LES THÉORIES DES ENSEMBLES ET L'ENSEMBLE DE TOUS LES ÉTANTS

I. Considérations préliminaires concernant l'ontologie sous-jacente au système de Zermelo (1908).[1].

Elle est très répandue parmi les experts de la théorie des ensembles l'opinion générale que, tandis que les systèmes «typés» de la théorie sont d'une inspiration «philosophique», les systèmes qui s'appuient, afin de résoudre les paradoxes, sur l'axiome de séparation de Zermelo ont plutôt un caractère technique destiné à fonder les mathématiques et ne touchent pas les grands enjeux philosophiques de la théorie des types. On divise donc la théorie en deux grandes branches. La première, qui correspond aux systèmes d'inspiration russellienne, est appelée «théorie des types», et la seconde, qui correspond aux systèmes d'inspiration zermelienne[2], écoute sous le nom de «théorie des ensembles». La précédente est une distinction qu'on trouve même parmi les spécialistes les plus distingués. Jean van Heijenoort, par exemple, caractérise ainsi le système de Zermelo de 1908, dans son introduction à l'article de la même année:

«The two responses are extremely different, the former [*théorie des types*] is a far-reaching theory of great significance for logic and even ontology, while the latter is an immediate answer to the pressing needs of the working mathematician»[3].

De prime abord, nous avons à constater qu'en ce qui concerne la dévalorisation générale du côté philosophique des systèmes zermeliens leur diversité les rend parfois philosophiques aussi. Reindhardt, pour prendre un exemple, a formulé un système[4], utilisant des notions comme «propriété imaginaire» et «univers en expansion», qui ne peut pas être considéré comme s'adressant uniquement au *mathématicien en travail.* Un autre argument qui va contre cette

1. E. Zermelo, Untersushungen über die Grundlagen der Mengenlehre, I, *Mathemaschische Annalen,* 59, 1908, pp. 261-281.

2. Par «inspiration zermelienne», nous faisons entendre non seulement Z., Z.F. et leurs variantes mais aussi tout système, ayant comme axiome ou théorème la *séparation.*

3. J. van Heijenoort, *From Frege to Gödel, A source Book in Mathematical Logic, 1879-1931,* Cambridge Mass., Harvard UP, 1967, p. 199.

4. W. Reinhardt, Set existence principles of Shoenfield, Ackermann, and Powell, *Fundamenta Mathematicae,* 84, 1971, pp. 5-34.

caractérisation générale est que les systèmes d'inspiration zermelienne qui utilisent la notion de classe (dans le sens de Bernays-Quine, et non celui de von Neumann), sont d'une inspiration métathéorique[5], qui intéresse très peu le *mathématicien en travail*.

Les raisons précédentes mises à part, notre objection principale est qu'on trouve cet aphorisme à l'introduction même de l'article historique de Zermelo. Nous voulons dire que, tandis que la théorie de Zermelo est pour le *mathématicien en travail* standardisée sous la forme que Fraenkel et Skolem lui ont attribuée, ce premier système de Zermelo est plein de considérations philosophiques, qui couvrent un domaine très vaste.

Les controverses parmi les théoriciens sont dues non pas à un désaccord clair et précis, mais à une différence concernant le prisme par lequel ils abordent les questions. Nous allons argumenter en faveur du caractère philosophique du système de Zermelo, et interpréter les «corrections», qui lui ont été effectuées, comme dérivantes de l'exigence du mathématicien (très justifiée de son point de vue), de vouloir fonder la mathématique, en ayant le moindre possible affaire avec la philosophie. L'examen de deux points de désaccord va illustrer ce qui vient d'être postulé: i) les *individus (Urelemente)* et ii) la *propriété définie (definite Eigenschaft)*.

i) Très peu de mathématiciens utilisent encore la notion d'*individu* de Zermelo. Et à juste titre. En tant que moyen pour fonder les mathématiques, cette notion est indispensable, seulement si on suit la notation initiale de Zermelo par rapport au processus d'engendrement des nombres naturels par l'ensemble vide: ...$\{\{0\}\}$... Si, par contre, on utilise la notation de von Neumann à ce propos ...$\{0,\{0\}\}$..., et moyennant l'axiome de remplacement et l'axiome d'infini, selon sa formulation par le même auteur, non seulement on ne les a plus besoin, mais on peut accéder à des cardinalités supérieures à $\aleph_\omega$, point limite pour le système de Zermelo. C'est-à-dire que, tandis que la notation de von Neumann semble au début moins économique à celle de Zermelo, la récompense d'avoir accepté ce surplus est énorme. Non seulement on est en position de fonder tout ce que Zermelo a pu fonder, mais nous pouvons (moyennant l'axiome de remplacement et celui d'infini), aller beaucoup plus loin; et tout ça sans avoir aucun besoin des individus, et des problématiques *ensembles impurs*, qui vont avec. Le mathématicien donc et ce logicien qui considère la logique comme les mathématiques d'ordre(s) inférieur(s) sont tout à fait justifiés de laisser complètement tomber les individus. Il ne faut pas quand même procéder à la confusion suivante. Rejeter un concept, parce qu'on n'en a pas besoin, est quelque chose tout à fait différente de le rejeter parce qu'il est un mauvais concept. Ou, pour dire mieux, rejeter un outil si on n'en a pas besoin est justifié, seulement si on s'est mis d'accord sur ce qu'on a besoin à faire.

5. Voir *infra*.

L'outil rejeté peut être inutile pour un certain but, mais indispensable pour un autre. Les *Urelemente* sont inutiles en ce qui concerne la fondation de l'analyse classique, mais tout à fait indispensables, si notre point de vue est celui du logicisme au sens large, et ils vont fatalement réapparaître dans le discours, si nous considérons à franchir des questions plus générales concernant, par exemple, le monde[6]. Nous faisons l'hypothèse que l'horizon de la problématique zermelienne était bien plus vaste, et que ses individus font preuve à ça. Nous croyons que cette hypothèse va être renforcée par la considération de sa définition de la *propriété définie.*

ii) La définition de la *propriété définie* par Zermelo est la partie la plus controversée de son système. Presque aucun parmi les logiciens n'était content de cette notion, et Zermelo de sa part n'était pas content des résultats des efforts pour son amélioration. La définition de la «question», «assertion», «énoncé ouvert» *bien définis* est la suivante:

«Une question ou une assertion E est dite *«bien définie»,* si les relations fondamentales du domaine permettent de décider, sans aucun arbitraire, de sa validité ou de sa non-validité à l'aide des axiomes et des lois logiques universellement valides. *De même un «énoncé ouvert [Klassenaussage] E(x), dans lequel la variable x peut parcourir tous les individus d'une classe K, sera dit »bien définie«, s'il est bien défini pour chacun des individus x de la classe K.* Par conséquent, la question de savoir si $a \in b$ ou non, est toujours bien définie, de même que celle de savoir si $M \subset N$ ou non»[7].

Nous allons exposer notre hypothèse tout de suite: La définition en elle-même est irréprochable, et la notion claire et stable. Le problème est qu'elle est de caractère *métathéorique,* chose qui la rend pratiquement inutile pour la recherche mathématique[8].

Elle est *métathéorique* dans le sens, où elle ne fait pas partie de la *syntaxe,* qui détermine les normes de formation des *propriétés bien définies,* mais elle précise quelle *doit* être cette *syntaxe,* quelle *doit* être sa tâche. Elle précise donc quelle condition une définition de la *propriété définie* doit remplir, afin d'être «satisfaisante», et de ce point de vue elle est une définition *de la définition* de la propriété définie. L'espace reste ouvert à la question concernant la formulation exacte de cette dernière, mais n'importe quelle qu'elle soit, elle doit valider comme *définies* les *énoncés ouverts* qui sont «définis» pour tout élément x

---

6. Quine exprime cette nécessité de la manière épigrammatique suivante: «Zermelo's system and others are sometimes rendered as pure set theory: set theory without individuals. [...] *However, set theory plays a part in our overall theory of the world only by going impure».* W. O. QUINE, *The Roots of Reference,* Illinois, Open Court, 1974.

7. F. RIVENC – Ph. de ROUILHAN, *Logique et fondements des mathématiques. Anthologie 1850-1934,* Paris, Payot, 1992, p. 372. C'est nous qui soulignons.

8. Tout de même nous n'allons pas jusqu'à attribuer à Zermelo l'intention précise de vouloir formuler un critère métathéorique.

d'une classe K et rejeter le reste. La seule chose qui suit est que l'énoncé est «défini», si et seulement s'il est «défini» pour chaque élément de la classe. Mais quand est-ce qu'on peut affirmer qu'un énoncé est défini pour un élément? Cela ne devient pas explicite avant la formulation de l'axiome de séparation du même article, mais même dans la définition que nous avons exposée, Zermelo énonce deux conditions métathéoriques d'importance majeure: (i) que l'énoncé ouvert n'a pas un domaine de signification illimité mais une classe, et (ii) qu'«être défini pour la classe» équivaut à «être défini pour chacun de ses éléments».

C'est vrai que cette définition *de définition* ne peut servir à rien au *mathématicien au travail*. On ne peut pas prendre un par un tous les éléments de tout ensemble et attendre jusqu'ils s'épuisent, afin de tester chaque propriété, et en plus, on ne possède aucun critère rigoureux pour savoir quand est-ce que l'énoncé est défini pour un élément. Mais cela n'était nullement ce à quoi Zermelo visait; il n'a pas formulé un critère empirique, mais un critère *méta*-théorique. Il n'avait pas lancé le mathématicien à l'épreuve impossible de tester empiriquement toute propriété possible, mais il avait chargé le théoricien à procurer une définition, qui serait conforme aux exigences de sa *méta*-définition.

De l'autre côté, on lui ferait un mauvais procès, si on lui attribuait que par le «à l'aide des axiomes et des lois logiques», il entend que le travail est déjà achevé. Il entend par là, nous croyons, exactement le contraire: que c'est précisément cette *aide* que nous devons chercher, afin de préciser la définition[9].

Dans ce sens, Fraenkel (i) et Skolem (ii), en essayant de trouver la forme précise de cette définition (se situant, c'est-à-dire, dans le niveau théorique), ils étaient dans la ligne de Zermelo plus qu'ils ne le croyaient.

(i) «To arrive at a precise formulation of Axiom III (axiom of separation), which as formulated in *Zermelo 1908a* contains the imprecise notion «definite», we shall understand by *function* a rule of the following kind: an object $\varphi(x)$ shall be formed from a («variable») object x that can range over the elements of a set, and possibly from further given («constant») objects, by means of a prescribed application (repeated only a finite number of times, of course, and denoted by φ) of Axioms II-IV [*du système de Zermelo* 1908]»[10].

(ii) «*By a definite proposition we now mean a finite expression constructed from elementary propositions of the form a* ε *b or a=b by means of the five operations mentioned.* [*conjonction, disjonction, négation, quantificateurs*

9. Par ailleurs, il est explicite, dans le même article, par rapport à son caractère inachevé: «J'espère, à tout le moins, que le travail préliminaire présenté ici se révélera utile pour des investigations ultérieures, traitant de tels problèmes de fond.», J. VAN HEIJENOORT, *op. cit.* p. 371.

10. A. FRAENKEL, Der Begriff «definit» und die Unabhängigkeit des Auswahlaxioms, *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, 1922, pp. 253-257; J. VAN HEIJENOORT, *op. cit.*, pp. 286.

universel et existentiel] This is a completely clear notion and one that is sufficiently comprehensive to permit us to carry out all ordinary set-theoretic proofs. I therefore adopt this conception as a basis here»[11].

Ça va de soi que pour le mathématicien la notion de *fonction* est plus utile que celle de *propriété,* ou de *condition*[12]. Et maintenant, quand on arrive à la définition de Skolem, le contentement est général. On peut pour n'importe quelle formule, dire à l'avance, si elle est *bien définie* ou pas, et, en plus, on se dispose d'une très bonne *syntaxe* pour la construction de toute *propriété définie*.

C'était le tour de Zermelo de ne pas être content[13]. De son refuge *méta*-théorique, il pouvait continuer de jouer le rôle du régulateur. Selon ses ambitions, la notion de *fonction* ne devrait pas être primitive (comme c'est le cas chez Fraenkel), mais elle devrait émaner de la notion de *propriété* et d'*énoncé ouvert*. Par ailleurs, c'était à cause de la même «préférence», qu'il n'a pas été content de l'axiome de remplacement. En plus, il reprochait à Fraenkel que la notion de nombre était chez lui présupposée. Elle, non plus, ne devrait pas, pour Zermelo, être considérée comme notion primitive. Comment donc une théorie axiomatique, qui ne compte pas les notions de *fonction* et de *nombre* parmi ses notions primitives, peut être une *réponse immédiate aux* «pressing needs» *du mathématicien au travail?*

Maintenant, en ce qui concerne la définition de Skolem, tandis qu'elle détermine exhaustivement la forme normative des propriétés, elle les délimite aux propriétés du premier ordre; limitation qui n'empêche pas la fondation des mathématiques, mais dont la seule raison d'être est précisément qu'on peut fonder les mathématiques sans des propriétés d'ordre supérieur. Ici encore les exigences d'un logicisme au sens large nous amèneraient éventuellement à considérer de telles propriétés[14]. Dans le même sens, la définition de Fraenkel avait déjà délimitée les propriétés à celles qui remplissent des *conditions fonctionnelles,* tandis que celle de Zermelo, qui utilise le terme d'*énoncé ouvert,* laissait la possibilité même pour la considération des relations, qui ne sont pas des fonctions.

---

11. Th. SKOLEM, Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre, *Conférences au 5ème congrès des Mathématiciens Scandinaves,* 1922, pp. 218-232; J. VAN HEIJENOORT, *op. cit.*, pp. 292-293.

12. Voir A. FRAENKEL – Y. BAR-HILLEL, *Foundations of Set Theory,* Amsterdam, North-Holland, 1959, p. 119.

13. Voir E. ZERMELO, Ueber den Begriff der Definitheit in der Axiomatik, *Fundamenta Mathematicae,* 14, 1929.

14. L'effet mis à part, nous ne voyons pas d'autre affinité entre la façon, dont Frege rend les fonctions de niveaux supérieurs inutiles, et la façon, dont Skolem défend la formulation de propriétés d'ordre supérieur. Chez Frege on peut représenter les fonctions des niveaux élevés par leurs *objets*: les fonctions inférieures (d'où une instance du «paradoxe de Frege»). Chez Skolem nous n'avons qu'une restriction *ad hoc*.

Et nous arrivons à cette différence de point de vue, dont nous parlions au début. C'est à juste titre que Fraenkel a remplacé la notion d'*énoncé ouvert bien défini* par celle de *fonction,* et à juste titre que Skolem a délimité les propriétés au premier ordre. La notion de fonction ne présente aucune ambiguïté, et une logique ayant des propriétés d'ordre supérieur, bien que possible, demeure, pour l'essentiel inutile. D'abord tout, elle est inutile pour fonder les mathématiques, mais nous avons déjà considéré dans quel contexte «inutile» et «mauvais» sont synonymes.

Si tous les trois (Zermelo, Fraenkel, et Skolem) visaient vers le même but: la fondation des mathématiques, la définition de Zermelo est (en tant qu'inutile), une mauvaise définition, mais si, selon notre hypothèse, Zermelo envisageait un but encore plus général, sa définition peut servir en tant que principe régulateur pour des recherches ultérieures, puisqu'en tant que *métathéorique,* reste valable dans tout contexte logique (même non-standard)[15].

Afin de conclure avec cette partie visant à revendiquer le «caractère philosophique» de la *théorie des ensembles,* nous allons faire citer un passage de *Set Theory,* selon lequel, il ne faut pas, en général, faire de distinctions rigoureuses entre la théorie des types et la *théorie des ensembles*:

«Quantification theory proper came to be called the first-order predicate calculus. It was a regrettable trend. Along with obscuring the important cleavage between logic and the »theory of types« (meaning set theory with types), *it fostered an exaggerated if fogy notion of the difference between the theory of types and »set theory« (meaning set theory without types) – as if the one did not involve outright assumption of sets the way the other does»*[16].

II. La solution des paradoxes selon les *théories des ensembles.*

L'axiome central du système de Zermelo de 1908 est l'axiome de séparation, et l'axiome central de la Z.F. est celui de remplacement. Nous allons pendant ce chapitre voir comment ces deux axiomes parviennent à empêcher les paradoxes de Russell et Cantor.

L'axiome de séparation est en lien profond avec la *propriété définie;* on peut dire qu'il en écoule. L'affaiblissement de l'axiome d'abstraction est effectué par la condition imposant que la propriété $f(x)$ doit être conçue, non pas comme ayant un domaine de signification illimité ou restreint dans un *type logique* (théorie des types), mais un *ensemble,* dont l'existence est déjà donnée. Reprenons la formulation de Zermelo (1908):

---

15. En faite, la définition de Fraenkel tombe dans une logique multivalente, dans le sens, où l'énoncé peut être «défini» pour un élément sans être ni vrai, ni faux. Celle de Zermelo reste valide, parce que l'axiome de séparation n'y figure pas, et donc le tiers exclu non plus.

16. W. O. QUINE, *Set Theory and its Logic,* Cambridge, Mass, Harvard UP, 1963; 2ème éd. révisée, 1969; et *The Roots of Reference, op. cit.,* p. 172.

«Si l'énoncé ouvert *E est bien défini* pour tous les éléments d'un ensemble *M*, *M* a toujours un sous-ensemble $M\varepsilon$ qui contient tous les éléments *x* de *M*, pour lesquels *E(x)* est vrai, et ceux-là seulement.» [...] En premier lieu, avec cet axiome, on n'a jamais le droit de *définir* des ensembles *de façon indépendante*, mais seulement en tant que sous-ensembles obtenus par *séparation* d'ensembles déjà donnés,[...]»[17].

La propriété donc doit être non seulement *bien définie*, mais, en plus, prendre ses valeurs d'un ensemble *M*, déjà donné.

Voyons maintenant l'axiome dans sa forme standard symbolique[18]:

$(\exists y)\,(\forall x)\,(x\in y \leftrightarrow x\in z \wedge f(x))$

L'ensemble y est constitué par les éléments qui satisfont *f*(x), *et qui* appartiennent à *z*. Autrement dit, *f*(x) prend ses arguments parmi les éléments de l'ensemble *z*, et l'ensemble des arguments qui satisfont f(…) est l'ensemble *y*. On doit par conséquent *séparer, extraire*, l'ensemble visé, d'un ensemble, qui se trouve «déjà là». Quand, par exemple, on veut établir l'existence d'un ensemble correspondant à une fonction, on procède très souvent (c'est une routine dans le formalisme)[19], comme par la suite: i) On trouve un ensemble *déjà donné* qui contiendrait tous les éléments de l'ensemble visé (sans présupposer ce dernier dans la définition de l'ensemble qui le contient), et ii) on applique en toute sécurité l'ancien axiome d'abstraction. Cet affaiblissement est suffisant, afin d'empêcher les paradoxes:

i) En ce qui concerne le paradoxe de Russell, bien que l'idée de l'axiome soit assez intuitive[20], l'empêchement en est moins. Par conséquent un minimum de formalisme est ici indispensable:

L'application de l'axiome en prenant *ne pas appartenir à soi-même* en tant que *f*(x), et «w» en tant que l'«ensemble» qu'elle définit va donner non pas $w\in w\leftrightarrow w\notin w$, mais $w\in w\leftrightarrow w\in z \wedge w\notin w$. Formule, qui n'est pas contradictoire, puisque rien n'empêche que $w\notin z$. Et en effet, on peut en déduire directement que: $\neg(\exists z)(w\in z)\wedge(\forall x)(x\in w\leftrightarrow x\notin x)$. Cette dernière formule équivaut à la non-existence de w. L'idée intuitive sous-jacente est que w est si grand qu'on ne peut pas trouver aucun ensemble pour le faire *séparer*[21]. Dans le cas du paradoxe de Russell, quand même, cela n'est pas évident dès le début comme c'est le cas avec celui de Cantor. Cette façon indirecte d'attaquer le problème a été un argument contre le système et en faveur de la théorie des types, qui, elle,

---

17. F. Rivenc – Ph. de Rouilhan, *Logique et fondements des mathématiques. Anthologie 1850-1934*, Paris, Payot, 1992, p. 373.

18. En réalité il s'agit d'un schéma infini dénombrable d'axiomes.

19. Voir P. Suppes, *Axiomatic Set Theory*, Dover, 1972, p. 62; (1ère éd. 1960).

20. Ici, et dans des contextes pareils, nous entendons par intuitif *ce qui peut être saisi sans que la récurrence au formalisme soit indispensable.*

21. W. O. Quine, *Set Theory and…, op. cit.*, p. 278, et P. Suppes, *op. cit.*, p. 202.

s'attaque directement à la notion d'auto-appartenance[22].

ii) Contrairement au paradoxe de Russell, la dissolution du paradoxe de Cantor peut être conçue de façon non pas formelle: Pour démonter l'existence de l'ensemble de tous les ensembles, il faut le concevoir comme sous-ensemble d'un autre ensemble, déjà donné. Tâche impossible. Nous ne pouvons pas faire extraire l'ensemble universel d'aucun autre. C'est l'inverse qui devrait, à la limite, avoir toujours lieu. Une fois que la multiplicité *de tous les ensembles* n'existe pas[23], le théorème de Cantor est inapplicable pour la même multiplicité. En ce qui concerne la solution de ce paradoxe les deux théories (types, *ensembles*) ont un point en commun: elles empêchent sa formulation en s'attaquant directement à la notion d'ensemble universel.

Le même lien qui unit l'axiome de séparation avec la définition zermelienne de la *propriété définie* associe la définition fraenkelienne de fonction à l'axiome de remplacement. L'idée de l'axiome est que chaque multiplicité, afin de pouvoir être ensemble, doit pouvoir être mise en bijection avec un ensemble, ou une partie d'un ensemble déjà existant.

«For any set a, if $\mathbf{P}(t,x)$ is a functional condition on a then there exists a set which contains exactly those elements x for which $\mathbf{P}(t,x)$ holds for some $t \in a$»[24]. Et sous sa forme standard symbolique[25]:

$(\forall x)\,(\forall y)\,(\forall z)\,(x \in A \wedge \varphi(x,y) \wedge \varphi(x,z) \rightarrow y=z) \rightarrow (\exists B)\,(\forall y)\,(y \in B \leftrightarrow (\exists x)\,(x \in A) \wedge \varphi(x,y)))$.

L'ensemble déjà existant est ici l'ensemble *A*, et la fonction qui le met en bijection avec *B* (l'ensemble qu'on veut obtenir), est représentée par φ. La condition «$x \in A \wedge \varphi(x,y) \wedge \varphi(x,z) \rightarrow y=z$» est mise, afin de nous empêcher de prendre ...∈ ___ comme φ, et obtenir l'ensemble universel, ou, en général, de prendre comme φ une relation non pas fonctionnelle et obtenir ainsi des multiplicités plus grandes que *A*.

L'affinité entre Zermelo et Fraenkel est évidente: la taille de l'ensemble *déjà existant* garantit que l'ensemble *séparé - remplacé* ne dépassera pas cette taille. Voilà, par conséquent, une grande différence entre la théorie des types et les *théories des ensembles.* Ces dernières sont clairement des mises en œuvre de la stratégie que Russell appelait *limitation de taille,* et se basent sur l'idée que les problèmes se présentent à de *très grandes multiplicités,* qu'il faut faire bannir du calcul ensembliste.

---

22. Par ailleurs, l'axiome de fondation, bien qu'il s'attaque directement à l'expression «$x \in x$», vise plutôt vers l'élimination d'autres paradoxes (cf. Mirimanoff).

23. Selon les formules laconiques de Halmos: *«nothing contains everything»*, et *«there is no universe»;* dans P. Halmos, *Naive Set Theory,* Princeton, van Nostrand, 1960, pp. 6-7.

24. F. Rivenc – Ph. de Rouilhan, *Logique et fondements des mathématiques. Anthologie 1850-1934,* Paris, Payot, 1992, pp. 50-51.

25. Ici aussi il s'agit d'un schéma infini dénombrable d'axiomes.

Selon l'axiome de remplacement nous avons deux voies à empêcher les paradoxes. On peut: (i) soit en déduire l'axiome de séparation et procéder comme par-dessus, soit (ii) empêcher leur formation par l'axiome de remplacement tout seul.

Contrairement à l'axiome de séparation, le blocage du paradoxe de Cantor effectué par l'axiome de remplacement ne peut pas être saisi intuitivement, et cela parce que l'«inexistence» de l'ensemble universel n'est pas évidente. Tout au contraire, selon l'intuition, on pourrait essayer de trouver une fonction pour mettre en bijection une multiplicité de très grande taille (univers) avec un ensemble de taille égale. Le problème est qu'on ne peut pas construire un ensemble si grand. Si nous pouvions, par exemple, démonter l'existence de l'«ensemble de tous les ordinaux», nous pourrions démontrer par la suite que toute multiplicité de *très grande taille* constitue un ensemble. Le problème est que nous n'en avons aucune pour commencer.

Afin de conclure, soulignons l'importance d'un processus que la théorie des types partage en commun avec les *théories des ensembles*: c'est le célèbre processus d'*itération*[26]. Putnam et Benacerraf semblent postuler qu'il est le principe fondamental de la théorie. Dans leur recueil classique[27] sur la philosophie des mathématiques, ils ont choisi d'inclure au chapitre qui s'occupe de la théorie des ensembles trois travaux qui traitent du sujet d'itération. Suivons comment Boolos définit ce processus:

«That is a rough statement of the iterative conception of set. According to this conception, no set belongs to itself, and hence there is no set of all sets; for every set is formed at some earliest stage and has as members only individuals or sets formed at still earlier stages»[28].

Et voilà comment Parsons exprime la même idée:

«If we suppose that the elements of a set must be »given« before the set, then no set can be an element of itself, and there can be no universal set. The reasoning leading to the Russell and Cantor paradoxes is cut off»[29].

Dans le cadre de la *théorie naïve*, le processus s'arrête une fois que l'ensemble universel est obtenu. Dans la théorie des types et celle des *ensembles*, il s'agit d'un processus qui ne s'arrête nul part. Étant donné n'importe quel ensemble, il

---

26. Le terme fait plutôt partie de la *philosophie* de la théorie, et tente d'exprimer le principe général sous-jacent à la théorie des types, ainsi qu'à la *théorie des ensembles*, en tant que point de différenciation, par rapport à la *théorie naïve*.

27. P. Benecerraf – H. Putnam, *Philosophy of Mathematics*, Cambridge UP, 1983^2, pp. 447-570.

28. G. Boolos, The iterative conception of set, dans P. Benecerraf – H. Putnam, *Philosophy...*, *op. cit.*, p. 493.

29. C. Parsons, What is the iterative conception of set?, dans P. Benecerraf – H. Putnam, *Philosophy...*, *op. cit.*, p. 505.

y a des éléments, qui ne lui appartiennent pas. Le grand avantage des *théories des ensembles,* par rapport aux types, est que l'itération impose aux ensembles *typés* des restrictions de quantification. Comme nous avons vu, la quantification universelle est défendue, sous peine d'obtenir les paradoxes. Aux *théories des ensembles,* par contre, on peut quantifier sans prendre en compte le niveau d'itération de chaque ensemble. L'expression, par exemple, «$\forall x$», quand x est un ensemble[30], a du sens dans une formule, sans que le niveau d'itération de x soit précisé par des indices[31]. La même chose n'est possible dans le formalisme *typé,* que par le moyen de l'*ambiguïté systématique*[32].

III. Les paradoxes dans le cadre des théories avec *classes.*

Les noms «ensemble» et «classe» peuvent, dans le langage courant, être considérés comme synonymes. Le fait qu'il y a deux termes exprimant la même notion a été exploité par les théoriciens des ensembles, qui ont procédé à faire une distinction rigide entre les multiplicités. Cette distinction, comme nous avons vu au début, date des travaux de Cantor et de Russell. Quand Cantor distingue, dans les lettres de 1899 à Dedekind, entre les *multiplicités consistantes* (ensembles) et les *multiplicités inconsistantes,* tout en signalant que toutes les deux sont *complètement déterminées,* il anticipe la possibilité que toutes les multiplicités ne peuvent pas être bien ordonnées, et, par conséquent, ne pas être considérées comme des unités, sans pour autant cesser de constituer des objets bien définis. La distinction que nous allons évoquer ne s'identifie pas à la précédente, mais elle lui doit, sans doute, beaucoup[33].

En 1925, John von Neumann a, afin de compléter le calcul, inventé une manière de le généraliser. Dans son article historique de la même année[34], il a essayé de remplir un «trou noir» de la théorie en introduisant la notion de *classe.* Le problème était que les théories des ensembles (la théorie des types y incluse), parvenaient à surpasser les paradoxes et à fonder sans contradiction une très grande partie de l'analyse classique, tout en étant obligées de rejeter complètement un très grand nombre d'expressions, à savoir, les expressions imprédicatives. Une grande partie de ces expressions, bien qu'elle soit trouvée coupable d'avoir amené aux paradoxes, continuait à posséder dans le langage, et parfois même dans le langage formel, un sens très précis et bien fondé. Prenons, par exemple, le cas classique du prédicat *être identique à soi-même.* Non seulement il est un prédicat stratifié, mais il constitue une des expressions

30. Exprimé dans l'idiome *Bourbaki:* $\forall_{(\text{ens } x)} x$.

31. Toutefois, quelques foncteurs sont nécessités, afin de bannir les expressions non stratifiées.

32. W. O. Quine, *Set Theory and..., op. cit.,* pp. 253-254.

33. Peut-être pas dans le sens d'influence directe. Les lettres de Cantor de juillet-août 1899 n'ont été publiées qu'en 1932.

34. J. Von Neumann, Eine Axiomatisierung der Mengenlehre, *Journal für reine und angewandte Mathematik,* 1925; J. Van Heijenoort, *op. cit.,* pp. 394-413.

le plus souvent répandues dans la logique. En tant, quand même, que prédicat d'abstraction de classe, il se prouve non légitime, parce qu'il aurait du avoir comme extension l'ensemble universel. Voyons un autre exemple: On n'a aucune raison de faire bannir du formalisme l'expression $\cap\varnothing$, sauf qu'elle dénote l'ensemble universel[35]. Nous avions encore signalé que la *théorie des ensembles* est une réalisation d'un projet conçu par Russell, et codifié sous le nom de *limitation de taille.* En effet, les antinomies pourraient être ainsi résolues. Les multiplicités problématiques n'étaient que celles de *très grande taille,* et le problème était que, par l'acte même de définition, on leur faisait introduire en tant que membres propres, des éléments, qui n'y existaient pas avant cet acte[36].

L'ensemble universel est imprédicatif, parce qu'en lui appliquant le théorème de Cantor, on lui «introduit» des membres nouveaux. D'où le paradoxe concernant sa cardinalité. Mais, ceci mis à part, qu'avons-nous à reprocher au prédicat *être identique à soi-même*? Le même vaut pour le prédicat *être nombre ordinal.* Si nous n'avions pas dans la main le paradoxe de Burali-Forti, que pourrions-nous avoir contre le prédicat *être nombre ordinal*?

L'idée de von Neumann était d'accepter comme légitimes les notions qui définissent des multiplicités de *très grande taille,* et appeler ces multiplicités des *classes.* Ce n'était pas juste une manœuvre esthétique. Ça aurait été le cas, si ces multiplicités étaient aliénées du calcul, et considérées comme des pluralités étranges, ne pouvant pas recevoir un sens exact dans le sein de la théorie. Le pas crucial dans l'idée de von Neumann est que ces multiplicités peuvent, dans un certain degré, constituer des objets de la théorie, sans pour autant amener aux paradoxes. Et le «dans un certain degré», qui est, en outre, notre garantie qu'elles ne vont pas amener à des paradoxes, est la restriction suivante, qui peut être considérée comme définition de la notion de classe: Les classes peuvent avoir des ensembles comme éléments, mais, contrairement à ces derniers, ne peuvent pas appartenir, ni à des ensembles, ni à des classes. Ainsi le prédicat *être identique à soi-même* dénote la classe universelle. L'univers de la Z.F. est l'union de la classe universelle ($\cup$U), qui est, elle aussi, une classe. L'argument est que, (i) puisque les classes n'appartiennent nul part, et, puisque l'union d'une classe est définie comme la classe de tous les éléments de ses éléments, les classes ne peuvent pas appartenir à $\cup$U. De l'autre côté, (ii) puisque chaque ensemble peut appartenir à un autre, et U contient tous les ensembles possibles, tous les ensembles dans U appartiennent nécessairement à un autre, et par conséquent $\cup$U contient tous les ensembles.

---

35. Voir P. SUPPES, *op. cit.*, p. 41.

36. Voir comment Russell exprime le même problème dans F. RIVENC – Ph. de ROUILHAN, *Logique et fondements des mathématiques. Anthologie 1850-1934,* Paris, Payot, 1992, p. 315: «L'inexistence d'une telle classe résulte du fait que la supposition du contraire fait immédiatement surgir de nouvelles classes».

Voyons comment cette restriction empêche les deux paradoxes que nous examinons dès le début.

En ce qui concerne le paradoxe de Russell, le prédicat *ne pas appartenir à soi-même* détermine une multiplicité de très grande taille, mais cette multiplicité n'amène plus à aucun paradoxe. Afin de pouvoir formuler le paradoxe de Russell, il faut que nous puissions poser la question, si le prédicat s'applique à soi-même. Question privée de sens, à cause du fait qu'elle correspond (du point de vue extensionnel), à la question si la multiplicité définie par le prédicat *ne pas appartenir à soi-même,* appartient à elle-même. Cette multiplicité quand même est une classe, et la question devient privée de sens.

En ce qui concerne le paradoxe de Cantor, afin de le formuler, il faut que nous appliquions le théorème de Cantor, application qui aurait donné la phrase suivante: «$U \in PU \in PPU$...»; tandis que U devrait contenir la totalité des ensembles. Par contre «U» ne peut pas appartenir nul part, parce qu'il est une classe. Autrement dit, le théorème de Cantor n'est applicable que pour les ensembles; s'il pouvait être appliqué aux classes, il aurait contredit leur définition. Sans pouvoir appliquer le théorème, on ne peut pas obtenir aucun paradoxe.

Une fois que les classes peuvent être intégrées dans la théorie, le profit est énorme. Nous n'allons pas rentrer à des points techniques, mais nous allons mentionner quelques uns parmi leurs effets favorables: i) on n'a plus envie des schémas d'axiomes (les axiomes de séparation et remplacement peuvent être exprimés autrement), ii) on n'a plus besoin de bannir par des arguments *ad hoc* les expressions dénotant des classes, iii) on sauvegarde une très grande partie (par exemple l'axiome d'abstraction) de la théorie naïve, qui reste préférable du point de vue intuitif, en lui attribuant une validité illimitée dans le domaine des classes et une validité restreinte dans celui des ensembles, iv) on pose des critères rigoureux d'abstraction (séparation) d'ensembles *par* les classes. L'axiome de séparation, pour prendre un exemple, prend la formulation suivante «$x \cap A \in U$»[37]. Formule qui exprime que toute intersection entre un ensemble et une classe et de nouveau un ensemble. v) on peut prouver les axiomes de séparation et de remplacement, qui deviennent des théorèmes du système[38].

Bernays abandonne la distinction rigide entre ensemble et classe, et procède ainsi: Chaque expression prédicative[39] dénote une classe. Quelques classes (celles qui ne sont pas de très grande taille) sont aussi des ensembles.

Finalement, c'est à Quine que nous devons l'idée de *classe virtuelle.* Par le

---

37. W. O. QUINE, *Set Theory and..., op. cit.*, p. 315.

38. P. BERNAYS – A. FRAENKEL, *Axiomatic Set Theory,* Amsterdam, North-Holland, 1968, p. 69; (1ère éd. 1958).

39. Selon la terminologie de Bernays, en tant qu'«expressions prédicatives» sont entendues toutes les expressions de la forme ...*est*___, même celles qui, dans le sens de Russell, sont non prédicatives.

biais de cette notion, on peut même faire abstraction d'ensemble (virtuel) en utilisant une expression imprédicative, à la seule réserve que dans le formalisme l'ensemble virtuel ne doit jamais figurer avant le prédicat binaire d'appartenance[40]. Dès qu'on introduit l'appartenance comme prédicat primitif, la simulation d'existence pour les *classes virtuelles* doit être abandonnée.

Le but ultime de toutes ces manœuvres est de fonder un calcul généralisé, qui intégrera toutes les multiplicités (prédicatives ou non). Tout de même, Z.F. reste une partie autonome du nouveau système, et elle peut être conçue en tant que modèle. Il a été prouvé à plusieurs reprises que la consistance de la Z.F. équivaut la consistance du système généralisé (VN.B.)[41].

Avant de conclure nous voudrions faire une remarque, qui sera exploitée pendant la partie suivante: Bien que dans le système de von Neumann ces *multiplicités inconsistantes* (dans le sens de Cantor) fassent partie du symbolisme, et peuvent constituer des objets de la théorie, le système a été prouvé *trop fort*. Cela veut dire qu'en acceptant les classes comme objets propres, on ne peut pas empêcher de pouvoir prouver des formules fausses. Bernays[42] et Quine[43] ont, afin d'éliminer ces désavantages, réformé le système, selon l'idée que les classes ne constituent pas des objets propres de la théorie, mais des objets métathéoriques, qui fixent les normes exactes, pour que ses «objets propres» (ensembles) soient bien déterminés.

Voilà ce que dit Fraenkel, par rapport à la nature de cette restriction des «objets propres» de la théorie aux ensembles:

«More consistent [*à celle de Von Neumann*] and radical solutions to the problem of developing a system of set theory in which classes occur for the sake of convenience, while sets are still believed to be the only mathematical objects that exist, are handled in the next subsection. [...] Bernays introduced an axiom system B which differs from VNB, in addition to pure technical matters as follows. While B retains the full formalism that VNB has for handling and also retains class variables, its language does not admit quantifiers over class variables. [...] An even more radical attitude was advocated by Quine. [...] The free class variables are interpreted as metamathematical variables for class abstracts»[44].

IV. Les deux systèmes de Quine.

Considérons maintenant la méthode *zig-zag*. Celle-ci vise à une élimination

---

40. W. O. Quine, *Set Theory and..., op. cit.*, § 2-3.

41. *Ibid*, pp. 319-322.

42. Voir P. Bernays – A. Fraenkel, *Axiomatic,..., op. cit.*, pp. 61-64.

43. Voir W. O. Quine, *Set Theory and..., op. cit.*, § 6.

44. A. Fraenkel – Y. Bar-Hillel, *Foundations of Set Theory*, Amsterdam, North-Holland, 1959, pp. 145-147.

des propriétés, qu'en étant «trop compliquées», ne peuvent pas dénoter des multiplicités consistantes. Ce n'est pas la méthode de la théorie des types extensionnelle. Elle, elle a comme but de déterminer des conditions rigoureuses d'appartenance, selon les niveaux des types. En d'autres termes, elle s'occupe plutôt de la partie gauche de l'équivalence de l'axiome d'abstraction: $(\forall x)\,(\exists y)\,(x \in y \leftrightarrow \varphi(x))$. Une fois qu'on aurait une *syntaxe typée* complète, la partie droite ne nécessiterait aucune correction. Cela est garanti par l'équivalence: une bonne formule à gauche donne une bonne formule à droite, et si l'hierarchie des types est respectée, «$\varphi(x)$» devient stratifié. Si, quand même, notre effort avait été concentré à la partie droite, la partie gauche (pour exactement la même raison) serait automatiquement *typifiée*[45]. C'est exactement cela, ce à quoi a pensé Quine, en 1936, en formulant le système de l'article *New Foundations for Mathematical Logic*[46].

«I will now suggest a method of avoiding the contradictions without accepting the theory of types or the disagreeable consequences which it entails. Whereas the theory of types avoids the contradictions by excluding unstratified formulas from the language altogether, we might gain the same end by continuing to countenance unstratified formulas but simply limiting R3 [i.e.: φ *dans* $(\exists x)\,(\forall y)\,((y \in x) \leftrightarrow \varphi))$] explicitly to stratified formulas»[47].

Et dans *Set Theory:*

«The alternative departure from the theory of types that suggests itself, finally, is this: that we reconstrue unindexed variables as truly general variables (rather than as typical ambiguous), but yet keep the restriction on (i) that the particular formulas put for «*F*(x)» must be stratified. Such is the system called NF»[48].

La réalisation de ce projet est d'une importance extrême pour notre propos à cause du fait suivant. *Être identique à soi-même* se prouve parfaitement stratifié comme F(x), et tout à coup l'ensemble universel réapparaît. Puisque l'existence est, dans ce système, définie en tant qu'appartenance à l'univers, nous avons la formule suivante, qui exprime que l'univers est un ensemble: «$U \in U$». Contrairement aux systèmes avec *classe universelle*, l'univers est ici un *objet* de la théorie (dans le sens frégéen). Il est, tout de même, un objet bizarre. Voyons quelques unes de ses particularités: i) le théorème de Cantor y est non plus inapplicable, mais faux, ii) postuler même qu'il peut être ordonné selon un bon

45. Quine utilise le terme «stratified» pour les formules valides dans tous les deux côtés de l'équivalence. Nous allons utiliser le terme «stratifié» pour le côté droit et «typifée» pour le côté gauche.

46. W. O. QUINE, New Foundations for Mathematical Logic, *American Mathematical Monthly*, 44, 1937.

47. W. O. QUINE, *From a Logical Point of View*, Cambridge, Mass, 1961^2, p. 92.

48. W. O. QUINE, *Set Theory and…*, *op. cit.*, p. 288.

ordre est impossible[49]. Ces particularités, parmi d'autres, le rendent finalement un ensemble difficile à manier. On doit donc procéder ainsi. On remplace la distinction classique entre multiplicité de très grande taille (classe) et multiplicité cohérente (ensemble) de la stratégie de limitation de taille, par une tripartition: i) ensembles normaux, ii) ensembles (cantoriens) de très grande taille, iii) ensembles non cantoriens[50]. Dans les cas (i) et (ii) le théorème de Cantor est valable, et non pas applicable, respectivement. Dans le cas (iii), il est faux. Cette tripartition rend le système inélégant et difficile à manipuler autant que la taille augmente.

Par contre, et pour le mathématicien qui préfère travailler dans des contextes *typés,* il est bien plus usuel que la théorie ramifiée. Cela est dû au fait qu'on peut en déduire l'analyse classique sans s'occuper du tout avec les ensembles non cantoriens[51]. Pour le métathéoricien, quand même, qui doit faire constituer la *syntaxe* fixant quelles sont les «*F(*x)» stratifiés, le travail est bien plus laborieux[52]. Etant donné la brusque «mise entre parenthèse» des ensembles non cantoriens, le seul avantage indiscutable du système est l'élimination des indices de type. Cet avantage, tout de même, constitue un avantage essentiel seulement par rapport à la théorie des types ramifiée. En ce qui concerne la théorie simple, l'indexation est bien plus facile à traiter, et en ce qui concerne les *théories des ensembles,* elle est inexistante.

Tous ces paramètres ont conduit Quine à faire additionner au système les classes, ou (pour suivre son terme), les *classes ultimes.* «Ultimes» à cause du fait que Quine appelle (dans *Mathematical Logic*[53]) «classe» ce que d'habitude est appelé ensemble (set). Śinon, les *classes ultimes* coïncident avec les classes de la VN.B. Ce système généralisé est exposé dans *Mathematical Logic*, et codifié en tant que M.L. L'adjonction des classes dans le système a comme conséquence non seulement la disparition des ensembles non cantoriens, mais une convergence générale du formalisme (qui continue d'être d'inspiration *zig-zag*), vers les *théories des ensembles* avec classes. L'affiliation n'est pas accidentelle. Comme VN.B. constitue une généralisation de Z.F., M.L. est une génération de N.F. Ce qui est étonnant est que M.L. est plus proche par rapport à VN.B. que ne l'est N.F. par rapport à Z.F.[54].

L'univers, une fois que les classes lui ont été additionnées, il ne peut plus appartenir à soi-même, et donc, selon les critères de la théorie, il n'existe pas. «∪U» constitue une classe (ultime) et non pas un ensemble. Les paradoxes

49. *Ibid,* § 41.

50. A. Fraenkel – Y. Bar-Hillel, *Foundations of Set Theory,* Amsterdam, North-Holland, 1959, p. 163.

51. *Ibid,* pp. 163-164.

52. *Ibidem.*

53. W. O. Quine, *Mathematical Logic*, New York, Harper, 1962; (1ère éd. 1940).

54. W. O. Quine, *Set Theory and..., op. cit.*, § 44.

donc de Russell et de Cantor sont réfutés de manière presque identique à celle de la VN.B.

Contrairement à la N.F., où le paradoxe de Cantor a été démontré faux (pour les ensembles non cantoriens), dans le cas de la M.L. on l'empêche parce que «U $\notin$ U», et par conséquent U n'existe pas, et $\cup$U est une classe ultime, et alors exactement comme à VN.B.

En ce qui concerne le paradoxe de Russell, la ligne est pareille. Tandis que dans N.F., il est retenu parce que x $\notin$ x n'est pas stratifié, dans M.L., x $\notin$ x est stratifié (si x est un ensemble), mais la multiplicité définie n'appartient pas à $\cup$U, et détermine donc une classe ultime[55], et (si x est une classe)[56], $\{x: x \notin x\}$ n'appartient pas à U, et donc n'existe pas[57]; il est une classe virtuelle.

Doukas KAPANTAÏS
(Athènes)

ΟΙ ΘΕΩΡΙΕΣ ΣΥΝΟΛΩΝ ΚΑΙ ΤΟ ΣΥΝΟΛΟ ΤΩΝ ΟΝΤΩΝ

Περίληψη

Τὸ παρὸν ἄρθρο ἀποτελεῖ μέρος μιᾶς εὐρύτερης μελέτης σχετικῆς μὲ τὶς ὀντολογικοῦ περιεχομένου καταβολὲς τῶν δύο κύριων σχολῶν τοῦ συνολοθεωρητικοῦ λογισμοῦ. Στόχος εἶναι νὰ καταδειχθεῖ ὅτι, πίσω ἀπὸ τὶς λεπτομέρειες τοῦ λογισμοῦ αὐτοῦ, λανθάνει μία μεταφυσική. Ὁ προβληματισμὸς εἶναι ἐμφανὴς στὸν Φρέγκε καὶ στὸν Κάντορ. Τὸ ὅτι αὐτὸς ὁ λογισμὸς εἶναι θεμελιακὰ ὀντολογικὸς συνάγεται ἐπίσης ἀπὸ τὴν Ρασσελιανὴ «θεωρία τῶν τύπων». Τὸ πρωταρχικὸ ἐρώτημα εἶναι ποιὸ μπορεῖ νὰ εἶναι τὸ ἰσχνότερο πεδίο γιὰ τὴν δεσμευμένη μεταβλητή, καί, συνεκδοχικά, ποιὸς μπορεῖ νὰ εἶναι ὁ πιὸ οἰκονομικὸς φορμαλισμός. Οἱ ὀντολογικὲς καταβολὲς τῆς θεωρίας ἀρχίζουν νὰ γίνονται πιὸ δυσδιάκριτες ἀπὸ τὸν Zermelo καὶ ἔπειτα. Καὶ ἂν ἐκεῖ ἡ παρουσία τῶν *Urelemente* ἀπηχεῖ ἀκόμα τὶς ὀντολογικὲς καταβολὲς τῆς θεωρίας, ἀπὸ τὸν Fraenkel καὶ μετὰ αὐτὲς τοῦτες οἱ καταβολὲς εἶναι, ἐν πολλοῖς, ζητούμενο. Τὸ ἴδιο ἰσχύει καὶ γιὰ τὶς δύο θεωρίες τοῦ Quine, ἂν καὶ ἐκεῖ ἡ κατάφαση τῆς ὕπαρξης τοῦ καθολικοῦ συνόλου ἐπαναφέρει τρόπον τινὰ τὶς ὀντολογικὲς συζητήσεις.

Δούκας ΚΑΠΑΝΤΑΗΣ

55. Et donc comme dans VN.B.
56. Soit une classe, soit une classe ultime.
57. W. O. QUINE, *Set Theory and…, op. cit.*, p. 309.