Cantor et les nombres transfinis

par Jean-Pierre Belna
chargé de cours à l'université Paris VIII Saint-Denis
et à l'École Supérieure d'Électricité,
chercheur associé à l'équipe SPHERE-REHSEIS du CNRS
(Université Paris VII- Denis Diderot)

***Figure 1 :** Georg Cantor (1845-1918), vers 1894.*

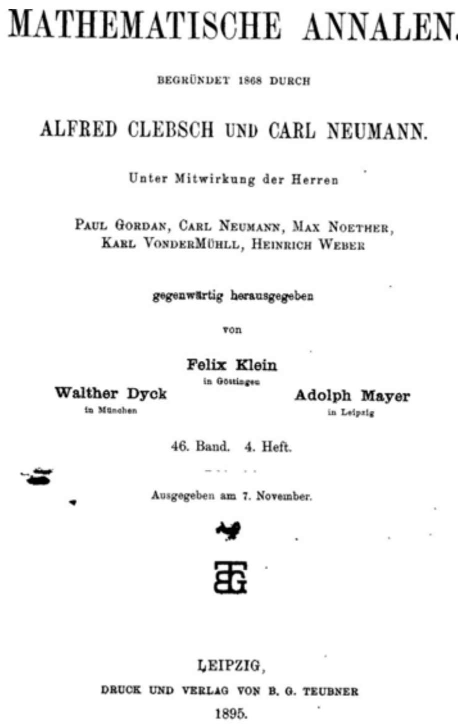
En 1895 et 1897, le mathématicien allemand Georg Cantor, né en 1845 à Saint-Petersbourg et mort à Halle en 1918, publie dans les *Mathematische Annalen* les deux parties d'un article intitulé *Contributions au fondement de la théorie des ensembles transfinis* (*Beiträge zur Begründung der transfiniten Mengenlehre*[1]). La fin du titre est trompeuse car il y expose surtout l'essentiel des résultats qu'il a obtenus sur les *nombres* transfinis, c'est-à-dire les nombres (cardinaux et ordinaux) que sa théorie permet d'associer aux ensembles infinis. C'est le dernier travail mathématique publié par Cantor, qui est, avec Richard Dedekind (1831-1916) et, dans une perspective autre, Gottlob Frege (1848-1925), le découvreur de la théorie des ensembles. Il y travaille depuis 1874,

1. Cantor aurait envisagé en 1899 la publication d'une troisième partie, consacrée au théorème du bon ordre et aux paradoxes de la théorie des ensembles (cf. section VII), mais le projet n'a pas abouti.

avec la particularité de s'intéresser spécialement à la notion de nombre transfini, introduite en 1883 et dont il tente de développer tous les aspects.

Si l'article est en deux parties, alors que la seconde est pour l'essentiel prête dès 1895, c'est notamment que Cantor espère un temps pouvoir résoudre un problème interne à sa théorie sur lequel il bute depuis longtemps, et dont on reparlera. Faute d'y parvenir, il se décide à publier cette seconde partie. L'article originel compte environ soixante-dix pages. Nous n'en avons retenu que les six premiers paragraphes (§1 à 6), les plus simples, qui constituent les débuts de la théorie cantorienne.

L'article de Cantor est très bien reçu par la communauté mathématique de l'époque, ce qui n'a pas toujours été le cas ; il est très vite traduit, en français notamment. La traduction française, supervisée par Jules Tannery (1848-1910), est due à Francisque Marotte (1873-1945), qui n'est pas un mathématicien de grand renom. Elle paraît en 1899 dans les *Mémoires de la Société des Sciences physiques et naturelles de Bordeaux*, avant d'être reprise sous forme de brochure la même année. Usant d'un vocabulaire mathématique daté, elle est surtout souvent très approximative. Nous avons donc choisi de proposer notre propre traduction du début de l'article.


MATHEMATISCHE ANNALEN.

BEGRÜNDET 1868 DURCH

ALFRED CLEBSCH UND CARL NEUMANN.

Unter Mitwirkung der Herren

PAUL GORDAN, CARL NEUMANN, MAX NOETHER,
KARL VONDERMÜHLL, HEINRICH WEBER

gegenwärtig herausgegeben

von

Felix Klein
in Göttingen

Walther Dyck
in München

Adolph Mayer
in Leipzig

46. Band. 4. Heft.

Ausgegeben am 7. November.

LEIPZIG,
DRUCK UND VERLAG VON B. G. TEUBNER
1895.

Figure 2 : La couverture des Mathematische Annalen *(1895)* contenant l'article de Cantor. *Cette revue (qui existe toujours) avait été fondée en 1868 par les mathématiciens allemands Alfred Clebsch (1833-1872) et Carl Gottfried Neumann (1832-1925). On relève aussi, parmi les précédents éditeurs, le nom de Félix Klein (1849-1925) (qui en avait été éditeur à partir de 1876).*

I - La théorie élémentaire des ensembles

La théorie des ensembles repose sur le concept d'ensemble, dont Cantor dit ceci :

> *Par « ensemble », nous entendons tout rassemblement M en une totalité d'objets m de notre intuition ou de notre pensée, déterminés et bien différenciés (qui seront appelés les « éléments » de M).*

Quel est le statut de cet énoncé ? L'ensemble étant le concept premier de la théorie des ensembles, il ne peut être défini à l'intérieur de celle-ci. Dans son article, Cantor ne fait pas de distinction explicite entre ce qui constitue une vraie définition, comme celle relative à la relation « plus petit que » entre deux nombres cardinaux, et les énoncés faisant intervenir des termes extérieurs à la théorie, dont la signification n'a pas été fixée auparavant. Certains mathématiciens préfèrent alors parler d'« explication », destinée à faire comprendre un mot nouveau par d'autres mots de la langue courante et dont le sens est supposé connu, comme c'est le cas ici.

Mais Cantor utilise justement des termes dont la signification peut faire débat, en particulier « intuition » et « pensée ». Sans développer ce point, on voit que pour lui, les ensembles nous sont donnés au travers d'un processus psychique ou mental. Par ailleurs, « rassemblement » peut sembler n'être qu'un synonyme d'« ensemble », d'où l'apparence de circularité de l'énoncé cantorien, dont il ressort :

1) qu'un ensemble est toujours une totalité, ce qui est fondamental lorsqu'il est infini ;
2) qu'un ensemble est constitué d'éléments, ce qui situe Cantor parmi ceux qui privilégient l'aspect extensionnel de la notion d'ensemble plutôt que son aspect intentionnel. Les premiers donnent la liste des éléments de l'ensemble, par exemple {2, 4, 6, 8, 10}, les seconds une propriété qui leur est commune, ici « entier naturel pair compris entre 2 et 10 » ;
3) qu'il est impossible pour Cantor de définir l'ensemble vide.

Il définit ensuite la réunion d'ensembles, toujours disjoints, et la partie ou sous-ensemble d'un ensemble donné, aujourd'hui partie ou sous-ensemble propre. S'il ne définit pas les relations d'appartenance et d'inclusion, il note que la relation « partie de » est transitive.

Le nombre cardinal d'un ensemble est le nombre de ses éléments, compte non tenu de l'ordre dans lequel ils sont donnés (dans le cas contraire, on parle de nombre ordinal). Cantor utilise aussi le mot « puissance », introduit en 1878. Il ne désigne pas alors un nombre à proprement parler mais sert à distinguer deux types d'ensembles infinis : après avoir montré en 1874 qu'il n'existe pas de bijection entre $\mathbb{N}$ et $\mathbb{R}$, Cantor appelle dénombrables ou ayant la puissance du dénombrable les ensembles en bijection avec $\mathbb{N}$, et continus ou ayant la puissance du continu ceux qui sont en bijection avec $\mathbb{R}$ [2].

$$\aleph_0$$

Figure 3 : La lettre ℵ (aleph), première lettre de l'alphabet hébraïque. *Elle a été choisie par Cantor pour symboliser les cardinaux d'ensembles infinis (cf. ci-après). $\aleph_0$, ici représenté, est le cardinal de l'ensemble des entiers naturels. Ce qui suit, comme l'article de Cantor, trouve bien évidemment son principal intérêt dans le cas des cardinaux d'ensembles infinis.*

Cardinaux, ordinaux et exemples d'ensembles

On peut trouver d'autres exemples que les simples nombres (entiers, rationnels, réels,...) pour illustrer les notions de cardinaux transfinis.
Définissons sur N × N (les couples d'entiers naturels positifs) la relation d'ordre suivante :

$$(0,0) \triangleleft (0,1) \triangleleft (0,2) \triangleleft (0,3) \triangleleft \ldots \triangleleft (1,0) \triangleleft (1,1) \triangleleft (1,2) \triangleleft (1,3) \triangleleft \ldots$$

L'élément (1,0) est le plus petit à venir après une infinité d'éléments : on convient d'indexer sa position par ω ; le couple (1,1) vient juste après et est indexé par (ω +1) ; le couple (2,0) est indexé par $\omega \times 2 = \omega + \omega$, etc.
L'ensemble peut donc être indexé (énuméré) comme suit :
1, 2, 3, ...n,... ω, ω +1, ω + 2,..., ω + n,..., $\omega \times 2$, $\omega \times 2$ + 1, ..., $\omega \times 3$,...
ω est appelé un ordinal – ceci correspond à la notion d'adjectif ordinal : (0,0) occupe la première place (*première* est un adjectif ordinal), (n,0)

2. Cette démonstration, qui est à l'origine de la théorie cantorienne des cardinaux transfinis, n'est pas reprise dans l'article de 1895-1897. Il n'est fait qu'allusion à la distinction entre dénombrable et continu dans la démonstration de la formule (11) du § 4 (cf. section IV). C'est que Cantor n'a pas de solution au problème crucial du continu dont nous parlons dans la section VII ci-après.

la $(n+1)^{\text{ème}}$ place (ordinal aussi), et (1,0) occupe la place ω (ordinal par extension).
L'adjectif numérique cardinal, quant à lui, ne prend pas d'ordre en considération : un ensemble à un élément (cardinal 1), … à *n* éléments (cardinal *n*), un ensemble infini à $\aleph_0$ éléments, etc. Le cardinal de l'ensemble ci-dessus ($\mathbb{N} \times \mathbb{N}$) est $\aleph_0$ (il est équipotent à $\mathbb{N}$).

Pour définir le nombre cardinal d'un ensemble, Cantor demande de procéder à un double acte d'abstraction, ce qui explique le symbole qu'il utilise, aujourd'hui remplacé par « *card* » : d'abord de l'ordre dans lequel sont donnés les éléments (pour lui, les éléments d'un ensemble sont naturellement donnés dans un certain ordre), ensuite de ce qui caractérise leur nature. Et d'expliquer que cela résulte d'un processus mental, psychologique ou intellectuel, lié à « notre faculté active de pensée ». Le vocabulaire est contestable et l'idée peu aisée à mettre en œuvre. Il faut comprendre qu'une fois effacé leur ordre et éliminée leur nature, les éléments de l'ensemble ne sont plus que ce que Cantor appelle des « uns », dont il ne reste que le nombre. Soit, par exemple, quatre vaches, données dans l'ordre croissant de leur poids, quatre pays, donnés par ordre alphabétique et quatre personnes, données par ordre de taille croissante. Faisons abstraction, dans chaque cas, de l'ordre donné puis de leur nature respective de vaches, de pays et d'individus : chaque ensemble n'est plus composé que de quatre « uns ».

Mais que sont, d'un point de vue mathématique, ces « purs uns » censés exister « dans notre esprit comme image intellectuelle ou projection de l'ensemble donné » ? Rendus indiscernables, puisque plus rien ne permet de les distinguer, ils doivent pourtant être différents pour pouvoir être « comptés » : Cantor utilise les expressions « différents uns » et « un particulier » pour démontrer le théorème *(8)* (si deux ensembles ont le même cardinal, ils sont équipotents).

Ses explications montrent qu'il tient à faire du nombre cardinal d'un ensemble un autre ensemble, qui lui est équipotent (théorème *(9)* : *M* est équipotent à l'ensemble card *M*), car il veut assurer à l'ensemble et à son cardinal un statut identique, ce qui est à ses yeux capital dans le cas infini : les deux sont des « objets » intellectuels.

Ce que Cantor expose un peu avant aurait pu lui fournir une définition plus simple. Il définit la relation d'équipotence, qu'il appelle « équivalence », entre ensembles : deux ensembles sont équipotents s'il existe une bijection de l'un sur l'autre, c'est-à-dire une application (une « correspondance » ou une « loi de correspondance », dit-il) de l'un sur l'autre et réciproquement, ce qu'il appelle une « correspondance biunivoque ». Mais bien que Cantor montre que la relation d'équipotence entre deux ensembles est une relation d'équivalence et en fait « la condition nécessaire et suffisante de l'égalité de leurs nombres cardinaux », il ne s'en sert pas, contrairement à certains de ses contemporains, pour définir le nombre cardinal comme une classe d'équivalence d'ensembles en bijection mutuelle : 1, par exemple, peut être défini comme la classe d'équivalence de tous les singletons, tous nécessairement en bijection. D'une manière générale, Cantor répugne à utiliser le procédé qui, à partir d'une relation d'équivalence sur un ensemble d'objets, en définit de nouveaux, à savoir les éléments de l'ensemble quotient correspondant[3].

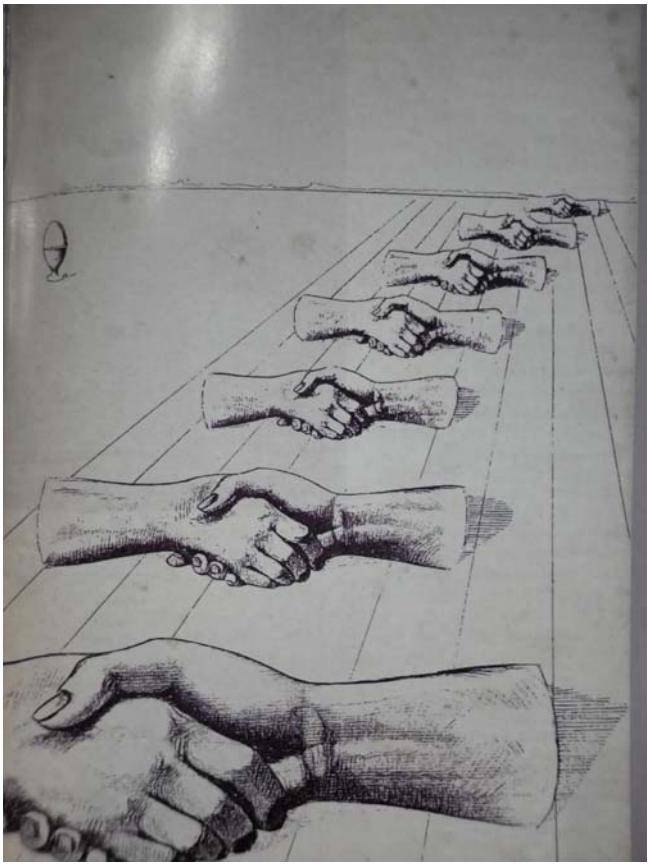


***Figure 4 :* Une vision de la bijection par le peintre et dessinateur Max Ernst (1891-1976)** *[Poèmes visibles, dessins de Max Ernst dans la revue* Le Minotaure*, 1934-1936, in D. Guedj,* L'Empire des nombres*, Gallimard, 1996, p. 118-119, ©AGADP].*

3. Étant donnés un ensemble *E* et une relation d'équivalence $\mathfrak{R}$ sur cet ensemble, l'ensemble quotient $E/\mathfrak{R}$ est l'ensemble des classes d'équivalence de *E* pour la relation $\mathfrak{R}$.

II - LE THÉORÈME DE CANTOR-BERNSTEIN

Au § 2 de son article, Cantor définit une relation d'ordre sur les cardinaux, qu'il désigne du même signe que pour les entiers naturels, sans vérifier qu'elle en est bien une extension. La définition elle-même ne pose pas de difficulté particulière puisque Cantor prend soin de préciser que le résultat ne dépend pas des ensembles choisis pour « représenter » les nombres cardinaux donnés.

Suivent trois propositions, dont voici les démonstrations, que Cantor ne donne qu'en partie. M et N sont deux ensembles tels que $a = \text{card } M$ et $b = \text{card } N$, $a < b$ est défini par deux conditions :

1°) il n'y a aucune partie de M qui soit équivalente à N,
2°) il y a une partie N_1 de N telle que $N_1 \sim M$.

<u>Première proposition</u> : $a < b$ exclut $a = b$. Si $a = b$, alors $M \sim N$ et il existe une bijection f telle que $N = f(M)$. Il y aurait alors une partie M_1 de M, à savoir la partie $f^{-1}(N_1)$, telle que $M_1 \sim N_1 \sim M$ (d'après 2° ci-dessus) $\sim N$, ce qui contredit la condition 1°) ci-dessus de la définition de $a < b$.

<u>Deuxième proposition</u> : $a < b$ exclut $b < a$. Soient 1°) et 2°) (cf. ci-dessus) les conditions satisfaites par $a < b$ et 1') et 2') celles satisfaites par $b < a$:

- condition 1) sur M et N : $\forall M_1$ *t.q.* $M_1 \subset M$, $\neg (M_1 \sim N)$[4],
- condition 2) sur M et N : $\exists N_1$ *t.q.* $N_1 \subset N$ et $N_1 \sim M$,
- condition 1') sur N et M : $\forall N_1$ *t.q.* $N_1 \subset N$, $\neg (N_1 \sim M)$,
- condition 2') sur N et M : $\exists M_1$ *t.q.* $M_1 \subset M$ et $M_1 \sim N$.

1') contredit 2) et 2') contredit 1).

<u>Troisième proposition</u> : Si $P_1 \subset P$, de $a < \text{card } P_1$ suit $a < \text{card } P$. On a $\text{card } M < \text{card } P_1$. Les conditions qui définissent $\text{card } M < \text{card } P$ sont satisfaites :

a. par définition, aucune partie de M n'est équipotente à P_1, et donc *a fortiori* à P (s'il existait une partie M_1 de M équipotente à P, il existerait aussi une partie M_2 de M_1, et donc de M, équipotente à P_1) ;
b. par définition aussi, il existe une partie de P_1, donc de P, équipotente à M.

@@@@@@@

4. La notation de logique mathématique $\neg(M_1 \sim N)$ signifie qu'on nie la proposition $(M_1 \sim N)$.

Ce que Cantor explique ensuite est particulièrement intéressant. S'il est exact que « chacune des trois relations » entre cardinaux a et b « exclut les deux autres », il n'est pas possible, à ce « stade » de sa théorie, comme il le dit, de prouver la comparabilité universelle des cardinaux, c'est-à-dire le théorème A selon lequel pour « deux nombres cardinaux quelconques a et b, l'une des trois relations [$a = b$, $a < b$ ou $b > a$] doive nécessairement être vérifiée ». Cantor prévoit de revenir plus tard sur la démonstration de ce théorème, mais il ne le fera jamais.

Est alors présentée comme une conséquence de ce théorème A le théorème B :

> *Si deux ensembles M et N sont tels que M est équivalent à une partie N_1 de N et N à une partie M_1 de M, M et N sont aussi équivalents.*

Ce théorème porte désormais le nom de théorème de Cantor-Bernstein (parfois de Bernstein-Schröder). En 1883, Cantor l'énonce pour la première fois[5], sans preuve, sous la forme du théorème C :

> *Si M_1 est une partie d'un ensemble M, M_2 une partie de l'ensemble M_1 et si les ensembles M et M_2 sont équivalents, M_1 est aussi équivalent aux ensembles M et M_2.*

Schröder donne en 1896 la première démonstration, en partie fautive, de (B), avant que Bernstein n'en publie la première preuve correcte en 1898. En réalité, Dedekind en avait déjà trouvé une en 1887, mais sans, curieusement, la publier alors. La démonstration ci-dessous ne fait pas appel au théorème A, mais au théorème C, qui en est donc un lemme[6].

Démonstration du théorème de Cantor-Bernstein

Démonstration du théorème B à partir du théorème C, qui dit que si $M_2 \subset M_1 \subset M$ et $M \sim M_2$, alors $M_1 \sim M$ et $M_1 \sim M_2$.

Les hypothèses du théorème B sont $M_1 \subset M$ et $N_1 \subset N$, de façon que $M \sim N_1$ et $N \sim M_1$. Il existe donc une bijection de N sur M_1, et donc une bijection de N_1 sur une partie propre de M_1, que nous appelons M_2. On a donc $M_2 \subset M_1 \subset M$ et $M_2 \sim N_1 \sim M$. Le théorème C permet de conclure

5. G. Cantor, « Über unendliche lineare Punktmannigfaltigkeiten », *Mathematische Annalen* 21, 1883, p. 545-586 ; repris sous la forme d'un ouvrage dont le titre est : *Grundlagen einer allgemein Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen*, Leipzig, Teubner, 1883.
6. La démonstration est empruntée à E. Kamke, *Theory of Sets*, trad. F. Bagemihl, New York, Dover Publications, 1950, p. 23-25.

que $M \sim M_1$. Comme $M_1 \sim N$, on a aussi $M \sim N$. Le théorème B est donc établi à partir du théorème C.

<u>*Démonstration du théorème C :*</u>
Les hypothèses de (C) sont $M_2 \subset M_1 \subset M$ et $M \sim M_2$ (on verra que ces hypothèses font de M un ensemble infini). Posons $A = M_2$, $B = M_1 - M_2$, $C = M - M_1$ (d'où $A \cup B = M_1$ et $A \cup B \cup C = M$).
Le théorème (C) s'écrit alors sous la forme (C*) :

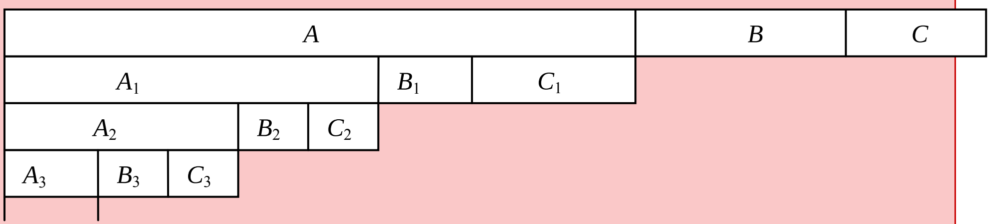
Si A, B et C sont des ensembles disjoints, alors $(A \cup B \cup C) \sim A$ entraîne $(A \cup B \cup C) \sim (A \cup B)$

Puisque $(A \cup B \cup C) \sim A$, il existe une bijection f de $(A \cup B \cup C)$ sur A. Posons $A_1 = f(A)$, $B_1 = f(B)$, $C_1 = f(C)$. Comme A, B et C sont des parties de $(A \cup B \cup C)$, et que ce dernier ensemble est en bijection avec A, on a $A_1 \subset A$, $B_1 \subset A$ et $C_1 \subset A$. D'où :

(1a) $A_1 \cup B_1 \cup C_1 = f(A) \cup f(B) \cup f(C) = f(A \cup B \cup C)$[7] $= A$,

(1b) $A \sim A_1$, $B \sim B_1$, $C \sim C_1$,

(1c) A_1, B_1 et C_1 disjoints (car si A_1 et B_1, par exemple, avaient un élément commun x, il serait l'image par f d'un élément de A et d'un élément de B, nécessairement distincts puisque A et B sont disjoints ; comme $A_1 \subset A$, x est un élément de A et serait donc l'image par f de deux éléments distincts de $A \cup B \cup C$, ce qui contredit le fait que f est une bijection $A \cup B \cup C$ sur A).



Puisque f est une bijection de A sur A_1, elle applique, comme précédemment, A_1, B_1 et C_1 sur des sous-ensembles A_2, B_2 et C_2 de A_1 [*cf*. schéma ci-dessus]. D'où :

(2a) $A_2 \cup B_2 \cup C_2 = A_1$,

(2b) $A_1 \sim A_2$, $B_1 \sim B_2$, $C_1 \sim C_2$,

(2c) A_2, B_2 et C_2 disjoints.

Du fait que $A \sim A_1 \sim A_2 \sim \ldots$, on peut itérer ce processus autant de fois qu'on veut, et on a en particulier : $C \sim C_1 \sim C_2 \sim \ldots$.

Posons alors D $= A \cap A_1 \cap A_2 \cap \ldots$.
Comme $B_1 \cup C_1 = A - A_1$, $B_2 \cup C_2 = A_1 - A_2$, $B_3 \cup C_3 = A_2 - A_3$, ...,

$$D \cup B_1 \cup C_1 \cup B_2 \cup C_2 \cup \ldots = D \cup (A - A_1) \cup (A_1 - A_2) \cup (A_2 - A_3) \cup \ldots$$
$$= D \cup A$$

7. $f(A) \cup f(B) \cup f(C) = f(A \cup B \cup C)$ est une propriété des bijections.

$= A$ (car $D \subset A$)

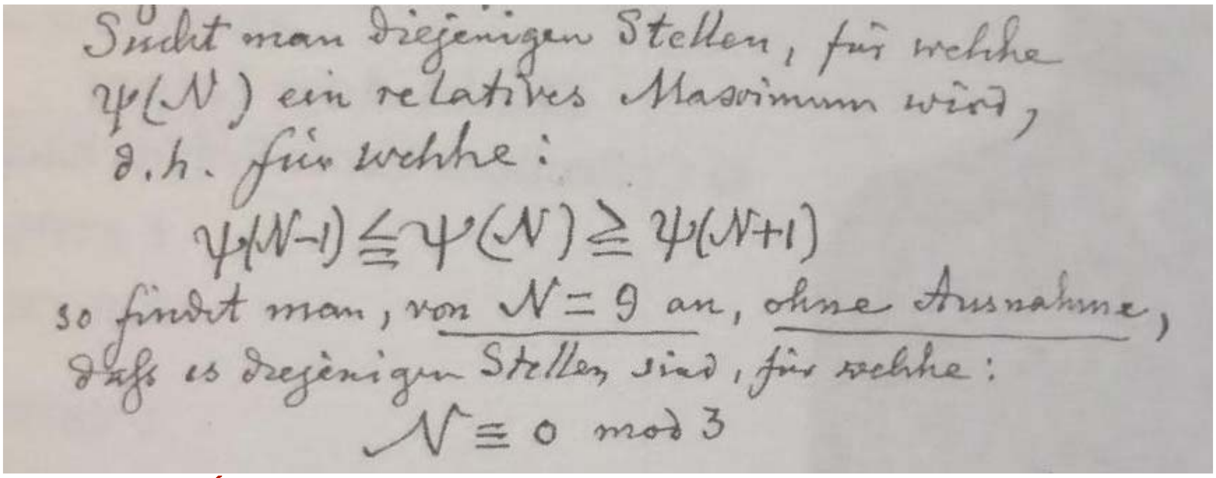
Puisque $A = D \cup B_1 \cup C_1 \cup B_2 \cup C_2 \cup \ldots$, on peut écrire :
$A \cup B \cup C = D \cup B \cup C \cup B_1 \cup C_1 \cup B_2 \cup C_2 \cup \ldots$, d'où :
$A \cup B = D \cup B \cup B_1 \cup C_1 \cup B_2 \cup C_2 \cup \ldots$
Comme $C \sim C_1 \sim C_2 \sim \ldots$, il vient $(A \cup B \cup C) \sim (A \cup B)$.
On a donc prouvé le théorème (C*), donc le théorème (C) et par conséquent le théorème (B).

III - Les opérations sur les nombres cardinaux

Cantor définit trois opérations sur les nombres cardinaux : l'addition, la multiplication et l'exponentiation (soustraction et division font problème dans le cas transfini). Là encore, il prend soin de préciser que le résultat ne dépend pas des ensembles choisis pour définir l'opération en question. L'addition et la multiplication de deux cardinaux résultent respectivement de la réunion de deux ensembles disjoints et de leur produit cartésien. Cantor, qui n'utilise pas le mot « couple », appelle cet ensemble « ensemble-liaison ». Il donne une seconde définition du produit de deux cardinaux, déjà proposée en 1887, évidemment équivalente à la première : elle consiste à remplacer le couple (m, n) par l'élément m_n, où m et n sont deux éléments quelconques de, respectivement, M et N. Par exemple, si $M = \{a, b, c\}$ et $N = \{1, 2\}$, on remplace 1 par $\{a_1, b_1, c_1)$ et 2 par $\{a_2, b_2, c_2\}$, d'où $S = \{a_1, b_1, c_1, a_2, b_2, c_2\}$. Les explications de Cantor semblent bien compliquées pour quelque chose qui nous paraît très simple, mais il est le premier à introduire la notion de produit cartésien de deux ensembles. Si Cantor utilise celle-ci pour définir le produit de deux cardinaux, c'est qu'elle lui permet de définir l'exponentiation dans le cas d'un exposant transfini. Lorsque b est fini, a^b est égal au produit de b facteurs égaux à a, mais, lorsque b est infini, cette définition est inutilisable. Cantor veut une définition générale, ce qui n'est possible qu'en partant de celle donnée précédemment pour le produit. À cet effet, il utilise la notion d'application d'un ensemble N sur un ensemble M, qu'il assimile à une correspondance entre N et M et appelle « recouvrement » de N par M. Ce terme a aujourd'hui une autre signification, mais la notion est alors nouvelle et le vocabulaire pas encore fixé[8]. Pour un lecteur d'aujourd'hui, celui de

8. Aujourd'hui, un recouvrement $\mathfrak{R}$ d'un ensemble E est un sous-ensemble de $\mathfrak{P}(E)$ tel que $E = \bigcup_{D \in \mathfrak{R}} D$. Le terme allemand actuel pour « application », qu'on trouve à l'époque chez Dedekind, est « *Abbildung* ».

Cantor est relâché puisque l'application désigne tantôt la loi elle-même – « par recouvrement [...], nous entendons une loi par laquelle... » –, tantôt l'image de l'ensemble – « le recouvrement correspondant de N sera appelé $f(N)$ » –, voire l'image d'un élément – « l'élément de M lié à n est en quelque sorte une fonction univoque de n et peut, par exemple, être désigné par $f(n)$ ». L'exponentiation ou élévation à une puissance est classiquement définie à l'aide de l'ensemble des applications de N sur M, que Cantor note $(N \mid M)$: si a = card M et b = card N, a^b = card $(N \mid M)$.



Sucht man diejenigen Stellen, für welche
$\psi(N)$ ein relatives Maximum wird,
d.h. für welche:
$$\psi(N-1) \leqq \psi(N) \geqq \psi(N+1)$$
so findet man, von $N = 9$ an, ohne Ausnahme,
daß es diejenigen Stellen sind, für welche:
$$N \equiv 0 \mod 3$$

Figure 5 : Écriture manuscrite de Cantor, lettre au mathématicien français Charles Hermite.

IV - Une bijection entre $\mathbb{R}$ et l'ensemble des applications de $\mathbb{N}$ sur $\{0, 1\}$

Alors que Cantor n'a pas encore parlé spécifiquement des cardinaux transfinis, il les utilise pour montrer l'intérêt de ses définitions, prouvant que si c est le cardinal de l'intervalle [0, 1] de $\mathbb{R}$ et $\aleph_0$ celui de $\mathbb{N}$, alors $c = 2^{\aleph_0}$, ce qui revient à montrer que [0, 1] et l'ensemble des applications de $\mathbb{N}$ sur {0, 1} sont équipotents. Avant d'exposer sa démonstration, observons que $\mathbb{R}$ est équipotent à]0, 1[et donc à [0, 1] : l'application de]0, 1[vers $\mathbb{R}$ définie par $x \mapsto \dfrac{1-2x}{x^2-x}$ est une bijection. Cantor prouve donc que $\mathbb{R}$ et $(\mathbb{N} \mid \{0, 1\})$ sont équipotents, mais le fait de ne considérer que les réels de partie entière égale à 0 simplifie la démonstration.

La démonstration de Cantor

Comme le montre sa formule (12), Cantor utilise la mise sous forme binaire ou développement dyadique des réels x de l'intervalle X = [0, 1] :

$$x = \frac{f(1)}{2} + \frac{f(2)}{2^2} + \ldots + \frac{f(\nu)}{2^\nu} + \ldots \quad [\text{où } f(\nu) = 0 \text{ ou } 1]$$

soit $x = b_1 b_2 \ldots b_\nu \ldots$, où les b_ν sont égaux à 0 ou à 1.
$2^{\aleph_0}$ est le cardinal de l'ensemble des applications de $\mathbb{N}$ dans {0, 1}, soit, puisqu'un développement dyadique est de fait une application de $\mathbb{N}$ dans {0, 1}, celui de l'ensemble $\mathfrak{D}$ de tous les développements dyadiques des réels x de X = [0, 1].

La difficulté est qu'il n'y a pas bijection entre [0, 1] et $\mathfrak{D}$. La plupart des réels de [0, 1] ont un unique développement dyadique, sauf les rationnels dont le numérateur est impair et le dénominateur égal à une puissance de deux, qui ont deux développements dyadiques. 0,5, par exemple, a un double développement dyadique : 10000… et 01111… ($1/2 = 1/2^2 + 1/2^3 + \ldots + 1/2^n + \ldots$)[9]. Le premier développement est dit « propre », le second « impropre ».

Il y a donc surjection de $\mathfrak{D}$ sur [0, 1] mais pas injection. Pour contourner l'obstacle, Cantor fait remarquer que l'ensemble des développements dyadiques « impropres », qu'il note $\{s_\nu\}$, est dénombrable, afin de montrer que sa cardinalité n'influe pas sur le résultat. Puisque D est égal à l'union de l'ensemble des développements dyadiques propres, identique à X = [0, 1], et de $\{s_\nu\}$, on a $\mathfrak{D} = X \cup \{s_\nu\}$, d'où $2^{\aleph_0} = \text{card } (X \cup \{s_\nu\})$.

Reste à prouver que X et $X \cup \{s_\nu\}$ sont équipotents. Pour cela, Cantor considère un sous-ensemble dénombrable $\{t_\nu\}$ de X, et écrit $X_1 = X - \{t_\nu\}$. On a donc : $X = X_1 \cup \{t_\nu\} = X_1 \cup \{t_{2\nu-1}\} \cup \{t_{2\nu}\}$[10] et $X \cup \{s_\nu\} = X_1 \cup \{t_\nu\} \cup \{s_\nu\}$.
Comme $X_1 \sim X_1$, $\{t_{2\nu-1}\} \sim \{t_\nu\}$, $\{t_{2\nu}\} \sim \{s_\nu\}$, il vient : $X \sim (X \cup \{s_\nu\})$.
D'où : $c = \text{card } (\mathbb{N} | \{0, 1\}) = 2^{\aleph_0}$.

D'autres démonstrations, voisines, sont possibles. Certaines utilisent un résultat que Cantor a prouvé en 1891, à savoir que card $\mathfrak{P}(E) = 2^{\text{card } E}$, d'où card $\mathfrak{P}(\mathbb{N}) = 2^{\text{card } \mathbb{N}} = 2^{\aleph_0}$ ($\mathfrak{P}(E)$ désigne l'ensemble des parties de E). Il faut

9. Un même réel peut aussi avoir deux développements décimaux, par exemple 0,1 et 0,099999999… .
10. {t2ν-1} et {t2ν} sont donc respectivement constitués des éléments de {tν} d'indices impairs pour le premier ensemble, pairs pour le second.

alors montrer que [0, 1] et $\mathfrak{P}(\mathbb{N})$ sont équipotents, en prenant en compte, comme ci-dessus, le fait que certains réels ont un double développement décimal.

Cantor indique d'autres résultats sur la puissance du continu, notamment ses formules (13) et (14) : pour tout ν cardinal fini quelconque (*i.e.* entier naturel positif) $c^{\nu} = c$ et $c^{\aleph_0} = c$. Comme l'égalité des cardinaux exprime l'équipotence entre ensembles dont ils sont les cardinaux, ces résultats ont bien la signification suivante :

> *Le continu de dimension ν comme le continu de dimension $\aleph_0$ ont la puissance du continu unidimensionnel.*

$c = c^{\nu}$ est un résultat que Cantor a déjà prouvé en 1878, mais par une méthode plus compliquée empruntant à la géométrie et à la topologie, dans l'article dont il donne les références. C'est à ce propos qu'il avait écrit à Dedekind, car le résultat semblait remettre en cause l'idée de dimension : « Je le vois, mais je ne le crois pas ».

Même si l'égalité numérique $c = 2^{\aleph_0}$ n'exprime elle aussi rien d'autre que l'existence d'une bijection entre $\mathbb{R}$ et $(\mathbb{N} \mid \{0, 1\})$, c'est un résultat extrêmement important aux yeux de Cantor, car il exhibe un lien entre la puissance *c* du continu et la puissance $\aleph_0$ du dénombrable, seule relation que le mathématicien établira jamais, comme on le verra le plus loin.

V - Une « théorie » des entiers naturels ?

La théorie cantorienne des entiers naturels n'est pas à la hauteur du reste de l'article, surtout lorsqu'on la compare à ses devancières proposées par Frege en 1884, Dedekind en 1888 et Giuseppe Peano (1858-1932) en 1889. Pour Cantor, les entiers naturels sont des nombres cardinaux associés aux ensembles finis, mais il ne définit pas la notion d'ensemble fini. Il ne peut affirmer ni que c'est un ensemble en bijection avec une partie finie de $\mathbb{N}$ ni que c'est un ensemble non infini, puisqu'il n'a encore rien dit là-dessus. Il propose donc une approche « inductive » des entiers naturels, présente pour partie dans la théorie des nombres transfinis qu'il a ébauchée en 1883.

Pour Cantor, comme pour beaucoup de ses contemporains, la suite des entiers naturels commence à 1, défini comme le cardinal d'un singleton quelconque $\{e_0\}$:

À une unique chose, si nous la subsumons sous le concept d'un ensemble $E_0 = (e_0)$, correspond comme nombre cardinal ce que nous appelons « un » et désignons par 1.

Les entiers naturels sont ensuite définis à l'aide de l'opération successeur soit, pour les ensembles, par réunion d'un ensemble et d'un singleton :

Tout nombre cardinal (autre que 1) est la somme de son prédécesseur immédiat et de 1.

C'est ce qu'exprime la formule (6) : si $E_{n-1} = \{e_0, e_1, ..., e_{n-1}\}$ et si e_n n'appartient pas à E_{n-1}, alors $E_n = E_{n-1} \cup \{e_n\}$ et card E_n = card $E_{n-1} + 1$.

Introduire 0 sans le définir, comme le fait Cantor, ne fait pas problème : 0 n'est pas ici un nombre mais un simple indice. Comme tous ceux qui définissent les nombres à partir d'ensembles constitués d'éléments, il ne peut d'ailleurs pas, sans artifice, définir 0. Mais on peut lui reprocher de définir 1 après avoir expliqué, au tout début de son article, que

puisque chaque élément isolé m*, abstraction faite de sa nature, devient un « un », le nombre cardinal* $\overline{\overline{M}}$ *est lui-même un ensemble déterminé constitué de purs uns.*

On peut également lui objecter de ne pas respecter ses propres notations : les ensembles apparaissent ici avec des parenthèses, comme des unions d'ensembles, et non avec des accolades, comme des collections d'éléments. C'est ainsi que les formules (2) et (5) sont incorrectes. Il est vrai que dès qu'il traite d'ensembles ordonnés, Cantor en place toujours les éléments entre parenthèses.

Plus grave est le fait que sa théorie des entiers naturels fait explicitement usage de la définition par induction et du principe d'induction complète sans aucunement les justifier. C'est le cas dans la formule (6), selon laquelle card E_n = card E_{n-1}, et dans la démonstration du théorème E, où Cantor écrit :

Il sera supposé l'exactitude du théorème jusqu'à un certain ν et conclu alors comme suit à sa validité pour le suivant $\nu + 1$.

Le cas de ce qu'on appelle aujourd'hui « définition récursive » avait pourtant été traité par Dedekind en 1888, dans *Was sind und was sollen die*

Zahlen ? : il y énonce un théorème qui justifie de telles définitions et lui sert à introduire les nombres cardinaux finis[11]. Le principe d'induction complète, qui justifie le raisonnement par récurrence, est chez lui un théorème, tandis qu'il est un axiome de l'arithmétique chez Peano[12]. Interrogé par ce dernier, Cantor répond que la définition des ensembles finis et le principe d'induction complète sont implicites dans sa théorie et que les nombres finis sont ceux qui satisfont aux conditions de l'induction mathématique.

VI - LE PREMIER CARDINAL TRANSFINI $\aleph_0$

Le § 6 est le plus original et le plus novateur de ceux analysés ici puisqu'il y est spécifiquement question des nombres cardinaux transfinis, même si Cantor se limite au premier d'entre eux, le cardinal de $\mathbb{N}$, appelé aleph-0 et noté $\aleph_0$. Le choix de ce signe, qui date de 1893, n'est pas le fruit du hasard : $\aleph$ est la première lettre de l'alphabet hébreu en même temps qu'il désigne le nombre un. La définition de Cantor, qui écrit que $\aleph_0$ est le nombre cardinal de « la totalité de *tous les nombres cardinaux finis* », donc de $\mathbb{N}$, souffre de ce que sa théorie des entiers naturels est elle-même fautive. Il avait pourtant une autre possibilité, mais qui supposait d'agencer son article différemment : non pas tant définir $\aleph_0$ comme la classe d'équivalence de tous les ensembles dénombrables, car on a vu que ce type de définition lui répugne, mais plutôt dériver la notion de nombre cardinal de celle d'ordinal, ce qu'il fait en 1897 pour $\aleph_1$, le successeur de $\aleph_0$.

Le premier résultat d'arithmétique transfinie est $\aleph_0 + 1 = \aleph_0$ (formule (2)), ce qui revient à dire que le successeur de $\aleph_0$ est $\aleph_0$. Il existe en effet une bijection de $\mathbb{N} \cup \{e_0\}$ ($e_0 \notin \mathbb{N}$) sur $\mathbb{N}$: on associe par exemple 1 à e_0 et $n + 1$ à n. Une formule telle que (2) est évidemment spécifique des nombres transfinis. Si on y ajoute les formules (3) et (4), on caractérise $\aleph_0$ comme le plus petit nombre cardinal transfini.

Cantor ne démontre pas le théorème A – « Tout ensemble transfini a des parties de nombre cardinal $\aleph_0$ » – avec toute la rigueur souhaitable. Il fait observer qu'après avoir supprimé un nombre fini d'éléments dans un ensemble

11. Dedekind, *Que sont et à quoi servent les nombres ?*, trad. Hourya Benis Sinaceur, dans Dedekind, *La création des nombres*, Vrin, 2008.
12. Peano, *Arithmetices prinicipia, nova methodo exposita*, Turin, Fratres Bocca, 1889.

infini, on peut toujours en supprimer un autre, et que l'ensemble de tous les éléments ainsi éliminés est dénombrable. C'est intuitivement vrai, mais affirmer qu'il est possible, après avoir choisi $n - 1$ éléments dans un ensemble, d'en choisir un *n*ième suppose que cet ensemble est bien ordonné, c'est-à-dire qu'il est muni d'une relation ordre telle que toute partie non vide a un élément initial. Il y a ici utilisation implicite du principe du bon ordre, selon lequel tout ensemble peut être bien ordonné.

Les formules (5) à (8) donnent d'autres résultats d'arithmétique transfinie : (5) et (7) devraient être précédées d'un raisonnement par récurrence ; (6) résulte de ce que $\mathbb{N}$ est égal à la réunion de l'ensemble des entiers naturels pairs et de l'ensemble des entiers naturels impairs, ensembles disjoints dont chacun est en bijection avec $\mathbb{N}$. La preuve de (8) renvoie à une construction que Cantor a déjà proposée en 1878, dans un article dont la référence est donnée à la fin du § 6. On regroupe tous les couples dont la somme des termes est égale à un même entier naturel. Ainsi ordonné, l'ensemble de tous ces couples est en bijection avec $\mathbb{N}$.

Les théorèmes C – « Tout ensemble fini *E* est tel qu'il n'est équivalent à aucune de ses parties » – et D – « Tout ensemble transfini *T* est tel qu'il a des parties qui lui sont équivalentes » – sont la version cantorienne de ce que Dedekind a donné en 1888 sous la forme d'une définition : tout ensemble est infini s'il est équipotent à une de ses parties propres et fini dans le cas contraire. Chez Cantor, la démonstration des deux théorèmes est fondée sur le théorème A, dont on a dit que la démonstration est imparfaite, et donc sur la notion de cardinalité. Cela est peu satisfaisant, mais cohérent avec le fait qu'il a déjà utilisé, au début de ce paragraphe consacré à $\aleph_0$, la distinction ensemble fini/ensemble infini en recourant à la notion de nombre cardinal. Comme Cantor le dit, ces deux « théorèmes mettent le plus distinctement en lumière la différence essentielle entre les ensembles finis et transfinis », déjà mentionnée dans l'article de 1878, où il écrivait :

> *Une partie d'un ensemble fini a toujours une puissance plus petite que celle de l'ensemble lui-même ; ce rapport cesse totalement avec les ensembles infinis, c'est-à-dire les ensembles constitués d'un nombre infini d'éléments.*

Il y eut un temps une querelle de priorité entre Cantor et Dedekind sur cette question, qu'il est facile de résoudre au profit de ce dernier. Certes, la définition de Dedekind date de 1888, mais en 1878, comme en 1895, Cantor ne donne pas une définition mais une propriété caractéristique des ensembles infinis. De plus, alors que ce dernier établit expressément un lien avec l'idée, non primitive, de cardinalité, Dedekind la laisse délibérément de côté. Enfin, on sait aujourd'hui que celui-ci possédait sa définition avant 1878.

***Figure 6 :** Le mathématicien allemand Richard Dedekind (1831-1916)*

La fin du § 6 présente la suite de l'article. Cantor prévoit d'étudier les cardinaux supérieurs à $\aleph_0$ et de montrer que l'ensemble de tous ces cardinaux, munis de la relation d'ordre prolongeant celle de $\mathbb{N}$, constitue un « ensemble bien ordonné ». C'est la première mention explicite de cette notion dans l'article, mais sans la définition, qui n'est donnée que dans sa seconde partie. Cantor poursuit en expliquant qu'une « loi déterminée » permet de construire la suite croissante des aleph, avec des indices prolongeant celle des entiers naturels : $\aleph_\omega$ en particulier (ω est l'ordinal de $\mathbb{N}$), est « le nombre cardinal immédiatement supérieur à tous les $\aleph_\nu$ ». Tout cela nécessite la théorie des types d'ordre et la théorie des nombres ordinaux, qu'il présente ensuite. La première date de 1884, mais sa publication avait été bloquée par le directeur de la revue à laquelle l'article l'exposant était destiné (il a été retrouvé et publié en 1970). La seconde est plus ancienne, puisqu'elle a été publiée en 1883 dans l'article dont Cantor donne la référence. Sur cette hiérarchie des aleph, sa tentative a en partie échoué, et il s'est limité à l'étude de $\aleph_1$, dont il donne la définition en 1897 :

c'est le nombre cardinal de l'ensemble de tous les types d'ordre des ensembles bien ordonnés de cardinal $\aleph_0$.

Il y a une autre manière de construire une suite croissante de nombres cardinaux, de la façon suivante : étant donné un nombre cardinal *a*, il est le cardinal d'un ensemble *A*, à partir duquel on peut former l'ensemble $\mathfrak{P}(A)$, dont le cardinal est strictement supérieur à celui de *A*. Mais ces deux modes de construction ne se recouvrent pas : la théorie des ensembles ne permet pas de montrer que $\aleph_1$, qui procède de $\aleph_0$ selon le premier, est égal à card $\mathbb{R}$, lui-même égal à card $\mathfrak{P}(\mathbb{N})$, et qui en procède également, mais selon le second. C'est à ce propos que Cantor aurait pu énoncer le théorème qui porte désormais son nom, et qui n'apparaît nulle part dans l'article, alors qu'il l'a publié en 1891 : pour tout ensemble *A*, card $\mathfrak{P}(A)$ > card *A*, de sorte que pour tout nombre cardinal *a*, il en existe un autre qui lui est strictement supérieur (c'est sous cette dernière forme que Cantor énonce son théorème).

VII - THÉORIE NAÏVE ET THÉORIE AXIOMATIQUE DES ENSEMBLES

La théorie cantorienne des ensembles souffre de plusieurs défauts, qui n'enlèvent rien à son immense mérite et dont certains ont déjà été évoqués. L'un d'eux est l'usage implicite par Cantor, dans cet article comme dans d'autres de ses écrits, du théorème du bon ordre, qu'il n'a jamais justifié de façon satisfaisante bien qu'il permette de prouver la comparabilité universelle des cardinaux. La première démonstration est due à Ernst Zermelo (1871-1953) en 1904. Elle repose sur l'*axiome du choix*, dont une formulation est la suivante :

> *Pour tout ensemble* E, *il existe une fonction* f *de* $\mathfrak{P}$(E) *dans* E, *dite* fonction de choix, *qui associe à toute partie non vide* X *de* E *un unique élément de* X.

La fonction *f*, en distinguant dans chaque partie non vide de *E* un élément particulier, fait de celui-ci l'élément initial de celle-là : *E* peut donc être bien ordonné. Dans le cas de $(\mathbb{N}, \leq)$, *f* associe à chaque partie de $\mathbb{N}$ son plus petit élément[13]. Mais cela n'est pas toujours aussi simple, et l'axiome du choix ne dit pas, par exemple, comment faire de $\mathbb{R}$ un ensemble bien ordonné.

---

13. Le théorème G du § 5 de l'article de Cantor est une formulation du théorème du bon ordre pour $\mathbb{N}$.

La démonstration de Zermelo a soulevé très vite des objections, et la validité universelle de l'axiome du choix, et donc du théorème du bon ordre, a donné lieu à une vive controverse. Accompagnée de la découverte simultanée de paradoxes dans la théorie cantorienne des ensembles, elle a donné naissance à ce qu'on appelle la « crise des fondements », qui pose cette question : comment donner aux mathématiques une base plus solide que ne le fait la théorie de Cantor ? Une réponse aujourd'hui largement admise est celle donnée par Zermelo en 1908. Elle consiste à donner une liste d'axiomes, qui définit la notion d'ensemble et évite les paradoxes connus, tout en permettant de bâtir l'essentiel des mathématiques. Cette théorie *axiomatique* des ensembles, par opposition à celle, dite *naïve*, c'est-à-dire préaxiomatique, de Cantor, a ensuite été amendée par Abraham Fraenkel (1891-1965), et porte le nom de ZF, ou ZFC si on y inclut l'axiome du choix.

Figure 7 : Le mathématicien allemand Ernst Zermelo (1871-1953) *(Oberwolfach Photo Collection)*

Elle ne permet pas de résoudre un problème dont Cantor ne parle pas dans son article et qui a contribué à retarder la publication de sa seconde partie, précisément parce que, bien que l'ayant formulé à de multiples reprises, il n'est pas parvenu à lui trouver une solution. C'est le problème du continu, qui consiste à déterminer le cardinal *c* de l'ensemble des nombres réels. Après avoir défini $\aleph_1$, Cantor aurait pu énoncer ce qu'on appelle aujourd'hui l'hypothèse du continu (HC), dont il a toujours été convaincu de la vérité : $c = \aleph_1$. Il a montré que $c = 2^{\aleph_0}$, mais HC dit autre chose, à savoir que $\aleph_1 = 2^{\aleph_0}$. On sait aujourd'hui que l'axiomatique ZFC ne permet ni de prouver ni de réfuter HC : HC est une proposition indécidable relativement à ZFC. Cela ne signifie pas que HC soit *absolument* indécidable : les recherches se poursuivent et il n'est pas exclu, sous réserve de dépasser les capacités de ZFC, qu'on puisse parvenir à une solution au problème du continu.

Et c'est ainsi que Cantor, en développant une véritable arithmétique des nombres transfinis, a radicalement modifié le visage des mathématiques.

■

(mai 2012)