Logique du premier ordre et théorie de Zermelo-Fraenkel

Jérôme Lapuyade-Lahorgue

Logique du premier ordre et théorie de Zermelo-Fraenkel

Jérôme Lapuyade-Lahorgue - LITIS

1 Logique du premier ordre

1.1 Langage de premier ordre

La logique du premier ordre permet de démontrer des formules écrites sur un langage du type :

$$\mathcal{L} = \mathcal{S}_C \cup \mathcal{S}_R \cup \mathcal{S}_F,$$

où $\mathcal{S}_C$, $\mathcal{S}_R$ et $\mathcal{S}_F$ sont respectivement des "ensembles" de symbôles dit de constantes, de relations et de fonctions. On considère également l'ensemble $\mathcal{V}$ des symbôles de variable, communs à tous les langages du premier ordre.
Les termes d'un langage du premier ordre sont de la forme :

$$\mathcal{T} = \mathcal{S}_C | \mathcal{V} | \mathcal{S}_F(\mathcal{T}, \dots, \mathcal{T}).$$

L'ensemble des variables d'un terme est défini par induction :

- Si t est un symbôle de constante c, alors $\mathcal{V}(c)$ est $\{\}$ (vide).
- Si t est un symbôle de variable x, alors $\mathcal{V}(x)$ est $\{x\}$ (une seule variable qui est x).
- Si t est le terme $f(t_1, \dots, t_n)$, une variable de t est une variable de l'un des t_i et les variables d'un t_i sont variables de t.

Si un terme ne contient pas de variables, on dit que le terme est clos.
Les formules atomiques d'un langage du premier ordre sont de la forme :

$$\mathcal{F}_0 = \mathcal{S}_R(\mathcal{T}, \dots, \mathcal{T}).$$

1.2 Modèle et valeurs de vérité d'une formule atomique

Un modèle $\mathcal{M}$ est la donnée d'un ensemble de base M et d'un environnement e, fonction de $\mathcal{V}$ dans M. A chaque symbôle de constance c, on associe un élément $c_{\mathcal{M}}$ de M. A chaque symbôle de variable x, on associe un élément $e(x)$ de M. A chaque symbôle de fonction f à n arguments, on associe une fonction $f_{\mathcal{M}}$ de M^n dans M. A chaque symbôle de relation R à n arguments, on associe un sous-ensemble $R_{\mathcal{M}}$ de M^n.
La valeur d'un terme est définie inductivement par :

- Si c est un symbôle de constante, $\mathrm{Val}_{\mathcal{M},e}(c) = c_{\mathcal{M}}$.

– Si x est un symbôle de variable, $\mathrm{Val}_{\mathcal{M},e}(x) = e(x)$.
– Si t est le terme $f(t_1, \ldots, t_n)$, $\mathrm{Val}_{\mathcal{M},e}(t) = f_{\mathcal{M}}(\mathrm{Val}_{\mathcal{M},e}(t_1), \ldots, \mathrm{Val}_{\mathcal{M},e}(t_n))$.

La valeur d'une formule atomique est définie par :
$\mathrm{Val}_{\mathcal{M},e}(R(t_1, \ldots, t_n)) = 1$ si et seulement si $(\mathrm{Val}_{\mathcal{M},e}(t_1), \ldots, \mathrm{Val}_{\mathcal{M},e}(t_n)) \in R_{\mathcal{M}}$, et 0 sinon. Lorsque la valeur vaut 1, on dira que la formule est vraie.
Soit Γ un ensemble de formules, on dira que $\mathcal{M}, e$ satisfait Γ lorsque toutes les formules de Γ sont vraies. On notera alors :

$$\mathcal{M}, e \vDash \Gamma.$$

1.3 Règle de démonstration

Soit Γ un ensemble de formule et F une formule. On va donner un sens au fait que Γ prouve F. Γ prouve F sera noté :

$$\Gamma \vdash F.$$

Une règle de démonstration sera notée :

$$\frac{\Gamma \vdash F}{\Gamma' \vdash F'},$$

et signifiera si Γ prouve F alors nécessairement Γ' prouve F.
Afin de donner un sens au mot "prouver". Nous devons introduire la notion de règle axiomatique. Celle-ci stipule que si F est une formule de Γ, alors nécessairement Γ prouve F. Cette règle peut s'écrire :

$$\Gamma, F \vdash F$$

On voit que si $\mathcal{M}, e \vDash \Gamma$ et si $\Gamma \vdash F$ au travers de la règle d'axiome, alors nécessairement $\mathcal{M}, e \vDash F$.
On peut également considérer la règle d'affaiblissement :

$$\frac{\Gamma \vdash B}{\Gamma, A \vdash B}.$$

Au final, $\Gamma \vdash F$ si c'est la conséquence d'un nombre fini d'étapes issues de la règle d'axiome.

1.4 Logique minimale finie du premier ordre

En logique minimale, nous considérons uniquement l'implication et les deux règles la définissant :

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B},$$
$$\frac{\Gamma \vdash A \to B; \Gamma \vdash A}{\Gamma \vdash B}.$$

Ainsi les formules de la logique minimale du premier ordre sont :

$$\mathcal{F} = \mathcal{F}_0 | \mathcal{F} \to \mathcal{F}.$$

La valeur de vérité de la nouvelle formule $A \to B$ sera définie de façon à ce que si $\mathcal{M}, e \vDash \Gamma$ et si $\Gamma \vdash A \to B$, alors $\mathcal{M}, e \vDash A \to B$. Cette valeur de vérité sera alors une fonction des valeurs de vérité de A et de B. Plus précisément, nous exigerons que si une formule est prouvable à partir d'une base axiomatique vraie, alors cette formule est encore vraie.
De la règle d'axiome, on déduit facilement que $\Gamma \vdash A \to A$ indépendamment de Γ, ainsi $A \to A$ est vraie indépendamment de la valeur de vérité de A.
De la règle d'introduction de l'implication, on déduit facilement que :

$$B \vdash A \to B.$$

En effet, $B, A \vdash B$ par la règle d'axiome. Ainsi, si B est vraie alors $A \to B$ est vraie peu importe la valeur de vérité de A.
De la règle d'élimination de l'implication, on en déduit facilement que :

$$A, A \to B \vdash B.$$

Ainsi, si A et $A \to B$ sont tous les deux vrais, B est nécessairement vraie. Ainsi si A est vraie et B est fausse, $A \to B$ est obligatoirement fausse. On en déduit les valeurs de vérités de $A \to B$:

A	B	$A \to B$	$B \to A$
0	0	1	1
0	1	1	0
1	0	0	1
1	1	1	1

Par construction des valeurs de vérité, si une formule est prouvable à partir d'une base axiomatique vraie alors elle est également vraie. La réciproque n'est cependant pas valide, il existe des formules vraies et non prouvable dans la logique minimale du premier ordre.
On peut montrer que toutes les formules composées de A et de B et du symbôle logique $\to$ ont mêmes valeurs de vérité (fonction de celles de A et de B) que les formules A, B, $A \to A$, $A \to B$, $B \to A$ et $(A \to B) \to B$, ces valeurs de vérité sont résumées dans le tableau suivant :

A	B	$A \to A$	$A \to B$	$B \to A$	$(A \to B) \to B$
0	0	1	1	1	0
0	1	1	1	0	1
1	0	1	0	1	1
1	1	1	1	1	1

Il est important de remarquer que l'on peut bien sûr composer des formules avec non seulement A et B mais aussi avec plus de deux formules de base, cependant l'implication met en relation seulement deux formules, ainsi on se ramène toujours en composant à une formule du type $A \to B$. Le point important est de considérer combien de fonction des valeurs de vérité de A et de B peut-on créer avec seulement l'implication (il devrait y en avoir au maximum 16 mais ici ce n'est pas le cas).
Du tableau précédent, on en déduit la liste complète des tautologies (formule vraie indépendamment du modèle) de la logique minimale formées à partir des

formules du tableau. Nous indiquons également si la tautologie est prouvable ou non en logique minimale :

1. $A \to A$ est prouvable en logique minimale.
2. $(A \to B) \to [A \to (A \to B)]$ et $[A \to (A \to B)] \to (A \to B)$ sont prouvables en logique minimale.
3. $A \to (B \to A)$ est prouvable en logique minimale.
4. $A \to [(A \to B) \to B]$ et $B \to [(A \to B) \to B]$ sont prouvables en logique minimale.
5. $A \to [(A \to B) \to A]$ est prouvable en logique minimale.
6. Sa réciproque $[(A \to B) \to A] \to A$ n'est pas prouvable en logique minimale.
7. $[(A \to B) \to (B \to A)] \to (B \to A)$ est prouvable en logique minimale.
8. $\{(A \to B) \to [(A \to B) \to B]\} \to [(A \to B) \to B]$ est prouvable en logique minimale.
9. $\{(B \to B) \to [(A \to B) \to B]\} \to [(A \to B) \to B]$ n'est pas prouvable en logique minimale.
10. $[(A \to B) \to B] \to [(B \to A) \to A]$ n'est pas prouvable en logique minimale.

A partir de ces tautologies de base, on forme toutes les tautologies de la logique minimale de la façon suivante. Si F est une tautologie, alors :

- $A \to F$ est une tautologie.
- $(F \to A) \to A$ est une tautologie.
- En remplaçant A ou B dans une tautologie déjà formée par n'importe quelle autre formule, on obtient encore une tautologie.

On peut montrer également que toute tautologie a été construite de la manière précédente.
De plus, si F est une tautologie prouvable, les tautologies ainsi formées sont prouvables. Par conséquent, pour que toutes les tautologies soient prouvables, il suffit que les tautologies de base le soient.

1.5 Logique intuitionniste finie du premier ordre

Les formules sont également obtenues uniquement avec le symbôle logique $\to$, cependant nous ajoutons des règles afin que les formules vraies dans n'importe quel modèle (tautologies) soient prouvables pour n'importe quel ensemble d'hypothèses.
Nous ajoutons la règle (dite de Peirce) :

$$\Gamma \vdash [(A \to B) \to A] \to A.$$

A ce stade, nous pouvons alors prouver $[(A \to B) \to B] \to [(B \to A) \to A]$.
Il suffit de montrer $(A \to B) \to B, B \to A \vdash A$. $(A \to B) \to B, B \to A \vdash A$ car $(A \to B) \to B, B \to A \vdash [(A \to B) \to A] \to A$ (règle de Peirce) et

$(A \to B) \to B, B \to A \vdash (A \to B) \to A$. $(A \to B) \to B, B \to A \vdash (A \to B) \to A$ car $(A \to B) \to B, B \to A, A \to B \vdash A$. $(A \to B) \to B, B \to A, A \to B \vdash A$ car $(A \to B) \to B, B \to A, A \to B \vdash B \to A$ (axiome) et $(A \to B) \to B, B \to A, A \to B \vdash B$. $(A \to B) \to B, B \to A, A \to B \vdash B$ car $(A \to B) \to B, B \to A, A \to B \vdash A \to B$ (axiome) et $(A \to B) \to B, B \to A, A \to B \vdash (A \to B) \to B$ (axiome).
Les autres tautologies de base sont prouvables, ainsi toute tautologie est prouvable.

1.6 Logique classique finie du premier ordre

Dans ce paragraphe, nous introduisons de nouveaux connecteurs logiques. Nous montrerons également que ces nouveaux connecteurs peuvent être déduits de la logique intuitionniste. Cependant, il est préférable d'introduire ces nouveaux connecteurs indépendamment par des règles de démonstration. En effet, on peut s'imaginer l'existence d'une logique d'ordre 0 dans laquelle ces connecteurs ne peuvent être déduits de la logique intuitionniste. Par exemple, nous verrons qu'en logique du second ordre, on peut déduire tous ces connecteurs à partir seulement de la logique minimale.

1.6.1 Faux universel

Le faux universel est parfois défini en logique intuitionniste. Nous préférons l'introduire en logique classique, car il n'est pas nécessaire pour en déduire que toutes les tautologies présentées précédemment sont prouvables en logique intuitionniste. La définition du faux universel s'appuie sur le fait suivant : si $\bot_{\mathcal{M},e}$ est une formule fausse dans le modèle $\mathcal{M}, e$, alors $\bot_{\mathcal{M},e} \to B$ est vraie dans le modèle $\mathcal{M}, e$. Ce fait est indépendant de la formule B. Cependant, ce résultat n'est pas une tautologie car dépend du modèle. On définit une nouvelle tautologie en introduisant le faux universel, qui correspond à une formule fausse indépendamment du modèle. Cette formule ne peut bien sûr pas se déduire des formules de la logique intuitionniste. L'existence d'une formule universellement fausse n'a pas de réelle justification à part qu'elle nous permet de retrouver les 16 combinaisons possibles de valeurs de vérité à partir des valeurs de vérité de A et de B.
Le faux universel peut être défini via la règle d'absurdité intuitionniste :

$$\Gamma \vdash \bot \to B$$

Après introduction de la formule fausse universelle, nous constatons que toute formule formée à partir de A, B, $\to$ et $\bot$ ont pour valeur de vérité les valeurs de vérités des 6 formules présentées précédemment, des 6 formules $\bot$, $A \to \bot$, $B \to \bot$, $(A \to B) \to \bot$, $(B \to A) \to \bot$, $[(A \to B) \to B] \to \bot$, des deux formules composées $A \to (B \to \bot)$ et $(A \to B) \to [(B \to A) \to \bot]$ et des deux dernières formules $[A \to (B \to \bot)] \to \bot$ et $\{(A \to B) \to [(B \to A) \to \bot]\} \to \bot$.
Les valeurs de vérité des différentes formules sont présentées dans les tableaux suivants :

Formules de base :

A	B	$A \to A$	$A \to B$	$B \to A$	$(A \to B) \to B$
0	0	1	1	1	0
0	1	1	1	0	1
1	0	1	0	1	1
1	1	1	1	1	1

Négation des formules de base :

$A \to \bot$	$B \to \bot$	$\bot$	$(A \to B) \to \bot$	$(B \to A) \to \bot$	$[(A \to B) \to B] \to \bot$
1	1	0	0	0	1
1	0	0	0	1	0
0	1	0	1	0	0
0	0	0	0	0	0

Formules composées :

$A \to (B \to \bot)$	$(A \to B) \to [(B \to A) \to \bot]$
1	0
1	1
1	1
0	0

Négation des formules composées :

$[A \to (B \to \bot)] \to \bot$	$\{(A \to B) \to [(B \to A) \to \bot]\} \to \bot$
0	1
0	0
0	0
1	1

Apparaissent également d'autres tautologies. Il y a notamment des tautologies concernant les formules du type $F \to \bot$, où F une formule de base de la logique intuitionniste. Notamment, si F est une tautologie, alors $(F \to \bot) \to \bot$ est encore une tautologie ; si F est prouvable, $(F \to \bot) \to \bot$ est encore prouvable. De plus, si $F_1 \to F_2$ est une tautologie, alors $(F_2 \to \bot) \to (F_1 \to \bot)$ est une tautologie. On en déduit alors que toutes les tautologies concernant les douzes premières formules (les 6 de la logique intuitionniste et leur négation) sont prouvables.
Il suffit alors de montrer que les tautologies concernant les formules faisant intervenir à la fois les formules de base et leur négation sont prouvables.

1.6.2 Négation d'une formule

La négation d'une formule est définie par les deux règles d'introduction et d'élimination :

$$\frac{\Gamma, A \vdash \bot}{\Gamma \vdash \text{non}(A)},$$

$$\frac{\Gamma \vdash A; \Gamma \vdash \text{non}(A)}{\Gamma \vdash \bot}.$$

De la même façon, on définit la valeur de vérité de façon à ce que si $\mathcal{M}, e \vDash \Gamma$ et si $\Gamma \vdash \text{non}(A)$, alors $\mathcal{M}, e \vDash \text{non}(A)$. De la règle d'introduction, on déduit

$A \to \bot \vdash \text{non}(A)$, ainsi si $A \to \bot$ est vrai alors non(A) est vrai. De la règle d'élimination, on montre que non$(A) \vdash A \to \bot$, ainsi si non(A) est vrai alors $A \to \bot$ est vrai, d'où si $A \to \bot$ est faux, nécessairement non(A) est faux. La table de vérité de la négation est alors donnée par :

A	$A \to \bot$	non(A)
0	1	1
1	0	0

On déduit également que les formules vraies $(A \to \bot) \to \text{non}(A)$ et non$(A) \to (A \to \bot)$ sont prouvables. On aurait pu déduire la définition de la négation de la logique intuitionniste en posant :

$$\text{non}(A) : A \to \bot.$$

La nilpotence de la négation peut être aussi prouvée. La formule $A \to \text{non}(\text{non}(A))$ est vraie et prouvable. En effet, on a $A \vdash \text{non}(\text{non}(A))$ car $A, \text{non}(A) \vdash \bot$, on en déduit la prouvabilité par l'élimination de la négation.
De même, non$(\text{non}(A)) \vdash A$ est prouvable. En effet, il suffit de prouver non$(\text{non}(A)) \vdash (A \to \bot) \to \bot$ et non$(\text{non}(A)) \vdash [(A \to \bot) \to \bot] \to A$. On a non$(\text{non}(A)) \vdash (A \to \bot) \to \bot$ car non$(\text{non}(A)), A \to \bot \vdash \bot$, car non$(\text{non}(A)), A \to \bot \vdash \text{non}(\text{non}(A))$ (axiome) et non$(\text{non}(A)), A \to \bot \vdash \text{non}(A)$ (déjà montré). On a non$(\text{non}(A)) \vdash [(A \to \bot) \to \bot] \to A$ car $\vdash [(A \to \bot) \to \bot] \to A$ car $(A \to \bot) \to \bot \vdash A$ car $(A \to \bot) \to \bot \vdash \bot \to A$ (règle) et $(A \to \bot) \to \bot \vdash (\bot \to A) \to A$ (déjà montré).
De cela, on en déduit l'absurdité classique. Supposons $\Gamma, \text{non}(A) \vdash \bot$, alors $\Gamma \vdash \text{non}(A) \to \bot$, et comme $\Gamma \vdash (\text{non}(A) \to \bot) \to \text{non}(\text{non}(A))$, alors $\Gamma \vdash \text{non}(\text{non}(A))$ et d'après précédemment $\Gamma \vdash A$.
Reste à montrer que l'on peut prouver toute formule vraie (dans n'importe quel modèle) composée de $\to$ et de non. Ceci est trivial car c'est la conséquence du fait que non(A) est équivalente à $A \to \bot$ et aurait pu alors être défini à partir de la logique intuitionniste.

1.6.3 Conjonction

La conjonction est définie au travers des trois règles :

$$\frac{\Gamma \vdash A; \Gamma \vdash B}{\Gamma \vdash A \wedge B}.$$

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A}.$$

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}.$$

On définit avec les mêmes arguments sa valeur de vérité.
De la première règle, on déduit $A, B \vdash A \wedge B$, ainsi si A et B sont vraies alors nécessairement $A \wedge B$ est vraie. De la deuxième règle, on déduit $A \wedge B \vdash A$, ainsi si $A \wedge B$ est vraie alors nécessairement A est vraie, donc si A est fausse, alors nécessairement $A \wedge B$ est fausse. De la troisième règle, on peut déduire que si

B est fausse alors $A \wedge B$ est fausse. Les valeurs de vérité sont représentées dans le tableau suivant :

A	B	$A \wedge B$	$(A \to (B \to \bot)) \to \bot$
0	0	0	0
0	1	0	0
1	0	0	0
1	1	1	1

Les formules $A \wedge B \to [(A \to (B \to \bot)) \to \bot]$ et $[(A \to (B \to \bot)) \to \bot] \to A \wedge B$ sont vraies dans n'importe quel modèle. Montrons qu'elles sont prouvables pour n'importe quelles hypothèses.
On a $A \wedge B \vdash (A \to (B \to \bot)) \to \bot$ car $A \wedge B, A \to (B \to \bot) \vdash \bot$ car $A \wedge B, A \to (B \to \bot) \vdash B$ (élimination du $\wedge$ et affaiblissement) et $A \wedge B, A \to (B \to \bot) \vdash B \to \bot$, car $A \wedge B, A \to (B \to \bot) \vdash A \to (B \to \bot)$ (axiome) et $A \wedge B, A \to (B \to \bot) \vdash A$ (élimination du $\wedge$ et affaiblissement).
On a $(A \to (B \to \bot)) \to \bot \vdash A \wedge B$ car $(A \to (B \to \bot)) \to \bot \vdash A$ et $(A \to (B \to \bot)) \to \bot \vdash B$. $(A \to (B \to \bot)) \to \bot \vdash A$ car $(A \to (B \to \bot)) \to \bot, \text{non}(A) \vdash \bot$, car $(A \to (B \to \bot)) \to \bot, \text{non}(A) \vdash (A \to (B \to \bot)) \to \bot$ (axiome) et $\text{non}(A) \vdash A \to (B \to \bot)$ car $\text{non}(A), A \vdash B \to \bot$ car $\text{non}(A), A, B \vdash \bot$ (élimination du non). L'assertion $(A \to (B \to \bot)) \to \bot \vdash B$ se montre de manière similaire.
On en déduit que la conjonction aurait pu être introduite de manière intuitionniste via :

$$A \wedge B : (A \to (B \to \bot)) \to \bot.$$

Ainsi, toute formule vraie dans n'importe quel modèle s'écrivant avec les symbôles $\to$, non, $\bot$ et $\wedge$ est prouvable sous n'importe quelle hypothèse.

1.6.4 Disjonction

La disjonction est introduite grammaticallement via les règles :

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B}.$$

$$\frac{\Gamma \vdash B}{\Gamma \vdash A \vee B}.$$

$$\frac{\Gamma \vdash A \vee B; \Gamma, A \vdash C; \Gamma, B \vdash C}{\Gamma \vdash C}.$$

On définit de la même manière sa valeur de vérité. Des deux premières règles, on déduit que si A est vraie alors $A \vee B$ est nécessairement vraie et que si B est vraie alors $A \vee B$ est nécessairement vraie. De la dernière règle, on déduit que $A \vee B, \text{non}(A) \vdash B$ et donc si $A \vee B$ est vraie et A est fausse, alors B est nécessairement vraie. Ainsi, si A et B sont fausses, $A \vee B$ est nécessairement fausse. On a alors la table de vérité :

A	B	$A \vee B$	$(A \to B) \to B$
0	0	0	0
0	1	1	1
1	0	1	1
1	1	1	1

De même, les formules $A \vee B \to [(A \to B) \to B]$ et $[(A \to B) \to B] \to A \vee B$ sont vraies indépendamment du modèle. Montrons qu'elles sont prouvables indépendamment des hypothèses.
$A \vee B \vdash (A \to B) \to B$ car $A \vee B, A \to B \vdash B$, car $A \vee B, A \to B \vdash A \vee B$ (axiome), $A \vee B, A \to B, B \vdash B$ (axiome) et $A \vee B, A \to B, A \vdash B$ (élimination de l'implication).
$(A \to B) \to B \vdash A \vee B$ car $(A \to B) \to B, \text{non}(A \vee B) \vdash \bot$, car (utilisant élimination du non et affaiblissement) $\text{non}(A \vee B) \vdash \text{non}[(A \to B) \to B]$. $\text{non}(A \vee B) \vdash \text{non}[(A \to B) \to B]$ car $\text{non}(A \vee B) \vdash [(A \to B) \wedge \text{non}(B)] \to \text{non}[(A \to B) \to B]$ (utiliser le fait que toute formule écrite avec $\wedge$, non et $\to$ vraie dans n'importe quel modèle est prouvable) et $\text{non}(A \vee B) \vdash (A \to B) \wedge \text{non}(B)$. On a $\text{non}(A \vee B) \vdash (A \to B)$ car $\text{non}(A \vee B), A \vdash B$, car $\text{non}(A \vee B), A, \text{non}(B) \vdash \bot$, car $\text{non}(A \vee B), A, \text{non}(B) \vdash \text{non}(A \vee B)$ (axiome) et $\text{non}(A \vee B), A, \text{non}(B) \vdash A \vee B$ (introduction de $\vee$). De même, $\text{non}(A \vee B) \vdash \text{non}(B)$ car $\text{non}(A \vee B), B \vdash \bot$: utiliser ensuite même méthode.
On voit ainsi que la disjonction aurait également pu être définie de manière intuitionniste et par conséquent toutes les formules vraies dans n'importe quelle modèle peuvent être prouvables.

1.7 Logiques infinies

Dans les logiques précédentes (minimales, intuitionniste et classique finies), nous ne pouvons prouver une formule qu'à partir d'un nombre fini de formules. Afin de pouvoir prouver une formule à partir d'un nombre infini de formules, on introduit les quantifieurs universel et existentiel $\forall$ et $\exists$. Pour cela, on définit l'ensemble des variables libres et muettes d'une formule de manière récursive :

- L'ensemble des variables de $\bot$ est vide, donc l'ensemble des variables libres et l'ensemble des variables muettes aussi.
- Si la formule est atomique, l'ensemble des variables libres est l'ensemble des variables de la formule et l'ensemble des variables muettes est vide.
- Si la formule est équivalente à une formule du type $F_1 \to F_2$, l'ensemble des variables libres (resp. muettes) est la réunion des ensembles des variables libres (resp. muettes) des formules F_1 et F_2.
- Si $F = QxG$, où Q est l'un des quantifieurs $\forall$ ou $\exists$, alors les variables libres de F sont les variables libres de G sauf x et les variables muettes de F sont les variables muettes de G et x.

Une formule sans variable libre est appelée "formule close" et un ensemble de formules closes est appelée "théorie".
La substitution d'une variable x par un terme t est également définie de la manière suivante :

- Si c est un symbôle de constante, $c[x := t]$ est le même symbôle de constante.
- $x[x := t]$ est le terme t.
- Si y est un symbôle de variable différent de x, alors $y[x := t]$ est le symbôle de variable y.
- $f(t_1, \ldots, t_n)[x := t]$ est le terme $f(t_1[x := t], \ldots, t_n[x := t])$.
- $R(t_1, \ldots, t_n)[x := t]$ est la formule atomique $R(t_1[x := t], \ldots, t_n[x := t])$.
- $(F_1 \to F_2)[x := t]$ est la formule $F_1(x := t) \to F_2(x := t)$.
- Si y symbôle de variable différent de x, $(QyF)[x := t]$ est la formule $Qy, F[x := t]$.
- $(Qx)[x := t]$ est la formule QxF.

Les quantifieurs universel et existentiel sont définis au travers des règles :

$$\frac{\Gamma \vdash A}{\Gamma \vdash \forall x, A},$$

où x n'est pas une variable libre dans les formules de Γ.

$$\frac{\Gamma \vdash \forall x, A}{\Gamma \vdash A[x := t]}.$$

$$\frac{\Gamma \vdash A[x := t]}{\Gamma \vdash \exists x, A},$$

$$\frac{\Gamma \vdash \exists x, A; \Gamma, A \vdash C}{\Gamma \vdash C},$$

où x n'est ni libre dans C ni dans les formules de Γ.
Finalement, on définit les logiques minimales, intuitionnistes et classiques comme les logiques finies respectives auquelles on a inclu les quantifieurs.

1.8 Théorème de complétude de la logique du premier ordre

Par définition des règles de démonstration et construction de la notion de vérité, nous avons le théorème de correction :
Si $\mathcal{M}, e \vDash \Gamma$ et si $\Gamma \vdash F$, alors $\mathcal{M}, e \vDash F$.
En particulier, si Γ admet un modèle (on dira alors que Γ est non contradictoire), alors $\Gamma \nvdash \bot$ (on dira que Γ est consistant).
Nous considérerons ici que Γ est une théorie T. Auquel cas, les valeurs de vérité ne dépendent pas de l'environnement e et nous noterons $\mathcal{M} \vDash T$ lorsque tous les axiomes de la théorie sont vrais. On s'intéresse dans cette sous-section à la réciproque du théorème de correction. Notamment nous répondrons aux questions :

1. Si T est non contradictoire et soit $\mathcal{M}$ un modèle de T. Si $\mathcal{M} \vDash F$, a-t-on $T \vdash F$?
2. Si T est consistante, T est-elle non contradictoire ?

Pour répondre à la première question, nous introduisons la définition suivante :

Definition 1 (Théorie complète)**.** *Une théorie T est complète si pour toute formule F écrite sur son langage, on a ou bien $T \vdash F$ ou bien $T \vdash non(F)$.*

On a alors la proposition suivante :

Proposition 1. *Soit T une théorie consistante. Alors T est complète si et seulement si pour tout modèle $\mathcal{M}$ de T et toute formule F écrite sur le langage de T, $\mathcal{M} \vDash F$ ssi $T \vdash F$.*

Démonstration. Si $T \vdash F$ on a trivialement $\mathcal{M} \vDash F$ peu importe si la théorie est complète ou non.
Supposons que T soit complète et que $\mathcal{M} \vDash F$. Si on avait $T \nvdash F$, alors par complétude, $T \vdash \text{non}(F)$ et donc $\mathcal{M} \vDash \text{non}(F)$, ce qui est impossible.
Supposons que T vérifie pour tout modèle $\mathcal{M}$ de T et toute formule F écrite sur le langage de T, $\mathcal{M} \vDash F$ ssi $T \vdash F$. Montrons que T est complète. Si $T \nvdash F$, alors d'après l'hypothèse $\mathcal{M} \nvDash F$ et donc $\mathcal{M} \vDash \text{non}(F)$ et donc d'après l'hypothèse $T \vdash \text{non}(F)$. □

Le théorème de complétude suivant permet de répondre à la deuxième question :

Theorem 1 (Théorème de complétude)**.** *Si T est une théorie consistante, alors elle est non contradictoire. De plus il existe une théorie Th complète et consistante telle que $Th \vdash T$.*

Démonstration. On supposera que le langage de la théorie est au plus dénombrable ainsi que la théorie.
Première étape : Quitte à rajouter des constantes supplémentaires au langage de T, on va construire une théorie T_∞ consistante telle $T_\infty \vdash T$ et telle que si $T \vdash \exists x, F(x)$ alors $T_\infty \vdash F(c_F)$. Cela revient à donner un sens aux x qui existent (s'ils existent au sens logique, c'est qu'ils existent concrétement).
Pour cela on étend l'ensemble des constantes de la manière suivante. On pose $\mathcal{C}_0 = \mathcal{S}_C$ et :

$$\mathcal{C}_{n+1} = \mathcal{C}_n \cup \{c_F : F \text{ a une seule variable libre sur } \mathcal{C}_n\}.$$

On pose ensuite $\mathcal{C}_\infty = \bigcup_{n \geq 0} \mathcal{C}_n$. On construit récursivement la théorie T_n écrite avec les constantes $\mathcal{C}_n$ de la manière suivante. On pose $T_0 = T$ et :

$$T_{n+1} = T_n \cup \{F(c_F) : F \text{ a une seule variable libre sur } \mathcal{C}_n\}.$$

On pose $T_\infty = \bigcup_{n \geq 0} T_n$. On vérifie qu'elle est bien consistante et possède les propriétés requises.
Deuxième étape : On construit Th. Pour cela, on construit la suite de théorie suivante. Soit $(F_p)_{p \in \mathbb{N}}$ l'ensemble dénombrable des formules pouvant être écrites sur $\mathcal{C}_\infty$. On pose $K_0 = T_\infty$. Si K_n est complète, on pose $K_{n+1} = K_n$ et sinon soit p le plus petit entier tel que $K_n \nvdash F_p$ et $K_n \nvdash \text{non}(F_p)$, on pose $K_{n+1} =$

$K_n \cup \{F_p\}$. On pose $Th = \bigcup_{n \geq 0} K_n$ et on vérifie facilement que Th a les propriétés requises.

Troisième étape : On construit un modèle pour Th. Cela sera obligatoirement un modèle pour T. On construit de modèle récursivement. L'ensemble de base est $M = \mathcal{C}_\infty$.

- Interprétation des constantes : $\mathrm{Val}_{\mathcal{M}}(c) = c$.
- Interprétation des fonctions et des termes : c'est trivial.
- Interprétation des formules atomiques : $\mathrm{Val}_{\mathcal{M}}(R(t_1, \ldots, t_n)) = 1$ si et seulement si $Th \vdash R(t_1, \ldots, t_n)$. On voit ici que l'hypothèse de complétude est très importante.

Quatrième étape : Montrons que c'est bien un modèle de Th. Soit A axiome de Th. Si A est une formule atomique, elle est trivialement vraie. Si A est équivalente à $F_1 \to F_2$, on en déduit le résultat par récursion sur la complexité de la formule. Si $A = \exists x, F$, il suffit de montrer qu'il existe un terme t tel que $F[x := t]$ soit vraie. De part construction, on a nécessairement $Th \vdash F(c_F)$ et donc $F[x := c_F]$ d'où le résultat. Si $A = \forall x, F$, suffit de montrer que pour tout terme t, $F[x := t]$ soit vraie. On a $Th \vdash F[x := t]$ pour tout terme t, d'où le résultat. □

2 Théories formelle et ambiante

Avant de détailler la théorie de Zermelo-Fraenkel, précisons par ce qu'on entend par théorie formelle et théorie ambiante. Une théorie formelle du premier ordre est un ensemble de formules closes écrites sur un langage du premier ordre. Tandis que la théorie ambiante correspondante consiste d'une part à donner un sens au quantificateur existentiel. On donne ainsi un sens "existentiel" aux variables figurant devant $\exists$. Afin de construire la théorie ambiante à partir de la théorie formelle, on procède de la manière suivante. On construit à partir du langage $\mathcal{L}$ de la théorie le langage $\mathcal{L}_\infty$ comme précisé dans le théorème de complétude. La théorie ambiante est alors par définition la théorie T_∞ écrite sur le langage $\mathcal{L}_\infty$. Ainsi, lorsque la théorie formelle prouve $\exists x, F(x)$ alors la théorie ambiante prouve $F(c_F)$. Le langage de la théorie ambiante possède plus de constante que la théorie formelle car il contient entre autre toutes les constantes que la théorie formelle est amenée à créer.

Lorsque le langage de la théorie contient le symbôle de relation $=$ et que la théorie prouve $t = t$ et $(t = s) \wedge F[x := t] \to F[x := s]$, alors on peut également ajouter des symbôles de fonctions au langage de la théorie. On appelle formule fonctionnelle à n variables toute formule $F(x_1, \ldots, x_n, y)$ ayant $x_1, \ldots, x_n, y$ pour variables libres et vérifiant $F(x_1, \ldots, x_n, y := t) \wedge F(x_1, \ldots, x_n, y := s) \to t = s$. A chaque formule fonctionnelle, on ajoute au langage un nouveau symbôle de fonction f_F. On ajoute également à la théorie la formule $F(x_1, \ldots, x_n, y) \to y = f_F(x_1, \ldots, x_n)$.

La théorie ambiante de celle de Zermelo-Fraenkel que nous allons voir maintenant est écrite sur un langage contenant presque toutes les constantes mathématiques

existantes et presque toutes les fonctions existantes. Il y a tout de même un problème : la théorie formelle de Zermelo-Fraenkel est écrite sur un langage fini, ainsi le langage ambiant est au plus dénombrable et donc il existe des constantes mathématiques qui ne figurent pas dans l'ensemble des constantes du langage ambiant.

3 Théorie de Zermelo-Fraenkel

La théorie de Zermelo-Fraenkel est écrite sur le langage $\mathcal{S}_C = \{\}$ (pas de symbôles de constante), $\mathcal{S}_C = \{\}$ (pas de symbôles de fonction) et $\mathcal{S}_R = \{\in, =\}$ (seulement deux symbôles de relation).

3.1 Axiome de la théorie

Elle est constituée des axiomes suivants permettant de construire un ensemble à partir d'ensembles existants ainsi qu'un axiome énonçant l'existence d'au moins un ensemble, ce dernier axiome sera détaillé dans la suite. Pour simplifier, on introduit les formules :

$$A \leftrightarrow B : (A \to B) \wedge (B \to A),$$

$$A \subset B : \forall x, (x \in A) \to (x \in B),$$

et :

$$x \notin y : \text{non}(x \in y).$$

On remarque que $\leftrightarrow$ est un symbôle logique pouvant être défini à partir des symbôles déjà existants, tandis que $\subset$ est un symbôle de relation qui aurait pu être ajouté au langage. Ainsi, si on définit $\leftrightarrow$ à partir de règles d'inférence logique, nous aboutissons à une tautologie tandis que si on définit $\subset$ par un axiome, nous aboutissons à un théorème concernant $\subset$.

- Axiome d'extensionnalité :

$$\forall x, \forall y, (x = y) \leftrightarrow [\forall z, (z \in x) \leftrightarrow (z \in y)]. \tag{1}$$

- Axiome de la paire :

$$\forall x, \forall y, \exists z, \forall t, (t \in z) \leftrightarrow [(t = x) \vee (t = y)]. \tag{2}$$

- Axiome de la réunion :

$$\forall x, \exists y, \forall z, (z \in y) \leftrightarrow [\exists t, (t \in x) \wedge (z \in t)]. \tag{3}$$

- Axiome des parties :

$$\forall x, \exists y, \forall z, (z \in y) \leftrightarrow (z \subset x). \tag{4}$$

- Axiome de compréhension : Soit $F(z)$ une formule ayant pour variable libre z,

$$\forall x, \exists y, \forall z, (z \in y) \leftrightarrow [(z \in x) \wedge F(z)]. \tag{5}$$

– Axiome de remplacement : Soit $F(t,z)$ une formule ayant pour variables libres t et z telle que $[F(t,z:=u_1)\wedge F(t,z:=u_2)]\rightarrow u_1=u_2$ (on dit que F est fonctionnelle), alors :

$$\forall x,\exists y,\forall z,(z\in y)\leftrightarrow[\exists t,(t\in x)\wedge F(t,z)]. \tag{6}$$

– Axiome de l'infini :

$$\exists x, OI(x), \tag{7}$$

où $OI(x)$ est la formule présentée dans le chapitre suivant et exprimant qu'un ensemble est un ordinal infini.
Avant d'aborder les ordinaux, voyons ce que l'on peut déjà déduire des axiomes de Zermelo-Fraenkel. Nous avons vu que nous pouvions déduire de nouveaux symbôles de fonctions ainsi que les règles les régissant à partir de relation fonctionnelle. La théorie de Zermelo-Fraenkel va plus loin en interprétant les fonctions comme étant des ensembles. Pour cela, on doit définir le produit cartésien de deux ensembles. Soient x et y deux ensembles, de l'axiome de la paire on peut construire l'ensemble dont les éléments sont x et y. Si a et b apparaissent comme étant des constantes du langage ambiant, on en déduit la nouvelle constante $\{a,b\}$ qui correspond alors à l'ensemble dont les éléments sont a et b. On déduit également la constante $\{a,\{a,b\}\}$ que l'on notera (a,b) (couple). Soient E et F deux ensembles (constantes figurant dans le langage ambiant), à $x\in E$ fixé, $F_x(y,z):z=(x,y)$ est une relation fonctionnelle liant y et z ainsi il existe un ensemble dont les éléments z sont tel qu'il existe un $y\in F$ vérifiant $F_x(y,z)$. Ceci nous amène à définir la constante notée $\{x\}\times F$. De même, $F(x,z):z=\{x\}\times F$ est une relation fonctionnelle, ainsi on peut introduire la nouvelle constante $E\times F$, appelé produit cartésien de E et de F.
Soit maintenant $F(x,y)$ une relation fonctionnelle, par l'axiome de compréhension, on peut construire le sous-ensemble G de $E\times F$ des éléments vérifiant $F(x,y)$. Ainsi, une fonction peut être interprétée comme un triplet (E,F,G) où $G\subset E\times F$ vérifie si $(x,y)\in G$ et $(x,y')\in G$ alors $y=y'$. On peut alors interpréter une relation fonctionnelle non plus comme un nouveau symbôle de fonction mais comme un symbôle de constante dans le langage ambiant. En fait, la relation fonctionnelle peut être interprétée comme plusieurs symbôles de constante selon le choix de E et de F. Finallement, le notion de fonction dépend non seulement de la relation mais également des ensembles de définition et d'arrivée. Il n'est d'ailleurs pas toujours judicieux d'interpréter une relation fonctionnelle comme un symbôle de constante. En effet, il faut créer une instance pour chaque couple (E,F) si on veut interpréter les fonctions comme des symbôles de constante. Si on interprète les fonctions comme des symbôles de fonctions, seule les règles d'inférence suffisent pour définir entièrement la fonction. Ce procédé est analogue aux classes template du langage C++ dans lesquelles on définit les fonctions membres pouvant être utilisées par différents types d'objet plutôt que de définir ces fonctions dans toutes les classes les utilisant.

3.2 Théorie des ordinaux

3.2.1 Définition formelle des ordinaux

Un ordinal α est défini par la formule $O(\alpha) = O_{\text{strict}}(\alpha) \wedge O_{\text{total}}(\alpha) \wedge O_{\text{ordre}}(\alpha) \wedge O_{\text{bon}}(\alpha) \wedge O_{\text{transit}}(\alpha)$, où :

- $O_{\text{strict}}(\alpha)$:
$$\forall x, (x \in \alpha) \rightarrow (x \notin x). \tag{8}$$
- $O_{\text{total}}(\alpha)$:
$$\forall x, \forall y, [(x \in \alpha) \wedge (y \in \alpha)] \rightarrow [(x = y) \vee (x \in y) \vee (y \in x)]. \tag{9}$$
- $O_{\text{ordre}}(\alpha)$:
$$\forall x, \forall y, \forall z, [(x \in \alpha) \wedge (y \in \alpha) \wedge (z \in \alpha) \wedge (x \in y) \wedge (y \in z)] \rightarrow (x \in z). \tag{10}$$
- $O_{\text{bon}}(\alpha)$:
$$\forall x, [(x \subset \alpha) \wedge (\exists t, (t \in x))] \rightarrow [\exists t, (t \in x) \wedge (\forall z, (z \in t) \rightarrow (z \notin x))]. \tag{11}$$
- $O_{\text{transit}}(\alpha)$:
$$\forall x, (x \in \alpha) \rightarrow (x \subset \alpha). \tag{12}$$

L'axiome de l'infini nous assure l'existence d'au moins un ordinal. De plus, de l'axiome de compréhension, on en déduit l'existence d'un ensemble vide à partir de cet ordinal. L'universalité du faux nous assure l'unicité de l'ensemble vide. L'ensemble vide vérifiant n'importe quelle propriété est bien sûr un ordinal.

3.2.2 Propriétés des ordinaux

Les démonstrations des différentes propriétés se feront en langue française pour éviter d'alourdir la compréhension. On traduira alors $x \in y$ par x appartient à y ou bien est un élément de y. On traduira $x \subset y$ par x est une partie de y. Si les 4 premières propriétés des ordinaux sont satisfaites, on dira que la relation $\in$ est une relation de bon ordre sur l'ensemble en question.

1. Si α est un ordinal, alors $\alpha \notin \alpha$. (trivial)
2. Tous les éléments de α sont des ordinaux :
   Soit x un élément de α, alors c'est une partie de α. Toute partie de x est aussi une partie de α, ainsi $\in$ est une relation de bon ordre sur x. Reste à montrer qu'un élément de x est une partie de x. Soit $y \in x$ et soit $t \in y$. Comme x est une partie de α, ainsi $y \in \alpha$ et donc y est une partie de α, on en déduit que t est également un élément de α. Comme t, y et x sont des éléments de α et que $t \in y$ et $y \in x$, alors on en déduit $t \in x$ ainsi y est bien une partie de x.
3. Si α est un ordinal non vide, son plus petit élément pour la relation $\in$ est l'ensemble vide :
   Soit x le plus petit élément de α. Si x était non vide, alors il existe $y \in x$, mais dans ce cas y est encore plus petit : contradictoire.

4. Si α et β sont deux ordinaux, alors $\alpha \in \beta$ ou $\beta \in \alpha$ ou $\alpha = \beta$:
   Si un des deux ordinaux est vide, c'est trivial. Supposons alors que les deux ordinaux sont non vides et supposons $\alpha \neq \beta$. Pour fixer les choses, on va supposer que $\alpha \backslash \beta$ est non vide. D'après la propriété de bon ordre, $\alpha \backslash \beta$ admet un plus petit élément, noté γ. Soit $z \in \gamma$. Comme $\gamma \in \alpha$, alors $z \in \alpha$, et comme γ est le plus petit élément de $\alpha \backslash \beta$, alors $z \in \beta$. Ainsi $\gamma \subset \alpha \cap \beta$. Soit $z \in \alpha \cap \beta$, z et γ appartiennent tout deux à l'ordinal α donc ou bien $z = \gamma$ ou bien $\gamma \in z$ ou bien $z \in \gamma$. Si $z = \gamma$, alors $\gamma \in \alpha \cap \beta$ ce qui est contradictoire. Si $\gamma \in z$, comme $z \in \beta$ alors $\gamma \in \beta$ qui est également contradictoire. Ainsi $z \in \gamma$. On a donc $\gamma = \alpha \cap \beta$. Supposons que $\beta \backslash \alpha$ soit également non vide, alors son plus petit élément est également $\alpha \cap \beta$; ceci est contradictoire avec le fait que $\alpha \backslash \beta$ et $\beta \backslash \alpha$ ne peuvent pas avoir d'éléments communs. On en déduit que $\beta \subset \alpha$ et que $\gamma = \beta$ ainsi $\beta \in \alpha$.
5. Si α et β sont deux ordinaux distincts tels que $\beta \subset \alpha$, alors le plus petit élément de $\alpha \backslash \beta$ est β. (trivial à partir de la preuve précédente).
6. Soient α et β deux ordinaux, alors $\alpha \subset \beta$ si et seulement si $\alpha = \beta$ ou $\alpha \in \beta$ (de la preuve précédente).
7. Si α est un ordinal, en utilisant l'axiome de la paire et celui de la réunion, on peut construire l'ensemble $\alpha \cup \{\alpha\}$ dont les éléments sont ceux de α et α lui-même. $\alpha \cup \{\alpha\}$ est également un ordinal. (preuve facile).
8. Soit α et β deux ordinaux. Si $\beta \in \alpha$ alors $\beta \cup \{\beta\} \subset \alpha$ (preuve facile).
9. Il n'existe pas d'ordinaux strictement plus grand que α et strictement plus petit que $\alpha \cup \{\alpha\}$ (preuve facile).
10. Un ordinal α admet un plus grand élément γ si et seulement si $\alpha = \gamma \cup \{\gamma\}$ (preuve facile).

On montre très facilement qu'il existe des ordinaux non vides ayant un plus grand élément. C'est le cas de $\{\emptyset\}$, noté 1. Par contre, sans l'axiome de l'infini, nous ne pouvons déduire qu'il existe aussi des ordinaux qui n'ont pas de plus grand élément. Introduisons les définitions suivantes :

Definition 2. *On dit qu'un ordinal non vide est :*
- *Successeur : s'il admet un plus grand élément.*
- *Limite : s'il n'admet pas de plus grand élément.*
- *Infini : s'il est limite ou si un de ses éléments est un ordinal limite.*

Finallement on définit la formule $OI(\alpha) = O(\alpha) \wedge I(\alpha)$, où $I(\alpha)$ est la formule décrivant un ordinal infini.
On sait créer un ordinal à partir d'un ordinal déjà existant, ceci peut se faire un associant à un ordinal α son successeur $\alpha \cup \{\alpha\}$. Les trois propositions suivantes nous permettent également de créer de nouveaux ordinaux à partir d'ordinaux déjà existants.

Proposition 2. *Soit A un ensemble non vide d'ordinaux, alors $\cap A$ défini comme l'ensemble des éléments qui appartiennent à tous les éléments de A est un ordinal. De plus, c'est le plus petit élément de A, noté $min(A)$.*

Démonstration. D'après l'axiome de la réunion, on peut définir l'ensemble $\cup A$, de plus $\cap A = \{x \in \cup A : \forall y \in A, x \in y\}$, donc d'après l'axiome de compréhension, $\cap A$ est bien un ensemble.
Montrons que c'est un ordinal. Trois éléments de $\cap A$ appartiennent forcément au même ordinal donc $\in$ est une relation d'ordre total et strict sur $\cap A$. Soit B une partie non vide de $\cap A$, B est alors une partie non vide d'un ordinal donc admet un plus petit élément. Soit $x \in \cap A$ et $y \in x$. Soit $\alpha \in A$, on a $x \in \alpha$ donc $x \subset \alpha$ et donc $y \in \alpha$ ainsi $y \in \cap A$. On en déduit que $\cap A$ est un ordinal.
Montrons que c'est le plus petit élément de A. Soit $\beta \in A$, si on avait $\beta \in \cap A$, alors $\beta \in \beta$ ce qui est contradictoire. □

Proposition 3. *Soit C une classe non vide d'ordinaux (non nécessairement un ensemble), alors C admet un plus petit élément.*

Démonstration. Soit $\alpha \in C$ et $A = \{\beta \in \alpha : \beta \in C\}$, comme α est un ensemble alors A est un ensemble par l'axiome de compréhension. Si A est vide, alors α est le plus petit élément de C. Si A est non vide, d'après précédemment, A admet un plus petit élément qui est également plus petit élément de C. □

Proposition 4. *Soit A un ensemble d'ordinaux, alors $\cup A$ est un ordinal. De plus, $\cup A$ est le plus petit ordinal supérieur ou égal aux éléments de A. On le note* $\sup A$.

Démonstration. Les ordinaux pouvant être comparés, trois éléments de $\cup A$ appartiennent au même ordinal. Ainsi $\in$ est une relation d'ordre total strict sur $\cup A$. Soit B partie non vide de $\cup A$, alors B est en particulier un ensemble non vide d'ordinaux donc possède un plus petit élément. Soit $x \in \cup A$ et $y \in x$, alors il existe $\alpha \in A$ tel que $x \in \alpha$, on en déduit $y \in \alpha$ donc $y \in \cup A$. Ainsi $\cup A$ est un ordinal.
Soit β ordinal tel que pour tout $\alpha \in A$, $\beta = \alpha$ ou $\alpha \in \beta$. Si $\beta \in \cup A$, alors il existe $\alpha \in A$ tel que $\beta \in \alpha$. On en déduit alors que $\beta \in \beta$ ce qui est contradictoire. □

Finalement, la proposition suivante nous donne une condition nécessaire et suffisante pour qu'un ordinal soit une limite :

Proposition 5. *Soit α un ordinal. Alors α est un ordinal limite si et seulement si $\alpha = \sup \alpha$.*

Démonstration. Supposons $\alpha = \sup \alpha = \bigcup_{\gamma \in \alpha} \gamma$. Si α possédait un plus grand élément δ, alors d'une part $\delta = \sup \alpha$ et d'autre part $\alpha = \delta \cup \{\delta\}$, d'où $\delta = \delta \cup \{\delta\}$, ce qui est contradictoire. Ainsi, α est un ordinal limite
Si $\alpha \neq \sup \alpha$, alors soit $\alpha \in \sup \alpha$, ce qui est impossible car α appartiendrait alors à un de ses éléments. Ainsi $\sup \alpha \in \alpha$ et donc $\sup \alpha$ est le plus grand élément de α. □

3.2.3 Ensembles bien ordonnés et récurrence transfinie

Les ordinaux sont des cas particuliers d'ensembles bien ordonnés. On dit qu'un ensemble est bien ordonné s'il est muni d'une relation d'ordre strict et totale telle que toute partie non vide a un plus petit élément.
Un segment initial S d'un ensemble ordonné X est un sous-ensemble tel que si $y \in S$ et si $x < y$, alors $x \in S$. Dans un ensemble bien ordonné, un segment initial est soit égal à X soit à $S_x = \{y \in X : y < x\}$.
Un segment initial d'un ordinal est un ordinal. De manière réciproque, un élément d'un ordinal est un segment initial.
Les ensembles bien ordonnés ont de belles propriétés, notamment nous pouvons généraliser la célèbre propriété de récurrence.

Proposition 6 (Propriété de récurrence)**.** *Soit X un ensemble bien ordonné et soit $P(x)$ une propriété ayant pour variable libre x. Si on a :*

$$\forall x, \{[\forall y < x, P(y)] \to P(x)\},$$

alors on a :

$$\forall x, P(x).$$

Démonstration. Soit $A = \{x \in X : \text{non}(P(x))\}$. Si A est vide, c'est trivial. Si A est non vide, alors il possède un plus petit élément x. Si $y < x$, étant plus petit que x, alors $y \notin A$ et donc $P(y)$, on en déduit alors $P(x)$, ce qui est contradictoire. □

La récurrence transfinie a de nombreuses conséquences très intéressantes dont celles-ci :

Proposition 7. *Soit X un ensemble bien ordonné et f une fonction strictement croissante de X dans X, alors pour tout $x \in X$, $x \leq f(x)$.*

Démonstration. Soit x tel que pour tout $y < x$, on a $y \leq f(y)$. Montrons alors que $x \leq f(x)$. Supposons que $x > f(x)$. Alors posant $y = f(x)$, on a $f(x) \leq f^2(x)$. Or puisque $x > f(x)$ et f croissante alors $f(x) > f^2(x)$ ce qui est contradictoire. On conclut par la propriété de récurrence transfinie. □

Proposition 8. *Soit X un ensemble bien ordonné et W un segment initial de X. Si $f : X \to W$ est un isomorphisme (ie. bijection croissante), alors $X = W$ et f est l'application identité.*

Démonstration. Pour tout $x \in X$, on a $f(x) \in W$. f étant un isomorphisme, alors f est srtictement croissante et donc $x \leq f(x)$. W étant un segment initial, donc $x \in W$. f^{-1} est également un isomorphisme, donc $x \leq f^{-1}(x)$ et par composition avec f, on a $f(x) \leq x$. □

Proposition 9. *Tout ensemble bien ordonné est isomorphe à un unique ordinal au travers d'un unique isomorphisme.*

Démonstration. Montrons d'abord que si un ensemble bien ordonné est isomorphe à un ordinal, alors l'ordinal est unique et l'isomorphisme est également unique. Soit X un tel ensemble et α et β deux ordinaux isomorphes avec X. Alors α et β sont isomorphes. Si on avait par exemple $\alpha \in \beta$, alors α serait segment initial de β et donc $\alpha = \beta$. On en déduit que $\alpha = \beta$. De plus, si f est l'isomorphisme de X sur α et g est l'isomorphisme de X sur β, alors $g \circ f^{-1}$ est l'identité de α, l'isomorphisme est donc unique.
Montrons maintenant l'existence. Soit Y l'ensemble des segments initiaux de X isomorphes chacun à un ordinal. Y est non vide car contient les singletons. Pour $I \in Y$, on note α_I l'unique ordinal isomorphe à I et f_I l'unique isomorphisme de I sur α_I. Soit $J \subset I$ un segment initial. Alors $f_I(J)$ est un segment initial de α_I, c'est donc un ordinal. La corestriction de f_I à J induit un isomorphisme de J sur $f_I(J)$, ainsi $J \in Y$, $f_I(J) = \alpha_J$ et f_J est la restriction de f_I à J.
Soit K la réunion des éléments de Y. On montre facilement que K est un segment initial. Soit $\alpha = \sup_{I \in Y}(\alpha_I)$. On définit f de K sur α par $f(x) = f_I(x)$ si $x \in I$. Montrons que f est un isomorphisme. Soit $\beta \in \alpha$, alors il existe $I \in Y$ tel que $\beta \in \alpha_I$, ainsi il existe $x \in I$ tel que $\beta = f_I(x) = f(x)$ et donc f est surjective. Si $x < y$ alors si $y \in I$, on a $x \in I$ car I est un segment initial et donc $f_I(x) < f_I(y)$ ainsi $f(x) < f(y)$, f est donc injective et est un morphisme. K est alors le plus grand élément de Y.
Supposons que $K \neq X$, alors il existe $x \in X$ tel que $K = S_x = \{y \in X : y < x\}$. $K \cup \{x\}$ est encore un segment initial. Soit l'application g définie par :

$$g : K \cup \{x\} \quad \to \quad \alpha \cup \{\alpha\}$$
$$y \to \begin{cases} f(y) \text{ si } y \in K \\ \alpha \text{ si } y = x \end{cases}$$

On montre que g est un isomorphisme, ce qui contredit la maximalité de K. □

3.2.4 Ordinaux infinis et arithmétique des ordinaux

L'axiome de l'infini affirme qu'il existe un ordinal infini. Il serait alors intéressant d'en construire. Soit α un ordinal infini et soit A l'ensemble des éléments de α qui sont infinis. Si A est vide, on montre facilement que α est le plus petit des ordinaux infinis, on le note ω. Si A est non vide, alors il possède un plus petit élément, on montre que ce plus petit élément est le plus petit des ordinaux infinis. On précisera que ω est la constante du langage ambiant définie à partir de la formule traduite en français par : "Il existe un plus petit ordinal infini".
Soit α un ordinal fini, on ne peut pas avoir $\omega \in \alpha$, car sinon α serait infini, donc $\alpha \in \omega$. Réciproquement, si $\alpha \in \omega$, alors α est nécessairement fini. Ainsi, ω est l'ensemble des ordinaux finis, c'est donc un ordinal limite.
Nous avons vu que tout ordinal non vide possède un plus petit élément et que ce plus petit élément est nécessairement l'ensemble vide. L'existence de l'ensemble vide permet de définir la constante 0 dans le langage ambiant. L'axiome de la paire permet ensuite de définir la constante 1 qui correspond à l'ensemble possédant pour unique élément 0.

Il est aussi utile de définir des nouveaux symbôles de fonctions et les règles les définissant :

1. Successeur : De la relation fonctionnelle $y = x \cup \{x\}$ (qui est en fait le résumé de la formule beaucoup plus lourde à écrire qui exprime que les éléments de y sont exactement les éléments de x ou x), on introduit dans le langage ambiant le symbôle de fonction S. On introduit dans la théorie ambiante, la formule $(y = x \cup \{x\}) \to (y = Sx)$.
2. Addition : on ajoute le symbôle $+$ et on définit les règles :
   - $\alpha + 0 = \alpha$.
   - $\alpha + 1 = S\alpha$.
   - $\alpha + (\beta + 1) = (\alpha + \beta) + 1$.
   - Si β est un ordinal limite, $\alpha + \beta = \bigcup_{\gamma \in \beta} (\alpha + \gamma)$.
3. Multiplication : on ajoute le symbôle $.$ et on définit les règles :
   - $\alpha.0 = 0$.
   - $\alpha.(\beta + 1) = \alpha.\beta + \alpha$.
   - Si β est un ordinal limite, $\alpha.\beta = \bigcup_{\gamma \in \beta} (\alpha.\gamma)$.
4. Puissance : on ajoute le symbôle $.^{\cdot}$ et on définit les règles :
   - $\alpha^0 = 1$.
   - $\alpha^{\beta+1} = \alpha^\beta.\alpha$.
   - Si β est un ordinal limite, $\alpha^\beta = \bigcup_{\gamma \in \beta} (\alpha^\gamma)$

Nous voyons que ces symbôles de fonctions ont été définis à travers de formules fonctionnelles. Pour s'en convaincre, écrivons la relation fonctionnelle définissant l'addition :

$$\begin{aligned} F_+((\alpha, \beta), \gamma) \quad &: \quad [(\beta = 0) \to (\gamma = \alpha)] \\ &\wedge \quad [(\beta = 1) \to (\gamma = S\alpha)] \\ &\wedge \quad [(\beta = S\delta) \to (\gamma = S(\alpha + \beta))] \\ &... \end{aligned}$$

Reste à vérifier que ce sont bien des relations fonctionnelles. Pour cela, on vérifie que $+$, $.$ et $.^{\cdot}$ sont des fonctions, c'est à dire qu'à un couple (α, β) correspond un unique résultat. Pour cela, on va utiliser la propriété de récurrence transfinie sur l'ordinal $\beta \cup \{\beta\}$.
Pour l'addition. Soit $\gamma \in \beta \cup \{\beta\}$, supposons que pour tout $\delta \in \gamma$, $\alpha + \delta$ soit bien définie, montrons que $\alpha + \gamma$ l'est également. Si γ est égal à 0 ou 1 : pas de problème. Si $\gamma = \delta + 1$, alors $\alpha + \gamma = (\alpha + \delta) + 1$ et comme $\alpha + \delta$ est bien définie, alors $\alpha + \gamma$ aussi. Si γ est un ordinal limite, alors $\alpha + \gamma = \bigcup_{\delta \in \gamma} (\alpha + \delta)$, comme les $\alpha + \delta$ sont bien définies pour tout $\delta \in \gamma$ alors $\alpha + \gamma$ est bien définie. Par récurrence, pour tout $\gamma \in \beta \cup \{\beta\}$, $\alpha + \gamma$ est bien définie. On en déduit que $\alpha + \beta$ est bien définie et donc $+$ est bien une fonction.
Le raisonnement est le même pour la multiplication et la puissance.

3.2.5 Lemme de Hartogs

L'addition, la multiplication et la puissance présentées ci-dessus permettent d'obtenir des ordinaux de plus en plus grand au sens de $\in$. Cependant, comme nous allons le voir par la suite, les ordinaux construits à partir de ω par récurrence transfinie en utilisant ces opérations sont tous dénombrables. C'est d'ailleurs également le cas du plus petit ordinal vérifiant $\omega^{\epsilon_0} = \epsilon_0$ qui est le plus petit "grand ordinal". On peut alors se demander s'il existe des ordinaux non dénombrables. Le lemme de Hartogs donne la réponse à cette question.

Proposition 10 (Lemme de Hartogs)**.** *Soit X un ensemble (non forcément ordonné), il existe un ordinal qui ne s'injecte pas dans X. Le plus petit ordinal ne s'injectant pas dans X est appelé "ordinal de Hartogs" de X.*

Démonstration. Soit Y l'ensemble des couples $(S, \leq)$ où $S \subset X$ et $\leq$ est une relation de bon ordre sur S. En particulier Y est une partie de $\mathcal{P}(X) \times \mathcal{P}(X \times X)$ et donc est un ensemble. Chaque S muni de $\leq$ est alors isomorphe à un unique ordinal. Ceci définit une fonction de Y vers la classe des ordinaux. En utilisant l'axiome de remplacement, on construit l'image de cette fonction, soit α_X cet ensemble. On montre alors que α_X est un ordinal et qu'il est l'ensemble des ordinaux qui s'injectent dans X ; c'est donc le plus petit ordinal ne s'injectant pas dans X. □

Avant d'aborder la théorie des cardinaux. Il est important de montrer qu'un ordinal de Hartogs d'un ordinal infini est un ordinal limite. Commençons par la proposition suivante :

Proposition 11. *Soit α ordinal infini, alors $\alpha + 1$ s'injecte dans α.*

Démonstration. Soit A l'ensemble des éléments infinis de $\alpha + 1$. On montre facilement que A est bien ordonné. On va ainsi montrer par récurrence transfinie que pour tout $\beta \in A$, $\beta + 1$ s'injecte dans β. En particulier, on aura que $\alpha + 1$ s'injecte dans α. Soit alors $\beta \in A$ tel que pour tout $\gamma \in A$ tel que $\gamma \in \beta$, on a $\gamma + 1$ s'injecte dans γ. Montrons que $\beta + 1$ s'injecte dans β.
Si β est un ordinal limite, alors :

$$\begin{aligned} f : \beta + 1 &\to \beta \\ \beta &\to \emptyset \\ \gamma \in \beta &\to \gamma + 1, \end{aligned}$$

est clairement une injection (non forcément surjective car β peut avoir des éléments limite).
Si $\beta = \gamma + 1$. On a $\gamma \in \beta$ et est encore infini car sinon β serait fini. Ainsi $\beta = \gamma + 1$ s'injecte dans γ. Soit f l'injection de β dans γ. Alors l'application :

$$\begin{aligned} g : \beta + 1 &\to \beta = \gamma + 1 \\ \beta &\to \gamma \\ \delta \in \beta &\to f(\delta) \neq \gamma \end{aligned}$$

est injective. □

Du théorème de Schröder-Bernstein, on déduit même que si α est infini, alors α et $\alpha + 1$ sont en bijection.

Theorem 2 (Théorème de Schröder-Bernstein)**.** *Soient X et Y deux ensembles. S'il existe une injection de X dans Y et une injection de Y dans X, alors il existe une bijection entre X et Y.*

Démonstration. A rédiger. □

Nous pouvons finalement en déduire la proposition suivante :

Proposition 12. *Nous avons équivalence entre les trois assertions suivantes :*

1. *α est un ordinal infini.*
2. *α et $\alpha + 1$ sont équipotents.*
3. *Pour tout $x \in \alpha$, α et $\alpha - \{x\}$ sont équipotents.*

Démonstration. 1. implique 2. a déjà été prouvé. Montrons que 2. implique 3. Soit f bijection de $\alpha + 1$ sur α, par restriction, f induit une bijection de α sur $\alpha - \{f(\alpha)\}$. Soit $x \neq f(\alpha)$, l'application :

$$\begin{array}{rcl} g : \alpha - \{f(\alpha)\} & \to & \alpha - \{x\} \\ x & \to & f(\alpha) \\ y \neq x & \to & y \end{array}$$

est clairement une bijection. On en déduit que α et $\alpha - \{x\}$ sont équipotents. Montrons que 3. implique 2. Soit f bijection de $\alpha - \{x\}$ sur α, alors l'application :

$$\begin{array}{rcl} g : \alpha & \to & \alpha + 1 \\ x & \to & \alpha \\ y \neq x & \to & f(y) \end{array}$$

est clairement une bijection.
Montrons que 2. implique 1. Il suffit de montrer que si α est un ordinal fini, alors α et $\alpha + 1$ ne sont pas équipotents. Nous procédons par récurrence sur ω. 0 est clairement non équipotent à 1. Soit $\alpha \in \omega$ tel que $\alpha \neq 0$ et pour tout $\beta \in \alpha$, β et $\beta + 1$ ne sont pas équipotents. Comme α est fini, alors il existe β tel que $\alpha = \beta + 1$. Il s'ensuit que α et $\beta = \alpha - \{\beta\}$ ne sont pas équipotents. D'après précédemment, on en déduit que α et $\alpha + 1$ ne sont pas équipotents. □

Comme conséquence, nous avons la proposition suivante :

Proposition 13. *Si α est un ordinal infini, alors son ordinal de Hartogs α^+ est un ordinal limite (réciproque vraie).*
Si α est un ordinal fini, alors $\alpha^+ = \alpha + 1$ est également fini (réciproque vraie).

Démonstration. Comme α est infini, alors α^+ est également infini. Supposons que $\sup(\alpha^+) \in \alpha^+$. Alors $\sup(\alpha^+)$ est obligatoirement infini. Ainsi $\sup(\alpha^+) + 1$ s'injecte dans $\sup(\alpha^+)$. De plus, $\sup(\alpha^+) \in \alpha^+$ entraîne que $\sup(\alpha^+)$ s'injecte dans α. On en déduit que $\sup(\alpha^+)+1$ s'injecte aussi dans α et donc $\sup(\alpha^+)+1 \in \alpha^+$, ce qui contredit la maximalité de $\sup(\alpha^+)$. □

3.3 Théorie des cardinaux

La théorie des ordinaux permet d'établir des classes d'équivalence pour les ensembles bien ordonnés. Nous avons vu qu'un ordinal est plus petit qu'un autre s'il est élément de ce dernier. Cependant, nous aimerions comparer les ordinaux en terme d'injection. Ceci peut se faire grâce à des ordinaux particuliers appelés "cardinaux". De plus, l'axiome du choix présenté ci-dessous nous permet de comparer tous les ensembles en terme d'injection.

3.3.1 Axiome du choix

Proposition 14. *Sous ZF, nous avons équivalence entre :*

1. *Pour tout ensemble non vide X, il existe une application φ, appelée fonction de choix, de $\mathcal{P}(X)-\{\emptyset\}$ dans X telle que pour tout $Y \subset X$, $\varphi(Y) \in Y$.*
2. *Pour toute famille $(X_i)_{i\in I}$ d'ensembles non vides, le produit cartésien (infini) $\prod_{i\in I} X_i$, défini comme l'ensemble des fonctions de I dans $\bigcup_{i\in I} X_i$ telles que $f(i) \in X_i$, est également non vide.*

Démonstration. Supposons 1. Alors il existe une fonction de choix de $\mathcal{P}\left(\bigcup_{i\in I} X_i\right) - \{\emptyset\}$ dans $\bigcup_{i\in I} X_i$. On pose alors $f(i) = \varphi(X_i)$ pour tout $i \in I$ et on vérifie que $f \in \prod_{i\in I} X_i$.

Supposons 2. D'après 2., l'ensemble $\prod_{\mathcal{P}(X)-\{\emptyset\}} X$ est non vide, ces éléments sont des fonctions de choix. □

L'axiome du choix est l'énoncé suivant :
Axiome du choix : Tout ensemble non vide admet une fonction de choix.
Nous montrerons qu'il est indépendant des axiomes de Zermelo-Fraenkel. Très souvent, cet axiome est utilisé dans une de ses formes équivalentes :
Principe du bon ordre : Tout ensemble admet une structure d'ensemble bien ordonné.
Lemme de Zorn : Tout ensemble ordonné non vide dans lequel toute partie totallement ordonnée est majorée admet un plus grand élément.

Proposition 15. *On a équivalence entre :*

1. *Lemme de Zorn.*
2. *Principe du bon ordre.*
3. *Axiome du choix.*

Démonstration. Supposons 1. Soit X un ensemble. On définit Y l'ensemble des couples $(S, \leq)$ où $S \subset X$ et $\leq$ une relation de bon ordre sur S. De même Y est clairement non vide. On munit Y de la relation d'ordre partiel $(S, \leq) < (S', \leq')$

si $S \subset S'$, $\leq'$ prolonge $\leq$ et S est un segment initial de S' pour l'ordre induit. Soit $(S_i, \leq_i)_{i\in I}$ une partie totallement ordonnée de Y. On pose $S = \bigcup_{i\in I} S_i$ qui est une réunion filtrante et on munit S de la relation d'ordre $x \leq y$ si et seulement si $x \leq_i y$ où $x \in S_i$ et $y \in S_i$. On montre clairement que $\leq$ est un ordre total sur S (car S est une réunion filtrante). On a $S_i \subset S$ pour tout $i \in I$, $\leq$ prolonge $\leq_i$ pour tout $i \in I$. Si $x \in S_i$ et $y \in S$ tel que $y \leq x$, alors il existe d'une part j tel que $y \in S_j$. Soit $S_j \subset S_i$, dans ce cas $y \in S_i$, ou bien $S_i \subset S_j$ et donc $x \in S_j$ ainsi $y \leq_j x$ et comme S_i est segment initial de S_j donc $y \in S_i$. On en déduit que S_i est segment initial de S. Il reste à montrer que $\leq$ est une relation de bon ordre sur S. Soit A une partie non vide de S, soit $x \in A$, alors il existe $i \in I$ tel que $x \in A \cap S_i$. $A \cap S_i$ est alors partie non vide de S_i donc admet un plus petit élément, noté x_i. Soit x plus petit élément de A, comme $A \cap S_i \subset A$, alors $x \leq x_i$. Soit j tel que $x \in S_j$. Si $S_j \subset S_i$ alors $x \in A \cap S_i$ et donc $x_i \leq x$. Ou bien $S_i \subset S_j$ comme S_i est segment initial de S_j et $x \leq x_i$ donc $x \in S_i$, on en déduit $x_i \leq x$. Ainsi $(S, \leq)$ majore $(S_i, \leq_i)_{i\in I}$ dans Y. D'après le lemme de Zorn, Y a donc un plus grand élément pour $<$, noté $(S_0, \leq_0)$. Supposons $S_0 \neq X$. Alors il existe $a \in X - S_0$. Sur $S_0 \cup \{a\}$, on définit $\leq'$ par $b <' a$ pour tout $b \in S_0$. On montre facilement que $(S_0 \cup \{a\}, \leq')$ est un élément de Y. On a $S_0 \subset S_0 \cup \{a\}$, $\leq'$ prolonge $\leq$. On montre facilement que S_0 est segment initial de $S_0 \cup \{a\}$ ce qui est contradictoire avec la maximalité de S_0.
Supposons 2. Soit X ensemble non vide, il admet alors une relation de bon ordre. On définit alors la fonction de choix :

$$\begin{aligned} \varphi : \mathcal{P}(X) - \{\emptyset\} &\to X \\ A &\to \min(A) \end{aligned}$$

Supposons 3. Soit X un ensemble ordonné dont tout sous-ensemble totalement ordonné est majoré. Supposons que X n'admette pas de plus grand élément. Soit α l'ordinal de Hartogs de X et soit φ une fonction de choix pour X. On définit par récurrence transfinie l'application suivante de α dans X. $f(0) = \varphi(X)$, $f(\beta+1) = \varphi(R_{\beta+1})$, où $R_{\beta+1}$ est l'ensemble des majorants stricts de $f(\beta)$, celui-ci est non vide car X n'a pas de plus grand élément, et si β est un ordinal limite $f(\beta) = \varphi(R_\beta)$ où $R_\beta = \bigcup_{\gamma\in\beta} R_\gamma$. Par récurrence transfinie, on montre facilement que f est strictement croissante donc injective, ce qui est contradictoire avec la définition de l'ordinal de Hartogs. □

3.3.2 Cardinaux

Definition 3 (Cardinaux)**.** *Un cardinal est un ordinal qui n'est équipotent avec aucun de ses éléments.*

De la proposition suivante, on déduit qu'il existe des cardinaux.

Proposition 16. *Un ordinal fini est un cardinal.*

Démonstration. L'ensemble vide est clairement un cardinal.
Soit α ordinal fini et non vide. Il existe alors β tel que $\alpha = \beta + 1$. Supposons que α soit équipotent à un $\gamma \in \alpha$. Alors $\gamma \neq \beta$ car sinon β serait infini. On a alors $\gamma \in \beta \in \alpha$ et donc γ s'injecte dans β, on en déduit que $\alpha = \beta + 1$ s'injecte dans β ce qui est contradictoire avec la finitude de β. □

On a également :

Proposition 17. *L'ordinal ω (ensemble des ordinaux finis) est un cardinal.*

Démonstration. Soit $\beta \in \omega$. Si ω était équipotent à β via une bijection f, alors :

$$\begin{aligned} g : \omega + 1 &\to \beta + 1 \\ \omega &\to \beta \\ \alpha \in \omega &\to f(\alpha) \end{aligned}$$

est clairement une bijection entre $\omega + 1$ et $\beta + 1$. Comme ω et $\omega + 1$ sont équipotents, alors β et $\beta + 1$ sont équipotents, ce qui contredit la finitude de β. □

Sans l'axiome du choix, on ne peut pas en général associer à un ensemble un cardinal. Cependant, la proposition suivante nous permet de définir le cardinal d'un ordinal et plus généralement d'un ensemble bien ordonné.

Proposition 18 (Cardinal d'un ordinal)**.** *Un ordinal est équipotent à un unique cardinal.*

Démonstration. Existence : Soit α un ordinal. Soit A l'ensemble des éléments de α qui lui sont équipotents. Si $A = \emptyset$, alors α est un cardinal. Si $A \neq \emptyset$, soit β son plus petit élément. Si $\gamma \in \beta$, alors α n'est pas équipotent à γ et a fortiori β n'est pas équipotent à γ. Ainsi, β est un cardinal.
Unicité : Supposons que α soit équipotent aux cardinaux β_1 et β_2, alors par définition des cardinaux, $\beta_1 = \beta_2$. □

On notera désormais $|\alpha|$ le cardinal de α. Comme α est équipotent à $|\alpha|$, alors $|\alpha| \leq \alpha$. On remarquera qu'un ordinal n'est pas équipotent à un unique ordinal. La proposition précédente se généralise en :

Proposition 19 (Cardinal d'un ensemble bien ordonné)**.** *Un ensemble X est équipotent à un ordinal si et seulement si il admet un bon ordre.*

Démonstration. Se déduit facilement de la proposition précédente et du fait qu'un ensemble bien ordonné est isomorphe et donc équipotent à un ordinal. □

Remarque : Nous avons vu qu'un ensemble bien ordonné est isomorphe à un unique ordinal, cependant il n'est pas équipotent à un unique ordinal.

Proposition 20. *Soit α un ordinal. Nous avons l'équivalence :*

1. *α est infini.*

2. *Son cardinal $|\alpha|$ est infini.*
3. *Son cardinal $|\alpha|$ est un ordinal limite.*

Démonstration. On a trivialement 3. $\rightarrow$ 2. $\rightarrow$ 1.. Supposons 1., si on avait $|\alpha| = \gamma + 1$, alors comme $|\alpha|$ est un cardinal, on a γ n'est pas équipotent à $|\alpha| = \gamma + 1$, il s'ensuit que $|\alpha|$ est un ordinal fini. Ainsi $|\alpha| = \alpha$ est un ordinal fini, ce qui est contradictoire. Donc $|\alpha|$ est un ordinal limite. □

Parmi les ordinaux limites, nous avons les ordinaux de Hartogs d'un ordinal. On peut se demander si les ordinaux de Hartogs sont des cardinaux.

Proposition 21. *Soit X un ensemble quelquonque, alors son ordinal de Hartogs α_X est un cardinal (appelé cardinal de Hartogs).*

Démonstration. Supposons qu'il existe $\beta \in \alpha_X$ tel que β et α_X soient équipotents. Alors β ne s'injecte pas dans X, ce qui est contradictoire avec $\beta \in \alpha_X$. □

Ainsi, à n'importe quel ensemble X, on peut associer un cardinal de Hartogs. Cependant, ce cardinal de Hartogs n'est pas équipotent à X. On peut alors se demander s'il existe un cardinal équipotent à un ensemble quelquonque. Sans l'axiome du choix, la réponse est négative en général. Plus exactement :

Theorem 3 (Théorème de Zermelo)**.** *On a l'équivalence entre :*

1. *Tout ensemble est équipotent à un (unique) cardinal.*
2. *Pour tout couple d'ensembles X, Y, soit X s'injecte dans Y ou bien Y s'injecte dans X.*
3. *L'axiome du choix.*

Démonstration. 1. implique 2. : Soit $|X|$ et $|Y|$ les cardinaux respectifs de X et de Y. Il est trivial que soit $|X|$ s'injecte dans $|Y|$ ou bien $|Y|$ s'injecte dans $|X|$.
2. implique 3. : Soit X un ensemble et soit α son ordinal de Hartogs. Alors nécessairement, X s'injecte dans α. Soit f l'injection de X dans α. X est bien ordonné via la relation d'ordre $x < y$ si et seulement si $f(x) \in f(y)$.
3. implique 1. : Du fait que l'axiome du choix implique que tout ensemble admet un bon ordre. □

Nous avons vu que tous les ordinaux de Hartogs sont des cardinaux. En fait, nous allons voir qu'il y a deux types de cardinaux infinis : les cardinaux de Hartogs appelés cardinaux successeurs et les cardinaux limites (qui ne sont pas de Hartogs). De la proposition suivante, on montre qu'un ordinal de Hartogs est nécessairement l'ordinal de Hartogs d'un cardinal.

Proposition 22. *Soit α un ordinal et soit $|\alpha|$ son cardinal, alors $\alpha^+ = |\alpha|^+$.*

Démonstration. On a $|\alpha| \leq \alpha \in \alpha^+$. Comme α^+ est un cardinal, alors α^+ ne s'injecte pas dans $|\alpha|$, ainsi $|\alpha|^+ \leq \alpha^+$. Soit $\beta \in \alpha^+$, alors β s'injecte dans α et donc dans $|\alpha|$, ainsi $\beta \in |\alpha|^+$. □

Il existe aussi des cardinaux qui ne sont pas des ordinaux de Hartogs. En effet, nous avons :

Proposition 23. *Un cardinal κ est de Hartogs si et seulement si l'ensemble des éléments cardinaux de κ a un plus grand élément.*

Démonstration. Supposons κ cardinal de Hartogs, alors il existe un cardinal $|\alpha|$ tel que $\kappa = |\alpha|^+$. Montrons que $|\alpha|$ est le plus grand élément cardinal de κ. Soit $\beta \in \kappa$ un cardinal, si on avait $|\alpha| \in \beta$, alors β ne s'injecte pas dans $|\alpha|$, ce qui est contradictoire avec $\beta \in \kappa$.
Si κ a un cardinal maximal α. Montrons que $\kappa = \alpha^+$. κ ne s'injecte pas dans α, donc $\alpha \in \alpha^+ \leq \kappa$. Soit $\beta \in \kappa$, par maximalité de $|\alpha|$, on a $|\beta| \leq \alpha$, ainsi $|\beta|$ s'injecte dans α, il en est de même de β. Ainsi $\beta \in \alpha^+$. □

L'existence de cardinaux limites est alors garantie par la proposition suivante :

Proposition 24. *Soit A un ensemble de cardinaux, alors* $\sup A$ *est également un cardinal.*

Démonstration. Notons $\alpha = \sup A$. Soit $\beta \in \alpha$, montrons que α n'est pas équipotent à β. Il suffit de montrer que $|\alpha|$ n'est pas équipotent à $|\beta|$ et donc de montrer que $|\beta| \in |\alpha|$. Comme $\beta \in \alpha$, donc il existe $\kappa \in A$ tel que $\beta \in \kappa$. On en déduit que $|\beta| \in \kappa \leq |\alpha|$ et donc $|\beta| \in |\alpha|$ d'où le résultat. □

On définit par récurrence transfinie la suite des cardinaux Aleph :
- $\aleph_0 = \omega$.
- $\aleph_{\beta+1} = \aleph_\beta^+$.
- Si α est un ordinal limite, $\aleph_\alpha = \sup_{\beta\in\alpha} \aleph_\beta$.

On a alors :

Proposition 25. *Si α est un ordinal limite, alors $\aleph_\alpha$ est un cardinal limite (réciproque vraie).*

Démonstration. Supposons que $\aleph_\alpha$ possède un plus grand élément cardinal κ. Ainsi il existe $\beta \in \alpha$ tel que $\kappa \in \aleph_\beta$. Comme α est un ordinal limite, alors $\beta+1 \in \alpha$ et par construction $\aleph_\beta \in \aleph_{\beta+1}$, on en déduit $\aleph_\beta \in \aleph_\alpha$, ce qui contredit la maximalité de κ. □

On a également :

Proposition 26. *Pour tout ordinal α, $\alpha \leq \aleph_\alpha$.*

Démonstration. $0 \leq \omega$ est clair.
On procède par récurrence sur $\alpha + 1$. Soit $\beta \in \alpha + 1$ tel que pour tout $\gamma \in \beta$, $\gamma \leq \aleph_\gamma$, montrons $\beta \leq \aleph_\beta$. Si $\beta = \gamma + 1$, on a $\gamma \leq \aleph_\gamma$ donc $\gamma \in \aleph_{\gamma+1}$ ainsi $\gamma+1 \leq \aleph_{\gamma+1}$. Si β est un ordinal limite, soit $x \in \beta$, il existe $\gamma \in \beta$ tel que $x \in \gamma$, comme $\gamma \leq \aleph_\gamma$, ainsi $x \in \aleph_\gamma$ et donc $x \in \bigcup_{\gamma\in\beta} \aleph_\gamma = \aleph_\beta$. □

Finallement, la proposition suivante nous certifie que tous les cardinaux sont soit finis ou soit Aleph :

Proposition 27. *Tout cardinal infini est de la forme $\aleph_\alpha$ pour un unique ordinal α.*

Démonstration. Montrons l'existence. Soit κ un cardinal infini, alors $\aleph_0 \leq \kappa$. Comme $\kappa \leq \aleph_\kappa$, alors il existe un plus petit α tel que $\kappa \leq \aleph_\alpha$. Si $\alpha = 0$, alors $\kappa = \aleph_0$. Si $\alpha = \beta + 1$, alors $\kappa \leq \aleph_{\beta+1} = \aleph_\beta^+$. Supposons $\kappa \in \aleph_\beta^+$, donc κ s'injecte dans $\aleph_\beta$ et par conséquent $\kappa \leq \aleph_\beta$, ce qui contredit la minimalité de $\beta + 1$. Si α est un ordinal limite, alors pour tout $\gamma \in \alpha$, $\aleph_\gamma \in \kappa$, on en déduit $\aleph_\alpha \leq \kappa$. □

3.3.3 Arithmétique des cardinaux

Proposition 28. *Soit κ et λ deux ordinaux. Les ensembles $\kappa \coprod \lambda = (\kappa \times \{0\}) \cup (\kappa \times \{1\})$ et $\kappa \times \lambda$ sont bien ordonnables.*

Démonstration. Sur $\kappa \coprod \lambda$, on définit la relation d'ordre $(\alpha, 0) < (\beta, 1)$ pour tout α, β, $(\beta_1, 0) < (\beta_2, 0)$ si et seulement si $\beta_1 < \beta_2$ et $(\beta_1, 1) < (\beta_2, 1)$ si et seulement si $\beta_1 < \beta_2$. On vérifie que c'est bien une relation de bon ordre, appelée "ordre de concaténation".
Sur $\kappa \times \lambda$, on définit la relation d'ordre $(\alpha_1, \beta_1) < (\alpha_2, \beta_2)$ si et seulement si $\max(\alpha_1, \beta_1) < \max(\alpha_2, \beta_2)$ ou $(\max(\alpha_1, \beta_1) = \max(\alpha_2, \beta_2)$ et $\alpha_1 < \alpha_2)$ ou $(\max(\alpha_1, \beta_1) = \max(\alpha_2, \beta_2)$ et $\alpha_1 = \alpha_2$ et $\beta_1 < \beta_2)$. C'est une relation d'ordre total, appelée "ordre lexicographique". Soit A partie non vide de $\kappa \times \lambda$. On pose :

$$\begin{aligned} \gamma_0 &= \min\left[\max(\alpha, \beta) : (\alpha, \beta) \in A\right]. \\ \alpha_0 &= \min\left[\alpha : \exists \beta, \gamma_0 = (\alpha, \beta), (\alpha, \beta) \in A\right]. \\ \beta_0 &= \min\left[\beta : \gamma_0 = (\alpha_0, \beta), (\alpha_0, \beta_0) \in A\right]. \end{aligned}$$

On vérifie que $(\alpha_0, \beta_0) = \min A$ et donc $\kappa \times \lambda$ est bien ordonnable. □

$\kappa \coprod \lambda$ et $\kappa \times \lambda$ sont alors équipotents à des cardinaux (ils sont même isomorphes à des uniques ordinaux).

Definition 4 (Addition et multiplication des cardinaux)**.** *Soient κ et λ deux ordinaux. On définit la somme et le produit cardinal par :*

$$\begin{aligned} |\kappa| + |\lambda| &= \left|\kappa \coprod \lambda\right| \\ |\kappa|.|\lambda| &= |\kappa \times \lambda| \end{aligned}$$

Remarque : A la différence avec les opérations ordinales, les opérations cardinales sont commutatives.
On notera $\kappa + \lambda$ (resp. $\kappa.\lambda$) lorsque l'on désignera la somme (resp. produit) ordinale et $|\kappa| + |\lambda|$ (resp. $|\kappa|.|\lambda|$) lorsque l'on désignera la somme (resp. produit) cardinale. En générale ils ne coïncident pas. Par contre, nous avons :

Proposition 29. *Soit α et β deux ordinaux, alors :*

$$|\alpha + \beta| = |\alpha| + |\beta|$$
$$|\alpha.\beta| = |\alpha|.|\beta|$$

Démonstration. On montre que :

$$\begin{aligned} f : \alpha + \beta &\rightarrow \alpha \coprod \beta \\ \gamma \in \alpha &\rightarrow (\gamma, 0) \\ \alpha + \gamma &\rightarrow (\gamma, 1) \end{aligned}$$

est une bijection. Ainsi $|\alpha| + |\beta| = |\alpha + \beta|$.
On montre que :

$$\begin{aligned} f : \alpha \times \beta &\rightarrow \alpha.\beta \\ (\gamma, \delta) &\rightarrow \alpha.\delta + \gamma \end{aligned}$$

est une bijection. □

On en déduit alors :

Proposition 30. *Pour les cardinaux finis, la somme et le produit cardinaux coïncident avec la somme et le produit ordinaux.*

Démonstration. Si α et β sont des ordinaux finis, alors $\alpha + \beta$ et $\alpha.\beta$ sont finis (se montre par récurrence). Ce sont donc des cardinaux, et plus exactement $|\alpha + \beta| = \alpha + \beta$ et $|\alpha.\beta| = \alpha.\beta$. □

Remarque : La proposition précédente permet d'interpréter la somme (resp. le produit) de nombres entiers aussi bien comme une somme (resp. produit) ordinale que cardinale.
On a vu que si α est un ordinal infini, alors α est équipotent à $\alpha + 1$. La proposition suivante généralise ce fait :

Proposition 31. *Soient κ et λ deux cardinaux dont au moins un est infini, alors $|\kappa| + |\lambda| = |\kappa|.|\lambda| = \max(|\kappa|, |\lambda|)$.*

Démonstration. Pour tout couple d'ordinaux α, β on munit $\alpha \times \beta$ de la relation d'ordre lexicographique. Montrons que $\max(|\kappa|, |\lambda|) = |\kappa|.|\lambda|$ par récurrence transfinie sur $\max(|\kappa|, |\lambda|)$. On suppose que $\kappa \geq \lambda$ et donc κ est infini. Soit $S \subset \kappa \times \kappa$ un segment initial propre pour l'ordre lexicographique. Il existe $(\alpha, \beta) \in \kappa \times \kappa$ tel que $S = S_{(\alpha,\beta)}$. On montre que $S \subset (\alpha + 1) \times (\beta + 1)$ et $\alpha, \beta \in \kappa$. Comme un cardinal infini est un ordinal limite, alors : $\alpha + 1, \beta + 1 \in \kappa$. Si α et β sont finis, alors $|S| \leq |\alpha + 1|.|\beta + 1| = (\alpha + 1).(\beta + 1) \in \omega \leq \kappa$. Sinon, soit $\mu = \max(|\alpha + 1|, |\beta + 1|)$. Alors $\mu \leq \max(\alpha + 1, \beta + 1) \in \kappa$ et par hypothèse de récurrence, $\mu = |\alpha + 1|.|\beta + 1|$. On en déduit que le cardinal d'un segment propre de $\kappa \times \kappa$ est strictement plus petit que κ.
Tout segment propre de $\kappa \times \kappa$ s'injectant strictement dans κ, $\kappa \times \kappa$ est alors

isomorphe à un segment initial de κ. Ainsi $|\kappa|.|\kappa| \leq |\kappa|$. Or $|\kappa| = |1|.|\kappa| \leq |\kappa|.|\kappa|$ d'où l'égalité.
Ainsi :

$$\kappa = |\kappa|.|1| \leq |\kappa|.|\lambda| \leq |\kappa|.|\kappa| = \kappa$$
$$\kappa = |\kappa| + |0| \leq |\kappa| + |\lambda| \leq |\kappa| + |\kappa| = |\kappa|.|2| = \max(\kappa, 2) = \kappa.$$

□

Outre l'addition et la multiplication des cardinaux, on définit également la puissance. Pour cela, on suppose que l'axiome du choix est vrai.

Definition 5. *Soient κ et λ deux cardinaux. Sous l'hypothèse que l'axiome du choix est vrai, on définit $|\kappa|^{|\lambda|}$ comme le cardinal des fonctions de λ dans κ.*

Sous l'axiome du choix, on montre qu'il existe $\alpha > 0$ tel que $2^{\aleph_0} = \aleph_\alpha$. $2^{\aleph_0}$ est le cardinal de l'ensemble des parties de ω. Le fait que $\alpha > 0$ est la conséquence de la proposition suivante :

Proposition 32. *Soit X un ensemble et $\mathcal{P}(X)$ l'ensemble de ses parties. Alors X et $\mathcal{P}(X)$ ne sont pas équipotents.*

Démonstration. Supposons qu'il existe f bijection de X dans $\mathcal{P}(X)$. Soit $A = \{x \in X : x \notin f(x)\}$. Soit alors x_0 antécédent de A par f. Soit $x_0 \in A$, d'où $x_0 \notin f(x_0)$, c'est à dire $x_0 \notin A$. Soit $x_0 \notin A$ et donc $x_0 \in f(x_0)$, ainsi $x_0 \in A$, ce qui est contradictoire. □

De la même façon que pour l'addition et la mutliplication, la puissance des cardinaux coïncide avec la puissance des ordinaux pour les cardinaux finis :

Proposition 33. *Soient κ et λ deux cardinaux finis. Alors $|\kappa|^{|\lambda|} = \kappa^\lambda$.*

Démonstration. Appelons Y l'ensemble des applications de λ dans κ. On montre (modulo un peu d'arithmétique dans $\mathbb{N}$) que :

$$\begin{aligned} F : Y &\to \kappa^\lambda \\ f &\to f(0).\kappa^{\lambda-1} + f(1).\kappa^{\lambda-2} + \ldots + f(\lambda-2).\kappa + f(\lambda-1), \end{aligned}$$

est une bijection. □

On a également les propriétés suivantes :

Proposition 34. *Soient κ, λ et μ trois cardinaux.*

1. $\left(|\kappa|^{|\lambda|}\right)^{|\mu|} = |\kappa|^{|\lambda|.|\mu|}$.
2. $|\kappa|^{|\lambda|}.|\kappa|^{|\mu|} = |\kappa|^{|\lambda|+|\mu|}$.
3. *Si $0 < n < \omega$ et κ infini, alors $|\kappa|^n = |\kappa|$.*
4. *Si κ est infini, $2^{|\kappa|} = |\kappa|^{|\kappa|}$.*

Démonstration. 1. A toute application f de $\lambda \times \mu$ dans κ, on fait correspondre bijectivement l'application g qui à $\beta \in \mu$ associe l'application qui à $\alpha \in \lambda$ associe $f(\alpha, \beta)$.
2. A toute application f de $\lambda \coprod \mu$ dans κ, on fait correspondre bijectivement le couple d'application (f_1, f_2) tel que si $\alpha \in \lambda$, $f_1(\alpha) = f(\alpha, 0)$ et tel que si $\beta \in \mu$, $f_2(\beta) = f(\beta, 1)$.
3. On montre par récurrence sur $n > 0$. Supposons que pour tout $0 < m \leq n$, on a $|\kappa|^m = |\kappa|$, alors $|\kappa|^{n+1} = |\kappa|^n.|\kappa| = |\kappa|.|\kappa| = |\kappa|^2$, d'où le résultat.
4. A toute application f de κ dans 2, il est facile d'associer injectivement par injection canonique une application de κ dans κ. Ainsi $2^{|\kappa|} \leq |\kappa|^{|\kappa|}$. κ étant infini ainsi $\omega \leq \kappa$. A f application de κ dans κ, on associe injectivement (mais non forcément surjectivement) l'application $g = Ff$ de κ dans 2 définie par :

$$\begin{aligned} g : \kappa &\to 2 \\ f(0) + f(1) + \ldots + f(n) + n &\to 1 \text{ où } n \in \omega \\ \alpha \neq f(0) + f(1) + \ldots + f(n) + n &\to 0 \end{aligned}$$

ainsi $|\kappa|^{|\kappa|} \leq 2^{|\kappa|}$. □

3.3.4 Somme et produit infinis de cardinaux

Definition 6 (Réunion disjointe et produit d'ensembles)**.** *Soit $(X_i)_{i \in I}$ une famille d'ensembles. La réunion disjointe est définie par :*

$$\coprod_{i \in I} X_i = \bigcup_{i \in I} X_i \times \{i\},$$

et le produit cartésien $\prod_{i \in I} X_i$ par l'ensemble des fonctions f de I dans $\bigcup_{i \in I} X_i$ telles que $f(i) \in X_i$.

Sous l'axiome du choix, le produit infini d'ensembles non vides est non vide.

Definition 7 (Somme et produit infinis de cardinaux)**.** *Soit α un ordinal. Pour tout $\beta \in \alpha$, soit κ_β un cardinal.*
Sous ZF, $\coprod_{\beta \in \alpha} \kappa_\beta$ est bien ordonnable via $(\gamma, \beta) < (\tilde{\gamma}, \tilde{\beta})$ si $\beta \in \tilde{\beta}$ et $(\gamma, \beta) < (\tilde{\gamma}, \beta)$ si et seulement si $\gamma \in \tilde{\gamma}$. Ainsi la somme infinie de cardinaux est définie par :

$$\sum_{\beta \in \alpha} |\kappa_\beta| = \left| \coprod_{\beta \in \alpha} \kappa_\beta \right|.$$

Sous ZFC, $\prod_{\beta \in \alpha} \kappa_\beta$ est bien ordonnable et le produit infini de cardinaux est alors défini par :

$$\prod_{\beta \in \alpha} |\kappa_\beta| = \left| \prod_{\beta \in \alpha} \kappa_\beta \right|.$$

3.3.5 Cardinaux réguliers et cofinalité

Definition 8 (Cofinalité d'un ordinal limite)**.** *Soit α un ordinal limite, sa cofinalité est définie par :*

$$cf(\alpha) = \min\{|A| : A \subset \alpha, \sup(A) = \alpha\}.$$

Un cardinal infini κ est régulier si et seulement si $\kappa = cf(\kappa)$, autrement il est dit singulier.

On a :

Proposition 35. *Le cardinal $\aleph_0$ est régulier.*

Démonstration. En effet, toute partie de ω est soit finie ou dénombrable. □

Nous avons également les propriétés suivantes :

Proposition 36. *Soit α un ordinal limite.*

1. *$cf(\alpha)$ est le plus petit ordinal γ tel qu'il existe fonction $f : \gamma \to \alpha$ dont l'image n'est pas strictement majorée.*
2. *$cf(\alpha)$ est un cardinal régulier.*
3. *Un cardinal κ est régulier si et seulement si pour tout $\lambda \in \kappa$ et toute famille $(X_\alpha)_{\alpha\in\lambda}$ d'ensembles tels que $|X_\alpha| < \kappa$ pour tout $\alpha \in \lambda$, on a $\left|\bigcup_{\alpha\in\lambda} X_\alpha\right| < \kappa$.*
4. *Soit κ un cardinal, $cf(\kappa)$ est le plus petit ordinal λ tel que κ soit la réunion de λ ensembles de cardinal strictement inférieur à κ.*
5. *Tout cardinal infini de Hartogs est régulier.*
6. *Si α est un ordinal limite, alors $cf(\alpha) = cf(\aleph_\alpha)$.*

Démonstration. 1. Soit $A(\alpha)$ le plus petit ordinal γ tel qu'il existe une fonction $f : \gamma \to \alpha$ dont l'image n'est pas strictement majorée. Soit A_0 partie de α telle que $cf(\alpha) = |A_0|$ et $\sup A_0 = \alpha$. La bijection f de $cf(\alpha)$ dans A_0 induit une application de $cf(\alpha)$ dans α dont l'image n'est pas strictement majorée. On en déduit que $A(\alpha)$ existe et $A(\alpha) \leq cf(\alpha)$.
Montrons l'inégalité réciproque. Soit f de $A(\alpha)$ dans α dont l'image n'est pas strictement majorée. On construit par récurrence transfinie l'application g de $A(\alpha)$ dans $\alpha + 1$ par :

- $g(0) = f(0)$.
- Soit $\beta \in A(\alpha)$, s'il existe δ tel que pour tout $\gamma \in \beta$, on a $g(\gamma) \in f(\delta)$, alors :
$$g(\beta) = f\left(\min\left[\delta : \forall\gamma \in \beta, f(\delta) > g(\gamma)\right]\right).$$
Sinon, $g(\beta) = \alpha$.

La fonction g est croissante. S'il existe $\beta \in A(\alpha)$ tel que $g(\beta) = \alpha$, alors soit θ le plus petit élément de $A(\alpha)$ tel que $g(\theta) = \alpha$. La restriction $g_{|\theta} : \theta \to \alpha + 1$ est strictement croissante et son image n'est pas strictement majorée. Ainsi $A(\alpha) \leq \theta$ ce qui est contradictoire avec $\theta \in A(\alpha)$. On en déduit que pour tout β, $g(\beta) \neq \alpha$, ainsi $\sup A(\alpha) = \alpha$ et donc $cf(\alpha) \leq |A(\alpha)| \leq A(\alpha)$ d'où le résultat.
2. On a $cf(cf(\alpha)) \leq cf(\alpha)$. Montrons l'inégalité réciproque. Il existe une fonction f de $cf(cf(\alpha))$ dans $cf(\alpha)$ dont l'image n'est pas strictement majorée dans $cf(\alpha)$ et une fonction g de $cf(\alpha)$ dans α dont l'image n'est pas strictement majorée dans α. On montre facilement que l'image de $g \circ f$ n'est pas strictement majorée dans α et donc $cf(\alpha) \leq cf(cf(\alpha))$.
3. On rappelle qu'un cardinal κ est régulier si pour toute partie $A \subset \kappa$ telle que $|A| < \kappa$, on a $\sup A \in \kappa$. Supposons qu'un cardinal κ vérifie la propriété 3. Soit $A \subset \kappa$ de cardinal $\lambda < \kappa$. Soit $f : \lambda \to A$ une bijection. Pour tout $\alpha \in \lambda$, on pose :

$$X_\alpha = \{\xi \in A : \xi \in f(\alpha)\}.$$

Soit $\xi \in X_\alpha$ et $\beta \in \xi$, alors ξ est une partie de $f(\alpha)$ donc de A et comme $\beta \in \xi$, ainsi $\beta \in A$ et donc X_α est un segment initial (strict) de κ (et donc un ordinal), ainsi $|X_\alpha| < \kappa$. On en déduit alors que $|\cup_{\alpha\in\lambda} X_\alpha| < \kappa$. Puisque $\cup_{\alpha\in\lambda} X_\alpha$ est aussi un segment initial de κ. Soit $\cup_{\alpha\in\lambda} X_\alpha = \kappa$, ce qui est contradictoire avec $|\cup_{\alpha\in\lambda} X_\alpha| < \kappa$, ainsi $\cup_{\alpha\in\lambda} X_\alpha \in \kappa$. On a alors $\sup(\cup_{\alpha\in\lambda} X_\alpha) \in \kappa$ ($\cup_{\alpha\in\lambda} X_\alpha$ étant segment initial strict). Soit $\xi \in A$, si $\xi = 0$, alors considérons B l'ensemble des éléments non nuls de A. Si $B = \emptyset$, on a $\sup A = 0 \in \kappa$. Si $B \neq \emptyset$, soit η le plus petit élément de B et $\alpha \in \lambda$ tel que $\eta = f(\alpha)$, alors on a $\xi = 0 \in X_\alpha$. Si $\xi \neq 0$, on considère B l'ensemble des éléments de A strictement plus grand. Si $B = \emptyset$, alors tous les éléments de A sont égaux à ξ ou appartiennent à ξ, on en déduit que $\sup A = \xi \in \kappa$. Si $B \neq \emptyset$, soit η son plus petit élément et α tel que $\eta = f(\alpha)$, on en déduit que $x_i \in X_\alpha$. Ainsi $A \subset \cup_{\alpha\in\lambda} X_\alpha$ et donc $\sup A \in \kappa$.
Réciproquement, supposons que κ soit un cardinal régulier. Soit un cardinal $\lambda < \kappa$ et une famille $(X_\alpha)_{\alpha\in\lambda}$ une famille d'ensembles telle que $|X_\alpha| < \kappa$ pour tout $\alpha \in \lambda$. On construit par récurrence transfinie une suite de sous-ensembles deux à deux disjoints de κ telle que $|Y_\alpha| = |X_\alpha|$. (Prendre $Y_0 = |X_0|$, $Y_1 = (|X_0| + |X_1|)\backslash|X_0|$, etc...). On considère la fonction f de λ dans κ définie par $f(\alpha) = \sup Y_\alpha$. Comme $\lambda \in \kappa$ et κ est régulier, alors l'image de f a un majorant dans κ, ainsi il existe $\beta \in \kappa$ tel que :

$$\gamma \in \bigcup_{\alpha\in\lambda} Y_\alpha \Rightarrow \gamma \leq \beta.$$

En particulier, $|\cup Y_\alpha| \leq |\beta| \in \kappa$ et donc :

$$\left|\bigcup_{\alpha\in\lambda} X_\alpha\right| \leq \left|\bigcup_{\alpha\in\lambda} Y_\alpha\right| < \kappa.$$

4. Soit λ un ordinal tel que κ soit la réunion de λ ensembles $(X_\xi)_{\xi\in\lambda}$ de cardinal strictement inférieur à κ. Pour tout $\xi \in \lambda$, on peut poser $f(\xi) = \sup X_\xi$. On montre alors que $f(\xi) \in \kappa$. L'image de λ par f ne peut être strictement majorée

dans κ et donc $\lambda \geq cf(\kappa)$.
Soit $\gamma = cf(\kappa)$ et $f : \gamma \to \kappa$ d'image non strictement majorée dans κ. On pose pour $\alpha \in \gamma$, $X_\alpha = \{\xi \in \kappa : \xi \leq f(\alpha)\}$. X_α est un segment initial strict du cardinal κ et donc $|X_\alpha| < \kappa$. Le fait que l'image ne soit pas strictement majorée impose que $\kappa = \cup X_\alpha$.
5. Supposons κ cardinal de Hartogs, il existe alors un cardinal λ tel que $\kappa = \lambda^+$. Soit $\gamma = cf(\kappa)$, il existe une famille d'ensembles $(X_\xi)_{\xi\in\gamma}$ de cardinal inférieur strictement à κ (donc inférieur ou égal à λ) tel que :

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

On en déduit que :

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

On en déduit $\kappa \leq \gamma = cf(\kappa)$. □