La naissance de la théorie des ensembles

La théorie des ensembles a été initiée au début du XX$^{\text{ème}}$ siècle sous l'impulsion de Georg Cantor. Elle fût contestée à ses débuts à cause de sa complexité et aussi parce qu'elle était source de paradoxes. Néanmoins, elle connut un formidable essor car elle permit de construire les objets mathématiques fondamentaux comme les nombres ou les fonctions.

Les infinis de Cantor

Lorsque Georg Cantor (1845-1918) étudie la convergence des séries de Fourier, il est amené à préciser et à approfondir la notion d'ensemble : il va poser les bases de la théorie des ensembles que nous connaissons. Au début, Cantor découvre l'existence de plusieurs infinis :

- le dénombrable qui réunit les ensembles qui sont équipotents (i.e. en bijection) avec l'ensemble des entiers naturels $\mathbb{N}$. Cantor démontre notamment que l'ensemble des nombres rationnels $\mathbb{Q}$ est dénombrable.

- le continu qui réunit les ensembles équipotents avec la droite réelle $\mathbb{R}$. Cantor observe en particulier que l'ensemble $\mathcal{P}(\mathbb{N})$ des parties de $\mathbb{N}$ est continu.

Cantor, sans doute mal inspiré par une interprétation géométrique, présume que l'ensemble $\mathbb{R}^2$ des couples de réels est un infini strictement plus grand que celui de $\mathbb{R}$. A son grand étonnement, il démontre en fait que l'ensemble $\mathbb{R}^2$ est équipotent à $\mathbb{R}$. A Dedekind, à qui il demande confirmation de ses travaux, il écrira : « *Je le vois mais je ne le crois pas.* »

Cantor hiérarchise ces infinis en introduisant la notion de *cardinal transfini* d'un ensemble infini. Il met aussi en place l'arithmétique cardinale afin d'opérer sur ses différents infinis. Il observe que si E est un ensemble infini, l'ensemble $\mathcal{P}(E)$ de ses parties est un infini strictement plus grand, c'est le théorème de Cantor :

$$\operatorname{Card}\mathcal{P}(E) = 2^{\operatorname{Card}E} > \operatorname{Card}E$$

Le cardinal de l'ensemble $\mathbb{N}$ (et de tout ensemble dénombrable) est noté $\aleph_0$ (la lettre $\aleph$ se lit *aleph*, c'est la première lettre de l'alphabet hébraïque). Le cardinal $2^{\aleph_0}$ de l'ensemble $\mathbb{R}$ est noté c (pour *continuum.*)

La théorie des ensembles de Cantor fait malheureusement apparaître un paradoxe :

> Si l'on considère l'ensemble E de tous les ensembles, celui-ci contient chacune de ses parties et est donc au moins aussi grand que $\mathcal{P}(E)$!

Cette théorie est alors très controversée, certains n'y voit que du « *brouillard dans le brouillard*. »

Le paradoxe de Russell

Le logicien Bertrand Russell (1872-1970) intervient dans le débat en introduisant le paradoxe suivant :

> Considérons A l'ensemble des ensembles qui n'appartiennent pas à eux-mêmes, autrement dit : $A = \{X \mid X \notin X\}$. Posons le problème : A s'appartient-il ou non ?
> Si $A \in A$ alors $A \notin A$, et si $A \notin A$ alors $A \in A$. Dans les deux cas il y a contradiction.

Par ce paradoxe, Russell simplifie la problématique et précise ce qui ne va pas dans la théorie de Cantor : la notion d'ensemble n'est pas correctement posée. Il est vrai que jusqu'alors celle-ci était restée relativement imprécise :

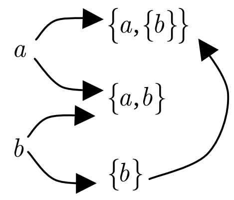
> « Un ensemble est une réunion d'objets, bien déterminés, que nous appelons éléments de cet ensemble. »

Le paradoxe de Russell est à l'origine d'une crise profonde dans l'histoire des mathématiques. Ce n'est pas la première. La découverte de l'incommensurabilité du nombre $\sqrt{2}$ par les Pythagoriciens a entraîné une redéfinition de la notion de nombre. Les paradoxes de Zénon d'Elée ont forcé les mathématiciens à préciser les concepts d'infini, de limite et de continuité. Le paradoxe de Russell nécessite une redéfinition du concept d'ensemble : il ne sera désormais plus possible de parler de « l'ensemble de tous les ensembles. »

Russell, et son professeur Whitehead, essaient dès lors de construire une axiomatique empêchant l'apparition du paradoxe. Celle-ci ne sera jamais considérée comme entièrement satisfaisante. Une solution sera apportée par Ernst Zermelo (1871-1953) et améliorée par Abraham Fraenkel (1891-1965) au début des années 1920. Celle-ci est désormais connue sous le nom de l'axiomatique de Zermelo et Fraenkel, ZF en abrégé. Donnons-en une présentation telle qu'elle est manipulée de nos jours.

Axiomatique ZF

On suppose disposer d'un univers $\mathcal{U}$ dont les objets sont appelés ensembles. Certains de ces objets sont reliés à d'autres par une relation appelée relation d'appartenance : on note $x \in y$ pour signifier que l'ensemble x appartient à l'ensemble y. Soulignons au passage, que nous ne distinguons pas éléments et ensembles, en effet un ensemble peut se voir comme étant un élément d'un ensemble d'ensembles, aussi tous les objets de notre univers peuvent être appelés ensembles et, généralement, ils sont désignés par une lettre minuscule. L'univers $\mathcal{U}$ peut être visualisé comme un graphe orienté où les sommets seraient les ensembles et où les flèches visualiseraient les relations d'appartenance.



Nous supposons que notre univers, et sa relation d'appartenance, vérifient les axiomes suivants :

- *Axiome d'extensionnalité :* Deux ensembles constitués des mêmes éléments sont égaux.

Par suite, il est fréquent d'établir l'égalité de deux ensembles en constatant que ceux-ci sont formés des mêmes éléments, par exemple en raisonnant par double inclusion.

- *Axiome de la paire :* Etant donnés deux ensembles a et b, il existe un ensemble dont les éléments sont a et b, et eux seuls. Par extensionnalité, cet ensemble est unique, on le note $\{a,b\}$.

Cet axiome permet bien entendu de définir les ensembles à deux éléments mais aussi les singletons qui sont obtenus quand $a=b$. Cet axiome permet de surcroît de définir les couples (a,b) comme correspondant aux objets $\{\{a\},\{a,b\}\}$. A titre d'exercice, on pourra vérifier l'implication : $(a,b)=(a',b') \Rightarrow (a=a' \text{ et } b=b')$

- *Axiome de la réunion :* Si l'on se donne un ensemble a, il existe un ensemble dont les éléments sont les éléments de a. Par extensionnalité, celui-ci est unique, on le note $\bigcup_{x\in a} x$.

A partir de deux ensembles a,b, il est désormais possible de définir leur union $a\cup b$, comme étant l'ensemble des éléments des ensembles de $\{a,b\}$. Cela permet de définir ensuite l'ensemble $\{a,b,c\}$ comme étant l'ensemble $\{a,b\}\cup\{c\}$ et plus généralement par récurrence : $\{a_1,\dots,a_n\}=\{a_1,\dots,a_{n-1}\}\cup\{a_n\}$.

- *Schéma de compréhension :* Pour tout ensemble a et tout énoncé $P(x)$ dont la valeur de vérité dépend d'une variable x, il existe un ensemble dont les éléments sont les éléments de a qui vérifient l'énoncé $P(x)$. Encore une fois, par extensionnalité, un tel ensemble est unique, on le note $\{x\in a / P(x)\}$.

Ici on ne parle plus d'un axiome, mais d'un schéma d'axiomes, car il forme, en fonction de l'énoncé $P(x)$, une infinité d'axiomes. C'est ce schéma d'axiomes qui permet définir les parties d'un ensemble formés des éléments vérifiant une propriété donnée. Par exemple, à partir d'ensembles a et b, on peut définir les ensembles $a\cap b=\{x\in a / x\in b\}$ et $a\setminus b=\{x\in a / x\notin b\}$. On peut aussi définir l'ensemble vide par : $\varnothing=\{x\in a / x\neq x\}$, celui-ci est unique, par extensionnalité.

Notons que cet axiome implique l'inexistence de l'ensemble des ensembles. En effet si ce dernier existait, notons-le a et introduisons l'ensemble $b=\{x\in a / x\notin x\}$. On ne peut avoir $b\in b$ car cela implique $b\notin b$, ni

$b \notin b$ car alors $b \in b$. C'est absurde. On retrouve ici l'idée du paradoxe de Russell, mais, cette fois-ci, celui-ci ne sert plus à contredire la théorie mais à établir un résultat : l'ensemble de tous les ensembles n'existe pas. Quel est alors l'objet constitué de tous les ensembles ? C'est l'univers $\mathcal{U}$ introduit au départ et duquel nous sommes en train de définir les propriétés.

- *Axiome de l'ensemble des parties :* Pour tout ensemble a, il existe un ensemble dont les éléments sont les parties de a. Par extensionnalité, cet ensemble est unique, on le note $\mathcal{P}(a)$.

En exploitant les axiomes qui précèdent, on peut définir le produit cartésien de deux ensembles a et b comme étant le sous-ensemble de $\mathcal{P}(\mathcal{P}(a \cup b))$ formé des éléments qui sont de la forme (x, y) i.e. $\{\{x\}, \{x, y\}\}$ avec $x \in a$ et $y \in b$.

- *Schéma d'axiomes de substitution :* C'est le plus compliqué de tous. Pour le présenter, considérons un énoncé fonctionnel $E(x, y)$, c'est à dire un énoncé tel que pour tout ensemble x, il existe au plus un ensemble y tel que l'énoncé $E(x, y)$ soit vrai. Par exemple, $E(x, y)$ peut être l'énoncé « $y = \{x\}$ » ou encore, en introduisant un ensemble z, l'énoncé « $y \subset z$ et $z \setminus y = x$ ». Le schéma d'axiomes de substitution exige que pour tout ensemble a, il existe un ensemble b dont les éléments sont les y vérifiant l'énoncé $E(x, y)$ pour au moins un $x \in a$.

D'un certain sens, un énoncé fonctionnel $E(x, y)$ permet de définir une « fonction » qui à certains ensembles x associe un ensemble y. Soulignons que cette fonction n'est pas nécessairement définie pour tout ensemble x. Le schéma d'axiomes de substitution permet alors de considérer l'ensemble des valeurs prises par cette fonction sur un ensemble a.

En considérant un énoncé $P(x)$ et en définissant $E(x, y)$ comme étant l'énoncé « $x = y$ et $P(x)$ », on observe que le schéma d'axiomes de substitution implique le schéma de compréhension. Ce dernier axiome est donc redondant et pourrait être supprimé de notre axiomatique sans que cela ait de conséquences. De même, on peut montrer que l'axiome de la paire est redondant.

A partir des notions qui précèdent, on peut définir les entiers naturels :

on pose $0 = \varnothing$, $1 = \{\varnothing\}$, $2 = 1 \cup \{1\} = \{\varnothing, \{\varnothing\}\}$ et plus généralement : $n + 1 = n \cup \{n\}$.

La relation d'ordre naturelle sur les entiers se confond alors avec l'inclusion des ensembles.

- *Axiome de l'infini :* Il existe un ensemble dont les entiers naturels sont éléments.

Cet axiome permet de construire l'ensemble $\mathbb{N}$ des entiers naturels et il est alors possible de construire les ensembles numériques classiques : $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$. Nous disposons ainsi d'une base solide pour faire des mathématiques.

Une base solide ?

Cependant il reste deux problèmes à résoudre :

- Notre axiomatique n'est-elle pas contradictoire ?

- Un tel univers $\mathcal{U}$ existe-t-il ?

Bien entendu, une réponse positive au premier problème résout immédiatement le second !

Commençons par étudier le problème de non-contradiction (on parle de *consistance*) : on veut que les axiomes précédemment cités n'impliquent pas qu'un énoncé soit vrai ainsi que sa négation. En effet, si tel était le cas, puisque la logique de base veut que le faux implique à la fois le vrai et le faux, on parvient à démontrer que tout énoncé est à la fois vrai et faux : l'édifice s'écroule. Pour l'instant il tient encore...

L'idéal serait que les axiomes précédemment cités impliquent l'inexistence de formules contradictoires. Les travaux de Kurt Gödel (1906-1978) nous obligent à abandonner cet idéal. En effet, il établit que pour toute axiomatique permettant de définir au moins l'ensemble $\mathbb{N}$, il existe des énoncés *indécidables*, c'est à dire des énoncés dont il est impossible de démontrer la véracité ou la fausseté. Plus précisément, Gödel montre que si l'axiomatique n'est pas contradictoire il existe des énoncés vrais impossibles à démontrer. Ce résultat est connu sous le nom de premier théorème d'incomplétude de Gödel, il fût publié en 1930. L'année suivante, Gödel établit que parmi ces indécidables figure l'énoncé exprimant la consistance de cet axiomatique, c'est le deuxième théorème d'incomplétude de Gödel. Il faut donc abandonner tout espoir d'une axiomatique qui permettrait de construire $\mathbb{N}$ et justifierait aussi sa consistance.

Evoquons maintenant le problème de l'existence de l'univers $\mathcal{U}$. Commençons par un parallèle plus simple : les nombres réels. On peut introduire l'ensemble $\mathbb{R}$ comme étant un corps totalement ordonné et vérifiant la propriété de la borne supérieure. On présente ainsi l'ensemble $\mathbb{R}$ par ses axiomes de définition, cela ne signifie par pour autant que cet objet existe. Pour assurer l'existence de celui-ci, il conviendrait de déterminer un modèle de la droite réelle, c'est à dire une construction d'un ensemble qui satisferait aux propriétés précédentes. Ce problème a été résolu dans la deuxième partie du XIX$^{\text{ème}}$ siècle et on connaît plusieurs démarches qui, à partir de l'ensemble $\mathbb{Q}$, permettent de construire un modèle de la droite réelle. Soulignons au passage que l'obtention de ces constructions a aussi permis de mieux comprendre ce qu'était la droite réelle.

En ce qui concerne l'univers $\mathcal{U}$, en prouver l'existence en en introduisant un modèle, revient à introduire un objet muni de sa propre axiomatique permettant de construire un modèle de $\mathcal{U}$. Cet objet sera alors confronté aux mêmes problèmes que $\mathcal{U}$: incomplétudes et existence présupposée.

L'axiome du choix

De nos jours, plus personne ne discute l'axiomatique ZF présentée ci-dessus. En revanche, on y adjoint très souvent un axiome qui reste encore source de polémiques : l'axiome du choix. Il existe de multiples manières de le présenter, l'une des plus simples est peut-être la suivante :

- *Axiome du choix :* Etant donnée une famille $(a_i)_{i\in I}$ d'ensembles non vides, le produit cartésien $\prod_{i\in I} a_i$ est non vide.

Obtenir un élément dans le produit cartésien $\prod_{i\in I} a_i$ revient à savoir choisir un élément dans chaque a_i.

Avec cet énoncé, cet axiome paraît très raisonnable, on peut même se demander s'il n'est pas conséquence de l'axiomatique ZF. En fait, l'axiome du choix fait partie des indécidables de l'axiomatique ZF. En 1938, Gödel a démontré que si ZF est consistante alors ZF + l'axiome du choix l'est encore. En 1963, Paul Cohen (1934-) a démontré que ZF + la négation de l'axiome de choix l'est aussi.

L'axiome du choix est utile pour définir la notion de cardinal d'un ensemble, pour justifier que tout espace vectoriel possède une base ou encore pour définir des suites dont les termes sont choisis à l'intérieur d'ensembles qui varient avec l'indice de la suite... Il est usuel de noter ZFC l'axiomatique constituée par ZF + l'axiome du choix, c'est celle couramment utilisée.

L'axiome du choix semble raisonnable, il n'est pas source de contradictions, il est utile. Pourquoi est-il controversé ? Parce qu'il est lourd de conséquences :

Le paradoxe de Banach-Tarski

En 1923, Stefan Banach (1892-1945) et Alfred Tarski (1902-1983) mettent à jour le paradoxe suivant :

> Deux parties bornées d'intérieurs non vides de $\mathbb{R}^3$ peuvent se décomposer en un nombre fini de parties disjointes deux à deux identiques à un déplacement près.

Ce paradoxe signifie qu'il est théoriquement possible de découper une boule en un nombre fini de morceaux et de recombiner ceux-ci afin de former une boule dont le rayon est le double de la première ! De même on peut aussi découper la boule initiale pour former deux boules identiques à la première !

Suite à un tel paradoxe, on peut être tenté de réfuter l'axiome du choix. Pourtant la plupart des mathématiciens le conservent. Il est vrai que le paradoxe de Banach-Tarski est contraire au sens commun mais rappelons que les mathématiques ne sont pas une représentation de la réalité. Prenons, par exemple, la notion de droite. Une droite est un objet théorique sans épaisseur et de longueur infinie. Un tel objet n'existe pas dans la réalité, néanmoins on le conserve car il est essentiel en géométrie... Pour l'axiome du choix, c'est presque pareil. Nous savons que celui-ci est indécidable, il n'est donc pas plus exact de le conserver que de le réfuter. Comme il est utile, bon nombre le conservent.

Encadré : Variantes du paradoxe de Russell

En 1918, Russell vulgarise son paradoxe par le paradoxe du barbier : dans une ville, un barbier rase tous les gens qui ne se rasent pas eux-mêmes. Rien de plus normal, mais, qui rase le barbier ? On peut aussi le considérer comme une variante du paradoxe du menteur : quelqu'un dit « je suis un menteur », est-il menteur ou non ?

Encadré : Un indécidable : l'hypothèse du continu

Lors de sa construction de la théorie des infinis, Cantor avait l'intuition qu'il n'existait pas d'infinis intermédiaires entre le dénombrable et le continu. Cet énoncé est appelé l'hypothèse du continu. Gödel et Cohen ont établi que cet énoncé est indécidable, autrement dit l'axiomatique ZF ne permet pas de qualifier de vrai ou de faux celui-ci. On peut donc ajouter à l'axiomatique ZF l'hypothèse du continu ou bien sa négation !

Encadré : Ecriture quantifiée des axiomes

axiome d'extensionnalité : $\forall a \forall b \left[\forall x \left(x \in a \Leftrightarrow x \in b\right)\right] \Rightarrow a = b$

axiome de la paire : $\forall a \forall b \exists c \forall x \left[x \in c \Leftrightarrow \left(x \in a \text{ ou } x \in b\right)\right]$

axiome de la réunion : $\forall a \exists b \forall x \left[x \in b \Leftrightarrow \exists t \left(x \in t \text{ et } t \in a\right)\right]$

schéma de compréhension : $\forall a \exists b \forall x \left[x \in b \Leftrightarrow \left(x \in a \text{ et } P(x)\right)\right]$

axiome de l'ensemble des parties : $\forall a \exists b \forall x \left(x \in b \Leftrightarrow \left[\forall t \left(t \in x \Rightarrow t \in a\right)\right]\right)$

schéma d'axiomes de substitution :

$\left[\forall x \forall y \forall y' \left(E(x,y) \text{ et } E(x,y') \Rightarrow y = y'\right)\right] \Rightarrow \forall a \exists b \forall y \left[y \in b \Leftrightarrow \exists x \left(x \in a \text{ et } E(x,y)\right)\right]$

axiome de l'infini : $\exists a \left[\varnothing \in a \text{ et } \forall x \left(x \in a \Rightarrow x \cup \{x\} \in a\right)\right]$

Encadré possible :

« Car, pour forger, il faut un marteau, et pour avoir un marteau, il faut le fabriquer. Ce pourquoi on a besoin d'un autre marteau et d'autres outils, et pour les posséder, il faut encore d'autres instruments, et ainsi infiniment. Et par ce raisonnement il serait vain d'essayer de prouver que les hommes n'ont aucun moyen de forger. Mais, au début, les hommes, avec des instruments naturels, réussirent certains objets très faciles, bien qu'avec difficulté et imparfaitement, et ceux-ci fabriqués, ils confectionnèrent d'autres objets plus difficiles avec moins de travail et plus de perfection, et ainsi, graduellement des travaux les plus simples aux outils, des outils à d'autres travaux et d'autres outils, ils arrivèrent à exécuter de nombreux et très difficiles ouvrages et sans beaucoup de travail. »
Spinoza, *Œuvres complètes,* Paris, La Pléiade, 1954, p. 167.

Encadré possible :

Les paradoxes de Zénon d'Elée

Références :

Pour aller plus loin : http://www.dma.ens.fr/culturemath, rubrique dossiers, « Logique, fondements des Mathématiques »

Pour aller beaucoup plus loin : Théorie des ensembles, Jean-Louis Krivine, Ed. Cassini.