COMPLÉMENTS MATHÉMATIQUES

Cours 2011/2012 d'Arnaud DEBUSSCHE [1]
à l'ENS Cachan Antenne de Bretagne

Notes de cours [2]

1. Excepté pour le chapitre 3, fait par Gilles Vilmart.
2. Tapé par ROSTAM Salim, sous licence cc-by-sa comme dirait l'autre ; voir http://creativecommons.org/licenses/by-sa/3.0/.

Table des matières

Chapitre 1

Logique et axiomes de la théorie des ensembles

I Introduction

On travaille sur des propriétés P,Q,... qui peuvent être vraies ou fausses. $P \Rightarrow Q$ veut dire $\overline{P} \vee Q$. *Rappel :* $(\forall Q, P \Rightarrow Q) \Rightarrow P$

Définition 1.1 Une **théorie** est un ensemble de propriétés ainsi que des axiomes. Cette théorie est dite **contradictoire** si elle est telle que $\exists P, P \wedge \overline{P}$.

Remarque : Si $P \wedge \overline{P}$ alors toute propriété est vérifiée. En effet, soit Q une propriété. On a alors $P \Rightarrow (P \vee Q)$ donc $P \vee Q$ est vraie (car P l'est). Or, $(P \vee Q) \equiv (\overline{P} \Rightarrow Q)$. Donc comme on a $\overline{P}$ on a également Q.

II Les cinq 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.

II.1 Axiome d'extensionnalité

Axiome
$\forall a, b$ ensembles, $(a = b) \Leftrightarrow (\forall x, x \in a \text{ \underline{ssi} } x \in b)$

Définition 1.2 (Inclusion) Soient a et b deux ensembles. On note $a \subset b$ <u>ssi</u> $\forall x, (x \in a \Rightarrow x \in b)$

Théorème 1.3 (Double inclusion)
Soient a et b deux ensembles. Alors $a = b \Leftrightarrow (a \subset b \text{ et } b \subset a)$

II.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 en 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 !

Paradoxe de Russel : *Il n'existe pas d'ensemble contenant tous les ensembles.* En effet, soit a un tel ensemble. Alors d'après l'axiome de sélection, $b = \{x \in a/x \notin x\}$ existe : on a $b \in b \Leftrightarrow b \notin b$ absurde !

Théorème 1.4 (Ensemble vide)

Il existe un **unique** ensemble n'ayant aucun élément. On le note $\emptyset$ et on a $\forall x, \emptyset \subset x$.

Démonstration : Soit a un ensemble. $\emptyset = \{x \in a/x \neq x\}$ n'a aucun élément. $\forall x, (y \in \emptyset \Rightarrow y \in x)$ donc $\forall x, \emptyset \subset x$. Soit Ω un ensemble n'ayant aucun élément. D'après ce qui précède, on a $\emptyset \in \Omega$. On a également $\forall x, (y \in \Omega \Rightarrow y \in \emptyset)$ donc $\forall x, \Omega \subset \emptyset$. Finalement, d'après le théorème 1.3 de double inclusion (page 3) on déduit que $\Omega = \emptyset$.

Définition 1.5 (Intersection) Soit a un ensemble non vide. On définit **l'intersection des éléments de** a par :

$$\bigcap_{x \in a} x = \{y \in c/x \in a \Rightarrow y \in x\}$$

où $c \in a$.

Remarque : Cette définition ne dépend pas de l'ensemble $c \in a$ choisi. En effet, soient $c_1, c_2 \in a$. Soit $z \in \{y \in c_1/x \in a \Rightarrow y \in x\}$. Comme $c_2 \in a$ on déduit que $z \in c_2$: donc $z \in \{y \in c_2/x \in a \Rightarrow y \in x\}$. En permutant 1 et 2 on obtient l'inclusion inverse.

On peut aussi définir la différence, la différence symétrique et le complémentaire d'un ensemble par rapport à un autre.

II.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 $\{\emptyset, \{\emptyset\}\}$.

II.4 Axiome de la réunion

Axiome

$\forall a, \exists b, [x \in b \Leftrightarrow (\exists c \in a, x \in c)]$. b est alors noté $\boldsymbol{\bigcup_{c \in a} c}$.

Exemple : $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ existe : c'est $\{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\}$.

Définition 1.6 Si x est un ensemble, on définit $\boldsymbol{x^+ = x \cup \{x\}}$.

II.5 Axiome de l'ensemble des parties

Axiome

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

Définition 1.7 (Couple) Soient a, b deux ensembles. Le **couple** (a, b) est définit par $(a, b) = \{\{a\}, \{a, b\}\}$ ($= \{\{a\}\}$ si $a = b$).

Proposition 1.8

Si $(a, b) = (c, d)$ alors $a = c$ et $b = d$.

Définition 1.9 (Produit cartésien) Soient A et B deux ensembles. On définit le **produit cartésien** de A et B par $A \times B = \{x \in \mathcal{P}(\mathcal{P}(A \cup B)) / \exists a \in A, \exists b \in B, x = (a, b)\}$

Remarque : Si $a \in A$ alors $\{a\} \in \mathcal{P}(A) \subset \mathcal{P}(A \cup B)$. On a également $\{a, b\} \in \mathcal{P}(A \cup B)$ donc $x = (a, b) \underset{\text{(def)}}{=} \{\{a\}, \{a, b\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$.

Définition 1.10 (Relation binaire) Soient A et B deux ensembles. Une **relation binaire** $\mathcal{R}$ entre les éléments de A et B est une **partie de** $A \times B$. Pour $x \in A$ et $y \in B$ on note $x \mathcal{R} y$ <u>ssi</u> $(x, y) \in \mathcal{R}$.

Exemple : Soit A un ensemble. On pose $\mathcal{R} = \{(x, y) \in A \times A / x = y\}$ (que l'on raccourci abusivement, puisqu'on ne respecte pas l'axiome de sélection, en $\{(x, x) / x \in A\}$). On a $x \mathcal{R} y$ <u>ssi</u> $x = y$.

On peut alors définir les relations d'ordre et d'équivalence, ainsi que les relations fonctionnelles.

Définition 1.11 (Fonction, application) Soit $\mathcal{R}$ une relation binaire sur $A \times B$. $\mathcal{R}$ est une **relation fonctionnelle** <u>ssi</u> $\forall x \in A, \forall y, z \in B, [(x \mathcal{R} y$ et $x \mathcal{R} z) \Rightarrow y = z]$. On note $\operatorname{dom} \mathcal{R} = \{x \in A / \exists y \in B, x \mathcal{R} y\}$. Si de plus $\operatorname{dom} \mathcal{R} = A$, $\mathcal{R}$ est une **application**.

Définition 1.12 (Ensemble quotient) Soit A un ensemble et $\mathcal{R}$ un relation d'équivalence sur A (comprendre sur $A \times A$). Pour $x \in A$, on note $\bar{x} = \{y \in A / x\mathcal{R}y\}$ (élément de $\mathcal{P}(A)$). On définit **l'ensemble quotient** par $\boldsymbol{A/\mathcal{R}} = \{\bar{x}/x \in A\}$.

Application : Soit $f : A \to B$. On pose $x\mathcal{R}y$ <u>ssi</u> $f(x) = f(y)$. On définit alors trois applications : $\left| \begin{array}{l} i \longrightarrow f(A) \;\; y, \\ B \longmapsto y \end{array} \right| \begin{array}{l} \hat{f} \longrightarrow A/\mathcal{R} \;\; f(x), \\ f(A) \longmapsto \bar{x} \end{array} \left| \begin{array}{l} \pi \longrightarrow A \qquad \bar{x}. \\ A/\mathcal{R} \longmapsto x \end{array} \right.$ La **décomposition canonique** de f est $\boldsymbol{f = i \circ \hat{f} \circ \pi}$.

Remarques :
- $x \in x$ n'est pas impossible ;
- on peut avoir $x \in y$ et $y \in x$.

On ajoute un axiome pour éviter ces choses "bizarres" : **l'axiome de fondement**.

III Deux autres axiomes

III.1 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 $\boldsymbol{x \cap \{x\} = \emptyset}$. Donc si $x \in x$, alors comme $x \in \{x\}$ on déduit $x \in (x \cap \{x\})$ ce qui est absurde ! Donc **$\boldsymbol{x \in x}$ 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 tout aussi absurde : donc **$\boldsymbol{x \in y}$ et $\boldsymbol{y \in x}$ est impossible**.

III.2 Axiome de l'infini

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

Axiome

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

On note alors :
- $0 = \emptyset$;
- $1 = 0 \cup \{0\} = \emptyset \cup \{\emptyset\} = \{\emptyset\}$;
- $2 = 1 \cup \{1\} = \{\emptyset, \{\emptyset\}\}$;
- $3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$...

Théorème 1.13

Il existe un **plus petit élément vérifiant l'axiome précédent** : on le note $\mathbb{N}$.

Démonstration : D'après l'axiome précédent, il existe un ensemble W tel que (1) $\emptyset \in W$ et $x \in W \Rightarrow x^+ \in W$. On pose :

$$\mathbb{N} = \bigcap_{\substack{c \subset W \\ c \text{ vérifie } (1)}} c.$$

Soit X vérifiant la propriété (1). $X \cap W$ vérifie encore (1) : comme $X \cap W \subset W$ on a donc $\mathbb{N} \subset X \cap W$ (par définition de $\mathbb{N}$). Donc $\mathbb{N} \subset \boldsymbol{X}$: $\mathbb{N}$ est donc bien le plus petit ensemble vérifiant la propriété (1).

Corollaire (Récurrence)

Soit $S \subset \mathbb{N}$, tel que $\emptyset \in S$ et $x \in S \Rightarrow x^+ \in S$. Alors $\boldsymbol{S = \mathbb{N}}$.

Théorème 1.14

On a les assertions suivantes :

1. $\forall n \in \mathbb{N}, \forall x \in n, n \not\subset x$
2. $\forall n \in \mathbb{N}, n \notin n$
3. $\forall n \in \mathbb{N}, \forall x \in n, x \subset n$
4. $\forall n \in \mathbb{N}, \forall m \in \mathbb{N}, [m \subset n \Rightarrow (m = n \text{ ou } m \in n)]$

Remarques :

– Le deuxième item est une conséquence de l'axiome de fondement.

– Le quatrième item correspond à :

$$\forall n \in \mathbb{N}, \forall m \in \mathbb{N}, [m \leqslant n \Rightarrow (m = n \text{ ou } m \in n)].$$

Démonstration : On démontre 1. par récurrence (les autres se démontrent de la même façon). Pour $n \in \mathbb{N}$ on considère $\mathcal{P}(n)$: $\forall x \in n, n \not\subset x$.

- Vrai pour $n = 0$ car comme $x \in \emptyset$ est faux on a $x \in \emptyset \Rightarrow \emptyset \not\subset x$.
- Pour $n \in \mathbb{N}$ on suppose $\mathcal{P}(n)$. Soit $x \in n^+ = n \cup \{n\}$: on a soit $x = n$ soit $x \in n$. Si $x = n$, alors comme $n \notin n$ on a $n \notin x$ et donc $n^+ = n \cup \{n\} \not\subset x$. Si $x \in n$, alors par hypothèse de récurrence on a $n \notin x$ donc $n^+ = n \cup \{n\} \not\subset x$. Dans tous les cas, on retrouve $\mathcal{P}(n^+)$.

Par récurrence, la propriété est donc prouvée !

Corollaire (Existence et unicité du prédécesseur)

$\forall n \in \mathbb{N}, (n \neq 0 \Rightarrow \exists!\, m \in \mathbb{N}, m^+ = n)$.

Démonstration :
Existence : Si pour $n \in \mathbb{N} - \{\emptyset\}, \nexists m \in \mathbb{N}, m^+ = n$, alors $\mathbb{N} - \{\emptyset\}$ vérifie encore la propriété (1) : ($\emptyset \in \mathbb{N} - \{\emptyset\}$ et ($x \in \mathbb{N} - \{\emptyset\} \Rightarrow x^+ \in \mathbb{N} - \{\emptyset\}$) ce qui est impossible par minimalité de $\mathbb{N}$!

Unicité : Si ${m_1}^+ = {m_2}^+$ alors $m_1 \cup \{m_1\} = m_2 \cup \{m_2\}$ donc $\begin{cases} m_1 \subset m_2 \\ m_2 \subset m_1 \end{cases}$

et d'après le 3. du théorème précédent on peut affirmer que $\begin{cases} m_1 \subset m_2 \\ m_2 \subset m_1 \end{cases}$ et donc $\boldsymbol{m_1 = m_2}$.

Théorème 1.15
$\forall n, m \in \mathbb{N}, (m = n$ ou $m \in n$ ou $n \in m)$ et un seul est vérifié.

Démonstration : On va encore raisonner par récurrence avec $\mathcal{P}(n)$: $\forall m \in \mathbb{N}, m = n$ ou $m \in n$ ou $n \in m$.

- $n = 0$: si $m = 0$ on a $n = m$ et sinon on démontre par récurrence sur m que $0 \in m$.
- Pour $n \in \mathbb{N}$ on suppose $\mathcal{P}(n)$. Soit $m \in \mathbb{N}$: on a trois cas :
  1. si $m = n$, alors $m \in n \cup \{n\} = n^+$;
  2. de même si $m \in n$ alors $m \in n^+$;
  3. si $n \in m$, alors on a $\{n\} \subset m$ et d'après le 3. du théorème précédent on a également $n \subset m$: finalement, on a $n^+ = n \cup \{n\} \subset m$ et d'après le 4. du théorème précédent $n^+ = m$ ou $n^+ \in m$. On retrouve bien $\mathcal{P}(n^+)$.

Par récurrence, la propriété est donc prouvée !

Remarque : On peut donc définir une relation d'ordre total sur $\mathbb{N}$.

Théorème 1.16
Si $\emptyset \subsetneq A \subset \mathbb{N}$ alors A a un plus petit élément.

Démonstration : On peut montrer la contraposée par récurrence : soit A est une partie de $\mathbb{N}$ qui n'a pas de plus petit élément. Montrons que A est vide ; on considère la propriété $\mathcal{P}(n)$: $\forall k \leqslant n, k \notin A$.

- Comme A n'a pas de plus petit élément, $0 \notin A$ et on a donc $\mathcal{P}(0)$.
- Pour $n \in \mathbb{N}$ on suppose $\mathcal{P}(n)$. On a donc $\forall k \leqslant n, k \notin A$. Si $n^+ \in A$, alors n^+ est le plus petit élément de A ce qui est impossible car A n'a pas de plus petit élément. Donc $n^+ \notin A$ et on retrouve donc $\mathcal{P}(n^+)$.

Par récurrence, on a donc $\forall n \in \mathbb{N}, \forall k \leqslant n, k \notin A$ donc **$\boldsymbol{A}$ est vide**. (Si A possède un élément m, alors c'est absurde car d'après $\mathcal{P}(m)$, qui est vraie d'après la récurrence, on a $m \notin A$.)

Définition 1.17 (Addition, multiplication) Soit $k \in \mathbb{N}$. On peut définir deux application s_k et p_k :

1. $\begin{cases} s_k(0) = k \\ \forall n \in \mathbb{N}, s_k(n^+) = s_k(n)^+ \end{cases}$
2. $\begin{cases} p_k(0) = 0 \\ \forall n \in \mathbb{N}, p_k(n^+) = s_k(p_k(n)) \end{cases}$.

Moralement, on a $s_k(n) = n + k$ et $p_k(n) = kn$.

Théorème 1.18 (Suites récurrentes)
Soit E un ensemble, $f : E \to E$ une application et $a \in E$. Alors :

$$\exists! u : \mathbb{N} \to E, \begin{cases} u(0) = u_0 = a \\ \forall n \in \mathbb{N}, u(n+1) = u_{n+1} = f(u_n) \end{cases}$$

IV Construction de $\mathbb{Z}$ et $\mathbb{Q}$

IV.1 Construction de l'anneau $\mathbb{Z}$ des entiers

On munit $\mathbb{N} \times \mathbb{N}$ de la relation suivante :

$$(a,b)\mathcal{R}(c,d) \text{ \underline{ssi} } a + d = b + c$$

$\mathcal{R}$ est une relation d'équivalence compatible avec l'addition sur $\mathbb{N} \times \mathbb{N}$. On va définir sur $\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\mathcal{R}$ deux lois, l'addition puis la multiplication.

1. L'addition $+ : \overline{(a,b)}, \overline{(c,d)} \longmapsto \overline{(a,b)} + \overline{(c,d)} = \overline{(a+d, c+b)}$. Sur $\mathbb{Z}$, la loi $+$ est une loi de composition interne, est associative, commutative, admet un élément neutre $\overline{(0,0)} = \overline{(a,a)}$ ($\forall a \in \mathbb{N}$) ; le symétrique d'un élément $\overline{(a,b)}$ est $\overline{(b,a)}$: on a en effet $\overline{(a,b)} + \overline{(b,a)} = \overline{(a+b, a+b)} = \overline{(0,0)}$. Finalement, $(\mathbb{Z}, +)$ **est un groupe abélien**.
   $\left| \begin{array}{l} \mathbb{N} \longrightarrow \mathbb{Z} \\ n \longmapsto \overline{(n,0)} \end{array} \right.$ est une injection : en identifiant $\mathbb{N}$ à son image, on obtient $\mathbb{N} \subset \mathbb{Z}$.
   On note $-n = \overline{(0,n)}$: on a donc $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$.
2. La multiplication $\times : a, b \in \mathbb{Z} \longmapsto a \times b =$ produit des signes $\cdot |a| \cdot |b|$ où $|n| = \begin{cases} n \text{ si } n \in \mathbb{N} \\ -n \text{ si } -n \in \mathbb{N} \end{cases}$.
   $(\mathbb{Z}, +, \times)$ est alors un **anneau commutatif unitaire intègre**, d'unité $1 = \overline{(1,0)}$.

IV.2 Construction du corps $\mathbb{Q}$ des fractions de $\mathbb{Z}$

On va définir $\mathbb{Q}$ encore comme un ensemble quotient. On munit $\mathbb{Z} \times \mathbb{Z}^*$ de la relation suivante :

$$(a, b)\mathcal{R}(c, d) \text{ \underline{ssi} } ad = cb.$$

On définit deux lois sur $\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z}^*)/\mathcal{R}$:

1. $\overline{(a, b)} + \overline{(c, d)} = \overline{(ad + bc), bd)}$
2. $\overline{(a, b)} \cdot \overline{(c, d)} = \overline{(ac, bd)}$

On vérifie que ces deux lois sont indépendantes des représentants choisis et qu'elles munissent $\mathbb{Q}$ d'une structure **d'anneau commutatif unitaire intègre**. On a de plus $0_{\mathbb{Q}} = \overline{(0, a)}$ $\forall a \in \mathbb{Z}^*$ et $1_{\mathbb{Q}} = \overline{(a, a)}$ $\forall a \in \mathbb{Z}^*$. Par conséquent, si $a \neq 0$, $\overline{(a, b)}$ a un inverse qui est $\overline{(b, a)}$ puisque $\overline{(a, b)} \cdot \overline{(b, a)} = \overline{(ab, ba)} = \overline{(ab, ab)} = 1_{\mathbb{Q}}$. Tout élément non nul possède un inverse dans l'anneau $\mathbb{Q}$ donc $\mathbb{Q}$ **est un corps (commutatif)**. Finalement, on plonge $\mathbb{Z}$ dans $\mathbb{Q}$ par l'injection $n \longmapsto \overline{(n, 1)}$.

Chapitre 2

Axiome du choix

Introduction

Après tous les axiomes qu'on vient de définir, on pourrait se dire qu'on en a assez. Mais il se trouve qu'on a assez vite besoin de quelque chose d'autre pour prouver certaines choses qu'on désirerait être vraies, comme l'existence d'une base dans toute espace vectoriel ou bien celle d'un idéal maximal dans un anneau, pour exhiber une partie de $\mathbb{R}$ non Lebesgue-mesurable ou tout simplement pour avoir qu'un produit quelconque d'ensembles non vides et non vide ! Pour ce faire, on introduit un autre axiome : **l'axiome du choix**. On peut démontrer (mais cela ne sera pas fait ici !) que l'apport de cet axiome à l'axiomatique précédente n'est pas contradictoire (on n'aboutit à aucune contradiction), et qu'on obtient la même théorie au bout du compte.

I Différents énoncés de l'axiome du choix

I.1 Avec le produit

Soit $(A_i)_{i\in I}$ une famille d'ensembles non vides. Tout d'abord, qu'est-ce que c'est ? C'est une application qui à $i \in I$ associe un ensemble A_i : mais cela ne suffit pas. En effet, dans le chapitre précédent on a vu qu'une application possède un ensemble d'arrivée. La famille $(A_i)_{i\in I}$ est donc une application $\left| \begin{array}{l} I \longrightarrow E \\ i \longmapsto A_i \end{array} \right.$ avec $\forall i, A_i \in E$. On s'intéresse maintenant au produit des A_i :

$$\prod_{i\in A} A_i = \{(a_i)_{i\in I} / \forall i \in I, a_i \in A_i\}$$

avec $\forall i,\ \ (a_i)_{i\in I} : \left| \begin{array}{l} I \longrightarrow \cup_{i\in I} A_i \\ i \longmapsto a_i \end{array} \right.$. On a supposé que les A_i étaient non vides :

on a en revanche aucun outil qui nous permet d'affirmer que $\prod_{i \in A} A_i \neq \emptyset$.

Axiome (Axiome du choix, version 1)
Si $\forall i, A_i \neq \emptyset$, alors $\prod_{i \in I} A_i \neq \emptyset$.

I.2 Avec une fonction de choix

Axiome (Axiome du choix, version 2)
Soit X un ensemble non vide d'ensembles non vides : $X \neq \emptyset$ et $\forall A \in X, A \neq \emptyset$. Alors il existe une application $g : X \longrightarrow \bigcup_{A \in X} A$, avec $\forall A \in X, g(A) \in A$. g s'appelle une **fonction de choix**.

Remarque : Si X est dénombrable, ie $X = \{A_k / k \in \mathbb{N}\}$ alors cet axiome est **inutile**, puisqu'on peut définir $g(A_k)$ par **récurrence**. En effet, pour $k \in \mathbb{N}$, on a pas besoin de cet axiome pour choisir un $k \in A_k$. En résumé, l'axiome du choix est nécessaire quand on doit **choisir dans un nombre non dénombrable d'ensembles**, car on ne peut alors plus construire la fonction de choix par récurrence. [1]

Démonstration : (Équivalence entre la version 1 et la version 2)

$\underline{\Longrightarrow :}$ Soit X un ensemble d'ensembles non vides. On pose $I = X$ et $E = X$. On a une application $\left| \begin{array}{l} I \longrightarrow E \\ A \longmapsto A \end{array} \right.$. L'axiome du choix version 1 nous garantit qu'il existe un $a \in \prod_{A \in X} A$: on a $a = (a_A)_{A \in X}$. L'application $g : \left| \begin{array}{l} X \longrightarrow \cup_{A \in X} A \\ A \longmapsto a_A \end{array} \right.$ est une fonction de choix.

$\underline{\Longleftarrow :}$ Soit $(A_i)_{i \in I}$ une famille d'ensembles non vides ($I \neq \emptyset$). On pose $X = \{A_i / i \in I\}$: c'est un ensemble non vide d'ensembles non vides. D'après la version 2 de l'axiome du choix, on peut trouver une fonction de choix $g : X \longrightarrow \bigcup_{i \in I} A_i$; on a $\forall i \in I, g(A_i) \in A_i$. Pour $i \in I$, on pose $a_i = g(A_i)$, puis $a = (a_i)_{i \in I}$: ce dernier est un élément de $\prod_{i \in I} A_i$ qui est par conséquent non vide !

1. Bertrand Russell disait à propos de l'axiome du choix : "Pour choisir une chaussette plutôt que l'autre pour chaque paire d'une collection infinie, on a besoin de l'axiome du choix. Mais pour les chaussures, ce n'est pas la peine". En effet, quand on est en face d'une paire de chaussette on peut choisir laquelle on va mettre en premier, mais il est impossible d'avoir un procédé général nous permettant de choisir à chaque fois laquelle on va mettre en premier ! Avec les chaussures, on peut en revanche décréter qu'on va toujours mettre la gauche en premier. (D'après Wikipédia)

I.3 Avec un théorème d'inversibilité pour les fonctions

Axiome (Axiome du choix, version 3)
Pour tous ensembles A et B, pour toute fonction $f : A \longrightarrow B$ surjective, il existe une fonction $g : B \longrightarrow A$, $f \circ g = \mathrm{Id}_B$.

Démonstration : $\underline{\Longrightarrow}$: Soit $f : A \to B$ une surjection. Par définition, on a donc $\forall b \in B, \exists a \in A, b = f(a)$. Quelque chose de facile ici (c'est une grosse arnaque) est de poser $g(b) = a$ et de dire qu'on a alors bien $f \circ g = \mathrm{Id}_B$. Mais **on a pas le droit de faire cela car on ne définit pas ainsi une application !** En effet, $b \in B$ peut a priori avoir plusieurs antécédents dans A. Voici la bonne façon de procéder ; on va utiliser l'axiome du choix. $\forall b \in B, f^{-1}(\{b\}) \neq \emptyset$ (car f est surjective). On pose $X = \{f^{-1}(\{b\})/b \in B\}$; c'est un ensemble non vide d'ensembles non vides. L'axiome du choix nous fourni une fonction de choix $\gamma : \left| \begin{array}{l} X \longrightarrow A \\ f^{-1}(\{b\}) \longmapsto \gamma[f^{-1}(\{b\})] \in f^{-1}(\{b\}) \end{array} \right.$. On pose alors $\forall b \in B, g(b) = \gamma[f^{-1}(\{b\})]$ et on a inversé à droite notre fonction f !

$\underline{\Longleftarrow}$: Soit $X \neq \emptyset$ un ensemble d'ensembles non vides. On pose :

$$Y = \bigcup_{A \in X} A \times \{A\}.$$

Remarque : Si $A = \{x, y\}$ alors $A \times \{A\} = \{(x, A), (y, A)\}$.

On a (on revient dans le cas général) $(a, A) \in Y \iff \begin{cases} A \in X \\ a \in A \end{cases}$.

$f : \left| \begin{array}{l} Y \longrightarrow X \\ (a, A) \longmapsto A \end{array} \right.$ est une surjection car X est un ensemble d'ensembles non vides donc $\forall A \in X, A \neq \emptyset$: ainsi $\exists a \in A, A = f[(a, A)]$. Par hypothèse, $\exists g : X \longrightarrow Y, g(A) = (g_1(A), g_2(A))$ avec $f \circ g = \mathrm{Id}_X$. On a donc $\forall A \in X, f \circ g(A) = f[(g_1(A), g_2(A))] = \boldsymbol{g_2(A) = A}$. Or, g est à valeurs dans Y : on a donc $\forall A \in X, g_1(A) \in g_2(A)$ ie $g_1(A) \in A$. Finalement, $g_1 : \left| \begin{array}{l} X \longrightarrow \cup_{A \in X} A \\ A \longmapsto g_1(A) \in A \end{array} \right.$ est une fonction de choix !

II Lemme de Zorn

On a déjà donné plusieurs formulations équivalentes de l'axiome du choix. Dans ce paragraphe, on va en établir une autre qui est moins "évidente" et qui est souvent utilisée quand l'axiome du choix est accepté [2].

2. Par exemple, on utilise le lemme de Zorn pour prouver l'existence des idéaux maximaux d'un anneau ou bien pour prouver l'existence d'une base dans un espace vectoriel (cet exemple sera repris dans la suite). Cependant, ce sont uniquement des résultats d'existence car le lemme de Zorn ne fournit pas de construction.

Pour commencer, quelques définitions ; X désigne un ensemble ordonné par la relation $\leqslant$.

II.1 Définitions et lemme

Définition 2.1 (Plus grand élément) On dit que $a \in X$ est un **plus grand élément** de X si $\forall b \in X, b \leqslant a$.

Définition 2.2 (Élément maximal) On dit que $a \in X$ est un **élément maximal** si $\forall b \in X, (b \geqslant a \Rightarrow b = a)$.

Remarques :

– Un plus grand élément n'existe pas forcément, un élément maximal non plus.

– Un plus grand élément est un élément maximal.

– Si $a \in X$ est un élément maximal, a n'est pas forcément un plus grand élément : en effet, rien ne garantit qu'on peut comparer tous les éléments de X à a (l'ordre $\leqslant$ n'est pas supposé total). La réciproque de la remarque précédente n'est donc pas vérifiée.

– On définit de manière analogue un plus petit élément et un élément minimal.

Pour illustrer la différence entre les notions de plus grand élément et d'élément maximal, il faut prendre un ordre partiel : par exemple $\mathbb{N}$ avec la relation $|$. Les nombres premiers sont des éléments minimaux de $\mathbb{N} \setminus \{0, 1\}$ (par définition des nombres premiers) mais ce ne sont pas des plus petits éléments : par exemple on a $\forall n \geqslant 2, (n|3 \Rightarrow n = 3)$ donc 3 est un élément minimal mais ce n'est pas un plus petit élément car on a pas $3|2$ (ni $3|4$ par ailleurs).

Définition 2.3 (Majorant) Si Y est une partie de X, on dit que $x \in X$ est un **majorant** de Y si $\forall y \in Y, y \leqslant x$.

Définition 2.4 (Ensemble inductif) Soit X en ensemble non vide. On dit que X est **inductif** si toute partie **totalement ordonnée** de X possède un majorant dans X.

Exemple : Soit G un groupe non réduit à $\{0\}$. On considère $X = \{H \subset G/$H est un sous-groupe propre de $G\}$ muni de l'inclusion.

- X est non vide car $\{0\} \in X$ (c'est un sous-groupe propre car G n'est pas réduit à $\{0\}$).
- Soit Y une partie totalement ordonnée de X : un élément maximal (dans X) de Y est donné par $H = \bigcup_{K \in Y} K$. C'est bien un élément de X car Y est totalement ordonné : si $h_1, h_2 \in H$ alors $h_i \in Y_i \subset Y$

et comme Y est totalement ordonné on a $Y_1 \subset Y_2$ ou $Y_2 \subset Y_1$ d'où $h_1h_2 \in Y_1$ ou Y_2 et donc $h_1h_2 \in H$. De plus, on a $\forall K \in Y, K \subset H$ donc H est un majorant (dans X) de Y.

Finalement, X est inductif !

Lemme 2.5 (Lemme de Zorn)

Tout ensemble inductif possède un élément maximal.

Cet énoncé est **équivalent** à l'axiome du choix.

II.2 Deux applications

II.2.1 Le théorème de la base

Ce théorème dit : si E un espace vectoriel sur $\mathbb{K}$, $L \subset E$ une famille libre et $G \subset E$ une famille génératrice, alors $\boldsymbol{\exists B \subset E, L \subset B \subset G \cup L, B}$ **est une base de** $\boldsymbol{E}$.

Démonstration : Soit $X = \{F \subset E / L \subset F \subset G \cup L$ et F est une famille libre$\}$ muni de l'inclusion. C'est une partie non vide de E car $L \in F$.

Soit $Y \subset X$ une partie totalement ordonnée. On pose $\overline{F} = \bigcup\limits_{F \in Y} F$, et soit $(\lambda_a)_{a \in F} \in \mathbb{K}^{(\overline{F})}$ une famille de scalaires à support fini [3] telle que $\sum\limits_{a \in \overline{F}} \lambda_a a = 0$. Par hypothèse, $S = \{a \in \overline{F} / \lambda_a \neq 0\}$ est fini. Pour chaque a dans S, il existe un élément de Y qui contienne a : comme S est fini et que Y est totalement ordonnée on en déduit que $\exists F \in Y, F \supset S$. Par conséquent, la somme précédente peut être interprétée comme une combinaison linéaire d'éléments de F qui est une famille libre : les λ_a sont donc tous nuls. Finalement, $\overline{F}$ est une famille libre !

L'ensemble X est donc inductif : le lemme de Zorn nous garantit alors l'existence d'un élément $B \in X$ maximal. $B \in X$ donc B est libre et $L \subset B \subset G \cup L$. Si B n'est pas génératrice, alors il existe un élément $g \in G$ tel que g n'est pas une combinaison linéaire [4] d'éléments de B. Ainsi, $B \cup \{g\}$ est libre et $L \subset B \cup \{g\} \subset G \cup L$ contradiction de la maximalité de B (on a bien évidemment $B \subsetneq B \cup \{g\}$) ! Donc B est une famille génératrice et comme B est une famille libre c'est une base.

II.2.2 Autour d'une équation fonctionnelle

Un exercice classique consiste à rechercher toutes les fonctions $f : \mathbb{R} \to \mathbb{R}$ telles que :

(*i*) $\forall x, y \in \mathbb{R}, f(x+y) = f(x) + f(y)$;

3. C'est à dire que, sauf pour un nombre fini d'entre eux, ils sont tous nul.
4. Finie, par définition.

(ii) f est continue au moins en un point de $\mathbb{R}$.

Généralement, on procède ainsi :

- $f(0+0) = f(0) = f(0) + f(0)$ donc $f(0) = 0$;
- on montre par récurrence pour n entier supérieur à 1 que $\forall x \in \mathbb{R}, f(nx) = nf(x)$ et on en déduit $\forall n \in \mathbb{N}^*, f(n) = nf(1)$;
- $\forall x \in \mathbb{R}, f(x - x) = f(0) = 0 = f(x) + f(-x)$ donc f est impaire ;
- on en déduit l'image de tous les entiers relatifs ;
- on étend le résultat précédent aux rationnels en remarquant que $qf(\frac{p}{q}) = f(q\frac{p}{q}) = f(p) = pf(1)$ et donc $f(\frac{p}{q}) = \frac{p}{q}f(1)$: autrement dit, f est $\mathbb{Q}$-linéaire ;
- on sait que f est continue en $x_0 \in \mathbb{R}$: si $x \in \mathbb{R}, \lim_{h\to 0} f(x+h) = f(x) + \lim_{h\to 0} f(h) = f(x) + \lim_{h\to 0} f(x_0 + h) - f(x_0) = f(x) + f(x_0) - f(x_0) = f(x)$ donc f est continue sur $\mathbb{R}$;
- par densité de $\mathbb{Q}$ dans $\mathbb{Q}$ on conclut que f est $\mathbb{R}$-linéaire !

Cependant, le théorème de la base nous fourni un contre exemple si on ne suppose plus (ii) : $\mathbb{R}$ est un $\mathbb{Q}$-espace vectoriel, soit B une base, et $b_0 \in B$ fixé. B est une base donc $\forall x \in \mathbb{R}, \exists (\lambda_b)_{b\in B} \in \mathbb{Q}^{(B)}, x = \sum\limits_{b\in B} \lambda_b b$. On pose alors $f(x) = \lambda_{b_0}$: f est $\mathbb{Q}$ linéaire.

Soit $b_1 \in B \setminus \{b_0\}$. $-\frac{b_0}{b_1} \in \mathbb{R}$ donc $\exists (\lambda_n) \in \mathbb{Q}^{\mathbb{N}}, \lambda_n \underset{n\to\infty}{\longrightarrow} -\frac{b_0}{b_1}$. $(x_n = b_0 + \lambda_n b_1)_n$ tend vers 0 ; $f(x_n) = \lambda_{b_0} = 1 \neq f(0) = 0$ donc f n'est pas continue donc f n'est pas linéaire.

II.3 Démonstration de l'équivalence entre l'axiome du choix et le lemme de Zorn

II.3.1 L'axiome du choix implique le lemme de Zorn

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

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

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

$$c : \left| \begin{array}{l} \mathcal{P}(X) \setminus \emptyset \longrightarrow X \\ A \longmapsto c(A) \in A \end{array} \right. .$$

Pour $A \in \mathcal{E}$ on note $m(A) = c(\{\text{majorants stricts de } A\})$.

On va d'abord faire une première 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 $\emptyset \in \mathcal{E}$ et on pose alors $\boldsymbol{x_0 = m(\emptyset)}$. On construit alors la suite $(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$ telle que $\forall k \in \mathbb{N}, x_0 < ... < x_k$ et $x_{k+1} = m(\{x_0, ..., x_k\})$ (à moins que le processus ne s'arrête, auquel cas on aurait 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 possède un majorant $M \in X$ avec $M \notin E$. Par conséquent, E a au moins un majorant strict ie $E \in \mathcal{E}$; on note $x_\infty = m(E)$.

On voit bien qu'à faire des récurrences on ne va jamais finir ! On va donc attaquer le problème sous un autre angle : d'abord, quelques définitions.

Définition 2.6 (Segment) Soient S et A deux parties de X. On dit que S est un **segment** de A ssi

- $S \subset A$
- $\forall (x, y) \in A \times S, (x \leqslant y \Rightarrow x \in S)$.

Définition 2.7 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 le plus petit majorant strict de S dans B.

Exemples :

– $\{x_0, ..., x_k\}$ est un bon ensemble, dont les segments sont les $\{x_0, ..., x_l\}$ pour $0 \leqslant l \leqslant k$;

– $\{x_k/k \in \mathbb{N}\}$ est un bon ensemble, dont les segments sont les $\{x_0, ..., x_k\}$ pour $k \in \mathbb{N}$;

– $\{x_k/k \in \mathbb{N}\} \cup \{l\}$ est un bon ensemble, dont les segments sont les $\{x_0, ..., x_k\}$ pour $k \in \mathbb{N} \cup \{\infty\}$ (avec $x_\infty = l$).

On pose $U = \cup_{\text{B bon ensemble}} B$. On va montrer dans les paragraphes suivants que U est un bon ensemble. On aura alors :

- U est totalement ordonné (car c'est un bon ensemble), il admet donc un majorant (car $X \supset U$ est inductif) ;
- si tous les majorants sont dans U, alors il n'y en a qu'un (car U est totalement ordonné), m : si $X \ni n > m$ alors n est un majorant strict de U ce qui est absurde car $n \notin U$, donc m est un élément maximal de X ;
- sinon, U a au moins un majorant strict $m \in X \setminus U$: $U \cup \{m\}$ est alors un bon ensemble qui contient strictement U ce qui est absurde car c'est U le plus gros.

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

II.3.1.1 Les bons ensembles sont emboités

Soient A et B deux bons ensembles de X. Alors ***A* est un segment de *B* ou *B* est un segment de *A***.

On suppose A et B non vides. On a alors $m(\emptyset) \in A \cap B$ et donc $A \cap B \neq \emptyset$: soit $a \in A \cap B$. On note $I_{a,A} = \{x \in A/x < a\}$ et $I_{a,B} = \{x \in B/x < a\}$.

$I_{a,A}$ est un segment de A : en effet, il est inclus dans A et si, pour $x \in A, A \ni y \leqslant x$ alors $y \leqslant x < a$ donc $y < a$ et $y \in I_{a,A}$! De même, $I_{a,B}$ est un segment de B.

On pose $C = \{a \in A \cap B / I_{a,A} = I_{a,B}\}$: on va démontrer que C est un segment de A. Soit $x \in C, y \in A, y \leqslant x$. Il y a deux cas :

– Si $y = x$ alors comme $x \in C$ on a $y \in C$.
– Si $y \neq x$, on procède en deux étapes :

1. On a alors $y < x$: donc $y \in I_{x,A} = I_{x,B}$ car ce sont des segments, et donc $y \in B$: donc $\boldsymbol{y \in A \cap B}$.
2. On pose $I_{y,A} = \{z \in A / z < y\}$. Si $z \in I_{y,A}$, on a $z < y < x$ donc $z < x$ donc $z \in I_{x,A} = I_{x,B}$: on a donc $z \in B$ et $z < y$ donc $z \in I_{y,B}$. Finalement, $I_{y,A} \subset I_{y,B}$ et de même on a l'inclusion inverse. On a donc $\begin{cases} I_{y,A} = I_{y,B} \\ y \in A \cap B \end{cases}$ donc $\boldsymbol{y \in C}$.

C est donc bien un segment de A. De même, on démontre que C est aussi un segment de B.

On suppose que $C \neq A$ et $C \neq B$. Comme A est un bon ensemble, $m(C) \in A$ et c'est le plus petit majorant strict de C dans A ; de même, B est un bon ensemble donc $m(C) \in B$ et c'est le plus petit majorant strict de C dans B. On a donc $I_{m(C),A} = I_{m(C),B}$ donc $m(C) \in C$ ce qui est absurde ! Donc $C = A$ ou $C = B$ et comme C est un segment inclus dans $A \cap B$ on en déduit que **$\boldsymbol{A}$ est un segment de $\boldsymbol{B}$ ou $\boldsymbol{B}$ est un segment de $\boldsymbol{A}$**.

II.3.1.2 U est totalement ordonné

Soit $x, y \in U$. Par définition de U, il existe B_x et B_y des bons ensembles, tels que $x \in B_x$ et $y \in B_y$. D'après le paragraphe précédent, B_x est un segment de B_y ou B_y est un segment de B_x. On a donc $B_x \subset B_y$ ou $B_y \subset B_x$ et donc $x, y \in B_x$ ou $x, y \in B_y$: comme B_x et B_y sont des bons ensembles, ils sont totalement ordonnés et on a donc $x \leqslant y$ ou $y \leqslant x$. Donc U est totalement ordonné.

II.3.1.3 Les bons ensembles de U sont des segments

Soit A un bon ensemble de U. Soit $x \in A, y \in U$ avec $y \leqslant x$. Par définition de U, il existe un bon ensemble $B \ni y$. D'après le paragraphe II.3.1.1 on a deux cas :

– si A est un segment de B, alors A est un segment et $y \leqslant x$ donc $y \in A$;
– si B est un segment de A, alors $y \in B \subset A$ donc $y \in A$.

Finalement, $y \in A$ et **$\boldsymbol{A}$ est un segment de $\boldsymbol{U}$**.

II.3.1.4 U est un bon ensemble

On sait déjà que U est bien ordonné. Soit S un segment de U, $S \neq U$. Soit $z \in U \setminus S$. D'après la définition de U, il existe un bon ensemble A, tel que

$z \in A$. Par définition d'un segment, z est un majorant strict de S : S a donc des majorants stricts. Reste à montrer que $m(S)$ est le plus petit majorant strict de S dans U.

1. Soit $x \in S, y \in A, y \leqslant x$. $A \subset U$ donc $y \in U$ et comme S est un segment de U on a $y \in S$.
2. De plus, $U \supset S \ni x < z \in A$. A est un bon ensemble, c'est donc un segment de U d'après le paragraphe précédent : donc $x \in A$. Finalement, on a montré que $S \subset A$!

Finalement, 1 et 2 fournissent que S est un segment de A.

On a $A \ni z \in U \setminus S$ donc on a $S \neq A$. A est un bon ensemble, S est un segment de A différent de A donc $m(S)$ existe et $m(S) \in A \subset U$. Soit y un majorant strict de S dans U. Si $y < m(S)$, alors $y \in A$ (car A est un bon ensemble donc un segment de U, et $m(S) \in A$) ce qui est impossible car $m(S)$ est le plus petit majorant strict de S dans A. Donc $y \geqslant m(S)$ et $m(S)$ est le plus petit majorant strict de S dans U. Donc U est un bon ensemble !

II.3.2 Le lemme de Zorn implique l'axiome du choix

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

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

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

- $C \neq \emptyset$ donc $\exists a \in C$; on a de plus $a \neq \emptyset$ donc $\exists x \in a$. On pose $A = \{a\}$ et $c : \left| \begin{array}{l} A \longrightarrow \cup_{B \in C} B \\ a \longmapsto x \end{array} \right.$. On a $(A, c) \in X$ donc $\boldsymbol{X \neq \emptyset}$.
- Ordre sur X : $[(A_1, c_1) \leqslant (A_2, c_2)]$ <u>ssi</u> $[A_1 \subset A_2$ et $c_2|_{A_1} = c_1]$. Soit Y une partie totalement ordonnée de X. On pose $B = \cup_{A/\exists c,(A,c) \in Y} A$ et $d : B \to \cup_{A \in C} A$ définie ainsi : si $a \in B$, alors $\exists (A, c) \in Y, a \in A$; on pose alors $d(a) = c(a)$. 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, $\boldsymbol{(B, d)}$ **majore** $\boldsymbol{Y}$.

X est donc inductif, et d'après le lemme de Zorn X admet un élément maximal (A, c). Si $A = C$, c est une fonction de choix sur C. Supposons $A \neq C$: alors $\exists \tilde{a} \in C \setminus A$. $\tilde{a} \neq \emptyset$ donc $\exists x \in \tilde{a}$. On pose $\tilde{A} = A \cup \{\tilde{a}\}$, puis $\tilde{c} : \tilde{A} \to \cup_{B \in C} B$ avec $\tilde{c}(a) = \begin{cases} c(a) \text{ si } a \in A \\ x \text{ si } a = \tilde{a} \end{cases}$. $\tilde{c}$ est un prolongement strict de c : c'est une contradiction avec la maximalité de (A, c) !

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

II.4 Comparabilité d'ensembles

On souhaiterait avoir une sorte de relation d'ordre sur les ensembles : on veut noter $\text{card}(A) \leqslant \text{card}(B)$ si il existe une fonction $f : A \to B$ injective [5].

Remarque : On a : $\exists f : A \to B$ injective ssi $\exists g : B \to A$ surjective (le sens de droite à gauche nécessite, comme on l'a vu en I.3, l'axiome du choix).

II.4.1 "Anti-symétrie" : théorème de Cantor-Bernstein

Que se passe-t-il s'il existe deux fonction injectives $f : A \to B$ et $g : B \to A$? Le théorème de Cantor-Bernstein affirme qu'alors il existe une fonction $h : A \to B$ bijective [6] (on va le voir d'ici quelques lignes).

Lemme 2.8
Soit E un ensemble, et $\phi : \mathcal{P}(E) \to \mathcal{P}(E)$ une application croissante [7]. Alors $\boldsymbol{\phi}$ **possède un point fixe** : $\exists C \in \mathcal{P}(E), \phi(C) = C$.

Démonstration : On note $\mathcal{E} = \{M \in \mathcal{P}(E) / M \subset \phi(M)\}$. $\emptyset \subset \phi(\emptyset)$ donc $\emptyset \in \mathcal{E}$ donc $\mathcal{E} \neq \emptyset$.

On pose $C = \cup_{M \in \mathcal{E}} M$.

- 1. $\forall M \in \mathcal{E}, M \subset \phi(M)$ d'où $\boldsymbol{C \subset \bigcup_{M \in \mathcal{E}} \phi(M)}$.
  2. Pour $M \in \mathcal{E}, M \subset C$ donc $\phi(M) \subset \phi(C)$ (car ϕ est croissante). On a donc $\boldsymbol{\bigcup_{M \in \mathcal{E}} \phi(M) \subset \phi(C)}$.

  Finalement on a $\boldsymbol{C \subset \phi(C)}$ donc $C \in \mathcal{E}$.
- ϕ est croissante : on a donc $\phi(C) \subset \phi(\phi(C))$. Cela signifie que $\phi(C) \in \mathcal{E}$. Par définition de C, on a donc $\boldsymbol{\phi(C) \subset C}$.

Par double inclusion on a donc $C = \phi(C)$, c'est-à-dire que $\boldsymbol{\phi}$ **possède un point fixe**.

Théorème 2.9 (Théorème de Cantor-Bernstein)
Soient A et B deux ensembles tels qu'il existe deux fonction $f : A \to B$ et $g : B \to A$ injectives. Alors il existe une fonction $h : A \to B$ bijective.

Démonstration : La fonction f induit une bijection sur $f(A)$. On a en outre pour toute partie $E \subset A$, f induit une bijection de E sur $f(E)$. On va

5. En fait, on ne pourra jamais avoir de relation d'ordre car il n'y a pas d'ensemble de tous les ensembles (comme le montre le paradoxe de Russel, qu'on a cité en II.2) et donc on ne peut pas définir de relation sur les ensembles.

6. Si l'on suppose les fonctions f et g non plus injectives mains surjectives, on arrive à la même conclusion mais elle nécessite cette fois l'axiome du choix (conséquence directe de la remarque précédente).

7. Au sens de l'inclusion : $A \subset B \Rightarrow \phi(A) \subset \phi(B)$.

chercher une partie E telle que g induise une bijection de $[f(E)]^c$ sur E^c : si on trouve une telle partie, alors la fonction $h : A \to B$ définie par $\forall x \in A, h(x) = \begin{cases} f(x) \text{ si } x \in E \\ g^{-1}(x) \text{ si } x \notin E \end{cases}$ sera une bijection !

On cherche une partie $E \subset A$ qui vérifie $g([f(E)]^c) = E^c$, soit $[g([f(E)]^c)]^c = E$. Posons $\phi : \left| \begin{array}{l} \mathcal{P}(A) \longrightarrow \mathcal{P}(A) \\ E \longmapsto [g([f(E)]^c)]^c \end{array} \right.$. Par composition de deux fonctions décroissantes [8], on en déduit que ϕ est croissante. D'après le lemme précédent, ϕ admet un point fixe E qui vérifie donc $g([f(E)]^c) = E^c$. La fonction $g|_{[f(E)]^c}$ reste injective, et induit une bijection sur $g([f(E)]^c) = E^c$: par conséquent, g induit une bijection de $[f(E)]^c$ sur E^c et c'est ce qu'on voulait !

II.4.2 "Définition" : comparer deux ensembles

Soient A et B deux ensembles : sont-ils comparables ?

Théorème 2.10

Soient A et B deux ensembles. Alors il existe une fonction $f : A \to B$ injective ou $g : B \to A$ injective.

La réponse à la question est donc : oui !

Démonstration : On suppose que A et B sont non vides, sinon le résultat est trivial. On note $X = \{(C, f)/C \subset A \text{ et } f : C \to B \text{ injective }\}$. On munit X de l'ordre usuel, qu'on a déjà utilisé : $[(C_1, f_1) \leqslant (C_2, f_2)]$ <u>ssi</u> $[C_1 \subset C_2$ et $f_2|_{C_1} = f_1]$. Comme A et B sont non vides, $\exists a \in A, \exists b \in B$. Avec $C = \{a\}$ et $f(a) = b$ on a bien $(C, f) \in X$ donc $X \neq \emptyset$. En refaisant la même démonstration que pour le lemme de Zorn (au II.3.2, page 19) on obtient que X est un ensemble inductif : il admet donc un élément maximal (C, f).

- Si $C = A$, alors f est une injection de A dans B et c'est gagné !
- Sinon, $C \subsetneq A$. Si $f(C) \neq B$, alors $\exists a \in A \setminus C, \exists b \in B \setminus f(C)$ et $\exists \tilde{f} : \tilde{C} = C \cup \{a\} \to B$ définie par $\tilde{f}(x) = \begin{cases} x \text{ si } x \in C \\ b \text{ si } x = a \end{cases}$ reste injective ce qui est impossible par maximalité de (C, f) ! Donc $f(C) = B$: f induit donc une bijection de C sur B, et en notant g son inverse, $g : B \to C$ est une bijection et donc g induit une injection de B sur A. C'est gagné !

8. Le passage au complémentaire étant décroissant : $A \subset B \Rightarrow A^c \supset B^c$.

Chapitre 3

Construction des nombres réels

Introduction [1]

On suppose construits $\mathbb{N}$, $\mathbb{Q}$, ce dernier étant un corps ordonné. Une approche intuitive de la construction des nombres réels est de dire qu'un nombre réel s'écrit sous la forme $n + 0.a_1a_2...a_i$ avec $n \in \mathbb{Z}$ et $a_i \in [\![0, 9]\!]$. L'inconvénient de cette description est qu'on a des choses du type $0.333... + 0.666... = 0.999... = 1$! [2]

Il existe dans la littérature deux constructions célèbres :

– les coupures de Dedekind (1858) ;

– l'utilisation des suites de Cauchy.

Nous allons voir entre autre que chaque construction a ses avantages.

I Coupures de Dedekind

Avantage : On va obtenir gratuitement la propriété de la borne supérieure.

I.1 Définition

Idée : Chaque $r \in \mathbb{Q}$ coupe $\mathbb{Q}$ en deux parties :

$$A_r = \{q \in \mathbb{Q}/q < r\} \text{ et } B_r = \{q \in \mathbb{Q}/q > r\}.$$

On appelle alors (A_r, B_r) une **coupure rationnelle** de $\mathbb{Q}$. On peut encore couper $\mathbb{Q}$ en (A, B) avec $Q = A \cup B$ et $\forall (a, b) \in A \times B, a < b$. Par exemple, $A = \{q \in \mathbb{Q}/q < 0 \text{ ou } (q \geqslant 0 \text{ et } q^2 \leqslant 2)\}$.

1. Ce chapitre est une copie pure et simple de l'article de Wikipédia intitulé "Construction des nombres réels". On pourra donc s'y référer pour de plus amples détails.

2. En effet, si on note $a = 0.999...$ alors on a $10a = 9.999... = 9 + a$ et donc $9a = 9$ puis $a = 1$.

Définition 3.1 (Coupure de Dedekind) Une **coupure de Dedekind** dans $\mathbb{Q}$ est un couple (A, B) de deux parties de $\mathbb{Q}$ tel que :
- $A \cup B = \mathbb{Q}$;
- $\forall (a, b) \in A \times B, a < b$ (on a donc $A \cap B = \emptyset$) ;
- A ne possède pas de plus grand élément.

On remarquera que $\begin{cases} A \cup B = \mathbb{Q} \\ A \cap B = \emptyset \end{cases} \Rightarrow (A, B)$ est une partition de $\mathbb{Q}$.

Définition 3.2 On définit $\mathbb{R}$ comme **l'ensemble des coupures de Dedekind**.

Remarques :

– A chaque rationnel r correspond la coupure $(A_r, \{r\} \cup B_r)$.

– Si (A, B) est une coupure, B a un plus petit élément $\Leftrightarrow (A, B)$ est une coupure rationnelle.

– $(A, B) = (A, A^c)$: une coupure de Dedekind est donc entièrement déterminée par la première composante du couple. Dorénavant, on désignera donc une coupure par cette seule composante.

I.2 Ordre, opérations

I.2.1 Ordre

On définit la relation $\leqslant$ par $\boldsymbol{A \leqslant B \Leftrightarrow A \subset B}$.

I.2.2 Addition

On définit l'addition par $\boldsymbol{A + B = \{a + b/(a, b) \in A \times B\}}$. En d'autres termes, $c \in A + B \Leftrightarrow \exists (a, b) \in A \times B, c = a + b$.

Remarque : $(\mathbb{R}, +)$ est alors un groupe abélien, d'élément neutre $A_0 = \{q \in \mathbb{Q}/q < 0\}$. L'opposé d'un élément A est :
- A_{-r} si $A = A_r$;
- $-(A^c)$ sinon (où $-$ et ^{c} sont les opérations ensemblistes).

I.2.3 Multiplication

Définition 3.3 On dit que A est un réel positif si $A \geqslant A_0$.

On définit alors la multiplication pour des réels A et B positifs par $\boldsymbol{AB = A_0 \cup \{ab/(a, b) \in (A \cap \mathbb{Q}_+) \times (B \cap \mathbb{Q}_+)\}}$. En d'autres termes, $c \in AB \Leftrightarrow \exists (a, b) \in (A \cap \mathbb{Q}_+) \times (B \cap \mathbb{Q}_+), c \leqslant ab$. La règle des signes permet de multiplier des réels quelconques.

Propriété 3.4

L'ensemble $\mathbb{R}$ des coupures muni des lois $+$ et $\times$ est un **corps totalement ordonné** et qui vérifie la **propriété de la borne supérieure**, c'est à dire que tout ensemble non vide majoré inclus dans $\mathbb{R}$ possède une borne supérieure.

Démonstration : On démontre d'abord l'existence d'un inverse pour toute coupure non nulle puis la propriété de la borne supérieure.

– Soit A une coupure strictement positive : son inverse pour la loi $\times$ est donné par
$$A^{-1} = \{q \in \mathbb{Q}/\exists a \in A \cap Q_+, aq < 1\}.$$
Si A est une coupure strictement négative, $-A$ est une coupure strictement négative et on a $A^{-1} = -[(-A)^{-1}]$.

– Soit $E \subset \mathbb{R}$ une partie non vide majorée. On remarque que $\bigcup_{A \in E} A$ est une borne supérieure !

II Avec les suites de Cauchy

Avantage : On va obtenir gratuitement que $\mathbb{R}$ est un corps complet.
Problème : Comment parler de suites de Cauchy sans parler de $\mathbb{R}$? Par exemple, on ne peut pas écrire $\forall \varepsilon > 0$.

II.1 Définition

Définition 3.5 (Suite de Cauchy) Soit K un corps totalement ordonné. On dit que $(a_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}}$ est une **suite de Cauchy**

$$\underline{\text{ssi}}\ \forall \varepsilon \in K_+^*, \exists N \in \mathbb{N}, \forall m \geqslant n \geqslant N, |a_m - a_n| \leqslant \varepsilon.$$

Définition 3.6 (Convergence) On reprend les notations de la définition précédente. On dit que la suite **converge** vers $l \in K$

$$\underline{\text{ssi}}\ \forall \varepsilon \in K_+^*, \exists N \in \mathbb{N}, \forall n \geqslant N, |a_n - l| \leqslant \varepsilon.$$

Remarque : On définit pour $x \in \mathbb{Q}, |x| = \max\{x, -x\}$.

On note $\mathcal{C}$ l'ensemble des suites de Cauchy à valeurs dans $\mathbb{Q}$. On peut voir que $(\mathcal{C}, +, .)$ est un anneau (où les opérations se font termes à termes). Cependant l'ensemble $\mathcal{C}$ est trop vaste ! Par exemple pour $q \in \mathbb{Q}$, il existe une infinité de suites de Cauchy qui convergent vers q. Pour réduire cet ensemble, on va le quotienter par une relation d'équivalence.

Relation d'équivalence sur $\mathcal{C}$: Si u et v sont deux suites à valeurs dans $\mathbb{Q}$, on définit

$$(u_n)\mathcal{R}(v_n) \Longleftrightarrow \lim_{n\to+\infty} u_n - v_n = 0.$$

On vérifie la réflexivité, la symétrie et la transitivité de $\mathcal{R}$ qui est donc une relation d'équivalence.

Définition 3.7 On définit $\mathbb{R}$ comme **l'ensemble des classes d'équivalence par $\mathcal{R}$ sur $\mathcal{C}$**. Autrement dit, $\mathbb{R} = \boldsymbol{\mathcal{C}/\mathcal{R}}$.

II.2 Opérations, structure, ordre

II.2.1 Opérations, structure

Les lois sur $\mathcal{C}$ passent au quotient ; autrement dit, si $a, b \in \mathbb{R}$ avec $a = \overline{u}$ et $b = \overline{v}$, alors on a :

- $a + b = \overline{u+v} = \overline{(u_n + v_n)_{n\in\mathbb{N}}}$;
- $ab = \overline{uv} = \overline{(u_n v_n)_{n\in\mathbb{N}}}$.

On vérifie que ces définitions ne dépendent pas des représentants choisis. De plus, $\mathbb{R}$ est un **anneau commutatif unitaire** : il hérite ces propriétés de la structure d'anneau commutatif unitaire de $\mathcal{C}$.

Proposition 3.8
$\mathbb{R}$ est un corps.

Démonstration : Soit $a \in \mathbb{R}, a \neq 0$; soit $(a_n)_{n\in\mathbb{N}} \in \mathcal{C}$ telle que $a = \overline{(a_n)}$. $a \neq 0$ donc $(a_n)\not\mathcal{R}(0)$ ie $\exists \varepsilon \in \mathbb{Q}_+^*, \forall N \in \mathbb{N}, \exists n \in N, |a_n| > \varepsilon$. Autrement dit, il existe une infinité de valeurs de (a_n) qui sont strictement supérieures à ε en valeur absolue. Comme la suite (a_n) est de Cauchy, $\exists N \in \mathbb{N}, \forall m, n \geqslant N, |a_m - a_n| \leqslant \frac{\varepsilon}{2}$. On fixe un tel N : on sait qu'il existe un $m \geqslant N, |a_m| > \varepsilon$. On a donc :

$$\forall n \in \mathbb{N}, |a_n| = |a_m - (a_m - a_n)| \geqslant |a_m| - |a_m - a_n| \geqslant \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

On pose

$$b_n = \begin{cases} \frac{1}{a_n} \text{ si } n \geqslant N \\ \frac{203}{100} \text{ si } n < N \end{cases}.$$

On[3] a bien $a_n b_n \underset{n\to+\infty}{\longrightarrow} 1$. Reste à démontrer que $(b_n) \in \mathcal{C}$ pour conclure $\overline{(a_n)(b_n)} = \overline{(1)} = 1$ et que donc (a_n) est inversible.

Soit $\eta \in \mathbb{Q}_+^*$. $\forall m, n \geqslant N, |b_m - b_n| = \left|\frac{1}{a_m} - \frac{1}{a_n}\right| = \frac{|a_m - a_n|}{|a_m||a_n|} \leqslant \frac{4}{\varepsilon^2}|a_m - a_n|$. La suite (a_n) est de Cauchy donc $\exists N' \in \mathbb{N}, \forall m, n \geqslant N', |a_m - a_n| \leqslant \frac{\varepsilon^2}{4}\eta$.

3. 2.03 est la constante de Rostam : se référer à l'annexe pour de plus amples informations.

Finalement, en posant $M = \max\{N, N'\}$ on a :

$$\forall m, n \geqslant M, |b_m - b_n| \leqslant \frac{4}{\varepsilon^2}\frac{\varepsilon^2}{4}\eta = \eta.$$

Donc (b_n) est de Cauchy et on conclut.

II.2.2 Ordre

Définition 3.9 (Réels positifs, négatifs) On définit $\mathbb{R}_+$ comme le sous-ensemble des classes contenant au moins une suite à valeurs dans $\mathbb{Q}_+^*$. On définit de même $\mathbb{R}_-$, l'ensemble des classes contenant au moins une suite à valeurs dans $\mathbb{Q}_-^*$.

On définit la relation d'ordre suivante, pour x et y dans $\mathbb{R}$:

$$x < y \overset{\text{def}}{\Longleftrightarrow} y - x \in \mathbb{R}_+.$$

On vérifie la réflexivité et la transitivité ; l'antisymétrie résulte du fait que $\mathbb{R}_+ \cap \mathbb{R}_- = \{\overline{(0)}\}$. $<$ est donc une **relation d'ordre**, et l'ordre est **total** car $\mathbb{R}_+ \cup \mathbb{R}_- = \mathbb{R}$.

Remarque : On peut plonger $\mathbb{Q}$ dans $\mathbb{R}$ par l'injection $\left|\begin{array}{l} \mathbb{Q} \longrightarrow \mathbb{R} \\ q \longmapsto \overline{(q)_{n\in\mathbb{N}}} \end{array}\right.$. La relation d'ordre définie précédemment restreinte à $\mathbb{Q}$ y coïncide avec l'ordre usuel. Dans la suite, on confondra $\mathbb{Q}$ et son image dans $\mathbb{R}$ par l'injection précédente.

II.3 Complétude

Théorème 3.10
$\mathbb{Q}$ est dense dans $\mathbb{R}$.

Démonstration : Soit $A \in \mathbb{R}$: on a donc $A = \lim\limits_{n\to+\infty} a_n$ où $(a_n) \in \mathcal{C}$. En s'inspirant de l'injection précédente, on note, pour $n \in \mathbb{N}, A_n = \overline{(a_n)_{m\in\mathbb{N}}}$ ($\in \mathbb{Q}$). Montrons que $(A_n)_{n\in\mathbb{N}}$ converge vers A.

Soit $\varepsilon \in \mathbb{Q}_+^*$. On cherche $N \in \mathbb{N}$, tel que $\forall n \geqslant N, |A_n - A| \leqslant \varepsilon$ (c'est à dire $\varepsilon - |A_n - A| \in \mathbb{R}_+$). La suite (a_n) est de Cauchy donc $\exists N \in \mathbb{N}, \forall m, n \geqslant N, |a_n - a_m| \leqslant \varepsilon$. En passant à la limite sur m on a bien $\varepsilon - |A_n - A| \geqslant 0$.

Théorème 3.11
$\mathbb{R}$ est complet (pour $|\cdot|$).

Démonstration : Soit (U_n) une suite de Cauchy de $\mathbb{R}^{\mathbb{N}}$. On va utiliser le théorème précédent de densité : pour chaque $n \in \mathbb{N}$, on choisit $a_n \in \mathbb{Q}$ tel que $|U_n - a_n| \leqslant \frac{1}{n}$. Comme $a_n = \underbrace{(a_n - U_n)}_{\substack{\text{CV par} \\ \text{construction}}} + \underbrace{U_n}_{\substack{\text{de} \\ \text{Cauchy}}}$ (on appelle cette égalité (1)), la suite (a_n) est une suite de Cauchy (comme somme de deux suites de Cauchy[4]) de rationnels donc elle converge dans $\mathbb{R}$ (par définition). On note $U = \overline{(a_n)}$. Comme $a_n - U_n \underset{n\to+\infty}{\longrightarrow} 0$, on obtient en passant à la limite dans (1) : $U_n \underset{n\to+\infty}{\longrightarrow} U$.

Passons à présent à une petite définition ainsi qu'à un théorème ; soit K un corps totalement ordonné.

Définition 3.12 (Corps archimédien) K est dit **archimédien** si $\forall a, b \in K_+^*, \exists n \in \mathbb{N}, Na > b$.

Théorème 3.13
Soit K un corps totalement ordonné. Les propriétés suivantes sont équivalentes.

- (i) Toute partie non vide majorée de K possède une borne supérieure.
- (ii) K est complet et archimédien.
- (iii) K est complet et vérifie le théorème des suites adjacentes.
- (iv) K est isomorphe à $\mathbb{R}$.

Démonstration : On a $(ii) \Rightarrow (iii)$ car deux suites adjacentes sont de Cauchy, donc convergentes si l'espace est complet[5].

III Développement décimal

On reprend maintenant notre ensemble $\mathbb{R}$ qu'on connait (on peut oublier la manière dont il a été construit). Soit $x \in \mathbb{R}$. En notant $n = \lfloor x \rfloor$, on a l'encadrement :

$$n + 0.a_1...a_{k-1} \leqslant x \leqslant n + 0.a_1...a_k$$

pour $k \geqslant 2$ et $a_i \in [\![0, 9]\!]$.

Propriété 3.14
Cet encadrement est unique.

Définition 3.15 (Développement décimal) Cet encadrement définit un **développement décimal** de x, donné par $n.a_1...a_k....$

Réciproquement, si $x = n + \displaystyle\sum_{k=1}^{+\infty} \frac{a_k}{10^k}$ avec $a_k \in [\![0, 9]\!]$, on a deux cas :

4. En effet, une suite convergente est toujours de Cauchy.
5. Pour le reste de la démonstration, on peut soit chercher soi-même soit aller jeter un œil sur l'article de Wikipédia !

– soit la suite $(a_k)_{k\in\mathbb{N}}$ converge vers 9 : on peut alors simplifier le développement ;
– soit la suite $(a_k)_{k\in\mathbb{N}}$ ne converge pas vers 9 : on retrouve le développement précédent.

Chapitre 4

Dualité

Introduction

Dans ce chapitre, E désigne un $\mathbb{R}$-espace vectoriel, $f : E \to \mathbb{R}$ est une application linéaire (autrement dit, f est une forme linéaire). Pour $x \in E$ on note $f(x) =< f, x >$. $E^* = \mathcal{L}(E, \mathbb{R})$ est **l'espace dual**. Si E est est normé, on note $E' = \mathcal{L}_c(E, \mathbb{R})$ **l'espace dual topologique**.

I Dimension finie

On note $n = \dim E < +\infty$. Toute application linéaire $f : E \to \mathbb{R}$ est alors continue : on a donc $\boldsymbol{E^* = E'}$.

I.1 Base de E^*

Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de E, et $\varphi_i : x = \sum_{j=1}^n x_j e_j \mapsto x_i$. On va démontrer que $(\varphi_i)_{1 \leqslant i \leqslant n}$ est une base de E^* :

- $\varphi_i \in E^*$ et $\forall j \in [\![1, n]\!], \varphi_i(e_j) = \delta_{i,j}$;
- si $l \in E^*, l(x) = l(\sum_{i=1}^n x_i e_i) = \sum_{i=1}^n x_i l(e_i) = \sum_{i=1}^n x_i < l, e_i >$ donc $l = \sum_{i=1}^n < l, e_i > \varphi_i$ donc $(\varphi_i)_{1 \leqslant i \leqslant n}$ est génératrice ;
- si $\sum_{i=1}^n \lambda_i \varphi_i = 0$ alors $\forall j, \sum_{i=1}^n \lambda_i \varphi_i(e_j) = \lambda_j = 0$ donc $(\varphi_i)_{1 \leqslant i \leqslant n}$ est libre.

Finalement, $(\varphi_i)_{1 \leqslant i \leqslant n}$ est une base de E^*, appelée **base duale** de E^* associée à $(e_i)_{1 \leqslant i \leqslant n}$: E^* est de dimension finie et $\boldsymbol{\dim E^* = \dim E}$.

I.2 Notation matricielle

On fixe une base $(e_i)_{1 \leqslant i \leqslant n}$ de E, on note $(\varphi_i)_{1 \leqslant i \leqslant n}$ la base duale associée. On considère les deux applications :

- $x = \sum_{i=1}^n x_i e_i \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$;
- $l = \sum_{i=1}^n l_i \varphi_i \mapsto \begin{pmatrix} l_1 & \cdots & l_n \end{pmatrix}$.

On a alors $< l, x >= \sum_{i=1}^n l_i x_i = \begin{pmatrix} l_1 & \cdots & l_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Il faut donc faire attention à **ne pas identifier $\boldsymbol{E}$ et $\boldsymbol{E^*}$**, puisque les éléments de E sont représentés par des vecteurs colonne et ceux de E^* par des vecteurs ligne.

II Problème de E^* en dimension infinie

II.1 La "base duale" n'est plus une base

II.1.1 Exemple des polynômes

On considère $E = \mathbb{R}[X]$, et la forme linéaire l définie par

$$\forall P \in E, < l, P = \sum_{n \in \mathbb{N}} a_n X^n >= \sum_{n \in \mathbb{N}} a_n = P(1).$$

l est continue pour $\|P\|_1 = \sum_{n \in \mathbb{N}} |a_n|$ mais pas pour $\|P\|_2 = \sup_{n \in \mathbb{N}} |a_n|$ (il suffit de considérer la suite $(P_n = \sum_{i=1}^n X^i)_{n \in \mathbb{N}}$ qui est constante pour $\|\cdot\|_2$ mais où $(l(P_n))_n$ tend vers $+\infty$).

$(X^n)_{n \in \mathbb{N}}$ est une base de E : pour $i \in \mathbb{N}$ on pose $\varphi_i \in E^*$ définie par $< \varphi_i, P = \sum_{n \in \mathbb{N}} a_n X^n >= a_i$. On a alors $< \varphi_i, X^n >= \delta_{i,n}$. On vérifie encore une fois que $(\varphi_i)_{i \in \mathbb{N}}$ est libre, mais elle n'est cette fois plus génératrice : on a en effet $l \notin \mathrm{vect}(\varphi_i)_{i \in \mathbb{N}}$. On a envie d'écrire $l = \sum_{n \in \mathbb{N}} \varphi_n$ mais c'est une somme infinie et donc on ne peut pas en déduire $l \in \mathrm{vect}(\varphi_i)_{i \in \mathbb{N}}$.

Démonstration : Le fait que $(\varphi_i)_{i \in \mathbb{N}}$ est libre se démontre exactement de la même façon que dans le cas de la dimension finie. Si $l \in \mathrm{vect}(\varphi_i)_{i \in \mathbb{N}}$ alors $l = \sum_{i \in \mathbb{N}} \lambda_i \varphi_i$ avec $(\lambda_i)_{i \in \mathbb{N}} \in \mathbb{R}^{(\mathbb{N})}$! C'est ça qui va tout changer : posons d'abord $N = \max_{i \in \mathbb{N}} \{\lambda_i \neq 0\}$. On a $l(X^{N+1}) = 1 \neq \sum_{i \in \mathbb{N}} \lambda_i \varphi_i (X^{N+1}) = \sum_{i=0}^N \lambda_i \varphi_i (X^{N+1}) = 0$ ce qui est absurde !

II.1.2 Généralisation à un espace de dimension infinie

Soit E un espace vectoriel qui n'est pas de dimension finie [1] ; on fixe $\mathcal{B} = (e_i)_{i \in I}$ une base de E (qui existe, en admettant le théorème de la base (cf chapitre sur l'axiome du choix)). On définit pour $i \in i$ la forme

1. Une définition rigoureuse d'un espace de dimension finie est : E est de dimension finie si $\{\mathrm{card}(e_i)_{i \in I} / (e_i)_{i \in I}$ est une famille libre de $E\}$ est majoré. On dit que E n'est pas de dimension finie si cet ensemble n'est pas majoré (ce qui est ici équivalent à infini) !

linéaire φ_i définie par $< \varphi_i, x = \sum_{j\in I} x_j e_j >= x_i$. On démontre à partir de $< \varphi_i, e_j >= \delta_{i,j}$ le fait que la famille $(\varphi_i)_{i\in I}$ est libre. En revanche, en considérant encore une fois la forme linéaire $< \varphi, x = \sum_{i\in I} >= \sum_{i\in I} x_i$ (on a $(x_i)_{i\in I} \in I^{(\mathbb{N})}$ donc c'est une somme finie) on a $\varphi \in E^*$ mais $\varphi \notin \text{vect}(\varphi_i)_{i\in I}$ donc $\boldsymbol{(\varphi_i)_{i\in I}}$ **n'est pas génératrice** !

II.2 Le bidual

Définition 4.1 (Bidual) Le **bidual** de E est $E^{**} = (E^*)^* = \mathcal{L}(E^*, \mathbb{R})$.

On considère l'application $\Phi : \left| \begin{array}{l} E \longrightarrow E^{**} \\ x \longmapsto \Phi_x \end{array} \right.$ définie par

$$\forall l \in E^*, < \Phi_x, l >_{E^{**},E^*} = < l, x >_{E^*,E}.$$

En remarquant que Φ est linéaire, il est facile de démontrer que Φ est **injective**. On démontrera plus loin que, quand E est de dimension infinie, Φ n'est **jamais surjective**.

II.3 Orthogonalité

II.3.1 Dans E

Définition 4.2 (Orthogonal d'une partie) Si $A \subset E$, on définit **l'orthogonal de $\boldsymbol{A}$** par :

$$A^{\perp} = \{\varphi \in E^* / \forall x \in A, < \varphi, x >= 0\}.$$

Remarques :

– Si $A \subset E$, on a $A^{\perp} = (\text{vect } A)^{\perp}$;

– si $x \in E$, on note $x^{\perp} = \{x\}^{\perp}$;

– si A et B sont des sous-espaces vectoriels de E, on a $A^{\perp} \cap B^{\perp} = (A \cup B)^{\perp} = (\text{vect}(A \cap B))^{\perp} = (A + B)^{\perp} = (A + B)^{\perp}$.

II.3.2 Dans E^*

Définition 4.3 (Orthogonal d'une partie) Si $F \subset E^*$, on définit **l'orthogonal de $\boldsymbol{F}$** par :

$$F^{\circ} = \{x \in E / \forall \varphi \in F, < \varphi, x >= 0\}.$$

Remarques :

– Si $F \subset E^*$, on a $F^{\circ} = \cap_{\varphi\in F} \text{Ker } \varphi = (\text{vect } F)^{\circ}$;

– si $\varphi \in E^*$, on note $\varphi^{\circ} = \{\varphi\}^{\circ}$;

– si C et D sont des sous-espace vectoriels de F, on a $C^{\circ} \cap D^{\circ} = (C + D)^{\circ}$.

II.3.3 Propriétés

Propriété 4.4

Si A est un sous-espace vectoriel de E, alors $\boldsymbol{A = (A^{\perp})^{\circ}}$.

Démonstration : On procède par double inclusion.

$\boldsymbol{A \subset (A^{\perp})^{\circ}}$: Si $x \in A$ alors $\forall \varphi \in A^{\perp}, < \varphi, x >= 0$ donc $x \in (A^{\perp})^{\circ}$.

$\boldsymbol{A \supset (A^{\perp})^{\circ}}$: Soit $B \subset E$ une partie telle que $B \oplus A = E$. (Une telle partie existe : il suffit d'appliquer le théorème de la base (II.2.1, page 15) avec L une base de A (qui existe d'après ce même théorème) et $G = E$.) Démontrons que $E^* = A^{\perp} \oplus B^{\perp}$:

- $A^{\perp} \cap B^{\perp} = (A + B)^{\perp} = E^{\perp} = \{0\}$;
- si $l \in E^*, l = l \circ p_B + l \circ p_A$: comme $A^{\perp} \cap B^{\perp} = \{0\}, l \circ p_B \in A^{\perp}$ et $l \circ p_A \in B^{\perp}$ donc $l \in A^{\perp} + B^{\perp}$ puis $E^* = A^{\perp} + B^{\perp}$.

Soit $x \in (A^{\perp})^{\circ}$. On peut écrire (de façon unique) $x = a + b$ avec $a \in A$ et $b \in B$. Par la première étape de la démonstration on a $a \in (A^{\perp})^{\circ}$. Comme $(A^{\perp})^{\circ}$ est un espace vectoriel, $b = x - a \in (A^{\perp})^{\circ}$; on a également $b \in B \subset (B^{\perp})^{\circ}$. Finalement, $b \in (A^{\perp})^{\circ} \cap (B^{\perp})^{\circ} = (A^{\perp} + B^{\perp})^{\circ} = (E^*)^{\circ} = \{0\}$ donc $b = 0$ et $x = a \in A$.

Propriété 4.5

On suppose que E est de dimension infinie. Si B est un sous-espace vectoriel de E^*, alors $\boldsymbol{B \subsetneq (B^{\circ})^{\perp}}$.

Démonstration : On procède en deux étapes.

$\boldsymbol{B \subset (B^{\circ})^{\perp}}$: Si $l \in B$ alors $\forall x \in B^{\circ}, < l, x >= 0$ donc $l \in (B^{\circ})^{\perp}$.

$\boldsymbol{B \neq (B^{\circ})^{\perp}}$: On va exhiber un contre-exemple. Soit $\mathcal{B} = (e_i)_{i \in I}$ une base de E, et $\varphi_i(x = \sum_{i \in I} x_i e_i) = x_i$. On pose $B = \text{vect}(\varphi_i)_{i \in I}$: on a déjà vu, pour peu que $\dim E = +\infty$, $B \subsetneq E$. On a :

$$\begin{array}{rcl} B^{\circ} & = & \{x \in E / \forall \varphi \in B, < \varphi, x >= 0\} \\ & \subset & \{x \in E / \forall i \in I, < \varphi_i, x >= 0\} \\ & \subset & \{x \in E / \forall i \in I, x_i = 0\} \\ & \subset & \{x \in E / x = 0\} \\ B^{\circ} & \subset & \{0\} \end{array}$$

et donc $B^{\circ} = \{0\}$ puis $(B^{\circ})^{\perp} = E^*$. On a donc bien $B \subsetneq (B^{\circ})^{\perp}$.

Corollaire

On suppose $\dim E = +\infty$. $\Phi : \left| \begin{array}{l} E \longrightarrow E^{**} \\ x \longmapsto \Phi_x \end{array} \right.$ définie par $\forall l \in E^*, < \Phi_x, l >_{E^{**}, E^*} = < l, x >_{E^*, E}$ n'est **jamais surjective**.

Démonstration : On suppose que Φ est surjective : on a déjà vu qu'elle est injective, elle est donc bijective. Intuitivement, on vient de voir que les orthogonaux ne sont pas les mêmes donc il y a a priori une absurdité. Soit $B \subset E^*$. On a $B^{\perp_{E^*,E^{**}}} = \Phi(B^\circ)$:

- Soit $L \in E^{**}$: Φ est bijective donc il existe $x \in E, L = \Phi_x$.

$$\begin{array}{rcl} L \in \varphi(B^\circ) & \Leftrightarrow & \exists y \in B^\circ, L = \Phi_y \\ & \Leftrightarrow & \exists y \in B^\circ, \Phi_x = \Phi_y \\ L \in \varphi(B^\circ) & \Leftrightarrow & x \in B^\circ \text{ car } \Phi \text{ est injective.} \end{array}$$

- On termine :

$$\begin{array}{rcl} L \in B^{\perp_{E^*,E^{**}}} & \Leftrightarrow & \forall l \in B, < L, l >_{E^{**},E^*} = 0 \\ & \Leftrightarrow & \forall l \in B, < \Phi_x, l >_{E^{**},E^*} = 0 \\ & \Leftrightarrow & \forall l \in B, < l, x >_{E^*,E} = 0 \\ & \Leftrightarrow & x \in B^\circ \\ L \in B^{\perp_{E^*,E^{**}}} & \Leftrightarrow & L \in \varphi(B^\circ) \text{ d'après le point précédent.} \end{array}$$

De plus,

$$\begin{array}{rcl} l \in (\varphi(B^\circ))^{\circ_{E^{**},E^*}} & \Leftrightarrow & \forall L \in \varphi(B^\circ), < L, l >_{E^{**},E^*} = 0 \\ & \Leftrightarrow & \forall x \in B^\circ, < \Phi_x, l >_{E^{**},E^*} = 0 \\ & \Leftrightarrow & \forall x \in B^\circ, < l, x >_{E^*,E} = 0 \\ l \in (\varphi(B^\circ))^{\circ_{E^{**},E^*}} & \Leftrightarrow & l \in (B^\circ)^\perp \end{array}$$

donc $(\varphi(B^\circ))^{\circ_{E^{**},E^*}} = (B^\circ)^\perp$ puis $\boldsymbol{(B^{\perp_{E^*,E^{**}}})^{\circ_{E^{**},E^*}} = (B^\circ)^\perp}$. D'après la propriété 4.4 on a $(B^{\perp_{E^*,E^{**}}})^{\circ_{E^{**},E^*}} = B$ donc $B = (B^\circ)^\perp$ ce qui est faux car $\dim E = +\infty$ et donc on a $B \subsetneq (B^\circ)^\perp$. On arrive donc a une absurdité, donc Φ n'est pas surjective.

Propriété 4.6

Soit A un sous-espace vectoriel de E. Alors $\boldsymbol{E^*/A^\perp}$ **est isomorphe à** $\boldsymbol{A^*}$.

Remarque : Si $\varphi, \psi \in E^*, \overline{\varphi} = \overline{\psi}$ <u>ssi</u> $\varphi - \psi \in A^\perp$ <u>ssi</u> $\varphi|_A = \psi|_A$.

Démonstration : On considère l'application $f : \left| \begin{array}{l} E^* \longrightarrow A^* \\ \varphi \longmapsto \varphi|_A \end{array} \right.$. f est linéaire, et $\mathrm{Ker} f = A^\perp$: en effet, $\varphi \in \mathrm{Ker} f \Leftrightarrow \forall x \in A, < \varphi, x >= 0 \Leftrightarrow \varphi \in A^\perp$. Par conséquent, $\hat{f} : \left| \begin{array}{l} E^*/A^* \longrightarrow A^* \\ \overline{\varphi} \longmapsto f(\varphi) \end{array} \right.$ est un isomorphisme [2] !

2. $\hat{f}$ est bien définie : $f(\varphi) = f(\psi) \Leftrightarrow \varphi|_A = \psi_A \Leftrightarrow \overline{\varphi} = \overline{\psi}$.

Propriété 4.7

Soit A un sous-espace vectoriel de E. Alors $\boldsymbol{(E/A)^*}$ **est isomorphe à** $\boldsymbol{A^\perp}$.

Remarque : On peut retenir que E/A c'est "presque comme" un supplémentaire de A dans E.

Démonstration : On pose $f : \left| \begin{array}{l} (E/A)^* \longrightarrow A^\perp \\ \varphi \longmapsto \varphi \circ \pi \end{array} \right.$ avec $\pi : \left| \begin{array}{l} E \longrightarrow E/A \\ x \longmapsto \overline{x} \end{array} \right.$. f est bien définie[3], f est linéaire et $\varphi \in \mathrm{Ker} f \Leftrightarrow \forall x \in E, \varphi(\pi(x)) = 0 \Leftrightarrow \forall x \in E, \varphi(\overline{x}) = 0 \Leftrightarrow \varphi = 0$ donc f est injective. De plus, si $\psi \in A^\perp$ alors $\overline{x} = \overline{y} \Rightarrow x - y \in A \Rightarrow \psi(x-y) = 0 \Rightarrow \psi(x) = \psi(y)$: on peut donc définir $\varphi \in (E/A)^*$ par $\varphi(\overline{x}) = \varphi(\pi(x)) = \psi(x)$. On a donc $\psi = f(\varphi)$, donc f est surjective. Finalement, f est un isomorphisme.

III On met de la topologie : le dual topologique

Rappel : Si E est un $\mathbb{R}$-espace vectoriel normé, $E' = \mathcal{L}_c(E, R)$ est appelé le dual topologique.

On pose $J : \left| \begin{array}{l} E \longrightarrow E'' \\ x \longmapsto Jx \end{array} \right.$ définie par $< Jx, l >=< l, x >$. On a $|< l, x >| = |l(x)| \leqslant \|l\|_{E'} \|x\|_E$ donc $|< Jx, l >| \leqslant \|x\|_E \|l\|_{E'}$. Comme Jx est linéaire, Jx est donc continue et $Jx \in E''$: J est bien définie. On a de plus $\|Jx\|_{E''} \leqslant \|x\|_E$: comme J est linéaire, J est continue et $\|J\|_{\mathcal{L}_c(E,E'')} \leqslant 1$.

On peut se demander si, comme l'application Φ précédente, J est injective. On fixe une base $(e_i)_{i \in I}$ de E. $Jx = 0 \Rightarrow \forall l \in E', < Jx, l >= 0 \Rightarrow \forall l \in E', < l, x >= 0$ mais cela n'implique pas $x = 0$ car on n'est pas sûr que la projection sur la coordonnée numéro i soit continue pour la norme $\|\cdot\|_E$. Cependant, on va prouver que J est injective.

Théorème 4.8 (Théorème de Hahn-Banach)

Soit E un espace vectoriel normé, F un sous-espace vectoriel de E ; soit $f \in F'$. Alors :

$$\exists \boldsymbol{g} \in \boldsymbol{E'}, \boldsymbol{g}|_{\boldsymbol{F}} = \boldsymbol{f} \text{ et } \|\boldsymbol{g}\|_{\boldsymbol{E'}} = \|\boldsymbol{f}\|_{\boldsymbol{F'}}.$$

Remarque : Ceci est en fait un corollaire du théorème original.

Application : On va démontrer que J est une injection, puis que c'est une isométrie !

Soit $x \in E \setminus \{0\}$. On pose $F = \mathrm{vect}(x) = \mathbb{R}x$ puis $f : \left| \begin{array}{l} F \longrightarrow \mathbb{R} \\ \lambda x \longmapsto \lambda \end{array} \right.$. f est linéaire et $\forall y \in F \setminus \{0\}, |f(y)| = |f(\lambda x)| = |\lambda| = \frac{1}{\|x\|_E} \|\lambda x\|_E$ donc f est

3. En effet, $\forall x \in A, (\varphi \circ \pi)(x) = \varphi(\pi(x)) = \varphi(0) = 0$ donc $\varphi \circ \pi \in A^\perp$.

continue : $f \in F'$ et $\|f\|_{F'} = \frac{1}{\|x\|_E}$. D'après le théorème, $\exists g \in E', \|g\|_{E'} = \|f\|_{F'} = \frac{1}{\|x\|_E}$ et $g|_F = f$.

- Si $\forall l \in E', < l, x >= 0$ alors $x = 0$ car si $x \neq 0$ alors on a $1 =< g, x >= 0$ ce qui est absurde ! Donc ***J* est injective**.
- Si $x \neq 0, < l, x > \leqslant \|x\| \, \|l\|$ et on a égalité pour $l = g$ donc :

$$\boldsymbol{\|x\|_E = \max_{l \in E' \setminus \{0\}} \frac{< l, x >}{\|l\|_{E'}}}$$

(marche aussi pour $x = 0$). On a alors :

$$\|J\| = \sup_{x \neq 0} \frac{\|Jx\|}{\|x\|} = \sup_{x \neq 0} \frac{1}{\|x\|} \left(\sup_{l \neq 0} \frac{< Jx, l >}{\|l\|} \right) = \sup_{x \neq 0} \frac{1}{\|x\|} \left(\sup_{l \neq 0} \frac{< l, x >}{\|l\|} \right) = \sup_{x \neq 0} \frac{1}{\|x\|} \|x\| = 1$$

donc ***J* est une isométrie** !

Démonstration : On va maintenant démontrer le théorème d'Hahn-Banach. On pose $X = \{(H, h)/H$ sous-espace vectoriel de $E, H \supset F, h \in H', h|_F = f, |h|_{H'} = |g|_{F'}\}$ (muni de l'ordre qu'on a déjà utilisé plusieurs fois).

- $X \neq \emptyset$ car $(F, f) \in X$.
- X est inductif : soit $(H_i, h_i)_{i \in I}$ est une famille totalement ordonné de X. On pose $H = \cup_{i \in I} H_i$: c'est un sous-espace vectoriel de E car les H_i forment une suite croissante. Reste maintenant à déterminer une fonction h qui convient : soit $x \in H$. Par définition de $H, \exists i \in I, x \in H_i$. Si $x \in H_j$, alors d'après l'ordre qu'on utilise on a soit ($H_i \subset H_j$ et $h_i = h_j|_{H_i}$) soit ($H_i \supset H_j$ et $h_i|_{H_j} = h_j$). Dans tous les cas on a $h_i(x) = h_j(x)$. $h_i(x)$ est donc indépendant du choix de i (pourvu que $x \in H_i$) et on peut donc poser $h(x) = h_i(x)$.
  On a $|h(x)| = |h_i(x)| \leqslant |h_i|_{H_i'} |x|_E = |f|_{F'} |x|_E$ donc

$$|h|_{H'} \leqslant |f|_{F'}.$$

  De plus,

$$|h|_{H'} = \sup_{\substack{x \in H \\ x \neq 0}} \frac{< h, x >}{|x|_E} \geqslant \sup_{\substack{x \in F \\ x \neq 0}} \frac{< h, x >}{|x|_E} = \sup_{\substack{x \in F \\ x \neq 0}} \frac{< f, x >}{|x|_E} = |f|_{F'}$$

  donc finalement $|h|_{H'} = |f|_F$. On en déduit que $(H, h) \in X$ et que (H, h) est un majorant de $(H_i, h_i)_{i \in I}$ dans X.

D'après le lemme de Zorn, X admet un élément maximal (H, h). On va supposer $H \neq E$. On peut donc trouver $x_0 \in E \setminus H$; on pose $\tilde{H} = H \oplus \mathbb{R} x_0$. On pose également $\tilde{h} : \tilde{H} \to \mathbb{R}$ définie par $\tilde{h}(x + \lambda x_0) = h(x) + \lambda \alpha$. Ainsi, $\tilde{h}$ est linéaire, $\tilde{h}(x_0) = \alpha \in \mathbb{R}$ et $\tilde{h}|_H = h$.

On veut avoir

$$\forall x \in H, \forall \lambda \in \mathbb{R}, \left|\tilde{h}(x+\lambda x_0)\right| = |h(x)+\lambda\alpha| \leqslant |g|_{F'}|x+\lambda x_0|_E.$$

On remarque qu'on peut enlever la valeur absolue en considérant $x \leftarrow -x$ et $\lambda \leftarrow -\lambda$.

Cas $\boldsymbol{\lambda = 0}$: On a déjà le résultat car $(H,h) \in X$.

Cas $\boldsymbol{\lambda > 0}$: En divisant par λ on trouve

$$h\left(\frac{x}{\lambda}\right) + \alpha \leqslant |g|_{F'}\left|\frac{x}{\lambda}+x_0\right|_E$$

et donc

$$\forall x \in H, \forall \lambda > 0, \alpha \leqslant |g|_{F'}\left|\frac{x}{\lambda}+x_0\right|_E - h\left(\frac{x}{\lambda}\right).$$

C'est équivalent, avec $y = \frac{x}{\lambda}$, à

$$\forall y \in H, \alpha \leqslant |g|_{F'}|y+x_0|_E - h(y).$$

Cas $\boldsymbol{\lambda < 0}$: En divisant par $-\lambda$ on trouve

$$-h\left(\frac{x}{\lambda}\right) - \alpha \leqslant |g|_{F'}\left|-\frac{x}{\lambda}-x_0\right|_E$$

ce qui équivaut, avec $z = -\frac{x}{\lambda}$, à

$$\forall z \in H, \alpha \geqslant h(z) - |g|_{F'}|z-x_0|_E.$$

On doit donc avoir

$$\forall y,z \in H, |g|_{F'}|y+x_0|_E - h(y) \geqslant h(z) - |g|_{F'}|z-x_0|_E$$

ce qui équivaut à

$$\forall y,z \in H, h(y)+h(z) \leqslant |g|_{F'}(|y+x_0|_E + |z-x_0|_E).$$

Cette dernière inégalité est vérifiée car $h(y)+h(z) = h(y+z) \leqslant |h|_{H'}|y+z|_E$ et comme $|h|_{H'} = |g|_{F'}$, en utilisant l'inégalité triangulaire on a $h(y)+h(z) \leqslant |g|_{F'}(|y+x_0|_E + |z-x_0|_E)$. Finalement,

$$\inf_{y\in H}[|g|_{F'}|y+x_0|_E - h(y)] \geqslant \sup_{z\in H}[h(z) - |g|_{F'}|z-x_0|_E]$$

et un α pris entre les deux membres convient !

On a donc :

$$\forall x \in \tilde{H}, \left|\tilde{h}(x)\right| \leqslant |g|_{F'}|x|_E.$$

Comme $\tilde{h}$ est linéaire, on en déduit que $\tilde{h}$ est continue et que $\left|\tilde{h}\right|_{\tilde{H}'} \leqslant |g|_F$. Par la même démonstration que dans le cas de h on a l'inégalité réciproque : donc $\left|\tilde{h}\right|_{\tilde{H}'} = |g|_F$. Finalement, $\boldsymbol{(\tilde{H},\tilde{h}) \in X}$ et donc $(H,h) < (\tilde{H},\tilde{h})$ ce qui est absurde car (H,h) est maximal ! Donc $\boldsymbol{H = E}$ et le théorème est démontré !

Définition 4.9 (Transposée d'une application) Soit F un $\mathbb{R}$-espace vectoriel normé et f une application de $E \to F$. On note $\boldsymbol{f^T}$ la transposée de f, fonction de $F' \to E'$ définie par :

$$\forall l \in F', \forall x \in E, < f^T(l), x >_{E',E} = < l, f(x) >_{F',F}.$$

Corollaire

Les résultats suivants se déduisent du théorème de Hahn-Banach :

- $\forall x \in E, |x|_E = \sup\limits_{\substack{l \in E' \\ l \neq 0}} \dfrac{< l, x >}{|l|_{E'}}$.
- Soit F un espace vectoriel normé et $f \in \mathcal{L}_c(E, F)$. Alors $f^T \in \mathcal{L}_c(F', E')$ et :

$$\left| f^T \right|_{\mathcal{L}_c(F',E')} = |f|_{\mathcal{L}_c(E,F)}.$$

Avant de démontrer ce corollaire, un petit lemme d'interversion de bornes sup.

Lemme 4.10

Soient I et J deux ensembles, et $(a_{i,j})$ une famille de réels indexée par $I \times J$. Alors :

$$\sup_{i \in I} \sup_{j \in J} a_{i,j} = \sup_{j \in J} \sup_{i \in I} a_{i,j}.$$

Démonstration : Pour $(i_0, j_0) \in I \times J, a_{i_0,j_0} \leqslant \sup_{j \in J} a_{i_0,j}$. En passant au sup sur i on a $\sup_{i \in I} a_{i,j_0} \leqslant \sup_{i \in I} \sup_{j \in J} a_{i,j}$ et en passant au sup sur j on trouve $\sup_{j \in J} \sup_{i \in I} a_{i,j} \leqslant \sup_{i \in I} \sup_{j \in J} a_{i,j}$. Par symétrie on a l'égalité !

Démonstration : On va à présent démontrer le corollaire du théorème d'Hahn-Banach ; on a déjà démontré le premier point, dans l'application qui a suivit le théorème. De plus, $< f^T(l), x >_{E',E} = < l, f(x) >_{F',F} \leqslant |l|_{F'} |f|_{\mathcal{L}_c(E,f)} |x|_E$ donc $f^T(l) \in E'$ et $\left| f^T(l) \right|_{E'} \leqslant |f|_{\mathcal{L}_c(E,F)} |l|_{F'}$. Donc $f^T \in \mathcal{L}_c(F', E')$ et $\left| f^T \right|_{\mathcal{L}_c(E',F')} \leqslant |f|_{\mathcal{L}_c(E,F)}$. Finalement,

$$\begin{aligned}
\left| f^T \right|_{\mathcal{L}_c(F',E')} &= \sup_{l \in F' \setminus \{0\}} \frac{\left| f^T(l) \right|_{E'}}{|l|_{F'}} \\
&= \sup_{l \in F' \setminus \{0\}} \left(\sup_{x \in E \setminus \{0\}} \frac{< f^T(l), x >_{E',E}}{|l|_{E'} |x|_E} \right) \\
&= \sup_{l \in F' \setminus \{0\}} \left(\sup_{x \in E \setminus \{0\}} \frac{< l, f(x) >_{F',F}}{|l|_{E'} |x|_E} \right)
\end{aligned}$$

On permute les sup :

$$\begin{aligned}
\left|f^T\right|_{\mathcal{L}_c(F',E')} &= \sup_{x\in E\setminus\{0\}}\left(\sup_{l\in F'\setminus\{0\}}\frac{< l, f(x) >_{F',F}}{|l|_{E'}|x|_E}\right)\\
&= \sup_{x\in E\setminus\{0\}}\frac{|f(x)|_F}{|x|_E}\\
\left|f^T\right|_{\mathcal{L}_c(F',E')} &= |f|_{\mathcal{L}_c(E,F)}
\end{aligned}$$

Corollaire

L'application $J : \left| \begin{array}{l} E \longrightarrow E'' \\ x \longmapsto Jx \end{array}\right.$ définie par $< Jx, l >_{E'',E'} = < l, x >_{E',E}$ est **une isométrie**.

Démonstration : Il suffit d'appliquer le premier point du corollaire précédent : $|J|_{\mathcal{L}_c(E,E'')} = \sup_{l\neq 0}\frac{<Jx,l>_{E'',E'}}{|l|_{E'}} = \sup_{l\neq 0}\frac{<l,x>_{E',E}}{|l|_{E'}} = |x|_E$.

Définition 4.11 (Espace réflexif) Si J est surjective, on dit que E est **réflexif**.

Remarque : On a vu que J est toujours injective, et que $J : E \to E^{**}$ n'est jamais surjective si $\dim E = +\infty$.

Si $\dim E < +\infty$, alors pour F un sous-espace vectoriel de E on a : si $\{\forall l \in E', ([\forall x \in F, < l, x >= 0] \Rightarrow l = 0)\}$ alors $F = E$. Si $\dim E = +\infty$, cette même condition implique non plus que $F = E$ mais que ***F* est dense**.

Démonstration : On suppose $\overline{F} \neq E$. On peut donc trouver un $x_0 \in E\setminus\overline{F}$; on note $G = \overline{F} \oplus \mathbb{R}x_0$. On définit alors sur G la forme linéaire [4] $f : x + \lambda x_0 \mapsto \lambda$ pour $x \in \overline{F}$ et $\lambda \in \mathbb{R}$; f est continue car $\overline{F}$ est fermé. D'après le théorème de Hahn-Banach, f se prolonge en une forme linéaire $l \in E'$. On a : $\forall x \in \overline{F}, < l, x >=< f, x >= 0$ donc par hypothèse $l = 0$. C'est absurde car $l \neq 0$ car $< l, x_0 >= 1$!

Remarque : Pour $f \in \mathcal{L}_c(E, F)$, peut-on avoir $\mathrm{Im}(f^T) = (\mathrm{Ker} f)^\perp$? Non, car $\mathrm{Im}\,(f^T) \subset (\mathrm{Ker} f)^\perp$ [5].

4. G est bien un sous-espace vectoriel car $\overline{F}$ en est un en tant qu'adhérence d'un sous-espace vectoriel.

5. J'imagine que c'est vrai par ce que $\mathrm{Im}\left(f^T\right)$ n'a a priori pas de raison d'être fermé, mais si quelqu'un sait vraiment qu'il me le dise !

IV Espaces ℓ^p [6]

Pour $p \in [1,+\infty[$ on définit $\ell^p = \left\{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} / \sum_{n=0}^{+\infty} |x_n|^p < +\infty\right\}$ qu'on munit de la norme $|x|_{\ell^p} = \left(\sum_{n=0}^{+\infty} |x_n|^p\right)^{\frac{1}{p}}$. Pour $p = +\infty$ on définit $\ell^\infty = \left\{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} / \exists M \in \mathbb{R}_+, \forall n \in \mathbb{N}, |x_n| \leqslant M\right\}$ qu'on muni de la norme $|x|_{\ell^\infty} = \sup_{n\in\mathbb{N}} |x_n|$.

Théorème 4.12
Les espaces ℓ^p pour $p \in [1,+\infty]$ sont des **espaces de Banach**.

Théorème 4.13
$(\ell^p)'$ est en **isométrie bijective** avec ℓ^q ssi $\frac{1}{p} + \frac{1}{q} = 1$ et $p, q \neq 1$. De plus, si $p \neq 1$ alors ℓ^p est **réflexif**.

Théorème 4.14
On a $\ell^\infty = (\ell^1)'$ et $(\ell^\infty)' \supsetneq \ell^1$. On a en fait $\ell^1 = (c_0)'$ où c_0 est l'ensemble des suites qui tendent vers 0.

6. Cette partie n'est plus utile puisqu'on s'y intéresse ensuite dans les cours de LEFO et d'EVNO ; c'est pourquoi je vais juste énoncer les résultats, sans les démonstrations (quand elles y étaient).

Chapitre 5

Hyperplans

Dans tout ce chapitre, E désigne un $\mathbb{R}$-espace vectoriel, et H est un sous-espace vectoriel de E.

Définition 5.1 (Hyperplan) H est un **hyperplan** ssi il est **maximal pour l'inclusion** parmi les sous-espaces vectoriels propres ssi $H \neq E$ et $\forall F$ sous-espace vectoriel de $E, [H \subset F \Rightarrow (H = F \text{ ou } F = E)]$.

Proposition 5.2
H est un hyperplan ssi $\dim E/H = 1$.

Démonstration : On note Π la projection 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 un 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 l'application $\widetilde{\Pi} : \left| \begin{array}{l} A \longrightarrow \widetilde{A} := \{\text{sev de } E/H\} \\ F \longmapsto \widetilde{\Pi}(F) = \Pi(F) \end{array} \right.$. Il est bon ici de faire attention : $\widetilde{\Pi}$ prend en argument des éléments de A (qui sont des sous-espaces vectoriels de E), contrairement à Π qui prend en argument des vecteurs de E. Ainsi, $\widetilde{\Pi}$ n'est plus linéaire, mais $\widetilde{\Pi}$ hérite quand même de la surjectivité de Π. On va maintenant montrer que $\widetilde{\Pi}$ est injective.

Soient $F_1, F_2 \in A$, tels que $\widetilde{\Pi}(F_1) = \widetilde{\Pi}(F_2)$. On a donc :

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

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 dans F_2 ! Donc $F_1 \subset F_2$; par symétrie, on a l'inclusion réciproque et par conséquent l'égalité. Finalement, $\widetilde{\Pi}$ est injective ; on sait qu'elle est surjective, c'est donc une bijection. On a donc une bijection entre $A = \{$sous-espaces vectoriels de E contenant $H\}$ et $\widetilde{A} = \{$sous-espaces vectoriels de $E/H\}$.

H est donc un hyperplan ssi $\operatorname{card} A = 2$ ssi $\operatorname{card} \widetilde{A} = 2$ ssi $\dim E/H = 1$.

Proposition 5.3

On a les équivalences suivantes :

(i) H est un hyperplan

(ii) $\exists e \notin H, E = H \oplus \mathbb{R}e$

(iii) $\forall e \notin H, E = H \oplus \mathbb{R}e$

(iv) $\exists \varphi \in E^* \setminus \{0\}, H = \operatorname{Ker} \varphi$

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

Démonstration :

$(i) \Rightarrow (iv)$ On suppose que H est un hyperplan. D'après la proposition précédente, $\dim E/H = 1$ donc on a un isomorphisme $\theta : E/H \to \mathbb{R}$. En notant Π la projection canonique $E \to E/H$ et en posant $\varphi = \theta \circ \Pi$, on a bien $\varphi \in E^* \setminus \{0\}$; de plus, $x \in \operatorname{Ker} \varphi \Leftrightarrow \theta(\Pi(x)) = 0 \Leftrightarrow \Pi(x) = 0 \Leftrightarrow x \in H$ donc $\operatorname{Ker} \varphi = H$.

$(iv) \Rightarrow (iii)$ On suppose que $H = \operatorname{Ker} \varphi$ pour $\varphi \in E^* \setminus \{0\}$. Soit $e \notin H$: on a donc $\varphi(e) \neq 0$. Soit $x \in E$; on pose $\lambda = \frac{\varphi(x)}{\varphi(e)}$. On a $x = \underbrace{x - \lambda e}_{\in H} + \underbrace{\lambda e}_{\in \mathbb{R}e}$; donc $E = H + \mathbb{R}e$.

Si $x \in H \cap \mathbb{R}e$, 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, $E = H \oplus \mathbb{R}e$.

$(iii) \Rightarrow (ii)$ La démonstration est laissée au lecteur !

$(ii) \Rightarrow (i)$ Cela résulte de la définition d'un hyperplan (cf maximalité).

Théorème 5.4

On suppose maintenant que E est un espace vectoriel normé. On a alors l'équivalence suivante :

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

Démonstration : Le sens de droite à gauche de l'implication découle directement de la continuité de φ. Pour l'autre sens, on peut faire le raisonnement élémentaire qui suit [1].

On suppose que H est fermé : $E \setminus H$ est donc un ouvert. Soit $e \in E \setminus H$; $E \setminus H$ est ouvert donc $\exists r > 0, B(e, r) \subset E \setminus H$. D'après la proposition précédente, on sait qu'il existe $\varphi \in E^*$ telle que $H = \operatorname{Ker} \varphi$. On a donc :

$$(*) \, \forall x \in B(e, r), \varphi(x) \neq 0.$$

On peut supposer $\varphi(e) > 0$ (quitte à considérer $-e$) : on a donc $\varphi|_{B(e,r)} > 0$ car $B(e, r)$ est convexe et φ est linéaire. En effet, s'il existe $f \in B(e, r)$ tel

1. On peut aussi considérer l'application $\widehat{\varphi} : E/H \to \mathbb{R}$ induite par φ qui reste est linéaire, et comme l'espace de départ est de dimension finie (de dimension 1 par une proposition précédente), cette application est continue et donc φ également (par composition).

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 :

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

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

Chapitre 6

Filtres

Préambule

Les filtres généralisent en quelque sorte la notion de suite ; on retrouvera ainsi des résultats qu'on connait pour les suites (compacité, continuité,...), qui seront vrais ici pour les filtres. Tout au long de ce chapitre, on fera le parallèle avec les suites.

I Définition, premières propriétés

I.1 Définitions

Dans toute cette partie, E désigne un ensemble.

Définition 6.1 (Filtre) Soit $\mathcal{F} \in \mathcal{P}(E)$. $\mathcal{F}$ est un **filtre** si :

i) $\emptyset \notin \mathcal{F}$;

ii) $\forall B \in \mathcal{F}, \forall A \in E, [B \subset A \Rightarrow A \in \mathcal{F}]$;

iii) $\forall A, B \in \mathcal{F}, A \cap B \in \mathcal{F}$.

Remarques :

– Si $\mathcal{F}$ est un filtre et si $A \in \mathcal{F}$ alors $A^c \notin \mathcal{F}$; ainsi, $\mathcal{P}(E) \setminus \{\emptyset\}$ 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 $\emptyset \subsetneq A \subset E$ 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 Fréchet**.

Définition 6.2 (Base de filtre) Une partie $\mathcal{B}$ de E est une **base de filtre** si :

i) $\emptyset \notin \mathcal{B}$;

ii) $\forall A, B \in \mathcal{B}, \exists C \in \mathcal{B}, C \subset A \cap B$.

Exemple : Les bases de voisinages sont des bases de filtres.

Définition 6.3 (Filtre engendré) Soit $\mathcal{B}$ 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 :

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

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

Définition 6.4 Soient $\mathcal{F}_1, \mathcal{F}_2$ deux filtres sur E. On dit que $\mathcal{F}_1$ est **plus fin** que $\mathcal{F}_2$ si $\mathcal{F}_1 \supset \mathcal{F}_2$.

Remarques :

– En faisant le parallèle avec les suites, on dirait que u est plus fine que v si u est une suite extraite de v.

– On peut faire l'analogie avec les topologies : τ_1 est plus fine que τ_2 si $\tau_1 \supset \tau_2$; il y a des suites qui convergent pour τ_2 alors qu'elles ne convergent pas pour τ_1.

Définition 6.5 (Ultrafiltre) Soit $\mathcal{F}$ un filtre. On dit que $\mathcal{F}$ est un **ultrafiltre** s'il n'y a pas de filtre plus fin, ie pour tout filtre $\mathcal{G}$, $[\mathcal{G} \supset \mathcal{F} \Rightarrow \mathcal{G} = \mathcal{F}]$.

Remarques :

– Une suite u serait une "ultrasuite" si on ne pouvait plus extraire de suite de u : cette notion est propre aux filtres.

– A priori, il n'est pas évident que les ultrafiltres existent !

I.2 Image d'un filtre

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

Propriété 6.6

On a les propriétés suivantes.

- Soit $\mathcal{B}$ une base de filtre sur E. Alors $f(\mathcal{B})$ est une base de filtre sur F.
- Soit $\mathcal{C}$ une base de filtre sur F. Alors $f^{-1}(\mathcal{C})$ est une base de filtre sur E.

Définition 6.7 (Filtre image) $f(\mathcal{B})$ s'appelle l'**image** de $\mathcal{B}$ par f.

Propriété 6.8

Soit $\mathcal{F}$ un filtre sur E. Alors $f(\mathcal{F})$ est une base de filtre qui **engendre** le filtre image de $\mathcal{F}$ par f. Par abus de notation, on notera parfois $f(\mathcal{F})$ le filtre image.

Exemple : Soit $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. On note $\mathcal{F}$ le filtre de Fréchet, de base $\mathcal{B} = \{\{p/p \geqslant n\}/n \in \mathbb{N}\}$. $x(\mathcal{B}) = \{\{x_p/p \geqslant n\}/n \in \mathbb{N}\}$ est une base de filtre sur E. Le filtre engendré par $x(\mathcal{B})$ est $\{A \subset E/\exists n \in \mathbb{N}, \{x_p/p \geqslant n\} \subset A\} = \{A \subset E/\exists J_{\text{fini}} \subset \mathbb{N}, \forall p \in \mathbb{N} \setminus J, x_p \in A\}$. C'est le **filtre associé** à (x_n).

I.3 Convergence d'un filtre

Dans ce paragraphe, on considère un espace topologique (E, τ).

Définition 6.9 (Convergence d'un filtre) 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 $\mathcal{F} \to a$, et on dit alors que $\mathcal{F}$ est **convergent** et a est sa **limite**.

Exemple : Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de E ; soit $a \in E$. On note $\mathcal{F}$ le filtre associé à $(x_n)_n$, ie $\mathcal{F} = \{A \subset E/\exists n \in \mathbb{N}, \{x_p/p \geqslant n\} \subset A\}$. On a alors :

$$\begin{aligned}
(x_n) \underset{n\to+\infty}{\longrightarrow} a \quad &\Leftrightarrow \quad \forall V \in \mathcal{V}_a, \exists n \in \mathbb{N}, \forall p \geqslant n, x_p \in V \\
&\Leftrightarrow \quad \forall V \in \mathcal{V}_a, \exists n \in \mathbb{N}, \{x_p/p \geqslant n\} \subset V \\
&\Leftrightarrow \quad \forall V \in \mathcal{V}_a, V \subset \mathcal{F} \\
&\Leftrightarrow \quad \mathcal{V}_a \subset \mathcal{F} \\
(x_n) \underset{n\to+\infty}{\longrightarrow} a \quad &\Leftrightarrow \quad \mathcal{F} \to a
\end{aligned}$$

Propriété 6.10

Si $\mathcal{F}$ est un filtre convergent vers a et si $\mathcal{G}$ est un filtre plus fin que $\mathcal{F}$, alors $\mathcal{G}$ converge également vers a.

Remarque : Cette propriété est à mettre en parallèle avec la propriété suivante sur les suites : si u est une suite convergente et si v est une suite extraite de u, alors v converge également, vers la même limite que u.

Exemple : Soit $(x_n)_{n\in\mathbb{N}}$ une suite, soit $\mathcal{F}$ son filtre associé ; soit $(x_{n_k})_{k\in\mathbb{N}}$ une suite extraite, et soit $\mathcal{G}$ sont filtre associé. Alors $\mathcal{G} \subset \mathcal{F}$; c'est bien cohérent avec la propriété précédente.

Proposition 6.11

Si l'espace topologique (E, τ) est séparé, alors un filtre a au plus une limite.

Démonstration : On suppose (E, τ) séparé ; on suppose qu'il existe un filtre $\mathcal{F}$ possédant deux limites $a \neq \widehat{a}$. Par définition, on a donc $\mathcal{V}_a \subset \mathcal{F}$ et $\mathcal{V}_{\widehat{a}} \subset \mathcal{F}$. (E, τ) est séparé, donc comme $a \neq b$ on peut trouver deux ouverts disjoints $\omega, \widehat{\omega} \subset \tau$ tels que $a \in \omega$ et $\widehat{a} \in \widehat{\omega}$. ω (respectivement $\widehat{\omega}$) est un voisinage de a (respectivement de $\widehat{a}$) donc $\omega \in \mathcal{F}$ (respectivement $\widehat{\omega} \in \mathcal{F}$). Le filtre $\mathcal{F}$ contient deux ensembles disjoints : c'est absurde ! Donc $a = \widehat{a}$.

Remarque : On retrouve encore une fois l'analogie avec les suites.

Théorème 6.12

Soit $A \subset E$; on note $\overline{A}$ son adhérence. On a l'équivalence suivante :

$$a \in \overline{A} \ \underline{\text{ssi}} \ \exists \mathcal{F} \text{ filtre} \to a, \exists B \text{ base de } \mathcal{F}, \text{tels que } \forall B \in \mathcal{B}, B \subset A.$$

Démonstration :

$\Longrightarrow$ On suppose $a \in \overline{A}$: on a donc $\forall V \in \mathcal{V}_a, V \cap A \neq \emptyset$. L'ensemble $\mathcal{B} = \{A \cap V / V \in \mathcal{V}_a\}$ est une base de filtre.

- On a bien $\forall B \in \mathcal{B}, B \subset A$.
- Soit $\mathcal{F}$ le filtre engendré par $\mathcal{B}$ (on rappelle que $\mathcal{F} = \{A \subset E / \exists B \in \mathcal{B}, B \subset A\}$). Si $V \in \mathcal{V}_a$, alors $V \supset A \cap V \in \mathcal{B}$ donc $V \in \mathcal{F}$. Finalement, $\mathcal{V}_a \subset \mathcal{F}$ donc $\mathcal{F} \to a$.

$\Longleftarrow$ Soit $V \in \mathcal{V}_a$. Comme $\mathcal{F} \to a, \mathcal{F}$ contient tous les voisinages de a et donc en particulier $V \in \mathcal{F}$. Par définition de $\mathcal{F}$, on a $A \in \mathcal{F}$: en effet, $\forall B \in \mathcal{B}, B \subset A$ donc en particulier $\exists B \in \mathcal{B}, B \subset A$. On a donc $A \cap V \in \mathcal{F}$ et par conséquent $A \cap V \neq \emptyset$. Finalement, $a \in \overline{A}$.

Remarque : Ceci généralise la caractérisation séquentielle de l'adhérence dans les espaces métriques.

Proposition 6.13

Soient (E, τ) et $(\widehat{E}, \widehat{\tau})$ deux espaces topologiques, soient $u : (E, \tau) \to (\widehat{E}, \widehat{\tau})$ une application et $a \in E$. Alors :

$$u \text{ est continue en } a \ \underline{\text{ssi}} \ \forall \mathcal{F} \text{ filtre sur } E, \mathcal{F} \to a, \text{on a } u(\mathcal{F}) \to u(a).$$

Remarques :

– On désigne par $u(\mathcal{F})$ le filtre image de $\mathcal{F}$ par u.

– Ceci généralise la caractérisation séquentielle de la continuité dans les espaces métriques.

Démonstration :

$\underline{\Longrightarrow}$: On suppose u continue en a ; soit $\mathcal{F} \to a$. Soit $V \in \mathcal{V}_{u(a)}$. Comme u est continue en a, il existe $U \in \mathcal{V}_a$ tel que $u(U) \subset V$. Le filtre $\mathcal{F}$ converge vers a donc $\mathcal{F}$ contient tous les voisinages de a donc en particulier $U \in \mathcal{F}$: donc $\exists B \in \mathcal{F}, B \subset B$. Finalement, V est dans le filtre image de $\mathcal{F}$ par u, donc $\mathcal{V}_{u(a)} \in u(\mathcal{F})$ ie $u(\mathcal{F}) \to u(a)$.

$\Longleftarrow$: On pose $\mathcal{F} = \mathcal{V}_a$; on a bien $\mathcal{F} \to a$. On a $u(\mathcal{F}) = \{B \subset F / \exists U \in \mathcal{F}, u(U) \subset B\}$: par hypothèse $u(\mathcal{F}) \to u(a)$ donc $\forall V \in \mathcal{V}_a, V \in u(\mathcal{F})$. Par conséquent, pour chaque $V \in \mathcal{V}_a$ il existe $U \in \mathcal{F} = \mathcal{V}_a$ tel que $u(U) \subset V$: on a démontré que u est continue !

Théorème 6.14
(E, τ) espace topologique séparé. Alors E est **compact** ssi $\forall \mathcal{F}$ filtre sur E, il existe un filtre $\mathcal{G} \supset \mathcal{F}$ qui converge.

Remarque : On retrouve « de toute suite on peut extraire une sous-suite qui converge ».

Démonstration :

$\Longrightarrow$: Soit $\mathcal{F} = (A_i)_{i \in I}$ un filtre sur E. On a $\forall J_{\text{fini}} \subset I, \cap_{j \in J} \overline{A_j} \neq \emptyset$ donc $\cap_{i \in I} \overline{A_i} \neq \emptyset$ (c.f. compacité de E). Soit $a \in \cap_{i \in I} \overline{A_i}$. Si $B = A_i$ et $V \in \mathcal{V}_a, a \in \overline{A_i}$ donc $B \cap V \neq \emptyset$. Soit G le filtre engendré par $\{B \cap V / B \in \mathcal{F}, V \in \mathcal{V}_a\}$. Avec $B \leftarrow E$ on obtient $\mathcal{V}_a \subset \mathcal{G}$ donc $\mathcal{G}$ converge et avec $V \leftarrow E$ on obtient $\mathcal{F} \subset \mathcal{G}$ CQFD.

$\Longleftarrow$: Soit $(F_i)_{i \in I}$ une famille de fermés tels que $\cap_{i \in I} F_i = \emptyset$. Par l'absurde, supposons que pour toute partie $J \subset I$ finie on ait $\cap_{j \in J} F_j \neq \emptyset$. Ainsi, $\mathcal{B} := \{\cap_{j \in J} F_j / J_{\text{fini}} \subset I\}$ est une base de filtre. Soit $\mathcal{F}$ le filtre engendré par $\mathcal{B}$. Par hypothèse, il existe un filtre $\mathcal{G} \supset \mathcal{F}$ qui converge vers a. Comme $\mathcal{V}_a$, on a $\forall V \in \mathcal{V}_a, V \in \mathcal{G}$; de plus, $\forall i \in I, F_i \in \mathcal{G}$. Ainsi, $\forall i \in I, \forall V \in \mathcal{V}_a, F_i \cap V \neq \emptyset$ donc $\forall i \in I, F_i \in \mathcal{V}_a$ et donc $a \in \overline{F_i} = F_i$. On a alors $\{a\} \subset \cap_{i \in I} F_i = \emptyset$ ce qui est absurde ! Donc il existe $J \subset I$ fini tel que $\cap_{j \in J} F_j = \emptyset$ donc finalement E est compact.

II Théorème de Tychonov

II.1 Ultrafiltres

Théorème 6.15
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 inductif.

– X est non vide car $\mathcal{F} \in X$.
– Soit $(\mathcal{G}_i)_{i \in I}$ une famille totalement ordonnée de X. On pose $\mathcal{G} := \cup_{i \in I} \mathcal{G}_\rangle$: $\mathcal{G}$ contient tous les $\mathcal{G}_\rangle$ et donc aussi $\mathcal{F}$. Il reste à montrer que c'est un filtre !
  • Soit $A \in \mathcal{G}$ et soit $B \subset E$ contenant A. Il existe $i \in I$ tel que $A \in \mathcal{G}_\rangle$ qui est un filtre donc $B \in \mathcal{G}_i$ donc $B \in \mathcal{G}$.
  • $\forall i \in I, \emptyset \notin \mathcal{G}_i$ donc $\emptyset \notin \mathcal{G}$.

- Soient A et B dans $\mathcal{G}$. Il existe $i, j \in I$ tels que $A \in \mathcal{G}_i, B \in \mathcal{G}_j$; la famille $(\mathcal{G}_k)_{k\in I}$ étant totalement ordonnée, on peut supposer 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 inductif : il contient donc un élément maximal, qui est donc un ultrafiltre.

Corollaire

(E, τ) séparé. Alors (E, τ) est **compact** ssi tout **ultrafiltre** est **convergeant**.

Démonstration : On utilise le théorème précédent ainsi que le théorème 6.14.

Proposition 6.16

Soit $\mathcal{U}$ un filtre. Alors $\mathcal{U}$ est un **ultrafiltre** ssi $\forall A \subset E, (A \in \mathcal{U}$ ou $A^c \in \mathcal{U})$.

Démonstration : Soit $\mathcal{U}$ un filtre.

$\Longrightarrow$: Soit $A \notin \mathcal{U}$. Comme $\forall B \in \mathcal{U}, B \not\subset A$ on a $B \cap A^c \neq \emptyset$. Ainsi, $\mathcal{B} := \{B \cap A^c / B \in \mathcal{U}\}$ est une base de filtre ; on note $\mathcal{F}$ le filtre qu'elle engendre. On a $\forall B \in \mathcal{U}, B \cap A^c \subset B$ donc $B \in \mathcal{F}$. Ainsi, $\mathcal{U} \subset \mathcal{F}$ donc $\mathcal{U} = \mathcal{F}$ car $\mathcal{U}$ est un ultrafiltre. En prenant $B = E$ on récupère donc $A^c \in \mathcal{U}$.

$\Longleftarrow$: Supposons qu'il existe un ultrafiltre $\mathcal{F}$ tel que $\mathcal{U} \subsetneq \mathcal{F}$; il existe donc un A dans $\mathcal{F} \setminus \mathcal{U}$. Par hypothèse, $A^c \in \mathcal{U} \subset \mathcal{F}$ ce qui est impossible car $A \in \mathcal{F}$.

Proposition 6.17

Soit $\mathcal{U}$ un filtre. Alors $\mathcal{U}$ est un **ultrafiltre** ssi $\forall A, B \subset E, [A \cup B \in \mathcal{U} \implies (A \in \mathcal{U}$ ou $B \in \mathcal{U})]$.

Démonstration : Soit $\mathcal{U}$ un filtre.

$\Longrightarrow$: Soit $A \cup B \in \mathcal{U}$; on suppose que $A \notin \mathcal{U}$. Par la proposition précédente, on a donc $A^c \in \mathcal{U}$. Or, $(A \cup B)^c = A^c \cap B^c \notin \mathcal{U}$ donc $B^c \notin \mathcal{U}$. En réutilisant la proposition précédente on conclut donc que $B \in \mathcal{U}$.

$\Longleftarrow$: $\forall A \subset E, A \cup A^c = E \in \mathcal{U}$ donc $\forall A, (A \in \mathcal{U}$ ou $A^c \in \mathcal{U})$ donc $\mathcal{U}$ est un ultrafiltre par la proposition qui précède.

Proposition 6.18

Soit $\mathcal{B}$ une base de filtre. Alors $\mathcal{B}$ engendre un **ultrafiltre** ssi $\forall A \subset E, \exists B \in \mathcal{B}, (B \subset A$ ou $B \subset A^c)$.

Théorème 6.19
Soit $\mathcal{B}$ une base de filtre sur E telle que le filtre engendré soit un ultrafiltre, soit $f : E \to F$. Alors $f(\mathcal{B})$ **engendre un ultrafiltre**.

Démonstration : Soit $A \subset F$. $\mathcal{B}$ est une base d'ultrafiltre donc comme $f^{-1}(A) \subset E$ par la proposition précédente on a $\exists B \in \mathcal{B}, (B \subset f^{-1}(A)$ ou $B \subset f^{-1}(A)^c)$. Ainsi, $f(B) \subset A$ ou $f(B) \subset A^c$. D'après la proposition précédente, $f(\mathcal{B})$ est donc une base d'ultrafiltre.

II.2 Théorème de Tychonov

Soit $(E_i, \tau_i)_{i\in I}$ une famille d'espaces topologiques (non vides) ; on considère $E := \prod_{i\in I} E_i$ (qui est non vide, c.f. le chapitre sur l'axiome du choix) muni de la topologie produit τ. On rappelle que les éléments de τ (les ouverts de E) sont engendrés par les $\prod_{j\in J} A_i \times \prod_{i\in I\setminus J} E_i$ où $J_{\text{fini}} \subset I$ et $A_j \in \tau_j$. De plus, $A \subset E$ est un voisinage de $a \in E$ <u>ssi</u> $\exists J_{\text{fini}} \subset I, \forall j \in J, \exists A_j \in \mathcal{V}_{E_j}(a_j), \prod_{j\in J} A_i \times \prod_{i\in I\setminus J} E_i \subset A$.

Théorème 6.20
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 6.21
Soit $\mathcal{F}$ un filtre sur E, $p_i : E \to E_i$ 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.

Démonstration :
- $\underline{\Longrightarrow}$: C'est vrai car les p_i sont continues et que $a_i = p_i(a)$ (c'est la proposition 6.13).
- $\underline{\Longleftarrow}$: Soit $U \in \mathcal{V}_E(a)$; par définition, U contient un ouvert qui contient a, i.e. $\exists J_{\text{fini}} \subset I, \forall j \in J, \exists A_j \in \tau_j, a_j \in A_j, \prod_{j\in J} A_j \times \prod_{i\in I\setminus J} E_i \subset U$. Chaque A_j est un voisinage de a_j dans E_j donc A_j est dans le filtre engendré par $p_j(\mathcal{F})$: en effet, ce filtre converge vers a_j donc contient tous les voisinages de a_j. Il existe donc $B \in \mathcal{F}$ tel que $p_j(B) \subset A$ donc $B \subset p_j^{-1}(A)$: on a donc $p_j^{-1}(A) \in \mathcal{F}$. Ainsi, $\cap_{j\in J} p_j^{-1}(A_j) \in \mathcal{F}$. Or, $p_j^{-1}(A_j) = A_j \times \prod_{i\in I\setminus\{j\}} E_i$ donc $\cap_{j\in J} p_j^{-1}(A_j) = \prod_{j\in J} A_j \times \prod_{i\in I\setminus J} E_i \subset U$. On a donc $U \in \mathcal{F}$.

Place maintenant à la démonstration du théorème de Tychonov.

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 $\exists U_{i_0} \in \mathcal{V}_{E_{i_0}}(a_{i_0}), \exists V_{i_0} \in \mathcal{V}_{E_{i_0}}(b_{i_0}), U_{i_0} \cap V_{i_0} = \emptyset$. On pose maintenant $U := U_{i_0} \times \prod_{i \neq i_0} E_i$ et $V := V_{i_0} \prod_{i \neq i_0} E_i$: on a $U \in \mathcal{V}_E(a)$, $V \in \mathcal{V}_E(b)$ et $U \cap V = \emptyset$. Et voilà !

On reprend nos applications coordonnées $p_i : \left| \begin{array}{l} E \longrightarrow E_i \\ a = (a_i)_{i \in I} \longmapsto a_i \end{array} \right.$. Soit $\mathcal{U}$ 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 II.1). $\mathcal{U}$ est un ultrafiltre donc d'après le théorème 6.19, $p_i(\mathcal{U})$ engendre un ultrafiltre. L'espace (E_i, τ_i) est compact donc cet ultrafiltre est convergeant (encore une fois le corollaire II.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, 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 $K := \{f : E \to \mathbb{R} / \forall x \in E, |f(x)| \leqslant M_x\} \subset \mathbb{R}^E = \prod_{x \in E} \mathbb{R}$. Alors K est compact pour la topologie produit (i.e. la topologie de la convergence simple) ! En effet, on a en fait $K = \prod_{x \in E} [-M_x, M_x]$ qui est compact car chaque $[-M_x, M_x]$ est compact (pour la topologie usuelle de $\mathbb{R}$).

On a en particulier le résultat suivant.

Théorème 6.22 (Banach-Alaoglu)

Soit E un espace vectoriel normé, E' son dual topologique ; on note $\sigma(E', E)$ la topologie de E' induite par la topologie produit de $\mathbb{R}^E$ ($\sigma(E', E)$ est appelée topologie faible-*). Alors $B_E := \{l \in E' / |l|_{E'} \leqslant 1\}$ est **compacte** pour la topologie $\sigma(E', E)$.

Remarque : C'est un résultat digne d'intérêt car si E est de dimension infinie alors B_E n'est pas compacte pour la topologie de la norme.

Annexe : Constante de Rostam

La constante de Rostam vaut 2.03. Elle se distingue des autres nombres par plusieurs propriétés :

– quand on l'ajoute à elle-même, on obtient pile son double ;
– quand on la multiplie par elle même, on obtient pile son carré ;
– quand on la retranche de 100 on obtient 97.97 ;
– c'est la hauteur des portes de l'internat du lycée Chateaubriand à Rennes ;
– c'est la taille en mètres de la diagonale verticale des lits de l'internat de ce même lycée (du haut de la tête de lit au bas du matelas) ;
– $2.03 = \lfloor \log(\sqrt{11500}) * 100 \rfloor / 100$ (formule de Riffaut-Rostam) ;
– c'est la taille en mètres du judoka Teddy Riner (d'après France 2 lors des championnat du monde de judo 2011) ;
– quand on calcule 203^3, on obtient exactement chaque chiffre de 2 à 8 ;
– dans les restaurants universitaires de Rennes pour l'année scolaire 2012/2013, dix tickets repas coûtent 31 euros, ce qui ramené en francs s'élève à 203F (et une trentaine de centimes) ;
– et encore d'autres certainement encore plus intéressantes à découvrir...

De plus :

– le 2 mars est le Rostam-day ;
– l'agent secret 003 est décidément le meilleur (merci Antonin !) ;
– et encore d'autres anecdotes encore plus intéressantes à découvrir !