Introduction à la Logique Mathématique

Première partie: Théorie des ensembles

Itaï Ben Yaacov

Thomas Blossier

Julien Melleray

Avant-Propos.

Ce document sert de support à la première partie du cours de Logique Mathématique donné en M1 à l'Université Lyon 1. Cette version est celle du cours de printemps 2011, comportant des modifications assez importantes par rapport à la version de 2010. Ces notes contiennent sans aucun doute des erreurs, coquilles, approximations, contradictions, assertions non justifiées, etc. Nous encourageons donc nos lecteurs à exercer leur sens critique durant leur lecture, et leur serions reconnaissants de bien vouloir nous signaler tout problème de cette nature qu'ils remarqueraient.

Table des matières

CHAPITRE 1

Les axiomes de Zermelo-Fraenkel

On va commencer par essayer de décrire brièvement le cadre de la théorie axiomatique de Zermelo-Fraenkel, (ZF).
Nous avons tous une notion intuitive d'ensemble, comme « collection d'objets». De même, nous avons une notion intuitive de ce que signifie appartenir à un tel ensemble. Mais pour faire des mathématiques, on a besoin que ces ensembles aient des propriétés qui correspondent à notre intuition ; par exemple on voudrait pouvoir former l'union d'un ensemble d'ensembles. On pourrait se contenter de s'autoriser ces manipulations intuitives (après tout on « voit» bien ce que cela signifie que d'appartenir à la réunion d'un ensemble d'ensembles). Mais le problème est alors que ce qui semble « intuitif» ne l'est pas forcément. Par exemple, pourquoi ne pourrait-on-pas considérer l'ensemble A formé par les ensembles x tels que $x \notin x$? On arrive alors à une situation désagréable: si $A \in A$ alors, par définition de A, on doit avoir $A \notin A$; on se dit que ce n'est pas grave et que cela signifie simplement que $A \notin A$. Hélas, dans ce cas la définition de A entraîne que $A \in A$... Autrement dit, si l'on veut que le principe du tiers exclu soit vrai, il est nécessaire que la « définition » de A ci-dessus ne soit en fait pas une définition mathématiquement acceptable ; dit autrement, il faut que A ne soit pas un ensemble [i]. Mais alors, qu'est-ce qu'un ensemble ?

Ce type de considérations a provoqué la naissance de l'approche *axiomatique* de la théorie des ensembles: plutôt que de répondre à la question « qu'est-ce qu'un ensemble ? », on voudrait spécifier les axiomes que doivent vérifier les ensembles « mathématiques» (par opposition aux ensembles intuitifs) et dériver, à partir de notre liste d'axiomes, les propriétés des ensembles qui nous permettent de mener des raisonnements mathématiques.

Bien sûr, il faut se donner un point de départ ; pour nous, c'est un *univers*, c'est-à-dire un ensemble au sens intuitif $\mathcal{U}$, non vide. Les éléments de cet ensemble intuitif sont les ensembles mathématiques, et on voudrait pouvoir spécifier les propriétés de ces objets. Dans la suite, on utilisera, pour éviter les confusions, le terme « collection» pour parler d'un ensemble intuitif [ii] et le terme « ensemble» sera réservé aux ensembles mathématiques, c'est-à-dire aux éléments de la collection $\mathcal{U}$.

On a aussi besoin d'une relation binaire $\in$ définie sur $\mathcal{U}$, dont on voudrait qu'elle corresponde à l'idée que l'on se fait de la relation d'appartenance. Par exemple, on voudrait que si x, y sont des ensembles ayant les mêmes éléments alors $x = y$. Mais quels énoncés peut-on écrire dans le langage de la théorie des ensembles ? Eh bien, simplement ceux que l'on peut former avec les quantificateurs $\forall$ et $\exists$, les conjonctions et disjonctions, ainsi que les opérations de substitution et de restriction appliquées en partant des relations $=$ et $\in$. Pour le théoricien des modèles que vous serez, à n'en pas douter, devenu(e) en fin de semestre, il s'agit des énoncés du premier ordre dans le langage à deux éléments $\{\in, =\}$.

Un énoncé sans variable libre est dit *clos* ; il est soit vrai soit faux dans l'univers $\mathcal{U}$ où on s'est placé (c'est le principe du tiers exclu). Les axiomes de (ZF) sont des énoncés clos, et un modèle de ZF est un univers ou chacun de ces axiomes est vérifié. Donnons quelques exemples:

- Si $a \in \mathcal{U}$ on peut former l'enoncé à une variable libre et un paramètre $x = a$.
- On peut aussi formuler l'énoncé à trois variables libres $R(x, y, z)$ défini par

$$\forall t\, (t \in z) \Leftrightarrow (t = x \text{ ou } t = y)\ .$$

- Par exemple, à partir de l'énoncé R ci-dessus, on peut former l'énoncé clos

$$\forall x\, \forall y\ \exists z\ R(x, y, z)\ .$$

Chaque énoncé $R(x_1, \ldots, x_k)$ à exactement k variables libres définit une *relation*, qui est la collection des k-uplets $(a_1, \ldots, a_k)$ d'éléments de $\mathcal{U}$ tels que $R(a_1, \ldots, a_k)$ soit vrai ; rappelons que $R(a_1, \ldots, a_k)$ est l'énoncé clos obtenu en substituant $a_1, \ldots, a_k$ aux variables libres de R.

i. Il s'agit là du « paradoxe de Russell».

ii. Notons que ceci ne correspond pas à l'usage en théorie des ensembles, où on réserve le terme « collection » aux ensembles intuitifs consistant d'ensembles satisfaisant une formule du premier ordre dans le langage de la théorie des ensembles.

Certaines relations (éventuellement avec des paramètres $a_1, \ldots, a_k$) sont particulièrement importantes: ce sont les *relations fonctionnelles* ; une relation $R(x, y)$ à exactement deux variables libres est une relation fonctionnelle (à 1 argument) si

$$\forall x \forall y \forall z \, (R(x, y) \text{ et } R(x, z) \Rightarrow y = z) \ .$$

On définit de même les relations fonctionnelles à n arguments ; étant donné une relation fonctionnelle $R(x, y)$ on définit son *domaine* comme la collection des $x \in \mathcal{U}$ tels que $\exists z R(x, z)$ et son *image* comme la collection des $z \in \mathcal{U}$ tels que $\exists x R(x, z)$.

Venons-en à l'enoncé des axiomes de (ZF).

1. Axiome d'extensionnalité

On est habitué à penser que deux ensembles sont égaux si, et seulement si, ils ont les mêmes éléments. Dans le langage de la théorie des ensembles, cet énoncé s'écrit ainsi:

$$\forall x \forall y \ (\forall z \, (z \in x) \Leftrightarrow (z \in y)) \Rightarrow x = y \ .$$

En particulier, cet axiome implique que l'ensemble vide, s'il existe, est unique, et on le notera $\varnothing$. L'existence de l'ensemble vide est une conséquence des autres axiomes.

2. Axiome de fondation

Cet axiome dit que pour tout ensemble non vide x, il existe un ensemble $y \in x$ et tel que $y \cap x = \varnothing$. En particulier l'axiome de fondation interdit l'existence d'ensembles x tels que $x \in x$, ou l'existence de suites $(x_n)_{n \in \omega}$ telles que $x_{n+1} \in x_n$ pour tout n.

3. Axiome de la réunion

Dans le langage usuel, cet axiome dit que si $(X_i)_{i \in I}$ est une famille d'ensembles (i.e I est un ensemble et chaque X_i aussi) alors on peut former un nouvel ensemble dont les éléments sont exactement ceux qui appartiennent à un X_i. On a déjà dit que pour un théoricien des ensembles tout objet mathématique est un ensemble ; ainsi cet axiome doit dire que pour tout ensemble a il existe un ensemble b dont les éléments sont exactement les éléments des éléments de a. La formule correspondante est:

$$\forall a \exists b \forall x \, ((x \in b) \Leftrightarrow (\exists y (y \in a \text{ et } x \in y))) \ .$$

4. Axiome de l'ensemble des parties

Pour tout ensemble X on veut pouvoir former un ensemble dont les éléments sont les parties de X, autrement dit on a besoin de l'énoncé suivant:

$$\forall x \exists y \forall z \, [(z \in y) \Leftrightarrow (\forall t (t \in z) \Rightarrow t \in x)] \ .$$

Cet ensemble sera notée $\mathcal{P}(X)$.

5. Axiome de l'infini

Cet axiome dit qu'il existe un ensemble x tel que $\varnothing \in x$ (donc en particulier, l'ensemble vide $\varnothing$ existe), et que pour tout y, si $y \in x$ alors $y \cup \{y\} \in x$ (i.e., il existe $z \in x$ tel que $\forall t (t \in z \Leftrightarrow (t \in y \vee t = y))$, et ce z est nécessairement unique par extensionnalité). Dans le prochain chapitre on verra que cela peut être énoncé d'une manière équivalente comme « il existe un ordinal infini ».

6. Schéma d'axiomes de remplacement

Il s'agit en fait d'une infinité d'axiomes. Le schéma d'axiomes de remplacement nous permet, à partir d'une relation fonctionnelle et d'un ensemble, de former un nouvel ensemble. Par exemple il nous permet de former l'image d'un ensemble par une fonction (en tant qu'*ensemble*, et pas seulement comme une collection).

Formellement, ce schéma dit que si $E(x, y, a_1, \ldots, a_k)$ est un énoncé à paramètres $a_1, \ldots, a_k$ qui définit une relation fonctionelle à 1 variable, et a est un ensemble, alors on peut considérer l'ensemble b dont les éléments sont les images par la relation fonctionnelle E des éléments de a appartenant au domaine de notre relation fonctionnelle.

Alors le schéma d'axiomes de remplacement consiste en la liste, paramétrée par tous les énoncés $E(x, y, x_1, \ldots, x_k)$ sans paramètres et à au moins deux variables libres des énoncés suivants:

$$\forall x_1 \ldots \forall x_k \ ((\forall x \forall y \forall y' \ (E(x, y, x_1, \ldots, x_k) \text{ et } E(x, y', x_1, \ldots, x_k)) \Rightarrow y = y'))$$

$$\Rightarrow \forall t \exists u \forall y \, (y \in u \Leftrightarrow \exists x (x \in t \text{ et } E(x, y, x_1, \ldots, x_k))) \ .$$

Il nous manque encore un axiome pour obtenir toute la liste d'axiomes de Zermelo-Fraenkel. Notons déjà que d'autres énoncés (couramment cités comme des axiomes de ZF) découlent des axiomes précédents.

Schéma d'axiomes de compréhension

Ce schéma découle directement du schéma d'axiomes de remplacement ; il dit que tous les éléments d'un ensemble a qui satisfont une formule du premier ordre à une variable libre (éventuellement avec paramètres) forment un ensemble. Pour le montrer, considérons un énoncé $A(x, x_1, \dots, x_k)$ sans paramètres et à au moins 1 variable libre x. Alors on a

$$\forall x_1, \dots, x_k \, \forall x \, \exists z \, \forall y \, (y \in z) \Leftrightarrow (y \in a \text{ et } A(x, x_1, \dots, x_k)) \ .$$

Cet énoncé découle du schéma de substitution appliqué à la relation

$$x = y \text{ et } A(x, x_1, \dots, x_k) \ .$$

Axiome de la paire

Cet axiome dit que, étant donnés deux ensembles x,y il existe un ensemble z dont les éléments sont exactement x et y. En formules:

$$\forall x \forall y \, \exists z \, \forall t (t \in z) \Leftrightarrow (t = x \text{ ou } t = y) \ .$$

Ceci définit la paire $\{x, y\}$. Notons qu'avec l'axiome de l'ensemble des parties on peut former l'ensemble $\{\varnothing\}$ qui n'a qu'un seul élément $(\varnothing)$ et de même on peut former l'ensemble des parties de $\{\varnothing\}$ qui, grâce à l'axiome d'extensionnalité, a deux éléments: $\varnothing$ et $\{\varnothing\}$. On vient donc de prouver l'existence de $\{\varnothing, \{\varnothing\}\}$. Soient maintenant deux ensembles x, y quelconques. Définissons une relation fonctionnelle $R(a, b)$ par

$$(a = \varnothing \text{ et } b = x) \text{ ou } (a = \{\varnothing\} \text{ et } b = y) \ .$$

En appliquant le schéma de substitution à cette relation et à l'ensemble $\{\varnothing, \{\varnothing\}\}$ on obtient un ensemble qui n'a que x, y comme éléments.

Une paire $\{x, y\}$ est *non ordonnée* – d'après l'extensionnalité, $\{x, y\} = \{y, x\}$. Nous définissons une *paire ordonnée* formée de x et de y, ainsi que le leur *produit cartésien*, par:

$$(x, y) = \{\{x, y\}, \{x\}\},$$
$$x \times y = \{(u, v) \colon u \in x \text{ et } v \in y\}.$$

C'est un bon exercice de voir qu'il découle de nos axiomes que si x et y sont des ensembles, alors les ensemble $x \cup y$, $x \cap y$, (x, y) et $x \times y$ existent, et que:

$$(x, y) = (u, v) \qquad \Longleftrightarrow \qquad x = u \text{ et } y = v.$$

Munis de ces notions, nous pouvons représenter des objets mathématiques tels que des relations, des fonctions, etc., par des ensembles. Par exemple, une relation binaire R sur un ensemble X n'est qu'une partie de $R \subseteq X \times X$ (et on peut écrire $x \, R \, y$ au lieu de $(x, y) \in R$). De la même façon, une fonction $f \colon X \to Y$ n'est qu'une partie $f \subseteq X \times Y$ vérifiant que pour tout $x \in X$ il existe un unique $y \in Y$ tel que $(x, y) \in f$ (et on écrit $f(x) = y$).

Il serait malhonnête de conclure cette section sans évoquer le problème suivant: existe-il un univers $\mathcal{U}$ dans lequel nos axiomes sont vérifiés ? De façon malheureuse, mais peu surprenante, croire qu'il en existe un est un acte de foi. La fameux théorème de Gödel affirme en effet qu'il est impossible de démontrer (avec des théorèmes de (ZF)) que (ZF) est consistante, c'est-à-dire que ses axiomes n'entraînent pas de contradiction. Toute théorie suffisamment complexe pour permettre de développer les mathématiques classiques se trouvant dans le même cas, la solution n'est pas de changer nos axiomes ; il nous faut simplement espérer que la théorie n'est pas contradictoire.

Notes bibliographiques. Une partie de ce chapitre a été reprise dans l'excellent livre de Krivine [**Kri98**].

CHAPITRE 2

Les ordinaux

1. Bons ordres

On va se placer dans le cadre général de la théorie dite de Zermelo-Fraenkel (ZF), dont on ne sortira pas dans ce cours. Il est très vraisemblable qu'il s'agisse du cadre axiomatique que vous avez toujours utilisé, même sans le savoir, pour faire des mathématiques.

Commençons par apprendre à compter... Il est facile de compter le nombre d'éléments d'un ensemble fini en les énumérant : (zéro, si l'ensemble est vide), un, deux... et on s'arrête quand il n'y en a plus. On associe ainsi à chaque ensemble fini un entier, qui est son nombre d'éléments. Mais comment faire quand on considère un ensemble infini ? Il n'est pas clair qu'on puisse l'énumérer ; plutôt que de considérer tous les ensembles, on va commencer par considérer des ensembles munis d'un ordre permettant une énumération.

Définition 1.1. Soit X un ensemble. Un *bon ordre* sur X est une relation d'ordre $\leq$ sur X tel que tout sous-ensemble non vide de X a un plus petit élément.

Exemple. (Intuitivement, puisqu'on n'a pas encore défini formellement ces notions)
- Tout ensemble ordonné fini est bien ordonné. En particulier, l'ensemble vide est bien ordonné !
- L'ensemble des entiers naturels, avec l'ordre habituel, $(\mathbf{N}, \leq)$, est bien ordonné. En quelque sorte, c'est le « cas modèle » d'un ensemble bien ordonné (infini) – ce sont ses propriétés que l'on cherche à reproduire.
- L'ensemble $S(\mathbf{N}) = \mathbf{N} \cup \{\mathbf{N}\}$, où $\mathbf{N} > n$ pour tout $n \in \mathbf{N}$, est également bien ordonné. *infini.*
- L'ensemble des rationnels $(\mathbf{Q}, \leq)$ n'est *pas* bien ordonné.

Un *isomorphisme* entre deux ensembles bien ordonnés $(X, \leq_X)$ et $(Y, \leq_Y)$ est une bijection qui préserve l'ordre : clairement, l'un est bien ordonné si et seulement si l'autre l'est.

Définition 1.2. On dit que $S \subseteq X$ est un *segment initial* si

$$\forall x, y \in X \ \ (y \in S \text{ et } x \leq y) \Rightarrow (x \in S) \ .$$

Si $x \in X$ on notera S_x le segment initial $\{y \in S \colon y < x\}$; on l'appellera « le segment initial strict associé à x ».

On vérifie aisément que toute partie d'un ensemble bien ordonné est elle aussi bien ordonnée, avec l'ordre induit. En particulier, tout segment initial d'un ensemble bien ordonné est bien ordonné. Notons que, dans un ensemble bien ordonné X, tout segment initial propre (c'est à dire, différent de X) est de la forme S_x pour un unique $x = \min(X \smallsetminus S)$. Par conséquent, X et isomorphe à l'ensemble des ses segments initiaux propres (ordonnés par l'inclusion).

L'idée, dans notre optique de comptage, est que pour énumérer un ensemble bien ordonné X, on commence au plus petit élément, puis on prend le plus petit des autres, etc. Or, cette énumération risque d'être infinie, voire « transfinie », quand un membre de X ne peut être énuméré qu'une fois qu'une infinité d'autres membres l'ont déjà été : c'est le cas de $X = \mathbf{N} \cup \{\mathbf{N}\}$. Ceci a-t-il donc un sens ?

Nous justifions la réponse positive par le principe de preuve récurrence, qui est l'outil de base pour démontrer que tous les entiers naturels vérifient une propriété donnée P. On le connaît sous deux formes:

(R)	Si $P(0)$ et $\forall n\, P(n) \Rightarrow P(n+1)$	alors	$\forall n\, P(n)$
(R')	Si $\forall n\, [\forall k < n\, P(k)] \Rightarrow P(n)$	alors	$\forall n\, P(n)$

Théorème 1.3. *(Démonstration par récurrence transfinie) Soit P une propriété*[i] *et X un ensemble bien ordonné, tel que*

$$\forall x \in X\, [\forall X \ni y < x\, P(y)] \Rightarrow P(x)$$

Alors $P(x)$ est vraie pour tout $x \in X$.

DÉMONSTRATION. Soit $Y = \{x \in X \colon \neg P(y)\}$. Si $Y = \varnothing$ tout est bien. Sinon, Y admet un plus petit élément, y. Autrement dit, pour tout $z < y$ on $P(y)$, d'où par hypothèse $P(y)$, une contradiction. ■1.3

À titre d'exemple, montrons :

Proposition 1.4. *Soit $(X, \leq)$ un ensemble bien ordonné et $f \colon X \to X$ une application strictement croissante. Alors pour tout $x \in X$ on a $f(x) \geq x$.*

DÉMONSTRATION. D'après le principe de démonstration par récurrence transfinie, il suffirait de montrer que si $x \in X$ est tel que $\forall y < x\, (f(y) \geq y)$, alors $f(x) \geq x$. Par l'absurde, supposons que $f(x) < x$. D'un côté, pour $y = f(x) < x$, nous avons $f^2(x) = f(y) \geq y = f(x)$. D'un autre, puisque f est strictement croissante, $f^2(x) < f(x)$, une contradiction. ■1.4

Ceci permet d'obtenir un résultat de rigidité des ensembles bien ordonnés.

Proposition 1.5. *Soit $(X, \leq)$ un ensemble bien ordonné, $W \subseteq X$ un segment initial et $f \colon X \to W$ un isomorphisme. Alors $W = X$ et pour tout $x \in X$ on a $f(x) = x$. Par conséquent, si deux segments initiaux de X sont isomorphes alors ils sont égaux.*

DÉMONSTRATION. Montrons tout d'abord que $W = X$. Pour cela, prenons $x \in X$. On a $f(x) \in W$, et $f(x) \geq x$ d'après la proposition précédente. Comme W est un segment initial, on en déduit que $W = X$. Pour conclure, il suffit de remarquer qu'alors f est une bijection, dont l'inverse f^{-1} est un isomorphisme de $(X, \leq)$ sur $(X, \leq)$. Par conséquent on a $f^{-1}(x) \geq x$ pour tout x, ce qui en composant par f donne $x \geq f(x)$ et donc $f(x) = x$ pour tout $x \in X$. ■1.5

À un bon ordre sur X correspond donc, intuitivement, une énumération transfinie des membres de X. Quelle est la « longueur » de cette énumération ? Si deux ensembles bien ordonnés sont isomorphes nous dirons, toujours intuitivement, que les énumérations correspondantes sont de la même longueur, et par conséquent, une notion de longueur d'un ensemble bien ordonné X pourrait tout simplement être sa classe d'isomorphisme. Or, ce serait plus « joli » de un représentant *canonique* pour chaque telle classe.

Commençons par les longueurs finies. La longueur du vide, « zéro », sera naturellement représentées par l'ensemble vide : $0 = \varnothing$. Munis de cette représentation de 0, nous pouvons représenter la longueur « un » par $1 = \{0\} = \{\varnothing\}$, et ainsi de suite : $2 = \{0, 1\} = \{\varnothing, \{\varnothing\}\}$, ..., $n = \{0, 1, \ldots, n-1\}$, ... Effectivement, chaque tel n est bien ordonné, l'ordre strict habituel étant donné par $\in$, ou encore par $\subsetneq$. Comment étendre ceci aux longueurs infinies ? Puisque $(X, \leq) \cong (\{\text{segments initiaux propres}\}, \subseteq)$, et puisque deux segments initiaux distincts ne sont pas isomorphes, nous pouvons imaginer que le représentant canonique de la classe d'isomorphisme de X serait

$$\text{représentant canonique de } X = \{\text{représentant canonique de } S \colon S \text{ segment initial propre de } X\},$$

ce qui est d'ailleurs exactement le cas des nombres finies $n = \{0, 1, \ldots, n-1\}$. Or, c'est définition est auto-référentielle, un problème que l'on ne pourra résoudre que plus tard, avec le principe de *construction par récurrence transfinie*. Attaquons donc par un autre angle.

2. Ordinaux

Définition 2.1. Un ensemble X est dit *transitif* si

$$\forall x (x \in X \Rightarrow x \subseteq X)$$

Autrement dit, un ensemble X est transitif si, dès qu'on a $x \in z \in X$ alors on a $x \in X$ (d'où la terminologie employée).

Evidemment, on se doute que la plupart des ensembles ne sont pas transitifs ; cela dit, il existe tout de même des ensembles transitifs, comme $\varnothing$, $\{\varnothing, \{\varnothing\}\}$...

Le lemme suivant est une conséquence immédiate de la définition.

i. Là encore la notion de propriété est floue ; disons simplement qu'une propriété est quelque chose qu'on peut exprimer par un énoncé du premier ordre écrit en utilisant le langage de la théorie des ensembles.

Lemme 2.2. *Une réunion d'ensemble transitifs est encore un ensemble transitif; une intersection d'ensembles transitifs est encore un ensemble transitif.*

Définition 2.3. Un ensemble α est un *ordinal* si α est transitif et strictement bien ordonné par la relation $\in$. La classe de tous les ordinaux est notée ON. (Ceci est donc une collection d'ensemble définie par une propriété de premier ordre. Nous verrons plus tard que ON est une *classe propre*, c'est-à-dire qu'elle est trop grande pour être un ensemble.)

Notre but est de montrer que:

(i) Tout ensemble bien ordonné est isomorphe à un ordinal unique, et par un isomorphisme unique.

(ii) Les ordinaux ont de « jolies propriétés »...

Il sera plus facile de commencer par le deuxième.

Lemme 2.4. *Soit α un ordinal et $Y \subseteq \alpha$. Alors Y est transitif si et seulement si c'est un segment initial de α.*

Démonstration. Supposons que Y est transitif et $\gamma', \gamma \in \alpha$, $\gamma' < \gamma \in Y$. Alors $\gamma' \in \gamma \in Y$ d'où $\gamma' \in Y$.

Supposons maintenant que Y est un segment initial et $\gamma' \in \gamma \in Y$. Alors $\gamma' \in \gamma \in \alpha$ d'où $\gamma' \in \alpha$ et $\gamma' < \gamma \in Y$, donc $\gamma' \in Y$. ■2.4

Lemme 2.5. *Soit α un ordinal et β quelconque. Alors sont équivalents:*

(i) $\beta \in \alpha$.

(ii) β *est un segment initial propre de* α.

(iii) β *est transitif,* $\beta \subsetneq \alpha$.

En outre $\alpha \notin \alpha$ et tout les membres de α sont des ordinaux.

Démonstration. (i) $\implies$ (ii). Soit $\beta \in \alpha$. Par transitivité de α nous avons $\beta \subseteq \alpha$, et pour $\gamma \in \alpha$ nous avons $\gamma \in \beta \iff \gamma < \beta$, d'où $\beta = \{\gamma \in \alpha \colon \gamma \in \beta\} = \{\gamma \in \alpha \colon \gamma < \beta\} = S_\beta$. En particulier, β est bien un segment initial propre de α.

(ii) $\implies$ (i). Soit β un segment initial propre de α. Alors $\beta = S_\gamma$ où $\gamma = \min(\alpha \smallsetminus \beta)$. Or, d'après l'argument précédent $S_\gamma = \gamma \in \alpha$.

(ii) $\iff$ (iii). Découle de Lemme 2.4. ■2.5

Maintenant, soit $\beta \in \alpha$. Alors $\beta \subsetneq \alpha$, d'où $\beta \neq \alpha$, et β est bien ordonné par $\in$. Comme en outre β est transitif, c'est un ordinal.

Lemme 2.6. *Soit A un ensemble non vide d'ordinaux, et soit $\alpha = \bigcap A$. Alors $\alpha \in A$, c'est donc en particulier un ordinal, et $\alpha \subseteq \beta$ pour tout $\beta \in A$.*

Démonstration. L'ensemble α est transitif en tant qu'intersection d'ensembles transitifs, et par définition $\alpha \subseteq \beta$ pour tout $\beta \in A$. Si $\alpha \notin A$, c'est que $\alpha \subsetneq \beta$ pour tout $\beta \in A$. D'après Lemme 2.5, $\alpha \in \beta$ pour tout $\beta \in A$, donc $\alpha \in \alpha$, absurde. Donc $\alpha \in A$. ■2.6

Proposition 2.7. *La relation $\in$ est un ordre strict total sur ON, où elle coïncide avec $\subsetneq$. En d'autres mots, pour tous ordinaux α, β, γ :*

(i) $\beta \in \alpha \iff \beta \subsetneq \alpha$.

(ii) *Anti-réflexivité :* $\alpha \notin \alpha$.

(iii) *Transitivité : Si $\gamma \in \beta$ et $\beta \in \alpha$ alors $\gamma \in \alpha$ (l'antisymétrie $\beta \in \alpha \implies \alpha \notin \beta$ découle de* (ii), (iii)*).*

(iv) *Ou bien $\alpha \in \beta$ ou bien $\beta \in \alpha$ ou bien $\alpha = \beta$.*

Démonstration. (i) Si $\beta \in \alpha$, c'est un segment initial propre, d'où $\beta \subsetneq \alpha$. Et si $\beta \subsetneq \alpha$, puisque β est transitif c'est un segment initial de α, et propre, donc $\beta \in \alpha$.

(ii) Déjà observé.

(iii) Par transitivité de l'ensemble α.

(iv) Il découlé de (ii) et (iii) que ces possibilités sont mutuellement exclusives. Montrons qu'au moins l'une d'elle est toujours vraie. Par (i), il suffirait de montrer que $\alpha \subseteq \beta$ ou $\beta \subseteq \alpha$. Soit maintenant $\gamma = \alpha \cap \beta$. D'après Lemme 2.6, $\gamma = \alpha$ ou $\gamma = \beta$, i.e., $\alpha \subseteq \beta$ ou $\beta \subseteq \alpha$.

■2.7

À partir de maintenant nous noterons cet ordre sur les ordinaux par $<$ (ou $\leq$):

(i) $\alpha < \beta \iff \alpha \in \beta \iff \alpha \subsetneq \beta$.

(ii) $\alpha \leq \beta \iff (\alpha \in \beta \text{ ou } \alpha = \beta) \iff \alpha \subseteq \beta$.

Nous avons ainsi bien généralisé la représentation des entiers $n = \{0, 1, \ldots, n-1\}$, car en effet pour tout ordinal α :

$$\alpha = \{\beta \in ON : \beta < \alpha\}.$$

Proposition 2.8. *Soit A un ensemble non vide d'ordinaux. Alors $\bigcap A = \min A$.*

En particulier, l'ordre sur ON est bon : tout ensemble non vide d'ordinaux admet un élément minimal. Mieux que ça : toute classe non vide d'ordinaux admet un élément minimal.

Démonstration. D'après Lemme 2.6, nous avons $\bigcap A \in A$, c'est donc le minimum. Si C est une classe non vide d'ordinaux, soit $\alpha \in C$ et $C' = C \cap \alpha = \{\beta \in C : \beta < \alpha\}$. Alors C' est un ensemble (par l'axiome de compréhension). Si $C' = \varnothing$, c'est que $\alpha = \min C$. Sinon, $\min C = \min C'$. ■2.8

Lemme 2.9. *Tout ensemble transitif d'ordinaux est un ordinal.*

Démonstration. Car l'ordre $\in$ sur les ordinaux est bon. ■2.9

Proposition 2.10. *Soit A un ensemble d'ordinaux. Alors $\bigcup A$ est un ordinal De plus, $\bigcup A = \sup A$ est la borne supérieure de A : c'est le plus petit ordinal α tel que $\alpha \geq \beta$ pour tout $\beta \in A$. (Par contre, il se peut bien que $\sup A \notin A$.)*

Démonstration. L'ensemble $\bigcup A$ est un ordinal, car il est transitif (en tant que réunion d'ensembles transitifs), et aussi un ensemble d'ordinaux (car chaque $\beta \in A$ l'est). C'est clairement le plus petit ordinal tel que $\alpha \supseteq \beta$ pour tout $\beta \in A$. ■2.10

Montrons maintenant qu'il y a « beaucoup » d'ordinaux. D'abord,

Lemme 2.11. *Il n'existe pas un plus grand ordinal. Plus exactement, pour tout ordinal α, $S(\alpha) = \alpha \cup \{\alpha\} = \{\beta : \beta \leq \alpha\}$ est lui aussi un ordinal, le successeur de α (le plus petit qui soit strictement plus grand que α).*

Démonstration. Si $\gamma \in \beta \in S(\alpha)$, alors ou bien $\beta = \alpha$ ou bien $\beta < \alpha$, and dans les deux cas $\gamma \in \alpha \subseteq S(\alpha)$, d'où la transitivité. $S(\alpha)$ est ainsi un ensemble transitif d'ordinaux, donc un ordinal. Finalement, si $\alpha < \beta$ alors $S(\alpha) \subseteq \beta$, $S(\alpha)$ est donc bien le successeur de α. ■2.11

Il existe le plus petit ordinal, c'est $0 = \varnothing$. Son successeur est $1 = S(0) = \{0\}$, puis on a $2 = S(1) = \{0, 1\}$, $3 = S(2) = \{0, 1, 2\}$, et ainsi de suite.

Notons le fait suivant, qui confirme qu'il faut faire attention à ce qui est un ensemble au sens mathématique et ce qui n'en est pas un.

Proposition 2.12. *La classe des ordinaux ON n'est pas un ensemble.*

Démonstration. Si ON était un ensemble alors $\alpha = \sup ON$ serait le plus grand ordinal, un absurde. ■2.12

On dit alors que la classe ON, est une *classe propre* : une collection définie par une propriété de premier ordre (être un ordinal est bien une propriété de premier ordre) mais qui est « trop grande » pour être un ensemble.

Maintenant, nous avons tout ce qu'il faut pour démontrer que :

Théorème 2.13. *Tout ensemble bien ordonné est isomorphe à un unique ordinal, et par un isomorphisme unique.*

Démonstration. L'unicité découle de Proposition 1.5, on démontre l'existence. Soit X un ensemble bien ordonné, et soit Y l'ensemble des segments initiaux (non nécessairement propres) de X qui sont isomorphes à des ordinaux. Par unicité, chaque $I \in Y$ détermine un unique tel isomorphisme $f_I : I \cong \alpha_I$.

Si $J \subseteq I$ est un autre segment initial de X alors $f_I(J)$ est un segment initial de α_I, donc un ordinal, et $f_I{\restriction_J} : J \cong f_I(J)$ est encore un isomorphisme – donc $J \in Y$, $\alpha_J = f_I(J)$ et $f_J = f_I{\restriction_J}$.

Soit maintenant $K = \bigcup Y$, $\beta = \sup\{\alpha_I\}_{I \in Y}$, et $g = \bigcup_{I \in Y} f_I$. Alors K est encore un segment initial de X et $g : K \cong \beta$ – en d'autres mots, $K \in Y$ est maximal. Si $K \neq X$, c'est que $K = S_x$ pour un certain $x \in X$. Posons $h = g \cup \{x \mapsto \beta\}$. Alors $h : K \cup \{x\} \cong S(\beta)$, donc $K \cup \{x\} \in Y$, contredisant la maximalité de K. Du coup $K = X$, et X et bien isomorphe à un ordinal. ■2.13

Introduisons un peu de terminologie.

Définition 2.14. Un ordinal α est *successeur* s'il existe un ordinal β tel que $\alpha = S(\beta)$; si α et ni zéro ni successeur, on dit que α est un *ordinal limite* .

Notons que, si A est un ensemble d'ordinaux qui n'a pas de plus grand élément, et α est la borne supérieure de A (autrement dit, l'union des élements de A) alors α est nécessairement un ordinal limite: comme α majore A on sait que $\alpha \notin A$ puisque A n'a par hypothèse pas de plus grand élément, et si jamais on avait $\alpha = S(\beta)$ alors les propriétés du successeur nous garantissent que β majorerait aussi A, ce qui contredirait la définition de α.

Exercice 2.1. Montrer qu'un ordinal β est limite si, et seulement si, $\beta = \sup\{\eta\colon \eta < \beta\} \neq 0$.

Définition 2.15. Un ordinal α est dit *fini* si tout ordinal tel que $0 < \beta \leq \alpha$ est successeur.

On remarque que si n est un ordinal fini et $m < n$, alors m et $S(n)$ sont fini eux aussi. Notons que jusqu'à maintenant nous n'avons pas utilisé l'axiome de l'infini (ni, d'ailleurs, l'axiome de fondation...)

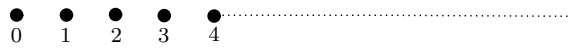
Proposition 2.16. *Modulo les autres axiomes de ZF (dont on peut même enlever l'axiome de fondation), l'axiome de l'infini est équivalente à l'énoncé : il existe l'ensemble de tous les ordinaux finis. Cet ensemble, noté ω, est un ordinal : c'est le plus petit ordinal fini, et aussi le plus petit ordinal limite*

DÉMONSTRATION. Il est clair est 0 est fini, et que si α est fini alors $S(\alpha)$ l'est aussi. Ainsi, si ω existe, l'axiome de l'infini est bien vérifié. Réciproquement, soit X un témoin que l'axiome de l'infini est bien vrai : $0 \in X$, et pour tout $x \in X$ on a aussi $S(x) = x \cup \{x\} \in X$. Montrons que X contient tous les ordinaux finis.

En effet, sinon, il existe un plus petit ordinal fini qui n'appartient pas à X, appelons le n. Alors $n > 0$ (car $0 \in X$), c'est donc un successeur : $n = S(m)$. Or m est nécessairement fini, donc $m \in X$, d'où $n \in X$, une contradiction.

Par compréhension, la collection de tous les membres de X qui sont des ordinaux finis est bien un ensemble, et tout ordinal fini y appartient, c'est donc bien ω.

Une fois que ω existe : Puisque tout ordinal plus petit qu'un ordinal fini est lui aussi fini, ω est transitif, c'est donc un ordinal. L'ordinal ω n'est pas fini, et comme tout ordinal infini est plus grand que (donc, contient) tout ordinal fini, ω est le plus petit ordinal infini. Aussi, ω n'est pas successeur (car si $\omega = S(\alpha)$ alors $\alpha < \omega$, donc α est fini, et ω aussi). C'est donc le plus petit ordinal limite. ■2.16



L'ordinal ω

Les ordinaux finis forment un modèle de l'arithmétique de Péano, et sont aussi appelés « entiers naturels ». Par conséquent, ω est l'ensemble des entiers naturels, que l'on dénote habituellement par $\mathbf{N}$.

Avant de passer à l'arithmétique des ordinaux, récapitulons les propriétés qu'il faut particulièrement retenir pour pouvoir les manipuler.

- Tout ordinal est un ensemble bien ordonné, et tout ensemble bien ordonné est isomorphe à un ordinal unique. En particulier deux ordinaux isomorphes sont nécessairement égaux ; de plus, pour deux ordinaux α, β on a soit $\alpha < \beta$, soit $\alpha = \beta$, soit $\beta < \alpha$.
- Pour tout ordinal α, on a $\alpha = \{\beta \in ON\colon \beta < \alpha\}$.
- La réunion d'un ensemble d'ordinaux E est un ordinal, qui est la borne supérieure de E.
- L'intersection d'un ensemble d'ordinaux E est un ordinal, qui est le plus petit élément de E.
- Il existe deux types d'ordinaux: les ordinaux successeurs (ceux qui ont un plus grand élément) et les ordinaux limites (ceux qui n'ont pas de plus grand élément).

3. Récurrence transfinie et arithmétique des ordinaux

Rappelons encore une fois les deux formes du principe de récurrence sur les entiers :

(R)	Si $P(0)$ et $\forall n\, P(n) \Rightarrow P(n+1)$	alors	$\forall n\, P(n)$
(R')	Si $\forall n\, [\forall k < n\, P(k)] \Rightarrow P(n)$	alors	$\forall n\, P(n)$

Au lieu de considérer tous les ensembles bien ordonnés, nous pouvons nous restreindre aux ordinaux.

Théorème 3.1. *(Démonstration par récurrence transfinie bis) Soit P une propriété et α un ordinal, tels que l'une des deux conditions suivantes est vérifiée :*

(i) *Pour tout $\beta < \alpha$: si $\beta = 0$ alors $P(\beta)$; si $\beta = S(\gamma)$ alors $P(\gamma) \Rightarrow P(\beta)$; et si β est limite alors $[\forall \gamma < \beta, P(\gamma)] \Rightarrow P(\beta)$.*

(ii) *Pour tout $\beta < \alpha$: $[\forall \gamma < \beta, P(\gamma)] \Rightarrow P(\beta)$.*

Alors $P(\beta)$ pour tout $\beta < \alpha$.

La même chose est d'ailleurs vrai quand on remplace « pour tout $\beta < \alpha$ » par « pour tout $\beta \in ON$ ».

DÉMONSTRATION. La première condition implique la deuxième, et on a déjà vu que la deuxième suffit. ■3.1

Un outil plus fort est celui de la construction par récurrence : pour construire une application $f\colon \mathbf{N} \to X$, il suffit de préciser d'abord $f(0)$, puis, pour tout n, comment construire $f(n+1)$ à partir de $f(n)$. Ou encore (et c'est équivalent !), il suffit de préciser une loi qui permet de construire $f(n)$ à partir de la restriction de f à $\{0, \dots, n-1\}$ $(= n...)$ C'est cette deuxième forme que nous allons généraliser.

Théorème 3.2. *Soit $(X, \leq)$ un ensemble bien ordonné, et G une loi qui associe à chaque fonction g dont le domaine est un segment initial propre de X un élément (c'est à dire, un ensemble) $G(g)$. Alors il existe une unique fonction f de domaine X telle que l'on ait, pour tout $x \in X$,*

$$f(x) = G(f\restriction_{S_x}) \ .$$

DÉMONSTRATION. Puisque chaque ensemble bien ordonné X est isomorphe à un ordinal, nous pouvons réduire le problème au cas où $X = \xi$ est un ordinal. L'hypothèse de récurrence est que le théorème est vrai pour tout ordinal $\alpha < \xi$. Autrement dit, pour tout $\alpha < \xi$ il existe une unique h_α de domaine α telle que

$$(*_\alpha) \qquad h_\alpha(\gamma) = G(h_\alpha\restriction_\gamma) \qquad \forall \gamma < \alpha.$$

Nous observons que si $\beta < \alpha < \xi$ alors $h_\alpha\restriction_\beta$ vérifie $(*_\beta)$, d'où $h_\beta = h_\alpha\restriction_\beta$.

Pour $\alpha < \xi$, posons $f(\alpha) = G(h_\alpha)$. Si $\beta < \alpha < \xi$ alors $f(\beta) = G(h_\beta) = G(h_\alpha\restriction_\beta) = h_\alpha(\beta)$. Ainsi $f\restriction_\alpha = h_\alpha$, et $f(\alpha) = G(h_\alpha) = G(f\restriction_\alpha)$, comme voulu.

Pour l'unicité, supposons que f' possède aussi la propriété $f'(\alpha) = G(f'\restriction_\alpha)$. On démontre alors par récurrence sur β que $f(\beta) = f'(\beta)$ pour tout $\beta < \xi$, i.e., $f = f'$, et la preuve est complète. ■3.2

Exemple 3.3. Nous avons déjà vu un exemple d'une construction par récurrence : c'était le cas du Théorème 2.13, qui dit que tout ensemble bien ordonné est isomorphe à un ordinal. Certes, nous avons ce fait dans la preuve de Théorème 3.2, mais ceci ne sert qu'à rendre la preuve plus élégante, et n'est surtout pas indispensable (ça pourrait d'ailleurs être un bon exercice de compréhension le faire directement).

Maintenant, soit G la loi associant à une fonction f son image. Alors d'après Théorème 3.2, pour tout X bien ordonné il existe une fonction f de domaine X telle que $f(x) = \operatorname{img} f\restriction_{S_x} = \{f(y)\colon y < x\}$. Montrons que chaque $f(x)$ est un ordinal, par récurrence sur x. En effet, par l'hypothèse de récurrence, $f(x) = \{f(y)\colon y < x\}$ est un ensemble d'ordinaux, et il suffirait de montrer qu'il est transitif. Pour cela, si $\gamma \in \beta \in f(x)$ alors $\beta = f(y)$ pour un $y < x$ et $\gamma = f(z)$ pour un $z < y$, d'où $\gamma \in f(x)$. Le même argument montre que $\alpha = \{f(x)\colon x \in X\}$ est un ordinal. En outre, $f(y) \in f(x)$ si et seulement si $y < x$ (« si » par définition de f, « seulement si » car si $y \geq x$ alors $f(y) \supseteq f(x)$, ce qui empêche $f(y) \in f(x)$). Ainsi $f\colon X \to \alpha$ est un isomorphisme.

Le plus souvent on utilisera ce principe quand $X = \xi$ est un ordinal, et très rarement sous cette forme abstraite. L'idée du principe est que construire une fonction $f\colon \xi \to Y$ c'est dire, pour chaque $\alpha < \xi$, comment construire $f(\alpha)$ à partir de la restriction $f\restriction_\alpha$. Et pour cela, il suffit de préciser :

(i) Cas zéro : $f(0)$.

(ii) Cas successeur : quand $\alpha = S(\beta)$, comment obtenir $f(\alpha)$ à partir de $f(\beta)$.

(iii) Cas limite : quand α est limite, comment obtenir $f(\alpha)$ à partir de $f\restriction_\alpha$.

Plutôt que de donner une preuve du théorème de construction par récurrence transfinie, donnons un exemple de preuve rédigée en utilisant un objet construit par récurrence transfinie.

Proposition 3.4. *Soit $(X, <)$ un ensemble bien ordonné, et $A \subseteq X$ un sous-ensemble non vide. Alors $(A, <)$ est encore bien ordonné, et est isomorphe à un segment initial de X.*

Notons que cette proposition implique en particulier que, étant donnés deux ordinaux α, β, $\alpha \leq \beta$ si, et seulement si, il existe une application strictement croissante de α dans β (et dont l'image n'est pas forcément un segment initial).

Démonstration. Il est immédiat que $(A, <)$ est bien ordonné. Pour construire un isomorphisme de A sur un segment initial de X, on procède par récurrence transfinie, en définissant $f\colon A \to X$ de la façon suivante:

- On pose $f(\min(A)) = \min(X)$.
- Si a est le successeur (dans A) de a', alors on définit $f(a)$ comme le successeur (dans X) de $f(a')$.
- Si a est limite (dans A) alors on pose $f(a) = \sup\{f(a')\colon a' \in A \text{ et } a' < a$.

On vérifie aisément (par récurrence transfinie!) que $f(a) \leq a$ pour tout $a \in X$ et que f est strictement croissante ; pour voir que son image est un segment initial de X, on note pour $a \in A$ S_a^A le segment initial associé à a dans A, et on va vérifier par récurrence transfinie que pour tout a dans A $f(S_a^A) = S_{f(a)}$. La propriété est vraie par construction pour $a = \min(A)$. On suppose maintenant que $a \in A \setminus \{\min(A)\}$ est tel que notre propriété est vraie pour tout $a' < a$.

- Si a est le successeur (dans A) de a', alors $f(S_a^A) = S_{a'} \cup \{f(a)\}$, et $f(a)$ est le plus petit élément de $X \setminus f(S'_a)$, donc le fait que $f(S_{a'}^A)$ soit un segment initial entraîne que $f(S_a^A)$ est aussi un segment initial de X.
- Si $a \neq \min(A)$ est limite dans A, alors

$$f(S_a^A) = \bigcup_{a'<a} f(S_{a'}^A) \ .$$

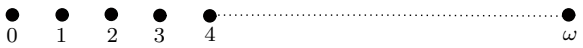
  Comme chaque $f(S_{a'}^A)$ est, par hypothèse de récurrence, un segment initial, et qu'une réunion de segments initiaux est encore un segment initial, la propriété est vraie au rang a.

Ceci achève la démonstration [ii]. □

On pourrait définir les opérations ordinales en décrivant des opérations sur les bons ordres ; pour gagner du temps dans ces notes, on va simplement énoncer une définition par récurrence transfinie. Rappelons qu'on note $S(\beta)$ le successeur d'un ordinal β, c'est-à-dire le plus petit ordinal strictement plus grand que β.

Définition 3.5. (addition ordinale) Soit α un ordinal. On pose $\alpha + 0 = \alpha$, puis on définit par récurrence transfinie sur $\beta \in ON$ l'addition ordinale $\alpha + \beta$ en posant:

$$\alpha + \beta = \begin{cases} S(\alpha + \gamma) & \text{si } \beta = S(\gamma) \\ \sup\left(\{\alpha + \xi\colon \xi < \beta\}\right) & \text{si } \beta \text{ est limite} \end{cases}$$



L'ordinal $\omega + 1$

Par exemple, on a $1 + \omega = \sup\{1 + n\colon n < \omega\} = \omega$. Par contre, $\omega + 1 \neq \omega$ puisque $\omega + 1$ a un plus grand élément ; *l'addition ordinale n'est donc pas commutative*. Intuitivement, l'addition de deux ordinaux correspond à mettre "bout à bout" α et β ; l'ordre dans lequel on « recolle» les deux ordinaux est important!

Exemple. Utilisons une démonstration par récurrence transfinie pour montrer que l'addition est associative, et que si $\alpha \neq \beta$ alors pour tout δ on a $\delta + \alpha \neq \delta + \beta$.

On veut commencer par montrer que, étant donnés trois ordinaux α, β, γ on a $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$. Raisonnons par récurrence sur γ ; autrement dit, on va essayer de démontrer que pour tout ordinal γ la propriété $P(\gamma)$ définie par « Pour tous les ordinaux α, β on a $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$» est vraie.

Notons qu'il n'y a rien à montrer si $\gamma = 0$; ensuite supposons que γ est tel que $P(\eta)$ est vrai pour tout $\eta < \gamma$. Si γ est le successeur d'un certain δ, alors on a pour toute paire d'ordinaux (α, β) (en utilisant la définition de l'addition ordinale et notre hypothèse de récurrence):

$$(\alpha + \beta) + \gamma = S((\alpha + \beta) + \delta) = S(\alpha + (\beta + \delta)) = \alpha + S(\beta + \delta) = \alpha + (\beta + \gamma)$$

ii. Notons qu'on n'avait pas vraiment besoin de distinguer le cas $a = \min(A)$ des autres cas limite, on l'a simplement fait pour éviter au lecteur de se poser des problèmes de zérologie.

On voit donc que $P(\gamma)$ est vraie; reste à traiter le cas où γ est un ordinal limite. Dans ce cas on a (toujours en utilisant la définition de l'addition, notre hypothèse de récurrence, et le fait que $\beta+\gamma$ est limite si γ l'est, ce qui est une conséquence directe de la définition de l'addition ordinale):

$$\begin{aligned}(\alpha+\beta)+\gamma &= \sup\{(\alpha+\beta)+\eta\colon \eta<\gamma\} = \sup\{\alpha+(\beta+\eta)\colon \eta<\gamma\}\\ &= \alpha+\sup\{\beta+\eta\colon \eta<\gamma\} = \alpha+(\beta+\gamma)\ .\end{aligned}$$

On voit donc que $P(\gamma)$ est vraie, et on a fini de prouver que l'addition ordinale est associative; ici le lecteur attentif devrait se rendre compte que, même s'il n'y a pas de difficulté particulière dans le raisonnement, il faut apporter un certain soin à la rédaction pour qu'elle soit correcte; par conséquent il faut s'entraîner à écrire ce type de démonstration!
Venons-en à la deuxième propriété ci-dessus; fixons δ et α et essayons de montrer que pour tout $\beta>\alpha$ on a $\delta+\alpha<\delta+\beta$. Raisonnons par récurrence sur β; si $\beta=S(\alpha)$ alors notre propriété est vraie puisque pour tout ordinal γ on a $S(\gamma)>\gamma$. Maintenant si $\beta>S(\alpha)$ est tel que notre propriété est vraie pour tout $\eta<\beta$, alors:

- Si $\beta=S(\eta)$ on a $\delta+\beta=\delta+S(\eta)=S(\delta+\eta)>\delta+\eta>\delta+\alpha$.
- Si β est limite alors on a $\delta+\beta=\sup\{\delta+\eta\colon \eta<\beta\}>\delta+S(\alpha)>\delta+\alpha$.

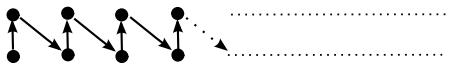
Ceci achève la démonstration; notons pour rassurer le lecteur que la rédaction ci-dessus est particulièrement lourde et détaillée, et que par la suite on évitera de trop rentrer dans le détail de raisonnements élémentaires comme celui-ci. Mais il faut vérifier qu'un raisonnement d'apparence élémentaire ne comporte pas de difficulté cachée, et c'est ce que nous avons fait ci-dessus. □

Dans la suite on utilisera toujours la notation $\alpha+1$ pour désigner le successeur d'un ordinal α. Répétons une dernière fois que $\alpha+1$ est simplement l'ordinal obtenu en rajoutant à α un élément qui majore tous les éléments de α; dans le monde un peu étrange des ordinaux, cela signifie que $\alpha+1=\alpha\cup\{\alpha\}$.

Définition 3.6. (multiplication ordinale) Soit α un ordinal. On pose $\alpha.0=0$, puis on définit par récurrence transfinie sur $\beta\in ON$ la multiplication ordinale $\alpha.\beta$ en posant:

$$\alpha.\beta=\begin{cases}(\alpha.\gamma)+\alpha & \text{si } \beta=\gamma+1\\ \sup\left(\{\alpha.\xi\colon \xi<\beta\}\right) & \text{si } \beta \text{ est limite}\end{cases}\ .$$

Cette fois on a $2.\omega=\omega$; l'idée de la multiplication ordinale est que "faire le produit de α par β, c'est mettre bout à bout β copies de α". Le dessin suivant essaie de justifier graphiquement l'égalité $2.\omega=\omega$.



$2.\omega=\omega$

Exercice 2.2. Utiliser une démonstration par récurrence transfinie pour montrer que la multiplication est associative, et que si $\alpha>0$ alors pour tout $\gamma>1$ on a $\alpha<\alpha.\gamma$. Pourver aussi que $\alpha(\beta+\gamma)=\alpha.\beta+\alpha.\gamma$.

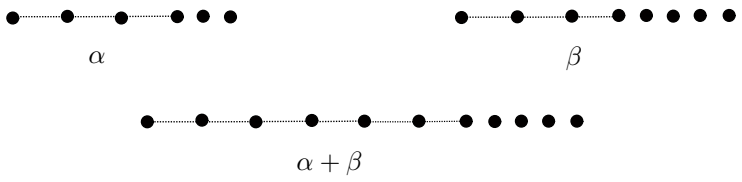
Les deux opérations définies ci-dessus sont associatives, on a bien comme attendu $\alpha+\alpha=\alpha.2$, par contre attention encore à la non-commutativité: on a vu que $1+\omega=\omega$ tandis que $\omega+1\neq\omega$ puisque $\omega+1$ est successeur; de même $2.\omega=\omega$ tandis que $\omega.2=\omega+\omega>\omega$.

Exercice 2.3.
Décrire des opérations sur les bons ordres qui donnent naissance à l'addition et à la multiplication des ordinaux (pour la somme ordinale, on pourra s'inspirer du dessin ci-dessous).

Un exercice pour vous entraîner aux démonstrations par récurrence transfinie:

Exercice 2.4. Montrer que tout ordinal α peut s'écrire de façon unique sous la forme $\alpha=\beta+n$, où β est un ordinal limite et n est fini.

Il nous reste à definir une dernière opération arithmétique sur les ordinaux: l'exponentiation.



La somme de deux ordinaux

Définition 3.7. Soit α un ordinal. On pose $\alpha^0 = 1$, puis on définit, par récurrence transfinie sur $\beta \in ON$, α^β en posant:

$$\alpha^\beta = \begin{cases} \alpha^\gamma.\alpha & \text{si } \beta = \gamma + 1 \\ \sup\left(\{\alpha^\xi : \xi < \beta\}\right) & \text{si } \beta \text{ est limite} \end{cases}$$

Par récurrence transfinie, on vérifie les propriétés suivantes.
- Etant donnés trois ordinaux α, β, γ, on a $\left(\alpha^\beta\right)^\gamma = \alpha^{\beta.\gamma}$.
- Etant donnés trois ordinaux α, β, γ, on a $\alpha^\beta.\alpha^\gamma = \alpha^{\beta+\gamma}$.
- Etant donnés trois ordinaux α, β, γ, si $\beta > \gamma$ alors $\alpha^\beta > \alpha^\gamma$.

Attention, les ordinaux et leur arithmétique ont beaucoup de propriétés contre-intuitives, et il faut donc toujours vous assurer que vous savez démontrer ce que vous affirmez à leur sujet. Par exemple, montrons qu'il existe un ordinal β tel que $\omega^\beta = \beta$: partons par exemple de $\beta_0 = \omega$, puis définissons par récurrence $\beta_{n+1} = \omega^{\beta_n}$. La troisième propriété ci-dessus nous permet de vérifier que cette suite est strictement croissante; définissons β comme la borne supérieure des β_n. C'est un ordinal limite (toute borne supérieure d'une suite infinie strictement croissante est limite) et on a donc, par définition de l'exponentiation aux ordinaux limite,

$$\omega^\beta = \sup\{\omega^{\beta_n} : n < \omega\} = \sup\{\beta_{n+1} : n < \omega\} = \beta \ .$$

Question. L'ordinal ω jouait-il un rôle particulier dans le raisonnement ci-dessus, ou peut-il être remplacé par d'autres ordinaux α ? Et que pensez-vous de l'existence d'un ordinal $\beta > 1$ tel que $\beta = \beta^\omega$?

Notes bibliographiques. La présentation étant complètement standard, il serait un peu vain de présenter des sources bibliographiques ; le lecteur intéressé par une approche intuitive de la théorie des ensembles est invité à consulter [**Hal74**]. On pourra aussi avec profit consulter les notes de cours de Tuna Altinel des années précédentes, ainsi que les notes de cours de Patrick Dehornoy (on trouvera des liens vers ces notes sur la page web du cours).

CHAPITRE 3

L'axiome du choix

Nous avons vu qu'un bon ordre sur un ensemble X correspond (d'une façon unique) à une bijection entre X et un (unique) ordinal α, ce qui reviens à une énumération de X de longueur α : $X = \{x_\beta\}_{\beta<\alpha}$, où $\beta \mapsto x_\beta$ est l'isomorphisme. Si, par exemple, X est un ensemble de contraintes qu'une construction devrait satisfaire, ceci permet de les traiter « une par une » par récurrence transfinie. Mais tout ensemble X admet-il un bon ordre ? Autrement dit, le principe suivant est-il vrai ?

Principe du bon ordre. *Tout ensemble admet un bon ordre.*

Passons à autre chose. Soit $\{X_i\}_{i\in I}$ une famille d'ensembles indexée par un ensemble I. Nous définissons son *produit cartésien* :

$$\prod_i X_i = \left\{ \text{fonctions } f\colon I \to \bigcup_i X_i \text{ t.q. } f(i) \in X_i \, \forall i \in I \right\}.$$

Supposons que $X_i \neq \varnothing$ pour tout i (sinon le produit est évidemment vide) et posons une question qui peut paraître étrange : le produit est-il non vide ? Autrement dit, peut-on choisir un élément de chaque X_i *simultanément* ? Si $X_i = X$ pour tout i, le produit est encore égal à la puissance

$$X^I = \{\text{fonctions } f\colon I \to X\},$$

qui est non vide (contient les fonctions constantes, par exemple). Si I est fini (isomorphe à un ordinal fini), nous pouvons encore montrer que $\prod_i X_i \neq \varnothing$: autrement dit, nous pouvons toujours effectuer un nombre fini de choix. Mais qu'en est-il d'un nombre infini ? Énonçons-le en tant que principe :

Axiome du Choix (AC). *Le produit d'une famille d'ensembles non vides est non vide.*

Pour un ensemble X, soit $\mathcal{P}(X)^- = \mathcal{P}(X) \setminus \{\varnothing\}$. Une *fonction de choix* pour X est une fonction $f\colon \mathcal{P}^-(X) \to X$ telle que pour tout $\varnothing \neq Y \subseteq X$: $f(Y) \in Y$. Il n'est pas difficile de vérifier que AC est équivalent à : tout ensemble admet une fonction de choix.

Finalement, considérons un principe fortement utilisé dans plusieurs domaines des mathématiques (pour démontrer Hahn-Banach, l'existence d'idéaux maximaux, etc...)

Lemme de Zorn. *Soit $(S, \leq)$ un ensemble partiellement ordonné où toute chaîne est majorée (un ensemble* inductif*). Alors S admet un élément maximal.*

Nous démontrerons que modulo ZF, ces trois principes sont équivalents.

Lemme 0.8 (Hartogs)**.** *Soit X un ensemble quelconque. Alors il existe un ordinal α qui ne s'injecte pas dans X.*

Le plus petit tel ordinal s'appelle l'*ordinal de Hartogs* de X.

DÉMONSTRATION. Soit α l'ensemble de tous les ordinaux qui s'injectent dans X.

D'abord, pourquoi est-ce un ensemble ? Soit Y l'ensemble des paires $(S, \leq)$ où $S \subseteq X$ et $\leq$ est un bon ordre sur S. Alors $Y \subseteq \mathcal{P}(X) \times \mathcal{P}(X \times X)$ est un ensemble par compréhension. Chaque élément de Y est isomorphe à un ordinal unique, et α est exactement l'ensemble de tous ces ordinaux. Ainsi, par l'axiome de remplacement, α est bien un ensemble.

Il est clair de la définition de α que si $\gamma < \beta \in \alpha$ alors $\gamma \in \alpha$. Ainsi α est transitif, donc lui-même un ordinal. Puisque $\alpha \notin \alpha$, α ne s'injecte pas dans X, et c'est d'ailleurs le plus petit ordinal possédant cette propriété. $\blacksquare_{0.8}$

Théorème 0.9. *Modulo ZF, dont équivalents :*

(i) *Le Lemme de Zorn.*

(ii) *Le Principe du bon ordre.*

(iii) *L'Axiome du choix.*

DÉMONSTRATION. (i) $\Longrightarrow$ (ii). Soit X donné, et soit Y l'ensemble de $(S, \leq)$ où $S \subseteq X$ et $\leq$ est un bon ordre sur S, comme dans la preuve précédente. Disons que $(S, \leq) < (S', \leq')$ si $(S, \leq)$ est un segment initial de $(S', \leq')$ (avec l'ordre induit). Ceci est bien un ordre partiel.

Soit $C = \{(S_i, \leq_i)\}$ une chaîne dans Y, et posons $S = \bigcup S_i$, $\leq = \bigcup \leq_i$. Puisque C est une chaîne, $\leq$ est bien un ordre total sur S : sa restriction à chaque S_i est $\leq_i$, et pour tous $x, y, z \in S$ il existe un $(S_i, \leq_i) \in C$ tel que $x, y, z \in S_i$... De la même manière, chaque S_i est un segment initial de S. Soit maintenant $\varnothing \neq A \subseteq S$, disons $x \in A$, donc $x \in S_i$ pour un certain i. Alors $A \cap S_i$ admet un minimum, et puisque S_i est un segment initial de chaque S c'est le minimum de A. Ainsi $(S, \leq) \in Y$ majore C.

D'après le Lemme de Zorn il existe $(S, \leq) \in Y$ maximal. Il suffit donc de montrer que $S = X$: sinon, soit $a \in X \smallsetminus S$, $S' = S \cup \{a\}$, et $\leq'$ défini sur S' par $a >' b$ pour tout $b \in S$. Alors $(S', \leq') \in Y$, une contradiction à la maximalité.

(ii) $\Longrightarrow$ (iii). Soit X un ressemble quelconque et $\leq$ un bon ordre sur X. Nous définissons $\varphi\colon \mathcal{P}^-(X) \to X$ par :

$$\varphi(A) = \min A.$$

Alors φ est une fonction de choix pour X.

(iii) $\Longrightarrow$ (i). Soit $(S, \leq)$ un ensemble inductif, et soit $\varphi\colon \mathcal{P}^-(S) \to S$ une fonction de choix pour S. Soit α l'ordinal de Hartogs de S, et $* \notin S$. On définit par récurrence transfinie une application $g\colon \alpha \to S \cup \{*\}$ comme suit. Si $* \notin \operatorname{img} g\restriction_\beta$, et $\operatorname{img} g\restriction_\beta$ admet un majorant strict dans S, soit R l'ensemble de tous les majorants stricts de $\operatorname{img} g\restriction_\beta$ et $g(\beta) = \varphi(R)$. Sinon, $g(\beta) = *$.

En particulier, si $\gamma < \beta < \alpha$ et $g(\beta) \neq *$ alors $g(\beta) > g(\gamma)$. Ainsi, si $* \notin \operatorname{img} g$ on obtient une injection $\alpha \hookrightarrow S$, ce qui n'est pas possible. Il existe dont un plus petit β tel que $g(\beta) = \times$. Alors $\operatorname{img} g\restriction_\beta \subseteq S$ est une chaîne, et admet un majorant $s \in S$, mais aucun majorant strict. En conséquence, s est maximal dans S. ■$_{0.9}$

L'axiome du choix a de nombreuses conséquences en mathématiques, dont certaines paraissent pathologiques. L'exemple le plus connu est sans doute l'existence de parties non Lebesgue-mesurables dans $\mathbf{R}$. Certains mathématiciens refusent de ce fait l'axiome du choix ; notons tout de même que, contrairement à une idée reçue, celui-ci n'est *pas* équivalent à l'existence de parties non Lebesgue-mesurables ; autrement dit, supposer que toute partie de $\mathbf{R}$ est Lebesgue-mesurable est plus fort que supposer que l'axiome du choix est faux. Il en va de même du paradoxe de Banach-Tarski: c'est une conséquence de l'axiome du choix qui ne lui est pas équivalente (ce qui ne fait sans doute que renforcer l'envie de refuser l'axiome du choix!).
Par ailleurs, l'axiome du choix a de nombreuses conséquences qui, elles, paraissent très utiles: théorème de la base incomplète ou lemme de Krull pour les algébristes, théorème de Tychonov pour les analystes... Et bien sûr on a vu que la théorie des ensembles devient très vite très compliquée [i] si on n'a pas l'axiome du choix, puisqu'il est déjà difficile de compter le nombre d'éléments d'un ensemble quelconque. Un autre exemple de difficulté liée à l'absence de l'axiome du choix se trouve dans l'exercice suivant.

Exercice 3.1. Montrer que l'axiome du choix est équivalent à l'énoncé suivant: si X, Y sont deux ensembles et $f\colon X \to Y$ est une surjection, alors il existe $g\colon Y \to X$ telle que $f(g(y)) = y$ pour tout $y \in Y$.

Dans la suite de ces notes, on utilisera sans vergogne l'axiome du choix sous ses différentes formes. Ceci ne correspond pas forcément aux usages actuels en théorie des ensembles, où l'on se contente souvent d'utiliser des formes plus faibles de l'axiome du choix, suffisantes pour faire de l'analyse mais n'impliquant pas que tous les ensembles sont bien ordonnables.

Ainsi, on pourrait être tenté de se contenter de l'*axiome du choix dénombrable*. Cet axiome, qui dit que si X_n est non vide pour chaque $n < \omega$ alors $\prod_{n<\omega} X_n \neq \varnothing$, (ou, de manière équivalente, qu'on peut choisir de manière simultanée un point dans chaque membre d'une famille *dénombrable* d'ensembles non vides), est fondamental pour le développement de l'analyse. Par exemple, montrer que les deux définitions classiques de la continuité pour des fonctions de $\mathbf{R}$ dans $\mathbf{R}$ (par les suites/image inverse d'un fermé est fermé) sont équivalentes requiert l'axiome du choix dénombrable...

i. Ce qui n'est pas forcément une mauvaise chose!

Exercice 3.2. Montrer que l'axiome du choix dénombrable entraîne que toute réunion dénombrable d'ensembles dénombrables est dénombrable (rappelons qu'un ensemble est dénombrable s'il est équipotent à ω).
Montrer que si toute réunion dénombrable d'ensembles dénombrables est dénombrable alors tout produit dénombrable de parties dénombrables non vides est non vide[ii].

En réalité, l'axiome du choix dénombrable n'est pas suffisant pour les analystes. En effet, en analyse on a souvent besoin de construire des suites en utilisant le principe suivant: supposons qu'étant donnés $x_1, \ldots, x_n$ tel que $P(\{x_1, \ldots, x_n\})$ est satisfaite (où P est une certaine propriété des ensembles finis) j'arrive à trouver un x tel que $\{x_1, \ldots, x_n, x\}$ a la propriété P; alors je suis capable de construire une *suite* $(x_n)_{n\in\mathbf{N}}$ tel que pour tout n on ait $P(\{x_1, \ldots, x_n\})$.

Ce procédé est à la base de beaucoup de constructions par « approximations successives» et devient légal quand on s'autorise à appliquer l'*axiome des choix dépendants*.

Définition 0.10. L'*axiome des choix dépendants* est l'énoncé suivant:
Soit X un ensemble et R une relation binaire sur X telle que pour tout $a \in X$ il existe $b \in X$ satisfaisant $a\, R\, b$. Alors il existe une suite $(x_n)_{n\in\mathbf{N}}$ d'éléments de X tels que $x_n\, R\, x_{n+1}$ pour tout n.

Notons que l'axiome du choix implique l'axiome des choix dépendants, qui implique à son tour l'axiome du choix dénombrable; on peut montrer qu'aucune des implications réciproques n'est vraie. Enfin, remarquons que l'axiome des choix dépendants, s'il est suffisant pour développer l'analyse classique, ne permet pas de démontrer l'existence d'ensembles non Lebesgue-mesurables; il semble raisonnable d'affirmer que cet axiome est accepté par une grande majorité des mathématiciens contemporains, y compris ces êtres étranges que sont les théoricien(ne)s des ensembles.

Pour simplifier l'exposition dans la suite, on utilisera l'axiome du choix « classique ». Il est en tous les cas important de savoir quand la démonstration d'un théorème utilise l'axiome du choix.

ii. et ce fait est indépendant de ZF.

CHAPITRE 4

Cardinaux

1. Définition des cardinaux

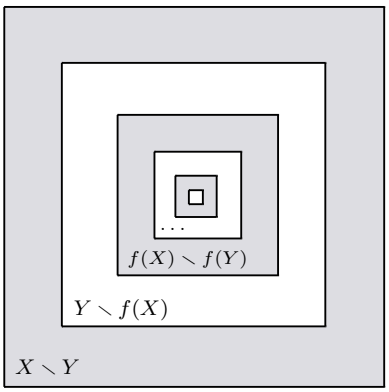
Nous cherchons ici à mesurer les « tailles » des ensembles, ce qui revient à les comparer.

Définition 1.1. On dit que la *cardinalité* (ou, dans certains textes, la *puissance*) d'un ensemble X est inférieure à celle de Y, et on note $|X| \leq |Y|$, s'il existe une injection de X dans Y, et on dit que X et Y ont la même cardinalité, ou qu'ils sont *équipotents* (noté $|X| = |Y|$), s'il existe une bijection de X sur Y.

Notez bien que pour l'instant nous n'avons pas donné un sens à la cardinalité « $|X|$ » en dehors d'une telle comparaison. Montrons déjà que les deux notions (égalité et ordre sur les cardinalités) sont bien compatibles :

Théorème 1.2. (Schröder-Bernstein)
Si $|X| \leq |Y|$ *et* $|Y| \leq |X|$ *alors* $|Y| = |X|$.

Démonstration. Soit X, Y deux ensembles et $f\colon X \to Y$, $g\colon Y \to X$ deux injections. Bien sûr, on a $X \supseteq g(Y) \supseteq g(f(X))$, et $g \circ f$ est une injection de X dans X.
On voit donc qu'il suffit de prouver que, si X est un ensemble, $f\colon X \to X$ une injection et $Y \subseteq X$ est tel que $f(X) \subseteq Y \subseteq X$ alors il existe une bijection de X sur Y.
En réfléchissant à ce cas, on est amené à considérer le dessin suivant:



On voit apparaître des « couronnes»: $X \setminus Y$, $Y \setminus f(X)$, $f(X) \setminus f(Y)$, etc.
Les couronnes « d'ordre impair»(en blanc sur le dessin) sont toutes contenues dans Y ; tandis que seule la première couronne d'ordre pair n'est pas contenue dans Y, et f envoie chaque couronne d'ordre pair sur la couronne suivante. Pour construire la bijection recherchée, on n'a donc qu'à laisser tous les points blancs fixes, et décaler les points gris d'une couronne en utilisant f.
Formellement, on définit une suite d'ensembles disjoints $X_i \subseteq X$ en posant $X_i = f^i(X \setminus Y)$ $(= f^i(X) \setminus f^i(Y))$; puis on définit une fonction $g\colon X \to Y$ en posant

$$g(x) = \begin{cases} f(x) & \text{si } x \in \bigcup X_i \\ x & \text{sinon} \end{cases}$$

Par définition il est clair que g est une injection dont l'image est contenue dans Y, d'autre part il est facile de vérifier, en utilisant le fait que $g(X_i) = f(X_i) = X_{i+1}$ pour tout i, que $g(X) = Y$. □

Autrement dit, s'il existe une injection de X dans Y et une injection de Y dans X alors il existe une bijection de X sur Y, et nos notations sont bien cohérentes et définissent un ordre partiel sur les cardinalités, c'est à dire sur les classes d'équipotence. Nous nous posons deux questions naturelles :

cet ordre est-il total, et deux, comment trouver pour chaque X un représentant canonique de sa classe d'équipotence ? D'une certain manière, les deux se heurtent aux même obstacle, l'axiome du choix.

Considérons d'abord la question du représentant.

Définition 1.3. Un *cardinal* est un ordinal qui n'est équipotent à aucun ordinal strictement plus petit.

Si deux cardinaux κ et λ sont équipotents ils sont donc égaux, et un ensemble est équipotent tout au plus à un unique cardinal. En outre, si X et Y sont équipotents à des cardinaux κ et λ, respectivement, alors $|X| \leq |Y|$ si et seulement si $\kappa \leq \lambda$ et $|X| = |Y|$ si et seulement si $\kappa = \lambda$. Nous pouvons donc définir le *cardinal de* X, noté $|X|$, comme étant l'unique cardinal équipotent à X, *si un tel existe*, et ceci est compatible avec nos définitions précédentes. Reste la question d'existence :

Lemme 1.4. *Soit X un ensemble. Alors X est équipotent à un (unique) cardinal si est seulement si X admet un bon ordre.*

DÉMONSTRATION. Exercice. ■1.4

Théorème 1.5. *Sont équivalents :*

(i) *Tout ensemble est équipotent à un (unique) cardinal.*

(ii) *Les cardinaux de tous deux ensembles sont comparables : $|X| \leq |Y|$ ou $|Y| \leq |X|$.*

(iii) *L'axiome du choix.*

DÉMONSTRATION. (i) $\Longrightarrow$ (ii). Soient $\kappa = |X|$ et $\lambda = |Y|$ leurs cardinaux. Puisque ce sont des ordinaux, il sont comparables.

(ii) $\Longrightarrow$ (iii). Soit $|X|$ un ensemble, et α son ordinal de Hartogs. Alors, par définition de α, il est impossible que $|\alpha| \leq |X|$. Donc $|X| \leq |\alpha|$, et l'injection de X dans α induit un bon ordre sur X. Ainsi tout ensemble admet un bon ordre, ce qui équivaut l'axiome du choix.

(iii) $\Longrightarrow$ (i). D'après le Lemme précédent et l'équivalence entre l'axiome du choix et le principe du bon ordre. ■1.5

Il est immédiat que si α est un ordinal alors $|\alpha| \leq \alpha$.

Exercice 4.1. Tous les ordinaux finis sont des cardinaux. L'ordinal ω est un cardinal.

Par contre, $\omega + 1$ n'est pas un cardinal, pas plus que $\omega + \omega$, $\omega.\omega$... Ces trois derniers ordinaux sont tous *dénombrables*, i.e équipotents à ω.

Il existe pour tout κ des ordinaux qui ne sont pas équipotents à une partie de κ (e.g., l'ordinal de Hartogs associé à κ), et donc des cardinaux λ tels que $\kappa < \lambda$. Notons que toute classe non vide de cardinaux a un plus petit élément (puisque c'est en particulier une classe non vide d'ordinaux).

Définition 1.6. Étant donné un cardinal κ, on note κ^+ le plus petit cardinal strictement plus grand que κ.

Exercice 4.2. Montrer que L'ordinal de Hartogs associé à un ensemble quelconque X est un cardinal. Montrer que pour un cardinal κ :

$$\kappa^+ = \text{l'ordinal de Hartogs associé à } \kappa = \{\alpha \in ON : |\alpha| \leq \kappa\}.$$

Définition 1.7. Si $\kappa > 0$ est un cardinal de la forme λ^+ pour un certain cardinal λ, on dit que κ est un *cardinal successeur* ; sinon, on dit que α est un *cardinal limite*.

Ici, attention à la terminologie: tous les cardinaux infinis sont des *ordinaux* limites ; par contre, ce ne sont pas tous des *cardinaux* limites. Notons que, si κ est un cardinal et s'il existe un plus grand cardinal $\lambda < \kappa$ alors on a $\kappa = \lambda^+$ et κ est donc un cardinal successeur ; par contre, si κ est limite, alors κ est égal à la réunion des cardinaux qui lui sont strictement inférieurs.

Lemme 1.8. *Soit A un ensemble de cardinaux. Alors $\sup A$ dans le sens des ordinaux (qui est égal à $\bigcup A$) est aussi un cardinal.*

DÉMONSTRATION. Soit $\alpha = \sup A$. Il suffit de montrer que si $\beta < \alpha$ alors $|\beta| < |\alpha|$. En effet, si $\beta < \alpha$ il existe $\kappa \in A$ tel que $\beta < \kappa$. Or $\kappa = |\kappa| \leq |\alpha|$, d'où $|\beta| \leq \beta < |\alpha|$. ■1.8

Définition 1.9. *(Alephs)*
On définit par récurrence transfinie $\aleph_\alpha$, pour tout ordinal α, en posant $\aleph_0 = \omega$ puis

$$\aleph_\alpha = \begin{cases} \aleph_\beta^+ & \text{si } \alpha = \beta + 1 \\ \sup_{\beta<\alpha} \aleph_\beta & \text{si } \alpha \text{ est limite} \end{cases}$$

On voit à partir de la définition que, si $\alpha < \beta$ sont deux ordinaux, alors $\aleph_\alpha < \aleph_\beta$. Avec le résultat précédent, on démontre par récurrence transfinie que $\aleph_\alpha$ est un cardinal pour tout α. Par récurrence transfinie, on peut également vérifier la propriété suivante.

Proposition 1.10. *Pour tout ordinal α, on a $\alpha \leq \aleph_\alpha$.*

Notons qu'il est possible que l'inégalité précédente soit une égalité: aussi contre-intuitif que cela puisse paraître, il existe des ordinaux α tels que $\alpha = \aleph_\alpha$ [i].

Proposition 1.11. *Tout cardinal infini est de la forme $\aleph_\alpha$ pour un unique ordinal α.*

DÉMONSTRATION. L'unicité est connue, montrons l'existence. Soit κ un cardinal infini : $\kappa \geq \aleph_0$. Puisque $\kappa \leq \aleph_\kappa$, il existe un plus petit α tel que $\kappa \leq \aleph_\alpha$. Si $\alpha = 0$ nous avons $\kappa \geq \aleph_0$; si $\alpha = \beta + 1$ alors $\aleph_\beta < \kappa$, d'où $\kappa \geq \aleph_\beta^+ = \aleph_\alpha$; et si α est limite $\aleph_\beta < \kappa$ pour tout $\beta < \alpha$, d'où $\aleph_\alpha = \sup_{\beta<\alpha} \aleph_\beta \leq \kappa$. Ainsi, $\kappa = \aleph_\alpha$. ■1.11

En conclusion, la classe des cardinaux consiste en les ensemble des cardinaux finis, plus les $\aleph_\alpha$,

$$\{\text{cardinaux}\} = \{\text{cardinaux finis}\} \cup \{\aleph_\alpha : \alpha \in ON\} = \omega \cup \{\aleph_\alpha : \alpha \in ON\}.$$

2. Arithmétique des cardinaux

Commençons par définir la somme et le produit de deux cardinaux. Avant cela, on a besoin d'un peu de terminologie: si X, Y sont deux ensembles alors on définit leur *union disjointe* $X \amalg Y$ par

$$X \amalg Y = \big(X \times \{0\}\big) \cup \big(Y \times \{1\}\big).$$

Définition 2.1. Soit κ et λ deux cardinaux. On définit :

$$|X| + |Y| = |X \amalg Y|, \qquad |X| \cdot |Y| = |X \times Y|.$$

Il faudra tout d'abord vérifier que ces opération son compatibles avec l'équipotence, et croissantes :

Proposition 2.2. *Supposons que $|X| \leq |X'|$ et $|Y| \leq |Y'|$. Alors $|X \amalg Y| \leq |X' \amalg Y'|$ et $|X \times Y| \leq |X' \times Y'|$. De la même façon, si $|X| = |X'|$ et $|Y| = |Y'|$ alors $|X \amalg Y| = |X' \amalg Y'|$ et $|X \times Y| = |X' \times Y'|$.*

Démonstration. Soit $f : X \to X'$ et $g : Y \to Y'$ deux injections (respectivement, bijections). Alors on peut définir une fonction $F : X \amalg Y \to X' \amalg Y'$ en posant

$$\begin{cases} F(x,0) &= (f(x),0) \\ F(y,1) &= (g(y),1) \end{cases}.$$

La vérification que F est bien une injection (bijection) est immédiate. De même, on peut définir une injection (bijection) $G : X \times Y \to X' \times Y'$ en posant $G(x,y) = (f(x), g(y))$. □

En particulier, si κ et λ sont des cardinaux, alors $\kappa + \lambda = |\kappa \amalg \lambda|$ et $\kappa \cdot \lambda = |\kappa \times \lambda|$, et ce sont également des cardinaux (pensez à la somme et au produit des ordinaux pour voir que $\kappa \amalg \lambda$ et $\kappa \times \lambda$ sont bien ordonnables, même sans l'axiome du choix). Par contre, ces opérations sont le plus souvent distinctes de la somme et le produit des ordinaux : par exemple, la somme ordinale $\omega + \omega$ est strictement plus grande que ω, or nous verrons que la somme cardinale $\aleph_0 + \aleph_0$ vaut $\aleph_0$. Notez d'ailleurs que nous écrivons ω quand nous pensons à l'ordinal et $\aleph_0$ quand nous pensons au cardinal, bien que $\omega = \aleph_0$.

Lemme 2.3. *Soient X, Y et Z des ensembles. L'addition et la multiplication des cardinalités sont commutatives et associatives, et la multiplication est distributive au-dessus de l'addition :*

$$|X| + |Y| = |Y| + |X|, \qquad (|X| + |Y|) + |Z| = |X| + (|Y| + |Z|),$$
$$|X| \cdot |Y| = |Y| \cdot |X|, \qquad (|X| \cdot |Y|) \cdot |Z| = |X| \cdot (|Y| \cdot |Z|),$$
$$(|X| + |Y|) \cdot |Z| = |X| \cdot |Z| + |Y| \cdot |Z|$$

i. C'est d'ailleurs un bon exercice ; pour le montrer, inspirez-vous de la preuve du fait qu'il existe un ordinal β tel que $\beta = \omega^\beta$

DÉMONSTRATION. Exercice. ■2.3

Restreignons-nous aux opérations arithmétiques sur les cardinaux, pour l'instant.

Lemme 2.4. *Soient α et β deux ordinaux. Malgré l'ambiguïté, notions $\alpha+\beta$ leur somme ordinale, et par $|\alpha|+|\beta|$ la somme cardinale de leur cardinaux, et de façon semblable pour le produit. Alors*

$$|\alpha+\beta| = |\alpha| + |\beta|, \qquad |\alpha\beta| = |\alpha||\beta|.$$

DÉMONSTRATION. D'après la représentation de $\alpha+\beta$ comme une concaténation d'ordre et de $\alpha\beta$ comme un ordre lexicographique. ■2.4

Puisque tout ordinal fini est un cardinal, et puisque la somme et le produit de deux ordinaux fini (pourquoi est-ce vrai ?), nous obtenons :

Lemme 2.5. *Sur les cardinaux finis, la somme et le produit ordinaux coïncident avec la somme et le produit cardinaux.*

L'addition et la multiplication des cardinaux sont déjà connues, donc. Si par contre au moins l'un de κ et λ est infini, l'addition et la multiplication deviennent assez inintéressantes.

Lemme 2.6. *Soit κ,λ deux cardinaux non nuls, dont au moins un est infini. Alors on a $\kappa+\lambda = \kappa\cdot\lambda = \max(\kappa,\lambda)$.*

DÉMONSTRATION. On définit d'abord une relation d'ordre sur $ON \times ON$:

$$\big((\alpha,\beta) \preceq (\alpha',\beta')\big) \iff \begin{cases} \max(\alpha,\beta) < \max(\alpha',\beta') & \text{ou} \\ \max(\alpha,\beta) = \max(\alpha',\beta') \text{ et } \alpha < \alpha' & \text{ou} \\ \max(\alpha,\beta) = \max(\alpha',\beta') \text{ et } \alpha = \alpha' \text{ et } \beta \leq \beta' \end{cases}$$

Il est facile de vérifier que $\preceq$ est une relation d'ordre total sur $ON \times ON$; pour voir que c'est un bon ordre, soit A un ensemble non vide contenu dans $ON \times ON$. Posons :

$$\begin{aligned} \gamma &= \min\big\{\max(\alpha,\beta) \colon (\alpha,\beta) \in A\big\}, \\ \alpha_0 &= \min\big\{\alpha \in ON \colon \exists\beta \text{ t.q. } \gamma = \max(\alpha,\beta) \text{ et } (\alpha,\beta) \in A\big\}, \\ \beta_0 &= \min\big\{\beta \in ON \colon \gamma = \max(\alpha_0,\beta) \text{ et } (\alpha_0,\beta) \in A\big\}. \end{aligned}$$

Il est facile de vérifier que $(\alpha_0,\beta_0) = \min A$.

Maintenant, montrons que $\kappa\cdot\lambda = \max(\kappa,\lambda)$ par récurrence transfinie sur $\max(\kappa,\lambda)$. Nous pouvons supposer que $\kappa \geq \lambda$, et en particulier que κ est infini. Soit $S \subseteq \kappa\times\kappa$ un segment initial propre. Nous avons alors $S = S_{(\alpha,\beta)}$ pour une certain paire $(\alpha,\beta) \in \kappa\times\kappa$. Par définition, $S \subseteq (\alpha+1)\times(\beta+1)$, et $\alpha,\beta < \kappa$, d'où $\alpha+1, \beta+1 < \kappa$ (car tout cardinal infini est un ordinal limite). Si α et β sont finis alors $|S| \leq (\alpha+1)(\beta+1) < \omega \leq \kappa$. Sinon, soit $\mu = \max\big(|\alpha+1|, |\beta+1|\big)$. Alors $\mu \leq \max(\alpha+1,\beta+1) < \kappa$, et d'après l'hypothèse de récurrence $|\alpha+1|\cdot|\beta+1| = \mu$, et encore $|S| \leq \mu < \kappa$. Ainsi, le cardinal d'un segment initial propre de $\kappa\times\kappa$ est strictement plus petit que κ.

Or, puisque tout deux ordinaux sont comparable, ou bien κ est isomorphe à un segment initial propre de $\kappa\times\kappa$, ou bien $\kappa\times\kappa$ est isomorphe à un segment initial (non nécessairement propre) de κ. Puisqu'on a exclu la première possibilité, c'est nécessairement la deuxième, d'où $\kappa\cdot\kappa \leq \kappa$. Or $\kappa = 1\cdot\kappa \leq \kappa\cdot\kappa$, d'où l'égalité.

Maintenant nous pouvons conclure par :

$$\begin{aligned} \kappa = \kappa\cdot 1 \leq \kappa\cdot\lambda \leq \kappa\cdot\kappa = \kappa, \\ \kappa = \kappa + 0 \leq \kappa + \lambda \leq \kappa + \kappa = \kappa\cdot 2 = \max(\kappa, 2) = \kappa. \end{aligned}$$

■2.6

Remarque 2.7. Il n'est pas vrai, sans l'axiome du choix, que $|X|\cdot|X| = |X|$ pour tout ensemble, même non bien ordonnable. En effet, on peut démontrer que si $|X|\cdot|X| = |X|$ pour tout X alors l'axiome du choix est vrai.

Au final, la somme et le produit des cardinaux ne sont pas bien intéressants... Définissons quelques opérations un peu plus complexes.

Nous définissons les opérations correspondantes pour les cardinaux. À partir de maintenant il convient de supposer que l'axiome du choix est vrai – sans cela, même si X et Y sont bien ordonnables, X^Y pourrait ne pas l'être, et bien que ce ne soit par trop gênant, nous préférons simplifier un peu la situation.

Définition 2.8. Pour des cardinaux κ, λ nous définissons κ^λ comme le cardinal de l'ensemble des fonctions de λ dans κ (i.e., avec notre notation ambiguë, $\kappa^\lambda = |\kappa^\lambda|...$).

Exercice 4.3. Si X_0, X_1 (resp. Y_0, Y_1) sont des ensembles équipotents, alors $X_0^{Y_0}$ et $X_1^{Y_1}$ sont équipotents.

Notons que, en associant à une partie d'un ensemble X sa fonction caractéristique, on obtient une bijection de $\mathcal{P}(X)$ sur 2^X. En particulier, le cardinal de l'ensemble des parties d'un cardinal κ est égal à 2^κ.

On retrouve sans difficulté les propriétés usuelles de l'exponentiation.

Exercice 4.4. Montrer que pour trois cardinaux κ, λ, μ on a $(\kappa^\lambda)^\mu = \kappa^{\lambda.\mu}$, et $\kappa^\lambda.\kappa^\mu = \kappa^{\lambda+\mu}$.

Exercice 4.5. Montrer que pour tout $0 < n < \omega$ et κ infini : $\kappa^n = \kappa$.

Souvent, on peut utiliser le théorème de Schröder-Bernstein, en conjonction avec les propriétés de l'arithmétique cardinale, pour montrer des égalités entre cardinaux. L'exercice suivant fournit un exemple d'une telle situation.

Exercice 4.6. Montrer que pour tout cardinal infini κ on a $2^\kappa = \kappa^\kappa$.

D'une certaine façon, l'opération arithmétique la plus mystérieuse/la plus intéressante sur les cardinaux est l'exponentiation.

Théorème 2.9. (Cantor) *Pour tout ensemble X il n'existe pas de surjection $f\colon X \to \mathcal{P}(X)$. Avec nos notations cela signifie que pour tout cardinal κ on a $\kappa < 2^\kappa$.*

Démonstration. Par l'absurde, soit $f\colon X \to \mathcal{P}(X)$ une surjection, et soit

$$Y = \{x \in X : x \notin f(x)\} .$$

On doit avoir $Y = f(x_0)$ pour un certain $x_0 \in X$, mais alors on vérifie que $(x_0 \in Y) \Leftrightarrow (x_0 \notin Y)$, et on arrive donc à une contradiction. □.

On sait donc produire une classe strictement croissante et non bornée de cardinaux, en répétant l'opération $\kappa \mapsto 2^\kappa$ et en prenant le sup aux ordinaux limite. Y a-t-il des cardinaux qui n'apparaissent pas dans cette énumération ?

Définition 2.10. *L'hypothèse du continu* (HC) est l'énoncé $2^{\aleph_0} = \aleph_1$.
L'hypothèse du continu *généralisée* est l'énoncé affirmant que pour tout ordinal α on a $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

L'idée sous-jacente de l'hypothèse du continu est qu'on peut « voir » $\mathbf{N}$, de cardinal $\aleph_0$, et $\mathbf{R}$, de cardinal $2^{\aleph_0}$, mais on ne voit pas d'ensemble de réels qui soit de cardinal intermédiaire. La question est donc: en existe-t-il ?
Pendant longtemps cette hypothèse a paru naturelle ; Gödel a prouvé qu'elle était consistante avec les axiomes de ZFC. Mais dans les années 60, Paul Cohen a montré, en utilisant la méthode du *forcing*, que la négation de l'hypothèse du continu était *aussi* consistante avec ZFC, autrement dit (HC) est indépendante de ZFC.
Aujourd'hui, la plupart des théoriciens des ensembles considèrent qu'il n'y a aucune raison de limiter la richesse de la théorie en imposant arbitrairement que l'hypothèse du continu soit vérifiée ; il existe des axiomes (« grands cardinaux») menant à une théorie très riche dans laquelle l'hypothèse du continu est fausse.

On peut aussi vouloir faire une somme/produit d'une infinité de cardinaux ; dans ce cas, le recours à l'axiome du choix s'avère indispensable ; on laisse la vérification de la propriété suivante en exercice.

Définition 2.11. Soit $\{X_i\}_{i\in I}$ une famille d'ensembles, indexée par un ensemble I. Nous définissons sa *réunion disjointe* (ou *co-produit*) comme suit, et rappelons à l'occasion la définition d'un produit cartésien :

$$\coprod_{i\in I} X_i = \bigcup_{i\in I} X_i \times \{i\},$$

$$\prod_{i\in I} X_i = \{\text{fonctions } f\colon I \to \bigcup X_i \text{t.q.} f(i) \in X_i\}.$$

Ainsi $X \times Y = \coprod_{y\in Y} X$ et $X^Y = \prod_{y\in Y} X$.

Définition 2.12. Soit $\alpha \in ON$ et pour $\beta < \alpha$ soit κ_β un cardinal. Alors

$$\sum_{\beta<\alpha} \kappa_\beta = \left|\coprod_{\beta<\alpha} \kappa_\beta\right|, \qquad \prod_{\beta<\alpha} \kappa_\beta = \left|\prod_{\beta<\alpha} \kappa_\beta\right|,$$

où, comme pour l'exponentiation, par $\left|\prod_{\beta<\alpha} \kappa_\beta\right|$ nous entendons le cardinal du produit cartésien infini.

Exercice 4.7. A l'aide de l'axiome du choix, vérifier que si $(X_\alpha)_{\alpha<\lambda}$ et $(Y_\alpha)_{\alpha<\lambda}$ sont tels que $|X_\alpha| = |Y_\alpha|$ pour tout $\alpha < \lambda$ alors on a

$$\left|\coprod_{\alpha<\lambda} X_\alpha\right| = \left|\coprod_{\alpha<\lambda} Y_\alpha\right| \text{ et } \left|\prod_{\alpha<\lambda} X_\alpha\right| = \left|\prod_{\alpha<\lambda} Y_\alpha\right| .$$

Exercice 4.8. Soit α un ordinal et X_β un ensemble pour tout $\beta \leq \alpha$. Montrer que

$$\left|\coprod_{\beta\leq\alpha} X_\beta\right| = \left|\left(\coprod_{\beta<\alpha} X_\beta\right) \amalg X_\alpha\right|, \qquad \left|\prod_{\beta\leq\alpha} X_\beta\right| = \left|\left(\prod_{\beta<\alpha} X_\beta\right) \times X_\alpha\right|.$$

3. Dénombrabilité

Dans cette section, on va détailler un peu une notion fondamentale en mathématiques: la dénombrabilité. On a pu croire un temps que cette notion n'était omniprésente qu'à cause de limites techniques, mais elle semble toujours aussi importante aujourd'hui malgrè l'avancée des mathématiques [ii].

Définition 3.1. Un ensemble X est *dénombrable* si $|X| \leq \aleph_0$.

(Selon certains usages on réserve ce terme aux ensembles infinis, et exige donc que $|X| = \aleph_0$.)

Lemme 3.2. *Soit X dénombrable et $f\colon X \to Y$ une application surjective. Alors Y est dénombrable.*

Démonstration. On peut supposer que $X \subseteq \omega$. Dans ce cas nous définissons une application injective $g\colon Y \to \omega$ par $g(y) = \min f^{-1}(\{y\})$, d'où $|Y| \leq \aleph_0$. ■3.2

Lemme 3.3. *Soit X un ensemble dénombrable. Alors l'ensemble des parties finies de X, que l'on notera $\mathcal{P}^{fin}(X)$ est dénombrable.*

Démonstration. Il suffirait de montrer que l'ensemble des parties finies de ω est dénombrable. En effet, pour $k < \omega$, notons par $\mathcal{P}^k(\omega)$ l'ensemble des partie de ω de cardinal k. Chaque membre de $\mathcal{P}^k(\omega)$ admet une énumération croissante de longueur k, ce qui donne une injection de $\mathcal{P}^k(\omega)$ dans ω^k. Ainsi :

$$|\mathcal{P}^{\text{fin}}(\omega)| = \sum_{k<\omega} |\mathcal{P}^k(\omega)| \leq \sum_{k<\omega} |\omega^k| = \sum_{k<\omega} \aleph_0 = \aleph_0 \cdot |\omega| = \aleph_0 \cdot \aleph_0 = \aleph_0.$$

■3.3

Remarque 3.4. Avec ce qu'on a vu, l'égalité $\sum_{k<\omega} |\omega^k| = \sum_{k<\omega} \aleph_0$ exige l'axiome du choix (car, pour chaque k, il faudrait choisir une bijection encore ω^k et ω). Or, nous avons déjà vu qu'il existe une bijection canonique entre κ^2 et κ pour chaque cardinal κ, et donc en particulier pour $\kappa = \aleph_0 = \omega$. On en construit (par récurrence sur k) une bijection canonique entre ω^k et ω pour chaque $0 < k < \omega$, et l'axiome du choix n'est plus nécessaire. Modulo un peu d'arithmétique dans $\mathbf{N}$, nous pouvons donner une preuve encore plus directe : l'application qui envoie $Y \in \mathcal{P}^{\text{fin}}(\omega)$ à $\sum_{n\in Y} 2^n$ est une bijection entre $Y \in \mathcal{P}^{\text{fin}}(\omega)$ et ω.

L'importance de la dénombrabilité vient, au moins en partie, du fait que beaucoup des notions que l'on considère en mathématiques s'expriment à partir d'énoncés finis dans un langage fini ou dénombrable, ce qui entraîne que beaucoup de structures « engendrées » par des ensembles dénombrables restent dénombrables. Ce phénomène sera particulièrement utile dans la partie du cours consacrée à la théorie des modèles. Donnons simplement un exemple, pour manipuler un peu cette notion de dénombrabilité.

Exemple. Soit G un groupe, et $A \subseteq G$ une partie dénombrable. Alors le sous-groupe de G engendré par A est dénombrable.

ii. Il est peut-être pertinent de rappeler ici la fameuse citation de Weyl ([**Dug03**]), faisant entre autres allusion à la définition des filtres censés éliminer l'usage des suites: « Avec le recul que donnent les quarante dernières années, on sourira sans doute du zèle que j'apportais à l'expulsion du dénombrable: chassé par la porte, il a fini par rentrer par la fenêtre» .

Démonstration. Supposons que A est dénombrable ; alors $A^{-1} = \{a^{-1} : a \in A\}$ est aussi dénombrable (il est équipotent à A) et donc $A \cup A^{-1}$ aussi. Par suite, quitte à remplacer A par $A \cup A^{-1}$, on peut donc supposer, pour se simplifier la vie, que A est symétrique (i.e stable par l'application inverse). Alors, le groupe engendré par A est égal à l'ensemble

$$\{a_1 \dots a_k : k < \omega \text{ et } a_1, \dots, a_k \in A\}$$

Notons que, par convention, le produit vide, obtenu quand $k = 0$, est égal à l'élément neutre de G.
En particulier, le groupe engendré par A est égal à la réunion, pour $k < \omega$, des ensembles

$$A_k = \{a_1 \dots a_k : a_1, \dots, a_k \in A\}$$

Chacun des ensembles A_k est l'image de A^k par la fonction qui associe $a_1 \dots a_k$ à $(a_1, \dots, a_k)$, par conséquent chaque A_k est dénombrable et donc le sous-groupe engendré par A est lui aussi dénombrable. □

Les ensembles dénombrables sont stables par d'autres types d'opérations, par exemple celles liées à l'arithmétique des ordinaux.

Exercice 4.9. Montrer que, si α et β sont des ordinaux dénombrables, alors $\alpha+\beta$, $\alpha.\beta$ et α^β sont encore des ordinaux dénombrables. Montrer que $\omega^{\omega_1} = \omega_1$, où ω_1 est le plus petit ordinal non dénombrable, et qu'il existe pour tout ordinal *dénombrable* α un ordinal *dénombrable* $\beta \geq \alpha$ tel que $\omega^\beta = \beta$.

Remarquons que, quand on y pense À ω_1 comme à un cardinal, on lui a donné un autre nom: $\aleph_1$. Il n'est pas trop difficile de prouver que $\mathbf{Q}$ est dénombrable. Ceci nous permet de calculer le cardinal de $\mathbf{R}$, comme le montre l'exercice suivant.

Exercice 4.10. Montrer que $|\mathbf{Q}| = \aleph_0$, puis montrer que $|\mathbf{R}| = 2^{\aleph_0}$. Pour le second point, on pourra considérer l'application $f : \mathbf{R} \to \mathcal{P}(\mathbf{Q})$ définie par

$$f(x) = \{q \in \mathbf{Q} : q < x\} \ .$$

Finissons cette section par une petite question d'apparence innocente: si on oublie l'axiome du choix, tous les ensembles infinis contiennent-ils un sous-ensemble dénombrable ? Mais, au fait, qu'est-ce qu'un ensemble infini ?

Définition 3.5. Un ensemble est *infini* s'il n'est équipotent à aucun ordinal fini. Un ensemble X est *Dedekind-infini* s'il existe une injection non surjective $f : X \to X$.

Une autre définition possible d'un ensemble infini serait: un ensemble qui contient un sous-ensemble équipotent à ω. On a en fait déjà introduit cette définition, comme le montre l'exercice suivant.

Exercice 4.11. Montrer qu'un ensemble est Dedekind-infini si, et seulement si, il contient un sous-ensemble dénombrable.

Exercice 4.12. En utilisant l'axiome du choix dénombrable, montrer que tout ensemble infini est Dedekind-infini.

L'équivalence entre ces deux notions (ensemble infini/ensemble Dedekind-infini) est en fait indépendante de ZF! On peut prendre cela comme une confirmation du fait que l'axiome du choix dénombrable est relativement naturel.

Notes bibliographiques.
En ce qui concerne l'axiome du choix, il existe une véritable encyclopédie [**HR98**] présentant ses multiples formes ; on pourra y trouver des références sur certains résultats énoncés sans référence dans le corps du chapitre ci-dessus. On pourra aussi consulter [**Jec73**], et le livre de S. Wagon [**Wag85**] est également très instructif.

4. Cardinaux réguliers et cofinalité

Avant de conclure ce chapitre sur les cardinaux, nous allons évoquer une notion importante dans l'étude des propriétés des cardinaux ; la *régularité*.

Définition 4.1. Soit α un ordinal. Nous définissions sa *cofinalité*, $\mathrm{cf}(\alpha)$ comme étant le plus petit cardinal d'une partie non majorée de α. Autrement dit,

$$\mathcal{F}(\alpha) = \min\{|A| : A \subseteq \alpha \text{ et } \sup A = \alpha\}.$$

Définition 4.2. Un cardinal infini κ est dit *régulier* si $\mathrm{cf}(\kappa) = \kappa$, i.e., si pour toute partie $X \subseteq \kappa$ de cardinal strictement inférieur à κ on a $\sup X < \kappa$. Un cardinal qui n'est pas régulier est dit *singulier*.

Ainsi, $\aleph_0$ est régulier, alors que $\aleph_\omega$ est singulier.

Dans les deux exercices suivants, que vous traiterez en TD, on étudie quelques propriétés de cette notion.

Exercice 4.13. (i) Montrer que $\mathrm{cf}(\alpha)$ est le plus petit ordinal γ tel qu'il existe une fonction $f : \gamma \to \alpha$ dont l'image ne soit pas strictement majorée.

(ii) Montrer que, pour tout ordinal α, $\mathrm{cf}(\alpha)$ est un cardinal régulier.

(iii) Montrer que si α est limite alors $\mathrm{cf}(\alpha) = \mathrm{cf}(\aleph_\alpha)$.

Exercice 4.14. (i) Montrer qu'un cardinal κ est régulier si, et seulement si, pour tout $\lambda < \kappa$ et toute famille $(X_\alpha)_{\alpha\in\lambda}$ d'ensembles tels que $|X_\alpha| < \kappa$ pour tout $\alpha < \lambda$, on a $|\bigcup X_\alpha| < \kappa$.

(ii) Soit κ un cardinal ; montrer que $\mathrm{cf}(\kappa)$ est le plus petit cardinal λ tel que κ soit la réunion de λ ensembles de cardinal strictement inférieur à κ.

(iii) On appelle *inaccessible* un cardinal non dénombrable à la fois limite et régulier. Montrer qu'un tel cardinal α doit vérifier $\alpha = \aleph_\alpha$. La réciproque est-elle vraie[iii] ?

Proposition 4.3. *Tout cardinal successeur (infini) est régulier.*

Démonstration. Fixons un cardinal successeur κ, et λ tel que $\kappa = \lambda^+$. Soit alors $\mu = \mathrm{cf}(\kappa)$; c'est un cardinal tel qu'il existe des ensembles (X_ξ) de cardinal strictement inférieur à κ (donc inférieur ou égal à λ) et tels que

$$\kappa = \bigcup_{\xi<\mu} X_\xi \ .$$

On en déduit que

$$\kappa = \left|\bigcup_{\xi<\mu} X_\xi\right| \leq \sum_{\xi<\mu} |X_\xi| \leq \sum_{\xi<\mu} \lambda = \mu.\lambda = \max(\mu, \lambda)$$

Puisque $\kappa > \lambda$, on obtient finalement $\kappa \leq \mu$, ce qu'il fallait démontrer. □

On peut, après la proposition ci-dessus, avoir l'impression que tout cardinal est régulier: mais les cardinaux limites sont bien souvent singuliers. C'est par exemple le cas de $\aleph_\omega$, qui est une union dénombrable d'ensembles de cardinal strictement plus petit que lui. On est amené à se poser la question suivante: existe-t-il un cardinal différent de $\aleph_0$ qui soit à la fois régulier et limite ? Un tel cardinal est dit *inaccessible*. Une version plus forte de la même définition est souvent utilisée : on dit qu'un cardinal $\kappa > \aleph_0$ est *fortement limite* si $\lambda < \kappa$ implique $2^\lambda < \kappa$ (et non seulement $\lambda^+ < \kappa$), et qu'il est *fortement inaccessible* s'il est à la fois fortement limite et régulier. Tout cardinal fortement limite est limite, et tout cardinal fortement inaccessible est inaccessible.

Il se trouve que l'existence d'un cardinal inaccessible n'est pas démontrable dans ZFC : en fait, à partir de ZFC + "il existe un cardinal inaccessible" on peut démontrer que ZFC est cohérent, ce que, d'après le théorème de Gödel, on ne peut pas démontrer du système ZFC seul ! Il est plus facile de voir ceci pour les cardinaux inaccessibles.

Exercice 4.15. Définissons pour chaque ordinal α :

$$V_\alpha = \begin{cases} \varnothing & \alpha = 0 \\ \mathcal{P}(V_\beta) & \alpha = \beta + 1 \\ \bigcup_{\beta<\alpha} V_\beta & \alpha \text{ limite.} \end{cases}$$

Montrer que si λ est fortement inaccessible alors V_λ (qui est un ensemble), muni de la relation usuelle $\in$, est un modèle de ZFC.

Nous sommes maintenant équipés pour prouver le dernier théorème de ce chapitre, le *lemme de König*. Avant de l'énoncer, notons qu'il est assez difficile d'obtenir des inégalités strictes dans l'arithmétique des cardinaux. A titre d'exemple, bien que $2 < \aleph_0$ et que l'on s'attendrait à ce qu'une somme soit plus petite qu'un produit, on a :

iii. Pour traiter cette question, on pourra d'abord lire la discussion ci-dessous concernant l'existence de cardinaux inaccessibles.

$$\sum_{i<\omega} 2 = \aleph_0 = \sum_{i<\omega} \aleph_0,$$
$$\prod_{i<\omega} 2 = 2^{\aleph_0} = \aleph_0^{\aleph_0} = \prod_{i<\omega} \aleph_0,$$
$$\sum_{i<\omega} 2^{\aleph_0} = 2^{\aleph_0} = \prod_{i<\omega} 2^{\aleph_0}.$$

Pour obtenir une inégalité stricte qui tient en toute généralité, on est ramené au résultat suivant. L'unique autre inégalité stricte que l'on a déjà vue, $2^\kappa > \kappa$, en est d'ailleurs un cas particulier (comment ?)

Théorème 4.4. *Soient $(\kappa_i)_{i\in I}$ et $(\lambda_i)_{i\in I}$ deux familles de cardinaux tels que pour tout i on ait $\kappa_i < \lambda_i$. Alors on a*

$$\sum_{i\in I} \kappa_i < \prod_{i\in I} \lambda_i \ .$$

DÉMONSTRATION. Montrons d'abord l'inégalité large. En effet, pour chaque $i \in I$ et $\alpha < \kappa_i$, posons

$$g_{i,\alpha} \colon I \to ON, \qquad g(j) = \begin{cases} 0 & i \neq j, \\ \alpha + 1 & i = j. \end{cases}$$

Alors $(\alpha, i) \mapsto g_{i,\alpha}$ est bien une injection $\coprod \kappa_i \to \prod \lambda_i$.

Pour prouver que l'inégalité est stricte, on raisonne par l'absurde et on suppose que $\sum \kappa_i = \prod \lambda_i$, c'est à dire qu'il existe une bijection $f \colon \coprod \kappa_i \to \prod \lambda_i$. Pour chaque $j \in I$, soit $\pi_j \colon \prod_i \lambda_i \to \lambda_j$ la projection sur la $j$ème coordonnée. Considérons l'application $h_j \colon \kappa_j \to \lambda_j$ donnée par $h_j(\alpha) = \pi_j \circ f(\alpha, j)$. Puisque $\kappa_i < \lambda_i$, elle n'est pas surjective, il existe donc un plus petit $\beta_j < \lambda_j$ qui n'est pas dans l'image. Nous avons $x = (\beta_i)_{i\in I} \in \prod_i \lambda_i$, est il doit exister donc $(\alpha, j) \in \coprod_i \kappa_i$ (c'est à dire, $j \in I$ et $\alpha < \kappa_j$) tel que $f(\alpha, j) = x$. Or, dans ce cas :

$$\beta_j \neq h_j(\alpha) = \pi_j \circ f(\alpha, j) = \pi_j\big((\beta_i)_i\big) = \beta_j.$$

Cette contradiction montre qu'une telle bijection f ne peut exister, et complète la preuve. ■$_{4.4}$

Le raisonnement ci-dessus est un bon exemple de *raisonnement diagonal* : pour chaque $j \in I$ on a choisit β_j tel que, à la fin, $f(\alpha, j)$ ne pourra être égal à x pour aucun $\alpha < \kappa_j$.

Exercice 4.16. En utilisant le fait que tout réel admet un développement décimal [iv], prouver à l'aide d'un raisonnement diagonal que $\mathbf{R}$ n'est pas dénombrable.

Corollaire 4.5. *Pour tout cardinal infini κ, on a $\kappa < \kappa^{\mathrm{cf}(\kappa)}$.*

Démonstration. Fixons $\kappa_i < \kappa$, $i < \mathrm{cf}(\kappa)$, tels que $\kappa = \sum_{i<\mathrm{cf}(\kappa)} \kappa_i$.
Ces κ_i existent par définition de $\mathrm{cf}(\kappa)$. Alors on a, d'après le lemme de König:

$$\kappa = \sum_{i<\mathrm{cf}(\kappa)} \kappa_i < \prod_{i<\mathrm{cf}(\kappa)} \kappa = \kappa^{\mathrm{cf}(\kappa)} \ .$$

Ceci conclut la preuve. □

On en déduit immédiatement un autre corollaire.

Corollaire 4.6. *Pour tout cardinal infini κ, on a $\mathrm{cf}(2^\kappa) > \kappa$, et en particulier $2^\kappa > \kappa$.*

Démonstration. Appliquons le corollaire précédent à 2^κ: on obtient

$$2^\kappa < (2^\kappa)^{\mathrm{cf}(2^\kappa)} = 2^{\kappa.\,\mathrm{cf}(2^\kappa)} \ .$$

Ceci n'est possible que si $\kappa < \kappa.\,\mathrm{cf}(2^\kappa) = \max(\kappa, \mathrm{cf}(2^\kappa))$. □

Ceci a pour conséquence une restriction à la négation de l'hypothèse du continu : si $2^{\aleph_0} = \kappa$ alors il faut que $\mathrm{cf}(\kappa) > \aleph_0$. Il s'agit en fait essentiellement de la seule obstruction que l'on puisse démontrer dans ZFC, mais démontrer ce fait est largement hors de notre portée dans ce cours.

iv. Attention tout de même: parfois il en existe deux!

Notes bibliographiques.
Encore une fois, ce chapitre reprend pour l'essentiel, avec plus de détails, les notes du cours de M2 « théorie descriptive des groupes ». Le lecteur intéressé est de nouveau invité à consulter [**Hal74**] s'il cherche une présentation intuitive de la théorie, et [**Mos06**] ou [**KM76**] pour une présentation plus formelle. Le lecteur anglophobe souhaitant se documenter sur les cardinaux pourra consulter avec profit la traduction française du livre de Kuratowski sus-cité ou le livre de Jean-Louis Krivine [**Kri98**].

Enfin, comme source bibliographique et comme référence concernant les résultats plus récents de théorie des ensembles (forcing, etc.), le lecteur est invité à consulter [**Jec03**].

CHAPITRE 5

Filtres et ultrafiltres

Dans ce chapitre, on va présenter quelques résultats élémentaires concernant les filtres et ultrafiltres, qui sont des objets centraux de la théorie des ensembles modernes et sont aussi utilisés aujourd'hui dans diverses branches des mathématiques: théorie des modèles bien sûr, mais aussi topologie, algèbres de von Neumann, systèmes dynamiques, géométrie des espaces de Banach...

1. Définitions, premières propriétés

Définition 1.1. Soit X un ensemble. Un *filtre* sur X est une famille $\mathcal{F} \subset \mathcal{P}(X)$ vérifiant les propriétés suivantes:

(i) $A \in \mathcal{F}$ et $B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$;

(ii) $A \in \mathcal{F}$ et $B \supseteq A \Rightarrow B \in \mathcal{F}$;

(iii) $\varnothing \notin \mathcal{F}$.

Exemple.
- L'exemple le plus simple de filtre sur X est $\{X\}$; l'exemple suivant est à peine moins inintéressant: pour tout $x \in X$, l'ensemble $\mathcal{F}_x = \{A : x \in A\}$ est un filtre. En fait, pour toute partie non vide $S \subseteq X$, l'ensemble des parties qui contiennent S est un filtre; on appelle un tel filtre un *filtre principal*. Quand X est fini, tout filtre est principal (pourquoi ?)
- Quand X est infini on a un exemple plus intéressant: l'ensemble

$$\mathcal{F} = \{A \subseteq X : \text{ le complémentaire de } A \text{ est fini}\}$$

est un filtre sur X , appelé *filtre de Fréchet* sur X. C'est un filtre non principal.

L'exercice suivant explique pourquoi on n'est pas vraiment intéressé par les filtres sur des ensembles finis.

Exercice 5.1. Soit $\mathcal{F}$ un filtre contenant une partie *finie* A. Alors $\mathcal{F}$ est principal.

Définition 1.2. On dit qu'une famille $\mathcal{A} \subseteq \mathcal{P}(X)$ est une *base de filtre* si toutes les intersections finies d'éléments de $\mathcal{A}$ sont non vides.

Proposition 1.3. *Pour toute base de filtre $\mathcal{A}$, il existe un filtre contenant $\mathcal{A}$; le plus petit tel filtre est défini par*

$$\mathcal{F} = \left\{ B \subseteq X : \exists A_1, \ldots, A_n \in \mathcal{A} \text{ t.q. } B \supseteq \bigcap_{i=1}^{n} A_i \right\}.$$

Démonstration. Soit $\mathcal{A}$ une base de filtre; il est clair que l'ensemble $\mathcal{F}$ défini ci-dessus contient $\mathcal{A}$ et que tout filtre contenant $\mathcal{A}$ doit contenir $\mathcal{F}$, donc il nous suffit de prouver que $\mathcal{F}$ est bien un filtre.
On voit tout de suite que $\mathcal{F}$ satisfait les points 1 et 2 de la définition d'un filtre; d'autre part, comme toute intersection finie d'éléments de $\mathcal{A}$ est non vide, on voit que $\varnothing \notin \mathcal{F}$ et donc $\mathcal{F}$ est bien un filtre. □

L'exemple suivant est très important en théorie des modèles, en particulier si l'on souhaite démontrer le théorème de compacité en utilisant des filtres.

Exemple. Soit I un ensemble infini, et X l'ensemble des parties finies de I [i]. Alors la famille des parties $\mathcal{B}$ de la forme $\{A \in X : A \supseteq F\}$, où F est une partie finie non vide de X, est une base de filtre sur X. En effet, une intersection finie d'élements de $\mathcal{B}$ est de la forme

$$\{A \in X : A \supseteq F_1 \text{ et } \ldots \text{ et } A \supseteq F_n\},$$

i. Savez-vous calculer le cardinal de X en fonction de celui de I ?

où les F_i sont des parties finies de X. Cet ensemble peut aussi s'écrire

$$\left\{ A \in X \colon A \supseteq \bigcup_{i=1}^{n} F_i \right\},$$

et ce dernier ensemble est non vide puisque la réunion des F_i est un ensemble fini.

Notons que le filtre engendré par $\mathcal{B}$ est non principal (parce que I est infini!).

Définition 1.4. Un filtre maximal s'appelle un *ultrafiltre*.

On vérifie facilement que l'ensemble des filtres contenant un filtre donné, ordonné par l'inclusion, est un ensemble ordonné inductif. Par conséquent, le lemme de Zorn garantit qu'il existe des ultrafiltres contenant tout filtre donné ; cet axiome, appelé *axiome de l'ultrafiltre*, est une forme faible d'axiome du choix.

Notons en tous les cas que, en présence de l'axiome de l'ultrafiltre, il existe des ultrafiltres non principaux sur tout ensemble infini X, puisqu'il existe des ultrafiltres contenant le filtre de Fréchet sur X.

Proposition 1.5. *Un filtre $\mathcal{F}$ est un ultrafiltre si, et seulement si, pour tout $A \in X$ on a $A \in \mathcal{F}$ ou $X \setminus A \in \mathcal{F}$.*

Démonstration. Supposons que $\mathcal{F}$ soit un filtre et qu'il existe $A \subseteq X$ tel que ni A ni $X \setminus A$ n'appartiennent à $\mathcal{F}$. Alors on va montrer que $\mathcal{G} = \mathcal{F} \cup \{A\}$ est une base de filtre, ce qui garantira l'existence d'un filtre contenant strictement $\mathcal{F}$ et montrera donc que $\mathcal{F}$ n'est pas un ultrafiltre.
Soit donc $B_1, \ldots, B_n \in \mathcal{F}$; on doit montrer que $B_1 \cap \ldots \cap B_n \cap A$ ne peut être vide. Raisonnons par l'absurde: si cette intersection est vide, alors $B_1 \cap \ldots \cap B_n \subseteq X \setminus A$, ce qui montre que $X \setminus A \in \mathcal{F}$ et cela contredit notre hypothèse. Donc $\mathcal{G}$ est bien une base de filtre et $\mathcal{F}$ n'est pas un ultrafiltre.
Réciproquement, supposons que $\mathcal{F}$ soit un filtre qui ne soit pas un ultrafiltre. Alors il existe un filtre $\mathcal{G}$ contenant strictement $\mathcal{F}$; considérons $A \in \mathcal{G} \setminus \mathcal{F}$. On ne peut avoir $X \setminus A \in \mathcal{G}$ puisque $\mathcal{G}$ est un filtre, a fortiori il est impossible que $X \setminus A \in \mathcal{F}$ et donc ni A ni $X \setminus A$ n'appartiennent à $\mathcal{F}$. □

Exercice 5.2. Soit $\mathcal{U}$ un ultrafiltre sur X. Montrer que soit $\mathcal{U}$ contient le filtre de Fréchet sur X, soit $\mathcal{U}$ est principal.

En topologie, les ultrafiltres peuvent être utilisés pour généraliser la notion de convergence de suite ; le lecteur intéressé est invité à consulter les feuilles de TD des années précédentes pour des exercices sur la question.

Avec l'axiome du choix, on sait qu'il existe des ultrafiltres non principaux sur tout ensemble infini ; mais certaines propriétés combinatoires de ces ultrafiltres sont elles-mêmes indépendantes de ZFC! Discutons un exemple, important pour les théoriciens des ensembles contemporains, avant de passer à la théorie des modèles. Cet exemple nous sert surtout de prétexte à manipuler un peu des ordinaux, des cardinaux, et des filtres, et donner l'idée que la théorie des ensembles modernes est en grande partie une forme de combinatoire infinie.

2. Utilisation des filtres en topologie

On va expliquer pourquoi les filtres et ultrafiltres peuvent être utiles en topologie ; la justification de l'introduction des filtres dans ce contexte est que dans certains espaces les points n'ont pas de base dénombrable de voisinages, et alors on ne peut plus se contenter d'utiliser des suites pour caractériser les notions habituelles de topologie (fonctions continues, ensembles fermés, etc.). Pourtant il est agréable de raisonner séquentiellement ; on peut alors utiliser des *suites généralisées*, comme le font généralement les anglo-saxons, ou bien des filtres. Voyons comment fonctionne cette deuxième approche.

Commençons par remarquer que, si X est un espace topologique et $x \in X$ alors la famille des voisinages de x, notée $\mathcal{V}_x$, forme un filtre. Ce filtre est l'analogue dans le contexte des espaces topologiques du filtre $\mathcal{F}_x$ défini plus haut.

Définition 2.1. Soit X un espace topologique, $\mathcal{F}$ un filtre sur X et $x \in X$. On dit que $\mathcal{F}$ *converge vers* x si $\mathcal{F}$ contient le filtre $\mathcal{V}_x$ des voisinages de x.

Exercice 5.3. Soit X un espace topologique. Montrer que X est séparé si, et seulement si, tout filtre convergent sur X a une limite unique.

Si l'on veut pouvoir utiliser nos filtres pour faire de la topologie, il faut qu'on comprenne ce qui arrive à un filtre quand on lui applique une fonction f. Si l'on considère simplement l'ensemble des images par f des parties contenues dans notre filtre, on n'obtient en général pas un filtre, tout bêtement parce que f n'est a priori pas surjective! Par contre on obtient bien une base de filtre.

Définition 2.2. Soit X, Y deux ensembles, $\mathcal{F}$ un filtre sur X et $f\colon X \to Y$ une fonction. Alors $\{B \subseteq Y\colon \exists A \in \mathcal{F}\ B = f(A)\}$ est une base de filtre, et on appelle *filtre image* de $\mathcal{F}$ par f le filtre engendré par cette base de filtre.
Notons que A appartient au filtre image de $\mathcal{F}$ par f si, et seulement si, $f^{-1}(A)$ appartient à $\mathcal{F}$.

On laisse en exercice le fait de prouver que la famille introduite ci-dessus est bien une base de filtre.

Proposition 2.3. *Le filtre image d'un ultrafiltre sur X est un ultrafiltre sur Y.*

Démonstration. Soit X, Y deux ensembles, $f\colon X \to Y$ une fonction et $\mathcal{U}$ un ultrafiltre sur X. On sait que $f(\mathcal{U})$ est un filtre. Pour prouver qu'il s'agit en fait d'un ultrafiltre, fixons une partie A de Y dont on suppose qu'elle n'appartient pas à $f(\mathcal{U})$. Alors on sait que $f^{-1}(A)$ n'appartient pas à $\mathcal{U}$, par conséquent $X \setminus f^{-1}(A) \in \mathcal{U}$ et donc $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$ appartient à $\mathcal{U}$. Ceci montre bien que $Y \setminus A$ appartient à $f(\mathcal{U})$, et donc $f(\mathcal{U})$ est un ultrafiltre. □

La proposition ci-dessous explique comment les notions que nous avons introduites permettent de caractériser les fonctions continues.

Proposition 2.4. *Soit X, Y deux espaces topologiques, $x \in X$ et $f\colon X \to Y$ une fonction.*
Alors f est continue en x si, et seulement si, $f(\mathcal{F})$ converge vers $f(x)$ pour tout filtre $\mathcal{F}$ qui converge vers x.

Démonstration. Commençons par supposer f continue en x, et considérons un filtre $\mathcal{F}$ qui converge vers x. Soit V un voisinage de $f(x)$. Comme f est continue en x, $f^{-1}(V)$ est un voisinage de x, par conséquent $f^{-1}(V) \in \mathcal{F}$ puisque $\mathcal{F}$ raffine le filtre des voisinages de x, et donc $V \in f(\mathcal{F})$. Ainsi, $f(\mathcal{F})$ converge vers $f(x)$.
Intéressons-nous maintenant à la réciproque: soit V un ouvert contenant $f(x)$, et $\mathcal{V}$ le filtre des voisinages de x. On sait que $f(\mathcal{V})$ converge vers $f(x)$, par conséquent $V \in f(\mathcal{V})$, ce qui signifie que $f^{-1}(V) \in \mathcal{V}$, et donc $f^{-1}(V)$ est un voisinage de x. Autrement dit, il existe un ouvert U contenant x et contenu dans $f^{-1}(V)$, c'est-à-dire un ouvert U tel que $f(U) \subseteq V$, et on vient de prouver que f est continue en x. □

Continuons à avancer vers une preuve du théorème de Tychonoff ; pour cela il nous faut comprendre la convergence des filtres dans les espaces produits. Rappelons que la topologie produit sur $Y = \prod X_i$ est la topologie la moins fine rendant toutes les projections $\pi_i\colon Y \to X_i$ continues ; une base d'ouverts pour cette topologie est donnée par les ensembles de la forme

$$\{(x_i) \in Y\colon \forall j \in J\ x_j \in U_j\}$$

où J est une partie *finie* de I et chaque U_j est ouvert dans X_j. Il est alors facile de voir qu'une suite (y_n) converge dans Y si, et seulement si, chaque $\pi_i(y_n)$ converge. La proposition suivante généralise ce fait aux filtres.

Proposition 2.5. *Soit $(X_i)_{i\in I}$ une famille d'espaces topologiques, et $X = \prod X_i$ muni de la topologie produit. Un filtre $\mathcal{F}$ sur X est convergent si, et seulement si, chacun des filtres image $\pi_i(\mathcal{F})$ est convergent.*

Démonstration. Notons déjà que, puisque chaque projection $\pi_i\colon X \to X_i$ est continue, on sait que $\pi_i(\mathcal{F})$ est convergent dès que $\mathcal{F}$ l'est. Nous n'avons donc qu'une implication à démontrer.
Supposons maintenant que $\mathcal{F}$ est un filtre sur X tel que chaque $\pi_i(\mathcal{F})$ converge vers $x_i \in X_i$. On va montrer que $\mathcal{F}$ converge vers $x = (x_i)_{i\in I}$. Pour cela, fixons un voisinage de x, dont on peut supposer qu'il est de la forme

$$U = \{y \in X\colon \forall j \in J\ \pi_j(y) \in U_j\}\ ,$$

où $J \subseteq I$ est un ensemble fini et chaque U_j est un ouvert de X_j qui contient x_j.
Par hypothèse, on sait que chaque $\pi_i(\mathcal{F})$ converge vers x_i ; en particulier, pour tout $j \in J$ on doit avoir $U_j \in \pi_j(\mathcal{F})$, c'est-à-dire qu'il existe $V_j \in \mathcal{F}$ tel que $\pi_j(V_j) \subseteq U_j$. Introduisons $V = \cap_{j\in J} V_j$; comme $\mathcal{F}$ est un filtre on sait que $V \in \mathcal{F}$, et de plus on a pour tout $j \in J$ que

$$\pi_j(V) \subseteq \pi_j(V_j) \subseteq U_j\ .$$

Ceci prouve que $V \subseteq U$, et donc $U \in \mathcal{F}$. On vient donc de prouver que tout voisinage de x appartient à $\mathcal{F}$, i.e que $\mathcal{F}$ converge vers x. □

Notons pour plus tard une caractérisation très utile de la convergence des ultrafiltres.

Proposition 2.6. *Soit X un espace topologique, $\mathcal{U}$ un ultrafiltre sur X et $x \in X$. Alors $\mathcal{U}$ converge vers x si, et seulement si,*

$$x \in \bigcap \mathcal{A}, \text{ avec } \mathcal{A} = \{A \subset X : A \in \mathcal{U} \text{ et } A \text{ est fermé}\} .$$

Démonstration. Commençons par supposer que $\mathcal{U}$ converge vers $x \in X$. Alors x appartient à A pour tout $A \in \mathcal{U}$, et on n'a donc essentiellement rien à prouver.
Réciproquement, supposons que x appartienne à l'intersection des éléments de $\mathcal{U}$ qui sont fermés dans X, et fixons un ouvert V contenant x.
On veut montrer que V appartient à $\mathcal{U}$. Si ce n'est pas le cas, on sait que $X \setminus V$ doit appartenir à $\mathcal{U}$, puisque $\mathcal{U}$ est un ultrafiltre. Comme $X \setminus V$ est fermé, on aboutit à une contradiction. □

Encore un dernier effort pour arriver au théorème de Tychonoff: cette fois-ci il nous faut exprimer un critère de compacité en termes de filtre. Ce critère n'est valide qu'en présence de l'axiome du choix.

Proposition 2.7. *Soit X un espace topologique séparé. Alors X est compact si, et seulement si, tout ultrafiltre sur X est convergent.*

Démonstration. Supposons tout d'abord que X n'est pas compact, et considérons un recouvrement (O_i) de X par des ouverts qui ne contienne pas de sous-recouvrement fini. Alors la famille formée par les complémentaires des O_i est une base de filtre, et cette famille se trouve donc contenue (modulo l'axiome du choix) dans un ultrafiltre $\mathcal{U}$. Cet ultrafiltre ne peut converger vers aucun $x \in X$: en effet, pour tout $x \in X$ on a $x \in O_i$ pour au moins un $i \in I$, et comme $O_i \notin \mathcal{U}$ on voit que pour tout $x \in X$ il existe un voisinage de x qui n'appartient pas à $\mathcal{U}$, et donc $\mathcal{U}$ ne converge pas vers x.
Réciproquement, supposons X compact, et considérons un ultrafiltre $\mathcal{U}$ sur X. Alors la famille formée par les élements de $\mathcal{U}$ qui sont fermés dans X a la propriété d'intersections finies non vides (puisque $\mathcal{U}$ est un filtre), et donc a une intersection non vide. Fixons x dans cette intersection ; la proposition 2.6 dit exactement que $\mathcal{U}$ converge vers x. □

A vous maintenant de recoller les morceaux et de vous convaincre qu'on a bien tous les outils en main pour établir [ii] le théorème de Tychonoff, dont l'énoncé est rappelé ci-dessous.

Théorème 2.8. *Soit $(X_i)_{i \in I}$ une famille d'espaces topologiques non vides, et $X = \prod X_i$ muni de la topologie produit. Alors X est compact si, et seulement si, chacun des X_i est compact.*

Notons qu'en fait le théorème de Tychonoff pour une famille d'espaces topologiques séparés X_i se trouve être (un peu) plus faible que l'axiome du choix.

3. Un exemple combinatoire: les ultrafiltres de Ramsey

Définition 3.1. Un ultrafiltre $\mathcal{F}$ sur ω, non principal, est un *ultrafiltre de Ramsey* si, pour toute partition $\{A_n : n < \omega\}$ de ω en $\aleph_0$ morceaux tels que $A_n \notin \mathcal{F}$ pour tout n, il existe $X \in \mathcal{F}$ tel que $|X \cap A_n| \leq 1$ pour tout n [iii].

On dira qu'un filtre (éventuellement principal) a la *propriété de Ramsey* s'il satisfait la seconde condition de la définition d'un ultrafiltre de Ramsey. Notons que, si $\mathcal{F} \subseteq \mathcal{G}$ sont deux filtres et $\mathcal{F}$ a la propriété de Ramsey, alors $\mathcal{G}$ a aussi la propriété de Ramsey.

Cette notion semble arbitraire ; il se trouve pourtant que l'existence d'ultrafiltres de Ramsey a des conséquences importantes sur la structure des ensembles. On sait aujourd'hui que l'existence d'ultrafiltres de Ramsey est indépendante de ZFC. C'est par contre une conséquence (dans ZFC) de l'hypothèse du continu, comme le montre le théorème suivant.

Théorème 3.2. *Si $2^{\aleph_0} = \aleph_1$ alors il existe un ultrafiltre de Ramsey.*

Avant de prouver ce théorème, établissons un lemme simple.

ii. Avec l'axiome du choix!
iii. On peut remplacer, sans changer la notion, cette condition par $|X \cap A_n| = 1$ pour tout n ; pourquoi ?

Lemme 3.3. *Il y a $2^{\aleph_0}$ partitions de ω en $\aleph_0$ morceaux.*

Preuve. Il n'existe que $\aleph_0$ parties finies dans ω, par conséquent il y a $2^{\aleph_0}$ parties de ω infinies et de complémentaire infini. Pour toute telle partie A, on obtient une partition $P(A) = \{B_n\}$ de ω obtenue en énumérant le complémentaire de A sous la forme $\{b_i : 1 \leq i < \omega\}$ et en posant $B_0 = A$, $B_i = \{b_i\}$ pour $1 \leq i < \omega$. L'application $A \mapsto P(A)$ est injective (on retrouve A dans $P(A)$ comme le seul morceau infini de $P(A)$), par conséquent il y a au moins $2^{\aleph_0}$ partitions de ω en $\aleph_0$ morceaux.
Pour voir l'inégalité réciproque, notons que l'ensemble des partitions de ω en $\aleph_0$ morceaux s'injecte naturellement dans $\mathcal{P}(\omega)^{\aleph_0}$, qui est de cardinal $(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$. □

Preuve du théorème 3.2.
Si $2^{\aleph_0} = \aleph_1$, alors on peut énumérer les partitions de ω en $\aleph_0$ morceaux comme une suite $(\mathcal{A}_\alpha)_{\alpha<\omega_1}$ indexée par ω_1.

Construisons maintenant par récurrence une suite indexée par ω_1 de sous-ensembles infinis de ω: on part de $X_0 = \omega$. Supposons maintenant X_β construit pour tout $\beta < \alpha$.

- Si $\alpha = \beta + 1$, deux cas sont possibles: si X_β est d'intersection non vide avec une infinité d'éléments A de la partition $\mathcal{A}_\alpha$, on peut choisir X_α infini, contenu dans X_β, et qui soit tel que $|X_\beta \cap A| \leq 1$ pour tout $A \in \mathcal{A}_\alpha$. Sinon, c'est que X_β est contenu dans la réunion d'un nombre fini d'éléments de A_α, et on peut choisir X_α infini, contenu dans X_β, et tel que $X_\beta \subseteq A$ pour un certain $A \in \mathcal{A}_\alpha$. Pour nous simplifier la vie par la suite, on s'assure aussi que pour tout $i < \omega$ on a $i \notin X_i$.
- Si α est limite, alors on choisit X_α de telle façon que $X_\alpha \smallsetminus X_\beta$ soit fini pour tout $\beta < \alpha$. Le fait qu'il est bien possible de faire ça sera justifié par le lemme 3.5 à la fin de la preuve.

Montrons que la famille $\{X_\alpha : \alpha < \omega_1\}$ est une base de filtre: en effet, si on considère $X_{\alpha_1}, \ldots, X_{\alpha_n}$ et qu'on fixe un ordinal limite dénombrable qui majore strictement $\alpha_1, \ldots, \alpha_n$ alors on sait par construction que $X_\gamma \smallsetminus X_{\alpha_i}$ est fini pour tout $i \in \{1, \ldots, n\}$; par conséquent $X_\gamma \smallsetminus (\cap X_{\alpha_i})$ est fini et, comme X_γ est infini, ceci prouve que $\cap X_{\alpha_i}$ est infini (donc non vide!). En fait, le raisonnement précédent nous donne un meilleur résultat.

Lemme 3.4. *Pour tout $\alpha < \beta < \omega_1$, $X_\beta \smallsetminus X_\alpha$ est fini.*

Preuve du Lemme 3.4.
On raisonne par récurrence transfinie: on va prouver que pour tout β la propriété « pour tout $\alpha < \beta$, $X_\beta \smallsetminus X_\alpha$ est fini » est vraie.
Cette propriété est trivialement vraie pour $\beta = 0$. Si elle est vraie au rang β, elle est vraie aussi au rang $\beta + 1$, puisque $X_{\beta+1} \subseteq X_\beta$. Il nous reste simplement à vérifier notre propriété aux ordinaux limites. Soit donc β un ordinal limite, et $\alpha < \beta$. Alors X_β a été construit de telle façon que $X_\beta \smallsetminus X_\alpha$ soit fini, ce qui conclut la preuve du lemme. □

Appelons maintenant $\mathcal{F}$ le filtre engendré par la famille $\{X_\alpha : \alpha < \omega_1\}$; on a vu qu'il ne contient que des parties infinies, et il est facile de voir qu'il a la propriété de Ramsey: si on a une partition de ω en $\aleph_0$ morceaux, cette partition apparaît sous la forme $\mathcal{A}_\alpha$ pour un certain ordinal successeur α ; si aucun élément de $\mathcal{F}$ n'appartient à la partition, c'est qu'en particulier $X_{\alpha+1}$ n'est inclus dans aucun élément de cette partition. Notre construction nous dit alors qu'on a choisi $X_{\alpha+1}$ de telle façon que $|X_\alpha \cap A| \leq 1$ pour tout $A \in \mathcal{A}$. Comme $X_{\alpha+1} \in \mathcal{F}$, on a bien montré que $\mathcal{F}$ a la propriété de Ramsey.

Notons également que $\mathcal{F}$ ne peut, par construction, pas être contenu dans un filtre principal. Pour cela, il suffit de prouver que pour toute partie $A \subseteq \omega$ il existe un élément de $\mathcal{F}$ qui ne contient pas A.
Si $|A| \geq 2$, on partitionne A en morceaux finis A_i tels que A_0 est de cardinal ≥ 2, et on étend cette partition en une partition de ω en $\aleph_0$ morceaux finis. Aucun des éléments de la partition ne peut appartenir à $\mathcal{F}$, ce qui nous donne, puisque $\mathcal{F}$ a la propriété de Ramsey, l'existence de $X \in \mathcal{F}$ tel que $|X \cap A_0| \leq 1$, en particulier X ne contient pas A.
Il nous reste à voir qu'il ne peut pas exister un entier $n < \omega$ tel que tous les X_α contiennent n. Mais le début de notre construction a justement garanti que $i \notin X_i$.

Finalement, on a donc construit un filtre $\mathcal{F}$ qui a la propriété de Ramsey et n'est contenu dans aucun filtre principal ; tout ultrafiltre le contenant est un ultrafiltre de Ramsey, ce qui conclut la preuve, modulo la justification du fait que notre construction peut effectivement être menée à bien aux ordinaux limites. Cette justification se base sur le fait suivant, souvent utilisé en combinatoire infinie.

Lemme 3.5. *Soit $\{Y_i\}_{i\in I} \subseteq \mathcal{P}(\omega)$ une famille dénombrable de sous-ensembles de ω tels que $\bigcap_{j\in J} Y_j$ soit infini pour toute partie finie $J \subseteq I$. Alors il existe une partie $Y \subseteq \omega$ infinie et telle que $Y \smallsetminus Y_i$ soit fini pour tout i.*

Comment appliquer ce lemme pour mener à bien notre construction ? Eh bien, si α est dénombrable, limite et qu'on a construit X_β pour tout $\beta < \alpha$ en respectant les propriétés imposées par notre construction, alors pour tout $\beta \leq \gamma < \alpha$ on sait (en reprenant le raisonnement du Lemme 3.4) que $X_\gamma \smallsetminus X_\beta$ est fini. Mais pour tout ensemble fini d'ordinaux $\beta_1, \ldots, \beta_n < \alpha$, si on pose $\beta = \max\{\beta_i \colon i = 1, \ldots, n\}$ alors la construction assure que

$$X_\beta \smallsetminus \left(\bigcap_{i=1}^{n} X_{\beta_i}\right) = \bigcup_{i=1}^{n} (X_\beta \smallsetminus X_{\beta_i}) \text{ est fini } .$$

Puisque X_β est infini, ceci impose bien que $\bigcap_{i=1}^{n} X_{\beta_i}$ est infini. En appliquant le lemme 3.5 à la famille $\{X_\beta\}_{\beta<\alpha}$, on obtient donc une partie Y telle que $Y \smallsetminus X_\beta$ est fini pour tout $\beta < \alpha$, et on peut finalement poser $X_\alpha = Y$.

Preuve du Lemme 3.5.
On peut bien sûr supposer que $I = \omega$ et alors, quitte à remplacer chaque Y_i par $\bigcap_{j=1}^{i} Y_j$, supposer que la suite (Y_i) est une suite décroissante d'ensembles infinis. Comme les Y_i sont infinis, on peut construire une suite strictement croissante $(y_i)_{i<\omega}$ telle que $y_i \in Y_i$ pour tout i, et $Y = \{y_i\}_{i<\omega}$ satisfait les conditions du lemme.

Ceci conclut la preuve du lemme, qui était tout ce qu'il nous manquait pour finir de justifier l'existence d'un ultrafiltre de Ramsey dans un univers où les axiomes de ZFC et l'hypothèse du continu sont vrais. □

Bibliographie

[Dug03] Pierre Dugac. *Histoire de l'Analyse*. Vuibert, Paris, 2003.

[Hal74] Paul R. Halmos. *Naive set theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1974. Reprint of the 1960 edition.

[HR98] Paul Howard and Jean E. Rubin. *Consequences of the axiom of choice*, volume 59 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998.

[Jec73] Thomas J. Jech. *The axiom of choice*. Studies in Logic and the Foundations of Mathematics, Vol. 75. North-Holland Publishing Co., Amsterdam, 1973.

[Jec03] Thomas Jech. *Set theory: The third millennium edition, revised and expanded*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003.

[KM76] Kazimierz Kuratowski and Andrzej Mostowski. *Set theory, with an introduction to descriptive set theory*, volume 86 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1976.

[Kri98] Jean-Louis Krivine. *Théorie des ensembles*. Nouvelle Bibliothèque mathématique. Cassini, Paris, 1998.

[Mos06] Yiannis Moschovakis. *Notes on set theory*. Undergraduate Texts in Mathematics. Springer, New York, second edition, 2006.

[Wag85] Stan Wagon. *The Banach-Tarski paradox*, volume 24 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1985.

Index