1. Axiomes de ZF.

La version la plus courante de la théorie des ensembles est celle qui fut proposée en 1908 par Ernst Zermelo, à laquelle on rajoute un axiome (Schéma Axiome de Remplacement) dû à Abraham Fraenkel et Thoralf Skolem. Cette théorie est désignée par le sigle ZF sans l'Axiome du Choix, par ZFC si on y adjoint l'Axiome du Choix.

Le langage de la théorie des ensembles comporte (outre l'égalité) deux symboles : une relation binaire notée $\in$, et un symbole de constante (parfois omis) noté $\varnothing$. La relation $x \in y$ se lit "x est un élément de y", ou bien "x est dans y". Le symbole $\varnothing$ s'appelle l'ensemble vide.

Il est habituel d'utiliser des "quantificateurs restreints" : $\forall t \in x(P(t))$ signifie $\forall t((t \in x) \Longrightarrow (P(t)))$, et $\exists t \in x(P(t))$ signifie $\exists t((t \in x) \wedge (P(t)))$.

Les axiomes sont :

Axiome d'Extension. *Si x et y ont les mêmes éléments alors $x = y$.*

$$\forall x \forall y((\forall z(z \in x \Longleftrightarrow z \in y)) \Longrightarrow (x = y))$$

Discussion : On notera $x \subseteq y$ (ou $y \supseteq x$) pour dire que tous les éléments de x sont dans y, i.e. $\forall z(z \in x \Longrightarrow z \in y)$: on dit alors que x est une partie de y, ou que x est un sous-ensemble de y. ou que x est inclus dans y. L'Axiome d'Extension s'écrit alors

$$\forall x \forall y(((x \subseteq y) \wedge (x \supseteq y)) \Longrightarrow (x = y))$$

Axiome du Vide. *L'ensemble vide n'a aucun élément.*

$$\forall x(\neg x \in \varnothing)$$

Axiome des Paires. *Donnés deux ensembles x et y il existe un ensemble ayant x et y pour seuls éléments.*

$$\forall x \forall y \exists z \forall t((t \in z) \Longleftrightarrow ((t = x) \vee (t = y)))$$

Discussion : Cet ensemble (unique par Extension) sera not'e $\{x, y\}$. Dans le cas où $x = y$, on note $\{x\}$. On notera de plus $(x, y) = \{\{x\}, \{x, y\}\}$ (et on dit qu'il s'agit du couple dont la première coordonnée est x et la seconde y). On vérifie aisément que $(x, y) = (x', y')$ ssi $x = x'$ et $y = y'$. On définit également $(x, y, z) = ((x, y), z)$, et plus généralement, par récurrence, $(x_1, \ldots, x_n) = ((x_1, \ldots, x_{n-1}), x_n)$.

Axiome de l'Union. *Donné un ensemble x il existe un ensemble y dont les éléments sont exactement les éléments des éléments de x.*

$$\forall x \exists y \forall z((z \in y) \Longleftrightarrow (\exists t \in x(z \in t)))$$

Discussion : Cet ensemble (unique par Extension) se note $\bigcup x$ ou bien $\bigcup_{t \in x} t$. On dira qu'il s'agit de l'union des éléments de x. Si $x = \{u, v\}$, on notera $u \cup v$, et on dit qu'il s'agit de l'union de u et de v.

Schéma Axiome de Sélection. *Soit $\varphi(t, p)$ un prédicat ayant (au plus) deux variables libres. Pour tous ensembles p et x il existe un y dont les éléments sont exactement les éléments t de x vérifiant $\varphi(t, p)$.*

$$\forall p \forall x \exists y \forall t((t \in y) \Longleftrightarrow ((t \in x) \wedge (\varphi(t, p))))$$

Discussion : On notera qu'il s'agit là d'une infinité dénombrable d'axiomes, un pour chaque prédicat φ. L'ensemble (unique par Extension) ainsi défini se note $\{t \in x : \varphi(t, p)\}$. On trouve souvent le Schéma Axiome de Sélection énoncé avec plusieurs paramètres $p_1, \ldots, p_n$ au lieu d'un. Ces deux formes sont équivalentes car

$$\{t \in x : \varphi(t, p_1, \ldots, p_n)\} = \{t \in x : \hat{\varphi}(t, (p_1, \ldots, p_n))\}$$

où $\hat{\varphi}(t, p)$ vaut par définition :

$$\exists p_1 \cdots \exists p_n((p = (p_1, \ldots, p_n)) \wedge \varphi(t, p_1, \ldots, p_n))$$

Si x est un ensemble *non vide*, u un élément de x, on définit l'intersection des éléments de x, notée $\bigcap x$ ou $\bigcap_{t \in x} t$ par :

$$\{z \in u : \forall t \in x(z \in t)\}$$

Cet ensemble est unique par Extension, et ne dépend pas du choix de u. Lorsque $x = \{u, v\}$, on note $u \cap v$. Lorsque $u \cap v = \varnothing$, on dit que u et v sont disjoints.

Axiome des Parties. *Donné un ensemble x il existe un ensemble y dont les éléments sont les parties de x.*

$$\forall x \exists y \forall t((t \in y) \iff (t \subseteq x))$$

Discussion : Cet ensemble (unique par Extension) se note $\mathcal{P}(x)$ et s'appelle ensemble des parties de x.

Axiome de l'Infini. *Il existe un ensemble contenant $\varnothing$ et stable par $t \mapsto t \cup \{t\}$.*

$$\exists x((\varnothing \in x) \wedge (\forall t \in x(t \cup \{t\} \in x)))$$

Discussion : Un tel ensemble s'appelle inductif. On vérifie immédiatement qu'une intersection quelconque d'ensembles inductifs est inductive. Il existe donc un plus petit ensemble inductif. On le notera ω : ainsi, ω est inductif, et si $I \subseteq \omega$ est inductif alors $I = \omega$ ("postulat de récurrence"). L'ensemble ω se note aussi $\mathbb{N}$, et ses éléments s'appellent les entiers naturels ou les ordinaux finis.

Schéma Axiome de Remplacement. *Soit $\psi(u, v, p)$ un prédicat ayant (au plus) trois variables libres. Pour tout ensemble p, si $\psi(u, v, p)$ est fonctionnel*[*] *en v alors pour tout ensemble x il existe un y dont les éléments sont exactement les images des éléments de x.*

$$\begin{aligned}&\forall p((\forall u \forall v \forall w(((\psi(u, v, p)) \wedge (\psi(u, w, p))) \Longrightarrow (v = w)))\\ &\Longrightarrow (\forall x \exists y \forall v((v \in y) \iff (\exists u \in x(\psi(u, v, p))))))\end{aligned}$$

Discussion : Comme pour le Schéma Axiome de Sélection, on peut en déduire une version à plusieurs paramètres. On remarquera que le Schéma Axiome de Sélection découle du Schéma Axiome de Remplacement : il suffit pour le voir de considérer le prédicat $\psi(u, v, p)$ donné par $(u = v) \wedge (\varphi(u, p))$, lequel est clairement fonctionnel en v.

Axiome de Régularité. *Tout ensemble x non vide a un élément y qui lui est disjoint.*

$$\forall x((x \neq \varnothing) \Longrightarrow (\exists y \in x(y \cap x = \varnothing)))$$

Discussion : Cet axiome n'est jamais utilisé en mathématique, mais il est très utile en *méta*mathématique. Il ne coûte rien de l'ajouter à notre liste, puisqu'on peut (aisément) montrer qu'il est consistant relatif aux autres. Entre autres conséquences de l'Axiome de Régularité, on ne peut pas avoir $t \in t$ (sans quoi l'ensemble $\{t\}$ ne vérifierait pas l'axiome), ni, plus généralement,

$$t_0 \ni t_1 \ni t_2 \ni \cdots$$

(considérer l'ensemble $\{t_0, t_1, \ldots\}$).

Le dernier axiome, l'Axiome du Choix, ne fait pas partie de la théorie ZF, mais de ZFC. On peut l'énoncer ainsi :

Axiome du Choix. *Si x est un ensemble dont les éléments sont non vides et deux à deux disjoints il existe un ensemble y dont l'intersection avec tout élément de x est réduite à un singleton.*

$$\forall x(((\forall t \in x(t \neq \varnothing)) \wedge (\forall u \in x \forall v \in x(u \cap v = \varnothing))) \Longrightarrow (\exists y \forall t \in x \exists z(t \cap y = \{z\})))$$

Discussion : *Cet axiome est un peu particulier, puisqu'il affirme l'existence d'un ensemble sans le définir, i.e. d'un ensemble qui n'est pas unique*[†]*. Il a donné lieu à bien des controverses au XIXème siècle (notamment de la part de l'école française) ; en tout cas, il est assez important de savoir ce qu'on peut et ce qu'on ne peut pas faire sans l'Axiome du Choix.*

Dans toute la suite, nous signalerons explicitement les résultats qui dépendent de l'Axiome du Choix.

---

[*] *Cela signifie que pour tout u il existe au plus un v (l'*image* de u) tel que $\psi(u, v, p)$.*

[†] *Il est vrai que c'était aussi le cas de l'Axiome de Régularité. D'ailleurs, ces deux axiomes ont quelques points communs.*

2. Bon ordre et ordinaux.

Nous supposons connues toutes les notions classiques de la théorie naïve des ensembles (relation, relation d'équivalence, relation d'ordre, fonctions, injections, surjections, bijections, etc.)

On dira qu'une relation d'ordre $\leq$ sur un ensemble W est une relation de bon ordre (ou que W est bien ordonné par $\leq$) ssi toute partie non vide de W admet un plus petit élément pour $\leq$. En considérant les parties à deux éléments, on voit aussitôt que $\leq$ est un ordre total.

Si T est un ensemble totalement ordonné, alors pour tout $a \in T$ on note $s(a) = \{t \in T : t < a\}$ (le segment initial de a).

Le théorème suivant, bien que trivial, est fondamental :

Théorème de l'induction transfinie. *Soit W un ensemble bien ordonné, et $I \subseteq W$ tel que pour tout $a \in W$, $s(a) \subseteq I \Longrightarrow a \in I$. Alors $I = W$.*

Démonstration : Dans le cas contraire, considérer le plus petit élément a de $W \setminus I$.

Il existe également un théorème permettant la définition par induction transfinie, que nous laissons le soin au lecteur de formuler et de démontrer.

Un ensemble T est dit transitif ssi $T \subseteq \mathcal{P}(T)$, soit $\forall x \in T (x \subseteq T)$. Il s'ensuit que la relation $\in$ restreinte à T est transitive.

Un ordinal est un ensemble transitif α tel que pour tous $x, y \in \alpha$ distincts on ait $x \in y$ ou bien $y \in x$.

Proposition. *Soit α un ordinal. Alors la relation $\in$ restreinte à α est une relation de bon ordre strict sur α.*

Démonstration : Elle est transitive par définition. D'après l'Axiome de Régularité, elle est irréflexive et antisymétrique. Il s'agit donc d'une relation d'ordre strict. Par définition, cet ordre est total. Soit enfin $A \subseteq \alpha$ non vide. D'après l'Axiome de Régularité, il existe $x \in A$ tel que $x \cap A = \varnothing$. Or ceci signifie précisément que x est minimal dans A ; comme l'ordre est total, x est le plus petit élément de A, et α est bien ordonné.

L'importance des ordinaux vient du théorème suivant :

Théorème. *Tout ensemble bien ordonné est isomorphe (en tant qu'ensemble ordonné) à un unique ordinal. De plus, l'isomorphisme en question est unique.*

(Ce théorème se montre par induction transfinie, et c'est facile, quoique assez pénible.)

Pour deux ordinaux α, β distincts, il revient au même de dire que $\alpha \in \beta$, que $\alpha \subset \beta$, ou que α est isomorphe à un segment initial de β. Lorsque l'une quelconque de ces affirmations est vérifiée, on dit que α est inférieur à β, et on note $\alpha < \beta$. Un ordinal n'est autre que l'ensemble des ordinaux strictement plus petits que lui.

Tout ensemble d'ordinaux admet un plus petit élément, qui n'est autre que l'intersection de ses éléments, une borne supérieure (le plus petit de ses majorants), qui n'est autre que l'union de ses éléments, et une borne supérieure stricte (le plus petit de ses majorants stricts). La borne supérieure stricte de $\{\alpha\}$, où α est un ordinal, est le plus petit ordinal strictement supérieur à α : on vérifie sans mal qu'il s'agit de $\alpha \cup \{\alpha\}$, on le note $\alpha + 1$, et on l'appelle successeur de α. Ainsi, la borne supérieure stricte d'un ensemble d'ordinaux n'est autre que la borne supérieure des successeurs des ordinaux en question. Un ordinal qui n'est pas le successeur d'un autre ordinal s'appelle un ordinal limite.

Le plus petit ordinal est $\varnothing$, que l'on note aussi 0. Son successeur est $\{\varnothing\}$, que l'on note aussi 1. Son successeur est $\{\varnothing, \{\varnothing\}\}$, que l'on note aussi 2. Et ainsi de suite. Le plus petit ordinal successeur autre que 0 est ω.

On définit des opérations sur les ordinaux par induction transfinie sur β : l'addition par :

$$\alpha + \beta = \begin{cases} \alpha \text{ si } \beta = 0 \\ (\alpha + \gamma) + 1 \text{ si } \beta = \gamma + 1 \\ \sup\limits_{\gamma < \beta} \alpha + \gamma \text{ si } \beta \text{ est limite} \end{cases}$$

la multiplication par :

$$\alpha\beta = \begin{cases} 0 \text{ si } \beta = 0 \\ (\alpha\gamma) + \alpha \text{ si } \beta = \gamma + 1 \\ \sup\limits_{\gamma<\beta} \alpha\gamma \text{ si } \beta \text{ est limite} \end{cases}$$

et l'exponentiation ordinale par :

$$\alpha^{[\beta]} = \begin{cases} 1 \text{ si } \beta = 0 \\ \alpha^{[\gamma]}\alpha \text{ si } \beta = \gamma + 1 \\ \sup\limits_{\gamma<\beta} \alpha^{[\gamma]} \text{ si } \beta \text{ est limite} \end{cases}$$

Ces opérations ne jouissent pas des propriétés sympathiques qu'on leur connaît sur les naturels. En particulier, l'addition et la multiplication ne sont pas commutatives en général. Certaines propriétés sont néanmoins valables, par exemple, l'associativité de l'addition et de la multiplication, la neutralité (à gauche et à droite) de 0 pour l'addition et de 1 pour la multiplication et la distributivité à gauche de la multiplication sur l'addition (i.e. $\alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma$).

3. Cardinaux et alephs.

Si A et B sont deux ensembles, on dit que A et B sont équipotents, et on note $|A| = |B|$ ss'il existe une bijection entre A et B. On dit que A est subpotent à B, et on note $|A| \leq |B|$ (ou $|B| \geq |A|$) ssi A est équipotent à une partie de B. (La notation $|A| \leq |B|$ est justifiée par le fait que ceci reste vrai si on remplace A, B par deux ensembles A', B' qui leur sont respectivement équipotents.) On a le théorème fondamental suivant :

Théorème de Cantor-Schröder-Bernstein. *Si $|A| \leq |B|$ et $|A| \geq |B|$ alors $|A| = |B|$.*

(Nous n'en donnerons pas la démonstration ici ; il faut cependant signaler qu'elle n'utilise *pas* l'Axiome du Choix.)

Il serait possible de donner une définition de $|A|$ pour un ensemble A quelconque, en utilisant l'Axiome de Régularité (à savoir, l'ensemble des ensembles équipotents à A et de rang minimal). Nous n'en aurons cependant pas besoin. Nous dirons que $|A|$ est le cardinal de A, sans trop préciser ce que c'est. La relation $\leq$ est une relation d'ordre entre les cardinaux (en l'absence de l'Axiome du Choix, elle n'a aucune raison d'être totale). Bien sûr, on notera $|A| < |B|$ (ou $|B| < |A|$) pour dire $(|A| \leq |B|) \wedge \neg(|A| = |B|)$. Le cardinal du vide sera noté 0.

Si A est l'image de B par une surjection, on notera $|A| \trianglelefteq |B|$, ou $|B| \trianglerighteq |A|$, et on dira que A est une projection de B.

Proposition. *Si $|A| \leq |B|$ avec $|A| > 0$ alors $|A| \trianglelefteq |B|$.*

> **Démonstration :** Soit $f: A \to B$ injective. Puisque $|A| > 0$, on peut trouver $x \in A$. On définit $g: B \to A$ par $g(y) = f^{-1}(y)$ si $y \in \operatorname{im} f$, et $g(y) = x$ sinon. g est clairement surjective, donc $|A| \trianglelefteq |B|$.

Un ordinal α tel que $|\beta| < |\alpha|$ pour tout $\beta < \alpha$ s'appelle un aleph. Par exemple, ω est un aleph, et en tant que tel, on le note $\aleph_0$. Il est d'usage d'identifier un aleph avec son cardinal, et on notera donc $|A| = \aleph_0$ pour dire $|A| = |\omega|$. Un tel ensemble A est dit dénombrable. Tout ordinal fini (élément de ω) est un aleph.

Proposition. *Si $|F| \triangleleft \aleph_0$ alors $|F| \leq \aleph_0$.*

> **Démonstration :** Soit $g: \omega \to F$ surjective. On définit $f: F \to \omega$ par :
>
> $$f(a) = \min\{n \in \omega : g(n) = a\}$$
>
> f est manifestement injective.

Un ensemble F tel que $|F| = n$ pour un certain $n \in \omega$ s'appelle <u>fini</u>. Un sous-ensemble d'un ensemble fini est fini, et de cardinal inférieur.

Proposition. *Pour tout ensemble F, les affirmations suivantes sont équivalentes :*
(i) $|F| = n$ *pour un certain* $n \in \omega$ *(i.e. F est fini).*
(ii) Toute partie non vide de $\mathcal{P}(F)$ admet un élément maximal pour l'inclusion.
(iii) $|F| < \aleph_0$.

Démonstration : (i)$\Longrightarrow$(ii) : Si $\mathcal{A} \subseteq \mathcal{P}(F)$, tous les éléments de $\mathcal{A}$ ont pour cardinal un certain $m \in \omega$, avec $m \leq n$. Considérer un élément de $\mathcal{A}$ ayant un cardinal maximal. (ii)$\Longrightarrow$(iii) : Considérer le $\mathcal{A} \subseteq \mathcal{P}(F)$ formé des $E \subseteq F$ tels que $|E| < \aleph_0$. On a $\varnothing \in \mathcal{A}$ donc $\mathcal{A}$ n'est pas vide. Soit $M \in \mathcal{A}$ maximal pour l'inclusion. Si $M \neq F$, alors il existe $x \in F \setminus M$, et on voit aisément que $M \cup \{x\} < \aleph_0$. Ceci contredit la maximalité de M. (iii)$\Longrightarrow$(i) : Puisque $|F| < \aleph_0$, on peut supposer $F \subseteq \omega$. On définit alors $f(0)$ comme le plus petit élément de F, $f(1)$ comme le plus petit élément de $F \setminus \{f(0)\}$ et ainsi de suite, jusqu'à ce que ce ne soit plus possible. Ceci définit soit $f\colon \omega \to F$ bijective, ce qui n'est pas possible, soit $f\colon n \to F$ bijective, et alors $|F| = n$.

Un ensemble F tel que tout $\mathcal{A} \subseteq \mathcal{P}(F)$ *totalement ordonné pour l'inclusion* admet un élément maximal s'appelle T-fini.

Proposition. *Pour tout ensemble F, les affirmations suivantes sont équivalentes :*
(i) Il n'existe pas de suite strictement croissante $A_0 \subsetneqq A_1 \subsetneqq \cdots$ *dans* $\mathcal{P}(F)$.
(i') Il n'existe pas de suite strictement décroissante $A_0 \supsetneqq A_1 \supsetneqq \cdots$ *dans* $\mathcal{P}(F)$.
(ii) Toute suite croissante $A_0 \subseteq A_1 \subseteq \cdots$ *dans* $\mathcal{P}(F)$ *est stationaire.*
(ii') Toute suite décroissante $A_0 \supseteq A_1 \supseteq \cdots$ *dans* $\mathcal{P}(F)$ *est stationaire.*
(iii) $|F| \not\trianglerighteq \aleph_0$.

Démonstration : (i)$\Longleftrightarrow$(i') et (ii)$\Longleftrightarrow$(ii') : Passer au complémentaire. (i)$\Longrightarrow$(ii) : Donnée une suite croissante non stationaire, en éliminant les répétitions, on obtient une suite strictement croissante. (ii)$\Longrightarrow$(i) : Trivial. (i)$\Longrightarrow$(iii) : Si $g\colon F \to \omega$ est surjective, alors en posant $A_n = g^{-1}(\{0, \ldots, n-1\})$, on obtient une suite strictement croissante dans $\mathcal{P}(F)$. (iii)$\Longrightarrow$(i) : Si (A_n) est une suite strictement croissante dans $\mathcal{P}(F)$, alors on pose $g(x) = n$ lorsque $x \in A_{n+1} \setminus A_n$, et $g(x) = 0$ si x n'est dans aucun des A_{n+1} : $g\colon F \to \omega$ ainsi définie est clairement surjective.

Un ensemble qui vérifie les conditions équivalentes de la proposition sera dit P-fini.

Proposition. *Pour tout ensemble F, les affirmations suivantes sont équivalentes :*
(i) F n'a pas de sous-ensemble dénombrable.
(ii) $|F| \not\geq \aleph_0$.

Démonstration : Trivial.

Un ensemble qui vérifie les conditions équivalentes de la proposition sera dit D-fini.

Proposition. *(a) Un ensemble fini est T-fini. (b) Un ensemble T-fini est P-fini. (c) Un ensemble P-fini est D-fini.*

Démonstration : (a) : Trivial (utiliser la caractérisation (iii) des ensembles finis). (b) : S'il existe une suite strictement croissante (A_n) dans $\mathcal{P}(F)$, alors $\{A_n : n \in \omega\}$ est une partie totalement ordonnée de $\mathcal{P}(F)$ sans élément maximal. (c) : On sait que si $|\omega| \leq |F|$ alors $|\omega| \trianglelefteq |F|$ (car $|\omega| > 0$).