CHAPITRE I

Le type « ensemble »

RÉSUMÉ. • *Fréquemment utilisés, les ensembles sont des objets mathématiques spécifiques, munis d'opérations et de propriétés qui leur sont propres. Il est donc naturel d'en élaborer une théorie.*
• *Autre argument en faveur d'une théorie des ensembles: l'existence de nombreuses questions mettant en jeu les ensembles et leur taille (équipotence), notamment le problème du continu sur l'existence ou non d'ensembles de taille intermédiaire entre celles de* $\mathbb{N}$ *et de* $\mathbb{R}$.
• *Muni des opérations* $\cup$ *et* $\cap$, *tout ensemble du type* $\mathfrak{P}(A)$ *a une structure d'algèbre de Boole; inversement, toute algèbre de Boole finie est de ce type, ce qui, en un sens, achève l'étude des ensembles finis.*
• *Faute de pouvoir définir commodément les ensembles à partir d'objets plus primitifs, on recourt à une approche axiomatique.*
• *L'axiome d'extensionnalité affirme qu'un ensemble est déterminé par ses éléments.*
• *Première étape (Cantor): axiome de compréhension affirmant que toute propriété donne naissance à un ensemble.*
• *Seconde étape (Frege): échapper au paradoxe de Berry en restreignant la compréhension aux propriétés exprimables par une formule du premier ordre.*
• *Troisième étape (Zermelo): échapper au paradoxe de Russell en remplaçant la compréhension par la séparation; il faut alors réintroduire d'autres axiomes d'existence: paire, union, parties.*
• *Reste à choisir une signature adaptée: l'option minimale est de se restreindre aux ensembles purs et à l'unique relation d'appartenance ; on obtient ainsi le système de Zermelo fini; on peut alors aussi utiliser les opérations et relations définissables à partir de* $\in$.
• *La théorie ainsi obtenue n'apparaît pertinente que pour les ensembles purs, construits à partir de l'ensemble vide. La question de l'existence de suffisamment de tels ensembles reste ouverte.*

► Ce chapitre comporte trois parties. Dans la première, on justifie l'opportunité de développer une théorie des ensembles par la description de quelques problèmes liés aux ensembles et aux comparaison de cardinalité. Dans la seconde, on introduit les opérations ensemblistes usuelles, union, intersection, *etc.* pour constater que celles-ci ne posent pas de problème majeur. En particulier, la notion d'algèbre de Boole capture toutes les propriétés dans le cas fini. Dans la troisième partie, on commence à ébaucher une théorie des ensembles. Comme une définition *ex nihilo* est malaisée, on recourt à une approche axiomatique, et on montre comment échapper aux paradoxes de Berry et de Russell mène au système de Zermelo, base de tous les développements ultérieurs. ◄

▷ *L'objet de ce chapitre est de planter le décor des développements ultérieurs en discutant les bases possibles sur lesquelles édifier une théorie des ensembles. Le texte contient davantage de*

discussions informelles que de démonstrations, ce qui est naturel puisque le choix d'un système axiomatique est affaire de consensus et non de preuve. Néanmoins, sauf à se contenter d'une démarche dogmatique, il est difficile d'en faire l'économie car la justification de la plupart des points de vue adoptés ultérieurement repose sur la réflexion esquissée ici.

Les mathématiques étudient des objets appartenant à des types variés : entiers, réels, points, fonctions, etc. , chacun muni d'opérations et de relations qui lui sont spécifiques. Les ensembles constituent un tel type d'objet, et élaborer une théorie des ensembles signifie organiser en une suite cohérente nos connaissances sur ceux des objets mathématiques qui se trouvent être des ensembles, à la façon dont on développe une théorie des nombres ou une théorie des fonctions réelles. ◁

1. Pourquoi une théorie des ensembles?

► Après une brève discussion de la notion d'ensemble et de son utilité, on examine les problèmes liés à la comparaison de la taille des ensembles. Dans le cas fini, on établit facilement les résultats combinatoires élémentaires. Par contre, on rencontre rapidement des problèmes lorsqu'on passe aux ensembles infinis, en particulier le problème du continu de Cantor. ◄

▷ *Même si on convainc rapidement de l'utilité des ensembles en mathématiques, il n'est pas a priori évident qu'il soit utile ou nécessaire d'en développer une théorie : par exemple, les suites aussi sont des objets très utiles en mathématiques, et pourtant il n'existe pas à proprement parler de théorie générale des suites. Ce qui a rendu nécessaire la construction d'une théorie des ensembles, c'est l'apparition, à la fin du XIXe siècle et au début du XXe, de problèmes ouverts difficiles et naturels mettant en jeu les ensembles infinis. On peut alors espérer qu'une théorie cohérente les résoudra ou, au moins, les éclairera.* ◁

1.1. La notion d'ensemble.

► On discute informellement la notion, en rappelant les définitions et notations usuels. ◄

▷ *Plutôt que des objets mathématiques isolés, par exemple l'entier* 2 *ou le réel* π*, il arrive qu'on considère plusieurs objets simultanément sans vouloir ou pouvoir référer à l'un d'entre eux spécifiquement, par exemple les solutions d'une équation, les entiers pairs ou les réels transcendants : l'usage est alors de nommer cette réunion d'objets, donc, ce faisant, de l'introduire comme un nouvel objet mathématique. Ainsi, lorsqu'on on dit : « Soit* P *l'ensemble des entiers pairs », à côté des entiers* $0, 2, 4...$ *pris individuellement, on introduit, et en particulier on nomme, un nouvel objet référant à tous ces entiers pris collectivement. On déclare alors « n appartient à* P *», ou encore « n est élément de* P *», noté* $n \in P$*, comme une autre façon d'exprimer que n a la propriété définissant* P*, ici être un nombre pair.*

La notation traditionnelle pour l'ensemble dont les éléments sont les objets x *possédant une certaine propriété* $\mathcal{P}(x)$ *est* $\{x\,;\,\mathcal{P}(x)\}$ [1]*, et celle pour l'ensemble dont les éléments sont des objets explicitement énumérés* a_1*, ...,* a_n *est* $\{a_1, ..., a_n\}$*. Par exemple* $\{-1, 1\}$ *est l'ensemble dont les deux éléments sont les entiers* -1 *et* 1*. Noter que cet ensemble est aussi* $\{1, -1\}$*, et également* $\{-1, 1, 1\}$*, puisque seuls comptent les éléments indépendamment de tout ordre ou multiplicité. Diverses représentations graphiques sont utilisées, par exemple les diagrammes de Venn (figure 1).*

On pourrait confondre ensemble et propriété en posant qu'un ensemble est simplement la propriété qui le définit. Ce n'est pas le point de vue adopté : l'ensemble ne retient que le résultat final, c'est-à-dire les objets sélectionnés, et non la propriété utilisée pour opérer la sélection. De la sorte, un ensemble est complètement déterminé par ses éléments, et par eux seuls (propriété dite d'extensionnalité) : deux propriétés équivalentes conduisent à sélectionner les mêmes objets, et elles définissent donc un seul ensemble. Par exemple, pour un entier, être le double d'un entier

[1] notation à laquelle on se tiendra ici ; on trouve également $\{x : \mathcal{P}(x)\}$ et $\{x|\mathcal{P}(x)\}$

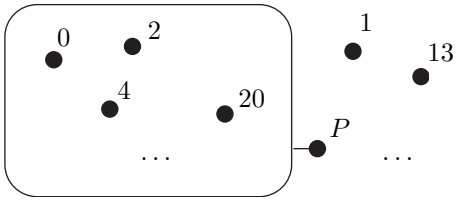

FIGURE 1. Représentation graphique d'un ensemble, ici l'ensemble P des nombres pairs : un cadre entoure les éléments, et laisse les non-éléments à l'extérieur ; dans le cas présent, la représentation est forcément incomplète puisqu'il existe une infinité d'entiers.

est équivalent à être congru à 0, 2, 4, 6, *ou* 8 *modulo* 10 *: les deux propriétés sont distinctes, mais l'ensemble qu'elles définissent est le même. Ainsi, un ensemble est associé à une classe d'équivalence de propriétés plutôt qu'à une propriété spécifique.*

Dans l'approche très libérale qui est celle de la théorie des ensembles telle que développée ici, on franchit une étape supplémentaire en oubliant même éventuellement les propriétés définissantes, et en envisageant abstraitement des ensembles indépendamment de toute propriété. Par exemple, une expression telle que « soit A un ensemble d'entiers » indique seulement que A est un objet tel que $n \in A$ est soit vrai, soit faux pour chaque entier n, et que A est entièrement déterminé par les entiers n qui vérifient $n \in A$, indépendamment du fait qu'on ait donné ou non une propriété explicite carac-térisant ceux-ci. ◁

1.2. Utilité des ensembles.

► L'introduction d'ensembles est non seulement commode, mais aussi utile pour exprimer des propriétés collectives ne faisant pas sens au niveau des objets individuels. ◄

▷ *La fréquence des termes « ensemble », « famille », « collection » dans les textes mathématiques contemporains ou anciens montre que l'introduction d'ensembles est au moins une convention commode : séparer les entiers pairs des impairs permet ensuite de travailler avec ceux-ci de manière uniforme, indépendamment de la propriété utilisée pour les séparer. De plus, de nombreux objets mathématiques sont définis comme des domaines, c'est-à-dire des ensembles, munis d'une structure additionnelle, algébrique, topologique, différentielle... et il serait donc malcommode de se passer d'ensembles pour définir un groupe, un corps, ou une variété différentiable.*

Il y a plus important. Au-delà de la commodité de formulation, l'introduction d'ensembles permet surtout de saisir des propriétés globales qui ne font pas sens pour chaque élément pris individuellement. Par exemple, affirmer que les multiples de 5 *forment un sous-groupe de* $\mathbb{Z}$ *dit quelque chose de plus que les propriétés individuelles des entiers* 10 *ou* -15*.*

Illustrons par un exemple. Soit à montrer qu'aucun entier n plus grand que 1 *ne divise* $3^n - 2^n$*. Soit p le plus petit facteur premier de n. Si p est* 2 *ou* 3*, alors p, donc a fortiori n, ne divise pas* $3^n - 2^n$*. Supposons $p \geqslant 5$. Alors les classes $\overline{2}$ et $\overline{3}$ de* 2 *et* 3 *dans* $\mathbb{Z}/p\mathbb{Z}$ *sont inversibles. L'ensemble* $\{k \in \mathbb{Z}\,;\, \overline{2}^k = \overline{3}^k\}$ *est un sous-groupe de* $\mathbb{Z}$*, donc il est de la forme* $m\mathbb{Z}$ *pour un certain m positif. Le petit théorème de Fermat implique* $\overline{2}^{p-1} = \overline{3}^{p-1} = \overline{1}$*, donc* $p - 1 \in m\mathbb{Z}$*, et* $m \leqslant p - 1$*. Donc m ne peut diviser n, puisque p est son plus petit facteur premier. On a donc* $n \notin m\mathbb{Z}$*, et p, donc n, ne divise pas* $3^n - 2^n$*. Dans la démonstation précédente, le point essentiel est l'introduction de l'ensemble* $\{k \in \mathbb{Z}\,;\, \overline{2}^k = \overline{3}^k\}$*, et le fait que tout sous-groupe de* $\mathbb{Z}$ *est de la forme* $m\mathbb{Z}$*. On pourrait s'en passer en redémontrant le résultat dans le cas particulier, en l'occurrence en considérant le plus petit entier m vérifiant* $\overline{2}^m = \overline{3}^m$*, et en montrant par division euclidienne que tout entier k vérifiant* $\overline{2}^k = \overline{3}^k$ *est multiple de m. Mais on perdrait*

ainsi la compréhension et l'économie apportées par le résultat structurel sur les sous-groupes de $\mathbb{Z}$, donc par l'introduction d'un ensemble. ◁

1.3. Premiers résultats, premiers problèmes.

► On établit des résultats simples de comparaison de taille entre ensembles, sous la forme d'existence ou de non-existence de bijections ou d'injections, en particulier le théorème de Cantor et le théorème de Cantor-Bernstein. On constate l'apparition de questions naturelles, telle le problème du continu sur l'existence de tailles intermédiaires entre celles de $\mathbb{N}$ et de $\mathbb{R}$. ◄

▷ *Même si on tient pour acquis que les ensembles sont utiles, il n'est pas évident qu'il faille construire une théorie générale des ensembles : les suites ou les fonctions sont aussi des outils importants et le besoin d'une théorie générale des suites ou des fonctions ne s'est pas fait ressentir jusqu'à présent. Ce qui a rendu l'élaboration d'une théorie des ensembles nécessaire, ou, au moins, opportune, c'est l'apparition, depuis la fin du XIXe siècle, de nombreux problèmes ouverts mettant en jeu des ensembles infinis et, typiquement, la comparaison de leur taille.*

La plupart des structures mathématiques vont de pair avec une notion de morphisme: homomorphismes pour les structures algébriques, fonctions continues et homéomorphismes pour les structures topologiques, etc. Dans le cas des ensembles, aucune structure additionnelle n'est considérée : les morphismes naturels sont donc les applications, et les isomorphismes sont les bijections, qui correspondent à l'intuition d'ensembles de même taille. Parmi les premières questions sur les ensembles figure donc la classification à bijection près, c'est-à-dire la comparaison des tailles. Si le cas des ensembles finis ne pose guère de problème, il n'en est pas de même du cas des ensembles infinis, inauguré par le résultat de Cantor affirmant la non-dénombrabilité de $\mathbb{R}$, puis l'existence d'une infinité de tailles d'infini distinctes. ◁

DÉFINITION 1.1. (équipotence) On dit que deux ensembles A, B sont *en bijection*, ou encore sont *équipotents*[2], s'il existe une bijection de A sur B.

L'identité, l'inverse d'une bijection, et la composée de deux bijections sont des bijections, donc l'équipotence est une relation d'équivalence.

▷ *L'existence d'une bijection entre A et B correspond à la possibilité de faire se correspondre un à un les éléments de A et de B, et elle exprime l'intuition que A et B ont le même nombre d'éléments, ou encore la même taille. De même, l'existence d'une injection de A dans B, donc d'une bijection de A sur un sous-ensemble de B, exprime l'intuition que la taille de A est au plus celle de B. L'utilisation d'un vocabulaire d'ordre est rendue cohérente par le résultat suivant.* ◁

PROPOSITION 1.2. (théorème de Cantor–Bernstein) *Supposons que A et B sont deux ensembles tels qu'il existe une injection de A dans B et une injection de B dans A. Alors A et B sont en bijection.*

DÉMONSTRATION. (Figure 2) Soient $f : A \to B$ et $g : B \to A$ des injections. Posons $A' = \mathrm{Im}(f)$ et $B' = \mathrm{Im}(g)$. Alors f est une bijection de A sur A', et g une bijection de B sur B'. On va construire une bijection h de A sur B'. Pour cela, posons $A_0 = A \setminus B'$, puis, inductivement, $A_1 = g \circ f[A_0]$, $A_2 = g \circ f[A_1]$, *etc.* Par construction, $g \circ f$ envoie $\bigcup_{i \geqslant 0} A_i$ sur $\bigcup_{i \geqslant 1} A_i$, lequel ensemble est inclus dans B'. Définissons alors h par

$$h(a) = \begin{cases} g \circ f(a) & \text{pour } a \in \bigcup_{i \geqslant 0} A_i, \\ a & \text{pour } a \notin \bigcup_{i \geqslant 0} A_i. \end{cases}$$

[2]terme un peu tombé en désuétude

Par construction, l'image de h est B', et, d'autre part, h est injective car obtenue en recollant deux injections dont les images sont disjointes. Donc h est une bijection de A sur B', et, finalement, $g^{-1} \circ h$ est une bijection de A sur B. □

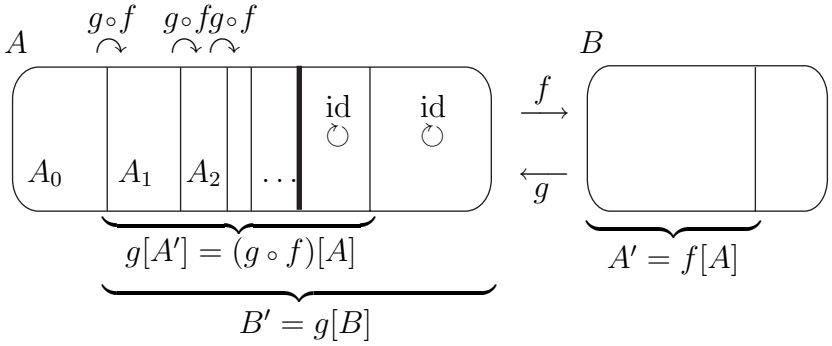

FIGURE 2. Démonstration du théorème de Cantor–Bernstein

La notion duale de celle d'injection est celle de surjection. L'existence d'une injection de A dans B entraîne celle d'une surjection de B sur A : si A est non vide et si $f : A \to B$ est injective, on définit une surjection $g : B \to A$ en choisissant un élément a de A et en posant $g(b) = f^{-1}(b)$ pour b dans $\mathrm{Im}(f)$ et $g(b) = a$ sinon.

QUESTION 1.3. *L'existence d'une injection de A dans B est-elle équivalente à celle d'une surjection de B sur A ?*

▷ *La réponse n'est pas évidente, et on y reviendra au chapitre IV.*
De nombreuses questions apparaissent lorsqu'on distingue les ensembles finis et infinis. Plusieurs définitions sont possibles ; ici on part de l'idée qu'un ensemble fini est un ensemble dont on peut numéroter les éléments du premier au dernier. ◁

DÉFINITION 1.4. (fini, infini) Un ensemble est dit *fini* s'il est en bijection avec un intervalle de $\mathbb{N}$ de la forme $\{1, \dots, p\}$; un ensemble qui n'est pas fini est dit *infini*.

PROPOSITION 1.5. (bijection) *Toute injection d'un ensemble fini dans lui-même est une bijection.*

DÉMONSTRATION. Il suffit de montrer le résultat pour les intervalles $\{1, \dots, p\}$. On raisonne par récurrence sur $p \geqslant 1$. Pour $p = 1$, l'intervalle $\{1, \dots, p\}$ est le singleton $\{1\}$, et la seule application de $\{1\}$ dans lui-même est l'identité, qui est une bijection. Supposons $p \geqslant 2$, et soit f une injection de $\{1, \dots, p\}$ dans lui-même. Soit $q = f(p)$. On définit une application g de $\{1, \dots, p-1\}$ dans lui-même en posant $g(i) = f(i)$ pour $f(i) < q$ et $g(i) = f(i) - 1$ pour $f(i) > q$. Alors g est injective puisque q n'appartient pas à $f[\{1, \dots, p-1\}]$. Par hypothèse de récurrence, g est surjective. Par construction, on a $\mathrm{Im}(f) = \mathrm{Im}(g) \cup \{p\}$, d'où $\mathrm{Im}(f) = \{1, \dots, p\}$. □

COROLLAIRE 1.6. *Tout ensemble fini est en bijection avec un unique intervalle $\{1, \dots, p\}$ de $\mathbb{N}$.*

On peut donc sans ambiguïté définir le cardinal d'un ensemble fini A comme l'unique entier p tel que A soit en bijection avec $\{1, \dots, p\}$. On a alors une classification complète des ensembles finis à l'aide des entiers : deux ensembles finis sont en bijection si et seulement si ils ont le même cardinal, et il existe une injection de A dans B si et seulement le cardinal de A est inférieur ou égal à celui de B.

▷ *La proposition 1.5 exprime qu'une partie propre d'un ensemble fini est strictement plus petite que celui-ci. Ce résultat ne s'étend pas aux ensembles infinis : par exemple, l'application « successeur » qui, pour tout entier n, envoie n sur $n+1$ définit une bijection de $\mathbb{N}$ sur la partie propre de $\mathbb{N}$ composée des entiers non nuls, ce qui montre que $\mathbb{N}$ privé de 0 n'est pas plus petit que $\mathbb{N}$. De même, l'application qui, pour tout entier n, envoie n sur $2n$ définit une bijection de $\mathbb{N}$ sur l'ensemble des nombres pairs, ce qui montre que ce dernier, qui pourtant ne contient qu'un entier sur deux, n'est pas plus petit que $\mathbb{N}$. On peut même aller plus loin et établir d'autres égalités de taille encore plus paradoxales. On notera l'utilisation répétée du théorème de Cantor–Bernstein dans les démonstrations : celui-ci simplifie les arguments en remplaçant la construction d'une bijection par celle de deux injections indépendantes.* ◁

PROPOSITION 1.7. (carré) *Le produit $\mathbb{N} \times \mathbb{N}$ est en bijection avec $\mathbb{N}$.*

DÉMONSTRATION. Comme tout entier non nul s'écrit de façon unique comme le produit d'une puissance de 2 et d'un entier impair, l'application $(p, q) \mapsto 2^p(2q+1)$ est une bijection de $\mathbb{N} \times \mathbb{N}$ sur $\mathbb{N} \setminus \{0\}$, donc $(p, q) \mapsto 2^p(2q+1) - 1$ est une bijection de $\mathbb{N} \times \mathbb{N}$ sur $\mathbb{N}$.

Une autre bijection est fournie par $(p, q) \mapsto (p+q)(p+q+1)/2+q$, qui correspond au parcours suivant des diagonales successives de $\mathbb{N} \times \mathbb{N}$ supposé écrit dans un quadrant (Figure 3). □

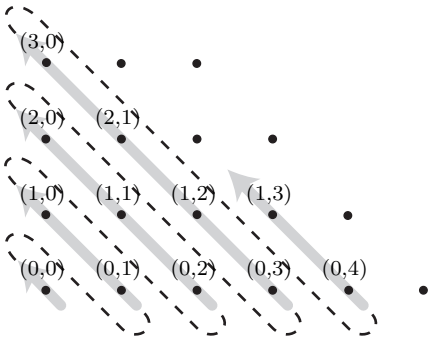

FIGURE 3. Une numérotation des couples d'entiers par des entiers

COROLLAIRE 1.8. (*i*) *Les ensembles $\mathbb{Z}$ et $\mathbb{Q}$ sont en bijection avec $\mathbb{N}$.*
(*ii*) *Le produit $\mathbb{R} \times \mathbb{R}$ est en bijection avec $\mathbb{R}$.*

DÉMONSTRATION. (*i*) L'inclusion fournit une injection triviale de $\mathbb{N}$ dans $\mathbb{Z}$. D'un autre côté, $\mathbb{Z}$ peut être construit comme quotient de $\mathbb{N} \times \mathbb{N}$, d'où on tire une injection de $\mathbb{Z}$ dans $\mathbb{N} \times \mathbb{N}$ donc dans $\mathbb{N}$ par la proposition 1.7, et on conclut par le théorème de Cantor–Bernstein. L'argument est le même pour $\mathbb{Q}$, qui peut être injecté dans $\mathbb{Z} \times \mathbb{Z}$.

(*ii*) Puisque $\mathbb{R}$ est en bijection avec $\mathfrak{P}(\mathbb{N})$, le produit $\mathbb{R} \times \mathbb{R}$ est en bijection avec $\mathfrak{P}(\mathbb{N}) \times \mathfrak{P}(\mathbb{N})$. Ce dernier s'injecte dans $\mathfrak{P}(\mathbb{N} \times \mathbb{N})$, donc, par la proposition 1.7, dans $\mathfrak{P}(\mathbb{N})$. Donc $\mathbb{R} \times \mathbb{R}$ s'injecte dans $\mathbb{R}$. Comme $\mathbb{R}$ s'injecte trivialement dans $\mathbb{R} \times \mathbb{R}$, le théorème de Cantor–Bernstein donne à nouveau le résultat. □

▷ *A ce point, on peut se demander si* tous *les ensembles infinis sont deux à deux en bijection, c'est-à-dire s'il existe plusieurs tailles d'infini distinctes. Une réponse formelle à cette question n'est apparue qu'assez récemment, précisément lorsque Georg Cantor a montré en 1874 le résultat suivant, qu'on peut prendre comme acte de naissance de la théorie des ensembles.* ◁

PROPOSITION 1.9. (diagonal) *L'ensemble* $\mathbb{R}$ *n'est pas dénombrable.*

DÉMONSTRATION. Il est suffisant de montrer que l'intervalle $[0, 1]$ de $\mathbb{R}$ n'est pas dénombrable. Or, soit A l'ensemble des nombres réels compris entre 0 et 1 dont un développement en base 3 ne contient que des 0 et des 1, et soit $f : \mathbb{N} \to A$ quelconque. On va montrer que f n'est pas surjective, à savoir qu'il existe au moins un réel dans A distinct de chacun des réels $f(1)$, $f(2)$, ... Pour cela, écrivons le développement en base 3 de $f(i)$ sous la forme $0, a_{i,0}a_{i,1}\ldots$ où les chiffres $\alpha_{i,j}$ sont 0 ou 1. Posons $\overline{0} = 1$, $\overline{1} = 0$, et considérons le réel a dont le développement est $0, \overline{a_{0,0}}\,\overline{a_{1,1}}\ldots$. Alors a, qui est dans A par construction, ne peut être, quel que soit i, égal à $f(i)$, car le i-ème chiffre du développement de $f(i)$ est $a_{i,i}$, alors que celui de a est $\overline{a_{i,i}}$, et que le développement en base 3 sans chiffre 2 est unique quand il existe. Donc f n'est pas surjective. □

▷ *La démonstration précédente repose sur ce qu'on appelle l'argument diagonal, qui conjugue autoréférence (ici les éléments diagonaux* $a_{i,i}$*) et négation (ici l'application* $\overline{\ }$*). Le recours à la base* 3 *plutôt que* 2 *n'est là que pour pallier le manque d'unicité du développement binaire pour les rationnels dyadiques.*

Il est facile d'établir une bijection entre $\mathbb{R}$ *et l'ensemble* $\mathfrak{P}(\mathbb{N})$*, donc une autre démonstration de la proposition 1.9 est fournie par le résultat suivant, également dû à Cantor et reposant à nouveau sur un argument diagonal. Ce résultat est plus général que le précédent puisqu'il montre l'existence d'une infinité d'ensembles infinis deux à deux non équipotents.* ◁

PROPOSITION 1.10. (théorème de Cantor) *Soit* A *un ensemble quelconque. Alors il n'existe pas de surjection de* A *sur* $\mathfrak{P}(A)$ *; en particulier,* A *et* $\mathfrak{P}(A)$ *ne sont pas en bijection.*

DÉMONSTRATION. Soit $f : A \to \mathfrak{P}(A)$ quelconque. Posons $B = \{a \in A\,;\, a \notin f(a)\}$. Pour tout a dans A, si a est dans B, on a $a \in B \setminus f(a)$, donc $f(a) \neq B$, et si a n'est pas dans B, on a $a \in f(a) \setminus B$ donc, à nouveau, $f(a) \neq B$. Par conséquent B n'appartient pas à l'image de f, et f n'est pas surjective. □

La taille de l'ensemble $\mathbb{R}$ est donc strictement supérieure à celle de l'ensemble $\mathbb{N}$. Parmi les sous-ensembles infinis de $\mathbb{R}$, certains sont en bijection avec $\mathbb{N}$ (par exemple $\mathbb{N}$ lui-même), d'autres sont en bijection avec $\mathbb{R}$ (par exemple $\mathbb{R}$ lui-même.

QUESTION 1.11. (problème du continu) *Tout sous-ensemble infini de* $\mathbb{R}$ *est-il en bijection soit avec* $\mathbb{N}$*, soit avec* $\mathbb{R}$ *?*

Une réponse positive est appelée l'*hypothèse du continu*, souvent abrégée en HC.

▷ *Posée par Cantor à la fin du XIXe siècle, la questions est celle de l'existence de tailles intermédiaires entre celle de* $\mathbb{N}$ *(le dénombrable) et celle de* $\mathbb{R}$ *(traditionnellement appelée le continu). Plus d'un siècle après que Cantor a soulevé la question, et que Hilbert l'a placée en première position de sa célèbre liste de problèmes ouverts en 1900, le sort du problème du continu est toujours loin d'être réglé. Tout au long du XXe siècle, il a été un des moteurs principaux du développement de la théorie des ensembles, et il le reste aujourd'hui avec notamment les récentes avancées dues à Hugh Woodin.* ◁

Bien d'autres questions se posent au sujet des tailles des ensembles infinis, et on va mentionner maintenant quelques autres problèmes sur lesquels on reviendra au chapitre IV. D'abord, on a montré l'existence de bijections entre $\mathbb{N}$ et $\mathbb{N} \times \mathbb{N}$, et entre $\mathbb{R}$ et $\mathbb{R} \times \mathbb{R}$.

QUESTION 1.12. *Tout ensemble infini A est-il en bijection avec $A \times A$?*

Ensuite, s'il existe une injection f de $\mathbb{N}$ dans un ensemble A, alors A est infini, car l'application envoyant $f(n)$ sur $f(n+1)$ pour tout n et laissant fixes les éléments non dans l'image de f est une injection non surjective de A dans lui-même.

QUESTION 1.13. *Existe-t-il une injection de $\mathbb{N}$ dans tout ensemble infini ?*

Enfin, on termine avec un problème ne mettant pas en jeu la taille des ensembles, mais simplement le fait de savoir s'ils sont vides ou non. On a utilisé ci-dessus le produit de deux ensembles A_1, A_2, défini comme ensemble des couples (a_1, a_2) avec $a_1 \in A_1$ et $a_2 \in A_2$. La notion s'étend naturellement à une famille quelconque d'ensembles :

DÉFINITION 1.14. (produit) Soit $(A_i)_{i \in I}$ une famille d'ensembles. On appelle *produit* des A_i, et on note $\prod_{i \in I} A_i$, l'ensemble des suites $(s_i)_{i \in I}$ vérifiant $s_i \in A_i$ pour chaque i.

Si $A_1, \ldots, A_n$ sont des ensembles non vides, le produit $A_1 \times \cdots \times A_n$ est non vide, car, en choisissant un élément a_1 dans A_1, puis un élément a_2 dans A_2, et ainsi de suite jusqu'à a_n dans A_n, on obtient un n-uplet $(a_1, \ldots, a_n)$ qui, par définition, est dans le produit $A_1 \times \cdots \times A_n$. Dans le cas d'une famille infinie $(A_t)_{t \in I}$, l'argument précédent ne s'adapte pas directement.

QUESTION 1.15. *Tout produit d'ensembles non vides est-il non vide ?*

▷ *A ce point, on constate donc l'existence de divers problèmes mettant en jeu des ensembles infinis, problèmes qu'on peut qualifier de purement ensemblistes car ils ne mettent en jeu aucune structure additionnelle. C'est en particulier le cas du problème du continu. Résoudre ces problèmes est la plus puissante des motivations pour développer une théorie des ensembles.*

Anticipant sur les résultats du chapitre III, on peut aussi mentionner dès à présent une motivation supplémentaire liée à la possibilité d'utiliser les ensembles comme type privilégié dans lequel tous les autres types d'objets mathématiques peuvent se représenter, et, partant, de leur faire jouer un rôle essentiel dans la fondation de l'édifice mathéma-tique. ◁

2. Opérations ensemblistes

► Cette brève partie établit quelques propriétés des opérations d'union et d'intersection d'ensembles. On montre que tout ensemble des parties $\mathfrak{P}(A)$ a une structure d'algèbres de Boole, et qu'inversement toute algèbre de Boole finie est de ce type. ◄

▷ *De même que d'une notion de morphisme, chaque type d'objet mathématique est accompagné d'opérations et de relations qui lui sont propres. Dans le cas des ensembles, de multiples opérations et relations s'introduisent naturellement: inclusion, union, intersection, différence... Une approche naïve pourrait faire penser que ces opérations sont le cœur de la théorie des ensembles, et que celle-ci consiste essentiellement à manipuler des expressions compliquées à base de* $\cup$ *et de* $\cap$*. En fait, ce n'est pas le cas : au moins dans le cas fini, les propriétés des opérations ensemblistes sont complètement décrites par le fait qu'elles définissent une algèbre de Boole, ce qui, en un sens, trivialise et clôt l'étude des ensembles finis.* ◁

2.1. Le treillis des parties d'un ensemble.

► On rappelle le vocabulaire usuel concernant l'inclusion et les opérations d'union et intersection. On montre que la relation d'inclu-sion est une relation d'ordre, et que sa restriction à tout ensemble du type $\mathfrak{P}(A)$ est un treillis distributif et complémenté. ◄

DÉFINITION 2.1. (inclusion, parties) Si A et B sont des ensembles, on dit que A est *inclus* dans B, ou encore que B est une *partie* de A, noté $A \subseteq B$ [3], si tout élément de B est élément de A. On note $\mathfrak{P}(A)$ l'ensemble de toutes les parties de A.

DÉFINITION 2.2. (opérations ensemblistes) Si A et B sont des ensembles, on définit

$$\begin{aligned} A \cup B &= \{t\,;\, t \in A \text{ ou } t \in B\}, && \textit{union} \text{ de } A \text{ et } B,\\ A \cap B &= \{t\,;\, t \in A \text{ et } t \in B\}, && \textit{intersection} \text{ de } A \text{ et } B,\\ A \setminus B &= \{t \in A\,;\, t \notin B\}, && \textit{différence} \text{ de } A \text{ et } B,\\ A \vartriangle B &= A \setminus B \cup B \setminus A, && \textit{différence symétrique} \text{ de } A \text{ et } B. \end{aligned}$$

La figure 4 illustre les opérations ensemblistes dans la représentation par diagrammes de Venn (*cf.* figure 1).

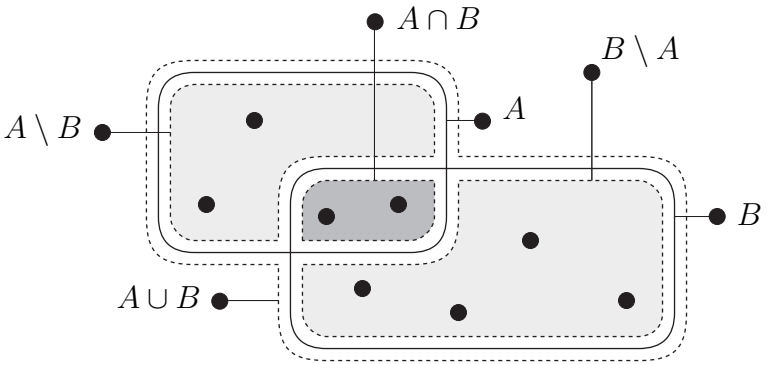

FIGURE 4. Opérations ensemblistes représentées à l'aide de diagrammes de Venn

[3] la notation « ordre large » est plus cohérente que $\subset$, qui suggère un ordre strict

Dans le contexte où un ensemble Ω est spécifié et où on s'intéresse exclusivement aux parties de Ω, il est d'usage d'introduire l'opération de *complémentaire* définie par $A^c = \Omega \setminus A$.

LEMME 2.3. *La relation d'inclusion est une relation d'ordre.*

DÉMONSTRATION. La réflexivité et la transitivité sont immédiates. L'antisymétrie provient de la propriété d'extensionnalité qui affirme que deux ensembles ayant les mêmes éléments coïncident. □

DÉFINITION 2.4. (treillis, algèbre de Boole) Un ensemble ordonné $(X, \leqslant)$ est appelé *treillis* si chaque paire d'éléments de X possède une borne supérieure et une borne inférieure. Un treillis est dit *distributif* si l'opération sup est distributive par rapport à l'opération inf, et *vice versa* ; il est dit *complémenté* s'il possède un minimum 0, un maximum 1 et si, pour tout x, il existe un élément x^c, appelé *complément* de c, vérifiant $\sup(x, x^c) = 1$ et $\inf(x, x^c) = 0$. Un treillis distributif et complémenté est appelé *algèbre de Boole*.

EXEMPLE 2.5. L'ensemble des entiers non nuls muni de la relation de divisibilité est un treillis distributif — le vérifier — possédant un élément minimum, à savoir 1, mais pas d'élément maximum, et ce n'est donc pas une algèbre de Boole. Par contre, l'ensemble des diviseurs d'un entier fixé N n'ayant que des facteurs premiers simples est une algèbre de Boole.

LEMME 2.6. *Supposons que $(X, \leqslant)$ est un treillis distributif possédant un minimum 0 et un maximum 1. Alors, pour tout élément a de X, il existe au plus un élément b vérifiant $\inf(a, b) = 0$ et $\sup(a, b) = 1$.*

DÉMONSTRATION. Supposons qu'on a à la fois $\inf(a, b) = 0$ et $\sup(a, b) = 1$, et $\inf(a, c) = 0$ et $\sup(a, c) = 1$. On veut montrer $b = c$. Posons $d = \sup(b, c)$. D'abord, $\sup(a, b) = 1$ entraîne $\inf(\sup(a, b), d) = \inf(1, d) = d$. Par ailleurs, par distributivité, on a aussi

$$\inf(\sup(a, b), d) = \sup(\inf(a, d), \inf(b, d)).$$

Or, par distributivité encore, on a

$$\inf(a, d) = \inf(a, \sup(b, c)) = \sup(\inf(a, b), \inf(a, c)) = \sup(0, 0) = 0,$$

et, par construction, $\inf(b, d) = b$. Nous obtenons donc $d = \inf(\sup(a, b), d) = \sup(0, b) = b$, soit $\sup(b, c) = b$, et donc $b \geqslant c$. Un raisonnement symétrique donne $c \geqslant b$, d'où $b = c$. □

Donc, dans une algèbre de Boole, le complément est nécessairement unique.

PROPOSITION 2.7. (algèbre de Boole) *Pour tout ensemble A, l'ensemble $\mathfrak{P}(A)$ muni de $\subseteq$ est une algèbre de Boole; l'opération* sup *est l'union, l'opéra-tion* inf *est l'intersection, le minimum est l'ensemble vide, le maximum est A, et le complément est le complémentaire.*

DÉMONSTRATION. Une vérification directe est facile.

Une démonstration alternative plus rapide consiste à remarquer que les algèbres de Boole sont définies par la satisfaction d'identités algébriques (voir paragraphe suivant), d'où il résulte

que tout produit d'algèbres de Boole est une algèbre de Boole. Or, définissons un ordre $\leqslant$ sur $\{0,1\}$ en posant $0<1$. Alors $(\{0,1\},\leqslant)$ est une algèbre de Boole. Maintenant, l'application

$$F : \mathfrak{P}(A) \to \{0,1\}^A$$

définie par $F(X)(x)=1$ pour $x\in X$ et $F(X)(x)=0$ pour $x\notin X$ établit un isomorphisme entre les structures $(\mathfrak{P}(A),\subseteq,\cup,\cap,\emptyset,A,{}^c)$ et $(\{0,1\},\leqslant,\sup,\inf,0,1,x\mapsto 1-x)^A$. □

La figure 5 montre les diagrammes de Hasse des algèbres de Boole $(\mathfrak{P}(A),\subseteq)$ pour A fini avec au plus 4 éléments.

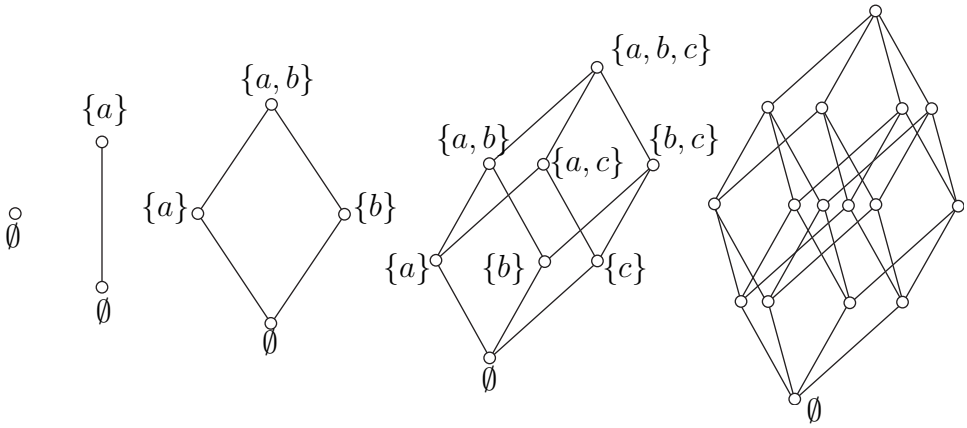

FIGURE 5. Diagramme de Hasse des algèbres de Boole $(\mathfrak{P}(A),\subseteq)$ pour A fini à n éléments avec $n\leqslant 4$; on y reconnaît la projection d'un n-cube

2.2. Les algèbres de Boole comme structures algébriques.

► Dans la section précédente, les algèbres de Boole ont été définies comme structures ordonnées. On établit ici l'équivalence avec une définition algébrique en termes de lois satisfaites par les opérations sup et inf. ◄

PROPOSITION 2.8. (axiomatisation 1) *(i) Soit $(T,\leqslant)$ un treillis. Pour a, b dans B, posons $a\vee b=\sup(a,b)$ et $a\wedge b=\inf(a,b)$. Alors $(T,\vee,\wedge)$ vérifie les lois*

$$\begin{aligned}
x\vee x&=x, && (I_0) & \qquad x\wedge x&=x, && (I'_0)\\
x\vee y&=y\vee x, && (I_1) & \qquad x\wedge y&=y\wedge x, && (I'_1)\\
x\vee(y\vee z)&=(x\vee y)\vee z, && (I_2) & \qquad x\wedge(y\wedge z)&=(x\wedge y)\wedge z, && (I'_2)\\
x\vee(x\wedge y)&=x, && (I_3) & \qquad x\wedge(x\vee y)&=x. && (I'_3)
\end{aligned}$$

De plus, $a\leqslant b$ est équivalent à $a\vee b=b$ et à $a\wedge b=a$.

(ii) Inversement, supposons que $(T,\vee,\wedge)$ satisfait aux lois (I_1), (I_2), (I_3), et (I'_1), (I'_2), (I'_3); pour $a,b\in T$, notons $a\leqslant b$ pour $a\vee b=b$; alors $(T,\leqslant)$ est un treillis, et $\vee$ (resp. $\wedge$) en est l'opération sup *(resp.* inf*).*

DÉMONSTRATION. (i) Pour tous a, b, c dans T, on a $\sup(a,a) = a$, $\sup(a,b) = \sup(b,a)$ et $\sup(a, \sup(b,c)) = \sup(a,b,c)$, donc $\sup(a, \sup(b,c)) = \sup(\sup(a,b), c)$. De plus, $a \leqslant b$ équivaut à $\sup(a,b) = b$, et à $\inf(a,b) = a$. Donc, comme on a toujours $a \leqslant \sup(a,b)$, on a $\inf(a, \sup(a,b)) = a$, et de même, $\inf(a,b) \leqslant a$ entraîne $\sup(a, \inf(a,b)) = a$. Donc les lois (I_0), (I_1), (I_2), (I_3) et (I'_3) sont satisfaites, et (I'_1), (I'_2) et (I'_3) sont obtenues par un argument symétrique.

(ii) Notons d'abord que les opérations $\vee$ et $\wedge$ sont idempotentes, c'est-à-dire que les lois (I_0) et (I'_0) sont conséquences de $(I_1), \dots, (I'_3)$. Soit a quelconque. Par (I_3) et (I'_3), nous avons $a = a \vee (a \wedge (a \vee a)) = a \vee a$, et, de même, $a \wedge (a \vee (a \wedge a)) = a \wedge a$.

On montre d'abord que $\leqslant$ est une relation d'ordre. Pour tout a dans T, on a $a \vee a = a$, donc $a \leqslant a$, et $\leqslant$ est réflexive. Supposons $a \leqslant b$ et $b \leqslant a$. Appliquant la commutativité de $\vee$, on obtient $a = b \vee a = a \vee b = b$, et $\leqslant$ est antisymétrique. Supposons $a \leqslant b \leqslant c$. Appliquant l'associativité de $\vee$, on obtient $a \vee c = a \vee (b \vee c) = (a \vee b) \vee c = b \vee c = c$, donc $a \leqslant c$, et $\leqslant$ est transitive.

Montrons ensuite que l'élément $a \vee b$ est borne supérieure de a et b vis-à-vis de $\leqslant$. D'abord on a $a \vee (a \vee b) = (a \vee a) \vee b = a \vee b$, donc $a \leqslant a \vee b$, et $b \vee (a \vee b) = b \vee (b \vee a) = (b \vee b) \vee a = b \vee a = a \vee b$, donc $b \leqslant a \vee b$, et $a \vee b$ est un majorant commun à a et b. Supposons ensuite que c est un majorant commun à a et b. On a donc $a \vee c = c$ et $b \vee c = c$, donc $(a \vee b) \vee c = a \vee (b \vee c) = a \vee c = c$, soit $a \vee b \leqslant c$, et on conclut que $a \vee b$ est le plus petit de tous les majorants communs à a et b.

Montrons maintenant que $a \leqslant b$, c'est-à-dire $a \vee b = b$, est équivalent à $a \wedge b = a$. Supposons $a \vee b = b$. Alors, par (I'_4), nous avons $a \wedge b = a \wedge (a \vee b) = a$. Inversement, supposons $a \wedge b = a$. Par (I_2) et (I_4), nous avons $a \vee b = (a \wedge b) \vee b = b \vee (b \wedge a) = b$.

Dès lors que $\vee$ et $\wedge$ jouent des rôles symétriques par rapport à l'ordre $\leqslant$, le raisonnement montrant que $a \vee b$ est borne supérieure de a et b montre *ipso facto* que $a \wedge b$ est borne inférieure de a et b, et on conclut que $(T, \leqslant)$ est un treillis. □

PROPOSITION 2.9. (axiomatisation 2) *(i) Supposons que $(B, \leqslant)$ est une algèbre de Boole de minimum 0 et de maximum 1. Pour $a, b \in B$, posons $a \vee b = \sup(a,b)$, $a \wedge b = \inf(a,b)$, et notons $\overline{a}$ l'unique élément vérifiant $a \vee \overline{a} = 1$ et $a \wedge \overline{a} = 0$. Alors $(B, \vee, \wedge, 0, 1, \overline{\ })$ vérifie les lois $(I_0), \dots, (I'_3)$ de la Proposition 2.8 (axiomatisation I), et, de plus,*

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z), \quad (I_4) \qquad x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z), \quad (I'_4)$$

$$x \vee 0 = x, \quad x \vee 1 = 1, \quad (I_5) \qquad x \wedge 0 = 0, \quad x \wedge 1 = x, \quad (I'_5)$$

$$x \vee \overline{x} = 1, \quad (I_6) \qquad x \wedge \overline{x} = 0. \quad (I'_6)$$

(ii) Inversement, supposons que $(B, \vee, \wedge, 0, 1, \overline{\ })$ satisfait aux lois $(I_1), \dots, (I'_6)$; notons $a \leqslant b$ pour $a \vee b = b$; alors $(B, \leqslant)$ est une algèbre de Boole, $\vee$ (resp. $\wedge$) est l'opération sup *(resp.* inf*) associée, 1 (resp. 0) est le maximum (resp. le minimum).*

DÉMONSTRATION. D'abord $(B, \leqslant)$ est un treillis, donc les lois $(I_0), \dots, (I'_3)$ sont satisfaites dans B, et, réciproquement, dès lors que $(I_1), \dots, (I'_3)$ sont vérifiées, on sait par la Proposition 2.8 (axiomatisation I) que $(B, \leqslant)$ est un treillis. Ensuite (I_4) et (I'_4) traduisent directement le fait que le treillis est distributif, (I_5) et (I'_5) traduisent le fait que 0 est minimum et 1 est maximum, et (I_6) et (I'_6) le fait que $\overline{a}$ est un complément pour a. □

On peut donc appeler algèbre de Boole aussi bien un treillis distributif et complémenté qu'une structure algébrique $(B, \vee, \wedge, 0, 1, \overline{\ })$ vérifiant les lois de la Proposition 2.9.

On conclut avec une caractérisation alternative des algèbres de Boole, toujours à l'aide de lois algébriques, mais cette fois dans le langage des anneaux.

DÉFINITION 2.10. (anneau de Boole) On appelle *anneau de Boole* un anneau commutatif et idempotent, c'est-à-dire un anneau dont la multiplication est commutative et où on a $x^2 = x$ pour tout x.

Il est facile de vérifier que, pour tout ensemble A, l'ensemble $\mathfrak{P}(A)$ muni des opérations Δ et $\cap$ est un anneau de Boole. Ce résultat n'est qu'un cas particulier du résultat général suivant exprimant l'équivalence entre algèbre de Boole et anneau de Boole. La démonstration est une vérification facile (exercice 6).

PROPOSITION 2.11. (équivalence) *(i) Supposons que $(B, \vee, \wedge, 0, 1, \overline{\ })$ est une algèbre de Boole. Alors $(B, \Delta, \wedge)$ est un anneau de Boole, où Δ est définie par $a \Delta b = (a \wedge \overline{b}) \vee (\overline{a} \wedge b)$.*

(ii) Inversement, supposons que $(B, +, \cdot)$ est un anneau de Boole. Alors $(B, \vee, \wedge, 0, 1, \overline{\ })$ est une algèbre de Boole, où $\vee$, $\wedge$ et $\overline{\ }$ sont définies par $a \vee b = a + b + ab$, $a \wedge b = ab$, et $\overline{a} = 1 + a$.

2.3. Algèbres de Boole finies.

▶ On montre que toute algèbre de Boole finie est isomorphe à une algèbre du type $\mathfrak{P}(A)$. ◀

DÉFINITION 2.12. (atome) Supposons que $(A, <)$ est un ensemble ordonné possédant un élément minimum 0. On dit que a est un *atome* de $(A, <)$ si a est un successeur immédiat de 0, c'est-à-dire qu'on a $0 < a$, mais qu'il n'existe pas d'élément b vérifiant $0 < b < a$.

▷ *Dans une algèbre de Boole de type $(\mathfrak{P}(A), \subseteq)$, les atomes sont les singletons, et, par conséquent, tout élément est union, c'est-à-dire borne supérieure, d'atomes. Si toute algèbre de Boole finie est isomorphe à une algèbre de type $\mathfrak{P}(A)$, il doit être vrai en particulier que, dans toute algèbre de Boole finie, tout élément est borne supérieure d'atomes. C'est ce qui suggère l'idée de la démonstration ci-dessous.* ◁

PROPOSITION 2.13. (algèbres de Boole finies) *Toute algèbre de Boole finie est isomorphe à une algèbre du type $(\mathfrak{P}(A), \subseteq)$.*

DÉMONSTRATION. Supposons que $(B, \vee, \wedge, 0, 1, \overline{\ })$ est une algèbre de Boole finie. On note A l'ensemble des atomes de B, et on définit $F : B \to \mathfrak{P}(A)$ par $F(b) = \{a \in A \,;\, a \leqslant b\}$. On va montrer que F établit l'isomorphisme cherché de $(B, \vee, \wedge, 0, 1, \overline{\ })$ sur $(\mathfrak{P}(A), \cup, \cap, \emptyset, A, ^c)$. Notons déjà qu'on a $F(0) = \emptyset$ et $F(1) = A$ par construction.

Montrons d'abord que $b \neq 0$ entraîne $F(b) \neq \emptyset$. En effet, soit $(b_0 = b, b_1, b_2, \dots)$ une chaîne strictement décroissante partant de b et de longueur maximale. Les éléments b_i sont deux à deux distincts, donc la longueur de la chaîne est au plus le cardinal de B. Il existe donc n vérifiant $b_n = 0$. Alors l'élément b_{n-1}, qui est un minorant de b par construction, est un atome de B.

En effet, s'il existait c vérifiant $0 < c < b_{n-1}$, la chaîne $(b_0, b_1, \dots, b_{n-1}, c, b_n)$ contredirait la maximalité de $(b_0, b_1, \dots, b_{n-1}, b_n)$.

Soit b quelconque dans B. Soit a un atome. Si a minorait à la fois b et $\overline{b}$, il minorerait $b \wedge \overline{b}$, qui est 0, ce qui est impossible, donc $F(b)$ et $F(\overline{b})$ sont disjoints. D'un autre côté, supposons $a \notin F(b)$, c'est-à-dire $a \not\leqslant b$. On a $a \wedge b \leqslant a$ et $a \wedge b \neq a$ (sinon on aurait $a \leqslant b$), donc, par définition d'un atome, $a \wedge b = 0$. On obtient $a = a \wedge 1 = a \wedge (b \vee \overline{b}) = (a \wedge b) \vee (a \wedge \overline{b}) = 0 \vee (a \wedge \overline{b}) = a \wedge \overline{b}$, donc $a \leqslant \overline{b}$. Par conséquent, $a \notin F(b)$ entraîne $a \in F(\overline{b})$, et on déduit $F(\overline{b}) = F(b)^c$.

Soient b et c quelconques de B. Par définition de la borne inférieure, un atome minore $b \wedge c$ si et seulement si il minore b et il minore c, ce qui donne $F(b \wedge c) = F(b) \cap F(c)$. Appliquant ceci à $\overline{b}$ et $\overline{c}$, ainsi que l'égalité $F(\overline{x}) = F(x)^c$, nous obtenons : $F(b \vee c) = F(\overline{\overline{b} \vee \overline{c}})^c = F(\overline{b} \wedge \overline{c})^c = (F(\overline{b}) \cap F(\overline{c}))^c = F(\overline{b})^c \cup F(\overline{c})^c = F(b) \cup F(c)$. A ce point, on a donc montré que F est un homomorphisme de $(B, \vee, \wedge, 0, 1, \overline{\ })$ dans $(\mathfrak{P}(A), \cup, \cap, \emptyset, A, {}^c)$.

Il reste à montrer que F est bijectif. Soient b et c deux éléments distincts de B. L'une au moins des deux relations $b \leqslant c$, $c \leqslant b$ est fausse. Supposons par exemple $b \not\leqslant c$, donc $b \neq b \wedge c$. Comme on a $b = b \wedge 1 = b \wedge (c \vee \overline{c}) = (b \wedge c) \vee (b \wedge \overline{c})$, on doit avoir $b \wedge \overline{c} \neq 0$, donc $F(b \wedge \overline{c}) \neq \emptyset$. Il existe donc un atome a minorant b et $\overline{c}$, donc ne minorant pas c, c'est-à-dire appartenant à $F(b)$ et non à $F(c)$: ces ensembles sont donc distincts,et F est injectif.

Finalement, soit X un sous-ensemble quelconque de A. Puisque B est fini, A l'est aussi, et on peut écrire $X = \{a_1, \dots, a_n\}$. Posons $b = a_1 \vee \cdots \vee a_n$ (comme $\vee$ est une opération associative, il n'y a pas d'ambiguïté à supprimer les parenthèses). Soit a un atome quelconque. Par distributivité de $\wedge$ vis-à-vis de $\vee$, on a $a \wedge b = (a \wedge a_1) \vee \cdots \vee (a \wedge a_n)$: si a est l'un des a_i, on obtient $a \wedge b = a$, soit $a \leqslant b$, ou $a \in F(b)$; sinon, on obtient $a \wedge b = 0$, donc $a \notin F(b)$. On a donc $F(b) = X$, et F est surjective. □

▷ *La proposition 2.13 montre que la notion d'algèbre de Boole capture toutes les propriétés des ensembles $\mathfrak{P}(A)$ finis : tant qu'on ne s'intéresse qu'à des ensembles finis, les axiomes de la proposition 2.9 caractérisent complètement les opérations ensemblistes. En un sens, ce résultat clôt la partie élémentaire de la théorie des ensembles, celle qui se concentre sur les manipulations d'union, d'intersection et de complémentaire dans le cas fini, et il explique que la partie non triviale de la théorie concerne surtout l'étude des ensembles infinis*[4]*. La situation est complètement différente avec les algèbres de Boole infinies, et, par exemple, le quotient de $\mathfrak{P}(\mathbb{N})$ par l'idéal des ensembles finis a une structure très riche.* ◁

3. Ébauche d'une théorie des ensembles

► On cherche à construire une théorie formelle des ensembles. Constatant la difficulté de définir les ensembles, on recourt à une démarche axiomatique basée sur les principes d'extensionnalité et de compréhension. Les paradoxes de Berry et de Russel obligent à affiner l'approche, et on parvient au système de Zermelo comme axiomatisation possible des ensembles purs. ◄

▷ *On a développé dans les parties précédentes quelques résultats et démonstrations simples concernant les ensembles, sans avoir introduit ceux-ci autrement que de façon très informelle. Si on veut progresser et développer une théorie élaborée, il est nécessaire de fixer un point de départ plus formel.*

La première étape semble devoir être de définir les ensembles précisément, y compris les ensembles infinis puisque les résultats de la section 2 ont suggéré que c'était à ce niveau que les questions profondes se posent. C'est ce qu'on se propose de faire dans cette partie. L'analyse au demeurant restera incomplète, dans la mesure où elle nous mènera à considérer une famille d'ensembles particuliers, les ensembles purs, dont il n'est pas évident que l'étude soit pertinente.

[4] ce qui, bien sûr, n'est pas dire que la combinatoire finie est une discipline triviale

Ce sera la tâche des chapitres II et III de montrer que la restriction aux ensembles purs ne limite en rien le champ d'application de la théorie, et de légitimer ainsi les options prises dans ce chapitre. ◁

3.1. Une tentative naïve.

▶ Comme pour n'importe quel autre type d'objet mathématique, il est naturel de débuter une théorie des ensembles par une définition des ensembles. Ceci n'est pas impossible — c'est même l'option prise dans la plupart des langages de programmation — mais requiert que toutes les propriétés des fonctions soient garanties. ◀

▷ *Les objets mathématiques relèvent de types divers: entiers, points, droites, fonctions, etc. Dans une approche élémentaire, et notamment dans les langages de programmation informatiques, les ensembles n'apparaissent pas comme un type de base unique, mais plutôt comme des types dérivés : pour chaque type mathématique τ, on introduit un nouveau type $\mathbf{Ens}_\tau$ formé par les ensembles d'objets de type τ, de la même façon qu'on peut introduite d'autres types voisins comme les suites (analogues aux ensembles, mais en tenant compte de l'ordre des facteurs), ou les multi-ensembles (analogues aux ensembles, mais en tenant compte d'éventuelles répétitions).*

La première question est donc de définir le type $\mathbf{Ens}_\tau$. Si le type « fonction » est présent, précisément si, pour chaque paire de types τ, τ', il existe un type formé des fonctions allant des objets de type τ vers les objets de type τ', alors on peut identifier les ensembles à des fonctions indicatrices : se donner un ensemble A d'objets de type τ, c'est spécifier, pour chaque objet de type τ, s'il est ou non dans A, donc se donner une fonction associant à tout objet de type τ soit la valeur VRAI*, soit la valeur* FAUX*.* ◁

Dans ce chapitre, on utilise la notation $x{:}\tau$ pour indiquer qu'un objet x est de type τ : par exemple, $x{:}\mathbf{Ent}$ indique que x est un entier. Si τ, τ' sont des types, on note $\tau{\to}\tau'$ le type des fonctions allant des objets de type τ vers les objets de type τ'. Par ailleurs, on note **Bool** (comme booléen) le type constitué de ces deux seules valeurs VRAI et FAUX.

« DÉFINITION » 3.1. (ensemble) Pour tout type τ, le type $\tau{\to}\mathbf{Bool}$ est noté $\mathbf{Ens}_\tau$, et les objets de type $\mathbf{Ens}_\tau$ sont appelés *ensembles* d'objets de type τ. Pour x de type τ, et A de type $\mathbf{Ens}_\tau$, on dit que x est *élément* de A, noté $x \in_\tau A$, si on a $A(x) =$ VRAI.

On peut alors commencer à établir des propriétés des ensembles, et en particulier montrer qu'ils obéissent aux principes informels dégagés dans la première partie de ce chapitre :

« PROPOSITION » 3.2. (propriétés) (*i*) *Un objet de type $\mathbf{Ens}_\tau$ est déterminé par ses éléments.*
(*ii*) *Pour tous $a_1, \ldots, a_n$ de type τ, il existe un objet de type $\mathbf{Ens}_\tau$ ayant $a_1, \ldots, a_n$ pour éléments.*
(*iii*) *Pour chaque propriété $\mathcal{P}(x{:}\tau)$ des objets de type τ, il existe un objet de type $\mathbf{Ens}_\tau$ dont les éléments sont exactement les éléments de type τ satisfaisant $\mathcal{P}$.*

DÉMONSTRATION. Le point (*i*) découle de ce que le type **Bool** ne contient que deux valeurs: si A, A' sont deux ensembles avec les mêmes éléments, on a $A(x) = A'(x) =$ VRAI pour tout x de type τ appartenant à A, et donc $A(x) = A'(x) =$ FAUX pour tout autre x de type τ. Les fonctions A et A' coïncident donc (pour autant que deux fonctions prenant les mêmes valeurs

partout coïncident). Pour (*ii*), on définit un objet A de type $\mathbf{Ens}_\tau$ en posant $A(a_1) = \ldots = A(a_n) =$ VRAI, et $A(x) =$ FAUX pour tout autre objet x de type τ. Pour (*iii*), on définit de même $A{:}\tau{\to}\mathbf{Bool}$ par $A(x) =$ VRAI si $\mathcal{P}(x)$ est vraie, et $A(x) =$ FAUX sinon. □

▷ *On doit bien sentir que ce qui précède n'est pas satisfaisant, et en tout cas pas suffisant. Même formellement acceptable, la définition 3.1 ne fait que reporter la définition des ensembles sur celle des fonctions, laquelle n'est pas donnée — et, si on se rappelle qu'une fonction est souvent définie comme ensemble de couples, on sent poindre le cercle vicieux. De même, la démonstration de la proposition 3.2 n'est qu'une vérification de la cohérence du vocabulaire : faute d'avoir indiqué comment spécifier une fonction, l'existence des fonctions mentionnées n'est en rien établie.* ◁

3.2. Le système de Cantor.

► A défaut de définir les ensembles, on cherche à les axiomatiser. Le système de Cantor repose sur deux principes : l'axiome d'extensionnalité qui affirme qu'un ensemble est déterminé par ses éléments, et l'axiome de compréhension qui affirme que toute propriété donne naissance à un ensemble. ◄

▷ *Le problème rencontré ci-dessus est usuel: qu'il s'agisse de débuter l'arithmétique, la géométrie, ou toute autre théorie concernant des objets basiques, on bute sur la définition des objets premiers. Or, même si la nature en soi des objets mathématiques peut être importante pour le philosophe, elle n'influe pas directement sur les démonstrations et ne concerne donc que peu le mathématicien : peu importe ce que* sont *les nombres entiers, ce qui lui importe pour démontrer de nouveaux théorèmes est de savoir comment ils se comportent, c'est-à-dire quelles sont leurs propriétés. On peut donc se contenter d'une approche axiomatique, consistant, à défaut de définir les objets qu'on souhaite étudier, à en énumérer des propriétés de base, puis à déduire de celles-ci, utilisées comme axiomes, des conséquences nouvelles. On sait par exemple que le système de Peano constitue un point de départ raisonnable pour l'arithmétique, tout comme le système d'Euclide en constitue un pour la géométrie.*

On se propose donc de développer ce type d'approche axiomatique pour les ensembles. Dans un premier temps, il s'agit de recenser les propriétés de base des ensembles, celles qu'on retiendra comme axiomes. On reviendra vers la fin de l'ouvrage sur les questions délicates de choix d'axiomes — questions qui ne peuvent faire l'objet que de consensus et non de démonstrations — mais, pour le moment, il est aisé de débuter en s'appuyant sur l'analyse effectuée dans la première partie du chapitre, qui assigne aux ensembles deux types de propriétés de base — celles-là même qu'on a « démontrées » dans la proposition 3.2. Le premier est un principe d'unicité, à savoir qu'un ensemble est déterminé par ses éléments. Le second est un principe d'existence, à savoir qu'on peut spécifier un ensemble soit en énumérant ses éléments, soit en donnant une propriété caractéristique de ceux-ci. ◁

DÉFINITION 3.3. (extensionnalité, extension, compréhension) On appelle *axiome d'extensionnalité* pour les objets de type τ l'assertion

$$\forall A, A'{:}\mathbf{Ens}_\tau\,(A = A' \Leftrightarrow \forall x{:}\tau\,(x \in A \Leftrightarrow x \in A')). \tag{3.1}$$

Pour $a_1, \ldots, a_n$ de type τ, on appelle *axiome d'extension* pour $a_1, \ldots, a_n$ l'assertion

$$\exists A{:}\mathbf{Ens}_\tau\ \forall x{:}\tau\ \ (x \in A \Leftrightarrow (x = a_1 \text{ ou } \ldots \text{ ou } x = a_n)). \tag{3.2}$$

Pour $\mathcal{P}(x)$ propriété faisant sens pour les objets de type τ, on appelle *axiome de compréhension* en $\mathcal{P}$ l'assertion

$$\exists A{:}\mathbf{Ens}_\tau\ \forall x{:}\tau\ \ (x \in A \Leftrightarrow \mathcal{P}(x) \text{ est vraie}). \tag{3.3}$$

Lorsque (3.1) est satisfait, les ensembles dont (3.2) et (3.3) affirment l'existence sont notés — ainsi qu'on l'a déjà dit — respectivement $\{x_1, \dots, x_n\}$ et $\{x{:}\tau\,;\,\mathcal{P}(x)\}$ (Figure 6). Remarquer qu'on peut s'affranchir des axiomes d'extension, qui n'ont été mentionnés que pour suivre l'usage : une définition par extension est un cas particulier de définition par compréhension puisque, pour $x_1, \dots, x_n$ de type τ, on a $\{x_1, \dots, x_n\} = \{x{:}\tau; x = x_1 \text{ ou } \dots \text{ ou } x = x_n\}$.

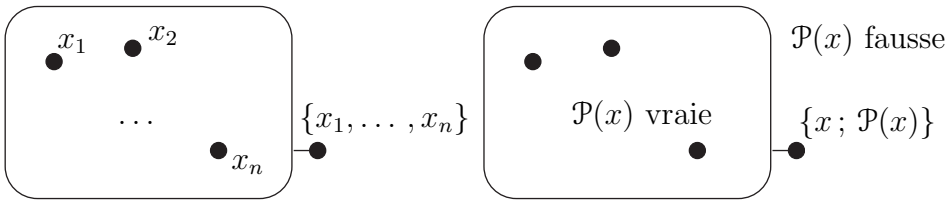

FIGURE 6. Deux façons usuelles de spécifier un ensemble: par extension, c'est-à-dire en énumérant les éléments (supposés en nombre fini), et par compréhension, c'est-à-dire en donnant une propriété caractéristique des éléments

▷ *L'idée à ce point est donc de poser que, pour chaque type τ, les objets de type $\mathbf{Ens}_\tau$ obéissent aux axiomes d'extensionnalité, d'extension, et de compréhension, et d'étudier les conséquences de ces axiomes. Nous sommes alors à peu près au point de départ proposé par Cantor vers 1890.* ◁

3.3. Le paradoxe de Berry.

► Le système précédent n'est pas tenable : il postule l'existence d'objets qui se révèlent contradictoires. On restreint donc le champ d'application des axiomes de compréhension aux propriétés exprimables dans une logique formelle. ◄

PROPOSITION 3.4. (paradoxe de Berry) *Soit $\mathcal{P}(n)$ la propriété « n est un entier définissable par une phrase française d'au plus cent caractères ». Alors il ne peut exister d'ensemble des entiers possédant la propriété $\mathcal{P}$.*

DÉMONSTRATION. Soit A l'ensemble $\{n{:}\mathbf{Ent}\,;\,\mathcal{P}(n)\}$, supposé exister. Il n'y a, en comptant les blancs, qu'au plus 27^{100} phrases françaises d'au plus 100 caractères, et chaque telle phrase ne peut définir qu'au plus un entier, puisque dire qu'une phrase définit n signifie que n est le seul entier satisfaisant la propriété exprimée par la phrase. Par conséquent, l'ensemble A a au plus 27^{100} éléments, et son complémentaire est non vide. L'ordre des entiers est un bon ordre, c'est-à-dire que tout ensemble non vide d'entiers possède un plus petit élément (*cf.* chapitre II), et donc le complémentaire de A doit posséder un plus petit élément, soit n_0. Puisqu'il appartient à A par hypothèse, l'entier n_0 est non définissable par une phrase française d'au plus 100 caractères. Mais « je suis le plus petit entier non définissable par une phrase française d'au plus cent caractères » est une définition pour n_0, qui comporte 96 caractères. Cette contradiction montre que l'hypothèse que A existe est à rejeter. □

▷ *Le système de Cantor n'est donc pas tenable, puisqu'il postule l'existence d'objets qui se révèlent contradictoires. On est donc mené à le modifier le système en espérant échapper aux contradictions. Renoncer à considérer des ensembles d'entiers est difficile, dans la mesure où le*

type « entier » est l'un des premiers pour lequel on souhaite construire des ensembles. Renoncer aux définitions par compréhension, et se cantonner aux définitions par extension, est une solution drastique — c'est plus ou moins le point de vue informatique — qui ôte à la notion d'ensemble l'essentiel de son intérêt mathématique. Ne reste donc qu'à restreindre le champ des propriétés permises. On sent bien que le paradoxe de Berry tient à ce que la propriété « être définissable par une phrase française d'au plus cent caractères » n'est pas une véritable propriété mathématique, parce qu'elle fait appel à la notion de phrase française, laquelle ne correspond à aucune définition précise. La solution naturelle est de restreindre le champ de la compréhension à des propriétés exprimables dans un langage assez souple pour laisser le plus de richesse d'expression possible, mais assez restrictif pour échapper au paradoxe de Berry.

Par exemple, on tient à l'existence d'un ensemble des entiers qui sont somme de deux carrés. Or une différence claire entre la propriété « être définissable par une phrase française d'au plus cent caractères » et la propriété « être somme de deux carrés » est que la seconde s'exprime par la formule

$$\exists p, q\,(n = p \times p + q \times q), \tag{3.4}$$

là où une traduction de la première est problématique. L'idée est de ne considérer désormais que des propriétés qui peuvent s'exprimer par des formules comme (3.4).

Il existe de nombreuses logiques formelles (voir partie B), et donc de nombreuses options sont possibles à ce stade. Pour le moment, on ne va pas entrer dans une discussion précise, mais simplement délimiter un peu le contexte, les définitions formelles étant renvoyées au chapitre VII. Ce qui importe ici est de savoir quelles formules peuvent être légitimement utilisées dans des définitions par compréhension. Le principe retenu est de faire appel aux formules dites du premier ordre à un seul type d'objet. Ces formules sont essentiellement les formules mathématiques usuelles, écrites avec des variables et des quantificateurs divers, soumises à quelques contraintes additionnelles qu'on va décrire maintenant.

Le cadre général consiste à fixer une signature *consistant en une liste de types et d'opérations et de relations susceptibles de relier des objets de ces types, puis à définir inductivement les formules en la signature Σ comme des suites finies de symboles qui sont soit des variables avec indication de type, soit des opérations et relation de Σ, soit le signe $=$, soit des connecteurs logiques (non, et, ou, $\Rightarrow$, $\Leftrightarrow$), des quantificateurs ($\exists$, $\forall$), ou des parenthèses. Les règles de construction seront données — ou plutôt rappelées : il s'agit ni plus ni moins des uages mathématiques — plus loin ; ici, il sera suffisant de mentionner que*

$$((\forall == n)))(+ + 01 \times p2 \Rightarrow\Rightarrow$$

n'est pas une formule, parce que les symboles n'y sont pas assemblés dans un ordre correct, alors que

$$\forall n{:}\mathbf{Ent}\ \exists p, q{:}\mathbf{Ent}\ \ (n = p \times p + q \times q) \tag{3.5}$$

en est une vis-à-vis de la signature Σ_1 comportant un unique type d'objet **Ent**, *et deux symboles d'oppération binaires $+$ et $\times$ — ainsi que vis-à-vis de toute signature incluant Σ_1. On notera que la formule* (3.5) *est considérée comme parfaitement légitime alors que la propriété qu'elle exprime est fausse : il existe des entiers qui ne sont pas somme de deux carrés. C'est l'occasion d'affirmer déjà une distinction qui sera très importante dans la partie B, à savoir qu'une formule n'est un objet syntaxique, un mot donc, et que l'éventuelle valeur de vérité qu'on peut lui attacher n'apparaît que relativement à un contexte extérieur d'évaluation qui n'est pas présent dans la formule : par exemple,* (3.5) *est fausse lorsque* **Ent** *réfère au type « entier », et que les symboles $+$ et $\times$ réfèrent à l'addition et à la multiplication des entiers.*

On parle spécifiquement de formules du premier ordre lorsque les seules variables portent sur les éléments des types déclarés dans Σ, à l'exclusion des ensembles de tels éléments ; par opposition, on parle de formules du second ordre si on autorise des variables et des quantifications de type τ et $\mathbf{Ens}_\tau$ et l'usage de $\in_\tau$, du troisième type si de même on autorise τ, $\mathbf{Ens}_\tau$, $\mathbf{Ens}_{\mathbf{Ens}_\tau}$ for chaque type τ de Σ, etc. Par exemple, la formule (3.5) *est du premier ordre par rapport à la signature Σ_1, alors que la formule*

$$\forall A{:}\mathbf{Ens}_{\mathbf{Ent}}\,((0 \in A \text{ et } \forall n{:}\mathbf{Ent}\,(n \in A \Rightarrow n+1 \in A)) \Rightarrow \forall n{:}\mathbf{Ent}\,(n \in A)) \tag{3.6}$$

— qu'exprime cette formule ? — est du second ordre par rapport à la signature Σ_1' obtenue à ajoutant à Σ_1 les deux symboles de constante 0 *et* 1. *Noter par contre que* (3.6) *est du premier*

ordre par rapport à la signature Σ_2 comportant les deux types d'objets **Ent** *et* $\mathbf{Ens}_{\mathbf{Ent}}$ *et, outre les symboles de Σ_1, le symbole de relation $\in$ entre objets de types* **Ent** *et* $\mathbf{Ens}_{\mathbf{Ent}}$.

Ayant ainsi (informellement) défini les formules du premier ordre relatives à une signature, donc relatives à un (ou des) type d'objet et à un choix d'opérations ou relations spécifiques à ces types, et en considérant comme intuitive (?) la notion de satisfaction d'une formule $\mathsf{F}(x)$ par un objet a de type τ, on peut revenir à la construction des ensembles, et réintroduire l'axiome de compréhension sous une forme restreinte. ◁

DÉFINITION 3.5. (compréhension, version réduite) Pour $\mathsf{F}(x, x_1, \dots, x_n)$ formule du premier ordre en une signature comportant l'unique type τ [5], on appelle *axiome de compréhension* en F l'assertion

$$\forall a_1, \dots, a_n{:}\tau\ \exists A{:}\mathbf{Ens}_\tau\ \forall x{:}\tau\,(x \in A \Leftrightarrow \mathsf{F}(x, a_1, \dots, a_n)). \tag{3.7}$$

▷ *Le nouveau système ainsi obtenu est un sous-système du système initial de Cantor. On l'appellera ici système de Frege, du nom d'un des pionniers de la logique formelle qui a proposé un tel système vers 1893.*

On notera qu'on a fait figurer explicitement dans (3.7) *des éventuels paramètres $a_1, \dots, a_n$ figurant dans la formule F, mais ne correspondant pas nécessairement à des symboles de constante de la signature Σ. Par exemple, si nous considérons le type « entier » et si Σ est la signature réduite aux quatre opérations arithmétiques de base, la formule $\exists p\,(n = p + 3$ ou $n = 5 \times p)$ est une formule du premier ordre en Σ avec un variable libre, à savoir n, et deux paramètres, à savoir les entiers 3 et 5. L'axiome de compréhension associé permet alors d'affirmer l'existence d'un ensemble tel que*

$$\{n{:}\mathbf{Ent}\,;\, \exists p\,(n = p + 3 \text{ ou } n = 5 \times p)\},$$

ainsi que le réclame l'usage.

Le système de Frege échappe au paradoxe de Berry, tout au moins sous la forme où on l'a énoncé: la conclusion de la proposition 3.4 devient simplement que la propriété d'être définissable par une phrase française d'au plus cent caractères n'est pas exprimable par une formule du premier ordre [6]. ◁

Quelle que soit la signature choisie, les formules $x = x$ et $x \neq x$ (négation de $x = x$) sont des formules du premier ordre en cette signature. Appliquant les axiomes de compréhension associés, on obtient, pour chaque type τ, deux ensembles particuliers :

DÉFINITION 3.6. (plein, vide) Soit τ un type quelconque. L'ensemble *plein* (*resp. vide*) de type τ est l'ensemble $\{x{:}\tau\,;\, x = x\}$ (*resp.* $\{x{:}\tau\,;\, x \neq x\}$); on le note Ω_τ (*resp.* $\emptyset_\tau$).

Certaines notations sont traditionnelles : $\mathbb{N}$ pour l'ensemble $\Omega_{\mathbf{Ent}}$ de tous les entiers, $\mathbb{R}$ pour l'ensemble $\Omega_{\mathbf{Reel}}$ de tous les réels, *etc.* Les définitions impliquent que, dans tous les cas, être un objet de type τ est équivalent à appartenir à l'ensemble Ω_τ. Par conséquent, la notation $\{x{:}\tau\,;\, \mathsf{F}(x)\}$ peut être remplacée par la notation équivalente $\{x \in \Omega_\tau\,;\, \mathsf{F}(x)\}$, comme dans $\{n \in \mathbb{N}\,;\, \exists p, q\,(n = p^2 + q^2)\}$.

[5] *A priori*, rien n'oblige à se restreindre ici à une signature à un seul type d'objet, mais, *de facto*, c'est l'option qui sera retenue dans la suite du texte.

[6] La phrase précédente reste vague car l'exprimabilité par une formule du premier ordre est relative au choix d'une signature. En l'occurrence, il s'agirait de tout choix d'opérations et de relations sur les entiers pour lequel le principe de récurrence s'applique, impliquant que tout ensemble non vide a un plus petit élément

3.4. Le paradoxe de Russell.

▶ Le système de Frege se révèle à son tour contradictoire, à cause du paradoxe de Russell sur l'ensemble de tous les ensembles. On est donc amené à de nouvelles restrictions, remplaçant les axiomes de compréhension généraux par les axiomes de séparation, qui en sont des cas particuliers. ◀

▷ *Quel que soit le type τ, les objets de type $\mathbf{Ens}_\tau$ ont en commun la propriété d'être des ensembles, et, à ce titre, ils partagent un certain nombre de propriétés. De même, les relations d'appartenance relatives aux divers types peuvent être considérées comme des restrictions d'une unique relation d'appartenance générale $\in$. Il apparaît donc naturel d'introduire un type « ensemble » général* **Ens** *englobant tous les types particuliers $\mathbf{Ens}_\tau$, en espérant en particulier qu'il permette d'uniformiser l'étude des divers types d'ensembles a priori distincts. On se trouvera ainsi en particulier libéré d'un contexte de type qui, pour intuitif qu'il est, n'a pas été défini rigoureusement.*

Les problèmes surviennent rapidement. En effet, s'il existe un type **Ens** *dont relèvent tous les ensembles, et si l'axiome de compréhension est valide pour les formules contenant la relation d'appartenance, alors une contradiction apparaît avec l'objet $\Omega_{\mathbf{Ens}}$, c'est-à-dire avec l'ensemble de tous les ensembles.* ◁

PROPOSITION 3.7. (paradoxe de Russell) *L'existence d'un ensemble de tous les ensembles non éléments d'eux-mêmes est une hypothèses contradictoire.*

DÉMONSTRATION. Supposons $A = \{X{:}\mathbf{Ens}\,;\, X \notin X\}$ [7], c'est-à-dire supposons que A est un ensemble tel que, pour tout ensemble X, on ait l'équivalence $X \in A \Leftrightarrow X \notin X$. Alors, en particulier, $X \in A$ est soit vrai, soit faux pour chaque ensemble X et, A étant lui-même un ensemble, l'assertion $A \in A$ doit être soit vraie, soit fausse. Or $A \in A$ entraînerait $A \notin A$ par définition de A, et, de même, $A \notin A$ entraînerait $A \in A$. Chacune des deux possibilités étant contradictoire, c'est que l'existence de A est contradictoire. □

Une autre version de la même difficulté apparaît pour le type **Ens** lorsqu'on le compare avec le sous-type $\mathbf{Ens}_{\mathbf{Ens}}$ formé par les ensembles d'ensembles. Par construction, $\mathbf{Ens}_{\mathbf{Ens}}$ est un sous-type de **Ens**, ce qui revient à dire qu'il existe une injection de l'ensemble $\Omega_{\mathbf{Ens}_{\mathbf{Ens}}}$, qui est aussi $\mathfrak{P}(\Omega_{\mathbf{Ens}})$, dans l'ensemble $\Omega_{\mathbf{Ens}}$, ce qui contredit le théorème de Cantor (proposition 1.10).

▷ *Deux solutions s'offrent pour échapper au paradoxe de Russell. La première est de renoncer à introduire un type « ensemble » général et de s'en tenir à un univers typé dans lequel on distingue des objets de base, puis des ensembles d'objets de base, puis des ensembles d'ensembles d'objets de base, etc. De la sorte, la relation d'appartenance ne fait sens qu'entre un objet de type τ et un objet de type $\mathbf{Ens}_\tau$, et des formules comme $X \in X$ ou $X \notin X$ n'ont pas de sens, donc ne peuvent être ni vraies ni fausses. Ce point de vue, qui est celui de la théorie de types de Russell, mène à un développement assez compliqué du fait de l'hétérogénéité des objets introduits, et la difficulté de maniement d'un tel système a, jusqu'à présent, limité les résultats en termes de théorie des ensembles proprement dite, par exemple dans l'analyse du problème du continu.*

Par ailleurs, même dans le cadre d'une théorie particulière, par exemple l'arithmétique, il est souvent utile de considérer des ensembles d'entiers dont la définition fait elle-même appel à des ensembles d'entiers, utilisant ainsi des formules qui, par rapport à la signature de l'arithmétique, ne sont pas du premier ordre. Si une relation d'appartenance générale est disponible, de telles définitions peuvent être formalisées, par contre c'est moins simple avec une théorie typée restreinte.

La seconde solution, qui est celle retenue par la théorie des ensembles classique, consiste à restreindre à nouveau le champ d'application de la compréhension pour échapper au paradoxe

[7] Dans toute la suite, $x \notin y$ dénote la négation de $x \in y$.

de Russell. L'idée est d'attribuer le paradoxe au fait que l'ensemble de tous les ensembles est un objet trop grand pour être un ensemble, et de réserver l'appellation d'ensemble à ceux des objets définis par compréhension qui sont, en un sens à préciser, assez petits.

Pratiquement, le principe est de renoncer à la forme générale de l'axiome de compréhension, c'est-à-dire de ne plus postuler l'existence, pour chaque formule $\mathsf{F}(x, x_1, \dots, x_n)$ *et chaque choix de* $a_1, \dots, a_n$*, de l'ensemble* $\{x\,;\,\mathsf{F}(x, a_1, \dots, a_n)\}$*, pour ne le conserver que pour les formules du type* $x \in A$ et $\mathsf{F}(x, a_1, \dots, a_n)$*, c'est-à-dire de postuler, pour chaque ensemble* A *et chaque formule, l'existence de l'ensemble*

$$\{x\,;\,x \in A \text{ et } \mathsf{F}(x, a_1, \dots, a_n)\}.$$

Il ne s'agit donc plus de former un ensemble ex nihilo, mais simplement de séparer *à l'intérieur d'un ensemble* A *préexistant les éléments qui vérifient* F*, et on parlera donc d'*axiome de séparation *pour ce cas particulier d'axiome de compréhension.* ◁

Définition 3.8. (séparation) Pour $\mathsf{F}(x, x_1, \dots, x_n)$ formule du premier ordre en une signature comportant l'unique type τ, on appelle *axiome de séparation* en F l'assertion

(3.8)

$$\forall a_1, \dots, a_n{:}\tau\ \forall A{:}\mathbf{Ens}_\tau\ \exists B{:}\mathbf{Ens}_\tau\ \forall x{:}\tau\ (x \in B \Leftrightarrow (x \in A \text{ et } \mathsf{F}(x, a_1, \dots, a_n))).$$

On note $\{x \in A\,;\,\mathsf{F}(x, a_1, \dots, a_n)\}$ l'ensemble B dont l'existence est affirmée par l'axiome de séparation (3.8)

▷ *Pour un type* τ *donné, deux situations sont alors possibles. Ou bien il existe un ensemble* Ω_τ *formé par tous les objets de type* τ*, et alors compréhension et séparation mènent aux mêmes définitions puisque tout ensemble* $\{x{:}\tau\,;\,\mathsf{F}(x)\}$ *défini par compréhension est aussi défini par séparation comme* $\{x \in \Omega_\tau\,;\,\mathsf{F}(x)\}$*. Ou bien un tel ensemble n'existe pas, et alors le champ des définitions est restreint. C'est par exemple le cas du type « ensemble ».* ◁

Proposition 3.9. (ensemble de tous les ensembles) *Les objets de type « ensemble » ne forment pas un ensemble.*

Démonstration. S'il existait un ensemble de tous les ensembles, alors, appliquant l'axiome de séparation associé à la formule $x \notin x$, on déduirait l'existence d'un ensemble de tous les ensembles qui ne sont pas éléments d'eux-mêmes, contredisant la proposition 3.7. □

3.5. Paire, union, et parties.

► Les axiomes de séparation ne permettent pas de débuter la construction des ensembles. On réintroduit donc certains axiomes de compréhension particuliers, les axiomes de la paire, de l'union, et des parties. ◄

▷ *En prenant comme point de départ l'axiome d'extensionnalité et la famille infinie de tous les axiomes de séparation en chacune des formules du premier ordre — en une signature et dans un contexte de type qui restent à spécifier — on se met à l'abri des paradoxes de Berry et de Russell. Mais une nouvelle difficulté apparaît: à la différence des axiomes de compréhension, les axiomes de séparation ne permettent pas de construire des ensembles ex nihilo. Par exemple, ils ne permettent même pas de garantir l'existence des ensembles définis par extension, pourtant réclamée par l'intuition et la pratique mathématique. On est donc conduit à réintroduire explicitement les définitions par extension et, plus généralement, des opérations ensemblistes de base dont la séparation seule ne garantirait pas qu'elles soient partout définies.*

Plutôt que de poser un axiome pour chaque définition par extension, il est usuel de se contenter d'un axiome posant l'existence de paires, c'est-à-dire autorisant les définitions par extension d'ensembles à deux éléments au plus, et d'un axiome général pour l'union d'une

famille d'ensembles. Par ailleurs, à partir du moment où on adopte les axiomes de séparation, il suffit, pour garantir par exemple l'existence d'une paire $\{a, b\}$, d'être assuré de l'existence d'un ensemble A contenant a et b, puisqu'ensuite l'axiome de séparation associé à la formule $x = a$ ou $x = b$ permet de séparer dans A la paire $\{a, b\}$. ◁

DÉFINITION 3.10. (paire, union) On appelle axiomes *de la paire* et *de l'union* pour le type τ les assertions

$$\forall a, b{:}\tau\ \exists A{:}\mathbf{Ens}_\tau\ (a \in A \text{ et } b \in A), \tag{3.9}$$

$$\forall A{:}\mathbf{Ens}_{\mathbf{Ens}_\tau}\ \exists B{:}\mathbf{Ens}_\tau\ \forall x{:}\tau\ (\exists X{:}\mathbf{Ens}_\tau(x \in X \text{ et } X \in A) \Rightarrow x \in B). \tag{3.10}$$

En présence des axiomes d'extensionnalité, qui garantit l'unicité, et de séparation, qui permettent d'extraire les ensembles souhaités, on note respectivement $\{a, b\}$ et $\bigcup A$ l'unique ensemble dont les éléments sont a et b, et l'unique ensemble dont les éléments sont les éléments des éléments de A.

▷ *Les axiomes précédents légitiment les définitions par extension dans le cas d'ensembles à un ou deux éléments. En appliquant l'axiome de l'union à une paire d'ensembles $\{A, B\}$, on obtient la réunion $A \cup B$ de A et de B, c'est-à-dire l'ensemble des éléments qui sont dans A ou dans B. La formulation plus générale donnée ici permet d'introduire l'union de familles quelconques d'ensembles, et pas seulement celle de deux ensembles. Pour ce qui est des définitions par extension, on peut alors les légitimer sans avoir à introduire de nouvel axiome.* ◁

LEMME 3.11. *L'axiome d'extension pour le type τ est conséquence des axiomes d'extensionnalité, de la paire, de l'union, et de séparation pour le types τ et de l'axiome de la paire pour le type $\mathbf{Ens}_\tau$.*

DÉMONSTRATION. On montre d'abord par récurrence sur n que, pour toute famille $a_1, ..., a_n$ d'objets de type τ, il existe un ensemble contenant $a_1, ..., a_n$. On procède par récurrence sur n. Pour $n = 1$ et $n = 2$, cela résulte de l'axiome de la paire pour le type τ. Supposons $n \geqslant 3$. L'hypothèse de récurrence garantit l'existence d'un ensemble A de type $\mathbf{Ens}_\tau$ contenant $a_1, ..., a_{n-1}$, et l'axiome de la paire pour le type τ garantit celle d'un ensemble B de même type contenant a_n. Alors l'axiome de la paire pour le type $\mathbf{Ens}_\tau$ garantit l'existence d'un ensemble C de type $\mathbf{Ens}_{\mathbf{Ens}_\tau}$ contenant A et B, puis l'axiome de l'union pour le type τ garantit celle d'un ensemble D de type $\mathbf{Ens}_\tau$ contenant les éléments des éléments de C, donc en particulier les éléments de A et ceux de B, soit $a_1, ..., a_{n-1}$ d'une part et a_n d'autre part. Finalement, l'existence de l'ensemble $\{a_1, ..., a_n\}$ s'obtient en séparant dans un ensemble quelconque contenant $a_1, ..., a_n$ les éléments x vérifiant $x = a_1$ ou ... ou $x = a_n$, et son unicité résulte de l'axiome d'extensionnalité pour le type τ. □

▷ *Si A est un ensemble, les axiomes de séparation ne permettent pas de sortir de A, et, en particulier, rien ne permet a priori d'affirmer l'existence d'un ensemble de toutes les parties de A. On est donc conduit à considérer un nouvel axiome affirmant l'existence d'un ensemble des parties $\mathfrak{P}(A)$ pour tout ensemble A ou, ce qui revient au même en présence des axiomes de séparation, d'un ensemble contenant toutes les parties de A. Le point de vue adopté ici, dit imprédicatif, consiste à considérer toutes les parties de A, que celles-ci soient définies ou non par des formules ou tout autre moyen explicite. La distinction est importante, car, par exemple, l'ensemble $\mathfrak{P}(\mathbb{N})$ de toutes les parties de $\mathbb{N}$ est non dénombrable, alors que l'ensemble des parties de $\mathbb{N}$ qui sont définissables, par exemple dans le langage de l'arithmétique, est dénombrable, car la famille des formules pouvant servir de définition l'est. Le principe ici est de choisir le cadre le plus large possible, à l'intérieur duquel pourront être développées d'autres théories plus restrictives, comme une théorie prédicative des ensembles où on se restreindrait aux parties définissables.* ◁

DÉFINITION 3.12. (parties) On appelle axiome *des parties* pour le type τ l'assertion

$$\forall A{:}\mathbf{Ens}_\tau\ \exists B{:}\mathbf{Ens}_{\mathbf{Ens}_\tau}\ \forall X{:}\mathbf{Ens}_\tau\,(X \subseteq A \Rightarrow X \in B). \tag{3.11}$$

En présence des axiomes d'extensionnalité et de séparation, l'axiome des parties garantit l'existence d'un unique ensemble dont les éléments sont les ensembles inclus dans A ; ainsi qu'on l'a dit dans la section 2, cet ensemble est noté $\mathfrak{P}(A)$.

▷ *En restreignant la compréhension au cas particulier des axiomes de séparation, on espère être à l'abri des paradoxes, et, en réintroduisant les axiomes de la paire, de l'union et des parties, disposer de suffisamment d'ensembles pour que la (ou les) théorie ainsi introduite ait du corps. C'est en tout cas sur ces bases qu'on se propose de développer dans la suite la théorie des ensembles.* ◁

3.6. Ensembles purs et système de Zermelo.

► On introduit le système de Zermelo fini $\mathsf{Z}_{\mathsf{fini}}$ comme la spécification des axiomes précédents correspondant au choix d'une signature à un seul type d'objet, les ensembles, et une seule relation, l'appartenance. Ce système est proposé comme première approximation d'une axiomatisation des ensembles purs, qui sont tous les ensembles obtenus à partir de l'ensemble vide à l'aide des opérations ensemblistes. ◄

▷ *A ce point de l'analyse, le principe est de fonder une théorie des ensembles sur l'axiome d'extensionnalité, la famille des axiomes de séparation, et les axiomes de la paire, de l'union et des parties — ou de faire cela pour chaque type τ et chaque choix d'une signature spécifique adaptée à τ. Développer autant de théories des ensembles qu'il existe de signatures, par exemple développer une théorie des ensembles d'entiers et une autre théorie pour les ensembles de réels, semble pénible et redondant, et il est naturel de chercher à l'éviter en choisissant une fois pour toutes un type d'objet permettant une théorie unifiée.*

Or même l'introduction d'un type général « ensemble » dont relèverait tous les ensembles ne semble pas suffisant. L'intuition immédiate et la pratique mathématique suggèrent qu'il existe de nombreux objets mathématiques qui ne sont pas des ensembles, à commencer par les nombres entiers, et on se retrouve au minimum avec deux types d'objets distincts, les ensembles et les non-ensembles. D'un point de vue technique, on se persuadera aisément a posteriori du caractère préjudiciable de cette situation qui entraîne notamment que les axiomes ne sont pas des formules mono-sortées, en particulier à cause de l'existence de puissants théorèmes de logique comme le théorème de complétude de Gödel (chapitre VII) qui s'appliquent aux signatures mono-sortées, mais pas, ou pas directement, à leurs extensions à plusieurs types d'objets.

Or il existe une solution pour sortir du dilemme entre ensembles et objets généraux et rétablir un contexte homogène où tous les objets sont de même type : se restreindre à un univers dont les objets soient des ensembles qui sont aussi ensembles d'ensembles, ensembles d'ensembles d'ensembles, et ainsi de suite. Appellant purs *de tels ensembles — dont une définition formelle reste à donner — on a alors que tout ensemble d'ensembles purs est lui-même pur et qu'inversement tout élément d'un ensemble pur est un ensemble pur. C'est ce type d'ensemble très particulier qu'on va considérer dans la suite.*

A ce point, il semble clair que les ensembles purs, s'ils existent, sont clos par rapport à toutes les opérations ensemblistes précédemment considérées: union, intersection, ensemble des parties, etc. Par contre, il n'est pas a priori évident que de tels ensembles purs existent. En fait, il en existe au moins un, à savoir l'ensemble vide $\emptyset$: c'est bien un ensemble, et, par défaut, tous ses éléments, éléments d'élements, etc. sont des ensembles, et même des ensembles purs. De proche en proche, on en déduit qu'il existe une infinité d'ensembles purs, par exemple

$$\{\{\emptyset\}\} \cup \{\mathfrak{P}(\emptyset)\}, \quad \mathfrak{P}(\mathfrak{P}(\emptyset)), \quad \mathfrak{P}(\mathfrak{P}(\mathfrak{P}(\emptyset \cup \{\emptyset, \{\emptyset\}\}))) \cup \{\emptyset\}, \tag{3.12}$$

et autres ensembles ejusdem farinae.

Si on se propose d'étudier, non pas les ensembles quelconques, mais seulement les ensembles purs, alors la question, laissée ouverte pour le moment, du choix de la signature mise en jeu dans les axiomes de séparation peut être aisément résolue. En effet, il est naturel de considérer un seul type d'objet, les ensembles purs, et comme seules opérations et relations les opérations et relations ensemblistes telles que $\in$, $\subseteq$, $\cup$, $\mathfrak{P}$, *etc.*

La dernière difficulté est que la liste précédente est mal délimitée, mais cette difficulté-là est aisée à résoudre. Toutes les relations et opérations ensemblistes considérées jusqu'à présent ont en effet la propriété de pouvoir être définies *à partir de la seule relation d'appartenance* $\in$. *Une définition formelle sera donnée au chapitre VII, mais l'idée est simple et naturelle. Par exemple, la relation d'inclusion* $\subseteq$ *est définissable à partir de l'appartenance puisque* $x \subseteq y$ *équivaut à*

$$\forall t(t \in x \Rightarrow t \in y)\,;$$

de même l'opération d'union $\bigcup$ *l'est puisque* $y = \bigcup x$ *équivaut à*

$$\forall z\,(z \in y \Leftrightarrow \exists t\,(z \in t \text{ et } t \in x)).$$

Or, il est facile de montrer que, si on adopte les axiomes de séparation relativement à une certaine signature Σ, *alors automatiquement sont valides tous les axiomes de séparation relatifs à une signature obtenue en étendant* Σ *par des opérations et relations définissables à partir de* Σ *(voir la proposition 2.5 pour un énoncé et une démonstration plus précis). Le choix dès lors est clair: pour les formules légitimes dans les axiomes de séparation, on se restreindra à l'option minimale d'une signature réduite à la seule relation* $\in$. *A l'ajout près de l'axiome de l'infini qu'on verra au chapitre II, le système obtenu est celui proposé par Zermelo en 1903.* $\triangleleft$

DÉFINITION 3.13. (formule ensembliste) On note Σ_{ens} la signature comportant un unique type d'objet $\mathbf{Ens_{pur}}$ et un unique symbole de relation binaire $\in$, et on appelle *formule ensembliste* toute formule du premier ordre en la signature Σ_{ens}.

Autrement dit, on appelle formule ensembliste toute formule obtenue en assemblant à l'aide de négations, conjonctions, disjonctions, implications et quantifications des formules de la forme $x = y$ et $x \in y$. Notons que, dès lors qu'un seul type d'objet est concerné, il n'est plus nécessaire de typer les variables, qui, par défaut, sont considérées de type $\mathbf{Ens_{pur}}$, et on peut du coup alléger les parenthèses. Ainsi, par exemple, $\forall x\,\exists y\,\forall z\,(z \in y \Leftrightarrow z \in x)$ et $z \in x$ sont des formules ensemblistes, alors que $\forall x \exists y\,(x > y)$ et $\exists x\,\forall n{:}\mathbf{Ent}\,(n \in x)$ n'en sont pas, puisqu'y interviennent respectivement la relation $>$ et le type $\mathbf{Ent}$.

DÉFINITION 3.14. (système $\mathsf{Z}_{\mathsf{fini}}$) On note $\mathsf{Z}_{\mathsf{fini}}$ (comme « Zermelo fini ») le système consistant en les axiomes d'extensionnalité, de la paire, de l'union, des parties et, pour chaque formule ensembliste F, de l'axiome de séparation en F.

Le système $\mathsf{Z}_{\mathsf{fini}}$ consiste donc en la liste (infinie) des axiomes suivants[8] :

(Ext) $\quad \forall a, b\,(\,\forall x(x \in a \Leftrightarrow x \in b) \Rightarrow a = b),$

(Paire) $\quad \forall a, b\,\exists c\,(a \in c \text{ et } b \in c),$

(Un) $\quad \forall a\,\exists b\,\forall x\,(\,\exists y(x \in y \text{ et } y \in a) \Rightarrow x \in b\,),$

(Par) $\quad \forall a\,\exists b\,\forall x\,(\,\forall y(y \in x \Rightarrow y \in a) \Rightarrow x \in b\,),$

[8] il n'y a en particulier plus lieu d'utiliser plusieurs typographies différentes pour les variables puisqu'elle réfèrent toutes ici à un même type d'objet, à savoir les ensembles purs

et, pour chaque formule ensembliste $\mathsf{F}(x, a_1, \dots, a_\mathsf{n})$ où les variables a et b n'apparaissent pas (au moins comme variables libres),

$$(\mathsf{Sep}_\mathsf{F}) \qquad \forall a, a_1, \dots, a_\mathsf{n}\, \exists b\, \forall x\, (x \in b \Leftrightarrow (x \in a \text{ et } \mathsf{F}(x, a_1, \dots, a_\mathsf{n}))).$$

▷ *Comme la signature ne contient aucun nom pour un ensemble particulier, on ne peut pas déduire des axiomes ci-dessus l'existence d'un ensemble, même vide. On peut considérer comme implicite l'axiome affirmant « il existe des ensembles », sans lequel la théorie serait vide, ou l'ajouter explicitement, par exemple sous la forme* $\exists a\,(a = a)$*. On peut encore ajouter* $\emptyset$ *comme nom dans la signature* Σ_{ens} *: s'il existe au moins un ensemble, alors, par séparation en la formule* $x \neq x$*, il existe un unique ensemble vide, sans qu'il soit besoin d'ajouter un quelconque axiome. Cette question mineure et purement formelle se trouvera de toute façon réglée au chapitre II quand l'axiome affirmant l'existence d'un ensemble infini sera ajouté.*

Ce qui est proposé à ce point, c'est d'utiliser le système $\mathsf{Z}_{\mathsf{fini}}$ *comme point de départ axiomatique pour l'étude des ensembles purs. Ce système contient un grand nombre d'axiomes d'existence, et il affirme donc l'existence d'un grand nombre d'ensembles purs — dont on rappelle qu'aucune définition précise n'a été donnée pour le moment. Il doit être clair que l'analyse précédente n'est pas terminée puisque, au départ, notre but n'était pas de tout de nous restreindre aux ensembles purs, mais de bâtir au contraire une théorie suffisamment générale pour éclairer le statut de questions comme le problème du continu, qui met en jeu des ensembles d'entiers et de réels, donc a priori pas des ensembles purs. Il n'est donc pas clair que l'étude qu'on va entamer ait une portée très vaste et qu'elle constitue autre chose qu'une première étape en direction d'une théorie plus générale restant à définir.*

En fait, un petit miracle va se produire. On va en effet montrer dans les chapitres II et III qu'il existe une telle profusion d'ensembles purs qu'il est possible de représenter à l'intérieur de ceux-ci la plupart des objets mathématiques, qu'il s'agisse d'ensembles purs ou non, ou même d'objets qui, a priori, ne sont pas des ensembles. Du coup, l'étude des ensembles purs qui, au départ, paraissait restrictive et artificielle, devient le cadre naturel de la théorie des ensembles, voire même en un sens de toutes les mathématiques, et le système $\mathsf{Z}_{\mathsf{fini}}$ *est alors un point de départ pertinent.*

Par contre, on constatera vite que ce système doit être complété de nouveaux axiomes. Typiquement, il s'agira d'étudier si le système axiomatique est suffisant pour rendre compte de l'existence et des propriétés de tous les ensembles (purs) dont l'intuition suggère l'existence, la situation rêvée étant celle d'un système suffisamment complet pour que toutes les propriétés envisageables puissent y être soient démontrées, soit réfutées. Si ce n'est pas le cas, et qu'on échoue à trancher pour une certaine propriété, il s'agira de se demander s'il existe une évidence intuitive, et des argumenst techniques, recommandant d'en faire un nouvel axiome. Ceci se produira à plusieurs reprises dans la suite, à brève échéance pour certaines questions simples qui nous mèneront successivement au système de Zermelo Z *et à celui de Zermelo–Fraenkel* ZFC*, puis, beaucoup plus tard, pour des questions bien plus sophistiquées qui nous mèneront sur la voie des développements récents de la théorie des ensembles.*

Une dernière remarque : on a écarté la définition 3.1 des ensembles qui requérait que les fonctions pré-existent aux ensembles, et préféré une approche axiomatique. Notons que, toute axiomatique qu'elle soit, l'approche développée maintenant requiert que les formules, quelles qu'elles soient, pré-existent aux ensembles. On peut signaler l'approche de Bourbaki, qui essaie de se libérer de la contrainte en construisant simultanément ensembles et formules, mais on verra dans la partie III que ce point de vue entraîne davantage de limitations que d'avantages, et on ne le suivra pas ici. ◁

Exercices

Exercice 1. (cardinal) Montrer que, pour $p \neq q$, il ne peut exister de bijection de $\{1, \dots, p\}$ sur $\{1, \dots, q\}$. En déduire que tout ensemble fini est en bijection avec un unique intervalle $\{1, \dots, p\}$.

Exercice 2. (algèbre de Boole) Montrer que, pour tout ensemble A, l'ensemble $\mathfrak{P}_f(A)$ formé des parties de A qui sont soit finies, soit co-finies (c'est-à-dire de complémentaire fini) est une algbère de Boole.

EXERCICE 3. (algèbre de Boole) Montrer que, pour tout ensemble A, $(\mathfrak{P}(A), \subseteq)$ est une algèbre de Boole complète, c'est-à-dire que toute partie (finie ou infinie) admet une borne supérieure et une borne inférieure.

EXERCICE 4. (algèbre de Boole) Soit $F : B \to B'$ un homomorphisme d'algèbres de Boole. On apelle *noyau* de F l'ensemble noté $\mathrm{Ker}(f)$ des éléments x de B vérifiant $F(x) = 0$.

(*i*) L'ensemble $\mathrm{Ker}(f)$ est-il stable par l'opération $\vee$? par $\wedge$? par complément ?

(*ii*) Montrer que, pour x, y dans B, la relation $x = y$ est équivalente à $x \wedge \overline{y} = \overline{x} \wedge y = 0$.

(*iii*) A l'aide de (*ii*), montrer que F est injectif si et seulement si $\mathrm{Ker}(f)$ est réduit à $\{0\}$.

(*iv*) En déduire une nouvelle preuve pour l'injectivité de F dans la proposition 2.13.

EXERCICE 5. (formules de de Morgan) Toute algèbre de Boole satisfait aux lois

$$\overline{a \vee b} = \overline{a} \wedge \overline{b}, \qquad \overline{a \wedge b} = \overline{a} \vee \overline{b}.$$

EXERCICE 6. (anneau de Boole) (*i*) Montrer que tout anneau de Boole est de caractéristique 2.

(*ii*) Démontrer la proposition 2.11. On pourra utiliser le cas particulier des algèbres $\mathfrak{P}(A)$ pour deviner la définition des opérations d'anneau, et utiliser les lois de de Morgan (exercice 5).

(*iii*) Redémontrer le résultat de l'exercice 4(*iii*) en passant par les anneaux de Boole.