Chapitre 1

Ordinaux, Cardinaux, Axiome du Choix

Ce cours traitera de théorie descriptive des ensembles, qui est une forme de "combinatoire infinie" ; avant de pouvoir faire de la combinatoire, il faut déjà apprendre à compter. On va donc discuter quelques notions élémentaires de théorie des ensembles avant de s'attaquer au sujet du cours proprement dit. Je n'essaierai pas de présenter le formalisme général de la théorie des ensembles ; 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.

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 considérer des ensembles munis d'un ordre permettant une énumération.

1.1 Bons ordres et ordinaux

Définition 1.1.1. 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.
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 $\{y \in S \colon y < x\}$.

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 ensemble bien ordonnés est fondé sur le résultat suivant :

Proposition 1.1.2. *Soit $(X, \leq)$ un ensemble bien ordonné et $f \colon X \to X$ une application strictement croissante. Alors pour tout $x \in X$ on a $f(x) \geq x$.*

Preuve.
Supposons qu'il existe $x \in X$ tel que $f(x) < x$, et appelons x_0 le plus petit élément ayant cette propriété. Alors on a, pour tout $x < x_0$, $f(x) \geq x$.
Puisque f est strictement croissante, on en déduit que pour tout $x < x_0$ on a $f(x_0) > x$.
Comme l'ordre $\leq$ est total, cela implique en particulier que $f(x_0) \geq x_0$, ce qui contredit la définition de x_0. □

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

Proposition 1.1.3. *Soit $(X, \leq)$ un ensemble bien ordonné, $W \subseteq X$ un segment initial et $f \colon X \to W$ un isomorphisme. Alors $W = X$ et pour tout $x \in X$ on a $f(x) = x$.*
Par conséquent, si deux segments initiaux de X sont isomorphes alors ils sont égaux.

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

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

Remarquons que le théorème 1.1.3 entraîne que $X \sim X'$ si, et seulement si, $X \preceq X'$ et $X' \preceq X$.

On a dit qu'on souhaitait pouvoir enumérer tous les ensembles bien ordonnés ; mais quelle notion de "longueur" utiliser ?

Théorème 1.1.4. *Soit X, Y deux ensembles bien ordonnés. Alors une et une seule des assertions suivantes est vraie :*

(a) $X \prec Y$

(b) $Y \prec X$

(c) $X \sim Y$

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

Preuve.
Notons $\tilde{X}$ l'ensemble des segments initiaux de X, ordonné par l'inclusion. On vérifie facilement que c'est un ensemble bien ordonné. Si tout $S \in \tilde{X}$ est isomorphe à un segment initial de Y alors c'est en particulier le cas de X, et la preuve est finie. Sinon, appelons S le plus petit élément qui ne soit pas isomorphe à un segment initial de Y.
Soit $x < x' \in S$. Alors $f_y \circ f_x^{-1}(f_x(S_x)) = f_y(S_x)$, et comme $f_y \circ f_x^{-1}$ est un isomorphisme entre deux segments initiaux de Y on en déduit que $f_y \circ f_x^{-1}(y) = y$ pour tout $y \in f_x(S_x)) = f_y(S_x)$.
Si jamais il existe $x \in S$ tel que $f(\{x' \colon x' \leq x\}) = f(S_x) = Y$ alors il n'y a rien à démontrer ; sinon pour tout $x \in S$ il existe une injection croissante $f_x \colon S_x \to Y$ d'image un segment initial de Y. Mais alors, comme $S = \bigcup_{x \in S} S_x$, on peut utiliser l'observation précédente 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. □

Corollaire 1.1.5. Soit X, Y deux ensembles bien ordonnés. Alors $X \preceq Y$ si, et seulement si, il existe une injection croissante de X dans Y.

Théorème 1.1.6. *Soit $\mathcal{W} = \{W_i \colon i \in I\}$ une famille d'ensembles bien ordonnés. Alors il existe $W \in \mathcal{W}$ tel que $W \preceq W'$ pour tout $W' \in \mathcal{W}$.*

Preuve.
Soit $W_0 \in \mathcal{W}$. Si $W_0 \preceq W'$ pour tout $W' \in \mathcal{W}$, il n'y a rien à démontrer. Sinon, l'ensemble $\{x \in W_0 \colon S_x \text{ est isomorphe à un élément de } \mathcal{W}\}$ 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. Mais une définition formelle pose quelques difficultés métamathématiques, qui ne correspondent pas à nos préoccupations dans ce cours. On va donc se contenter d'une présentation intuitive.
L'idée est que les ordinaux doivent permettre de "représenter" les ensembles bien ordonnés, au sens où tout ordinal soit un ensemble bien ordonné et pour tout ensemble bien ordonné il y ait un ordinal unique qui lui soit isomorphe ; c'est cet ordinal-là qui doit représenter la "longueur" d'un ensemble bien ordonné. Admettons que cela soit posssible (et pensons donc intuitivement à un ordinal comme à une classe d'isomorphisme d'ensembles bien ordonnés).
Allons plus loin et notons que si α est un ordinal, alors tout ordinal plus petit que α est isomorphe à un (unique) segment initial de α ; et les segments initiaux stricts de α s'identifient naturellement aux éléments de α.
On a donc envie d'identifier les ordinaux strictement inférieurs à α aux éléments de α, et donc d'effectuer notre choix de représentants de classes d'isomorphisme de bons ordres de telle façon que chaque ordinal α soit égal à l'ensemble des ordinaux strictement inférieurs à α.
Ceci impose une contrainte : si $\beta < \alpha$ on doit en même temps identifier les ordinaux strictement inférieurs à β aux éléments de β, ce qui amène à vouloir que l'ensemble des éléments strictement inférieurs à β (c'est-à-dire β) soit contenu dans α. Finalement, on a donc envie que tout élément d'un ordinal soit en fait *inclus* dans cet ordinal.
On n'est toujours pas tout à fait satisfait : si on a une famille d'ordinaux, alors on voudrait pouvoir "compter strictement plus loin" que tous les ordinaux de cette famille, ce qui imposerait que la réunion de notre famille d'ordinaux soit un ordinal. On rajoute cela dans les conditions qu'on demande aux ordinaux.

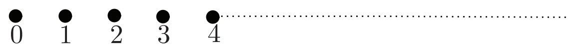
Voilà, on sait maintenant quelles propriétés attendre d'un ordinal, et on sait même comment effectuer leur construction : en effet, il n'y a pas d'élement plus petit que 0, donc 0 doit être l'ensemble vide. De même, $1 = \{0\} = \{\emptyset\}$, pour tout ordinal fini (i.e tout entier naturel !) on doit avoir $n = \{0, 1, \ldots, n-1\}$, etc. On va se contenter d'admettre qu'une telle construction des ordinaux est possible dans le cadre de la théorie axiomatique de Zermelo-Fraenkel, et reprendre le fil de ce cours.

Les relations $\prec$, $\preceq$ correspondent à des opérations sur les ordinaux notées cette fois $<, \leq$. On notera maintenant ON la classe[1] des ordinaux ; on utilisera dans la suite (entre autres) les propriétés suivantes des ordinaux :

– Tout ensemble bien ordonné est isomorphe à un ordinal unique.
– Pour tout ordinal α, on a $\alpha = \{\beta \in ON \colon \beta < \alpha\}$.
– La réunion (resp. l'intersection) d'un ensemble d'ordinaux est un ordinal.

Avant de continuer, introduisons un peu de terminologie.

Définition 1.1.7. Un ordinal α est *successeur* s'il existe un ordinal β tel que $\beta < \alpha$ et pour tout ordinal γ on ait soit $\gamma \leq \beta$ soit $\alpha \leq \gamma$; sinon on dit que α est un *ordinal limite* .
On notera ω le plus petit ordinal infini, qui est aussi le plus petit ordinal limite.



L'ordinal ω

Vous êtes habitués 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 1.1.8. *(Démonstration par récurrence transfinie)*
Soit P une propriété[2] 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.

1. On dit classe parce que ce n'est pas un ensemble, sinon sa réunion serait un ordinal plus grand que tous les ordinaux, or tout ordinal a un successeur, ce qui fait un point commun entre les ordinaux et les hommes politiques (j'espère)

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

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

Théorème 1.1.9. *Soit $(X, \leq)$ un ensemble bien ordonné, Y un ensemble, et $\mathcal{F}$ l'ensemble de toutes les fonctions dont le domaine est un segment initial de X et dont l'image est contenue dans Y. Pour toute fonction $G \colon \mathcal{F} \to Y$, il existe une unique fonction $f \colon X \to Y$ telle que l'on ait, pour tout $x \in X$,*

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

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

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

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

Cette relation est à nouveau une relation d'équivalence, qui étend $\sim$ et est compatible avec $\leq$.
Soit maintenant ξ un ordinal quelconque. Pour $\alpha < \xi$, on définit une relation d'équivalence $\sim_\alpha$ compatible avec $\leq$ par récurrence transfinie, en respectant les trois points suivants :
(a) $(x \sim_0 y) \Leftrightarrow (x = y)$
(b) Si α est le successeur de β, alors $\sim_\alpha = D(\sim_\beta)$.
(c) Si $\alpha = \sup(\{\beta \colon \beta < \alpha\})$ alors $\sim_\alpha = \bigcup_{\beta < \alpha} \sim_\beta$.
Intuitivement, on a "épluché X ξ fois" : on a commencé par identifier tous les points tels que $[x, y]$ est fini et formé ainsi un nouvel ensemble ordonné, auquel on a appliqué la même construction, et on a répété le procédé pendant ξ étapes.

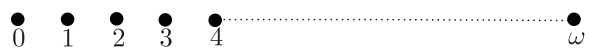
Par exemple, si on applique cette construction à $\mathbb{N}$ muni de son ordre usuel, on a $\sim_1 = X \times X$; par contre, si on l'applique à $\mathbb{Q}$ muni de son ordre usuel, on a $\sim_1 = \sim_0$ et donc $\sim_\alpha = \sim_0$ pour tout ordinal α.
On est amené à se poser un certain nombre de questions : est-ce qu'on peut continuer à éplucher X indéfiniment sans jamais s'arrêter ? Au contraire, est-ce que X est "épluchable", autrement dit ne reste-il plus rien au bout d'un nombre assez grand d'étapes ? Ou tombe-t-on sur un noyau, c'est-à-dire est-ce que $\sim_n$ arrête de grossir au bout d'un moment ? Nous reviendrons sur ces questions après l'introduction des cardinaux.

Exercice 1.1.10. $\star$ Montrer que si X n'est pas épluchable du tout, c'est-à-dire si $\sim_1 = \sim_0$, alors l'ordre de X est dense.
En déduire avec la méthode de va-et-vient que le seul ordre dénombrable totalement non épluchable est $\mathbb{Q}$, avec peut-être un plus grand et un plus petit élément. En déduire que si un ordre dénombrable contient un noyau alors il contient un sous-ensemble isomorphe à $\mathbb{Q}$ (en utilisant un peu de théorie des cardinaux, on verra qu'un ordre dénombrable doit soit être épluchable soit contenir un noyau).

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. On note $s(\beta)$ le successeur d'un ordinal β, c'est-à-dire le plus petit ordinal strictement plus grand que β.

Définition 1.1.11. (addition ordinale) Soit α un ordinal. On pose $\alpha + 0 = 0$, puis on définit par récurrence transfinie sur $\beta \in ON$ l'addition ordinale $\alpha + \beta$ en posant :

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



L'ordinal $\omega + 1$

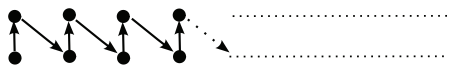
Par exemple, on a $1 + \omega = \sup\{1 + n \colon n < \omega\} = \omega$. Intuitivement, l'addition de deux ordinaux correspond à mettre "bout à bout" α et β.

Exercice 1.1.12. Utiliser une démonstration par récurrence transfinie pour montrer que l'addition est associative, et que si $\alpha \neq \beta$ alors pour tout δ on a $\delta + \alpha \neq \delta + \beta$.

Définition 1.1.13. (multiplication ordinale) Soit α un ordinal. On pose $\alpha.0 = 0$, puis on définit par récurrence transfinie sur $\beta \in ON$ la multiplication ordinale $\alpha.\beta$ en posant :

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

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

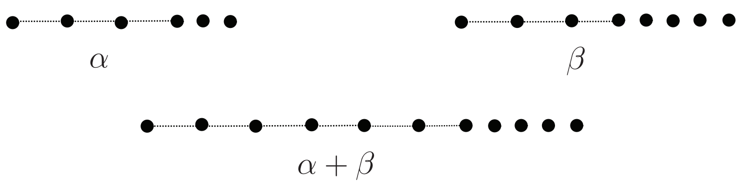


$2.\omega = \omega$

Exercice 1.1.14. Utiliser une démonstration par récurrence trasnfinie pour montrer que la multiplication est associative, et que si $\alpha > 0$ alors pour tout $\gamma > 1$ on a $\alpha. < \alpha.\gamma$.

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

Exercice 1.1.15.
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

Exercice 1.1.16. Comment définiriez-vous α^β pour deux ordinaux α, β ? Donner d'abord une description par récurrence transfinie, puis essayer de décrire une opération sur les ordres qui donne naissance à cette opération (ce n'est pas si facile !)

Exercice 1.1.17. On reprend les notations introduites lors de la définition de la dérivation de Hausdorff. Montrer que, si on prend un ordinal dénombrable α et qu'on considère $X = \omega^\alpha$, alors $\sim_\alpha$ est la relation grossière.

On n'utilisera pratiquement pas dans la suite d'arithmétique des ordinaux ; un exercice pour s'entraîner à la récurrence transfinie :

Exercice 1.1.18. 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.

1.2 Cardinaux et axiome du choix.

On a vu comment énumérer des ensembles bien ordonnés ; par contre, contrairement aux ensembles finis, un ensemble peut 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 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.

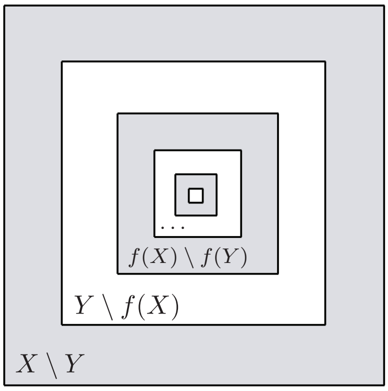
Définition 1.2.1. On dit que X a un cardinal inférieur à Y, et on note $|X| \leq |Y|$, s'il existe une injection de X dans Y, et on dit que X et Y ont même cardinal, ou sont équipotents (noté $|X| = |Y|$), s'il existe une bijection de X sur Y.

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

Théorème 1.2.2. *(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.
On voit donc qu'il suffit de prouver que, si X est un ensemble, $f : X \to X$ 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 $g : X \to Y$ 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. De plus, s'il existe une injection de X dans Y alors il existe une surjection de Y sur X, donc toutes nos définitions possibles pour comparer les cardinaux d'ensembles s'entendent bien, et on obtient ainsi un quasi-ordre sur les ensembles. Notre préoccupation maintenant est de savoir si deux ensembles sont nécessairement comparables pour ce quasi-ordre.

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

Exercice 1.2.3. Montrer (sans utiliser l'axiome de Zermelo) que pour tout ensemble X il existe un plus petit ordinal non équipotent à une partie de X. Montrer que cet ordinal est en fait un cardinal, appelé *cardinal de Hartogs* de X

Exercice 1.2.4. Donner une autre preuve du théorème de Schröder-Bernstein dans le cas où X et Y sont des ensembles bien ordonnables.

Les discussions ci-dessus servaient de prologue à la définition suivante.

Définition 1.2.5. 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 $\kappa(X)$ et que de plus si α est un ordinal alors $\kappa(\alpha) \leq \alpha$. Par exemple, tous les ordinaux finis sont des cardinaux, ainsi que ω ; par contre, $\omega + 1$ n'est pas un cardinal. Notons également que par définition deux cardinaux distincts ne peuvent pas être équipotents. Enfin, dans la suite, si X est un ensemble bien ordonnable, la notation $|X|$ désignera l'unique cardinal équipotent à X.
Il existe pour tout κ des ordinaux qui ne sont pas équipotents à une partie de κ, et donc des cardinaux λ tels que $\kappa < \lambda$. On note κ^+ le plus petit tel cardinal.

Définition 1.2.6. (Alephs)
On définit par récurrence transfinie $\aleph_\alpha$, pour tout ordinal α, en posant $\aleph_0 = \omega$ puis

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

Il est alors possible de vérifier assez facilement que tous les $\aleph_\alpha$ sont des cardinaux, et réciproquement (avec l'axiome du choix !) que tout cardinal est égal à $\aleph_\alpha$ pour un certain α. De plus $\alpha \leq \aleph_\alpha$ pour tout ordinal α.

Exercice 1.2.7.
Montrer qu'il existe des ordinaux (donc des cardinaux !) tels que $\alpha = \aleph_\alpha$.

Les cardinaux nous permettent de prouver que soit un ensemble ordonné dénombrable est épluchable, soit il contient un noyau (et donc un sous-ensemble isomorphe à $(\mathbb{Q}, \leq)$).

Exercice 1.2.8. Soit $(X, <)$ un ensemble dénombrable. On note $\sim_\alpha$ la α-ième relation d'équivalence apparaissant dans la dérivation de Hausdorff de X ; montrer qu'il existe $\alpha \leq \omega_1$ tel que $\sim_\alpha = \sim_{\alpha+1}$.

Pour définir une notion satisfaisante du cardinal d'un ensemble, on a eu besoin d'utiliser le théorème de Zermelo. Cet axiome 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.* L'énoncé intuitif de cet axiome, 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".
Avant de citer trois énoncés équivalents de l'axiome du choix , rappelons qu'un ensemble ordonné $((X, \leq)$ est *inductif* si tout sous-ensemble totalement ordonné admet un majorant. Nous dirons aussi qu'un ensemble X admet une *fonction de choix* s'il existe une fonction $f \colon \mathcal{P}(X) \to X$ telle que pour toute partie $A \subseteq X$ non vide on ait $f(A) \in A$.

Définition 1.2.9. On introduit les énoncés suivants :

1. (Axiome du choix) Tout ensemble X admet une fonction de choix.
2. (Axiome de Zorn) Tout ensemble ordonné inductif non vide a au moins un élément maximal.
3. (Axiome 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 nombreux mathématiciens du dix-neuvième siècle et paraît plutôt "naturel". il est plus difficile de se faire une idée intuitive du second énoncé, dont on voit qu'il permet d'une certaine façon de faire de l'analyse en évitant la théorie des ordinaux et 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[3] : "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. Alors tout sous-ensemble totalement ordonné $M \subseteq X$ admet un majorant strict noté $\psi(M)$. Soit alors φ une fonction de choix sur X et 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) $x_{\alpha+1} = \varphi(\{y \in X \colon y > x_\alpha\})$;

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

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

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

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. (REFS)
Par ailleurs, l'axiome du choix a de nombreuses conséquences qui, elles, paraissent très utiles : théorème de la bae 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 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 1.2.10. 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 utlisera sans vergogne l'axiome du choix sous ses différentes formes. En fait, comme on ne considèrera jamais d'ensembles de très grand cardinal, 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 recquiert l'axiome du choix dénombrable... Plus élémentairement, on peut noter le fait suivant, qui peut paraître surprenant pour certains qui l'utilisent tout en récusant l'axiome du choix :

Exercice 1.2.11. On dit qu'un ensemble infini est *Dedekind-infini* s'il contient un sous-ensemble infini dénombrable [4]. Montrer que l'axiome du choix dénombrable est équivalent à l'énoncé suivant :
Tout ensemble infini est Dedekind-infini ; de plus, toute réunion dénombrable d'ensembles dénombrable est dénombrable..

En réalité, l'axiome du choix dénombrable n'est pas suffisant pour les raisonnements que nous devrons effectuer dans ce cours.

4. ce n'est pas la définition habituelle, mais au-dessus de ZF cet énoncé est équivalent à la définition classique.

En effet, on aura besoin de construire des suites en utilisant le principe suivant : supposons qu'étant donné $x_1, \ldots, x_n$ 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 toutes les constructions par "approximation successives" et devient légal quand on s'autorise à appliquer l'*axiome des choix dépendants*.

Définition 1.2.12. 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.

Avant de présenter une autre forme de l'axiome des choix dépendants, disons qu'un ensemble T muni d'une relation binaire R antisymétrique, sans cycles et connexe (i.e deux points sont reliés par une suite d'éléments en relation R) est un *arbre* . Une *branche infinie* de T est une suite injective (x_n) d'éléments de T telle que $x_n R x_{n+1}$ pour tout n.

Exercice 1.2.13. Montrer que l'axiome des choix dépendants est équivalent à l'énoncé suivant :
Soit T un arbre infini tel que chaque élément a un nombre fini de voisins. Alors T a une branche infinie.

Pour nous, l'axiome des choix dépendants sera particulièrement important, puisqu'il se trouve en fait être équivalent (au dessus de ZF) à l'énoncé "le théorème de Baire est vrai dans tout espace complet", et le théorème de Baire (dont l'énoncé sera rappelé au prochain chapitre) est d'une certaine façon la pierre angulaire de la théorie descriptive des ensembles.
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 ; on peut considérer que cet axiome est accepté par une grande majorité des mathématiciens contemporains.

Revenons à nos cardinaux ; à l'aide de l'axiome du choix, on peut vérifier que si $(X_\alpha)_{\alpha<\lambda}$ et $(Y_\alpha)_{\alpha<\lambda}$ sont tels que $|X_\alpha| = |Y_\alpha|$ pour tout $\alpha < \lambda$ alors on a

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

De même si $|X| = |X_1|$ et $|Y| = |Y_1|$ alors $\left|X^Y\right| = \left|X_1^{Y_1}\right|$. On peut alors définir les opérations arithmétiques cardinales :

Définition 1.2.14. (arithmétique cardinale) Soit α un ordinal et $(\kappa_\alpha)_{\alpha<\lambda}$ une suite de cardinaux indexée par un ordinal λ. On définit $\sum_{\alpha<\lambda} \kappa_\lambda$ comme l'unique cardinal équipotent à $\bigsqcup \kappa_\alpha$.
De même, $\prod_{\alpha<\lambda} \kappa_\lambda$ est l'unique cardinal équipotent à $\bigsqcup \kappa_\alpha$.
Enfin, si κ et λ sont deux cardinaux on définit κ^λ comme l'unique cardinal équipotent à l'ensemble des fonctions de λ dans κ ; en particulier $\left|2^X\right| = |\mathcal{P}(X)|$ pour tout ensemble X.

La somme et le produit de cardinaux sont des opérations commutatives et associatives ; attention au fait que la sommme/produit de deux cardinaux diffère selon qu'on les considère comme des cardinaux ou comme des ordinaux...

Exercice 1.2.15. Montrer que $|\mathbb{Q}^2| = \aleph_0$, puis montrer que $|\mathbb{R}| = 2^{\aleph_0}$. Pour la dernière question, on pourra considérer l'application $f\colon \mathbb{R} \to \mathcal{P}(\mathbb{Q}^2)$ définie par

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

L'addition et la multiplication d'un nombre fini de cardinaux sont simples à comprendre, comme le montre le théorème suivant.

Théorème 1.2.16.
Soit κ,λ deux cardinaux infinis. Alors on a $\kappa + \lambda = \kappa.\lambda = \max(\kappa, \lambda)$

Preuve.
Il suffit de montrer que pour tout cardinal infini κ on a $\kappa.\kappa = \kappa$, autrement dit il nous faut montrer que tout ensemble infini X est équipotent à $X \times X$ [5]
On raisonne par récurrence transfinie : soit donc κ un cardinal infini, dont on suppose que pour tout cardinal infini $\lambda < \kappa$ on a $\lambda.\lambda = \lambda$.
On va munir $\kappa \times \kappa$ d'un bon ordre dont tous les segments initiaux stricts se plongent dans κ. Ce bon ordre sera isomorphe à un certain ordinal, qui est alors plus petit que κ (un ordinal est la réunion de ses segments initiaux), et permettra de définir une injection de $\kappa \times \kappa$ dans κ.
Pour définir notre bon ordre, on pense à l'énumération classique de $\mathbb{N} \times \mathbb{N}$, et on pose, pour $(\alpha, \beta), (\alpha', \beta') \in \kappa^2$:

$$\big((\alpha, \beta) \preceq (\alpha', \beta')\big) \Leftrightarrow \big((\alpha + \beta < \alpha' + \beta') \text{ ou } (\alpha + \beta = \alpha' + \beta' \text{ et } \alpha \leq \alpha')\big)$$

On vérifie que ceci définit bien un bon ordre sur κ^2 ; de plus, si $(\alpha, \beta) \in \kappa^2$ alors le $\preceq$-segment initial associé à (α, β) se plonge dans $(\alpha + \beta) \times (\alpha + \beta)$.

5. voilà encore un énoncé équivalent à l'axiome du choix !

Par notre hypothèse de récurrence, et puisque α, β sont de cardinal $< \kappa$, on obtient que $|\alpha + \beta|$ est soit fini soit infini strictement inférieur à κ ; par conséquent, $|(\alpha \times \beta) \times (\alpha \times \beta)| < \kappa$.
On voit donc que tous les segments initiaux du bon ordre $\preceq$ sur $\kappa \times \kappa$ sont de cardinal strictement inférieur à κ, ce qui termine la preuve. $\square$

D'une certaine façon, l'opération arithmétique la plus mystérieuse sur les cardinaux est l'exponentiation ; le premier résultat à son sujet a été obtenu par Cantor.

Théorème 1.2.17. *(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

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

On doit avoit $Y = f(x_0)$ 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 $\kappa \mapsto 2^\kappa$. Y a-t-il des cardinaux qui ne sont pas de cette forme ?

Définition 1.2.18. *L'hypothèse du continu* 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 ardinal 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.
Âujourd'hui, la plupart des théoriciens des ensembles considèrent qu'il n'y a aucune raison de limiter la richesse de la théorie en imposant arbitrairement que l'hypothèse du continu soit vérifiée ; il existe des axiomes ("grands cardinaux") menant à une théorie très riche dans laquelle l'hypothèse du continu est fausse.

On n'en dira pas plus sur l'hypothèse du continu dans ce cours ; pas plus qu'on n'aura besoin de notions supplémentaires d'arithmétique cardinale. Notons cependant que dans le cadre de ZFC peu de théorèmes "simples" peuvent être établis, car beaucoup d'énoncés sur les cardinaux se trouvent être indépendants de ZFC. Mais nous n'allons pas manipuler des ensembles quelconques, mais des ensembles munis d'une topologie "sympathique", dont nous ne manipulerons que des sous-ensembles "définissables".

Notes bibliographiques.
Ce chapitre ne saurait constituer une introduction complète à la théorie des ensembles élémentaire ; à ce sujet, le lecteur intéressé est invité à consulter [Hal74] s'il cherche une présentation intuitive de la théorie, et [Mos06] ou [KM] pour une présentation plus formelle et orientée vers les sujets qui apparaîtront dans les chapitres suivants de ces notes.
Le lecteur anglophobe souhaitant se documenter sur le sujet pourra consulter avec profit la traduction française du livre de Kuratowski sus-cité.
En ce qui concerne l'axiome du choix, il existe une véritable encyclopédie [HR98] présentant ses multiples formes ; on pourra y trouver des références sur certains résultats énoncés sans référence dans le corps du chapitre ci-dessus. Le livre de S. Wagon [Wag85] est également très instructif.
Enfin, comme source bibliographique concernant les résultats plus récents de théorie des ensembles (forcing, etc) le lecteur est invité à consulter [Jec03].