

# Introduction (facultative) à la théorie des ensembles : Axiomes de Zermelo-Fraenkel

## I Un peu d'Histoire

Vers la fin du XIX<sup>ème</sup> siècle, mathématicien(ne)s et philosophes se sont trouvé face à une «crise des fondements» : après des millénaires de calculs, de géométrie, de physique, pour laquelle on a constaté par exemple l'efficacité des notions mathématiques de dérivée et d'intégrales ; on se rend compte que l'on est parfaitement incapable de définir clairement et proprement les objets manipulés. Entre autres, impossible de dire ce qu'est le nombre 1 (ou pire : 0). Euclide et Aristote ont bien essayé, mais ce n'était plus convaincant arrivé à cette époque : voir par exemple *Les Fondements de l'Arithmétique*<sup>1</sup> de G. Frege, un peu difficile à lire mais très instructif. Frege tente de donner des axiomes de base pour l'arithmétique mais l'un d'entre eux, essentiel pour Frege, est tout-de-même problématique. En effet, Russell en tire un célèbre paradoxe : l'axiome de Frege implique l'existence d'un ensemble  $X$ , formé des ensembles qui ne se contiennent pas eux-même, en terme de quantificateurs,

$$\exists X \forall x [(x \notin x) \Leftrightarrow (x \in X)] \quad \text{ou encore} \quad X = \{x \mid x \notin x\}.$$

Le paradoxe est le suivant : on a deux possibilités, soit  $X$  se contient lui-même, soit pas. Si  $X$  se contient lui-même, comme  $X$  ne contient que des ensembles qui ne se contiennent pas eux-même,  $X$  ne doit pas non-plus se contenir lui-même, or on a supposé l'inverse : contradiction. Ainsi,  $X$  ne doit pas se contenir lui-même ; c'est donc un ensemble qui ne se contient pas lui-même et, par définition, ceci nous dit que  $X$  appartient à  $X$ , donc est élément de lui-même : une nouvelle contradiction. Dans les deux cas, on obtient une contradiction et donc la théorie de Frege n'est pas cohérente...

D'autre part, on pensait jusque là que toute courbe était représentable par une fonction continue «sans points anguleux». Or, Weierstrass a construit une fonction continue partout et nulle-part dérivable (on peut même démontrer que «l'immense majorité» des fonctions continues ne sont dérivables nulle-part).

Enfin, grosse question qui donne des sueurs froides à Hilbert<sup>2</sup> : l'arithmétique (au sens large, comprendre les maths) est-elle cohérente ? Autrement dit, les axiomes de l'arithmétique telle que l'on la connaît mènent-ils à une contradiction ? Si c'était le cas, ce serait terrible car cela signifierait que les résultats que l'on connaît se contredisent les uns les autres ; rien n'est vrai, toutes les maths sont à mettre à la poubelle !

Bref, il est temps de faire le ménage et de repartir de zéro : reconstruire les maths depuis le début, avec seulement la logique pour donnée. La première idée est la suivante : pas de fonctions, d'opérations, de droites etc, seulement des ensembles ! On avait déjà bien remarqué que la notion d'ensemble est inextricable : il faut savoir ce qu'est un ensemble. Ensuite, toutes les autres notions (fonctions...) sont construites uniquement à partir des ensembles.

La première théorie des ensembles qui tienne debout date des années 1922–1925 : la théorie de Zermelo et Fraenkel, abrégée en ZF. On peut y ajouter «sans problème» un axiome bien pratique qu'est l'axiome du choix, obtenant la théorie ZFC. Ce solide paradigme est encore très largement utilisé dans la communauté mathématique et est suffisant pour les chercheurs non spécialistes du sujet. Signalons qu'il existe plusieurs autres théories des ensembles, certaines se permettant de traiter non seulement des ensembles mais aussi des «classes» : des collections d'objets a priori plus grandes que des ensembles. Mais nous allons voir que même dans la théorie de ZF, il apparaît des collections d'objets qui ne sont pas des ensembles.

Malheureusement, l'espoir de cohérence de Hilbert fût brisé par les Théorèmes d'incomplétude de Gödel en 1931 : informellement, dans toute théorie contenant l'arithmétique usuelle, il existe des énoncés indécidables (on ne peut les démontrer ni les nier) et, comme conséquence, on ne peut affirmer la cohérence d'une théorie à partir de ses axiomes. Ainsi, la réponse au deuxième problème de Hilbert est non.

Étant donné que la théorie ZFC n'a pas encore produit d'incohérence et que de toute façon, on ne peut pas prouver qu'elle est cohérente, on continue de l'utiliser tant qu'on ne trouve pas de contradiction.

1. titre original : Die Grundgesetze der Arithmetik

2. deuxième problème de sa célèbre liste de 23 problèmes (1900), dont la moitié est résolue

## II Les axiomes de Zermelo-Fraenkel

Nous allons ici introduire les cinq (schéma) d'axiomes constituant la théorie de ZF. L'ajout de l'axiome du choix donne la théorie ZFC. Une bonne référence pour ce qui suit est le livre *Théorie des ensembles* de Jean-Louis Krivine. Rappelons que son contenu est largement hors programme, ainsi que ce polycopié.

### II. 1 Univers, ensembles et trois premiers axiomes (extensionnalité, réunion, parties)

#### Donnée

Nous considérons et fixons une fois pour toute une collection (au sens intuitif) d'objets  $\mathcal{U}$ , que nous appellerons **univers**. Les objets de cet univers sont appelés **ensembles**.

L'univers  $\mathcal{U}$  est muni d'une relation binaire, notée  $\in$  et appelée **relation d'appartenance**. Pour deux ensembles  $x$  et  $y$ , on dira que « $x$  appartient à  $y$ » ou que « $x$  est élément de  $y$ » s'ils vérifient la relation  $x \in y$ .

Dans la suite, on réservera le mot «appartenir» à cette relation, de même pour le mot «élément». Par exemple, on ne dira pas qu'un ensemble «appartient» à  $\mathcal{U}$  ou qu'il est «élément» de  $\mathcal{U}$ , puisqu'un ensemble ne peut appartenir qu'à un autre ensemble et non à un univers. On dira donc plutôt qu'un ensemble est dans  $\mathcal{U}$ .

Les axiomes de ZF sont les propriétés élémentaires que l'on impose à l'univers  $\mathcal{U}$  et à la relation  $\in$ . Avant tout, il nous reste à définir l'égalité.

#### Définition

Deux ensembles  $x$  et  $y$  dans  $\mathcal{U}$  sont dits **égaux** s'ils ont les mêmes éléments et appartiennent aux mêmes ensembles. En formule,

$$(x = y) \stackrel{\text{def}}{\iff} (\forall z[z \in x \iff z \in y] \wedge \forall w[x \in w \iff y \in w]).$$

Pour l'intuition, on pourra dire que «deux ensembles sont égaux s'ils sont les mêmes». Autrement dit, on dit que deux ensembles sont «égaux» s'il est impossible de les discerner à l'aide de la relation  $\in$ .

Dans la suite, pour éviter un discours vide de sens, on suppose que l'univers  $\mathcal{U}$  est non vide, en ce sens que

$$\exists x(x = x),$$

c'est-à-dire qu'il existe au-moins un ensemble égal à lui-même, i.e. il existe au-moins un ensemble.

Le premier axiome de ZF est l'axiome d'extensionnalité, qui nous dit que si deux ensembles ont les mêmes éléments, alors ils appartiennent en fait aux mêmes ensembles, i.e. sont égaux.

#### Axiome 1 : Axiome d'extensionnalité

Dans  $\mathcal{U}$ , deux ensembles sont égaux si et seulement s'ils ont les mêmes éléments. En formule,

$$\forall x \forall y [\forall z (z \in x \iff z \in y) \Rightarrow x = y].$$

Étant donnés deux ensembles  $x$  et  $y$ , on peut former leur **paire**  $\{x, y\}$ , un ensemble dont les seuls éléments sont  $x$  et  $y$ . C'est un axiome dans la théorie originale de Zermelo, mais dans la théorie ZF, il résulte du schéma d'axiome de substitution et de l'axiome de l'ensemble des parties, que nous verrons plus bas. En prenant  $x = y$ , on peut construire le **singleton**  $\{x\}$ , ensemble dont le seul élément est l'ensemble  $x$ .

On peut ensuite former la **paire ordonnée**  $(x, y) := \{\{x\}, \{x, y\}\}$ ; de la sorte, on a  $(x, y) = (x', y') \iff (x = x' \wedge y = y')$ . On retrouve bien la notion de «paire» telle qu'utilisée en mathématiques, par exemple les éléments du plan  $\mathbb{R}^2$  sont les paires  $(a, b)$  avec  $a, b \in \mathbb{R}$ . On peut ensuite itérer la construction et parler de triplet :  $(x, y, z) := (x, (y, z))$ , puis ensuite de  $n$ -uplet...

Le deuxième axiome est celui de la réunion (ou de la somme).

### Axiome 2 : Axiome de la réunion

Pour tout ensemble  $x$ , il existe un ensemble (et un seul, d'après l'Axiome 1) dont les éléments sont les éléments des éléments de  $x$ , autrement dit,

$$\forall x \exists y \forall z [z \in y \Leftrightarrow \exists t (z \in t \wedge t \in x)].$$

Cet ensemble est noté  $\bigcup_{a \in x} a$  ou plus rarement  $\bigcup x$ .

Si  $x, y, z$  sont trois ensembles, il existe un ensemble  $t$  dont les éléments sont  $x, y, z$  et eux seulement : c'est la réunion de l'ensemble  $\{\{x, y\}, \{z\}\}$ , que l'on note  $\{x, y, z\}$ .

Si  $x$  et  $y$  sont deux ensembles, la réunion de  $\{x, y\}$  est appelée **réunion** de  $x$  et de  $y$  et on la note  $x \cup y := \bigcup \{x, y\}$ . Pour trois ensembles  $x, y, z$ , on a  $x \cup (y \cup z) = (x \cup y) \cup z = \bigcup \{x, y, z\}$  et on note cette réunion  $x \cup y \cup z$ . Ainsi, une fois que l'on aura le schéma d'axiome de substitution et l'axiome de l'ensemble des parties, on pourra parler de paire, de paire ordonnée et de réunion de deux (ou d'un nombre fini d') ensembles.

### Définition

On dit qu'un ensemble  $x$  est **inclus** dans un ensemble  $y$  et on note  $x \subset y$  si tout élément de  $x$  est aussi élément de  $y$ , i.e.

$$(x \subset y) \stackrel{\text{def}}{\iff} \forall z [z \in x \Rightarrow z \in y].$$

### Axiome 3 : Axiome des parties

Cet axiome exprime qu'étant donné un ensemble  $x$ , il existe un ensemble dont les éléments sont les objets de  $\mathcal{U}$  qui sont inclus dans  $x$  :

$$\forall x \exists y \forall z [z \in y \Leftrightarrow z \subset x].$$

D'après le premier axiome, cet ensemble est unique et on le note  $\mathcal{P}(x)$ .

## II. 2 Schéma d'axiomes de substitution et conséquences

On en arrive maintenant à un axiome un peu particulier, mais indispensable. En effet, ce n'est pas un axiome comme les précédents, car il contient une infinité d'axiomes à la fois, mais qui sont tous de la même forme<sup>3</sup>. On parle alors de «schéma d'axiomes».

Étant donnés des ensembles fixés  $x_1, \dots, x_k$ , un **énoncé**  $E(x_1, \dots, x_k, x, y)$  est une proposition, qui peut être soit vraie, soit fausse, portant sur les ensembles fixés  $x_1, \dots, x_k$  et sur deux ensembles «libres»  $x$  et  $y$  qui peuvent varier. On dit que  $E(x_1, \dots, x_k, x, y)$  est un énoncé à  $k$  paramètres et deux arguments.

**Exemple.** Un tel énoncé (pour  $k = 3$ ) est

$$E(x_1, x_2, x_3, x, y) \Leftrightarrow (x_1 \subset x_2 \cup y) \wedge (x_3 \subset x_2 \cup x).$$

### Définition

On dit qu'un énoncé  $E(x_1, \dots, x_k, x, y)$  à  $k$  paramètres et deux arguments est **fonctionnel** si

$$\forall x_1 \dots \forall x_k (\forall x \forall y \forall y' [E(x_1, \dots, x_k, x, y) \wedge E(x_1, \dots, x_k, x, y') \Rightarrow y = y']).$$

Autrement dit, pour tous ensembles  $x_1, \dots, x_k, x$ , il existe au-plus un seul ensemble  $y$  tel que  $E(x_1, \dots, x_k, x, y)$ .

3. Montague a montré dans sa thèse en 1957 qu'il était impossible d'axiomatiser ZF à l'aide d'un nombre fini d'axiomes...

**Exemple.** L'énoncé de l'exemple précédent n'est pas fonctionnel.

Le terme «fonctionnel» provient du fait que c'est la bonne notion de fonction. Nous détaillerons ceci après avoir énoncé le schéma d'axiome de substitution et introduit la notion de produit de deux ensembles. Le schéma d'axiomes est le suivant :

#### Axiome 4 : Schéma d'axiomes de substitution

Si  $E(x_1, \dots, x_k, x, y)$  est un énoncé fonctionnel à  $k$  paramètres et deux arguments et  $t$  un ensemble quelconque, alors il existe un ensemble  $z$  dont les éléments sont les images par la relation  $E$  des éléments  $x$  de  $t$  qui sont dans le domaine de  $E$  (i.e. la collection des  $x \in t$  tels qu'il existe un ensemble  $a$  vérifiant  $E(x_1, \dots, x_k, x, a)$ ). Autrement dit,

$$\forall x_1 \dots \forall x_k \forall t \exists z \forall v [v \in z \Leftrightarrow \exists u (u \in t \wedge E(x_1, \dots, x_k, u, v))].$$

**Remarque.** Si l'on veut inclure le fait que  $E$  est fonctionnel, le schéma d'axiomes précédent se réécrit

$$\begin{aligned} \forall x_1 \dots \forall x_k \{ \forall x \forall y \forall y' [E(x_1, \dots, x_k, x, y) \wedge E(x_1, \dots, x_k, x, y') \Rightarrow y = y'] \\ \Rightarrow \forall t \exists z \forall v (v \in z \Leftrightarrow \exists u [u \in t \wedge E(x_1, \dots, x_k, u, v)]) \}. \end{aligned}$$

Ce schéma d'axiomes a une conséquence importante, appelée schéma d'axiomes de compréhension. Même si c'est un résultat et non un axiome à proprement parler, on continue de l'appeler «axiome» car, historiquement, c'est cet axiome qui fût introduit en premier par Zermelo, avant d'être amélioré en le schéma d'axiomes 4 par Fraenkel.

#### Corollaire : Schéma d'axiomes de compréhension

Si  $A(x_1, \dots, x_k, x)$  est un énoncé à  $k$  paramètres et un argument et si  $x$  est un ensemble quelconque, alors il existe un ensemble  $z$  dont les éléments sont les éléments  $u$  de  $x$  qui vérifient  $A(x_1, \dots, x_k, u)$ . Ceci s'écrit

$$\forall x_1 \dots \forall x_k \forall x \exists z \forall u [u \in z \Leftrightarrow (u \in x \wedge A(x_1, \dots, x_k, u))],$$

cet ensemble  $z$  sera noté  $\{u \in x \mid A(x_1, \dots, x_k, u)\}$ .

*Démonstration.* On applique le schéma 4 à l'énoncé fonctionnel défini par

$$E(x_1, \dots, x_k, u, v) \stackrel{\text{def}}{\Leftrightarrow} (u = v \wedge A(x_1, \dots, x_k, u)).$$

□

#### Remarques.

▷ Considérons l'énoncé à un argument sans paramètre défini par

$$A(u) \stackrel{\text{def}}{\Leftrightarrow} u \neq u$$

et soit  $x$  un ensemble quelconque. L'ensemble  $\{u \in x \mid A(u)\}$ , dont l'existence est assurée par le schéma d'axiomes de compréhension, ne contient aucun élément : aucun ensemble  $u$  ne vérifie  $u \neq u$ . Par le premier axiome, cet ensemble est unique (en particulier, indépendant de  $x$ ), on l'appelle l'**ensemble vide** et on le note  $\emptyset$ . Nous venons de démontrer que les (schémas d') axiomes 1 et 4 impliquent l'existence et l'unicité de l'ensemble vide !

▷ Le schéma d'axiomes de compréhension ressemble à l'axiome introduit par Frege et que nous avons mentionné dans la première section. Cependant, ce n'est pas le même axiome ! L'axiome de Frege dit que pour toute propriété, il existe un ensemble dont les éléments sont les ensembles vérifiant cette propriété et cela mène au paradoxe de Russell. Ici c'est un peu différent : on affirme que pour toute propriété **et tout ensemble**  $x$ , il existe un ensemble formé **des éléments de**  $x$  vérifiant la propriété ; et ceci supprime le paradoxe. En effet, on ne peut pas parler de l'ensemble des  $u$  tels que  $u \notin u$  puisqu'aucun ensemble «ambiant» n'est disponible. Par contre, étant donné un ensemble  $x$  fixé, on peut bien parler de l'ensemble  $z$  des  $u \in x$  tels que  $u \notin u$ . Si  $z \in x$  alors  $z \in z$  ou  $z \notin z$  et on retrouve le paradoxe de Russell. On en déduit donc que  $z \notin x$  et donc  $z \notin z$  et... c'est tout ! plus d'incohérence.

- ▷ Du point précédent, on tire que la collection des ensembles  $x$  tels que  $x \notin x$  n'est **pas** un ensemble, c'est seulement une partie (au sens intuitif, pas celui introduit plus haut à l'aide du symbole  $\subset$ ) de l'univers  $\mathcal{U}$ . On peut montrer aussi que l'univers  $\mathcal{U}$  lui-même n'est pas un ensemble. Ainsi, les axiomes présentés suggèrent déjà qu'il doit exister des entités plus grandes et sauvages que les ensembles...

On peut maintenant prouver l'existence des paires (et donc des paires ordonnées et des singletons) à partir des (schémas d') axiomes 1, 3 et 4 :

**Lemme : Existence de la paire**

Si  $x$  et  $y$  sont deux ensembles, alors il existe un unique ensemble ayant  $x$  et  $y$  pour seuls éléments. On note cet ensemble  $\{x, y\}$ .

*Démonstration.* L'unicité provient de l'axiome 1 ; montrons donc l'existence. Toute partie de  $\emptyset$  est vide, donc l'ensemble des parties de  $\emptyset$  est  $\mathcal{P}(\emptyset) = \{\emptyset\}$ . Ce dernier ensemble n'a qu'un seul élément, donc si  $x \subset \{\emptyset\}$ , alors  $x = \emptyset$  ou  $x = \{\emptyset\}$ , soit  $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$  et cet ensemble a donc exactement deux éléments. Étant donnés  $x$  et  $y$ , définissons la relation

$$E(u, v) \stackrel{\text{def}}{\iff} (u = \emptyset \wedge v = x) \vee (u = \{\emptyset\} \wedge v = y).$$

C'est une relation fonctionnelle à 0 paramètre et à un argument et le schéma d'axiomes 4 appliqué à l'ensemble  $t = \{\emptyset, \{\emptyset\}\}$  donne un ensemble  $z$  qui a bien pour seuls éléments  $x$  et  $y$ . En effet, pour tout ensemble  $v$ , on a

$$\begin{aligned} v \in z &\iff \exists u \{u \in \{\emptyset, \{\emptyset\}\} \wedge E(u, v)\} \\ &\iff \exists u \{[u = \emptyset \vee u = \{\emptyset\}] \wedge [(u = \emptyset \wedge v = x) \vee (u = \{\emptyset\} \wedge v = y)]\} \\ &\iff \exists u \{(u = \emptyset \wedge v = x) \vee (u = \{\emptyset\} \wedge v = y)\} \\ &\iff v = x \vee v = y. \end{aligned}$$

□

Comme annoncé plus haut, une fois que l'on sait définir les paires, on peut définir le *singleton*  $\{x\}$ , la *paire ordonnée*  $(x, y) := \{\{x\}, \{x, y\}\}$  et l'union  $x \cup y := \bigcup \{x, y\}$  (axiome 2).

On peut à présent définir le produit de deux ensembles  $x$  et  $y$ . Dans la théorie naïve, le produit  $x \times y$  est l'ensemble des paires  $(u, v)$  où  $u \in x$  et  $v \in y$ . Le schéma d'axiomes de compréhension permet de formaliser ceci.

**Lemme**

Étant donnés deux ensembles  $x$  et  $y$  dans  $\mathcal{U}$ , il existe un unique ensemble dont les éléments sont les paires ordonnées  $(u, v)$  où  $u \in x$  et  $v \in y$ . On le note  $x \times y$ .

*Démonstration.* Considérons l'énoncé à un argument

$$X(w) \stackrel{\text{def}}{\iff} \exists u \exists v [w = (u, v) \wedge u \in x \wedge v \in y].$$

La collection des ensembles  $w$  de  $\mathcal{U}$  vérifiant  $X(w)$  est un ensemble. En effet, si  $u \in x$  et  $v \in y$ , alors  $(u, v) \stackrel{\text{def}}{=} \{\{u\}, \{u, v\}\} \in \mathcal{P}(\mathcal{P}(x \cup y))$ , donc

$$\forall w [X(w) \iff (w \in \mathcal{P}(\mathcal{P}(x \cup y)) \wedge X(w))]$$

et donc, par le schéma d'axiomes de compréhension, il existe un ensemble  $z$  tel que

$$\forall w [w \in z \iff X(w)],$$

qui est bien l'ensemble cherché. L'unicité découle de l'axiome 1.

□

**Remarque.** Soient  $x$  un ensemble et  $y \subset x$  une partie de  $x$ . On peut définir le **complémentaire de  $y$  dans  $x$**  par la collection  $C(u) \stackrel{\text{def}}{\iff} (u \in x \wedge u \notin y)$ . Par le schéma d'axiomes de compréhension, ceci définit un ensemble, qui est une partie de  $x$ . On note cet ensemble  $x \setminus y$ .

## II. 3 Notion d'application entre ensembles, réunion, intersection et produit d'une famille d'ensembles

Nous pouvons en revenir à la définition d'une *fonction*. Au final, qu'est-ce qu'une fonction sinon son graphe ? Par exemple, une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  peut être identifiée avec le sous-ensemble  $\{(a, b) \in \mathbb{R} \times \mathbb{R} ; b = f(a)\}$  de  $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ . La propriété importante d'une fonction est qu'étant donné un élément de l'ensemble de départ, il admet une unique image par la fonction dans l'ensemble d'arrivée. Pour l'exemple d'une fonction réelle  $f : \mathbb{R} \rightarrow \mathbb{R}$ , on peut définir la relation binaire  $F$  suivante sur  $\mathbb{R}^2$  :

$$F(a, b) \Leftrightarrow b = f(a).$$

Cette relation est fonctionnelle à un argument : si on a  $F(a, b)$  et  $F(a, b')$ , alors  $b = f(a) = b'$ . On se sert de ceci pour redéfinir en général la notion de fonction. Dans ce cadre généraliste, on parle plutôt d'«application», mais c'est une simple question de vocabulaire.

### Définition

Soient  $x$  et  $y$  deux ensembles. On dit qu'une partie  $f \subset x \times y$  est une **application de  $x$  dans  $y$**  (ou une fonction définie sur  $x$ , à valeurs dans  $y$ , ou encore famille d'éléments de  $y$  indexée par  $x$ ) et on note  $f : x \rightarrow y$  si la relation d'appartenance à  $f$  est une relation fonctionnelle à un argument. Autrement dit,  $f : x \rightarrow y$  si on a

$$\begin{cases} f \in \mathcal{P}(x \times y), \\ \forall u \forall v \forall v' [(u, v) \in f \wedge (u, v') \in f \Rightarrow v = v'], \\ \forall u [u \in x \Rightarrow \exists v (v \in y \wedge (u, v) \in f)]. \end{cases}$$

La dernière ligne se réécrit usuellement en  $v = f(u)$ .

### Remarques.

- ▷ Si  $A(f)$  désigne l'énoncé « $f$  est une application de  $x$  dans  $y$ », alors puisque toute application  $x \rightarrow y$  est une partie de  $x \cup y$ , on a

$$\forall f [A(f) \Leftrightarrow (A(f) \wedge f \in \mathcal{P}(x \cup y))]$$

et donc, par le schéma d'axiomes de compréhension, les application de  $x$  dans  $y$  forment un ensemble, que l'on note  $y^x$ .

On peut maintenant définir la réunion, l'intersection et le produit d'une famille (presque) quelconque d'ensembles. Soit  $R(u, v)$  un énoncé fonctionnel à deux arguments et sans paramètre, dont le domaine défini par l'énoncé  $\exists v R(u, v)$  est un ensemble  $I$  : on parle alors de **famille d'ensembles indexée par  $I$** . Par le schéma d'axiomes de compréhension, ceci est la même donnée que celle d'une application  $x : I \rightarrow X$  de  $I$  dans un certain ensemble  $X$ . Si l'on note  $x_i := x(i) \in X$  (i.e.  $(i, x_i) \in x$ ), on retrouve la définition usuelle d'une famille d'ensemble  $(x_i)_{i \in I}$  indexée par  $I$ .

On appelle **réunion** de la famille  $(x_i)_{i \in I}$  la réunion de l'image de  $x$ , autrement dit, c'est un ensemble  $y$  tel que

$$\forall u [u \in y \Leftrightarrow \exists i (i \in I \wedge u \in x_i)].$$

Cet ensemble est noté  $\bigcup_{i \in I} x_i$ .

L'**intersection** de la famille  $(x_i)_{i \in I}$  est la collection définie par l'énoncé

$$N(u) \stackrel{\text{def}}{\Leftrightarrow} \forall i [i \in I \Rightarrow u \in x_i].$$

Si  $I = \emptyset$ , alors  $N$  définie la collection  $\mathcal{U}$ , donc n'est pas un ensemble. Si  $I \neq \emptyset$ , soit  $i_0 \in I$  et alors

$$N(u) \Leftrightarrow (u \in x_{i_0} \wedge \forall i [i \in I \Rightarrow u \in x_i])$$

et donc la collection  $N$  est un ensemble, d'après le schéma d'axiomes de compréhension, que l'on note  $\bigcap_{i \in I} x_i$ .

Enfin, on définit le **produit** de la famille  $(x_i)_{i \in I}$ . On considère la collection  $A$  des application  $f : I \rightarrow \bigcup_{i \in I} x_i$  telles que  $f(i) \in x_i$  pour tout  $i \in I$ . Une telle application est élément de  $\left(\bigcup_{i \in I} x_i\right)^I$ , donc

$$\forall f \left[ A(f) \Leftrightarrow \left( A(f) \wedge f \in \left(\bigcup_{i \in I} x_i\right)^I \right) \right]$$

et donc la collection  $A$  est un ensemble par le schéma d'axiomes de compréhension. Cet ensemble est le produit de la famille  $(x_i)_{i \in I}$ , que l'on note  $\prod_{i \in I} x_i$ .

## II. 4 Entiers naturels, axiome de l'infini et arithmétique de Peano

La définition et l'étude des entiers naturels d'un point de vue ensembliste requiert une notion un peu subtile : la notion d'*ordinal* et de *cardinal*, qui sont des types d'ensembles particuliers. Cependant, nous n'entrerons pas dans ces considérations ici et nous contenterons de définir les entiers sans «voir» leur structure d'ordinaux.

Une notion important dégagée par Peano est celle de **successeur**  $S(n)$  d'un entier  $n$ . Bien-sûr, intuitivement on a que  $S(n) = n + 1$  est le premier entier après  $n$ , d'où le nom de successeur. Il nous faut formaliser un peu ceci.

Il s'agit donc de définir l'entier 0 et, étant donné un entier, on doit pouvoir définir son successeur, ce qui doit se faire avec une procédure générale utilisant uniquement les règles de construction déjà introduites plus haut.

Quel candidat donc pour 0? Pour le deviner, on peut raisonner en terme de «cardinal», au sens intuitif (combien y a-t-il d'éléments dans un ensemble fini). L'entier 0 devrait être de cardinal zéro, i.e. ne doit pas contenir d'élément. Le bon ensemble à considérer est donc l'ensemble vide  $\emptyset$ . On pose alors  $0 := \emptyset$ . Ensuite, toujours avec les cardinaux, le successeur d'un entier doit contenir un élément de plus que l'entier considéré. Par exemple, l'entier 1 doit être un ensemble à un élément et on peut par exemple considérer le singleton  $1 := \{\emptyset\}$ , ensemble dont le seul élément est l'ensemble vide. Maintenant, l'entier 2 doit avoir deux éléments. L'ensemble des parties de  $1 = \{\emptyset\}$  est un tel ensemble :  $2 := \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$  et on remarque que  $\mathcal{P}(\{\emptyset\}) = \{\emptyset\} \cup \{\{\emptyset\}\} = 1 \cup \{1\}$ . On note également que  $1 = \{\emptyset\} = 0 \cup \{0\}$ . et on commence déjà à deviner le schéma général. Définissons encore l'entier 3 : il doit avoir un élément de plus que 2 et un bon candidat est  $3 := \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = 2 \cup \{2\}$ . On en arrive à la définition suivante :

### Définition

Étant donné un ensemble  $x$ , on définit son successeur  $S(x)$  par

$$S(x) := x \cup \{x\}.$$

On appelle **zéro** et on note 0 l'ensemble vide :

$$0 := \emptyset.$$

On définit les entiers naturels par successeurs :  $1 := S(0)$ ,  $2 := S(1)$ ,  $3 := S(2)$ ,  $4 := S(3)$ , etc.

On ne peut pas déduire des axiomes déjà introduits que la collection des entiers naturels est un ensemble (cet énoncé est dit *indépendant* des axiomes introduits). Il nous faut donc ajouter cet axiome.

### Axiome 5 : Axiome de l'infini

a collection des entiers naturels est un ensemble ; autrement dit

$$\exists x[0 \in x \wedge \forall y(y \in x \Rightarrow S(y) \in x)].$$

Cet ensemble est noté  $\mathbb{N}$ .



## Remarques.

- ▷ L'entier 0 n'est successeur d'aucun ensemble, en effet, si un tel ensemble  $u$  existait, alors on aurait  $\emptyset = S(u) = u \cup \{u\}$  et on aurait alors  $u \in \emptyset$  : impossible.
- ▷ L'ensemble  $\mathbb{N}$  n'est successeur d'aucun entier.
- ▷ En utilisant la théorie des ordinaux, on peut montrer que deux entiers naturels ayant même successeur sont égaux.
- ▷ On a une relation d'ordre naturelle sur les entiers induite par la relation  $n < S(n)$ .
- ▷ Utilisant toujours les ordinaux, on montre que si une propriété  $P$  définie sur  $\mathbb{N}$  est vraie pour 0 et si  $P(n) \Rightarrow P(n+1)$  pour tout entier naturel  $n \in \mathbb{N}$ , alors  $P(n)$  est vraie pour tout  $n \in \mathbb{N}$  : c'est le **principe de récurrence**.

Rappelons que les axiomes de l'arithmétique de Peano (1889) sont les suivants :

1. L'élément appelé zéro et noté 0 est un entier naturel.
2. Tout entier naturel admet un successeur.
3. Aucun entier n'admet zéro pour successeur.
4. Deux entiers ayant le même successeur sont égaux.
5. Si un ensemble d'entiers naturels contient 0 et le successeur de chacun de ses éléments, alors cet ensemble contient  $\mathbb{N}$ .

D'après les remarques précédentes, ces propriétés sont bien vérifiées par les constructions précédentes et on retrouve bien l'arithmétique usuelle.

**Remarque.** Pour définir rigoureusement l'addition et la multiplication des entiers naturels, il faut faire appel à la notion supplémentaire de *cardinal*, définie à partir des ordinaux. Cette difficulté mise à part, on peut définir intuitivement la somme et le produit de deux entiers  $x, y \in \mathbb{N}$  (vus comme des ensembles!) par

$$x + y := (x \times \{0\}) \cup (y \times \{1\}) \quad \text{et} \quad xy := x \times y.$$

On peut ensuite montrer que ces opérations vérifient bien les propriétés usuelles de l'addition et de la multiplication<sup>4</sup>.

**La théorie de Zermelo-Fraenkel (ZF) est la donnée de l'univers (non vide)  $\mathcal{U}$  et de la relation d'appartenance  $\in$ , sujets aux (schéma d') axiomes 1,2,3,4,5 (extensionnalité, réunion, parties, schéma de substitution, infini).**

## II. 5 L'axiome du choix

Il s'agit d'un axiome qui a longtemps fait (et fait encore parfois) débat dans la communauté mathématique. D'un côté, il est très intuitif et permet de démontrer l'existence de beaucoup d'objets mathématiques, impossible à prouver autrement ; d'un autre côté, il a des implications très étranges qui peuvent être considérées comme des paradoxes.

### Axiome 6 : Axiome du choix

Pour tout ensemble non vide  $x$ , il existe une application  $f : \mathcal{P}(x) \setminus \{\emptyset\} \rightarrow x$  telle que, pour toute partie non vide  $y$  de  $x$ , on a  $f(y) \in y$ .

**La théorie de Zermelo-Fraenkel, à laquelle on ajoute l'axiome du choix, est appelée la théorie ZFC.** C'est la théorie couramment employée.

## Remarques.

- ▷ Autrement dit, dans toute partie non vide  $y$  de  $x$ , on peut *choisir* un élément  $f(y)$  ; c'est très intuitif mais la subtilité réside dans le fait que ceci doit aussi être possible pour des ensembles infinis.
- ▷ Une formulation équivalente est la suivante : le produit d'une famille d'ensembles non vides est non vide. Là encore c'est intuitif, mais c'est le cas d'un produit infini d'ensembles qui est subtil.

4. pour être précis : l'addition et la multiplication sont associatives, commutatives, la multiplication est distributive sur l'addition et de plus,  $x + 0 = x$  et  $x \times 1 = x$ .



- ▷ Bertrand Russell illustre cet axiome avec la phrase suivante : "Pour sélectionner une chaussette plutôt qu'une autre dans chaque paire d'un ensemble infini de paires, j'ai besoin de l'axiome du choix. Mais pour les chaussures, je n'en ai pas besoin." Explications : les chaussettes d'une paire sont indiscernables l'une de l'autre, il me faut donc faire un choix jour après jour ; je ne peux pas faire de procédure générale pour sélectionner mes chaussettes. Pour les chaussures c'est différent : il y a les chaussures de droite et les chaussures de gauche et je peux les discerner. Je n'ai donc pas besoin de «choisir» les chaussures, je peux par exemple ne garder que les chaussures de droite.

Un exemple célèbre de paradoxe lié à l'axiome du choix est le paradoxe de Banach-Tarski :

#### **Théorème : Banach–Tarski, 1924**

Il est possible de découper une boule de volume 1 dans  $\mathbb{R}^3$  en cinq parties (non «mesurables»), telles que l'on puisse réarranger ces parties pour former deux boules de volume 1.

On vient littéralement de dédoubler la boule unité<sup>5</sup> ! Bien que très contre-intuitif et reposant de façon cruciale sur l'axiome du choix, les cinq parties de la boule de départ sont très bizarres : lorsque l'on dit qu'elles ne sont pas «mesurables», cela veut dire que leur volume n'est pas un nombre bien défini<sup>6</sup>. En particulier, rassurez-vous : impossible de dédoubler un gâteau ; les parties considérées ici ne sont pas physiquement constructibles. On peut donc mettre ce résultat étrange sur le dos de l'infini<sup>7</sup>, dont l'usage mène déjà à des résultats contre-intuitifs sans faire appel à l'axiome du choix (exemple, l'hôtel de Hilbert).

Pour les conséquences heureuses maintenant ! À l'aide de l'axiome du choix, Peano a montré en 1890 qu'il existait toujours des solutions aux équations différentielles<sup>8</sup>. On peut également déduire de cet axiome que tout espace vectoriel (de dimension finie ou non) admet une base. On s'en sert aussi pour montrer qu'étant donnée une famille de polynômes, on peut construire un ensemble qui en contient toutes les racines. On peut enfin en tirer qu'étant donnés deux ensembles, on peut toujours comparer leur cardinaux (soit l'un à moins d'éléments que l'autre, soit l'inverse, même pour les ensembles infinis).

Cet axiome est donc un peu déroutant et il est difficile de décider s'il faut l'accepter ou non. Le résultat suivant, dit de **consistance relative**, nous dit que, que l'on accepte ou non l'axiome du choix, cela n'apportera pas de nouvelle incohérence dans la théorie : on peut l'affirmer ou le nier sans faire tomber la théorie.

#### **Théorème : Consistance relative de l'axiome du choix et de sa négation**

L'axiome du choix et sa négation sont relativement consistants avec ZF. En d'autres termes, s'il y a une incohérence dans ZFC, alors il y en a aussi une dans ZF et, s'il y en a une dans ZF avec négation de l'axiome du choix, alors il y en a aussi une dans ZF.

## **II. 6 Quelques axiomes supplémentaires et remarques**

Un axiome que l'on trouve parfois inclus dans les axiomes de ZF suivant les auteurs est le suivant :

#### **Axiome 7 : Axiome de fondation**

Tout ensemble non vide a un élément qui n'a aucun élément commun avec cet ensemble, en d'autres termes

$$\forall x[x \neq \emptyset \Rightarrow \exists y(y \in x \wedge y \cap x = \emptyset)].$$

#### **Remarques.**

- ▷ Conséquence : aucun ensemble n'est élément de lui-même, ce qui est intuitif. Ainsi, la collection des ensembles  $x$  tels que  $x \notin x$  est donc tout l'univers  $\mathcal{U}$ , qui n'est pas un ensemble.
- ▷ Tout comme l'axiome du choix, on montre que cet axiome et sa négation sont relativement consistants avec ZF.

5. voir à ce sujet l'excellente vidéo : <https://www.youtube.com/watch?v=fzyd02CXf-I>

6. en fait, on ne peut montrer l'existence de parties non mesurables qu'avec l'axiome du choix

7. comme en physique quantique, lorsque l'on parle de l'infini, il faut se méfier sérieusement de son intuition !

8. plus précisément, aux équations explicites et continues

Rappelons un peu de vocabulaire sur les applications :

### Définition

Soient  $x, y$  deux ensembles et  $f : x \rightarrow y$  une application. On dit que  $f$  est

- **injective** si tout élément de  $y$  a au-plus un antécédent par  $f$  :

$$\forall u \forall u' [u \in x \wedge u' \in x \wedge \exists v (v \in y \wedge (u, v) \in f \wedge (u', v) \in f) \Rightarrow u = u'].$$

- **surjective** si tout élément de  $y$  a au-moins un antécédent par  $f$  :

$$\forall v [v \in y \Rightarrow \exists u (u \in x \wedge (u, v) \in f)].$$

- **bijective** si  $f$  est à la fois injective et surjective, i.e. tout élément de  $y$  admet exactement un antécédent par  $f$  :

$$\forall v [v \in y \Rightarrow \exists! u (u \in x \wedge (u, v) \in f)].$$

On peut réécrire respectivement ces propriétés à l'aide de la notation  $v = f(u)$  au-lieu de  $(u, v) \in f$  :

$$\begin{cases} f \text{ injective} & \iff \forall u, u' \in x, f(u) = f(u') \Rightarrow u = u', \\ f \text{ surjective} & \iff \forall v \in y, \exists u \in x ; f(u) = v, \\ f \text{ bijective} & \iff \forall v \in y, \exists! u \in x ; f(u) = v. \end{cases}$$

À partir de ces définitions, on peut former un nouvel axiome, appelé **hypothèse du continu**. Définissons l'ensemble des **réels** par  $\mathbb{R} := \mathcal{P}(\mathbb{N})$ .

### Axiome 8 : Hypothèse du continu

Si  $x$  est un ensemble et si on a des injections  $f : \mathbb{N} \rightarrow x$  et  $g : x \rightarrow \mathbb{R}$ , alors ou bien  $f$  ou bien  $g$  est en fait une bijection.

Cet axiome signifie qu'il n'existe pas d'ensemble *strictement* compris entre les entiers naturels et les réels, en ce sens que tout tel ensemble est en fait en bijection avec  $\mathbb{N}$  ou avec  $\mathbb{R}$  : il a soit autant d'éléments que d'entiers, soit autant d'éléments que les réels. Ceci semble un peu étrange mais quand on y pense, tous les sous-ensembles des réels que l'on connaît (entiers relatifs, nombres décimaux, rationnels, réels positifs, réels non nuls, irrationnels...) sont en bijection soit avec  $\mathbb{N}$  (les trois premiers) soit avec  $\mathbb{R}$  (les trois derniers).

Tout comme l'axiome du choix et l'axiome de fondation, cet axiome et sa négation sont relativement consistants avec ZF. On peut démontrer que tout énoncé «arithmétique» (grossièrement, ne portant que sur les entiers) qui serait démontrable à l'aide de l'axiome du choix et/ou de l'hypothèse du continu, est en fait démontrable sans ces axiomes.

Pour finir, signalons deux résultats utiles (mais pas pour le programme) :

### Théorème : Cantor

Pour tout ensemble  $x$ , il n'existe aucune injection  $\mathcal{P}(x) \rightarrow x$  ni aucune surjection  $x \rightarrow \mathcal{P}(x)$ .

### Théorème : Cantor–Bernstein

Si  $x$  et  $y$  sont deux ensembles tels qu'il existe une injection  $x \rightarrow y$  et une injection  $y \rightarrow x$ , alors il existe une bijection  $x \rightarrow y$  (et donc aussi sa bijection réciproque  $y \rightarrow x$ ).

## III Et après ?

Une fois qu'on a construit les entiers naturels, les opérations  $+$  et  $\times$  et admis l'existence (via l'axiome de l'infini) de l'ensemble  $\mathbb{N}$  des entiers naturels, on peut construire les ensembles «plus gros» que l'on connaît. Ces constructions font appel à de l'*algèbre*<sup>9</sup> : l'étude des structures telles que  $+$  et  $\times$  et, plus

9. Un de mes enseignants définit informellement l'algèbre comme étant "l'étude de liens abstraits entre objets concrets"

tard, à la *topologie*, qui est l'étude des notions de limite, de voisinage, de distances... Bien que de différente saveur, ces notions se formulent uniquement en termes d'ensembles.

Signalons, sans être trop précis, les idées qui entrent en jeu dans la construction de ces ensembles.

Une fois qu'on a les entiers naturels, on cherche à construire les **entiers relatifs** : on veut ajouter des signes devant les entiers. Étant donné un entier naturel  $n \in \mathbb{N}$ , l'entier relatif  $-n$  est caractérisé par la propriété  $(-n) + n = n + (-n) = 0$  : l'entier  $-n$  est l'*opposé* de l'entier  $n$ . On peut construire alors rigoureusement l'ensemble  $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$  des entiers et de leurs opposés. Notons que  $-0 = 0$  et on peut alors énumérer  $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$ . On dit que le triplet  $(\mathbb{Z}, +, \times)$  est un *anneau* : une riche structure algébrique.

Maintenant que l'on sait ce qu'est un entier négatif, il faut savoir *diviser* par un entier non nul. Étant donné  $n \in \mathbb{Z} \setminus \{0\}$ , la fraction  $1/n$  est caractérisée par la relation  $(1/n) \times n = n \times (1/n) = 1$ . Avec cette propriété, on voit qu'on a par exemple  $1/(-n) = -1/n$  et on peut former toutes les fractions  $m/n := m \times (1/n)$  pour  $m \in \mathbb{Z}$ . Cet ensemble de fractions (que l'on peut là encore construire explicitement) est l'ensemble des rationnels  $\mathbb{Q}$  et on peut montrer qu'il est en bijection avec  $\mathbb{N}$ . On peut de plus voir  $\mathbb{Z}$  comme un sous-ensemble de  $\mathbb{Q}$  en voyant un entier  $m \in \mathbb{Z}$  comme la fraction  $m/1 \in \mathbb{Q}$ . On dit alors que le triplet  $(\mathbb{Q}, +, \times)$  est un *corps* : une structure algébrique encore plus riche que celle d'anneau.

Nous avons maintenant l'ensemble  $\mathbb{Q}$  des rationnels. Cet ensemble est très confortable pour faire de l'algèbre, mais beaucoup moins pour faire de l'analyse : impossible de définir la notion de dérivée ou d'intégrale et donc impossible de donner un sens aux calculs de physique. Le problème est qu'il existe des suites de rationnels dont les termes se rapprochent autant que l'on veut<sup>10</sup>, mais qui ne convergent pas vers un rationnel. Autrement dit, il y a des «trous» dans les rationnels ; on ne peut pas passer continûment d'un rationnel à l'autre. Pour résoudre ceci, on construit l'ensemble  $\mathbb{R}$  des **réels** comme l'ensemble des limites de suites de rationnels dont les termes se rapprochent autant que l'on veut et là, il n'y a plus de trou. On a donc construit  $\mathbb{R}$  à l'aide de notions de topologie (ou, si l'on préfère, d'analyse), mais on peut montrer que cette définition coïncide avec celle donnée plus haut : en effet, il existe une bijection entre l'ensemble  $\mathbb{R}$  tel que construit ici et l'ensemble  $\mathcal{P}(\mathbb{N})$  des parties de  $\mathbb{N}$ . Il est à noter que l'on peut également construire  $\mathbb{R}$  à l'aide de la notion de *coupure de Dedekind*, qui ne fait appel qu'aux ensembles. L'avantage de la construction dont nous parlons ici est qu'elle permet d'étendre facilement les opérations  $+$  et  $\times$  à  $\mathbb{R}$  et de prouver que  $(\mathbb{R}, +, \times)$  est aussi un corps. L'avantage de la méthode de Dedekind est qu'elle permet de montrer facilement que  $\mathbb{R}$  a la *propriété de la borne supérieure*. Les deux constructions sont néanmoins équivalentes, bien-entendu.

Dernier problème : il existe dans  $\mathbb{R}$  des équations polynomiales qui n'ont pas de solution, par exemple l'équation  $a^2 + 1 = 0$  n'a pas de solution dans  $\mathbb{R}$ . On a donc une nouvelle fois recours à l'algèbre en construisant (de manière ensembliste à nouveau) un ensemble  $\mathbb{C}$ , qui est simplement  $\mathbb{R}$  auquel on a «ajouté» un élément  $i$  tel que  $i^2 + 1 = 0$ . On montre ensuite que ce nouvel ensemble  $\mathbb{C}$  des **nombres complexes** contient en fait toutes les solutions de toutes les équations polynomiales réelles et que l'on peut étendre les opérations  $+$  et  $\times$  à  $\mathbb{C}$ , de telle sorte que  $(\mathbb{C}, +, \times)$  soit à nouveau un corps, donc chaque élément  $z$  s'écrit  $z = a + bi$  avec  $a, b \in \mathbb{R}$ . Il se trouve que toutes les équations polynomiales à coefficients dans  $\mathbb{C}$  ont aussi toutes leurs solutions dans  $\mathbb{C}$ , donc inutile de réitérer la procédure : en ajoutant à  $\mathbb{C}$  les racines de ses polynômes, on n'ajoute en fait aucun élément : elles sont toutes déjà dans  $\mathbb{C}$ . On peut enfin montrer que cet ensemble est en bijection avec  $\mathbb{R}$ .

Ainsi, on peut en effet construire tous les sur-ensembles courants de  $\mathbb{N}$  en n'utilisant que les notions et axiomes de ZF, mais en introduisant de nouveaux concepts, afin d'alléger les notations, de faciliter l'intuition et d'utiliser à bon escient les structures portées par les ensembles.

Nous avons là en quelque sorte l'essence des mathématiques : à partir des règles de logique usuelles et d'un nombre le plus restreint possible d'axiomes raisonnables, construire ou prouver l'existence de nouveaux objets, s'habituer à les manipuler, en dégager les structures et les propriétés et ainsi, établir des théories les plus abouties et élégantes possible. Un but possible, mais non nécessaire, étant de fournir un outil solide et fiable aux Sciences.

---

10. de telles suites sont dites *de Cauchy*