Introduction à la Logique Mathématique

Première partie : Théorie des ensembles

Thomas Blossier & Julien Melleray

Avertissement.

Ces notes sont encore susceptibles d'évoluer. Attentions aux typos et autres erreurs.

Table des matières

Chapitre 1

Les ordinaux

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

Remarquons qu'un bon ordre est nécessairement total : si $x, y \in X$ alors l'ensemble $\{x, y\}$ a un plus petit élément, ce qui impose que $x \leq y$ ou $y \leq x$. Notons de plus, que toute partie d'un bon ordre est bien ordonné pour l'ordre induit.

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

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

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

Dans un ensemble bien ordonné X, tout segment initial S 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 1.3. *Soit $(X, \leq)$ un ensemble bien ordonné et $f : X \to X$ une application strictement croissante. Alors pour tout $x \in X$ on a $f(x) \geq x$.*

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 1.4. *Soit $(X, \leq)$ un ensemble bien ordonné, $W \subseteq X$ un segment initial et $f : X \to W$ 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 la proposition 1.4 entraîne que $X \sim X'$ si, et seulement si, $X \preceq X'$ et $X' \preceq X$. On a dit qu'on souhaitait pouvoir énumérer tous les ensembles bien ordonnés ; mais quelle notion de « longueur »utiliser ?

Théorème 1.5. *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 $x < y \in S$ 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 $f_y \circ f_x^{-1}(z) = z$ pour tout $z \in f_x(S_x) = f_y(S_x)$. 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 1.6. *Soit $\mathcal{W} = \{W_i : i \in I\}$ 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 $\{x \in W_0 : S_x \text{ est isomorphe à un élément de } \mathcal{W}\}$ 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 possible (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'élément 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 1.7. 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).

Évidemment, 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 1.8. *Une réunion d'ensemble transitifs est encore un ensemble transitif ; une intersection d'ensembles transitifs est encore un ensemble transitif.*

Définition 1.9. 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 [i], les ordinaux se comportent bien relativement à $\in$.

Lemme 1.10. *Pour tout ordinal α, on a $\alpha \notin \alpha$ et les éléments de α sont des ordinaux.*

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

i. L'axiome de fondation 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) telles que $x_{n+1} \in x_n$ pour tout n. Cet axiome n'est pas une conséquence des axiomes de (ZF)

déduirait que $\alpha \notin \alpha$, ce qui est bien sûr une contradiction.
Soit $\delta \in \alpha$. Par transitivité de α, on a $\delta \subset \alpha$ et ainsi $\in$ est également une relation de bon ordre strict sur δ. Vérifions que δ est également transitif. Soit $y \in x \in \delta$. Par transitivité de α, on a $x, y \in \alpha$ et par transitivité de l'ordre strict $\in$ dans α, on obtient $y \in \delta$. $\square$

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

Proposition 1.11. *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 (car β est inclus dans les ordinaux dont il est l'intersection). Par conséquent, β est bien un ordinal. $\square$

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 1.12. *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 par transitivité $\delta \in \beta \setminus \alpha$, 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 inclusion, 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 $\alpha = \gamma$. $\square$

Corollaire 1.13. Tout ordinal est égal à l'ensemble de ses segments initiaux stricts.

Preuve.
Soit α un ordinal. Si $\beta \in \alpha$, alors par le lemme 1.10, β est un ordinal et $\beta \neq \alpha$. Par transitivité de α et β, il suit que $\beta \subseteq \alpha$ et β est un segment initial de α.
Réciproquement, si β est un segment initial strict de α alors β est transitif et bien ordonné par $\in$ (comme segment initial d'un ensemble transitif bien ordonné par $\in$) donc β est un ordinal tel que $\beta \subset \alpha$ et $\beta \neq \alpha$, d'où $\beta \in \alpha$ d'après le résultat ci-dessus. $\square$

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 1.14. *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.

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 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 $\alpha = \beta$. $\square$

Ainsi, si α et β sont deux ordinaux, on a les équivalences suivantes :
— $\alpha \prec \beta$ si et seulement si $\alpha \in \beta$,
— $\alpha \preceq \beta$ si et seulement si $\alpha \subseteq \beta$,
— $\alpha \sim \beta$ si et seulement si $\alpha = \beta$.

L'intuition étant souvent un peu réfractaire au fait que $\in$ soit une relation d'ordre, on va maintenant utiliser $<$ pour dénoter la relation d'ordre entre ordinaux (autrement dit, on écrira $\alpha < \beta$ comme synonyme de $\alpha \in \beta$). De même, on écrira $\alpha \leq \beta$ pour signifier que $\alpha \in \beta$ ou $\alpha = \beta$, ce qui est équivalent à dire que $\alpha \subseteq \beta$.

Proposition 1.15. *1. Tout ensemble d'ordinaux est bien ordonné.*

2. L'union d'un ensemble d'ordinaux est encore un ordinal et correspond à la borne supérieure de cet ensemble.

3. Tout ensemble non vide d'ordinaux a un plus petit élément, qui est égal à l'intersection des éléments de cet ensemble.

Preuve.

1. Soit E un ensemble d'ordinaux ; d'après ce qui précède $\leq$ est un ordre total 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 $(\alpha, \leq)$ est bien ordonné) est aussi le plus petit élément de A : en effet, si $\beta \in A \cap \alpha$ alors on doit avoir $\gamma \leq \beta$ par définition de γ, et si $\beta \in A \setminus \alpha$ alors on doit avoir $\alpha \leq \beta$ et donc $\gamma < \beta$.
   Si maintenant $\alpha \cap A = \emptyset$, alors pour tout $\beta \in A$ on ne peut avoir $\beta \in \alpha$, d'où $\alpha \leq \beta$ et on voit que dans ce cas α est le plus petit élément de A.
2. 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 bien ordonné. De plus on sait que α est transitif, par conséquent α est un ordinal.

Notons que α contient tous les éléments de E, c'est donc un majorant de E. Si γ est un majorant de E, alors γ contient chaque élément de E et donc α.

3. Soit E un ensemble non vide d'ordinaux. Par le premier point, E possède un plus petit élément, qui est ainsi inclus dans tous les éléments de E. Ce plus petit élément correspond donc à l'intersection des éléments de E. □

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 [ii]. La définition formelle de ce successeur est présentée ci-dessous.

Proposition 1.16. *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 1.12, 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 1.17. *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 un ordinal α, qui est la borne supérieure de cet ensemble. Comme $S(\alpha) > \alpha$, on ne peut avoir $S(\alpha) \in \mathcal{O}$. □

À partir de maintenant, on notera ON la collection des ordinaux.

Proposition 1.18. *Pour tout ensemble X il existe un plus petit ordinal non équipotent à une partie de X. On l'appelle l'ordinal de Hartogs de X.*

Preuve.
Soit $A = \{\alpha \in \mathrm{ON} \colon$ il existe une injection de α dans $X\}$. Vérifions que A est un ensemble : on note d'abord que

$$\{(S, \leq) \in \mathcal{P}(X) \times \mathcal{P}(X \times X) \colon \text{ la relation } \leq \text{ est un bon ordre sur } S\}$$

est un ensemble. Comme tout ensemble ordonné est au plus isomorphe à un ordinal (car deux ordinaux isomorphes sont égaux), on en déduit que A est

ii. ce qui fait un point commun entre les ordinaux et les hommes politiques (espérons-le).

également un ensemble [iii]. Comme A est un ensemble d'ordinaux, il existe donc un ordinal β qui n'appartient pas à A. Alors le plus petit élément de $S(\beta) \setminus A$ convient. (On peut également vérifier que A est lui même un ordinal et en déduire que c'est le plus petit ordinal non équipotent à une partie de X.) $\square$

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

Preuve.
Soit $(X, <)$ un ensemble bien ordonné. On considère α l'ordinal de Hartogs de X. Par le théorème 1.5 et la définition de l'ordinal de Hartogs, on a nécessairement X isomorphe à un segment initial de α, donc à un ordinal. Cet ordinal est unique car deux ordinaux isomorphes sont égaux. $\square$

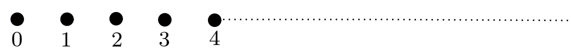
Introduisons un peu de terminologie.

Définition 1.20. 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* .

Un ensemble non vide d'ordinaux n'a pas en général de plus grand élément (pensez aux ordinaux finis !). 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 éléments 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 1.21. Montrer qu'un ordinal β est limite si, et seulement si, $\beta = \sup\{\eta\colon \eta < \beta\}$.

Définition 1.22. 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 :

iii. On utilise ici le schéma de remplacement de la théorie ZF, qui dit qu'étant donné un ensemble, les images de ses éléments par une relation fonctionnelle forment également 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 $\beta < \alpha$.
- Par rigidité, l'isomorphisme entre un ensemble bien ordonné et un ordinal, est également unique. On peut donc utiliser cette correspondance sans risque d'ambiguïté.
- Pour tout ordinal α, on a $\alpha = \{\beta \in \mathrm{ON} \colon \beta < \alpha\}$.
- La réunion d'un ensemble d'ordinaux E est un ordinal, qui est la borne supérieure de E.
- L'intersection d'un ensemble d'ordinaux E est un ordinal, qui est le plus petit élément de E.
- Il existe deux types d'ordinaux : les ordinaux successeurs (ceux qui ont un plus grand élément) et les ordinaux limites (ceux qui n'ont pas de plus grand élément).

1.2 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 1.23. *(Démonstration par récurrence transfinie)*
Soit P une propriété[iv] *des ordinaux telle que pour tout ordinal α on ait*

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

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

iv. La notion de propriété est floue ; disons en anticipant un peu sur le cours qu'une propriété est quelque chose qu'on peut exprimer par un énoncé du premier ordre écrit en utilisant le langage de la théorie des ensembles.

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 1.24. *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 utilise 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 1.25. *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. On voudrait définir $f \colon A \to X$ de la façon suivante :

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

Mais on n'est a priori pas sûr que ce soit possible ; quitte à rédiger de manière un peu lourde puisque c'est notre première construction de cette sorte, on introduit un symbole $*$ qui n'appartient pas à X. Puis on procède comme ci-dessus, sauf que si jamais à une étape on ne peut pas définir $f(a)$ satisfaisant les conditions ci-dessus, alors on pose $f(a) = *$ et ensuite pour tout $b \geq a$ on pose $f(b) = *$.

Maintenant, prouvons par récurrence transfinie, que pour tout $a \in A$ on a $f(a) \neq *$ et que $f(a) \leq a$. C'est clair pour $a = \min(A)$. Si a est successeur de a' dans A, alors $f(a') \leq a' < a$, donc a est un majorant strict de $f(a')$ qui a donc un successeur $\leq a$; on a donc la fois $f(a) \in X$ et $f(a) \leq a$. Reste le cas où a est limite dans A, donc aussi dans X ; alors pour tout $a' < a$ dans A on a $f(a') \leq a' < a$, donc $\sup_X\{f(a') \colon a' \in A \text{ et } a' < a\}$ est bien défini et plus petit que a. On a donc bien $f(a) \in X$ et $f(a) \leq a$.

Il nous reste simplement à montrer que l'image de f est un segment initial de X. Pour cela, on pourrait procéder par l'absurde, ou montrer par récurrence transfinie que, si S_a^A est le segment initial associé à a dans A alors pour tout a dans A on a $f(S_a^A) = S_{f(a)}^X$. 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

$$f(S_a^A) = f(S_{a'}^A \cup \{a'\}) = S_{f(a')}^X \cup \{f(a')\}$$

et $f(a)$ est le successeur de $f(a')$ dans X, donc $f(S_a^A) = S_{f(a)}^X$.

— Si $a \neq \min(A)$ est limite dans A, alors

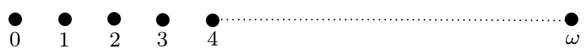
$$f(S_a^A) = \bigcup_{a'<a} f(S_{a'}^A) = \bigcup_{a'<a} S_{f(a')}^X = S_{\sup\{f(a')\colon a'<a\}}^X = S_{f(a)}^X \ .$$

Ceci achève la démonstration [v]. ☐

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

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



L'ordinal $\omega + 1$

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

Exemple. Utilisons une démonstration par récurrence transfinie pour montrer que l'addition est associative, et que si $\alpha \neq \beta$ alors pour tout δ on a $\delta+\alpha \neq \delta+\beta$. Commençons par le deuxième point ; 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 $\delta + \beta = \sup\{\delta + \eta\colon \eta < \beta\}$; comme $\beta > \alpha$ et β est limite on peut trouver η tel que $\alpha < \eta < \beta$, et par définition d'un sup $\delta + \beta \geq \delta + \eta > \delta + \alpha$ par hypothèse de récurrence.

v. 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.

A partir de cette propriété, il n'est pas difficile de démontrer que si A est un ensemble d'ordinaux, alors pour tout ordinal α on a

$$\alpha + \sup(A) = \sup(\{\alpha + \gamma \colon \gamma \in A\}) \ .$$

.

Raisonnons par récurrence sur γ par montrer que, étant donnés trois ordinaux α, β, γ on a $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$. 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 la remarque faite juste avant le début de notre récurrence)) :

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

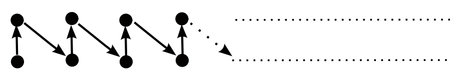
On voit donc que $P(\gamma)$ est vraie, et on a fini de prouver que l'addition ordinale est associative ; ici le lecteur attentif devrait se rendre compte que, même s'il n'y a pas de difficulté particulière dans le raisonnement, il faut apporter un certain soin à la rédaction pour qu'elle soit correcte ; par conséquent il faut s'entraîner à écrire ce type de démonstration ! □

Dans la suite on utilisera toujours la notation $\alpha+1$ pour désigner le successeur d'un ordinal α.

Définition 1.27. (multiplication ordinale) Soit α un ordinal. On pose $\alpha.0 = 0$, puis on définit par récurrence transfinie sur $\beta \in \mathrm{ON}$ la multiplication ordinale $\alpha.\beta$ en posant :

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

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

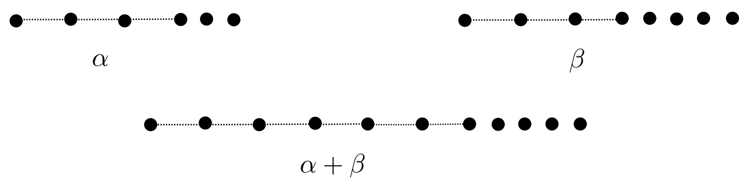


$2.\omega = \omega$

Exercice 1.28. 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 $\alpha < \alpha.\gamma$. Prouver 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 1.29.
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 1.30. 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 1.31. Soit α un ordinal. On pose $\alpha^0 = 1$, puis on définit, par récurrence transfinie sur $\beta \in \mathrm{ON}$, α^β en posant :

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

Par récurrence transfinie, on vérifie les propriétés suivantes.
— Etant donnés trois ordinaux α, β, γ, on a $\left(\alpha^\beta\right)^\gamma = \alpha^{\beta \cdot \gamma}$.
— 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 $\beta = \beta^\omega$?

Chapitre 2

Cardinaux et axiome du choix.

2.1 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 2.1. 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.

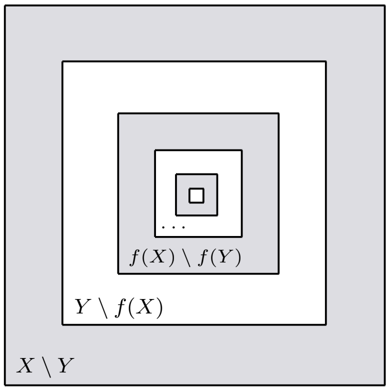
Ainsi, on cherche à étendre les notions intuitives de comptage, qui marchent pour les ensembles finis [i], à tous les ensembles. Déjà, il faut s'assurer que ces notions sont bien compatibles entre elles ; au début, tout se passe bien.

Théorème 2.2. (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.

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

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 $g(X) = Y$. $\square$

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 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 on a déjà vu que ce n'est pas le cas en construisant l'ordinal de Hartogs de X. Par conséquent, avec nos méthodes, on aura besoin de l'axiome de Zermelo pour avoir une notion satisfaisante de cardinal d'un ensemble.

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

Définition 2.3. 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 2.4. *Étant 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 2.5. Étant donné un cardinal κ, on note κ^+ le plus petit cardinal qui n'est pas équipotent à une partie de κ.

Définition 2.6. 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 κ 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 2.7. *(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 2.8. *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 2.9. *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. $\square$

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 2.10. *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

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

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. $\square$

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 ».

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 $\beta = \omega^\beta$

2.2 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 $f : \mathcal{P}(X) \to X$ telle que pour toute partie $A \subseteq X$ non vide on ait $f(A) \in A$.

Définition 2.11. 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) $\Rightarrow$ (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) $\Rightarrow$ (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 $\alpha = \sup\{\beta : \beta < \alpha\}$

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

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

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

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 $A = X$. □

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 2.12. 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.

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 2.13. Montrer que l'axiome du choix dénombrable entraîne que toute réunion dénombrable d'ensembles dénombrables est dénombrable (rappelons

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

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 2.14. 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.

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

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

Exercice 2.16. Montrer qu'un ensemble est Dedekind-infini si, et seulement si, il contient un sous-ensemble dénombrable.

Exercice 2.17. 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.

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.

---

v. et ce fait est indépendant de ZF.

2.3 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

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

Définition 2.18. 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 2.19. *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 $f \colon X \to X'$ et $g \colon Y \to Y'$ deux bijections. Alors on peut définir une fonction $F \colon X \sqcup Y \to X' \sqcup Y'$ 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 $G \colon X \times Y \to X' \times Y'$ 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 2.20.
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.

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

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

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 $|\alpha|.|\beta|$.

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 γ qui s'écrivent 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 infini 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 $(\aleph_\alpha \times \aleph_\alpha, \preceq)$. 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 $\kappa \leq \lambda$. 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 2.21. *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

vii. ou la réunion, si vous préférez !

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. $\square$

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

Définition 2.22. Soit κ, λ deux cardinaux ; on définit κ^λ comme le cardinal de l'ensemble des fonctions de λ dans κ.

Exercice 2.23. 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 2.24. Montrer que pour trois cardinaux κ, λ, μ on a $\left(\kappa^\lambda\right)^\mu = \kappa^{\lambda\cdot\mu}$, 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 2.25. Montrer que pour tout cardinal infini κ on a $2^\kappa = \kappa^\kappa$.

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

Théorème 2.26. (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. $\square$.

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

Définition 2.27. *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.

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 2.28. 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

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

Définition 2.29. Soit β un ordinal et $(\kappa_\alpha)_{\alpha<beta}$ une suite de cardinaux indexée par λ. On définit $\sum_{\alpha<\beta}\kappa_\alpha$ comme l'unique cardinal équipotent à $\bigsqcup \kappa_\alpha$.
De même, $\prod_{\alpha<\beta}\kappa_\alpha$ est l'unique cardinal équipotent au produit cartésien des κ_α.

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

Proposition 2.30. *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 fixons aussi [viii] une famille $(f_\alpha)_{\alpha<\kappa}$ telle pour tout $\alpha < \kappa$, f_α soit une bijection 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. □

2.4 Cardinaux réguliers et cofinalité

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

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

Définition 2.31. 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 2.32. Soit $(X, <)$ un ensemble totalement ordonné. On dit qu'une partie Y de X est *cofinale* dans X si tout élément de X a un majorant dans Y : $\forall x \in X\, \exists y \in Y\, x \leq y$.

La *cofinalité* d'un ordinal limite α est le plus petit ordinal β qui est équipotent à une partie cofinale de α. On note alors $\beta = \mathrm{cof}(\alpha)$. Remarquons que $\mathrm{cof}(\alpha)$ correspond au plus petit cardinal d'une partie non majorée de α, et qu'un cardinal infini κ est régulier si et seulement si $\mathrm{cof}(\kappa) = \kappa$.

On ne s'intéresse qu'à la cofinalité des ordinaux limites : à chaque fois qu'on écrit $\mathrm{cof}(\alpha)$, on suppose implicitement que α est un ordinal limite.

Les propriétés suivantes de cette notion seront vérifiées en exercices.

Proposition 2.33. *Soit α un ordinal limite.*

1. $\mathrm{cof}(\alpha)$ *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. $\mathrm{cof}(\alpha)$ *est le plus petit ordinal γ tel qu'il existe une fonction strictement croissante $f\colon \gamma \to \alpha$ dont l'image ne soit pas (strictement) majorée. En particulier, il existe une suite strictement croissante $(\lambda_i)_{i<\mathrm{cof}(\alpha)}$ d'éléments de α telle que $\alpha = \sup\{\lambda_i \colon i < \mathrm{cof}(\alpha)\}$.*
3. *On a* $\mathrm{cof}(\mathrm{cof}(\alpha)) = \mathrm{cof}(\alpha)$.

Proposition 2.34. *Soit κ un cardinal infini.*

1. *Si κ est un cardinal limite, il existe une suite strictement croissante $(\kappa_i)_{i<\mathrm{cof}(\kappa)}$ de cardinaux telle que $\kappa = \sup\{\kappa_i \colon i < \mathrm{cof}(\kappa)\}$.*
2. *Toute partie de κ de cardinal κ est cofinale dans κ.*
3. $\mathrm{cof}(\kappa)$ *est le plus petit ordinal γ tel que κ soit la réunion de γ ensembles de cardinal strictement inférieur à κ.*
4. *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 $|\cup X_\alpha| < \kappa$.*

Proposition 2.35. *Tout cardinal infini et successeur est régulier.*

Preuve. Fixons un cardinal infini et 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. $\square$

On peut, après la proposition ci-dessus, avoir l'impression que tout cardinal est régulier : mais les cardinaux limites sont bien souvent singuliers. C'est par exemple le cas de $\aleph_\omega$, qui est une union dénombrable d'ensembles de cardinal strictement plus petit que lui. On est amené à se poser la question suivante : existe-t-il un cardinal différent de $\aleph_0$ qui soit à la fois régulier et limite ? Un tel cardinal est dit *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 2.36. 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 [ix] 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 nos capacités dans ce cours.

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

Lemme 2.37. *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'ensembles 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 un élément $a_i \in A_i$, et 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

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

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

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 $|B_i| = \kappa_i < \lambda_i = |A_i|$, l'ensemble $Y_i = A_i \setminus \pi_i(B_i)$ 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 2.38. En utilisant le fait que tout réel admet un développement décimal[x], prouver à l'aide d'un raisonnement diagonal que $\mathbb{R}$ n'est pas dénombrable.

Corollaire 2.39. 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.

Corollaire 2.40. 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.

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

Chapitre 3

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.

3.1 Définitions, premières propriétés

Définition 3.1. 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 3.2. Soit $\mathcal{F}$ un filtre contenant une partie *finie* A. Alors $\mathcal{F}$ est principal.

Définition 3.3. 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 [i].

Proposition 3.4. *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 [ii]. 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 (précisément parce que I est infini !).

Définition 3.5. 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.

i. on trouve aussi la terminologie « base de filtre » pour désigner une famille de parties ne contenant pas l'ensemble vide et stable par intersections finies

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

Proposition 3.6. *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 $\mathcal{F}$. $\square$

Exercice 3.7. Soit $\mathcal{U}$ un ultrafiltre sur X. Montrer que soit $\mathcal{U}$ contient le filtre de Fréchet sur X, soit $\mathcal{U}$ est principal.

3.2 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 3.8. 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 3.9. 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 3.10. 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.

Une autre façon équivalente de définir le filtre image : 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 le filtre image est bien un filtre.

Proposition 3.11. *Le filtre image d'un ultrafiltre sur X est un ultrafiltre sur Y.*

Preuve.
Soit X, Y deux ensembles, $f: X \to Y$ 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. $\square$

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

Proposition 3.12. *Soit X, Y deux espaces topologiques, $x \in X$ et $f: X \to Y$ 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. $\square$

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 $\pi_i : Y \to X_i$ 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 3.13. *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 $\pi_i : X \to X_i$ 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

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

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. $\square$

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

Proposition 3.14. *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,*

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

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. $\square$

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 3.15. *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 3.14 dit exactement que $\mathcal{U}$ converge vers x. $\square$

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

Théorème 3.16. *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.

iii. Avec l'axiome du choix !

Index