JEAN CAVAILLÈS

Philosophie mathématique

Préface de Raymond Aron

Introduction de Roger Martin



HERMANN

115, BOULEVARD SAINT-GERMAIN, PARIS VI

© 1962 HERMANN PARIS

Tous droits de reproduction, même fragmentaire, sous quelque forme que ce soit, y compris la photographie, réservés pour tous pays.

TABLE DES MATIÈRES

PREFACE 11

INTRODUCTION 17

REMARQUES SUR LA FORMATION DE LA THEORIE ABSTRAITE DES ENSEMBLES

Introduction 25

Chapitre I. — **Préhistoire : Nombres irrationnels, série trigonométriques, théorie de la croissance** 31

1. La démonstration Bolzano du théorème fondamental de l'analyse 32

2. Les définitions des nombres irrationnels 35

a) définition Dedekind 36
b) définition Weierstrass 39
c) définition Heine-Cantor 41

3. Les séries trigonométriques 44

a) la notion de fonction arbitraire : origine du problème 44
b) les travaux de Dirichlet-Lipschitz 47
c) le mémoire de Riemann 53
1. nouvelle définition de l'intégrale 53
2. les conditions nécessaires pour la représentation 55
d) le problème de l'unicité du développement d'une fonction :
1. convergence uniforme et intégration terme à terme 56
2. travaux de Heine et de Cantor 58

4. Théorie de la croissance de Du Bois Reymond 61

a) idée du continu infinitaire 61
b) première utilisation du procédé de la diagonale 63

Chapitre II. — **La création cantorienne** 67

1. Pressentiments d'une théorie des ensembles..... 67
a) les *Paradoxes de l'infini* de Bolzano..... 67
b) les *Propositions générales sur les espaces* de Dedekind.... 71
2. La découverte des deux puissances..... 72
a) puissance de l'ensemble des nombres réels, distingués de celle de l'ensemble des nombres entiers..... 73
b) invariance de la puissance dans les continus à un nombre quelconque de dimensions..... 74
3. L'opération de la dérivation et la mise en rapport des deux puissances 79
4. La création de 1882-1884..... 83
a) les ordinaux transfinis..... 85
b) ensembles fermés et ensembles parfaits..... 91
5. Le procédé général de la diagonale..... 94
6. Systématisation de la théorie..... 96
a) théorie cardinale..... 97
b) les ensembles finis..... 101
c) les types d'ordre..... 103
d) théorie ordinale..... 107
1. les ensembles bien ordonnés..... 107
2. la récurrence transfinie..... 110
3. forme canonique des nombres ordinaux..... 112
4. les nombres critiques ε..... 114
7. Le paradoxe de Burali-Forti : essai d'une solution au problème du bon ordre (les multiplicités inconsistantes)..... 116
Chapitre III. — **Dedekind et la chaîne — les ordinalisations.** 119
1. Préparation d'une théorie abstraite par Dedekind..... 119
a) définition du matériel de la théorie générale. — La *chaîne*. 121
b) Les ensembles *simplement infinis*..... 124
1. théorème d'existence de l'infini..... 125
2. la suite des entiers et l'ordre..... 126
3. fondement des définitions par récurrence..... 127
c) les ensembles finis..... 132
1. première définition (*Was sind u. was sollen die Zahlen ?*). 132
2. l'ordre cyclique 133
2. Les deux démonstrations du bon ordre, grâce à la *chaîne* : l'axiome de choix..... 135
3. L'axiomatisation de Zermelo..... 140
a) l'antinomie Russell..... 141
b) les axiomes de Zermelo..... 143

4. Elaboration critique (Fraenkel) 146
a) précision du domaine de la théorie : l'axiome de séparation. 146
b) l'indépendance de l'axiome de choix 150
5. Les ordinaux transfinis (von Neumann) 152
a) théorie générale. Bon ordre et nombrement 153
b) fondement des définitions par récurrence transfinie 154
c) l'axiomatique von Neumann : nouvelle démonstration du bon ordre (par le paradoxe Burali-Forti) 155
6. Le paradoxe de Skolem et le manque de catégoricité dans toute axiomatisation de la théorie 158
Conclusion 161
Bibliographie 165
Index des termes employés 175

CORRESPONDANCE CANTOR-DEDEKIND

Avertissement 179
Note du Traducteur 186
Cantor à Dedekind 187
Notes de Dedekind sur les lettres de 1873 193
Cantor à Dedekind 196
Dedekind à Cantor 197
Cantor à Dedekind 198
Dedekind à Cantor 199
Cantor à Dedekind 200
Dedekind à Cantor 203
Cantor à Dedekind 204
Dedekind à Cantor 214
Cantor à Dedekind 216
Dedekind à Cantor 219
Cantor à Dedekind 220
Dedekind à Cantor 224
Cantor à Dedekind 226
Cantor à Dedekind 238
Cantor à Dedekind 244
Dedekind à Cantor 246
Cantor à Dedekind 247
Commentaires et Notes de Zermelo aux lettres de 1899 250

TRANSFINI ET CONTINU 253

NOTE DES EDITEURS

Cette deuxième édition de textes de Jean CAVAILLÈS comporte des corrections d'ordre typographique ou des rectifications de détail dans les références. L'auteur n'eût pas manqué de procéder lui-même à cette mise au point. Dans tous les cas où les éditeurs ont cru devoir modifier ou compléter le texte, les modifications ou additions sont signalées par des crochets [...].

PRÉFACE

Les ouvrages de Jean Cavaillès, réunis dans ce volume, n'ont pas besoin d'être présentés au public scientifique : déjà classiques, indispensables aux spécialistes, ils sont accessibles à tous ceux qui s'intéressent au fondement de la théorie des ensembles. Au reste, je ne possède aucune qualification pour présenter ces travaux d'histoire réflexive des idées, de philosophie mathématique. « Les amis de Jean Cavaillès », — association fondée en 1945 pour honorer la mémoire de celui qui fut, selon le mot de Gabrielle Ferrières, philosophe et combattant, — ont souhaité qu'en leur nom je témoigne de notre fidélité à un homme, dont la vie et la mort furent exemplaires.

La Résistance a été autre chose qu'un épisode d'une guerre impitoyable, autre chose que des réseaux de renseignement ou des maquis. La guerre aurait été gagnée par la coalition antihitlérienne même si les peuples d'Europe avaient subi passivement le joug hitlérien. La contribution matérielle de l'armée de l'intérieur ne mesure pas exactement la portée de l'action entreprise, dès le lendemain de la défaite, par des hommes parmi lesquels notre ami figure au premier rang. La Résistance fut d'abord et avant tout un moment de la conscience française et Jean Cavaillès fut, par excellence, un Résistant par conscience.

Les conditions de la guerre clandestine font des soldats sans uniforme une troupe singulière. Volontaires, ils se sont mobilisés eux-mêmes. Contraints de dissimuler leur nom et leur lutte, ils savent qu'aussi longtemps que l'ennemi n'aura pas été chassé, ils n'ont pas de récompense à espérer, pas de gloire à recueillir. L'obscurité est la loi du combat et le destin des combattants. Le risque qu'ils affrontent n'est pas celui-là seul qui est commun à tous les soldats ; plus que par la mort, ils se sentent menacés par la défaillance. Le Résistant savait que « tomber » pour lui pouvait ne pas signifier tomber au champ d'honneur.

Plus que beaucoup d'autres Jean Cavaillès aurait eu le droit de ne pas s'engager. Il avait fait courageusement la malheureuse campagne de France, il s'était échappé, il avait repris son enseignement d'abord à Clermont-Ferrand où s'était repliée l'université de Strasbourg, puis à la Sorbonne. Par sa thèse, il s'était d'un coup imposé comme un des quelques philosophes dans le monde capables d'accéder aux plus hautes spéculations de l'esprit. Pour la science, pour son pays, pour ses disciples, il représentait une valeur irremplaçable. Il le savait, en dépit de sa modestie, et pourtant il n'hésita pas. Le Professeur Pol Le Cœur évoque ce choix en un témoignage cité par Gabrielle Ferrières : « Je ne pourrais pas non plus oublier le ton de paisible et profond mépris dont il me parla de « ce normalien qui veut conserver son cerveau pour la France ». Il n'est que trop évident que celui-ci avait raison du point de vue pratique, puisque le cerveau de Cavaillès est perdu sans recours, alors que la France sans lui eût été sauvée tout de même, mais le sacrifice d'un

Cavaillès est à la fois cause de la libération de la France, et signe de sa force. » Là est, en effet, la question que nous, ses amis, n'avons cessé de nous poser au long des années. Or, nous savons, d'une certitude immédiate, irrésistible, que la décision de Jean Cavaillès était fondée en raison, nous savons que sa mort a eu un sens, qu'elle a été une sorte d'accomplissement, bien qu'il n'y ait pas de commune mesure entre l'apport matériel du Résistant à la cause de la France et ce qu'eût été l'apport du philosophe à la théorie des sciences ou de la morale.

La Résistance française est devenue peu à peu l'équivalent des armées de guérilla qui, aux quatre coins du monde, s'intitulent depuis quinze ans « armées de libération » et mènent la lutte soit contre une puissance impériale soit contre un gouvernement légal. La Résistance française a eu deux tâches principales, renseignement et action. Elle renseignait l'état-major allié sur le déplacement de troupes, l'organisation des défenses allemandes, l'installation des V1. Elle créait un climat d'inquiétude et de danger autour des armées d'occupation, elle exécutait parfois des destructions d'usines ou d'installations militaires, évitant ainsi les raids de l'aviation alliée et, du même coup, épargnant des pertes aux populations civiles. Enfin elle a créé des maquis, capables d'opérations militaires à ciel ouvert, attaques par embuscades, obstacles élevés sur les routes et les voies ferrées en vue de ralentir les déplacements des convois, etc. Mais si la Résistance débouchait finalement sur une forme de guerre, avait-elle besoin d'un Cavaillès ? Demeurait-elle l'expression du refus moral, en marge de tout calcul de rendement ?

Plus que jamais et sans la moindre hésitation, je répon-

drai que l'acte moral de la Résistance ne pouvait se réaliser comme tel qu'en se soumettant à la loi d'efficacité. L'occupation, la guerre créaient cette apparente antinomie : des hommes dont l'existence n'était pas vouée à la politique, passablement indifférents aux idéologies, éprouvaient le devoir de dire *non* à l'occupant, *non* à la collaboration et ne reculaient devant aucun risque pour traduire en actes de combat cet engagement inconditionnel.

Le langage dans lequel ces hommes exprimaient leur engagement n'était pas toujours identique, même si nous laissons de côté les « politiques », communistes, socialistes ou démocrates. Certains évoquaient l'honneur national, d'autres, tout simplement, l'évidence qu'une vieille nation telle la nation française *doit* se dresser contre l'occupant. Jean Cavaillès, d'après le souvenir que je garde de nos dernières conversations, insistait sur la *nécessité,* qui commandait aux impératifs pratiques aussi bien qu'aux propositions scientifiques, « je suis spinoziste », disait-il, « il faut résister, combattre, affronter la mort. Ainsi l'exige la vérité, la raison ». A partir de cette nécessité spirituelle, le philosophe devenait chef de guerre, attentif aux détails, d'un sang-froid de tous les instants, collectionnant les dates, les lieux, les noms dans sa mémoire infaillible. Comme la charité de Vincent de Paul se manifestait par des organisations multiples, par des *actions* infatigables, ainsi le refus moral de la passivité ou de la collaboration menait les Résistants à faire la guerre, une guerre à la fois plus passionnante mais aussi plus terrible que celle des armées régulières. C'est précisément à une telle guerre que le philosophe Jean Cavaillès ne pouvait pas se dérober. Si quelque ami lui avait objecté que la

France aurait besoin de philosophes et que d'autres pouvaient prendre sa place dans les réseaux, je ne crois pas qu'il aurait invoqué l'argument du Professeur Pol Le Cœur : « Le peuple dont les philosophes se font tuer est un peuple effrayant, donc un peuple redoutable, un peuple fort. » Non, la nécessité était d'ordre spirituel, elle n'était pas subordonnée à la considération des conséquences bénéfiques de l'exemple laissé aux vivants. La Résistance, telle que Cavaillès l'a vécue dès 1940, était militaire et politique en ses implications et ses objectifs, mais elle était morale en son inspiration. Et c'est pourquoi elle ne saurait être entamée par les impuretés du mouvement historique ou les péripéties de l'après-guerre.

De temps à autre surgissent des situations dans lesquelles la conduite de l'individu paraît chargée d'une signification éternelle. Tout se passe comme si certains choix étaient intrinsèquement bons ou mauvais, comme si l'âme était soudain confrontée par l'alternative de se perdre ou de se sauver. La décision de résistance, chez les quelques uns qui l'ont prise en toute conscience, était de ce type. Elle était décision proprement religieuse, au sens où ce terme dépasse toutes les confessions et toutes les églises. Et c'est pourquoi le sacrifice de Cavaillès n'est pas justifié par le profit que la nation a pu en tirer durablement. La justification du sacrifice est donnée dans le sacrifice lui-même, comme la justification de la charité dans la charité même, comme la justification de l'amour dans l'amour même.

Si le souvenir de Jean Cavaillès reste aussi lumineux parmi nous qu'au jour de 1945 où nous apprîmes qu'il était l'inconnu n° 5 du cimetière d'Arras, c'est précisé-

ment que le sens de sa décision ne dépend pas des jugements portés rétrospectivement sur vichystes et gaullistes. Ce que Cavaillès nous rappelle, plus encore que le devoir d'engagement, ce sont les limites de la politique considérée comme combinaison de moyens efficaces en vue de fins souhaitables. Lui aussi a dit un jour : « Voilà ma position, je ne peux pas faire autrement. »

Nous sommes quelques uns qui de cette affirmation, tranquille et héroïque, avons tiré un enseignement dont nous n'épuiserons pas la signification.

Raymond Aron

INTRODUCTION

Les deux études qui accompagnent dans ce recueil l'édition (établie en collaboration avec Emmy Noether) de la correspondance échangée par Cantor et Dedekind constituent la contribution de Jean Cavaillès à l'histoire de la théorie des ensembles. L'importance de cette dernière ne saurait aujourd'hui être surestimée : si certaines de ses parties restent moins utilisées que d'autres, les réserves formulées à sa naissance, contre l'axiome du choix en particulier, ont cessé de gêner les mathématiciens et elle est devenue pour eux un outil indispensable. Parvenue à l'état de théorie classique au terme d'une évolution assez proche de nous, elle permet d'autre part d'étudier dans les meilleures conditions la genèse de ce qui constitue actuellement l'assise la plus fondamentale de la mathématique.

A cet égard, les *Remarques sur la formation de la théorie abstraite des ensembles* (1938) présentent un double mérite. Elles sont tout d'abord une étude historique minutieusement conduite des étapes par lesquelles s'est constituée peu à peu la théorie : anticipations formulées à l'occasion des recherches concernant les nombres irrationnels, les séries trigonométriques et la théorie de la croissance; création cantorienne ensuite et contribution de Dedekind ; enfin axiomatisations successives rendues nécessaires par

l'apparition des paradoxes et par le désir de disposer d'une théorie à la fois aussi sûre et aussi forte que possible. Le but poursuivi par J. Cavaillès n'est cependant pas exclusivement historique et le souci de jeter les bases d'une philosophie des mathématiques reste toujours présent : la théorie des ensembles est considérée comme un exemple privilégié sur lequel on peut essayer de savoir comment « les inflexions inattendues du devenir mathématique » n'excluent pas une objectivité due au fait que « les sinuosités du processus de révélation sont en relation avec la structure des parties révélées ». De telles remarques aboutissent naturellement à une réflexion sur le sens des efforts d'axiomatisation qui atteignent seulement un demi-succès : en effet les systèmes obtenus ne sont pas catégoriques (paradoxe de Skolem) et on ne peut en démontrer la non-contradiction absolue (théorème de Gödel). De ce fait le résultat des recherches poursuivies après 1938 et qui s'inscrivent nécessairement dans les limites établies par ces théorèmes ne sauraient infirmer les conclusions prudentes auxquelles s'arrête l'auteur : le succès d'une théorie générale qui rencontre des difficultés ne se mesure pas à sa capacité de satisfaire sur-le-champ à des exigences fixées a priori mais à la possibilité qu'elle trouve de s'imposer au mathématicien préoccupé de problèmes plus particuliers : « Rejeter ou fonder une théorie n'est travail définitif ni dépourvu de degrés ; (...) il ne peut s'opérer par simple investigation logique; (...) les considérations pragmatistes du mathématicien militant ont le dernier mot. »

La correspondance Cantor-Dedekind dont la traduction fait suite ici aux *Remarques...* s'étend sur la période 1872-1899. Elle est traduite de l'édition publiée

par E. Noether et J. Cavaillès chez Hermann en 1937, à l'exception des lettres de l'année 1899 qui sont traduites du texte publié par Zermelo dans son édition des œuvres de Cantor. On y voit donc apparaître la plupart des idées fondamentales de la théorie des ensembles et on peut y saisir à la fois ce que les exigences de rationalité et de systématisation donnent de nécessaire à une découverte et ce qu'il y subsiste d'incertitude ou de contingence. La correspondance Cantor-Dedekind permet ainsi de juger sur pièces certains thèmes des *Remarques...* et prend par là un intérêt qui vient s'ajouter à sa valeur historique propre.

L'étude consacrée à *Transfini et Continu* enfin complète les *Remarques...* en commentant un théorème publié par Gödel en 1939 et dont celles-ci n'avaient pu faire état. Cette proposition établit que si la théorie des ensembles axiomatisée à la manière de von Neumann modifiée par Bernays puis par Gödel est non-contradictoire, l'adjonction au système de l'axiome du choix et de l'hypothèse du continu ne peuvent y introduire aucune contradiction. Comme la démonstration met en jeu la notion d'ordinal constructible, J. Cavaillès est amené d'une part à rappeler les recherches de Church et de Kleene sur la constructibilité, d'autre part à réfléchir sur la signification philosophique du finitisme de stricte observance et sur la théorie de l'intuition qui se lie à lui naturellement. Il en vient ainsi à rejeter l'idée kantienne d'une intuition immuable au sein de laquelle se construirait l'être mathématique et à souligner au contraire la possibilité d'une « transformation de la zone intuitive » caractéristique du progrès de la pensée mathématique. Observations qui complètent sans les infirmer les conclusions des *Remar-*

ques... puisque, ici encore, « la nécessité apparaît après coup. Aucune analyse de la conscience des actes ne permet de prévoir, pas plus qu'elle ne procure une permanence quelconque. Intuitif est synonyme de conscience effective (ou effectuante), transcendantal de constitutif (ou constituant) relativement à un système conceptuel donné. Pas plus qu'il n'y a de système clos, il n'y a de transcendantal absolu. »

La réédition d'œuvres consacrées à la présentation de questions scientifiques ne va jamais sans soulever des scrupules : les informations qu'elles fournissent n'ont-elles pas perdu de leur actualité ? Les commentaires proposés semblent-ils justifiés dans l'état actuel de la recherche ? Quelle que soit la piété dont s'entoure à bon droit le souvenir de leur auteur, les textes de J. Cavaillès ne sauraient échapper à ces questions mais sont capables d'en supporter l'épreuve. Travail d'historien et de philosophe réfléchissant sur l'histoire, les *Remarques*... n'ont pas vieilli et n'ont perdu ni l'intérêt ni la valeur que leur reconnaissait en 1938 A. Fraenkel. Les pages consacrées à *Transfini et Continu* ont souffert davantage des conditions dans lesquelles elles ont été rédigées et publiées. Ecrites en 1940, publiées en 1947, elles n'ont vraisemblablement pu tenir compte du mémoire détaillé publié par Gödel à Princeton en 1939. A fortiori, elles n'ont pu bénéficier des commentaires ultérieurement donnés par Gödel lui-même. Quelques misères de ce genre, marque d'une œuvre tragiquement interrompue, n'ont cependant pas paru exiger qu'on écarte de la réédition un texte où se trouvent exposés avec beaucoup de densité quelques-uns des thèmes fondamentaux de la pensée de J. Cavaillès. Par ailleurs il faut rappeler qu'au moment où ces pages furent rédigées,

leur auteur était, dans un pays où la recherche logique semble poursuivie par une perpétuelle malchance, un de ceux qui en avaient le mieux compris la portée et le premier qui acceptât de lui donner une place dans son enseignement. Cette clairvoyance, jointe aux qualités d'intelligence et de cœur que l'on sait, méritait que la publication de ce recueil fût aussi un hommage.

Roger Martin

REMARQUES SUR LA FORMATION DE LA THÉORIE ABSTRAITE DES ENSEMBLES

INTRODUCTION

Il est peu d'histoires aussi traversées que celle de la théorie des ensembles. Après un départ difficile, dans une méfiance justifiée par l'échec des spéculations antérieures sur l'infini actuel (1), à peine la construction terminée reçoit-elle la consécration du monde mathématique, que la découverte des antinomies vient en secouer les bases d'un ébranlement qui menace de s'étendre à l'analyse, contre lequel aucun remède ne s'est jusqu'ici révélé efficace. En 1925, VON NEUMANN concluait un nouvel essai d'axiomatisation par la mise en évidence du caractère irréductiblement non catégorique (2) de tout système d'axiomes qui voudrait sauvegarder l'essentiel de la théorie : « nous ne pouvons faire autre chose que constater qu'ici réside une nouvelle objection contre la théorie des ensembles et que pour l'instant aucun moyen de réhabilitation n'est en vue » (3). Aussi bien ces entreprises ont-elles eu pour résultat de provoquer une juxtaposition de théories plus ou moins différentes de l'œuvre cantorienne primitive :

(1) Par exemple la *Géométrie de l'Infini* par FONTENELLE, tissu de raisonnements incorrects.

(2) C'est-à-dire que le système d'axiomes considéré comme définition implicite ne détermine pas de façon univoque son objet. Nous renvoyons pour cette notion à J. CAVAILLÈS, *Méthode axiomatique et formalisme*. Paris, Hermann, 1937, ch. II.

(3) VON NEUMANN, *Eine Axiomatisierung der Mengenlehre. Journ. f. Math.*, t. 154, p. 240.

il n'y aurait plus une théorie des ensembles, mais celle de ZERMELO-FRAENKEL, celle de VON NEUMANN, celle de BROUWER. Or, les utilisations, au moins dans les mathématiques classiques, exigent l'unité. Ici apparaît un autre fait qui n'est pas le moindre paradoxe de cette histoire, les recours toujours plus nombreux à une théorie contestée dans les mathématiques incontestables. Dès 1873, dans une lettre à DUBOIS-REYMOND, WEIERSTRASS se servait de la dénombrabilité des nombres algébriques, — que CANTOR venait de démontrer, — pour étendre un théorème sur les fonctions de variables réelles (1). A partir de 1884 avec les mémoires de POINCARÉ sur la structure du domaine d'existence des fonctions automorphes, de MITTAG-LEFFLER sur la possibilité de construction de fonctions avec des singularités données, se développent une série d'applications toujours plus importantes en analyse (2). Contrairement à l'attente de CANTOR, tourné vers les fonctions analytiques, c'est surtout à l'étude des fonctions de variables réelles que la théorie fait faire les progrès les plus considérables. Les travaux de BAIRE, BOREL, LEBESGUE donnent des modes nouveaux d'engendrement ou de mise en correspondance dans l'espace cartésien et transforment la théorie de l'intégration. La crise théorique déclenchée en 1905 ne fait qu'embarrasser de restrictions et de scrupules les raisonnements sans arrêter leur puissance interne d'expansion. Aujourd'hui restrictions et scrupules passent au second plan quoique le problème du fondement semble moins près que jamais d'une solution, et trois disciplines manifestent une dépendance directe et essentielle par rapport à la théorie des ensembles : l'analyse fonctionnelle ou topologie générale fondée en 1909 par FRÉCHET, qui elle-

(1) Lettre du 15 décembre 1874, *Acta Mathematica*, t. 39 (1924), p. 206.

(2) Cf. A. DENJOY, *Introduction à la théorie des fonctions de variables réelles*. Paris, Hermann, 1937, p. 12 : « La plus grande révolution que la mathématique ait connue depuis l'invention du calcul différentiel et intégral venait de s'ouvrir. »

même trouve chaque jour de nouveaux champs d'application, la théorie de la mesure et de l'intégration développée par BOREL, LEBESGUE et DENJOY, enfin la topologie combinatoire. Sans doute ne s'agit-il que d'utilisations chaque fois nettement délimitables, et certains mathématiciens ont-ils même soin de séparer les résultats obtenus à l'aide des parties les plus contestées de la théorie, par exemple les différents énoncés plus ou moins forts de l'axiome de choix. Mais, une méthode n'est pas chose isolable arbitrairement pour les besoins d'un problème. L'exigence même de sécurité oblige à la dépouiller de son vêtement accidentel, pris dans un cas particulier, à préciser les conditions nécessaires et suffisantes de son application. Si elle réussit avec certains ensembles, ce ne peut être qu'en vertu de propriétés définissables en eux : il n'y aura mathématique rigoureuse que lorsqu'on aura, par l'investigation même des procédés, défini le champ d'objets auxquels ils conviennent. D'autre part, ces procédés s'enchaînent dans un complexe organique : c'est un état d'esprit qui constitue la base secrète d'où ils jaillissent avec une apparente imprévisibilité, à quoi il faut revenir si l'on veut vraiment comprendre et non plus exposer mais continuer. Parenté profonde qui unit — sans qu'ils en aient même pleine conscience, comme une entente implicite — les chercheurs d'une même discipline au moins dans une période donnée, *vis a tergo*, élan objectif de la recherche elle-même. Cette double unité, justification, pour une théorie, de son existence mathématique, la théorie des ensembles la possède-t-elle ? Rechercher la première est travail technique : on a vu à quelles difficultés il se heurte. Mais peut-être, ne fut-ce que pour évaluer ses chances d'aboutir, est-il possible d'essayer de déceler la seconde là où elle existe, sous les juxtapositions, peut-être factices, d'une construction arbitraire. C'est ici que l'histoire intervient.

L'histoire mathématique semble, de toutes les histoires, la moins liée à ce dont elle est véhicule ; s'il y a lien, c'est

a parte post, servant uniquement pour la curiosité, non pour l'intelligence du résultat : l'après explique l'avant. Le mathématicien n'a pas besoin de connaître le passé, parce que c'est sa vocation de le refuser : dans la mesure où il ne se plie pas à ce qui semble aller de soi par le fait qu'il est, dans la mesure où il rejette autorité de tradition, méconnaît un climat intellectuel, dans cette mesure seule il est mathématicien, c'est-à-dire révélateur de nécessités. Cependant avec quels moyens opère-t-il ? L'œuvre négatrice d'histoire s'accomplit dans l'histoire. Double liaison : avec les problèmes posés et étudiés dans un temps — choix de la rébellion —, avec les méthodes déjà existantes, matière où forger le nouvel instrument. Dans les deux cas, l'arbitraire individuel ou le style d'un milieu ne suffisent pas à expliquer : même si l'on concevait les mathématiques comme système en soi, les sinuosités du processus de révélation seraient en relation avec la structure des parties révélées. Autrement dit, il y a une objectivité, fondée mathématiquement, du devenir mathématique ; c'est l'exigence d'un problème qui oblige à dépouiller une méthode d'accidents qu'aucune réflexion n'apercevait inutiles, c'est la vigueur interne d'une méthode qui dépasse son champ primitif d'application et pose de nouveaux problèmes. Connexion du système mathématique d'un temps, pour lequel la réciprocité d'actions exclut aussi bien les lacunes intérieures, que l'imagination d'un vide extérieur où se situerait la fixité d'un modèle.

Dans cette indétermination, le départ est difficile à faire entre enchaînements psychologiques ou sociaux et nécessité mathématique — on le voit à propos de ces découvertes simultanées, explicables à la fois par la présence de l'instrument ou l'acuité d'une question, et parce que l'attention y était socialement portée. Les éléments indubitablement objectifs sont ces nœuds que constituent les théories : ici dans l'enchevêtrement entre notions et méthodes disparaissent les liens de causalité au profit des relations d'in-

telligibilité. Processus dont l'exposé systématique — axiomatisation par exemple — est manifestation mais non indispensable. Il est possible d'apercevoir l'unité interne dans l'intuition centrale d'où procèdent les enchaînements, de la déceler aussi méthodiquement par la seule description du phénomène. Suivre la genèse des notions, préciser surtout leurs liens effectifs avec les problèmes et isoler les procédés généraux dont le geste intérieur rapproche le plus de l'intuition centrale impossible à décrire, tel est le travail — à la fois soumis à l'histoire et critique de celle-ci au nom de ses résultats — qui a quelque chance d'aboutir à un résultat objectif. Devant cet effort pour situer au sein du système mathématique, en même temps que caractériser une structure autonome, devraient disparaître les architectures factices, peut-être même aussi se deviner la stérilité de certaines questions. C'est un étonnement toujours renouvelé que les inflexions inattendues du devenir mathématique, son ironique abandon des voies qu'un essai de prévision systématique ouvrait devant lui. A ce problème d'une explication de l'inégale fécondité des méthodes, une étude critique suffisamment ample de théories voisines pourrait peut-être apporter quelque lumière. Mais il ne s'agit ici que d'une application limitée du procédé à la question spéciale posée à propos de la théorie des ensembles : apparition nécessaire ou, et dans quelle mesure, fantaisie historique, structure autonome ou pluralité enrobée dans un système contingent.

CHAPITRE I

PRÉHISTOIRE NOMBRES IRRATIONNELS SÉRIES TRIGONOMÉTRIQUES

Le XIX^e siècle mathématique s'ouvre par une crise de l'infini : aux premières objections contre infiniment grand et infiniment petit, le XVIII^e avait fait succéder une confiance aveugle, pour le passage à la limite, par exemple, et l'emploi des séries. En 1826, ABEL pouvait se plaindre encore qu'on raisonne sur des séries divergentes : « elles sont quelque chose de bien fatal et c'est une honte qu'on ose y fonder aucune démonstration... Ce sont elles qui ont fait tant de malheurs et causé tant de paradoxes (1) ». Des malheurs et des paradoxes la méthode d'exhaustion avait su se garder tant que réservée à la géométrie, garantie contre les absurdités par l'intuition sensible. La révolution cartésienne, en assurant un parallélisme complet entre nombres et étendue, en inaugurant l'étude des variations de fonctions, incorporait le continu à l'algèbre : celle-ci, devenue autonome, devait définir par ses propres moyens les objets employés, préciser, étendre elle-même les modes d'emploi. D'où deux sortes de problèmes : donner un fondement purement ana-

(1) ABEL, *Untersuchungen über die Reihe* $1 + \frac{m}{1} x + \frac{m.(m-1)}{1.2} x^2 + \ldots$ *Journal de Crelle*, t. 1 (1826), p. 311.

lytique aux notions et propriétés assurées jusque-là seulement par l'évidence de leur corrélat géométrique ; développer les unes et les autres hors des limites arbitraires imposées par l'intuition. Etablissement rigoureux de la théorie des fonctions continues, extension des opérations analytiques, double travail du XIX^e siècle qui exigeait, pour remplacer l'intuition défaillante, la création des notions essentielles de la théorie des ensembles.

En 1799, BOLZANO examine (1) les différentes démonstrations du théorème que « toute fonction algébrique rationnelle d'une variable peut être décomposée en facteurs réels du 1^er ou du 2^e degré » ; aucun raisonnement employé n'échappe à l'alternative : cercle vicieux, recours à l'intuition géométrique. Ainsi ce postulat posé au préalable (et d'où le théorème se déduit aisément) « qu'une ligne continue à courbure simple dont les coordonnées sont d'abord positives puis négatives doit rencontrer l'axe des abscisses au moins en un point situé entre ces ordonnées ». Rien ne manque à son évidence sensible ; si on veut rester dans l'analyse, il faut traduire — toute fonction continue de x qui dans un intervalle prend des valeurs positives et négatives doit au moins une fois s'annuler — et « chercher un fondement objectif ». La définition de la continuité est la définition moderne. BOLZANO utilise un lemme important aussi bien par son énoncé que par son mode de démonstration. Soit un caractère M que possèdent une partie des nombres (2) de l'intervalle (0 ... 1) ; en particulier tous les nombres inférieurs à u. On peut déterminer dans (0...1) un nombre U qui soit le plus grand des nombres tels que

(1) BOLZANO (I), p. 4-14. Les numéros renvoient à la bibliographie méthodique.

(2) Nous employons dès maintenant le mot nombre pour désigner un élément quelconque de l'intervalle arithmétique (0...1). Pour BOLZANO et ses contemporains, c'était encore le mot « valeur » ou « grandeur » qui s'imposait, tant était encore géométrique la notion de continu.

u. Soit, en effet, $u + v$ un des nombres (existant par hypothèse) qui ne possèdent pas la propriété M ; on examinera si tous les nombres entre u et $u + \frac{v}{2}$ possèdent M ; en cas de réponse négative, si tous entre u et $u + \frac{v}{2^2}$, etc. Deux éventualités : ou on arrive à une réponse affirmative pour $u + \frac{v}{2^m}$, ou l'on a toujours des réponses négatives. Dans le 2e cas $U = u$; dans le 1er cas, on recommence le même raisonnement pour l'intervalle :

$$\left(u + \frac{v}{2^m} \ldots u + \frac{v}{2^m} + \frac{v}{2^{m+n}}\right)$$

S'il est possible de poursuivre ainsi indéfiniment dans des intervalles emboîtés, le nombre U sera défini comme la limite de la suite convergente.

$$u, u + \frac{v}{2^m}, u + \frac{v}{2^m} + \frac{v}{2^{m+n}}, \ldots$$

dans le cas contraire, il coïncide avec un des termes de la suite.

CANTOR (1) donne une analyse générale de ce procédé pour étudier une propriété *distributive*, c'est-à-dire telle que si elle appartient aux éléments d'un intervalle (ou à un ensemble contenu dans cet intervalle), elle appartient à ceux d'au moins une portion de cet intervalle divisé en un nombre fini de portions. Cette polychotomie finie qui se poursuit de proche en proche dans des intervalles emboîtés de plus en plus petits, est « une méthode qui peut être remplacée difficilement par une autre ; l'essentiel en est très vieux, et, dans les temps modernes, utilisé par LAGRANGE, LEGENDRE, DIRICHLET, CAUCHY » (2). On verra l'usage qu'en

(1) CANTOR (XV), p. 212. La pagination est, de même que pour les mémoires de DEDEKIND, empruntée aux recueils CANTOR (XXVIII) et DEDEKIND (V).
(2) *Ibid.*

feront non seulement CANTOR lui-même, mais PEANO et, en particulier, BOREL et LEBESGUE pour la démonstration de leur lemme. Si BOLZANO n'est pas ici novateur, du moins l'application est-elle originale : les nombres de l'intervalle sont des éléments d'ensemble pour lesquels on ne suppose rien sinon qu'ils sont ordonnés et qu'on peut décider s'ils possèdent ou non la propriété M.

Le théorème principal peut être alors établi : soit a un nombre rendant la fonction $f(x)$ négative ; puisque $f(x)$ est continue, il existe un nombre η tel que :

$$\text{pour } \varepsilon < \eta \quad f(a + \varepsilon) < 0$$

(caractère M du lemme pour les nombres ε) ; soit U le nombre du lemme : on ne peut avoir ni

$$f(a + U) < 0$$

puisqu'en vertu de la continuité $f(x + U + \varepsilon)$ serait encore négatif pour ε assez petit, contrairement à la définition de U, ni

$$f(a + U) > 0$$

pour les mêmes raisons relativement à $f(a + U - \varepsilon)$. Donc

$$f(a + U) = 0.$$

Mais une rigueur complète n'est toutefois pas atteinte : il faut admettre que la suite convergente :

$$u,\ u + \frac{v}{2^m},\ u + \frac{v}{2^m} + \frac{v}{2^{m+n}},\ldots$$

donne l'existence de U. BOLZANO croit le démontrer en prouvant que, la suite étant donnée, on peut calculer sa limite avec une approximation arbitraire, comme si la suite des approximations n'était pas à son tour une suite convergente. Ici encore subsistent donc irréductiblement cercle vicieux ou recours à l'intuition : seule la définition complète du nombre réel permet de les éviter.

* * *

Il est assez remarquable qu'en 1876 LIPSCHITZ — qui devait d'ailleurs adopter en 1877 une définition voisine pour son traité d'analyse, le « premier rigoureux » au dire de PASCH (1) — refuse encore toute originalité à la théorie de DEDEKIND dans *Stetigkeit und irrationalen Zahlen* : « je ne nie pas la rectitude de votre définition, mais j'estime qu'elle ne se distingue que dans la forme de l'expression, et non pour le fond, de celle que les anciens ont posée... Aussi souhaiterais-je que vous laissiez l'affirmation que des propositions comme $\sqrt{2}.\sqrt{3} = \sqrt{6}$ n'ont jusqu'ici pas été démontrées. Je crois, en effet, qu'en particulier les lecteurs français seront convaincus avec moi que le ... [V^e] livre d'EUCLIDE contient les principes nécessaires et suffisants pour la démonstration de ce théorème » (2). Si grande était la difficulté de distinguer à ce degré de généralité entre évidence d'un acte et démonstration d'une possibilité dans un domaine abstrait. « EUCLIDE, répond DEDEKIND, peut appliquer sa définition de l'égalité des rapports à toutes les grandeurs... dont l'existence est admise pour de bonnes raisons. Mais pour fonder l'arithmétique sur la notion de rapport de grandeurs (ce qui n'a jamais été le but d'EUCLIDE) elle ne suffit absolument pas (3). » En réalité, le processus d'abstraction nécessaire se heurtait à la représentation traditionnelle des deux actes générateurs des objets mathématiques : compter et mesurer. ILLIGENS reprochait en 1889 dans les *Mathematische Annalen* (4) à la définition de CANTOR de ne concerner que des signes abstraits qui, n'étant ni multiplicités ni grandeurs, n'avaient rien de comparable

(1) PASCH (7), p. 1.
(2) *Correspondance Lipschitz Dedekind* in DEDEKIND (V), t. III, p. 469.
(3) *Ibid.*, p. 477.
(4) ILLIGENS (I).

aux nombres véritables. Pourtant, l'introduction des nombres complexes avait déjà nécessité l'abandon du « concept obscur de grandeur ». De HAMILTON à GRASSMANN le calcul symbolique rendait familières ces définitions, par leurs propriétés formelles, d'opérations effectuables sur un matériel abstrait quelconque. Le recours à la notion d'ensemble est inévitable : c'est d'elle que doit partir une mathématique qui ne veut plus reconnaître l'absolu d'un acte. La quasi simultanéité des différentes définitions du continu numérique entre 1870 et 1880 est, à cet égard, suffisamment instructive.

Les trois méthodes de DEDEKIND-TANNERY (1), de WEIERSTRASS et de HEINE-CANTOR ne se distinguent guère que par la présentation : DEDEKIND procède de façon plus descriptive en caractérisant par ses propriétés l'ensemble des nombres réels, WEIERSTRASS analyse la notion de nombre pour y trouver motif et possibilité de l'extension cherchée, enfin HEINE et CANTOR se préoccupent davantage des procédés de calcul. Dans tous les cas on « prend comme seule base, sur laquelle on doit s'être mis d'accord, l'arithmétique des nombres rationnels. » (2)

La notion fondamentale pour DEDEKIND est celle d'inextensibilité ; les nombres rationnels constituent un *corps* (système invariant par rapport aux 4 opérations) soumis à un ordre que caractérisent trois propriétés : 1° la définition habituelle de l'ordre (relation existant entre 2 éléments quelconques, asymétrique, non réfléchie, transitive) ; 2° entre deux éléments quelconques l'ordre en situe une infinité d'autres ; 3° tout élément sépare le système en deux classes non vides telles que tout élément de la première précède dans l'ordre tout élément de la deuxième (propriété distin-

(1) Nous n'exposons pas la méthode de Tannery parce qu'identique et — quoique trouvée de façon indépendante — postérieure à celle de Dedekind.
(2) L. de DEDEKIND à LIPSCHITZ, 1876, *l. c.*, p. 470.

guant l'ordre linéaire d'un ordre circulaire, ou plan). L'ensemble des nombres réels sera le corps maximum possédant ces trois propriétés : son inextensibilité s'appellera *continuité*. Pour l'obtenir il suffit d'exiger par rapport à l'opération impliquée dans la 3e propriété la même fermeture que par rapport aux 4 opérations arithmétiques : toute répartition en deux classes soumises aux conditions précédentes (ou coupure) est déterminée par un élément de l'ensemble (dernier élément de la 1re classe ou 1er élément de la 2e) : on voit qu'il est alors impossible d'ajouter un nouvel élément (1). L'ensemble L des nombres rationnels n'est pas inextensible : la répartition

$$x^2 < D < (x+1)^2,$$

D n'étant pas un carré, est une coupure dont les deux classes n'ont ni premier ni dernier élément. Considérons l'ensemble Λ des coupures de l'ensemble L ; on peut y établir un ordre : une coupure en précède une autre si la 1re classe de la 2e contient des éléments de la 2e classe de la 1re. On démontre (2) que cet ordre possède les trois propriétés énoncées plus haut. De plus : tout élément a de L possède un représentant α dans Λ en vertu de la propriété 3° de L (l'ordre est évidemment conservé) ; entre deux éléments quelconques de Λ il y a une infinité d'éléments α représentants d'éléments de L (si, en effet, dans L, deux classes correspondant à deux coupures diffèrent par plus d'un élément, elles diffèrent par une infinité d'éléments en vertu de la propriété 2°). Enfin apparaît pour Λ une 4e propriété : toute

(1) En effet son insertion dans l'ordre déjà établi provoquerait une coupure dans le système primitif : il coïnciderait donc, en vertu de la propriété 2°, avec l'élément déterminant la coupure.

(2) La propriété 1° est évidemment présente ; entre 2 coupures de L il y a une infinité de coupures, celles que déterminent les éléments de L ; enfin une coupure de L répartit bien les autres en deux classes, celles qui la précèdent et celles qui la suivent ; chacune des classes comportant évidemment des éléments (les coupures déterminées par les éléments de L en particulier).

coupure Π est déterminée de façon univoque par un élément et un seul (premier élément de la classe Π_2 ou dernier élément de la classe Π_1). En effet, la coupure Π définit une coupure P dans L (les deux classes sont formées respectivement des éléments dont les représentants sont dans Π_1 et Π_2) donc un élément Π de Λ : Π ne peut être que le 1er élément de Π_2 ou le dernier de Π_1, sans quoi il serait précédé dans Π_2, ou suivi dans Π_1 d'éléments représentants d'éléments de L, ce qui est impossible. L'ensemble Λ est un corps si on définit les 4 opérations arithmétiques à partir des mêmes opérations effectuées sur les nombres rationnels des classes correspondant aux coupures (la somme de deux coupures est la coupure dont la 1re classe est formée de tous les nombres rationnels inférieurs à la somme de deux nombres quelconques appartenant respectivement aux premières classes des deux sommants). Une suite convergente détermine bien ici un nombre fixe qui diffère aussi peu qu'on veut du terme général de u_n de la suite pour n assez grand : la répartition des nombres de Λ en classe de ceux pour lesquels, quel que soit N, on peut trouver un $n > N$ tel que u_n les dépasse, et classe des autres, constitue une coupure ; le nombre séparatif est la limite cherchée.

Le passage du système L au système Λ est produit par la « puissance créatrice » (1) de notre esprit : autrement dit, d'une part les éléments de Λ possèdent une existence aussi authentique (et de même ordre) que les nombres rationnels, d'autre part ils s'en distinguent justement par l'étape du processus créateur qui fait passer du nombre à la coupure (si elle est déterminée par un nombre rationnel) puis de la coupure au nombre (réel). Le modèle — et sans doute l'idée première — de semblables créations se trouve dans les idéaux de KUMMER dont l'introduction, « conduite avec assez de rigueur, doit être considérée comme absolument légi-

(1) Dedekind l. à Weber, 24 janvier 1888, *in* DEDEKIND (V), t. III, p. 489.

time » (1). C'est en acte la méthode axiomatique dans une de ses premières utilisations : « une infaillible méthode pour l'analyse [de toutes les suppositions explicites ou tacites] ... est de remplacer toutes les expressions nouvelles par des mots arbitraires dépourvus de sens, l'édifice doit, s'il est bien construit, ne pas en être ébranlé, et j'affirme que par exemple ma théorie des nombres réels soutient l'épreuve. » (2)

Au contraire, WEIERSTRASS (3) se place au point de vue génétique : « Le nombre, dans le cas le plus simple, est la représentation synthétique de l'adjonction de l'unité à l'unité. » De toute façon, il se rapporte à une multiplicité : « un nombre complexe est la représentation synthétique de groupes différents d'éléments homogènes entre eux. » (4) Ainsi le nombre résulte, de toutes façons, d'un processus d'abstraction appliqué à une collectivité : il y faut distinguer les éléments qui peuvent appartenir à des espèces différentes et leur agrégat (ANZAHL) fini ou infini (5). Opérations et relations quoique spécialisées pour chaque espèce, sont soumises à des principes généraux, où l'on peut trouver à peu près exactement les définitions correspondantes formulées plus tard par CANTOR pour la théorie des ensembles. La somme de 2 nombres est l'agrégat formé de la réunion de leurs éléments, le produit de α par β, l'agrégat obtenu en remplaçant chaque élément de β par un agrégat équivalent à α. On peut, dans un nombre, changer les éléments fondamentaux, sans changer le nombre : ainsi deux nombres sont égaux si, par un changement convenable, on

(1) *Ibid.*, p. 490.

(2) DEDEKIND l. à LIPSCHITZ, 27 juillet 1876, *ibid.*, p. 479.

(3) La théorie de WEIERSTRASS a été exposée en 1872 dans KOSSAK (I).

(4) *Ibid.*, p. 16.

(5) Lorsqu'un nombre est représenté par un agrégat fini, avec comme éléments des nombres rationnels, par exemple, il coïncide avec la somme de ses éléments.

peut les amener à constituer un agrégat identique des mêmes éléments. Mais il y a une définition plus générale de l'égalité s'appliquant aussi aux agrégats infinis : deux nombres sont égaux si tout nombre formé d'un agrégat fini d'éléments de l'un d'eux est partie intégrante de l'autre. Un nombre a dénotant un agrégat infini est fini s'il existe un nombre b formé d'un agrégat fini tel que tout nombre formé d'éléments de a (en agrégat fini) est contenu dans lui, la réciproque étant impossible (ce qui donne, d'après les définitions précédentes, $a < b$). Tel est le nombre irrationnel agrégat infini d'éléments qui sont des nombres rationnels (que l'on suppose déjà définis). KOSSAK qui rapporte cette définition, pressé d'arriver à la théorie des nombres complexes (agrégats avec différentes espèces d'éléments), ne précise pas davantage. En 1883, CANTOR (1) la reprend — d'après « les leçons de WEIERSTRASS sur les fonctions analytiques » (2) — comme introduction à la sienne qu'il réexpose de façon plus systématique. « Le moment générateur qui lie l'agrégat avec le nombre défini par lui réside dans la sommation » (des éléments de l'agrégat). Cependant cette sommation n'est jamais que finie ; il y aurait la même « faute logique » à définir le nombre irrationnel comme somme infinie que comme limite de nombre rationnel puisque « la définition de la somme infinie n'est obtenue que par identification avec le nombre irrationnel auparavant défini » (3). L'inconvénient que le même nombre peut être représenté par une infinité d'agrégats s'évite aisément : il suffit de spécialiser le mode de constitution des agrégats, d'utiliser seulement, par exemple, comme éléments les fractions décimales ou les fractions continues. Mais il y a une autre difficulté : la sommation finie des éléments d'un agrégat (nécessaire, par exemple, pour la comparaison avec un autre nombre) doit

(1) CANTOR (XV), p. 190.
(2) *Ibid.*, p. 183.
(3) P. 184-185.

s'effectuer dans un ordre déterminé si l'agrégat est constitué par les termes d'une série non absolument convergente. Or la définition de WEIERSTRASS ne peut faire intervenir aucun ordre entre les éléments.

Aussi vaut-il mieux substituer à la notion d'agrégat celle de suite fondamentale introduite simultanément, en 1871, par CANTOR (1) et HEINE. Avec toujours comme base l'ensemble des nombres rationnels, on considère l'ensemble des suites convergentes formées de ces nombres. Ce sera l'ensemble des nombres irrationnels d'ordre 1. On voit que tout nombre rationnel a y a un représentant : la suite $(a, a...)$. L'ordre et les quatre opérations sont aisés à définir : deux éléments sont égaux si la différence $a_n - a'_n$ des termes de même rang des suites fondamentales tend vers o avec $\frac{1}{n}$; un élément en précède un autre si $a_n - a'_n$, à partir d'un certain rang, reste plus grand qu'un nombre positif rationnel : un de ces deux cas se présente forcément pour tout couple d'éléments. La somme, la différence, le produit de deux éléments ont une suite fondamentale composée respectivement des sommes, différences ou produits des termes de même rang dans les suites. D'une façon générale toute relation algébrique $\Phi(\alpha, \beta, \gamma, ...)$ entre éléments se traduit en relations algébriques $F(a_n, b_n, c_n...)$ entre nombres rationnels. Si des éléments $\alpha_1, \alpha_2, \ldots \alpha_n \ldots$ forment une suite convergente (2), il est toujours possible de former avec les nombres rationnels de leurs suites respectives une seule suite fondamentale qui définit un nouvel élément β appelé limite de la suite $\alpha_1, \alpha_2 \ldots$: on voit que $\beta - \alpha_n$ peut être, d'après les définitions données, rendu aussi petit qu'on veut pour n assez grand. Toutefois, si l'on considère l'aspect « conceptuel » des opérations, la

(1) CANTOR (V), p. 92-96 ; HEINE (II), p. 172.

(2) Une suite d'éléments α_i est convergente si la différence $\alpha_{m+n} - \alpha_m$ — qu'on a appris à définir — est plus petite qu'un nombre donné quelconque pour n assez grand.

suite convergente $\alpha_1, \alpha_2, \ldots \alpha_n \ldots$ engendre en réalité un élément d'un autre ordre : on superpose donc au domaine des nombres irrationnels d'ordre 1, un domaine d'ordre 2 formé des suites fondamentales d'éléments d'ordre 1 ou d'ordre 0 (nombres rationnels), et ainsi de suite indéfiniment, les domaines d'ordre $\geqslant 1$ ayant la propriété de se recouvrir tous exactement (chaque élément de l'un d'eux a un représentant dans tous les autres). CANTOR voit à sa définition l'avantage qu'elle décrit exactement le procédé qui donne pratiquement les nombres irrationnels dont s'occupe l'analyse. Ainsi pour « décrire le contenu intellectuel de la proposition $\sin \frac{\pi}{2} = 1$ on peut se donner $\frac{\pi}{2}$ et ses puissances par les suites (1) :

pour $\frac{\pi}{2}$: (α_ν)

pour $\left(\frac{\pi}{2}\right)^{2m+1}$: (α_ν^{2m+1}) où $\alpha_\nu = 2 \sum_{n=0}^{\nu} \frac{(-1)^n}{2n+1}$

Enfin : $\sin \frac{\pi}{2}$: $\left(\sum_{m=0}^{p} (-1)^m \frac{\left(\frac{\pi}{2}\right)^{2m+1}}{2m+1} \right)$

c'est-à-dire $\sin \frac{\pi}{2}$ est défini grâce à une suite fondamentale de 2e ordre et par cette proposition est exprimée l'égalité entre le nombre rationnel 1 et un nombre $\sin \frac{\pi}{2}$ donné par une suite fondamentale de 2e ordre. De façon analogue, le contenu intellectuel de formules compliquées comme celles de la théorie des fonctions θ se peut décrire avec précision et une relative simplicité (2). »

Mais cette « distinction conceptuelle » entre les différents

(1) CANTOR et HEINE représentent une suite fondamentale de terme général a_n par (a_n).

(2) CANTOR (XV), p. 189.

ordres de nombres irrationnels avait, dans le mémoire de 1871, un but précis. Définissons comme points rationnels sur un axe les points dont l'abscisse est donnée par un nombre rationnel. Tout point de l'axe « s'il nous est connu par une construction quelconque » (1), il est possible de donner une suite fondamentale (a_n), telle que les points d'abscisse a_n se rapprochent indéfiniment de lui. Inversement, ce sera un axiome « parce qu'il est dans sa nature de ne pouvoir être démontré de façon générale » (2) — l'axiome de CANTOR — qu'à tout nombre irrationnel correspond un point de la droite ayant ce nombre pour abscisse. Si l'on considère maintenant l'ensemble A' des points limites d'un ensemble A (son 1er dérivé, c'est-à-dire l'ensemble des points tels qu'il y ait, aussi près d'eux qu'on veut, toujours des points de cet ensemble) (3), puis l'ensemble A'' des points limites de A' (2e dérivé) et ainsi de suite, l'ensemble A pourra être caractérisé par le nombre possible de ces opérations successives de *dérivation* : on devra s'arrêter à un ensemble $A^{(\nu)}$ s'il ne comporte qu'un nombre fini de points (donc pas de point limite). L'ensemble A sera dit de ν^e espèce. « Un exemple d'ensemble de ν^e espèce est déjà offert par un seul point, si son abscisse est donnée par un nombre irrationnel d'ordre ν, satisfaisant à des conditions faciles à poser. Si l'on résoud, en effet, ce nombre en ses éléments, les termes d'ordre $\nu-1$ de la suite fondamentale correspondante, puis ceux-ci dans leurs éléments constitutifs d'ordre $\nu-2$, etc...., on obtient enfin une infinité de nombres rationnels, l'ensemble des points correspondant est un ensemble de ν^e espèce (4). »

Ainsi c'est bien la théorie des ensembles naissante qui dicte à CANTOR cette transformation de la définition de

(1) CANTOR (V), p. 96.

(2) P. 97.

(3) Pour la définition des points limites et de l'opération de dérivation, voir ci-après, p. 56 sq.

(4) P. 98.

WEIERSTRASS. La notion de suite fondamentale a comme intérêt direct de donner le corrélat analytique de l'opération de dérivation. Celle-ci elle-même se trouvait exigée à son tour par les recherches faites à la suite de RIEMANN et de HEINE sur l'unicité de la représentation d'une fonction par une série trigonométrique.

* * *

On sait l'importance centrale des travaux sur les séries trigonométriques dans le développement de l'analyse au XIX^e siècle : c'est grâce à eux que la notion générale de fonction a pu se dégager, c'est à leur occasion que les principales extensions de l'intégration ont été effectuées de LEJEUNE-DIRICHLET à RIEMANN et même à LEBESGUE (1). D'où la mise en jeu de la théorie des ensembles, à propos tant du but des recherches — théorie générale des fonctions, — que des moyens utilisés — intégrales de FOURIER.

HANKEL fait dater les mathématiques modernes « de l'instant où DESCARTES, à partir du traitement purement algébrique des équations, s'avança jusqu'à étudier les variations de grandeur que subit une expression algébrique lorsqu'une grandeur désignée en elle de façon générale, parcourt une suite continue de valeurs. » (2) C'est cependant LEIBNIZ qui introduisit le mot même de fonction, et s'il semble en avoir eu une notion assez générale, comme celle de dépendance entre deux grandeurs géométriques données, — « parametri seu rectæ magnitudine constantes... censentur... indifferentiabiles, quemadmodum et ipsa recta tangens vel aliæ nonnullæ functiones ab ea pendentes, v. g.

(1) Dès la préface de la 1^re édit. de ses *Leçons sur l'intégration* (1907), LEBESGUE signale comme une « très importante application » de sa définition « l'étude du développement trigonométrique des fonctions ». En effet, « pour qu'une fonction admette une série de FOURIER, il faut et il suffit qu'elle soit sommable ». LEBESGUE (II), p. 31.

(2) HANKEL (I), p. 45.

perpendiculares ad tangentiam ab axe ad curvam ductæ» [1], — du moins les grandeurs elles-mêmes, puisqu'il s'agit de courbes connues, étaient-elles définies par des expressions analytiques simples, de sorte que JOHANN BERNOULLI pouvait se réclamer de lui pour sa définition, reprise ensuite textuellement par EULER et classique pour le XVIII^e siècle : « on appelle fonction d'une grandeur variable, une quantité composée de quelque manière que ce soit de cette grandeur variable et de constantes » [2]. Ici c'est la composition qui définit la fonction, non la dépendance posée entre termes donnés extra-analytiquement. D'où la limitation des fonctions aux expressions analytiques connues : mais le décalage avec la première définition exigeait un élargissement. Déjà EULER distinguait [3] entre *curvae continuae*, celles dont le domaine coïncide avec le domaine des expressions analytiques, et *curvae discontinuae, seu mixtae seu irregulares*, les « courbes mécaniques » de DESCARTES données graphiquement, par exemple formées de morceaux bout à bout de courbes continues. Le procédé même semblait exclure une définition analytique quelconque (non plus seulement donc limitée aux expressions connues) puisque, comme le note LEBESGUE, « on admettait que deux expressions analytiques égales dans un intervalle sont égales partout » [4]. Mais, à ce renoncement devant le mécanique, le mouvement spontané des mathématiques, en tant qu'expression — et solution — de tout problème posé précisément obligeait de passer outre.

En 1753, Daniel BERNOULLI et EULER étudient le mouvement des cordes vibrantes. Il s'agit de donner à l'équation aux dérivées partielles

(1) *Acta Erudit. Lips.*, 1692.
(2) Joh. BERNOULLI, *Opera omnia*, II, p. 241 (cité par HANKEL, *ibid.*).
(3) *Opusc. math.*, t. I, p. 41, cité par LEBESGUE (II), p. 21 n.
(4) LEBESGUE, *ibid.*

$$\frac{\partial^2 y}{\partial t^2} = K^2 \frac{\partial^2 y}{\partial x^2},$$

la solution la plus générale satisfaisant à un certain nombre de conditions aux limites. En particulier, pour $t = 0$, la fonction $y = f(x, t)$ doit représenter la position initiale de la corde, or, remarque d'ALEMBERT, « la manière ordinaire, pour ne pas dire l'unique, de faire sortir une corde de son état de repos, c'est de la prendre par un de ses points et de la tendre en la tirant, ce qui lui donne la figure de deux lignes droites qui font un angle entre elles » (1). Il est donc « contre les règles de l'analyse » de prétendre, comme faisait EULER, qu'une série trigonométrique pourrait fournir la solution. Cependant en 1807 FOURIER donnait un procédé régulier pour construire cette série ; si l'on pose :

$$f(x) = \sum_{K=0}^{\infty} (a_k \cos k\,x + b_k \sin k\,x)$$

en multipliant les deux membres par $\sin kx$ ou $\cos k\,x$ et en intégrant terme à terme de $-\pi$ à $+\pi$, on obtient :

$$b_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\alpha)\,d\alpha \qquad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\cos k\alpha\,d\alpha$$

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\sin k\alpha\,d\alpha.$$

L'originalité de FOURIER, remarque SACHSE (2), n'est pas — contrairement à ce que dit RIEMANN — dans le procédé de calcul des coefficients, déjà utilisé par LAGRANGE et EULER, aux travaux desquels d'ailleurs FOURIER renvoie lui-même, mais d'avoir vu « qu'une série trigonométrique, avec des coefficients ainsi déterminés, peut représenter une fonction arbitraire quelconque » ; la méthode semblait tellement sûre pour résoudre le problème que pendant un demi-siècle

(1) LEBESGUE, *ibid.*
(2) SACHSE (I), p. 47.

séries trigonométriques et séries de FOURIER furent identifiées, jusqu'à ce que HEINE, en 1870, s'avisât le premier de faire la distinction. D'où la mise au premier plan, pendant presque toute cette période, des études sur l'intégration : non seulement le calcul des coefficients mais la convergence de la série s'y trouvent subordonnés pour DIRICHLET. Quelles conditions, remplies par la fonction arbitraire, suffisent pour qu'une certaine intégration soit possible, tel est le problème dont lui et ses disciples s'occupent. Que résulte-t-il pour la marche d'une fonction de cette propriété d'être représentable par une série trigonométrique quelconque — et cette représentation est-elle univoque, c'est la seconde question posée, de RIEMANN à CANTOR, et qui, attaquant de front la notion de fonction, devait mener directement à la théorie des ensembles.

FOURIER lui-même ne s'est pas préoccupé de démontrer la convergence de sa série : il lui suffit que l'on puisse voir dans des cas particuliers que la série converge effectivement vers la valeur de la fonction correspondante. La première démonstration due à CAUCHY repose sur une confusion entre convergence et absolue convergence : la série de FOURIER converge parce que le rapport entre son terme général et $\frac{A \sin n x}{n}$ (A étant une constante) tend vers $+ 1$ lorsque n croît indéfiniment. « Mais, objecte DIRICHLET en 1829 [1], il est facile de donner l'exemple de séries, dont les termes ne sont pas tous positifs, telles que l'une soit convergente, l'autre divergente quoique le rapport des termes correspondants tende vers $+ 1$:

$$\sum_n \frac{(-1)^n}{\sqrt{n}} \qquad \text{et} \qquad \sum_n \frac{(-1)^n}{\sqrt{n}} \left[1 + \frac{(-1)^n}{\sqrt{n}}\right]. \text{ »}$$

La méthode de DIRICHLET consiste à mettre la somme des n premiers termes de la série de FOURIER sous la forme :

(1) LEJEUNE-DIRICHLET (I), p. 158.

$$\frac{1}{\pi}\int_{-\pi}^{+\pi} \varphi(\alpha)\,\frac{\sin\dfrac{2n+1}{2}(\alpha-x)}{2\sin\dfrac{\alpha-x}{2}}\,d\alpha \qquad (1)$$

($\varphi(\alpha)$ étant la fonction arbitraire). « Tout se réduit à déterminer la limite dont cette intégrale approche quand n croît indéfiniment ». Lorsque $\varphi(\alpha)$ est continue, monotone et finie, la limite est bien $\varphi(x)$. Mais la notion d'intégrale telle que CAUCHY l'avait précisée permet d'aller plus loin :

1° l'intégrale d'une fonction dans un intervalle étant la somme des intégrales étendues à un nombre fini d'intervalles partiels en lesquels le premier est divisé, on peut accorder à $\varphi(x)$ un nombre fini de maxima et de minima ;

2° Si $\int_a^x f(\alpha)\,d\alpha$ tend vers une limite déterminée lorsque x tend vers b, point singulier pour f (point de discontinuité ou point où f est infinie), CAUCHY définit par cette limite l'intégrale $\int_a^b f(\alpha)\,d\alpha$. En particulier si $f(x)$ tend vers une limite déterminée, $f(b-0)$ ou $f(b+0)$ (notation de DIRICHLET), pour x tendant vers b en croissant ou décroissant, c'est-à-dire si b est un point de discontinuité de 1re espèce (condition toujours réalisée, ainsi que le remarque RIEMANN lorsque $f(x)$ n'a qu'un nombre fini de maxima et de minima), les deux intégrales $\int_a^{b-o} f(\alpha)\,d\alpha$ et $\int_{b+o}^c f(\alpha)\,d\alpha$ auront des valeurs bien déterminées, l'intégrale $\int_a^c f(\alpha)\,d\alpha$ (avec $a < b < c$) sera définie comme leur demi-somme.

En utilisant donc seulement les résultats déjà connus, DIRICHLET peut affirmer que la série (1) est convergente si $\varphi(x)$ remplit les trois conditions :

1° rester finie dans l'intervalle $-n \ldots +n$;
2° n'avoir qu'un nombre fini de discontinuités (de 1re espèce) ;
3° n'avoir qu'un nombre fini de maxima et de minima.

Pour que la série représente effectivement $\varphi(x)$, RIEMANN remarque (1) qu'il est nécessaire d'ajouter une 4e condition : $\varphi(x)$ est égale à la valeur moyenne de ses deux valeurs limites $\varphi(x+0)$ et $\varphi(x-0)$ aux points de discontinuité ; sans quoi il faut renoncer à la coïncidence entre la série et la fonction en ces points (2). En revanche, aucune des trois premières conditions n'est irréductible : « il nous resterait à considérer les cas où les suppositions, que nous avons faites sur le nombre des solutions de continuité et sur celui des valeurs maxima et minima, cessent d'avoir lieu. Ces cas singuliers peuvent être ramenés à ceux que nous venons de considérer... Mais la chose, pour être faite avec toute la clarté qu'on peut désirer, exige quelques détails liés aux principes fondamentaux de l'analyse infinitésimale et qui seront exposés dans une autre note (3). »

Cette promesse de révision des principes, c'est-à-dire d'une nouvelle théorie de l'intégration, n'a pas été tenue. La seule restriction levée fut en 1837 (4) celle dont justement il n'était pas question en 1829, l'obligation de rester finie pour $\varphi(x)$. Si $\int_x^y \varphi(\alpha)\, d\alpha$ est une fonction finie et continue des variables x et y, ($\varphi(\alpha)$ pouvant devenir infinie entre x et y), quelles que soient les valeurs prises par x et y entre $-\pi$ et π (c'est-à-dire si $\varphi(\alpha)$ est intégrable au sens CAUCHY entre $-\pi$ et $+\pi$), l'intégrale (1) prise non entre $-\pi$ et π mais entre des valeurs arbitrairement proches encadrant la valeur rendant $\varphi(x)$ infinie, peut être rendue aussi petite qu'on le veut. Il suffit alors, suivant la remarque de LIPSCHITZ, s'il n'y a qu'un nombre fini de points où $\varphi(x)$ est infinie, de les entourer d intervalles arbitrairement petits : leur contribution

(1) RIEMANN (I).

(2) Plus tard, SCHÄFFLI (I) et DU BOIS REYMOND (V) devaient revenir sur le caractère indépendant (par rapport aux autres) de cette nouvelle condition.

(3) LEJEUNE-DIRICHLET (I), p. 169.

(4) LEJEUNE-DIRICHLET (III).

dans l'intégrale (1) prise de $-\pi$ à π est négligeable. La condition 1° se transforme donc en : n'être infinie qu'en un nombre fini de points. Pour l'affaiblissement de la condition 2° DIRICHLET donne une indication dès 1829 : « il est nécessaire que la fonction $\varphi(x)$ soit telle que, si l'on désigne par a et b deux quantités quelconques comprises entre $-\pi$ et π, on puisse toujours placer entre a et b d'autres quantités r et s assez rapprochées pour que la fonction reste continue dans l'intervalle de r à s. On sentira facilement la nécessité de cette restriction en considérant que les différents termes de la série sont des intégrales définies et en remontant à la notion fondamentale des intégrales. On verra alors que l'intégrale d'une fonction ne signifie quelque chose qu'autant que la fonction satisfait à la condition précédemment énoncée. On aurait un exemple d'une fonction qui ne remplit pas cette condition, si l'on supposait $\varphi(x)$ égale à une constante déterminée c lorsque la variable x obtient une valeur rationnelle, et égale à une autre constante d lorsque cette variable est irrationnelle » (1). L'exemple ne suffit pas à éclaircir l'obscurité du passage : s'il est évident qu'une fonction dont les discontinuités sont partout denses dans un intervalle n'est pas intégrable (au sens CAUCHY ou RIEMANN) dans cet intervalle, la densité (la condition imposée revient à dire que l'ensemble des points de discontinuité est non dense entre $-\pi$ et π) ne joue cependant pas de rôle dans l'intégration telle qu'on l'envisageait alors. Le but était d'enfermer les points de discontinuité dans des intervalles dont la contribution puisse être négligée dans l'intégrale totale. Il semble qu'il y ait eu dans l'esprit de DIRICHLET interférence entre deux considérations — que la théorie des ensembles devait distinguer : — celle de la mesure, qui devait mener à la notion de groupe intégrable, introduite par DU BOIS REYMOND ; celle de la répartition des points de

(1) LEJEUNE-DIRICHLET (I), p. 169.

discontinuité autour de points limites, notion d'ensemble réductible précisée par CANTOR. Mais pour utiliser la 1re, l'intégrale de RIEMANN était indispensable. Aussi dans son commentaire de ce passage LIPSCHITZ choisit-il la seconde, sans même apercevoir, ainsi que le note MONTEL (1), qu'il transforme la condition de DIRICHLET. L'explication n'est d'ailleurs guère précise : en supposant que l'ensemble des points de discontinuité ne comporte qu'un nombre fini de points limites, on peut enfermer ceux-ci dans un nombre fini d'intervalles de longueur arbitrairement petite ; dans les intervalles qui subsistent, la fonction n'a qu'un nombre fini de discontinuités ; pour les premiers intervalles les intégrales correspondantes peuvent être rendues aussi petites qu'on le veut ; en effet « chacune de ces intégrales est la somme d'une infinité d'intégrales constituées de manière que cette somme même puisse être rendue aussi petite qu'on le veut » (2). On doit voir là, sans doute comme le propose LEBESGUE (3), l'application d'une définition : l'intégrale d'une fonction $f(x)$ entre deux points limites α, β de ses points de discontinuité est :

$$\int_{\alpha}^{\beta} f(x)\,dx = \lim_{\substack{\varepsilon = o,\\ \eta = o}} \int_{\alpha+\varepsilon}^{\beta-\eta} f(x)\,d$$

lorsque cette limite existe, et il n'y a d'intégrale que si cette limite existe.

Contrairement à l'opinion exprimée par RIEMANN (4), LIPSCHITZ montre également (5) que la 3e condition peut être supprimée si en tous les points où la fonction oscille, l'oscillation $\omega = | f(x + \delta) - f(x) |$ est telle que :

$$\omega < k\,\delta\alpha$$

(1) LIPSCHITZ (I), p. 284, n. 2.
(2) *Ibid.*, p. 284.
(3) LEBESGUE (I), p. 10.
(4) RIEMANN (I), p. 224 : « Si la dernière supposition (que la fonction n'a pas une infinité de maxima et de minima) est abandonnée, les deux théorèmes de DIRICHLET sont insuffisants pour décider de la question. »
(5) LIPSCHITZ (I), p. 285-295.

(k étant une constante, α un nombre > 0), exigence encore trop forte puisque, comme le fait remarquer SACHSE (1), LIPSCHITZ se sert seulement de celle-ci :

$$\lim_{\delta=o} \omega \log \delta = o$$

— « hac in re sunt nervi demonstrationis nostræ (2) » — adoptée ensuite par U. DINI. Ici encore interviennent des notions de la théorie des ensembles : « ou bien on peut intercaler entre a et b des nombres $r - \delta$ et $r + \delta$ tels qu'il y ait un nombre fini de maxima et minima dans $(a, r - \delta)$ et $(r + \delta, b)$, mais un nombre infini entre $r - \delta$ et $r + \delta$, quelque petite que soit la distance 2δ de ces deux nombres ; ou bien quelles que soient les quantités r et s situées dans l'intervalle (a, b) et à une distance finie l'une de l'autre, on ne peut jamais trouver un nombre fini de maxima et de minima dans l'intervalle (r, s) ; ou bien l'intervalle considéré contient un nombre fini d'intervalles non nuls dans lesquels l'une ou l'autre hypothèse est réalisée (3). » Répartition en trois cas, qui, comme le note MONTEL (4), ne joue d'ailleurs aucun rôle dans la démonstration. Mais elle éclaire la confusion précédente entre ensemble à 1er dérivé fini et ensemble non dense. Si, en effet, on remplace la définition du premier cas par celle de DIRICHLET : ensemble des maxima ou minima non dense, et dans celle du troisième cas le mot fini par « fini ou infini », la répartition est cette fois exhaustive. C'est la considération précise du rapport entre points et intervalles les enfermant qui fait défaut à DIRICHLET et LIPSCHITZ : elle ne devait s'éclaircir, grâce à la notion d'ensemble parfait non dense, par exemple, qu'après les travaux de RIEMANN sur l'intégration.

(1) SACHSE (I), p. 62.

(2) LIPSCHITZ (I), p. 294, texte original de LIPSCHITZ ; les p. renvoient à la traduction MONTEL.

(3) P. 286.

(4) *Ibid.*, p. 286, n. 1.

La partie positive du mémoire de RIEMANN commence, en effet, par une nouvelle définition de l'intégrale, dont les propriétés permettent de démontrer à la fois un des lemmes fondamentaux, et directement, quelques-uns des principaux résultats, comme le théorème que lorsque la fonction est intégrable dans ce nouveau sens, les coefficients de la série trigonométrique correspondante deviennent infiniment petits. RIEMANN transforme, ainsi que le note LEBESGUE (1), le procédé opératoire donné par CAUCHY pour calculer l'intégrale d'une fonction continue en définition de l'intégrale d'une fonction quelconque. Si la somme :

$$S = \delta_1 f(x_1) + \delta_2 f(x_2) + \ldots + \delta_n f(x_n)$$

(les δ_i représentent des intervalles partiels en lesquels est divisé l'intervalle d'intégration $(a\ b)$, les x_i des points quelconques à l'intérieur des δ_i) tend vers une limite, quels que soient les intervalles δ_i « lorsqu'ils deviennent tous infiniment petits, sa valeur s'appelle :

$$\int_a^b f(x)\, dx.$$

Si... [la somme S] ... n'a pas cette propriété $\int_a^b f(x)\, d\alpha$ n'a aucune signification (2). » D'où une conséquence qui montrera l'élargissement par rapport à CAUCHY : si on appelle ω_i l'oscillation de $f(x)$ dans l'intervalle δ_i (c'est-à-dire la différence entre la limite supérieure et la limite inférieure dans cet intervalle), la condition nécessaire et suffisante pour que S ait une limite, c'est-à-dire pour que $f(x)$ soit intégrable, lorsque $f(x)$ reste finie dans $(a\ b)$ est que l'on puisse trouver une division de $(a\ b)$ en intervalles δ telle que les δ_i pour lesquels l'oscillation ω_i correspondante est supérieure à un nombre quelconque donné σ aient une longueur totale aussi petite qu'on veut. Ainsi se trouve préci-

(1) LEBESGUE (I), p. 24.
(2) RIEMANN (I), p. 225.

sée et étendue l'idée de DIRICHLET-LIPSCHITZ : non seulement il suffit que les points de discontinuité d'une fonction intégrable puissent être enfermés dans un nombre fini d'intervalles de longueur totale arbitrairement petite, c'est-à-dire forment ce que DU BOIS REYMOND devait appeler un groupe intégrable (ensemble évidemment non dense, mais en plus de mesure nulle, ce que n'avait pas distingué DIRICHLET, en tous cas plus général que l'ensemble, à 1er dérivé fini, auquel s'était borné LIPSCHITZ), mais encore il n'est nécessaire que de considérer l'importance des discontinuités relativement aux intervalles où elles apparaissent : si les points où les sauts (1) sont inférieurs à une quantité donnée forment un groupe intégrable, l'intégrale existe. Ainsi la série :

$$f(x) = \sum_{n=1}^{\infty} \frac{(n\,x)}{n^2},$$

où $(n\,x)$ représente l'excès de $n\,x$ sur l'entier le plus voisin (on pose $(n\,x) = 0$ lorsque $n\,x$ est à égale distance de deux entiers), convergente pour tout x, est discontinue pour toute valeur rationnelle de x qui, sous forme de fraction irréductible, a un dénominateur pair ; les points de discontinuité sont partout denses. Mais, d'une part, les discontinuités sont de première espèce, d'autre part le nombre des sauts supérieurs à une grandeur donnée est fini. En réalité, RIEMANN n'a pas directement besoin du groupe intégrable : tout ce qui concerne son intégrale (lorsque la fonction à intégrer est finie) se traite, comme le montre son travail, à l'aide de propriétés d'intervalles. Si la théorie des ensembles est ici latente, elle n'a pas trouvé encore obligation d'apparaître.

(1) Le saut de $f(x)$ au point x est la limite vers laquelle tend

$$f(x+\varepsilon) - f(x-\eta)$$

lorsque ε et η tendent vers 0 : $f(x+0) - f(x-0)$.

A l'étude des conditions suffisantes — à laquelle s'étaient exclusivement consacrés ses prédécesseurs, — RIEMANN se propose de substituer celle des conditions nécessaires pour la représentabilité d'une fonction par une série trigonométrique, « et parmi celles-ci ensuite de choisir les conditions suffisantes ». SACHSE observe que le problème traité en fait est différent (et ceci expliquerait l'incorrection de la phrase puisque les conditions suffisantes comprennent évidemment toutes les conditions nécessaires) : il s'agit des conditions nécessaires pour qu'une fonction soit égale à une série trigonométrique partout où celle-ci converge. Mais peut-on parler encore de représentation, s'il y a, par exemple, dans un intervalle une infinité de points où la série ne converge pas ? SACHSE propose de se restreindre à un nombre fini d'exceptions. Le résultat fondamental obtenu par RIEMANN est que les conditions qu'il trouve sont également suffisantes, et « que, lorsque les coefficients de la série deviennent infiniment petits $\left[\text{avec } \frac{1}{n}\right]$, sa convergence pour une certaine valeur de x ne dépend que du comportement de la fonction au voisinage immédiat de cette valeur ». D'où la nécessité, d'une part, d'analyser ce voisinage — de faire donc intervenir implicitement la notion de point limite, — d'autre part, pour la représentation d'une fonction donnée, de rechercher des conditions suffisantes. RIEMANN montre que si les trois conditions de DIRICHLET transformées (1. la fonction ne doit pas être infinie ; 2. elle doit être intégrable au sens RIEMANN ; 3. elle doit ne pas comporter une infinité de maxima et de minima) sont remplies, les conditions nécessaires le sont également, les coefficients de la série sont ceux de FOURIER et la série converge en tous les points qui n'ont pas dans leur voisinage une infinité de points de discontinuité de la fonction. Reste l'hypothèse que les coefficients ne deviennent pas infiniment petits avec $\frac{1}{n}$ pour toute va-

leur de x : on a vu qu'elle ne peut pas se présenter lorsque $f(x)$ est intégrable. Dans les autres cas, et si $f(x)$ ne comporte pas une infinité de maxima et de minima, on peut appliquer, par un artifice immédiat, les raisonnements précédents : la série ne peut converger qu'en un nombre fini de points ; on ne peut guère parler de représentation.

Mais est-ce la propriété d'être égale, là où elle converge, à une fonction ne possédant qu'un nombre fini de maxima et de minima, qui oblige une série trigonométrique dont les coefficients ne tendent pas vers o, à ne converger qu'en un nombre fini de points ? Plus généralement, que résulte-t-il de cette propriété pour l'aspect même de la série ? D'autre part, y a-t-il une différence de portée entre les raisonnements de RIEMANN pour l'établissement des conditions nécessaires — qui concernent un développement trigonométrique quelconque — et les démonstrations des conditions suffisantes, qui font intervenir les propriétés des coefficients de FOURIER ? Double façon d'arriver au problème de l'unicité du développement. Le procédé même de FOURIER garantit la réponse, sous réserve qu'on puisse intégrer terme à terme, ce qu'admettaient FOURIER et CAUCHY. Avec la découverte de la condition de convergence uniforme un nouvel examen s'impose.

En 1822, CAUCHY énonçait dans son *Cours d'Analyse* que « lorsque les différents termes d'une série sont des fonctions d'une même variable x, continues par rapport à cette variable dans le voisinage d'une valeur particulière pour laquelle la série est convergente, la somme S de la série est aussi, dans le voisinage de cette valeur particulière, fonction continue de x » (1). Théorème dont, en 1826, ABEL montrait « qu'il souffre des exceptions » (2), par exemple la série :

(1) CAUCHY (I), p. 132.
(2) ABEL (I), p. 312.

$$f(\varphi) = \sin\varphi - \frac{\sin 2\varphi}{2} + \frac{\sin 3\varphi}{3} \cdots$$

discontinue pour toute valeur $\varphi = (2m + 1)\pi$ (m étant un nombre entier). En 1848, SEIDEL remarquait qu'il était également « en contradiction avec les travaux de DIRICHLET puisque la série trigonométrique converge même quand elle représente une fonction discontinue » (1). D'une façon générale, « un agrégat de grandeurs continues ne pourrait jamais représenter des grandeurs discontinues si on fait intervenir la forme de l'infini ». En réalité, « si on a une série convergente qui représente une fonction discontinue de x et dont les termes sont des fonctions continues de x, on doit, dans le voisinage de la valeur où la fonction *saute*, trouver des valeurs de x pour lesquelles la série converge *arbitrairement lentement* ». Si la vitesse de convergence est, au contraire, égale pour tous les points d'un intervalle, c'est-à-dire, si quel que soit x, on a :

$$\text{pour } n > N \qquad |\, u_{n+m}(x) - u_n(x) \,| < \varepsilon$$

(ε étant un nombre arbitraire donné auquel on fait correspondre l'entier N, m étant quelconque), il y a convergence uniforme (2) : WEIERSTRASS démontrait peu après que c'était la condition nécessaire et suffisante pour qu'on puisse intégrer terme à terme. En complément du théorème de SEIDEL, DU BOIS REYMOND (3) établissait que, même aux points de continuité d'une fonction, la convergence de sa série peut n'être pas uniforme : de sorte que, même pour les fonctions continues, l'unicité du développement trigonométrique n'est pas garantie par le procédé de FOURIER.

(1) SEIDEL (I), p. 382.

(2) CAUCHY rectifia lui-même son erreur dans une note aux *Comptes rendus*, en 1853. « Il est facile de voir comment on doit modifier l'énoncé du théorème pour qu'il n'y ait plus lieu à aucune exception » et il donne exactement la définition de la convergence uniforme (CAUCHY (II), p. 455).

(3) DU BOIS REYMOND (X).

HEINE (1) introduit la notion de convergence uniforme « en général », c'est-à-dire à l'exception des voisinages immédiats d'un nombre fini de points de discontinuité. Une fonction finie, ne possédant pas une infinité de maxima et de minima et continue « en général » (c'est-à-dire sauf en un nombre fini de points) ne peut être développée que d'une façon en série trigonométrique qui soit, en général, uniformément convergente. Autrement dit, une série trigonométrique en général uniformément convergente ne peut représenter o que si tous ses coefficients sont nuls. HEINE le démontre en intégrant deux fois terme à terme dans les intervalles $(x_n \ldots x_{n+1})$, à l'intérieur desquels il y a convergence uniforme, tandis qu'elle cesse pour les extrémités : on obtient des fonctions linéaires $F(x) = a_n x_n + b_n$.

Or, en vertu d'une des conditions nécessaires établies par RIEMANN, on doit avoir :

$$\lim_{\alpha = o} \frac{F(x+\alpha) + F(x-\alpha) - 2F(x)}{2\alpha} = o \qquad (2)$$

Il en résulte que tous les a_n et les b_n sont égaux entre eux, et par suite tous nuls (puisque pour $x = o$ on a $b_o = o$). La série qui représente $F(x)$ est donc nulle partout : comme elle est également partout uniformément convergente, on peut lui appliquer le procédé de FOURIER, tous ses coefficients sont nuls.

L'hypothèse de la convergence uniforme n'est pas essentielle au raisonnement si on utilise un autre résultat de RIEMANN, établi avec la seule supposition que les coefficients tendent vers o avec $\frac{1}{n}$.

CANTOR (2) commence par démontrer que c'est le cas pour toute série trigonométrique convergente dans tout un in-

(1) HEINE (II), p. 353.
(2) CANTOR (I) et (III).

tervalle (à l'exception de points isolés) répondant ainsi à la première question laissée en suspens par RIEMANN. Mais le théorème n'est même pas indispensable comme fit observer KRONECKER [1] en reprenant un artifice que RIEMANN avait d'ailleurs déjà utilisé. Une des conditions nécessaires de RIEMANN est que, si on pose :

$$F(x) = b_0 \frac{x^2}{4} - \sum_{n=1}^{\infty} \frac{a_n \sin x + b_n \cos n x}{n^2} \qquad (3)$$

(série obtenue par deux intégrations consécutives terme à terme de la série donnée $f(x)$) on a pour la dérivée seconde généralisée de $F(x)$:

$$\lim_{\alpha = 0} \frac{F(x + 2\alpha) + F(x - 2\alpha) - 2F(x)}{4\alpha^2} = f(x) \qquad (4)$$

Ici la limite doit être o. Or SCHWARZ [2] démontre que lorsque la dérivée seconde généralisée d'une fonction est nulle, la fonction est linéaire : on retrouve donc les conditions où se plaçait HEINE. S'il n'y a pas de points d'exception (points de discontinuité — c'est-à-dire ici points où $f(x)$ cesse d'être = o — ou points de non convergence pour la série), on voit sans peine que l'équation (3) devient :

$$K = \sum_{n=1}^{\infty} \frac{a_n \sin n x + b_n \cos n x}{n^2}$$

le second membre est évidemment uniformément convergent lorsque les coefficients a_n et b_n tendent vers o avec $\frac{1}{n}$, on peut donc appliquer le procédé FOURIER, et l'on trouve $a_n = b_n = o$ quel que soit n. S'il y a des exceptions, le raisonnement de HEINE permet à CANTOR, grâce à la relation (2), de jeter en quelque sorte un pont entre les intérieurs de deux intervalles séparés par l'une de ces exceptions : les deux

(1) CANTOR (III), p. 84.
(2) SCHWARZ (I).

fonctions linéaires correspondantes sont identiques. Donc un intervalle ne comportant à son intérieur qu'un nombre fini de points exceptionnels correspondra à une seule fonction linéaire : ce seront alors les points limites des points exceptionnels qui sépareront deux intervalles semblables, et qui seront surmontés de la même façon grâce à la relation (2) : et ainsi de suite. « Une fonction discontinue $f(x)$ qui pour toutes les valeurs de x, qui correspondent aux points d'un ensemble de ν^e espèce donné dans l'intervalle $(0 \ldots 2\pi)$, est différente de o ou indéterminée, pour toutes les autres valeurs de x, est nulle, ne peut pas être représentée par une série trigonométrique (1). »

Tel est l'acte de naissance, au moins de la théorie cantorienne des ensembles. On a vu comment s'y inscrit la définition du continu numérique. Mais l'important n'est pas là : la notion d'ensemble, la distinction entre les modes de répartition des points dans un intervalle avaient été imposées antérieurement par la théorie des fonctions. Aux environs de 1870 les travaux se multiplient qui s'en servent : HANKEL, à propos des *Fonctions infiniment souvent oscillantes et discontinues*, énonce son principe de condensation des singularités et définit oscillation en un point et oscillation dans un intervalle. DU BOIS REYMOND prétend avoir, dès cette époque, utilisé la distinction entre *pantachies* (ensembles partout denses) et *apantachies* (ensembles non denses). Il n'y a théorie des ensembles qu'à l'apparition, non de notions, mais d'un mode de raisonnement original : tel est celui dont CANTOR faisait pour la première fois usage en généralisant le théorème de HEINE ; de lui — que l'on retrouvera sous le nom d'itération transfinie — dépend une partie essentielle de l'édifice futur.

(1) CANTOR (V), p. 101.

* * *

Un autre thème intervient cependant avec les travaux de DU BOIS REYMOND sur la croissance des fonctions. A propos de l'intégrale de FOURIER :

$$\int_{-\pi}^{\pi} f(x) \cos n\,(x - a)\,dx$$

dans le cas où n est une grandeur continue, il est amené à comparer les limites vers lesquelles tendent des expressions lorsque les variables y deviennent infinies de façon indépendante. Pour $z = f(xy)$, par exemple, il faudra distinguer les deux limites (1) :

$$z_1 = \lim_{x\infty} \lim_{y\infty} f(xy) \quad \text{et} \quad z_2 = \lim_{y\infty} \lim_{x\infty} f(xy)$$

De même pour les façons de tendre vers o. Après un essai de traitement géométrique, il jette les bases d'un *calcul infinitaire* permettant de comparer la croissance de deux fonctions différentes. L'ordre entre elles sera établi par les relations :

$$f(x) \prec \varphi(x) \qquad \text{si } \lim_{x\infty} \frac{f(x)}{\varphi(x)} = 0$$

$$f(x) \sim \varphi(x) \qquad \text{si } \lim_{x\infty} \frac{f(x)}{\varphi(x)} = \text{une quantité finie}$$

Ainsi pourrait se constituer un continu infinitaire, premier exemple d'ensemble, étudié mathématiquement, dont les éléments ne soient ni des points ni des nombres (2). « Nous nous proposons de développer quelques vues générales... concernant la suite continue des fonctions ordonnées sui-

(1) DU BOIS REYMOND (I).

(2) A peu près à la même époque THOMAE étudiait les ordres de décroissance de fonctions tendant vers o.

vant les limites de leurs quotients et l'analogie de cette suite avec la suite des nombres réels (1). » Aux nombres positifs correspondent les fonctions devenant infinies, aux nombres négatifs les fonctions devenant nulles, à 0 les fonctions restant finies. Il est possible de jalonner ce continu, comme au moyen d'entiers, avec ce que BOREL (2) devait appeler une « échelle de types croissants ». Il suffit de prendre une fonction croissante, l'exponentielle, par exemple, et son inverse la fonction logarithmique, et d'itérer indéfiniment l'une et l'autre :

$$l\,x = \underbrace{\log \log \cdots \log}_{\nu}\, x \qquad \text{et} \qquad e^{e^{\cdot^{\cdot^{\cdot^{e^x}}}}}$$

Entre deux ordres infinitaires différents on peut toujours insérer une infinité d'ordres intermédiaires, il suffit de multiplier la fonction d'ordre inférieur par une suite de fonctions croissant encore moins vite ; par exemple x^μ multiplié par lx, $lx.l_2x$,... donnera des ordres tous inférieurs à celui de $x^{\mu+1}$. Mais ici s'arrête l'analogie avec le continu numérique. Le mode de raisonnement original donnant ici un début de théorie des ensembles, le procédé de la diagonale, montre en effet, que d'abord l'axiome d'ARCHIMÈDE n'est pas satisfait : étant donnée une suite infinie quelconque d'ordres infinitaires, il est toujours possible de construire une fonction d'ordre inférieur à tous les éléments de la suite. Si l'on prend une suite de courbes croissantes, en les coupant par des parallèles à oy, on voit que la courbe joignant l'intersection de la 1re courbe et de la 1re parallèle avec celle de la 2e courbe et de la 2e parallèle, etc... répond à la question (fig. 1). D'autre part (3), étant données deux suites

(1) DU BOIS REYMOND (II), p. 338.
(2) BOREL (I), p. 114.
(3) DU BOIS REYMOND (VII).

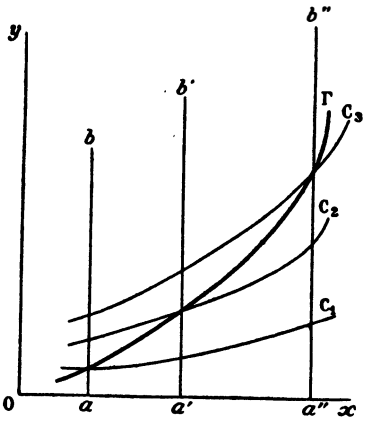


FIG. 1.

de fonctions balayant le même intervalle infinitaire, il est impossible de représenter la même valeur de cet intervalle dans les deux suites. Soit, par exemple :

$$x^a \text{ et } (x \log x)^b \left| \begin{array}{l} \text{pour } a \text{ et } b < \frac{1}{10} \text{ les deux fonctions sont} \prec x^{\frac{1}{4}} \\ \text{pour } a \text{ et } b > \frac{9}{10} \text{ toutes deux} \succ x^{\frac{3}{4}} \end{array} \right.$$

cependant quels que soient a et b on ne peut jamais avoir :

$$x^a \sim (x \log x)^b$$

D'où l'affirmation, assez confuse, par DU BOIS REYMOND [1] que le continu fonctionnel est plus riche, de puissance plus élevée que le continu numérique, et une discussion encore plus obscure avec PRINGSHEIM [2] (qui prétendait le continu fonctionnel analogue à un continu numérique pluridimensionnel). La seule justification de cette affirmation se trouverait dans les travaux contemporains ou postérieurs de

(1) DU BOIS REYMOND (VII).
(2) PRINGSHEIM (I).

STOLZ (1), BETTAZZI (2), HAHN (3) sur les systèmes non archimédiques : ceux-ci peuvent, en effet, toujours se représenter par des nombres complexes dont les unités, en général en nombre infini, forment une suite ordonnée. Mais le but du calcul infinitaire est de comparer les uns aux autres les degrés de croissance : la pluridimensionalité serait absurde.

BOREL clarifie le débat en montrant (4), d'une part, que l'ensemble des ordres de croissance, ayant même puissance que celui des fonctions continues, a la puissance du continu linéaire ; d'autre part, que la comparaison avec celui-ci ne vaut que si on part d'une échelle donnée de types croissants et si on se borne à la compléter uniquement par multiplication de fonctions qui y figurent, exactement comme on obtient les nombres rationnels à partir des entiers. Dans ce cas il sera même possible de définir des coupures, et une *fonction idéale* correspondante. DU BOIS REYMOND en avait déjà eu l'idée à propos de l'intégrale de FOURIER : il s'agissait d'examiner la convergence d'intégrales de la forme $I = \int_{\varepsilon}^{\alpha} \frac{d\alpha}{\varphi(\alpha)}$ lorsque ε tend vers o.

On voit que si on a :

$$\varphi(\alpha) \sim \frac{1}{l\frac{1}{\alpha} \cdot l_2 \frac{1}{\alpha} \dots l_r \frac{1}{\alpha}} \qquad \lim I = \infty$$

Si, au contraire,

$$\varphi(\alpha) \sim \frac{1}{l\frac{1}{\alpha} \cdot l_2 \frac{1}{\alpha} \dots \left(l_r \frac{1}{\alpha}\right)^{1+\mu}} \qquad \lim I \text{ est finie}$$

(1) STOLZ (I).
(2) BETTAZZI (I).
(3) HAHN (I).
(4) BOREL (I), p. 117-119.

(μ étant un nombre positif quelconque). La coupure entre les degrés infinitaires faisant diverger et converger l'intégrale, définira la fonction idéale $\tau(\alpha)$. On démontre qu'il est impossible que $\tau(\alpha)$ soit de même ordre infinitaire qu'une fonction logarithmique quelconque. Pour $\varphi(\alpha) = \tau(\alpha)$, l'intégrale ne diverge ni ne converge, «représentation choquante pour le sens commun ». Mais « le miracle n'est pas plus grand... que d'admettre de nouveaux nombres entre les nombres rationnels pourtant denses à volonté » (1).

Ainsi échoue à demi cette première tentative d'application de la théorie des ensembles à un problème mathématique. L'ambition de DU BOIS REYMOND était d'embrasser tout le système des ordres de croissance dans un seul « continu » d'ordre nouveau. Mais il aurait fallu posséder des moyens de caractériser celui-ci. Or la notion claire du continu linéaire n'était même pas encore dégagée : le procédé de la diagonale seul se trouvait insuffisant. D'autres instruments propres à la théorie des ensembles étaient nécessaires, que justement à cette époque, dans un tout autre domaine, CANTOR et DEDEKIND travaillent à mettre au point.

(1) DU BOIS REYMOND (IX), p. 165.

CHAPITRE II

LA CRÉATION CANTORIENNE

Deux écrits méritent de figurer avant ceux de CANTOR dans une histoire de la théorie parce qu'essayant les premiers de soumettre à un traitement systématique les ensembles en tant que tels : les *Paradoxes de l'Infini* de BOLZANO et un fragment de DEDEKIND, non publié de son vivant : *Propositions générales sur des espaces.*

« BOLZANO, écrit CANTOR en 1883, est peut-être le seul chez lequel les nombres proprement infinis jouissent d'un certain droit de cité ; au moins en parle-t-il à plusieurs reprises (1). » C'est par l'isolement de la notion d'ensemble que BOLZANO peut, en effet, distinguer de l'infini improprement dit — l'infini de l'analyse qui ne concerne qu'une possibilité de variation et représente des nombres finis pris aussi grands qu'on veut — le véritable infini, sur lequel jusque là rien que de vague ou d'absurde n'avait été dit. Outre la définition générale d'un ensemble, le livre apporte la notion de correspondance biunivoque ; une ébauche d'arithmétique des ensembles, un essai de définition abstraite du continu. « Un certain système (INBEGRIFF) de choses est un tout consistant en certaines parties... Un système dans lequel l'ordre des parties est indifférent s'appelle un ensemble (2). » La notion de parties est encore vague : si ce sont « les unités d'une certaine espèce A..., c'est-à-dire des objets subsumés sous un concept A » l'ensemble est dit une

(1) CANTOR (XV), p. 180.
(2) BOLZANO (II), p. 4.

« multiplicité » (1) (VIELHEIT). Ainsi l'ensemble est considéré comme une réalité existante en soi, indépendamment de nos moyens de l'atteindre. Il suffit qu'on puisse décider si un élément donné lui appartient ou non. « Nous avons déterminé de la façon la plus complète l'ensemble... infini des points entre m et n, aussitôt que nous avons déterminé les deux points m et n. Car alors, par ce peu de mots, il est déjà décidé exactement pour tout autre point s'il appartient ou non à cet ensemble (2). » D'où la possibilité qu'un ensemble soit infini en acte, c'est-à-dire que nous ne puissions l'embrasser du regard et même que nous ne sachions l'engendrer. Des relations entre de tels ensembles — un calcul de l'infini — peuvent être établies grâce d'abord à la mise en correspondance biunivoque : « deux ensembles sont dans un tel rapport l'un avec l'autre que pour toute partie a du premier A, nous pouvons, d'après une certaine règle, trouver une partie b du second B et que toutes les paires $(a + b)$ ainsi formées contiennent toute chose qui se trouve dans A ou dans B et chacune seulement une fois » (3). L'addition et la multiplication (contrairement à ce que croyait CANTOR) (4) sont correctement définies : « il peut y avoir des systèmes dont les parties... sont elles-mêmes des systèmes. Parmi eux tels systèmes que rien en eux d'essentiel n'est changé si on considère les parties des parties comme parties du tout. Je les appelle... sommes » (5). On n'introduit la multiplication que pour l'ensemble des entiers, représenté par la notation :

(1) *Ibid.*

(2) *Ibid.*, p. 38.

(3) P. 31.

(4) « Je ne m'accorde en rien avec la façon dont il les traite [les nombres infinis], sans pouvoir donner d'eux une définition correcte et je considère par exemple les §§ 29-33 [où se trouve la définition du produit] comme inconsistants et erronés ». CANTOR (XV), p. 180.

(5) *Ibid.*, p. 4.

$$\overset{0}{N} = 1^0 + 2^0 + 3^0 + \ldots$$

(les ^{0} et les $+$ indiquent qu'il ne s'agit pas d'une somme arithmétique), le produit de cet ensemble par lui-même sera :

$$(\overset{0}{N})^2 = 1^0.\overset{0}{N} + 2^0.\overset{0}{N} + 3^0.\overset{0}{N} + \ldots$$

à chaque élément du multiplicateur est substitué un ensemble équivalent au multiplicande [1]. Ici, comme de façon générale, il n'y a pas dégagement suffisant du procédé par rapport aux exemples où il est saisi intuitivement. D'où un flottement à propos d'abord de la notion de partie — sous ensemble — et de l'égalité de deux ensembles. Les prédicats *fini* et *infini* ne sont déterminés que pour les multiplicités (VIELHEITEN) dont on connaît la loi d'engendrement, qui constituent des séries : lorsque la multiplicité peut être obtenue tout entière par adjonction successive d'un élément, elle est finie ; « une multiplicité telle que tout ensemble fini en représente une partie est appelée multiplicité infinie. » [2] Quoique le premier exemple d'ensemble infini ne soit pas emprunté aux mathématiques — ensemble de toutes les propositions — l'image de la suite des nombres (avec un ordre) est dominante ; la mise en correspondance biunivoque est ici masquée par le processus de numération. Bien plus, le paradoxe des ensembles infinis — que DEDEKIND transformera en définition de ces ensembles — amène à diminuer son importance, là où elle est expressément utilisée : on peut mettre en correspondance biunivoque les éléments d'un ensemble et ceux d'une de ses parties. Donc l'égalité de deux ensembles devra être définie autrement : « une autre considération doit intervenir..., savoir que les deux ensembles ont mêmes principes de détermination, par

(1) P. 45.
(2) P. 6.

exemple le même mode d'engendrement [1] ». Ici encore c'est la référence aux exemples qui fait faire fausse route : les deux ensembles *A* (intervalle de nombres 0...5) et *B* (intervalle 0...12) ne sont pas égaux parce que « la distance que deux grandeurs ont entre elles dans *B* est autre (plus grande) que la distance qui sépare dans *A* les deux grandeurs correspondantes ». Mais comment avoir une théorie autonome si une définition des structures empruntée à d'autres théories doit intervenir dès la relation d'égalité ?

A propos du continu, même flottement. Le but est abstrait : il s'agit de caractériser le continu comme ensemble infini ayant un « certain ordre » établi entre ses éléments : « il n'y a de continu que là où se trouve un système d'objets simples (de points dans le temps ou dans l'espace, ou encore de substances) placés de telle sorte que chacun d'eux possède, à toute distance si petite soit-elle, un voisin appartenant au système » [2]. La deuxième partie de la définition la restreint aux espaces métriques — et ne laisse aucune place à des considérations de voisinage : il ne s'agit pas de densité en soi, comme on a cru depuis [3]. On le voit par l'exemple du point 0, *isolé* dans le segment 0...1 qui n'est donc plus un continu si on garde 0 et si on en retranche les points d'abscisse $\frac{1}{2^n}$, parce que « pour toutes les distances de forme $\frac{1}{2_n}$ il n'y a pas de voisin pour 0 ». La définition n'est donc qu'un premier énoncé de l'axiome de CANTOR [4]. — La dimension est définie par récurrence : « un étendu (c'est-à-dire un continu) est *simplement étendu* ou ligne, si tout point, pour toute distance si petite soit-elle, y a un ou plusieurs

(1) P. 31.

(2) P. 73.

(3) Le voisin ne doit pas se trouver à l'*intérieur* d'un système de rayon quelconque mais *sur la surface* de cette sphère.

(4) A tout nombre réel correspond sur une droite un point ayant ce nombre pour abscisse.

voisins, mais jamais assez pour que leur système à lui seul constitue déjà un étendu », il a deux dimensions « si tout point pour toute distance... a une ligne entière de points comme voisins, etc... (1) ». Enfin on dit qu'un étendu est « *d'un seul tenant* s'il n'a aucune partie formant elle-même un étendu dans laquelle il n'y ait au moins un point qui, compté aux autres parties, ne forme avec elles également un étendu (2) ».

L'intérêt des dernières définitions est réduit par l'étroitesse de celle du continu. Le fragment de DEDEKIND (3) marque un progrès en introduisant la notion de voisinage (défini métriquement) : il s'agissait, à l'occasion des leçons de DEDEKIND sur le potentiel, et, « pour fonder plus rigoureusement le principe de DIRICHLET », de développer une *théorie des domaines* (GEBIETSLEHRE) « *ab ovo* sans faire appel à l'intuition géométrique (4) ». On définit ensembles ouverts, points intérieurs, extérieurs et frontières : « un système P de points p, p'... forme un *corps* [ensemble ouvert] si pour tout point p on peut donner une longueur δ telle que tous les points distants de p de moins de δ appartiennent à P. Les points p, p'... sont à l'intérieur de P » (5). De même « si P est un corps et m un point autour duquel on peut décrire une sphère dont tous les points sont non-intérieurs à P, le point m est *à l'extérieur* de P ». Enfin la *frontière* d'un corps est l'ensemble des points ni extérieurs ni intérieurs. — Quelques théorèmes : s'il y a un point extérieur à un corps, il y en a une infinité et leur système forme un corps ; le système des points frontières d'un corps ne peut consti-

(1) P. 80, n.

(2) P. 95.

(3) DEDEKIND (I), t. II, p. 353-355.

(4) L. à CANTOR, 19 janvier 1879 : « je ne me permettrais pas un tel jugement, si il y a de nombreuses années lorsque je voulais encore éditer la théorie de DIRICHLET sur le potentiel... je ne m'étais déjà assez occupé de ces questions ». CANTOR-DEDEKIND (I), p. 47.

(5) DEDEKIND (I), t, II, p. 353.

tuer un corps. « Pris par la mise au point de la théorie des nombres de DIRICHLET », DEDEKIND abandonne ces définitions qui, en 1879, lui semblent encore « constituer une bonne base » : quoique sans prolongement — au moins immédiat — elles sont une sorte de préface aux travaux de DEDEKIND sur les ensembles, échelonnés parallèlement à ceux de CANTOR à partir de 1872. L'échange entre les deux mathématiciens qui se poursuit tout au long de la création de CANTOR les situe donc logiquement avant celle-ci.

* * *

De 1873 à 1884 — jusqu'aux mémoires *Sur les ensembles infinis et linéaires de points*, premier exposé en forme de la théorie — la création cantorienne se caractérise par son intime subordination à l'analyse. C'est d'ensembles de points ou de nombres — non d'ensembles abstraits — qu'il s'agit, ce sont les applications analytiques qui importent, et si de nouveaux êtres sont introduits, avec une arithmétique qui leur est propre, leur justification est que seuls ils permettent « d'aller plus loin (1) » dans cette branche de la théorie des fonctions qu'est l'étude des ensembles de points. D'où l'enrobement des méthodes nouvelles dans une gaine intuitive sous laquelle il faut chercher leur véritable portée ; mais aussi dans la fraîcheur de leur apparition une sorte de naïveté qui laisse apparents les liens, permet de suivre la continuité d'un mouvement nécessaire. La création se fait en deux étapes : de 1873 à 1877, découverte et distinction de deux puissances ; de 1878 à 1884, exploration du continu, établissement de la théorie avec, comme complément attardé, en 1890, l'article où apparaît pour la première fois le procédé de la diagonale.

La première période n'est guère que préparation, mais

(1) CANTOR (XV), p. 165.

décisive. BOLZANO avait abouti à une impasse : pour explorer l'infini, l'application (1), seul procédé précis (généralisation de la numération), se révélait impropre puisque confondant tout et partie. Avec les deux résultats de CANTOR l'application est réhabilitée, un nouvel infini conquis au raisonnement mathématique : non seulement négativement on peut démontrer qu'il y a plus grand que l'ensemble des entiers, mais positivement, quoique l'on dépasse toute loi de série, on peut égaler deux espèces distinctes de ce plus grand: la mathématique sort cette fois de l'indéfini, une théorie de l'infini est fondée. Le point initial des deux recherches se trouve dans la définition des nombres irrationnels : d'une part, représentation du continu comme ensemble de toutes les limites superposées de suites de nombres rationnels ; d'autre part, emploi systématique de la réduction de suites à n^{uple} entrée en suites simples ou *linéarisation*, — procédé qui dans les deux cas ouvre le problème.

En 1873, CANTOR se demande si l'ensemble des nombres réels ne peut être mis sous forme de suite simple : « à première vue on se dit non, c'est impossible, car (n) [l'ensemble des entiers] est formé de parties discrètes, tandis que (x) [l'ensemble des nombres réels] constitue un continu ; seulement cette objection n'avance à rien... Ne serait-on pas aussi enclin, à première vue, à affirmer que (n) ne se laisse pas coordonner de façon biunivoque au système $(\frac{p}{q})$ de tous les nombres rationnels positifs $\frac{p}{q}$? (2). » Cependant, non seulement l'ensemble des nombres rationnels, mais, comme le montre DEDEKIND (3), l'ensemble des nombres algébriques se laisse mettre sous forme de suite simple et, de façon plus générale, tout ensemble ayant pour éléments $a_{n_1, n_2 \ldots n}$ dont les indices en nombre fini parcourent indépendamment les

(1) Nous dirons désormais *application* (ABBILDUNG) d'un ensemble sur un autre au lieu de mise en correspondance *biunivoque* de leurs éléments.

(2) L. à DEDEKIND, 29 novembre 1873. CANTOR-DEDEKIND (I), p. 13.

(3) CANTOR-DEDEKIND (I), p. 18.

uns des autres la suite des entiers. L'intérêt de la question, ajoute CANTOR comme pour se justifier, est qu'une réponse négative « donnerait une nouvelle démonstration au théorème de LIOUVILLE qu'il existe des nombres transcendants » (1). Il s'agit donc bien encore d'analyse. Le 7 décembre la réponse est trouvée grâce au procédé de l'emboîtement des intervalles, déduit directement de deux des propriétés du continu : entre deux éléments il y en a toujours d'autres (densité) ; toute suite convergente d'éléments définit un élément du système (compacité en soi). De la 1re propriété découle la présence entre deux termes de la suite :

(1) $a_1, a_2 \ldots a_n \ldots$

— où seraient rangés les nombres réels — d'autres termes de (1), et ainsi indéfiniment : constitution d'une suite d'intervalles emboîtés tels que tous les termes de (1) dont le rang est inférieur à celui des extrémités d'un intervalle soient extérieurs à cet intervalle. De la 2e propriété suit qu'il y a au moins un point m intérieur à tous les intervalles, le point limite commun des deux suites d'extrémités, si la longueur des intervalles tend vers o : m ne peut donc appartenir à (1).

La linéarisation n'intervient que négativement ; le théorème montre qu'elle est inefficace devant le continu : on ne peut pas coordonner (2) à (n) le système $(a_{n_1, n_2} \ldots)$ où les indices sont en nombre infini. Immédiatement — avec la même attente de réponse négative — se repose le problème pour le passage du continu simple au continu à n dimensions. « Dans le même ordre d'idées, écrit CANTOR le 5 janvier 1874, se présente aussi la question suivante : peut-on coordonner de façon... [biunivoque] ... une surface et une ligne ?... Il me semble encore pour l'instant — quoique

(1) L. à DEDEKIND, 2 décembre 1873, *ibid.*, p. 13.
(2) Nous représentons, comme fait CANTOR, une suite dont le terme général est n par (n).

l'on se sente tellement contraint ici au *non*, que l'on voudrait tenir une démonstration presque pour superflue — que la réponse présente de grandes difficultés (1). » Absurde question, déclarent les mathématiciens de Berlin, car « il va de soi que deux variables indépendantes ne peuvent se laisser ramener à une (2) ». Pourtant trois ans après, la linéarisation l'emporte (3) : un système de n fractions décimales indéfinies (les n coordonnées du point dans l'espace à n dimensions) peut se mettre sous forme d'une seule fraction décimale indéfinie (abscisse du point de la droite). Mais il faut que le procédé donne une correspondance biunivoque et exhaustive, difficulté signalée par DEDEKIND : si on exclut les fractions terminées par une suite infinie de 0 (pour obtenir l'unicité de la représentation décimale) les points de la droite dont l'abscisse comporte un développement avec périodicité de 0 (de période inférieure ou égale à n) n'auront pas de correspondant dans l'espace à n dimensions (4).

(1) CANTOR-DEDEKIND (I), p. 20.

(2) L. à DEDEKIND, 18 mai 1874, *ibid.*, p. 21.

(3) L. à DEDEKIND, 20 juin 1877, *ibid.*, p. 25 : si on a le système :

$$\left|\begin{array}{l} x_1 = \alpha_{11}\frac{1}{10} + \alpha_{1,2}\frac{1}{10^2} + \ldots + \alpha_{1,\nu}\frac{1}{10^\nu} + \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x_\rho = \alpha_{\rho,1}\frac{1}{10} + \alpha_{\rho,2}\frac{1}{10^2} + \ldots + \alpha_{\rho,\nu}\frac{1}{10^\nu} + \ldots \end{array}\right.$$

on en tire : $y = \beta_1\frac{1}{10} + \beta_2\frac{1}{10^2} + \ldots + \beta_\nu\frac{1}{10^\nu} + \ldots$ en posant :

$$\beta(n-1)\rho + \sigma = \alpha_{\sigma,n}.$$

(4) L. DEDEKIND à CANTOR, 22 juin 1877, *ibid.*, p. 27. Si on prend pour $n = 2$:

$$\left|\begin{array}{l} x_1 = \frac{\alpha_1}{10} + \frac{\alpha_2}{10^2} + \ldots \\ x_2 = \frac{\beta_1}{10} + \frac{\beta_2}{10^2} + \ldots \end{array}\right.$$

$y = \frac{\gamma_1}{10} + \frac{\gamma_2}{10^2} + \ldots$ avec $\gamma_1 = \alpha_1 \quad \gamma_2 = \beta_1 \quad \gamma_3 = \alpha_2 \quad \gamma_4 = \beta_2$, etc.
des fractions comme $y = 0{,}310507090_8\, 0\alpha_7\, 0\alpha_8 0\ldots$ ne seront pas représentées puisqu'on aurait $x_2 = \frac{3}{10} + 0 + 0 \ldots$ ce qui est exclu par convention.

Dans une nouvelle démonstration, le procédé de linéarisation est appliqué seulement aux systèmes de nombres irrationnels définis sans ambiguïté par leur développement en fraction continue. Mais qu'ajoute aux infinis en présence le système des nombres rationnels ? ici doit porter l'effort, uniquement d'ailleurs pour dégager les notions, car aucune méthode nouvelle n'est nécessaire. CANTOR avance avec timidité, en référence constante à l'intuition, comme pour se rassurer devant le résultat : « je le vois mais je ne le crois pas » (1). Trois étapes :

1° l'ensemble A, continu linéaire privé des nombres rationnels est appliqué sur l'ensemble B, continu linéaire privé d'une suite croissante (ε_ν) de nombres irrationnels : l'équivalence est démontrée entre deux variables qui parcourent A et B (et non pas directement entre A et B) et dont on voit qu'on peut faire correspondre les valeurs ;

2° l'ensemble B est appliqué sur le continu linéaire complet M par utilisation itérée, pour chaque intervalle $(\varepsilon_\nu \ldots \varepsilon_{\nu+1})$, du théorème que « un nombre y qui peut prendre toutes les valeurs de l'intervalle (0...1) à l'unique

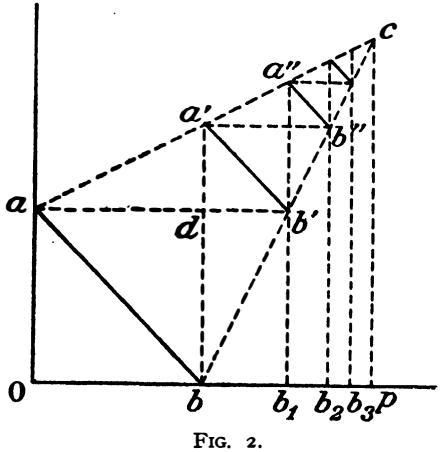


FIG. 2.

(1) L. à DEDEKIND, 29 juin 1877, *ibid.*, p. 34.

exception de la valeur o, peut être coordonné de façon univoque à un nombre x qui prend toutes les valeurs de (o...1) sans exception ». Ici la démonstration n'est même plus analytique mais graphique, grâce à la « courbe remarquable » (1) à laquelle CANTOR attache tant d'importance qu'il la reproduit de sa lettre à DEDEKIND à son article du *Journal de Crelle*, alors que figure dans celui-ci une autre démonstration complète qui tient en quelques lignes ;

3° enfin, puisque tout continu linéaire complet M est applicable sur l'ensemble linéaire irrationnel A, le continu à n dimensions M_n, défini par n continus linéaires, s'appliquera sur l'ensemble irrationnel à n dimensions A_n donc, par l'intermédiaire de l'ensemble A, sur le continu linéaire M.

Enfantement pénible de la notion de puissance qu'à la fois la transitivité des applications et l'étude des rapports entre ensemble et sous-ensemble obligent à dégager complètement. « Si deux ensembles bien définis M et N se laissent coordonner l'un à l'autre, élément par élément, de façon univoque et complète, je me sers de l'expression... qu'ils ont *égale puissance* ou qu'ils sont équivalents. Par une *partie intégrante* (BESTANDTHEIL) d'un ensemble M nous entendons tout autre ensemble M' dont les éléments sont en même temps éléments de M. Si deux ensembles M et N ne sont pas d'égale puissance, ou bien M a une puissance égale à celle d'une partie intégrante de N, ou N celle d'une partie intégrante de M ; dans le premier cas nous nommons la puissance de M *plus petite*, dans le deuxième cas nous la nommons *plus grande* que la puissance de N (2). » Formule (et

(1) La courbe y est constituée par les segments ab, $a'b'$, $a''b''$... et le point c, en sont exclus les points b, b', $b^{(\nu)}$ ... : on voit que x prend toutes les valeurs de o à 1, au contraire y toutes les valeurs de o à 1 sauf o.

(2) CANTOR (VIII), p. 119. Cf. plus tard : « j'ai emprunté le mot *puissance* à J. STEINER qui l'emploie dans un sens tout spécial, cependant voisin pour exprimer que deux figures sont coordonnées l'une à l'autre *projectivement* de telle sorte qu'à tout élément de l'une correspond un et un seul élément de l'autre ; dans le concept de puissance ici employé on conserve la corres-

solution supposée) déjà du problème de la comparabilité. Ainsi un calcul des puissances se trouve ébauché. La puissance est un nombre : elle se confond avec le nombre entier positif pour les ensembles finis « puisque en effet deux tels ensembles ont et n'ont puissance égale que si le nombre de leurs éléments est le même ». Pour les deux puissances de l'ensemble des entiers et du continu, non seulement la relation d'inégalité est démontrée, mais le résultat d'additions et de multiplications :

1° « Si M', M'', M''' est une suite finie ou infinie d'ensembles ayant chacun la puissance de la suite des entiers positifs, l'ensemble M formé de la réunion de ces ensembles a aussi la même puissance (1). »

2° L'addition de la puissance des entiers à la puissance du continu (ou la soustraction de la 1re de la 2e) donne la puissance du continu (parties 1 et 2 du théorème).

3° La multiplication de la puissance du continu par elle-même donne la puissance du continu. C'était au moins un des aspects sous lesquels se présentait le problème : « comme si l'infini des points d'une surface était obtenu à partir des points d'une ligne par élévation au carré, celui des points d'un volume par élévation au cube (2) ». Toutefois ce calcul — en particulier le dernier résultat — n'est pas encore expressément dégagé : pour que les puissances deviennent effectivement des nombres, il faut que leur mode d'engendrement, en particulier d'abord le passage du dénombrable au continu, soit découvert. C'est donc l'analyse du continu qui passe au premier plan, avec d'ailleurs son autorité concrète : après avoir en partie détruit la notion intuitive (pour la dimension) il s'agit de reconstruire. Mais dans la joie d'avoir dominé « cette force étonnante qui réside dans

pondance biunivoque, mais en revanche n'est imposée aucune restriction à la loi de coordination ». CANTOR (XIII), p. 151.

(1) CANTOR (VIII), p. 120.

(2) L. à DEDEKIND, 20 juin 1877. CANTOR-DEDEKIND (I), p. 25.

les nombres réels rationnels et irrationnels, capable de déterminer les éléments d'un système continu à ρ dimensions [ou à une infinité de dimensions, comme le montre une extension immédiatement faite de la linéarisation] de façon univoque à l'aide d'une seule coordonnée » (1), disparaît toute timidité devant l'infini : « tous les ensembles linéaires imaginables de points ou de nombres réels » sont répartis en classes déterminées d'après leur puissance et déjà se formule le problème central : « par un procédé d'induction — qui ne sera pas précisé ici — est approché le théorème que le nombre des classes donné par cette répartition est fini et égal à deux (2) ».

Ainsi s'oriente définitivement la théorie : parallèlement aux recherches de topologie proprement dite— sur la dimension et la connexion — s'ouvre une série de travaux (théorie des ensembles de points mêlée avec théorie abstraite des ensembles) qui n'ont dès maintenant d'autre but que de prouver la validité de l'hypothèse et d'établir une relation entre les deux puissances. Le moyen d'action est emprunté au dernier mémoire *Sur les séries trigonométriques* dont l'influence souterraine semble n'avoir pas cessé (3) : c'est l'opération de la dérivation, opération centrale, — que U. DINI a utilisée « dans son dernier ouvrage pour une série de remarquables généralisations de théorèmes connus d'analyse... Sur cette opération, comme nous le montrerons plus tard, est fondée la plus simple et en même temps la plus complète explication et détermination du continu (4) ». C'est d'elle en effet que proviennent également les deux espèces d'objets essentiels à l'édifice cantorien et grâce aux-

(1) L. 25 juin 1877, *ibid.*, p. 34.
(2) CANTOR (VIII), p. 132.
(3) « Pour ce qui concerne les nombres ordinaux transfinis il me paraît vraisemblable que j'en avais déjà une représentation en 1871 ». L. à JOURDAIN, 31 août 1905 *in* JOURDAIN (I), p. 28.
(4) CANTOR (X), p. 139.

quels le problème sera attaqué sinon résolu, les nombres transfinis et les ensembles fermés et parfaits.

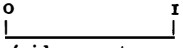
En 1871, au moins dans le mémoire publié, on ne distinguait que deux *genres* (GATTUNGEN) d'ensembles : ceux qu'un nombre fini de dérivations réduisaient à o ; ceux « pour lesquels la suite indéfinie des dérivés ne s'arrête pas». En 1877, dans la *classe* des ensembles de 1re puissance, figure le 1er *genre* d'ensembles. Mais comment se comportent, du point de vue de la puissance, les ensembles de 2e genre ? A une première confrontation de ces deux modes de répartition sont consacrés les mémoires échelonnés entre 1879 et 1882.

D'abord deux nouvelles notions qui serviront d'instruments : la densité et le recouvrement par des intervalles. La densité avait été définie en 1872 par DEDEKIND en tant que propriété commune aux systèmes des nombres rationnels et réels. CANTOR marque seulement sa parenté avec l'opération de la dérivation : être dense dans un intervalle A c'est, pour un ensemble, posséder des points dans tout intervalle déterminé au sein de A : donc le premier dérivé d'un ensemble dense coïncide avec l'intervalle tout entier. Un ensemble dense est nécessairement de 2e genre (la réciproque n'étant pas vraie). Un ensemble partout non dense est tel que dans tout intervalle on puisse trouver un intervalle où il ne possède aucun point. La notion de recouvrement par intervalles est empruntée à DU BOIS REYMOND: une fonction discontinue peut être intégrée si l'ensemble de ses points de discontinuité forme un *groupe intégrable*, c'est-à-dire peut être recouvert par une suite finie d'intervalles dont la longueur totale est nulle. Un tel ensemble est évidemment partout non dense. Le résultat important obtenu par CANTOR est que tout ensemble P dont le 1er dérivé P' est dénombrable forme un groupe intégrable : en effet, ni cet ensemble, ni son dérivé, ni la somme $P + P'$ ne peuvent être denses dans aucun intervalle; chaque point ex-

térieur à $P + P'$ peut donc être entouré d'un intervalle dont aucun point sauf les extrémités n'appartient à $P + P'$; on démontre que la somme de ces intervalles *contigus* à $P + P'$ est égale à la longueur totale de l'intervalle global (dans lequel P et P' sont définis), en utilisant une propriété élémentaire des fonctions continues (1) ; il sera donc possible d'enfermer $P + P'$ dans un nombre fini d'intervalles de longueur arbitrairement petite. Le raisonnement par considération des intervalles contigus est nouveau (dû en partie à DU BOIS REYMOND) et resservira pour les ensembles parfaits.

En second lieu précision et extension des notions déjà acquises : la dérivation est généralisée au cas de l'espace à n dimensions ; on démontre, en transformant par rayons vecteurs réciproques, l'espace à n dimensions en une surface (d'aire finie) à $n + 1$ dimensions, et en utilisant ensuite le procédé de l'*échelle dénombrable* (il n'y a qu'un nombre fini d'intervalles d'aire supérieure à un nombre rationnel donné puisque la somme totale est finie), qu'il n'y a qu'une infinité dénombrable d'intervalles (à n dimensions) non empiétants, — théorème nécessaire pour la suite. La puissance est définie pour des ensembles non plus de points, mais d'objets quelconques « lignes, surfaces, corps ». Pour tout couple d'ensembles bien définis « on doit considérer comme intrinsèquement déterminé s'ils ont ou non même puissance ». Un ensemble d'éléments est « *bien défini*, si sur la

(1) La fin du raisonnement peut être ainsi donnée : soit $f(x) = x - \sigma$ σ étant la somme des intervalles contigus entre o et x ($P + P'$ est supposé situé sur l'intervalle o...1) :



$f(x)$ est continue car on a évidemment :

$$|f(x + L) - f(x)| \leqq L.$$

Soit l la longueur totale, des intervalles contigus : $f(x)$ varie de o à $1 - l$; or, à l'intérieur d'un intervalle contigu $f(x)$, est constante, comme $P + P'$ est dénombrable elle ne prend donc qu'une infinité de valeurs, comme elle est continue elle prend toutes les valeurs entre o et $1 - l$, donc $l = 1$.

base de sa définition, et en conséquence du principe logique du tiers exclu, on doit considérer comme intrinsèquement déterminé, aussi bien si un objet quelconque appartenant à la même sphère conceptuelle fait partie ou non de l'ensemble, qu'également si deux objets appartenant à l'ensemble sont identiques malgré les différentes façons dont ils sont donnés (1) ». De ce processus d'abstraction et de projection en quelque sorte dans un domaine objectif, le problème du continu reçoit une acuité plus grande : il doit être intrinsèquement déterminé si les ensembles de points ne peuvent avoir que les deux puissances connues ; d'autre part une définition abstraite de ces puissances et de leur rapport est possible si les deux types d'ensembles auxquels elles s'appliquent peuvent être suffisamment caractérisés conceptuellement. La dérivation va-t-elle permettre cette caractérisation ? L'analyse mécanique, nécessaire en quelque sorte, qu'elle produit dans un ensemble le fait espérer : on sait déjà qu'un ensemble dont le $n^{ième}$ dérivé est nul a la première puissance ; mais on peut aller plus loin : il suffit de remarquer qu'à partir du 2e dérivé tous les dérivés sont partie intégrante de leur prédécesseur. Si donc tous les dérivés finis ont des points, on considérera comme dérivé de rang $\infty, p\infty$, l'ensemble des points communs à toute la suite. En dérivant à nouveau on introduit les symboles $\infty + 1, \infty + 2, \ldots, \infty + n$, puis une nouvelle partie commune donne $p^2\infty, \ldots p^{\infty^2}, \ldots p^{\infty^\infty}, p^{\infty^{\infty^\infty}} \ldots$
« Nous voyons ici une génération dialectique de concepts qui conduit toujours plus loin, et libre de tout arbitraire reste nécessaire en soi et conséquente (2). »

Grâce à elle, en effet, sont procurés immédiatement au sujet de la puissance trois résultats d'extension croissante :

1° Si l'on appelle isolé un ensemble qui n'a pas de point commun avec son dérivé (par exemple la différence $P' - P''$

(1) CANTOR (XIII), p. 150.
(2) CANTOR (XI), p. 148.

entre le 1er et le 2e dérivé), on remarque que les dérivations successives à partir de la 2e ont pour effet de retrancher un ensemble isolé. Or tout ensemble isolé est dénombrable : chacun de ses points peut être, en effet, enfermé dans un intervalle (à n dimensions si l'ensemble est situé dans un espace à n dimensions) non empiétant sur les autres et comme on a vu qu'il n'y a qu'une infinité dénombrable de tels intervalles l'ensemble est dénombrable.

2° Si le 1er dérivé P' d'un ensemble P est dénombrable, P est dénombrable : en effet, un ensemble se compose d'un ensemble isolé et de la partie commune avec son dérivé.

3° Tout ensemble dont le dérivé d'ordre α (α étant soit un nombre fini, soit un quelconque des précédents symboles transfinis) est nul et dénombrable. La démonstration est immédiate lorsque α est fini. Lorsque α est un symbole transfini, CANTOR déclare « qu'elle doit se faire exactement de même, à l'aide de l'induction complète (1) », sans même remarquer l'extension donnée du même coup au procédé d'induction. C'est que le mouvement dialectique engendrant la suite des ∞ lui apparaît avec l'évidence contraignante de la logique. Mais ici se manifeste aussi sa limitation : l'induction complète ne peut prouver la dénombrabilité qu'à condition de ne comporter elle-même qu'une infinité dénombrable d'étapes. Si l'adhérence reste entre induction et génération des ∞, impossible de dépasser le dénombrable, la liaison avec le continu est manquée. Si l'on veut atteindre celle-ci il faut que les symboles ∞ soient autre chose que les étapes d'une génération progressive, mais les représentants d'une réalité immanente que des moyens nouveaux permettent d'atteindre, il faut s'engager dans le domaine véritable du transfini.

Tel est le dilemme devant lequel se trouvait CANTOR à la

(1) CANTOR (XIV), p. 160.

fin de l'été 1882 : jusque-là rien que de normal dans l'édifice construit ; si les résultats sont étonnants, les méthodes employées sont celles de l'analyse ordinaire, aucun mathématicien, si empiriste soit-il, n'y peut rien trouver à redire. Dans le mémoire de mars 1882 (1) le ton a sans doute déjà changé : mais l'affirmation d'un en soi des vérités mathématiques n'a pas de conséquence pratique. La décision effective pour l'égalité de puissance de deux ensembles est comparée à la décision pour savoir si un nombre donné est ou non transcendant, « un des plus difficiles problèmes » de l'analyse (2). C'est plutôt l'idée que la théorie des ensembles domine et logiquement précède le reste des mathématiques qui pourrait annoncer un changement de point de vue : « la théorie des ensembles, dans la conception ici exposée, embrasse — si nous ne considérons que le mathématique et laissons provisoirement de côté les autres sphères conceptuelles — les domaines de l'arithmétique, de la théorie des fonctions et de la géométrie ; elle les réunit sur la base du concept de puissance en une unité supérieure. *Discontinu* et *continu* se trouvent de la sorte considérés du même point de vue et mesurés avec une mesure commune (3) ». Quels peuvent être cette mesure, ce point de vue communs, sinon ceux que procure une refonte de la notion de nombre ? La phrase étrange sur « les autres sphères conceptuelles » semble se référer aux enquêtes philosophiques sur l'infini, qui tiendront la première place dans le mémoire d'octobre. Toutefois mathématiquement rien encore n'apparaît : dans la correspondance avec DEDEKIND jusqu'aux premiers jours d'octobre il n'est question que de la définition topologique du continu. C'est, semble-t-il, dans l'exaspération de l'échec, après l'inefficacité reconnue de la dérivation, que CANTOR se

(1) CANTOR (XIII) daté 31 mars 1882.
(2) *Ibid.*, p. 150.
(3) P. 152.

décide à franchir le pas : dans une soudaine illumination, il construit, en quelques jours sans doute, le mémoire d'octobre, il se lance à corps perdu dans l'infini actuel : « Justement depuis nos dernières rencontres à Harzburg et Eisenach, écrit-il le 5 novembre à DEDEKIND (1), Dieu tout-puissant a voulu que je parvienne aux lumières les plus étonnantes et les plus inattendues sur la théorie des ensembles et sur la théorie des nombres, ou, bien plus, que je trouve ce qui depuis des années a fermenté en moi, ce vers quoi j'ai fait de longues recherches (2). » L'exaltation du mémoire lui-même, ce désordre de considérations mathématiques pêle-mêle avec les références philosophiques les plus confuses — la règle d'addition des ordinaux $\omega + 1 > \omega$ est présentée comme solution au problème spinoziste du rapport entre modes infinis et modes finis, — aussi bien que les excuses réitérées dans l'article et dans la lettre, montrent assez que la rupture avec les mathématiques habituelles est profondément sentie (3). « Je ne vois vraiment pas ce qui pourrait nous retenir dans cette activité créatrice de nouveaux nombres, aussitôt qu'il apparaît que pour le progrès de la science l'introduction d'une nouvelle, parmi ces innombrables classes de nombres, est devenue souhaitable ou même indispensable ; et la dernière éventualité me semble se produire effectivement en théorie des ensembles, peut-être même aussi encore dans un domaine beaucoup plus étendu, du moins *sans* cette extension je ne peux plus aller de l'avant, *avec* elle j'atteins toute sorte d'inattendu. »

(1) D'après une note manuscrite de DEDEKIND (brouillons de *Was sind und was sollen die Zahlen ?*) l'entrevue de HARZBURG eut lieu du 7 au 12 septembre (DEDEKIND communiqua la 1re esquisse de son travail). La création de CANTOR fut donc très rapide, puisque le mémoire est daté d'octobre.

(2) CANTOR-DEDEKIND (I), p. 55.

(3) « Je ne me dissimule aucunement qu'avec cette entreprise je me mets dans une opposition certaine avec des vues répandues sur l'infini mathématique et avec des opinions souvent défendues sur l'essence de la grandeur numérique. » CANTOR (XV), p. 165.

De même le début du mémoire : sans cette extension du concept de nombre « il me serait impossible de faire le plus petit pas en avant dans la théorie des ensembles ; puisse cette circonstance me justifier, ou si c'est nécessaire, m'excuser d'introduire des idées en apparence étranges dans mes développements (1) ».

Le surprenant n'est pas tant, en effet, dans l'usage de nouveaux symboles donnant lieu à un calcul spécial — comme toutes les extensions du nombre, et plus récemment les travaux de GRASSMANN et de KUMMER en avaient donné l'exemple (2) — ni même dans la considération de l'infini actuel, invoqué déjà en particulier pour compléter le plan d'une variable complexe, que dans la décomposition logique de la notion de nombre et la superposition de classes d'objets, hors des prises de toute atteinte effective. Définition du nombre comme caractère émanant d'un système posé en soi d'objets et se dissociant en deux concepts : « celui de la puissance... indépendante de l'ordre imposé à l'ensemble et celui du nombre ordinal (ANZAHL) qui est nécessairement lié à l'ordre imposé à l'ensemble par une règle, au moyen de laquelle l'ensemble devient... *bien ordonné*... Si je redescends de l'infini au fini je vois avec la même clarté et la même beauté comment les deux concepts redeviennent « un » et se *fondent* dans le concept du nombre entier fini (3). » La notion de *bon ordre* pour un ensemble, — être ordonné de telle façon qu'à tout sous-ensemble corresponde dans l'ensemble global un élément

(1) L'excuse est reprise vingt pages plus loin : « sans les concepts susdits [nombre ordinal et cardinal] on ne va — j'en ai la conviction — pas plus loin dans la théorie des ensembles ; et il en est de même, semble-t-il, des domaines qui dépendent de la théorie des ensembles ou ont avec elle le contact le plus intime comme, par exemple, la théorie moderne des fonctions d'une part, la logique et la théorie de la connaissance d'autre part ». *Ibid.*, p. 180-181.

(2) *Ibid.* p. 183.

(3) P. 181.

qui suit immédiatement — provient en effet de l'extension de l'opération de la numération telle que la provoque la dérivation : au premier principe d'engendrement des numéros d'ordre finis — adjonction d'une unité — doit être joint pour le numérotage des dérivés un second principe qui consiste, « quand on se trouve en présence d'une succession déterminée quelconque de nombres entiers réels définis parmi lesquels aucun n'est le plus grand, a créer un nouveau nombre qui sera considéré comme la *limite* de ces nombres, c'est-à-dire comme le nombre immédiatement plus grand qu'eux tous (1) ». L'analogie avec la définition des nombres irrationnels (2) — reprise d'ailleurs dans le même mémoire — est évidente : seulement les nombres irrationnels sont justifiés en tant que nombres par le double caractère d'être insérables dans un ordre, où figurent d'ailleurs les représentants des nombres rationnels, et de permettre la définition entre eux des 4 opérations arithmétiques — obéissant aux lois ordinaires (et donnant les résultats ordinaires pour les représentants des nombres rationnels). Pour les nombres transfinis l'ordre est défini par leur processus d'engendrement : *le 2e principe n'est autre que la condition du bon ordre.* Donc ensemble bien ordonné dénoté par un ordinal transfini et ensemble des ordinaux inférieurs à cet ordinal sont deux réalités coïncidant du point de vue de la théorie. Quant aux règles d'arithmétique, elles se déduisent immédiatement de la signification intuitive des opérations : ajouter un ensemble à un autre est former un nouvel ensemble où tous les éléments du second suivent tous les élé-

(1) CANTOR (XV), p. 196.

(2) L'analogie est d'ailleurs explicitement signalée dans la lettre à LASSWITZ du 15 février 1884 (CANTOR (XXIII), p. 395) : « Les nombres transfinis sont dans un certain sens même de *nouvelles irrationalités* et en fait la meilleure méthode à mes yeux pour définir les nombres irrationnels finis est... je pourrais même dire dans le principe la même que ma méthode pour l'introduction des nombres transfinis. On peut dire absolument : les nombres transfinis *restent ou tombent* avec les irrationnels finis ».

ments du premier, l'ordre intérieur à chaque ensemble étant conservé. Multiplier un ensemble A par un ensemble B c'est remplacer tout élément de B par un ensemble équivalent à A et ordonné de même façon, l'ordre de B étant conservé. Sur les procédés de calcul, la décomposition en produit de nombres premiers transfinis, le mémoire ne donne encore que de brèves indications.

Mais le but à atteindre n'est pas une théorie de ces nombres : il s'agit des puissances dont le caractère numérique avait été posé dès l'abord. Or à la première difficulté, passage d'une des deux puissances connues à l'autre, s'en ajoute maintenant une seconde, comparabilité des puissances — qui avait été postulée sans plus en 1877. De la relation (1) d'inégalité définie alors —

$$\text{puissance } M < \text{puissance } N \text{ si } M \sim N',$$

N' étant une partie de N et M non équivalent à N— résulte que si M est applicable sur N' partie de N et N sur M' partie de M on devrait pouvoir démontrer que M et N ont même puissance ; ou encore en appelant M'' la partie de M' sur laquelle l'application de N sur M' applique N' :

$$M \sim N' \sim M'' \qquad (1) ;$$

$$\text{si puissance } M'' \leqq \text{puissance } M' \leqq \text{puissance } M \qquad (2)$$

on a :

$$\text{puissance } M' = \text{puissance } M = \text{puissance } M'' \qquad (3).$$

C'est donc la suite (2) d'inégalités qu'il faudrait démontrer. « Vous vous rappelez, écrit CANTOR à DEDEKIND le 5 novembre (2), que je vous disais à Harzburg qu'il ne m'était pas possible de démontrer le théorème suivant : si M' est partie d'un ensemble M, M'' partie de M' et si... M et M'' ont même puissance, M' a aussi même puissance

(1) Nous nous servons du symbole $\sim$ pour exprimer l'équivalence de deux ensembles.

(2) CANTOR-DEDEKIND (I), p. 55.

de M et M''. Maintenant j'ai trouvé la source de ce théorème. » Ce sont les nombres transfinis qui réduisent l'une et l'autre difficulté.

En premier lieu ils permettent d'engendrer régulièrement des puissances supérieures. On a vu que les symboles transfinis du précédent mémoire ne permettaient pas de dépasser le dénombrable ; les ordinaux nés des deux principes (que l'on appellera ordinaux de la classe II, la classe I étant réservée aux ordinaux finis), n'étant que leur nouveau déguisement ne dénoteront donc jamais que des ensembles dénombrables. Mais leur ensemble n'est pas dénombrable : en effet, à toute suite dénombrable d'ordinaux de la classe II correspond soit un ordinal maximum appartenant à la suite, soit un ordinal de la classe II immédiatement supérieur, c'est-à-dire tel que tout ordinal inférieur est dépassé par un terme de la suite ; si donc l'ensemble de tous les ordinaux pouvait se mettre sous forme de suite dénombrable, ou bien il y aurait un terme γ plus grand que tous les autres, alors $\gamma + 1$ n'appartiendrait pas à la suite, ou bien il y aurait un nombre β de la classe II immédiatement supérieur à la suite. On reconnaît dans ce raisonnement un nouveau procédé de dépassement des suites dénombrables : le premier, pour le continu, utilisait la propriété que toute suite convergente d'éléments définit un nouvel élément ; celui-ci que toute suite d'ordinaux de la classe II peut être bien ordonnée (donc définit un nouvel ordinal de la même classe), méthode qui donnera le paradoxe de BURALI-FORTI (1). Donc est brisé le cercle magique qui enfermait les symboles transfinis. La classe II sera suivie d'autres classes à condition d'ajouter aux deux principes d'engendrement un *principe d'arrêt*, qui clôture les classes, donc donne une nouvelle application aux principes moteurs :

(1) V. ci-après, p. 94.

n'appartiennent à une classe que les ordinaux dénotant des ensembles ayant la puissance de la classe précédente ; ainsi, pour la classe II, des ensembles dénombrables. Donc le 2e principe moteur définira un ordinal immédiatement supérieur à ceux de la classe II, Ω premier élément de la classe III. « Sur cette voie... on peut, avec la plus grande sécurité et évidence, parvenir à des classes de nombres toujours nouvelles, et avec elles à toutes les puissances successives différentes qui se présentent dans la nature corporelle et spirituelle (1). » En effet, l'échelle ainsi obtenue permet de passer d'une puissance à la puissance immédiatement supérieure ; le théorème est démontré pour la classe II : tout ensemble d'ordinaux de la classe II est ou fini, ou dénombrable ou a la puissance de la classe II ; *quartum non datur*. Donc le procédé est pleinement satisfaisant si on est sûr d'atteindre ainsi toutes les puissances.

C'est ici que reparaît la deuxième question. D'abord un postulat ou une promesse : « qu'il soit toujours possible de mettre tout ensemble *bien défini* sous forme d'un ensemble *bien ordonné*, sur cette loi, me semble-t-il, fondamentale et féconde de notre pensée, particulièrement remarquable par son universalité, je me propose de revenir dans un mémoire ultérieur (2) » — promesse non tenue. Puis une démonstration, restreinte au cas de la classe II. Le théorème de l'équivalence est immédiat si M a la puissance de la classe II : en effet, M' ne peut être que fini dénombrable ou de puissance de la classe II ; comme il possède une partie M'' ayant cette puissance, il lui est équivalent. Ici encore promesse de solution ultérieure pour le cas général : dans la lettre à DEDEKIND on a vu que CANTOR croyait déjà la tenir.

Mais, même si on admet « la loi fondamentale du bon ordre », le problème initial n'est pas résolu : dans l'échelle

(1) CANTOR (XV), p. 199.
(2) *Ibid.*, p. 169.

des puissances, quelle place a la puissance du continu ? Ici se produit, symétriquement à la création des ordinaux, un événement qui restreint au moins la recherche, l'apparition des ensembles parfaits dans le processus de dérivation. On a vu que tout ensemble, dont le 1er dérivé est dénombrable, est lui-même dénombrable. Le mémoire d'octobre 1882 introduit un nouveau résultat, d'ailleurs inexact partiellement : si le dérivé $P^{(1)}$ n'est pas dénombrable il est décomposable d'une seule manière en un ensemble réductible R et un ensemble parfait S ; un ensemble parfait étant un ensemble identique à son premier dérivé. L'opération de la dérivation renouvelée un nombre suffisant de fois (nombre qui est un des ordinaux de la classe II) réduit tout ensemble dont le 1er dérivé n'est pas dénombrable à un ensemble parfait. Le théorème était inexact, ainsi que le montra BENDIXSON (1) au sujet de R qui peut être irréductible, mais il est toujours dénombrable et sans point commun avec ses dérivés à partir d'un certain rang, de sorte que l'essentiel du résultat subsiste. C'est le mémoire de novembre 1883 qui donne la démonstration et étudie la puissance des ensembles parfaits (2). D'abord un ensemble parfait ne peut être dénombrable : preuve le procédé des intervalles emboîtés ; le raisonnement est à peu près le même que pour le continu. On remplace la densité par la densité en soi (tout point de l'ensemble est point limite, propriété que possède évidemment le continu) ; les intervalles ne seront plus pris entre deux points, mais autour d'un point. En second lieu, tout ensemble dont le 1er dérivé est dénombrable est lui-même réductible. Il suffit de se rappeler la décomposition opérée par la dérivation et que l'introduction des nombres ordinaux, avec Ω, rend plus simple :

(1) « Dans une lettre qu'il m'adressa (mai 1882). Sur ma demande il a publié les résultats trouvés à cette occasion dans les *Acta mathematica*, t. II, p. 415 ». CANTOR (XVIII), p. 224. Cf. BENDIXSON (I).

(2) CANTOR (XVIII).

$$P^{(1)} = \sum_{\gamma} [P^{(\gamma)} - P^{(\gamma+1)}] + P^{(\Omega)} \qquad (1)$$

(où γ parcourt tous les nombres des classes I et II). Si aucune des différences sous le Σ n'était nulle, le Σ aurait la puissance de la classe II ; donc elles s'annulent toutes à partir d'un certain γ ; d'autre part, on ne peut avoir :

$$P^{(\gamma)} = P^{(\gamma+1)},$$

l'ensemble $P^{(\gamma)}$ étant parfait et contenu dans $P^{(\gamma)}$ celui-ci ne pourrait être dénombrable. Donc à partir d'un certain γ tous les dérivés $P^{(\gamma)}$ sont = 0. Enfin si le 1er dérivé d'un ensemble a une puissance supérieure au dénombrable, il existe des points communs à tous les dérivés et leur ensemble est parfait. Le procédé utilisé pour la démonstration est une spécialisation du principe de la propriété distributive utilisée par BOLZANO dite ici propriété r (propriété telle que, si un ensemble la possède dans toute division finie de cet ensemble, une des parties au moins la possède et toute somme de cet ensemble avec un autre la possède également) : on démontre que dans tout ensemble M possédant une propriété r il existe au moins un point g dont tous les voisinages si petits soient-ils renferment des parties de M possédant r. Avoir tous ses dérivés différents de 0 est pour l'ensemble P une propriété r : le point g correspondant est donc point limite de tous les dérivés (puisqu'ils ont tous une portion contenue dans tous les voisinages de g) ; comme à partir de $P^{(1)}$ ils contiennent tous leurs points limites, g leur appartient à tous. On voit d'autre part sans peine que l'ensemble des points tels que g est parfait. Il est donc possible de décomposer $P^{(1)}$ en un ensemble dénombrable R et un ensemble parfait S :

$$P^{(1)} = R + S \qquad (2)$$

La formule (1) dans laquelle $P^{(\Omega)}$ est évidemment un ensemble parfait puisque commun à tous les dérivés, oblige

à identifier R et $\Sigma\,[P^{(\gamma)} - P^{(\gamma+1)}]$, donc γ ne parcourt qu'une infinité dénombrable de nombres de la classe II, il y a un plus petit ordinal α pour lequel :

$$P^{(\alpha)} = P^{(\alpha+1)} = S$$

Ainsi l'aboutissement à un ensemble parfait remplace le parcours effectif de la totalité de la classe II : rapport qui pouvait sembler de favorable augure pour l'hypothèse sur le continu. En effet un dernier résultat donnait au problème son acuité maximum : tout ensemble parfait a la puissance du continu. La démonstration s'effectue sur le cas particulier de l'ensemble parfait non dense (découvert par DU BOIS REYMOND) : les intervalles contigus à cet ensemble P forment dans l'intervalle global un ensemble dénombrable partout dense et chaque point de P peut être défini de façon univoque par une suite convergente de ces intervalles (il est limite de suites d'extrémités). Si donc on met en correspondance biunivoque les intervalles contigus à P et l'ensemble des nombres rationnels de l'intervalle (0...1), à tout nombre réel dans (0...1), défini par une suite de nombres rationnels, correspondra de façon univoque un point de P défini par la suite homologue d'intervalles — et réciproquement. D'autre part tout ensemble fermé — c'est-à-dire contenant ses points limites — peut être considéré comme le 1er dérivé d'un ensemble : il est donc ou dénombrable ou de puissance du continu. Pour toute une catégorie d'ensembles est donc démontrée l'hypothèse de la dualité des puissances. Puisque la puissance de la classe II vient immédiatement après le dénombrable, si elle était identique à celle du continu, il n'y en aurait bien que deux. D'autre part si l'on pouvait construire effectivement un ensemble fermé ayant la puissance de la classe II, en vertu de l'égalité (2) le problème serait résolu. « Comme vous le voyez, écrit CANTOR à MITTAG-LEFFLER en 1884, tout revient à définir seulement un seul ensemble fermé

ayant la 2e puissance (1). » C'est à quoi il s'acharne pendant toutes ces années, attaquant le problème surtout du biais de la théorie des ensembles de points, privilégiant parmi ceux-ci toujours les ensembles fermés, traversant d'ailleurs parfois des crises de doute qui lui font contester la possibilité d'égaler continu et puissance d'une classe quelconque de nombres transfinis (2).

Dans cette période de regroupement et de systématisation, plutôt que de création, apparaît cependant un résultat nouveau sur les puissances (3) : la démonstration par le procédé de la diagonale que l'ensemble des fonctions d'une variable a une puissance supérieure à l'ensemble de variation de son argument. On a vu que Du Bois Reymond employait déjà de façon intuitive le procédé de la diagonale en 1876. Cantor part du cas où l'ensemble de variation de l'argument est l'ensemble des entiers : le système des fonctions prenant deux valeurs sur cet ensemble est alors celui desnombres réels, nouvelle démonstration du théorème de 1874. Mais le procédé peut se généraliser à n'importe quel ensemble base ; c'est le troisième procédé de dépassement, plus efficace que les deux autres puisque non restreint à une suite dénombrable, comme le premier, ou à une suite bien ordonnée comme le deuxième, qui utilise l'indépendance relative des deux variables : l'ensemble des fonctions étant supposé coordonné à celui de l'argument, il est toujours possible de définir une nouvelle fonction par la propriété d'avoir, pour un argument quelconque, une valeur différente de celle prise par la fonction coordonnée à cet argument. Cantor n'envisage comme

(1) L. du 26 août 1884, Schœnflies (II), p. 17.

(2) L. du 14 novembre à Mittag Leffler : « Heureusement je n'ai utilisé nulle part cette affirmation [que le continu a la puissance de la classe II] comme moyen de démonstration... J'ai trouvé une démonstration rigoureuse que le continu n'a pas la puissance de la classe II et bien plus qu'elle n'a aucune puissance exprimable en général par un nombre. » Cf. Cantor-Dedekind (I), p. 5.

(3) Cantor (XXV).

autre application que le dépassement du continu : c'est l'ensemble de toutes les fonctions d'une variable réelle. Qu'il y ait dans le procédé exactement la réponse à la question posée en 1882 par les résultats mêmes : y a-t-il un moyen régulier pour engendrer des puissances successives, en particulier pour le passage entre les deux puissances connues (dont la mise en présence exigeait qu'il y ait détermination de leur rapport ?) donc inutilité de faire le détour des classes d'ordinaux transfinis, il ne semble pas le remarquer. Bien plus, il s'appuie, pour confirmer « que les puissances n'ont pas de maximum », sur la démonstration, obtenue « par de tout autres moyens » que « le système de toutes les puissances si nous les envisageons ordonnées suivant leur grandeur forme un ensemble bien ordonné de sorte qu'à toute puissance correspond dans la nature une puissance immédiatement supérieure, mais aussi à toute suite croissante illimitée de puissances une qui la dépasse immédiatement (1) ». Tant étaient forts à la fois l'attrait de l'opération du bon ordre et la confiance dans l'hypothèse du continu. Aucun des trois principes présidant à l'engendrement des classes d'ordinaux transfinis ne permet d'ailleurs de prouver la dernière affirmation, c'est-à-dire d'atteindre seulement la classe d'ordinaux qui a pour numéro d'ordre ω. C'est l'obsession imaginative du processus engendré par l'opération de la dérivation: le lien entre les développements de la théorie qui en naissent et le problème effectivement posé par la découverte des deux puissances apparaît donc cette fois contingent ; il s'agit de deux manières radicalement différentes d'envisager les ensembles, aucun problème effectivement posé n'assure qu'il y ait communication de l'une à l'autre.

(1) *Ibid.*, p. 280.

* * *

De là une dualité latente, un malaise sous-jacent dans la systématisation de la science de l'infini à laquelle est consacré le reste des travaux cantoriens. Il s'agit, comme le note ZERMELO (1), uniquement « de développement » et de précision d'idées dont la « quintessence » se trouve dans les six mémoires de 1883-1884. D'une part théorie de l'infini actuel comme fondement de l'édifice, d'autre part les deux arithmétiques cardinale et ordinale du transfini, avec, s'ajoutant à la seconde, la nouvelle théorie des types d'ordre. La caractéristique de ces derniers exposés est d'achever l'identification entre puissances, nombres transfinis d'une part, nombres de l'arithmétique ordinaire d'autre part : les règles de calcul simplement indiquées en 1883 sont développées et précisées, le rapport aux collections d'objets est posé comme le même que pour les nombres finis.

Théorie réalisée de l'infini actuel qui se poursuit à travers la correspondance rassemblée dans les *Communications sur la doctrine du transfini*. Comme le fait observer FRAENKEL, à aucun moment il ne s'agit d'une construction arbitraire de théorie abstraite justifiée par la seule non contradiction de ses axiomes. CANTOR avait bien fait, en 1883, la distinction entre *réalité immanente* ou *intrasubjective* et réalité *transiente* ou *transsubjective*, la seconde réservée aux objets de la physique, tandis que les objets de la mathématique « dont l'essence est la liberté » (2) ne peuvent prétendre qu'à la première, puisque leur création n'est liée que par « la restriction, qui va de soi, que leurs concepts soient non contradictoires ». Mais dans le même passage apparaît aussi un nouveau critère pour qu'à « l'arbitraire soit laissé le moindre jeu..., chaque concept mathématique porte en

(1) CANTOR (XXVIII), p. 246.
(2) CANTOR (XV), p. 182.

lui le correctif nécessaire ; est-il infructueux ou incommode, il le manifeste bientôt par son caractère inutilisable (1) ». Cette fécondité ou stérilité possible n'a-t-elle pas son fondement dans la nature, même la nature physique ? A côté de ses applications à la mécanique, à la physique, à la biologie — sur lesquelles CANTOR revient à plusieurs reprises — la justification de la théorie est dans le fait de l'existence d'ensembles infinis, comme les points rationnels ou irrationnels d'une droite. Des définitions de la puissance, puis des nombres ordinaux et cardinaux le montrent assez : « j'entends par *puissance* ou nombre cardinal d'un ensemble M, écrit-il à LASSWITZ en 1884, le concept universel ou générique (universale) que l'on obtient en faisant abstraction pour l'ensemble aussi bien de la constitution de ses éléments que de toutes les relations que ces éléments ont entre eux ou avec d'autres choses donc en particulier aussi de l'ordre qui règne entre eux, et ne considère que ce qui est commun à tous les ensembles *équivalents* à M (2). » Il ne s'agit même pas d'une définition par postulats, relative à une opération — l'application — comme semble l'indiquer la dernière phrase. Ailleurs — même mémoire — la puissance de l'ensemble est le « concept général dans lequel les éléments comme des sortes d'*uns* sont réunis en un tout unifié, de sorte que chacun d'eux n'a pas de rang privilégié (3) ». Définition reprise en 1897 — contre laquelle pouvait s'élever ZERMELO (4) parce qu'impropre à fonder une science abstraite — mais caractéristique de la théorie en tant que réaliste : « le nombre cardinal est lui-même un ensemble déterminé formé de la réunion de simples *uns* qui a son existence dans notre esprit comme image intellectuelle ou pro-

(1) *Ibid.*
(2) Lettre publiée *in* CANTOR (XXIII), p. 387.
(3) *Ibid.*, p. 379.
(4) CANTOR (XXVIII), p. 351.

jection de l'ensemble M (1) ». Si les *uns* forment une pluralité quoique rien ne les distingue plus entre eux, c'est qu'ils ne sont que la silhouette incolore des objets concrets eux-mêmes pour lesquels on tient compte du fait qu'ils sont distincts, en négligeant le contenu de cette distinction. D'où l'importance de la question qu'il y ait effectivement un infini actuel, la recherche des cautions dans l'histoire, depuis PYTHAGORE. D'où également le rejet de l'infiniment petit actuel dont STOLZ (2), DU BOIS REYMOND (3) et VÉRONÈSE (4) essayaient de faire la théorie. Non pas seulement les difficultés techniques de celle-ci, l'impossibilité, par exemple, d'utiliser la division par nombres transfinis puisqu'on n'obtient pas de résultat univoque, mais surtout l'origine intuitive de la théorie des ensembles et l'exigence de ces méthodes protestent contre toute assimilation (5). Elle part, note FRAENKEL, « du concept dans une certaine mesure intuitif d'ensemble... la mise en ordre [des nombres transfinis] et les lois du calcul sont par suite non pas décidées par des actes arbitraires de définition, mais dans une certaine mesure décelées dans la nature, avec leur structure imposée. Ainsi apparaissent les règles de calcul très hétérogènes entre elles pour diverses espèces de nombres qu'il n'y aurait eu, *a priori*, aucune raison de poser ainsi et... [auxquelles] ... par des considérations philosophiques sur l'infiniment grand on ne serait certainement pas venu (6) ». Parmi ces lois, l'une des plus importantes est celle qui pose une limite à toute suite de nombres croissants, loi qui est une sorte d'extension de l'acte de compter :

(1) CANTOR (XXVI), p. 283.
(2) STOLZ (II).
(3) DU BOIS REYMOND (IX).
(4) VÉRONÈSE (I).
(5) « Le fait de l'existence des nombres infiniment grands est donc si peu une raison pour l'existence des grandeurs infiniment petites que bien plus justement avec l'aide de ceux-là l'impossibilité de ceux-ci est démontrée. » CANTOR (XXIII), p. 408.
(6) FRAENKEL (X), p. 118-119.

définir après toute suite infinie un nombre immédiatement supérieur, affirmer en quelque sorte l'empire sans borne des nombres, axiome d'ARCHIMÈDE prolongé. D'où la seconde raison pour rejeter l'infiniment petit : il ne s'insère que dans un système non archimédique. « Ce qu'on appelle l'*axiome d'Archimède* n'est pas un axiome, mais une proposition découlant avec une nécessité logique du concept de grandeur linéaire (1). »

Les lois naturelles du calcul avec les nombre cardinaux supposent donc au préalable une définition des rapports des ensembles entre eux lorsqu'on fait abstraction de l'ordre : ainsi sont introduites en 1895 les notions de partie d'ensemble — restreinte au cas de la vraie partie (c'est-à-dire distincte du tout) — de réunion, ou ensemble somme (restreinte à la réunion d'ensembles sans éléments communs), d'ensemble produit de deux ensembles M et N : ensemble de tous les couples d'éléments pris l'un dans M l'autre dans N (définition évidemment équivalente à celle qui avait été donnée en 1883) ; enfin de l'exponentiation d'un ensemble par un autre : ensemble de toutes les *façons de recouvrir* (BELEGUNGEN) un ensemble par les éléments d'un autre, c'est-à-dire de toutes les fonctions dont l'argument est pris dans le premier ensemble et la valeur dans le second. Opérations concrètes, que complète l'acte également concret de la coordination biunivoque des éléments de deux ensembles équivalents. En regard, les rapports qui s'en déduisent pour les nombres cardinaux de ces ensembles : de l'équivalence se déduit l'égalité, ce qui doit être démontré puisque le processus d'abstraction définissant le nombre est posé comme différent du processus d'application (différence imaginative, qu'il est caractéristique de voir CANTOR maintenir ici). De la relation entre ensemble et partie, les inégalités numériques : « s'il n'y a aucune partie de M équiva-

(1) CANTOR (XXIII), p. 409.

lente à N et s'il y a une partie N_1 de N équivalente à M », on pose (1) :

$$\overline{\overline{M}} < \overline{\overline{N}}$$

transitivité, exclusion mutuelle des 3 relations $=, >, <$ sont démontrées. En revanche « il ne va nullement de soi et ne pouvait être ici démontré que pour tout couple de nombres cardinaux une de ces 3 relations doive être nécessairement réalisée (2) ». C'est la première fois que la comparabilité universelle ne fait pas l'objet d'une simple affirmation ; les corollaires, dont l'établissement est réservé avec elle, car « ils s'en déduisent immédiatement », précisent la distinction des deux autres cas qui peuvent se présenter d'un point de vue purement logique, dans les relations entre deux ensembles :

« il y a une partie N_1 de N équivalente à M et il y a aussi une partie M_1 de M équivalente à N ;

il n'y a ni une partie de N équivalente à M ni une partie de M équivalente à N (3). »

L'un des corollaires exclut la deuxième hypothèse ; l'autre affirme que dans la première $\overline{\overline{M}} = \overline{\overline{N}}$: c'est le théorème de l'équivalence dont SCHRÖDER et BERNSTEIN devaient donner la démonstration sans invoquer la comparabilité (4).

(1) CANTOR représente la puissance d'un ensemble M par $\overline{\overline{M}}$ pour rappeler la double abstraction (propriétés des éléments et leur ordre) et son type d'ordre par $\overline{M}$ — puisqu'on ne fait plus qu'une abstraction.

(2) CANTOR (XXVI), p. 285.

(3) L. à DEDEKIND, 30 août 1899, CANTOR (XXVIII), p. 450.

(4) La démonstration est d'ordre élémentaire (cf. l'exposé qu'en donne, en 1898, E. BOREL (I), p. 104-106). Soient M, T et P :

$$P \sim M \text{ et } P \subset T \subset M.$$

Posons $M - T = R \quad T - P = Q$

L'application f de M sur P donne :

$$P = f(P + Q + R) = f(P) + f(Q) + f(R) = P_1 + Q_1 + R_1$$

$P_1 = f(P_1 + Q_1 + R_1) = f(P_1) + f(Q_1) + f(R_1) = P_2 + Q_2 + R_2$ etc. en posant $P_1 = f(P)$, etc.

Si on pose P_ω = intersection de $P. P_1. P_2 \ldots$ on a :

$$P = P_\omega + Q_1 + R_1 + Q_2 + R_2 + Q_3 + R_3 + \ldots \sim$$

$$\sim P_\omega + Q + R_1 + Q_1 + R_2 + Q_2 + R_3 + \ldots = Q + P = T.$$

puisque $P_{n-1} - P_n = Q_n + R_n$

(HAUSDORFF (I), p. 50).

Pour celle-ci, déclare CANTOR, « sa vérité ne pourra être reconnue que plus tard lorsque nous aurons un aperçu de la suite croissante des nombres cardinaux transfinis et une idée de leur enchaînement », c'est-à-dire après l'exposé de la théorie des nombres ordinaux. Difficulté dans le plan qui en manifeste une plus grave : le développement, — et la justification interne de la théorie telle qu'elle est présentée exige que l'on ne fasse appel qu'aux intuitions évoquées par les définitions des rapports entre ensembles. Utiliser la théorie du bon ordre est, sinon commettre un cercle vicieux — puisque les résultats invoqués dans la théorie ordinale sont préalables à l'affirmation de la comparabilité — du moins porter atteinte aux fondements de l'édifice : il ne subsiste que si les intuitions naïves originelles lui suffisent. D'où le recours ultérieur à l'axiomatisation : c'est la difficulté de la comparabilité, avant les paradoxes, qui en est la première raison.

Les résultats effectifs de la théorie cardinale ne peuvent être qu'assez pauvres : on se borne à retrouver les propriétés habituelles de l'addition et de la multiplication (commutativité, associativité, distributivité), de l'exponentiation :

$$a^b . a^c = a^{b+c} \qquad (a^b)^c = a^{bc}$$

Pour aller plus loin la première distinction qui s'impose est celle des opérations effectuées sur un nombre fini ou un nombre infini de termes, c'est-à-dire, puisque la théorie prétend être préalable à toute discipline mathématique, l'étude des ensembles finis et celle de l'infini dénombrable.

La théorie des ensembles finis manque de rigueur. Dans les *Communications sur la doctrine du transfini* (1) (1884), la définition donnée suppose déjà une notion d'ordre : « nous entendons par ensemble fini un ensemble M qui est engendré par une telle adjonction successive d'éléments à

(1) CANTOR (XXIII) ; le mémoire a été publié en 1887 mais est daté de 1884.

un élément originairement donné, que *inversement* par éloignement successif des éléments dans l'ordre opposé, l'élément origine peut être retrouvé (1) ». Cette évidente difficulté est évitée par suppression de toute définition générale : les nombres cardinaux finis sont obtenus chacun à partir du précédent par adjonction d'un objet à l'ensemble qu'il dénotait. On empêche par artifice que, comme indice du nouvel objet, figure le nombre cardinal dans sa propre définition, mais on ne peut éviter que l'emploi de l'induction complète pour les propriétés élémentaires — comme : toute partie d'un ensemble ayant pour nombre cardinal ν possède pour nombre cardinal un de ceux qui précèdent ν — représente un cercle vicieux. La seule démonstration effective est celle de la proposition que si un ensemble M n'est équivalent à aucune de ses parties, l'ensemble $\{M,e\}$ formé par adjonction à M d'un nouvel élément e possède la même propriété (2). En 1884 la propriété est signalée comme « caractéristique des ensembles finis », donc pouvant servir de définition, ce que faisait déjà DEDEKIND. Mais la notion d'engendrement est essentielle chez CANTOR. Le plus petit infini est construit : « l'ensemble de tous les nombres cardinaux finis nous offre l'exemple le plus proche d'un ensemble transfini ». Il est celui dont le nombre cardinal est le plus petit ; en effet tout ensemble transfini possède un sous-ensemble qui lui est équivalent : il suffit d'observer que la suppression dans un ensemble transfini d'un nombre fini d'éléments laisse toujours la possibilité d'en supprimer un autre ; l'ensemble de toutes les suppressions est donc l'ensemble des nombres cardinaux finis, c'est-à-dire le dénombrable. La propriété caractéristique du transfini est intuitive pour le dénombrable : on peut supprimer un élément de

(1) *Ibid.*, p. 415*n*.

(2) Le raisonnement employé est exactement celui qui sert à DEDEKIND *in* DEDEKIND (III).

l'ensemble sans que sa puissance change, il est donc équivalent à une de ses parties. Propriété qu'il transmet à tous les ensembles transfinis puisqu'il en est sous-ensemble. Enfin si on désigne par $\aleph_0$ son nombre cardinal on a les équations :

$$\aleph_0 + n = \aleph_0 \quad \aleph_0 + \aleph_0 = \aleph_0 \quad \aleph_0 . n = \aleph_0 \quad \aleph_0 . \aleph_0 = \aleph_0$$

Le continu, puisque ensemble, par exemple, de tous les développements dyadiques illimités, c'est-à-dire de toutes les fonctions dont l'argument parcourt un ensemble dénombrable et qui prennent les valeurs 0 et 1, a pour nombre cardinal $2^{\aleph_0}$. D'où en particulier une nouvelle démonstration de l'invariance de la puissance des continus quel que soit leur nombre de dimensions (fini ou infini dénombrable) :

$$(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 . \aleph_0} = 2^{\aleph_0}$$

L'étude n'est pas poussée plus loin ; aussitôt se pose la question du nombre cardinal suivant immédiatement $\aleph_0$: renvoi à la notion d'ordre.

La première mention des types d'ordre est faite dans le mémoire *Sur différents points de vue relatifs à l'infini actuel*: le point de vue abstrait nous permet de saisir l'infini actuel « sous forme... des nombres transfinis ou sous la forme encore plus générale des *types d'ordre transfinis* » (1). A la fin des *Communications sur la doctrine du transfini* se trouve esquissée une étude générale de ces types, dont les nombres ordinaux ne sont qu'une espèce particulière : le type est « le concept universel qui se dégage d'un ensemble M — si nous faisons abstraction seulement de la constitution des éléments, mais conservons les relations d'ordre entre eux ». L'élargissement de la théorie ordinale (qui ne considérait que les ensembles bien ordonnés) en théorie des

(1) CANTOR (XXII), p. 372.

types est nécessaire pour une caractérisation abstraite du continu : la dissociation entre topologie et arithmétique transfinie s'achève entre 1884 et 1887. S'il est possible d'édifier de façon autonome une théorie des puissances où s'insère le continu, il faut que celui-ci puisse être, en tant que nombre, défini indépendamment de toute considération géométrique ; aussi bien l'opération simple de la dérivation s'était-elle montrée insuffisante : la solution sera cherchée dans l'analyse, qu'elle amorçait, de l'ordre entre éléments. Le mémoire de 1887 ne s'occupe que de distinguer entre ordres multiples, contre-coup des efforts pour définir la notion de dimension. En 1895 CANTOR revient au seul ordre linéaire et démontre les deux importants théorèmes qui caractérisent les types de l'ensemble des nombres rationnels et du continu.

Au préalable, une théorie générale des types : l'opération correspondant à l'application dans la théorie cardinale et définissant l'égalité des types est l'application *semblable*, conservant les relations d'ordre entre éléments. Puisqu'elle est cas particulier de la 1re, on voit qu'à tout nombre cardinal $\mathfrak{A}$ correspond la classe $[\mathfrak{A}]$ de tous les types qu'une application non semblable transforme les uns dans les autres. De là deux nouvelles notions : celle de *type inverse*, ou conjugué — déjà introduite en 1887 — obtenu en inversant toutes les relations d'ordre dans un ensemble : ainsi l'inverse du type ω de l'ensemble des nombres entiers rangés par ordre de grandeur est $^*\omega$ type de l'ensemble des nombres entiers négatifs rangés par ordre de grandeur. Celle d'*automorphisme* : application semblable d'un ensemble sur lui-même distincte de l'identité ; certains types ω (en général tous les types d'ensembles bien ordonnés) ne permettent pas l'automorphisme, au contraire celui de l'ensemble de tous les entiers (négatifs et positifs), des nombres rationnels, etc. (en général ceux qui ne définissent ni premier ni dernier élément) sont applicables d'une infi-

nité de manières sur eux-mêmes. L'addition et la multiplication se définissent comme pour les ordinaux en 1883 ; pour le produit nouvelle présentation : si M est le multiplicande, N le multiplicateur, on considère l'ensemble des éléments m_n parcourant pour chaque indice n un ensemble semblable à M, l'indice parcourant N et l'ordre introduit par lui l'emportant sur l'ordre entre 2 éléments m et m' de M ; on peut ainsi généraliser au produit de k facteurs, l'élément est de la forme $m_{n_p q \ldots}$. Comme on voit immédiatement, l'addition n'est pas commutative non plus que la multiplication ; toutes deux sont associatives et la multiplication possède la distributivité du multiplicateur :

$$\overline{M}.\ (\overline{N} + \overline{P}) = \overline{M}.\ \overline{N} + \overline{M}.\overline{P}$$

Le *type* η de l'ensemble des nombres rationnels rangés par ordre de grandeur est défini par les 3 propriétés : 1° appartenir à la classe $[\aleph_0]$; 2° n'avoir ni premier ni dernier élément ; 3° être tel qu'entre 2 éléments de l'ensemble l'ordre en insère d'autres : *densité*. La 3e propriété est la première transposition abstraite d'une propriété topologique. La démonstration que tout ensemble possédant ces trois propriétés est *semblable* à l'ensemble des nombres rationnels s'effectue grâce au procédé qui avait procuré l'application du continu sur tout ensemble parfait non dense : la première étape réalisait une application semblable de l'ensemble des nombres rationnels sur celui des intervalles contigus à l'ensemble parfait non dense. Ici ce dernier ensemble est remplacé par un ensemble quelconque M possédant les trois propriétés de la définition : la 3e propriété, densité, permet de coordonner progressivement (la possibilité d'une coordination progressive venant de la 1re propriété : les deux ensembles peuvent être rangés en une suite numérotée par les entiers positifs), en respectant l'ordre, à tout nombre rationnel un élément de M et un seul, et réciproquement

de prouver que tout élément de M est ainsi pris dans la coordination.

Pour atteindre le continu, CANTOR décalque de même dans l'abstrait ses raisonnements sur les nombres réels : aux suites convergentes fondamentales définissant un nombre irrationnel correspondent ici des suites fondamentales abstraites croissantes ou décroissantes : l'ordre d'un ensemble étant conservé entre les éléments des ensembles partiels distingués en son sein, ceux-ci possèdent par là même un type d'ordre déterminé (en général distinct du type de l'ensemble total) ; on appellera suites fondamentales croissantes les ensembles partiels de type ω, suites fondamentales décroissantes les ensembles de type $^*\omega$. Deux suites, croissantes ou décroissantes toutes deux, sont conjuguées si à tout élément de l'une correspond dans l'autre un élément supérieur dans l'ordre. Lorsque l'une est croissante et l'autre décroissante, elles sont conjuguées : si tout élément de la 1re précède tout élément de la 2e et si entre elles l'ordre insère au maximum un élément de l'ensemble total M étranger à chacune d'elles. Une suite fondamentale possède dans M un élément limite — ou élément principal dans M — s'il existe un élément m_o tel que pour tout élément a_ν de la suite on ait :

$a_\nu \prec m_o$ (suite croissante) ou $a_\nu \succ m_o$ (suite décroissante) et que, si pour tout élément $m \prec m_o$ dans le 1er cas, $m \succ m_o$ dans le 2e, on puisse trouver dans la suite un rang ν_o tel que pour tout $\nu \geqq \nu_o$ on ait :

$$m \prec a_\nu \text{ dans le 1er cas} \qquad m \succ a_\nu \text{ dans le 2e.}$$

On voit qu'une suite fondamentale — avec toutes les suites qui lui sont conjuguées — ne peut posséder qu'un seul élément limite et que dans l'application semblable les caractères suite fondamentale et élément limite sont invariants. Enfin les notions de densité en soi, fermeture, perfection subissent la même transformation abstraite : un ensemble

ordonné est *dense en soi* si tous ses éléments sont éléments principaux, *fermé* s'il contient un élément limite pour toute suite fondamentale, *parfait* s'il est à la fois dense en soi et fermé. D'où la caractérisation du *type* θ du continu linéaire : c'est un ensemble ordonné M : 1° parfait ; 2° contenant un ensemble S dénombrable dense dans M. En supprimant de S le premier et le dernier élément s'ils existent, ce qui ne change pas la propriété de S utilisée dans l'énoncé, on voit que S a le type η. La démonstration reproduit la 2e partie de la démonstration appliquant le continu sur un ensemble parfait non dense. L'application semblable de S sur l'ensemble R des nombres rationnels étant réalisée d'après le 1er théorème, on détermine la correspondance entre les autres éléments par celle déjà établie entre les suites fondamentales d'éléments de R et de S dont ils sont respectivement éléments limites. Comme le remarque ZERMELO (1), la démonstration est plus simple avec la définition DEDEKIND du continu comme ensemble ordonné sans saut (c'est-à-dire dense) ni lacune (c'est-à-dire tel que toute coupure soit déterminée par un élément). D'autre part la 1re propriété exigée de θ est trop forte ; il suffit de le définir comme fermé puisque d'après la 2e propriété il est dense et que la densité entraîne évidemment la densité en soi (2).

Le premier objectif de la théorie du bon ordre est l'établissement de l'universelle comparabilité des ensembles bien ordonnés. Il est atteint grâce à la notion de *segment* : ensemble des éléments précédant dans l'ordre un élément d'un ensemble ; le complémentaire du segment est son *reste*. Le caractère unique d'un segment dans un ensemble — impossibilité d'établir une application semblable entre

(1) CANTOR (XXVIII), p. 354.

(2) Si entre deux éléments d'un ensemble il y en a toujours d'autres, tout point sera point limite.

lui et, soit l'ensemble tout entier, soit un autre segment — est démontré un peu longuement par CANTOR : ZERMELO simplifia depuis, grâce au théorème que, dans toute application semblable d'un ensemble bien ordonné sur une de ses parties, aucun élément ne peut être appliqué sur un élément qui le précède (1). Comme la similitude est transitive, il n'y a qu'un moyen d'appliquer semblablement un ensemble M sur un autre N ou sur un de ses segments N. La mise en correspondance des segments d'un ensemble M avec ceux d'un ensemble N — par application semblable des segments homologues l'un sur l'autre — définit une application semblable de M sur N et réciproquement.

Enfin le résultat cherché ; étant donnés deux ensembles bien ordonnés M et N, trois cas seulement sont possibles : ou M est semblable à N, ou M est semblable à un segment de N, ou N à un segment de M. Qu'ils soient chacun semblables à un segment de l'autre est impossible en vertu des théorèmes précédents. Qu'ils ne soient ni semblables entre eux, ni l'un semblable à un segment de l'autre est également exclu : en effet si tous les segments de M ont des homologues semblables dans N et réciproquement on a vu que la similitude de M et N s'en déduirait ; si un segment de M n'a pas d'homologue semblable dans N il y en a un plus petit M_1, jouissant de cette propriété : ou bien tout segment de N a un segment semblable dans M (donc dans M_1) donc M_1 est semblable à N ; ou bien il y a également des segments de N sans homologue semblable dans M, soit

(1) En effet, le sous-ensemble des éléments jouissant de cette propriété posséderait un premier élément a, donc on aurait (f étant la fonction application) :

$$f(a) = a' \prec a ;$$

donc comme l'application est semblable $f(a') \prec f(a) = a'$ donc l'élément a' précédant a jouirait de la propriété, contrairement à la définition de a. On voit qu'une application semblable d'un segment sur un autre segment du même ensemble produirait ce résultat.

N_1 le plus petit : M_1 et N_1 sont semblables (puisque par hypothèse tous les segments plus petits, donc tous leurs segments, s'appliquent semblablement les uns sur les autres) contrairement à leur définition.

Par la comparabilité des ensembles, est en même temps définie celle des types d'ordre ou nombres ordinaux, qui les dénotent: une et une seule des trois relations [1] $=, >, <$ existe nécessairement entre deux d'entre eux. L'arithmétique transfinie qui en découle procure les trois derniers objectifs du mémoire: démonstration que tout ensemble de nombres ordinaux rangés par ordre de grandeur est bien ordonné ; règles précises de calcul et de décomposition canonique d'un nombre en un produit de facteurs irréductibles ; enfin exploration de la classe II avec l'un des résultats les plus intéressants, définition des *nombres* ε. Mais dès ce moment se manifeste la gêne introduite par la découverte du paradoxe de BURALI-FORTI : à aucun des raisonnements faits ici n'est accordée sa portée véritable ; sauf la définition générale des opérations, qui reproduit les définitions données pour les types d'ordre, avec en plus chaque fois la démonstration que l'ensemble résultat est également bien ordonné (ce qui est évident aussi bien pour le produit que pour la somme), toutes les considérations sont restreintes aux nombres ordinaux de la classe II, alors que la plupart seraient immédiatement valables pour les ordinaux de n'importe quelle classe. Ainsi l'ensemble de tous les ordinaux inférieurs à un nombre ordinal donné α est bien ordonné et de type α : soit M un ensemble bien ordonné de type α, à tout nombre $\alpha' < \alpha$ correspond un segment de M et réciproquement, les éléments déterminent les segments de M, et les nombres inférieurs à α peuvent donc être mis en correspondance biunivoque, application qui conserve évidemment l'ordre. Si on considère

(1) Nous remplacerons désormais la relation $\prec$ par $<$ puisque le nombre ordinal a tous les caractères d'un nombre.

un ensemble quelconque P de nombres ordinaux transfinis (en particulier la classe II), l'intersection de chacune de ses parties (et de P lui-même) avec le segment déterminé par un élément de P possède un 1er élément qui est évidemment aussi 1er élément de la partie (ou de P) donc P est bien ordonné.

Les opérations n'avaient en principe été définies que pour un nombre fini de termes : on voit sans peine que la réunion d'une suite infinie d'ensembles bien ordonnés est bien ordonnée ; de là, en particulier, les notions de suite fondamentale (c'est-à-dire croissante) de nombres ordinaux et de limite correspondante : soit une suite $\alpha_1, \alpha_2, \ldots$; si on fait la somme :

$$\beta = \alpha_1 + (\alpha_2 - \alpha_1) + (\alpha_3 - \alpha_2) + \ldots$$

β est un nombre ordinal supérieur immédiatement à la suite, c'est-à-dire tel que tout $\alpha_\nu < \beta$ et que pour tout nombre $\gamma < \beta$ il y a des $\alpha_\nu > \gamma$, on posera :

$$\beta = \lim_{\nu} \alpha_\nu .$$

Tout nombre de la classe II est obtenu soit par adjonction de 1 à un nombre déjà connu (il est dit de 1re *espèce*), soit par passage à la limite d'une suite fondamentale (2e *espèce*) : de chaque façon on a donc le nombre qui suit immédiatement les nombres déjà posés (on reconnaît les deux principes d'engendrement du mémoire de 1883).

L'exponentiation n'avait été définie que pour le cas d'exposants finis, au delà l'intuition ne peut servir. C'est alors qu'intervient pour la première fois la récurrence transfinie. « Il y a une et une seule fonction complètement déterminée $f(\xi)$ de la variable ξ [parcourant la classe II] qui satisfait aux conditions suivantes (γ et δ étant des constantes, $\delta > 0$, $\gamma > 1$) :

1) $f(0) = \delta$

2) si $\xi' < \xi''$ $\quad f(\xi') < f(\xi'')$

3) pour tout ξ $\quad f(\xi + 1) = f(\xi).\gamma$

4) si $\{\xi_\nu\}$ est une suite fondamentale quelconque, $\{f(\xi_\nu)\}$ est aussi une suite fondamentale et si on a :

$$\lim_\nu \xi_\nu = \xi \qquad \text{on a } \lim_\nu f(\xi_\nu) = f(\xi). \text{ »}$$

La démonstration se fait par induction transfinie : pour les ξ finis ($\xi = 1, 2, 3\ldots$) on obtient de proche en proche $f(\xi)$ grâce à la condition 3) ; et la condition 2) procure pour $f(\omega)$ et de façon générale pour tout nombre de 2^e espèce des suites fondamentales qui déterminent f de façon univoque. Le bon ordre garantit l'extension du procédé à toute la classe II : si des nombres y échappaient, il y en aurait un plus petit α, mais les deux procédés 3) et 4) permettraient alors de calculer $f(\alpha)$. On voit la dépendance de la nouvelle induction par rapport aux deux principes d'engendrement (qui impliquent la notion de bon ordre) : toute propriété qu'ils laissent invariante sera démontrable par elle.

Si on pose $\delta = 1$ on écrit :

$$f(\xi) = \gamma^\xi ;$$

en effet dans le cas de ξ fini on retrouve bien la définition ordinaire ; d'autre part les propriétés caractéristiques sont conservées :

$$\gamma^\eta . \gamma^\xi = \gamma^{\eta+\xi} \qquad (\gamma^\eta)^\xi = \gamma^{\eta.\xi} ;$$

pour démontrer la 1re égalité il suffit en effet de poser $\delta = \gamma^\eta$ et de considérer la fonction $f(\xi) = \gamma^{\eta+\xi}$: on voit que $\gamma^\eta.\gamma^\xi$ satisfait aux quatre conditions ; on a donc l'égalité en vertu de l'unicité ; de même pour la 2^e, si on pose $\delta = 1$ et si on remplace γ par γ^η. D'une façon générale toute fonction satisfaisant aux quatre conditions est de la forme :

$$f(\xi) = \delta . \gamma^\xi .$$

Les règles du calcul se caractérisent par leur harmonieuse simplicité, qui découle de la distributivité du multiplica-

teur dans la multiplication et de la propriété de la somme.

$$\alpha + \beta = \beta \qquad \text{si } \alpha < \beta$$

on en déduit d'abord la forme normale univoque sous laquelle peut se mettre tout nombre de la classe II.

$$\alpha = \omega^{\alpha_0} k_0 + \omega^{\alpha_1} k_1 + \ldots + \omega^{\alpha_r} k_r \qquad (1)$$

où les k_0, k_1 ... k_r sont des entiers finis, les α_0 ... α_r des nombres de la 1re ou de la 2e classe tels que :

$$\alpha \geqq \alpha_0 > \alpha_1 > \ldots > \alpha_r \geqq 0$$

Il suffit de procéder de proche en proche : parmi les nombres ξ tels que $\omega^\xi > \alpha$, il y en a un plus petit β, qui est de 1re espèce, on pose $\beta = \alpha_0 + 1$; on a donc :

$$\omega^{\alpha_0} \leqq \alpha < \omega^{\alpha_0} \cdot \omega$$

et la 2e inégalité montre qu'il y a un plus grand nombre fini k_0 tel que :

$$\omega^{\alpha_0} k_0 \leqq \alpha$$

on obtient donc :

$$\alpha = \omega^{\alpha_0} k_0 + \alpha' \quad \text{avec} \quad \alpha' < \omega^{\alpha_0} < \alpha$$

et ainsi de suite : l'opération ne peut se renouveler qu'un nombre fini de fois puisque la suite $\alpha > \alpha' > \ldots$ est décroissante.

α_0 est appelé le degré de α. D'où la définition des quatre opérations effectuées sur les formes normales ; dans l'addition ne sont conservés du 1er sommant que les termes de degré supérieur au degré du 2e :

$$\alpha + \beta = \omega^{\alpha_0} k_0 + \ldots + \omega^{\alpha_\rho} k_\rho + \omega^{\beta_0} \lambda_0 + \ldots + \omega^{\beta_\sigma} \lambda_\sigma.$$

Dans la multiplication si le multiplicateur est de 2e espèce ($\beta_\sigma > 0$) on a :

$$\alpha\,\beta = \omega^{\alpha_0 + \beta_0} \lambda_0 + \ldots + \omega^{\alpha_0 + \beta_\sigma} \lambda_\sigma = \omega^{\alpha_0} \beta \,;$$

si le multiplicateur est de 1re espèce, en vertu de la distri-

butivité tous les termes du multiplicande paraissent : ($\beta_\sigma = 0$)

$$\alpha\beta = \omega^{\alpha_0} + \beta_0\lambda_0 + \ldots + \omega^{\alpha_0} + \beta_{\sigma-1}\lambda_{\sigma-1} + \omega^{\alpha_0}k_0\lambda_\sigma + \omega^{\alpha_1}k_1 + \ldots + \omega^{\alpha_\tau}k_\tau$$

Enfin pour l'exponentiation si l'exposant est de 2e espèce :

$$\alpha^\beta = \omega^{\alpha_0\beta}$$

si l'exposant est de 1re espèce: $\beta = \beta' + \lambda_\sigma$, β' étant de 2e espèce,

$$\alpha^\beta = \omega^{\alpha_0\beta'} \cdot \alpha^{\lambda_0}$$

De la 2e règle de la multiplication se déduit la décomposition univoque d'un nombre en ses facteurs premiers ; on peut, en effet, écrire α sous forme du produit :

$$\alpha = (\omega^{\alpha_1}k_1 + \ldots + \omega^{\alpha_\tau}k_\tau)\ (\omega^{\alpha_0 - \alpha_1}k_0 + 1)$$

En itérant jusqu'à α_τ on obtient donc :

$$\alpha = \omega^{\alpha_\tau}k_\tau\ (\omega^{\alpha_{\tau-1} - \alpha_\tau}k_{\tau-1} + 1) \ldots (\omega^{\alpha_0 - \alpha_1}k_0 + 1)$$

Et comme, suivant la même règle :

$$(\omega^\lambda k + 1) = (\omega^\lambda + 1)\ k$$

si on pose $\gamma_0 = \alpha_\tau\ \gamma_1 = \alpha_{\tau-1} - \alpha_\tau$, etc., on obtient la décomposition canonique :

$$\alpha = \omega^{\gamma_0}k_\tau\ (\omega^{\gamma_1} + 1)\ k_{\tau-1}\ (\omega^{\gamma_2} + 1)\ k_{\tau-2} \ldots (\omega^{\gamma_\tau} + 1)\ k_0$$

Si on ne retient des conditions pour l'exponentiation que 2) et 4) on définit ce que HAUSDORFF (1) devait appeler les *fonctions normales*. Elles se caractérisent par l'intéressante propriété de posséder des *nombres critiques*. Pour tout argument ξ d'une telle fonction on a :

$$f(\xi) \geqq \xi$$

(sans quoi f déterminerait une application semblable de la

(1) HAUSDORFF (I), p. 114.

classe II sur elle-même où un élément serait appliqué sur un élément qui le précède). Les nombres critiques sont ceux pour lesquels :

$$f(\xi) = \xi$$

CANTOR n'étudie que la seule fonction γ_ξ — mais ses raisonnements valent pour toute fonction normale — : les nombres critiques correspondants sont les *nombres-ε*. Leur existence est démontrée par une construction (et résulte immédiatement des propriétés des suites fondamentales) : soit un nombre α qui ne soit pas critique :

$$f(\alpha) > \alpha$$

en posant $\alpha_1 = f(\alpha)$ $\alpha_2 = f(\alpha_1)$, etc., ... on obtient une suite fondamentale (propriété 2). Sa limite (propriété 4) est donnée par :

$$f(\beta) = f\left(\lim_\nu \alpha_\nu\right) = \lim_\nu f(\alpha_\nu) = \lim_\nu \alpha_{\nu+1} = \beta$$

β est le 1er nombre critique supérieur à α. On voit immédiatement (1) en effet qu'il ne peut y avoir d'autre nombre critique entre α et β, d'autre part que le 1er nombre critique après β est obtenu par le même procédé en posant $\alpha = \beta + 1$, enfin que toute suite fondamentale de nombres critiques a pour limite un nombre critique. L'ensemble des nombres critiques relatifs à une fonction donnée s'obtient donc par deux procédés semblables aux deux principes d'engendrement des nombres de la classe II : il est bien ordonné et semblable à celle-ci. Pour la fonction γ^ξ, quel que soit γ, en vertu

(1) En effet, on aurait, comme $\alpha < \gamma < \beta$, $f(\alpha) = \alpha_1 < f(\gamma) = \gamma$ de même $\alpha_2 < \gamma$, etc., ... donc $lim_\nu\, \alpha_\nu = \beta < \gamma$. Par suite, comme $\beta + 1$ suit immédiatement β, s'il y avait un nombre critique entre β et ce nouveau nombre, il serait également entre $\beta + 1$ et lui, ce qu'on vient de voir être impossible. Enfin soit la suite $\beta_1\, \beta_2 \ldots \beta_\nu \ldots$ de nombres critiques, on sait qu'on a :

$$f\,(lim_\nu\, \beta_\nu) = lim_\nu\, f\,(\beta_\nu) = lim_\nu\, \beta_\nu$$

donc $lim_\nu\, \beta_\nu$ est un nombre critique.

de la forme canonique des nombres de la classe II, il n'y a qu'un ensemble de nombres critiques, dont on obtient le plus petit ε_0 en posant $\gamma = \omega$: ε_0 est évidemment la limite de la suite :

$$\omega,\ \omega^{\omega},\ \omega^{\omega^{\omega}} \ldots$$

Pour tout nombre α inférieur à un ε on a les trois égalités : (1)

$$\alpha + \varepsilon = \varepsilon \qquad \alpha\varepsilon = \varepsilon \qquad \alpha^{\varepsilon} = \varepsilon.$$

Telle est cette théorie des ordinaux transfinis — dernier étage de l'édifice cantorien — dont l'harmonie, la pureté de lignes devaient exciter l'admiration d'HILBERT. On a vu qu'elle découle tout entière — une fois admises les définitions intuitives de la somme et du produit de deux ensembles bien ordonnés — d'un approfondissement des relations établies par les deux principes générateurs des ordinaux. La simplicité des résultats vient de l'unité interne des méthodes. Mais de là aussi un divorce avec la théorie cardinale des ensembles, avec les problèmes posés initialement : l'application semblable est décidée par une loi ; la notion de puissance est purement descriptive ; il s'agit de prouver une corrélation biunivoque possible entre système de tous les objets jouissant des propriétés *A* et *B*. Sans doute la définition des ordinaux de la classe II permet-elle d'atteindre une nouvelle puissance : grâce au 3e principe d'engendrement on a l'espoir de construire une échelle exhaustive des puissances. Mais ce principe, on l'a vu, n'est que négatif ; en réalité la définition de la suite des puissances obéit à deux nouveaux principes : le 1er fondé sur le théorème que toute classe de tous les ordinaux dénotant les ensembles

(1) Soit α_0 le *degré* de α, le degré de $\varepsilon = \omega^{\varepsilon}$ est ε, donc α a un degré inférieur à ε, donc $\alpha + \varepsilon = \varepsilon$, de même $\alpha_0 + \varepsilon = \varepsilon$. D'autre part :

$$\alpha\,\varepsilon = \alpha\,\omega^{\varepsilon} = \omega^{\alpha_0 + \varepsilon} = \omega^{\varepsilon} = \varepsilon$$

donc aussi : $\alpha_0\,\varepsilon = \varepsilon$. Enfin :

$$\alpha^{\varepsilon} = \alpha^{\omega^{\varepsilon}} = \omega^{\alpha_0 \omega^{\varepsilon}} = \omega^{\alpha_0 \varepsilon} = \omega^{\varepsilon} = \varepsilon.$$

d'une certaine puissance a possède la puissance immédiatement supérieure ; le 2e, transposition du 2e principe d'engendrement des ordinaux : toute suite croissante infinie de nombres cardinaux détermine de façon univoque un nombre cardinal immédiatement supérieur à tout nombre de la suite (ainsi se définit $\aleph_\omega$). Mais l'utilité de l'échelle ainsi constituée n'apparaît que si l'on est sûr qu'y figure tout nombre cardinal, si d'autre part on a un moyen d'y fixer sa place. Or la théorie ordinale seule ne permet de répondre ni à l'une ni à l'autre question : l'universelle comparabilité, en tant que démontrée entre les ordinaux, ne servirait que si la 1re question était résolue, loin qu'elle aide à la résoudre ; quant au 1er principe, au lieu de permettre une assimilation entre $\aleph_1$ et le continu, il repose sur un procédé qui ouvre l'ère des paradoxes menaçant de ruiner toute la théorie.

En 1895, d'après le témoignage de BERNSTEIN, (1) CANTOR eut la première idée du paradoxe que BURALI-FORTI devait exposer en 1897 : l'ensemble de tous les nombres ordinaux (tels que les peuvent procurer les deux principes d'engendrement et le principe d'arrêt) rangés par ordre de grandeur forme un ensemble bien ordonné, qui a donc pour type un nombre ordinal α ; mais celui-ci ne peut figurer dans la suite Ω car à son rang il y dénoterait un segment de type α (l'ensemble de tous les ordinaux le précédant) et Ω serait semblable à un de ses segments. Antinomie qui oblige CANTOR à restreindre son mémoire de 1897 à la considération de la classe II, à laisser l'édifice inachevé. En 1899 quelques lettres adressées à DEDEKIND (2) essaient de la rendre inoffensive, bien plus de l'utiliser pour résoudre le problème de la comparabilité : « déjà depuis deux ans je suis

(1) BERNSTEIN (I), p. 187.
(2) Publiées *in* CANTOR (XXVIII), p. 443-450.

en possession d'une démonstration qu'il n'y a pas d'autre [puissance que les alephs] ; de sorte que par exemple un aleph déterminé correspond comme nombre cardinal au continu linéaire arithmétique.

Le principe est la distinction entre *multiplicités inconsistantes* « telles que l'hypothèse d'une réunion (ZUSAMMENSEIN) de *tous* leurs éléments conduit à une contradiction, de sorte qu'il est impossible de considérer la multiplicité comme une unité, une « chose finie », et *multiplicités consistantes*, non contradictoires en tant que *choses unes*. Les secondes sont les *ensembles* dont s'occupe la théorie. Exemples des premières sont le système Ω de tous les ordinaux transfinis et celui de tous les alephs, ℶ, parce que semblable à Ω. Deux systèmes équivalents sont ou tous deux consistants ou tous deux inconsistants ; toute partie d'un ensemble est un ensemble, tout ensemble d'ensembles est un ensemble. Enfin, si un système V n'a comme puissance aucun aleph « on reconnaît facilement qu'on peut projeter sur lui tout le système Ω, c'est-à-dire qu'il doit exister un système partiel V' de V qui est équivalent à Ω, V' est inconsistant » (1), donc V aussi ; donc tout système consistant a pour puissance un aleph. ZERMELO aperçoit dans le « on reconnaît facilement » un renvoi à l'intuition d'une bonne ordination transfinie : « visiblement CANTOR se représente les éléments de V successivement et arbitrairement coordonnés aux nombres de Ω, de sorte que tout élément de V n'est employé qu'une fois. Le procédé devrait ou bien se terminer une fois par l'épuisement de V, et V serait coordonné à un segment de Ω et sa puissance serait un aleph, contre l'hypothèse. Ou, au contraire, V reste inépuisable, donc contient une partie équivalente à Ω » (2). On verra pourtant que cette solution au problème de la

(1) *Ibid.*, p. 447.
(2) CANTOR (XXVIII), p. 451.

comparabilité et l'évasion corrélative de l'antinomie annoncent un travail analogue de VON NEUMANN (1). Mais il fallait pour l'entreprendre abandonner les considérations intuitives. D'où l'échec de la tentative de 1897 ; un mois après sa première lettre, CANTOR fait lui-même l'objection décisive : « on doit soulever la question d'où je puis bien savoir que les multiplicités bien ordonnées... auxquelles j'attribue les nombres cardinaux :

$$\aleph_0,\ \aleph_1\ \ldots\ \aleph_{\omega_0},\ldots\ \aleph_{\omega_1}\ldots$$

sont aussi réellement des... multiplicités consistantes. Ne serait-il pas concevable que déjà *ces* multiplicités soient inconsistantes et que la contradiction de l'hypothèse d'une « réunion » de tous leurs éléments ne se soit *seulement pas encore fait remarquer* ? Ma réponse est que cette question est également à étendre aux multiplicités finies... En d'autres termes : le fait de la « consistance » des multiplicités finies est une vérité simple indémontrable, c'est l'« axiome de l'arithmétique » (au sens ancien du mot). Et également la « consistance » des multiplicités auxquelles j'attribue les alephs comme nombres cardinaux est « l'axiome de l'arithmétique transfinie élargie » (2). »

A ces axiomes « au sens ancien » qu'aucune évidence ne garantissait plus, les successeurs de CANTOR s'efforceront, pour maintenir l'essentiel de la théorie abstraite, de substituer dans la mesure du possible la cohésion d'une axiomatique moderne. Comme point de départ pour leurs efforts se présentaient naturellement la construction systématique et les méthodes par lesquelles l'esprit algébriste de DEDEKIND avait ébauché un fragment de théorie des ensembles qui dût le moins possible à l'intuition.

(1) V. ci-après, p. 132.
(2) L. du 28 août 1899, *ibid.*, p. 447-448.

CHAPITRE III

DEDEKIND ET LA CHAINE LES AXIOMATISATIONS

« Il arrive fréquemment que dans un système donné Ω d'objets ou d'éléments, tout élément ω est remplacé, d'après une certaine loi, par un élément déterminé lui correspondant ; on a l'habitude d'appeler un tel acte une substitution... On s'exprime de façon encore plus commode si, comme nous le ferons, on considère cette substitution comme une application du système Ω et, de la sorte, appelle ω' l'image de ω, Ω' l'image de Ω. » « Sur cette capacité de l'esprit de comparer une chose ω avec une chose ω', ou de mettre en relation ω avec ω'... sans laquelle aucune pensée en général n'est possible, repose aussi, comme j'essaierai de le montrer ailleurs, toute la science des nombres (1). » Tel est le premier énoncé publié de l'idée centrale de *Was sind und was sollen die Zahlen* ? probablement aussi — quoique les trois premiers brouillons soient antérieurs (échelonnés entre 1872 et 1878) — l'origine de l'entreprise, sous l'influence des travaux arithmétiques de DIRICHLET. La différence avec le point de vue de CANTOR est assez nette : il ne s'agit pas pour résoudre des problèmes mathématiques de développer une nouvelle science, mais de fonder la mathématique déjà existante. C'est une œuvre d'analyse des notions, de critique logique, non de création : DEDEKIND s'insère de lui-même (2) dans la suite, inaugurée par SCHRÖ-

(1) DIRICHLET (V). Supplément XI (rédigé par DEDEKIND), p. 470 et n. (paru en 1879).

(2) « La parution de ses mémoires [SCHRÖDER, *Lehrbuch der Arithmetik*

DER, KRONECKER et HELMHOLTZ — que devaient continuer FREGE et PEANO —, des essais de réduction des mathématiques à la logique. « Les nombres sont de libres créations de l'esprit humain, ils servent comme moyen pour saisir la diversité des choses plus facilement et avec plus de précision. » Telle est la réponse à la question du titre. Fonder « cette partie de la logique qui traite des nombres... [en n'admettant] rien de démontrable sans démonstration », tel est l'objet de l'ouvrage. Toutefois il ne s'agit pas d'« un essai de systématisation scolaire » comme ZERMELO devait caractériser plus tard le *Formulaire* de PEANO (1), mais la griffe du mathématicien se montre dès la méthode choisie. Par « une décomposition raisonnée des suites de pensées sur lesquelles reposent les lois des nombres » (2), par la suppression de ces recours à l'« intuition intérieure » qui n'est qu'un raccourci, et surtout la systématisation de toutes les démarches élémentaires autour d'un acte unique, « relier des choses à d'autres », sera possible un travail de portée mathématique parce qu'on ne sort pas de l'activité mathématique. « Ἀεὶ ὁ ἄνθρωπος ἀριθμητίζει ... dès notre naissance... par cet exercice incessant déjà dans nos premières années..., nous acquérons un trésor de vérités proprement arithmétiques auxquelles plus tard nos premiers maîtres se réfèrent comme à quelque chose de simple, évident, donné dans l'intuition intérieure (3). » Ainsi cette expérience arithmétique, analysée conceptuellement, mais aussi « ramenée à une base unifiée » peut-elle servir de départ pour la création de notions nouvelles qui dépasseront l'arithmétique. Si « en opposition à CAN-

und der Algebra, KRONECKER *Über den Zahlbegriff*, HELMHOLTZ *Zählen und Messen* Leipzig, 1887] est l'occasion qui m'a déterminé à mettre en avant ma conception sous maints rapports semblable quoiqu'essentiellement différente pour le fondement. » DEDEKIND (II), t. III, p. 335.

(1) ZERMELO (II), p. 115.

(2) DEDEKIND (II), t. III, p. 337.

(3) *Ibid.*

TOR » [1] on peut montrer que « maints concepts en réalité très complexes (comme par exemple celui du nombre [ANZAHL] des objets) sont tenus a tort pour simples », leurs arrangements ou leurs simplifications ne permettront-ils pas des extensions imprévues ?

Telle est la singularité décisive dans l'histoire de la théorie des ensembles, que ses notions, quelques uns de ses résultats essentiels aient été retrouvés, comme involontairement, au détour d'une recherche sur le fondement de l'arithmétique. Si le transfini, parce qu'échappant à l'évidence intuitive, réclame consolidation et définition du domaine mathématique, systématiquement la simple investigation critique de celui-ci pour le fini seul amène au transfini. Mais dans un cas c'est l'infini concret d'êtres saisissables et s'engendrant au sein d'une activité qui ne veut pas de bornes, dans l'autre c'est l'indétermination de l'abstraction. Avec DEDEKIND il n'y a plus d'ensembles effectifs de la réalité, mais systèmes de relations entre ensembles possibles quelconques et méthodes pour les définir. La seule analyse procure cependant trois résultats principaux : une organisation de la théorie générale autour de la notion d'application, une définition de l'ordre et corrélativement le fondement des définitions par récurrence, enfin la théorie complète des ensembles finis.

La définition du matériel de la théorie générale est simple précision du bien commun de cette époque. « Une *chose* est tout objet de notre pensée. » Lorsque « différentes choses *a*, *b*, *c*.... conçues d'un même point de vue, sont réunies dans l'esprit, on dit qu'elles forment un *système S* ». Les choses *a*, *b*, *c*... sont les éléments du système qui « comme objet de notre pensée est également une chose ». Enoncé essentiel,

(1) Note marginale sur le premier brouillon de la préface (datée : 8 juillet 1887).

mais déjà fait par BOLZANO et CANTOR : un système S « est complètement déterminé si, pour toute chose, il est déterminé qu'elle soit ou non élément de S ». Donc deux systèmes qui ont mêmes éléments sont identiques ; on distingue évidemment entre un élément et le système formé de ce seul élément. Cependant, malgré les hésitations à ce sujet, dont les brouillons de 1872 à 1878 portent trace, le même signe dénote la relation ensemble-partie et ensemble-élément (1). Ou plutôt la seconde semble reléguée en dehors de la théorie ; d'où, à la fin surtout, quelques flottements. C'est qu'au départ au moins, comme les éléments abstraits ne sont pas distinguables, seules les relations entre ensembles sont objets d'étude : réunion, intersection et, avant tout, application. Pour celle-ci la différence avec l'attitude cantorienne apparaît nettement : il ne s'agit pas d'un simple élargissement de l'opération géométrique de la projection — mise en correspondance de deux réalités données — mais d'une généralisation du rapport fonctionnel. Deux conséquences : 1° l'application n'est pas forcément biunivoque (ce qu'elle était chez CANTOR puisqu'elle ne servait qu'à définir la puissance) ; 2° surtout la seule application considérée — étudiable *in abstracto* — sera l'application d'un système en lui-même par rapport à quoi il est une ***chaîne***. Quel autre mode de séparation, en effet, entre ensembles, indiscernables puisque abstraits, que la relation de partie à tout. A l'exploration des propriétés de la ***chaîne*** est consacré tout l'ouvrage.

En premier lieu, fondement sur elle du raisonnement d'induction complète. Le caractère de chaîne — relativement

(1) Nous emploierons les signes modernes : $m \in M$: m est élément de M ; $N \subset M$: N est *partie* de M ; $M + N$: ***réunion*** (système de tous les éléments contenus dans M et N) de M et N ; $M.N$ ***intersection*** de M et N (partie commune). En principe le système formé de l'élément a sera $\{a\}$ Pratiquement, quand il n'y aura pas d'ambiguïté, le système sera aussi désigné par a. Ainsi la ***chaîne*** du système $\{a\}$ sera désignée par a_0 au lieu de $\{a\}_0$.

à une application donnée φ — est invariant par rapport à cette application (l'image d'une chaîne est une chaîne), la réunion et l'intersection. D'où la possibilité de définir une sorte de système maximum ou d'enveloppe d'un sous-ensemble relativement à l'application : c'est la *chaîne propre* (1) M_o, intersection de toutes les chaînes dont M est partie. Il importe de noter le caractère transfini — c'est-à-dire abstrait — de la définition : la totalité des chaînes contenant M n'est donnée par aucun caractère, mais purement posée *in abstracto*. Le rôle maximal s'en déduit, qui apparaît mieux avec la propriété caractéristique de la *chaîne-image* : l'image de la chaîne propre d'un ensemble est la chaîne propre de son image (2) :

$$(M')_o = (M_o)' = M'_o$$

On en tire la décomposition (3) :

$$M_o = M + M'_o$$

M'_o représente donc le maximum de ce que l'application itérée φ permet d'atteindre à partir de M (il y a chaque fois débordement sur le domaine base puisque M n'est pas une chaîne). Le principe d'induction complète se formule alors : si une propriété est vraie pour M et si l'application φ la transmet, elle est aussi vraie pour la chaîne propre M_o de M par rapport à φ. La démonstration est immédiate à partir des définitions et propriétés précédentes (4).

(1) Pour abréger — sauf dans les cas prêtant à la confusion — nous disons chaîne de M.

(2) En effet $M \subset M_o$ donc $M' \subset (M_o)'$ or $(M_o)'$ est une chaîne donc $(M')_o \subset (M_o)'$; inversement $M' \subset (M')_o$ donc $M + (M')_o$ est une chaîne, donc $M_o \subset (M')_o + M$ donc :

$$(M_o)' \subset [(M')_o]' + M' \subset (M')_o.$$

(3) On a $M \subset M_o$ et puisque $M' \subset M_o$: $(M')_o \subset M_o$; avec l'inégalité de la note précédente on en tire l'égalité.

(4) En effet : qu'une propriété soit vraie de M s'écrit $M \subset \Sigma$ (Σ = système des objets possédant la propriété). Par hypothèse, si $G = M_o . \Sigma$, on a $G' \subset \Sigma$, donc $G' \subset G$, donc G est une chaîne, donc $M_o \subset G$ donc $M_o = G \subset \Sigma$.

Telle est la théorie minima, celle qui ne dépend que des notions corrélatives d'ensemble et d'application (en général, c'est-à-dire non biunivoque). Pour aller plus loin il est immédiat de faire intervenir l'application biunivoque et formuler une hypothèse sur le comportement à son égard de la relation partie-tout : ou bien un ensemble peut être appliqué de façon biunivoque sur une vraie partie de lui-même, ou la chose est impossible. Bifurcation entre ensembles finis et infinis qui s'imposait avec la notion d'application et à laquelle DEDEKIND prétend avoir été amené par ses propres réflexions : « plusieurs années avant la parution du mémoire de CANTOR et à une époque où le nom même de l'œuvre de BOLZANO m'était complètement inconnu » (1). Pour inaugurer la théorie des ensembles infinis il faut un théorème d'existence ; ici la rencontre avec BOLZANO se précise : l'exemple d'ensemble infini sera « le monde de mes pensées » dont chacune possède une image, la pensée de cette pensée, application évidemment biunivoque. BOLZANO avait, pour prouver qu'il y a, « dans le règne des choses qui sont simplement possibles..., incontestablement des ensembles infinis », invoqué « l'ensemble des propositions et vérités en soi » (2). Sa définition de l'infini est — on l'a vu — pourtant différente : pour toute vérité, nous en pouvons énoncer une autre, savoir que la première est vraie; ainsi avons-nous une loi permettant « de déduire de la proposition A celle-ci différente d'elle... B, et de B à nouveau C et ainsi sans fin ». Cette suite de propositions est donc semblable à la suite des nombres, c'est-à-dire « possède une multiplicité plus grande que tout nombre, donc est infinie ». Mais le principe est le même, expliquant non seulement la rencontre, mais le recours à l'univers psychologique, étrange dans un écrit comme celui de DEDEKIND :

(1) *Ibid.*, p. 342.
(2) BOLZANO (II), p. 14.

il s agit du pouvoir indéfini de thématisation décrit de façon analogue par les deux auteurs (faire d'une proposition — ou pensée — le sujet d'une autre proposition, ou pensée) et source véritable de l'infini mathématique. L'itération simple n'en est qu'un cas particulier, dans la numération, par exemple, où la collection finie des objets déjà numérotés est rassemblée dans un système, analogue à l'unité, à laquelle donc on peut adjoindre l'unité. Comme on voit, l'acte fécond n'est pas l'adjonction, mais l'acte de rassemblement préalable. Lorsque, en 1911, après huit ans d'hésitations, DEDEKIND se décide à rééditer son livre malgré « l'importance et la justesse partielle des doutes... exprimés sur la solidité de ses fondements », doutes qui portaient en premier lieu contre la démonstration de l'infini, il se borne en justification à affirmer sa confiance « qu'une investigation rigoureuse de la puissance créatrice de l'esprit capable de faire à partir d'éléments déterminés, un nouvel élément, leur système, qui est nécessairement différent de chacun de ces éléments, doit conduire à reconstituer de façon inattaquable les fondements de mon travail » (1). Former un système, par exemple, à partir d'un élément le système de cet élément, c'est bien thématiser, rendre une pensée sujet (logique) d'une autre pensée. Le théorème de l'infini est partie essentielle de l'œuvre, justification unique aussi d'un infini mathématique dans n'importe quelle théorie. Mais BOLZANO ne prouvait qu'un infini virtuel: ce n'est que par la notion d'un en soi des propositions qu'il pouvait confondre dans sa définition infini syncatégorématique — ou indéfini — avec infini catégorématique. Il est remarquable que la démonstration de DEDEKIND ne prouve pas davantage : pour atteindre l'infini par l'application il faudrait que le système des pensées de pensées soit vraie partie du système des pensées ; or c'est « mon propre moi » (2) qui est

(1) Préface de la 3e éd. DEDEKIND (V), p. 343.
(2) *Ibid.*, p. 357. Cf. L. inédite à KEFERSTEIN (8-2-1890). « Je tiens l a

exemple d'une pensée qui ne serait pas pensée de pensée. Seul moment extra-mathématique du raisonnement, mais décisif, chute dans ce que HILBERT appelait le transcendantal, ou réalisme psychologique. En fait, la reconnaissance du pouvoir de thématisation ne permet d'affirmer qu'un infini défini par l'ordre. La propriété prise comme départ par DEDEKIND étant plus forte, permet de déduire l'ordre.

On appelle système simplement infini un système N qui soit la chaîne propre d'un de ses éléments par rapport à une application biunivoque, φ, de N sur une de ses vraies parties. L'élément ainsi distingué — élément fondamental — n'appartient évidemment pas à l'image. Tout système infini contient un système simplement infini : il suffit de prendre la chaîne d'un des éléments extérieurs à l'image dans une des applications transformant le système en une de ses vraies parties. Un système simplement infini est ordonné par l'application qui le définit comme tel ; en effet, d'une part tout élément est distinct de son image, d'autre part étant donnés deux éléments quelconques m et n, on a entre leurs chaînes propres et leurs chaînes images une et une seule des 2 relations :

$$n_0 \subset m'_0 \qquad m_0 \subset n'_0$$

qui s'écriront

$$m < m \qquad n < m$$

On voit que les caractéristiques de l'ordre sont présentes : asymétrie (les 2 relations précédentes s'excluent) ; non réflexivité : il est impossible que l'on ait :

$$n_0 \subset n'_0 ;$$

enfin transitivité :

$$n_0 \subset m' \quad \text{et} \quad m_0 \subset p'_0 \quad \text{entraînent} \quad n_0 \subset p'_0$$

démonstration pour rigoureusement correcte : mon moi peut être objet de ma pensée... mais n'est pourtant pas lui-même une pensée. »

On voit immédiatement que tout élément précède, dans l'ordre, son image. Cet ordre est un bon ordre : en effet toute chaîne du système est la chaîne propre d'un élément ; considérons la chaîne T_0 d'une partie quelconque T du système : T_0 étant la chaîne de l'élément k on a :

$$T_0 = k + k'_0$$

T contient k (sans quoi la chaîne de T serait k'_0) et tout élément de T distinct de k est contenu dans K'_0, donc le suit dans l'ordre ; toute partie T a un premier élément. L'élément fondamental du système tout entier précède tous les autres : on le désigne par 1 ; les éléments du système sont les nombres ordinaux naturels ; le système entier, la suite N de ces nombres. La notion de segment Z_n, ensemble des éléments $\leqq n$, n étant un élément donné, permet de retrouver les propriétés classiques (tout nombre a, dans l'ordre, un prédécesseur et un suivant immédiat, toute partie d'un segment possède un plus grand nombre ; toute partie du système qui ne possède pas de plus grand nombre est simplement infinie). Le raisonnement employé en général pour les démonstrations est l'induction complète qui revêt ici son aspect habituel. Pour qu'une propriété soit vraie de tous les éléments de la chaîne $1_0 = N$, il suffit de la démontrer pour 1, et de prouver que l'ordre (l'application φ) la transmet d'un élément à son successeur immédiat (son image).

La conséquence immédiate et la plus importante de cette théorie de l'ordre est la justification des définitions par récurrence, apport original de DEDEKIND. Etant donné un système quelconque Ω appliqué sur lui-même par l'application θ (biunivoque ou non), une définition par récurrence consiste à établir une correspondance entre N et Ω telle que :

1° l'image de 1 soit un élément donné ω de Ω ;

2° l'image de tout nombre n' soit l'image par rapport à θ de l'image dans Ω de son prédécesseur n. Il y a une sorte de transport de la mise en ordre φ sur l'application θ. De telles définitions sont justifiées si l'on démontre qu'étant donnés ω et θ il existe une et une seule application Ψ satisfaisant aux 3 conditions :

$$1.\ \Psi\ (N) \subset \Omega$$
$$2.\ \Psi\ (1) = \omega$$
$$3.\ \Psi\ (n') = \theta\ \bar{\Psi}\ (n)$$

La démonstration s'effectue grâce au *procédé de fusionnement*. On prouve d'abord que pour tout nombre n il existe une et une seule fonction Ψ_n satisfaisant aux 3 conditions :

$$1'.\ \Psi_n\ (Z_n) \subset \Omega$$
$$2'.\ \Psi_n\ (1) = \omega$$
$$3'.\ \Psi_n\ (m') = \theta\ \Psi_n\ (m) \qquad \text{pour tout } m < n.$$

Par induction complète : la fonction Ψ_1 existe évidemment (il suffit de la définir par la propriété 2' : 1' est alors satisfaite, 3' disparaît). Si la fonction Ψ_n existe et est déterminée de façon univoque, on construit $\Psi_{n'}$, en posant :

$$\Psi_{n'}\ (m) = \Psi_n\ (m) \quad \text{pour tout } m \leqq n$$
et
$$\Psi_{n'}\ (n') = \theta\ \Psi_n\ (n)$$

$\Psi_{n'}$ est complètement définie et satisfait aux conditions 1' — 3'.

Enfin on obtient Ψ en écrivant :

$$\Psi\ (n) = \Psi_n\ (n)$$

Définition évidemment univoque et garantissant l'existence (puisque pour tout n il existe un et un seul Ψ_n).

Ainsi s'introduisent les opérations élémentaires si on prend pour système Ω le système N lui-même ; pour l'addition de $+\ m$: $n + m$ (n étant un nombre choisi) on a :

$$\omega = m \qquad \theta\ (n) = n'$$

Pour la multiplication par m, on pose :

$$\omega = m \qquad \theta\,(n) = n + m$$

Enfin, l'exponentiation du nombre a :

$$\omega = a \qquad \theta\,(n) = n \times a$$

Les propriétés classiques sont immédiatement retrouvées par induction complète. Il importe de noter, comme fait DEDEKIND, le caractère spécial des définitions par récurrence : elles ne sont fondées que par rapport au système simplement infini N (si on suppose sa non contradiction comme système ordonné $1_o = N$) ; tandis que l'induction complète est utilisable, comme on l'a vu, pour toute chaîne relative à une application quelconque (biunivoque ou non). Si au lieu de N on prend une telle chaîne, le procédé de définition par récurrence aboutit, en général, à des contradictions (1).

La théorie précédente a son utilisation immédiate dans la théorie des groupes : c'est grâce à elle que prend un sens la notion de $n^{ième}$ itération d'une opération. En prenant un système Ω d'objets quelconques où la liaison de deux éléments ω et ν est déterminée (représentée par le symbole $\omega\nu$), on peut considérer que tout élément choisi au hasard, ω, détermine une application de Ω en lui-même (à tout ν correspond $\omega\nu$). L'application n fois itérée de Ω par ω sera définie par les conditions :

$$\text{II.}\quad \omega' = \omega$$
$$\text{III.}\quad \omega^{n'} = \omega\,\omega^{n}$$

(1) DEDEKIND prend l'exemple du système S composé de 2 éléments a et b, appliqué sur lui-même : $a' = b$ $b' = a$; donc $a_o = b_o = S$. Si on choisit pour Ω le système formé des 3 éléments α, β, γ avec l'application θ telle que :

$$\theta(\alpha) = \beta,\ \theta(\beta) = \gamma,\ \theta(\gamma) = \alpha$$

on voit qu'il est impossible de satisfaire aux conditions 1-3 : si $\Psi\,(a) = \alpha$, on aura :

$$\Psi\,(b) = \beta \text{ et } \Psi\,(b' = a) = \gamma \neq \alpha.$$

elle résulte donc d'une application (suivant le procédé de récurrence) de la suite des nombres N sur Ω. Ainsi peut s'interpréter la notion de chaîne propre : si l'on considère le système Ω de toutes les applications d'un système S en lui-même, les conditions II et III définissent bien toutes les itérations d'une certaine application ω ; elles ont, en effet, la forme des conditions 2 et 3 d'une définition par récurrence si on pose $\theta = \omega$; or ceci est possible en vertu des propriétés élémentaires de la chaîne : si

$$\omega\ (S) \subset S \qquad \text{(vrai par hypothèse)}$$

on voit que :

$$\omega\ [\omega\ (S)] \subset \omega\ (S)$$

Donc la liaison $\omega\ \omega$ fait passer d'un élément de Ω à un autre élément de Ω (application de S en lui-même), $\omega\ (\nu)$ (ν étant un élément de Ω) est donc une application de Ω en lui-même, définissant θ. Considérons maintenant la chaîne propre $\omega_0\ (A)$ d'une partie quelconque A de S ; on voit que quel que soit n :

$$\omega^n\ (A) \subset \omega_0\ (A)$$

Si l'on pose :

$$B = \sum_n \omega^n\ (A) + A$$

d'une part, donc :

$$B \subset \omega_0\ (A)$$

d'autre part, B est une chaîne par rapport à ω ; en effet :

$$\omega\ (B) = \omega\ (A) + \sum_n \omega^n\ (A) \subset \sum_n \omega^n\ (A)$$

donc, en vertu de la définition de la chaîne propre :

$$\omega_0\ (A) \subset B \qquad \omega_0\ (A) = B$$

Ainsi l'utilisation de la chaîne propre coïncide avec le procédé cantorien du passage à la limite (1) : prendre l'intersec-

(1) D'où la coïncidence pour le fond entre la démonstration BERNSTEIN SCHRÖDER du théorème de l'équivalence et celle que Zermelo, et, sans qu'il le sache, avant lui DEDEKIND (V. CANTOR-DEDEKIND (I), p. 8 ; CANTOR-

tion de toutes les chaînes revient à considérer ce qui résulte de toutes les itérations d'une même application. Le même recours au transfini est impliqué. Il importe de remarquer que si la présentation de DEDEKIND est plus simple — puisque partant de principes préalables à la définition de la suite des entiers (et procurant celle-ci), — elle ne permet le passage aux autres présentations que si on considère comme démontré le théorème d'existence d'un ensemble infini (*sur quoi repose la non contradiction de la définition d'un ensemble simplement infini*). La notion de chaîne propre serait donc plus générale (puisque indépendante de ce théorème d'existence). Il est vrai que la non contradiction de sa définition devait elle-même apparaître plus tard problématique.

(XXVIII), p. 449, avec les remarques de Zermelo p. 450, enfin DEDEKIND (V), t. III, p. 447) obtiennent par la notion de chaîne. La démonstration DEDEKIND s'énonce ainsi, avec les notations déjà employées : si $M \sim P$ avec :

$$P \subset T = P + Q \subset M = P + Q + R$$

on doit avoir $M \sim T$.

Par rapport à l'application φ de M sur P, soit R_0 la chaîne de R, posons

$$M = R_0 + V ;$$

on a, puisque $P = M'$:

$$P = R'_0 + V'$$

il en résulte :

$V \subset T$ (puisque $R \subset R_0$ et $R_0 + V = T + R$)

et $R'_0 + V \subset T$ (car $R'_0 \subset T$, puisque $P = R'_0 + V' \subset T$) ;

comme d'autre part :

$$T \subset R_0 + V = M = R + R'_0 + V,$$

et que T est par hypothèse extérieur à R, on a :

$$T \subset R'_0 + V \quad \text{donc} \quad T = R'_0 + V$$

Donc, puisque :

$$R'_0 \sim R_0 \text{ (ou } V' \sim V) : T \sim M \sim P.$$

Si l'on pose maintenant :

$$R_0 = R + R' + R'' + \ldots \quad (R'' = \varphi\,\varphi\, R, \text{ etc...})$$

on a :

$$M = V + R + R' + R'' + \ldots$$
$$P = V' + R' + R'' + \ldots \subset T$$

donc, puisque R est extérieur à T, tous les R', R''... sont extérieurs à R ; on peut écrire $M - R = T = V + R' + R'' + \ldots$ ce qui est évidemment équivalent à P et à M.

Le même procédé du *fusionnement* permet de démontrer la proposition capitale de la théorie des ensembles finis : la condition nécessaire et suffisante pour qu'un ensemble S ne soit pas fini est que tout segment Z_n de la suite des nombres soit applicable sur lui de façon biunivoque. La condition est évidemment nécessaire, puisque tout ensemble infini contient un ensemble simplement infini (donc applicable de façon biunivoque sur la suite N). La démonstration qu'elle est suffisante se fait par la construction d'une suite d'applications Ψ_n biunivoques définies respectivement pour chaque Z_n et telles que chacune d'elles soit contenue dans (coïncide avec) toutes les applications de numéro d'ordre plus élevé. Autrement dit pour $m < n$ et pour tout $t \leqq m$ $(t \in Z_m)$ on a :

$$\Psi_n(t) = \Psi_m(t)$$

La définition des Ψ_n est faite par récurrence (1) grâce aux applications α_n données des Z_n. On en déduit une application biunivoque de toute la suite N en posant :

$$\chi(n) = \Psi_n(n)$$

on a bien :

$$\chi(N) \subset S$$

donc S est infini. L'exemple montre la symétrie entre le procédé de fusionnement et le procédé de la diagonale pour les systèmes ordonnés : dans les deux cas définition d'une nou-

(1) Soit Ω le système de toutes les applications biunivoques de tous les systèmes Z_n sur S : à toute application β d'un système Z_n on fait correspondre une application γ de $Z_{n'}$ en posant $\gamma(p) = \beta(p)$ pour tout $p \in Z_n$ et $\gamma(n^*) = \alpha_{n'}(k)$, $\alpha_{n'}$ étant l'application, donnée dans l'hypothèse, de $Z_{n'}$ sur S, k étant le plus petit nombre pour lequel l'image donnée par $\alpha_{n'}$ n'appartient pas à $\beta(Z_n)$ (il en existe évidemment puisque ces applications sont biunivoques, et que Z_n et $Z_{n'}$ ne peuvent être mis en correspondance biunivoque). Cette correspondance bien définie entre β et γ applique donc Ω sur lui-même (on voit que γ est biunivoque), c'est l'application θ des définitions par récurrence ; l'élément ω sera α_1 : chacune des valeurs de la fonction récurrente ainsi définie est un ψ_n.

velle fonction d'après l'ensemble des valeurs des fonctions d'une suite pour l'argument égal à leur numéro d'ordre, ici unissant toutes les fonctions, là distincte d'elles toutes. Du théorème découle que tout ensemble fini T est applicable de façon biunivoque sur un et un seul Z_n : le nombre n déterminant le segment sera le *nombre cardinal* de T.

Ainsi, d'une part la théorie cardinale se trouve dans la dépendance de la théorie ordinale, d'autre part celle-ci n'est établie que grâce à la considération de l'ensemble simplement infini : elle fonde le fini sur l'infini. Sans doute est-il possible, à partir de la théorie générale de l'application, de déduire déjà deux propriétés : toute partie d'un ensemble fini est finie, l'ensemble formé par adjonction d'un élément à un ensemble fini est fini ; en revanche, les propriétés caractéristiques comme : l'ensemble des parties d'un ensemble fini est fini, exigent le recours à l'ordre. Caractère boiteux de cette théorie (tolérable parce qu'elle n'est qu'un accessoire dans l'ouvrage dont le but est l'étude de la suite indéfinie des nombres) que DEDEKIND a lui-même senti : une autre définition du fini, basée également sur l'application et la chaîne, permet de déduire l'ordre sans passer par l'infini.

C'est l'ordre « cyclique » : un ensemble est fini lorsqu'il existe une application φ de l'ensemble sur lui-même telle qu'aucune des vraies parties de l'ensemble ne soit appliquée sur elle-même (1). La notion de chaîne se transpose en *quasi chaîne* : toute partie H_b de T telle que l'image de tout élément de H_b soit contenue dans H_b sauf de celle d'un de ses éléments b ; à la chaîne propre correspond l'*arc* $a\ b$: in-

(1) Préface de la 2e édition DEDEKIND (II), t. III, p. 342. Dans un fragment resté inédit de son vivant (DEDEKIND (V), t. III, p. 450-458), DEDEKIND avait commencé à développer les conséquences de cette définition, en particulier donné la notion d'arc et défini l'ordre : on peut en déduire les propriétés ci-dessus (J. CAVAILLÈS (I), p. 143-147) ; la démonstration que toute partie d'un ensemble fini est finie s'effectue en particulier sans recours à l'induction complète.

tersection de toutes les quasi chaînes H_b contenant l'élément a. On voit que $a\ b$ est également une quasi-chaîne et, grâce à cette notion, il est facile de démontrer que l'application φ est biunivoque, tout élément distinct de son image, enfin que — exactement comme la chaîne propre — un arc se décompose :

$$a\ b = a + a'b$$

Etant donnés deux éléments a et b, on a une et une seule des relations :

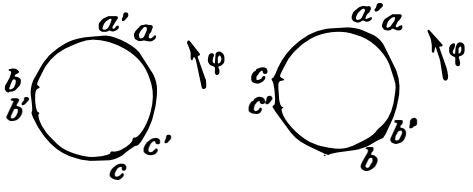
$$a \in a'b' \qquad \text{ou} \qquad b \in a'b'$$

qui permet de définir l'ordre (1) :

$$a > b \qquad \text{ou} \qquad b > a$$

avec les propriétés habituelles. De là se déduisent le principe d'induction complète, les propriétés que toute partie est finie, que l'ensemble de toutes les parties est fini et qu'il e. t impossible d'appliquer de façon biunivoque un ensemble fini sur une de ses vraies parties. La réciproque de la dernière proposition, qui établirait l'équivalence des deux définitions, ne semble pas pouvoir être démontrée sans que l'on fasse intervenir, sous une forme ou une autre, le transfini.

Tel est le phénomène significatif auquel aboutit la théorie des ensembles finis : toutes les autres définitions adop-



(1) Nous simplifions; en réalité on définit du même coup l'ordre entre 3 éléments, a, b, c : on a :

ou bien $b'c = b'a + a'c$ et $c'b = c'a.a'b$
ou bien $b'c = b'a.a'c$ et $c'b = c'a + a'b$

tées ultérieurement, définition ZERMELO (1) analogue à la 2e définition de DEDEKIND, définition RUSSELL-WHITEHEAD (2) par le principe ordinaire d'induction complète (un ensemble A est fini s'il appartient comme élément à toute classe K contenant l'ensemble vide et telle que si $B \in K$ et $a \in A$, $B + a \in K$), définitions TARSKI (3) et JANISZEWSKI par les éléments *saturé* et *irréductible* (un ensemble A est fini si tout ensemble M non vide de ses parties contient un élément saturé, c'est-à-dire qui ne soit partie d'aucun autre élément de M, ou irréductible, c'est-à-dire dont aucune partie ne soit élément de M) sont équivalentes entre elles, ainsi que l'a montré TARSKI (4). Or le moyen employé uniformément pour démontrer les équivalences est l'induction complète : il semble donc que, sous-jacent à chacune de ces définitions, se trouve un renvoi à la définition arithmétique ordinaire, la construction progressive de l'ordre, intuition irremplaçable, comme le voulait Poincaré. Seule la propriété prise par DEDEKIND comme 1re définition est pure de toute notion d'ordre : elle est démontrable à partir des autres définitions, mais ne permet pas de démontrer les propriétés énoncées dans celles-ci, à moins que l'on n'ait recours à un instrument sur lequel concentraient leurs attaques au début du siècle la plupart des mathématiciens hostiles à la théorie des ensembles, l'axiome de choix, principe qui se révélait fondamental pour l'étude abstraite du transfini.

On en trouve, note FRAENKEL (5), la première formule dans un écrit de PEANO de 1890 sur les équations différen-

(1) ZERMELO (IV).

(2) RUSSEL-WHITEHEAD (I).

(3) TARSKI (I) où se trouvent passées en revue et comparées les différentes définitions.

(4) *Ibid.*

(5) FRAENKEL (XI), p. 33.

tielles : « mais comme on ne peut pas appliquer une infinité de fois une loi arbitraire par laquelle à une classe *a* on fait correspondre un individu de cette classe, on a formé ici une loi déterminée par laquelle à chaque classe *a*, sous des hypothèses convenables, on fait correspondre un individu de cette classe ». En 1902, BEPPO LEVI l'énonce comme principe séparé à l'occasion d'une démonstration de BERNSTEIN (1). En 1904, sur la suggestion d'Erhard SCHMIDT, ZERMELO le prend comme point de départ — et pierre angulaire — pour sa démonstration que tout ensemble peut être bien ordonné. Aussitôt apparaissent les objections, à la tête desquelles, par une curieuse rencontre, celle de PEANO. La critique de BOREL est particulièrement claire. Si la démonstration ZERMELO prouve l'équivalence entre les deux problèmes : mettre un ensemble *M* sous forme bien ordonnée, et faire correspondre à toute partie *M'* de *M* un élément distingue, on ne peut considérer qu'elle résolve le premier puisqu'elle laisse le second intact. « Un tel raisonnement ne me paraît pas mieux fondé que le suivant : pour bien ordonner un ensemble *M* il suffit d'y choisir arbitrairement un élément auquel on attribuera le rang 1, puis un autre auquel on attribuera le rang 2 et ainsi de suite *transfinimment*... De tels raisonnements sont en dehors des mathématiques (2). » Ils reviennent, en fait, à supposer le bon ordre déjà donné dans l'ensemble considéré : il n'y a pas raisonnement mais cercle vicieux.

La réponse de ZERMELO est triple : 1° le problème spécial posé n'était pas : donner un moyen de bien ordonner tout ensemble, mais démontrer que tout ensemble peut être mis sous forme bien ordonnée, problème spécifique du cantorisme qui, dans la théorie cardinale, part d'en-

(1) ZERMELO (II), p. 113 n.
(2) BOREL (II), p. 195.

sembles donnés concrètement comme le continu et veut effectuer la synthèse avec le développement conceptuel de la théorie ordinale.

2° Une solution effective est donnée à ce problème grâce à une utilisation de la théorie abstraite de DEDEKIND qui remplace justement un procédé de *construction* transfinie par une détermination *abstraite* de l'ordination. Il y a gain puisque dans le 1er cas on ne peut assurer que l'ensemble *tout entier* est épuisé par la mise en ordre progressive, ce qui est évidemment le seul point en question. La démonstration elle-même se borne à utiliser le procédé de *fusionnement* pour les parties bien ordonnées de l'ensemble M : si l'on suppose coordonné à toute partie de M un *élément distingué*, il suffit d'exiger des parties bien ordonnées M_γ que chacun de leurs éléments soit élément distingué du complémentaire $M-A$ du segment A qu'il y détermine. Il existe de telles parties : l'ensemble $\{m\}$ formé de l'élément distingué de M. Etant données deux telles parties, l'une est identique à un segment de l'autre (avec même bon ordre) ; la réunion de toutes ces parties constitue un ensemble bien ordonné auquel appartient tout élément de M : conséquences immédiates de la notion de bon ordre. Le seul reproche que l'on pourrait faire au raisonnement est qu'il suppose établie déjà la théorie du bon ordre, c'est-à-dire non contradictoire la notion d'ensemble bien ordonné transfini : l'exemple prouvant qu'il existe des parties M_γ est un ensemble bien ordonné fini. Dans cette mesure l'objection de BOREL est justifiée : ce qu'il faut montrer — et qui n'apparaît pas ici — c'est que la théorie du bon ordre peut être fondée avec les seules ressources de la théorie abstraite des ensembles (et l'application).

D'où : 3° la dernière réponse de ZERMELO (1), une nou-

n),

ZERMELO (II), p. 107.

velle démonstration qui ne suppose préalablement fondées que les notions d'ensemble (avec les opérations réunion intersection), de partie d'ensemble et d'application, filiation cette fois directe à partir de DEDEKIND. « Un ensemble M est dit bien ordonné si à chacun de ses éléments a correspond de façon univoque une partie $R(a)$ de M, son *reste*, et si toute partie non vide P de M contient un et un seul élément p dont le reste $R(p)$ contient P », p est dit le *premier* élément de P. On voit qu'il y a ainsi un ordre entre les éléments de M puisque pour tout couple $\{m, n\}$ l'un des éléments est le *premier* des deux : on écrit $m < n$, relation équivalente (1) avec $n \in R(m)$, qui exclut $m \in R(n)$; donc relation transitive : si $m < n < p$, on a forcément (1)

$$\{m, n, p\} \subset R(m) \text{ donc } p \in R(m) : m < p ;$$

enfin dans une partie quelconque le premier élément précède bien tous les autres : la définition habituelle est retrouvée. Le théorème de ZERMELO s'énonce alors : tout choix dans M (c'est-à-dire toute coordination à chaque partie A de M d'un élément *distingué* a de A) entraîne une bonne ordination de M, savoir celle où le reste $R(p)$ d'un élément p a pour élément distingué p. En effet, considérons une application φ de l'ensemble $\mathfrak{M}$ des parties de M en lui-même, faisant correspondre à toute partie P la partie $P' = P - \{p\}$ obtenue par retranchement de l'élément distingué de P, on appellera θ-*chaîne* par rapport à φ, toute chaîne telle que l'intersec-

(1) En effet : 1° $m < n$, d'après sa définition, entraîne $\{m, n\} \subset R(m)$ donc $n \in R(m)$; 2° si $n \in R(m)$ on a $\{m, n\} \subset R(m)$ et m est le *premier* des deux éléments.

2° Par définition $\{m, n, p\}$ est partie d'un des trois restes $R(m)$, $R(n)$, $R(p)$; or les deux autres inclusions: $\{m, n, p\} \subset R(n)$ et $\{m, n, p\} \subset R(p)$ sont impossibles par hypothèse (puisque $m < n$ on ne peut avoir $m \in R(n)$, etc...)

tion d'un système quelconque de ses éléments lui appartienne. Soit M_0 la θ-*chaîne propre* de $\{M\}$, on voit que pour tout couple d'éléments N et P de M_0 on a ou $N \subset P$ ou $P \subset N$ (en effet le système des éléments jouissant de cette propriété dans M_0 constitue une θ-chaîne propre contenant M). Etant donnée une partie quelconque Q de M, appelons *enveloppe* Q_1 de Q l'intersection de tous les éléments de M_0 qui contiennent Q (il en existe au moins un, M), Q_1 est le reste $R(q)$ de l'élément distingué de Q. En effet, l'élément distingué de Q, est nécessairement dans Q (sans quoi $Q'_1 \subset Q_1$ élément de M_0, contiendrait aussi Q) et Q_1 est le seul élément de M_0 à jouir de la propriété ; soit d'autre part Z l'enveloppe de $\{q\}$, on a évidemment $Z \subset Q_1$ (puisque Z est élément de M_0), considérons $Q' = Q - \{q\}$, l'enveloppe de Q' est l'image Q'_1, de Q_1 dans l'application φ, comme Q'_1 et Z sont éléments de M_0 on a $Q'_1 \subset Z \subset Q_1$ et comme Q'_1 et Q_1 ne diffèrent que par q, contenu dans Z, $Z = Q_1$. La partie Q_1 de M satisfait donc aux conditions exigées du reste $R(q)$ dans la définition du bon ordre. C'est d'ailleurs la seule : tout système de parties de M possédant les mêmes propriétés contiendrait M et formerait une θ-chaîne par rapport à φ, donc serait contenu dans M.

L'élargissement par rapport aux notions de *Was sind und was sollen die Zahlen* ? est marqué dans le passage de la chaîne à la θ-chaîne : l'application seule ne permet pas d'atteindre dans l'ordre plus loin que ω. La chaîne propre fait la somme d'une suite infinie d'opérations : elle ne fonde que l'induction complète ordinaire. Pour avoir l'induction complète transfinie — équivalente au bon ordre — il faut introduire une nouvelle opération, ici l'intersection d'un système d'éléments de M_0 : grâce à elle est garanti que toute partie P de M possède bien une enveloppe dans M_0. Ainsi se retrouve donc l'indépendance relative des deux principes générateurs des nombres ordinaux chez CANTOR : à l'adjonction d'une unité correspond l'appli-

cation, au passage à la limite, après lequel on peut à nouveau ajouter une unité, l'intersection. Mais si le raisonnement de la chaîne se trouve ainsi compliqué, aucune notion nouvelle n'est introduite, sauf peut-être celle d'ensemble des parties d'un ensemble, que DEDEKIND n'avait pas expressément isolée. La théorie abstraite permet donc d'unifier dans une organisation systématique toute l'œuvre cantorienne ; c'est à quoi s'emploie ZERMELO en 1907.

* * *

Le travail est imposé par une double urgence : l'insertion du théorème du bon ordre dans la théorie classique, la neutralisation des antinomies récemment découvertes. L'axiome de choix est indispensable dans la théorie classique : ZERMELO montre (1) notamment que sans lui on ne sait ni fonder la théorie des ensembles finis sur la 1re définition de DEDEKIND (démontrer le théorème que tout ensemble fini peut toujours être ordonné de telle sorte que chacune de ses parties possède un premier et un dernier élément) (2), ni prouver les propositions essentielles suivantes : l'ensemble réunion d'un système S d'ensembles disjoints a une puissance supérieure ou égale à celle de S ; les sommes d'ensembles disjoints équivalents sont équivalentes ; une somme dénombrable d'ensembles finis ou dénombrables est dénombrable (3) ; enfin le produit de

(1) ZERMELO (II), p. 113-114.

(2) En effet, ce théorème n'est autre que le théorème DEDEKIND vu plus haut : tout ensemble fini peut être appliqué de façon biunivoque sur un et un seul segment Z_n, conséquence immédiate du théorème que si on peut appliquer de façon biunivoque tout segment Z_n sur un ensemble S, S est infini ; on a vu que dans la démonstration on prenait comme base une application α_n pour chaque Z_n, donc un choix dans l'ensemble des applications de Z_n sur S (signalé pour la 1re fois par HESSENBERG (I)).

(3) Il faut également supposer une application simultanée de tous les éléments de la somme respectivement les uns sur les autres (donc le choix d'une application pour chaque élément).

puissances ne s'annule que si un des facteurs est nul : énoncé équivalent à l'axiome de choix, comme on voit immédiatement (1). Ainsi avec l'axiome de choix (et le bon ordre) il faut savoir que la théorie à peu près tout entière subsiste ou s'écroule. Mais l'écroulement menaçait déjà par ailleurs.

L'antinomie BURALI-FORTI gênait le libre essor de la théorie ordinale cantorienne. En 1903, paraît dans la théorie cardinale l'antinomie RUSSELL (2), à laquelle ZERMELO s'était déja heurté (3). Ici encore c'est un procédé essentiel de la théorie qui, défini abstraitement, procure une contradiction : pour l'antinomie BURALI-FORTI le 2e principe générateur des ordinaux, pour l'antinomie RUSSELL, le procédé de la diagonale. A toute application d'un ensemble M sur l'ensemble $\mathfrak{P}M$ de ses parties correspond un élément de $\mathfrak{P}M$ extérieur à l'application : c'est l'ensemble des éléments de M qui n'appartiennent pas à la partie à laquelle ils sont coordonnés. Si M est l'ensemble de tous les ensembles, on devrait pourtant avoir $\mathfrak{P}M \subseteq M$: contradiction. La situation est la même dans les deux cas : une théorie s'est constituée pour répondre à des problèmes effectivement posés par la mathématique existante ; les notions, les procédés nouveaux forgés permettent d'obtenir des résultats indépendants de la structure particulière des objets mathématiques dont on partait. Une théorie abstraite s'impose, dont le seul élément est l'ensemble, « réunion en un tout d'objets déterminés et distincts de notre

(1) Un produit d'ensembles M, N, P... a pour éléments les systèmes $\{m, n, p...\}$ composés d'éléments appartenant respectivement à M, N, P... : chaque système représente donc un choix.

(2) RUSSELL-WHITEHEAD (I), p. 363-368.

(3) « J'avais trouvé moi-même cette antinomie indépendamment de RUSSELL et l'avais déjà communiquée... au professeur HILBERT avant 1903. » ZERMELO (II), p. 119 n.

intuition ou de notre pensée » (1). Dès que l'on dit « tous les objets qui... », on a un ensemble. Mais l'ensemble de tous les ordinaux est contradictoire ; mais l'ensemble de tous les ensembles est contradictoire. La faute n'en est pas aux procédés, simple pouvoir exécutif des notions, ce sont les notions bases mêmes qui disparaissent. Qu'est-ce maintenant qu'un ensemble ? et, si l'on change la définition, que reste-t-il de la théorie naïve ? On voit la chute après les espoirs que celle-ci avait fait naître, la gravité pour le système entier des mathématiques, l'impossibilité d'un recours étranger, puisqu'elle est préalable à tout : « La théorie des ensembles, commence ZERMELO, est la branche des mathématiques à laquelle il revient d'étudier mathématiquement les concepts fondamentaux de nombre, d'ordre et de fonction dans leur simplicité primitive, et par là, de développer les bases logiques de l'arithmétique tout entière et de l'analyse (2). »

La seule solution doit être cherchée dans un examen critique : au lieu de poser une nouvelle définition *a priori*, on obtiendra une définition implicite en scrutant raisonnements et résultats qui tiennent : « partant de la « théorie des ensembles » historiquement existante, il faut chercher les principes qui sont nécessaires pour la fondation de cette discipline mathématique.. J'ai l'intention de montrer comment toute la théorie créée par CANTOR et DEDEKIND se laisse ramener à quelques définitions et à 7 axiomes qui semblent indépendants les uns des autres » (3). Travail à la fois d'explicitation (4) et de limitation : il s'agit de « res-

(1) Définition de CANTOR (XXVII), p. 282.
(2) ZERMELO (III), p. 261.
(3) *Ibid.*
(4) ZERMELO l'avait déjà ébauché en tête de sa seconde démonstration du bon ordre, pour prévenir les objections ; outre l'axiome de choix se trouvent énoncés deux postulats qui garantissent l'existence des parties d'ensembles, de l'ensemble des parties d'un ensemble, du complémentaire et de l'intersection d'un système d'ensembles (ZERMELO (II), p. 107).

treindre assez les principes pour exclure toutes les contradictions et ... garder pourtant ce qui a une valeur ». La méthode est celle dont HILBERT avait donné le modèle pour la géométrie : entre les objets indéterminés qui constituent le champ de la théorie (corrélatif des axiomes et non à son tour objet de la théorie, c'est-à-dire non ensemble) l'axiomatique définit par ses propriétés une relation fondamentale ϵ : élément-ensemble (1). Les seconds termes de cette relation seront les ensembles : ou plutôt la théorie ne s'occupe que de déduire logiquement de nouvelles propriétés à partir des propriétés initiales de la relation, sans s'inquiéter de la réalité ou de la structure des termes. En fait un seul axiome, l'axiome I — de *détermination* — définit une relation : la relation d'égalité suite des relations ϵ et $\subset$ ($M = N$ si $M \subset N$ et $N \subset M$, ou si pour tout $x \epsilon M$, $x \epsilon N$ et réciproquement) tous les autres axiomes sont des affirmations d'existence. La limitation est donc obtenue de façon réaliste : sont ensembles les seuls objets posés comme tels par les axiomes.

Deux sortes d'affirmations : 1° catégoriques : l'axiome II — des *ensembles élémentaires* — pose comme ensembles l'ensemble vide (simple moyen de supprimer des restrictions dans les raisonnements, mais non indispensable), les ensembles constitués de un et de deux éléments (2) ; l'axiome VII — *de l'infini* — : il existe « un ensemble Z

(1) La relation $M \subset N$ (M *partie* de N) est une abréviation pour : de $x \epsilon M$ suit toujours $x \epsilon N$.

(2) Il s'agit d'affirmation catégorique relativement à l'existence d'ensembles : a étant une chose quelconque du domaine — pas forcément un ensemble — dont l'existence en tant qu'ensemble n'a donc pas à être supposée, $\{ a \}$ est un ensemble ; pour $\{ a, b \}$, il y a cette fois l'hypothèse que a et b sont distincts, c'est-à-dire que la question $a \epsilon \{ b \}$ peut être décidée au moyen des axiomes. Seule l'existence de l'ensemble vide est une affirmation absolument catégorique (v. ci-après son importance dans l'axiomatique de FRAENKEL).

qui contient comme éléments l'ensemble vide et ... avec tout élément a l'élément de forme $|a|$ [1] » ;

2° Hypothétiques : si un ensemble existe, existent également d'autres ensembles définissables à partir du premier. Axiome III — de *séparation* — : « si une proposition $\mathfrak{E}(x)$ est *définie* pour tous les éléments d'un ensemble M, M possède toujours une partie $M\mathfrak{E}$ constituée par tous les éléments x de M pour lesquels $\mathfrak{E}(x)$ est vrai [2] ». Une proposition est *définie* si « sa validité... peut être décidée sans arbitraire au moyen des axiomes et des lois logiques universelles ». De l'axiome III résultent l'existence de parties [3] d'un ensemble M, de son complémentaire par rapport à un ensemble N qui le contient, de l'intersection d'un système d'ensembles. — L'axiome IV et l'axiome V posent respectivement l'existence de l'*ensemble des parties* $\mathfrak{P}M$ d'un ensemble donné et de l'*ensemble réunion* $\mathfrak{S}M$ formé par les éléments des éléments de M. Enfin l'axiome de *choix*, VI, garantit que le produit n'est pas vide lorsqu'aucun des facteurs n'est vide ; « si T est un ensemble dont tous les éléments sont des ensembles disjoints non vides, leur réunion $\mathfrak{S}T$ contient au moins une partie S_1 qui a, avec chaque élément de T, un et un seul élément commun [4] ». Le choix général, dans le cas où les ensembles éléments de T ne sont pas disjoints, s'en déduit, car les axiomes précédents permettent de remplacer toujours un ensemble quelconque T ayant pour éléments des ensembles A, B, C,... par un ensemble équivalent T', dont les éléments A' B', C', ... sont disjoints et respectivement équivalents à A, B, C,...

A part l'axiome de choix et l'axiome de l'ensemble des parties, ces énoncés reproduisent exactement — aux

(1) ZERMELO (III), p. 266.

(2) *Ibid.*, p. 263.

(3) A l'exception des parties comme $|a|$, $|a, b|$ — a et b étant des éléments de l'ensemble — dont l'existence est posée par l'axiome II.

(4) *Ibid.*, p. 266.

termes près — les définitions placées par DEDEKIND dans le § 1 de son ouvrage. L'axiome de l'infini remplace le théorème de l'infini dont la démonstration utilise l'ensemble illégitime de toutes les pensées. On est donc certain de retrouver la théorie classique. ZERMELO en donne la preuve pour la théorie de l'équivalence. Si les démonstrations sont légèrement compliquées, aucun procédé nouveau n'apparaît : le raisonnement de la diagonale et celui de la chaîne se laissent traduire. D'autre part l'antinomie RUSSELL disparaît : de l'axiome III, de *séparation*, résulte que tout ensemble M contient au moins une partie qui n'est pas élément de M ; c'est le procédé de la diagonale qui la donne : soit M la partie formée de tous les éléments qui ne sont pas éléments d'eux-mêmes (propriété *définie*, donc M_0 existe) ; M_0 ne peut être élément de M, sous peine de se contenir comme élément, ce qui serait contradictoire. Or dans l'ensemble de tous les ensembles, s'il existait, toute partie serait en même temps élément. D'une façon plus générale les axiomes ne permettent de poser comme ensemble le système de tous les objets possédant une propriété *définie*, que si les objets sont éléments d'un ensemble déjà reconnu comme tel : ni l'ensemble de tous les ensembles, ni l'ensemble de tous les ordinaux ne peuvent satisfaire à la condition.

Cependant, la théorie des ensembles n'est pas fondée pour cela : en effet, l'œuvre de ZERMELO reste inachevée ; il n'est montré ni que la théorie de l'ordre et des nombres ordinaux peut se déduire des axiomes, ni que ceux-ci sont indépendants entre eux et non contradictoires. D'autre part — et ceci en liaison avec l'absence de démonstration de non contradiction — la forme même des axiomes prête à critique : l'axiome de *séparation* fait intervenir la possibilité d'une décision d'après « les axiomes posés et les lois de la logique ». Or les antinomies de la théorie des ensembles en éveillent d'autres dans le domaine de l'épistémolo-

gie générale, comme celle de RICHARD sur les définitions en termes finis (1). Toutes les difficultés du problème de la définition mathématique et de celui de la décision se trouvent là, malencontreusement greffées. A combler cette double lacune en modifiant les axiomes et précisant leurs rapports entre eux et avec l'édifice cantorien, s'est attaché FRAENKEL dans une série de mémoires close récemment (2). Trois résultats principaux peuvent en être dégagés :

1° Transformations des notions fondamentales de l'axiomatique — et par suite de l'énoncé même des axiomes ; α) transformation du matériel de la théorie : la notion d'un champ général des objets de la théorie, parmi lesquels les ensembles devraient être distingués, est abandonnée ; on exclut tous les objets non-ensembles (à l'exception de l'ensemble vide). Seuls les ensembles sont considérés, et, d'abord — pour la théorie générale, ou théories de l'équivalence, des ensembles ordonnés et bien ordonnés — de façon hypothétique. Il n'y a pour cette théorie aucun axiome d'existence inconditionnée. L'axiome de *détermination* devient : « si a, b et A sont des ensembles et si $a = b$ et $a \in A$ on a aussi $b \in A$ » (l'égalité entre deux ensembles a et b est définie par la relation : pour tout ensemble x, $x \in a$ entraîne $x \in b$ et réciproquement). L'axiome de détermination garantit à cette relation les propriétés de l'égalité mathématique ordinaire : substituabilité de deux objets égaux dans un énoncé de la théorie. L'axiome des *ensembles élémen-*

(1) Si l'on considère l'ensemble des fractions décimales définissables par un nombre fini de mots : on peut les ranger en une suite dénombrable bien ordonnée et le procédé de la diagonale procure aussitôt une nouvelle fraction extérieure à la suite et pourtant définissable également avec un nombre fini de mots : le procédé de la diagonale n'est pas responsable de la contradiction, mais l'incertitude de la « définition en termes finis ». Ces antinomies « épistémologiques » ne concernent pas directement la théorie des ensembles.

(2) V. en particulier FRAENKEL (VII) et (X) où les résultats de ces mémoires sont résumés.

taires se réduit à un axiome d'*accouplement* (affirmation d'existence hypothétique) : si a et b sont des ensembles distincts, il existe un ensemble ayant pour éléments a et b : $|a, b|$;

β) transformation principale : la notion de fonction définie inductivement remplace la notion de propriété *définie*. L'axiome de *séparation* s'énonce : s'il existe un ensemble M et deux fonctions données $\varphi(x)$ et $\psi(x)$, il existe une partie de M comprenant tous les éléments x de M pour lesquels $\varphi(x) \in \psi(x)$ ou $\psi(x) \in \varphi(x)$. Comme fonctions de l'ensemble x sont posées : d'abord les constantes (c'est-à-dire les ensembles admis comme existant inconditionnellement), x lui-même, l'ensemble des parties de x ($\mathfrak{P}x$), l'ensemble réunion $\mathfrak{S}x$, l'ensemble couple $|\varphi(x),\psi(x)|$ (si $\varphi(x)$ et $\psi(x)$ sont déjà connus) enfin l'ensemble fonction de fonction $\varphi(\psi(x))$. On le voit, le progrès sur ZERMELO a consisté à interpréter la phrase : « propriété décidable au moyen des axiomes de la théorie », en termes mêmes de la théorie (au moyen de la relation fondamentale $\in$ et des fonctions que les axiomes II, IV, V de ZERMELO procurent ; l'axiome VI ne définissant pas de façon univoque, aucun ensemble de choix ne peut intervenir). Mais le procès inductif ne s'arrête évidemment pas là : si dans la définition de l'ensemble M ou des fonctions $\varphi(x)$, $\psi(x)$ intervient une autre variable y, la partie d'ensemble :

$$M_{\varphi \ldots \in \psi \ldots}$$

dépend aussi de y, elle constituera donc aussi une fonction de y. A la vérité une difficulté apparaît dès le début du procès si, par exemple, pour la définition d'une partie de M l'ensemble $\mathfrak{P}M$ des parties de M est utilisé : c'est là, note FRAENKEL, « un exemple classique de définition non prédicative » (1). La solution est celle qu'admettent les mathé-

(1) FRAENKEL (VII), p. 112.

maticiens non intuitionnistes en pareil cas : une construction de la partie en question est impossible, mais il peut y avoir un sens mathématique dans la caractérisation d'un élément d'un ensemble, fût-ce au moyen de la totalité de l'ensemble (par exemple la borne supérieure ou inférieure en analyse).

2° Sectionnement de la théorie en fragments rattachés à un nombre croissant d'axiomes. Grâce aux axiomes précédents auxquels sont joints sans modification les axiomes de l'ensemble des parties, de l'ensemble réunion, et du choix (IV, V, VI de ZERMELO) peut être construite la théorie générale des ensembles (1) : théorie de l'équivalence, théorie des ensembles ordonnés et bien ordonnés dont FRAENKEL a opéré le développement. En particulier sont assurées, au cas où en général un ensemble existe (ce qui n'a pas encore été posé), l'existence de l'ensemble vide et celle de l'ensemble formé d'un élément (2). L'ordre est défini par le procédé de subsomption des restes (introduit par HESSENBERG (3)) : un ensemble M est ordonné s'il existe une partie de $\mathfrak{T}$ de l'ensemble des parties $\mathfrak{P}M$ de M, telle que : 1° de deux éléments de $\mathfrak{T}$ l'un est partie de l'autre et 2° $\mathfrak{T}$ est une partie saturée de $\mathfrak{P}M$ possédant cette propriété (c'est-à-dire la relation $\mathfrak{T} \subset \mathfrak{T}'$, $\mathfrak{T}'$ possédant la propriété 1° entraîne $\mathfrak{T} = \mathfrak{T}'$) (4). Avec l'axiome de l'infini sous la forme

(1) Note de FRAENKEL in VON NEUMANN (IV), p. 392.

(2) FRAENKEL (V), p. 256.

(3) HESSENBERG (I) ; cf. aussi FRAENKEL (X) et (VII), p. 137 et KURATOWSKI (I).

(4) On dit que $\mathfrak{T}$ ordonne M : la 1re propriété établit un ordre entre tous les éléments de M tels que deux quelconques d'entre eux a et b appartiennent respectivement à deux éléments distincts $T(a)$ et $T(b)$ de $\mathfrak{T}$; on a :

$$a \prec b \text{ si } T(a) \supset T(b)$$

(une des 2 relations entre $T(a)$ et $T(b)$ existe toujours ; relation transitive) ; la 2e propriété garantit que l'ordination est complète (étendue à tous les éléments de M).

donnée par ZERMELO sont posées les existences classiques : ensembles élémentaires, dénombrable, continu et ensembles dont les puissances sont de forme $2^{2^{\aleph_0}}$, $2^{2^{2^{\aleph_0}}}$, etc... ... En revanche, l'existence de l'ensemble dénombrable

$$\{Z, \mathfrak{P}Z, \mathfrak{P}\mathfrak{P}Z, \mathfrak{P}\mathfrak{P}\mathfrak{P}Z, \ldots\}$$

où Z représente la suite des entiers, ne peut être assurée, non plus que la théorie des ordinaux transfinis, ce qui était déjà le cas dans l'axiomatique de ZERMELO. FRAENKEL énonce dans ce but un axiome VIII de *substitution* (1) : « Si M est un ensemble et $\varphi(x)$ une fonction, il existe un ensemble $\mathfrak{M}$ déduit de M par substitution de $\varphi(x)$ à chacun de ses éléments x ». Il n'y a gain, ainsi que l'a montré VON NEUMANN (2), que si on élargit la notion de fonction : $\varphi_y(x)$ et $\psi_y(x)$ étant deux fonctions de x dépendant du paramètre y, « si pour tout y (parcourant un certain ensemble) existe l'ensemble $\Phi(y)$ des x tels que :

$$\varphi_y(x) \in \psi_y(x),$$

$\Phi(y)$ est une fonction de y ». L'élargissement vient de ce que les ensembles des x, $\Phi(y)$, ne sont plus — comme dans les précédentes définitions — astreints à être parties d'un certain ensemble majorant T, déjà posé, le même pour tous les y, c'est-à-dire tel qu'on aurait, pour chaque y :

$$\Phi(y) = T_{\varphi_y(x) \in \psi_y(x)}$$

On obtient donc davantage de fonctions (3). Telles sont les trois principales étapes dans le développement axiomatique de la théorie cantorienne. Doit-il y en avoir d'autres, si l'on

(1) FRAENKEL (X), p. 309.

(2) VON NEUMANN (IV), p. 377 : avec les fonctions définies comme précédemment VON NEUMANN montre que l'axiome de substitution est une conséquence des autres axiomes.

(3) VON NEUMANN (IV), p. 376.

veut retrouver tous les résultats que la continuation des méthodes naïves (1) a permis d'atteindre après CANTOR ? la question est encore en suspens.

3° Démonstration d'indépendance relative entre les axiomes (2) : il n'y a là, semble-t-il, qu'exigence esthétique pour l'édification de la théorie. Aussi bien se trouve-t-elle satisfaite en gros par la façon même dont ils sont posés les uns après les autres comme garantie indispensable pour l'existence (hypothétique ou catégorique) de certains ensembles. Seule l'indépendance de l'axiome de choix concerne directement le contenu de la théorie classique : les objections contre la démonstration de ZERMELO venaient, on l'a vu, de la faible distance entre hypothèse et conclusion. Au lieu de s'attaquer à la validité de l'hypothèse — donc de rejeter simultanément les deux propositions, puisque équivalentes — une autre voie serait de la supprimer en tant qu'hypothèse, parce que démontrable à partir des axiomes déjà posés. Le particulier de son énoncé même y invite : puisqu'il garantit, étant donné un ensemble M d'ensembles, l'existence d'une certaine partie de l'ensemble réunion $\mathfrak{S}M$; il devrait tomber sous la juridiction de l'axiome de séparation (par quoi sont posées en général les existences de partie d'ensembles): en progressant assez dans la construction des fonctions introduites par les définitions de FRAENKEL, pourrait-on atteindre un ensemble de choix ? FRAENKEL a réussi sous certaines conditions à montrer qu'il n'en est rien. Il distingue quatre énoncés, de forces en gros croissantes : VI_1 existence d'un ensemble de choix pour un ensemble dénombrable d'ensembles finis ; VI_2 pour un ensemble

(1) Cf. les travaux de MAHLO (I) et (II), H. HAUSDORFF (I) et BAER (I) sur les nombres initiaux réguliers à indices limites, cités par FRAENKEL (X), p. 310.

(2) Les trois étapes précédentes n'indiquent nullement un ordre essentiel entre les axiomes : chaque axiome est indépendant des précédents et des suivants.

dénombrable d'ensembles quelconques ; VI_3 pour un ensemble quelconque d'ensembles finis ; VI_4 pour un ensemble quelconque d'ensembles quelconques. L'indépendance pour eux se définit ainsi : VI_1 peut n'être pas satisfait lorsque tous les axiomes *I-V* et *VII* de ZERMELO-FRAENKEL sont satisfaits ; VI_2, VI_4 peuvent chacun n'être pas satisfaits lorsque *I-V*, VI_1 (éventuellement aussi VI_3), et *VII* sont satisfaits. La démonstration a été effectuée par FRAENKEL pour un domaine différent de celui que considère son axiomatique. Au lieu de se restreindre aux seuls objets ensembles, il considère — comme faisait primitivement ZERMELO — d'autres objets de la théorie, qui ne sont pas des ensembles, et à partir desquels l'axiome d'accouplement en particulier permet d'en former. La définition de l'égalité (pour ces objets elle doit exiger l'identité) doit être modifiée en conséquence. En revanche les autres amendements FRAENKEL (nouvelle forme de l'axiome de séparation) sont maintenus et jouent un rôle essentiel dans la démonstration : on procède par induction complète (1) sur la construction de tous les ensembles qui peuvent être obtenus dans la théorie à partir d'une certaine base définie : la base et tous les ensembles

(1) On construit une pseudo-théorie où figure, pour le cas VI_1 un ensemble dénombrable de couples d'objets qui ne sont pas eux-mêmes des ensembles:

$$P = \{(a_1 b_1), (a_2 b_2), \ldots\}$$

et on démontre que tout ensemble M de la théorie est invariant si on permute entre eux les éléments accouplés dans une certaine partie P_m de P comprenant presque tous les couples de P (c'est-à-dire tous sauf un nombre fini). Ainsi l'ensemble $\{a\}$ est invariant par rapport aux permutations effectuées dans les couples de l'ensemble comprenant tous les couples de P, sauf $(a\ b)$. Donc il ne peut exister pour P d'ensemble de choix (ayant un élément et un seul commun avec chacun des couples) : il ne serait invariant par rapport à aucune partie de P. Pour les axiomes VI_2, VI_4 on remplace dans P les couples par des ensembles quelconques, pour les axiomes VI_3 et VI_4, P n'est pas assujetti à être dénombrable. Cf. FRAENKEL (VII), p. 165 sq. et (XIII), p. 42-51. La démonstration pour VI_1 a été publiée in FRAENKEL (III). Dans FRAENKEL (XIV) vient d'être donnée la démonstration détaillée pour l'indépendance de l'axiome général (VI_4) par rapport au système des axiomes I-V, VII et VI_1.

qui en dérivent possèdent une propriété qui exclut l'existence de l'ensemble de choix pour certains ensembles d'ensembles. Pour un domaine où apparaissent les seuls ensembles aucune démonstration n'est en vue : preuve que cette élimination des objets non ensembles n'est pas inoffensive : elle semble nécessaire si l'on veut préciser le caractère mathématique de la théorie.

A cet achèvement de l'œuvre zermelienne, manquait encore une définition des nombres ordinaux transfinis. Les travaux de Von NEUMANN (1) à ce sujet, outre leur intérêt intrinsèque, permettent de fonder les définitions par récurrence transfinie et — par le détour d'une autre axiomatique — ouvrent de nouvelles perspectives sur les rapports entre elles des différentes parties de la théorie cantorienne. Le nombre ordinal peut être caractérisé (de façon traduisible dans le système ZERMELO-FRAENKEL) comme un ensemble bien ordonné dont chaque élément est égal au segment qu'il détermine dans l'ensemble. On retrouve le procédé naïf de la numération : compter les éléments d'un système est bien remplacer chaque système partiel déjà compté par un nombre, à la fois élément succédant aux nombres précédents, et représentant de leur collectivité. L'ordre d'un nombre ordinal, ou d'un système de nombres ordinaux est l'ordre de subsomption, introduit par HESSENBERG (2) : les relations $P < Q$ et $P \subset Q$ sont équivalentes (immédiat d'après la définition). A tout ensemble bien ordonné correspond un nombre ordinal et un seul, le nombre ordinal qui lui est semblable. Par induction complète transfinie : d'une part deux nombres ordinaux semblables sont identiques, dans l'application de l'un sur l'autre le premier élément qui serait distinct de son image est égal au segment qui le pré-

(1) VON NEUMANN (I), (III) et (IV).
(2) HESSENBERG (II).

cède, segment par hypothèse identique à son image. D'autre part la propriété pour un ensemble bien ordonné d'être semblable à un nombre ordinal se transmet de l'ensemble à ce segment (par suite de l'identité entre un segment d'un segment et un segment de l'ensemble) et des segments à l'ensemble. Pour la deuxième transmission — qui reprend le procédé de DEDEKIND pour appliquer la suite des entiers sur un ensemble lorsque tout segment de cette suite y est applicable — l'axiome de substitution (non ici l'axiome de choix puisque l'application est déterminée de façon univoque, un ensemble bien ordonné ne pouvant être semblable qu'à un seul nombre ordinal) doit être utilisé : des fonctions bien déterminées appliquant chaque segment de M sur un nombre ordinal, se déduit une fonction $f(x)$ faisant correspondre à l'élément x, définissant un segment, le nombre ordinal sur lequel il est appliqué. Les valeurs de $f(x)$ constituent, en vertu de l'axiome de substitution, un ensemble dont on voit sans peine qu'il possède les propriétés d'un nombre ordinal et qu'il est semblable à M. L'induction complète transfinie achève le raisonnement : la *nombrabilité* (1) (similitude avec un nombre ordinal) se transmet le long du bon ordre.

Deux propriétés jouent surtout un rôle dans la théorie : pour aucun ensemble qui soit nombre ordinal, x, on ne peut avoir $x \in x$ (conséquence immédiate de la théorie du bon ordre) ; on peut prendre comme propriété caractéristique du nombre ordinal : ensemble P de nombres ordinaux tel que tout élément soit en même temps une partie (si $x \in P$, $x \subset P$). Les procédés d'engendrement de CANTOR — dont on a vu qu'ils étaient englobés dans la définition du bon ordre — peuvent alors être déduits. Si M est un ensemble de nombres ordinaux, $\mathfrak{S}M$ sera un nombre ordinal (en vertu

(1) En allemand : ***Zählbarkeit*** ; à distinguer de la ***dénombrabilité*** (Abzählbarkeit).

de la définition précédente). Le passage à la limite, après une suite N bien ordonnée de nombres ordinaux, consiste alors à considérer l'ensemble réunion $\mathfrak{S}N$; ce nouveau nombre ordinal suit bien immédiatement tous les éléments de la suite (si elle n'a pas de dernier élément) : on sait en effet que si x est un nombre ordinal et :

$$x \prec \mathfrak{S}N \quad \text{on a } x \subset \mathfrak{S}N \text{ donc } \quad x \in N.$$

L'adjonction d'une unité se fait par considération de l'ensemble x qui est un nombre ordinal (également en vertu de la définition précédente). On démontre aisément que pour tout nombre ordinal existe l'alternative :

$$x = \mathfrak{S}x \text{ (nombre de } 2^{e} \text{ espèce) ou } x = \mathfrak{S}x + \overline{\mathfrak{S}x}$$

(nombres de 1^{re} espèce) : on voit que si l'ensemble des nombres précédant x a un dernier élément y, on a $y = \mathfrak{S}x$; on peut donc écrire :

$$x = y + 1$$

La théorie des définitions par récurrence transfinie (1) se déroule dans une symétrie parfaite avec celle de DEDEKIND pour les définitions par récurrence finie. Il s'agit de prouver l'existence d'une fonction et d'une seule $f(x)$ définie pour tout nombre ordinal x et satisfaisant à :

$$f(x) = \varphi\,(F(f, x), x) \qquad (1)$$

φ étant une fonction donnée et $F(f, x)$ représentant l'ensemble des valeurs de f pour les éléments du nombre ordinal x. On démontre par induction complète transfinie que pour tout nombre ordinal y existe un *élément de fonction* $f_y(x)$ et un seul, défini pour tout $x \prec y$ et satisfaisant pour ces valeurs à (1) : la propriété se transmet en effet d'un ensemble bien ordonné à tous ses segments et de ceux-ci à l'ensemble. Si l'on fusionne en posant :

(1) Le procédé de définition par récurrence transfinie tel que l'employaient HAUSDORFF et FRAENKEL n'était pas encore rigoureusement fondé.

$$f(x) = f_x(x)$$

on a bien la fonction cherchée.

Les résultats précédents ne dépendent que de la théorie du bon ordre, sans théorème du bon ordre. Ils ne font intervenir ni l'axiome de choix ni la théorie de l'équivalence, au moins dans sa totalité. D'où l'idée d'inverser le développement traditionnel du système cantorien, idée que CANTOR semble avoir eue lui-même. L'axiomatique établie par VON NEUMANN se distingue en outre d'une part en considérant comme objets des fonctions et des arguments, au lieu d'ensembles et d'éléments (différence technique accessoire, destinée seulement à simplifier) ; d'autre part en ne limitant pas le champ de la théorie par exclusion des objets « trop gros », sources des paradoxes connus, mais en imposant seulement que toutes les fonctions ne puissent devenir arguments. Les contradictions n'apparaissent en effet que lorsque l'ensemble de tous les ensembles ou de tous les ordinaux sont traités en éléments possibles d'ensembles. Tous les arguments peuvent être fonctions (simplification (1) analogue à celle de FRAENKEL par rapport au système ZERMELO), mais la condition nécessaire et suffisante pour qu'une fonction ne puisse être argument sera qu'à tout argument du domaine corresponde de façon bi-univoque un argument pour lequel cette fonction a une valeur différente d'un argument A donné une fois pour toutes. Si l'on appelle *élément x du système S*, tout argument pour lequel la fonction :

$$f(x) \neq A,$$

on voit que pourront être éléments et eux seuls, les systèmes ne pouvant être appliqués de façon bi-univoque sur le système de tous les arguments (*Axiome N*) : les systèmes pouvant être éléments sont les ensembles.

(1) Simplification que ne fait d'ailleurs VON NEUMANN ni dans (II) ni dans (III) mais seulement dans (V).

Cet axiome remplace à la fois, dans l'axiomatique ZERMELO-FRAENKEL, les axiomes d'accouplement, de séparation, de substitution et de choix. Il écarte évidemment l'antinomie RUSSELL : le système de tous les ensembles n'est pas un ensemble. Quant à l'antinomie BURALI-FORTI, elle se transforme en *théorème du bon ordre,* suivant le procédé que CANTOR avait essayé avec la distinction entre ensembles consistants et inconsistants (1). Le système Ω^{**} de tous les ordinaux ne peut être un ensemble : en effet on a vu que tout ensemble d'ordinaux dont les éléments sont à la fois ses parties (condition évidemment remplie) est un nombre ordinal : donc ce nombre ordinal serait son propre élément, ce qu'on a vu également être impossible) (2). S'il n'est pas un ensemble, il est, en vertu de l'axiome N, applicable sur le système Ω de tous les éléments (ou arguments). Cette application bi-univoque communique à Ω le bon ordre de Ω^{**} : comme tout ensemble est une vraie partie de Ω, il est bien ordonnable.

Le raisonnement a l'avantage de mettre nettement en lumière le caractère spécial de l'antinomie : c'est une propriété particulière aux nombres ordinaux (découlant de la théorie du bon ordre: un segment d'un ensemble bien ordonné ne peut être semblable à un segment qui en est une vraie partie) qui empêche Ω^{**} d'être un ensemble. En revanche, la conclusion est entachée de tout l'arbitraire de l'axiome N: celui-ci pose l'existence d'une fonction qui applique le système non ensemble (ici Ω^{**}) sur le système Ω de tous les arguments ; d'elle se déduit la fonction φ bien ordonnante de toute partie de Ω, et également la

(1) V. ci-dessus, p. 95.

(2) Dans (III) VON NEUMANN admet qu'un système (non ensemble) puisse être un nombre ordinal : mais tous ses segments doivent être des ensembles puisqu'ils figurent comme argument dans la fonction faisant correspondre au segment l'élément qui le détermine. Ω^{**} est alors le seul nombre ordinal qui ne soit pas un ensemble : rien n'est changé au raisonnement du texte.

fonction ψ de choix qui distingue, pour chaque ensemble, le premier élément dans le bon ordre imposé par φ. La théorie de l'équivalence se peut ensuite facilement développer : on appelle nombre cardinal un nombre ordinal qui ne peut être appliqué de façon bi-univoque sur aucune de ses vraies parties ; l'universelle comparabilité des puissances se déduit de celle des nombres ordinaux.

On retrouve donc en pleine lumière ce qui n'était que latent chez CANTOR : il n'y a, en réalité, qu'une seule catégorie de nombres, les nombres ordinaux ; le caractère numérique est lié aux propriétés de l'acte de compter d'où sortent directement bon ordre et nombres ordinaux. Ce n'est que l'intervention de considérations étrangères (de nouvelles fonctions), qui parmi les ordinaux en distingue certains doués d'une propriété particulière, les nombres initiaux de classe. Mais ici encore apparaît un accident dans la création : le troisième principe d'arrêt. On peut faire toute la théorie des ordinaux sans y avoir recours, développer toute l'arithmétique ordinale transfinie non seulement sans lui, mais sans même s'inquiéter d'une pluralité de puissances.

C'est, en réalité, la mathématique classique qui exige leur affirmation. Mais ici apparaît une dernière difficulté que Von NEUMANN a contribué à préciser ; les comparaisons de puissances — et d'une façon générale les propriétés de bon ordre, etc.. — sont relatives à l'axiomatique considéré. Si, à un système Σ_1 satisfaisant aux axiomes de Von NEUMANN, on superpose un système Σ_2, soumis pour l'essentiel aux mêmes axiomes, mais comportant davantage de fonctions (ce qu'il est possible d'obtenir par des modifications accessoires des axiomes), le même objet a pourra être dans Σ_1 un ensemble non dénombrable et se trouver dénombrable dans Σ_2 : ainsi également pour la finitude et le bon ordre : « il n'est pas exclu que l'ensemble a paraisse fini et bien ordonné dans le système Σ_1 et se révèle dans le système « plus fin » Σ_2

non bien ordonné (infini) simplement parce qu'une partie de *a*, qui n'a pas de premier élément, n'était pas un ensemble dans Σ_1, donc n'avait pas été remarqué, tandis qu'elle est un ensemble dans Σ_2 » (1). Donc relativité aussi bien vers le bas (finitude et bon ordre) que vers le haut.

Même relativité dévoilée antérieurement par le paradoxe de SKOLEM : chaque système d'axiomes est satisfait par une infinité dénombrable d'objets (par exemple des ensembles que procurent les axiomes ZERMELO-FRAENKEL) (2). Il s'agit évidemment d'une autre dénombrabilité que de celle qui est définissable au sein de l'axiomatique. Mais le divorce même est significatif : entre le mouvement naïf des créations mathématiques et les systèmes axiomatisés la jonction immédiate est impossible. On a l'impression de microcosmes, ingénieuses miniatures « qui représentent fidèlement la théorie des ensembles jusque dans tous les détails.. Il y paraît toutes les puissances... Mais aussitôt que l'on met en jeu des instruments plus fins tout s'évanouit. De toutes les puissances ne restent que le fini et le dénombrable. Cela seul a un sens réel, tout le reste est fiction formaliste (3) ». Au moins l'utilisation dans des problèmes mathématiques en est compliquée.

L'axiomatisation de la théorie des ensembles n'était, dans l'intention, que codification de démarches effectives pour l'élimination des antinomies. On ne peut guère prouver que le résultat soit atteint : si les ensembles de RUSSELL et de BURALI-FORTI disparaissent, MIRIMANOFF a donné l'exem-

(1) VON NEUMANN (II), p. 240.

(2) La démonstration rigoureuse ne peut s'effectuer que dans un système entièrement formalisé, quoique SKOLEM ait montré que le résultat est indépendant du mode de formalisation ; la construction récurrente d'ensembles, telle qu'elle s'effectue à partir du seul axiome d'existence absolue, axiome de l'infini, dans le système ZERMELO-FRAENKEL, grâce à l'axiome de séparation modifié, en peut donner idée.

(3) VON NEUMANN, *ibid.*, p. 239.

ple (1) d'un ensemble descendant, impossible à exorciser avec les axiomes connus si l'on ne veut pas se borner aux ensembles construits par les axiomes (c'est-à-dire à une infinité dénombrable) comme cela est possible dans les systèmes FRAENKEL et VON NEUMANN, et sur lequel aucun calcul n'est possible. Seule une démonstration de non contradiction donnerait la sécurité : en 1907 ZERMELO avouait n'y être pas encore arrivé (2). Toutes les démonstrations de non contradiction connues consistent à fabriquer un modèle de la théorie, dans une théorie mathématique déjà fondée : pour la théorie des ensembles, le procédé est évidemment inapplicable. En revanche la difficulté du relativisme enlève aux démarches axiomatisées leur efficacité concrète : si l'on veut garder sa place à la théorie, il faut que, pierre d'angle des mathématiques, elle n'en soit pas séparée dans l'opération d'axiomatisation. C'est à une reconstruction complète de l'édifice qu'il importerait de se livrer, qu'elle aboutisse au moyen d'une axiomatisation rigoureuse, ou négativement — avec l'intuitionnisme — par définition nouvelle des enchaînements admissibles : dans tous les cas le problème de la théorie des ensembles devient problème du fondement des mathématiques.

(1) MIRIMANOFF (I) : il s'agit de l'ensemble $a = \{a\}$ avec $\{a\} = \{\{a\}\}$, etc...
(2) ZERMELO (III). p. 262.

CONCLUSION

Il est possible de donner au procès d'axiomatisation une signification moins absolue : sorte de condensation de la matière active d'une théorie autour de quelques procédés intuitivement saisissables et pour lesquels la question d'une approbation, sinon d'une justification, est plus facile à résoudre. De ce point de vue les recherches de ZERMELO-FRAENKEL et Von NEUMANN sont aussi utiles pour le développement de la théorie, que significatives dans le dessin de son histoire. Grâce à elles semblent se fixer — du moins autant qu'il se peut actuellement — portée du système et liaisons véritables.

La théorie de l'équivalence est née au sein de l'analyse, avec des méthodes empruntées à l'analyse. Si BOLZANO pressentait déjà les notions de la réunion et du produit d'ensembles, WEIERSTRASS, à propos des nombres complexes, en donnait la définition précise que CANTOR n'eut qu'à transposer (1) : opérations simples, résultats pauvres, tout l'intérêt de la théorie est dans la manipulation qu'elle rend plus aisée, dans les mises en rapport qu'elle établit pour des ensembles déjà procurés par l'analyse. Si on la veut préalable à celle-ci, elle n'existe que dans la mesure où sa liaison avec elle, de contiguïté devient continuité, sous peine de s'évanouir dans le vide logique. Nombre, fonction, ces actes intuitifs de l'esprit sont des instruments qui se forgent eux-mêmes, car rien n'existe avant eux. La duperie de l'application comme opération logique antérieure à la fonction se manifeste dans la floraison des paradoxes : c'est, après ZERMELO, la précise définition de l'opération de séparation (2), par construction récurrente de fonctions, qui

(1) V. ci-dessus, p. 18.
(2) SKOLEM s y était déjà employé in SKOLEM (II).

permet à FRAENKEL de relever l'édifice de Dedekind. Après les fonctions réunion, ensemble des parties, etc., le moment non prédicatif permet une génération illimitée qui ouvre les perspectives de la théorie. Peut-elle aller très loin avec ses seules ressources (sans axiome de substitution) ? Il est difficile d'en décider : outre la définition de l'ordre et du bon ordre, elle n'a guère obtenu de résultats après les théorèmes de BERNSTEIN et de KOENIG. Pour être fécondée il semble qu'il lui faille des objets — ensembles ou fonctions — définis dans la mathématique classique : fonction additive, mesure dans l'intégration abstraite, notion de voisinage pour les espaces topologiques. Résultats plus nombreux, sécurité gagnée tiennent à la solidarité mathématique intérieure des actes qu'ils réclament, au contrôle effectif auquel les soumet leur caractère plus concret. Il n'y a plus alors développement singulier d'une théorie autonome. Le problème de l'universelle comparabilité est assez révélateur de l'imagination, qui sous-tendait les travaux de CANTOR, d'un en-soi des ensembles : comment affirmer l'existence de la fonction appliquante, si les deux ensembles en présence ne sont définis par rien d'autre que leur existence ? il semble que les opérations de l'équivalence ne permettent pas de l'atteindre. Au lieu de problème il n'y a sans doute là qu'un nouvel axiome, dont HARTOGS (1) a montré qu'il équivaut à l'affirmation du bon ordre (et par suite à l'axiome de choix) : son indépendance par rapport aux autres axiomes résulte donc des travaux de FRAENKEL sur l'axiome de choix.

Avec la théorie des ordinaux, apparaît un mouvement de pensée radicalement nouveau — caractérisé par l'axiome de substitution — plus propre, semble-t-il, à procurer le développement d'une véritable théorie. On a vu que sa subordination aux considérations d'équivalence résulte d'un accident historique. Mais il faut encore noter que ce n'est pas le bon ordre qui est préalable au nombrement, mais le nombre-

(1) Il est évident que du bon ordre résulte la comparabilité (puisque tous les ensembles bien ordonnés sont comparables). HARTOGS a prouvé la réciproque. HARTOGS (I).

ment au bon ordre (1). Aucun ensemble mathématique effectif, hors la suite des entiers, ne donne exemple de bon ordre ; s'il est définissable avec les axiomes de l'équivalence, aucun moment intérieur de cette théorie n'y conduirait si l'idée n'en était fournie par l'engendrement des ordinaux. Il semble donc que — quels que soient par ailleurs ses avantages — la présentation axiomatique ne fasse ici que reprendre, à peine masqués, les enchaînements naïfs : la validité de celle-là, en l'absence de toute démonstration de non contradiction, ne pourrait s'appuyer que sur la réalité de ceux-ci ; retour aux principes d'engendrement de CANTOR. Le remarquable est d'ailleurs qu'une axiomatisation comme celle de Von NEUMANN d'une part élimine toute une série de raisonnements de la théorie de l'équivalence (démonstration Bernstein, démonstration Zermelo du bon ordre), d'autre part n'apporte aucune lumière sur le problème du continu : s'il est possible d'y prouver que la puissance d'un ensemble M est inférieur à celle de l'ensemble de ses parties $\mathfrak{P}M$ on n'a aucun moyen de comparer $2^{\aleph_0}$ et $\aleph_1$. Certaines existences posées par l'axiome N — ou par l'axiome de substitution de FRAENKEL — et celles données par d'autres axiomes, comme celui de l'ensemble des parties ou de l'ensemble réunion, semblent impossibles à mettre en rapport autrement que par un nouvel axiome. C'est aussi bien ce que rend vraisemblable la possibilité d'une deuxième hypothèse du continu, reprise par LUSIN (2), et qui, jusqu'ici, n'a donné lieu à aucune contradiction, bien qu'elle exclue la première : $2^{\aleph_0} = 2^{\aleph_1}$.

Sa place privilégiée, les difficultés rencontrées, donnent à l'histoire de la théorie des ensembles sa portée philosophique : l'ambition d'un développement préalable au reste des mathématiques devait permettre de préciser la notion de théorie mathématique, isolée comme *in vitro* ; l'obliga-

(1) D'autre part nous ne connaissons aucun moyen d'ordonner un ensemble autrement que par le bon ordre, tandis que nous connaissons des ensembles simplement ordonnés.

(2) Cf. LUSIN (I).

tion, après les paradoxes, d'énumérer exhaustivement les hypothèses, précise la notion même de démonstration : si les résultats ne sont qu'une transformation immédiate des axiomes (1), peut-on encore parler de théorie ? La réalité d'un raisonnement se trouve en quelque sorte mesurée par son brassage de matière mathématique, par l'unification qu'il procure entre objets — procédés — déjà mathématiquement existant. D'où la difficulté de divisions de droit dans un tissu vivant, par rapport auquel les affirmations ne sont que superstructures des procédés, où l'imprévisible déviation d'un geste procure à la fois les nouveautés et maintient les liaisons. Rejeter ou fonder une théorie n'est travail définitif, ni dépourvu de degrés ; surtout, comme pour les caractérisations d'unité, il ne peut s'opérer par simple investigation logique. Si les axiomatiques semblent dissocier deux théories : équivalence, nombres ordinaux (et peut-être une troisième où serait étudié le problème du continu), dont la sécurité va décroissant, les considérations pragmatistes du mathématicien militant ont le dernier mot. C'est dans l'analyse entière que se pose le problème d'une existence et d'un affermissement, pour lesquels une définition du travail mathématique en général est nécessaire.

(1) Il ne s'agit ici bien entendu en aucun cas des importantes démonstrations données par ZERMELO, FRAENKEL, VON NEUMANN au cours de leurs axiomatisations respectives. Mais il y a peu de théorèmes.

BIBLIOGRAPHIE [1]

OUVRAGES GÉNÉRAUX [2]

(cités dans tout le cours du livre).

CANTOR G. (XXVIII). — *Gesammelte Abhandlungen*, herausggben v. E. Zermelo, nebst einem Lebenslauf Cantors von A. Fraenkel. Berlin, Springer, 1932.

DEDEKIND R. (V). — *Gesammelte mathematische Werke*, herausggben v. R. Fricke, E. Noether u. O. Ore. Braunschweig Vieweg u. Sohn, 1931, 3 vol.

CANTOR-DEDEKIND (I). — *Briefwechsel* herausggben von. E. Noether u. J. Cavaillès, Paris, Hermann, 1937.

FRAENKEL A. (X). — *Einleitung in die Mengenlehre*, 3e ed. Berlin, Springer, 1928, t. 1, 420 p.

— (XI). — *Georg Cantor*. Jahresber. d. deutsch. Math. Verein., 1930.

HAUSDORFF F. (I). — *Grundzüge der Mengenlehre*. Leipzig Veit, u. C., 1914.

— (II). — *Mengenlehre* (2e éd. modifiée du précédent). Berlin W. de Gruyter, 1927.

HESSENBERG G. (I). — *Grundbegriffe der Mengenlehre* (Sonderdruck aus den *Abhandl. d. Friesschen Schule*). Göttingen, 1906.

SCHOENFLIES A. (I). — *Die Entwickelung der Lehre von den Punktmannigfaltigkeiten* I. Teil : *Jahresb. d. deutsch. Math. Verein.* t. 8 (1900), II. Teil: 2 Ergänzgsbd z. d. *Jahresb. d. deutsch Math. Verein.* (1908). Leipzig.

CHAPITRE I

§ 1-2. — ***Le continu et l'introduction des nombres irrationnels.***

BOLZANO (I). — *Rein analytischer Beweis des Lehrsatzes dass zwischen je zwei Werten die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege. Abhdl. d. Kgl. Ges.*

1. Il ne s'agit ici que d'ouvrages en liaison directe avec le texte. On trouvera une bibliographie complète sur la théorie des ensembles proprement dite (jusqu'à 1928) in Fraenkel (X).

2. La pagination de tous les mémoires cités de Cantor et de Dedekind est empruntée à celle des deux recueils Cantor (XXVIII) et Dedekind (V).

der Wiss. Prag., 1817. Herausggben v. Ph. E. B. Jourdain (Ostwald's Klassiker). Leipzig Engelmann, 1905.

CANTOR (V). — *Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen. Math. Annalen*, t. 5 (1872), p. 123-132.

CANTOR (XV). — *Über unendliche lineare Punktmannigfaltigkeiten* 5. *Math. Annalen*, t. 21 (1883), p. 545-591.

CANTOR (XXIV). — *Bemerkung mit Bezug auf den Aufsatz : zur Weierstrass-Cantorschen Theorie der Irrationalenzahlen. Math. Ann.*, t. 33 (1889), p. 476.

DEDEKIND (II). — *Stetigkeit und irrationale Zahlen*, 1re édit. Braunschweig Vieweg, 1888, 6e éd., 1930.

GRASSMANN (I). — *Lehrbuch der Arithmetik*, 1861 (Sämmt. W., II, 1, p. 295-349).

HEINE (III). — *Die Elemente der Functionenlehre. Journal f. reine u. angew. Mathematik*, t. 74 (1872), p. 172.

HELMHOLTZ (I). — *Zählen u. Messen erkenntnisstheoretisch betrachtet.* Phil. Aufsätze E. Zeller gewidm., 1887, p. 17-52.

ILLIGENS (I). — *Zur Weierstrass-Cantorschen Theorie d. irrationalen Zahlen Math. Annalen*, t. 33 (1889), p. 155-166.

— (II). — *Über die Definition der irrationalen Zahlen*, ibid., t. 35.

KOSSAK (I). — *Die Elemente der Arithmetik.* Programm des Werderschen Gymnasiums, 1872, p. 1-29.

LIPSCHITZ (I). — *Grundlage der Analysis*, Bonn, 1877.

MÉRAY (I). — *Nouveau Précis d'analyse infinitésimale.* Paris, Savy, 1872.

PASCH (I). — *Einleitung in die Differential u. Integralrechnnung.* Leipzig, 1882.

— (II). — *Über die Einführung der irrationalen Zahlen. Math. Annalen*, t. 40 (1892), p. 149-152.

TANNERY J. (I). — *Introduction à l'étude des fonctions d'une variable.* Paris, Hermann, 1886.

Pour une bibliographie complète sur les nombres irrationnels. V. Pringsheim-Molk. *Nombres irrationnels* in *Encyclopédie des sciences mathématiques*, I, 3.

§ 3. — *Séries trigonométriques.*

A. — Ouvrages généraux et histoires.

HANKEL (I). — *Untersuchungen über die unendlich oft oszillierenden u. unstetigen Funktionen* Gratulationsprogramm d. Tübinger Universität (1870), reprod. in *Math. Annalen*, t. 20 (1882), p. 63-112.

LEBESGUE (I). — *Leçons sur l'intégration et la recherche des fonctions primitives*, 1re éd. Paris, Gauthier-Villars, 1903, 2e éd. augmentée, 1928.

— (II). — *Leçons sur les séries trigonométriques.* Paris, Gauthier-Villars, 1905.

RIEMANN (I). — *Über die Darstellbarkeit einer willkürlichen Function durch eine trigonometrische Reihe. Abhandl Gölting Gesellsch.*, t. 13 (Habilitationschrift 1854). *Math. Werke*, p. 228.

SACHSE A. (I). — *Versuch einer Geschichte der Darstellung willkürlicher Functionen einer Variabeln durch trigonometrische Reihen*, Inaugural Dissertation Göttingen 1879, trad. *in Bulletin d. Sces mathématiques*, t. 4 (1880), p. 43-64 et 83-112.

B. — Les travaux autour de Dirichelet et de Riemann.

ASCOLI (I). — *Sulla serie di Fourier. Annali di Math.* (II) t. 4 (1873).

DU BOIS REYMOND P. (V). — *Über die sprungweise Werthveränderungen analytischer Functionen. Math. Ann.*, t. 7 (1873), p. 241.

— (V). — *Über die Darstellung stetiger Functionen durch trigonometrische Reihen. Götting Nachrichten*, t. 1 (1873), p. 572-574.

— (X). — *Beweis dass die Coefficienten der trigonometrischen Reihen... Abhdl. d. bayer. Akad. Math. Nat. Wiss Kl.* t. XII, 1875), p. 117-167.

— (XI). — *Untersuchungen über die Convergenz u. Divergenz der Fourierschen Darstellungsformeln*, ibid., IIte Abteilung (1876), p. 1.

DUHAMEL (I). — *Note sur la discontinuité des séries et sur les moyens de la reconnaître. Journ. Louville*, t. 12 (1854) p. 112-120.

HEINE E. (I). — *Handbuch d. Kugelfunctionen.* Berlin, 1861.

— (II). — *Über die trigonometrische Reihen. Journ f. reine u. angew. Math.*, t. 71 (1870), p. 352.

GENOCCHI A. (I). — *Interno ad alcune serie. Atti d. reale Acc. d. Sc. d. Torino*, t. 10 (1875).

LEJEUNE DIRICHLET (I). — *Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données. Journ. f. reine u. angew. Math.*, t. 4 (1829), p. 157.

— (II). — *Über die Darstellung willkurlicher Funktionen durch trigonometrischen Reihen. Dove's Repertorium.* I (1837).

— (III). — *Sur les séries dont le terme général dépend de deux angles. Journ. f. r. u. angew. Math.* t. 17 (1837).

— (IV). — *Sur l'usage des intégrales définies dans la sommation des séries finies ou infinies, ibid.*

LIPSCHITZ (I). — *De explicatione per series trigonometricas instituenda functionum unius variabilis arbitrarium et praecipue earum quae per variabilis spatium finitum valorum maximorum et minimorum numerum habent infinitum disquisitio. Journ. f. r. u. angew. Math.* t. 63 (1864) p. 296 — trad. française av. notes par P. Montel in *Acta mathematica* t. 36 (1913) p. 281 (les notes du texte renvoient à la trad. Montel).

SCHÄFFLI (I). — *Einige Zweifel an d. allgemeinen Darstellbarkeit einer willkürlicher periodischen Function einer variabeln durch eine trigonometr. Reihe*, Berne Univ. Programme, 1874.

— (II). — *Über die partielle Differential Gleichungen* $\frac{\partial \omega}{\partial t} = \frac{\partial^2 \omega}{\partial x^2}$. *Journ. f. r. u. angew. Math.*, t. 72 (1870), p. 283.

C. — Unicité du développement et notion de convergence uniforme.

ABEL (I). — *Untersuchungen über die Reihe* $1 + \frac{m}{1} x + \frac{m(m-1)}{1\cdot 2} x^2 + \ldots$ *Journ. f. r. u. angew. Math.*, t. I (1826), p. 311.

CANTOR G. (I). — *Über einen die trigonometrische Reihen betreffenden Lehrsatz. Journ. f. r. u. angew. Math.*, t. 72 (1870), p. 130-8.

— (II). — *Beweis dass eine für jeden reellen Wert von x durch eine trigonometrische Reihe gegebene Funktion f (x) sich nur auf eine einzige Weise in dieser Form darstellen lässt.*, ibid., t. 72 (1870), p. 139-142.

— (III). — *Notiz zu dem Aufsatze : Beweis dass eine...*, ibid., t. 73 (1871), p. 292-294.

— (IV). — *Über trigonometrische Reihen. Math. Annalen*, t. 4 (1871), p. 139-143.

— (V). — *Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen. Math. Annalen*, t. 5 (1872), p. 123-132.

CAUCHY (I). — *Cours d'analyse algébrique.* Paris, 1821.

— (II). — *Notes sur les séries convergentes dont les divers termes sont des fonctions continues d'une variable réelle ou imaginaire entre les limites données. C. R. Acad. des Sciences*, t. 36 (1853), p. 454-459.

DINI U. (I). — *Sulla serie die Fourier. Annali di Matematica* (II), t. 6 (1873), p. 21 et 298.

HEINE E. (III). — *Die Elemente der Functionenlehre. Journ. f. r. u. angew. Math.*, t. 74 (1872), p. 172.

SCHWARZ (I). — *Beweis eines für die Theorie der trigonometrischen Reihen in Betracht kommenden Hülfsatzes. Ges. Abhandl.* Berlin, 1890, t. II, p. 341-343.

SEIDEL (I). — *Note über eine Eigenschaft von Reihen welche discontinuirliche Functionen darstellen. Abh. d. bayer. Akad.*, 1847-1849.

THOMÉ (I). — *Über hypergeometrische Kettenbrüche. Journ. f. r. u. angew. Math.*, t. 66, p. 334

et les ouvrages cités en A) et B).

§ 4. — *Théorie de la croissance.*

BOREL E. (I). — *Leçons sur la théorie des fonctions*, 1re édit., Paris. Gauthier-Villars, 1898 ; 3e éd. augm., 1928. Note II sur la théorie de la croissance — (les notes du texte renvoient à la 3e éd.).

— (II). — *Leçons sur les fonctions entières*, 1re édit. Paris. Gauthier-Villars, 1900 ; 2e éd. 1920.

— (III). — *Leçons sur les séries à termes positifs*, Paris. Gauthier-Villars, 1902.

— (IV). — *Leçons sur la théorie de la croissance.* Paris. Gauthier-Villars, 1909.

DU BOIS REYMOND P. (I). — *Bemerkungen über die verschiedenen Werthe*

welche eine Function zweier reellen Variabeln erhält, wenn man diese Variabeln entweder nacheinander oder Gewissen Beziehungen gemäss gleichzeitig verschwinden lässt. Journ. f. r. u. angew. Math., t. 70 (1868), p. 10-45.

— (II). — *Sur les infinis des fonctions. Annali di matematica*, t. 4 (1871).

— (III). — *Théorème général concernant la grandeur relative des infinis des fonctions et de leurs dérivées. Journ. f. r. u. angew. Math.*, t. 74 (1871).

— (IV). — *Neue Theorie der Convergenz u. Divergenz von Reihen mit positiven Gliedern*, ibid., t. 76 (1872), p. 60.

— (VII). — *Über asymptotische Werthe, infinitäre Approximationen und infinitäre Auflösung von Gleichungen. Math. Ann.*, t. 8 (1874), p. 362.

— (VIII). — *Zwei Sätze über Grenzwerthe von Functionen zweier Veränderlichen*, ibid., t. 11 (1877), 145-8.

— (IX). — *Über die Paradoxen des Infinitärcalcüls*, ibid., t. 11 (1877), p. 149.

— (XII). — *Über Integration und Differentiation infinitärer Relationen Math. Ann.*, t. 14, p. 498-501.

— (XIII). — *Allgemeine Funktionen Theorie* Tübingen (1882).

HAHN H. (I). — *Über die nichtarchimedischen Grössensysteme. Sitz. Ber. d. kaiser. Akad. d. Wiss. Math. Nat. Wiss. Kl*, t. 116, II *a* (1907), p. 601.

HARDY G. H. (I). — *Orders of Infinity*. London, 1910.

PRINGSHEIM (I). — *Über die sogenannte Grenze u. die Grenzgebiete zwischen Convergenz u. Divergenz. Sitz. Ber. d. bayer. Akad. Math. Nat. Wiss Kl*,, t. 26 (1896).

— (II). — *Über die Du Bois Reymond'sche ConvergenzGrenze und eine besondere Form der Convergenz Bedingung für unendliche Reihen*, ibid., t. 27 (1897).

— (III). — *Allgemeine Theorie d. Convergenz u. Divergenz. Math. Ann.*, t. 35 (1890) et 39 (1891).

STOLZ O. (I). — *Über die Grenzwerthe der Quotienten. Math. Ann.*, t. 14 (1878), p. 231-240.

— (II). — *Zur Geometrie der Alten, insbesondere über ein Axiom des Archimedes. Math. Ann.*, t. 22 (1883), p. 504.

CHAPITRE II

§ 1. — *Bolzano et Dedekind.*

BOLZANO (II). — *Paradoxien des Unendlichen* herausggben von Prihonsky, Berlin, Mayer et Müller, 1899.

CANTOR (XV). — Voir § suivant.

DEDEKIND (I). — *Allgemeine Sätze über Räume* inédit publié par E. Noether in Dedekind (V), t. II, p. 353-355.

§§ 2-5. — ***La première création cantorienne.***

BENDIXSON (I). — *Quelques théorèmes de la théorie des ensembles de points Acta mathematica*, t. II (1883), p. 414.

CANTOR (VI). — *Über eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen. Journ. f. r. u. angew. Math.*, t. 77 (1874), p. 258-263.

— (VII). — *Ein Beitrag zur Mannigfaltigkeitslehre*, ibid., t. 84 (1877), p. 242.

— (VIII). — *Über einen Satz aus der Theorie der stetigen Mannigfaltigkeiten. Götting. Nachrichten*, 1879, p. 127-135.

— (IX). — *Über unendliche lineare Punktmannigfaltigkeiten*, I, *Math. Annalen*, t. 15 (1870), p. 1-18.

— (X). — *Über unendliche lineare Punktmannigfaltigkeiten*, II, ibid., t. 17 (1880), p. 355-359.

— (XI). — *Bemerkungen über trigonometrische Reiben. Math. Ann.*, t. 16 (1880), p. 113-115 ; *Fernere Bemerkungen über trigonometrische Reihen*, ibid., p. 267-269.

— (XII). — *Über ein neues u. allgemeines Kondensationsprinzip der Singularitäten von Funktionen. Math. Ann.*, t. 19 (1882), p. 588-594.

— (XIII). — *Über unendliche lineare Punktmannigfaltigkeiten*, III. *Math. Ann.*, t. 20 (1882), p. 183.

— (XIV). — *Über unendliche lineare Punktmannigfaltigkeiten*, IV, ibid., (1883), p. 51.

— (XV). — *Über unendliche lineare Punktmannigfaltigkeiten*, V. ibid., p. 545-586.

— (XVI). — *Sur divers théorèmes de la théorie des ensembles de points dans un espace continu à n dimensions. Acta Math.*, t. II (1883). p. 409-414.

— (XVII). — *Gundlagen einer allgemeinen Mannigfaltigkeitslehre*, Leipzig Teubner, 1883 (réunion des mémoires XI, X, XII, XIV, XV).

— (XVIII). — *Über unendliche lineare Punktmannigfaltigkeiten*, VI, *Math. Ann.*, t. 23 (1884), p. 453-488.

— (XIX). — *De la puissance des ensembles parfaits de points. Acta Math.*, t. 4 (1884), p. 381-392.

— (XXIII). — *Mitteilungen zur Lehre von Transfiniten*, I et II, *Zeitschr. f. Philos. u. phil. Kritik*, t. 91 (1887) p. 81-125 et 252-270 et t. 92 (1888), p. 240-265.

— (XXV). — *Über eine elementare Frage der Mannigfaltigkeitslehre Jahresber. d. deutsch. Math. Verein.*, t. 1 (1892), p. 75-78.

DU BOIS REYMOND (I). — *Allgemeine Funktionentheorie* (déjà cité).

GRASSMANN (II). — *Ausdehnungslehre*, 2e éd. 1862.

JOURDAIN (I) — *The development of the theory of transfinite numbers*, I, *Arch. d. Math. u Physik* (3) : I, t. 10 (1905), p. 244-281 ; II, t. 14 (1909) p. 289-311 ; III : t. 16 (1910) p. 21-43 ; IV : t. 22 (1914) p. 1-21.

SCHŒNFLIES (II). — *Die Krisis in Cantors mathematischen Schaffen. Acta math.*, t. 50 (1928), p. 1, sq.

§ 6. — *Systématisation de la théorie.*

BOREL E. (I). — *Leçons sur la théorie des fonctions* (cité pour ch. I, § 4).

CANTOR (XX). — *Zum Problem des aktualen unendlichen. Zeitschr. f. Philos. u. phil. Kritik*, t. 88 (1885) p. 224-233.

— (XXI). — *Über verschiedene Theoreme aus der Theorie der Punktmengen in einem n-fach ausgedehnten stetigen Raume G_n. Acta Math.*, t. 7 (1885), p. 105-124.

— (XXII). — *Über die verschiedenen Ansichten in Bezug auf die aktual unendlichen Zahlen. Stockholm Bihang till kongl. svenska Vetenskaps Akad. Handlingar*, t. 11 (1886) reproduction avec quelques modifications de (XX).

— (XXIII) *Mitteilungen zur Lehre von Transfiniten* I et II (déjà cité in § 2-5).

— (XXVI) *Beiträge zur Begründung der transfiniten Mengenlehre I. Math. Ann.* t. 46 (1895) p. 481-512.

— (XXVII). — *Beiträge zur Begründung der transfiniten Mengenlehre* II, ibid , t. 49 (1897), p. 343-347.

— (XXVI) et (XXVII) trad. en français par Marotte in *Mémoires de la Société des sciences physiques et naturelles de Bordeaux* (5) t. 3 (1899), p. 343-437, trad. en anglais par Ph. E. B. Jourdain, Chicago 1915.

— (XXIX). — *Sui numeri transfiniti (Estratto d'una lettera di Georg Canteor a G. Vivanti) Rivista di Matemat.*, t. 5 (1895), p. 104-108.

— (XXX). — *Lettera di Georg Canter a G. Peano, ibid.*, p. 108-109.

DEDEKIND R. (IV). — *Zweite Definition des Endlichen und Unendlichen*, opuscule inachevé publié *in* Dedekind (V) t. III.

DU BOIS REYMOND (IX). — *Über die Paradoxen des Infinitärkalküls* (cité pour ch. I, § 4).

— (XII). — *Über Integration u Differentiation infinitärer Relationen* (id).

STOLZ (II). — *Zur Geometrie der Alten, insbesondere über ein Axiom des Archimedes* (id).

VERONESE (I). — *Fundamenti di Geometria.* Padua, 1891.

§ 7. — *Le paradoxe Burali-Forti.*

BERNSTEIN F. (I). — *Über die Reihe der transfiniten Ordnungszahlen. Math. Ann.*, t. 60 (1905), p. 187-193.

BURALI-FORTI C. (I). — *Una questione sui numeri transfiniti. Rendiconti del Circolo Matemat. di Palermo*, t. 11 (1897), p. 154.

CHAPITRE III

§ 1. — *L'œuvre de Dedekind.*

A. — La création de 1887 et les ouvrages qui la préparent.

DEDEKIND R. (III). — *Was sind und was sollen die Zahlen ?* 1re éd. Braunschweig Vieweg u. Sohn 1887, 3e éd., 1911.

HELMHOLTZ (I). — *Zählen und messen erkenntnis theoretisch betrachtet*, Leipzig, 1887.

KRONECKER (I). — *Über den Zahlbegriff*, Leipzig, 1887.
LEJEUNE-DIRICHLET (V). — *Vorlesungen über Zahlentheorie*, 3e éd. 1879.
SCHRÖDER E. (I). — *Lehrbuch der Arithmetik und Algebra*, 1er Band. Leipzig, 1873.
ZERMELO (II). — V. ci-après, § 2-3.

B. — Les ensembles finis.

CAVAILLÈS J. (I). — *Sur la 2e définition des ensembles finis donnée par Dedekind. Fund. Mathematicae*, t. 19 (1932), p. 143.
DEDEKIND R. (IV). — *Zweite Definition des Endlichen und Unendlichen*, opuscule inachevé publié *in* Dedekind (V).
VON NEUMANN J. (IV). — *Die Axiomatisierung der Mengenlehre. Math. Zeitschr.*, t. 27 (1928), p. 669.
TARSKI A. (II). — *Sur les ensembles finis. Fund. Mathematicae*, t. 6 (1925), p. 45.
WEBER H. (I). — *Arithmetik, Algebra, Analysis* in Weber-Wellstein *Enzyklopädie der Elementarmathematik*, t. 1, 4e éd. Leipzig et Berlin, 1922.
WHITEHEAD and RUSSELL (I). — *Principia Mathematica*, Cambridge, 1re éd., t. I (1910), t. II (1912), t. III (1913), 2e éd., t. I (1925); t. II, III non modifiés (1927).
ZERMELO E. (IV). — *Sur les ensembles finis et le principe d'induction complète Acta Mathematica*, t. 32 (1909), p. 185.

§ 2. — *L'axiome de choix et le théorème du bon ordre.*

BERNSTEIN F. (II). — *Über den Beweis der Wohlordnung Math. Ann.*, t. 60 (1905), p. 463.
— (III) *Untersuchungen aus der Mengenlehre*, ibid., t. 61 (1905), p. 117-165 (Reproduction de : Inaugural Dissertation Göttingen, 1901).
BOREL E. (II). — *A propos d'un théorème de M. Zermelo. Math. Ann.*, t. 60 (1905).
FRAENKEL A. (III). — *Über den Begriff « definit » und die Unabhängigkeit des Auswahlaxioms.* V. ci-après § 3-5.
— (XIII). — *Sur l'axiome du choix, L'Enseignement math.*, 1935, p. 32.
— (XIV) *Uber eine abgeschwächte Fassung des Auswahlaxioms.* V. ci-après § 3-5.
KÖNIG J. (II). — *Über die Grundlagen der Mengenlehre und das Kontinuumproblem* I : *Math. Ann.*, t. 61 (1905), p. 156-160. II : *ibid.*, t. 63 (1907), p. 217, trad. in *Acta Mathematica*, t. 30 (1906), p. 329 et t. 31 (1908), p. 89.
LÉVI B. (I). — *Intorno alla teoria degli aggregati. R. Instituto Lombardo di scienze e lettere Rendic.* (2), t. 35 (1902), p. 863.
TARSKI (I). — *Sur quelques théorèmes qui équivalent à l'axiome de choix. Fund Mathematicae*, t. 5 (1924, p. 47.
ZERMELO E. (I). — *Beweis dass jede Menge wohlgeordnet werden kann. Math. Ann.*, t. 59 (1904), p. 514-516.
— (II) *Neuer Beweis für die Wohlordnung*, ibid., t. 65 (1908), p. 107-128.

§ 3-5. · ***L'Antinomie Russell et l'axiomatisation Zermelo-Fraenkel.***

BAER R. (I). — *Zur Axiomatik der Kardinalzahlarithmetik. Math. Zeitschr.* t. 29 (1929).

FRAENKEL (I). — *Axiomatische Begründung der transfiniten Kardinalzahlen*, I. *Math. Zeitschrift*, t. 13 (1922), p. 153-188.

— (II). — *Zu den Grundlagen der Cantor-ZermeloschenMengenlehre Math. Annalen*, t. 86 (1922), p. 230-237.

— (III). — *Über den Begriff « definit » und die Unabhängigkeit des Auswahlaxioms. Sitzungsb. d. Preuss. Akad. d. Wiss. Phys.-math. Kl.* 1922, p. 253-257.

— (IV). — *Die neueren Ideen zur Grundlegung der Analysis und Mengenlehre. Jahresb. d. dtsch. Math. Verein*, t. 33 (1924), p. 97-103.

— (V). — *Untersuchungen über die Grundlagen der Mengenlehre Math. Zeitschrift* t. 22 (1925), p. 250-273.

— (VI). — *Axiomatische Theorie der Geordneten Mengen Journ. f. Mat.* t. 155 (1926), p. 129-158.

— (VII). — *Zehn Vorlesungen über die Grundlegung der Mengenlehre.* Leipzig u. Berlin Teubner 1927.

— (VIII). — *Über die Gleichheitsbeziehung in der Mengenlehre. Journ. f. Math.* t. 157 (1927), p. 79-81.

— (IX). — *Über die Ordnungsfähigkeit beliebiger Mengen. Sitzungsb. d. Preuss. Akad. d. Wiss. Phys. math. Kl.* 1928.

— (XII). — *Axiomatische Theorie der Wohlordnung. Journ. f. Math.*, t. 167 (1932), p. 1-11.

— (XIII). — *Sur l'axiome du choix.* Conférence faite à Genève en 1934. *L'enseignement mathématique*, 1935, p. 32-51.

— (XIV). — *Über eine abgeschwächte Fassung des Auswahlaxioms. The Journal of symbolic Logic*, t. 2 (1937), p. 1-25.

HESSENBERG G. (I). — *Grundbegriffe der Mengenlehre* Sonderdruck aus den *Abhandl d. Friesschen Schule* Bd 1 Heft 4 Göttingen, 1906.

KURATOWSKI C. (I). — *Sur la notion de l'ordre dans la théorie des ensembles. Fund. Mathematicae* t. 2 (1921), p. 161-171.

MAHLO P. (I). — *Über lineare transfinite Mengen. Ber. über d. Verhandl. d. K. Sächs. Ges. d. Wiss. z. Leipzig Math. phys. Kl. t.* 63 (1911), p. 187-225.

— (II). — *Zur Theorie und Anwendung der ϱ_0-Zahlen*, ibid. t. 64 (1912), p. 108-112 et t. 65 (1913), p. 268-282.

VON NEUMANN J. (V). — *Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre* avec un complément de A. Fraenkel. *Math. Ann.* t. 99 (1928), p. 373-393.

RUSSELL B. (I). — *The principles of mathematics* I Cambridge 1903.

ZERMELO (I). — *Beweis dass jede Menge wohlgeordnet werden kann.* Math. Annalen, t. 59 (1904), p. 514-516.

— (II). — *Neuer Beweis für die Wohlordnung*, ibid., t. 65 (1908), p. 107-128.

— (III). — *Untersuchungen über die Grundlagen der Mengenlehre* I, ibid. t. 65 (1908), p 261-281.

§ 5. — *Les ordinaux transfinis.*

HESSENBERG G. (II) *Kettentheorie und Wohlordnung. Journ. f. r. u. angew. Math.* t. 135 (1909), p. 81-133.

VON NEUMANN J. (I). — *Zur Einführung der transfiniten Ordnungszahlen. Acta Univ. Franz. Jos. Szeged* 1 (1923), p. 199-208.

— (II). — *Eine Axiomatisierung der Mengenlehre. Journ. f. r. u. angew. Math.*, t. 154 (1925), p. 219-240.

— (III). — *Die Axiomatisierung der Mengenlehre. Math. Zeitschr.*, t. 27 (1928), p. 669-752.

— (IV). — *Über die Definition durch transfinite Induktion* etc. *Math. Ann.*, t. 99 (1928), p. 373.

— (V). — *Die Widerspruchsfreiheit eines Axiomensystems der Mengenlehre. Journ. f. r. u. angew. Math.* t. 160.

§ 6. — *L'absence de catégoricité, le paradoxe de Skolem.*

EKLUND H. (I). — *Über Mengen die Elemente ihrer selbst sind. Nyt Tidsskr. f. Mathematik.* Avd. 13, t. 29 (1918), p. 8-28.

MIRIMANOFF D. (I). — *Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles. L'enseignement math.*, t. 19 (1917), p. 37-52.

— (II). — *Remarques sur la théorie des ensembles et les antinomies cantoriennes*, I, ibid, t. 19 (1917), p. 209-217 ; II, ibid, t. 21 (1920), p. 29-52.

VON NEUMANN (II) et (V).

SKOLEM Th. (I). — *Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze. Skr. utg. av. Videnskaps. Kristiania* 1919, I, *Math. Naturv.* Kl., p. 1-37.

— (II). — *Einige Bemerkungen zur axiomatischen Begründung der Mengerlehre. 5 Kongress d. skandinav. Math. Helsingfors*, 1922, p. 217.

— (III) *Über einige Grundlagenfragen der Mathematik. Skr. Nor. Vid. Akad. in Oslo, Mat-Naturv. Kl.*, 1929, n° 4.

CONCLUSION

HARTOGS F. (I). — *Über das Problem der Wohlordnung. Math. Ann.*, t. 76 (1915), p. 438-443.

KÖNIG J. (I). — *Zum Kontinuumproblem. Math. Ann.*, t. 60 (1905), p, 177-180 et 462. 1er redact. in *Verh. d. 3. internat. Math. Kongress in Heidelberg*, (1904).

LUSIN N. (I). — *Sur les ensembles analytiques nuls. Fund. Math.*, 185 (1935), p. 109-131.

INDEX DES TERMES EMPLOYÉS (1)

A

Aleph (puissance d'un ensemble bien ordonné) : 81, 94.
Apantachie : 39.
Application : 51 n.
Axiomatique (méthode) : 18.
Axiome-N : 133.

B

Bien ordonné (ensemble) : 64.
Biunivoque (correspondance) : 46.
Bon ordre : 64.

C

Cardinal (nombre) : 75.
Catégoricité : 5 n.
Chaîne : 100.
Chaîne propre (d'un ensemble) : 101.
θ-Chaîne : 117.
Choix (axiome de) : 113.
Classes I, II, III des nombres ordinaux transfinis : 67.
Compacité en soi : 52.
Comparabilité des puissances : 66.
Consistante (multiplicité) : 95.
Continuité : 16.
Coupure : 16.
Critiques (nombres) : 91.
Croissance (ordre de) : 40.
Cyclique (ordre) : 111.

D

Définie (propriété) (Zermelo) : 123.
Dénombrable (ensemble) : qui peut être mis en correspondance biunivoque avec l'ensemble des entiers : 50.
Dense (ensemble) dans un intervalle : 58.
Dense en soi (ensemble) : 69, 85.
Dense (partout non) : 58.
Densité : 52.
Dérivation (d'un ensemble) : 22.
Diagonale (procédé de la) : 42, 72.
Distributive (procédé de la propriété) : 12, 70.

E

Emboîtement des intervalles (procédé de l') : 52.
Equivalence (théorème de l') : 78, 109 n.
Equivalents (ensembles) : 55, 75.
Espèce (d'un ensemble de points) : 22.
Espèce d'un nombre ordinal (1re ou 2e espèce) : 88.
Exponentiation d'un ensemble par un autre : 77.

F

Fermé (ensemble) : 71, 85.
Fusionnement (procédé de) : 106.

G

Genre (d'un ensemble) : 58.
Groupe intégrable : 58.

I

Intersection de 2 ou un nombre quelconque d'ensembles : 100 n.
Infinitaire (calcul) : 40.
— (ordre) : 41.
Irréductible (élément) d'un ensemble : 113.
Isolé (ensemble) : 60.

L

Linéarisation : 51.

N

Nombrabilité (d'un ensemble) : 131
Nombrement : 131.

(1) Les pages indiquées sont celles où le terme est défini ou (en l'absence d'une définition explicite, qui est alors donnée ici) utilisé pour la première fois.

Nombres-ε : 92.
Normales (Fonctions) : 91.

O

Ordinal (nombre) : 64.

P

$\mathfrak{P}M$: ensemble des parties de l'ensemble M : 119.
Pantachie : 39.
Parfait (ensemble) : 69, 85.
Partie d'un ensemble : 55.
Principal (élément), d'un ensemble orddnné : 84.
Principes d'engendrement, d'arrêt : 65, 67.
Produit d'ensembles : 77.
Propriété distributive ou propriété Γ (procédé de la) : 12, 70.
Puissance d'un ensemble : 55, 75.

Q

Quasi-Chaîne : 111.

R

Récurrence (définition par) : finie : 105.
Réductible (ensemble) : dont le $\nu^{ième}$ dérivé (ν étant un nombre fini ou un ordinal transfini de la classe II) est fini : 68.
Reste d'un ensemble ordonné (par rapport à un segment) : 85.
Réunion (ensemble) : 77.

S

$\mathfrak{S}M$ (ensemble-réunion des ensembles éléments de M) : 122.
Saturé (élément) d'un ensemble : 113.
Segment, dans un ensemble ordonné : 85.
Semblable (application) : 82.
Séparation (axiome de) : 122.
Sous ensemble : V. partie d'ensemble.
Subsomption (ordre de) : 130.
Substitution (axiome de) : 127.

T

Type d'ordre : 81.

CORRESPONDANCE CANTOR-DEDEKIND

AVERTISSEMENT

En 1872 Cantor faisait, au hasard d'un voyage en Suisse, la connaissance de Dedekind. De cette rencontre date un échange de lettres qui, avec des interruptions, se poursuivit jusqu'en 1899 (fin de la période productive pour Cantor) et où s'éprouvèrent à peu près toutes les idées fondamentales de la théorie des ensembles.

On sait le destin par instants dramatique de Georg Cantor. Il est peu d'exemples, remarque Fraenkel [1], de vie aussi étroitement mêlée à une œuvre, de théorie, œuvre aussi exclusivement personnelle d'un seul chercheur. Dans une lettre qu'il écrivait à 17 ans, Cantor parle d' « une voix inconnue et mystérieuse » qui, contre la volonté des siens, l'appelle aux mathématiques [2]. En 1872, son collègue de Halle, Heine, l'incite à l'étude des séries trigonométriques. Ce sont alors en quelques années ces découvertes étonnantes — étonnantes pour leur auteur le premier : « je le vois mais ne le crois point » écrivait-il en 1877 à Dedekind à propos de l'une d'elles —, ces notions radicalement nouvelles, puissance des ensembles abstraits, débuts de topologie, arithmétique des nombres transfinis, édifice dont la hardiesse et la beauté faisaient dire à Hilbert : « qu'il représente une des plus belles créations de l'esprit mathématique » mais dont la nouveauté suscita dès le début défiance, chez Weierstrass, hostilité déclarée chez Kronecker.

(1) A. Fraenkel, *Georg. Cantor. Jahresb. d. Deutsch. Math. Verein.*, t. 39 (1930), p. 189-266, reprod. partiellement in *G. Cantor. Ges. Abhandlungen*, hrgg. v. Zermelo. Berlin, Springer, 1932, p. 452. Nous lui empruntons nombre de renseignements. C'est d'ailleurs A. Fraenkel qui nous signala le premier cette correspondance.

(2) *Ibid.*, p. 453.

En 1877 dans quelques lettres (que nous n'avons pas cru devoir reproduire ici) Cantor s'inquiète du retard apporté par le *Journal de Crelle* (dirigé par Kronecker) à publier sa démonstration sur l'identité de puissance des continus à nombre quelconque de dimensions. On y croyait voir — Cantor le premier au début — une atteinte à la notion classique de dimension. Inquiet, déjà sensible à l'atmosphère défavorable, il envisage de retirer son article, de le faire paraître en brochure séparée comme Dedekind pour *Stetigkeit u. irrationale Zahlen*. Cette fois l'alerte était vaine, mais ce fut la dernière collaboration de Cantor au *Journal de Crelle*. Les articles fondamentaux — où l'essentiel de la théorie était déjà développé — parurent dans les *Mathematische Annalen* de 1875 à 1883, au milieu d'une hostilité grandissante. Même Weierstrass, d'après le témoignage de Schoenflies (1), marquait de la froideur. Le seul appui vint de Mittag-Leffler qui ouvrit à Cantor les *Acta Mathematica* qu'il venait de fonder. On sait que, dès le tome II, y parut une traduction française des Mémoires des *Mathematische Annalen*, faite, raconte Mittag-Leffler (2), par le groupe de jeunes mathématiciens gravitant autour d'Hermite ; en particulier l'article *Quelques théorèmes sur les ensembles infinis et linéaires de points* aurait été traduit par Poincaré. C'était donc une reconnaissance partielle que confirma le succès de deux mémoires de Mittag-Leffler, *Sur la représentation analytique des fonctions d'une variable indépendante* et *Démonstration nouvelle du théorème de Laurent*, où les notions cantoriennes étaient pour la première fois utilisées. Mais l'irritabilité de Cantor ne faisait que s'accroître ; il savait que Kronecker le traitait de « corrupteur de la jeunesse », dénigrait publiquement ses travaux et lui interdisait l'accès de Berlin. Il s'en plaint, parfois avec une verve amusante — Kronecker est désigné par le sobriquet de « Méré », type de l'esprit étroit qui s'oppose à l'inattendu du progrès mathématique —, parfois avec un énervement maladif dans ses lettres à Mittag-Leffler autour de 1884 (3). C'est qu'à cette date apparaît l'autre aspect, le plus important du drame.

(1) A. Schoenflies, *Die Krisis in Cantors mathematischem Schaffen. Acta Math.* 50 (1928), p. 1-23.
(2) Mittag Leffler, *ibid.*, p. 25.
(3) Publiées par Schoenflies, *loc. cit.*

L'édifice des ensembles était bien à grands traits ébauché dans les mémoires des *Mathematische Annalen*, mais il manquait l'indispensable couronnement, le théorème sur le continu. « De ce qui précède, on peut à l'aide des théorèmes démontrés au n° 5 conclure que le continu linéaire a la puissance de la classe II. » Ainsi s'achevait en 1882 le dernier article *Sur les ensembles infinis linéaires de points.* Au paragraphe 5 avait été prouvé que la puissance de la classe II (des ordinaux transfinis dénotant un ensemble bien ordonné dénombrable) suivait immédiatement le dénombrable. D'autre part, dans la théorie des ensembles de points on n'avait eu affaire qu'aux deux puissances dénombrable et continu (ou puissance des ensembles parfaits) : en particulier tout ensemble fermé se décompose en ensemble parfait et ensemble dénombrable ; d'où la supposition, formulée dès 1878, que ce sont là les deux seules puissances possibles. Par la solution du problème se trouverait donc effectuée la liaison entre les deux branches de la théorie des ensembles, topologie et arithmétique transfinie. L'année 1884 fut celle de l'effort le plus intense sur ce point. Le 26 octobre Cantor triomphe dans une lettre à Mittag-Leffler : « Je suis maintenant en possession d'une démonstration extrêmement simple pour ce théorème le plus important de la théorie des ensembles (1). » Mais le 4 novembre : « heureusement je n'ai usé nulle part de cette affirmation comme moyen de démonstration... j'ai trouvé ces jours derniers une démonstration rigoureuse que le continu n'a pas la puissance de la classe II et bien plus qu'il n'a pas une puissance exprimable par un nombre ». Enfin le 15 désaveu de cette lettre — qui supposait que l'on puisse donner un exemple d'un ensemble fermé ayant la deuxième puissance — et la poursuite reprend de ce but toujours fuyant. La véritable souffrance des alternatives dans cette recherche vaine, l'isolement moral, les attaques des adversaires provoquent une crise de dépression nerveuse qui l'oblige à se réfugier un temps dans une clinique, qui le détache momentanément de son œuvre. « Ce n'est que l'idée... d'une application possible à la théorie des organismes vivants... qui m'a fait entreprendre ce travail fatigant et ingrat sur les ensembles de points. » Il demande au ministre prussien de l'instruction publique l'autorisation d'enseigner désormais la philosophie et ne s'occupe

(1) A. Schoenflies, *loc. cit.*

bientôt plus que de savoir si François Bacon est bien l'auteur des drames de Shakespeare. Crise dont les répercussions se prolongent une douzaine d'années pendant lesquelles il se considère surtout comme philosophe et cherche à l'infini actuel des cautions jusque dans le Moyen Age.

Cependant son œuvre véritable mûrissait : en 1897 parait le nouvel exposé des *Contributions à la fondation de la théorie des ensembles transfinis*, plus systématique, plus rigoureux avec un développement complet de l'arithmétique transfinie. Fraenkel y voit l'influence de l'esprit ordonné de Husserl qui débutait alors à Halle. Cette fin de siècle marque aussi la reconnaissance officielle de son œuvre : au Congrès de Zurich (1897) tour à tour Hadamard, Hurwitz, Hilbert lui rendent hommage, et la même année les *Leçons sur la théorie des fonctions* de Borel en présentent un exposé partiel et une utilisation désormais classique. En reprenant ses échanges avec Dedekind, Cantor chasse les mauvais fantômes. « Je voudrais que nous restions désormais en correspondance régulière... La question Bacon-Shakespeare me laisse maintenant complètement calme ; elle ne m'a fait perdre que trop de temps et d'argent. » Cependant la deuxième grande crise est déjà là : après l'impossibilité du problème du continu, voici la découverte des paradoxes. Dès 1895 — deux ans avant Burali-Forti — Cantor se heurte à l'antinomie de l'ensemble des nombres ordinaux : il essaie en 1899 de l'éviter par sa distinction entre ensemble consistant et ensemble inconsistant, par là aussi de rattraper le problème du continu. A quel point cette dernière question restait irritante pour lui, l'incident Koenig au Congrès de Heidelberg (1904) le montre assez : par un ingénieux théorème lié à tort avec un résultat de Bernstein mal interprété, Koenig semblait démontrer que la puissance du continu ne pouvait être un alef. Schonflies a raconté [1] l'émouvante scène où Cantor adjura ses amis plus jeunes de trouver la faute, de réfuter une affirmation qui ruinait son œuvre. Cette fois le succès ne fit pas défaut : la même année Zermelo, en démontrant la possibilité de bien ordonner le continu, rendait du même coup impossible des affirmations comme celle de Koenig. Mais le problème restait ; et en 1905 le paradoxe de Russell menaçait cette fois les bases même

[1] A. Schonflies, *Georg Cantor* in Mitteldeutsche Lebensbilder III (1928), p. 548-563.

de la théorie abstraite des ensembles. On sait que Dedekind en fut si touché qu'il refusa longtemps une réédition de *Was sind u. was sollen die Zahlen* ?

Théorie solitaire, aujourd'hui encore, dans sa partie abstraite — celle à quoi tenait le plus Cantor — aussi inachevée, aussi incertaine aux yeux de beaucoup, bras tendu vers le ciel dans une hardiesse qui lui vaut la complaisance des hilbertiens, l'hostilité des empiristes. Mais Cantor ne s'y attachait pas comme à son œuvre propre. Contre Kronecker, « réclamant l'impartialité pour ses élucubrations..., pour mon travail c'est la partialité que j'exige, non partialité pour ma personne terrestre, mais partialité pour la vérité qui est éternelle » (1). Si « l'essence de la mathématique est la liberté » le mathématicien ne fait que découvrir un monde, immanent à son esprit, résidant en soi dans une objectivité absolue : « je ne suis quant au contenu de mes travaux que rédacteur et fonctionnaire » (2).

Ce sont les premières étapes de la découverte de ce monde que suit la correspondance avec Dedekind entre 1872 et 1883. Fraenkel y voit l'opposition entre le mathématicien de « Sturm und Drang » et le classique épris de symétrie. D'un côté, fougue créatrice, quelquefois hasardeuse, de l'autre, rigueur et critique. En effet, dans ces lettres, les uns après les autres, tous les résultats pressentis établis par Cantor, sont d'abord soumis au jugement de Dedekind : celui-ci relève des démonstrations erronées (à propos de l'invariance de puissance des continus), en simplifie d'autres, précise enfin leur portée (comme pour le nombre des dimensions à propos de quoi il fait remarquer que c'est la bi-continuité de la mise en correspondance qui doit garantir l'invariance de la dimension). Mais il n'y a pas seulement une épreuve — à laquelle Cantor attachait tant d'importance qu'après son résultat sur la puissance des continus de dimensions quelconques, il ne déclare s'y fier, quoique sûr de la démonstration, qu'après avis de Dedekind. Outre l'autorité de son œuvre déjà importante à cette date, Dedekind apportait dans cet échange l'expérience de réflexions antérieures sur le même sujet : non seulement la définition de la continuité, donnée dans *Stetigkeit u. irrationale*

(1) L. à Mittag Leffler, 26 janvier 1884 in A. Schonflies, *loc. cit.*, *Acta Math. etc...*

(2) « Nur Berichterstatter und Beamter », *ibid.*

Zahlen, mais les ébauches restées inédites de cette *Systemenlehre*, à la fois topologie et théorie des ensembles abstraits. A ses travaux sur la première il fait allusion au détour d'une lettre ; Mlle Noether a pu en retrouver quelques-uns pour l'édition des *Œuvres complètes*. L'essentiel des réflexions sur la seconde se trouve dans les esquisses successives de *Was sind und was sollen die Zahlen* ? (de 1872 à 1887). — Esprit secret défiant de soi, « Treppenverstand » comme il se nomme, Dedekind mûrissait longuement ses idées les plus hardies, gardant au besoin des années dans son tiroir les résultats déjà obtenus. En 1899, à propos de la distinction proposée par Cantor entre ensembles consistants et inconsistants (destinée à neutraliser le paradoxe de Burali-Forti), pour laquelle il « avoue n'être pas encore au clair », la mention du théorème de l'équivalence lui fait évoquer un souvenir : « lorsque le jeune Felix Bernstein me rendit visite à Harzburg pour la Pentecôte 1897, il me parla du théorème B, p. 7, de la traduction Marotte [le théorème de l'équivalence que Bernstein venait de démontrer] et fut un peu saisi lorsque je lui dis ma conviction que ce théorème était *facile* à démontrer avec mes moyens (was sind und was sollen die Zahlen ?) ; cependant, l'entretien n'alla pas plus avant sur sa ou ma démonstration. Après son départ je m'y attelai et construisis la démonstration ci-jointe ». En 1932, prenant connaissance de cette correspondance, Zermelo eut la surprise d'y trouver exactement, aux termes près, la démonstration que lui-même découvrit et publia en 1908. Elle a sur celle de Bernstein (dont Schröder avait donné une ébauche incomplète un an avant), l'avantage de ne pas faire intervenir explicitement la suite des nombres entiers. Le curieux est que Dedekind avait déjà spontanément rédigé cette démonstration dix ans avant la visite de Bernstein : nous avons pu en effet retrouver dans ses manucrits une feuille datée par lui 1887-7-11, qui donne, en notations cette fois de *Was sind und was sollen die Zahlen* (1) le même raisonnement. Preuve nouvelle de la fécondition de la notion de chaîne à l'aide de laquelle en 1907 Zermelo devait démontrer le théorème du bon ordre.

La présente édition était prête il y a quatre ans. Retardée par diverses circonstances, elle paraît aujourd'hui exactement

(1) Reproduite in *Dedekind Gesammelte Werke* Braunschweig Vieweg u. Sohn. 1932. t. III.

telle que l'avait revue avec nous M^lle Noether, souvenir vers ces journées de Göttingen où il nous avait été donné après bien d'autres de connaître la bonté joyeuse de son accueil, l'intense rayonnement de son esprit. Je tiens à remercier les professeurs Hasse, directeur de l'Institut mathématique de Göttingen, et Sperner, éditeur du *Jahresbericht d. deutschen Mathematiker Vereinigung* pour l'amabilité avec laquelle ils m'ont accordé les autorisations nécessaires.

J. C.

NOTE DU TRADUCTEUR

La traduction en français de la « Correspondance Cantor-Dedekind » présentait cette difficulté que la plupart des concepts mathématiques qui en font l'objet étaient encore en formation à cette époque, et que les termes employés n'ont donc pas toujours aujourd'hui le même sens. On a donc traduit le texte aussi littéralement que possible, sauf lorsque la traduction française littérale a une signification différente, ou a reçu depuis une acceptation différente en mathématiques ; et, chaque fois que cela a paru utile, on a ajouté en note l'expression employée dans le texte allemand.

La « Correspondance » est traduite de l'édition publiée par E. Noether et J. Cavaillès chez Hermann en 1937, à l'exception des lettres de l'année 1899, que E. Zermelo avait déjà publiée dans son édition des *Œuvres complètes* de Cantor, d'où nous les avons traduites.

Comme le signalaient les éditeurs de 1937 : « Une grande partie de ces lettres est passée presque sans modification dans les publications ultérieures. La correspondance provient de l'héritage de Dedekind ; les lettres de Dedekind à Cantor sont des brouillons. »

Les notes des auteurs des lettres sont indiquées par des numéros. Celles des éditeurs par des numéros, et placées entre [ ]. Celles de Zermelo sont indiquées par des lettres capitales. Celles du traducteur par des lettres minuscules.

CANTOR A DEDEKIND

Halle sur la S., le 28 avril 1872

Je vous remercie très vivement pour l'envoi de votre traité sur la continuité et les nombres irrationnels. Comme j'avais déjà pu m'en convaincre, le point de vue auquel je suis arrivé sur ce sujet depuis quelques années, à partir de préoccupations d'arithmétique, coïncide en fait avec vos conceptions : il n'y a de différence que dans *l'introduction conceptuelle* des grandeurs numériques. Que vous ayiez bien mis en évidence ce qui constitue l'essence de la continuité, j'en suis absolument convaincu.

Halle, le 29 novembre 1873

Permettez-moi [1] de vous soumettre une question qui a pour moi un certain intérêt théorique, mais à laquelle je ne puis répondre ; peut-être le pourrez-vous, et serez-vous assez bon pour m'écrire à ce sujet. Voici de quoi il s'agit.

Prenons l'ensemble [a] de tous les individus entiers positifs n, et représentons-le par (n) ; puis considérons l'ensemble de toutes les grandeurs numériques réelles positives x, et représentons-le par (x) ; la question est simplement de savoir si (n) peut être mis en correspondance avec (x) de telle manière qu'à chaque individu d'un des ensembles corresponde un individu et un seul de

[1] [Pour les lettres du 29 novembre 1873 au 27 décembre 1873, cf. p. 11 les remarques de Dedekind, visiblement écrites plus tard, qui les résument.]

[a] *Ensemble* : en allemand *Inbegriff*, terme employé exclusivement par Cantor comme d'ailleurs par Dedekind dans les lettres de cette période. *Inbegriff* dérive de *Begriff*, concept : un ensemble est un système d'objets répondant à un concept.

l'autre. A première vue, on se dit que ce n'est pas possible, car (n) est composé de parties discrètes, tandis que (x) forme un continu ; mais il n'y a rien à gagner avec cette objection, et, si fort que j'incline à penser qu'il n'y a pas de correspondance univoque entre (x) et (n), je ne peux pourtant pas en trouver la raison et c'est d'elle que je me préoccupe — peut-être est-elle très simple.

Ne serait-on pas aussi tenté de conclure au premier abord que (n) ne peut être mis en correspondance univoque avec l'ensemble $\left(\frac{p}{q}\right)$ de tous les nombres rationnels $\frac{p}{q}$? et pourtant, il n'est pas difficile de montrer que (n) peut être mis en correspondance univoque, non seulement avec cet ensemble, mais aussi avec l'ensemble plus général :

$$(a_{n_1, n_2, \ldots, n_\nu})$$

n_1, n_2, ..., n_ν étant des indices entiers positifs illimités en nombre quelconque ν.

Halle, le 2 décembre 1873

J'ai été particulièrement heureux de recevoir aujourd'hui votre réponse à ma dernière lettre. Si je vous ai soumis ma question, c'est parce que je me suis posé celle-ci il y a déja plusieurs années et je me suis toujours demandé si la difficulté que j'y trouvais était de nature subjective, ou bien si elle tenait au problème lui-même. Puisque vous me dites être, vous aussi, hors d'état d'y répondre, je peux supposer que c'est la deuxième éventualité qui est exacte. — Du reste, je voudrais ajouter que je ne m'en suis jamais occupé sérieusement, parce qu'elle n'a pas d'intérêt pratique particulier pour moi, et je suis bien de votre avis quand vous dites que, pour cette raison, elle ne mérite pas qu'on y consacre beaucoup de peine. Ce serait pourtant beau, si l'on pouvait y répondre ; si, par exemple, la réponse était *négative*, on disposerait par là même d'une nouvelle démonstration du théorème de Liouville affirmant l'existence de nombres transcendants.

Votre démonstration du fait que (n) peut être mis en correspondance univoque avec le corps des nombres algébriques est à

peu près la même que celle par laquelle je prouve mon affirmation contenue dans ma dernière lettre. Je pose $n_1^2 + n_2^2 + \ldots + n_\nu^2 = \mathfrak{M}$ et j'ordonne ensuite les éléments.

N'est-il pas en soi bon et commode que, comme vous l'avez si bien fait ressortir, l'on puisse parler du n^{ieme} nombre algébrique, de telle manière que chacun d'eux figure une fois dans la suite ?

Comme vous le notez très justement, notre question peut se formuler comme suit : « Peut-on mettre (n) en correspondance univoque avec un ensemble

$$(a_{n_1, n_2, \ldots}),$$

$n_1, n_2, \ldots$ étant des indices entiers positifs illimités, en nombre infini ? »

Halle, le 7 décembre 1873

Ces derniers jours, j'ai eu le temps d'étudier, d'une façon un peu plus suivie, ma conjoncture dont je vous avais parlé ; c'est seulement aujourd'hui que j'en ai terminé, me semble-t-il, avec cette affaire ; si je devais pourtant me tromper, je ne trouverais certainement pas de juge plus indulgent que vous. Je prends donc la liberté de soumettre à votre jugement ce que j'ai couché sur le papier, dans toute l'imperfection de ce premier jet.

Supposons que l'on puisse ranger tous les nombres positifs $\omega < 1$ en une suite :

(I) $$\omega_1, \omega_2, \omega_3, \ldots, \omega_n, \ldots$$

Après ω_1 soit ω_α le premier terme plus grand, après celui-ci soit ω_β le premier terme plus grand, et ainsi de suite. Posons : $\omega_1 = \omega_1^1$, $\omega_\alpha = \omega_1^2$, $\omega_\beta = \omega_1^3$, etc..., et extrayons de (I) la suite infinie [a] :

$$\omega_1^1, \omega_1^2, \omega_1^3, \ldots, \omega_1^n, \ldots$$

Dans la suite restante, soit ω_2^1 le premier terme, ω_2^2 le pre-

[a] Le terme allemand que nous traduisons par « *suite* » est « *Reihe* », sauf mention expresse du contraire (actuellement, on dit « *Reihe* » pour « *série* », et « *Folge* » pour « *suite* »).

mier terme suivant ω_2^1 qui soit plus grand, etc..., et extrayons la 2e suite :

$$\omega_2^1,\ \omega_2^2,\ \omega_2^3,\ \ldots,\ \omega_2^n,\ \ldots$$

En continuant ainsi, on reconnaît que la suite (I) peut être décomposée en une infinité de suites :

(1) $\omega_1^1,\ \omega_1^2,\ \omega_1^3,\ \ldots,\ \omega_1^n,\ \ldots$

(2) $\omega_2^1,\ \omega_2^2,\ \omega_2^3,\ \ldots,\ \omega_2^n,\ \ldots$

(3) $\omega_3^1,\ \omega_3^2,\ \omega_3^3,\ \ldots,\ \omega_3^n,\ \ldots$

mais dans chacune d'elles les termes vont en croissant de gauche à droite ; on a :

$$\omega_k^\lambda < \omega_k^{\lambda+1}.$$

Choisissons maintenant un intervalle $(p \ldots q)$ de telle sorte qu'aucun élément de la suite (1) ne s'y trouve contenu ; on peut, par exemple, prendre $(p \ldots q)$ à l'intérieur de l'intervalle $(\omega_1^1 \ldots \omega_1^2)$; il se peut alors que tous les termes de la deuxième suite, ou de la troisième, se trouvent à l'extérieur de $(p \ldots q)$; mais il faut qu'il y ait une première suite, disons la $k^{\text{ième}}$, dont tous les termes ne se trouvent pas à l'extérieur de $(p \ldots q)$; (car sinon, les nombres qui se trouvent à l'intérieur de $(p \ldots q)$ ne seraient pas contenus dans (I), contrairement à l'hypothèse) ; on peut alors déterminer, à l'intérieur de $(p \ldots q)$, un intervalle $(p' \ldots q')$ tel que les termes de la $k^{\text{ième}}$ suite soient tous à l'extérieur de celui-ci ; $(p'' \ldots q')$ se comporte évidemment de la même manière relativement aux suites précédentes ; mais, parmi les suivantes, on arrivera à une $k'^{\text{ième}}$ suite dont les termes ne sont pas tous à l'extérieur de $(p' \ldots q')$, et l'on choisit alors, à l'intérieur de $(p' \ldots q')$, un troisième intervalle $(p'' \ldots q'')$ tel que tous les éléments de la $k'^{\text{ième}}$ suite soient à l'extérieur de cet intervalle.

On voit ainsi qu'il est possible de former une suite infinie d'intervalles :

$$(p \ldots q),\quad (p' \ldots q'),\quad (p'' \ldots q''),\ \ldots$$

dont chacun contient les suivants ; ces intervalles relativement à nos suites (1), (2), (3), ..., se comportent comme suit :

Les termes de la 1re, de la 2e, ..., de la $k - 1^{\text{ième}}$ suite sont à l'extérieur de $(p \ldots q)$.

Les termes de la k^{ieme}, ..., de la $k' - 1^{\text{ieme}}$ suite sont à l'extérieur de $(p' \ldots q')$.

Les termes de la k'^{ieme}, ..., de la $k'' - 1^{\text{ieme}}$ suite sont à l'extérieur de $(p'' \ldots q'')$.

Mais il y a toujours *au moins* un nombre, que j'appellerai η, qui se trouve à l'intérieur de tous ces intervalles ; on voit immédiatement que ce nombre η, qui est évidemment $\begin{matrix} > 0 \\ < 1 \end{matrix}$, ne peut appartenir à aucune de nos suites (1), (2), ..., (*n*), ... Ainsi, en partant de l'hypothèse que tous les nombres $\begin{matrix} > 0 \\ < 1 \end{matrix}$ sont contenus dans (I), on arriverait à ce résultat opposé qu'un nombre déterminé $\eta \begin{matrix} > 0 \\ < 1 \end{matrix}$ n'appartient *pas* à (I) : par conséquent, cette hypothèse est incorrecte.

Je crois donc être finalement arrivé à la raison pour laquelle l'ensemble appelé (*x*) dans mes précédentes lettres *ne* peut *pas* être mis en correspondance univoque avec celui désigné par (*n*).

Halle, le 9 décembre 1873

J'ai déja trouvé, pour le théorème démontré dernièrement, une démonstration simplifiée, pour laquelle on n'a plus besoin de décomposer la suite (I) en (1), (2), (3).

Je montre directement que, en partant d'une suite

$$\omega_1, \omega_2, \ldots, \omega_n, \ldots,$$

je peux déterminer, dans *tout* intervalle donné $(\alpha \ldots \beta)$ un nombre η n'appartenant pas à (I). Cela suffit pour en déduire que l'ensemble (*x*) ne peut pas être mis en correspondance univoque avec l'ensemble (*n*), et j'en conclus qu'il y a, entre les systèmes et ensembles de valeurs [a], des différences essentielles, dont je ne savais pas tout récemment encore sonder les causes.

Je vous prie maintenant de m'excuser d'avoir tellement pris de votre temps pour cette question.

[a] *Systèmes et ensembles de valeurs* : en allemand *Inbegriffe und Werthmengen*.

Halle, le 10 décembre 1873

En vous accusant réception de vos lignes amicales du 8 courant, je tiens à vous assurer que rien ne peut me réjouir plus que l'intérêt que j'ai eu le bonheur d'éveiller chez vous pour certaines question de l'analyse ; permettez-moi d'ajouter que rien ne peut aussi m'inciter davantage à poursuivre mes efforts, et laissez-moi vous prier de continuer à l'avenir à me faire parvenir vos remarques. La mise sous la forme $\omega_1, \omega_2, \ldots, \omega_\nu, \ldots$ ne devrait-elle pas donner de bons résultats pour votre théorie des nombres algébriques ? (1).

Berlin, le 25 décembre 1873

Bien que je n'aie pas eu l'intention de publier pour le moment la question que j'ai récemment débattue pour la première fois avec vous, j'ai été amené tout-à-coup à le faire. J'ai en effet communiqué mes résultats le 22 à M. Weierstrass ; mais le temps manquait alors pour en parler d'une façon précise ; j'ai eu, dès le 23, la grande joie de recevoir sa visite, ce qui m'a permis de lui communiquer les démonstrations ; il a été d'avis qu'il fallait que je publie cette affaire, pour autant qu'elle concerne les nombres algébriques. J'ai donc écrit un petit article sous le titre : *Sur une propriété de l'ensemble de tous les nombres algébriques réels* (a), et l'ai adressé à M. le professeur Borchardt pour qu'il en considère la publication dans le *Journal f. Math.*

Pour la rédaction de cet article vos remarques et votre mode d'expression m'ont été, comme vous le verrez, très utiles. C'est ce que j'ai voulu vous faire savoir.

Berlin, le 27 décembre 1873

Avec l'habitude prise, ces derniers temps, vous écrire m'est devenu si naturel que mes réponses me viennent à la plume,

(1) [En fait, dans son travail, *Sur les permutations du corps de tous les nombres algébriques* (*Œuvres complètes*, vol. II, p. 278), Dedekind a appliqué ce théorème, cependant que, pour les corps finis de nombres algébriques, l'existence de la base finie permet de se passer du bon ordre.]

(a) *Ueber eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen.*

sans que j'aie beaucoup à y réfléchir ; et j'en oublie presque d'en motiver et excuser la fréquence ; peut-être cela tient-il à la similitude de nos intérêts et à ce que nous prenons également à cœur le développement de la science au profit de tous.

Si j'ai donné, pour la publication, un caractère restreint à mon exposé, cela est en partie motivé par les conditions régnant ici, et dont je vous parlerai peut-être un jour de vive voix ; mais, d'autre part, je crois qu'il importe d'appliquer d'abord mon raisonnement dans un cas particulier (comme celui des nombres algébriques réels) ; les extensions dont il est susceptible — et je peux déjà en apercevoir une quantité — ne devraient pas demander ensuite beaucoup de peine ; que ce soit moi qui les donne ou quelqu'un d'autre, peu importe ; j'ai donc, après une courte introduction, rédigé deux paragraphes ; dans le premier il est démontré que l'ensemble des nombres algébriques réels peut être mis en correspondance univoque avec l'ensemble des nombres entiers positifs ; dans le second, que, étant donné une suite $\omega_1, \omega_2, \ldots, \omega_n, \ldots$, on peut définir dans tout intervalle des nombres η qui ne sont pas contenus dans cette suite.

NOTES DE DEDEKIND SUR LES LETTRES DE 1873

29.11.1873

Monsieur G. Cantor (Halle) me pose la question suivante : l'ensemble (n) de tous les entiers positifs (nombres naturels) peut-il être mis en correspondance avec l'ensemble (x) de toutes les grandeurs numériques réelles positives, de telle sorte que, à un individu d'un des ensembles corresponde un individu et un seul de l'autre ? — Il conclut en ces termes : « Ne serait-on pas aussi tenté de conclure au premier abord que (n) ne peut être mis en correspondance univoque avec l'ensemble $\left(\frac{b}{a}\right)$ de tous les nombres rationnels positifs $\frac{b}{a}$? et pourtant, il n'est pas difficile de montrer que (n) peut être mis en correspondance univoque,

non seulement avec cet ensemble, mais aussi avec l'ensemble plus général :

$$(a_{n_1, n_2, \ldots, n_\nu}),$$

n_1, n_2, ..., n_ν étant des indices entiers positifs illimités en nombre quelconque ν. »

J'ai répondu par retour du courrier que je ne savais pas décider de la première question, mais en même temps j'ai formulé et complètement démontré que même l'ensemble de tous les nombres algébriques peut être mis en correspondance de la manière indiquée avec l'ensemble (n) des nombres naturels (peu de temps après, ce théorème et sa démonstration ont été reproduits presque littéralement, y compris l'emploi du terme technique « *hauteur* », dans l'article de Cantor paru dans le *Journal de Crelle*, tome 77 [1]), mais à la différence près, maintenue malgré mon conseil, que seulement l'ensemble de tous les nombres algébriques *réels* est pris en considération). Mais l'opinion formulée par moi que la première question ne méritait pas qu'on y consacre beaucoup de peine parce qu'elle n'avait aucun intérêt pratique particulier a été réfutée de façon flagrante avec la démonstration qu'a donnée Cantor de l'existence de nombres transcendants (*Crelle*, tome 77).

2.12.1873

C. insiste sur l'importance de la première question, car, dans le cas d'une réponse négative, l'existence de nombres transcendants (Liouville) serait démontrée d'une autre manière. Il poursuit : « Vous démontrez que (n) peut être mis en correspondance univoque avec le corps des nombres algébriques de la même manière à peu près que je prouve ce que j'affirmais dans ma dernière lettre. Je pose $n_1^2 + n_2^2 + \ldots + n_\nu^2 = N$, et j'ordonne ensuite les éléments. N'est-il pas en soi bon et commode que, comme vous l'avez fait ressortir, l'on puisse parler du $n^{\text{ième}}$ nombre algébrique, de telle manière que chacun d'eux figure une fois dans la suite ? Comme vous le notez très justement, notre question peut se formuler comme suit : « Peut-on mettre (n) en correspondance univoque avec un ensemble :

$$(a_{n_1, n_2, \ldots}),$$

[1] [*Œuvres complètes*, p. 116.]

$n_1, n_2, \ldots$ étant des indices entiers positifs illimités, en nombre infini ? »

7.12.1873

Cantor me communique une démonstration rigoureuse, qu'il a trouvée le même jour, du théorème suivant lequel l'ensemble de tous les nombres positifs $\omega < 1$ *ne peut pas* être mis en correspondance univoque avec l'ensemble (n).

Je réponds le même jour à cette lettre, reçue le 8 décembre, en le félicitant pour ce beau succès, tout en « réfléchissant », sous une forme très simplifiée, la partie essentielle de sa démonstration (qui était encore bien compliquée) ; cette présentation a été transcrite, presque mot pour mot, dans l'article de Cantor (*Crelle*, tome 77) ; toutefois, l'expression employé par moi « d'après le principe de continuité » a été évitée là où elle se trouvait (p. 261, lignes 10-14) !

9.12.1873

Cantor m'écrit en hâte qu'il a trouvé une démonstration simplifiée du théorème. Comme il ne mentionne pas ma lettre, il ne doit l'avoir reçue que plus tard.

10.12.1873

Cantor accuse réception de ma lettre du 8 décembre, sans mentionner la présentation simplifiée de la démonstration qui s'y trouvait ; il me remercie pour l'intérêt que je porte à cette affaire.

25.12.1873

Cantor écrit (de Berlin) qu'il a rédigé (à l'incitation de Weierstrass) un petit article sous le titre : *Sur une propriété de l'ensemble de tous les nombres algébriques réels.* « Pour la rédaction de cet article, vos remarques et votre mode d'expression m'ont été, comme vous le verrez, très utiles. »

Je réponds, par retour du courrier, en lui conseillant d'abandonner la limitation au corps des nombres algébriques *réels.*

27.12.1873

Cantor écrit (de Berlin) : « Si j'ai donné, pour la publication, un caractère restreint à mon exposé, cela est en partie motivé par les conditions régnant ici, et dont je vous parlerai peut-être un jour de vive voix ; mais d'autre part, je crois qu'il importe d'appliquer d'abord mon raisonnement dans un cas particulier (comme celui des nombres algébriques réels). »

Je n'ai jamais eu d'éclaircissement sur les « conditions berlinoises » ; nous ne sommes pas non plus revenus sur l'article (*Crelle*, tome 77) par la suite.

CANTOR A DEDEKIND

Halle, le 5 janvier 1874

A propos des questions qui m'ont occupé ces derniers temps, je m'aperçois que, dans cet ordre d'idées, se présente aussi la suivante :

Est-ce qu'une surface (par exemple un carré, frontière comprise) peut être mise en relation univoque avec une courbe (par exemple un segment de droite, extrémités comprises), de telle sorte qu'à tout point de la surface corresponde un point de la courbe, et réciproquement à tout point de la courbe un point de la surface ?

Il me semble encore en ce moment que la réponse à cette question présente de grosses difficultés — bien que, ici aussi, l'on soit si fortement enclin à une réponse *négative*, qu'on pourrait presque tenir une démonstration pour superflue.

Halle, le 28 janvier 1874

Dans les comptes rendus de l'année dernière, je trouve un article d'Hermite où est exposée, avec l'art le plus grand dans le maniement de l'analyse, une méthode de développement simultané de n grandeurs exponentielles e^{ax}, e^{bx}, ..., e^{hx}, et, ce qui semble

le plus important, il fonde là-dessus une démonstration complètement sûre de la transcendance du nombre *e*. Hermite reconnaît s'être beaucoup occupé de la preuve de la transcendance de π, mais y renonce pour ce nombre, en observant qu'il se réjouirait beaucoup si un autre y réussissait.

Halle, le 18 mai 1874

Éprouvant le besoin de m'entretenir avec vous de sujets scientifiques, et d'entrer en relations personnelles plus proches avec vous, je souhaiterais vous rendre visite un jour à Brunswick cet été.

... Si vous voulez bien à l'occasion me répondre à ce propos, j'aimerais savoir si vous trouvez la même difficulté que moi à la question que je vous ai posée en janvier sur la correspondance entre une surface et une ligne, ou si je me suis laissé induire en erreur à ce sujet ; à Berlin, mon ami, à qui j'ai exposé la même difficulté, m'a déclaré que la chose était pour ainsi dire absurde, car il se comprend de soi-même que deux variables indépendantes ne peuvent être ramenées à une seule.

DEDEKIND A CANTOR

... L'objection de Monsieur I., que vous me communiquez, ne peut, si je la comprends bien, atteindre ma présentation ; le principe de continuité, exposé au § 3, p. 18 (1), ne doit naturellement être compris que comme le *complément* nécessaire des lois I, II, III énoncées au § 2, p. 15, et en II est dit exactement ce dont vous ou Monsieur I. semblez regretter l'absence ; l'objection n'aurait vraisemblablement pas été faite, si j'avais mis le numéro IV devant le principe de la p. 18, comme cela a été effectivement fait pour les lois analogues, au § 5, p. 25. Ou ai-je mal compris l'objection ? alors, veuillez me donner des éclaircissements.

Brunswick, le 11 mai 1877

(1) [Il s'agit naturellement de la première édition de *Continuité et nombres irrationnels*) (*Stetigkeit und irrationale Zahlen*).]

CANTOR A DEDEKIND

Halle, le 17 mai 1877

En vous remerciant vivement de votre réponse, je dois reconnaître qu'à la page 25 de votre écrit sur les nombres irrationnels vous énoncez, sous les numéros I, II, III et IV (1), des propriétés qui sont tout à fait caractéristiques pour le domaine de tous les nombres réels, de telle sorte qu'*aucun* autre système de valeurs de nombres réels ne puisse avoir comme celui-là toutes ces propriétés.

Cependant, permettez-moi, je vous prie, de faire cette remarque : l'accent que vous mettez, en divers endroits de votre travail, *expressément* sur la propriété IV, comme constituant l'*essence* de la continuité, ne peut-il pas donner lieu à des malentendus, qui ne pourraient pas se produire, à mon avis, au sujet de votre théorie, n'était cette prépondérance donnée à IV (comme en constituant l'essence propre). En particulier, vous dites dans l'avant-propos que l'axiome indiqué par moi est entièrement équivalent à celui que vous présentez au § 3 comme l'*essence* de la continuité. Mais vous comprenez par là cette même propriété qui, à la page 25, est donnée sous le numéro IV ; or, cette propriété appartient aussi au système de tous les nombres entiers, qui peut pourtant être considéré comme un prototype de la discontinuité.

Dans l'intérêt d'une affaire qui m'est devenue chère à moi aussi, je vous prie, si vous en trouvez le temps, d'examiner de plus près mes réserves.

P. S. : Je m'explique de la façon suivante pourquoi vous insistez particulièrement sur IV : c'est dans cette propriété que réside ce qui distingue le domaine complet des nombres du domaine des nombres rationnels ; et pourtant il me semble, pour les raisons dites plus haut, que l'on ne peut attribuer à la propriété IV le nom employé par vous : « Essence de la continuité ».

(1) [I : Ordre ; II : entre deux nombres, il y a une infinité de nombres différents ; III : coupure ; IV : à chaque coupure correspond un nombre.]

DEDEKIND A CANTOR

D'après votre dernière lettre, il me semble que nous courons le risque de discuter plus sur des mots que sur des choses. Tout lecteur attentif de mon travail comprendra certainement mon opinion sur la continuité ainsi : des domaines dont les éléments s'opposent et se complètent, comme cela est exprimé par les propriétés I et II dans les § 1, p. 14, § 2, p. 15, § 5, p. 25 (III est une conséquence de I, et n'est ajouté que pour préparer IV), ne sont pas encore nécessairement des domaines continus ; de tels domaines obtiennent la propriété de continuité par l'adjonction de la propriété IV (p. 18 sans numéro, et p. 25) et seulement par cette propriété. Et c'est pour cela que cette propriété est désignée comme l'essence de la continuité.

Vous me dites dans votre carte du 10 courant que ma définition de la continuité n'est pas complète, et me faites une proposition d'amélioration pour pallier ce défaut. En réponse, je rejette cette proposition, en attirant votre attention sur II, qui contient ce dont vous regrettez l'absence. Sur quoi, vous m'accordez dans votre dernière lettre que rien n'a en effet été omis dans ma définition ; si je dis par exemple : « Les domaines qui possèdent les propriétés I et II sont dits continus, s'ils possèdent aussi la propriété IV », vous ne contesterez nullement, si j'ai bien compris votre dernière lettre, le *caractère complet* d'une telle affirmation ; mais vous préféreriez, semble-t-il, que la propriété II passe de la proposition relative à la proposition conditionnelle : « Les domaines dont les éléments s'opposent conformément à I, sont dits continus s'ils possèdent en même temps les propriétés II *et* IV », et vous craignez que l'accent mis par moi exclusivement sur IV, comme étant *la* propriété dans laquelle s'exprime l'*essence* de la continuité, puisse conduire à des malentendus. Je ne partage pas cette crainte ; je suis fermement convaincu que tout lecteur attentif de mon travail comprendra que mon point de vue est celui que j'ai exprimé plus haut en commençant, et alors votre exemple du système de tous les nombres entiers rationnels ne donne naturellement lieu à aucune objection. En ce qui concerne la modification ci-dessus de la définition, je ne peux pas dire qu'elle me plaise, et je conclus de votre post-

scriptum que vous-même, si vous essayiez seulement de *remanier* mon travail dans ce sens, m'accorderiez certainement que ce travail, qui a pour sujet principal le progrès de l'arithmétique du rationnel à l'irrationnel, perdrait, si l'on émoussait ce qui constitue spécifiquement sa pointe, qui réside précisément dans la place prépondérante donnée à IV, car II est déjà présent dans le domaine rationnel, discontinu. Mais si quelqu'un préfère la définition transformée, je n'ai absolument rien à objecter contre la *légitimité* d'une telle méthode, surtout si celle-ci présente des avantages pour certaines recherches ; mais ma formulation originelle me plaît bien davantage, et je tiens pour plus approprié, en ce qui concerne l'essence de la continuité, de mettre l'accent exclusivement sur IV, et de discuter précédemment la propriété II, avant qu'il ne soit question de continuité ou de discontinuité. En tout cas, je conteste absolument la *nécessité* de cette transformation de la définition ; si c'est vraiment cela que l'on réclame, on pourrait aussi bien poser la question, dont je me suis aussi déjà occupé, s'il ne convient pas de faire passer aussi la propriété I, autant que possible, de la proposition relative à la conditionnelle. Cette question n'est absolument pas dénuée d'intérêt, mais cela me conduirait trop loin d'en parler maintenant. Je crois vraiment que nous ne différons d'opinion tout au plus que sur ce qui convient le mieux, et non sur ce qui est nécessaire, et qu'en conséquence, si nous poursuivions le débat, il n'en sortirait plus grand'chose.

Brunswick, le 18 mai 1877

CANTOR A DEDEKIND

Halle, le 20 juin 1877

En vous remerciant de votre lettre du 18 mai, avec laquelle je suis entièrement d'accord, reconnaissant que notre désaccord était uniquement superficiel, j'ai de nouveau une prière à vous adresser. Vous voyez que les intérêts théoriques qui nous lient ont le désavantage pour vous que je vous importune plus souvent que vous ne le voudriez peut-être.

Je voudrais savoir si vous estimez qu'une méthode de démonstration appliquée par moi est arithmétiquement rigoureuse ?

Il s'agit de montrer que les surfaces, les volumes et même les variétés [a] continues à ρ dimensions peuvent être mis en correspondance univoque avec des courbes continues, donc avec des variétés à *une seule* dimension, que les surfaces, les volumes, les variétés à ρ dimensions elles aussi ont donc la même *puissance* que les courbes ; cette opinion paraît opposée à celle généralement répandue, notamment parmi les représentants de la nouvelle géométrie, suivant laquelle on parle de variétés simplement, doublement, triplement, ... ρ fois infinies ; on se représente même parfois les choses comme si l'on obtenait l'infinité de points d'une surface en quelque sorte en élevant au carré, celle d'un cube en élévant au cube l'infinité des points d'une ligne.

Comme les variétés d'un nombre égal de dimensions peuvent être rapportées l'une à l'autre *analytiquement*, il me semble que l'on peut mettre les questions plus générales dont j'ai parlé sous la forme purement arithmétique suivante :

« Soient $x_1, x_2, \ldots, x_\rho$, ρ grandeurs variables réelles indépendantes, dont chacune peut prendre toutes les valeurs $\geqq 0$ et $\leqq 1$. Soit y une $\rho + 1^{\text{eme}}$ grandeur variable réelle avec le même domaine de variation [b] $\left(y \begin{matrix} \geqq 0 \\ \leqq 1 \end{matrix}\right)$.

Est-il alors possible de faire correspondre les ρ grandeurs $x_1, x_2, \ldots, x_\rho$ à la seule grandeur y de telle sorte qu'à chaque système de valeurs déterminé $(x_1, x_2, \ldots, x_\rho)$ corresponde une valeur déterminée y, et que, réciproquement, à chaque valeur déterminée y corresponde un système de valeurs déterminé et un seul $(x_1, x_2, \ldots, x_\rho)$? »

A ce qu'il me semble, il faut répondre *affirmativement* à cette question, *bien que j'aie tenu depuis des années le contraire pour exact*, et ce pour les raisons suivantes :

Tout nombre $x \begin{matrix} \geqq 0 \\ \leqq 1 \end{matrix}$ peut être représenté d'une manière et d'une seule sous la forme d'une fraction décimale illimitée, soit :

$$x = \alpha_1 \cdot \frac{1}{10} + \alpha_2 \cdot \frac{1}{10^2} + \ldots + \alpha_\nu \cdot \frac{1}{10^\nu} + \ldots$$

[a] *Variété* : en allemand *Gebilde*.
[b] *Domaine de variation* : en allemand *Spielraum*.

où les α_ν sont des nombres entiers $\geqslant 0$ et $\leqslant 9$. Tout nombre x détermine ainsi une suite infinie $\alpha_1, \alpha_2, \ldots$ et réciproquement.

Nous pouvons donc écrire :

$$x_1 = \alpha_{1,1} \cdot \frac{1}{10} + \alpha_{1,2} \cdot \frac{1}{10^2} + \ldots + \alpha_{1,\nu} \cdot \frac{1}{10^\nu} + \ldots$$

$$x_2 = \alpha_{2,1} \cdot \frac{1}{10} + \alpha_{2,2} \cdot \frac{1}{10^2} + \ldots + \alpha_{2,\nu} \cdot \frac{1}{10^\nu} + \ldots$$

$$x_\rho = \alpha_{\rho,1} \cdot \frac{1}{10} + \alpha_{\rho,2} \cdot \frac{1}{10^2} + \ldots + \alpha_{\rho,\nu} \cdot \frac{1}{10^\nu} + \ldots$$

De ces ρ nombres on peut déduire un $(\rho + 1)^{\text{eme}}$ nombre y :

$$y = \beta_1 \cdot \frac{1}{10} + \beta_2 \cdot \frac{1}{10^2} + \ldots + \beta_\nu \cdot \frac{1}{10^\nu} + \ldots$$

où l'on pose :

(I) $$\beta_{(n-1)\rho+1} = \alpha_{1,n} ; \quad \beta_{(n-1)\rho+2} = \alpha_{2,n} ; \ldots$$
$$\beta_{(n-1)\rho+\sigma} = \alpha_{\sigma,n} ; \ldots \beta_{(n-1)\rho+\rho} = \alpha_{\rho,n}.$$

Comme tout nombre entier positif ν peut se mettre d'une manière et d'une seule sous la forme :

$$\nu = (n - 1)\rho + \sigma, \quad \text{où} \quad \sigma \begin{matrix} > 0 \\ \leqq \rho \end{matrix},$$

on voit que la suite $\beta_1, \beta_2, \ldots$, et par conséquent aussi y, sont pleinement déterminés par les équations (I) ; mais réciproquement, si l'on part du nombre y, et par conséquent de la suite $\beta_1, \beta_2, \ldots$, les ρ suites

$$\begin{array}{l} \alpha_{1,1}, \quad \alpha_{1,2}, \ldots \\ \ldots\ldots\ldots\ldots \\ \alpha_{\sigma,1}, \quad \alpha_{\sigma,2}, \ldots \\ \ldots\ldots\ldots\ldots \\ \alpha_{\rho,1}, \quad \alpha_{\rho,2}, \ldots \end{array}$$

sont déterminées de façon *univoque* par les équations (I), et par conséquent aussi les ρ nombres $x_1, x_2, \ldots, x_\rho$.

DEDEKIND A CANTOR

La seule objection que je puisse trouver pour le moment à votre intéressante argumentation, et que vous écarterez peut-être sans difficulté, est la suivante. Vous dites : « Tout nombre x ($\geqslant 0$ et $\leqslant 1$) peut être représenté d'une manière et d'une seule sous la forme d'une fraction décimale *infinie*, de telle sorte que

$$x = \frac{\alpha_1}{10} + \frac{\alpha_2}{10^2} + \ldots + \frac{\alpha_\nu}{10^\nu} + \ldots$$

où les α_ν sont des nombres entiers $\geqslant 0$ et $\leqslant 9$. Chaque nombre x détermine ainsi une suite infinie $\alpha_1, \alpha_2, \ldots$, et réciproquement. » Le fait que vous ayiez souligné le mot « *infinie* » me laisse supposer que vous excluez le cas d'une fraction finie, c'est-à-dire telle qu'un α_ν différent de 0 ne soit suivi que par des chiffres $\alpha_{\nu+1} = \alpha_{\nu+2} = \text{etc.} = 0$, et que vous demandez d'écrire, au lieu de

$$x = \frac{\alpha_1}{10} + \frac{\alpha_2}{10^2} + \ldots + \frac{\alpha_\nu}{10^\nu} + \frac{0}{10^{\nu+1}} + \frac{0}{10^{\nu+2}} + \ldots + \frac{0}{10^{\nu+\nu'}} + \ldots$$

constamment

$$x = \frac{\alpha_1}{10} + \frac{\alpha_2}{10^2} + \ldots + \frac{\alpha_\nu - 1}{10^\nu} + \frac{9}{10^{\nu+1}} + \frac{9}{10^{\nu+2}} + \ldots + \frac{9}{10^{\nu+\nu'}} + \ldots$$

excluant ainsi toute possibilité d'une double représentation d'un seul et même nombre x (le nombre $x = 0$ lui-même serait néanmoins représenté sous la forme 0,0000 ... ; mais $x = \frac{3}{10}$ sous la forme 0,2999 ...).

Si c'est là votre opinion (— on pourrait naturellement aussi bien exclure le cas où, à partir d'un certain rang, ne se présente que le chiffre 9 ; mais les conséquences seraient similaires —),

alors mon objection (1) est la suivante. Je me limite pour la simplicité au cas $\rho = 2$ et pose :

$$x = \frac{\alpha_1}{10} + \frac{\alpha_2}{10^2} + \ldots = 0, \alpha_1\alpha_2 \ldots \alpha_\nu \ldots$$

$$y = \frac{\beta_1}{10} + \frac{\beta_2}{10^2} + \ldots = 0, \beta_1\beta_2 \ldots \beta_\nu \ldots$$

et je construis comme vous, à partir de ces deux nombres x, y, un troisième nombre

$$z = 0, \gamma_1\gamma_2\gamma_3 \ldots$$

où

$$\gamma_1 = \alpha_1, \qquad \gamma_2 = \beta_1, \qquad \gamma_3 = \alpha_2,$$
$$\gamma_4 = \beta_2, \ldots, \gamma_{2\nu-1} = \alpha_\nu, \qquad \gamma_{2\nu} = \beta_\nu, \ldots,$$

de telle sorte que z est une fonction bien déterminée des deux variables continues x et y, et contenue dans le même intervalle $(0 \leqslant \tau \leqslant 1)$. Mais il y a alors une infinité de vraies fractions auxquelles z ne sera *jamais* égal, par exemple :

$$0{,}478310507090\ \alpha_7\ 0\ \alpha_8\ 0\ \alpha_9\ 0\ \ldots\ \alpha_\nu\ 0\ \ldots$$

et de même toute fraction $0, \gamma_1\gamma_2\gamma_3 \ldots$, pour laquelle, à partir d'un endroit déterminé, soit $\gamma_{2\nu-1}$, soit $\gamma_{2\nu}$ est toujours $= 0$; car, si l'on voulait tirer inversement la valeur de x et de y d'une telle valeur de z, on arriverait à une valeur de x ou de y absente (exclue).

Je ne sais pas si mon objection est d'une importance essentielle pour votre idée, mais je n'ai pas voulu la garder pour moi.

Brunswick, le 22 juin 1877.

CANTOR A DEDEKIND

(Carte ; timbre de la poste du 23-6-77)

Vous avez malheureusement tout à fait raison avec votre objection ; heureusement, celle-ci n'atteint que la démonstration, et

(1) [Reproduite dans l'ouvrage de Cantor, *Contribution à la théorie des multiplicités* (*Ein Beitrag zur Mannigfaltkeitslehre*) § 7 (*Journal de Crelle*, 84) (*Œuvres complètes*, p. 130).]

non la chose elle-même ; en effet, je démontre *d'une certaine manière plus* que je ne voulais démontrer, puisque je mets un système $x_1, x_2, \ldots, x_\rho$ de variables réelles ($\geqq 0$ et $\leqq 1$) sans autre restriction en correspondance univoque avec une variable y, qui est contenue dans le même intervalle et y prend toutes les valeurs à l'exception de certaines d'entre elles y'' ; mais les valeurs y' qu'elle prend effectivement, elle ne les prend *qu'une seule fois*, et cela est, me semble-t-il, l'essentiel. Car je peux alors mettre y' en relation univoque avec une autre grandeur t qui prend toutes les valeurs $\geqq 0$ et $\leqq 1$.

Je suis seulement content que vous n'ayiez jusqu'à présent rien d'autre à objecter ; je me permettrai prochainement de vous écrire plus en détail sur ce sujet.

Halle, le 25 juin 1877

Sur une carte postale que je vous ai adressée avant-hier, j'ai reconnu la lacune découverte par vous dans ma démonstration, et remarqué en même temps que j'étais en état de la combler, bien que je ne puisse me retenir de regretter quelque peu que cette question ne se laisse pas régler sans faire intervenir des considérations plus compliquées ; mais cela tient sans doute à la nature même du sujet, il faut donc que je me console ; peut-être se trouvera-t-il plus tard que ce qui manque dans cette démonstration peut se traiter plus simplement qu'il ne serait en mon pouvoir de faire en ce moment. Mais comme je tiens avant tout à vous convaincre, si possible, de l'exactitude de mon théorème, savoir :

(A) « Une multiplicité continue à e dimensions (a) peut être mise en correspondance univoque avec une multiplicité continue à une dimension, ou (ce qui n'est qu'une autre forme du même théorème) les points (éléments) d'une multiplicité à ρ dimensions peuvent se déterminer par une coordonnée réelle t de telle sorte qu'à chaque valeur réelle de t dans l'intervalle (0 ... 1) corresponde un point de la multiplicité, mais aussi réciproquement, qu'à chaque point de la multiplicité corresponde une valeur déterminée de t dans l'intervalle (0 ... 1). »

(a) *Une multiplicité continue à e dimensions* : en allemand *Eine nach e Dimensionen ausgedehnte stetige Mannigfaltigkeit.*

je me permets de vous en soumettre une autre démonstration (1), sur laquelle j'étais même tombé plus tôt que sur celle-là.

Je pars du théorème suivant lequel tout nombre irrationnel $e \begin{matrix} > 0 \\ < 1 \end{matrix}$ peut être représenté d'une manière complètement déterminée par une fraction continue :

$$e = \cfrac{1}{\alpha_1 + \cfrac{1}{\alpha_2 + \ddots + \cfrac{1}{\alpha_\nu + \ddots}}} = (\alpha_1, \alpha_2, \ldots, \alpha_\nu, \ldots)$$

où α_ν est un nombre entier rationnel positif. A chaque nombre irrationnel $\begin{matrix} > 0 \\ < 1 \end{matrix}$ correspond une suite infinie déterminée α_ν, et, réciproquement, à chaque suite illimitée α_ν correspond un nombre irrationnel déterminé $e \begin{matrix} > 0 \\ < 1 \end{matrix}$.

Si $e_1, e_2, \ldots, e_\rho$, sont alors ρ grandeurs indépendantes les unes des autres, dont chacune peut prendre toutes les valeurs irrationnelles de l'intervalle (0 ... 1) et ces valeurs seulement, posons :

$$\begin{aligned} e_1 &= (\alpha_{1,1}, \alpha_{1,2}, \ldots, \alpha_{1,\nu}, \ldots) \\ e_2 &= (\alpha_{2,1}, \alpha_{2,2}, \ldots, \alpha_{2,\nu}, \ldots) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_\rho &= (\alpha_{\rho,1}, \alpha_{\rho,2}, \ldots, \alpha_{\rho,\nu}, \ldots) \end{aligned}$$

et définissons à partir de ces nombres un $\overline{\rho + 1}^{\text{ième}}$ nombre irrationnel :

$$\vartheta = (\beta_1, \beta_2, \ldots, \beta_\nu, \ldots)$$

par le système d'équations :

$$\text{(I)} \qquad \beta_{(n-1)\rho+1} = \alpha_{1,n}, \ \ldots \ \beta_{(n-1)\rho+\sigma} = \alpha_{\sigma,n}, \ \ldots \ \beta_{n\rho} = \alpha_{\rho,n}$$

(1) [Publiée sous une forme semblable dans le travail *Contribution à la théorie des multiplicités* (*Ein Beitrag zur Mannigfaltiglkeitslehre*) *Crelle*, 84 (1878) (*Œuvres complètes*, p. 122).]

alors réciproquement, chaque nombre irrationnel $\vartheta \begin{matrix} > 0 \\ < 1 \end{matrix}$ engendrera, par l'intermédiaire de (I), un système déterminé $e_1, e_2, \ldots, e_\rho$.

On ne rencontre pas ici, me semble-t-il, l'obstacle que vous avez relevé pour ma précédente démonstration.

Il s'agit alors de démontrer le théorème suivant :

(B) « Un nombre variable e, qui peut prendre toutes les valeurs *irrationnelles* de l'intervalle (0 ... 1), peut être mis en correspondance *univoque* avec un nombre x, qui prend *toutes* les valeurs de cet intervalle sans exception. »

Car, une fois ce théorème (B) démontré, on met alors en correspondance univoque une par une les variables désignées plus haut par $e_1, e_2, \ldots, e_\rho$ et ϑ respectivement avec d'autres variables

$$x_1, x_2, \ldots, x_\rho, y$$

qui ont toutes un domaine de variation (*) sans restrictions dans l'intervalle (0 ... 1). On a ainsi défini également une relation univoque et réciproque entre le système

$$(x_1, x_2, \ldots, x_\rho),$$

d'une part, et l'unique variable y, de l'autre, ce qui conduit à la démonstration du théorème (A).

Pour démontrer maintenant (B), on met d'abord tous les nombres *rationnels* de l'intervalle (0 ... 1) (extrémités comprises) sous forme d'une suite, soit :

$$r_1, r_2, \ldots, r_\nu, \ldots$$

Les valeurs que peut prendre la variable e sont alors *toutes* celles de l'intervalle (0 ... 1) à *l'exception* des nombres r_ν.

On prend ensuite arbitrairement dans l'intervalle (0 ... 1) une suite infinie ε_ν de nombres *irrationnels*, qui satisfont seulement aux conditions $\varepsilon_\nu < \varepsilon_{\nu+1}$ et $\lim(\varepsilon_\nu) = 1$ pour $\nu = \infty$, et l'on désigne par f une grandeur variable qui peut prendre

(*) *Domaine de variation* : en allemand *Spielraum*.

toutes les valeurs réelles $\begin{matrix}\geqq 0\\ \leqq 1\end{matrix}$ à l'exception des valeurs ε_ν. Les deux grandeurs variables e et f soumises aux restrictions indiquées peuvent alors être mises en relation univoque et réciproque par les définitions suivantes :

si f n'est égal à aucun r_ν, soit alors pour le e correspondant :

$$e = f ;$$

si $f = r_\nu$, soit alors pour le e correspondant $e = \varepsilon_\nu$; on se convaincra facilement que, réciproquement : si e n'est égal à aucun ε_ν, le f correspondant $= e$ et si $e = \varepsilon_\nu$, alors $f = r_\nu$.

Le théorème (B) est alors ramené au suivant :

(C) « Un nombre f, qui peut prendre toutes les valeurs de l'intervalle (0 ... 1), à l'exception de certaines d'entre elles ε_ν, satisfaisant aux conditions suivantes : $\varepsilon_\nu < \varepsilon_{\nu+1}$ et $\lim \varepsilon_\nu = 1$, peut être mis en correspondance univoque avec une variable continue x qui prend toutes les valeurs de l'intervalle (0 ... 1) sans exception. »

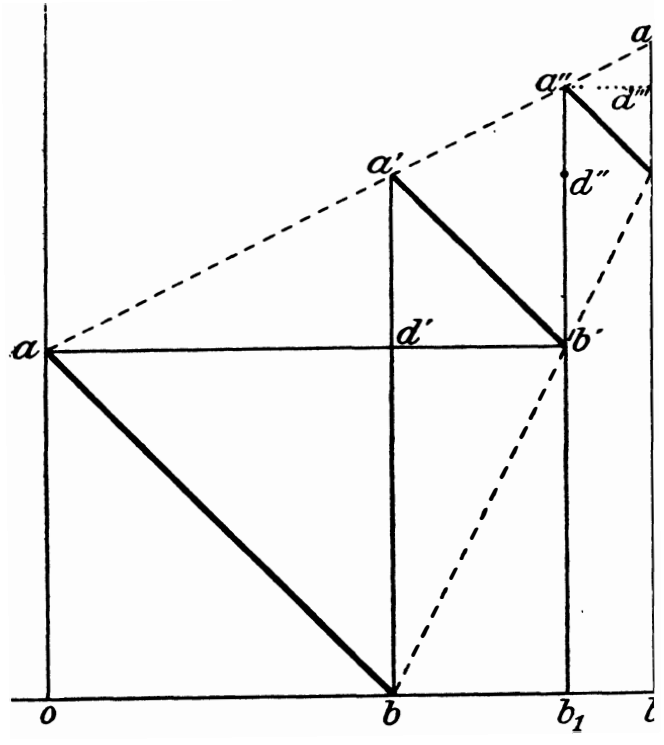
Nous utiliserons ici le fait que les points ε_1, ε_2, ... forment une *suite* (a), et que, par conséquent, l'intervalle (0 ... 1) est divisé par eux en une infinité d'intervalles partiels.

Comme vous ne manquerez pas de le voir, ce théorème (C) peut alors se démontrer par application successive du théorème suivant :

(D) « Un nombre y, qui peut prendre toutes les valeurs de l'intervalle (0 ... 1) à la seule exception de la valeur 0, peut être mis en correspondance univoque avec un nombre x qui prend toutes les valeurs de cet intervalle sans exception. »

Ce dernier théorème (D) peut alors être reconnu comme vrai par la considération de la curieuse courbe ci-jointe. Les coordonnées d'un point courant m de cette courbe sont mes grandeurs x et y dont l'une est une fonction univoque de l'autre ; mais, tandis que x prend toutes les valeurs de l'intervalle (0 ... 1), le domaine de variation de y est le même intervalle *à la seule exception de la valeur* 0.

(a) *Une suite* : en allemand *Eine Folge*.



La courbe se compose d'une infinité de segmen parallèles, devenant toujours plus petits, $\overline{ab}$, $\overline{a'b'}$ et du point *c*. *Les extrémités* b, b′, b″, . . ., *ne sont pa comme appartenant à la courbe.* Les longueurs des seg

$$\overline{op} = \overline{pc} = 1\,; \quad \overline{ob} = \frac{1}{2}\,; \quad \overline{bb_1} = \frac{1}{4}, \quad \overline{b_1b_2} = \frac{1}{8}, \quad \overline{b_2}$$

$$\overline{oa} = \frac{1}{2}\,; \quad \overline{a'd'} = \frac{1}{4}\,; \quad \overline{a''d''} = \frac{1}{8}\,; \quad \overline{a'''d'''} = \frac{1}{16}, \ldots$$

J'ai suivi avec intérêt depuis plusieurs années les efforts que l'on a consacrés, après Gauss, Riemann, Helmholtz et d'autres, à la clarification des questions qui touchent aux premières hypothèses de la géométrie. Il m'est apparu à cet égard que toutes les recherches faites dans ce domaine partent *elles-mêmes* d'une hypothèse non démontrée, qui ne m'apparaît pas comme allant de soi, mais bien plutôt comme ayant besoin d'être fondée. Je veux parler de l'hypothèse selon laquelle une multiplicité continue ρ fois étendue [a] nécessite pour la détermination de ses éléments ρ coordonnées réelles indépendantes entre-elles, le nombre de ces coordonnées ne pouvant être, pour une même multiplicité, ni augmenté ni diminué.

J'en étais venu à croire moi aussi à cette hypothèse, j'étais presque persuadé de son exactitude ; mon point de vue différait seulement de tous les autres en ceci, que je considérais cette hypothèse comme un théorème qui nécessitait au plus haut point une démonstration, et j'avais précisé mon point de vue sous forme d'une question que j'avais soumise à quelques collègues, en particulier aussi à l'occasion du jubilé Gauss, à Göttingen, savoir la question suivante :

« Une variété [b] continue à ρ dimensions, avec $\rho > 1$, peut-elle être mise en relation univoque avec une variété continue à une dimension, de telle sorte qu'à un point de l'une corresponde un point et un seul de l'autre ? »

La plupart de ceux à qui j'ai soumis cette question se sont beaucoup étonnés que j'aie seulement pu la poser, car il *se comprenait de soi* que, pour la détermination d'un point dans une extension [c] à ρ dimensions, il fallait toujours employer ρ coordonnées indépendantes. Celui qui, pourtant, pénétrait le sens de la question, devait reconnaître qu'il fallait au moins démontrer pourquoi la réponse était « évidemment » *non*. Comme je l'ai dit, je faisais partie de ceux, qui tenaient pour *vraisemblable* que la réponse fût négative — jusqu'au moment tout récent où, par une succession assez complexe de pensées, je suis arrivé à la conviction que la réponse était *affirmative* sans aucune restriction.

([a]) *Une multiplicité continue* ρ *fois étendue* : en allemand *eine* ρ *fach ausgedehnte stetige Mannigfaltigkeit.*
([b]) *Variété* : en allemand *Gebilde.*
([c]) *Extension* : en allemand *Ausgedehntheit.*

Peu après, je trouvai la démonstration que vous avez aujourd'hui sous les yeux.

On voit ici quelle force prodigieuse il y a dans les nombres réels habituels, rationnels et irrationnels, si bien que par elle on peut déterminer de façon univoque, *à l'aide d'une seule coordonnée*, les éléments d'une multiplicité continue ρ fois étendue ; et je veux ajouter tout de suite que leur force va plus loin encore, puisque, comme vous ne manquerez pas de le voir, ma démonstration peut s'étendre, sans que les difficultés en soient sensiblement accrues, à des multiplicités à un nombre infini de dimensions, pourvu que ces dimensions en nombre infini prennent la forme d'une suite simplement infinie.

Il me semble donc que toutes les déductions philosophiques ou mathématiques qui utilisent cette hypothèse erronée sont inadmissibles. Il faut plutôt rechercher la différence qui existe entre deux variétés [a] à un nombre *différent* de dimensions, dans quelque raison tout autre que celle, généralement tenue pour caractéristique, du nombre de coordonnées indépendantes.

Halle, le 29 juin 1877

Veuillez excuser mon zèle pour cette affaire, si je fais appel tellement souvent à votre bonté et à votre peine ; ce que je vous ai communiqué tout récemment est pour moi-même si inattendu, si nouveau, que je ne pourrai pour ainsi dire pas arriver à une certaine tranquillité d'esprit avant que je n'aie reçu, très honoré ami, votre jugement sur son exactitude. Tant que vous ne m'aurez pas approuvé, je ne puis que dire : Je le vois, mais je ne le crois pas [b]. C'est pourquoi je vous prie de m'envoyer une carte postale pour me dire quand vous pourriez avoir terminé l'examen de la chose et si je peux compter vous voir exaucer ma demande, certainement bien exigeante.

La démonstration du théorème (C) sera considérablement facilitée par l'emploi du symbolisme suivant :

Si a et b sont deux grandeurs variables qui peuvent être mises en relation univoque l'une avec l'autre, on écrit :

$$a \sim b.$$

[a] *Variétés* : en allemand *Gebilde*.
[b] En français dans le texte.

Si alors $a \sim b$ et $b \sim c$, on a aussi :

$$a \sim c.$$

Si de plus a', a'', ... est une suite finie ou infinie de variables bien définies ou de constantes, qui ne prennent deux à deux aucune valeur commune, mais telles que la réunion de leurs domaines de variations soit exactement celui de la variable unique a, on pose :

$$a \equiv (a', a'', \ldots).$$

On a alors le théorème suivant :

(E) « Si : $a \equiv (a', a'', \ldots)$
$b \equiv (b', b'', \ldots)$

et si de plus :

$$\begin{aligned} a' &\sim b' \\ a'' &\sim b'' \\ a''' &\sim b''' \\ &\ldots\ldots\ldots \end{aligned}$$

on a alors aussi :

$$a \sim b.$$

Grâce aux substitutions :

$$y = \frac{z - \alpha}{\beta - \alpha}; \qquad x = \frac{u - \alpha}{\beta - \alpha}$$

on déduit de (D) la généralisation suivante :
(F) « Un nombre z, qui peut prendre toutes les valeurs d'un intervalle $(\alpha \ldots \beta)$ à l'exception de α, peut être mis en correspondance univoque avec un nombre u, qui prend toutes les valeurs de l'intervalle $(\alpha \ldots \beta)$ sans exception. »

De là résulte immédiatement le théorème suivant :
(G) « Un nombre w, qui prend toutes les valeurs de l'intervalle $(\alpha \ldots \beta)$, à l'exception des deux valeurs extrêmes α, β, peut être mis en correspondance univoque avec un nombre variable u qui prend toutes les valeurs de l'intervalle $(\alpha \ldots \beta)$. »

Démonstration. Soit un nombre $\gamma \begin{matrix} < \alpha \\ > \beta \end{matrix}$.

w' une variable qui prend toutes les valeurs de l'intervalle $(\alpha \ldots \gamma)$ à l'exception de α *et* de γ ; w'' une variable qui prend toutes les valeurs de l'intervalle $(\gamma \ldots \beta)$ à l'exception de la seule extrémité β.

On a alors :

$$w \equiv (w', w''). \tag{1}$$

Si l'on désigne alors par u'' une variable qui prend toutes les valeurs de l'intervalle $(\gamma \ldots \beta)$ sans exception, par z une variable qui prend toutes les valeurs de l'intervalle $(\alpha \ldots \beta)$ à l'exception de α, on a :
d'après (F) :

$$w'' \sim u'', \tag{2}$$

et, à cause de (1) et (E) :

$$w \sim (w', u'').$$

Mais : $(w', u'') \equiv z$, donc :

$$w \equiv z.$$

Mais d'après (F) on a aussi :

$$z \sim u, \quad \text{donc enfin :}$$
$$w \sim u, \quad \text{q. e. d.}$$

Pour démontrer maintenant (C), nous décomposons f en variables f', f'', ... et la valeur isolée 1, f' prenant toutes les valeurs de l'intervalle $(0 \ldots \varepsilon_1)$ à l'exception de ε_1, $f^{(\nu)}$ toutes les valeurs de l'intervalle $(\varepsilon_{\nu-1} \ldots \varepsilon_\nu)$ à l'exception des valeurs extrêmes $\varepsilon_{\nu-1}$ et ε_ν. On a alors :

$$f \equiv (f', f'', f''', \ldots f^{(\nu)}, \ldots, 1).$$

Soit x'' une variable qui prend toutes les valeurs de $(\varepsilon_1 \ldots \varepsilon_2)$ sans exception, $x^{(IV)}$ une variable qui prend toutes les valeurs de $(\varepsilon_3 \ldots \varepsilon_4)$ sans exception, $x^{(2\nu)}$ une variable qui prend toutes

les valeurs de l'intervalle $(\varepsilon_{2\nu-1} \ldots \varepsilon_{2\nu})$ sans exception, alors on a, à cause de (G) :

$$\begin{aligned} f'' &\sim x'' \\ f^{IV} &\sim x^{IV} \\ &\ldots\ldots\ldots \\ f^{(2\nu)} &\sim x^{(2\nu)} \\ &\ldots\ldots\ldots \end{aligned}$$

par suite :

$$f \sim (f', x'', f''', x^{IV}, \ldots, f^{(2\nu-1)}, x^{(2\nu)}, \ldots, 1).$$

Mais :

$$(f', x'', f''', x^{IV}, \ldots, f^{(2\nu-1)}, x^{(2\nu)}, \ldots, 1) \equiv x,$$

donc :

$$f \sim x.$$

DEDEKIND A CANTOR

J'ai examiné encore une fois votre démonstration, et n'y ai pas trouvé de lacune ; je suis convaincu que votre intéressant théorème est exact et je vous en félicite. Mais je voudrais, comme je vous l'ai déjà annoncé sur ma carte postale, faire une remarque, qui est dirigée *contre* les conséquences que vous avez ajoutées dans votre lettre du 25 juin à la communication et à la démonstration du théorème, et qui ont trait au concept d'une multiplicité continue à ρ dimensions. D'après ce que vous dites, il pourrait sembler — mais mon opinion peut être inexacte — que vous vouliez, à partir de votre théorème, mettre en doute la signification ou l'importance de ce concept ; vous dites par exemple à la fin de cette lettre : « *Il me semble donc* que toutes les déductions philosophiques ou mathématiques qui utilisent cette hypothèse erronée » [celle du caractère déterminé du nombre de dimensions] « sont inadmissibles. Il faut plutôt rechercher la différence qui existe entre deux variétés à un nombre différent de dimensions dans quelque raison tout autre que celle, généralement tenue pour caractéristique, du nombre de coordonnées indépendantes. »

A l'encontre de ce point de vue, je déclare être convaincu (malgré votre théorème, ou plutôt par suite des considérations occasionnées par votre théorème) ou croire (je n'ai pas encore eu le temps de faire même seulement l'essai d'une démonstration) que le nombre de dimensions d'une multiplicité continue est, après comme avant, le premier et le plus important de ses invariants, et je dois prendre ici la défense de tous ceux qui ont écrit sur ce sujet jusqu'à ce jour. Je vous concède volontiers que cette constance du nombre de dimensions exige absolument une démonstration, et que, aussi longtemps que cette démonstration n'a pas été faite, l'on est en droit de mettre en doute la proposition. Mais je ne doute pas de cette constance bien qu'elle semble anéantie par votre théorème. Mais tous les auteurs ont évidemment fait l'hypothèse tacite, tout à fait naturelle, que pour une nouvelle détermination des points d'une multiplicité continue à l'aide de nouvelles coordonnées, ces dernières doivent être aussi des fonctions *continues* (en général) des anciennes coordonnées, afin que ce qui apparaît comme étant continuement connexe ([a]) dans la première détermination de position) reste continuement lié ([b]) dans la deuxième détermination de position. Je crois donc provisoirement à l'exactitude du théorème suivant : « Si l'on réussit à établir une correspondance complète univoque et réciproque ([c]) entre les points d'une multiplicité continue A à a dimensions d'une part, et les points d'une multiplicité continue B à b dimensions, d'autre part, alors, si a et b sont *inégaux*, cette correspondance est nécessairement partout *discontinue*. Ce théorème expliquerait aussi le phénomène qui s'est manifesté lors de votre première démonstration de votre théorème, à savoir justement l'insuffisance de cette démonstration ; la relation que vous vouliez y établir (à l'aide de fractions décimales) entre les points d'un domaine à ρ dimensions et les points d'un « segment à une dimension » aurait été, si je ne me trompe, *continue*, pourvu seulement qu'elle ait englobé *tous* les points du segment à une dimension ; de même, dans votre démonstration actuelle, la correspondance *initiale* entre les points du segment à ρ dimen-

([a]) *Continuement connexe* : en allemand *als stetig zusammenhängend.*
([b]) *Continuement lié* : en allemand *stetig verbunden.*
([c]) *Une correspondance complète, univoque et réciproque* : en allemand *eine gegenseitige eindeutige und vollständige Correspondenz.*

sions dont les coordonnées sont toutes irrationnelles et les points à une coordonnée également irrationnelle du segment à une dimension est, me semble-t-il, aussi continue que possible en un certain sens (petitesse des changements) ; mais, pour remplir les lacunes, vous êtes obligé d'introduire dans la correspondance une discontinuité, à donner le vertige, une discontinuité qui réduit tout en atomes, telle que toute partie continuement connexe, si petite qu'elle soit, de l'un des domaines a une image complètement déchirée, discontinue.

J'espère m'être expliqué assez clairement ; cette lettre n'a d'autre but que de vous prier de ne pas entreprendre publiquement des polémiques contre les articles de foi admis jusqu'à présent de la théorie des multiplicités avant d'avoir soumis mon objection à un examen approfondi.

Brunswick, le 2 juillet 1877

CANTOR A DEDEKIND

(Carte ; timbre de la poste du 2-7-77)

Je suis très heureux que vous ayiez étudié mon travail et l'ayez trouvé exact. Je vous prie, de persévérer dans votre projet et de me communiquer vos idées sur la signification du résultat, en détail et avec précision ; d'après elles j'aimerais former mon jugement sur la façon de poursuivre toute l'affaire.

Halle, le 4 juillet 1877

J'ai été très heureux de recevoir votre lettre du 2 juillet, et je vous remercie de vos remarques précises et extraordinairement pertinentes.

A la fin de ma lettre du 25 juin, j'ai, contre ma volonté, donné l'impression de vouloir, par ma démonstration, m'opposer au concept même de multiplicité continue ρ fois étendue, alors que tous mes efforts n'ont d'autre but que de clarifier ce concept

et de lui donner un fondement correct. Lorsque je disais : « Il me semble donc que toutes les déductions philosophiques ou mathématiques qui utilisent cette hypothèse erronée sont inadmissibles », *ce n'est pas* au « caractère déterminé du nombre de dimensions » que se rapportait cette hypothèse, mais bien au caractère déterminé des coordonnées indépendantes dont le nombre est supposé par certains auteurs égal en toute circonstance au nombre de dimensions, alors que, si l'on prend le concept de coordonnées *dans sa généralité*, sans aucune hypothèse sur la nature des fonctions utilisées, le nombre de coordonnées indépendantes, univoques, complètes peut, comme je l'ai démontré, être ramené à tout nombre donné. Je suis aussi de votre avis que, si l'on impose à la correspondance d'être continue, on ne peut plus alors mettre en relation univoque que des variétés à un nombre égal de dimensions, et que, de cette manière, le nombre de coordonnées indépendantes constitue un invariant qui pourrait conduire à la définition du nombre de dimensions d'une variété continue.

Je ne suis pourtant pas encore capable de reconnaître à quel degré peuvent s'élever les difficultés dans cette voie (celle qui conduit au concept du nombre de dimensions), parce que je ne sais pas si l'on est en état de délimiter le concept de *correspondance continue en général* (a). Mais de la possibilité d'une telle délimitation me paraît tout dépendre dans cette voie.

Je crois apercevoir une autre difficulté en ceci que cette voie n'aboutirait sans doute pas dès que la variété cesse d'être partout continue ; et pourtant l'on aimerait avoir dans ce cas aussi quelque chose qui corresponde au nombre de dimensions, d'autant que, pour les multiplicités qui se présentent dans la nature, leur continuité partout semble difficile à établir.

Je voudrais seulement indiquer par ces lignes que je suis bien loin d'avoir voulu utiliser inconditionnellement mon résultat contre les articles de foi de la théorie des multiplicités, et, bien au contraire, que je souhaite pouvoir contribuer avec son aide à en raffermir autant que possible les théorèmes. Je ne voudrais pas vous déranger davantage aujourd'hui ; je vous prie, si vous en trouvez le temps, d'examiner les questions qui s'imposent, de

(a) *Correspondance continue en général* : en allemand *Im Allgemeinen stetigen Correspondenz.*

ne pas dédaigner ce travail, et de me faire connaître ensuite vos résultats.

Halle, le 23 octobre 1877

... Sur les recherches que j'ai faites l'été dernier, il y a chez Monsieur Borchardt, depuis un trimestre, un travail que j'ai rédigé sous le titre « Une contribution à la théorie des multiplicités ». J'espère qu'il va très bientôt être publié. Comme il m'a été permis d'y mettre à profit vos conseils amicaux, cela vous intéressera peut-être de savoir que j'ai trouvé pour l'un de mes théorèmes une démonstration encore plus simple (1). Si deux multiplicités bien définies peuvent être mises en correspondance univoque et complète, élément par élément, l'une avec l'autre, je me sers alors du mode d'expression qu'elles ont *même puissance* ou encore qu'elles sont *équivalentes* ; j'appelle aussi *équivalentes* deux variables réelles a et b, lorsqu'elles peuvent être mises en correspondance univoque et complète l'une avec l'autre, et j'écris dans ce cas, comme vous le savez :

$$a \sim b.$$

Il s'agit alors du théorème suivant :

« Si e est une variable, qui doit prendre toutes les valeurs irrationnelles $\begin{matrix} > 0 \\ < 1 \end{matrix}$, et x une variable qui prend toutes les valeurs rationnelles et irrationnelles $\geqq 0$ et $\leqq 1$, on a :

$$e \sim x.$$

Démonstration : Soit φ_ν le terme général d'une suite, qui se compose de tous les nombres rationnels $\geqslant 0$ et $\leqslant 1$; soit η_ν le terme général d'une suite formée de n'importe quels nombres irrationnels inégaux $\begin{matrix} > 0 \\ < 1 \end{matrix}$, par exemple :

$$\eta_\nu = \frac{\sqrt{2}}{2^\nu}$$

(1) [*Contribution à la théorie des multiplicités* (*Ein Beitrag z. Mannigfaltigfaltigkeitslehre*) § 6 (*Œuvres complètes*, p. 129).]

h désignera la variable qui prend toutes les valeurs de l'intervalle (0 ... 1) à l'exception des φ_ν et des η_ν.

Alors

$$x \equiv \{ h, \eta_\nu, \varphi_\nu \} \tag{1}$$
$$e \equiv \{ h, \eta_\nu \}.$$

Pour la dernière formule, nous pouvons aussi écrire :

$$e \equiv \{ h, \eta_{2\nu-1}, \eta_{2\nu} \}. \tag{2}$$

Si l'on compare les formules (1) et (2) et que l'on remarque que :

$$h \sim h ; \quad \eta_\nu \sim \eta_{2\nu-1} ; \quad \varphi_\nu \sim \eta_{2\nu},$$

il en résulte que :

$$x \sim e, \quad \text{C. Q. F. D.}$$

Peut-être avez-vous davantage étudié la question de savoir si, pour la détermination du concept de la multiplicité n-uple continue, la condition de la correspondance continue suffit pour que le concept soit affermi en lui-même, assuré contre toute contradiction ?

P. S. : J'ai regardé le nouveau manuel d'analyse de Lipschitz ; vous plaît-il ?

DEDEKIND A CANTOR

Le 27-10-77

... Mais je *crois* toujours que le concept du nombre de dimensions reçoit réellement son caractère d'invariant par la condition de la correspondance *continue.*

Il y a dans l'ouvrage de Lipschitz, pour autant que j'aie pu l'examiner jusqu'à présent, beaucoup de bon et d'intéressant ; conformément à ma nature critique, j'ai, il est vrai, quelques objections à y faire sur certains points ; mais je trouve cependant très réjouissant qu'une tentative sérieuse soit faite pour faire régner la rigueur mathématique même dans un *manuel.* En ce qui concerne le fondement de la théorie des irrationnels,

pour laquelle il adopte presque tout à fait votre présentation, publiée pour la première fois par M. Heine, j'ai tout au plus à objecter que (p. 46), une *hypothèse* à vrai dire tout à fait nouvelle (« Alors la valeur limite $\mathfrak{S}$ tombe etc... ») est présentée apparemment comme une *conséquence* évidente de ce qui précède ; de plus, je ne puis considérer comme exacte la remarque finale du § 14 sur les Grecs [1]...

CANTOR A DEDEKIND

Halle sur la Saale, la 29 décembre 1878

... Vous devez sans doute être entré en possession du livre : « Fondamenti per la teorica delle funzioni di variabili reali di *Ulisse Dini*, Pisa 1878 [a] ». Cet ouvrage me semble être rédigé par quelqu'un qui connaît son sujet et est très habile ; pour l'introduction des nombres, il se sert de votre méthode. — Bien que je sois entièrement d'accord avec cette dernière, je crois néanmoins que celle indiquée par moi dans un travail sur les séries trigonométriques [2] lui est équivalente, et que, par la distinction formelle de grandeurs numériques d'ordre différent, par quoi j'ai *seulement* voulu exprimer les différentes manières de les définir par des suites simplement infinies (dont les termes d'indice croissant se rapprochent indéfiniment *l'un de l'autre*), il n'y a pas de danger que l'on puisse croire que j'aie voulu étendre le domaine des nombres réels. Cette bévue, je n'ai *jamais* envisagé, *même de loin*, de la commettre ; je dis expressément dans mon travail que tout nombre que je désigne par c peut être égal à un nombre b. Du reste, cette bévue a été réellement faite d'un autre côté, si inouï que cela puisse paraître ; je ne sais si vous connaissez l'esquisse d'une théorie des fonctions complexes et des fonctions Theta de Thomae ; dans la deuxième édition, à la page 9, on trouve des

[1] [Cf. la correspondance de Dedekind avec Lipschitz dans les *Œuvres complètes* de Dedekind, tome III, pp. 469-479.]

[a] En italien dans le texte.

[2] [*Sur l'extension d'un théorème de la théorie des séries trigonométriques* (*Ueber die Ausdehnung eines Satzes aus d. Theorie d. trigonometrischen Reihen*) *Math. Ann.*, vol. V (1872), p. 123 (*Œuvres complètes*, p. 95.)]

nombres (1) qui sont (horribile dictu) plus petits que tout nombre réel concevable, et cependant différents de zéro (2).

La question de savoir si des multiplicités continues d'un nombre différent de dimensions peuvent être mises en correspondance univoque et continue, ou plutôt sur le théorème suivant lequel cela n'est pas possible, a fait l'objet, depuis la parution de mon travail sur la théorie des multiplicités, de mémoires de Thomae, Lüroth, Jürgens et, il y a quelques jours, de Netto, dans le journal de Borchardt (3) ; il me semble pourtant que la question n'est pas encore complètement réglée.

Halle sur la Saale, le 17 janvier 1879

Je crois avoir réglé maintenant de la manière la plus simple et la plus rigoureuse la question de l'application univoque et *continue* des multiplicités, qui s'est posée naturellement à la suite de mes recherches ; je la ramène au théorème fondamental bien connu de l'analyse, d'après lequel :

(I) Une fonction continue d'*une* variable continue t, qui, pour $t = t_0$, prend une valeur négative, pour $t = t_1$, une valeur positive, s'annule au moins une fois entre ces valeurs.

Les tentatives respectives de Thomae et de Netto sont incomplètes, comme vous l'avez peut-être remarqué ; ainsi, par exemple, Thomae s'appuie sur un théorème de Riemann indémontrable *pour lui* (4) (*Œuvres complètes*, page 450, (1) : Un

(1) Il s'agit des ordres d'annulation d'une fonction ; dans leur description — en tant que domaine non archimédien de grandeurs —, Thomae se rencontre de manière curieuse avec les recherches de Du Bois Raymond, parues peu avant (*Annali di Mat.*, IV, 1871).

(2) Sur cette exclusion, souvent soulignée (*Œuvres complètes*, p. 156, p. 172, etc...) des grandeurs infiniment petites actuelles, cf. la lettre à Weierstrass du 16 mai 1887 (*Contributions à la théorie des nombres transfinis*, VI, *Œuvres complètes*, p. 408), où Cantor croit avoir trouvé une démonstration. En réalité, la seule raison en était sa définition du continu comme système connexe — c'est-à-dire soumis à l'axiome d'Archimède.

(3) [Thomae, *Théorèmes de la théorie des fonctions*, *Göttinger Nachrichten* (1878), Lüroth, *Application de multiplicités de différentes dimensions l'une sur l'autre*, Erlangen, *Phys. med. Soc.*, *Sitz-Ber*, 10 (1878), Jürgens, *Sur l'application continue et univoque de multiplicités*, *Tagebl. d. Versamme. Deutsch. Naturforscher u. Aetzte* in Cassel 1878. Netto. *Zur Mannigfaltigkeitslehre. Borch. J.* 86 (1879).]

(4) Il l'appelle : « axiome ».

segment [a] à moins de $n - 1$ dimensions, etc., etc...) ; alors que, je l'ai maintenant clairement reconnu, ce théorème de Riemann est, dans une certaine mesure, équivalent à celui qu'il s'agit maintenant de démontrer dans le cas où $\nu = n - 1$; mais comme $n - 1$ n'est pas un nombre moins général que n, Thomae fait une sorte de cercle vicieux avec sa démonstration.

La démonstration générale que je vais donner maintenant, je la connais à vrai dire déjà depuis longtemps, en fait depuis plus d'un an ; mais je ne la croyais pas rigoureuse jusqu'à présent, et c'est pourquoi je me gardais bien d'en parler. La découverte que j'ai faite il y a quelques jours réside donc au fond seulement en ce que cette démonstration est rigoureuse. Si je pensais le contraire, c'est que les relations qui s'y présentent ne sont pas univoques *des deux côtés.* Mais leur *caractère multivoque ne* nuit *pas* au résultat, parce qu'il n'intervient que dans le passage d'une variété [b] à un nombre supérieur de dimensions à une variété [b] à un nombre inférieur de dimensions [1].

Dans ce qui suit, j'entends par sphère [c] du p^{ieme} ordre la variété [b] continue du p^{ieme} ordre qui est extraite par le moyen de l'équation :

$$(x_1 - a_1)^2 + (x_2 - a_2)^2 + \ldots + (x_{p+1} - a_{p+1})^2 = r^2$$

de la multiplicité $p + 1$ uple, dont les coordonnées sont x_1, x_2, ..., x_{p+1}, de telle sorte que par exemple la circonférence dans le plan est une sphère du 1er ordre.

Le théorème à démontrer est, sous une forme élargie, le suivant :

« Une M_μ continue et une M_ν continue, si $\mu < \nu$, ne peuvent être mises en correspondance *continue* de telle sorte qu'à chaque élément de M_μ corresponde *un seul* élément de M_ν, et à chaque élément de M_ν *un ou plusieurs* éléments de M_μ. »

Pour le cas de $\nu = 1$, le théorème est immédiat. Pour le

[a] *Segment :* en allemand *Vielstreck.*

[b] *Variété* : en allemand *Gebilde.*

[1] [Il est pourtant bien connu que ce théorème est faux dans le cas d'une correspondance multivoque, comme suffit à le montrer la courbe de Peano. La démonstration ci-dessus a été publiée : *Sur un théorème de la théorie des multiplicités continues* (*Ueber einen Satz aus der Theorie der stetigen Mannigfaltigkeiten*) Göttinger Nachrichten. 1879 (*Œuvres complètes*, p. 134).]

[c] *Sphère* : en allemand *Kugel.*

prouver dans le cas général, supposons-le vrai pour $\nu = n - 1$, et montrons qu'il est alors aussi vrai pour $\nu = n$.

A cette fin, supposons l'existence d'une relation continue entre une M_μ et une M_n, avec $\mu < n$, de telle sorte que le passage de M_μ à M_n soit *univoque* ; et montrons qu'il y a une contradiction interne à la base de cette hypothèse, ou plutôt une contradiction avec le théorème fondamental I énoncé plus haut.

Soit a et b deux points intérieurs de M_μ, A et B les points correspondants de M_n.

Autour de A comme centre, construisons dans M_n une sphère du $\overline{n-1}^{\text{eme}}$ ordre, K_{n-1}, qui soit assez petite pour que le point B soit à l'extérieur de l'espace qu'elle délimite.

Autour de a comme centre, construisons de même dans M_μ une sphère du $\overline{\mu-1}^{\text{eme}}$ ordre $K_{\mu-1}$, qui soit assez petite pour que :

1) le point b soit à l'extérieur de cette sphère ;

2) la variété continue (a) du $\overline{\mu-1}^{\text{eme}}$ ordre correspondant dans M_n à cette sphère soit tout entière contenue à l'intérieur de l'espace délimité par la sphère K_{n-1}, ce qui peut être réalisé, à cause de la continuité de la relation dans le voisinage des points a et A.

Soit z un point quelconque de $K_{\mu-1}$, ζ le point correspondant de $G_{\mu-1}$; traçons le rayon rectiligne Aζ ; prolongé dans cette direction, il rencontre la sphère K_{n-1} en un point et un seul, Z.

Ainsi, *au moyen de cette construction*, à chaque point z de K_{n-1}, correspond *un* point Z de $K_{\mu-1}$ qui varie de façon continue avec z.

Mais le point Z ne peut pas atteindre n'importe quel point de K_{n-1}, *car cela irait contre notre théorème, qui a été supposé démontré dans le cas* $\nu = n - 1$.

Nous concluons donc avec certitude qu'il y a des points P de la sphère K_{n-1} qui ne sont pas atteints par le point Z ; si l'on joint A à un tel point P, ce rayon AP ne rencontrera *pas* la variété $G_{\mu-1}$.

Si l'on joint alors P encore au point B, qui est à l'extérieur de K_{n-1}, par une courbe continue contenue dans l'espace M_n, on obtient

(II) une ligne continue composée APB, *qui n'a aucun point commun avec la variété* $G_{\mu-1}$.

(a) *Variété continue* : en allemand *Stetige Gebilde.*

A cette ligne, par l'intermédiaire de la relation continue posée au départ entre M_μ et M_n, correspondent une ou plusieurs courbes situées dans M_μ qui joignent continûment a à b et qui (III) à cause du théorème fontamental I, doivent rencontrer *nécessairement* la sphère $K_{\mu-1}$ *une fois au moins.*

Ces deux faits, II et III, se contredisent mutuellement ; donc l'hypothèse d'une relation continue entre M_μ et M_n est incorrecte, et notre théorème est démontré pour le cas $\nu = n$.

Comme l'académie de Göttingen m'a récemment fait membre correspondant, et que Thomae a publié sa démonstration dans les *Göttinger. Anzeigen*, je pense envoyer également à ce même périodique cette démonstration ; c'est pourquoi j'aimerais bien savoir ce que vous en pensez.

DEDEKIND A CANTOR

J'ai étudié votre démonstration avec soin, et je n'y ai rencontré qu'un détail qui pourrait soulever le doute. Après avoir construit autour du point A de M_n une sphère K_{n-1}, et autour du point a de M_μ une sphère $K_{\mu-1}$ dont l'image dans M_n est désignée par $G_{\mu-1}$, vous en déduisez une application de $K_{\mu-1}$ sur K_{n-1} : au point z de $K_{\mu-1}$ correspond un point déterminé ζ de $G_{\mu-1}$, l'intersection Z du rayon Aζ avec K_{n-1} sera alors considérée comme une nouvelle image de z. Mais il est alors concevable que ζ coïncide avec A, puisque vous admettez qu'à *plusieurs* points de M_μ correspond un seul et même point de M_n ; dans ce cas, l'image Z serait en général *indéterminée.* Cette difficulté est visiblement facile à écarter lorsque le nombre de points a' dans M_μ auxquels correspond le même point A dans M_n est *fini*, car il suffit alors de prendre le rayon de la sphère $K_{\mu-1}$ assez petit pour que les autres points a' tombent à l'extérieur. Mais si le nombre de points a' est infini, je vois pour le moment dans cette circonstance une véritable difficulté, et celle-ci intervient encore à une deuxième place dans votre démonstration. Vous dites en effet qu'à la ligne APB correspondront une ou plusieurs lignes dans M_μ qui y joignent continûment a à b ; ceci exigerait à mon avis d'être expliqué et fondé de façon plus précise, *même* lorsque B est seulement l'image d'un point ou d'un nombre *fini* de points distincts b' de M_μ ; mais si une infinité de points b' de M_μ pos-

sèdent la même image B dans M_μ, l'existence d'une ligne joignant continûment a à b, et dont l'image doit être la ligne APB, devient encore un peu plus douteuse. Je *crois* d'ailleurs que le théorème reste encore vrai si l'on admet que chaque point de M_μ peut être l'image d'une infinité de points distincts de M_μ. Peut-être parviendrez-vous à écarter immédiatement les deux difficultés en question. Pour une publication, je tiendrais pour souhaitable que soient exactement définis les noms ou expressions techniques de la théorie des multiplicités (à ce propos, je donnerais décidément la préférence au terme de « domaine » [a], également employé par Riemann, comme étant bien plus court que le terme pesant de « multiplicité » [b]) ; il serait très méritoire de développer ab ovo toute cette « théorie des domaines », sans faire appel à l'intuition géométrique, et il faudrait alors définir d'une façon nette et précise par exemple le concept d'une ligne joignant continûment le point a au point b à l'intérieur du domaine G. Les définitions de Netto (dont le mémoire me plaît beaucoup, et dont la démonstration devient, à ce que je crois, tout à fait correcte avec quelques modifications) contiennent un bon germe, mais elles me semblent susceptibles d'être simplifiées et en même temps complétées. Je ne me permettrais pas un tel jugement, si je ne m'étais pas déjà beaucoup occupé de questions de ce genre, il y a un bon nombre d'années, lorsque je voulais encore éditer les leçons de Dirichlet sur la théorie du potentiel, et, à cette occasion, fonder plus rigoureusement ce qu'on appelle le principe de Dirichlet. J'ai moi-même quelques définitions de ce genre [1], qui me semblent fournir une base satisfaisante, mais j'ai par la suite laissé toute l'affaire de côté, et ne pourrais en ce moment donner que quelque chose d'incomplet, car je suis entièrement pris par le remaniement de la théorie des nombres de Dirichlet. Toutefois, votre communication, j'ai à peine besoin de le dire, m'a intéressé au plus haut point, et, en vous remerciant très vivement, je reste avec mes salutations cordiales...

Brunswick, le 19 janvier 1879

(a) *Domaine* : en allemand *Gebiet.*

(b) *Multiplicité* : en allemand *Mannigfaltigkeit.*

(1) [Cf. « Théorèmes généraux sur les espaces » (Allgmeine Sätze über Räume) extraits des œuvres posthumes (*Œuvres complètes*, vol. II, p. 352).]

CANTOR A DEDEKIND

(Carte postale : timbre de la poste du 20-1er-79)

Dans la démonstration que je vous ai envoyée le 17, je trouve un point, *peu important* il est vrai, à améliorer.

Il est *préférable* de partir de deux points A et B intérieurs *de* M_n, auxquels correspondent des points intérieurs de M_μ.

Soit *a l'un* des points intérieurs correspondant au point A et *b*, *b'*, *b''*, ... *tous* les points correspondant au point B. La sphère $K_{\mu-1}$ de centre *a* sera supposée assez petite pour que $G_{\mu-1}$ soit à l'intérieur de K_{n-1} et qu'en même temps *tous les points b*, *b'*, *b''*, ... soient à *l'extérieur de* $K_{\mu-1}$. Ceci est possible à cause de la continuité, *qui empêche* les points *b*, *b'*, *b''*, ... de se rapprocher indéfiniment du point *a*.

Il ne devrait alors plus rester aucun doute que l'*une* des courbes correspondant dans M_μ à la ligne APB va de *a* à l'*un* des points *b*, *b'*, *b''*, ... et doit donc rencontrer la sphère $K_{\mu-1}$.

Après avoir écrit cette carte, je reçois votre lettre pour laquelle je vous remercie beaucoup. J'y répondrai plus tard, car je dois partir tout de suite pour une soirée.

Carte : timbre de la poste 21-1-79)

Hier soir, sur le point de sortir, je n'ai pu faire plus que d'accuser réception de votre lettre sur ma carte déjà terminée. J'ai répondu par avance dans cette carte à l'une de vos objections ; quant à l'autre, suivant laquelle la variété $G_{\mu-1}$ peut contenir le point A, il me semble en effet qu'il y a là une difficulté que je ne puis momentanément surmonter complètement, mais qu'on pourra peut-être lever en choisissant convenablement le point A, pour lequel on dispose d'une si grande latitude. Quoi qu'il en soit, j'avais bien l'intention de donner au caractère multivoque du passage de la multiplicité supérieure à la multiplicité inférieure une *telle généralité* qu'à un point de M_n puisse éventuellement correspondre *une infinité* de points de M_μ ; par suite cela *ne m'arrangerait pas du tout*, si je devais, pour sauver ma démonstration, restreindre cette hypothèse. Quoi qu'il en soit, je n'envisage une

publication *que dans le cas* où je parviendrais à régler également ce point.

Halle sur la Saale, le 22 décembre 1879

Vous aurez sans doute reçu il y a quelque temps de ma part une petite note (1) sur une démonstration d'un Monsieur Appell. Les explications suivantes vous intéresseront peut-être. Appell applique le théorème suivant : « Si $f(n, x)$ pour chaque valeur de $x \geqq \alpha$ et $\leqq \beta$ devient ∞ petit pour $n = \infty$, et que $f(n, x)$ est une fonction continue de x ayant pour maximum absolu B_n, on a alors Lim $B_n = 0$ pour $n = \infty$. » Or ce théorème est faux (en général), comme le montre l'exemple suivant :

$$f(n, x) = e^{-\left(\frac{nx}{1-x} + \frac{1-x}{nx}\right)}, \text{ pour } 0 < x < 1,$$

avec $f(n, 0) = f(n, 1) = 0$. Cette fonction est une fonction continue de x pour $0 \leqq x \leqq 1$; en outre, pour chaque valeur particulière de x telle que $x \geqq 0$ et $\leqq 1$: Lim $f(n, x) = 0$ pour $n = \infty$. On voit facilement que le maximum de $f(n, x)$, pour chaque valeur particulière de n, est $= e^{-2}$, donc $B_n = e^{-2}$; *donc B_n ne devient pas ∞ petit.* Que dans le cas :

$$f(n, x) = a_n \sin nx + b_n \cos nx$$

B_n devient infiniment petit lorsque n croît, cela *résulte* seulement de ma démonstration, mais cela ne peut être *pris comme hypothèse*, sinon la démonstration se réduirait à un circulus vitiosus, comme c'est le cas chez Appell.

Les résultats sur la théorie des fonctions publiés au cours de ces derniers mois dans les comptes rendus par Em. Picard m'ont beaucoup plu ; si ce jeune mathématicien continue de cette manière, nous pouvons attendre de lui encore beaucoup de belles choses.

Je constate après coup que l'exemple

$$f(n, x) = \frac{nx(1 - x)}{n^2x^2 + (1 - x)^2}$$

$$\text{pour } 0 \leqq x \leqq 1$$

(1) [*Une autre remarque sur les séries trigonométriques* (*Fernere Bemerkung über trigonometrische Reihen*), *Math. Ann.*, XVI (1880) (*Œuvres complètes*, p. 104).]

remplit le même but que celui donné plus haut, et est plus simple !

Berlin, le 7 avril 1882

... Je me permets d'attirer votre attention sur le résultat curieux suivant, que vous avez peut-être déjà remarqué vous-même.

Si l'on a un domaine A continûment connexe n fois étendu [a], et dans ce domaine un ensemble de points [b] (M_ν) partout dense [c], mais *dénombrable*, le domaine $\mathcal{A}$ que l'on obtient en retranchant (M_ν) de A est alors pour $n \geqslant 2$ lui aussi continûment connexe, dans ce sens que deux points quelconques N et N′ de ce domaine peuvent être joints par une infinité de lignes continues, définissables analytiquement, situées à l'intérieur de ce domaine $\mathcal{A}$, sur lesquelles ne se trouve donc pas un seul point de l'ensemble (M_ν). Le *mouvement* est donc possible en un certain sens même dans de tels espaces $\mathcal{A}$ [1].

La validité générale de ce théorème se reconnaît le plus facilement si l'on part de mon théorème, suivant lequel, quand on a une suite de grandeurs réelles :

$$\omega_1, \omega_2, \ldots, \omega_\nu, \ldots, \ldots,$$

il y a dans tout intervalle (α, β) des grandeurs η qui ne sont égales à aucun ω_ν. En effet, supposons $n = 2$, donc A plan ; on peut alors tout d'abord joindre les points N et N′ par une ligne continue L sans se préoccuper de (M_ν), puis prendre sur L un nombre fini de points $N_1, N_2, \ldots, N_k$ n'appartenant pas à (M_ν) de telle sorte que les segments NN_1, N_1N_2, ..., N_kN' tombent entièrement à l'intérieur de A, enfin *remplacer* ces segments par des arcs de cercle ayant les mêmes extrémités et sur lesquels ne se trouve aucun point de (M_ν).

[a] *Un domaine continûment connexe n fois étendu* : en allemand *ein n-fach ausgedehntes stetig zusammenhängendes Gebirt.*

[b] *Ensemble de points* : en allemand *Punctmenge.*

[c] *Parsout dense ;* en allemand *überalldicht verbreitet.*

[1] [Publié sous une forme analogue dans le mémoire « Sur les multiplicités infinies linéaires de points », § 3 (*Ueber unendliche lineare Punktmannigfaltigkeiten*, § 3). *Math. Ann.*, XX (*Œuvres complètes*, p. 154).]

Prenons par exemple le segment NN_1 ; par les points N et N_1 passe une famille de cercles simplement infinie dont les centres se trouvent sur une droite g et peuvent être déterminés sur celle-ci par une coordonnée u. On peut déterminer un intervalle (α, β) tel que les arcs de cercle correspondant aux valeurs de u comprises dans cet intervalle se trouvent entièrement à l'intérieur de A. Aux cercles de la famille qui passent non seulement par N et N_1 mais aussi par les points

$$M_1, M_2, M_3, \ldots, M_\nu, \ldots$$

correspondent des valeurs de u que l'on désignera par

$$\omega_1, \omega_2, \ldots, \omega_\nu, \ldots$$

Si l'on prend alors à l'intérieur de (α, β) une valeur η qui ne soit égale à aucun ω_ν, on obtient alors, en faisant $u = \eta$, un arc de cercle joignant les points N et N_1, entièrement contenu dans A, et sur lequel *ne se trouve aucun point* M_ν, C. Q. F. D.

W. Berlin, le 15 avril 1882

J'aimerais connaître votre réaction devant le contenu mathématique de ma dernière lettre, à supposer qu'elle soit parvenue entre vos mains, ce dont je doute.

Voici pourquoi je m'y intéresse : à la suite de nos recherches, et des résultats auxquels nous sommes parvenus, vous et moi, indépendamment l'un de l'autre, depuis pas mal de temps, concernant le concept de nombre irrationnel et sa nature idéale, au sens de Kummer, il était établi que, pour construire le concept d'espace, il n'y a aucune nécessité interne de se représenter ce dernier comme partout continu ; vous attirez expressément l'attention sur ce point dans votre mémoire sur la continuité. Il semblait alors naturel de déduire la continuité d'un espace de raisons extérieures, notamment de la possibilité du déplacement continu, c'est-à-dire du mouvement, et c'était là en effet pendant assez longtemps mon opinion.

Mais celle-ci devient maintenant insoutenable, puisque dans un espace aussi discontinu partout que celui que j'ai désigné par $\mathcal{A}$ dans ma lettre, un déplacement continu de n'importe quel

point jusqu'à n'importe quel autre est possible ; et l'on pourrait donc concevoir une mécanique modifiée, qui serait valable pour des espaces $\mathcal{A}$?..

Halle sur la Saale, 15 septembre 1882

... Ci-joint (1) une tentative de formulation d'une question qui m'intéresse depuis longtemps, savoir ce qu'il faut entendre sous le nom de *continu* ; si vous ne le trouvez pas inutile, nous pourrons la discuter de vive voix.

J'ai essayé de généraliser votre concept de coupure et de l'employer pour la définition générale du continu, mais je n'y ai pas réussi. Par contre, mon point de départ, les « suites fondamentales » (a) dénombrables (par « suites fondamentales », j'entends des suites dont les éléments s'approchent indéfiniment *l'un de l'autre*), semble s'adapter sans difficulté à cette tentative.

Pour des ensembles de points (b) linéaires, c'est-à-dire contenus dans une droite, je peux montrer facilement, avec l'aide de mes théorèmes, que seul un intervalle complet satisfait aux conditions A et B (+).

De même, on devrait arriver à démontrer, pour des ensembles de points qui se trouvent dans un plan, que si A et B sont remplies, ce sont, soit des lignes fermées d'un seul tenant, soit des surfaces limitées par de telles courbes (+) ...

(+) à supposer que l'on prenne pour $[m, n]$ la distance des deux points m, n.

P. S. : Un ensemble dénombrable ne peut sans doute être considéré comme un continu relativement à aucun ordre. Par contre, tout ensemble non dénombrable doit vraisemblablement pouvoir être considéré comme un continu relativement à certains ordres.

I (2). Soit M *n'importe quel* ensemble (c) bien défini se composant d'une infinité d'éléments m, n, p.

A toute combinaison de deux éléments m et n de l'ensemble, faisons correspondre d'une manière déterminée (*mais complè-*

(1) [Voir à la fin de la lettre après le post-scriptum.]
(a) *Suites fondamentales* : en allemand *Fundamentalreihen.*
(b) *Ensembles de points* : en allemand *Punctmannichfaltigkeiten.*
(c) *Ensemble* : en allemand *Mannigfaltigkeit.*
(2) [Feuille séparée, à laquelle Cantor se réfère au début de cette lettre et de la suivante.]

tement arbitraire) un nombre réel positif (*différent de zéro*), qui sera désigné par $[m, n]$; il peut être considéré en quelque sorte comme fonction de la combinaison $[m, n]$. Réciproquement, cette fonction $[m, n]$ introduira certaines relations entre les éléments de M, de telle sorte que m, n, p, ..., qui sont originellement dépourvus de rapports mutuels, apparaissent maintenant mis dans un certain ordre 0 ; nous appellerons cet ordre 0 l'ordre induit [a] sur M par la fonction $[m, n]$. Toute autre fonction $[m, n]'$, pourvu qu'elle diffère de $[m, n]$ par un terme au moins, induira un autre ordre 0′ sur le *même ensemble* M.

II. M sera appelé *continu*, *relativement* à l'ordre 0 induit par la fonction $[m, n]$, si les deux conditions suivantes sont remplies :

A. m, n étant deux éléments quelconques de M et ε une grandeur donnée quelconque, on peut toujours trouver des éléments $m_1, m_2, \ldots, m_k$ de M en nombre fini tels que :

$$[m, m_1], [m_1, m_2], \ldots, [m_{k-1}, m_k], [m_k, n],$$

soient tous plus petits que ε.

B. $m_1, m_2, \ldots, m_\nu$ étant un ensemble infini dénombrable quelconque d'éléments de M ayant la propriété suivante :

$$\lim [m_{\nu+\mu}, m_\nu] = 0 \quad \text{pour} \quad \nu = \infty,$$

alors il y a toujours *un* élément m de M et *un seul* tel que :

$$\lim [m, m_\nu] = 0 \quad \text{pour} \quad \nu = \infty.$$

III. Il résulte clairement de ce qui précède qu'un même ensemble M peut constituer un continu relativement à *un* ordre 0, un discontinu relativement à un *autre* ordre 0′. Ainsi un carré, sa frontière incluse, est un continu relativement à l'ordre pour lequel $[m, n]$ est la distance en ligne droite des points m et n ; le même carré est au contraire un discontinu relativement à l'ordre que l'on obtient si on l'applique de façon biunivoque [b] et complète sur un *segment de droite*, de telle sorte qu'aux points m, n, p, ..., correspondent les points m', n', p',

(a) *Induit* : en allemand *bewirkte*.
(b) *Biunivoque* : en allemand *Gegenseitig eindeutig*.

... du segment, et que l'on prenne $[m, n]$ égal à la distance des deux points correspondants $[m', n']$.

Halle, 15 sept. 82

Halle sur la Saale, 30 septembre 1882

... Si le feuillet sur le concept du continu vous tombe sous la main, n'oubliez pas de biffer le dernier passage, car il repose sur une erreur.

Le carré est visiblement encore un continu avec l'ordre que j'y considérais, seulement c'est un continu à une dimension.

Il serait par contre facile d'établir un autre ordre, avec lequel le carré serait un discontinu. — Les idées fondamentales sont que l'on ne peut parler d'un continu que *relativement* à un ordre déterminé des éléments induit par une fonction $[m, n]$; puis les *deux* propriétés caractéristiques du continu, relativement à un ordre donné, exposées *en A* et *B*. J'espère trouver le temps très bientôt d'élucider davantage ces idées, et peut-être de les publier.

(Carte, timbre de la poste 2.10.82)

Je constate qu'il faut, aux conditions *A* et *B* du continu *relativement* à un *ordre donné* 0, en ajouter encore une *C*, qui n'est pas du tout en général une conséquences des précédentes ; à savoir :

C. (Réciproque de *B*). Si m est un élément de M, m_ν un ensemble dénombrable infini d'éléments tels que :

$$\lim [m, m_\nu] = 0 \quad \text{pour} \quad \nu = \infty,$$

on doit alors toujours avoir : $\lim [m_{\nu+\mu}, m_\nu] = 0$ pour $\nu = \infty$ et μ quelconque.

Mais, avec cela, on devrait avoir épuisé tout ce qui peut être exigé d'un *continu*.

Halle sur la Saale, 5 novembre 1882

... La lacune m'est d'autant plus sensible que, depuis bon nombre d'années, je me suis habitué à soumettre à votre jugement éprouvé mes expériences intérieures mathématiques, et, juste-

ment, depuis nos dernières rencontres à Harzburg et à Eisenach, il a plu à Dieu tout-puissant que je parvienne aux révélations les plus étonnantes et les plus inattendues dans la théorie des ensembles ([a]) et la théorie des nombres, ou plutôt que je trouve ce qui, depuis des années, a fermenté en moi, ce que j'ai longtemps cherché. — Il ne s'agit pas de la définition générale d'un continu ponctuel ([1]), dont nous avons parlé, et pour laquelle je crois avoir progressé, mais de quelque chose de beaucoup plus général et donc plus important. —

Vous vous souvenez qu'à Harzburg je vous ai dit ne pas pouvoir démontrer le théorème suivant :

« Si M′ est une partie ([b]) d'un ensemble ([c]) M, M″ une partie de M′, et si M et M″ peuvent être mis en correspondance biunivoque l'un avec l'autre, c'est-à-dire si M et M″ ont la même puissance, alors M′ a lui aussi la même puissance que M et M″. »

J'ai maintenant trouvé la source de ce théorème, et je puis le démontrer rigoureusement et avec la généralité nécessaire, ce qui comble une lacune essentielle dans la théorie des ensembles.

J'y arrive par une extension naturelle, une continuation de la suite des nombres entiers réels, qui me conduit successivement avec la plus grande sécurité aux puissances ascendantes, dont la définition précise m'a manqué jusqu'à ce jour, à l'exception de la première, celle qui est donnée par la suite des nombres 1, 2, 3, ..., ν, ... —

Je nomme la suite 1, 2, 3, ..., ν, ... première classe de nombres entiers réels, et la désigne par (ν). —

J'arrive comme suit à la *deuxième* classe de nombres entiers réels :

de même que le nombre ν exprime que l'on pose l'unité un nombre ([d]) déterminé de fois et que l'on réunit ces unités, je crée d'abord un nouveau nombre ω, qui doit exprimer que l'en-

([a]) *Théorie des ensembles* : en allemand *Mannigfaltigkeitslehre.*

([1]) [Cf. le travail daté d'octobre 1882, qui traite ces mêmes thèmes de façon détaillée : *Sur les ensembles de points linéaires infinis* (*Ueber unendliche lineare Punktmannigfaltigkeiten*), n° 5. *Math. Ann.*, 21 (*Œuvres complètes*, p. 165).]

([b]) *Partie* : en allemand *Bestandtheil.*

([c]) *Ensemble* : en allemand *Mannigfaltigkeit.*

([d]) *Nombre* : en allemand ici *Anzahl.*

semble (a) (ν) tout entier est donné ; je peux me représenter ω comme la limite des nombres ν, si par là on entend seulement que ω est le premier nombre entier créé après *tous* les ν, c'est-à-dire le *premier* qui doive être appelé plus grand que tous les ν.

Si j'applique maintenant l'addition d'une unité à ω, comme précédemment à ν, j'obtiens un nouveau nombre $\omega + 1$, qui exprime que ω est d'abord posé, *puis* l'unité ajoutée et réunie avec ω en un nouveau nombre. Je nomme *premier mode de génération* la transition d'un nombre ν ou ω au nombre *consécutif* ; par contre, *deuxième mode de génération* la transition d'un ensemble ininterrompu de nombres entiers sans plus grand élément au nombre consécutif plus grand qu'eux tous.

La formation du nombre ω résulte donc du *deuxième* mode de génération, celle du nombre $\omega + 1$ du *premier*.

Si l'on applique alors ces *deux* modes de façon répétée, on aboutit à une extension de notre suite de nombres qui se poursuit dans un ordre de succession déterminé :

$$1, 2, 3, \ldots, \nu, \ldots \quad \omega, \omega + 1, \ldots, \omega + \nu, \ldots \quad 2\omega, 2\omega + 1, \ldots,$$
$$\mu\omega + \nu, \ldots \quad \lambda\omega^2 + \mu\omega + \nu, \ldots,$$
$$\sigma\omega^k + \rho\omega^{k-1} + \ldots + \mu\omega + \nu, \ldots$$
$$\ldots \text{ etc., etc.}$$

La première impression que vous fera cette suite sera sans doute que l'on ne voit pas comment, en la continuant, on peut arriver à une sorte d'arrêt, qui serait pourtant nécessaire, si nous devons obtenir par ce moyen une *nouvelle* puissance *déterminée*, à savoir la puissance de la *deuxième* classe, immédiatement *consécutive* à la puissance de la première classe. —

Pour réaliser cet arrêt, un *troisième* mode, que je nomme mode de *limitation*, vient s'ajouter aux deux modes de génération définis plus haut ; ce troisième mode consiste en l'exigence de ne créer un nouveau nombre entier à l'aide de l'un des deux autres modes que si la totalité des nombres antécédents *est dénombrable* dans une classe de nombres déjà existante et connue dans toute son étendue.

Dans cette voie l'on peut, en suivant tous ces trois modes, parvenir avec la plus grande sécurité à des classes de nombre tou-

(a) *Ensemble* : en allemand ici *Inbegriff*.

jours nouvelles et à leurs puissances, et les nouveaux nombres obtenus de cette manière sont alors toujours déterminés d'une façon aussi concrète et ont la même réalité que les anciens. Je ne sais donc vraiment pas ce qui pourrait nous retenir dans cette activité de former de nouveaux nombres entiers, dès qu'il s'avère que, pour le progrès de la science, l'introduction d'une de plus parmi ces innombrables classes de nombres est devenue souhaitable ou même indispensable ; et c'est ce qui me paraît être le cas dans la théorie des ensembles, et peut-être aussi dans un domaine beaucoup plus vaste ; du moins, je ne peux plus avancer *sans* cette extension, *avec* elle j'obtiens beaucoup de résultats complètement inattendus. —

Il nous faut naturellement tout d'abord en rester à la *deuxième* classe de nombres, que je nomme (α) ; elle contient à son début la première classe (ν). D'après ce qui précède, (α) est caractérisée par le fait que les nombres $1, 2, 3, \ldots \omega, \omega + 1, \ldots \omega^2, \ldots$ qui *précèdent* un nombre déterminé α forment *toujours* un ensemble qui (abstraction faite de son ordination naturelle) est dénombrable dans la *première* classe.

On peut alors démontrer en toute rigueur que l'ensemble [a] (α) *n'est pas lui-même* dénombrable dans la première classe.

L'ensemble [a] de nombres (α) a donc une *autre* puissance que celle de (ν), celle qui est *immédiatement* supérieure ; car je peux démontrer rigoureusement le théorème suivant :

« Si (α') est une partie quelconque de (α), alors, *ou bien* (α') est fini, *ou bien* elle est dénombrable dans la première classe, *ou bien* elle peut être mise en correspondance univoque et réciproque avec l'ensemble [a] (α) lui-même, c'est-à-dire que (α') est *dénombrable* dans la *deuxième* classe ; quartum non datur. »

De plus, je crois pouvoir démontrer rigoureusement que l'ensemble [a] de tous nos nombres *réels*, rationnels *et* irrationnels, peut être mis en correspondance univoque avec (α), ce qui, compte tenu du théorème précédent, établit le *théorème des deux classes* pour les ensembles linéaires infinis (ou ceux qui s'y ramènent par une application). —

Peut-être vous étonnerez-vous de la hardiesse que j'ai de nommer eux aussi nombres *entiers* les objets $\omega, \omega + 1, \ldots, \alpha, \ldots,$ et précisément nombres *entiers réels* de la deuxième classe, alors

[a] *Ensemble* : en allemand *Inbegriff*.

que, jusqu'à présent, lorsque je m'en suis servi (dans les *Annalen*, vol. 17, 20 et 21), je les ai appelées modestement : symboles d'infinité.

Mais la liberté que je prends s'explique par cette remarque que, entre les objets de pensée α que je nomme nombres entiers réels de la deuxième classe, existent des relations qui peuvent se ramener aux opérations fondamentales. —

Il est vrai que les lois auxquelles ces opérations obéissent sont essentiellement autres et plus compliquées, plus difficiles à trouver par induction que dans notre théorie des nombres, celle qui se rapporte aux anciens nombres. Déjà pour l'addition on trouve que la loi de *commutativité* ne s'applique pas en général ; $\alpha + \beta$ ***n'est pas*** en général $= \beta + \alpha$.

On s'en aperçoit facilement en constatant que $1 + \omega = \omega$, alors que $\omega + 1$ est bien distinct de ω.

Lorsque $\beta > \alpha$, on définit la soustraction par l'équation :

$$\alpha + (\beta - \alpha) = \beta,$$

mais *pas* par l'équation : $(\beta - \alpha) + \alpha = \beta$, qui n'est pas en général résoluble en $(\beta - \alpha)$. Si α est le multiplicande, β le multiplicateur, il y a aussi un nombre déterminé dans (α) qui est le produit :

$$\alpha\beta ;$$

mais ici aussi on trouve que $\beta\alpha$ est généralement différent de $\alpha\beta$.

Cependant, la loi d'*associativité* est encore valable dans la deuxième classe, tant pour l'addition que pour la multiplication ; on a :

$$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$$
$$\alpha.(\beta.\gamma) = (\alpha.\beta).\gamma.$$

Une moitié seulement du principe de distributivité est valable ; on a :

$$(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma.$$

Parmi les nombres de (α), une distinction naturelle s'établit entre les nombres α qui ont été engendrés au moyen du *premier* mode de génération, ceux-là donc qui ont un *précédent*, que

je nomme α_{-1} (il est en général distinct de $\alpha - 1$, car $\alpha - 1 = \alpha$ pour $\alpha \geqq \omega$), et les nombres α qui n'ont aucun précédent dans la suite, et pour lesquels par conséquent α_{-1} n'a pas de sens. Parmi les nombres de première espèce, il y a lieu de distinguer des autres ceux qui ne sont pas décomposables, et qu'on pourrait bien appeler nombres premiers (1). —

Pour autant que je puisse voir, nos nombres irrationnels finis peuvent se déterminer de manière relativement simple avec l'aide des nombres α ; je vais poursuivre mes recherches sur ce point.

Appelons P l'ensemble des nombres contenus dans la formule :

$$z = \alpha_1 \cdot \frac{1}{3} + \alpha_2 \cdot \frac{1}{3^2} + \ldots + \alpha_\nu \cdot \frac{1}{3^\nu} + \ldots$$

dans laquelle les α_ν ne prennent que les valeurs 0 et 2.

P a alors les deux propriétés suivantes :

1) $P \equiv P'$.

2) P n'est partout dense dans aucun intervalle.

Problème : Soit M un ensemble (a) quelconque d'éléments ;

(1) Il me semble (mais je n'en suis pas encore tout à fait sûr) (1) que tout nombre α est représentable d'une manière unique sous la forme :

$$\alpha = c_0 . \pi . c_1 . \pi' . c_2 . \pi'' \ldots c_{\nu-1} . \pi^{(\nu-1)} . c_\nu . \rho$$

où $c_0, c_1, \ldots, c_{\nu-1}, c_\nu$ sont des nombres entiers positifs de (ν), π, π', π'', ... des nombres premiers de (α), enfin ρ un nombre de (α) de deuxième espèce (donc tel que ρ_{-1} n'existe pas), et tel qu'il ne soit *divisible par aucun nombre de première espèce.* Ces nombres de deuxième espèce ont un caractère bien particulier, car par exemple $\omega = \omega . \nu$, où ν est un nombre de (ν). C'est pourquoi il ne peut pas être question d'une décomposition *déterminée* de ρ. — C'est pourquoi l'on ne peut pas non plus parler réellement de nombres premiers de *deuxième* espèce.

Il faut encore remarquer que, lorsque je dis que α est divisible par β, j'affirme que l'équation

$$\alpha = \beta\gamma,$$

où β est le *multiplicande*, est possible. Cette détermination du concept de divisibilité me semble nécessaire dans la classe de nombres (α).

(a) *Ensemble* : en allemand *Mannigfaltigkeit.*

(1) [On sait que cette décomposition a été pour la première fois exposée et démontrée dans les *Contributions à la fondation de la théorie des ensembles* (*Beiträge zur Begründung der Mengenlehre*), travail publié en 1897 (*Œuvres complètes*, p. 343).]

M_1 une vraie partie de M, M_2 une vraie partie de M_1 ; on suppose qu'il existe une relation biunivoque entre M et M_2 ; il faut démontrer que M et M_1, et par conséquent aussi M_1 et M_2 ont la *même puissance* ?

(Carte : timbre de la poste 6.11.82)

Je crois que j'ai fait une faute matérielle qui pourrait être l'occasion d'une erreur dans mon analyse d'hier.

Dans le produit

$$\alpha = \beta\gamma,$$

c'est β qui est pour moi le *multiplicateur*, γ le multiplicande, je dis que α est divisible par β si α peut-être mis sous forme d'un produit dans lequel β est le *multiplicateur*.

G. C.

En ce sens doit aussi être comprise la décomposition que je crois universellement valable:

$$\alpha = c_0\pi c_1\pi' \ldots c_{\nu-1}\pi^{(\nu-1)}\, c_\nu . \rho.$$

CANTOR A DEDEKIND

Halle, le 28 juillet 1899

... Vous savez que, depuis beaucoup d'années déjà, je suis arrivé à une suite bien ordonnée de puissances ou nombres cardinaux transfinis que je nomme les « alephs » :

$$\aleph_0, \aleph_1, \aleph_2, \ldots \aleph_{\omega_0}, \ldots$$

$\aleph_0$ signifie la puissance des ensembles « dénombrables » au sens usuel, $\aleph_1$ est le nombre cardinal suivant par ordre de grandeur.

$\aleph_2$ est encore le suivant par ordre de grandeur, etc. ; $\aleph_{\omega_0}$ est le premier qui suive tous les $\aleph_\nu$ (c'est-à-dire le suivant par ordre de grandeur), égal à

$$\lim_{\nu \to \omega_0} \aleph_\nu,$$

etc.

La grande question était de savoir s'il existe, en dehors des alephs, encore d'autres puissances d'ensembles ; il y a déjà deux ans que je suis en possession d'une démonstration de ce qu'il n'en existe pas, si bien que, par exemple, un aleph déterminé correspond, comme nombre cardinal, au continu linéaire arithmétique (la totalité [a] de tous les nombres réels).

En partant du concept d'une multiplicité [b] déterminée d'objets (un système, une classe [c]), il m'est apparu nécessaire de distinguer deux sortes de multiplicités (il s'agit toujours de multiplicités *définies*).

En effet, une multiplicité peut être ainsi constituée que l'hypothèse d'une « existence simultanée » [d] de *tous* ses éléments conduit à une contradiction, de telle sorte qu'il est impossible de concevoir cette multiplicité comme une unité, comme « un objet achevé ». Je nomme de telles multiplicités *multiplicités absolument infinies* ou *inconsistantes*.

Par exemple, on se persuadera facilement que la « classe [c] de tout ce qui est pensable » est une telle multiplicité ; d'autres exemples se présenteront plus tard.

Si, au contraire, la totalité [a] des éléments d'une multiplicité peut être pensée comme « existant simultanément » [d], de telle sorte qu'il soit possible de la concevoir comme « *un seul objet* », je la nomme une *multiplicité consistante* ou un « ensemble » [e]. (En français et en italien ce concept s'exprime avec précision par les mots « ensemble » [f] et « insieme »).

Deux multiplicités équivalentes sont, ou toutes deux des « ensembles », ou toutes deux inconsistantes.

Toute multiplicité partie d'un ensemble est un ensemble.

Tout ensemble d'ensembles, si on décompose ces derniers en leurs éléments, est aussi un ensemble.

Étant donné un ensemble M, le concept général qui est propre à cet ensemble [g] et à tous les ensembles équivalents, et

[a] *Totalité* : en allemand *Gesamtheit.*
[b] *Multiplicité* : en allemand *Vielheit.*
[c] *Classe* : en allemand *Inbegriff.*
[d] « *Existence simultanée* » : en allemand « *Zusammensein* » (les guillemets sont dans le texte).
[e] « *Ensemble* » : en allemand « *Menge* » (guillemets dans le texte).
[f] En français dans le texte.
[g] *Le concept général qui est propre à cet ensemble* : en allemand *der Allgemeinbegriff, welcher dieser Menge zukommt.*

à eux seuls, je l'appelle son *nombre cardinal* ou encore sa *puissance*, et je le désigne par m. J'arrive maintenant de la manière suivante au système de toutes les puissances dont on verra plus tard que c'est une multiplicité inconsistante.

Une multiplicité s'appelle « simplement ordonnée » quand il existe entre ses éléments un ordre tel que, de deux quelconques de ses éléments, l'un soit antérieur, l'autre ultérieur, et que, de trois quelconques de ses éléments, l'un soit le premier, un autre au milieu, et que celui qui reste soit le dernier par son rang.

Si la multiplicité simplement ordonnée est un *ensemble*, j'entends par son *type* μ le concept général qui englobe cet ensemble ainsi que tous les ensembles ordonnés *semblables*, et eux seuls. (Le concept de *similitude* est employé par moi dans un sens plus restreint que chez vous ; je nomme *semblables* deux multiplicités simplement ordonnées, si elles peuvent être mises en correspondance biunivoque, de telle sorte que la relation entre les rangs soit la même pour des éléments correspondants).

Une multiplicité est dite bien ordonnée si elle satisfait à la condition que chacune de ses *parties* a un *premier* élément ; pour abréger, j'appelle une telle multiplicité une « suite » (a).

Chaque partie d'une « suite » est une « suite ».

Si maintenant une suite F est un ensemble, le type de F, je l'appelle son « *nombre ordinal* » ou plus brièvement son « *nombre* » ; donc, dans la suite, si je parle de nombres, sans plus, il s'agira de nombres ordinaux, c'est-à-dire de types d'ensembles bien ordonnés.

Je considère maintenant le système de tous les nombres et le désigne par Ω.

Dans les *Math. Annalen*, vol. 49, p. 216, (*Œuvres complètes*, III, 9, p. 320), il est établi que, si l'on a deux nombres différents α et β, il y en a toujours un qui est le plus petit, l'autre le plus grand, et que, si l'on a trois nombres α, β, γ, tels que $\alpha < \beta$, $\beta < \gamma$, alors $\alpha < \gamma$.

Ω est donc un système simplement ordonné.

Mais il résulte aussi facilement des théorèmes démontrés au § 13 sur les ensembles bien ordonnés que toute multiplicité de nombres, c'est-à-dire toute partie de Ω, contient *un plus petit nombre.*

(a) « *Suite* » : en allemand « *Folge* » (les guillemets sont dans le texte).

Le système Ω *constitue donc avec son ordre de grandeur naturelle une « suite ».*

Adjoignons encore à cette suite comme élément le 0, et plaçons-le à la première place ; nous obtenons une suite Ω′ :

$$0, 1, 2, 3, \ldots, \omega_0, \quad \omega_0 + 1, \ldots, \quad \gamma, \ldots ;$$

on se convaincra facilement que *chaque* nombre γ de cette suite est le *type de la suite de tous les éléments précédents* (0 compris). (La suite Ω n'a cette propriété qu'à partir de $\omega_0 + 1$.)

Ω′ ne peut être une multiplicité *consistante* (et par conséquent Ω non plus) ; si Ω′ était consistante, en tant qu'ensemble bien ordonné un nombre δ lui serait associé, qui serait plus grand que tous les nombres du système Ω ; mais le nombre δ doit être un élément du système Ω qui contient *tous* les nombres ; δ serait alors plus grand que δ, ce qui est contradictoire. Donc : A. *Le système* Ω *de tous les nombres est une multiplicité inconsistante, absolument infinie.*

Comme la *similitude* des ensembles bien ordonnés entraîne en même temps leur *équivalence*, à chaque nombre γ correspond un nombre cardinal déterminé $\aleph(\gamma) = \overline{\gamma}$, à savoir le nombre cardinal de l'ensemble bien ordonné dont le type est γ.

Les nombres cardinaux qui correspondent en ce sens aux nombres *transfinis* du système Ω, je les nomme des « alephs », et j'appelle $\beth$ (*tau*, dernière lettre de l'alphabet hébreu) le *système de tous les alephs.*

Le système de tous les nombres γ qui correspondent à un même nombre cardinal $\mathfrak{c}$, je le nomme une « *classe de nombres* », la classe des nombres $Z(\mathfrak{c})$. On voit facilement qu'il y a, dans chaque classe de nombres, un plus petit nombre γ_0, et qu'il y a un nombre γ_1, situé à l'extérieur de $Z(\mathfrak{c})$, et tel que la condition

$$\gamma_0 \leqq \gamma < \gamma_1$$

équivaut à l'appartenance du nombre γ à la classe de nombres $Z(\mathfrak{c})$. Chaque classe de nombres est donc un « segment » déterminé de la suite Ω [1].

[1] Ici est utilisé de nouveau le théorème déjà mentionné plus haut suivant lequel *chaque* classe de nombres, donc *chaque* multiplicité partie de Ω a un *minimum*, un *plus petit nombre.*

Certains nombres du système Ω forment chacun pris isolément une classe de nombres, ce sont les nombres « finis » 1, 2, 3, ..., ν, ..., auxquels correspondent les divers nombres cardinaux « finis »

$$\bar{1}, \bar{2}, \bar{3}, \ldots, \bar{\nu}, \ldots$$

Soit ω_0 le plus petit nombre transfini ; je nomme $\aleph_0$ l'aleph correspondant, de telle sorte que

$$\aleph_0 = \bar{\omega}_0 ;$$

$\aleph_0$ est *le plus petit* aleph et détermine la classe de nombres

$$Z(\aleph_0) = \Omega_0.$$

Les nombres α de $Z(\aleph_0)$ remplissent la condition

$$\omega_0 \leqslant \alpha < \omega_1,$$

et cette condition les caractérise ; ω_1 est ici le plus petit nombre transfini dont le nombre cardinal n'est pas égal à $\aleph_0$. Si l'on pose

$$\bar{\omega}_1 = \aleph_1,$$

non seulement $\aleph_1$ est distinct de $\aleph_0$, mais c'est l'aleph immédiatement supérieur, car l'on peut démontrer qu'il n'y a aucun nombre cardinal entre $\aleph_0$ et $\aleph_1$. On obtient ainsi la classe de nombres $\Omega_1 = Z(\aleph_1)$ immédiatement contiguë à Ω_0. Elle comprend tous les nombres β qui satisfont à la condition

$$\omega_1 \leqslant \beta < \omega_2 ;$$

ω_2 est ici le plus petit nombre transfini, dont le nombre cardinal est différent de $\aleph_0$ et de $\aleph_1$.

$\aleph_2$ est l'aleph immédiatement supérieur à $\aleph_1$ et détermine la classe de nombres $\Omega_2 = Z(\aleph_2)$ suivant immédiatement Ω_1 ; elle est composée de tous les nombres γ qui sont $\geqslant \omega_2$ et $< \omega_3$, où ω_3 est le plus petit nombre transfini, dont le nombre cardinal est différent de $\aleph_0$, $\aleph_1$ et $\aleph_2$, etc... Je signale encore :

$$\bar{\bar{\Omega}}_0 = \aleph_1, \quad \bar{\bar{\Omega}}_1 = \aleph_2 \ldots \bar{\bar{\Omega}}_\nu = \aleph_{\nu+1},$$

$$\sum_{\nu'=0,1,2,\ldots,\nu} \aleph_{\nu'} = \aleph_\nu,$$

tout cela est facile à démontrer.

Parmi les nombres transfinis du système Ω auxquels aucun des $\aleph_\nu$ (avec ν fini) ne correspond comme nombre cardinal, il y en a de nouveau un qui est le plus petit ; nous l'appelons ω_{ω_0}, et nous obtenons ainsi un nouvel aleph :

$$\aleph_{\omega_0} = \overline{\omega}_{\omega_0},$$

qui peut aussi être défini par l'égalité

$$\aleph_{\omega_0} = \sum_{\nu\ 0,1,2,\ldots\ldots} \aleph_\nu$$

et que l'on reconnaît être le nombre cardinal *immédiatement supérieur* à tous les $\aleph_\nu$.

On voit que ce procédé de construction des alephs et des classes de nombres du système Ω qui leur correspondent est *absolument* sans limites.

B. *Le système* ℸ *de tous les alpehs*

$$\aleph_0,\ \aleph_1,\ \ldots\ \aleph_{\omega_0},\ \aleph_{\omega_0+1},\ \ldots\ \aleph_{\omega_1},\ \ldots,$$

rangé par ordre de grandeur, forme une suite semblable au système Ω, *donc aussi une suite inconsistante, absolument infinie.*

La question se pose maintenant de savoir si *ce système* ℸ contient *tous les nombres cardinaux transfinis.* Y a-t-il, en d'autres termes, un *ensemble* dont la puissance *n'est pas un aleph* ?

La réponse est *négative*, et la raison doit en être cherchée dans *l'inconsistance*, reconnue par nous, des systèmes Ω et ℸ.

Démonstration : Prenons une multiplicité déterminée V et supposons qu'*aucun aleph* ne lui corresponde *comme nombre cardinal*, nous en concluons que V doit être *inconsistante.*

Car on reconnaît facilement (a) que, avec l'hypothèse que nous avons faite, le système Ω tout entier peut être projeté dans la multiplicité V, c'est-à-dire qu'il doit exister une partie de V, soit V', qui est équivalente au système Ω.

V' est inconsistante, puisque Ω l'est, il doit donc en être de même de V [cf. plus haut].

Par conséquent, toute *multiplicité consistante* transfinie, tout ensemble transfini, doit avoir un aleph déterminé pour nombre cardinal. Donc :

C. *Le système* ℸ *de tous les alephs n'est autre que le système de tous les nombres cardinaux transfinis.*

(a) Voir Note A de Zermelo.

Tous les ensembles sont donc « *dénombrables* », en un sens élargi, en particulier tous les « *continus* ».

Nous reconnaissons en outre, à l'aide de C, l'exactitude de ce théorème, énoncé dans les *Math. Annalen*, vol. 46 (*Œuvres complètes*, p. 285) :

« Si $\mathfrak{a}$ et $\mathfrak{b}$ sont des nombres cardinaux quelconques, on a, ou bien $\mathfrak{a} = \mathfrak{b}$, ou bien $\mathfrak{a} < \mathfrak{b}$, ou bien $\mathfrak{a} > \mathfrak{b}$. »

Car les *alephs* possèdent, comme nous l'avons vu, cette propriété des grandeurs.

CANTOR A DEDEKIND

Hahnenklee, le 28 août 1899

... On doit se demander, d'où je sais que les multiplicités bien ordonnées ou suites auxquelles j'assigne les nombres cardinaux

$$\aleph_0, \aleph_1, \ldots \aleph_{\omega_0}, \ldots \aleph_{\omega_1}, \ldots$$

sont réellement des « ensembles » dans le sens du mot qui a été expliqué, c'est-à-dire des « multiplicités consistantes ». Ne serait-il pas concevable que déjà *ces* multiplicités-*ci* soient « inconsistantes », mais que la contradiction qu'il y a à supposer une « existence simultanée (a) de tous leurs éléments » ne se soit pas encore fait remarquer ? Ma réponse est que cette question peut être *étendue également aux multiplicités finies*, et que, si l'on y réfléchit exactement, on aboutit alors au résultat : même pour les multiplicités finies, *il n'y a pas* de démonstration de leur « consistance ». En d'autres termes : le fait de la « consistance » des multiplicités finies est une vérité simple, indémontrable, c'est « l'*axiome* de l'arithmétique » (dans l'ancien sens du mot). Et, de même, la « consistance » des multiplicités auxquelles j'attribue comme nombres cardinaux les alephs, c'est « l'axiome de l'arithmétique transfinie élargie ».

CANTOR A DEDEKIND

Hahnenklee, 31 août 1899

... Plaçons les « ensembles » équivalents dans une seule et même *classe* de puissance, les ensembles non équivalents dans des classes différentes, et considérons le système

(a) *Existence simultanée* : en allemand *Zusammensein*.

S de toutes les classes concevables.

Par $\mathfrak{a}$, j'entends à la fois le nombre cardinal et la puissance des ensembles de la classe considérée, qui est la même pour tous ces ensembles.

Soit $M_{\mathfrak{a}}$ n'importe quel ensemble déterminé de la classe $\mathfrak{a}$.

J'affirme que le système déterminé bien défini S n'est *pas un « ensemble »*.

Démonstration : Si S était un ensemble, il en serait de même de

$$T = \Sigma M_{\mathfrak{a}},$$

la somme étant étendue à toutes les classes $\mathfrak{a}$; T appartiendrait alors à une classe déterminée, disons à la classe $\mathfrak{a}_0$.

Mais nous avons le théorème suivant :

« Si M est un ensemble quelconque de nombre cardinal $\mathfrak{a}$, on peut toujours en déduire un autre ensemble M', dont le nombre cardinal $\mathfrak{a}'$ est plus grand que $\mathfrak{a}$. »

J'ai démontré ce théorème dans le *premier* volume des Comptes rendus [a] de l' « Association des mathématiciens allemands » (*Œuvres complètes*, p. 278), dans les cas qui sont pour nous les plus familiers, savoir lorsque $\mathfrak{a}$ est égal à $\aleph_0$ (dénombrabilité au sens usuel du terme), ou est égal à $\mathfrak{c}$, $\mathfrak{c}$ étant la puissance du continu arithmétique, et cela *par un procédé uniforme*. Ce procédé se laisse étendre sans aucune difficulté à un $\mathfrak{a}$ quelconque. La signification de cette méthode s'exprime simplement par la formule

$$2^{\mathfrak{a}} > \mathfrak{a}.$$

Soit donc $\mathfrak{a}'_0$ n'importe quel nombre cardinal plus grand que $\mathfrak{a}_0$. Alors T, de puissance $\mathfrak{a}_0$, contient comme partie l'ensemble $M_{\mathfrak{a}'_0}$, de puissance $\mathfrak{a}'_0$ plus grande, ce qui est contradictoire.

Donc le système T, et par suite aussi le système S *ne sont pas des ensembles*. Il existe donc des multiplicités déterminées, qui *ne sont pas en même temps des unités*, c'est-à-dire des multiplicités telles qu'une réelle « existence simultanée [b] de tous

[a] *Comptes rendus de l' « Association des mathématiciens allemands »* : en allemand *Berichte der « Deutschen Mathematikervereinigung »*.

[b] *Existence simultanée* : en allemand *Zusammensein*.

leurs éléments » est *impossible*. Ce sont celles-là que je nomme « système inconsistants », et les autres « ensembles ».

DEDEKIND A CANTOR

29 août 1899

Théorème de la théorie des systèmes.

Si le système U est une partie du système T, et celui-ci une partie du système S, et si S est semblable à U, S est alors aussi semblable à T.

Démonstration [a] : Le théorème est trivial si T est identique à S ou à U. Dans le cas contraire, où T est une vraie partie de S, soit A le système de tous les éléments de S qui ne sont pas contenus dans T. On a donc (avec la notation de Dedekind, Cantor, Schröder) :

$$S = \mathfrak{M}(A, T) = (A, T) = A + T.$$

Par hypothèse, S est semblable à la partie propre U de T ; il y a donc une application semblable φ de S dans lui-même, par laquelle S vient sur $S' = \varphi(S) = U$; soit A_0 la « chaine de A » [b] (§ 4 de mon mémoire : « Que sont et à quoi servent les nombres ? [c] »), on a :

$$A_0 = A + A_0' ;$$

comme A_0 est une partie de S, $A_0' = \varphi(A_0)$ est une partie de $S' = \varphi(S) = U$, donc A_0' est aussi une partie propre de T, par suite A et A_0' n'ont aucun élément commun, et A_0 est aussi une vraie partie de S. Soit B le système de tous les éléments de S qui ne sont pas contenus dans A_0, on a

$$S = A + T = A_0 + B, \quad T = A_0' + B,$$

[a] Voir note B de Zermelo.
[b] Voir note C de Zermelo.
[c] « *Que sont et à quoi servent les nombres ?* » : en allemand « *Was sind und was sollen die Zahlen ?* »

où A_0', comme partie de A_0, n'a aucun élément commun avec B. Définissons alors une application ψ de S, en posant :

$$\psi(s) = \varphi(s) \quad \text{ou} \quad \psi(s) = s$$

suivant que l'élément s de S est contenu dans A_0 ou dans B. Cette application ψ de S est semblable ; car si s_1, s_2 sont des éléments différents de S, *ou bien* ils sont contenus dans A_0 ; alors $\psi(s_1) = \varphi(s_1)$ est différent de $\psi(s_2) = \varphi(s_2)$ parce que φ est une application semblable de S (ce qui est utilisé ici pour la première et seule fois) ; *ou bien*, ils sont contenus dans B ; alors $\psi(s_1) = s_1$ est différent de $\psi(s_2) = s_2$; *ou encore* l'un des deux éléments s_1 est contenu dans A_0, l'autre s_2 dans B ; alors $\psi(s_1) = \varphi(s_1)$ est différent de $\psi(s_2) = s_2$, parce que $\psi(s_1)$ est contenu dans A_0' et s_2 dans B.

Par cette application semblable ψ, $S = A_0 + B$ est appliqué sur

$$\psi(S) = \psi(B) + \psi(A_0) = \varphi(A_0) + B = T,$$

parce que $\psi(A_0) = \varphi(A_0) = A_0'$ et $\psi(B) = B$, C. Q. F.D.

CANTOR A DEDEKIND

Hahnenklee, 30 août 1899

Grand merci pour votre aimable lettre que j'ai reçue encore hier soir, en particulier pour l'exposé de la démonstration simple que vous avez faite, avec les moyens de votre ouvrage : « Que sont et à quoi servent les nombres », du théorème C (et par là même du théorème B) de mon mémoire (a). La forme mise à

(a) [Cantor se réfère ici à son mémoire de 1895 et 1897 : « Contributions à la fondation de la théorie des nombres transfinis ». Les énoncés des théorèmes A, B, C, D, E se trouvent page 285 des *Œuvres complètes*. Les voici (comme Cantor le rappelle lui-même un peu plus loin, pour lui « partie » veut dire « partie propre ») :

A. « *Si* $\mathfrak{a}$ *et* $\mathfrak{b}$ *sont deux nombres cardinaux quelconques, on a, ou bien* $\mathfrak{a} = \mathfrak{b}$, *ou bien* $\mathfrak{a} < \mathfrak{b}$, *ou bien* $\mathfrak{a} > \mathfrak{b}$. »

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

C. « *Si* M_1 *est une partie d'un ensemble* M, M_2 *une partie de l'ensemble* M_1 *et si* M *et* M_2 *sont équivalents,* M_1 *est équivalent aux ensembles* M *et* M_2. »

D. « *Si, pour deux ensembles* M *et* N, *la condition suivante est remplie* : N

part, cette démonstration coïncide (si je ne me trompe) avec celle communiquée d'abord par Schr¨der à l'automne 1896 à la Réunion des savants (b) de Francfort-sur-le-Main, et publiée un an et demi après dans un mémoire de la Léopoldine, que le jeune M. Félix Bernstein a redonné indépendamment au séminaire de Halle, aux alentours de Pâques 1897. Ce serait cependant *très utile* si, avec les mêmes méthodes, vous démontriez aussi le théorème principal A (dont on déduit facilement les autres B, C, D, E comme des corollaires).

Voyons clairement ce qui resterait à faire à cette fin en plus de la démonstration déjà faite de B !

D'un point de vue purement logique, avec deux ensembles M et N, on peut rencontrer *quatre* cas s'excluant mutuellement :

I. Il y a une partie de N, qui est équivalente (avec votre terminologie « semblable ») à M, et il n'y a aucune partie de M qui soit équivalente à N.

II. Il n'y a aucune partie de N qui soit équivalente à M, mais il y a une partie M_1 de M qui est équivalente à N.

III. Il y a une partie N_1 de N qui est équivalente à M, et il y a aussi une partie M_1 de M qui est équivalente à N (1).

IV. Il n'y a aucune partie de N qui soit équivalente à M, ni aucune partie de M qui soit équivalente à N.

Si l'on désigne par $\mathfrak{a}$ et $\mathfrak{b}$ les nombres cardinaux de M et N, on a, d'après la définition que j'ai donnée du « plus petit » et du « plus grand »

$$\text{dans le cas I : } \mathfrak{a} < \mathfrak{b},$$
$$\text{dans le cas II : } \mathfrak{a} > \mathfrak{b}.$$

Donc, ce qui est à démontrer, c'est que, *aussi bien* dans le cas III *que* dans le cas IV, les ensembles M et N sont équivalents, donc que $\mathfrak{a} = \mathfrak{b}$. Dans le cas III, ce résultat a été établi par vous, et par MM. Schröder et F. Bernstein, avec la démonstration di-

n'est équivalent ni à M, ni à une partie de M, il existe alors une partie N_1 de N qui est équivalente à M. »

E. « *Si deux ensembles M et N ne sont pas équivalents, et s'il existe une partie N_1 de N qui est équivalente à M, il n'y a alors aucune partie de M qui soit équivalente à N.* »]

(1) *N. B.* — Quand je parle d'une « partie », je veux toujours dire une « partie propre ».

(b) ***Réunion des savants*** : en allemand *Naturforscherversammlung*.

recte du théorème C. *Il reste donc à démontrer le théorème suivant* : « Si deux ensembles M et N sont tels que M n'est pas équivalent à une partie de M (dans votre langage une « partie propre »), ni N à une partie de M, alors M et N sont équivalents (et donc tous deux des ensembles finis). »

Schröder dit expressément qu'il ne sait pas démontrer ce théorème ; quant à moi, je n'ai pas réussi davantage à en donner une démonstration avec les *moyens simples* que vous avez utilisés pour la démonstration de C et B ; je peux *seulement* le démontrer *indirectement* à partir de A, dont j'ai esquissé la démonstration dans ma lettre du 3 août.

COMMENTAIRES ET NOTES DE ZERMELO AUX LETTRES DE 1899 (*)

La partie reproduite ici de la correspondance (jusqu'ici inédite) des deux savants forme un complément essentiel et indispensable aux travaux de Cantor, surtout parce que celui-ci y exprime ses dernières idées sur le système de *tous* les nombres ordinaux et de *tous* les alephs, ainsi que sur les totalités « consistantes » et « inconsistantes ». De même les discussions sur le « théorème de l'équivalence » et la question de la « comparabilité » d'ensembles quelconques pourraient intéresser le lecteur moderne, surtout l'essai de démonstration du théorème suivant lequel toute puissance est un aleph. La démonstration que donne Dedekind du théorème de l'équivalence (qui, il y a peu de temps, était encore inconnue, et se trouve imprimée ici pour la première fois) peut aujourd'hui encore être considérée comme classique, cependant que la tentative de Cantor d'établir que toute puissance est un aleph a visiblement été reconnue par la suite comme insuffisante par l'auteur lui-même.

A) C'est justement ici que se trouve la faiblesse de la démonstration esquissée. Que l'on puisse « projeter » dans toute multiplicité V, dont le nombre cardinal n'est pas un aleph, la suite de nombres Ω tout entière, cela est *non pas* démontré, mais tiré d'une intuition assez vague. Visiblement Cantor se représente les nombres de Ω mis en correspondance avec des éléments successifs et arbitraires de V de manière à ce que chaque élément de V ne soit utilisé qu'*une seule fois*. *Ou bien*, avec ce procédé, on parviendrait à épuiser tous les éléments de V, et V serait alors mis en correspondance avec une section de la suite des nombres : sa puissance serait donc un aleph, contrairement à l'hypothèse ; *ou* alors V resterait inépuisable et contiendrait donc une partie équivalente à Ω tout entier, donc inconsistante. Ici l'on applique

(*) *Œuvres complètes* de Cantor, page 451.

donc l'intuition du temps à un procédé qui dépasse toute intuition et l'on imagine un être, qui pourrait faire des choix *successifs* arbitraires, et par conséquent définir une partie V′ de V qui, par les conditions posées, *n'est* précisément *pas* définissable. C'est seulement par application de l' « axiome du choix », qui postule la possibilité d'un choix *simultané* et que Cantor applique partout inconsciemment et instinctivement, mais ne formule nulle part expressément, que V′ pourrait être défini comme partie de V. Même alors, il resterait toujours les scrupules engendrés par le fait que la démonstration opère avec des multiplicités « inconsistantes », peut-être même avec des concepts contradictoires et serait déjà pour cette raison logiquement inadmissible. C'est ainsi que des scrupules de cette sorte ont déterminé l'éditeur, quelques années plus tard, à fonder sa démonstration du théorème du bon ordre (*Math. Ann.*, 59, p. 524, 1904) sur l'axiome du choix seulement, *sans* utiliser des multiplicités inconsistantes.

B) La présente démonstration du théorème formulé ici, et identique en fait au « théorème d'équivalence » de Schröder-Bernstein, opère de façon purement logique avec les concepts de la théorie des « chaînes » de Dedekind, et n'est pas essentiellement différente de celle que l'éditeur a publiée en 1908, sans avoir connaissance de celle de Dedekind, dans les *Math. Annalen*, vol. 65, pp. 271-272. Pourquoi ni Dedekind, ni Cantor ne se sont décidés à l'époque à publier cette démonstration dont l'importance est certaine, c'est ce qui, aujourd'hui, n'est pas bien compréhensible.

C) Si un système (= ensemble) S est appliqué de façon biunivoque (« semblable ») sur une partie de lui-même S′, et si A est une partie quelconque de S, Dedekind entend alors par « chaîne de A » l'intersection de toutes les parties A_1 de S qui 1) contiennent A lui-même et 2) avec chacun de leurs éléments s contiennent toujours aussi l'élément image s'. Cette « chaîne » A_0 n'est alors pas autre chose que la réunion des ensembles A, A′, A″, ..., qui s'obtiennent à partir de A par l'emploi répété de l'application φ de S sur S′.

TRANSFINI ET CONTINU

AVERTISSEMENT DE L'ÉDITEUR

Cette œuvre posthume de Jean CAVAILLÈS a été écrite en 1940 ou 1941 et fut remise à la *Revue Philosophique* vers Septembre 1941. Mais elle ne put paraître alors, cette Revue ayant été interdite par les Allemands.

Nous devons cette publication à Madame LAUTMAN qui conserva pieusement l'exemplaire, annoté par CAVAILLÈS, et qu'il avait confié à son ami LAUTMAN. LAUTMAN, comme CAVAILLÉS, devait payer de sa vie sa foi en son idéal.

Nous remercions également Messieurs Henri CARTAN et CHEVALLEY, fidèles amis de Jean CAVAILLÈS, qui ont bien voulu revoir et corriger les épreuves de cette étude.

Le problème du continu reste comme au temps de Leibniz, une « croix » ou la croix de la philosophie mathématique.

Il semblait que l'arithmétisation effectuée au XIX^e siècle réussissait à en donner une définition dépourvue de tout élément intuitif. Mais, pour une critique plus avertie, le doute subsiste d'une part si une telle prétention est justifiée, c'est-à-dire si l'effectif de la définition n'utilise pas subrepticement des intuitions sensibles, d'autre part s'il n'y a pas là plutôt représentation que réduction, c'est-à-dire si, en fait, ne reste comme base une image géométrique latente, fondant l'originalité et l'unité du représenté par rapport aux relations arithmétiques accidentellement réunies pour lui dans un enchaînement extrinsèque.

Si l'on veut considérer « un nombre irrationnel quelconque entre 0 et 1 » ou la totalité des nombres irrationnels de cet intervalle, l'arbitraire de la loi d'engendrement suffit-il à épuiser le contenu de la notion, c'est-à-dire les moyens qu'elle offre à une utilisation mathématique ? Il ne semble pas : on voit bien la différence avec les cas où les nombres en question sont donnés par un procédé de calcul. En fait dans l'analyse habituelle, l'imagination spatiale reste aussi bien un guide pour l'enrichissement des procédés et le développement des raisonnements qu'une garantie pour leur évidence [1]. Pourtant ç'avait été l'intention décisive des créations de Cantor et de Dedekind d'éliminer le géométrique de l'analyse grâce à la notion d'ensemble. On se rappelle la phrase célèbre de Dedekind : « Les nombres sont des libres créations de l'esprit humain.... Grâce à l'édification purement logique de la science des nombres et grâce au domaine continu des nombres qu'elle nous procure, nous sommes en état d'analyser nos relations de l'espace et du temps en les mettant en relation avec ce domaine des nombres [2]. »

Il s'en faut que le programme ait pu être rempli. D'une part les difficultés internes de la notion d'ensemble et des procédés de son utilisation, d'autre part la séparation apparue dès les premiers travaux de

1. Cf. les remarques de P. BERNAYS à ce sujet dans les *Entretiens de Zurich* (déc. 1938, pp. 135-136 et 152).
2. *Was sind und was sollen die Zahlen* ? Préface de la 1^re éd., p. 1.

Cantor entre théorie cardinale et théorie ordinale en ont rendu longtemps l'exécution problématique. La critique contemporaine dissocie l'axiome de choix et l'hypothèse du continu du groupe des autres axiomes de la théorie ; et par ailleurs, prise des mêmes scrupules devant les démarches naïves que Dedekind confondait sous les vocables pour lui apparentés de logique et d'arithmétique, elle s'efforce, en continuant son œuvre, de préciser des critères applicables à une pensée contrôlante dont le développement canonisé servirait d'assurance à une partie des mathématiques. Telle est la direction dans laquelle des travaux récents de Gödel, Church, Kleene ont réalisé de sensibles progrès. Plus récemment, par des procédés analogues quoique plus étendus, Gödel a pu démontrer qu'il n'y a pas contradiction entre le groupe B formé par l'axiome de choix et l'hypothèse du continu et le groupe A des autres axiomes, faisant ainsi décidément franchir à la philosophie mathématique une étape que nous allons essayer de marquer.

La comparaison de la puissance des ensembles apparaît dès la naissance et comme élément décisif de la théorie. Deux ensembles ont même puissance s'il existe entre leurs éléments une correspondance univoque et réciproque. Ainsi, à son grand étonnement, Cantor aperçoit que l'ensemble des points d'une droite et celui des points d'un plan ont même puissance. En revanche la puissance de l'ensemble des entiers, égale à celle de l'ensemble des nombres rationnels, est plus petite que celle de l'ensemble des nombres réels, c'est-à-dire que d'une part l'hypothèse d'une correspondance bi-univoque entre ces deux ensembles mène à une contradiction, que d'autre part le premier est une partie du second. Ce résultat particulier amène donc non seulement à définir la relation d'inégalité entre les puissances (un ensemble A est de puissance inférieure à un ensemble B lorsque d'une part l'égalité est impossible et que d'autre part A a même puissance qu'une partie de B) mais encore à construire une échelle indéfinie de puissances superposées.

Le continu ou l'ensemble des nombres réels peut être considéré comme l'ensemble des fractions décimales illimitées c'est-à-dire de suite indéfinies formées de nombres entiers. C'est donc l'*ensemble des parties* de l'ensemble des entiers (ou des fonctions y prenant les valeurs 0 et 1). Si $\aleph_0$ désigne la puissance de celui-ci, celle du continu sera $c = 2^{\aleph_0}$. Et l'on peut poursuivre en prenant l'ensemble des parties du continu (ou des fonctions numériques définies sur le continu) : il aura pour puissance : $2^c = 2^{2^{\aleph_0}} > 2^{\aleph_0}$ etc.... La structure des ensembles ainsi comparés reste inconnue : dès l'ensemble des nombres réels il n'y a aucune loi d'engendrement ; c'est la notion du *quelconque* qui

donne sa base logique à l'unité de l'ensemble et à la réalité de sa puissance. On conçoit qu'elle ait éveillé des doutes, que nombre de mathématiciens n'aient voulu voir dans le continu qu'une virtualité pure l'aient rejeté hors des existences mathématiques pour lesquelles sont valables les raisonnements classiques. Mais c'était déjà amputer une bonne partie de l'analyse, tout autant que ruiner la théorie cardinale.

Dans la même série de mémoires, en étudiant les discontinuités d'une fonction, Cantor était amené à créer la théorie ordinale pour numéroter les dérivés successifs d'un ensemble de points. L'opération fondamentale est ici celle du *nombrement* qui se définit par les étapes mêmes de son exécution progressive : si nous supposons obtenu un système de résultats, le nouvel acte consiste à poser à son tour ce système comme résultat ou nombre ordinal. Ainsi pourrait-on compter dans le fini : l'objet 2 sera la collectivité $\{1\}$ posée après 1, 3 la collectivité $\{1, 2\}$ ou $\{1,\{1\}\}$ et ainsi de suite. La collectivité sans fin $\{1, 2, 3,...\}$ définit un nouveau nombre ordinal ω_0 qui, à son tour, entre dans une collectivité $\{1, 2, 3..., \omega_0\}$, dénotée par $\omega_0 + 1$. Le processus abstrait ainsi défini ne comporte aucune limitation. En même temps il dote les ordinaux de trois propriétés caractéristiques :

1° la relation d'inclusion établit entre eux un ordre (relation non réfléchie, dissymétrique, transitive) : on voit que $\{1, 2\}$ est contenu dans $\{1, 2, 3\}$ donc 3 précède 4 etc... ; de même tout nombre fini est contenu dans $\{1, 2, 3...\}$ c'est-à-dire ω_0, donc précède ω_0 qui à son tour précède $\omega_0 + 1$. Cet ordre est un bon ordre, c'est-à-dire que toute partie, isolée par une propriété quelconque dans la suite ordonnée, possède un premier élément (on voit immédiatement que toute partie est elle-même bien ordonnée). Ainsi ω_0 est le premier élément dans la suite des ordinaux transfinis (c'est-à-dire non finis), il succède immédiatement aux ordinaux finis.

2° Tout ensemble bien ordonné est *semblable* à la suite complète, ou *segment*, des ordinaux précédant un certain ordinal (la *similitude* est une correspondance bi-univoque qui conserve l'ordre entre les éléments homologues) : il sera nombré par cet ordinal. Comme aucun segment de la suite des ordinaux n'est semblable à un autre segment de la même suite, ce nombrement pour tout ensemble ordonné est déterminé de façon univoque.

3° Les ordinaux peuvent être rangés en classes, d'après leur puissance (la puissance du système qu'ils représentent) : la classe I comprend les ordinaux finis (en réalité il devrait y avoir autant de classes que d'ordinaux finis), la classe II, les ordinaux ayant la puissance du dénombrable. A partir de la classe II, on distingue deux espèces d'ordinaux : ceux qui ont un prédécesseur immédiat, c'est-à-dire qui repré-

sentent un système ayant un dernier élément : on peut les écrire $\alpha + n$ (n étant un entier fini) et on les nomme *ordinaux de première espèce* ; les *ordinaux de seconde espèce* représentent un système n'ayant pas de dernier élément ; on les considère aussi comme limite supérieure du système qu'ils posent : ils viennent en effet dans l'ordre immédiatement après lui. Et, quand ils appartiennent à la classe II, on peut leur faire correspondre une suite, semblable à ω_0, d'ordinaux plus petits qui les caractérise de façon univoque et qui est appelée leur *suite fondamentale*. On voit aisément que la classe II a une puissance supérieure à celle de la classe I : en effet le système de ses éléments mis en ordre

$$\{\omega_0,\ \omega_0 + 1,\ldots\ n\omega_0,\ldots\ \omega_0^{\omega_0},\ldots\}$$

définit un nombre ordinal qui ne peut faire partie du système, sans quoi il y aurait deux segments différents semblables. On l'appelle ω_1 et on désigne sa puissance par $\aleph_1$: $\aleph_1 > \aleph_0$. Les ordinaux supérieurs à ω_1 et ayant pour puissance $\aleph_1$, constituent la classe III, et ainsi de suite. Les alephs : $\aleph_1, \aleph_2, \aleph_3\ldots$ ainsi engendrés forment une nouvelle échelle de puissances.

La difficulté fondamentale devant laquelle se trouvait placé Cantor était évidemment de préciser le rapport entre ces deux séries de puissances. Même départ, le dénombrable $\aleph_0$; mais dès le continu et $\aleph_1$, Cantor, malgré son acharnement à trouver une solution, devait s'avouer désarmé. L'hypothèse la plus simple était d'identifier les deux suites, en posant :

$$2^{\aleph_0} = \aleph_1 \quad \text{(hypothèse du continu)}$$

et en général

$$2^{\aleph_n} = \aleph_{n+1} \quad \text{(hypothèse générale du continu).}$$

Ainsi se trouverait unifié l'édifice des ensembles et assurée une domination de tout ensemble naturel (c'est-à-dire présenté par les problèmes d'autres théories mathématiques) grâce au procédé rationnel du nombrement. Mais dans l'état primitif de la théorie il était même impossible d'affirmer que les puissances de deux ensembles quelconques soient comparables.

Un progrès décisif pour cette question fut accompli en 1904 par Zermelo. C'est une expérience intuitive courante que des ensembles naturels puissent être ordonnés de façons différentes. Ainsi l'ensemble des nombres rationnels, rangés par ordre de grandeur croissante, est ordonné mais non bien ordonné (par ex. la partie constituée par les nombres rationnels dont le carré est $> \frac{1}{2}$ n'a pas de premier élé-

ment). Mais si, après les avoir mis sous forme de fraction irréductible, on les range d'après la grandeur de la somme du numérateur et du dénominateur, et, pour une même somme, d'après leur grandeur, on obtient un ensemble bien ordonné semblable à la suite des ordinaux finis, donc nombré par ω_0. On pourrait aussi les ranger d'après la grandeur du numérateur, puis, à numérateur égal, d'après celle du dénominateur : on obtiendrait l'ensemble bien ordonné

$$\frac{1}{1}, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots, \frac{2}{2}, \frac{2}{3}, \ldots, \frac{2}{n}, \ldots, \frac{3}{3}, \frac{3}{4}, \ldots$$

semblable non plus à ω_0 mais à ω_0^2. Le changement d'ordre ne change évidemment pas la puissance d'un ensemble. Donc si tout ensemble peut être bien ordonné, en vertu de la deuxième propriété des ordinaux, il sera semblable à un ordinal, donc aura une puissance situable dans l'échelle des alephs. Par là même se trouvera généralement assurée la comparabilité des puissances de deux ensembles quelconques.

Tel est le double résultat atteint par Zermelo, avec cependant une restriction essentielle : ses deux démonstrations (données en 1904 et 1908) supposent la possibilité, étant donné un système d'ensembles sans élément commun deux à deux, de raisonner sur un ensemble ayant avec chaque ensemble du système un élément commun et un seul. C'est l'ensemble de *choix* qui fut aussitôt pris à parti dans des discussions que le vague de la notion d'existence rendait confuses. Zermelo n'a jamais supposé pour son théorème du bon ordre qu'on puisse construire effectivement l'ensemble de choix, pas plus qu'il n'a donné le moyen de bien ordonner un ensemble. Mais, étant admis qu'on puisse considérer comme donné l'ensemble de choix au même titre que les ensembles initiaux M, il a prouvé qu'on peut considérer comme également donnée une application des éléments de M sur les parties de M, qui pose entre les éléments la relation possédant les propriétés caractéristiques du bon ordre.

Par là se trouvent manifestés en pleine lumière à la fois les réquisits de la théorie des ensembles et les difficultés de la notion générale d'existence pour les objets mathématiques. Les axiomes de la théorie, formulés initialement par Zermelo lui-même, se bornent à affirmer soit l'existence inconditionnée de certains ensembles comme l'ensemble Z comparable à la suite des entiers, soit l'existence hypothétique d'autres ensembles. « Si l'ensemble M existe, l'ensemble N existe aussi » (par ex. l'*ensemble des parties* de M ou l'*ensemble réunion* des éléments de M). L'axiome de choix appartient à cette deuxième espèce. On peut s'étonner qu'il ait concentré sur lui tant d'attaques. Son énoncé présente évidemment le double défaut de lier deux existences sans donner

le moyen de passer de la première à la seconde (mais l'axiome de l'ensemble des parties et celui de l'ensemble réunion sont exactement dans le même cas) et surtout de ne pas caractériser de façon univoque l'existence subordonnée (il existe un ensemble de choix indéterminé tandis que l'ensemble des parties et l'ensemble réunion sont uniques). On s'est appliqué à séparer dans les travaux contemporains les démonstrations qui en faisaient usage et celles qui pouvaient s'en passer.

Mais le rejeter peut signifier deux choses : ou qu'il est une conséquence des autres axiomes, ou qu'il est en contradiction avec eux. La première possibilité semble devoir être éliminée après les travaux de Fraenkel. Il distingue divers cas suivant les puissances respectives du système S d'ensembles M_i sans élément commun : ou S est fini et les M_i quelconques, ou S est dénombrable et les M_i finis, ou S est dénombrable ou quelconque avec des restrictions pour les M_i, ou enfin S et les M_i sont quelconques. Dans le premier cas seulement il n'y a pas d'axiome : l'existence de l'ensemble de choix est démontrable (si S ne contient qu'un ensemble M de puissance d'ailleurs quelconque, c'est une conséquence de la logique classique : intuitivement cela revient à dire qu'on peut raisonner sur un exemple, sur un élément d'un ensemble non vide ; et on passe au cas où S contient n ensembles M_i par induction complète). Les autres cas exigent des axiomes de force croissante. Pour des modèles modifiés de la théorie des ensembles, Fraenkel a démontré l'indépendance, par rapport au groupe A, de l'axiome C_1 correspondant au cas 2, et celle par rapport à la réunion $A + C_1$ de l'axiome C_4 correspondant au cas général. Quoique la dernière démonstration ait exigé quelques changements dans A, il est infiniment peu probable qu'on puisse déduire C_4 de $A + C_1$. En revanche le problème de la non-contradiction reste entier.

De même pour l'hypothèse H du continu : doit-elle être posée comme un axiome, c'est-à-dire est-elle indépendante non seulement du groupe A (comme il est vraisemblable) mais même de l'axiome de choix ? C'est là une question jusqu'ici toujours ouverte. Sierpinski a consacré un livre à étudier les conséquences de H ou les propositions qui lui sont équivalentes, peut-être avec le secret espoir de trouver parmi elles une absurdité. C'eût été en tout cas le seul résultat qu'une telle recherche eût pu apporter sur la question (puisque d'un système d'axiomes contradictoires on peut tout déduire). En fait, ni pour son indépendance ni pour sa non-contradiction avec les groupes A et $A + C_4$ aucune indication n'avait été fournie avant Gödel. Lusin avait bien tiré des conséquences plausibles d'une autre hypothèse $2^{\aleph_0} = 2^{\aleph_1}$, mais sans, bien entendu, prouver sa non-contradiction avec A ou $A + C$ (ce qui

aurait pour conséquence de prouver l'indépendance de H par rapport à A ou A + C).

Quant à la non-contradiction interne du système A, elle n'est pas davantage établie. Ackermann a seulement démontré celle du système sans l'axiome d'infinité, ce qui est presque trivial. En fait une démonstration de non-contradiction n'a de sens que relativement aux moyens qu'elle utilise. C'est un rapport entre une zone absolue de sécurité et le nouveau domaine qu'on veut fonder sur elle. Mais y a-t-il une zone de sécurité absolue ? On sait que Hilbert avait cru la délimiter comme le royaume de la pensée concrète et finie, méta-mathématique qu'il opposait au développement entièrement formalisé de la mathématique. En fait des recherches ultérieures ont montré que la séparation ne pouvait se faire ni en une seule fois, ni de façon simple. Gödel donne un critère pour l'efficacité des démonstrations de non-contradiction : lorsqu'elles s'appliquent à un formalisme contenant l'arithmétique, elles ne doivent pas être formalisables en lui. Or une telle impossibilité n'existe sûrement pas pour les démarches situées par Hilbert et ses disciples dans la zone intuitive et finie.

D'une façon générale le mémoire de Gödel de 1931 montre que toutes les relations ou propriétés méta-mathématiques comme : « être la démonstration d'une proposition », « être une proposition démontrable » sont du même ordre que des relations récurrentes, peuvent être exprimées par elles, si on remplace les symboles et les groupes de symboles de formalisme par des chiffres ou numéros comme il est facile de le faire. Ainsi d'une part la pensée qui doit assurer les mathématiques est elle-même mathématique, d'autre part si elle domine des théories infinies, elle doit dépasser leur infinité. La démonstration de Gentzen pour la non-contradiction de l'arithmétique ordinaire emploie une induction transfinie. Paradoxe de la théorie des ensembles, qui, après avoir ébranlé l'édifice entier par l'exorbitant de ses développements, reparaît avec ce même infini comme instrument indispensable pour fonder les parties les moins contestées des mathématiques telles que l'arithmétique ordinaire. A la vérité il ne s'agit pas de toute la théorie des ensembles, mais justement le départ est instructif entre la portion abstraite — où se situent les paradoxes — et certains domaines des nombres transfinis dont les travaux récents de Gentzen, de Church et de son école montrent qu'ils possèdent et procurent une aussi grande sécurité que celle qu'exigeaient Hilbert et les plus résolus finitistes. Ici le régulier prolonge l'intuitif concret, le calculable le fini.

La séparation s'opère à l'intérieur des ordinaux de la classe II. La démonstration de Gentzen n'utilise que le segment dénoté par le

nombre ε_0, le premier ordinal x tel que $x^x = x$. Si on représente les ordinaux par les moyens que nous fournit l'arithmétique courante : somme, produits, élévation à la puissance, ε_0 est le premier nombre qui échappera à cette représentation. La formule

$$\alpha = k_0\omega^{x_1} + k_1\omega^{x_2} + \ldots\ldots + k_{n-1}\omega^{x_n} + k_n$$

(où $x_1 > x_2 > \ldots > x_n$ sont des ordinaux $< \alpha$ et où $k_0, k_1, \ldots k_n$ sont des entiers finis) ne peut servir pour lui. Il est justement le système ordonné de ces nombres, ou la limite de la suite fondamentale ω, ω^ω, ω^{ω^ω},.. Gentzen fait appel à une induction transfinie étendue à ce segment. Son principe — quel que soit le segment — est l'axiome du bon ordre. La proposition : « tout système d'ordinaux possédant une certaine propriété P a un premier élément » entraîne immédiatement : « si le fait que tous les ordinaux précédant un ordinal α possèdent une propriété a pour conséquence que α la possède, alors tous les ordinaux possèdent cette propriété » (en effet il y aurait sans cela un premier ordinal β parmi ceux qui ne la possèdent pas, et comme tous ceux qui le précèdent en sont pourvus, il devrait aussi la posséder). Mais Gentzen [1] et, après lui, Bernays se sont attachés à montrer que l'utilisation du principe en tant que limité au segment ε_0 revêt un caractère particulier et quasi fini, pourrait-on dire, qui tient à la simplicité de ces premiers ordinaux.. Bernays [2], grâce aux nombres premiers, a fourni une ingénieuse démonstration du principe d'induction complète pour le segment ε_0 par le moyen de l'induction ordinaire. On a déjà vu que ω_0^2 représentait un certain ordre des nombres rationnels finis. Ici il s'agit de superposer des ordres d'entiers finis : l'ordre 0 sera l'ordre naturel. L'ordre 1 ($\underset{1}{<}$) est établi d'après l'exposant du plus grand (dans l'ordre naturel) des facteurs premiers qui n'ont pas le même exposant dans deux nombres. Ainsi $8 \underset{1}{<} 3$ puisque 3 a l'exposant 0 dans le premier nombre. Avec cet ordre les nombres se trouvent rangés comme suit :

$$\begin{array}{l} 1,\ 2,\ 2^2,\ \ldots.\ 2^n,\ \ldots. \\ 3,\ 3.2,\ 3.2^2,\ \ldots\ldots \\ 3^2,\ 3^2.2,\ 3^2.2^2,\ \ldots. \\ \ldots\ldots\ldots\ldots\ldots \\ 5,\ 5.2,\ 5.2^2,\ \ldots. \\ \ldots\ldots\ldots\ldots\ldots \end{array}$$

1. GENTZEN (G.), *Die Widerspruchsfreiheit der reinen Zahlentheorie* (*Math. Ann.*, **112** (1936), pp. 493-565.

2. Le lecteur se reportera utilement à l'exposé donné par BERNAYS lui-même dans : HILBERT (D.), BERNAYS (P.), *Grundlagen der Mathematik*, Bd 2. Berlin, Springer, 1939, pp. 360 et suiv.

On voit qu'on a ainsi un bon ordre, où chaque suite introduite par un nouveau facteur premier doit se dilater en quelque sorte de toutes les suites semblables aux suites qui la précèdent. ω_0 sera coordonné à 3, $2\omega_0$ à 3^2, ... ω_0^2 à 5 et finalement $\omega_3^{\omega_0}$ à toute la suite.

L'ordre 2 rangera les nombres d'après la grandeur naturelle des exposants du nombre premier le plus grand dans l'ordre 1 qui n'a pas le même exposant dans deux nombres. Ainsi cette fois nous aurons :

$$11 \underset{2}{<} 7 \quad \text{ou} \quad 11.7.2 \underset{2}{<} 7^2$$

puisque 7 est le troisième nombre premier, 11 le 4^e et qu'on a dans l'ordre 1 : $4 \underset{1}{<} 3$.

L'ordinal $\omega_3^{\omega_0^{\omega_0}}$ dénote cet ordre, etc.... Bernays démontre que si le principe d'induction vaut pour l'ordre n il vaut pour l'ordre $n+1$. Comme il vaut pour l'ordre 0, qui est l'ordre naturel, cette induction complète ordinaire fonde bien l'induction transfinie pour le segment ε_0. Bernays et Gentzen se bornent là pour l'arithmétique ordinaire. Mais il faudrait aller plus loin si l'on voulait par exemple assurer, comme l'espère Gentzen, par le même procédé la non-contradiction de l'analyse. Jusqu'où peut s'étendre dans la classe II cette définition quasi finitiste des nombres ordinaux, tel est donc un problème intéressant à la fois la technique méta-mathématique et en elle-même la notion de transfini.

Déjà au début du siècle on avait essayé divers algorithmes pour assurer une notation uniforme au moins à un segment de ces nombres. Ainsi, en 1908, O. Veblen [1] par une généralisation des fonctions donnant les nombres ε_x : on voit que si on pose $y_1 = \omega^{x_1}$, $y_2 = \omega^{x_2}$,... la limite supérieure de $\omega^{y_n} = \lim. y_{n+1}$. Le $x^{\text{ième}}$ nombre obtenu par ce procédé est appelé ε_x. Mais les fonctions introduites par Veblen ne répondent pas exactement aux exigences d'un point de vue finitiste. On se souvient qu'en 1905 Borel et Lebesgue avaient discuté sur les notions de nommable et d'effectif. « Un objet, disait Lebesgue, est défini ou donné quand on a prononcé un nombre fini de mots s'appliquant à cet objet et à celui-là seulement [2]. » Le vague de cette exigence — qui devait limiter le domaine de la mathématique authentique — tient à l'indétermination des *mots*. Si parmi ceux-ci figurent « tous les objets qui... il y a un objet qui... », où est la finitude ? Le développement des formalismes, en particulier grâce aux méthodes de Gödel, permet d'éviter cet écueil. En obligeant un calcul à être effectif, Lebesgue se

1. VEBLEN (O.), *Continuous increasing functions of finite and transfinite ordinals.* (*Trans. Amer. Math. Soc.*, vol. **9** (1908), pp. 280-292).

2. LEBESGUE (H.), *Sur les fonctions représentables analytiquement* (*Journ. de Math. pures et appl.*, 1905, p. 205).

trouvait amené à le transformer en calcul effectué, s'il voulait chasser l'indéterminé de « un calcul ». Mais c'eût été rendre impossible tout progrès. Les notions corrélatives de calcul régulier et de fonctions quasi récurrentes introduites grâce aux travaux de Gödel échappent à ce dilemme.

Un calcul est régulier s'il s'accomplit dans un système formel où signes, groupes de signes, c'est-à-dire formules, et suites de formules sont rigoureusement numérotables de façon récurrente, où, bien plus, des fonctions ou des relations récurrentes permettent de donner, pour toute équation dont le deuxième membre est un nombre naturel, le numéro de cette équation dans le formalisme, pour toute fonction $f(x) = m$, le numéro de l'équation lorsque x prend la valeur particulière n (n étant un entier fini). Il y a par le moyen de ce numérotage et de ces relations récurrentes un contrôle permanent de l'effectivité du calcul : chacune de ces étapes tombe sous nos prises puisque nous sommes sûrs de l'atteindre au moyen d'une suite canonique de démarches. La détermination réside dans le processus de récurrence. A la vérité celui-ci n'est pas fixé une fois pour toutes, ce qui évite justement de borner les mathématiques aux mathématiques actuelles. Mais une fonction par exemple sera dite *régulièrement calculable*, si, pour elle, il est possible d'édifier un formalisme spécial qui réponde aux conditions posées.

Les fonctions régulièrement calculables coïncident avec les fonctions dites généralement récurrentes ou *quasi récurrentes* [1] dont Herbrand avait eu le premier l'idée. On connaît le schéma de la récurrence ordinaire pour les fonctions à une variable x : deux équations donnant l'une la valeur de la fonction pour $x = 0$, l'autre pour $x = n + 1$ (n étant un entier fini) au moyen d'expressions où n'entrent que des fonctions récurrentes déjà introduites et (dans la deuxième équation) la valeur de la fonction pour $x = n$. Il se produit ainsi une édification progressive de structures fonctionnelles à partir de l'unique fonction du suivant $sn = n + 1$. Par exemple les équations de définitions de l'addition et de la multiplication seront :

$$a + 0 = a \qquad a \times 0 = 0$$
$$a + sn = s(a+n) \qquad a \times sn = a \times n + a.$$

La généralisation, envisagée par Herbrand et précisée par Gödel et Kleene, consiste à poser un système d'équations simultanées où interviennent différents signes fonctionnels f de telle sorte que, par simples substitutions itérées, on puisse déterminer de façon univoque

1. Nous empruntons cette dénomination à Bernays qui donne un excellent exposé de la question dans un supplément au t. II de ses *Grundlagen d. Mathematik* (en collaboration avec Hilbert) Berlin, Springer, 1940 (pp. 392-421).

la valeur de f (en général à plusieurs arguments) pour toute valeur donnée du système de ses arguments. La longueur et la complication du calcul sont arbitraires : ils font intervenir les autres f et les valeurs de f pour d'autres systèmes de valeurs de ces arguments. L'important est que n'apparaissent comme opérations que les trois espèces de substitutions : substitution d'un système de valeurs numériques pour un système d'arguments dans une fonction, substitution dans une expression, à la fonction f, de sa valeur pour un système donné d'arguments, substitution dans une expression A, à l'expression partielle B, de l'expression C si l'équation B = C a pu être déduite. Ce sont exactement les mêmes opérations qui régissent les formalismes du calcul régulier : Church et Kleene ont montré [1] qu'elles peuvent toujours s'effectuer sous les conditions précédentes de numérotage et que réciproquement ces conditions, pour de tels formalismes, ne constituent pas autre chose que le système E des équations de définition d'une fonction quasi récurrente.

Il est naturel d'employer les nombres transfinis à ordonner les degrés de complication des fonctions quasi récurrentes. L'idée de Church et de Kleene [2] a été inversement d'utiliser celles-ci pour une définition des ordinaux *constructibles*. La notion de constructibilité, remarque Church [3], doit être dans une certaine mesure absolue, c'est-à-dire indépendante de la référence à un système formel donné si nous voulons garder son caractère intuitif à la représentation qu'elle précise. Puisqu'elle renvoie à un progrès à étapes jalonnables depuis les nombres finis, on peut considérer qu'elle pose les exigences suivantes : définir pour tout nombre α un système de notations tel que :

a) il donne une notation unique à tout ordinal β précédant α.

b) il procure 3 processus effectifs grâce auxquels :

1° la notation d'un ordinal α étant donnée (donc de tout ordinal $< \alpha$) on puisse déterminer s'il est de première ou de deuxième espèce.

2° pour un ordinal de première espèce on ait la notation de son pré-

1. Church, *An insolvable Problem of number theory* (*Amer. journal of math.* **58** (1936) pp. 345-363) ; — Kleene. *A theory of positive integers in formal logic* (*ibid.* **57** (1935), pp. 153-173 et 219-244).

2. Church (A.) et Kleene (S. C.), *Formal definitions in the theory of ordinal numbers* *Fundamenta mathematicae*, **28** (1937), pp. 11-21).

Kleene (S. C.), *On notation for ordinal numbers* (*The Journ. for Symb. Log.*, **3** (1938), pp. 150-155).

3. Church (A.), *The constructive second number class* (*Bull. of Amer. Math. Soc.*, vol. **44** (1938), pp. 224-232).

décesseur immédiat.

3° pour tout ordinal γ de deuxième espèce on ait une méthode, effective aussi, pour calculer la notation de chaque terme d'une suite fondamentale définissant γ.

Ces exigences ont également pour conséquence qu'étant donnés deux ordinaux $< \alpha$ on ait un processus effectif pour déterminer lequel est le plus grand et que, pour tout ordinal de deuxième espèce $< \alpha$, on ait un moyen de choisir une suite fondamentale parmi toutes celles qui le définissent. Il est bien entendu que le système de notations doit être fixé chaque fois pour le nombre particulier à construire.

Toute la définition repose comme on le voit sur la notion de processus effectif mais celle-ci est maintenant éclaircie grâce à la théorie des fonctions quasi récurrentes ou régulièrement calculables. Ou bien, en effet, le processus porte sur des valeurs numériques : il devra alors utiliser une suite de fonctions quasi récurrentes. Ou bien il consiste dans un enchaînement de formules et, grâce au procédé Gödel de numérotage récurrent, nous sommes renvoyés au cas précédent. Ainsi les trois opérations de substitution peuvent-elles également donner une forme canonique à toute définition générale des nombres ordinaux constructibles. Aussi bien ceux-ci coïncident-ils avec les ordinaux définissables dans le formalisme de la *conversion* (c'est-à-dire de ces trois opérations) établi par Church et Kleene. On voit en effet que ce formalisme permet de représenter le nombre 0, étant donné un nombre, celui qui le suit, enfin étant donnée une suite d'ordinaux représentée par une fonction f appartenant au formalisme et coordonnée à la suite des entiers ($f(0)$, $f(1)$ ... $f(n)$ ...), la limite supérieure de cette suite. On montre ensuite que les exigences de la constructibilité sont bien satisfaites : Church et Kleene construisent des fonctions telles que $K(n)$, égale à 1 si n est de première espèce, égale à 2 si n est de deuxième espèce.

Toutes les fonctions ordinaires, addition, multiplication, élévation à la puissance sont représentables dans le formalisme. Tous les nombres jusqu'à ε_0 sont donc constructibles, bien plus tout nombre ε_α, si α est lui-même constructible : donc tous les nombres du segment déterminé par le premier nombre x tel que $\varepsilon_x = x$. Cependant les nombres constructibles ne forment qu'une infime minorité dans la classe II : en effet leur ensemble est non seulement dénombrable, mais énumérable puisqu'on peut numéroter leur définition dans le formalisme Church-Kleene. Or la classe II a une puissance supérieure au dénombrable : les nombres de la classe II non constructibles forment un ensemble ayant lui-même la puissance de la classe II. Il est d'ailleurs

impossible par essence de préciser où se fait la séparation : le plus petit nombre non constructible μ est évidemment de deuxième espèce, mais si l'on pouvait lui assigner une suite fondamentale, comme elle est par hypothèse composée de nombres constructibles, μ le serait également.

Tel est le point extrême jusqu'où s'avance pour le moment une exploration finitiste de la classe II. Aussi bien, quelles que soient les modifications ultérieures, les conditions de constructibilité ne peuvent permettre d'aller sensiblement plus loin, puisque liées au dénombrable. Mais la partie non dénombrable de la classe reste-t-elle comme une virtualité, extérieurement et antérieurement à laquelle se déterminent les objets mathématiques véritables ? La définition précise du calcul effectif d'une fonction, les développements formels qui y sont rattachés présentent intrinsèquement un intérêt mathématique incontestable. Il est en outre possible que le sens même de certains problèmes exige que les êtres qui y interviennent soient soumis à des conditions de constructibilité. Philosophiquement on peut se demander pourquoi l'extension d'évidence obtenue au-delà du fini s'arrêterait à cette étape. En particulier pour la théorie générale des ensembles, il ne saurait être question de lier la démonstration de sa non-contradiction aux conditions précédentes. Cependant le théorème publié récemment [1] à ce sujet par Gödel s'appuie sur des raisonnements dont la parenté avec les développements effectifs au sens Church se traduit par leur insertion possible dans un système intuitionniste. En effet ils n'utilisent pas de définition non-prédicative. Or comme le montrait Gödel [2] à propos du formalisme intuitionniste codifié par Heyting et d'une arithmétique intuitionniste qu'on y peut joindre (définie par des axiomes que Herbrand a formulés) le rejet du tiers exclu ne change pas le volume des propositions démontrables : toute proposition démontrable avec le tiers exclu étant traduisible en une proposition démontrable dans le formalisme intuitionniste. C'est seulement avec les définitions prédicatives qu'apparaissent les différences essentielles. Ainsi l'actuelle démonstration peut-elle être considérée comme constructive « au sens que si une contradiction se manifestait dans le système A + B, elle peut être actuellement manifestée dans le système A ».

Le raisonnement consiste à fabriquer, en n'utilisant que les axiomes

1. Consistency proof for the generalized continuum hypothesis (*Proc. of the National Academy of Sc. of the U. S. A.*, an. 1939, pp. 220-224). Le lecteur se reportera utilement au mémoire de GÖDEL, *The consistency of the axiom of choice and of the generalized continuum hypothesis with the axioms of set theory* (Princeton, 1940), et à l'article de GÖDEL, *What is Cantor's continuum problem ?* (*The Amer. Math. Monthly*, **54** (1947), pp. 515-525).

2. *Ergebnisse eines math. Kolloquiums* (MENGER), 1933, heft 4, p. 35.

du groupe A, un modèle de théorie des ensembles dans lequel l'axiome de choix et l'hypothèse du continu se trouvent vérifiés. Le principe est de substituer à l'indétermination radicale de l'*ensemble des parties* d'un ensemble, une édification progressive de cet ensemble en supposant formalisées toutes les définitions possibles de parties. Si l'on considère un système formalisable ou formalisé de la théorie (par adjonction du formalisme logique aux axiomes), une partie d'un ensemble M ne sera déterminée effectivement que par une fonction propositionnelle étendue à M, c'est-à-dire par une proposition formelle portant sur un individu variable parcourant tout l'ensemble : la partie sera constituée par les éléments de M pour lesquels la proposition est vraie. On pourra donc définir autant de parties de M qu'il est possible d'établir de fonctions propositionnelles portant sur M. Or une fonction propositionnelle ne consiste que dans un certain assemblage de signes en nombre restreint, que l'on a déterminés une fois pour toutes : les complications des architectures que l'on peut fabriquer avec eux ne dépassent pas une infinité dénombrable.

Si on part, comme ensemble initial par exemple, de la suite des entiers, on pose ainsi l'ensemble M_1 des parties définissables par toutes les fonctions propositionnelles possibles portant sur M. Puis à nouveau l'ensemble M_2 des parties définissables sur M_1 de la même façon. Le procédé est considéré comme se poursuivant ainsi transfiniment : l'ensemble M_{ω_0} et, de façon générale, tout ensemble dénoté par un ordinal de deuxième espèce sera la réunion de tous les ensembles qui le précèdent. Qu'il soit d'ailleurs dénoté par un ordinal de première ou de deuxième espèce, tout ensemble contient comme éléments tous les ensembles qui le précèdent et il est lui-même à la fois élément et partie d'un ensemble qui le suit. On dit qu'il est d'ordre α s'il est élément de $M_{\alpha+1} - M_\alpha$.

Le modèle envisagé comprend tous les ensembles dits *constructibles* figurant ainsi dans l'ensemble M_{ω_ω}[1]. On y peut démontrer en effet une sorte d'axiome de réductibilité qui empêche les axiomes de Zermelo-Fraenkel ou de Von Neumann d'exiger l'existence d'ensembles situés hors de M_{ω_ω} Ces axiomes (ceux de l'ensemble des parties ou de la réunion) sont de la forme : « il existe un ensemble a tel que tous ses éléments x soient définis par une certaine fonction propositionnelle $\varphi(x)$ » ; donc si les x appartiennent à l'ensemble M_α, a appartiendra à $M_{\alpha+1}$. Mais Gödel établit le théorème fondamental T : si un ensemble

1. Nous simplifions : pour atteindre des ordinaux assez grands il faut poser un axiome dit de *remplacement*. Le modèle pour la théorie ainsi complétée doit s'étendre jusqu'à M_Ω, Ω étant le premier ordinal *inaccessible*.

constructible est partie d'un ensemble M_{ω_α} son ordre est inférieur à $\omega_{\alpha+1}$. Comme M_{ω_ω} est la réunion de tous les M_{ω_α} pour $\alpha = 1, 2, .. n...$, quel que soit le M_{ω_α} auquel appartiennent les x, a appartiendra toujours à M_{ω_ω}. Le système des éléments de M_{ω_ω} vérifie donc les axiomes du groupe A.

D'autre part leur définition et les démonstrations présentées, en particulier celle du théorème T, sont effectuées uniquement avec le groupe A : elles peuvent se refaire dans une théorie des ensembles ne considérant comme objets que les éléments de M_{ω_ω}. Bien plus si on prend ceux-ci comme point de départ (et non plus comme tout à l'heure la suite des entiers), on retrouve exactement les mêmes ensembles ; autrement dit l'opération de construction, telle qu'elle a été définie, est absolue par rapport à la restriction au modèle M_{ω_ω} Mais, dans ce modèle, il n'y a pas d'autres ensembles que les ensembles constructibles. Or pour eux l'axiome de choix est déduisible des autres axiomes : il est toujours possible de numéroter au moyen des ordinaux transfinis tous les éléments d'un ensemble (en procédant de proche en proche en suivant à partir d'un ensemble base aux éléments numérotés comme la suite des entiers et en numérotant les fonctions propositionnelles à chaque étape d'après le nombre et l'ordre des signes qui interviennent). Il suffira alors de prendre comme ensemble de choix, pour un système d'ensembles, l'ensemble des éléments ayant le numéro 1 dans chaque ensemble du système.

De même l'hypothèse du continu est vérifiée : la puissance d'un ensemble M_{ω_α} est par hypothèse $\aleph_\alpha$, d'autre part, puisqu'il n'y a pas d'autres ensembles que les ensembles de M_{ω_ω} ceux d'entre eux que notre procédé construit sur M_{ω_α} formant l'ensemble des parties de M_{ω_α} ont pour puissance $2^{\aleph_\alpha}$. Or chacune de ces parties est, en vertu du théorème T, d'ordre plus petit que $\omega_{\alpha+1}$, leur ensemble étant en partie de $M_{\omega_{\alpha+1}}$ a donc une puissance $\leqslant \aleph_{\alpha+1}$, donc $2^{\aleph_\alpha} \leqslant \aleph_{\alpha+1}$. Comme $2^{\aleph_\alpha} > \aleph_\alpha$, on a bien l'égalité cherchée : $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

Ces raisonnements n'utilisent encore que le groupe A d'axiomes : B est bien pour ce modèle conséquence du groupe A. S'il existe une contradiction dans la théorie des ensembles, elle doit être localisée dans le groupe A.

On ne saurait exagérer l'importance de ce résultat. Non seulement il clôture définitivement les controverses auxquelles l'énoncé de Zermelo avait donné occasion, mais il achève l'unification vainement

poursuivie par Cantor entre la théorie cardinale et la théorie ordinale des ensembles. Tous les théorèmes ingénieusement rassemblés par Sierpinski et Lusin par exemple dans la théorie des ensembles analytiques, comme conséquences possibles de l'hypothèse du continu, en reçoivent confirmation : leur validité peut dépendre d'un nouvel axiome, elle n'est plus exposée à être rejetée par l'apparition d'une contradiction inaperçue. Comme le remarque Gödel, il suffit d'ailleurs d'ajouter (pour obtenir à la fois l'axiome de choix et l'hypothèse du continu) le seul axiome : « il n'y a pas d'autres objets-ensembles que les ensembles constructibles au sens défini plus haut ». Comme élément initial, on peut se borner à l'ensemble vide, la suite des nombres est obtenue immédiatement grâce à l'axiome de l'infini. La construction du modèle M_{ω_ω} prouve que le nouvel axiome ne présente pas de contradictions avec le groupe A.

D'autre part une telle adjonction se situe dans la ligne d'évolution normale de la théorie : il est naturel que pour la soumettre aux prises de la critique on ait cherché à uniformiser la manipulation concrète de ces objets, c'est-à-dire à l'incorporer à un formalisme logique. Ainsi se présentent, quoique laissant un champ plus large aux objets-ensembles, les modifications apportées par Von Neumann et Fraenkel au système primitif de Zermelo. En particulier l'effort de Fraenkel consistait à préciser l'axiome trop vague posant l'existence de certaines parties d'un ensemble. « Si une proposition $\mathcal{E}(x)$ est définie pour tous les éléments d'un ensemble M, il existe une partie $M_{\mathcal{E}}$ de M constituée par tous les éléments pour lesquels $\mathcal{E}(x)$ est vraie. » A quoi Zermelo ajoutait l'unique éclaircissement : « Une proposition est définie si sa validité peut être décidée sans arbitraire au moyen des axiomes et des lois logiques universelles. » Les lois logiques universelles n'empêchaient pas les antinomies de se manifester.

Fraenkel fait intervenir la notion de fonction d'un ensemble : « S'il existe un ensemble M et deux fonctions données $\varphi(x)$ et $\psi(x)$, il existe une partie de M comprenant tous les éléments x de M pour lesquels $\varphi(x)$ est élément de $\psi(x)$ ou inversement. » La seule différence avec l'axiome suggéré par Gödel réside dans la définition de $\varphi(x)$ et $\psi(x)$. Tandis que pour Gödel, les fonctions propositionnelles (qui remplacent l'unique proposition $\varphi(x) \in \psi(x)$) [1] sont uniquement constituées avec les symboles logiques, les constantes représentant les ensembles déjà construits et le symbole $\in$, il intervient chez Fraenkel par exemple comme fonction d'ensemble Px, l'ensemble des parties de x dont l'existence se trouve posée par un autre axiome. Autrement dit l'ensemble des parties joue lui-même un rôle dans la définition

d'une partie d'ensemble, « exemple classique de moment non prédicatif » [2] mais nécessaire pour conserver toute son ampleur à la théorie. Gödel cependant l'évite, sans restreindre semble-t-il sensiblement, en poursuivant son processus de construction assez loin dans le transfini. Du théorème T découle l'énoncé de réductibilité, qui assure comme on l'a vu la validité aussi bien de l'axiome de séparation que de celui de l'ensemble des parties.

Ainsi se manifeste une sorte de réciprocité entre superposition transfinie et moment non prédicatif : de même qu'en géométrie projective, l'intervention d'une nouvelle dimension de l'espace rend inutiles les axiomes de congruence (en passant de la droite au plan, du plan à trois dimensions), de même l'élévation jusqu'à un ordre transfini suffisamment élevé remplace les définitions non prédicatives et les justifie par un processus prédicatif. C'est l'idée même qui guide Gentzen dans son effort pour démontrer la non-contradiction de l'analyse classique. L'originalité de Gödel est d'avoir poussé jusqu'à la classe ω_ω et du même coup affermi ce que l'on croyait depuis l'échec de la méta-mathématique hors des possibilités d'un affermissement quelconque.

Est-il raisonnable d'exiger ou seulement d'espérer davantage ? Les conditions posées par les finitistes de stricte observance à l'existence mathématique manifestent une méconnaissance de ce qui fait la réalité même de l'objet. Il n'y a pas d'absolu au départ : que dire du point, du continu, du nombre entier, sinon qu'ils sont éléments et origines d'enchaînements où ils prennent leur sens et qui les dépassent ? Une intuition irréductible n'est qu'un arrêt sans pensée. La théorie kantienne des mathématiques obsède encore logiciens et mathématiciens, mais par ce qu'elle présente de superficiel et d'illégitimement interprétable en termes psychologiques. Prise d'ailleurs dans son authenticité elle reste inadéquate en tant qu'elle confond le moment dialectique de la position du concept et le moment transcendantal de sa schématisation. La notion de construction dans l'intuition doit concilier l'exigence de domination par le concept avec l'insertion de son objet dans une multiplicité qui lui procure une sorte d'indépendance par rapport à tout contenu actuel de pensée. Il y a dans l'objet mathématique singulier comme une affinité avec ce qui n'est pas lui, par quoi il échappe à sa définition et justifie le progrès hors de lui. Mais comment, dans la théorie kantienne du concept, poser en termes intellectuels une liaison qui semble comporter un élément irréductible

1. a ϵ b signifie a est élément de b.

2. FRAENKEL, *Untersuchungen über die Grundlagen der Mengenlehre* (*Math. Zeitschrift*, 22 (1925), p. 256).

d'extériorité ? D'où l'idée d'une connaissance intuitive où multiplicité et extériorité subsistent par essence, et d'une construction intuitive, qui soumet cette extériorité à l'unité sans l'abolir comme telle, parce qu'elle reste à la fois à l'intérieur et à l'extérieur du concept. Sa rationalité virtuelle, les liens de celle-ci avec les concepts d'un intellect fini sont autant de problèmes : on peut se demander ce que comporte le concept de *quantum* comme contenu intelligible indépendamment de son schème numérique. Par ailleurs apparaît le malentendu qui fait donner la construction elle-même comme représentable ; c'est la confusion déjà signalée entre procès effectif et procès effectué. Puisque le temps de la construction est le temps transcendantal, il devrait échapper à toute représentation et par suite à toute condition dans l'accomplissement de sa synthèse. Mais serait-il encore un temps ? et pourrait-on encore parler d'accomplissement ? C'est l'ambiguïté fondamentale de la philosophie transcendantale de poser les conditions de la connaissance relativement à une activité synthétique qui serait définissable et appréhendable directement dans une connaissance échappant elle-même à la critique : la condition de la représentation se confondant avec les représentations de condition. D'où l'exclusion du nombre infini : comment construire dans le temps ce qui suppose une infinité d'étapes ? mais c'est justement dissocier ces étapes de leur unique fonction d'*éléments* pour en faire des *événements*, les objets exigeant une représentation singulière et n'existant que par elle. Alors la mathématique est impossible et l'éléatisme a raison.

Si l'on veut déjà donner un rôle au *quelconque* dans la définition d'un procès synthétique, aussi bien au *quelconque* apparaissant dans le discernement des étapes, que s'opposant à la singularité de celles-ci et du résultat final, il faut que le procès soit lui-même défini comme règle, unité rationnelle d'une certaine multiplicité, que la référence à l'intuition suppose une relativité de cette intuition à d'autres schèmes, d'autres enchaînements réglés. A chaque étape du progrès mathématique de nouvelles notions se voient refuser leur droit de cité, parce qu'elles ne peuvent procurer de représentation dans le système intuitif préexistant. Ainsi pour le nombre négatif, pour les imaginaires, pour l'infiniment petit et l'infiniment grand à l'époque de Kant. La solution chaque fois n'est pas, comme on le croit superficiellement, la découverte d'une traduction dans une intuition spatiale immuable (représentation des nombres négatifs sur une droite, des imaginaires dans le plan) : aussi bien faudrait-il donner un sens à la notion confuse de traduction. La solution consiste dans une transformation de la zone intuitive c'est-à-dire des règles qui en posent l'emploi de pensée, ou le

système schématique. L'exemple le plus frappant est le continu numérique avec la définition donnée par Dedekind. Autrement dit l'intuition dans sa quiddité progresse parallèlement à l'enchaînement dialectique des concepts. Elle n'est que la manifestation pour la conscience empirique, d'une indépendance relative des méthodes et des théories, qui permet des élaborations autonomes provoquant par leurs résultats rencontres et renversements. Le schème transcendantal représente le pouvoir moteur d'une synthèse nouvelle — postérieure à de tels renversements — la force interne d'une méthode, l'unité sous-jacente d'une théorie. Dans la mesure où il schématise une intuition, il fixe la théorie dans un aspect rigide, lui procure une sorte d'uniforme dont l'intelligibilité même le fait croire définitif. Ainsi l'arithmétisation de l'analyse, et cette intuition du nombre pur dont parle Poincaré, schéma en réalité du raisonnement d'induction complète.

Mais les exigences d'actualisation sont chaque fois relatives à l'étage intuitif, c'est-à-dire à la multiplicité inhérente au système conceptuel mathématique : si la suite des nombres est une intuition, il est aussi impossible de s'y représenter tel nombre particulier (comme d'imaginer un chiliogone) que d'y décider leurs propriétés (problèmes non résolus de l'arithmétique). Les travaux de Church, Kleene sur la constructibilité des fonctions et des nombres transfinis, représentent une étape postérieure : en utilisant les fonctions quasi récurrentes, en parcourant un segment de la classe II, ils exigent une intuition plus compliquée à quoi s'applique le schématisme des fonctions quasi-récurrentes. De même l'induction transfinie non restreinte du théorème de Gödel : le schème est aussi précis puisqu'il permet de démontrer, la zone intuitive est constituée par le système des schèmes (à un étage inférieur) du formalisme logico-mathématique de la théorie des ensembles. Cantor disait « Je me représente un ensemble comme un abîme. » Le progrès par superposition intuitive a précisé cette intuition dans la multiplicité des modes d'emploi du signe-ensemble au sein des systèmes Zermelo-Fraenkel ou Von Neumann : elle est bien ici encore mobilité réglée, relation à l'indétermination des problèmes non résolus.

Le lien entre cette superposition intuitive et la dialectique du concept reste le problème fondamental de la philosophie mathématique. La description est bien connue de ces renversements historiques où le résultat fait éclater la méthode et le système tout entier dont il est issu : les procédés exigés par la solution d'un problème provoquent, dans l'actualisation même qui leur donne un sens, un tel changement d'éclairage qu'il faut déjà abandonner les notions qui forment leur

structure. Mais les liaisons intellectuelles dépassent l'histoire empirique : c'est leur développement dialectique qui assure à la fois le mouvement de celles-ci et par elles-mêmes la permanence de leur validité. Ce qui marque l'histoire est la soumission du transcendantal à ses étapes : l'obligation du passage est reconnue dans un échec, la nécessité du progrès dans l'indétermination d'une découverte. La nécessité apparaît après coup. Aucune analyse de la conscience des actes ne permet de prévoir, pas plus qu'elle ne procure une permanence quelconque. Intuitif est synonyme de conscience effective (ou effectuante), transcendantal de constitutif (ou constituant), relativement à un système conceptuel donné. Pas plus qu'il n'y a de système clos, il n'y a de transcendantal absolu : l'unité formelle est la continuité d'un enchaînement d'actes, le concret est la nouveauté d'unification d'un multiple reconnu encore conceptuellement comme tel. Mais s'il n'y a de multiplicité qu'au départ, toute tentative de régularisation inconditionnée est frappée de caducité : ainsi pour l'arithmétisme fin de Kronecker, pour la théorie weierstrassienne du calcul infinitésimal, pour le finitisme hilbertien.

" Ἀεὶ ὁ ἄνθρωπος ἀριθμητίζει "

Si l'on veut garder la devise de Dedekind, il faut dilater la notion de numération plus loin que l'*application*, plus loin qu'un transfini déterminé, plus loin même qu'un nombrement quelconque, pour qu'elle marque la solidarité effective d'un développement réglé dont la justification intelligible est à la fois hors de la condition humaine — pour être intégrale — et pourtant — au moins en idée — peut-être une de ses caractéristiques essentielles.