REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère De L'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE KASDI MERBAH
OUARGLA

FACULTÉ DES MATHÉMATIQUES ET DES SCIENCES DE LA

MATIÈRE

Département DE MATHÉMATIQUES

MASTER

Spécialité : Mathématiques

Option : Analyse fonctionnelle

Présenté par : IMANE Dida

Thème

Le Lemme de Zorn et la Séparation des Espaces Vectoriels (Etude et applications)

La date : 10/09/2018
Devant le jury composé de :

B. TELLAB	Dr.Université Kasdi Merbah-Ouargla	Président
S. BADIDJA	Dr. Université Kasdi Merbah- Ouargla	Rapporteur
K. KALICHE	Dr.Université Kasdi Merbah- Ouargla	Examinateur
A. AMARA	Dr.Université Kasdi Merbah- Ouargla	Examinateur

Année universitre : 2017/2018

Table des matières

Remerciments

Tout d'abord, je remercie Dieu le bien car il a donné de la force et de la patience à long terme.
À la suite de ce travail, je tiens à remercier mon fidèle directeur , le Dr. SALIM Badijah
pour sa confiance et ses efforts considérables pour terminer ce travail.
Je voudrais remercier l'Université de Kasdi Merbah Ourgla(Algérie) car elle m'a contenu et
m'a donné les meilleurs professeurs pour mon enseignement et ma formation, et d'ici je remercie tous les
professeurs, en particulier les professeurs BADIJAH Salim, AMIR et ASSILA Mustafa , pour avoir laissé
une impression claire dans ma vie pratique
Enfin, je remercie l'étudiante en chimie organique Asma ABAID pour ses efforts dans ce travail

Imane DIDA

Dédicace

Je dédie ce modeste travail...
♠ a ma très chère mère *FATMA*
affable, honorable, aimable : Tu représentes pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement qui n'a pas cessé de m'encourager et du prier pour moi.
Je te dédie ce travail en témoignage de mon profond amour. Puisse DIEU, letout puissant, te préserver et t'accorder santé, longue vie et bonheur.
♠ a mon trés cher pére *M.SAYAH*
a l'homme de ma vie, mon exemple éternel ma source de joie et bonheur, rien au monde ne vaut les efforts fournis jour et nuit pour mon éducation être.
Ce travail est le fruit de tes innombrable sacrifies que tu as consentis pour mon éducation et ma formation, que DIEU te garde dans son vaste paradis.
♠ a mes trés chers fères
({*TOUFIK*, *MOUNIR*, *A.ELRAHMANE*, *TAHAR*, *ROUDOINE*, *A.ELATIFE*) et adorable sœur *ISRA*
les mots ne suffisent guère pour exprimer l'attachement, l'amour et l'affection que je porte pour vous, mes anges gardiens et mes fidèles accompagnants dans des moments les plus délicats de cette vie mystèrieuse , je vous dédie ce travail avec tous mes voeé et de réussite.
♠ a m grand mère *KHADRA*
que DIEU le protége.
♠ a mes chers ancles, mes chères tantes, cousins et cousines
que DIEU les protèges.
♠ a toute la famille de DIDA
♠ a mes belles sœurs :*ASMA*, *SAFA*, *FAIZA*, *ZINEB*, *LILA*.
Je vous souhaite tout ce qui est bon.

Introduction

•Lemme de Zorn (ou théorème de Zorn, ou parfois lemme de Kuratowski-Zorn) est un théorème de théorie des ensembles qui affirme que si un ensemble ordonné est tel que tout chaîne (sous-ensemble totalement ordonné) possède un majorant, alors il possède un élément maximal.
Lemme de Zorn n'adintérêt que pour les ensembles ordonnés qui ne sont pas totalement ordonnés. On dit qu'un ensemble ordonn'e est inductif lorsqu'il vérifie l'hypothèse du lemme de Zorn.
Il doit son nom au mathématicien Max Zorn(1906-1993) qui, dans un article de 1935, en donnait le premier un grand nombre d'applications, en redémontrant des résultats connus d'algèbre. Cependant Kazimierz Kuratowski en avait déjà publié une version en 1922, et plusieurs mathématiques, à commencer par Felix Hausdorff en 1907, avaient introduit des principes de maximalité proches du lemme de Zorn.
Des principes de maximalité plus ou moins proches du lemme de Zorn ont été découverts et redécouverts de nombreuses fois, sur une période 1930. Zorn lui-même ne revendiquait d'ailleurs pas la paternité du résultat. En 1928, Salomon Bochner, dans un article sur les surfaces de Riemann, démontre un lemme dont l'énoncé est celui, usuel aujourd'hui de lemme de Zorn pour un ensemble ordonné.
Le lemme de Zorn est équivalent à l'axiome du choix modulo les axiomes de la théorie des ensembles de Zermelo-Fraenkel.
Le lemme de Zorn permet d'utiliser l'axiome du choix sans recourir à la théorie des ordinaux (ou à celle des bons ordres via le théorème de Zermelo).
En effet, sous les hypothèses du lemme de Zorn, on peut obtenir un élément maximal par une définition par récurrence transfinie, la fonction itérée étant obtenue par axiome du choix.
Cependant, les constructions par récurrence transfinie sont parfois plus intuitives (quoique plus logues) et plus informatives.
Le lemme de Zorn a des applications aussi bien en topologie, comme le théorème de Tychonov, qu'en analyse fonctionnelle, comme le théorème de Hahn-Banach et théorème de Baire, ou en algèbre comme le théorème de Krull ou l'existence d'une clôture algébrique.
Le lemme de zorn peut être utilisé pour montrer que chaque graphe connexe a un arbre couvrant. L'ensemble de tous les sous-graphes qui sont des arbres est ordonné par inclusion et l'union d'une chaîne est une limite supérieurieure. Le lemme de Zorn dit qu'un arbre maximal doit exister, qui est un arbre couvrant puisque le graphe est connecté. Le lemme de Zorn n'est pas nécessaire pour les graphes finis.

•Il n'est pas indispensable pour un analyste de connaître la démonstration du le lemme de Zorn, par contre il essentiel de bien comprendre l'énoncé et de savoir l'utiliser. Le lemme de Zorn a de nombreuses et très importantes applications en Analyse, c'est un outil indispensable pour établir certains résultats d'existence.

•Ce mémoire se décompose de trois chapitres partagés de la manière suivante :

• Dans le premier chapitre, nous rappelons quelques définitions et présentons le lemme de Zorn et la démontion de l'équivalent entre le lemme de Zorn, l'axime de choix et l'axiome de Zermelo.

- : Dans le duxième chapitre, nous présentons les applications de ce lemme dans l'espace de Banach.

- Dans le troisième chapitre, nous séparons les espaces vectoriels par les hyperplans.

Chapitre 1

Préliminaires sur les théories des ensembles

1.1 Ensembles ordonnés

1.1.1 Relation d'ordre

Définition 1.1. (**relation d'ordre**)[**1**]. Soit E un ensemble. Une relation biniare $\preceq$ sur E est un sous-ensemble de E×E. On note $x \preceq y$ pour signifier que $(x, y) \in \mathbb{R}$ et $x \npreceq y$ pour signifier que $(x, y) \notin \mathbb{R}$.

Exemple 1.1.1. La relation $\leq$ est une relation d'ordre sur $\mathbb{R}$.

Définition 1.2. (**classsification**)[**1**]. Soit $\preceq$ une relation binaire sur E. On dit que $\preceq$ est

- réflexive quand $\forall x \in E, x \preceq x$;
- irréflexive quand $\forall x \in E, x \npreceq x$;
- symetrique quand $\forall x, y \in E, x \preceq y \Rightarrow y \preceq x$;
- antisymetréque quand $\forall x, y \in E, x \preceq y$ et $y \preceq x \Rightarrow x = y$;
- transitive quand $\forall x, y, z \in E, x \preceq y$ et $y \preceq z \Rightarrow x \preceq z$.

Définition 1.3. (**ordre**)[**1**]. une relation binaire est un ordre (ou une relation d'ordre) quand elle est réflexive, antisymétrique et transitive.

Exemple 1.1.2. Sur $E = \mathbb{R}$, la rolation $x \mathfrak{R} y \Leftrightarrow \sin x = \sin y$ est une rolation réflexive, symétrique et transitive.

1.1.2 Ensemble ordonné

Définition 1.4. [**1**]. Soit E un ensemble et $\preceq$ une relation d'ordre sur E. On dit que (E,$\preceq$) est un ensemble ordonné.

Exemple 1.1.3. $(\mathbb{R}, \leq), (\mathcal{P}(E), \subset)$ et $(\mathbb{N}^*, |)$ sont des ensembles ordonnés.

Définition 1.5. (**ordre strict**)[**1**]. Une relation binaire est un ordre strict (ou une relation d'ordre strict) quand elle irréflexive transitive.

Exemple 1.1.4. $<$ et $\nsubseteq$ sont les ordres stricts associés à $\leq$ et $\subset$.

Définition 1.6. (**ensemble strectement ordonné**) [**2**]. Soit E un ensemble et $\prec$ une relation d'ordre strict sur E. On dit que (E,$\prec$) est une ensemble strictement ordonné.

Remarque : Attention, les définition ci-dessus correspondent à l'ordre large mais pas à l'ordre strict :(R,$<$) où $<$ est l'ordre strict usule sur les réels n'est pas un ensemble ordonné (c'est un ensemble strictement ordonné).

1.1.3 Ensemble bien ordonné

Un ensemble ordonné est dit bien ordonné si tout sous-ensemble non vide de cet ensemble possède un plus petit élément.

Exemple 1.1.5. L'ensemble $(\mathbb{N}, \leq)$ des entiers naturels, muni de son ordre usuel, est bien ordonné.

1.2 Ensembles inductifs

Définition 1.7. (**ordre total**)[**2**]. Un ordre $\preceq$ sur E est dit total si deux éléments sont toujours comparables :$\forall x, y \in E$, $x \preceq y$ ou $y \preceq x$. Un ordre qui n'est pas total est dit partiel.

Exemple 1.2.1. L'ensemble des lettres d'un alphabet est totalement ordonné par un ordre alphabétique.

Définition 1.8. (**majorant , plus grand élément**)[**2**]. Soit [E,$\preceq$] un ensemble ordonné et F une partie non vide de E. On dit que $x \in E$ est un majorant de F si tout élément de F est plus petit que x pour $\preceq$:$\forall y \in F$, $y \preceq x$. Si la majorant de F est un élément de F on dit que c'est le plus grand élément de F.

Exemple 1.2.2. Pour l'intervalle $]0; 10[$, partie de l'ensemble $\mathbb{R}$ des nombres réels ordonné par l'ordre usuel $\leq$: 10 et 11 sont des majorants.

Définition 1.9. (**ensemble inductif**)[**2**]. Soit E un ensemble partiellement ordonné, E est dit inductif si toute partie de E non vide et totalement ordonné posséde un majorant.

Exemple 1.2.3. L'ensemble

$$\mathbf{P} = \left\{ h \,\middle|\, \begin{array}{l} h : D(h) \subset E \to \mathbb{R} \text{ avec } D(h) \text{ sous espace vectorial de } E.h \text{ linéaire} \\ G \subset D(h) \text{ prolonge } g \text{ et } h(x) \leq p(x)\ \forall x \in D(h) \end{array} \right\}$$

P est muni de rolation d'ordre

$$(h_1 \leqslant h_2) \Leftrightarrow (\mathbf{D}(h_1) \subset \mathbf{D}(h_2) \text{ et } h_2 \text{ prolonge } h_1).$$

P est inductif.

1.3 Lemme de Zorn

1.3.1 Énoncé du lemme

Définition 1.10. (**élément maximal**)[**2**]. Soit (E,$\preceq$) un ensemble ordonné et F une partie nono vide de E. Un élément $x \in F$ est un élément maximal de F quand aucun élément de F n'est strictement plus grand, pour $\preceq$, que x : $\forall y \in F$, $x \preceq y \Rightarrow y = x$.

Lemme 1.1.

(Lemme de Zorn)[3] : Tout ensemble non vide et inductif possède un élément maximal.

1.3.2 Démonstration de lemme de Zorn

Démontrons maintenant le lemmede Zorn, à cette fin nous intrduisons un peu de vocabulaire :

- Siot (E,$\preceq$) un ensemble ordonné. on appelle chaîne de E un sous-ensemble totalement ordonné de (E,$\preceq$). en particulier l'ensemble vide est une chaîne de E. Tout singleton est une chaîne de E.

- Nous dirons qu'une chaîne C $\subset$ (E,$\preceq$) est ultime si elle n'admet aucun majorant strict(c'est-à-dire, si il n'y a aucun $m \in$ (E,$\preceq$) pour lequel $\forall c \in C$; $c \prec x$), et continuable dans le cas contraire.
- Pour C $\subset$(E,$\preceq$) une chaîne, on dit qu' un sous-ensemble strict I$\subset$C en est un début si tous les éléments de C $\setminus$ I sont supériurs à tous les éléments de I. Si en outre I $\neq$ C, ce début est dit strict .

- Une chaîne C $\subset$ (E,$\preceq$) est dite bonne quand, pour tout début strict I de C, m(I) appartient à C et le plus petit élément de C $\setminus$ I .

- Une chaîne C $\subset$ (E,$\preceq$) est dite bonne quand, pour tout début strict I de C, m(I) appartient à C et est le plus petit élément de C $\setminus$ I.
- Un élélemnt m de (E,$\preceq$) est un élément maximal s'il ne possède aucun majorant strict : $\forall x \in E$ $(m \preceq x \Rightarrow m = x)$.

- Une chaîne qui n'est strictement contenue dans aucune chaîne est appelée chaîne maximale. Si une îtelle chaîne posséde un majorant, celui-ci est alors un élément maximal de l'ensemble E.

Lemme 1.2. [**1**] Tout un ensemble ordonné (E,$\preceq$) admet une chaîne ultime.

Démonstration. Pour toute chaîne continuable C$\subset$ $(E, \preceq)$, choisissons une fois pour toutes un majorant strict de C, que notons $m(C)$.
Nous allons montrer maintenant que si C_1 et C_2 sont deux bonnes chaînes de(E,$\preceq$),alors l'une est un début de l'autre.
Considérons les $x \in C_1 \cap C_2$ tels que $\{y \in C_1 \cup C_2; y \prec x\} \subset C_1 \cap C_2$;notons C^* l'ensemble qu'ils forment. La définition de C^*assure qu'il s'agit d'un débt à la fois des chaînes C^1 et C^2 ; par conséquent, si C^* coïncide avec C^1 ou C^2, la chaîne en qustion est un début de l'autre. Il ne reste alors plus qu' à montrer qu'il est impossible que C^* serait un débnt strict de C^1 et de C^2, etdonc, puisque ces chaînes sont bonnes, $m(C^*)$ serait dans chacune des C_i, et serait le plus petit élément de chaque $C_i \setminus C^*$, une contradiction.
Considérons alors l'ensemble $\bar{\mathrm{C}}$ $\subset$ (E,$\preceq$) défini comme la réunion de toutes les bonnes chaînes de (E,$\preceq$), i.e $\bar{\mathrm{C}}$ est l'ensemble des points de (E,$\preceq$) qui appartiennent à une bonne chaîne au moins. Comme, de deux bonnes chaînes, l'une est le début de l'autre, on en déduit que $\bar{\mathrm{C}}$ est une chaîne.
$\star$Montrons que la chaîne $\bar{\mathrm{C}}$ est bonne.
Soit I un début strict de $\bar{\mathrm{C}}$. Soit $j \in \bar{\mathrm{C}} \setminus I$; par définition de $\bar{\mathrm{C}}$, il existe une bonne chaîne de (E,$\preceq$) contenant j : considérons une telle chaîne J. On observe alors que I$\subsetneq$J : en effet, pour tout $h \in I$, on sait que h appartiant à au moins une bonne chaîne H ; et comme H et J deux bonnes chaînes, l'une est le début de l'autre, et donc dans tous les cas $h \in J$. Puisque I $\subsetneq$ J, que I est un début de $\bar{\mathrm{C}}$, etque J $\subset\bar{\mathrm{C}}$ (par définition de $\bar{\mathrm{C}}$), on voit que I est un début strict de J ; et puisque J est une bonne chaîne, il s'ensuit que $m(I) \in J \subset \bar{\mathrm{C}}$, ce qui prouve bien que la chaîne $\bar{\mathrm{C}}$ est bonne.
Pour finir, $\bar{\mathrm{C}}$ ne peut être continuable, car sinon $\bar{\mathrm{C}} \cup \{m(\bar{\mathrm{C}})\}$ serait une bonne chaîne, et donc en revenant à la définition de $\bar{\mathrm{C}}$, $m(\bar{\mathrm{C}})$ appartendrait à $\bar{\mathrm{C}}$, ce quiest absurde. La chaîne $\bar{\mathrm{C}}$ que nous avons construite est donc ultime. $\square$

1.4 Axiome du choix

1.4.1 Axiome(choix)

.

Définition 1.11. [**1**] Soit X un ensemble d'ensembles tous non vides. Alors il existe une fonction f, définie sur X et à valeur dans $\cup_{x\in X}x$ telle que pour tout $x \in X, f(x) \in x$ (une telle fonction sera appelée fonction de choix)

1.4.2 les premiers axiomes

On ne définit pas la notion d'ensemble. On a une collection d'ensembles et on peut dire qu' un ensemble appartient à un autre. On suppose enfin qu' il existe au moins un ensemble.

1. **Axiome d'extensionnalité**

   Axiome

   $\forall a, b$ ensembles, $(a = b) \Leftrightarrow (\forall x, x \in a$ ssi $x \in b)$

2. **Axiome de sélection**

   Axiome

   Soit $P(x)$ une propriété dépendant de x. Alors $\forall a, \exists b, x \in b \Leftrightarrow (x \in a$ et$P(x))$. On a outre b=$\{x \in a, P(x)\}$.
   Remarque : On ne peut pas définir $\{x/P(x)\}$. En effet, ce serait contradictoire : si $P(x)$:$x \notin x$ et b=$\{x/P(x)\}$ alors b $\in$ b ssi b $\notin$ b absurde!

3. **Axiome de la paire**

   Axiome

   $\forall a, b, \exists c, [x \in c \Leftrightarrow (x = a$ ou $x = b)]$. c est noté $\{a, b\}$ si $a \neq b$ et $x = \{a\}$ si $a = b$
   Exemple : $\{\emptyset\}$ existe, puis $\{\emptyset, \{\emptyset\}\}$

4. **Axiome de réunion**

   Axiome

   $\forall a, \exists b, [x \in b \Leftrightarrow (\exists c \in a, x \in c)]$. b est alors noté $\bigcup\limits_{c\in a} c$
   . *Exemple* : $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ existe : c'est $\{\emptyset\{\emptyset\}\} \bigcup \{\{\emptyset, \{\emptyset\}\}\}$

5. **Axiome de l'ensemble des partie**

   Axiome

   $\forall a, \exists b, (c \in b \Leftrightarrow c \subset a)$.On note $b = \mathcal{P}(a)$

6. **Axiome de fondement**

   Axiome

   Si $a \neq \emptyset, \exists x \in a, x \cap a = \emptyset$
   Résultats :
   - Si $a = \{x\}$ cela donne : $\exists z \in a, z \cap a = \emptyset$. Or,$z \in a = \{x\} \Rightarrow z = x$ et donc de $z \cap a = \emptyset$ on déduit $x \cap \{x\} = \emptyset$.Donc si$x \in x$,alors comme $x \in \{x\}$ on déduit $x \in (x \cap \{x\})$ ce qui est absurde ! Donc $x \in x$ est **impossible**.
   - Si $y \in x$ et $x \in y$, posons $a = \{x, y\}$: donc $\exists z \in a, z \cap a = \emptyset$. Si $z = x$, alors $z \cap a = x \cap \{x, y\} = \emptyset$ ce qui est absurde car $y \in (x \cap \{x, y\}$!le cas $z = y$ est tut aussi absurde : donc $x \in y$ et $y \in x$ est **impossible**.

7. **Axiome de l'infini**

   Rappel : Si x est un ensemble, on définit $x^+ = x \cup \{x\}$

   Axiome

   $\exists a, [\emptyset \in a$ et $x \in a \Rightarrow x^+ \in a]$.

1.5 Équivalence entre l'axiome du choix et le lemme de Zorn

1.5.1 lemme de Zorn implique l'axiome du choix

On suppose que tout ensemble inductif posséde élément maximal. Soit C ensemble non vide d'ensemble non vides. On pose :

$$X = \{(A, c)/A \subset C et F : A \to \bigcup_{B \in C} B, \forall a \in A, F(a) \in a\}$$

et on va montrer que X est inductif, pour un certain ordre que l'on expliciterz.

- $C \neq \emptyset$ donc $\exists a \in C$; on a de plus $a \neq \emptyset$ donc $\exists x \in a$. On pose $A = \{a\}$ et

$$F: \begin{array}{c} A \to \bigcup_{B \in C} B \\ a \mapsto x \end{array}$$

  . On a $(A, F) \in X$ donc $X \neq \emptyset$.
- Ordre sur X : $[(A_1, F_1) \preceq (A_2, F_2)]$ ssi $[A_1 \subset A_2$ et $F_2|_{A_1} = F_1]$. Soit Y une partie totalement ordonnée de X. On pose $B = \bigcup_{A/\exists F,(A,F) \in Y} A$ et $D : B \to \bigcup_{A \in C} A$ définie ainsi : si $a \in B$, alors $\exists (A, F) \in Y, a \in A$; on pose alors $D(a) = f(a)$. Cette définition est indépendante de l'ensemble A choisi car Y est totalement ordonné et la définition de l'ordre nous garanti la définition de D. Finalement, (B, D) **majore Y**.

X est donc inductif, et dáprés le lemme de Zorn X admet un élément maximal (A, F). Si $A = C$, c'est une fonction de choix sur C. Supposons $A \neq C$: alors $\exists b \in C \setminus A$. ã $\neq \emptyset$ donc $\exists x \in$ ã. Onpose Ã$=\{A \cup b\}$, puis

F̃ :Ã$\to \bigcup_{B \in C} B$ avec F̃$= \begin{cases} F(a) & \text{si } a \in A \\ x & \text{si } a = b \end{cases}$ .F̃ est un prolongement strict de F : c'est une contradiction avec la maximalité de (A, F) !

Finalement, on a bien trouvé unefonction de choix sur C, qui est un ensemble quelconque non vide d'ensembles non vides : c'est l'axiome du choix.

1.5.2 L'axiome du choix implique le lemme de Zorn

Soit $(X, \preceq)$ un ensemble inductif. On pose :

$$\Theta = \{A \subset X/A \text{ } a \text{ } au \text{ } moins \text{ } un \text{ } majorant \text{ } strict\}$$

$\mathcal{P}(X) \backslash \emptyset$ est un ensemble non vide dr parties non vides. L'axiome du choix garantit donc lÃ©existence d'une fonction de choix

$$\mathcal{P}(X) \setminus \emptyset \to X, A \mapsto F(A) \in A$$

Pour $A \in \Theta$ on note $m(A) = F(\{\text{ majorants stricts de } A\})$.
On va d'abord faire une premiére approche, en utilisant une récurrence. $\emptyset \subset X, X \neq \emptyset$ donc $\forall x \in X, \{x\}$ est un majorant strict de $\emptyset$. On a donc $\emptyset \in \Theta$ et on pose alors $x_0 = m(\Theta)$. On construit alors la suit $(x_n)_{n \in \mathbb{N}} \in X^{\mathbb{N}}$ telle que $\forall k \in \mathbb{N}, x_0 < \cdots < x_k$ et $x_{k+1} = m(\{x_0, \ldots, x_k\})$ (à moins que le processus ne s'arrête, auquel cas on autrait trouvé un élément maximal).
$E = \{x_n/n \in \mathbb{N}\}$ est (par construction) une partie de X totalement ordonnée. X est inductif donc E posséde un majorant $M \in X$ avec $M \notin E$. Par conséquent, E a au moins un majorant strict *ie* $E \in \Theta$; on note $x_\infty = m(E)$.
Onvoit bien qu'à faire des récurrences on va jamais finir ! On va donc attaquer le probléme sousun autre angle : d'abord, quelques définition s.

Définition 1.12. (**Segment**) Soient S et A deux parties de X. Ondit que S est un segment de A ssi
- $S \subset A$
- $\forall (x, y) \in A \times S, (x, \preceq y \Rightarrow x \in S)$.

Définition 1.13. (**Bon ensemble**) $B \subset X$ est un bon ensemble ssi
- B est totalement ordonné
- Pour tout segment S de B, si $S \neq B$ alors S a des majorants stricts dans B et $m(S)$ est plus petit majorant strict de S dans B.

Exemples :
- $\{x_0, \ldots, x_k\}$ est un bon ensemble, dont les segments sont les $\{x_0, \ldots, x_k\}$ pour $0 \leqslant l \leqslant k$;
- $\{x_k/k \in \mathbb{N}\}$ est un bon ensemble, dont les segment sont les $\{x_0, \ldots, x_k\}$ pour $k \in \mathbb{N}$
- $\{x_k/k \in \mathbb{N}\} \cup \{l\}$ est un bon ensemble, dont les segments sot les $\{x_0, \ldots, x_k\}$ pour $k \in \mathbb{N} \cup \{\infty\}$ (avec $x_\infty = l$)

On pose $\mathbf{U} = \cup_{B \quad bon \quad ensemble} B$. On va montrer dans les suivants que $\mathbf{U}$ est un bon ensemble. On aura alors :
- $\mathbf{U}$ est totalement ordonné (car c'est bon ensemble), il admet donc un majorant(car $X \supset \mathbf{U}$ est inductif) ;
- Si tous les majorants sont sont dans $\mathbf{U}$, alors il n'y en a qu' un (car $\mathbf{U}$ est totalement ordonné), m : si $X \ni n > m$ alors n est un majorant strict de$\mathbf{U}$ ce qui est absurde car $n \notin \mathbf{U}$, donc m est un élément maximal de X ;
- Sinon, $\mathbf{U}$ a au moins un majorant strict $m \in X$ $\mathbf{U}$: $\mathbf{U} \cup \{m\}$ est alors un bon ensemble qui contient strictement $\mathbf{U}$ ce qui est absurde car c'est $\mathbf{U}$ le plus gros.

Donc X admet un élément maximal ce qui démontre le lemme de Zorn !

1.6 Équivalence entre lemme de Zorn et axiome de Zermelo

En mathémathiques, l'axiome de Zermelo, appelé aussi théorème du bon ordre, est un résultat de théorie des ensembles, démontré en 1904 par *Ernst Zermelo*, qui affirme :

1.6.1 Axiome de Zermelo

Tout ensemble peut être muni d'une structure de bon ordre, c'est-à-dire d'un ordre tel que toute partie non vide admette un plus petite élément[**1**].

1. **Lemme de Zorn implique l'axiome de Zermelo**

   Soit E un ensemble, soit M l'ensemble des rolations de bon ordre sur une partie de E. M luimême peut être muni d'un ordore partiel : on dit qu' un bon ordre **01** est inférieur ou égal à un bon ordre **02** si **01** est un segment initial de **02**. On vérifie ensuite que M muni de cette relation est un ensemble inductif. Toute châne de M admet un majorant (qui est même une borne supérieure) : la rolation dont le graphe est la réunion des graphes des ordres de la châne. On vérifie que cette relation est bien une relation de bon ordre (on exploite le fait que la châne est ordonnée par segment initial). Donc M admet un élément maximal. Un tel élément maximal est alors un bon ordre sur tout E(on pourrait sinon le prolonger en un bon ordre successeur, ce qui contredirait la maximalité)

2. **L'axiome de Zermelo implique l'axiome du choix**

   Soient E un ensemble bien ordonné, et $P(E)$ l'ensemble de ses parties. Alors, on définit une fonction dechoix sur $P(E)\backslash\{\emptyset\}$ en associant, à chaque partie non vide de E, son plus petit élément(l'existence d'une telle fonction est un des énoncés possibles de l'axiome du choix)
   L'axiome de Zermelo implique donc lemme de Zorn(une preuve directe, voir **1.5.2**)

Chapitre 2

Application de lemme de Zorn en topologie

2.1 Théorème de Tychonov

On commence par une définition fondamental.

Définition 2.1. (Espase topologique)[**4**]. Soit X un ensemble et désibnons par $\mathcal{P}(X)$ l'ensemble de ses parties. Une topologie sur X est un sous-ensemble $\tau \subset \mathcal{P}(X)$ qui vérifie :

1. $\emptyset$, $X \in \tau$
2. Si $\{U_i\}_{i\in I} \subset \tau$, alors $\bigcup_{i\in I} U_i \in \tau$
3. Si $U_1, ..., U_N \in \tau$, alors $\bigcap_{j=1}^{N} U_j \in \tau$

Exemple 2.1.1. X=0,1, $\tau = \{\emptyset, \{0,1\}, \{0\}\}$

Définition 2.2. (Espace métriques)[**4**]. Soit X un ensemble ; une distance ou métrique sur X est une application $d : X \times X \longrightarrow \mathbb{R}$ qui vérifie les propriétés suivantes :

1. $\forall x, y \in$, $d(x,y) \geqslant 0$ et $d(x,y) = 0$ si et seulement si $x = y$.
2. $d(x,y) = d(y,x)$ (symétrie)
3. $\forall x, y, z \in X$, $d(x,y) \leqslant d(x,z) + d(z,y)$ (inégalité du triangle)

On dit que le couple (X, d) est un espace métrique.

Exemple 2.1.2. La distance euclidienne sur $\mathbb{R}^n$ en fait un espace métrique.

Définition 2.3. (Voisinages)[**4**]. Soit (X, τ) un espace topologique, $x \in X$. On dit que $V \subset X$ est un voisinage de x s'il existe un ouvert U_x tel que $x \in U_x \subset V$. On not $\mathcal{U}_x$ l'ensemble des voisinages de x

Exemple 2.1.3.
- $\mathbb{R}$ est un voisinage de chacun de ses point.
- Plus généralement, tout ouvert, donc en particulier tout intervalle ouvert est un voisinage de chaucn de ses points.

Définition 2.4. (Espace topologiques séparés)[**4**]. Soit X un espace topologique ; on dit qu'il est séparé si pour tout $x, y \in X$, $x \neq y$, il existe deux ouverts U_x et U_y tel que $x \in U_x$, $y \in U_y$ et $U_x \cap U_y = \emptyset$.

On dit que les ouverts U_x et U_y séparent les points x et y.

Exemple 2.1.4. Tout espace métrique (X, d) est séparé : si $x, y \in X$ et $x \neq y$, alors $r = d(x,y) > 0$ et on prondre $U_x = B(x, \frac{r}{2})$, $U_y = B(y, \frac{y}{2})$.

2.1.1 Filtres

Dans cette partie, E et F désignent deux ensembles ;f est une application de E à valeurs dans F.

Définition 2.5. (**Filtre**)[4] Soit $\mathcal{F} \in \mathcal{P}(E)$. $\mathcal{F}$ est un **filtre** si :

1. $\theta \notin \mathcal{F}$;
2. $\forall B \in \mathcal{F}, \forall A \in E, [B \subset A \Longrightarrow A \in \mathbf{F}]$;
3. $\forall A, B \in \mathcal{F}, A \cap B \in \mathcal{F}$.

Exemples 2.1. • Si $\mathcal{F}$ est filtre et si $A \in \mathcal{F}$ alors $A^c \notin \mathcal{F}$; ainis, $\mathcal{P}(E) \backslash \{\theta\}$ n'est pas un filtre. Plus généralement, si $\mathcal{F}$ est un filtre, alors $\mathcal{F}$ ne contient pas deux ensembles disjoints.

- Si (E, τ) est un espace topologique et si $a \in E$, alors $\mathcal{V}_a$ (ensemble des voisinages de a) est un filtre. Plus généralement, si $\theta \subsetneq A \subset E$ alors $\mathcal{V}_a$ est un filtre.
- Cas $E = \mathbb{N}$: l'ensemble $\mathcal{F} = \{A \subset \mathbb{N} / A^c \text{est fini}\}$ est un filtre ; c'est le filtre de `Frécht`.
- Soit E un ensemble non vide et X_0 un sous-ensemble non vide de E. L'ensemble

$$\mathcal{F}_{X_0} = \{X \in P(E) \mid X_0 \subseteq X\}$$

est un filtre, qu'on dit être un filtre principale.

Définition 2.6. (**Base de filtre**)[4] Une partie $\mathcal{B}$ de E est une base de filtre si :

1. $\theta \notin \mathcal{B}$;
2. $\forall A, B \in \mathcal{B}, \exists C \in \mathcal{B}, C \subset A \cap B$.

Exemples 2.2. 1. Les bases de voisinage sont des bases de filtre.

2. plus généralement, soit E un espace métrique et x un point de E, l'ensemble des boules ouvertes ou fermées de centre x et de rayon $r > 0$ est une base du filtre des voisinage de x.

Définition 2.7. (Filtre engendré)[4] Soit $\mathcal{F}$ une base de filtre. Le plus petit filtre contenant $\mathcal{B}$ est donné par $\mathcal{F} = \{A \subset E / \exists B \in \mathcal{B}, B \subset A\}$ on dit que $\mathcal{B}$ est la base de $\mathcal{F}$.

Exemples 2.3. 1. Si $a \in (E, \tau)$, alors une base de voisinage de a est base de $\mathcal{V}$

2. Une base du filtre de Fréchet est $\mathcal{B} = \{\{p \in \mathbb{N} / p \geqslant n\} / n \in \mathbb{N}\}$.

Définition 2.8. Soint $\mathcal{F}_1, \mathcal{F}_2$[4] deux filtres sur E. On dit que $\mathcal{F}_1$ est **plus fin** que $\mathcal{F}_2$ si $\mathcal{F}_1 \supset \mathcal{F}_2$.

2.1.2 Ultrafiltre

Définition 2.9. (**Ultrafiltre**)[4]. Soit $\mathcal{F}$ un filtre.On dit que $\mathcal{F}$ est un **ultrafiltre** s'il n'y a pas de filtre plus fin, *i.e* pour tout filtre $\mathcal{G}, [\mathcal{G} \supset \mathcal{F} \Longrightarrow \mathcal{G} = \mathcal{F}]$.

Définition 2.10. (**Filtre image**)[4]. $f(\mathcal{B})$ s'appelle l'image de $\mathcal{B}$ par f.

Définition 2.11. (**Convergence d'un filtre**)[4]. Soit $\mathcal{F}$ un filtre sur E, soit $a \in E$. On dit que $\mathcal{F}$ converge vers a si $\mathcal{F} \supset \mathcal{V}_a$. On note ce fait $\mathcal{F} \longrightarrow a$, et on dit alors que $\mathcal{F}$ est convergent et a est sa **limite**

Théorème 2.1. Soit $\mathcal{F}$ un filtre sur E ; alors il existe un ultrafiltre $\mathcal{U}$ sur E tel que $\mathcal{U} \supset \mathcal{F}$.

Démonstration. On va utiliser le lemme de Zorn ; soit $\mathcal{F}$ un filtre sur E. On pose $X := \{\mathcal{G} / \mathcal{G}$ est un filtre contenant $\mathcal{F}\}$ muni de l'inclusion. Démontrons que X est inducatif.

- X est non vide car $\mathcal{F} \in X$.

- Soit $(\mathcal{G}_i)_i \in I$ une famille totalement ordonnée de X. On pose $\mathcal{G} := \{\cup_{i\in I}\mathcal{G}_i/\mathcal{G}$ contient toutes les $\mathcal{G}_i$ et donc aussi $\mathcal{F}$. Il reste à montrer que c'est un filtre
  1. Soit $A \in \mathcal{G}$ soit $B \subset E$ contenant A.Il existe $i \in I$ tel que $A \in \mathcal{G}_i$ qui est un filtre donc$B \in \mathcal{G}_i$ donc $B \in \mathcal{G}$.
  2. $\forall i \in I, \theta \notin \mathcal{G}_i$ donc $\theta \notin \mathcal{G}$.
  3. Soient A et B dans $\mathcal{G}$. Il existe $i, j \in I$ telque $A \in \mathcal{G}_i, B \in \mathcal{G}_j$; la famille $(\mathcal{G}_k)_{k\in I}$ étant totalement ordonnée, on peut suppouser que $\mathcal{G}_j \subset \mathcal{G}_i$. Ainsi, A et B sont dans le filtre $\mathcal{G}_i$ donc $A \cap B \in \mathcal{G}_i \subset \mathcal{G}$.

  On a démontré que $\mathcal{G}$ est un filtre ; c'est un majorant de la famille $(\mathcal{G}_i)_{i\in I}$.
  Ainsi, X est inducatif : il contient donc un élément maximal, qui est donc un ultrafiltre.

□

Corollaire 2.1. (E, τ) séparé. Alors (E, τ) est **compact** <u>*ssi*</u> tout **ultrafiltre** est **convergeant**.

Théorème 2.2. [**4**]Soit $\mathcal{B}$ une base de filtre sur E telle que le filtre engendré soit un ultrafiltre, soit $f : E \longrightarrow F$. Alors $f(\mathcal{B})$ **engdendre un ultrafiltre**.

2.1.3 Théorème de Tychonov

Soit $(E_i, \tau_i)_i \in I$ une famille d'espaces topoloiques (non vides) ; on considère $E := \Pi_{i\in I}E_i$ (qui est non vide,l'axiome du choix) muni dela topologie produit τ. On rappelle ques les éléments de τ (les ouverts de E) sont engendrés par les $\Pi_{j\in J}A_i \times \Pi_{i\in I\setminus J}E_i$ où $J_{fini} \subset I$ et $A_j \in_j$. De plus, $A \subset E$ est voisinage de $a \in E$ <u>*ssi*</u> $\exists J_{fini} \subset I, \forall j \in J, \exists A_j \in \mathcal{V}_{E_j}(a_j), \Pi_{j\in J}A_i \times \Pi_{i\in I\setminus J}E_i \subset A$.

Théorème 2.3. [**5**]Avec les notations précédentes, on suppose que tous les (E_i, τ_i) sont compacts. Alors (E, τ) est **compact**.

Un petit lemme avant de se lancer dans la démonstration.

Lemme 2.1. [**5**]Soit $\mathcal{F}$ un filtre sur E, $p_i : E \longrightarrow E_i$ les applications coordonnées et $a = (a_i)_{i\in I} \in E$. Alors $\mathcal{F}$ converge vers a <u>*ssi*</u> $\forall i$, le filtre engendré par $p_i(\mathcal{F})$ converge vers a_i.

On passe maintenaint à la démonstration du théoréme deTychonov.

Démonstration. Montrons d'abord que (E, τ) est séparé. Soient $a \neq b \in E$, et soit i_0 tel que $a_{i_0} \neq b_{i_0}$. L'espace (E_{i_0}, τ_{i_0}) est compact donc séparé (par définition) donc $\exists \mathbf{U}_{i_0} \in \mathcal{V}_{E_{i_0}}(a_{i_0}), \exists \mathbf{V}_{i_0} \in \mathcal{V}_{E_{i_0}}(b_{i_0}), \mathbf{U}_{i_0} \cap \mathbf{V}_{i_0} = \emptyset$. On pose maintenant $\mathbf{U} := \mathbf{U}_{i_0} \times \Pi_{i\neq i_0}E_i$ et $\mathbf{V} := \mathbf{V}_{i_0} \times \Pi_{i\neq i_0}E_i$; on a $\mathbf{U} \in \mathcal{V}_E(a)$, $\mathbf{V} \in \mathcal{V}_E(b)$ et $\mathbf{U} \cap \mathbf{V} = \emptyset$.
on reprend nos applications coordonnées

$$p_i : E \longrightarrow E_i, \qquad a = (a_i)_{i\in I} \longmapsto a_i$$

$\mathcal{U}$ un ultrafiltre sur E : comme (E, τ) est séparé, il suffit de montrer que $\mathcal{U}$ converge et on aura montré que (E, τ) est compact (corollaire**2.1**). $\mathcal{U}$ est un ultrafiltre donc d'aprés le théorème **2.2**, $p_i(\mathcal{U})$ engendre un ultrafiltre. L'espace (E_i, τ_i) est compact donc cet ultrafiltre est convergeant (encore une fois le corollaire **2.1**) :on note $a_i \in E_i$ sa limite. Ainsi, $\forall i \in I$ le filtre engendré par $p_i(\mathcal{U})$ converge vers a_i donc par le lemme **2.1**, le filtre $\mathcal{U}$ converge vers $a = (a_i)_{i\in I}$ ce qui conclut la démonstration. □

Application : Soit E un ensemble, soit $(M_x)_{x\in E}$ une famille de réels positifs et $K := \{f : E \longrightarrow \mathbb{R}/\forall \in E, \mid f(x) \mid \leqslant M_x\} \subset \mathbb{R}^E = \Pi_{x\in E}\mathbb{R}$. Alors $\mathbf{K}$ est compact pour la topologie produit (*i.e.* la topologie de la convergence simple). En effet, on a en fait $\mathbf{K} = \Pi_{x\in E}[-M_x, M_x]$ qui est compact car chaque $[-M_x, M_x]$ est compact (pour la topologie de la usuelle de $\mathbb{R}$).

2.2 Théorème de Hahn-Banach

2.2.1 Hyperplans

Définition 2.12. Un **hyperplan**[7] (affine) est un ensemble de la forme

$$\mathbf{H} = \{x \in E; f(x) = \alpha\}$$

où f est un forme linéaire sur E, non identiquement nulle et $\alpha \in \mathbb{R}$. On dit que $\mathbf{H}$ est l'hyperplan d'équation $[f = \alpha]$.

Corollaire 2.2. (Equation d'un hyperplan).

1. Soit $\mathfrak{B} = (e_1, ..., e_n)$ une base de E, et $a_1, ..., a_n \in \mathbb{K}$ non tous nuls, alors l'ensemble des x appartenant à E vérifiant :
$$\sum_{i=1}^{n} a_i x_i = 0 \qquad (*)$$
relativementà la base est un hyperplan.
2. Tout hyperplan de E admet une équation de la forme$(*)$ qui est unique à constante multiplication.

Proposition 2.1. *H est un hyperplan <u>ssi</u> $\dim E/H = 1$.*

Démonstration. On note Π la projiction canonique de E sur E/H ; Π est surjective. On pose $A = \{F \subset E/F$ est un sous-espace vectoriel de E contenant H $\}$. Si $F \in A$, alors $\Pi(F)$ est sous-espace vectoriel de E/H ; inversement, si G est un sous-espace vectoriel de E/H, alors $\Pi^{-1}(F) \in A$.
On définit lapplication

$$\tilde{\Pi} : A \longrightarrow \tilde{A} := \{\text{sev de E/H}\}$$
$$F \longmapsto \tilde{\Pi}(F) = \Pi(F).$$

Il est bon ici de faire attention :$\tilde{\Pi}$ prend en argument des éléments de A (qui sont des sous-espace vectoriels de E), contrairement à Π qui prend en argument des vecteurs de E. Ainsi, $\tilde{\Pi}$ n'est plus linéaire, mais $\tilde{\Pi}$ hérite quand même dela surjectivé de Π. On va maintenant montrer que $\tilde{\Pi}$ est surjective.
Soient $F_1, F_2 \in A$, tels que $\tilde{\Pi}(F_1) = \tilde{\Pi}(F_2)$. On a donc :

$$\forall a \in F_1, \exists y \in F_2, \Pi(x) = \Pi(y).$$

Comme $\Pi(x) = \Pi(y) \Leftrightarrow x - y \in H$ et que H est contenu dans F_1, F_2, on déduit que pour tout $x \in F_1, x$ est aussi dansF_2. Donc $F_1 \subset F_2$; par symétrie, on a l'inclusion réciproque et par conséquent l'égalité. Finalement, $\tilde{\Pi}$ est injective ; on sait qu' elle est surjective, c'est donc une bijection. On a donc une bijective entre $A = \{$ sous-espace vectoriel de E contenant H $\}$ et
$\tilde{A} = \{$ sous-espace vectoriel de $E/H\}$.
H est donc un hyperplan <u>ssi</u> card$A = 2$ <u>ssi</u> card$\tilde{A} = 2$ <u>ssi</u> $\dim E/H = 1$. □

Proposition 2.2. *On a les équivalences suivantes :*

1. H est un hyperplan
2. $\exists e \notin H, E = H \oplus \mathbb{R}e$
3. $\forall e \notin H, E = H \oplus \mathbb{R}e$
4. $\exists \varphi \in E^* \ \{0\}, H = \mathbf{Ker}\varphi$

De plus, si $H = \mathbf{Ker}\varphi_1 = \mathbf{Ker}\varphi_2$ alors $\exists \lambda \in \mathbb{R}^*, \varphi_1 = \lambda\varphi_2$.

Démonstration. (1) $\Rightarrow$ (4) On suppose que H est un hyperplan. D'aprés la proposition précédent, dimE/H=1 donc on a un isomorphisme $\theta : E/H \longrightarrow \mathbb{R}$. En notont Π la projection canonique $E \longrightarrow E/H$ et en posant $\varphi = \theta \circ \Pi$, on a bien $\varphi \in E^*$ $\{0\}$; deplus, $x \in \mathbf{Ker}\varphi \Leftrightarrow \theta(\Pi(x)) = 0 \Leftrightarrow \Pi(x) = 0 \Leftrightarrow x \in H$ donc $\mathbf{Ker}\varphi = H$.
(4) $\Rightarrow$ (3) On suppose que $H = \mathbf{Ker}\varphi$ pour $\varphi \in E^*$ $\{0\}$. Soit $e \notin H$: on a donc $\varphi(e) \neq 0$. Soit $x \in E$; on pose $\lambda = \dfrac{\varphi(x)}{\varphi(e)}$. On a $x = \underbrace{x - \lambda e}_{\in H} + \underbrace{\lambda e}_{\in \mathbf{R}e}$ donc $E = H + \mathbf{R}e$.
Si $x \in H \cap \mathbf{R}e$, alors on peut écrire $x = \lambda e$. $x \in H$ donc $\varphi(x) = 0 = \lambda\varphi(e)$ donc $\lambda = 0$ (car $\varphi(e) \neq 0$), donc $x = 0$. Finalement, $E = H \oplus \mathbb{R}e$.
(3) $\Rightarrow$ (2) La démonstration est laissée au lecteur.
(2) $\Rightarrow$ (1) Cela résulte de la définition d'un hyperplan (cf maximalité).

□

Théorème 2.4. [**5**]On suppose maintenant que E est un espace vectoriel normé. On a alors l'équivalence suivante :

$$\text{H est un hyperplan fermé} \Leftrightarrow \exists \varphi \in E', H = \mathbf{Ker}\varphi.$$

Démonstration. Le sens de droit à gauche de l'implication découle directement de la continuité de φ. Pour l'autre sens, on peut faire le raisonnement élémentaire qui suit .
On suppose que H est fermé : E H est donc un ouvert. Soit $e \in E$ H; E H est ouvert donc $\exists r > 0, B(e, r) \subset E$ H. D'aprés la proposition précédente, on sait qu'il existe $\varphi \in E^*$ telle que $H = \mathbf{Ker}\varphi$. On a donc :

$$\forall x \in B(e, r), \varphi(x) \neq . \qquad (*)$$

On peut supposer $\varphi(e) > 0$ (quitte à considérer $-e$) : on a donc $\varphi \mid_{B(e,r)} > 0$ car $B(e, r)$ est convexe et φ est linéaire. En effet, s'il existe $f \in B(e, r)$ tel que $\varphi(f) < 0$, alors comme $0 \in [\varphi(f), \varphi(e)]$ il existe un $\lambda \in [0, 1]$ tel que $0 = (1 - \lambda)\varphi(f) + \lambda\varphi(e)$; la boule est convexe donc $g = (1 - \lambda)f + \lambda e$ y reste, et comme φ est linéaire on en déduit que $0 = \varphi(g)$ ce qui contredit (*).
Soit $h \in E$ de norme 1. On a donc :

$$0 < \varphi(e - \frac{r}{2}h) = \varphi(e) - \frac{r}{2}\varphi(h)$$

donc $\varphi(h) \leqslant \dfrac{2}{r}\varphi(e)$. La transformation $h \longleftarrow -h$ donne $-\varphi(h) \leqslant \dfrac{2}{r}\varphi(e)$ et on a donc $\mid \varphi(h) \mid \leqslant \dfrac{2}{r}\varphi(e)$.
Finalement, φ est bornée sur la sphère unité : φ est linéaire, elle est donc continue. □

Exemples 2.4. Sur les hyperplan :

1. Dans le K-espace vectoriel des matrices carrées d'ordre n à cofficients dans un corps K, l'ensemble des matrices de trace nulle est un hyperplan
2. Dans le K-espace vectoriel $K[X]$ des polynômes à une indéterminée, l'ensemble des polynômes divisibles par X est un hyperplan, car c'est le noyau de la forme linéaire $P \to P(0)$.

2.2.2 Forme analytique du théorème de Hanhn-Banach

Soit E un espace vectoriel sur $\mathbb{R}$. Rappelons qu' une forme linéaire est une aplication linéaire définie sur E, ou sur un sous espace vectoriel de E, à valeurs dans $\mathbb{R}$. Le résultat essentiel théorème de Hanh-Banach(forme analytique) concerne le prolongement d'une forme linéaire définie sur un sous-espace vectoriel de E en une forme linéaire défine sur E tout entier.

Définition 2.13. Soit E un $\mathbf{K}$-espace vectoriel dont le corps des scalaires $\mathbf{K}$ est $\mathbf{R}$ ou $\mathbf{C}$, une application $p : E \longrightarrow \mathbf{R}$ est dite sous-linéaire si elle est

1. positivement homogène de dégré 1, i.e. $p(\lambda x) = \lambda p(x)$ pour tout $\lambda > 0$ et tout $x \in E$.
2. sous-additive, i.e. $p(x+y) \leqslant p(x) + p(y)$ pour tout $x, y \in E$.

Définition 2.14. Soit E un $\mathbf{K}$-espace vectoriel dont le corps des scalaires $\mathbf{K}$ est $\mathbf{R}$ ou $\mathbf{C}$, une application $p : E \longrightarrow \mathbf{R}$ est dite semi-norme si

1. $p(x) \geqslant 0$ pour tout $x \in E$,
2. $p(\lambda x) =| \lambda | p(x)$ por tout $\lambda \in \mathbf{K}$ et tout $x \in E$,
3. pest sous-additive.

Théorème 2.5. Hahn – Banach(version réelle)[**7**]. Soit E un $\mathbf{R}$-espace vectoriel rt soit p une application sous-linéaire, tout forme linéaire $\varphi : F \longrightarrow \mathbf{R}$ défine sur un sous-espace vectoriel $F \subset E$ vérifiant $\varphi \leqslant p$ sur F,se prolonge en une application linéaire $\tilde{\varphi} \leqslant p$ sur E .

Théorème 2.6. Hahn – Banach(version complexe)[**7**]. Soit E un $\mathbf{C}$-espace vectoriel et soit p une application sous-linéaire, tout forme linéaire $\varphi : F \longrightarrow \mathbf{C}$ défine sur un sous-espace vectoriel $F \subset E$ vérifiant $| \varphi |\leqslant p$ sur F,se prolonge en une application linéaire $| \tilde{\varphi} |\leqslant p$ sur E .

Démonstration. (théorème de Hahn-Banach cas réelle). On considère l'ensemble

$$\mathbf{P} = \left\{ h \,\middle|\, \begin{array}{l} h : D(h) \subset E \to \mathbb{R} \text{ avec } D(h) \text{ sous espace vectorial de } E.h \text{ linéaire} \\ G \subset D(h) \text{ prolonge } g \text{ et } h(x) \leq p(x)\ \forall x \in D(h) \end{array} \right\}$$

P est muni de rolation d'ordre

$$(h_1 \leqslant h_2) \Leftrightarrow (\mathbf{D}(h_1) \subset \mathbf{D}(h_2) \text{ et } h_2 \text{ prolonge } h_1).$$

P n'est pas vide puisque $\varphi \in \mathbf{P}$. D'autre part, **P** est inductif. En effet soit $\mathbf{Q} \subset \mathbf{P}$ un sous-ensemble totalement ordonné ; on note $\mathbf{Q} = (h_i)_{i \in I}$. On définit

$$\mathbf{D}(h) = \bigcup_{i \in I} \mathbf{D}(h_i) \qquad et \qquad h(x) = h_i(x) \qquad si \qquad x \in \mathbf{D}(h_i).$$

On vérifie que cette définition a bien un sens, que $h \in \mathbf{P}$ et que h est un majorant de **Q**. Il résulte du lemme de Zorn que **P** admet un élément maximal noté φ. Prouvous que $\mathbf{D}(\tilde{\varphi}) = E$ -ce qui achévera la démonstration du théorème 2.4. Raisonnons par l'absurde et supposons que $\mathbf{D}(\varphi) \neq E$. Soit $x_0 \notin \mathbf{D}(\varphi)$; posons $\mathbf{D}(h)\mathbf{D}(\varphi) + \mathbb{R}x_0$ etpour $x \in \mathbf{D}(\varphi)$, $h(x + x_0) = \varphi(x) + t\alpha$ $(t \in \mathbb{R})$ où α est une constante qui sera fixée ultérieurement de maniére à ce que $h \in \mathbf{P}$. On doit donc s'assurer que

$$\varphi(x) + t\alpha \leqslant p(x + tx_0) \qquad \forall x \in \mathbf{D}(\varphi), \qquad \forall t \in \mathbb{R}.$$

Grâceà p est sous-linéaire($p(\lambda x) = \lambda p(x)$) il suffit devérifier que

$$\begin{cases} \varphi(x) + \alpha \leqslant p(x + x_0) & \forall x \in \mathbf{D}(\varphi) \\ \varphi(x) - \alpha \leqslant p(x - x_0) & \forall x \in \mathbf{D}(\varphi). \end{cases}$$

Autrement dit, il faut choisir α tel que

$$\sup_{y \in \mathbf{D}(\varphi)} \{\varphi(y) - p(y - x_0)\} \leqslant \alpha \leqslant \inf_{x \in \mathbf{D}(\varphi)} \{p(x + x_0) - \varphi(x)\}.$$

Un tel choix est possible puisque

$$\varphi(y) - p(y - x_0) \leqslant p(x + x_0)_\varphi(x) \qquad \forall x \in \mathbf{D}(\varphi), \forall y \in \mathbf{D}(\varphi);$$

en effet on notera que

$$\varphi(x) + \varphi(y) \leqslant p(x + y) \leqslant p(x + x_0) + p(y - x0)$$

grâce à (p(x+y)⩽p(x)+p(y) $\quad \forall x, y \in E$).
On conclut que φ est majorée par h et que $\varphi \neq h$; ceci contredit la maximalité de φ.

□

On peut maintenant passer à la preuve de la version complexe du théorème de Hahn-Banach.

Démonstration. (téorème de Hahn-Banach cas complexe) Considérons donc un **C**-espace vectoriel E et $\varphi : F \longrightarrow \mathbf{C}$ une forme linéaire sur un sous-espace vectoriel F de E bornée par p.
-Notez que

$$\varphi(x) = g(x) + ih(x)$$

tel que g, h sont $\mathbb{R}$-linéaires

$$g, h : F \longrightarrow \mathbb{R}$$

-D'autre parte, selon l'hypothèse

$$\begin{cases} \mid g(x) \mid \leqslant \mid \varphi(x) \mid \leqslant p(x) & \forall x \in F \\ \mid h(x) \mid \leqslant \mid \varphi(x) \mid \leqslant p(x) & \forall x \in F \end{cases}$$

donc

$$\begin{aligned} g(ix) + ih(ix) = \varphi(ix) = i\varphi(x) \\ = i(g(x) + ih(x)) \\ = ig(x) - h(x) \end{aligned}$$

c'est-à-dire

$$h(x) = -g(ix) \qquad \forall x \in F$$
$$\Rightarrow \varphi(x) = g(x) - ig(x)$$

d'aprés le théorème de Hanan-Banach (cas réelle) la forme g (reélle) admet une extension $\tilde{g}$ ($\tilde{g}$ forme linéaire réelle) sur E

$$\tilde{g}(x) \leqslant p(x) \qquad \forall x \in E$$
$$-\tilde{g}(x) = \tilde{g}(-x) \leqslant p(-x) = p(x)$$

c'est-à-dire

$$\mid \tilde{g}(x) \mid p(x)$$

On va définit $\tilde{\varphi}$ comme suit

$$\tilde{\varphi}(x) = \tilde{g}(x) - i\tilde{g}(ix)$$

On a

$$\begin{aligned} \tilde{\varphi}(ix) = \tilde{g}(ix) - i\tilde{g}(-x) \\ = \tilde{g}(ix) + i\tilde{g}(x) = i\tilde{\varphi}(x) \end{aligned}$$

c'est-à-dire $\tilde{\varphi}$ est linéaire complexe, nous prouvons que $\tilde{\varphi}$ est une extension de φ.
On a pour tout $x \in F$

$$\begin{aligned} \tilde{\varphi}(x) = \tilde{g}(x) - i\tilde{g}(x) \\ = g(x) - g(ix) \\ g(x) + ih(x) = \varphi(x) \end{aligned}$$

c'est-à-dire $\tilde{\varphi}$ est extension de φ
On a

$$\mid \tilde{\varphi}(x) \mid \leqslant p(x) \qquad \forall x \in E$$

Nous avons

$$\tilde{\varphi}(x) = \tilde{g}(x) - i\tilde{g}(ix) = re^{-i\theta}$$

donc

$$| \tilde{\varphi}(x) |= r = e^{-i\theta}\tilde{\varphi}(x) = \tilde{\varphi}(e^{-i\theta}x) \qquad (\star)$$

d'aprés $(\star)$ nous concluons que $\tilde{\varphi}(e^{-i\theta}x)$ est réelle positive
d'autre part

$$\tilde{\varphi}(e^{i\theta}x) = \tilde{g}(e^{i\theta}x) - i\tilde{g}(ie^{i\theta}x)$$

c'est-à-dire

$$| \tilde{\varphi}(x) |= \tilde{g}(e^{i\theta}x) - ig(i\widetilde{e^{i\theta}}x)$$

Et de la il sera

$$| \tilde{\varphi}(x) |=| \tilde{g}(e^{i\theta}x) |\leqslant p(e^{i\theta}x) =| e^{i\theta} |= p(x)$$

$$| \tilde{\varphi}(x) |\leqslant p(x) \qquad x \in E.$$

□

Sur un espace vectoriel normé, en choisisant pour p un multiple de norme, on déduit aisément le corollaire qui suit.

Corollaire 2.3. Soit E un $\mathbf{K}$-espace vectoriel normé, dont le corps des scalaires $\mathbf{K}$ est $\mathbf{R}$ ou $\mathbf{C}$, toute forme linéaire $\varphi : F \longrightarrow \mathbf{K}$ continue sur un sous-espace vectoriel $F \subset E$ se prolonge en une application linéaire $\tilde{\varphi} : E \longrightarrow \mathbf{K}$ contiune.

On peut maintenant passer à la preuve de la version réelle du théorème de Hahn-Banach.

2.2.3 Forme géométriques du théorème de Hahn-Banach

Dans toute la suite E désigne un e.v.n.

Proposition 2.3. L'hyperplan d'équation $[f = \alpha]$ est fermé <u>ssi</u> f est continue.

Démonstration. si f est continue alors $\mathbf{H}$ est fermé. Réciproquement ; supposons que $\mathbf{H}$ est fermé. Le complémentaire $\complement\mathbf{H}$ de $\mathbf{H}$ est ouvert et non vide puique $f \not\equiv 0$. Soit $x_0 \in \complement\mathbf{H}$ et supposons que $f(x_0) < \alpha$. Soit $r > 0$ tel que $B(x_0, r) \subset \complement\mathbf{H}$ où

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

On a

$$f(x) < \alpha \qquad \forall x \in B(x_0, r). \qquad (*)$$

En effet supposons que $f(x_1) > \alpha$ pour un certain $x_1 \in B(x_0, r)$. Le segment

$$\{x_t = (1 - t)x_0 + tx_1; t \in [0, 1]\}$$

est contenu dans $B(x_0, r)$ et donc $f(x_t) \neq \alpha$ $\forall t \in [0, 1]$; par ailleurs $f(x_t) = \alpha$ pour $t = \dfrac{f(x_1) - \alpha}{f(x_1) - f(x_0)}$, ce qui est absurde et donc (*) est démontré. Il résulte de (*) que

$$f(x_0 + rz) < \alpha \qquad \forall z \in B(0, 1).$$

Par conséquent f est continue et $\| f \|\leqslant \dfrac{1}{r}(\alpha - f(x_0))$. □

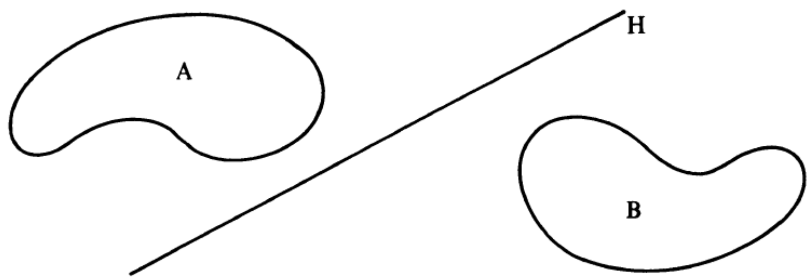
Définition 2.15. Soient $A \subset E$ et $B \subset E$. On dit que l'hyperplan H d'équation $[f = \alpha]$ sépare A et B au sens large si l'on a

$$f(x) \leqslant \alpha \qquad \forall x \in A \qquad et \qquad f(x) \geqslant \alpha \qquad \forall x \in B.$$

On dit que H sépare A et B ausens strict s'il $\exists \varepsilon > 0$ tel que

$$f(x) \leqslant \alpha - \varepsilon \qquad \forall x \in A \qquad et \qquad f(x) \geqslant \alpha + \varepsilon \qquad \forall x \in B.$$

Géométriquement la séparation exprime que A et B se situent(de part et d'autre de H)



Rappelons enfin qu' un ensemble $A \subset E$ est **convexe** si

$$tx + (1-t)y \in A \qquad \forall x, y \in A, \qquad \forall t \in [0, 1].$$

Théorème 2.7. (Hahn-Banach, première forme géométrique)[**7**]. Soient $A \subset E$ et $B \subset E$ deux ensembles convexes, non vides et disjoints. On suppose que A est ouvert. Alors il existe un hyperplan fermé qui sépare A et B au sens large.

La démonstration du cette théorème est basée sur les deux lemme suivants

Lemme 2.2. (Jauge d'un convexe)[**7**]. Soit $C \subset E$ un convexe ouvert avec $0 \in C$. Pour tout $x \in E$ on pose :

$$p(x) = \inf\{\alpha > 0; \alpha^{-1}x \in C\}$$

(on dit que p est la jauge de C).
Alors p est positivement homogène de dégrie 1 et sous-additive,et

$$\exists M \text{ tel que} 0 \leqslant p(x) \leqslant M \parallel x \parallel \qquad \forall x \in E$$

$$C = \{x \in E; p(x) < 1\}.$$

Lemme 2.3. [**7**]Soit $C \subset E$ un convexe ouvert non vide et soit $x_0 \in E$ avec $x_0 \notin C$ Alors il existe $f \in E'$ tel que $f(x) < f(x_0) \forall x \in C$. En particulier l'hyperplan d' équation $[f = f(x_0)]$ sépare x_0 et C au sens large.

Démonstration. (Théorème 2.7) On pose $C = A - B$ de sort que C est convexe (vérification facile), C est ouvert (noter que $C = \bigcup_{y \in B}(A - y)$) et $0 \notin C$ (puisque $A \cap B = \varnothing$). D'aprés lelemme 2.2 il existe $f \in E'$ tel que

$$f(z) < 0 \qquad \forall z \in C$$

c'est-à-dire

$$f(x) < f(y) \qquad \forall x \in A \qquad \forall y \in B.$$

On fixe $\alpha \in \mathbb{R}$ avec

$$\sup_{x \in A} f(x) \leqslant \alpha \leqslant \inf_{y \in B} f(y)$$

et donc l'hyperplan d'équation $[f = \alpha]$ sépare au sens large A et B. $\square$

Théorème 2.8. (Hahn-Banach, deuxiéme forme géométrique)[**7**]. Soient $A \subset E$ et $B \subset E$ deux ensembles convexes, non vides, disjoints. On suppose que A est fermé et que B est compact. Alors il existe un hyperplan fermé qui sépare A et B au sens strict

Démonstration. (Théorème 2.8) Pour $\varepsilon > 0$ on pose $A_\varepsilon = A + B(0, \varepsilon)$ et $B_\varepsilon = B + B(0, \varepsilon)$ de sorte que A_ε et B_ε sont convexes, ouverts et non vides. De plus, pour $\varepsilon > 0$ assez petit, A_ε et B_ε sont disjoints (sinon on pourrait trouver des suites $\varepsilon_n \longrightarrow 0, x_n \in A$ et $y_n \in B$ telles que $\| x_n - y_n \| < 2\varepsilon_n$; on pourrait ensiute extraire une sous-suit $y_n \longrightarrow y \in A \cap B$). D'aprés le théorème 2.4, il existe un hyperplan fermé d'équation $[f = \alpha]$ qui sépare A_ε et B_ε au sens large. On a donc

$$f(x + \varepsilon z) \leqslant \alpha \leqslant f(y + \varepsilon z) \qquad \forall x \in A, \qquad \forall y \in B, \qquad \forall z \in B(0, 1).$$

Il n résulte que

$$f(x) + \varepsilon \| f \| \leqslant \alpha \leqslant f(y) - \varepsilon \| f \|, \qquad \forall x \in A, \qquad \forall y \in B.$$

On conclut que A et B sont séparés au sens strict par l'hyperplan $[f = \alpha]$ puisque $\| f \| \neq 0$. $\square$

2.3 Théorème de Baire

Théorème 2.9. [**7**]Soit X un espace métrique complet non vide. Soit $(X_n)_{n \geqslant 1}$ une suite de fermés. On suppose que

$$\mathbf{Int} X_n = \varnothing \text{ pour chaque} \qquad n \geqslant 1.$$

Alors

$$\mathbf{Int}\left(\bigcup_{n=1}^{\infty} X_{n_0} \right) \neq \varnothing.$$

dautre définition
•Soit (E, d) un espace métrique complet.

1. Si $(X_n)_{n \geq 1}$ est suite d'ouverts denses de E, alors $\cap_{n \geq 1} X_n$ et encore dense dans E.
2. Si $(X_n)_{n \geq 1}$ est une suite de fermés d'intérieur vide de E, alors $\cup_{n \geq 1} X_n$ est encore d'intérieur vide dans E

Remarques 2.1. Le théorème de Baire est en général utilisé sous la forme suivant. Soit X un espace métrique complet non vide. Soit $(X_n)_{n \geqslant 1}$ une suite de fermés telle que

$$\bigcup_{n=1}^{\infty} X_n = X.$$

Alors il existe n_0 tel que $\mathbf{Int} X_{n_0} \neq \varnothing$

Démonstration. On pose $O_n = \complement X_n$ de sorte que O_n est un overt dence. Il s'agit de montrer que :

$$\mathbf{G} = \bigcap_{n=1}^{\infty} O_n,$$

est dense dans X.
Soit ρ un overt non vide de X ; on va prouver que $\rho \cap \mathbf{G} \neq \varnothing$.
On not

$$\mathbf{B}(x, r) = \{y \in X; \qquad d(y, x) < r\}.$$

On choisit $x_0 \in \rho$ et $r_0 > 0$ arbitraires tel que

$$\overline{\mathbf{B}}(x_0, r_0) \subset \rho.$$

On choisit ensuite $x \in \mathbf{B}(x_0, r_0) \cap O_1$ et $r_1 > 0$ tels que

$$\begin{cases} \overline{\mathbf{B}}(x_1, r_1) \subset \mathbf{B}(x_0, r_0) \cap O_1 \\ 0 < r_1 < \dfrac{r_0}{2} \end{cases}$$

ceci est possible puisque O_1 est ouvert et dence. Ainsi de suite, on construit par récurrence deux suites (x_n) et (r_n) telles que

$$\begin{cases} \overline{\mathbf{B}}(x_{n+1}, r_{n+1}) \subset \mathbf{B}(x_n, r_n) \cap O_{n+1} \qquad \forall n \geqslant 0 \\ 0 < r_{n+1} < \dfrac{r_n}{2}. \end{cases}$$

Il en résulte que la suite (x_n) est de Cauchy ; soit $x_n \longrightarrow L$. Comme $x_{n+p} \in \mathbf{B}(x_n, r_n)$ pour tout $n \geqslant 0$ et tout $p \geqslant 0$, on obtient à la limite (quand $p \longrightarrow \infty$) :

$$L \in \overline{\mathbf{B}}(x_n, r_n) \qquad \forall n \geqslant 0.$$

En particulier $L \in \rho \cap \mathbf{G}$. $\square$

Chapitre 3

Séparation des espaces vectoriels par les hyperplans

3.1 Formes qoudratiques

Dans tout ce chapitre, E désingne un $\mathbb{R}$-espace vectoriel de dimonsion finie n.

3.1.1 Formes bilinéaires symétriques

Définition 3.1. [**8**] Une application
$$\varphi : E \times E \longrightarrow \mathbb{R}$$
appelée une forme bilinéaire quand

$$\forall x_1, x_2, y \in E,\ \forall \lambda \in \mathbb{R} \qquad \varphi(x_1 + \lambda x_2, y) = \varphi(x_1, y) + \lambda \varphi(x_2, y)$$

$$\forall x, y_1, y_2 \in E,\ \forall \lambda \in \mathbb{R} \qquad \varphi(x, y_1 + \lambda y_2) = \varphi(x, y_1) + \lambda \varphi(x, y_2)$$

(bilinéarité= linéarité à gauche +linéarité à droite).
On dit que φ est symétrique quand

$$\forall x, y \in E \qquad \varphi(x, y) = \varphi(y, x).$$

Remarques 3.1. Si φ est une forme bilinéaire sur E. alors , pour tout $x \in E$, $\varphi(0, x) = \varphi(x, 0) = 0$.

Exemples 3.1. 1. $E = \mathbb{R}$. La multiplication $(x, y) \longmapsto xy$ est une forme bilinéaire symétrique sur $E \times E$.

2. $E = \mathbb{R}^2$. Le produit scalaire usuel

$$\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \longmapsto x_1 y_1 + x_2 y_2$$

est une forme bilinéaire symétrique sur $\mathbb{R}^2 \times \mathbb{R}^2$.

Proposition 3.1. 1. Pour tout scalaire λ et tout vectoeur ϑ, $\varphi(\lambda\vartheta) = \lambda^2 \varphi(\vartheta)$.

2. Deux vecteurs ϑ, et ω sont orthogonaux par rapport à E ssi $\varphi(\vartheta + \omega) = \varphi(\vartheta) + \varphi(\omega)$
3. φ obéit à la régel du parallélogramme : $\varphi(\vartheta + \omega) + \varphi(\vartheta - \omega) = 2\varphi(\vartheta) + 2\varphi(\omega)$

Proposition 3.2. L'ensemble des formes quadratiques sur $\mathbb{R}$-espace vectoriel E est un $\mathbb{R}$-espace vectoriel.

Définition 3.2. (**Formes quadratiques**).[8] Une application $q : E \longrightarrow \mathbb{R}$ est appelée forme quadratique sur E s'il existe une forme bilinéaire symétrique φ sur $E \times E$ telle que

$$\forall x \in E \qquad q(x) = \varphi(x, x)$$

La forme quadratique q est dite associée à la forme bilinéaire symétrique φ

Exemples 3.2.
1. $x \longmapsto x^2$ (sur $\mathbb{R}^2$),
2. $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto x_1^2 + x_2^2$ (sur $\mathbb{R}^2$),
3. $f \longmapsto \int_{-1}^{1} f(t)^2 dt$ (sur $\mathcal{C}^0([-1, 1], \mathbb{R})$).

3.1.2 Écriture matricielle

Soit $(e_1, ..., e_n)$ unebase de E.
Soint x et y deux vecteurs de E de coordonnées respectives $(x_i)_{1 \leq i \leq n}$ et $(y_j)_{1 \leq j \leq n}$ dans la base $(e_1, ..., e_n)$.
Soit φ une forme bilinéaire symétrique définie sur E. On a alors par bilinéairité de φ :

$$\begin{aligned}\varphi(x, y) &= \varphi\left(\sum_{i=1}^{n} x_i e_i \sum_{j=1}^{n} y_j e_j\right) \\ &= \sum_{1 \leq i,j \leq n} x_i y_j \varphi(e_i, e_j)\end{aligned}$$

Réciproquement, soit $(a_{ij})_{1 \leq i,j \leq n}$ une famille de réels telle que $a_{ij} = a_{ji}$ pour $1 \leq i, j \leq n$; alors l'application $(x, y) \longmapsto \sum_{1 \leq i,j \leq n} a_{ij} x_i y_j$ est bilinéaire symétrique.

Définition 3.3. [8] Soit φ une forme bilinéaire symétrique définie sur E et soit $(e_1, ..., e_n)$ une base de E. La matrice M de $\mathcal{M}_n(\mathbb{R})$ définie par $M_{ij} = \varphi(e_i, e_j)$ s'appelle la matrice de φ dans la base $(e_1, ..., e_n)$

Définition 3.4. [8] Soit q une forme quadratique. La matrice de la forme bilinéaire symétrique associée à q dans $\mathfrak{B}$ s'appelle la matrice de q dans la base $\mathfrak{B}$.

Définition 3.5. [8] Deux matrices M et M' de $\mathfrak{M}_n(\mathbb{K})$ sont dites congruentes s'il existe une matrice $P \in GL_n(\mathbb{K})$ telle que $M' = P^t M P$.
Deux matrices sont donc congruentes si elles représentent la même forme bilinéaire dans deux bases différentes de E.

Exemple 3.1.1. La forme quadratique

$$q(x_1, x_2, x_3) = x_1^2 + 7x_2^2 + 6x_1x_2 - 2x_1x_3 + 8x_2x_3$$

a pour matrice $\begin{pmatrix} 1 & 3 & -1 \\ 3 & 7 & 4 \\ -1 & 4 & 0 \end{pmatrix}$.

3.1.3 Recherche de la forme bilinéaire associée à forme quadratique

Soit $(e_1, ..., e_n)$ une base de E. Une forme bilinéaire symétrique φ est une application de $E \times E$ dans $\mathbb{R}$ définie par $\varphi(x, y) = X^t M Y = \sum_{i,j} m_{ij} x_i y_j$ où M est matrice symétrique réelle définie par

$m_{ij} = \varphi(e_i, e_j)$.
Une forme quadratique s'écrit donc sous la forme :

$$q(x) = \sum_{1\leq i,j\leq n} m_{ij}x_ix_j = \sum_{i=1}^{n} m_{ii}x_i^2 + 2\sum_{1\leq i<j\leq n} m_{ij}x_ix_j.$$

Réciproquement, si on donne une forme quadratique q, on a alors

$$q(x) = \sum_{i=1}^{n} m_{ii}x_i^2 + 2\sum_{1\leq i<j\leq n} m_{ij}x_ix_j.$$

Pour retrouver la forme bilinéaire associéée φ à q, on utilise la règle du dédoublement des termes :

- on remplace les termes x_i^2 par x_iy_i
- on remplace le terme x_ix_j par $\frac{1}{2}(x_iy_j + x_jy_i)$

On vérifie que, pour φ ainsi contruite, on a bien $\varphi(x, y) = \frac{1}{2}[q(x + y) - q(x) - q(y)]$.

3.1.4 Rang d'une forme bilinéaire

Soient φ une forme bilinéaire définie sur un espace vectoriel E de dimension finie et x et y deux vecteurs de E.
On définit deux formes linéaires φ_x φ_y de E par

$$\forall y \in E, \qquad \varphi_x = \varphi(x, y)$$

$$\forall x \in E, \qquad \varphi_y = \varphi(x, y)$$

Notons E' le dual de E (c'est-à-dire l'ensemble des formes linéaires définies sur E). Les deux applications de E dans E' définies par

$$x \mapsto \varphi_x$$

et

$$y \mapsto \varphi_y$$

sont linéaires de E dans E'.
Soint $(e_1, ..., e_n)$ une base de E, M la matrice de φ dans cette base et $(e*_1, ..., e*_n)$ la base dual. On a, pour tout $1 \leq i, j \leq n$, $m_{ij} = \varphi(e_i, e_j)$ donc la matrice M^t (respectivement M) représente l'endomorphiseme $x \mapsto \varphi_x$ (respectivement $y \mapsto \varphi_y$) de la base $(e_1, ..., e_n)$ dans la base $(e*_1, ..., e*_n)$.
En effet, la $j$éme colonne de la matrice représentant l'endomorphisme $x \mapsto \varphi_x$ dans les base définies précédemment est la matrice-colonne des coordonnées de φ_{e_j} dans la base $(e*_1, ..., e*_n)$. Posons $\varphi_{e_j} = \sum_{i=1}^{n} \lambda_i e_i^*$.
Comme $\varphi_{e_j}(e_k) = \sum_{i=1}^{n} \lambda_i e_i^*(e_k) = \lambda_k = \varphi(e_j, e_k)$,la matrice représentant l'endomorphisme $x \mapsto \varphi_x$ de la base $(e_1, ..., e_n)$ dans la base $(e*_1, ..., e*_n)$ est donc bien M^t. De même, pour $y \mapsto \varphi_y$.

Définition 3.6. (**Noyau**)[**8**]. On appelle noyau de la forme quadratique q, et on note **Ker**(q), l'ensemble

$$\mathbf{Ker}(q) = \{y \in E; \varphi(x, y) = 0\}.$$

Exemple 3.1.2. La forme quadratique q sur $\mathbb{R}^2$ définie par

$$q(x_1, x_2) = x_1^2 - x_2^2$$

a pour matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
Son noyau est réduit à $\{0\}$.

Définition 3.7. (**Rang d'une forme bilinaire**)[8]. On appelle rang d'une forme bilinéaire φ définie sur un espace vectoriel E de dimension finie n est le nombre maximal de vecteurs lignes (ou colonnes) linéairement indépandnts ou le plus grand des ordres des matrices carrées inversibles extraites de la matrice A qu'associe φ

Exemple 3.1.3.
$q(x,y,z) = x^2 - 2y^2 + 3z^2 + 4xy + 6xz$
$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -2 & 0 \\ 3 & 0 & 3 \end{pmatrix} \qquad rg(q) = 2$

$$Ker(q) = Vect\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

3.2 La séparation des espaces topologiques

Définition 3.8. (**Sous-espace supplémentaire**). Deux sous-espace vectoriels V et W de E sont dits supplémentaires si, et sulement si, E est la somme directe de V et W, soit

$$E = V \oplus W$$

Remarques 3.2. Si H est un hyperplan et φ une forme linéaire sur E de noyau H, on a

$$E = H + \mathbf{Ker}\varphi \Longleftrightarrow \varphi(e) \neq 0e \notin H$$

ainsi, toute vectorielle e non contenue dans l'hyperplan H est un supplémentaire de E

Remarques 3.3. Soit E un $\mathbb{K}$-espace vectoriel de dimonsion finie $n \in \mathbb{N}^*$, H un hyperpaln de E et D une droit vectorielle de E.
À quelle condition H et D sont-ils supplémentaires dans E ?
Solution
Si $D \subset H$ alors H et D ne sont pas supplémentaires car

$$H \cap D = D \neq \{0_E\}.$$

Supposons $D \not\subseteq H$.
Soit $x \in D \cap H$. Si $x \neq 0_E$ alors $D = Vect(x) \subset H$ ce qui exclu.
Nécessairemant $D \cap H = \{0_E\}$.
De plus $dimE = dimH + dimD$ donc

$$E = H \oplus D.$$

Théorème 3.1. Pour toute forme quadratique sur espace de dimension finie, il existe une base formée de vecteurs deux à deux orthogonaux.

Démonstration. Si φ est la forme nulle, n'importe quelle base convient. Sinon, il existe un vecteur $(Vect(u))$ tel que la forme linéaire est donc non nulle, son noyau est un hyperplan □

3.2.1 Exemples sur la séparation par hyperplan

Exemple 3.2.1. Dans $E = \mathbb{R}_2[X]$, on considére : $F = \{P \in E \mid P(1) + P'(0) = 0\}$.
Justifier que F est un hyperplan de E, en déduire sa dimension, puis en donner une base, toutes ses équations et tous ses supplémentaires dans E.

Solution.

1. F est un hyperplan. La forme linéaire

$$\varphi : \mathbb{R}_2[X] \longrightarrow \mathbb{R}$$
$$P \longmapsto P(1) + P'(0)$$

est une forme linénaire. En effet, si $P, Q \in \mathbb{R}_2[X]$ et $\lambda \in \mathbb{R}$ alors

$$\begin{aligned}\varphi(P + \lambda Q) &= (P + \lambda Q)(1) + (P + \lambda Q)'(0)\\ &= P(1) + \lambda Q(1) + P'(0) + \lambda Q'(0)\\ &= P(1) + P'(0) + \lambda(Q(1) + Q'(0))\\ &= \varphi(P) + \lambda\varphi(Q).\end{aligned}$$

De plus φ est non-nulle,
par exemple on a $\varphi(X) = 1 + 1 = 2$.
Et donc

$$Ker(\varphi) = F \text{ est un hyperplan de } E.$$

2. Dimension de F. Puisque F est un hyperplan de $\mathbb{R}_2[X]$, qui est de dimension 3, F est de dimension 2.
3. Equation de F. Tout polynôme $P \in \mathbb{R}_2[X]$ s'écrit sous la forme

$$aX^2 + bX + c \text{ avec } a, b, c \in \mathbb{R}.$$

On a donc

$$\varphi(P) = 0 \Longleftrightarrow a + b + c + b = 0 \Longleftrightarrow a + 2b + c = 0.$$

Toutes les équations de F sont de la forme

$$\lambda(a + 2b + c) = 0 \text{ avec } \lambda \in \mathbb{K}^*.$$

4. Base de F. Les polynômes appartenant à F sont dela forme :

$$(-2b - c)X^2 + bX + c = b(-X^2 + X) + c(-X^2 + 1).$$

Ainsi la famille composée de polynômes $-2X^2 + X$ et $-X^2 + 1$ est une famile génératricee F, de plus elle est libre, c'est donc une base de F.
5. Supplémentaire de F. Comme F est hyperplan, tous les supplémentaires de F sont de la forme $Vect(u)$ où $u \notin F$. En d'autres termes, si $P = aX^2 + bX + c$ vérifie $a \neq -2b - c$ alors $Vect(u)$ est supplémentaire de F

Exemple 3.2.2. Dans $E = \mathbb{R}^4$, muni de sa base canonique $\mathfrak{B}_0 = (e_1, e_2, e_3, e_4)$, on considère :

$$F = \{(x, y, z, t) \in E \mid x + y - z - t = 0\}.$$

Justifier que F est hyperplan de E, en déduire sa dimonsion, puis en donner une base, toutes ses équations et tous ses supplémentaires dans E.
Solution

1. F est un hyperplan, la forme linéaire

$$\varphi \quad : \quad \mathbb{R}^4 \longrightarrow \quad \mathbb{R}$$
$$(x, y, z, t) \longmapsto x + y - z - t$$

est non nulle car $\varphi(1, 0, 0, 0) = 1 \neq 0$ et $F = Ker(\varphi)$. Ainsi F est un hyperplan de E.

2. Dimension de F. Puisque F est un hyperplan de $\mathbb{R}^4$, qui est dimension 4, F est de dimension 3.
3. Equations de F. Toutes les équations de F sont de la forme

$$\lambda(x + y - z - t) = 0 \text{ avec } \lambda \in \mathbb{K}^*.$$

4. Base de F. Les vecteurs de F sont de la forme :

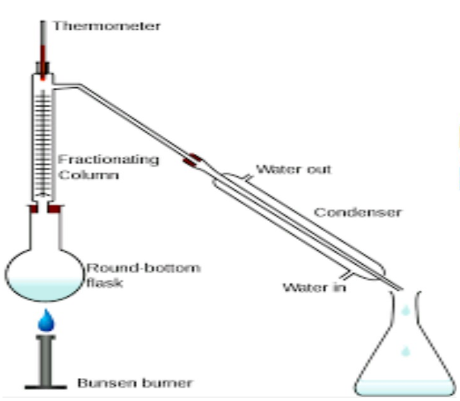
$$\begin{pmatrix} z+t-y \\ y \\ z \\ t \end{pmatrix} = Y \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \text{ où } y, z, t \in \mathbb{R}.$$

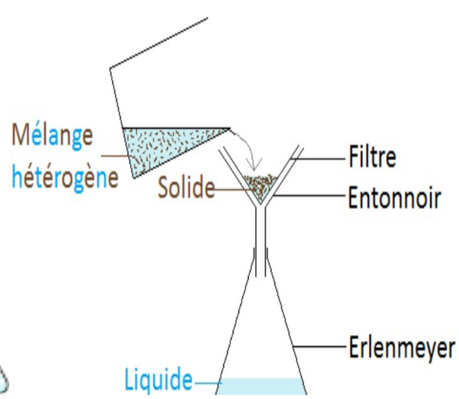
On voit donc que la famille composée des vecteurs $\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ est la famille génératrice de F, de plus elle est libre, ainsi c'est unebase de F.
5. Supplémentaire de F. Comme F est un hyperplan, tous les supplémentaires de F sont de forme $Vect(u)$ où $u \notin F$. En d'autres termes, si $u = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ est tel que $x + y - z - t \neq 0$ alors $Vect(u)$ est supplémentaire de F.

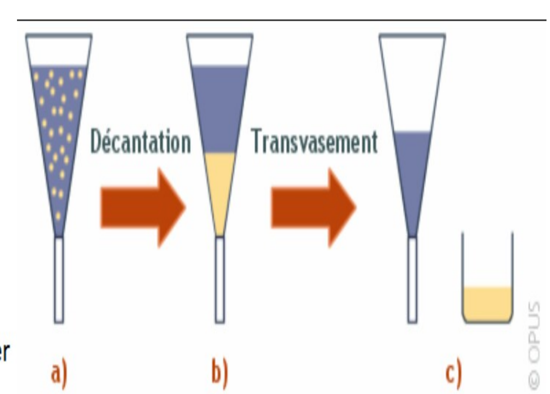
3.3 L'avantage de la séparation

En chimie et physique, un procédé de séparation est une technique ou une technologie permettant de transformer un mélange de substances en deux ou plusieurs composants distincts. Les buts de ce type de procédé peuvent être divers :
•Purification : des impuretés doivent être extraites du composé d'intérêt .
•Concentration : élimination d'une partie du solvant.
•Fractionnement : séparation d'un mélange complexe en plusieurs mélanges différents.

Le principe d'un procédé de séparation est d'utiliser une différence de propriétés entre le composé d'intérêt et le reste du mélange. Plus la différence de propriété sera grande, plus la séparation sera aisée. Ainsi, le choix du procédé de séparation commence par une bonne connaissance de la composition du mélange et des propriétés des différents composants.







3.3.1 Les SVM

•Les modèles chimiométriques peuvent s'inscrire dans le cadre de la théorie de l'apprentissage statistique qui intègre explicitement l'optimisation du compromis entre biais et variance. Nous insistons sur les développements récents en classification supervisèe à l'origine des méthodes de séparateurs à vaste marge, également appelées machines à vecteurs de support (SVM, support vector machines). L'algorithme d'optimisation est par nature adapté aux données spectroscopiques, la formulation du problème reposant d'une part sur l'écriture d'un nombre relativement limité de contraintes dans un espace de grande dimension et d'autre part sur la notion de produit scalaire. Dans la lignée des travaux effectués sur les réseaux de neurones artificiels, nous présentons les résultats obtenus récemment pour la prédiction de propriétés physico-chimiques par la mesure du spectre proche infrarouge. Les performances des modèles SVM de classification sont discutées, l'accent étant mis sur l'optimisation de l'étape d'apprentissage et l'interprétation des résultats.

•Les SVM sont des algorithmes d'apprentissage basés sur la recherche de l'hyperplan optimal, celui qui sépare correctement les données tout en étant le plus éloigné possible des observations pour assurer de bonnes capacités de généralisation. Cette définition se comprend facilement pour des données séparables. Dans le cas de la discrimination de données non séparables, deux astuces de calcul vont permettre de ramener la recherche de surfaces séparatrices non linéaires à un problème de classification linéaire. La première idée consiste à définir l'hyperplan séparateur comme solution d'un problème d'optimisation sous contrainte. La fonction de coût possède alors la particularité de ne s'exprimer qu'à l'aide de produits scalaires. La deuxième idée réside dans l'introduction d'une fonction noyau dans le produit scalaire. Cela induit implicitement une transformation non linéaire de l'espace initial des données vers un espace intermédiaire de plus grande dimension dans lequel est résolu un problème de classification linéaire.
•Le principe fondateur des SVM est lié aux développements récents en apprentissage statistique, en particulier le fait que la capacité de généralisation d'un modèle puisse être optimisée en contrôlant sa complexité, c'est-à-dire finalement le nombre de vecteurs de support. Les méthodes SVM ont d'abord été définies pour la discrimination puisque le modèle de classification SVM est une généralisation de l'algorithme du perceptron. Elles ont ensuite été étendues en régression. La notion clé est ici la notion de marge, au sens d'une marge de sécurité prise lors de la décision ; les échantillons doivent non seulement être bien classés mais également se trouver à une distance suffisante de la frontière.

3.3.2 Hyperplan séparateur optimaux (cas des classes non séparables)

On se place comme précédemment dans le cas d'un échantillon statistique $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ de loi inconnue avec X dans $\mathbb{R}^p$ et Y dans $\{-1, 1\}$
Le problème se pose comme la recherche d'une frontière de décision, fonction discriminante $f(x)$, dans l'espace des valeurs de X. Cet hyperplan est défini par

$$\{x : f(x) = x^T\beta + \beta_0\}$$

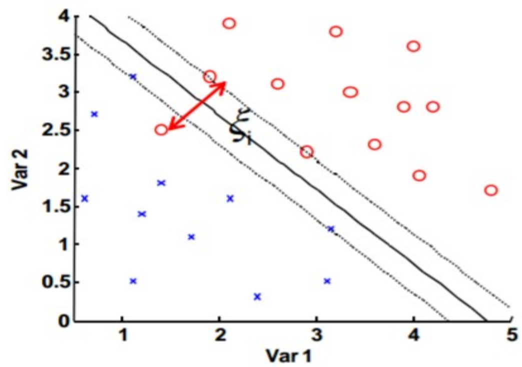
où β vérifie $||\beta|| = 1$. Le plan séparateur optimal est celui qui sépare les deux classes tout en maximisant la marge (la distance au point échantillon le plus proche pour chaque classe).
les équation (3.1),(3.2) et (3.3) définissant l'hyperplan optimal.

$$\beta = \sum_{i=1}^{n} \alpha_i y_i x_i \tag{3.1}$$

$$\sum_{i=1}^{n} \alpha_i y_i x^T x_i + \beta_0 = 0 \qquad (\alpha > 0) \tag{3.2}$$

$$G(x) = signe(x^T\beta + \beta_0) \tag{3.3}$$

L'approche précédente peut être généralisée aux cas pour lesquels les classes ne sont pas séparables, au sens d'une séparation linéaire, c'est à dire lorsque les distributions d'échantillons se recouvrent dans l'espace des variables



L'idée est la suivante : continuer de maximiser la marge pour avoir la meilleure généralisation possible, mais tolérer certains points du mauvais coté de la frontière. Un terme pénalisant les erreurs individuelles ξ_i est donc introduit dans l'expression des contraintes (3.4).
Ce terme reprsente la proportion d'échantillons que l'on accepte dans la marge, ou du mauvais coté de la frontière, c'est à dire l'erreur totale de pré édiction en apprentissage.

$$y_1(X_1^T\beta + \beta_0) \geq d(1-\xi_i), \xi_i \geq 0, \sum_{i=1}^{n} \leq cte \qquad et\forall i = 1, ..., n \tag{3.4}$$

Une formulation équivalente du problème est donnée par leséquations (3.5) et (3.6).

$$\underset{(\beta,\beta_0)}{\text{Minimiser}} \frac{1}{2} \parallel \beta \parallel^2 + C\sum_{i=1}^{n} \xi_i \tag{3.5}$$

soumis à

$$y_i(X_i^T\beta + \beta_0) \geq 1 - \xi_i, \xi_i \geq 0 \qquad \forall i = 1, ..., n \tag{3.6}$$

Il s'agit comme précédemment d'un problème quadratique d'optimisation sous contrainte. On notera l'introduction du terme de pénalité et du paramètre C. Il s'agit d'un paramètre d'adaptation (méta-paramètre) qui doit être optimisé lors de l'apprentissage. Plus C est grand, plus les erreurs de classification ($\xi_i > 0$) sont pénalisées au profit de l'ajustement mais potentiellement au détriment de la généralisation. A l'inverse, des valeurs faibles de C favorisent les marges larges.
Les étapes principales des calculs pour la résolution du probléme (3.5)(3.6) sont reprises en annexe 3 [9].
L'interprétation des résultats peut être calquée sur celle présentée dans le cas des données séparables.
La solution est de la forme (3.1). Les observations pour lesquelles les contraintes sont actives ($\alpha_i > 0$) sont les vecteurs de support, et β ne s'exprime qu'en fonction d'eux. Parmi ces vecteurs, certains sont disposés sur la marge ($\xi_i = 0$) avec $0 < \alpha_i < C$, les autres correspondent à $\xi_i > 0$ avec $\alpha_i = C$. La fonction de décision (3.3) reste inchangée.

3.4 Conclusion

Dans ce travail, nous traitons le problème de la séparation des espaces de Banach en utilisant les hyperplans comme un model, comme on a vu quelques aspects chimie qui parement l'importance de la séparation. Par l'ide de lemme du Zorn on pente créer des autres modèles de séparation dans des espaces quelconques.

Bibliographie

[1] P.RÉMI. Le lemme de Zorn. 17 avril 2008.

[2] M. EIDEHEIT. Zur theorie der konvexen Mengen in linearen normiertrn Raümen . Studia Mathematica. vol. 6.1936. p.104-111.

[3] G.PHILIPPE . Ciarlet. Linear and Nonlinear Functional Analysis with Applications.City University of Hong Kong. Siam.2013. 5-7.

[4] F. RONGA. Topologique et géomtrie. Genéve. MMVIap. J.-C.3-31.

[5] COMPLÉMENTS MATHÉMATIQUES. Cours 2011/2012 d'Arnaud DEBUSSCHE[1] à l'ENS Cachan Antenne de Bretagne. Note de cours[2].45-49.

[6] S. RIVAUD, Formes Linéaires et Hyperplans en dimension finie. Exemples et applications. 7 novembre 2012.

[7] H.BREZIS. Analyse fonctionnelle, théorie et application. MASSON Paris New York Barcelone Milan Mexico Sao Paulo . (1987). 1-16.

[8] J. HAVET. Algèbre bilinéaire et géométrie euclidienne. BP 6759, 45067 Orléans Cedex 2, France. Janvier 2013. 4-18.

[9] C. RUCKEBUSCH. Résolution et modélisation chimiométrique en spectroscopie moléculaire.HAL.France.2008. 52-75.

Résumé

Dans ce mémoire nous avons étudié le lemme de Zorn et nous présentions les applications de cette lemme dans l'espace de Banach.
Et nous avons séparé les espaces vectoriels par les hyperplans, pour cela, nous avons donné Des exemples et applications de séparation des espaces
Les mots clés : lemme de Zorn, hyperplan , séparation.

Abstract

in this memoire we studied Zorn's lemma and we applied in the Banach spaces.
And we separated the vector spaces by hyper plane, for this we give exemplas and application of separate spaces
Keywords : Zorn's lemma, hyper plane, separate.