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 1
1. Théorie des ensembles intuitive, et ses problèmes 1
2. Une théorie axiomatique des ensembles 2

Chapitre 2. Les ordinaux 5
1. Bons ordres 5
2. Ordinaux 7
3. Récurrence transfinie et arithmétique des ordinaux 11

Chapitre 3. L'axiome du choix 15

Chapitre 4. Cardinaux 19
1. Définition des cardinaux 19
2. Arithmétique des cardinaux 21
3. Dénombrabilité 24
4. Cardinaux réguliers et cofinalité 25

Chapitre 5. Filtres et ultrafiltres 29
1. Définitions, premières propriétés 29
2. Utilisation des filtres en topologie 30
3. Un exemple combinatoire: les ultrafiltres de Ramsey 32

Bibliographie 35

CHAPITRE 1

Les axiomes de Zermelo-Fraenkel

1. Théorie des ensembles intuitive, et ses problèmes

Un ensemble, pour paraphraser un logicien lyonnais, j'espère que vous savez ce que c'est – ou, au moins, que vous croyez savoir ce que c'est. En effet, puisque nous proposons d'utiliser les ensembles comme objets fondamentaux, les objets mathématiques les plus basiques, à partir desquels nous allons pouvoir définir ou construire tout autre objet mathématique, nous ne pourrons pas définir ce que sont les ensemble par le moyen d'objets encore plus simples.

Plus tard, les *axiomes* de la théorie des ensembles nous fournirons certaines propriétés des ensembles, sans pour autant prétendre les définir. Pour l'instant, nous partirons (très) naïvement de l'idée qu'un ensemble n'est qu'une collection d'objets, et nous verrons où cela nous mène. Cette idée implique en particulier les deux principes suivants:

(i) *Compréhension (non restreinte)* : toute collection imaginable (ça veut dire quoi, ça?) d'objets est un ensemble.

(ii) *Extensionnalité* : puisqu'un ensemble n'est que la collection de ses membres, deux ensembles ayant les même membres sont égaux.

Par exemple, pour tous deux objets x et y il existe l'ensemble paire $\{x, y\}$. Cette paire est *non ordonnée* – d'après l'extensionnalité, $\{x, y\} = \{y, x\}$ (remarquons également que si $x = y$ alors $\{x, y\} = \{x, x\} = \{x\}$). Nous définissons une *paire ordonnée* formée de x et de y par :

$$(x, y) = \big\{\{x, y\}, \{x\}\big\}.$$

Exercice 1.1.1. Montrer que :

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

Pour deux ensembles X et Y, nous pouvons définir leur *produit cartésien* par:

$$X \times Y = \big\{(x, y) \colon x \in X \text{ et } y \in Y\big\}.$$

Une *application* est un ensemble f dont tous les membres sont des paires ordonnés, et tel que pour tout x il existe au plus un y tel que $(x, y) \in f$, auquel cas nous disons que $f(x) = y$ (autrement dit, nous représentons une application par son graphe). Nous définissons le *domaine* et l'*image* d'une application f par

$$\operatorname{dom} f = \{x \colon \exists y\, (x, y) \in f\}, \qquad \operatorname{img} f = \{y \colon \exists x\, (x, y) \in f\} = \{f(x) \colon x \in \operatorname{dom} f\}.$$

La notation $f \colon X \to Y$ signifie que f est une application, $X = \operatorname{dom} f$ et $Y \supseteq \operatorname{img} f$.

Comment représenter les nombres par des ensembles? Commençons (comme toujours) par les nombres les plus basiques, les entiers naturels. D'une façon récursive, l'entier n sera représenté par l'ensemble $\{0, 1, \ldots, n-1\}$:

$$0 = \varnothing, \qquad 1 = \{0\} = \{\varnothing\}, \qquad 2 = \{0, 1\} = \{\varnothing, \{\varnothing\}\}, \ldots$$

La collection de tous les ensembles de cette forme sera notée par $\mathbf{N}$ (nous rendrons tout ceci plus formel plus tard). À partir de $\mathbf{N}$ nous pouvons construire l'ensemble des entiers relatifs $\mathbf{Z}$ (par exemple, comme l'ensemble des paires ordonnées $(m, n) \in \mathbf{N} \times \mathbf{N}$ telles que $m = 0$ ou $n = 0$, où $(n, 0)$ représente n et $(0, n) = -n$). À partir de $\mathbf{Z}$ nous pouvons construire $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, et ainsi de suite...

Plus tard, nous pouvons essayer de comparer les tailles, ou *cardinalités*, des ensembles. Nous pourrons démontrer qu'il existe une application bijective $f \colon \mathbf{N} \to \mathbf{Q}$, donc $\mathbf{N}$ et $\mathbf{Q}$ ont la même cardinalité, ou encore, $\mathbf{Q}$ est un ensemble *dénombrable*. Par contre, il n'existe pas d'application surjective $f \colon \mathbf{N} \to \mathbf{R}$ (le Théorème de Cantor), l'ensemble $\mathbf{R}$ est donc non dénombrable.

Jusqu'ici, tout va bien.

Démontrons la variante suivante du Théorème de Cantor.

NOTATION 1.1.2. Pour un ensemble X, nous notons par $\mathcal{P}(X)$ l'*ensemble des parties* de X:

$$\mathcal{P}(X) = \{Y : Y \subseteq X\}.$$

THÉORÈME 1.1.3 (Cantor). *Soit X un ensemble. Alors il n'existe pas d'application surjective $f : X \to \mathcal{P}(X)$.*

DÉMONSTRATION. Soit $f : X \to \mathcal{P}(X)$, et posons:

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

Supposons, par l'absurde, que $Y \in \operatorname{img} f$. Alors $Y = f(x_0)$ pour un certain $x_0 \in X$. Par définition de Y, ceci donne $x_0 \in Y \iff x_0 \notin f(x_0) = Y$, ce qui est absurde. Ainsi, aucune application $f : X \to \mathcal{P}(X)$ ne peut être surjective. ■$_{1.1.3}$

Mais attention : soit X l'ensemble de tous les ensembles. Alors $\mathcal{P}(X) \subseteq X$, et nous pouvons définir une application $f : X \to \mathcal{P}(X)$ par $f(x) = x$ si $x \in \mathcal{P}(X)$, $f(x) = \varnothing$ sinon, et c'est surjectif ! Nous nous trouvons obligés d'admettre que nos deux principes (compréhension et extensionnalité) sont trop forts, au point d'être contradictoires. Cette contradiction est exprimée d'une façon plus concise par le célèbre Paradoxe de Russel :

> Soit $X = \{A : A \notin A\}$, l'ensemble de tous les ensembles qui ne sont pas leur propre membres. Alors, si $X \in X$, c'est que nécessairement $X \notin X$; et si $X \notin X$, alors par définition de X, on a bien $X \in X$...

2. Une théorie axiomatique des ensembles

Le paradoxe de Russel provient du principe de compréhension, pourtant assez intuitif, et nous nous trouvons dans l'obligation d'accepter que toute collection ne peut pas nécessairement être un ensemble. Nous allons remplacer ce principe par un système d'axiomes, les *axiomes* de Zermelo-Fraenkel (ZF), où la compréhension non restreinte est remplacée par plusieurs axiomes affirmant que *certaines* collections sont des ensembles (intuitivement, car elles sont suffisamment petites pour l'être).

Dans cette approche *axiomatique* nous allons éviter entièrement la question de « qu'est-ce qu'un ensemble ». Plutôt, nous allons supposer que nous travaillons avec un « univers », noté le plus souvent V, qui consiste en des objets que l'on appelle « ensembles », muni d'une relation binaire $\in$, l'*appartenance* (ainsi que de la relation égalité $=$, qui n'est vérifiée qu'entre un ensemble et lui-même), et qui vérifient les axiomes que nous allons énumérer. Nous remarquons au passage qu'une conséquence de cette approche axiomatique particulière est que tout objet est un ensemble.

À partir de maintenant nous allons éviter le mot « collection » et parler plutôt de « classe » : une *classe* est toute collection d'ensembles que nous pouvons définir par une propriété de ses membres, c'est à dire de la forme

$$C = \{x : P(x)\}$$

où P est une propriété d'ensembles (formellement, il faudrait préciser : P est une propriété que l'on peut exprimer en *en logique du premier ordre*, mais en ce moment nous n'allons pas insister sur ce point). Nous remarquons que tout ensemble x est une classe : $x = \{t : t \in x\}$, mais toute classe n'est pas nécessairement un ensemble – une telle classe sera appelée une *classe propre*. Par exemple l'univers des ensembles V est une classe propre (car sinon, le paradoxe de Russel revient...)

I. Axiome d'extensionnalité

> Deux ensembles sont égaux dès qu'ils ont les mêmes éléments. En langage formel,
>
> $$\forall x \forall y \Big[\big(\forall z\,(z \in x) \leftrightarrow (z \in y)\big) \to x = y\Big].$$

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.

II. Axiome de fondation

> Pour tout ensemble non vide x, il existe un ensemble $y \in x$ et tel que $y \cap x = \varnothing$.

Cet axiome peut paraître un peu obscur, ceci n'est pas particulièrement gênant, on y reviendra plus tard.

III. Axiome de la réunion

La réunion d'un ensemble (une famille) d'ensembles est un ensemble : pour tout x, la réunion $\bigcup_{y\in x} y$ est un ensemble. Formellement,

$$\forall x\,\exists z\,\forall t\big[(t\in z)\leftrightarrow(\exists y(y\in x \text{ et } t\in y))\big].$$

IV. Axiome de l'ensemble des parties

Pour tout X, la classe $\mathcal{P}(X)=\{y\colon y\subseteq X\}$ est un ensemble :

$$\forall x\,\exists y\,\forall z\,\big[(z\in y)\leftrightarrow(\forall t(t\in z)\to t\in x)\big].$$

V. Axiome de l'infini

Pour un ensemble y, notons $S(y)=y\cup\{y\}$, que l'on appelle le *successeur* de y.

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 $S(y)\in x$ (i.e., il existe $z\in x$, nécessairement unique, tel que $\forall t\big[t\in z\leftrightarrow(t\in y\vee t=y)\big]$).

Dans le prochain chapitre on verra que cela peut être énoncé d'une manière équivalente comme « il existe un ordinal infini », ou encore « la classe de tous les ordinaux finis est un ensemble ».

Exercice 1.2.1. Montrer à partir des axiomes (y compris ceux que l'on énonce plus bas) que $S(y)$ est un ensemble, quelque soit l'ensemble y.

Vérifier que si nous représentons les entiers naturels par des ensembles comme dans la section précédente, alors $S(n)=n+1$.

VI. Axiome de remplacement

Disons qu'une propriété de paires (ordonnées) $R(x,y)$ est une *relation fonctionnelle* si pour tout x il existe au plus un seul y tel que $R(x,y)$ est vrai. Une telle relation définit une loi F_R, où $F_R(x)=y$ si y est l'unique tel que $R(x,y)$, et s'il n'existe pas un tel y alors $F_R(x)$ n'est pas défini.

Pour toute relation fonctionnelle R, et pour tout x, l'image de x par la loi F_R est un ensemble. Autrement dit, la classe suivante est un ensemble :

$$\{F_R(t)\colon t\in x \text{ et } F_R(t) \text{ est défini}\}=\{z\colon \exists t\in x \text{ tel que } R(t,z)\}.$$

Notons que F_R n'est pas nécessairement une application, c'est à dire que la classe $\{(t,z)\colon R(t,z)\}$ n'est pas nécessairement un ensemble – pour cela il nous faudrait la compréhension non restreinte.

De ces axiomes découlent certains principes qui des fois sont aussi considérés comme axiomes (superflus) :

Axiome de compréhension restreinte (ou de séparation)

Pour tout x et propriété P, la classe de membres de x qui vérifient P est un ensemble :

$$\{t\in x\colon P(t)\}.$$

Ceci découle de l'axiome du remplacement, où la relation $R(t,z)$ est : « $t=z$ et $P(t)$ ».

Axiome de la paire

Pour tous x et y, la classe $\{x,y\}$ est un ensemble :

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

En effet, avec l'axiome de l'ensemble des parties on peut former les ensembles $\mathcal{P}(\varnothing)=\{\varnothing\}$ et $\mathcal{P}(\{\varnothing\})=\{\varnothing,\{\varnothing\}\}$. Soient maintenant x et y deux ensembles quelconques, et soir $R(t,z)$ la relation fonctionnelle

$$(t=\varnothing \text{ et } z=x) \text{ ou } (t=\{\varnothing\} \text{ et } z=y).$$

En appliquant l'axiome de replacement à $\{\varnothing,\{\varnothing\}\}$ et à cette relation, on obtient que $\{x,y\}$ est un ensemble.

Exercice 1.2.2. Vérifier que tout ce qu'on a fait dans la première section, avec la compréhension non restreinte, reste valable dans ZF. En particulier, pourquoi le produit cartésien de deux ensembles est-il un ensemble ?

Il serait malhonnête de conclure cette section sans évoquer le problème suivant: existe-il un univers V 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 (à partir de ZF) que le système ZF est consistant, 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

Ayant énoncé les axiomes de ZF, autrement dit, ayant défini le cadre axiomatique dans lequel nous travaillerons, commençons à faire des maths. En un premier temps, 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 ? Considérons d'abord des ensembles munis d'un ordre permettant une énumération.

Définition 2.1.1. Soit X un ensemble. Un *bon ordre* sur X est une relation d'ordre $\leq$ sur X tel que toute partie 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.
- 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 2.1.2. On dit que $s \subseteq X$ est un *segment initial* si

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

Si $x \in X$ on notera $X_{<x}$ le segment initial $\{y \in X : y < x\}$; 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 $X_{<x}$ pour un unique $x = \min(X \setminus s)$. 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 par 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 2.1.3 (Démonstration par récurrence transfinie). *Soit P une propriété*[i] *et X un ensemble bien ordonné, tel que pour tout $x \in X$:*

$$\big[\forall y \in X \ y < x \to P(y)\big] \to P(x)$$

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

i. Comme pour l'axiome de remplacement, nous devrions exiger que la propriété puisse être exprimée par la logique du premier ordre.

DÉMONSTRATION. Soit $Y = \{x \in X : \neg P(y)\}$. Si $Y = \varnothing$ tout est bien. Sinon, Y admet un plus petit élément, y. Autrement dit, pour tout $z < y$ on a $P(y)$, d'où par hypothèse $P(y)$, une contradiction. ■2.1.3

À titre d'exemple, montrons :

COROLLAIRE 2.1.4. *Soit $(X, \leq)$ un ensemble bien ordonné et $f : X \to X$ 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. ■2.1.4

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

COROLLAIRE 2.1.5. *Soit $(X, \leq)$ un ensemble bien ordonné, $Y, W \subseteq X$ deux segments initiaux, et $f : Y \to W$ un isomorphisme. Alors $W = Y$ et $f = \mathrm{id}_Y$.*

DÉMONSTRATION. D'abord, puisque W et Y sont des segments initiaux, on a $W \subseteq Y$ ou $Y \subseteq W$. Quitte à remplacer f par f^{-1}, on peut supposer que $W \subseteq Y$. Montrons tout d'abord que $W = Y$. Pour cela, prenons $x \in Y$. On a $f(x) \in W$, et $f(x) \geq x$ d'après Corollaire 2.1.4. Comme W est un segment initial, on en déduit que $x \in W$ et $W = Y$. Maintenant, $f^{-1} : Y \to Y$ est aussi un isomorphisme, d'où $f^{-1}(x) \geq x$ pour tout $x \in Y$, ce qui en composant par f donne $x \geq f(x)$ et donc $f(x) = x$ pour tout $x \in Y$. ■2.1.5

THÉORÈME 2.1.6 (Construction par récurrence transfinie). *Soit $(X, \leq)$ un bon ordre, 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 $x \in X$,*

$$f(x) = G(f{\restriction}_{X_{<x}}).$$

DÉMONSTRATION. D'abord, montrons qu'une telle fonction, si elle existe, doit être unique. En effet, supposons que f et g vérifient la condition. Pour chaque $x \in X$, si $f(y) = g(y)$ pour tout $y < x$, c'est que $f{\restriction}_{X_{<x}} = g{\restriction}_{X_{<x}}$ est donc $f(x) = g(x)$. L'unicité suit par récurrence transfinie.

Montrons maintenant par récurrence que pour tout $x \in X$:

($*_x$) il existe une unique fonction f_x de domaine $X_{<x}$ telle qu'on ait

$$f_x(y) = G(f_x{\restriction}_{X_{<y}}) \qquad \forall y < x.$$

Supposons que ($*_z$) est vrai pour tout $z < x$, et montrons pour x. L'unicité est déjà connue. Si $z < y < x$ alors $f_y{\restriction}_{X_{<z}}$ vérifie bien la condition de ($*_z$), d'où $f_z = f_y{\restriction}_{X_{<z}}$, ou encore, $f_y(z) = G(f_z)$. Posons maintenant $f_x(y) = G(f_y)$ pour chaque $y < x$. Alors, pour chaque $z < y < x$ on a $f_y(z) = G(f_z) = f_x(z)$, d'où $f_y = f_x{\restriction}_{X_{<y}}$, et donc $f_x(y) = G(f_x{\restriction}_{X_{<y}})$. L'existence de f_x est ainsi démontrée.

Maintenant, nous posons $f(x) = G(f_x)$ pour chaque $x \in X$. Comme plus haut, $f{\restriction}_{X_{<x}} = f_x$ par l'unicité de f_x, d'où $f(x) = G(f{\restriction}_{X_{<x}})$, et la preuve est finie. ■2.1.6

COROLLAIRE 2.1.7. *Soient $(X,)$ et $(Y, \leq)$ deux bons ordres. Alors exactement l'une des option suivante est vraie :*

(i) *X et Y sont isomorphes.*

(ii) *X est isomorphe à un unique segment initial propre de Y.*

(iii) *Y est isomorphe à un unique segment initial propre de X.*

DÉMONSTRATION. Soit $* \notin Y$, et définissons une application $f : X \to Y \cup \{*\}$ par récurrence. Pour chaque $x \in X$:

– Si $Y \smallsetminus \operatorname{img} f{\restriction}_{X_{<x}} \neq \varnothing$, nous posons $f(x) = \min\big(Y \smallsetminus \operatorname{img} f{\restriction}_{X_{<x}}\big)$.
– Sinon, $f(x) = *$.

(Par la suite, nous n'allons pas s'occuper de cet $*$, et juste dire: « nous posons $f(x) = \min\big(Y \smallsetminus \operatorname{img} f{\restriction}_{X_{<x}}\big)$ tant que ceci est possible, puis on s'arrête ».)

– Si on a réussi à définir $f : X \to Y$ (sans recours à $*$), c'est que f est un isomorphisme entre X et un segment initial $Y_0 = f(X) \subseteq Y$. Ce segment initial est unique par Corollaire 2.1.5, et peut être propre ou non.

– Sinon, soit $x \in X$ le plus petit tel que $f(x) = *$. Alors $f\restriction_{X_{<x}} : X_{<x} \to Y$ est un isomorphisme entre un segment initial propre de X et Y. L'unicité de ce segment initial de X est encore donnée par Corollaire 2.1.5.

■2.1.7

EXERCICE 2.1.8. 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. Ce segment initial est-il nécessairement propre ?

2. Ordinaux

Le Corollaire 2.1.7 rend plus précise l'idée qu'un bon ordre correspond à une énumération : si X et Y sont bien ordonnés, nous pouvons les énumérer côte à côte, obtenant un isomorphisme entre l'énumération la plus courte et un segment initial de la plus longue, ou constater que les deux énumérations sont de la même longueur.

Maintenant, pouvons nous associer à un bon ordre un objet qui représente sa longueur ? Par exemple, dans chaque classe d'isomorphisme de bons ordres (i.e., pour chaque longueur), pouvons nous distinguer un représentant « canonique » ?

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 la longueur de X serait

$$\text{longueur de } X = \{\text{longueur de } S \colon S \text{ segment initial propre de } X\}, \tag{1}$$

ce qui est d'ailleurs exactement le cas des nombres finies $n = \{0, 1, \ldots, n-1\}$.

Quelles seraient les propriétés de telles « longueurs » ?

– Soit α la longueur de X, et supposons que $\gamma \in \beta \in \alpha$. Alors β est la longueur d'un segment initial propre S_1 de X, et γ est la longueur d'un segment initial propre S_2 de S_1. Or, S_2 est aussi un segment initial propre de X, d'où $\gamma \in \alpha$.
– L'application qui à $x \in X$ associe la longueur de $X_{<x}$ est une bijection $f \colon X \to \alpha$, et

$$x < y \iff X_{<x} \subsetneq X_{<y} \iff f(x) \in f(y).$$

Ainsi, $\in$ est un bon ordre strict sur α.

Formalisons cette idée:

DÉFINITION 2.2.1. 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.

LEMME 2.2.2. *Une réunion d'ensemble transitifs est encore un ensemble transitif ; une intersection d'ensembles transitifs est encore un ensemble transitif.*

DÉFINITION 2.2.3. 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 2.2.4. *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 $\gamma' \in Y$. ■2.2.4

LEMME 2.2.5. *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) $\iff$ (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 $\beta = \{\gamma \in \alpha \colon \gamma \in \beta\} = \{\gamma \in \alpha \colon \gamma < \beta\} = \alpha_{<\beta}$. En particulier, β est bien un segment initial propre de α.

Réciproquement, soit β un segment initial propre de α. Alors $\beta = \alpha_{<\gamma}$ où $\gamma = \min(\alpha \setminus \beta)$. Or, d'après l'argument précédent $\alpha_{<\gamma} = \gamma \in \alpha$.

(ii) $\iff$ (iii). Découle de Lemme 2.2.4. ■2.2.5

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 2.2.6. *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 2.2.5, $\alpha \in \beta$ pour tout $\beta \in A$, donc $\alpha \in \alpha$, absurde. Donc $\alpha \in A$. ■2.2.6

PROPOSITION 2.2.7. *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 2.2.6, $\gamma = \alpha$ ou $\gamma = \beta$, i.e., $\alpha \subseteq \beta$ ou $\beta \subseteq \alpha$.

■2.2.7

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) $\alpha \leq \beta \iff (\alpha \in \beta$ 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 :

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

PROPOSITION 2.2.8. *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 2.2.6, 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, $\min C = \min C'$. ■2.2.8

LEMME 2.2.9. (i) *Tout ensemble transitif d'ordinaux est un ordinal.*

(ii) *Tout ensemble qui est un segment initial de ON est un ordinal.*

DÉMONSTRATION. (i) Car l'ordre $\in$ sur les ordinaux est bon.

(ii) Si $A \subseteq ON$ est un segment initial, alors il est transitif.

■2.2.9

PROPOSITION 2.2.10. *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 $\beta \in A$. ■2.2.10

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

LEMME 2.2.11. *Pour tout ordinal α, $S(\alpha) = \alpha \cup \{\alpha\}$ est aussi un ordinal, et c'est le plut petit ordinal $\gamma > \alpha$ (donc, c'est le* successeur *de α). En particulier, il n'existe pas un plus grand ordinal.*

DÉMONSTRATION. L'ensemble $S(\alpha) = \{\beta : \beta \leq \alpha\}$ est un segment initial de ON, donc un ordinal, et si $\gamma > \alpha$ c'est que $\gamma = \{\beta : \beta < \gamma\} \supseteq S(\alpha)$. ■2.2.11

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

DÉMONSTRATION. Si ON était un ensemble alors $\alpha = ON$ serait un ordinal (segment initial de ON), ce qui donnerait $\alpha < \alpha$, un absurde. ■2.2.12

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.

Introduisons un peu de terminologie.

DÉFINITION 2.2.13. 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 2.2.14. Montrer qu'un ordinal β est limite si, et seulement si, $\beta = \sup\{\eta : \eta < \beta\} \neq 0$.

DÉFINITION 2.2.15. 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...)

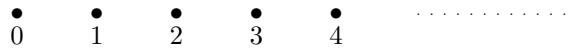
PROPOSITION 2.2.16. *Modulo les autres axiomes de ZF (auxquels on peut même enlever l'axiome de fondation), l'axiome de l'infini est équivalente à l'énoncé : il existe l'ensemble de tous les ordinaux finis. Cet ensemble, noté ω, est un ordinal : c'est le plus petit ordinal infini, et aussi le plus petit 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. ■2.2.16



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.

- La classe des ordinaux ON est totalement ordonnée et bien ordonné par l'ordre strict $\in$, équivalent à $\subsetneq$.
- Pour tout ordinal α on a $\alpha = \{\beta \in ON : \beta < \alpha\}$. Il est encore bien ordonné par l'ordre induit de ON.
- La réunion d'un ensemble d'ordinaux E est un ordinal, qui est la borne supérieure de E (non nécessairement atteinte).
- L'intersection d'une classe non vide d'ordinaux E est un ordinal, qui est le plus petit élément de E.
- Il existe trois types d'ordinaux: zéro (vide), les ordinaux successeurs (ceux qui ont un plus grand élément) et les ordinaux limites (ceux qui n'ont pas de plus grand élément).

Montrons qu'il existe des ordinaux « arbitrairement grands ».

THÉORÈME 2.2.17 (Hartogs). *Soit X un ensemble quelconque. Alors il existe un ordinal α qui ne s'injecte pas dans X. En particulier, il existe un plus petit ordinal possédant cette propriété, que l'on 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 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, au plus, 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. ■2.2.17

Nous avons maintenant tout ce qu'il faut pour « énumérer » tout ensemble bien ordonné :

COROLLAIRE 2.2.18. *Tout ensemble bien ordonné est isomorphe à un unique ordinal, et par un isomorphisme unique.*

DÉMONSTRATION. Soit X bien ordonné et soit α l'ordinal de Hartogs associé à X. Par Corollaire 2.1.7, X est isomorphe à un segment initial de α, ou α est isomorphe à un segment initial propre de X. Or, la deuxième possibilité est exclue par définition de α. Ainsi, X est isomorphe à un segment initial de α, c'est à dire à un ordinal. L'unicité de encore par Corollaire 2.1.7. ■2.2.18

EXERCICE 2.2.19. Soit $(X, <)$ un bon ordre.

(i) Montrer que existe une unique fonction de domaine X telle que pour tout $x \in X$:

$$f(x) = \mathrm{img}(f\restriction_{X_{<x}}).$$

(ii) Montrer que $\mathrm{img}\, f$ est un ensemble transitif.

(iii) Que peut-on dire d'autre de $\mathrm{img}\, f$ (et de f) ?

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 2.3.1 (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. ■2.3.1

D'une manière semblable, quand nous construisons une fonction f sur les ordinaux, par récurrence transfinie, nous allons souvent considérer trois cas:

(i) Cas zéro : nous précisons $f(0)$.

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

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

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 2.3.2 (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:

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

Par exemple, on a $1 + \omega = \sup\{1 + n : n < \omega\} = \omega$. 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!

L'ordinal $\omega + 1$

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):

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

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 $\delta+\beta = \sup\{\delta+\eta\colon \eta<\beta\} > \delta+S(\alpha) > \delta+\alpha$.

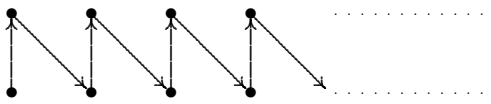
Ceci achève la démonstration ; notons pour rassurer le lecteur que la rédaction ci-dessus est particulièrement lourde et détaillée, et que par la suite on évitera de trop rentrer dans le détail de raisonnements élémentaires comme celui-ci. Mais il faut vérifier qu'un raisonnement d'apparence élémentaire 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 2.3.3. (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:

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

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 $2.\omega = \omega$.

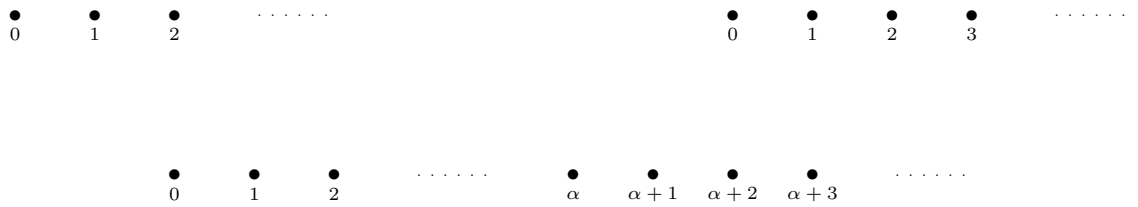


$$2 \cdot \omega = \omega$$

EXERCICE 2.3.4. 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 2.3.5.
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).



$$\alpha + \beta = \{0, 1, 2, \ldots, \alpha, \alpha+1, \alpha+2, \alpha+3, \ldots\}$$

Un exercice pour vous entraîner aux démonstrations par récurrence transfinie:

EXERCICE 2.3.6. 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.

DÉFINITION 2.3.7. 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 $\alpha > 1$ et $\beta > \gamma$ 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.

THÉORÈME 3.0.8. *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, soit α l'ordinal de Hartogs correspondant, et soit

$$S = \{\text{application injective } f\colon \operatorname{dom} f \in ON \text{ et } \operatorname{img} f \subseteq X\}.$$

Alors $S \subseteq \mathcal{P}(\alpha \times X)$, donc c'est bien un ensemble, que l'on munit de l'ordre partiel $\subseteq$. Soit $C = \{f_i\}_{i\in I} \subseteq S$ une chaîne : alors $\beta = \sup \operatorname{dom} f_i$ est un ordinal et $f = \bigcup f_i$ est une application injective, $f\colon \beta \to X$. Donc, $f \in S$ majore C.

D'après le Lemme de Zorn il existe $f \in S$ maximal, de domaine $\beta = \operatorname{dom} f$. Si $\operatorname{img} f \neq X$, c'est qu'il existe $a \in X \setminus \operatorname{img} f$, et nous pouvons étendre (injectivement) f à $\beta + 1$, envoyant $\beta \mapsto a$. Ceci contredit la maximalité de f, d'où $\operatorname{img} f = X$, et nous avons un bon ordre sur X défini par:

$$a <_f b \quad \Longleftrightarrow \quad f^{-1}(a) < f^{-1}(b) \text{ dans } \beta.$$

(ii) $\Longrightarrow$ (iii). Soit X un ensemble 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 donc un plus petit β tel que $g(\beta) = *$. 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. ■3.0.8

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 3.0.9. 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...

Exercice 3.0.10. 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*.

i. Ce qui n'est pas forcément une mauvaise chose!
ii. et ce fait est indépendant de ZF.

DÉFINITION 3.0.11. 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.

CHAPITRE 4

Cardinaux

1. Définition des cardinaux

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

DÉFINITION 4.1.1. 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 4.1.2 (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 la décomposition suivante :

$$\begin{aligned} X &= \big(X \smallsetminus Y\big) \cup \big(Y \smallsetminus f(X)\big) \cup \big(f(X) \smallsetminus f(Y)\big) \cup \ldots \cup \bigcap_n f^n(X), \\ &= \left[\bigcup_{n\geq 0} f^n\big(X \smallsetminus Y\big)\right] \cup \left[\bigcup_{n\geq 0} f^n\big(Y \smallsetminus f(X)\big) \cup \bigcap_n f^n(X)\right], \end{aligned}$$

où nous remarquons que $\bigcap f^n(X) = \bigcap f^n(Y)$. La partie de X définie par la deuxième parenthèse est contenue dans Y, et on n'aura pas besoin d'y toucher. Quant à la première parenthèse, $X \smallsetminus Y$ est disjoint de Y, tandis que pour $n \geq 1$, $f^n(X \smallsetminus Y) \subseteq Y$. Nous pouvons donc utiliser f pour envoyer $X \smallsetminus Y$ sur $f(X \smallsetminus Y)$, envoyer $f(X \smallsetminus Y)$ sur $f^2(X \smallsetminus Y)$, et ainsi de suite, ce qui donne la définition suivante:

$$g(x) = \begin{cases} f(x) & \text{si } x \in \bigcup f^n(X \smallsetminus Y) \\ x & \text{sinon} \end{cases}$$

En considérant l'action de g sur X selon la décomposition de X donnée plus haut, il est facile à vérifier que g est une bijection de X avec Y. ■4.1.2

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 4.1.3. 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 4.1.4. *Soit X un ensemble. Alors X est équipotent à un (unique) cardinal si est seulement si X admet un bon ordre.*

DÉMONSTRATION. Exercice. ■4.1.4

THÉORÈME 4.1.5. *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. ■4.1.5

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

EXERCICE 4.1.6. 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 4.1.7. Étant donné un cardinal κ, on note κ^+ le plus petit cardinal strictement plus grand que κ.

EXERCICE 4.1.8. Montrer que L'ordinal de Hartogs associé à un ensemble quelconque X est un cardinal. Montrer que pour un cardinal κ :

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

DÉFINITION 4.1.9. 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 4.1.10. *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 $|\beta| \leq \beta < |\alpha|$. ■4.1.10

DÉFINITION 4.1.11. *(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 4.1.12. *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].

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 $\beta = \omega^\beta$

PROPOSITION 4.1.13. *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, $\kappa = \aleph_\alpha$. ■4.1.13

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 4.2.1. 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 4.2.2. *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 4.2.3. *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 :*

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

DÉMONSTRATION. Exercice. ■4.2.3

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

LEMME 4.2.4. *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. ■4.2.4

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

LEMME 4.2.5. *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 finis 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 4.2.6. *Soit κ un cardinal infini. Alors $\kappa \cdot \kappa = \kappa$.*

DÉMONSTRATION. On le démontre par récurrence sur κ. On définit d'abord une relation d'ordre sur $\kappa \times \kappa$:

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

Il est facile de vérifier que $\preceq$ est une relation d'ordre total sur $\kappa \times \kappa$; pour voir que c'est un bon ordre, soit A un ensemble non vide contenu dans $\kappa \times \kappa$. 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 $(\alpha_0,\beta_0) = \min A$.

Soit $S \subseteq \kappa \times \kappa$ un segment initial propre. Nous avons alors $S = (\kappa \times \kappa)_{<(\alpha,\beta)}$ 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 $|S| \leq |\alpha+1| \cdot |\beta+1| = \mu \times \mu = \mu < \kappa$. 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. ■4.2.6

REMARQUE 4.2.7. Il en découle, à l'aide de l'axiome du choix, que $|X \times X| = |X|$ pour tout ensemble X. En fait la réciproque est vraie aussi: $|X \times X| = |X|$ pour tout X implique AC.

PROPOSITION 4.2.8. *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 peut supposer que $\kappa = \max(\kappa,\lambda)$. Alors

$$\begin{aligned} \kappa = \kappa \cdot 1 \leq \kappa \cdot \lambda \leq \kappa \cdot \kappa = \kappa, \\ \kappa = \kappa + 0 \leq \kappa + \lambda \leq \kappa + \kappa = \kappa \cdot 2 = \max(\kappa, 2) = \kappa. \end{aligned}$$ ■4.2.8

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

DÉFINITION 4.2.9. Soit X et Y deux ensembles. Nous définissons

$$X^Y = \{\text{fonctions } f\colon Y \to X\}.$$

EXERCICE 4.2.10. 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.

À partir de maintenant il convient de supposer l'axiome du choix – 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 4.2.11. Pour des cardinaux κ, λ nous définissons la *puissance cardinale* κ^λ comme le cardinal de l'ensemble des fonctions de λ dans κ (i.e., avec notre notation ambiguë, $\kappa^\lambda = |\kappa^\lambda|$...).

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

EXERCICE 4.2.12. Montrer que pour trois cardinaux κ, λ, μ on a $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$, et $\kappa^\lambda . \kappa^\mu = \kappa^{\lambda+\mu}$.

EXERCICE 4.2.13. 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 4.2.14. 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. Nous remarquons d'abord que pour tout κ nous avons $2^\kappa = |\mathcal{P}(\kappa)| \geq \kappa$ (comment obtenir cette dernière inégalité?) Puis, d'après le Théorème de Cantor (Théorème 1.1.3), $\kappa < 2^\kappa$.

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:

$$\beth_0 = \aleph_0, \qquad \beth_{\alpha+1} = 2^{\beth_\alpha}, \qquad \beth_\alpha = \sup_{\beta<\alpha} \beth_\beta \text{ pour } \alpha \text{ limite.}$$

Y a-t-il des cardinaux qui n'apparaissent pas dans cette énumération?

DÉFINITION 4.2.15. L'*hypothèse du continu* (HC) est l'énoncé $2^{\aleph_0} = \aleph_1$.

L'*hypothèse du continu généralisée* (GCH) est l'énoncé affirmant que $2^\kappa = \kappa^+$ pour tout cardinal infini κ, ou encore, que $\aleph_\alpha = \beth_\alpha$ pour tout α.

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 4.2.16. 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 4.2.17. 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 4.2.18. 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 4.2.19. 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 4.3.1. 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 4.3.2. *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 $|Y| \leq \aleph_0$. ■4.3.2

LEMME 4.3.3. *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 :

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

■4.3.3

REMARQUE 4.3.4. 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.

Démonstration. Supposons que A est dénombrable ; alors $A^{-1} = \{a^{-1}\colon a \in A\}$ 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

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

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

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

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.

---

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 » .

EXERCICE 4.3.5. 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 4.3.6. Montrer que $|\mathbf{Q}| = \aleph_0$, puis montrer que $|\mathbf{R}| = 2^{\aleph_0}$. Pour le second point, on pourra considérer l'application $f\colon \mathbf{R} \to \mathcal{P}(\mathbf{Q})$ définie par

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

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 4.3.7. 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 $f\colon X \to X$.

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

EXERCICE 4.3.9. 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 4.4.1. 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,

$$\mathrm{cf}(\alpha) = \min\{|A|\colon A \subseteq \alpha \text{ et } \sup A = \alpha\}.$$

DÉFINITION 4.4.2. 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 4.4.3.
(i) Montrer que $\mathrm{cf}(\alpha)$ est le plus petit ordinal γ tel qu'il existe une fonction $f\colon \gamma \to \alpha$ 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 4.4.4.
(i) Soit κ un cardinal ; montrer que $\mathrm{cf}(\kappa)$ est le plus petit cardinal μ tel que κ soit la réunion de μ ensembles de cardinal strictement inférieur à κ.

(ii) Montrer que pour un cardinal κ, $\mathrm{cf}(\kappa)$ est le plus petit μ tel qu'il existe $\lambda_i < \kappa$ pour $i < \mu$ avec $\sum_{i<\mu} \lambda_i = \kappa$.

(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 4.4.5. *Tout cardinal successeur infini est régulier.*

DÉMONSTRATION. Fixons un cardinal successeur infini κ^+, et soit $\mu = \mathrm{cf}(\kappa^+)$. Alors il existe $\lambda_i < \kappa^+$, donc $\leq \kappa$, tels que

$$\kappa^+ = \sum_{i<\mu} \lambda_i \leq \sum_{i<\mu} \kappa = \mu\kappa = \max \mu, \kappa,$$

d'où $\mu = \kappa^+$. ■4.4.5

Par contre, 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 4.4.6. 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 *Théorème 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 :

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

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 4.4.7 (König). *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

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

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

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. ■4.4.7

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 4.4.8. 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 4.4.9. *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 4.4.10. *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.

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**].

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

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 5.1.1. 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 $\mathcal{F}_x = \{A \colon x \in A\}$ 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

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

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

DÉFINITION 5.1.3. 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 5.1.4. *Pour toute base de filtre $\mathcal{A}$, il existe un filtre contenant $\mathcal{A}$; le plus petit tel filtre est défini par*

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

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 $\{A \in X \colon A \supseteq F\}$, 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

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

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 5.1.5. 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 5.1.6. *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 5.1.7. 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 5.2.1. 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 5.2.2. 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 5.2.3. 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 5.2.4. *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 5.2.5. *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 5.2.6. *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 5.2.7. *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 5.2.8. *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 5.2.7 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 5.2.9. *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 5.3.1. 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 5.3.2. *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 5.3.3. *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 5.3.2.
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 5.3.5 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 5.3.4. *Pour tout $\alpha < \beta < \omega_1$, $X_\beta \smallsetminus X_\alpha$ est fini.*

Preuve du Lemme 5.3.4.
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 5.3.5. *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 5.3.4) que $X_\gamma \smallsetminus X_\beta$ est fini. Mais pour tout ensemble fini d'ordinaux $\beta_1, \ldots, \beta_n < \alpha$, si on pose $\beta = \max\{\beta_i : i = 1, \ldots, n\}$ 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 5.3.5 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 5.3.5.
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.