CHAPITRE III

Le système de Zermelo–Fraenkel

RÉSUMÉ. • *On introduit les ensembles purs comme obtenus récursivement à partir de* $\emptyset$ *avec l'aide de* $\mathfrak{P}$ *et* $\bigcup$.
• *Les ordinaux finis satisfont aux axiomes de Peano, et peuvent être pris comme représentation des entiers naturels par des ensembles.*
• *On peut représenter les couples par des paires, puis les fonctions par des ensembles de couples.*
• *A partir de là, de proche en proche, on peut définir pour chaque objet mathématique* x *une copie* $\underline{x}$ *qui est un ensemble pur. Il y a donc beaucoup d'ensembles purs, ce qui légitime l'option de se restreindre à ce type d'ensembles.*
• *Pour les ensembles purs, la signature minimale réduite à* $\in$ *est raisonnable ; on part donc d'une base axiomatique centrée sur la séparation pour les formules associées à cette signature.*
• *Pour garantir l'existence de l'ordinal* ω, *on ajoute un axiome spécifique dit de l'infini, et on obtient le système de Zermelo* Z.
• *Pour établir le théorème de comparaison et garantir l'existence d'ordinaux comme* $\omega + \omega$ *ou* ω_1, *on ajoute les axiomes de remplacement qui affirment que l'image d'une correspondance fonctionnelle dont le domaine est un ensemble est aussi un ensemble.*
• *L'introduction d'ensembles par des définitions récursives ordinales est alors valide.*
• *Le système de Zermelo–Fraenkel* ZF *est obtenu à partir de* Z *en ajoutant les axiomes de remplacement, plus un axiome dit de fondation exprimant que tout ensemble est pur. C'est la base de référence pour la théorie des ensembles.*
• *Rien ne nécessite ni ne justifie de considérer que les objets représentés par des ensembles sont des ensembles.*

► On revient dans ce chapitre sur la question du choix d'une base axiomatique pour les ensembles, laissée ouverte à la fin du chapitre I. En partant des ordinaux construits au chapitre II, on montre l'existence de suffisamment d'ensembles purs pour représenter les entiers et, de là, pratiquement tous les objets mathématiques, et on légitime donc l'option de limiter la théorie des ensembles à l'étude des ensembles purs. L'examen des constructions du chapitre II amène à rafiner le système Z de Zermelo en le système ZF de Zermelo–Fraenkel, base usuellement retenue.

L'organisation du chapitre est la suivante. Dans la première partie, on donne une définition des ensembles purs, on observe que les ordinaux sont des ensembles purs, et que les ordinaux finis constituent une copie des entiers naturels dans le monde des ensembles purs. De là, on montre comment représenter par des ensembles purs la plupart des objets mathématiques.

La deuxième partie examine l'axiomatisation des ensembles purs à partir du système de Zermelo $\mathsf{Z}_{\mathsf{fini}}$. Reprenant les arguments du chapitre II, on complète le système $\mathsf{Z}_{\mathsf{fini}}$ de nouveaux axiomes, menant au système de Zermelo–Fraenkel ZF.

La troisième partie légitime les définitions récursives de suites dans le système ZF, montrant qu'aucun axiome additionnel n'est nécessaire pour justifier l'existence de telles suites.

La dernière partie fait le point sur le système ZF et la position de la théorie des ensembles dans les mathématiques, revenant notamment sur la notion d'ensemble pur et discutant quelques alternatives possibles. ◀

▷ *A la fin du chapitre I, on a laissé plusieurs questions ouvertes, à commencer par celle d'une définition précise des ensembles purs, ainsi que celle de l'existence de suffisamment d'ensembles purs pour que leur étude soit signifiante. Les constructions du chapitre II vont permettre de répondre à ces questions de façon satisfaisante, et légitimer ainsi l'approche du chapitre I.*

Les résultats de ce chapitre sont importants d'un point de vue historique et épistémologique, car ce sont eux qui expliquent le rôle spécifique joué par la théorie des ensembles dans la construction de l'édifice mathématique : dès lors que tous les objets mathématiques peuvent être représentés comme des ensembles purs, l'étude des ensembles purs recouvre celles des objets les plus variés, et l'intégralité des mathématiques peut être développée à l'intérieur du cadre d'une théorie des ensembles. Pour autant, on observera la distinction entre une représentation par des ensembles et une identification à des ensembles, et l'absence de justification à prolonger l'une en l'autre. ◁

1. Représentation par des ensembles purs

▶ On pose une définition précise des ensembles purs, et on montre que tous les objets mathématiques usuels peuvent être représentés par de tels ensembles : pour (presque) chaque objet x, il existe un ensemble pur $\underline{x}$ qui en constitue une copie, au sens où chaque propriété satisfaite par x a une contrepartie satisfaite par $\underline{x}$ et exprimée exclusivement en termes ensemblistes. On aborde successivement le cas des entiers, celui des couples, celui des fonctions, puis celui des divers objets dérivés. ◀

1.1. Ensembles purs.

▶ On définit les ensembles purs à partir de l'ensemble vide et du passage à l'ensemble des parties. On montre que les ordinaux sont purs, et que la famille des ensembles purs est close par les opérations ensemblistes usuelles. ◀

▷ *Au chapitre I, on a introduit informellement l'idée d'un type « ensemble pur » correspondant à des ensembles construits exclusivement à l'aide d'opérations ensemblistes. On a observé qu'au moins l'ensemble vide devrait être un ensemble pur, et, de là, que seraient purs tous les ensembles obtenus à partir de $\emptyset$ par des opérations ensemblistes telles que $\bigcup$ ou $\mathfrak{P}$. Ayant désormais à disposition la suite des ordinaux, on peut préciser cette approche. Au demeurant, la définition proposée utilise la notion de définition récursive d'une suite indexée par les ordinaux, qui ne sera justifiée formellement que dans la quatrième partie de ce chapitre. On se convaincra facilement qu'il n'y a pas de cercle vicieux dans la construction.* ◁

Définition 1.1. (ensemble pur, rang) On dit qu'un ensemble a est *pur* s'il appartient à au moins un des ensembles V_α récursivement définis par les clauses

$$V_{\underline{0}} = \emptyset, \quad V_{\alpha+\underline{1}} = \mathfrak{P}(V_\alpha), \quad V_\lambda = \bigcup_{\alpha<\lambda} V_\alpha \quad \text{pour } \lambda \text{ limite.} \tag{1.1}$$

Si a est un ensemble pur, le plus petit ordinal α tel que a appartienne à $V_{\alpha+\underline{1}}$ est appelé *rang* de a, et noté $\mathrm{rang}(a)$.

Par construction, tout V_λ avec λ limite est réunion des V_α pour $\alpha < \lambda$: donc le plus petit α tel qu'un ensemble appartienne à V_α est toujours un ordinal successeur, ce qui légitime la définition du rang.

Exemple 1.2. (ensembles V_α) On trouve successivement $V_{\underline{0}} = \emptyset$, puis
$V_{\underline{1}} = \mathfrak{P}(V_{\underline{0}}) = \{\emptyset\} = \{\underline{0}\}$,
$V_{\underline{2}} = \mathfrak{P}(V_{\underline{1}}) = \{\emptyset, \{\emptyset\}\} = \{\underline{0}, \underline{1}\}$,
$V_{\underline{3}} = \mathfrak{P}(V_{\underline{2}}) = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\{\emptyset\}\}\} = \{\underline{0}, \underline{1}, \underline{2}, \{\underline{1}\}\}$.
Ensuite, $V_{\underline{4}}$ a 16 éléments, et $V_{\underline{5}}$ en a 2^{16}. De proche en proche, on voit que chaque ensemble $V_{\underline{n}}$ est fini, avec un nombre d'éléments qui est une tour d'exponentielles de base 2 de hauteur n. Il en résulte que l'ensemble V_ω est infini, de même que chacun des ensembles V_α pour $\alpha \geqslant \omega$.

▷ *La définition* (1.1) *des ensembles* V_α *correspond bien au principe d'ensembles construits à partir de l'ensemble vide à l'aide d'opérations ensemblistes. Ensuite, déclarer purs tous les éléments de tels ensembles répond à l'option de définir le cadre le plus large possible : dès que* α *est un ordinal transfini, l'ensemble* V_α *est infini, et son ensemble des parties, qui est non dénombrable, contient davantage d'ensembles qu'il n'existe de phrases pour les spécifier. Autrement dit, il existe des ensembles purs qui ne sont spécifiables par aucune opération, ensembliste ou autre.* ◁

On établit maintenant qu'il existe beaucoup d'ensembles purs en montrant que les ordinaux sont purs, et que la famille des ensembles purs est close par les opérations ensemblistes de base.

Lemme 1.3. (i) *Pour tout* α, *l'ensemble* V_α *est transitif.*
(ii) *La suite des ensembles* V_α *est croissante :* $\alpha \leqslant \beta$ *entraîne* $V_\alpha \subseteq V_\beta$.

Démonstration. (i) On utilise une induction sur α. Le résultat est vrai pour $\alpha = \underline{0}$ puisque les éléments des éléments de $\emptyset$ sont éléments de $\emptyset$. Pour le cas où α est successeur, le lemme II.2.2 affirme que $\mathfrak{P}(a)$ est transitif dès que a l'est. Enfin, pour le cas où α est limite, le même lemme II.2.2 affirme que toute union d'ensembles transitifs est transitive.

(ii) On fixe α, et on montre par induction sur $\beta \geqslant \alpha$ que V_α est inclus dans V_β. Le résultat est vrai pour $\beta = \alpha$. Supposons $\beta = \gamma + \underline{1}$: par hypothèse d'induction, on a $V_\alpha \subseteq V_\gamma$, et, puisque V_γ est transitif, $V_\gamma \subseteq \mathfrak{P}(V_\gamma) = V_\beta$, d'où $V_\alpha \subseteq V_\beta$. Enfin, supposons β limite. Par définition, V_β est la réunion des V_γ pour $\gamma < \beta$, donc chacun de ces ensembles, et en particulier V_α, est inclus dans V_β. □

Proposition 1.4. (ordinaux) *Tout ordinal* α *est un ensemble pur, et on a* $\mathrm{rang}(\alpha) = \alpha$.

Démonstration. On va montrer que, pour tout α, les ordinaux appartenant à V_α sont les ordinaux strictement plus petits que α. Puisque V_α est transitif, et que $\alpha < \beta$ entraîne $\alpha \in \beta$, il suffit, pour montrer que $\beta \geqslant \alpha$ entraîne $\beta \notin V_\alpha$, de montrer $\alpha \notin V_\alpha$. Au total, il suffit donc de montrer $\alpha \subseteq V_\alpha$ et $\alpha \notin V_\alpha$ et, pour cela, on utilise une induction sur α.

Pour $\alpha = \underline{0}$, le résultat est vrai puisque $\emptyset \subseteq x$ est toujours vrai, et $x \in \emptyset$ toujours faux. Supposons $\alpha = \beta + \underline{1}$, soit $\alpha = \beta \cup \{\beta\}$. Par hypothèse d'induction, on a $\beta \subseteq V_\beta$, donc *a fortiori* $\beta \subseteq V_\alpha$, et $\{\beta\} \subseteq V_\alpha$, d'où $\alpha = \beta \cup \{\beta\} \subseteq V_\alpha$. Par contre, supposons $\alpha \in V_\alpha$, soit

$\beta \cup \{\beta\} \in \mathfrak{P}(V_\beta)$: on déduit $\{\beta\} \subseteq V_\beta$, donc $\beta \in V_\beta$, contredisant l'hypothèse d'induction. Donc on a $\alpha \notin V_\alpha$.

Supposons finalement α limite. Soit $\beta \in \alpha$. Alors on a $\beta < \alpha$, donc, par hypothèse d'induction, $\beta \in V_{\beta+1}$, d'où $\beta \in V_\alpha$ par le lemme 1.3, et donc $\alpha \subseteq V_\alpha$. Enfin, supposons $\alpha \in V_\alpha$. Puisque α est un ordinal limite, il existe β vérifiant $\beta < \alpha$ et $\alpha \in V_\beta$. Puisque V_β est transitif, ceci entraîne $\beta \in V_\beta$, contredisant l'hypothèse d'induction. On a donc $\alpha \notin V_\alpha$. □

Pour $\beta > \alpha$, l'ordinal α appartient à V_β mais pas à V_α, et on déduit donc :

COROLLAIRE 1.5. *La suite des ensembles V_α est strictement croissante: $\alpha < \beta$ entraîne $V_\alpha \subsetneq V_\beta$ (figure 1).*

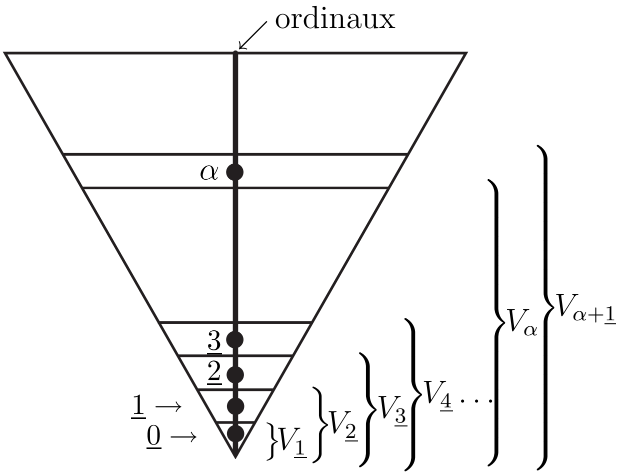

FIGURE 1. Structure en cornet de glace des V_α : les ordinaux jouent le rôle d'une colonne vertébrale et V_α correspond aux éléments situés au-dessous du niveau de l'ordinal α

PROPOSITION 1.6. (clôture) *La famille des ensembles purs est close par passage à un élément, à un sous-ensemble, à l'ensemble des parties, à l'union et l'intersection, et par formation d'ensembles finis ; de plus, pour a, a_1, ... , a_{n} purs, les relations suivantes sont vérifiées*

$$b \in a \Rightarrow \mathrm{rang}(b) < \mathrm{rang}(a), \tag{1.2}$$

$$b \subseteq a \Rightarrow \mathrm{rang}(b) \leqslant \mathrm{rang}(a), \tag{1.3}$$

$$\mathrm{rang}(\mathfrak{P}(a)) = \mathrm{rang}(a) + 1, \tag{1.4}$$

$$\mathrm{rang}(\bigcup a) \leqslant \mathrm{rang}(a) \leqslant \mathrm{rang}(\bigcup a) + \underline{1}, \tag{1.5}$$

$$\mathrm{rang}(\bigcap a) < \mathrm{rang}(a), \tag{1.6}$$

$$\mathrm{rang}(\{a_1, \dots, a_{\mathsf{n}}\}) = \sup(\mathrm{rang}(a_1), \dots, \mathrm{rang}(a_{\mathsf{n}})) + \underline{1}. \tag{1.7}$$

DÉMONSTRATION. Soit a un ensemble pur, et soit α le rang de a. Par définition, on a $a \in V_{\alpha+\underline{1}} = \mathfrak{P}(V_\alpha)$, donc $a \subseteq V_\alpha$. Donc $b \in a$ entraîne $b \in V_\alpha$: on en déduit que b est pur, et qu'on a $\mathrm{rang}(b) < \alpha$, donc $\mathrm{rang}(b) < \mathrm{rang}(a)$.

Soit maintenant b une partie de a. Comme précédemment, on a $a \subseteq V_\alpha$, d'où, par transitivité de l'inclusion, $b \subseteq V_\alpha$, et donc $b \in V_{\alpha+\underline{1}}$: ceci montre que b est pur, et, de plus, que (1.3) est vérifiée. Ensuite, puisque chaque partie de a appartient à $V_{\alpha+\underline{1}}$, l'ensemble $\mathfrak{P}(a)$ est inclus dans $V_{\alpha+\underline{1}}$, donc il est élément de $V_{\alpha+\underline{2}}$: ceci montre que $\mathfrak{P}(a)$ est pur, et donne l'inégalité $\mathrm{rang}(\mathfrak{P}(a)) \leqslant \mathrm{rang}(a) + 1$. En fait, on a l'égalité (1.4), car $a \in \mathfrak{P}(a)$ entraîne $\mathrm{rang}(a) < \mathrm{rang}(\mathfrak{P}(a))$ par (1.2).

Considérons maintenant $\bigcup a$. Soit c un élément quelconque de $\bigcup a$: il existe b vérifiant $c \in b \in a$, donc b, puis c, sont purs, et on a $\mathrm{rang}(c) < \mathrm{rang}(b) < \mathrm{rang}(a)$, donc $\mathrm{rang}(c) < \alpha$, et $c \in V_\alpha$. Ceci entraîne $\bigcup a \subseteq V_\alpha$, et donc $\bigcup a \in V_{\alpha+\underline{1}}$, montrant que $\bigcup a$ est pur, et qu'on a $\mathrm{rang}(\bigcup a) \leqslant \mathrm{rang}(a)$. D'un autre côté, soit $b \in a$: alors tous les éléments de b sont des éléments de $\bigcup a$, et donc on a $b \subseteq \bigcup a$, soit $b \in \mathfrak{P}(\bigcup a)$, ce qui montre qu'on a toujours $a \subseteq \mathfrak{P}(\bigcup a)$. Appliquant (1.4), on en déduit $\mathrm{rang}(a) \leqslant \mathrm{rang}(\mathfrak{P}(\bigcup a)) = \mathrm{rang}(\bigcup a) + \underline{1}$, d'où (1.5).

Supposons a non vide, et soit b un élément quelconque de a. Alors, par définition, $\bigcap a$ est inclus dans b : ceci entraîne que $\bigcap a$ est pur, et, par (1.3), donne (1.6).

Enfin, supposons $a_1, \ldots, a_\mathsf{n}$ pur de rangs respectifs $\alpha_1, \ldots, \alpha_\mathsf{n}$. Posons $\alpha = \sup(\alpha_1, \ldots, \alpha_\mathsf{n})$. Alors, puisque la suite des ensembles V_β est croissante vis-à-vis de l'inclusion, les ensembles a_1, $\ldots$, a_n sont tous dans $V_{\alpha+\underline{1}}$. L'ensemble $\{a_1, \ldots, a_\mathsf{n}\}$ est donc inclus dans $V_{\alpha+\underline{1}}$, et il appartient donc à $V_{\alpha+\underline{2}}$: ceci montre que $\{a_1, \ldots, a_\mathsf{n}\}$ est pur, et qu'on a $\mathrm{rang}(\{a_1, \ldots, a_\mathsf{n}\}) \leqslant \sup(\mathrm{rang}(a_1), \ldots, \mathrm{rang}(a_\mathsf{n})) + \underline{1}$. En fait, on a l'égalité (1.7), car $\{a_1, \ldots, a_\mathsf{n}\} \in V_{\gamma+1}$ entraîne $\{a_1, \ldots, a_\mathsf{n}\} \subseteq V_\gamma$, donc $a_i \in V_\gamma$ pour tout i, et, de là, $\mathrm{rang}(a_i) < \gamma$. $\square$

Corollaire 1.7. *Un ensemble est pur si et seulement si tous ses éléments sont purs.*

Démonstration. Si a est pur, tous les éléments de a sont purs d'après (1.2). Inversement, supposons que a est un ensemble dont tous les éléments sont purs. Soit X l'ensemble des rangs des éléments de a. Alors X est un ensemble d'ordinaux, donc il possède une borne supérieure α (appartenant ou non à X). Comme la suite des V_α est croissante pour l'inclusion, tous les éléments de a sont dans V_α. Donc a est inclus dans V_α, et donc est élément de $V_{\alpha+\underline{1}}$. $\square$

1.2. Représentation des entiers naturels.

► On montre qu'il existe à l'intérieur des ensembles purs une copie des entiers naturels, à savoir les ordinaux finis. ◄

▷ *Comme noté au chapitre I, l'intuition ne dit pas clairement ce que sont les entiers naturels, mais par contre suggère une description de leur comportement par les propriétés des opérations et relations spécifiques qui les mettent en jeu, addition, multiplication, ordre, etc. Faute d'une définition, on peut donc fonder l'arithmétique sur une approche axiomatique, et on sait que le système de Peano est considéré comme adéquat, c'est-à-dire comme reflétant de façon satisfaisante notre intuition. Ce système se présente comme une liste d'axiomes où apparaissent l'entier* 0*, l'opération successeur* S*, et deux opérations binaires, l'addition et la multiplication. Les deux premiers axiomes expriment que* S *est injective et que seul* 0 *n'est pas dans son image ; les suivants définissent l'addition et la multiplication récursivement ; le dernier est l'axiome d'induction.* ◁

Définition 1.8. (système de Peano) On appelle *système de Peano* PA la liste de formules suivante [1] :

(Succ$_1$) $\forall x\,(x \neq 0 \Leftrightarrow \exists y(x = S(y))$

(Succ$_2$) $\forall x, y\,(x \neq y \Rightarrow S(x) \neq S(y))$

(Add$_1$) $\forall x\,(x + 0 = x)$

(Add$_2$) $\forall x, y\,(x + S(y) = S(x + y))$

(Mult$_1$) $\forall x\,(x \cdot 0 = 0)$

(Mult$_2$) $\forall x, y\,(x \cdot S(y) = x \cdot y + x)$

(Ind) $\forall X\,((X(0)$ et $\forall x(X(x) {\Rightarrow} X(S(x)))) \Rightarrow \forall x(X(x)))$.

On a construit au chapitre II une suite ordonnée d'ensembles purs, à savoir les ordinaux, qui commence par une copie de $\mathbb{N}$, à savoir les ordinaux finis $\underline{n}$. Pour éviter toute ambiguïté avec les notations de la section 2.4, on utilise ici $\underline{\mathbb{N}}$ pour l'ensemble des ordinaux $\underline{n}$ pour n dans $\mathbb{N}$ — c'est-à-dire pour l'ensemble ω du chapitre II.

Lemme 1.9. *La structure* $(\underline{\mathbb{N}}, \underline{0}, S, +, \cdot)$ *satisfait aux axiomes de Peano.*

Démonstration. Par construction, l'application $n \mapsto \underline{n}$ établit une bijection entre $\mathbb{N}$ et $\underline{\mathbb{N}}$. De plus, on a $\underline{S(n)} = S(\underline{n})$ pour tout n, et on a observé au chapitre II que, pour tous n, p, on a $\underline{(n{+}p)} = \underline{n}{+}\underline{p}$, et $\underline{(n{\cdot}p)} = \underline{n}{\cdot}\underline{p}$. Donc l'application $n \mapsto \underline{n}$ est un isomorphisme de $(\mathbb{N}, 0, +, \cdot)$ sur $(\underline{\mathbb{N}}, \underline{0}, +, \cdot)$. L'hypothèse que $(\mathbb{N}, 0, +, \cdot)$ satisfait aux axiomes de Peano entraîne que $(\underline{\mathbb{N}}, \underline{0}, +, \cdot)$ y satisfait aussi — et, plus généralement, que les deux structures satisfont les mêmes formules (de quelque logique que ce soit). □

▷ *Ainsi, munis des opérations ordinales, les ordinaux finis se comportent exactement comme les entiers naturels : puisqu'ils satisfont les axiomes du système de Peano, ils satisfont aussi à toutes les propriétés qui se déduisent de ceux-ci, donc à toutes les propriétés usuelles des entiers naturels. L'existence de cette copie des entiers à l'intérieur du monde des ensembles purs invite évidemment à* représenter *chaque entier naturel n par l'ordinal $\underline{n}$, et l'ensemble $\mathbb{N}$ des entiers par l'ordinal ω. Faut-il en déduire que les entiers* sont *les ordinaux finis, et considérer que la formule $\underline{2} = \{\underline{0}, \underline{1}\}$, qui résulte de la construction particulière proposée au chapitre II, exprime que l'entier* 2 est *l'ensemble $\{0, 1\}$? On y reviendra dans la section 4, mais on peut déjà noter que rien dans les résultats précédents ne le justifie.* ◁

1.3. Représentation des couples.

► On construit à l'intérieur du monde des ensembles purs des contreparties pour d'autres types d'objets, ici les couples. ◄

▷ *Notre but est de définir, pour chaque objet mathématique x un ensemble pur $\underline{x}$ qui en soit une copie. Le passage aux ensembles ne pose pas de problème : si a est un ensemble, et que $\underline{x}$ a été défini pour chaque x dans a, il est naturel de poser*

$$\underline{a} := \{\underline{x}\,;\, x \in a\},$$

[1] On notera que les six premiers axiomes de PA sont exprimés par des formules du premier ordre relativement à la signature constituée de $0, S, +$, et $\cdot$; par contre, l'axiome Ind est exprimé par une formule qui est du second ordre relativement à cette signature, puisqu'y figure une variable X référant non à un entier naturel, mais à un ensemble d'entiers naturels, ou, ce qui revient au même, à une relation (unaire) sur les entiers.

et le corollaire 1.7 garantit alors que $\underline{a}$ *est un ensemble pur dès que chacun des* $\underline{x}$ *en est un, tandis que la propriété d'extensionnalité garantit que* $\underline{a}$ *détermine* a *pourvu que* $\underline{x}$ *détermine* x *pour chaque* x *dans* a*. C'est en particulier ce qu'on a fait en prenant* ω*, ensemble des ordinaux finis, comme représentant de* $\mathbb{N}$*, c'est-à-dire comme valeur de* $\underline{\mathbb{N}}$*.*

D'autres opérations de base sont plus délicates, à commencer par la formation de couples. Comme dans le cas des entiers naturels, ce qu'est un couple n'est pas clair : le seul point important est que, quels que soient x *et* y*, il existe un couple* (x, y) *et qu'à la différence d'une paire un couple détermine un premier et un second élément, autrement dit que* $(x, y) = (x', y')$ *équivaut à la conjonction de* $x = x'$ *et* $y = y'$*. Ceci revient à dire que l'opération « couple »* C *est partout définie et déterminée par l'axiome*

$$\forall x, y, x', y' (C(x, y) = C(x', y') \Leftrightarrow (x = x' \text{ et } y = y')). \tag{1.8}$$

Si on trouve une opération ensembliste C *satisfaisant* (1.8)*, alors on peut utiliser* $C(x, y)$ *comme contrepartie ensembliste pour le couple* (x, y)*, ou plutôt, supposant que* $\underline{x}$ *et* $\underline{y}$ *sont les contreparties de* x *et* y*, utiliser* $C(\underline{x}, \underline{y})$ *comme contrepartie à* (x, y)*. Or il est facile de trouver une telle opération.* ◁

Lemme 1.10. *Posons* $C(x, y) = \{\{x, y\}, \{x\}\}$*. Alors* C *satisfait à* (1.8)*.*

Démonstration. Il est clair que la condition est suffisante. Réciproquement, supposons $\{\{x, y\}, \{x\}\} = \{\{x', y'\}, \{x'\}\}$. Supposons d'abord $x = y$. Alors, $\{\{x, y\}, \{x\}\}$ est le singleton $\{\{x\}\}$. Donc $\{\{x', y'\}, \{x'\}\}$ est un singleton, ce qui entraîne $\{x', y'\} = \{x'\}$, donc $x' = y'$. Il vient alors $\{\{x\}\} = \{\{x'\}\}$, donc $\{x\} = \{x'\}$, donc $x = x'$, et, finalement, $y = x = x' = y'$.

Supposons maintenant $x \neq y$. Alors $\{\{x, y\}, \{x\}\}$ n'est pas un singleton, et, par l'argument précédent $\{\{x', y'\}, \{x'\}\}$ non plus n'est pas un singleton, et, par conséquent, on a $x' \neq y'$. Ensuite, le singleton $\{x\}$ est élément de $\{\{x, y\}, \{x\}\}$, donc de $\{\{x', y'\}, \{x'\}\}$, ce qui est dire qu'on a soit $\{x\} = \{x', y'\}$, soit $\{x\} = \{x'\}$. Le premier cas est impossible, puisque $\{x\}$ a un élément, alors que $\{x', y'\}$ en a deux. On a donc nécessairement $\{x\} = \{x'\}$, donc $x = x'$. Ensuite, par le même raisonnement, on a $\{x, y\} = \{x', y'\} = \{x, y'\}$, et donc en particulier $y \in \{x, y'\}$. Puisque y n'est pas x, la seule possibilité est $y = y'$. □

Définition 1.11. (couple, produit) Pour a, b ensembles purs, on pose

$$\underline{(}a, b\underline{)} := \{\{a, b\}, \{a\}\}, \quad \text{et} \quad a \mathbin{\underline{\times}} b := \{\underline{(}x, y\underline{)}\,;\, x \in a \text{ et } y \in b\}. \tag{1.9}$$

Etant définies en termes purement ensemblistes, les opérations précédentes ne font pas sortir du cadre des ensembles purs :

Proposition 1.12. (clôture 2) *La famille des ensembles purs est close par les opérations* $\underline{(}.,.\underline{)}$ *et* $\underline{\times}$ *; pour* a, b *purs, on a*

$$\operatorname{rang}(\underline{(}a, b\underline{)}) = \sup(\operatorname{rang}(a), \operatorname{rang}(b)) + \underline{2}, \tag{1.10}$$

$$\operatorname{rang}(a \mathbin{\underline{\times}} b) \leqslant \sup(\operatorname{rang}(a), \operatorname{rang}(b)) + \underline{2}. \tag{1.11}$$

Démonstration. Soit a et b des ensembles purs, de rangs respectifs α et β. Par la proposition 1.6, le singleton $\{a\}$ est pur de rang $\alpha + \underline{1}$, la paire $\{a, b\}$ est pure de rang $\sup(\alpha, \beta) + \underline{1}$, et l'ensemble $\underline{(}a, b\underline{)}$ est donc pur de rang $\sup(\alpha, \beta) + \underline{2}$.

Supposons maintenant $t \in a \times b$. Par définition, il existe x dans a et y dans b tels qu'on ait $t = \underline{(}x, y\underline{)}$. Alors x et y sont purs de rang au plus $\sup(\alpha, \beta)$, et $\underline{(}x, y\underline{)}$ est pur de rang au plus $\sup(\alpha, \beta) + \underline{2}$. Donc $a \mathbin{\underline{\times}} b$ est inclus dans $V_{\sup(\alpha,\beta)+\underline{2}}$, et il est donc élément de $V_{\sup(\alpha,\beta)+\underline{3}}$, ce qui montre qu'il est pur, et que son rang est au plus $\sup(\alpha, \beta) + \underline{2}$. □

▷ *Les résultats précédents montrent qu'on peut sans danger utiliser l'opération* $\underline{(.,.)}$ *pour représenter la formation de couples, c'est-à-dire, pour* a, b *ensembles purs, prendre l'ensemble pur* $\underline{(a,b)}$ *comme représentation du couple* (a, b) *dans le monde des ensembles purs, et, plus généralement, si* $\underline{x}$ *et* $\underline{y}$ *sont les représentations comme ensembles purs de deux objets* x, y*, prendre* $\underline{(\underline{x},\underline{y})}$ *comme représentation du couple* (x, y)*, soit, en d'autres termes, poser* $\underline{\underline{(x,y)}} := \underline{(\underline{x},\underline{y})}$. ◁

1.4. Représentation des fonctions.

▶ L'étape suivante est la représentation des fonctions par des ensembles purs. ◀

▷ *La solution est bien connue : on représente une fonction* f *par son graphe, c'est-à-dire par l'ensemble des couples* $(x, f(x))$*, ou plutôt l'ensemble des ensembles* $\underline{\underline{(x, f(x))}}$ *qui représentent ces derniers.*

Pour usuelle qu'elle soit, la représentation d'une fonction par son graphe peut difficilement passer pour une identification, tout au moins du point de vue de l'intuition : lorsqu'on pense à la fonction « carré » sur $\mathbb{N}$*, on pense à l'algorithme « prendre un nombre et le multiplier par lui-même » — c'est-à-dire à la recette définissant la fonction — bien plus qu'à l'ensemble infini des couples* (n, n^2)*, qui contient des quantités de couples auxquels personne n'a probablement jamais pensé explicitement, comme par exemple* $(2857, 8162449)$ *et* $(5731, 32844361)$*.*

Pour la suite, le point important est que les objets attachés à une fonction f *puissent se lire à l'aide d'opérations ensemblistes à partir de l'ensemble* $\underline{f}$*, et que les représentations ainsi définies ne fassent pas sortir du cadre des ensembles purs, ce qu'on montrera facilement. Par contre, on notera qu'il est douteux qu'on puisse en général retrouver à partir de* $\underline{f}$ *les propriétés de complexité de la fonction* f *vue comme règle de calcul, typiquement le nombre d'étapes d'un algorithme évaluant* $f(x)$ *en fonction de la taille de* x [2]*.* ◁

Notation 1.13. Le domaine d'une fonction f est noté $\mathrm{Dom}(f)$, et son image $\mathrm{Im}(f)$. On note $\mathrm{Fonc}(A, B)$ pour l'ensemble de toutes les fonctions de A dans B, et ${}^{A}B$ ou $\mathrm{Applic}(A, B)$ pour l'ensemble des applications de A dans B, c'est-à-dire l'ensemble des fonctions de A dans B partout définies sur A.

Définition 1.14. (fonction) Pour a, b ensembles purs et f fonction de a dans b, on pose

$$\underline{f} = \{\underline{(x, f(x))}\, ; \, x \in \mathrm{Dom}(f)\}, \tag{1.12}$$

$$\underline{\mathrm{Fonc}}(a, b) := \{\underline{f}\, ; \, f \in \mathrm{Fonc}(a, b)\}. \tag{1.13}$$

Lemme 1.15. *Soient* a, b, c *des ensembles purs. Alors la propriété pour* c *de représenter une fonction de* a *dans* b*, puis, si c'est le cas, la valeur de cette fonction en un point, ainsi que son domaine et son image s'expriment à partir de* c *à l'aide d'opérations ensemblistes.*

Démonstration. D'abord, c représente une fonction de a dans b si et seulement si on a

$$c \subseteq a \underline{\times} b \text{ et } \forall x{\in}a\, \forall y, y'{\in}b\, ((\underline{(x,y)} \in c \text{ et } \underline{(x,y')} \in c) \Rightarrow y = y').$$

Ensuite supposons $c = \underline{f}$. Alors $f(x) = y$ équivaut à $\underline{(x,y)} \in c$. D'autre part, on a $\mathrm{Dom}(f) = \{x\, ; \, \exists y(\underline{(x,y)} \in c)\}$ et $\mathrm{Im}(f) = \{y\, ; \, \exists x(\underline{(x,y)} \in c)\}$. □

[2]La question est elle-même floue, car, dès que le domaine de f est infini, on ne saurait spécifier l'ensemble des couples $(x, f(x))$ autrement que par une définition synthétique, typiquement une recette d'évaluation.

Proposition 1.16. (clôture 3) *La famille des ensembles purs est close par formation de fonction et opération* $\underline{\mathrm{Fonc}}$ *; pour* a, b *purs et* f *fonction de* a *dans* b*, on a*

$$\mathrm{rang}(\underline{f}) \leqslant \sup(\mathrm{rang}(a), \mathrm{rang}(b)) + \underline{3}, \tag{1.14}$$

$$\mathrm{rang}(\underline{\mathrm{Fonc}}(a,b)) \leqslant \sup(\mathrm{rang}(a), \mathrm{rang}(b)) + \underline{4}. \tag{1.15}$$

Démonstration. Soient a et b des ensembles purs, de rang respectif α et β. Soit f une fonction de a dans b. Par la proposition 1.12 (clôture 2), tous les ensembles $\underline{(x,y)}$ avec $x \in a$ et $y \in b$ sont purs, donc tous les éléments de $\underline{f}$ sont des ensembles purs de rang au plus $\sup(\alpha,\beta)+\underline{2}$, et, par le corollaire 1.7, $\underline{f}$ est lui-même un ensemble pur, de rang au plus $\sup(\alpha,\beta) + \underline{3}$.

Ensuite, tout ensemble $\underline{f}$ avec f fonction de a dans b étant pur de rang au plus $\sup(\alpha,\beta)+\underline{3}$, l'ensemble $\underline{\mathrm{Fonc}}(a,b)$ de tous ces ensembles $\underline{f}$ est inclus dans $V_{\sup(\alpha,\beta)+\underline{3}}$, donc est élément de $V_{\sup(\alpha,\beta)+\underline{4}}$. □

1.5. Représentation de tous les objets mathématiques.

▶ Dès lors qu'on a une représentation par des ensembles purs pour les entiers naturels, les fonctions, et les opérations qui en dérivent, on représente de proche en proche la plupart des objets mathématiques. ◀

« Proposition » 1.17. (représentation) *Tous les objets mathématiques usuels peuvent être représentés par des ensembles purs.*

Démonstration. Il s'agit d'associer à chaque objet x un ensemble pur $\underline{x}$ qui en soit une copie suffisamment conforme. Pour les entiers naturels, on a déjà justifié le choix $\underline{\mathbb{N}} = \omega$.

On passe aux entiers relatifs. On sait que ceux-ci peuvent être représentés comme classes d'équivalence de couples d'entiers naturels pour la relation $\equiv$ définie par $(x,y) \equiv (x',y')$ pour $x + y' = x' + y$. Par le lemme 1.15, on peut utiliser l'ensemble pur $\underline{\omega \times \omega}$ comme représentation de $\mathbb{N} \times \mathbb{N}$. Les classes d'équivalence pour la relation $\underline{\equiv}$ sur $\omega \underline{\times} \omega$ analogue à $\equiv$ sont elles-mêmes des ensembles purs, de même que l'ensemble $\underline{\mathbb{Z}}$ de ces classes, qu'on peut prendre comme contrepartie de $\mathbb{Z}$.

Le passage des entiers relatifs aux rationnels est analogue : on peut construire un rationnel comme classe d'équivalence de couples d'entiers relatifs, et on obtient une représentation $\underline{\mathbb{Q}}$ de $\mathbb{Q}$ en considérant des classes d'équivalence sur $\underline{\mathbb{Z}} \underline{\times} \underline{\mathbb{Z}}$.

Ensuite les réels peuvent se construire à partir des rationnels soit en utilisant les coupures de Dedekind (complétion de l'ordre), soit en utilisant les suites de Cauchy (complétion topologique). Dans la première approche, un réel est un couple formé de deux sous-ensembles adjacents de $\mathbb{Q}$, donc un sous-ensemble de $\mathbb{Q} \times \mathbb{Q}$. En remplaçant ce dernier ensemble par sa copie $\underline{\mathbb{Q}} \underline{\times} \underline{\mathbb{Q}}$, on obtient pour chaque nombre réel a une contrepartie $\underline{a}$ qui est un ensemble pur, et de là un ensemble pur $\underline{\mathbb{R}}$ formé par tous les $\underline{a}$ avec a dans $\mathbb{R}$.

Les nombres complexes pouvant être construits comme couples de réels, on peut poser $\underline{\mathbb{C}} = \underline{\mathbb{R}} \underline{\times} \underline{\mathbb{R}}$.

Ces structures de base construites, les autres objets s'en déduisent à leur tour. Par exemple, le lemme 1.15 permet de représenter toutes les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ par des ensembles purs, et d'obtenir un ensemble pur $\mathrm{Fonc}(\mathbb{R},\mathbb{R})\_$ de toutes ces copies de fonctions. Et ainsi de suite... □

▷ *Les résultats précédents légitiment l'approche proposée au chapitre I. Dès lors que pratiquement tous les objets mathématiques, ensembles ou non, admettent des copies bien définies à l'intérieur du monde des ensembles purs, restreindre l'étude des ensembles aux ensembles purs* *devient une option raisonnable : ce qui, au départ, apparaissait comme une restriction artificielle et de nature à altérer la portée de l'étude, définit en fait un champ très vaste puisqu'il inclut une réplique de tous les objets usuels.*

Comme on l'a souligné pour les entiers naturels, il n'y a aucune raison de considérer que la possibilité de représenter les objets mathématiques par des ensembles (purs) indique que ces objets soient eux-mêmes des ensembles. Par contre, mais uniquement à des fins de simplification des notations [3]*, il sera commode, dans ce texte principalement consacré aux ensembles, d'adopter pour la suite les conventions suivantes.* ◁

Convention 1.18. Par défaut, toutes les variables a, b,... x, y *etc.* réfèrent à des ensembles purs, les objets qui ne sont pas des ensembles purs étant en principe désignés par des caractères typographiques spéciaux tels que $\mathbb{N}$, $\mathbb{R}$,... $\mathbf{A}$, $\mathbf{B}$,... n, p,... F, G, *etc.*

On utilise (x, y) pour $\underline{(x, y)}$, et $x \times y$ pour $x \underline{\times} y$, c'est-à-dire qu'on identifie les couples avec leur représentation ensembliste — ou, plus exactement, dans ce texte traitant principalement d'ensembles, on décide d'écrire « couple » comme un raccourci pour « ensemble représentant le couple ». Au même sens, on identifie les fonctions et leurs représentations par des ensembles de couples, ainsi que les entiers et leurs représentations par des ordinaux finis. Par conséquent, on notera simplement 0 pour $\underline{0}$, 1 pour $\underline{1}$, *etc.* , et on appelle *entiers* les éléments de l'ordinal ω. Néanmoins, on conservera la notation $\mathbb{N}$ pour référer aux « vrais » entiers du discours, appelés ici *intuitifs* et notés typiquement i, n, p,... par opposition à i, n, p,... pour des entiers, ordinaux finis. Cette distinction, qu'on peut oublier pour le moment, deviendra importante lorsqu'on abordera la théorie axiomatique des ensembles à partir du chapitre IX.

2. Axiomatisation des ensembles purs

► On revient pour la préciser sur l'axiomatisation des ensembles purs introduite au chapitre I. L'examen des démonstrations du chapitre II et de la première partie du présent chapitre amène à compléter le système $\mathsf{Z}_{\mathsf{fini}}$ en le système ZF de Zermelo–Fraenkel. ◄

▷ *Le programme envisagé à la fin du chapitre I est maintenant clarifié. On vient de constater qu'il existe assez d'ensembles purs pour que l'étude de ces derniers soit un but légitime et une approximation suffisante de la théorie des ensembles. Dès lors, la première tâche est de reprendre le système axiomatique esquissé au chapitre I et d'examiner si les axiomes de ce système suffisent à justifier l'existence et les propriétés de tous les ensembles précédemment introduits, et plus spécifiquement celles de tous les ensembles purs intervenant dans la représentation des objets mathématiques. Il s'agit donc de vérifier si, sur la base des axiomes déjà formulés, on peut justifier l'existence et les propriétés des ordinaux, en particulier le fait que les ordinaux finis forment une copie des entiers naturels, ainsi que celles des couples et des fonctions. A chaque fois que les axiomes initiaux apparaîtront insuffisants — et cela arrivera — il conviendra de se demander si la propriété qu'on échoue à démontrer est cruciale et, dans l'affirmative, d'envisager d'amender le système en ajoutant un ou plusieurs nouveaux axiomes. C'est ainsi qu'on parviendra au système de Zermelo–Fraenkel, point de départ usuel pour une théorie générale des ensembles.* ◁

[3] et parce que cela ne créera pas d'ambiguïté, *cf.* discussion de la section 4.4

2.1. Extensions par définition.

▶ On montre qu'il est légitime d'utiliser dans les constructions d'ensembles par séparation des relations et des opérations définies à partir de l'appartenance. ◀

On rappelle d'abord la définition du système $\mathsf{Z}_{\mathsf{fini}}$ introduit au chapitre I. Afin d'alléger les notations, on adopte dans toute la suite la convention suivante :

Convention 2.1. (suite) La notation $\vec{a}$ désigne une suite finie $(a_1, ..., a_{\mathsf{n}})$ de longueur non spécifiée ; dans ce contexte, $\vec{a} \in A$ signifie $a_1 \in A$ et ... et $a_{\mathsf{n}} \in A$.

Définition 2.2. (système $\mathsf{Z}_{\mathsf{fini}}$) On note $\mathsf{Z}_{\mathsf{fini}}$ la liste de formules suivante :

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

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

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

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

et, pour chaque formule ensembliste $\mathsf{F}(x, \vec{c})$ où a et b n'apparaissent pas comme variables libres,

($\mathsf{Sep}_{\mathsf{F}}$) $$\forall a\,\forall \vec{c}\,\exists b\,\forall x\ \ (x \in b \Leftrightarrow (x \in a \text{ et } \mathsf{F}(x, \vec{c}))).$$

▷ *Noter que toutes les formules du système $\mathsf{Z}_{\mathsf{fini}}$ sont elles-mêmes des formules ensemblistes. Noter aussi que le système $\mathsf{Z}_{\mathsf{fini}}$ contient tous les axiomes de séparation, indépendamment de toute signification des formules: on peut énumérer le système en énumérant systématiquement toutes les formules ensemblistes ayant une variable libre t (rangées par exemple par longueur croissante) et en écrivant mécaniquement les axiomes de séparation associés. Si, par exemple, les trois premières formules de l'énumération sont*

$\mathsf{F}_1 = x \in x$, $\mathsf{F}_2 = x \notin c$, $\mathsf{F}_3 = \exists y(y \in x \Rightarrow x \in y)$,

alors les trois premiers axiomes de séparation sont

$\mathsf{Sep}_{\mathsf{F}_1}$: $\forall a\,\exists b\,\forall x\,(x \in b \Leftrightarrow (x \in a \text{ et } x \in x))$,

$\mathsf{Sep}_{\mathsf{F}_2}$: $\forall a, c\,\exists b\,\forall x\,(x \in b \Leftrightarrow (x \in a \text{ et } x \notin c))$,

$\mathsf{Sep}_{\mathsf{F}_3}$: $\forall a\,\exists b\,\forall x\,(x \in b \Leftrightarrow (x \in a \text{ et } \exists y(y \in x \Rightarrow x \in y)))$,

qui font donc partie du système $\mathsf{Z}_{\mathsf{fini}}$ même si le sens des formules, en particulier F_3, n'est pas clair.

On rappelle que les formules ensemblistes mises en jeu dans les axiomes de séparation sont les formules du premier ordre relativement à la signature comportant l'unique symbole de relation binaire $\in$ — dont la définition formelle (une induction sans surprise) sera donnée au chapitre VII. Pour le moment, il suffit de bien noter que les formules sont des mots formés sur un certain alphabet contenant des variables et des symboles divers, donc des objets purement syntaxiques, c'est-à-dire sans aucune notion de sens ou de valeur de vérité incluse.

Plusieurs fois dans ce qui précède, et ainsi qu'il est d'usage dans n'importe quelle partie des mathématiques, on a introduit des notations spécifiques pour des objets, opérations ou relations dérivées de notions plus primitives. Par exemple, l'axiome de l'union affirmant l'existence, pour chaque ensemble x, d'un nouvel ensemble z dont les éléments sont les éléments des éléments de x, et l'axiome d'extensionnalité garantissant l'unicité de z, on a introduit la notation $\bigcup x$ pour z.

Comme toujours, il est loisible d'introduire toutes les notations qu'on souhaite. Le point spécifique est qu'ici le rôle des formules est plus que sténographique, dans la mesure où elles apparaissent explicitement dans les axiomes retenus. Or, introduire une nouvelle opération, par exemple ici l'opération unaire $\bigcup$, revient à étendre la signature. La question est alors de savoir si le principe de séparation, initialement posé pour les formules en $\in$ seule, reste valable lorsque

$\bigcup$ *est adjoint, c'est-à-dire s'il est légitime d'utiliser* $\bigcup$ *dans des définitions par séparation et de même pour les diverses relations et opérations qui vont s'introduire dans la suite.*

Comme on l'a annoncé au chapitre I, c'est effectivement le cas, pour autant que les notions nouvelles soient définissables *à partir de la relation* $\in$. *De façon précise, introduire un objet nouveau accompagné d'une notation revient à poser un axiome les définissant. Par exemple, introduire l'opération* $\bigcup$ *revient à poser un axiome*

$$(\mathsf{Intro}_{\bigcup}) \qquad \forall x, y\,(y = \bigcup x \Leftrightarrow \mathsf{Def}_{\bigcup}(x, y)),$$

où $\mathsf{Def}_{\bigcup}(x, y)$ *est une définition de l'union, par exemple la formule*

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

De même pour l'introduction de relations ou de noms. Par exemple, on introduit la relation d'inclusion $\subseteq$ *à l'aide de l'axiome*

$$(\mathsf{Intro}_{\subseteq}) \qquad \forall x, y\,(x \subseteq y \Leftrightarrow \mathsf{Def}_{\subseteq}(x, y)),$$

où $\mathsf{Def}_{\subseteq}(x, y)$ *est une définition de l'inclusion, par exemple* $\forall t\,(t \in x \Rightarrow t \in y)$. *On introduit de même le nom* $\emptyset$ *à l'aide de l'axiome*

$$(\mathsf{Intro}_{\emptyset}) \qquad \forall x\,(x = \emptyset \Leftrightarrow \mathsf{Def}_{\emptyset}(x)),$$

où $\mathsf{Def}_{\emptyset}(x)$ *est une définition de l'ensemble vide, par exemple la formule* $\forall y\,(y \notin x)$. ◁

Définition 2.3. (formule étendue) On appelle formule ensembliste *étendue* toute formule du premier ordre en la signature comprenant, outre la relation binaire $\in$, la relation binaire $\subseteq$, l'opération binaire $\{\cdot, \cdot\}$, les opérations unaires $\bigcup$ et $\mathfrak{P}$, et le nom (opération 0-aire) $\emptyset$.

Lemme 2.4. *Il existe un algorithme associant à toute formule ensembliste étendue* F *une formule ensembliste* F' *telle que, en présence des axiomes de* $\mathsf{Z}_{\mathsf{fini}}$ *et de* $\mathsf{Intro}_{\bigcup}, \dots, \mathsf{Intro}_{\emptyset}$, *la formule* F *est équivalente à* F'.

Démonstration. La formule F' s'obtient à partir de F en remplaçant chacun des symboles autres que $\in$ par leur définition: pour $\subseteq$, on remplace chaque sous-formule du type $x \subseteq y$ par $\mathsf{Def}_{\subseteq}(x, y)$; pour $\emptyset$, on remplace chaque occurrence du symbole $\emptyset$ dans une sous-formule atomique $\mathsf{F}(\dots\emptyset\dots)$ de F par $\exists x(\mathsf{Def}_{\emptyset}(x)$ et $\mathsf{F}(\dots x \dots))$, x étant une variable sans occurrence dans F. Alors, $\mathsf{F}(\dots\emptyset\dots)$ étant équivalente à $\exists x(x = \emptyset$ et $\mathsf{F}(\dots x \dots))$, la formule obtenue est équivalente à F en présence des autres axiomes, qui garantissent l'existence et l'unicité de $\emptyset$. La procédure est la même pour les autres notions: on remplace par exemple $\mathsf{F}(\dots \bigcup a \dots)$ par $\exists x(\mathsf{Def}_{\bigcup}(a, x)$ et $\mathsf{F}(\dots x \dots))$, où x est une variable sans occurrence dans F. □

On en déduit :

Proposition 2.5. (extension de signature) *Tout axiome de séparation relatif à une formule ensembliste étendue est conséquence des axiomes de* $\mathsf{Z}_{\mathsf{fini}}$ *augmenté des axiomes d'introduction* $\mathsf{Intro}_{\bigcup}, \dots, \mathsf{Intro}_{\emptyset}$— *, et donc peut être utilisé légitimement.*

Démonstration. Soit F une formule ensembliste étendue. Par le lemme 2.4, il existe une formule ensembliste F' telle que l'équivalence $\mathsf{Sep}_{\mathsf{F}} \Leftrightarrow \mathsf{Sep}_{\mathsf{F}'}$ soit conséquence de $\mathsf{Z}_{\mathsf{fini}}$ et des axiomes d'introduction concernés. Comme $\mathsf{Sep}_{\mathsf{F}'}$ est dans $\mathsf{Z}_{\mathsf{fini}}$, il en résulte que $\mathsf{Sep}_{\mathsf{F}}$ est conséquence de $\mathsf{Z}_{\mathsf{fini}}$ et des axiomes d'introduction. □

▷ *Le procédé peut être répété pour d'autres définitions, et, par conséquent, on pourra introduire autant de nouvelles opérations et relations qu'on le souhaite : dès lors que celles-ci sont définies par des formules ensemblistes, on peut en faire un libre usage dans les définitions par séparation. Dans toute la suite, la mention aux axiomes d'introduction sera implicite, et, lorsqu'on parlera du système $\mathsf{Z}_{\mathsf{fini}}$ — ou d'un autre système analogue — on supposera toujours qu'on y a inclus les axiomes d'introduction requis.*

Le lemme 2.4 montre également qu'on peut empiler les définitions successives : ainsi, à partir du moment où l'inclusion $\subseteq$ a été définie à partir de l'appartenance, on peut l'utiliser à son tour dans des définitions, par exemple en définissant $y = \mathfrak{P}(x)$ par $\forall t(t \in y \Leftrightarrow t \subseteq x)$. Il n'est pas nécessaire de redescendre à chaque fois à la signature de base constituée de l'unique relation d'appartenance. ◁

2.2. Construction d'ensembles.

▶ On montre que les constructions mettant en jeu les couples et les fonctions ne nécessitent pas d'autre axiome que ceux de $\mathsf{Z}_{\mathsf{fini}}$. ◀

▷ *Dans la section 1, on a proposé de représenter le couple (a, b) par l'ensemble $\{\{a, b\}, \{a\}\}$. Le problème typique est donc le suivant: les axiomes de $\mathsf{Z}_{\mathsf{fini}}$ garantissent-ils l'existence, pour chaque a, b, du couple (a, b), c'est-à-dire en l'occurrence de l'ensemble $\{\{a, b\}, \{a\}\}$ utilisé comme représentation de ce couple ? La réponse est claire :* ◁

Lemme 2.6. ($\mathsf{Z}_{\mathsf{fini}}$) *Quels que soient a, b, le couple (a, b) et le produit cartésien $a \times b$ existent et ont les propriétés décrites dans le lemme 1.10*[4].

Démonstration. Puisque, par convention, (a, b) désigne l'ensemble $\{\{a, b\}, \{a\}\}$, la question est de savoir si les axiomes de $\mathsf{Z}_{\mathsf{fini}}$ garantissent l'existence, pour chaque a, b, de l'ensemble $\{\{a, b\}, \{a\}\}$, c'est-à-dire celle d'un ensemble dont les éléments soient exactement la paire $\{a, b\}$ et le singleton $\{a\}$. La réponse est immédiatement positive : les axiomes d'extensionnalité, de la paire et de séparation garantissent l'existence de $\{a, b\}$ et de $\{a\}$, puis celle de $\{\{a, b\}, \{a\}\}$. Ensuite, le fait que $\{\{a, b\}, \{a\}\}$ vérifie la propriété du lemme 1.10 résulte du seul axiome d'extensionnalité.

De la même façon, la question est de savoir si les axiomes de $\mathsf{Z}_{\mathsf{fini}}$ garantissent l'existence d'un ensemble dont les éléments soient exactement tous les couples (c, d) avec c dans a et d dans b. L'argument est un peu plus délicat. D'abord, l'ensemble $a \cup b$ existe comme union de la paire $\{a, b\}$. Supposons $c \in a$ et $d \in b$. Comme c et d appartiennent à $a \cup b$, les paires $\{c, d\}$ et $\{c\}$ appartiennent à $\mathfrak{P}(a \cup b)$, et donc le couple (c, d), c'est-à-dire la paire $\{\{c, d\}, \{c\}\}$, appartient à $\mathfrak{P}(\mathfrak{P}(a \cup b))$. On peut donc former l'ensemble $a \times b$ à l'aide de la définition:

$$a \times b = \{z \in \mathfrak{P}(\mathfrak{P}(a \cup b))\,;\, \exists x{\in}a\, \exists y{\in}b\, (z = (x, y))\},$$

donc en utilisant dans l'ensemble $\mathfrak{P}(\mathfrak{P}(a \cup b))$ l'axiome de séparation $\mathsf{Sep}_{\mathsf{F}}$, où $\mathsf{F}(z, u, v)$ est la formule ensembliste étendue $\exists x{\in}u\, \exists y{\in}v\, (z = (x, y))$ prise ici avec les paramètres $u = a$, $v = b$. En vertu de la proposition 2.5, l'utilisation d'une telle formule étendue dans une définition par séparation est légitime dans le système $\mathsf{Z}_{\mathsf{fini}}$. Si on tient à alléguer une formule ensembliste stricte, on peut remplacer l'opération couple par une de ses définitions, passant d'abord par

[4] Dans toute la suite, la notation entre parenthèses d'un système axiomatique au début d'un énoncé indique que l'énoncé en question peut être démontré à partir des axiomes du système — et par une démonstration elle-même formalisable suivant les règles de déduction usuelles telles que décrites au chapitre VII ; ici, la proposition affirme donc qu'on peut montrer l'existence de (a, b) et $a \times b$ à partir des seuls axiomes de $\mathsf{Z}_{\mathsf{fini}}$. Par ailleurs, on rappelle que, par défaut, toutes les variables réfèrent désormais à des ensembles purs ; ici, l'hypothèse est donc que a et b sont des ensembles purs

$\exists x{\in}u\, \exists y{\in}v\, (t = \{\{x,y\},\{x\}\})$,
puis $\exists x{\in}u\, \exists y{\in}v\, \exists p,q\, (p = \{x,y\}$ et $q = \{x,x\}$ et $t = \{p,q\})$,
et enfin quelque chose comme

$$\exists x{\in}u\, \exists y{\in}v\, \exists p,q\, \forall z\, ((z \in p \Leftrightarrow (z = x \text{ ou } z = y)) \text{ et } (z \in q \Leftrightarrow z = x) \text{ et } (z \in t \Leftrightarrow (z = p \text{ ou } z = q))).$$ □

Lemme 2.7. ($\mathsf{Z}_{\mathsf{fini}}$) *(i) Pour toute fonction f, le domaine* $\mathrm{Dom}(f)$ *et l'image* $\mathrm{Im}(f)$ *de f sont des ensembles, définissables à partir de f.*

(ii) Quels que soient a, b, l'ensemble des fonctions $\mathrm{Fonc}(a,b)$ *et l'ensemble des applications* $\mathrm{Applic}(a,b)$ *existent.*

Démonstration. (i) Supposons que f est une fonction, donc un ensemble de couples. Pour (x,y), c'est-à-dire $\{\{x,y\},\{x\}\}$, dans f, l'ensemble $\{x\}$ est un élément d'un élément de f, donc est élément de $\bigcup f$, et x est élément de $\{x\}$, donc est élément de $\bigcup\bigcup f$. On a alors

$$\mathrm{Dom}(f) = \{x \in \bigcup\bigcup f\, ;\, \exists y\, ((x,y) \in f)\}\ :$$

l'axiome de l'union garantit l'existence de $\bigcup\bigcup f$, et un axiome de séparation garantit donc l'existence de $\mathrm{Dom}(f)$. De même, on a

$$\mathrm{Im}(f) = \{y \in \bigcup\bigcup f\, ;\, \exists x\, ((x,y) \in f)\}.$$

(ii) Toute fonction de a dans b est un sous-ensemble de $a \times b$, donc on peut définir $\mathrm{Fonc}(a,b)$ par séparation par

$$\mathrm{Fonc}(a,b) = \{f \in \mathfrak{P}(a \times b)\, ;\, f\colon a \to b\},$$

où $f\colon a \to b$ est une formule ensembliste exprimant que f est une fonction de a dans b, laquelle existe par le lemme 1.15. L'ensemble $\mathrm{Applic}(a,b)$ des applications de a dans b est alors défini par séparation dans $\mathrm{Fonc}(a,b)$ par

$$\mathrm{Applic}(a,b) = \{f \in \mathrm{Fonc}(a,b)\, ;\, \mathrm{Dom}(f) = a\}.$$

Le lemme 1.15 à nouveau garantit que l'opération Dom est exprimable par une formule ensembliste, donc la définition est légitime dans $\mathsf{Z}_{\mathsf{fini}}$. □

2.3. Construction des ordinaux.

► On vérifie que le système $\mathsf{Z}_{\mathsf{fini}}$ est suffisant pour introduire les ordinaux et montrer leurs propriétés essentielles, en particulier le principe d'induction. ◄

Lemme 2.8. *La propriété « α est un ordinal » est exprimable par une formule ensembliste.*

Démonstration. Les quatre conditions de la définition II.2.3 se traduisent en la formule suivante, dont α est la seule variable libre :

$$\forall x{\in}\alpha\, (x \subseteq \alpha) \text{ et } \forall x{\in}\alpha\, (x \notin x) \text{ et } \forall x,y,z{\in}\alpha\, ((x \in y \text{ et } y \in z) \Rightarrow x \in z) \text{ et } \forall t{\subseteq}\, \alpha\, (t \neq \emptyset \Rightarrow \exists x{\in}t\, \forall y{\in}x\ (y = x \text{ ou } x \in y)). \tag{2.1}$$

Les symboles non logiques $\emptyset$ et $\subseteq$ figurent dans (2.1) qui n'est donc pas une formule ensembliste au sens strict. En remplaçant $x \subseteq \alpha$ par $\forall z{\in}x(z \in \alpha)$, et $t \neq \emptyset$ par $\exists z(z \in t)$, on obtient une formule équivalente qui, cette fois, est une formule ensembliste, notée $\mathbf{Ord}(\alpha)$ dans la suite. □

Le fait que la formule $\mathbf{Ord}(\alpha)$ introduite ci-dessus soit une formule ensembliste implique qu'on peut à son tour l'utiliser dans des définitions par séparation. Par exemple, pour tout ensemble a, le système $\mathsf{Z}_{\mathsf{fini}}$ garantit l'existence de l'ensemble des ordinaux éléments de a, c'est-à-dire de $\{\alpha \in a\,;\, \mathbf{Ord}(\alpha)\}$.

LEMME 2.9. ($\mathsf{Z}_{\mathsf{fini}}$) *L'ordinal* 0 *(c'est-à-dire $\emptyset$) existe, et, pour chaque ordinal α, l'ordinal $S(\alpha)$ existe.*

DÉMONSTRATION. On a déjà vu que l'existence de l'ensemble vide est garantie [5] en séparant dans un ensemble quelconque les éléments x vérifiant $x \neq x$ ou toute autre formule non satisfaisable.

Ensuite on a posé $S(\alpha) := \alpha \cup \{\alpha\}$. L'axiome de la paire et celui de l'union garantissent que l'opération $\cup$ est partout définie. Par composition, il en est de même de l'opération S. ☐

De proche en proche, on déduit l'existence de chacun des ordinaux $S^{\mathsf{n}}(\alpha)$ pour n entier intuitif, c'est-à-dire celle des ordinaux notés ici 1, 2, *etc.* Ensuite, pour les résultats des sections II.2.2 et II.2.3, on n'utilise aucune propriété spécifique sinon celles qui figurent dans la définition des ordinaux.

▷ *Le seul point à mentionner est le lemme II.2.8 et la proposition II.2.13 (tout ensemble d'ordinaux possède un plus petit élément), où on introduit une intersection $\bigcap$. Or, si a est un ensemble non vide, l'existence de $\bigcap a$ est garantie par séparation, puisqu'on a par exemple $\bigcap a = \{t \in \bigcup a\,;\, \forall x{\in}a(t \in x)\}$, définition qui a l'avantage de faire sens même si a est vide.* ◁

En particulier, on obtient la formulation précise suivante pour le principe d'induction ordinale :

PROPOSITION 2.10. (induction ordinale) ($\mathsf{Z}_{\mathsf{fini}}$) *Supposons que F est une formule ensembliste telle que, quel que soit α, si $\mathsf{F}(\beta, a_1, \dots, a_{\mathsf{n}})$ est vraie pour tout $\beta < \alpha$, alors $\mathsf{F}(\alpha, a_1, \dots, a_{\mathsf{n}})$ est vraie. Alors $\mathsf{F}(\alpha, a_1, \dots, a_{\mathsf{n}})$ est vraie pour tout ordinal α.*

DÉMONSTRATION. La démonstration est celle du chapitre II. Si $\mathsf{F}(\alpha, a_1, \dots, a_{\mathsf{n}})$ est en défaut, alors l'existence de l'ensemble $\{\beta \in S(\alpha)\,;\, \mathrm{non}\mathsf{F}(\beta, a_1, \dots, a_{\mathsf{n}})\}$ est garantie par séparation. ☐

2.4. L'axiome de l'infini.

▶ Faute de réussir à légitimer l'existence de l'ordinal ω, on ajoute un nouvel axiome, dit *de l'infini*, en postulant explicitement l'existence. On parvient ainsi au système Z de Zermelo. ◀

▷ *On passe à la section II.2.4. Le premier résultat est la proposition II.2.16 (tout ensemble d'ordinaux possède une borne supérieure). Il ne pose pas de problème, puisque l'axiome de l'union garantit l'existence de $\bigcup a$ pour tout ensemble a. L'argument de Burali–Forti montre alors qu'il n'existe pas d'ensemble contenant tous les ordinaux, donc, par séparation, qu'il n'existe pas d'ensemble de tous les ordinaux.*

Le point suivant est l'introduction de l'ordinal ω comme borne supérieure des ordinaux $\underline{\mathsf{n}}$ pour n entier naturel. Il y a ici un problème, car, pour pouvoir appliquer la proposition II.2.16, il faut savoir qu'il existe un ensemble A dont les éléments sont les ordinaux $\underline{\mathsf{n}}$. Le problème est double : rien a priori ne garantit l'existence d'un ensemble contenant tous les ordinaux $\underline{\mathsf{n}}$; ensuite, à supposer que cette existence soit garantie, rien non plus ne garantit l'existence d'un ensemble

[5] pour autant qu'il existe au moins un ensemble

contenant exactement les ordinaux $\underline{\mathsf{n}}$*, le problème étant que, pour arguer d'une séparation, il faudrait être assuré que les ordinaux* $\underline{\mathsf{n}}$ *peuvent être caractérisés par une formule ensembliste.*

On verra dans la suite qu'aucun de ces deux problèmes ne peut être contourné dans le cadre de $\mathsf{Z}_{\mathsf{fini}}$*. La situation est alors typique : l'introduction de l'ordinal* ω *semble techniquement utile, et, d'autre part, l'intuition la recommande, en tout cas, recommande l'introduction d'ensembles infinis*[6]*, alors qu'on échoue à la justifier à l'aide des axiomes déjà introduits. La solution naturelle est de considérer que la liste des axiomes est incomplète et d'ajouter un nouvel axiome.*

On verra plus loin qu'il s'agit ni plus, ni moins de poser l'existence d'ensembles infinis. Pour éviter tout appel à des objets externes comme les entiers naturels mentionnés dans la définition de la finitude, on adopte une formulation directement adaptée au contexte des ordinaux. ◁

Définition 2.11. (axiome de l'infini) La formule

$$(\mathsf{Inf}) \qquad \exists a(\emptyset \in a \text{ et } \forall x{\in}a(S(x) \in a)).$$

est appellée *axiome de l'infini.* On définit le *système de Zermelo* Z comme le système obtenu en ajoutant à $\mathsf{Z}_{\mathsf{fini}}$ l'axiome de l'infini.

▷ *Sous la forme ci-dessus, l'axiome de l'infini n'est pas une formule ensembliste puisqu'y figurent le nom* $\emptyset$ *et à l'opération unaire* S*. En remplaçant ceux-ci par des définitions, on obtient une formule équivalente modulo les axiomes de* $\mathsf{Z}_{\mathsf{fini}}$ *et les axiomes d'introduction, par exemple*

$$\exists a\, \exists e\, (e{\in}a \text{ et } \forall t(t{\in}e{\Rightarrow}t \neq t)$$
$$\text{et } \forall x(x{\in}a{\Rightarrow}\, \exists y\, (y{\in}a \text{ et } \forall t(t{\in}y \Leftrightarrow (t{\in}x \text{ ou } t = x))))),$$

authentique formule ensembliste, parfaitement illisible.

L'axiome de l'infini ne postule pas directement l'existence de l'ordinal ω*, plus petit ordinal infini, mais il est facile de déduire celle-ci des autres axiomes de Zermelo. Pour cela, il est commode d'introduire la notion auxiliaire d'ensemble récurrent.* ◁

Définition 2.12. (récurrent) On appelle *récurrent* tout ensemble contenant $\emptyset$ et clos par l'application $S : x \mapsto x \cup \{x\}$.

L'axiome de l'infini affirme l'existence d'un ensemble récurrent.

Lemme 2.13. ($\mathsf{Z}_{\mathsf{fini}}$) *S'il existe un ensemble récurrent, alors il existe un plus petit ordinal récurrent, et ce dernier est inclus dans tout ensemble récurrent.*

Démonstration. On note que toute intersection d'ensembles récurrents est récurrente. Supposons que a est un ensemble récurrent. Soit X l'intersection de tous les ensembles récurrents inclus dans a. D'après la remarque précédente, X est récurrent, et il est inclus dans tout ensemble récurrent, inclus ou non dans a, car, si b est récurrent, on a $X \subseteq a \cap b \subseteq b$. Par ailleurs, l'ensemble a' des ordinaux appartenant à a est récurrent, donc on a $X \subseteq a'$, et, par conséquent, X est un ensemble d'ordinaux. Soit alors ω le plus petit ordinal qui n'est pas dans X. Par construction, tout ordinal plus petit que ω est dans X, donc on a $\omega \subseteq X$. Par ailleurs, on a $\underline{0} \in \omega$, c'est-à-dire $\omega > \emptyset$, puisque $\emptyset$ est dans X ; d'autre part, $\alpha \in \omega$, c'est-à-dire $\omega > \alpha$, entraîne $\alpha \in X$, donc $S(\alpha) \in X$ puisque X est récurrent, donc $\omega > S(\alpha)$, c'est-à-dire $S(\alpha) \in \omega$. Par conséquent, ω est récurrent, et on a donc $X \subseteq \omega$, d'où, finalement $\omega = X$. ☐

Définition 2.14. (omega) On note ω le plus petit ordinal récurrent.

En vertu de la relation $\emptyset \notin \emptyset$, l'ensemble vide n'est pas récurrent, et une récurrence montre qu'aucun des ordinaux $\underline{\mathsf{n}}$ ne l'est. On a donc $\underline{\mathsf{n}} < \omega$ pour tout entier naturel n.

[6] un point sur lequel on pourra réfléchir et qu'on discutera plus loin

QUESTION 2.15. *A-t-on* $\omega = \sup\{\underline{\mathsf{n}}\,;\, \mathsf{n} \in \mathbb{N}\}$ *?*

▷ *La question peut paraître étrange, puisqu'au chapitre II on a défini* ω *comme borne supérieure des ordinaux* $\underline{\mathsf{n}}$. *Le point est qu'il n'est pas évident que la définition 2.14 de* ω *soit équivalente à la définition II.2.18, et c'est précisément la question posée, la difficulté tenant à la référence à un ensemble* $\mathbb{N}$ *extérieur.*

On y reviendra au chapitre IX. Pour le moment, le point important est qu'on peut travailler avec l'ordinal ω *ainsi (re)défini exactement comme on l'escompte. Vu le point de vue adopté ici, il est naturel de (re)définir les notions d'entier et d'ensemble fini pour les rendre cohérentes avec la définition de* ω. ◁

DÉFINITION 2.16. (fini, dénombrable) On appelle *entiers* les éléments de ω. Un ensemble est dit *fini* (*resp. dénombrable*) s'il est en bijection avec un entier (*resp.* avec ω).

▷ *De la sorte, et quelle que soit la réponse à la question 2.15, l'ordinal* ω *est la borne supérieure des entiers. Noter que, puisque les fonctions, donc en particulier les bijections, sont définies en termes ensemblistes, il en est de même des relations « être fini » et « être infini ».*

Comme le stipule la convention 1.18, on utilise dans la suite les notations traditionnelles n, k, i, *etc. pour les entiers. Il résulte de la construction spécifique des ordinaux adoptée ici que tout ordinal* α *coïncide avec le segment initial* $[0, \alpha[$ *qu'il détermine : ceci vaut en particulier pour les entiers, et on a donc* $n = [0, n[$ *pour tout entier* n.

Faute d'une réponse positive démontrée à la question 2.15, on continuera à utiliser une notation distincte, à savoir $\mathsf{n}, \mathsf{k}, \mathsf{i}$, *etc. , pour les entiers naturels, qu'on appellera intuitifs, qui interviennent dans les définitions extérieures au monde des ensembles, typiquement les formules mathématiques. Ceci ne doit pas ouvrir des abîmes de perplexité : pour le moment, la distinction typographique est sans importance réelle, et ce n'est qu'à partir du chapitre IX qu'il deviendra nécessaire d'y prendre garde.*

Par contre, dans la mesure où la définition des entiers a été reformulée — même si elle vise à définir les mêmes objets — il est nécessaire de vérifier que les résultats précédemment énoncés et mettant en jeu des entiers restent valables dans le nouveau cadre formel. Ci-dessous, on examine successivement le schéma d'induction sur les entiers, puis les résultats affirmant que tout entier non nul est un successeur, que toute injection d'un ensemble fini dans lui-même est une bijection, et enfin que l'ensemble des entiers fournit un modèle pour le système de Peano. ◁

PROPOSITION 2.17. (induction sur les entiers) *Supposons que* F *est une formule ensembliste telle que* $\mathsf{F}(0, a_1, \ldots, a_\mathsf{k})$ *est vraie et que, si* $\mathsf{F}(n, a_1, \ldots, a_\mathsf{k})$ *est vraie, alors* $\mathsf{F}(S(n), a_1, \ldots, a_\mathsf{k})$ *l'est aussi. Alors* $\mathsf{F}(n, a_1, \ldots, a_\mathsf{k})$ *est vraie pour tout entier* n.

DÉMONSTRATION. Soit $A = \{n \in \omega\,;\, \mathsf{F}(n, a_1, \ldots, a_\mathsf{n})\}$. Un axiome de séparation justifie l'existence de A. L'hypothèse signifie que A est récurrent. Par le lemme II.2.13, A doit inclure ω, donc lui être égal. □

LEMME 2.18. *Tout entier différent de* 0 *est un successeur.*

DÉMONSTRATION. Soit $\mathsf{F}(x)$ la formule « $x = 0$ ou $\exists y (x = S(y))$ ». Alors $\mathsf{F}(0)$ est vraie, et $\mathsf{F}(n)$ entraîne $\mathsf{F}(S(n))$. Par induction, $\mathsf{F}(n)$ est donc vraie pour tout entier n. □

PROPOSITION 2.19. (fini) (Z) *Toute injection d'un ensemble fini dans lui-même est bijective. Par contre, pour tout ordinal infini* α, *il existe une injection non surjective de* α *dans lui-même.*

Démonstration. Pour la première propriété, comme au chapitre I, il suffit de considérer le cas d'un entier, c'est-à-dire d'un ordinal fini. On utilise alors une induction sur ω. Par défaut, toute injection de 0, c'est-à-dire de $\emptyset$, dans lui-même est bijective. Ensuite, soit n un ordinal fini quelconque, et soit f une injection de $S(n)$ dans lui-même. On a $S(n) = n \cup \{n\}$. Soit $k = f(n)$. On définit g de n dans lui-même en posant $g(i) = f(i)$ pour $f(i) < k$, et $g(i) =$ l'unique ordinal fini j vérifiant $S(j) = f(i)$ pour $f(i) > k$. Le lemme 2.18 entraîne que g est partout définie. L'existence de g comme ensemble est garantie par le fait qu'elle est définie à partir des paramètres f, n, et k, et de là justifiée par un axiome de séparation à l'intérieur de $\mathrm{Fonc}(n, n)$. Ensuite, on vérifie que g est injective, donc bijective par hypothèse d'induction, et on conclut que l'image de f est $S(n)$ entier.

Pour la seconde propriété, on considère d'abord le cas de ω. On sait que l'opération successeur S fournit une injection non surjective de ω dans lui-même, mais il s'agit de justifier le fait que la restriction $S\restriction_\omega$, vue comme famille de couples, est bien un ensemble. Comme S est définie par une formule ensembliste, c'est facile :

$$S\restriction_\omega = \{x \in \omega \times \omega\,;\, \exists n{\in}\omega(c = (n, n \cup \{n\}))\}$$

convient, et fournit une injection non surjective de ω dans lui-même. Pour $\alpha \geqslant \omega$, on justifie de façon analogue l'existence d'une application (injective et non surjective) de α dans lui-même qui coïncide avec $S\restriction_\omega$ sur ω et avec l'application identité sur le complémentaire $[\omega, \alpha[$. □

Proposition 2.20. (Peano) (Z) *La structure* $(\omega, 0, S, +, \cdot)$ *satisfait aux axiomes du système de Peano* PA.

Démonstration. Le lemme 1.9 n'était qu'une tautologie dans la mesure où on prenait l'égalité $\omega = \underline{\mathbb{N}}$ comme point de départ. Désormais la règle du jeu a changé : on ne s'autorise plus à utiliser l'hypothèse $\omega = \underline{\mathbb{N}}$ — laquelle ne fait tout simplement pas sens dans l'univers des ensembles purs — mais seulement celle que ω est le plus petit ordinal récurrent.

Les propriétés des opérations successeurs, addition, et multiplication ne posent pas de problème. Par contre, ce qui pourrait poser problème est la clôture de ω, c'est-à-dire le fait que, si n, p sont des entiers, alors il en est de même de $S(n)$, $n + p$, et np. Pour le successeur, la propriété résulte de la définition d'un ensemble récurrent.

Pour l'addition, on montre par induction sur p que, pour tout ordinal fini n, l'ordinal $n{+}p$ est fini. Pour $p = 0$, on a $n{+}p = n{+}0 = n < \omega$ par hypothèse. Supposons $p > 0$. Par le lemme 2.18, il existe un ordinal fini q tel que p est $S(q)$. On a alors $n + p = n + S(q) = S(n + q) < \omega$, puisque, par hypothèse d'induction, $n + q$ est un ordinal fini.

Ensuite, il reste à vérifier la légitimité de $+\restriction_\omega$ en tant qu'ensemble de couples dans $(\omega \times \omega) \times \omega$. Or $+\restriction_\omega$ est l'ensemble des $((p, q), r)$ dans $(\omega \times \omega) \times \omega$ tels qu'il existe un isomorphisme de $(p, \in) \sqcup (q, \in)$ sur $(r, \in)$: cette dermière propriété étant exprimable par une formule ensembliste, l'existence de l'ensemble $+\restriction_\omega$ résulte d'un axiome de séparation.

Les vérifications sont analogues pour la multiplication, une fois qu'on sait que ω est clos par addition. □

2.5. Les axiomes de remplacement.

▶ Pour établir le théorème de comparaison qui affirme que tout ensemble bien ordonné est isomorphe à un (unique) ordinal, on introduit une nouvelle famille d'axiomes, les axiomes de remplacement. On parvient ainsi essentiellement au système de Zermelo–Fraenkel, à partir duquel tous les résultats du chapitre II sont légitimés. ◀

▷ *On reprend l'examen des résultats sur les ordinaux établis au chapitre II, avec la section II.2.5 et le théorème de comparaison (proposition II.2.15) qui affirme que tout ensemble bien ordonné* *est isomorphe à un (unique) ordinal. Ce théorème est ensuite utilisé de nombreuses fois, tant pour construire les opérations de l'arithmétique ordinale que pour justifier l'existence d'ordinaux non dénombrables.*

Un nouveau problème apparaît. Avec les notations du chapitre II, il s'agit de partir d'un ensemble bien ordonné $(A,<)$ *et de construire un isomorphisme de* $(A,<)$ *sur un ordinal en considérant la correspondance* F *qui met en relation un élément* a *de* A *et un ordinal* α *si et seulement si le segment initial* $I(a)$ *de* $(A,<)$ *déterminé par* a *est isomorphe à l'ordinal* α*. Que* F *soit fonctionnelle et injective, et que le domaine et l'image de* F *soient clos par minorant ne pose pas de problème. Par contre, pour conclure que le domaine de* F *est* A *entier, il faut pouvoir rejeter l'hypothèse que la suite des ordinaux est l'image de* F *ou un segment initial de celle-ci, c'est-à-dire rejeter l'hypothèse que l'image de* F *contient tous les ordinaux. On l'a fait au motif qu'aucun ensemble ne contient tous les ordinaux. La validité du dernier point n'est pas mise en cause, mais on ne peut conclure que si on sait que l'image de* F *est un ensemble, ce que rien pour le moment ne garantit. En effet, on sait seulement que* F *est fonctionnelle et que son domaine est inclus dans l'ensemble* A*. Comme* F *est définie par une formule ensembliste, l'existence du domaine de* F *comme sous-ensemble de* A *est garantie par un axiome de séparation puisqu'on a*

$$\mathrm{Dom}(F) = \{x \in A\,;\, \exists\alpha\ \exists f{:}I(x){\to}\alpha\ (f \textit{ isomorphisme de } (I(x),<) \textit{ sur } (\alpha,\in))\}$$ [7],

et que les propriétés « $f{:}I(x){\to}\alpha$ » et « f isomorphisme de $(I(x),<)$ sur $(\alpha,\in)$ » sont exprimables par des formules ensemblistes (avec les paramètres A, x, et $<$) : pour la première, on l'a vu dans le lemme 2.7, pour la seconde les détails sont laissés au lecteur. ◁

QUESTION 2.21. *L'image d'une correspondance (fonctionnelle, ou même bijective) dont le domaine est un ensemble est-elle nécessairement un ensemble ?*

▷ *Si on s'en tient à l'intuition que seule la taille excessive d'un éventuel ensemble de tous les ensembles est cause du paradoxe de Russel, il est naturel de postuler une réponse positive. Pour autant, il ne serait pas étonnant que les axiomes de séparation échouent à la démontrer puisqu'ils ne permettent pas de s'extraire d'un ensemble supposé préexister. On est donc conduit à introduire un nouvel axiome. Pour énoncer celui-ci à l'aide de formules ensemblistes, on ne peut mentionner directement une correspondance* F *qui, a priori n'est pas donnée comme ensemble*[8] *; dans l'exemple du théorème de comparaison, la correspondance* F *est définie par une formule ensembliste, au sens où il existe une formule ensembliste* $\mathsf{F}(x,\alpha,A,<)$ *telle qu'on peut montrer à partir de* Z *que, pour tout* x*, il existe au plus un* α *vérifiant* $\mathsf{F}(x,\alpha,A,<)$*. C'est cette situation qu'on prend comme modèle.* ◁

DÉFINITION 2.22. (remplacement, $\mathsf{ZF}_\bullet$) Pour $\mathsf{F}(x,y,\vec{c})$ formule ensembliste où a et b n'apparaissent pas comme variables libres, on appelle axiome de *remplacement* pour F l'énoncé Remp_F suivant :

$$\forall a\,\forall\vec{c}\,((\forall x,y,z((\mathsf{F}(x,y,\vec{c}) \text{ et } \mathsf{F}(x,z,\vec{c})) \Rightarrow y=z) \\ \Rightarrow \exists b\,\forall y\,(\,\exists x{\in}a(\mathsf{F}(x,y,\vec{c})) \Rightarrow y \in b))).$$

On note $\mathsf{ZF}_\bullet$ le système obtenu en ajoutant à Z la famille de tous les axiomes de remplacement[9].

[7] on utilise ici $\exists f{:}A{\to}B(...)$ comme abréviation pour $\exists f(f \in \mathrm{Fonc}(A,B)$ et $...)$

[8] si on sait que F, en tant que composée de couples, est un ensemble, alors on déduit directement du lemme 2.7 que l'image de F est un ensemble

[9] Comme dans le cas du système de Zermelo avec les axiomes de séparation, on ajoute dans le système $\mathsf{ZF}_\bullet$ tous les axiomes de remplacement, quelle que soit la signification des formules considérées. On pourrait proposer de restreindre la liste des axiomes de remplacement aux seules formules F qui sont fonctionnelles en leur second argument. Ceci supposerait de savoir reconnaître *a priori* de telles formules, ce qui dépend des autres axiomes et de la possibilité de

L'axiome de remplacement pour F exprime que, si , pour des valeurs $\vec{c}$ des éventuels paramètres, $\mathsf{F}(x, y, \vec{c})$ définit une correspondance fonctionnelle en y, alors, pour tout ensemble a, il existe un ensemble b contenant les images des éléments de a par $\mathsf{F}(\cdot, \cdot, \vec{c})$, c'est-à-dire les éléments y vérifiant $\mathsf{F}(x, y, \vec{c})$ pour au moins un élément x de a. En présence des axiomes de séparation et d'extensionnalité, les axiomes de remplacement garantissent l'existence d'un unique ensemble formé par les images des éléments de a par la correspondance définie par F, ensemble sera naturellement noté $\{y\,;\,\exists x{\in}a(\mathsf{F}(x, y, \vec{c}))\}$.

$\triangleright$ *Ayant ainsi enrichi la base axiomatique, on revient à la construction des ordinaux. D'abord, on peut maintenant conclure la démonstration du théorème de comparaison.* $\triangleleft$

PROPOSITION 2.23. (comparaison) ($\mathsf{ZF}_\bullet$) *Tout ensemble bien ordonné est isomorphe à un ordinal.*

DÉMONSTRATION. On a fait ce qu'il fallait : les cas à exclure sont ceux où l'image de la correspondance F de la proposition II.2.15 contiendrait tous les ordinaux. Comme ladite correspondance est définie par une formule ensembliste, et qu'on a montré à partir des axiomes du système Z qu'elle est fonctionnelle, un axiome de remplacement assure que son image est un ensemble, et donc, en vertu du paradoxe de Burali–Forti, ne peut contenir tous les ordinaux. Le seul cas possible est donc celui où il existe un ordinal α tel que l'ensemble bien ordonné $(A, <)$ de départ est isomorphe à $(\alpha, \in)$. $\square$

$\triangleright$ *On passe à la dernière partie du chapitre II, avec l'introduction des opérations de l'arithmétique ordinale, et la démonstration de l'existence d'ordinaux non dénombrables.*

La construction, pour chaque couple d'ensembles bien ordonnés $\mathcal{A}, \mathcal{B}$, de l'ensemble-somme $\mathcal{A} + \mathcal{B}$, ainsi que la démonstration du fait que la somme de deux bons ordres est un bon ordre, ne pose pas de problème. Notons alors $\mathbf{Add}(\alpha, \beta, \gamma)$ la formule ensembliste

$$\mathbf{Ord}(\alpha) \text{ et } \mathbf{Ord}(\beta) \text{ et } \mathbf{Ord}(\gamma) \text{ et } (\alpha, \in) + (\beta, \in) \cong (\gamma, \in).$$

Cette formule définit l'addition ordinale au sens où $\mathbf{Add}(\alpha, \beta, \gamma)$ équivaut à $\alpha + \beta = \gamma$. Il en résulte en particulier que, pour tout ordinal θ, il existe par séparation un ensemble $\{(\alpha, \beta, \gamma) \in \theta^3\,;\, \mathbf{Add}(\alpha, \beta, \gamma)\}$, soit encore $\{(\alpha, \beta, \gamma) \in \theta^3\,;\, \alpha + \beta = \gamma\}$, qu'on peut noter simplement $+\restriction_\theta$. Jusqu'à présent, on n'a pas fait usage d'axiome de remplacement, et l'existence des ensembles $+\restriction_\theta$ est garantie à partir des axiomes de Z. Par contre, l'énoncé $\forall \alpha, \beta\, \exists \gamma\ (\gamma = \alpha + \beta)$ affirmant que l'addition est partout définie requiert a priori le théorème de comparaison, et donc l'usage du remplacement. Par conséquent, l'introduction d'un ordinal tel que $\omega + \omega$ n'est pour le moment garantie qu'à partir du système $\mathsf{ZF}_\bullet$[10].

Les cas de la multiplication et de l'exponentiation ordinale sont similaires, et donc, dans le système $\mathsf{ZF}_\bullet$, on peut montrer l'existence d'ordinaux tels que $\omega \cdot \omega$ ou ω^ω, et, plus généralement, démontrer toutes les propriétés des opérations établies au chapitre II. En particulier, le théorème de convergence des suites de Goodstein, qui utilise toute l'arithmétique ordinale, est démontrable à partir des axiomes du système $\mathsf{ZF}_\bullet$.

Considérons enfin les ordinaux non dénombrables de la section II.3.5 : il s'agit de réexaminer la proposition II.3.20. $\triangleleft$

---

décider si une formule donnée est ou non conséquence des axiomes: comme on le verra plus loin, ce n'est possible ni en pratique, ni même en théorie, sauf à détruire la possibilité d'énumérer effectivement les axiomes de ZF.

[10]on verra plus loin que l'usage du remplacement est effectivement indispensable : on ne peut pas montrer à partir du système de Zermelo l'existence de $\omega + \omega$; pour le moment, on ne peut que constater qu'on ne sait pas montrer cette existence à partir de Z, ce qui bien sûr ne prouve pas qu'elle ne pourrait pas l'être avec davantage de travail

Proposition 2.24. (non injection) *Pour chaque ordinal infini α, l'ensemble des ordinaux s'injectant dans α est un ordinal, et c'est le plus petit ordinal qui ne s'injecte pas dans α.*

Démonstration. Le problème est de justifier l'existence d'un ensemble Θ formé par tous les ordinaux ne s'injectant pas dans α. Ensuite, le fait que l'ordinal $\bigcup\Theta$ ne s'injecte pas dans α est facile. Le problème est que, comme les ordinaux ne forment pas un ensemble, l'existence de Θ ne peut *a priori* pas être établie par séparation. Par contre, s'il existe une bijection f entre un ordinal β et α ou un segment initial de α, alors transporter par f le bon ordre de β fournit un bon ordre R sur α ou un sous-ensemble de α. Or un bon ordre sur une partie de α est un sous-ensemble de $\mathfrak{P}(\alpha \times \alpha)$, et, la propriété d'être un bon ordre étant exprimable par une formule ensembliste, un axiome de séparation garantit l'existence de

$$W = \{R \in \mathfrak{P}(\alpha \times \alpha)\,;\, \text{« } R \text{ est un bon ordre sur une partie de } \alpha \text{ »}\}.$$

Soit alors $\mathsf{F}(R, \beta)$ la formule ensembliste

$$\mathbf{Ord}(\beta) \text{ et } (\mathrm{Dom}(R), R) \cong (\beta, \in).$$

Pour toute relation R dans W, il existe au plus un ordinal β tel que l'ensemble bien ordonné $(\mathrm{Dom}(R), R)$ soit isomorphe à $(\beta, \in)$. Par conséquent, la formule $\mathsf{F}(R, \beta)$ définit une correspondance fonctionnelle, c'est-à-dire que la première partie de l'axiome de remplacement pour F est satisfaite. Appliquant cet axiome dans le cas de l'ensemble W, on obtient l'existence d'un ensemble de tous les ordinaux β tels que $(\beta, \in)$ soit isomorphe à un bon ordre sur une partie de α, c'est-à-dire précisément l'existence de l'ensemble Θ de tous les ordinaux s'injectant dans α, qui était le point à justifier. □

Par conséquent, le système $\mathsf{ZF}_\bullet$ garantit l'existence des ordinaux non dénombrables $\omega_1, \omega_2, \ldots$, ce qui achève la validation des résultats du chapitre II.

▷ *On peut se demander si l'introduction des axiomes de remplacement permet de décider la question 2.15. La réponse est négative. Certes, la correspondance qui associe à chaque entier naturel* n *l'ordinal fini* $\underline{\mathsf{n}}$ *est fonctionnelle, mais le problème est que l'ensemble* $\mathbb{N}$ *des entiers naturels n'est pas un ensemble pur, et qu'aucun axiome de* $\mathsf{ZF}_\bullet$ *ne fonde a priori l'existence de cet ensemble extérieur à l'univers des ensembles purs. On montrera au chapitre IX que le passage de* Z *à* $\mathsf{ZF}_\bullet$ *ne change rien à la question.* ◁

3. Définitions récursives

▶ On montre que le système $\mathsf{ZF}_\bullet$ légitime la construction d'ensembles par le biais de définitions récursives ; plus précisément, pour chaque définition récursive, on établit à partir des axiomes de Zermelo l'existence pour chaque ordinal θ, d'une suite indexée par θ satisfaisant la clause de récursion, et, à partir des axiomes de $\mathsf{ZF}_\bullet$ et moyennant l'introduction d'une notion de classe convenable, l'existence d'une telle suite indexée par tous les ordinaux. ◀

▷ *On a, dans la partie précédente, revu et légitimé sur la base des axiomes de* $\mathsf{ZF}_\bullet$*, tous les résultats établis au chapitre II et dans la première partie du chapitre III, à une exception près : la définition des ensembles purs à partir des ensembles* V_α*. Le problème est le suivant : les axiomes de* $\mathsf{ZF}_\bullet$ *— et même déjà de* Z *— affirment l'existence pour chaque ensemble* a *de l'ensemble des parties* $\mathfrak{P}(a)$ *ou celle de l'union* $\bigcup a$*, mais est-ce suffisant pour justifier l'existence d'une suite définie en appliquant les opérations* $\mathfrak{P}$ *en* $\bigcup$ *non pas une fois, mais de façon itérée, donc typiquement un nombre infini d'opérations dans le cas des ensembles* V_α *avec* α *transfini ?*

La réponse est positive dans le cas de la suite des V_α, mais également d'une façon générale pour toute suite définie par récursion ordinale. Ce point est très important pour les développements ultérieurs, et il mérite d'être étudié en détail : c'est l'objet de cette section. ◁

Notation 3.1. Dans la suite, on utilise à la fois le vocabulaire des fonctions et celui des suites : ainsi, $(x_i)_{i\in I}$ est une notation alternative pour l'application f de domaine I vérifiant $f(i) = x_i$. Dans le cas d'une suite indexée par un ordinal θ, on peut aussi noter $(x_\alpha)_{\alpha<\theta}$ puisque θ coïncide avec l'ensemble des ordinaux α vérifiant $\alpha < \theta$; on parle alors de *suite de longueur* θ, ou encore de suite indexée par θ, ou encore, simplement, de θ-suite. L'ensemble de toutes les suites de longueur θ à valeurs dans A est noté A^θ [11].

3.1. Récursion sur les entiers.

► On montre que les axiomes de Z légitiment la construction de suites de longueur ω, c'est-à-dire de suites indexées par les entiers, par une définition récursive. ◄

▷ *Dans la pratique mathématique, il est fréquent d'introduire un objet, par exemple une suite $(x_n)_n$ indexées par les entiers, non par une définition directe $x_n := F(n)$, mais par une définition récursive $x_n = F(n, x_{n-1})$ où la valeur en n dépend non seulement de n mais aussi de la valeur en $n-1$, ou plus généralement de la famille des valeurs en tous les entiers inférieurs à n. La question posée ici est de savoir si les axiomes de Z autorisent de telles constructions, ou si au contraire l'introduction de nouveaux axiomes paraît nécessaire pour justifier de telles constructions. La réponse est que Z est suffisant. Comme dans ce qui précède et dans toute la suite, les entiers dont nous parlons ici sont évidemment les ordinaux finis, éléments de ω.* ◁

Proposition 3.2. (récursion sur les entiers) *Soit A un ensemble, et F une application de $\omega \times \bigcup_{n<\omega} A^n$ dans A [12]. Alors il existe une unique application G de ω dans A vérifiant pour tout entier n*

$$G(n) = F(\,n, G\restriction_{[0,n[}\,), \tag{3.1}$$

c'est-à-dire encore une unique suite $(x_n)_{n\in\omega}$ vérifiant pour tout entier n

$$x_n = F(\,n, (x_p)_{p<n}\,). \tag{3.2}$$

Démonstration. On utilise la notation de fonction. Pour p entier, appelons *p-solution* toute fonction (partielle) g de ω dans A dont le domaine est l'intervalle $[0, p[$ et qui satisfait, en tout point n de son domaine, la clause de récursion $g(n) = F(n, g\restriction_{[0,n[})$. Noter que, pour $q \leqslant p$ et par définition, la restriction à $[0, q[$ d'une p-solution est une q-solution. On va montrer, par induction sur p, que, pour tout p, il existe une unique p-solution g_p.

Pour $p = 0$, la fonction vide est la seule fonction de domaine vide, et le résultat est donc valide avec $g_0 := \emptyset$.

Supposons $p > 0$. Alors il existe q vérifiant $p = q + 1$, et, par hypothèse d'induction, il existe une unique q-solution g_q. Posons $g_p := g_q \cup \{(q, F(q, g_q))\}$. Alors g_p est une p-solution.

---

[11] On remarquera que, pour tout entier intuitif n, il existe une bijection évidente entre le produit cartésien $A \times ... \times A$, n termes, et l'ensemble A^{n} des suites de longueur n à valeurs dans A.

[12] Noter que l'existence de l'ensemble $\bigcup_{n\in\omega} A^n$ ne requiert pas d'axiome de remplacement : on l'obtient en séparant à l'intérieur de l'ensemble $\mathrm{Fonc}(\omega, A)$ les fonctions dont le domaine est un intervalle du type $[0, n[$.

En effet, g_p est une fonction de ω dans A dont le domaine est $[0,q[\,\cup\{q\}$, c'est-à-dire $[0,p[$, et la condition $g_p(n) = F(n, g_p{\restriction}_{[0,n[})$ est vérifiée pour tout $n < p$: pour $n < q$, cela résulte des égalités $g_p(n) = g_q(n)$ et $g_p{\restriction}_{[0,n[} = g_q{\restriction}_{[0,n[}$ et de ce que g_q est une q-solution, et, pour $n = q$, cela résulte de la construction de g_p.

Inversement, soit g une p-solution quelconque. Alors $g{\restriction}_{[0,q[}$ est une q-solution, donc, par hypothèse d'induction, on a $g{\restriction}_{[0,q[} = g_q$, donc aussi $g{\restriction}_{[0,q[} = g_p{\restriction}_{[0,q[}$. Par ailleurs, puisque g est une p-solution, on doit avoir $g(q) = F(q, g{\restriction}_{[0,q[}) = F(q, g_q) = g_p(q)$, et, finalement, $g = g_p$.

Posons alors $G := \bigcup_{p<\omega} g_p$. L'existence de G doit être légitimée, et il ne suffit pas d'alléguer l'axiome de l'union, car ce dernier ne garantit l'existence de G que pour autant que l'ensemble $\{g_p\,;\, p < \omega\}$ existe lui-même. Mais ceci ne pose pas de problème, puisque ce dernier ensemble, par exemple, peut être spécifié comme

$$\{g : \omega \to A\,;\, \exists p{\in}\omega(\mathrm{Dom}(g) = [0,p[\text{ et } \forall n{<}\, p(g(n) = F(n, g{\restriction}_{[0,n[})))\},$$

dont l'existence est garantie par séparation. Comme, par hypothèse d'induction, les fonctions partielles g_p sont deux à deux compatibles sur l'intersection de leurs domaines, G est une fonction, et, puisqu'il existe une p-solution pour tout p, le domaine de G est l'ensemble ω entier. Ensuite G satisfait (3.3) pour tout n puisqu'on a

$$G(n) = g_{n+1}(n) = F(n, g_{n+1}{\restriction}_{[0,n[}) = F(n, G{\restriction}_{[0,n[}).$$

Enfin, supposons que G' est une application quelconque satisfaisant (3.3). Pour tout entier p, la restriction $G'{\restriction}_{[0,p[}$ est une p-solution, donc on a $G'{\restriction}_{[0,p[} = g_p = G{\restriction}_{[0,p[}$, d'où $G' = G$, et, par conséquent, G est l'unique solution du problème. □

▷ *Le résultat précédent est naturel et simple : on ne l'a établi avec soin que pour montrer que le recours aux ordinaux finis ne pose aucun problème, et d'autre part pour rendre explicite l'utilisation des axiomes affirmant l'existence d'un ensemble. Une fois la démonstration précédente bien comprise, l'extension aux ordinaux transfinis ne devrait apparaître que comme une variante facile.*

On pourra noter que, vue la construction des ordinaux adoptée ici, l'entier n coïncide avec l'intervalle $[0,n[$, et on peut donc également écrire la condition de récursion (3.3) *sous la forme $G(n) = F(G{\restriction}_n)$; malgré tout, on s'en tiendra ici à la notation $[0,n[$ qui rend plus visible le principe de la récursion, à savoir que ce qui est pris en compte est la suite de toutes les valeurs antérieurement obtenues.*

Avant de passer aux ordinaux , notons que la forme restreinte de récursion « $n \to n+1$ » sur les entiers se déduit immédiatement de la forme générale qu'on vient d'établir. ◁

Corollaire 3.3. (Z) *Soit A un ensemble, a un élément de A, et F une application de $\omega \times A$ dans A. Alors il existe une unique application G de ω dans A vérifiant*

$$G(0) = a \quad \text{et, pour tout entier } n, \quad G(n+1) = F(n, G(n)), \tag{3.3}$$

c'est-à-dire encore une unique suite $(x_n)_{n\in\omega}$ vérifiant

$$x_0 = a \quad \text{et, pour tout entier } n, \quad x_{n+1} = F(n, x_n). \tag{3.4}$$

Démonstration. Appliquer la proposition 3.2 à la fonction F' définie par $F'(0,\emptyset) := a$ et $F'(n+1, g) = F(n, g(n))$ pour $n \geqslant 0$ et g application de $[0,n+1[$ dans A. □

3.2. Récursion ordinale.

▶ Dans ce qui précède, l'ordinal ω ne joue pas de rôle particulier, et on peut étendre la construction précédente pour y remplacer ω par un ordinal θ quelconque. ◀

PROPOSITION 3.4. (récursion ordinale) *Soit A un ensemble, θ un ordinal, et F une application de $\theta \times \bigcup_{\alpha<\theta} A^\alpha$ dans A. Alors il existe une unique application G de $[0,\theta[$ vérifiant pour tout ordinal $\alpha < \theta$*

$$G(\alpha) = F(\,\alpha, G{\restriction}_{[0,\alpha[}\,), \tag{3.5}$$

c'est-à-dire encore une unique suite $(x_\alpha)_{\alpha<\theta}$ vérifiant pour tout ordinal $\alpha < \theta$

$$x_\alpha = F(\,\alpha, (x_\beta)_{\beta<\alpha}\,). \tag{3.6}$$

DÉMONSTRATION. L'argument est essentiellement le même que pour la proposition 3.2. Vue son importance, on le répète néanmoins, *mutatis mutandis.* On utilise à nouveau la notation de fonction, et, pour $\beta \leqslant \theta$, on appelle *β-solution* toute fonction (partielle) g de θ dans A dont le domaine est l'intervalle $[0,\beta[$ et qui satisfait, en tout α de son domaine, la clause de récursion $g(\alpha) = F(\alpha, g{\restriction}_{[0,\alpha[})$. A nouveau, on note que, pour $\gamma \leqslant \beta$ et par définition, la restriction à $[0,\gamma[$ d'une β-solution est une γ-solution. On va montrer, par induction sur β, que, pour tout $\beta \leqslant \theta$, il existe une unique β-solution g_β.

Pour $\beta = 0$, la fonction vide est la seule fonction de domaine vide, et le résultat est donc valide avec $g_0 := \emptyset$.

Supposons $\beta > 0$ avec β successeur, soit $\beta = \gamma + 1$. Par hypothèse d'induction, il existe une unique γ-solution g_γ. Posons $g_\beta := g_\gamma \cup \{(\gamma, F(\gamma, g_\gamma))\}$. Alors g_β est une β-solution. En effet, g_β est une fonction de θ dans A dont le domaine est $[0,\gamma[\cup\{\gamma\}$, c'est-à-dire $[0,\beta[$, et la condition $g_\beta(\alpha) = F(\alpha, g_\beta{\restriction}_{[0,\alpha[})$ est vérifiée pour tout $\alpha < \beta$: pour $\alpha < \gamma$, cela résulte des égalités $g_\beta(\alpha) = g_\gamma(\alpha)$ et $g_\beta{\restriction}_{[0,\alpha[} = g_\gamma{\restriction}_{[0,\alpha[}$ et de ce que g_γ est une γ-solution, et, pour $\alpha = \gamma$, cela résulte de la construction de g_β.

Inversement, soit g une β-solution quelconque. Alors $g{\restriction}_{[0,\gamma[}$ est une γ-solution, donc, par hypothèse d'induction, on a $g{\restriction}_{[0,\gamma[} = g_\gamma = g_\beta{\restriction}_{[0,\gamma[}$. Par ailleurs, puisque g est une β-solution, on doit avoir $g(\gamma) = F(\gamma, g{\restriction}_{[0,\gamma[}) = F(\gamma, g_\gamma) = g_\beta(\gamma)$, soit, finalement, $g = g_\beta$.

Supposons maintenant β limite non nul. Par hypothèse d'induction, il existe une unique γ-solution pour chaque $\gamma < \beta$, et on a

$$\{g_\gamma\,;\, \gamma < \beta\} = \{g : \theta \to A\,;\, \exists\gamma\,{\in}\,\beta(\mathrm{Dom}(g) = [0,\gamma[\text{ et } \forall\alpha{<}\,\gamma\,(g(\alpha) = F(\alpha, g{\restriction}_{[0,\alpha[})))\},$$

ce qui, par séparation, garantit l'existence de cet ensemble. Par l'axiome de l'union, on peut alors introduire $g_\beta := \bigcup\{g_\gamma\,;\, \gamma < \beta\}$, c'est-à-dire $g_\beta = \bigcup_{\gamma<\beta} g_\gamma$. Toujours par hypothèse d'induction, les fonctions partielles g_γ sont deux à deux compatibles sur l'intersection de leurs domaines, donc g_β est une fonction, et, puisqu'il existe une γ-solution pour tout $\gamma < \beta$, le domaine de g_β est $[0,\beta[$ entier. Ensuite g_β est une β-solution. En effet, pour $\alpha < \beta$, on a $g_\beta(\alpha) = g_{\alpha+1}(\alpha) = F(\alpha, g_{\alpha+1}{\restriction}_{[0,\alpha[}) = F(\alpha, g_\beta{\restriction}_{[0,\alpha[})$.

Inversement, soit g une β-solution quelcoinque. Alors, pour tout γ vérifiant $\gamma < \beta$, la fonction $g{\restriction}_{[0,\gamma[}$ est une γ-solution, donc, par hypothèse d'induction, on a $g{\restriction}_{[0,\gamma[} = g_\gamma = g_\beta{\restriction}_{[0,\gamma[}$, d'où $g = \bigcup_{\gamma<\beta} g_\gamma = g_\beta$, soit, finalement, $g = g_\beta$.

Il suffit alors de poser $G := g_\theta$ pour conclure.[13] □

[13] On notera une légère différence entre la fin de cette démonstration et celle de la proposition 3.2: ici, les ordinaux limites sont traités sytématiquement dans le corps de la démonstration, et on conclut en appliquant simplement le résultat général à l'ordinal θ (qui, du reste, peut être

▷ *Le résultat précédent est très puissant, puisqu'il permet de construire des suites de longueur arbitrairement grande, par exemple des suites indexées par l'ordinal non dénombrable ω_1. Il y a une sorte de petit miracle à ce que l'existence de telles suites soit établie par un argument aussi simple que celle des suites indexées par les entiers, et la démonstration peut même sembler une sorte de tour de passe-passe : tout ce qu'on fait est d'écrire les conditions souhaitées à chaque étape, et le principe d'induction fait le reste et en particulier assure l'existence. En fait, ceci illustre la puissance du principe de séparation : à partir du moment où il y a unicité de l'objet souhaité, écrire une formule ensembliste le spécifiant suffit à en garantir l'existence.*

En distinguant le cas de 0*, des ordinaux successeurs et celui des ordinaux limites, on obtient la forme suivante, souvent utilisée en pratique :* ◁

Corollaire 3.5. *Soit A un ensemble, θ un ordinal, a un élément de A, F une application de $\theta \times A$ dans A, et F^* une application de $\theta \times \bigcup_{\lambda<\theta,\,\lambda\text{limite}} A^\lambda$ dans A. Alors il existe une unique application G de θ dans A vérifiant, pour tout α plus petit que θ et tout λ limite plus petit que θ,*

$$G(0) = a,\ G(\alpha+1) = F(\alpha, G(\alpha)),\ \text{et } G(\lambda) = F^*(\,\lambda, G\restriction_{[0,\lambda[}\,) \text{ pour } \lambda \text{ limite} \tag{3.7}$$

c'est-à-dire encore une unique suite $(x_\alpha)_{\alpha<\theta}$ vérifiant

$$x_0 = a, \quad x_{\alpha+1} = F(\alpha, x_\alpha), \quad \text{et} \quad x_\lambda = F^*(\,\lambda, (x_\alpha)_{\alpha<\lambda}\,) \quad \text{pour } \lambda \text{ limite.} \tag{3.8}$$

Démonstration. Appliquer la proposition 3.4 à la fonction F' définie par $F'(0, \emptyset) := a$, $F'(\alpha+1, g) = F(\alpha, g(\alpha))$ pour $\alpha \geqslant 0$ et g application de $[0, \alpha+1[$ dans A, $F'(\lambda, g) = F^*(\lambda, g)$ pour λ limite et g application de $[0, \lambda[$ dans A. □

3.3. Vocabulaire des classes.

► A côté des ensembles, il est commode d'introduire un nouveau type d'objet, *les classes*, pour référer à des collections définies par compréhension mais dont les axiomes ne garantissent pas qu'ils forment des ensembles, ou même montrent qu'ils n'en forment pas. ◄

▷ *Le paradoxe de Russel montre qu'il est impossible d'introduire, pour chaque formule ensembliste* F*, un objet $\{x\,;\,\mathsf{F}(x)\}$ et, dans le même temps, de donner un sens (vrai ou faux) à la relation $X \in Y$ pour chaque choix de X et Y. On a échappé au paradoxe en renonçant à la forme générale des axiomes de compréhension. et en se contentant des axiomes de séparation. Ceci revient à considérer que toutes les formules ensemblistes ne définissent pas un ensemble : par exemple, la formule $x \notin x$, ou encore la formule $\mathbf{Ord}(\alpha)$ qui exprime que α est un ordinal, ne définissent pas d'ensemble.*

Pour autant, et quelle que soit la formule F*, rien n'empêche pas d'introduire un objet $\mathbf{C}_\mathsf{F}$ correspondant à l'usage souhaité pour $\{x\,;\,\mathsf{F}(x)\}$, c'est-à-dire ne dépendant que des x qui satisfont* $\mathsf{F}(x)$*, puis, comme dans le cas des ensembles, d'utiliser $x \in \mathbf{C}_\mathsf{F}$ comme synonyme au fait que x satisfait* F *et donc aussi toute autre formule* F *équivalente à* F*. Ces objets sont utilisés en théorie des ensembles sous le nom de classes.*

Il s'agit donc d'introduire, à côté du type « ensemble » **Ens***, un nouveau type « classe »* **Cl***, ainsi qu'une relation $\in$ entre objets de type* **Ens** *et objets de type* **Cl***, dont l'utilisation est régie par un axiome d'extensionnalité*

$$\forall \mathbf{C}, \mathbf{C}'{:}\mathbf{Cl}\,(\mathbf{C} = \mathbf{C}' \Leftrightarrow \forall x{:}\mathbf{Ens}\,(x \in \mathbf{C} \Leftrightarrow x \in \mathbf{C}'))$$

---

limite ou successeur), alors que, dans le cas des entiers, faute d'introduire l'ordinal ω, la construction de la fonction finale G apparaissait, au moins typographiquement, comme de nature différente à celle des fonctions g_p.

et des axiomes de compréhension associés à chaque formule ensembliste F

$$\forall a_1, \ldots, a_{\mathsf{n}}{:}\mathbf{Ens}\ \exists \mathbf{C}{:}\mathbf{Cl}\ \forall x{:}\mathbf{Ens}\,(x \in \mathbf{C} \Leftrightarrow \mathsf{F}(x, a_1, \ldots, a_{\mathsf{n}})).$$

Pour une définition formelle, les propriétés précédentes sont garanties si on pose : ◁

Définition 3.6. (classe) Deux formules ensemblistes F, F' à une variable libre et d'éventuels paramètres sont dites *co-extensionnelles* si le système $\mathsf{ZF}_{\bullet}$ prouve que les ensembles a vérifiant $\mathsf{F}(a)$ et $\mathsf{F}'(a)$ coïncident. On appelle *classe* associée à F sa classe d'équivalence, notée $\{x\,;\,\mathsf{F}(x)\}$, vis-à-vis de la relation de co-extensionnalité. Si $\mathbf{C}$ est la classe associée à F, on note $a \in \mathbf{C}$ pour $\mathsf{F}(a)$.

Exemple 3.7. (classe) La classe $\{x\,;\,x = x\}$ est la classe de tous les ensembles ; la classe $\{x\,;\,x \notin x\}$ est la classe de tous les ensembles qui ne sont pas éléments d'eux-mêmes, *etc.*

Lemme 3.8. *Pour tout ensemble a, notons $\mathbf{C}_a$ la classe $\{x\,;\,x \in a\}$.*

(i) La correspondance associant à tout ensemble a la classe $\mathbf{C}_a$ est injective.

(ii) Pour tout ensemble b, il y a équivalence entre $b \in a$ et $b \in \mathbf{C}_a$.

(iii) Si une formule ensembliste F définit un ensemble pur a par compréhension, c'est-à-dire s'il existe a vérifiant $\forall x(x \in a \Leftrightarrow \mathsf{F}(x))$, alors la classe $\{x\,;\,\mathsf{F}(x)\}$ coïncide avec la classe $\mathbf{C}_a$.

Démonstration. (i) On a $\mathbf{C}_a = \mathbf{C}_b$ si et seulement si les mêmes ensembles x satisfont $x \in a$ et $x \in b$, donc, par l'axiome d'estensionnalité, si et seulement si a et b sont égaux.

(ii) Par définition, $b \in \mathbf{C}_a$ est satisfait si et seulement si b satisfait la formule $x \in a$, c'est-à-dire si $b \in a$ est vrai.

(iii) Supposons $\forall x(x \in a \Leftrightarrow \mathsf{F}(x))$. Alors, par définition, les formules $x \in a$ et $\mathsf{F}(x)$ sont co-extensionnelles, et donc définissent la même classe. ☐

▷ *Le lemme 3.8 rend cohérente l'option, a priori étonnante, d'utiliser les mêmes notations pour les définitions en compréhension de classes et d'ensembles, ainsi que pour l'appartenance « ensemble–ensemble » et l'appartenance « ensemble–classe », et il montre qu'il est sans danger d'identifier tout ensemble a et la classe $\mathbf{C}_a$ correspondante, c'est-à-dire à considérer les ensembles comme des classes particulières — et c'est ce qu'on fera désormais. Ce qui, dans ce contexte, évite la réapparition du paradoxe de Russel est le fait qu'on ne considère pas d'ensemble de classes: si $\mathbf{C}$ est une classe et a un ensemble (ou une classe), la relation $\mathbf{C} \in a$ n'est ni vraie, ni fausse, elle n'a pas de sens — pas davantage que n'a de sens $\mathsf{F} \in a$ avec F formule et a ensemble.*

Par contre, on notera que l'inclusion entre classes fait sens : pour $\mathbf{C} = \{x\,;\,\mathsf{F}(x)\}$ et $\mathbf{C}' = \{x\,;\,\mathsf{F}'(x)\}$, on peut sans danger écrire $\mathbf{C} \subseteq \mathbf{C}'$ pour indiquer que tout ensemble qui est dans $\mathbf{C}$ est dans $\mathbf{C}'$, c'est-à-dire que $\mathsf{F}(x)$ entraîne $\mathsf{F}'(x)$.

A partir de ce point, deux options sont possibles pour développer la théorie des ensembles. Ou bien on considère délibérement deux types d'objets, les ensembles et les classes, pour lesquels on pose une base axiomatique et qu'on étudie en parallèle. C'est notamment le point de vue adopté dans les théories de Gödel–Bernays et de Kelley–Morse. Ou bien — et c'est le point de vue qu'on adoptera ici — les ensembles restent le seul type étudié pour lui-même, et l'introduction des classes n'est qu'une facilité de rédaction permettant d'écrire par exemple $\alpha \in \mathbf{Ord}$ plutôt que $\mathbf{Ord}(\alpha)$ pour exprimer que α est un ordinal, et, de même, de pouvoir référer globalement aux ordinaux comme « la classe des ordinaux », puisque ceux-ci ne forment pas un ensemble. En fait, les deux points de vue sont essentiellement équivalents, et on montre en particulier que toute propriété des ensembles démontrable à partir du système de Gödel–Bernays (et donc en utilisant des classes) l'est aussi à partir du système de Zermelo–Fraenkel (donc sans utiliser de classes).

Dans toute la suite, on s'en tient à la convention que, par défaut, les variables réfèrent à des ensembles, les classes n'intervenant que comme variables non quantifiées et en général représentées par des caractères gras tels que $\mathbf{C}, \mathbf{F}$*, etc. Noter qu'introduire la classe* $\mathbf{C}$ *associée à une formule* F *revient à ajouter* $\mathbf{C}$ *comme nouvelle relation unaire à la signature, accompagnée de l'axiome* $\mathsf{Intro}_{\mathbf{C}} : x \in \mathbf{C} \Leftrightarrow \mathsf{F}(x)$*. On a observé qu'il est alors légitime d'utiliser les relations ainsi introduites dans des axiomes de séparation ou de remplacement. Il en est donc de même avec les classes : une formule où figurent comme paramètre des classes n'est pas une formule ensembliste au sens strict, mais on obtient une formule ensembliste équivalente en remplaçant chaque classe par une formule la définissant. Noter qu'il est également légitime d'utiliser des classes comme paramètres dans la définition de nouvelles classes : ainsi, on peut utiliser la classe des ordinaux* $\mathbf{Ord}$ *comme paramètre pour définir des sous-classes et par exemple définir la classe des ordinaux limites par*

$$\mathbf{Lim} := \{\lambda \in \mathbf{Ord}\,;\, \lambda = \bigcup \lambda\}.$$

Enfin, on remarquera que les axiomes de séparation prennent une forme spécialement simple lorsqu'on les énonce en termes de classe: en effet, séparer dans un ensemble a *les* x *qui vérifient* $\mathsf{F}(x)$*, c'est prendre l'intersection de* a *et de la classe associée à* F*; les axiomes de séparation peuvent donc s'énoncer sous la forme:*
« L'intersection d'un ensemble et d'une classe est un ensemble ». ◁

3.4. Récursion ordinale généralisée.

▶ Le vocabulaire des classes permet de parler de suites indexées par les ordinaux, et on peut alors étendre la construction par définition récursive à de telles suites. ◀

▷ *De même qu'il est commode de pouvoir parler de la classe de tous les ordinaux, il est commode de pouvoir parler de suites indexées par (tous) les ordinaux, c'est-à-dire de suites* $(x_\alpha)_{\alpha \in \mathbf{Ord}}$*, dont les indices ne sont pas limités aux ordinaux plus petits qu'un certain ordinal* θ*, mais parcourent la suite complète des ordinaux.*

Moyennant l'identification d'une fonction avec un ensemble de couples, ceci peut se faire facilement : il s'agit simplement de considérer des classes dont les éléments sont des couples. Il est néanmoins commode d'adopter un vocabulaire spécifique. ◁

Définition 3.9. (classe fonctionnelle) On appelle *fonctionnelle* une classe $\mathbf{F}$ dont les éléments sont des couples et qui est telle que, pour tout x, il existe au plus un y tel que (x, y) appartienne à $\mathbf{F}$. On note alors $y = \mathbf{F}(x)$ pour $(x, y) \in \mathbf{F}$.

Exemple 3.10. (classe fonctionnelle) Soit $\mathsf{F}(z)$ la formule $\exists x, y(z = (x, y)$ et $y = \mathfrak{P}(x))$. Alors $\mathsf{F}(z)$ définit[14] une classe fonctionnelle $\mathbf{F}$. Comme $y = \mathbf{F}(x)$ équivaut à $y = \mathfrak{P}(x)$, on peut identifier $\mathbf{F}$ à $\mathfrak{P}$ et parler de la (classe) fonctionnelle $\mathfrak{P}$.

▷ *D'une façon générale, introduire une classe fonctionnelle revient à introduire une nouvelle opération définissable dans la signature, et les possibilités d'utilisation dans des définitions ultérieures sont les mêmes qu'avec les classes. Noter que, dans le contexte des classes fonctionnelles, les axiomes de remplacement peuvent se reformuler en :*

[14] moyennant la définition de $\mathfrak{P}$ et en supposant l'axiome d'extensionnalité vérifié : le caractère fonctionnel est relatif à un cadre axiomatique donné, ici le système $\mathsf{ZF}_\bullet$ ou une extension de celui-ci

« Si le domaine d'une classe fonctionnelle est un ensemble, alors son image en est un également. »

Un cas particulier de classe fonctionnelle est celui de suite indexée par les ordinaux, c'est-à-dire une classe fonctionnelle de la forme $(u_\alpha)_{\alpha\in\mathbf{Ord}}$ dont le domaine est la classe des ordinaux. Le cadre étant ainsi défini, on peut revenir sur les définitions récursives et étendre le résultat pour englober le cas de suites indexées par les ordinaux, et non plus par un ordinal, c'est-à-dire un segment initial de la classe des ordinaux. $\triangleleft$

Proposition 3.11. (récursion ordinale généralisée) ($\mathsf{ZF}_\bullet$) *Soit* $\mathbf{A}$ *une classe, et* $\mathbf{F}$ *une classe fonctionnelle à valeurs dans* $\mathbf{A}$ *dont le domaine inclut* $\{\alpha\}\times\mathbf{A}^\alpha$ *pour tout ordinal* α. *Alors il existe une unique classe fonctionnelle* $\mathbf{G}$ *définie sur* $\mathbf{Ord}$ *et à valeurs dans* $\mathbf{A}$ *vérifiant, pour tout ordinal* α,

$$\mathbf{G}(\alpha) = \mathbf{F}(\,\alpha, \mathbf{G}{\restriction}_{[0,\alpha[}\,), \tag{3.9}$$

c'est-à-dire encore une unique suite $(x_\alpha)_{\alpha\in\mathbf{Ord}}$ *vérifiant*

$$x_\alpha = \mathbf{F}(\,\alpha, (x_\beta)_{\beta<\alpha}\,). \tag{3.10}$$

Si $\mathbf{A}$ *est un ensemble, la conclusion s'obtient à partir du système* Z.

Démonstration. L'argument est le même que pour la proposition 3.4. Comme tous les ensembles introduits y étaient définis par des formules ensemblistes explicites, on obtient ici des contreparties sous forme de classes. Supposons que la classe $\mathbf{A}$ est définie par la formule $\mathsf{A}(x,\vec{c})$, et, de même, que $x=\mathbf{F}(\alpha,g)$ est défini par une formule ensembliste $\mathsf{F}(\alpha,g,x,\vec{c})$ [15]. Comme précédemment, pour chaque ordinal β, on appelle β-solution toute fonction g de domaine $[0,\beta[$ vérifiant la condition $g(\alpha)=\mathbf{F}(\alpha, g{\restriction}_{[0,\alpha[})$ pour tout α dans son domaine.

Observons d'abord que les solutions forment une classe, qu'on notera $\mathbf{S}$. En effet, on a

$$\mathbf{S} = \{g : \mathbf{Ord}\to\mathbf{A}\,;\, \exists\beta(\mathrm{Dom}(t)=[0,\beta[\text{ et } \forall\alpha{<}\,\beta(g(\alpha)=\mathbf{F}(\alpha,g{\restriction}_{[0,\alpha[})))\},$$

soit, en remplaçant les classes $\mathbf{A}$ et $\mathbf{F}$ par des formules les définissant

$$\begin{aligned}\mathbf{S} = \{g\,;\, \exists\beta(&\text{« } g \text{ est une fonction »} \text{ et } \mathrm{Dom}(g)=[0,\beta[\\ &\text{et } \forall x{\in}\mathrm{Im}(g)(\mathsf{A}(x,\vec{c})) \text{ et } \forall\alpha{<}\,\beta(\mathsf{F}(\alpha,g{\restriction}_{[0,\alpha[},g(\alpha),\vec{c}))\},\end{aligned}$$

qui est explicitement ensembliste dès lors que A et F le sont, et donc définit bien une classe. Comme précédemment, on montre par induction sur β que, pour tout ordinal β, il existe une unique β-solution g_β. Pour $\beta=0$, l'unique possibilité est $g_\beta=\emptyset$. Pour $\beta=\gamma+1$, on a $g_\beta = g_\gamma\cup\{(\gamma,a)\}$, où a est l'unique élément vérifiant $\mathsf{A}(a,\vec{c})$ et $\mathsf{F}(\alpha,g_\gamma,a,\vec{c})$, c'est-à-dire $a\in\mathbf{A}$ et $a=\mathbf{F}(\alpha,g_\gamma)$. Pour β limite, l'hypothèse d'induction garantit que, pour chaque $\gamma<\beta$, il existe une unique γ-solution g_γ. C'est dire que la formule à deux variables γ et g

$$\gamma<\beta \text{ et } g\in\mathbf{S} \text{ et } \mathrm{Dom}(g)=[0,\gamma[$$

est fonctionnelle en γ. Un axiome de remplacement garantit alors que l'image de la correspondance ainsi définie, qui est $\{g_\gamma\,;\,\gamma<\beta\}$, est un ensemble. Il est donc légitime de poser $g_\beta := \bigcup\{g_\gamma\,;\,\gamma<\beta\}$, et on vérifie comme dans la démonstration de la proposition 3.4 que g_β est la seule β-solution.

Soit alors $\mathbf{G}$ la classe fonctionnelle telle que $y=\mathbf{G}(\beta)$ est définie par

$$\exists g{\in}\mathbf{S}(\mathrm{Dom}(g)=[0,\beta+1[\text{ et } y=g(\beta)),$$

soit, sans utiliser de classe,

[15] Noter que, quitte à prendre une réunion, on peut toujours supposer que les éventuels paramètres figurant dans les définitions des classes $\mathbf{A}$ et $\mathbf{F}$ sont les mêmes

$$\exists g(\text{ « } g \text{ est une fonction » et } \mathrm{Dom}(g) = [0, \beta+1[$$
$$\text{et } \forall x{\in}\mathrm{Im}(g)(\mathsf{A}(x, \vec{c})) \text{ et } \forall\alpha{<}\ \beta+1(\mathsf{F}(\alpha, g\restriction_{[0,\alpha[}, g(\alpha), \vec{c}) \text{ et } y = g(\beta)).$$

Dans ce qui précède, la seule utilisation d'un axiome de remplacement est pour justifier l'existence de l'ensemble $\{g \in \mathbf{S}\,;\, \exists\gamma{<}\ \beta(\mathrm{Dom}(g) = [0,\gamma[\ )\}$. Si $\mathbf{A}$ est un ensemble, ce dernier ensemble est un sous-ensemble de l'ensemble des fonctions de β dans $\mathbf{A}$, et il suffit d'invoquer un axiome de séparation pour en justifier l'existence. Dans ce cas, la démonstration n'utilise que les axiomes de Z. □

Si $\mathbf{G}$ est une classe fonctionnelle, alors, par remplacement, toute restriction de $\mathbf{G}$ à un ensemble A est un ensemble[16]. En particulier, toute restriction d'une suite définie par récursion ordinale généralisée à un segment initial $[0, \theta[$ de la suite des ordinaux est un ensemble, et on retrouve le résultat de la proposition 3.4.

Pour terminer, on reformule le résultat précédent en séparant ordinaux successeurs et limites :

Corollaire 3.12. ($\mathsf{ZF}_\bullet$) *Soit* $\mathbf{A}$ *une classe,* a *un élément de* $\mathbf{A}$*, et* $\mathbf{F}, \mathbf{F}^*$ *deux classes fonctionnelles à valeurs dans* $\mathbf{A}$ *dont les domaines incluent respectivement* $\mathbf{Ord} \times \mathbf{A}$ *et* $\{\lambda\} \times \mathbf{A}^\lambda$ *pour tout ordinal limite* λ*. Alors il existe une unique classe fonctionnelle* $\mathbf{G}$ *définie sur* $\mathbf{Ord}$ *et à valeurs dans* $\mathbf{A}$ *vérifiant pour tous* α, λ

$$\mathbf{G}(0) = a,\ \mathbf{G}(\alpha+1) = \mathbf{F}(\alpha, \mathbf{G}(\alpha)),\ \text{et } \mathbf{G}(\lambda) = \mathbf{F}^*(\,\lambda, \mathbf{G}\restriction_{[0,\lambda[}\,) \text{ pour } \lambda \text{ limite,}$$

c'est-à-dire encore une unique suite $(x_\alpha)_{\alpha\in\mathbf{Ord}}$ *vérifiant pour tous* α, λ

$$x_0 = a, \quad x_{\alpha+1} = \mathbf{F}(\alpha, x_\alpha), \quad \text{et} \quad x_\lambda = \mathbf{F}^*(\,\lambda, (x_\alpha)_{\alpha<\lambda}\,) \quad \text{pour } \lambda \text{ limite.}$$

Si $\mathbf{A}$ *est un ensemble, la conclusion s'obtient à partir du système* Z.

3.5. La suite des ensembles V_α.

► On légitime l'existence de la suite des ensembles V_α utilisés dans la section 1. ◄

▷ *Tous les résultats du chapitre II ainsi que de la section 1 du présent chapitre ont été justifiés sur la base des axiomes de* $\mathsf{ZF}_\bullet$*, à l'exception de l'existence de la suite des ensembles* V_α *mis en jeu dans la définition des ensembles purs. Avec les résultats précédents, il est maintenant immédiat d'achever la tâche.* ◁

Proposition 3.13. (suite des V_α) ($\mathsf{ZF}_\bullet$) (*i*) *Il existe une unique suite d'ensembles* $(V_\alpha)_{\alpha\in\mathbf{Ord}}$ *indexée par les ordinaux et vérifiant pour tous* α, λ

$$V_0 = \emptyset, \quad V_{\alpha+1} = \mathfrak{P}(V_\alpha), \quad V_\lambda = \bigcup_{\alpha<\lambda} V_\alpha \text{ pour } \lambda \text{ limite.} \tag{3.11}$$

(*ii*) *Les ensembles purs forment une classe.*

[16] si l'image de $\mathbf{F}\restriction_A$ est *a priori* connue comme incluse dans un ensemble B, alors il suffit d'un axiome de séparation à l'intérieur de l'ensemble $\mathfrak{P}(A \times B)$

Démonstration. Le point (i) est une application directe du corollaire 3.12, les deux classes fonctionnelles mises en jeu correspondant aux opérations $\mathfrak{P}$ et $\bigcup$.

Le point (ii) en résulte, puisque « x est pur » peut être défini par la formule ensembliste $\exists\beta(x \in V_\beta)$, qu'on peut si on préfère expliciter par exemple en

$$\begin{aligned}\exists\beta, g\,(\,&\mathbf{Ord}(\beta)\ \text{ et }\ \text{« } g \text{ est une fonction »}\ \text{ et }\ \mathrm{Dom}(g) = [0, \beta+1[\ \text{ et }\ g(0) = \emptyset\\ &\text{et }\ \forall\alpha{<}\,\beta(g(\alpha+1) = \mathfrak{P}(g(\alpha)))\ \text{ et }\ \forall\lambda\,\text{limite}{<}\,\beta\,(g(\lambda) = \textstyle\bigcup(g{\restriction}_{[0,\lambda[}))\\ &\text{et }\ x \in g(\beta)\ ).\end{aligned}$$

□

▷ *Le résultat précédent signifie non seulement que, pour chaque ordinal α, le système* $\mathsf{ZF}_\bullet$ *garantit l'existence de l'ensemble V_α, mais plus généralement que $x = V_\alpha$ est une formule ensembliste en les deux variables α et x, de surcroît fonctionnelle en α. On peut donc légitimement utiliser cette formule dans toute définition par séparation ou remplacement : il existe une formule* uniforme *globale définissant la suite des V_α. En d'autres termes, la suite des ensembles V_α est une classe. Par ailleurs, pour tout ordinal θ, la suite $(V_\alpha)_{\alpha<\theta}$ est un ensemble.* ◁

Notation 3.14. (ensembles purs) On note $\mathbf{V}$ la classe des ensembles purs [17].

L'opération « rang » est une classe fonctionnelle définie sur la classe $\mathbf{V}$, puisque $\mathrm{rang}(a) = \alpha$ équivaut à $a \in V_{\alpha+1} \setminus V_\alpha$.

▷ *La fin des vérifications de la section 1 ne pose alors pas de problème, au sens où les axiomes de* $\mathsf{ZF}_\bullet$ *permettent de légitimer l'existence de tous les ensembles introduits et de démontrer tous les résultats annoncés, en particulier les résultats de clôture de la classe* $\mathbf{V}$ *par diverses opérations ensemblistes, et la proposition 1.17 qui en découle.* ◁

4. La théorie des ensembles

▶ On introduit le système de Zermelo–Fraenkel en ajoutant à $\mathsf{ZF}_\bullet$ l'axiome de fondation qui exprime que tout ensemble est pur, et on observe qu'en vertu de la possibilité de représenter tout objet mathématique par un ensemble pur ce système est une base réaliste pour développer la théorie des ensembles. On discute quelques points concernant la base axiomatique ainsi obtenue, les options qu'elle privilégie, et les approches alternatives. ◀

4.1. Une caractérisation des ensembles purs.

▶ Les ensembles purs ont été définis ci-dessus en termes de la suite des ensembles V_α. On en établit une caractérisation alternative intrinsèque. ◀

On rappelle qu'un ensemble a est dit transitif si les éléments des éléments de a sont des éléments de a. Tout ensemble n'est pas transitif, mais on observe que tout ensemble est inclus dans un ensemble transitif distingué.

Lemme 4.1. *Pour tout ensemble A, il existe un plus petit ensemble transitif $\widetilde{A}$ incluant A ; de plus, $\widetilde{A}$ est pur si et seulement si A l'est.*

[17] La notation simple V sera aussi utilisée plus loin ; pour le moment, on insiste sur le fait que $\mathbf{V}$ n'est pas un ensemble.

Démonstration. Par la proposition 3.2, il existe une suite $(A_n)_{n\in\omega}$ définie récursivement par $A_0 = A$ et $A_{n+1} = \bigcup A_n$. Posons $\widetilde{A} = \bigcup_{n\in\omega} A_n$. Supposons $x \in y \in \widetilde{A}$. Il existe n tel que y appartient à A_n, et alors x appartient à A_{n+1}, donc à $\widetilde{A}$. Par conséquent $\widetilde{A}$ est transitif. Inversement, supposons que B est transitif et inclut A. Alors une induction sur n montre $A_n \subseteq B$ pour tout n dans ω, d'où $\widetilde{A} \subseteq B$.

Supposons A pur. Par la proposition 1.6, les ensembles purs sont clos par l'opération $\bigcup$, donc, inductivement, chacun des ensembles A_n est pur, puis leur union l'est [18]. Inversement, si $\widetilde{A}$ est pur, il en est de même de A, puisque A est inclus dans $\widetilde{A}$. □

Définition 4.2. (clôture transitive) Le plus petit ensemble transitif incluant A est appelé *clôture transitive* de A et noté $\mathrm{Clot}_\in(A)$.

Par construction, un ensemble est transitif si et seulement si il est égal à sa clôture transitive. Noter la formule

$$\mathrm{Clot}_\in(A) = A \cup \bigcup\{\mathrm{Clot}_\in(a)\,;\, a \in A\},$$

qui résulte du fait que $a \in A$ entraîne $\mathrm{Clot}_\in(a) \subseteq \mathrm{Clot}_\in(A)$.

On rappelle (Définition II.1.1) qu'une relation binaire R sur un ensemble A est dite *bien fondée* si toute partie non vide X de A contient un élément R-minimal, c'est-à-dire un élément x tel qu'aucun élément y de X ne vérifie yRx.

Proposition 4.3. (ensemble pur) *Un ensemble A est pur si et seulement si la restriction de $\in$ à $\mathrm{Clot}_\in(A)$ est une relation bien fondée.*

Démonstration. Supposons A pur. Alors, par le lemme 4.1, $\mathrm{Clot}_\in(A)$ est également pur. Soit X une partie non vide de $\mathrm{Clot}_\in(A)$. L'ensemble des rangs des éléments de X est une partie non vide de **Ord**, donc il possède un plus petit élément. Il existe donc dans X un élément a de rang minimal. Alors a est $\in$-minimal dans X, car $x \in a$ entraîne $\mathrm{rang}(x) < \mathrm{rang}(a)$, donc $x \notin X$. Donc la relation $\in\restriction_{\mathrm{Clot}_\in(A)}$ est bien fondée.

Inversement, supposons $\in\restriction_{\mathrm{Clot}_\in(A)}$ bien fondée. Si $\mathrm{Clot}_\in(A)$ est vide, alors A est l'ensemble vide, donc est pur. Sinon, posons $X = \{x \in \mathrm{Clot}_\in(A)\,;\, x \notin \mathbf{V}\}$, et supposons X non vide. Alors, par hypothèse, il existe un élément $\in$-minimal a dans X. Par construction, tous les éléments de a sont dans $\mathrm{Clot}_\in(A)$, qui est transitif, mais pas dans X, puisque a est $\in$-minimal dans X. C'est dire que tous les éléments de a sont dans $\mathbf{V}$. Par le corollaire 1.7, cela implique que a lui-même est dans $\mathbf{V}$, contrairement à l'hypothèse. C'est donc que X est vide, c'est-à-dire que $\mathrm{Clot}_\in(A)$ est inclus dans $\mathbf{V}$, et donc, par le corollaire 1.7 à nouveau, $\mathrm{Clot}_\in(A)$ est dans $\mathbf{V}$, et par conséquent A aussi. □

Corollaire 4.4. *En présence des axiomes de $\mathsf{ZF}_\bullet$, l'assertion « tout ensemble est pur » équivaut à « tout ensemble non vide possède un élément $\in$-minimal », c'est-à-dire à la condition*

$$\forall a(a \neq \emptyset \Rightarrow \exists b{\in}a(b \cap a = \emptyset)). \tag{4.1}$$

Démonstration. D'abord, dire qu'un élément b de a est $\in$-minimal dans a signifie qu'aucun x vérifiant $x{\in}b$ n'appartient à a, donc que l'intersection de b et a est vide.

Supposons alors (4.1), et soit a un ensemble quelconque. Par (4.1), la restriction de $\in$ à $\mathrm{Clot}_\in(a)$ est bien fondée, et donc, par la proposition 4.3, l'ensemble a est pur.

[18] Noter que, plus précisément, si A est dans V_α, alors chaque A_n est dans V_α, et $\widetilde{A}$ également

Inversement, soit a pur quelconque. Alors, par la proposition 4.3, la restriction de $\in$ à $\mathrm{Clot}_\in(a)$ est bien fondée, et, *a fortiori*, la restriction de $\in$ à a l'est et, en particulier, a lui-même possède un élément $\in$-minimal s'il est non vide. □

4.2. Le système de Zermelo–Fraenkel.

► On définit (enfin!) le système de Zermelo–Faenkel ZF, en ajoutant aux axiomes de $\mathsf{ZF}_\bullet$ l'axiome de fondation qui exprime que tout ensemble est pur. ◄

▷ *On a montré dans la section 1 comment représenter tous les objets mathématiques usuels, donc en particulier tous les ensembles, par des ensembles purs. Quitte à remplacer les objets par leur représentation, on ne limite donc pas le champ de la théorie des ensembles en limitant celle-ci à l'étude des ensembles purs : par exemple, tout résultat sur le problème du continu mettant en jeu les copies* $\underline{\mathbb{N}}$ *et* $\underline{\mathbb{R}}$ *induit automatiquement un résultat semblable pour les originaux* $\mathbb{N}$ *et* $\mathbb{R}$. *Il est dès lors naturel de se placer dans le monde des ensembles purs, ce qui revient à ajouter au système des axiomes de base l'axiome « tout ensemble est pur ».*

Le statut de cet axiome est différent de celui des autres axiomes de $\mathsf{ZF}_\bullet$ *: l'adoption de ces derniers a été justifiée par le fait que l'intuition que nous avons des ensembles et de leur utilisation rend cette adoption naturelle et utile. Aucune intuition ne vient recommander l'axiome « tout ensemble est pur », et, au contraire, il semble naturel de considérer que les entiers naturels ne sont pas des ensembles et donc qu'un ensemble tel que* $\mathbb{N}$ *n'est pas pur. L'axiome « tout ensemble est pur » n'est donc pas ajouté parce qu'il est considéré comme intuitivement vérifié, mais simplement comme une hypothèse technique indiquant qu'on restreint l'étude des ensembles à celle des ensembles purs, hypothèse justifiée par le fait qu'on n'introduit ainsi aucune limitation sérieuse à la portée de la théorie et par le bénéfice technique apportée par la condition de pureté, à savoir la possibilité d'utiliser le rang comme paramètre d'induction dans les démonstrations.*

D'après le corollaire 4.4, affirmer que tout ensemble est pur équivaut à poser la condition (4.1)*, et c'est celle-ci qu'il est d'usage d'introduire comme axiome.* ◁

Définition 4.5. (axiome de fondation, système ZF) On appelle axiome de *fondation* la condition (4.1), c'est-à-dire

$$\forall a(a \neq \emptyset \Rightarrow \exists b{\in}a(b \cap a = \emptyset)). \tag{AF}$$

On appelle *système de Zermelo–Fraenkel*, ou ZF, le système obtenu en ajoutant AF à $\mathsf{ZF}_\bullet$, c'est-à-dire la liste (infinie) des axiomes suivants : extensionnalité, paire, union, parties, infini, fondation, ainsi que tous les axiomes de séparation et de remplacement.

▷ *La règle du jeu est maintenant claire : les axiomes du système* ZF *étant pris comme point de départ, il s'agit d'en explorer les conséquences, et, en particulier, de chercher s'ils résolvent ou au moins éclairent les divers problèmes ouverts mentionnés au chapitre I.*

Le système axiomatique ZF *est un outil pour explorer le monde des ensembles. Un consensus s'accordant à considérer les axiomes de* ZF *comme intuitivement vrais (ou, dans le cas de l'axiome de fondation, comme bénin et techniquement utile), on pourra tenir pour établie toute propriété déduite des axiomes de* ZF [19]*. Les problèmes les plus urgents et les plus naturels à aborder sont ceux qu'on a mentionnés au chapitre I, par exemple le problème du continu sur l'existence d'ensembles de taille intermédiaire entre celles de* $\mathbb{N}$ *et* $\mathbb{R}$*. Bien sûr, ce qu'étudiera la théorie des ensembles est l'existence d'ensembles purs de taille intermédiaire entre celles des ensembles purs* $\underline{\mathbb{N}}$ *et* $\underline{\mathbb{R}}$*, mais, comme on l'a dit, dans la mesure où on tient la copie des objets divers par des ensembles purs pour fidèle, on pourra considérer que toute réponse mettant en*

[19] pour autant qu'il y ait également consensus sur la validité de la notion de déduction utilisée

jeu les ensembles purs $\underline{\mathbb{N}}$ *et* $\underline{\mathbb{R}}$ *est un argument probant pour le résultat analogue mettant en jeu les véritables ensembles* $\mathbb{N}$ *et* $\mathbb{R}$ *— quels que soient ces derniers.*

A ce point, rien n'indique que les axiomes de ZF *épuisent notre intuition de la notion d'ensemble. Il n'y aura donc rien d'étonnant à ce que certaines propriétés ne puissent être ni démontrées ni réfutées à partir de* ZF*. Dans un tel cas, et comme on l'a déjà fait en introduisant l'axiome de l'infini ou les axiomes de remplacement, il s'agira à chaque fois de s'interroger sur l'opportunité d'ajouter de nouveaux axiomes, et certainement pas de tenir le problème pour indécidable en un sens plein de mystère.* ◁

4.3. La théorie des ensembles comme base de l'édifice mathématique.

▶ Du fait de la possibilité de représenter tous les objets comme ensembles purs, le système ZF devient un cadre global dans lequel fonder l'intégralité de l'édifice mathématique. ◀

▷ *L'intérêt de la représentation des objets mathématiques quelconques par des ensembles purs est considérable. Au départ, les mathématiques apparaissent comme un domaine diffus où les investigations portent sur une multiplicité d'objets de natures apparemment très diverses. La possibilité de représenter tous les objets mathématiques usuels par des ensembles permet d'unifier ce cadre et, au moins en théorie, de ramener les mathématiques à l'étude d'un seul type d'objet, les ensembles.*

Un des bénéfices de cette unification concerne les questions de fondement des mathématiques. L'objectif est de décrire un système formel à l'intérieur duquel toutes les constructions puissent être légitimées et qu'on espère à l'abri des contradictions. Au départ, la multiplicité des types d'objets semble rendre le problème insoluble. La possibilité de représenter tous les objets mathématiques par des ensembles (purs) ramène le problème général au problème, apparemment plus simple, du fondement de la seule théorie des ensembles. Comme, pour cette dernière, on vient de proposer un cadre axiomatique adapté, à savoir le système ZF*, le problème de fondement des mathématiques se trouve ramené au seul problème de l'absence de contradiction dans le système* ZF*. On peut donc énoncer le résultat suivant (nécessairement informel) :* ◁

Proposition 4.6. (non contradiction) *Si le système* ZF *est non contradictoire, il en est de même de l'intégralité de l'édifice mathématique usuel.*

▷ *Au passage, on obtient un résultat analogue pour les systèmes axiomatiques proposés pour les divers types d'objets réprésentés par des ensembles et dont les axiomes sont prouvables à partir de ceux de* Z *ou de* ZF*.* ◁

Proposition 4.7. (non contradiction) *Si le système* Z *est non contradictoire, donc a fortiori si le système* ZF *est non contradictoire, il en est de même du système de Peano* PA*.*

▷ *En fait, pour ce qui est de la cohérence de l'édifice mathématique, et contrairement aux espoirs naïfs de Hilbert il y a un siècle, il n'est pas certain que la représentation des objets mathématiques comme ensembles soit un gain décisif. Au vu des résultats précédents, le problème fondamental devient* ◁

Question 4.8. *Le système* ZF *est-il non contradictoire?*

▷ *Or, Kurt Gödel a montré au début des années 1930 des résultats négatifs limitant drastiquement les espoirs dans cette direction. Comme on le verra au chapitre VIII, le (second) théorème d'incomplétude affirme que l'absence de contradiction dans le système* ZF *ne peut pas être établie à l'intérieur de ce système (sauf s'il est contradictoire). Il est donc possible d'adopter* ZF *comme point de départ unifié de la théorie des ensembles et des mathématiques, mais il ne saurait être exclu que le système recèle une contradiction. L'argument de Gödel n'illustre pas une faiblesse* *spécifique du système* ZF*: la même limitation vaut pour tout système suffisamment puissant pour que l'arithmétique puisse y être représentée, et il serait donc illusoire d'espérer y échapper en ajoutant de nouveaux axiomes ou en modifiant les axiomes existants. Aussi longtemps qu'on accepte le cadre métamathématique des démonstrations dans la logique dite du premier ordre, il ne saurait y avoir de démonstration de cohérence intrinsèque du monde mathématique. Les limitations précédentes relativisent donc l'intérêt théorique de la représentation par des ensembles en termes de fondements.* ◁

4.4. Le point de vue « tout est ensemble ».

► L'introduction des ensembles purs permet de développer une théorie des ensembles en se plaçant dans un monde exclusivement constitué d'ensembles, et même d'ensembles purs. Par le biais de la représentation des autres objets par des ensembles purs, les résultats de la théorie obtenue ont une valeur universelle. Mais ce n'est pas une raison pour confondre comportement et essence, ni pour extrapoler le fait que les ensembles fournissent des contreparties aux autres types d'objet en la conclusion hasardeuse que toute objet est un ensemble. ◄

▷ *On a montré plus haut que la plupart des objets mathématiques peuvent être adéquatement représentés par des ensembles purs. Par exemple, chaque entier naturel* n *se trouve représenté par un certain ensemble pur, l'ordinal fini* $\underline{\mathsf{n}}$. *Pour des raisons de lisibilité, on a proposé de noter x pour $\underline{x}$, écrivant par exemple simplement 2 à la place de $\underline{2}$, ce qui ne crée pas d'ambiguïté puisqu'on se place résolument dans un monde d'ensembles et que, donc, on ne considère en fait que les ensembles-représentants, et pratiquement jamais les objets originaux — à l'exception des formules et des entiers intuitifs pour lesquels, précisément, on conserve une notation spécifique. Ceci semble proche d'une* identification *de x et $\underline{x}$. Il y a là en fait deux approches possibles qui, même si elles mènent aux mêmes développements techniques, procèdent de conceptions assez éloignées.*

Ou bien on identifie résolument x et $\underline{x}$ pour tout x, considérant que l'objet x est *l'ensemble $\underline{x}$, c'est-à-dire prenant la contruction de $\underline{x}$ comme une définition de x. Dans cette approche, où tous les objets mathématiques* sont *des ensembles, les entiers naturels sont définis comme étant les ordinaux finis, les fonctions sont identifiées à leur graphe, etc. Par exemple, l'entier 2 est défini comme étant la paire $\{0, 1\}$, puisqu'il se trouve que, dans la construction de von Neumann des ordinaux, on a $\underline{2} = \{\underline{0}, \underline{1}\}$. Même si la beauté formelle de la construction a pu la rendre séduisante aux yeux du jeune Bourbaki, il ne semble pas que l'égalité $2 = \{0, 1\}$ soit une évidence intuitive que les humains partageraient. La poser comme axiome relève alors d'une approche dogmatique difficilement justifiable.*

De façon plus profonde, l'option « tout est ensemble » nécessite que tous les objets soient des ensembles, y compris les objets métamathématiques comme les formules et les entiers qui y figurent[20] *: dans l'approche du traité de Bourbaki, les formules, les entiers, les ensembles vivent tous dans le même monde, dont on n'a qu'une vue interne puisqu'il n'y a pas d'extérieur. Cette approche qui confond le niveau des ensembles et celui des formules interdit ou, au moins, rend extrêmement périlleux, le passage au point de vue des modèles de* ZFC *abordé dans la partie C de ce texte, point de vue qui seul a pu permettre les grandes avancées du XXe siècle sur les problèmes ouverts comme le problème du continu. L'option « tout est ensemble » paraît donc peu défendable aujourd'hui car elle induit un biais qui limite intrinsèquement la portée de la théorie.*

Ou bien on continue à faire la distinction entre les objets et leur représentation, entre un extérieur et un intérieur du monde des ensembles. Peut-être moins fascinante car moins directement théorie du grand tout, cette option a le mérite d'être soutenue par l'intuition et d'être proche de la pratique mathématique, tout en autorisant les développements ultérieurs. On a noté qu'une telle approche n'interdit pas de poser des conventions simplificatrices, de sorte par exemple que, du point de vue technique interne de la théorie des ensembles, il est commode de faire comme si *le couple (a, b) était la paire $\{\{a, b\}, \{a\}\}$, comme si une fonction était son graphe,*

[20] sans parler des classes, même considérées comme simples auxiliaires de rédaction

plus généralement comme si tout objet mathématique [21] *était un ensemble (pur). Mais ceci ne signifie pas qu'on oublie qu'il ne s'agit que d'une représentation, et qu'il existe des objets hors du monde des ensembles, typiquement les formules logiques comme celles qui sont explicitement mentionnées dans les axiomes du système* ZF, *et les entiers qui y figurent comme indices et paramètres de longueur, et de là paramètres de récurrence. Ces entiers externes sont ceux qu'on a proposé d'appeler* intuitifs *pour les distinguer des entiers internes qui sont les ordinaux finis éléments de* ω. *C'est cette option qui est aujourd'hui privilégiée, parfois de façon implicite, dans la plupart des ouvrages récents — et ce sera notre approche ici.*

Le lecteur hésitant entre les deux options ci-dessus pourra toujours noter qu'un des points de vue est en quelque sorte inclus dans l'autre, au moins tant qu'on n'aborde pas les développements décrits dans la partie C de ce texte. En effet, une fois posées les conventions simplificatrices, la seule différence entre le point de vue « tout est ensemble » et le point de vue « il y a un à-côté des ensembles » est que le premier sur-ajoute l'hypothèse, techniquement indifférentes, que les représentations sont des définitions. Comme on l'a dit plusieurs fois, la question de la nature des objets mathématiques n'influe guère sur les théorèmes qu'on peut démontrer à leur propos, et le fait que l'entier 2 *soit ou non la paire* $\{0,1\}$ *ne change rien aux théorèmes d'arithmétique. On peut donc toujours laisser la question de côté.*

Une dernière remarque est que, du point de vue des fondements, il n'est nullement nécessaire d'identifier les objets à leur représentation pour pouvoir utiliser la théorie des ensembles comme base de l'édifice mathématique : si la théorie des ensembles est cohérente, elle assure la cohérence de la copie du monde mathématique incluse dans les ensembles, donc celle du monde mathématique lui-même. ◁

4.5. Autres approches.

► La théorie des ensembles n'est pas la seule à offrir un cadre fondationnel, et d'autres approches sont possibles. ◄

▷ *Une autre raison de contester le dogme « tout est ensemble » et de plaider pour le mainten d'une distinction entre les objets mathématiques et leur représentation par des ensembles est que la théorie des ensembles n'est pas la seule à permettre une représentation des autres types d'objets et, partant, d'offrir un cadre pour une théorie des fondements.*

Le lambda-calcul (ou λ*-calcul), et les théories de types qui s'y rattachent, sont des approches où « tout est fonction ». Les fonctions, considérées comme procédures d'évaluation et non comme collections de couples, sont les objets de départ, et les opérations de base sont l'application d'une fonction à son argument (qui est une autre fonction), c'est-à-dire l'opération* $(f,x) \mapsto f(x)$, *et l'abstraction qui à toute spécification explicite d'une correspondance associe une fonction, c'est-à-dire l'opération* $(x, f(x)) \mapsto f$, *notée* $(\lambda x)(f(x))$. *On peut alors représenter les entiers par des fonctions, et, de là, d'autres types d'objets. L'approche par le lambda-calcul n'est pas exactement de même nature que celle par les ensembles, dans la mesure où seuls des objets effectivement définissables sont considérés : par exemple, on ne pourra et on ne cherchera à représenter que des ensembles définissables d'entiers, et non tous les ensembles d'entiers comme on le fait — de la façon qu'on a qualifiée d'imprédicative — en théorie des ensembles. Ainsi, cette approche ne prétend pas à la même universalité que la théorie des ensembles, mais, en contrepartie, elle prend mieux en compte les questions d'effectivité.*

La théorie des catégories est une approche où « tout est morphisme ». Elle place la notion de morphisme à la base de la construction, les opérations de base étant la composition des flèches. Là encore, le point de vue n'est pas exactement comparable à celui de la théorie des ensembles, mais on peut définir une très grande diversité d'objets mathématiques comme catégories.

Mentionnons encore qu'il a été suggéré de faire jouer aux probabilités un rôle fondateur, même si, pour le moment, aucun système concret n'a vraiment émergé dans cette direction.

Il n'y a pas lieu d'opposer ces diverses approches, dont chacune est adaptée à certains aspects d'un édifice mathématique divers et complexe. Pour sa part, la théorie des ensembles permet une exploration conceptuelle de la notion d'infini que les autres approches ne semblent pouvoir atteindre, tandis que les théories de types sont mieux adaptées pour étudier les questions d'effectivité, et que les théories de catégories semblent mieux aptes à mettre en évidence des phénomènes géométriques où la théorie des ensembles reste pataude. Il y a donc complémentarité

[21] à l'exception de ceux qui sont trop gros telle la classe de tous les ordinaux

davantage qu'antinomie entre des approches dont aucune n'a vocation à éclipser les autres, et entre lesquelles du reste divers ponts ont été établis. En tout cas, l'existence de ces approches diverses est un argument supplémentaire soulignant qu'il n'y a ni nécessité, ni bénéfice à poser le dogme « tout est ensemble », risqué et daté. $\triangleleft$

Exercices

Exercice 1. (ensembles V_α) Montrer que, si λ est un ordinal limite, alors l'ensemble V_λ est clos par passage à un élément, à un sous-ensemble, à l'ensemble des parties, à l'union et l'intersection, et par formation d'ensembles finis de fonctions, et d'ensemble des fonctions.

Exercice 2. (ensembles V_n) Montrer que, pour tout entier n, l'ensemble V_n est fini.

Exercice 3. (ordinaux finis) (i) Montrer, par induction sur n, qu'on a $n + \omega = \omega$ pour tout entier n. [On rappelle qu'on se place dans le système de Zermelo, où la seule hypothèse est que ω est le plus petit ordinal récurrent.] En déduire en utilisant uniquement la définition récursive de l'addition que, pour tous n, p entiers, on a $n + p < \omega$, puis $n + p = p + n$.

(ii) Traiter de même la multiplication.

Exercice 4. (représentation par des ensembles) Montrer que tout entier relatif est (représenté par) un élément de $V_{\omega+3}$, et que $\mathbb{Z}$ est (représenté par) un élément de $V_{\omega+4}$. Montrer de même que $\mathbb{Q}$ est (représenté par) un élément de $V_{\omega+8}$, puis $\mathbb{R}$ (par) un élément de $V_{\omega+13}$, et $\mathbb{C}$ (par) un élément de $V_{\omega+15}$. Que dire de l'ensemble de toutes les fonctions de $\mathbb{R}$ dans $\mathbb{R}$? Et de l'ensemble des fonctions continues ?

Exercice 5. (théorème de Cantor–Bendixson) Vérifier que la démonstration du théorème de Cantor-Bendixson donnée au chapitre II est valide dans le cadre de ZF.

Exercice 6. (remplacement) Montrer que, en présence des axiomes d'extensionnalité, des parties et de séparation, les axiomes de remplacement entraînent l'axiome de la paire. [Montrer l'existence de l'ensemble $\mathfrak{P}(\mathfrak{P}(\emptyset))$, et utiliser le remplacement associé à la formule ($\boldsymbol{x} = \emptyset$ et $\boldsymbol{y} = a$) ou ($\boldsymbol{x} = \{\emptyset\}$ et $\boldsymbol{y} = b$).]

Exercice 7. (remplacement) Appelons axiome de remplacement fort pour F l'axiome obtenu en remplaçant la condition « $\Rightarrow y \in b$ » par « $\Leftrightarrow y \in b$ ». Montrer que les axiomes de remplacement forts impliquent les axiomes de séparation. [Pour obtenir la séparation associée à $\mathsf{F}(\boldsymbol{x})$, considérer le remplacement associé à $\boldsymbol{y} = \boldsymbol{x}$ et $\mathsf{F}(\boldsymbol{x})$.]