Introduction à la Logique Mathématique

Première partie : Théorie des ensembles

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 I au semestre de printemps 2010. Ces notes contiennent sans aucun doute des erreurs, coquilles, approximations, contradictions, assertions non justifiées, etc. Nous encourageons donc nos lecteurs à exercer leur sens critique durant leur lecture, et leur serions reconnaissants de bien vouloir nous signaler tout problème de cette nature qu'ils remarqueraient.

Table des matières

Chapitre 1

Les axiomes de Zermelo-Fraenkel

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

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

Bien sûr, il faut se donner un point de départ ; pour nous, c'est un *univers*,

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

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

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

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

- Si $a \in \mathcal{U}$ on peut former l'enoncé à une variable libre et un paramètre $x = a$.
- On peut aussi formuler l'énoncé à trois variables libres $R(x, y, z)$ défini par
$$\forall t\, (t \in z) \Leftrightarrow (t = x \text{ ou } t = y)\ .$$
- Par exemple, à partir de l'énoncé R ci-dessus, on peut former l'énoncé clos
$$\forall x\, \forall y\ \exists z\ R(x, y, z)\ .$$

Chaque énoncé $R(x_1, \ldots, x_k)$ à exactement k variables libres définit une *relation*, qui est la collection des k-uplets $(a_1, \ldots, a_k)$ d'éléments de $\mathcal{U}$ tels que $R(a_1, \ldots, a_k)$ soit vrai ; rappelons que $R(a_1, \ldots, a_k)$ est l'énoncé clos obtenu en substituant $a_1, \ldots, a_k$ aux variables libres de R.
Certaines relations (éventuellement avec des paramètres $a_1, \ldots, a_k$) sont particulièrement importantes : ce sont les *relations fonctionnelles* ; une relation $R(x, y)$ à exactement deux variables libres est une relation fonctionnelle (à 1 argument) si
$$\forall x\, \forall y\, \forall z\, (R(x, y) \text{ et } R(x, z) \Rightarrow y = z)\ .$$

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

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

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

1. Axiome d'extensionnalité

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

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

2. Axiome de la réunion

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

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

3.Axiome de l'ensemble des parties

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

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

3.Schéma d'axiomes de remplacement

Il s'agit en fait d'une infinité d'axiomes. Le schéma d'axiomes de remplacement nous permet, à partir d'une relation fonctionnelle et d'un ensemble, de former un nouvel ensemble. Par exemple il nous permet de former l'image d'un ensemble par une fonction (en tant qu'*ensemble*, et pas seulement comme une collection).
Formellement, ce schéma dit que si $E(x, y, a_1, \ldots, a_k)$ est un énoncé à paramètres $a_1, \ldots, a_k$ qui définit une relation fonctionelle à 1 variable, et a est

un ensemble, alors on peut considérer l'ensemble b dont les éléments sont les images par la relation fonctionnelle E des éléments de a appartenant au domaine de notre relation fonctionnelle.
Alors le schéma d'axiomes de remplacement consiste en la liste, paramétrée par tous les énoncés $E(x, y, x_1, \ldots, x_k)$ sans paramètres et à au moins deux variables libres des énoncés suivants :

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

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

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

Schéma d'axiomes de compréhension

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

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

Cet énoncé découle du schéma de substitution appliqué à la relation

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

Axiome de l'ensemble vide

Cet axiome dit qu'il existe un ensemble et un seul qui n'a aucun élément. L'unicité est une conséquence directe de l'axiome d'extensionnalité. Pour prouver l'existence d'un tel ensemble, il suffit d'appliquer le schéma de compréhension à un ensemble a de $\mathcal{U}$ et à l'énoncé $x \neq x$: en effet, on obtient qu'il existe un ensemble noté $\emptyset$ tel que

$$\forall x \, x \in \emptyset \Leftrightarrow (x \in a \text{ et } x \neq x)$$

Par conséquent, $\forall x \, x \notin \emptyset$.

Axiome de la paire

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

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

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

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

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

C'est un bon exercice de voir qu'avec nos axiomes on peut former produit et intersection d'une famille d'ensembles.
Il nous reste encore un axiome à énoncer ; avant de lire l'énoncé de cet axiome il faut avoir lu la définition des ordinaux (leur existence ne dépend que des axiomes déjà énoncés).

4.Axiome de l'infini

Cet axiome dit simplement : il existe un ordinal non fini. Ou encore, il existe un ensemble bien ordonné qui n'ait pas de plus grand élément. La signification de cet axiome devrait être claire après le prochain chapitre ; on est obligé de faire attention en l'écrivant et de ne pas se contenter d'un énoncé du type « il existe un ensemble infini »parce que la notion même d'ensemble infini n'est pas claire et peut avoir plusieurs définitions qui ne sont pas équivalentes dans ZF.

Une fois que vous aurez lu le prochain chapitre, vous devriez être capable de voir qu'une façon équivalente d'énoncer cet axiome est : la collection des ordinaux finis est un ensemble ; ou encore : il existe un ordinal limite.

Axiome de fondation

Cet axiome dit que pour tout ensemble non vide x, il existe un ensemble $y \in x$ et tel que $y \cap x = \emptyset$. En particulier l'axiome de fondation interdit l'existence d'ensembles x tels que $x \in x$, ou l'existence de suites $(x_n)_{n\in\omega}$ telles que $x_{n+1} \in x_n$ pour tout n.
Cet axiome, noté AF, n'est *pas* une conséquence des axiomes de (ZF) ; il est courant de se placer dans le cadre axiomatique ZF+AF.
Si on part d'un univers $\mathcal{U}$ satisfaisant les axiomes de ZF, et qu'on définit $V_0 = \emptyset$, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$ [iii] et $V_\alpha = \cup_{\beta<\alpha} V_\beta$ pour α limite, on obtient une collection d'ensembles dont on peut former la réunion (au sens naïf). Notons V cette réunion ; on peut montrer que si $\mathcal{U}$ est un modèle de (ZF) alors V est un modèle de ZF+AF.

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

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

iii. Pour suivre cette discussion, il faut d'abord avoir lu la définition des ordinaux présentée au chapitre suivant.

Chapitre 2

Les ordinaux

2.1 Bons ordres et définition des ordinaux.

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

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

Définition 2.1. Soit X un ensemble. Un *bon ordre* sur X est une relation d'ordre $\leq$ sur X tel que tout sous-ensemble non vide de X a un plus petit élément. On dit que $S \subseteq X$ est un *segment initial* si

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

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

Notons que, dans un ensemble bien ordonné X, tout segment initial différent de X est de la forme S_x pour un unique $x = \min(X \setminus S)$.

L'idée, dans notre optique de comptage, est que pour énumérer un ensemble bien ordonné, on commence au plus petit élément, puis on prend le

plus petit des autres, etc. ; mais s'arrête-t-on un jour ?

L'essentiel de la théorie des ensembles bien ordonnés est fondé sur le résultat suivant :

Proposition 2.2. *Soit $(X, \leq)$ un ensemble bien ordonné et $f\colon X \to X$ une application strictement croissante. Alors pour tout $x \in X$ on a $f(x) \geq x$.*

Preuve.
Supposons qu'il existe $x \in X$ tel que $f(x) < x$, et appelons x_0 le plus petit élément ayant cette propriété. Alors on a, pour tout $x < x_0$, $f(x) \geq x$.
Puisque f est strictement croissante, on en déduit que pour tout $x < x_0$ on a $f(x_0) > x$.
Mais alors $f(x_0) < f(x_0)$, ce qui est absurde. □

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

Proposition 2.3. *Soit $(X, \leq)$ un ensemble bien ordonné, $W \subseteq X$ un segment initial et $f\colon X \to W$ un isomorphisme. Alors $W = X$ et pour tout $x \in X$ on a $f(x) = x$.*
Par conséquent, si deux segments initiaux de X sont isomorphes alors ils sont égaux.

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

Notation. Si X, X' sont deux ensembles bien ordonnés, on note $X \preceq X'$ si X est isomorphe à un segment initial de X', et $X \sim X'$ si X et X' sont isomorphes. On utilisera la notation $X \prec X'$ pour signifier que $X \preceq X'$ et $X \not\sim X'$, autrement dit si X est isomorphe à un segment initial strict de X'.

Remarquons que le théorème 2.3 entraîne que $X \sim X'$ si, et seulement si, $X \preceq X'$ et $X' \preceq X$. On a dit qu'on souhaitait pouvoir enumérer tous les ensembles bien ordonnés ; mais quelle notion de « longueur »utiliser ?

Théorème 2.4. *Soit X, Y deux ensembles bien ordonnés. Alors une et une seule des assertions suivantes est vraie :*

(a) $X \prec Y$;

(b) $Y \prec X$;

(c) $X \sim Y$.

Ce théorème dit qu' une notion de « longueur »possible d'un ensemble bien ordonné est l'ensemble lui-même, où on compare deux longueurs par la relation « être isomorphe à un segment initial ». Restera ensuite à choisir un représentant dans chaque classe d'isomorphisme...

Preuve.
Notons $\tilde{X}$ l'ensemble des segments initiaux de X, ordonné par l'inclusion. On vérifie facilement que c'est un ensemble bien ordonné. Si tout $S \in \tilde{X}$ est isomorphe à un segment initial de Y alors c'est en particulier le cas de X, et la preuve est finie. Sinon, appelons S le plus petit élément qui ne soit pas isomorphe à un segment initial de Y.
Si jamais S a un plus grand élément x, alors il doit exister un isomorphisme f_x de S_x sur un segment initial de Y, et si $f_x(S_x) \neq Y$ alors on voit facilement qu'on peut prolonger f_x en un isomorphisme de S sur un segment initial de Y (en envoyant X sur le plus petit élément de $Y \setminus f_x(S_x)$), ce qui contredit la définition de S ; par conséquent, si S a un plus grand élément alors la preuve est finie.
Il nous reste donc à traiter le cas où S n'a pas de plus grand élément ; pour tout $x \in S$ on note toujours f_x l'unique isomorphisme de S_x sur un segment initial de Y. Alors, pour $x < x' \in S$ on a $f_y \circ f_x^{-1}(f_x(S_x)) = f_y(S_x)$, et comme $f_y \circ f_x^{-1}$ est un isomorphisme entre deux segments initiaux de Y on en déduit que $f_y \circ f_x^{-1}(y) = y$ pour tout $y \in f_x(S_x) = f_y(S_x)$. Autrement dit, f_y et f_x coïncident sur S_x.
Mais alors, comme $S = \bigcup_{x \in S} S_x$, on peut « recoller »toutes ces fonctions pour définir une fonction strictement croissante $f \colon S \to Y$ d'image $\cup_{x \in S} f_x(S_x)$ (en posant $f(y) = f_x(y)$ dès que $y \in S_x$). L'image de f est une union de segments initiaux de Y, et est donc un segment initial de Y, ce qui contredit le choix de S. □

Théorème 2.5. *Soit $\mathcal{W} = \{W_i \colon i \in I\}$ une famille d'ensembles bien ordonnés. Alors il existe $W \in \mathcal{W}$ tel que $W \preceq W'$ pour tout $W' \in \mathcal{W}$.*

Preuve.
Soit $W_0 \in \mathcal{W}$. Si $W_0 \preceq W'$ pour tout $W' \in \mathcal{W}$, il n'y a rien à démontrer. Sinon, l'ensemble $\{x \in W_0 \colon S_x$ est isomorphe à un élément de $\mathcal{W}\}$ est non vide. Appelons w le plus petit élément de cet ensemble, et prenons $W \in \mathcal{W}$ qui soit isomorphe à S_w (vu dans W_0). Pour tout $W' \in \mathcal{W}$, il est

impossible par définition que W' soit isomorphe à un segment initial strict de S_w, par conséquent on a $W \preceq W'$ pour tout $W' \in \mathcal{W}$. □

Maintenant, il faudrait définir rigoureusement les *ordinaux* ; l'idée est qu'on veut compter à partir de 0 jusqu'à l'infini, et au-delà. On pourrait simplement les définir, comme Cantor l'a fait, comme les classes d'isomorphisme de bons ordres ; ci-dessous on va plutôt décrire l'approche de von Neumann, qui peut paraître arbitraire mais a beaucoup d'avantages une fois qu'on l'a comprise. Commençons par essayer d'introduire intuitivement (autant que possible...) cette approche.

L'idée est que les ordinaux doivent permettre de « représenter »les ensembles bien ordonnés, au sens où tout ordinal soit un ensemble bien ordonné et pour tout ensemble bien ordonné il y ait un ordinal unique qui lui soit isomorphe ; c'est cet ordinal-là qui doit représenter la « longueur »d'un ensemble bien ordonné. Admettons que cela soit posssible (et pensons donc intuitivement à un ordinal comme à une classe d'isomorphisme d'ensembles bien ordonnés).

Allons plus loin et notons que si α est un ordinal, alors tout ordinal plus petit que α est isomorphe à un (unique) segment initial de α ; et les segments initiaux stricts de α s'identifient naturellement aux éléments de α.

On a donc envie d'identifier les ordinaux strictement inférieurs à α aux éléments de α, et donc d'effectuer notre choix de représentants de classes d'isomorphisme de bons ordres de telle façon que chaque ordinal α soit *égal* à l'ensemble des ordinaux strictement inférieurs à α.

Ceci impose une contrainte : si $\beta < \alpha$ on doit en même temps identifier les ordinaux strictement inférieurs à β aux éléments de β, ce qui amène à vouloir que l'ensemble des éléments strictement inférieurs à β (c'est-à-dire β) soit contenu dans α. Finalement, on a donc envie que tout élément d'un ordinal soit en fait *inclus* dans cet ordinal.

On n'est toujours pas tout à fait satisfait : si on a une famille d'ordinaux, alors on voudrait pouvoir « compter strictement plus loin »que tous les ordinaux de cette famille, ce qui imposerait que la réunion de notre famille d'ordinaux soit un ordinal. On rajoute cela dans les conditions qu'on demande aux ordinaux.

Voilà, on sait maintenant quelles propriétés attendre d'un ordinal, et on sait même comment effectuer leur construction : en effet, il n'y a pas d'élement plus petit que 0, donc 0 doit être l'ensemble vide. De même, $1 = \{0\} = \{\emptyset\}$, pour tout ordinal fini (i.e tout entier naturel !) on doit avoir $n = \{0, 1, \ldots, n-1\}$, etc.

Passons maintenant à la construction rigoureuse des ordinaux, basée sur

l'idée que le plus petit ordinal est l'ensemble vide, et que *tout ordinal est égal à l'ensemble des ordinaux qui le précèdent.*

Définition 2.6. 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 $\emptyset$, $\{\emptyset, \{\emptyset\}\}$...

Le lemme suivant est une conséquence immédiate de la définition.

Lemme 2.7. *Une réunion d'ensemble transitifs est encore un ensemble transitif ; une intersection d'ensembles transitifs est encore un ensemble transitif.*

Définition 2.8. Un ensemble α est un *ordinal* si α est transitif et strictement bien ordonné par la relation $\in$.

Par exemple, $\emptyset$ est un ordinal ; $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ est un ordinal (exercice !) Même si on n'a pas supposé que l'axiome de fondation était vérifié, les ordinaux se comportent bien relativement à $\in$.

Lemme 2.9. *Pour tout ordinal α, on a $\alpha \notin \alpha$.*

Preuve.
Si α est un ordinal, alors $\in$ munit α d'une structure de bon ordre strict, en particulier pour tout $x \in \alpha$ on doit avoir $x \notin x$. Par conséquent, de $\alpha \in \alpha$ on déduirait que $\alpha \notin \alpha$, ce qui est bien sûr une contradiction. □

Notons également que la définition nous donne tout de suite la propriété suivante.

Proposition 2.10. *L'intersection d'un ensemble d'ordinaux est un ordinal.*

Preuve.
Soit $\{\alpha_i\}_{i \in I}$ un ensemble d'ordinaux, et β son intersection. Alors β est un ensemble transitif comme intersection d'ensembles transitifs, et il est également clair que $\in$ munit encore β d'une structure de bon ordre strict. Par conséquent, β est bien un ordinal. □

Le lemme suivant paraît contre-intuitif, mais est simplement une manifestation du fait qu'on souhaite qu'un ordinal soit égal à l'ensemble des ordinaux qui le précèdent.

Lemme 2.11. *Soit α, β deux ordinaux tels que $\alpha \subseteq \beta$ et $\alpha \neq \beta$. Alors $\alpha \in \beta$.*

Preuve.
Définissons γ comme le plus petit élément de $\beta \setminus \{\alpha\}$. On va montrer que $\gamma = \alpha$; pour cela, par extensionnalité, il nous suffit de montrer que α et γ ont les mêmes éléments.
S'il existe $\delta \in \gamma \setminus \alpha$, alors on doit avoir $\delta \in \beta \setminus \alpha$, par conséquent on a à la fois $\delta \in \gamma$ et $\delta \in \beta \setminus \alpha$, ce qui contredit la définition de γ.
Soit maintenant $\delta \in \alpha$. Alors on a aussi $\delta \in \beta$ par transitivité, et comme $\in$ est un ordre total on a trois possibilités : $\delta \in \gamma$, $\gamma \in \delta$ ou $\delta = \gamma$. La première possibilité est ce qu'on souhaite obtenir ; la deuxième entraînerait que $\gamma \in \delta \in \alpha$ et donc $\gamma \in \alpha$ (puisque α est un ensemble transitif) ce qui contredit le choix de γ. La troisième possibilité nous donne encore $\gamma \in \alpha$ et doit donc également être exclue. Par suite, tout élément de α appartient à γ, et on a enfin fini de montrer que $\alpha = \gamma$. □

Poursuivons nos investigations ; le premier point de la proposition suivante nous dit que, étant donnés deux ordinaux distincts, l'un est nécessairement un élément de l'autre - dans le monde des ordinaux, c'est une manifestation du théorème qu'on a vu précédemment et qui dit que, étant donné deux ensembles bien ordonnés, il y en a nécessairement un qui est isomorphe à un segment initial de l'autre.

Proposition 2.12.
1. *Si α, β sont deux ordinaux, alors on a soit $\alpha = \beta$, soit $\alpha \in \beta$, soit $\beta \in \alpha$ (et les trois propriétés précédentes sont exclusives)*
2. *Deux ordinaux isomorphes sont égaux.*
3. *Tout ensemble d'ordinaux est (strictement) bien ordonné par $\in$.*
4. *L'union d'un ensemble d'ordinaux est encore un ordinal.*

Preuve.

1. On a vu que $\alpha \cap \beta$ est un ordinal ; si $\alpha \cap \beta \neq \alpha$ alors comme $\alpha \cap \beta \subseteq \alpha$ on a $\alpha \cap \beta \in \alpha$ d'après le lemme précédent. De même si $\alpha \cap \beta \neq \beta$ alors on a $\alpha \cap \beta \in \beta$; Par conséquent, si $\alpha \cap \beta \neq \alpha$ et $\alpha \cap \beta \neq \beta$ alors on doit avoir $\alpha \cap \beta \in \alpha \cap \beta$, ce qui est impossible puisque c'est un ordinal.
2. Soit α, β deux ordinaux. Si $\alpha \in \beta$ alors α est un segment initial strict de β (puisque α est un ensemble transitif) et donc α ne peut pas être isomorphe à β. Il en va de même si $\beta \in \alpha$. Donc, d'après le point précédent, α et β ne peuvent être isomorphes que si $\alpha = \beta$.

3. Soit E un ensemble d'ordinaux ; d'après le point (1), $\in$ est un ordre total (strict) sur E. Reste à montrer que c'est un bon ordre. Soit donc A une partie de E non vide, et $\alpha \in A$. Si $\alpha \cap A$ est non vide, on vérifie que le plus petit élément γ de $\alpha \cap A$ (qui existe puisque $(\alpha, \in)$ est bien ordonné) est aussi le plus petit élément de E : en effet, si $\beta \in A \cap \alpha$ alors on doit avoir $\gamma \subseteq \beta$ par définition de γ, et si $\beta \in A \setminus \alpha$ alors on doit avoir (toujours par le point (1)) $\alpha \subseteq \beta$ et donc $\gamma \in \beta$.
   Si maintenant $\alpha \cap A = \emptyset$, alors pour tout $\beta \in A$ on ne peut avoir $\beta \in \alpha$, d'où $\alpha \subseteq \beta$ et on voit que dans ce cas α est le plus petit élément de A.
4. Soit E un ensemble d'ordinaux, et α la réunion des éléments de E. Puisque α est un ensemble dont les éléments sont des ordinaux, le point précédent nous dit que α est strictement bien ordonné par $\in$. De plus on sait que α est transitif, par conséquent α est un ordinal. □

Comme on l'a dit plus haut, on veut compter jusqu'à l'infini, et au delà ; en particulier pour tout ordinal il doit exister des ordinaux plus grands que lui, et en fait tout ordinal doit avoir un successeur [i]. La définition formelle de ce successeur est présentée ci-dessous.

Proposition 2.13. *Pour tout ordinal α, $S(\alpha) := \alpha \cup \{\alpha\}$ est encore un ordinal ; de plus pour tout ordinal γ tel que $\alpha \subseteq \gamma \subseteq S(\alpha)$ on doit avoir $\alpha = \gamma$ ou $S(\alpha) = \gamma$.*
On appelle $S(\alpha)$ l'ordinal successeur *de α.*

Preuve.
Il est facile de vérifier que $S(\alpha)$ est transitif, et bien ordonné par $\in$ (et on le laisse comme exercice pour vérifier que vous avez bien compris ces notions). C'est donc bien un ordinal. Reste à vérifier que $S(\alpha)$ a bien la propriété décrite ci-dessus. Pour cela, fixons un ordinal γ tel que $\alpha \subsetneq \gamma \subseteq S(\alpha)$. Alors, on doit avoir $\alpha \in \gamma$ d'après le lemme 2.11, et par conséquent $S(\alpha) \subseteq \gamma$. □

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 2.14. *Il n'existe pas d'ensemble de tous les ordinaux (autrement dit, la collection formée par les ordinaux est un ensemble au sens intuitif mais pas au sens mathématique).*

Preuve.
Soit $\mathcal{O}$ un ensemble d'ordinaux ; alors la réunion des éléments de $\mathcal{O}$ est encore

i. ce qui fait un point commun entre les ordinaux et les hommes politiques (j'espère).

un ordinal d'après une proposition précédente. Notons-le α. Par définition, $\gamma \in \alpha \Leftrightarrow \exists \beta \in O \gamma \in \beta$. Appliquons cette équivalence à $S(\alpha)$: si jamais $S(\alpha)$ appartenait à 0, comme on a $\alpha \in S(\alpha)$ on aurait $\alpha \in \alpha$, ce qui est impossible puisque α est un ordinal. □

Notre travail précédent sur les bons ordres, a pour conséquence que, comme on le souhaitait, les ordinaux fournissent un « modèle »pour tous les bons ordres.

Théorème 2.15. *Tout ensemble bien ordonné est isomorphe à un ordinal unique.*

Preuve.
Soit $(X, <)$ un ensemble bien ordonné. Si X n'est pas isomorphe à un ordinal, notons S le plus grand segment initial de X (pour l'ordre déjà utilisé sur les segments initiaux) qui soit isomorphe à un ordinal α. Cet ensemble est non vide puisque $\emptyset$ est un segment initial de X et est un ordinal. Appelons x le minimum de $X \setminus S$. Alors $S \cup \{x\}$ est isomorphe à $S(\alpha)$, ce qui contredit la définition de S.
L'unicité est immédiate puisqu'on a vu précédemment que deux ordinaux isomorphes sont égaux. □

Pour éviter les confusions, on notera la plupart du temps dans la suite $\alpha < \beta$ pour dire que l'ordinal α est un élément de l'ordinal β ; autrement dit, le bon ordre associé à α est un segment initial du bon ordre associé à β. De même, on écrira $\alpha \leq \beta$ pour signifier que $\alpha \in \beta$ ou $\alpha = \beta$, ce qui est équivalent à dire que $\alpha \subseteq \beta$.

On notera maintenant ON la collection des ordinaux. Deux points sont particulièrement à retenir sur cette collection et sur l'ordre $<$.

Proposition 2.16. *Tout ensemble non vide d'ordinaux E a un plus petit élément, qui est égal à l'intersection des éléments de E.*

Preuve.
On a déjà vu que $\in$ est un bon ordre strict sur E, ce qui montre que E a un plus petit élément α. Reste à voir que α est l'intersection de tous les éléments de E : pour cela, appelons cette intersection β et notons que, puisque $\alpha \in E$, on a tout de suite $\beta \subseteq \alpha$. Réciproquement, on vient de voir que $\alpha \subseteq \delta$ pour tout $\delta \in E$, et par définition de l'intersection ceci impose $\alpha \subseteq \beta$. Par extensionnalité, on a donc bien $\alpha = \beta$. □

Par contre, un ensemble non vide d'ordinaux n'a pas en général de plus grand élément (pensez aux ordinaux finis !) ; la proposition suivante est donc le meilleur résultat que l'on puisse espérer.

Proposition 2.17. *Tout ensemble E d'ordinaux a une borne supérieure, qui est la réunion des éléments de E.*

Preuve.
Soit E un ensemble d'ordinaux, et α sa réunion. Alors α est un ordinal, et pour tout $\beta \in E$ on a $\beta \subseteq \alpha$; donc α est un majorant de E. Si maintenant $\gamma \in \alpha$, alors on a (par définition de la réunion) $\gamma \in \beta$ pour un certain $\beta \in E$. Rappelons que ceci signifie $\gamma < \beta$, et donc, puisque $\beta \in E$, γ n'est pas un majorant de E. Ceci prouve que α est la borne supérieure de E. □

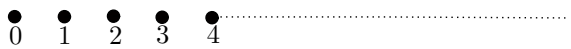
Introduisons un peu de terminologie.

Définition 2.18. Un ordinal α est *successeur* s'il existe un ordinal β tel que $\alpha = S(\beta)$; sinon 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 2.19. Montrer qu'un ordinal β est limite si, et seulement si, $\beta = \sup\{\eta \colon \eta < \beta\}$.

Définition 2.20. Un ordinal α est dit *fini* si tout ordinal tel que $0 < \beta \leq \alpha$ est successeur.
On notera ω le plus petit ordinal infini, qui est aussi le plus petit ordinal limite. L'existence d'un tel ordinal est un axiome de ZF.



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 $\mathbb{N}$.

Notons que l'axiome qui justifie l'existence de ω, l'axiome de l'infini, est aussi équivalent à l'assertion selon laquelle les ordinaux finis forment un ensemble : en effet, si c'est le cas cet ensemble d'ordinaux doit avoir une borne supérieure, qui est alors un ordinal limite puisque l'ensemble des ordinaux finis n'a pas de plus grand élément.

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

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

2.2 Récurrence transfinie et arithmétique des ordinaux.

Vous êtes habitué(e)s à utiliser des démonstrations par récurrence pour montrer, par exemple, que tous les entiers satisfont une certaine propriété ; le principe de la démonstration par récurrence est de dire : si une propriété (P) est telle que pour tout entier naturel n

$$(\forall k < n \; P(k)) \Rightarrow P(n)$$

alors P est vraie pour tout n (notons que l'hypothèse ci-dessus implique en particulier que $P(0)$ est vraie !). Ce principe s'applique dans tout ensemble bien ordonné (à vous d'en faire une démonstration, ce qui ne devrait pas être trop difficile) et on obtient le résultat suivant :

Théorème 2.21. *(Démonstration par récurrence transfinie)*
Soit P une propriété[ii] *des ordinaux telle que pour tout ordinal α on ait*

$$(\forall \beta < \alpha \; P(\beta)) \Rightarrow P(\alpha) \; .$$

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

Alors $P(\alpha)$ est vraie pour tout α.

Bien sûr, cette propriété pourrait s'énoncer dans tout ensemble bien ordonné : si A est un ensemble bien ordonné, et P est une propriété telle que $\forall a \in A\,((\forall a' < a\ P(a')) \Rightarrow P(a))$, alors $P(a)$ est vraie pour tout $a \in A$.

On sait maintenant, au moins en théorie, comment démontrer des énoncés par récurrence transfinie ; il est aussi courant en analyse et en combinatoire infinie qu'on soit amené à *construire* un objet par récurrence transfinie ; c'est une construction facile à comprendre mais à l'énoncé assez aride.

Théorème 2.22. *Soit $(X, \leq)$ un ensemble bien ordonné, Y un ensemble, et $\mathcal{F}$ l'ensemble de toutes les fonctions dont le domaine est un segment initial de X et dont l'image est contenue dans Y. Pour toute fonction $G\colon \mathcal{F} \to Y$, il existe une unique fonction $f\colon X \to Y$ telle que l'on ait, pour tout $x \in X$,*

$$f(x) = G(f_{|_{S_x}}) \ .$$

On n'utilise jamais cet énoncé sous cette forme très abstraite ; mais on utlise fréquemment ce principe pour construire des objets. L'idée est que construire un objet par récurrence transfinie (en ξ étapes, pour ξ un certain ordinal), c'est dire ce qu'on fait au rang 0, puis donner une procédure pour passer de l'étape α à l'étape $\alpha+1$, et enfin donner une procédure pour passer aux ordinaux limites, jusqu'à ce qu'on atteigne ξ. C'est le « cas limite »qui est nouveau par rapport au schéma de récurrence classique.
Plutôt que de donner une preuve du théorème de construction par récurrence transfinie, donnons un exemple de preuve rédigée en utilisant un objet construit par récurrence transfinie.

Proposition 2.23. *Soit $(X, <)$ un ensemble bien ordonné, et $A \subseteq X$ un sous-ensemble non vide. Alors $(A, <)$ est encore bien ordonné, et est isomorphe à un segment initial de X.*

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

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

- On pose $f(\min(A) = \min(X)$.
- Si a est le successeur (dans A) de a', alors on définit $f(a)$ comme le successeur (dans X) de $f(a')$.

– Si a est limite (dans A) alors on pose $f(a) = \sup\{f(a') : a' \in A \text{ et } a' < a$.

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

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

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

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

Ceci achève la démonstration [iii]. □

Présentons un autre exemple.

Exemple : la dérivation de Hausdorff.

Soit $(X, \leq)$ un ensemble ordonné, et $\sim$ une relation d'équivalence compatible avec $\leq$ (c'est-à-dire, telle que $\leq$ passe au quotient par $\sim$). Alors on peut définir une nouvelle relation, notée $D(\sim)$, en posant

$$xD(\sim)y \Leftrightarrow (\text{ il existe un nombre fini de } \sim\text{-classes entre } x \text{ et } y) \ .$$

Cette relation est à nouveau une relation d'équivalence, qui étend $\sim$ et est compatible avec $\leq$.

Soit maintenant ξ un ordinal quelconque. Pour $\alpha < \xi$, on définit une relation d'équivalence $\sim_\alpha$ compatible avec $\leq$ par récurrence transfinie, en respectant les trois points suivants :

(a) $(x \sim_0 y) \Leftrightarrow (x = y)$
(b) Si α est le successeur de β, alors $\sim_\alpha = D(\sim_\beta)$.

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

(c) Si $\alpha = \sup(\{\beta : \beta < \alpha\})$ alors $\sim_\alpha = \bigcup_{\beta<\alpha} \sim_\beta$.

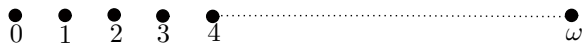
Intuitivement, on a « épluché X ξ fois » : on a commencé par identifier tous les points tels que $[x, y]$ est fini et formé ainsi un nouvel ensemble ordonné, auquel on a appliqué la même construction, et on a répété le procédé pendant ξ étapes.
Par exemple, si on applique cette construction à $\mathbb{N}$ muni de son ordre usuel, on a $\sim_1 = X \times X$; par contre, si on l'applique à $\mathbb{Q}$ muni de son ordre usuel, on a $\sim_1 = \sim_0$ et donc $\sim_\alpha = \sim_0$ pour tout ordinal α.

On est amené à se poser un certain nombre de questions : est-ce qu'on peut continuer à éplucher X indéfiniment sans jamais s'arrêter ? Au contraire, est-ce que X est « épluchable », autrement dit ne reste-il plus rien au bout d'un nombre assez grand d'étapes ? Ou tombe-t-on sur un noyau, c'est-à-dire est-ce que $\sim_n$ arrête de grossir au bout d'un moment ? Nous reviendrons sur ces questions après l'introduction des cardinaux.

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 2.24. (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 :

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



L'ordinal $\omega + 1$

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 !

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 2.25. (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 :

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

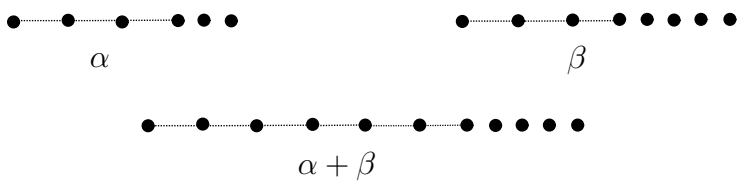
Cette fois on a $2.\omega = \omega$; l'idée de la multiplication ordinale est que "faire le produit de α par β, c'est mettre bout à bout β copies de α". Le dessin suivant essaie de justifier graphiquement l'égalité $2.\omega = \omega$.

$2.\omega = \omega$

Exercice 2.26. 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 2.27.
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).



La somme de deux ordinaux

Un exercice pour vous entraîner aux démonstrations par récurrence transfinie :

Exercice 2.28. 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 2.29. Soit α un ordinal. On pose $\alpha^0 = 1$, puis on définit, par récurrence transfinie sur $\beta \in ON$, α^β en posant :

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

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

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

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

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

Notes bibliographiques. Ce chapitre reprend presque verbatim, en développant la définition des ordinaux, le début du premier chapitre du cours de M2 « théorie descriptive des groupes ». 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

Cardinaux et axiome du choix.

3.1 Définition des cardinaux

On a vu comment énumérer des ensembles bien ordonnés ; mais un ensemble infini peut (doit ?) admettre des bons ordres non isomorphes : c'est par exemple le cas de $\mathbb{N}$.
Cela n'empêche pas d'associer à un ensemble bien ordonnable (c'est-à-dire, un ensemble qu'on peut munir d'un bon ordre) un certain nombre ordinal uniquement déterminé : le plus petit ordinal α tel qu'il existe un bon ordre $<$ sur X avec $(X, <)$ isomorphe à α. Cela permettrait de développer une théorie satisfaisante des cardinaux des ensembles bien ordonnables ; mais comment faire si on a sous la main un ensemble X qui ne nous est pas fourni avec une structure de bon ordre ? La solution fournie par l'axiome de Zermelo est de dire : autorisons-nous à munir tout ensemble d'un bon ordre. Dans ce cas, on saura définir le cardinal d'un ensemble en utilisant des ordinaux, comme expliqué ci-dessus.

A première vue, l'axiome de Zermelo peut paraître excessif ; essayons de nous en passer. On peut définir le fait que X et Y ont « le même nombre d'éléments »sans utiliser de bon ordre, comme le montre la définition suivante.

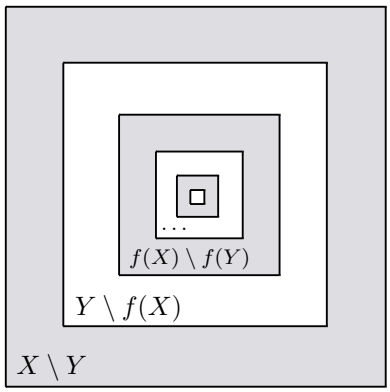
Définition 3.1. On dit que X a un cardinal inférieur à Y, et on note $|X| \leq |Y|$, s'il existe une injection de X dans Y, et on dit que X et Y ont même cardinal, ou sont équipotents (noté $|X| = |Y|$), s'il existe une bijection de X sur Y.

Ainsi, on cherche à étendre les notions intuitives de comptage, qui marchent pour les ensembles finis [i], à tous les ensembles. Déjà, il faut s'assurer que ces notions sont bien compatibles entre elles ; au début, tout se passe bien.

i. C'est-à-dire : équipotents à un ordinal fini, également connu sous le nom d'*entier naturel*

Théorème 3.2. (Schröder-Bernstein)
Si $|X| \leq |Y|$ *et* $|Y| \leq |X|$ *alors* $|Y| = |X|$.

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



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

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

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

Autrement dit, s'il existe une injection de X dans Y et une injection de Y dans X alors il existe une bijection de X sur Y, et nos notations sont bien cohérentes et définissent un quasi-ordre sur l'univers des ensembles. Notre

préoccupation maintenant est de savoir si deux ensembles sont nécessairement comparables pour ce quasi-ordre.

On peut déjà subodorer un problème : si X, Y sont deux ensembles, Y est bien ordonnable et $|X| \leq |Y|$, alors il existe une injection de X dans Y, qu'on peut utiliser pour munir X d'un bon ordre. Autrement dit, si les cardinalités de deux ensembles sont toujours comparables, et s'il existe un ensemble X qui ne peut pas être muni d'un bon ordre, alors on doit avoir $|X| > |Y|$ pour tout ensemble bien ordonnable, et en particulier pour tout ordinal. Donc tout ordinal s'injecte dans X ; mais alors on pourrait utiliser les axiomes de la théorie des ensembles pour prouver que les ordinaux forment un ensemble, et on sait que cela n'est pas possible. Par conséquent, avec nos méthodes, on aura besoin de l'axiome de Zermelo pour avoir une notion satisfaisante de cardinal d'un ensemble.

La proposition qui suit formalise l'objection exposée ci-dessus.

Proposition 3.3. *Pour tout ensemble X il existe un plus petit ordinal non équipotent à une partie de X.*

Preuve.
Soit X un ensemble. Alors on peut considérer l'ensemble

$$Y = \{R \subseteq X^2 \colon R \text{ est un bon ordre sur} X\} \ .$$

C'est un ensemble (la propriété « être un bon ordre »s'exprime par une formule, et on peut donc appliquer le schéma d'axiomes de compréhension dans $\mathcal{P}(X^2)$ à cette propriété), et pour tout $R \in Y$ il existe un unique ordinal α tel que α soit isomorphe (en tant qu'ensemble bien ordonné) à R. Mais alors, puisqu'un bon ordre est isomorphe à un ordinal unique, le schéma d'axiome de remplacement nous permet d'affirmer que

$$\{\alpha \in ON \colon \alpha \text{ est équipotent à une partie de } X\}$$

est un ensemble. Puisqu'on sait que ON n'est pas un ensemble, cela prouve qu'il existe des ordinaux non équipotents à une partie de X. Si l'on considère un tel ordinal α, alors il existe un plus petit ordinal $\beta \leq \alpha$ qui n'est pas équipotent à une partie de X (parce que $S(\alpha)$ est bien ordonné par $<$) ; on vérifie, puisque l'ordre $\leq$ est total, que α est le plus petit ordinal qui n'est équipotent à une partie de X. □

Les discussions précédentes servaient, entre autres, de prologue à la définition suivante.

Définition 3.4. Soit α un ordinal. On dit que α est un *cardinal* si aucun ordinal strictement inférieur à α n'est équipotent à α.

Il est alors immédiat que tout ensemble bien ordonnable X est équipotent à un unique cardinal noté $|X|$ et que de plus si α est un ordinal alors $|\alpha| \leq \alpha$. Par exemple, tous les ordinaux finis sont des cardinaux, ainsi que ω ; 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.

Notons également que par définition deux cardinaux distincts ne peuvent pas être équipotents.

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

Proposition 3.5. *Etant donné un ensemble X, le plus petit ordinal non équipotent à une partie de X est en fait un cardinal, et on l'appelle le* cardinal de Hartogs *de X.*

Preuve.
Soit X un ensemble, et α le plus petit ordinal non équipotent à une partie de X. Par définition de α, tout ordinal $\beta < \alpha$ doit être équipotent à une partie de X, par conséquent un tel β ne peut pas être équipotent à α, ce qui prouve que α est bien un cardinal. □

Définition 3.6. Etant donné un cardinal κ, on note κ^+ le plus petit cardinal qui n'est pas équipotent à une partie de κ.

Définition 3.7. Si κ 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.

Définition 3.8. *(Alephs)*
On définit par récurrence transfinie $\aleph_\alpha$, pour tout ordinal α, en posant $\aleph_0 = \omega$ puis

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

On voit à partir de la définition que, si $\alpha < \beta$ sont deux ordinaux, alors $\aleph_\alpha < \aleph_\beta$. Par récurrence transfinie, on peut également vérifier la propriété suivante.

Proposition 3.9. *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 $\alpha = \aleph_\alpha$ [ii].

Proposition 3.10. *Pour tout ordinal α, $\aleph_\alpha$ est un cardinal.*

Preuve.
Raisonnons par récurrence transfinie ; $\aleph_0$ est bien un cardinal, puisque tous les ordinaux $< \aleph_0$ sont finis. Supposons maintenant que α soit un ordinal tel que, pour tout $\beta < \alpha$, $\aleph_\beta$ soit un cardinal. On doit considérer deux cas :
- α est successeur, c'est-à-dire $\alpha = \beta + 1$ pour un certain ordinal β. Alors on sait que $\aleph_\beta$ est un cardinal, et que $\aleph_\alpha = \aleph_\beta^+$, qui est un cardinal par définition.
- α est limite. Il nous faut maintenant prouver que $\aleph_\alpha$ est un cardinal. Raisonnons par l'absurde et supposons que $\aleph_\alpha$ soit équipotent à un ordinal $\gamma < \aleph_\alpha$. Puisque $\gamma < \aleph_\alpha = \bigcup_{\beta<\alpha} \aleph_\beta$, on voit que $\gamma < \aleph_\beta$ pour un certain $\beta < \alpha$. En particulier, $|\gamma| \leq \aleph_\beta$; comme $\beta < \alpha$ on a aussi $|\aleph_\beta| \leq |\aleph_\alpha| = |\gamma|$, et alors le théorème de Schröder-Bernstein nous dit que γ et $\aleph_\beta$ sont équipotents ; ceci est impossible puisque $\gamma < \aleph_\beta$ et que notre hypothèse de récurrence affirme que $\aleph_\beta$ est un cardinal. □

Il est maintenant naturel de se demander si tous les les cardinaux infinis sont de cette forme ; la proposition suivante affirme que c'est bien le cas.

Proposition 3.11. *Tout cardinal infini est de la forme $\aleph_\alpha$ pour un unique ordinal α.*

Preuve.
Si ce n'est pas le cas, il existe un plus petit cardinal infini κ qui ne s'écrive pas sous la forme $\aleph_\alpha$; bien sûr, $\kappa > \aleph_0$. Si $\kappa = \lambda^+$ pour un certain cardinal λ, alors puisqu'il existe β tel que $\lambda = \aleph_\beta$ on obtient que $\kappa = \aleph_{\beta+1}$, ce qui est impossible.
Donc κ doit être un cardinal limite, c'est-à-dire qu'on doit avoir

$$\kappa = \bigcup_{\{\lambda\colon \lambda \text{ cardinal infini} < \kappa\}} \lambda \ .$$

ii. 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 $\beta = \omega^\beta$

Pour tout cardinal $\lambda < \kappa$, on a un unique α_λ tel que $\lambda = \aleph_{\alpha_\lambda}$; cette famille est strictement croissante, et si on pose $\alpha = \sup\{\alpha_\lambda\}$ alors α est limite, et donc par définition on a

$$\aleph_\alpha = \bigcup_{\{\lambda \text{ cardinal infini} < \kappa\}} \aleph_{\alpha_\lambda} = \bigcup_{\lambda < \kappa} \lambda = \kappa \ .$$

Dans les deux cas, on obtient donc une contradiction, ce qui prouve que tout cardinal est de la forme $\aleph_\alpha$ pour un (unique) ordinal α. □

Pour définir une notion satisfaisante du cardinal d'un ensemble, on a eu besoin d'utiliser l'axiome de Zermelo. Celui-ci réapparaît quand on essaie de traiter l'arithmétique des cardinaux, mais sous une forme différente. Il paraît donc raisonnable de faire une pause dans notre exposition et de nous arrêter sur cet axiome, connu généralement sous le nom d'*axiome du choix*. Son énoncé intuitif, dans sa version la plus connue, est : « si on me donne une famille d'ensembles non vides, alors je peux choisir simultanément un élement dans chaque ensemble de cette famille ».

3.2 L'axiome du choix

Avant de citer trois énoncés équivalents de l'axiome du choix , rappelons qu'un ensemble ordonné $(X, \leq)$ est *inductif* si tout sous-ensemble totalement ordonné admet un majorant. Nous dirons aussi qu'un ensemble X admet une *fonction de choix* s'il existe une fonction $f\colon \mathcal{P}(X) \to X$ telle que pour toute partie $A \subseteq X$ non vide on ait $f(A) \in A$.

Définition 3.12. On introduit les énoncés suivants :

1. (Axiome du choix) Tout ensemble X admet une fonction de choix.
2. (Lemme de Zorn) Tout ensemble ordonné inductif non vide a au moins un élément maximal.
3. (Lemme de Zermelo) Tout ensemble peut être bien ordonné.

Ces trois énoncés sont équivalents. Le premier d'entre eux est l'enoncé « historique »de l'axiome du choix ; sous cette forme il a été introduit par Zermelo en 1904. Cet axiome était implicitement utilisé par de très nombreux mathématiciens du dix-neuvième siècle et paraît plutôt « naturel ». il est plus difficile de se faire une idée intuitive du second énoncé, qui est communément utilisé en analyse comme alternative à l'utilisation de la théorie des ordinaux et de la récurrence transfinie. Le dernier énoncé paraît, lui, assez arbitraire, et dit qu'en fait on peut ramener les raisonnements de théorie

des ensembles à des raisonnements sur les ordinaux. Un résumé fameux, mais apocryphe [iii] : « il est clair que l'axiome du choix est vrai et que l'axiome de Zermelo est faux ; quant au théorème de Zorn, qui sait ? »

Preuve que les trois énoncés ci-dessus sont équivalents.
Toutes les implications entre les axiomes ci-dessus sont instructives à démontrer, et c'est un exercice vivement recommandé ; ici on va se contenter d'expliquer rapidement pourquoi (Zermelo) implique (Choix), (Choix) implique (Zorn) et (Zorn) implique (Zermelo).
(Zermelo) ⇒ (Choix) :
C'est l'implication la plus facile des trois : en effet, si $(X, \leq)$ est bien ordonné alors on peut obtenir une fonction de choix sur $\mathcal{P}(X)$ en posant simplement $f(A) = \min(A)$.
(Choix) ⇒ (Zorn) :
Soit $(X, \leq)$ un ensemble ordonné inductif, dont on suppose qu'il n'a pas d'élément maximal. Fixons une fonction de choix φ sur X ; pour tout ensemble totalement ordonné $M \subset X$ il doit exister un majorant strict de M, autrement dit l'ensemble des majorants stricts de M est non vide. En appliquant φ à cet ensemble, on obtient une fonction ψ qui associe à tout sous-ensemble totalement ordonné M de X un majorant strict $\psi(M)$ de M.
Soit maintenant κ un ordinal non équipotent à une partie de X. Soit $x \in X$; par récurrence transfinie, on peut construire une suite indexée par κ d'éléments de X en posant, pour tout $\alpha < \kappa$:

(a) $x_0 = x$;

(b) $x_{\alpha+1} = \varphi(\{y \in X : y > x_\alpha\})$;

(c) $x_\alpha = \psi\big(\{x_\beta : \beta < \alpha\}\big)$ si $\alpha = \sup\{\beta : \beta < \alpha\}$

La suite qu'on vient de construire nous donne une injection de κ dans X, ce qui est impossible par définition de κ.
(Zorn) ⇒ (Zermelo) :
Introduisons l'ensemble

$$\mathcal{A} = \{(A, \leq) : A \subseteq X \text{ et } (A, \leq) \text{ est bien ordonné}\}$$

On peut munir $\mathcal{A}$ d'une structure d'ordre en posant $(A, \leq_A) \preceq (B, \leq_B)$ si, et seulement si, $A \subseteq B$ et $\leq_B$ étend $\leq_A$.
Alors on peut vérifier que $(\mathcal{P}(X), \preceq)$ est un ensemble ordonné inductif non vide, qui a par conséquent un élément maximal $(A, \leq)$. Reste à remarquer que la maximalité de $(A, \leq)$ a pour conséquence que $A = X$. □

iii. Wikipedia l'attribue à un certain Jerry Bona.

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 $\mathbb{R}$. 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 $\mathbb{R}$ 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[iv] 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 3.13. 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 qu'un produit dénombrable d'ensembles non vides est non vide (ou, de manière équivalente, qu'on peut choisir de manière simultanée un point dans chaque élement 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 $\mathbb{R}$ dans $\mathbb{R}$ (par les suites/image inverse d'un fermé est fermé) sont équivalentes requiert l'axiome du choix dénombrable... De même on a besoin d'une forme d'axiome du choix pour justifier qu'une union dénombrable d'ensembles dénombrables est dénombrable, comme le montre l'exercice suivant.

Exercice 3.14. Montrer que l'axiome du choix dénombrable entraîne que toute réunion dénombrable d'ensembles dénombrables est dénombrable (rap-

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

pelons 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[v].

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* $(x_n)_{n\in\mathbb{N}}$ tel que pour tout n on ait $P(\{x_1, \ldots, x_n\})$.

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

Définition 3.15. 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 aRb. Alors il existe une suite $(x_n)_{n\in\mathbb{N}}$ 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.

3.3 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* $X \sqcup Y$ par

$$X \sqcup Y = \{(x, 0) \colon x \in X\} \cup \{(y, 1) \colon y \in Y\} \ .$$

v. et ce fait est indépendant de ZF.

Définition 3.16. Soit κ, λ deux cardinaux. Alors on définit $\kappa + \lambda$ comme le cardinal de $\kappa \sqcup \lambda$, et $\kappa.\lambda$ comme le cardinal de $\kappa \times \lambda$.

A titre de remarque, notons qu'on n'a pas besoin de l'axiome du choix pour cette définition : en effet, dès que X, Y sont bien ordonnables il en va de même de $X \sqcup Y$ et de $X \times Y$ (rappelez-vous la somme et le produit d'ordinaux !), et ces ensembles ont donc bien un cardinal. Par contre, on peut se demander si on n'a pas un problème de compatibilité dans nos définitions : si X, X' (resp. Y, Y') sont équipotents, il n'est pas forcément immédiat que $X \sqcup Y$ et $X' \sqcup Y'$ le sont également, et le même problème semble pouvoir arriver avec $X \times Y$ et $X' \times Y'$. Rassurons-nous tout de suite.

Proposition 3.17. *Soit X, X' (resp. Y, Y') deux ensembles équipotents. Alors $X \sqcup Y$ et $X' \sqcup Y'$ sont équipotents, et il en va de même de $X \times Y$ et $X' \times Y'$.*

Preuve.
Soit $f : X \to X'$ et $g : Y \to Y'$ deux bijections. Alors on peut définir une fonction $F : X \sqcup Y \to X' \sqcup Y'$ 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 bijection est immédiate. De même, on peut définir une bijection $G : X \times Y \to X' \times Y'$ en posant $G(x,y) = (f(x), g(y))$ et là encore il est immédiat qu'il s'agit bien d'une bijection. □

Notons que la somme et le produit de cardinaux sont des opérations commutatives et associatives (exercice !) ; attention tout de même au fait que la sommme/produit de deux cardinaux diffèrent selon qu'on les considère comme des cardinaux ou comme des ordinaux [vi]...

L'addition et la multiplication de deux cardinaux sont simples à comprendre, comme le montre le théorème suivant.

Théorème 3.18.
Soit κ,λ deux cardinaux infinis. Alors on a $\kappa + \lambda = \kappa.\lambda = \max(\kappa, \lambda)$

Preuve.
Commençons par montrer que pour tout ordinal α on a $\aleph_\alpha.\aleph_\alpha \leq \aleph_\alpha$. Sans surprise, on va raisonner par récurrence transfinie sur α.

vi. Notons tout de même que le cardinal de $\alpha + \beta$ est égal à $|\alpha| + |\beta|$, et que le cardinal de $\alpha.\beta$ est égal à $|\alpha|.|\beta|$.

On définit une relation de bon ordre sur les paires d'ordinaux, en posant pour deux paires d'ordinaux $(\alpha,\beta),(\alpha',\beta')$:

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

Il est facile de vérifier que $\preceq$ est une relation d'ordre total sur $ON\times ON$; pour voir que c'est un bon ordre, soit A un ensemble non vide contenu dans $ON\times ON$. On commence par noter que l'ensemble des ordinaux qui s'écrivent γ sous la forme $\max(\alpha,\beta)$ pour un certain $(\alpha,\beta)\in A$ est non vide, et a donc un plus petit élément γ_0. Alors l'ensemble des ordinaux α pour lesquels il existe un ordinal β satisfaisant à la fois $(\alpha,\beta)\in A$ et $\max(\alpha,\beta)=\gamma_0$ est lui aussi non vide, et on appelle son plus petit élément α_0. Enfin, l'ensemble des ordinaux β tels que $(\alpha_0,\beta)\in A$ et $\max(\alpha_0,\beta)=\gamma_0$ est lui aussi non vide ; appelons β_0 son plus petit élément, et notons que (α_0,β_0) est le plus petit élément de A pour $\preceq$.

Pour tout ordinal limite α, $(\alpha\times\alpha,\preceq)$ n'a pas de plus grand élément : c'est donc un ordinal limite. Comme la définition impose aussi que les segments initiaux stricts de $(\omega\times\omega,\preceq)$ sont tous finis, on en déduit que $(\omega\times\omega,\preceq)$ est un ensemble bien ordonné dont tous les segments initiaux stricts sont finis. Il est donc isomorphe à ω, ce qui nous donne une bijection de $\omega\times\omega$ sur ω et prouve par conséquent que $\aleph_0.\aleph_0=\aleph_0$.

Supposons maintenant que α soit un ordinal >0 tel que $\aleph_\beta.\aleph_\beta=\aleph_\beta$ pour tout $\beta<\alpha$. Considérons à nouveau les segments initiaux stricts de $(\aleph_\alpha\times\aleph_{\alpha_0},\preceq)$. Pour cela, fixons une paire $(\theta,\beta)\in\aleph_\alpha\times\aleph_\alpha$, et notons $\gamma=\max(\theta,\beta)+1$. Alors la définition de $\preceq$ impose que le $\preceq$-segment initial associé à (θ,β) est contenu dans $\gamma\times\gamma$. Mais puisque $\gamma<\aleph_\alpha$, on a $|\gamma|=\aleph_\delta$ pour un certain $\delta<\alpha$, par conséquent notre hypothèse de récurrence nous donne $|\gamma\times\gamma|=\aleph_\delta$. On voit donc que les segments initiaux stricts de $(\aleph_\alpha\times\aleph_\alpha,\preceq)$ sont tous de cardinal strictement inférieur à $\aleph_\alpha$.

Appelons $\Gamma(\alpha)$ l'unique ordinal isomorphe à $(\aleph_\alpha\times\aleph_\alpha,\preceq)$; on a dit que c'est un ordinal limite, donc c'est la borne supérieure [vii] de ses segments initiaux stricts. Ceux-ci étant tous inférieurs à $\aleph_\alpha$ d'après ce qui précède, on en déduit que $\Gamma(\alpha)\subseteq\aleph_\alpha$. Puisque $\aleph_\alpha\times\aleph_\alpha$ et $\Gamma(\alpha)$ sont isomorphes, ils sont en particuliers équipotents et on vient donc de prouver que $\aleph_\alpha.\aleph_\alpha\leq\aleph_\alpha$. Il est bien clair que $\aleph_\alpha\leq\aleph_\alpha.\aleph_\alpha$, par conséquent le théorème de Schröder-Bernstein nous permet d'affirmer que $\aleph_\alpha.\aleph_\alpha=\aleph_\alpha$.
Soit maintenant deux cardinaux infinis κ,λ, dont on suppose que $\kappa\leq\lambda$.

vii. ou la réunion, si vous préférez !

Alors on peut écrire, en utilisant le résultat qu'on vient de démontrer :

$$\lambda \leq \kappa + \lambda \leq \kappa.\lambda \leq \lambda.\lambda = \lambda \ .$$

Ceci prouve bien que $\kappa + \lambda = \kappa.\lambda = \lambda$. □

Notons que, en présence de l'axiome du choix, le résultat ci-dessus nous dit que pour tout ensemble infini X les ensembles X et $X \times X$ sont équipotents ; il se trouve que cet énoncé est *équivalent* à l'axiome du choix ! On n'a pas besoin de l'axiome du choix pour ajouter/multiplier deux cardinaux ; par contre, on en a besoin pour parler du cardinal d'un ensemble X quelconque, et appliquer l'arithmétique cardinale à cet ensemble.

Proposition 3.19. *Soit κ, λ, μ trois cardinaux. Alors $\kappa.(\lambda + \mu) = \kappa.\lambda + \kappa.\mu$.*

Preuve.
Il nous faut montrer que $\kappa \times (\lambda \sqcup \mu)$ et $(\kappa \times \lambda) \sqcup (\kappa \times \mu)$ sont équipotents. En revenant à notre définition de $\sqcup$, on voit facilement que la fonction f suivante est une bijection entre les deux ensembles qui nous intéressent :

$$\begin{cases} f(\alpha, (\beta, 0)) & = ((\alpha, \beta), 0) \\ f(\alpha, (\beta, 1)) & = ((\alpha, \beta), 1) \end{cases}$$

Ceci conclut la preuve. □

Comme pour les ordinaux, il nous reste à définir une dernière opération arithmétique : l'exponentiation.

Définition 3.20. Soit κ, λ deux cardinaux ; on définit κ^λ comme le cardinal de l'ensemble des fonctions de λ dans κ.

Exercice 3.21. Si X_0, X_1 (resp. Y_0, Y_1) sont des ensembles équipotents, alors $X_0^{Y_0}$ et $X_1^{Y_1}$ sont équipotents.

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

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

Exercice 3.22. Montrer que pour trois cardinaux κ, λ, μ on a $(\kappa^\lambda)^\mu = \kappa^{\lambda.\mu}$, et $\kappa^\lambda.\kappa^\mu = \kappa^{\lambda+\mu}$.

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

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

Théorème 3.24. (Cantor) *Pour tout ensemble X il n'existe pas de surjection $f\colon X \to \mathcal{P}(X)$. Avec nos notations cela signifie que pour tout cardinal κ on a $\kappa < 2^\kappa$.*

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

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

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

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

Définition 3.25. *L'hypothèse du continu* (HC) est l'énoncé $2^{\aleph_0} = \aleph_1$.
L'hypothèse du continu *généralisée* est l'énoncé affirmant que pour tout ordinal α on a $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

L'idée sous-jacente de l'hypothèse du continu est qu'on peut « voir »$\mathbb{N}$, de cardinal $\aleph_0$, et $\mathbb{R}$, 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.

Exercice 3.26. 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

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

Définition 3.27. Soit λ un ordinal et $(\kappa_\alpha)_{\alpha<\lambda}$ une suite de cardinaux indexée par λ. On définit $\sum_{\alpha<\lambda} \kappa_\alpha$ comme l'unique cardinal équipotent à $\bigsqcup \kappa_\alpha$.
De même, $\prod_{\alpha<\lambda} \kappa_\alpha$ est l'unique cardinal équipotent au produit cartésien des κ_λ.

Vérifions que ces sommes/produits infinis ont bien les propriétés attendues.

Proposition 3.28. *Soit κ, λ deux cardinaux. Alors on a $\sum_{\alpha<\kappa} \lambda = \kappa.\lambda$ et $\prod_{\alpha<\kappa} \lambda = \lambda^\kappa$.*

Preuve.
On va simplement montrer (avec l'axiome du choix) la première égalité, la deuxième se démontrant sur le même modèle. Fixons une famille $(X_\alpha)_{\alpha<\kappa}$ d'ensembles deux à deux disjoints de cardinal λ, et pour tout $\alpha < \kappa$ fixons aussi [viii] une bijection f_α de X_α sur λ.
Alors, on peut définir une fonction $F\colon \bigcup X_\alpha \to \kappa \times \lambda$ en posant, pour tout α et tout $x \in X_\alpha$,

$$F(x) = (\alpha, f_\alpha(x)) .$$

Cette fonction est bijective, puisqu'elle a pour inverse la fonction

$$(\alpha, \beta) \mapsto f_\alpha^{-1}(\beta) .$$

Ceci conclut la preuve. □

3.4 Dénombrabilité

Commençons par introduire une notation.

viii. C'est là qu'on utilise l'axiome du choix

Définition 3.29. On appelle ω_1 le plus petit ordinal non dénombrable. Remarquons que, quand on y pense comme à un cardinal, on lui a donné un autre nom : $\aleph_1$.

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 [ix].

Rappelons ce que nous avons vu précédemment : un ensemble est dénombrable si, et seulement si, il est équipotent à ω. A l'aide de l'axiome du choix, on voit également facilement que tout ensemble infini contient un sous-ensemble dénombrable : si X est infini et bien ordonnable alors X est équipotent à un cardinal κ, et κ contient une copie de ω. Ce fait, allié au théorème de Schröder-Bernstein, nous dit qu'un ensemble infini X est dénombrable si, et seulement si, X s'injecte dans ω ou encore si, et seulement si, ω se surjecte sur X.

On a déjà vu que, modulo l'axiome du choix dénombrable, une réunion dénombrable d'ensembles dénombrables est dénombrable. Ceci a une conséquence importante.

Lemme 3.30. *Soit X un ensemble dénombrable. Alors l'ensemble des parties finies de X est dénombrable.*

Preuve.
Etant donné ce qu'on vient de dire sur la réunion dénombrable d'ensembles dénombrables, il nous suffit de prouver que pour tout entier k l'ensemble des parties de X de cardinal k est dénombrable. Pour cela, on va utiliser le fait que ω^k est dénombrable (qui est une conséquence de nos raisonnements précédents sur l'arithmétique des cardinaux, appliqués au cardinal $\aleph_0$). Fixons une bijection de ω sur X ; alors l'ensemble X_k des parties de X de cardinal k est l'image de ω par la bijection

$$(n_1, \ldots, n_k) \mapsto (f(n_1), \ldots, f(n_k)) \ .$$

Ceci prouve que X_k est bien dénombrable, et achève la preuve du fait que l'ensembles des parties finies de X est dénombrable. □

ix. 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 ».

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 (au plus) dénombrable. Alors le sous-groupe de G engendré par A est (au plus) dénombrable.

Preuve.
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 par la fonction qui associe $a_1 \dots a_k$ à $(a_1, \dots, a_k)$, par conséquent chaque A_k est dénombrable et donc le sous-groupe engendré par A est lui aussi dénombrable. □

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

Exercice 3.31. 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 $\omega^{\omega_1} = \omega_1$, et qu'il existe pour tout ordinal *dénombrable* α un ordinal *dénombrable* $\beta \geq \alpha$ tel que $\omega^\beta = \beta$.

Il n'est pas trop difficile de prouver que $\mathbb{Q}$ est dénombrable. Ceci nous permet de calculer le cardinal de $\mathbb{R}$, comme le montre l'exercice suivant.

Exercice 3.32. Montrer que $|\mathbb{Q}^2| = \aleph_0$, puis montrer que $|\mathbb{R}| = 2^{\aleph_0}$. Pour le second point, on pourra considérer l'application $f \colon \mathbb{R} \to \mathcal{P}(\mathbb{Q}^2)$ définie par

$$f(x) = \left\{(q, q') \in \mathbb{Q}^2 \colon q < x < q'\right\} \ .$$

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

Définition 3.33. Un ensemble est *infini* s'il n'est équipotent à aucun ordinal fini. Un ensemble X est *Dedekind-infini* s'il existe une injection non surjective $f : X \to X$.

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

Exercice 3.34. Montrer qu'un ensemble est Dedekind-infini si, et seulement si, il contient un sous-ensemble dénombrable.

Exercice 3.35. 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.

3.5 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 3.36. Un cardinal infini κ est dit *régulier* 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. Pour comprendre cette notion, on va introduire un nouvel invariant combinatoire, la *cofinalité* d'un cardinal.

Définition 3.37. Soit α, β deux ordinaux. On dit que α *est cofinal à* β s'il existe une fonction $f : \beta \to \alpha$ strictement croissante et dont l'image n'est pas strictement majorée dans α (autrement dit, pour tout $\gamma \in \alpha$ il existe $\delta \in \beta$ tel que $f(\delta) \geq \gamma$).

La *cofinalité* de α est le plus petit ordinal β tel que α soit cofinal à β. On note alors $\beta = \mathrm{cof}(\alpha)$.

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

Exercice 3.38.
1. Montrer que $\mathrm{cof}(\alpha)$ 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.
2. Montrer que, pour tout ordinal α, $\mathrm{cof}(\alpha)$ est un cardinal.
3. Montrer que $\mathrm{cof}(\mathrm{cof}(\alpha)) = \mathrm{cof}(\alpha)$ pour tout ordinal α.
4. Montrer qu'un cardinal λ infini est régulier si et seulement si $\mathrm{cof}(\lambda) = \lambda$.

Exercice 3.39.
1. Montrer qu'un cardinal κ est régulier si, et seulement si, pour tout $\lambda < \kappa$ et toute famille $(X_\alpha)_{\alpha\in\lambda}$ d'ensembles tels que $|X_\alpha| < \kappa$ pour tout $\alpha < \lambda$, on a $|\cup X_\alpha| < \kappa$.
2. Soit κ un cardinal ; montrer que $\mathrm{cof}(\kappa)$ est le plus petit ordinal γ tel que α soit la réunion de γ ensembles de cardinal strictement inférieur à κ.
3. On appelle *faiblement 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 [x] ?

Proposition 3.40. *Tout cardinal successeur est régulier.*

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

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

On en déduit que

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

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

On peut, après la proposition ci-dessus, avoir l'impression que tout cardinal est régulier : mais les cardinaux limites sont bien souvent singuliers. C'est par exemple le cas de $\aleph_\omega$, qui est une union dénombrable d'ensembles

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

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 *faiblement inaccessible*.

Il se trouve que l'existence d'un cardinal faiblement inaccessible est un énoncé indépendant de ZFC : on ne peut ni le prouver ni prouver sa négation à partir des axiomes de ZFC. Comme remarque culturelle, notons qu'il en va de même des cardinaux *fortement inaccessibles* ; par définition, un cardinal non dénombrable λ est fortement inaccessible s'il est faiblement inaccessible et pour tout $\kappa < \lambda$ on a $2^\kappa < \lambda$.

Exercice 3.41. Supposons que λ soit un cardinal fortement inaccessible. Montrer qu'alors l'ensemble des parties de λ de cardinal strictement inférieur à λ est un univers dans lequel les axiomes de ZFC sont vérifiés.

En particulier, le théorème de Gödel entraîne immédiatement que[xi] dans ZFC on ne peut pas prouver l'existence d'un cardinal fortement inaccessible ; sinon, ZFC démontrerait sa propre consistance et c'est impossible. Il se trouve qu'on peut aussi démontrer que l'existence d'un cardinal fortement inaccessible est indépendante de ZFC ; mais ce théorème est bien au-delà de ce qu'on peut démontrer avec les outils de ce cours.

Nous sommes maintenant équipés pour prouver le dernier théorème de ce chapitre, le *lemme de König*.

Lemme 3.42. *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*

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

Ce lemme peut paraître évident : on voit mal comment une union d'ensemble de petit cardinal pourrait coïncider avec un produit d'ensembles de gros cardinal. Mais, si l'on repense à l'axiome du choix, on se dit que les propriétés des produits infinis sont assez contre-intuitives... Et, effectivement, le lemme de König est plus subtil qu'il n'y paraît.

Preuve du lemme de König.
Montrons déjà (avec l'axiome du choix !) que l'inégalité large est vraie : considérons une famille d'ensembles deux à deux disjoints B_i de cardinal κ_i, et des ensembles A_i de cardinal λ_i. Fixons pour tout i[xii] un élément $a_i \in A_i$, et

xi. pour peu que ZFC soit consistant, bien sûr...

xii. Cest un des endroits où l'axiome du choix intervient dans cette preuve. Quels sont les autres ?

une injection $f_i\colon B_i \to A_i \setminus \{a_i\}$. On peut maintenant définir une injection $f\colon \bigsqcup B_i \to \prod A_i$ en définissant pour $x \in B_{i_0}$ $f(x)$ par

$$f(x)(i) = \begin{cases} a_i & \text{si } i \neq i_0 \\ f_i(x) & \text{si } i = i_0 \end{cases}$$

Il est facile de vérifier que cette fonction est injective, et ceci signifie bien que $\sum \kappa_i \leq \prod \lambda_i$.

Pour prouver que l'inégalité est stricte, on raisonne par l'absurde et on suppose que $\sum \kappa_i = \prod \lambda_i$. Cela signifie qu'on peut trouver deux familles d'ensembles $(A_i)_{i\in I}$ et $(B_i)_{i\in I}$ telles que :
• $|A_i| = \lambda_i$,
• $B_i \subset \prod_{j\in I} A_j$, $|B_i| = \kappa_i$, et
• $\cup_{i\in I} B_i = \prod A_i$.
Considérons maintenant la projection π_i sur la i-ième coordonnée, qui envoie B_i dans A_i ; puisque $|B_i| \leq |B_i| < |A_i|$, l'ensemble $Y_i = 1_i \setminus \pi_i(B_i)$ est non vide pour tout $i \in I$. L'axiome du choix nous autorise alors à considérer un élément $x = (x_i) \in \prod_{i\in I} Y_i$. Cet élement est construit de telle sorte que $\pi_i(x) = x_i \notin \pi_i(B_i)$, par conséquent $x \notin \cup B_i$ et ceci contredit le fait que $\bigcup_{i\in I} B_i = \prod_{i\in I} A_i$. □

Le raisonnement ci-dessus est un bon exemple de *raisonnement diagonal* : on construit x de telle façon que la i-ième coordonnée de X assure que x n'appartient pas à B_i.

Exercice 3.43. En utilisant le fait que tout réel admet un développement décimal [xiii], prouver à l'aide d'un raisonnement diagonal que $\mathbb{R}$ n'est pas dénombrable.

Corollaire 3.44. Pour tout cardinal infini κ, on a $\kappa < \kappa^{\mathrm{cof}(\kappa)}$.

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

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

Ceci conclut la preuve. □

On en déduit immédiatement un autre corollaire.

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

Corollaire 3.45. Pour tout cardinal infini κ, on a $\mathrm{cof}(2^\kappa) > \kappa$.

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

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

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

Ceci a pour conséquence une restriction à la négation de l'hypothèse du continu : pour tout ordinal limite dénombrable $\alpha < 2^{\aleph_0}$, il est impossible que l'on ait $2^{\aleph_0} = \aleph_\alpha$. En effet, la cofinalité de $\aleph_\alpha$ pour un tel α est égale à $\aleph_0$. 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.

Retour sur une erreur dans une version précédente des notes. Il était écrit que, si X et Y sont deux ensembles et s'il existe une surjection de X sur Y, alors il existe une injection de Y dans X. C'est faux sans l'axiome du choix (de même que le théorème öbtenu en remplaçant « injection »par « surjection »dans l'énoncé du théorème de Schöder-Bernstein). Ce qui est correct sans l'axiome du choix, et facile à démontrer, est que s'il existe une injection de X dans Y alors il existe une surjection de Y sur X.

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 [KM] 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].

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.

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

Chapitre 4

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... Pour simplifier un peu l'exposition, on ne parlera que de filtres sur des ensembles infinis ; ainsi, la lettre X désignera dans toute la suite un ensemble infini.

4.1 Définitions, premières propriétés

Définition 4.1. Soit X un ensemble infini. Un *filtre* sur X est une famille $\mathcal{F} \subset \mathcal{P}(X)$ vérifiant les propriétés suivantes :

1. $A \in \mathcal{F}$ et $B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$;
2. $A \in \mathcal{F}$ et $B \supseteq A \Rightarrow B \in \mathcal{F}$;
3. $\emptyset \notin \mathcal{F}$.

Exemple. Rappelons que X est un ensemble infini.

- 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*.
- Nettement 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 4.2. Soit $\mathcal{F}$ un filtre contenant une partie *finie* A. Alors $\mathcal{F}$ est principal.

Définition 4.3. 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 4.4. *Pour toute base de filtre $\mathcal{A}$, il existe un filtre contenant $\mathcal{A}$; le plus petit tel filtre est défini par*

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

Preuve.
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 $\emptyset \notin \mathcal{F}$ et donc $\mathcal{F}$ est bien un filtre. □

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

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

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

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

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

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 !).

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

Définition 4.5. On dit qu'un filtre $\mathcal{F}$ est un *ultrafiltre* si pour tout filtre $\mathcal{G}$ on a

$$\mathcal{F} \subseteq \mathcal{G} \Rightarrow \mathcal{F} = \mathcal{G} \ .$$

Les ultrafiltres sont donc exactement les filtres maximaux pour l'inclusion. 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 4.6. *Un filtre $\mathcal{F}$ est un ultrafiltre si, et seulement si, pour tout $A \in X$ on a $A \in \mathcal{F}$ ou $X \setminus A \in \mathcal{F}$.*

Preuve.
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 4.7. 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.

4.2 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 4.8. 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 4.9. 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 4.10. Soit X, Y deux ensembles, $\mathcal{F}$ un filtre sur X et $f\colon X \to Y$ une fonction. Alors $\{B \subseteq Y : \exists A \in \mathcal{F}\ B = f(A)\}$ 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 4.11. *Le filtre image d'un ultrafiltre sur X est un ultrafiltre sur Y.*

Preuve.
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 4.12. *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.

Preuve.
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

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

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 4.13. *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.*

Preuve.
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 4.14. *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,*

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

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

Preuve.
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 4.14 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 4.16. *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.

4.3 Un exemple combinatoire : les ultrafiltres de Ramsey

Définition 4.17. 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

ii. Avec l'axiome du choix !

$A_n \notin \mathcal{F}$ 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 4.18. *Si $2^{\aleph_0} = \aleph_1$ alors il existe un ultrafiltre de Ramsey.*

Avant de prouver ce théorème, établissons un lemme simple.

Lemme 4.19. *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 4.18.
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,

iii. On peut remplacer, sans changer la notion, cette condition par $|X \cap A_n| = 1$ pour tout n ; pourquoi ?

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 $X_\alpha \setminus X_\beta$ soit fini pour tout $\beta < \alpha$. Le fait qu'il est bien possible de faire ça sera justifié par le lemme 4.21 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 $X_\gamma \setminus X_{\alpha_i}$ est fini pour tout $i \in \{1, \ldots, n\}$; par conséquent $X_\gamma \setminus (\cap X_{\alpha_i})$ 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 4.20. *Pour tout $\alpha < \beta < \omega_1$, $X_\beta \setminus X_\alpha$ est fini.*

Preuve du Lemme 4.20.
On raisonne par récurrence transfinie : on va prouver que pour tout β la propriété « pour tout $\alpha < \beta$, $X_\beta \setminus X_\alpha$ 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 $X_\beta \setminus X_\alpha$ 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 4.21. *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 $Y \setminus Y_i$ 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 4.20) que $X_\gamma \setminus X_\beta$ est fini. Mais pour tout ensemble fini d'ordinaux $\beta_1, \ldots, \beta_n < \alpha$, si on pose $\beta = \max\{\beta_i \colon i = 1, \ldots, n\}$ alors la construction assure que

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

Puisque X_β est infini, ceci impose bien que $\bigcap_{i=1}^n X_{\beta_i}$ est infini. En appliquant le lemme 4.21 à la famille $\{X_\beta\}_{\beta<\alpha}$, on obtient donc une partie Y telle que $Y \setminus X_\beta$ est fini pour tout $\beta < \alpha$, et on peut finalement poser $X_\alpha = Y$.

Preuve du Lemme 4.21.
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.

[KM] Kazimierz Kuratowski and Andrzej Mostowski. *Set theory, with an introduction to descriptive set theory*. Studies in Logic and the Foundations of Mathematics.

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

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

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

Index