ces deux lignes que la continuité implique toujours que le graphe est fermé :

2.4.2 Proposition

Soit $T : X \to Y$ un opérateur linéaire entre espaces vectoriels normés X et Y. Si T est continu alors $\Gamma(T)$ est fermé.

La réciproque est non triviale et nécessite la complétude des deux espaces X et Y :

2.4.3 Théorème (Le théorème du graphe fermé)

Soit $T : X \to Y$ un opérateur linéaire entre espaces de Banach X et Y. Si $\Gamma(T)$ est fermé alors T est borné.

Démonstration: La somme directe $X \oplus_1 Y$ est un espace de Banach (Exercice **??**). Le graphe $\Gamma(T)$ est un sous-espace vectoriel fermé de $X \oplus_1 Y$, par conséquent, $\Gamma(T)$ est un espace de Banach.

Considérons l'opérateur linéaire

$$u : \Gamma(T) \to X, \quad u(x, Tx) := x.$$

Alors u est un opérateur linéaire borné, surjectif et injectif entre deux espaces de Banach. Par le théorème d'isomorphisme de Banach, u^{-1} est bornée. Cela signifie qu'il existe un certain nombre M tel que

$$\|X\| + \|Tx\| = \|(x, Tx)\| \leq M\|x\| \quad \text{pour tout } x \in X.$$

L'inégalité $\|Tx\| \leq M\|x\|$ implique que T est bornée. ∎

2.4.1 Interprétation et un exemple

———————————————————-

Rappelant l'interprétation de la continuité (2.4.1) et la propriété du graphe fermé (2.4.2). Donc, pour vérifier la continuité d'un opérateur linéaire T en utilisant la définition de (2.4.1) , *on peut toujours supposer que Tx_n converge* . La vérification de la continuité ne nécessite plus que la limite existe, elle se réduit à vérifier la *cohérence des limites de x_n et Tx_n.*

A titre d'exemple, considérons le plus simple opérateur différentiel

$$T = \frac{d}{dt}, \qquad T : C^1[0,1] \to C[0,1]$$

où $C^1[0,1]$ est considéré comme un sous-espace de $C[0,1]$, c'est à dire par rapport à la norme uniforme.

2.4.5 LEMME
L'opérateur différentiel T a un graphe fermé.

Démonstration: Soit $f_n \to f$ et $Tf_n = f_n' \to g$ dans $C[0,1]$, i.e. la convergence est uniforme. Alors, par le théorème de dérivation à la limite,[3] $(\lim_n f_n)' = \lim_n f_n'$, c'est à dire que $g = Tf$. Ceci termine la preuve. ■

Néanmoins, comme nous le savons l'opérateur différentiel n'est pas continue, par exemple $f_n(x) = \frac{\sin(nx)}{n}$ converge uniformeément vers 0, mais f_n' ne converge même pas simplement. Cela ne contredit pas le théorème du graphe fermé, car $C^1[0,1]$ n'est pas complet par rapport à la norme-sup. Si l'on considère $C^1[0,1]$ muni de la norme naturelle $\|f\|_\infty + \|f'\|_\infty$, il est alors un espace de Banach (Exercice 1.6.15), alors l'opérateur différentiel sera évidemment continu.

2.4.2 opérateurs symétriques sur les espaces de Hilbert

———————————————————-

Une application remarquable du théorème du graphe fermé est que la propriété de symétrie d'un opérateur implique toujours borné :

2.4.7 THÉORÈME (HELLINGER-TOEPLITZ)
Soit $T : H \to H$ un opérateur linéaire sur un espace de Hilbert H. Supposons que

$$\langle x, Tx \rangle = \langle Tx, y \rangle \quad \text{pour tout } x, y \in H. \tag{2.4.3}$$

Alors T est bornée.

3. Le théorème sur la dérivation de la limite on a que $(\lim_n f_n)' = \lim_n f_n'$ sous la condition que f_n' converge uniformément et que $f_n(t_0)$ converge pour un certain point t_0.

Démonstration: Par le théorème du graphe fermé, il suffit de vérifier que le graphe de T est fermé. pour cela, nous choisissons suites convergentes $x_n \to x$, $Tx_n \to y$ dans H. Nous voudrions montrer que $y = Tx$. Il suffit de montrer que

$$\langle z, y \rangle = \langle z, Tx \rangle \quad \text{pour tout } z \in H.$$

(Pourquoi ?) Ce qui suit à l'aide de la continuité du produit scalaire et (2.4.3) :

$$\langle z, y \rangle = \lim_n \langle z, Tx_n \rangle = \lim_n \langle Tz, x_n \rangle = \langle Tz, x \rangle = \langle z, Tx \rangle.$$

La preuve est terminée. ■

Le théorème de Hellinger-Toeplitz identifie la source de certaines difficultés en physique mathématique. De nombreux opérateurs naturels, telles que la différentielle, satisfont la condition de symétrie (2.4.3) mais ne sont pas bornés. Le théorème de Hellinger-Toeplitz affirme que ces opérateurs ne peuvent pas être définis partout *sur l'espace de Hilbert. Par exemple, on ne peut jamais définir une notion interressante de différentielle qui rendrait toutes les fonctions de L^2 dérivables.*

C'est ce qui explique que travailler avec un opérateur non borné on doit toujours faire attention à leurs domaines de définition. Par exemple, un opérateur linéaire T sur un espace de Hilbert H est appelé symétrique *si le domaine de T est dense dans H, et la (2.4.3) est satisfaite. Un exemple d'un tel opérateur symétrique est l'opérateur différentiel sur $L^2[0,1]$*

$$T = i\frac{d}{dt}$$

avec pour domaine

$$\operatorname{Dom} T = \{f \in L^2[0,1] : \ f \in C^1[0,1],\ f(0) = f(1) = 1\}.$$

2.5 Espaces séparables

2.5.1 Définition

Soit (E, d) un espace métrique.

On dit que E est **séparable** s'il existe une partie dénombrable et dense $\mathbb{D}$ de E i.e toute boule de E rencontre $\mathbb{D}$.

2.5.2 Exercice Tout sous-espace d'un espace métrique séparable est séparable.

2.5.3 EXEMPLE. Les espaces $\mathbb{R}$ et $\mathbb{C}$ sont séparables.
$\mathbb{D} = \mathbb{Q}$ pour le premier et $\mathbb{D} = \mathbb{Q} + i\mathbb{Q}$ pour le second.

2.5.4 EXEMPLE. L'espace euclidienne $\mathbb{R}$ est séparable : un exemple de sous-ensemble dense et dénombrable est donné par $\mathbb{Q}^n$, qui consiste de tous les points dont toutes les coordonnées sont rationnelles. ▲

2.5.5 EXEMPLE. Plus généralement, tout espace normé V de dimension finie est séparable sur $\mathbb{K}$. Soit $\mathcal{B}$ une base de V :

$$\mathcal{B} = \{e_1, e_2, \ldots, e_n.\}$$

Notons

$$A = \left\{ \sum_{i=1}^{n} q_i e_i \colon q_i \in \mathbb{D} \right\}.$$

où $\mathbb{D} = \mathbb{Q}$ ou $\mathbb{Q} + i\mathbb{Q}$.
Comme, le sous-ensemble $A \subseteq V$ est en bijection avec $\mathbb{D}^n$, il est donc dénombrable. Au même temps, A est dense dans V. Notons

$$L = \max_{1 \leq i \leq n} \| e_i \|.$$

Soient $x \in V$ et $\varepsilon > 0$. On peut écrire $x = \sum_{i=1}^{n} \lambda_i e_i$. Pour chaque $i = 1, 2, \ldots, n$ choisissons $q_i \in \mathbb{Q}$ de telle manière que

$$|\lambda_i - q_i| < \frac{\varepsilon}{n(L+1)}.$$

Notons

$$a = \sum_{i=1}^{n} q_i e_i \in A.$$

On a

$$\begin{aligned}
\| x - a \| &= \left\| \sum_{i=1}^{n} \lambda_i e_i - \sum_{i=1}^{n} q_i e_i \right\| \\
&\leq \sum_{i=1}^{n} \| \lambda_i e_i - q_i e_i \| \\
&= \sum_{i=1}^{n} |\lambda_i - q_i| \| e_i \| \\
&< n \cdot \frac{\varepsilon}{n(L+1)} \cdot L \\
&< \varepsilon.
\end{aligned}$$

2.5.6 EXEMPLE. Un espace métrique dont la fonction distance, d, prend seulement les valeurs 0 et 1 est séparable si et seulement si X est dénombrable.

L'implication $\Leftarrow$ est triviale. Pour montrer $\Rightarrow$, supposons que X est séparable. Alors il existe un sous-ensemble $A \subseteq X$ dénombrable et dense. Cet ensemble, A, rencontre chaque singleton $\{x\}$, $x \in X$, parce que $\{x\}$ est ouvert et non-vide. On conclut que $A = X$, donc X est dénombrable. ▲

2.5.7 Exercice Tout espace métrique compact est séparable

2.5.8 EXEMPLE.
1. Pour tout $1 \leq p < +\infty$, l'espace de Banach $\ell^p(\mathbb{N}) = \{(x_n)_{n\in\mathbb{N}}; \sum_{n\in\mathbb{N}} |x_n|^p < +\infty\}$ est séparable (exercice)
2. Par contre, l'espace de Banach $\ell^\infty(\mathbb{N}) = \{(x_n); \|(x_n)\|_\infty = \sup_n |x_n| < +\infty\}$ n'est pas séparable.

Preuve :

On note, pour tout $A \subset \mathbb{N}$, par χ_A l'élément de $\ell^\infty(\mathbb{N})$ défini par $\chi_A(n) = \begin{cases} 1 & \text{si } n \in A \\ 0 & \text{si } n \notin A \end{cases}$

On remarquera que $A \neq B$ si et seulement si $\|\chi_A - \chi_B\|_\infty = 1$

Maintenant, si on suppose que $\ell^\infty(\mathbb{N})$ est séparable, il existe alors un sous-ensemble dénombrable $D \subset \ell^\infty(\mathbb{N})$ tel que $\overline{D} = \ell^\infty(\mathbb{N})$, en particulier pour tout $A \in \mathcal{P}(\mathbb{N})$ il existe $d_A \in D$ tel que $\|d_A - \chi_A\|_\infty < \frac{1}{2}$, On définit une application $\phi : \mathcal{P}(\mathbb{N}) \to D$ en associant à $A \in \mathcal{P}(\mathbb{N})$ une des suites $d_A \in D$. Cette application est injective, car si $A \neq B$, l'inégalité triangulaire nous donne

$$\|d_A - d_B\|_\infty \geq \|\chi_A - \chi_b\|_\infty - \|d_A - \chi_A\|_\infty - \|d_B - \chi_B\|_\infty > 1 - \frac{1}{2} - \frac{1}{2} = 0$$

et donc $d_A \neq d_B$ i.e. l'application ϕ est injective.

Ainsi, $Card(D) \geq Card\mathcal{P}(\mathbb{N})$ mais comme $\mathcal{P}(\mathbb{N})$ est équipotent à $\mathbb{R}$, ceci contredit l'hypothèse D dénombrable.

2.5.9 EXEMPLE.
1. Si $1 \leq p < \infty$ alors $L^p(X)$ est séparable.
2. Si X est un ouvert non vide, $L^\infty(X)$ n'est pas séparable et les espaces $C_c^k(X)$, $k \in \mathbb{N} \cup \{+\infty\}$, ne sont pas denses dans $L^\infty(X)$.

Démonstration:
1. Si $1 \leq p < \infty$, l'espace $C_c(X)$ des fonctions continues à support compact, est dense dans $L^p(X)$.
   Alors pour tout $f \in L^p(X)$ et $\varepsilon > 0$, il existe $g \in C_c(X)$, telle que $\|f - g\|_p < \varepsilon$
   On utilise le théorème de Stone-Weirstrass, pour approchée la fonction continue à support compact g par une fonction polynomiale P
   i.e. $\|g - P\|_\infty < \varepsilon$.
   Ainsi $\|f - P\|_p \leq \|f - g\|_p + \|g - P\|_p \leq \|f - g\|_p + |g - P\|_\infty \leq 2\varepsilon$. Donc l'ensemble des fonctions polynômiales est dens e dans L^p
   qui est donc une partie totale dénombrable, donc $fL^p(X)$ est séparable.

2. On va maintenant montrer que L^∞ n'est pas séparable.

   En effet, la famille $\{f_r = \mathbb{1}_{B(0,r)}\}_{r\in\mathbb{R}_+}$ est n'est pas dénombrable, contenue dans L^∞ et $\|f_r - f_s\|_\infty = \|\mathbb{1}_{B(0,r)} - \mathbb{1}_{B(0,s)}\|_\infty = 1$ si $r \neq s$.

   S'il existe une suite (g_n) dense dans L^∞, alors pour tout $r > 0$, il existerait $n_r \in \mathbb{N}$ tel que $\|g_{n_r} - f_r\|_\infty < \frac{1}{3}$. Ceci entraîne que l'application $\mathbb{R}_+ \to \mathbb{N}$ $r \to n_r$ est injective i.e. le cardinal de $\mathbb{N}$ est supérieur ou égal à celui de $\mathbb{R}_+$ ce qui est absurde. ∎

2.5.1 Théorème de Stone-Weierstarss

2.5.2 Sous-algèbres fermées de $C^*_{\mathbb{R}}(X)$

2.5.11 PROPOSITION

On considère la suite récurrente $(P_n)_{n\in\mathbb{N}}$ de fonctions polynomiales de $\mathbb{R}$ dans $\mathbb{R}$ ainsi définie :

$$\begin{cases} P_0 = 0 \\ P_{n+1} = P_n + \frac{1}{2}(x - P_n^2) \end{cases}$$

Alors $(P_n)_{n\in\mathbb{N}}$ converge uniformément vers la fonction $x \mapsto \sqrt{x}$ plus précisément : pour tout $x \in [0,1]$, $\quad 0 \leq \sqrt{x} - P_n(x) \leq \frac{2}{n+1}$.

Démonstration: On montre par récurrence sur n que pour tout $x \in [0,1]$, $0 \leq \sqrt{x} - P_n(x) \leq \frac{2\sqrt{x}}{n\sqrt{x}+2}$. ∎

Ainsi, si X est un espace topologique, si $a > 0$ et $f : X \to [0,a]$, alors la suite $\sqrt{a}(P_n \circ \frac{f}{a})_{n\in\mathbb{N}}$ converge uniformément vers $\sqrt{f}$.

En particulier :

2.5.13 COROLLAIRE

Soit X un espace topologique. Toute sous-algèbre fermée de $C^*_{\mathbb{R}}(X)$ (des applications continues et bornées de X dans $\mathbb{R}$) est stable par les opérations $|.|$, $\sqrt{.}$ sup et inf.

Démonstration: En effet, $|f| = \sqrt{f^2}$, $\inf(f,g) = \frac{1}{2}(f + g - |f-g|)$ et $\sup(f,g) = \frac{1}{2}(f + g + |f-g|)$. ∎

2.5.3 Le théorème de Stone-Weierstrass réel

Etant donné un ensemble X, une partie A de $\mathbb{R}^X$ est séparante lorsque pour tous éléments distincts x, y de X, il existe $u \in A$ tel que $u(x) \neq u(y)$. Si A est un sous-espace vectoriel de $\mathbb{R}^X$, séparant et contenant la fonction constante $\mathbb{1}_X$, alors pour tous éléments distincts x, y de X, et tous réels α, β, il existe $u \in A$ tel que $u(x) = \alpha$ et $u(y) = \beta$.

En effet, il existe $h \in A$ tel que $h(x) \neq h(y)$, d'où le système $\begin{cases} \alpha & = \lambda h(x) + \mu \\ \beta & = \lambda h(y) + \mu \end{cases}$ a une solution unique (λ, β). Alors $u = \lambda h + \beta$ convient.

2.5.15 THÉORÈME

Soit K un espace topologique compact. Toute sous-algèbre unitaire et séparante de $C_{\mathbb{R}}(K)$ (algèbre des applications continues de K dans $\mathbb{R}$ munie de la norme uniforme) est dense dans $C_{\mathbb{R}}(K)$.

Démonstration: Soit A l'adhérence dans $C_{\mathbb{R}}(K)$ d'une telle algèbre. On va montrer que A est dense dans $C_{\mathbb{R}}(K)$ (donc égale à $C_{\mathbb{R}}(K)$). Soit $f \in C_{\mathbb{R}}(K)$ et $\varepsilon > 0$. Pour tout élément $x \in K$, notons A_x le sous-ensemble $\{u \in A \,:\, u(x) = f(x)\}$. Pour tous éléments $x, y \in K$, on note $A_{x,y}$ l'ensemble $A_x \cap A_y$. La preuve "usuelle" semble utiliser l'axiome du choix car on commence par "choisir" pour chaque couple $(x, y) \in K \times K$, une fonction arbitraire $u_{x,y} \in A_{x,y}$. Mais on peut éviter le recours à l'axiome du choix en considérant toutes les fonctions $u \in A_{x,y}$ (au lieu d'en choisir une).

$\boxed{\text{Etape 1 :}}$ Pour tout $z \in K$, il existe $g \in A_z$ tel que $f - \varepsilon \leq g$. Soit $z \in K$. On considère le recouvrement ouvert suivant du compact K :

$$K \subset \bigcup_{y \in K,\, u \in A_{y,z}} \{x \in K :\, f(x) - \varepsilon < u(x)\}$$

De ce recouvrement ouvert, on extrait un sous-recouvrement fini : il existe donc une partie finie F de K, et pour chaque $y \in F$, une partie finie G_y de $A_{y,z}$ telles que :

$$K \subset \bigcup_{y \in F} \bigcup_{u \in G_y} \{x \in K :\, f(x) - \varepsilon < u(x)\}$$

Alors la fonction $g := \sup_{y \in F,\, u \in G_y} u$ appartient à A (car la sous-algèbre fermée A est close par "sup fini") et $g \in A_z$. De plus, si $x \in K$, soit $y \in F$ et $u \in A_{y,z}$ tels que $f(x) - \varepsilon < u(x)$. Alors $g(x) \geq u(x) > f(x) - \varepsilon$.

$\boxed{\text{Etape 2 :}}$ il existe $g \in A$ tel que $f - \varepsilon \leq g \leq f + \varepsilon$. On considère maintenant le recouvrement ouvert suivant de K :

$$K \subset \bigcup_{u \in A,\, f - \varepsilon \leq u} \{x \in K :\, u(x) < f(x) + \varepsilon\}.$$

On en extrait un sous-recouvrement fini : soit F une partie finie de $\{u \in A : f - \varepsilon \leq u\}$ telle que

$$K \subset \bigcup_{u \in F} \{x \in K : u(x) < f(x) + \varepsilon\}.$$

Alors la fonction $g := \inf_{u \in F} u$ appartient à A et $f - \varepsilon \leq g$. De plus, si $x \in K$, soit $u \in F$ tel que $u(x) < f(x) + \varepsilon$. Alors $g(x) \leq u(x) < f(x) + \varepsilon$. ∎

2.5.17 EXEMPLE. 1. Soit A l'ensemble des fonctions lipschitziennes de X à valeurs dans $\mathbb{R}$, c'est l'ensemble des fonctions $h : X \to \mathbb{R}$ telle qu'il existe $C > 0$ (dépendant de h) tel que $|h(x) - h(y)| \leq Cd(x, y)$ pour tout $x, y \in X$.

Clairement A est une sous-algèbre uniraire et séparante car pour $x \neq y$, la fonction $u : z \mapsto d(z, x)$ vérifie $u(x) = 0 \neq d(y, x) = u(y)$ et est lipschitzienne. D'après le théorème 2.5.15 A est dense dans $C_{\mathbb{R}}(X)$.

2. Soit X un compact de $\mathbb{R}^d$ et A l'ensemble des fonctions polynomiales (en d variables) de X dans $\mathbb{R}$,

$$A = \mathbb{R}[X_1, \ldots, X_d].$$

Il est clair, que A est une algèbre unitaire. Si $x \neq y$, alors il existe $j \in \{1, \ldots, d\}$ tel que $x_j \neq y_j$, d'où X_j prend des valeurs différente en x et y, donc A est séparante et donc dense dans $C_{\mathbb{R}}(X)$.

Par conséquent, la suite de monômes $\{1, X, X^2, \ldots, X^n, \ldots\}$ est dénombrable et totale d'où $C_{\mathbb{R}}(X)$ est séparable.

2.5.18 REMARQUE

Dans le théorème précédent on ne peut pas remplacer $C_{\mathbb{R}}(X)$ par $C_{\mathbb{C}}(X)$. Soit $\mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$. L'ensemble A des fonctions polynomiales de $\mathbb{U}$ dans $\mathbb{C}$ est une algèbre unitaire et séparante, mais n'est pas dense dans $C_{\mathbb{C}}(\mathbb{U})$. En effet, on a pour tout $n \in \mathbb{N}$, $\int_0^{2\pi} e^{int} e^{it}\, dt = 0$, d'où pour tout $f \in A$, $\int_0^{2\pi} f(e^{it}) e^{it}\, dt = 0$. En prenant la limite uniforme, on aura $\int_0^{2\pi} f(e^{it}) e^{it}\, dt = 0$, pour tout $f \in \bar{A}$. Mais pour la fonction $g : z \mapsto \bar{z}$, on a $\int_0^{2\pi} g(e^{it}) e^{it}\, dt = \int_0^{2\pi} e^{-it} e^{it}\, dt = 2\pi$. Donc $g \notin \bar{A}$ c-à-d A n'est pas dense dans $C_{\mathbb{C}}(\mathbb{U})$.

Une condition supplémentaire est nécessaire dans le cas complexe.

2.5.4 Le théorème de Stone-Weierstrass complexe

2.5.19 THÉORÈME

Soit K un espace topologique compact. Toute sous-algèbre A unitaire et séparante et stable par conjugaison (i.e. $f \in A \Rightarrow \bar{f} \in A$) de $C_{\mathbb{C}}(X)$ (algèbre des applications continues de K dans $\mathbb{C}$ munie de la norme uniforme) est dense dans $C_{\mathbb{C}}(X)$.

Démonstration: Soit $A_{\mathbb{R}} = \{f \in A : f(x) \in \mathbb{R}$, pour tout $x \in X\}$, alors clairement $A_{\mathbb{R}}$ est une sous-algèbre unitaire de $C_{\mathbb{R}}(X)$. Maintenant, si $f \in A$, alors la partie réelle
$\mathfrak{Re}(f) = \frac{f+\bar{f}}{2}$ et $\mathfrak{Im}(f) = \frac{f-\bar{f}}{2i}$ sont des éléments de $A_{\mathbb{R}}$, car A est stable par conjugaison.

Montrons que $A_{\mathbb{R}}$ est séparante. Soit $x \neq y$, comme A est séparante, il existe $f \in A$ tel que $f(x) not= f(y)$, ce qui entraîne que $Re(f)(x) \neq Re(f)(y)$ ou $\mathfrak{Im}(f)(x) \neq \mathfrak{Im}(f)(y)$.

D'après Stone-Weierstrass réel, est dense dans $A_{\mathbb{R}}$ dans $C_{\mathbb{R}}(X)$. Soit $f \in A$ et $\varepsilon > 0$, alors il existe f_1 et f_2 dans $A_{\mathbb{R}}$ telles que $\|Re(f) - f_1\|_\infty < \varepsilon$ et $\|Im(f) - f_2\|_\infty < \varepsilon$, ainsi $\|f - (f_1 + i f_2)\|_\infty < 2\varepsilon$ ce qu'il fallait démontrer. ■

2.5.21 EXEMPLE.
1. Soit A l'ensemble des fonctions lipschitziennes de X à valeurs dans $\mathbb{C}$, le même que dans le cas réel et en remarquant la stabilité de A par conjugaison, d'après le théorème 2.5.19 A est dense dans $C_{\mathbb{C}}(X)$.
2. Soit X un compact de $\mathbb{C}^d$ et A l'ensemble des fonctions polynomiales (en z_i et $\bar{z}_i$) de X dans $\mathbb{C}$
$$A = \mathbb{C}[Z_1, \ldots, Z_d, \bar{Z}_1, \ldots, \bar{Z}_d].$$
A est dense dans $C_{\mathbb{C}}(X)$. En particulier, si $d = 1$ et $X = \mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$ on voit que A est un sous-espace engendreé par les fonctions $z \mapsto z^n$, $n \in \mathbb{Z}$. On remarquera que $\bar{Z} = Z^{-1}$.

2.5.5 Bases de Schauder

La notion de base de Hamel, que nous avons étudié dans la section 1.2.4, a un inconvénient. Dans tout espace de Banach de dimension infinie, les bases de Hamel ne sont pas dénombrables. *Il est donc difficile d'utiliser des bases de Hamel dans la pratique. Il existe une notion alternative de base, qui est plus adaptée aux besoins de l'analyse :*

2.5.22 DÉFINITION (BASE DE SCHAUDER)
Une suite $(x_k)_{k=1}^\infty$ dans un espace de Banach X est appelée une *base de Schauder* de X si chaque vecteur $x \in X$ peut être uniquement exprimée en une série convergente
$$x = \sum_{k=1}^{\infty} a_k x_k \tag{2.5.1}$$
pour certains scalaires a_k.

2.5.23 EXEMPLE. Montrer que seuls les espaces séparables peuvent avoir une base de Schauder.

Si (x_k) est une base de Schauder alors la suite (x_k) est clairement linéaire, indépendante et totale Toutefois, la propriété de base est beaucoup plus forte que l'indépendance linéaire et la totalité. La totalité signifie que pour un vecteur arbitraire $x \in X$ et pour un $\varepsilon > 0$ arbitraire , on peut trouver une combinaison linéaire finie d'éléments de la base qui approche x à $\varepsilon > 0$ près i.e. :

$$\Big\|x - \sum_{k=1}^{n} a_k x_k\Big\| \leq \varepsilon. \tag{2.5.2}$$

Toutefois, les coefficients $a_k = a_k(\varepsilon)$ peuvent dépendre de ε. La limite $\lim_{\varepsilon\to 0} a_k(\varepsilon)$ n'existe pas en général, comme on le verra dans l'exercice suivant. En revanche, la propriété de base garantie que l'on peut approchement (2.5.2) avec les mêmes coefficients *a_k indépendant de ε uniquement en augmentant le nombre de termes n dans la combinaison linéaire.*

2.5.24 Exemple (Complétudeet la propriété base). Dans un espace de Hilbert H, trouver une suite totale de vecteurs linéairement indépendants (x_k), qui ne soit pas une base de Schauder. Indication : construire x_k afin qu'elle converge vers un vecteur non nul dans H ; montrer que ces suites ne sont jamais des bases de Schauder.

2.5.25 Exemple (Bases dans les espaces de suites). Une base orthogonale d'un espace de Hilbert est une base de Schauder. (Vérifiez l'unicité de la represenation de x !)

Dans les espaces de suites ℓ^p, $1 \leq p < \infty$ et c_0, la suite canonique (e_n) forme une base de Schauder.

Dans ℓ^∞, il n'existe pas de base de Schauder parce que cet espace n'est pas séparable.

2.5.26 Exemple (Base de l'espace des fonctions continues). Sur $C[0,1]$, les candidats naturels ne parviennent pas à être des bases de Schauder. La base de Fourier n'est pas une base de Schauder - sinon cela impliquerait que la série de Fourier de toute fonction continue converge dans $C[0,1]$, ce qui contredirait le théorème 2.2.9.

La suite de monômes $1, t, t^2, \ldots$ n'est pas une base de Schauder $C[0,1]$. (Prouver le !)

La base de Schauder la plus connue de $C[0,1]$ est le *système de Schauder d'ondelette. Son ondelette mère $\phi(t)$ est obtenue par intégration de l'ondelette mère du système de Haar c'est à dire*

$$\phi(t) = \int_0^t h(s)\,ds = \begin{cases} t, \text{ si } t \in [0,1/2) \\ 1-t, \text{ si } t \in [1/2,1) \end{cases}$$

Nous considérons les translatés et dilatés de l'ondelette mère :

$$\phi_{kl}(t) = \phi(2^k t - l), \quad k = 0,1,2,\ldots, \quad l = 0,1,2,\ldots,2^k - 1.$$

Avec la fonction constante 1*, le système de fonctions* $\phi_{kl}(t)$ *est appelé* système de Schauder*. Il constitue une base de Schauder de* $C[0,1]$*.*

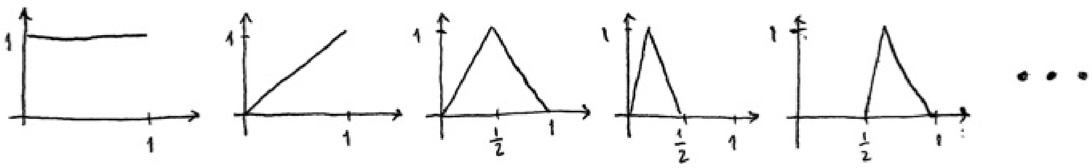


FIGURE 2.2 – Les premières fonctions du système de Schauder

2.5.27 EXEMPLE (VÉRIFICATION DES BASES DE SCHAUDER). Faire toutes les vérifications dans les exemples précédents. Pour $C[0,1]$, on note que l'espace linéaire engendré se compose de fonctions linéaires par morceaux sur $[0,1]$ dont les noeuds sont des points dyadiques.

Une propriété importante des bases de Schauder est la borne uniforme sur les sommes partielles :

2.5.28 THÉORÈME (SOMMES PARTIELLES D'UNE BASE DE SCHAUDER)
Soit (x_k) est une base de Schauder d'un espace de Banach X. Alors il existe un certain nombre M appelé *constante de base de* (x_k) *avec la propriété suivante. Les sommes partielles de* (2.5.1) *dans la base de Schauder de chaque* $x \in X$ *vérifie*

$$\Big\| \sum_{k=1}^{n} \alpha_k x_k \Big\| \leq M\|x\|, \quad n = 1, 2, \dots$$

Démonstration: Considérons l'espace des suites

$$E := \Big\{ a = (a_k)_{k=1}^{\infty} : \sum_{k=1}^{\infty} a_k x_k \text{ converge dans } X \Big\} \tag{2.5.3}$$

muni de la norme

$$\|a\|_E := \sup_{n \in \mathbb{N}} \left\| \sum_{k=1}^{n} a_k x_k \right\|. \tag{2.5.4}$$

Il s'agit d'un exercice de vérifier que E est un espace de Banach. Nous allons montrer que E et X sont isomorphes. [4]

4. Pour un repère orthonormé base (x_k) dans un espace de Hilbert X, cette déclaration signifie que X est isomorphe à $E = \ell^2$. Nous l'avons prouvé dans la section **??**

pour cela, nous considérons la synthèse opérateur $T : E \to X$ définie par

$$Ta = \sum_{k=1}^{\infty} a_k x_k.$$

Par construction, T est un opérateur linéaire borné :

$$\|Ta\| \leq \|a\|_E.$$

Comme (x_k) est une base de Schauder, T est surjective et injective. Par le théorème d'isomorphisme de Banach, T est un isomorphisme. Par conséquent, on peut trouver un certain nombre M de telle sorte que

$$\|a\|_E \leq M\|Ta\|, \quad a \in E.$$

Mais cela signifie que

$$\sup_{n \in \mathbb{N}} \Big\| \sum_{k=1}^{n} a_k x_k \Big\| \leq M\|x\|, \quad x \in X,$$

. ■

2.5.30 EXEMPLE (ESPACE DE COEFFICIENTS). Soit $(x_k)_{k=1}^{\infty}$ une suite de vecteurs non nuls dans un espace de Banach X. Définir l'espace des coefficients E par (2.5.3) et (2.5.4). Montrer que E est un espace de Banach.

Considérons les sommes partielles d'une expansion de base (2.5.1) :

$$S_n(x) := \sum_{k=1}^{n} a_k x_k.$$

D'après le théorème 2.5.28, on voit que S_n est une projection de X sur $\mathrm{Vect}(x_1, \dots, x_n)$, *et que S_n est uniformément bornée (par la constante de base) :*

$$\sup_{n \in \mathbb{N}} \|S_n\| = M < \infty.$$

De plus, les coefficients $a_k = a_k(x)$ de l'expansion de base (2.5.1) sont des formes linéaires sur X qu'ont appelle formes biorthogonales *de la base (x_k) et notées x_k^*, c'est à dire*

$$x_k^*(x) = a_k.$$

Avec cette notation, l'expansion dans la base de $x \in X$ est

$$x = \sum_{k=1}^{\infty} x_k^*(x) x_k$$

Cela ressemble à la série de Fourier par rapport à des bases orthogonales dans un espace de Hilbert, sauf que maintenant nous somme dans des espaces de Banach généraux.

2.5.31 COROLLAIRE (FORMES BIORTHOGONALES)
Les formes biorthogonales (x_k^*) d'une base de Schauder (x_k) sont uniformément bornées :

$$\sup_{k \in \mathbb{N}} \|x_k^*\| \, \|x_k\| < \infty.$$

Démonstration: Nous estimons que le$k$-ième terme de l'expansion

$$\|x_k^*(x)x_k\| = \|S_k(x) - S_{k-1}(x)\| \leq \|S_k(x)\| + \|S_{k-1}(x)\| \leq 2M\|x\|$$

où M est la constante de base. D'autre part, $\|x_k^*(x)x_k\| = |x_k^*(x)| \, \|x_k\|$. Ceci termine bien la preuve. ■

2.5.33 EXEMPLE. Les formes biorthogonales pour une base orthonormale dans un espace de Hilbert coïncident avec les vecteurs de base, c'est à dire $x_k^* = x_k$. Les formes biorthogonales pour la norme canonique (coordonnées) base de ℓ^p, $1 \leq p < \infty$ sont les vecteurs coordonnés de ℓ^q, où q et p sont des exposants conjugués.

2.5.34 REMARQUE (PROBLÈME DE LA BASE DE BANACH)
Tous les espaces de Banach séparables classiques sont connus pour avoir des bases de Schauder. Cependant, il existe des espaces de Banach séparables qui n'ont pas de bases de Schauder. Les premières constructions de ces espaces ont été données par P.Enflo *A counterexample to the approximation problem in Banach spaces, Acta Math. 130 (1973), 309–317.* comme une réponse négative au problème de base de Banach.

Chapitre 3

Dualité et théorème de Hahn-Banach

3.1 Formes linéaires (fonctionnelles)

3.1.1 Définition et exemples

Pour le moment, la topologie n'a pas d'importance, nous définissons les formes linéaires sur un espace vectoriel.

3.1.1 DÉFINITION (FORMES LINÉAIRES)
Soit E un espace vectoriel sur $\mathbb{C}$ (le cas réel est similaire). Un *forme linéaire* sur E est un opérateur linéaire $f : E \to \mathbb{C}$. De manière équivalente, une fonction $f : E \to \mathbb{C}$ est une forme linéaire si

$$f(ax + by) = af(x) + bf(y) \quad \text{pour tout} x, y \in E, \; a, b \in \mathbb{C}$$

3.1.2 EXEMPLE (INTÉGRATION). L'intégrale d'une fonction intégrable est un exemple de base d'une forme linéaire. Plus précisément, l'application

$$F(g) = \int g(t)\, d\mu$$

est clairement une forme linéaire sur $L^1(\Omega, \Sigma, \mu)$.

De même, pour un poids fixé $w \in L^1[0,1]$, l'application

$$F(g) = \int_0^1 g(t)w(t)\, dt$$

définie une forme linéaire sur $L^\infty[0,1]$.

3.1.3 EXEMPLE (FORME D'ÉVALUATION EN UN POINT). Pour un nombre fixe $t_0 \in [0,1]$, l'application

$$F(g) = g(t_0) \tag{3.1.1}$$

est clairement une forme linéaire sur $C[0,1]$. Elle est appelée la *fonctionnelle d'évaluation en t_0*.

Les physiciens voient la fonction d'évaluation en un point comme un cas particulier de l'intégration avec le poids :

$$g(t_0) = \int_0^1 g(t)\delta(t - t_0)\, dt \tag{3.1.2}$$

Le poids est ici donné par *la fonction de Dirac $\delta(t)$*, qui est nulle pour toutes les valeurs t, sauf pour t_0, $\delta(t_0) = \infty$, et telle que $\int_0^1 \delta(t)\ dt = 1$. La fonction de Dirac n'existe pas en tant que fonction de $[0,1] \to \mathbb{R}$, et devrait être comprise comme une *forme linéaire*. " L'intégrale" doit être comprise comme la fonction d'évaluation en un point (3.1.1).

3.1.4 EXEMPLE (FORMES LINÉAIRES SUR $\mathbb{K}^n$). Montrer que toute forme linéaire f sur $\mathbb{K}^n$ est de la forme

$$f(x) = \sum_{k=1}^{n} x_k y_k = \langle x, y \rangle, \quad x = (x_1, \ldots, x_n),$$

pour $y = (y_1, \ldots, y_n) \in \mathbb{K}^n$.

3.1.2 Continuité et bornitude. Espace dual

Soit $(X, \|.\|)$ un espace vectoriel normé sur $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$).

Le dual (topologique) de X est l'espace $\mathcal{L}(X, \mathbb{K})$ des formes linéaires continues sur X. On le note par X^* ou X'. On définit sur X^* la norme subordonnée à la norme de X.

On notera souvent, par x^* un élément de X^* et $\langle x^*, x \rangle$ l'action de la forme x^* sur le vecteur x i.e.

$$\begin{array}{rcl} x^* : X & \longrightarrow & \mathbb{K} \\ x & \mapsto & \langle x^*, x \rangle := x^*(x) \end{array}$$

De même on défnit le bidual de X, l'espace dual de X^*, noté X'' ou X^{**}.

On a alors, une action d'un élément $x \in X$ sur $x^* \in X^*$, par l'évaluation de la forme x^* au point x :

$$\begin{array}{rcl} x : X^* & \longrightarrow & \mathbb{K} \\ x^* & \mapsto & \langle x, x^* \rangle := x^*(x) \end{array}$$

Ainsi, tout point de $x \in X$ peut être considéré comme une forme linéaire sur X^*, on a donc une application canonique :

*$J_X : X \to X^{**}$ définie par $J_X(x)$ est la forme linéaire sur X^* définie par :
pour tout $x^* \in X^*$ on pose $J_X(x)(x^*) = x^*(x) = \langle x, x^* \rangle$.*

On montrera par la suite, à l'aide du théorème de Hahn-Banach, que J_X est une isométrie, et donc X est isométriquement isomorphe à un sous-espace de son bidual.

3.1.5 Définition (espace Dual)

Soit X un espace vectoriel normé.

L'espace dual est un espace vectoriel normé, muni de la norme subordonnée à la norme de X, définie par :

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)|, \quad f \in X^*.$$

3.1.6 Remarque

La définition implique l'inégalité suivante :

$$|f(x)| \leq \|f\| \, \|x\| \quad \text{pour tout} x \in X, \, f \in X^*.$$

De plus, $\|f\|$ est le plus petit nombre dans cette inégalité qui en fait valable pour tout $x \in X$.

3.1.7 Proposition

Soit E un espace normé réel de dimension finie. Pour tout $x \in E$, $x \neq 0$ il existe une forme linéaire bornée $\phi \colon E \to \mathbb{R}$ telle que $\phi(x) \neq 0$.

Comme on dit, les formes linéaires bornées séparent les éléments *de l'espace E.*

Démonstration: Il existe une base finie, X, de l'espace vectoriel E qui contient x. Définissons ϕ par

$$E \ni y = \sum_{z \in X} \lambda_z z \mapsto \phi(y) = \lambda_x \in \mathbb{R}.$$

Cette forme est évidemment linéaire, donc bornée, et $\phi(x) = 1$. ■

Cet argument ne fonctionne pas pour les espaces de dimension infinie.

3.1.9 Proposition

Tout espace normé de dimension infinie admet au moins une forme linéaire discontinue.

Démonstration: Soit E un espace normé de dimension infinie. Choisissons une base X de E. Alorsque $\dim E = \infty$, la base X est infinie. Elle contient alors une suite

$(x_n)_{n=1}^{\infty}$ d' éléments deux à deux distincts. Définissons une fonction $f\colon X \to \mathbb{R}$ de façon suitante : pour $x \in X$, posons

$$f(x) = \begin{cases} n\| x_n \|, & \text{si } x = x_n, \\ 0, & \text{sinon.} \end{cases}$$

La fonction f admet un prolongement unique, $\tilde{f}$, sur E, par la règle suivante : si $x = \sum_{i=1}^{n} \lambda_i x_i$, où $x_i \in X$ et $\lambda_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$, on pose

$$\tilde{f}\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i f(x_i).$$

Évidemment, la forme $\tilde{f}\colon E \to \mathbb{R}$ est linéaire. De plus, $\tilde{f}$ n'est pas bornée : Alorsque pour tous $n = 1, 2, \ldots$, on a

$$|\tilde{f}(x_n)| = n \cdot \| x_n \|,$$

il n'existe donc aucune valeur $K \geq 0$ telle que pour tous $x \in E$ on ait

$$|\tilde{f}(x)| \leq K \cdot \| x \|.$$

■

3.1.11 Définition (Continuité, bornée)

Soit f une forme linéaire sur un espace vectoriel normé X.

(i) f est *continue* si

$$x_n \to x \in X \quad \text{implique} \quad f(x_n) \to f(x).$$

(ii) f est dite *bornée* s'il existe un nombre C tel que

$$|f(x)| \leq C\|x\| \quad \text{pour tout } x \in X.$$

3.1.12 Proposition

Une forme linéaire est continue si est seulement si elle est bornée

Démonstration: Supposons que f est bornée, et soit $x_n \to x$. Alors

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq C\|x_n - x\| \to 0.$$

Ainsi, f est continue.

Inversement, supposons que f ne soit pas bornée. Alors, nous pouvons trouver une suite (x_n) de vecteurs X (non nuls) telle que

$$|f(x_n)| \geq n\|x_n\|, \quad n = 1, 2, \ldots$$

Divisant les deux côtés par $n\|x_n\|$, nous obtenons

$$\left\|f(\frac{x_n}{n\|x_n\|})\right| \geq 1, \quad n = 1, 2, \ldots$$

D'autre part, $\frac{x_n}{n\|x_n\|} \to 0$ comme la norme de ce vecteur est égale à $1/n$. Cela implique que f n'est pas continue. ■

3.1.14 EXEMPLE. Montrer que si f est continue en un point $x_0 \in X$, alors f est continue (partout sur X).

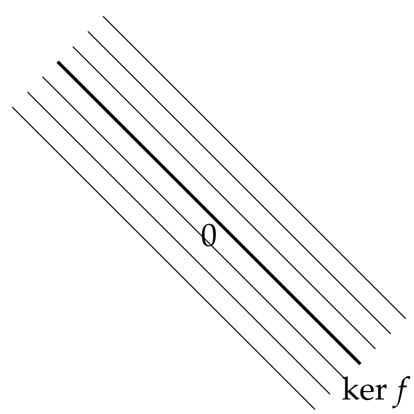
3.1.3 Les hyperplans comme des ensembles de niveau de formes linéaires

Les formes linéaires f, sur un espace vectoriel X vectoriel peuvent être visualisées en décrivant les niveaux

$$\{x \in X : f(x) = c\}$$

pour différentes valeurs $c \in \mathbb{C}$. Le niveau correspondant à $c = 0$ est le noyau de f, noté $\ker f$.

Il s'avère que $\ker f$ *est un* hyperplan, *c'est à dire un sous-espace de X de codimension 1. Tous les autres ensembles de niveaux de f sont évidemment les translatés de* $\ker f$. *De plus, il existe une correspondance canonique entre les formes linéaires et les hyperplans de X. Cela est précisé dans la proposition suivante.*



Ensembles de niveaux d'une forme linéaire f sur un espace vectoriel E

3.1.15 PROPOSITION (FORMES LINÉAIRES ET HYPERPLANS)
Soit E un espace vectoriel.

(i) Pour chaque forme linéaire f sur E, $\ker f$ est un hyperplan dans E, ie $\operatorname{codim}(\ker f) = 1$.

(ii) Si $f, g \neq 0$ sont des formes linéaires sur E telles que $\ker f = \ker g$, alors $f = ag$ pour certains scalaire un $a \neq 0$.

(iii) Pour tout hyperplan $H \subset E$, il existe une forme linéaire $f \neq 0$ tel que $\ker f = H$.

Démonstration: (i) Découle une version linéaire du théorème fondamental sur homomorphismes. En effet, le injectivisation $\tilde{f} : E/\ker f \to \mathbb{C}$ de f établit une bijection linéaire (isomorphisme) entre $E/\ker f$ et $\mathbb{C}$ de f. Ainsi $\dim(E/\ker f) = \dim(\mathbb{C}) = 1$, donc $\ker f$ est un hyperplan dans E.

(ii) Comme $\ker f = \ker g = H$, les injectivisations $\tilde{f}, \tilde{g} : E/H \to \mathbb{C}$ sont linéaires sur l'espace E/H de dimension 1. Ces formes linéaires doivent être égales à un facteur constant a, soit $\tilde{f} = a\tilde{g}$. Alors $\tilde{f}[x] = a\tilde{g}[x]$ pour tout $x \in E$. D'autre part, par la construction de l'injectivisation, $f(x) = \tilde{f}[x]$ et $g(x) = \tilde{g}[x]$. Donc $f(x) = ag(x)$ comme souhaité.

(iii) Comme $\dim(E/H) = 1$, nous avons

$$E/H = \{a[x_0] : \; a \in \mathbb{C}\}$$

pour $x_0 \in E$. Soit $x \in E$ arbitraire, alors $[x] = a[x_0]$ pour $a = a(x) \in \mathbb{C}$, ce qui implique $x = ax_0 + h$ pourt un certain $h \in H$. Nous allons définir f sur E par $f(x) = a$. Alors f est une forme linéaire , et on a bien $\ker f = H$. ■

3.1.17 PROPOSITION
Soit f est un forme linéaire continue, c'est à dire $f \in X'$, alors $\ker f$ est un fermé de X.

Démonstration: $\ker f$ est la pré-image de l'ensemble fermé $\{0\}$ par l'application continue f, donc il doit être fermé. ■

3.1.19 REMARQUE
Utilisant l'injectivisation de f, on peut montrer que la réciproque est vraie aussi. Ainsi, une forme linéaire f est bornée si et seulement si $\ker f$ est fermé. Il en résulte que le noyau d'une forme linéaire est fermé ou dense dans X. (Pourquoi ?)

3.2 Le théorèmes de représentation des formes linéaires

Dans les espaces de Banach concrèts, les formes linéaires bornées ont généralement une forme spécifique et pratique. De manière générale, toutes les formes linéaires sur les espaces *de fonctions (telle que L^p et $C(K)$) sont représentée par l'intégration d'une fonction (par rapport à un certain poids ou mesure). De même, toutes les formes linéaires sur les espaces de suite (tels que ℓ^p et c_0) qui agissent par des somme avec des poids.*

3.2.1 Le dual de L^p

Une version du théorème de représentation de Riesz, est valable pour tous les espaces L^p . En résumé, il affirme que $(L^p)' = L^q$ où p et q sont des exposants conjugués comme dans l'inégalité de Hölder, c'est à dire

$$\frac{1}{p} + \frac{1}{q} = 1, \quad 1 \leq p, p' \leq \infty,$$

et $q = \infty$ si $p = 1$.

Pour les espaces ℓ^p, on a :

3.2.1 PROPOSITION ($(\ell^p)' = \ell^q$)
Soit $1 \leq p < \infty$ et q est l'exposant conjugué de p.

(i) Pour tout y $\in \ell^q$, somme dont le poids

$$G(x) := \sum_{k=1}^{\infty} x_k y_k, \quad x \in \ell^p$$

est une forme linéaire bornée sur ℓ^p, et sa norme est $\|G\| = \|y\|_q$.

(ii) Inversement, toute forme linéaire bornée $G \in (\ell^p)'$ peut être représenté en tant que somme de poids pour une masse unique $y \in \ell^q$. De plus, $\|G\| = \|y\|_q$.

Démonstration: Nous ne prouvons que le cas où $1 < p, q < \infty$, et le cas $p = 1, q = \infty$ est en exercice.

(i) Par l'inégalité H ölder, nous avons

$$|G(x)| = \Big|\sum_k x_k y_k\Big| \leq \|x\|_p \|y\|_q.$$

Il en résulte que F est une forme linéaire bornée sur ℓ^p, et $\|G\| \leq \|y\|_q$. Pour prouver l'inégalité inverse, notons que l'inégalité H "est plus forte. Plus précisément, pour tout y $\in \ell^q$, il existe une telle $x \in \ell^p$ qui

$$|G(x)| = \Big|\sum_k x_k y_k\Big| = \|x\|_p \|y\|_q.$$

En effet, on peut vérifier que cela est vrai pour $x = (x_k)$ définie par

$$x_k = e^{-i \arg(y_k)} |y_k|^{p'-1}.$$

Pour ce x, il s'ensuit que $\|G\| \geq \|y\|_q$, donc la partie (i) du théorème est démontré.

(ii) Considérons $g \in (\ell^p)'$. Soit EK désigne comme d'habitude les vecteurs de coordonnées dans ℓ^p, c'est- $e_k = (0, \ldots, 0, 1, 0, \ldots)$ avec k-ième de coordonner l'égalité 1. Par la linéarité et la continuité de G, nous avons

$$G(x) = G\Big(\sum_{k=1}^{\infty} x_k e_k\Big) = \sum_{k=1}^{\infty} x_k G(e_k), \quad x = (x_k) \in \ell^p.$$

Nous affirmons que la conclusion de (ii) suit de

$$y = (y_k) := (G(e_k)).$$

De toute évidence, $G(x) = \sum_k x_k y_k$ tel que requis. Pour prouver que $y \in \ell^q$, nous considérons

$$y^{(n)} := \sum_{k=1}^{n} y_k e_k \quad \text{et} \quad x^{(n)} := \sum_{k=1}^{n} e^{-i \arg(y_k)} |y_k|^{p'-1} e_k, \quad n = 1, 2, \ldots$$

Alors, en utilisant le cas de l'égalité de l'inégalité de Hölder, nous voyons que

$$\|G\| \|x^{(n)}\|_p \geq |G(x^{(n)})| = \Big|\sum_{k=1}^{n} x_k y_k\Big| = \|x^{(n)}\|_p \, \|y^{(n)}\|_q.$$

en simplifiant les $\|x^{(n)}\|_p$ des deux côtés, nous concluons que

$$\|y^{(n)}\|_q = \Big(\sum_{k=1}^{n} |y_k|^q\Big)^{1/q} \leq \|G\|.$$

soit $n \to \infty$, nous concluons que $y \in \ell^q$ et que $\|y\|_{p'} \leq \|G\|$. Par la partie (i), nous avons effectivement $\|y\|_q = \|G\|$. Ceci termine la preuve. ∎

3.2.3 Exemple ($c_0' = \ell^1$). double Montrer que $c_0' = \ell^1$. La signification de ceci est le même que dans le corollaire **??**, c'est à dire des formes linéaires sur c_0 sont donnés par sommation avec un poids de ℓ^1.

Le cas général

3.2.4 Théorème ($(L^p)' = L^q$)
Doit l'espace $L^p = L^p(\Omega, \Sigma, \mu)$, où μ est une mesure σ-finie et $1 \leq p < \infty$. Soit q est l'exposant conjugué de p.

(i) Pour chaque fonction $g \in L^q$, l'intégration avec le poids g

$$\phi(g)(f) := \int f g \, d\mu, \quad f \in L_p$$

est une forme linéaire bornée sur L^p, et sa norme est $\|\phi(g)\| = \|g\|_q$.

(ii) Réciproquement, toute forme linéaire bornée $G \in L^p$ peut être représenté par $\phi(g)$ pour une unique fonction $g \in L^q$. De plus, $\|G\| = \|\phi(g)\| = \|g\|_q$.

Démonstration: On va commencer par montrer que ϕ est une isométrie. Soit $g \in L^q(X)$ d'après l'inégalité de Hölder on a $\|\phi(g)(f)\|_1 \leq \|f\|_p \|g\|_q$ pour tout $f \in L^p(X)$; d'où $\|\phi(g)\| \leq \|g\|_q$.

Pour l'inégalité inverse, on va considérer le cas $1 < p < +\infty$, Alors le cas $p = 1$.

Soit $1 < p < +\infty$. Si $\|g\|_q = 0$, alors $g = 0$ p.p. et donc $\phi(g)$ est la forme nulle.

Si $\|g\|_q \neq 0$, on pose $f_0 = sgn(g).|g|^{q-1}$ où la fonction "signe de g", est définie par

$$sgn(g)(x) = \begin{cases} 0 & \text{si } g(x) = 0 \\ \frac{\overline{g(x)}}{|g(x)|} & \text{si } g(x) \neq 0 \end{cases}$$

Alors, $|f_0|^p = |g|^{p(q-1)} = |g|^q$, d'où $\|f_0\|_p = \|g\|_q^{\frac{q}{p}}$. Ainsi, $f_0 \in L^p(X)$ et $\|f_0\|_p > 0$.

On en déduit, $\|\phi(g)\| = \sup\limits_{f \neq 0} \dfrac{|\phi(g)(f)|}{\|f\|_p} \geq \dfrac{|\phi(g)(f_0)|}{\|f_0\|_p} = \dfrac{|\int_X f_0 g \, d\mu|}{\|f_0\|_p} = \dfrac{\left|\int_{g\neq 0} |g|^q \, d\mu\right|}{\|g\|_q^{\frac{q}{p}}}$

d'où $\|\phi(g)\| \geq \dfrac{\|g\|_q^q}{\|g\|_q^{\frac{q}{p}}} = \|g\|_q$.

Maintenant on va traiter le cas $p = 1$. Soit $g \in L^\infty(X)$. Si $\|g\|_\infty = 0$ il est clair que $\phi(g) = 0$. Supposons que $\|g\|_\infty > 0$, alors pour tout $0 < \epsilon < \|g\|_\infty$, on pose $X_\epsilon = \{x \in X \,|\, \|g\|_\infty - \epsilon \leq |g(x)| \leq \|g\|_\infty\}$.

Par définition du supremum essentiel $\mu(X_\epsilon) > 0$, comme μ est σ-finie, il existe $Y_\epsilon \subset X_\epsilon$ tel que $0 < \mu(Y_\epsilon) < +\infty$. On pose $f_\epsilon = \dfrac{\mathbb{1}_{Y_\epsilon}}{\mu(Y_\epsilon)} sgn(g)$. Alors, $\|f_\epsilon\|_1 = 1$ et

$\|\phi(g)\| = \sup\limits_{f \neq 0} \dfrac{|\phi(g)(f)|}{\|f\|_1} \geq \dfrac{|\phi(g)(f_\epsilon)|}{\|f_\epsilon\|_1} = \left|\displaystyle\int_X f_\epsilon g \, d\mu\right| = \dfrac{1}{\mu(Y_\epsilon)} \displaystyle\int_{Y_\epsilon} |g| \, d\mu \geq \|g\|_\infty - \epsilon.$

En faisant tendre ϵ vers 0, on obtient $\|\phi(g)\| \geq \|g\|_\infty$. On a ainsi montrer, que si $1 \leq p < +\infty$, l'application ϕ est une isométrie, donc injective, il reste à montrer qu'elle est surjective.

Ceci est une consuite, d'un résultat sur la théorie de la mesure à valeurs complexes, connu sous le nom de "Théorème de Radon-Nikodym".

Considérons deux mesures μ, ν sur le même σ-algèbre. Rappelons que ν est dite *absolument continue* par rapport à μ, en abrégé $\nu \ll \mu$, si

$$\mu(A) = 0 \quad \text{implique} \quad \nu(A) = 0$$

pour les ensembles mesurables A.

3.2.6 THÉORÈME (RADON-NYKODYM)
Soit (X, Σ, μ) un espace mesuré et soit ν une mesure à valeurs complexes, qui est absolument continue par rapport à μ . Alors il existe une unique fonction $f \in L^1(X, \mu)$ telle que pour tout $A \in \Sigma$ on a :

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

3.2.7 REMARQUE
En général, $(L^\infty)'$ contient strictement L^1, car L^∞ n'est pas séparable.

3.2.2 Exercices supplémentaires

3.2.8 EXEMPLE (L'ESPACE ENGENDREÉ PAR LES FORMES D'ÉVALUATIONS EN UN POINT).
Calculer la norme de la forme linéaire suivante sur $C[0,1]$:

$$F(f) = \sum_{k=1}^{n} a_k f(t_k)$$

où a_k sont des scalaires fixes et $t_1, \ldots, t_n$ sont des points fixes distincts dans $[0,1]$.

En déduire que l'espace engendré par les formes linéaires d'évaluations aux points $\delta_{t_1}, \ldots, \delta_{t_n}$ de $(C[0,1])'$ est isométrique à $\mathbb{R}^n$.

3.3 Le théorème de Hahn-Banach

Le théorème de Hahn-Banach permet de d'étendre les formes linéaires continues f sur un sous-espace vectoriel normé, tout en préservant la continuité de f. Le théorème de Hahn-Banach est un outil majeur dans l'analyse fonctionnelle. Ce résultat a des applications dans divers domaines des mathématiques, de l'informatique, de l'économie et de l'ingénierie.

Soit X un espace vectoriel normé, et soit X_0 un sous-espace de X. Considérons une forme linéaire bornée f_0 définie sur X_0, soit $f_0 \in X_0'$. Une extension *de f_0 à tout l'espace X est une forme linéaire bornée $f \in X'$ dont la restriction à X_0 coïncide avec f_0, soit*

$$f|_{X_0} = f_0|_{X_0}, \quad \text{ce qui signifie que } f_0(x_0) = f(x) \text{ pour tout } x_0 \in X_0.$$

Construire des extensions est un problème non trivial en raison de l'exigence de la continuité de f.

3.3.1 EXEMPLE. Montrer que si l'on n'impose pas la continuité de f, on peut construire f en utilisant une base Hamel.

3.3.2 DÉFINITION
Soit X un espace vectoriel sur $\mathbb{R}$. Une application $p : X \to \mathbb{R}$ est une fonctionnelle sous-linéaire si :

(i) $p(\lambda x) = \lambda p(x)$ pour tout $\lambda \in \mathbb{R}_+$, $x \in X$

(ii) $p(x+y) \leq p(x) + p(y)$ pour tout $x, y \in X$.

3.3.3 EXEMPLE. (1) Toute forme linéaire est une fonctionnelle sous-linéaire.
Toute semi-norme, et en particulier toute norme, est une fonctionnelle sous-linéaire.

(2) **Fonctionnelle de Minkowski**
Soit C un convexe d'un espace vectoriel normé $(X, \|.\|)$ tel que $0 \in C^0$.
On pose pour tout $x \in X$, $p_C(x) := \inf\{\alpha > 0 : \frac{x}{\alpha} \in C\}$.
Alors p_C est une fonctionnelle sous-linéaire, appelée Jauge ou fonctionnelle de Minkowski de C.

Démonstration: Comme $0 \in C^0$, il existe une boule fermée $\bar{B}(0,r) \subset C$, d'où pour tout $x \in X - \{0\}$ on a
$\frac{rx}{\|x\|} \in \bar{B}(0,r) \subset C$ Ainsi, l'application $p_C : X \to \mathbb{R}$ est bien définie.

(i) Soit $\lambda > 0$ et $x \in X$, $p_C(\lambda x) := \inf\{\alpha > 0 : \frac{\lambda x}{\alpha} \in C\} = \lambda \inf\{\frac{\alpha}{\lambda} > 0 : \frac{\lambda x}{\alpha} \in C\} = \lambda p_C(x)$.

(ii) Soient $x, y \in X$ et α_1 et α_2 tels que $\frac{x}{\alpha_1} \in C$ et $\frac{y}{\alpha_2} \in C$.
Alors $\frac{x+y}{\alpha_1+\alpha_2} = \frac{\alpha_1}{\alpha_1+\alpha_2}\frac{x}{\alpha_1} + \frac{\alpha_2}{\alpha_1+\alpha_2}\frac{y}{\alpha_2} \in C$ par convexité. Ainsi $p_C(x+y) \leq \alpha_1 + \alpha_2$ et en passant à l'infimum sur tels α_1 et α_2, on obtient :
$p_C(x+y) \leq p_C(x) + p_C(y)$. ∎

3.3.5 Exercice 1. Montrer que si $C = \bar{B}(0,r)$, alors $p_C = \|.\|$.

2. Montrer que si C est ouvert (reps. fermé) alors $C = \{x \in X : p_C(x) < 1\}$ (resp. $C = \{x \in X : p_C(x) \leq 1\}$.

3. Soit C un convexe fermé et $0 \in C$, d'un espace de dimension finie. Montrer que C est borné si et seulement si "$p_C(x) = 0 \Leftrightarrow x = 0$."

3.3.6 THÉORÈME (DE HAHN-BANACH)
Soit E un espace vectoriel sur $\mathbb{R}$ et $p : E \to \mathbb{R}$ une fonctionnelle sous-linéaire.

Soit $F \subset E$ un sous-espace vectoriel et $f : F \to \mathbb{R}$ une forme linéaire, telle que $f(x) \leq p(x)$ pour tout $x \in F$.

Alors il existe une forme linéaire $g : E \to \mathbb{R}$ telle que

1. $g = f$ sur F i.e. $g(x) = f(x)$ pour tout $x \in F$.
2. $f \leq p$ sur E i.e. $f(x) \leq p(x)$ pour tout $x \in E$.

Démonstration: On suppose que $F \neq E$, sinon il n'y a rien à montrer.

étape 1 : Extension par une seule dimension.

Soit $z \in F^c = E \backslash F$ et $x, y \in F$. Alors $f(x) - f(y) \leq p(x - y \leq p(x+z) + p(-y-z)$ d'où $-p(-y-z) - f(y) \leq p(x+z) - f(x)$. Ainsi pour tout $x \in F$,

$$s = \sup_{y \in F}\{-p(-y-z) - f(y)\} \leq p(x+z) - f(x),$$

d'où

$$s \leq \inf_{x \in F}\{p(x+z) - f(x)\}.$$

donc il exists un hombre t tel que pour tout $y \in F$,

$$-p(-y-z) - f(y) \leq t \leq p(y+z) - f(y). \tag{3.3.1}$$

On pose alors $F_z = \{x + \alpha z : x \in F, \alpha \text{ real}\}$. D'où si $w \in F_z$ on a $w = x + \alpha z$ et la décomposition est unique. On définit la fonction h sur F_z par

$$h(w) = f(x) + \alpha t.$$

Alors h est linéaire sur F_z et est une extension de f de F à F_x. On remarque alors que

$$h(w) \leq p(w) \quad \text{sur } F_z.$$

Ceci est une conséquence directe de (3.3.1) en posant $y = x/\alpha$, $\alpha \neq 0$ et en considérant le cas $\alpha > 0$ Alors $\alpha < 0$.

étape 2 : Induction transfinie. Cet argument repose sur le lemme norme Zorn **??** de la même manière que dans la preuve de la proposition 1.2.6 sur l'existence des bases de Hamel.

Si $F_z = E$ on a alors fini ; sinon on répète le processus d'extension, mais comment garantir, si l'espace est de dimension infinie, qu'on Alorsse ainsi obtenir une entension à l'espace tout entier ? c'est à ce niveau qu'on a besoin du lemme de Zorn.

Soit $\mathcal{E}$ l'ensemble des paires ordonnées (F', h'), où F' est un sous-espace vectoriel contenant F et h' est une extension de f à F' telle que $h' \leq p$ sur F'. L'ordre partiel sur $\mathcal{E}$ est défini par

$$(F', h') \preccurlyeq (F'', h'') \quad \text{si et seulement si } F' \subset F'', h' = h'' \ \text{ sur } \ F'.$$

Soit $C = \{(F_i, g_i)$
$, \alpha \in I\}$ une chaîne de $\mathcal{E}$ i.e. une partie totalement ordonnée de $\mathcal{E}$.

On pose $F_C = \cup_\alpha F_i$ et $f_C : F_C \to \mathbb{R}$ la forme linéaire définie par :

$$f_C(x) = g_i(x) \quad \text{for } x \in F_i.$$

Alors, (G, g) est un majorant de C.

($\cup F_i$ est un sous-espace car C est totalement ordonnée.)

Donc $(\mathcal{E}, \preccurlyeq)$ est un ensemble ordonné inductif et d'après le lemme de Zorn, il admet un élément maximal $(\mathscr{G}, g)$.

Tout ce qui reste à vérifier pour terminer la preuve, est que $\mathscr{G} = E$.

Supposons le contraire i.e. $\mathscr{G} \neq E$. Alors, d'après la première partie de la preuve, on peut faire une extension $(\mathscr{G}_z, {}_z) \succcurlyeq (\mathscr{G}, g)$. Comme $(\mathscr{G}, g)$ est maximal on doit avoir $\mathscr{G}_z = \mathscr{G}$, ce qui contredit $\mathscr{G}$ est un sous-espace strict de $\mathscr{G}_z$. Ainsi $\mathscr{G} = E$ et g est l'extension souhaitée.

3.3.8 Corollaire (Extension par continuité)

Soit $(E, \|.\|)$ un espace vectoriel normé sur $\mathbb{K}$ (=$\mathbb{R}$ ou $\mathbb{C}$.)

Soit F un sous-espace vectoriel et $f : F \to \mathbb{K}$ une forme linéaire continue.

Alors il existe une forme linéaire continue $g : E \to \mathbb{K}$ telle que $\|g\| = \|f\|$.

Démonstration: (i) $\boxed{\mathbb{K} = \mathbb{R}}$.

On pose $C = \|f\|$ et $p = C\|.\|$.

Alors $f \leq p$ sur F et d'après le théorème de hahn-Banach, il existe une forme linéaire continue $g : E \to \mathbb{R}$ telle que $g = f$ sur F et $g \leq p$ sur E.

Ceci entraîne que pour tout $x \in E$, $g(x) \leq C\|x\|$ et $g(-x) \leq C\| - x\| = C\|x\|$ i.e. $|g(x)| \leq C\|x\|$. D'où lg $\| \leq C = \|f\|$ et comme $|g(x)| = |f(x)|$ sur F, on aura $\|g\| \geq \|f\|$ ce qui donne finalement $\|g\| = \|f\|$.

(ii) $\boxed{\mathbb{K} = \mathbb{C}}$.

On écrit, $f = f_1 + if_2$, où $f_1 = \mathfrak{Re}(f)$ et $f_2 = \mathfrak{Im}(f)$. Alors f_1 et f_2 sont des formes linéaires réelles sur F considéré comme espace vectoriel sur $\mathbb{R}$, et $f_2(x) = -f_1(ix)$.

D'après le cas réel, il existe une forme linéaire continue $g_1 : E \to \mathbb{R}$ telle que $g_1 = f_1$ sur F et $\|g_1\| = \|f_1\|$.
On pose, pour tout $x \in E$, $g(x) = g_1(x) - ig_1(ix)$. Alors $g(x) = f(x)$ sur F, on aussi $g(ix) = ig(x)$ et comme g_1 est $\mathbb{R}$-linéaire, g est $\mathbb{C}$-linéaire sur l'espace vectoriel complexe E.
Il reste à montrer que $\|g\| \leq \|f\|$, ce qui entraînera $\|g\| = \|f\|$.
Soit $x \in E$ tel que $g(x) \neq 0$ et $\alpha = \arg(g(x))$, d'où

$$|g(x)| = g(x)e^{-i\alpha} = g(xe^{-i\alpha}) = g_1(xe^{-i\alpha}) = \|g_1(xe^{-i\alpha})\| \leq \|g_1\|\|x\| = \|f\|\|x\|$$

donc $\|g\| \leq \|f\|$. ■

3.3.10 COROLLAIRE
Soit $E \neq \{0\}$ un espace vectoriel normé sur $\mathbb{K}$ (=$\mathbb{R}$ ou $\mathbb{C}$). Soit $x_0 \in E - \{0\}$. alors il existe une forme linéaire continue $f : E \to \mathbb{R}$ telle que

$$\begin{cases} f(x_0) = \|x_0\| \\ \|f\| = 1 \end{cases}$$

Démonstration: Soit $x_0 \neq 0$, on pose $F = \{\alpha x_0 : \alpha \in \mathbb{K}\}$. Alors F set un sous-espace et on définit, $f_0 : F \to \mathbb{K}$, par $f_0(\alpha x_0) = \alpha\|x_0\|$. Alors f_0 est une forme linéaire continue sur F de norme $\|f_0\| = 1$ (Alorsque $|f_0(x)| = |\alpha|\,\|x_0\| = \|x\|$ pour tout $x \in F$). D'après le lemme précédent, f_0 admet une extension linéaire continue $f : E \to \mathbb{K}$ qui préserve la norme i.e. $\|f\| = \|f_0\| = 1$ et $f = f_0$ sur F, entraine que $f(x_0) = \|x_0\|$. ■

3.3.12 REMARQUE
Si $E \neq \{0\}$ un espace vectoriel normé réel ou complexe, alors il existe toujours une forme linéaire continue et non identiquement nulle sur E i.e. $E' \neq \{0\}$.

3.3.13 COROLLAIRE
Soit E un espace vectoriel normé réel ou complexe. Soit $x \in E$ tel que $f(x) = 0$ pour tout $f \in X^*$. Alors $x = 0$.

Démonstration: Sinon $x \neq 0$ et d'après le corollaire précédent, il existe $f \in X^*$ telle que $f(x) > 0$. ■

3.3.15 Corollaire (X^* sépare les points de X)
Pour tous vecteurs $x_1 \neq x_2$ dans un espace vectoriel normé X, il existe $f \in E^*$ telle que $f(x_1) \neq f(x_2)$.

Démonstration: Par hypothèse $x = x_1 - x_2 \neq 0$, alors on prend $f \in E^*$ telle que $f(x) \neq 0$ convient. ■

Comme application du théorème de Hahn-Banach on a le théorème de représentation de Riesz pour les éléments de $(C[a,b])'$.

3.3.17 Théorème (Le théorème de représentation de Riesz pour $C[0,1]$)
Soit $f \in (C[0,1])'$, alors il existe une fonction à variation bornée g, sur $[0,1]$, telle que

$$f(x) = \int_0^1 x(t)dg(t) \quad \text{pour tout } \; x \in C[0,1],$$

$$\|f\| = \bigvee_0^1(g) = \text{ la variation totale de } g \text{ on [0,1]}$$

où $\bigvee_0^1(g) = \sup_{\substack{n\geq 1 \\ 0\leq t_0<\ldots<t_n\leq 1}} \sum_{i=1}^{n} |g(t_i) - g(t_{i-1})|$

et l'intégrale est au sens de Riemann.

Démonstration: Avant de commencer la démonstration, on remarque qu'on suppose connue certain propriétés de l'intégrale de Riemann-Stieltjes $\int_0^1 x(t)dg(t)$.

Pour commencer, on a besoin de connaître que si $x \in C[0,1]$ et $g \in BV[0,1]$ (l'espace des fonctions à variations bornées sur $[0,1]$) alors $\int_0^1 x(t)dg(t)$ existe et

$$\lim_{\mu(P)\to} \sum_{i=1}^{n} x(c_i)(g(t_i) - g(t_{i-1})) = \int_0^1 x(t)dg(t),$$

où $P\{0 = t_0, t_1, \cdots, t_n = 1\}$, $t_0 < t_1 < \cdots < t_n$ est une partition de $[0,1]$; $\mu(P) = \max(t_i - t_{i-1}), 1 \leq i \leq n$ et $c_i \in [t_{i-1}, t_i]$. La limite est prise sur toutes les partitions P et tous les choix de c_i.

On considère $C[0,1]$ comme un sous-espace de l'espace de Banach M des fonctions bornées sur $[0,1]$ muni de la nome $\|x\| = \sup\{|x(t)| : 0 \leq t \leq 1\}$, $x \in M$. Si $f \in (C[a,b])'$, d'après le théorème de Hahn-Banach, il existe une extension F à M telle que $f(x) = F(x)$ sur $C[0,1]$ et $\|F\| = \|f\|$.

Soit χ_t la fonction caractéristique de $[0,t]$, $0 \leq t \leq 1$, i.e. $\chi_t(t) = 1$ $(0 \leq t \leq t)$, $\chi_t(y) = 0$ $(t \leq t \leq 1)$. Alors $\chi_t \in M$ pour tout $t \in [0,1]$ et on écrit $F(\chi_t) = g(t)$. On pose maintenant $\varepsilon_i = \mathrm{sgn}[g(t_i) - g(t_{i-1})]$. Alors $|g(t_i) - g(t_{i-1})| = [g(t_i) - g(t_{i-1})]\varepsilon_i$ d'où

$$\sum_{i=1}^{n} |g(t_i) - g(t_{i-1})| \leq \|F\| \left\| \sum_{i=1}^{n} [\chi_{t_i} - \chi_{t_{i-1}}]\varepsilon_i \right\|.$$

si $y \in [0,1]$ alors $y \in]t_{j-1}, t_j]$ pour un certain j, ainsi $\sum_{i=1}^{n} [\chi_{t_i}(y) - \chi_{t_{i-1}}\varepsilon_i] = \varepsilon_j$. Par conséquent

$$\sum_{i=1}^{n} |g(t_i) - g(t_{i-1})| \leq \|F\|, \tag{3.3.2}$$

ainsi $g \in BV[0,1]$.

On définit

$$z_n = \sum_{r=1}^{n} x(r/n)(\chi_{r/n} - \chi_{(r-1)/n}).$$

Alors si $u \in]0,1[$ on a $u \in](r-1)/n, r/n[$ pour un certain r d'où $z_n(u) = x(r/n)$. Comme x est continue et donc uniformément continue sur $[0,1]$, il s'en suit que $\|x - z_n\| \to 0$ lorsque $n \to \infty$. Par conséquent, $F(z_n) \to F(x) = f(x)$ d'où

$$f(x) = \lim_{n\to\infty} \sum_{r=1}^{n} x(r/n) \left(g(r/n) - g\left(\frac{r-i}{n}\right) \right) = \int_0^1 x(t) dg(t). \tag{3.3.3}$$

De (3.3.3), on obtient que

$$|f(x)| \leq \max |x(t)| \bigvee_0^1 (g) = \|x\| \bigvee_0^1 (g),$$

ainsi $\|f\| \leq \bigvee_0^1(g)$, et (3.3.2), donnent $\|f\| = \bigvee_0^1(g)$. ■

3.3.19 REMARQUE

Dans le théorème 3.3.17 la fonction $g \in BV$ n'est pas unique, Par example, on peut rajouter use constant arbitraire à g et obtenir la même représentation pour f. Donc on ne peut pas identifier $(C[a,b])'$ avec $BV[0,1]$.

Pour obtenir l'identification de $(C[a,b])'$ on a besoin de deux propriétés des fonctions g de $BV[0,1]$. La première est que la limite à droite, $g(t+0) = \lim_{x\to t^+} g(x)$ existe pour tout $t \in]0,1[$, et la seconde est que l'ensemble des points de discontinuités de g est au plus dénombrable.

On écrit

$$\widehat{BV}[0,1] = \{G \in BV[0,1] : G(0) = 0 \text{ and } G(t+0) = G(t), 0 < t < 1\}.$$

Alors $\widehat{BV}$ est un sous-espace de BV. On définit la fonction G par $G(0) = 0$, $G(1) = g(1) - g(0)$ et $G(t) = g(t+0) - g(0)$, $0 < t < 1$. Alors $G(t) = g(t) - g(0)$ est *continue sur* $[0,1]$. *D'où pour tout* $x \in C[0,1]$ *on a* $\int_0^1 x(t)dg(t) = \int_0^1 x(t)dG(t)$. *Il est maintenant clair que* $G \in \widehat{BV}$. *Ainsi à tout* $f \in (C[a,b])'$ *on associe* $G \in \widehat{BV}$ *telle que*

$$f(x) = \int_0^1 x(t)dG(t) \quad \textit{for all } x \in C[0,1].$$

La function associée G *est unique et* $\|f\| = \bigvee_0^1(g)$, *ceci complète l'identification de* $(C[a,b])'$ *avec* $\widehat{BV}[0,1]$.

Une autre application du théorème de Hahn-Banach est la relation de l'espace normé X *avec son bidual. Le bidual sera noté* $X'' := (X')'$ *(ou* $X^{**} = (X^*)^*$. *On va montrer que* X *est identifié avec un sous-espace* $\widetilde{X}$ *de* X^{**}. *On peut donc voir* X *comme plongé naturellement dans son bidual* X^{**}.

3.3.20 Théorème

Soit X un espace vectoriel normé. Alors X est isométriquement isomorphe à un sous-espace $\widetilde{X}$ de son bidual X^{**}.

Démonstration: On note un élément de X par x, et un élément de X^* par x^*. On définie la forme linéaire par $\tilde{x}$ on X^* by

$$\tilde{x}(x^*) = x^*(x).$$

Alors pour tout, $x^* \in X^*$ on associe la valeur de x^* au point $x \in X$. $\tilde{x}$ est linéaire et continue sur X^* :

$$|\tilde{x}(x^*)| = |x^*(x)| \leq \|x^*\| \, \|x\|,$$

ainsi $\|\tilde{x}\| \leq \|x\|$. Alors l'application $J_X : x \to \tilde{x}$ définie une application linéaire continue de X dans X^{**}. On note $\widetilde{X}$ l'image de X par J_X, i.e, $\widetilde{X} = J_X(X)$. Il nous reste à montrer que J_X est une isométrie. D'après le corollaire 3.3.10 to the il existe $y^* \in X^*$ telle que $y^*(x) = \|x\|$ et $\|y^*\| = 1$. D'où

$$\|J_X(x)\| = \|\tilde{x}\| = \sup_{x^* \neq 0} \frac{|\tilde{x}(x^*)|}{\|x^*\|} \geq |\tilde{x}(y^*)| = |y^*(x)| = \|x\|,$$

donc $\|J_X(x)\| = \|\tilde{x}\| \leq \|x\|$, ainsi $\|J_X(x)\| = \|x\|$. Ceci termine la preuve du théorème.

L'isométrie $J_X : X \to J(X) \subset X^{**}$ *est appelée l'application canonique de* X *dans* X^{**}. *En général on a* $J(X) = \widetilde{X} \subset X^{**}$. *Si on a égalité pour les espaces* **reflexifs**

3.3.22 Définition

Un espace vectorial normé X est dit reflexif si $J(X) = X^{**}$.

3.3.23 REMARQUE
Tout espace normé réflexif est un espace de Banach. La réciproque est en général fausse.

Les espaces ℓ^p $(1 < p < \infty)$, $L^p[a, b]$ $(p > 1)$ and toute espace de Hilbert sont des espaces reflexifs, mais les espaces de Banach suivant ne le sont pas : ℓ^1, $L^1[a, b]$ et $C[a, b]$.

Tout espace de dimension finie est un espace réflexif

3.3.24 PROPOSITION
Soit X un espace réflexif. Alors, chaque fonction $f \in X^*$ atteint sa norme sur X.

Démonstration: Soit $f \in X^*$. D'après le lemme 3.3.10, il existe $x \in X^{**} = X$ tel que $x(f) = \|f\|$ i.e $f(x) = \|f\|$, comme demandé. ∎

La réciproque de la proposition 3.3.24 est également vraie. Si toute forme linéaire $f \in X^*$ sur un espace de Banach X atteint sa norme, alors X est réflexive. C'est le théorème de James (1971).

3.3.26 THÉORÈME (DUAL ET SÉPARABILITÉ)
Soit X un espace vectoriel normé. Si X^* est séparable il en est de même de X.

Démonstration: Soit $\{f_n\}_{n\in\mathbb{N}}$ une suite dense d'élément de X^*. Soit $\{x_n\}_{n\in\mathbb{N}}$ une suite de X telle que pour tout n :

$$\begin{cases} \|f_n(x_{n_0})\| \geq \frac{1}{2}\|f_n\| \\ \|x_n\| = 1 \end{cases}$$

On va montrer que $\{x_n\}$ est totale dans X.

Supposons le contraire i.e. $\overline{\text{Vect}\{x_n\}} \neq X$. Alors, par le théorème de Hahn-Banach, il existe $f \in X^*$ telle que

$$\begin{cases} f(x_n) = 0 & \text{pour tout } n \in \mathbb{N} \\ \|f\| = 1 \end{cases}$$

Par densité de la suite $\{f_n\}_{n\in\mathbb{N}}$, il existe $n_0 \in \mathbb{N}$ tel que $\|f_{n_0} - f\| \leq \frac{1}{4}$.

Alors, $\|f_{n_0}\| \geq \|f\| - \|f_{n_0} - f\| \geq \frac{3}{4}$ et donc $0 = f(x_{n_0}) = f_{n_0}(x_{n_0}) + (f - f_{n_0})(x_{n_0})$, d'où

$$0 \geq |f_{n_0}(x_{n_0})| - |(f - f_{n_0})(x_{n_0})| \geq \frac{1}{2}\|f_n\| - \|f_{n_0} - f\| \geq \frac{1}{2}\frac{3}{4} - \frac{1}{4} = \frac{1}{8} > 0$$

ce qui est absurde, donc $\overline{\text{Vect}\{x_n\}} = X$ i.e. X est séparable.

3.3.28 REMARQUE
La réciproque à la proposition 3.3.26, est en générale fausse ; par exemple : $(L^1)^* = L^\infty$; L^1 est séparable, mais pas sont dual L^∞ n'est pas séparable

3.3.1 Exercices supplémentaires

3.3.29 Exercice D'après le théorème 3.2.4, $(L^p)^* = L^q$ où $1 \leq p < \infty$ et q est l'exposant conjugué de p. Par conséquent, L^p est un espace réflexif pour $1 < p < \infty$.

D'après la remarque 3.3.28, L^1 et L^∞ ne sont pas des espaces réflexifs.

- *Soit X un espace vectoriel normé. Supposons que $x_1, x_2, \cdots, x_k$ soient des vecteurs linéairement indépendant de X, $\alpha_1, \alpha_2, \cdots, \alpha_k$ des scalaires dans $\mathbb{K}$. Montrer qu'il existe $f \in X^*$ avec $f(x_\nu) = \alpha_\nu$ ($\nu = 1, 2, \cdots, k$) et que $\|f\| \leq M$ si et seulement si pour tout $t_1, t_2, \cdots, t_k \in \mathbb{K}$*

$$\left|\sum_{\nu=1}^{k} t_\nu \alpha_\nu\right| \leq M \left\|\sum_{\nu=1}^{k} t_\nu x_\nu\right\|$$

- *Soit X un espace vectoriel normé et Z un sous-espace vectoriel de X. Soit $x_0 \in X$ tel que $d(x_0, Z) = \inf_{z \in Z} d(x_0, z) > 0$. Montrer qu'il existe $f \in X^*$ telle que*
  (1) $f(z) = 0$ si $z \in Z$;
  (2) $f(x_0) = d(x_0, Z)$;
  (3) $\|f\| = 1$.
- *Montrer que le dual d'un espace de dimension finie est de dimension finie*

3.3.30 Exercice Montrer que le dual d'un espace réflexif est séparable.

3.4 Convergences forte et faible

3.4.1 DÉFINITION
Soit $\{x_n\}$ une suite d'un espace vectoriel normé X. on dit que $\{x_n\}$ converge faiblement vers $x \in X$ lorsque $n \to \infty$, et on écrit $x_n \rightharpoonup x$ lorsque $n \to \infty$, si

$$f(x_n) \to f(x) \quad \text{as } n \to \infty, \text{ pour tout } f \in X^*.$$

On dit que x est la limite faible de $\{x_n\}$.

Si $x_n \to x$ en norme, lorsque $n \to \infty$ (*i.e.* $\|x_n - x\| \to 0$ *lorsque* $n \to \infty$), on dira que $\{x_n\}$ converge fortement vers x lorsque $n \to \infty$.

3.4.2 Remarque

Si $\{x_n\}$ converge (fortement) vers x lorsque $n \to \infty$, alors, comme

$$|f(x_n) - f(x)| \leq \|f\| \, \|x_n - x\|$$

pour tout $f \in X^*$, on a $x_n \rightharpoonup x$ quand $n \to \infty$. Donc la convergence en norme entraîne la convergence faible, mais la réciipropque est en générale fausse (cf. voir le théorème savant).

3.4.3 Théorème

(i) La limite faible est unique.
(ii) En général, la convergence faible n'est pas équivalente à la convergence forte.
(iii) Si $x_n \rightharpoonup x$ lorsque $n \to \infty$ alors $\sup_n \|x_n\| < \infty$ *(i.e. $\{x_n\}$ est bornée).*

Démonstration: (i) Soit $x_n \rightharpoonup x$ et $x_n \rightharpoonup y$ lorsque $n \to \infty$. Alors $f(x-y) = f(x-x_n) + f(x_n - y) \to 0$ lorsque $n \to \infty$, d'où $f(x-y) = 0$ pour tout $f \in X^*$. D'après le corollaire 3.3.13 on a $x = y$.

(ii) Considérons $\ell^p (1 < p < \infty)$. D'après le théorème **??**, on sait que $f \in (\ell^p)^*$ peut s'écrire comme $f(x) = \sum_k a_k x_k$ pour tout $x = \{x_k\} \in \ell^p$, où $a = \{a_k\} \in \ell^q$ and $1/p + 1/q = 1$. Soit

$$e_n = (\underbrace{0, \cdots, 0}_{n-1}, 1, 0, \cdots) \quad \text{for all } n \in \mathbb{N},$$

alors, pour la norme de ℓ^p on a $\|e_n - e_m\|_p = 2^{1/p}$ $(n \neq m)$, ainsi $\{e_n\}$ n'est pas fortement convergente. D'autre part, $\{e_n\}$ est faiblement convergent vers 0 car, pour tout $f \in (\ell^p)^*$ on a $f(e_n) = a_n$, et $a = \{a_n\} \in \ell^p$ entraîne $a_n \to 0$ lorsque $n \to \infty$.

Dans le cas $p = 1$ la convergence faible et forte coïncident.

(iii) $x_n \rightharpoonup x$ quand $n \to \infty$ entraîne $f(x_n - x) \to 0$ lorsque $n \to \infty$. D'après la remarque 3.3.12 il existe $f_n \in X^*$ telle que $f_n(x_n - x) = \|x_n - x\|$ et $\|f_n\| = 1$. Pour tout $f \in X^*$ on définit $F_n(f) = f(x_n - x)$. Alors $\{F_n\}$ est une suite de formes linéaires continues de l'espace de Banach X^*. Comme $f(x_n - x) \to 0$ lorsque $n \to \infty$, par conséquent $\limsup_{n\to\infty} |F_n(f)| < \infty$ on X^*. le théorème de Banach-Steinhaus, permet d'affirmer que $M = \sup_n \|F_n\| < \infty$. Ainsi

$$\|x_n - x\| = |f(x_n - x)| = |F_n(f_n)| \leq \|F_n\| \, \|f_n\| \leq M,$$

d'où $\|x_n\| \leq M + \|x\|$ pour tout n. ∎

Un autre mode de convergence, peut être défini sur l'espace dual d'un espace vectoriel normé X appelé convergence faible* (faible étoile).

3.4.5 Définition

Soit X un espace vectoriel normé et X^* son espace dual. Soit $\{f_n\}$ une suite de formes linéaires continues X^*. S'il existe $f \in X^*$ telle que

(1) $\|f_n - f\| \to 0$ lorsque $n \to \infty$, on dit que $\{f_n\}$ converge fortement vers f.

(2) pour tout $x \in X$, $|f_n(x) - f(x)| \to 0$ lorsque $n \to \infty$, on dit alors que $\{f_n\}$ converge faiblement* (ou converge pour la topologie faible*) vers f lorsque $n \to \infty$.

(3) pour tout $F \in X^{**}$, $F(f_n) \to F(f)$ lorsque $n \to \infty$, on dit alors que $\{f_n\}$ converge faiblement vers f lorsque $n \to \infty$.

En general, les topologies forte, faible et faible* sont distinctes. Si X est reflexif alors les topologies faible et faible* coïncident.

3.4.1 Exercices supplémentaires

- *Soit X un espace vectoriel de dimension finie.*
  Montrer que les convergences faible et forte coîncident.
- *Soit M un sous-espace fermé d'un espace vectoriel normé X. Soit une suite $\{x_n\} \subset M$ qui converge faiblement vers x_0. Montrer que $x_0 \in M$.*
- *Soit X un espace de banach séparable et M un sous-ensemble borné de X^*. Montrer que pour toute suite de M admet une sous-suite faiblement* convergente.*
- *Soit X un espace de Banach et $T : X \to X$ use application linéaire. Montrer que T est continue si et seulement si pour tout $x_n \rightharpoonup x$ ($n \to \infty$), on a $Tx_n \rightharpoonup Tx$ lorsque $n \to \infty$.*
- *Soit $\{x_n\}$ une suite de $C[a,b]$ et $x_0 \in C[a,b]$. Montrer que $\{x_n\}$ converge faiblement vers x_0 si et seulement si il existe une constante $M > 0$ telle que $\|x_n\| \leq M$ pour tout $n \in \mathbb{N}$, et $\lim_{n\to\infty} x_n(t) = x_0(t)$ pour tout $t \in [a,b]$.*

3.5 La séparation des ensembles convexes

Le théorème de Hahn-Banach a des conséquences géométriques remarquables, qui sont regroupées sous le nom de théorèmes de séparation. *Sous certaines conditions topologiques, ces résultats garantissent que deux ensembles convexes A, B peuvent toujours être séparés par un hyperplan. Comme nous le savons à partir de la section 3.1.3, les hyperplans correspondent à des ensembles de niveau de formes linéaires f. Par conséquent, Nous nous* *attendons à ce qu'un théorème de séparation pour A, B soit donnét par une forme linéaire f et un nombre C tels que*

$$f(a) \leq C \leq f(b), \quad a \in A, b \in B.$$

Dans ce cas, les ensembles A et B sont séparés par l'hyperplan $\{x : f(x) = C\}$.

Commençons par le cas plus simple lorsque l'un des deux ensembles est un point.

3.5.1 Théorème (Séparation d'un point par rapport à un ensemble convexe) Soit K un ouvert convexe d'un espace vectoriel normé X, et un point $x_0 \notin K$. Alors il existe une fonction $f \in X^*$, $f \neq 0$, telle que

$$f(x) < f(x_0), \quad x \in K.$$

Démonstration: Quitte à translater K, on peut supposer sans perte de généralité que $0 \in K$. (Pourquoi ?). Par conséquent, la fonctionnelle de Minkowski de K, p_K, est une fonctionnelle sous-linéaire sur X.

Comme $0 \in K$ et K est ouvert, il contient une boule centrée $B_X(0, r)$ de rayon $r > 0$ (Voir la figure). L'inclusion ensembliste $B_X(0, r) \subseteq K$ implique l'inégalité pour la fonctionnelle de Minkowski :

$$\frac{1}{r}\|x\| \geq p_K(x), \quad x \in K.$$

(Pourquoi ?)

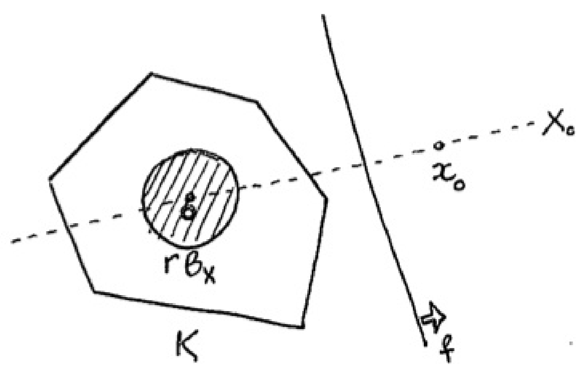


FIGURE 3.1 – Séparation d'un point x_0 de l'ensemble K par f fonctionnelle

Considérons le sous-espace de dimension 1

$$X_0 = \text{Vect}(x_0)$$

et on définit une forme linéaire f_0 sur X_0^* par

$$f_0(tx) = tp_K(x), \quad t \in \mathbb{R}.$$

Alors f_0 est dominée par p_K sur X_0, pour $t \geq 0$, nous avons

$$f_0(tx_0) = p_K(tx_0); \;\; f_0(-tx_0) = -tf_0(x_0) = -tp_K(x_0) \leq 0 \leq p_K(tx_0). \qquad (3.5.1)$$

Par le théorème de Hahn-Banach, f_0 admet une extension f définie sur X telle que la domination est préservée, à savoir

$$f(x) \leq p_K(x), \quad x \in X.$$

Pour terminer la preuve, nous avons besoin de vérifier que f est bornée et qu'elle sépare x_0 de K. La continuité de f découle de l'inégalité

$$f(x) \leq p_K(x) \leq \frac{1}{r}\|x\|, \quad x \in X,$$

Donc $f \in X^*$. Pour vérifier la séparation, on prend un $x \in K$. Comme $x_0 \notin K$, nous avons

$$f(x) \leq p_K(x) < 1 \leq \|x_0\| = f_0(x_0) = f(x_0).$$

Ceci termine la preuve. ■

3.5.3 THÉORÈME (SÉPARATION DES ENSEMBLES CONVEXES OUVERTS)
Soit A, B des convexes disjoints d'un espace vectoriel normé X.

(i) Supposons que A soit ouvert. Alors il existe une forme linéaire $f \in X^*$ et un nombre $C \in \mathbb{R}$ de telle sorte que

$$f(a) < C \leq f(b), \quad a \in A, b \in B.$$

(ii) Supposons que A et B soient ouverts, on a alors une inégalité plus forte, : il existe une forme linéaire $f \in X^*$ et un nombre $C \in \mathbb{R}$

$$f(a) < C < f(b), \quad a \in A, b \in B.$$

Démonstration: (i) Considérons la *différence de Minkowski*

$$K = A - B := \{a - b : \; a \in A, b \in B\}.$$

L'ensemble K est un convexe ouvert. Comme A et B sont disjoints, $0 \notin K$.

En utilisant le théorème 3.5.1 , on obtient une fonction $f \in X^*$, $f \neq 0$ tel que

$$f(a-b) \leq f(0) = 0, \quad a \in A, b \in B.$$

Donc $f(a) \leq f(b)$ pour tout $a \in A$, $b \in B$, d'où en posant $C := \sup_{a \in A} f(a)$, on obtient

$$f(a) \leq C \leq f(b), \quad a \in A, b \in B.$$

Comme A est ouvert, en considérant un voisinage de a dans A, on obtient l'inégalité stricte

$$f(a) < C \leq f(b), \quad a \in A, b \in B.$$

(ii) Cette partie est similaire, en considérant de petits voisinages de a dans A et de b dans B. ∎

3.5.5 EXEMPLE. Ecrire les détails de la preuve du théorème 3.5.3.

3.5.6 COROLLAIRE (SÉPARATION DES ENSEMBLES CONVEXES FERMÉS)
Soit A, B des sous-ensembles convexes, disjoints et fermés d'un espace vectoriel normé X. Supposons que B soit compact. Alors il existe une fonction $f \in X^*$ et un nombre $C \in \mathbb{R}$ tels que

$$\sup_{a \in A} f(a) \leq C < \inf_{b \in B} f(b).$$

Démonstration: Soit

$$r = \operatorname{dist}(A, B) := \inf_{a \in A, b \in B} \|a-b\|.$$

par hypothèse, $r > 0$. (Pourquoi ?) Par conséquent, $A_{r/3} := \{x \in X : d(x, A) < r/3\}$ voisinage de A et $B_{r/3} := \{x \in X : d(x, B) < r/3\}$ voisinage de B sont disjoints, ouverts et convexes (pourquoi ?). D'après le théorème 3.5.3, il existe une fonction $f \in X^*$ et $C \in \mathbb{R}$ telle que les $A_{r/3}$ et $B_{r/3}$:

$$f(a) < C < f(b) \quad \text{pour tout } a \in A_{r/3}, b \in B_{r/3}$$

l'inégalité est vraie a fortiori pour A et B.

Comme B est compacte, $\inf_{b \in B} f(b) > C$.

Ainsi

$$\sup_{a \in A} f(a) \leq C < \inf_{b \in B} f(b).$$

3.5.8 EXEMPLE. Ecrire les détails dans la preuve du corollaire 3.5.6.

3.5.9 REMARQUE
Supposons que dans le théorème 3.5.1 , l'ensemble K est soit ouvert (c'est à dire comme dans l'énoncé) ou fermé. Alors, la séparation est stricte :

$$f(x) < f(x_0), \quad x \in K.$$

En effet, pour des ensembles ouverts c'est une conséquence du théorème 3.5.3, tandis que pour les fermés c'est une conséquence du corollaire 3.5.6.

3.5.1 Les ensembles convexes sont des intersections de demi-espaces

3.5.10 COROLLAIRE
Chaque sous-ensemble convexe fermé K d'un espace vectoriel normé X est l'intersection de tous les demi-espaces fermés qui contiennent K.

Rappelons qu'un demi-espace est ce qui se trouve d'un côté d'un hyperplan ; donc un demi-espaces est de la forme

$$\{X \in X : f(x) \leq a\}$$

pour $f \in X^*$, $a \in \mathbb{R}$. Voir la figure illustrant le corollaire.

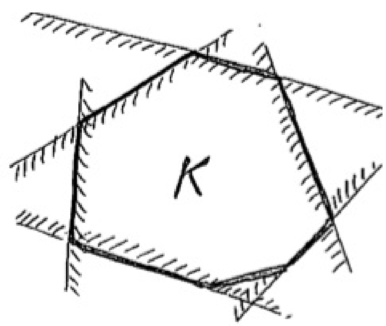


FIGURE 3.2 – l'ensemble convexe K est l'intersection de demi-espaces

Démonstration: K est trivialement contenu dans l'intersection des demi-espaces qui contiennent K. Pour prouver l'autre inclusion, on choisit un point $x_0 \notin K$ et on utilise le théorème de séparation 3.5.6 pour $A = K$ et $B = \{x_0\}$. On obtient ainsi une fonction $f \in X^*$ et on pose

$$a := \sup_{x \in K} f(x) < f(x_0).$$

Il s'ensuit que le demi-espace $\{x \in X : \ f(x) \leq a\}$ contient K, mais pas x_0. Ceci termine la preuve. ∎

3.5.2 Exercices supplémentaires

3.5.12 Exercice (Ensembles convexes fermés qui ne peuvent pas être strictement séparés) Montrer que l'hypothèse de compacité dans le corollaire 3.5.6 est essentielle. Construire deux ensembles convexes fermés du plan $\mathbb{R}^2$ qui ne peuvent pas être strictement séparés.

3.5.13 Exercice (Ensembles convexes qui ne peuvent pas être séparés) Montrer que l'hypothèse d'ouverture dans le théorème 3.5.3 est essentielle.

Pour cela, considérons l'espace vectoriel $\mathbb{R}[X]$ des polynômes à une variable et à coefficients réels. Soit le sous-ensemble A composé de polynômes à coefficient dominant négatif, et le sous-ensemble B constitué des polynômes dont tous les coefficients ne sont pas négatifs. Montrer que A et B sont des parties convexes disjointes de $\mathbb{R}[X]$, et qu'il n'y a pas exister une valeur non nulle linéaire f fonctionnel sur $\mathbb{R}[X]$ telle que

$$f(a) \le f(b) \quad \text{pour tout } a \in A, b \in B.$$

(Indication : supposons que pour environ $C \in \mathbb{R}$ *on a* $f(a) \le C \le f(b)$, $a \in A$, $b \in B$ *; déduire de* $0 \in B$ *que* $C \le 0$ *et en considérant les monômes que* $C \ge 0$.

3.6 Théorème de Krein-Milman

3.6.1 DÉFINITION

Un *sous-ensemble extrémal* S d'un convexe K est un sous-ensemble $S \subset K$ tel que

1. S est un convexe non vide.
2. Si $x \in S$ et $x = ty + (1-t)z$ avec $y, z \in K$ alors $y, z \in S$.

Un *point extrémal* de K est un point $x \in K$ tel que $\{x\}$ soit un sous-ensemble extrémal.

On rappel le théorème classique de Carathéodory :

3.6.2 THÉORÈME (DE CARATHÉODORY)

Tout convexe compact K de $\mathbb{R}^n$ a des points extrémaux, et tout point de K est une combinaison convexe d'au plus $n+1$ points extrémaux.

3.6.3 Exercice Démontrer le théorème de Carathéodory (utiliser une récurence sur n).

3.6.4 Théorème (de Krein et Milman)

Soit X un espace vectoriel normé, K un sous-ensemble compact, convexe non vide de X. Alors

1) K a au moins un point extrémal.

2) K est égal à l'adhérence de l'enveloppe convexe de ses points extrémaux.

Démonstration: On considère $\mathcal{E}$ l'ensemble des sous-ensembles extrémaux, fermés et non vides de K. Comme $K \in \mathcal{E}$ cet ensemble n'est pas vide. On ordonne $\mathcal{E}$ par l'inclusion inverse i.e. $A \leq B$ si est seulement si $B \subset A$. On veut appliquer le lemme de Zorn **??** pour montrer que $\mathcal{E}$ a un élément "maximal", i.e., un sous-ensemble qui est *minimal* par rapport à l'inclusion.

Soit $\mathcal{T} = \{A_i : i \in I\} \subset \mathcal{E}$ une chaîne i.e. $\mathcal{T}$ est une partie totalement ordonnée. On doit montrer que $\mathcal{T}$ admet un majorant.

Alors $A = \bigcap_{i \in I} A_i$ est le candidat "naturel". On doit montrer que $A \in \mathcal{E}$ i.e. A est un sous-ensembles extrémal, fermé et non-vide de K..

Clairement, A est un fermé, come intersection de fermés. On vérifie facilement qu'une intersection non-vide, de sous-ensembles extrémaux est un sous-ensemble extrémal de K.

Il ne reste plus qu'a montrer que A n'est pas vide.

Supposons que ce ne soit pas le cas, alors $\emptyset = A = \bigcap_{i \in I} A_i$, par suite $K = \bigcup_{i \in I} A_i^c$ est un recouvrement ouvert de K par les ouverts $A_i^c = K \setminus A_i$. Comme K est compact, on peut extraire un sous-recouvrement fini, soit $K = \bigcup_{i=1}^N A_i^c$. Par passage au complémentaire, on aura $\emptyset = A = \bigcap_{i=1}^n A_i$. Mais, comme $\mathcal{T}$ est une partie totalement ordonnée, toute intersection finie d'éléments de $\mathcal{T}$ est non-vide, d'où une contradiction.

Ainsi $A = \bigcap_{i \in I} A_i$ est un majorant de $\mathcal{T}$.

D'après le lemme de Zorn, $\mathcal{E}$ admet un élément maximal E. On va montrer que E est un singleton.

Sinon, E, a au moins 2 points (il doit contenir le segment qui les joint). D'après le corollaire 3.5.6, il existe une forme linéaire continue f sur X, qui sépare deux points de E. Comme E est compact, f atteint son maximum sur E.

Soit $M \subset E$ l'ensemble

$$M = \{x \in E \,:\, f(x) = \max_{z \in E} f(z)\}.$$

Alors M est un sous-ensemble fermé et non-vide de E et $M \neq E$. Il est clairement convexe et on vérifie facilement que c'est un sous-ensemble extrémal de E. D'où M est un sous-ensemble extrémal de K mais comme $M \subsetneq E$ alors E n'est pas un élément maximal de $\mathcal{E}$. D'où contradiction, ainsi E est un singleton.

Ceci montre que 1), K a au moins un point extrémal.

Soit K_e l'ensemble des points extrémaux de K et $\widehat{K}_e$ son enveloppe convexe, et $\check{K}_e$ l'adhérence de $\widehat{K}_e$. On vérifie facilement que $\check{K}_e$ est convexe. (Exercice)

Clairement, $\widehat{K}_e \subset K$, et comme K est fermé $\check{K}_e \subset K$.

Soit $z \notin \check{K}_e$ on va montrer que $z \notin K$.

Comme $z \notin \check{K}_e$, d'après la théorème 3.5.6 il existe une forme linéaire continue ℓ et $c \in \mathbb{R}$ tel que

$$\ell(y) \leq c < \ell(z) \quad \text{pour tout } y \in \check{K}_e.$$

Comme K est compact ℓ atteint son maximum sur K. Soit $E = \{x \in K : \ell(x) = \max_{t \in E} \ell(t)\}$. Alors E est fermé, convexe et extrémal. D'après ce qui précède E contient un point extrémal $p \in K_e$. Ainsi $\ell(p) \leq c$, mais $\ell(p) = \max_{x \in K} \ell(x)$, donc pour tout $x \in K$, $\ell(y) \leq c$. Comme $\ell(z) > c$, $z \notin K$. ∎

Chapitre 4

Espaces de Hilbert

Dans ce chapitre tous les espaces vectoriels sont sur le corps des complexes $\mathbb{C}$ ou réels $\mathbb{R}$. Pour $\lambda = a + ib \in \mathbb{C}$, $\bar{\lambda} = a - ib$ désigne son conjugué.

4.0.6 Définition

Un *espace préhilbertien est la donnée d'un espace vectoriel H sur le corps $\mathbb{K}$ (des nombres complexes $\mathbb{C}$ ou réels $\mathbb{R}$) et d'une application application de $H \times H \to \mathbb{K}$, $(x, y) \to \langle x, y \rangle$, appelée* **produit scalaire***, vérifiant les propriétés suivantes :*

(I1) $\langle y, x \rangle = \overline{\langle x, y \rangle}$. *(le produit est Hermitien)*

(I2) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ *pour tout $x, y, z \in H$.*

(I3) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ *pour tout $x, y \in H$ et $\alpha \in \mathbb{C}$.*

(I4) $\langle x, x \rangle \geq 0$ *pour tout $x \in H$.*

(I5) $\langle x, x \rangle = 0 \iff x = 0$.

4.0.7 Remarque

1. Une application qui vérifie les conditions (I2) et (I3) est appelée forme sesquilinéaire.
2. Une application qui vérifie les conditions (I1), (I2) et (I3) est appelée forme hermitienne.
3. Si de plus elle vérifie (I4) est on dit que c'est une forme hermitienne positive.
4. Si elle vérifie (I1), (I2), (I3), (I4) et (I5) on dit que c'est une forme hermitienne positive et non-dégénérée ou tout simplement un produit scalaire.
5. L'application de H dans $\mathbb{R}_+$ définie par $\|x\| = \sqrt{\langle x, x \rangle}$, est appelée la norme induite (du produit scalaire). On va montrer que c'est bien une norme.

On notera que $\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle$, on dit que le produit scalaire est sur un espace vectoriel complexe est linéaire par rapport à la première variable et anti-linéaire par rapport à la seconde variable. Sur $\mathbb{R}$, on a la symétrie $\langle x, y \rangle = \langle y, x \rangle$, et $(H, \langle \cdot, \cdot \rangle)$ est appelé espace préhilbertien réel.

4.0.8 EXEMPLE. *L'espace $\mathbb{R}^n$ est un espace préhilbertien réel, où le produit scalaire sur $\mathbb{R}^n$ est défini par*

$$\langle x, y\rangle = \sum_{k=1}^{n} \xi_k \eta_k, \tag{4.0.1}$$

pour tout $x = (\xi_1, \cdots, \xi_n), y = (\eta_1, \cdots, \eta_n) \in \mathbb{R}^n$. Le produit scalaire défini par(4.0.1) est appelé produit scalaire stetard sur $\mathbb{R}^n$.

4.0.9 EXEMPLE. *L'espace $\mathbb{C}^n$ est un espace préhilbertien où le produit scalaire est défini par*

$$\langle x, y\rangle = \sum_{k=1}^{n} \xi_k \bar{\eta}_k, \tag{4.0.2}$$

pour tout $x = (\xi_1, \cdots, \xi_n), y = (\eta_1, \cdots, \eta_n) \in \mathbb{C}^n$. Le produit scalaire défini par (4.0.2) est le produit scalaire stetard sur $\mathbb{C}^n$.

4.0.10 EXEMPLE. Soit $A = (a_{ij})_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{C})$ une matrice hermitienne, i.e. $a_{ij} = \overline{a_{ji}}$.

Alors l'application f_A de $\mathbb{C} \times \mathbb{C}$ dans $\mathbb{C}$ définie par $f_A(x, y) = \langle Ax, y\rangle$ est une forme hermitienne.

- f_A est positive si et seulement si les valeurs propres de A sont positives.
- f_A est un produit scalaire si et seulement si les valeurs propres de A sont strictement positives.

4.0.11 EXEMPLE. Soit $\ell^2(\mathbb{N}, \mathbb{C}) = \{x = \{\xi_k\}_{k\in\mathbb{N}} : \xi_k \in \mathbb{C}, \sum_{k=1}^{\infty} |\xi_k|^2 < \infty\}$.

D'après l'inégalité $|\xi_k\bar{\eta}_k| \leq |\xi_k|^2 + |\bar{\eta}_k|^2$,

$$\sum_{k=1}^{\infty} |\xi_k\bar{\eta}_k| \leq \sum_{k=1}^{\infty} |\xi_k|^2 + \sum_{k=1}^{\infty} |\bar{\eta}_k|^2$$

pour tout $x = \{\xi_k\}, y = \{\eta_k\} \in \ell^2(\mathbb{N})$, on en déduit que la série $\sum_{k=1}^{\infty} \xi_k\bar{\eta}_k$ converge absolument. Ceci nous permet de définir

$$\langle x, y\rangle = \sum_{k=1}^{\infty} \xi_k \bar{\eta}_k,$$

il est facile de voir que $\langle x, y\rangle$ est un produit scalaire, qu'on appelera le produit scalaire stetard de $\ell^2(\mathbb{N}, \mathbb{C})$. Ainsi $\ell^2(\mathbb{N})$ est un espace préhilbertien .

La norme induite par ce produit scalaire est $\|x\|_2 = \left(\sum_{k=1}^{\infty} |\xi_k|^2\right)^{1/2}$.

4.0.12 EXEMPLE. L'espace de Lebesgue $L^2([a,b])$ des classes d'équivalences de fonctions mesurables $x : [a,b] \to \mathbb{C}$, pour la relation d'équivalence définie d'égalité presque partout, et telle que $\int_a^b |x(t)|^2\,dt < \infty$, est un espace préhilbertien, muni du produit scalaire

$$(x,y) \mapsto \langle x,y\rangle = \int_a^b x(t)\,\overline{y(t)}\,dt.$$

D'après l'inégalité $|x(t)\,\overline{y(t)}| \leq |x(t)|^2 + |\overline{y(t)}|^2$

$$\int_a^b |x(t)\,\overline{y(t)}|\,dt \leq \int_a^b |x(t)|^2\,dt + \int_a^b |\overline{y(t)}|^2\,dt < \infty$$

pour tout $x, y \in L^2([a,b])$, on peut alors définir

$$\langle x,y\rangle = \int_a^b x(t)\,\overline{y(t)}\,dt,$$

il est alors facile de vérifier que $\langle x,y\rangle$ est un produit scalaire, qu'on appelera produit scalaire stetard de $L^2([a,b])$. Ainsi $L^2([a,b])$ est un espace préhilbertien . Ce produit scalaire induit la norme $\|x\|_2 = \left(\int_a^b |x(t)|^2\,dt\right)^{1/2}$.

Deux vecteurs x et y d'un espace préhilbertien H sont dit **orthogonaux** *si $\langle x,y\rangle = 0$, on note ceci par $x \perp y$.*

Pour deux vecters orthogonaux x et y on a (Théorème de Pythagore) :
$\|x+y\|^2 = \|x\|^2 + \|y\|^2$

Soit $y \in H$ un vecteur unitaire ($\|y\| = 1$), la projection de $x \in H$ le long de y est donnée par $x_\parallel = \langle x,y\rangle y$ et on pose $x_\perp = x - x_\parallel$.

Alors, $x_\perp$ est orthogonal à $x_\parallel$, car

$$\langle x_\perp, x_\parallel\rangle = \langle x - \langle x,y\rangle y, \langle x,y\rangle y\rangle = \overline{\langle x,y\rangle}\langle x,y\rangle - \langle x,y\rangle\overline{\langle x,y\rangle}\langle y,y\rangle = 0$$

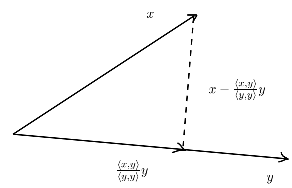
Ainsi d'après le théorème de Pythagore :

$$\|x\|^2 = \|x_\parallel + x_\perp\|^2 = \|x_\parallel\|^2 + \|x_\perp\|^2 = |\langle x,y\rangle|^2 + \|x_\perp\|^2$$

d'où on a $\|x\| \geq |\langle x,y\rangle|$. Comme conséquence on obtient l'inégalité de Cauchy-Schwarz :

4.0.13 THÉORÈME (INÉGALITÉ DE CAUCHY-SCHWARZ)
Soit H un espace préhilbertien. Alors pour tout $x, y \in H$ on a :

$$|\langle x,y\rangle| \leq \|x\|.\|y\|$$



Démonstration: Si $y = 0$ l'inégalité est évidente. Sinon, on pose $z = \frac{y}{\|y\|}$ et de ce qui précède on aura $\|x\| \geq |\langle x, z\rangle| = \frac{|\langle x,y\rangle|}{\|y\|}$. ■

4.0.15 Remarque

La différence entre un espace préhilbertien réel et complexe est dans la conjugaison (I1) de la définition 4.0.6. Mise à part la présence de nombre complexe conjugés, la plupart des résultats et démonstrations sont valables pour les réels et les complexes.

Nous considérerons, sauf mention contraire, que les espaces vectoriels sont complexes, en sous-entendant que la partie imaginaire doit être ignorée si l'espace est réel.

4.0.16 Proposition

Tout sous-espace vectoriel d'un espace préhilbertien est un espace préhilbertien

Démonstration: Il suffit de voir que les propriétés (I1)-(I5) sont encore satisfaites par la restriction du produit scalaire à un sous-espace vectoriel. ■

4.0.18 Théorème

Soit H un espace préhilbertien sur $\mathbb{K}$, et $x, y \in H$. Alors

(1) si $\langle x, z\rangle = \langle y, z\rangle$ ou $\langle z, x\rangle = \langle z, y\rangle$ pour tout $z \in H$, alors $x = y$;
(2) la fonction $\|\cdot\| : H \to \mathbb{R}_+$ définie par $\|x\| = \sqrt{\langle x, x\rangle}$, est une norme sur H.

Proof : *(1) D'après la définition 4.0.6,*

$$0 = \langle x, z\rangle - \langle y, z\rangle = \langle x, z\rangle + (-1)\langle y, z\rangle = \langle x, z\rangle + \langle -y, z\rangle = \langle x - y, z\rangle$$

pour tout $z \in H$, en particulier pour $z = x - y$, d'où $\langle x - y, x - y\rangle = 0$. Il résulte de (I1) que $x - y = 0$. De même $\langle z, x\rangle = \langle z, y\rangle$ $(\forall\, z \in H) \Longrightarrow x = y$.

(2) On va vérifier les axiomes d'une norme, d'après les propriétés du produit scalaire on a :

1° $\|x\| = \sqrt{\langle x, x\rangle} \geq 0$, *et* $\|x\| = 0 \Longleftrightarrow \sqrt{\langle x, x\rangle} = 0 \Longleftrightarrow x = 0$;
2° $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha\overline{\alpha}\langle x, x\rangle} = \sqrt{|\alpha|^2\langle x, x\rangle} = |\alpha|\sqrt{\langle x, x\rangle} = |\alpha|\|x\|$;

3° *Comme*

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \|x\|^2 + 2\mathfrak{Re}\langle x,y\rangle + \|y\|^2 \\
&\leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \quad \textit{(d'après l'inégalité de Cauchy-Schwarz (2))} \\
&= (\|x\| + \|y\|)^2,
\end{aligned}$$

d'où l'inégalité triangulaire. ∎

4.0.19 Remarque

La norme $\|x\| = \sqrt{\langle x,x\rangle}$ *définie dans le théorème précédent sur le préhilbertien* H *est appelé norme induite par le produit scalaire* $\langle\cdot,\cdot\rangle$. *Le théorème montre que tout espace préhilbertien peut être considéré comme un espace vectoriel normé. Dorénavant les propriétés topologiques d'un espace préhilbertien sont entendues pour la norme associée.*

Les espaces préhilbertiens sont des espaces vectoriels normés particuliers. Comme tout espace préhilbertien est un espace vectoriel normé, il est natural de se demeter sous quelles conditions un espace normé est un espace préhilbertien, c-à-d sa norme est induite par un produit scalaire. Le résultat suivant nous donne une condition nécessaire.

4.0.20 Lemme

Soit H un espace préhilbertien muni du produit scalaire $\langle\cdot,\cdot\rangle$. alors pour tout $x, y \in H$ on a :

(a) *(l'identité du parallélogramme)*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2);$$

(b) *(l'identité de polarisation)* Si H est réel alors

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right);$$

(c) *(l'identité de polarisation)* Si H est complexe alors

$$\langle x,y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k\|x+i^k y\|^2.$$

Démonstration: Comme,

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \tag{4.0.3}$$

et

$$\|x-y\|^2 = \langle x-y, x-y\rangle = \|x\|^2 - \langle x,y\rangle - \langle y,x\rangle + \|y\|^2, \tag{4.0.4}$$

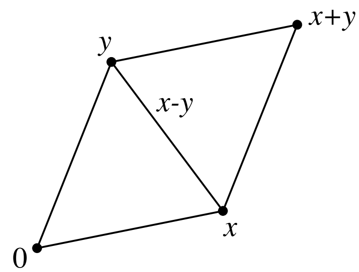


FIGURE 4.1 – l'identité du parallélogramme

d'où l'identité (a) par (4.0.3) + (4.0.4). i.e. dans le parallélogramme formé par les vecteurs x et y, la somme des carrés des diagonales est égale à la somme des carrés des côtés.

D'autre part, (4.0.3) − (4.0.4) $\Longrightarrow$

$$\|x+y\|^2 - \|x-y\|^2 = 2(\langle x,y\rangle + \langle y,x\rangle), \tag{4.0.5}$$

Alors si E est réel, on a (b). Si E est complexe, en remplçant y dans (4.0.5) par iy, on aura

$$\|x+iy\|^2 - \|x-iy\|^2 = 2(\langle x,iy\rangle + \langle iy,x\rangle). \tag{4.0.6}$$

finalement, (4.0.5) et (4.0.6) entraîne

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

d'où (c). ■

4.0.22 THÉORÈME (VON NEUMANN)
Soit $(E, \|\cdot\|)$ un espace vectoriel normé.

E est un espace préhilbertien si et seulement si sa norme vérifie l'identité du parallélogramme 4.0.20

Démonstration: (voir le TD) .

4.0.24 EXEMPLE. Soit $\mathbb{C}^n = \{z = (z_1, \ldots, z_n) : z_i \in \mathbb{C}\}$ l'espace vectoriel complexe de dimension n, muni de la norme

$$\|z\|_1 = \sum_{i=1}^{n} |z_i|.$$

Montrer que cette norme, n'est pas issue d'un produit scalaire.

Démonstration: Si $\|\cdot\|_1$ est issue d'un produit scalaire, alors l'identité du parallélogramme, doit être vérifiée, pour tout $x, y \in \mathbb{C}^n$. Soit H tel que $x_1 = 1$ et les autres coordonnées sont nulles. Soit y tel que $y_2 = 1$ et et les autres coordonnées sont nulles . Alors

$$\|x\|_1 = \|y\|_1 = 1,$$

et

$$\|x+y\|_1 = \|x-y\|_1 = 2.$$

Comme $8 \neq 4$ l'identité du parallélogramme n'est pas satisfaite. ∎

4.0.26 EXEMPLE. Soient $a < b$ des réels. On munit l'espace des fonctions continues sur $[a, b]$, $C([a, b])$, de sa norme stetard $\|.\|_\infty$, de la convergence uniforme, définie par $\|x\|_\infty = \max_{a \le t \le b} |x(t)|$.

La norme $\|.\|_\infty$ sur $C([a.b])$ n'est pas induite par un produit scalaire.

En effet, on considère les fonctions $x, y \in C([a, b])$ définies par $x(t) = 1$, $y(t) = \frac{t-a}{b-a}$, $t \in [a, b]$. Alors

$$\|x\|_\infty = \|y\|_\infty = 1$$

et

$$\|x+y\|_\infty = \max_{a \le t \le b} \left|1 + \frac{t-a}{b-a}\right| = 2$$

$$\|x-y\|_\infty = \max_{a \le t \le b} \left|1 - \frac{t-a}{b-a}\right| = 1,$$

Alors $\|x+y\|_\infty^2 + \|x-y\|_\infty^2 \neq 2(\|x\|_\infty^2 + \|y\|_\infty^2)$, donc l'identité du parallélogramme n'est pas satisfaite et par suite la norme ne peut être induite par un produit scalaire.

4.0.27 EXEMPLE. *The space $\ell^p(\mathbb{N})$ ($p \ge 1, p \neq 2$) n'est pas un espace préhilbertien,*

où $\ell^p(\mathbb{N}) = \{x = \{\xi_k\} : \xi_k \in \mathbb{C}, k \in \mathbb{N}, \sum_{k=1}^{\infty} |\xi_k|^p < \infty\}$ muni de la norme

$\|x\|_p = (\sum_{k=1}^{\infty} |\xi_k|^p)^{1/p}$ pour tout $x = \{x_k\} \in \ell^p$.

En effet, on prend $x = (1, 1, 0, 0, \cdots) \in \ell^p$, $y = (1, -1, 0, 0, \cdots) \in \ell^p$ et on calcule

$$\|x\|_p = \|y\|_p = 2^{1/p}, \qquad \|x+y\|_p = \|x-y\|_p = 2.$$

Donc l'identité du parallélogramme n'est pas satisfaite, par suite ℓ^p n'est pas un espace. Qu'en est-il de $L^p([0, 1]) = \{x : [0, 1] \to \mathbb{K} : \int_0^1 |x(t)|^p \, dt < \infty\}$?

4.0.28 EXEMPLE. Montrer que l'inégalité du parallélogramme n'est pas vérifiée pour $L^1([0,1])$ ce n'est donc pas un espace de préhilbertien.

Démonstration: Soit $f = \mathbb{1}_{[0,1/2]}$ et $g = \mathbb{1}_{[1/2,1]}$. Alors $\|f\|_1 = \|g\|_1 = \frac{1}{2}$, et $\|f+g\|_1 = \|f-g\|_1 = 1$. Mais

$$1^2 + 1^2 = 2 \neq 1 = 2\left(\frac{1}{2}\right)^2 + 2\left(\frac{1}{2}\right)^2.$$

Comme la norme est une fonction continue, le produit scalaire est alors aussi une fonction continue de l'espace produit $H \times H$.

4.0.30 THÉORÈME
Soit H un espace préhilbertien , $\{x_n\}$ et $\{y_n\}$ deux suites de H qui convergent respectivement, vers $x \in H$ et $y \in H$. Alors $\langle x_n, y_n\rangle \to \langle x, y\rangle$ quet $n \to \infty$.

Démonstration: D'après l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &\leq |\langle x_n, y_n\rangle - \langle x, y_n\rangle| + |\langle x, y_n\rangle - \langle x, y\rangle| \\ &= |\langle x_n - x, y_n\rangle| + |\langle x, y_n = y\rangle| \\ &\leq \|x_n - x\| \, \|y_n\| + \|x\| \, \|y_n - y\|. \end{aligned}$$

Comme $\{y_n\}$ est convergente, $\{\|y_n\|\}$ est bornée. D'où $\langle x_n, y_n\rangle \to \langle x, y\rangle$ quet $n \to \infty$.

On peut aussi, en utilisant l'identité de polarisation, voir que le produit scalaire est une fonction continue car il s'écrit comme une somme finie de normes. ∎

Les espace préhilbertiens les plus importants sont ceux qui sont complets (en tant qu'espaces métriques), ces espaces sont appelés des espaces de Hilbert.

4.0.32 DÉFINITION
Un espace préhilbertien H est un **espace de Hilbert** s'il est complet par rapport à la distance induite.

4.0.33 EXEMPLE. *Les espaces $\mathbb{R}^n$ et $\mathbb{C}^n$ sont des espaces de Hilbert.*

Plus généralement, tout espace préhilbertien de dimension finie est un espace de Hilbert.

4.0.34 EXEMPLE. *L'espace de suites* $\ell^2(\mathbb{N}) = \{x = \{\xi_k\}_{k\in\mathbb{N}} : \xi_k \in \mathbb{C},, \sum_{k=1}^{\infty} |\xi_k|^2 < \infty\}$ *est un espace de Hilbert.*

Soit $x^n = (x_k^n)_{k\in\mathbb{N}}$ *une suite de Cauchy dans* $\ell^2(\mathbb{N})$.
Soit $\varepsilon > 0$, *il existe alors un entier* $N(\varepsilon) > 0$ *tel que pour* $m, p \geq N(\varepsilon)$ *on ait* $\|x_k^m - x_k^p\|_2 \leq \varepsilon$, *i.e.*

$$\sum_{k=0}^{+\infty} |x_k^m - x_k^p|^2 \leq \varepsilon^2 \tag{4.0.7}$$

Ainsi, pour tout $k \in \mathbb{N}$ *et tout* $m, p \geq N(\varepsilon)$, $|x_k^m - x_k^p| \leq \varepsilon$, *ce qui veut dire que pour tout* $k \in \mathbb{N}$, *la suite* $(x_k^m)_{m\in\mathbb{N}}$ *est une suite de Cauchy, par la complétude de* $\mathbb{C}$, *elle converge vers un point* $x_k \in \mathbb{C}$.

On pose $x = (x_k)_{k\in\mathbb{N}}$. *D'après* 4.0.7, *pour tout* $M \in \mathbb{N}$ *et* $m, p \geq N(\varepsilon)$, *on a*

$$\sum_{k=0}^{M} |x_k^m - x_k^p|^2 \leq \sum_{k=0}^{+\infty} |x_k^m - x_k^p|^2 \leq \varepsilon^2$$

En passant à la limite lorsque $m \to +\infty$, *on obtient*

$$\sum_{k=0}^{M} |x_k - x_k^p|^2 = \sum_{k=0}^{M} \lim_{m\to+\infty} |x_k^m - x_k^p|^2 = \lim_{m\to+\infty} \sum_{k=0}^{M} |x_k^m - x_k^p|^2 \leq \varepsilon^2$$

et finalement en faisant tendre M *vers* $+\infty$ *on obtient : pour tout* $p \geq N(\varepsilon)$

$$\sum_{k=0}^{+\infty} |x_k - x_k^p|^2 \leq \varepsilon^2$$

i.e. $\|x - x^p\|_2 \leq \varepsilon$ *d'où* $x - x^p \in \ell^2(\mathbb{N})$
Comme $\epsilon > 0$ *est arbitraire, on aura* $\lim_{n\to+\infty} x^n = x$.

En conclusion, on a $\begin{cases} x = x^{N(\varepsilon)} + (x - x^{N(\varepsilon)}) \in \ell^2(\mathbb{N}) \\ \lim_{n\to+\infty} x^n = x \end{cases}$.

4.0.35 EXEMPLE. *L'espace* $L^2([a,b])$ *est un espace de Hilbert, conséquence du théorème de Riesz* 1.5.10

4.0.36 EXEMPLE. *Soit* $C([-1,1])$ *l'espace vectoriel des fonctions continues à valeurs dans* $\mathbb{C}$. *Pour tout* $x, y \in H$, *on pose*

$$\langle x, y\rangle = \int_{-1}^{1} x(t)\,\overline{y(t)}\,dt,$$

Alors $\langle x, y\rangle$ *est un produit scalaire, de norme induite* $\|x\|_2 = \left(\int_{-1}^{1} |x(t)|^2\,dt\right)^{1/2}$. *H est un espace préhilbertien mais pas un espace de Hilbert.*

4.0.37 Exercice Trouver une suite de cauchy de $C([-1,1])$ pour la norme $\|.\|_2$ qui ne converge pas.

4.0.38 DÉFINITION
Soit $(E, \|.\|_E$ et $(F, \|.\|_F)$ deux espaces vectoriels normés. Une application $\Phi : E \to F$ est une **isométrie** si : pour tout $x, y \in E$ on a

$$\|\Phi(x) - \Phi(y)\|_F = \|x - y\|_E.$$

4.0.39 REMARQUE
Si Φ est de plus linéaire, la condition précédente est équivalente à, pour tout $x \in E$, $\|\Phi(x)\|_F = \|x\|_E$.

4.0.40 DÉFINITION
Soit $(H_1, \langle,\rangle_1$ et $(H_2, \langle,\rangle_2)$ deux espaces préhilbertiens.

1. Une application linéaire $\Phi : H_1 \to H_2$ est une **isométrie** si : pour tout $x \in H_1$, $\|\Phi(x)\|_2 = \|x\|_1$.
2. L'application est dite **unitaire** si elle est en plus bijective.

4.0.41 Exercice Montrer qu'une application linéaire $\Phi : H_1 \to H_2$ entre espaces préhilbertiens est une isométrie si et seulement si pour tout $x, y \in H_1$ on a

$$\langle \Phi(x), \Phi(y)\rangle_2 = \langle x - y\rangle_1.$$

Tout espace préhilbertien peut être plongé dans un espace de Hilbert par complétion.

4.0.42 THÉORÈME
Soit $(H, \langle,\rangle)$ un espace préhilbertien.

Il existe un espace de Hilbert $(\mathcal{H}, \langle,\rangle)$ et une isométrie linéaire $\Phi : H \to \mathcal{H}$ tels que $\Phi(H)$ soit dense dans $\mathcal{H}$.

On dit que $\mathcal{H}$ est un complété de H. Ce complété est unique à isomorphisme isométrique près.

Démonstration: C'est une conséquence des théorèmes de complétion d'un espace métrique et de l'unicité du prolongement des applications uniformément continues, voir 1.6.10. ∎

4.1 Orthogonalité

D'après l'inégalité de Cauchy-Schwarz pour des espaces réels, si x, y sont tous les deux non nuls, alors

$$-1 \leq \frac{\langle x, y\rangle}{\|x\|\,\|y\|} \leq 1,$$

et donc l'angle entre x et y peut être défini

$$\theta = \cos^{-1}\left(\frac{\langle x, y\rangle}{\|x\|\,\|y\|}\right).$$

Pour les espaces complexes, le problème est plus difficile, comme le produit scalaire $\langle x, y\rangle$ est à valeurs complexes il n'est pas clair ce que "angle" signifie dans ce cas. Néanmoins, un cas particulier, très important, peut être considéré, à savoir le cas $\langle x, y\rangle = 0$.

4.1.1 Définition

Soit H un espace préhilbertien sur $\mathbb{K}$. Deux vecteurs $x, y \in H$ sont dit orthogonaux si $\langle x, y\rangle = 0$, et qu'on note $x \perp y$.

On va généraliser cette notion, à des ensembles.

4.1.2 Définition

Soit H un préhilbertien et soit A un sous-ensemble de H. un vecteur $x \in H$ et l'ensemblet A sont dit orthogonaux dans H si x est orthogonal à tout vecteur de A, qu'on note $x \perp A$. Le complémentaire orthogonal de A est l'ensemble

$$A^{\perp} = \{x \in H : x \perp A\} (= \{x \in H : x \perp a \,\forall\, a \in A\} = \{x \in H : \langle x, a\rangle = 0, \,\forall\, a \in A\}),$$

i.e. l'ensemble $A^{\perp}$ formé des vecteurs de H qui sont orthogonaux à tout vecteur de A (Si $A = \emptyset$ alors $A^{\perp} = H$).

Soient M et N deux sous-ensembles de H. M et N sont orthogonaux si pour tout $x \in M, y \in N$, $x \perp y$, et qu'on note $M \perp N$.

4.1.3 Exemple. *Si $E = \mathbb{R}^3$ et $A = \{(a_1, a_2, 0) : a_1, a_2 \in \mathbb{R}\}$, alors $A^{\perp} = \{(0, 0, x_3) : x_3 \in \mathbb{R}\}$.*

4.1.4 Proposition

Soit E est un préhilbertien alors :

(a) Pour tout $x, y \in E$

1. $x \perp y \Leftrightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2$; si $\mathbb{K} = \mathbb{R}$ *(Le théorème de Pythagore)*

2. $x \perp y \Leftrightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$ et $\|x+iy\|^2 = \|x\|^2 + \|y\|^2$; si $\mathbb{K} = \mathbb{C}$

(b) $0 \in A^\perp$.

(c) Si $0 \in A$ alors $A \cap A^\perp = \{0\}$, autrement $A \cap A^\perp = \emptyset$.

(d) $\{0\}^\perp = E$, $E^\perp = \{0\}$.

(e) Si A contient une boule ouverte $B(a,r)$ pour un certain $a \in E$ et $r > 0$, alors $A^\perp = \{0\}$; en particulier, si A est un ouvert non vide alors $A^\perp = \{0\}$.

(f) Si A est dense dans E, i.e. $\bar{A} = E$, alors $A^\perp = \{0\}$.

(g) Si $B \subset A$, alors $A^\perp \subset B^\perp$.

(h) $A^\perp$ est un sous-espace fermé de E.

(i) $A \subset (A^\perp)^\perp$.

Démonstration: (a) Comme $\|z\|^2 = \|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 = \|x\|^2 + \|y\|^2$ car $x \perp y$ (i.e. $\langle x,y\rangle = \langle y,x\rangle = 0$).

(b) Comme $\langle 0, a\rangle = 0$ pour tout $a \in A$, on a $0 \in A^\perp$.

(c) D'après (b), $\{0\} \subset A \cap A^\perp$. On suppose que $x \in A \cap A^\perp$. Alors $\langle x,x\rangle = 0$ et donc $x = 0$ d'après la définition du produit scalaire.

(d) Si $A = \{0\}$, $x \in E$ on a $\langle x,a\rangle = \langle x,0\rangle = 0$ pour tout $a \in A$, d'où $x \in A^\perp$ et par suite $A^\perp = E$. Si $A = E$ et $x \in A^\perp$ alors $\langle x,a\rangle = 0$ pour tout $a \in X$. En particulier, en posant $a = x$ donne $\langle x,x\rangle = 0$, qui entraîne que $x = 0$, donc $X^\perp \subset \{0\}$ et $X^\perp = \{0\}$.

(e) Soit $x \in A^\perp$. Si $x = 0$ il n'y arien à faire. Si $x \neq 0$, on note que $A \supset B(a,r)$, alors $a \in A$ qui donne $\langle x,a\rangle = 0$. Mais on a aussi $a + \frac{rx}{2\|x\|} \in B(a,r) \subset A$, on doit donc avoir

$$0 = \left\langle x, a + \frac{rx}{2\|x\|} \right\rangle = \langle x,a\rangle + \frac{r}{2\|x\|}\langle x,x\rangle,$$

qui entraîne $r = 0$, qui est contradictoire. Ainsi $A^\perp = \{0\}$.

(f) Supposons que $x \in A^\perp$. Comme $\bar{A} = X$, il s'ensuit qu'il existe une suite $\{x_n\}$ de A telle que $x_n \to x$ as $n \to \infty$. Notons que $\langle x, x_n\rangle = 0$ pour tout $n \in \mathbb{N}$, d'où, par continuité du produit scalaire (cf. Theorem 4.0.30),

$$\langle x,x\rangle = \lim_{n\to\infty} \langle x, x_n\rangle = 0,$$

et alors, $x = 0$.

(g) Soit $x \in B^\perp$ et $a \in A$. Alors $a \in B$ (car $A \subset B$), d'où $\langle x,a\rangle = 0$ pour tout $a \in A$, on aurra $x \in A^\perp$, donc $B^\perp \subset A^\perp$. Si $B \subset A$, alors $A^\perp \subset B^\perp$.

(h) Soit $x, y \in A^\perp$, $\alpha, \beta \in \mathbb{K}$. Alors

$$\langle \alpha x + \beta y, a\rangle = \alpha\langle x,a\rangle + \beta\langle y,a\rangle = 0,$$

et $\alpha x + \beta y \in A^\perp$, et donc $A^\perp$ est un sous-espace vectoriel de E. Soit $\{x_n\}$ une suite de $A^\perp$ telle que

$\{x_n\}$ converge vers $x \in E$. La continuité nous donne

$$\langle x, a\rangle = \lim_{n\to\infty} \langle x_n, a\rangle = 0.$$

pour tout $a \in A$. Ainsi $x \in A^\perp$ et $A^\perp$ est fermé.

(i) Soit $a \in A$. Pour tout $x \in A^\perp$, $\langle a, x\rangle = \overline{\langle x, a\rangle} = 0$, d'où $a \in (A^\perp)^\perp$ c-à-d $A \subset (A^\perp)^\perp$. ■

4.1.6 Proposition

Soit F un sous-espace vectoriel d'un préhilbertien E. Alors $x \in F^\perp$ si et seulement si $\|x - y\| \geq \|x\|$ pour tout $y \in F$.

Démonstration: ($\Rightarrow$) Soit $x \in F^\perp$, pour tout $y \in F$, $x \perp y$, et $x \perp (-y)$. D'après (a) du Lemme 4.1.4,

$$\|x - y\|^2 = \|x\|^2 + \|(-1)y\|^2 = \|x\|^2 + \|y\|^2 \geq \|x\|^2.$$

($\Leftarrow$) Supposons que $\|x - y\|^2 \geq \|x\|^2$ pour tout $y \in F$. Soit $y \in F$. Si $y = 0$ alors $\langle x, y\rangle = 0$. Si $y \neq 0$, alors $\alpha y \in F$ pour tout $\alpha \in \mathbb{K}$ car F est un sous-espace vectoriel. Ainsi

$$\begin{aligned}\|x\|^2 &\leq \|x - \alpha y\|^2 \\ &= \langle x - \alpha y, x - \alpha y\rangle \\ &= \|x\|^2 + |\alpha|^2\|y\|^2 - \alpha\overline{\langle x, y\rangle} - \overline{\alpha}\langle x, y\rangle.\end{aligned}$$

En posant $\alpha = \langle x, y\rangle / \|y\|^2$ dans l'inégalité, on aura

$$\frac{|\langle x, y\rangle|^2}{\|y\|^4}\|y\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2} - \frac{|\langle x, y\rangle|^2}{\|y\|^2} \geq 0,$$

qui donne $\langle x, y\rangle = 0$ pour tout $y \in F \backslash \{0\}$, alors pour tout $y \in F$. D'où $x \in F^\perp$. ■

4.2 Projection hilbertienne et conséquences

4.2.1 Définition

Un sous-ensemble C d'un espace vectoriel E est dit convexe si, pour tout $x, y \in C$ aet tout $\lambda \in [0, 1]$, on a $\lambda x + (1 - \lambda)y \in C$.

En d'autre termes, C est convexe si, pour tout $x, y \in C$, le segment dans E joignant x est y est contenu dans C.

4.2.2 Théorème (Projection sur un convexe fermé)

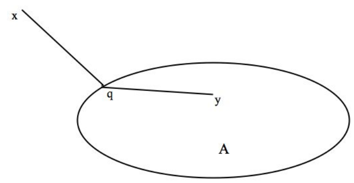
Soit H un espace de Hilbert et C un convexe fermé non vide de H.

Alors pour tout $x \in H$, il existe un **unique** $q \in C$, noté $P_C(x)$, tel que

$$\|x-q\| = d(x,C) = \inf_{a\in C} \|x-a\|.$$

De plus, $P_C(x)$ est caractérisé par les propriétés

$$\begin{cases} P_C(x) \in C \\ \mathfrak{Re}\langle x - P_C(x), y - P_C(x)\rangle \leq 0 \quad \forall y \in C \end{cases} \tag{4.2.1}$$



Cette inégalité traduit le fait que l'angle entre les vecteurs $x - P_C(x)$ et $y - P_C(x)$ est obtus.

Démonstration: Soit $\gamma = \inf\{\|x-a\| : a \in C\}$. Par définition de γ, il existe une suite $\{q_n\}$ de C telle que

$$\gamma_n = \|x - q_n\| \to \gamma \quad (n \to \infty). \tag{4.2.2}$$

Comme C est convexe, il s'en suit que

$$\frac{q_n + q_m}{2} \in C \quad \text{pour tout } n, m \in \mathbb{N}.$$

D'où

$$\left\| x - \frac{q_n+q_m}{2} \right\| \geq \gamma \quad \text{pour tout } n, m \in \mathbb{N}.$$

L'identité du parallélogramme appliquée à $x - q_n$ et $x - q_m$, nous donne

$$\begin{aligned} \|q_n - q_m\|^2 &= \|(q_n - x) + (x - q_m)\|^2 \\ &= 2\|q_n - x\|^2 + 2\|x - q_m\|^2 - \|(q_n - q_m) - 2x\|^2 \\ &= 2\|q_n - x\|^2 + 2\|x - q_m\|^2 - 4\left\| x - \frac{q_n+q_m}{2}\right\|^2 \\ &\leq 2\gamma_n^2 + 2\gamma_m^2 - 4\gamma^2. \end{aligned} \tag{4.2.3}$$

Ainsi $\{q_n\} \subset C$ est une suite de Cauchy (4.2.2), et donc converge vers un point $q \in C$ car C est un fermé de H, donc est lui même complet. Par suite $\|x - q\| \geq$

γ.D'après la continuité de la norme (4.2.2), on aura $\|x-q\| = \gamma$. Ce qui donne l'existence de q .

Pour l'unicité on suppose qu'un $w \in C$ vérifie $\|x-w\| = \gamma$. Alors $(q+w)/2 \in C$ car C est convexe, et donc $\|x - \frac{1}{2}(q+w)\| \geq \gamma$. L'identité du parallélogramme appliquée à $x-w$ et $x-q$ donne

$$0 \leq \|q-w\| \leq 4\gamma^2 - 4\gamma^2 = 0$$

Alors $w = q$, d'où l'unicité.

Enfin, étant donné $y \in C$, pour tout $\lambda \in]0,1]$ le vecteur $(1-\lambda)P_C(x) + \lambda y - x$, de sorte que $0 \leq \|(1-\lambda)P_C(x) + \lambda y - x\|^2 - \|P_C(x) - x\|^2$
$= \|(P_C(x) - x) + \lambda(y - P_C(x))\|^2 - \|P_C(x) - x\|^2$
$= -2\lambda\mathfrak{Re}\langle x - P_C(x), y - P_C(x)\rangle + \lambda^2\|y - P_C(x)\|^2$ d'où $\mathfrak{Re}\langle x - P_C(x), y - P_C(x)\rangle \leq \frac{\lambda}{2}\|y - P_C(x)\|^2$.

En faisant tendre λ vers 0^+, on en déduit que $\mathfrak{Re}\langle x - P_C(x), y - P_C(x)\rangle \leq 0$.

Inversement, si $\mathfrak{Re}\langle x-w, y-w\rangle \leq 0, \forall y \in C$, pour un certain $w \in C$, alors on obtient $\|(1-\lambda)w + \lambda y - x\|^2 - \|w-x\|^2 \geq 0$ et en faisant tendre λ vers 1^-, on obtient $\|y-x\|^2 - \|w-x\|^2 \geq 0, \forall y \in C$, c-à-d que $w = P_C(x)$.

En particulier si $x = 0$ on a :

4.2.4 Corollaire

Tout ensemble convexe fermé d'un espace de Hilbert a un **unique** élément de norme minimale.

Si F est un sous-espace vectoriel de H, alors F est convexe, d'où

4.2.5 Corollaire (Théorème de la projection hilbertienne)

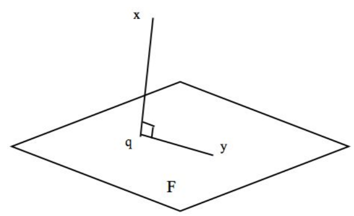
Soit F un sous-espace vectoriel fermé de l'espace de Hilbert H et soit $x \in H$.

Alors, il existe un **unique** $q \in F$, noté $P_F x$, tel que

$$\|x-q\| = d(x,F) := \inf_{a \in F} \|x-a\|.$$

De plus, $P_F x$ est caractérisé par les propriétés

$$\begin{cases} P_F x \in V \\ \langle x - P_F x, y\rangle = 0 \quad \forall y \in V \end{cases} \tag{4.2.4}$$



Démonstration: L'existence de P_F est assuré par le théorème précédent. Il reste à établir la caractérisation.

Soit $x \in H$ et $q = P_F x$. Pour tout $y \in F$, $y + q \in F$ en appliquant la caractérisation du théorème précédent, on a $\mathfrak{Re}\langle x - q, y\rangle = \mathfrak{Re}\langle x - q, (y + q) - q\rangle \leq 0$

En remplaçant y par $-y$, iy et $-iy$ on obtient $\mathfrak{Re}\langle x - q, y\rangle = 0$ pour tout $y \in F$.

Inversement, si $\mathfrak{Re}\langle x - w, y\rangle = 0$, $\forall y \in F$, pour un certain $w \in V$, alors $\mathfrak{Re}\langle x - w, y - w\rangle = 0$ et $\|(1-\lambda)w + \lambda y - x\|^2 - \|w - x\|^2 \geq 0$ et en faisant tendre λ vers 1^-, on obtient $\|y - x\|^2 - \|w - x\|^2 \geq 0$, $\forall y \in F$, c-à-d que $w = P_F x$.

La même démonstration est valable si on suppose seulement que H est un espace préhilbertien par contre F doit être un sous-espace complet

4.2.7 Corollaire (Théorème de la projection : le cas préhilbertien)

Soit F un sous-espace vectoriel complet de l'espace préhilbertien H et soit $x \in H$.

Alors, il existe un **unique** $q \in F$, noté $P_F x$, tel que

$$\|x - q\| = d(x, F) := \inf_{a \in F} \|x - a\|.$$

De plus, $P_F x$ est caractérisé par les propriétés

$$\begin{cases} P_F x \in V \\ \langle x - P_F x, y\rangle = 0 \quad \forall y \in V \end{cases} \tag{4.2.5}$$

Cas particuliers : *Soit H un espace préhilbertien.*
Soit F un sous-espace vectoriel de dimension finie n et on suppose fixée une base $\mathcal{B} = \{\phi_1, \ldots, \phi_n\}$ *de F.*

Alors pour tout $x \in H$*, la caractérisation de la projection de x sur F, 4.2.4 devient*

$$\begin{cases} \textit{Trouver } u = \sum_{i=1}^n u_i \phi_i \in F \textit{ solution de} \\ Au = B \end{cases} \tag{4.2.6}$$

où $A = (a_{ij})_{1 \leq i,j \leq n}$ *avec* $a_{ij} = \langle \phi_i, \phi_j\rangle$*, est la matrice de Gram du produit scalaire dans la base* $\mathcal{B}$ *et* $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ *où* $b_j = \langle x, \phi_j\rangle$.

En effet, la caractérisation 4.2.4 $\langle x-u,y\rangle = 0$, $\forall y \in F$ est équivalente à : pour tout $j \in \{1,\dots,n\}$, $\langle u - \sum_{i=1}^n u_i\phi_i, \phi_j\rangle = 0$ i.e. $\sum_{i=1}^n u_i\langle \phi_i,\phi_j\rangle = \langle x,\phi_j\rangle$ qui n'est autre que le système d'équations $Au = B$.

4.2.8 Exercice Montrer que la matrice de Gram est définie positive.

4.2.9 EXEMPLE. Approximation par les moindres carrés.

Soit n nombres réels strictement positifs $p_i > 0$, $i \in \{1,\dots,n\}$ fixés, appelés "poids".

Etant donné un nuage de n points $(x_i,y_i)_{1\le i\le n}$ de $\mathbb{R}^2$, que l'on désire "ajuster" au mieux par une droite affine, par rapport à la norme et relativement aux poids.

On cherche alors à minimiser la fonction $f : \mathbb{R}^2 \to \mathbb{R}_+$ définie par $f(s,t) = \sum_{i=1}^n p_i\,(y_i - sx_i - t)^2$. On va utiliser pour cela le résultat sur les projections hilbertiennes.

On prend pour H l'espace des fonctions définies sur l'ensemble fini $\{x_1,\dots,x_n\}$ et à valeurs dans $\mathbb{R}$, i.e. $H = \{f : \{x_1,\dots,x_n\} \to \mathbb{R}\}$, et comme produit scalaire l'application définie par :

$\langle f,g\rangle := \sum_{i=1}^n p_i f(x_i)g(x_i)$ pour tous $f,g \in H$

4.2.10 Exercice Montrer que $(H, \langle,\rangle)$ est un espace de Hilbert

Comme on veut approcher ce nuage de points par des droites affines, on va choisir pour F l'espace vectoriel des droites affines, i.e. $F = \{f \in H | \exists c,d \in \mathbb{R}, f = c\phi_1 + d\phi_0\}$ où $\phi_0(x) = 1$ et $\phi_1(x) = x$ pour tout $x \in \{x_1,\dots,x_n\}$.

La fonction qui nous intéresse est la fonction $f \in H$ définie par $f(x_i) = y_i$ et sa projection sur F, $P_F f = a\phi_1 + b\phi_0$ est solution du système 4.2.5.

On est donc ramener à résoudre le système

$$\begin{pmatrix} \langle\phi_1,\phi_1\rangle & \langle\phi_0,\phi_1\rangle \\ \langle\phi_1,\phi_0\rangle & \langle\phi_0,\phi_0\rangle \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \langle f,\phi_1\rangle \\ \langle f,\phi_0\rangle \end{pmatrix}$$

ce qui revient à

$$\begin{pmatrix} \sum_{i=1}^n p_ix_i^2 & \sum_{i=1}^n p_ix_i \\ \sum_{i=1}^n p_ix_i & \sum_{i=1}^n p_i \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n p_iy_ix_i^2 \\ \sum_{i=1}^n p_iy_i \end{pmatrix}$$

d'où

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n p_ix_i^2 & \sum_{i=1}^n p_ix_i \\ \sum_{i=1}^n p_ix_i & \sum_{i=1}^n p_i \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n p_iy_ix_i^2 \\ \sum_{i=1}^n p_iy_i \end{pmatrix}.$$

4.2.11 EXEMPLE. Soit $a < b$ deux réels et $p :]a,b[\to \mathbb{R}$ une fonction continue et intégrable sur $[a,b]$. Soit $H = C^0([a,b])$ l'espace des fonctions continues sur $[a,b]$, muni du produit scalaire $\langle f,g\rangle := \int_a^b p(x)f(x)g(x)\,dx$ (c'est un espace préhilbertien).

Soit $F_n := \mathbb{R}_n[X]$ l'espace des polynômes de degré $\leq n$. F est de dimension $n+1$ et $\{1, x \ldots, x^n\}$ est une base de F .

Alors pour tout $f \in H$, la projection orthogonale $P_F f = \sum_{i=0}^{n} a_i x^i$, de f sur F est déterminée par le système

$$\begin{pmatrix} \langle 1,1 \rangle & \langle x,1 \rangle & \cdots & \langle x^n,1 \rangle \\ \vdots & \cdots & \cdots & \vdots \\ \langle 1,x^n \rangle & \langle x,x^n \rangle & \cdots & \langle x^n,x^n \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \langle f,1 \rangle \\ \vdots \\ \langle f,x^n \rangle \end{pmatrix}.$$

Par exemple si $p \equiv 1$ et $[a,b] = [0,1]$, la matrice de Gram du produit scalaire dans la base $\{1, x \ldots, x^n\}$ est la matrice $A = (a_{ij})_{0 \leq i,j \leq n}$ où $a_{ij} = \frac{1}{i+j+1}$ cette matrice est appelée matrice de Hilbert. En analyse numérique, cette matrice est connue pour être “mal conditionnée” Grosso modo cela veut dire, que si l'on perturbe le second membre, l'ordre de l'erreur relative sur la solution du sytème est très grand par rapport à l'erreur relative sur le second membre.

C'est pour cela qu'il est parfois utile de prendre des bases orthogonales ou orthonormées, une fois qu'on les a déterminé, les calculs dans ces bases sont plus faciles (voir 4.6.10).

4.3 Somme directe

4.3.1 DÉFINITION

Soient F et G deux sous-espace vectoriels d'un espace vectoriel E tels que $F \cap G = \{0\}$ et tout élément $x \in E$ peut s'écrire de façons unique sous la forme $x = y + z$ pour un certain $y \in F$ et $z \in G$. Dans ce cas on dit que E est la somme directe algébrique de F et G.

Dans ce cas on a la projection sur F parallèlement à G, notée P_F et la projection sur G parallèlement à F, notée P_G, et relation $P_F + P_G = I$.

Alors E est la somme directe algébrique de F et G si et seulement si l'application de $E \to F \times G$ définie par $x \mapsto (P_F(x), P_G(x))$ est un isomorphisme (algébrique)

Si E est un espace vectoriel normé, on dit que E est la somme directe topologique de F et G si l'application $E \to F \times G$ définie par $x \mapsto (P_F(x), P_G(x))$ est un isomorphisme (topologique).

4.3.2 THÉORÈME

Soit F un sous-espace vectoriel **fermé** d'un espace de Hilbert H. On a les propriétés suivantes :

1. Tout $x \in H$ a une décomposition unique :
$$x = P_F x + (I - P_F)x$$
avec $P_F x \in V$ et $P_{F^\perp} x = (I - P_F)x \in F^\perp$ i.e. $H = F \oplus F^\perp$.
2. $\|x\|^2 = \|P_F x\|^2 + \|P_{F^\perp} x\|^2$
3. L'application $P_F : H \longrightarrow F$ est linéaire continue et pour tout $x \in H$, $\|P_F x\| \leq \|x\|$.
$\|P_F\| = 1$ si $F \neq \{0\}$.
4. $P_F \circ P_F = P_F$ (i.e. P_F est idempotent)
5. Pour tout $x, y \in H$ on a $\langle P_F x, y\rangle = \langle x, P_F y\rangle$ (i.e. P_F est un opérateur auto-adjoint)
L'application $P_F : H \to F$ est appelée la projection orthogonale de H sur F.

Démonstration: 1. D'après le théorème de la projection hilbertienne, $\forall x \in H$ $\langle x - P_F x, y\rangle = 0$ pour tout $y \in F$, donc $x - P_F x$ est orthogonal à F.
Ainsi $x = P_F x + (x - P_F x) \in F + F^\perp$ d'où $H = F + F^\perp$.
La somme est directe car si $x \in F \cap F^\perp$ alors $\|x\|^2 = \langle x, x\rangle = 0$, donc $F \cap F^\perp = \{0\}$.
2. Comme $x \in H$ s'écrit $x = P_F x + (x - P_F x)$ avec $P_F x \in F$ et $x - P_F x \in F^\perp$, d'après Pythagore on aura $\|x\|^2 = \|P_F x\|^2 + \|x - P_F x\|^2$
3. On montre que l'application P_F est linéaire en utilisant la caractérisation 4.2.4 et l'unicité de la projection. Ainsi $P_F x + P_F y$ est le projeté de $x + y$ et par unicité il est égal à $P_F(x + y)$, de même $\lambda P_F x = P_F(\lambda x)$.
D'après $\|x\|^2 = \|P_F x\|^2 + \|x - P_F x\|^2$, on a $\|x\| \geq \|P_F x\|$, d'où P_F est continue et de norme $\|P_F\| \leq 1$.
Si $F \neq \{0\}$, on a pour $x \in F - \{0\}$, $P_F x = x$ d'où $\|P_F x\| = \|x\|$, ce qui entraîne $\|P_F\| \geq 1$. En conclusion, les deux inégalités nous donne $\|P_F\| = 1$.
4. Clairement, pour tout $x \in H$, on a $P_F(P_F x) = P_F x$.
5. Soient $x, y \in H$, d'après la caractérisation 4.2.4 et $P_F y \in F$ on a $\langle x - P_F x, P_F y\rangle = 0$ i.e. $\langle x, P_F y\rangle = \langle P_F x, P_F y\rangle$.
De même $\langle P_F x, P_F y\rangle = \overline{\langle P_F y, P_F x\rangle} = \overline{\langle y, P_F x\rangle} = \langle P_F x, y\rangle$.
Ainsi $\langle x, P_F y\rangle = \langle P_F x, y\rangle$. ∎

4.3.4 Remarque

Les mêmes propriétés sont valables pour $(I - P_F) = P_{F^\perp}$, où I est l'application identité de H.

Maintenant qu'on a étudier l'orthogonal $F^{\perp}$ de F, on peut passer à celle de $F^{\perp\perp}$ l'orthogonal de $F^{\perp}$.

4.3.5 Corollaire (Décomposition orthogonale)

Soit H un espace de Hilbert et F un sous-espace fermé. Alors, tout vecteur $x \in H$ peut être représenté de manière unique comme

$$x = y + z, \quad y \in F, \quad z \in F^{\perp}.$$

Cette décomposition orthogonale est notée

$$H = F \oplus F^{\perp}.$$

c'est une somme directe topologique

4.3.6 Définition (Projection orthogonale)

Dans les hypothèses du corollaire 4.3.5, l'application

$$P_F : H \to H, \quad P_F x = y$$

est appelée la *projection orthogonale de* H sur F.

4.3.7 Corollaire

Soit F un sous-espace vectoriel fermé d'un espace de Hilbert H. Alors $F^{\perp\perp} = F$.

Démonstration: D'après (i) du Lemme 4.1.4 on a $F \subset F^{\perp\perp}$. Maintenant on suppose que $x \in F^{\perp\perp}$, alors , il existe un unique $y \in F$ et $z \in F^{\perp}$ tel que $x = y + z$. Comme $y \in Y$ et $x \in Y^{\perp\perp}$, $0 = \langle y, z \rangle = \langle x, z \rangle$. Alors

$$0 = \langle x, z \rangle = \langle y + z, z \rangle = \langle y, z \rangle + \langle z, z \rangle = \|z\|^2.$$

d'où $z = 0$ et donc $x = y \in F$. Ainsi $F^{\perp\perp} \subset F$, ce qui termine la preuve. ■

Si F n'est pas fermé, le résultat précédent n'est plus vrai car $F^{\perp\perp}$ est toujours fermé. Neanmoins on a :

4.3.9 Corollaire

Soit F un sous-espace vectoriel d'un espace de Hilbert H. Alors

1. $F^{\perp\perp} = \overline{F}$
2. F est dense dans H si et seulement si $F^{\perp} = \{0\}$.

Démonstration: Comme $F \subset \overline{F}$, il s'en suit de (g) du Lemme 4.1.4 que $\overline{F}^{\perp} \subset F^{\perp}$ et donc $F^{\perp\perp} \subset \overline{F}^{\perp\perp}$. D'après le corollaire précédent, $\overline{F}^{\perp\perp} = \overline{F}$, et donc $F^{\perp\perp} \subset \overline{F}$. D'autre part, d'après (i) du Lemme 4.1.4, $F \subset F^{\perp\perp}$, mais $F^{\perp\perp}$ est un fermé $\overline{F} \subset F^{\perp\perp}$. D'où $F^{\perp\perp} = \overline{F}$. ■

4.3.11 REMARQUE
D'après le corollaire précédent, $F^{\perp\perp}$ est le plus petit fermé contenant F.

4.3.12 Exercice Soit H un espace de Hilbert, M un sous-espace fermé, et N un sous-espace de dimension finie de H. Montrer que $M + N$ est un sous-espace fermé de H.

4.3.13 DÉFINITION
Soit H un espace de Hilbert. Un endomorphisme $P : H \to H$ est un opérateur **auto-adjoint** (ou **hermitien**) si $\langle Px, y\rangle = \langle x, Py\rangle$ pour tout $x, y \in H$.

4.3.14 Exercice Soit P un opérateur auto-adjoint d'un espace de Hilbert H tel que $P^2 = P$. Montrer que P est la projection orthogonale sur un sous-espace fermé de H.

4.4 Le théorème de représentation de Riesz

Les espaces de Hilbert possèdes des propriétés de dualités remarquables.

4.4.1 Définition

Soit $(E, \|.\|)$ un espace vectoriel normé. On note par E^* son dual topologique, i.e. l'espace vectoriel des formes linéaires continues $f : E \to \mathbb{K}$.

Cet espace est muni de la norme

$$\|f\| = \sup\{|f(x)|; \|x\| \leq 1\}.$$

Soit H un espace de Hilbert, pour $y \in H$ on associe la forme linéaire sur H, $\phi_y : H \to \mathbb{C}$ définie par $\phi_y(x) = \langle x, y\rangle$. D'après l'inégalité de Cauchy-Schwarz on a $|\phi_y(x)| = |\langle x, y\rangle| \leq \|y\|\|x\|$ d'où la forme linéaire ϕ_y est continue, donc un élément de H^*.

On définit alors une application $\Phi : H \to H^*$, qui à $y \in H$ associe $\Phi(y) = \phi_y$.

4.4.2 Théorème (Le théorème de Représentation de Riesz)

Soit H un espace de Hilbert. Alors

1) L'application Φ est une isométrie i.e. pour tout $y \in H$, $\|\Phi(y)\| = \|y\|$.
2) Soit f est forme linéaire continue i.e. $f \in H^*$, alors il existe un unique $y \in H$ tel que $f = \phi_y$ et $\|f\| = \|y\|$.

Démonstration: 1) Il suffit de supposer que $y \neq 0$, alors de l'inégalité de Cauchy-Schwarz on a $\|\phi_y\| \leq \|y\|$, d'autre part $\phi_y(y) = \|y\|^2$, entraîne $\|\phi_y\| \geq \|y\|$, d'où $\|\phi_y\| = \|y\|$.

2) Soit $f \in H^*$, si $f = 0$, alors $f = \phi_0$. Maintenant, si $f \neq 0$, on pose $F = \ker f$, comme f est continue $\ker f$ est un sous-espace vectoriel fermé de H, d'après le théorème 4.3.2, $F^\perp \neq \{0\}$. Il en résulte l'existence de $y_0 \in F^\perp - \{0\}$, alors $f(y_0) \neq 0$. On peut supposer quitte à diviser par le scalaire $f(y_0)$, que $f(y_0) = 1$.

Soit $x \in H$, on vérifie directement que $x - f(x)y_0 \in F$ d'où $\langle x - f(x)y_0, y_0\rangle = 0$ et ainsi $\langle x, y_0\rangle = f(x)\|y_0\|^2$, finalement $f(x) = \langle x, \frac{y_0}{\|y_0\|^2}\rangle = \langle x, y\rangle$ avec $y = \frac{y_0}{\|y_0\|^2}$.

4.4.4 Corollaire

Soit H un espace de Hilbert. L'application $\Phi : H \to H^*$, $y \mapsto \Phi(y)$ définie par : pour tout $x \in H$, $\Phi(y)(x) = \langle x, y\rangle$ est une isométrie bijective et antilinéaire de H sur H^*.

Si H est réel, c'est un isomorphisme isométrique de H sur H^*.

4.4.1 Adjoint d'un opérateur

Soient H et K deux espaces de Hilbert et $A \in \mathcal{L}(H,K)$ i.e une application linéaire continue.

Il existe une unique application linéaire continue $A^* \in \mathcal{L}(K,H)$ telle que pour tout $x \in H$ et $y \in K$ on a :

$$\langle Ax, y\rangle_K = \langle x, A^*y\rangle_H$$

De plus $\|A^*\| = \|A\|$ et $A^{**} = A$.

L'application A^* est appelée l'application **adjointe** *de A (ou l'opérateur adjoint de l'opérateur A).*

Démonstration: Soit $y \in K$ fixé, l'application de $H \to \mathbb{C}$ définie par $x \mapsto \langle Ax, y\rangle_K$ est une forme linéaire continue, d'après le théorème de représentation de Riesz, il existe un unique élément de H, noté A^*y tel que $\langle Ax, y\rangle_K = \langle x, A^*y\rangle_H$. Grâce à l'unicité, on définit ainsi une application $A^* : K \to H$. En utilisant encore l'unicité, on montre que A^* est linéaire.

On va maintenant établir sa continuité. D'après l'inégalité de Cauchy-Schwarz on a

$$|\langle x, A^*y\rangle_H| = |\langle Ax, y\rangle_K| \leq \|Ax\|\|y\| \leq \|A\|\|x\|\|y\|.$$

D'où la norme de la forme linéaire $x \mapsto \langle x, A^*y\rangle_H$, à savoir $\|A^*y\| \leq \|A\|\|y\|$. On en déduit la continuité de A^* et l'inégalité $\|A^*\| \leq \|A\|$.

Maintenant par unicité de l'opérateur adjoint, on aura $A^{**} = (A^*)^* = A$ et d'après ce qui précède $\|A\| = \|A^{**}\| \leq \|A^*\|$, d'où $\|A^*\| = \|A\|$. ∎

4.5 Systèmes orthonormés

4.5.1 DÉFINITION

Un système $\{x_i\}_{i\in I}$ de vecteurs d'un espace vectoriel H sur $\mathbb{K}$ est un **système libre** si pour toute partie finie $J \subset I$ la combinaison linéaire finie $\sum_{j\in J} \lambda_j x_j = 0$ si et seulement si $\lambda_j = 0$ pour tout $j \in J$.

En d'autres termes, le vecteur nul 0 a une écriture unique dans

$$\mathrm{vect}\{\{x_i\}_{i\in I}\} := \left\{ \sum_{j\in J} \lambda_j x_j \mid \text{pour toute partie finie } J \subset I \text{ et } \lambda_j \in \mathbb{K} \right\}.$$

4.5.2 DÉFINITION
Un système $\{x_i\}_{i\in I}$ de vecteurs d'un espace préhilbertien H sur $\mathbb{K}$ est un **système orthogonal** si, $\langle x_i, x_j\rangle = 0$ pour $i, j \in I$ tels que $i \neq j$; en d'autres termes, le système est constitué de vecteurs de H orthogonaux deux à deux.

4.5.3 DÉFINITION
Un système orthogonal $\{x_i\}_{i\in I}$ de vecteurs d'un préhilbertien H est un **système orthonormé** si $\|x_i\| = 1$ pour tout $i \in I$.

4.5.4 REMARQUE
Dans le cas particulier où $I = \mathbb{N}$ (ou dénombrable), un système orthonormé (respectivement orthogonal) $\{x_n\}_{n\in\mathbb{N}}$ est appelé parfois suite orthonormée (resp. orthogonale).

4.5.5 EXEMPLE. *Dans $\mathbb{R}^n$ (ou $\mathbb{C}^n$) muni du produit scalaire standard le système $\{e_1, \cdots, e_n\}$, donné par*

$$e_1 = (1, 0\cdots, 0), e_2 = (0, 1, 0\cdots, 0), \ldots, e_k = (\underbrace{0, \cdots, 0}_{k-1}, 1, 0, \cdots, 0), \ldots, e_n = (0, \ldots, 1)$$

est une suite orthonormée.

4.5.6 EXEMPLE. *L'espace de Hilbert $\ell^2(\mathbb{N})$, muni du produit scalaire $\langle x, y\rangle = \sum_{k=0}^{\infty} \xi_k \overline{\eta_k}$, le système $\{e_n\}_{n\in\mathbb{N}}$, donné par*

$$e_n = (\underbrace{0, \cdots, 0}_{n-1}, 1, 0, \cdots) \quad \textit{pour tout } n \in \mathbb{N},$$

$$e_n = (\delta_{k,n})_{k\in\mathbb{N}^*} \text{ où } \delta_{ij} = \begin{cases} 1 & si\ i = j \\ 0 & si\ i \neq j \end{cases} \text{(le symbole de Kronecker),}$$

est une suite orthonormée.

4.5.7 EXEMPLE. *L'espace de Hilbert $L^2[-\pi, \pi]$ muni du produit scalaire $\langle x, y\rangle = \int_{-\pi}^{\pi} x(t)\overline{y(t)}\, dt$, le système $\{e_n\}$, pour $n \in \mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$, où la fonction $e_n : [-\pi, \pi] \to \mathbb{C}$ est définie par*

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int} \quad \textit{pour tout } t \in [-\pi, \pi].$$

Alors pour tout $n, m \in \mathbb{Z}$, on a

$$\langle e_n, e_m\rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)t}\, dt = \begin{cases} 1 & si\ n = m \\ 0 & si\ n \neq m. \end{cases}$$

Ainsi $\{e_n\}_{n\in\mathbb{Z}}$ *est une suite orthonormée, appelée suite trigonométrique.*

On considère $\{f_0, f_1, g_1, f_2, g_2, \cdots\}$, *où la fonction* $f_0 : [-\pi, \pi] \to \mathbb{R}$ *est définie par*

$$f_0(t) = \frac{1}{\sqrt{2\pi}} \quad \text{pour tout } t \in [-\pi, \pi,],$$

et pour tout $n \in \mathbb{N}^*$, *les fonctions* $f_n, g_n : [-\pi, \pi] \to \mathbb{R}$ *sont définie par*

$$f_n(t) = \frac{1}{\sqrt{\pi}} \cos nt \quad \text{et} \quad g_n(t) = \frac{1}{\sqrt{\pi}} \sin nt \quad \text{pour tout } t \in [-\pi, \pi].$$

on montre aussi que $\{f_0, f_1, g_1, f_2, g_2, \cdots\}$ *est une suite orthonormée appelée suite trigonométrique réelle.*

Ces fonctions seront considérées plus en détails au chapitre sur les séries de Fourier.

4.5.8 PROPOSITION (IDENTITÉ DE PYTHAGORE ET INÉGALITÉ DE BESSEL)
Soit H un espace préhilbertien .

(i) Soit $\{x_1, \ldots, x_N\}$ un système de vecteurs orthogonaux alors on **l'identité de Pythagore** suivante :

$$\|x_1 + \ldots + x_N\|^2 = \|x_1\|^2 + \ldots + \|x_N\|^2$$

(ii) Soit $\{e_n\}_{n\in\mathbb{N}}$ une suite orthonormée de H, on a alors :

(a) $$\begin{cases} \forall x \in H \\ \forall N \in \mathbb{N} \end{cases} \quad \|x\|^2 = \sum_{k=0}^{N} |\langle x, e_k\rangle|^2 + \|x - \sum_{k=0}^{N} \langle x, e_k\rangle e_k\|^2$$

(b) **l'inégalité de Bessel**

$$\sum_{n\geq 0} |\langle x, e_n\rangle|^2 \leq \|x\|^2, \qquad \forall x \in H$$

Démonstration: (i) Par récurrence à partir de l'identité de Pythagore pour 2 vecteurs.

(ii) (a) On écrit $x_{\|} = \sum_{k=0}^{N} \langle x, e_k\rangle e_k$ et $x_{\perp} = x - \sum_{k=0}^{N} \langle x, e_k\rangle e_k$.

Alors, $x_{\|} \perp x_{\perp}$ et comme $x = x_{\|} + x_{\perp}$ on aura d'après l'identité de Pythagore $\|x\|^2 = \|x_{\|}\|^2 + \|x_{\perp}\|^2$.

(b) D'après (a), pour tout N entier ≥ 1 et en minorant simplement $\|x_{\perp}\|^2$ par 0 on aura $\sum_{n=0}^{N} |\langle x, e_n\rangle|^2 \leq \|x\|^2$, et par passage à la limite quet $N \to +\infty$ on obtient l'inégalité de Bessel. ∎

Une conséquence du lemme précédent est, si on se donne dans un espace préhilbertien H et une suite orthonormée $\{e_n\}_{n\in\mathbb{N}}$, alors pour tout $x \in H$, la série de terme général $(\langle x, e_n\rangle e_n)$ est absolument convergente, et si H est en plus complet, la série est convergente.

4.5.10 PROPOSITION
Soit H un espace de Hilbert et $\{e_n\}_{n\in\mathbb{N}}$ une suite orthonormée.
Soit $(\alpha_n)_{n\in\mathbb{N}}$ une suite de scalaire.
La série $\sum\limits_{n\in\mathbb{N}} \alpha_n e_n$ converge dans H si et seulement si $(\alpha_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$.

Démonstration: Alorsque H est complet la série converge si et seulement si elle vérifie le critère de Cauchy.

Pour tout $m, p \in \mathbb{N}$, $p > m$, l'identité de Pythagore nous donne :

$$\|\sum_{n=0}^{p} \alpha_n e_n - \sum_{n=0}^{m} \alpha_n e_n\|^2 = \|\sum_{n=m+1}^{p} \alpha_n e_n\|^2 = \sum_{n=m+1}^{p} |\alpha_n|^2$$

ainsi la série $\sum_{n\in\mathbb{N}} \alpha_n e_n$ est convergente si et seulement si la série réelles à termes positifs $\sum_{n\in\mathbb{N}} |\alpha_n|^2$ vérifie le critère de Cauchy, donc converge car $\mathbb{R}$ est complet. ∎

4.5.12 REMARQUE
Dans le cas particulier où la suite $(\alpha_n)_{n\in\mathbb{N}}$ est la suite $(\langle x, e_n\rangle)_{n\in\mathbb{N}}$, l'inégalité de Bessel nous assure de l'appartenance de cette suite à $\ell^2(\mathbb{N})$.

Ainsi, on définit une application $\Phi : H \to \ell^2(\mathbb{N})$, par $\Phi(x) = (\langle x, e_n\rangle)_{n\in\mathbb{N}}$. Cette application est linéaire, car le produit scalaire est linéaire par rapport à x et l'inégalité de Bessel assure la continuité de cette application :

$$\|\Phi(x)\|_2 = \sum_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2 \leq \|x\|^2.$$

Finalement, la proposition précédente entraîne sa surjectivité, en effet, pour toute suite $(\alpha_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$, le vecteur $\sum_{n\in\mathbb{N}} \alpha_n e_n \in H$ et vérifie $\Phi(\sum_{n\in\mathbb{N}} \alpha_n e_n) = (\alpha_n)_{n\in\mathbb{N}}$.

4.5.1 Sommes et bases hilbertiennes

4.5.13 DÉFINITION
Soit H un espace de Hilbert et $\{E_n\}_{n\in\mathbb{N}}$ une suite de sous-espaces vectoriels fermés de H.

On dit que H est la somme hilbertienne des E_n, si :

1) $E_m \perp E_n$ pour tous entiers m, n tels que $m \neq n$.

2) l'espace engendré par les E_n est dense dans H.

On note dans ce cas, $H = \bigoplus_n E_n$.

4.5.14 THÉORÈME

Soit H un espace de Hilbert et $\{E_n\}_{n\in\mathbb{N}}$ une suite de sous-espaces vectoriels fermés telle que $H = \bigoplus_n E_n$. Soit P_{E_n} la projection orthogonale sur E_n.

Alors pour tout $x \in H$ on a :

1) $x = \sum\limits_n P_{E_n}(x)$

2) $\|x\|^2 = \sum\limits_n \|P_{E_n}(x)\|^2$

Démonstration: Pour $k \in \mathbb{N}$, on pose $S_k = \sum_{n=0}^k P_{E_n}$. S_k est alors un opérateur linéaire continu sur H. Comme le système $\{P_{E_n}(x)\}_{1\le n\le k}$ est orthogonal, d'après la proposition 4.5.8, on a

$$\|S_k(x)\|^2 = \sum_{n=0}^{k} \|P_{E_n}(x)\|^2$$

or pour tout $n \in \mathbb{N}$, on a $\langle x, P_{E_n}(x)\rangle = \|P_{E_n}(x)\|^2$, il s'ensuit que pour tout $k \in \mathbb{N}$ on a $\langle x, S_k(x)\rangle = \|S_k(x)\|^2$, d'où par Cauchy-Scwarz, $\|S_k(x)\| \le \|x\|$ i.e. $\|S_k\| \le 1$.

Soit $E = \mathrm{Vect}\{E_n \,|\, n \in \mathbb{N}\}$. Par hypothèse, pour tout $\varepsilon > 0$, il existe $y \in E$ tel que $\|x - y\| \le \varepsilon$.

Mais $y \in E$, entraîne l'existence de $N \in \mathbb{N}$ et $y_n \in E_n$ tels que $y = \sum_{n=0}^N y_n$. Comme les E_n sont deux à deux orthogonaux, on a pour $k \ge m$, $S_k(y) = y$.

On a alors pour $k \ge m$:

$$\|x - S_k(x)\| \le \|x-y\| + \|y - S_k(y)\| + \|S_k(y) - S_k(x)\| \le\le \|x-y\| + 0 + \|x-y\| \le 2\varepsilon$$

ainsi $\lim_{k\to+\infty} S_k(x) = x$ ce qui démontre 1).

Pour 2), en utilise la continuité du produit scalaire :

$$\|x\|^2 = \lim_{k\to+\infty} \langle x, S_k(x)\rangle = \lim_{k\to+\infty} \sum_{n=0}^{k} \|P_{E_n}(x)\|^2.$$

4.5.16 DÉFINITION

Soit H un espace de Hilbert. Une suite orthonormée $\{e_n\}_{n\in\mathbb{N}}$ est une **base hilbertienne** si elle est totale i.e. $\begin{cases} \langle x, e_n\rangle = 0 \\ \forall\, n \in \mathbb{N} \end{cases} \quad \Rightarrow \; x = 0.$

4.5.17 REMARQUE
Une suite orthonormée est une base hilbertienne si et seulement si l'application Φ de la remarque 4.5.12 est injective.

4.5.18 THÉORÈME
Soit H un espace de Hilbert et $\{e_n\}_{n\in\mathbb{N}}$ est une suite orthonormée.
Les conditions suivantes sont équivalentes :

(i) $\{e_n\}_{n\in\mathbb{N}}$ est une base hilbertienne
(ii) $\forall x \in H, x = \sum_{n\in\mathbb{N}} \langle x, e_n\rangle e_n$.
(iii) $\forall x \in H, \|x\|^2 = \sum_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2$. (**l'identité de Parseval**)

Démonstration: 1. $(i) \Rightarrow (ii)$ Pour tout $j \in \mathbb{N}$, $\langle x - \sum_{n\in\mathbb{N}} \langle x, e_n\rangle e_n, e_j\rangle = \langle x, e_j\rangle - \langle x, e_j\rangle = 0$, alors l'hypothèse (i) entraîne $x - \sum_{n\in\mathbb{N}} \langle x, e_n\rangle e_n = 0$ i.e. $x = \sum_{n\in\mathbb{N}} \langle x, e_n\rangle e_n$.

2. $(ii) \Rightarrow (iii)$ D'après le (ii)-(a) du lemme 4.5.8, on a pout tout $N \in \mathbb{N}$, $\|x\|^2 = \sum_{0\leq k\leq N} |\langle x, e_k\rangle|^2 + \|x - \sum_{0\leq k\leq N} \langle x, e_k\rangle e_k\|^2$. Ainsi, lorsque $N \to +\infty$ on obtient l'identité de Parseval.

3. $(iii) \Rightarrow (i)$
Comme par hypothèse on a $\begin{cases} \langle x, e_n\rangle = 0 \\ \forall\, n \in \mathbb{N} \end{cases}$ de l'identité de Parseval, on aura que $\|x\|^2 = 0$ i.e. $x = 0$.

4.5.20 REMARQUE
Si $\{e_n\}_{n\in\mathbb{N}}$ est une base hilbertienne de H, l'identité de Parseval entraîne que $\|\Phi(x)\|_2 = \sum_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2 = \|x\|^2$ d'où Φ est une isométrie, donc injective et par la remarque précédente, on obtient ainsi une application unitaire de H dans $\ell^2(\mathbb{N})$.

4.5.21 PROPOSITION
Soit $\{x_i\}_{i\in I}$ un système orthonormé d'un préhilbertien H. Alors
(a) Pour tout sous-ensemble fini $J \subset I$ on a

$$\|\sum_{i\in J} x_i\|^2 = \sum_{i\in J} \|x_i\|^2;$$

(b) Tout système orthonormé $\{x_i\}_{i\in I}$ est un sytème linéairement indépendant de H, i.e. pour tout $J \subset I$ sous-ensemble fini, $\{x_i\}_{i\in J}$ sont linéairement indépendents.

Une conséquence de cette proposition est, comme une suite orthogonale (ou orthonormée) est un système linéairement indépendant, alors si H contient un système orthogonal infini, H est de dimension infinie. La réciproque est donnée par le procédé d'orthonormalisation de Gram-Schmidt 4.6.10.

4.6 Base Hilbertienne

4.6.1 Définition

Soit H un espace préhilbertien . Une **base hilbertienne** de H est un système orthonormé $\{e_i\}_{i\in I}$ total c-à-d que $H = \overline{vect\{e_i\}_{i\in I}}$.

4.6.2 Remarque

– *Si H est un espace de hilbert alors un sytème orthonormé* $\{e_i\}_{i\in I}$ *est une base Hilbertienne si et seulement si* $\begin{cases} \langle x, e_i\rangle = 0 \\ \text{pour tout } i \in I \end{cases} \Rightarrow x = 0.$

– *En effet, si* $\langle x, e_i\rangle = 0$, *pour tout* $i \in I$, *alors* $x \in vect\{e_i\}_{i\in I}^{\perp} = \overline{vect\{e_i\}_{i\in I}}^{\perp} = H^{\perp} = \{0\}$.

4.6.3 Proposition

Soit H un espace préhilbertien et soit $\{f_i\}_{i\in I}$ un système orthogonal dans H. Alors $S = \sum_{\alpha\in I} f_i$ existe si et seulement si $\sum_{i\in I} \|f_i\|^2 < \infty$.

On suppose que $\sum_{i\in I} \|f_i\|^2 < \infty$, alors

1. $\|S\|^2 = \sum_{i\in I} \|f_i\|^2$
2. Pour tout $x \in H$, on a $\langle x, S\rangle = \sum_{i\in I} \langle x, f_i\rangle$.

Voir l'Annexe 4.7.3 pour la notion de famille sommable.

4.6.4 Corollaire

Soit H un espace préhilbertien et soit $\{e_i\}_{i\in I}$ un système orthonormé dans H. On pose $F = \overline{vect\{e_i\}_{i\in I}}$.

Alors pour tout $x, y \in H$ on a :

1. $P_F x = \sum_{i\in I} \langle x, e_i\rangle e_i$
2. $\|P_F x\|^2 = \sum_{i\in I} |\langle x, e_i\rangle|^2$

3. $\langle P_F x, y\rangle = \sum_{i\in I}\langle x, e_i\rangle\langle e_i, y\rangle$

Démonstration: D'après l'inégalité de Bessel, pour tout $x \in H$, $\sum_{i\in I}|\langle x, e_i\rangle|^2 \leq \|x\|^2$ et daprès la proposition précédente $Px := \sum_{i\in I}\langle x, e_i\rangle e_i$ existe dans H, et pour tout $y \in H$ on a

$$\langle Px, y\rangle = \langle \sum_{i\in I}\langle x, e_i\rangle e_i, y\rangle = \sum_{i\in I}\langle x, e_i\rangle\langle e_i, y\rangle.$$

En particulier si $y = e_j$, alors $\langle Px, e_j\rangle = \langle x, y\rangle$, i.e. $\langle x - Px, e_j\rangle = 0$, comme $j \in I$ est quelconque, on a pour tout $y \in vect\{e_i\}_{i\in I}$, $\langle x - Px, y\rangle = 0$ et par continuité du produit scalaire, $\langle x - Px, y\rangle = 0$ pour tout $y \in F$, d'après le théorème de la projection on a $P = P_F$.

Maintenant, $P_F^2 = P_F$ et P_F auto-adjoint entraînent : $\|P_F x\|^2 = \langle P_F x, P_F x\rangle = \langle P_F^2 x, x\rangle = \langle P_F x, x\rangle = \sum_{i\in I}\langle x, e_i\rangle\langle e_i, x\rangle = \sum_{i\in I}\langle x, e_i\rangle\overline{\langle x, e_i\rangle} = \sum_{i\in I}|\langle x, e_i\rangle|^2$. ■

Le théorème suivant est une généralisation de 4.5.18

4.6.6 Théorème

Soit H un espace de Hilbert et $\{e_i\}_{i\in I}$ un système orthonormé.

Les conditions suivantes sont équivalentes :

1. $\{e_i\}_{i\in I}$ est une base hilbertienne.
2. Pour tout $x \in H$, $x = \sum_{i\in I}\langle x, e_i\rangle e_i$
3. Pour tout $x, y \in H$,

$$\langle x, y\rangle = \sum_{i\in I}\langle x, e_i\rangle\langle e_i, y\rangle = \sum_{i\in I}\langle x, e_i\rangle\overline{\langle y, e_i\rangle}$$

4. Pour tout $x \in H$, $\|x\|^2 = \sum_{i\in I}|\langle x, e_i\rangle|^2$ (l'identité de Parseval)

Démonstration:
- (1)⇒ (2), D'après le corollaire 4.6.4, appliqué à $F = H$ et donc P_H est égale à l'identité, on aura pour tout $x \in H$: $x = P_H x = \sum_{\alpha\in I}\langle x, e_i\rangle e_i$
- (2)⇒(3), conséquence de la proposition précédente
- (3)⇒(4), Il suffit de prendre $y = x$
- (4)⇒(1), Si $x \in \overline{vect\{e_i\}_{i\in I}}^{\perp}$, d'après 4), $\|x\| = 0$, i.e. $x = 0$, ceci entraîne que $\overline{vect\{e_i\}_{i\in I}}^{\perp} = \{0\}$ et donc $\overline{vect\{e_i\}_{i\in I}} = \{0\}^{\perp} = H$. ■

On va maintenant s'intéresser au problème de l'existence d'une base hilbertienne. On va commencer par considérer le cas dénombrable.

4.6.1 Le cas séparable

4.6.8 DÉFINITION
Soit (E, d) un espace métrique.

On dit que E est **séparable** s'il existe une partie dénombrable $\mathbb{D}$ sous-ensemble de E qui soit dense dans E i.e $\overline{\mathbb{D}} = E$.

4.6.9 EXEMPLE.
1. Les espaces $\mathbb{R}$ et $\mathbb{C}$ sont séparables.
   $\mathbb{D} = \mathbb{Q}$ pour le premier et $\mathbb{D} = \mathbb{Q} + i\mathbb{Q}$ pour le second.
2. Pour tout $1 \leq p < +\infty$, l'espace de Banach
   $\ell^p(\mathbb{N}) = \{(x_n)_{n\in\mathbb{N}}; \sum_{n\in\mathbb{N}} |x_n|^p < +\infty\}$ est séparable (exercice)
3. Par contre, l'espace de Banach $\ell^\infty(\mathbb{N}) = \{(x_n)_{n\in\mathbb{N}}; \|(x_n)\|_\infty = \sup_n |x_n| < +\infty\}$ n'est pas séparable.

Etant donnée une suite libre $\{f_n\}_{n\in\mathbb{N}}$ d'un espace préhilbertien H on voudrait construire une suite orthonormée $\{e_n\}_{n\in\mathbb{N}}$ telle que pour tout $p \in \mathbb{N}$ on ait
$\mathrm{Vect}\{e_0, \ldots, e_p\} = \mathrm{Vect}\{f_0, \ldots, f_p\}$, *pour cela on utilise :*

4.6.10 THÉORÈME (LE PROCÉDÉ D'ORTHONORMALISATION DE GRAM-SCHMIDT)
Soit H un espace préhilbertien de dimension infinie.

Soit $\{f_n\}_{n\in\mathbb{N}}$ une suite libre de vecteurs de H.

On pose pour $p \in \mathbb{N}$, $V_p := \mathrm{Vect}\{f_0, \ldots, f_p\}$.

Alors, la suite $\{e_n\}_{n\in\mathbb{N}}$ définie par une récurrence :

$e_0 = \frac{f_0}{\|f_0\|}$ et pour $p \geq 1$, $e_p = \frac{f_p - P_{V_{p-1}}(f_p)}{\|f_p - P_{V_{p-1}}(f_p)\|}$

est une suite orthonormée telle que $\mathrm{vect}\{e_0, \ldots, e_p\} = V_p$.

En fait, on a pour $p \geq 0$, $e_{p+1} = \dfrac{f_{p+1} - \sum_{i=0}^{p} \langle f_p, e_i\rangle f_i}{\|f_{p+1} - \sum_{i=0}^{p} \langle e_p, e_i\rangle f_i\|}$.

Démonstration: On raisonne par récurrence : Comme la suite est libre $f_0 \neq 0$ par suite $e_0 = \frac{f_0}{\|f_0\|}$ vérifie bien $\|e_0\| = 1$.

Supposons que $e_0, \ldots, e_p$ sont déjà obtenus et que $V_p = \mathrm{Vect}\{e_0, \ldots, e_p\}$.

Soit $P_{V_p}(f_{p+1})$ la projection orthogonale de f_{p+1} sur V_p, d'après la caractérisation de la projection on a

$\langle f_{p+1} - P_{V_p}(f_{p+1}), y\rangle = 0$ pour tout $y \in V_p$, en particulier, $f_{p+1} - P_{V_p}(f_{p+1})$ est orthogonal aux vecteurs $e_0, \ldots, e_p$ et on obtient un système orthonormé $e_0, \ldots, e_p, e_{p+1}$ en posant $e_{p+1} = \frac{f_{p+1} - P_{V_p}(f_{p+1})}{\|f_{p+1} - P_{V_p}(f_{p+1})\|}$.

Il reste à determiner $P_{V_p}(f_{p+1})$ dans la base orthonormée $\{e_1, \dots, e_p\}$.

Si on écrit $P_{V_p}(f_{p+1}) = \sum_{0=1}^{p} \lambda_i e_i$, alors $\langle P_{V_p}(f_{p+1}), e_i \rangle = \lambda_i$ et de la caractérisation précédente on a $\langle P_{V_p}(f_{p+1}), e_i \rangle = \langle f_{p+1}, e_i \rangle$, finalement

$P_{V_p}(f_{p+1}) = \sum_{i=0}^{p} \lambda_i e_i = \sum_{i=0}^{p} \langle f_{p+1}, e_i \rangle e_i$.

4.6.12 EXEMPLE. Soit $H = C^0([-1,1])$ muni du produit scalaire $\langle f, g \rangle = \int_0^1 p(x) f(x) g(x)\, dx$ avec $p(x) = \frac{1}{\sqrt{1-x^2}}$.

Alors la suite orthonormée obtenue à partir de la suite des monômes $\{1, x, x^2, \dots, x^n, \dots\}$ est la suite des polynôme de Tchebychev $P_n = \cos(n \text{Arcos}(x))$, $n \in \mathbb{N}$.

4.6.13 THÉORÈME

Un espace préhilbertien est séparable si et seulement s'il admet une base hilbertienne dénombrable.

Dans ce cas toutes les bases hilbertiennes sont dénombrables.

Démonstration: Soit $\mathbb{D} = \{u_n\}_{n \in \mathbb{N}}$ une suite dense dans H. On peut se ramener au cas où la suite est libre. D'après le procédé de Gram-Schmidt 4.6.10, il existe une suite orthonormée $\{e_n\}_{n \geq 1}$ telle que pour tout $n \in \mathbb{N}$, $\text{vect}\{e_0, \cdots, e_n\} = \text{vect}\{u_0, \cdots, u_n\}$. Il nous reste à montrer que la suite $\{e_n\}_{n \in \mathbb{N}}$ est totale.

Soit $x \in H$ et $\varepsilon > 0$, comme $\mathbb{D} = \{u_n\}_{n \in \mathbb{N}}$ est dense dans H, il existe $N \in \mathbb{N}$ tel que $\|x - u_N\| < \varepsilon$. Comme $u_N \in \text{Vect}\{e_0, \dots, e_N\}$ il existe des scalaires $\lambda_0, \dots, \lambda_N$ tels que $\|x - \sum\limits_{k=0}^{N} \lambda_k e_k\| < \varepsilon$ cqfd.

Réciproquement, soit $\{e_n\}_{n \in \mathbb{N}}$ est une base hilbertienne de H. On pose

$$A = \text{Vect}_{\mathbb{Q}+i\mathbb{Q}}\{e_n\}_{n \in \mathbb{N}} = \{\sum_{n \in I} \lambda_n e_n;\ I \subset \mathbb{N},\ card(I) < +\infty \text{ et } \lambda_n \in \mathbb{Q} + i\mathbb{Q}\}.$$

On vérifie que A est dénombrable et dense dans H. Soit $x \in H$ et $\varepsilon > 0$, alors il existe $\sum_{k=0}^{N} a_k e_k \in \text{Vect}\{e_n\}_{n \in \mathbb{N}}$ tel que $\|x - \sum\limits_{k=0}^{N} a_k e_k\| < \dfrac{\varepsilon}{2}$. Comme $\mathbb{Q} + i\mathbb{Q}$ est dense dans $\mathbb{C}$, il existe pour tout $k \in \{0, \dots, N\}$, $\lambda_k \in \mathbb{Q} + i\mathbb{Q}$ tel que $|a_k - \lambda_k| < \frac{\varepsilon}{2(N+1)}$.

Ainsi $y_\varepsilon = \sum_{k=0}^{N} \lambda_k e_k \in A$ vérifie, $\|x - y_\varepsilon\| \leq \frac{\varepsilon}{2} + \sum_{k=0}^{N} \frac{\varepsilon}{2(N+1)} \leq \varepsilon$ cqfd.

Finalement, soit $\{e_n\}_{n \in \mathbb{N}}$ une base hilbertienne dénombrable de H (on vient de montrer l'existence) et $\beta = \{u_i\}_{i \in I}$ une autre base hilbertienne de H.

On pose, pour tout n, $B_n = \{u \in \beta; \langle e_n, u \rangle \neq 0\}$ et $B = \bigcup\limits_{n \in \mathbb{N}} B_n$.

Comme la famille $\beta = \{u_i\}_{i \in I}$ est une base hilbertienne de H, pour tout n, la série $\sum_{i\in I}\langle e_n, u_i\rangle u_i$ est sommable, ainsi B_n dénombrable (l'annexe : proposition 4.7.17) et par suite B est dénombrable comme réunion dénombrable d'ensembles dénombrables.

Soit $u \in \beta \setminus B$, alors pour tout $n \in \mathbb{N}$, $\langle u, e_n\rangle = 0$ et comme $\{e_n\}_{n\in\mathbb{N}}$ est une base hilbertienne, $u = 0$, ce qui est impossible car $\|u\| = 1$, donc nécessairement $\beta \setminus B = \emptyset$ et par suite $\beta = B$ est dénombrable. ■

4.6.15 EXEMPLE. *La suite orthonormée $\{e_n\}_{n\in\mathbb{N}}$ de $\ell^2(\mathbb{N})$ donnée dans Exemple 4.5.6 est une base hilbertienne, qu'on appelle la base standard de $\ell^2(\mathbb{N})$. On rappelle que*

$$e_n = (\underbrace{0,\cdots,0}_{n-1},1,0,\cdots) \quad \textit{pour tout } n \in \mathbb{N}.$$

En effet, pour tout $x = \{\xi_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$, alors,

$$\|x\|^2 = \sum_{n=1}^{\infty} |\xi_n|^2 = \sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2.$$

d'après le théoreme 4.6.6, $\{e_n\}_{n\in\mathbb{N}}$ est une base hilbertienne. Comme elle est dénombrable, l'espace de Hilbert $\ell^2(\mathbb{N})$ est alors séparable.

Tout espace vectoriel normé de dimension finie n est algebriquement isomorphe et homéomorphe à $\mathbb{K}^n$. Le théorème suivant étend ce résultat aux espaces de Hilbert séparables de dimension infinie

4.6.16 THÉORÈME

Soit $\mathcal{H}$ un espace de Hilbert séparable de dimension finie et $\{e_n\}_{n\in\mathbb{N}}$ une base hilbertienne. Alors, l'application $\varphi : \mathcal{H} \to \ell^2(\mathbb{N})$, définie par $\varphi(x) = \{\langle x, e_n\rangle\}_{n\in\mathbb{N}}$ est un isomorphisme isométrique .

Démonstration: D'après l'inégalité de Bessel 4.5.8, φ est bien définie, i.e. $\varphi(x) \in \ell^2(\mathbb{N})$. Pour tout $x, y \in \mathcal{H}$ et $\alpha, \beta \in \mathbb{K}$,

$$\varphi(\alpha x + \beta y) = \{\langle \alpha x + \beta y, e_n\rangle\}_{n\in\mathbb{N}} = \{\alpha\langle x, e_n\rangle + \beta\langle y, e_n\rangle\}_{n\in\mathbb{N}}$$

$$= \alpha\{\langle x, e_n\rangle\}_{n\in\mathbb{N}} + \beta\{\langle y, e_n\rangle\}_{n\in\mathbb{N}} = \alpha\varphi(x) + \beta\varphi(y),$$

d'ou φ est linéaire. D'après, l'identité de Parseval, φ est une isométrie (donc injective) ; car $\|\varphi(x)\|^2 = \sum_{i\in I} |\langle x, e_n\rangle|^2 = \langle x, x\rangle = \|x\|^2$. On en déduit aussi que φ est surjective, en effet si $(a_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$ alors $x = \sum_{n\in\mathbb{N}} a_n e_n$ est un élément de $\mathcal{H}$, car $\langle x, e_n\rangle = a_n$.

Ainsi φ est un isomorphisme linéaire et isométrique. ■

4.6.2 Le cas général

On va maintenant montrer que tout espace de Hilbert (même s'il n'est pas séparable) admet toujours une base hilbertienne. Ce résultat est basé sur le lemme de Zorn (voir **??***).*

4.6.18 THÉORÈME
Tout espace de Hilbert $\mathcal{H}$ admet une base hilbertienne.

Démonstration: On désigne par $\mathcal{B}$ l'ensemble des systèmes orthonormés de $\mathcal{H}$, ordonné par l'inclusion ($\subset$). Montrons que $(\mathcal{B}, \subset)$ est inductif.

Soit $\mathcal{C} = \{B_i, i \in I\}$ chaîne de $\mathcal{B}$ i.e. une partie totalement ordonnée de $(\mathcal{B}, \subset)$ Alors pour montrer que $\bigcup_{i\in I} B_i$ est un majorant, il suffit de montrer que $\bigcup_{i\in I} B_i$ est un sytème orthonormé. En effet, soit $x, y \in \bigcup_{i\in I} B_i$, $x \neq y$, alors il existe $j \in I$ tel que $x, y \in B_j$, d'où $\|x\| = \|y\| = 1$ et $\langle x, y\rangle = 0$, ainsi on a montrer que $\bigcup_{i\in I} B_i$ est un sytème orthonormé.

D'après le lemme de Zorn, il existe un élément maximal B dans $\mathcal{B}$. Alors B est une base hilbertienne de $\mathcal{H}$, Il suffit de montrer que B est total i.e $\overline{\mathrm{Vect}(B)} = \mathcal{H}$. Sinon, $\mathcal{H} \setminus \overline{\mathrm{Vect}(B)} \neq \emptyset$ et il existerait un x de norme 1 orthogonal à $\overline{\mathrm{Vect}(B)}$. Alors, $B \cup \{x\}$ est un sytème orthonormé qui contient strictement B, ceci contredit le caractère maximal de B. Donc B est total et par suite une base hilbertienne de $\mathcal{H}$.

4.6.20 PROPOSITION
Dans un espace de Hilbert deux bases hilbertiennes sont équipotentes.
Leur cardinal s'appelle la dimension hilbertienne de l'espace de hilbert.

Démonstration: Soit $\mathcal{H}$ un espace de hilbert de dimension infinie. Soient B et B' deux bases de $\mathcal{H}$.

Pour tout $e \in B$, l'ensemble $D_e = \{f \in B'; \langle e, f\rangle \neq 0\}$ est dénombrable d'après , car $e = \sum_{f\in B'} \langle e, f\rangle f$. D'autre part pour tout $f \in B'$, il existe $e \in B$ tel que $\langle e, f\rangle \neq 0$, car $e = \sum_{f\in B'} \langle e, f\rangle f \neq 0$. Ainsi $B' = \bigcup_{e\in B} D_e$, il existe donc une injection de B' dans $\mathbb{N} \times B$. Comme $\mathbb{N} \times B$ est équipotent B (car B est infini), il existe donc une injection de B' dans B. Un raisonnement identique donne une injection de B dans B'.

Finalement, le théorème de Cantor-Bernstein donne une bijection de B sur B'.■

4.6.22 EXEMPLE. Soit I est un ensemble non dénombrable alors,
$\ell^2(I, \mathbb{C}) = \{x = \{x_i\}_{i\in I} : x_i \in \mathbb{K}, \{|x_i|^2\}_{i\in I}$ est sommable $\}$ muni du produit scalaire

$$\langle \{x_i\}_{i\in I}, \{y_i\}_{i\in I}\rangle = \sum_{i\in I} x_i \overline{y_i}$$

est un espace de Hilbert non séparable (voir 4.0.34).
Une base hilbertienne de $\ell^2(I, \mathbb{K})$ est donnée par la famille de vecteurs $\{e_i\}_{i\in I}$ avec $e_i = (\delta_{ij})_{j\in I}$ où δ_{ij} est le symbole de Kronecker défini par $\delta_{ij} = 0$ si $j \neq i$ et $\delta_{ii} = 1$.

Ce théorème est une généralisation de 4.6.16, sa démonstration est identique à celle de 4.6.16

4.6.23 THÉORÈME

Soit $\mathcal{H}$ un espace de Hilbert de dimension hilbertienne I et soit $\{e_i\}_{i\in I}$ une base hilbertienne. Alors, l'application $\varphi : \mathcal{H} \to \ell^2(I)$, définie par $\varphi(x) = \{\langle x, e_n\rangle\}_{i\in I}$ est un isomorphisme isométrique .

4.7 Applications

4.7.1 Séries de Fourier

Soit T un réel > 0. Une fonction $f : \mathbb{R} \to \mathbb{C}$ est dite de période T ou T-périodique si $f(x+T) = f(x)$ pour tout $x \in \mathbb{R}$. On a alors, $f(x+nT) = f(x)$ pour tout $x \in \mathbb{R}$ et tout $n \in \mathbb{Z}$. Il est souvent plus commode de voir les fonctions de période T comme des fonctions sur l'espace quotient $\mathbb{T} = \mathbb{R}/T\mathbb{Z}$ de $\mathbb{R}$ sur $T\mathbb{Z}$ par la relation $x \sim y$ si et seulement si $x - y \in T\mathbb{Z}$ i.e. il existe $n \in \mathbb{Z}$ tel que $x = y + nT$.

Soit $\pi : \mathbb{R} \to \mathbb{T}$ l'application qui associe à un point de $\mathbb{R}$ sa classe d'équivalence modulo $T\mathbb{Z}$.

Alors si $\tilde{f} : \mathbb{R} \to \mathbb{C}$ est une fonction T-périodique, il existe une unique fonction $f : \mathbb{T} \to \mathbb{C}$ telle que $f = \tilde{f} \circ \pi$. Réciproquement, étant donné une fonction $f : \mathbb{T} \to \mathbb{C}$ alors $\tilde{f} = f \circ \pi : \mathbb{R} \to \mathbb{C}$ est une fonction T-périodique. Ainsi les fonctions sur $\mathbb{T}$ correspondent aux fonctions sur $\mathbb{R}$ admettant la période T ; ou encore aux fonctions f définies sur un intervalle $[a, a+T]$ telles que $f(a) = f(a+T)$ quel que soit a.
$\mathbb{T}$ est muni de la topologie quotient i.e. la topologie telle que $f : \mathbb{T} \to \mathbb{C}$ est continue si et seulement si $f \circ \pi : \mathbb{R} \to \mathbb{C}$; ainsi l'espace $C(\mathbb{T})$ des fonctions continues sur $\mathbb{T}$ s'identifie avec l'espace des fonctions continues et T-périodiques sur $\mathbb{R}$.

Soit $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$ le cercle unité. L'application $\phi : \mathbb{R} \to \mathbb{S}^1$ définie par $\phi(t) = e^{i\omega t}$ est surjective continue où $\omega = \frac{2\pi}{T}$. L'application induite $\psi : \mathbb{T} \to \mathbb{S}^1$ qui à la classe d'équivalence de $t \in \mathbb{R}$ associe $e^{i\omega t}$ est une bijection continue donc un homéomorphisme car $\mathbb{T}$ est compact, et $\psi \circ p = \phi$ i.e. le diagramme suivant est commutatif

$$\begin{tikzcd} \mathbb{R} \arrow[rr, "\phi"] \arrow[dr, "p"'] & & \mathbb{S}^1 \\ & \mathbb{T} \arrow[ur, "\psi"'] & \end{tikzcd}$$

Ainsi pour toute fonction $f : \mathbb{R} \to \mathbb{C}$, T-périodique et continue il existe une unique fonction continue $\tilde{f} : \mathbb{S}^1 \to \mathbb{C}$ telle que, pour tout $t \in \mathbb{R}$, $f(t) = \tilde{f}(e^{i\omega t})$.

Polynômes trigonométriques

Soit $T > 0$, on pose $\omega = \frac{2\pi}{T}$. Pour $n \in \mathbb{Z}$, on note e_n, la fonction T-périodique de $\mathbb{R}$ à valeurs dans $\mathbb{C}$ définie par $e_n(t) := e^{i\omega nt}$.

Soit $n \in \mathbb{N}$; un **polynôme trigonométrique de degré** *n est une expression de la forme $\sum_{k=-n}^{n} \lambda_k e_k$ avec $\lambda_k \in \mathbb{C}$ et λ_n ou $\lambda_{-n} \neq 0$. On note par $\mathcal{P}_n = Vect\{e_k \,|\, 0 \leq k \leq n\}$ l'espace des polynômes trigonométriques de degré $\leq n$ et par $\mathcal{P}$ l'espace des polynômes trigonométriques i.e. $\mathcal{P} = \bigcup_{n \geq 0} \mathcal{P}_n$.*

4.7.1 Définition (coefficients de Fourier)

On définit les **coefficients de Fourier** de f pour une fonction intégrable, $\widehat{f}(n)$, par :

$$\widehat{f}(n) = \frac{1}{T} \int_0^T f(t) e^{-i\omega nt}\, dt, \quad \forall n \in \mathbb{Z}.$$

La suite $\left(\widehat{f}(n)\right)_{n \in \mathbb{Z}}$ est appelée la suite des coefficients de Fourier de f.

La série de Fourier de f est la série $Sf = \sum_{n \in \mathbb{Z}} \widehat{f}(n) e_n$ et la suite des sommes partielles $S_N f = \sum_{n=-N}^{N} \widehat{f}(n) e_n$.

La série de Fourier de f est convergente si la suite des sommes partielles $(S_n f)$ converge.

4.7.2 Remarque

Si f est une fonction paire i.e. $f(-x) = f(x)$ alors $\widehat{f}(-n) = \widehat{f}(n)$ et sa série de Fourier peut s'écrire $Sf(x) = \widehat{f}(0) + 2\sum_{n \geq 1} \widehat{f}(n) \cos \omega nx$ Si f est une fonction impaire i.e. $f(-x) = -f(x)$ alors $\widehat{f}(-n) = -\widehat{f}(n)$ et sa série de Fourier peut s'écrire $Sf(x) = 2i \sum_{n \geq 1} \widehat{f}(n) \sin \omega nx$

4.7.3 Proposition

Pour $p \in [1, +\infty]$, $f \in L^p(\mathbb{T})$ et $n \in \mathbb{Z}$ on a

$$|\widehat{f}(n)| \leq \|f\|_p$$

les coefficients de fouriers sont bien définis.

Démonstration: Soit q l'exposant conjugué de p, alors l'inégalité de Hölder et $|e^{-i\omega nt}| = 1$, nous donne :

$$|\widehat{f}(n)| \leq \frac{1}{T} \int_0^T |f(t) e^{-i\omega nt}|\, dt \leq \|f\|_p \|e_{-n}\|_q = \|f\|_p.$$

On va étudier les questions suivantes :
- *Sous quelles conditions la suite des coefficients de Fourier est un élément de $\ell^2(\mathbb{Z})$?*
- *La suite $(S_N f)_{N\in\mathbb{N}}$ converge-t-elle ? si oui, dans quelle sens ?*
- *sous quelles conditions (et dans quel sens) a-t-on l'égalité*

$$f = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e_n$$

4.7.2 Série de Fourier d'une fonction de $L^2([a,b])$

Soit $a < b$. On pose $T = b - a$ et $\omega = \frac{2\pi}{T}$. On munit $L^2([a,b])$ du produit scalaire $\langle f, g\rangle = \dfrac{1}{T}\displaystyle\int_a^b f(t)\overline{g}(t)dt$.
Alors

4.7.5 THÉORÈME
1) $L^2([a,b])$ est un espace de hilbert séparable.

2) La suite $(e_n)_{n\in\mathbb{Z}}$ est une base hilbertienne de $L^2([a,b])$.

Démonstration: **1)** La complétude est une conséquence du théorème de Riesz-Fischer 1.5.10.
2) La suite $(e_n)_{n\in\mathbb{Z}}$ est orthonormée.
En, effet $\langle e_n, e_m\rangle = \dfrac{1}{T}\displaystyle\int_a^b e_n(t)\overline{e}_m(t)dt = \dfrac{1}{T}\displaystyle\int_a^b e^{i\omega(n-m)t}dt = \begin{cases} 1 & \text{si } n = m \\ \left[\frac{e^{i\omega(n-m)t}}{i\omega(n-m)}\right]_a^b = 0 & \text{si } n \neq m \end{cases}$
On va montrer que c'est une base hilbertienne. Soit $f \in L^2([a,b])$ et $\varepsilon > 0$. On prolonge f par 0 en dehors de $[a,b]$, d'après le théorème de densité, il existe $g \in C_c(\mathbb{R},\mathbb{C})$ telle que $\|f - g\|_2 \leq \frac{\varepsilon}{2}$. On peut supposer, quitte à modifier g au voisinage des point a et b que $g(a) = g(b) = 0$. On peut donc prolonger g par T-périodicité à $\mathbb{R}$. On a ainsi une fonction $g : \mathbb{R} \to \mathbb{C}$ continue et T- périodique. Soit $\tilde{g} : \mathbb{S}^1 \to \mathbb{C}$, la fonction continue telle que $\tilde{g}(e^{i\omega t}) = g(t)$ pour tout $t \in \mathbb{R}$. D'après le théorème de Stone-Weierstrass complexe 2.5.21, l'ensemble des polynômes à coefficients complexes en les variables z et $\overline{z}$ sont denses dans $C(\mathbb{S}^1,\mathbb{C})$. Ainsi, il existe $\tilde{P}(z) = \displaystyle\sum_{0\leq n,m\leq N} a_{n,m}z^n\overline{z}^m$ tel que $\|\tilde{g} - \tilde{P}\|_\infty \leq \frac{\varepsilon}{2}$. Si on pose, pour tout $t \in \mathbb{R}$, $P(t) = \tilde{P}(e^{i\omega t})$. Alors P est un polynôme trigonométrique et $\|g - P\|_\infty \leq \frac{\varepsilon}{2}$.
Ainsi, $\|f - P\|_2 \leq \|f - g\|_2 + \|g - P\|_2 \leq \|f - g\|_2 + \|g - P\|_\infty \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. ∎

En identifiant une fonction T-périodique, $f : \mathbb{T} \to \mathbb{C}$ avec sa restriction sur l'intervalle $[0,T]$ on a comme conséquence du théorème précédent et des propriétés d'une base hilbertienne :

4.7.7 THÉORÈME

L'application $\Phi : L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$, $f \mapsto (\widehat{f}(n))_{n\in\mathbb{Z}}$ est un isomorphisme isométrique donc linéaire bijective et pour tout $f \in L^2(\mathbb{T})$, $\|f\|_2 = \|\widehat{f}(n)\|_2$. Autrement dit, pour toute fonction $f : \mathbb{T} \to \mathbb{C}$ T-périodique et telle que $\int_0^T |f(t)|^2\,dt$ on a :

(i) $f(t) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{i\omega nt}$ dans $L^2(\mathbb{T})$, i.e. $\lim\limits_{N\to\infty} \|f - S_N f\|_2 = 0$,

où $S_N f = \sum\limits_{-N}^{N} \widehat{f}(n)e^{i\omega nt}, \quad N \in \mathbb{N}$.

(ii) l'identité de Parseval

$$\frac{1}{T}\int_0^T |f(t)|^2\,dt = \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2.$$

(iii) Pour tout $f, g \in L^2(\mathbb{T})$.

$$\langle f, g\rangle = \frac{1}{T}\int_0^T f(t)\overline{g(t)}\,dt = \sum_{n=-\infty}^{+\infty} \widehat{f}(n)\overline{\widehat{g}(n)}.$$

4.7.8 REMARQUE

Ce résultat ne dit rien sur la convergence simple de la série de Fourier d'une fonction de $L^2(\mathbb{T})$, on a pour cela le théorème suivant :

4.7.9 THÉORÈME (CARLESON(1965))

Si $f \in L^2(\mathbb{T})$, la série de Fourier de f converge pour presque tout x vers f.

D'autres conséquences de l'existence d'une base hilbertienne

4.7.10 PROPOSITION

Tout espace de Hilbert de dimension infinie admet un sous-espace isomorphe à $\ell^2(\mathbb{N})$.

Démonstration: En effet, un espace de Hilbert de dimension infinie, admet une base hilbertienne (infinie) et donc une suite orthonormée $(e_n)_{n\in\mathbb{N}}$ et $E = \overline{(e_n)_{n\in\mathbb{N}}}$ est alors isomorphe à $\ell^2(\mathbb{N})$. ■

En retrouve un résultat valable pour les espaces de Banach, dans le cas des espace de Hilbert, mais sans utilisé le théorème de Baire.

4.7.12 COROLLAIRE
Si un espace de Hilbert admet une base de Hamel dénombrable, alors il est de dimension finie.

Démonstration: Si l'espace de Hilbert est de dimension infinie, il admet d'après la proposition un sous-espace isomorphe à $\ell^2(\mathbb{N})$.

Il suffit de montrer que $\ell^2(\mathbb{N})$ n'admet pas de base de Hamel dénombrable.

Pour $\alpha \in]0,1[$ On pose $x_\alpha = (1+\alpha, \dots, \alpha^n, \dots)$.

Alors $x_\alpha \in \ell^2(\mathbb{N})$ et la famille (non dénombrale) $\{x_\alpha\}_{\alpha\in]0,1[}$, est une famille libre de $\ell^2(\mathbb{N})$, ce qui montre qu'une base de Hamel de $\ell^2(\mathbb{N})$ ne peut pas être dénombrable.

Comme $\|x_\alpha\|^2 = \sum_{n=0}^{+\infty} \alpha^{2n} = \frac{1}{1-\alpha^2} < +\infty$, ona bien $x_\alpha \in \ell^2(\mathbb{N})$.

Soit maintenant, $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ et α_i distincts tels que $\sum_{i=1}^{k} x_{\alpha_i} = 0$.

Alors, pour tout $n \in \mathbb{N}$, on a $\sum_{i=1}^{k} \alpha_i^n = 0$.

En particulier, en considérant, $0 \le n \le k-1$, on obtient le système

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_{k-1} \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_{k-1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \dots & \alpha_{k-1}^{k-1} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

qui pour solution unique $\lambda_1 = \dots = \lambda_k = 0$, Alorsque les α_i sont distincts et que la matrice de Vandermonde a pour déterminant $\prod_{i<j}(\alpha_j - \alpha_i)$. ■

4.7.3 Annexe

Familles sommables

4.7.14 DÉFINITION
Soit $(E, \|.\|)$ un espace vectoriel normé. Soit $(x_i)_{i\in I}$ une famille d'éléments de E et $S \in E$.
On dit que la famille $(x_i)_{i\in I}$ est **sommable** et de somme S et on écrit $S = \sum_{i\in I} x_i$, si pour tout $\varepsilon > 0$, il existe une partie finie $J_\varepsilon \subset I$ telle que pour toute partie finie $K \subset I$ contenant J_ε on ait $\|S - \sum_{i\in K} x_i\| < \varepsilon$.

4.7.15 REMARQUE

1. La somme d'une famille sommable est unique.
2. Soit $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ sont des familles sommables de somme respectivement S et S'. Alors pour tout λ , $\mu \in \mathbb{K}$ on la famille $(\lambda x_i + \mu y_i)_{i\in I}$ est sommable et de somme $\lambda S + \mu S'$.

4.7.16 Exercice Soit $(E, \|.\|_E)$ et $(F, \|.\|_F)$ deux espaces vectoriels normés et $f \in \mathcal{L}(E, F)$. Soit $(x_i)_{i \in I}$ une famille sommable de E, de somme S.
Montrer que $(f(x_i))_{i \in I}$ est une famille sommable et que $f(S) = \sum_{i \in I} f(x_i)$.

Dans une famille sommable il n'y a qu'un nombre dénombrable de termes non nuls

4.7.17 PROPOSITION
Soit $(E, \|.\|)$ un espace vectoriel normé. Si $(x_i)_{i \in I}$ est une famille sommable de E, alors l'ensemble $D = \{i \in I;\ x_i \neq 0\}$ est (au plus) dénombrable.

Démonstration: D'après la définition de famille sommable, pour tout entier $n \geq 1$, il existe un sous-ensemble fini J_n de I, tel que pour tout sous-ensemble fini $L \subset I - J_n$ on ait $\| \sum_{i \in L} x_i \| < \frac{1}{n}$. En particulier, pour tout $i \in I - J_n$ on a $\|x_i\| \leq \frac{1}{n}$. D'autre part, $D = \displaystyle\bigcup_{n \geq 1} \{i \in I; \|x_i\| \geq \frac{1}{n}\}$ (en effet, si $x_i \neq 0$, alors il existe $n_0 \geq 1$ tel que $\|x_i\| \geq \frac{1}{n_0}$ i.e. $x_i \in \{i \in I; \|x_i\| \geq \frac{1}{n_0}\}$.

Finalement, $D \subset \bigcup_{n \geq 1} J_n$ est au plus dénombrable (contenu dans une réunion dénombrable d'ensembles finis). ∎

4.7.19 DÉFINITION (CRITÈRE DE CAUCHY)
Soit $(E, \|.\|)$ un espace vectoriel normé.

On dit que la famille $(x_i)_{i \in I}$ vérifie le critère de Cauchy, si pour tout $\varepsilon > 0$, il existe une partie finie $J_\varepsilon \subset I$ telle que pour toute partie finie $L \subset I$ disjointe de J_ε on ait $\| \sum_{i \in L} x_i \| < \varepsilon$.

4.7.20 THÉORÈME
(i) Toute famille sommable vérifie le critère de Cauchy.
(ii) Si $(E, \|.\|)$ est un espace de Banach, on a réciproquement, toute famille vérifiant le critère de Cauchy est sommable

---

4.8 La transformation de Fourier

La théorie des séries de Fourier s'applique aux fonctions périodiques ou de façon équivalente aux fonctions sur le cercle unité $\mathbb{S}^1$*, elle associe à une fonctions périodique une suite, ses coefficients de Fourier.*

Dans cette partie, on développe une théorie analogue, valable pour les fonctions sur la droite réelle, non nécéssairement périodiques, on va associé à une fonction intégrable, une autre fonction, qui est une "version continue" des coéfficients de Fourier.

Soit $f \in L^1(\mathbb{R})$*, alors pour tout* $y \in \mathbb{R}$ *on a* $|f(x)e^{-2i\pi xy}| = |f(x)|$*, d'où* $\int_{\mathbb{R}} |f(x)e^{-2i\pi xy}|\,dx = \int_{\mathbb{R}} |f(x)|\,dx = \|f\|_1 < +\infty$*, on peut donc définir :*

4.8.1 Définition

1. Soit $f \in L^1(\mathbb{R})$. On appelle **transformée de Fourier**, la fonction $\widehat{f}$ définie par

$$\widehat{f}(y) = \int_{\mathbb{R}} f(x)e^{-2i\pi x.y}\,dx, \quad \forall y \in \mathbb{R}.$$

2. La **transformation de Fourier** est l'application

$$\begin{array}{rcl} \widehat{\ }: L^1(\mathbb{R}) & \longrightarrow & \mathbb{C}^{\mathbb{R}} \\ f & \longmapsto & \widehat{f} \end{array}$$

4.8.2 Remarque

1. En utilisant $e^{-2i\pi xy} = \cos(2\pi xy) - i\sin(2\pi xy)$, on a les implications :
   f est paire presque partout $\Rightarrow \widehat{f}$ est réelle
   f est impaire presque partout $\Rightarrow \widehat{f}$ est imaginaire.
2. On a $\overline{\widehat{f}}(y) = \overline{\int_{\mathbb{R}} f(x)e^{-2i\pi xy}\,dx} = \int_{\mathbb{R}} \overline{f}(x)e^{2i\pi xy}\,dx = \widehat{\overline{f}}(-y)$.
3. On utilise aussi la notation $\mathscr{F}(f)$ pour désigner $\widehat{f}$.

Dans la partie qui va suivre on va énoncer les propriétés de base de la transformation de Fourier.

Propriétés de la transformée de Fourier

Les fonctions considérées dans cette section, seront dans $L^1(\mathbb{R})$.

$\boxed{P_1}$ $\|\widehat{f}\|_\infty \leq \|f\|_1$.

Démonstration: En effet, pour tout $y \in \mathbb{R}$, $|\widehat{f}(y)| = \left|\int_{\mathbb{R}} |f(x)e^{-2i\pi xy}|\,dx\right| \leq \int_{\mathbb{R}} |f(x)|\,dx = \|f\|_1$. ∎

$\boxed{P_2}$ *L'application $\widehat{f} : \mathbb{R} \to \mathbb{C}$ est continue.*

Démonstration: Soit (y_k) une suite dans $\mathbb{R}$ telle que $\lim_{k\to+\infty} y_k = y$. Alors pour presque tout $x \in \mathbb{R}$, on a $\lim_{k\to+\infty} f(x)e^{-2i\pi xy_k} = f(x)e^{-2i\pi xy}$, et comme $|f(x)e^{-2i\pi xy_k}| = |f(x)|$ est intégrable et ne dépend pas de y_k, d'après le théorème de convergence dominée, $\lim_{k\to+\infty} \widehat{f}(y_k) = \widehat{f}(y)$. ∎

$\boxed{P_3}$ $\widehat{f(0)} = \int_{\mathbb{R}} f(x)\,dx$.

$\boxed{P_4}$ *Si $f, g \in L^1(\mathbb{R})$, alors $f * g \in L^1(\mathbb{R})$ et $\widehat{f * g} = \widehat{f}\,\widehat{g}$, où $f * g = \int_{\mathbb{R}} f(x-z)g(z)\,dz$ est le produit de convolution de f et g.*

Démonstration: D'après l'inégalité de Young (voir 4.8.43) , $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$, d'où $f * g \in L^1(\mathbb{R})$.
Pour $y \in \mathbb{R}$ fixé, on a, à l'aide du théorème de Fubini,
$\widehat{f * g}(y) = \int_{\mathbb{R}} f * g(x)e^{-2i\pi xy}\,dx = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x-z)g(z)e^{-2i\pi xy}\,dz\,dx$
$= \int_{\mathbb{R}} g(z)e^{-2i\pi xz}\left(\int_{\mathbb{R}} f(x-z)e^{-2i\pi(x-z)y}\,dx\right)dz = \widehat{f}(y)\int_{\mathbb{R}} g(z)e^{-2i\pi xz})\,dz = \widehat{f}(y)\,\widehat{g}(y)$. ∎

4.8.6 Définition

Soit $f : \mathbb{R} \to \mathbb{C}$.

1. Soit $a \in \mathbb{R}$. On définit l'opérateur translation τ_a par :

$$\tau_a f(x) = f(x+a).$$

2. Pour $\lambda \in \mathbb{R}^*$. On définit l'opérateur dilatation σ_λ par

$$\sigma_\lambda f(x) = f(\lambda x).$$

$\boxed{P_5}$ *(a)* $\widehat{\tau_a f}(y) = e^{2i\pi ay}\widehat{f}(y)$.

(b) $\widehat{\sigma_\lambda f}(y) = \frac{1}{|\lambda|}\widehat{f}(\frac{y}{\lambda})$.

Démonstration: (a) $\widehat{\tau_a f}(y) = \int_{\mathbb{R}} f(x+a)e^{-2i\pi xy}\,dx = \int_{\mathbb{R}} f(x+a)e^{-2i\pi(x+a)y}e^{2i\pi ay}\,dx = e^{2i\pi ay}\int_{\mathbb{R}} f(x+a)e^{-2i\pi(x+a)y}\,dx = e^{2i\pi ay}\widehat{f}(y)$.

(b) En utilisant le changement de variable $z = \lambda x$, on obtient
$\widehat{\sigma_\lambda f}(y) = \int_{\mathbb{R}} f(\lambda x)e^{-2i\pi xy}\,dx = \frac{1}{|\lambda|}\int_{\mathbb{R}} f(z)e^{-2i\pi z\frac{y}{\lambda}}\,dz$ ∎

4.8.8 LEMME (DE RIEMANN-LEBESGUE)

$\boxed{P_6}$ Si $f \in L^1(\mathbb{R})$ alors $\lim_{|y|\to+\infty} \widehat{f}(y) = 0$.

Démonstration: D'après la propriété $\boxed{P_5}$, $\int_{\mathbb{R}} f(x+z)e^{-2i\pi xy}\,dx = \widehat{\tau_z f}(y) = e^{2i\pi zy}\widehat{f}(y)$. Si on pose $z = \frac{1}{2y}$, alors $e^{2i\pi zy} = e^{i\pi} = -1$.

D'où, $2\widehat{f}(y) = \int_{\mathbb{R}} (f(x) - f(x+z))\,e^{-2i\pi xy}\,dx$. Ainsi, $2|\widehat{f}(y)| \leq \int_{\mathbb{R}} |f(x) - f(x+z)|\,dx = \|f - \tau_z f\|_1$.

D'où, $\displaystyle\lim_{|y|\to+\infty} |\widehat{f}(y)| \leq \frac{1}{2}\lim_{z\to 0} \|f - \tau_z f\|_1 = 0.$ ∎

$\boxed{P_7}$ *(Formule de transfert) Si $f, g \in L^1(\mathbb{R})$, alors :*

$$\int_{\mathbb{R}} \widehat{f}(x)g(x)\,dx = \int_{\mathbb{R}} f(x)\widehat{g}(x)\,dx$$

Démonstration: Comme $f, g \in L^1(\mathbb{R})$, $\widehat{f}$ et $\widehat{g} \in C_0 \subset L^\infty(\mathbb{R})$, d'où $\widehat{f}g$ et $f\widehat{g} \in L^1$. Ainsi les deux intégrales sont bien définies.
Par le théorème de Fubini on a :

$$\int_{\mathbb{R}} \widehat{f}(x)g(x)\,dx = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(z)e^{-2i\pi zx}\,dz\right) g(x)\,dx$$
$$= \int_{\mathbb{R}} f(z) \left(\int_{\mathbb{R}} g(x)e^{-2i\pi zx}\,dx\right)\,dz = \int_{\mathbb{R}} f(z)\widehat{g}(z)\,dz.$$

4.8.11 EXEMPLE. (i) soit $a > 0$ et $f = \mathbb{1}_{[-a,a]}$ Alors, si $y \neq 0$, $\widehat{f}(y) = \int_{-a}^{a} e^{-2i\pi xy}\,dx = 2\int_0^a \cos(2\pi xy)\,dx = 2\dfrac{\sin(2\pi ay)}{2\pi y}$ et $\widehat{f}(0) = \int_{-a}^{a} dx = 2a$.
(ii) $g(x) = e^{-x}\mathbb{1}_{\mathbb{R}_+}$. Alors $\widehat{f}(y) = \int_0^{+\infty} e^{-x}e^{-2i\pi xy}\,dx = \int_0^{+\infty} e^{-x(1+2i\pi y)}\,dx = \frac{1}{1+2i\pi y}$.

On peut résumer les propriétés $\boxed{P_1}$, $\boxed{P_2}$ *et* $\boxed{P_6}$ *par*

4.8.12 THÉORÈME

La transformation de Fourier $\widehat{\ }: L^1(\mathbb{R}) \longrightarrow C_0(\mathbb{R})$ est une application linéaire continue de norme 1 de $L^1(\mathbb{R})$ muni de la norme $\|.\|_1$ dans $C_0(\mathbb{R})$ muni de la norme $\|.\|_\infty$.

On peut se poser les questions suivantes, l'opérateur précédent est-il injectif ? surjectif ?

On montre qu'il existe des fonctions dans $C_0(\mathbb{R})$ qui ne sont pas les transformées de Fourier d'aucune fonction intégrable i.e. l'opérateur n'est pas surjectif. Voir 4.8.5 pour une formule d'inversion, sous certaines conditions, de la transformation de Fourier dans L^1.

Dans la section suivante, on va montrer que cette opérateur est un isomorphisme l'espace de Schwartz et admet une extension en un isomorphisme isométrique à $L^2(\mathbb{R}, \mathbb{C})$.

4.8.1 Transformation de Fourier sur $L^2(\mathbb{R}, \mathbb{C})$

4.8.2 L'espace de Schwartz

4.8.13 DÉFINITION

1. Une fonction $f \in C^\infty(\mathbb{R})$ est à dite **à décroissance rapide** si pour tout $k \in \mathbb{N}$, il existe une constante $C_k > 0$ telle que pour tout $x \in \mathbb{R}$ on ait $|x|^k |f(x)| \leq C_k$.
2. L'espace de Schwartz $\mathcal{S}(\mathbb{R})$ est l'espace de toutes les fonctions de classe C^∞, à décroissances rapides ainsi que toutes leurs dérivées partielles i.e.

$$\mathcal{S}(\mathbb{R}) = \left\{ f \in C^\infty(\mathbb{R}) \,\middle|\, \text{pour tout } k, l \in \mathbb{N}, \sup_{x \in \mathbb{R}} |x^k f^{(l)}(x)| < +\infty \right\}.$$

4.8.14 REMARQUE

1. Soit $f \in \mathcal{S}(\mathbb{R})$, alors pour tout $k \in \mathbb{N}$ et $l \in \mathbb{N}$, $\lim_{|x| \to +\infty} |x|^k f^{(l)}(x) = 0$.
   Comme, il existe $C > 0$ telle que $\sup_{x \in \mathbb{R}} |x|^{k+1} |f^{(l)}(x)| \leq C$
   alors $\lim_{|x| \to +\infty} |x|^k f^{(l)}(x) \leq \lim_{|x| \to +\infty} \frac{C}{|x|} = 0$.
2. Pour tout $P \in \mathbb{R}[x]$ et $f \in \mathcal{S}(\mathbb{R})$, la fonction produit $P.f \in \mathcal{S}(\mathbb{R})$
   En effet, $P.f$ est de classe C^∞, et $x^k (Pf)^{(l)}(x)$ est une combinaison linéaire finie de de termes de la forme $x^{k'} f^{(l')}(x)$, est donc bornée, Alorsque $f \in \mathcal{S}(\mathbb{R})$.
3. On a $C_c^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R}) \subset L^p(\mathbb{R})$ pour tout $p \in [1, +\infty]$.
   Si $f \in C_c^\infty(\mathbb{R})$, $x^k f^{(l)}(x)$ est a support compact, donc bornée.
   Si $f \in \mathcal{S}(\mathbb{R})$, alors f est bornée, d'où $\mathcal{S}(\mathbb{R}) \subset L^\infty(\mathbb{R})$.
   Soit $p \in [1, +\infty[$ et $f \in \mathcal{S}(\mathbb{R})$, d'après 1), il existe $C > 0$
   telle que $(1 + |x|^2)|f(x)| \leq C$, d'où
   $\int_{\mathbb{R}} |f(x)|^p \, dx \leq C \int_{\mathbb{R}} \frac{1}{(1+|x|^2)^p} \, dx < +\infty$ car $2p > 1$.
   Donc $f \in L^p(\mathbb{R})$.
   On a $\mathcal{S}(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$ pour $1 \leq p < +\infty[$, car $C_c^\infty(\mathbb{R})$ est dense dans $L^p(\mathbb{R})$.

4.8.15 EXEMPLE. $e^{-a|x|^2}$ avec $a > 0$, $e^{-|x|^{2k}}$ avec $k \in \mathbb{N}$, $x^5 e^{x^4 |x|^2}$ et $\frac{1}{\cos h(x)}$ sont des éléments de $\mathcal{S}(\mathbb{R})$.

Par contre, $e^{-|x|}$, $e^{-|x|^3}$, $\sin(x)$ et toute fonction polynomiale non identiquement nulle, ne sont pas des éléments de $\mathcal{S}(\mathbb{R})$.

4.8.16 PROPOSITION
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors pour tout $k \in \mathbb{N}$, $f^{(k)}$ et $x^k f$ sont dans $\mathcal{S}$.

Démonstration: En effet, d'après la formule Leibnitz $(x^k f)^{(l)}(x)$ est une combinaison linéaire finie de termes de la forme $x^{k'} f^{(l')}(x)$ est donc bornée sur $\mathbb{R}$, pour tout $k, l \in \mathbb{N}$, par suite $x^k f \in \mathcal{S}(\mathbb{R})$. ■

4.8.18 Proposition (Transformation de Fourier et différentiabilité)
Soit $f \in \mathcal{S}(\mathbb{R})$. Alors

1. pour tout $k \in \mathbb{N}$ et tout $y \in \mathbb{R}$:

$$\left(\widehat{f}\right)^{(k)}(y) = \widehat{(-2i\pi x)^k f}(y)$$

2. pour tout $k \in \mathbb{N}$ et tout $y \in \mathbb{R}$:

$$\widehat{f^{(k)}}(y) = (2i\pi y)^k \widehat{f}(y)$$

Démonstration: On va démontrer les résultats pour $k = 1$, i.e ; pour les dérivées partielles d'ordre 1, le cas général se fait par récurrence.

1. On pose $g(x, y) = f(x)e^{-2i\pi xy}$, alors $\dfrac{\partial g}{\partial y} = -2i\pi x f(x)e^{-2i\pi xy}$ et $|\dfrac{\partial g}{\partial y}| \leq 2\pi|x||f(x)|$ qui ne dépend pas de y et qui est dans L^1 d'après le proposition précédente.

   D'après le théorème de dérivation sous le signe $\int$,

$$\begin{aligned}\widehat{f}'(y) &= \frac{\partial}{\partial y}\int_{\mathbb{R}} f(x)e^{-2i\pi xy}\,dx = \int_{\mathbb{R}} \frac{\partial}{\partial y}\left(f(x)e^{-2i\pi xy}\right)\,dx \\ &= \int_{\mathbb{R}} (-2i\pi x) f(x)e^{-2i\pi xy}\,dx.\end{aligned}$$

2. Comme $f \in \mathcal{S}(\mathbb{R})$, $\lim_{x\to+\infty} f(x)e^{-2i\pi xy}$ et $\lim_{x\to+\infty} f(x)e^{-2i\pi xy}$ existent et sont nécessairement égales à 0 et une intégration par parties donne alors :

$$\widehat{f'}(y) = \int_{\mathbb{R}} \frac{\partial f}{\partial x}(x)e^{-2i\pi xy}\,dx = \lim_{\substack{B\to+\infty \\ A\to-\infty}} \left[f(x)e^{-2i\pi xy}\right]_A^B + 2i\pi y\int_{\mathbb{R}} f(x)e^{-2i\pi xy}\,dx = 2i\pi y\widehat{f}(y).$$

4.8.20 Corollaire
$\mathcal{S}(\mathbb{R})$ est stable par la transformation de Fourier i.e. $f \in \mathcal{S}(\mathbb{R})$ alors $\widehat{f} \in \mathcal{S}(\mathbb{R})$.

Démonstration: Soit $f \in \mathcal{S}(\mathbb{R})$, d'après la proposition précédente, $k \in \mathbb{N}$, $f^{(k)}$ et $x^k f$ sont dans $L^1(\mathbb{R})$. Donc $\widehat{f}$ est à décroissance rapide.

D'après la proposition, pour tous k et l dans $\mathbb{N}$ on a

$$y^k(\widehat{f})^{(l)}(y) = y^k\widehat{(-2i\pi x)^l f}(y) = \frac{1}{(2i\pi)^k}\widehat{\left((-2i\pi x)^l f\right)^{(k)}}(y)$$

d'où $y^k(\widehat{f})^{(l)}$ est bornée Alorsque c'est la transformée de Fourier d'une fonction dans L^1. ∎

Exemple fondamental : La gaussienne $e^{-\pi x^2}$ et le noyau K_δ

4.8.22 LEMME
$\int_{\mathbb{R}} e^{-\pi x^2}\,dx = 1.$

Démonstration: En effet, un changement de variable en coordonnée polaires donne
$\left(\int_{\mathbb{R}} e^{-\pi x^2}\,dx\right)^2 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\pi(x^2+y^2)}\,dx = \int_0^{2\pi}\int_0^{+\infty} e^{-\pi r^2}\,r dr\,d\theta$
$= 2\pi\left[\frac{e^{-\pi r^2}}{2\pi}\right]_0^{+\infty} = 1.$ ∎

4.8.24 LEMME
Si $\phi(x) = e^{-\pi x^2}$ alors : $\widehat{\phi} = \phi$

Démonstration: On a $\phi'(x) = -2\pi x\phi(x)$, d'où $\widehat{\phi'} = \widehat{-2\pi x\phi}$ ce qui donne $2i\pi y\widehat{\phi} = -i\widehat{\phi}'$. Ainsi $\widehat{\phi}$ est solution de l'équation différentielle, $-2\pi y\widehat{\phi} = i\widehat{\phi}'$ i.e. $\widehat{\phi}(y) = ce^{-\pi y^2}$, et $c = \widehat{\phi}(0) = \int_{\mathbb{R}} e^{-\pi x^2}\,dx = 1$. Finalement, $\widehat{\phi}(y) = e^{-\pi y^2}$. ∎

4.8.26 COROLLAIRE
Si $\delta > 0$ et $K_\delta = \frac{e^{\frac{-\pi x^2}{\delta}}}{\sqrt{\delta}}$. Alors $\widehat{K_\delta}(y) = e^{-\pi\delta y^2}$ et $\widehat{\widehat{K_\delta}} = K_\delta$.

Démonstration: En remarquant que $K_\delta = \frac{1}{\sqrt{\delta}}\sigma_{\frac{1}{\sqrt{\delta}}}\phi$, on obtient en utilisant les propriétés de la transformation de Fourier que :
$\widehat{K_\delta}(y) = \frac{1}{\sqrt{\delta}}\widehat{\sigma_{\frac{1}{\sqrt{\delta}}}\phi}(y) = \frac{1}{\sqrt{\delta}}.\sqrt{\delta}.\widehat{\phi}(\sqrt{\delta}y) = \phi(\sqrt{\delta}y) = e^{-\pi\delta x^2}.$ ∎

4.8.28 PROPOSITION (PROPRIÉTÉ DU NOYAU K_δ)
.

1. $K_\delta \geq 0$
2. $\int_{\mathbb{R}} K_\delta(x)\,dx = 1.$
3. Pour tout $\eta > 0$, on a $\lim_{\delta \to 0^+} \int_{|x| \geq \eta} K_\delta(x)\,dx = 0.$

Démonstration: 1. $\frac{e^{\frac{-\pi x^2}{\delta}}}{\sqrt{\delta}} \geq 0.$

2. Par le changement de variable $x = \sqrt{\delta} t$ on a $\int_{\mathbb{R}} K_\delta(x)\,dx = \int_{\mathbb{R}} e^{-\pi t^2}\,dt = 1.$
3. Le même changement de variable $x = \sqrt{\delta} t$, donne, pour $\eta > 0$, $\lim_{\delta \to 0^+} \int_{|x| \geq \eta} K_\delta(x)\,dx = \lim_{\delta \to 0^+} \int_{|t| \geq \frac{\eta}{\sqrt{\delta}}} e^{-\pi t^2}\,dt = \lim_{r \to +\infty} \int_{|t| \geq r} e^{-\pi t^2}\,dt = 0$
car $\int_{\mathbb{R}} e^{-\pi t^2}\,dt$ est convergente. ■

4.8.30 COROLLAIRE
La famille $\{K_\delta\}_{\delta > 0}$ est une identité approchée.

4.8.31 COROLLAIRE
Pour tout $f \in \mathcal{S}(\mathbb{R})$ et tout $x \in \mathbb{R}$
$\lim_{\delta \to 0^+} f * K_\delta(x) = f(x).$

4.8.32 THÉORÈME (D'INVERSION DE FOURIER)
Soit $f \in \mathcal{S}(\mathbb{R})$ alors pour tout $x \in \mathbb{R}$,
$\widehat{\widehat{f}}(-x) = f(x)$ i.e $f(x) = \int_{\mathbb{R}} \widehat{f}(y) e^{2i\pi xy}\,dy$ (Formule d'inversion).

Par conséquent, la transformation de Fourier est un isomorphisme linéaire de $\mathcal{S}(\mathbb{R})$ sur lui-même.

Démonstration: On va commencer par montrer la formule pour $x = 0$ i.e. $f(0) = \int_{\mathbb{R}} \widehat{f}(y)\,dy$.

D'après, la corollaire 4.8.26 et la formule de transfert :

$$\int_{\mathbb{R}} f(x)K_\delta(x)\,dx = \int_{\mathbb{R}} f(x)\widehat{\widehat{K_\delta}}\,dx = \int_{\mathbb{R}} \widehat{f}(y)\widehat{K_\delta}(y)\,dx = \int_{\mathbb{R}} \widehat{f}(y)e^{-\pi\delta y^2}\,dx.$$

Comme $|\widehat{f}(y)e^{-\pi\delta y^2}| \leq |\widehat{f}(y)|$ qui est dans $\mathcal{S} \subset L^1$, d'après le le théorème de convergence dominée on a

$$\lim_{\delta\to 0^+} \int_{\mathbb{R}} \widehat{f}(y)e^{-\pi\delta y^2}\,dy = \int_{\mathbb{R}} \widehat{f}(y)(\lim_{\delta\to 0^+} e^{-\pi\delta y^2})\,dy = \int_{\mathbb{R}} \widehat{f}(y)\,dy$$

D'autre part, le corollaire 4.8.31, donne

$$\lim_{\delta\to 0^+} \int_{\mathbb{R}} f(x)K_\delta(x)\,dx = \lim_{\delta\to 0^+} f * K_\delta(0) = f(0).$$

Ainsi, on a montré que $f(0) = \int_{\mathbb{R}} \widehat{f}(y)\,dy$. Pour le cas, de x quelconque (fixé), on pose pour tout $t \in \mathbb{R}.$, $F(t) = f(t+x)$. Alors

$$f(x) = F(0) = \int_{\mathbb{R}} \widehat{F}(y)\,dy = \int_{\mathbb{R}} \widehat{f}(y)e^{2i\pi xy}\,dy.$$

Comme $\widehat{\ }: \mathcal{S} \to \mathcal{S}$, on peut appliquer la formule d'inversion de Fourier, de sorte que pour tout $f \in \mathcal{S}$ on a $\widehat{\widehat{f}}(-x) = f(x)$, d'où la bijectivité de $\widehat{\ }\mathcal{S}$ sur lui-même. ∎

4.8.34 COROLLAIRE

si $f, g \in \mathcal{S}$ alors :

1) $\widehat{f * g} = \widehat{f}\widehat{g}$

2) $\widehat{f} * \widehat{g} = \widehat{fg}$

Démonstration:

1. On l'a déjà montrer pour L^1, propriété $\boxed{P_4}$, elle est donc valable en particulier pour $\mathcal{S}$.
2. Par la bijectivité de la transformation de Fourier sur $\mathcal{S}$, il existe $u, v \in \mathcal{S}$ tels que $f = \widehat{u}$ et $g = \widehat{v}$ alors, $\widehat{fg}(x) = \widehat{\widehat{u}\widehat{v}}(x) = \widehat{\widehat{u * v}}(x) = u * v(-x)$ d'où $\widehat{fg}(x) = \widehat{f} * \widehat{g}$. ∎

4.8.36 COROLLAIRE (L'IDENTITÉ DE PARSEVAL)

1. Pour tout $f \in \mathcal{S}(\mathbb{R})$ on a $\|\widehat{f}\|_2 = \|f\|_2$
2. Pour tout $f, g \in \mathcal{S}(\mathbb{R})$, $\langle \widehat{f}, \widehat{g} \rangle = \langle f, g \rangle$.
   Ainsi, la transformation de Fourier est un isomorphisme isométrique de $\mathcal{S}(\mathbb{R})$ sur lui-même.

Démonstration: 1. Soit $f, g \in \mathcal{S}(\mathbb{R})$, alors $= \widehat{fg} = \widehat{f} * \widehat{g}$ en particulier pour $y = 0$ on a :

$$\int_{\mathbb{R}} f(x)g(x)\,dx = \widehat{fg}(0) = \widehat{f} * \widehat{g}(0) = \int_{\mathbb{R}} \widehat{f}(z)\widehat{g}(-z)\,dz.$$

Si on pose $g = \overline{f}$, alors $\widehat{g}(-z) = \widehat{\overline{f}}(-z) = \int_{\mathbb{R}} \overline{f}(x)e^{2i\pi xz}\,dx = \overline{\int_{\mathbb{R}} f(x)e^{-2i\pi xz}\,dx} = \overline{\widehat{f}}(z)$. On obtient finalement,

$$\|f\|_2^2 = \int_{\mathbb{R}} f(x)\bar{f}(x)\,dx = \int_{\mathbb{R}} \widehat{f}(z)\overline{\widehat{f}}(z)\,dz = \|\widehat{f}\|_2^2.$$

2. On utilise l'identité de polarisation. ∎

4.8.3 Transformation de Fourier sur $L^2(\mathbb{R}, \mathbb{C})$

On va maintenant étendre la transformation de Fourier de $\mathcal{S}(\mathbb{R})$ à $L^2(\mathbb{R})$. On remarquera qu'on ne peut pas étendre directement la formule de la transformation de Fourier, à $L^2(\mathbb{R})$, en effet si $f \in L^2(\mathbb{R}) \setminus L^1(\mathbb{R})$ alors $\int_{\mathbb{R}} f(x)e^{2i\pi xy}\,dx$ diverge pour tout $y \in \mathbb{R}$; donc $\widehat{f}$ n'a pas de sens en général pour $f \in L^2(\mathbb{R})$. Pour surmonter cette difficulté, on va utiliser la densité de l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ dans $L^2(\mathbb{R})$ et le théorème de prolongement des application uniformément continues (voir 4.7.3), pour étendre de façon unique l'isomorphisme isométrique $\widehat{\ }: \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ en un isomorphisme isométrique $\mathscr{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$, appelé transformation de Fourier-Plancherel.

4.8.38 THÉORÈME (FOURIER-PLANCHEREL)

Pour tout $f \in L^2(\mathbb{R})$, on associe une fonction $\mathscr{F}(f) \in L^2(\mathbb{R})$ telle que :

1. $\mathscr{F}|_{\mathcal{S}(\mathbb{R})} = \widehat{\ }|_{\mathcal{S}(\mathbb{R})}$
2. Pour tout $f \in L^2(\mathbb{R})$, $\mathscr{F}(\mathscr{F}(f))(x) = f(-x)$ p.p.
3. (l'identité de Parseval) Pour $f \in L^2(\mathbb{R})$, $\|\mathscr{F}(f)\|_2 = \|f\|_2$,
   et pour f, g éléments de $L^2(\mathbb{R})$ on a :

$$\langle \mathscr{F}(f), \mathscr{F}(g) \rangle = \langle f, g \rangle$$

4. Si $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, alors $\mathscr{F}(f) = \widehat{f}$
   i.e. les transformations de Fourier et de Fourier-Plancherel coïncident dans $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.
5. L'application $f \mapsto \mathscr{F}(f)$ est une isomorphisme isométrique de l'espace de Hilbert $L^2(\mathbb{R})$ sur lui-même.
6. Pour $f \in L^2(\mathbb{R})$ on pose

$$\varphi_n(y) = \int_{\mathbb{R}} \mathbb{1}_{\{\|x\| \leq n\}} f(x) e^{-2i\pi xy}\, dx, \quad \forall y \in \mathbb{R},$$

   et

$$\psi_n(x) = \int_{\mathbb{R}} \mathbb{1}_{\{\|x\| \leq n\}} \mathscr{F}(f)(t) e^{2i\pi xt}\, dt, \quad \forall x \in \mathbb{R},$$

   alors $\|\varphi_n - \mathscr{F}(f)\|_2 \to 0$ et $\|\psi_n - f\|_2 \to 0$ quand $n \to \infty$,

---

Démonstration: 1. Comme $\mathcal{S}(\mathbb{R})$ est dense dans $L^2(\mathbb{R})$ et que $\widehat{\ }: \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ est une application unitaire d'après le théorème de prolongement, il existe une et une seule application linéaire continue $\mathscr{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ telle que pour tout $f \in \mathcal{S}(\mathbb{R})$ on ait $\mathscr{F}(f) = \widehat{f}$.
   Soit $f \in L^2(\mathbb{R})$, il existe une suite $f_k \in \mathcal{S}(\mathbb{R})$, telle que $\lim_{k\to+\infty} \|f_k - f\|_2 = 0$. Alors $\mathscr{F}(f) = \lim_{k\to+\infty} \widehat{f_k}$.

2. Alors $\mathscr{F}(\mathscr{F}(f))(x) = \mathscr{F}(\mathscr{F}(\lim_{k\to+\infty} f_k))(x) = \lim_{k\to+\infty} \widehat{\widehat{f_k}}(x) = \lim_{k\to+\infty} f_k(-x) = f(-x)$.
3. Comme $\mathscr{F}$ et $\|.\|_2$ sont continues on a

$$\|\mathscr{F}(f)\|_2 = \|\lim_{k\to+\infty} \widehat{f_k}\|_2 = \lim_{k\to+\infty} \|\widehat{f_k}\|_2 = \lim_{k\to+\infty} \|f_k\|_2 = \|\lim_{k\to+\infty} f_k\|_2 = \|f\|_2,$$

   d'où l'identité de Parseval.
4. Soit $\phi_n \in C_c^\infty$, telle que $0 \leq \phi_n \leq 1$ et $\phi_n \equiv 1$ dans $B(0, n)$.
   Alors pour tout $g \in \mathcal{S}$ on a $\langle \widehat{f}\phi_n, g\rangle = \langle \widehat{f}, \phi_n g\rangle = \int_{\mathbb{R}} \widehat{f}\, \overline{\phi_n g} = \int_{\mathbb{R}} f\, \overline{\widehat{\phi_n g}} = \int_{\mathbb{R}} f\, \overline{\mathscr{F}(\phi_n g)} = \int_{\mathbb{R}} \mathscr{F}(f) \overline{\phi_n g} = \langle \mathscr{F}(f), \phi_n g\rangle = \langle, \mathscr{F}(f)\phi_n, g\rangle$.
   D'où $\langle (\mathscr{F}f - \widehat{f})\phi_n, g\rangle = 0$ pour tout $g \in \mathcal{S}$ i.e. $(\mathscr{F}f - \widehat{f})\phi_n \in \mathcal{S}^\perp$. Comme $\mathcal{S}$ est dense dans L^2, $(\mathscr{F}f - \widehat{f})\phi_n = 0$ ce qui entraîne que $\mathscr{F}f = \widehat{f}$ sur $B(0, n)$ pour tout n, i.e. $\mathscr{F}f = \widehat{f}$. ∎

4.8.40 REMARQUE

On voit que l'espace $L^2(\mathbb{R})$ se comporte mieux par rapport à la transformation de Fourier, Alorsque c'est une application unitaire de $L^2(\mathbb{R})$ sur lui-même, alors que pour $L^1(\mathbb{R})$ il n'est pas facile de caractériser l'image de la transformation de Fourier.

Cependant, l'expression de la transformée de Fourier $\mathscr{F}(f)$ est moins simple pour $f \in L^2(\mathbb{R})$, que pour $f \in L^1(\mathbb{R})$, Alorsqu'elle se calcule par passage à la limite.

4.8.41 EXEMPLE. Soit $f = e^{-x}\mathbb{1}_{\mathbb{R}_+} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Comme $\widehat{f}(y) = \frac{1}{1+2i\pi y} \in L^2(\mathbb{R}) \setminus L^1(\mathbb{R})$. La transformation de Fourier-Plancherel nous donne alors : $\mathscr{F}(\mathscr{F}(f))(x) = f(-x)$ i.e.

$$\mathscr{F}(\widehat{f})(x) = \mathscr{F}\left(\frac{1}{1+2i\pi y}\right)(x) = \begin{cases} 0 & \text{si } x > 0 \\ e^x & \text{si } x \leq 0 \end{cases}$$

(iii) $h(x) = xe^{-x}\mathbb{1}_{\mathbb{R}_+}$. Alors $h(x) = xg(x)$, d'où

$$\widehat{h}(y) = \widehat{xg}(y) = \frac{-1}{2i\pi}\frac{\partial \widehat{g}}{\partial y}(y) = \frac{-1}{2i\pi}\frac{-2i\pi}{(1+2i\pi y)^2} = \frac{1}{(1+2i\pi y)^2}.$$

Complément : La transformation de Laplace.

Une autre transformation, très importante pour les applications est la *transformation de Laplace*, qui est de la même espèce que la transformation de Fourier. Étant donnée une fonction $f(t)$ définie sur l'intervalle $[0, +\infty[$, on appelle transformée de Laplace de f la fonction

$$\mathcal{L}(f)(z) = \int_0^{+\infty} f(t)\, e^{-zt}\, dt \tag{6.1}$$

losque l'intégrale précédente converge.

Voyons quelques exemples

4.8.42 EXEMPLE.
1. Si f est à croissance au plus exponentielle, i.e. il existe des constantes positives M et α telles que $\forall t \geq 0$, $|f(t)| \leq M\, e^{\alpha t}$, alors la transformée de Laplace $\mathcal{L}(f)(z)$ est une fonction analytique de z dans le demi-plan $\mathfrak{Re}(z) > \alpha$. Pour $\mathfrak{Re}(z) \leq \alpha$, l'intégrale sera en général divergente, mais cela n'interdit pas que $\mathcal{L}(f)(z)$ puisse avoir un prolongement analytique au-delà du demi-plan $\mathfrak{Re} z > \alpha$.
2. La fonction $f(t) = e^{t^2}$ par exemple, n'a pas de transformée de Laplace.
3. Soit $f(t) = t^{\alpha-1}$. Cette fonction est bien à croissance au plus exponentielle, sa transformée de Laplace est $\mathcal{L}(f)(z) = \int_0^{+\infty} t^{\alpha-1}\, e^{-zt}\, dt$. Cette fonction est analytique dans le domaine $\mathbb{C}$ privé de la demi-droite $]-\infty, 0]$. Mais l'intégrale n'est convergente que pour $\mathfrak{Re} z > 0$ et par conséquent ne représente une fonction analytique que dans ce demi-plan.

4. Soit $f(t) = e^{at}$, a étant un nombre complexe quelconque. En prenant $M = 1$ et $\alpha = \mathfrak{Re}(a)$ on a bien l'inégalité $|f(t)| \leq M e^{\alpha t}$, et on s'attend donc à ce que $F(z)$ soit analytique dans le demi-plan $\mathfrak{Re} z > \alpha$. Le calcul direct donne $\mathcal{L}(f)(z) = 1/(z-a)$; cette fonction est en effet analytique dans le demi-plan $\mathfrak{Re} z > \alpha$, mais se prolonge à $\mathbb{C} - \{a\}$.

Il y a une parenté entre la transformation de Laplace et la transformation de Fourier. En effet, soit $f(t) = e^{-\alpha t}$ une fonction définie sur l'intervalle $]0, +\infty[$ et $\mathcal{L}(f)(z)$ sa transformée de Laplace, définie et analytique au moins dans le demi-plan $\mathfrak{Re} z > \alpha$. Posons alors

$$\varphi(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ f(t) & \text{si } t > 0 \end{cases}$$

Il est clair que la transformée de Fourier $\widehat{\varphi}(y)$ est égale à $\mathcal{L}(f)(-2i\pi y)$. Autrement dit, les valeurs de la transformée de Laplace $\mathcal{L}(f)(z)$ le long de la droite $\mathfrak{Re} z = 0$ représentent la transformée de Fourier de la fonction $\varphi(t)$; plus généralement, les valeurs de $\mathcal{L}(f)(z)$ le long de la droite $\mathfrak{Re} z = a$ représentent la transformée de Fourier de la fonction $\varphi(x)\, e^{-ax}$; en effet :

$$\mathcal{L}(f)(a + 2i\pi y) \;=\; \int_{-\infty}^{+\infty} \varphi(t)\, e^{-at}\, e^{-2i\pi t y}\, dt.$$

D'après la formule d'inversion, on en déduit que :

$$\varphi(t)\, e^{-at} \;=\; \int_{-\infty}^{+\infty} \mathcal{L}(f)(a + 2i\pi y)\, e^{2i\pi t y}\, dy$$

ou encore

$$\varphi(t) \;=\; \int_{-\infty}^{+\infty} \mathcal{L}(f)(a + 2i\pi y)\, e^{t(a+2i\pi y)}\, dy \tag{6.4}$$

On peut interpréter le membre de droite de comme l'intégrale obtenue par paramétrage de

$$\frac{1}{2\pi} \int_{\Gamma_a} \mathcal{L}(f)(z)\, e^{tz}\, dz$$

où Γ_a est le chemin (infini) constitué par la droite $\mathfrak{Re} z = a$ parcourue du bas vers le haut.

4.8.4 Annexe

4.8.43 Théorème (l'inégalité de young)

Soit p, q et $r \in [1, +\infty]$ tels que, $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$.

Alors pour f et g des fonctions mesurables de $\mathbb{R}^n \to \mathbb{C}$ on a :

$$\|f * g\|_r \leq \|f\|_p \|h\|_q$$

En particulier, si $f \in L^p(\mathbb{R}^n)$ et $g \in L^q(\mathbb{R}^n)$, alors $f * g \in L^r(\mathbb{R}^n)$.

Démonstration: On a $|f*g(x)| \leq |f|*|g|(x) = \int_{\mathbb{R}^n} |f(x-y)||g(y)|\,dy$

$$= \int_{\mathbb{R}^n} \left(|f(x-y)|^p|g(y)|^q\right)^{\frac{1}{r}} |f(x-y)|^{p(\frac{1}{p}-\frac{1}{r})}|g(y)||^{q(\frac{1}{q}-\frac{1}{r})}\,dy$$

Comme $1 = (\frac{1}{p}-\frac{1}{r}) + (\frac{1}{q}-\frac{1}{r}) + \frac{1}{r}$, l'inégalité de Hölder nous donne :

$$|f*g(x)| \leq \|\left(|f|^p(x-.)|g|^q\right)^{\frac{1}{r}}\|_r.\||f|^{p(\frac{1}{p}-\frac{1}{r})}(x-.)\|_{\frac{pr}{r-p}}.\||g|^{q(\frac{1}{q}-\frac{1}{r})}\|_{\frac{qr}{r-q}}$$

d'où

$$|f*g(x)| \leq \left(|f|^p*|g|^q(x)\right)^{\frac{1}{r}}\|f\|_p^{p(\frac{1}{p}-\frac{1}{r})}\|g\|_q^{q(\frac{1}{q}-\frac{1}{r})}.$$

Alors

$$\int_{\mathbb{R}^n} |f*g(x)|^r\,dx \leq \left(\int_{\mathbb{R}^n} |f|^p*|g|^q(x)\,dx\right)\|f\|_p^{r-p}\|g\|_q^{r-q}$$

$$\leq \left(\||f|^p\|_1\||g|^q\|_1\right)\|f\|_p^{r-p}\|g\|_q^{r-q} = \left(\|f\|_p^p\|g\|_q\right)\|f\|_p^{r-p}\|g\|_q^{r-q}$$

finalement, $\|f*g\|_r \leq \|f\|_p\|g\|_q$ ∎

4.8.45 THÉORÈME (LE THÉORÈME DE PROLONGEMENT)

Soit (E, d_E) un espace métrique, $D \subset E$ une partie dense et (F, d_F) un espace métrique complet.

Soit $f : A \to F$ une application uniformément continue.

Alors, il existe une unique application uniformément continue $\tilde{f} : E \to F$ telle que $\tilde{f}|_D = f$.

En particulier

4.8.46 COROLLAIRE

$(E, \|.\|_E)$ est un espace vectoriel normé, D un sous-espace vectoriel dense de E et $(F, \|.\|_F)$ est un espace de Banach. Alors pour toute application linéaire continue $f : D \to F$, il existe une unique application linéaire continue $\tilde{f} : E \to F$ telle que, $\tilde{f}|_D = f$.

Démonstration: Comme f est linéaire continue, pour tout $x, x' \in E$, $\|f(x)-f(x')\|_F \leq \|f\|.\|x-x'\|_E$, elle est en particulier uniformément continue et d'après le théorème de prolongement, il existe une unique application uniformément continue $\tilde{f} : E \to F$ telle que $\tilde{f}|_D = f$.

Soit $x, x' \in E$ α et $\beta \in \mathbb{C}$. Par densité, il existe $\{x_n\}$, $\{x'_n\}$ dans D telles que $\lim_{n\to+\infty} x_n = x$ et $\lim_{n\to+\infty} x'_n = x'$. Alors, par continuité on a $\tilde{f}(\alpha x + \beta x') = \tilde{f}\left(\lim_{n\to+\infty}(\alpha x_n + \beta x'_n)\right) = \lim_{n\to+\infty} \tilde{f}(\alpha x_n + \beta x'_n) = \lim_{n\to+\infty} f\left((\alpha x_n + \beta x'_n)\right) = \lim_{n\to+\infty} \alpha f(x_n) + \beta f(x'_n) = \alpha\tilde{f}(x) + \beta\tilde{f}(x')$. D'où $\tilde{f}$ est linéaire. ∎

4.8.5 Complément : Le théorème d'inversion pour $L^1(\mathbb{R}, \mathbb{C})$

4.8.48 THÉORÈME (THÉORÈME D'INVERSION)
Soit $f \in L^1(\mathbb{R})$ telle que $\widehat{f} \in L^1(\mathbb{R})$, alors $\widehat{\widehat{f}}(x) = f(-x)$ p.p.

Démonstration: Soit $\phi : \mathbb{R} \to \mathbb{R}$ telle que $\phi(x) = e^{-\pi\|x\|^2}$. Alors $\int_{\mathbb{R}} \phi(x)\, dx = 1$ et pour tout $x \in \mathbb{R}$, $\lim_{t\to 0^+} \sigma_t \phi(x) = \lim_{t\to 0^+} \phi(tx) = 1$.

Par définition $\widehat{\widehat{f}}(x) = \int_{\mathbb{R}} \widehat{f}(y) e^{-2i\pi xy}\, dy$

Comme $|\widehat{f}(y)e^{-2i\pi xy}\phi(ty) \leq |\widehat{f}(y)|$ qui est par hypothèse dans L^1, d'après le le théorème de convergence dominée on a $\widehat{\widehat{f}}(x) = \lim_{t\to 0^+} \int_{\mathbb{R}} \widehat{f}(y)e^{-2i\pi xy}\phi(ty)\, dy$.

D'autre part, en utilisant le lemme de transfert, on a

$$\int_{\mathbb{R}} \widehat{f}(y)e^{-2i\pi xy}\phi(ty)\, dy = \int_{\mathbb{R}} \widehat{\tau_x(f)}(y)\phi(ty)\, dy = \int_{\mathbb{R}} \tau_x(f)(y)\widehat{\sigma_{\frac{1}{t}}\phi}(y)\, dy$$

$$= t^{-n} \int_{\mathbb{R}} \tau_x(f)(y)\phi(\frac{y}{t})\, dy = \int_{\mathbb{R}} f(z-x)\phi_t(z)\, dz = f * \phi_t(-x).$$

Comme $\phi_t(z) = t^{-n}\phi(\frac{y}{t})$ est une approximation de l'identité ,
$\lim_{t\to 0^+} f * \phi_t(-x) = f(-x)$ dans L^1. D'après 1.5.10, il existe une sous-suite $\{f * \phi_{t_k}\}_{k\in\mathbb{N}}$ qui converge presque partout i.e. $\lim_{k\to+\infty} f * \phi_{t_k}(-x) = f(-x)$ p.p.
Finalement,

$$\widehat{\widehat{f}}(x) = \lim_{k\to+\infty} \int_{\mathbb{R}} \widehat{f}(y)e^{-2i\pi xy}\phi(t_k y)\, dy = \lim_{k\to+\infty} f * \phi_{t_k}(-x) = f(-x) \text{ p.p.}$$

4.8.50 COROLLAIRE (THÉORÈME D'UNICITÉ)
La transformation de Fourier est injective i.e. si $f, g \in L^1(\mathbb{R})$, tels que $\widehat{f} = \widehat{g}$, alors $f(x) = g(x)$ p.p.

Démonstration: La transformation de Fourier étant linéaire, il suffit de montrer que $\widehat{f} = 0 \implies f = 0$. Comme $\widehat{f} = 0 \in L^1$, le théorème d'inversion s'applique et donne $f = 0$ p.p. ■

4.8.52 COROLLAIRE
La transformation de Fourier $\widehat{\ }: L^1(\mathbb{R}) \longrightarrow C_0(\mathbb{R})$ est un monomorphisme d'algèbre de Banach, où $L^1(\mathbb{R})$ est muni de la norme $\|.\|_1$ et du produit de convolution, alors que $C_0(\mathbb{R})$ est muni de la norme $\|.\|_\infty$ et de la multiplication ordinaire.

4.8.53 COROLLAIRE (L'IDENTITÉ DE PARSEVAL)
Si $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ alors $\widehat{f} \in L^2(\mathbb{R})$ et

$$\|\widehat{f}\|_2 = \|f\|_2$$

4.8.54 EXEMPLE. Si $f(x) = e^{-|x|}$, alors $\widehat{f}(y) = \frac{2}{(1+4\pi^2 y)^2}$. Comme $\widehat{f} \in L^1(\mathbb{R})$, on peut appliquer le théorème d'inversion pour obtenir,

$\widehat{\widehat{f}}(x) = f(-x)$ pour tout $x \in \mathbb{R}$ (car, f est continue). En utilisant la parité de f, on obtient l'identité : $\displaystyle\int_0^{+\infty} \frac{cos(2\pi xy)}{1+4\pi^2 y^2}\, dy = \frac{e^{-|x|}}{4}$.

Chapitre 5

Opérateurs compacts et théorie spectrale sur les espaces de Hilbert

5.1 Opérateurs compacts

Opérateurs compacts constituent une classe importante d'opérateurs linéaires bornés. D'une part, ils sont presque des opérateurs de rang fini (i.e. d'image de dimension finie).

D'autre part, la classe des opérateurs compacts est suffisamment large pour inclure iles opérateur à noyau continus où dans L^2.

5.1.1 Définition

Soient E et F deux espaces vectoriels normés. Une application $T \in \mathcal{L}(E, F)$ est dite **compacte** (ou un opérateur compact) si : $\overline{T(\bar{B}_E)}$ est un compact de F,

où $\bar{B}_E$ est la boule unité fermée de E.

Notation : On note pat $K(E, F)$ l'ensemble des opérateurs compacts de E dans F.

Dans le cas $E = F$ on note simplement cet espace par $K(E)$.

5.1.2 Proposition

Soient E et F deux espaces vectoriels normés et $T \in \mathcal{L}(E, F)$.

Les conditions suivantes sont équivalentes :

i) T est compact.

ii) Pour tout $A \subset E$ borné, $\overline{T(A)}$ est compact.

iii) Toute suite bornée $(x_n)_{n\in\mathbb{N}}$ de E, $(Tx_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence.

Démonstration: 1. $i) \Rightarrow ii)$ Soit $A \subset E$ borné, alors il existe $r > 0$ tel que $A \subset r.\bar{B}_E$ d'où $\overline{T(A)} \subset r\overline{T(\bar{B}_E)}$. Ainsi, $\overline{T(A)}$ est compact, comme fermé du compact $r.\overline{T(\bar{B}_E)}$.

$ii) \Rightarrow iii)$ Il suffit de poser $A = \{x_n,\ n \in \mathbb{N}\}$.

$iii) \Rightarrow i)$ Soit $(y_n)_{n\in\mathbb{N}}$ une suite de $\overline{T(\bar{B}_E)}$.

pour tout $n \in \mathbb{N}$ il existe $z_n \in T(\bar{B}_E)$ tel que $\|y_n - z_n\|_F \leq 2^{-n}$.

Comme par hypothèse $(z_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence, il en est de même pour $(y_n)_{n\in\mathbb{N}}$. ■

5.1.1 Propriétés de base des opérateurs compacts

5.1.4 PROPOSITION (PROPRIÉTÉS DES OPÉRATEURS COMPACTS)

Soit X et Y deux espaces de Banach.

(i) L'ensemble des opérateurs compacts $K(X,Y)$ est un sous-espace vectoriel fermé de $\mathcal{L}(X,Y)$.

(ii) Soient $S \in \mathcal{L}(X,Y)$ et $T \in \mathcal{L}(Y,Z)$, alors si S ou T est compact, $T \circ S \in K(X,Z)$.

En particulier, $K(X)$ est un *idéal bilatère de* $\mathcal{L}(X)$.

Démonstration: i) Soient $\lambda, \beta \in \mathbb{K}$ et $S, T \in K(X,Y)$. Soit (x_n) une suite bornée de X. Comme S et T sont compacts, il existe une sous-suite $(x_{\phi(n)})$ telle que $S(x_{\phi(n)})$ et $T(x_{\phi(n)})$ convergent, ainsi $\lambda S(x_{\phi(n)}) + \beta T(x_{\phi(n)})$ converge. Donc, toute suite bornée, son image par $\lambda S + \beta T$ admet une valeur d'adhérence i.e. $\lambda S + \beta T \in \mathbb{K}(X,Y)$.

Fermeture : *Soit $T \in \overline{K(X,Y)}$, alors pour tout $\varepsilon > 0$, il existe $T_\varepsilon \in K(X,Y)$ tel que $\|T - T_\varepsilon\| \leq \frac{\varepsilon}{2}$, cela signifie que*

$$\|Tx - T_\varepsilon x\| \leq \frac{\varepsilon}{2} \quad \textit{pour tout } x \in \bar{B}_X.$$

Comme T_ε est compact, $T_\varepsilon(\bar{B}_X)$ est précompact, il exise donc $N_\varepsilon \in \mathbb{N}$ et $\{y_1, \ldots, y_{N_\varepsilon}\}$ tels que

$$T_\varepsilon(\bar{B}_X) \subset \bigcup_{i=1}^{N_\varepsilon} B(y_i, \frac{\varepsilon}{2}).$$

Ainsi, pour tout $x \in \bar{B}_X$, il existe $i_0 \in \{1, \ldots, N_\varepsilon\}$ tel que $\|T_\varepsilon x - y_{i_0}\| \leq \frac{\varepsilon}{2}$, alors $\|Tx - y_{i_0}\| \leq \|Tx - T_\varepsilon x\| + \|T_\varepsilon x - y_{i_0}\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. D'où $T(\bar{B}_X) \subset \bigcup_{i=1}^{N_\varepsilon} B(y_i, \varepsilon)$, par suite $\overline{T(\bar{B}_X)}$ est compact car Y est complet. On a donc montrer que $T \in K(X,Y)$. Finalement $\overline{K(X,Y)} = K(X,Y)$.

ii) Supposons $S \in K(X,Y)$. Comme $T(S(\bar{B}_X)) \subset T(\overline{S(\bar{B}_X)})$, ce dernier est compact, comme image du compact $\overline{S(\bar{B}_X)}$ par l'application continue T.
Ainsi $\overline{T \circ S(\bar{B}_X)}$ est compact i.e. $T \circ S \in K(X,Y)$. ■

5.1.6 Corollaire (Les isomorphismes ne sont pas compacts)

Soit X un espace vectoriel normé de dimension infinie. Alors, l'opérateur identité de X n'est pas compact. Plus généralement, tout isomorphisme $T : X \to X$ n'est pas compact.

Démonstration: Pour l'opérateur identité I sur X, ona $\overline{I(\bar{B}_X)} = \bar{B}_X$, qui ne peut être compacte car la dimension de X est infinie. Quant à l'assumption générale, si un isomorphisme $T : X \to X$ est compact alors l'opérateur identité $I = T^{-1} \circ T$ serait compact, ce qui serait une contradiction. ■

5.1.8 Définition

Soient X et Y deux espaces vectoriels normés. Une application $T \in \mathcal{L}(X,Y)$ est dite **de rang fini** si la dimension de l'image de L est finie i.e. $\dim L(X) < +\infty$.

5.1.9 Remarque

Tout opérateur de rang fini est compact. En effet, $\overline{T(\bar{B}_X)}$ est un fermé borné de l'espace de dimension finie $L(E)$, est donc compact.

Comme $K(X,Y)$ est fermé, il s'ensuit que tout opérateur qui peut être approché par des opérateurs de rang fini est également compact :

5.1.10 Corollaire (Opérateurs de rang presque fini sont compacts)

Soit X et Y deux espaces de Banach. Soit $T \in \mathcal{L}(X,Y)$, tel qu'il existe $T_n \in L(X,Y)$, T_n de rang fini, et $\lim_{n\to+\infty} \|T_n - T\| = 0$. Alors T est compact.

On a la réciproque, tout opérateur compact est limite d'une suite d'opérateurs de rang fini, pour les espaces de Hilbert.

5.1.11 Proposition

Soient H_1 et H_2 deux espaces de Hilbert. Alors :

$T \in K(H_1, H_2)$ si et seulement s'il existe une suite $T_n \in L(X,Y)$, T_n de rang fini, telle que $\lim_{n\to+\infty} \|T_n - T\| = 0$.

Démonstration: Soit $T \in K(H_1, H_2)$ et $\varepsilon > 0$. Comme T est compact, $T\left(\bar{B}_X\right)$ est pré-compact, il exise donc $N_\varepsilon \in \mathbb{N}$ et $\{y_1, \ldots, y_{N_\varepsilon}\}$ tels que

$$T\left(\bar{B}_X\right) \subset \bigcup_{i=1}^{N_\varepsilon} B(y_i, \varepsilon). \tag{5.1.1}$$

On pose $F_\varepsilon = \text{Vect}\{y_1, \ldots, y_{N_\varepsilon}\}$, et $P_\varepsilon : H_2 \to H_2$ la projection orthogonale sur F_ε. Soit $T_\varepsilon = P_\varepsilon \circ T$. Comme $T_\varepsilon(H_1) = P_\varepsilon \circ T(H_1) \subset F_\varepsilon$, T_ε est de rang fini. D'autre part, d'après 5.1.1, pour tout $x \in \bar{B}_X$,

$$\|Tx - T_e x\| = \|Tx - P_\varepsilon(Tx)\| = \inf_{y \in F_\varepsilon} \|Tx - y\| \leq \varepsilon.$$

D'où $\|T - T_\varepsilon\| \leq \varepsilon$. ■

5.1.2 Exemples

L'exemple suivant est l'une des principale motivation pour étudier les opérateurs compacts.

5.1.13 Proposition (Les opérateurs intégraux sont compacts)
Soit $C[0,1]$ munit de la norme $\|.\|_\infty$.
Considérons l'opérateur intégral $T : C[0,1] \to C[0,1]$ défini par

$$(Tf)(t) = \int_0^1 k(t,s) f(s)\, ds$$

avec un "noyau" $k \in C([0,1] \times [0,1])$. Alors T est un opérateur compact.

Démonstration: Nous devons montrer que $K := T(B_{C[0,1]})$ est un sous-ensemble pré-compact de $C[0,1]$. D'après le théorème d'Ascoli -Arzelà 5.3.10, cela découle de la bornitude et l'equicontinuoity de l'ensemble K.

La bornitude : Soit $f \in B_{C[0,1]}$, alors $\|f\|_\infty \leq 1$, d'où

$$\|Tf\|_2^2 = \int_0^1 \left| \int_0^1 k(t,s) f(s)\, ds \right|^2 dt \leq \|f\|_\infty^2 \int_0^1 \left| \int_0^1 k(t,s)\, ds \right|^2 dt \leq \int_0^1 \left| \int_0^1 k(t,s) f(s)\, ds \right|^2 dt = C.$$

D'où $K \subset \bar{B}(0, \sqrt{C})$ est donc borné.

Équicontinuité : Comme k est continue sur le compact $[0,1] \times [0,1]$, pour tout $\varepsilon > 0$ et on choisit $\delta > 0$ tel que

$$|t_1 - t_2| \leq \delta \quad \text{entraîne} \quad |k(t_1,s) - f(t_2,s)| \leq \varepsilon \quad \text{pour tout } s \in [0,1].$$

Maintenant, pour chaque $f \in B_{C[0,1]}$, on obtient par l'inégalité triangulaire que

$$\left|(Tf)(t_1) - (Tf)(t_2)\right|| \leq \int_0^1 |k(t_1,s) - f(t_2,s)|\, |f(s)|\, ds \leq \varepsilon$$

car $|f(s)| \leq 1$ pour tous s. Cela montre que l'ensemble K est équicontinu. Ainsi K est précompact. ■

5.1.15 Exercice Soit une suite de nombres réels $(\lambda_k)_{k\in\mathbb{N}}$, et pour $x = (x_n)_{n\in\mathbb{N}} \in \ell^2$ on pose $Tx := (\lambda_k x_k)_{k\in\mathbb{N}}$.
Sous quelles conditions sur la suite (λ_k) l'opérateur T est bien défini ? continu ? compact ?

Opérateurs de Hilbert-Schmidt

Il s'agit de la classe la plus utilisé d'opérateurs compacts dans les espaces de Hilbert.

5.1.16 Définition
Soit $\mathcal{H}$ un espace de Hilbert séparable, et soit (x_k) une base hilbertienne de $\mathcal{H}$. Un opérateur linéaire $T : \mathcal{H} \to \mathcal{H}$ est un *opérateur de Hilbert-Schmidt* si

$$\sum_{k=1}^{\infty} \|Tx_k\|^2 < \infty.$$

Le réel

$$\|T\|_{\mathrm{HS}} = \left(\sum_{k=1}^{\infty} \|Tx_k\|^2\right)^{1/2}$$

est appelé la *norme de Hilbert-Schmidt* de T.

5.1.17 Exemple. Pour les opérateurs des espaces de dimension finie $T : \mathbb{C}^m \to \mathbb{C}^n$, (de façon équivalente, pour les matrices $n \times m$) , la norme de Hilbert-Schmidt est la norme $\|T\| = \sqrt{\mathrm{tr}\,(T^*T)}$.

5.1.18 Proposition
La définition de l'opérateur de Hilbert-Schmidt et de la norme de Hilbert-Schmidt ne dépendent pas du choix d'une base hilbertienne de $\mathcal{H}$.

Démonstration: Supposons que $\sum_k \|Tx_k\|^2 < \infty$ pour une base hilbertienne (x_k) de $\mathcal{H}$. En utilisant l'identité de Parseval deux fois, on obtient

$$\sum_k \|Tx_k\|^2 = \sum_{k,j} |\langle Tx_k, x_j\rangle|^2 = \sum_{k,j} |\langle x_k, T^*x_j\rangle|^2 = \sum_k \|T^*x_j\|^2. \tag{5.1.2}$$

Soit (x'_k) une autre base hilbertienne de $\mathcal{H}$. Alors un argument similaire donne

$$\sum_j \|T^*x_j\|^2 = \sum_{j,k} |\langle x'_k, T^*x_j\rangle|^2 = \sum_{j,k} |\langle Tx'_k, x_j\rangle|^2 = \sum_k \|Tx'_k\|^2.$$

Ceci termine la preuve. ∎

5.1.20 REMARQUE

1. Une conséquence de la démonstration, on obtient dans (5.1.2) que $\|T^*\|_{\text{HS}} = \|T\|_{\text{HS}}$.
2. $\|T\| \leq \|T\|_{\text{HS}}$.
   Soit x un vecteur de norme 1 et $B = \{x_k\}_{k\in\mathbb{N}}$ une base hilbertienne de $\mathcal{H}$ telle que $x \in B$. Alors $\|Tx\| \leq \sum_k \|Tx_k\|^2 = \|T\|_{\text{HS}}$. En appliquant cette inéhgalité à tous les x tel que $\|x\| = 1$, on obtient le résultat.

5.1.21 PROPOSITION
Tout opérateur de Hilbert-Schmidt T est compact .

Démonstration: Soit une base hilbertienne (x_k) de $\mathcal{H}$. on pose $c = \sum_k \|Tx_k\|^2$.

Soit $\varepsilon > 0$. Alors il existe $N_\varepsilon \in \mathbb{N}$ tel que $\sum_{k>N_\varepsilon} \|Tx_k\|^2 \leq \varepsilon^2$.

On pose $F_\varepsilon = \text{Vect}\{y_1, \ldots, y_{N_\varepsilon}\}$, et $P_\varepsilon : H_2 \to H_2$ la projection orthogonale sur F_ε.

Alors $T_\varepsilon = P_\varepsilon \circ T$ est de rang fini et

$$\|T - T_\varepsilon\|^2 \leq \|T - T_\varepsilon\|^2_{\text{HS}} = \sum_{k>N_\varepsilon} \|Tx_k\|^2 \leq \varepsilon^2.$$

Ainsi T est limite d'opérateurs de rang fini il est donc compact.

Un exemple d'opérateur de Hilbert-Schmidt :

5.1.23 PROPOSITION (OPÉRATEURS INTÉGRAUX DE HILBERT-SCHMIDT)
Considérons l'opérateur intégral $T : L^2([0,1]) \to L^2([0,1])$ défini par

$$(Tf)(t) = \int_0^1 k(t,s)f(s)\,ds$$

avec un noyau $k(t,s) \in L^2([0,1]^2)$. Alors T est un opérateur de Hilbert-Schmidt, et

$$\|T\|_{\text{HS}} = \|k\|_2.$$

Démonstration: Nous allons voir l'intégrale dans la définition de T, comme le produit scalaire de f avec le noyau k. Plus précisément, considérons la fonction $k_t(s) = k(t,s)$, Alors

$$(Tf)(t) = \langle k_t, f\rangle \quad \text{pour tout } t \in [0,1].$$

Fixons une base orthonormée (x_k) de $L^2[0,1]$. Alors

$$\begin{aligned}
\|T\|^2_{\text{HS}} &= \sum_k \|Tx_k\|_2^2 = \sum_k \int_0^1 |(Tx_k)(t)|^2\,dt = \sum_k \int_0^1 |\langle k_t, x_k\rangle|^2\,dt \\
&= \int_0^1 \sum_k |\langle k_t, x_k\rangle|^2\,dt \quad \text{(par le théorème de convergence monotone)} \\
&= \int_0^1 \|k_t\|_2^2\,dt \quad \text{(par l'identité de Parseval)} \\
&= \|k\|_2^2 \quad \text{(Par définition de } k_t \text{ et le théorème de Fubini)}
\end{aligned}$$

Ceci termine la preuve. ∎

5.1.3 Alternative de Fredholm

5.1.25 PROPOSITION

Soit $\mathcal{H}$ un espace de Hilbert sur $\mathbb{K}$ et $T \in \mathcal{L}(\mathcal{H})$. Alors

(a) $\ker T = (\operatorname{Im} T^*)^{\perp}$, $\ker T^* = (\operatorname{Im} T)^{\perp}$.

(b) $\overline{\operatorname{Im} \mathrm{T}} = (\operatorname{Ker} \mathrm{T}^*)^{\perp}$, $\overline{\operatorname{Im} T^*} = (\ker T)^{\perp}$.

(c) $\ker T^* = \{0\}$ si et seulement si $\operatorname{Im} T$ est dense dans $\mathcal{H}$,

où $\ker T = \{x \in \mathcal{H} : Tx = 0\}$, *Im* $T = \{y \in \mathcal{H} : y = Tx, x \in \mathcal{H}\}$ *et* $\overline{\operatorname{Im} T}$ *est l'adhérence de* $\operatorname{Im} T$ *dans l'espace* $\mathcal{H}$.

Démonstration: (a) Soit $x \in \ker T$, alors $Tx = 0$, alors que pour tout y $\in \mathcal{H}$

$$\langle x, T^*y\rangle = \langle Tx, y\rangle = 0,$$

ce qui signifie que $x \in (\operatorname{Im} T^*)^{\perp}$ et ainsi $\ker T \subset (\operatorname{Im} T^*)^{\perp}$. Si $x \in (\operatorname{Im} T^*)^{\perp}$ alors pour tout $y \in \mathcal{H}$

$$\langle Tx, y\rangle = \langle x, T^*y\rangle = 0,$$

ce qui implique $x \in \ker T$ et donc $(\operatorname{Im} T^*)^{\perp} \subset \ker T$. Donc $\ker T = (\operatorname{Im} T^*)^{\perp}$.

Par ce qui précède et du théorème **??** nous avons

$$\ker T^* = (((\operatorname{Im} T^*)^*))^{\perp} = (\operatorname{Im} T)^{\perp}.$$

(b) D'après la partie (a), nous avons

$$\begin{array}{rcccl} \overline{\operatorname{Im} \mathrm{T}} & = & ((\operatorname{Im} T)^{\perp})^{\perp} & = & (\ker T^*)^{\perp} \\ \overline{\operatorname{Im} T^*} & = & ((\operatorname{Im} T^*)^{\perp})^{\perp} & = & (\ker T)^{\perp}. \end{array}$$

(c) Si $\ker T^* = \{0\}$, alors par partie (b)

$$\overline{\operatorname{Im} T^*} = (\ker T^*)^{\perp} = \{0\}^{\perp} = \mathcal{H}.$$

Par conséquent Im T est dense dans $\mathcal{H}$. Inversement, si Im T est dense dans $\mathcal{H}$ Alors par (a),

$$\ker T^* = (\operatorname{Im} T)^{\perp} = (((\operatorname{Im}, T)^{\perp})^{\perp})^{\perp} = (\mathcal{H})^{\perp} = \{0\}.$$

5.1.27 COROLLAIRE

Soit $\mathcal{H}$ un espace de Hilbert, et $T \in \mathcal{L}(\mathcal{H})$. Alors la décomposition orthogonale devient :

$$\mathcal{H} = \overline{\operatorname{Im} T} \oplus \ker T^* = \overline{\operatorname{Im} T^*} \oplus \ker T.$$

5.1.28 THÉORÈME (ALTERNATIVE DE FREDHOLM)
Soit $\mathcal{H}$ un espace de Hilbert sur $\mathbb{K}$ et $T \in K(\mathcal{H})$ un opérateur compact.
Alors

i) $\ker(I - T)$ est de dimension finie.

ii) $\operatorname{Im}(I - T)$ est fermé.

iii) $\ker(I - T) = 0$ si et seulement si $\operatorname{Im}(I - T) = \mathcal{H}$

Démonstration: i) Supposons que $\mathcal{M} := \ker(I - T)$ est de dimension infinie. Comme le noyau d'un opérateur linéaire borné est fermé, l'espace $\mathcal{M}$ est alors un espace de Hilbert de dimension infinie, et il existe une suite orthonormée $\{e_n\}$ dans $\mathcal{M}$. Il s'ensuit que $Te_n = e_n$ pour tout $n \in \mathbb{N}$. Comme $\{e_n\}$ est orthonormée, la suite $\{e_n\}$ n'a pas de sous-suite convergente. Ceci contredit la compacité de T. Donc $\dim \ker(I - T) < +\infty$.

ii) Soit $\{y_n\}$ une suite de $\operatorname{Im}(I - T)$ avec $y_n \to y$ lorsque $n \to \infty$. Il s'agit de montrer que $y \in \operatorname{Im}(I - T)$.
Pour tout $n \in \mathbb{N}$, il existe $x_n \in \mathcal{H}$ tel que $y_n = (I - T)x_n$. D'après ce qui précède $\ker(I - T)$ est fermé, il résulte du théorème 4.3.2 que $x_n = u_n + v_n$ avec $u_n \in \ker(I - T)$ et $v_n \in \ker(I - T)^\perp$. Nous affirmons que la suite $\{v_n\}$ est bornée. Sinon, quitte à prendre une sous-suite, on peut supposer que $\|v_n\| \neq 0$ pour tout n et que $\|v_n\| \to \infty$. Soit $w_n := v_n / \|v_n\|$, alors $w_n \in \ker(I - T)^\perp$ avec $\|w_n\| = 1$ pour tout $n \in \mathbb{N}$ et

$$(I - T)w_n = \frac{y_n}{\|v_n\|} \to 0 \quad \text{lorsque } n \to \infty$$

car $\{y_n\}$ est bornée. D'après, la compacité de T, $\{Tw_n\}$ converge (quitte à prendre une sous-suite si nécessaire). Ainsi, on obtient que $\{w_n\}$ converge vers un élément $w \in \mathcal{H}$. Il est clair que $\|w\| = 1$ et

$$(I - T)w = \lim_{n \to \infty} (I - T)w_n = 0,$$

ce qui donne $w \in \ker(\lambda I - T)$. Cependant $w_n \in \ker(\lambda I - T)^\perp$ implique

$$\|w - w_n\|^2 = \langle w - w_n, w - w_n \rangle = 1 + 1 = 2,$$

ce qui contredit $w_n \to w$ lorsque $n \to \infty$. D'où la suite $\{v_n\}$ est bornée.
Maintenant, par la compacité de T on peut supposer que $\{Tv_n\}$ est convergente (quitte à prendre une sous-suite si nécessaire). D'après $v_n = ((I - T)v_n + Tv_n) = (y_n + Tv_n)$ pour tout $n \in \mathbb{N}$, il s'ensuit que la suite $\{v_n\}$ converge vers $v \in \mathcal{H}$,.
Alors

$$y = lim_{n \to \infty} y_n = lim_{n \to \infty} (\lambda I - T)v_n = (\lambda I - T)v,$$

Cela prouve que $\operatorname{Im}(\lambda I - T)$ est fermé.

iii) Soit $V := I - T$ et supposons que $V^{-1}\{0\} = \{0\}$. On pose $E_n := V^n(\mathcal{H})$ et notez que

$$E_n = V^n(\mathcal{H}) = V^{n-1}(V(\mathcal{H})) \subset V^{n-1}(\mathcal{H}) = E_{n-1} \quad \forall\, n \geq 1.$$

D'après i) , nous savons que E_n est fermé pour tous $n \geq 1$. Supposons que $E_{n+1} \neq E_n$ pour tous $n \geq 1$. Alors il existe une suite $\{x_n\} \subset E_n \subset$ avec $\|x_n\| = 1$ tel que

$$\|Tx_n - Tx_m\|_{\mathcal{H}} > \frac{1}{2} \quad \text{pour tout } \; n \neq m,$$

en contradiction avec la compacité de T. Donc, $E_{n+1} = E_n$ pour un certain $n \geq 1$. Nous allons montrer que $\mathcal{H} = E_0 = E_1$. Supposons que ce n'est pas vrai, c'est à dire $E_0 \neq E_1$. Soit $m \geq 1$ le plus petit entier positif tel que

$$E_{m-1} \neq E_m = E_{m+1}.$$

Nous choisissons $y \in E_{m-1} \backslash E_m$, alors $V(y) \in E_m = E_{m+1}$. Par conséquent, nous pouvons trouver $z \in E_m$ tel que

$$V(y) = V(z) \quad \text{et} \quad y \neq z$$

car $y \notin E_m = E_{m+1}$. Donc $V(y - z) = 0$ et ainsi de $y - z \in \ker V$, en contradiction avec l'hypothèse que $V^{-1}(\{0\}) = \{0\}$. Donc V est surjective.
Réciproquement, si $\mathrm{Im}(I - T) = \mathcal{H}$, alors $\{0\} = \mathrm{Im}(I - T)^{\perp} = \ker(I - T^*)$ d'où $\mathrm{Im}(I - T^*) = \mathcal{H}$ alors $\mathrm{Im}(I - T^*)^{\perp} = \{0\}$ d'où $\ker(I - T) = \{0\}$. ∎

5.2 Théorème spectral pour les opérateurs auto-adjoints compacts

Rappelons que si A est une matrice carrée $n \times n$, un nombre complexe λ est une valeur propre de A si et seulement si il existe un $x \in \mathbb{R}^n$ avec $x \neq 0$ tel que $Ax = \lambda x$, ce qui signifie que $(\lambda I - A)x = 0$, c'est-à $\lambda I - A$ n'est pas inversible, où I est la matrice identité sur $\mathbb{R}^n$. Comme les valeurs propres ont de nombreuses applications en dimension finie, il est naturel d'essayer d'étendre ces notions à des espaces de dimension infinie. C'est le but de cette partie. Comme l'adjoint est un outils qui aide à déterminer quand un opérateur est inversible dans les espaces de Hilbert, nous limiterons notre étude aux espaces de Hilbert, bien que les définitions que nous allons présenter peuvent facilement être étendues à des espaces de Banach.

5.2.1 Le spectre d'un opérateur

5.2.1 DÉFINITION
Soit E est un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et T un endomorphisme de E.
On appelle

- Spectre de T est l'ensemble

$$\sigma_p(T) := \{\lambda \in \mathbb{K} : \lambda I - T \text{ ne soit pas inversible}\}$$

- Spectre ponctuel de T est l'ensemble

$$\sigma_p(T) := \{\lambda \in \sigma(T) : \ker(\lambda I - T) \neq \{0\}\}$$

i.e. l'ensemble des valeurs propres de T. On appelle multiplicité de la valeur propre λ, la dimension du sous-espace propre $\ker(\lambda I - T)$.

5.2.2 REMARQUE

a) On a toujours $\sigma_p(T) \subset \sigma(T)$.

b) Les vecteurs propres correspondant à des valeurs propres distinctes sont linéairement indépendants. Explicitement, soit E un espace vectoriel et T un opérateur linéaire sur E. Soit $\lambda_1, \lambda_2, \cdots, \lambda_k$ des valeurs propres distinctes de T et pour chaque $1 \leq j \leq k$, x_j un vecteur propre correspondant à λ_j. Alors $x_1, x_2, \cdots, x_k$ sont linéairement indépendants.

Démonstration: Supposons qu'ils soient linéairement dépendants et soit une combinaison

$$\alpha_1 x_1 + \sum_{j=2}^{k} \alpha_j x_j = 0$$

avec $\alpha_1 \neq 0$. Soit P un polynôme tel que $P(\lambda_1) = 1$ et $P(\lambda_j) = 0$ pour $j \geq 2$. On remarque que $P(T)x_j = P(\lambda_j)x_j$, $P(\lambda)$ est une valeur propre de l' opérateur $P(T)$. Par application de $P(T)$, nous obtenons

$$0 = \alpha_1 P(T)x_1 + \sum_{j=2}^{k} \alpha_j P(T)x_j = x_1\alpha_1.$$

Ainsi, $\alpha_1 = 0$ car $x_1 \neq 0$, ce qui est une contradiction. Nous répétons le même processus pour le reste des α_i. ■

5.2.4 EXEMPLE. *Soit I est l'opérateur identité de l'espace de Hilbert $\mathcal{H}$. Il est clair que $\sigma(I) = \{1\}$, car $\lambda I - I = (\lambda - 1)I$ est inversible si $\lambda - 1 \neq 0$. De même, si $\mu \in \mathbb{K}$, alors $\sigma(\mu I) = \{\mu\}$.*

5.2.5 EXEMPLE. *Soit $\mathcal{V}$ est un espace vectoriel normé de dimension finie. Alors, chaque opérateur linéaire T sur $\mathcal{V}$ peut être représenté par une matrice carrée A, et que l'opérateur linéaire $(\lambda I - T)$ est inversible précisément lorsque la matrice $(\lambda I - A)$ est inversible. Il est clair que pour chaque $\lambda \in \mathbb{C}$ on a l'un des cas suivants :*

(1) *λ est une valeur propre de A ;*

(2) *$\lambda I - A$ est une matrice inversible, c'est à dire la matrice $(\lambda I - A)^{-1}$ existe.*

Il en résulte que le spectre $\sigma(T) = \sigma_p(T)$ est précisément l'ensemble des valeurs propres de la matrice A, et la dimension de l'espace des vecteurs propres correspondant à chaque valeur propre est finie.

Nous voyons dans l'exemple ci-dessus que le spectre d'un opérateur linéaire dans d'un espace de dimension finie a une structure simple, mais pour un opérateur général dans un espace de dimension infinie le spectre peut être très différent et plus complexe.

5.2.6 EXEMPLE (OPÉRATEUR DIAGONAL SUR $\ell^2(\mathbb{N},\mathbb{C})$). Soit $(\lambda_n)_{n\in\mathbb{N}}$ une suite dans $\mathbb{C}\setminus\{0\}$ telle que $\lim\limits_{n\to+\infty}\lambda_n = 0$. On définit l'opérateur T sur ℓ^2 par

$$T((x_n)_{n\in\mathbb{N}}) = (\lambda_n x_n)_{n\in\mathbb{N}}.$$

Comme $(T-\lambda I)x = ((\lambda_k-\lambda)x_k)_{k\in\mathbb{N}}$, nous aurons $(T-\lambda I)^{-1}y = \big(\frac{y_k}{\lambda_k-\lambda}\big)_{k\in\mathbb{N}}$.

Il en résulte que $(T-\lambda I)^{-1}$ est un opérateur borné si et seulement si λ n'est pas dans l'adhérence de $\{\lambda_k\}_{k\in\mathbb{N}}$, qui n'est autre que $\{\lambda_k\}_{k\in\mathbb{N}}\cup\{0\}$.

Tous les λ_k sont clairement des valeurs propres de T comme $Te_k = \lambda_k e_k$ pour e_k élément de la base canonique de ℓ^2. Mais 0 n'est pas valeur propre car T est injective (comme tous les $\lambda_k \neq 0$). Notre conclusion est la suivante :

$$\sigma(T) = \{\lambda_k\}_{k\in\mathbb{N}}\cup\{0\}, \qquad \sigma_p(T) = \{\lambda_k\}_{k\in\mathbb{N}}.$$

5.2.7 EXEMPLE (OPÉRATEUR DE MULTIPLICATION SUR $L^2[0,1]$). Considérons l'opérateur de multiplication $T : L^2[0,1]\to L^2[0,1]$ définie par $(Tf)(t) = tf(t)$.

Comme $(T-\lambda I)f(t) = (t-\lambda)f(t)$, nous avons

$$(T-\lambda I)^{-1}y(t) = \frac{1}{t-\lambda}y(t). \tag{5.2.1}$$

Si $\lambda \notin [0,1]$ alors la fonction $t\mapsto \frac{1}{t-\lambda}$ est bornée, donc $(T-\lambda I)^{-1}$ est un opérateur borné.

Inversement, si $\lambda\in[0,1]$, alors $\frac{1}{t-\lambda}\notin L^2[0,1]$ en raison de la singularité non-intégrable en $t=\lambda$. D'où $T-\lambda I$ n'est pas inversible (prendre $y(t)\equiv 1$). Par conséquent, $\sigma(T) = [0,1]$.

Supposons que λ soit une valeur propre de T avec f un vecteur propre dans $L^2[0,1]$. Cela signifie que l'identité suivante est vérifiée

$$(t-\lambda)f(t) = 0 \quad \text{pour tout } t\in[0,1].$$

Il en résulte que $f = 0$ dans $L^2[0,1]$. Par conséquent, T *n'a pas de valeurs propres*. Notre conclusion est la suivante :

$$\sigma(T) = [0,1], \qquad \sigma_p(T) = \emptyset.$$

5.2.8 EXEMPLE (OPÉRATEUR SHIFT). Considérons les opérateurs de décalage à droite R et à gauche L sur ℓ^2, agissent sur un vecteur $x = (x_1, x_2, \dots)$ par

$$R(x) = (0, x_1, x_2, \dots), \quad L(x) = (x_2, x_3, \dots,).$$

comme R est clairement injectif et comme $\operatorname{Im} R$ n'est pas dense dans ℓ^2, 0 est dans le spectre de R.

Montrer que

$$\sigma(R) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}, \qquad \sigma_p(R) = \emptyset,$$
$$\sigma(L) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}, \qquad \sigma_p(L) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}.$$

Le spectre est un compact

5.2.9 LEMME (VON NEUMANN)
Soit X un espace de Banach.
Considérons un opérateur $U \in \mathcal{L}(X)$ telle que $\|U\| < 1$. Alors $I - U$ est inversible et son inverse peut s'exprimer comme une série convergente dans $L(X)$:

$$(I-U)^{-1} = \sum_{k=0}^{\infty} U^k, \quad \|(I-U)^{-1}\| \leq \frac{1}{1 - \|U\|}.$$

Démonstration: La série $\sum_{k=0}^{\infty} U^k$ est absolument converge car $\|U^k\| \leq \|U\|^k$, et $\|U\| < 1$, donc convergente car $\mathcal{L}(X)$ est un Banach. De plus,

$$(I-U)\sum_{k=0}^{\infty} U^k = \sum_{k=0}^{\infty} U^k(I-U) = \sum_{k=0}^{\infty}(U^k - U^{k+1}) = I$$

comme série télescopique. Enfin,

$$\|(I-U)^{-1}\| \leq \sum_{k=0}^{\infty} \|U\|^k \leq \frac{1}{1-\|U\|}.$$

Ceci termine la preuve. ∎

5.2.11 PROPOSITION
Soit $\mathcal{H}$ un espace de Hilbert et $T \in \mathcal{L}(\mathcal{H})$.
Le spectre $\sigma(T)$ de T est un sous-ensemble compact de $\mathbb{K}$ contenu dans
$\{\lambda \in \mathbb{C} : \ |\lambda| \leq \|T\|\}$.

Démonstration: Soit $\lambda \in \mathbb{K}$ tel que $|\lambda| > \|T\|$.

On a $T - \lambda I = \lambda(\lambda^{-1}T - I)$. Comme $\|\lambda^{-1}T\| < 1$, le lemme de von Neumann 5.2.9 implique que l'opérateur alors $T - \lambda I$ est inversible, donc $\lambda \notin \sigma(T)$ i.e. $\sigma(T) \subseteq \{\lambda \in \mathbb{C} : \ |\lambda| \leq \|T\|\}$.

Il reste à montrer que $\sigma(T)$ est fermé. Soit $\lambda_0 \notin \sigma(T)$, alors $S = T - \lambda_0 I$ est inversible.

D'où $T - \lambda I = S - (\lambda_0 - \lambda)I = S(I - (\lambda_0 - \lambda)S^{-1})$, d'après le lemme 5.2.9, $T - \lambda I$ est inversible si $\|(\lambda_0 - \lambda)S^{-1}\| = |\lambda - \lambda_0|.\|S^{-1}\| < 1$, donc si $|\lambda - \lambda_0| < \frac{1}{\|S^{-1}\|}$. Ainsi le disque $\{\lambda \in \mathbb{K}, |\lambda - \lambda_0| < \frac{1}{\|S^{-1}\|}\}$ est contenu in $\mathbb{K} \setminus \sigma(T)$. Ce qui montre que $\mathbb{K} \setminus \sigma(T)$ est ouvert et donc $\sigma(T)$ est fermé.

5.2.13 Théorème (Propriétés du spectre d'un opérateur compact)

Soit $\mathcal{H}$ est un espace de Hilbert sur $\mathbb{K}$ et T un opérateur compact sur $\mathcal{H}$, alors :

(i) Si $\mathcal{H}$ est de dimension infinie, $0 \in \sigma(T)$.

(ii) $\sigma(T) = \sigma_p(T) \cup \{0\}$ et toute valeur propre est de multiplicité finie.

(iii) Soit $\delta > 0$, alors l'ensemble des valeurs propres deux à deux disjointes et de module $\geq \delta$ est fini. Par conséquent, si $\sigma(T)$ contient une suite d'éléménts deux à deux disjoints, alors cette suite converge vers 0.

Démonstration: (i) Si $0 \notin \sigma(T)$ alors T est un isomorphisme, contredit T compact d'après 5.1.6.

(ii) Soit $\lambda \neq 0$, si $\lambda \notin \sigma_p(T)$, alors $T - \lambda I = -\lambda(I - \lambda^{-1}T)$ est injectif et d'après 5.1.28 $T - \lambda I$ est sujectif, donc $\lambda \notin \sigma(T)$.

(iii) Soit $\delta > 0$. Supposons qu'il existe une suite Soit $(\lambda_n)_{n\in\mathbb{N}} \subset \sigma_p(T) \setminus \{0\}$, formée d'éléménts deux à deux disjoints tels que $|\lambda_n| \geq \delta$.
Pour tout $n \in \mathbb{N}$, soit x_n un vecteurs propres associé à la valeur propre λ_n i.e. $T(x_n) = \lambda_n x_n$.
On pose $H_n = \mathrm{Vect}\{x_0, \dots, x_n\}$. Alors $T(H_n) \subset H_n$ et $(T - \lambda_n I)H_n \subset H_{n-1}$.
Par le prcédé d'ortonormalisation, on construit une suite $(e_n)_{n\in\mathbb{N}}$ telle que $\|e_n\| = 1$ et $e_n \perp H_{n-1}$ pour tout $n \in \mathbb{N}^*$.
Alors, pour $n > m$, on a $Te_n - Te_m = \lambda_n e_n + \underbrace{(Te_n - \lambda_n e_n) - Te_m}_{\in H_{n-1}}$, d'où

$$\|Te_n - Te_m\| \geq \inf_{z\in H_{n-1}} \|\lambda_n e_n - z\| = \|\lambda_n e_n\| = |\lambda_n| \geq \delta > 0.$$

Ceci entraîne que $(Te_n)_{n\in\mathbb{N}}$ n'a pas de valeur d'adhérence, contredit ainsi la compacite de T. ∎

5.2.2 Opérateurs auto-adjoints

Définition et exemples

Soit T est un opérateur linéaire borné sur un espace de Hilbert, soit $T \in \mathcal{L}(\mathcal{H})$. Rappelons 4.4.1, que l'opérateur adjoint $T^* \in \mathcal{L}(\mathcal{H})$ est définie par $\langle T^*x, y\rangle = \langle x, Ty\rangle$ pour $x, y \in \mathcal{H}$.

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

$$\langle Tx, y\rangle = \langle x, Ty\rangle, \quad x, y \in \mathcal{H}.$$

5.2.16 EXEMPLE. Des exemples de opérateurs auto-adjoints suivants :

(i) L'opérateur identité sur un espace de Hilbert est auto-adjoint.

(ii) opérateurs linéaires sur $\mathbb{C}^n$ donné par des matrices hermitiennes (a_{ij}), c'est à dire telles que $a_{ij} = \overline{a_{ji}}$;

(iii) opérateur intégral $(Tf)(t) = \int_0^1 k(s,t)f(s)\,ds$ sur $L^2[0,1]$ avec un noyau hermitien, c'est à dire tel que $k(s,t) = \overline{k(t,s)}$;

(iv) Les projections orthogonales P sur $\mathcal{H}$. (Pourquoi ?)

Chaque opérateur linéaire borné peut être décomposée en deux opérateurs auto-adjoints :

5.2.17 LEMME
Chaque opérateur $A \in \mathcal{L}(\mathcal{H})$ peut être représenté de manière unique comme

$$A = T + iS$$

où $T, S \in \mathcal{L}(\mathcal{H})$ sont opérateurs auto-adjoints.

Démonstration: Si $A = T + iS$, alors $A^* = T - iS$. La résolution de ces deux équations, nous voyons que le lemme avec $T = \frac{A+A^*}{2}$ et $S = \frac{A-A^*}{2i}$. ∎

5.2.19 EXEMPLE. Montrer que l'ensemble des opérateurs auto-adjoints forme un sous-espace linéaire fermé de $\mathcal{L}(\mathcal{H})$.

5.2.20 DÉFINITION (SOUS-ESPACE INVARIANT)
Un sous-espace E de $\mathcal{H}$ est un *sous-espace invariant* par T si $T(E) \subseteq E$.

5.2.21 EXEMPLE. Chaque sous-espace propre de T est invariant. Plus généralement, l'espace engendré par n'importe quel sous-ensemble de vecteurs propres de T est un sous-espace invariant.

5.2.22 PROPOSITION
Soit $T \in \mathcal{L}(\mathcal{H})$ est auto-adjoint. Si $E \subseteq \mathcal{H}$ est un sous-espace invariant par T alors $E^{\perp}$ est aussi un sous-espace invariant par T.

Démonstration: Soit $x \in E^{\perp}$; nous allons vérifier que $Tx \in E^{\perp}$. Soit $y \in E$ arbitrairement. Alors $\langle Tx, y\rangle = \langle x, Ty\rangle = 0$ car $x \in E^{\perp}$ et $y \in E$, donc $Ty \in E$. ∎

5.2.24 REMARQUE
Si $\mathcal{H}$ est un espace de Hilbert complexe et $T \in \mathcal{L}(\mathcal{H})$ est un opérateur auto-adjoint, Alors

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

Ainsi $\langle Tx, x\rangle$ est un nombre réel.

5.2.25 LEMME
Si $\mathcal{H}$ est un espace de Hilbert et $T \in \mathcal{L}(\mathcal{H})$ est un opérateur auto-adjoint. Alors toutes les valeurs propres de T sont réelles (c.-à-d $\sigma_p(T) \subset \mathbb{R}$) et les vecteurs propres correspondant à des valeurs propres différentes sont orthogonaux.

Démonstration: Soit λ une valeur propre de T avec un vecteur propre x ($\neq 0$). Alors

$$\langle Tx, x\rangle = \langle \lambda x, x\rangle,$$

donc, à partir Remarque 5.2.24, nous avons

$$\lambda = \frac{\langle Tx, x\rangle}{\langle x, x\rangle} = \frac{\langle Tx, x\rangle}{\|x\|^2} \in \mathbb{R}.$$

Aussi, si μ est une autre valeur propre de T avec un vecteur propre y, alors nous avons

$$\lambda\langle x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \overline{\mu}\langle x, y\rangle = \mu\langle x, y\rangle$$

car μ est réel. Il en résulte que $(\lambda - \mu)\langle x, y\rangle = 0$. Parce $\lambda \neq \mu$, nous concluons que $\langle x, y\rangle = 0$, ce qui signifie que $x \perp$ y. ∎

5.2.27 LEMME (NORME D'UN OPÉRATEUR AUTO-ADJOINT)
Soit $\mathcal{H}$ et T comme dans *lemme* 5.2.25 . Alors

$$\|T\| = \sup_{x \in \mathcal{H}\ \|x\|=1} |\langle Tx, x\rangle|.$$

Démonstration: Soit $\alpha := \sup\{|\langle Tx, x\rangle| : x \in \mathcal{H},\ \|x\| \leq 1\}$. Il suffit de prouver que

$$|\langle Tx, y\rangle| \leq \alpha\|x\|\|y\|$$

pour tous x et y. On peut évidemment supposer que x et y sont non nul. De plus, nous pouvons multiplier y par un nombre complexe β de module un ($|\beta| = 1$) pour obtenir $\mathfrak{Re}\langle Tx, \beta y\rangle = |\langle Tx, y\rangle|$, on peut donc supposer que $\langle Tx, y\rangle \geq 0$. Pour chaque $x, y \in \mathcal{H}$, comme $T^* = T$, nous avons

$$\begin{aligned}\langle T(x+y), x+y\rangle &= \langle Tx, x\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Ty, y\rangle \\ &= \langle Tx, x\rangle + \langle Tx, y\rangle + \overline{\langle Tx, y\rangle} + \langle Ty, y\rangle \\ &= \langle Tx, x\rangle + 2\,\mathfrak{Re}\langle Tx, y\rangle + \langle Ty, y\rangle,\end{aligned}$$

De même,

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

Alors

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

si

$$|\langle Tx, y\rangle| \leq \frac{\alpha}{4}(\|x+y\|^2 + \|x-y\|^2) = \frac{\alpha}{2}(\|x\|^2 + \|y\|^2),$$

Maintenant, on applique cette inégalité à $\sqrt{\|y\|/\|x\|}\,x$ à la place de x et $\sqrt{\|x\|/\|y\|}\,y$ à la place de y.

5.2.29 Définition

Un opérateur auto-adjoint $T \in \mathcal{L}(\mathcal{H})$ est dit positif si $\langle Tx, x\rangle \geq 0$, pour tout $x \in \mathcal{H}$.

5.2.30 Lemme

Soit $T \in \mathcal{L}(\mathcal{H})$ un opérateur auto-adjoint positif alors :

$$\|Tx\|^2 \leq \|T\|\langle Tx, x\rangle$$

Démonstration: Si $Ax = 0$, il n'y a rien à prouver. Supposons donc $Ax \neq 0$. D'après l'inégalité de Cauchy-Schwarz appliquée à la forme hermitienne positive $(u, v) \mapsto \langle Tu, v\rangle$ on a

$$|\langle Tu, v\rangle|^2 \leq \langle Tu, u\rangle.\langle Tv, v\rangle$$

En particulier, pour $u = x$ et $v = \frac{Tx}{\|Tx\|}$ on obtient l'inégalité

$$\|Tx\|^2 \leq \langle Tx, x\rangle.\langle Tv, v\rangle \leq \|T\|\langle Tx, x\rangle.$$

5.2.32 LEMME
Soit $\mathcal{H}$ est un espace de Hilbert et $T \in K(\mathcal{H})$ un opérateur auto-adjoint et compact. Soit $M := \sup_{x\in\mathcal{H},\,\|x\|=1}\langle Tx, x\rangle$. Si M est non nul, l'opérateur T admet M comme valeur propre.

Démonstration: On suppose $M \neq 0$

Soit $(x_n) \in \mathcal{H}$ telle que $\|x_n\| = 1$ et $\lim_{n\to+\infty}\langle Tx_n, x_n\rangle = M$. Comme T est compact, quitte à prendre une sous-suite, on peut supposer que (Tx_n) converge vers $y \in \mathcal{H}$. Alors l'opérateur $MI - T$, est auto-adjoint positif et d'après le lemme précédent

$$\|(MI-T)x_n\|^2 = \|MI-T)\|.\langle (MI-T)x_n, (MI-T)x_n\rangle \leq M^2 + M^2 - 2M\,\langle Tx_n, x_n\rangle,$$

Il en résulte que $\lim\limits_{n\to+\infty} \|(M\,I - T)x_n\| = 0$. Puisque (Tx_n) converge vers y, on déduit que la suite (x_n) converge vers $z = \frac{y}{M}$. D'où $\|z\| = 1$ et $Tz = Mz$, ainsi $M \in \sigma(T)$. ■

5.2.34 COROLLAIRE
Soit T un opérateur auto-adjoint compact sur un espace de Hilbert $\mathcal{H}$. Alors T admet une valeur propre λ tel que $|\lambda| = \|T\|$.

Démonstration: Si $T = 0$, tous les vecteurs non nuls sont associés à la valeur propre $0 = \|T\|$.

Supposons $T \neq 0$. D'après le lemme précédent

$$\sup_{x\in\mathcal{H},\,\|x\|=1} \langle T^2x, x\rangle = \sup_{x\in\mathcal{H},\,\|x\|=1} \langle Tx, Tx\rangle = \sup_{x\in\mathcal{H},\,\|x\|=1} \|Tx\|^2 = \|T\|^2$$

est valeur propre de T^2.

Ainsi, il existe un vecteur $v \in \mathcal{H} - \{0\}$ tel que :

$$0 = T^2v - \|T\|^2v = (T + \|T\|) \circ (T - \|T\|)v.$$

Si $(T - \|T\|)v = 0$, alors $\|T\|$ est valeur propre de T, sinon $v_1 = (T - \|T\|)v \neq 0$ et $(T + \|T\|)v_1 = 0$, d'où $-\|T\|$ est valeur propre de T. ■

Diagonalisation des opérateurs auto-adjoints compacts

5.2.36 THÉORÈME (DIAGONALISATION DES OPÉRATEURS AUTO-ADJOINTS COMPACTS)
Soit $\mathcal{H}$ est un espace de Hilbert sur $\mathbb{K}$ et $T \in K(\mathcal{H})$ un opérateur auto-adjoint compact. Alors, il existe une base hilbertienne de $\mathcal{H}$ formée de vecteurs propres de T.

Démonstration: On désigne par $\mathcal{B}$ l'ensemble des parties U de $\mathcal{H}$ qui vérifient les conditions $(*)$

$$\begin{cases} x \in U \Rightarrow \|x\| = 1 \\ x, y \in U \text{ et } x \neq y \Rightarrow \langle x, y \rangle = 0 \\ x \in U \Rightarrow Tx \in \mathbb{K}x = \mathrm{Vect}\{x\} \end{cases} \qquad (*)$$

ordonné par l'inclusion $(\subset)$ des parties de $\mathcal{H}$.

Montrons que $(\mathcal{B}, \subset)$ est inductif.

Soit $\mathcal{C} = \{B_i, i \in I\}$ chaîne de $\mathcal{B}$ i.e. une partie totalement ordonnée de $(\mathcal{B}, \subset)$ Alors pour montrer que $\bigcup_{i \in I} B_i$ est un majorant de C, il suffit de montrer que $\bigcup_{i \in I} B_i$ vérifie $(*)$.

En effet, soit $x, y \in \bigcup_{i \in I} B_i$, $x \neq y$, alors il existe $j \in I$ tel que $x, y \in B_j$, d'où $\|x\| = \|y\| = 1$, $Tx \in \mathbb{K}x$, $Ty \in \mathbb{K}y$ et $\langle x, y \rangle = 0$. Ainsi on a montrer que $\bigcup_{i \in I} B_i$ vérifie $(*)$.

D'après le lemme de Zorn, il existe un élément maximal B dans $\mathcal{B}$. Alors B est une base hilbertienne de $\mathcal{H}$ formée de vecteurs propres : En effet, les deux premières conditions de $(*)$ montrent que les éléments de B forment un système orthonormé et que la troixième condition qu'ils sont des vecteurs propres de T.

Il suffit de montrer que B est total i.e $\overline{\mathrm{Vect}(B)} = \mathcal{H}$. Sinon, $H_0 = \overline{\mathrm{Vect}(B)}^{\perp} \neq \emptyset$. On note $T_0 = T|_{H_0}$.

D'après le lemme 5.2.22, H_0 est stable par T, d'où $T_0 : H_0 \to H_0$ définit un opérateur auto-adjoint compact.

D'après le corollaire 5.2.34, on peut trouver $v_0 \neq 0$ qui soit vecteur propre de T_0 associé à une valeur propre λ telle que $|\lambda| = \|T_0\|$.

Mais alors $B \cup \{v_0\}$ est un sytème vérifiant $(*)$ et qui contient strictement B, ceci contredit le caractère maximal de B. Donc B est total et par suite une base hilbertienne de $\mathcal{H}$.

Maintenant, nous pouvons enoncer ce qu'on appelle **le théorème spectrale** pour opérateurs auto-adjoints compacts sur un espace de Hilbert séparable.

5.2.38 Théorème

Soient $\mathcal{H}$ est un espace de Hilbert de dimension infinie séparables et $T \in \mathcal{L}(\mathcal{H})$ un opérateur auto-adjoint compact. Alors il existe une base orthonormée $\{e_n\}_{n \in \mathbb{N}}$ de $\mathcal{H}$ formée de vecteurs propres de T, de sorte que

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n \quad \text{pour tout } x \in \mathcal{H},$$

où λ_n est la valeur propre associée à e_n pour chaque $n \in \mathbb{N}$.

Démonstration: D'après le lemme 5.2.36, nous savons qu'il existe une base hilbertienne de $\mathcal{H}$ formée de vecteurs propres de T. Cette base est dénombrable car $\mathcal{H}$ est séparable. On la note $\{e_n\}_{n\in\mathbb{N}}$. Pour chaque $x \in \mathcal{H}$ et $m > k \geq 1$, on a

$$\left\|\sum_{n=k}^{m} \lambda_n \langle x, e_n \rangle e_n\right\|^2 = \sum_{n=k}^{m} |\lambda_n \langle x, e_n \rangle|^2 \leq \|T\| \sum_{n=k}^{m} |\langle x, e_n \rangle|^2 \to 0 \quad \text{que } \ k, m \to \infty.$$

Donc $\sum_{n=k}^{m} \lambda_n \langle x, e_n \rangle e_n$ est convergente dans $\mathcal{H}$.

De plus, si $x \in \mathcal{H}$ avec $\|x\| \leq 1$ alors pour tout $m \geq 1$, nous avons

$$\left\|\sum_{n=1}^{m} \lambda_n \langle x, e_n \rangle e_n\right\|^2 \leq \|T\| \sum_{n=1}^{m} |\langle x, e_n \rangle|^2 = \|T\|^2 \|x\|^2. \tag{5.2.2}$$

Par conséquent, si nous définissons

$$Lx = \sum_{n=1}^{m} \lambda_n \langle x, e_n \rangle e_n,$$

à partir de (5.2.2), nous constatons que $L \in \mathcal{L}(\mathcal{H})$. Notez que $L(e_n) = T(e_n)$ pour tout $n \in \mathbb{N}$. Ainsi, par linéarité et continuité, on aura $T = L$. ∎

5.2.3 Exercices supplémentaires

5.2.40 Exercice

1. Soit $\mathcal{H}$ un espace de Hilbert et $T \in \mathcal{L}(\mathcal{H})$. Montrer que
$$\sigma(T + \mu I) = \sigma(T) - \mu \quad \text{pour tout } \mu \in \mathbb{R}.$$
2. Soit $\mathcal{H}$ un espace de Hilbert et $y, z \in \mathcal{H}$. Montrer que $T \in \mathcal{L}(\mathcal{H})$ définie par $Tx = \langle x, y \rangle z$ est compact.
3. On définit les opérateurs S et $T \in \mathcal{L}(\ell^2)$ par
$$Sx = (0, \frac{x_1}{1}, \frac{x_2}{2}, \frac{x_3}{3}, \cdots), \quad Tx = (\frac{x_1}{1}, \frac{x_2}{2}, \frac{x_3}{3}, \cdots).$$
Montrer que ces opérateurs sont compacts, $\sigma(S) = \{0\}$, $\sigma_p(S) = \emptyset$ et $\sigma(T) = \sigma_p(T) = \{0\}$. Montrer également que $\mathfrak{Im}S$ n'est pas dense dans ℓ^2, mais que $\mathfrak{Im}T$ est dense.
4. Soit $\mathcal{H}$ un espace de Hilbert et Soit $T \in \mathcal{L}(\mathcal{H})$. Montrer que $\langle Tx, x \rangle \in \mathbb{R}$ pour l'ensemble $X \in \mathcal{H}$ si et seulement si T est auto-adjoint

5. Soit $\mathcal{H}$ un Hilbert l'espace sur $\mathbb{K}$ et $S \in \mathcal{L}(\mathcal{H})$ auto-adjoint. Montrer que S^n est auto-adjoint pour chaque $n \in \mathbb{N}$ et conclure que $\|S^n\| = \|S\|^n$.
6. Soit $T \in \mathcal{L}(\mathcal{H})$ un opérateur auto-adjoint sur un espace de Hilbert complexe $\mathcal{H}$. Montrer que si $\sigma(T)$ contient exactement un point λ alors $T = \lambda I$.
7. Montrer qu'une suite orthonormale $\{e_n\}$ dans un espace de Hilbert $\mathcal{H}$ n'a pas de sous-suite convergente.
8. Soit $\mathcal{H}$ un espace de Hilbert de dimension infinie avec une base orthonormée $\{e_n\}$ et $T \in \mathcal{L}(\mathcal{H})$. Montrer que si T est compact alors $\lim_{n\to\infty} \|Te_n\| = 0$.
9. Soit $\mathcal{H}$ un espace de Hilbert de dimension infinie et soit $\{e_n\}$ et $\{\widehat{e}_n\}$ deux suites orthonormales de $\mathcal{H}$. Soit $\{\alpha_n\}$ est suite dans $\mathbb{C}$, on définit un opérateur linéaire $T : \mathcal{H} \to \mathcal{H}$ par $Tx = \sum_{n=1}^{\infty} \alpha_n \langle x, e_n\rangle \widehat{e}_n$. Montrer que
   a) T est continu si et seulement si la suite $\{\alpha_n\}$ est bornée.
   b) T est compact si et seulement si $\lim_{n\to\infty} \alpha_n = 0$
   c) $y \in \operatorname{Im} T$ si et seulement si $y = \sum_{n=1}^{\infty} \alpha_n \xi_n \widehat{e}_n$ pour certains $(\xi_n)_{n\in\mathbb{N}} \in \ell^2$, et en déduire que si une infinité de α_n sont non nuls et $\lim_{n\to\infty} \alpha_n = 0$, alors $\operatorname{Im} T$ n'est pas fermé.
10. Soit $\mathcal{H}$ est un Hilbert espace, $T \in \mathcal{L}(\mathcal{H})$ un opérateur auto-adjoint compact sur $\mathcal{H}$ et $\mathcal{N}$ est un sous-espace vectoriel fermé de $\mathcal{H}$, qui est invariant par T (ie $T(\mathcal{N}) \subset \mathcal{N}$). Soit $T|_{\mathcal{N}}$ désigne la restriction de T à $\mathcal{N}$. Montrer que $T|_{\mathcal{N}}$ est un opérateur auto-adjoint compact sur l' espace de Hilbert $\mathcal{N}$.

5.3 Annexe

5.3.1 Ensembles compacts dans les espaces normés

La compacité est un substitut utile à la dimension dimension finie. Nous supposons que le lecteur est familier avec la notion de compacité d'un cours de base en topologie.

5.3.2 Rappel sur la compacité

Par définition, un sous-ensemble A d'un espace topologique (X, τ) est compact si de tout recouvrement ouvert de A on peut extraire un sous-recouvrement fini. Plus précisément, si $A \subseteq \bigcup_{i\in I} U_i$ pour une collection d'ensembles ouverts $U_i, i \in I$, alors $A \subseteq \bigcup_{k=1}^{n} U_{i_k}$ pour certains sous-famille finie $\{i_1, \ldots, i_n\} \subset I$. (propriété de Borel-Lebesgue).
Quelques propriétés fondamentales des ensembles compacts :

(i) un sous-ensemble compact d'un espace séparé est fermé ;

(ii) un sous-ensemble fermé d'un compact est compact ;

(iii) l' images d'un compact par une application continue est un compact ;

(iv) Les fonctions continues sur un compact sont uniformément continues et elles atteignent leur maximum et minimum.

Dans un espace métrique (X, d), une autre description, pratique, des ensembles compacts A est donnée en termes de ε -réseaux.

Soit $\varepsilon > 0$. Un sous-ensemble $\mathcal{N}_\varepsilon$ est un ε-*réseau* de A, si pour tout $x \in A$, il existe $y \in \mathcal{N}_\varepsilon$ tels que $d(x, y) \leq \varepsilon$. De façon équivalente, $\mathcal{N}_\varepsilon$ est un ε-réseau de A si A peut être recouvert par des boules de rayon ε centrée en des points de $\mathcal{N}_\varepsilon$. On note par $N(A, \varepsilon)$ le cardinal minimum pour un ε-réseau de (A, d).

5.3.1 Définition

Soit un espace métrique (X, d).

Une partie A est **précompacte** si pour tout $\varepsilon > 0$, il existe un ε-réseau fini de A i.e. il existe un sous-ensemble fini $\mathcal{N}_\varepsilon$ tel que $A \subset \bigcup_{x \in \mathcal{N}_\varepsilon} B(x, \varepsilon)$.

5.3.2 Théorème

Pour un sous-ensemble A d'un espace métrique X ,les conditions suivantes sont équivalentes :

(i) A est précompact ;

(ii) Si toute suite (x_n) dans A admet une sous-suite de Cauchy.

(iii) Pour tout $\varepsilon > 0$, il existe un ε-réseau de cardinal fini de A.

5.3.3 Corollaire (Compacité dans les espaces métriques)

Pour un sous-ensemble A d'un espace métrique X ,les conditions suivantes sont équivalentes :

1. A est compact
2. Si toute suite de E admet une valeur d'adhérence
3. précompact et complet

En conséquence, les ensembles précompact dans les espaces métriques sont bornés.

5.3.4 Théorème (Heine-Borel)

Un sous-ensemble A d'un espace vectoriel normé de dimension finie X est précompact si et seulement si A est bornée.

5.3.5 Exercice (Dimension métrique) Soit (A,d) un espace métrique compact.
On appelle dimension métrique le nombre

$$\dim(A) = \overline{\lim}_{\varepsilon\to 0^+} \frac{\ln(N(X,\varepsilon))}{\ln(1/\varepsilon)} \in [0,+\infty]$$

(i) Soit B_X la boule unité fermée d'un espace vectoriel normé $(X, \|.\|)$, de dimension finie n.
Monter que $(\frac{1}{\varepsilon})^n \leq N(B_X,\varepsilon) \leq (1+\frac{2}{\varepsilon})^n$. En déduire que $\dim(B_X) = n$.

(ii) Si $C = \left\{ x \in [0,1];\ \sum_{n=1}^{+\infty} \frac{a_n}{3^n},\ a_n \in \{0,2\} \right\}$ est l'ensemble triadique de Cantor, alors $\dim C = \frac{\ln 2}{\ln 3}$.

5.3.3 Compacité en dimension infinie

En dimension infinie espaces normés, théorème de Heine-Borel échoue. Par exemple, une base orthonormée (e_k) de ℓ^2 est un ensemble borné, mais il n'est pas précompact, car il n'a pas de sous-suite convergente (Comme $\|e_k - e_j\| = \sqrt{2}$ pour $k \neq j$).

Ensembles compacts sont presque finies dimensionnelles. Cette heuristique, qui est fait de rigueur dans le résultat suivant, sous-tend de nombreux arguments dans l'analyse :

5.3.6 Lemme (Approximation par des sous-espaces de dimensions finies)
Un sous-ensemble A d'un espace vectoriel normé X est précompact si et seulement si A est bornée et, pour chaque $\varepsilon > 0$, il existe un sous-espace de dimension finie Y de X, ce qui soit un ε-réseau de A.

Démonstration: Nécessité. Supposons A précompact. Choisissez un ε-réseau fini $\mathcal{N}_\varepsilon$ de A, Alors le sous-espace $Y := \mathrm{Vect}(\mathcal{N}_\varepsilon)$ est de dimension finie et forme un ε-réseau de A.

Suffisance. Comme A est bornée, $A \subseteq rB_X$ pour un certain $r > 0$. Comme Y est un ε-réseau de A, il s'ensuit que $(r+\varepsilon)B_Y$ est aussi un ε-réseau de A. De plus, comme Y est de dimension finie, l'ensemble $(r+\varepsilon)B_Y$ est précompact par le théorème de Heine-Borel. Donc, nous avons trouvé un ε-réseau fini de A. Par conséquent A est lui-même précompact. ■

Par le théorème de Heine-Borel, la boule unité fermée B_X, d'un espace vectoriel normé de dimension finie X, est compact. Ceci n'est pas vrai en dimension infinie :

5.3.8 THÉORÈME (F. RIESZ)
La boule unité fermée B_X d'un espace vectoriel normé de dimension infinie X n'est jamais compacte.

Démonstration: Supposons B_X est compact. Par le lemme d'approximation 5.3.6, on peut trouver un sous-espace de dimension finie Y de X qui forme un $\frac{1}{2}$-réseau de B_X, c'est à dire

$$\operatorname{dist}(x, y) \leq \frac{1}{2} \quad \text{pour tout } x \in B_X. \tag{5.3.1}$$

Comme X est de dimension infinie et Y est de dimension finie, l'espace quotient X/Y est non nul. Ainsi, nous pouvons trouver un classe d'équivalence $[x] \in X/Y$ avec $\|[x]\| = 0,9$. Alors que, par définition $\|[x]\| = \inf_{x \in [x]} \|x\|$, nous pouvons encore choisir un représentant $x \in [x]$ tel que $\|x\| \leq 1$. En résumé, $\operatorname{dist}(x, y) = \|[x]\| = 0,9$ et $x \in B_X$. Cela contredit (5.3.1) et achève la démonstration. ■

Enfin, nous mentionnons sans démonstration un critères de compacité dans l'espace de fonctions continues sur un compact.

5.3.10 THÉORÈME (ASCOLI-ARZELÀ)
Soit (X, d) un espace métrique compact.
L'adhérence d'une partie $A \subseteq C(X, \mathbb{R})$ est compacte si et seulement si

1. A est borné et
2. *équicontinue i.e.*
   pour tout $\varepsilon > 0$ *il existe* $\delta > 0$ *tel que*

$$d(s,t) \leq \delta \quad \textit{entraîne} \quad |f(s) - f(t)| \leq \varepsilon \quad \textit{pour tout } f \in A.$$

5.3.11 EXEMPLE. Le théorème d'Arzelà-Ascoli entraîne que l'ensemble des fonctions dérivables f avec $\|f'\|_\infty \leq 1$ est compact dans $C[0, 1]$.

5.3.12 EXEMPLE. *Considérons le problème linéaire des valeurs propres*

$$\begin{aligned} -\Delta u(x) &= \lambda u(x) \quad x \in \Omega \\ u(x) &= 0 \\ x &\in \partial\Omega \end{aligned} \tag{5.3.2}$$

avec $\lambda \in \mathbb{R}$, où Ω est un domaine borné de $\mathbb{R}^n$. Évidemment $u = 0$ est une solution de (5.3.2) dans les sens, que nous appelons la solution triviale de (5.3.2).

Nous disons $\lambda \in \mathbb{R}$ est une valeur propre de l'opérateur $-\ Delta$ si (5.3.2) a une solution non triviale faible $u \in H_0^1(\Omega)$ dans le sens suivant

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \lambda \int_\Omega u \, v \, dx \quad \text{pour tout } v \in H_0^1(\Omega).$$

Les solutions non nulles sont les fonctions propres correspondant à la valeur propre λ.

Il suit la Poincaré-Friedrichs inégalité (cf. [**?**]) que

$$\langle u, v \rangle_1 = \int_\Omega \nabla u \cdot \nabla v \, dx \quad (\text{pour tout } u, v \in H_0^1(\Omega))$$

est un produit équivalent intérieure sur $H_0^1(\Omega)$, et pour chaque $U \in L^2(\Omega)$, la fonction $v \mapsto \int_\Omega u \, v \, dx$ est bornée linéaire sur $H_0^1(\Omega)$. Ainsi, par le théorème de représentation de Riesz (théorème **??**), nous pouvons définir $Tu \in H_0^1(\Omega)$ tel que

$$\langle Tu, v \rangle_1 = \int_\Omega u \, v \, dx \quad \text{pour tout } \ v \in H_0^1(\Omega).$$

La cartographie $T : L^2(\Omega) \to H_0^1(\Omega) \subset L^2(\Omega)$ est clairement linéaire, borné et auto-adjoint. De plus, l'exploitation de la compacité de l'intégration $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ (cf. [**?**]), on peut facilement vérifier que T est compact. Utilisation de l'opérateur T, un problème (5.3.2) peut être équivalente réécrite sous la forme $\langle u, v \rangle_1 = \langle \lambda Tu, v \rangle_1$ pour tout $v \in H_0^1(\Omega)$. Donc

$$Tu = \frac{1}{\lambda} u.$$

D'après T est auto-adjoint compact sur l'espace de Hilbert $\mathcal{H} = L^2(\Omega)$, on peut appliquer les résultats précédents pour déterminer la valeurs propres et les fonctions propres du problème (5.3.2). Nous les conclure dans la suivante quoi. Notez que, dans notre situation, $\langle Tu, u \rangle_1 > 0$ pour tout u $\neq 0$. Par conséquent, tous valeurs propres de T sont positifs. Pour la preuve de la simplicité de la première valeur propre λ_1, on peut voir par exemple [**?**].

➠ ➤ Problème (5.3.2) dénombrable a valeurs propres de nombreuses

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots$$

avec $\lambda_n \to +\infty$ et une base orthonormale $\{u_n\}$ de $L^2(\Omega)$ des fonctions propres, avec $\{u_n / \ sqrt\lambda_n\}$ étant une base orthonormée de $H_0^1(\Omega)$. Les valeurs propres $\{\lambda_n\}$ sont les caractérisations variationnelles suivantes :

$$\lambda_n = \inf_{Y \in \mathcal{L}_n} \sup_{u \in Y u \neq 0} \frac{\langle u, u \rangle_1}{\langle Tu, u \rangle_1} = \min_{Y \in \mathcal{L}_n} \max_{u \in Y u \neq 0} \frac{\int_\Omega |\nabla u|^2 \, dx}{\int_\Omega u^2 \, dx}, \tag{5.3.3}$$

où $\mathcal{L}_n$ est l'ensemble de tous les n Y de sous-espace de dimension de $H_0^1(\Omega)$. Alternativement λ_n peut aussi être caractérisé par

$$\lambda_n = \min_{u \in H_0^1(\Omega) u \perp \{u_1, \cdots, u_{n-1}\}} \frac{\int_\Omega |\nabla u|^2 \, dx}{\int_\Omega u^2 \, dx} \quad \text{et} \quad \lambda_1 = \min_{u \in H_0^1(\Omega) u \neq 0} \frac{\int_\Omega |\nabla u|^2 \, dx}{\int_\Omega u^2 \, dx}.$$

La première valeur propre λ_1 est simple (c'est à dire l'espace propre correspondant à λ_1 est de dimension 1) et la fonction propre correspondante ne change pas de signe que nous Alorssions choisir u_1 tel que $u_1 \geq 0$.

5.4 Annexe

Inverses of Operators. The Banach Theorem

One way to solve a matrix equation

$$Ax = y$$

is to find the inverse matric A^{-1} (if it exists) and then the solution is $x = A^{-1}y$.

In this section we will extent this to the infinite-dimensional case. First we need the concept of negative powers for some operators, namely the operators T for which we will make sense of the inverse T^{-1} of T.

5.4.1 DÉFINITION

Suppose that X is a normed linear space over $\mathbb{K}$. A linear operator $T \in \mathcal{B}(X)$ is said to be invertible if there exists a linear operator $S \in \mathcal{B}(X)$ such that

$$TS = I = ST,$$

où $I : X \to X$ is the identity linear operator on X, so that $Ix = x$ for all $x \in X$. In this case, we say S is the inverse linear operator of T, and write $B = T^{-1}$.

5.4.2 REMARQUE

In general, a mapping T is invertible if it is bijective (i.e. injective and surjective). Here, we need more, namely that T^{-1} is required to belong to $\mathcal{B}(X)$; in the other words, we require T^{-1} to be linear and continuous, or bounded.

5.4.3 EXEMPLE. *Consider the linear operator T from ℓ^2 into ℓ^2 defined by $Tx = (0, x_1, x_2, \cdots)$ for all $x = \{x_n\} \in \ell^2$. This continuous linear operator is not invertible since it is not surjective. Consider next the linear operator $S : \ell^2 \to \ell^2$ defined by $S(x) = (x_2, x_3, \cdots)$ for all $x = \{x_n\} \in \ell^2$. This continuous linear operator is also not invertible since it is not injective. Note here that $ST \neq TS$.*

5.4.4 EXEMPLE. *Consider the linear operator $T : L^2[0,1] \to L^2[0,1]$ defined by*

$$(T(f))(t) = (t+1)f(t) \quad \text{for all } f \in L^2[0,1] \text{ and } t \in [0,1].$$

For a possible inverse to T, let us consider the continuous linear operator $S : L^2[0,1] \to L^2[0,1]$ defined by

$$(S(g))(t) = (1+t)^{-1}g(t) \quad \text{for all } g \in L^2[0,1] \text{ and } t \in [0,1].$$

Note that

$$((ST)(f))(t) = (S(T(f)))(t) = (1+t)^{-1}(T(f))(t) = (1+t)^{-1}(1+t)f(t) = f(t)$$

for all $f \in L^2[0,1]$ and $t \in [0,1]$. Similarly, $((TS)(g))(t) = g(t)$ for all $g \in L^2[0,1]$ and $t \in [0,1]$. Hence $ST = I = TS$, so that T and S are both invertible.

Note that if we define the linear operator $\widetilde{T} : L^[0,1] \to L^2[0,1]$ as $(\widetilde{T}(f))(t) = tf(t)$ for all $f \in L^2[0,1]$ and $t \in [0,1]$, then $\widetilde{T}$ is injective, for if $tf(t) = tg(t)$ for almost all $t \in [0,1]$, we get $f(t) = g(t)$ for almost all $t \in [0,1]$. But $\widetilde{T}$ is not surjective since the only possible candidate for f such that $(\widetilde{T}(f))(t) = 1$ for almost all $t \in [0,1]$ must satisfy $f(t) = t^{-1}$ for almost all $t \in [0,1]$. However, this candidate f is not square integrable over $[0,1]$, so that $f \notin L^2[0,1]$. Hence the linear operator $\widetilde{T}$ is not invertible.

We do not have an effectual technique yet to determine whether an operator is invertible. We knew that the determinant of a matrix can be used to determine if the matrix is invertible, but there is no obvious generalization of this to infinite-dimensional case. We therefore have to find other methods to determine whether an operator is invertible.

5.4.5 LEMME
If X is a normed linear space and $T_1, T_2 \in \mathcal{B}(X)$ are invertible then T_1^{-1} and T_2^{-1} are invertible with inverse T_1 and T_2, respectively. Moreover T_1T_2 is invertible with inverse $T_2^{-1}T_1^{-1}$.

5.4.6 THÉORÈME
Suppose that X is a Banach space and that $T \in \mathcal{B}(X)$. Suppose further that $\|T\| < 1$. Then $I - T \in \mathcal{B}(X)$ is invertible and

$$(I-T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

Démonstration: We here give a sketch of the proof. First, the convergence of series $\sum_{n=0}^{\infty} \|T^n\| = \sum_{n=0}^{\infty} \|T\|^n$ (since $\|T\| < 1$) implies that the series $\sum_{n=0}^{\infty} T^n$ converges since $\mathcal{B}(X)$ is a Banach space (cf. Theorem **??**), i.e. the sequence of partial sums $\{S_k\}$ of the series $\sum_{n=0}^{\infty} T^n$ converges to $S = \sum_{n=0}^{\infty} T^n$. Second, it follows from $\|(I-T)S_k - I\| \leq \|T\|^{k+1}$ that $(I-T)S = \lim_{k\to\infty}(I-T)S_k = I$. Finally, we also have $S(I-T) = I$ by using the same argument, then the conclusion. ■

5.4.8 COROLLAIRE
Suppose that X is a Banach space. Then the set $\mathcal{J}$ of all invertible operators in $\mathcal{B}(X)$ is an open set in $\mathcal{B}(X)$.

Démonstration: Let $T \in \mathcal{J}$ and let $\eta = \|T^{-1}\|^{-1}$, we will show that $\mathcal{J}$ contains the ball $B(T,\eta)$ in $\mathcal{B}(X)$. In order to do this, we let $S \in B(T,\eta)$, so that $\|T-S\| < \eta$. Then

$$\|(T-S)T^{-1}\| \leq \|T-S\|\|T^{-1}\| < \|T^{-1}\|\,\|T^{-1}\| = 1,$$

which $I - (T-S)T^{-1} = ST^[-1]$ is invertible by Theorem 5.4.6. Hence $S = ST^{-1}T$ is invertible by Lemma 5.4.5. So $S \in \mathcal{J}$ and then $B(T,\eta) \subset \mathcal{J}$. ■

5.4.1 Application : théorie ergodique

Les théorèmes ergodiques permettent de calculer les moyennes spatiales comme des moyennes temporelles. Nous commençons par prouver une forme préliminaire du théorème ergodique de Von Neumann ; son interprétation suivra.

5.4.10 THÉORÈME (THÉORÈME ERGODIQUE DE VON NEUMANN)
Soit T un opérateur unitaire sur un espace de Hilbert H i.e. $TT^* = T^*T = I$.
Soit P la projection orthogonale sur le sous-espace invariant $\{x \in H : \ Tx = x\}$.
Alors, pour tout $x \in H$, nous avons

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T^n x = Px.$$

Démonstration: Il suffit de démontrer le résultat pour $x \in \ker P$ et pour $x \in \operatorname{Im} P$, car alors le résultat pour tout $x \in H$, suivra grâce à la décomposition orthogonale $H = \ker P \oplus \operatorname{Im} P$.

Pour $x \in \operatorname{Im} P$ le résultat est trivial car dans ce cas $T^n x = Tx = x$ et $x = Px$. Soit $x \in \ker P$. Nous allons d'abord trouver une représentation convenable de $\ker P$. Par définition, $\operatorname{Im} P = \ker(I - T) = \ker(I - T^*)$, car pour les opérateurs unitaires, $Tx = x$ si et seulement si $T^* x = x$ (à vérifier !). Par conséquent, en utilisant la dualité entre les noyaux et les images, le corollaire 5.1.27, nous aurons

$$\ker P = (\ker(I - T^*))^{\perp} = \overline{\operatorname{Im}(I - T)}.$$

Par conséquent, chaque $x \in \ker P$ peut être approchée arbitrairement bien par des vecteurs de la forme $(T - I)y$. Pour $x = (I - T)y$, on obtient une somme télescopique

$$\frac{1}{N} \sum_{n=0}^{N-1} T^n x = \frac{1}{N}(x - T^N x) \to 0 \quad \text{lorsque } N \to \infty.$$

Ce porte sur les vecteurs $x \in \overline{\operatorname{Im}(I - T)}$ par un argument simple approximation. La preuve est terminée. ■

5.4.12 Exemple. Montrer l'argument d'approximation dans cette preuve.

Maintenant nous allons explorer les conséquences du théorème 5.4.10 pour le temps et l'espace des moyennes. Considérons d'abord le cas simple d'un système dynamique discret étudié par Weyl. Soit $\mathbb{T}$ désigne le tore (le cercle unité). Nous avons mis une particule sur le cercle, et on considère ses rotations consécutives à angle fixé $\theta \in (0, 2\pi)$. De manière équivalente, on considère la suite

$$n\theta \mod 2\pi, \quad n \in \mathbb{N}.$$

Cette suite est finie si $\theta/2\pi$ est rationnel et infinie si $\theta/2\pi$ est irrationnel. De plus, dans ce dernier cas l'expérimentation suggère que les valeurs deviennent uniformément réparties sur le cercle pour N grand, voir la figure.

Cette observation est formulée par le théorème ergodique de Weyl :

5.4.13 Théorème (Théorème ergodique de Weyl)

Pour chaque nombre $\theta \in (0, 2\pi)$ tel que $\theta/2\pi$ est irrationnel, et chaque sous-ensemble mesurable $A \subseteq [0, 2\pi]$, on a

$$\frac{|\{n \leq N : \ n\theta \mod 2\pi \in A\}|}{N} \to \frac{\mu(A)}{2\pi} \quad \text{lorsque } N \to \infty. \tag{5.4.1}$$

où μ est la mesure de Lebesgue.

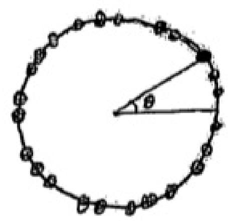


FIGURE 5.1 – Rotations par l'angle θ sont équiréparties sur le cercle

---

Le théorème 5.4.10 implique une forme très générale du théorème ergodique de Weyl, pour une mesure ergodique arbitraire, préservant la transformation d'un espace de probabilité (au lieu d'une rotation irrationnelle du cercle).

5.4.14 DÉFINITION (TRANSFORMATION ERGODIQUE)
Soit (Ω, Σ, μ) un espace de probabilité. On dit q'une transformation $T : \Omega \to \Omega$ *préserve la mesure* si

$$\mu(T^{-1}(A)) = \mu(A)$$

pour tout sous-ensemble mesurable $A \subseteq \Omega$.[1] Une transformation T, injective et qui préserve la mesure est dite *ergodique si les seules fonctions* $f \in L^2(\Omega, \Sigma, \mu)$ *qui satisfont* $f(Tx) = f(x)$ *pour presque tout des* $x \in \Omega$ *sont les fonctions constantes.*

5.4.15 Exercice Montrer que T est ergodique si et seulement si pour tout sous-ensemble mesurable $A \subseteq \Omega$, $T^{-1}(A) = A$ implique $\mu(A) = 0$ ou $\mu(A) = 1$.

5.4.16 THÉORÈME
Soit T une transformation ergodique qui préserve la mesure sur un espace de probabilité (Ω, Σ, μ). Pour chaque $f \in L^2(\Omega, \Sigma, \mu)$ on a

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n \Omega) = \int_\Omega f \, d\mu \qquad (5.4.2)$$

où la convergence est pau sens de la norme L^2.

---

1. Ici $T^{-1}(A) = \{B \in \Omega : TB \in A\}$ est la préimage de A par T.

Démonstration: Pour voir le lien avec le théorème ergodique 5.4.10, on définit un opérateur $U : L^2(\Omega, \Sigma, \mu) \to L^2(\Omega, \Sigma, \mu)$ par

$$Uf = f \circ T, \quad \text{soit} \quad (Uf)(x = f(Tx), \quad x \in \Omega.$$

Comme T est une mesure de préservation, U est un opérateur unitaire. (vérifier !) Comme T est ergodique, le sous-espace invariant de U est l'espace des constantes. La projection orthogonale P dans $L^2(\Omega, \Sigma, \mu)$ sur le sous-espace des constantes est l'intégrale dans la partie droite de (5.4.2). Le résultat est une conséqunce alors du théorème 5.4.10. ■

Le théorème ergodique de Weyl 5.4.13 est un cas particulier du théorème 5.4.16 pour Ω étant le cercle unité, T étant la rotation du cercle par un angle irrationnel θ (pourquoi T est ergodique ?) et pour $f = \mathbb{1}_A$.