# La théorie des ensembles.

On se place dans la logique du 1er ordre avec $\mathscr{L} = \{\in, =\}$. On se place dans un univers $\mathcal{U}$ non vide, le modèle, dont les éléments sont appelés des *ensembles*.

Il faudra faire la différence entre les ensembles « naïfs » (les ensembles habituels), et les ensembles « formels » (les éléments de $\mathcal{U}$).

On a le paradoxe de Russel. On peut l'écrire

> « On a un barbier qui rase tous les hommes qui ne se rasent pas eux-mêmes. Qui rase le barbier ? ».

Si $\mathcal{U}$ est l'ensemble de tous les ensembles, alors

$$a := \{\, x \in \mathcal{U} \mid x \notin x \,\}$$

vérifie $a \in a \iff a \notin a$, ***paradoxe***. Pour éviter ce paradoxe, on choisit donc de ne pas faire $\mathcal{U}$ un ensemble.

## 1 Les axiomes de la théorie de Zermelo-Fraenkel.

**ZF 1.** *Axiome d'extensionnalité* : deux ensembles sont égaux ssi ils ont les mêmes éléments

$$\forall x\, \forall y\, \Big(\forall z\, (z \in x \leftrightarrow z \in y) \leftrightarrow x = y\Big).$$

- *Axiome de la paire*[1] : il existe une paire $\{x, y\}$ pour tout élément $x$ et $y$

$$\forall x\, \forall y\, \exists z\, \forall t \Big(t \in z \leftrightarrow (t = x \vee t = y)\Big).$$

[*continué plus tard*...]

1. On verra plus tard que cet axiome est une conséquence des autres (de ZF 3 et ZF 4).

**Remarque 1.** Cela nous donne l'existence du *singleton* $\{x\}$ si $x$ est un ensemble. En effet, il suffit de faire la paire $\{x, x\}$ avec l'Axiome de la paire.

**Définition 1.** Si $a$ et $b$ sont des ensembles, alors $(a, b)$ est l'ensemble $\{\{a\}, \{a, b\}\}$. Ainsi, $(a, a)$ est l'ensemble $\{\{a\}\}$.

**Lemme 1.** Pour tous ensembles $a, b, a', b'$, on a $(a, b) = (a', b')$ ssi $a = a'$ et $b = b'$.

**Preuve.** En exercice. □

**Définition 2.** On peut construire des 3-uplets $(a_1, a_2, a_3)$ avec $(a_1, (a_2, a_3))$, et ainsi de suite pour les $n$-uplets.

**Notation.** On utilise les raccourcis
- ▷ $t = \{a\}$ pour $\forall x\ (x \in t \leftrightarrow x = a)$ ;
- ▷ $t = \{a, b\}$ pour $\forall x\ (x \in t \leftrightarrow (x = a \lor x = b))$ ;
- ▷ $t \subseteq a$ pour $\forall x\ (z \in t \to z \in a)$.

**ZF 3.** *Axiome des parties* : l'ensemble des parties $\wp(a)$ existe pour tout ensemble $a$

$$\forall a\, \exists b\, \forall t\ (t \in b \leftrightarrow t \subseteq a).$$

**ZF 2.** *Axiome de la réunion* : l'ensemble $y = \bigcup_{z \in x} z$ existe

$$\forall x\, \exists y\, \forall t(t \in y \leftrightarrow \exists z(t \in z \land z \in x)).$$

**Remarque 2.** Comment faire $a \cup b$ ? La paire $x = \{a, b\}$ existe par l'Axiome de la paire, et $\bigcup_{z \in x} z = a \cup b$ est un ensemble par ZF 2.

**ZF 4'.** *Schéma de compréhension* : pour toute formule $\varphi(y, v_1, \ldots, v_n)$, on a l'ensemble $x = \{\, y \in v_{n+1} \mid \varphi(y, v_1, \ldots, v_n) \,\}$

$$\forall v_1\ \ldots\ \forall v_n\, \exists x\, \forall y\, \Big(y \in x \leftrightarrow (y \in v_{n+1} \land \varphi(y, v_1, \ldots, v_n))\Big).$$

**Remarque 3.** Peut-on faire le paradoxe de Russel ? On ne peut pas faire $a := \{z \in \mathscr{U} \mid z \notin z\}$ car $\mathscr{U}$ n'est pas un ensemble ! Et, on ne peut pas avoir de paradoxe avec $b := \{z \in E \mid z \notin z\}$, car on a l'ajout de la condition $b \in E$.

**Définition 3.** Une *relation fonctionnelle* en $w_0$ est une formule $\varphi(w_1, w_2, a_1, \ldots, w_n)$ *à paramètres* (où les $a_i$ sont dans $\mathscr{U}$) telle que

$$\mathscr{U} \models \forall w_0\, \forall w_1\, \forall w_2\, \big(\varphi(w_0, w_1, a_1, \ldots, a_n) \wedge \varphi(w_0, w_2, a_1, \ldots, w_n) \to w_1 = w_2\big).$$

En termes naïfs, c'est une fonction partielle. On garde le terme *fonction* quand le domaine et la collection d'arrivée sont des *ensembles*, autrement dit, des éléments de $\mathscr{U}$.

**ZF 4.** *Schéma de substitution/de remplacement* : « la collection des images par une relation fonctionnelle des éléments d'un ensemble est aussi un ensemble ». Pour tout $n$-uplet $\bar{a}$, si la formule à paramètres $\varphi(w_0, w_1, \bar{a})$ définit une relation fonctionnelle $f_{\bar{a}}$ en $w_0$ et si $a_0$ est un ensemble alors la collection des images par $f_{\bar{a}}$ des éléments de $a_0$ est un ensemble nommé $a_{n+1}$

$$\begin{gathered}\forall a_0 \cdots \forall a_n \\ \big(\forall w_0\, \forall w_1\, \forall w_2 (\varphi(w_0, w_1, a_1, \ldots, a_n) \wedge \varphi(w_0, w_2, a_1, \ldots, a_n)) \to w_1 = w_2\big) \\ \downarrow \\ \exists a_{n+1}\, \forall a_{n+2} (a_{n+2} \in a_{n+1} \leftrightarrow \exists w_0\, w_0 \in a_0 \wedge \varphi(w_0, a_{n+2}, v_1, \ldots, v_n)).\end{gathered}$$

**Théorème 1.** Si ZF 1, ZF 2, ZF 3 et ZF 4 sont vrais dans $\mathscr{U}$, il existe (dans $\mathscr{U}$) un et un seul ensemble sans élément, que l'on notera $\emptyset$.

**Preuve.** ▷ *Unicité* par ZF 1.

▷ *Existence.* On procède par compréhension : l'univers $\mathscr{U}$ est non vide, donc a un élément $x$. On considère la formule $\varphi(w_0, w_1) := \bot$ qui est une relation fonctionnelle. Par ZF 4 (avec la formule $\varphi$ et l'ensemble $a_0 := x$) un ensemble $a_{n+1}$ qui est vide.

□

**Proposition 1.** Si ZF 1, ZF 2, ZF 3 et ZF 4 sont vrais dans $\mathscr{U}$, alors l'Axiome de la paire est vrai dans $\mathscr{U}$.

**Preuve.** On a $\emptyset$ dans $\mathscr{U}$ et également $\wp(\emptyset) = \{\emptyset\}$ et $\wp(\wp(\emptyset)) = \{\emptyset, \{\emptyset\}\}$ par ZF 3.

Étant donné deux ensemble $a$ et $b$, on veut montrer que $\{a, b\}$ est un ensemble avec ZF 4

$$\varphi(w_0, w_1, a, b) := (w_0 = \emptyset \wedge w_1 = a) \vee (w_0 = \{\emptyset\} \wedge w_1 = b),$$

où

▷ $w_0 = \emptyset$ est un raccourci pour $\forall z\, (z \notin w_0)$ ;

▷ $w_0 = \{\emptyset\}$ est un raccourci pour $\forall z\, (z \in w_0 \leftrightarrow (\forall t\; t \notin z))$.

Ces notations sont compatibles avec celles données précédemment.

Comme $\varphi$ est bien une relation fonctionnelle et $\{a, b\}$ est l'image de $\{\emptyset, \{\emptyset\}\}$. □

**Proposition 2.** Si ZF 1, ZF 2, ZF 3 et ZF 4 sont vrais dans $\mathscr{U}$, alors ZF 4' est vrai dans $\mathscr{U}$.

**Preuve.** On a la formule $\varphi(y, v_1, \ldots, v_n)$ et on veut montrer que

$$\mathscr{U} \models \forall v_1 \;\cdots\; \forall v_{n+1}\, \exists x\, \forall y\, \Big(y \in x \leftrightarrow (y \in v_{n+1} \wedge \varphi(y, v_1, v_n))\Big).$$

On considère la formule $\psi(w_0, w_1, \bar{v}) := w_0 = w_1 \wedge \varphi(w_0, \bar{v})$, qui est bien une relation fonctionnelle en $w_0$. La collection

$$\{\, y \in v_{n+1} \mid \varphi(y, v_1, \dots, v_n) \,\}$$

est l'image de $v_{n+1}$ par $\psi$ par ZF 4. $\square$

**Remarque 4.** La réciproque du théorème précédent est fausse ! Les axiomes ZF 4 et ZF 4' ne sont pas équivalents. On le verra en TD (probablement).

**Proposition 3.** Le produit ensembliste de deux ensembles est un ensemble.

**Preuve.** Soient $v_1$ et $v_2$ deux ensembles. On considère

$$X := v_1 \times v_2 = \{\, (x, y) \mid x \in v_1 \text{ et } y \in v_2 \,\} \text{ (en naïf)} .$$

La notation $(x, y)$ correspond à l'ensemble $\{\{x\}, \{x, y\}\} \in \wp(\wp(v_1 \cup v_2))$.

On applique ZF 4' dans l'ensemble ambiant $\wp(\wp(v_1 \cup v_2))$, on définit le produit comme la compréhension à l'aide de la formule

$$\varphi(z, v_1, v_2) := \exists x \, \exists y \left(z = \{\{x\}, \{x, y\}\} \wedge x \in v_1 \wedge y \in v_2\right).$$

C'est bien un élément de $\mathcal{U}$. $\square$

**Définition 4.** Une *fonction* (sous-entendu *totale*) d'un ensemble $a$ dans un ensemble $b$ est un sous-ensemble de $a \times b$ qui vérifie la propriété

$$\varphi(f, a, b) := \begin{pmatrix} f \subseteq a \times b \\ \wedge \\ \forall x\, \forall y\, \forall y'\, (x, y) \in f \wedge (x, y') \in f \to y = y' \\ \wedge \\ \forall x\, x \in a \to \exists y\, y \in b \wedge (x, y) \in f \end{pmatrix}.$$

On identifie ainsi $f$ et son graphe.

Une *fonction partielle* d'un ensemble $a$ dans un ensemble $b$ est un sous-ensemble de $a \times b$ qui vérifie la propriété

$$\varphi(f, a, b) := \begin{pmatrix} f \subseteq a \times b \\ \wedge \\ \forall x\, \forall y\, \forall y'\, (x, y) \in f \wedge (x, y') \in f \to y = y' \end{pmatrix}.$$

On note $b^a$ la collection des fonctions partielles de $a$ dans $b$.

**Proposition 4.** La collection $b^a$ est un ensemble, *i.e.* si $a$ et $b$ sont dans $\mathcal{U}$ alors $b^a$ aussi.

**Preuve.** En exercice. □

**Remarque 5 (Réunion indexée).** Soit $a$ une famille d'ensemble indexée par l'ensemble $I$, *i.e.* $a$ est une fonction de domaine $I$. Si $i \in I$, on note $a_i$ pour $a(i)$.

**Proposition 5.** Si $I$ est un ensemble et $a$ est une fonction de domaine $I$, alors $\bigcup_{i \in I} a_i$ est un ensemble. Autrement dit, si dans $\mathcal{U}$, ZF 1, ZF 2, ZF 3, ZF 4 sont vraies, et que $I$ et $a$ sont dans $\mathcal{U}$, et $a$ est une fonction, alors la collection définie naïvement par $\bigcup_{i \in I} a_i$ appartient à $\mathcal{U}$.

**Preuve.** On pose $b := \{a_i \mid i \in I\}$. C'est bien un ensemble car $b$

est l'ensemble des images des éléments de $I$ par $a$. On peut écrire $a$ comme relation fonctionnelle :

$$\varphi(w_0, w_1, a) := (w_0, w_1) \in a.$$

On a donc que $b$ est un ensemble avec ZF 4.

Et, $\bigcup_{i \in I} a_i = \bigcup_{z \in b} z$ donc on conclut par ZF 2. $\square$

**Proposition 6 (Propriété d'intersection).** Si $I$ est un ensemble non vide et $a$ est une fonction de domaine $I$ alors $\bigcap_{i \in I} a_i$ est un ensemble.

**Preuve.** On pose $c := \bigcup_{i \in I} a_i$ qui est un ensemble par ZF 2. On considère

$$\varphi(x, a, I) := \forall i\ i \in I \to x \in a_i.$$

Par compréhension (ZF 4') on construit l'ensemble

$$\bigcap_{i \in I} a_i := \{x \in c \mid \varphi(x, a, I)\}.$$

$\square$

**Proposition 7.** Si $I$ est un ensemble et $a$ une fonction de domaine $i$ alors $\prod_{i \in I} a_i$ est un ensemble.

**Preuve.** La collection $\prod_{i \in I} a_i$ est l'ensemble des fonctions de $I$ dans $\bigcup_{i \in I} a_i$ telles que $f(i) \in a_i$ pour tout $i$. $\square$

**ZF 5** *Axiome de l'infini* : il existe un ensemble ayant une infinité d'élément

$$\exists x\, (\emptyset \in x \land \forall y\, (y \in x \to y \cup \{y\} \in x)).$$

On encode les entiers avec des ensembles :

- ▷ $0 \rightsquigarrow \emptyset$
- ▷ $1 \rightsquigarrow \{\emptyset\}$

- ▷ $2 \rightsquigarrow \{\emptyset, \{\emptyset\}\}$
- ▷ $\vdots$
- ▷ $n+1 \rightsquigarrow n \cup \{n\}$
- ▷ $\vdots$

Ainsi, on a bien $n = \{0, 1, \ldots, n-1\}$.

**Remarque 6.** Si on retire $\emptyset \in x$, on peut avoir $x = \emptyset$ qui satisfait la version modifiée de ZF 5.

Cependant, sans retirer $\emptyset \in x$, on peut quand même avoir un ensemble fini s'il existe un ensemble fini $y$ tel que $y \in y$. Ceci est impossible avec l'axiome de bonne fondation.

**Remarque 7.** Les français sont les seuls à considérer que l'axiome de bonne fondation ne fait pas partie de la théorie de ZF.

# 2 Ordinaux et induction transfinie.

**Théorème 2 (Cantor).**
1. Soient $A$ et $B$ deux ensembles et supposons qu'il existe des injections $A \to B$ et $B \to A$ alors il existe une bijection $A \to B$.
2. Il n'existe pas de surjection de $A \to \wp(A)$.

**Preuve.** En TD. □

**Définition 5.** Deux ensembles sont *équipotents* s'il existe une bijection entre-eux.

**Définition 6.** Soit $A$ un ensemble. Un *ordre* (*partiel*, *strict*) sur $A$ est une relation binaire $<$ (donnée par un sous-ensemble de $A \times A$) telle que

1. *transitivité* : $\forall x\, \forall y\, \forall z\; x < y \to y < z \to x < z$ ;

2. *anti-réflexif* : $\forall x, x \not< x$.

**Notation.** On note $x \leq y$ pour $x < y$ ou $x = y$.

**Exemple 1.** L'ordre $\subsetneq$ sur $\wp(\mathbb{N})$ est partiel. Les ordres $<_\mathbb{N}$, $<_\mathbb{R}$, $<_\mathbb{Z}$ sur $\mathbb{N}, \mathbb{R}, \mathbb{Z}$ sont totaux.

**Définition 7.** Soit $(A, <)$ ordonné. Soient $a, a' \in A$ et $B \subseteq A$. On dit que

- $a$ est un plus petit élément de $B$ si $a \in B$ et pour tout $b \in B$ et si $b \neq a$ alors $b > a$ ;
- $a$ est un élément minimal de $B$ si $a \in B$ et pour tout $b \in B$, $b \not< a$ ;
- $a$ est un minorant de $B$ si pour tout $b \in B$, $a \leq b$ ;
- de la même manière, on définit plus grand élément, élément maximal, majorant ;
- $a$ est une borne inférieure de $B$ si $a$ est un plus grand élément de l'exemple des minorants de $B$ ;
- $a$ et $a'$ sont incomparables si $a \neq a'$ et $a \not< a'$ et $a' \not< a$ ;
- un ordre est bien fondé si toute partie non vide de $A$ a un élément minimal ;
- un bon ordre est un ordre total bien fondé.

**Proposition 8.** Un ordre total est bien fondé ssi il n'existe pas de suite infinie décroissante.

**Preuve.** En exercice. □

**Exemple 2.**

- L'ordre $<$ sur $\mathbb{N}$ est bien fondé.
- L'ordre $<$ sur $\mathbb{Z}$ n'est pas bien fondé.
- L'ordre $\subsetneq$ sur $\wp_{\text{finies}}(\mathbb{N})$ est bien fondé.
- L'ordre $\subsetneq$ sur $\wp(\mathbb{N})$ n'est pas bien fondé.

**Définition 8.**
- Deux ensembles ordonnés sont *isomorphes* s'il existe une bijection préservant l'ordre de l'un vers l'autre. On note $A \simeq B$.
- Soit $X$ totalement ordonné. Un sous-ensemble $J \subseteq X$ est un *segment initial* si pour tous $a, b \in X$ avec $a < b$ alors $b \in J$ implique $a \in J$.
- Un ensemble $X$ est *transitif* si pour tout $x \in X$ et $y \in x$ alors $y \in X$.
- Un ensemble $X$ est un *ordinal* s'il est transitif et que $\in$ défini un bon ordre sur $X$.
- On note $\mathbb{O}$ la *classe des ordinaux*, et on note indifféremment les relations $\in$ et $<$.

**Exemple 3.** Les entiers de Von Neumann sont des ordinaux.

**Proposition 9.** Soient $\alpha$ et $\beta$ des ordinaux. On a les propriétés suivantes :

1. $\emptyset$ est un ordinal ;
2. si $\alpha \neq \emptyset$ alors $\emptyset \in \alpha$ ;
3. $\alpha \notin \alpha$ ;
4. si $x \in \alpha$ alors $x = \{y \in \alpha \mid y < x\}$.
5. si $x \in \alpha$ alors $x$ est un ordinal (on a l'abus de notation $x \in \mathbb{O}$) ;
6. $\beta \leq \alpha$ ssi $\beta = \alpha$ ou $\beta \in \alpha$ ;
7. $x = \alpha \cup \{\alpha\}$ est un ordinal noté $\alpha + 1$.

**Preuve.**
1. C'est vrai.
2. La relation $\in$ est un bon-ordre sur $\alpha$, soit $\beta$ le plus petit élément. Si $\beta \neq \emptyset$ alors il contient au moins un élément $\gamma$ d'où $\gamma < \beta$ et $\gamma \in \alpha$ (par transitivité), *absurde* car $\beta$ minimal.
3. Le reste sera vu en TD ou en exercice.

□

**Proposition 10.**
- Si $\alpha$ et $\beta$ sont des ordinaux et que l'on a $\alpha \leq \beta \leq \alpha + 1$ alors $\beta = \alpha$ ou $\beta = \alpha + 1$.
- Si $X$ est un ensemble non vide d'ordinaux alors $\bigcap_{\alpha \in X} \alpha$ est le plus petit élément de $X$.

□

**Définition 9.** Soit $\beta$ un ordinal.
- S'il existe $\alpha$ tel que $\beta = \alpha + 1$ alors on dit que $\beta$ est un ordinal successeur ;
- Sinon, c'est un ordinal limite.

**Exemple 4.** Quelques ordinaux limites :

$$\omega \qquad \omega \cdot 2 \qquad \omega \cdot 3 \qquad \omega \cdot \omega$$

$$\omega^2 + \omega \qquad \omega^\omega \qquad \omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega}}}}}}.$$

**Lemme 2.** Soit $X$ un ensemble d'ordinaux. Le plus petit élément de $X$ est $\bigcap_{\alpha \in X} \alpha$.

**Théorème 3.** Si $\alpha$ et $\beta$ sont des ordinaux alors une et une seule de ces propriétés est vérifiée :

$$\alpha = \beta \qquad \alpha \in \beta \qquad \alpha \ni \beta.$$

**Preuve.** Soit $X = \{\alpha, \beta\}$, on sait que $\alpha \cap \beta$ est le plus petit élément de $X$.
- Si $\alpha \cap \beta = \alpha$ alors $a \subseteq \beta$ donc $\alpha = \beta$ ou $\alpha \in \beta$.
- Si $\alpha \cap \beta = \beta$ alors $a \supseteq \beta$ donc $\alpha = \beta$ ou $\alpha \ni \beta$.

□

**Proposition 11.** Soit $X$ un ensemble d'ordinaux. Alors l'ensemble $b := \bigcup_{\alpha \in X} \alpha$ est un ordinal. On le note $b = \sup_{\alpha \in X} \alpha$. De plus si $\gamma \in b$ alors il existe un certain $\alpha \in X$ tel que $\gamma \in \alpha$.

**Preuve.** En exercice. □

**Proposition 12.** Soit $\lambda$ un ordinal non vide. On a :

$$\overbrace{\lambda \text{ est limite}}^{(1)} \iff \overbrace{\lambda = \bigcup_{\alpha \in \lambda} \alpha}^{(2)}\,.$$

**Preuve.**

1. Par contraposée, si $\lambda$ n'est pas limite, c'est donc un successeur d'un certain ordinal $\beta$ et donc $\lambda = \beta \cup \{\beta\}$. On a
$$\bigcup_{\alpha \in \lambda} \alpha = \beta \cup \bigcup_{\alpha \in \beta} \alpha = \beta \neq \lambda.$$
2. Soit $\lambda$ limite. Montrons qu'il n'a pas de plus grand élément $\beta$. Sinon, $\lambda = \beta \cup \{\beta\}$. Donc, pour tout $\alpha \in \lambda$ il existe un certain $\gamma \in \lambda$ tel que $\alpha < \gamma$, *i.e.* $\alpha \in \gamma$. On en conclut que $\lambda = \bigcup_{\gamma \in \lambda} \gamma$.

□

**Théorème 4 (Induction transfinie).** Soit $\mathscr{P}$ une propriété sur les ordinaux. On suppose que :

- ▷ $\emptyset$ satisfait $\mathscr{P}$ ;
- ▷ pour tout ordinal $\alpha$ tel que, pour tout $\beta < \alpha$ satisfait $\mathscr{P}$, alors $\alpha$ satisfait $\mathscr{P}$ :
$$\forall \alpha,\ (\forall \beta < \alpha,\ \mathscr{P}(\beta)) \implies \mathscr{P}(\alpha)\ ;$$

alors tous les ordinaux satisfont $\mathscr{P}$.

**Preuve.** Par l'absurde, soit $\alpha$ ne satisfaisant pas $\mathscr{P}$. Soit $\beta$ le plus petit ordinal de $\alpha \cup \{\alpha\}$ ne satisfaisant pas $\mathscr{P}$. Tous les ordinaux plus petit que $\beta$ satisfont $\mathscr{P}$, d'où $\mathscr{P}(\beta)$, ***absurde***. On en conclut que $\alpha$ n'existe pas. □

**Remarque 8.** En pratique on décompose :
- ▷ on montre pour $\emptyset$ ;
- ▷ on montre pour $\alpha$ successeur ;
- ▷ on montre pour $\alpha$ limite.

**Définition 10.** Un ordinal $\alpha$ est *fini* si $\alpha = \emptyset$ ou si $\alpha$ et sous ses éléments sont des successeurs.

**Proposition 13.** L'ensemble des ordinaux finis $\omega$ est un ordinal. C'est le plus petit ordinal limite.

**Preuve.** En exercice. □

**Lemme 3.** Soit $f : \alpha \to \alpha'$ une fonction strictement croissante entre deux ordinaux $\alpha$ et $\alpha'$. Alors $f(\beta) \geq \beta$ pour tout $\beta \in \alpha$. De plus, on a $\alpha' \geq \alpha$. Aussi si $f$ est un isomorphisme alors $\alpha = \alpha'$ et $f$ est l'identité.

**Preuve.**
- ▷ Soit $\beta_0$ le plus petit élément tel que $f(\beta_0) < \beta_0$. Comme $f$ strictement croissante, on a $f(f(\beta_0)) < f(\beta_0) < \beta_0$ absurde car $\beta_0$ est le plus petit.
- ▷ Soit $\beta \in \alpha$. On a $f(\beta) \in \alpha'$ et $\beta \leq f(\beta)$ donc $\beta \in \alpha'$, donc $\beta \in \alpha'$, d'où $\alpha \subseteq \alpha'$ et donc $\alpha \leq \alpha'$.
- ▷ Si $f$ est un isomorphisme alors $f^{-1}$ est strictement croissante. On applique le point précédent à $f^{-1}$, d'où $\alpha = \alpha'$.
- ▷ Montrons que $f$ est l'identité. On sait que, pour tout $\beta \in \alpha$, on a $f_{|\beta}$ strictement croissante de $\beta$ dans $f(\beta)$ et bijective, d'où $\beta = f(\beta)$ par le point précédent. D'où $f$ est l'identité.

□

**Théorème 5.** Tout ensemble bien ordonné est isomorphe à un ordinal. Cet ordinal ainsi que l'isomorphisme sont uniques.

**Preuve.** Cette preuve ressemble à une induction sans en être une. On aura le droit d'en faire quand on aura le théorème.

Si l'isomorphisme existe, il est unique grâce au lemme précédent. En effet, s'il y en a deux $f$ et $g$ alors $f \circ g^{-1}$ est un isomorphisme entre deux ordinaux égaux, donc c'est l'identité.

Notons $\mathscr{P}(x)$ la propriété « il existe un ordinal $\alpha_x$ et un isomorphisme $f_x : S_{\leq x} \to \alpha_x$ » où $S_{\leq x} := \{y \in X \mid y \leq x\}$. Pour montrer $\mathscr{P}(x)$ pour tout $x \in X$, on pose

$$Y := \{x \in X \mid \mathscr{P}(x) \text{ est vraie}\}.$$

et on montre $Y = X$.

Supposons $Y \neq X$ et soit $a = \min(X \setminus Y)$.

- Si $Y$ a un plus grand élément $b$, alors il existe un isomorphisme $f_b : S_{\leq b} \to \alpha_b$ (car $\mathscr{P}(b)$). Or, $S_{\leq a} = S_{\leq b} \cup \{a\}$ (il faudrait montrer que $Y$ est un segment initial de $X$). Et, on construit $f_a : S_{\leq a} \to \alpha_b \cup \{\alpha\}_b$ qui est un isomorphisme, donc $a \in Y$, ***absurde***.
- Si $Y$ n'a pas de plus grand élément, on considère $\alpha := \bigcup_{x \in Y} \alpha_x$ un ordinal. Pour tout $x < \alpha$ il existe un isomorphisme de $S_{\leq x}$ dans $\alpha_x$. Si on prend $x < y < \alpha$ alors $(f_y)_{|S_{\leq x}}$ est un isomorphisme et, par unicité, on a donc que $f_y$ prolonge $f_x$. On peut définir un isomorphisme $f_{\leq \alpha}$ comme limite des $f_x$ pour $x < \alpha$. C'est un isomorphisme de $S_{<a}$ dans $\beta := \bigcup_{x<\alpha} \alpha_x$. On peut prolonger $f$ en $f_X : S_{\leq \alpha} \to \beta + 1$ où $\alpha \mapsto \beta$, d'où $\alpha \leq Y$, ***absurde***.

$\square$

**Lemme 4** (Définition récursive transfinie des fonctions)**.** Soient

- $\alpha$ un ordinal ;

- ▷ $S$ une collection ;
- ▷ $\mathscr{F}$ la collection des applications définies sur les ordinaux $\beta \leq \alpha$ et prenant leur valeurs dans $S$ ;
- ▷ $F$ une relation fonctionnelle de domaine $\mathscr{F}$ à valeur dans $S$.

Alors il existe une fonction $f$ dans $\mathscr{F}$ (et une unique définie sur $\alpha$) telle que :

$$(\star) \qquad \text{pour tout } \beta < \alpha \qquad f(\beta) = F(f_{|\beta}).$$

**Preuve.** **Unicité.** Soient $f$ et $g$ satisfaisant $(\star)$. Montrons $\mathscr{P}(\beta)$ : « si $\beta < \alpha$ alors $f(\beta) = g(\beta)$ » par induction transfinie.

- ▷ On a directement $\mathscr{P}(\emptyset)$ car il y a une unique fonction de $\emptyset \to \emptyset$.
- ▷ Supposons que $f(\gamma) = g(\gamma)$ pour tout $\gamma < \beta$. Alors $f_{|\beta} = g_{|\beta}$ et donc par $(\star)$ on a $f(\beta) = g(\beta)$.

**Existence.** Soit $\tau$ l'ensemble des ordinaux $\gamma \in \alpha$ tels qu'il existe $f_\gamma \in \mathscr{F}$ définie sur $\gamma$ et vérifiant $(\star)$. Alors $\tau$ est un segment initial de $\alpha$ donc un ordinal. Par unicité si $\gamma < \gamma'$ alors $f_{\gamma'}$ prolonge $f_\gamma$. On définit $f_\tau$ par $f_\tau(\gamma) := F(f_\gamma)$ si $\gamma < \tau$.

- ▷ Si $\tau \in \alpha$ alors $f_\tau$ prolonge tous les $f_\gamma$ et donc $\tau \in \tau$, ***impossible***.
- ▷ D'où $\tau = \alpha$.

□

**Exercice 1.** Généraliser la preuve ci-dessus en replaçant $\alpha$ par la classe de tous les ordinaux $\mathscr{O}$ (et remplacer $\mathscr{F}$ par autre chose).

**Proposition 14.**
1. La classe $\mathscr{O}$ n'est pas un ensemble.
2. Il n'existe pas de relation fonctionnelle bijective entre $\mathscr{O}$ et un ensemble $a$.

**Preuve.**
1. En effet, supposons $\mathscr{O}$ un ensemble. On a que $\mathscr{O}$ est transitif et $\in$ y définit un ordre total, donc $\mathscr{O}$ est un ordinal.

D'où $\mathfrak{O} \in \mathfrak{O}$ ce qui est impossible pour un ordinal.

2. Sinon $\mathfrak{O}$ serait un ensemble.

□

# 3 Axiome de choix et variantes équivalentes.

Les axiomes du choix sont exprimable au premier ordre !

**AC 1.** Le produit d'une famille d'ensembles non vides est non vide.

**AC 2.** Pour tout ensemble $a$ non vide, il existe une fonction de $\mathscr{P}(a)$ dans $a$ tel que si $x \subseteq a$ est non vide alors $f(x) \in x$. (C'est une fonction de choix.)

**AC 3.** Si $a$ est un ensemble dont tous les éléments sont non vides et deux à deux disjoints alors il existe un ensemble $c$ tel que, pour tout $x \in a$, $x \cap c$ a exactement un élément.

**(Lemme de) Zorn.** Tout ensemble non vide partiellement ordonné et inductif admet un élément maximal.

On rappelle qu'un ensemble partiellement ordonné $X$ est inductif si $X \neq \emptyset$ et si tout sous-ensemble $Y \subseteq X$ totalement ordonné admet un majorant dans $X$.

**(Lemme de) Zermelo.** Tout ensemble non vide peut être mini d'un bon ordre (*i.e.* un ordre total où toute partie non vide a un plus petit élément).

En supposant les axiomes ZF 1, . . . , ZF 5, on va montrer les implications suivantes :

$$\begin{tikzcd}
\text{AC 3} \arrow[rr, Rightarrow, "3.3"] & & \text{AC 1} \arrow[dl, Rightarrow, "3.1"] \\
& \text{AC 2} \arrow[ul, Rightarrow, "3.2"] \arrow[dl, Rightarrow, "3.7"] \arrow[dr, Rightarrow, bend left, "3.4"] & \\
\text{Zorn} \arrow[uu, Rightarrow, "3.6"] & & \text{Zermelo} \arrow[ul, Rightarrow, bend left, "3.5"]
\end{tikzcd}$$

## 3.1 AC 1 implique AC 2.

On rappelle que l'ensemble $\prod_{i \in I} a_i$ est l'ensemble de fonctions $f$ de la forme $f : I \to \bigcup_{i \in I} a_i$ tel que $f(i) \in a_i$ pour tout $i$.

Soit $a$ non vide. On considère $\prod_{\emptyset \neq x \subseteq a} x$ qui est non vide par AC 1. Soit $f$ un de ces éléments. On a, pour tout $\emptyset \neq x \subseteq a$, que $f(x) \in x$ donc $f$ est une fonction de choix.

## 3.2 AC 2 implique AC 3.

Soit $a$ un ensemble dont les éléments sont non vides et deux à deux disjoints. On considère $b = \bigcup_{x \in a} x$ qui est un ensemble. Par AC 2, on a une fonction de choix $f$ sur $\wp(b)$. On prend $c = \{f(x) \mid x \in a\}$. Comme les $x$ sont disjoints, on obtient la propriété recherchée.

## 3.3 AC 3 implique AC 1.

Soit $X = \prod_{i \in I} a_i$ un produit d'ensemble non vides. On considère $A := \{\, \{i\} \times a_i \mid i \in I \,\}$. Par AC 3, il existe $c$ tel que, pour tout $x \in A$, $x \cap c$ a exactement un élément. D'où, $c$ peut s'écrire

$$c = \{(i, d_i) \mid i \in I \text{ et } d_i \in a_i\}.$$

On a donc $c \in \prod_{i \in I} a_i$ (c'est le graphe d'une fonction) et donc $\prod_{i \in I} a_i \neq \emptyset$.

## 3.4 AC 2 implique Zermelo.

Soit $a$ un ensemble non vide.

> **Remarque 9 (Idée).** L'idée est qu'on utilise $f : \mathscr{P}(a) \to a$ une fonction de choix pour définir l'ordre. On peut imaginer définir $x \leq y$ ssi $f(\{x, y\}) = x$ (on traite donc $f$ comme la fonction minimum), mais on n'a pas la transitivité. Il faut être plus futé.
>
> On va construire en partant du plus petit une bijection entre $a$ et un ordinal.

> - $h(0) := f(a)$
> - $h(1) := f(a \setminus \{h(0)\})$
> - $h(2) := f(a \setminus \{h(0), h(1)\})$
> - $\vdots$
> - $h(\omega) = f(a \setminus \{h(0), h(1), \dots\})$
> - $\vdots$
>
> On s'arrête quand $a \setminus \{h(0), h(1), \dots\}$ est vide.

Soit $\theta \notin a$ (pour « détecter » quand l'ensemble $a \setminus \{h(0), \dots\}$ est vide). Il existe car $a$ est un ensemble donc pas l'univers tout entier.

On définit

$$F(\alpha) := \begin{cases} f(a \setminus \{F(\beta) \mid \beta < \alpha\}) & \text{si } a \setminus \{F(\beta) \mid \beta < \alpha\} \neq \emptyset \\ \theta & \text{sinon} \end{cases}.$$

D'après le dernier lemme de la section précédente, on peut construire l'application $F$ ainsi et elle est unique.

S'il n'existe pas d'ordinal $\alpha$ tel que $F(\alpha) = \theta$ alors $F$ est une injection de $\mathbb{O}$ dans $a$. ***Absurde***. Il existe donc $\alpha$ tel que $F(\alpha) = \theta$.

Le sous-ensemble $\{\beta \in \alpha \mid F(\beta) = \theta\}$ a un plus petit élément $\beta$. Montrons que $F_{|\beta}$ est une bijection de $\beta$ dans $\alpha$.

- D'une part, on sait que c'est une injection.
- D'autre part, $F(\beta) = \theta$ implique $\{F(\gamma) \mid \gamma < \beta\} = a$ donc $F_{|\beta}$ est une surjection.

On définit le bon ordre $x \prec y$ ssi $F^{-1}(x) < F^{-1}(y)$.

## 3.5 Zermelo implique AC 2.

Soit $a$ non vide. Il existe un bon-ordre $<$ sur $a$. Soit $\emptyset \neq x \subseteq a$. On définit $f(x) = \min x$, c'est bien une fonction de choix.

## 3.6 Zorn implique AC 3.

Soit $a$ un ensemble dont les éléments sont disjoints et non vides. On pose :

$$b := \bigcup_{x \in a} x \quad \text{et} \quad X := \{c \subseteq b \mid \forall x \in a,\ |c \cap x| \leq 1\}.$$

Montrons que l'ensemble $(X, \subsetneq)$ est inductif.

Soit $Y \subseteq X$ est totalement ordonné. Montrons que $Y$ a un majorant dans $X$. On pose $z = \bigcup_{y \in Y} y$ qui majore $Y$. On a bien $z \in X$ (on ne duplique pas les éléments). On en conclut que $X$ est inductif. Soit $d$ un élément maximal de $X$ (il existe par Zorn).

- ▷ S'il existe $x \in a$ tel que $x \cap d = \emptyset$ alors prenons $u \in x$ et posons $d_1 := d \cup \{u\}$. D'où $d_1 \in X$ et $d \subsetneq d_1$ donc $d$ non maximal, ***absurde***.
- ▷ Pour tout $x \in a$, on a $|d \cap x| = 1$, d'où $d$ est l'ensemble recherché (appelé $c$ dans AC 3).

## 3.7 AC 2 implique Zorn.

Soit $a$ un ensemble inductif.

> **Remarque 10 (Idée).** On construit une chaîne dans $a$ (*i.e.* un ensemble totalement ordonné) de taille maximale (c'est ici qu'on utilise la fonction de choix de AC 2). Elle a un majorant car $a$ est inductif. Ce majorant va être l'élément maximal.

Soit $f : \wp(a) \to a$ une fonction de choix donnée par AC 2. Si $x \subseteq a$ on appelle majorant strict de $x$ un $y \in a$ tel que $z < y$ pour tout $y \in x$.

> **Remarque 11 (Idée – suite).**
> - ▷ On part de $\emptyset$ : tout élément de $a$ a un majorant strict de $\emptyset$ en choisissant $a_1 = f(a)$.
> - ▷ Soit $X_1$ l'ensemble des majorants de $\{a_1\}$. On pose $a_2 := f(X_1)$.

> ▷ Soit $X_2$ l'ensemble des majorants de $\{a_1, a_2\}$. On pose $a_3 := f(X_2)$.

Formellement, soit $C := \{x \subseteq a \mid x \text{ a un majorant strict dans } a\}$. On a $\emptyset \in C$. On définit

$$\begin{aligned} m : C &\longrightarrow a \\ x &\longmapsto f(\{y \in a \mid y \text{ est un majorant strict de } x \text{ dans } a\}). \end{aligned}$$

On définit par induction la chaîne maximale (dernier lemme de la section précédente). Soit $\theta \notin a$. On définit :

$$F(\alpha) := \begin{cases} m(\{F(\beta) \mid \beta < \alpha\}) & \text{si } \{F(\beta) \mid \beta < \alpha\} \in C \\ \theta & \text{sinon} \end{cases}.$$

La fonction $F$ n'est pas une injection de $\mathfrak{O}$ dans $a$ donc il existe un ordinal $\alpha$ tel que $F(\alpha) = \theta$. Comme $\alpha + 1$ est un ordinal, l'ensemble $\{\beta \in \alpha + 1 \mid F(\beta) = \theta\}$ a un plus petit élément $\alpha_0$. D'où l'ensemble $\{F(\beta) \mid \beta < \alpha_0\}$ n'a pas de majorant strict mais a un majorant $M$ car $a$ inductif. Et, $a$ n'a pas d'élément plus grand que $a$. Ainsi $M$ est maximal dans $a$.

## 3.8 Indépendance de ZF et de l'axiome du choix.

On a les deux théorèmes suivants (que l'on admet).

**Théorème 6 (Gödel, 1938).** S'il est cohérent, ZF ne réfute pas l'axiome du choix.

**Théorème 7 (Cohen, 1963).** S'il est cohérent, ZF ne montre pas l'axiome du choix.

Ainsi l'axiome du choix est indépendant de ZF. Cependant, il existe des versions plus faibles : axiome du choix dépendant (ACD), axiome du choix dénombrable ($\mathsf{AC}_\omega$).