Premier exposé : rappels de logique et de théorie des ensembles

David A. Madore

11 octobre 1998

Table des matières

1 Rappels de logique du premier ordre **1**

2 Logique du second ordre **2**

3 Théorie des ensembles **4**
3.1 Les axiomes de ZF . . . 4
3.2 Ensembles bien ordonnés et ordinaux . . . 5
3.3 Cardinalités . . . 8
3.4 L'axiome du choix . . . 9
3.5 Exponentiation cardinale . . . 10
3.6 L'axiome de régularité . . . 11

4 Modèles de la théorie des ensembles **12**

1 Rappels de logique du premier ordre

Sans donner de définition rigoureuse, rappelons simplement qu'une formule du premier ordre dans un langage L donné est formée à partir des connecteurs du calcul propositionnel (qui sont $\Rightarrow$, $\wedge$, $\top$, $\vee$ et $\bot$), des quantificateurs $\forall$ et $\exists$, de variables (dites variables d'objets), et des symboles du langage, qui peuvent être des symboles de constantes, de fonctions ou de relations.

Une L-structure est un ensemble M muni d'un élément c_i pour chaque symbole de constante $\bar{c}_i$ du langage L, d'une fonction $f_i\colon M^n \to M$ pour chaque symbole de fonction $\bar{f}_i$ d'arité n du langage, et d'une relation $R_i \subseteq M^n$ pour chaque symbole de relation $\bar{R}_i$ d'arité n du langage. Lorsqu'on a une telle L-structure M, on peut alors définir une notion de satisfaction des formules du premier ordre, par induction sur la complexité de celles-ci. Par exemple, on a $M \vDash \forall x \exists y \bar{R}(x, y, \bar{c})$ (avec $\bar{R}$ une relation d'arité 3 du langage, et $\bar{c}$ une constante du langage) si et seulement si pour tout x de M il existe y de M tel que $R(x, y, c)$, où $R \subseteq M^3$ est la fonction qui représente le symbole $\bar{R}$, et $c \in M$ l'élément qui représente le symbole $\bar{c}$.

Par ailleurs, on a une notion de démonstration formelle, dont nous ne rappelons pas les mécanismes (ceux-ci seront étudiés en détail et beaucoup plus élégamment que de la façon "classique" dans la partie du groupe de travail consacrée à la logique catégorique). Voici cependant les règles fondamentales :

$$A \vdash A \text{ (identité)}$$

$$\frac{\Gamma \vdash A}{\Gamma, \Delta \vdash A} \text{ (affaiblissement)} \qquad \frac{\Gamma, C, C \vdash A}{\Gamma, C \vdash A} \text{ (rasoir)}$$

$$\frac{\Gamma \vdash A \qquad \Delta, A \vdash B}{\Gamma, \Delta \vdash B} \text{ (coupure)}$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \Rightarrow B} \text{ (théorème de déduction)} \qquad \frac{\Gamma \vdash A \Rightarrow B}{\Gamma, A \vdash B} \textit{ (modus ponens)}$$

$$\frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} (\wedge\text{intro}) \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} (\wedge\text{elimL}) \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B} (\wedge\text{elimR})$$

$$\Gamma \vdash \top \ (\top\text{intro})$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} (\vee\text{introL}) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} (\vee\text{introR}) \qquad \frac{\Gamma, A \vdash C \quad \Gamma, B \vdash C}{\Gamma, A \vee B \vdash C} (\vee\text{elim})$$

$$\Gamma, \perp \vdash A \ (\perp\text{elim})$$

$$\frac{\Gamma \vdash A(x)}{\Gamma \vdash \forall x(A(x))} (\forall\text{intro}, x \text{ pas libre dans } \Gamma) \qquad \frac{\Gamma \vdash \forall x(A(x))}{\Gamma \vdash A(t)} (\forall\text{elim})$$

$$\frac{\Gamma \vdash A(t)}{\Gamma \vdash \exists x(A(x))} (\exists\text{intro}) \qquad \frac{\Gamma, A(x) \vdash B}{\Gamma, \exists x(A(x)) \vdash B} (\exists\text{elim}, x \text{ pas libre dans } \Gamma \text{ ou } B)$$

$$\frac{\Gamma, A \Rightarrow \perp \vdash \perp}{\Gamma \vdash A} \text{ (par l'absurde)}$$

On a le théorème très important suivant :

Théorème 1.1 (de complétude). *Si Γ est un ensemble de formules closes du premier ordre dans un langage L et A une telle formule, alors $\Gamma \vdash A$ si et seulement si pour toute M-structure qui vérifie $M \vDash \Gamma$ on a $M \vDash A$.*

Une M-structure telle que $M \vDash \Gamma$ est appelée un modèle de Γ. Autrement dit, une formule A est un théorème de la théorie Γ si et seulement si elle est vraie dans tout modèle. Le "seulement si" signifie que la structure sémantique apportée par les modèles est pertinente à la logique étudiée, et le "si" signifie que cette structure sémantique (lorsqu'on considère *tous* les modèles) est complète.

D'autre part, on rappelle le

Théorème 1.2 (de compacité). *Une théorie Γ est consistante (i.e. $\Gamma \nvdash \perp$) si et seulement si toute partie finie de Γ est consistante.*

Et enfin le

Théorème 1.3 (de Löwenheim-Skolem). *Si $\mathfrak{M}$ est une structure d'un langage L, alors il existe un sous-modèle élémentaire $\mathfrak{M}_0 \subseteq \mathfrak{M}$ (c'est-à-dire que les formules du premier ordre avec constantes dans $\mathfrak{M}_0$ ont la même valeur de vérité dans $\mathfrak{M}_0$ et dans $\mathfrak{M}$) dont le cardinal est* $\operatorname{card} L + \aleph_0$.

La démonstration du théorème de Löwenheim-Skolem se fait de la façon suivante : pour toute formule $\varphi(x, x_1, \ldots, x_n)$ à $n+1$ variables libres, on trouve une fonction $h_\varphi \colon \mathfrak{M}^n \to \mathfrak{M}$ telle que, pour tous $a_1, \ldots, a_n$ de $\mathfrak{M}$, *si* $\mathfrak{M} \vDash \exists x \varphi(x, a_1, \ldots, a_n)$, alors $\mathfrak{M} \vDash \varphi(h_\varphi(a_1, \ldots, a_n), a_1, \ldots, a_n)$. Autrement dit, on choisit un témoin de chaque formule existentielle. On dit que h_φ est une fonction de Skolem pour la formule φ. On construit alors $\mathfrak{M}_0$ comme la plus petite partie de $\mathfrak{M}$ close sous l'opération de toutes les formules de Skolem.

Précisons un dernier point de terminologie : une partie A d'une structure $\mathfrak{M}$ d'un langage L est dite définissable si et seulement si il existe une formule $\varphi(x, x_1, \ldots, x_n)$ et des éléments $a_1, \ldots, a_n \in A$ tels que $a \in A$ si et seulement si $\mathfrak{M} \vDash \varphi(a, a_1, \ldots, a_n)$. On dit alors que A est définissable à partir de $a_1, \ldots, a_n$. On dit que $a \in \mathfrak{M}$ est définissable à partir de $a_1, \ldots, a_n$ ssi $\{a\}$ l'est.

2 Logique du second ordre

Dans la logique du second ordre, on autorise la quantification non seulement sur les objets mais aussi sur les prédicats (d'arité quelconque). Par exemple, le principe de récurrence, qui en logique du premier ordre doit se formuler comme une infinité d'axiomes (un pour chaque formule, ce qui revient à n'affirmer le principe que pour les parties définissables de $\mathbb{N}$), s'énonce en logique du second ordre comme

$$\forall P((P(0) \wedge \forall n(P(n) \Rightarrow P(Sn))) \Rightarrow (\forall n(P(n))))$$

La question est alors de savoir comment on définit un modèle en logique du second ordre. Lorsque l'on décide que la quantification sur les prédicats concerne toutes les parties du modèle, on parlera de modèle plein. Par exemple, la formule écrite ci-dessus s'interprète dans un modèle plein $\mathfrak{M}$ comme signifiant que toute partie A de $\mathfrak{M}$ qui contient l'élément qui représente le symbole 0 et contient l'image par la fonction qui représente le symbole S de tout élément qu'elle contient, est égale à $\mathfrak{M}$. Mais on peut aussi décider de considérer des modèles dans lesquels la quantification sur les prédicats ne concerne pas toutes les parties de $\mathfrak{M}$. Il faut alors rajouter, dans les données qui définissent une structure $\mathfrak{M}$, un ensemble de parties de $\mathfrak{M}$; le cas des modèles pleins est celui où cet ensemble est justement $\mathscr{P}(\mathfrak{M})$.

D'autre part, il faut se demander comment la notion de démonstration formelle se transpose en logique du second ordre. Les règles habituelles restent bien entendu valables ; en particulier, si on a démontré une formule du type $\exists P(\cdots)$, avec P une variable de prédicat, on peut la concrétiser et "choisir un tel P". La règle $\exists$elim ne pose donc pas de problème, pas plus que la règle $\forall$intro (qui ne peut, bien sûr, que s'appliquer à des variables de prédicats). Les règles $\forall$elim et $\exists$intro, en revanche, posent plus de problèmes sur les variables du second ordre, car on ne sait pas quoi prendre comme "termes" de prédicats. On veut bien sûr pouvoir utiliser n'importe quelle formule ayant une variable-objet libre à cette place. La solution est alors de formuler un schéma d'axiomes (comme on fait en logique du premier ordre pour le principe de récurrence), le schéma d'axiomes de compréhension : pour toute formule libre $\varphi(x, x_1, \ldots, x_n)$, on postule

$$\forall t_1 \cdots \forall t_n \exists P \forall x (P(x, x_1, \ldots, x_n) \Leftrightarrow \varphi(x, x_1, \ldots, x_n))$$

Une façon légèrement différente, et néanmoins équivalente, de procéder, consiste à autoriser la formation de termes de prédicats (plutôt que simplement postuler leur existence, comme ci-dessus) avec une notation λ : si on a une formule $\varphi(x, x_1, \ldots, x_n)$, on peut former un terme $\lambda x. \varphi(x, x_1, \ldots, x_n)$ qui est un prédicat.

On se doute bien qu'avec un tel système de démonstration, les théories du deuxième ordre ne vont pas être plus puissantes que les théories du premier ordre correspondantes. Cependant, elles sont beaucoup plus expressives, puisqu'elles permettent de parler non seulement d'éléments de leur domaine, mais aussi de parties de celui-ci.

Les relations entre démonstrations et sématique sont plus complexes en logique du deuxième ordre qu'en logique du premier ordre. Avant toute chose, on note que tout modèle du deuxième ordre (plein ou non) vérifie l'axiome suivant

$$\forall P \forall Q((\forall x(P(x) \Leftrightarrow Q(x))) \Rightarrow P = Q)$$

dit axiome d'extension. On le prendra donc comme axiome de la logique du second ordre, à côté de l'axiome de compréhension.

Malgré cette addition, les modèles pleins continuent à ne pas former une structure sémantique complète pour la logique du second ordre. Pour le voir, considérons les axiomes de Peano du second ordre, c'est-à-dire la théorie sur le langage $(0, S)$ avec les axiomes suivants

$$\forall x(\neg(Sx = 0))$$

$$\forall x \forall y((Sx = Sy) \Rightarrow (x = y))$$

$$\forall P((P(0) \land \forall n(P(n) \Rightarrow P(Sn))) \Rightarrow (\forall n(P(n))))$$

Contrairement à ce qui se passe en logique du premier ordre, on n'a ici pas besoin de prendre $+$ et $\times$ comme symboles du langage pour capturer l'arithmétique : en effet, la formule $u + v = w$ peut se traduire par

$$\forall Q(((\forall x(Q(x, 0, x))) \land (\forall x \forall y \forall z(Q(x, y, z) \Rightarrow Q(x, Sy, Sz)))) \Rightarrow (Q(u, v, w)))$$

Cette arithmétique du second ordre a de remarquable par rapport à sa consœur du premier ordre qu'elle est catégorique : autrement dit, elle a un seul modèle plein (à isomorphisme près), et c'est

$\mathbb{N}$ (muni de son zéro et de sa fonction successeur habituels). En effet, si $(\mathfrak{M}, 0, S)$ est un modèle des axiomes écrits ci-dessus, on peut définir $\psi \colon \mathbb{N} \to \mathfrak{M}$ par $\psi(n) = S^n 0$. Il est facile de voir que ψ est un isomorphisme de $\mathbb{N}$ sur une sous-structure de $\mathfrak{M}$, et le seul point restant à montrer est que ψ est surjectif ; mais pour cela on applique le troisième axiome à l'image de ψ, et on a immédiatement le résultat. Quand on se rappelle qu'en logique du premier ordre toutes les théories catégoriques ont des modèles finis, cela a de quoi surprendre.

Quoi qu'il en soit, cela suffit à prouver que le théorème de complétude ne tient pas : *toute affirmation arithmétique vraie dans* $\mathbb{N}$ est vrai dans tous les modèles de l'arithmétique du second ordre (puisqu'il n'y en a qu'un justement, et c'est $\mathbb{N}$). Mais toutes sont loin d'être démontrables (ne serait-ce que parce que le système de démonstration que nous avons proposé est récursif, donc il ne peut certainement pas énumérer les affirmations vraies dans $\mathbb{N}$).

3 Théorie des ensembles

3.1 Les axiomes de ZF

La théorie des ensembles (ZF) est une théorie sur le langage ayant un seul symbole, une relation binaire notée $\in$.

Les axiomes sont :

Axiome 3.1 (d'extension). *Si X et Y ont les mêmes éléments alors $X = Y$.*

(La réciproque est conséquence des propriétés logiques de l'égalité.)

Axiome 3.2 (de la paire). *Pour tous x et y il existe un ensemble, noté $\{x, y\}$, dont les éléments sont x et y et seulement eux.*

Schéma axiome 3.3 (de séparation). *Soit $\varphi(x, x_1, \dots, x_n)$ une formule. Pour tout ensemble X et tous $a_1, \dots, a_n$ il existe un ensemble dont les éléments sont exactement les a tels que $a \in X$ et que $\varphi(a, a_1, \dots, a_n)$. On le note $\{x \colon \varphi(x, a_1, \dots, a_n)\}$.*

Il faut naturellement formuler un tel axiome pour chaque formule φ du premier ordre. En fait, on peut se contenter de formules ayant deux variables libres (puisqu'on peut coder tout n-uplet $(a_1, \dots, a_n)$ dans un seul ensemble). Si on formule ZF en logique du second ordre, on n'a alors plus naturellement qu'un seul axiome de séparation, qui quantifie sur tous les prédicats, et c'est l'axiome de compréhension qui fait le travail.

Axiome 3.4 (de l'union). *Pour tout X il existe un ensemble Y (noté $\bigcup X$) dont les éléments sont précisément les éléments des éléments de X.*

Axiome 3.5 (de l'ensemble des parties). *Pour tout X il existe un ensemble (noté $\mathscr{P}(X)$) dont les éléments sont précisément les parties de X.*

Axiome 3.6 (de l'infini). *La classe des naturels est un ensemble.*

Il y a beaucoup de façons de formuler cet axiome. On dit qu'un ensemble est inductif si et seulement si il contient $\varnothing$ et contient $x \cup \{x\}$ pour chacun de ses éléments x. Un ensemble est appelé entier naturel si et seulement s'il appartient à tout ensemble inductif. La classe des naturels ω est donc l'intersection de tous les ensembles inductifs. L'axiome de l'infini postule que cette classe est un ensemble. Cela revient en fait (grâce à l'axiome de séparation) à postuler qu'il existe un ensemble inductif.

Accessoirement, l'axiome de l'infini est la seule façon dont on dispose pour prouver l'existence d'un ensemble. Toutefois, si on considère la théorie des ensembles sans l'axiome de l'infini, on postulera explicitement l'existence d'un ensemble (ou, ce qui revient au même, de l'ensemble vide).

Schéma axiome 3.7 (de remplacement). *Soit $\varphi(x, y, x_1, \dots, x_n)$ une formule. Soient X et $a_1, \dots, a_n$ des ensembles. On suppose que pour tout $u \in X$ il existe* au plus un *ensemble v tel que $\varphi(u, v, a_1, \dots, a_n)$. Alors il existe un ensemble dont les éléments sont exactement les v tels que $\varphi(u, v, a_1, \dots, a_n)$ pour un certain $u \in X$.*

Axiome 3.8 (de régularité (ou de fondation)). *Si X est un ensemble non vide, alors il existe un élément Y de X tel que $Y \cap X = \varnothing$.*

En présence de l'axiome du choix dénombrable (voir plus loin), l'axiome de régularité équivaut à l'inexistence de suites $X_0 \ni X_1 \ni X_2 \ni \cdots$ d'ensembles. Une formulation beaucoup plus constructive de l'axiome de régularité, qui a l'avantage d'éviter les recours au tiers exclu, est la suivante

$$\forall X \forall X_0((\forall z \in X(z \cap X \subseteq X_0 \Rightarrow z \in X_0)) \Rightarrow X_0 \supseteq X)$$

Elle a l'inconvénient que les gens qui essayent de la comprendre finissent en général par se cogner la tête contre les murs. Pourtant, c'est un simple principe d'induction : si X_0 contient un élément z de X dès qu'il contient tous les éléments de z qui sont aussi dans X, alors X_0 contient tous les éléments de X. Lorsque l'ensemble X est transitif, ce qui signifie $\forall z \in X(z \subseteq X)$ (ou encore, $X \subseteq \mathscr{P}(X)$), alors le principe s'énonce plus simplement :

$$\forall X_0((\forall z \in X(z \subseteq X_0 \Rightarrow z \in X_0)) \Rightarrow X_0 \supseteq X)$$

L'axiome suivant ne fait pas partie de ZF mais de ZFC :

Axiome 3.9 (du choix). *Si $\mathfrak{X}$ est un ensemble ne contenant pas l'ensemble vide, alors il existe une fonction qui à chaque élément X de $\mathfrak{X}$ associe un élément de X.*

Ou encore, le produit cartésien d'une famille d'ensembles non vide est non vide. C'est cet axiome qui permet de remplacer la suite de quantificateurs $\forall x \in X \exists y \in Y(P(x,y))$ par la suite $\exists f : X \to Y \forall x \in X(P(x, f(x)))$.

Pour le moment, on travaille sans l'axiome du choix et sans l'axiome de régularité.

Bien que les seuls objets permis en théorie des ensembles soient les ensembles, on se permettra de parler de classes d'ensembles : par une classe d'ensembles, on entend une expression formelle $\{x : \varphi(x, a_1, \ldots, a_n)\}$, où $\varphi(x, x_1, \ldots, x_n)$ est une certaine formule du premier ordre, et $a_1, \ldots, a_n$ des ensembles. (On dit alors que la classe en question est définissable à partir de $a_1, \ldots, a_n$.) Par exemple, si a est un ensemble, alors en prenant $x \in x_1$ pour $\varphi(x, x_1)$ et a pour a_1, on voit que $\{x : x \in a\}$ est une classe. Grâce à l'axiome d'extension, on peut l'identifier à a. Comme autre classe importante, on a la classe universelle $V = \{x : x = x\}$. La théorie des ensembles ZF ne permet pas vraiment de manipuler les classes. Une formule portant sur une classe précise ne pose pas de problème, mais la quantification est plus ennuyeuse. En fait, la seule chose qu'on puisse plus ou moins faire, c'est quantifier universellement sur les classes, ce qui se traduit par une infinité de formules ; par exemple, l'axiome de séparation traduit précisément le fait que l'intersection d'une classe et d'un ensemble est un ensemble, et il y a bien quantification universelle sur les classes, au prix d'une infinité de formules. La quantification existentielle sur les classes n'est pas possible, cependant : une affirmation telle que "Il existe un bon ordre sur la classe V." n'est pas formulable dans ZF (car un bon ordre sur V serait une classe), même si pour une classe $<$ donnée l'affirmation "La classe $<$ est un bon ordre sur la classe V." a un sens dans ZF. (Il s'agit là de versions fortes de l'axiome du choix.)

3.2 Ensembles bien ordonnés et ordinaux

On dit qu'un ensemble ordonné W est bien ordonné ssi toute partie non vide de W a un plus petit élément. Clairement, W est alors totalement ordonné. Une formulation plus constructive de la définition est le principe suivant :

Théorème 3.10 (principe d'induction transfinie). *Si W_0 est un ensemble, et que pour tout $z \in W$ tel que $\{t \in W : t < z\} \subseteq W_0$ on a $z \in W_0$, alors $W_0 \supseteq W$.*

Un ensemble totalement ordonné vérifie le principe d'induction transfini si et seulement si il est bien ordonné. (Si W est bien ordonné, considérer le plus petit élément de W qui n'est pas dans W_0. Si W vérifie le principe d'induction transfinie et $P \subseteq W$ n'a pas de plus petit élément, considérer pour W_0 le complémentaire de P.)

On appelle segment initial d'un ensemble bien ordonné W une partie W_0 de W telle que si $x \in W_0$ et $y < x$ alors $y \in W_0$. On montre facilement que tout segment initial de W est soit W soit est le segment initial strict d'un $x \in W$, c'est-à-dire $s(x) = \{y \in W : y < x\}$.

Les ensembles totalement ordonnés forment une catégorie rigide : si W est bien ordonné, le seul automorphisme de W (bijection croissante de W sur lui-même) est l'identité. Par conséquent, un isomorphisme entre deux ensembles bien ordonnés est unique s'il existe. Plus généralement, si W est bien ordonné et $f\colon W \to W$ est injective croissante alors $f(x) \geq x$ pour tout $x \in W$. Si W_1 et W_2 sont deux ensembles bien ordonnés, alors soit W_1 et W_2 sont isomorphes (et cet isomorphisme est unique), soit W_1 est isomorphe (par un isomorphisme unique) au segment initial $s(b) = \{y \in W_2 : y < b\}$ d'un $b \in W_2$, lequel est alors uniquement déterminé, soit W_2 est isomorphe (par un isomorphisme unique) au segment initial $s(a) = \{x \in W_1 : x < a\}$ d'un $a \in W_1$, lequel est alors uniquement déterminé, et de plus ces trois possibilités sont mutuellement exclusives. (L'idée clef dans la démonstration est d'introduire l'ensemble I des $a \in W_1$ tels que $s(a)$ soit isomorphe à un certain $s(b)$ avec $b \in W_2$ (lequel est alors clairement uniquement déterminé), et l'ensemble J des $b \in W_2$ tels que $s(b)$ soit isomorphe à un certain $s(a)$ avec $a \in W_1$ uniquement déterminé. Il est clair que I et J sont des segments initiaux. Le seul cas à exclure est celui où tous deux sont stricts, disons $I = s(u)$ et $J = s(v)$. Mais en associant à $a \in I$ l'unique $b \in J$ tel que $s(a)$ et $s(b)$ soient isomorphes, on définit un isomorphisme entre $I = s(u)$ et $J = s(v)$ donc $u \in I$ et $v \in J$, contradiction.)

Un ensemble X est dit <u>transitif</u> ssi $X \subseteq \mathscr{P}(X)$, autrement dit ssi $\forall z \in X (z \subseteq X)$, soit $\forall z \in X \forall t \in z (t \in X)$, soit encore $\bigcup X \subseteq X$. Grâce à l'axiome de l'infini, si X est un ensemble quelconque, on peut considérer la suite définie par récurrence $X_0 = X$ et $X_{n+1} = \bigcup X_n$. L'union de tous les X_i est le plus petit ensemble transitif contenant X et on l'appelle la clôture transitive de X.

Si X est transitif, on peut considérer la relation $\in$ entre les éléments de X. Si cette relation est une relation de bon ordre strict, on dit que X est un <u>ordinal</u>. En fait, si on suppose vérifié l'axiome de régularité, il suffit de supposer que la relation est transitive, autrement dit que tout élément de X est lui-même transitif.

Théorème 3.11 (du comptage (Cantor)). *Tout ensemble bien ordonné est isomorphe à un ordinal unique (et l'isomorphisme est lui-même unique).*

Si α et β sont deux ordinaux, il revient au même de dire que $\alpha \in \beta$, que $\alpha \subseteq \beta$ avec $\alpha \neq \beta$, ou encore que α est isomorphe (en tant qu'ensemble bien ordonné) à un segment initial strict de β. Dans ces conditions, on dit que $\alpha < \beta$. Avec cette relation d'ordre, la classe On des ordinaux est elle-même une classe transitive bien ordonnée.

Il existe des ordinaux : $\varnothing$ en est un, noté 0. Si α est un ordinal, alors $\alpha \cup \{\alpha\}$ est un ordinal, noté $\alpha+1$ et appelé successeur de α, qui est le plus petit ordinal strictement supérieur à α. Tout ensemble d'ordinaux est bien ordonné, et de plus l'union de ses éléments en est la borne supérieure dans la classe des ordinaux ; l'union des successeurs de ses éléments est le plus petit ordinal strictement supérieur à tous les éléments de l'ensemble (la borne supérieure stricte de l'ensemble).

Tout ordinal n'est pas le successeur d'un autre ordinal : la preuve, 0 n'en est pas un. Un ordinal autre que 0 qui n'est pas un successeur est dit <u>limite</u>. La plus petite classe contenant 0 et clos par l'opération successeur est notée ω. C'est soit un ordinal soit la classe de tous les ordinaux : l'axiome de l'infini assure que $\omega < On$, autrement dit que ω est un ordinal (est un ensemble).

Si α et β sont deux ordinaux, on peut définir la somme de α et β de la façon suivante : on ordonne l'union disjointe de α et de β en mettant les éléments de α avant ceux de β, ce qui définit un bon ordre, et on appelle $\alpha + \beta$ l'unique ordinal qui est isomorphe à ce bon ordre (en vertu du théorème du comptage de Cantor). De même, en mettant l'ordre lexicographique sur le produit cartésien qui donne le poids le plus fort aux éléments de β, on définit $\alpha\beta$. Enfin, en considérant les fonctions presque nulles de β dans α, et en les munissant de l'ordre lexicographique qui donne le poids le plus fort aux valeurs lointaines (sur les grands éléments de β), on définit α^β (exponentiation ordinale, à ne pas confondre avec l'exponentiation cardinale qui sera définie ultérieurement). L'addition et la multiplication sont associatives, et 0 et 1 y sont respectivement élément neutre bilatère. On a de plus $\alpha(\beta+\beta')=\alpha\beta+\alpha\beta'$ et $\alpha^{\beta+\beta'}=\alpha^{\beta}\alpha^{\beta'}$. Pour tout ordinal α, l'ordinal $\alpha+\omega$ est le plus petit ordinal limite strictement supérieur à α.

Signalons au passage que le théorème du comptage utilise l'axiome de remplacement de manière essentielle : sans lui (i.e. avec seulement l'axiome de séparation), on ne peut pas montrer que $\omega 2$ est un ensemble — et de fait il est consistant que $\omega 2 = On$.

Le principe d'induction transfinie sert non seulement à démontrer des résultats sur les ensembles bien ordonnés, mais aussi à y définir des fonctions : de même que si dans un ensemble S on se donne $a\in S$ et $f\colon S\to S$ on peut définir une fonction $u\colon \mathbb{N}\to S$ telle que $u(0)=a$ et $u(n+1)=f(u(n))$, et que cette fonction est unique — et d'ailleurs cette propriété caractérise $\mathbb{N}=\omega$ —, de même si W est bien ordonné, on peut définir une fonction $W\to S$ en supposant à chaque étape la fonction déjà définie à toutes les étapes précédentes. Pour être précis, si W est un ensemble bien ordonné et S un ensemble, et que pour chaque $a\in W$ on a une fonction $f_a : S^{s(a)}\to S$, où $s(a)=\{x\in W : x<a\}$, alors il existe une unique fonction $u\colon W\to S$ telle que pour chaque $a\in W$ on ait $u(a)=f_a(u|_{s(a)})$. Ce principe de définition par induction transfinie peut même servir sur des classes : si $\mathbf{W}$ est une classe bien ordonnée dont tous les segments initiaux sont des ensembles, $\mathbf{S}$ une classe, et $\mathbf{f}$ une fonction classe définie sur toutes les fonctions de $s(a)$ vers $\mathbf{S}$ (pour un certain $a\in\mathbf{W}$), alors il existe une unique fonction classe $\mathbf{u}\colon \mathbf{W}\to\mathbf{S}$ telle que $\mathbf{u}(a)=\mathbf{f}(\mathbf{u}|_{s(a)})$ pour chaque a. Le principal problème dans cette assertion est la quantification existentielle sur une classe ($\mathbf{u}$), mais en fait l'affirmation est quand même formulable dans ZF, puisqu'on peut expliciter $\mathbf{u}$, et c'est la classe des couples (a,u_a) tels qu'il existe une fonction $u\colon s(a)\to\mathbf{S}$ vérifiant $u(x)=\mathbf{f}(u|_{s(x)})$ pour tout $x\in s(a)$, et que de plus $u_a=\mathbf{f}(u)$.

Cette définition par induction transfinie est surtout utilisée pour définir des fonctions sur la classe des ordinaux. Par exemple, $\alpha+\beta$ peut se définir par induction transfinie de la façon suivante :

$$\left\{\begin{array}{rcl} \alpha+0 &=& \alpha\\ \alpha+(\beta+1) &=& (\alpha+\beta)+1\\ \alpha+\beta &=& \sup_{\gamma<\beta}\alpha+\gamma \text{ si } \beta \text{ est limite}\end{array}\right.$$

(ou, plus succintement, $\alpha+\beta$ est le plus petit ordinal au moins égal à α et strictement supérieur à $\alpha+\gamma$ pour tout $\gamma<\beta$: ceci capture les trois cas). De même, pour la multiplication et l'exponentiation,

$$\left\{\begin{array}{rcl} \alpha 0 &=& 0\\ \alpha(\beta+1) &=& \alpha\beta+\alpha\\ \alpha\beta &=& \sup_{\gamma<\beta}\alpha\gamma \text{ si } \beta \text{ est limite}\end{array}\right. \qquad \left\{\begin{array}{rcl} \alpha^0 &=& 1\\ \alpha^{\beta+1} &=& \alpha^\beta\alpha\\ \alpha^\beta &=& \sup_{\gamma<\beta}\alpha^\gamma \text{ si } \beta \text{ est limite}\end{array}\right.$$

Tout ordinal α peut être muni d'une topologie définie par son ordre naturel. Les points d'accumulation de cet espace topologique sont les ordinaux limites inférieurs à α. L'espace est compact si et seulement si α est un ordinal successeur. Si α est limite et $(\gamma_\iota)_{\iota<\alpha}$ est une famille d'ordinaux indicée par les éléments de α (appelée une α-séquence d'ordinaux), on dit comme d'habitude que γ_ι tend vers β lorsque ι tend vers α si et seulement si pour tous $\beta_1<\beta<\beta_2$ il existe $\alpha_0<\alpha$ tel que $\beta_1<\gamma_\iota<\beta_2$ dès que $\iota\geq\alpha_0$. Lorsque β est successeur, cela signifie que la séquence est stationnaire de valeur β. Si la séquence est croissante, elle a toujours une limite (la borne supérieure de tous ses termes).

On dit qu'une α-séquence d'ordinaux (pour α limite) est continue ssi sa limite en tout ordinal limite $\beta<\alpha$ est précisément sa valeur en ce point. Une α-séquence à la fois continue et strictement croissante est appelée normale. Ces notions sont encore valables pour une séquence indicée par la classe On tout entière. Si $(\gamma_\iota)_{\iota\in On}$ est une On-séquence normale d'ordinaux, alors il existe des ordinaux ι arbitrairement grands tels que $\gamma_\iota=\iota$ (en effet, si α est un ordinal, on considère $\alpha_1=\gamma_\alpha$, $\alpha_2=\gamma_{\alpha_1}$ et ainsi de suite, et on appelle ι la limite de la séquence α_n). On en déduit que pour tout α il existe des β arbitrairement grands tels que $\alpha+\beta=\beta$, que si $\alpha>0$ il existe des β arbitrairement grands tels que $\alpha\beta=\beta$, et que si $\alpha>1$ il existe des β arbitrairement grands tels que $\alpha^\beta=\beta$. Le ι-ième point fixe de la fonction $\beta\mapsto\omega^\beta$ est appelé ε_ι.

On définit un ordre sur On^2 (la classe des couples d'ordinaux) de la façon suivante :

$$\begin{array}{rcl}(\alpha_1, \alpha_2) < (\beta_1, \beta_2) & \Leftrightarrow & \max(\alpha_1, \alpha_2) < \max(\beta_1, \beta_2) \\ & \text{ou} & \max(\alpha_1, \alpha_2) < \max(\beta_1, \beta_2) \text{ et } \alpha_1 < \beta_1 \\ & \text{ou} & \max(\alpha_1, \alpha_2) < \max(\beta_1, \beta_2) \text{ et } \alpha_1 = \beta_1 \text{ et } \alpha_2 < \beta_2\end{array}$$

Cet ordre est un bon ordre sur On^2, et tous les α^2 pour $\alpha \in On$ en sont des segments initiaux. En notant $\Gamma(\alpha, \beta)$ l'ordinal du segment initial strict de (α, β) pour l'ordre en question, on en déduit une bijection Γ entre On^2 et On, qui par ailleurs se restreint à une bijection entre α^2 et $\gamma(\alpha) = \Gamma(\alpha, \alpha)$ pour tout ordinal α. La fonction $\alpha \mapsto \gamma(\alpha)$ est normale donc il existe des α arbitrairement grands tels que $\gamma(\alpha) = \alpha$, i.e. tels que Γ induise une bijection entre α^2 et α. Le plus petit tel α est ω.

Si α est un ordinal limite, une partie $S \subseteq \alpha$ est dite cofinale ssi pour tout $\gamma < \alpha$ il existe $\gamma' \in S$ tel que $\gamma < \gamma'$. Une β-séquence $(\gamma_\iota)_{\iota<\beta}$ strictement croissante d'éléments de α est dite cofinale ssi elle tend vers α : il y a évidemment correspondance bijective entre une partie cofinale S de α et une β-séquence cofinale, où β est l'ordinal de S comme ensemble bien ordonné. Le plus petit tel β s'appelle la cofinalité de α et se note $\operatorname{cf} \alpha$. On a $\operatorname{cf} \alpha \leq \alpha$ et $\operatorname{cf} \operatorname{cf} \alpha = \operatorname{cf} \alpha$ puisque si $\beta = \operatorname{cf} \alpha$, disons $(\gamma_\iota)_{\iota<\beta}$ tend vers α, et que $(\eta_\iota)_{\iota<\delta}$ tend vers β, alors $(\gamma_{\eta_\iota})_{\iota<\delta}$ tend vers α. Il est donc raisonable d'appeler régulier un ordinal limite α tel que $\operatorname{cf} \alpha = \alpha$.

3.3 Cardinalités

Si X et Y sont deux ensembles, et qu'il existe une bijection entre X et Y, on dit que X et Y ont même cardinalité, ou qu'ils sont équipotents, et on note $|X| = |Y|$. Il s'agit là clairement d'une relation d'équivalence sur l'univers.

On pourrait utiliser l'axiome de régularité pour donner un sens à $|X|$ (comme l'ensemble des ensembles équipotents à X et de rang minimal) ou encore l'axiome du choix (comme on le verra plus loin). Ce n'est pas absolument nécessaire : on peut se contenter de dire que les cardinalités ne sont qu'une commodité de langage, et traduire les assertions qui en parlent en termes d'ensembles.

Si X est équipotent à une partie de Y, on dit que la cardinalité de X est inférieure ou égale à celle de Y, et on note $|X| \leq |Y|$. L'intérêt de cette définition est justifiée par le théorème bien connu suivant :

Théorème 3.12 (de Cantor-Bernstein). *Si $|X| \leq |Y|$ et $|Y| \leq |X|$ alors $|X| = |Y|$.*

Autrement dit, les cardinalités sont partiellement ordonnées. On notera bien entendu $|X| < |Y|$ lorsque $|X| \leq |Y|$ mais que $|X| \neq |Y|$ (i.e. non $|X| = |Y|$). On emploiera sans plus d'explications des signes tels que $>$, $\geq$, $\not\geq$ entre cardinalités.

Si un ensemble X se surjecte sur un ensemble Y, on dira que $|Y|$ est une projection de $|X|$ et on notera $|Y| \trianglelefteq |X|$. Il va de soi que si $|Y| \leq |X|$ et que $|Y| > 0$ (i.e. $Y \neq \varnothing$) alors $|Y| \trianglelefteq |X|$.

La cardinalité de l'union de X et Y, dans la mesure où ceux-ci sont disjoints, ne dépend que des cardinalités de X et Y : cette cardinalité est notée $|X| + |Y|$. La cardinalité du produit cartésien de X et Y définit encore une opération sur les cardinalités, notée $|X|\,|Y|$. Enfin, in en va de même de la cardinalité de l'ensemble des fonctions de Y dans X, notée $|X|^{|Y|}$. On identifie un entier naturel (un élément de ω) à sa cardinalité : ainsi, $|\mathscr{P}(Y)| = 2^{|Y|}$.

On note de plus $\aleph_0$ la cardinalité de ω. On dira qu'un ensemble X est fini ssi $|X| < \aleph_0$, ou, ce qui revient au même, ssi $|X| = n$ pour un certain $n \in \omega$. On dira qu'un ensemble X est D-fini (pour "Dedekind-fini") ssi $|X| \not\geq \aleph_0$, autrement dit ss'il n'existe pas de suite d'éléments deux à deux distincts de X ; cela revient à exiger que $|X| > |Y|$ pour toute partie $Y \subseteq X$ autre que X.

Le résultat suivant est dû à Cantor :

Théorème 3.13 (Cantor). *Pour tout ensemble X, on a $2^{|X|} \ntrianglelefteq |X|$ (et en particulier, $2^{|X|} \not\leq |X|$) ; comme de plus $|X| \leq 2^{|X|}$, on en déduit $|X| < 2^{|X|}$.*

Si α et β sont des ordinaux, il est évident que $\alpha \leq \beta$ implique $|\alpha| \leq |\beta|$. Si κ est un ordinal tel que $|\alpha| < |\kappa|$ dès que $\alpha < \kappa$, autrement dit que κ est le plus petit ordinal ayant sa cardinalité, alors on dit que κ est un cardinal. Il est clair que tout ordinal, donc tout ensemble bien ordonné, est équipotent à un unique cardinal (et, réciproquement, tout ensemble équipotent à un cardinal peut être bien ordonné) : les cardinaux sont les cardinalités des ensembles bien ordonnés. Cela étant, pour κ un cardinal, on identifiera κ à $|\kappa|$. On vérifie facilement (whatever it's worth for) que si $|X| \trianglelefteq \kappa$, où κ est un cardinal, alors $|X| \leq \kappa$.

Si X est un ensemble quelconque, on appellera nombre de Hartogs de X le plus petit ordinal α tel qu'on n'ait pas $|\alpha| \leq |X|$, autrement dit le plus petit ordinal qui n'est pas l'ordinal d'un bon ordre sur une partie de X (un tel ordinal existe car les bons ordres sur des parties de X forment clairement un ensemble). Le nombre de Hartogs $h(X)$ de X est clairement un cardinal. Si $|X|$ était déjà un cardinal (autrement dit si X peut être bien ordonné), alors $h(X) > |X|$. Ceci montre que tout les cardinaux sont cofinaux dans les ordinaux (tout ordinal est majoré par un cardinal) : en particulier, pour tout cardinal κ il existe un plus petit cardinal supérieur à κ noté κ^+. Par ailleurs, il est facile de vérifier que le sup (en tant qu'ordinal) d'une famille de cardinaux est un cardinal.

Cela étant, on peut légitimement définir $\aleph_0 = \omega$, pour tout ordinal successeur $\alpha + 1$, poser $\aleph_{\alpha+1} = \aleph_\alpha^+$, et pour tout ordinal limite α, poser $\aleph_\alpha = \sup_{\beta<\alpha} \aleph_\beta$. Ainsi, les $\aleph_\alpha$ sont des cardinaux, et décrivent même tous les cardinaux infinis. Pour cette raison, lorsque la distinction entre cardinalités et cardinaux est cruciale, on préférera dire "un aleph" plutôt que "un cardinal" (infini), ce qui revient au même. En tant que séquence ordinale, la fonction $\alpha \mapsto \aleph_\alpha$ est normale (i.e. strictement croissante et continue), donc elle possède des points fixes arbitrairement grands. Un aleph $\aleph_\alpha$ dont l'indice α est un ordinal successeur (resp. limite) est appelé un cardinal successeur (resp. limite). Lorsqu'on veut parler de $\aleph_\alpha$ en tant qu'ordinal, on note plutôt ω_α ; ainsi, $\omega_0 = \omega$, et ω_1 est le plus petit ordinal indénombrable, qui est aussi l'ordinal de l'ensemble des façons de bien ordonner ω.

Revenant à la bijection canonique Γ établie entre On^2 et On, on a vu qu'il existait des ordinaux α arbitrairement grands tels que $\gamma(\alpha) = \alpha$, c'est-à-dire que Γ induit une bijection entre α^2 et α. On montre assez facilement par induction transfinie sur γ que c'est le cas de $\aleph_\gamma$ pour tout γ, i.e. que tout cardinal $\kappa = \aleph_\gamma$ vérifie $\kappa^2 = \kappa$ (il s'agit là de l'exponentiation cardinale et non de l'exponentiation ordinale pour laquelle on a $\alpha^2 > \alpha$ pour tout α). Par conséquent, pour tout cardinaux κ et λ dont au moins un est infini, on a

$$\kappa\lambda = \kappa + \lambda = \max(\kappa, \lambda)$$

Tout cardinal infini est un ordinal limite. On a défini précédemment, pour tout ordinal limite α, la cofinalité $\operatorname{cf}\alpha$ de α comme le plus petit ordinal limite β pour lequel il existe une β-séquence d'éléments de α (qu'on peut toujours supposer strictement croissante) qui tend vers α. Cette cofinalité est nécessairement un cardinal, car si $\operatorname{cf}\alpha$ est en bijection avec un ordinal β, alors en composant une $\operatorname{cf}\alpha$-séquence de limite α avec la bijection sur β, on voit qu'il existe une β-séquence de limite α, donc que $\beta \geq \operatorname{cf}\alpha$. Ainsi, tout ordinal régulier est un cardinal, et on parle donc plutôt de cardinaux réguliers. La réciproque n'est pas vraie : $\aleph_\omega$, étant la limite de la ω-séquence $(\aleph_n)_{n<\omega}$, est un cardinal mais n'est pas régulier, et on dit alors que c'est un cardinal singulier. En fait, en l'absence de l'axiome du choix, on ne sait pas si on peut montrer la régularité d'un cardinal indénombrable quelconque (et on pense qu'on ne peut pas).

On a vu (théorème de Cantor) que $2^\kappa > \kappa$ pour toute cardinalité κ, et en particulier pour tout cardinal κ. Une autre inégalité importante dans ce genre est la suivante :

Théorème 3.14. *Si κ est un cardinal infini, alors $\kappa^{\operatorname{cf}\kappa} > \kappa$.*

Démonstration. On a clairement $\kappa^{\operatorname{cf}\kappa} \geq \kappa$. Il s'agit de montrer que cette inégalité est stricte. Supposons donc par l'absurde qu'on ait une énumération $(f_\alpha)_{\alpha<\kappa}$ des fonctions de $\operatorname{cf}\kappa$ dans κ. Soit $(\gamma_\iota)_{\iota<\operatorname{cf}\kappa}$ une suite strictement croissante d'éléments de κ qui tend vers κ. On définit $g\colon \operatorname{cf}\kappa \to \kappa$ comme suit. Si $\iota < \operatorname{cf}\kappa$ alors les $f_\alpha(\iota)$, pour $\alpha \leq \iota$ ne remplissent pas κ, et on appelle $g(\iota)$ le plus petit élément de κ qui diffère de tous ces éléments. La fonction $g\colon \operatorname{cf}\kappa \to \kappa$ est alors différente de tous les f_α. □

Cependant, en l'absence de l'axiome du choix, il y a peu d'autres choses qu'on puisse dire sur les κ^λ. Entre autres, on ne peut pas montrer que κ^λ soit un aleph.

Un cardinal est dit faiblement inaccessible ss'il est limite et régulier et indénombrable. Entre autres, puisque $\operatorname{cf}\aleph_\alpha = \operatorname{cf}\alpha$ lorsque α est limite, un cardinal faiblement inaccessible λ doit vérifier

$\aleph_\lambda = \lambda$. Le plus petit point fixe de la fonction $\alpha \mapsto \aleph_\alpha$ est la limite de la suite $\aleph_0, \aleph_\omega, \aleph_{\omega_\omega}, \aleph_{\omega_{\omega_\omega}}, \cdots$, et il est donc de cofinalité ω (donc il est "très loin" d'être inaccessible).

3.4 L'axiome du choix

L'axiome du choix à la conséquence très importante suivante :

Théorème 3.15 (du bon ordre). *Tout ensemble peut être bien ordonné.*

Démonstration. Soit X un ensemble. D'après l'axiome du choix, on peut trouver une fonction f qui à toute partie non vide de X associe un élément de cette partie. On définit alors par induction sur l'ordinal ι la fonction $x_\iota = f(X \setminus \{x_\xi : \xi < \iota\})$ tant que $\{x_\xi : \xi < \iota\}$ n'est pas tout X. Cette fonction est injective donc par l'axiome de remplacement elle ne peut pas s'étendre à tous les ordinaux, ce qui montre qu'il existe un ordinal α tel que $\{x_\iota : \iota < \alpha\}$ soit exactement X, et alors on a une bijection entre X et α, donc X est bien ordonnable. $\square$

La démonstration a été donnée pour montrer à quel point les ordinaux sont un outil puissant : on évite ici le recours au lemme de Zorn. D'ailleurs, grâce aux ordinaux, le lemme de Zorn se montre aussi très simplement :

Lemme 3.16 (Kuratowski-Zorn). *Soit P un ensemble partiellement ordonné dont toute chaîne (i.e. sous-ensemble bien ordonné) a un majorant. Alors P a un élément maximal.*

Démonstration. Supposons par l'absurde que P n'ait pas d'élément maximal. Soit x_0 un élément de P. On définit une séquence x_ι d'éléments de P comme suit : $x_{\iota+1}$ est un élément de P strictement supérieur à x_ι (qui existe car x_ι n'est pas maximal), et si ι est limite alors x_ι est un majorant strict de la chaîne $\{x_\xi : \xi < \iota\}$. L'axiome du choix permet de faire ces choix. On a alors une fonction injective des ordinaux dans P, ce qui est une contradiction. $\square$

Grâce à l'axiome du choix, il est possible d'identifier cardinalités et cardinaux : toute cardinalité infinie est un aleph (c'est précisément le contenu du théorème du bon ordre). Ceci a plusieurs conséquences agréables. Premièrement, deux cardinalités sont toujours comparables : autrement dit, si X et Y sont deux ensembles, il existe toujours soit une injection de X dans Y soit une injection de Y dans X. Le fait d'être fini ou D-fini coïncide. Ensuite, grâce à ce qu'on a déjà montré sur les cardinaux, on voit que la somme ou le produit de deux cardinalités dont une au moins est infinie est le maximum de ces deux cardinalités. Entre autres, X^2 est équipotent à X lorsque X est infini. Aussi, dès que $X \neq \varnothing$, dire que $|X| \trianglelefteq |Y|$ ou $|X| \leq |Y|$ revient au même : toute surjection admet une section. Enfin, l'axiome du choix permet de montrer le résultat suivant :

Proposition 3.17. *Le cardinal de $\bigcup_{i \in I} X_i$ est majoré par $|I| \sup_{i \in I} |X_i|$.*

Corollaire 3.18. *Tout cardinal successeur, i.e. tout $\aleph_{\alpha+1}$, est régulier.*

(Sans quoi $\omega_{\alpha+1}$ serait la réunion de au plus $\aleph_\alpha$ ensembles de cardinal au plus $\aleph_\alpha$, ce qui contredit la proposition vu que $\aleph_\alpha^2 = \aleph_\alpha$.)

En particulier, le plus petit cardinal singulier est $\aleph_\omega$, qui est de cofinalité ω.

Par ailleurs, l'axiome du choix permet de définir un sens à la somme ou au produit d'une famille quelconque de cardinaux, comme suit. Si κ_i sont des cardinaux, X_i des ensembles deux à deux disjoints de cardinal κ_i, alors le cardinal de $\bigcup_i X_i$ ne dépend pas du choix des X_i et on le note $\sum_i \kappa_i$. De même, le produit cartésien permet de définir le produit d'une famille de cardinaux. Pour les calculer, on dispose du lemme suivant :

Lemme 3.19. *Soit $(\kappa_\iota)_{\iota<\lambda}$ une λ-séquence de cardinaux tous non nuls, où λ est un cardinal infini. Alors $\sum_{\iota<\lambda} \kappa_\iota = \lambda \sup_{\iota<\lambda} \kappa_\iota$. Si de plus la séquence (κ_ι) est croissante alors $\prod_{\iota<\lambda} \kappa_\iota = (\sup_{\iota<\lambda} \kappa_\iota)^\lambda$.*

On a par ailleurs le résultat suivant :

Théorème 3.20 (König). *Si $\kappa_i < \lambda_i$ pour tout $i \in I$ alors $\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i$.*

En particulier, comme $1 < 2$, on retrouve $\kappa < 2^\kappa$, et en écrivant κ comme somme de $\operatorname{cf} \kappa$ cardinaux $\kappa_i < \kappa$, on retrouve $\kappa < \kappa^{\operatorname{cf} \kappa}$.

3.5 Exponentiation cardinale

On suppose toujours vérifié l'axiome du choix.

L'axiome du choix décrète péremptoirement que κ^λ est un aleph (i.e. peut être bien ordonné). La question naturelle est alors de se demander quelle est sa valeur précise. On ne peut pas répondre à cette question : en fait, on ne peut même pas dire ce que vaut $2^{\aleph_0}$. On peut cependant trouver certains résultats positifs. On sait déjà que $2^{\aleph_0} > \aleph_0$, et donc que $2^{\aleph_0} \geq \aleph_1$ (cette dernière inégalité fait appel à l'axiome du choix). D'autre part, par exemple, $2^{\aleph_0}$ ne peut pas être $\aleph_\omega$ puisque $(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0^2} = 2^{\aleph_0}$ tandis que $(\aleph_\omega)^{\aleph_0} > \aleph_\omega$ car $\operatorname{cf} \aleph_\omega = \omega$.

Si on connait les fonctions $\kappa \mapsto 2^\kappa$ (la fonction du continu) et $\kappa \mapsto \gimel(\kappa) = \kappa^{\operatorname{cf} \kappa}$ (la fonction gimel), alors on peut calculer κ^λ pour tous κ et λ d'après le théorème suivant :

Théorème 3.21 (calcul inductif de $\aleph_\alpha^{\aleph_\beta}$). *Soit $\aleph_\beta$ un cardinal infini fixé. Alors $\aleph_\alpha^{\aleph_\beta}$ peut être calculé inductivement comme suit :*

1. *Si $\alpha \leq \beta$ alors $\aleph_\alpha^{\aleph_\beta} = 2^{\aleph_\beta}$.*
2. *Si pour un certain $\gamma < \alpha$ on a $\aleph_\gamma^{\aleph_\beta} \geq \aleph_\alpha$, alors $\aleph_\alpha^{\aleph_\beta} = \aleph_\gamma^{\aleph_\beta}$.*
3. *Si $\alpha > \beta$ et $\aleph_\gamma^{\aleph_\beta} < \aleph_\alpha$ pour tous $\gamma < \alpha$, alors*
   (a) *Si $\operatorname{cf} \aleph_\alpha > \aleph_\beta$ (entre autres si $\aleph_\alpha$ est régulier), alors $\aleph_\alpha^{\aleph_\beta} = \aleph_\alpha$.*
   (b) *Si $\operatorname{cf} \aleph_\alpha \leq \aleph_\beta$ alors $\aleph_\alpha^{\aleph_\beta} = \aleph_\alpha^{\operatorname{cf} \aleph_\alpha}$.*

En fait, on montre même que la fonction du continu est elle-même déterminée par la fonction gimel.

La question est alors de savoir quels comportements sont possibles pour la fonction du continu et la fonction gimel, moyennant les axiomes de la théorie des ensembles. L'hypothèse la plus simple, et aussi la plus célèbre, et la suivante :

Hypothèse généralisée du continu 3.22. *Pour tout cardinal κ, on a $2^\kappa = \kappa^+$.*

Cette hypothèse détermine alors complètement la fonction gimel et du même coup toute l'exponentiation des cardinaux.

Théorème 3.23. *Sous l'hypothèse généralisée du continu, $\aleph_\alpha^{\aleph_\beta}$ vaut*
- $\aleph_\alpha$ *si* $\aleph_\beta < \operatorname{cf} \aleph_\alpha$.
- $\aleph_{\alpha+1}$ *si* $\operatorname{cf} \aleph_\alpha \leq \aleph_\beta \leq \aleph_\alpha$.
- $\aleph_{\beta+1}$ *si* $\aleph_\alpha \leq \aleph_\beta$.

En fait, l'hypothèse du continu est indépendante des axiomes de la théorie des ensembles. Mais on a beaucoup plus : il est consistant, par exemple, que $2^{\aleph_0} = \aleph_{42}$, $2^{\aleph_1} = \aleph_{42}$, $2^{\aleph_2} = \aleph_{\omega_{1729}}$ et $2^{\aleph_3} = \aleph_{\omega_{1729}+\varepsilon_0^2+666}$. Le théorème précis est le théorème d'Easton, qui affirme que toute fonction croissante $\mathbf{F}$ des cardinaux réguliers vers les cardinaux, qui vérifie $\operatorname{cf} \mathbf{F}(\kappa) > \kappa$ pour tout κ, est susceptible d'"être" la fonction du continu. (Cet énoncé est un peu vague : on renvoie au Jech pour une formulation précise.) La fonction du continu (ou la fonction gimel) sont beaucoup plus difficiles à contrôler sur les cardinaux singuliers. Par exemple, si l'hypothèse du continu généralisée est vraie sur tous les cardinaux inférieurs à $\aleph_{\omega_1}$, alors elle est vraie en $\aleph_{\omega_1}$ (un théorème de Silver, théorème 23 dans le Jech). Ceci est encore vrai pour $\aleph_\omega$ en l'absence d'une hypothèse très forte, l'existence de $0^\sharp$ (qui est par exemple impliquée par l'existence d'un cardinal mesurable). D'autre part, si la fonction du continu est stationnaire en-dessous d'un cardinal singulier, alors sa valeur sur ce cardinal est la valeur à laquelle elle stationne (théorème de Bukovský-Hechler).

3.6 L'axiome de régularité

L'axiome de régularité permet une forme d'induction : si $\mathbf{C}$ est une classe telle que $x \subseteq \mathbf{C} \Rightarrow x \in \mathbf{C}$ (autrement dit, $\mathbf{C}$ contient un ensemble x dès qu'elle contient tous les éléments de x), alors $\mathbf{C} = V$ (induction bien-fondée).

Comme tout principe d'induction, l'induction bien-fondée donne aussi un principe de définition par induction bien-fondée. Précisément : si $\mathbf{S}$ est une classe, et $\mathbf{f}$ une fonction classe de $V \times V$

vers $\mathbf{S}$, alors il existe une unique fonction classe $\mathbf{u}\colon V \to \mathbf{S}$ telle que $\mathbf{u}(x) = \mathbf{f}(x, \mathbf{u}|_x)$ pour chaque x. Autrement dit, on peut, quand on définit une fonction sur l'univers (ou sur une classe ou un ensemble transitif en général) supposer qu'elle est déjà définie sur tous les éléments de x avant de la définir sur x.

Comme application, on définit une application "rang" de V vers les ordinaux par

$$\mathrm{rg}(x) = \sup\{\mathrm{rg}\, y + 1 : y \in x\}$$

Autrement dit, le rang de x est le plus petit ordinal strictement supérieur au rang de tous les éléments de x. (On trouve parfois la définition alternative $\mathrm{rg}(x) = \sup\{\mathrm{rg}\, y : y \in x\} + 1$, qui est toujours de 1 supérieure à la définition que nous donnons.) On a ainsi $\mathrm{rg}(\varnothing) = 0$, $\mathrm{rg}(\{\varnothing\}) = 1$, et plus généralement $\mathrm{rg}(\alpha) = \alpha$ pour tout ordinal α, $\mathrm{rg}(\mathscr{P}(x)) = \mathrm{rg}(x) + 1$.

Si $\mathbf{C}$ est une classe, on peut définir une classe $\mathscr{H}\mathbf{C}$, par induction bien-fondée : on pose $\mathscr{H}\mathbf{C} = \{x \in \mathbf{C}\colon x \subseteq \mathscr{H}\mathbf{C}\}$. Il s'agit là de la plus grande classe transitive contenue dans $\mathbf{C}$: on dit que ses éléments sont les ensembles "héréditairement dans $\mathbf{C}$" (ils sont dans $\mathbf{C}$, ainsi que tous leurs éléments, les éléments de leurs éléments et ainsi de suite).

(Un autre exemple, plus amusant, de définition par induction bien-fondée est la fonction de Grundy : sa valeur sur un ensemble x est le plus petit ordinal qui n'est pas la fonction de Grundy d'un élément de x. La fonction de Grundy de x vaut alors 0 si et seulement si le jeu dans lequel chaque joueur tour à tour choisit un élément de l'ensemble proposé par le précédent, en commençant par x, est gagnant pour le deuxième joueur — i.e. est un jeu nul au sens de Conway. L'axiome de régularité assure précisément que ce jeu termine toujours.)

Pour tout ordinal α, on appelle V_α l'ensemble des ensembles dont le rang est strictement inférieur à α. Pour montrer qu'il s'agit bien d'un ensemble, il suffit de remarquer que V_α peut encore se définir par induction transfinie sur α par

$$\left\{\begin{array}{rcl} V_0 & = & \varnothing \\ V_{\alpha+1} & = & \mathscr{P}(V_\alpha) \\ V_\alpha & = & \bigcup_{\beta<\alpha} V_\beta \text{ si } \beta \text{ est limite} \end{array}\right.$$

On note $V_\infty = \bigcup_{\alpha \in On} V_\alpha$: cette classe est bien définie puisque les V_α peuvent se définir uniformément pour tout α par une formule. L'axiome de régularité montre que $V_\infty = V$. En fait, les V_α et V_∞ peuvent se définir même en l'absence de l'axiome de régularité, et l'axiome de régularité est justement équivalent à l'assertion $V_\infty = V$.

L'axiome de régularité a une vague similarité avec l'axiome du choix en ce sens : il permet pour chaque classe de choisir un ensemble d'éléments de celle-ci qui est non vide si la classe était non vide. En effet, si $\mathbf{C}$ est une classe, on peut appeler $\hat{\mathbf{C}}$ l'ensemble des éléments de $\mathbf{C}$ ayant rang minimal. En quelque sorte, on choisit donc "un petit nombre" (seulement un ensemble !) d'éléments de $\mathbf{C}$. Ceci permet par exemple de définir les cardinalités en l'absence de l'axiome du choix : on appelle cardinalité d'un ensemble A l'ensemble de tous les ensembles équipotents à A et ayant rang minimal pour cette propriété.

4 Modèles de la théorie des ensembles

Un modèle de la théorie des ensembles est bien entendu un ensemble $\mathfrak{M}$ muni d'une relation ε tel que $\mathfrak{M} \vDash \Phi$ pour tout axiome Φ de la théorie ZF (ou ZFC selon le cas). Il faut remarquer que l'assertion "$(\mathfrak{M}, \varepsilon)$ est un modèle de ZF" est traduisible dans ZF (bien qu'il y ait une infinité d'axiomes de ZF). D'après le théorème d'incomplétude de Gödel, on ne peut pas montrer dans ZF l'existence d'un couple $(\mathfrak{M}, \varepsilon)$ qui soit un modèle de ZF.

On peut cependant généraliser la définition d'un modèle aux classes : si $\mathfrak{M}$ est une classe et ε est une relation binaire dessus, on peut définir $\mathfrak{M} \vDash \Phi$ de la façon évidente, et définir ce que signifie le fait que $(\mathfrak{M}, \varepsilon)$ soit un modèle de ZF. Le cas le plus évident est celui où la classe $\mathfrak{M}$ est précisément la classe universelle V, et où ε est justement la relation $\in$ sur celle-ci. Ainsi, ZF n'est pas assez fort pour montrer qu'un ensemble est modèle de ZF, mais il peut montrer qu'une classe est modèle de ZF (et, comme on le verra, V n'est pas le seul exemple en général).

Il faut toutefois noter le fait important suivant : bien que pour une formule Φ quelconque fixée, l'assertion "$\mathfrak{M} \vDash \Phi$" soit toujours traduisible dans ZF, elle ne l'est pas forcément uniformément en Φ lorsque $\mathfrak{M}$ est une classe, c'est-à-dire que pour une numérotation de Gödel Φ_n des formules, l'assertion "$\mathfrak{M} \vDash \Phi_n$" (dans laquelle n est une variable libre) ne peut pas se traduire dans le langage de ZF.

On prêtera une attention tout particulière aux modèles transitifs de la théorie des ensembles, c'est-à-dire aux modèles $(\mathfrak{M}, \varepsilon)$ dans lesquels $\mathfrak{M}$ est un ensemble ou une classe transitif, et ε est la relation $\in$ sur cette classe. C'est le cas entre autres de V.

On dit qu'une relation ε sur une classe $\mathfrak{N}$ est bien-fondée si et seulement si

- Pour tout $x \in \mathfrak{N}$ la classe $\mathrm{ext}_\varepsilon(x)$ des $y \in \mathfrak{N}$ tels que $y \varepsilon x$ est un ensemble.
- Si $S \subseteq \mathfrak{N}$ est un ensemble quelconque, alors il existe $x \in S$ tel que $\mathrm{ext}_\varepsilon(x) \cap S = \varnothing$ (autrement dit, un $x \in S$ minimal pour la relation ε).

On dit de plus que la relation ε est extensionnelle ssi $\mathrm{ext}_\varepsilon(x) = \mathrm{ext}_\varepsilon(y)$ implique $x = y$ pour tous $x, y \in \mathfrak{N}$ (autrement dit, ssi $((N), \varepsilon)$ modélise l'axiome d'extension). Ainsi, l'axiome de régularité équivaut à l'affirmation que la classe V munie de la relation $\in$ est bien-fondée, et l'axiome d'extension à l'affirmation que cette relation est extensionnelle. Réciproquement, on a le résultat suivant :

Théorème 4.1 (d'écrasement (Mostowski)). *Soit $\mathfrak{N}$ une classe munie d'une relation binaire ε bien-fondée et extensionnelle. Alors il existe une classe transitive $\mathfrak{M}$ unique telle que les structures $(\mathfrak{N}, \varepsilon)$ et $(\mathfrak{M}, \in)$ soient isomorphes. De plus, l'isomorphisme entre celles-ci est unique.*

Si Φ est une formule, comportant éventuellement des variables libres, et $\mathfrak{M}$ une classe transitive, on notera $\Phi^{\mathfrak{M}}$ la formule obtenue en restreignant tous les quantificateurs de Φ à $\mathfrak{M}$. Par exemple, si Φ est la formule $\forall z((\forall t \in z(t \in x)) \Rightarrow z \in y)$, la formule $\Phi^{\mathfrak{M}}$ sera $\forall z \in \mathfrak{M}((\forall t \in z(t \in x)) \Rightarrow z \in y)$ (il n'est pas nécessaire de restreindre le second quantificateur puisqu'il est déjà restreint à z, qui est un élément donc une partie de $\mathfrak{M}$). On a alors $\mathfrak{M} \vDash \Phi$ si et seulement si $\Phi^{\mathfrak{M}}$. Autrement dit, il est très facile de tester si une formule est vérifiée dans une classe transitive. Ceci est d'autant plus facile lorsque la formule Φ ne comporte que des quantificateurs déjà restreints (on dit que Φ est une formule restreinte), car alors on a $\mathfrak{M} \vDash \Phi$ si et seulement si Φ est vraie (i.e. $V \vDash \Phi$). On dit que les formules restreintes sont "absolues". Par exemple, en prenant pour Φ la formule $\forall z \in x(z \in y)$, on voit que si $\mathfrak{M}$ est une classe transitive et $x, y \in \mathfrak{M}$, on a $\mathfrak{M} \vDash x \subseteq y$ si et seulement si $x \subseteq y$.

Ceci nous permet de voir qu'une classe transitive $\mathfrak{M}$ modélise l'axiome de la paire si et seulement si pour tous $x, y \in \mathfrak{M}$ on a $\{x, y\} \in \mathfrak{M}$. De même, $\mathfrak{M}$ modélise l'axiome de l'union si et seulement si pour tout $x \in \mathfrak{M}$ on a $\bigcup x \in \mathfrak{M}$. Et $\mathfrak{M}$ modélise l'axiome de l'infini si et seulement si $\omega \in \mathfrak{M}$. En ce qui concerne l'axiome des parties, il faudra faire plus attention : $\mathfrak{M}$ le modélise si et seulement si pour tout $x \in \mathfrak{M}$ on a $\mathscr{P}(x) \cap \mathfrak{M} \in \mathfrak{M}$.

On rappelle qu'on a défini précédemment la classe V_∞ des ensembles réguliers. Au sujet de cette classe, on montre, sans utiliser l'axiome de régularité

Théorème 4.2. *La classe V_∞ est un modèle de ZF (et de ZFC si l'axiome du choix tient) avec l'axiome de régularité. Entre autres, l'axiome de régularité est consistant avec les autres axiomes de la théorie des ensembles.*

Démonstration. Toute classe transitive modélise l'axiome d'extension. Si α est soit un ordinal limite soit ∞, alors V_α contient $\{x, y\}$ dès qu'il contient x et y, donc V_α modélise l'axiome de la paire. De même V_α modélise alors l'axiome de l'ensemble des parties puisque si $x \in V_\beta$, on a $\mathscr{P}(x) \cap V_\beta = \mathscr{P}(x) \in V_{\beta+1}$. Pour tout $\alpha \in On \cup \{\infty\}$, la classe V_α modélise l'axiome de l'union.

Pour l'axiome de séparation, il s'agit de montrer que si $\varphi(x, x')$ est une formule (on peut se contenter de deux variables libres) et $y \in V_\alpha$, alors pour tout $x' \in V_\alpha$ on a $\{x \in y : V_\alpha \vDash \varphi(x, x')\} \in V_\alpha$. Mais ceci est évidemment vrai (pour tout $\alpha \in On \cup \{\infty\}$) puisque l'ensemble en question est une partie de y.

L'axiome de l'infini est vérifié dès que $\alpha > \omega$. L'axiome de régularité est vérifié pour tout α. Le cas de l'axiome du choix (si nécessaire) est également facile.

Reste l'axiome de remplacement. Si $x \in V_\alpha$ et que y est un ensemble obtenu en remplaçant chaque élément de x par un certain élément de V_α, on voudrait montrer que y est dans V_α. Ceci est clairement le cas pour $\alpha = \infty$ (puisqu'un ensemble est régulier si et seulement si tous ses éléments sont réguliers). C'est encore le cas pour $\alpha = \omega$ (puisque x est alors fini et que le rang de y est donc le sup d'un ensemble fini de nombres finis). Plus généralement, si κ est un cardinal régulier et tel que $2^\lambda < \kappa$ dès que $\lambda < \kappa$, alors $\alpha = \kappa$ convient (on dit que κ est régulier et fortement limite, ou encore fortement inaccessible). □

On notera au passage que V_ω (l'ensemble des ensembles hériditairement finis) est un modèle de tous les axiomes de la théorie des ensembles sauf l'axiome de l'infini, ce qui montre que l'axiome de l'infini est indépendant des autres axiomes. Et que $V_{\omega 2}$ est un modèle de tous les axiomes sauf l'axiome de remplacement, ce qui montre que l'axiome de remplacement est indépendant des autres axiomes. La classe V_∞ est un modèle de ZF(C), et c'est aussi le cas des ensembles V_κ où κ est un cardinal fortement inaccessible (voir plus haut). Comme on ne peut pas montrer l'existence d'un modèle ensemble de la théorie des ensembles, ceci montre que l'existence de cardinaux inaccessibles n'est pas une conséquence de ZF(C). En fait, on montre assez facilement que si κ est le premier cardinal inaccessible, alors V_κ est un modèle de “ZF + il n'existe aucun cardinal fortement inaccessible”. Si IC désigne l'axiome “il existe un cardinal fortement inaccessible”, on voit que ZF+IC montre la consistance de ZF ; or ZF+IC ne peut pas montrer la consistance de ZF+IC d'après le théorème d'incomplétude de Gödel ; ainsi, la consistance de ZF+IC ne peut pas être démontrée, même relativement à celle de ZF.

À partir de maintenant, nous supposerons sans aucun scrupule vérifié l'axiome de régularité.

On termine par le résultat suivant :

Théorème 4.3 (principe de réflexion). *Soient $\varphi_i(x_1, \ldots, x_n)$, pour $i = 1, \ldots, m$ des formules. Alors pour chaque ensemble M_0 il existe un ensemble $M \supseteq M_0$ tel que*

$$\forall i \forall x_1, \ldots, x_n(((M, \in) \vDash \varphi_i(x_1, \ldots, x_n)) \Leftrightarrow \varphi_i(x_1, \ldots, x_n))$$

(on dit que M reflète les formules $\varphi_1, \ldots, \varphi_m$).

1. *De plus, on peut choisir un tel ensemble M* transitif. *Plus précisément, on peut trouver α un ordinal limite tel que $V_\alpha \supseteq M_0$ reflète les formules $\varphi_1, \ldots, \varphi_m$.*
2. *Supposant l'axiome du choix, on peut trouver M de sorte que* $\operatorname{card} M \leq \operatorname{card} M_0 + \aleph_0$. *En particulier, il existe M dénombrable qui reflète $\varphi_1, \ldots, \varphi_m$.*

(On peut choisir M soit transitif soit dénombrable, mais pas les deux à la fois.)

En particulier, si $\sigma_1, \ldots, \sigma_m$ sont des formules closes vraies dans l'univers, on peut trouver un modèle ensemble de celles-ci. A fortiori, la théorie ZF n'est pas finiment axiomatisable (sans quoi on aurait démontré l'existence d'un modèle ensemble de celle-ci).

Le principe de réflexion est extrêmement puissant quand on le combine à la remarque que toute démonstration mathématique n'utilise qu'un nombre fini d'axiomes.

Le principe de réflexion est à mettre en parallèle avec le théorème de Löwenheim-Skolem : on construit en quelque sorte un sous-modèle élémentaire de l'Univers pour ce qui concerne les fonctions $\varphi_1, \ldots, \varphi_m$. La démonstration en est semblable : pour le deuxième point (utilisant l'axiome du choix), on choisit des fonctions de Skolem pour toutes les formules φ_m (écrites sans le quantificateur universel) et toutes leurs sous-formules. En l'absence d'axiome du choix, on utilise l'axiome de régularité pour choisir un ensemble de témoins pour chaque formule existentielle (plutôt qu'un seul témoin comme dans une fonction de Skolem). On renvoie au Jech pour plus de détails.