CAHIERS DU SÉMINAIRE D'HISTOIRE DES MATHÉMATIQUES

GUSTAVE CHOQUET

Épistémologie du transfini

Cahiers du séminaire d'histoire des mathématiques 1re série, tome 1 (1980), p. 1-17
<http://www.numdam.org/item?id=CSHM_1980__1__1_0>

EPISTEMOLOGIE DU TRANSFINI*

par Gustave CHOQUET

Introduction.

Pour le mathématicien d'aujourd'hui, du moins pour le théoricien des ensembles, l'utilisation du transfini est devenue quotidienne ; le transfini est maintenant à la fois un instrument et un sujet d'étude en soi. Chacun des tomes d'une revue aussi importante que les Fundamenta Mathematicae[1] est peuplé d'un monde d'alephs qui semblent avoir, pour ceux qui les manipulent, une réalité aussi concrète que les fonctions analytiques pour les mathématiciens d'une autre génération. Bien mieux, l'intuition que ces théoriciens ont acquise, leur semble assez sûre pour qu'ils n'hésitent pas, comme W. Sierpinski, à partir d'hypothèses non encore démontrées[2], et à examiner très longuement leurs conséquences logiques.

Qu'il n'en ait pas toujours été ainsi, il suffit, pour s'en convaincre, de relire dans l'ouvrage d'Emile Borel, Leçons sur la théorie des fonctions[3], les discussions passionnées sur les problèmes que souleva, au début de ce siècle, l'introduction du transfini.

Que des notions d'allure purement mathématique n'aient pas été acceptées par certains mathématiciens et qu'elles le soient quasi-unanimement trente ans plus tard, peut sembler étonnant. On comprendra mieux une telle situation en se souvenant des crises analogues qui se sont produites dans l'histoire des mathématiques ; signalons seulement, pour mémoire, celles qu'ont suscitées l'introduction des infiniment petits, ou l'introduction d'éléments idéaux comme les nombres complexes.

De telles crises se produiront encore toutes les fois que des mathématiciens voudront préciser des notions d'un type nouveau tirées de leur intuition, que celle-ci soit d'origine physique, mathématique ou logique ; et qu'ils ne disposeront pas d'un langage adapté à la formulation mathématique de ces nouvelles notions.

D'ailleurs les mathématiques n'ont pas le privilège de telles crises. Les mêmes raisons expliquent des crises de la physique moderne, comme celles crées par le principe d'indétermination.

Je n'ai pas l'intention de retracer ici l'histoire de la dernière crise des mathématiques, causée par le développement rapide de la théorie des ensembles. L'examen des problèmes qu'elle a posés n'est pas achevé et ne le sera sans doute jamais. L'étude de ces problèmes a exigé une réédification de la logique, parallèle à celle des mathématiques. Ce travail de reconstruction s'est fait dans des directions vraiment très variées. Qu'il suffise de rappeler que les logiciens de cette reconstruction ont été jusqu'à -

---

* Ce texte est celui d'une conférence prononcée lors d'un Colloque de Logique et Philosophie en 1945, à Paris. Son intérêt ne saurait être actuellement qu'historique.

non pas mettre en doute - mais considérer la logique classique comme une à côté d'autres logiques qu'il est difficile de rejeter *a priori* : Le choix de l'une d'elles semble n'être qu'une question de goût et de commodité. La question de leur adéquation au monde physique est autre chose ; mais justement il semble que certaines s'adaptent mieux que la logique classique à certaines théories physiques.

Qu'il soit difficile à des mathématiciens,habitués à la logique classique, de raisonner en suivant une logique différente - et même de lui accorder une plus grande valeur - n'est pas étonnant. Sur un plan différent, c'est une réaction du même type qu'avaient les mathématiciens du XVIII^e siècle devant le schéma d'une géométrie non-euclidienne, et ils ne furent pleinement rassurés que lorsqu'ils surent construire dans le cadre euclidien des modèles non-euclidiens.

Les mathématiciens modernes ne seront, eux aussi, rassurés que lorsqu'ils auront pu construire dans le cadre de la logique classique des modèles isomorphes aux modèles proposés par les logiques nouvelles. C'est la possibilité d'un tel modèle pour la logique intuitionniste que nous montrera, je pense, M. Ky Fan, dans sa conférence.

Les règles du raisonnement dans ces nouvelles logiques n'ont en général été mathématisées (disons plutôt formalisées) qu'après de patients efforts. Les débuts de la logique intuitionniste en particulier ont été fort nébuleux et il reste difficile de suivre la pensée de ses premiers logiciens. La raison en est, je crois, qu'une logique n'a de sens que lorsqu'elle est formalisée ; essayer de donner un sens vulgaire à ses règles de déduction est une entreprise chimérique puisque notre langage quotidien n'est bien adapté qu'à la logique classique. Une telle traduction n'a en tout cas de chance que lorsque la théorie est déjà développée et approfondie.

La logique intuitionniste est née des difficultés de la théorie des ensembles, et pourtant les mathématiciens continuent à étudier les ensembles et à jongler avec les alephs sans avoir eu à modifier pour celà les règles de leur logique. C'est que la logique classique, cette fois encore, a paru suffisante aux mathématiciens pour continuer leur oeuvre mathématique. Pour lever leurs difficultés, il leur a suffi de partir de définitions plus claires, et de mieux préciser les axiomes de départ. Nées ensemble, la logique intuitionniste et la logique classique se sont dit, presque dès le départ, un fraternel adieu[4] - ou peut-être seulement un "au revoir" car même un mathématicien ne peut jurer de rien.

Je n'essaierai pas de construire une théorie du transfini dans le cadre intuitionniste. Revenant définitivement à la logique classique, j'esquisserai la manière dont on a assuré les fondements de la théorie du transfini ; puis j'examinerai quelle est actuellement la place du transfini, et les problèmes qu'il pose.

Fondements de la théorie du transfini.

Mot à mot, "transfini" signifie ce qui est au-delà du fini, ou de façon plus precise, ce qui n'est pas fini. En fait, l'étude du transfini est synonyme aujourd'hui de l'étude de cette partie de la théorie des ensembles qui traite des nombres cardinaux et des nombres ordinaux. Cette étude englobe donc aussi l'étude des ensembles finis.

Il a semblé tout d'abord plus simple d'étudier les ensembles finis que d'étudier les ensembles quelconques. Aussi a-t-on commencé, dans un effort pour ramener les notions fondamentales au minimum par essayer de définir les nombres entiers ordinaires.[5] Peano[6] cherche à définir les entiers en se basant sur le fait qu'ils forment un ensemble ordonné. Il part des idées primitives 0, le nombre, le successeur. Et voici ses propositions primitives : 1) 0 est un nombre. 2) Le successeur d'un nombre est un nombre. 3) Deux nombres différents ne peuvent avoir le même successeur. 4) 0 n'est le successeur d'aucun nombre. 5) Toute propriété qui appartient à 0, ainsi qu'au successeur d'un nombre qui possède cette propriété, appartient à tous les nombres.

Ces cinq propriétés marquaient une étape très intéressante dans la recherche des notions fondamentales. En particulier, la cinquième proposition qui constitue le principe de l'induction mathématique mettait en évidence que ce principe, d'apparence jusque là mystérieuse[7], n'était qu'une des propriétés qui permettent de définir l'ensemble des entiers.

Qu'il y ait des ensembles ordonnés qui vérifient les 4 premières propositions sans vérifier la 5[e] est évident ; il suffit de considérer l'ensemble $\{0,1,2,3,\dots\} \cup \{a_i,$ où $i =$ entier quelconque $\geqslant 0\}$, en convenant que a_{i+1} est le successeur de a_i, et que tout a_i est $\geqslant$ tout entier n. Il est immédiat que les 4 premières propriétés sont vérifiées, et non la cinquième.

Donc le principe d'induction mathématique est ici à juste titre considéré comme complétant la définition des nombres entiers.

Nous retrouverons une circonstance analogue dans l'étude des fondements de la théorie des ensembles.

On peut montrer que deux systèmes d'êtres mathématiques vérifiant ces 5 propositions sont isomorphes, c'est-à-dire qu'il existe entre eux une correspondance biunivoque conservant la relation "successeur". Mais, sans axiome assez fort, on ne sait pas montrer qu'il existe bien un système de nombres vérifiant ces 5 propositions.

D'autre part ces propositions nous donnent au fond seulement la notion de suite, mais aucunement la notion de nombre entier cardinal.

C'est la théorie des ensembles qui, parmi d'autres choses, donnera à la notion d'entier sa pleine valeur, par le recours aux ensembles finis.

Nous dirons tout à l'heure plus précisément ce qu'il faut entendre par le mot "ensemble" mais en anticipant un peu, pour n'avoir plus à revenir sur la question des nombres entiers, disons tout de suite qu'en partant de la notion d'ensemble et de celle de nombre cardinal qui s'en déduit, on obtient une définition très satisfaisante des entiers : Un nombre entier sera le nombre cardinal d'un ensemble fini. Voici la définition qu'a donnée Tarski, dans le tome 6 des Fundamenta Mathematicae[8], des ensembles finis : "L'ensemble A est fini, lorsqu'à toute classe K de ses sous-ensembles appartient comme élément au moins un ensemble B dont aucun vrai sous-ensemble n'appartient à K."

Dans son mémoire, Tarski étudie d'abord longuement (et en se basant sur les cinq axiomes de Zermelo, que nous énoncerons plus loin - ceux du choix et de l'infini étant exclus) les propriétés des ensembles finis à partir de cette définition. Il montre en particulier le théorème important : "Si A est un ensemble fini et B un vrai sous-ensemble de A, A n'a pas la même puissance que B." Et il remarque qu'on ne sait pas démontrer sans utiliser l'axiome du choix la réciproque de ce théorème.

Il est temps maintenant que nous abordions l'étude des fondements de la théorie des ensembles.

Zermelo met à la base de la théorie sept axiomes, qui comprennent en particulier le fameux axiome du choix, et l'axiome de l'infini qui assure l'existence d'un ensemble infini[9].

La nécessité d'énoncer explicitement les cinq premiers de ces axiomes n'a pas été ressenti tout de suite par les fondateurs de la théorie des ensembles[10]. On ne disait pas ce qu'est un ensemble, on se bornait à montrer quelques ensembles et à demander à l'esprit l'effort d'abstraction nécessaire pour concevoir la notion générale d'ensemble. Que cette méthode soit hasardeuse est suffisamment démontré par les polémiques qui ont divisé les premiers théoriciens de la théorie des ensembles.

Mais est-il donc nécessaire de savoir ce qu'est l'essence de la notion d'ensemble pour pouvoir raisonner à leur sujet ? Le physicien daltonien peut aussi bien qu'un autre étudier le rayonnement solaire dès qu'il connaît les lois du comportement des lumières rouge et verte à travers un prisme. Le mathématicien, pour raisonner sur les ensembles, n'aura besoin que de connaître les lois élémentaires qui régissent les ensembles.

Les axiomes de départ doivent être pour lui les seules choses qu'il sache sur les ensembles. Ils constituent l'inventaire exact des propriétés qu'ils possèdent.

A la base de la construction de Zermelo il y a des êtres mathématiques, entre certains desquels existent des relations : La relation d'égalité $a=b$ et la relation d'appartenance $a \in b$. Encore une fois il ne faut temporairement ajouter aucun sens intuitif à ces relations. Les axiomes vont consister à décrire les liens entre les êtres étudiés et les relations d'égalité et d'appartenance. (La négation de $a \in b$ se note $a \notin b$.)

En fait Zermelo admet comme intuitive la notion d'égalité. Aussi admet-il comme allant de soi que si $x \in z$ et $x=y$, on a $y \in z$.

Si l'on ne veut pas admettre comme intuitive la notion d'égalité, on posera comme axiome cette ligne.

Et voici maintenant les axiomes de Zermelo. Pour simplifier l'exposé, nous conviendront de dire que la relation $a\in b$ se lit : a est élément de b.

Axiome 1 ou axiome de l'identité des ensembles : Si deux ensembles a et b ont mêmes éléments, on a $a=b$. (Si pour $x\in a$ on a $x\in b$ et réciproquement.)

Axiome 2: α) Il y a un ensemble vide, c'est-à-dire un a tel qu'on n'ait pas $b\in a$, quel que soit b.

D'après l'axiome 1, l'ensemble vide est unique.

β) Pour tout a, il existe un b tel que pour tout $c\in b$, on ait $a=c$; cet être b est forcément unique et se note $\{a\}$ ou ensemble formé du seul élément a.

γ) Pour tout a, b donnés, il existe un c tel que si $d\in c$, on ait soit $d=a$, soit $d=b$.

Cet être c unique se note $\{a,b\}$.

Disons maintenant ce qu'est un sous-ensemble b d'un ensemble a : On note $b\subset a$ et on dit que b est inclus dans a, si pour tout $x\in b$ on a aussi $x\in a$.

Axiome 3 : a étant un être, chaque propriété des éléments de a définit un être b tel que $b\subset a$.

Axiome 4 : Pour tout a il existe un b dont les éléments sont les sous-ensembles de a. (Axiome de l'ensemble des parties.)

Axiome 5 : Pour tout a, il y a un b dont les éléments sont ceux qui appartiennent à l'un au moins des c tels que $c\in a$. (Axiome de la réunion d'une famille d'ensembles.)

A partir de ces cinq premiers axiomes beaucoup de résultats peuvent être obtenus. Nous avons tout à l'heure indiqué que Tarski en déduit des propriétés importantes des ensembles finis.

Il permettent aussi à Zermelo de résoudre l'un des paradoxes de la théorie des ensembles, né de la considération de l'ensemble des ensembles.

Zermelo montre en effet facilement qu'il n'existe pas (ou plus précisément qu'il est contradictoire qu'il existe),parmi les êtres qu'il étudie et qui satisfont aux cinq premiers axiomes, d'être a tel que tout être de sa collection appartienne à a. (Car si a est un tel ensemble, la relation $x\notin x$ définit une propriété des éléments de a, donc une partie $b\subset a$, et on ne peut avoir $b\in b$, ni $b\notin b$.)

Axiome 6 (axiome de l'infini) : Il y a un a tel que $\emptyset$ lui appartienne, et tel en outre que chaque fois que $b\in a$, on ait aussi $\{b\}\in a$. (D'où existence d'un ensemble a formé de $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, etc.

Axiome 7 du choix : Pour tout a tel que ses éléments soient sans éléments communs deux à deux et non vides, il existe dans leur réunion un sous-ensemble b qui a en commun avec chacun des éléments de a un élément et un seul.

Il reste dans la formulation des axiomes de Zermelo des éléments non satisfaisants. En particulier dans l'axiome 3, Zermelo introduit la notion de propriété des éléments de a, notion qui n'est certainement pas absolument claire.

On arrivera à une construction plus satisfaisante en formalisant complètement l'énoncé de ces axiomes, et en opérant en même temps une reconstruction de la logique. On trouvera les détails de cette reconstruction dans une très intéressante conférence de M. Henri Cartan, publiée dans la <u>Revue Scientifique</u> en 1943[11].

Il est possible maintenant d'introduire la notion de nombre cardinal. On définit d'abord l'égalité de puissances de deux ensembles. Les ensembles a et b ont même puissance lorsqu'il existe un c dont les éléments sont des couples $\{\alpha, \beta\}$ d'éléments de a et b, tels que tout élément de a (ou de b) entre dans la composition d'un couple et d'un seul.

On écrit alors $a \sim b$. Cette relation est réflexive, symétrique et transitive. Le nombre cardinal d'un ensemble est alors défini comme la classe des ensembles qui ont même puissance que cet ensemble.

Les opérations sur les nombres cardinaux se définissent ensuite de la façon classique.

On montre qu'il n'existe pas d'ensemble, de nombre cardinal plus grand que tous les autres, en montrant, par le procédé diagonal, que l'ensemble des parties d'un ensemble a une puissance supérieure à celle de l'ensemble.

Vient la question importante : De deux nombres cardinaux donnés, peut-on dire qu'ils soient égaux ou que l'un est inférieur à l'autre ?

Pour en décider, il faut faire appel à l'axiome de choix. L'axiome de choix permet en effet de montrer que tout ensemble peut être bien ordonné ; or il résulte de la théorie des ensembles bien ordonnés, édifiée par Cantor, que, de deux ensembles bien ordonnés, toujours l'un est un segment de l'autre. La comparabilité de deux nombres cardinaux quelconques en résulte et de plus on voit que les nombres cardinaux ordonnés par grandeur forment un ensemble bien ordonné. Le plus petit nombre cardinal infini se note $\aleph_0$, le second $\aleph_1$.

Il résulte du théorème de Zermelo, que le nombre cardinal $2^{\aleph_0}$, qui peut encore être défini comme la puissance de l'ensemble des parties d'un ensemble de nombre cardinal $\aleph_0$, a sa place dans l'ensemble bien ordonné des cardinaux.

La question de savoir si $\aleph_1 = 2^{\aleph_0}$ est une de celles qui ait le plus intrigué les mathématiciens.

Hilbert dans un célèbre mémoire, traduit dans les <u>Acta Mathematica</u> sous le titre <u>Sur l'infini</u>[12], affirme pouvoir le démontrer et indique effectivement le schéma de sa démonstration. Elle consiste essentiellement à considérer comme identiques l'ensemble des nombres réels et l'ensemble des fonctions arithmétiques d'entiers définies à partir de fonctions élémentaires par des récurrences générales.

Cette démonstration ne semble pas, à cause de cette affirmation, avoir reçu l'accord des mathématiciens ; du moins a-t-elle donné un vif essor à l'étude des fonctions arithmétiques définies par des récurrences.

Certains mathématiciens pensent qu'il faut considérer cette proposition comme un nouvel axiome.

Dans ce cas, c'est que la proposition rentrerait dans la catégorie des propositions ni vraies ni fausses lorsqu'on admet seulement les axiomes de Zermelo. Cette position semble excessive. Je cite Arnaud Denjoy qui, après avoir démontré un théorème d'allure très voisine de la proposition sur l'hypothèse du continu, et relatif à certaines ordinations du dénombrable, dit ceci[13] :

"Rien ne saurait mieux montrer combien il est artificiel de rejeter dans la métaphysique un sujet différant aussi peu de ceux que les mathématiques les plus certaines traitent immédiatement."

Je clos ici l'examen des problèmes classiques soulevés par le transfini pour passer à l'étude de questions moins connues. Je parlerai d'abord des nombres ordinaux constructibles.

Nombres ordinaux constructibles.

L'une des raisons des polémiques qui agitèrent il y a une trentaine d'années le monde mathématique était la difficulté que trouvaient les mathématiciens à s'entendre sur le sens de mots tels que "ensemble bien défini, nombres constructibles", etc. Plusieurs étaient d'accord pour limiter les mathématiques à l'étude des êtres définissables et clairement conçus ; seulement il ne l'étaient plus sur ce qu'il fallait mettre sous ces vocables. On a reconnu bien vite que toute borne imposée au champ des êtres mathématiques était vite dépassée et que les singularités dont on voulait exclure l'étude venaient s'imposer au coeur même des théories les plus classiques.

Mais les mathématiciens contemporains ont repris d'un autre point de vue la tentative de discerner dans une classe étendue d'êtres mathématiques ceux que l'on peut plus immédiatement concevoir. Ce n'est plus qu'ils veuillent nous imposer l'étude exclusive de ces sous-ensembles privilégiés. Mais le mathématicien ne se sent pleinement satisfait que lorsqu'il a pu mathématiser entièrement des notions d'abord très vagues nées de son intuition: telles que la constructibilité effective.

C'est l'effort fait dans ce sens en théorie des ordinaux que je vais tâcher d'exposer brièvement.

Il n'y a pas de mathématicien qui ne considère comme clairement concevables les premiers nombres transfinis de seconde classe. Tous sont même d'accord sur la notation à adopter : $1, 2, \ldots, n, \ldots, \omega, \omega+1, \ldots, \omega\times 2, \ldots, \omega^2, \ldots, \omega^2\times 2, \ldots, \omega^3, \ldots, \omega^\omega, \ldots, \omega^{(\omega^\omega)}, \ldots, \omega^{\left((\omega^\omega)^\omega\right)}, \ldots, \varepsilon_1, \ldots$.

Pour tous les nombres $< \varepsilon_1$, l'absence d'ambiguïté dans les notations vient d'ailleurs du fait suivant : Tout nombre ordinal $< \varepsilon_1$ peut se mettre d'une façon et d'une seule sous la forme : $(\omega^{\alpha_o} \nu_o + \omega^{\alpha_1} \nu_1 + \ldots + \omega^{\alpha_p} \nu_p)$, chacun des nombres α_i

étant effectivement inférieur au nombre ordinal donné (avec $\alpha_o > \alpha_1 > \dots > \alpha_p$, les ν_j étant finis et non nuls). Sur chaque α_i on peut recommencer la même décomposition. Si l'on remarque que toute suite strictement décroissante de nombres ordinaux est toujours finie, on voit que tout nombre $\alpha < \varepsilon_1$ peut, par un nombre fini de sommes, produits, exponentiations, s'exprimer au moyen des seuls ordinaux 1, ω .

D'autre part, la notation adoptée pour chacun de ces nombres permet de voir immédiatement s'il est de première ou de seconde espèce ; il est de première espèce si $\alpha_p \neq 0$, sinon il est de seconde espèce. S'il est de première espèce, on a aussitôt la notation pour l'ordinal qui le précède et, s'il est de seconde espèce, on a une suite canonique bien définie de type d'ordre ω dont ce nombre est la limite. C'est :

$\omega^{\alpha_o}\nu_o + \dots + \omega^{\alpha_{p-1}}\nu_{p-1} + \omega^{\alpha_p}(\nu_p - 1) + \omega^{(\alpha_p - 1)}\lambda$ où $\lambda = 1, 2, 3, \dots, n, \dots$ si α_p est de première espèce et $\omega^{\alpha_1}\nu_1 + \dots + (\nu_p - 1)\,\omega^{\alpha_p} + \omega^{\beta_i}$ où β_i est la suite canonique pour α_p si α_p est de seconde espèce.

Remarquons que j'ai pu assigner ici d'une façon univoque et suivant une règle simple un antécédent ou une suite d'antécédents à tout nombre $\alpha < \varepsilon_1$, mais un autre mathématicien pris dans la salle aurait peut-être fait ce choix univoque d'une autre façon.

En tout cas, une chose importante à retenir est que j'ai pu faire ce choix, et que j'ai pu donner une formule qui permet de trouver l'antécédent ou la suite d'antécédents de tout nombre $\alpha < \varepsilon_1$. En particulier, à tout nombre $\alpha < \varepsilon_1$ de seconde espèce j'ai pu attacher une suite canonique croissante ayant cet élément pour limite.

Les mathématiciens ont cherché si l'on pouvait plus généralement, à tout ordinal de la deuxième classe et de seconde espèce, attacher univoquement une suite canonique croissante ayant cet ordinal pour limite. On peut y parvenir pour des classes étendues d'ordinaux (c'est évident pour les ordinaux de la forme : $\alpha = \beta + \omega$) mais quel que soit le raffinement du procédé, il y a toujours des ordinaux qui y échappent.

D'autres mathématiciens ont alors cherché à résoudre simultanément ce problème et celui de la recherche d'une notation bien définie pour les ordinaux.

Nous avons vu que nous avons pu résoudre ce problème pour les ordinaux $< \varepsilon_1$. Pour les ordinaux $> \varepsilon_1$, les notations deviennent déjà moins universellement adoptées mais par exemple il est clair qu'on pourra toujours s'entendre, dans un groupe de mathématiciens, pour les notations de tous les nombres ordinaux $< \varepsilon_1^{\varepsilon_1^{\cdot^{\cdot^{\cdot^{\varepsilon_1^{\cdot^{\cdot^{\cdot}}}}}}}}$.

Est-il vrai que, pour tous les ordinaux < un ordinal α donné, les mathématiciens pourraient s'entendre pour trouver un système de notations aussi commode que celui adopté pour les nombres $< \varepsilon_1$? Ce n'est pas certain du tout. Nous voyons au contraire apparaître plus clairement la notion de ce qu'on va appeler ordinal constructif.

Voici la définition, encore volontairement vague, qui a été proposée par M.Alonzo Church (<u>Bull. Am. Math. Soc.</u>, April 1938)[14] :

"Un ordinal ξ est "constructible" si l'on peut trouver un système de notation qui assigne une notation unique à tout ordinal inférieur ou égal à ξ et, associés avec le système de notation, trois procédés effectifs, grâce auxquels respectivement : 1) Etant donnée la notation pour un ordinal on puisse déterminer si l'ordinal est de première ou de seconde espèce. 2) Etant donnée la notation pour un ordinal de première espèce on puisse trouver la notation pour son prédécesseur. 3) Etant donnée la notation pour un ordinal de seconde espèce, on puisse trouver une suite fondamentale pour cet élément - dans le sens d'un procédé effectif pour le calcul des notations des termes successifs de cette suite."

Nous devons remarquer que le système de notation associé à la donnée d'un ordinal constructible ξ varie avec cet ordinal.On voit donc comment M. Church a pu éviter les difficultés qui se présentaient dans la construction progressive d'un système de notation pour les ordinaux. Il ne procède pas de bas en haut en appelant constructible tout nombre que sa notation pourra atteindre. Il ne cherche pas un système de notation unique et valable pour tous les nombres constructibles.

Revenons à la définition de M. Church. Un seul point semble tout d'abord encore peu satisfaisant : C'est que cette définition implique la notion d'un "procédé effectif" sans qu'on sache ce qu'on entend par là. Qu'est-ce qu'un procédé effectif de calcul ? Diverses définitions ont été proposées, en particulier par Turing, Kleene et Church.

Pour A. M. Turing[15] un procédé de calcul est effectif si l'on peut imaginer pour faire le calcul une machine de type suivant : Elle comporte un nombre fini de pièces de dimension finie, et elle est traversée par un ruban infini, sur lesquels sont imprimés des symboles en nombre fini, qui sont les directives données à la machine ; lorsque la machine fonctionne, les résultats du calcul s'impriment sur le ruban.

Dans le cas du calcul d'un nombre réel, on a une définition équivalente à la précédente en disant qu'un nombre est constructible si l'on peut donner à un bureau de calcul des instructions en nombre fini qui permettent ensuite aux calculatrices - que l'on dispense de savoir résoudre un problème quel qu'il soit - de calculer ce nombre avec une approximation aussi bonne qu'on désire (les instructions restant les mêmes, indépendamment de l'approximation désirée).

Dans le cas d'un procédé qui, appliqué à un nombre entier, donne un autre nombre entier, une autre définition d'un procédé effectif consiste à dire qu'un procédé est effectif, s'il correspond à une fonction arithmétique qui soit récurrente au sens le plus général - cette récurrence devant être entendue au sens d'Herbrand et de Gödel[16].

Cette condition de récurrence pour la fonction arithmétique peut aussi être remplacée par une condition d'une autre forme introduite par M. Church dans sa formalisation de la théorie des fonctions d'entiers - ce qu'il appelle λ-formalisme. Cette condition revient à dire que la fonction envisagée doit être λ-définissable.

Ces trois définitions des procédés effectifs peuvent être appliquées à la définition des ordinaux constructibles. Il est remarquable alors que la classe des ordinaux constructibles est la même, que l'on adopte l'une ou l'autre des trois défintions de l'effectivité. C'est ce qui fait espérer à M. Church que la seule définition verbale des nombres constructibles, que nous avons donnée tout à l'heure, indépendamment du système formel de logique qui permet de la préciser, a une valeur en soi, et peut être considérée comme définissant sans ambiguïté la classe des nombres constructibles de deuxième classe.

Il est évident que tout ordinal inférieur à un ordinal constructible est constructible ; et on montre aisément que l'ensemble des ordinaux constructibles est dénombrable : cet ensemble forme donc un segment dans l'ensemble des ordinaux de deuxième classe. Soit ω_1 l'ordinal qui vient immédiatement après tous ceux là ; ω_1 est le plus petit des ordinaux non constructibles. Il joue, par rapport à l'ensemble de ceux-ci le même rôle que Ω par rapport aux ordinaux de la deuxième classe. On montre aisément que ω_1 est un nombre de seconde espèce en se basant sur le fait que si <u>a</u> et b sont constructibles, (a+b) l'est aussi. Donc ω_1 peut être défini par une suite fondamentale croissante de nombres constructibles, mais on n'a aucun moyen de calculer effectivement les termes d'une telle suite. (ω_1 est un nombre très complexe. Notons par exemple que ω_1 est forcément un nombre ε , c'est-à-dire tel que $\omega^x = x$, et même que c'est une solution de $\varepsilon_x = x$.)

On peut montrer que si a,b sont constructibles, (a+b) l'est aussi. On montre aussi que a.b , a^b le sont aussi ; il en est de même aussi de ε_a , ε_b (un nombre ε étant défini par la condition : $\omega^\varepsilon = \varepsilon$). Il en résulte en particulier que tous les ordinaux inférieurs à la première solution de $\varepsilon_x = x$ sont constructibles. On voit donc que les ordinaux constructibles de Church contiennent bien l'ensemble de ceux que l'on considère comme intuitivement constructibles.

Du fait qu'aucun des nombres $\geqslant \omega_1$ ne soit constructible ne veut pas dire que l'on ne puisse pas définir univoquement d'autres ordinaux de la deuxième classe : ω_1 en est le premier exemple, aussi ω_1^2 , etc.

L'intérêt de la construction de Church est avant tout d'avoir défini correctement la classe de ce que l'on pourrait appeler les "petits ordinaux", ou les ordinaux boréliens, en un mot ceux que l'on peut atteindre en atteignant en même temps tous ceux qui les précèdent.

Place du transfini en mathématiques. Peut-on l'éliminer des démonstrations ?

Je viens de parler de quelques grandes questions soulevées par le transfini et dont se sont occupé des mathématiciens de valeur incontestée. Le transfini occupe donc une grande place dans les mathématiques modernes. Mais cette place n'est-elle pas usurpée ? Les mathématiques ne pourraient-elles subsister sans leur utilisation ?

Si l'on se souvient que leur introduction en mathématiques n'a pas été le fruit d'une création artificielle, mais le résultat de l'évolution naturelle de la théorie des fonctions, on sera dès l'abord peu tenté de le croire.

Il y a d'abord toute une partie de la théorie des ensembles d'où le transfini n'a aucune chance d'être éliminé, parce que justement c'est lui qui fait l'objet de l'étude. Que l'on conteste toute valeur mathématique à l'étude du transfini en soi, cela est affaire de goût. Mais on ne peut empêcher personne de choisir le sujet de son étude, et les alephs sont devenus pour beaucoup de mathématiciens des réalités très concrètes.

Parmi les théorèmes dont l'objet véritable n'est pas l'étude du transfini, mais dans lesquels ce dernier intervient cependant, je distinguerai deux catégories : Il y a les théorèmes dans l'énoncé desquels le transfini intervient explicitement ; puis il y a ceux où le transfini n'intervient que dans la démonstration et que l'on peut espérer démontrer un jour sans son usage.

Considérons la question suivante : Recherche des conditions pour que deux ensembles fermés, dénombrables, situés sur le segment $[0,1]$ pour fixer les idées, soient homéomorphes (c'est-à-dire tels qu'on puisse établir entre eux une correspondance biunivoque et bicontinue). Dans l'énoncé du problème, le transfini n'intervient aucunement. Il peut être compris d'un mathématicien n'ayant aucune notion du transfini. Or la réponse est la suivante : Soient E, E' ces deux ensembles. Notons d'abord qu'à tout fermé dénombrable X de $[0,1]$ est associé un couple (β,p) où β est un ordinal de deuxième classe et p un entier $\geqslant 1$ tel que le dérivé X_β d'ordre β de X ait exactement p points ; on montre d'ailleurs que le couple (β,p) ainsi obtenu est absolument quelconque. Notons alors $(\alpha,n),(\alpha',n')$ les couples associés ainsi respectivement à E et E'. Eh bien, la condition d'homéomorphie cherchée peut s'exprimer par les deux égalités $\alpha=\alpha'$; $n=n'$.

Voici donc un problème dont l'énoncé ne fait pas intervenir le transfini, et dont la solution le fait obligatoirement intervenir.

Signalons encore l'exemple suivant, déjà souligné par Lebesgue dans ses Leçons sur l'intégration et la recherche des fonctions primitives[17]. Le théorème de Baire : "Toute fonction de classe 1 est ponctuellement discontinue sur tout ensemble parfait, et réciproquement" peut être démontré sans l'emploi du transfini. Lebesgue lui-même l'a montré. Mais le problème de Baire, à savoir : "Etant donnée une fonction ponctuellement discontinue sur tout ensemble parfait, trouver une série de fonctions continues dont elle est la somme" a été résolu par Baire en faisant explicitement intervenir le transfini ; son procédé opératoire ne saurait se passer du transfini. Cela ne veut pas dire d'ailleurs que d'autres procédés opératoires pour résoudre le problème de Baire ne puissent pas se passer du transfini, contrairement à ce que semble penser Lebesgue. Nous reparlerons de cette question tout à l'heure.

Il en est de même du procédé opératoire utilisé par Arnaud Denjoy pour définir la totalisation ; le transfini est au coeur même de son procédé, on ne saurait l'en éliminer. A. Denjoy a démontré en effet sur des exemples que toutes les étapes de sa totalisation peuvent être effectivement nécessaires.

Etudions maintenant la seconde espèce de théorèmes pour lesquels le transfini n'intervient que comme un outil au cours de la démonstration, pour disparaître entièrement dans l'énoncé du théorème. Peut-on, dans ce cas, remplacer le raisonnement par un autre ne faisant pas appel au transfini ? Il est aussi difficile de répondre à une question aussi générale, que de concevoir la forme de tous les problèmes que l'on pourra se poser en mathématiques.

Toutefois en restreignant suffisamment la classe des problèmes envisagés, M.Kuratowski a pu donner (dans le tome 3 des Fund. Math.) un procédé général d'élimination[18] des nombres transfinis dans les raisonnements mathématiques. Son procédé s'apparente à celui que Lebesgue et quelques autres mathématiciens avaient utilisé pour éliminer le transfini dans quelques cas particuliers. Mais il est plus général et plus simple.

J'insisterai un peu ici sur sa méthode. Cette méthode se rattache étroitement à l'idée de chaîne de Dedekind, développée par Zermelo dans la seconde démonstration du théorème qui porte son nom. Son intérêt philosophique vient de ce qu'elle permet de démontrer certains théorèmes d'un type très général , à partir des axiomes mis par Zermelo à la base de la théorie des ensembles, sans utiliser l'axiome supplémentaire indépendant sur l'existence des nombres transfinis.

Considérons les schémas suivants :

Schémas. E est un ensemble donné. Et G est une fonction qui, à tout sous-ensemble $X \subset E$, associe un ensemble $G(X) \subset X$.

A est un sous-ensemble de E. (Sa définition ainsi que celle de E et de G ne font pas appel aux nombres transfinis.)

On pose $A_0=A$; $A_{\alpha+1}=G(A_\alpha)$; $A_\alpha= \prod_{\xi<\alpha} A_\xi$ pour α de seconde espèce.

Soit $\mathcal{A}(A)$ la classe de tous les A_α.

Envisageons alors dans E toute classe Z qui vérifie les conditions :

1) Les éléments de Z sont des sous-ensembles de E.

2) $A \in Z$.

3) Si $X \in Z$ on a $G(X) \in Z$.

4) Pour toute famille (X_i) d'éléments de Z , l'intersection des X_i est un élément de Z .

Il y a de telles classes Z (exemple : la classe de tous les sous-ensembles de E).

La partie commune à toutes les classes Z est une classe Z . C'est donc la plus petite de toutes ces classes Z . Soit M (A) cette classe.

La méthode d'élimination des nombres transfinis consiste à remplacer,dans tout procédé représenté par le schéma ci-dessus,la définition de la classe A(A) par celle de M (A).

En effet, M. Kuratowski montre aisément que $A(A) \equiv M(A)$. Comme la définition de $M(A)$ n'a pas utilisé les nombres transfinis, on a bien remplacé la construction transfinie de A (A) par un procédé débarrassé du transfini.

M. Kuratowski a pu appliquer avec succès son principe général à la démonstration de plusieurs résultats classiques : C'est ainsi qu'il a résolu le problème de Baire sans faire usage du transfini.

Remarquons tout de suite que le passage d'une démonstration utilisant le transfini à une démonstration ne l'utilisant pas n'est pas toujours immédiat et la seconde démonstration reste souvent moins intuitive. La construction de Baire, faite pas à pas, même lorsque le chemin à parcourir est transfini, peut satisfaire davantage certains esprits.

Et avant tout il reste vrai que même ces théorèmes dont la démonstration peut se faire sans utiliser le transfini ont été découverts en l'utilisant. Même là où il n'est pas indispensable, le transfini peut donc être un outil précieux. La situation est analogue à celle présentée par la géométrie du compas. On sait qu'un grand nombre de constructions que l'on fait d'ordinaire en géométrie avec la règle et le compas peuvent être faites en utilisant simplement le compas. N'empêche qu'il sera presque toujours préférable d'utiliser d'abord règle et compas, quitte ensuite à examiner si le compas seul aurait suffi.

Nous venons d'étudier des cas où , indispensable ou non, le transfini intervenait comme un outil puissant de recherche et de démonstration.

Examinons maintenant un rôle non négligeable qu'il joue dans les mathématiques modernes, en topologie notamment : C'est son utilisation dans la construction de contre-exemples.Dans l'étude des espaces topologiques, un résultat prend souvent sa valeur du fait qu'il prolonge un résultat valable sous des hypothèses beaucoup plus restrictives, dans un espace cartésien, par exemple. Le chercheur en face d'un énoncé de théorème qu'il cherche à étendre à un espace très général se trouve pris entre deux alternatives : Démontrer le théorème ou trouver un exemple qui l'infirme. Ce lui est alors infiniment commode de pouvoir puiser dans l'arsenal des nombres transfinis pour construire un contre-exemple paradoxal. Sa construction n'aura peut-être aucune réalité, aux yeux d'un autre mathématicien, rejetant <u>a priori</u> la notion du transfini, mais il devra admettre cependant qu'elle démontre l'impossibilité de trouver une démonstration au théorème en question. Prenons un exemple simple . Il s'agit de savoir si le théorèmes suivant est exact : "Tout espace topologique de Hausdorff connexe, dont tout point possède un voisinage homéomorphe à un intervalle linéaire, est soit homéomorphe à une circonférence, soit homéomorphe à la droite réelle."

Or on montre qu'en plus des deux types topologiques d'espace indiqués dans l'énoncé du théorème il existe deux autres types topologiques d'espace vérifiant les conditions du théorème et deux seulement , ces deux types se construisant aisément en utilisant la totalité des nombres transfinis de la deuxième classe.

Voici donc un cas dans lequel il était indispensable, pour construire un contre-exemple mettant le théorème en défaut, d'utiliser le transfini.

Je terminerai cette revue rapide de l'utilisation du transfini en donnant l'énoncé d'un problème qui a l'intérêt de toucher une question qui suscita en son temps diverses polémiques, et de montrer comment la totalité des nombres transfinis de deuxième classe peut intervenir au coeur même des théories classiques.

En vertu d'un théorème bien connu de Paul du Bois-Reymond, il est impossible de trouver une suite de fonctions croissantes $f_1, f_2, \ldots, f_i, \ldots$ de type d'ordre α (α étant un nombre transfini de la deuxième classe), telle que pour toute fonction croissante f il y ait une fonction f_k de la suite qui croisse plus vite que f, c'est-à-dire telle que $\frac{f_k(x)}{f(x)} \to \infty$ avec x[19].

Or, l'application du théorème de Zermelo permet de montrer qu'il existe de telles suites transfinies, satisfaisant même à la condition supplémentaire que, si $\beta_2 > \beta_1$, la fonction f_{β_2} croît plus vite que f_{β_1}.

Le problème est le suivant : Peut-on imposer à une telle suite transfinie d'avoir pour type d'ordre Ω ? Si l'hypothèse du continu est vérifiée, la réponse est oui. Mais il est possible que l'on puisse répondre à cette question sans résoudre l'énigme de l'hypothèse du continu.

Me voici au terme de cette conférence. Je m'excuse auprès des philosophes d'avoir donné à ma conférence un aspect peut-être trop mathématique.

Mon but était de montrer quelle est, aux propres yeux du mathématicien qui l'utilise, la justification du transfini, et de montrer pourquoi les nombres transfinis sont devenus pour lui des outils usuels, comme d'autres plus traditionnels.

NOTES DE LA REDACTION

1 Le tome 1 de la revue Fundamenta Mathematicae paraît en 1920, avec S. Mazurkiewicz (1888-1945) et W. Sierpinski (1882-1969) comme rédacteurs. En 1945 paraît le tome 33, avec comme rédacteurs K. Kuratowski et W. Sierpinski.

2 C'est dans son livre Hypothèse du continu, Warszawa(Monogr. Matem., tome IV), que W. Sierpinski avait utilisé l'hypothèse du continu, et il écrit dans sa Préface (p.III) : "La question si l'ainsi dite hypothèse du continu est vraie ou non appartient aux problèmes les plus difficiles de la mathématique contemporaine. La présente monographie ne tend point à résoudre ce problème ; elle a pour but de faire connaître au lecteur les conséquences que l'hypothèse du continu implique."

Sur l'historique de cette question on peut consulter l'article de M. Guillaume : Axiomatique et logique, p.315-430 du tome II de l'Abrégé d'histoire des mathématiques, publié sous la direction de J. Dieudonné, Paris(Hermann), 1978. Voir, en particulier (p.418), la solution qu'a donné P. Cohen en 1963 à ce problème.

3 Il s'agit de la Note IV, parue dans la 2e édition des Leçons sur la théorie des fonctions de Borel, p.135-181, Paris(Gauthier-Villars), 1914, et intitulée : Les polémiques sur le transfini et sur la démonstration de M. Zermelo. Cette Note contient en particulier (p.150-160) les fameuses Cinq lettres sur la théorie des ensembles, écrites en 1904, de Baire, Borel, Hadamard et Lebesgue.

4 L'adieu n'a pas été, comme c'est souvent le cas en histoire des mathématiques, très fraternel. L. Kronecker (1821-1891), le premier intuitionniste, écrivait déjà en 1870 à H.A. Schwarz (P. Dugac, Eléments d'analyse de Karl Weierstrass, Archive for History of Exact Sciences, 10(1973), p.77) que "les conclusion de Bolzano sont évidemment de faux arguments". En 1885, il écrit encore au même correspondant (ibid. p.146) que,"s'il lui reste assez d'années et de forces", il espère montrer "la fausseté de toutes ces conclusions avec lesquelles travaille maintenant la soi-disant analyse". Cette "soi-disant analyse" est celle de Weierstrass, de Dedekind et de Cantor.

5 Ce fut R. Dedekind (1831-1916) - dans un essai qu'il a commencé en 1872 et publié en 1888 : Was sind und was sollen die Zahlen ? (Que sont et que représentent les nombres ?), Braunschweig(Vieweg) - qui, le premier, a tenté d'élaborer une théorie des nombres entiers, en partant de la notion primitive d'ensemble, basée sur le concept d'application.

6 G. Peano (1858-1932) a d'abord exposé sa théorie des nombres entiers dans Arithmetices principia novo methodo exposita, Turin(Bocca), 1889 = Opere scelte, tome II, Roma(Ed. Cremonese), 1958, p.20-55 = traduit en anglais par H.C. Kennedy dans G. Peano : Selected Works, p.101-134, Toronto(University Press), 1973. Puis dans Sul concette di numero, Rivista di Matematica, 1(1891), 87-102, 256-267 = Opere scelte, tome III, 1959, p.80-109. Sous la forme donnée ici, les "idées primitives" et les "propositions primitives" ont été exposées (p.216-217) dans le Formulaire de mathématiques, tome II,

Turin(Bocca), 1898 = Opere scelte, tome III, p.215-231 : Fondamenti dell'aritmetica.

7 Le livre de Dedekind, Was sind und was sollen die Zahlen ?, avait aussi pour but de donner à l'induction mathématique une base rigoureuse fondée sur la notion clé de la mathématique dedekindienne : la chaîne.

8 A. Tarski : Sur les ensembles finis, Fundamenta Mathematicae, 6(1924), 45-95. La définition des "ensembles finis" se trouve p.46, et le "théorème important" mentionné ici est énoncé p.72.

9 E. Zermelo (1871-1953) a formulé ces axiomes dans son article : Untersuchungen über die Grundlagen der Mengenlehre (Recherches sur les fondements de la théorie des ensembles), Math. Annalen, 65(1908), 261-281. Les sept axiomes sont énoncés p.263-267 sous une forme et dans un ordre différents de ceux présentés ici, présentation qui suit celle de l'article mentionné dans la note 11.

10 E. Zermelo écrit p.261 de son mémoire qu'il se propose d'y montrer comment la théorie des ensembles "créée conjointement par G. Cantor et R. Dedekind " peut se ramener à un petit nombre de définitions et à sept axiomes.

11 Il s'agit de l'article Sur le fondement logique des mathématiques, La Revue Scientifique, 81(1943), 3-11. H. Cartan y écrit p.3 : "Notre but est de montrer comment la logique peut servir de base à tout l'édifice des mathématiques, contrairement à l'opinion des "intuitionnistes". Qu'il soit bien entendu, dès le début, que nous nous plaçons ici uniquement au point de vue du mathématicien qui désire savoir si les fondements de sa science sont assurés et cherche à prendre une conscience exacte de la nature et de la portée des démarches qu'il accomplit lorsqu'il fait des mathématiques."

12 D. Hilbert (1862-1943) : Sur l'infini, Acta Mathematica, 48(1926), 91-122, traduit par A. Weil = Ueber das Unendliche, Math. Annalen, 95(1926), 161-190. Hilbert y écrit p.99 : "Grâce au gigantesque travail collectif de Frege, Dedekind et Cantor, l'infini fut finalement élevé sur le trône, et il connut son plus grand triomphe. L'infini, d'un vol audacieux, était parvenu à un succès vertigineux."

13 A. Denjoy (1884-1974) écrit encore dans son livre L'énumération transfinie, tome I, p.35, Paris(Gauthier-Villars), 1946 : "Peut-on admettre que nous entrions aussitôt dans le domaine de la pure métaphysique, parce que les ordinations ainsi particularisées sont les bonnes ordinations, dont chacune correspond à un nombre transfini de la seconde classe ? On ne manquera pas de se demander s'il est vraiment légitime de rejeter hors des mathématiques un ordre de sujets si étroitement mêlés à ceux que traite couramment l'analyse."

On peut trouver dans ce livre (p.XXV-XXXVII) une excellente bibliographie sur le transfini établie par G. Choquet.

14 A. Church : The constructive second number class, Bulletin of the Amer. Math. Society, 44(1938), 224-232. La définition en question se trouve p.226.

15 Sur A.M. Turing (1912-1954) voir p.411 et 429 de l'article de M. Guillaume cité dans la note 2.

16 Voir p.399-400, 407-412 et 422-423 de l'article de M. Guillaume.

17 Voir p.338 des <u>Leçons sur l'intégration et la recherche des fonctions primitives</u>, 2^{e} édition, Paris(Gauthier-Villars), 1950.

18 K. Kuratowski, <u>Une méthode d'élimination des nombres transfinis des raisonnements mathématiques</u>, Fund. Math., 3(1922), 76-108.

19 Sur ce théorème de du Bois-Reymond on peut consulter le livre de J. Cavaillès, <u>Philosophie mathématique</u>, Paris(Hermann), 1962, p.61-65.