IREM

Institut de Recherche sur l'Enseignement des Mathématiques des Pays de la Loire

NANTA IREMICA N° 7

Éléments d'analyse fonctionnelle

Jean Dhombres

1976

ISBN 2-86300-011-X

INTRODUCTION

Dans les pages qui suivent, nous avons voulu rassembler des éléments d'analyse fonctionnelle utiles dans un grand nombre de questions d'intérêt théorique ou pratique : que ce soit l'étude des équations aux dérivées partielles, l'analyse harmonique, la stabilité des équations différentielles, l'approximation des fonctions ou l'analyse numérique.

Ces éléments sont prévus pour servir à des exposés orientés vers des sujets aussi différents et il n'est pas question d'être exhaustif mais d'introduire le vocabulaire communément utilisé en donnant suffisamment d'exemples explicatifs. Au fond, nous avons voulu essentiellement "géométriser" l'analyse fonctionnelle et donc faire jouer un rôle primordial aux espaces de HILBERT.

Par souci de ne pas renvoyer sans cesse le lecteur à d'autres ouvrages, nous avons regroupé en un chapitre 0 tout le matériel algébrique, à vrai dire bien modeste, requis par la suite.

L'attention du lecteur est attirée sur la présence, en fin de texte, d'un index des expressions. Une bibliographie plus que succincte permettra de s'avancer plus avant dans la jungle des textes parus sur l'Analyse Fonctionnelle.

Le présent texte a été utilisé avec des stagiaires de l'Institut de Recherche sur l'Enseignement des Mathématiques et de nombreux ingénieurs en recyclage à l'Ecole Nationale Supérieure des Techniques Avancées et l'Ecole Nationale Supérieure de l'Aéronautique et de l'Espace.

Nous avons surtout voulu fournir un texte utilisable pour un travail indépendant, d'où le grand nombre d'exercices, travail qui serait éventuellement épaulé par un cours oral assez bref. Le niveau requis est volontairement élémentaire (réminiscences des deux premières années de mathématiques de l'Université).

Les solutions des exercices proposés sont fournies dans un autre volume de la collection Nanta-Iremica. (Vol. N° 8)

NANTES 1975.

NANTA-IREMICA

LISTE DES PUBLICATIONS DE L'I.R.E.M. de NANTES

N° 1	Introduction à la logique (EPUISE)	MM. DURAND VAN DEN BOSSCHE
N° 2	Introduction à la Théorie des Ensembles (EPUISE)	MM. DURAND VAN DEN BOSSCHE
N° 3 *	Etude épistémologique et historique de la notion de Nombre réel et de Mesure des grandeurs	M. DHOMBRES
N° 4 *	Documents relatifs au volume N° 3 (EPUISE)	M. DHOMBRES
N° 5	Le langage BASIC (EPUISE)	M. BELHACHE
N° 6 **	Echec en Mathématiques (EPUISE)	M. BIGARD
N° 7	Eléments d'Analyse Fonctionnelle	M. DHOMBRES
N° 8	Exercices d'Analyse Fonctionnelle	M. DHOMBRES
N° 9	Algèbre linéaire et géométrie vectorielle (EPUISE)	M. SEROUX
N° 10 ***	Analyse et Topologie (EPUISE)	Melle VENARD - DHOMBRES
N° 11	Méthodes mathématiques modernes utilisées en théorie de l'Approximation	M. DHOMBRES
N° 12	Introduction à la géométrie et propositions pour la classe de 4è (EPUISE)	M. LETOURNEUX
N° 13	Le PL/1 Optimiseur (EPUISE)	M. BELHACHE
	Quelques difficultés pédagogiques dans l'enseignement de l'Analyse : Second cycle des lycées (EPUISE)	MM. FOUQUES-SEROUX
	Trigonométrie, Algèbre linéaire optique (EPUISE)	M. SEROUX
N° 14	Mathématiques et langage informatique (BASIC)(EPUISE)	M. BETREMA
N° 15	Mathématiques-Physique (spécial IREM de NANTES) (EPUISE)	M. SEROUX
N° 16	Sensibilisation aux structures de données	M. VIVET
	Mathématiques comparées : Angleterre, Québec, Allemagne Fédérale, Chine	M. BIGARD
N° 17	Brochure HP 25	M. QUELFETER
N° 18	Nombres entiers naturels (P.E.N.)	Groupe Ecole Normale
N° 19	Nantinfo 78	M. QUELFETER
N° 20	Mathématiques en LEP	M. PAPIN
N° 21	Liaison CM2-6è	M. CHARLOT

.../...

Pour se procurer ces volumes, s'adresser au Secrétariat :

I.R.E.M. de NANTES
2, chemin de la Houssinière
44072 NANTES CEDEX

* Cet ouvrage a fait l'objet d'une autre édition : Nombres, Mesure et Continu - Epistémologie et Histoires. Collection CEDIC F. NATHAN Paris 1978

** Cet ouvrage a fait l'objet d'une autre édition : Echec en Mathématiques Collection CEDIC Paris 1976

*** Cet ouvrage fera prochainement l'objet d'une édition en librairie.

TABLE DES MATIERES

- CHAPITRE 0 -

RAPPELS D'ALGEBRE

1. NOTATIONS.

Tout au long de ce texte, nous utilisons le vocabulaire usuel de la théorie des ensembles.

Rappelons toutefois les notations les plus courantes : soit un ensemble formé d'éléments x , y ,... Etant donnés deux sous-ensembles A et B de l'ensemble E :

$x \in A$ signifie que x est un élément de A ,

$x \notin A$ signifie que x n'est pas un élément de A ,

$A \cap B$ intersection de A et B est le sous-ensemble des points de E appartenant à la fois à A et B ,

$A \cup B$ réunion de A et B est le sous-ensemble des points de E appartenant à l'un au moins des ensembles A, B.

$\complement A$ complémentaire de A est le sous-ensemble des points de E n'appartenant pas à A .

Si $A \cap B = A$, on dit que A est inclus dans B et on note $A \subset B$, ce qui signifie que tout point de A est dans B . Si l'on veut spécifier que $A \subset B$ mais qu'il existe un élément $b \in B$ n'appartenant pas à A (ou encore que l'on n'a pas $B \subset A$), on précise :

$$A \subset B \quad \text{strictement}.$$

Une application f , définie sur un ensemble E à valeurs dans un ensemble E' , est une correspondance qui permet d'associer à tout élément

x de E un élément unique x' de E'. On note :

$$f : x \longrightarrow f(x) = x'$$

$$E \longrightarrow E'$$

(toutefois, dans ce texte, nous utilisons rarement le mot application mais plutôt les mots fonction, opérateur)

Une application est surjective si tout élément de E' est image d'un élément de E, injective si deux éléments distincts de E ont pour image des éléments distincts de E'.

Une application bijective est une application à la fois injective et surjective.

Si A est un sous-ensemble de E, $f(A)$ désigne le sous-ensemble de E' constitué par les éléments x' de E' associés à au moins un élément x de A par f selon :

$$x' = f(x)$$

$f(A)$ est appelé l'image directe de A.

Si A' est un sous-ensemble de E', $\overset{-1}{f}(A')$ désigne le sous-ensemble de E constitué par les éléments x de E tels que $f(x) \in A'$.

$\overset{-1}{f}(A')$ est appelée l'image inverse de A'.

On rappelle que si E, E' et E'' sont trois ensembles, f une application définie sur E à valeurs dans E' et g une application définie sur E' à valeurs dans E'', l'application composée $g \circ f$ est l'application définie sur E à valeurs dans E'', qui à x dans E fait correspondre $g(f(x))$ dans E'' :

$$\underbrace{E \xrightarrow{f} E' \xrightarrow{g} E''}_{g \circ f}$$

Lorsque f est une application bijective, on note $\overset{-1}{f}$ l'application inverse (ou réciproque) de f :

$$E \xrightarrow{f} E'$$

$$E' \xrightarrow[\overset{-1}{f}]{} E$$

$$f o \overset{-1}{f}(x') = x' \quad \text{et} \quad \overset{-1}{f} o f(x) = x$$

Remarque : Soit f une application définie sur un ensemble E à valeurs dans un ensemble E' . Il importe de bien distinguer entre la fonction f , correspondance entre les ensembles E et E' , et la valeur $f(x)$ assignée au point x par la correspondance que définit la fonction f .
On notera donc :

Soit f la fonction $x \longrightarrow f(x)$, de E dans E' .

Nous allons assez rapidement rappeler dans ce Chapitre 0 des notions fondamentales d'algèbre linéaire. Nous nous intéressons aux espaces vectoriels et nous démontrons de nombreuses propriétés de ces espaces en essayant de distinguer nettement les propriétés des espaces de dimension finie de cellesdes espaces de dimension infinie.

Par ailleurs, nous éviterons l'emploi des bases en essayant de donner des raisonnements intrinsèques. Ce double point de vue est nécessaire quant à l'analyse fonctionnelle car les espaces considérés plus loin sont tous de dimension infinie et il n'y a aucun intérêt en général à privilégier une base.

Le lecteur retrouvera de nombreux théorèmes connus sous une forme quelque peu différente. Pour permettre la comparaison, nous avons rappelé des résultats en dimension finie, et par rapport à une base donnée, dans un paragraphe réservé aux matrices.

2. ESPACES VECTORIELS SUR UN CORPS.

Dans ce qui suit, K désigne exclusivement soit le corps des nombres réels $\mathbb{R}$, soit le corps des nombres complexes $\mathbb{C}$. K s'appelle le corps des scalaires. Pour un élément λ de K , $|\lambda|$ est soit le module (si $K = \mathbb{C}$, $|\lambda| = |\lambda_1 + i\lambda_2| = \sqrt{\lambda_1^2 + \lambda_2^2}$), soit la valeur absolue (si $K = \mathbb{R}$).

2.1. Définition.

Un espace vectoriel E sur le corps K est un ensemble sur lequel sont données :

a) une loi interne notée additivement telle que :

$x+(y+z) = (x+y)+z$	associativité
$x+0 = 0+x = x$	0 est élément neutre
$x+(-x) = (-x)+x = 0$	$(-x)$ est élément symétrique de x
$x+y = y+x$	commutativité

(E,+) forme alors un groupe commutatif additif.

b) une loi externe qui à tout élément λ du corps K et à tout élément x de E fait correspondre un élément de E noté λx tel que :

$\lambda(x+y) = \lambda x + \lambda y$	distributivité
$(\lambda+\mu)x = \lambda x + \mu x$	distributivité
$(\lambda\mu)x = \lambda(\mu x)$	associativité
$1.x = x$	1 étant l'unité de K

Les éléments d'un espace vectoriel sur K sont appelés des vecteurs.

Il est facile de démontrer que l'élément neutre 0 est unique, ainsi que l'élément symétrique $(-x)$ d'un élément x . De même, on obtient sans peine que $0x=0$ $(\lambda x = (\lambda + 0)x = \lambda x + 0x)$, et également $\lambda.0 = 0$

(il faut remarquer que nous notons de la même façon l'élément neutre 0 de l'espace vectoriel E et l'élément neutre 0 du corps K).

Lorsque $K = \mathbb{R}$, on dit que l'espace vectoriel est réel.

Lorsque $K = \mathbb{C}$, on dit que l'espace vectoriel est complexe.

2.2. Exemples.

a) L'ensemble des vecteurs libres de la géométrie euclidienne usuelle, considérés à une équipollence près, constitue un espace vectoriel. Si à tout vecteur libre $\vec{U}$, on fait correspondre le vecteur équipollent dont l'origine est fixée en un point 0 , on obtient ainsi l'ensemble des vecteurs liés de même origine 0 . Un vecteur est alors déterminé par un seul point : son extrémité. L'ensemble des points de la géométrie ordinaire est ainsi structuré en un espace affine.

b) Soit S une suite formée de n éléments du corps K : $S = (x_1, x_2, \ldots x_n)$.
L'ensemble de toutes les suites de n éléments constitue un ensemble noté K^n .

On peut définir une addition et une multiplication par un scalaire sur K^n :

$$S + S' = (x_1 + x_1', x_2 + x_2', \ldots, x_n + x_n')$$

$$\lambda S = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$

Muni de ces deux lois, on vérifie que K^n constitue un espace vectoriel sur le corps K . K^n s'appelle l'espace des vecteurs d'ordre n sur K .
En particulier K^1 s'identifie avec K .

c) Appelons $\mathcal{F}[0,1]$ l'ensemble des fonctions numériques (c'est-à-dire à valeurs réelles) définies sur le segment $[0,1]$ de l'axe réel $\mathbb{R}$.

On peut définir une addition et une multiplication par un scalaire réel sur $\mathcal{F}[0,1]$:

$$f + g \; : \; x \longrightarrow f(x) + g(x) \qquad \text{sur } [0,1]$$

$$\lambda f \; : \; x \longrightarrow \lambda f(x) \qquad \text{sur } [0,1]$$

Muni de ces deux lois, il est facile de vérifier que $\mathcal{F}[0,1]$ constitue un espace vectoriel sur le corps des nombres réels.

2.3. Sous-espaces vectoriels.

Un sous-ensemble F d'un espace vectoriel E sur K est un sous-espace lorsqu'il constitue lui-même un espace vectoriel pour l'addition et la multiplication par un scalaire définies sur E.

Pour que $F \neq \emptyset$ constitue un sous-espace de E, il faut et il suffit que pour tous les éléments x et y de F, $x - y$ appartienne également à F.

Ainsi, reprenant l'exemple (c) du paragraphe 2.2, on constate que l'ensemble $C[0,1]$ formé par les fonctions numériques définies et continues sur $[0,1]$ est un sous-espace vectoriel de $\mathcal{F}[0,1]$.

Soient F_1 et F_2 deux sous-espaces vectoriels d'un même espace vectoriel E, l'intersection $F_1 \cap F_2$ constitue un sous-espace vectoriel non vide de E (l'élément nul 0 de E appartient à tout sous-espace vectoriel, donc à toute intersection de tels sous-espaces). Si $F_1 \cap F_2 = \{0\}$ on convient de dire que les deux sous-espaces sont "disjoints", ou mieux encore complémentaires.

On appelle somme de deux sous-espaces, le sous-ensemble de E constitué par tous les vecteurs de E de la forme $x + y$ où $x \in F_1$ et $y \in F_2$. La somme de deux sous-espaces vectoriels est encore un sous-espace vectoriel.

Lorsque les sous-espaces F_1 et F_2 sont disjoints, leur somme est dite somme directe et on la note $F_1 + F_2$ ou mieux $F_1 \oplus F_2$.

Dans ce cas, un élément x de $F_1 \oplus F_2$ s'écrit d'une façon unique sous la forme $x_1 + x_2$ où $x_1 \in F_1$ et $x_2 \in F_2$.

On dit que deux sous-espaces vectoriels F_1 et F_2 d'un espace vectoriel E sur K sont supplémentaires lorsque ces deux espaces sont disjoints et que $F_1 \oplus F_2 = E$.

Il revient au même de dire que tout vecteur x de E se décompose d'une façon unique sous la forme $x_1 + x_2$ où $x_1 \in F_1$ et $x_2 \in F_2$.

2.4. Système de générateurs et linéaire indépendance.

Soit F un sous-espace vectoriel d'un espace vectoriel E sur le corps K . On dit qu'un ensemble A de vecteurs de F est un système de générateurs pour F si tout vecteur de F peut s'écrire sous forme d'une combinaison linéaire finie de vecteurs de A , c'est-à-dire si pout tout $x \in F$ il existe des scalaires λ_1 , λ_2 ... λ_n du corps K tels que :

$$(1) \qquad \sum_{i=1}^{i=n} \lambda_i \; x_i = x$$

où x_1 , x_2 ... et x_n appartiennent à A . Le nombre n variant en général avec chaque élément x de F .

On dit aussi que F est le sous-espace vectoriel engendré par A . Il est naturel de rechercher les systèmes de générateurs comportant le moins d'éléments possibles. Le problème est lié au problème de l'unicité de l'écriture (1) pour un vecteur x quelconque de F . Pour que cette écriture soit unique, il faut et il suffit qu'une égalité de la forme (2) :

$$(2) \qquad \sum_{i=1}^{i=k} \lambda_i \; x_i = 0$$

où les x_i appartiennent à A , entraîne nécessairement $\lambda_i = 0$ pour $i = 1, 2 \ldots k$.

Définition : Un ensemble de vecteurs A d'un sous-espace vectoriel E est dit libre (ou linéairement indépendant) lorsque toute égalité de la forme :

$$\sum_{i=1}^{i=k} \lambda_i \, x_i = 0 \qquad \text{où } \lambda_i \in K \text{ et } x_i \in A \;, \quad k \geqslant 1$$

entraîne $\lambda_i = 0$, pour $i = 1, 2, \ldots k$.

On démontre alors le premier résultat suivant en utilisant l'axiome du choix.

Théorème 1 - Dans tout espace vectoriel, il existe au moins un système libre de générateurs.

Définition : Un système libre de générateurs d'un espace vectoriel s'appelle une base.

Le théorème précédent revient donc à assurer que dans tout espace vectoriel E , il existe un ensemble A de vecteurs (linéairement indépendants) tels que tout vecteur x de E s'écrive d'une façon unique sous la forme d'une combinaison linéaire finie d'éléments de A .

Soient k vecteurs donnés x_1 , x_2 ... x_k d'un espace vectoriel E . Si l'on considère l'ensemble des éléments x qui en sont des combinaisons linéaires,

$$x = \sum_{i=1}^{i=k} \lambda_i \, x_i$$

on obtient un sous-espace vectoriel F , visiblement engendré par les vecteurs x_1 , x_2 ... x_k .

Théorème 2 - Dans le sous-espace F engendré par un système de k générateurs (x_1 .... x_k), tout système linéairement indépendant (y_1 , y_2 .... y_n) comporte un nombre n de vecteurs tel que $n \leqslant k$.

La méthode de démonstration est basée sur le théorème dit de l'échange. On prend les k générateurs (x_1 ... x_k) auxquels on ajoute le vecteur y_1 . Le système (y_1 , x_1 ... x_k) est un système de $k+1$ vecteurs, qui engendre F et qui est nécessairement linéairement dépendant puisque y_1 s'écrit comme combinaison linéaire des x_i , il existe donc une relation à coefficients non tous nuls :

$$\alpha_1 y_1 + \beta_1 x_1 + \cdots + \beta_k x_k = 0$$

où les β_i ne peuvent pas être tous nuls (car $y_1 \neq 0$, puisqu'il appartient à un système libre). L'un des β_i est donc différent de 0 et le x_i correspondant dépend des k vecteurs y_1 , x_1 $\ldots, x_{i+1}$, x_{i+1} , $\ldots, x_k$.
Le système (y_1 , x_1 , x_2 $\ldots, x_{i-1}$, x_{i+1}, $\ldots, x_k$) engendre F . On peut ainsi continuer le processus en adjoignant y_2 et finalement on obtient, si $n \leqslant k$, un système de générateurs de F : (y_1 , y_2 $\cdots$ y_n , x_{n+1} $\cdots$ x_k).
Mais on ne peut avoir $k \leqslant n$ car le procédé assurerait que (y_1 , y_2 $\cdots$ y_k) engendre F et par suite $y_{k+1}, \ldots, y_n$ seraient des combinaisons linéaires des précédents. Ceci est contraire à l'hypothèse de linéaire d'indépendance des $(y_1 \cdots y_n)$.

2.5. Dimension d'un espace vectoriel.

Définition : On dit qu'un espace vectoriel E est de dimension finie s'il existe une base de E comportant un nombre fini de vecteurs.

Autrement , on dit que E est de dimension infinie.

Théorème 3 - Dans un espace vectoriel E de dimension finie, toutes les bases de E sont finies et ont un même nombre de vecteurs. Ce nombre s'appelle la dimension de l'espace.

Soient A_1 et A_2 deux bases de l'espace vectoriel E . D'après le théorème 2, le nombre des vecteurs contenus dans A_1 ou A_2 est nécessairement inférieur ou égal au nombre fini de vecteurs contenus dans une certaine base dont l'existence est assurée par l'hypothèse de la dimension finie.

Il y a donc k_1 vecteurs dans A_1 et k_2 vecteurs dans A_2 . Appliquant de nouveau le théorème 2 à A_1 et A_2 , on trouve :

$$k_1 \leqslant k_2 \quad \text{et} \quad k_2 \leqslant k_1 \quad , \quad \text{soit} \quad k_1 = k_2$$

Remarque : Le théorème précédent permet en particulier d'assurer qu'un espace vectoriel est de dimension infinie lorsqu'on a trouvé un système infini de vecteurs linéairement indépendants.

Exemples :

L'espace vectoriel K^n , des vecteurs d'ordre n sur K , est de dimension n .

L'espace des vecteurs de la géométrie ordinaire est à trois dimensions.

L'espace $\mathbb{C}[0,1]$ est de dimension infinie. Pour cela, il suffit de vérifier que le système (f_0 , f_1 ,..., f_n ,...)

$$\begin{array}{llll} f_0 & : & x & \longrightarrow 1 \\ f_1 & : & x & \longrightarrow x \\ \vdots & & & \\ f_n & : & x & \longrightarrow x^n \\ \vdots & & & \end{array}$$

etc.

constitue un système infini de vecteurs linéairement indépendants. Supposons qu'il existe une relation :

$$\sum_{i=0}^{i=n} \lambda_i f_i = 0$$

c'est-à-dire :

$\sum_{i=0}^{i=n} \lambda_i f_i(x) \equiv 0$,pour tout x réel compris entre 0 et 1 . Ou encore

$\sum_{i=0}^{i=n} \lambda_i x^i_i \equiv 0$ pour tout x réel dans $[0,1]$. En dérivant, on vérifie

facilement que $\lambda_i = 0$ pour $i = 0, 1, 2 \dots n$. Ceci démontre la linéaire indépendance des f_i . Dans la suite de ce texte, nous aurons essentiellement à nous occuper d'espaces vectoriels de dimension infinie dont les vecteurs sont des fonctions, comme c'est le cas pour $\mathbb{C}[0,1]$. De tels espaces s'appellent des espaces fonctionnels et leur étude constitue l'analyse fonctionnelle.

Enonçons maintenant le théorème de la base incomplète :

Théorème 4 - Soit E un espace vectoriel de dimension n et soient p éléments de E linéairement indépendants ($p < n$). On peut alors adjoindre à ces P éléments, ($n - p$) autres éléments de E afin d'obtenir une base de E .

Soient (e_1 , e_2 ... e_p) les p vecteurs linéairement indépendants de E . Il existe dans E au moins un vecteur f_1 tel que le système (e_1 , e_2 ... e_p , f_1) soit linéairement indépendant (grâce à l'hypothèse $p < n$). On peut répéter le raisonnement jusqu'à introduire ($n - p$) nouveaux vecteurs de telle sorte que (e_1 , e_2 ,... e_p , f_1 , f_2 , f_{n-p}) soit une base de E .

Par ce procédé, si l'on appelle E_1 le sous-espace vectoriel engendré par (e_1 , e_2 ,... e_p) et E_2 le sous-espace vectoriel engendré par (f_1 , f_2 ,... f_{n-p}), on constate que :

$$E = E_1 \oplus E_2 \qquad \dim(E) = n = \dim(E_1) + \dim(E_2)$$
$$= p + (n-p)$$

3. HOMOMORPHISMES D'ESPACES VECTORIELS.

Soient E_1 et E_2 deux espaces vectoriels définis sur un même corps K et p une application définie sur E_1 à valeurs dans E_2 .

3.1. Opérateurs linéaires.

On dit que p est un homomorphisme d'espaces vectoriels (ou encore un opérateur linéaire) si p vérifie :

$$\begin{cases} P(x+y) = P(x) + P(y) \\ P(\lambda x) = \lambda P(x) \end{cases}$$

pour tous x et y dans E et tout λ dans K .

On remarque immédiatement que :

$$\underline{P(0) = 0} \qquad [P(x+0) = P(x) + P(0) = P(x)]$$

Lorsque P est une application bijective de E_1 sur E_2 et que P est un homomorphisme, on dit que P est un <u>isomorphisme d'espaces vectoriels</u>.

On remarque en particulier que si P est un isomorphisme, alors $\overset{-1}{P}$ est aussi un opérateur linéaire :

$$\overset{-1}{P}(x'+y') = \overset{-1}{P}(x') + \overset{-1}{P}(y') \quad \text{et} \quad \overset{-1}{P}(\lambda x') = \lambda \overset{-1}{P}(x')$$

pour tous x' et y' dans E_2 et tout λ dans K .

3.2. Terminologie.

On appelle $\ker P$ et on prononce <u>noyau de l'opérateur linéaire P</u> le sous-ensemble de E_1 constitué par les vecteurs x tels que $P(x)=0$.

$\ker P$ est visiblement un sous-espace vectoriel de E_1 réduit au seul vecteur 0 si et seulement si P est une application injective.

On appelle $\operatorname{Im} P$ et on prononce <u>image de l'opérateur linéaire P</u> le sous-ensemble de E_2 constitué par l'image directe de E_1 selon P (Cf. § 1).

$\operatorname{Im} P$ est visiblement un sous-espace vectoriel de E_2 , coïncidant avec E_2 si et seulement si P est une application surjective.

3.3. Théorème d'isomorphisme.

Théorème 5 - <u>Pour que deux espaces vectoriels de dimension finie soient isomorphes, il faut et il suffit que ces deux espaces aient même dimension</u>.

Supposons d'une part que les espaces vectoriels E et F aient même dimension. Soit $(e_1, \dots, e_n)$ une base de E et $(f_1, f_2 \dots f_n)$ une base de F .

La correspondance qui au vecteur $x = \sum_{i=1}^{i=n} x_i e_i$ associe le vecteur

$y = \sum_{i=1}^{i=n} x_i f_i$ est une bijection de E dans F . C'est évidemment un isomorphisme qui transforme la base (e_1 , ..., e_n) en la base (f_1 ,..., f_n) ; cet isomorphisme dépend du choix initial d'une base sur E et d'une base sur F .

Réciproquement, supposons donné un isomorphisme P entre les espaces vectoriels de dimension finie E et F . Soit (e_1, e_2 ... e_n) une base de E et appelons $f_i = P(e_i)$, le transformé de e_i par l'isomorphisme P .

Le vecteur $x = \sum_{i=1}^{i=n} x_i e_i$ est transformé par P en un vecteur $y = \sum_{i=1}^{i=n} x_i \, P(e_i)$

(linéarité de P) et l'on remarque que lorsque x décrit E , y décrit F (P est surjectif). Par conséquent, le système des ($P(e_1)$, $P(e_2)$,...., $P(e_n)$) constitue un système de générateurs pour l'espace vectoriel F . La dimension de F est donc inférieure à la dimension de E . Par suite du rôle symétrique de E et de F , on déduit que la dimension de E est égale à la dimension de F .

Ainsi, dans un isomorphisme, toute base de l'un des espaces devient une base de l'autre et l'isomorphisme est entièrement déterminé dès que l'on connaît deux bases correspondantes.

Exemple : Tout espace vectoriel sur le corps K à n dimensions est isomorphe à l'espace vectoriel K^n . Sur cet espace (Cf § 2.2, exemple b), il existe une base canonique :

$$
\begin{aligned}
e_1 &= (1, 0, \ldots\ldots 0) \\
e_2 &= (0, 1, \ldots\ldots 0) \\
&\vdots \\
e_n &= (0, 0 \ldots\ldots 1)
\end{aligned}
$$

Tout $x = (x_1 , x_2 \ldots x_n)$ de K^n s'écrit sous la forme (unique) :

$$x = \sum_{i=1}^{i=n} x_i \, e_i$$

Pour l'étude algébrique d'un espace vectoriel de dimension finie n, on peut donc se ramener à l'étude de K^n repéré dans sa base canonique.

3.4. Propriété des opérateurs linéaires dans un espace vectoriel de dimension finie.

Soit E un espace vectoriel de dimension n sur K et F un espace vectoriel quelconque sur K. On a la relation suivante pour un opérateur P, défini sur E, à valeurs dans F.

Théorème 6 -

$$\boxed{\dim(\ker P) + \dim(\operatorname{Im} P) = n}$$

Du point de vue de la terminologie, la dimension de $\operatorname{Im} P$ s'appelle fréquemment le rang de l'opérateur linéaire P et la dimension de $\operatorname{Ker} P$ s'appelle la codimension de l'opérateur linéaire P.

Soit k la dimension du noyau et $(e_1, e_2, \dots e_k)$ une base de ce sous-espace vectoriel. D'après le théorème de la base incomplète, démontré en 2.5 (théorème 4), il existe $(n-k)$ vecteurs $e_{k+1}, e_{k+2}, \dots e_n$ tels que $(e_1, e_2, \dots, e_k, e_{k+1}, \dots, e_n)$ forment une base de E.

Les vecteurs $P(e_1), P(e_2), \dots, P(e_n)$ engendrent le sous-espace vectoriel $\operatorname{Im} P$ (Cf. démonstration du théorème 5, § 3.3).
Comme $P(e_1) = P(e_2) = \dots = P(e_k) = 0$, on en déduit que les vecteurs $(P(e_{k+1}), \dots\dots, P(e_k))$ engendrent le sous-espace vectoriel $\operatorname{Im} P$.
Or ces $(n-k)$ vecteurs sont indépendants. En effet, s'il existait une relation :

$$\sum_{i=k+1}^{i=n} \alpha_i P(e_i) = 0,$$

on aurait :

$$\sum_{i=k+1}^{i=n} P(\alpha_i e_i) = P\left(\sum_{i=k+1}^{i=n} \alpha_i e_i\right) = 0,$$

donc $\sum_{i=k+1}^{i=n} \alpha_i e_i$ appartiendrait à $\ker P$. D'après le choix de la base, la seule possibilité est :

$$\alpha_{k+1} = 0 \qquad \alpha_{k+2} = 0 \quad \dots \quad \text{et } \alpha_n = 0$$

Finalement, l'espace $\mathrm{Im}\ P$ est engendré par la base $(P(e_{k+1}), \dots\dots, P(e_n))$ comportant $n-k$ vecteurs. Donc la dimension de $\mathrm{Im}\ P$ est $n-k$, ce qui assure la relation énoncée.

Tirons une conséquence directe de cette relation.

Corollaire : Entre deux espaces de même dimension E_1 et E_2, tout opérateur linéaire injectif (c'est-à-dire tel que $\ker P = 0$) est aussi surjectif ; et tout opérateur linéaire surjectif (c'est-à-dire tel que $\mathrm{Im}\,P = E_2$) est aussi injectif.

Cela signifie que pour un opérateur linéaire défini et à valeurs sur un même esapce de dimension finie, l'hypothèse d'injectivité ou celle de surjectivité entraîne ipso facto la bijectivité.

La situation d'un espace vectoriel de dimension infinie est tout à fait différente et le corollaire précédent est inexact (Cf. § 3.10).

3.5. Quelques définitions algébriques.

Nous aurons encore besoin, dans la suite, de quelques notions supplémentaires d'algèbre générale.

Anneau : Un ensemble A est un anneau lorsque deux lois internes, notées + et x sont données sur A et vérifient les hypothèses suivantes :

a) Pour la loi +, A est un groupe additif commutatif (ou abélien)

$(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3)$	associativité
$x + 0 = 0 + x = x$	élément neutre
$x + (-x) = 0$	élément symétrique
$x + y = y + x$	commutativité

b) Pour la loi x, on a :

$$x_1\,(x_2\,x_3) = (x_1\,x_2)\,x_3 \qquad \text{associativité}$$

$$\left.\begin{aligned}(x_2 + x_3)\,x_1 &= x_2\,x_1 + x_3\,x_1\\ x_1\,(x_2 + x_3) &= x_1\,x_2 + x_1\,x_3\end{aligned}\right\}\quad \text{distributivité}$$

Si $x_1\,x_2 = x_2\,x_1$, on dit que A est un <u>anneau commutatif</u>.

L'ensemble des nombres rationnels, l'ensemble des nombres réels constituent des exemples d'anneau.

Lorsque A possède un élément neutre pour la multiplication, on dit que l'on a un <u>anneau unifère</u>.

<u>Algèbre</u> : Un ensemble A est dit une <u>algèbre sur le corps K</u> lorsque A est muni de deux lois internes +, x et d'une multiplication extérieure par les éléments du corps K pour lesquelles A est à la fois un anneau et un espace vectoriel (sur le corps K) satisfaisant l'hypothèse supplémentaire :

$$\lambda\mu\,(x_1\,x_2) = (\lambda x_1)\,(\mu x_2) \qquad \text{pour } x_1 \text{ et } x_2 \text{ dans } A\,,\ \lambda \text{ et } \mu \text{ dans } K\,.$$

3.6. <u>Espace des opérateurs linéaires</u>.

Soient E et F deux espaces vectoriels sur un corps K . L'ensemble des opérateurs linéaires définis sur E , à valeurs dans F , se note $\mathrm{Hom}\,(E,F)$

Soient P_1 et P_2 deux éléments de $\mathrm{Hom}\,(E\,,F)$. Il est naturel de définir la somme de P_1 et P_2 par :

$$P_1 + P_2 : x \longrightarrow (P_1 + P_2)(x) = P_1(x) + P_2(x)$$

et la multiplication de P par un scalaire λ de K par :

$$\lambda P \;:\; x \longrightarrow (\lambda P)(x) = \lambda P(x)$$

Muni de ces deux lois, Hom(E , F) constitue un espace vectoriel sur le même corps K , comme il est facile de le vérifier.

Remarques : Lorsque l'on prend $F = K$, on remplace souvent la notation Hom (E , K) par la notation E^* qui se prononce : dual algébrique de l'espace vectoriel E .

Un opérateur linéaire défini sur E , à valeurs dans K , s'appelle une forme linéaire sur E .

Si $K = \mathbb{C}$, on dit que la forme linéaire est complexe.

Si $K = \mathbb{R}$, on dit que la forme linéaire est réelle.

3.7. Algèbre des opérateurs linéaires.

Soit E un espace vectoriel sur le corps K . Considérons l'espace vectoriel Hom (E , E) constitué par les opérateurs linéaires définis sur l'espace E et à valeurs dans l'espace E .

On peut définir une loi multiplicative sur Hom (E , E) en utilisant la loi de composition des applications. Si P_1 et P_2 sont dans Hom (E , E) on définit $P_1 \; P_2$ par $P_1 \circ P_2$:

$$P_1 \; P_2 \qquad \begin{array}{l} x \longrightarrow P_1 \circ P_2(x) = P_1(P_2(x)) \\ E \longrightarrow E \end{array}$$

$P_1 \; P_2$ est aussi un élément de Hom (E , E) comme il est facile de le vérifier On définit les puissances de P (notées P^n) et les polynômes formés à partir de P .

On vérifie que Hom (E , E) constitue ainsi une algèbre sur le corps K . Cette algèbre est unifère puisque l'application identique I de E sur E (I : $x \longrightarrow x$) est un élément unité pour la multiplication.

L'algèbre Hom (E, E), et certaines sous-algèbres particulières, seront très souvent utilisées dans la suite de ce texte. Nous donnerons comme exemple typique l'algèbre des matrices carrées d'ordre n, puis un exemple fonctionnel (Cf § 3.9 et 3.10).

3.8. Inverse d'un opérateur.

Soit E un espace vectoriel sur K et P un opérateur linéaire appartenant à Hom (E, E). On dit que P' est un inverse de P (à gauche) lorsque P' est un élément de Hom (E, E) tel que :

$$\boxed{P'P = I}$$

Cette définition coïncide avec la définition générale d'un inverse dans une algèbre (ici Hom (E, E)).

Tout d'abord, si P' existe, P est nécessairement injectif. En effet, si $x_1 \neq x_2$, on a $P(x_1) \neq P(x_2)$ (sinon on aurait $P'P(x_1) = P'P(x_2) = x_1 = x_2$)

1er cas : E est de dimension finie.

Le corollaire du théorème 6 implique que si P' existe, alors P est une bijection, c'est-à-dire un automorphisme (isomorphisme d'un espace vectoriel sur lui-même). Ce résultat entraîne en particulier que P' est unique et que $P'P = PP' = I$ (inverse à gauche et à droite).

On note en général l'inverse de P par P^{-1}. Bien entendu, il existe en général des opérateurs linéaires n'admettant pas d'inverse (Hom (E, E) n'est pas un corps). Dans l'exemple 3.9, nous reprendrons ces questions sur une représentation analytique d'un opérateur linéaire : la théorie des matrices.

2ème cas : E est de dimension infinie.

L'existence d'un inverse à gauche n'entraîne pas la surjectivité de l'opérateur P. Nous allons donner un contre-exemple.

Soit E l'ensemble constitué par toutes les suites de nombres réels. Un élément de E est une suite $x = (x_1, x_2, \ldots, x_n, \ldots)$. On peut définir l'addition de deux suites :

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n, \ldots)$$

et la multiplication d'une suite par un scalaire :

$$\lambda x = (\lambda x_1 , \lambda x_2 , \dots \lambda x_n \dots)$$

Muni de ces deux lois, E constitue un espace vectoriel sur le corps des nombres réels (quelquefois noté R^{N}). Définissons deux opérateurs linéaires sur E , l'opérateur P et l'opérateur P' :

$$x = (x_1 , x_2 , \dots , x_n , \dots)$$

$$Px = (x_1 , 0 , x_2 , 0 , x_3 , 0 , x_4 , 0 , \dots , 0 , x_n , 0 \dots)$$

$$P'x = (x_1 , x_3 , x_5 , x_7 , \dots x_{2n-1} , x_{2n+1} , \dots)$$

Visiblement P et P' sont des opérateurs linéaires définis sur E et l'on a la relation

$$P'P = I$$

Toutefois $P(E) \neq E$ et l'opérateur P n'est pas un opérateur surjectif (mais est injectif). On remarquera aussi que P n'est pas un inverse à gauche de P', en ce sens que l'on a :

$$PP' \neq I$$

Ce contre-exemple est destiné à bien marquer les notables différences entre le cas de la dimension finie et celui de la dimension infinie.

• Dans le cas général, on dispose cependant du théorème suivant :

<u>Théorème 7</u> - <u>Soient E et F deux espaces vectoriels sur le même corps K et P un opérateur linéaire, élément de Hom (E, F). Si P est une application bijective, l'application inverse P^{-1} est un opérateur linéaire, élément de Hom (F, E).</u>

En effet, montrons que P^{-1} est un opérateur linéaire et vérifie :

$$P^{-1}(\alpha y_1 + \beta y_2) = \alpha P^{-1}(y_1) + \beta P^{-1}(y_2)$$

Posons :

$$x_i = P^{-1}(y_i) \qquad (i = 1 \text{ ou } 2)$$

On a :

$$P(\alpha x_1 + \beta x_2) = \alpha P(x_1) + \beta P(x_2)$$

grâce à la linéarité de P .

Par conséquent $\alpha x_1 + \beta x_2$ est le correspondant de $\alpha y_1 + \beta y_2$ par P^{-1} , ce qui donne la linéarité de P^{-1} :

$$P^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha P^{-1}(y_1) + \beta P^{-1}(y_2)$$

3.9. Algèbre des matrices carrées.

Nous allons examiner maintenant le cas d'un espace vectoriel E de dimension finie n sur le corps K . Au lieu de faire des raisonnements intrinsèques, nous nous plaçons dans une base (e_1 , e_2 , ..., e_n) de E choisie a priori : ceci va nous permettre de retrouver les théorèmes précédents dans le cadre de la théorie des matrices et sous un aspect analytique.

Un opérateur linéaire P , défini sur E et à valeurs dans E (élément de Hom (E , E)) est déterminé par les vecteurs ($P(e_1)$, $P(e_2)$, ..., $P(e_n)$). En effet si x est un vecteur de E , il existe n constantes x_i dans K :

$$x = \sum_{i=1}^{i=n} x_i \quad e_i$$

et

$$P(x) = \sum_{i=n}^{i=n} x_i \; P(e_i)$$

grâce à la linéarité de P (Cf § 3.3).

Posons alors :

$$P(e_i) = \sum_{j=1}^{j=n} a_{ji} \quad e_j$$

où les a_{ji} sont des constantes, en nombre n^2 , appartenant au corps des scalaires K . C'est-à-dire que nous décomposons les n vecteurs $P(e_i)$ selon la base donnée (e_1 , e_2 , ... e_n).

On peut aussi écrire la décomposition de $P(x)$ selon cette même base :

$$P(x) = \sum_{j=1}^{j=n} x_j' \, e_j$$

On obtient :

$$P(x) = \sum_{j=1}^{j=n} x_j' \, e_j = \sum_{i=1}^{i=n} x_i \sum_{j=1}^{j=n} a_{ji} \, e_j$$

$$= \sum_{j=1}^{j=n} \left(\sum_{i=1}^{i=n} a_{ji} \, x_i \right) e_j$$

En identifiant les deux décompositions, suivant la base $(e_1, e_2, \ldots, e_n)$, du vecteur $P(x)$, il vient :

$$x_j' = \sum_{i=1}^{i=n} a_{ji} \, x_i$$

ce que l'on note suivant un tableau carré d'ordre n , ou matrice $n \times n$

$$\begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ a_{21} & a_{22} \cdots & a_2 \\ \vdots & & \\ a_{n1} & a_{n2} & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Ainsi, dans une base déterminée de E , un opérateur linéaire P se représente par une matrice $[P]$

$$[P] = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ a_{21} & a_{22} \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & a_{n2} \cdots & a_{nn} \end{bmatrix}$$

Réciproquement, une matrice $[P]$ étant donnée dans une base $(e_1, e_2, \ldots, e_n)$, cette matrice représente un opérateur linéaire P appartenant à $\mathrm{Hom}(E, E)$, tel que :

$$P(e_i) = \sum_{j=1}^{j=n} a_{ji}\, e_j$$

La structure d'espace vectoriel sur K de $\mathrm{Hom}(E, E)$ lorsque E est de dimension n se transpose sur l'ensemble des matrices $n \times n$

$$\begin{cases} [P] + [P'] = [P + P'] \\ \lambda [P] = [\lambda P] \end{cases}$$

De même, la structure d'algèbre sur K de $\mathrm{Hom}(E, E)$ se transpose sur l'espace vectoriel des matrices $n \times n$. On obtient ainsi l'algèbre des matrices carrées d'ordre n :

$$[P][P'] = [PP']$$

Du point de vue des "tableaux", ceci donne avec :

$$[P] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{et} \quad [P'] = \begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ \vdots & & & \vdots \\ a'_{n1} & a'_{n2} & \cdots & a'_{nn} \end{bmatrix}$$

$$[P + P'] = \begin{bmatrix} a_{11} + a'_{11} & a_{12} + a'_{12} & \cdots & a_{1n} + a'_{1n} \\ \vdots & & & \vdots \\ a_{n1} + a'_{n1} & a_{n2} + a'_{n2} & \cdots & a_{nn} + a'_{nn} \end{bmatrix}$$

$$[\lambda P] = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \vdots & & & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn} \end{bmatrix}$$

et :

$$[PP'] = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

où l'on a posé : $$C_{ij} = \sum_{k=1}^{k=n} a_{ik} \; a'_{kj}$$

A l'application identique I , élément neutre de Hom (E,E), correspond la matrice unité :

$$[I] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix}$$

Changement de base :

Comme nous venons de le voir, une matrice $[P]$ constitue la représentation, relativement à une base donnée, d'un opérateur linéaire P , élément de l'espace Hom(E,E). Si l'on change de base, l'opérateur linéaire P se représente dans la nouvelle base par une nouvelle matrice. Nous nous proposons de calculer les éléments de la nouvelle matrice en fonction des éléments de l'ancienne matrice.

Soit $(e_1 , e_2 \ldots e_n)$ l'ancienne base et $(f_1 , f_2 \ldots f_n)$ la nouvelle base de E . Dans l'ancienne base, P se représente par la matrice :

$$[P] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = [a_{ij}]$$

Dans la nouvelle base, P se représente par la matrice :

$$[P] = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{1n} & \cdots & b_{nn} \end{bmatrix} = [b_{lj}]$$

Nous savons que :

(1) $$P(e_k) = \sum_{j=1}^{j=n} a_{jk} \, e_j$$ et (2) $$P(f_i) = \sum_{l=1}^{l=n} b_{li} \, f_l$$

Posons maintenant la relation de passage d'une base à l'autre :

$$(3) \qquad f_i = \sum_{k=1}^{k=n} \omega_{ki}\, e_k$$

et :

$$(4) \qquad e_j = \sum_{l=1}^{l=n} \rho_{lj}\, f_l$$

Il vient, grâce à (3) :

$$P(f_i) = \sum_{k=1}^{k=n} \omega_{ki}\, P(e_k)$$

On utilise alors (1) pour obtenir :

$$P(f_i) = \sum_{k=1}^{k=n} \omega_{ki} \sum_{j=1}^{j=n} a_{jk}\, e_j$$

Puis on utilise (4) :

$$P(f_i) = \sum_{k=1}^{k=n} \omega_{ki} \sum_{j=1}^{j=n} a_{jk} \sum_{l=1}^{l=n} \rho_{lj}\, f_l = \sum_{\substack{l=1\\ j=1\\ k=1}}^{\substack{l=n\\ j=n\\ k=n}} \rho_{lj}\, a_{jk}\, \omega_{ki}\, f_l$$

comparant avec (2), on déduit grâce à l'écriture unique dans une base

$$\boxed{b_{li} = \sum_{\substack{j=1\\ k=1}}^{\substack{k=n\\ j=n}} \rho_{lj}\, a_{jk}\, \omega_{ki}}$$

Ces formules résolvent le problème, mais on peut mettre ces formules sous une forme plus agréable à manier, grâce aux notations matricielles.

Appelons $[\Omega]$ la matrice dont les éléments sont :

$$[\Omega] = \begin{bmatrix} \omega_{11} & \cdots & \omega_{1n} \\ \omega_{21} & \cdots & \omega_{2n} \\ \vdots & & \vdots \\ \omega_{n1} & \cdots & \omega_{nn} \end{bmatrix}$$

et $[\Omega]^{-1}$ la matrice dont les éléments sont :

$$[\Omega]^{-1} = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1n} \\ \vdots & & \vdots \\ \rho_{n1} & \cdots & \rho_{nn} \end{bmatrix}$$

On a alors, en posant $[A] = [a_{ij}]$ et $[B] = [b_{ij}]$

$$\boxed{[B] = [\Omega]^{-1}[A][\Omega]}$$

Il convient maintenant d'interpréter la signification de $[\Omega]$ et de $[\Omega]^{-1}$

Nous avons posé :

(3) $$f_i = \sum_{k=1}^{k=n} \omega_{ki}\, e_k$$

Un vecteur x se décompose suivant la nouvelle base et suivant l'ancienne base en :

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

$$x = \sum_{i=1}^{i=n} X_i\, f_i$$

et l'on a :

$$x = \sum_{i=1}^{i=n} X_i \sum_{k=1}^{k=n} \omega_{ki} e_k = \sum_{k=1}^{k=n} e_k \left(\sum_{i=1}^{i=n} \omega_{ki} X_i \right)$$

Soit :

(5) $$\boxed{x_k = \sum_{i=1}^{i=n} \omega_{ki} X_i}$$

ce qui se note matriciellement :

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \omega_{11} & \cdots & \omega_{1n} \\ \vdots & & \vdots \\ \omega_{n1} & \cdots & \omega_{nn} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

ou encore :

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [\Omega] \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

Ainsi la matrice Ω est la matrice du changement de base, dont les colonnes représentent la décomposition des nouveaux vecteurs de base dans l'ancienne base. Cette remarque peut être mémorisée.

De même, avec la matrice $[\Omega]^{-1}$, on aura en changeant d'indice dans (4) :

$$(4') \qquad e_i = \sum_{k=1}^{k=n} \rho_{ki} f_k$$

et il vient :

$$x = \sum_{i=1}^{i=n} x_i \sum_{k=1}^{k=n} \rho_{ki} f_k$$

Soit :

$$\boxed{X_k = \sum_{i=1}^{i=n} \rho_{ki} \, x_i} \qquad \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \vdots & & & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

ou encore :

$$\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = [\Omega]^{-1} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

En comparant les formules, on voit que :

$$[\Omega]\,[\Omega]^{-1} = [\Omega]^{-1}[\Omega] = [I]$$

c'est-à-dire que $[\Omega]^{-1}$ est l'inverse de $[\Omega]$ au sens de la multiplication dans l'algèbre des matrices carrées d'ordre n.

Remarque : En algèbre des déterminants, on démontre les formules suivantes relatives à l'inverse d'une matrice $[\Omega]$:

Si $[\Omega] = [\omega_{ij}]$, alors :

$$[\Omega]^{-1} = [\rho_{ij}]$$

avec les relations suivantes concernant les coefficients :

$$\rho_{ij} = (-1)^{i+j}\,\frac{\text{Mineur}(\omega_{ji})}{\det[\Omega]} = \frac{\text{Cofacteur}(\omega_{ji})}{\det[\Omega]}$$

où :

$$\det[\Omega] = \begin{vmatrix} \omega_{11} & \cdots & \omega_{1n} \\ \vdots & & \vdots \\ \omega_{1n} & \cdots & \omega_{nn} \end{vmatrix} \neq 0$$

et :

$$\text{Mineur}(\omega_{ij}) = \begin{vmatrix} \omega_{11} & \cdots & \omega_{1,j-1}, & \omega_{1,j+1} & \cdots & \omega_{1n} \\ \omega_{21} & \cdots & \omega_{2,j-1}, & \omega_{2,j+1} & \cdots & \omega_{2n} \\ \vdots & & & & & \vdots \\ \omega_{i-1,1} & \cdots & \omega_{i-1,j-1}, & \omega_{i-1,j+1} & \cdots & \omega_{i-1,n} \\ \omega_{i+1,1} & \cdots & \omega_{i+1,j-1}, & \omega_{i+1,j+1} & \cdots & \omega_{i+1,n} \\ \vdots & & & & & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,j-1}, & \omega_{n,j+1} & \cdots & \omega_{n,n} \end{vmatrix}$$

C'est le déterminant de la matrice $[\Omega]$ dont on a éliminé la $i^{\text{ième}}$ ligne et la $j^{\text{ième}}$ colonne.

Conclusion : Soit $[A]$ la matrice représentant un opérateur linéaire dans une base $(e_1 \ldots e_n)$ et $[B]$ la matrice représentant cet opérateur linéaire dans une base, $(f_1, f_2, \ldots, f_n)$.
Si l'on appelle $[\Omega]$ la matrice du changement de base, c'est-à-dire la matrice dont la colonne d'ordre j

représente la décomposition du vecteur f_j dans la base $(e_1 \dots e_n)$, on a la formule :

$$\boxed{[B] = [\Omega]^{-1}[A][\Omega]}$$

Deux matrices reliées par la relation précédente sont dites semblables. La relation de similitude est une relation d'équivalence dans l'algèbre des matrices carrées d'ordre n.

Il importe de remarquer que l'algèbre $\mathrm{Hom}\,(E, E)$ n'est pas commutative. Dans le cas des matrices, il est facile de trouver deux matrices qui ne commutent pas :

ainsi les matrices :

$$[A] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{et} \quad [B] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

satisfont : $\quad -[A][B] = +[B][A] = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}$

Le résultat du paragraphe 3.2 se retrouve par la théorie des matrices. Dire que $[A]$ est injective est équivalent à dire que $\det[A] \neq 0$ et par conséquent équivalent à l'existence d'un inverse $[A]^{-1}$. Dans ce cas, $[A]$ est donc associée à un opérateur bijectif de $\mathrm{Hom}\,(E, E)$.

Terminologie :

Plaçons-nous dans une base $(e_1, e_2, \dots, e_n)$ choisie a priori. Dans cette base, la matrice $[P]$ associée à un opérateur linéaire P a pour éléments les scalaires a_{ik}. On associe à cette matrice, diverses autres matrices. On appelle :

Matrice transposée de $[P]$ la matrice ${}^t[P]$ qui, dans la base choisie, a pour éléments a_{ki}.

Matrice conjuguée de $[P]$ la matrice $\overline{[P]}$ qui, dans la base choisie, a pour éléments les conjuguées $\overline{a_{ik}}$ et a_{ik}.

Matrice adjointe de $[P]$ la matrice $[P]^*$ qui, dans la base choisie, a pour éléments les $\overline{a_{ki}}$.

Si, dans une certaine base, deux matrices sont transposées l'une de l'autre (ou conjuguées, ou adjointes), elles ne le sont plus en général après un changement de coordonnées. Il s'agit donc de relations entre matrices, valables dans un certain repère, mais qui ne sont pas des relations entre les opérateurs. Ces notions sont cependant importantes en elles-mêmes, et nous verrons que, si l'on réduit la généralité des changements de base autorisés, il peut arriver que certaines relations entre matrices deviennent alors indépendantes de la base, et équivalentes à des relations entre opérateurs. Mais pour le moment, nous laissons ce point de vue de côté, et nous nous occupons de propriétés formelles des matrices, valables dans une base choisie. Nous y reviendrons au Chapitre 1, paragraphe B, paragraphe 6.

On voit facilement que :

$$[P]^* = {}^t\overline{[P]}$$

l'adjointe est la conjuguée de la transposée.

On vérifie aussi que :

$$\left([P]^*\right)^* = [P] \qquad \text{(propriété involutive)}$$

$${}^t\left([B]\,[A]\right) = {}^t[A]\,{}^t[B] \qquad , \qquad \overline{[B]\,[A]} = \overline{[B]}\;\overline{[A]}$$

$$\left([B]\,[A]\right)^* = [A]^*[B]^*$$

On dit que par rapport à une base, une matrice carrée est :

réelle si elle est égale à sa conjuguée $[P] = \overline{[P]}$

symétrique si elle est égale à sa transposée $[P] = {}^t[P]$

hermitienne si elle est égale à son adjointe $[P] = [P]^*$

On voit immédiatement que si une matrice est

réelle, a_{ik} est réel

symétrique $a_{ik} = a_{ki}$

hermitienne $a_{ik} = \overline{a_{ki}}$

La propriété d'être réelle, hermitienne ou symétrique n'est pas une propriété invariante par changement de base (ce n'est pas une propriété des classes d'équivalence de matrices semblables) et par conséquent ce n'est pas une propriété des opérateurs.Par contre par rapport à certains changements de coordonnées privilégiés, ces propriétés deviennent invariantes.

Etant donné une matrice P quelconque

$[P] + {}^t[P]$ est symétrique

$[P] + [P]^*$ est hermitienne

$[P]\,[P]^*$ est hermitienne.

3.10. Exemple fonctionnel.

Reprenons l'espace vectoriel $\mathbb{C}\,[0,1]$ (§ 2.5) des fonctions continues à valeurs réelles définies sur le segment $[0,1]$. Définissons sur $\mathbb{C}\,[0,1]$ un opérateur P , dit de Volterra :

$$P : \; f \to \int_0^x f(t)\,dt \qquad \text{pour} \quad f \in \mathbb{C}\,[0,1]$$

L'opérateur est visiblement linéaire, à valeurs dans $\mathbb{C}\,[0,1]$

$$P(f_1 + f_2) = \int_0^x \big(f_1(t) + f_2(t)\big)dt = \int_0^x f_1(t)\,dt + \int_0^x f_2(t)\,dt = P(f_1) + P(f_2)$$

P appartient à $\mathrm{Hom}\,\big[\mathbb{C}\,[0,1],\ \mathbb{C}\,[0,1]\big]$

Cherchons le noyau de P .

$\ker P$ est l'ensemble des fonctions f telles que $\int_0^x f(t)\,dt \equiv 0$ pour tout x compris entre 0 et 1. Par dérivation, celà entraîne $f(x) \equiv 0$ et par conséquent $\ker P = 0$. L'opérateur P est injectif.

Cherchons l'image de P.

$\mathrm{Im}\, P$ est l'ensemble des fonctions g de la forme $g(x) = \int_0^x f(t)\,dt$
On note tout de suite deux propriétés des fonctions g :

$$g(0) = 0$$

et la dérivée $g'(x)$ existe et est continue.

Réciproquement une fonction g, nulle à l'origine, dont la dérivée est continue, peut se mettre sous la forme $\int_0^x g'(t)\,dt = g(x)$

Par conséquent, $\mathrm{Im}\, P$ est différent de $\mathbf{C}\,[0,1]$ (opérateur non surjectif) et constitué par l'ensemble des fonctions g définies sur $[0,1]$, nulles à l'origine, dérivables et de dérivée continue.

L'opérateur P', défini sur $\mathrm{Im}\,P \subset \mathbf{C}\,[0,1]$ par :

$$P'g = \frac{dg}{dx}$$

vérifie la relation d'un inverse à gauche :

$$P'P = I$$

Mais P' n'est pas défini sur tout $\mathbf{C}\,[0,1]$ (il existe des fonctions continues non dérivables !). Cet exemple constitue un nouvel exemple, dans le cas de la dimension infinie, d'un opérateur injectif, non surjectif.

4. <u>ESPACE DUAL ALGEBRIQUE</u>.

En 3.6, nous avons défini $\mathrm{Hom}\,(E, F)$ et en particulier $E^* = \mathrm{Hom}(E, K)$ espace vectoriel des formes linéaires définies sur E à valeurs dans K. Cet espace s'appelle le <u>dual algébrique de E</u>. Cet espace joue un grand rôle en algèbre et de même en géométrie élémentaire, où la théorie de la dualité (équation ponctuelle, équation tangentielle) donne une double interprétation de certaines équations.

4.1. Notation.

Soit y^* un élément quelconque de E^*. Nous notons $< x , y^* >$ la valeur de la forme linéaire y^* sur E au point x de E. Cette notation a un sérieux avantage :

- d'une part elle marque la linéarité en x, résultant de la définition de la forme y^*:

$$\langle \lambda x_1 + \mu x_2, y^* \rangle = \lambda \langle x_1, y^* \rangle + \mu \langle x_2, y^* \rangle$$

- d'autre part elle marque la linéarité en y^*, résultant de la définition des opérations sur E^* :

$$\langle x, \lambda y_1^* + \mu y_2^* \rangle = \lambda \langle x, y_1^* \rangle + \mu \langle x, y_2^* \rangle$$

Bien entendu $\langle x , y^* \rangle$ est un scalaire (élément de K).

L'application $x \longrightarrow \langle x, y^* \rangle$ (y^* étant fixé) est la forme liné-aire sur E, appelée y^*.

L'application $y^* \longrightarrow \langle x , y^* \rangle$ (x étant fixé) est une forme liné-aire sur E^*, élément de $\mathrm{Hom}(E^*, E^*) = (E^*)^*$. A chaque x fixé de E on fait ainsi correspondre une forme linéaire $y^* \longrightarrow \langle x , y^* \rangle$ sur E^*.

L'espace E^{**} s'appelle le bidual algébrique de E : c'est un espace vectoriel et ce que nous venons de voir permet d'assurer que E s'identifie à un sous-espace vectoriel de E^{**} puisque tout élément de E définit une forme linéaire sur E^* et que deux éléments distincts de E définissent deux formes linéaires distinctes sur E^* (Cf 4.2 in fine).

En général, on a $E \subset E^{**}$ strictement. Cependant, dans le cas particulier d'un espace vectoriel de dimension finie, E et E^{**} sont canoniquement identifiés, comme nous allons le voir. Auparavant donnons quelques définitions.

Orthogonal d'un sous-espace vectoriel :

Soit E un espace vectoriel sur K et F un sous-espace vectoriel. On appelle orthogonal de F le sous-espace F° de E^{*} constitué par toutes les formes linéaires définies sur E qui s'annulent sur F .

Si $y^{*} \in F^{\circ}$, $\langle x , y^{*} \rangle = 0$ pour tout élément x de F :

- F° constitue un espace vectoriel
- $0^{\circ} = E^{*}$
- $E^{\circ} = 0$

4.2. Espace vectoriel de dimension finie.

Soit E un espace vectoriel de dimension n sur K . A chaque base $B(e_1, e_2, \ldots, e_n)$, nous pouvons associer une base $B^{*}(e_1^{*}, e_2^{*}, \ldots, e_n^{*})$ de l'espace E^{*}. En effet, décomposons un vecteur x suivant la base B:

$$x = \sum_{i=1}^{i=n} x_i \, e_i$$

Soit y^{*} un élément quelconque de $E^{*} = \mathrm{Hom}(E, K)$. Calculons $y^{*}(x) = \langle x, y^{*} \rangle$ (notation de 4.1). Posons auparavant $\langle e_i, y^{*} \rangle = a_i$

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

Considérons alors les n formes linéaires $e_1^{*}, e_2^{*}, \ldots, e_n^{*}$ définies sur E par :

$$\begin{array}{lcl} e_1^{*} : & x \longrightarrow \langle x, e_1^{*} \rangle & = x_1 \\ e_2^{*} : & x \longrightarrow \langle x, e_2^{*} \rangle & = x_2 \\ \vdots & & \\ e_n^{*} : & x \longrightarrow \langle x, e_n^{*} \rangle & = x_n \end{array}$$

Visiblement, $e_1^{*}, \ldots, e_n^{*}$ appartiennent à E^{*} . De plus :

$$(1) \qquad \langle e_j , e^*_i \rangle = \delta_{ij} = \begin{vmatrix} 0 & \text{si } i \neq j \\ & \\ 1 & \text{si } i = j \end{vmatrix} \qquad \text{(symbole de Kronecker)}$$

Par conséquent :

$$\langle x , y^* \rangle = \sum_{i=1}^{i=n} \langle x, e^*_i \rangle \; a_i$$

$$\langle x , y^* \rangle = \langle x, \sum_{i=1}^{i=n} a_i \; e^*_i \rangle$$

c'est-à-dire :

$$(2) \qquad \boxed{y^* = \sum_{i=1}^{i=n} a_i e^*_i}$$

En résumé, tout élément y^* de E^* s'obtient comme combinaison linéaire des n éléments e^*_i de E^*. Les $[e^*_i]$ forment un système de générateurs de E^*.

Montrons que les n formes $[e^*_i]$ sont linéairement indépendantes en tant que vecteurs de E^*. Supposons qu'il existe une relation :

$$\sum_{i=1}^{i=n} \alpha_i e^*_i = 0$$

C'est-à-dire :

$$\sum_{i=1}^{i=n} \alpha_i \langle x, e^*_i \rangle \equiv 0$$

pour tout x dans E.

Prenons en particulier $x = e_j$, il vient $\alpha_j = 0$ pour $j = 1, 2, \ldots, n$ en utilisant la relation de dualité (1), ce qui démontre la linéaire indépendance des $[e^*_i]$. En conclusion, on obtient le théorème :

<u>Théorème 9</u> - ($e^*_1, e^*_2, \ldots, e^*_n$) <u>forment une base de E^*. L'espace E^* est un espace de dimension n lorsque E est de dimension n sur K</u>

$$\boxed{\dim E = \dim E^*}$$

Nous savons (théorème 5, § 3.3) que deux espaces de même dimension (finie) sont isomorphes. Nous avons plus pour E et E^*; on connaît un isomorphisme associé à la donnée d'une base quelconque sur E. La base $(e_1^*, \ldots, e_n^*)$ est dite <u>base conjuguée</u> de la base $(e_1, \ldots, e_n)$.

<u>Remarque</u> : Le noyau de la forme linéaire e_i^* est le sous-espace vectoriel de E^* engendré par $(e_1, e_2, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n)$.

L'orthogonal du sous-espace vectoriel engendré par e_i est le sous-espace vectoriel de E^* engendré par $(e_1^*, e_2^*, \ldots, e_{i-1}^*, e_{i+1}^*, \ldots, e_n^*)$.

Cherchons maintenant à caractériser l'<u>espace bidual $(E^*)^*$</u> lorsque E est de dimension n. A priori, nous savons que :

$$\dim E = \dim E^* = \dim (E^*)^*$$

donc que ces espaces sont isomorphes. On peut préciser l'isomorphisme.

Soit $(e_1, e_2 \ldots e_n)$ une base de E et $(e_1^*, e_2^*, \ldots e_n^*)$ la base conjuguée de $(e_1, e_2 \ldots e_n)$ définie par l'analyse du théorème précédent. La base conjuguée de $(e_1^*, e_2^* \ldots e_n^*)$, pour la dualité entre E^* et $(E^*)^*$, s'identifie à $(e_1, e_2 \ldots e_n)$. En effet, e_i^{**} est la forme linéaire sur E^* faisant correspondre à y^* sa composante dans la base $(e_1^*, e_2^* \ldots e_n^*)$. Cette composante vaut $\langle e_i, y^* \rangle$ et par conséquent :

$$e_i^{**} : \qquad y^* \longrightarrow \langle e_i, y^* \rangle$$

On peut donc identifier e_i et e_i^{**}. Ceci revient à identifier, par un isomorphisme, E et E^{**}.

<u>En fait, cet isomorphisme est canonique</u>, c'est-à-dire indépendant du choix d'une base dans E. D'une part, nous savons que $E \subset (E^*)^*$ d'après ce qui a été vu en 4.1 et cette injection est canonique.

Explicitons cette injection (canonique) L de E dans $(E^*)^*$. Par définition, L est la correspondance entre E et $(E^*)^*$ qui à x dans E fait correspondre Lx dans $(E^*)^*$:

$$\begin{array}{ccc} x & \longrightarrow & Lx \\ E & \longrightarrow & (E^*)^* \end{array}$$

où Lx est la forme linéaire définie sur E^* par :

$$\langle y^*, Lx \rangle = \langle x, y^* \rangle = y^*(x)$$

Visiblement la correspondance $L : x \longrightarrow Lx$ est linéaire et injective. Comme E et $(E^*)^*$ ont la même dimension (finie), cela prouve que L est une bijection et donc un isomorphisme entre E et $(E^*)^*$.

Théorème - E et $(E^*)^*$ sont canoniquement isomorphes lorsque E est de dimension finie. On les identifiera désormais.

4.3. Applications transposées.

Soient E et F deux espaces vectoriels sur un même corps K . Soit P un opérateur linéaire de E dans F , élément de $\mathrm{Hom}(E,F)$. Envisageons l'application ${}^tP$, définie sur F^*, qui à tout élément y^* de F^* fait correspondre l'élément de E^* : $y^* o P = {}^tP(y^*)$

$$y^* \longrightarrow {}^tP(y^*)$$

$$\begin{tikzcd} E \arrow[r, "P"] \arrow[dr, "{}^tP(y^*)"'] & F \arrow[d, "y^*"] \\ & K \end{tikzcd}$$

$y^* o P$ est un élément de E^* dont la valeur en $x \in E$ est :

$$(y^* o P)(x) = y^*\big(P(x)\big) = \langle P(x), y^* \rangle$$

(1) $$\boxed{\langle P(x), y^* \rangle = \langle x, {}^tP(y^*) \rangle}$$

L'application ${}^tP$ (prononcer transposé de l'opérateur P) est une application linéaire de F^* dans E^*

$$E \xrightarrow{\ P\ } F$$

$$F^* \xrightarrow{\ {}^tP\ } E^*$$

Donnons quelques propriétés du transposé d'un opérateur P .

Propriété 1 $$\boxed{\ker({}^tP) = (\mathrm{Im}\, P)^\circ}$$

Soit $y^* \in \ker({}^tP)$ et soit un élément x de E. Par définition, on a :

$\langle x, {}^tP(y^*) \rangle = 0$ donc $\langle P(x), y^* \rangle = 0$ d'après (1) et réciproquement, ce qui prouve bien l'égalité cherchée. On démontrerait de même :

Propriété 2 $\boxed{\mathrm{Im}({}^tP) = (\ker P)^\circ}$

4.4. Interprétation matricielle.

Soit E un espace vectoriel de dimension n et P un élément de $\mathrm{Hom}(E, E)$, représenté par une matrice $[P] = [a_{ij}]$ dans une base $(e_1, e_2 \dots e_n)$ de E.

L'opérateur transposé ${}^tP$ est représenté par la matrice transposée ${}^t[P]$ dans la base conjuguée $(e_1^*, e_2^* \dots e_n^*)$ de E^*. Cela revient à dire que $[{}^tP] = {}^t[P]$ d'après les notations de 4.3 et celles de 3.9 (in fine).

Pour démontrer cette assertion, nous allons utiliser la formule de définition de l'opérateur transposé d'un opérateur linéaire.

Posons :

(1) $$x = \sum_{i=1}^{i=n} x_i e_i \qquad \text{et} \qquad [P] = [a_{ij}]$$

(2) $$P(x) = \sum_{i=1}^{i=n} \big(P(x)\big)_i e_i = \sum_{i=1}^{i=n} \Big(\sum_{j=1}^{j=n} a_{ij} x_j\Big) e_i = \sum_{\substack{i=1\\ j=1}}^{\substack{j=n\\ i=n}} a_{ij} x_j e_i$$

Posons également : $$y^* = \sum_{j=1}^{j=n} \alpha_j e_j^* \qquad \text{et} \quad [{}^tP] = [b_{ij}]$$

(décomposition de y^* ($\in E^*$) dans la base conjuguée $(e_1^* \dots e_n^*)$ et représentation dans la base $(e_1^* \dots e_n^*)$ par une matrice $[b_{ij}]$ de l'opérateur ${}^tP$) :

$$ {}^tP[y^*] = \sum_{j=1}^{j=n} \big({}^tP(y^*)\big)_j e_j^* = \sum_{j=1}^{j=n} e_j^* \Big(\sum_{i=1}^{i=n} b_{ji} \alpha_i\Big) = \sum_{\substack{i=1\\ j=1}}^{\substack{i=n\\ j=n}} b_{ji} \alpha_i e_j^*$$

Grâce aux formules de dualité, on obtient :

$$\left\langle P(x), y^* \right\rangle = \left\langle \sum_{\substack{i=1\\j=1}}^{\substack{i=n\\j=n}} a_{ij} x_j e_i \,,\ \sum_{k=1}^{k=n} \alpha_k \, e_k^* \right\rangle$$

$$= \sum_{i,j,k} a_{ij} x_j \, \alpha_k \left\langle e_i , e_k^* \right\rangle$$

Mais : $\left\langle e_i , e_k^* \right\rangle = \delta_{ik} = \begin{cases} 1 & \text{si } i = k \\ 0 & \text{si } i \neq k \end{cases}$

D'où :

(3) $$\left\langle P(x), y^* \right\rangle = \sum_{i,j} a_{ij} x_j \, \alpha_i$$

De même :

$$\left\langle x, {}^tP(y^*) \right\rangle = \left\langle \sum_{k=1}^{k=n} x_k e_k \,, \sum_{\substack{i=1\\j=1}}^{\substack{i=n\\j=n}} b_{ji} \alpha_i \, e_j^* \right\rangle$$

$$= \sum_{i,j,k} x_k \, b_{ji} \, \alpha_i \left\langle e_k , e_j^* \right\rangle$$

Soit :

(4) $$\left\langle x, {}^tP(y^*) \right\rangle = \sum_{i,j} b_{ji} x_j \, \alpha_i$$

En comparant les deux expressions (3) et (4) :

$$\sum_{i,j} a_{ij} x_j \alpha_i = \sum_{i,j} b_{ji} x_j \alpha_i$$

on obtient aussitôt :

$$b_{ji} = a_{ij} \quad \text{c'est-à-dire que} \quad \boxed{[{}^tP] = {}^t[P]}$$

<u>L'opérateur transposé de P est représenté dans la base conjuguée ($e_1^*, \dots e_n^*$) de E^* par la matrice transposée de celle qui représente l'opérateur P dans la base ($e_1, \dots, e_n$) de E</u> .

Remarque 1 : Si $[P] = [a_{ij}]$ dans la base $(e_1, e_2, \dots, e_n)$ et $[{}^tP] = [b_{ij}]$ dans la base $(e_1^*, e_2^*, \dots, e_n^*)$, on vérifie que :

$$a_{ij} = \langle P(e_j), e_i^* \rangle$$

De même :

$$b_{ij} = \langle e_i, {}^tP(e_j^*) \rangle$$

Par conséquent :

$$b_{ij} = \langle e_i, {}^tP(e_j^*) \rangle = \langle P(e_i), e_j^* \rangle = a_{ji}$$

ce qui démontre le résultat de 4.4 d'une façon plus élégante.

Remarque 2 : On peut donner les formules de changement de coordonnées dans E^*, correspondant à un changement de coordonnées dans E . On démontre que ce sont les formules de changement de vecteurs de base dans E . En fait, lorsque l'on effectue un changement de base dans E , les formules qui font passer des nouvelles coordonnées aux anciennes coordonnées d'un même vecteur x , sont obtenues en écrivant l'invariance du vecteur x :

(a) $$x = \sum_{i=1}^{i=n} x_i e_i = \sum_{i=1}^{i=n} x'_i e'_i$$

Un changement de base dans E implique un changement des bases conjuguées dans E^*. Pour calculer les nouvelles coordonnées d'un vecteur y^* de E^* , on écrit alors l'invariance du vecteur y^* , comme forme linéaire sur E :

(b) $$\langle x, y^* \rangle = \sum_{i=1}^{i=n} x_i \alpha_i = \sum_{i=1}^{i=n} x'_i \alpha'_i$$

où

$$x = \sum_{i=1}^{i=n} x_i e_i$$

et :

$$y^* = \sum_{i=1}^{i=n} \alpha_i e_i^*$$

Formellement, les calculs déduits de (a) ou de (b) sont identiques, c'est-à-dire que les nouvelles coordonnées x'_i du vecteur y^* s'obtiennent en fonction des anciennes coordonnées x_i , de la même façon que e'_i (vecteur de la nouvelle base) se déduit des e_i (ancienne base).

EXERCICES SUR LE CHAPITRE 0

EXERCICE N° 1 :

Soient E et F deux expaces vectoriels sur le même corps K , de dimension m et n respectivement. Quelle est la dimension de l'espace vectoriel $Hom(E, F)$? (Réponse : $m.n$).

En particulier, quelle est la dimension de l'espace vectoriel E^* ?

EXERCICE N° 2 :

Soient E et F deux espaces vectoriels de même dimension finie n sur le corps K. Soit P un opérateur linéaire de E dans F ; démontrer l'équivalence des propriétés suivantes :

1) P admet un inverse à gauche
2) P est injective
3) P est surjective
4) L'image de toute base de E est une base de F
5) Si $(P(x_1), P(x_2), \dots, P(x_n))$ est une base de F , alors $(x_1 , x_2 , \dots, x_n)$ est une base de E .

EXERCICE N° 3 :

Soient P_1 et P_2 deux opérateurs linéaires sur un espace vectoriel E. Nous supposons que $(I - P_1 P_2)$ admet un inverse noté $(I - P_1 P_2)^{-1}$

Première question : supposons que E soit de dimension finie :

a) en utilisant au besoin l'exercice n° 2, montrer que $(I - P_2 P_1)$ admet un inverse,

b) démontrer les formules suivantes :

$$(I - P_1 P_2)^{-1} = I + P_1 (I - P_2 P_1)^{-1} P_2$$

$$(I - P_2 P_1)^{-1} = I + P_2 (I - P_1 P_2)^{-1} P_1$$

Deuxième question : se servant de b), faut-il faire une hypothèse supplémentaire pour assurer l'existence d'un inverse pour $(I - P_2 P_1)$ dans le cas où E n'est pas de dimension finie ?

EXERCICE N° 4 :

On considère l'espace vectoriel E des suites de nombres réels (Cf $\mathbb{R}^{\mathbb{N}}$, § 3.8) et F le sous-espace vectoriel des suites de nombres réels telles que seul un nombre fini de termes soient différents de zéro. Caractériser l'espace dual algébrique du sous-espace vectoriel F .

EXERCICE N° 5 :

On considère l'ensemble E des fonctions réelles f d'une variable réelle, définie sur $\mathbb{R}$, indéfiniment dérivables, et telles que pour tous les entiers p et q , les intégrales :

$$\int_{-\infty}^{+\infty} |x|^p \left| \frac{d^q f}{d x^q}(x) \right| d x$$

soient convergentes.

Première question : montrer que E est un espace vectoriel de dimension infinie. Montrer que si $f \in E$, f tend vers zéro lorsque x tend vers l'infini.

Deuxième question : à toute fonction f de E , on fait correspondre la fonction g :

$$g : \quad x \longrightarrow g(x) = s(x)\, f(x) + m \frac{df}{dx}(x)$$

où m une constante donnée et s une fonction donnée, indéfiniment dérivable telle que pour tout n entier $\geqslant 0$, il existe un entier k, un nombre réel M et un nombre positif X_0 tels que pour tout x, $|x| \geqslant X_0$ on ait

$$\left|\frac{d^n s}{d x^n}(x)\right| \leqslant M \left| x^k \right|$$

a) Montrer que A est un opérateur linéaire défini sur E.

b) ψ étant donnée dans E, chercher la fonction ξ de E, dépendant uniquement de ψ, telle que :

$$\int_{-\infty}^{+\infty} A\varphi(x)\psi(x)\, dx = \int_{-\infty}^{+\infty} \varphi(x)\xi(x)\, dx$$

pour toute fonction φ appartenant à E.

c) Montrer que la correspondance $\psi \longrightarrow \xi$ définit un opérateur linéaire sur E, noté A^*.

d) Déterminer une fonction s, nulle en 0, de sorte que l'on ait :

$$A A^* - A^* A = I$$

ou I désigne l'opérateur unité de $\mathrm{Hom}(E, E)$.

e) Montrer que les valeurs propres de $A^* A$ sont non négatives.

f) En outre, si $A A^* - A^* A = I$ alors si λ est valeur propre $\neq 1$ de $A^* A$, $(\lambda - 1)$ et $(\lambda + 1)$ sont encore des valeurs propres de $A^* A$.

EXERCICE N° 6 :

E_{n+1} désigne l'ensemble formé par les polynômes à coefficients réels $P(x)$ dont le degré est inférieur ou égal à n :

$$P(x) = \sum_{0}^{n} a_k x^k = a_0 + a_1 x + \dots + a_n x^n$$

1°) montrer que E_{n+1} est un espace vectoriel sur le corps des réels de dimension $(n+1)$, dont une base est $\left\{1, x, x^2, \dots, x^n\right\}$

2°) Donner la dimension du sous-espace vectoriel formé par les polynômes P multiples d'un polynôme donné Q (le degré de Q est supposé égal à k avec $k \leqslant n$).

3°) Par x_0^* , on désigne l'application qui fait correspondre à tout élément P de E_{n+1} le nombre réel $P(x_0)$.[x_0 est un nombre réel] .

Cette application est-elle linéaire sur E_{n+1} ?

4°) En déduire une base de l'espace vectoriel E^*_{n+1} , espace dual de l'espace vectoriel E_{n+1} .

(on pourra montrer que si ($x_0, x_1, \dots, x_n$) sont ($n+1$) nombres réels distincts , (x_0^* , $x_1^* \dots x_n^*$) sont ($n+1$) formes linéaires indépendantes de E^*_{n+1}).

5°) En déduire qu'il existe ($n+1$) constantes ($\alpha_0 \dots \alpha_n$) telles que pour tout élément $P(x)$ de E_{n+1} , on ait :

$$\int_0^1 P(x)dx = \alpha_0 P(x_0) + \text{-------} + \alpha_n P(x_n)$$

Déterminer ces constantes.

EXERCICE N° 7.

Soit E un espace vectoriel de dimension finie sur le corps K et soit F un sous-espace vectoriel de E . Démontrer la relation suivante :

$$\dim F + \dim F^\circ = \dim E$$

EXERCICE N° 8.

Soit E un espace vectoriel de dimension finie sur K . Démontrer les deux théorèmes suivants : (théorèmes de Rouché)

1°) Un opérateur linéaire P et son transposé ${}^tP$ ont même rang (Cf § 3.4).

2°) Pour qu'une équation linéaire $P(x) = y_0$ soit résoluble (P opérateur linéaire, y_0 vecteur fixe), il faut et il suffit que y_0 soit orthogonal à toutes les solutions de l'équation homogène transposée :

$$^{t}P(x^{*}) = 0$$

EXERCICE N° 9.

lère partie :

On considère la fonction de deux variables réelles $F(x,y)$, définie selon

$$F(x,y) = \mathrm{Log}[(1-x)(1+y)] \text{ pour } -1 < x \leqslant y < +1$$

et $F(x,y) = F(y,x)$ pour tous x , y réels qui satisfont les inégalités

$$-1 < x < +1 \quad ; \quad -1 < y < +1$$

1°) Pour tout entier $n \geqslant 0$, calculer l'intégrale

$$f_n(y) = \int_{-1}^{+1} F(x,y)\, x^n \, dx$$

à partir de la fonction $g_{n+1}(y) = \int_{-1}^{y} \frac{x^{n+1}}{1-x} dx$ laquelle est définie par $-1 \leqslant y < +1$

2°) Calculer $g_{n+1}(y)$

3°) En déduire que f_n est un polynôme de degré n

4°) Calculer explicitement f_0 , f_1 , f_2 et f_3 en ordonnant ces polynômes par puissances croissantes.

2ème partie :

On pose $G(x,y) = -\frac{1}{2} F(x,y) + \left(\mathrm{Log}\, 2 - \frac{1}{2}\right)$

où F désigne la fonction déjà définie en I. On désigne par E l'espace vectoriel formé par les polynômes à coefficients réels, de degré 3 au plus. A tout polynôme P de E, on fait correspondre la fonction Q définie sur $]-1,+1[$ selon

$$Q(y) = \int_{-1}^{+1} G(x,y)\ P(x)\, dx$$

1°) Etablir que Q est la restriction, à $]-1,+1[$, d'un polynôme encore noté Q. Montrer que l'application $\Phi : P \longrightarrow Q$ est une application linéaire de E dans E.

2°) Soit T_i le polynôme $x \longrightarrow T_i(x) = x^i$ où $i = 0, 1, 2$ ou 3. La famille des $\{T_i\}$, i variant de 0 à 3, forme une base de E. Quelle est la matrice $[\Phi]$ représentant l'application Φ dans cette base de E ?

3°) Quelles sont les valeurs propres et les vecteurs propres de la matrice $[\Phi]$? Donner une base de vecteurs propres dont tout élément vérifie

$$P(1) = 1$$

3ème partie :

Soit λ un paramètre réel quelconque. On désigne par $\mathbb{J}_\lambda$ l'équation différentielle linéaire et du second ordre

$$(1-y^2)\frac{d^2f}{dy^2}(y) - 2y\frac{df}{dy}(y) + \lambda f(y) = 0$$

1°) Trouver le polynôme P, de degré 3 au plus, satisfaisant $\mathbb{J}_\lambda$ avec $\lambda = 12$ et vérifiant $P(1) = 1$

On admet dorénavant que toute fonction continue f sur $[-1,+1]$ et satisfaisant

$$f(y) = \lambda \int_{-1}^{+1} G(x,y)\ f(x)\, dx$$

satisfait aussi l'équation $\mathbb{J}_\lambda$ pour $\lambda \neq 0$

2°) Montrer qu'une condition nécessaire et suffisante pour que l'équation $\mathbb{J}_\lambda$ admette pour solution un polynôme est que $\lambda = n(n+1)$ où n désigne un entier positif ou nul.

3°) On appelle P un polynôme non identiquement nul satisfaisant $\mathbb{J}_\lambda$ pour $\lambda = n(n+1)$. Montrer que $P(1)$ est différent de zéro. En déduire qu'il existe un unique polynôme, noté P_n , de degré n , satisfaisant $1 = P_n(1)$ et l'équation $\mathbb{J}_\lambda$ pour $\lambda = n(n+1)$.

4°) Montrer que $\displaystyle\int_{-1}^{+1} P_n(x) P_m(x)\, dx = 0$ pour $n \neq m$.

4ème partie :

Soit f une fonction satisfaisant l'équation $\mathbb{J}_\lambda$ pour $\lambda = n_0(n_0+1)$ où n_0 est un entier positif ou nul, et pour tout x appartenant à $]a, b[$ où $a < -1$ et $b > 1$.

1°) Montrer que $\displaystyle\int_{-1}^{+1} f(x)\, P_n(x)\, dx = 0$ pour tout entier $n \neq n_0$.

2°) Montrer qu'il existe une constante α telle que $g(x) = f(x) - \alpha P_{n_0}(x)$ satisfasse

$$\int_{-1}^{+1} g(x)\, P_n(x)\, dx = 0 \quad \text{pour tout } n \geqslant 0$$

3°) Montrer que $\displaystyle\int_{-1}^{+1} g(x)\ P(x)\, dx = 0$ pour tout polynôme P de degré quelconque.

4°) Montrer que $g(x) = 0$ sur $[-1, +1]$ et donc que f est un multiple de P_{n_0}. Conclure.

(on raisonne par l'absurde en supposant qu'il existe un point x_0 de $[-1, +1]$ tel que $g(x) \neq 0$. Par homothétie, on peut se ramener à supposer $g(x_0) = a > 0$. Par continuité, il existe $1 > \delta > 0$ tel que $g(x) > \frac{a}{2}$ pour tout x satisfaisant $|x - x_0| \leqslant \delta$. On considère alors $P(x) = 1 - \frac{1}{2}(x^2 - \delta^2)$ de sorte que $P(x) \geqslant 1$ pour $|x| \leqslant \delta$ et $|P(x)| \leqslant 1$ pour $\delta \leqslant |x| \leqslant 2$. On décompose

$$\int_{-1}^{+1} g(x) \big(P(x - x_0)\big)^n dx$$

en deux intégrales, l'une restant bornée et l'autre tendant vers l'infini avec n . Il y a contradiction).

EXERCICE N° 10.

Soit $f(x) = 3x + 2$ et $g(x) = x^2 + x - 2$ deux polynômes à coefficients réels. On désigne par $Q_n(x) = a_0 + \cdots + a_n x^n$ le quotient de la division suivant les puissances croissantes à l'ordre n de $f(x)$ par $g(x)$.

1°) Calculer a_0 et a_1 et montrer que pour tout $n \geqslant 2$, on a

$$a_n = \frac{a_{n-1} + a_{n-2}}{2}$$

2°) La matrice $\begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix}$ admet deux valeurs propres λ_1, λ_2 $(\lambda_1 < \lambda_2)$. Soit e_1 (resp^t. e_2) le vecteur propre associé à λ_1(resp^t. λ_2) et dont la première composante dans la base canonique est 1. Calculer e_1 et e_2.

3°) Calculer A^p et en déduire a_n en fonction de n.

- CHAPITRE 1 -

ESPACES VECTORIELS NORMES

L'utilisation fréquente de l'expression "quel que soit un nombre aussi voisin de zéro qu'on le désire" évoque une notion de voisinage fort utile en analyse classique. Nous nous proposons de généraliser cette notion de voisinage à d'autres espaces que l'espace des nombres réels ou l'espace des nombres complexes. Tout d'abord, nous allons définir ces espaces et étudier leurs propriétés algébriques. Dans un deuxième temps, nous envisagerons des propriétés topologiques.

A. CONSIDERATIONS ALGEBRIQUES.

1. DISTANCES.

Donnons, a priori, des définitions que nous commenterons ensuite : soit E un ensemble d'éléments x, y, z, ... Une application d qui, à deux éléments quelconques x et y de E fait correspondre un nombre réel $d(x,y)$, définit une distance lorsque les quatre propriétés suivantes sont vérifiées :

1- $d(x,y) \geqslant 0$ pour tout couple d'éléments de E.

2- $d(x,y) = 0$ si, et seulement si, $x = y$.

3- $d(x,y) = d(y,x)$: relation de symétrie.

4- $d(x,z) \leqslant d(x,y) + d(y,z)$ pour trois éléments quelconques de E.

Cette dernière inégalité porte le nom d'inégalité triangulaire.

Remarque : Lorsque la propriété (2) n'a pas lieu, l'expression $d(x,y)$ définit un écart sur l'ensemble E.

2. ESPACES METRIQUES.

La donnée d'un ensemble E et d'une distance d définit un espace métrique, noté (E, d). Très souvent d'ailleurs, on se contentera de noter l'espace métrique par E. Les éléments de E sont appelés des points. Donnons quelques exemples familiers pour illustrer cette définition.

Exemples :

2.1. La droite réelle.

Prenons pour ensemble E l'ensemble des nombres réels R et prenons $d(x, y)$ égale à $|x-y|$ pour un couple quelconque (x, y) de nombres réels. Les quatre propriétés d'une distance sont bien vérifiées grâce aux propriétés connues de la valeur absolue. L'espace métrique ainsi constitué s'écrit $\mathbb{R}$ et se prononce droite réelle.

2.2. L'espace $\mathbb{R}^3$.

Prenons pour ensemble E l'espace ordinaire de la géométrie à trois dimensions, rapporté à trois axes de coordonnées orthonormés. Un point M de E est représenté par trois coordonnées x, y, z. On définit en particulier trois distances distinctes :

$$d_1(M,M') = \mathrm{Sup}(|x-x'|, |y-y'|, |z-z'|)$$

$$d_2(M,M') = |x-x'| + |y-y'| + |z-z'|$$

$$d_3(M,M') = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$$

On reconnaît en d_3 la distance euclidienne usuelle, et dans ce dernier cas la propriété (4) est la traduction analytique de l'assertion bien connue : dans un triangle, la somme des longueurs de deux côtés est supérieure ou égale à la longueur du troisième côté, d'où le nom d'inégalité triangulaire réservé à l'inégalité (4).

Ainsi, $(\mathbb{R}^3, d_1)$, $(\mathbb{R}^3, d_2)$, $(\mathbb{R}^3, d_3)$ sont trois espaces métriques distincts.

Remarque : En géométrie ordinaire, la différence des longueurs de deux côtés d'un triangle est toujours inférieure ou égale à la longueur du troisième côté. Plus généralement, dans un espace métrique quelconque, l'inégalité suivante est encore valable, comme il est facile de le vérifier :

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

2.3. Un exemple fonctionnel.

Un fonction f prenant des valeurs réelles, et définie sur le segment $[0,1]$ $(0 \leqslant x \leqslant 1)$, est bornée lorsqu'il existe un nombre M tel que $|f(x)| \leqslant M$ pour tout x de $[0\ 1]$. Le plus petit nombre M satisfaisant l'inégalité précédente pour tout x appartenant au segment $[0,1]$ est la borne supérieure de la fonction $|f|$ sur ce segment. On la note :

$$\sup_{x \in [0,1]} |f(x)| \qquad \text{ou encore } \|f\|_\infty$$

Soit E l'ensemble constitué par les fonctions f, g, ... définies sur le segment $[0,1]$, prenant des valeurs réelles, et bornées (bien entendu, la borne M de chaque élément f dépend de cet élément f). L'expression :

$$d(f,g) = \sup_{x \in [0,1]} |f(x) - g(x)|$$

a un sens et définit une distance sur l'ensemble E.

Les propriétés (1), (2) et (3) proviennent, en effet, de résultats élémentaires sur les bornes supérieures. De plus, à partir de l'inégalité :

$$|f(x) - h(x)| \leqslant |f(x) - g(x)| + |g(x) - h(x)|$$

et en passant aux bornes supérieures, on obtient :

$$(4) \qquad d(f,h) \leqslant d(f,g) + d(g,h)$$

3. SOUS-ESPACE METRIQUE.

Soit (E, d) un espace métrique et F un sous-ensemble de E. Restreinte à l'ensemble F, d définit encore une distance et donc une métrique sur F. Muni de cette métrique, F est un sous-espace (métrique) de l'espace métrique E. Dans ce qui suit, sauf mention contraire, tout sous-ensemble d'un espace métrique sera considéré comme sous-espace métrique.

4. ESPACES VECTORIELS NORMES.

4.1. Métrique sur un espace vectoriel.

Nous voulons construire, sur un espace vectoriel E, une distance $d(x,y)$,

- d'une part invariante par translation (ce qui lie la structure additive et la métrique) ; c'est-à-dire vérifiant :

(5) $d(x+z, y+z) = d(x,y)$ pour tout vecteur z de E ;

- d'autre part, positivement homogène pour la multiplication scalaire :

(6) $d(\lambda x, \lambda y) = |\lambda| d(x,y)$ pour tout élément λ de K (ce qui lie la structure multiplicative extérieure à la métrique).

Rappelons (Chapitre 0) que K est le corps de base de l'espace vectoriel E. C'est soit le corps des réels, soit le corps des complexes.

En posant $z = -x$, la propriété (5) permet d'écrire :

$$d(0,y-x) = d(x,y)$$

Quelles sont les propriétés de la fonction $d(0,x)$, notée $\|x\|$ et considérée comme fonction numérique de x ? Nous allons déduire ses caractéristiques des propriétés (1), (2), (3), (4) d'une distance (Cf. § A.1) et des propriétés (a) et (b) d'un espace vectoriel sur K (Cf. Chap. 0, § 2.1).
Il vient aisément :

(1') $\|x\|$ est positive ou nulle

(2') $\|x\| = 0$ si, et seulement si, $x = 0$

(3') $\|\lambda x\| = |\lambda| \|x\|$ pour tout λ de K

(4') $\|x + y\| \leqslant \|x\| + \|y\|$

(1'), (2'), (3') sont faciles à démontrer . Quant à (4'), majorons :

$$d(0, x+y) = \|x+y\|$$

$$d(0, x+y) = d(-x, y)$$

Donc, d'après l'inégalité triangulaire (§ A.1, propriété 4), on a la majoration :

$$d(-x, y) \leqslant d(-x, 0) + d(0, y)$$

Or :

$$d(-x, 0) = d(0, -x) \qquad \text{(§ A.1, propriété 3)}$$

et :

$$d(0, -x) = d(0, x) \qquad \text{(d'après (5) en faisant } \lambda = -1 \text{).}$$

(4') est alors acquise :

$$d(0, x+y) \leqslant d(0, x) + d(0, y)$$

soit :

$$\|x+y\| \leqslant \|x\| + \|y\|$$

Dès lors, il devient tout naturel de poser la définition suivante.

4.2. Norme.

On appelle <u>norme sur un espace vectoriel E</u>, une application numérique $x \longrightarrow \|x\|$, définie sur E et vérifiant les quatre propriétés

(1'), (2'), (3') et (4').

Une norme étant donnée sur un espace vectoriel E , on peut lui associer une distance par :

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

On vérifie aisément pour cette distance les propriétés supplémentaires (5) et (6) que nous avions posées comme base de notre recherche.

4.3. Espaces vectoriels normés.

4.3.1. définition.

Un espace vectoriel normé est la donnée d'un espace vectoriel sur K et d'une norme.

Un espace vectoriel normé est, en particulier, un espace métrique. On note $(E, \| . \|)$ un espace vectoriel normé.

On convient d'appeler vecteurs les éléments d'un espace vectoriel normé lorsque la structure vectorielle joue un rôle primordial, points les éléments d'un espace vectoriel normé lorsque la structure affine importe le plus. Dans ce dernier cas, l'élément neutre 0 de l'addition est fréquemment appelé l'origine de l'espace vectoriel E .

L'exemple 2.1 est en fait le cas d'une distance provenant d'une norme (communément appelée valeur absolue) sur l'espace vectoriel.

4.3.2. exemple fonctionnel.

De même, reprenons l'exemple 2.3. Sur l'ensemble $\mathcal{C}[0,1]$ des fonctions numériques continues sur le segment $[0,1]$, nous avons construit une distance et obtenu l'espace métrique $\mathcal{C}[0,1]$ comme sous-espace métrique de $\mathcal{F}[0,1]$ muni de la norme $\| \cdot \|_\infty$.

Pour deux éléments f et g de $\mathcal{C}[0,1]$, on peut définir :

$f+g$ comme la fonction $x \longrightarrow f(x)+g(x)$

λf comme la fonction $x \longrightarrow \lambda f(x)$ où λ désigne un scalaire réel.

Muni de ces deux lois, à savoir l'addition et la multiplication par un scalaire, $\mathscr{C}[0,1]$ constitue un espace vectoriel sur le corps des nombres réels, comme nous l'avons déjà constaté au Chapitre 0, § 2.3.

Montrons que $\mathscr{C}[0,1]$ est un espace vectoriel normé. Visiblement la distance :

$$d(f,g) = \operatorname*{Sup}_{x \in [0,1]} |f(x) - g(x)|$$

introduite en 2.3, est invariante par translation et satisfait (6) :

$$d(f+h, g+h) = d(f,g) \quad \text{pour tout } h \text{ de } \mathscr{C}[0,1]$$

et

$$d(\lambda f, \lambda g) = |\lambda| d(f,g) \quad \text{pour tout } \lambda \text{ de } \mathbb{R}$$

La normé associée s'écrit :

$$\| f \| = d(0,f) = \operatorname*{Sup}_{x \in [0,1]} | f(x)|$$

et s'appelle norme de la convergence uniformé.

4.4. Espace préhilbertien.

En analyse fonctionnelle, on rencontre très fréquemment des espaces vectoriels normés sur lesquels on peut définir des notions d'orthogonalité et disposer du théorème de Pythagore. De plus, ces espaces apparaissent comme la généralisation la plus naturelle des espaces euclidiens de dimension finie, ce qui permet de géométriser l'analyse. Nous procédons de façon didactique en donnant directement des définitions.

4.4.1. produit scalaire.

Soit E un espace vectoriel sur K (K désigne soit $\mathbb{R}$, soit $\mathbb{C}$). Un produit scalaire sur un espace vectoriel E est une application qui, à tout couple d'élément x, y de E , associe un élément de K , noté $\langle x, y\rangle$, de telle sorte que l'on ait :

(a) $\quad \langle x,y\rangle = \overline{\langle y,x\rangle}$ $\quad$ symétrie hermitienne (par la barre de conjugaison, on désigne l'imaginaire conjugué de $<y,x>$ lorsque le corps des scalaires est $\mathbb{C}$ et $<y,x>$ lui même, lorsque le corps des scalaires est $\mathbb{R}$).

(b) $$<x+z,y> = <x,y> + <z,y>$$

et $<\lambda x,y> = \lambda <x,y>$ pour tout λ du corps K (linéarité du produit scalaire en x)

(c) $<x,x> \geqslant 0$ positivité du produit scalaire

(d) $<x,x> = 0$ si, et seulement si $x = 0$

Remarque : grâce à la linéarité de $<x,y>$ en x , pour y fixé, (a) entraîne l'antilinéarité de $<x,y>$ en y , pour x fixé, c'est-à-dire :

$$<x,y+z> = <x,y> + <x,z> \quad \text{et} \quad <x,\lambda y> = \overline{\lambda}<x,y>$$

Au lieu de produit scalaire, on parle dans certains textes, plus explicitement, de forme bilinéaire hermitienne positive non dégénérée. Le mot hermitien vient de la propriété (a), les mots forme bilinéaire de la propriété (b) le mot positif de la propriété (c) et les mots non dégénérés de la propriété (d).

4.4.2. inégalité de Schwarz.

Un produit scalaire étant donné, on peut obtenir l'importante inégalité suivante, dite inégalité de Schwarz :

$$\boxed{|<x,y>|^2 \leqslant <x,x><y,y>}$$

L'égalité n'étant possible que s'il existe un λ_0 de K tel que :

$$x = \lambda_0 y \quad (\text{lorsque } y \neq 0\ ).$$

Par hypothèse, $<x,x>$ est non négatif. Donc, quel que soit le scalaire λ :

$$<x+\lambda y, x+\lambda y> \geqslant 0$$

on obtient en développant et en supposant que $K = \mathbb{C}$ pour fixer les idées :

$$<x+\lambda y, x+\lambda y> = <x,x> + \lambda \overline{<x,y>} + \bar{\lambda} <x,y> + \lambda\bar{\lambda} <y,y>$$

Ce trinôme se décompose, en supposant $<y,y> \neq 0$, sous la forme :

$$<y,y>\left(\lambda + \frac{<x,y>}{<y,y>}\right)\left(\bar{\lambda} + \frac{\overline{<x,y>}}{<y,y>}\right) + <x,x> - \frac{|<x,y>|^2}{<y,y>}$$

Il reste toujours positif et cela en particulier lorsque λ prend une valeur qui annule le premier terme . Il reste :

$$<x,x> - \frac{|<x,y>|^2}{<y,y>} \geqslant 0$$

ce qui est bien l'inégalité cherchée.

Cette inégalité devient une égalité s'il existe un $\lambda_0 (\in \mathbb{C})$ particulier tel que :

$$<x+\lambda_0 y, x+\lambda_0 y> = 0$$

. Grâce à la propriété (d) du produit scalaire, il existe entre les vecteurs x et y la relation de proportionnalité :

$$x = -\lambda_0 y$$

Dans tous les autres cas, il y a inégalité stricte (si $y \neq 0$).

L'inégalité de Schwarz permet d'obtenir une autre inégalité : l'inégalité de Minkowski,

4.4.3. Inégalité de Minkowski.

Pour tout x et y dans E, on a :

$$\boxed{\sqrt{<x+y, x+y>} \leqslant \sqrt{<x,x>} + \sqrt{<y,y>}}$$

Il y a égalité si et seulement si $x = \alpha y$ ou $y = \alpha x$ $(\alpha \geqslant 0)$

Partons de $<x+y, x+y> = <x,x> + <x,y> + <y,x> + <y,y>$

et utilisons l'inégalité de Schwarz pour majorer

$$\begin{aligned} <x+y, x+y> &\leqslant <x,x> + <y,y> + |<x,y>| + |<y,x>| \\ &\leqslant <x,x> + <y,y> + 2\sqrt{<x,x><y,y>} \\ &\leqslant \left(\sqrt{<x,x>} + \sqrt{<y,y>}\right)^2 \end{aligned}$$

Ce qui donne bien l'inégalité cherchée.

L'inégalité de Minkowski permet de construire une norme sur un espace vectoriel à partir de la donnée d'une forme bilinéaire hermitienne positive non dégénérée.

Théorème : l'expression $\|x\| = \sqrt{<x,x>}$ définit une norme sur l'espace vectoriel E muni d'un produit scalaire $<x,y>$.

Vérifions les quatre axiomes de définition d'une norme :

(1') $\|x\| \geqslant 0$

(2') $\|x\| = 0$ si, et seulement si $x = 0$: propriété (d) du produit scalaire

(3') $\|\lambda x\| = |\lambda| \, \|x\|$ pour tout λ de K : propriété (b) du produit scalaire

(4') $\|x+y\| \leqslant \|x\| + \|y\|$ grâce à l'inégalité de Minkowski.

Il est assez fréquent de pouvoir construire sur un espace vectoriel un produit scalaire (propriétés linéaires). On en déduit alors une norme. On convient de distinguer les espaces vectoriels normés dont la norme dérive d'un produit scalaire, donc de la définition suivante.

4.4.5. espaces préhilbertiens.

Un espace préhilbertien est un espace vectoriel sur K, muni d'un produit scalaire et normé par $\|x\| = \sqrt{<x,x>}$.

Vocabulaire.

Un espace préhilbertien sur le corps des réels est appelé un préhilbertien réel.

Un espace préhilbertien réel de dimension finie est un espace euclidien.

Un espace préhilbertien sur le corps des complexes est appelé un préhilbertien complexe.

Un espace préhilbertien complexe de dimension finie est un espace hermitien.

4.4.6. définition de l'orthogonalité.

Deux vecteurs x et y d'un espace préhilbertien sont orthogonaux lorsque $<x,y> = 0$.

Deux sous-espaces vectoriels E_1 et E_2 d'un préhilbertien sont orthogonaux lorsque tout vecteur x de E_1 est orthogonal à tout vecteur y de E_2. Le théorème suivant est évident, mais souvent utile.

Théorème : une condition nécessaire et suffisante pour que $x = 0$ est que $<x,y> = 0$ pour tout y d'un préhilbertien E.

4.4.7. théorème de Pythagore.

Si deux vecteurs x et y d'un espace préhilbertien sont orthogonaux on a l'égalité de Pythagore : $\boxed{\|x+y\|^2 = \|x\|^2 + \|y\|^2}$

4.4.8. espace dual algébrique d'un espace préhilbertien.

L'application $x \longrightarrow <x,y>$ (souvent notée $<x|y>$ par les physiciens), y étant fixé, définit une forme linéaire P_y de l'espace vectoriel préhilbertien E dans son corps de base K. La correspondance $y \longrightarrow P_y$ est une application linéaire de E dans son dual algébrique E^*. De plus, cette application est injective.

En effet, si $y_1 \neq y_2$, on a $P_{y_1} \neq P_{y_2}$ dans E^*, puisque l'égalité dans E signifierait $P_{y_1}(x) = P_{y_2}(x)$ pour tout élément x de E, donc :

$$<x,y_1> = <x,y_2> \quad \text{pour tout } x \text{ de } E .$$

Soit :

$$< x \; y_1 - y_2 > \; = 0$$

Faisant $x = y_1 - y_2$, on en déduit $\| y_1 - y_2 \| = 0$, soit $y_1 = y_2$ au sens du préhilbertien E (Cf. 4.4.6).

Finalement, on peut identifier E à un sous-espace vectoriel de E^* par la correspondance indiquée :

1er cas : lorsque E est un espace vectoriel de dimension finie, muni d'un produit scalaire donné a priori, E est en fait identifié à E^* par la correspondance $y \longrightarrow Py$. Nous verrons plus loin comment tout espace vectoriel de dimension finie peut être muni d'un produit scalaire lorsque l'on choisit une base a priori.

2e cas : lorsque E est un espace vectoriel de dimension infinie, muni d'un produit scalaire donné, on démontre que E est un sous-espace vectoriel strict de E^*.

En fait, les formes linéaires $Py : x \longrightarrow <x,y>$ de E^*, ne sont pas quelconques, elles satisfont à l'inégalité :

$$|Py(x)| \leqslant \|x\| \; M \qquad (\text{où } M = \|y\| \;)$$

qui est l'inégalité de Schwarz. Les formes linéaires qui vérifient une telle inégalité ont des propriétés remarquables, liées à l'analyse et non plus à l'algèbre pure. Nous les étudierons dans la partie B de ce chapitre et pratiquement nous ne rencontrerons pas d'autres formes linéaires.

4.4.9. Caractérisation d'un espace préhilbertien.

Un espace préhilbertien réel possède une norme $\|x\|$ qui satisfait l'égalité dite du parallélogramme (ou d'Apollonius) :

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

Réciproquement, un espace vectoriel normé (sur le corps des réels) dont la norme satisfait l'égalité précédente, est un espace préhilbertien réel.

Il suffit de vérifier que l'expression (Cf Exercice N° 14) :

$$<x,y> \;=\; \left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)\frac{1}{2}$$

constitue un produit scalaire sur E dont dérive la norme de E , selon la loi :

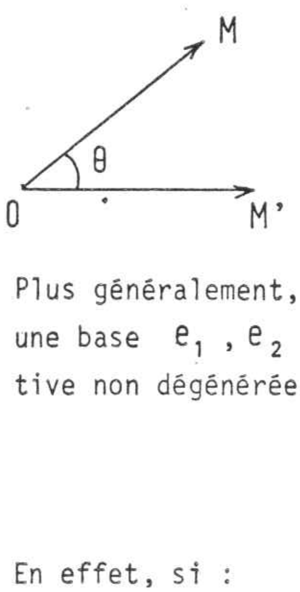
$$<x,x> \;=\; \|x\|^2$$

<u>L'égalité d'Apollonius caractérise donc les espaces préhilbertiens.</u>

4.5. Exemples.

4.5.1. Géométrie euclidienne.

Pour l'espace vectoriel à trois dimensions des vecteurs libres de la géométrie ordinaire, rapportés à une base orthonormale, on sait bien définir le produit scalaire de deux vecteurs $\overrightarrow{OM}$ et $\overrightarrow{OM'}$, produit qui vérifie les propriétés générales d'un produit scalaire. D'ailleurs, l'inégalité de Schwarz provient d'une propriété du cosinus de l'angle de deux vecteurs (le cosinus est borné en valeur absolue par 1) :



$$\cos.\,(\overrightarrow{OM},\overrightarrow{OM'}) = \frac{<\overrightarrow{OM},\overrightarrow{OM'}>}{\sqrt{<\overrightarrow{OM},\overrightarrow{OM}><\overrightarrow{OM'},\overrightarrow{OM'}>}} = \frac{\overrightarrow{OM}.\overrightarrow{OM'}}{|\overrightarrow{OM}|\,|\overrightarrow{OM'}|}$$

où $<\overrightarrow{OM},\overrightarrow{OM'}> \;=\; xx' + yy' + zz'$

Plus généralement, prenons un espace vectoriel réel à n dimensions. Dans une base $e_1, e_2, \ldots, e_n$ de E , une forme bilinéaire hermitienne positive non dégénérée est bien déterminée par les n^2 nombres :

$$<e_i\,,\,e_j> \;=\; g_{ij}$$

En effet, si :

$$x = \sum_{i=1}^{i=n} \alpha_i\, e_i$$

$$y = \sum_{j=1}^{j=n} \beta_j\, e_j$$

On a :

$$<x,y> = \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \alpha_i \beta_j g_{ij}$$

L'hypothèse de positivité (c) implique :

$$(1) \qquad \sum_{\substack{i=1 \\ j=1}}^{\substack{j=n \\ i=n}} \alpha_i \alpha_j g_{ij} = \sum_{i=1}^{i=n} g_{ii} \alpha_i^2 + 2 \sum_{i>j} \sum g_{ij} \alpha_i \alpha_j \geqslant 0$$

et l'hypothèse de non-dégénérescence assure que l'équation

$$(2) \qquad \sum_{\substack{i=1 \\ j=1}}^{n} \alpha_i \alpha_j g_{ij} = 0$$

ne possède que la solution : $\alpha_i = 0$ pour tout i de 1 à n.

En outre :

$$(3) \qquad g_{ij} = g_{ji} \qquad \text{(relation de symétrie puisque } K = \mathbb{R} \text{).}$$

Réciproquement, si les g_{ij} vérifient ces trois conditions (1), (2) et (3), la forme bilinéaire

$$<x,y> = \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \alpha_i \beta_j g_{ij}$$

définit un produit scalaire sur l'espace E , qui devient alors un espace euclidien, conformément au vocabulaire utilisé.

Ainsi, dans un espace vectoriel de dimension finie, il est aisé d'associer un produit scalaire à partir de la donnée d'une base. Le plus simple est de prendre $g_{ij} = 0$ si $i \neq j$, 1 si $i = j$.

4.5.2. Exemple fonctionnel.

Reprenons l'exemple 4.3.2 utilisant l'ensemble $\mathbb{C}[0,1]$, formé par les fonctions numériques continues définies sur $[0,1]$. Cet espace est un espace vectoriel réel pour les opérations ordinaires. On peut même le normer en espace $\mathscr{C}[0,1]$ par la norme :

$$\|f\| = \sup_{x \in [0,1]} |f(x)|$$

Cependant, nous allons construire sur $\mathbb{C}[0,1]$ un produit scalaire qui en fera un espace préhilbertien noté $\mathscr{C}[0,1]$. Pour deux éléments f et g de $\mathbb{C}[0,1]$ on pose :

$$< f,g> = \int_0^1 f(t)\, g(t)\, dt$$

f et g étant continues sur $[0,1]$, la fonction $x \longrightarrow f(x)\, g(x)$ est intégrable sur le segment $[0,1]$. On vérifie les propriétés :

(a) $<f,g>$ est réel et $<f,g> = <g,f>$

(b) la linéarité provient de la linéarité de l'intégrale définie.

(c) $<f,f> = \int_0^1 f^2(t)\, dt \geq 0$

(d) $<f,f> = 0$ si, et seulement si $f \equiv 0$. Démontrons cette assertion par l'absurde. Soit x_0 un point de l'intervalle $]0,1[$ où $f(x_0) \neq 0$. Grâce à la continuité de la fonction f sur le segment $[0,1]$, il existe un nombre $\varepsilon > 0$ de sorte que $f^2(x)$ soit strictement supérieur à une constante positive a^2 sur le segment $\left[x_0 - \frac{\varepsilon}{2}, x_0 + \frac{\varepsilon}{2}\right]$ On a alors la majoration :

$$\int_0^1 f^2(x)\, dx \geq \int_{x_0-\frac{\varepsilon}{2}}^{x_0+\frac{\varepsilon}{2}} f^2(x)\, dx > a^2\varepsilon > 0$$

ce qui est en contradiction avec l'hypothèse :

$$\int_0^1 f^2(x)\, dx = 0$$

(cf figure 1)

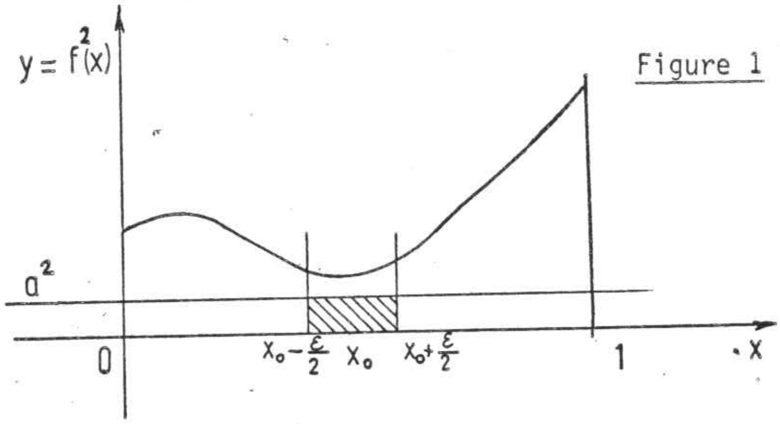


Figure 1

Finalement, $\langle f,g \rangle$ est un produit scalaire sur $\mathcal{C}[0,1]$ et l'espace vectoriel $\mathcal{C}[0,1]$, normé par :

$$\|f\|_2 = \left(\int_0^1 |f(t)|^2\, dt\right)^{1/2}$$

constitue un espace préhilbertien réel que nous notons : $\mathcal{C}[0,1]$.

L'inégalité de Schwarz n'est autre que l'inégalité classique :

$$\left(\int_0^1 f(t)\, g(t)\, dt\right)^2 \leqslant \left(\int_0^1 |f(t)|^2 dt\right)\left(\int_0^1 |g(t)|^2 dt\right)$$

L'égalité n'étant possible que si les deux fonctions f et g sont proportionnelles sur le segment $[0,1]$.

4.5.3. Espace l^2.

Soit $x = (x_1, x_2, \ldots, x_n, \ldots)$ une suite de nombres réels. Nous dirons que x est une suite de carré sommable lorsque :

$$\sum_{n=1}^{\infty} |x_n|^2 < +\infty$$

(ce qui signifie que la série converge vers une somme finie).

On appelle l^2 le sous-ensemble de l'ensemble des suites numériques constitué par toutes les suites de carré sommable.

a) l^2 est un espace vectoriel réel.

De façon naturelle, on définit l'addition de deux éléments de l^2 et la multiplication d'un élément de l^2 par un scalaire réel.

$$\left.\begin{aligned} x+y &= (x_1+y_1, x_2+y_2, \ldots, x_n+y_n, \ldots) \\ \lambda x &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n, \ldots) \end{aligned}\right\}$$

Il nous faut vérifier que ces opérateurs sont stables dans l^2.

Bien évidemment :

$$\sum_1^\infty |\lambda|^2 |x_n|^2 = |\lambda|^2 \sum_1^\infty |x_n|^2 < \infty \qquad \text{et } \lambda x \text{ appartient à } l^2$$

Par ailleurs, à l'ordre N :

$$\sum_{n=1}^{n=N} |x_n + y_n|^2 \leqslant \sum_{n=1}^{n=N} |x_n|^2 + \sum_{n=1}^{n=N} |y_n|^2 + 2 \sum_{n=1}^{n=N} |x_n y_n|$$

Mais l'on connaît l'inégalité de Schwarz, démontrée en 4.4.2 et adaptée au cas d'un espace vectoriel réel de dimension finie N muni de la norme hilbertienne : $\|x\|^2 = \sum_{n=1}^{n=N} |x_n|^2$

$$\left(\sum_{n=1}^{n=N} |x_n y_n|\right)^2 \leqslant \left(\sum_{n=1}^{n=N} |x_n|^2\right)\left(\sum_{n=1}^{n=N} |y_n|^2\right)$$

Pour tout N fini, grâce à la positivité des termes :

$$\sum_{n=1}^{N} |x_n + y_n|^2 \leqslant \sum_{n=1}^{\infty} |x_n|^2 + \sum_{n=1}^{\infty} |y_n|^2 + 2\sqrt{\left(\sum_{n=1}^{\infty} |x_n|^2\right)\left(\sum_{n=1}^{\infty} |y_n|^2\right)}$$

et la suite $S_N = \sum_{n=1}^{n=N} |x_n + y_n|^2$ est une suite croissante bornée par un nombre indépendant de N, donc converge lorsque N tend vers l'infini. Les autres propriétés, permettant d'assurer que l^2 est un espace vectoriel réel, sont de vérification facile.

L'élément neutre est évidemment $0 = (0, 0, \ldots, 0, \ldots)$.

b) <u>l'espace l^2 est un espace préhilbertien.</u>

L'expression $\sum_{n=1}^{\infty} x_n \; y_n$ a un sens si $x = (x_1, \ldots, x_n, \ldots)$ et $y = (y_1, y_2, \ldots, y_n, \ldots)$ sont des points de l^2 (utiliser l'égalité de Schwarz à l'ordre N et faire tendre N vers l'infini).

On vérifie alors les propriétés du produit scalaire pour $\sum_{n=1}^{\infty} x_n y_n$ défini sur l^2. On le note donc

$$<x, y> = \sum_{n=1}^{\infty} x_n y_n$$

et l'on obtient ainsi un espace l^2, préhilbertien réel.

D'ailleurs, conformément à la théorie générale, en posant

$$\|x\| = <x,x>^{1/2} = \left(\sum_{n=1}^{\infty} |x_n|^2\right)^{1/2}$$

on définit une norme sur l'espace l^2. Si l'on considérait des suites à valeurs complexes, au lieu de suites à valeurs réelles, on définirait de même un espace préhilbertien complexe, encore noté l^2. Sur cet espace, le produit scalaire est :

$$<x,y> = \sum_{n=1}^{\infty} x_n \overline{y_n}$$

l^2 réel apparaît comme la généralisation naturelle de l'espace euclidien à n dimensions et l^2 complexe comme la généralisation de l'espace hermitien à n dimensions.

Comme sur l'espace euclidien, nous voudrions disposer des outils usuels de l'analyse, à savoir la notion de limite, celle de continuité, de suite et de série convergente... Pour cela, nous allons maintenant développer l'aspect topologique de la théorie des espaces vectoriels normés.

B. CONSIDERATIONS TOPOLOGIQUES

La présence d'une norme sur un espace vectoriel E permet de mesurer la proximité de deux points x et y . Celle-ci est donnée par la valeur de $\|x-y\| = d(x,y)$ qui est une distance sur E . On sait que la notion de proximité sur l'axe réel conduit à la notion de limite et partant à toute l'analyse usuelle sur les nombres réels. Nous nous proposons de généraliser ces notions dans le cas d'un espace vectoriel normé. La théorie que nous ébauchons ici n'est qu'un cas particulier de la théorie générale des espaces vectoriels topologiques.

Sur un espace vectoriel normé, nous disposons de deux outils : la structure vectorielle et la norme. La structure additive fait jouer à tous les points de E un rôle identique en ce sens que, pour étudier la proximité de deux points, il suffit d'étudier la proximité de leur différence à zéro.

1. NOTION DE LIMITE.

Soit x_n une suite de vecteurs d'un espace vectoriel normé E sur le corps K . Par définition, on dit que la suite x_n converge vers l'élément x de E si, à tout nombre ε positif, on peut faire correspondre un entier naturel N tel que, pour tout n supérieur ou égal à N , on ait :

$$\|x_n - x\| \leqslant \varepsilon$$

On note :

$$\lim_{n\to\infty} x_n = x$$

En utilisant la notion de limite usuelle sur $\mathbb{R}$, cela signifie

$$\boxed{\lim_{n\to+\infty} x_n = x \text{ si, et seulement si } \lim_{n\to+\infty} \|x_n - x\| = 0}$$

Comme dans le cas de l'axe réel, on démontre que la limite d'une suite, si elle existe, est unique (ceci est lié à la propriété (2')).

Lorsqu'on a défini une notion de limite sur un espace E , on dit qu'on lui a donné une structure topologique et il devient un espace topologique, du moins si cette notion de limite satisfait certaines propriétés.

(Pour plus de détail, on pourra consulter les ouvrages classiques de topologie et au point de vue plus pédagogique dans Analyse et Topologie, Nanta Iremica Volume N° 10).

Exemple : prenons l'espace préhilbertien $\mathcal{C}\,[0,1]$, défini en 4.5.2

et considérons la suite de fonctions $f_o, f_1, f_2, \ldots f_n$

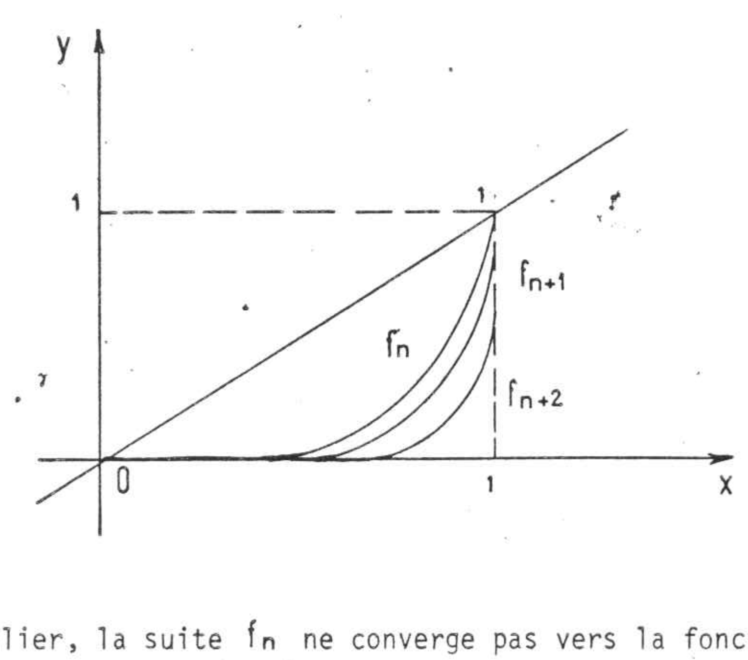
$$\begin{array}{lcl} f_o : & x \longrightarrow 1 \\ \vdots & \\ f_n : & x \longrightarrow x^n \end{array}$$

La suite des fonctions f_n converge, au sens de $\mathcal{C}\,[0,1]$, vers la fonction nulle. Calculons en effet :

$$\| f_n - 0 \|_2 = \sqrt{\int_0^1 (f_n(t) - 0)^2\, dt} = \frac{1}{\sqrt{2n+1}}$$

et $$\lim_{n \to \infty} \frac{1}{\sqrt{2n+1}} = 0$$

Par contre, il est clair que la suite f_n ne converge pas en chaque point x vers 0 puisque $f_n(1) = 1^n = 1$.



En particulier, la suite f_n ne converge pas vers la fonction 0 au sens de la norme définie sur $\mathcal{C}[0,1]$ (Cf A § 4.3), puisque

$$\| f_n - 0 \| = \sup_{x \in [0,1]} |f_n(x)| = 1$$

Cet exemple montre bien que la notion de limite est essentiellement liée à la norme utilisée sur l'espace vectoriel.

2. NOTION DE CONTINUITE.

Soient E et F deux espaces vectoriels normés sur le même corps K et f une application définie sur E à valeurs dans F.

Nous notons $\|.\|_1$ la norme dans E et $\|.\|_2$ la norme dans F.

Par définition, on dit que l'application f est continue au point x, de $(E, \|.\|_1)$ dans $(F, \|.\|_2)$ si, quelle que soit la suite x_n convergeant vers x, la suite $f(x_n)$ converge vers $f(x)$.

On note :

$$\boxed{\lim_{x_n \to x} f(x_n) = f(x)}$$

On dit que la fonction f est continue si elle est continue en chaque point x de E. Bien entendu, la notion de continuité dépend essentiellement des normes introduites sur les espaces E_1 et E_2. A titre d'exercice, on démontrera le théorème suivant :

Théorème de composition : si f est une fonction continue de $(E_1, \|.\|_1)$ dans $(E_2, \|.\|_2)$ et g une fonction continue de $(E_2, \|.\|_2)$ dans $(E_3, \|.\|_3)$, la fonction composée $g \circ f$ est continue de $(E_1, \|.\|_1)$ dans $(E_3, \|.\|_3)$.

3. OPERATEURS LINEAIRES CONTINUS.

Puisque nous avons affaire à des espaces vectoriels, il importe d'étudier particulièrement les opérateurs linéaires continus. Il existe une caractérisation de tels opérateurs, caractérisation que nous allons maintenant évoquer. Soient $(E, \|.\|_1)$ et $(F, \|.\|_2)$ deux espaces vectoriels normés sur le même corps K.

Lemme : pour qu'un opérateur linéaire P soit continu de E dans F, il faut et il suffit qu'il soit continu à l'origine.

Démonstration : la condition est évidemment nécessaire. Elle est suffisante puisque, si x_n converge vers x, alors $x_n - x$ converge vers 0, donc $P(x_n - x) = P(x) - P(x_n)$ converge vers $P(0) = 0$.

Théorème : pour qu'un opérateur linéaire P soit continu de E sur F, il faut et il suffit qu'il existe une constante M telle que l'on ait l'inégalité pour tout x de E :

$$(1) \qquad \boxed{\|P(x)\|_2 \leqslant M\,\|x\|_1}$$

Démonstration : il suffit d'examiner la continuité à l'origine grâce au lemme établi. La condition est suffisante puisque, si une suite x_n converge vers 0, alors

$$\|P(x_n)\|_2 \leqslant M\|x_n\|_1, \text{ donc } \lim_{n\to\infty} \|P(x_n)\|_2 \leqslant M \lim_{n\to\infty} \|x_n\|_1 = 0 \text{ et}$$

$P(x_n)$ converge vers $P(0) = 0$ lorsque n tend vers l'infini.

La condition est aussi nécessaire.

Supposons en effet qu'il n'existe pas de constante M finie pour laquelle (1) soit satisfaite : il existe alors une suite y_n de vecteurs non nuls de E pour lesquels on ait une inégalité du type :

$$\|P(y_n)\|_2 > n\,\|y_n\|_1$$

Considérons la suite $x_n = \frac{1}{\sqrt{n}} \frac{y_n}{\|y_n\|_1}$. Cette suite converge vers 0 au sens de E puisque :

$$\|x_n\|_1 = \frac{1}{\sqrt{n}},$$

mais $\|P(x_n)\|_2 > \sqrt{n}$, donc ne peut converger vers 0 contrairement à l'hypothèse de continuité de P. Il y a contradiction, ce qui assure l'existence de la constante M.

Indiquons la terminologie courante sur un espace vectoriel normé E .

Boule : on appelle boule (ouverte) de centre 0 et de rayon r (> 0), l'ensemble des points x tels que $\|x\| < r$.

La boule unité ouverte est la boule ouverte de rayon 1 : c'est l'ensemble des x tels que $\|x\| < 1$. La boule unité est l'ensemble des x tels que $x \leq 1$.

Ensemble borné : on appelle ensemble borné, un sous-ensemble de E inclus dans une boule de rayon fini.

Le théorème précédent peut se résumer de la façon suivante :

Pour qu'un opérateur linéaire P de $(E, \|.\|_1)$ dans $(F, \|.\|_2)$ soit continu, il faut et il suffit que tout ensemble borné de E soit transformé par P en un ensemble borné de F, ou encore que la boule unité de E soit transformée par P en un ensemble borné de F . On remarquera qu'une inégalité (propriété algébrique) équivaut à la continuité (propriété topologique).

Norme d'un opérateur linéaire continu : on appelle norme d'un opérateur linéaire continu la plus petite constante M réalisant l'inégalité (1). On notera cette norme $\|P\|$ ou mieux $\|P\|_{1,2}$ pour indiquer qu'elle dépend des normes et $\|.\|_1$ et $\|.\|_2$

Par définition :

$$(2) \qquad \boxed{\|P\|_{1,2} = \operatorname*{Sup}_{x \in E} \frac{\|P(x)\|_2}{\|x\|_1}}$$

Lemme : la norme d'un opérateur linéaire et continu P vaut

$$\|P\| = \operatorname*{Sup}_{\|x\|_1 \leq 1} \|P(x)\|_2 = \operatorname*{Sup}_{\|x\|_1 = 1} \|P(x)\|_2$$

Ces résultats sont faciles à démontrer grâce à l'homogénéité de P .

Exemple : soit $\mathcal{C}[0,1]$ l'espace vectoriel normé des fonctions réelles continues sur $[0,1]$ pour la norme uniforme (Cf A.2.3)

$$\|f\| = \operatorname*{Sup}_{x \in [0,1]} |f(x)|$$

Soit P l'opérateur linéaire (Cf Chapitre 0, § 3.10), défini de E dans E.

$$P : f \longrightarrow \int_0^x f(t)\,dt$$

Calculons sa norme :

$$\|P\| = \operatorname{Sup}_f \left(\frac{\operatorname{Sup}_{x \in [0,1]} \left| \int_0^x f(t)\,dt \right|}{\|f\|} \right)$$

Or :

$$\left| \int_0^x f(t)\,dt \right| \leqslant \int_0^x |f(t)|\,dt \leqslant \underbrace{\left(\int_0^1 dt \right) \operatorname{Sup}_{t \in [0,1]} |f(t)|}_{\leqslant \|f\|}$$

Par conséquent $\|P\| \leqslant 1$: l'opérateur P est un opérateur continu sur $\mathcal{C}[0,1]$

En outre, comme $P(1) = 1$, on a $\|P\| = 1$: l'opérateur P est de norme unité.

. On démontrerait de même que P est un opérateur linéaire continu et de norme 1 sur l'espace $\mathfrak{C}[0,1]$ (Cf A.4.5.2).

4. ESPACE DES OPERATEURS LINEAIRES CONTINUS.

Au Chapitre 0, § 3.6, nous avons montré que $\mathrm{Hom}(E, F)$ constitue un espace vectoriel sur le même corps que E et F. Lorsque ces deux espaces sont normés, on peut définir un sous-ensemble de $\mathrm{Hom}(E, F)$, l'ensemble des opérateurs linéaires continus de E dans F. On note $\mathcal{L}(E, F)$ cet ensemble, qui dépend des normes sur E et F.

$$\mathcal{L}(E, F) \subset \mathrm{Hom}(E, F)$$

Théorème : $\mathcal{L}(E, F)$ constitue un espace vectoriel normé sur K pour la norme $\| P \|$ définie en (2).

Il suffit de vérifier d'abord que $\mathcal{L}(E, F)$ est un sous-espace vectoriel, c'est-à-dire que $P_1 - P_2 \in \mathcal{L}(E,F)$ lorsque P_1 et $P_2 \in \mathcal{L}(E,F)$ Grâce à la caractérisation (1), il suffit donc de prouver que $\| P_1 - P_2 \|$ est finie.

Or :

$$\|P_1(x) - P_2(x)\|_2 \leqslant \|P_1(x)\|_2 + \|P_2(x)\|_2 \leqslant \|P_1\| \, \|x\|_1 + \|P_2\| \, \|x\|_1$$
$$\leqslant \left(\|P_1\| + \|P_2\|\right) \|x\|_1$$

Soit :

$$\|P_1 - P_2\| < \|P_1\| + \|P_2\|$$

Par ailleurs, il n'y a pas de difficulté à montrer que $P \longrightarrow \|P\|$ constitue une norme sur $\mathcal{L}(E, F)$.

On convient des notations suivantes :

| $\mathcal{L}(E, E)$ se note en général $\mathcal{L}(E)$. On a $\mathcal{L}(E) \subset \mathrm{Hom}(E, E)$
| $\mathcal{L}(E, K)$ se note en général E'. On a $E' \subset E^*$.

Dans les cas de dimension infinie, on démontre que les inclusions précédentes sont strictes (Cf. exemples ci-dessous).

L'espace E' s'appelle dual topologique de E pour le distinguer de E^*, dual algébrique. Dans ce qui suit, de nombreux raisonnements utilisent les propriétés relatives de E et E' et non celles de E et E^*. Cela constitue la théorie de la dualité topologique. Bien entendu, si x' est un élément de E', x' est une forme linéaire continue sur E dont la valeur au point x est notée :

$$(x, x')$$

conformément à la notation du Chapitre 0 § 4.1. De plus $|(x, x')| \leqslant \|x\| \, \|x'\|$ et cette inégalité, formellement voisine de celle de Schwarz, va jouer un rôle considérable.

Cas particuliers :

1er cas : soit $E_1 = E_2 = E$, un espace vectoriel muni de la norme $\|.\|$.
P est un opérateur continu sur E si, et seulement s'il existe un nombre positif a tel que :

(1) $$\|P(x)\| \leqslant a \|x\|$$

En posant :

$$\|P\| = \sup_{\|x\| \leqslant 1} \|P(x)\|$$

On voit que l'on dispose de l'inégalité (1), valable pour tout x de E :

(3) $$\|P(x)\| \leqslant \|P\| \|x\|$$

et de : $$\|P\| \leqslant a$$

$\|P\|$ est la meilleure constante possible en ce sens que toute constante inférieure à $\|P\|$ ne réalise pas l'inégalité (3). Réciproquement, l'inégalité (1) caractérise les opérateurs linéaires continus sur E .

2e cas : si $E_1 = E$ et $E_2 = K$ ($\mathbb{R}$ ou $\mathbb{C}$), P est une forme linéaire continue (définie sur E à valeurs dans K). Avec des notations analogues à celles de 5.3, on pose :

$$\|P\| = \sup_{\|x\| \leqslant 1} |P(x)|$$

on en déduit l'inégalité :

(2) $$|P(x)| \leqslant \|P\| \|x\|$$

$\|P\|$ est la meilleure constante possible réalisant l'inégalité (2).

5. ISOMORPHISMES D'ESPACES VECTORIELS NORMES.

5.1. Lorsque les espaces vectoriels E_1 et E_2 sont normés, sur le même corps K , on dit que P est un isomorphisme d'espaces vectoriels normés, si P constitue un isomorphisme algébrique tel que P et son inverse P^{-1} soient des opérateurs continus.

S'il existe un isomorphisme (d'espaces vectoriels normés) entre deux espaces vectoriels normés E_1 et E_2, on dit que ces deux espaces vectoriels normés sont isomorphes.

Du point de vue du calcul, on note que les conditions de continuité pour P et P^{-1} reviennent à l'existence de constantes a et b positives telles que :

$$\|P(x)\|_2 \leqslant a\|x\|_1$$

$$\|P^{-1}(y)\|_1 \leqslant b\|y\|_2$$

Remplaçons y par $P(x)$ dans cette dernière inégalité (ce qui est possible puisque P est une bijection) :

$$\|x\|_1 \leqslant b\|P(x)\|_2$$

On groupe ces inégalités sur une seule ligne pour obtenir l'écriture caractéristique des isomorphismes entre espaces vectoriels normés :

$$\boxed{\frac{\|x\|_1}{b} \leqslant \|P(x)\|_2 \leqslant a\|x\|_1}$$

Les meilleures constantes possibles sont :

$$a = \|P\|$$

$$b = \|P^{-1}\|$$

5.2. Comparaison des topologies normées.

Supposons que sur un espace vectoriel E, nous disposions de deux normes, notées $\|\cdot\|_1$ et $\|\cdot\|_2$. Différents cas intéressants sont à envisager.

1er cas : il existe une constante positive a telle que :

$$(1) \qquad \boxed{\|x\|_1 \leqslant a\|x\|_2}$$

pour tous les éléments x de E.

Ceci signifie que toute boule centrée en 0 pour la norme 1 contient une boule centrée en 0 pour la norme 2, grâce à l'inclusion :

$$\left\{ x \mid \|x\|_2 < \frac{\alpha}{a} \right\} \subset \left\{ x \mid \|x_1\| < \alpha \right\}$$

Soit I la transformation identique de E sur E. La relation (1), jointe à la caractérisation d'un opérateur continu, revient à dire que l'application identique est continue de (E , $\|.\|_2$) sur (E , $\|.\|_1$). On dit alors que la norme $\|.\|_2$ est plus fine que la norme $\|.\|_1$ ou encore que la topologie associée à la norme 2 est plus fine que la topologie associée à la norme 1 .

2e cas : $\|x\|_2 \leqslant b\|x\|_1$

La norme $\|.\|_1$ est plus fine que la norme $\|.\|_2$.

3e cas : $a\|x\|_1 \leqslant \|x\|_2 \leqslant b\|x\|_1$

Les deux normes sont respectivement plus fines l'une que l'autre et la transformation identique de (E , $\|.\|_1$) sur (E , $\|.\|_2$) constitue un isomorphisme d'espaces vectoriels normés. On dit que les topologies normées (ou les normes) sont équivalentes.

La relation : $\|.\|_1$ plus fine que $\|.\|_2$ est une relation d'ordre non total. Il existe en effet des normes non comparables (Cf. § 5.3.5).

Il importe de connaître les inégalités du type (1) entre les différentes normes que l'on considère sur un même espace vectoriel, et de rechercher les normes équivalentes afin d'effectuer les calculs sur la norme la plus facile à manier. De plus, il faut remarquer que pour deux normes équivalentes, toutes les notions de limite et de continuité coïncident. Du point de vue topologique, on identifie les deux normes.

5.3. Equivalence des topologies normées.

Une des tâches essentielles de l'analyse fonctionnelle est de fournir des moyens de comparaison des normes, donc des topologies normées. Nous énoncerons l'important théorème suivant.

5.3.1. Théorème : sur un espace vectoriel de dimension finie sur le corps des réels (ou sur le corps des complexes), toutes les normes possibles sont équivalentes.

Ce théorème fondamental est à la base de la plupart des théorèmes d'existence d'une meilleure approximation pour une fonction donnée et de bien d'autres théorèmes. Sa démonstration repose sur une notion de compacité que nous n'allons pas évoquer précisément.

Au Chapitre 0, paragraphe 3.3, nous avons établi que tout espace vectoriel de dimension finie N sur K est isomorphe (algébriquement) à l'espace vectoriel K^N. Pour fixer les idées, nous prenons pour K l'ensemble des nombres réels $\mathbb{R}$.

On peut munir l'espace $\mathbb{R}^N$ d'une norme particulière, notée $\|x\|$. Si y est un élément de $\mathbb{R}^N$ de la forme :

$$y = (y_1, y_2, \ldots y_N)$$

On pose :

$$\|y\| = \sup_{i=1,\ldots,N} |y_i|$$

On structure ainsi $\mathbb{R}^N$ en un espace vectoriel normé, noté $\mathbb{R}^N$, on peut alors énoncer le théorème suivant :

5.3.2. Théorème de Bolzano-Weierstrass : de toute suite bornée d'éléments de $\mathbb{R}^N$ on peut extraire une sous-suite convergente.

Dire qu'unesuite $y^{(n)}$ d'éléments de $\mathbb{R}^N$ est bornée signifie que cette suite est incluse dans une boule de $\mathbb{R}^N$, c'est-à-dire qu'il existe une constante positive A telle que :

$$\left| y_i^{(n)} \right| \leqslant A$$

pour tout i variant de 1 à N et pour tout n.

Si pour tout i fixé, nous pouvons montrer que l'on peut extraire une sous-suite $y_i^{(n_h)}$ telle que $\lim\limits_{h\to\infty} y_i^{(n_h)}$ existe, alors le théorème sera acquis.

En effet, construisons la sous-suite $y_1^{(n_h)}$ qui converge. Puis considérons la suite $y_2^{(n_h)}$ et extrayons de cette suite une sous-suite qui converge. On peut ainsi continuer de proche en proche jusqu'à l'indice N et le théorème est bien démontré.

Il reste donc à prouver que de toute suite bornée $y_1^{(n)}$ de nombres réels, on peut extraire une sous-suite convergente. On procède comme suit :

on a :

$$|y_1^{(n)}| \leqslant A$$

si les $y_1^{(n)}$ ne forment qu'un ensemble fini d'éléments, c'est que la suite $y_1^{(n)}$ prend une même valeur une infinité de fois, donc qu'il existe une sous-suite convergente.

On peut supposer dès lors que les $y_1^{(n)}$ forment un ensemble infini d'éléments distincts. Divisons le segment $[-A, +A]$ en deux parties égales $[-A, 0]$ et $[0, A]$. Une de ces deux parties au moins contient une infinité d'éléments de la suite $y_1^{(n)}$. Supposons que ce soit $[0, A]$. Divisons maintenant $[0, A]$ en deux parties $[0, \frac{A}{2}]$ et $[\frac{A}{2}, A]$.

Une de ces deux parties au moins contient une infinité d'éléments de la suite $y_1^{(n)}$. Supposons que ce soit $[0, \frac{A}{2}]$. Continuons de proche en proche. Au rang K, on aura une infinité d'éléments dans un intervalle du type $[A_K, B_K]$ où la longueur de $[A_K, B_K]$ est égale à $\frac{A}{2^{K-1}}$.

De plus, on a les inégalités :

$$\cdots \leqslant A_K \leqslant A_{K+1} \leqslant A_{K+2} \leqslant \cdots$$

$$\cdots \geqslant B_K \geqslant B_{K+1} \geqslant B_{K+2} \geqslant \cdots$$

La suite A_K et la suite B_K sont des suites de Cauchy puisque (Cf. remarque)

$$|A_{K+p} - A_K| \leqslant \text{longueur de } [A_K, B_K] = \frac{1}{2^{K-1}}$$

et de même pour B. Ces deux suites convergent donc, et la convergence a lieu vers le même nombre A_0 puisque : $|A_K - B_K| = \frac{1}{2^{K-1}}$

qui tend vers 0 lorsque K tend vers l'infini.

Si l'on prélève, à chaque rang K , un élément $y_1^{(n_k)}$ de la suite $y_1^{(n)}$ appartenant à l'intervalle $[A_K, B_K]$, on constate alors que la suite $y_1^{(n_k)}$ est convergente vers A_0 , ce qu'il fallait démontrer.

Remarque : dans cette démonstration, nous utilisons le théorème suivant : R est un espace complet (pour la définition, voir Chap.2). Ce théorème est démontré dans les cours d'analyse élémentaire, et apparaît comme essentiel lorsque l'on construit le corps des nombres réels par le procédé de G. Cantor.

5.3.3. Démonstration du théorème d'équivalence.

Considérons maintenant deux normes $\|.\|_1$ et $\|.\|_2$ sur un même espace vectoriel E de dimension finie N sur le corps des réels par exemple. Etre équivalentes pour deux normes est une relation transitive et par suite le théorème d'équivalence des normes sera acquis sur E si nous montrons que toute norme sur E est équivalente à une norme particulière. Choisissons une base algébrique $e_1, e_2, \ldots, e_N$ de E . Tout vecteur x de E s'écrit d'une façon unique :

$$x = \sum_{i=1}^{i=N} x_i e_i$$

Posons :

$$\|x\| = \operatorname{Sup}_{i=1 \ldots N} |x_i|$$

L'application $x \rightsquigarrow \|x\|$ définit une norme sur E .

Nous allons montrer que $\|.\|$ et $\|.\|_1$ sont en fait équivalentes.

Lemme 1 : il existe une constante b telle que l'on ait :

(1) $$\|x\|_1 \leqslant b \, \|x\|$$

En effet :

$$\|x\|_1 = \|\sum_{i=1}^{N} x_i e_i\|_1 \leqslant \sum_{i=1}^{N} |x_i| \, \|e_i\|_1 \leqslant \left(\operatorname*{Sup}_{i=1,\dots,N} |x_i|\right)\left(\sum_{i=1}^{N} \|e_i\|_1\right)$$
$$\leqslant b \, \|x\|$$

en posant :

$$b = \sum_{i=1}^{i=N} \|e_i\|_1$$

Lemme 2 : il existe une constante a telle que pour tout x de E

$$(2) \qquad \|x\| \leqslant a \, \|x\|_1$$

Supposons que (2) ne soit pas vérifiée. On peut alors construire une suite $x^{(n)}$ de vecteurs de E telle que

$$\|x^{(n)}\|_1 \leqslant 1 \quad \text{et} \quad \|x^{(n)}\| = \operatorname*{Sup}_{i=1,\dots,N} |x_i^{(n)}| \geqslant n$$

Posons donc :

$$\alpha_n = \operatorname*{Sup}_{i=1,\dots,N} |x_i^{(n)}| \quad \text{et formons} \quad y_i^{(n)} = \frac{x_i^{(n)}}{\alpha_n}$$

Appelons $y^{(n)}$ le vecteur de R^N dont les composantes sont les $y_i^{(n)}$
Puisque $|y_i^{(n)}| \leqslant 1$, le vecteur $y^{(n)}$ reste dans le cube unité de $\mathbb{R}^N$ et l'on peut appliquer le théorème de Bolzano-Weierstrass. Il existe une sous-suite $y^{(n_k)}$, convergente vers un élément y° de $\mathbb{R}^N$, c'est-à-dire qu'il existe N nombres réels y_i°, compris entre -1 et $+1$, tels que :

$$\lim_{k \to +\infty} |y_i^{(n_k)} - y_i^\circ| = 0 \qquad (i = 1, 2, \dots, N)$$

Avec un peu de réflexion, puisque $y^\circ \neq 0$, on notera que l'on peut choisir un nombre ε positif de sorte que :

$$\|\sum_{i=1}^{i=N} y_i^\circ e_i\|_1 - \varepsilon\sum_{i=1}^{N}\|e_i\|_1 = C$$

soit une constante (strictement) positive. Il existe un entier K tel que pour tous les entiers k supérieurs à K, on ait :

$$|y_i^{(n_k)} - y_i^0| < \varepsilon \text{ , i variant de } 1 \text{ à } N .$$

Donc, en utilisant l'inégalité triangulaire (Cf Remarque ; A.2.2), on a :

$$1 \geqslant \|\sum_{i=1}^{i=N} x_i^{(n_k)} e_i\|_1 = \alpha_{n_k}\|\sum_{i=1}^{N}\left(\frac{x_i^{(n_k)}}{\alpha_{n_k}} - y_i^\circ\right) e_i + \sum_{i=1}^{N} y_i^\circ e_i\|_1$$

$$\geqslant \alpha_{n_k}\left[\|\sum_{i=1}^{N} y_i^\circ e_i\|_1 - \varepsilon\sum_{i=1}^{i=N}\|e_i\|_1\right]$$

$$\geqslant \alpha_{n_k} C > C n_k .$$

Soit $1 > Cn_k$ ce qui est contradictoire puisque n_k tend vers l'infini et C est positif q.e.d.

En définitive, (1) et (2) impliquent :

$$\frac{1}{b}\|x_1\| < \|x\| \leqslant a\|x\|_1$$

c'est-à-dire l'équivalence des normes $\|.\|$ et $|.|_1$ et donc plus généralement l'équivalence de deux normes quelconques définies sur $\mathbb{E}$.

Finalement, l'étude tant topologique qu'algébrique d'un espace vectoriel de dimension finie sur K , se ramène à l'étude de $\mathbb{K}^n$ ($\mathbb{K}$ étant $\mathbb{R}$ ou $\mathbb{C}$, et n la dimension de l'espace).

Le théorème d'équivalence des normes est inexact en dimension infinie car le théorème de Bolzano-Weierstrass n'est plus valable lorsque la dimension n'est pas finie. Au paragraphe 5.3.5. nous construirons un contre-exemple significatif.

Corollaire : sur un espace vectoriel de dimension finie, tout opérateur linéaire est continu.

Lorsque l'opérateur P est bijectif ce corollaire se déduit directement du théorème d'équivalence des normes puisque $x \rightsquigarrow \|P(x)\|$ définit une norme.

Par passage au quotient de l'espace de départ par le noyau de l'opérateur P, on se ramène au cas précédent lorsque l'opérateur P n'est pas injectif.

5.3.4. Exemples usuels.

Sur l'espace vectoriel $\mathbb{R}^3$ (en fait, nous considérons plutôt l'espace affine associé en imposant une origine commune O aux vecteurs associés) nous avons construit trois distances distinctes en A.1.2.2. Ces trois distances proviennent des trois normes suivantes :

$$\|\overrightarrow{OM}\|_1 = \mathrm{Sup}(|x|,|y|,|z|)$$

$$\|\overrightarrow{OM}\|_2 = |x|+|y|+|z|$$

$$\|\overrightarrow{OM}\|_3 = \sqrt{x^2+y^2+z^2}$$

où $\overrightarrow{OM}$ désigne un vecteur de composantes x, y et z.

On obtient facilement les inégalités suivantes qui démontrent, dans ce cas particulier, le théorème d'équivalence pour les trois normes considérées :

$$a\,\|\overrightarrow{OM}\|_1 \leqslant \|\overrightarrow{OM}\|_3 < b\,\|\overrightarrow{OM}\|_1$$

$$c\,\|\overrightarrow{OM}\|_1 \leqslant \|\overrightarrow{OM}\|_2 \leqslant d\,\|\overrightarrow{OM}\|_1$$

et

$$e\,\|\overrightarrow{OM}\|_2 \leqslant \|\overrightarrow{OM}\|_3 < f\,\|\overrightarrow{OM}\|_2$$

où les constantes a, b, c, d, e et f ont les valeurs respectives :

$$a = 1\ ,\quad b = \sqrt{3}\ ,\quad c = 1\ ,\quad d = 3\ ,\quad e = \frac{1}{\sqrt{3}}\ ,\quad f = 1$$

Si l'on envisageait $\mathbb{R}^n$, au lieu de $\mathbb{R}^3$, le nombre 3 serait partout à remplacer par le nombre n .

5.3.5. Espace de polynômes.

a) soit E_n l'espace vectoriel constitué par les polynômes à une indéterminée, de degré inférieur ou égal à n et dont les coefficients sont réels.

Un élément de E_n est noté, en ordonnant ses puissances, selon

$$P(x) = a_0 + a_1 x + \text{-------} + a_n x^n$$

On définit la somme $P + Q$ et le produit λP comme à l'accoutumée :

$$P(x) + Q(x) = a_0 + b_0 + \text{-------} + (a_n + b_n)\, x^n$$

$$\lambda P(x) = \lambda a_0 + \text{--------} + \lambda a_n x^n$$

lorsque
$$P(x) = a_0 + a_1 x + \text{-------} + a_n x^n$$

et
$$Q(x) = b_0 + b_1 x + \text{-------} + b_n x^n$$

tandis que λ est un nombre réel quelconque.

Les polynômes $P_0(x) = 1$, $P_1(x) = x$, ... $P_n(x) = x^n$, constituent une base de l'espace vectoriel réel E_n dont la dimension est alors $n + 1$.

Construisons deux normes distinctes sur cet espace de dimension finie :

$$\|P\|_1 = \operatorname{Sup}\,(|a_0|, |a_1|, \text{-------}, |a_n|) = \operatorname{Sup}_{i = 0, \dots, n} |a_i|$$

et
$$\|P\|_2 = \operatorname{Sup}_{x \in [-1, +1]} |P(x)|$$

La première norme est classique sur un espace vectoriel de dimension ($n + 1$), rapporté à une base donnée (exemple A, § 1.2.2, norme 1) et nous avons déjà utilisé la deuxième norme (exemple 2.3 car l'on peut considérer E_n comme un sous-ensemble de $\mathcal{C}[-1, 1]$ et en faire ainsi un sous-espace métrique).

Le théorème d'équivalence des normes permet d'écrire a priori les deux inégalités suivantes, où a et b désignent des constantes positives :

$$a \, \|P\|_1 \leqslant \|P\|_2 \leqslant b \|P\|_1$$

Voyons ce qu'il en est directement.

La constante b est facile à déterminer grâce à l'inégalité :

$$\sup_{x \in [-1+1]} |P(x)| \leqslant (n+1) \sup_{i=,1..,n} |a_i|$$

donc on peut prendre $b \leqslant n+1$. En fait, la valeur $n+1$ est effectivement atteinte par le polynome :

$$P(x) = 1 + x + \text{-----} + x^n$$

pour la valeur $x = 1$. On prend donc $b = n+1$.

La constante a est moins facile à calculer et on admettra sa valeur (Cf. exercice N° 23)

$$a = 2^{1-n}$$

effectivement atteinte par un multiple scalaire du polynôme de Čebičev d'ordre n :

$$T_n(x) = \cos.(n \operatorname{Arc} \cos. x)$$

b) <u>extension</u> : on appelle ensemble E, la réunion des espaces vectoriels E_n :

$$E = \bigcup_{n=0}^{\infty} E_n$$

E est l'ensemble des polynômes à une indéterminée sur le corps des nombres réels (polynômes de degré quelconque). L'ensemble E est structuré en un espace vectoriel sur le corps des nombres réels, de la même manière que E_n. D'ailleurs E_n est un sous-espace vectoriel de E. Visiblement aussi, l'espace E n'est pas de dimension finie puisque les vecteurs

$$P_0(x) = 1 \;, \quad P_1(x) = x \;, \quad P_2(x) = x^2, \text{----}, P_n(x) = x^n \text{---}$$

sont linéairement indépendants

On remarque que E peut être considéré comme un sous-espace vectoriel de $\mathcal{C}[-1,+1]$.

Grâce à la connaissance de $\quad b = n+1$

et de $$\frac{1}{a} = 2^{n-1}$$

fonctions croissantes et tendant vers l'infini lorsque n croît, on peut constater que, sur E, les normes $\| P \|_1$ et $\| P \|_2$ ne sont pas équivalentes. En effet, les expressions
$$\frac{\| P \|_1}{\| P \|_2} \text{ et } \frac{\| P \|_2}{\| P \|_1}$$
ne peuvent pas rester bornées lorsque P parcourt l'ensemble E puisque déjà sur E_n, on a les estimations :

$$\sup_{P \in E_n} \frac{\| P \|_1}{\| P \|_2} = \frac{1}{a} = 2^{n-1}.$$

et

$$\sup_{P \in E_n} \frac{\| P \|_2}{\| P \|_1} = b = n+1$$

On peut ainsi constater que nous avons construit sur E deux normes $\| . \|_1$ et $\| . \|_2$ qui induisent des topologies normées non comparables. Tout provient de ce que E n'est pas de dimension finie.

5.4. Sous-espace vectoriels denses.

Au Chapitre 0, nous avons défini un sous-espace vectoriel. Les notions de topologie permettent de distinguer certains sous-espaces vectoriels particuliers.

Soit E un espace vectoriel normé sur K et F un sous-espace vectoriel de E (voire même un sous-ensemble quelconque).

5.4.1. Définition :

On dit que F est dense dans E lorsque tout vecteur x de F peut être approché, d'aussi près qu'on le désire, par un élément y de F .

Pour tout x dans E, et tout $\varepsilon > 0$, il existe y dans F tel que

$$\|x - y\| \leqslant \varepsilon$$

Il existe une définition équivalente.

Définition : F est dense dans E si, pour tout x donné dans E, il existe une suite x_n d'éléments de F telle que :

$$\lim_{n \to \infty} \|x - x_n\| = 0$$

Tout théorème de densité est un théorème d'approximation.

Exemple: le sous-ensemble des rationnels est dense dans l'espace des nombres réels.

Un sous-espace vectoriel F d'un espace E normé est évidemment dense dans lui-même.

Définition : F est dit fermé si toute suite convergente dans E et formée d'éléments de F, converge vers un élément de F.

Un sous-espace est fermé s'il ne permet pas d'approximer d'autres éléments que ses propres éléments.

Nous utiliserons ce vocabulaire commode d'espaces denses et d'espaces fermés. Bien entendu, ces notions restent relatives au choix d'une norme, c'est-à-dire d'une topologie. Toutefois, lorsque deux topologies sont comparables, on peut comparer les sous-espaces fermés. On dispose du théorème suivant :

5.4.2. Théorème : soit E un espace vectoriel muni de deux topologies normées $\|.\|_1$ et $\|.\|_2$. Supposons que $\|.\|_1$ soit plus fine que $\|.\|_2$ (c'est-à-dire qu'il existe une constante $a > 0$ telle que $\|x\|_2 \leqslant a\,\|x\|_1$) Tout sous-espace vectoriel fermé pour la topologie 2 est fermé pour la topologie 1.

Soit F un sous-espace vectoriel fermé pour la topologie 2 . Soit x_n une suite de F convergeant au sens de la norme $\|.\|_1$ vers un élément x de E . On a la majoration :

$$\| x_n - x \|_2 \leqslant a \| x_n - x \|_1$$

La suite x_n converge donc au sens de la norme 2 ce qui, joint à l'hypothèse, entraîne que x appartient à F. F est donc également fermé pour la topologie 1 .

Corollaire : pour deux normes équivalentes, les sous-espaces vectoriels fermés coïncident.

Ce corollaire permet, lors de la recherche de théorèmes d'approximation, d'utiliser parmi les normes équivalentes celle qui se prête le mieux aux calculs.

Définition : soit F un sous-ensemble d'un espace vectoriel normé E. La fermeture $\bar{F}$ de F est l'ensemble des éléments de E qui peuvent s'obtenir comme limite de suites d'éléments de F . La fermeture d'un sous-espace vectoriel est encore un sous-espace vectoriel. On peut montrer que $\bar{\bar{F}} = \bar{F}$, c'est-à-dire que la fermeture d'un sous-ensemble est un fermé.

5.4.3. Une technique de dualité.

Déterminer la fermeture d'un sous-espace vectoriel F d'un espace E , c'est déterminer les éléments de E approchables par des éléments de F . On dispose d'une technique de dualité très puissante pour ce faire, technique dont un usage remarquable sera fait au Chapitre III sur les méthodes hilbertiennes.

Théorème : soit F un sous-espace d'un espace normé E . L'espace F est dense dans E si et seulement si toute forme linéaire et continue définie sur E et nulle sur F est nécessairement la forme nulle.

Soit $\bar{F} = E$ et L est une forme linéaire et continue sur E , nulle sur F . Soit x un point de E , x est limite d'une suite $\{x_n\}$, $n \geqslant 0$, de points de F . Donc

$$\lim_{n \to +\infty} L(x_n) = L(x) \quad \text{mais} \quad L(x_n) = 0 \text{ pour tout } n \quad ,$$

donc $\quad L(x) = 0 \qquad$ et par suite $\quad L \equiv 0$

Inversement, soit $x \notin \overline{F}$, ce qui revient à dire, puisque $\overline{\overline{F}} = \overline{F}$, qu'il existe $a > 0$ et $\|x + y\| \geqslant a$ pour tout $y \in \overline{F}$. Par suite, définissons une forme linéaire L sur l'espace vectoriel $\mathcal{F}$ engendré par $\overline{F}$ et $\{x\}$ selon

$$L(\lambda x + y) = \lambda a$$

pour tout λ de $\mathbb{R}$ et tout y de $\overline{F}$. Par suite, si $\lambda \neq 0$, on a : $\|x + \frac{y}{\lambda}\| \geqslant a$ donc

$$|L(\lambda x + y)| = |\lambda| a \leqslant |\lambda| \|x + \frac{y}{\lambda}\| = \|\lambda x + y\|$$

et si $\lambda = 0$

$$0 = |L(y)| < \|y\|$$

On a donc toujours :

$$|L(\lambda x + y)| \leqslant \|\lambda x + y\|$$

c'est-à-dire que L est une forme linéaire et continue sur $\mathcal{F}$, nulle sur $\overline{F}$. On démontre (Cf théorème de Hahn-Banach ci-dessous) qu'on peut prolonger L en une forme linéaire (encore notée L) continue sur tout E.

On a $\quad L(F) = 0$ et $L(x) = a \neq 0$, ce qui est contradictoire.

L'étude de la densité (ou de la théorie de l'approximation) est donc simplifiée dès que l'on connaît bien le dual topologique de l'espace sur lequel on se place. Tel sera le cas pourles espaces de Hilbert que nous envisagerons au Chapitre III, ce qui explique une partie de leur succès. Terminons en énonçant le théorème utilisé :

Théorème de Hahn-Banach : soit E un espace normé et L une forme linéaire et continue sur un sous-espace vectoriel F de E (c'est-à-dire telle que $|L(y)| \leqslant \|L\| \|y\|$ pour tout y de F).

<u>Il existe une forme linéaire et continue $\widetilde{L}$ sur E, prolongeant L ($L(y) = \widetilde{L}(y)$ pour $y \in F$) et telle que $|\widetilde{L}(x)| \leqslant \|L\| \, \|x\|$ pour tout x de E.</u>

Nous ne démontrerons pas ce théorème et renvoyons le lecteur à la bibliographie indiquée en fin de volume.

EXERCICES SUR LE CHAPITRE I

EXERCICE N° 11:

Soit E un espace vectoriel de dimension n sur le corps des nombres complexes. On définit sur E une norme particulière liée au choix d'une base $(e_1, e_2, \ldots, e_n)$.

Si x s'écrit (d'une manière unique)

$$x = x_1 e_1 + \cdots + x_n e_n \text{ où } \quad x_K \in \mathbb{C}$$

On pose

$$\|x\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$$

$\|X\|$ est une norme provenant du produit scalaire $<x, y> = \sum_{k=1}^{k=n} x_k \overline{y}_k$ et E est un espace hermitien.

1°) Soit A un élément de $Hom(E, E)$ représenté par une matrice A dans la base $(e_1 \ldots e_n)$. Montrer que

$$\sup_{\|X\| = 1} \|Ax\| = \|A\|$$

définit une norme sur $Hom (E, E)$.

2°) Vérifier l'inégalité :

$$\|A B\| \leqslant \|A\| \|B\|$$

3°) Supposons la matrice A hermitienne sur E, de valeurs propres λ_i. Montrer que

$$\|A\| = \sup_{i = 1, \ldots, n} |\lambda_i|$$

Interpréter en termes de valeurs propres de $[B] = [g_{ij}]$ les conditions pour que n^2 nombres g_{ij} $(i = 1 \ldots, n ; j = 1, \ldots, n)$ permettent de définir un produit scalaire (cf A 4.5.1.)

4°) Pour une matrice quelconque, de coefficients a_{ij} , montrer que

$$\operatorname{Sup} |\lambda_i| \leqslant n \operatorname{Sup}_{i,j} |a_{ij}|$$

Que peut-on dire de $\|A\|$ par rapport aux valeurs propres de la matrice A ?

EXERCICE N° 12 :

On utilise les notations de l'exercice précédent.

Montrer que

$$\|A\|_2 = \sqrt{\operatorname{Tr}\left([A].[A]^*\right)}$$

(où $[A]^*$ désigne la matrice adjointe de $[A]$), est une norme sur $\operatorname{Hom}(E,E)$, (la trace notée Tr A d'une matrice A est la somme des coefficients de la diagonale de A).

EXERCICE N° 13 :

Soit E un espace vectoriel normé. Montrer que l'application $y \longrightarrow \|y\|$ est une application continue de E dans $\mathbb{R}$.

EXERCICE N° 14 :

Soit E un espace vectoriel réel normé. On suppose que cette norme $x \longrightarrow \|x\|$ satisfait l'égalité d'Apollonius.

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

Montrer que cette norme provient d'un produit scalaire.

EXERCICE N° 15 :

Soit E un espace vectoriel normé réel et P un opérateur linéaire défini sur E , à valeurs dans F ($P \in \operatorname{Hom}(E,F)$), où F est un espace vectoriel normé.

Montrer qu'une condition nécessaire de continuité pour P est que l'ensemble des vecteurs x tels que $P(x) = 0$ (noyau de P) soit fermé. Examiner si la condition est suffisante. (cas F quelconque et cas $F = \mathbb{R}$ ou $\mathbb{C}$).

Application : montrer que l'ensemble des fonctions f de $\mathcal{C}[0,1]$ telles que

$$\int_0^1 f(x)\, dx = 0 ,$$ est un ensemble fermé de $\mathcal{C}[0,1]$

Chapitre II

ESPACES NORMES COMPLETS

La droite réelle est le modèle qui a donné naissance aux espaces vectoriels topologiques plus généraux. Mais la droite réelle possède de nombreuses propriétés que ne possède pas tout espace normé. On classe alors les espaces normés selon les propriétés qu'ils ont en commun avec la droite réelle. Nous prendrons l'exemple du critère de CAUCHY afin de disposer d'une technique pour l'étude de la convergence.

II.1. CRITERE DE CAUCHY, ESPACES DE BANACH ET DE HILBERT.

Lorsque l'on se donne une suite $\{x_n\}$ de nombres réels, on peut chercher à savoir si cette suite est convergente, sans avoir cependant de moyen pour deviner quelle sera la valeur éventuelle de la limite . Il existe un théorème qui permet de démontrer qu'une suite de nombres réels est convergente, avant de savoir vers quel nombre elle converge (les termes de la suite forment alors des approximations successives de ce nombre).

II.1.1. Définition.

Une suite x_n de nombres réels est convergente si,et seulement si, à tout nombre ε positif, on peut faire correspondre un entier N tel que, pour tous les entiers m et n supérieurs ou égaux à N , on ait :

$$|x_m - x_n| < \varepsilon$$

Ce critère (condition nécessaire et suffisante de convergence d'une suite de nombres réels) est dû à CAUCHY et on veut le transcrire en termes d'espaces normés. Dans ce cadre plus général, définissons une suite de Cauchy :

Une suite de Cauchy dans un espace normé (E , $\|.\|$) est une suite $\{x_n\}$ de points de E telle que l'on puisse faire correspondre à tout nombre ε positif un entier naturel N de sorte que pour tous les entiers m et n supérieurs ou égaux à N , on ait :

$$\|x_m - x_n\| < \varepsilon$$

a) Montrons d'abord que sur un espace normé toute suite convergente constitue une suite de Cauchy. En effet, soit $\{x_n\}$ une suite convergeant vers un élément x_0 au sens de E . Par définition de la convergence, pour tout $\varepsilon > 0$, il existe un nombre N tel que pour tout n supérieur ou égal à N , on ait :

$$\|x_n - x_o\| < \frac{\varepsilon}{2}$$

Par conséquent, pour un couple m , n d'entiers naturels supérieurs ou égaux à N , on a la majoration de Cauchy :

$$\|x_m - x_n\| \leqslant \|x_m - x_o\| + \|x_o - x_n\| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

soit :

$$\|x_m - x_n\| \leqslant \varepsilon$$

b) Inversement, des exemples simples montrent qu'une suite de Cauchy n'est pas toujours convergente dans le cas d'un espace métrique quelconque. On pensera, par exemple, au corps des nombres rationnels, $\mathbb{Q}$, considéré comme sous-espace métrique du corps des nombres réels $\mathbb{R}$. Soit $\{x_n\}$ une suite de nombres rationnels convergeant vers le nombre $\sqrt{2}$. Par exemple la suite des approximations décimales de $\sqrt{2}$ ($x_1 = 1$; $x_2 = 1{,}4$; $x_3 = 1{,}41$; $x_4 = 1{,}414$) est de Cauchy au sens de $\mathbb{R}$ et a fortiori au sens de $\mathbb{Q} \subset \mathbb{R}$,puisque les x_n sont des nombres rationnels. Or $\sqrt{2}$ n'est pas un nombre rationnel. $\{x_n\}$ est une suite de Cauchy de nombres rationnels qui ne converge pas vers un nombre rationnel.

Cette remarque justifie la définition suivante :

II.1.2. Espaces normés complets : espaces de Banach.

Un espace normé est dit complet lorsque le critère de Cauchy s'applique, c'est-à-dire lorsque toute suite de Cauchy de cet espace converge au sens de (E , $\|\cdot\|$).

Théorème : la droite réelle est un espace complet (critère de Cauchy).

De même, l'espace $\mathbb{R}^n$, l'espace des nombres complexes $\mathbb{C}$, le produit $\mathbb{C}^n$, sont des espaces complets.

Bien des espaces importants de l'Analyse sont des espaces complets comme nous aurons l'occasion de le voir.

II.1.3. Exemple.

Reprenons l'espace $\mathcal{C}\,[0,1]$ et montrons qu'il s'agit d'un espace normé complet.

Soit $\{f_n\}$ une suite dans $\mathcal{C}\,[0,1]$, telle que, pour tout $\varepsilon < 0$, il existe un entier $N > 0$ assurant, pout tous les entiers m et n supérieurs à N , l'inégalité :

$$(1) \qquad \|f_n - f_m\| = \operatorname*{Sup}_{x \in [0,1]} |f_n(x) - f_m(x)|$$

Fixons x ; la suite numérique $\{f_n(x)\}$ est de Cauchy et converge donc vers un nombre noté $f(x)$. Maintenant $x \longrightarrow f(x)$ définit une fonction f de $[0,1]$ dans $\mathbb{R}$. De plus, pour tout x , l'inégalité :

$$|f_n(x) - f_m(x)| < \varepsilon$$

entraîne :

$$\lim_{m \to \infty} |f_n(x) - f_m(x)| \leq \varepsilon$$

soit :

$$|f_n(x) - f(x)| \leq \varepsilon$$

pour tout nombre x du segment $[0,1]$, c'est-à-dire, pour tout entier n supérieur ou égal à N :

$$(2) \qquad \operatorname*{Sup}_{x \in [0,1]} |f_n(x) - f(x)| \leq \varepsilon$$

ou encore : $\|f_n - f\| \leq \varepsilon$

Par conséquent, la suite de fonctions $\{f_n\}$ converge au sens de la convergence uniforme sur le segment $[0,1]$ vers la fonction f .

Démontrons enfin que f est une fonction continue.

Soit ε un nombre donné, x_0 un point donné de $[0,1]$.
Formons :

$$f(x) - f(x_0) = f(x) - f_n(x) + f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0)$$

Majorons

$$\begin{aligned} |f(x)-f(x_0)| &\leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &\leqslant 2\|f - f_n\| + |f_n(x) - f_n(x_0)| \end{aligned}$$

Choisissons n tel que $\|f - f_n\| \leqslant \varepsilon/4$, ce qui est possible d'après (2). Puis, n fixé, choisissons η tel que $|f_n(x) - f_n(x_0)| < \varepsilon/2$ pour $|x - x_0| < \eta$, ce qui est possible grâce à la continuité de f_n. Par suite, pour tout x satisfaisant $|x - x_0| < \eta$, on a :

$$|f(x) - f(x_0)| < \varepsilon$$

ce qui établit la continuité de f et termine la démonstration puisque $\{f_n\}$ est une suite de Cauchy, convergeant vers la fonction continue f , au sens $\mathcal{C}[0,1]$. Nous avons donc établi que :

$\mathcal{C}[0,1]$ est un <u>espace vectoriel normé complet</u>.

II.1.4. Exemple : l'espace l^2

Reprenons l'espace préhilbertien réel l^2 défini au Chap.IA,§4.5.3 et montrons que cet espace est un espace complet pour la norme :

$$\|x\| = \sqrt{\sum_{n=1}^{\infty} |x_n|^2}$$

Nous avons déjà dit que l'espace l^2 constitue une généralisation naturelle, pour une dimension infinie, de l'espace euclidien $\mathbb{R}^n$ de la géométrie usuelle. Dans ce cas, $(x_n)_{n\geqslant 1}$ sont les coordonnées du vecteur x .
Le langage géométrique, support de l'intuition, doit être conservé pour l^2

et nous allons montrer comment des propositions naturelles de géométrie ordinaire se traduisent dans ce cadre plus général.

On sait déjà que $\mathbb{R}^n$ est un espace complet. Montrons que l^2 est également un espace complet. Ce résultat est important car il permet de faire des calculs de séries sur l^2 comme dans le cas de $\mathbb{R}$.

Soit donc $x^{(m)}$ une suite de Cauchy d'éléments de l^2 , c'est-à-dire :

$$x^{(m)} = (x_1^{(m)}, x_2^{(m)}, \dots x_n^{(m)}, \dots)$$

Pour tout ε positif, il existe un N tel que pour tout m et tout p supérieurs à N , on ait :

$$\| x^{(m)} - x^{(p)} \| < \varepsilon$$

c'est-à-dire :

$$\sum_{n=1}^{\infty} |x_n^{(m)} - x_n^{(p)}|^2 < \varepsilon^2$$

En particulier, pour tout n donné

$$|x_n^{(m)} - x_n^{(p)}| < \varepsilon$$

pour m et p supérieurs à N .

Donc $x_n^{(m)}$, pour n fixé, constitue une suite de Cauchy de nombres réels et converge, puisque $\mathbb{R}$ est complet, vers un nombre réel noté x_n .

De plus, on peut écrire à l'ordre k :

$$\sum_{n=1}^{n=k} |x_n^{(m)} - x_n^{(p)}|^2 < \varepsilon^2$$

pour m et p supérieurs à N .

Faisons tendre p vers l'infini, k restant fixé. Pour chaque n , $x_n^{(p)}$ converge vers x_n . On peut alors écrire :

$$\sum_{n=1}^{n=k} |x_n^{(m)} - x_n|^2 \leq \varepsilon^2$$

Montrons tout d'abord que $x = (x_1, x_2, \ldots, x_n, \ldots)$ définit un élément de l^2. On dispose de

$$\sqrt{\sum_{n=1}^{n=k} |x_n|^2} \leqslant \sqrt{\varepsilon^2} + \sqrt{\sum_{n=1}^{n=k} |x_n^{(m)}|^2}$$

grâce à l'inégalité triangulaire à l'ordre k ; ce qui est l'inégalité triangulaire de la géométrie de $\mathbb{R}^k$.

Passant à la limite en k, il vient aussitôt :

$$\sqrt{\sum_{n=1}^{\infty} |x_n|^2} \leqslant \varepsilon + \sqrt{\sum_{n=1}^{\infty} |x_n^{(m)}|^2} < +\infty$$

De plus, de l'inégalité :

$$\sum_{n=1}^{k} |x_n^{(m)} - x_n|^2 \leqslant \varepsilon^2$$

lorsque $m > N$, on déduit en passant à la limite en k :

$$\sum_{n=1}^{\infty} |x_n^{(m)} - x_n|^2 \leqslant \varepsilon^2$$

soit :

$$\| x^{(m)} - x \| \leqslant \varepsilon$$

pour tout $m > N$.

Nous avons bien démontré que la suite de Cauchy $x^{(m)}$ de l^2 converge au sens de l^2 vers l'élément $x = (x_1, x_2, \ldots, x_n, \ldots)$ de l^2. L'espace l^2 est un espace préhilbertien normé complet.

II.1.5. Espaces de Banach et de Hilbert.

Un espace de Banach est un espace vectoriel normé complet.

Les espaces de Banach jouent un grand rôle en analyse tant par leurs remarquables propriétés que par leur fréquente réalisation.

Un espace K^n , l'espace $\mathcal{C}[0,1]$,etc. sont des espaces de Banach. De plus, au paragraphe suivant, nous allons donner une méthode visant à rendre complet tout espace vectoriel normé.

Un espace de Banach dont la norme dérive d'un produit scalaire s'appelle un espace de Hilbert. Ou encore, un espace préhilbertien complet est un espace de Hilbert.

Dorénavant, nous allons essentiellement considérer des espaces de Hilbert. Le Chapitre III étudie les propriétés générales des espaces de Hilbert et des opérateurs définis sur les espaces de Hilbert. Cependant, il importe de savoir construire de bons espaces fonctionnels qui soient des espaces de Hilbert. Une technique très générale va être décrite en II.3.

II.2. THEOREME DE PROLONGEMENT D'UN OPERATEUR CONTINU.

Auparavant, nous allons établir un théorème de prolongement montrant tous les avantages des espaces complets et mettant en évidence le rôle tout particulier des sous-espaces vectoriels denses pour l'étude des opérateurs linéaires.

Théorème de prolongement : soient E et F deux espaces vectoriels normés où F est supposé de Banach. Soit E_1 un sous-espace vectoriel dense de E . Tout opérateur linéaire continu P , défini sur E_1 et à valeurs dans F , se prolonge d'une manière unique en un opérateur linéaire continu $\bar{P}$, défini sur E à valeurs dans F .

Ceci signifie qu'il existe un opérateur linéaire continu unique $\bar{P}$ (élément de $\mathcal{L}(E, F)$) tel que :

$$\bar{P}(x) = P(x) \quad \text{pour tout } x \quad \text{dans } E_1$$

Démonstration : la continuité de P de E dans F assure l'existence d'une constante $\|P\|$ telle que :

$$\forall x \in E_1, \|P(x)\|_F \leqslant \|P\| \, \|x\|_{E_1}$$

(nous avons, bien sûr $\|x\|_{E_1} = \|x\|_E$: Cf. Chapitre 1;A § 1.3)

Soit y un élément de E . Puisque E_1 est dense dans E , il existe une suite $[y_n]_{n\geq 1}$ d'éléments de E_1 convergeant vers y . La suite $[y_n]_{n\geq 1}$ est de Cauchy et il en est de même pour $[P(y_n)]_{n\geq 1}$ dans F puisque :

$$\| Py_n - Py_m \|_F = \| P(y_n - y_m) \|_F \leq \|P\| \, \|y_n - y_m\|_{E_1}$$

Or F est de Banach, par conséquent $P(y_n)$ converge vers un élément noté $\bar{P}(y)$. Pour assurer la cohérence de cette notation, il convient de montrer qu'une autre suite $[y'_n]_{n\geq 1}$ de E_1 convergeant vers x , donne une suite $P(y'_n)$ qui converge dans F vers le même vecteur $\bar{P}(y)$. Or :

$$\| P(y'_n) - P(y_n) \|_F \leq \|P\| \, \|y'_n - y_n\|_E$$

et :

$$\lim_{n\to\infty} \|y'_n - y_n\|_E = 0$$

puisque par définition : $\lim\limits_{n\to\infty} y'_n = \lim\limits_{n\to\infty} y_n = y$ dans E .

En conséquence $\bar{P}(y)$ ne dépend que de y et non de la suite $\{y_n\}$

Nous avons ainsi défini une application $\bar{P}$:

$$x \rightsquigarrow \bar{P}(x)$$

qui, à tout x dans E , associe un élément $\bar{P}(x)$ de F . Bien entendu, si $x \in E_1$, $\bar{P}(x) = P(x)$ puisque la suite stationnaire $[y_n = x]_{n\geq 1}$ peut être utilisée pour définir $\bar{P}(x)$. $\bar{P}$ <u>prolonge l'opérateur</u> P .

Montrons que $\bar{P}$ est linéaire : soient y_n et y'_n deux suites (de Cauchy) dans E_1 , convergeant respectivement vers y et y' :

$$P(\lambda y_n + \mu y'_n) = \lambda P(y_n) + \mu P(y'_n)$$

Soit :

$$\bar{P}(\lambda y + \mu y') = \lim_{n\to\infty} (\lambda P(y_n) + \mu P(y'_n)) = \lambda \bar{P}(y) + \mu \bar{P}(y')$$

ce qui établit la linéarité.

Assurons maintenant la continuité de $\bar{P}$: ceci provient des inégalités :

$$\|P(y_n)\|_F \leqslant \|P\| \ \|y_n\|_E \quad \text{où} \quad y_n \in E_1$$

pour tout n , donc par prolongement :

$$\|\bar{P}(y)\|_F \leqslant \|P\| \ \|y\|_E$$

soit :

$$\|\bar{P}\| \leqslant \|P\|$$

mais un examen facile conduit à

$$\boxed{\|\bar{P}\| = \|P\|}$$

Reste à établir l'unicité du prolongement $\bar{P}$. Supposons qu'il existe un deuxième prolongement $\bar{\bar{P}}$. On aurait :

$$\bar{P}(x) = \bar{\bar{P}}(x) = P(x) \qquad \text{pour tout } x \text{ dans } E_1 \ .$$

Pour toute suite y_n de vecteurs de E_1 convergeant vers y dans E , on aurait grâce à la continuité du prolongement:

$$\bar{P}(y) = \lim_{n \to \infty} P(y_n) \qquad \text{et} \qquad \lim_{n \to \infty} P(y_n) = \bar{\bar{P}}(y)$$

d'où $\bar{P} = \bar{\bar{P}}$ c.q.f.d.

Le théorème que nous venons de démontrer est important car il permet de travailler sur un sous-espace dense, du moins en ce qui concerne un opérateur linéaire continu. Ce sera le théorème clef pour la construction de l'intégrale de Lebesgue.

Notons le rôle essentiel joué par la continuité de l'opérateur et par le fait que l'espace où l'opérateur prend ses valeurs est complet. Signalons qu'il est possible qu'une forme linéaire, continue sur un sous-espace dense, par exemple nulle sur ce sous-espace, se prolonge en une forme linéaire non identiquement nulle et donc non continue sur l'espace entier.

Nous serons ainsi amenés à étudier certains sous-espaces denses d'un espace de Banach (Cf. notion voisine de sous-ensemble total, Chapitre III). Venons-en, maintenant, à l'étude promise de la complétion.

II.3. COMPLETION D'UN ESPACE NORMÉ.

Nous avons noté au paragraphe précédent que l'espace $\mathbb{Q}$ des nombres rationnels ne forme pas un espace métrique complet. Cependant, il existe un espace normé complet bien familier : la droite réelle. Cet espace contient l'espace $\mathbb{Q}$ et tout point de $\mathbb{R}$ peut s'obtenir comme limite d'une suite de Cauchy de nombres rationnels. Ce procédé abstrait de construction des nombres réels au moyen de suites de Cauchy est dû au mathématicien G. Cantor*. Le procédé est généralisable à tout espace normé.

II.3.1. Théorème de complétion : A tout espace normé $(E, \|\cdot\|)$, on peut associer un espace normé complet $(\tilde{E}, \tilde{\|\cdot\|})$ tel que tout élément de $\tilde{E}$ s'obtienne comme limite d'une suite de Cauchy d'éléments de E.

De plus, $(E, \|\cdot\|)$ est un sous-espace métrique (dense) de l'espace $(\tilde{E}, \tilde{\|\cdot\|})$. En particulier, pour tout point x de E :

$$\tilde{\|x\|} = \|x\|$$

L'espace $(\tilde{E}, \tilde{\|\cdot\|})$ est appelé le complété de $(E, \|\cdot\|)$.

Nous ne ferons qu'indiquer brièvement l'idée de la démonstration, peu enrichissante en soi. On considère l'ensemble $\mathcal{E}$ des suites de Cauchy sur E. Un élément de $\mathcal{E}$ est donc une suite $x=[x_n]_{n\geq 1}$ de Cauchy d'éléments de E. On dit que deux suites x et x' sont équivalentes lorsque

$$\lim_{n\to\infty}\|x_n - x'_n\| = 0$$

On considère alors l'espace vectoriel $\tilde{E}$ des classes d'équivalence dans $\mathcal{E}$. Soit $\tilde{x}$ un élément de $\tilde{E}$ et $x=[x_n]_{n\geq 1}$ un représentant de $\tilde{x}$ dans $\mathcal{E}$. On pose

$$\tilde{\|\tilde{x}\|} = \lim_{n\to\infty}\|x_n\|$$

On montre sans peine que cette définition a un sens car $\|x_n\|$ forme une suite de Cauchy de nombres réels et ne dépend que de la classe $\tilde{x}$ de x. En outre, $\tilde{\|\cdot\|}$ constitue une norme sur $\tilde{E}$.

* Un historique de cette question est discuté dans le volume 3 de Nanta Iremica, publi aussi chez CEDIC/NATHAN : Nombre, mesure et continu, Epistémologie et histoire J. DHOMBRES (1978).

De plus E peut être envisagé comme sous-ensemble de $\tilde{E}$ en associant à l'élément x de E la classe d'équivalence de la suite stationnaire

$$[x, x, x, \dots].$$

En outre si $x \in E$, on a bien

$$\|\tilde{x}\| = \|x\|$$

Deux choses restent à montrer:

d'une part que E est dense dans $\tilde{E}$, ce qui n'est pas très difficile puisque, pour parler de façon imagée, un élément $\tilde{x}$ est la limite, qui manque peut-être dans E, d'une suite de Cauchy d'éléments de E. Le procédé de construction de $\tilde{E}$ a été de "boucher tous les trous possibles";

d'autre part, et c'est l'essentiel, que $\tilde{E}$ est un espace normé complet. En un mot qu'il n'y a pas de nouveaux "trous"! Ce point technique se vérifie sans trop de peine par extraction convenable d'une suite dans une suite double.

Pour résumer, $\tilde{E}$ apparaît comme l'ensemble des "limites" de suites de Cauchy d'éléments de E. On peut même facilement montrer que si E' est un espace normé complet ayant vis à vis de E les mêmes propriétés que $\tilde{E}$, alors E' et $\tilde{E}$ sont isométriquement isomorphes en tant qu'espaces de Banach (utiliser le paragraphe 2). Ceci justifie l'emploi de l'expression : $\tilde{E}$, le complété de E.

Au lieu d'insister sur ces généralités, donnons un remarquable exemple d'application. Signalons cependant que si $(E, \| \ \|)$ est un espace préhilbertien, on peut choisir pour $(\tilde{E}, \widetilde{\|\cdot\|})$ un espace de Hilbert, dont le produit scalaire $\widetilde{<\ ,\ >}$ prolonge le produit scalaire défini sur E.

II.3.2. Espace préhilbertien $\mathcal{C}^{\bullet}[0,1]$.

Sur l'espace vectoriel $\mathcal{C}[0,1]$ des fonctions continues réelles, nous avons défini deux topologies normées donnant les espaces $\mathcal{C}[0,1]$ et $\mathcal{C}^{\bullet}[0,1]$ (Cf Chapitre 1, A.4.5.2).

Comparons maintenant les deux topologies normées que nous avons définies sur l'espace vectoriel $\mathcal{C}[0,1]$. Grâce aux propriétés bien connues des intégrales définies, on démontre facilement :

$$\|f\|_2^2 = \int_0^1 |f^2(t)|dt \leq \sup_{t\in[0,1]} |f^2(t)| \int_0^1 dt = \left(\sup_{t\in[0,1]} |f(t)|\right)^2$$

En notant :

$$\| f \|_2 = \left(\int_0^1 |f(t)|^2 dt \right)^{1/2}$$

et

$$\| f \|_\infty = \operatorname*{Sup}_{t \in [0,1]} | f(t) |$$

il vient donc :

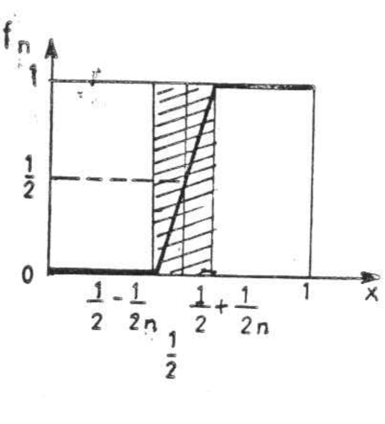
$$\boxed{\| f \|_2 \leqslant \| f \|_\infty}$$

Cette inégalité implique la continuité de l'opérateur identique de $\mathcal{C}[0,1]$ dans $\mathcal{C}[0,1]$ (Cf. § 5.2), ce qui signifie que la topologie de $\mathcal{C}[0,1]$ est plus fine que la topologie de $\mathcal{C}[0,1]$.

Par ailleurs, l'espace $\mathcal{C}[0,1]$, espace préhilbertien réel, n'est pas un espace normé complet en ce sens qu'une suite de Cauchy sur $\mathcal{C}[0,1]$ ne converge pas toujours vers un élément de $\mathcal{C}[0,1]$

On considère par exemple la suite de fonctions f_n (continues) définies par : les formules suivantes (cf figure)

$$f_n(x) \begin{cases} = 0 \text{ si } & x \leqslant \dfrac{1}{2} - \dfrac{1}{2n} \\ = 1 \text{ si } & x \geqslant \dfrac{1}{2} + \dfrac{1}{2n} \\ \text{linéaire entre } \dfrac{1}{2} - \dfrac{1}{2n} \text{ et } \dfrac{1}{2} + \dfrac{1}{2n} \end{cases}$$



La fonction $x \longrightarrow f_n(x)$ est continue et la suite f_n constitue une suite de Cauchy au sens de $\mathcal{C}[0,1]$ puisque pour $p \geqslant n$ (Cf Figure) :

$$\| f_n - f_p \|_2 \leqslant \frac{1}{\sqrt{n}}$$

Cependant, si la suite f_n convergeait vers une fonction continue au sens de la norme $\| \ \|_2$, $L(f_n)$ convergerait vers $L(f)$ pour toute forme linéaire continue L sur $\mathcal{C}[0, 1]$. En particulier $L(f) = \int_a^b f(x)dx$ pour $0 \leqslant a < b \leqslant 1$. Or, si $0 \leqslant a \leqslant b < 1/2$, on a $L(f_n) = 0$ pour n assez grand. Donc $0 = \int_a^b f(x)\,dx$, ce qui fournit par dérivation en b, $f(x) = 0$ pour $x \in [0, \frac{1}{2}[$ et par continuité pour $x \in [0, \frac{1}{2}]$. On aurait de même $f(x) = 1$ pour $x \in [1/2, 1]$. Mais $f(1/2) = 0 =$ ce qui est impossible et d'ailleurs, on note que f_n converge au sens de $\| \ \|_2$ vers f où $f(x) = 0$ si $0 \leqslant x < 1/2$ et $f(x) = 1$ si $1/2 < x \leqslant 1$. On peut même précise $f(1/2) = 1/2$ puisque $\lim_{n \to \infty} f_n(\frac{1}{2}) = \frac{1}{2}$, cependant cette valeur ne joue aucun rôle pour la convergence au sens de $\| \ \|_2$.

Ce dernier exemple permet d'assurer que les topologies normées définies sur $\mathcal{C}[0,1]$ par $\mathcal{C}[0,1]$ et $\mathfrak{C}[0,1]$ ne peuvent être équivalentes. La topologie de $\mathcal{C}[0,1]$ est <u>strictement plus fine</u> que la topologie de $\mathfrak{C}[0,1]$. En effet, s'il y avait équivalence des topologies, les deux espaces $\mathfrak{C}[0,1]$, et $\mathcal{C}[0,1]$ auraient les mêmes propriétés d'espaces normés et seraient ainsi, ou tous les deux complets, ou tous les deux non complets. Or, l'espace $\mathcal{C}[0,1]$ est complet, et l'espace $\mathfrak{C}[0,1]$ ne l'est pas!

II.3.3. <u>Espace L^2 [0,1] : intégrale de Lebesgue</u>.

Conformément à la théorie énoncée en II.3, proposons-nous de caractériser le plus petit espace normé complet contenant $\mathfrak{C}[0,1]$ comme sous-espace métrique. Nous cherchons donc l'espace de Hilbert complété du préhilbertien $\mathfrak{C}[0,1]$. On note $L^2[0,1]$ cet espace de Hilbert. Bien que tout élément f de $L^2[0,1]$ soit une limite en moyenne quadratique d'éléments f_n de $\mathfrak{C}[0,1]$ formant une suite de Cauchy dans $\mathfrak{C}[0,1]$, il ne paraît pas aisé de représenter concrètement les éléments de $L^2[0,1]$. On a cependant le théorème suivant, dont nous établirons les étapes essentielles.

<u>Théorème</u> :

<u>Soit f_n une suite de Cauchy dans $\mathfrak{C}[0,1]$. Il existe une sous-suite $\{f_{n_k}\}$ telle que $f_{n_k}(x)$ converge, en chaque point x n'appartenant pas à un certain ensemble exceptionnel E , vers une fonction $f(x)$</u> .

Partons de f_{n_0} pour n_0 assez grand. Pour f_{n_1} , choisissons la première fonction f_n , pour $n > n_0$, telle que :

$$\| f_{n_1} - f_{n_0} \|_2 \leqslant \frac{1}{2}$$

Plus généralement, f_{n_0}, f_{n_1}, ..., $f_{n_{k-1}}$ étant déterminées, choisissons pour f_{n_k} la première fonction, f_n, $n_k > n_{k-1}$ telle que :

$$\| f_{n_k} - f_{n_{k-1}} \|_2 \leqslant \frac{1}{2^k}$$

Ceci est possible puisque $\{f_n\}_{n \geqslant 1}$ constitue une suite de Cauchy dans $\mathcal{C}[0,1]$.

Formons alors la suite F_N, définie pour chaque valeur de x par

$$F_N(x) = \sum_{k=1}^{k=N} |f_{n_k}(x) - f_{n_{k-1}}(x)|$$

La suite de fonctions F_N (fonctions appartenant au préhilbertien $\mathcal{C}[0,1]$) est une suite croissante en chaque point x ($F_N(x) \leqslant F_{N+1}(x)$).

Ceci permet de définir une fonction $F(x)$ comme la limite

(1) $$F(x) = \lim_{N \to +\infty} F_N(x)$$

laquelle fonction F peut prendre des valeurs infinies. Montrons que ceci ne peut avoir lieu que sur un ensemble "petit", c'est-à-dire de mesure nulle, comme nous allons le préciser. Pour ce faire, on note que :

$$\| F_N(x) \|_2 \leqslant \sum_{k=1}^{k=N} \| f_{n_k}(x) - f_{n_{k-1}}(x) \|_2$$

d'après l'inégalité de Minkowski (propriété (3) de la norme).
Ce qui montre, qu'en moyenne quadratique, F_N reste bornée par 1

(2) $$\| F_N(x) \|_2 \leqslant \sum_{k=1}^{k=N} \frac{1}{2^k} \leqslant 1$$

Appelons donc E l'ensemble des points de $[0,1]$ où $F(x) = +\infty$.
On note par $E_{N,K}$ l'ensemble des points x de $[0,1]$ où $F_N(x) > K$ lorsque N et K sont des entiers positifs.

Il est clair avec $E_K = \{x \mid F(x) > K\} = \bigcup_{N=1}^{\infty} E_{N,K}$ que

$$E = \bigcap_{K=1}^{\infty} \left(\bigcup_{N=1}^{\infty} E_{N,K} \right) = \bigcap_{K=1}^{\infty} E_K$$

La continuité de la fonction F_N assure que le complémentaire de $E_{N,K}$ est un ensemble fermé (Cf. définition: I §B.5.4.1), ainsi d'ailleurs que le complémentaire de E_K. Un théorème assez facile de topologie sur $\mathbb{R}$, dont nous omettons cependant la démonstration, assure que tout complémentaire d'un fermé peut s'écrire comme une réunion, au plus dénombrable, d'intervalles ouverts disjoints $]a_n, b_n[$ avec $a_n < b_n$.(Cf Nanta Iremica, Vol 10). Ici

$$E_K = \bigcup_{N=1}^{\infty} E_{N,K} = \bigcup_{n=1}^{\infty} (]a_n \; b_n[)$$

L'intégrale $\int_{E_K} f(x)dx$ pour $f \in \mathcal{C}[0, 1]$ a un sens. On pose

$$\int_{E_K} f(x)dx = \sum_{n=1}^{\infty} \int_{a_n}^{b_n} f(x)dx.$$

En particulier

$$\int_{E_K} dx = \sum_{n=1}^{\infty} |b_n - a_n|$$

et ce nombre est appelé la <u>mesure</u> de E_K. De même pour $E_{N,K}$. (Il conviendrait, pour être cohérent, d'établir que ce nombre ne dépend pas de la décomposition de E_K). Lorsque $f \geqslant 0$ et $f \in \mathcal{C}[0, 1]$ on établit, puisque $E_{N,K} \subset [0, 1]$

$$\int_0^1 f(x)dx \geqslant \int_{E_{N,K}} f(x)dx$$

On en déduit Mesure $(E_{N,K}) \leqslant \frac{1}{K^2}$ puisque

$$1 \geqslant \int_0^1 F_N^2(x)dx \geqslant \int_{E_{N,K}} F_N^2(x)dx \geqslant K^2. \text{ Mesure } (E_{N,K}). \text{ En outre}$$

$E_{N,K} \subset E_{N+1,K} \subset \ldots \subset E_K$ et $E_K = \bigcup_{N=1}^{\infty} E_{N,K}$. Par suite, on admettra facilemen encore que ce soit moins aisé à prouver rigoureusement, que Mesure$(E_K) \leqslant \frac{1}{K^2}$

Donnons-nous maintenant un $\varepsilon > 0$ quelconque. Il existe un entier K tel que $\frac{1}{K^2} < \varepsilon$. Alors E est inclus dans E_K et E_K est une réunion dénombrable d'intervalles ouverts $]a_n , b_n[$ dont la somme totale des longueurs est inférieure à ε . On convient de la définition suivante :

définition : Un ensemble E est dit de mesure nulle si pour tout $\varepsilon > 0$ il existe une famille dénombrable d'intervalles $]a_n , b_n[$ telle que

$$E \subset \bigcup_{n=1}^{\infty}]a_n , b_n[$$

et

$$\sum_{n=1}^{\infty} |b_n - a_n| < \varepsilon$$

définition : Une propriété vraie partout sauf sur un ensemble de mesure nulle, est dite vraie presque partout.

Nous avons donc montré que $F(x)$ est une fonction presque partout finie. Or la fonction $F(x)$ permet, lorsqu'elle est finie, de définir une fonction f, en enlevant les valeurs absolues dans la série qui définit $F(x)$. De façon précise, la fonction f suivante est presque partout définie (et finie).

$$f(x) = \lim_{N \to +\infty} \sum_{k=1}^{k=N} \left(f_{n_k} - f_{n_{k-1}}\right) = \left(\lim_{N \to \infty} f_{n_N}(x)\right) - f_{n_0}(x)$$

La sous-suite $\{f_{n_k}(x)\}$ converge donc presque partout vers une fonction $f(x)$ et l'ensemble exceptionnel de l'énoncé est de mesure nulle (c.q.f.d). Voici qu'à tout élément de $L^2[0, 1]$,auquel est donc associée au moins une suite de Cauchy dans $\mathscr{C}[0, 1]$, nous associons une fonction f définie presque partout sur $[0, 1]$. Il importerait de démontrer que s'il existe une autre sous-suite $f_{n_{k'}}$ telle que $f_{n_{k'}}$ converge presque partout vers une fonction $f'(x)$, alors $f(x) = f'(x)$ du moins presque partout sur $[0,1]$. Nous admettrons ce point facile.

Dès lors, à toute suite de Cauchy $[f_n]_{n \geq 1}$ sur $\mathscr{C}[0,1]$, on associe une fonction $f(x)$, presque partout définie. Or, l'on sait que le complété de l'espace vectoriel normé $\mathscr{C}[0,1]$ est formé par les suites de Cauchy définies sur l'espace $\mathscr{C}[0\,1]$ où l'on identifie deux suites de Cauchy f_n et g_n telles que $\lim_{n \to \infty} \|f_n - g_n\|_2 = 0$. Il n'est pas difficile de montrer que deux suites de Cauchy équivalentes définissent deux fonctions f et g égales presque partout, donc presque partout, la même fonction.

Finalement, l'espace complété $L^2[0,1]$ s'identifie à un espace de fonctions définies presque partout, sur $[0, 1]$, et à valeurs réelles.

En outre, l'application $\mathcal{I} : f \longrightarrow \int_0 f(t)\,dt$ est une forme linéaire et continue sur $\mathcal{C}[0,1]$ puisque :

$$|I(f)| = \left|\int_0^1 f(t)\,dt\right| \leqslant \left(\int_0^1 |f(t)|\,dt\right)^{\frac{1}{2}} = \| f \|_2$$

Par le théorème de prolongement II.2, il existe une unique forme linéaire et continue $\overline{\mathcal{I}}$ prolongeant $\mathcal{I}$ à tout $L^2[0,1]$. Par définition $\overline{\mathcal{I}}$ s'appelle l'intégrale de Lebesgue de f sur $[0,1]$ et se note encore :

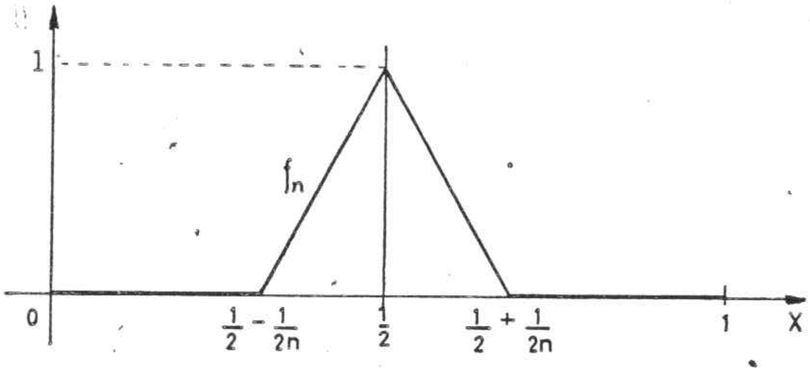
$$\overline{I}(f) = \int_0^1 f(t)\,dt$$

où f est une fonction définie presque partout et représentant un élément de $\mathcal{L}[0,$
Par définition du prolongement, l'intégrale de Lebesgue coïncide avec l'intégrale de Riemann pour les fonctions continues.

De fait, on peut donner une autre caractérisation de l'espace $L^2[0,1]$: ensemble des fonctions mesurables au sens de Lebesgue, définies presque partout et de carré intégrable au sens de Lebesgue, à condition de considérer comme égales deux fonctions égales presque partout. Nous y reviendrons après avoir donné un exemple et quelques résultats moins abstraits.

Exemple : la fonction $f(x) = 0$ si $x \neq \frac{1}{2}$ et 1 si $x = \frac{1}{2}$ est une fonction (mesurable) de carré intégrable, équivalente à la fonction 0. On voit bien d'ailleurs qu'il s'agit d'une fonction de $L^2[0,1]$ en la considérant comme la limite de la suite $\{f_n\}_{n\geqslant 1}$ de fonctions continues définie par :

$$f_n(x) \left| \begin{array}{ll} = 0 & \text{si } 0 < x \leqslant \frac{1}{2} - \frac{1}{2n} \quad \text{ou} \quad 1 \geqslant x > \frac{1}{2} + \frac{1}{2n} \\ = 1 & \text{si } x = \frac{1}{2} \\ \text{linéaire entre } \frac{1}{2} - \frac{1}{2n} \text{ et } \frac{1}{2} & \text{et entre } \frac{1}{2} \text{ et } \frac{1}{2} + \frac{1}{2n} \end{array} \right.$$



$\lim_{n \to \infty} f_n(x) = f(x)$ et par ailleurs f_n converge vers la fonction 0 au sens de l'espace $L^2[0,1]$ puisque $\|f_n\|_2 = \frac{1}{\sqrt{6n}}$ expression qui tend vers 0 lorsque $n \longrightarrow \infty$

Remarque : $\|f\|_2 = 0$ pour $f \in L^2[0,1]$ signifie $f = 0$ presque partout.

II.3.4. Théorème d'approximation.

Bien entendu, on a $\mathcal{C}[0,1] \subset L^2[0,1]$ et l'inclusion est stricte, (Cf. exemple précédent ou considération de fonctions telles que $\frac{1}{\sqrt[4]{x}}$ sur $]0,1]$). Mais nous avons un résultat d'approximation :

théorème : étant donnée une fonction f de $L^2[0,1]$, il existe une suite f_n de fonctions continues telle que

$$\lim_{n \to \infty} \|f_n - f\|_2 = 0$$

On peut aussi énoncer : $\mathcal{C}[0,1]$ est dense dans $L^2[0,1]$.

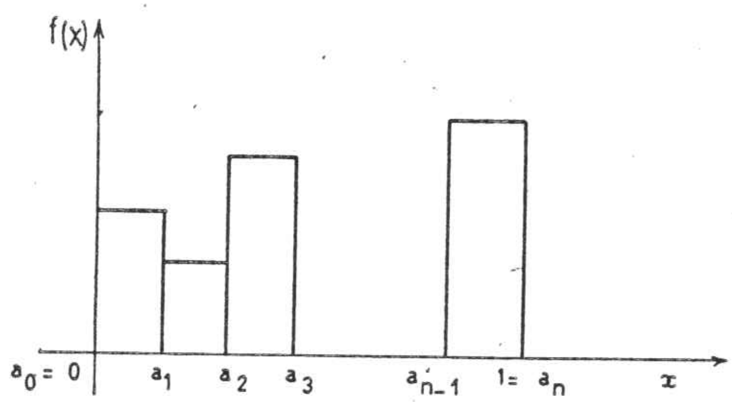
Ce résultat est d'une grande importance pratique et théorique ; il permet en fait d'effectuer les calculs sur $\mathcal{C}[0,1]$, et d'en déduire des résultats sur $L^2[0,1]$ par un passage à la limite ; de même que l'on peut obtenir de nombreuses propriétés de nombres réels à partir des propriétés des nombres rationnels. Malheureusement nous ne disposons pas encore d'un outil pratique pour construire f_n connaissant f. Ce sera l'outil de convolution.

Il est souvent commode, notamment dans les applications numériques, de remplacer les fonctions continues par une autre classe de fonctions : les fonctions en escalier.

On dit qu'une fonction f à valeurs réelles (ou complexes) est en escalier lorsqu'elle est de la forme :

$$f(x) = \sum_{i=0}^{i=n} \alpha_i \chi_i(x)$$

où les α_i sont des constantes et $\chi_i(x)$ est la fonction caractéristique de l'intervalle $[a_i, a_{i+1}]$, c'est-à-dire la fonction valant 1 sur $[a_i, a_{i+1}[$, 0 , ailleurs, lorsque $\{a_i\}$ est une subdivision de $[0,1]$.



L'ensemble $E\,[0,1]$ des fonctions en escalier est un sous-espace vectoriel de $\mathcal{F}\,[0,1]$. Il existe un théorème d'approximation d'une fonction continue au moyen de fonctions en escalier.

Théorème : étant donnée une fonction continue f , définie sur $[0,1]$, il existe une suite $\{f_n\}_{n\geq 1}$ de fonctions en escalier telle que

$$\boxed{\lim_{n \to \infty} \left[\sup_{x \in [0,1]} |f(x) - f_n(x)| \right] = 0}$$

c'est-à-dire que toute fonction continue sur $[0,1]$ est uniformément approchable par des fonctions en escalier.

La démonstration utilise exclusivement le théorème assurant qu'une fonction continue sur un intervalle $[0,1]$ est uniformément continue sur cet intervalle.

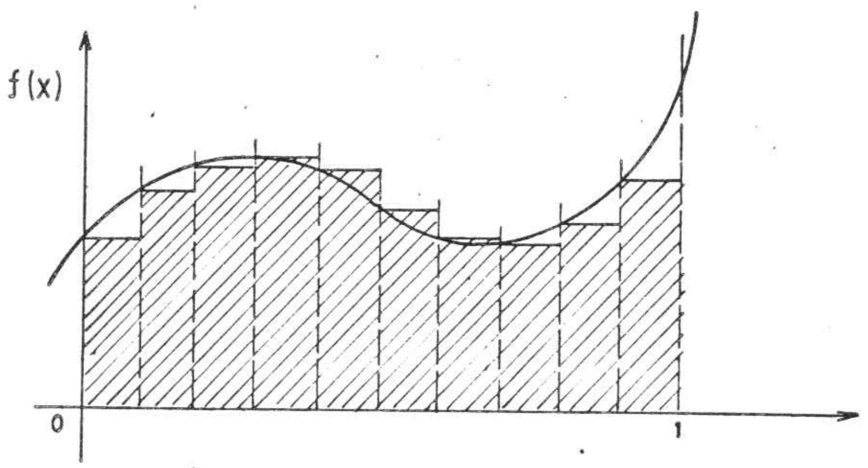
$$\forall \varepsilon > 0 \ \exists \eta > 0 \text{ tel que } |x_1 - x_2| < \eta \text{ entraîne}$$
$$|f(x_1) - f(x_2)| < \varepsilon$$

Fixons-nous $\varepsilon > 0$, il existe un nombre $\eta > 0$ et si l'on divise $[0,1]$ en n segments égaux $[a_i, a_{i+1}]$ de longueur $\eta = \frac{1}{n}$, la fonction :

$$f_n(x) = \sum_{i=1}^{i=n} \alpha_i \chi_i(x)$$

où $\alpha_i = f(a_i)$, est une fonction en escalier et

$$|f(x) - f_n(x)| < \varepsilon$$



Graphes de f et de f_n

On peut déduire de ce théorème un théorème d'approximation d'une fonction f de $L^2[0,1]$ par une fonction en escalier.

Théorème : Etant donnée une fonction f de $L^2[0\,1]$, il existe une suite $\{f_n\}$ de fonctions en escalier telle que

$$\lim_{n \to \infty} \| f - f_n \|_2 = 0$$

On peut encore énoncer :

$E[0,1]$ est dense dans $L^2[0,1]$

Ce théorème utilise le théorème précédent, le théorème d'approximation d'une fonction f de L^2 et le résultat II.3.1 assurant que sur $\mathbb{E}[0,1]$, $\mathfrak{C}[0,1]$ a une topologie moins fine que $\mathcal{C}[0,1]$.

En effet, soit f un élément de $L^2[0,1]$. D'après le théorème d'approximation II.3.3, il existe une suite g_k de fonctions continues telle que :

$$\lim_{k \to +\infty} \| f - g_k \|_2 = 0 .$$

D'après le théorème précédent, à toute fonction continue g_k, on peut faire correspondre une suite $g_{k,n}$ de fonctions en escalier telle que :

$$\sup_{x \in [0,1]} \left| g_k(x) - g_{k,n}(x) \right| \leq \frac{1}{n}$$

Mais on sait que pour une fonction f continue, ou en escalier, on a la relation :

$$\| f \|_2 \leqslant \sup_{x \in [0,1]} |f(x)|$$

On a donc :

$$\| g_k - g_{k,n} \|_2 \leq \frac{1}{n} .$$

Si l'on pose $g_{n,n} = f_n$, alors f_n est une fonction en escalier et

$$\lim_{n \to \infty} \| f - f_n \|_2 = 0$$

ce qui termine la démonstration.

II.3.5. Généralisations.

Soit $[a,b]$ un intervalle borné ou non de l'axe réel et $h(x)$ une fonction positive ($h(x) > 0$) sur ce segment. On peut définir de façon analogue à l'espace $L^2[0,1]$, un espace $L^2([a,b],h)$ (on doit supposer $h(x)$ mesurable).

C'est l'espace des fonctions (mesurables) définies sur $[a,b]$ et à valeurs complexes telles que :

$$\int_a^b |f(x)|^2 h(x)\, dx \qquad \text{existe}$$

La norme de cet espace vectoriel est naturellement :

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 h(x)\, dx\right)^{\frac{1}{2}}$$

Le produit scalaire dont dérive cette norme s'écrit :

$$<f, g> = \int_a^b f(x)\, \overline{g}(x)\, h(x)\, dx$$

Avec $a = 0$, $b = 1$ et $h \equiv 1$, on retrouve l'espace $L^2[0,1]$. On remarque encore que $\|f\|_2 = 0$ équivaut à $f = 0$ presque partout.

$L^2([a,b],h)$ est un espace de Hilbert et les théorèmes d'approximation II.3.4 sont encore exacts pour cet espace lorsque $[a,b]$ est un intervalle borné, et h une fonction intégrable.

Dans le cas d'un intervalle non borné, ce sont les fonctions continues ou les fonctions en escalier appartenant à l'espace $L^2([a,b],h)$ qu'il faut considérer pour avoir un résultat d'approximation.

II.4. COMPARAISON ENTRE DES TOPOLOGIES SUR $[(0,1)]$.

Sur l'espace $\mathbb{C}[0,1]$, nous connaissons au moins deux topologies normées, définissant des espaces vectoriels topologiques distincts :

$$\mathfrak{C}[0\ 1] \quad \text{pour la norme} \quad \|f\|_2 = \left(\int_0^1 |f(x)|^2 dx\right)^{\frac{1}{2}}$$

$$\mathcal{C}[0\ 1] \quad \text{pour la norme} \quad \|f\|_\infty = \sup_{x \in [0,1]} |f(x)|$$

Grâce à la relation $\|f\|_2 \leqslant \|f\|_\infty$ et d'après les résultats du paragraphe II.3.1, on constate que la topologie de $\mathcal{C}[0,1]$ est plus fine que la topologie de $\mathfrak{C}[0,1]$. On a même montré que cette topologie est strictement plus fine, c'est-à-dire qu'une relation de la forme

$$\|f\|_\infty \leqslant A\|f\|_2$$

est impossible.

Cependant, d'autres notions de convergence sont intéressantes sur l'espace $\mathcal{C}\,[0,1]$ et même si ces notions de convergence ne proviennent pas de l'existence d'une norme, il est important de faire des comparaisons. Sans insister sur la forme abstraite des résultats, nous allons donner quelques renseignements.

Soit f_n une suite de fonctions numériques définies sur $[0,1]$. On dit que $\{f_n\}$ <u>converge ponctuellement vers f</u> si pour tout $\varepsilon > 0$, et pour tout x fixé, il existe un nombre N tel que pour $n > N$ on ait :

$$|f_n(x) - f(x)| < \varepsilon$$

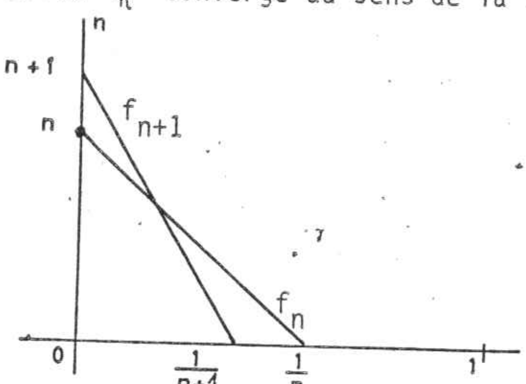
Bien entendu, cette convergence se distingue de la convergence uniforme par le fait qu'a priori N dépende de x. En particulier, la convergence uniforme implique la convergence ponctuelle.

Il n'y a pas de rapport simple entre la convergence ponctuelle et la convergence en moyenne quadratique, comme le montrent les exemples suivants :

<u>1er exemple</u> : soit f_n la fonction continue sur $[0,1]$, valant :

$$\begin{cases} 0 & \text{pour } \frac{1}{n} \leqslant x \leqslant 1 \\ n^2\left(\frac{1}{n} - x\right) & \text{pour } 0 \leqslant x \leqslant \frac{1}{n} \end{cases}$$

La suite f_n converge au sens de la topologie de la convergence ponctuelle sur $]0,1]$ vers la fonction nulle (elle ne converge pas au sens de la topologie de la convergence uniforme). Mais f_n ne peut pas converger au sens de la topologie de la convergence en moyenne quadratique sur $[0,1]$ puisque la norme $\|f_n - 0\|_2 = \|f_n\|_2$ vaut $\sqrt{\frac{n}{3}}$ et tend vers l'infini avec n.



<u>comparaison des topologies</u>

<u>2ème exemple</u> : <u>exemple de la bosse glissante</u> :

Ecrivons tout entier n naturel sous la forme $n = 2^h + k$ où h est un nombre entier positif et k un entier positif strictement inférieur à 2^h.

A tout entier n , on associe de façon unique un entier h et un entier k .

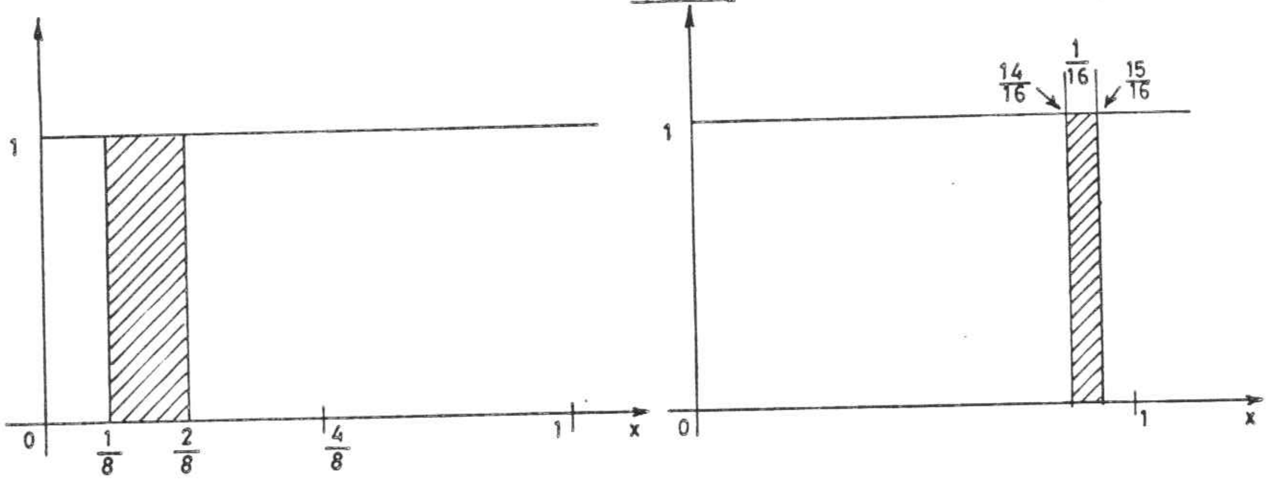
Par f_n , on désigne la fonction prenant sur $[0,1]$ les valeurs suivantes :

$$\begin{cases} f_n(x) = 1 \quad \text{si} \quad \dfrac{k}{2^h} \leqslant x \leqslant \dfrac{k+1}{2^h} \\ f_n(x) = 0 \quad \text{ailleurs} \qquad\qquad \text{(Cf. figures)} \end{cases}$$

<u>Fonction</u> $f_n(x)$ $\begin{cases} h = 3 \\ k = 1 \end{cases}$
$n = 2^3 + 1 = 9$

<u>Fonction</u> $f_n(x)$ $\begin{cases} h = 4 \\ k = 14 \end{cases}$
$n = 2^4 + 14 = 30$

<u>Figures</u>



La fonction $f_n(x)$ est une fonction ayant un seul créneau.

Lorsque n varie entre 2^h et $2^{h+1}-1$, le créneau de f_n garde la même largeur $\dfrac{1}{2^h}$ mais se déplace sur le segment $[0,1]$. Cependant lorsque n passe de la valeur $2^{h+1}-1$ à la valeur 2^{h+1} , la largeur du créneau diminue de

$$\frac{1}{2^h} \quad \text{à} \quad \frac{1}{2^{h+1}}$$

Ces remarques montrent que la suite $f_n(x)$ ne peut converger en aucun point x de $[0,1]$. En effet, N étant donné, il existe toujours deux entiers n_1 et n_2 supérieurs à N , tels que $f_{n_1}(x) = 0$ et $f_{n_2}(x) = 1$

Cependant, la suite f_n converge vers 0 au sens de la moyenne quadratique. Il suffit de calculer la norme quadratique de f_n :

$$\| f_n - 0 \| \leqslant \frac{1}{2^{h/2}}$$

Lorsque n tend vers l'infini, il en est de même pour h et donc $\|f_n\|_2$ converge vers 0 .

Le lecteur se convaincra facilement que l'on pourrait lisser les bords de f_n suffisamment pour que le contre-exemple de la bosse glissante soit donné avec des fonctions indéfiniment dérivables.
Bien que les convergences en moyenne quadratique et ponctuelle ne soient pas comparables, on dispose sur L^2 d'un théorème analogue au résultat du début du paragraphe II.3.3.

Théorème : De toute suite $\{f_n\}$ de fonctions de $L^2\,[0,1]$, convergeant en moyenne quadratique vers une fonction f , on peut extraire une sous-suite f_{n_k} qui converge presque partout vers la fonction f .

Pour effectuer la démonstration, on peut reprendre le raisonnement du paragraphe II.3.3 mais, au lieu d'utiliser l'ensemble $E_{N,K}$, on doit faire intervenir un résultat particulièrement important concernant l'intégrale de Lebesgue. Ce résultat est le suivant (avec les notations de II § 5), à partir duquel le théorème précédent est facile.

Théorème de Beppo-Levi : soit f_k une suite croissante de fonctions intégrables au sens de Lebesgue telles que les intégrales

$$\int_0^1 f_k(x)\,dx$$

restent bornées par un même nombre. Alors la fonction :

$$x \rightsquigarrow f(x)$$

où

$$f(x) = \lim_{k \to \infty} f_k(x)$$

est une fonction intégrable satisfaisant :

$$\lim_{k \to \infty} \int_0^1 f_k(x)\,dx = \int_0^1 f(x)\,dx < +\infty$$

Nous admettrons le théorème de Beppo-Levi, dont la démonstration n'est pas très difficile en utilisant le fait que $L^2[0,1]$ est complet. Nous fournirons plus loin cette démonstration (cf II 6.1.2.)

remarque : considérons le cas de la bosse glissante. En choisissant une sous-suite de $\{f_n\}$ déterminée par les valeurs de n pour lesquelles $k = 0$, par exemple, on constate sa convergence ponctuelle.

II.5. ESPACES $L^p[0,1]$ ET INEGALITES DE CONVEXITE.

On peut munir l'espace $\mathcal{C}[0\ 1]$ de multiples autres normes intéressantes. Nous allons définir des normes analogues aux normes L^2 . Nous avons cependant besoin de quelques résultats préliminaires de convexité concernant les inégalités célèbres de Hölder et de Minkowski.

Lemme 1 : la fonction $x \longrightarrow e^x$ est une fonction convexe.

Ceci signifie que pour deux nombres réels x_1 et x_2 , on a

(1) $$\exp\left(\frac{x_1 + x_2}{2}\right) \leqslant \frac{e^{x_1} + e^{x_2}}{2}.$$

Si l'on pose $e^{x_1} = a$ et $e^{x_2} = b$ cette inégalité devient :

$$\sqrt{a\,b} \leqslant \frac{a + b}{2}$$

Inégalité fondamentale et bien connue (comparaison de la moyenne arithmétique et de la moyenne géométrique).

Pour une fonction continue, on sait que (1) entraîne :

(2) $$\exp\,(\lambda\, x_1 + (1-\lambda)x_2) \leqslant \lambda e^{x_1} + (1-\lambda)e^{x_2} \qquad \text{où } 0 < \lambda < 1$$

Soit p un nombre réel appartenant à $[1, +\infty]$. On appelle conjugué de p le nombre q tel que

$$\frac{1}{q} + \frac{1}{p} = 1$$

Si $p = 1$ alors $q = +\infty$

Lemme 2 : soient a et b deux nombres positifs. On a l'inégalité suivante

$$(3) \qquad a\, b \leqslant \frac{a^p}{p} + \frac{b^q}{q}$$

où p et q sont deux nombres réels conjugués avec $+\infty > p > 1$

Il suffit de poser :

$$\begin{cases} x_1 = p \log a \\ x_2 = q \log b \end{cases}$$

et d'appliquer l'inégalité (2) où $\lambda = \frac{1}{p}$ et $1-\lambda = \frac{1}{q}$ grâce à la relation de conjuguaison (le cas $p = 1$ est inexact car alors q est infini).

Lemme 3 : inégalité de Hölder

Soit p un nombre réel tel que $1 < p < +\infty$ et q son conjugué.

Pour deux fonctions positives et continues f et g définies sur $[0,1]$, on a l'inégalité suivante :

$$(4) \qquad \boxed{\int_0^1 f(x)\, g(x)\, dx \leqslant \left(\int_0^1 f^p(x)\, dx\right)^{\frac{1}{p}} \cdot \left(\int_0^1 g^q(x)\, dx\right)^{\frac{1}{q}}}$$

Démonstration : Sauf si les dénominateurs sont nuls, auquel cas l'inégalité est facile (Cf I.A § 4.5), posons

$$f_1(x) = \frac{f(x)}{\left(\int_0^1 f^p(x)\, dx\right)^{\frac{1}{p}}} \quad \text{et} \quad g_1(x) = \frac{g(x)}{\left(\int_0^1 g^q(x)\, dx\right)^{\frac{1}{q}}}$$

Grâce à l'inégalité (3), on a :

$$f_1(x)\, g_1(x) \leqslant \frac{f_1^p(x)}{p} + \frac{g_1^q(x)}{q}$$

Soit en intégrant entre 0 et 1 :

$$\int_0^1 f_1(x)\, g_1(x)\, dx \leqslant \frac{1}{p}\int_0^1 f_1^p(x)\, dx + \frac{1}{q}\int_0^1 g_1^q(x)\, dx \qquad \frac{1}{p} + \frac{1}{q} = 1$$

Finalement, en revenant aux fonctions f et g, on obtient l'inégalité (4), dite inégalité de Hölder.

Lemme 4 : inégalité de Minkowski

Soit p un nombre réel appartenant à $[1,+\infty[$. On a pour deux fonctions continues et positives f et g sur $[0,1]$, l'inégalité suivante :

$$(5) \qquad \left(\int_0^1 (f(x)+g(x))^p dx\right)^{\frac{1}{p}} \leqslant \left(\int_0^1 f^p(x)\, dx\right)^{\frac{1}{p}} + \left(\int_0^1 g^p(x)\, dx\right)^{\frac{1}{p}}$$

Démonstration :

1er cas : $p = 1$. L'inégalité est évidente avec des fonctions positives

2e cas : $1 < p < +\infty$. On utilise l'inégalité (4) de Hölder selon le procédé suivant. On décompose d'abord $(f+g)^p$:

$$(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1}$$

Puis on applique (4) aux deux expressions du 2ème membre selon

$$\int_0^1 f(x)\left(f(x)+g(x)\right)^{p-1} dx \leqslant \left(\int_0^1 f^p(x)\, dx\right)^{\frac{1}{p}} \left(\int_0^1 \left(f(x)+g(x)\right)^{(p-1)q} dx\right)^{\frac{1}{q}}$$

$$\text{et} \quad \int_0^1 g(x)\left(f(x)+g(x)\right)^{p-1} dx \leqslant \left(\int_0^1 g^p(x) dx\right)^{\frac{1}{p}} \left(\int_0^1 \left(f(x)+g(x)\right)^{(p-1)q} dx\right)^{\frac{1}{q}}$$

Or $\frac{1}{p} + \frac{1}{q} = 1$, ce qui entraîne $(p-1)q = p$. D'où en additionnant

$$\int_0^1 \left(f(x)+g(x)\right)^p dx \leqslant \left(\int_0^1 \left(f(x)+g(x)\right)^p dx\right)^{\frac{1}{q}} \left[\left(\int_0^1 f^p(x)\, dx\right)^{\frac{1}{p}} + \left(\int_0^1 \left(g(x)\right)^p dx\right)^{\frac{1}{p}}\right]$$

ce qui donne bien l'inégalité (5) dite inégalité de Minkowski.

Remarque : une analyse facile montre que l'inégalité a lieu dans (4) si, et seulement si, il existe une constante λ telle que :

$$f^p(x) = \lambda\, g^q(x) \qquad (\ f \text{ et } g \text{ sont positives, } g \not\equiv 0\ )$$

De même l'inégalité a lieu dans (5) si, et seulement si, il existe une constante λ telle que

$$f(x) = \lambda g(x) \qquad (f \text{ et } g \text{ positives, } g \neq 0)$$

Théorème 1 : l'expression $\left(\int_0^1 |f(x)|^p dx\right)^{\frac{1}{p}}$ constitue une norme sur l'espace $\mathfrak{C}[0,1]$, pour $1 \leq p < +\infty$

Il suffit d'appliquer (5) à $|f|$ et $|g|$ pour l'inégalité triangulaire. On note $\mathfrak{C}_p[0,1]$ l'espace vectoriel normé ainsi constitué.

Comme l'espace $\mathfrak{C}[0,1]$ du paragraphe II.3.1, l'espace $\mathfrak{C}_p[0,1]$ n'est pas un espace complet (en outre $\mathfrak{C}_p[0,1]$, pour $p \neq 2$ n'est pas un espace préhilbertien).

Suivant alors les directives de II.3.3, c'est-à-dire l'exemple de l'espace $\mathfrak{C}[0,1]$, on peut définir le plus petit espace normé complet contenant $\mathfrak{C}_p[0,1]$ comme sous-espace dense. On appelle $L_p[0,1]$ l'espace complété ainsi constitué. C'est un espace de Banach.

De la même manière qu'en II.3.2, on démontre que $L_p[0,1]$ est un espace de classes de fonctions : c'est l'espace des fonctions mesurables et de puissance $p^{\text{ième}}$ intégrable, en considérant comme équivalentes deux fonctions qui sont égales presque partout.

Pour ces espaces $L_p[0,1]$, les théorèmes du paragraphe II.3.3 sont encore valables. De même, on a :

$\mathfrak{C}_p[0,1]$ est dense dans $L_p[0,1]$ Cf II.3.3

$E[0,1]$ est dense dans $L_p[0,1]$ Cf II.3.3

Ceci montre qu'il convient de travailler sur $\mathfrak{C}_p[0,1]$ ou $E[0,1]$ puis de passer à la limite, pour obtenir des résultats constructifs sur $L_p[0,1]$. En fait, on peut comparer les divers espaces obtenus :

Théorème 2 : on a les résultats suivants :

1°) la topologie de $\mathfrak{C}_p[0,1]$ est moins fine que la topologie de $\mathfrak{C}_{p'}[0,1]$ lorsque $p < p'$.

2°) <u>L'espace</u> $L_p[0,1]$ <u>contient l'espace</u> $L_{p'}[0,1]$ <u>lorsque</u>

$p < p'$

Démonstration :

Etablissons l'inégalité suivante pour une fonction f de $\mathbb{I}[0,1]$

(6) $$\|f\|_p \leqslant \|f\|_{p'} \quad \text{pour} \quad p' > p$$

On utilise l'inégalité (4) de Hölder, appliquée à $|f|^p$ et $g = 1$

$$\int_0^1 |f(x)|^p dx = \int_0^1 |f(x)|^p g(x)\,dx \leqslant \left(\int_0^1 |f(x)|^{p\alpha} dx\right)^{\frac{1}{\alpha}} \left(\int_0^1 dx\right)^{\frac{1}{\beta}}$$

avec $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ et $\alpha > 1$

Choisissons α tel que $p\alpha = p'$ ($p' > p$ donc $\alpha > 1$). On a

$$\int_0^1 |f(x)|^p dx \leqslant \left(\int_0^1 |f(x)|^{p'} dx\right)^{\frac{p}{p'}}$$

Soit l'inégalité (6) et conformément à la terminologie adoptée, ceci termine la démonstration de la première partie du théorème.

2°) Par densité, l'inégalité (6) reste valable pour une fonction quelconque de $L_p[0,1]$ d'où la deuxième assertion.

Il est donc clair que $L^1[0,1]$ est le plus grand des espaces $L^p[0,1]$ pour $p \geqslant 1$. <u>L'espace $L^1[0,1]$ s'appelle l'espace des fonctions intégrables au sens de Lebesgue</u> .

$\|f\|_1 = 0$ <u>signifie</u> $f = 0$, <u>presque partout</u>.

On se reportera à II § 6 pour des compléments sur $L^1[0,1]$.

<u>Notation</u> : on convient fréquemment de noter par $\|F\|_\infty$ la norme de la topologie de la convergence uniforme sur $\mathbb{I}[0,1]$, c'est-à-dire la norme de $\mathscr{C}[0,1]$:

$$\|F\|_\infty = \operatorname{Sup}_{x \in [0,1]} |F(x)|$$

Un résultat limite permet d'ailleurs de justifier cette notation (Ex. N° 23). Rappelons donc ici les résultats obtenus pour une fonction de $\mathbb{C}\,[0,1]$

$$\|F\|_1 \leqslant \|F\|_p \leqslant \|F\|_{p'} \leqslant \|F\|_\infty \quad \text{où} \quad +\infty > p' > p > 1$$

Toutefois, il est intéressant d'introduire l'espace des fonctions bornées contenues dans $L^2\,[0,1]$ mais non nécessairement continues. Le procédé qui consiste à compléter $\mathbb{C}\,[0\;1]$ pour la norme $\|.\|_\infty$ ne produit que $\mathcal{C}\,[0,1]$. On convient alors de noter par $L^\infty\,[0,1]$ l'espace des fonctions f définies sur $[0,1]$ et appartenant à $L^2\,[0,1]$ telles qu'il existe un ensemble de mesure nulle $E \subset [0,1]$ et que f soit bornée sur le complémentaire de E dans $[0,1]$.

Comme norme, on prend $\|f\|_\infty = \inf_E \left(\sup_{x \in [0,1]\setminus E} |f(x)| \right)$

où E parcourt tous les sous-ensembles de mesure nulle de $[0,1]$. On appelle $\|f\|_\infty$ le <u>Sup essentiel</u> de f dans $[0,1]$ et toute la complication vient de ce que f n'est pas définie partout, mais seulement presque partout. Toutefois, l'inégalité suivante reste valable

$$\left| \int_0^1 f(x)\, g(x)\, dx \right| \leqslant \|f\|_1 \, \|g\|_\infty$$

si $f \in L^1[0,1]$ et $g \in L^\infty[0,1]$

On conserve également

$$\|f\|_p \leqslant \|f\|_\infty \text{ pour tout } p,\ 1 \leqslant p < \infty$$

On laisse au lecteur le soin de vérifier, avec les notations de ce Chapitre II, § 3.5, que les inclusions suivantes subsistent sur un segment $[a,b]$ <u>borné</u> quelconque de $\mathbb{R}$:

$$L^1\,[a,b] \supset L^p\,[a,b] \supset L^{p'}\,[a,b] \supset L^\infty\,[a,b]$$

pour $1 < p < p' < \infty$, ainsi que le théorème 2. On définit aussi $L^p(\mathbb{R})$ lorsque $a = -\infty$ et $b = +\infty$. Cet espace peut apparaître comme le complété au sens de la norme $\|f\|_p = \left(\int_{-\infty}^{+\infty} |f(x)|^p\, dx\right)^{1/p}$ de l'espace des fonctions continues $\mathbb{R} \to \mathbb{C}$, nulles en dehors d'un sous-ensemble borné de $\mathbb{R}$, sous-ensemble qui dépend de la fonction considérée.

Par contre, il n'y a aucune inclusion, entre les divers espaces $L^p(\mathbb{R})$ dans un sens ou dans l'autre, lorsque l'on prend $a = -\infty$ ou $b = +\infty$. Il est indispensable que le lecteur vérifie ces faits en exerçant sa propre imagination.

II.6. INTEGRATION AU SENS DE LEBESGUE.

II.6.1. Théorèmes fondamentaux de Lebesgue.

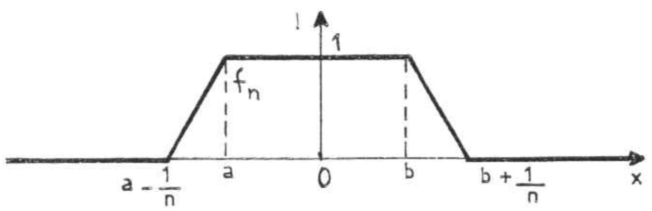
L'introduction du § 3.2 à l'intégrale de Lebesgue, pour économique qu'elle soit, est beaucoup trop abrupte et sèche et nous allons faire ici quelques commentaires rassurants.

Comme nous l'avons vu, un élément de $L^p\,[a,b]$ s'identifie à une classe de fonctions, la relation d'équivalence étant l'égalité presque partout. En outre, un représentant f d'une telle classe est limite presque partout d'une suite de fonctions continues (ou en escalier) : on dit qu'une telle fonction est mesurable au sens de Lebesgue. Pratiquement toute fonction rencontrée est mesurable. Pour vérifier alors qu'une fonction donnée f mesurable est intégrable au sens de Lebesgue, divers moyens sont possibles et nous allons donner des exemples.

II.6.1.1. Un moyen est de montrer que f est la limite ponctuelle (presque partout) d'une suite de Cauchy au sens $L^1\,[0,1]$ par exemple de fonctions continues. Ainsi f définie par

$$\left| \begin{array}{lll} f(x) = \dfrac{1}{\sqrt{x}} & x \neq 0 & x \in \,]0,1] \\ f(0) \text{ quelconque} & & \end{array} \right.$$

est limite presque partout de f_n où $f_n(x) = \dfrac{1}{\sqrt{x}}$ pour $x \in [\frac{1}{n}, 1]$ et $f_n(x) = \sqrt{n}$ pour $x \in [0, \frac{1}{n}]$. C'est une suite de Cauchy au sens de $L^1\,[0,1]$
Mais f, mesurable, a une "taille" trop grande pour appartenir à $L^2\,[0,1]$
Un autre exemple est la fonction caractéristique d'un intervalle borné $[a,b]$ qu'on approche par la suite f_n représentée ci-dessous (dans $L^1\,(\mathbb{R})$).



Plus généralement, une réunion finie d'intervalles bornés détermine une fonction de $L^1(\mathbb{R})$.

II.6.1.2. Un autre moyen est d'utiliser le théorème de Beppo-Lévi. Profitons en pour donner la démonstration de ce théorème. On calcule l'expression de Cauchy selon :

$$\| f_{k+p} - f_k \|_1 = \int_0^1 | f_{k+p}(x) - f_k(x) | dx = \int_0^1 (f_{k+p}(x) - f_k(x)) dx$$

Posons $M = \operatorname{Sup}_k \int_0^1 f_k(x) dx$ Par hypothèse M est finie et donc

$\| f_{k+p} - f_k \|_1 \leqslant M - \int_0^1 f_k(x) dx$ qui tend vers 0 lorsque k tend vers l'infini. Par suite, $\{f_k\}_{k \geqslant 1}$ est une suite de Cauchy dans $L^1[0,1]$ et elle converge ponctuellement (croissance de $f_k(x)$ en k pour chaque x fixé) vers une fonction f

$f(x) = \lim_{k \to \infty} f_k(x) = \operatorname{Sup}_k f_k(x)$. Donc, par construction, $f \in L^1[0,1]$ et par définition de l'intégrale de Lebesgue.

$$\lim_{k \to \infty} \int_0^1 f_k(x) dx = \int_0^1 f(x) dx$$

La démonstration se fait de façon identique dans $L^p([a\ b]$, $\infty > p \geqslant 1$ et pour un ensemble $[a, b]$, borné ou non. On dispose naturellement d'un théorème analogue pour une famille décroissante f_n telle que $\int f_n dx$ reste minoré.

II.6.1.3. Encore un autre moyen particulièrement efficace est le théorème suivant, dû à Lebesgue (Notations du § 3.5).

<u>Théorème de la convergence dominée de Lebesgue</u> :

<u>Soit</u> $\{f_n\}_{n \geqslant 1}$ <u>une suite de fonctions appartenant à</u> $L^1([a,b],h)$ <u>qui converge presque partout vers une fonction</u> f. <u>Supposons qu'il existe une fonction</u> g <u>dans</u> $L^1([a,b]\ h)$ <u>telle que</u>

$$| f_n(x) | \leqslant g(x) \qquad \text{presque partout dans } [a, b].$$

<u>Alors</u> f <u>appartient à</u> $L^1([a,b],h)$ <u>et</u>

$$\lim_{n \to \infty} \int_a^b f_n(x) h(x) dx = \int_a^b f(x) h(x) dx$$

La démonstration est simple en appliquant deux fois le théorème de Beppo-Lévi aux deux suites de fonctions

$$g_n(x) = \operatorname{Sup}_{k \geqslant n} f_k(x) \qquad h_n(x) = \operatorname{Inf}_{k \geqslant n} f_k(x)$$

En supposant f_k réel, g_n est décroissante, h_n est une suite croissante et on dispose par hypothèse de $\lim_{n\to\infty} f_n = \lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} h_n(x)$ presque partout.

(Le cas f_k complexe n'est pas difficile à déduire).

Ce théorème de la convergence dominée de Lebesgue est d'un emploi particulièrement simple. Donnons un exemple. Posons $f_n(x) = (1 - \frac{x}{n})^n \chi_n(x)$ où $\chi_n(x)$ est la fonction caractéristique de $[0,n]$. Alors on obtient, en prenant les logarithmes :

$$|f_n(x)| \leqslant e^{-x} \qquad \forall x \geqslant 0$$

Par suite

$$\lim_{n\to\infty} \int_0^n (1 - \frac{x}{n})^n dx = \int_0^\infty e^{-x} dx = 1$$

II.6.1.4. Une autre application est particulièrement utile et nous l'énonçons sous forme d'un théorème :

Théorème : Soit $f(x,t)$ une fonction de deux variables x et t. On suppose que x appartienne à $[a,b]$ (borné ou non) et que t appartienne à $[\alpha,\beta]$

1) Supposons que f soit continue en t sur $[\alpha,\beta]$ pour presque tout x sur $[a,b]$ et qu'il existe g dans $L^1[a\ b]$ telle que $|f(x,t)| \leqslant g(x)$ presque partout.

Alors

$$h(t) = \int_a^b f(x,t)\,dx$$

est une fonction continue en t sur $[\alpha,\beta]$

2) Supposons que

(a) $\frac{\partial f}{\partial t}(x,t)$ existe presque partout en x et pour tout $t \in [\alpha,\beta]$

(b) $\left| \frac{\partial f}{\partial t}(x,t) \right| \leqslant F(x)$ pour tout $t \in [\alpha,\beta]$ et presque partout en x.

Ici F appartient à $L^1[a,b]$

(c) $f(x,t)$ appartient pour tout $t \in [\alpha,\beta]$ à $L^1[a,b]$.

Alors $h(t) = \int_a^b f(x,t)\,dx$ est une fonction continue et dérivable, et

$$\frac{dh}{dt}(t) = \int_a^b \frac{\partial f}{\partial t}(x,t)\,dx \quad \text{pour tout } t \in [\alpha,\beta]$$

Démonstration : Pour abréger, on ne démontrera que (1). On prend une suite de points de $[\alpha,\beta]$ qui converge vers un t donné de $[\alpha,\beta]$. On pose

$$f_n(x) = f(t_n,x)$$

On a $|f_n(x)| \leq g(x)$ et $\lim_{n\to\infty} f_n(x) = f(t,x)$ presque partout en x. Par suite, grâce au théorème de la convergence dominée de Lebesgue.

$$h(t_n) = \int_a^b f_n(x)\,dx \qquad \text{converge vers} \quad \int_a^b f(t,x)dx = h(t)$$

c'est-à-dire que h est continue sur $[\alpha,\beta]$.

II.6.1.6. Pour une fonction continue, l'intégrale de Lebesgue coïncide avec l'intégrale de Riemann. En fait, on peut montrer que toute fonction intégrable au sens de Riemann, sur $[0,1]$ (au sens ordinaire, sinon voir § I.1.8) est intégrable au sens de Lebesgue. On peut même démontrer le très remarquable théorème suivant, qui n'est pas très difficile en utilisant ce qui précède.

Théorème de Lebesgue: Une fonction $f \in L^1[0,1]$ est intégrable au sens de Riemann si et seulement si l'ensemble de ses points de discontinuité est de mesure nulle.

II.6.1.7. Théorie de la mesure : Soit $\mathcal{T}$ la famille des sous-ensembles E de $[0,1]$ tels que la fonction caractéristique de E appartienne à $L^1[0,1]$. Il est facile de voir que $\mathcal{T}$ a les propriétés suivantes en utilisant le théorème de la convergence dominée

(1) $[0,1] \in \mathcal{T}$, $\phi \in \mathcal{T}$

(2) si $E_n \in \mathcal{T}$, alors $\bigcup_{n=1}^{\infty} E_n \in \mathcal{T}$

(3) si $E_n \in \mathcal{T}$, alors $\bigcap_{n=1}^{\infty} E_n \in \mathcal{T}$

(4) si $E \in \mathcal{T}$, alors $C E \in \mathcal{T}$

Tous les ensembles de mesure nulle appartiennent à $\mathcal{T}$ qui s'appelle la tribu de Lebesgue de $[0,1]$. En outre, on pose

$$m(E) = \int_0^1 \chi_E(x) \qquad \text{si } E \in \mathcal{T}$$

L'application m , dite mesure de Lebesgue, vérifie les trois propriétés :

(1) $m([a,b]) = b-a \qquad 1 \geqslant b \geqslant a \geqslant 0$

(2) $m(E_1 \cup E_2) + m(E_1 \cap E_2) = m(E_1) + m(E_2)$ pour $E_1, E_2 \in \mathscr{C}$

(3) $m(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m(E_n)$ si $[E_n]_{n\geqslant 1}$ est une suite d'ensembles de $\mathscr{C}$ deux à deux disjoints.

On voit qu'on a réussi à généraliser la notion de longueur d'un segment à tout ensemble de la tribu de Lebesgue. Bien sûr, si l'on donne une autre valeur compatible à la longueur d'un segment $[a,b]$, on imagine que l'on peut générer d'autres mesures : c'est la théorie de la mesure sur laquelle nous n'insisterons pas.

II.6.1.8. <u>Une remarque notable</u> : On démontre en analyse que (Dirichlet) :

$$\lim_{A\to\infty} \int_0^A \frac{\sin x}{x}\, dx = \frac{\pi}{2}$$

où l'intégrale est prise au sens de Riemann. On convient de noter $\int_0^{+\infty} \frac{\sin x}{x}\, dx = \frac{\pi}{2}$, ce que l'on appelle une <u>intégrale de Riemann généralisée</u>.
<u>Mais au sens de Lebesgue</u>

<u>l'intégrale</u> $\int_0^{\infty} \frac{\sin x}{x}$ <u>n'a pas de sens</u>.

En fait, si f est une fonction intégrable au sens de Lebesgue, il en est de même pour $|f|$. Ceci est facile à voir grâce à notre introduction par les suites de Cauchy puisque si $\{f_n\}_{n\geqslant 1}$ est de Cauchy, ainsi en est-il de $\{|f_n|\}_{n\geqslant 1}$ car

$$||f_n| - |f_m|| \leqslant |f_n - f_m|$$

Or $\int_0^{\infty} \left|\frac{\sin x}{x}\right| dx$ diverge !

Par suite, pour les intégrales de Riemann généralisées de fonctions de signe non constant, il faut se méfier.

(<u>Remarque</u> : le résultat de Dirichlet peut s'obtenir en dérivant en t la fonction $\int_0^{\infty} e^{-xt} \frac{\sin x}{x}\, dx$ et en vérifiant la validité des dérivations formelles effectuées).

II.6.2. Intégration de Lebesgue dans $\mathbb{R}^n$.

Considérons l'ensemble $C_{oo}(\mathbb{R}^n)$ des fonctions numériques continues définies sur $\mathbb{R}^n$ et nulles en dehors d'un ensemble borné (lequel dépend de la fonction considérée). On munit $C_{oo}(\mathbb{R}^n)$ d'une norme

$$\|f\|_1 = \int \cdots \int_{\mathbb{R}^n} |f(x_1, \ldots x_n)| dx_1 \ldots dx_n$$

où l'intégrale est prise au sens usuel de Riemann à n dimensions. On complète l'espace $C_{oo}(\mathbb{R}^n)$ pour obtenir un espace $L^1(\mathbb{R}^n)$. Les méthodes précédentes s'adaptent pour montrer que $L^1(\mathbb{R}^n)$ s'identifie à un ensemble de fonctions, mais où l'on considère comme équivalentes deux fonctions f et g égales presque partout. L'égalité presque partout signifie l'égalité partout, sauf sur un ensemble de mesure nulle. Un ensemble E de $\mathbb{R}^n$ est de mesure nulle si pour tout $\varepsilon > 0$, il existe une famille dénombrable de boules B_k dans $\mathbb{R}^n$ pour lesquelles on dispose de

$$E \subset \bigcup_{h=1}^{\infty} B_h \qquad \text{et} \quad \sum_{h=1}^{\infty} |B_h| \leqslant \varepsilon$$ où $|B_h|$ désigne le "volume" de B_k ,

soit $|B_k| = \int \cdots \int_{B_k} dx_1 \ldots dx_n$

Un ensemble de mesure nulle peut être très "bizarre" et contenir un "continu" de points.

Sur $L^1(\mathbb{R}^n)$, les théorèmes de Beppo-Lévi, de la convergence dominée de Lebesgue, de dérivation etc., restent exacts aux seuls changements de notation près. Pour un hypercube K , une boule ou un pavé de $\mathbb{R}^n$, on peut aussi définir $L^1(K)$. D'ailleurs on pourrait dire que $f \in L^1(K)$ si et seulement si $f\chi_K \in L^1(\mathbb{R}^n)$ où χ_K est la fonction caractéristique de K .

II.6.3. Théorème de Fubini.

Donnons dans $\mathbb{R}^2$, le théorème fondamental concernant l'intégration double. Sa généralisation à $\mathbb{R}^n$ est immédiate.

Théorème de Fubini : Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ (ou $\mathbb{C}$) une fonction appartenant à $L^1(\mathbb{R}^2)$. Alors les fonctions

$$f_1 : \quad y \longrightarrow \int_{\mathbb{R}} f(x,y)dx$$
$$f_2 : \quad x \longrightarrow \int_{\mathbb{R}} f(x,y)dy$$

sont définies presque partout et appartiennent à $L^1(\mathbb{R})$. De plus

$$\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = \int_{-\infty}^{+\infty} dx \left[\int_{-\infty}^{+\infty} f(x,y)\,dy \right] = \int_{-\infty}^{+\infty} dy \left[\int_{-\infty}^{+\infty} f(x,y)\,dx \right]$$

Nous ne ferons pas la démonstration, mais elle n'est pas très difficile en utilisant le théorème de la convergence dominée et le théorème de permutation des intégrales dans une intégrale double (au sens de Riemann) d'une fonction continue définie sur un domaine borné du plan.

Remarque : Le théorème de Lebesgue, valable dans $L^1(\mathbb{R}^n)$ fournit une caractérisation des fonctions intégrables au sens de Lebesgue sur $\mathbb{R}^n$.

Théorème : Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction mesurable. S'il existe $g \in L^1(\mathbb{R}^n)$ et $|f| \leq g$ (presque partout), alors f appartient à $L^1(\mathbb{R}^n)$

Corollaire : Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction mesurable. Cette fonction est intégrable au sens de Lebesgue si et seulement si $|f|$ l'est.

EXERCICES SUR LE CHAPITRE II.

EXERCICE N° 16.

Soit E un espace de Hilbert et F un sous-espace vectoriel de E
F est muni d'une structure préhilbertienne en tant que sous-espace de F .
Etablir l'assertion suivante :

Pour que F soit complet, il faut et il suffit que F soit fermé.

EXERCICE N° 17.

Soit l'espace vectoriel $\mathbb{R}^n$ muni d'une norme quelconque.
Soit $[A]$ une matrice réelle $n \times n$. On définit la matrice

$$S_n = [I] + t\,[A] + \cdots\cdots + t^n \frac{[A]^n}{n!}$$

où $[I]$ désigne la matrice identité et $[A]^n$ la puissance $n^{\text{ième}}$ de la matrice $[A]$.

1°) montrer que pour chaque vecteur x et chaque valeur t , la suite de vecteurs $S_n(x)$ converge vers un vecteur noté $e^{t[A]}(x)$

2°) $e^{t[A]}$ définit-il une matrice ?

Montrer que $e^{(t_1+t_2)[A]} = e^{t_1[A]}\, e^{t_2[A]}$

3°) calculer $e^{t[A]}$ pour

$$[A] = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

(on trouvera : $e^{tA} = \dfrac{[A]+[I]}{3}\, e^{2t} + \dfrac{2[I]-[A]}{3}\, e^{-t}$).

4°) en reprenant cet exercice, montrer que l'on peut parler d'exponentielle d'un opérateur (continu) A sur $\mathbb{R}$, et plus généralement sur un espace de Banach E . Dans le cas de $\mathbb{R}^n$, et pour une base donnée, montrer avec les notations du Chapitre 0 et du début de l'exercice que l'on a :

$$[e^{tA}] = e^{t[A]}$$

EXERCICE N° 18.

1°) Soit x (x_1 , x_2 ,..., x_n ,...) un élément de l^2 . Montrer que l'application Π_n , définie sur l^2 par :

$$\Pi_n(x) = x_n \qquad \text{(coordonnée d'ordre } n \text{)}$$

est une forme linéaire continue de norme 1.

EXERCICE N° 19.

Soit D le disque unité du plan complexe $\mathbb{C}$ ($|z| < 1$)
On appelle L^2 (D) l'espace vectoriel de Hilbert (Cf. II § 6), constitué par les fonctions (mesurables) et de carré intégrable muni du produit scalaire

$$< f,g > = \iint_D f(z)\, \bar{g}(z)\, dx\, dy \ , \quad \text{avec} \quad z = x + iy$$

Montrer que le sous-espace vectoriel $HL^2(D)$ composé des fonctions holomorphes est un sous-espace vectoriel fermé de $L^2(D)$.

EXERCICE N° 20.

Soit F le sous-espace vectoriel de l^2 constitué par les suites $x = \{x_1, x_2 ,..., x_n ,...\}$ telles que seul un nombre fini de coordonnées x_n soient différentes de 0.

Montrer que F est un sous-espace vectoriel dense de l^2 .

EXERCICE N° 21.

Soit $\mathscr{C}[0,1]$ l'espace vectoriel normé pour la norme de la convergence uniforme des fonctions continues réelles définies sur $[0,1]$.

On définit l'opérateur P_n sur $\mathscr{C}[0,1]$, à valeurs dans $\mathscr{C}[0,1]$

$$P_n(f)(x) = \sum_{k=0}^{k=n} C_n^k x^k (1-x)^{n-k} f(\tfrac{k}{n}) \qquad n = 0,1,2 \ldots$$

Montrer que P_n est un opérateur linéaire continu, positif, de norme 1 sur $\mathcal{C}[0,1]$ (polynôme de Bernstein).

EXERCICE N° 22.

h désigne un nombre réel (strictement) positif et $\mathcal{C}[0,1]$ l'espace vectoriel normé, pour la norme de la convergence uniforme, des fonctions continues réelles, définies sur $[0,1]$.

On définit la forme P_h sur $\mathcal{C}[0,1]$ par ses valeurs

$$P_h(f) = \frac{2}{\pi}\int_0^1 \frac{h}{h^2+x^2} f(x)\,dx$$

1°) montrer que $P_h(f)$ est une forme linéaire continue de norme 1 sur $\mathcal{C}[0,1]$

2°) montrer que, pour toute fonction f de $\mathcal{C}[0,1]$,

$$\lim_{h \to 0} P_h(f) = f(0)$$

Si l'on appelle P_0 la forme linéaire continue $f \longrightarrow f(0)$, définie sur $\mathcal{C}[0,1]$, peut-on assurer que P_h converge vers P_0 au sens de la norme du dual de $\mathcal{C}[0,1]$ lorsque h tend vers 0 ?

EXERCICE N° 23.

Soit f une fonction appartenant à $\mathcal{C}[0,1]$. Démontrer le résultat suivant :

$$\lim_{p \to +\infty} \|f\|_p \quad \text{existe et vaut} \quad \|f\|_\infty$$

EXERCICE N° 24.

En utilisant l'inégalité de Hölder, montrer que l'on peut identifier $L_q[0,1]$ à un sous-espace du dual topologique de $L_p[0,1]$ lorsque $1 < p < +\infty$

[On peut en fait démontrer que par l'identification précédente, l'espace $L_q[0,1]$ est isométriquement isomorphe à l'espace ($(L_p[0,1])'$)]

EXERCICE N° 25.

Donner la valeur exacte de la constante a_n qui intervient dans l'exemple 5.3.5. du Chapitre I. On démontrera le théorème suivant :

Théorème de Čebičev.

Parmi tous les polynômes de degré inférieur ou égal à n , dont le terme de plus haut degré a l'unité pour coefficient, le polynôme $\frac{1}{2^{n-1}} T_n(x)$ où T_n est le $n^{\text{ième}}$ polynôme de Čebičev est celui dont le maximum de la valeur absolue sur $[-1, +1]$ est la plus petite.

[Pour la définition des polynômes de Čebičev (du 1er type), on peut se reporter à l'exercice n° 36. Pour notre théorème, nous avons seulement besoin de quelques remarques :

$$T_n(x) = \cos\big(n(\operatorname{arc}\cos x)\big) = x^n + C_n^2\, x^{n-2}\,(x^2-1) + \ldots +$$

On notera $T_n(x) = 2^{n-1}x^n$ + termes de degré inférieur, en outre le polynôme de degré n , $T_n(x)$, a n racines simples aux points $x_h = \cos \frac{2h-1}{2n}\pi$ et prend $n+1$ valeurs extrêmes, alternées, en $n+1$ points $x'_h = \cos \frac{2h}{2n}\pi$.

Pour démontrer le théorème, on écrit : $Q(x) = \frac{T_n(x)}{2^{n-1}} - P(x)$ où

$P(x) = x^n$ + termes de degré inférieur, et $\sup\limits_{x\in[-1+1]} |P(x)| < \sup\limits_{x\in[-1+1]} \frac{|T_n(x)|}{2^{n-1}}$.

Q est de degré $n-1$; mais $Q(x'_h) = \frac{(-1)^h}{2^{n-1}} - P(x'_h)$ pour $h = 0, 1, 2, \ldots, n$

Par alternance, on en déduit que Q admet n zéros donc est nul ; donc

$$P(x) = \frac{T_n(x)}{2^{n-1}}$$

ce qui est impossible.]

EXERCICE N° 26.

Soit E l'espace des fonctions numériques définies sur $[0, +\infty[$, bornées et continues telles que $\lim\limits_{x \to +\infty} f(x)$ existe. On munit E de la norme de la convergence uniforme :

$$\| f \| = \sup_{x \in [0,+\infty[} |f(x)|$$

On définit les deux opérateurs suivants P et Q sur E

$$P : \qquad Pf(x) = \frac{1}{x} \int_0^x f(y)\, dy \qquad x \neq 0 \qquad \text{et} \quad Pf(0) = f(0)$$

$$Q : \qquad Qf(x) = \frac{2x}{\sqrt{\pi}} \int_0^{+\infty} e^{-(xy)^2} f(y)\, dy \qquad x \neq 0$$

a) étudier la continuité des opérateurs P et Q.

b) montrer que l'application $L : f \longrightarrow \lim\limits_{x \to +\infty} f(x)$ est une forme linéaire continue sur l'espace E,

c) montrer que $L \circ P = L$ c'est-à-dire

$$\lim_{x \to +\infty} Pf(x) = \lim_{x \to +\infty} f(x)$$

et montrer que

$$\lim_{x \to 0} Qf(x) = \lim_{x \to +\infty} f(x) = L(f)$$

EXERCICE N° 27.

Soit E_{n+1} l'espace des polynômes d'une variable complexe z et de degré inférieur ou égal à n. On pose comme norme sur E_{n+1}

$$\| P \| = \sup_{|z| \leq 1} |P(z)|$$

Soit L une forme linéaire sur E_{n+1}. On construit, à partir de L, un opérateur T, linéaire sur E_{n+1}, selon :

$$T(P)(z) = L(P_z) \quad \text{où} \quad P_z(y) = P(zy)$$

Montrer que T applique E_{n+1} dans E_{n+1} et à une norme égale à la norme α de la forme linéaire L .

EXERCICE N° 28.

On appelle E l'espace vectoriel constitué par toutes les suites infinies $x = \{x_n\}_{n>0}$ de nombres réels, telles que

$$\lim_{n \to +\infty} x_n = 0$$

On pose en outre

$$\|x\|_\infty = \operatorname{Sup}_{n \geq 0} |x_n|$$

lère question : montrer que $x \longrightarrow \|x\|_\infty$ constitue une norme sur E ,

montrer que E est un espace vectoriel de Banach pour cette norme.

2ème question : soit $x^{(n)}$ l'élément de E défini par $x_k^{(n)} = 0$ si $k \neq n$ et 1 si $k = n$.

Montrer que $\sum_{n=0}^{p} x_n x^{(n)}$ converge vers $x = \{x_n\}_{n \geq 0}$ au sens de E lorsque p tend vers l'infini.

3ème question : nous nous proposons de déterminer l'espace dual topologique de E .

a) soit y une suite de nombres réels $y = \{y_n\}_{n \geq 0}$ telle que

$$\sum_{n=0}^{+\infty} |y_n| \quad \text{converge}$$

montrer que l'application $x \longrightarrow \sum_{n=0}^{+\infty} x_n y_n$ est une forme linéaire continue sur E . Déterminer la norme de cette forme linéaire continue.

b) on appelle ℓ^1 l'espace vectoriel constitué par les suites infinies $y = \{y_n\}_{n \geq 0}$ de nombres réels, telles que $\sum_{n=0}^{\infty} |y_n|$ converge .

Montrer que $\|y\|_1 = \sum_{n=0}^{\infty} |y_n|$ définit une norme sur ℓ^1 .

c) montrer que E' est isométriquement isomorphe à ℓ^1.

4ème question : E' peut être considéré, sans sa norme, comme un sous-espace vectoriel de E. Montrer que sur E', la topologie de dual est strictement plus fine que la topologie de sous-espace de E.

5ème question : montrer que E'', bidual topologique de E, peut s'identifier à l'espace vectoriel constitué par les suites bornées réelles.

EXERCICE N° 29.

Soit E un espace vectoriel normé. Montrer qu'il ne peut exister deux applications linéaires et continues, P et Q, de $\mathcal{L}(E)$, telles que

$$PQ - QP = I$$

où I désigne l'opérateur identité sur E

En déduire que l'espace E de l'exercice N°5 n'est pas normable.

- CHAPITRE III -

TECHNIQUES HILBERTIENNES

Dans tout ce chapitre, $\mathcal{H}$ désigne un espace de Hilbert. Trois théorèmes fondamentaux vont être démontrés sur un tel espace et nous en examinerons ensuite les conséquences pratiques sur des réalisations concrètes d'espaces fonctionnels. Chacun de ces théorèmes généralise une propriété familière de l'espace euclidien et cette parenté géométrise considérablement l'analyse fonctionnelle, comme nous allons le voir.

Le produit scalaire dans $\mathcal{H}$ est noté $\langle x , y \rangle$.

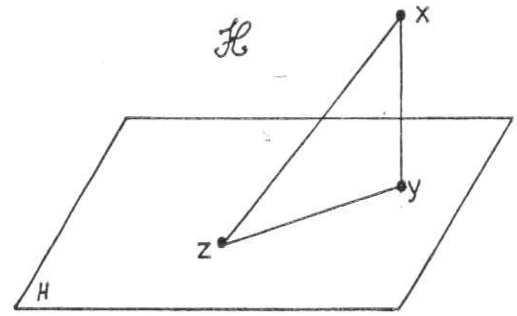
3.1. THEOREME DE LA PROJECTION.

3.1.1. Première forme du théorème de projection.

Soit H un sous-espace vectoriel fermé de $\mathcal{H}$. A tout point x de $\mathcal{H}$, il correspond un point y de H , unique, tel que

$$\boxed{\| x - y \| = \inf_{z \in H} \| x - z \|}$$

Le point y est appelé la projection du point x sur le sous-espace H. L'énoncé de ce théorème signifie qu'il existe un point y, et un seul dans H , tel que la distance $\| x - z \|$ soit la plus petite possible lorsque z parcourt H . Dans le cadre d'un espace euclidien, par exemple l'espace $\mathbb{R}^3$ de la géométrie ordinaire, ce théorème est bien connu. Soit H un plan et x un point de $\mathbb{R}^3$. De tous les segments joignant x aux points de H, il en est un qui est de longueur minimale : celui qui est porté par une droite passant par x et orthogonale à H . Le point y n'est autre que la projection orthogonale de x sur le plan H . Nous allons retrouver ces notions familières dans le cas d'un espace de Hilbert.



Projection orthogonale de x sur H

a) Unicité du point y :

Soit x un point arbitraire donné de $\mathcal{H}$. Montrons que si y existe, ce point est nécessairement unique. Supposons en effet que y_1 et y_2 réalisent le minimum

$$\| x - y_1 \| = \| x - y_2 \| = \inf_{z \in H} \| x - z \| = \delta$$

Rappelons l'identité du parallélogramme (Cf Chapitre I, § A, 4.4.9), identité caractéristique des espaces préhilbertiens :

$$\| x' - y' \|^2 + \| x' - y' \|^2 = 2 (\| x' \|^2 + \| y' \|^2)$$

On peut encore écrire, en posant $\begin{cases} x' = x - y_1 \\ y' = x - y_2 \end{cases}$

$$\begin{aligned} 2 (\| x - y_1 \|^2 + \| x - y_2 \|)^2 &= \| 2x - (y_1 + y_2) \|^2 + \| y_2 - y_1 \|^2 \\ &= 4 \left\| x - \frac{y_1 + y_2}{2} \right\|^2 + \| y_2 - y_1 \|^2 \\ 4\delta^2 &= 4 \left\| x - \frac{y_1 + y_2}{2} \right\|^2 + \| y_2 - y_1 \|^2 \end{aligned}$$

Mais $\frac{y_1 + y_2}{2}$ appartient à H comme y_1 et y_2, donc on dispose de l'inégalité

$\left\| x - \frac{y_1 + y_2}{2} \right\| \geqslant \delta$ Soit $\| y_1 - y_2 \| \leqslant 0$, d'où $y_1 = y_2$

La démonstration précédente n'utilise que la structure préhilbertienne de $\mathcal{H}$. On pourrait la généraliser (cf Nanta Iremica Vol N° 8 : Exercices)

b) Existence du point y :

Puisque $\inf_{z \in H} \| x - z \| = \delta$, il existe une suite de points y_n de H telle que $\| x - y_n \| = \delta_n$ où δ_n est une suite décroissante de nombres réels convergeant vers δ.

Une fois encore, écrivons l'identité du parallélogramme, avec
$\begin{cases} x' = x - y_n \\ y' = x - y_m \end{cases}$

On a : $$2 (\| x - y_n \|^2 + \| x - y_m \|)^2 = 4 \left\| x - \frac{y_n + y_m}{2} \right\|^2 + \| y_m - y_n \|^2$$

$$2 (\delta_n^2 + \delta_m^2) - 4 \left\| x - \frac{y_n + y_m}{2} \right\|^2 \geqslant \| y_m - y_n \|^2$$

donc :

$$0 \leqslant \| y_m - y_n \|^2 \leqslant 2 (\delta_n^2 + \delta_m^2 - 2\delta^2)$$

Soit :

$$\lim_{\substack{n \to +\infty \\ m \to +\infty}} \| y_m - y_n \| = 0$$

La suite y_n est une suite de Cauchy dans H , or H est un espace fermé donc complet comme sous-espace fermé d'un espace complet (Cf. Ex. N° 16). La suite $\{y_n\}_{n \geq 1}$ converge donc vers un élément y de H lequel réalise

$$\| x - y \| = \inf_{z \in H} \| x - z \|$$

Dans (b), nous avons maintenant eu besoin du fait que H est complet (c'est-à-dire ici de Hilbert).

Le théorème de la projection une fois démontré, il importe de donner quelques propriétés du point y . Pour fixer les idées et simplifier certains calculs, nous supposerons que $\mathcal{H}$ est un espace de Hilbert réel.

Montrons que $\langle x - y , z - y \rangle = 0$ pour tout élément z de H . Cette propriété caractérise même le point y parmi tous les points de H .

Pour tout λ tel que $0 < \lambda$, le point $y + \lambda (z - y)$ appartient à H , donc on a :

$$\| x - [y + \lambda (z - y)] \|^2 \geq \| x - y \|^2$$

ce qui s'écrit :

$$\| x - y \|^2 + \lambda^2 \| z - y \|^2 + 2 \lambda \langle x - y , z - y \rangle \geq \| x - y \|^2$$

ou encore :

$$2 \langle x - y , z - y \rangle \geq (- \lambda \| z - y \|^2)$$

Mais comme $\lambda > 0$ peut être arbitrairement petit, on a la relation :

$$\langle x - y , z - y \rangle \geq 0 \quad \text{pour tout z dans H .}$$

Changeant z - y en y - z , on aurait aussi $\langle x - y , y - z \rangle \geq 0$. Donc :

$$\langle x - y , z - y \rangle = 0 \quad \text{pour tout z dans H.}$$

Réciproquement, si y appartenant à H , est tel que $\langle x - y , z - y \rangle = 0$ pour tout z de H , on a :

$$\| x - [y + (z - y)] \|^2 = \| x - y \|^2 + \| z - y \|^2$$

donc :

$$\| x - z \|^2 \geq \| x - y \|^2 \quad \text{pour tout z de H .}$$

y est bien le point de H réalisant le minimum de $\|x-z\|$ lorsque z parcourt H .

Naturellement $<x-y, z-y>=0$ pour tout z de H équivaut à $<x-y, z>=0$ pour tout z de H .

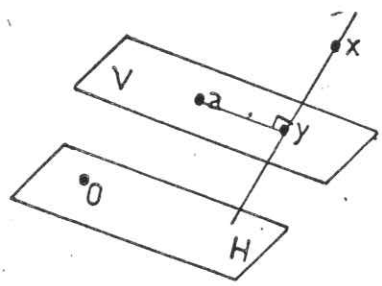
On peut donc énoncer une deuxième forme du théorème de la projection.

3.1.2. Deuxième forme du théorème de projection.

La projection y d'un point x sur un sous-espace fermé H de l'espace $\mathcal{H}$ est l'unique point y de H tel que $y-x$ soit orthogonal à H .

Remarque : On peut aussi projeter un élément x sur une variété linéaire V . Rappelons qu'une variété linéaire est l'ensemble des vecteurs de la forme $a+H$ où a est un vecteur fixe et H un sous-espace vectoriel de $\mathcal{H}$.

La projection sera également définie en minimisant $\|x-y\|$ où y parcourt V . On démontrerait de même que la projection y est telle que $x-y$ est orthogonale à tout vecteur de H et $y \in V$.



Donnons une autre formulation du théorème. La projection sur H étant unique, on a ainsi défini une application de $\mathcal{H}$ sur H . Appelons P_H l'opérateur $x \longrightarrow y = P_H(x)$. Cet opérateur possède certaines propriétés qui suffisent à le caractériser.

Définition : Un opérateur, élément de $\mathrm{Hom}(\mathcal{H},\mathcal{H})$ est un opérateur idempotent si

$$P^2 = P$$

3.1.3. Troisième forme du théorème de projection.

Soit H un sous-espace vectoriel fermé d'un espace de Hilbert $\mathcal{H}$. L'opérateur P_H est un opérateur linéaire, continu, de norme 1 et idempotent.

P_H est appelé le projecteur orthogonal de $\mathcal{H}$ sur H .

D'une part $P_H(x)$ est caractérisé, parmi les points de H, par la propriété :

$$<P_H(x) - x\ ,\ z> = 0$$

pour tout z de H d'après la deuxième forme du théorème de projection.

C'est-à-dire que $P_H(x)$ est <u>caractérisé</u>, en tant qu'élément de H, par :

$$<P_H(x), z> = <x, z>$$

pour tout z de H.

Par suite, considérons $P_H(\lambda_1 x_1 + \lambda_2 x_2)$, lequel est caractérisé par

$$<P_H(\lambda_1 x_1 + \lambda_2 x_2), z> = <\lambda_1 x_1 + \lambda_2 x_2, z>$$

Mais $\lambda_1 P_H(x_1) + \lambda_2 P_H(x_2)$ satisfait cette relation puisque

$$\begin{aligned} <\lambda_1 P_H(x_1) + \lambda_2 P_H(x_2), z> &= \lambda_1 <P_H(x_1), z> + \lambda_2 <P_H(x_2), z> \\ &= \lambda_1 <x_1, z> + \lambda_2 <x_2, z> \\ &= <\lambda_1 x_1 + \lambda_2 x_2, z> \end{aligned}$$

Donc par unicité :

$$\lambda_1 P_H(x_1) + \lambda_2 P_H(x_2) = P_H(\lambda_1 x_1 + \lambda_2 x_2)$$

L'opérateur P_H est linéaire.

De même on vérifie aussitôt l'idempotence de P_H.

• De plus H est image directe de P_H et les points de H sont caractérisés par la relation $P_H(x) = x$. Enfin, le noyau de P_H est l'ensemble des vecteurs orthogonaux à H.

Décomposons un vecteur quelconque x de $\mathcal{H}$ sous la forme :

$$x = P_H(x) + (x - P_H(x))$$

Il vient $\qquad P_H(x) \in H \quad$ et $\quad x - P_H(x) \in H^{\perp}$

en notant $H^{\perp}$ le sous-espace orthogonal à l'espace H (Cf Chapitre I, §A.4.4.6).

On dispose de

$$\|x\|^2 = \|P_H(x)\|^2 + \|x - P_H(x)\|^2$$

d'après le théorème de Pythagore, donc

$$\|P_H(x)\| \leqslant \|x\| \text{ et } \quad \|P_H\| = 1$$

Enfin, on note que $\mathcal{H}$ est somme directe de H et de $H^\perp$ ($H^\perp$ est un supplémentaire de l'espace H). L'existence d'un opérateur continu P_H fait dire que $\mathcal{H}$ est somme directe topologique de H et $H^\perp$:

$$\mathcal{H} = H \oplus H^\perp$$

$H^\perp$ est appelé le supplémentaire orthogonal de H (noter que $(H^\perp)^\perp = H$).

Le théorème de projection permet de démontrer le théorème de dualité suivant, dû simultanément à Fr. RIESZ et M. FRÉCHET, au tout début du siècle.

3.2. THEOREME DE DUALITE ET SES CONSEQUENCES.

3.2.1. Théorème.

Soit L une forme linéaire et continue sur un espace de Hilbert $\mathcal{H}$. Il existe un élément y de $\mathcal{H}$, unique, tel que pour tout vecteur x de $\mathcal{H}$

$$L(x) = \langle x , y \rangle \tag{1}$$

Au Chapitre I,A §4.4.8, nous avons vu que tout élément y de $\mathcal{H}$ définissait une forme linéaire $L_y : x \longrightarrow \langle x, y \rangle$ ce qui permet d'identifier $\mathcal{H}$ à un sous-espace vectoriel de $\mathcal{H}^*$, puisque $L_{\lambda_1 y_1 + \lambda_2 y_2} = \lambda_1 L_{y_1} + \lambda_2 L_{y_2}$ (cas réel).

Grâce à l'inégalité de Schwarz :

$$|L_y(x)| = |\langle x , y \rangle| \leqslant \| x \| \; \| y \|$$

on constate que L_y est une forme linéaire continue sur $\mathcal{H}$ et de norme $\| y \|$. En outre on a $L_{y_1} \neq L_{y_2}$ lorsque $y_1 \neq y_2$. Par conséquent $\mathcal{H}$ est identifiable à un sous-espace vectoriel du dual topologique $\mathcal{H}'$.

Le théorème de dualité permet la réciproque : c'est-à-dire enseigne que $\mathcal{H}$ et $\mathcal{H}'$ sont canoniquement identifiables par l'application $y \longrightarrow L_y$

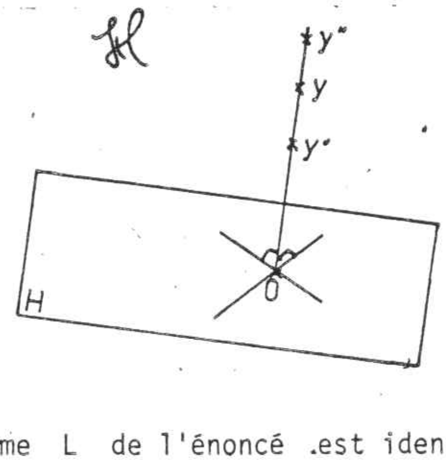
Théorème : Un espace de Hilbert (réel) $\mathcal{H}$ est isométriquement isomorphe à son dual.

Ce théorème a un grand intérêt dans la mesure où les formes linéaires continues, sur un espace vectoriel normé, jouent en dimension infinie le rôle des "coordonnées" (covariantes) dans le cas d'un espace vectoriel de dimension finie. Avant de passer à la démonstration, et pour éclairer celle-ci, interprétons le théorème de dualité dans le cas de l'espace $\mathbb{R}^3$, espace usuel de la géométrie.

Une forme linéaire sur $\mathbb{R}^3$ est nécessairement continue d'après le Chapitre I. Nous savons également (Cf Chapitre 0) qu'un espace vectoriel de dimension n est isomorphe à son dual mais il n'y a pas d'isomorphisme privilégié. Cependant lorsque l'on dote $\mathbb{R}^3$ d'une structure hilbertienne, en le rapportant par exemple à une base orthonormée, une forme linéaire L est déterminée par trois nombres a , b , c .

$$L(x) = ax_1 + bx_2 + cx_3$$

Ceci signifie que le plan passant par l'origine : $ax_1 + bx_2 + cx_3 = 0$,détermine la forme linéaire L . Les trois nombres a , b et c sont les coordonnées d'un vecteur V orthogonal au plan P et l'on sait bien qu'il y a correspondance biunivoque entre l'ensemble des plans contenant l'origine et l'ensemble des droites passant par l'origine par la relation d'orthogonalité (cf figure). En somme, l'isomorphisme de $\mathbb{R}^3$ et de son dual devient canonique, indépendant du choix d'une base... à condition d'avoir une structure hilbertienne. Cette situation est encore exacte dans le cadre général d'un espace de Hilbert comme nous allons l'établir ci-dessous. On supposera ici encore $\mathcal{H}$ réel pour simplifier.



Lorsque la forme L de l'énoncé est identiquement nulle, il suffit de prendre y = 0 dans la formule du théorème 3.2.1.

Soit L une forme linéaire continue non identiquement nulle définie sur $\mathcal{H}$. Soit H le noyau de L , c'est-à-dire l'ensemble des vecteurs x tels que $L(x)=0$ H est un sous-espace vectoriel de $\mathcal{H}$, fermé grâce à la continuité de L . (Cf exercice 15). Il existe un vecteur y' non nul tel que y' soit orthogonal à H . En effet, comme H est strictement inclus dans $\mathcal{H}$ puisque L n'est pas identiquement nulle, on peut prendre un vecteur x_1 de $\mathcal{H}$ n'appartenant pas à H et considérer sa projection orthogonale $P_H\ (x_1)$ sur H . Le vecteur $y' = x_1 - P_H\ (x)$, non nul, est orthogonal à H .

Posons
$$y'' = \frac{y'}{L(y')}$$
($L(y') \neq 0$ puisque y' est non nul et appartient à $H^{\perp}$).

Pour la même raison $\|y''\| \neq 0$. Envisageons $< x, y'' >$

- Pour x appartenant à H, $L(x) = 0$ donc vaut aussi $\langle x, y''\rangle$
- Pour x n'appartenant pas à H, $L(x - y''L(x)) = L(x) - L(x)L(y'')$
$$= L(x) - L(x) = 0$$

car $L(y'') = 1$. Donc $x - y''L(x)$ appartient à l'espace H. Ce qui nous fournit
$$0 = L(x - y''L(x)) = \langle x - y''L(x), y'\rangle$$
et ce qui donne à son tour
$$< x, y'' > = L(x) < y'', y'' >$$
d'où puisque $\|y''\|^2 = < y'', y''> \neq 0$, l'écriture :
$$L(x) = \left\langle x, \frac{y''}{\langle y'', y''\rangle}\right\rangle$$
En posant $y = \dfrac{y''}{\langle y'', y''\rangle}$, on obtient pour tout x de $\mathcal{H}$ la relation cherchée
$$L(x) = \langle x, y\rangle$$
De plus la correspondance entre L et l'élément y unique défini ci-dessus est linéaire et : $\|L\| \leqslant \|y\|$
d'après l'inégalité de Schwarz. Mais si l'on fait $x = y$, il vient
$$L(y) = \|y\|^2 \leqslant \|L\| \|y\| \quad \text{donc } \|L\| \geqslant \|y\|$$

D'où la relation
$$\|L\| = \|y\|$$
ce qui termine la démonstration de l'isométrie entre $\mathcal{H}$ et $\mathcal{H}'$.

Par contre, en général, $\mathcal{H}^*$, le dual algébrique, est beaucoup plus grand que $\mathcal{H}$. (Il faut noter que l'isométrie est antilinéaire lorsque $\mathcal{H}$ est complexe),
Examinons quelques conséquences du théorème de projection.

3.2.2. Critère de densité.

Nous avons déjà défini, au Chapitre 0, la notion algébrique d'orthogonal H^0 d'un sous-espace vectoriel H. Dans le cadre d'un espace de Hilbert, cette notion coïncide avec la notion d'orthogonal $H^{\perp}$, à condition de considérer la dualité $(\mathcal{H}, \mathcal{H}')$ au lieu de la dualité $(\mathcal{H}, \mathcal{H}^*)$.

Nous avions dit que la considération algébrique d'un orthogonal avait une grande utilité (théorie de la dualité algébrique). Dans le cadre d'un espace de Hilbert, il importe de considérer l'orthogonal comme sous-ensemble de $\mathcal{H}' \simeq \mathcal{H}$ (théorie de la dualité topologique). On remarque alors que l'on peut étudier soit H , soit $H^{\perp}$, puisque toute propriété sur l'un de ces espaces fournit une propriété sur l'autre espace (lorsque H est fermé, on a $(H^{\perp})^{\perp} = H$: le démontrer à titre d'exercice).

De même supposons que H_1 et H_2 soient deux sous-espaces vectoriels tels que $H_1 \subset H_2$. Supposons en outre que H_2 soit fermé et que H_1 et H_2 aient le même sous-espace orthogonal. Dans ce cas, on peut démontrer que H_1 est dense dans H_2 (le démontrer en exercice).

Ce résultat est d'usage courant et il donne un critère de densité. Nous l'utiliserons notamment pour démontrer qu'un sous-espace vectoriel est dense dans un Hilbert, sous la forme suivante :

Théorème : Soit E un sous-espace vectoriel d'un espace de Hilbert $\mathcal{H}$ tel que seul le vecteur nul soit orthogonal à E . Dans ce cas, E est alors dense dans $\mathcal{H}$.

La démonstration est simple : soit H la fermeture de E . Par hypothèse l'espace $E^{\perp}$, orthogonal de H, est réduit à 0. Par conséquent, d'après le théorème de projection, valable pour un fermé H , on a :

$$H \oplus [0] = \mathcal{H} \quad \text{soit} \quad H = \mathcal{H} \qquad \text{c.q.f.d.}$$

Ce résultat peut naturellement être considéré comme une simple conséquence du critère de densité examiné en I, B, 5.4.3, puisque nous avons canoniquement identifié les formes linéaires et continues sur $\mathcal{H}$ à des éléments de $\mathcal{H}$.

3.2.3. Application du théorème de projection au problème de Dirichlet .

Envisageons formellement l'application du théorème de projection au problème classique de Dirichlet.

Soit Ω un domaine de $\mathbb{R}^2$ et f une fonction définie sur Ω et sur sa frontière. Le problème de Dirichlet consiste en la recherche d'une fonction ψ , harmonique sur Ω , prenant les mêmes valeurs frontières que f . On va voir que formellement la solution ψ s'obtient en minimisant l'expression

$$I(g) = \iint_{\Omega} |\overrightarrow{\text{grad}}\, g|^2 dx\, dy$$

lorsque g parcourt les fonctions g , suffisamment régulières, prenant les mêmes valeurs frontières que f .

On définit pour cela un espace de Hilbert $\mathcal{H}$ tel que $\sqrt{I(g)}$ représente la norme d'un élément g de $\mathcal{H}$ et on doit minimiser cette norme sur la variété linéaire E des fonctions ayant les mêmes valeurs frontières que f . Il s'agit donc de projeter 0 orthogonalement sur E . La fonction minimisante, notée ϑ , satisfera alors (théorème de projection) :

$$\langle \vartheta, g-\vartheta\rangle = \iint_{\Omega} (\overrightarrow{\text{grad}}\, \vartheta)(\overrightarrow{\text{grad}}\,(g-\vartheta))\,dx\,dy = 0$$

ou encore :

$$\iint_{\Omega} (\overrightarrow{\text{grad}}\, \vartheta)(\overrightarrow{\text{grad}}\, V)\,dx\,dy = 0$$

pour toute fonction V dans $\mathcal{H}$ ayant une valeur frontière nulle. On note $\mathrm{Fr}\,\Omega$ la frontière de Ω .

Or la formule de Green donne (en la supposant exacte pour les éléments de $\mathcal{H}$) :

$$\langle V, \vartheta\rangle = \iint_{\Omega} (\overrightarrow{\text{grad}}\, V)(\overrightarrow{\text{grad}}\, \vartheta)\,dx\,dy = \iint_{\Omega} V\Delta\vartheta\,dx\,dy - \int_{\mathrm{Fr}\Omega} V\,\frac{\delta\vartheta}{\delta\eta}\,ds$$

Comme V est nul sur la frontière

$$\langle V, \vartheta\rangle = \iint_{\Omega} V\Delta\vartheta\,dx\,dy = 0$$

ce qui donne $\Delta\vartheta = 0$ puisque l'égalité précédente est exacte pour tout V .

Bien entendu, il faut procéder soigneusement pour vérifier toutes les assertions précédentes et préciser le vocabulaire. Il y a plusieurs difficultés puisqu'il faut définir correctement l'espace de Hilbert $\mathcal{H}$ d'une part et d'autre part il faut donner un sens à la valeur frontière d'une fonction f de $\mathcal{H}$.

En théorie de l'approximation, nous étudierons la méthode de Galerkine qui est une méthode par projections successives fort utilisée pour les équations aux dérivées partielles de type elliptique, et qui repose sur le théorème de dualité.

3.2.4. Exemple l^2 :

Le théorème de dualité assure que toute forme linéaire continue L sur ℓ^2 se représente sous la forme d'un élément y de ℓ^2 :

$$y = (y_1, y_2, \dots y_n, \dots)$$

$$\boxed{L(x) = \langle x, y\rangle = \sum_{n=1}^{\infty} x_n\, y_n}$$

3.2.5. Exemple $L^2[0,1]$:

Le théorème de dualité assure que toute forme linéaire continue L sur $L^2\,[0,1]$ se représente par une fonction g de $L^2\,[0,1]$ selon

$$\boxed{L(f) = \int_0^1 f(x)\,g(x)\,dx}$$

Sur $L^2([0,1],h)$, on aurait :

$$L(f) = \int_0^1 f(x)\,g(x)\,h(x)\,dx$$

(rappelons que nous avons supposé les espaces réels pour simplifier).

3.3. THEOREME DU CHOIX.

Dans un espace euclidien E , on sait que de toute suite infinie $\{x_n\}_{n\geq 1}$ d'éléments de E dont les normes restent bornées par un nombre fixe, on peut extraire une sous-suite $\{x_{n_k}\}$ convergeant vers un élément x de E . Ce résultat constitue le théorème de Bolzano-Weierstrass que l'on a démontré au Chapitre I.

Considérons maintenant l'exemple de l'espace ℓ^2 pour voir si le théorème de Bolzano-Weierstrass est encore valable.

Soit $\rho^{(1)}$ le vecteur de ℓ^2 défini par la suite $(1,0,0\ldots0,\ldots)$ et plus généralement soit $\rho^{(n)}$ le vecteur défini par la suite $(0,0,\ldots,1,0\ldots)$ (0 partout sauf 1 à la $n^{\text{ième}}$ place)

$$\begin{aligned}
\rho^{(1)} &= (1, 0, \ldots, 0, \ldots, 0\ldots) \\
\rho^{(2)} &= (0, 1, \ldots, 0, \ldots, 0\ldots) \\
&\vdots \\
\rho^{(n)} &= (0, 0, \ldots, 1, \ldots 0\ldots)
\end{aligned}$$

La norme de $\rho^{(n)}$ est 1 puisque $\|\rho^{(n)}\|^2 = \langle \rho^{(n)}, \rho^{(n)} \rangle = 1$, donc la suite $\rho^{(n)}$ de vecteurs de ℓ^2 , est incluse dans la boule unité de ℓ^2 . Par ailleurs, considérée comme une suite de vecteurs de ℓ^2 ,cette suite $\rho^{(n)}$ ne constitue pas une suite de Cauchy, et on ne peut pas en extraire une sous-suite de Cauchy.

En effet, on note que

$$\begin{aligned}
\|\rho^{(n)} - \rho^{(p)}\|^2 &= \langle \rho^{(n)} - \rho^{(p)}, \rho^{(n)} - \rho^{(p)} \rangle \\
&= \|\rho^{(n)}\|^2 + \|\rho^{(p)}\|^2 - 2\langle \rho^{(n)}, \rho^{(p)} \rangle
\end{aligned}$$

Or $\langle \rho^{(n)}, \rho^{(p)} \rangle = 0$ si $n \neq p$ et donc

$$\|\rho^{(n)} - \rho^{(p)}\| = \sqrt{2} \qquad \text{lorsque } n \neq p \quad (\text{et bien sûr } 0 \text{ si } n = p\ )$$

On ne peut donc pas extraire de la suite $\rho^{(n)}$ une sous-suite convergente au sens de ℓ^2 , c'est-à-dire qu'il n'existe pas de sous-suite $\rho^{(n_k)}$ et d'élément x de ℓ^2 tel que

$$\lim_{k\to\infty} \| \rho^{(n_k)} - x \| = 0$$

Le théorème de Bolzano-Weierstrass n'est donc pas exact dans le cas d'un espace de Hilbert de dimension infinie. Cependant il existe une forme affaiblie d'un tel théorème.

Prenons la suite $\rho^{(n)}$ de vecteurs de ℓ^2 et un vecteur quelconque y de ℓ^2 . Formons le produit scalaire $\langle \rho^{(n)}, y \rangle$.

Il vient :

$$\langle \rho^{(n)}, y \rangle = \sum_{k=1}^{\infty} \rho_k^{(n)} y_k = y_n$$

Or $\lim_{n\to\infty} y_n = 0$ puisque y, appartenant à l^2 , $\sum_{n=1}^{\infty} y_n^2$ converge .

Par conséquent :

$$\lim_{n\to\infty} \langle \rho^{(n)}, y \rangle = 0 \qquad \text{pour tout élément } y \text{ de } \ell^2 .$$

On note

$$\lim_{n\to\infty} \langle \rho^{(n)}, y \rangle = \langle 0, y \rangle = 0$$

Ce résultat est général. Nous pouvons énoncer le théorème suivant, dit <u>théorème du choix</u>.

3.3.1. Théorème.

<u>De toute suite infinie d'éléments d'un espace de Hilbert $\mathcal{H}$ dont les normes restent bornées par un nombre fixe, on peut extraire une sous-suite $x^{(n)}$ et trouver un élément x de l'espace $\mathcal{H}$ de sorte que pour tous les éléments y de $\mathcal{H}$, on ait</u> :

$$\boxed{\lim_{n\to\infty} \langle x^{(n)}, y \rangle = \langle x , y \rangle}$$

On n'obtient donc pas la convergence de la sous-suite $x^{(n)}$ (comme nous l'avons vu pour ℓ^2) mais seulement la convergence de $\langle x^{(n)}, y \rangle$ pour tout y de l^2 . On convient de dire qu'une suite $x^{(n)}$ de $\mathcal{H}$, <u>converge faiblement vers x</u> si, pour tous les éléments y de $\mathcal{H}$, on a :

$$\lim_{n\to+\infty} \langle x^{(n)}, y \rangle = \langle x , y \rangle$$

Nous n'allons pas donner la démonstration générale du théorème du choix, mais faire la démonstration dans un cas particulier : le cas de l'espace ℓ^2 . Ce cas couvre tous les cas usuels car il s'applique en fait à tous les espaces de Hilbert séparables (la définition de ce mot sera donnée plus loin ainsi que la démonstration du fait que ℓ^2 est isomorphe et isométrique à tout espace de Hilbert séparable).

Nous utilisons le vocabulaire du paragraphe suivant et on pourra entreprendre la lecture de cette démonstration après la lecture du paragraphe 3.4.

Nous nous servons essentiellement du fait que $\rho^{(n)}$ constitue une base hilbertienne (orthogonale) de l'espace de Hilbert ℓ^2 . Nous procéderons en quatre étapes.

Soit X_i une suite infinie de vecteurs de l'espace ℓ^2 telle que

$$\|X_i\| \leqslant 1$$

pour fixer les idées. Ici $i \in I$, où I est un ensemble quelconque d'indices.

Première étape : procédé diagonal :

Il est possible d'extraire de X_i une sous-suite, notée $x^{(n)}$ telle que

$$\lim_{n \to \infty} \langle x^{(n)}, \rho^{(k)} \rangle \quad \text{existe}$$

et est finie pour tout élément $\rho^{(k)}$ défini précédemment.

Considérons $\langle X_i , \rho^{(1)} \rangle = a_i$: c'est une famille infinie de nombres réels tels que

$$|a_i| \leqslant \|X_i\| \, \|\rho^{(1)}\| \leqslant 1$$

Il existe donc une sous-suite dénombrable d'éléments de la forme X_i où $i \in I$, sous-suite encore notée $X_j^{(1)}$ où $j = 1, 2, \ldots$ et telle que

$$\lim_{j \to +\infty} \langle X_j^{(1)}, \rho^{(1)} \rangle \quad \text{existe}$$

(théorème de Bolzano-Weierstrass sur la droite réelle).

Considérons maintenant $\langle X_j^{(1)}, \rho^{(2)} \rangle = b_j$. Il s'agit d'une suite infinie de nombres réels tels que $|b_j| \leqslant 1$. Il existe donc une sous-suite de $X_j^{(1)}$, notée $X_j^{(2)}$, telle que

$$\lim_{j \to +\infty} \langle X_j^{(2)}, \rho^{(2)} \rangle \quad \text{existe.}$$

On forme ainsi des suites $\chi_j^{(n)}$, chacune étant extraite de la suite précédente, et telles que

$$\lim_{j \to +\infty} \quad < \chi_i^{(n)}, \rho^{(n)} > \qquad \text{existe}$$

Considérons alors la suite diagonale notée $x^{(n)} = \chi_n^{(n)}$. Bien sûr $x^{(n)} \in l^2$.

Selon le procédé de formation, il est clair que :

$$\lim_{n \to +\infty} < x^{(n)}, \rho^{(k)} > \quad \text{existe pour tout entier } k \ (k \geq 1)$$

<u>Deuxième étape</u> : $\lim_{n \to +\infty} < x^{(n)}, y >$ <u>existe pour tout élément y de ℓ^2</u> .

Ecrivons un élément y sous la forme (cf théorème 3.5.1.)

$$y = \sum_{k=1}^{N} y_k \rho^{(k)} + \sum_{k>N} y_k \rho^{(k)} = Y_N + \sum_{k>N} y_k \rho^{(k)}$$

On peut choisir N de telle sorte que pour un ε donné

$$\| y - Y_N \|^2 = \sum_{k>N} y_k^2 < \varepsilon^2$$

Par conséquent, formons

$$< x^{(n)}, y > - < x^{(m)}, y >$$

$$<x^{(n)}, y>-<x^{(m)}, y> = <x^{(n)}, y - Y_N> + <x^{(n)}, Y_N> - <x^{(m)}, Y_N> - <x^{(m)}, y - Y_N>$$

Soit :

$$|<x^{(n)}, y>-<x^{(m)}, y>| \leq \|x^{(n)}\| \|y - Y_N\| + |<x^{(n)} - x^{(m)}, Y_N>| + \|x^{(m)}\| \|y - Y_N\|$$

ε étant donné, on choisit N de telle sorte que $\|y - Y_N\| < \varepsilon$. Puis on choisit N' de telle sorte que pour n et m supérieurs à N'

$$|<x^{(n)} - x^{(m)}, Y_N>| < \varepsilon$$

ce qui est possible d'après le résultat de la première étape car Y_N est une combinaison linéaire finie des $\rho^{(k)}$. Par conséquent, avec N" = Sup (N, N'),

pour n et m > N', $|<x^{(n)}, y> - <x^{(m)}, y>| \leq 3\varepsilon$

La suite $<x^{(n)}, y>$ est une suite de Cauchy de nombres réels et converge donc, ce qui démontre l'assertion de la deuxième étape.

Troisième étape : la forme $y \longrightarrow L(y) = \lim\limits_{n \to +\infty} <x^{(n)}, y>$ est une forme linéaire continue sur ℓ^2 .

Bien entendu, $y \longrightarrow \lim\limits_{n \to \infty} <x^{(n)}, y> = L(y)$ est une forme linéaire sur ℓ^2 puisque :

$$<x^{(n)}, \lambda y + \mu y'> \quad = \quad \lambda < x^{(n)}, y > + \mu < x^{(n)}, y'>$$

donc

$$L(\lambda y + \mu y') \quad = \quad \lambda L(y) + \mu L(y') \quad .$$

Il reste à montrer que L est une forme continue, c'est-à-dire qu'il existe une constante M telle que pour tout y de ℓ^2 on ait

(1) $$|L(y)| \quad < \quad M \| y \|$$

Ce résultat constitue un cas particulier d'un théorème général d'analyse fonctionnelle : le théorème de Banach-Steinhaus, qui est facile à démontrer dans ce cas très particulier.

Nous avons l'inégalité suivante pour chaque n et pour tout y :

$$|<x^{(n)}, y>| \leqslant \| x^{(n)} \| \, \| y \|$$

Or

$$\| x^{(n)} \| \leqslant 1 \qquad \text{par hypothèse,}$$

donc $$|L(y)| = \lim_{n \longrightarrow \infty} |<x^{(n)}, y>| \leqslant \| y \|$$

Cette inégalité n'est autre que l'inégalité cherchée :

(1) $$|L(y)| \leqslant \| y \| \qquad \text{où} \quad M = 1$$

Quatrième étape : il existe un vecteur x de ℓ^2 tel que

$$<x , y> \; = \lim_{n \to +\infty} < x^{(n)}, y>$$

Il suffit d'appliquer le théorème de dualité pour assurer que $L(y)$ se représente par un élément x de ℓ^2 tel que $\| x \| = \| L \|$ et

$$\boxed{<x , y> \; = \; L(y) \; = \; \lim_{n \to +\infty} < x^{(n)}, y >}$$

Ceci termine la démonstration du théorème du choix.

Nous allons maintenant définir les bases hilbertiennes qui vont remplacer les bases algébriques usuelles des espaces de dimension finie. En dimensic infinie, les bases algébriques ont trop d'éléments et sont peu utilisables.

3.4. FAMILLES ORTHOGONALE ET TOTALE : BASES HILBERTIENNES.

3.4.1. Orthonormalité.

Soit $\mathcal{H}$ un espace préhilbertien réel ou complexe, une famille $\{\rho_i\}$ de vecteurs est dite orthonormale lorsque

$$<\rho_i , \rho_j> = \delta_{ij} \qquad \text{où} \quad \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

orthogonale lorsque l'on a seulement

$$<\rho_i , \rho_j> = 0 \quad \text{si} \quad i = j$$

Deux éléments d'une famille orthogonale sont linéairement indépendants puisque $\lambda\rho_i + \mu\rho_2 = 0$ entraîne $\lambda < \rho_1,\rho_1 > + \mu < \rho_2,\rho_1 > = \lambda = 0$ et de même $\mu = 0$ en faisant le produit scalaire avec ρ_2.

En fait, les vecteurs d'une famille orthonormale $\{\rho_i\}$ possèdent une indépendance plus fine : une indépendance de nature topologique

Il n'existe pas de combinaisons linéaires finies des ρ_i $(i \neq j)$ approchant ρ_j d'aussi près qu'on le désire dans l'espace $\mathcal{H}$. Ceci peut aussi s'énoncer en disant que le sous-espace vectoriel fermé engendré par les ρ_i $(i \neq j)$ ne contient pas l'élément ρ_j ; C'est le mot fermé qui ajoute quelque chose par rapport à la seule indépendance linéaire.

La démonstration de ce fait peut s'effectuer en utilisant la théorie de la dualité. Montrons donc que ρ_j appartient à l'espace orthogonal du sous-espace H engendré par les vecteurs ρ_i $(i \neq j)$. Or ceci est évident puisque si

$$x = \sum_{\substack{i=1 \\ i \neq j}}^{n} x_i \rho_j \qquad \text{alors} \quad <x, \rho_j> = 0 \quad (j \neq i)$$

La fermeture $\overline{H_j}$ du sous-espace vectoriel H_j engendré par les ρ_i où $i \neq j$ et l'espace H_j ont même orthogonal: $H_j^{\perp} = (\overline{H_j})^{\perp}$. Mais ρ_j appartient à cet orthogonal $(\overline{H_j})^{\perp}$. Par suite $\rho_j \neq 0$ ne peut appartenir simultanément à $\overline{H_j}$ et à $(\overline{H_j})^{\perp}$. Une démonstration, par le calcul, que $\rho_j \notin \overline{H_j}$ est la suivante.

Supposons que pour tout $\varepsilon > 0$, on puisse écrire

$$\| \rho_j - \sum_{\substack{i \in J \\ i \neq j}}' a_i \rho_i \, \| < \varepsilon \quad \text{où } J \text{ est un ensemble fini.}$$

On devrait avoir d'après l'inégalité de Schwarz

$$\left| < \rho_j - \sum_{\substack{i \in J \\ i \neq j}}' a_i \rho_i , \rho_j > \right| \leqslant \varepsilon$$

Cependant cette inégalité se réduit à

$$1 = < \rho_j , \rho_j > < \varepsilon$$

ce qui est impossible.

3.4.2. Famille totale.

Soit $\{\rho_i\}_{i \in I}$ une famille de vecteurs dans un espace vectoriel de Hilbert où i appartient à un ensemble d'indices I. Formons l'espace vectoriel H engendré par les ρ_i , c'est-à-dire l'ensemble H des combinaisons linéaires finies x des vecteurs ρ_i :

$$x = \sum_{i \in J} a_i \rho_i$$

où J est un sous-ensemble fini de l'ensemble des indices I.

Les familles $\{\rho_i\}_{i \in I}$ pour lesquelles H est dense dans $\mathcal{H}$ sont particulièrement intéressantes. On convient de la définition :

Définition : on dit qu'une famille de vecteurs $\{\rho_i\}_{i \in I}$ constitue une famille totale sur un espace de Hilbert $\mathcal{H}$ lorsque l'espace vectoriel H engendré par les ρ_i est dense dans $\mathcal{H}$.

Une famille totale peut être caractérisée par une propriété d'orthogonalité, du moins dans un espace de Hilbert.

Théorème : une condition nécessaire et suffisante pour qu'une famille $\{\rho_i\}_{i \in I}$ soit totale dans un espace de Hilbert $\mathcal{H}$ est que tout vecteur orthogonal à tous les ρ_i est obligatoirement le vecteur nul.

Si $\{\rho_i\}$ est totale, un vecteur x orthogonal à tous les ρ_i ($< x, \rho_i > = 0$ pour tout i dans I) est orthogonal au sous-espace vectoriel H et donc à sa fermeture $\overline{H}$ dans $\mathcal{H}$. Or H est dense dans $\mathcal{H}$ par la définition d'une suite totale, donc x est orthogonal à lui même , donc nul.

Réciproquement, si $\{\rho_i\}_{i\in I}$ est une famille telle que, seul le vecteur nul lui est orthogonal, alors l'espace vectoriel H engendré par les ρ_i a un orthogonal réduit au vecteur nul. H est donc dense dans $\mathcal{H}$ et la famille $\{\rho_i\}$ est totale d'après ce qui a été vu au paragraphe 3.2.2. Seule cette réciproque exige que $\mathcal{H}$ soit de Hilbert au lieu de préhilbertien seulement.

Exemple : les vecteurs $\{e^{(n)}\}$ de l^2 constituent une famille totale puisque l'hypothèse $< x , e^{(n)} > = 0$ signifie $x_n = 0$ pour tout n et donc :

$$x = (x_1, x_2, ---, x_n, ---) = (0, 0, ---0, ---) = 0$$

Rappelons que $e^{(n)} = (0, 0, \ldots, 1, 0, \ldots, 0, \ldots, 0)$ partout sauf à la $n^{\text{ième}}$ place où l'on met 1.

3.4.3. Bases hilbertiennes.

Définition : on appelle base (orthonormale) hilbertienne, une famille $\{e_i\}$ à la fois totale et orthonormale dans un espace de Hilbert $\mathcal{H}$

Les vecteurs $\{e^{(n)}\}$ de l^2 constituent une base orthonormale hilbertienne. c'est d'ailleurs la base hilbertienne canonique de l^2.
A titre d'exercice, utilisons la condition du théorème pour montrer que dans l'espace euclidien à n dimensions $(\mathbb{R}^n)$, une base orthonormale hilbertienne $\{e_k\}$ est une base au sens algébrique de la théorie des espaces vectoriels (Cf Chapitre 0, paragraphe 2).

Soit donc $\{e_k\}_{k\in K}$ où K est un ensemble d'indices, une base hilbertienne de l'espace euclidien $\mathbb{R}^n$. Nous avons déjà noté que les $\{e_k\}_{k\in K}$ sont linéairement indépendants puisque la famille est orthonormale. Par suite, K est fini et de cardinal inférieur ou égal à n. Pour terminer la démonstration, il suffit d'établir que le cardinal de K est n exactement. Deux démonstrations au moins sont possibles.

Ou bien l'on remarque que l'espace vectoriel engendré par les $\{e_k\}_{k\in K}$ est fermé car K est fini. Or la famille est totale, donc cet espace est $\mathbb{R}^n$ tout entier soit $\text{Card } K = n$.

Ou bien l'on remarque que si x est un vecteur n'appartenant pas à l'espace vectoriel engendré par les $\{e_k\}_{k\in K}$, on peut poser

$$\lambda_k = < x, e_k >$$

et

$$y = x - \sum_{k\in K} \lambda_k \; e_k$$

Or y est orthogonal à tous les e_k lorsque $k \in K$ et ne saurait par hypothèse

Cependant, dans le cas d'une dimension infinie, une base orthonormale hilbertienne n'est jamais une base au sens de la théorie des espaces vectoriels. Précisément, tout vecteur ne s'écrit pas comme une combinaison linéaire finie des vecteurs de base mais, comme nous allons le préciser, selon une combinaison infinie .

Notation : on appelle coordonnée (hilbertienne) d'un vecteur x d'un espace hilbertien $\mathcal{H}$, par rapport à un vecteur e_i d'une base hilbertienne $\{e_i\}_{i\in I}$ le nombre : $\boxed{<x\ ,\ e_i>}$.

Ce nombre est réel si $\mathcal{H}$ est un hilbertien réel, complexe si $\mathcal{H}$ est un hilbertien complexe.

Théorème de la base hilbertienne :

Dans tout espace hilbertien, il existe une base hilbertienne orthonormale.

La démonstration utilise l'axiome du choix, tout comme la démonstration du théorème de la base au Chapitre 0 dans le cas d'un espace vectoriel. Nous admettrons ce théorème que nous utiliserons peu car dans tous les cas concrets nous exhiberons une base hilbertienne explicite.

Convention : Afin de ne pas compliquer inutilement les notations et pour éviter d'introduire de nouvelles notions, nous allons faire l'hypothèse suivante sur l'espace de Hilbert $\mathcal{H}$:

Nous supposerons que l'espace $\mathcal{H}$ contient une base hilbertienne au plus dénombrable. On dit dans ce cas que $\mathcal{H}$ est séparable.

Comme l'on démontre que deux bases hilbertiennes orthonormales d'un même espace de Hilbert $\mathcal{H}$ ont même cardinal, appelé dimension hilbertienne de $\mathcal{H}$, toute base hilbertienne d'un espace de Hilbert séparable est dénombrable.

Les espaces de Hilbert les plus utilisés, comme ℓ^2 et $L^2[0,1]$, sont des espaces séparables : pour le premier, c'est évident puisque nous avons construit la base hilbertienne $\{e_n\}_{n\geq 1}$; pour le second, c'est moins évident. Nous le démontrerons(cf 3.5.6.)en construisant effectivement une base dénombrable. Nous allons maintenant donner un moyen pour construire des familles orthonormales et même totales.

L'hypothèse de séparabilité de $\mathcal{H}$ est traditionnelle pour une introduction à l'analyse fonctionnelle. Malgré tout, nous essaierons de ne pas en faire trop grand usage parce que certains espaces de Hilbert utiles ne sont pas séparables (Exemple cas des fonctions presque-périodiques de Besicovitch).

3.4.4. Théorème d'orthonormalisation de Gram-Schmidt.

Soit $\mathcal{H}$ un espace de Hilbert.

Si $\{a_n\}$ est une famille de vecteurs linéairement indépendants, il existe une famille orthonormale $\{e_n\}$ de vecteurs de $\mathcal{H}$ telle que, pour tout entier p, le sous-espace vectoriel engendré par $a_1, a_2, \ldots, a_p$ soit identique au sous-espace vectoriel engendré par $e_1, e_2, \ldots, e_p$.

De plus, si la famille $\{a_n\}_{n\geqslant 1}$ est totale, la famille $\{e_n\}_{n\geqslant 1}$ constitue une base hilbertienne orthonormale.

Démonstration : Posons
$$e_1 = \frac{a_1}{\|a_1\|}$$
Visiblement $\|a_1\| \neq 0$ puisque les vecteurs $\{a_n\}_{n\geqslant 1}$ sont indépendants.

On a ainsi normé le vecteur e_1
$$\|e_1\| = 1$$
Posons maintenant
$$e_2 = \lambda e_1 + \mu a_2$$
où λ et μ sont des scalaires tels que
$$\begin{cases} < e_2, e_1 > = 0 \\ < e_2, e_2 > = 1 \end{cases}$$

Ceci donne deux conditions sur λ et μ (On suppose $\mathcal{H}$ réel pour simplifier).
$$\begin{cases} 0 = <e_2, e_1> = \lambda + \mu < a_2, e_1> \\ 1 = <e_2, e_2> = \lambda^2 + 2\lambda\mu <e_1, a_2> + \mu^2 \|a_2\|^2 \end{cases}$$

Si $< a_1, a_2> = 0$, il suffit de prendre $\lambda = 0$ et μ tel que $|\mu|^2 . \|a_2\|^2 = 1$

Si $< a_1, a_2> \neq 0$, $\mu^2(\|a_2\|^2 - <e_1, a_2>^2) = 1$. L'indépendance linéaire de e_1 et a_2 entraîne que $\|a_2\|^2 - <e_1, a_2>^2 \neq 0$ (critère d'égalité dans l'inégalité de Schwarz). Ceci permet de choisir convenablement μ et par conséquent λ avec
$$\lambda = -\mu < a_2, e_1>$$

e_1 et e_2 engendrent le même espace vectoriel que a_1 et a_2. Les vecteurs e_1 et e_2 étant choisis, on pourrait, par le calcul, vérifier qu'il existe un e_3, orthogonal à e_1 et e_2, de norme 1, et tel que (e_1, e_2, e_3) engendre le même espace vectoriel que (a_1, a_2, a_3). Nous allons obtenir ce résultat directement et sans supposer $\mathcal{H}$ réel.

Appelons E_n le sous-espace vectoriel engendré par $(a_1, a_2, \ldots, a_n)$.

Grâce à la linéaire indépendance des a_n, on a l'inclusion stricte :

$$E_n \subset E_{n+1}$$

Supposons avoir trouvé n vecteurs orthonormaux ($e_1, e_2, \ldots, e_n$), qui engendrent le même espace vectoriel E_n que les n vecteurs ($a_1, a_2, \ldots, a_n$).

L'espace vectoriel E_n est de dimension finie (n), donc fermé dans l'espace de Hilbert E_{n+1}. Grâce au théorème de projection et puisque E_n est strictement inclus dans E_{n+1}, il existe au moins un vecteur e'_{n+1} non nul, orthogonal à E_n et appartenant à E_{n+1}. (Prendre la différence entre e_n et sa projection sur E_n).

Posons
$$e_{n+1} = \frac{e'_{n+1}}{\| e'_{n+1} \|}$$

on constate que e_{n+1} est orthogonal à tous les e_k ($k \leq n$), de norme 1, et les conditions sur la dimension assurent que les ($e_1, e_2, \ldots, e_{n+1}$) engendrent l'espace vectoriel E_{n+1} .

Par récurrence, on construit ainsi la suite orthonormale (e_1 , e_2, $\ldots$, e_n , $\ldots$, suite éventuellement infinie si la suite des $\{a_n\}_{n \geq 1}$ l'est.

Si ($a_1, a_2, \ldots a_n, \ldots$) est supposée totale, la suite orthonormale $\{e_n\}_{n \geq 1}$ constitue une base hilbertienne. En effet, supposons que

$$< x, e_n > = 0 \quad \text{pour tout entier } n.$$

On a alors $< x, a_n > = 0$ pout tout entier n puisque a_n appartient au sous-espace vectoriel engendré par les ($e_1, e_2, \ldots e_n$).

Ceci entraîne, grâce à l'hypothèse de totalité des $\{a_n\}_{n \geq 1}$, que $x = 0$, d'où la totalité de $\{e_n\}_{n \geq 1}$.

Venons-en maintenant au théorème fondamental sur les bases hilbertiennes. Nous supposons toujours que $\mathcal{H}$ est séparable et de dimension infinie puisque le cas de la dimension finie est bien connu (cf chapitre 0 et § 3.4.3.)

3.5. PROPRIETES DE BASES HILBERTIENNES.

3.5.1. Théorème.

Soit $\{e_n\}_{n \geq 1}$ une base orthonormale hilbertienne sur un espace de Hilbert séparable complexe pour fixer les idées. Soit x un vecteur dont les coordonnées hilbertiennes sont :

$$\lambda_n = < x, e_n >$$

(1) La série $\sum_{n=1}^{n=+\infty} \lambda_n e_n$ converge au sens de l'espace $\mathcal{H}$ vers le vecteur x

(2) La série numérique $\sum_{n=1}^{n=\infty} |\lambda_n|^2$ converge vers $\|x\|^2$

(3) Réciproquement, si μ_n est une suite de nombres complexes telle que $\sum_{n=1}^{\infty} |\mu_n|^2$ soit convergente, la série $\sum_{n=1}^{n=+\infty} \mu_n e_n$ converge au sens de $\mathcal{H}$ vers un vecteur x de $\mathcal{H}$ tel que $< x , e_n > = \mu_n$

Démonstration : calculons en utilisant $\lambda_k = < x, e_k >$

$$\|x - \sum_{k=1}^{k=n} \lambda_k e_k\|^2 = < x - \sum_{k=1}^{k=n} \lambda_k e_k, x - \sum_{k=1}^{k=n} \lambda_k e_k >$$

Grâce aux relations d'orthonormalité, il vient

$$\|x - \sum_{k=1}^{k=n} \lambda_k e_k\|^2 = \|x\|^2 - \sum_{k=1}^{k=n} |\lambda_k|^2$$

Comme $\|x - \sum_{k=1}^{k=n} \lambda_k e_k\| \geqslant 0$, on en déduit pour tout entier $n \geqslant 1$

$$\boxed{\begin{aligned} \|x\|^2 &\geqslant \sum_{k=1}^{k=n} |\lambda_k|^2 \\ &\geqslant \sum_{k=1}^{k=n} |< x , e_k >|^2 \end{aligned}}$$

inégalité de Bessel

Ceci prouve donc que la série à termes positifs $\sum_{1}^{\infty} |\lambda_k|^2$ est convergente puisque ses sommes partielles sont bornées par $\|x\|^2$. Envisageons alors la suite $x^{(n)} = \sum_{k=1}^{k=n} \lambda_k e_k$. La suite de vecteurs $x^{(n)}$ est une suite de Cauchy sur $\mathcal{H}$ puisque pour $n > m$

$$\|x^{(n)} - x^{(m)}\|^2 = \sum_{m+1}^{n} |\lambda_k|^2 \quad \text{et que} \quad \sum_{k=1}^{k=\infty} |\lambda_k|^2 \quad \text{converge.}$$

Par conséquent $x^{(n)}$ converge vers un élément noté x^{∞} puisque $\mathcal{H}$ est complet.

Calculons $< x^{\infty}, e_k >$ pour tout entier $k \geqslant 1$. Pour chaque entier k, $x \to < x, e_k >$ est une forme linéaire continue sur $\mathcal{H}$. On dispose donc de

$$< x^{\infty}, e_k > = \lim_{n \to \infty} < x^{(n)}, e_k > = \lim_{n \to \infty} < \sum_{i=1}^{i=n} \lambda_i e_i , e_k > = \lambda_k$$

On constate donc que

$$< x^{\infty}, e_k > = < x, e_k >$$

pour tout nombre k, c'est-à-dire que pour tout $k \geqslant 1$:

$$< x^{\infty} - x , e_k > = 0$$

Or le système $\{e_k\}_{k\geq 1}$ est total, ce qui assure la nullité de $x^\infty - x$ et l'égalité $x^\infty = x$.

Nous avons ainsi démontré (1) et (2) (ainsi que la réciproque comme il est facile de s'en assurer puisque dans la deuxième partie de la démonstration, on a seulement utilisé la convergence de $\sum_{n=1}^{\infty} |\lambda_n|^2$).

Conséquences : la notion de base hilbertienne généralise, avec des sommes infinies, la notion usuelle de base orthonormale en dimension finie. Au risque de nous répéter, rappelons qu'une base hilbertienne, en dimension infinie, ne constitue pas une base au sens algébrique envisagé au Chapitre 0.

Donnons l'expression du produit scalaire $<x,y>$ de deux vecteurs x et y de $\mathcal{H}$ dont les coordonnées hilbertiennes sont :

$$\begin{cases} x_n = < x , e_n > \\ y_n = < x , e_n > \end{cases}$$

$$<x,y> = < \sum_{k=1}^{\infty} x_k e_k , \sum_{n=1}^{\infty} y_n e_n > = \sum_{n=1}^{\infty} x_n y_n$$

$$\boxed{<x,y> = \sum_{n=1}^{\infty} x_n \overline{y_n}} \qquad \text{formule de Parseval}$$

(l'espace $\mathcal{H}$ est supposé complexe pour fixer les idées).

Le théorème précédent établit un isomorphisme entre un Hilbert $\mathcal{H}$ séparable et l'espace ℓ^2 lorsqu'une base hilbertienne est donnée sur $\mathcal{H}$. Prenons un espace réel pour changer.

Soit T la correspondance qui à x , élément de $\mathcal{H}$, fait correspondre l'élément Tx de ℓ^2 selon :

$$T(x) = (x_1, x_2, \ldots, x_n = <x , e_n>, \ldots).$$

$T(x)$, d'après le théorème, est un élément de ℓ^2 et la correspondance $x \longrightarrow T(x)$ de $\mathcal{H}$ dans ℓ^2 est bijective et isométrique

$$\| x \|^2 = \| T(x) \|^2 = \sum_{n=1}^{\infty} |<x , e_n>|^2$$

On peut donc dire que tous les espaces de Hilbert séparables de dimension infinie sont isomorphes (l'isomorphisme étant lié au choix d'une base) et ceci justifie notre démonstration du théorème du choix envisagée seulement sur l'espace l^2.

Remarque : soit $\{e_n\}$ une famille orthonormale telle que si l'on pose

$$\lambda_n = \langle x, e_n \rangle$$

on ait

$$\boxed{\|x\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2}$$

La famille $\{e_n\}_{n\geq 1}$ est alors une base hilbertienne.

Il suffit de démontrer que le système $\{e_n\}_{n\geq 1}$ est total. Supposons donc que $\langle x, e_n \rangle = 0$ pour tout n. On en déduit que $\|x\|^2 = 0$, donc $x = 0$ ce qui justifie le théorème donné.

On peut réunir les propriétés précédentes dans un seul théorème.

3.5.2. Théorème des bases hilbertiennes :

Soit $\{e_n\}_{n\geq 1}$ une famille orthonormale d'un espace de Hilbert $\mathcal{H}$.

Les propriétés suivantes sont équivalentes :

a) la famille $\{e_n\}_{n\geq 1}$ est totale

b) pour tout $x \in \mathcal{H}$, la série $\sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$ converge dans $\mathcal{H}$ vers l'élément x. Cette série s'appelle la série de Fourier (généralisée en base $\{e_n\}_{n>1}$ de x.

c) pour tout x, on a $\|x\|^2 = \sum_{n=1}^{n=\infty} |\langle x, e_n \rangle|^2$ relation de Parseval

d) les conditions $\langle x, e_n \rangle = 0$ pour tout entier $n \geq 1$, entraînent $x = 0$.

Remarque : Les implications $b \Leftrightarrow c \Rightarrow d$; $a \Rightarrow d$ et $b \Rightarrow a$ ne nécessitent pas que $\mathcal{H}$ soit un espace de Hilbert mais seulement que $\mathcal{H}$ soit un préhilbertien. (Cf § 3.2.2)

3.5.3. Exemple. L'espace $L^2[-\pi, +\pi]$

Nous allons construire un système orthonormal total de grande importance sur l'espace $L^2[-\pi, +\pi]$. On convient de prendre comme norme sur $L^2[-\pi, +\pi]$ (espace réel pour fixer les idées)

$$\|f\|_2 = \left[\int_{-\pi}^{+\pi} |f(x)|^2 dx\right]^{\frac{1}{2}}$$

Envisageons la suite de fonctions continues, appelée système trigonométrique normé (pour la norme choisie) :

$$\frac{1}{\sqrt{2\pi}}\ ,\ \frac{\cos x}{\sqrt{\pi}}\ ,\ \frac{\sin x}{\sqrt{\pi}}\ ,\ ---\ ,\ \frac{\cos nx}{\sqrt{\pi}}\ ,\ \frac{\sin nx}{\sqrt{\pi}}\ ,\ ---$$

1) C'est un système orthonormal

Il suffit de vérifier les relations suivantes

$$\frac{1}{\pi}\int_{-\pi}^{+\pi}\cos nx\ \sin mx\ dx = 0$$

$$\left.\begin{array}{l}\displaystyle\frac{1}{\pi}\int_{-\pi}^{+\pi}\cos nx\ \cos mx\ dx = \delta_{n,m}\\ \displaystyle\frac{1}{\pi}\int_{-\pi}^{+\pi}\sin nx\ \sin mx\ dx = \delta_{n,m}\end{array}\right\}\ \text{où}\ \delta_{n,m}\left\{\begin{array}{l}= 0 \text{ si } n \neq m\\ = 1 \text{ si } n = m\end{array}\right.$$

2) C'est un système total

Nous allons procéder en plusieurs étapes. La démonstration qui suit est assez longue et peut être considérée comme un exercice sur les méthodes précédentes.

Première étape :

Admettons que pour tout δ, $0 < \delta < \pi$, on ait une famille de polynômes trigonométriques $T_n(x)$ vérifiant les conditions

$$T_n(x) = \sum_{k=0}^{n} a_k \cos kx + b_k \sin kx$$

(1) $|T_n(x)| \leqslant 1$ $\quad \forall\, x \in [-\pi,-\delta] \cup [+\delta,+\pi]$

(2) $T_n(x) \geqslant 0$ sur $[-\delta,+\delta]$

(3) la suite $\dfrac{1}{T_n(x)}$ converge uniformément vers 0 sur tout sous-ensemble de la forme $[-\delta',+\delta']$ où $0 < \delta' < \delta < \pi$

Soit f une fonction de $L^2[-\pi,+\pi]$ telle que pour tout entier $n \geqslant 0$.

$$\int_{-\pi}^{+\pi} f(x)\cos nx\ dx = 0$$

et

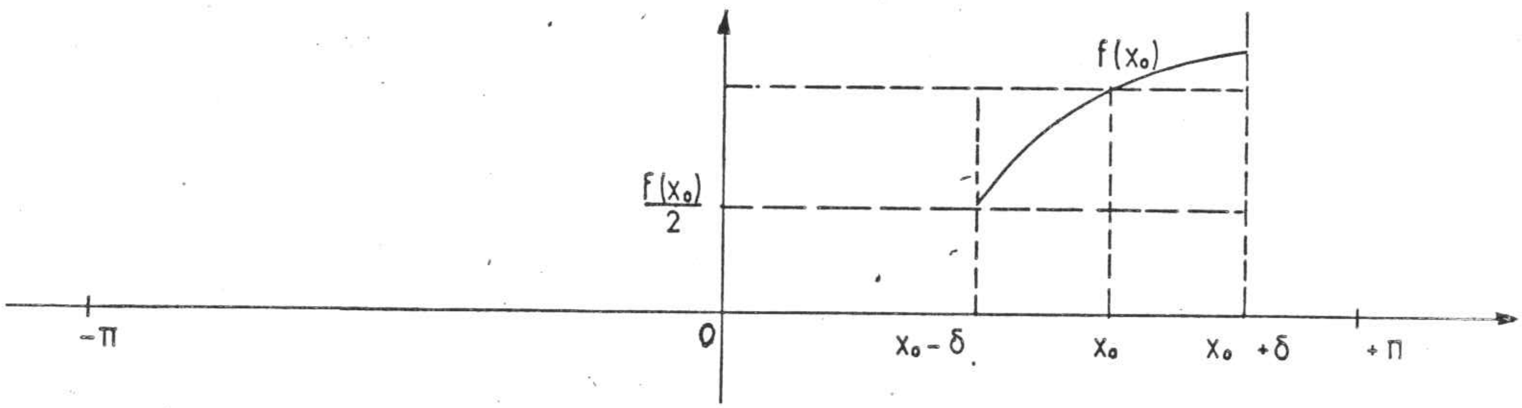
$$\int_{-\pi}^{+\pi} f(x)\sin nx\ dx = 0$$

($n = 0$ signifie $\int_{-\pi}^{+\pi} f(x)\,dx = 0$).

Nous voulons montrer que $\|f\|_2 = 0$. Raisonnons par l'absurde en supposant que $\|f\|_2 > 0$. Rappelons que f est une fonction de période 2π.

<u>Premier cas</u> : <u>faisons l'hypothèse supplémentaire de la continuité de</u> $f(x)$ <u>sur</u> $[-\pi, +\pi]$.

Si $\|f\|_2 > 0$, c'est qu'il existe un nombre x_0, que l'on peut supposer distinct de $+\pi$ et $-\pi$ tel que $|f(x_0)| > 0$. Bien entendu, quitte à multiplier f par une constante, on peut supposer $f(x_0) > 0$. Mais la continuité de $f(x)$ implique alors l'existence d'un nombre δ, $(0 < \delta < \pi)$ tel que, sur $[x_0 - \delta, x_0 + \delta]$ on ait $f(x) > \frac{f(x_0)}{2}$. Convenons aussi de prendre δ suffisamment petit pour que $[x_0 - \delta, x_0 + \delta] \subset [-\pi, +\pi]$ (cf figure)



$T_n(x)$ étant un polynôme trigonométrique, donc de la forme

$$a_0 + \sum_{k=0}^{k=n} a_n \cos kx + b_k \sin kx$$

on a grâce à l'hypothèse faite sur f,

$$\int_{-\pi}^{+\pi} f(x)\, T_n(x)\, dx = 0$$

et, bien entendu, puisque $T_n(x - x_0)$ est aussi un polynôme trigonométrique, comme $T_n(x)$

$$\int_{-\pi}^{+\pi} f(x)\, T_n(x - x_0)\, dx = 0$$

Minorons cette intégrale en la décomposant d'abord en trois parties :

$$\int_{-\pi}^{x_0-\delta} f(x) T_n(x - x_0)\, dx + \int_{x_0-\delta}^{x_0+\delta} f(x) T_n(x - x_0)\, dx + \int_{x_0+\delta}^{x_0+\pi} f(x)\, T_n(x - x_0)\, dx$$

La continuité de $f(x)$ implique l'existence d'un nombre M tel que

$$|f(x)| \leqslant M \quad \text{pour} \quad x \in [-\pi, +\pi]$$

Donc

$$\left|\int_{-\pi}^{x_0-\delta} f(x) T_n(x-x_0)\,dx\right| \leqslant M \int_{-\pi}^{x_0-\delta} |T_n(x-x_0)|\,dx$$

et de même

$$\left|\int_{x_0+\delta}^{+\pi} f(x) T_n(x-x_0)\,dx\right| \leqslant M \int_{x_0+\delta}^{+\pi} |T_n(x-x_0)|\,dx$$

Or sur l'ensemble

$[-\pi, -\delta] \cup [\delta, +\pi]$, le polynôme $T_n(x)$ est borné par 1 en valeur absolue. Par conséquent, puisque 2π est période de T_n, sur l'ensemble

$[-\pi, x_0 - \delta] \cup [x_0 + \delta, +\pi]$, le polynôme $T_n(x - x_0)$

est également borné par 1 en valeur absolue.
On dispose donc d'une majoration de deux des intégrales :

$$\left|\int_{-\pi}^{x_0-\delta} f(x) T_n(x-x_0)\,dx + \int_{x_0+\delta}^{+\pi} f(x) T_n(x-x_0)\,dx\right| \leqslant 2M(\pi-\delta) \leqslant 2M\pi$$

Par contre sur

$[x_0 - \delta', x_0 + \delta']$ où $0 < \delta' < \delta$; δ' fixé et

pour un nombre A donné aussi grand qu'on le désire, on peut trouver un nombre n_0 tel que pour tout $x \in [x_0 - \delta', x_0 + \delta']$ et tout $n \geqslant n_0$, on ait

$T_{n_0}(x - x_0) \geqslant A$.

Ce résultat provient de la convergence uniforme de $\dfrac{1}{T_n(x)}$ vers zéro sur $[-\delta' + \delta']$. Or

$$\int_{x_0-\delta}^{x_0+\delta} f(x) T_n(x-x_0)\,dx \geqslant \int_{x_0-\delta'}^{x_0+\delta'} f(x) T_n(x-x_0)\,dx$$

puisque $f(x)$ et $T_n(x - x_0)$ sont positifs sur $[x_0 - \delta, x_0 + \delta]$

Et donc, pour $n \geqslant n_0$

$$\int_{x_0-\delta'}^{x_0+\delta'} f(x) T_n(x-x_0)\,dx \geqslant \frac{f(x_0)}{2} A . 2\delta'$$

Donc pour tout $n \geqslant n_0$

$$\int_{x_0-\delta'}^{x_0+\delta'} f(x)\,T_n(x-x_0)\,dx \quad > \quad A\,\delta' f(x_0)$$

Utilisant la majoration des deux morceaux restants de l'intégrale, on obtient pour $n \geqslant n_0$, $\int_{-\pi}^{+\pi} f(x)\ T_n(x)dx \geqslant A\delta'\ f(x_0)\ -\ 2M\ \pi$

Par ailleurs
$$\int_{-\pi}^{+\pi} f(x)\,T_n(x)\,dx \ = \ 0$$
quel que soit n. En choisissant A de sorte que $A\delta'\ f(x_0)\ -\ 2M\pi > 0$, il y a une contradiction.

Donc $f(x)$ ne peut avoir aucune valeur distincte de 0 : $f \equiv 0$ et donc en particulier $\|f\|_2 = 0$

<u>Deuxième cas</u> : si f n'est pas supposée continue, le raisonnement fait au premier cas est inapplicable. On peut toutefois conclure mais nous devrons admettre une propriété particulière qui tient à ce que en fait L^2 n'est pas défini grâce à l'intégrale de Riemann mais en utilisant l'intégrale de Lebesgue.

Considérons donc $f \in L^2$ telle que

$$0 = \ < f,\cos nx > \ = \ < f,\sin nx >$$

pour tout n entier $\geqslant 0$

$$(n = 0 :\ < f, 1 > \ = \ 0)$$

et considérons
$$F(x) = \int_{-\pi}^{+x} f(t)\,dt$$

La fonction $x \longrightarrow F(x)$ est bien définie sur $[-\pi,+\pi]$ et de plus périodique et de période 2π

$$F(-\pi) = F(+\pi) = 0 \qquad \left(\text{puisque } \int_{-\pi}^{+\pi} f(t)\,dt = 0\right)$$

La fonction $x \longrightarrow F(x)$ est continue sur $[-\pi,+\pi]$: pour le démontrer, il suffit d'utiliser l'inégalité de Schwarz ; de

$$F(x) - F(y) \ = \ \int_y^x f(t)\,dt$$

on déduit ainsi :
$$|F(x)-F(y)| \leqslant \left|\int_y^x dt\right|^{\frac{1}{2}} \left(\left|\int_y^x |f(t)|^2 dt\right|\right)^{\frac{1}{2}}$$

d'où

$$|F(x) - F(y)| \leqslant \sqrt{|y-x|} \; \|f\|_{L^2}$$

Lorsque $y \longrightarrow x$, il est clair que $F(y) \longrightarrow F(x)$ (on remarque que lorsque $f \in L^2$, on ne peut pas dire que $F'(x) = f(x)$ en tous points x de $[-\pi + \pi]$ puisque $f(x)$ peut être infinie en certains points et $F(x)$ n'est pas partout dérivable : c'est ce genre de difficultés qui compliquent la démonstration).

Calculons

$$< F(x), \cos nx > \; = \int_{-\pi}^{+\pi} F(x) \cos nx \, dx$$

Par intégration par parties, pour $n \neq 0$, on a :

$$\left[\frac{F(x) \cos nx}{n} \right]_{-\pi}^{+\pi} - \frac{1}{n} \int_{-\pi}^{+\pi} f(x) \sin nx \, dx$$

En fait, on devrait raisonner en faisant une intégration double pour être rigoureux (la formule d'intégration par parties n'a pas été démontrée pour l'intégrale de Lebesgue, mais le théorème de Fubini est exact (Cf chap. II, § 6.3). Nous laissons au lecteur le soin d'aménager ce point de détail.
D'où $\quad < F(x), \cos nx > \; = 0$; Pour $n > 0$

De même

$$< F(x), \sin nx > \; = 0$$

pour toutes valeurs positives de n. Si $< F, 1 > = A_0$ et si $A_0 \neq 0$, il suffit de prendre $F(x) - A_0$ pour constater que cette nouvelle fonction continue satisfait l'orthogonalité à tous les polynômes trigonométriques. Par conséquent on déduit de l'étude du premier cas que :

$$F(x) \equiv A_0$$

Soit

$$\int_{-\pi}^{+x} f(t) \, dt \equiv A_0$$

Mais de

$$\int_{-\pi}^{+\pi} f(t) \, dt = 0$$

on déduit $A_0 = 0$, soit pour toutes les valeurs de x

$$\int_{-\pi}^{+x} f(t) \, dt = 0$$

Nous admettrons que ceci implique $f(x) = 0$ au sens L^2 (on démontre, assez difficilement, que sauf sur un ensemble de mesure nulle, on a $F'(x) = f(x)$, or $F'(x) = 0$, donc $f(x) = 0$ presque partout).

Remarque : En théorie de l'approximation, on définit le produit de convolution de deux fonctions de $L^2[-\pi,+\pi]$ et on montre que cette notion permet de construire une suite de fonctions continues $\{f_n\}_{n\geq 1}$ telle que (cf Nanta Iremica Vol N° 11).

(1) $$\lim_{n \to \infty} f_n = f \qquad \text{au sens de } L^2[-\pi,+\pi)$$

(2) le coefficient de Fourier d'ordre k de f_n est le produit du coefficient d'ordre k de f par une constante (qui dépend de n et de k).

Par conséquent, si f est orthogonal au système trigonométrique, il en est de même de f_n. Or f_n est identiquement nulle pour tout n d'après le résultat du premier cas donc f l'est au sens de $L^2[-\pi, +\pi]$. Ceci termine la démonstration du second cas en évitant le recours à un théorème délicat de théorie de la mesure.

Deuxième étape :

a) supposons avoir construit un polynôme trigonométrique $T(x)$ tel que l'on ait les inégalités suivantes

$$\begin{cases} T(x) > 1 & \text{sur } [-\delta',+\delta'] \text{ pour tout } \delta' \text{ ; } 0 < \delta' < \\ |T(x)| \leqslant 1 & \text{sur } [-\pi,-\delta] \cup [+\delta,+\pi] \end{cases}$$

$T(x)$ s'écrit

$$a_0 + \sum_{j=1}^{j=k} a_j \cos jx + b_j \sin jx$$

Envisageons $T^n(x)$, la puissance $n^{\text{ième}}$ de T où n est un entier. C'est encore un polynôme trigonométrique puisque des expressions comme $\cos^n jx$ et $\sin^n jx$ s'expriment comme combinaison linéaires de $\sin px$ et de $\cos px$

De plus

$$|T^n(x)| \leqslant |T(x)| \leqslant 1 \quad \text{sur } [-\pi,-\delta] \cup [+\delta,+\pi]$$

et sur

$$[-\delta',+\delta'] \qquad \text{avec} \qquad 0 < \delta' < \delta$$

on a

$$T^n(x) > \left(\inf_{x \in [-\delta' + \delta']} T(x) \right)^n$$

Or

$$\inf_{x \in [-\delta', + \delta']} T(x) > 1$$

Donc $\dfrac{1}{T^n(x)}$ converge uniformément vers zéro sur $[-\delta', + \delta']$

En posant $T_n(x) = T^n(x)$, on se conforme aux prescriptions de la première question.

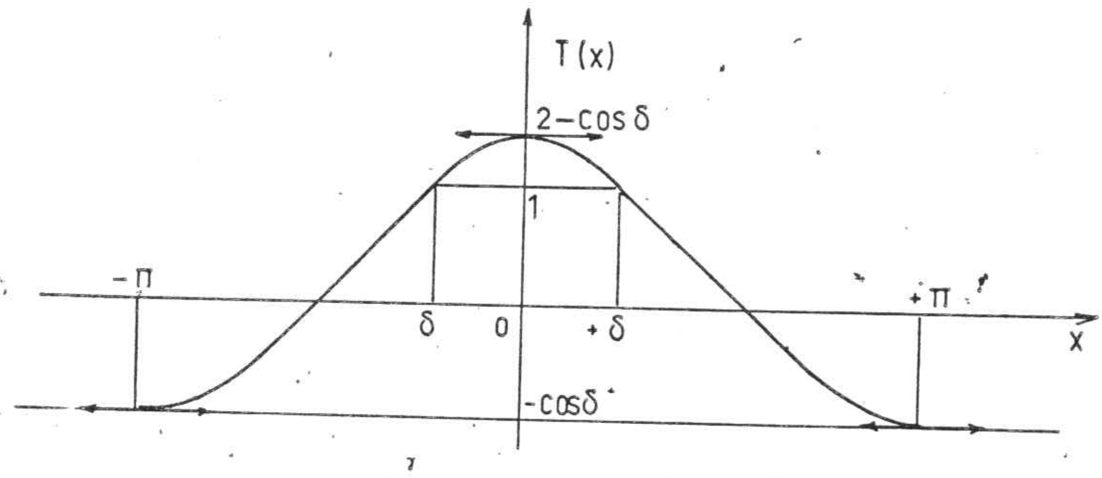
b) construisons maintenant $T(x)$, ce qui ne pose aucun problème ;

$$T(x) = 1 + \cos x - \cos \delta$$

$$T'(x) = -\sin x$$

$$T''(x) = -\cos x$$

x	$-\pi$		$-\frac{\pi}{2}$		$-\delta$	0	$+\delta$		$+\frac{\pi}{2}$		$+\pi$
T''	1	$+$	0	$-$		1		$-$	0	$+$	1
T'	0			$+$		0		$-$			0
T	$-\cos\delta$			$\nearrow$		$2-\cos\delta$		$\searrow$			$-\cos\delta$



On constate bien que pour tout δ', $0 < \delta' < \delta < \pi$, on a :

$$\begin{cases} T(x) > 1 & \text{sur } [-\delta', + \delta'] \\ |T(x)| \leqslant 1 & \text{sur } [-\pi, -\delta] \cup [+\delta, +\pi] \end{cases}$$

La démonstration est donc terminée.

Nous allons en déduire quelques rudiments d'analyse harmonique classique.

Nous avons démontré le théorème suivant :

Théorème : la suite $\frac{1}{\sqrt{2\pi}},\dots,\frac{\cos nx}{\sqrt{\pi}},\frac{\sin nx}{\sqrt{\pi}}$ est une base hilbertienne de $L^2[-\pi,+\pi]$

Ce théorème a de remarquables conséquences.(cf théorème des bases hilbertiennes)

Si les nombres réels $a_0, a_1, b_1,\dots, a_n, b_n,\dots$ sont tels que la série de terme général

$$|a_n|^2 + |b_n|^2$$

soit convergente, la série

$$(1) \qquad \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + ---- + a_n \cos nx + b_n \sin x + ---$$

converge en moyenne quadratique vers une fonction f, qui admet les quantités

$$\sqrt{\frac{\pi}{2}a_0}, ---, \sqrt{\pi a_n}, \quad \sqrt{\pi b_n}, ---$$

comme coordonnées hilbertiennes relativement à la base hilbertienne formée par le système trigonométrique :

$$\frac{1}{\sqrt{2\pi}}, ---, \frac{1}{\sqrt{\pi}}\cos nx, \quad \frac{1}{\sqrt{\pi}}\sin nx, ---$$

On a donc

$$\sqrt{\frac{\pi}{2}}a_0 = \int_{-\pi}^{+\pi}\frac{f(x)}{\sqrt{2\pi}}\,dx, --- \sqrt{\pi a_n} = \int_{-\pi}^{+\pi}\frac{\cos nx}{\sqrt{\pi}}f(x)\,dx$$

$$\sqrt{\pi b_n} = \int_{-\pi}^{+\pi}\frac{\sin nx}{\sqrt{\pi}}f(x)\,dx$$

ou

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{+\pi}f(x)\,dx, ---, a_n = \frac{1}{\pi}\int_{-\pi}^{+\pi}\cos nx\, f(x)\,dx$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{+\pi}\sin nx\, f(x)\,dx$$

Or ces formules ne sont autres que celles qui expriment $a_0,\dots, a_n, b_n,\dots$ comme <u>coefficients de Fourier</u> de f. <u>La série (1) s'appelle alors série de Fourier de f.</u>

Réciproquement, si $f(x)$ est mesurable et de carré intégrable sur $[-\pi,+\pi]$, on sait que la série (1) existe et converge en moyenne quadratique vers $f(x)$. Cela veut dire, rappelons-le, que

$$\lim_{n\to\infty} \int_{-\pi}^{+\pi} \left| f(x) - \frac{a_0}{2} - \cdots - a_n \cos nx - b_n \sin nx \right|^2 dx = 0$$

Ce résultat est équivalent au suivant :

Théorème : si $f \in L^2$, sur l'intervalle $[-\pi, +\pi]$, les coefficients de Fourier de $f(x)$ satisfont à l' égalité de Parseval.

$$\boxed{\frac{a_0^2}{2} + \cdots + a_n^2 + b_n^2 + \cdots = \frac{1}{\pi} \int_{-\pi}^{+\pi} |f(x)|^2 \, dx}$$

Cette formule est naturellement valable sur tout intervalle Δ de longueur 2π. Sur un intervalle Δ de longueur différente de 2π, il faut la modifier légèrement en construisant la suite de fonctions trigonométriques orthogonales et normées sur $L^2(\Delta)$. Si Δ a pour longueur T, c'est la suite.

$$\frac{1}{\sqrt{T}}, \cdots, \sqrt{\frac{2}{T}} \cos \frac{2\pi}{T} nx \; \sqrt{\frac{2}{T}} \sin \frac{2\pi}{T} nx, \cdots$$

Si l'on pose

$$a_n = \frac{2}{T} \int_0^T f(x) \cos \frac{2\pi}{T} nx \, dx \; ; \; b_n = \frac{2}{T} \int_0^T f(x) \sin \frac{2\pi}{T} nx \, dx$$

on a

$$\frac{a_0}{2} + \cdots + a_n^2 + b_n^2 + \cdots = \frac{2}{T} \int_0^T |f(x)|^2 dx$$

Remarque : naturellement, on peut souhaiter d'autres types de convergence de la série de Fourier de f lorsque f est suffisamment régulière. Nous supposerons ici que f est une fonction périodique et de période 2π.

1°) si $f(x)$ est continue et à variation bornée, on démontre que sa série de Fourier est uniformément convergente sur tout intervalle. Le fait que la série (1) converge aussi en moyenne quadratique vers $f(x)$ résulte donc d'un théorème élémentaire,

2°) si $f(x)$ est continue et dérivable, sauf en des points où f et f' peuvent avoir des sauts de première espèce, a_n et b_n tendent vers zéro au moins comme $\frac{1}{n}$ et l'on sait d'avance que la série $\sum_{n\geq 1} (a_n^2 + b_n^2)$ est convergente.

Dans ces deux cas, la série de Fourier converge donc vers f de plusieurs façons :

1°) elle converge vers $f(x)$ en tout point x où f est continue. On peut même préciser qu'en un point de discontinuité, la série de Fourier de f converge vers la demi-somme de la limite à droite et de la limite à gauche de la fonction f en x.
2°) elle converge uniformément sur tout intervalle fermé où $f(x)$ est continue

3°) elle converge en moyenne quadratique sur l'intervalle $[-\pi, +\pi]$.

Des théorèmes beaucoup plus précis sont développés dans tout cours d'analyse harmonique (par exemple : J. DHOMBRES, Stage de Perfectionnement, E.N.S.T.A. 1973 où l'aspect numérique est détaillé).

Application : théorème d'unicité.

Si f et g sont deux fonctions de L^2 $(-\pi, +\pi)$ ayant les mêmes coefficients de Fourier, elles sont égales au sens de L^2, c'est-à-dire que $f(x) = g(x)$ presque partout (Cf Chapitre II § 3.3)

(on peut remplacer "définies sur l'intervalle $[-\pi, +\pi]$ " par "périodique de période 2π ").

3.6. PROPRIETES DES POLYNOMES ORTHOGONAUX.

Soit un espace de fonctions muni d'un produit scalaire et contenant des polynômes. Nous cherchons à construire une base hilbertienne sur cet espace dont chaque élément soit un polynôme.

3.6.1. Polynômes de Legendre.

Pour fixer les idées, nous nous plaçons sur l'espace de Hilbert $L^2[-1, +1]$ et recherchons les polynômes réels P_n qui constituent une famille orthogonale dans $L^2[-1, +1]$ relativement au produit scalaire $<f, g> = \int_{-1}^{+1} f(x) g(x)\, dx$
Le but est donc, à partir de l'espace vectoriel constitué par les polynômes, dont une base est $1, x, x^2, \dots, x^n, \dots$ de déterminer un système orthonormal engendrant le même espace. Le procédé utilisé, dit d'orthonormalisation de Gram-Schmidt, permet de faire cette construction. Donnons un calcul effectif:

On part de $P_0(x) = \frac{1}{\sqrt{2}}$ (puisque $< \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} > = \sqrt{\left(\int_{-1}^{+1} \frac{dx}{2} \right)} = 1$

On prend pour $P_1(x)$ un polynôme quelconque du premier degré auquel on impose les deux conditions d'orthonormalité :

$$< P_1 , P_1 > = \int_{-1}^{+1} P_1^2 (x) dx = 1$$

et
$$< P_1 , P_0 > = \int_{-1}^{+1} P_1 (x) P_0 (x) dx = 0$$

En posant $P_1 (x) = ax + b$, il vient

$$a = \pm \sqrt{\frac{3}{2}} \quad \text{et} \quad b = 0$$

On prend maintenant pour $P_2(x)$ un polynôme quelconque du second degré auquel on impose les trois conditions d'orthonormalité

$$< P_2 , P_2 > = \int_{-1}^{+1} P_2^2(x) dx = 1$$
$$< P_2 , P_1 > = \int_{-1}^{+1} P_2(x) P_1(x) dx = 0$$
$$< P_2 , P_0 > = \int_{-1}^{+1} P_2(x) P_0(x) dx = 0$$

Si $P_2 (x) = ax^2 + bx + c$, cela détermine a, b et c (à un signe près)

$$a = \frac{3}{2} \sqrt{\frac{5}{2}} \qquad b = 0 \qquad \text{et} \qquad c = - \frac{1}{2} \sqrt{\frac{5}{2}}$$

Par récurrence, on voit que l'on peut ainsi construire une suite orthonormale de polynômes P_n (le degré de P_n est n). Cette suite n'est d'ailleurs pas déterminée, puisque l'on a un certain choix sur les signes des coefficients. Toutefois, le théorème d'orthonormalisation nous assure que l'on peut construire la suite infinie $\{P_n\}_{n \geqslant 1}$. L'usage cependant est de ne pas normer les polynômes P_n mais d'imposer une condition qui détermine complètement P_n, à savoir $P_n(1) = 1$. Nous verrons plus loin que cette condition est possible et détermine la suite $\{P_n\}_{n \geqslant 1}$. Il vient pour les premiers termes: $P_0(x) = 1$.

$$P_1(x) = x$$
$$P_2(x) = \frac{3x^2 - 1}{2}$$
$$P_3(x) = \frac{5x^3 - 3x}{2} \quad \text{etc...}$$

La famille ainsi construite possède les trois propriétés caractéristiques :

(1) $P_n(1) = 1$

(2) le degré de $P_n(x)$ est n

(3) $\int_{-1}^{+1} P_n(x) P_m(x) dx = 0$ si $n \neq m$

La famille orthogonale ainsi construite s'appelle la famille des polynômes de Legendre.

La suite orthonormale associée est la suite $\sqrt{\frac{2n+1}{2}}\, P_n(x)$ comme on le démontrera un peu plus loin.

Par construction, au rang n, la base orthogonale $P_0, P_1, \dots, P_n$ engendre le même espace vectoriel, de dimension $n+1$, que la base $1, x, x^2, \dots, x^n$

On peut donner une formule explicite de l'expression du $n^{\text{ième}}$ polynôme au moyen de déterminants, mais l'utilité de cette formule est réduite, aussi nous ne la démontrerons pas (Cf exercice n° 34).

Au contraire, nous allons montrer quelques types de relations entre polynômes orthogonaux d'une même famille. Nous restons sur l'espace $L^2[-1,+1]$ mais le lecteur s'apercevra bien vite que nos raisonnements s'adaptent facilement au cas très général d'un espace L^2 avec poids. Nous reviendrons d'ailleurs sur ces propriétés en théorie de l'approximation (Cf. J. DHOMBRES, Nanta Iremica Volume N° 11).

Nous gardons donc la famille orthonormale $[P_n]_{n\geqslant 1}$ de polynômes réels sur $L^2[-1,+1]$, c'est-à-dire satisfaisant les deux propriétés d'ailleurs non caractéristiques.

(1) le degré de P_n est n

$$\text{(2)} \int_{-1}^{+1} P_n(x)\, P_m(x)\, dx = 0 \qquad \text{si } n \neq m$$

$$= 1 \qquad \text{si } n = m$$

Ces relations ne déterminent pas complètement la famille $[P_n]_{n\geqslant 1}$. Toutefois si $[Q_n]_{n\geqslant 1}$ est une autre famille (réelle) satisfaisant (1) et (2), on doit avoir

$$P_n = \alpha_n Q_n \qquad \text{où} \quad \alpha_n = \pm 1$$

L'idée de la démonstration va être constamment exploitée. Puisque les Q_0, $Q_1, \dots, Q_n$ engendrent l'espace des polynômes de degré inférieur ou égal à n, on doit avoir

$$P_n = \sum_{k=0}^{n} \alpha_k Q_k$$

Mais pour $k < n$, $\langle P_n, Q_k\rangle = 0$ puisque Q_k à son tour est combinaison linéaire de P_i où $0 \leqslant i \leqslant k$. Donc par produit scalaire

$$P_n = \alpha_n Q_n$$

Cependant $1 = < P_n , P_n > = \alpha_n < Q_n , P_n > = \alpha_n^2 < Q_n , Q_n >$.

Soit la relation $\alpha_n^2 = 1$ qui conduit à $\alpha_n = \pm 1$ puisque nous sommes restés sur des polynômes réels.

3.6.2. Relations de récurrence.

Théorème : entre trois polynômes orthonormaux successifs, on a la relation de récurrence

(1) $$P_n(x) = (A_n x + B_n) P_{n-1}(x) - C_n P_{n-2}(x) \quad (n \geqslant 2)$$

De plus, si a_n est le terme de plus haut degré de $P_n(x)$, alors

$$A_n = \frac{a_n}{a_{n-1}} \qquad \text{et } C_n = \frac{A_n}{A_{n-1}} = \frac{a_n a_{n-2}}{a_{n-1}^2}$$

Choisissons $A_n = \frac{a_n}{a_{n-1}}$ de sorte que $P_n(x) - A_n x P_{n-1}(x)$ soit un polynôme de degré n-1. Remarquons que tout polynôme de degré n-1 s'exprime comme une combinaison linéaire de $P_0, P_1, \ldots, P_{n-1}$, par construction. Ainsi :

$$P_n(x) - A_n x P_{n-1}(x) = \sum_{k=0}^{n-1} \alpha_k P_k(x) \quad \text{où } \alpha_k = < P_n(x) - A_n x P_{n-1}(x), P_k(x) > .$$

Mais

$$\alpha_k = < P_n(x), P_k(x) > - A_n < P_{n-1}(x), x P_k(x) >$$

les deux produits scalaires sont nuls puisque P_{n-1} est orthogonal à tout polynôme de degré inférieur. dès lors que $k < n-2$. Par suite

$$P_n(x) - A_n x P_{n-1}(x) = \alpha_{n-1} P_{n-1}(x) + \alpha_{n-2} P_{n-2}(x)$$

ce qui est bien la formule requise avec $\alpha_{n-1} = B_n$ et $\alpha_{n-2} = -C_n$

En outre :

$$< P_n , P_{n-2} > = 0 = A_n < x P_{n-1}(x), P_{n-2}(x) > + B_n < P_{n-1}, P_{n-2} > - C_n < P_{n-2}, P_{n-2} >$$

$$A_n < x P_{n-1}(x), P_{n-2}(x) > = C_n \text{ puisque } < P_{n-2}, P_{n-2} > = 1$$

Mais $x P_{n-2}(x) = a_{n-2} x^{n-1} + \ldots$ les termes suivants étant de degré inférieur à n-1. Donc

$$< x P_{n-1}(x), P_{n-2}(x) > = < P_{n-1}(x), x P_{n-2}(x) >$$
$$= a_{n-2} < P_{n-1}(x), x^{n-1} >$$

Mais on a encore

$$< P_{n-1}(x), P_{n-1}(x) > = < P_{n-1}(x), a_{n-1} x^{n-1} >$$

donc

$$< x P_{n-1}(x), P_{n-2}(x) > = \frac{a_{n-2}}{a_{n-1}} < P_{n-1}(x), P_{n-1}(x) > = \frac{a_{n-2}}{a_{n-1}}$$

Soit :

$$C_n = \frac{a_{n-2}}{a_{n-1}} A_n = \frac{a_{n-2}}{a_{n-1}} \frac{a_n}{a_{n-1}} = \frac{a_n a_{n-2}}{a^2_{n-1}}$$

On vérifierait qu'en appelant b_n le terme de degré n-1 dans l'expression de P_n, on dispose de la relation

$$B_n = A_n\left(\frac{b_n}{a_n} - \frac{b_{n-1}}{a_{n-1}}\right) \quad \text{pour} \quad n = 2, 3 \ldots$$

3.6.3. Théorème de Darboux-Christoffel.

On a la relation de la récurrence

$$(2) \quad K(x,y) = \sum_{k=0}^{k=n} P_k(x) \; P_k(y) = \frac{a_n}{a_{n+1}} \quad \frac{P_{n+1}(x)\, P_n(y) - P_n(x) P_{n+1}(y)}{x-y}$$

On calcule $P_{n+1}(x) \; P_n(y) - P_n(x) \, P_{n+1}(y)$ à partir de la formule de récurrence.(1)

Il vient

$$P_{n+1}(x) \, P_n(y) - P_n(x) P_{n+1}(y) = A_{n+1}(x-y) P_n(x) P_n(y) + C_{n+1}\left(P(x) P_{n-1}(y) - P_{n-1}(x) \; P_n(y)\right)$$

D'où

$$\frac{1}{A_{n+1}} \frac{P_{n+1}(x)\, P_n(y) - P_n(x)\, P_{n+1}(y)}{x-y} = P_n(x) \; P_n(y) + \frac{1}{A_n} \; \frac{P_n(x) \; P_{n-1}(y) - P_n(y) \; P_{n-1}(x)}{x-y}$$

et par récurrence :

$$= P_n(x) \; P_n(y) + P_{n-1}(x) \; P_{n-1}(y) + \cdots + \frac{P_1(x)\, P_0(y) - P_1(y) \; P_0(x)}{A_1 \; (x-y)}$$

$$= P_n(x) \, P_n(y) + \cdots + P_1(x) \, P_1(y) + \frac{A_1}{A_1} a_0^2$$

Mais on a bien $\qquad a_0 = P_0(x) = P_0(y)$

Cette formule permet le calcul effectif de la projection orthogonale d'un fonction sur l'espace des polynômes de degré inférieur ou égal à n.

Conformément à la théorie, cette projection d'une fonction quelconque f de $L^2[-1,+1]$ est donnée par :

$$\sum_{k=0}^{k=n} < f, P_k > P_k(x) = \sum_{k=0}^{k=n} P_k(x) \int_{-1}^{+1} f(y) P_k(y)\, dy$$

$$= \int_{-1}^{+1} \Big(\sum_{k=0}^{k=n} P_k(x)\ P_k(y)\Big) f(y)\, dy$$

Soit la formule :

$$(3) \quad \sum_{k=0}^{k=n} < f, P_k > P_k(x) = \frac{1}{A_{n+1}} \int_{-1}^{+1} \frac{P_{n+1}(x)\ P_n(y) - P_n(x)\ P_{n+1}(y)}{x-y} f(y)\, dy$$

En faisant $f = 1$ dans (3) on obtient $\frac{1}{A_{n+1}} \int_{-1}^{+1} \frac{P_{n+1}(x)\ P_n(y) - P_n(x)\ P_{n+1}(y)}{x - y}\, dy = 1$

Donc on estime le reste selon :

$$(4) \quad R_n(f)(x) = f(x) - \sum_{k=0}^{k=n} < f, P_k > P_k(x)$$

$$= \frac{1}{A_{n+1}} \int_{-1}^{+1} \Big[P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)\Big] \frac{f(x)-f(y)}{x - y}\, dy$$

La fonction $K_n(x,y) = \sum_{k=0}^{n} P_k(x) P_k(y)$ s'appelle le noyau d'ordre n de la famille de polynômes orthogonaux P_n.

Théorème du noyau reproduisant :

On a la formule $\int_{-1}^{+1} P(x) K_n(x,y)\, dx = P(y)$ pour tout polynôme P de degré inférieur ou égal à n. Cette formule est caractéristique de la fonction K_n.

Grâce au théorème d'orthogonalisation, on peut écrire pour un polynôme P de degré $\leqslant n$:

$$P(x) = \sum_{k=0}^{n} < P, P_k > P_k(x)$$

donc

$$P(x) = \sum_{k=0}^{k=n} \Big(\int_{-1}^{+1} P(y)\ P_k(y)\, dy\Big)\ P_k(x)$$

$$= \int_{-1}^{+1} \Big(\sum_{k=0}^{k=n} P_k(x)\ P_k(y)\Big)\ P(y)\, dy$$

$$= \int_{-1}^{+1} K_n(x,y)\ P(y)\, dy$$

et bien entendu:

$$K_n(x,y) = K_n(y,x)$$

Nous laissons la réciproque à titre d'exercice.

Les propriétés des zéros des polynômes orthogonaux ont fait l'objet de travaux nombreux. En liaison avec la théorie de l'approximation, nous énoncerons

Théorème des coïncidences :

Soit f une fonction continue sur $[-1,+1]$. La projection orthogonale de f sur l'espace des polynômes de degré inférieur ou égal à n coïncide avec f au moins en $(n+1)$ points de $[-1,+1]$.

Comme nous l'avons déjà dit, cette projection s'écrit avec les polynômes orthonormaux P_n :

$$\sum_{k=0}^{n} < f, P_k > P_k(x)$$

et la différence avec f , notée $R_n f(x)$ s'écrit :

$$R_n f(x) = f(x) - \sum_{k=0}^{n} < f, P_k > P_k(x)$$

On caractérise $R_n f(x)$ par son orthogonalité à tout polynôme de degré inférieur ou égal à n. En particulier $\int_{-1}^{1} P_0(x) R_n f(x) dx = 0$ donc $R_n(f)$ s'annule au moins une fois dans $[-1,+1]$.

Soient $x_1, x_2, \ldots, x_k$ les seuls zéros distincts de R_n en lesquels R_n change de signe dans $[-1,+1]$, rangés par ordre croissant. $R_n(x)(x-x_1)\ldots(x-x_k)$ a un signe constant dans $[-1, +1]$ et pourtant si l'on suppose $k \leq n$

$$\int_{-1}^{+1} R_n(x)(x-x_1)\cdots(x-x_k)\,dx = 0.$$

Par suite $k > n$ et le théorème est démontré.

On tire d'intéressantes conséquences de ce fait en théorie de l'approximation (Cf. J. DHOMBRES, Nanta Iremica Volume 11).

3.6.4. Formule de Rodrigues.

Théorème : les polynômes orthogonaux de Legendre satisfont pour tout $n \geq 0$

$$P_n(x) = \frac{(-1)^n}{2^n(n!)} \frac{d^n}{dx^n}\left((1-x^2)^n\right)$$

Posons $Q_n(x) = \frac{d^n}{dx^n}\left((1-x^2)^n\right)$. Montrons que Q_n est un polynôme orthogonal à tout polynôme de degré inférieur à n. Soit P un tel polynôme, effectuons une intégration par parties :

$$\begin{aligned}\int_{-1}^{+1} Q_n(x) P(x) dx &= \int_{-1}^{+1} \frac{d^n((1-x^2)^n)}{dx^n} P(x) dx \\ &= \left[\frac{d^{n-1}(1-x^2)^n}{dx^{n-1}} P(x)\right]_{-1}^{+1} - \int_{-1}^{+1} \frac{d^{n-1}(1-x^2)^n}{dx^{n-1}} P'(x) dx\end{aligned}$$

Mais le terme tout intégré s'annule car on peut vérifier facilement que

$\frac{d^k((1-x^2)^n)}{dx^k}$ s'annule en 1 et en -1 pour tout k tel que $0 \leqslant k < n$.

En intégrant par parties n fois, il vient

$$\int_{-1}^{+1} Q_n(x)\,P(x)\,dx = (-1)^n \int_{-1}^{+1} (1-x^2)^n P^{(n)}(x)\,dx.$$

Mais $P^{(n)}(x) = 0$ puisque P est de degré au plus égal à n - 1.

Par suite, Q_n est nécessairement un multiple du polynôme de Legendre P_n, puisque seule la condition de normalité n'est pas satisfaite.

$P_n = \lambda_n Q_n$ et il importe maintenant de déterminer le coefficient λ_n.

Par définition des polynômes de Legendre $P_n(1) = 1$ (ces polynômes ne sont, en effet, pas normés). Donc $\lambda_n = \frac{1}{Q_n(1)}$. Pour calculer $Q_n(1)$, on utilise la formule de Leibniz, laquelle fournit la relation

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

où $\Gamma(n)$ représente la valeur de la fonction gamma au point n, à savoir

$$\Gamma(n) = \overline{n-1}\,!$$

En particulier

$$Q_n(1) = (-1)^n C_n^n 2^n \frac{(\Gamma(n+1))^2}{\Gamma(1)\,\Gamma(n+1)} = (-1)^n\, 2^n (n!)$$

donc $\lambda_n = \frac{(-1)^n}{2^n.(n!)}$ ce qui termine la démonstration.

<u>Corollaire</u> : <u>la norme du polynôme de Legendre d'ordre n est égale à</u> $\sqrt{\frac{2}{2n+1}}$

Il nous faut calculer $\int_{-1}^{+1} (P_n(x))^2 dx$. Si a_n désigne le terme de plus haut degré (n) dans l'expression de $P_n(x)$, on obtient :

$$\int_{-1}^{+1} (P_n(x))^2 dx = a_n \int_{-1}^{+1} P_n(x)\, x^n dx$$

En intégrant n fois par parties, on obtient comme précédemment

$$= a_n (-1)^n \left(\int_{-1}^{+1} (1-x^2)^n dx\right) (n\,!) . \frac{(-1)^n}{2^n.n\,!}$$

$$= \frac{a_n}{2^n} \int_{-1}^{+1} (1-x^2)^n dx$$

Un calcul classique fournit $I_n = \int_{-1}^{+1} (1-x^2)^n \, dx$ selon

$$I_n = \int_{-1}^{+1} (1 - x^2)^{n-1} \, dx - \int_{-1}^{+1} x^2(1 - x^2)^{n-1} \, dx$$

Après intégration par parties

$$I_n = I_{n-1} - \left[- \frac{(1 - x^2)^n}{2n} x \right]_{-1}^{+1} - \int_{-1}^{+1} \frac{(1 - x^2)^n}{2n} \, dx$$

Soit

$$I_n = \frac{2n}{2n+1} I_{n-1} = \frac{2n \,.\, (2n-2) \ldots\ 2}{(2n+1)\ (2n-1) \ldots\ 3} \,.\, 2$$

ou encore

$$I_n = 2^{2n+1} \frac{(n\,!)^2}{(2n+1)!}$$

Selon la formule de Rodrigues, on détermine aisément le coefficient a_n du terme de plus haut degré du n-ième polynôme de Legendre P_n :

$$a_n = \frac{2n \,.\, (2n-1) \ldots (n+1)}{2^n \,.\, n\,!}$$

Par suite

$$\int_{-1}^{+1} P_n^2(x) dx = \frac{2n\ (2n-1) \ldots (n+1)}{2^n \,.\, n\,!} \,.\, \frac{1}{2^n} \,.\, 2^{2n+1} \frac{(n\,!)^2}{(2n+1)!}$$

Soit $\int_{-1}^{+1} P_n^2(x) dx = \frac{2}{2\,n+1}$ ce qui démontre le corollaire

. On démontrerait de la même façon que le coefficient de x^{n-1} dans le développement de $P_n(x)$ est nul, ce que l'on pourrait établir à priori par un argument de symétrie.

Ces résultats permettent de préciser les coefficients intervenant dans la formule de récurrence. Les polynômes T_n, obtenus en normant les polynômes de Legendre, sont tels que $T_n(x) = \sqrt{\frac{2n+1}{2}} \, P_n(x)$. Donc le coefficient de x^n dans T_n vaut $\sqrt{\frac{2n+1}{2}} \, \frac{1}{2^n} \,.\, \frac{(2n)!}{(n!)^2}$ et le coefficient de x^{n-1} est nul. Par suite :

$$\left\{ \begin{array}{l} A_n = \sqrt{\frac{2n+1}{2n-1}} \,.\, \frac{2n-1}{n} \\ B_n = 0 \\ C_n = \frac{A_n}{A_{n-1}} \end{array} \right.$$

et ces relations fournissent la relation de récurrence relative aux polynômes de Legendre

$$\sqrt{\frac{2n+1}{2}} \quad P_n(x) = (A_n x + B_n)\; P_{n-1}(x)\sqrt{\frac{2n-1}{2}} - C_n\sqrt{\frac{2n-3}{2}}\; P_{n-2}(x)$$

qui devient

$$P_n(x) = \frac{(2n-1)}{n}\, x\, P_{n-1}(x) - \frac{n-1}{n} P_{n-2}(x)$$

que l'on préfère noter pour tout entier $n \geqslant 2$.

$$\boxed{n P_n(x) = (2n-1)\, x\; P_{n-1}(x) - (n-1)\; P_{n-2}(x)}$$

3.6.5. Equation différentielle.

Chaque polynôme de Legendre est solution d'une équation différentielle linéaire du second ordre d'un type très particulier (Laplace).

Théorème : les polynômes de Legendre satisfont l'équation :

$$(1-x^2)\; \frac{d^2 P_n(x)}{dx^2} - 2x\, \frac{d\, P_n(x)}{dx} + n(n+1)\; P_n(x) = 0$$

L'équation différentielle précédente peut encore s'écrire sous la forme

$$0 = \frac{d}{dx}\left((1-x^2)\, \frac{dP_n(x)}{dx}\right) + n(n+1)\; P_n(x)$$

Puisque P_n est un polynôme de degré n, $\frac{d}{dx}\left((1-x^2)\, \frac{d\,P_n(x)}{dx}\right)$ est un polynôme de degré n. Montrons qu'il est orthogonal à tout polynôme de degré inférieur à n. Soit P un tel polynôme :

$$\begin{aligned}
\int_{-1}^{+1}\frac{d}{dx}\left((1-x^2)\, \frac{d\,P_n(x)}{dx}(x)\right) P(x)\,dx &= \left[P(x)\,(1-x^2)\frac{dP_n(x)}{dx}\right]_{-1}^{+1} - \int_{-1}^{+1}(1-x^2)\,\frac{dP_n}{dx}(x)\,\frac{dP}{dx}(x)\,dx \\
&= -\int_{-1}^{+1}(1-x^2)\, \frac{dP_n(x)}{dx}\, \frac{dP(x)}{dx}\, dx \\
&= -\left[(1-x^2)\,\frac{dP}{dx}(x)\,P_n(x)\right]_{-1}^{+1} + \int_{-1}^{+1} P_n(x)\,\frac{d}{dx}\left((1-x^2)\,\frac{dP}{dx}(x)\right) \\
&= \int_{-1}^{+1} P_n(x)\,\frac{d}{dx}\left((1-x^2)\,\frac{dP(x)}{dx}\right) dx = 0
\end{aligned}$$

Car $(1-x^2)\,\frac{dP(x)}{dx}$ est au plus de degré n, donc sa dérivée est au plus de degré n-1 et P_n est orthogonal à tous les polynômes de degré inférieur à n.

Finalement, $\frac{d}{dx}\left((1-x^2)\frac{dP_n(x)}{dx}\right)$ est égal à P_n à une constante près. Pour déterminer cette constante, le plus simple est de comparer les termes de plus haut degré. A droite, ce terme vaut $-n(n+1)\,a_n$ et à gauche $\lambda_n a_n$ donc $\lambda_n = n(n+1)$ et l'on a bien l'équation différentielle dite.

Il faut remarquer que réciproquement cette équation différentielle où $n(n+1)$ est remplacé par λ n'a de solution polynômiale que pour $\lambda = n(n+1)$ et apparaît comme liée à un problème de valeur propre (l'exercice n° 9 introduit les polynômes de Legendre comme vecteurs propres d'une certaine équation intégrale).

3.6.6. Fonction génératrice.

Théorème : dans un voisinage de $z = 0$, on a la relation

$$\sum_{n=0}^{\infty} P_n(x)\, z^n = \frac{1}{\sqrt{1-2xz+z^2}}$$

où la racine carrée a comme détermination la valeur positive lorsque le radicande est positif

La démonstration est simple à partir du théorème de Cauchy. Si f est une fonction holomorphe à l'intérieur d'un domaine simplement connexe limité par une courbe C orientée positivement, on a pour x dans ce domaine :

$$\frac{d^n f}{dx^n}(x) = \frac{n!}{2\pi i}\int_C \frac{f(t)}{(t-x)^{n+1}}\, dt$$

En particulier

$$\frac{d^n}{dx^n}(1-x^2)^n = \frac{n!}{2\pi i}\int_C \frac{(1-t^2)^n}{(t-x)^{n+1}}\, dt$$

Donc :

$$P_n(x) = \frac{(-1)^n}{2^n}\,\frac{1}{2\pi i}\int_C \frac{(1-t^2)^n}{(t-x)^{n+1}}\, dt$$

où C est une courbe limitant un domaine simplement connexe ne contenant pas +1 et -1 en son intérieur

$$P_n(x) = \frac{1}{2\pi i}\int_C \left(\frac{(t^2-1)}{2(t-x)}\right)^n \frac{dt}{t-x}$$

On effectue alors le changement de variable $\frac{t^2-1}{2(t-x)} = \frac{1}{z}$ où $t = \frac{1}{z}\left(1-\sqrt{1-2xz+z^2}\right)$

Cette transformation applique un voisinage de $t = x$ sur un voisinage de $z = 0$.
On vérifie, en appelant C' la courbe C transformée dans le plan des z et orientée par cette transformation :

$$P_n(x) = \frac{1}{2\pi i}\int_{C'} \frac{dz}{z^{n+1}\sqrt{1-2zx+z^2}}$$

et le théorème de Cauchy donne (formule dite de Laplace)

$$P_n(x) = \frac{1}{n!}\frac{d^n}{dz^n}\left(\frac{1}{\sqrt{1-2zx+z^2}}\right)$$

ce qui correspond bien au développement en série donné dans le théorème.
La fonction suivante : $\dfrac{1}{\sqrt{1-2zx+z^2}}$

s'appelle la fonction génératrice des polynômes de Legendre et permet souvent d'effectuer des raisonnements sur ceux-ci.

Des applications intéressantes des polynômes orthogonaux existent en théorie de l'approximation uniforme et de l'interpolation ou du calcul mécanique des intégrales. (cf Nanta Iremica, Volume N° 11).

3.6.7. Tableau succinct des polynômes orthogonaux.

Nous donnons ci-après un tableau de quelques polynômes orthogonaux parmi les plus fréquemment utilisés. Ces polynômes sont dans un espace $L^2([a,b],h)$. Dans la première colonne, figure l'intervalle $[a,b]$, dans la seconde colonne le poids h, dans la troisième le nom usuel et dans la quatrième colonne l'équation différentielle satisfaite par ces polynômes.

Nous proposons en exercice (N° 36) la détermination explicite de quelques propriétés des polynômes de Čebičev de première espèce. L'exercice N°30 permet de retrouver quelques propriétés des polynômes d'Hermite en partant directement de la fonction génératrice de ces polynômes. On en déduit (exercice N° 31) une propriété d'invariance de la fonction d'Hermite par transformation de Fourier. Les valeurs explicites des coefficients des polynômes orthogonaux classiques peuvent se déduire des formules établies à l'exercice N° 34. On pourra aussi noter les exercices N° 9, 35, 37, 38 et 39 comme utilisations des polynômes orthogonaux.

En physique mathématique (Equations de Schrödinger, théorie des couches électroniques expliquant certaines propriétés chimiques) ces polynômes orthogonaux jouent un très grand rôle, qu'il est hors de propos de détailler ici.

INTERVALLE	POIDS	NOM	EQUATION DIFFERENTIELLE
$[-1,+1]$	1	LEGENDRE $P_n(x)$	$(x^2-1)\dfrac{d^2P_n(x)}{dx^2}+2x\dfrac{dP_n(x)}{dx}-n(n+1)P_n(x)=0$
$[-1,+1]$	$(1-x)^{\alpha}(1+x)^{\beta}$ $\alpha>-1$ $\beta>-1$	JACOBI $P_n^{\alpha,\beta}(x)$	$(1-x^2)\dfrac{d^2P_n^{\alpha,\beta}(x)}{dx^2}+\left[\beta-\alpha-(\alpha+\beta+2)x\right]\dfrac{dP_n^{\alpha,\beta}(x)}{dx}+$ $n(\alpha+\beta+n+1)P_n^{\alpha,\beta}(x)=0.$
$[-1,+1]$	$\dfrac{1}{\sqrt{1-x^2}}$	ČEBIČEV 1er type $P_n(x)$	$(1-x^2)\dfrac{d^2P_n(x)}{dx^2}-x\dfrac{dP_n(x)}{dx}+n^2P_n(x)=0$
$[-\infty,+\infty]$	e^{-x^2}	HERMITE $H_n(x)$	$\dfrac{d^2H_n(x)}{dx^2}-2x\dfrac{dH_n(x)}{dx}+2nH_n(x)=0$
$[0,+\infty]$	e^{-x}	LAGUERRE $L_n(x)$	$x\dfrac{d^2L_n(x)}{dx^2}+(1-x)\dfrac{dL_n(x)}{dx}+nL_n(x)=0$
$[-1,+1]$	$\sqrt{1-x^2}$	ČEBIČEV 2ème type $T_n(x)$	$(1-x^2)\dfrac{d^2T_n}{dx^2}(x)-3x\dfrac{dT_n}{dx}+n(n-2)T_n(x)=0$

3.7. JRS LINEAIRES SUR UN ESPACE DE HILBERT.

ıs allons maintenant nous intéresser à certaines classes d'opérateurs es sur les espaces de Hilbert. Pour ces opérateurs, on dispose de th de représentation qui conduisent à des théorèmes d'approximation en moıadratique. On peut même quelquefois obtenir des renseignements pour dergences plus fortes. Naturellement, le but de cette étude est la résoluquations linéaires dans des espaces fonctionnels (équations difféles ou intégrales, etc.)

3ijoint d'un opérateur.

Chapitre 0, nous avons défini le transposé ${}^tP$ d'un opérateur linéaır la formule

$$(1) \qquad < P(x), y^* > = < x, {}^tP(y^*) > \qquad \text{pour} \begin{cases} x \in \mathcal{H} \\ y^* \in \mathcal{H}^* \end{cases}$$

Dans l'un espace de Hilbert, nous avons vu que $\mathcal{H}$ est isomorphe à $\mathcal{H}'$. donc intéressant de voir ce que donne la notion de transposé d'un or linéaire continu.

Théorèit P un opérateur linéaire continu sur un espace de Hilbert $\mathcal{H}$. existe un opérateur linéaire continu, unique, P^*, satisfaisant ur tous x et y dans $\mathcal{H}$:

$$(2) \qquad \boxed{< P(x), y > \; = \; < x, P^*(y) >}$$

Si l'oifie $\mathcal{H}$ et $\mathcal{H}'$ dans le cas d'un Hilbert réel, on note donc que P^* esstriction à $\mathcal{H}$ de ${}^tP$.

Défini P^* s'appelle l'adjoint de l'opérateur P.

ontrons le théorème ci-dessus :

L'appli $x \longrightarrow < P(x), y >$ de $\mathcal{H}$ dans son corps de base K, est une application re en x, pour tout y fixé. Cette application est même une forme linéairınue, comme il résulte des majorations suivantes provenant de l'inéga Schwarz et de la définition de $\| P \|$

$$(3) \qquad \begin{aligned} |< P(x), y >| &\leqslant \| P(x) \| \, \| y \| \\ |< P(x), y >| &\leqslant \| P \| \, \| x \| \, \| y \| \end{aligned}$$

Le théorème de représentation implique qu'il existe un unique vecteur y de $\mathcal{H}$, noté $P^*(y)$, tel que

$$< P(x), y > \; = \; < x, P^*(y) >$$

Par ailleurs

$$< Px, \lambda y_1 + \mu y_2 > \; = \; \overline{\lambda} < Px, y_1 > + \; \overline{\mu} < Px, y_2 > \; =$$
$$< x, P^*(\lambda y_1 + \mu y_2) > \; = \; < x, \lambda P^*(y_1) + \mu P^*(y_2) >$$

ce qui entraîne la linéarité de P^*.

Enfin $$\| P^*(y) \| \leqslant \| P \| \| y \|$$

d'après l'inégalité (3), ce qui montre la continuité de l'opérateur P et $\|P^*\| \leqslant \|P\|$.

Pour démontrer l'unicité de P, il suffit de remarquer que si P'_1 et P'_2 satisfaisaient tous deux l'égalité (2), on aurait

$$< x, P'_1(y) - P'_2(y) > \; = \; 0$$

pour tout x et tout y dans $\mathcal{H}$. Cela entraîne

$$P'_1(y) = P'_2(y) \qquad \text{donc} \qquad P'_1 = P'_2$$

Les calculs précédents permettent de fournir la valeur de $\| P^* \|$. En effet, grâce à l'unicité, on a :

$$(P^*)^* = P$$

Donc, comme $\| P^* \| \leqslant \| P \|$, il vient $\| (P^*)^* \| \leqslant \| P^* \|$, ce qui donne

$$(4) \qquad \| P^* \| = \| P \|$$

3.7.2. Propriétés algébriques de l'adjoint.

$$(P^*)^* = P$$
$$(\lambda P)^* = \overline{\lambda} P^*$$
$$(P_1 P_2)^* = P_2^* P_1^*$$

Définition : un opérateur linéaire continu tel que $P = P^*$ est dit hermitien (ou auto-adjoint ou self-adjoint).

Si P est un opérateur linéaire continu :

$$P_1 = \frac{P + P^*}{2} \qquad \text{et} \qquad P_2 = \frac{i}{2}(P - P^*)$$

sont deux opérateurs hermitiens satisfaisant

$$P = P_1 - iP_2 \quad \text{et} \quad P^* = P_1 + i\,P_2$$

De plus P^*P et PP^* sont deux opérateurs continus hermitiens pour lesquels

$$\|PP^*\| = \|P^*P\| = \|P\|^2 = \|P^*\|^2$$

En effet, d'une part :

$$\|P^*P\| \leqslant \|P^*\|\,\|P\| = \|P\|^2 \qquad (\|P_1 P_2\| \leqslant \|P_1\|\,\|P_2\|)$$

d'autre part :

$$\|P(x)\|^2 = \langle P(x), P(x)\rangle = \langle x, P^*P(x)\rangle \leqslant \|x\|^2 \|P^*P\|$$

ce qui donne :

$$\|P\|^2 \leqslant \|P^*P\|$$

D'où l'égalité annoncée.

En particulier, pour un opérateur hermitien P, on a le théorème suivant :

<u>Théorème</u> : <u>si P est un opérateur hermitien ($P = P^*$), on a l'égalité</u>

$$\boxed{\|P^2\| = \|P\|^2}$$

<u>Remarque générale</u> : exprimée dans le cadre d'un espace euclidien (dimension finie) et par rapport à une base orthonormée, la représentation d'un opérateur hermitien est une matrice hermitienne ($a_{ij} = \overline{a_{ji}}$, ce qui équivaut à $[A] = [A^*]$). Même dans ce cas, l'avantage de la définition donnée d'un opérateur hermitien est de ne pas faire intervenir de base particulière.

Au Chapitre 0, paragraphe 3.7, nous avons remarqué que la propriété d'hermiticité pour une matrice n'est pas invariante par changement de base. Par contre, si l'on se restreint aux changements de base qui conservent la structure hilbertienne de E, c'est-à-dire aux changements de base U tels que $\langle Ux, Uy\rangle = \langle x, y\rangle$ pour tout x et tout y dans E (U est un <u>opérateur unitaire</u> dans l'espace E de dimension finie), alors la propriété d'hermiticité est conservée, comme on peut le constater directement.

3.7.3. <u>Un exemple d'opérateur intégral</u>.

Prenons l'espace préhilbertien $\mathcal{C}[0,1]$ dont le complété est l'espace $L^2[0,1]$.

Soit par ailleurs $K(x,y)$ une fonction continue de deux variables x et y, définie sur le carré $[0,1]\times[0,1]$ et à valeurs réelles. A toute fonction continue sur $[0,1]$, on peut faire correspondre la fonction Tf :

$$Tf : \qquad Tf(y) = \int_0^1 K(x,y)\, f(x)\, dx$$

La continuité uniforme de $K(x,y)$ sur $[0,1]^2$ permet de montrer que Tf est un élément de $\mathcal{C}\,[0,1]$ selon

$$|Tf(y_1) - Tf(y_2)| \leqslant \int_0^1 |K(x,y_1) - K(x,y_2)|\, |f(x)|\, dx$$

et il existe un $\eta > 0$ tel que $|y_1 - y_2| < \eta$ entraîne

$$|K(x,y_1) - K(x,y_2)| < \varepsilon$$

pour tout x. Donc, en utilisant l'inégalité de Schwarz

$$|Tf(y_1) - Tf(y_2)| \leqslant \varepsilon \int_0^1 |f(x)|\, dx \leqslant \varepsilon \|f\|_2$$

Il est clair que T est un opérateur linéaire sur $\mathcal{C}\,[0,1]$. Que peut-on dire de sa continuité relativement à la norme choisie sur $\mathcal{C}\,[0,1]$.

Grâce à l'inégalité de Schwarz, on peut écrire les majorations ;

$$\|Tf\|_2^2 = \int_0^1 dy \left| \int_0^1 K(x,y)\, f(x)\, dx \right|^2$$

$$\leqslant \int_0^1 dy \left(\int_0^1 |f(x)|^2 dx\right)\left(\int_0^1 |K(x,y)|^2 dx\right)$$

$$\leqslant \left(\int_0^1 |f(x)|^2 dx\right) \int_0^1 dy \int_0^1 |K(x,y)|^2 dx$$

$$\leqslant \|f\|_2^2 \iint_{[0,1]^2} |K(x,y)|^2\, dx\, dy$$

On obtient donc :

$$\|T\| \leqslant \left(\iint_{[0,1]^2} |K(x,y)|^2\, dx\, dy\right)^{1/2}$$

ce qui démontre la continuité de l'opérateur T. Une estimation précise de $\|T\|$ s'avère nettement plus délicate d'ailleurs.

Définition : on dit que l'opérateur T :

$$f \longrightarrow \int_0^1 f(x)\, K(x,y)\, dx \;=\; Tf(y)$$

est un opérateur intégral de noyau $K(x,y)$

Prolongement de l'opérateur T : l'opérateur T a été défini sur l'espace $\mathcal{C}[0,1]$ mais on peut définir un prolongement de T en un opérateur linéaire continu sur tout $L^2[0,1]$. Il suffit d'appliquer le théorème de prolongement démontré en 1.6.9. Il existe un opérateur linéaire $\overline{T}$, prolongeant T à l'espace $L^2[0,1]$ tel que : $\|\overline{T}(f)\|_2 \leqslant \|T\|\,\|f\|_2$. Cet opérateur $\overline{T}$ est unique :

$$\left(\overline{T}(f)\right) = \lim_{n\to\infty}\int_0^1 f_n(x) K(x,y)\, dx$$

où f_n est une suite de fonctions continues convergeant en moyenne quadratique vers f (Cf. 1.6.9).

On confondra $\overline{T}$ et T dorénavant en notant

$$(1) \qquad T(f) = \int_0^1 K(x,y)\, f(x)\, dx$$

pour toute fonction f de $L^2[0,1]$ puisque l'intégrale dans cette égalité est à prendre maintenant au sens de Lebesgue, et que nous n'avons fait que passer au complété de $\mathcal{C}[0,1]$, à savoir l'espace $L^2[0,1]$

Cherchons à déterminer l'adjoint de l'opérateur linéaire T, défini maintenant sur $L^2[0,1]$. On doit avoir

$$<Tf, g> \;=\; <f, T^*g>$$

pour tous f et g dans $L^2[0,1]$, que nous supposerons défini sur $\mathbb{C}$ pour fixer les idé-

$$\int_0^1 Tf(y)\overline{g(y)}\, dy = \int_0^1 f(x)\overline{T^*(g)(x)}\, dx$$

ou encore

$$\int_0^1 \overline{g(y)} dy \left(\int_0^1 k(x,y) f(x) dx\right) dy = \int_0^1 f(x)\, \overline{T^*(g)(x)}\, dx$$

Mais on peut intervertir les signes $\int$ dans le premier membre (théorème de Fubini, Chap. II, § 6.3).

Soit :

$$\int_0^1 f(x) \left(\int_0^1 k(x,y)\, \overline{g(y)}\, dy\right) dx = \int_0^1 f(x)\, \overline{T^*(g)(x)}\, dx$$

Il vient donc :

$$T^*(g)(x) = \int_0^1 \overline{K(x,y)}\, g(y)\, dy$$

L'opérateur adjoint de l'opérateur intégral T, de noyau $K(x,y)$ est un opérateur intégral, de noyau $\overline{K(y,x)}$

$$\boxed{\begin{aligned} Tf(y) &= \int_0^1 K(x,y)\, f(x)\, dx \\ T^*g(x) &= \int_0^1 \overline{K(x,y)}\, g(y)\, dy \end{aligned}}$$

En particulier, lorsque $\overline{K(x,y)} = K(y,x)$, l'opérateur intégral et continu T, est un opérateur hermitien, On pourra vérifier que l'on pourrait se contenter de supposer $K(x,y)$ dans $L^2[0,1]$ au lieu de le supposer continu.

3.8. OPERATEURS COMPACTS DANS UN ESPACE DE HILBERT.

Pour obtenir de bonnes propriétés de décomposition d'un opérateur on doit faire certaines hypothèses restrictives. La définition suivante s'avère satisfaisante, mais il faudra préciser la signification de ce satisfecit

3.8.1 Définition.

Un opérateur linéaire T défini sur un espace de Hilbert $\mathcal{H}$ est dit compact si de toute suite $\{x_n\}$ de vecteurs tels que $\|x_n\| \leqslant 1$, on peut extraire une sous-suite $\{x_{n_k}\}$ telle que :

$$\lim_{k \to \infty} T(x_{n_k}) \qquad \text{existe dans } \mathcal{H}$$

Un opérateur compact est nécessairement continu. En effet, il suffit de montrer que l'image de la boule unité par T est incluse dans une certaine boule. Or, s'il n'en était pas ainsi, on pourrait trouver une suite $\{x_n\}_{n \geqslant 1}$ de vecteurs tels que $\|x_n\| \leqslant 1$ et pour lesquels $\|T(x_n)\|$ tendrait vers l'infini. D'une telle suite x_n, on ne pourrait extraire une sous-suite x_{n_k} pour laquelle $T(x_{n_k})$ converge, ce qui contredit la compacité de l'opérateur T.

Remarque : L'application identique ne constitue pas un opérateur compact sur un espace de Hilbert de dimension infinie.

En effet, si $\mathcal{H}$ est de dimension infinie, le procédé d'orthonormalisation permet de construire une suite $\{e_n\}_{n\geq 1}$ orthonormale. De cette suite bornée on ne peut extraire une sous-suite convergente puisque la norme de la différence de deux éléments distincts quelconques de cette suite vaut $\sqrt{2}$.

Remarque : Plus généralement, F. RIESZ a établi un résultat valable dans un espace de Banach quelconque. La boule unité d'un tel espace est compacte si et seulement si l'espace est de dimension finie.

Donnons maintenant l'exemple le plus simple d'un opérateur compact

Théorème : Un opérateur linéaire continu de rang fini est un opérateur compact

Rappelons qu'un opérateur est de rang fini si son image est de dimension finie. Soit alors P un tel opérateur continu et de rang fini. Par continuité, l'image de la boule unité est un ensemble borné. Puisque P est de rang fini, cet ensemble borné est inclus dans une boule d'un sous-espace vectoriel de dimension n. Grâce au théorème de Bolzano-Weierstrass ceci implique la compacité de l'opérateur P.

La notion d'opérateur compact se conserve par passage à la limite au sens de la topologie de la norme sur l'espace $\mathcal{L}(\mathcal{H})$ des opérateurs linéaires continus sur $\mathcal{H}$ comme nous allons l'établir. Ce point important permet de construire des opérateurs compacts à partir des opérateurs de rang fini.

3.8.2 Soit T_n une suite d'opérateurs compacts convergeant vers un opérateur (continu) T ($\lim_{n\to\infty} \|T_n - T\| = 0$), alors T est un opérateur compact.

Le procédé de démonstration est une fois encore le procédé diagonal déjà utilisé pour la démonstration de la première étape du théorème du choix.

Soit x_n une suite de vecteurs de $\mathcal{H}$ tels que $\|x_n\| \leq 1$. On peut extraire de cette suite, une suite $x_n^{(1)}$ telle que $\lim_{n\to\infty} T_1(x_n^{(1)})$ existe, puisque T_1 est un opérateur compact. De même, on peut extraire de $x_n^{(1)}$ une suite $x_n^{(2)}$ telle que $\lim_{n\to\infty} T_2(x_n^{(2)})$ existe... et ainsi de suite. Envisageons alors la suite $x_n^{(n)}$ que l'on note y_n pour simplifier. Pour chaque valeur fixée de n, $\lim_{p\to+\infty} T_n(y_p)$ existe. Etudions la suite $T(y_n)$ où $T = \lim_{n\to+\infty} T_n$ Le théorème est acquis si l'on prouve la convergence de $T(y_n)$ puisque $\{y_n\}_{n\geq 1}$ est une suite extraite de $\{x_n\}_{n\geq 1}$.

Comme $\| y_p \| \leqslant 1$ et $\| y_q \| \leqslant 1$, on a la majoration :

$$\| T(y_q) - T(y_p) \| \leqslant 2 \| T - T_n \| + \| T_n(y_q - y_p) \|$$

ε étant donné (positif), on choisit n tel que $\| T - T_n \| \leqslant \varepsilon/4$, puis n étant fixé, on choisit N tel que $\| T_n(y_q - y_p) \| \leqslant \varepsilon/2$ pour p et $q \geqslant N$ (ce qui est possible puisque $\lim\limits_{p \to \infty} T_n(y_p)$ existe).

Par conséquent, pour p et q supérieurs ou égaux à N, $\| T(y_p) - T(y_q) \| \leqslant \varepsilon$
La suite $T(y_p)$ est de Cauchy et $\lim\limits_{p \to \infty} T(y_p)$ existe puisque $\mathcal{H}$ est un espace de Hilbert.

<u>Corollaire</u> : <u>L'opérateur intégral T : $Tf(y) = \int_0^1 k(x,y) f(x)\, dx$ où $k(x,y)$ est un noyau continu ou appartenant à $L^2[0,1]^2$, est un opérateur compact dans $L^2[0,1]$</u>.

Il suffit de montrer que l'opérateur T est une limite (uniforme) d'opérateurs compacts. On utilise les trois faits suivants :

Si $k(x,y)$ est une fonction en escalier sur $[0,1]^2$, l'opérateur T est un opérateur de rang fini, donc compact (cf chap II § 6.3).

Toute fonction $k(x,y)$ (dans $L^2[0,1]^2$) est limite, en moyenne quadratique, d'une suite de fonctions en escalier. Ce résultat provient de la construction même de $L^2[0,1]$ (Cf. Chapitre II).

Si $k_n(x,y)$ converge en moyenne quadratique vers $k(x,y)$, alors T_n, l'opérateur intégral associé à k_n, converge uniformément vers T, opérateur intégral associé à k (appliquer l'inégalité de Schwarz).

Disons tout de suite que ce procédé d'approximation, par l'intermédiaire de la compacité, est à la base des résultats que nous allons obtenir. C'est d'ailleurs une constante de ce cours que de montrer comment des résultats difficiles sur les espaces fonctionnels s'obtiennent à partir de certains résultats sur des sous-espaces denses. Bien entendu, on peut faire la théorie abstraite des opérateurs compacts dans un espace de Banach . F. RIESZ a établi magistralement la démarche essentielle de la théorie. Cependant, nous allons nous contenter de fournir le cas assez particulier d'un opérateur intégral hermitien et son application aux équations intégrales.

3.8.3. Equations intégrales

On appelle équation intégrale linéaire, ou équation de Fredholm, une équation de la forme

$$(1) \qquad \boxed{f(y) - \lambda \int_0^1 k(x,y)\, f(x)\, dx = g(y)}$$

où g est une fonction donnée sur $[0,1]$ tandis que $k(x,y)$ est continue sur $[0,1]^2$ et λ un nombre complexe donné. On peut même supposer que le noyau $k \in L^2[0,1]^2$

Résoudre cette équation, c'est chercher une fonction f satisfaisant (1). Un tel problème est mal posé : il y a quelques conditions de compatibilité à respecter.

D'une part, on ne peut chercher que des solutions pour lesquelles l'intégrale $\int_0^1 k(x,y)\, f(x)\, dx$ ait un sens.

Par ailleurs, si g est donnée, par exemple continue, il y a de fortes chances que cette propriété ait des répercussions sur la fonction f.

Avec ce qui précède, il paraît naturel de chercher à résoudre cette équation intégrale dans certains espaces fonctionnels. Par exemple :

- Résoudre (1) dans $\mathcal{C}[0,1]$ en supposant que $g \in \mathcal{C}[0,1]$, c'est chercher une fonction continue satisfaisant (1) lorsque g est une fonction continue.
- Résoudre (1) dans $L^2[0,1]$, c'est chercher une fonction de carré intégrable satisfaisant (1) lorsque g est de carré intégrable. Dans ce dernier cas, en appelant T l'opérateur intégral dont le noyau est k, on voit qu'il s'agit de résoudre, dans $L^2[0,1]$ l'équation :

$$(2) \qquad (I - \lambda T) f = g$$

Il s'agit donc d'inverser l'opérateur $I - \lambda T$ dans l'espace des opérateurs continus sur $L^2[0,1]$.

On pourra lire les paragraphes suivants en supposant que g appartient à $\mathcal{C}[0,1]$ et en cherchant à inverser $(-\lambda T + I)$ dans l'espace $\mathrm{Hom}(\mathcal{C}[0,1])$ On vérifiera que les raisonnements n'utilisent pas la propriété de complétion de $L^2[0,1]$ (lorsque $k(x,y)$ est continue) mais seulement la compacité de T (la propriété de compacité se définit sans problème sur un espace normé). Pratiquement, cela signifie que nous pouvons éviter l'intégrale de Lebesgue dans nos calculs, tous les raisonnements se poursuivant par complétion à l'espace $L^2[0,1]$.

Le problème étant visiblement linéaire, nous allons d'abord étudier l'équation homogène associée à (1), c'est-à-dire chercher une fonction f de $L^2[0,1]$ telle que :

(3) $$f(y) = \lambda \int_0^1 k(x,y)\, f(x)\, dx$$

En langage d'opérateurs, il s'agit de trouver f dans $L^2[0,1]$ telle que :

(4) $$\lambda T f = f$$

Convenons des définitions suivantes, déjà connues dans le cas des matrices.

Si $f \neq 0$ et si $Tf = \lambda' f$, on dit que f est un vecteur propre de l'opérateur T pour la valeur propre λ'.

(il est fort possible que $\lambda' = 0$, et dans ce cas $Tf = 0$, par conséquent l'opérateur T n'est pas injectif).

A partir de maintenant, nous allons supposer que T est un opérateur hermitien, c'est-à-dire que $k(x,y) = \overline{k(y,x)}$ conformément au résultat du paragraphe 3.7.3.

L'hypothèse d'hermiticité va nous fournir un certain nombre de propriétés algébriques qui, jointes à la compacité de l'opérateur laquelle s'exprime ici par une propriété d'approximation uniforme par des opérateurs de rang fini, permettra de donner une réponse satisfaisante au problème de la résolution de l'équation intégrale (1). Nous suivons la démonstration élégante de F. RIESZ et B. Sz.NAGY.

3.8.4. Popriétés algébriques des opérateurs hermitiens.

Donnons trois théorèmes cruciaux, dont deux sont bien connus.

3.8.4.1. Théorème : Les valeurs propres d'un opérateur hermitien sont réelles.

Si f est un vecteur propre pour la valeur propre λ de T, on peut écrire :

$$< Tf, f > \; = \; < f, Tf > \quad \text{grâce à l'hermiticité}$$

et donc $\quad \lambda < f, f > \; = \; \bar{\lambda} < f, f > \quad$ puisque $T(f) = \lambda f$

D'où $\lambda = \bar{\lambda}$ car $< f, f > = \|f\|^2$ est différent de 0 par hypothèse (vecteur propre).

3.8.4.2. <u>Théorème</u> : <u>Deux vecteurs propres correspondant à deux valeurs propres distinctes sont orthogonaux.</u>

Soient f et f' deux vecteurs propres qui correspondent aux valeurs propres λ et λ'. On écrit

$$< Tf, f' > \; = \; \lambda < f, f' >$$

$$< Tf, f' > \; = \; < f, Tf' > \; = \; \lambda' < f, f' >$$

D'où :

$$(\lambda - \lambda') < f, f' > \; = 0$$

ce qui entraîne $< f, f' > = 0$ puisque $\lambda \neq \lambda'$

3.8.4.3. <u>Théorème</u> : <u>Si T est un opérateur hermitien,</u> $< Tf, f >$ <u>est un nombre réel et l'on a l'égalité</u>

$$\|T\| = \sup_{\|f\|=1} \|T(f)\| = \sup_{\|f\|=1} |< Tf, f >|$$

Pour tout f, le nombre $< Tf, f >$ est réel puisque $< Tf, f > = < f, Tf > = \overline{< Tf, f >}$

Par ailleurs, posons $\quad t = \sup_{\|f\|=1} |< Tf, f >|$, on a l'inégalité

$$t \leqslant \|Tf\| \|f\| \leqslant \|T\| \|f\|^2 = \|T\|$$

Démontrons l'inégalité inverse : soit α un nombre positif quelconque; remarquons l'égalité combinatoire suivante :

$$\|Tf\|^2 = < Tf, Tf >$$

$$= \frac{1}{4} \left[< T\left(\alpha f + \frac{1}{\alpha} Tf\right), \alpha f + \frac{1}{\alpha} Tf > - < T\left(\alpha f - \frac{1}{\alpha} T(f)\right), \alpha f - \frac{1}{\alpha} T(f) > \right]$$

que l'on obtient grâce à la relation d'hermiticité

$$< T^2 f, f > \quad = \quad < Tf, Tf >$$

On a alors les majorations

$$\| T(f) \|^2 \leqslant \frac{1}{4} \left(t \| \alpha f + \frac{T(f)}{\alpha} \|^2 + t \| \alpha f - \frac{T(f)}{\alpha} \|^2 \right)$$

Mais l'égalité du parallélogramme permet d'écrire :

$$\| T(f) \|^2 \leqslant \frac{t}{2} \left(\alpha^2 \| f \|^2 + \frac{1}{\alpha^2} \| T(f) \|^2 \right)$$

En calculant la valeur de α assurant le minimum de l'expression du second membre, il vient

$$\| T(f) \|^2 < t \| T(f) \| \| f \| \quad \text{d'où} \quad \| T(f) \| \leqslant t \| f \|$$

et donc $\| T \| \leqslant t$, ce qui démontre le théorème. Ce dernier va nous assurer l'existence d'une valeur propre non nulle pour un opérateur hermitien compact.

Puisque $\| T \| = \sup\limits_{\| f \| = 1} |<Tf, f>|$, il existe une suite f_n d'éléments de $L^2[0,1]$ telle que

$$\| f_n \| = 1 \quad \text{et} \quad \lim_{n \to +\infty} |<Tf_n, f_n>| = \| T \|$$

Il est toujours possible de choisir la suite f_n de sorte que $<Tf_n, f_n>$ converge également vers un nombre réel noté $\lambda \neq 0$. On a bien sûr $\lambda = \pm \| T \|$.

Calculons $\| Tf_n - \lambda f_n \|^2$ en développant ce carré :

$$\| Tf_n - \lambda f_n \|^2 = \| T(f_n) \|^2 + \lambda^2 \| f_n \|^2 - 2\lambda <Tf_n, f_n>$$

Or lorsque n tend vers l'infini, $<T(f_n), f_n>$ converge vers λ et $\| T(f_n) \|^2$ reste inférieur à $\| T \|^2 = \lambda^2$.

Ceci prouve que $$\lim_{n \to +\infty} \| Tf_n - \lambda f_n \|^2 = 0$$

Les f_n constituent donc une <u>solution approchée</u> de l'équation propre

$$Tf = \lambda f$$

Utilisons maintenant le fait que T est compact. De la suite $\{f_n\}$, ($\| f_n \| = 1$), on peut extraire une sous-suite $\{f_{n_k}\}$ telle que $\lim\limits_{k \to \infty} T(f_{n_k})$ existe.

Mais la suite f_{n_k} elle-même est convergente dans $L^2[0,1]$ puisque

$$\lim_{n \to +\infty} \left(T(f_n) - \lambda(f_n)\right) = 0$$

Par conséquent, si l'on pose $f = \lim_{k \to +\infty} f_{n_k}$, il vient

$$\boxed{T(f) = \lambda f} \qquad \text{avec} \qquad |\lambda| = \|T\| = \|T(f)\|$$

Ceci permet d'énoncer le théorème suivant.

3.8.4.4. Théorème : Un opérateur hermitien compact non nul admet au moins une valeur propre non nulle.

Nous constatons d'ailleurs que nous avons trouvé une valeur propre λ ayant le plus grand module possible en cherchant une fonction f réalisant l'extrémum $\sup\limits_{\|f\|=1} |<Tf, f>|$. Le procédé utilisé est itératif et à partir des techniques hilbertiennes, nous allons décomposer l'opérateur T.

3.8.5. Théorème de décomposition d'un opérateur hermitien compact.

Soit donc T un opérateur hermitien compact sur un espace de Hilbert .
Le lemme suivant organise la démonstration par récurrence.

Soit φ une fonction propre relative à une valeur propre non nulle λ de l'opérateur T. Le sous-espace $\mathcal{H}_\varphi$ orthogonal à φ est stable par T. Ceci signifie que $T(\mathcal{H}_\varphi) \subset \mathcal{H}_\varphi$ où $\mathcal{H}_\varphi$ désigne le sous-espace orthogonal à Q.

$$\mathcal{H}_\varphi = \{f \mid f \in \mathcal{H} \ ; < f, \varphi > = 0\}$$

le lemme est facile à établir puisque si f $\mathcal{H}_\varphi$ on a

$$\begin{aligned} < T(f), \varphi > &= < f, T(\varphi) > = \lambda < f, \varphi > \\ &= 0 \end{aligned}$$

Par suite, $\mathcal{H}_\varphi$ est un sous-espace fermé de $\mathcal{H}$, donc un espace de Hilbert, sur lequel opère T. On recommence la recherche d'une valeur propre pour T restreint à $\mathcal{H}_\varphi$. On a ainsi gagné une dimension.

Précisons les notations. Par le procédé d'extrémalité, on dispose d'une fonction propre φ_1 relative à une valeur propre λ_1 où $|\lambda_1| = |T|$. Notons T_{φ_1} la restriction de T à $\mathcal{H}_{\varphi_1}$. Soit φ_2 une fonction propre (appartenant à $\mathcal{H}\varphi_1$) de T_{φ_1} et relative à une valeur propre λ_2 où $|\lambda_2| = \|T_{\varphi_1}\|$. Naturellement

$$|\lambda_2| \leqslant |\lambda_1|$$

Le procédé itératif permet de choisir un ensemble <u>totalement ordonné</u> d'indices I et deux familles $\{\varphi_i\}_{i \in I}$; $\{\lambda_i\}_{i \in I}$ où $\varphi_i \in \mathcal{H}$ et $\lambda_i \in \mathbb{R}$ de sorte que

$$\|\varphi_i\| = 1$$

$$< \varphi_i, \varphi_j > = 0 \qquad i \neq j$$

$$T(\varphi_i) = \lambda_i \varphi_i$$

$$\text{et } |< T(\varphi_i), \varphi_i >| = |\lambda_i| = \operatorname*{Sup}_{C} |< T\varphi, \varphi>|$$

Ici $\operatorname*{Sup}_{C}$ désigne la borne supérieure conditionnée par le fait que φ parcourt $\mathcal{H}$ satisfait $\|\varphi\| = 1$ et $< \varphi, \varphi_j > = 0$ pour tout j de I tel que $j < i$. On déduit que pour $i > j$, on a $|\lambda_i| \leqslant |\lambda_j|$

Si T est un opérateur sur un espace de Hilbert séparable, par exemple un opérateu intégral sur $L^2[0, 1]$, alors l'ensemble I est au plus dénombrable puisque la famille $\{\varphi_i\}_{i \in I}$ est orthonormale. Ce résultat est général (mais peut être laissé de côté en première lecture).

[Appelons en effet $J_a = \{i \mid i \in I ; |\lambda_i| \geqslant a\}$ où $a > 0$. Montrons que J_a est fini. S'il existait une infinité d'indices dans J_a, on saurait trouver, par compacité de T, une suite d'éléments distincts $\{\varphi_{i_n}\}_{n \geqslant 1}$ où $i_n \in J_a$ pour tout n, de sorte que $T(\varphi_{i_n})$ fût convergente dans $\mathcal{H}$ et donc constituerait une suite de Cauchy. Ceci serait contredit puisque pour $n \neq m$.

$$\begin{aligned}\|T(\varphi_{i_n}) - T(\varphi_{i_m})\|^2 &= \|\lambda_{i_n} \varphi_{i_n} - \lambda_{i_m} \varphi_{i_m}\|^2 \\ &= |\lambda_{i_n}|^2 + |\lambda_{i_m}|^2 \\ &\geqslant 2a^2\end{aligned}$$

Puisque J_a est fini, et comme $I = \bigcup_{n>1} J_{1/n}$, on constate que I est au plus dénombrable.]

Deux cas se présentent alors

ou bien I est fini , disons possède n éléments et l'on note $\varphi_1, \varphi_2, \ldots, \varphi_n$ $\lambda_1, \lambda_2, \ldots \lambda_n$. L'opérateur T est nul sur l'orthogonal $\mathcal{H}_n^{\perp}$ du sous-espace vectori $\mathcal{H}_n$ engendré par les n vecteurs propres $\varphi_1, \varphi_2, \ldots, \varphi_n$. Grâce au théorème de projection sur le sous-espace de dimension finie $\mathcal{H}_n$, on a pour tout f de $\mathcal{H}$ l'écriture unique

$$f = \sum_{k=1}^{n} < f, \varphi_k > \varphi_k + g \quad \text{où} \quad g \in \mathcal{H}_n^{\perp}$$

donc

$$\boxed{Tf = \sum_{k=1}^{n} \lambda_k < f, \varphi_k > \varphi_k}$$

On laisse au lecteur le soin d'établir que T n'a alors d'autres valeurs propres que les λ_k et éventuellement 0.

ou bien I est infini, et l'on note $\{\varphi_n\}_{n \geq 1}$, $\{\lambda_n\}_{n \geq 1}$. Montrons alors que $\lim_{n\to\infty} \lambda_n = 0$. Ce résultat est évident si l'on a suivi la démonstration du fait que J_a est un sous-ensemble fini. Mais on peut donner une précision lorsque T est un opérateur intégral, disons sur $L^2[0, 1]$ et de noyau $k(x,y)$. Dans ce cas, en effet $\sum_{n=1}^{\infty} |\lambda_n|^2$ est une série convergente (et donc $\lim_{n\to\infty} \lambda_n = 0$).

. On constate que $\lambda_k \varphi_k(y)$, pour y fixé, est le coefficient de Fourier de la fonction $k(x, y)$ (y fixé) par rapport à la fonction φ_k, élément du système orthogonal $\{\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots\}$. Par conséquent, on peut appliquer l'inégalité de Bessel pour toute valeur de l'entier n :

$$\sum_{k=1}^{k=n} |\lambda_k|^2 |\varphi_k^2(y)| \leqslant \int_0^1 |k(x,y)|^2 \, dx$$

En intégrant entre 0 et 1 cette inégalité, on obtient :

$$\sum_{k=1}^{n} |\lambda_k|^2 \leqslant \int_0^1\int_0^1 |k(x,y)|^2 dx \, dy$$

De ce résultat, on déduit deux choses :

D'une part que $\sum_{i=1}^{\infty} |\lambda_i|^2$ converge, par conséquent que $\lim_{i \to \infty} |\lambda_i| = 0$.

Par ailleurs, si $\varphi_1, \varphi_2, \dots, \varphi_n$ sont des vecteurs propres associés à une même valeur propre λ, on a

$$n |\lambda|^2 \leqslant \int_0^1 \int_0^1 |k(x,y)|^2 dx\, dy$$

Ce qui prouve que le nombre des vecteurs propres orthogonaux normés associés à la valeur λ est fini. Ce nombre est d'ailleurs appelé l'ordre de λ et donne la dimension finie du sous-espace propre associé à λ.

On pourrait facilement aménager la démonstration pour en déduire d'abord que la famille calculée par le procédé d'extrémalisation des valeurs propres non nulles de l'opérateur intégral de noyau $k(x, y)$ est au plus dénombrable et ce, sans utiliser le recours à l'ensemble J_a. Nous laissons ce soin au lecteur vigilant.

Revenons maintenant au cas plus général d'un opérateur hermitien compact T sur un espace de Hilbert, lorsque I est un ensemble infini. Nous devons calculer la différence entre la fonction f et son développement suivant les k premiers vecteurs de la famille orthonormale $\{\varphi_k\}_{k \geqslant 1}$

Pour ce faire, on calcule
$$\psi_n = f - \sum_{k=1}^{k=n} < f, \varphi_k > \varphi_k$$

Puisque ψ_n appartient à l'espace de Hilbert des vecteurs orthogonaux à $\varphi_1, \varphi_2, \dots, \varphi_n$, et grâce à la propriété d'extrémalité de $|\lambda_{n+1}|$ et à la propriété de la norme d'un opérateur, on dispose de l'inégalité :

$$\|T\psi_n\| \leqslant |\lambda_{n+1}| \; \|\psi_n\|$$

On sait que $\lim_{n \to \infty} \lambda_{n+1} = 0$. Il reste à majorer l'expression $\|\psi_n\|$.

De plus

$$\| \psi_n \|^2 = \| f \|^2 - \sum_{i=1}^{i=n} |<f, \varphi_i>|^2 \leqslant \| f \|^2$$

(inégalité de Bessel).

On a alors :

$$\| T(\psi_n) \| \leqslant |\lambda_{n+1}| \, \| f \|^2$$

ce qui montre que

$$\lim_{n \to \infty} T(\psi_n) = 0 \text{ puisque } \lim_{n \to \infty} \lambda_n = 0$$

Nous écrivons ce résultat sous la forme

$$\boxed{T(f) = \lim_{n \to \infty} \sum_{i=1}^{i=+n} \lambda_i < f, \varphi_i > \varphi_i}$$

Il n'est pas difficile d'établir, grâce à l'égalité ci-dessus, que la suite $\{\lambda_n\}_{n \geq 1}$ épuise la famille des valeurs propres non nulles de l'opérateur hermitien compact T.

On peut encore écrire

$$T(f) = \lim_{n \to +\infty} \sum_{i=1}^{i=n} < f, T(\varphi_i) > \varphi_i = \lim_{n \to +\infty} \sum_{i=1}^{i=n} < Tf, \varphi_i >$$

Nous avons donc établi le théorème suivant, qui résume les diverses propriétés de T et que nous énonçons dans le cadre d'un opérateur intégral, mais le théorème reste exact pour un opérateur hermitien compact quelconque.

3.8.5.1. Théorème : Soit T un opérateur intégral de noyau $k(x, y)$ tel que $k(x, y) = \overline{k(y, x)}$.

T admet au plus une infinité dénombrable de valeurs propres. Chaque valeur propre non nulle est d'ordre fini. Si la suite des valeurs propres est infinie, elle tend vers 0. De plus, tout élément $T(f)$ peut être développé suivant le système orthonormal des φ_i, vecteurs propres associés aux valeurs propres non nulles de T.

$$\boxed{T(f) = \sum_{i=1}^{\infty} < Tf, \varphi_i > \varphi_i = \sum_{i=1}^{\infty} \lambda_i < f, \varphi_i > \varphi_i}$$

3.8.5.2. Totalité du système des vecteurs propres.

Une question se pose aussitôt : puisque $\{\varphi_n\}_{n \geqslant 1}$ est un système orthonormal, qui permet de représenter tout élément de la forme $T(f)$, ce système est-il total dans l'espace $L^2[0,1] = \mathcal{H}$?

Théorème : Pour que la suite orthonormale des vecteurs propres φ_n correspondant aux valeurs propres $\lambda_n \neq 0$ soit une base hilbertienne de $L^2[0,1]$, il faut et il suffit que 0 ne soit pas valeur propre de T, c'est-à-dire que T soit un opérateur injectif.

En effet, soit f un élément quelconque de $L^2[0,1]$. La série $\sum_{n=1}^{\infty}(f,\varphi_n)\varphi_n$ converge dans $L^2[0,1]$ vers une fonction g telle que

$$T(g) = \sum_{n=1}^{\infty} \lambda_n (f,\varphi_n)\varphi_n$$

Par suite :

$$T(g) - T(f) = T(g-f) = 0$$

donc $g = f$ grâce à l'hypothèse.

Ce que prouve

$$\boxed{f = \sum_{n=1}^{\infty} \langle f,\varphi_n\rangle \varphi_n}$$

et assure que $\{\varphi_n\}$ constitue une base de $L^2[0,1]$ d'après le théorème des bases hilbertiennes. (§ 3.5.2). Réciproquement, supposons que $\{\varphi_n\}$ constitue une base hilbertienne de $L^2[0,1]$. Soit φ une fonction telle que $T\varphi = 0$. Alors

$$\langle \varphi,\varphi_n\rangle = \frac{1}{\lambda_n}\langle \varphi, T\varphi_n\rangle = \frac{1}{\lambda_n}\langle T\varphi,\varphi_n\rangle = 0$$

D'où $\varphi = 0$, ce qui termine la démonstration.

3.8.6. Application à la résolution de l'équation intégrale de Fredholm.

Soit k un noyau non nul tel que $k(x,y) = \overline{k(y,x)}$. Considérons l'équation associée à l'opérateur intégral T de noyau k :

(1) $$f(y) - \lambda \int_0^1 k(x,y)\, f(x)\, dx = g(y)$$

où g est dans $L^2[0,1]$ et f une fonction inconnue dans $L^2[0,1]$, λ est un nombre complexe non nul. On a le théorème suivant (alternative de Fredholm)

Théorème : Si $\frac{1}{\lambda}$ n'est pas une valeur propre de T, cette équation admet une solution unique.

Si $\frac{1}{\lambda}$ est une valeur propre de T, à laquelle correspondent p vecteurs propres $\varphi_{n_0}, \varphi_{n_0+1}, \dots, \varphi_{n_0+p-1}$ orthonormés, l'équation (1) admet une solution si, et seulement si, g est orthogonale aux fonctions $\varphi_{n_0}, \varphi_{n_0+1}, \dots, \varphi_{n_0+p-1}$. La solution n'est alors déterminée qu'à l'addition près d'une combinaison linéaire des $\varphi_{n_0}, \varphi_{n_0+1}, \dots, \varphi_{n_0+p-1}$

On cherche à résoudre dans $L^2[0, 1]$ l'équation

(1) $$f - \lambda T(f) = g$$

où g est donnée dans $L^2[0, 1]$ et où f est l'élément inconnu.

Par conséquent, en supposant l'équation réalisée, on calcule

$$< f , \varphi_n > = \lambda < T(f), \varphi_n > + < g, \varphi_n > = \lambda < f, T(\varphi_n) > + < g, \varphi_n >$$

D'où

$$(3) \quad (1 - \lambda \lambda_n) < f, \varphi_n > = < g, \varphi_n >$$

<u>1er cas</u>

Si donc $\lambda \neq \frac{1}{\lambda_n}$, on aura

$$< f, \varphi_n > = \frac{1}{1 - \lambda \lambda_n} < g, \varphi_n >$$

et

$$f = g + \lambda T(f) = g + \lambda \sum_{n=1}^{\infty} \lambda_n < f, \varphi_n > \varphi_n$$

$$\boxed{f = g + \lambda \sum_{n=1}^{\infty} \frac{\lambda_n}{1 - \lambda \lambda_n} < g, \varphi_n > \varphi_n}$$

Cette série converge au sens de $L^2 [0,1]$ et définit un élément de $L^2[0,1]$ satisfaisant (1) comme il est facile de s'en assurer.(parce que $\sum\limits_{n=1}^{\infty} \lambda_n^2$ est fini)

<u>2ème cas</u> Si $\lambda = \frac{1}{\lambda_{n_0}}$ pour un certain entier n_0, on doit avoir $< g, \varphi_n > = 0$ pour tout entier n tel que φ_n soit vecteur propre de valeur propre λ_{n_0}

$$T(\varphi_n) = \lambda_{n_0} \varphi_n$$

On peut toujours supposer qu'il en est seulement ainsi pour tout n tel que $n_0 \leqslant n < n_0 + p$ où p est l'ordre de la valeur propre λ_{n_0}. Visiblement, une solution de (1) est déterminée à une combinaison linéaire quelconque des $\varphi_{n_0}, \varphi_{n_0+1}, \ldots, \varphi_{n_0+p-1}$.

Pour $n \geqslant n_0 + p$ ou $n < n_0$, le calcul grâce à (2) convient.

D'où la solution générale de (1)

$$f = g + \sum_{n=1}^{n_0-1} \frac{\lambda_n}{1-\lambda \lambda_n} < g, \varphi_n > \varphi_n + \sum_{n_0+p}^{\infty} \frac{\lambda_n}{1-\lambda \lambda_n} < g, \varphi_n > \varphi_n$$

$$+ \sum_{n=n_0}^{n_0+p-1} a_n \varphi_n$$

où les a_n sont des scalaires quelconques.

Ceci termine la démonstration du théorème quelquefois appelé <u>alternative de Fredholm</u>.

Théorème (Hilbert-Schmidt) Sous les conditions de l'alternative de Fredholm, on a l'égalité suivante au sens de la norme de $L^2([0, 1]^2)$

$$(3) \qquad k(x, y) = \sum_{n=1}^{\infty} \lambda_n \overline{\varphi_n(x)} \varphi_n(y)$$

c'est à dire

$$\lim_{N\to\infty} \int_0^1\int_0^1 |k(x, y) - \sum_{n=1}^{N} \lambda_n \overline{\varphi_n(x)} \varphi_n(y)|^2 \, dxdy = 0$$

Grâce à l'orthogonalité des $\{\varphi_n\}_{n\geq 1}$ et à la convergence de la série $\sum_{n=1}^{\infty} |\lambda_n|^2$ il est facile de constater que le second membre de (3) définit un élément de $L^2([0,1]^2)$, noté k'

$$k' = \sum_{n=1}^{\infty} \lambda_n \Phi_n$$

où $\Phi_n(x, y) = \overline{\varphi_n(x)} \varphi_n(y)$.

Cet élément k' définit à son tour un opérateur T' sur $L^2[0, 1]$

$$T'f(y) = \int_0^1 k'(x, y) f(x)dx$$

On constate et justifie l'égalité dans $L^2[0, 1]$

$$T'f = \sum_{n=1}^{\infty} \lambda_n \varphi_n \int_0^1 f(x) \overline{\varphi_n(x)}dx$$

$$T'f = \sum_{n=1}^{\infty} \lambda_n \varphi_n < f, \varphi_n >$$

soit $\quad Tf = T'f \quad$ pour tout f de $L^2[0, 1]$

ou encore $k = k'$ presque partout sur $[0, 1]^2$ comme le lecteur pourra facilement s'en assurer.

Remarque : On peut généraliser l'alternative de Fredholm au cas d'un opérateur compact non nécessairement auto adjoint dans un espace de Hilbert. On se doute, comme dans le théorème de Rouché en dimension finie, du rôle joué par l'équation adjointe homogène $\quad f - \overline{\lambda}T^*(f) = 0$

Comme cet exposé se veut introductif, nous ne donnerons pas les théorèmes plus généraux, renvoyant le lecteur à l'index bibliographique.

Donnons pour terminer un raffinement du théorème de décomposition, pour le cas d'un noyau continu.

Théorème : Si $K(x, y)$ est un noyau continu sur le carré $[0, 1] \times [0,1]$, la série

$$Tf(y) = \int_0^1 K(x, y)\, f(x)dx = \sum_{n=1}^{\infty} \lambda_n \left(\int_0^1 f(x)\, \overline{\varphi_n(x)}dx\right) \varphi_n(y)$$

converge uniformément sur le segment $[0, 1]$.

La seule hypothèse que nous utiliserons est $\int_0^1 |K(x, y)|^2 dx < C^2$ pour tout y dans $[0,1]$. Puisque $f_n = \sum_{k=1}^{k=n} \lambda_k < f, \varphi_k > \varphi_k$ converge vers f dans $L^2[0, 1]$, alors Tf_n converge uniformément dans $[0, 1]$ grâce à la majoration

$$|Tf(y) - Tf_n(y)|^2 \leqslant \left(\int_0^1 |K(x, y)|^2 dx\right) \left(\int_0^1 |f(x) - f_n(x)|^2 dx\right) \leqslant C^2 \|f-f_n\|_2^2$$

et ceci démontre le théorème.

3.9 Problème de Sturm-Liouville

Nous allons appliquer la théorie précédente des équations intégrales au cas des équations différentielles pour certains problèmes aux limites, dits problèmes de Sturm-Liouville.

On considère l'équation différentielle linéaire du second ordre

(1) $x''(t) + (a(t) + \lambda\mu(t))x(t) = y(t)$

où $\begin{cases} \lambda \text{ est un paramètre complexe} \\ a, \mu \text{ et } y \text{ sont des fonctions continues données, à valeurs complexes,} \end{cases}$
sur un intervalle $[\alpha_0, \alpha_1]$ de l'axe réel ($\alpha_0 < \alpha_1$). Pour des raisons qui apparaîtront ultérieurement, on supposera toujours que μ est une fonction strictement positive sur $\Delta=[\alpha_0, \alpha_1]$ et que a est une fonction réelle.

On considère l'espace vectoriel E des fonctions x définies sur $\Delta = [\alpha_0, \alpha_1]$, continues, pourvues de dérivées premières et secondes continues sur Δ, et telles que

$$\begin{cases} A_0 x(\alpha_0) - B_0 x'(\alpha_0) = 0 \\ A_1 x(\alpha_1) - B_1 x'(\alpha_1) = 0 \end{cases}$$

où A_0, B_0 ; A_1 et B_1 sont des nombres réels tels que A_0, B_0 (et A_1, B_1) ne sont pas tous les deux nuls.

Le problème de Sturm-Liouville qui correspond à ces données est de chercher les solutions de (1) qui appartiennent à E.

Exemple : Prenons $a(t) = \pi^2$; $\mu(t) = 1$; $\alpha_0 = 0$ et $\alpha_1 = 1$; $y \equiv 0$ l'équation (1) devient :

$$x''(t) + \pi^2 x(t) + \lambda x(t) = 0$$

Déterminons l'espace E en posant $A_0 = 1$, $B_0 = 0$, $A_1 = 1$, $B_1 = 0$. On vérifie que le problème de Sturm-Liouville correspondant possède une solution non identiquement nulle, si et seulement si λ est de la forme $(n^2 - 1)\pi^2$ où n est un entier supérieur ou égal à 1. Il y a donc un phénomène de quantification.

Nous allons montrer que ce phénomène est général. Une première remarque est que les λ convenables sont tous situés à droite d'une même valeur.

Proposition 3.9.1 : Soit un problème de Sturm-Liouville relatif à une équation (1) homogène ($y \equiv 0$) et un espace E. Il existe une valeur $\lambda_0 \geqslant 0$ telle que la seule solution du problème de Sturm-Liouville pour $\lambda < -\lambda_0$ soit la solution nulle.

Démonstration : On commence par établir une majoration de la norme uniforme de x sur Δ en fonction de la norme $L^2(\Delta)$ de x et de x'.

Précisément on écrit

$$x(\alpha) = x(\beta) + \int_\beta^\alpha x'(t)dt \quad \text{où} \quad \alpha, \beta \in \Delta$$

L'inégalité de Schwarz fournit (cf. E.A.F IA 442)

$$(2) \qquad |x(\alpha)-x(\beta)|^2 \leqslant |\alpha-\beta| \int_{\alpha_0}^{\alpha_1} |x'(t)|^2 \, dt = |\alpha-\beta| \, \|x'\|_2^2$$

Par ailleurs, on dispose de l'égalité du parallélogramme

$$2 \left[|x(\alpha)-x(\beta)|^2 + |x(\beta)|^2\right] = |x(\alpha)|^2 + |x(\alpha) - 2x(\beta)|^2$$

Par suite, en omettant le dernier terme

$$|x(\alpha)|^2 \leqslant 2\left[|x(\alpha) - x(\beta)|^2 + |x(\beta)|^2\right]$$

Inégalité que l'on intègre en β entre α et $\alpha+\varepsilon$ où ε est un nombre positif arbitraire (ou entre $\alpha-\varepsilon$ et α pour $\alpha=\alpha_1$), en tenant compte de (2)

$$\varepsilon|x(\alpha)|^2 \leqslant 2\left(\int_{\alpha}^{\alpha+\varepsilon}(\beta-\alpha)d\beta\right)\|x'\|_2^2 + 2\|x\|_2^2$$

ou encore

$$|x(\alpha)|^2 \leqslant \varepsilon\|x'\|_2^2 + \frac{2}{\varepsilon}\|x\|_2^2$$

et donc

$$(3)\quad \|x\|_\infty \leqslant \varepsilon\|x'\|_2^2 + \frac{2}{\varepsilon}\|x\|_2^2$$

Majorons maintenant l'expression suivante pour $x \in E$

$$I = \int_{\alpha_0}^{\alpha_1}(x''(t) + a(t)\,x(t))\,\overline{x}(t)dt$$

On intègre par parties la première intégrale

$$\int_{\alpha_0}^{\alpha_1} x''(t)\,\overline{x}\,(t)dt = \left[x'\overline{x}\right]_{\alpha_0}^{\alpha_1} - \int_{\alpha_0}^{\alpha_1}|x'(t)|^2dt$$

d'où $$I = \left[x'\overline{x}\right]_{\alpha_0}^{\alpha_1} - \|x'\|_2^2 + \int_{\alpha_0}^{\alpha_1} a(t)\,|x(t)|^2dt$$

On calcule, en tenant compte de ce que $x \in E$ et en supposant $B_1 B_0 \neq 0$

$$\left[x'\overline{x}\right]_{\alpha_0}^{\alpha_1} = \frac{A_1}{B_1}\,|x(\alpha_1)|^2 - \frac{A_0}{B_0}\,|x(\alpha_0)|^2$$

donc $$\left|\left[x'\overline{x}\right]_{\alpha_0}^{\alpha_1}\right| \leqslant \left(\left|\frac{A_1}{B_1}\right| + \left|\frac{A_0}{B_0}\right|\right)\left(\varepsilon\|x'\|_2^2 + \frac{2}{\varepsilon}\|x\|_2^2\right)$$

Puisque a est une fonction continue réelle sur Δ , il existe une constante réelle A et $a(t) \leqslant A$ pour tout t de Δ. D'où la majoration de I.

$$I \leqslant \left[\left(\left|\frac{A_1}{B_1}\right| + \left|\frac{A_0}{B_0}\right|\right)\frac{2}{\varepsilon} + A\right] \|x\|_2^2 + \left[\left(\left|\frac{A_1}{B_1}\right| + \left|\frac{A_0}{B_0}\right|\right)\varepsilon - 1\right]\|x'\|_2^2$$

Il ne reste plus qu'à choisir ε assez petit pour que $\left(\left|\frac{A_1}{B_1}\right| + \left|\frac{A_0}{B_0}\right|\right)\varepsilon-1 < 0$ et poser $\Lambda = \left(\left|\frac{A_1}{B_1}\right| + \left|\frac{A_0}{B_0}\right|\right)\frac{2}{\varepsilon} + A$. Il vient

(4) $\quad I \leqslant \Lambda\|x\|_2^2$

Ces préliminaires de calcul acquis, nous pouvons prouver 3.9.1.

Si maintenant $x \in E$ est solution de l'équation (1) avec $y \equiv 0$, on dispose de

$$I = -\lambda\int_{\alpha}^{\alpha_1} |x(t)|^2\mu(t)dt$$

Si le problème de Sturm-Liouville n'a pas de solution non identiquement nulle pour $\lambda < 0$, on prend $\lambda_0 = 0$ et la proposition est acquise. Sinon, supposons que $\lambda < 0$ et qu'il existe une solution x_λ , non identiquement nulle, au problème de Sturm-Liouville. Puisque μ est continue et est strictement positive sur $[\alpha_0, \alpha_1]$ il existe un $m > 0$ et $I \geqslant -\lambda m \|x_\lambda\|_2^2$. D'où

$$-\lambda m \|x_\lambda\|_2^2 \leqslant \Lambda \|x_\lambda\|_2^2$$

Puisque x_λ n'est pas identiquement nulle, on déduit

$$(-\lambda) \leqslant \frac{\Lambda}{m}$$

et on peut choisir $\lambda_0 = +\frac{\Lambda}{m}$ pour assurer la proposition 3.9.1.

(On laisse au lecteur le soin de vérifier l'exactitude de cette proposition 3.9.1. lorsque $B_1 B_0 = 0$)

Quitte à modifier l'équation (1), en remplaçant a par $a - \lambda_0 - 1$, et λ par $\lambda + \lambda_0 + 1$ on peut faire l'hypothèse suivante pour la suite de ce chapitre

(H) <u>le problème de Sturm-Liouville relatif à E et à l'équation homogène associée à (1) n'a que la solution identiquement nulle pour $\lambda \leqslant 0$.</u>

Pour traiter le <u>problème normalisé</u> (H), on va d'abord le remplacer par un problème équivalent d'équation intégrale, à laquelle on appliquera ensuite la théorie de Fredholm.

Proposition 3.9.2.

Supposons l'hypothèse (H) satisfaite. Il existe une fonction $K : \Delta \times \Delta \longrightarrow \mathbb{R}$ continue et symétrique de sorte que l'assertion suivante soit vérifiée :

Une fonction continue $x : \Delta \longrightarrow \mathbb{R}$ satisfait au problème de Liouville relatif à E et à l'équation (1) si et seulement si x est solution de l'équation intégrale :

$$x(t) - \lambda \int_{\alpha_0}^{\alpha_1} K(\alpha, t)x(\alpha)\mu(\alpha)d\alpha = Y(t) \quad \text{où}$$

$$Y(t) = -\int_{\alpha_0}^{\alpha_1 0} K(\alpha, t)y(\alpha)d\alpha$$

Démonstration A)Il s'agit d'abord de construire la fonction K. Pour ce faire, on construit une fonction $x_0 : \Delta \longrightarrow \mathbb{R}$ telle que x_0 satisfasse l'équation homogène $x_0''(t)+a(t)\, x_0(t) = 0$ et la condition aux limites

$$(5) \qquad A_0\, x_0(\alpha_0) - B_0\, x_0'(\alpha_0) = 0$$

On peut choisir pour x_0 une fonction non identiquement nulle comme l'enseigne la théorie élémentaire des équations différentielles (cf. une démonstration au chapitre IV).

De même, on construit une fonction $x_1 : \Delta \longrightarrow \mathbb{R}$ solution de l'équation homogène $x_1''(t)+a(t)x_1(t) = 0$ et satisfaisant la condition aux limites

$$(6) \qquad A_1\, x_1(\alpha_1) - B_1\, x_1'(\alpha_1) = 0$$

Les fonctions x_1 et x_0 sont linéairement indépendantes grâce à l'hypothèse (H).

Par suite $h_0\, x_0 + h_1\, x_1$, où h_1 et h_2 sont des constantes réelles, est une solution de $x''(t)+a(t)x(t) = 0$. Soit $t \in \,]\alpha_0, \alpha_1[$.

Déterminons $h_0(t)$ et $h_1(t)$ de sorte que l'on ait à la fois $h_0(t)$ $x_0(t) - h_1(t)\; x_1(t) = 0$

et $\qquad h_0(t)\; x_0'(t) - h_1(t)\; x_1'(t) = 1$

Cela est possible, puisque $\delta = x_0(t)\; x_1'(t) - x_1(t)\; x_0'(t)$ est une constante non nulle. En effet :

$$\delta'(t) = x_0(t)\, x_1''(t) - x_1(t)\, x_0''(t) = 0$$

donc δ est une constante, nécessairement non nulle puisque x_0 et x_1 sont linéairement indépendantes.

D'où

$$\begin{cases} h_0(t) = -\dfrac{x_1(t)}{\delta} \\ h_1(t) = -\dfrac{x_0(t)}{\delta} \end{cases}$$

On pose pour $t \in \left]\alpha_0, \alpha_1\right[$ et pour tout $\alpha \in [\alpha_0, \alpha_1]$

$$(7) \quad \begin{cases} K(\alpha, t) = -\dfrac{x_1(t)\, x_0(\alpha)}{\delta} & \text{pour} \quad \alpha_0 \leq \alpha \leq t \\ K(\alpha, t) = -\dfrac{x_0(t)\, x_1(\alpha)}{\delta} & \text{pour} \quad t \leq \alpha \leq \alpha_1 \end{cases}$$

Il est clair que l'on peut prolonger par continuité K pour $t = \alpha_0$ ou $t = \alpha_1$ de sorte que K constitue une fonction continue réelle sur $\Delta \times \Delta$. On en vérifie aisément la symétrie : $K(\alpha, t) = K(t,\alpha)$. On peut résumer les propriétés de $K(\alpha, t)$, appelée la <u>fonction de Green</u> du problème de Sturm-Liouville considéré.

1) $K(\alpha, t) = K(t,\alpha)$

2) Pour chaque t fixé de $\left]\alpha_0, \alpha_1\right[$, $\alpha \longrightarrow K(t,\alpha)$ est deux fois continuement dérivable sur $\left]\alpha_0, t\right[$ et $\left]t, \alpha_1\right[$ et satisfait l'équation $x''(\alpha) + a(\alpha)x(\alpha) = 0$

3) Pour chaque t fixé de $\left]\alpha_0, \alpha_1\right[$, $\alpha \longrightarrow K(t,\alpha)$ appartient à l'espace E, sauf qu'elle n'a pas de dérivée première ou seconde en $\alpha = t$, mais est continue sur $\Delta \times \Delta$

4) $\dfrac{\partial K}{\partial \alpha}(\alpha, t)$ possède une limite à gauche et une limite à droite en $\alpha = t$, notées respectivement $\dfrac{\partial K}{\partial \alpha}(t-0, t)$ et $\dfrac{\partial K}{\partial \alpha}(t+0, t)$. On a

$$\frac{\partial K}{\partial \alpha}(t-0, t) - \frac{\partial K}{\partial \alpha}(t+0, t) = 1$$

B) Considérons maintenant $x(t) = -\int_{\alpha_0}^{\alpha_1} K(\alpha, t)y(\alpha)d\alpha$ où y est une fonction continue et à valeurs complexes sur Δ. Montrons que x est deux fois continuement dérivable sur Δ

$$x(t) = -\int_{\alpha_0}^{t} K(\alpha, t)\, y(\alpha)d\alpha - \int_{t}^{\alpha_1} K(\alpha, t)y(\alpha)d\alpha$$

$$x(t) = \frac{x_1(t)}{\delta} \int_{\alpha_0}^{t} x_0(\alpha)y(\alpha)d\alpha + \frac{x_0(t)}{\delta} \int_{t}^{\alpha_1} x_1(\alpha)y(\alpha)d\alpha$$

d'où

$$x'(t) = \frac{x_1'(t)}{\delta} \int_{\alpha_0}^{t} x_0(\alpha)y(\alpha)d\alpha + \frac{x_0'(t)}{\delta} \int_{t}^{\alpha_1} x_1(\alpha)y(\alpha)d\alpha$$

et

$$x''(t) = \frac{x_1''(t)}{\delta} \int_{\alpha_0}^{t} x_0(\alpha)y(\alpha)d\alpha + \frac{x_0''(t)}{\delta} \int_{t}^{\alpha_1} x_1(\alpha)y(\alpha)d\alpha + y(t)$$

soit

$$x''(t) = -a(t)\, x(t) + y(t)$$

et donc

(8) $\quad x''(t) + a(t)\, x(t) = y(t)$

Etablissons maintenant que x appartient à E.

$$A_0x(\alpha_0) - B_0x'(\alpha_0) = A_0\, \frac{x_0(\alpha_0)}{\delta} \int_{\alpha_0}^{\alpha_1} x_1(\alpha)y(\alpha)d\alpha - B_0\, \frac{x_0'(\alpha_0)}{\delta} \int_{\alpha_0}^{\alpha_1} x_1(\alpha)y(\alpha)d\alpha$$

$$= \frac{(A_0x_0(\alpha_0) - B_0x_0'(\alpha_0))}{\delta} \int_{\alpha_0}^{\alpha_1} x_1(\alpha)y(\alpha)d\alpha = 0$$

Par définition de x_0.

De même $A_1x(\alpha_1) - B_1x'(\alpha_1) = 0$

Nous avons montré que tout $x(t) = -\int_{\alpha_0}^{\alpha_1} K(\alpha, t)\, y(\alpha)d\alpha$ satisfait (8) et appartient à E.

Réciproquement, soit x un élément de E et solution de l'équation différentielle (8) $x''(t) + a(t)\,x(t) = y(t)$ où y est une fonction continue et à valeurs complexes sur Δ.

$$\int_{\alpha_0}^{t} x''(\alpha)K(\alpha,\,t)d\alpha = -\int_{\alpha_0}^{t} x''(\alpha)\,\frac{x_1(t)\,x_0(\alpha)}{\delta}\,d\alpha = \int_{\alpha_0}^{t} x(\alpha)a(\alpha)\,\frac{x_1(t)x_0(\alpha)}{\delta}\,d\alpha + $$

$$- \int_{\alpha_0}^{t} K(\alpha,\,t)\;y(\alpha)d\alpha$$

$$\int_{\alpha_0}^{t} x(\alpha)\quad \frac{\partial^2 K}{\partial \alpha^2}\,(\alpha,t)d\alpha = -\int_{\alpha_0}^{t} x(\alpha)\,\frac{x_1(t)x_0''(\alpha)}{\delta}\,d\alpha = \int_{\alpha_0}^{t} x(\alpha)\,\frac{x_1(t)a(\alpha)x_0(\alpha)}{\delta}\,d\alpha$$

d'où

$$\int_{\alpha_0}^{t}\Big[x''(\alpha)K(\alpha,\,t) - x(\alpha)\;\frac{\partial^2 K}{\partial \alpha^2}\,(\alpha,\,t)\Big]d\alpha = \int_{\alpha_0}^{t} K(\alpha,\,t)\;y(\alpha)d\alpha$$

Et le premier nombre s'écrit : $= \dfrac{x_1(t)}{\delta}\displaystyle\int_{\alpha_0}^{t}\big[x(\alpha)x_0''(\alpha) - x_0(\alpha)\;x''(\alpha)\big]d\alpha$

Soit en intégrant par parties

$$\int_{\alpha_0}^{t} K(\alpha,\,t)\;y(\alpha)d\alpha = \frac{x_1(t)}{\delta}\;\big[x(\alpha)\;x_0'(\alpha) - x_0(\alpha)\;x'(\alpha)\big]_{\alpha_0}^{t}$$

De la même façon, on calcule

$$\int_{t}^{\alpha_1} K(\alpha,\,t)\;y(\alpha)d\alpha = \frac{x_0(t)}{\delta}\;\big[x(\alpha)x_1'(\alpha) - x_1(\alpha)\;x'(\alpha)\big]_{t}^{\alpha_1}$$

soit

$$\int_{\alpha_0}^{\alpha_1} K(\alpha,\,t)y(\alpha)d\alpha = -x(t) + \frac{x_0(t)}{\delta}\,(x(\alpha_1)x_1'(\alpha_1) - x_1(\alpha_1)\;x'\;(\alpha_1))$$

$$- \frac{x_1(t)}{\delta}\,(x(\alpha_0)\;x_0'(\alpha_0) - x_0(\alpha_0)\;x'(\alpha_0))$$

Tenant compte de l'appartenance de x à E et de (5), on a :

$$A_0 x_0(\alpha_0) - B_0 x_0'(\alpha_0) = 0$$

et

$$A_0 x(\alpha_0) - B_0 x'(\alpha_0) = 0$$

Comme A_0 et B_0 ne sont pas tous les deux simultanément nuls, c'est que le déterminant $x_0(\alpha_0)\, x'(\alpha_0) - x_0'(\alpha_0) x(\alpha_0) = 0$. De même, on montre que $x(\alpha_1)\, x'(\alpha_1) - x_1(\alpha_1)\, x'(\alpha_1) = 0$. Soit

$$(9) \qquad - \int_{\alpha_0}^{\alpha_1} K(\alpha, t)\, y(\alpha) d\alpha = x(t)$$

Finalement <u>une fonction continue</u> $x : \Delta \longrightarrow \mathbb{C}$ <u>est solution de</u> (8) $x'' + ax = y$ <u>et</u> $x \in E$ <u>si et seulement si</u> x <u>satisfait l'équation</u> (9)

Dès lors, une fonction continue $x : \Delta \longrightarrow \mathbb{C}$ est solution de l'équation $x'' + ax + \lambda\mu x = y$ et $x \in E$ si et seulement si x satisfait l'équation (9) où y doit être remplacé par $y - \lambda\mu x$. Ce qui donne

$$x(t) = - \int_{\alpha_0}^{\alpha_1} K(\alpha, t) y(\alpha) d\alpha + \lambda \int_{\alpha_0}^{\alpha_1} K(\alpha, t)\, \mu(\alpha)\, x(\alpha) d\alpha$$

ou encore l'équation intégrale de Fredholm.

$$(10) \quad x(t) - \lambda \int_{\alpha_0}^{\alpha_1} K(\alpha, t)\, x(\alpha)\, \mu(\alpha) d\alpha = Y(t)$$

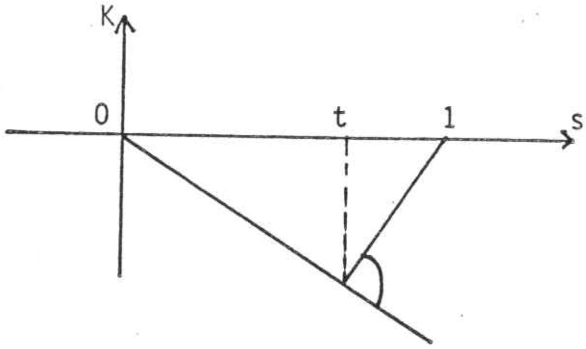
en posant

$$Y(t) = - \int_{\alpha_0}^{\alpha_1} K(\alpha, t)\, y(\alpha) d\alpha$$

Ceci termine la démonstration de la proposition 3.9.2

<u>Exemple</u> : Voici un exemple de transformation d'une équation différentielle en équation intégrale, dans lequel $\lambda = 0$ n'est pas valeur propre. Choisissons $a = 0$, $\mu = 1$, $\alpha_0 = 0$, $\alpha_1 = 1$, $A_0 = 1$, $B_0 = 0$, $A_1 = 1$, $B_1 = 0$:

$x'' - \lambda x = 0$



L'équation $x'' = 0$ a deux solutions :

$$x_0(t) = t \quad , \quad x_1(t) = (1-t)$$

respectivement nulles pour $t = 0$ et $t = 1$. On a donc pour déterminer K

$$At = B(1 - t)$$

$$\text{et} \quad A + B = 1$$

d'où

$$K(\alpha, t) = \alpha(1-t) \qquad 0 \leqslant \alpha \leqslant t$$

$$K(\alpha, t) = (1-\alpha)t \qquad t \leqslant \alpha \leqslant 1$$

La fonction $K(\alpha, t)$ est la <u>fonction de Green</u> de l'équation différentielle $x'' - \lambda x = 0$, pour les conditions aux limites choisies. Nous voyons que c'est une solution de l'équation

$$x'' = 0$$

<u>soit</u> sur l'intervalle $0 \leqslant \alpha < t$, <u>soit</u> sur l'intervalle $t < \alpha \leqslant 1$.

Nous sommes maintenant en mesure d'utiliser la théorie des opérateurs intégraux développée au paragraphe précédent en vue de résoudre le problème de Sturm-Liouville. Enonçons ces résultats sous forme d'un théorème. On notera toutefois que l'espace de Hilbert qui doit intervenir est l'espace $L^2[\Delta, \mu]$

Théorème de Sturm-Liouville

Soient a, μ et y des fonctions à valeurs complexes continues sur intervalle $\Delta = [\alpha_0, \alpha_1]$ de l'axe réel. On suppose μ strictement positive et a à valeurs réelles. On se donne l'espace vectoriel E défini par les quatre nombres réels A_0, A_1, B_0 et B_1 comme cela a été expliqué.

1) Il existe une suite infinie strictement croissante de nombres réels $\{\lambda_n\}_{n\geq 1}$ telle que $\sum\limits_{n=1}^{\infty} \frac{1}{\lambda_n^2}$ converge les valeurs de cette suite sont les seules valeurs de λ dans $\mathbb{C}$ pour lesquelles l'équation

$$(11) \quad x''(t) + a(t)\, x(t) + \lambda\mu(t)\, x(t) = 0$$

possède une solution, non identiquement nulle, dans E. Lorsque $\lambda = \lambda_n$, toutes les solutions de (11) qui appartiennent à E sont multiples d'une fonction x_n de E pour laquelle on impose une normalisation

$$\int_{\alpha_0}^{\alpha_1} |x_n(t)|^2\, \mu(t)dt = 1$$

En outre x_n est à valeurs réelles.

La famille des fonctions $\{x_n\}_{n\geq 1}$ constitue une base hilbertienne de l'espace $L^2(\Delta, \mu)$. En particulier pour $n \neq m$

$$\int_{\alpha_0}^{\alpha_1} x_n(t)\, x_m(t)\, \mu(t)dt = 0$$

2) Si λ n'est pas un élément de la suite $\{\lambda_n\}_{n\geq 1}$ il existe une unique solution x de l'équation

$$(12) \qquad x''(t) + a(t)\, x(t) + \lambda\mu(t)\, x(t) = y(t)$$

qui appartienne à E. Cette solution est donnée par la série absolument et uniformément convergente dans Δ

$$x(t) = \sum_{n=1}^{\infty} c_n\, x_n(t)$$

<u>où l'on a posé</u>

dV

$$c_n = \frac{\displaystyle\int_{\alpha_0}^{\alpha_1} x_n(t)\, y(t)dt}{\lambda - \lambda_n}$$

<u>Démonstration</u> : le problème de Sturm-Liouville revient à étudier l'opérateur linéaire $T : L^2\big[[\alpha_0, \alpha_1], \mu\big] \longrightarrow L^2\big[[\alpha_0, \alpha_1], \mu\big]$ défini par

$$Tx(t) = \int_{\alpha_0}^{\alpha_1} K(\alpha, t)\ x(\alpha)\ \mu(\alpha)d\alpha$$

Il est facile de voir que T est un opérateur linéaire continu hermitien et compact sur $L^2\big[[\alpha_0, \alpha_1], \mu\big]$. Rappelons que sur ce dernier espace, le produit scalaire est fourni par

$$< f, g > = \int_{\alpha_0}^{\alpha_1} f(t)\ \overline{g(t)}\ \mu(t)dt$$

La première assertion de 1) est donc facile à partir de la théorie du §8. En ce qui concerne le fait que chaque espace propre soit de dimension un (engendré par x_n), il suffit de noter que l'ensemble des solutions de l'équation (11) pour $\lambda = \lambda_n$ est un sous-espace vectoriel de dimension deux (la valeur de x et x' en α_0 est arbitraire) lequel ne peut coïncider avec le sous-espace propre.

De même, grâce au §8, la famille $\{x_n\}_{n\geq 1}$ est orthonormale dans $L^2(\Delta,\mu)$. Grâce au § 3.8.5.2, cette famille est une base hilbertienne si l'on peut établir que T est injectif. Or, notons $Z(t) = -Tx(t)$. Cette fonction satisfait l'équation différentielle $Z''(t) + a(t)\ Z(t) = x(t)\ \mu(t)$ grâce au calcul fait à la proposition 3.9.2. Si donc $Z \equiv 0$, on déduit $x(t)\ \mu(t) \equiv 0$. Or μ est strictement positive, donc $x \equiv 0$.

Au sens de $L^2(\Delta,\mu)$, on a l'écriture

$$x = \sum_{n=1}^{\infty} C_n\ x_n$$

où $C_n = \int_{\alpha_0}^{\alpha_1} x_n(t)\ x(t)\ \mu(t)dt$

On écrit puisque $x(t) - \lambda Tx(t) = Y(t) = - \int_{\alpha_0}^{\alpha_1} K(\alpha, t)\ y(\alpha)d\alpha$

$$(13) \qquad C_n = \lambda \int_{\alpha_0}^{\alpha_1} x_n(t)\ Tx(t)\ \mu(t)dt - \int_{\alpha_0}^{\alpha_1} x_n(t)\left[\int_{\alpha_0}^{\alpha_1} K(\alpha, t)\ y(\alpha)d\alpha\right]\mu(t)dt$$

$$\int_{\alpha_0}^{\alpha_1} x_n(t)\ Tx(t)\ \mu(t)dt = \int_{\alpha_0}^{\alpha_1} Tx_n(t)\ x(t)\ \mu(t)dt = \frac{1}{\lambda_n} \int_{\alpha_0}^{\alpha_1} x_n(t)x(t)\mu(t)dt = \frac{C_n}{\lambda_n}$$

La deuxième intégrale dans l'expression (13) est une intégrale double et on intègre d'abord en t puis en α

$$\int_{\alpha_0}^{\alpha_1} \left[\int_{\alpha_0}^{\alpha_1} K(\alpha, t)\ x_n(t)\ \mu(t)dt\right] y(\alpha)d\alpha = \frac{1}{\lambda_n} \int_{\alpha_0}^{\alpha_1} x_n(\alpha)\ y(\alpha)d\alpha$$

Finalement :

$$C_n(1 - \frac{\lambda}{\lambda_n}) = - \frac{\int_{\alpha_0}^{\alpha_1} x_n(\alpha)\ y(\alpha)d\alpha}{\lambda_n}$$

ce qui est bien l'expression donnée pour la valeur de C_n. Nous laisserons au lecteur le soin de démontrer l'uniforme convergence de la série

$$\sum_{n=1}^{\infty} C_n\ x_n(t)$$

EXERCICES SUR LE CHAPITRE III

Exercice n° 30

Polynôme d'Hermite

On part du développement

$$e^{-t^2-2tx} = \sum_{n=0}^{\infty} \frac{t^n}{n!} Qn(X)$$

1°) Montrer que les Qn sont des polynômes tels que

$$Qn(x) = e^{x^2} \frac{d^n(e^{-x^2})}{dX^n}$$

2°) Montrer que la suite $H_n(x) = e^{-\frac{x^2}{2}} Qn(X)$ est orthogonale sur $L^2[-\infty, +\infty]$ en intégrant par rapport à x le produit des séries

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(X) \qquad \text{et} \qquad \sum_{n=0}^{\infty} \frac{s^n}{n!} H_n(X)$$

$$\begin{aligned} \int_{-\infty}^{+\infty} H_n(x)\, H_m(x)\,dx &= 0 \qquad \text{pour } n \neq m \\ &= 2^n\, n!\sqrt{\pi} \quad \text{pour } n = m \end{aligned}$$

Exercices n° 31

Avec les notations de l'exercice précédent, calculer

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} \int_{-\infty}^{+\infty} e^{isx} H_n(s)ds$$

En déduire que $H_n(x)$ est solution de l'équation intégrale

$$\lambda H(X) = \int_{-\infty}^{+\infty} e^{isx} H(s)ds \qquad \text{pour} \quad \lambda_n = i^n\sqrt{2\pi}$$

Exercice n° 32

Montrer qu'il n'existe pas de poids h pour lequel les monômes x^n sont orthogonaux sur l'espace de Hilbert $L^2([-1, +1], h)$.

Exercice n° 33

Utilisation des bases orthogonales.

On désigne par $\mathbb{C}[0, 1]$ l'espace vectoriel des fonctions (à valeurs réelles ou complexes) d'une variable réelle, définies et continues sur l'intervalle fermé $[0, 1]$.

On structure $\mathbb{C}[0, 1]$ en espace préhilbertien noté $\mathcal{C}[0, 1] = E$ à l'aide du produit scalaire usuel :

$$<f, g> = \int_0^1 f(x)\,\overline{g(x)}dx$$

On désigne par H la partie de $\mathcal{C}[0, 1]$ formée des fonctions deux fois continûment dérivables (dérivée à droite en $x = 0$, à gauche en $x = 1$), et vérifiant les relation

$$(1) \qquad \begin{cases} f(0) + f(1) = 0 \\ f'(0) + f'(1) = 0 \end{cases}$$

On considère l'opérateur intégral $\mathbb{A}$, appliquant E dans E, qui à toute fonction f appartenant à E fait correspondre la fonction $Af = F$ définie par

$$F(x) = \int_0^1 (1-2\,|x-y|)\,f(y)dy$$

1°) démontrer que H est un sous-espace vectoriel de E,

2°) a- démontrer que, pour toute fonction f appartenant à E, $Af = F$ appartient à H et vérifie la relation :

$$F''(x) = -4f(x)$$

b- démontrer qu'inversement, quelle que soit la fonction F appartenant à H, il existe une fonction unique f, appartenant à E, telle que $Af = F$

3°) vérifier que l'opérateur $\mathbb{A}$ est hermitien,

4°) démontrer la relation : $\|Af\|^2 \leqslant \frac{1}{3}\|f\|^2$ valable pour toute fonction $f \in E$ et qui montre que l'opérateur $\mathbb{A}$ est borné, avec $\|A\| \leqslant 1/\sqrt{3}$

5°) déterminer les valeurs propres de l'opérateur $\mathbb{A}$ et montrer qu'à chaque valeur propre λ_n correspond un espace propre P_n de dimension 2.

6°) a- on désigne par $(f_{1,n}, f_{2,n})$ une base orthonormée de l'espace propre P_n.

Montrer que la fonction : $\varphi_n(x, y) = f_{1,n}(x), f_{1,n}(y) + f_{2,n}(x)\, f_{2,n}(y)$

est indépendante de la base orthonormée de P_n choisie.

b- soit $s(t)$ la fonction périodique de période 2, définie par :

$$s(t) = 1-2|t|$$

sur l'intervalle $[-1, +1]$. Calculer le développement en série de Fourier de $s(t)$. En déduire la relation :

$$1-2|x-y| = \sum_{n=0}^{\infty} \lambda n\, \varphi n(x, y)$$

Préciser le domaine de validité de cette relation dans le plan des (x, y).

Exercice n° 34

Calcul de l'orthonormalisation de Gram-Schmidt.

Soit $\mathcal{H}$ un espace préhilbertien réel et $x_1, x_2, \ldots x_n$, n vecteurs linéairement indépendants. Montrer qu'il existe n éléments $y_1, y_2, \ldots y_n$, formant un système orthonormal de sorte que l'espace vectoriel engendré par $y_1, \ldots y_p$ coïncide avec l'espace vectoriel engendré par $x_1, \ldots x_p$ (construire pas à pas le vecteur y_p), $p \leqslant n$).

$$\text{Soient} \quad G(x_1, x_2, \ldots x_p) = \begin{bmatrix} \langle x_1, x_1\rangle & \langle x_1, x_2\rangle & \langle x_1, x_p\rangle \\ \vdots & \vdots & \vdots \\ \langle x_p, x_1\rangle & \langle x_p, x_2\rangle & \langle x_p, x_p\rangle \end{bmatrix}$$

et $g(x_1, x_2, \ldots x_n)$ le déterminant de cette matrice. A un signe près, montrer que y_n peut s'écrire (sauf pour n = 1 auquel cas $y_1 = \dfrac{x_1}{\sqrt{g(x_1)}}$

$$y_n = \frac{1}{\sqrt{g(x_1, \ldots, x_{n-1})\, g(x_1, \ldots, x_n)}} \begin{vmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \ldots & \langle x_n, x_1\rangle \\ \vdots & \vdots & & \vdots \\ \langle x_1, x_{n-1}\rangle & \langle x_2, x_{n-1}\rangle & \ldots & \langle x_n, x_{n-1}\rangle \\ x_1 & x_2 & & x_n \end{vmatrix}$$

(on utilisera les règles de Cramer pour le calcul de y_n en remarquant qu'on obtient y_n en minimisant

$$\| x_n - (a_1 x_1 + \ldots + a_{n-1} x_{n-1}) \|$$

retrouver ainsi les trois premiers polynômes de Legendre.

Exercice n° 35

Pour une famille de polynômes orthonormaux, établir la relation

$$\sum_{n=0}^{n} P_n^2(x) = \frac{a_n}{a_{n+1}} (P'_{n+1}(x)\, P_n(x) - P'_n(x)\, P_{n+1}(x))$$

(utiliser la formule de Christoffel-Darboux en faisant tendre y vers x).

Exercice n° 36

Soient $T_n(x)$ les polynômes de Čebičev de premier type, c'est-à-dire les polynômes orthogonaux sur $L^2[(-1, +1), \frac{1}{\sqrt{1-x^2}}]$ (on remarque que ces polynômes sont des cas particuliers des polynômes de Jacobi avec $\alpha = \beta = -\frac{1}{2}$).

Vérifier les diverses propriétés suivantes (avec une normalisation convenable)

1°) $\displaystyle\int_{-1}^{+1} \frac{1}{\sqrt{1-x^2}} (T_n(x))^2 dx = \frac{\pi}{2}$ si $n \geqslant 1$ et π si $n = 0$

2°) $a_n = 2^{n-1}$

3°) $A_n = 2$, $B_n = 0$ et $C_n = 1$

4°) $T_n(\cos x) = \cos nx$

5°) $(1 - x^2) \dfrac{d^2 T_n(x)}{dx} - x \dfrac{d\, T_n(x)}{dx} + n^2 T_n(x) = 0$

6°) $$2^n \frac{\Gamma(n+\frac{1}{2})}{\Gamma[\frac{1}{2}]} T_n(x) = (-1)^n \sqrt{1-x^2} \frac{d^n((1-x^2)^{n-\frac{1}{2}})}{dx^n}$$

7°) $$T_n(x) = \frac{n}{2} \sum_{k=0}^{[\frac{n}{2}]=k} (-1)^k \frac{(n-k-1)}{k!(n-2k)!} (2x)^{n-2k}$$

où $[\frac{n}{2}]$ désigne la partie entière de $\frac{n}{2}$

8°) $|T_n(x)| \leqslant 1$ pour $-1 \leqslant x \leqslant +1$

Exercice n° 37

I°) Soit f une fonction développable en série de polynômes de Čebičev de première espèce selon

$$f(x) = \sum_{n=0}^{\infty} a_n T_n(x) \quad \text{où } |x| \leqslant 1 \quad \text{et} \quad \sum_{n=0}^{\infty} |a_n| < +\infty$$

Montrer que :

$$\int_{-1}^{x} f(t)dt = \lambda + \sum_{k=1}^{\infty} \left(\frac{a_{k-1} - a_{k+1}}{2k}\right) T_k(x) \qquad |x| \leqslant 1$$

2°) montrer que le développement d'une fonction $f(x)$ en série de polynômes de Čebičev de première espèce correspond au développement de $f(\cos x)$ en série de cosinus. En déduire le développement de $\frac{1}{1+x^2}$ en séries de polynômes de Čebičev.

$$\frac{1}{1+x^2} \sim \sqrt{2}\ \left(\tfrac{1}{2}(\sqrt{2}-1)^2\ T_2(x) + \ldots + (-1)^n (\sqrt{2}-1)^{2n}\ T_{2n}(x) + \ldots\right)$$

Exercice

Propri zéros des polynômes orthogonaux

Considé famille des polynômes orthogonaux de Legendre sur $[-1, +1]$ pour l'espace $+1]$

a) montrer s n racines de $P_n(x)$ sont réelles, distinctes et situées dans l'inte $[-1, +1]$.

(pour eff la démonstration, on remarque que $\int_{-1}^{+1} P_n(x)dx = 0$, donc que P_n admet au mo zéro dans $[-1, +1]$ autour duquel $P_n(x)$ change de signe. On raisonne al l'absurde en supposant qu'il existe ℓ racines distinctes $x_1, x_2, \ldots$ ans $[-1, +1]$ et en considérant $P_n(x)\ (x-x_1)\ldots(x - x_\ell)$

On en dédque nécessairement $\ell = n$)

b) soient $x_1, \ldots x_n$ les zéros de $P_n(x)$ et posons $x_0 = 1$ et $x_{n+1} = +1$

Dans chaque rvalle $[x_k, x_{k+1}]$ pour $k = 0, \ldots n$, il y a exactement un seul zéro de P_{n+1}

On pourra éduire une décomposition de la fraction rationnelle où y_k désigne les zé de $P_{n+1}(x)$

$$\frac{P_n(x)}{P_{n+1}(x)} = \sum_{k=0}^{n+1} \frac{\alpha k}{x-y_k} \qquad \text{avec} \qquad \alpha_k > 0$$

Les démonstrons de cet exercice sont généralisables à toute famille de polynômes orthoaux sur $L^2[(a, b), h]$ pour a et b finis.

Exercice n° 39

Soit $\{P_n\}$ un famille de polynômes orthonormaux sur $L^2([a, b], h(x))$.
Soit $\pi(x) = (x - x_1)\ldots(x - x_h)$ où $x_1, x_2, \ldots x_h$ sont des nombres distincts

de sorte que $\pi(x) \geqslant 0$ sur $[a, b]$.

Montrer que la famille Q_n où

$$Q_n(x) = \frac{1}{\pi(x)} \begin{vmatrix} P_n(x) & \dots & P_{n+h}(x) \\ P_n(x_1) & \dots & P_{n+h}(x_1) \\ P_n(x_h) & \dots & P_{n+h}(x_h) \end{vmatrix}$$

est orthogonale sur $L^2([a, b], \pi(x)\, h(x))$

Exercice n° 40

Soit E un espace de Hilbert et $\mathcal{U}$ un opérateur hermitien linéaire et continu de E dans E. On suppose que

$$\|\mathcal{U}^n\| \leqslant C \qquad \text{pour tout entier} \quad n \geqslant 0$$

où C est une constante positive.

En utilisant le théorème de projection, montrer que pour chaque x de E

$$P_\infty(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{h=1}^{n} \mathcal{U}^h(x)$$

existe au sens de l'espace E. Montrer que P_∞ définit un opérateur linéaire, projecteur orthogonal, tel que :

$$P_\infty \circ \mathcal{U} = \mathcal{U} \circ P_\infty = P_\infty$$

Exercice n° 41

Soit E un espace vectoriel préhilbertien sur le corps des complexes. Soient P et Q deux opérateurs hermitiens de E dans E tels que :

$$[P, Q] = PQ - QP = iI$$

où I est l'opérateur identité ; $[P, Q]$ s'appelle <u>crochet de Poisson</u> de P et Q.

On pose :

$$A = P + iQ$$

$$A^* = P - iQ$$

$$H = \frac{1}{2}(P^2 + Q^2)$$

1°) calculer $A A^*$ et $A^* A$ en fonction de H.

En déduire que : $[A, H] = A$

$$[H, A^*] = A^*$$

2°) x_λ étant un vecteur propre relatif à une valeur propre λ de H (s'il en existe), montrer que :

$$\|A^* X\|^2 = (2\lambda + 1)\|X_\lambda\|^2$$

$$\|A X\|^2 = (2\lambda - 1)\|X_\lambda\|^2$$

$$H A X_\lambda = (\lambda - 1) AX_\lambda$$

$$H A^* X_\lambda = (\lambda + 1) A^* X_\lambda$$

3°) déduire de ces formules que :

1) Si le spectre S(c'est-à-dire l'ensemble des valeurs propres de H) est non vide, il est borné inférieurement

2) Si $\lambda = 1/2$ n'appartient pas à S, S est vide

3) Si $\lambda = 1/2$ appartient à S, alors $\lambda_n = n + \frac{1}{2}$ sont des valeurs propres

4) Tous les vecteurs propres correspondant à la valeur propre λ_n sont de la forme :

$$X_{\lambda_n} = (A^*)^n X_{\lambda_0}$$

où X_{λ_0} est un vecteur propre correspondant à la valeur propre λ_0

5) $\lambda_0 = 1/2$ est valeur propre de H si, et seulement si, il existe un vecteur X tel que $AX = 0$. Quels sont les vecteurs propres $X_{1/2}$ de H ?

Exercice n° 42

On considère l'équation différentielle $xy'' + y' + \lambda xy = 0$ où λ est un nombre réel donné, et on appelle problème P_λ la recherche des solutions y non nulles et à valeurs réelles de cette équation satisfaisant les conditions C suivantes :

$$\begin{cases} y(0) = 1 \\ y(1) = 0 \end{cases}$$

1ère question

Montrer que pour deux valeurs distinctes de λ, les solutions (supposée exister) correspondantes des problèmes P_λ respectifs vérifient une propriété d'orthogonalité relativement à un certain produit scalaire. En déduire qu'elles sont linéairement indépendantes.

2ème question

Montrer que l'ensemble Λ des valeurs λ, pour lesquelles P_λ admet une solution, est formé de nombres positifs.

3ème question

On appelle E l'espace formé par les fonctions f d'une variable réelle et à valeurs réelles telles que $x(f(x))$ soit intégrable entre 0 et 1.

a) par analogie avec $L^2[0, 1]$, définir de façon simple un produit scalaire et une norme sur l'espace vectoriel E,

b) établir les relations d'inclusion entre les espaces E, $L^1[0, 1]$ et $L^2[0, 1]$. Comparer les normes définies sur ces trois espaces pour une fonction appartenant à l'intersection de ces trois espaces fonctionnels,

c) construire à partir des questions précédentes un système orthonormé dans l'espace E.

4ème (

Etquation différentielle que vérifie la fonction $y_\lambda(\frac{x}{\sqrt{\lambda}})$ où $y_\lambda(x)$ désigrlution du problème P_λ. En déduire que le problème P_λ possède au plus uon en utilisant quelques résultats classiques sur les fonctions de Bes

Doncaractérisation de l'ensemble Λ.

5ème q

Le $\{y_\lambda\}$ pour $\lambda \in \Lambda$ est-il total dans l'espace E ?
Calculeloppement de 1 en série de Fourier généralisée relativement à ce systèmer les relations de récurrence définies sur les fonctions de Bessel d'ordret de première espèce).

6ème qu

Mon un minimum de calcul qu'un intervalle de longueur finie de l'axe réel net qu'un nombre fini de valeurs λ appartenant à Λ.

Que en déduire pour l'ensemble Λ ?

Peutiser ce résultat ?

Exercic

Soit ous-espace vectoriel fermé d'un espace de Hilbert E.
1°) réc les propriétés de l'opérateur de projection orthogonale sur M,
2°) soi érateur linéaire et continu, idempotent et hermitien sur E.
Mon P est un projecteur orthogonal,
3°) mon $(M^\perp)^\perp = M$ et plus généralement que $(F^\perp)^\perp = \overline{F}$ lorsque F est un sous ectoriel quelconque de E.

- CHAPITRE IV -

THEOREMES DU POINT FIXE

Les chapitres 0, I et II nous ont essentiellement introduit des notations et des outils pratiques, notamment la densité, la complétion et le théorème de prolongement.

Le chapitre III nous a permis une approche géométrique et fourni des théorèmes d'existence pour des équations linéaires dans un espace fonctionnel de Hilbert, ainsi que des outils très souples comme les bases hilbertiennes.

Pour terminer ces rudiments d'Analyse Fonctionnelle, il convient de faire jouer un rôle essentiel à la propriété pour un espace d'être complet en ne gardant plus que la notion de norme ou de distance et non la notion de produit scalaire. Les théorèmes constructifs que nous allons obtenir tournent tous autour de la recherche d'un point fixe. Nous ne donnerons pas ici les applications qui sont fort nombreuses (théorème des fonctions implicites, théorème de Cauchy-Lipchitz en théorie des équations différentielles par exemple,...) les réservant à des cours plus spécialisés.

4.1 THEOREME DES APPROXIMATIONS SUCCESSIVES DE PICARD

Définition : soit (E, d) un espace métrique et U une application de E dans E. On dit que U est k -contractante s'il existe un nombre réel $k (0 < k < 1)$ tel que $d(Ux, Uy) \leqslant k\, d(x, y)$ pour tous x, y dans E.

Théorème : soit U une application k-contractante sur un espace métrique complet (E, d). Il existe un unique point x de E tel que $Ux = x$

Un tel point s'appelle un point fixe de l'application U.

<u>Démonstra</u> la démonstration de ce théorème est d'autant plus intéressante qu'elle fourı algorithme de calcul de x et précise même la rapidité de convergence.

L'unicité est évidente car si x et x' sont deux points fixes pour U, on doit avoiı

$$x, Ux') = d(x, x') \leqslant kd(x, x') \quad \text{et} \quad k < 1$$

donc $d(x = 0$, soit $x = x'$.

On part d'int x_0 arbitraire et on définit par récurrence une suite x_n de points de elon

$$x_{n+1} = U\, x_n$$

La suite st une suite de Cauchy grâce au calcul suivant

$$d(x_{n+p}, x_n) \leqslant_{+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \dots + d(x_{n+1}, x_n)$$

mais on a

$$d(x_m) \leqslant k\; d(x_n, x_{n-1}) \leqslant k^2 d(x_{n-1}, x_{n-2})$$

$$\leqslant k^n d(x_1, x_0)$$

donc

$$d(x_m) \leqslant (k^{n+p-1} + \dots + k^n)\; d(x_1, x_0)$$

$$\leqslant k^n \frac{1}{1-k}\, d(x_1, x_0)$$

puisque $0 <$ La majoration est indépendante de p et tend vers zéro lorsque n tend vers ini. La suite $\{x_n\}$ est donc de Cauchy et converge vers un point x de ıisque (E, d) est complet.

Etudions Upart de $x_{n+1} = Ux_n$ et puisque x_{n+1} converge vers x ainsi que x_n, tand U est continue (puisque k-contractante), on a

$$\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} Ux_n = U(\lim_{n\to\infty} x_n) \quad \text{soit} \quad x = Ux \quad \text{c.q.f.d.}$$

On a souvent besoin d'une légère généralisation de ce théorème

Théorème : soit (E, d) un espace métrique et soit U une application de E dans E. Supposons qu'il existe un entier $p (\geq 1)$ tel que la composée d'ordre p de U soit k-contractante, alors il existe un unique point fixe pour U.

On fait pour U^p ce qu'on a fait précédemment pour U en partant successivement de x_0, $U(x_0)$, ..., $U^{p-1}(x_0)$ où x_0 est un point arbitraire de E. Pour chaque entier Γ, où $0 \leq \Gamma < p$, alors $U^{kp+\Gamma}(x_0)$ converge vers l'unique point fixe x de U^p lorsque k tend vers l'infini. On en déduit bien que $U^n(x)$ converge vers x et donc satisfait $Ux = x$. En outre, comme $U^p x = x$ a une unique solution dans E, a fortiori il en est de même pour l'équation $Ux = x$.

Ce théorème de point fixe est particulièrement important en Analyse car il fournit un grand nombre de théorèmes d'existence. En outre, c'est un théorème constructif en ce sens qu'il donne un algorithme pratique pour déterminer le point fixe.

Donnons un exemple concernant les opérateurs intégraux de Volterra.

4.2 APPLICATION AUX OPERATEURS DE VOLTERRA

Proposition : soit $\Delta = [\alpha, \beta]$ un intervalle borné de l'axe réel et $K(t, s)$ une fonction continue et à valeurs complexes définie sur le produit $\Delta \times \Delta$. Soit $b(t)$ une fonction continue et à valeurs complexes définie sur Δ.

L'équation intégrale

$$(1) \qquad x(t) + \lambda \int_\alpha^t K(t, s)\, x(s) ds = b(t)$$

où λ désigne un paramètre complexe, possède une unique solution x continue sur Δ.

Démonstration : Rappelons que l'espace $E = C[\alpha, \beta]$ des fonctions continues et à valeurs complexes muni de la norme uniforme est un espace de Banach, c'est-à-dire en particulier un espace métrique complet. Considérons l'opérateur U défini sur $E = C[\alpha, \beta]$ par l'équation $(\alpha < \beta)$

$$U(x)(t) = -\lambda \int_{\alpha}^{t} K(t, s)\, x(s) ds + b(t)$$

Montrons d'abord que U est un opérateur défini de E dans E. Avec les hypothèses faites, $U(x)(t)$ est bien définie pour chaque valeur de t dans Δ, et comme nous l'avons déjà vu $U(x)$, considérée comme fonction de t, est continue.

[On peut vérifier que si $b \equiv 0$, alors U est un opérateur linéaire continu sur E. La continuité provient de la majoration suivante :

$$\sup_{t \in \Delta} |U(x)(t)| = \|U(x)\| \leqslant |\lambda| \sup_{(t,s) \in \Delta\times\Delta} |K(t, s)| \|x\| (\beta - \alpha)$$

En notant :

$$M = \sup_{(t,s) \in \Delta\times\Delta} |K(t,s)|$$

on a bien

$$\|U(x)\| \leqslant M|\lambda| \|x\| (\beta - \alpha)]$$

Reprenant le cas général, on dispose de la majoration

$$\|U(x) - U(y)\| \leqslant |\lambda|\, M\|x-y\| (\beta - \alpha).$$

Si $M|\lambda| (\beta - \alpha) < 1$, on se trouve dans le cas d'application du premier théorème du point fixe, mais en général, cette majoration n'est pas vraie. Nous allons alors étudier un composé d'ordre suffisamment grand de U.

Retenons quand même que pour λ assez petit, précisément si

$$|\lambda| < \frac{1}{(\beta-\lambda) \sup\limits_{(t,s)\in\Delta\times\Delta} |K(t,s)|} = \frac{1}{M(\beta - \alpha)}$$

alors l'équation

$$x(t) = - \lambda \int_{\alpha}^{t} K(t, s)\ x(s)ds + b(t)$$

admet une solution unique dans l'espace E.

Passons au composé d'ordre p de l'opérateur U. On a :

$$U(x)(t) - U(y)(t) = - \lambda \int_{\alpha}^{t} K(t, s)\ (x(s) - y(s))ds$$

et l'on a en notant toujours

$$M = \sup_{s,t\in\Delta\times\Delta} |K(t, s)|$$

$$|U(x)(t) - U(y)(t)| \leqslant |\lambda|\ M\|x-y\|(t - \alpha)$$

Montrons par récurrence que l'on a la majoration

$$|U^p(x)(t) - U^p(y)(t)| \leqslant |\lambda|^p\ M^p\ \frac{(t-\alpha)^p}{p!}\ \|x-y\|$$

Pour ce faire, on calcule $U^{p+1}(x) - U^{p+1}(y)$

$$U(U^p(x))(t) - U\ (U^p(y))(t) = -\lambda \int_{\alpha}^{t} K(t, s)\ [U^p(x)(s) - U^p(y)(s)]ds$$

par conséquent

$$|U^{p+1}(x)(t) - U^{p+1}(y)(t)| \leqslant |\lambda| \int_{\alpha}^{t} |K(t, s)|\ |U^p(x)(s) - U^p(y)(s)|ds$$

$$\leqslant |\lambda|\ M \int_{\alpha}^{t} |\lambda|^p\ M^p\ \frac{(s-\alpha)^p}{p!}\ ds$$

$$\leqslant |\lambda|^{p+1} M^{p+1} \frac{(t - \alpha)^{p+1}}{(p+1)!}$$

ce qui termine la démonstration par récurrence et fournit la majoration cherchée.

En majorant l'expression obtenue, il vient

$$\|U^p(x) - U^p(y)\| \leqslant |\lambda|^p M^p \frac{(\beta - \alpha)^p}{p!} \|x - y\|$$

Mais nous savons que $\lim\limits_{p \to +\infty} \frac{|\lambda|^p M^p(\beta-\alpha)^p}{p!} = 0$. Par suite, il existe un entier

p tel que U^p Soit contractante. Appliquant le deuxième théorème du point fixe, on assure que pour tout λ complexe, il existe une unique solution continue à l'équation de Volterra

$$x(t) = -\lambda \int_\alpha^t K(t, s)\ x(s)ds + b(t)$$

Ce résultat constitue le <u>théorème de Volterra</u>.

Lorsque $b(t) \equiv 0$, alors U est un opérateur linéaire continu. Visiblement $x \equiv 0$ est solution de l'équation

$$Ux = x$$

D'après ce qui précède, c'est la <u>seule</u> solution de cette équation. Par suite, $\int_\alpha^t K(t, s)\ x(s)ds$ est un exemple d'opérateur linéaire continu sur l'espace de Banach E, n'admettant <u>aucune valeur propre différente de zéro</u>.

Le cas $\lambda = +\infty$, soit $\int_\alpha^t K(t, s)\ x(s)ds \equiv 0$, doit être mis à part. En dérivant, on en déduit :

$$K(t, t)\ x(t) + \int_\alpha^t \frac{\partial K}{\partial t}(t, s)\ x(s)ds \equiv 0$$

Si l'on suppose $K(t,t) \neq 0$ en tout point t de Δ, et $\frac{\partial K}{\partial t}(t,s)=0$ pour $\alpha \leqslant s \leqslant t$, alors on dispose bien d'un <u>opérateur linéaire continu n'ayant aucune valeur propre sur l'espace de Banach E</u>. On aura remarqué que ces opérateurs de Volterra sont à l'opposé des opérateurs intégraux hermitiens (noyau non symétrique).

4.3 GENERALISATIONS POSSIBLES

4.3.1 Une idée de généralisation est de diminuer les hypothèses sur l'application U $E \longrightarrow E$ dans le théorème des approximations successives de Picard, quitte à restreindre la classe des espaces métriques (E, d) considérés.

Un exemple simple est celui où $E = [0, 1]$, muni de la métrique usuelle et $U : [0, 1] \longrightarrow [0, 1]$ une application continue quelconque. Considérons la fonction g, définie sur $[0, 1]$ par ses valeurs

$$g(x) = U(x) - x \qquad \forall x \in [0, 1]$$

Cette fonction s'annule une fois au moins, c'est-à-dire que U possède un point fixe. En effet de deux choses l'une

ou bien $g(0) = 0$ et 0 est point fixe de U

ou bien $g(0) > 0$

De même, de deux choses l'une :

ou bien $g(1) = 0$ et $x = 1$ est point fixe de U

ou bien $g(1) < 0$

On a donc à traiter le seul cas $g(0) > 0$, $g(1) < 0$. Mais puisque la fonction g est continue, elle prend toute valeur intermédiaire entre $g(0) > 0$ et $g(1) < 0$, en particulier la valeur 0. Il existe donc un point fixe de U. On peut généraliser ce résultat, mais par une démonstration tout à fait différente (cf Exercice n° 49 pour le cas de $\mathbb{R}^2$). Cela constitue un résultat important.

Théorème de Brouwer : Soit E un espace vectoriel de dimension finie sur le corps $\mathbb{R}$ (ou $\mathbb{C}$) et soit B une boule fermée de E. Soit $U : B \longrightarrow B$ une application continue. Il existe au moins un point fixe x de U dans B :

Dans la démonstration du théorème de Brouwer, deux types de propriétés interviennent pour les boules fermées : d'une part, la convexité, d'autre part, le fait d'être fermé et borné. On peut montrer que la propriété : E est fermé et borné, est identique à la propriété suivante dite de compacité (métrique) dans un espace vectoriel de dimension finie.

De toute suite $\{x_n\}_{n\geq 1}$ de E on peut extraire une sous-suite convergente dans E. Ce résultat provient du théorème de Bolzano-Weierstrass, mais pour des raisons pédagogiques nous n'avons pas voulu insister sur la compacité dans ce cours).

On dispose alors d'un théorème nettement plus fort que celui de Brouwer.

Théorème de Schauder-Tychonoff : Soit E un espace vectoriel normé sur $\mathbb{R}$(ou $\mathbb{C}$) Soit B un sous-ensemble convexe et compact non vide de E. Soit U : E $\longrightarrow$ E une application continue. Il existe au moins un point fixe x de U dans E.

Le théorème de Schauder-Tychonoff a été extrêmement généralisé par Leray et Schauder avec la notion de degré topologique en vue d'applications fort brillantes à la théorie des équations aux dérivées partielles.

4.3.2. Une généralisation beaucoup plus modeste concerne le cas d'une application U : B $\longrightarrow$ B, où B est un sous-ensemble convexe fermé, non vide, d'un espace de Banach E supposé réel, telle que $\|U_x - U_y\| \leq \|x-y\|$ pour tous x, y de B. On pourrait dire que U est 1-contractante. Mais le théorème de Picard ne s'applique pas. Toutefois, on peut définir quelquefois un nouvel opérateur V : B $\longrightarrow$ B pour lequel le théorème de Picard conviendra.

On pose $V_x = \frac{1}{2}(Ux + x)$. Grâce à la convexité $V(B) \subset B$. En outre, si l'on suppose, par exemple, $U(x)V(y) = k(x, y)\,(x-y)$ avec $-1 \leq k(x, y) \leq 0$, on calcule $U(x)-V(y) = \frac{x-y}{2}(1 + k(x, y))$

Soit $\|V(x)\ V(y)\| \leq \frac{\|x-y\|}{2}$ puisque $1 \geq 1 + k(x, y) \geq 0$

Donc V est $\frac{1}{2}$ -contractante. Le théorème de Picard s'applique puisque B est complet. Il existe un unique point fixe pour V, donc pour U.

Exemple : $B = [\sqrt{2}, \infty[$, $E = \mathbb{R}, U(x) = \frac{2}{x}$, $V_x = \frac{1}{2}(1 + \frac{2}{x})$

Le point fixe de U est $x = \sqrt{2}$ et l'algorithme du théorème de Picard pour V est un procédé de calcul rapide de $\sqrt{2}$.

EXERCICES SUR LE CHAPITRE IV

Exercice n° 44

Avec les notations de ce chapitre, résoudre l'équation

$$\int_0^s \frac{(s-t)^n}{n!}\, x(t)dt = x(s)$$

dans l'espace E des fonctions continues.

En déduire la formule de Taylor sous forme intégrale (on considèrera l'équation intégrale

$$\int_0^s \frac{(s-t)}{n!}\, x(t)dt = y(s)$$

où y est une fonction continue).

Exercice n° 45

Un espace métrique (E, d) est dit α-chaînable s'il existe un nombre $\alpha > 0$ fixe tel que, étant donné deux points quelconques x et y de E, il existe une suite finie $\{x_i\}_{i=0,\dots n}$ satisfaisant

a) $x_0 = x$; $x_n = y$ et $d(x_{i-1}, x_i) < \alpha \quad 1 \leqslant i \leqslant n$

Naturellement, l'entier n et les points x_i dépendent des deux points x et y.

1°) montrer que $\delta(x, y) = \mathrm{Inf}(\sum_{i=1}^{n} d(x_i, x_{i-1}))$, où la borne inférieure est à prendre sur l'ensemble des suites finies vérifiant (a), définit une distance sur l'espace E,

2°) supposons que (E, d) soit un espace métrique α-chaînable et complet. Supposons en outre que $f : E \longrightarrow E$ satisfasse

$\exists\, k < 1$ tel que $\forall x \in E$; $\forall y \in E$ et $d(x, y) < \alpha$ entraîne

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

Montrer que f possède un unique point fixe.

Exercice n° 46

Soit $[A] = [a_{ij}]$ une matrice $n \times n$, strictement diagonalement dominante, c'est-à-dire telle que

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| < a_{ii} \qquad \text{pour} \qquad i = 1, 2, \ldots, n .$$

Montrer que la matrice $[A]$ est inversible.

Exercice n° 47

Donner des contre-exemples montrant que les hypothèses du théorème des approximations successives de Picard sont essentielles.

Exercice n° 48

Démontrer le théorème suivant :

soit Ω un ouvert convexe d'un espace vectoriel normé $(X, \|.\|)$ de dimension finie. Soit $f : \Omega \longrightarrow \mathbb{R}$ une application convexe, c'est-à-dire telle que

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha)\, f(y) \qquad \begin{array}{l} \forall \alpha \; ; \; 0 < \alpha < 1 \\ \forall x \in \Omega \\ \forall y \in \Omega \end{array}$$

La fonction f est alors continue. En outre elle satisfait localement une condition de Lipschitz ; c'est-à-dire que pour tout a de Ω il existe une boule ouverte V centrée en a et une constante C (dépendant de a et de la boule ouverte) de sorte que

$$|f(x) - f(y)| \leqslant C\|x-y\| \qquad \forall x, y \in V$$

Exercice n° 49

Soit $\mathcal{D}$ un domaine simplement connexe du plan $\mathbb{R}^2$ limité par une courbe simple de Jordan Γ. Soit $\vec{V}(M)$ un champ de vecteurs, continu, défini dans le plan. On appelle rotation du champ $\vec{V}$ le long de Γ, et on note $\alpha(V, \Gamma)$, l'angle dont tourne le vecteur $\vec{V}(M)$ lorsque le point M décrit la courbe Γ dans le sens direct. Cette rotation n'est bien définie que si $\vec{V}(M)$ ne s'annule pas sur la courbe Γ. On remarque alors que

$$\alpha(V, \Gamma) = 2k\pi \quad \text{où } k \text{ est un entier relatif.}$$

On suppose désormais que le champ $\vec{V}$ est deux fois continûment différentiable et ne s'annule pas sur Γ.

Première question

φ désigne l'angle que fait $\vec{V}(M)$ avec une direction fixe. Exprimer $d\varphi$ en fonction de $d\vec{M}$ et en déduire une expression de $\alpha(V, \Gamma)$ sous forme d'une intégrale curviligne.

Deuxième question

A l'aide de la formule de Riemann, démontrer que si $\vec{V}(M)$ ne s'annule pas dans le domaine $\mathcal{D}$, la rotation $\alpha(V, \Gamma)$ est nulle.

Que peut-on dire si la rotation $\alpha(V, \Gamma)$ n'est pas nulle ? Cette dernière assertion admet-elle une réciproque ?

Troisième question

Etudier les cas suivants :

a) $\vec{V}(M) = \overrightarrow{AM}$, où A est un point fixe du plan ($\mathcal{D}$ est quelconque).

b) $\mathcal{D}$ est le cercle de centre O et de rayon 1. $\vec{V}(M) = |OM|^2\,\vec{V}$ où $\vec{V}$ est un vecteur fixe et $|OM|$ la longueur du vecteur $\overrightarrow{OM}$.

Quatrièmeon

Soit deux champs de vecteurs ne s'annulant pas sur Γ. On suppose de plus qu' ue point M de Γ l'angle non orienté formé par les deux vecteurs $\vec{V}(M)$ et $\vec{W}$ strictement inférieur à π.

Démonts calculs que : $\alpha(V, \Gamma) = \alpha(W, \Gamma)$

Cinquièmeon

Soit application du plan dans lui-même deux fois différentiable. On notera pa $\Phi(M)$ l'image de M par Φ. On suppose que pour $|OM| \leqslant 1$ on a $|OM'| \leqslant$:

Démon il existe un point P tel que $|OP| \leqslant 1$ et $\Phi(P) = P$ (théorèmnt fixe). Peut-on généraliser ce résultat ?

[Pourer ce théorème on considèrera les champs $\vec{V}(M) = \overrightarrow{MM'}$ et $\vec{W}'M) = \overrightarrow{MO}$, et on prnur $\mathfrak{D}$ le cercle de centre 0 et de rayon 1.]

Sixième

On ut représentation complexe du plan : $z = x + iy$

On as vecteur $\vec{V}(M)$, son affixe $Z = X + iY$. On suppose que la fonction f, qui de champ de vecteurs ($Z = f(z)$), est une fonction holomorphe.

er que :

$$\alpha(V, \Gamma) = - i \int_{\Gamma} \frac{f'((z)}{f(z)} dz$$

Calcte intégrale.

Ce rest-il plus précis que l'assertion démontrée à la deuxième question ?

INDEX DES EXPRESSIONS

Le premier numéro donne le numéro du chapitre, etc.

.../...

INDICATIONS BIBLIOGRAPHIQUES

La plupart des bibliographies commencent par dire qu'il existe tant de livres sur le sujet envisagé qu'on ne saurait donner que quelques titres. Tel est ici en encore le cas.

On a cherché à évoquer des ouvrages qui ne soient pas réservés au mathématicien professionnel. Les commentaires devraient aider le lecteur. Les livres classiques cités ont des Bibliographies généralement importantes.

Sur les opérateurs linéaires en dimension finie,

D.P. LACHAT   Géométrie des opérateurs linéaires ; Dunod 1966

Ce petit livre couvre, en dimension finie, et sous forme intrinsèque, les bases algébriques de la théorie des opérateurs.

Naturellement les cours universitaires du premier cycle contiennent tous une description de l'algèbre linéaire. Citons un seul titre

J. DIXMIER   Cours de mathématiques du premier cycle : lère et 2ème année
Gauthier-Villars 1969, 2ème édition 1975.

Une introduction compacte à l'algèbre linéaire est fournie en annexe (p 367-380) au livre de

J. DIEUDONNE   Fondements de l'Analyse Moderne
Eléments d'analyse Tome 1
Gauthier-Villars Paris 1969

Une introduction très élémentaire et dirigée vers le calcul (à deux ou trois dimensions) est développé en vue de la géométrie dans

R. SEROUX   Géométrie vectorielle et algèbre linéaire
Nanta-Iremica, volume n° 9, 1976

Pour l'algèbre linéaire en vue de l'enseignement du second degré, il existe beaucoup de publications réunies par les IREM ou dans l'ambiance des IREM. Citons notamment :

A. REVUZ   Mathématiques pour l'élève-professeur Cours de l'A.P.M.

Passons maintenant à l'analyse fonctionnelle elle-même.

Essentiellement conçu pour des ingénieurs, on peut citer

J. BASS — Cours de mathématiques Tomes I, II, III
Masson 4ème édition 1968 et édition de 71 pour le dernier tome
Nouvelle édition 1977.

Les deux premiers tomes contiennent le matériel d'analyse classique enseigné dans les classes préparatoires aux Grandes Ecoles et dans certaines de ces Ecoles. Le troisième volume traite de la théorie de l'intégrale de Lebesgue, de la topologie et des opérateurs intégraux.

A un niveau plus universitaire, mais constituant des ouvrages remarquablement clairs et qui sont devenus des classiques, on doit citer en premier :

F. RIESZ et B. Sz NAGY — Leçons d'analyse fonctionnelle - 3ème édition 1955
Paris Gauthier-Villars -Budapest Akadémiai Kiado

Ce livre étudie les espaces de Hilbert et de Banach après une étude de l'intégrale de Lebesque rédigée en un style clair et accessible. Le chapitre sur les opérateurs compacts (appelés complètement continus) est particulièrement intéressant.
De nombreuses applications sont proposées. Malgré la difficulté du sujet, la lecture de la première partie de ce long livre est sans doute la meilleure façon de s'informer des techniques de l'analyse fonctionnelle. Un second ouvrage, plus moderne et très intéressant :

W. RUDIN — Analyse complexe et réelle
traduit de l'anglais, Masson, 1975

Les derniers chapitres de cet ouvrage sont nettement spécialisés, mais les premiers chapitres couvrent la théorie de l'intégration et de la différentiation d'une manière particulièrement concise et complète. Rudin a également écrit un livre d'Analyse Fonctionnelle, non traduit en français.

Donnons maintenant d'autres titres, qui correspondent à des ouvrages plus ou moins accessibles. On pourra également consulter la bibliographie de ces divers ouvrages, bibliographie plus spécialisée et que nous ne reproduisons pas.

I.E. SEGAL ; R.A. KUNZE Integrals and Operators
Mc Graw Hill 1968. Réédition récente chez Spinger-Werlag avec des modifications.

Ouvrage remarquable sur le sujet, mais orienté pour le mathématicien.

G. BACHMAN ; L. NARICI Functional Analysis
Academic Press New York 1966

Plus facile d'accès, mais encore orienté pour le lecteur mathématicien

A.C. ZAANEN Linear Analysis
North Holland Publ C.
Amsterdam 1960

Même remarque que précédemment

Pour des introductions élémentaires à l'analyse fonctionnelle

J. DHOMBRES Analyse Harmonique, Analyse Fonctionnelle Paris 71, Paris 72
Techniques de l'Ingénieur, Rubrique mathématique

C. GOFFMAN & G. PREDICK First course in functional analysis - Prentice Hall, Englewood Cliffs NJ 1965

Introduction concise et agréable sinon très profonde à l'analyse fonctionnelle abstraite.

R.E. EDWARDS Functional analysis. Theory and applications - Holt Rinehart and Winston Inc. 1965

C'est un gros livre touffu, riche d'applications diverses, mais exigeant un goût certain pour l'abstraction. De plus, les notations sont souvent lourdes et un lecteur, non mathématicien professionnel, risque d'être découragé. Il y a une très riche bibliographie de 31 pages.

N.I. AKHIEZ ; I.M. GLAZMAN Theory of linear operators in Hilbert Spaces
2 volumes Ungar Publ. Co New York 1961-1963

C'est un classique sur le sujet décrit par le titre, de même que l'ouvrage suivant très recommandable pour le lecteur désireux de s'orienter vers les équations aux dérivées partielles et différentielles comme applications

K. YOSIDA Functional Analysis
Academic Press Inc. New York 1965

Plus s>t le livre de

A.E. TAYL Introduction to functional analysis
John Wiley and Sons Inc New York 1958

Citonsnant des ouvrages plus accessibles à un lecteur non mathématicien

A.N. KOLI – Measure, Lebesgue integrals and Hilbert space
S.V. FOMI

Academic Press New York 1961
– Elements of the theory of functions and functional analysis
(2 volumes) Graybock 1961

J.P. BER' Mathématiques pour l'informatique Tome 1 Analyse fonctionnelle
A. Colin 1970

H.L. ROY' Real analysis Mac Millan 1963

Livreure assez facile

S. LANG Analysis I (1968)
II (1969)
Addison Wesley
Traduction en français : Inter Editions S.A. 1976

Natur nous excluons de cette bibliographie les ouvrages sur des sujets plus spé comme la topologie, la théorie de la mesure et de l'intégration etc.

Nousrons pas ici des ouvrages d'applications de l'analyse fonctionnelle à diverhes des mathématiques, mais citerons sur un domaine très particulier

J. D Some aspects of functional equations
Lecture Notes, Chulalonghorn University 1979 ; 361 pages

(Etuatique des fonctions f satisfaisant $f(x+y) = f(x) + f(y)$ pour cerleurs de x et y avec applications à la géométrie des espaces de ll'analyse fonctionnelle.)

TABLE DES MATIERES

Le chapitre 5 sera publié comme volume indépendant dans la collection Nanta Iremica sous le N° 7