UNIVERSITÉ LAVAL

L'hypothèse du continu
Contexte et conséquences

Mémoire

Frédéric Morneau-Guérin

Maîtrise en mathématiques
Maître ès sciences (M.Sc.)

Québec, Canada

Résumé

En mettant à profit le lien étroit entre la théorie des nombres ordinaux et la cardinalité des ensembles bien ordonnables, on présente l'arithmétique des $\aleph$ puis, en exploitant l'astuce de Scott, on expose la théorie des nombres cardinaux en toute généralité. Enfin, on énonce l'hypothèse du continu.

Une fois cette mise en contexte effectuée, on se tourne vers les résultats exclusifs aux théories des ensembles ZF et ZFC, toutes deux enrichies de l'hypothèse du continu. On démontre d'abord que l'hypothèse du continu permet d'établir un principe de dualité entres les notions de mesure et de catégorie au sens de Baire. Puis, on observe comment des ensembles aux propriétés topologiques particulières - obtenus en supposant l'hypothèse du continu - apportent un éclairage différent sur des thèmes à saveur analytique comme le théorème d'Egoroff, le théorème de Fréchet de suite double et le problème de la mesure généralisée.

Table des matières

Résumé iii

Table des matières v

Liste des tableaux vii

Liste des figures ix

Remerciements xv

Introduction 1

1 Théorie axiomatique des ensembles 3
- 1.1 Système axiomatique de Zermelo–Fraenkel . . . 4
- 1.2 Relation d'ordre . . . 13
- 1.3 Axiome du choix et principe du bon ordre . . . 15

2 Théorie des nombres ordinaux 19
- 2.1 Nombres ordinaux . . . 20
- 2.2 Propriétés des nombres ordinaux . . . 25
- 2.3 Arithmétique ordinale . . . 30

3 Théorie des nombres cardinaux 35
- 3.1 Théorème de Cantor–Bernstein . . . 36
- 3.2 Nombres de Hartogs et alephs . . . 38
- 3.3 Arithmétique cardinale dans ZFC . . . 45
- 3.4 Hiérarchie cumulative et astuce de Scott . . . 47
- 3.5 Catastrophe en arithmétique cardinale dans ZF . . . 51
- 3.6 Hypothèse du continu . . . 53

4 Principe de dualité résultant de l'hypothèse du continu 57
- 4.1 Catégories de Baire . . . 58
- 4.2 Mesure . . . 62
- 4.3 Similarités et distinctions entres les classes $\mathcal{M}$ et $\mathcal{N}$ . . . 64
- 4.4 Principe de dualité de Sierpiński–Erdős . . . 66

5 Ensemble de Lusin 75
- 5.1 Existence d'un ensemble de Lusin . . . 75
- 5.2 Résultats d'équivalence liant les ensembles de Lusin à l'hypothèse du continu 77

5.3 Théorème d'Egoroff 78
5.4 Théorème de Fréchet de suite double 80
5.5 Problème généralisé de la mesure 82
5.6 Application du principe de dualité de Sierpiński–Erdős aux ensembles de Lusin 86

Conclusion **89**

Bibliographie **91**

Liste des tableaux

0.1 Notations 2

Liste des figures

1.1 Genèse ensembliste *ex nihilo*, conception itérative du domaine du discours de la théorie de Zermelo–Fraenkel 9

2.1 Représentation visuelle du premier ordinal transfini ω, l'ensemble des ordinaux finis. Source : http ://naturelovesmath.blogspot.ca/2012/05/infini-et-nombres-transfinis.html 24

2.2 Représentation des nombres ordinaux jusqu'à ω^ω. Chaque tour de spirale représente une puissance de ω. Source : http ://naturelovesmath.blogspot.ca/2012/05/infini-et-nombres-transfinis.html 25

3.1 Représentation des correspondances biunivoques induites par les deux injections. Source : adaptée de http ://commons.wikimedia.org/wiki/File :Cantor2.png 37

À Kiana
We are in this together.

Ce n'est point dans l'objet que réside le sens des choses, mais dans la démarche.

Antoine de Saint-Exupéry
(1900-1944)

Remerciements

Mes premiers remerciements vont à mon directeur de maîtrise, le professeur Javad Mashreghi, qui a embarqué avec moi dans cette audacieuse aventure hors de son domaine de prédilection et qui m'a guidé patiemment tout en me laissant une grande liberté d'action. Ses judicieux conseils, son écoute attentive et bienveillante ainsi que sa disponibilité ont été d'une valeur inestimable.

Je tiens à exprimer ma reconnaissance à Jérémie Rostand pour son efficacité, sa rigueur, son professionnalisme et son dévouement pédagogique admirable. Toutes ces qualités ont sans doute fait de lui un important catalyseur de mon coup de foudre pour la topologie générale. J'aimerais aussi exprimer toute ma gratitude à Thomas Ransford pour son style lumineux et particulièrement fluide. Je me réjouis d'avoir eu la chance et le privilège de côtoyer un tel homme d'exception.

Je me dois de souligner l'apport pédagogique de Line Baribeau, Claude Bélisle, Frédéric Gourdeau et Bernard Hodgson qui m'ont fait voir qu'au-delà des mathématiques en dimensions finies ou infinies, il y a aussi les mathématiques à dimension humaine. Je remercie spécialement Bernard Hodgson de m'avoir fait découvrir la poésie mathématique.

Merci à mes amis Thierry Fournier, Benoît Nolet et Pierre Truchon ainsi que mon professeur et ami Nicolas Doyon qui ont tous joué à leurs façons un rôle marquant dans ma vie personnelle et dans mon parcours académique. Je remercie profondément Martin Dufresne sans qui je n'aurais sans doute pas eu la confiance nécessaire en mes aptitudes mathématiques pour entreprendre des études dans ce domaine.

Sur une note plus personnelle, ma plus sincère reconnaissance va à mes parents Jacinthe et Mario ainsi qu'à Émilie et Francis pour la fierté et l'appui inconditionnel qu'ils m'ont toujours témoigné. Enfin, j'aimerais remercier de tout mon coeur ma chère Kiana dont l'amour et le soutien de tous les instants me comblent au-delà de toute espérance.

Introduction

L'hypothèse du continu est l'un des plus célèbres problèmes de la théorie des ensembles qui demeure en suspens. Étant désormais admis que la théorie axiomatique des ensembles constitue *un cadre adéquat pour fonder les mathématiques*[1], l'hypothèse du continu se révèle une question d'importance capitale tant pour des raisons mathématiques que philosophiques.

Comme le titre l'indique, ce mémoire se divise en deux parties. Dans un premier temps, nous proposons de situer l'*hypothèse du continu* dans le contexte qui en révèle le sens. D'abord, au chapitre 1, nous présentons les axiomes de la théorie ZFC à partir desquels tous les résultats subséquents sont obtenus. De cette façon, la théorie des ensembles est développée *ab initio* et, à ce titre, nous référons à la littérature avec parcimonie. Au chapitre 2, nous nous autorisons une digression importante nous permettant de jeter les bases de la théorie des nombres ordinaux. C'est que la théorie des nombres cardinaux, dont est issue l'hypothèse du continu, peut être construite avec ou sans l'aide des nombres ordinaux. Les deux types de constructions offrant leurs avantages et occasionnant des modifications subtiles, mais non négligeables dans la formulation de l'hypothèse du continu, nous avons choisi - pour maintenir autant que possible la concorde - de les présenter toutes deux en détail au chapitre 3.

Dans un deuxième temps, nous étudions quelques conséquences du l'hypothèse du continu. D'abord, il est question au chapitre 4 du principe de dualité de Sierpiński–Erdős, un résultat nécessitant l'hypothèse du continu qui établit une relation de dualité entre deux notions de petitesse en apparence indépendantes l'une de l'autre : les ensembles de premières catégories de Baire et les ensembles de mesure nulle au sens de Lebesgue. Enfin, le chapitre 5 traite des conséquences topologiques et analytiques de deux classes d'ensembles de nombres réels aux propriétés particulières - les ensembles de Lusin et les ensembles de Sierpiński - dont l'hypothèse du continu suffit à assurer l'existence.

Certes, les concepts présentés dans la première partie de ce mémoire sont de moins grande envergure. Quoi qu'il en soit, nous estimons qu'il est crucial d'y fixer les idées vouées à jouer un rôle important dans la seconde partie. Ce faisant, nous nous distançons de l'usage le plus généralement admis qui consiste à augmenter le nombre de pages au fur et à mesure qu'on se

1. Cela signifie que l'intégralité de la mécanique interne des mathématiques peut se définir en termes d'ensembles.

rapproche des résultats les plus récents et à construire en quelque sorte une pyramide posée sur la pointe.

Notations

Dans ce mémoire, nous utilisons - sauf mention contraire - les lettres capitales $A, B, C, ...$ pour représenter les ensembles et les lettres minuscules $a, b, c, ...$ pour désigner les éléments appartenant à ces ensembles. L'alphabet grec $\alpha, \beta, \gamma, ...$ est affecté à la notation des nombres ordinaux. Quant à l'écriture *fraktur* $\mathfrak{m}, \mathfrak{n}, \mathfrak{p}, ...$, elle est réservée pour les nombres cardinaux.

Expression	**Signification**
$\neg$	Symbole logique de négation
$\mathbb{N}$	Ensemble des nombres naturels non nuls 1, 2, 3, ...
$\mathbb{R}$	Ensemble des nombres réels.
$A \backslash B$	Différence d'ensembles ; ensemble des éléments appartenant à A mais pas à B.
$A \triangle B$	Différence symétrique d'ensembles ; ensemble des éléments appartenant à A ou à B mais non au deux.
$\mathscr{P}(A)$	Ensemble de tous les sous-ensembles de A.
$\overline{A}$	Fermeture ou adhérence de l'ensemble A.
$\text{int}(A)$	Intérieur de l'ensemble A.
$m(A)$	Mesure de Lebesgue de l'ensemble A.
∂E	Frontière ou bord de $tikzE$.
$\begin{array}{cccc} f: & X & \rightarrow & Y \\ & x & \mapsto & y \end{array}$	Fonction f de X à valeur dans Y définie par $f(x) = y$.
$f\big\|_A$	Restriction de la fonction f au domaine A.

TABLE 0.1 – Notations

Chapitre 1

Théorie axiomatique des ensembles

L'hypothèse du continu, tel qu'énoncée par Cantor en 1878, affirme qu'il n'existe aucun ensemble dont le cardinal est strictement compris entre le cardinal de l'ensemble des entiers naturels et celui des nombres réels. Le pendant rigoureux de cet énoncé sera formulé ultérieurement à la lumière de la théorie des nombres cardinaux.

L'attention soutenue que reçoit l'hypothèse du continu depuis 135 ans de la part de la communauté mathématique a tout à voir avec le fait qu'il s'agisse d'un énoncé indécidable dans le cadre du *système axiomatique de Zermelo–Fraenkel enrichi de l'axiome du choix* (sigle : ZFC), généralement considéré comme une formalisation adéquate de la théorie des ensembles de Cantor. Cela signifie que l'élévation de l'hypothèse du continu ou de sa négation au rang d'axiome - en addition à ceux de la théorie ZFC - n'engendre une théorie contradictoire que si la théorie ZFC l'est. La démonstration par le logicien autrichien Kurt Gödel de l'irréfutabilité de l'hypothèse du continu dans ZFC en 1938, puis celle de sa indémontrabilité par l'américain Paul Cohen en 1963, ne constituent pas le sujet de ce mémoire ; nous nous contenterons d'admettre ces résultats. Néanmoins, ce statut particulier nous obligera d'emblée à quelques précautions. Au cours de ce chapitre, nous tâcherons d'établir que l'hypothèse du continu est un candidat légitime au titre d'axiome de la théorie des ensembles en contrant l'impression largement répandue voulant que la ligne de démarcation entre justifications intrinsèques et extrinsèques - aussi imprécise et vague puisse-t'elle être - passe entre les axiomes de la théorie de Zermelo–Fraenkel (sigle : ZF) et le reste.

Étant globalement accessibles aux non-initiés, les axiomes de ZF sont abondement présentés dans les livres de référence. Cela peut porter à croire qu'ils disposent d'un statut épistémologique particulier - par rapport à ceux suggérés postérieurement comme l'axiome du choix, les axiomes des grands cardinaux ou l'hypothèse du continu - alors que tant s'en faut ! Les axiomes de ZF n'ont rien de sacré. C'est le propre d'un axiome que de découler d'un acte de foi. Un axiome est un principe en lequel il faut *croire* et non une vérité à *connaître*[1].

1. La nuance est subtile, mais capitale. La pensée axiomatique se caractérise par le rejet de l'idée que des

Afin d'éviter de se retrouver emprisonné dans un carcan de dogmatisme, nous allons présenter les axiomes de ZF en justifiant autant que faire se peut les actes de foi qui y sont inhérents. Cette stratégie vise à abolir les castes : la méthodologie menant à l'introduction de tous les axiomes que nous entendons utiliser sera uniformisée. Cette démarche mettra en lumière le fait qu'autant les raisons intrinsèques (cohérence) qu'extrinsèques (praticabilité, nécessité pour les applications) entrent en considération lorsque vient le temps d'élever un énoncé au rang d'axiome. Également, il convient de constater que toute tentative de développer la théorie des ensembles - comme nous entendons le faire au cours de ce mémoire - est futile si nous n'avons qu'une idée confuse en vertu de quelles primitives ou de quels principes premiers les résultats doivent être démontrés. Ainsi, passer en revue les axiomes de la théorie ZF sera l'occasion de nous familiariser avec le contenu de notre coffre à outils.

1.1 Système axiomatique de Zermelo–Fraenkel

Pour paraphraser Paul Halmos (30), nous adopterons au cours de ce chapitre une approche axiomatique de la théorie des ensembles, mais d'un point de vue naïf ; *axiomatique* au sens où nous entendons établir une liste d'axiomes de la théorie des ensembles desquels découlera la suite de nos travaux, mais *naïf* puisque nous nous en tiendrons au langage et à la notation mathématique informelle usuelle. Nous estimons en effet que le recours à un langage entièrement formalisé serait à la fois lourd et injustifié.

Ce mémoire n'apporte aucune définition rigoureuse de la notion d'ensemble. C'est en quelque sorte le constituant fondamental, l'atome de la théorie des ensembles. L'objectif de cette section est plutôt de décrire ce qu'il est possible de faire d'une manière convenable avec de tels objets.

Informellement, un *ensemble* est un agrégat, une collection d'objets de pensée définis et distincts, ces objets étant appelés les *éléments* de l'ensemble. Bien qu'étant un nom collectif au sens où l'entend la langue française, c'est-à-dire un terme singulier qui désigne une collection d'autres entités, un ensemble ne doit pas être vu comme une simple façon de référer au singulier aux éléments qui autrement devraient porter la marque du pluriel. En effet, un ensemble n'est pas qu'une *fusion*, il y a là plus qu'une somme des parties. Or, c'est précisément ce *quelque chose de plus*, cette structure additionelle, qui transcende notre capacité à définir les choses.

De nombreuses images mentales sont utilisées pour expliquer la notion d'ensemble, allant des diagrammes de Venn, à la célèbre phrase attribuée à Cantor « Un ensemble est une *multitude* qui se laisse concevoir comme un *Un* », en passant par la comparaison avec un sac contenant

énoncés puissent être assimilés à des certitudes universelles, à des énoncés vrais dans l'absolu, à des lois de la pensée et à des truismes. Les axiomes se doivent plutôt d'être vus comme des points de départs, des conventions ou des conditions : si une structure satisfait les axiomes, alors elle satisfait les théorèmes qui en découlent. La pertinence d'une axiomatique forme un critère de bon goût au sens où une théorie incohérente n'est pas perçue comme erronée, mais plutôt comme inintéressante de par la vacuité de son domaine d'application.

des objets. Cette dernière métaphore met - bien imparfaitement - en lumière la distinction entre un ensemble et une fusion : la relation d'appartenance est l'apanage des ensembles. Le premier axiome de la théorie ZF vise d'ailleurs à abstraire l'essence de cette relation pour en faire le pilier de la théorie des ensembles.

Axiome d'extensionalité. Deux ensembles sont égaux si et seulement s'ils ont les mêmes éléments.

Comme conséquence immédiate de cet axiome, on observe que la notion d'ensemble est exempte d'une quelconque *structure d'ordre* au sens intuitif du terme. Un ensemble est uniquement déterminé par son extension, c'est-à-dire par les éléments qu'il contient. Ainsi, l'axiome d'extensionalité assure l'unicité des ensembles dont l'existence sera affirmée dans l'énoncé des différents axiomes.

Il importe de souligner que l'axiome d'extensionalité ne fait pas état d'une banale propriété logique de l'égalité. Il s'agit au contraire d'un énoncé lourd de sens au sujet de la relation d'appartenance - l'âme de la théorie des ensembles - établissant un lien entre ce concept et celui d'égalité.

Remarquons finalement qu'inhérent à l'énoncé de l'axiome d'extensionalité se trouve le fait que le domaine du discours de la théorie des ensembles est exclusivement constitué d'ensembles. À l'inverse, on peut affirmer qu'une théorie sans l'axiome d'extensionalité n'est pas une théorie traitant des ensembles.

Schéma d'axiomes de compréhension. À tout ensemble A et toute condition P (exprimée dans le langage de la théorie des ensembles) correspond un ensemble B dont les éléments sont exactement les éléments a de A pour lesquels $P(a)$ est vraie.

Exception faite de l'axiome d'extensionalité, qui revêt un caractère de définition, les principes fondamentaux de la théorie des ensembles visent à construire de nouveaux ensembles. Soulignons que dans l'approche préaxiomatique de la théorie des ensembles (12), la portée de l'axiome de compréhension était significativement plus étendue. Celui-ci spécifiait que *toute propriété définit un ensemble.* Le paradoxe de Russell [2] (65) s'est chargé d'enseigner une leçon aux mathématiciens trop enthousiastes : rien ne s'obtient sans effort. L'existence en mathématique n'est certes pas très dispendieuse, mais elle n'est pas gratuite. Pour obtenir un nouvel ensemble, il ne suffit pas de prononcer une formule magique. Il est en effet impératif d'avoir un ensemble sous la main afin de pouvoir *sélectionner* certains éléments au moyen d'une règle

2. L'énoncé de ce paradoxe va comme suit : l'ensemble des ensembles n'appartenant pas à eux-mêmes appartient-il à lui-même ? Le seul moyen d'échapper à la contradiction est d'admettre l'impossibilité de concevoir l'existence d'un tel ensemble.

précise, c'est-à-dire au moyen d'un énoncé formalisable. Autrement dit, un ensemble ne peut être défini en compréhension de manière tout à fait libre ; on ne peut que *séparer*[3] dans un ensemble préexistant les éléments satisfaisant à une certaine condition.

La paresse étant une propension naturelle, on peut être enclin à ne voir que la lourde perte qu'essuie - d'un point de vue pratique - la théorie des ensembles lorsqu'on se voit contraint d'amender le schéma de compréhension. Ce serait oublier qu'en contrepartie le bénéfice est non-négligeable : on gagne un nouvel espoir de posséder une théorie cohérente.

Suite à la rectification du schéma d'axiome de compréhension, le paradoxe de Russell devient une démonstration du fait qu'il ne peut exister d'ensemble de tous les ensembles : rien ne contient tout.

Théorème 1.1. Il n'existe pas d'ensemble de tous les ensembles.

Démonstration. Si tel était le cas, c'est-à-dire s'il existait un ensemble $\mathscr{U}$ de tous les ensembles, on pourrait appliquer le schéma de compréhension à cet ensemble et à la condition $x \notin x$. On définirait ainsi un ensemble A composé des ensembles qui n'appartiennent pas à eux-mêmes :

$$A := \left\{x \in \mathscr{U} : \quad x \notin x\right\}.$$

De deux choses l'une : soit l'ensemble A est élément de A, soit il ne l'est pas.

— Si $A \in A$, alors l'élément A - étant assujetti à la condition définissant l'ensemble A - ne s'autoappartient pas. Autrement dit $A \in A$ implique que $A \notin A$;

— Si $A \notin A$, alors A est un ensemble qui ne s'appartient pas, ce qui signifie que A satisfait à la condition faisant en sorte que $A \in A$. Autrement dit, $A \notin A$ implique que $A \in A$.

En somme, $A \in A$ si et seulement si $A \notin A$. Cette évidente contradiction nous indique que notre hypothèse est insoutenable et l'on conclut qu'il ne peut exister d'ensemble $\mathscr{U}$ composé de tous les ensembles.

■

L'extrême dépouillement de la théorie des ensembles au moment où survient le paradoxe de Russell illustre le fait qu'il ne s'agit pas tant d'une attaque à la théorie des ensembles qu'une caractéristique, une curiosité avec laquelle il nous faudra composer si l'on désire discourir au sujet des ensembles et de la relation d'appartenance. La capacité des mathématiques de s'autoréguler s'étant manifestée effectivement, la relation d'appartenance - un concept fécond - est lavée des accusations d'incohérence qui pesait sur elle.

Pour la suite, lorsque nous rencontrerons des collections déterminables par une propriété dont nous ignorons à un certain stade s'il s'agit d'un ensemble ou non, nous désignerons celles-ci

3. Certains auteurs préfèrent d'ailleurs les appellations suivantes : *axiome de sélection* ou *axiome de séparation.*

sous le nom de *classe*. De plus, nous nous autoriserons par commodité à commettre un abus de language en étendant le concept de relation d'appartenance aux classes ainsi qu'à utiliser la notation usuelle. Selon ce nouveau vocable, les ensembles sont précisément les classes dont traite le système axiomatique de ZFC, alors que les *classes propres*, c'est-à-dire les classes qui ne sont pas des ensembles, sont les collections d'objets qu'on peut décrire dans le langage de la théorie, mais dont on ne doit parler qu'à titre informel. Reformulé en ces termes, le théorème qui précède nous apprend que la collection de tous les ensembles - le domaine du discours de la théorie des ensembles - est une classe propre.

Axiome de l'ensemble vide. Il existe un ensemble dont aucun ensemble n'est élément.

L'unicité de cet ensemble, garantie par l'axiome d'extensionalité, permet de le désigner sans ambiguïté comme étant l'*ensemble vide*, que nous noterons par $\emptyset$. L'introduction de l'axiome de l'ensemble vide est motivée par la nécessité d'établir que la théorie des ensembles que nous construisons n'est pas stérile et qu'elle n'a pas le néant comme domaine du discours.

Le prochain axiome, l'axiome de la paire, permet à la théorie des ensembles, jusqu'ici très pauvre en contenu, de véritablement prendre son envol. Si le schéma d'axiome de compréhension permet la création de nouveaux ensembles, les contraintes introduites dans le but de corriger les lacunes de la compréhension non restreinte stipulent qu'on ne peut créer que par sélection à partir d'ensembles que l'on connaît. C'est la création par attrition. Or, force est de constater qu'on connaît bien peu d'ensembles. Nous ne sommes pour ainsi dire assurés que de l'existence de l'ensemble vide.

Axiome de la paire. Soit A et B deux ensembles, alors il existe un ensemble contenant A et B et eux seuls comme éléments.

Le schéma d'axiome de compréhension non restreint avait comme avantage de faire en sorte que chaque ensemble puisse aisément devenir élément d'un autre ensemble. C'est une idée qui séduit puisque cela permet de construire, sans artifice, une infinité d'ensembles à partir du connu. En réintroduisant dans un nouvel axiome un peu de l'esprit du schéma de compréhension non restreinte, il faut nous assurer de ne pas tomber dans les mêmes excès ; il faut en quelque sorte limiter l'ampleur du Big Bang ensembliste.

Deux arguments psychologiques sont susceptibles de fortifier la confiance qu'on accorde aux axiomes introduits dans le seul dessein d'élargir le domaine du discours de la théorie des ensembles. D'abord, notre position se trouve quelque peu confortée par le fait qu'à ce jour la théorie axiomatique des ensembles ZFC n'a pas été démentie ; aucune antinomie n'a émergé. En contrepartie, on pourrait se permettre d'émettre des doutes quant à savoir si de sérieuses tentatives de pousser la théorie ZFC aux limites de ses capacités ont été entreprises ; on ne peut véritablement mesurer l'ampleur des contraintes qu'on fait subir à une structure ou à un système qu'au moment où celui-ci cède.

Le second argument est ce qu'on nous appellerons provisoirement le *principe de conception itérative des ensembles*. Il s'agit d'une représentation de ce qu'on pourrait appeler la Genèse ensembliste qui permet de visualiser, d'interpréter et dans une moindre mesure de motiver l'effet des axiomes ainsi que le comportement des mécanismes limitant adéquatement l'envergure de la création d'ensembles. Ce principe n'a pas la prétention de *démontrer* les axiomes de ZFC. Il se démarque plutôt pour sa valeur didactique. Voici donc en quoi consiste le principe de conception itérative des ensembles.

Suivant la description de George Boolos (5), nous dirons qu'un *ensemble* est une collection formée à l'une ou l'autre des étapes suivantes :

- Au commencement, il n'y a que l'ensemble vide.

- Au premier jour, on forme toutes les collections possibles à partir des ensembles existants. Concrètement, on obtient $\{\emptyset\}$.

- Au deuxième jour, on forme à nouveau toutes les collections possibles d'ensembles précédemment engendrés. Plus explicitement, aux ensembles $\emptyset, \{\emptyset\}$ viennent se greffer $\big\{\emptyset, \{\emptyset\}\big\}$ et $\big\{\{\emptyset\}\big\}$.

- On poursuit inductivement cette construction.

On remarque que chaque ensemble a une et une seule date de naissance suivant laquelle son existence se poursuit *ad vitam æternam*.

Au regard du principe de conception itérative, on vérifie aisément qu'étant donné des ensembles A et B, la paire formée des ensembles A et B voit le jour au lendemain de la naissance du plus jeune parmi A et B. Cette arrivée au monde si peu éloignée de celle de ses éléments porte à croire que l'ensemble $\big\{A, B\big\}$ doit être suffisamment restreint pour constituer un ensemble.

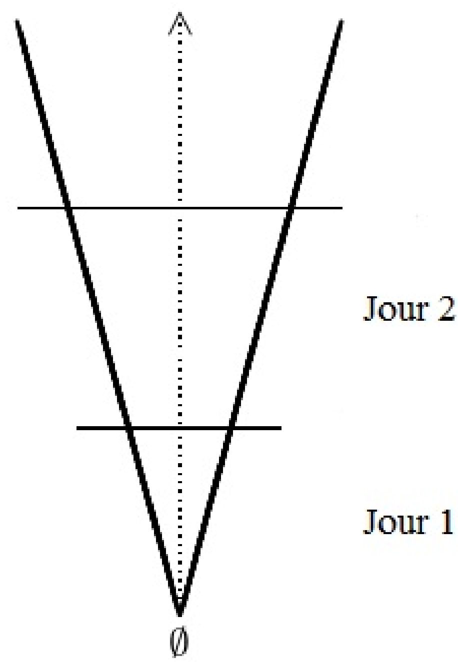

FIGURE 1.1 – Genèse ensembliste *ex nihilo*, conception itérative du domaine du discours de la théorie de Zermelo–Fraenkel

Axiome de la réunion. Pour toute collection d'ensembles, il existe un ensemble qui contient tous les éléments appartenant à au moins un ensemble de la collection donnée.

L'axiome de la réunion répond au même objectif que l'axiome de la paire : remédier aux manquements du schéma d'axiome de compréhension amendé, augmenter sa portée, réintroduire un peu de la puissance de la théorie préaxiomatique. Il demeure peu de choses à ajouter au sujet de l'axiome de la réunion sinon que du point de vue du principe de conception itérative, l'ensemble $A \cup B$ est formé le même jour que le plus jeune parmi A et B.

Rappelons en terminant que s'il nous faut de nouveaux axiomes pour créer des ensembles plus vastes à partir d'ensembles préexistants, une utilisation du schéma d'axiomes de compréhension avec une condition judicieusement choisie suffit à créer des ensembles plus restreints à partir d'ensembles préalablement connus. C'est pourquoi il est inutile d'introduire des axiomes pour l'intersection, la complémentation ou la différence symétrique.

Axiome d'infinité. Il existe un ensemble contenant 0 ainsi que le successeur de chacun de ses éléments.

Pris hors contexte, cet énoncé semble terriblement alambiqué ; sait-on seulement ce que signifient 0 et le successeur d'un élément ? La raison derrière cette apparente hypersophistication est que cet axiome répond à des considérations pratiques locales ; l'axiome d'infinité apporte une réponse à une question concernant un fragment du domaine du discours de la théorie des ensembles. Lorsque nous construirons au prochain chapitre les nombres naturels en termes ensemblistes, nous aspirerons à regrouper ceux-ci au sein d'un ensemble, l'ensemble des nombres naturels. Or, les axiomes que nous avons posés jusqu'ici se révèlent inaptes à garantir l'existence d'un tel ensemble.

Une façon intuitive de reformuler cet axiome serait de dire qu'il existe un ensemble contenant une représentation des nombres naturels. L'appellation « axiome d'infinité » est bien entendu justifiée par le fait que l'ensemble dont l'axiome postule l'existence contient une *infinité* d'éléments. Nous présenterons ultérieurement une définition rigoureuse des notions de *finitude* et d'*infinitude*.

En définitive, au regard du principe de conception itérative, l'axiome d'infinité pourrait s'énoncer ainsi : il existe un jour sans veille, c'est-à-dire un jour qui - partant du jour zéro - est au-delà de tout lendemain. Si l'introduction d'un jour limite donne un second souffle au processus de conception itérative, on vérifie sans peine que cette action n'occasionne aucune difficultée.

Axiome de l'ensemble des parties. Pour chaque ensemble, il existe une collection d'ensembles qui contient parmi ses éléments tous les sous-ensembles de l'ensemble donné.

Ayant déjà tous les outils pour considérer individuellement les sous-ensembles d'un ensemble A quelconque, il est naturel de se demander si la collection de tous les sous-ensembles de A forme elle aussi un ensemble. L'axiome de l'ensemble des parties nous permet de répondre par l'affirmative : l'ensemble des parties de A existe et l'axiome d'extensionalité nous en assure l'unicité ; nous le noterons $\mathscr{P}(A)$.

Étant donné un ensemble A contenant un nombre fini éléments, disons n éléments, imaginons que nous voulons extraire un sous-ensemble. Alors, pour chaque élément a de A, deux choix s'offrent à nous : inclure l'élément a, ou ne pas l'inclure. Ainsi, 2^n sous-ensembles de A distincts peuvent être construits. Cette remarque ne peut évidemment pas se généraliser directement aux ensembles possédant un nombre infini d'éléments. En effet, quel sens pourrait-on donner à l'expression 2^∞ ? Afin de contourner cette difficulté, il nous faut introduire un peu de notations : pour tous les ensembles A et B, l'ensemble des fonctions $f : A \to B$ est noté par B^A.

Définition 1.1. Soit A un ensemble. Pour tout sous-ensemble $B \subseteq A$, on définit la *fonction indicatrice* $\mathbb{1}_B$ de l'ensemble B comme suit :

$$\begin{aligned} \mathbb{1}_B : \quad A &\to \quad \{0,1\} \\ a &\mapsto \begin{cases} 0 & \text{si } a \notin B, \\ 1 & \text{si } a \in B. \end{cases} \end{aligned}$$

En associant à chaque élément de $\mathscr{P}(A)$ sa fonction caractéristique, on obtient une correspondance biunivoque entre $\{0,1\}^A$ et $\mathscr{P}(A)$. En d'autres termes, l'ensemble des fonctions indicatrices sur A est essentiellement la même chose que l'ensemble des parties de A

Il est d'usage d'écrire 2^A pour désigner l'ensemble des parties de A, c'est-à-dire l'ensemble des fonctions de A dans un ensemble à deux points. Cette notation, qui a l'avantage de réconcilier les cas fini et infini, met en lumière la raison pour laquelle l'ensemble des parties de A est aussi appelé l'*ensemble puissance de* A. De plus, nous verrons au prochain chapitre qu'associer l'ensemble $\{0,1\}$ au nombre naturel 2 est plus éclairé qu'il n'y paraît.

Enfin, du point de vue du principe de conception itérative des ensembles, on pourra vérifier que l'ensemble $\mathscr{P}(A)$ naît au lendemain de la venue au monde de A.

Pour être en mesure de présenter le prochain axiome, il nous faut préparer le terrain avec quelques commentaires préliminaires. Supposons que P est un prédicat à deux opérandes tels que pour chaque ensemble a il existe un unique ensemble b de sorte que $P(a,b)$ soit satisfait, alors il existe une formule F_P telle que $F_P(a) = b$ si et seulement si $P(a,b)$ est satisfaite. Intuitivement, la formule F_P possède tous les attributs d'une « relation fonctionnelle » à l'exception d'un seul : F_P est une classe propre et non pas un ensemble de couples. Nous dirons donc de F_P qu'elle est une *classe fonctionnelle.*

Considérons la classe - possiblement une classe propre - notée par B et définie de sorte que pour tout ensemble x, on ait $x \in B$ si et seulement s'il existe a dans l'ensemble A tel que $F_P(a) = x$. Alors, B est appelé l'*image de* A *par* F_P et l'on écrit $F_P(A) = B$.

Schéma d'axiomes de substitution. Étant donné un ensemble A, pour toute classe fonctionnelle F_P, l'image $F_P(A)$ est un ensemble.

Informellement, ce schéma d'axiomes énonce un principe de *petitesse* : si A est suffisamment restreint pour être un ensemble, alors toute chose intelligente qu'on peut faire aux éléments de A résulte en un ensemble ; on peut construire un nouvel ensemble à partir d'un ancien en substituant quelque chose de nouveau à chaque élément de l'ancien. Ainsi, derrière ce schéma d'axiomes se trouve l'idée que le nombre d'éléments est le seul critère déterminant si une classe est suffisamment restreinte pour constituer un ensemble.

Le schéma d'axiomes de substitution n'intervient pas immédiatement en mathématique. Tout l'intérêt de l'axiome réside dans les cas où la classe fonctionnelle ne correspond pas à une fonction au sens de la théorie des ensembles. Voici un exemple d'ensemble dont la construction découle de manière nécessaire du schéma d'axiomes de substitution.

Exemple 1.1. L'axiome d'infinité nous assure de l'existence d'un et d'un seul ensemble infini :

$$\omega := \{0, 1, 2, ...\}.$$

Tout autre ensemble infini qu'il est jusqu'ici possible de considérer doit pouvoir être obtenu par utilisations successives des axiomes présentés antérieurement. Néanmoins pour qu'un ensemble soit constructible, le processus de formation de paire, d'union, de passage à l'ensemble puissance et de sélection de sous-ensembles doit ultimement prendre fin. Par conséquent, l'ensemble

$$A := \{\omega, \mathscr{P}(\omega), \mathscr{P}(\mathscr{P}(\omega)), ...\},$$

se situe à l'extérieur du domaine des ensembles engendrable dans la théorie ZF privé du schéma d'axiomes de substitution.

En ce qui concerne le principe de conception itérative des ensembles, le schéma d'axiomes de substitution affirme l'existence d'ensembles naissant tardivement ; il agit comme une sorte d'axiome d'infinité plus sophistiqué.

Axiome de fondation. Tout ensemble non-vide A contient un élément a tel que A et a sont disjoints, c'est-à-dire tel que

$$A \cap a \ = \ \emptyset.$$

La quasi-totalité des mathématiques usuelles peut être développée sans jamais recourir à l'axiome de fondation. Ce dernier ne sert qu'à éliminer du portrait quelques cas pathologiques qui ne cadrent pas avec la conception intuitive que l'on a de la relation d'appartenance. L'axiome de fondation assure en effet qu'un ensemble et ses éléments n'ont pas le même statut, qu'aucun ensemble n'est élément de lui-même.

Théorème 1.2. Aucun ensemble ne s'auto-appartient.

Démonstration. Étant donné un ensemble A, alors l'axiome de la paire implique que $\{A\}$ est un ensemble. L'axiome de fondation nous dit qu'il existe un élément de $\{A\}$ qui est disjoint de $\{A\}$. Puisque A est l'unique élément de $\{A\}$, alors la conclusion s'impose :

$$\{A\} \cap A \ = \ \emptyset.$$

■

En corollaire du théorème précédent, il ne peut exister d'ensemble de tous les ensembles, car celui-ci devrait se contenir lui-même. Rappelons que le paradoxe de Russell, en présence du schéma d'axiome de compréhension, fournit également une réfutation de l'existence d'un ensemble universel. L'axiome de fondation conjointement avec l'axiome de la paire permet donc de fournir une deuxième preuve de ce fait. Il faut cependant s'abstenir de mal interpréter cette redondance : l'ajout de l'axiome de fondation à la théorie des ensembles naïve - dont le paradoxe de Russell a montré l'incohérence - ne suffit pas à en faire une théorie cohérente. En effet, ce n'est pas en éradiquant un symptôme qu'on assainira une théorie incohérente.

On dit d'un ensemble qu'il est *bien fondé* s'il ne se contient pas lui-même. L'axiome de fondation représente donc le choix conscient de restreindre notre domaine du discours aux seuls ensembles bien fondés. Cela ne nous force pas à croire que ceux-ci forment le seul type d'ensemble existant ; on choisit plutôt de faire l'économie d'une acclimatation aux curieuses aberrations pour le moins contre-intuitives que sont les ensembles qui se contiennent eux-mêmes en les maintenant à distance du domaine du discours de la théorie que nous nous employons à développer.

Les conséquences de l'axiome de fondation ne peuvent que renforcer la confiance que nous accordons au principe de conception itérative. D'abord, on observe que tous les ensembles engendrés par ce principe sont bien fondés, au sens où aucun d'entre eux ne s'autocontient. Ensuite, le principe de conception itérative s'articule autour d'un jour zéro, d'un commencement. Or, l'axiome de fondation va dans le même sens que cette représentation de la création de l'univers ensembliste puisqu'il permet d'enrayer la possibilité de voir surgir des descentes infinies d'inclusions

$$... \in x_5 \in x_4 \in x_3 \in x_2 \in x_1 \in x_0,$$

En effet, si une telle descente infinie d'inclusion existait, l'ensemble

$$X := \big\{x_n : \quad n \in \mathbb{N}\big\}$$

serait d'intersection non vide avec tous ses éléments. En effet, pour $n \in \mathbb{N}$ fixé,

$$X \cap x_n \ \supseteq \ x_{n+1}.$$

Or, il s'agit d'une situation que l'axiome de fondation ne saurait permettre ; un tel ensemble X entre en contradiction avec l'énoncé de cet axiome et, en corollaire, toute descente d'inclusion doit éventuellement prendre fin.

1.2 Relation d'ordre

Cette section vise à asseoir la théorie des relations binaires et de l'ordre de manière stable et solide dans le langage de la théorie des ensembles. Nous référerons aux fins détails de cette construction ensembliste des relations à de nombreuses reprises au cours de ce mémoire.

Définition 1.2. Un ensemble A est appelé un *couple* ou une *paire ordonnée* s'il existe des ensembles a et b tels que

$$A := \ \big\{\{a\}, \{a, b\}\big\}.$$

L'ensemble $\bigcup \bigcap A = a$ est appelé *première coordonnée* et l'ensemble $\bigcup \big(\bigcup A \backslash \{a\}\big) = b$ est appelé *deuxième coordonnée*. Enfin, le couple A de première coordonnée a et de deuxième coordonné b est noté (a, b).

Fort de cette définition - que l'on doit au mathématicien polonais Kazimierz Kuratowski (39) -, l'axiome de l'ensemble des parties permet de définir le *produit cartésien* de deux ensembles. On vérifie en effet que $(a,b) \in \mathscr{P}\big(\mathscr{P}(A \cup B)\big)$ chaque fois que $a \in A$ et que $b \in B$. Une fois ceci connu, c'est une simple question de routine que d'appliquer l'axiome de compréhension puis l'axiome d'extensionalité pour obtenir l'unique ensemble $A \times B$ constitué des seuls couples (a,b) avec $a \in A$ et $b \in B$. En utilisant les couples, nous pouvons formuler la théorie mathématique des relations (binaires) dans le language de la théorie ensembles.

Définition 1.3.

1. Étant donné des ensembles A et B, un sous-ensemble $R \subseteq A \times B$ est appelé une *relation sur* $A \times B$. L'*inverse de* R est une relation sur $B \times A$ définie par

$$R^{-1} := \ \big\{(b,a) \in B \times A : \ (a,b) \in R\big\};$$

2. Étant donné un autre ensemble C et une relation $S \subseteq B \times C$, le *produit de* S *et* R est défini par

$$SR := \ \big\{(a,c) \in A \times C : \ \exists b \in B \text{ tel que } (a,b) \in R \text{ et } (b,c) \in C\big\};$$

3. Si $R \subseteq A \times A$ est une relation, alors on appelle *espace relationnel* la paire (A,R) et l'on qualifie de *diagonale* l'ensemble

$$\Delta := \ \big\{(a,a) : \ a \in A\big\}.$$

Les propriétés suivantes permettent de caractériser différents types de relations d'ordre.

Définition 1.4. Soit (A,R) un espace relationnel. La relation R est dite :

1. *Réflexive* si $\Delta \subseteq R$;
2. *Antiréflexive* si $\Delta \cap R = \emptyset$;
3. *Symétrique* si $R^{-1} = R$;
4. *Antisymétrique* si $R \cap R^{-1} \subseteq \Delta$;
5. *Transitive* si $R^2 \subseteq R$;
6. *Comparative* si $R \cup R^{-1} \cup \Delta = A \times A$.

Définition 1.5. Étant donné un espace relationnel (A,R), alors la relation R est qualifiée de :

1. *Relation d'ordre strict* sur A si elle est antiréflexive et transitive ;
2. *Relation d'ordre* sur A si elle est réflexive, transitive et antisymétrique ;
3. *Relation d'ordre total (strict)* sur A si elle (anti)réflexive, transitive, antisymétrique et comparative. Une telle relation est également appelée *relation d'ordre linéaire*.

Définition 1.6.

1. Étant donné un espace relationnel (A, R), un élément $a \in A$ est dit *minimal* si $(a, b) \in R$ pour tout $b \in A \backslash \{a\}$. On dit aussi d'un tel élément a qu'il est le *plus petit élément par rapport à la relation* R ;
2. Étant donné un ensemble A, une relation R sur A telle que pour tout $B \subseteq A$ non vide le sous-ensemble B admet un élément minimal est appelée *relation de bon ordre.*

On vérifie sans difficulté que tout ensemble bien ordonné est également totalement ordonné.

1.3 Axiome du choix et principe du bon ordre

Choisir arbitrairement un objet afin de montrer qu'il satisfait à une certaine propriété en vue de conclure que cette propriété est partagée par tous les éléments de l'ensemble duquel l'élément a été extrait est une technique de preuve sans doute aussi ancienne que les mathématiques elles-mêmes. Avec le développement de l'analyse et de la théorie des nombres au 19e siècle, survinrent des situations où une infinité de choix sont nécessaires. Si les mathématiciens prévoyants prirent d'abord la peine de spécifier une règle dictant les choix, on ne tarda pas à s'abstenir de prendre cette précaution. C'est au cours de cette période d'insouciance téméraire et de laxisme endémique concernant la sélection d'une infinité d'éléments que Cantor prit inconsciemment un tournant décisif : il effectua en 1871 (46)(32), au cours d'une démonstration consacrant l'équivalence entre deux caractérisations de la continuité des fonctions réelles, une infinité de choix arbitraires pour lesquels aucune règle ne peut être définie. L'axiome du choix venait de se glisser subtilement dans l'histoire des mathématiques.

Axiome du choix I. Étant donné une famille non vide I d'ensembles non vides $\{A_i\}_{i \in I}$, le produit cartésien $\prod\limits_{i \in I} A_i$ est non vide.

Cette façon d'énoncer l'axiome du choix était celle privilégiée par Bertrand Russell (51) qui y référait sous le nom d'*axiome multiplicatif.* Si cet énoncé possède une valeur esthétique certaine, il ne fait cependant pas ressortir le lien avec la notion de « choix ». Pour cette raison, on lui préfère parfois la version alternative suivante :

Axiome du choix II. Étant donné un ensemble A et une famille non vide $B := \{B_i\}_{i \in I}$ de sous-ensembles disjoints et non vides de A, alors il existe une *fonction de choix*, c'est-à-dire une fonction $f : B \to A$ telle que pour tout $i \in I$ on ait $f(B_i) \in B_i$.

L'opinion d'un mathématicien au sujet de l'axiome du choix (sigle : AC) laisse transparaître sa philosophie concernant l'existence en mathématique. La constructibilité est-elle un critère nécessaire à l'existence d'objets mathématiques ? Est-il légitime d'admettre dans le champ des mathématiques des ensembles dont on ne peut caractériser les éléments ou des fonctions dont on ne peut spécifier la valeur en chaque point ? À celui qui estime que la possibilité de définir une fonction de choix soit tout à fait sans rapport avec la question de son existence, l'axiome du choix apparaît sans équivoque. Pour notre part, nous admettrons l'axiome du choix sans réserve particulière.

Enfin, on ne peut passer sous silence le rôle du *principe du bon ordre* - un énoncé équivalant à l'axiome du choix - dans le développement de la théorie des ensembles. Formulé par Cantor en 1883 alors qu'il s'affairait à poser les jalons de ce qui allait devenir la théorie des nombres ordinaux, le principe du bon ordre, qu'on appelle également *théorème de Zermelo*, conjecture que tout ensemble peut être muni d'une structure de bon ordre. Dans un élan d'enthousiasme, Cantor estima que le principe du bon ordre était un principe fondamental. N'ayant pu convaincre personne de la validité du principe qu'il croyait pourtant être d'une évidence désarmante, Cantor se résolut en 1895 à faire cesser la suspicion générale en démontrant le principe du bon ordre. Il semble qu'il ne fut jamais entièrement satisfait de sa démonstration puisqu'il s'opposa à sa publication.Croyant sans doute faire entrer le principe du bon ordre dans les bonnes grâces de la communauté mathématique, le mathématicien Ernst Zermelo proposa en 1904 (70) une démonstration reposant sur l'axiome de l'ensemble des parties et l'axiome du choix. La démarche de Zermelo eut plutôt pour effet de relancer la controverse. Courroucés par le théorème de Zermelo, certains mathématiciens s'attaquèrent à ce qu'ils estimaient être le maillon faible de sa preuve : l'axiome du choix.

Les débuts orageux de la théorie axiomatique des ensembles, pour intéressants qu'ils soient, ne sont pas au coeur de notre propos. Nous retiendrons pour l'essentiel que Zermelo (71) établit en 1908 les bases des la théorie axiomatique des ensembles et que certains ajouts et correctifs furent amenés par Abraham Fraenkel (26) et Thoralf Skolem (60) en 1922. Au demeurant, il sied de souligner que l'histoire de l'axiome du choix est intimement liée à celle de l'hypothèse du continu. En effet, au tournant du 20^e siècle, lors d'un congrès scientifique, le mathématicien Wacław Sierpiński constata avec stupéfaction que la demi-douzaine de mathématiciens polonais présents étaient incapables d'échanger sur leurs spécialités respectives tant leurs sujets de recherche étaient divergents. Il n'en fallut pas plus pour convaincre Sierpiński de l'importance de recentrer les projets de recherche de tous les mathématiciens polonais autour d'une seule et même branche des mathématiques afin de favoriser la coopération et d'assurer la pérennité de l'École polonaise de mathématiques.

Les universités de Cracovie, Lwów et Varsovie se virent toutes trois accorder l'exclusivité d'un certain domaine de recherche. De surcroît, l'Académie polonaise des sciences fonda une nouvelle revue mensuelle - le *Fundamenta Mathematicae* - portant sur la théorie des ensembles, les fondements des mathématiques ainsi que la topologie et ne publiant que des articles dans les langues véhiculaires qu'étaient alors l'anglais, le français, l'allemand et l'italien.

Rapidement, un nombre important de jeunes mathématiciens talentueux [4] vinrent grossir les rangs et firent de l'École polonaise l'une des plus florissantes de l'entre-deux-guerres.

Sierpiński, figure tutélaire du groupe, s'intéressait à l'axiome du choix depuis qu'il avait assisté à une conférence prononcée par Bertrand Russell sur ce sujet en 1911. C'est en s'appliquant à démontrer que certains énoncés dépendent de l'axiome du choix que l'attention de Sierpiński se porta progressivement sur la théorie des nombres cardinaux et sur l'hypothèse du continu. Les chapitres 4 et 5 de ce mémoire couvrent une partie des travaux de Sierpiński concernant l'hypothèse du continu et ses multiples conséquences.

4. La liste est pour le moins impresionnante : Stefan Banach (1892-1945), Bronisław Knaster (1893-1980), Kazimierz Kuratowski (1896-1980), Adolf Lindenbaum (1904-1941), Stanisław Mazur (1905-1981), Stefan Mazurkiewicz (1888-1945), Otto Nikodym (1887-1974), Stanisław Saks (1897-1942), Juliusz Schauder (1899-1943), Hugo Steinhaus (1887-1972), Edward Szpilrajn/Marczewski (1907-1976), Alfred Tarski (1901-1983), Stanisław Ulam (1909-1984), ...

Chapitre 2

Théorie des nombres ordinaux

Si la notion d'infini occupe un rôle important en analyse, c'est bien souvent au seul titre d'*infini potentiel*, à savoir une valeur limite de grandeurs variables finies. Cependant, héritier de Cantor, libéré de la confusion séculaire entre l'infini et l'indéfini, le mathématicien moderne est également prompt à accepter le *transfini*, l'*infini en acte*, et à concevoir qu'un ensemble infiniment grand puisse être quantifié. Cette attitude conciliante envers l'infini est l'aboutissement du développement et du perfectionnement de deux types d'outils permettant d'en saisir l'essence.

D'abord, en mettant à profit certaines identités de structure, on peut partitionner la classe des ensembles bien ordonnés, chaque classe d'équivalence étant représentée par un nombre dit *ordinal*. Cette caractérisation des types de bons ordres est d'autant plus utile dans le cadre de la théorie ZFC puisque le théorème de Zermelo affirme que tout ensemble peut être muni d'une structure de bon ordre, c'est-à-dire d'un ordre tel que tout sous-ensemble non vide admet un plus petit élément.

Ensuite, grâce aux notions de *bijection* et d'*équipotence*, on obtient un moyen adéquat et prolifique de mesurer la taille des ensembles ; *adéquat* puisque ces notions confèrent une mesure commune aux ensembles finis et infinis et *prolifique* puisque la découverte par Cantor du fait qu'il existe différentes tailles d'ensembles infinis laisse présager une riche théorie, la *théorie des nombres cardinaux*.

Les nombres naturels comblent convenablement les tâches de détailler la position des éléments d'une suite et de décrire la taille d'un ensemble fini. C'est dire que dans le cas fini, la distinction entre nombres ordinaux et nombres cardinaux ne tient qu'à des considérations linguistiques. Par contre, dans le cas infini il y a distinction effective. Étant donné un ensemble infini, on obtient deux théories distinctes selon la nature de l'abstraction que l'on effectue pour le *nombrer* ; d'une opération de neutralisation complète de la nature des éléments et des différentes structures résultera le nombre cardinal tandis qu'un filtrage perméable aux éventuelles informations sur l'ordre fournira le nombre ordinal associé à cet ensemble. L'objectif de ce chapitre est d'exposer les bases des théories des ordinaux, une théorie constituant une propédeutique indispensable à la théorie des cardinaux qui, pour sa part, fera l'objet du prochain chapitre.

2.1 Nombres ordinaux

Les nombres 0, 1, 2, etc., sont définis en termes ensemblistes de la même façon que sont définis en science les notions de mètre et de gramme : on choisi pour chaque nombre n un archétype d'ensemble réalisant la propriété d'avoir exactement n éléments puis, on assimile n à cet ensemble étalon. Il ne reste qu'à déterminer comment on entend choisir les ensembles étalons. Nos choix se doivent d'être guidés par des considérations de simplicité, d'économie et de cohésion. Le point de départ d'une méthode heuristique de construction des nombres 0, 1, 2, etc., s'impose de lui-même puisque, en vertu de l'axiome de l'ensemble vide et de l'axiome d'extensionalité, il existe un unique ensemble ne contenant aucun élément. Pour cette raison, il sied de poser

$$0 := \emptyset.$$

Pour obtenir un ensemble comportant un et un seul élément, une proposition raisonnable consiste à prendre l'ensemble contenant comme seul élément le nombre zéro précédemment défini. De même, étant donné un ensemble n, on définit le *successeur* de n, noté n^+, comme étant l'ensemble obtenu en adjoignant n aux éléments de n. En d'autres mots,

$$n^+ := n \cup \{n\}.$$

On aura par exemple

$$\begin{array}{ccccccc} 1 & := & 0^+ & = & \{0\} & = & \{\emptyset\} \\ 2 & := & 1^+ & = & \{0,1\} & = & \{\emptyset, \{\emptyset\}\} \\ 3 & := & 2^+ & = & \{0,1,2\} & = & \Big\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\Big\} \end{array}$$

Cette projection des nombres dans notre esprit fondée sur ce principe d'engendrement se poursuit *ad infinitum* et induit une relation de bon ordre canonique sur ceux-ci.

Définition 2.1. Un ensemble A est dit *inductif* s'il satisfait aux propriétés suivantes :

1. $0 \in A$;
2. $n^+ \in A$ pour tout $n \in A$.

Dans le cadre de la théorie ZFC, les images intellectuelles pour les nombres 0, 1, 2, etc. peuvent être réunies au sein d'un ensemble. En effet, l'axiome d'infinité assure l'existence d'au moins un ensemble inductif, mais il demeure cependant avare de détails concernant son contenu outre les éléments 0, 1, 2, etc. La proposition suivante soutient que l'ensemble obtenu par l'axiome d'infinité peut être épuré au sens où l'on peut le restreindre à un ensemble contenant 0, 1, 2, ... et rien d'autre.

Proposition 2.1. Il existe un ensemble inductif minimal ω.

Démonstration. Soit A l'ensemble inductif dont l'existence est affirmée par l'axiome d'infinité et soit $\mathcal{C}$ la collection de tous les sous-ensembles de A qui sont inductifs, ou de façon succincte :

$$\mathcal{C} := \big\{S \in \mathscr{P}(A) : \quad S \text{ est inductif}\big\}.$$

Nous soutenons que $\omega := \bigcap\limits_{S \in \mathcal{C}} S$ est inclus dans tout ensemble inductif. Voici quelques informations permettant d'étayer cette affirmation.

D'abord, il ne faut aucun doute que ω est un ensemble inductif. Ensuite, étant donné un ensemble inductif B quelconque, on a que $A \cap B \subseteq A$. On vérifie aisément que l'ensemble $A \cap B$ est inductif ; il s'agit donc d'un élément de $\mathcal{C}$. Comme ω est le plus petit élément de $\mathcal{C}$ qui est inductif, il s'ensuit que $\omega \subseteq B$.

■

Définition 2.2. Un *nombre naturel* est un élément de l'ensemble inductif minimal.

Ayant comme but avoué d'étendre notre capacité d'énumération au-delà des nombres naturels, nous devons immanquablement isoler la caractéristique fondamentale de ceux-ci. La notion d'ordre semble à première vue être ce qui distingue une énumération d'une simple accumulation. Par conséquent, la relation de bon ordre sur les nombres naturels apparaît être l'idée à abstraire en vue de procéder à une généralisation.

Définition 2.3. Étant donné un ensemble partiellement ordonné $(A, \leq)$ et un élément $a \in A$, la *section commençante déterminée par* a est l'ensemble

$$s(a) := \big\{x \in A : \quad x < a\big\}.$$

À la lumière de cette définition, la section commençante déterminée par a est l'ensemble des prédécesseurs [1] stricts de a.

Proposition 2.2. Tout nombre naturel est un ensemble bien ordonné tel que la section commençante déterminée par chaque élément est égale à cet élément.

1. Dans l'intention de soulager le texte, nous nous autorisons à parler des *prédécesseurs* ou d'*antécédent* de a pour désigner les éléments qui sont inférieurs à a dans la relation d'ordre considérée.

Démonstration. Soit n un nombre naturel. Par construction,

$$n := \{0, 1, 2, ..., k-1, k, k+1, ..., n-1\}.$$

Étant donné un élément k de n, la section commençante déterminée par k dans le bon ordre canonique est l'ensemble

$$s(k) \ = \ \{0, 1, 2, ..., k-1\},$$

qui est précisément la définition du nombre naturel k.

■

Bien que ce soit le mathématicien allemand Georg Cantor qui a posé les jalons de la théorie des nombres ordinaux, la définition standard des nombres ordinaux que voici est due au mathématicien américano-hongrois John von Neumann (67).

La propriété de minimalité de ω peut être exprimée ainsi : soit $S \subseteq \omega$ avec $0 \in S$ et tel que $n^+ \in S$ chaque fois que $n \in S$, alors $S = \omega$. Cette propriété, connue sous le nom de *principe d'induction mathématique*, caractérise ω de façon unique et permet de définir rigoureusement les nombres ordinaux.

Définition 2.4. Un *nombre ordinal* est un ensemble bien ordonné α tel que $s(\zeta) = \zeta$ pour tout ζ dans α.

Notons que toute relation d'ordre sur un ensemble est déterminée de façon unique par ses sections commençantes. Il s'ensuit que si tant est qu'il soit possible de définir une relation de bon ordre sur un ensemble de façon à en faire un nombre ordinal, il n'y a qu'un et un seul moyen de le faire.

Proposition 2.3. Les énoncés suivants sont équivalents :

1. L'ensemble α est un nombre ordinal ;
2. L'ensemble α est un ensemble *transitif*[2] bien ordonné par la relation d'appartenance.

Démonstration.

$(1 \Rightarrow 2)$ Comme α est un nombre ordinal, on a que $s(\zeta) = \zeta$ pour tout ζ dans α. Il en découle immédiatement que α est un ensemble transitif. En outre, comme les prédécesseurs de ζ dans la relation de bon ordre sur α sont précisément les éléments de ζ, la relation d'ordre en question est impérativement la relation d'appartenance.

2. Un ensemble A est dit *transitif* si tout élément a de A est également un sous-ensemble de A. De façon équivalente, A est dit transitif si $b \in a \in A$ implique que $b \in A$.

($1 \Leftarrow 2$) Par hypothèse, l'ensemble α est bien ordonné par la relation d'appartenance. Étant donné β et ζ des éléments de α, on a $\beta < \zeta$ si et seulement si $\beta \in \zeta$. Cela signifie que $s(\zeta) \subseteq \zeta$ et, plus précisément, on a

$$s(\zeta) \;=\; \zeta \cap \alpha. \tag{2.1}$$

Par transitivité, $\zeta \in \alpha$ implique que $\zeta \subseteq \alpha$. Ainsi, l'équation (2.1) revient à dire que $s(\zeta) = \zeta$.

■

En vertu de cette proposition, l'ensemble ω est un nombre ordinal. Lorsqu'il nous arrivera de vouloir parler spécifiquement des nombres ordinaux qui ne sont pas des nombres naturels, nous parlerons des *nombres transfinis.*

Proposition 2.4. L'ensemble successeur d'un nombre ordinal est un nombre ordinal.

Démonstration. Soit α un nombre ordinal et considérons l'ensemble successeur α^+. Il nous faut dans un premier temps définir une relation d'ordre sur α^+. La façon la plus naturelle de procéder est d'étendre la relation de bon ordre dont est muni α à l'ensemble α^+. Comme α^+ est obtenu de α en y adjoignant l'élément α, il suffit de poser $\zeta < \alpha$ pour tout $\zeta \in \alpha$.

Il ne reste qu'à vérifier que $s(\zeta) = \zeta$ pour tout $\zeta \in \alpha^+$. Pour ce faire, on doit distinguer deux cas :

1. Soit $\zeta \in \alpha$, auquel cas le fait que α soit un nombre ordinal implique que $s(\zeta) = \zeta$. Par définition de l'ordre sur α^+, cette section commençante de ζ dans α coïncide avec celle de ζ dans α^+ ;
2. Soit $\zeta = \alpha$ auquel cas la définition de l'ordre sur α^+ nous donne directement $s(\alpha) = \alpha$.

■

Comme nous l'avons maintes fois répété, la notion de nombre ordinal est fondamentalement issue du bon ordre ; les nombres ordinaux *se suivent* et sont définis en termes d'ordre. Il va donc de soi de se demander quel nombre ordinal l'ensemble ω suit-il ? La réponse a de quoi surprendre : ω suit tous les nombres naturels mais aucun en particulier au sens où il n'a pas de plus grand prédécesseur. En prévision de la découverte éventuelle d'autres nombres ordinaux partageant avec ω la particularité de n'être le successeur immédiat d'aucun autre nombre ordinal, nous dirons d'eux qu'ils sont des *ordinaux limites.*

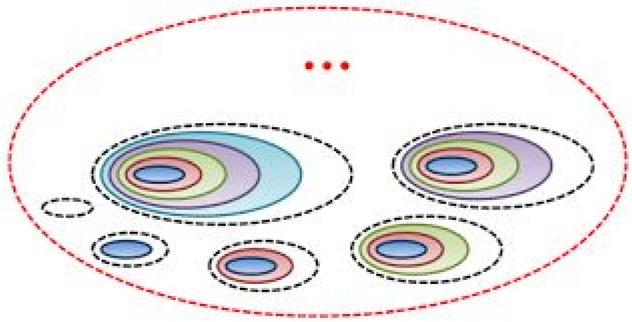

FIGURE 2.1 – Représentation visuelle du premier ordinal transfini ω, l'ensemble des ordinaux finis. Source : http ://naturelovesmath.blogspot.ca/2012/05/infini-et-nombres-transfinis.html

Ayant montré que ω est un nombre ordinal et que les nombres ordinaux sont clos sous la prise de successeur, le processus d'extension de notre capacité d'énumération s'étend naturellement pour inclure les successeurs de ω. On a donc

$$0, 1, 2, ..., \omega, \omega + 1, \omega + 2, ...,$$

où $\omega + n$ désigne le n-ième successeur de ω. À ce stade, la construction des nombres ordinaux semble atteindre un point stationnaire. La question quant à savoir s'il existe quelque chose au-delà des $\omega + n$ au même titre que ω se situe au-delà des nombres naturels appelle l'existence d'un ensemble contenant ω ainsi que tout élément pouvant être obtenu à partir de ω par formation répétée de successeurs. Malheureusement, l'axiome d'infinité se révèle cette fois inapte à résoudre la question. C'est plutôt le schéma d'axiomes de substitution qui permet de montrer l'existence d'un second ensemble inductif limite en construisant, dans un premier temps, l'ensemble suivant :

$$T := \{\omega, \omega + 1, \omega + 2, ...\}.$$

L'ensemble T n'est pas un nombre ordinal puisque l'élément ω n'est pas égal à sa section commençante. En effet, $s(\omega) = \emptyset$. Par contre, il suffit de considérer l'ensemble

$$\omega \cdot 2 := \; T \cup \omega \; = \; \{0, 1, 2, ..., \omega, \omega + 1, \omega + 2, ...\}$$

pour remédier à ce problème. Ce choix de notation quelque peu insolite prendra tout son sens lorsque nous définirons la multiplication ordinale. Pour faire suite à la formation de ce second nombre ordinal limite, l'enchaînement de formation du successeurs et de substitution se poursuit inlassablement.

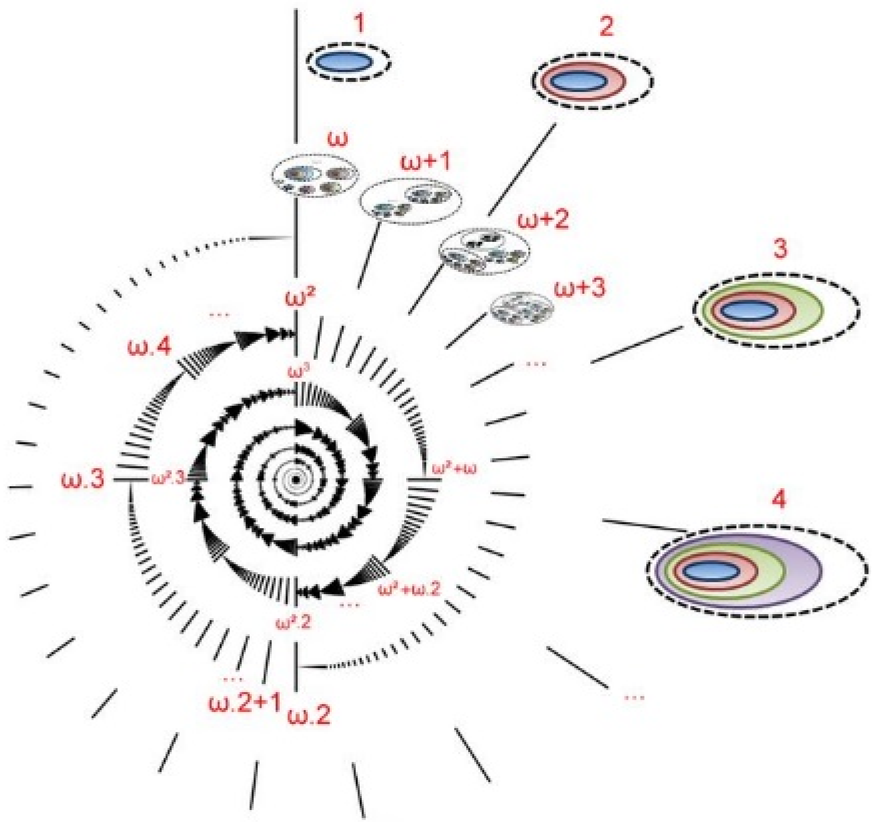

FIGURE 2.2 – Représentation des nombres ordinaux jusqu'à ω^ω. Chaque tour de spirale représente une puissance de ω. Source : http ://naturelovesmath.blogspot.ca/2012/05/infini-et-nombres-transfinis.html

2.2 Propriétés des nombres ordinaux

Proposition 2.5. Tout élément d'un nombre ordinal est lui-même un nombre ordinal.

Démonstration. Soit ζ un élément d'un nombre ordinal α. Munissons ζ de la relation de bon ordre obtenue par restriction du bon ordre sur α.

Soit $\eta \in \zeta$. La section commençante déterminée par η dans ζ est la même que la section commençante déterminée par η dans α. Puisque cette dernière n'est nulle autre que l'ensemble η lui-même, il en est de même pour la première. Ainsi, ζ est un nombre ordinal.

$\blacksquare$

Une autre façon d'interpréter ce résultat est de dire que toute section commençante d'un nombre ordinal est un nombre ordinal.

La prochaine propriété des nombres ordinaux digne de mention repose sur le principe d'induction transfinie, une extension aux ensembles bien ordonnés du principe d'induction mathématique. L'induction transfinie s'articule autour de la notion de section commençante :

Principe d'induction transfinie. Soit B un sous-ensemble d'un ensemble bien ordonné A. Supposons que, pour tout $a \in A$, l'appartenance à B de tous les prédécesseurs de a implique l'appartenance à B de a lui-même, alors nous avons impérativement que $B = A$. Formellement :

$$\text{Si } \forall a \in A \ \left[s(a) \subseteq B \Rightarrow a \in B \right] \text{ alors } B = A.$$

Démonstration. Supposons que l'hypothèse tienne, mais que la conclusion soit fausse. Alors, l'ensemble non vide $A \backslash B$ admet un plus petit élément, disons a. Ceci implique que tout élément de la section commençante $s(a)$ appartient à B. D'après l'hypothèse d'induction, on doit avoir $a \in B$. De cette contradiction évidente on conclut que $A \backslash B$ est vide.

■

Contrairement au principe d'induction mathématique, le principe d'induction transfinie ne requiert pas d'hypothèse sur un élément de départ. Si a_0 est le plus petit élément de A, alors la section commençante en a_0 est vide, et en conséquence, $s(a_0) \subseteq B$; l'hypothèse d'induction transfinie exige donc que a_0 appartienne à B.

Définition 2.5.

1. Une application $f : A \to B$ entre deux ensembles partiellement ordonnés *préserve l'ordre* si $a_1 \leq a_2$ dans A implique que $f(a_1) \leq f(a_2)$ dans B ;
2. Un *isomorphisme d'ordre* entre deux ensembles partiellement ordonnés est une bijection préservant l'ordre dont l'inverse f^{-1} préserve l'ordre. Pour alléger le texte, nous dirons que deux ensembles sont *semblables* - notée $A \cong B$ - s'il existe un isomorphisme d'ordre entre les deux.

Lemme de rigidité 2.6. Soit A et B des ensembles bien ordonnés et $f_1, f_2 : A \to B$ deux isomorphismes d'ordre. Alors $f_1 = f_2$.

Démonstration. Soit f_1 et f_2 deux isomorphismes d'ordre entre les ensembles bien ordonnés A et B qu'on peut supposer non vides. Soit A_2 l'ensemble des éléments a de A tels que $f_1(a) \neq f_2(a)$ et soit $A_1 := A \backslash A_2$. Comme l'image par n'importe quel isomorphisme d'ordre du plus petit élément de A se doit d'être le plus petit élément de B, l'ensemble A_1 est non vide. Posons $B_1 := f_1(A_1) = f_2(A_1)$.

Supposons que A_2 est non vide et que a_2 est son plus petit élément. Alors, $f_1(a_2)$ et $f_2(a_2)$ se caractérisent tous deux par le fait qu'ils sont le plus petit élément de $B \backslash B_1$. En corollaire, ils doivent être égaux, ce qui constitue une évidente contradiction. ■

Proposition 2.7. Si deux nombres ordinaux sont semblables, alors ils sont égaux.

Démonstration. Soit α et β des nombres ordinaux et $f : \alpha \to \beta$ un isomorphisme d'ordre. Nous montrerons que pour chaque ζ dans α on a $f(\zeta) = \zeta$. Pour ce faire, nous exploiterons le fait que pour chaque élément ζ d'un nombre ordinal la section commençante $s(\zeta)$ admet une borne supérieure, à savoir l'élément ζ lui-même. Posons

$$S := \left\{\zeta \in \alpha : \ f(\zeta) = \zeta\right\}.$$

Si $\zeta \in \alpha$ est tel que $s(\zeta) \subseteq S$, alors f envoie tout prédécesseur de ζ sur lui-même. Il va sans dire que les isomorsphimes d'ordres préservent les bornes supérieures. Ainsi, il résulte que ζ et $f(\zeta)$ sont tous deux égaux à la borne supérieure de $s(\zeta)$ et, conséquemment, $\zeta = f(\zeta)$. Ayant montré que $\zeta \in S$ chaque fois que $s(\zeta) \subseteq S$, le principe d'induction transfinie implique que $S = \alpha$ et il résulte de ceci que $\alpha = \beta$.

■

Proposition 2.8. Soit α et β des nombres ordinaux tel que $\beta \subseteq \alpha$. Alors, β est une section commençante de α. Si de plus $\beta \subsetneq \alpha$, alors $\beta \in \alpha$.

Démonstration.

1. Soit $\gamma \in \beta$ et $\zeta \in \alpha$ des éléments tels que ζ vienne avant γ dans le bon ordre sur α. Puisque la relation d'ordre sur les ordinaux est la relation d'appartenance, on a donc que $\zeta \in \gamma \in \beta$. Comme β est un ensemble transitif, il s'ensuit que $\zeta \in \beta$.
2. Supposons que $\beta \subsetneq \alpha$. Alors, l'ensemble $\alpha \backslash \beta$ est non vide. Soit γ l'élément minimal de $\alpha \backslash \beta$. Nous entendons montrer que $\beta = \gamma$.
   - ($\subseteq$) Supposons qu'il existe un élément $\zeta \in \beta$ qui n'appartient pas à γ. Comme ζ et γ appartiennent tous deux à α - un ensemble bien ordonné par la relation d'appartenance - on doit avoir que $\gamma \in \zeta \in \beta$. Puisque β est transitif, cela implique que $\gamma \in \beta$, contredisant la définition de γ ;
   - ($\supseteq$) Soit $\zeta \in \gamma$. On doit également avoir que $\zeta \in \beta$ sans quoi il s'agirait d'un élément de $\alpha \backslash \beta$ inférieur à γ. Cette éventualité doit être écartée puisqu'elle irait à l'encontre de la propriété de minimalité de γ.

■

La dernière proposition met en relief le principe de trichotomie que voici :

Proposition 2.9. Étant donné deux nombres ordinaux α et β, soit ceux-ci sont égaux, soit $\alpha \in \beta$, soit $\beta \in \alpha$.

Démonstration. Considérons l'ensemble $\alpha \cap \beta$. Il s'agit d'un sous-ensemble - possiblement vide - de α et β. Comme $\alpha \cap \beta$ est l'intersection de deux ensembles transitifs, il s'agit d'un ensemble transitif. De plus, en tant que sous-ensemble de α, l'ensemble $\alpha \cap \beta$ est ordonné par la relation d'appartenance. Puisqu'il satisfait aux deux propriétés de la caractérisation des nombres ordinaux donnée à la proposition 2.3, l'ensemble $\alpha \cap \beta$ est un nombre ordinal.

1. Si $\alpha \cap \beta = \alpha$, alors $\alpha \subseteq \beta$ et, par la proposition 2.8, on conclut que $\alpha \in \beta$ ou $\alpha = \beta$;
2. De façon similaire, $\alpha \cap \beta = \beta$ implique que $\beta \in \alpha$ ou $\beta = \alpha$;
3. Le seul autre cas possible est que $\alpha \cap \beta$ soit un sous-ensemble propre de α et de β. Par la proposition 2.8, on obtient alors que $(\alpha \cap \beta) \in \alpha$ et $(\alpha \cap \beta) \in \beta$. En conséquence, $(\alpha \cap \beta) \in (\alpha \cap \beta)$, ce qui est impossible en raison de l'axiome de fondation.

■

Nous avons glané suffisamment d'informations pour affirmer que tout ensemble de nombres ordinaux est totalement ordonné et nous sommes en mesure d'offrir trois interprétations équivalant à l'expression $\beta < \alpha$:

1. $\beta \in \alpha$;
2. $\beta \subset \alpha$;
3. β est une section commençante de α, ou α est une *continuation* de β.

Mais il y a plus :

Proposition 2.10. Tout ensemble de nombres ordinaux est bien ordonné par la relation d'appartenance. Par abus de langage, on dit que la classe des nombres ordinaux est bien ordonnée par la relation d'appartenance.

Démonstration. Pour ce qui suit, l'ordre $<$ est la relation d'appartenance $\in$. Supposons que Ω est un ensemble non vide de nombres ordinaux et que α est un élément de Ω. Alors, de deux choses l'une :

1. Soit $\alpha \leq \beta$ pour tout β dans Ω, auquel cas α est le premier élément de Ω ;
2. Soit il existe un élément β dans Ω tel que $\beta < \alpha$, c'est-à-dire tel que $\beta \in \alpha$. En d'autres termes, les ensembles α et Ω possèdent un ou des élément(s) commun(s). Puisque α est un ensemble bien ordonné, le sous-ensemble $\alpha \cap \Omega$ admet un premier élément, disons α_0. Si $\eta \in \Omega$, alors
   a) Ou bien $\alpha \leq \eta$ auquel cas $\alpha_0 < \eta$,
   b) Ou bien $\eta < \alpha$ auquel cas $\eta \in \alpha \cap \Omega$. Par minimalité de α_0, on a alors que $\alpha_0 \leq \eta$;

   Ainsi, dans un cas comme dans l'autre, Ω admet un premier élément, à savoir α_0.

■

Proposition 2.11. Tout ensemble de nombres ordinaux admet une borne supérieure.

Démonstration. Soit Ω une collection de nombres ordinaux et considérons l'union $\bigcup\limits_{\zeta\in\Omega}\zeta$ de tous les éléments de cette collection. Tous les $\zeta \in \Omega$ étant des ensembles transitifs, il apparaît clairement que $\bigcup\limits_{\zeta\in\Omega}\zeta$ en est également un.

On sait par la proposition 2.5 que tous les éléments d'un nombre ordinal sont des nombres ordinaux. Il en découle que $\bigcup\limits_{\zeta\in\Omega}\zeta$ est un ensemble d'ordinaux. Ensuite, par la proposition qui précède, on a que $\bigcup\limits_{\zeta\in\Omega}\zeta$ est bien ordonné par la relation d'appartenance.

Étant un ensemble transitif bien ordonné par la relation d'appartenance, l'union $\bigcup\limits_{\zeta\in\Omega}\zeta$ est lui-même un nombre ordinal. Sans surprise, $\bigcup\limits_{\zeta\in\Omega}\zeta$ constitue une borne supérieure des éléments de Ω. Il s'agit même de la plus petite borne supérieure de Ω. En effet, si β est une autre borne supérieure de Ω, alors $\zeta \subseteq \beta$ quel que soit $\zeta \in \Omega$. Cela signifie que $\bigcup\limits_{\zeta\in\Omega}\zeta$ est le suprémum de la collection Ω.

■

La proposition précédente entraîne qu'il ne saurait exister d'ensembles contenant exactement tous les nombres ordinaux. En effet, si un tel ensemble existait, nous pourrions former la borne supérieure de tous les nombres ordinaux. Or, cette borne supérieure serait un nombre ordinal supérieur ou égal à tout nombre ordinal. Cette éventualité est à exclure puisqu'un nombre ordinal admet toujours un successeur. Une conclusion s'impose : cela n'a pas de sens de parler de l'ensemble de tous les nombres ordinaux. Les ordinaux forment une classe propre que nous désignerons par Ord. Cette particularité est connue sous le nom de *paradoxe de Burali-Forti* (11). Entendons-nous, le mot « paradoxe » est à prendre au sens étymologique, c'est-à-dire « une proposition contraire au sens commun ». Il n'y aurait de véritable antinomie que dans un système permettant la construction ou affirmant l'existence d'un ensemble des ordinaux, ce qui n'est pas le cas dans le cadre de la théorie axiomatique ZFC. Il est simplement question ici d'une curiosité de la théorie des ensembles avec laquelle il nous faut composer.

Si l'on peut apprécier la théorie des nombres ordinaux pour ce qu'elle est, la construction tarabiscotée de ceux-ci fait craindre qu'ils n'aient qu'une utilité somme toute limitée, alors que tant s'en faut. Le théorème suivant affirme que tout ensemble bien ordonné est semblable à un unique nombre ordinal.

Théorème 2.12. Tout ensemble bien ordonné A est semblable à un et un seul nombre ordinal, appelé *type d'ordre* de A et noté $\mathrm{ord}(A)$.

Démonstration.

- L'unicité est immédiate puisque deux nombres ordinaux sont semblables si et seulement s'ils sont égaux.
- Étant donné un ensemble bien ordonné A, supposons qu'un élément a de A soit tel que la section commençante déterminée par chaque prédécesseur de a est semblable à un certain nombre ordinal (nécessairement unique). Par le schéma d'axiomes de substitution on forme, pour chaque prédécesseur x de a, le singleton

  $$\big\{\alpha \in \mathrm{Ord} : \quad \alpha \text{ est un nombre ordinal semblable à } s(x)\big\}.$$

  En effectuant l'union sur les prédécesseurs de a de ces singletons, on obtient un ensemble Φ constitué exactement des nombres ordinaux semblables à une section commençante déterminée par un des prédécesseurs de a.

  Puisque la borne supérieure des prédécesseurs de a est l'élément a lui-même et que les isomorphismes d'ordre préservent les bornes supérieures, il s'ensuit que $s(a)$ est semblable à la borne supérieure de la collection de nombres ordinaux Φ (lui-même un nombre ordinal en vertu de la proposition 2.11). Ce raisonnement prépare la voie à une application du principe d'induction transfinie. La conclusion est que chaque section commençante dans A est semblable à un certain nombre ordinal.

  De la même façon que précédemment, on forme par substitution l'ensemble Ω composé des nombres ordinaux semblables à une quelconque section commençante de A. La borne supérieure de cet ensemble d'ordinaux est un nombre ordinal semblable à la borne supérieure des sections commençantes, à savoir l'ensemble A.

■

2.3 Arithmétique ordinale

Les opérations arithmétiques sur les nombres naturels peuvent aisément être définies par récurrence. Par contre, pour le cas plus général de l'arithmétique ordinale, nous allons prévilégier une approche intrinsèque.

Addition

L'addition ordinale repose sur la concaténation : on veut mettre bout à bout des chaînes. Cela doit cependant être fait précautionneusement puisqu'on peut avoir affaire à des chaînes qui ne sont pas disjointes. Cette difficulté peut être contournée en peignant chaque chaîne d'une couleur différente de sorte à rendre chacun de leurs éléments distincts. Par exemple, étant donné la famille d'ensembles $\{A_i\}_{i\in I}$, on obtiendra une famille d'ensembles deux à deux disjoints en considérant plutôt $\{\widetilde{A}_i\}_{i\in I}$, où $\widetilde{A}_i := A_i \times \{i\}$.

Pour définir l'addition ordinale, on ne peut mettre à profit les caractéristiques propres aux nombres ordinaux puisque ceux-ci ne sont pas deux à deux disjoints et qu'en les peignant de façon à en faire des ensembles deux à deux disjoints la propriété définissant les nombres ordinaux - à savoir être un ensemble tel que $s(\zeta) = \zeta$ pour chaque élément ζ - est perdue. Par conséquent, nous devrons baser l'addition ordinale sur la seule propriété d'être un ensemble bien ordonné.

Définition 2.6. Étant donné $\{A_i\}_{i\in I}$ une famille d'ensembles bien ordonnés et deux à deux disjoints indicée par un ensemble bien ordonné I, on appel *somme ordinale de* $\{A_i\}_{i\in I}$ l'ensemble $\bigcup\limits_{i\in I} A_i$ muni de la relation de bon ordre suivante : étant donné des éléments a et b de $\bigcup\limits_{i\in I} A_i$ avec $a \in A_i$ et $b \in A_j$, alors $a < b$ si et seulement si l'un des cas suivants se réalise.

- Soit $i < j$ dans l'ordre donné sur I ;
- Soit $i = j$ et a précède b dans l'ordre donné sur A_i.

La définition de l'addition des nombres ordinaux s'obtient aisément à partir de la somme ordinale. Étant donné $\{\alpha_i\}_{i\in I}$ une famille de nombres ordinaux indicée par un ensemble bien ordonné I et $\{A_i\}_{i\in I}$ une famille d'ensembles bien ordonnés disjoints avec $\mathrm{ord}(A_i) = \alpha_i$ pour chaque $i \in I$, la somme de nombres ordinaux $\sum\limits_{i\in I} \alpha_i$ est définie comme étant le type d'ordre de la somme ordinale de la famille $\{A_i\}_{i\in I}$. Autrement dit,

$$\sum_{i\in I} \alpha_i := \mathrm{ord}\left(\bigcup_{i\in I} A_i\right).$$

Notons que la somme $\sum\limits_{i\in I} \alpha_i$ ne dépend pas du choix particulier des ensembles A_i. De plus, on vérifie aisément que dans le cas où tous les α_i sont des nombres naturels, cette définition de la somme des nombres ordinaux coïncide avec la somme intuitive des nombres naturels.

En étendant aux nombres ordinaux l'opération d'addition, nous acceptons de payer la rançon de la généralité, soit l'éventualité de voir succomber certaines propriétés usuelle de l'addition. Réjouissons-nous de la survivance des propriétés suivantes :

1. Le nombre ordinal 0 constitue l'identité additive ;
2. L'addition ordinale est associative ;
3. L'addition ordinale est strictement croissante et continue à droite. En particulier,

$$\alpha + 1 = \alpha^+.$$

Le comportement singulier de l'addition ordinale résulte presque toujours de ce que la loi commutative n'est pas vérifiée. Par exemple, si dans une addition le fini précède l'infini alors il y disparaît, mais s'il lui succède alors il subsiste et se combine avec lui en un infini nouveau. Cela s'explique par le fait qu'adjoindre un élément devant une suite infinie n'altère pas le type d'ordre de celle-ci, *a contrario* adjoindre un élément à la fin d'une suite infinie détruit la similitude ; l'ancien ensemble n'avait pas de dernier élément et avait un seul élément sans prédécesseur alors que le nouveau en a deux et admet un dernier élément.

Multiplication

L'étude des produits d'ensembles bien ordonnés est grandement facilitée par une utilisation judicieuse des sommes infinies. Étant donné des ensembles bien ordonnés A et B, on cherche à définir leur produit comme le résultat obtenu en ajoutant A à lui-même B fois. Pour donner un sens à ceci, on construit une famille d'ensembles bien ordonnés disjoints, dont chacun est semblable à A, indicée par l'ensemble bien ordonné B. En d'autres termes, on considère la famille $\{A_b\}_{b\in B}$.

Définition 2.7. Étant donné des ensembles bien ordonnés A et B, on appel *produit ordinal de A et B* le produit cartésien $A \times B$ muni de l'ordre lexicographique inverse, c'est-à-dire que si $a, c \in A$ et $b, d \in B$ alors $(a, b) < (c, d)$ signifie :

- Soit $b < d$;
- Soit $b = d$ et a précède c dans l'ordre donné sur A.

Fort de cette notion de produit ordinal de deux ensembles bien ordonnés, on peut maintenant définir le produit de deux nombres ordinaux α et β. Pour ce faire, il suffit de choisir des ensembles bien ordonnés A et B avec $\text{ord}(A) = \alpha$ et $\text{ord}(B) = \beta$ et de poser $\alpha\beta$ comme étant le nombre ordinal de $A \times B$, de sorte que

$$\alpha \cdot \beta := \text{ ord}(A \times B).$$

Le produit de nombres ordinaux est défini sans ambiguïté - c'est-à-dire indépendamment des choix arbitraires des ensembles bien ordonnés A et B - et conserve les propriétés de la multiplication usuelle suivantes :

1. Le nombre 0 est un élément absorbant pour la multiplication ;
2. Le nombre 1 est une identité multiplicative ;
3. La multiplication ordinale est associative ;
4. La multiplication ordinale est distributive à droite ;
5. Aucun nombre ordinal non nul n'est diviseur de zéro .

La loi de commutativité n'étant pas vérifié il faut faire preuve de prudence en traitant du produit de nombres ordinaux. Par exemple, il ne faut pas confondre $2 \cdot \omega$ - une suite de type ω de couples - avec $\omega \cdot 2$, un couple de suites infinies.

On a vu que l'addition répétée conduit à la définition des produits ordinaux. De même la multiplication répétée peut être utilisée pour définir les exposants ordinaux. Les détails étant plutôt sophistiqués et non requis pour la suite de notre propos, nous nous en tiendrons à souligner que de plus amples informations à ce sujet sont données dans (59).

Chapitre 3

Théorie des nombres cardinaux

Nous avons développé au chapitre précédent une théorie procurant une caractérisation raisonnable des ensembles bien ordonnés. Cependant, bien que formant une ingénieuse abstraction de la méthode d'énumération, les nombres ordinaux ne constituent pas un outil parfaitement adéquat de comparaison de la taille des ensembles ; leur utilité est ailleurs. Il ne suffit pas, pour comparer la tailles d'ensembles A et B, de les bien ordonner puis de comparer $\mathrm{ord}(A)$ et $\mathrm{ord}(B)$. En effet, la difficulté surgit du fait que certains ensembles peuvent être bien ordonnés d'une multitude de façons sans que le type d'ordre soit forcément préservé. En d'autres mots, le nombre ordinal d'un ensemble bien ordonné mesure le bon ordre plus qu'il ne mesure l'ensemble. Pour prendre un exemple concret, considérons l'ensemble des nombres naturels bien ordonnés de la façon évidente - un ordre de type ω - puis introduisons un nouvel ordre sur cet ensemble en plaçant 0 après tout le reste ; on obtient un bon ordre de type $\omega + 1$.

Bien que nous continuerons d'attacher quelque attention aux nombres ordinaux, nos esprits graviteront principalement autour d'autres thèmes ; nous nous mettrons en quête d'une méthode effective de comparaison de la taille des ensembles.

Cela dit, remarquons qu'étant donné des ensembles A et B munis d'un bon ordre, une condition nécessaire et suffisante pour que $\mathrm{ord}(A) < \mathrm{ord}(B)$ est qu'il existe une bijection préservant l'ordre - un isomorphisme d'ordre - entre A et une certaine section commençante de B. Par conséquent, la connaissance des nombres ordinaux des ensembles bien ordonnés A et B n'est pas requise pour comparer leurs tailles ordinales. Ce que nous essayons de mettre en évidence est que la comparaison qualitative au moyen du concept d'isomorphisme d'ordre est porteuse de sens ; les informations complémentaires issues de l'aspect quantitatif ne sont pas toujours probantes.

Si on laisse tomber toute allusion aux structures d'ordres, la notion de *correspondance biunivoque* - où *bijection* - se révèle être la clé de voute d'une méthode analogue permettant de comparer qualitativement la tailles des ensembles.

3.1 Théorème de Cantor–Bernstein

Informellement, la *cardinalité* d'un ensemble A est un outil d'évalutation $\operatorname{card}(A)$ tel que, étant donné des ensembles A et B, l'égalité $\operatorname{card}(A) = \operatorname{card}(B)$ est vérifiée si et seulement s'il existe une bijection entre A et B. On appelle *nombre cardinal* tout objet qui est la cardinalité d'un ensemble.

Pour définir rigoureusement la notion de nombres cardinaux, plusieurs avenues s'offrent à nous, chacune ayant ses avantages et ses inconvénients. De la même façon qu'il n'est pas nécessaire de connaître et de comprendre le mode de fonctionnement de la jauge à carburant d'un automobile afin d'y faire la lecture des mesures, il n'est pas toujours indispensable d'avoir une compréhension complète et approfondie de la façon dont la cardinalité représente la taille des ensembles. Ce qui prime est que cette notion s'harmonise avec la méthode qualitative et qu'elle soit cohérente avec notre conception de l'addition, de la multiplication et dans une moindre mesure de l'exponentiation.

Afin d'affranchir les nombres cardinaux de la démarche choisie pour les définir et afin d'illustrer clairement que les propriétés essentielles du concept en sont indépendantes, nous reportons à plus tard la présentation d'une construction *ad hoc* des nombres cardinaux.

Définition 3.1.

1. Deux ensembles A et B sont dits *équipotents*, noté $A \sim B$, s'il existe une bijection entre eux ;
2. Si A est équipotent à un sous-ensemble de B, nous écrirons $A \precsim B$, et dirons que A est *subpotent* à B ou que B *domine* A. Cela revient à dire qu'il existe une injection $\iota : A \hookrightarrow B$.

L'utilisation d'un symbolisme rappelant les relations d'ordre n'est pas fortuite. Cela suggère que le concept représenté jouit de certaines propriétés. Voyons voir ce qu'il en est :

- *Réflexivité* : Tout ensemble A est équipotent à un de ses sous-ensembles, à savoir l'ensemble A lui-même. Il s'ensuit que $A \precsim A$;
- *Transitivité* : Si f est une bijection entre A et un sous-ensemble de B et si g est une bijection entre B et un sous-ensemble de C, alors la composition de f et de la restriction de g à l'image de f est une bijection entre A et un sous-ensemble de C. En d'autres termes, si $A \precsim B$ et $B \precsim C$, alors $A \precsim C$;
- *Antisymétrie* : Il est absurde d'espérer que de $A \precsim B$ et $B \precsim A$ résulte $A = B$. Deux ensembles équipotents ne sont pas forcément identiques. Il est toutefois possible d'énoncer un résultat plus faible, mais tout de même capital.

Théorème 3.1. (Cantor–Bernstein, 1896) Si $A \precsim B$ et $B \precsim A$, alors $A \sim B$.

Démonstration. Soit f une application injective de A dans B et g une application injective de B dans A. Le problème consiste à construire une bijection entre A et B.

Nous dirons provisoirement de $a \in A$ qu'il est le *parent* de $f(a) \in B$ et de ce dernier qu'il est le *descendant* de a. De même $b \in B$ est le parent de $g(b) \in B$ et ce dernier est *descendant* de b. Chaque élément $a \in A$ a une suite infinie de descendants

$$f(a), g(f(a)), f(g(f(a))), ...$$

De façon analogue, on peut dresser la liste des termes dont un élément B est l'*ancêtre*. Pour chaque élément de A ou de B, trois cas peuvent survenir :

1. En remontant les ancêtres d'un élément, on arrive finalement à un élément de A qui est orphelin (un élément de $A \backslash g(B)$) ;
2. En remontant les ancêtres d'un élément, on arrive finalement à un élément de B qui est orphelin (un élément de $B \backslash f(A)$) ;
3. La lignée remonte *ad infinitum*.

Partitionnons A en trois sous-ensembles, notés A_A, A_B et A_∞, composés des éléments qui ont respectivement leur origine dans A, leur origine dans B, une généalogie infinie. Procédons de même pour former les sous-ensembles B_A, B_B et B_∞ de B. Observons que :

- Si $a \in A_A$, alors $f(a) \in B_A$ à la suite de quoi la restriction de f à A_A est une bijection entre A_A et B_A ;
- Si $a \in A_B$, alors a appartient au domaine de la fonction inverse g^{-1} et $g^{-1}(a) \in B_B$. Il s'ensuit que la restriction de g^{-1} à A_B est une bijection entre A_B et B_B ;
- Si $a \in A_\infty$, alors $f(a) \in B_\infty$ et la restriction de f à A_∞ est une bijection entre A_∞ et B_∞.

La combinaison de ces trois bijections forme une correspondance biunivoque entre A et B.

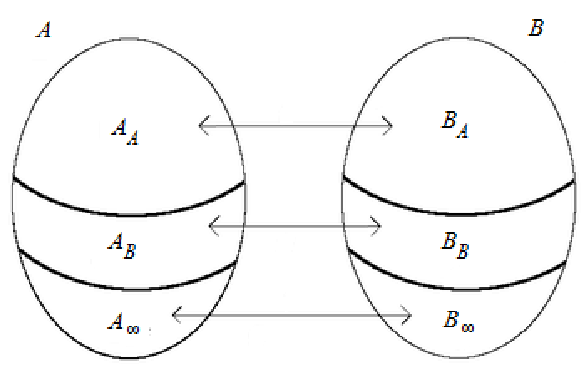

FIGURE 3.1 – Représentation des correspondances biunivoques induites par les deux injections. Source : adaptée de http ://commons.wikimedia.org/wiki/File :Cantor2.png

■

Ce théorème, aussi appelé bien injustement *théorème de Schröder–Bernstein*, fut d'abord énoncé (17) et démontré (19) par Georg Cantor. Toutefois, sa démonstration reposait indûment sur l'axiome du choix. Felix Bernstein, un étudiant de Cantor, présenta une démonstration plus convenable au cours d'un séminaire à Halle au printemps 1897, mais il fallut attendre plus d'un an pour que celle-ci soit publiée (6). Entretemps, Ernst Schröder, professeur à la Polytechnische Schule de Karlsruhe avait lui aussi fait paraître une démonstration sans choix (52). En 1911, Alwin Reinhold Korselt (37) porta à l'attention de Schröder que sa démonstration comportait une faille. Ce dernier avoua avoir lui aussi identifié la brèche depuis un moment déjà, mais il s'était bien gardé d'en faire part à quiconque (9).

L'équipotence, la subpotence ainsi que le théorème de Cantor–Bernstein peuvent être utilisé pour définir formellement la finitude et l'infinitude ainsi que pour exprimer élégamment quelques-uns des faits élémentaires sur les ensembles finis et infinis.

Définition 3.2. Un ensemble A est dit :

1. *Fini* s'il est équipotent à un certain nombre naturel. Autrement, A est dit *infini* ;
2. *Au plus dénombrable* si $A \precsim \omega$;
3. *Dénombrable* si $A \sim \omega$;
4. *Indénombrable* si $A \not\precsim \omega$.

Proposition 3.2. Étant donné des ensembles A et B,

1. Si $A \precsim B$ et si B est fini, alors A est fini ;
2. L'ensemble ω est infini ;
3. L'ensemble A est fini si et seulement si $A \prec \omega$;
4. Si A est dénombrable, alors tout sous-ensemble infini de A est dénombrable ;
5. Si A et B sont dénombrables, alors $A \cup B$ est dénombrable ;
6. Si A et B sont dénombrables, alors $A \times B$ est dénombrable.

La démonstration de ces propriétés étant élémentaire et guère éclairante, elle sera omise.

3.2 Nombres de Hartogs et alephs

Nous avons acquis une certaine connaissance factuelle de la théorie des nombres cardinaux sans même indiquer de façon précise la signification de ce qu'est un *nombre cardinal* ou de ce qu'est la *cardinalité d'un ensemble* A. Toutefois, tant au plan conceptuel que du point de vue de la notation, il y a des avantages à définir $\mathrm{card}\,(A)$, « le nombre d'éléments de A », comme un objet concret. Dans cette section et dans celle qui suit, nous ferons en sorte de démontrer - à l'aide de l'axiome du choix - l'hypothèse suivante :

Il existe des ensembles appelés nombres cardinaux ayant la propriété que pour tout ensemble A il existe un unique nombre cardinal $card(A)$ - appelé cardinalité de A - tel que :

1. *Les ensembles A et B sont équipotents si et seulement si $card(A) = card(B)$;*
2. *L'ensemble A est strictement subpotent à B si et seulement si $card(A) < card(B)$.*

Pour ce faire, nous allons démontrer l'existence d'un unique représentant pour chaque classe d'ensembles mutuellement équipotents.

Cela va de soi qu'afin de n'avoir aucun élément un ensemble A quelconque doit être l'ensemble vide. La propriété d'unicité de $\emptyset$ lui confère le rôle de nombre cardinal. Poursuivant en ce sens, suggérons - conjecture raisonnable - la candidature de l'ensemble n comme nombre cardinal des ensembles à n éléments.

Avant d'avoir exposé les ensembles infinis à des investigations, il ne serait pas déraisonnable de conjecturer que tous les ensembles infinis sont dénombrables. Or, nous allons montrer qu'il n'en est pas ainsi ; admettre l'axiome de l'ensemble des parties élargit incommensurablement le champ des possibles. C'est d'ailleurs ce résultat négatif qui rend la théorie des nombres cardinaux intéressante.

Théorème 3.3. (Cantor, 1891 (18)) Tout ensemble est strictement subpotent à son ensemble des parties. En d'autres termes,

$$A \prec \mathscr{P}(A).$$

Démonstration. Il existe une injection naturelle de A dans $\mathscr{P}(A)$, à savoir l'application qui associe à chaque élément a de A le singleton $\{a\}$. L'existence de cette application prouve que $A \precsim \mathscr{P}(A)$.

Il reste à montrer que A n'est pas équipotent à $\mathscr{P}(A)$. Supposons qu'il existe une application surjective $f : A \to \mathscr{P}(A)$. Considérons l'ensemble

$$B := \left\{a \in A : \ a \notin f(a)\right\}.$$

Il va sans dire que B est un élément de $\mathscr{P}(A)$. Par la surjectivité de l'application f, il existe un élément b dans A tel que $f(b) = B$. Or,

- Si $b \in B$ alors, par définition de B, nous devons avoir $b \notin B$ ce qui est impossible ;
- Si $b \notin B$ alors, toujours par la définition de B, nous devons avoir $b \in B$ ce qui est également impossible.

La contradiction mise en évidence réfute l'existence d'une fonction surjective f de A dans $\mathscr{P}(A)$.

■

Remarque. Il est d'usage de désigner la cardinalité des nombres réels par $\mathfrak{c}$. Ce choix de symbôle s'explique par le fait que l'ensemble des nombres réels est parfois qualifié de *continu* ou de *continuum*. Soulignons au passage qu'il est possible de construire une bijection entre les ensembles $\mathbb{R}$ et $\mathscr{P}(\omega)$. À ce sujet, voir par exemple (28) p. 54. Sur la base de cette affirmation, il résulte que l'ensemble $\mathbb{R}$ se présente à nous comme étant un premier cas concret d'ensemble indénombrable.

À la lumière du théorème de Cantor, l'ensemble des parties de tout ensemble A - fini ou infini - contient toujours plus d'éléments que A lui-même. Par formation successive de l'ensemble des parties, il est possible de former une suite d'ensembles infinis de plus en plus grand au sens de la cardinalité

$$A \prec \mathscr{P}(A) \prec \mathscr{P}\big(\mathscr{P}(A)\big) \prec \ldots$$

En particulier, s'il existait un ensemble C composé de tous les nombres cardinaux, alors $\mathscr{P}(C)$ serait un ensemble strictement plus grand. Ce paradoxe, communément appelé *paradoxe du plus grand cardinal*, est évoqué pour la première fois par Georg Cantor dans une lettre à David Hilbert en 1897. Il met en lumière le fait que la classe des nombres cardinaux est une classe propre que nous désignerons par Card.

Désormais assurés que le fini et le dénombrable n'offrent pas une description exhaustive de la classe Card, tâchons d'identifier d'autres nombres cardinaux. Suivant l'idée de Von Neumann (67), il peut être avantageux de recruter nos ensembles représentatifs parmi les nombres ordinaux. Cette idée est motivée par le fait que dans la théorie ZFC, tout ensemble peut être muni d'une structure de bon ordre et, par voie de conséquence, être mis en bijection avec un certain nombre ordinal. En présence de l'axiome du choix, la théorie des nombres cardinaux se résume donc à étudier la cardinalité des ensembles bien ordonnés.

Nous reviendrons ultérieurement sur ce qui se produit en l'absence de l'axiome du choix. Pour l'heure, notre tâche se limite à identifier un critère permettant de trouver quel nombre ordinal a la même cardinalité qu'un ensemble A donné. Ce faisant, on rencontre une difficulté : un ensemble peut être équipotent à plus d'un nombre ordinal. On entrevoit néanmoins une lueur d'espoir :

Proposition 3.4. Pour chaque ensemble A, les nombres ordinaux équipotents à A constituent un ensemble.

Démonstration. Soit γ un nombre ordinal équipotent à $\mathscr{P}(A)$ et α un nombre ordinal équipotent à A. Alors, par le théorème de Cantor, l'ensemble α est dominé strictement par l'ensemble γ, c'est-à-dire $\alpha \prec \gamma$. Il en résulte que la possibilité d'avoir un isomorphisme d'ordre entre γ et un sous-ensemble de α doit être exclue puisqu'il n'existe même pas de bijections entre ces deux ensembles. On a donc impérativement $\alpha < \gamma$. Ainsi, γ est un nombre ordinal strictement plus grand que tous les nombres ordinaux équipotents à A.

Puisque pour les nombres ordinaux $\alpha < \gamma$ signifie $\alpha \in \gamma$, nous avons trouvé un ensemble, à savoir γ, qui contient tous les nombres ordinaux équipotents à A. Ceci implique qu'on peut former par l'axiome de compréhension sur γ l'ensemble composé des nombres ordinaux équipotents à A et de ceux-ci seulement.

■

Étant donné un ensemble A, lequel parmi les nombres ordinaux qui lui sont équipotents mérite d'être singularisé et appelé un *nombre cardinal* ? Considérant que la morphologie des ensembles bien ordonnés laisse clairement entrevoir un élément saillant - le plus petit - la réponse s'impose d'elle-même.

Définition 3.3.

1. Un nombre ordinal α est appelé *ordinal initial* si, pour tout ordinal β tel que $\beta < \alpha$, on a $\beta \prec \alpha$;
2. Étant donné un ensemble bien ordonné A, la *cardinalité de A* ou le *nombre cardinal* de A est l'unique ordinal initial équipotent à A.

Comme nous l'avons présumé, les propriétés des nombres naturels leur octroient le statut de nombre ordinal initial, c'est-à-dire de représentant canonique pour les classes des ensembles finis mutuellement équipotents. À ce titre, ils constituent des nombres cardinaux. De plus, comme l'ensemble ω - à savoir l'ensemble bien ordonné infini dont chaque section commençante est finie - est le plus petit nombre ordinal infini, il satisfait à la propriété de minimalité caractérisant les nombres ordinaux initiaux. Par contre, aucun des autres nombres ordinaux construits à la section 2.1 n'est un ordinal initial puisqu'ils sont tous dénombrables. Par conséquent, nous manquons cruellement de nombres ordinaux pour refléter la connaissance acquise en démontrant le théorème de Cantor ; il est impératif d'en construire de nouveau.

Définition 3.4. Pour tout ensemble A - bien ordonnable ou non - on appel *ordinal de Hartogs de A* le plus petit nombre ordinal qui n'est équipotent à aucun sous-ensemble de A et on le désigne par $h(A)$.

Par définition, $h(A)$ est le plus petit nombre ordinal α qu'il n'est pas possible d'injecter dans A, c'est-à-dire le plus petit $\alpha \in \mathrm{Ord}$ tel que $\alpha \not\precsim A$.

Proposition 3.5. Pour tout ensemble A, l'ordinal de Hartogs $h(A)$ est un ordinal initial.

Démonstration. Supposons que $\beta \sim h(A)$ pour un certain ordinal $\beta < h(A)$. Alors, β est similaire - et en particulier équipotent - à un sous-ensemble strict de $h(A)$. Comme la relation d'équipotence est transitive, il s'ensuit que $h(A)$ est équipotent à un sous-ensemble strict de $h(A)$. Cela signifie que $h(A) < h(A)$, ce qui est impossible.

■

Jusqu'ici, nous avons évité la principale difficulté qui consiste à montrer l'existence de l'ordinal de Hartogs de A. Cette question fera l'objet du prochain théorème.

Théorème 3.6.

1. (Hartogs, 1915 (31)) Pour tout ensemble A, l'ordinal de Hartogs de A existe ;
2. (Sierpiński, 1947 (58)) On a $h(A) \precsim \mathscr{P}\big(\mathscr{P}(\mathscr{P}(A))\big)$.

Démonstration. On désire montrer qu'il existe un nombre ordinal α qui ne peut être injecté dans A. Pour ce faire, nous devons procéder par étapes :

1. Soit $\mathcal{C}$ une collection de sous-ensembles de A, autrement dit soit $\mathcal{C} \subseteq \mathscr{P}(A)$. Il peut survenir que les éléments de $\mathcal{C}$ soient bien ordonnés par inclusion. Par exemple, il tombe sous le sens que la collection $\mathcal{C} := \{\emptyset\}$ soit bien ordonnée par inclusion puisqu'elle ne comporte qu'un seul sous-ensemble de A. Notons par $\mathscr{B}$ l'ensemble de toutes les collections $\mathcal{C}$ bien ordonnées par inclusion et dont les éléments sont des sous-ensembles de A. Ainsi défini, il appert que $\mathscr{B} \subseteq \mathscr{P}\big(\mathscr{P}(A)\big)$;
2. Considérons l'ensemble de toutes les classes d'équivalences de $\mathscr{B}$ par la relation de similarité, à savoir l'ensemble $\mathscr{B}/\cong$. Chaque élément de $\mathscr{B}/\cong$ étant une classe d'équivalence, c'est-à-dire une collection de sous-ensembles de $\mathscr{B}$, on a
$$\mathscr{B}/\cong \;\subseteq\; \mathscr{P}(\mathscr{B}) \;\subseteq\; \mathscr{P}\big(\mathscr{P}(\mathscr{P}(A))\big);$$
3. Par le théorème 2.12, chacune de ces classes d'équivalences peut être associée à un unique nombre ordinal. En vertu du schéma d'axiome de substitution, l'image par la relation fonctionnelle associant à chaque élément de $\mathscr{B}/\cong$ son type d'ordre constitue un ensemble de nombres ordinaux que nous noterons α. Il va sans dire que α est bien ordonné par la relation d'appartenance et que
$$\alpha \;\precsim\; \mathscr{P}\big(\mathscr{P}(\mathscr{P}(A))\big);$$
4. Nous allons montrer que α est un ensemble transitif. La proposition 2.3 nous permettra alors de déduire que α est lui-même un nombre cardinal. Étant donné $\beta \in \alpha$, alors, par la définition de α, il existe au moins une collection $\mathcal{C}$ constituée de sous-ensembles de A pour laquelle l'inclusion forme une relation de bon ordre de type β. Soit ϕ un isomorphisme d'ordre de β sur $\mathcal{C}$.

   Par la propriété de transitivité des nombres ordinaux, on a que $\zeta \in \beta$ implique $\zeta \subseteq \beta$. L'application $\phi\big|_\zeta$ est donc un isomorphisme d'ordre de ζ à une sous-collection $\mathcal{C}'$ de sous-ensembles de A. Cette sous-collection est sans surprise munie d'un bon ordre - l'inclusion - obtenu par restriction de celui sur $\mathcal{C}$. Ainsi, cette collection $\mathcal{C}'$ est élément de l'ensemble $\mathscr{B}$ et son type d'ordre, c'est-à-dire ζ, appartient à α. Cela achève la démonstration de la transitivité de l'ensemble α ;
5. Montrons maintenant que $\alpha \not\precsim A$. Pour ce faire, nous allons supposer au contraire que α est équipotent à un certain sous-ensemble B de A. De cette correspondance biunivoque

avec le nombre ordinal α, l'ensemble B hérite d'une relation de bon ordre de type α. Notons par $\mathcal{S}$ l'ensemble des sections commençantes de B. Comme $\mathcal{S}$ est un ensemble bien ordonné par inclusion, il s'agit d'un élément de $\mathscr{B}$ et

$$\operatorname{ord}(B) \ = \ \operatorname{ord}(\mathcal{S}) \ = \ \alpha. \tag{3.1}$$

De plus, $\mathcal{S}$ appartient à une certaine classe d'équivalence de $\mathscr{B}/\cong$. C'est alors que vient le coup de grâce : par l'équation (3.1), la classe d'équivalence à laquelle appartient $\mathcal{S}$ est similaire à α. Il s'ensuit que $\alpha \in \alpha$, ce qui est proprement impossible en vertu de l'axiome de fondation. Notre hypothèse étant insoutenable, on doit avoir que $\alpha \not\precsim A$. ∎

Il importe de remarquer que le théorème de Hartogs est un résultat de la théorie de Zermelo–Fraenkel au sens où l'axiome du choix n'intervient en aucun temps dans la démonstration. Cela n'est pas fortuit puisqu'en démontrant ce théorème, le mathématicien allemand Friedrich Hartogs avait comme principale motivation de montrer l'équivalence suivante :

Théorème 3.7. Le *principe de trichotomie cardinale* [1] équivaut à l'axiome du choix.

Démonstration.

$\Rightarrow$) (Hartogs, 1915 (31)) Soit A un ensemble quelconque. Par le théorème de Hartogs, il existe un ordinal $h(A)$ tel que $h(A) \not\precsim A$. Par le principe de trichotomie cardinale, on doit avoir que $A \precsim h(A)$. De cette injection dans $h(A)$, l'ensemble A hérite d'une structure de bon ordre ;

$\Leftarrow$) (Cantor, 1899) En supposant l'axiome du choix, tout ensemble peut être muni d'une relation de bon ordre. Ainsi, tout ensemble est équipotent à un certain nombre ordinal. Le principe de trichotomie des nombres ordinaux assure que ceux-ci sont comparables par inclusion. Conséquemment, les ensembles qui leur sont équipotents sont également comparables (par subpotence).

∎

Si on applique le théorème de Hartogs à l'ensemble $A := \omega$, le nombre ordinal dont l'existence est affirmée, c'est-à-dire le nombre de Hartogs de ω, est le plus petit ordinal indénombrable. Le nombre ordinal possédant cette propriété de minimalité est noté ω_1. De façon équivalente, on peut dire de ω_1 qu'il est l'ensemble de tous les ordinaux au plus dénombrable. En itérant, on définit une chaîne de nombres ordinaux initiaux de plus en plus grands : pour $\alpha \in \mathrm{Ord}$ on pose

$$\omega_0 := \omega;$$
$$\omega_{\alpha+1} := h(\omega_\alpha).$$

1. Aussi appelé *principe de comparabilité cardinale*, le principe de trichotomie cardinale s'énonce comme suit : pour tous nombres cardinaux $\mathfrak{m}$ et $\mathfrak{n}$, soit $\mathfrak{m} < \mathfrak{n}$, soit $\mathfrak{m} = \mathfrak{n}$, soit $\mathfrak{m} > \mathfrak{n}$.

Nous prendrons soin d'introduire une nouvelle notation pour désigner les nombres ordinaux initiaux infinis dans leur rôle de nombres cardinaux. Cette décision s'explique par le fait que les opérations arithmétiques auront des significations différentes selon qu'on considère les nombres sous leur nature ordinale ou cardinale. Plutôt que d'utiliser la lettre grecque ω, nous emploierons $\aleph$ - dit *aleph* -, la première lettre de l'alphabet hébreu.

Définition 3.5. Pour tout $\alpha \in \mathrm{Ord}$, on pose $\aleph_\alpha := \omega_\alpha$.

En définitive, le théorème de Hartogs permet de voir qu'aucun nombre cardinal n'est plus grand - par rapport à la relation de subpotence - que tous les $\aleph$.

De par leur nature ordinale, les $\aleph$ héritent d'une structure d'ordre satisfaisant aux propriétés suivantes :

1. Deux nombres $\aleph$ sont toujours comparables ;
2. Toute collection de nombres $\aleph$ est bien ordonnée ;
3. Toute collection de nombres $\aleph$ admet un suprémum ;

Nous avons pris l'engagement de faire en sorte que notre définition des nombres cardinaux s'harmonise avec la méthode qualitative, au sens où la relation entre la taille de deux ensembles - que ce soit l'équipotence ou la subpotence - doit se refléter directement entre leurs nombres cardinaux respectivement par une égalité ou une inégalité. Le résultat qui suit marque la consécration de nos efforts ; les $\aleph$ se comportent exactement comme nous l'avions anticipé.

Proposition 3.8. Soit A et B des ensembles. Alors,

1. $\mathrm{card}(A) = \mathrm{card}(B)$ si et seulement si $A \sim B$;
2. $\mathrm{card}(A) < \mathrm{card}(B)$ si et seulement si $A \prec B$.

Démonstration.

1. $\Rightarrow$) Chaque ensemble est équipotent à son nombre cardinal. De plus, la relation d'équipotence est transitive ; le résultat s'ensuit,

   $\Leftarrow$) Supposons que $A \sim B$. Puisque $\mathrm{card}(A)$ est le plus petit nombre ordinal équipotent à A, il en résulte que $\mathrm{card}(A) \leq \mathrm{card}(B)$. La situation étant symétrique, l'inégalité inverse est également vérifiée ;

2. $\Rightarrow$) Par la définition de l'ordre sur les ordinaux, on a que $\mathrm{card}(A)$ est un sous-ensemble de $\mathrm{card}(B)$. Il en résulte que $A \precsim B$. Par l'énoncé 1, si nous avions $A \sim B$, nous aurions $\mathrm{card}(A) = \mathrm{card}(B)$. Il s'ensuit que nous devons avoir $A \prec B$;

   $\Leftarrow$) Supposons que $A \prec B$. Alors, il est impossible que $\mathrm{card}(A) \geq \mathrm{card}(B)$ puisque l'isomorphisme d'ordre implique l'équivalence et donc $\mathrm{card}(A) < \mathrm{card}(B)$.

■

3.3 Arithmétique cardinale dans ZFC

Addition
Étant donné des nombres cardinaux $\mathfrak{m}$ et $\mathfrak{n}$ ainsi que des ensembles M et N disjoints tels que $\mathrm{card}\,(M) = \mathfrak{m}$ et $\mathrm{card}\,(N) = \mathfrak{n}$, on définit la *somme* de $\mathfrak{m}$ et $\mathfrak{n}$ comme étant :

$$\mathfrak{m} + \mathfrak{n} := \ \mathrm{card}\,(M \cup N)\,.$$

Cette définition est libre d'ambiguïté puisqu'elle est indépendante du choix des ensembles M et N pour autant qu'ils soient disjoints. Les propriétés de l'union permettent aisément de vérifier que l'addition cardinale est associative et commutative.

Multiplication
Étant donné des nombres cardinaux $\mathfrak{m}$ et $\mathfrak{n}$ ainsi que des ensembles M et N disjoints tels que $\mathrm{card}\,(M) = \mathfrak{m}$ et $\mathrm{card}\,(N) = \mathfrak{n}$, on définit le *produit* de $\mathfrak{m}$ et $\mathfrak{n}$ comme étant :

$$\mathfrak{m} \cdot \mathfrak{n} := \ \mathrm{card}\,(M \times N)\,.$$

Cette fois encore, la définition est exempte d'ambiguïté. On montre sans difficulté - à partir des propriétés de l'union et du produit cartésien - que la multiplication cardinale est commutative, associative et distributive sur l'addition cardinale.

Exponentiation
Pour définir l'exponentiation cardinale, il nous faut utiliser la notion d'exponentiation d'ensembles abordée dans la discussion sur l'axiome de l'ensemble des parties. Étant donné des nombres cardinaux $\mathfrak{m}$ et $\mathfrak{n}$ ainsi que des ensembles M et N tels que $\mathrm{card}\,(M) = \mathfrak{m}$ et $\mathrm{card}\,(N) = \mathfrak{n}$, on définit sans ambiguïté :

$$\mathfrak{m}^{\mathfrak{n}} := \ \mathrm{card}\left(M^N\right).$$

Pour tous nombres cardinaux $\mathfrak{m}, \mathfrak{n}, \mathfrak{p}$ et $\mathfrak{q}$, l'exponentiation cardinale vérifie les identités suivantes :

1. $\mathfrak{m}^{\mathfrak{p}+\mathfrak{q}} = \mathfrak{m}^{\mathfrak{p}}\mathfrak{m}^{\mathfrak{q}}$;
2. $(\mathfrak{m}\mathfrak{n})^{\mathfrak{p}} = \mathfrak{m}^{\mathfrak{p}}\mathfrak{n}^{\mathfrak{p}}$;
3. $\mathfrak{m}^{\mathfrak{p}\mathfrak{q}} = (\mathfrak{m}^{\mathfrak{p}})^{\mathfrak{q}}$.

La démonstration est donnée dans (59) p. 141-143.

Comme conséquence digne d'intérêt de cette définition, observons qu'il est désormais possible de calculer le nombre cardinal de l'ensemble des parties de M : si $\mathrm{card}\,(M) = \mathfrak{m}$, alors en utilisant le fait que $\mathscr{P}(M)$ est équipotent à 2^M, on déduit que $\mathrm{card}\,(\mathscr{P}(M)) \ = \ 2^{\mathfrak{m}}$. On peut reformuler en ces termes les deux théorèmes majeurs obtenus à la section précédente :

Théorème 3.9. (Cantor) Pour tout nombre cardinal $\mathfrak{m}$, on a

$$\mathfrak{m} < 2^{\mathfrak{m}}.$$

Théorème 3.10. (Hartogs–Sierpiński) À chaque nombre cardinal infini $\mathfrak{m}$ est associé un aleph, noté $\aleph(\mathfrak{m})$, satisfaisant aux relations suivantes :

$$\aleph(\mathfrak{m}) \not\leq \mathfrak{m} \qquad \text{et} \qquad \aleph(\mathfrak{m}) \leq 2^{2^{2^{\mathfrak{m}}}}.$$

Propriétés arithmétiques des $\aleph$

Si de nombreuses propriétés de l'arithmétique usuelle sont vérifiées en arithmétique cardinale, les propositions qui suivent montrent que les nombres cardinaux infinis se comportent parfois de manière pour le moins inattendue.

Lemme 3.11. Soit $\mathfrak{m}, \mathfrak{n}$ et $\mathfrak{p}$ des nombres cardinaux. Si $\mathfrak{m} \leq \mathfrak{n}$, alors

$$\mathfrak{m} + \mathfrak{p} \ \leq \ \mathfrak{n} + \mathfrak{p} \qquad \text{et} \qquad \mathfrak{m} \cdot \mathfrak{p} \ \leq \ \mathfrak{n} \cdot \mathfrak{p}.$$

Démonstration. Soit M, N et P des ensembles de cardinalité $\mathfrak{m}, \mathfrak{n}$ et $\mathfrak{p}$ respectivement. Comme $\mathfrak{m} \leq \mathfrak{n}$, on peut choisir M et N tels que $M \subseteq N$ et la proposition s'en trouve automatiquement démontrée, car

$$M \cup P \ \subseteq \ N \cup P \qquad \text{et} \qquad M \times P \ \subseteq \ N \times P.$$

■

Théorème 3.12. Pour tout nombre ordinal α, on a

$$\aleph_\alpha \cdot \aleph_\alpha \ = \ \aleph_\alpha.$$

La démonstration de ce théorème étant fastidieuse et peu instructive, elle est donc omise. Pour les détails, voir (34) p. 134-136.

Corollaire 3.13. Pour tous nombres ordinaux α et β tels que $\alpha \leq \beta$ et tout nombre naturel n, on a :

1. $\aleph_\alpha \cdot \aleph_\beta \ = \ \aleph_\beta$;
2. $n \cdot \aleph_\alpha \ = \ \aleph_\alpha$.

Démonstration. D'une part, $\alpha \leq \beta$ implique que

$$\aleph_\beta \ = \ 1 \cdot \aleph_\beta \ \leq \ \aleph_\alpha \cdot \aleph_\beta.$$

D'autre part, le théorème 3.12 nous donne que

$$\aleph_\alpha \cdot \aleph_\beta \ \le \ \aleph_\beta \cdot \aleph_\beta \ = \ \aleph_\beta.$$

Par le théorème de Cantor–Bernstein, il s'ensuit que $\aleph_\alpha \cdot \aleph_\beta \ = \ \aleph_\beta$. La seconde égalité se démontre de façon similaire.

■

Corollaire 3.14. Pour tous nombres ordinaux α et β tels que $\alpha \le \beta$, on a :

1. $\aleph_\alpha + \aleph_\beta \ = \ \aleph_\beta$, et conséquemment $\aleph_\alpha + \aleph_\beta = \max\{\aleph_\alpha, \aleph_\beta\}$;
2. $n + \aleph_\alpha \ = \ \aleph_\alpha$.

Démonstration. Si $\alpha \le \beta$, alors

$$\aleph_\beta \ \le \ \aleph_\alpha + \aleph_\beta \ \le \ \aleph_\beta + \aleph_\beta \ = \ 2 \cdot \aleph_\beta \ = \ \aleph_\beta.$$

Le deuxième énoncé se démontre de manière similaire.

■

3.4 Hiérarchie cumulative et astuce de Scott

Dans cette section, nous étudierons la théorie des cardinaux dans le cadre axiomatique de Zermelo–Fraenkel, c'est-à-dire en l'absence de l'axiome du choix. Nous verrons que la définition des nombres cardinaux s'en trouve grandement complexifiée.

Supposons un instant que la théorie des ensembles soit l'univers idyllique imaginé par Cantor où les ensembles peuvent être définis de deux façons, à savoir :

1. En extension, soit en nommant chaque élément qui en fait partie ;
2. En compréhension, soit par une propriété spécifique commune à tous les éléments.

On pourrait notamment former par compréhension l'ensemble P de toutes les paires $\{a, b\}$ avec $a \neq b$. Puis, nous pourrions définir le nombre naturel 2 comme étant l'ensemble P. En d'autres termes, nous pourrions assimiler la propriété « avoir deux éléments » à l'ensemble P de tous les objets en jouissent. De la même façon, chaque classe d'équivalence par rapport à la relation d'équipotence constituerait un nombre cardinal.

Sitôt qu'on cesse de se bercer d'illusions concernant la possibilité de définir aveuglément des ensembles en compréhension, il nous faut rabattre nos prétentions et tenir compte du fait que la relation d'équipotence partitionne le domaine du discours de la théorie des ensembles en classes propres plutôt qu'en ensembles ; pour s'en convaincre, il suffit d'observer que l'union des éléments de l'ensemble P formé de toutes les paires imaginables d'ensembles distincts n'est rien autre que l'ensemble universel $\mathscr{U}$ dont l'existence est chimérique en vertu du paradoxe de Russell. C'est précisément pour remédier à cette situation regrettable que, lorsque nous pouvions compter sur l'axiome du choix, nous avons assimilé les nombres cardinaux à un représentant canonique de classe d'équipotence - l'ordinal initial - plutôt qu'à la classe d'équivalence elle-même. Cependant, en l'absence de l'axiome du choix, cette méthode ne résout plus complètement la question puisque nous ne sommes plus en mesure de garantir que tout ensemble puisse être muni d'une relation de bon ordre. Nous sommes alors dans l'obligation de choisir entre deux avenues possibles : désirons-nous que la cardinalité soit un indicateur de la taille des seuls ensembles bien ordonnables, ou désirons-nous plutôt qu'elle soit une mesure de la taille de tous les ensembles ? Les deux options sont viables.

Si on choisit de considérer la cardinalité comme une anomalie des ensembles bien ordonnables, alors il y a un lourd prix à payer :

- Il n'est plus possible d'attribuer un nombre cardinal à l'ensemble $\mathbb{R}$;
- L'exponentiation cardinale n'est plus définie. Il n'y a en effet aucune raison de croire que l'ensemble des fonctions d'un ensemble A dans un ensemble B est bien ordonnable ;
- Le théorème de Cantor - dans lequel intervient l'ensemble des parties d'ensembles arbitraires - devient caduc.

Ces fâcheuses conséquences nous poussent à explorer la seconde voie qui consiste à développer une notion de cardinalité qui ne dépend pas de l'aptitude d'un ensemble à être bien ordonné. On fait cependant face à certaines difficultés majeures : il est cohérent avec ZF que certaines classes d'ensembles mutuellement équipotents n'admettent aucun représentant canonique (35). Autrement dit, il n'existe pas de classe fonctionnelle F définie sur le domaine du discours de ZF attribuant à chaque ensemble son représentant de classe d'équivalence de telle sorte que pour tout ensemble X on ait simultanément :

1. $F(X) = F(Y)$ si et seulement si $\text{card}\,(X) = \text{card}\,(Y)$;
2. $\text{card}\,(F(X)) = \text{card}\,(X)$.

Le mathématicien américain Dana Scott a développé une méthode pragmatique - que nous appellerons *astuce de Scott*[2] - permettant de donner un sens à la cardinalité d'un ensemble A quelconque sans recourir à l'axiome du choix. Celle-ci tranche avec ce que nous avons vu jusqu'ici puisqu'il n'est pas question de représentant de classe d'équivalence. L'astuce de Scott s'inspire plutôt de la théorie naïve des ensembles de Cantor. En revanche, cette méthode utilise abondamment l'axiome de fondation. Le principe qui sous-tend l'astuce de Scott est de transformer les classes d'équivalence en des ensembles plutôt compliqués, mais qui ont l'avantage majeur de permettre de travailler aisément avec l'ordre partiel sur les cardinalités. L'astuce de Scott met à profit le principe de conception itérative décrit superficiellement au chapitre 1 et dont voici une définition formelle et détaillée :

2. Traduction libre de *Scott's trick*.

1. Soit $\mathcal{V}_0$ est l'ensemble vide :
$$\mathcal{V}_0 := \emptyset;$$
2. Pour tout ordinal limite α, soit $\mathcal{V}_\alpha$ l'union de tous les $\mathcal{V}_\zeta$ indicés par un nombre ordinal $\zeta < \alpha$:
$$\mathcal{V}_\alpha := \bigcup_{\zeta<\alpha} \mathcal{V}_\zeta;$$
3. Pour tout nombre ordinal α, soit $\mathcal{V}_{\alpha+1}$ l'ensemble des parties de $\mathcal{V}_\alpha$:
$$\mathcal{V}_{\alpha+1} := \mathscr{P}(\mathcal{V}_\alpha);$$
4. La classe $\mathscr{V}$ - appelée *hiérarchie cumulative des ensembles* ou *univers de Von Neumann* - est définie comme étant l'union de tous les $\mathcal{V}_\alpha$ avec $\alpha \in \mathrm{Ord}$:
$$\mathscr{V} := \bigcup \mathcal{V}_\alpha.$$

Notons que par construction, $\mathcal{V}_\alpha$ est un ensemble et ce, pour tout choix de $\alpha \in \mathrm{Ord}$. En outre, on montre sans difficulté que :

1. Pour chaque $\alpha \in \mathrm{Ord}$, l'ensemble $\mathcal{V}_\alpha$ est transitif ;
2. Si $\alpha < \beta$, alors $\mathcal{V}_\alpha \subset \mathcal{V}_\beta$;
3. Pour chaque $\alpha \in \mathrm{Ord}$, on a $\alpha \subseteq \mathcal{V}_\alpha$ et $\alpha \in \mathcal{V}_{\alpha+1}$.

Rappelons qu'un ensemble est dit *transitif* si tous ses éléments en sont également des sous-ensembles. Autrement dit, un ensemble est transitif si et seulement si $\bigcup A \subseteq A$ où $\bigcup A$ désigne l'union de tous les éléments de A. Étant donné un ensemble A qui n'est pas forcément transitif, on appel *clôture transitive de A* le plus petit ensemble transitif contenant A et on le note par $\mathrm{CT}(A)$. Pour l'obtenir, on définit inductivement pour $n \in \omega$ les ensembles :

$$\begin{aligned} A_0 &:= A, \\ A_{n+1} &:= \bigcup A_n. \end{aligned}$$

Enfin, on pose

$$\mathrm{CT}(A) := \bigcup_{n\in\omega} A_n = \bigcup \Big\{ \bigcup A, \bigcup\bigcup A, \bigcup\bigcup\bigcup A, ... \Big\}.$$

On a par exemple que $x_{n+1} \in A_{n+1}$ si et seulement s'il existe $x_0 \in A_0, x_1 \in A_1, ..., x_n \in A_n$ tels que $x_{n+1} \in ... \in x_1 \in x_0$. Rappelons qu'en raison de l'axiome de fondation, de telles chaînes décroissantes doivent toutesêtre finies. C'est pourquoi nous sommes en mesure d'affirmer que $\mathrm{CT}(A)$ possède les propriétés que nous lui prêtons.

Théorème 3.15. Pour tout ensemble A du domaine du discours de la théorie des ensembles ZFC, il existe un nombre ordinal α tel que $A \in \mathcal{V}_\alpha$.

Démonstration. En vue d'aboutir à une contradiction, supposons qu'il existe un ensemble A qui n'appartient pas à $\mathscr{V}$. Posons $W := \big\{x \in \mathrm{CT}\big(\{A\}\big) :\ x \notin \mathscr{V}\big\}$. Dit autrement, W est composé des éléments x de la clôture transitive de $\{A\}$ pour lesquels il n'existe pas de rang $\alpha \in \mathrm{Ord}$ tels que $x \in \mathcal{V}_\alpha$. Notons que W est non vide puisqu'il contient A.

Par l'axiome de fondation, il existe $w \in W$ tel que $W \cap w = \emptyset$. Comme $w \in W$, on a $w \notin \mathscr{V}$, ce qui implique en particulier que $w \neq \emptyset$. Or, pour tout $x \in w$, il doit exister un nombre ordinal minimal $\alpha(x)$ tel que $x \in \mathcal{V}_{\alpha(x)}$.

Par le schéma d'axiome de substitution, la collection $\big\{\alpha(x) \in \mathrm{Ord} :\ x \in w\big\}$ est un ensemble et $\alpha := \bigcup\limits_{x \in w_0} \alpha(x)$ est un nombre ordinal. Cela implique que $w \subseteq \mathcal{V}_\alpha$ et, par voie de conséquence, $w \in \mathcal{V}_{\alpha+1}$ contredisant le fait que $w \notin \mathscr{V}$.

■

Bien qu'en vertu de l'axiome de fondation, la classe $\mathscr{V}$ coïncide avec le domaine du discours de ZFC, on évite généralement de décrire $\mathscr{V}$ comme étant la « classe de tous les ensembles » puisque cette dernière formulation dépend du cadre axiomatique. Il existe en effet des théories des ensembles alternatives dans lesquelles l'axiome de fondation est remplacé par un axiome dit d'*anti-fondation*. Or, dans ces théories, la « classe de tous les ensembles » ne correspond pas $\mathscr{V}$.

Astuce de Scott

Scott a déduit du théorème 3.15 qu'en supposant l'axiome de fondation, il existe pour tout ensemble A au moins un rang dans la hiérarchie cumulative des ensembles pour lequel un ensemble ayant la même cardinalité que A surgit. Il suffit alors de définir la *cardinalité de A* comme étant l'ensemble :

$$\mathrm{card}\,(A) := \big\{B \in \mathcal{V}_{\beta_0} :\ \text{il existe une bijection entre } B \text{ et } A\big\},$$

où β_0 est le plus petit nombre ordinal tel que $\mathcal{V}_{\beta_0}$ contienne un ensemble B ayant la cardinalité A.

L'ensemble $\mathrm{card}\,(A)$ est appelé *nombre cardinal*. Il est à noter que A n'appartient pas nécessairement à l'ensemble $\mathrm{card}\,(A)$. Remarquons également que, conformément à notre idée intuitive, on a bien

1. $\mathrm{card}\,(A) = \mathrm{card}\,(B)$ si et seulement si $A \sim B$;
2. $\mathrm{card}\,(A) < \mathrm{card}\,(B)$ si et seulement s'il existe $B' \subseteq B$ tel que $\mathrm{card}\,(A) = \mathrm{card}\,(B')$. Autrement dit, $\mathrm{card}\,(A) < \mathrm{card}\,(B)$ si et seulement si $A \prec B$.

Bien que l'astuce de Scott s'applique universellement à tous les ensembles, il est d'usage d'adopter une définition qui est le produit d'un syncrétisme avec la définition tablant sur les nombres ordinaux initiaux. Cette approche hybride se démarque en alliant la souplesse de la méthode de Scott à l'esthétisme des $\aleph$ dont les définitions sont intégralement conservées.

3.5 Catastrophe en arithmétique cardinale dans ZF

Au cours de cette section, nous verrons que de nombreux résultats concernant l'ordre sur les nombres cardinaux ainsi que la plupart des résultats d'arithmétique cardinale obtenus à la section 3.3 échouent lamentablement lorsqu'on tente de les généraliser aux nombres cardinaux définis par l'astuce de Scott en l'absence de l'axiome du choix.

Théorème 3.16. Il résulte de la négation de l'axiome du choix qu'il existe des ensembles incomparables par la relation $\precsim$.

Démonstration. Soit A un ensemble qui n'est pas bien ordonnable et soit $h(A)$ son nombre de Hartogs. Alors, par définition, $h(A)$ ne peut être injecté dans A. De plus, il n'est pas possible d'injecter A dans le nombre ordinal $h(A)$ sans quoi A serait bien ordonnable. En somme, on a $h(A) \not\precsim A$ et $A \not\precsim h(A)$.

∎

Corollaire 3.17. Il découle de la négation de l'axiome du choix qu'il existe des nombres cardinaux incomparables.

C'est en prévision du théorème 3.16 que nous avions pris le soin d'éviter de définir l'indénombrabilité comme étant la propriété d'avoir une cardinalité strictement plus grande que $\aleph_0$.

Définition 3.6. Un ensemble A de cardinalité $\mathfrak{m}$ est dit :

1. *Dedekind-infini* si $\aleph_0 \leq \mathfrak{m}$ ou, de façon équivalente, si A est équipotent à l'un de ses sous-ensembles propres ;
2. *Dedekind-fini* s'il n'est pas Dedekind-infini.

Dans le cadre de la théorie ZF, il est plausible qu'un ensemble soit à la fois infini et Dedekind-fini. Ce serait notamment le cas de tout ensemble infini non comparable avec $\aleph_0$. C'est dire que lorsqu'on nous prive de l'axiome du choix, la frontière entre le fini et l'infini est passablement nébuleuse. Nous verrons que cela a de bien fâcheuses conséquences lorsque vient le temps d'énoncer l'hypothèse du continu.

La somme $\mathfrak{m} + \mathfrak{n}$, le produit $\mathfrak{m} \cdot \mathfrak{n}$ et l'exponentiation $\mathfrak{m}^{\mathfrak{n}}$ sont définis dans ZF de la même manière qu'ils le sont dans ZFC et il en est de même pour les sommes finies $\sum\limits_{i=0}^{n} \mathfrak{m}_i$ et les produits finis $\prod\limits_{i=0}^{n} \mathfrak{m}_i$. En revanche, étant donné une famille quelconque de cardinaux $\{\mathfrak{m}_i\}_{i \in I}$, le théorème suivant montre qu'en l'absence de l'axiome du choix la somme $\sum\limits_{i \in I} \mathfrak{m}_i$ et le produit $\prod\limits_{i \in I} \mathfrak{m}_i$ ne riment à rien .

Théorème 3.18. Dans certains modèles ZF il existe des suites $\{M_i\}_{i\in\omega}$ et $\{N_i\}_{i\in\omega}$ d'ensembles tels que $\operatorname{card}(M_i) = \operatorname{card}(N_i)$ pour tout $i \in \omega$ mais pour lesquelles

$$\operatorname{card}\left(\bigcup_{i\in\omega} M_i\right) \neq \operatorname{card}\left(\bigcup_{i\in\omega} N_i\right) \qquad \text{et} \qquad \operatorname{card}\left(\prod_{i\in\omega} M_i\right) \neq \operatorname{card}\left(\prod_{i\in\omega} N_i\right).$$

Démonstration. Posons $M_i := \{0_i, 1_i\}$ pour tout $i \in \mathbb{N}$. Il est compatible avec ZF qu'il existe une suite $\{N_i\}_{i\in\omega}$ d'ensembles à deux éléments dont le produit cartésien $\prod_{i\in\omega} N_i$ est vide. Alors,

1. $\operatorname{card}(M_i) = 2 = \operatorname{card}(N_i)$ quel que soit $i \in \omega$;
2. $\operatorname{card}\left(\prod_{i\in\omega} M_i\right) = 2^{\aleph_0} \neq 0 = \operatorname{card}\left(\prod_{i\in\omega} N_i\right)$.
3. $\operatorname{card}\left(\bigcup_{i\in\omega} M_i\right) = \aleph_0 \neq \operatorname{card}\left(\bigcup_{i\in\omega} N_i\right)$.

■

Poursuivant les travaux de Bernstein et de Sierpiński, le logicien et philosophe polonais Alfred Tarski montra en 1924 (63) que dans ZF, chacun des sept résultats d'arithmétique cardinale suivants est équivalent à l'axiome du choix.

1. $\mathfrak{m} \cdot \mathfrak{n} = \mathfrak{m} + \mathfrak{n}$ pour tous nombres cardinaux $\mathfrak{m}$ et $\mathfrak{n}$ infinis ;
2. $\mathfrak{m} = \mathfrak{m} \cdot \mathfrak{m}$ pour tout nombre cardinal $\mathfrak{m}$ infini ;
3. Si $\mathfrak{m}^2 = \mathfrak{n}^2$, alors $\mathfrak{m} = \mathfrak{n}$;
4. Monotonicité de l'addition cardinale : si $\mathfrak{m} < \mathfrak{n}$ et $\mathfrak{p} < \mathfrak{q}$, alors $\mathfrak{m} + \mathfrak{p} < \mathfrak{n} + \mathfrak{q}$;
5. Monotonicité de la multiplication cardinale : si $\mathfrak{m} < \mathfrak{n}$ et $\mathfrak{p} < \mathfrak{q}$, alors $\mathfrak{m} \cdot \mathfrak{p} < \mathfrak{n} \cdot \mathfrak{q}$;
6. Loi de simplification additive : si $\mathfrak{m} + \mathfrak{p} < \mathfrak{n} + \mathfrak{p}$, alors $\mathfrak{m} < \mathfrak{n}$;
7. Loi de simplification multiplicative : si $\mathfrak{m} \cdot \mathfrak{p} < \mathfrak{n} \cdot \mathfrak{p}$, alors $\mathfrak{m} < \mathfrak{n}$.

Que chacun de ces énoncés soient équivalent à l'axiome du choix ne signifie pas qu'il faille tout rejeter en bloc lorsqu'on choisi de se priver de AC. En effet, faisant suite aux travaux de Lorenz J. Halbeisen et Saharon Shelah (29), les mathématiciens thaïlandais Sapukun Panasawatwong et Pimpen Vejjajiva (48) sont parvenus à montrer que certains de ces énoncés sont vérifiés moyennant quelques conditions à imposer sur les nombres cardinaux :

1. Pour tout nombre cardinal Dedekind-fini $\mathfrak{p} > 1$, si les nombres naturels $\mathfrak{m}$ et $\mathfrak{n}$ sont tels que $\mathfrak{m} < \mathfrak{n}$, alors $\mathfrak{p}^m < \mathfrak{p}^n$;
2. Pour tout nombre cardinal Dedekind-fini $\mathfrak{p} > 1$, si $\mathfrak{p}^m < \mathfrak{p}^n$ où $\mathfrak{m}$ et $\mathfrak{n}$ sont des nombres naturels, alors $\mathfrak{m} = \mathfrak{n}$;

3. Monotonicité conditionnelle de l'addition et de la multiplication cardinale : pour tout nombre cardinal Dedekind-fini $\mathfrak{p}$, si $\mathfrak{m} < \mathfrak{n}$ où $\mathfrak{m}$ et $\mathfrak{n}$ sont soit des nombres naturel soit des $\aleph$, alors $\mathfrak{m} + \mathfrak{p} < \mathfrak{n} + \mathfrak{p}$ et $\mathfrak{m} \cdot \mathfrak{p} < \mathfrak{n} \cdot \mathfrak{p}$ pour autant que $\mathfrak{p} \neq 0$;
4. Lois de simplification additive et multiplicative partielles : pour tout nombre cardinal $\mathfrak{n}$ ainsi que tous nombres cardinaux Dedekind-finis $\mathfrak{m}$ et $\mathfrak{p}$, si $\mathfrak{m} < \mathfrak{n}$, alors $\mathfrak{m} + \mathfrak{p} < \mathfrak{n} + \mathfrak{p}$ et $\mathfrak{m} \cdot \mathfrak{p} < \mathfrak{n} \cdot \mathfrak{p}$ pour autant que $\mathfrak{p} \neq 0$.

3.6 Hypothèse du continu

Étant donné un nombre cardinal $\mathfrak{m}$, le théorème de Hartogs affirme qu'il existe un $\aleph$ minimal tel que $\aleph(\mathfrak{m}) \not\leq \mathfrak{m}$. En particulier, on montre en posant $\mathfrak{m} := \aleph_0$ l'existence d'un ensemble indénombrable $\aleph_1 := \aleph(\aleph_0)$. Le nombre cardinal $\aleph_1$ est tel qu'il ne saurait exister de nombre cardinal $\mathfrak{n}$ avec $\aleph_0 < \mathfrak{n} \leq \aleph_1$ car l'injection de $\mathfrak{n}$ dans $\aleph_1$ confèrerait à $\mathfrak{n}$ une structure de bon ordre, ce qui ferait de lui un $\aleph$. Or, cette dernière hypothèse est insoutenable puisqu'elle va à l'encontre de la propriété de minimalité de $\aleph_1$.

Le théorème de Cantor permet lui aussi de mettre le doigt sur un ensemble indénombrable. En effet, en appliquant ce théorème à $\aleph_0$, on obtient que $\aleph_0 < 2^{\aleph_0}$.

Ayant obtenus les nombres cardinaux $2^{\aleph_0}$ et $\aleph_1$ par des méthodes dissemblables, le mathématicien allemand Georg Cantor s'interrogea quant à savoir si ces deux nombres pouvaient avoir d'autres liens les unissant ou d'autres caractéristiques communes que le simple fait d'être tous deux indénombrables.

L'*hypothèse du continu* est une conjecture apportant une réponse concrète à la question soulevée par Cantor concernant la relation entre les nombres cardinaux $2^{\aleph_0}$ et $\aleph_1$. Elle se décline sous plusieurs formes, dont voici les trois plus fréquemment rencontrées dans la littérature scientifique :

CH_1 : (Cantor, 1878 (12)) Tout ensemble infini de nombres réels est soit dénombrable, soit il a la puissance du continu. De façon rigoureuse,

$$\aleph_0 < \mathfrak{m} \leq 2^{\aleph_0} \quad \Rightarrow \quad \mathfrak{m} = 2^{\aleph_0};$$

CH_2 : (Cantor, 1895 (19))

$$2^{\aleph_0} \;=\; \aleph_1;$$

CH_3 : Tout ensemble strictement plus grand, au sens de la cardinalité, que l'ensemble des naturels doit contenir un exemplaire - à un réétiquetage des éléments près - de l'ensemble des nombres réels. Formellement,

$$\aleph_0 < \mathfrak{m} \quad \Rightarrow \quad 2^{\aleph_0} \leq \mathfrak{m}.$$

Pour Cantor, l'hypothèse du continu est d'autant plus naturelle que les opérations d'arithmétique cardinale élémentaire nous dirigent avec insistance vers l'un ou l'autre de $\aleph_0$ et $2^{\aleph_0}$, laissant suspecter qu'il n'existe rien entre ces deux nombres cardinaux. On vérifie en effet que :

$$\aleph_0 \;=\; n+\aleph_0 \;=\; \aleph_0+\aleph_0 \;=\; \aleph_0\cdot n \;=\; \aleph_0^2 \;=\; \aleph_0^n$$

et

$$\mathfrak{c} \;=\; 2^{\aleph_0} \;=\; n^{\aleph_0} \;=\; \aleph_0^{\aleph_0}.$$

L'énoncé CH_2 est vraisemblablement le plus fidèle à l'intuition de Cantor mais il requiert un cadre conceptuel plus sophistiqué, ce qui le rend moins accessibles de prime abord que les formulations CH_1 et CH_3. En ce qui a trait aux liens unissant ces trois énoncés, on pourra vérifier que

$$\mathsf{CH}_3 \;\Rightarrow\; \mathsf{CH}_2 \;\Rightarrow\; \mathsf{CH}_1.$$

De plus, en présumant que $2^{\aleph_0}$ est un $\aleph$ - une hypothèse qui est notamment vérifiée lorsqu'on travaille dans la théorie ZF enrichie de l'axiome du choix - ces implications sont réversibles.

Cependant, il est compatible avec $\mathsf{ZF}+\neg\mathsf{AC}$ que CH_2 soit vérifiée mais que CH_3 échoue. Le fait que $2^{\aleph_0}=\aleph_1$ est en effet cohérent avec l'existence un nombre cardinal $\mathfrak{m}>\aleph_0$ incomparable avec $2^{\aleph_0}$, c'est-à-dire tel que $\mathfrak{m}\not\leq 2^{\aleph_0}$ et $\mathfrak{m}\not\geq 2^{\aleph_0}$. Par ailleurs, le mathématicien américain Robert M. Solovay (61) a construit un modèle de $\mathsf{ZF}+\neg\mathsf{AC}$ dans lequel CH_1 est vérifiée, mais où CH_2 et CH_3 sont infirmées. Comme nos travaux se situeront toujours dans la théorie ZFC, nous ne ferons pas la distinction entre l'une ou l'autre des interprétations possibles de l'hypothèse du continu ; le sigle CH référera concurremment à ces trois énoncés. Il importe néanmoins de garder en tête que ceux-ci ne sont pas équivalents par essence.

Les travaux des mathématiens Kurt Gödel (27) et Paul Cohen (21) stipulent que la théorie des ensembles ZFC est trop faible pour trancher la question à savoir si l'hypothèse du continu est vraie ou non. Autrement dit, ZFC ne permet pas de statuer sur l'hypothèse du continu. De plus, contrairement à ce qui se produit pour l'axiome du choix, les arguments empiriques et philosophiques ne permettent pas pour le moment de faire pencher la balance ni du côté de l'hypothèse du continu ni du côté de sa négation afin d'arriver à un consensus.

D'une part, en soutient à l'hypothèse du continu, Sierpiński (56) en a tiré 82 conséquences - dont certaines seront présentées dans la seconde partie de ce mémoire - faisant subir à la théorie des ensembles d'importantes simplifications ; en matière de commodité, la théorie $\mathsf{ZFC}+\mathsf{CH}$ revêt certains attraits. En comparaison, Martin et Solovay (45) ont souligné que la négation de l'hypothèse du continu ne semble en mesure de statuer sur aucune de ces 82 propositions.

D'autre part, l'argument psychologique suivant plaide en faveur de $\mathsf{ZFC}+\neg\mathsf{CH}$: une des raisons motivant l'introduction de l'axiome d'infinité puis du schéma d'axiomes de substitution est qu'il semble absurde de croire que le principe d'engendrement des nombres ordinaux puisse venir à bout de générer l'intégralité de la classe des ordinaux sans un petit coup de pouce, sans une impulsion nouvelle. De même, l'axiome de l'ensemble des parties se doit d'être un audacieux axiome permettant d'engendrer des ensembles beaucoup plus vastes que les simples ensembles obtenus par à-coups, c'est-à-dire par substitution et formation de successeur.

Chapitre 4

Principe de dualité résultant de l'hypothèse du continu

La cardinalité, les catégories de Baire et la mesure de Lebesgue suggèrent chacune à leur façon une certaine notion de taille, respectivement en terme de nombre d'éléments, de densité ou de superficie. Malgré ce support intuitif commun, les considérations ayant motivé le développement et l'introduction de ces trois concepts sont plutôt divergentes ; la cardinalité est un pur produit de la théorie des ensembles, les catégories de Baire interviennent principalement dans le cadre de la topologie générale et de l'analyse fonctionnelle tandis que les mesures ont dans un premier temps été développées en vue d'être appliquées aux théories de l'intégration ainsi qu'aux probabilités.

L'objectif de ce chapitre est de montrer comment l'hypothèse du continu permet d'établir un principe de dualité entre les concepts de mesure et de catégorie. Les deux premières sections seront consacrées à l'étude préalable de ces deux concepts. Dans la troisième section, nous mettrons en lumière certaines similarités entre la classe de sous-ensembles de $\mathbb{R}$ de mesure nulle et celle des sous-ensembles de $\mathbb{R}$ de première catégorie. Cette troisième section se terminera sur une note discordante puisque nous présenterons un résultat mettant en relief le caractère distinct de ces deux classes d'ensembles.

Loin de saper notre enthousiasme, la non-comparabilité par inclusion de ces deux classes d'ensembles est précisément ce qui rend digne d'intérêt le principe de dualité que nous obtiendrons dans la dernière section grâce à l'hypothèse du continu.

Notation. Dans l'unique but d'assurer la fluidité du texte, nous utiliserons l'expression *ensemble linéaire* en guise de synonyme de « sous-ensemble de $\mathbb{R}$ ».

4.1 Catégories de Baire

Avant de fournir une définition de ce qu'est un ensemble linéaire de première catégorie de Baire et d'exhiber les principales propriétés de tels ensembles, il est impératif de préciser quelques notions topologiques de base.

Définition 4.1. Soit (X, τ) un espace topologique et soit A un sous-ensemble de X.

1. On dit que A est *dense* dans X si, pour tout élément x dans X, tout voisinage de x contient au moins un point de A.
2. On dit que A est *nulle part dense* dans X si l'intérieur de la fermeture de A est vide, c'est-à-dire si
$$\mathrm{int}(\overline{A}) = \emptyset.$$

Intuitivement, un ensemble nulle part dense est *de taille infime* du point de vue topologique ; il ne s'apparente en rien à un ouvert, pas même si l'on considère sa fermeture.

Les deux propositions suivantes donnent des caractérisations des ensembles nulle part denses et des ensembles fermés nulle part denses qui seront utiles par la suite.

Proposition 4.1. Soit (X, τ) un espace topologique et soit N un sous-ensemble de X. Alors, les énoncés suivants sont équivalents :

1. N est nulle part dense ;
2. Le complément de la fermeture de N est dense dans X ;
3. Pour chaque ouvert non vide $U \subseteq X$, il existe un ouvert non vide $V \subseteq U \backslash N$.

Démonstration.

(1 ⇔ 2) L'ensemble N est nulle part dense dans X si et seulement si $\mathrm{int}(\overline{N}) = \emptyset$. Par la définition de l'intérieur, il découle que $\mathrm{int}(\overline{N}) = \emptyset$ si et seulement si tout ouvert de (X, τ) contient un point de $X \backslash \overline{N}$. Ainsi, $X \backslash \overline{N}$ est dense dans X ;

(1 ⇒ 3) Soit $x \in U$. Si pour tout voisinage de x dans U, noté V_x, on a $V_x \cap N \neq \emptyset$, alors $x \in \overline{N}$. Cela ne peut être le cas pour tout $x \in U$ puisqu'on aurait alors $U \subseteq \overline{N}$, contredisant le fait que N soit nulle part dense. Ainsi, il existe un point $x \in U$ et un voisinage V_x de x tel que $V_x \subseteq U \backslash N$;

(1 ⇐ 3) Nous allons montrer qu'aucun ensemble ouvert et non vide n'est inclus dans $\overline{N}$. Puisque $\mathrm{int}(\overline{N})$ est un sous-ensemble ouvert de $\overline{N}$, celui-ci devra alors être vide, démontrant du coup que N est nulle part dense.

Soit U un ouvert non vide. Par hypothèse, il existe un ouvert non vide $V \subseteq U \backslash N$. Soit $x \in V$. Supposons que $x \in \overline{N}$, alors tout voisinage de x contient un point de N. Cela contredit le fait que x appartienne à l'ouvert $V \subseteq U \backslash N$. Ainsi, l'hypothèse voulant que

$x \in \overline{N}$ est insoutenable ; on a donc montré que $V \nsubseteq \overline{N}$ et, par la même occasion, que $U \subseteq \overline{N}$.

■

Proposition 4.2. Soit (X, τ) un espace topologique. Un ensemble F fermé dans X est nulle part dense si et seulement s'il est la frontière d'un certain ensemble U ouvert dans X.

Démonstration.

($\Rightarrow$) Soit $F \subseteq X$ un ensemble fermé nulle part dense. Soit l'ouvert $U := X \backslash F$. Comme F est fermé et nulle part dense, la proposition précédente implique son complément est dense dans X. Ainsi,

$$\begin{aligned} F &= X \backslash U \\ &= \overline{U} \backslash \mathrm{int}(U) \\ &= \partial U; \end{aligned}$$

($\Leftarrow$) Soit U un ouvert dans X et supposons que $F = \partial U$. L'ensemble F est fermé puisque la frontière de U peut s'écrire comme l'intersection des fermés $\overline{U}$ et $\overline{X \backslash U}$. Donc,

$$\begin{aligned} \mathrm{int}(\overline{F}) &= \mathrm{int}(F) \\ &= \mathrm{int}(\overline{U} \cap \overline{X \backslash U}) \\ &\overset{a)}{=} \mathrm{int}(\overline{U}) \cap \mathrm{int}(\overline{X \backslash U}) \\ &= \mathrm{int}(\overline{U}) \cap \mathrm{int}(X \backslash U) \\ &\overset{b)}{=} \mathrm{int}(\overline{U}) \cap \big(X \backslash \overline{U}\big) \\ &\subseteq \overline{U} \cap \big(X \backslash \overline{U}\big) \\ &= \emptyset. \end{aligned}$$

Cela montre que F est nulle part dense. Notons que cette démonstration a nécessité l'utilisation des deux propriétés élémentaires de l'intérieur que voici : pour tous les sous-ensembles A et B d'un espace topologique (X, τ),

a) $\mathrm{int}(A \cap B) = \mathrm{int}(A) \cap \mathrm{int}(B)$,

b) $(X \backslash \overline{A}) = \mathrm{int}(X \backslash A)$.

■

La classe des ensembles nulle part denses possède des propriétés de clôture par rapport à certaines opérations qui confirment notre conception intuitive voulant que les ensembles nulle part denses soient en quelque sorte *négligeables.*

Proposition 4.3.

1. Tout sous-ensemble d'un ensemble nulle part dense est nulle part dense ;
2. La fermeture d'un ensemble nulle part dense est nulle part dense ;
3. L'union d'un nombre fini d'ensembles nulle part dense est nulle part dense.

Démonstration.

1. Le contraire contredirait la définition.
2. Cela découle directement du fait que pour tout ensemble A, on a $\overline{\overline{A}} = \overline{A}$;
3. Supposons que $N_1, ..., N_k$ soient des ensembles nulle part denses dans X. Étant donné un ensemble ouvert et non vide $U \subseteq X$, comme N_1 est nulle part dense, il existe par - la proposition 4.1 - un ouvert non vide $U_1 \subseteq U \backslash N_1$. De même, puisque N_2 est nulle part dense, il existe un ouvert non vide $U_2 \subseteq U_1 \backslash N_2$. De façon générale, pour chaque $i \in \{1, ..., k\}$, comme N_i est nulle part dense, il existe un ouvert non vide $U_i \subseteq U_{i-1} \backslash N_i$. Comme les ouverts U_i sont emboîtés et que $N_i \cap U_i = \emptyset$, on a que $N_i \cap U_k = \emptyset$ quel que soit $i \in \{1, ..., k\}$. Ainsi,

$$\begin{aligned} \left(\bigcup_{i=1}^{k} N_i \right) \cap U_k &= \bigcup_{i=1}^{k} (N_k \cap U_k) \\ &= \bigcup_{i=1}^{k} \emptyset \\ &= \emptyset. \end{aligned}$$

Donc, $U_k \subseteq U \backslash \left(\bigcup_{i=1}^{k} N_k \right)$ et, la proposition 4.1, on obtient la conclusion désirée.

■

On vérifie aisément - en considérant par exemple l'ensemble des nombres rationnels dans l'espace topologique euclidien $\mathbb{R}$ - que les unions dénombrables d'ensembles nulle part denses ne sont pas nécessairement nulle part denses. Par contre, la clôture de la classe des ensembles de mesure nulle d'un espace topologique (X, τ) par rapport aux unions dénombrables donne lieu à la définition suivante.

Définition 4.2. Soit (X, τ) un espace topologique et soit A un sous-ensemble de X.

1. On dit que A est *maigre* ou *de première catégorie* s'il peut s'écrire comme une union d'un nombre au plus dénombrable d'ensembles nulle part denses dans X.
2. On dit que A est *non maigre* ou *de seconde catégorie* s'il n'est pas de première catégorie ou, de façon équivalente, s'il peut s'écrire comme une intersection d'un nombre au plus dénombrable d'ensembles d'intérieur dense dans X.

Les propriétés de clôture dont jouit la classe des ensembles maigres d'un espace topologique (X, τ) étant observables chez d'autres classes d'ensembles, nous allons par souci de commodité isoler celle-ci au sein d'une définition.

Définition 4.3. Soit X un ensemble. On appel *σ-idéal* de X toute classe $\mathcal{C}$ de sous-ensembles de X satisfaisant aux propriétés suivantes :

1. La classe $\mathcal{C}$ contient l'ensemble vide ;
2. Étant donnés A et B des sous-ensembles de X, si $B \in \mathcal{C}$ et $A \subseteq B$, alors $A \in \mathcal{C}$. En d'autres termes, $\mathcal{C}$ contient tous les sous-ensembles de ses éléments ;
3. Étant donné une famille dénombrable d'ensembles A_n appartenant à la classe $\mathcal{C}$, alors, $\mathcal{C}$ contient également $\bigcup\limits_{n \in \mathbb{N}} A_n$.

Remarquons au passage que la classe des ensembles au plus dénombrables est un σ-idéal compris dans la classe des ensembles maigres d'un espace topologique.

Notation. Nous noterons dorénavant le σ-idéal des ensembles linéaires maigres par $\mathcal{M}$.

La définition suivante décrit un certain type d'ensembles, appelés F_σ, qui sont aux fermés ce que sont les ensembles maigres aux ensembles nulle part denses. Par complémentation, on obtient un type d'ensembles appelés G_δ.

Définition 4.4. Soit (X, τ) un espace topologique et soit A un sous-ensemble de X.

1. On dit que A est un F_σ s'il peut s'écrire comme une union dénombrable de sous-ensembles fermés dans X ;
2. On dit que A est un G_δ s'il peut s'écrire comme une intersection dénombrable de sous-ensembles ouverts dans X.

Bien que ces deux concepts ne soient pas directement liés aux catégories de Baire, ils interviendront accessoirement dans la démonstration des résultats de dualité que nous cherchons à obtenir.

Proposition 4.4. Tout ensemble maigre est contenu dans un F_σ maigre.

Démonstration. Soit (X, τ) un espace topologique et soit $M \subseteq X$ un ensemble maigre. Il existe une famille d'ensembles nulle part denses $\{N_i\}_{i \in \mathbb{N}}$ tels que

$$M = \bigcup_{i \in \mathbb{N}} N_i.$$

On sait, par la proposition 4.3, que la fermeture d'un ensemble nulle part dense est nulle part dense. Par conséquent, l'union $\bigcup\limits_{i \in \mathbb{N}} \overline{N_i}$, qui est un ensemble contenant M, est un F_σ maigre. ■

Théorème 4.5. (Théorème de Baire pour la droite réelle) Le complément relatif à $\mathbb{R}$ de tout ensemble linéaire maigre est dense dans $\mathbb{R}$.

Démonstration. Soit M un ensemble linéaire maigre. Il existe une famille dénombrable $\{N_l\}_{l\in\mathbb{N}}$ d'ensembles nulle part denses telle que

$$M = \bigcup_{l=1}^{\infty} N_l.$$

Étant donné un intervalle I quelconque, on considère un sous-intervalle fermé de $I\backslash N_1$, que nous noterons I_1. Puis, pour chaque $l \geq 2$, on considère un sous-intervalle fermé I_l de $I_{l-1}\backslash N_l$. Par le théorème des intervalles emboîtés, $\bigcap_{l\in\mathbb{N}} I_l$ est un sous-ensemble non vide de $I\backslash M$. Puisque ce raisonnement est exact pour tout intervalle $I \subseteq \mathbb{R}$, on a que $\mathbb{R}\backslash M$ est dense.

Comme nous avons dû effectuer une infinité de choix d'intervalles, il est nécessaire d'édicter une règle précisant comme effectuer ces choix. Considérons l'ensemble des intervalles linéaires fermés dont les extrémités sont rationnelles. Puisque ces intervalles sont entièrement déterminés par leurs extrémités, on vérifie qu'il en existe un nombre dénombrable. Considérons une énumération $\{J_l\}_{l\in\mathbb{N}}$ de ces intervalles. Pour effectuer notre choix d'intervalle I_l, il suffit alors de poser $I_0 := I$ et, pour $l > 0$, on prend I_l comme étant le premier J_k tel que $J_k \subseteq I_{l-1}\backslash N_l$. ■

Corollaire 4.6. Aucun intervalle de la droite réelle n'est maigre ; $\mathcal{M}$ ne contient pas d'intervalles.

Corollaire 4.7. L'intersection de toute famille dénombrable d'ensembles linéaires ouverts denses dans $\mathbb{R}$ est dense dans $\mathbb{R}$.

4.2 Mesure

Définition 4.5. Un ensemble linéaire A est dit *de mesure nulle* s'il peut être couvert par une famille d'intervalles dont la longueur totale est arbitrairement petite. En d'autres termes, un ensemble linéaire A est de mesure nulle si, pour tout $\varepsilon > 0$, on peut trouver une famille $\{I_n\}_{n\in\mathbb{N}}$ telle que

$$A \subseteq \bigcup_{n=1}^{\infty} I_n \qquad \text{et} \qquad \sum_{n=1}^{\infty} m(I_n) < \varepsilon.$$

On vérifie aisément que la classe des ensembles linéaires de mesure nulle forme un σ-idéal - que nous noterons par $\mathcal{N}$ - contenant le σ-idéal formé des ensembles au plus dénombrables. Voici un théorème suivi de son interprétation en terme de connaissance de $\mathcal{N}$.

Théorème 4.8. (Borel, (7)) Si une famille au plus dénombrable d'intervalles I_n recouvre un intervalle I, alors

$$\sum m(I_n) \;\geq\; m(I).$$

Démonstration. La démonstration se fait en deux temps :

- Supposons d'abord que $I = [a, b]$ et que les I_n soient des intervalles ouverts. Par le théorème de Heine–Borel, l'intervalle I est compact. Ainsi, le recouvrement de I par la famille d'ouverts I_n admet un sous-recouvrement fini $\{I_n\}_{n=1}^N$. Puisque la mesure de Lebesgue d'un ensemble est non négative, on peut se restreindre sans perte de généralité aux seules familles finies d'ouverts recouvrant I. Notons par (a_1, b_1) le premier intervalle contenant a. On a alors

$$a_1 < a < b_1.$$

  Si $b_1 \leq b$, notons par (a_2, b_2) le premier intervalle contenant b_1. De façon générale, si $b_{n-1} \leq b$, notons par (a_n, b_n) le premier intervalle contenant b_{n-1}. Puisque la famille est $\{I_n\}_{n=1}^N$ est finie et qu'il s'agit d'un recouvrement de I, ce procédé doit éventuellement prendre fin. On obtient ainsi une sous-famille

$$(a_1, b_1), ..., (a_k, b_k)$$

  telle que pour tout $n > 1$,

$$a_n < b_{n-1} < b_n \qquad \text{et} \qquad a_k < b < b_k.$$

  Ainsi,

$$b - a \;<\; b_k - a_1 \;=\; \left(\sum_{n=2}^{k}(b_n - b_{n-1})\right) + b_1 - a_1 \;\leq\; \sum_{n=1}^{k}(b_n - a_n). \tag{4.1}$$

  Par conséquent, le théorème est valide dans le cas considéré ;

- Dans le cas général, pour tout $\alpha > 1$, considérons un sous-intervalle $J \subseteq I$ fermé et tel que

$$m(J) = \frac{m(I)}{\alpha}.$$

  Soit également une famille $\{J_n\}_{n\in\mathbb{N}}$ d'intervalles ouverts contenant I_n avec

$$m(J_n) = \alpha m(I_n).$$

  Alors, J est recouvert par la famille $\{J_n\}_{n\in\mathbb{N}}$. On a montré précédemment que

$$\sum_{n\in\mathbb{N}} m(J_n) \;\geq\; m(J).$$

  Par conséquent,

$$\alpha \sum_{n\in\mathbb{N}} m(I_n) \;=\; \sum_{n\in\mathbb{N}} m(J_n) \;\geq\; m(J) \;=\; \frac{m(I)}{\alpha}.$$

  En laissant α tendre vers 1, on obtient la conclusion voulue.

■

L'équation (4.1) montre que la mesure de Lebesgue d'un intervalle n'est jamais nulle ; la classe $\mathcal{N}$ ne contient aucun intervalle.

Le concept de mesure nulle nous dit ce qu'il est possible de *négliger* en ce qui a trait à la notion de « longueur ».

Définition 4.6. La *mesure extérieure* d'un ensemble linéaire A est donnée par

$$m^*(A) := \inf \left\{ \sum_{n=1}^{\infty} \ell(I_n) : \quad \text{les } I_n \text{ sont des intervalles et } A \subseteq \bigcup_{n=1}^{\infty} I_n \right\},$$

où $\ell(I_k)$ désigne la longueur de l'intervalle I_k.

Proposition 4.9. Un ensemble linéaire est de mesure nulle si et seulement si sa mesure extérieure est nulle.

Démonstration. Pour simplifier la notation, posons

$$Z_A = \left\{ \sum_{n=1}^{\infty} \ell(I_n) : \quad \text{les } I_n \text{ sont des intervalles et } A \subseteq \bigcup_{n=1}^{\infty} I_n \right\}.$$

Ainsi, $m^*(A) = \inf Z_A$.

$\Rightarrow$) Supposons que $A \subseteq \mathbb{R}$ est de mesure nulle. On veut montrer que $\inf Z_A = 0$. Pour ce faire, nous allons montrer que pour tout $\varepsilon > 0$, on peut trouver $z \in Z_A$ tel que $z < \varepsilon$. Par définition, comme A est de mesure nulle on peut trouver une famille $\{I_n\}_{n\in\mathbb{N}}$ d'intervalles recouvrant A tel que

$$\sum_{n=1}^{\infty} l(I_n) < \varepsilon.$$

Cette sommation constitue l'élément z de Z_A que nous désirions ;

$\Leftarrow$) Si $A \subseteq \mathbb{R}$ est tel que $m^*(A) = 0$, alors par la définition de la mesure extérieure, pour tout $\varepsilon > 0$, il existe $z \in Z_A$ avec $z < \varepsilon$. Or, un élément de Z_A est la somme totale d'un recouvrement de A. Ainsi, il existe un recouvrement $\{I_n\}_{n\in\mathbb{N}}$ de A de longueur totale inférieure à ε. Par conséquent, A est de mesure nulle.

■

Cette proposition montre que la négation de « $A \subseteq \mathbb{R}$ est de mesure nulle » est « $A \subseteq \mathbb{R}$ est de mesure extérieure positive ».

4.3 Similarités et distinctions entres les classes $\mathcal{M}$ et $\mathcal{N}$

Voici quelques observations au sujet de la classe des ensembles linéaires maigres et de celle des ensembles linéaires de mesure nulle laissant entrevoir un certain parallélisme.

1. $\mathcal{M}$ et $\mathcal{N}$ sont des σ-idéaux ;
2. $\mathcal{M}$ et $\mathcal{N}$ contiennent strictement la classe des ensembles linéaires dénombrables. Pour démontrer que l'inclusion est stricte, il suffit de considérer l'ensemble de Cantor qui est maigre, de mesure nulle et qui a la puissance du continu ;
3. Ni $\mathcal{M}$, ni $\mathcal{N}$ ne contiennent d'intervalles. Le complément d'un ensemble linéaire appartenant à n'importe laquelle de ces deux classes est dense dans $\mathbb{R}$;
4. Les ensembles appartenant à l'une ou l'autre de ces deux classes possèdent certaines propriétés de *petitesse* :
   - D'un point de vue géométrique, on peut voir les ensembles nulle part denses comme étant clairesemés, c'est-à-dire densément perforés par des intervalles. Quant aux ensembles maigres, bien qu'ils puissent ne pas avoir de larges trous, ils ont toujours un ensemble dense d'interstices. Aucun intervalle ne saurait être représenté par une union dénombrable d'ensembles maigres,
   - Un ensemble linéaire de mesure nulle est perçu comme étant de taille infinie du fait qu'il peut être recouvert par une famille dénombrable d'intervalles dont la taille totale est arbitrairement petite.

Les propriétés communes de $\mathcal{M}$ et $\mathcal{N}$ ne doivent pas nous berner, aucune de ces classes ne contient l'autre. Le théorème suivant montre de façon sans équivoque qu'un ensemble linéaire peut être perçu comme étant de taille infime selon une interprétation, mais imposant selon l'autre.

Théorème 4.10. La droite réelle peut être partitionnée en deux ensembles complémentaires A et B de sorte que A soit un F_σ maigre et que B soit un G_δ de mesure nulle.

Démonstration. Soit $\{q_i\}_{i\in\mathbb{N}}$ une énumération des nombres rationnels (un ensemble linéaire dense). Pour $j \in \mathbb{N}$, soit $I_{i,j}$ l'intervalle ouvert de centre r_i et de longueur $\frac{1}{2^{i+j}}$. Posons

$$G_j := \bigcup_{i=1}^{\infty} I_{i,j} \qquad \text{et} \qquad B := \bigcap_{j=1}^{\infty} G_j.$$

Notons que B est un ensemble non vide puisque les nombres rationnels sont contenus dans tous les G_j et donc dans leurs intersections. Pour tout $\varepsilon > 0$, on peut trouver un $j(\varepsilon)$ tel que $\frac{1}{2^{j(\varepsilon)}} < \varepsilon$. Alors,

$$B \subseteq \bigcup_{i=1}^{\infty} I_{i,j(\varepsilon)} = G_{j(\varepsilon)}$$

et

$$\sum_{i=1}^{\infty} m(I_{i,j(\varepsilon)}) = \sum_{i=1}^{\infty} \frac{1}{2^{i+j(\varepsilon)}} = \frac{1}{2^{j(\varepsilon)}} < \varepsilon.$$

Ainsi, B est un ensemble de mesure nulle.

Pour tout $j \in \mathbb{N}$, l'ensemble G_j est ouvert et dense puisqu'il s'agit d'une union d'intervalles ouverts contenant les rationnels ; cela signifie en particulier que B est un G_δ et que les ensembles $\mathbb{R}\backslash G_j$ sont fermés. Pour $j \in \mathbb{N}$ fixé, la frontière de G_j est définie par

$$\partial G_j := \overline{G_j} \cap \overline{(\mathbb{R}\backslash G_j)} = \overline{G_j} \cap (\mathbb{R}\backslash G_j).$$

Comme G_j est dense, $\overline{G_j} = \mathbb{R}$. Ainsi, $\partial G_j = \mathbb{R}\backslash G_j$. Or, il découle des propriétés des ouverts que leur frontière est nulle part dense (propriété 4.2). Donc, les ensembles $\mathbb{R}\backslash G_j$ sont nulle part denses. Il s'ensuit que $\mathbb{R}\backslash B$ est un F_σ maigre puisque

$$\begin{aligned} \mathbb{R}\backslash B &= \mathbb{R}\backslash \bigcap_{j=1}^{\infty} G_j \\ &= \bigcup_{n=1}^{\infty} (\mathbb{R}\backslash G_j). \end{aligned}$$

■

Corollaire 4.11. Tout ensemble linéaire peut être exprimé comme l'union d'un ensemble maigre et d'un ensemble de mesure nulle.

4.4 Principe de dualité de Sierpiński–Erdős

Dans cette section, nous verrons que la théorie des catégories de Baire se distingue par ses affinités avec la théorie de la mesure. Nous présenterons certains résultats de dualité illustrant comment l'étude des catégories de Baire permet d'enrichir nos connaissances de la théorie de la mesure et d'ouvrir de nouvelles perspectives. La prochaine proposition, précédée d'un lemme, constitue un premier résultat de dualité s'articulant autour des concepts de mesure et de catégorie.

Lemme 4.12. Tout ensemble linéaire G_δ indénombrable contient un ensemble fermé nulle part dense et de mesure nulle dont l'image par une certaine fonction continue est $[0, 1]$. En particulier, cet ensemble fermé a la puissance du continu.

Démonstration. Soit $G \subseteq \mathbb{R}$ un G_δ indénombrable. Alors, il existe une famille dénombrable d'ouverts U_n tels que

$$G = \bigcap_{n=1}^{\infty} U_n.$$

Posons F l'ensemble des points de G dont tout voisinage contient un nombre indénombrable de points de G.

Nous soutenons que F est non vide. En effet, s'il en était autrement, la classe de tous les intervalles dont les extrémités sont rationnelles ne comportant qu'un nombre dénombrable de points de G constituerait un recouvrement de G. Or, comme la classe des intervalles à extrémité rationnelle est dénombrable, ce recouvrement contredirait l'hypothèse voulant que G soit indénombrable. Un argument similaire permet de voir que F n'a pas de points isolés.

Soit $I(0)$ et $I(1)$ deux intervalles fermés et disjoints de longueur au plus $\frac{1}{3}$ dont l'intérieur intersecte F et dont l'union est contenue dans U_1. En procédant inductivement, si on dispose de 2^n intervalles fermés et disjoints $I(i_1, ..., i_n)$, où $i_k \in \{0,1\}$ pour tout $k \in \mathbb{N}$, dont l'intérieur intersecte F et dont l'union est contenue dans U_n, alors on pose $I(i_1, ..., i_{n+1})$, où $i_{n+1} \in \{0,1\}$, comme étant des intervalles fermés et disjoints de longueur au plus $\left(\frac{1}{3}\right)^{n+1}$ contenus dans $U_{n+1} \cap I(i_1, ..., i_n)$ et dont l'intérieur intersecte F. L'existence de tels intervalles est assurée par le fait que F n'admet pas de points isolés et que $G \subseteq U_{n+1}$. Posons

$$C := \bigcap_{n \geq 1} \bigcup_{i_1, ..., i_n} I(i_1, ..., i_n).$$

Alors C est un sous-ensemble fermé et nulle part dense de G. De plus, C est de mesure nulle pour la même raison que l'ensemble de Cantor l'est ; en fait, on peut même montrer que C est homéomorphe à l'ensemble de Cantor. Pour tout $x \in C$, il y a une unique suite $\{i_n\}$, où $i_n \in \{0,1\}_{n \in \mathbb{N}}$, telle que $x \in I(i_1, ..., i_n)$ pour tout $n \in \mathbb{N}$. À l'inverse, toute telle suite correspond à un certain point de C.

L'application $f : C \to [0,1]$ qui envoie un point de C sur le nombre réel ayant le développement binaire $i_1, i_2, i_3, ...$ étant surjective, on trouve que C a la puissance du continu. Enfin, f est continue puisque

$$\left| f(x) - f(y) \right| \leq \frac{1}{2^n}$$

lorsque $x, y \in C \cap I(i_n, ..., i_n)$.

$\blacksquare$

Proposition 4.13.

1. Le complément d'un ensemble linéaire de mesure nulle contient un ensemble de mesure nulle de cardinalité $2^{\aleph_0}$.

2. Le complément d'un ensemble linéaire maigre contient un ensemble maigre de cardinalité $2^{\aleph_0}$.

Démonstration.

1. Soit A un ensemble linéaire tel que $m(A) = 0$. Cela signifie que $m(\mathbb{R}\backslash A) = \infty$. Par la propriété de régularité de la mesure de Lebesgue, $\mathbb{R}\backslash A$ contient un sous-ensemble fermé F tel que $m(F) > 0$; cela signifie que l'ensemble fermé F est indénombrable. Puisque dans un espace métrique, tout ensemble fermé est un G_δ, l'existence d'un sous-ensemble fermé $F' \subseteq F \subseteq (\mathbb{R}\backslash A)$ de mesure nulle et de cardinalité $2^{\aleph_0}$ découle du lemme 4.12.
2. Soit A un ensemble linéaire maigre. Alors

$$A \;=\; \bigcup_{n=1}^{\infty} A_n,$$

où les A_n sont des ensembles linéaires nulle part dense. Puisque $\bigcup\limits_{n=1}^{\infty} \overline{A}_n$ est un F_σ maigre contenant A, l'ensemble

$$\mathbb{R}\backslash \left(\bigcup_{n=1}^{\infty} \overline{A}_n \right) \;=\; \bigcap_{n=1}^{\infty} (\mathbb{R}\backslash \overline{A}_n)$$

est un ensemble G_δ comaigre[1] inclus dans $\mathbb{R}\backslash A$. Si on pouvait montrer que ce G_δ comaigre a la puissance du continu, on pourrait conclure par le lemme 4.12 que $\mathbb{R}\backslash \left(\bigcup\limits_{n=1}^{\infty} \overline{A}_n \right)$ contient un ensemble de nulle part dense (et, par le fait même, maigre) de cardinalité $2^{\aleph_0}$. Or, une construction similaire à celle du lemme 4.12 permet de montrer que $\mathbb{R}\backslash \left(\bigcup\limits_{n=1}^{\infty} \overline{A}_n \right)$ contient un exemplaire de l'ensemble de Cantor.

■

Encouragés par ce premier cas appréciable de dualité entre les σ-algèbres $\mathcal{M}$ et $\mathcal{N}$, nous sommes en droit de suspecter une certaine similarité entre elles. La nature de cette similarité reste cependant à préciser.

Théorème 4.14. (Sierpiński, 1934 (56)) Il résulte de l'hypothèse du continu qu'il existe une fonction biunivoque $f : \mathbb{R} \to \mathbb{R}$ telle que $f(A) \in \mathcal{N}$ si et seulement si $A \in \mathcal{M}$.

Cette proposition corrobore le principe de dualité suivant : *Soit P une proposition ne faisant appel qu'au concept d'ensemble maigre ainsi qu'à des concepts purement ensemblistes et soit P' la proposition obtenue de P en substituant « mesure nulle » à « maigre » chaque fois que ce terme est rencontré, alors en supposant l'hypothèse du continu, P implique P' et réciproquement.*

La question quant à savoir s'il existe un résultat plus fort, c'est-à-dire une fonction f envoyant chacune des deux classes sur l'autre simultanément a été résolue en affinant la démonstration de Sierpiński.

1. Le complément d'un ensemble maigre est dit *comaigre*.

Définition 4.7. Une involution $f : \mathbb{R} \to \mathbb{R}$ est appelée *application de Sierpiński–Erdős* si, pour tout $A \subseteq \mathbb{R}$,

1. $f(A) \in \mathcal{M}$ si et seulement si $A \in \mathcal{N}$;
2. $f(A) \in \mathcal{N}$ si et seulement si $A \in \mathcal{M}$.

Théorème 4.15. (Erdős, 1943 (25)) Il découle de l'hypothèse du continu qu'il existe une application de Sierpiński–Erdős.

Le résultat de Erdős renforce le principe de dualité énoncé plus tôt : *Soit P une proposition ne faisant intervenir que les concepts d'ensembles de mesure nulle, d'ensembles maigres ainsi que des concepts ensemblistes et soit P' la proposition obtenue de P en interchangeant les termes « maigre » et « mesure nulle » chaque fois que l'un d'eux est rencontré, alors en supposant l'hypothèse du continu, P implique P' et réciproquement ; la démonstration de P peut être traduit mutatis mutandis en une démonstration de P' et vice versa.*

Afin de démontrer le théorème de Erdős - et du même coup celui de Sierpiński, il nous faut préalablement obtenir quelques résultats intermédiaires :

Lemme 4.16. La classe des ensembles linéaires G_δ de mesure nulle, notée $\mathcal{N}_{G_\delta}$, a la puissance du continu.

Démonstration. Il n'est pas difficile de se convaincre que $\mathcal{N}_{G_\delta}$ a au moins la puissance du continu. En effet, il suffit de remarquer que chaque singleton de $\mathbb{R}$ est un G_δ de mesure nulle. Par contre, un effort un peu plus substantiel est requis afin de montrer que $\mathcal{N}_{G_\delta}$ a au plus la puissance du continu. Nous procéderons par étapes.

1. En premier lieu, remarquons que tout intervalle ouvert $(a, b) \subseteq \mathbb{R}$, où $a, b \in \mathbb{R}$ et $a < b$, est uniquement déterminé par ses extrémités. Il existe donc $2^{\aleph_0} \cdot 2^{\aleph_0}$ tels intervalles. Puisque la proposition 3.12 affirme que

$$2^{\aleph_0} \cdot 2^{\aleph_0} \;=\; 2^{\aleph_0},$$

   il ne saurait exister plus de $2^{\aleph_0}$ intervalles ouverts ;
2. Sachant que tout sous-ensemble ouvert de $\mathbb{R}$ peut s'exprimer comme une union dénombrable d'intervalles ouverts disjoints, on déduit qu'il existe $(2^{\aleph_0})^{\aleph_0}$ familles dénombrables d'intervalles ouverts. Or, par les propriétés de l'exponentiation cardinale et par la proposition 3.12, on a

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

   Il y a donc au plus $2^{\aleph_0}$ ensembles linéaires ouverts ;
3. Le même argument permet de montrer qu'il existe au plus $2^{\aleph_0}$ familles d'ensembles linéaires ouverts et, par voie de conséquence, $2^{\aleph_0}$ ensembles linéaires G_δ.

■

Lemme 4.17. La classe des ensembles linéaires F_σ maigres, notée $\mathcal{M}_{F_\sigma}$, a la puissance du continu.

Démonstration. Par la proposition 4.2, tout ensemble fermé et nulle part dense est la frontière d'un ensemble ouvert. Or, nous avons vu dans la preuve du lemme précédent qu'il existe exactement $2^{\aleph_0}$ ensembles linéaires ouverts. Il existe donc tout au plus $2^{\aleph_0}$ ensembles linéaires fermés et nulle part denses.

Comme un ensemble linéaire F_σ maigre est une union dénombrable d'ensembles fermés nulle part denses, la proposition 3.12 nous permet de conclure que $\mathcal{M}_{F_\sigma}$ a au plus la puissance du continu. En fait, $\mathcal{M}_{F_\sigma}$ a exactement la puissance du continu puisque tout singleton de $\mathbb{R}$ est un ensemble F_σ maigre.

■

Théorème 4.18. Soit X un ensemble de puissance $\aleph_1$, et soit $\mathcal{K}$ une classe de sous-ensembles de X satisfaisant aux propriétés suivantes :

1. La classe $\mathcal{K}$ est un σ-idéal ;
2. L'union des éléments de $\mathcal{K}$ est X ;
3. La classe $\mathcal{K}$ admet une sous-classe $\mathcal{G}$ de puissance $\aleph_1$ ayant la propriété que chaque élément de $\mathcal{K}$ est contenu dans un élément de $\mathcal{G}$;
4. Le complément de chaque élément de $\mathcal{K}$ contient un ensemble de puissance $\aleph_1$ appartenant à $\mathcal{K}$.

Alors, X peut être décomposé en $\aleph_1$ ensembles disjoints X_α, tous de puissance $\aleph_1$, tels qu'un sous-ensemble de X appartient à $\mathcal{K}$ si et seulement s'il est contenu dans une union dénombrable d'ensembles X_α.

Démonstration. Par hypothèse, la sous-classe $\mathcal{G}$ est de puissance $\aleph_1$. En corollaire, il existe une suite transfinie du type ω_1 formée de tous les ensembles de la classe $\mathcal{G}$. Posons, pour $\alpha < \omega_1$:

$$H_\alpha := \bigcup_{\zeta \leq \alpha} G_\zeta.$$

Supprimons de la suite transfinie

$$H_1, H_2, H_3, ..., H_\omega, H_{\omega+1}, ..., H_\zeta, ..., \qquad (\zeta < \omega_1)$$

tous les H_α pour lesquels l'ensemble $H_\alpha \setminus \bigcup_{\zeta<\alpha} H_\zeta$ est au plus dénombrable. Il résulte aisément des propriétés 3 et 4 que les termes conservés forment une sous-suite transfinie de type ω_1

$$H_{\alpha_1}, H_{\alpha_2}, H_{\alpha_3}, ..., H_{\alpha_\omega}, H_{\alpha_{\omega+1}}, ..., H_{\alpha_\zeta}, ..., \qquad (\zeta < \omega_1).$$

Pour tout $\beta \in \omega_1$, posons

$$X_\beta := H_{\alpha_\beta} \setminus \bigcup_{\zeta<\beta} H_{\alpha_\zeta}.$$

Ainsi définis, les ensembles X_β sont deux à deux disjoints et, par la propriété 1, ils appartiennent tous à $\mathcal{K}$. De plus, le critère retenu pour extraire la sous-suite $\{H_{\alpha_\beta}\}_{\beta\in\omega_1}$ nous assure que tous les ensembles X_β sont non dénombrables. Notons que pour tout $\gamma \in \omega_1$, on a que $\gamma < \alpha_\beta$ pour un certain $\beta \in \omega_1$, ce qui implique que

$$G_\gamma \subseteq H_\gamma \subseteq H_{\alpha_\beta} = \bigcup_{\zeta\leq\beta} X_\zeta.$$

Comme conséquence nécessaire de la propriété 3 et de ce qui a été dit, chaque élément de la classe $\mathcal{K}$ est contenu dans une union dénombrable d'ensembles X_β. Par la propriété 2, il s'ensuit que

$$X = \bigcup \mathcal{K} \subseteq \bigcup_{\beta\in\omega_1} X_\beta.$$

Ainsi, $\left\{X_\beta\right\}_{\beta\in\omega_1}$ est une décomposition de X satisfaisant aux propriétés désirées.

■

Théorème 4.19. Soit X un ensemble de puissance $\aleph_1$. Soit $\mathcal{K}$ et $\mathcal{L}$ deux classes de sous-ensembles de X satisfaisant aux propriétés 1 à 4 du théorème 4.18. Supposons de plus que X soit l'union de deux ensembles complémentaires K et L, avec $K \in \mathcal{K}$ et $L \in \mathcal{L}$. Alors, il existe une involution $f : X \to X$ telle que $f(A) \in \mathcal{L}$ si et seulement si $A \in \mathcal{K}$.

Démonstration. Suivant la notation adoptée dans la démonstration du théorème précédent, soit $\{X_\beta\}_{\beta\in\omega_1}$ une partition de X par rapport à la classe $\mathcal{K}$ et soit

$$G_1, G_2, G_3, ..., G_\omega, G_{\omega+1}, ..., G_\zeta, ..., \qquad (\zeta < \omega_1)$$

une suite transfinie de type ω_1 formée de tous les ensembles de la sous-classe $\mathcal{G} \subseteq \mathcal{K}$. Supposons que l'ensemble K appartienne à la sous-classe $\mathcal{G}$ et que G_0 soit choisi de sorte à être égal à K. Alors, $X_0 = K$ puisqu'il est impossible que l'ensemble K soit dénombrable ; en effet, cela entrerait en contradiction avec le fait que K et L sont des ensembles équipotents formant une partition de l'ensemble indénombrable X.

De façon similaire, soit $\{Y_\beta\}_{\beta\in\omega_1}$ une partition de X par rapport à la classe $\mathcal{L}$ avec $Y_0 = L$. Alors,

$$K = \bigcup_{0<\beta<\omega_1} Y_\beta \qquad \text{et} \qquad L = \bigcup_{0<\beta<\omega_1} X_\beta.$$

Par conséquent, les ensembles X_β et Y_β pour $0 < \beta < \omega_1$ constituent une partition de X en ensembles de puissance $\aleph_1$. Pour tout $0 < \beta < \omega_1$, soit f_β la correspondance biunivoque de X_β sur Y_β. Définissons l'application f comme étant f_β sur X_β et f_β^{-1} sur Y_β pour $0 < \beta < \omega_1$.

Alors, f est une bijection de X sur lui-même telle que $f = f^{-1}$ et $f(X_\beta) = Y_\beta$ pour tout $0 < \beta < \omega_1$. Comme

$$X_0 = \bigcup_{0<\beta<\omega_1} Y_\beta \qquad \text{et} \qquad Y_0 = \bigcup_{0<\beta<\omega_1} X_\beta,$$

on a également que $f(X_0) = Y_0$. Conséquemment, $f(X_\beta) = Y_\beta$ pour tout $0 \leq \beta < \omega_1$. Par les propriétés de X_β et Y_β énoncées au théorème 4.18, il s'ensuit que $f(A) \in \mathcal{L}$ si et seulement si $A \in \mathcal{K}$.

■

Démonstration du résultat de Erdös. Le théorème de Erdős est une conséquence immédiate du théorème 4.19. Il suffit de poser

$X := \mathbb{R}$;

$\mathcal{K} := \mathcal{M}$, la classe des ensembles linéaires maigres;

$\mathcal{L} := \mathcal{N}$, la classe des ensembles linéaries de mesure nulle.

Grâce au lemme 4.17 et à la proposition 4.4, on sait que $\mathcal{M}$ admet une sous-classe de puissance $2^{\aleph_0}$ - la classe des ensembles linéaires F_σ maigres - telle que tout élément de $\mathcal{M}$ est contenu dans un élément de cette sous-classe. De même, grâce au lemme 4.16 et à la propriété de régularité de la mesure de Lebesgue, on sait que $\mathcal{N}$ admet une sous-classe de puissance $2^{\aleph_0}$ - la classe des G_δ de mesure nulle - telle que tout élément de $\mathcal{N}$ est contenu dans un élément de cette sous-classe.

Nous savons d'emblée que $\mathcal{M}$ et $\mathcal{N}$ satisfont aux deux premières propriétés du théorème 4.18. En ce qui a trait à la troisième, elle est satisfaite sous l'hypothèse que $2^{\aleph_0} = \aleph_1$, c'est-à-dire sous l'hypothèse du continu. Enfin, la proposition 4.13 nous assure que les classes $\mathcal{M}$ et $\mathcal{N}$ satisfont à la quatrième propriété.

Pour ce qui est des ensembles $K \in \mathcal{K}$ et $L \in \mathcal{L}$ dont il est question dans l'énoncé du théorème 4.19, il suffit de considérer la partition de la droite réelle présentée au théorème 4.10.

■

S'il ne fait aucun doute que les applications de Sierpiński–Erdős maintiennent les propriétés ensemblistes de base des σ-idéaux $\mathcal{M}$ et $\mathcal{N}$, on est en droit de se demander si d'autres propriétés sont préservées par ces applications. Longtemps non résolu, le problème soulevé par le mathématicien polonais Czesław Ryll-Nardzewski, qui consiste à déterminer s'il existe une application de Sierpiński–Erdős qui soit additive, a récemment trouvé réponse. Adaptant une preuve de Tomek Bartoszyński (3), le polonais Marcin Kysiak a montré (40) qu'il n'existe pas d'application de Sierpiński–Erdős $f : \mathbb{R} \to \mathbb{R}$ telle que pour tout $x, y \in \mathbb{R}$ on ait

$$f(x + y) = f(x) + f(y).$$

Nous présenterons au prochain chapitre quelques résultats de dualité obtenus grâce au principe de Sierpiński–Erdős. Pour l'heure, nous nous appliquerons à montrer que le principe de dualité ne peut être étendu de sorte à faire de la mesurabilité et de la propriété de Baire des concepts duaux.

Définition 4.8. Un sous-ensemble A d'un espace topologique X jouit de la *propriété de Baire* s'il ne diffère d'un ouvert que par un ensemble maigre, c'est-à-dire s'il existe un ouvert $U \subseteq X$ tel que $A \triangle U$ est maigre. Les ensembles jouissant de la propriété de Baire sont également qualifiés d'ensembles *presque ouverts*.

Voici d'abord quelques faits laissant suspecter à tort qu'il puisse être possible d'étendre le principe de dualité de Sierpiński–Erdős :

i. La classe des ensembles linéaires jouissant de la propriété de Baire est une σ-algèbre qui contient la σ-algèbre de Borel.
Démonstration. (47) p. 19-20.

ii. L'ensemble de Vitali - une partie non mesurable de la droite réelle - ne jouit pas de la propriété de Baire.
Démonstration. (47) p. 22.

iii. Si tout sous-ensemble d'un ensemble linéaire A est mesurable, alors A est de mesure nulle. De même, si tout sous-ensemble d'un ensemble linéaire A jouit de la propriété de Baire, alors A est maigre.
Démonstration. (47) p. 24.

Lemme 4.20. Tout ensemble linéaire A admet un sous-ensemble dénombrable qui est dense dans A.

Démonstration. Ce résultat met à profit le fait que l'espace topologique $\mathbb{R}$ admet une base dénombrable : les intervalles ouverts aux extrémités rationnelles. Considérons la collection $\mathcal{I}$ de tous les intervalles ouverts aux extrémités rationnelles contenant un point de A. Pour chaque intervalle de $\mathcal{I}$, choisissons l'un de ses éléments qui appartienne également à A et notons-le par x_n. La collection $\{x_n\}_{n\in\mathbb{N}}$ constitue un sous-ensemble dénombrable et dense de A.

■

Théorème 4.21. (Szpilrajn, 1934 (62)) Il n'existe pas de bijection $f : \mathbb{R} \to \mathbb{R}$ telle que pour tout $A \subseteq \mathbb{R}$:

1. $f(A)$ est mesurable si et seulement si A jouit de la propriété de Baire ;
2. $f(A) \in \mathcal{N}$ si et seulement si $A \in \mathcal{M}$.

À strictement parler, la seconde condition est superflue puisqu'il s'agit d'une conséquence de la première en vertu du fait iii. mentionné plus haut et de sa réciproque.

Démonstration. Supposons que f soit une telle bijection. Posons $I := [0, 1]$ et $E := f^{-1}(I)$. Comme I est mesurable, E jouit de la propriété de Baire. Soit $\{x_n\}_{n\in\mathbb{N}}$ un sous-ensemble dénombrable qui est dense dans E et soit I_n un intervalle ouvert contenant x_n tel que

$$m\,(f(I_n) \cap I) \;<\; \left(\frac{1}{2}\right)^{n+1}.$$

Posons $G := \bigcup\limits_{n\in\mathbb{N}} I_n$. Alors, G est ouvert et $E \subseteq \overline{G}$. Ainsi,

$$E \;\subseteq\; (G \cap E) \cup (\overline{G}\backslash G).$$

Donc,

$$I \;=\; f(E) \;\subseteq\; f(G \cap E) \cup f(\overline{G}\backslash G) \;\subseteq\; \bigcup_{n\in\mathbb{N}} \big(f(I_n) \cap I\big) \cup f(\overline{G}\backslash G).$$

Comme $\overline{G}\backslash G$ est nulle part dense, $f(\overline{G}\backslash G)$ est un ensemble de mesure nulle. Conséquemment,

$$m(I) \;\leq\; \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{n+1} \;=\; \frac{1}{2},$$

ce qui est impossible.

■

Bien que le principe étendu de dualité ne soit pas valide dans l'absolu, Oxtoby (47) lui reconnaît néanmoins une certaine valeur heuristique puisque de nombreuses propriétés de la mesure ne reposent que sur des propriétés de la classe des ensembles mesurables qui sont communes à la classe des ensembles jouissant de la propriété de Baire.

Chapitre 5

Ensemble de Lusin

Parmi les diverses conséquences de l'hypothèse du continu identifiées par Sierpiński (56), il y a une - qu'il désigne par C_1 - qui est d'une importance capitale puisqu'elle se révèle être la pierre angulaire de tout un réseau de théorèmes, de propositions et de corollaires reposant de près ou de loin sur l'hypothèse du continu. Ce chapitre sera consacré à la proposition C_1 ainsi qu'à quelques autres conséquences de l'hypothèse du continu qui sont dans un rapport étroit avec la proposition C_1.

5.1 Existence d'un ensemble de Lusin

Définition 5.1.

1. Un ensemble linéaire jouit de la *propriété de Lusin* s'il intersecte tout ensemble linéaire de première catégorie en un nombre au plus dénombrable de points.

   De façon équivalente, un ensemble jouit de la propriété de Lusin si tous ses sous-ensembles indénombrables sont de deuxième catégorie.

2. Un ensemble linéaire L ayant la puissance du continu et jouissant de la propriété de Lusin est appelé *ensemble de Lusin.*

On vérifie aisément que la propriété de Lusin est *héréditaire* au sens ou elle appartient à tout sous-ensemble d'un ensemble qui en jouit.

Il est quelque peu regrettable qu'il soit entré dans l'usage d'accoler le nom du mathématicien russe Nikolaï Lusin à cette propriété et aux ensembles qui la possède puisque c'est le mathématicien allemand Friedrich Paul Mahlo qui, le premier, démontra l'existence - conditionnellement à ce que l'hypothèse du continu soit vérifiée - d'ensembles linéaires indénombrables intersectant tout ensemble de première catégorie en un nombre au plus dénombrable de points.

Théorème 5.1. (Mahlo, 1913 (44) et Lusin, 1914 (42)) Il découle de l'hypothèse du continu qu'il existe un ensemble de Lusin [1].

Démonstration. Pour l'avoir maintes fois mentionné au dernier chapitre, il est désormais admis que tout ensemble linéaire de première catégorie est contenu dans un ensemble F_σ de première catégorie. De plus, par le lemme 4.17, il existe exactement $2^{\aleph_0}$ ensembles F_σ linéaires de première catégorie.

Par l'hypothèse du continu, $2^{\aleph_0}$ coïncide avec le nombre cardinal $\aleph_1$. Ainsi, il existe une suite transfinie de type ω_1 formée de tous les ensembles linéaires F_σ de première catégorie. Définissons notre candidat au titre d'ensemble de Lusin comme suit :

$$L := \{x_\alpha\}_{\alpha\in\omega_1},$$

où, pour chaque α, l'élément x_α est n'importe quel nombre réel tel que

$$x_\alpha \in \mathbb{R}\backslash\Big(\bigcup_{\zeta<\alpha}(A_\zeta \cup \{x_\zeta\})\Big).$$

Trouver un tel x_α est toujours possible puisqu'en vertu du théorème de Baire la droite réelle n'est pas l'union d'une famille dénombrable d'ensembles linéaires de première catégorie.

Par sa construction, l'ensemble L a la puissance du continu puisque les x_α sont tous distincts. Enfin, L intersecte tout ensemble linéaire de première catégorie en un nombre au plus dénombrable de points. En effet, étant donné un ensemble linéaire de première catégorie M, on peut trouver un F_σ de première catégorie A_α qui contient M. Or,

$$L \cap A_\alpha \;\subseteq\; \{x_\zeta\}_{\zeta\leq\alpha}.$$

Ainsi, $L \cap M$ est compris dans l'ensemble $\{x_\zeta\}_{\zeta\leq\alpha}$ qui est dénombrable.

■

Proposition 5.2. Tout ensemble de Lusin est de mesure nulle.

Démonstration. Par le lemme 4.10, on peut décomposer $\mathbb{R}$ en une union d'ensembles disjoints M et N tels que M soit de première catégorie et N soit de mesure nulle. Soit L un ensemble de Lusin, alors

$$L \;=\; (L\cap M)\cup(L\cap N).$$

Puisque M est de première catégorie, il découle des propriétés définissant L que $(L \cap M)$ est dénombrable. Ainsi,

$$m(L) \;=\; m(L\cap M) + m(L\cap N) \;=\; 0.$$

■

1. Cet énoncé constitue la proposition C_1 de (56)

5.2 Résultats d'équivalence liant les ensembles de Lusin à l'hypothèse du continu

Lemme 5.3. L'hypothèse du continu est satisfaite si et seulement si $\mathbb{R}$ est l'union d'une chaîne croissante d'ensembles dénombrables, c'est-à-dire d'une famille $\{A_i\}_{i\in I}$ d'ensembles tels que $A_i \subseteq A_j$ chaque fois où $i \leq j$.

Démonstration.

$\Rightarrow$) Supposons que l'hypothèse du continu soit satisfaite, alors $\aleph_1 = 2^{\aleph_0}$. Cela signifie qu'il existe une suite transfinie de type ω_1

$$r_1, r_2, r_3, ..., r_\omega, r_{\omega+1}, ..., r_\zeta, ..., \qquad (\zeta < \omega_1)$$

formée de tous les nombres réels. Posons $A_\alpha := \{r_\zeta\}_{\zeta<\alpha}$. Alors,

$$\mathbb{R} \;=\; \bigcup_{\alpha<\omega_1} A_\alpha.$$

Cela signifie que $\mathbb{R}$ peut être représenté comme l'union d'une chaîne croissante d'ensembles dénombrables.

$\Leftarrow$) Supposons que $\mathbb{R}$ puisse être exprimée comme l'union d'une chaîne croissante d'ensembles dénombrables. En d'autres termes, supposons que $\mathbb{R} = \bigcup_{i\in I} A_i$, où

1) I est un ensemble bien ordonné d'indices ;

2) A_i est dénombrable quelque soit le choix de $i \in I$;

3) $A_i \subseteq A_j$ chaque fois que $i \leq j$.

Considérons un sous-ensemble $X \subseteq \mathbb{R}$ de puissance $\aleph_1$. Par hypothèse, pour tout $x \in X$, on a que $x \in A_{i(x)}$ pour un certain $i(x) \in I$. Il ne peut exister d'indice $j \in I$ tel que pour tout $x \in X$ on ait $i(x) < j$. En effet, puisque la famille $\{A_i\}_{i\in I}$ est une chaîne croissante ordonnée par l'inclusion, on aurait alors que $X \subseteq A_j$. Or, X est indénombrable et A_j est dénombrable. Donc, pour tout $j \in I$, il existe un indice $i(x_j) > j$ pour un certain $x_j \in X$. Ainsi,

$$\mathbb{R} \;=\; \bigcup_{i\in I} A_i \;\subseteq\; \bigcup_{i\in I} A_{i(x_j)}.$$

La cardinalité de ce dernier ensemble est au plus $\aleph_1 \cdot \aleph_0$. Or, par la proposition 3.12, on a que

$$\aleph_1 \cdot \aleph_0 \;=\; \aleph_1.$$

Sachant par le théorème de Cantor que $\mathbb{R}$ est indénombrable et ayant majoré sa cardinalité par $\aleph_1$, c'est-à-dire par le plus petit nombre cardinal indénombrable, on doit nécessairement avoir que $2^{\aleph_0} = \aleph_1$. Ainsi, l'hypothèse du continu est vérifiée.

■

Théorème 5.4. S'il existe un ensemble de Lusin et si tout ensemble linéaire de puissance strictement inférieure à celle du continu est de première catégorie, alors l'hypothèse du continu est vérifiée.

Démonstration. Supposons qu'il existe un ensemble de Lusin L et que tout ensemble linéaire de puissance strictement inférieure à celle du continu est de première catégorie. Soit

$$r_1, r_2, r_3, ..., r_\omega, r_{\omega+1}, ..., r_\zeta, ..., \qquad (\zeta < 2^{\aleph_0})$$

une suite transfinie formée de tous les nombres réels. Il peut sembler incongru d'indicer cette famille par un nombre cardinal plutôt que par un nombre ordinal. C'est qu'en l'absence de l'hypothèse du continu, on ignore a quel aleph correspond $2^{\aleph_0}$. Grâce à l'axiome du choix, nous sommes cependant assurés que $2^{\aleph_0}$ est bel et bien un aleph, c'est-à-dire un ordinal initial.

Posons

$$L_\alpha := L \cap \{r_\zeta\}_{\zeta \leq \alpha}.$$

Par hypothèse, pour tout nombre ordinal $\alpha < 2^{\aleph_0}$, la collection $\{r_\zeta\}_{\zeta \leq \alpha}$ est de première catégorie. Comme L est un ensemble de Lusin, on a donc que tous les L_α sont dénombrables. Montrons que

$$L = \bigcup_{\alpha < 2^{\aleph_0}} L_\alpha.$$

Considérons un élément quelconque $x \in L$. Comme $x \in \mathbb{R}$ on a $x = r_\alpha$ pour un certain nombre ordinal $\alpha < 2^{\aleph_0}$. Par suite, $x \in L_\alpha$ et $x \in \bigcup_{\alpha < 2^{\aleph_0}} L_\alpha$. Considérons maintenant un élément quelconque $y \in \bigcup_{\alpha < 2^{\aleph_0}} L_\alpha$. Alors, $y \in L_\alpha$ pour un certain L_α non-vide. Par la définition de L_α, on a que $y \in L \cap \{r_\beta\}_{\beta \leq \alpha}$ montrant du coup que $y \in L$.

En appliquant la même idée que dans la démonstration du lemme qui précède, on obtient que L est de cardinalité au plus $\aleph_1$. Or, tout ensemble de Lusin a par définition la puissance du continu. On a donc impérativement égalité entre $\aleph_1$ et $2^{\aleph_0}$. Ainsi, sous les hypothèses présentées dans l'énoncé du théorème, l'hypothèse du continu est vérifiée.

■

La conjonction des théorèmes 5.1 et 5.4 entraîne une quasi-équivalence entre l'hypothèse du continu et l'existence d'un ensemble de Lusin. Nous verrons en fin de chapitre un second résultat de quasi-équivalence entre ces deux assertions.

5.3 Théorème d'Egoroff

Le mathématicien italien Carlo Severini ainsi que le physicien et géomètre russe Dmitri Egoroff ont indépendamment établi la condition de convergence uniforme dans certains espaces mesurables que voici :

Théorème 5.5. (Severini, 1910 (53) et Egoroff, 1911 (23)) Soit $\{f_n\}_{n\in\mathbb{N}}$ une suite de fonctions à valeurs réelles mesurables sur l'espace mesuré $(X, \mathscr{X}, \mu)$ et soit A un ensemble mesurable de μ-mesure finie tel que $\{f_n\}_{n\in\mathbb{N}}$ converge μ-presque partout sur A vers une fonction limite f. Alors, pour tout $\varepsilon > 0$, il existe un ensemble mesurable $B \subseteq A$ avec $\mu(B) < \varepsilon$ tel que $\{f_n\}_{n\in\mathbb{N}}$ converge uniformément vers f sur $A\backslash B$.

Sierpiński (56) a démontré à l'aide de l'hypothèse du continu que le théorème d'Egoroff n'admet pas d'extension aux fonctions quelconques d'une variable réelle.

Théorème 5.6. (Sierpiński, 1928) Sous l'hypothèse du continu, il existe une suite de fonctions d'une variable réelle $\{f_n\}_{n\in\mathbb{N}}$ qui convergent non uniformément sur tout ensemble indénombrable.

Ce théorème est une conséquence immédiate de l'existence d'un ensemble de Lusin et de la proposition suivante.

Proposition 5.7. Étant donné un ensemble linéaire indénombrable Q jouissant de la propriété de Lusin, il existe une suite infinie $\{f_n\}_{n\in\mathbb{N}}$ de fonctions continues d'une variable réelle qui converge non uniformément vers 0 sur tout ensemble indénombrable de Q.

Démonstration. Soit Q un ensemble linéaire jouissant de la propriété de Lusin et soit $D = \{d_1, d_2, d_3, ...\}$ un sous-ensemble dénombrable de $\mathbb{R}\backslash Q$ dense dans l'espace ambiant $\mathbb{R}$; l'existence de l'ensemble D est assurée par le lemme 4.20. Posons

$$f(x) := \begin{cases} \frac{1}{m} & \text{si } x = d_m \text{ pour } m \in \mathbb{N}, \\ 0 & \text{si } x \in \mathbb{R}\backslash D. \end{cases}$$

On vérifie aisément que la fonction $f(x)$ est continue en tout point de $\mathbb{R}\backslash D$ et discontinue en tout point de D. Définissons

$$f_n(x) := \begin{cases} \frac{1}{m} & \text{si } x = d_m \text{ et } m \leq n, \\ 0 & \text{sinon.} \end{cases}$$

Il va sans dire que $f_n(x)$ converge vers f pour tout x réel. Reste à montrer que cette suite converge non uniformément sur tout sous-ensemble indénombrable de Q.

Soit N un sous-ensemble indénombrable arbitraire de Q. Comme l'ensemble Q jouit de la propriété de Lusin, N se doit d'être de deuxième catégorie. Cela signifie en particulier que N n'est pas nulle part dense. Par suite, il existe un intervalle I tel que N est dense dans I.

Supposons que la suite $\{f_n\}_{n\in\mathbb{N}}$ avec $n \in \mathbb{N}$ converge uniformément sur N. Puisque les fonctions f_n sont continues, il en résulte que la suite $\{f_n\}_{n\in\mathbb{N}}$ converge uniformément sur l'intervalle I

tout entier. La limite uniforme d'une suite de fonctions continues étant continue, la limite de la suite f_n se doit d'être une fonction continue sur I. Or, ce n'est pas le cas puisque la fonction f est par hypothèse discontinue en tout point de l'ensemble dense D.

■

5.4 Théorème de Fréchet de suite double

Entre 1904 et 1906, les mathématiciens français Henri-Léon Lebesgue et Maurice René Fréchet entretiennent une correspondance (64) au cours de laquelle ils discutent à quelques occasions de diverses versions du théorème suivant :

Théorème 5.8. Soit $\{f^m\}_{n\in\mathbb{N}}$ et $\{f_n^m\}_{m,n\in\mathbb{N}}$ des suites de fonctions mesurables et soit f une fonction mesurable. Supposons que

$$\lim_{m\to\infty} f^m(x) \;=\; f(x) \qquad \text{presque partout}$$

et que

$$\lim_{n\to\infty} f_n^m(x) \;=\; f^m(x) \qquad \text{presque partout.}$$

Alors, il existe deux sous-suites $\{n_k\}_{k\in\mathbb{N}}$ et $\{m_k\}_{k\in\mathbb{N}}$ telles que

$$\lim_{k\to\infty} f_{n_k}^{m_k}(x) \;=\; f(x) \qquad \text{presque partout.}$$

À l'aide de l'hypothèse du continu, Sierpiński (56) a pu exhiber une preuve du fait que le théorème de Fréchet de suite double n'admet pas d'extension aux fonctions arbitraires.

Théorème 5.9. Si l'hypothèse du continu est satisfaite, il existe une suite de fonctions d'une variable réelle $\{f^m\}_{m\in\mathbb{N}}$ et une double suite de fonctions d'une variable réelles $\{f_n^m\}_{m,n\in\mathbb{N}}$ telles que

1. Pour tout $m \in \mathbb{N}$, on a
$$\lim_{n\to\infty} f_n^m(x) \;=\; f^m(x);$$
2. On a
$$\lim_{m\to\infty} f^m(x) \;=\; 0;$$
3. Quelle que soit la suite croissante d'indices $m_1 < m_2 < m_3 < ...$ et la suite $n_1, n_2, n_3, ...$, l'égalité
$$\lim_{k\to\infty} f_{n_k}^{m_k}(x) \;=\; 0$$
ne se présente tout au plus que pour une infinité dénombrable de valeurs de x.

Démonstration. Nous prouverons que le théorème 5.6 implique le théorème 5.9.

Admettons qu'il existe une suite convergente $\{\varphi_n\}_{n\in\mathbb{N}}$ de fonctions d'une variable réelle qui converge non uniformément sur chaque ensemble indénombrable. Définissons la double suite de fonctions $f_n^m(x)$ comme suit : pour tout $m, n \in \mathbb{N}$,

$$f_n^m(x) := \begin{cases} 0 & \text{si } |\varphi_p(x) - \varphi_q(x)| < \frac{1}{m} \quad \text{pour} \quad p, q > n, \\ 1 & \text{sinon.} \end{cases}$$

Fixons m et x. La suite $\{\varphi_n\}_{n\in\mathbb{N}}$ étant par hypothèse convergente, il existe un indice $r := r(m, x)$ tel que

$$|\varphi_p(x) - \varphi_q(x)| < \frac{1}{m}$$

pour $p, q > r$. Nous en concluons tout de suite, par la définition, que $f_n^m(x) = 0$ pour $n \geq r$, ce qui donne

$$\lim_{n\to\infty} f_n^m(x) \; = \; 0.$$

Puisque cela est vrai indépendamment du choix de $m \in \mathbb{N}$ et de $x \in \mathbb{R}$, nous n'avons qu'à poser $f^m(x) \equiv 0$ quel que soit $m \in \mathbb{N}$; les deux premières conditions du théorème 5.9 s'en trouvent vérifiées.

Pour montrer la dernière condition, nous procéderons par contradiction. Soit $\{m_k\}_{k\in\mathbb{N}}$ une suite croissante de nombres naturels et admettons que

$$\lim_{k\to\infty} f_{n_k}^{m_k}(x) \; = \; 0$$

ait lieu pour un ensemble indénombrable N de valeurs de x. Les fonctions $f_n^m(x)$ ne prenant que deux valeurs 0 et 1, on en conclut qu'il existe pour tout nombre $x \in N$ un indice $\mu(x)$ tel que l'on ait

$$f_{n_k}^{m_k}(x) \; = \; 0$$

pour $k > \mu(x)$. Il en résulte que

$$|\varphi_p(x) - \varphi_q(x)| < \frac{1}{m_k}$$

pour $p, q > n_k, \; k > \mu(x)$ et $x \in N$. Puisque N est indénombrable et que l'ensemble des valeurs possibles de μ est dénombrable, il existe un nombre naturel s tel que l'égalité $\mu(x) = s$ se présente pour une infinité indénombrable de valeurs $x \in N$. Soit N_1 leur ensemble. On a donc

$$|\varphi_p(x) - \varphi_q(x)| < \frac{1}{m_k} \qquad \text{pour } p, q > n_k, \; k > s \text{ et } x \in N_1.$$

Comme la suite des indices $\{m_k\}_{k\in\mathbb{N}}$ est croissante, $m_k \geq k$. Il en résulte aussitôt que la suite de fonctions $\{\varphi_k\}_{k\in\mathbb{N}}$ converge uniformément sur l'ensemble indénombrable N_1, contrairement à la propriété admise pour cette suite. Ainsi, notre hypothèse est insoutenable et l'égalité

$$\lim_{k\to\infty} f_{n_k}^{m_k}(x) \; = \; 0$$

ne peut être vérifiée que pour les nombres réels formant un ensemble au plus dénombrable. La dernière condition du théorème 5.9 est donc réalisée.

■

5.5 Problème généralisé de la mesure

Henri-Léon Lebesgue (41) a appelé *problème de la mesure* la tâche consistant à définir une fonction $m : \mathscr{P}(\mathbb{R}) \to [0, \infty)$ vérifiant les trois conditions que voici :

I. Si $a, b \in \mathbb{R}$ avec $b \geq a$,

$$m\big([a,b]\big) \;=\; b - a;$$

II. La fonction m est invariante par translation, c'est-à-dire que si A et B sont des ensembles linéaires superposables, on a :

$$m(A) \;=\; m(B);$$

III. La fonction m est σ-additive, c'est-à-dire que si $A_1, A_2, ...$ est une suite d'ensembles deux à deux disjoints, on a :

$$m\left(\bigcup_{k=1}^{\infty} A_k\right) \;=\; \sum_{k=1}^{\infty} m(A_k).$$

Le mathématicien italien Giuseppe Vitali a montré en 1905 que ce problème n'admet pas de solutions en exhibant un ensemble de nombres réels qui est non mesurable au sens de Lebesgue (66). Quoi qu'il en soit, on peut s'interroger sur la possibilité de résoudre ce problème en imposant à la fonction m des conditions moins restrictives. Par exemple, on pourrait remplacer la σ-additivité par l'additivité finie, auquel cas Banach (1) a montré que le problème de mesure admet une solution. Une autre possibilité serait, étant donné un intervalle E ayant la puissance du continu, de chercher à définir une fonction $m(X)$ qui fasse correspondre à chaque ensemble X situé dans l'intervalle E un nombre réel $m(X) \geq 0$ de façon à vérifier les conditions suivantes :

i. Pour $X \subseteq E$ composé d'un seul élément, $m(X) = 0$;

ii. La fonction m est σ-additive ;

iii. La fonction m n'est pas identiquement 0.

Faisant écho au résultat obtenu par Vitali, les mathématiciens polonais Stefan Banach et Kazimierz Kuratowski ont répondu par la négative à cette généralisation du problème de la mesure ; pour y parvenir, il leur a cependant fallu admettre l'hypothèse du continu. Le problème généralisé de la mesure - dans lequel la notion de translation a été évacuée - a comme avantage de mettre en lumière le fait que l'impossibilité de trouver une solution au problème de la mesure n'est pas de caractère géométrique, mais bien un fait de la théorie des ensembles.

L'approche de Banach et Kuratowski (2) consiste à montrer que le théorème suivant est incompatible avec l'existence d'une solution au problème généralisé de la mesure.

Théorème 5.10. Soit un ensemble linéaire E ayant la puissance du continu. En supposant que l'hypothèse du continu soit vérifiée, il existe une double suite d'ensembles A_k^i avec $i, k \in \mathbb{N}$ telle que :

1. Pour chaque $i \in \mathbb{N}$ fixé, on a
$$E \ = \ \bigcup_{k=1}^{\infty} A_k^i;$$
2. Pour chaque $l \in \mathbb{N}$, les ensembles A_k^i sont deux à deux disjoints ;
3. Quelle que soit la suite d'entiers positifs $\{k_i\}_{i \in \mathbb{N}}$, le produit
$$\prod_{i=1}^{\infty} \left(\bigcup_{j=1}^{k_i} A_j^i \right)$$
est au plus dénombrable.

Supposons qu'il existe une fonction m satisfaisant aux conditions du problème de la mesure généralisée. Nous conclurons que le théorème 5.10 est en défaut.

On peut admettre que $m(E) \neq 0$ car, selon la condition iii, il existe un ensemble $E_1 \subseteq E$ tel que $m(E_1) \neq 0$. Donc, si $m(E) = 0$, on a conformément à la condition ii que $m(E \backslash E_1) \neq 0$. Au moins l'un des ensembles E_1 et $E \backslash E_1$ a la puissance du continu ; supposons sans perte de généralité que ce soit le cas de E_1. La fonction m étant assujettie aux conditions i à iii pour tout $X \subseteq E_1$, on en conclut aussitôt - en s'appuyant sur le fait que E et E_1 ont la même puissance - qu'il existe une fonction satisfaisant aux mêmes conditions pour tout $X \subseteq E$ et ne s'annulant pas pour $X = E$.

Soit donc $m(E) = a$, où $a \neq 0$. Considérons une double suite A_k^i satisfaisant aux conditions 1 et 2 du théorème 5.10. Il s'agit de prouver que la condition 3 est en défaut. En vertu de la condition 2, on a
$$m(E) \ = \ \sum_{k=1}^{\infty} m(A_k^1).$$
Il existe, par conséquent, un indice k_1 tel que
$$\left| \sum_{k=k_1+1}^{\infty} m(A_k^1) \right| \ \leq \ \frac{|a|}{4}.$$
En désignant l'union $\bigcup\limits_{k=k_1+1}^{\infty} A_k^1$ par R^1, on trouve que
$$\left| m(R^1) \right| \ \leq \ \frac{|a|}{4}.$$
En général, il existe une suite d'entiers $\{k_j\}_{j \in \mathbb{N}}$ tels qu'en posant
$$R^i := \bigcup_{k=k_i+1}^{\infty} A_k^i,$$

on a

$$\left| m\left(R^i \setminus \bigcup_{j=1}^{i-1} R^j\right)\right| \;\leq\; \frac{|a|}{2^{i+1}},$$

puisque la décomposition évidente

$$E \setminus \bigcup_{j=1}^{i-1} R^j \;=\; \bigcup_{k=1}^{\infty}\left(A_k^i \setminus \bigcup_{j=1}^{i-1} R^j\right)$$

entraîne

$$m\left(E \setminus \bigcup_{j=1}^{i-1} R^j\right) \;=\; \sum_{k=1}^{\infty} m\left(A_k^i \setminus \bigcup_{j=1}^{i-1} R^j\right),$$

donc pour k_i suffisamment grand :

$$\left|\sum_{k=k_i+1}^{\infty} m\left(A_k^i \setminus \bigcup_{j=1}^{i-1} R^j\right)\right| \;\leq\; \frac{|a|}{2^{i+1}}.$$

On parvient ainsi à la conclusion que

$$\begin{aligned}
\left| m\left(\bigcup_{i=1}^{\infty} R^i\right)\right| &= \left| m\left(\bigcup_{i=1}^{\infty} R^i \setminus \bigcup_{j=1}^{i-1} (R^j)\right)\right| \\
&= \left|\sum_{i=1}^{\infty} m\left(R^i \setminus \bigcup_{j=1}^{i-1} R^j\right)\right| \\
&\leq \frac{|a|}{2}
\end{aligned}$$

et comme

$$E \setminus \bigcup_{i=1}^{\infty} R^i \;=\; \prod_{i=1}^{\infty}\left(\bigcup_{j=1}^{k_i} A_j^i\right),$$

il s'ensuit que

$$\left| m\left(\prod_{i=1}^{\infty}\left(\bigcup_{j=1}^{k_i} A_j^i\right)\right)\right| \;\geq\; \frac{|a|}{2} \;\neq\; 0,$$

ce qui prouve que l'ensemble $\prod_{i=1}^{\infty}\left(\bigcup_{j=1}^{k_i} A_j^i\right)$ est indénombrable puisque, pour tout ensemble X fini ou dénombrable, on a conformément aux conditions i et ii - auxquelles la fonction m est assujettie par définition - que $m(X) = 0$.

La condition 3 du théorème 5.10 n'est donc pas réalisée. Il est ainsi établi que lorsque nous aurons montré le théorème 5.10 - à l'aide de l'hypothèse du continu, faut-il le rappeler - l'impossibilité de trouver une solution au problème de la mesure généralisée le sera aussi.

Démonstration. *Nous avons reproduit assez fidèlement la démonstration de Sierpiński (56) p. 55-56 en prenant cependant soin d'utiliser la notation moderne.*

Admettons le théorème 5.9 - qui requiert l'hypothèse du continu - et soit $f^m(x)$ ainsi que $f_n^m(x)$, des suites de fonctions satisfaisant aux conditions 1 à 3 du théorème 5.9.

Pour $i, k \in \mathbb{N}$ avec $k > 1$, posons

$$A_1^i := \bigcup_{n=1}^{\infty} \left\{ x \in E : \quad \left| f_n^i(x) - f^i(x) \right| < \frac{1}{i} \right\} \tag{5.1}$$

et

$$A_k^i := \prod_{n=k}^{\infty} \left(\left\{ x \in E : \quad \left| f_n^i(x) - f^i(x) \right| < \frac{1}{i} \right\} \setminus \left\{ x \in E : \quad \left| f_{k-1}^i(x) - f^i(x) \right| < \frac{1}{i} \right\} \right). \tag{5.2}$$

Nous allons montrer que la suite double d'ensembles A_k^i satisfait aux conditions 1 et 2 du théorème 5.10.

À cette fin, soit i un indice et x un nombre réel donné. D'après la condition 1 du théorème 5.9, il existe un indice p tel que

$$\left| f_n^i(x) - f^i(x) \right| \; < \; \frac{1}{i}$$

pour $n \geq p$. Soit k le plus petit de tels indices p. On déduit aisément des équations (5.1) et (5.2) que l'on a alors $x \in A_k^i$. La suite double A_k^i satisfait donc à la condition 1 de 5.10 et les formules (5.1) et (5.2) montrent d'elles-mêmes qu'elle remplit aussi la condition 2 de 5.10.

Soit maintenant $\{n_j\}_{j \in \mathbb{N}}$ une suite infinie de nombres naturels. En vertu des équations (5.1) et (5.2), on a

$$\bigcup_{j=1}^{m} A_{n_j}^m \;\; = \;\; \prod_{p=n_m}^{\infty} \left\{ x \in E : \quad |f_p^m(x) - f^m(x)| < \frac{1}{m} \right\}$$

et

$$\prod_{m=1}^{\infty} \left(\bigcup_{j=1}^{m} A_{n_j}^m \right) \;\; = \;\; \prod_{m=1}^{\infty} \prod_{p=n_m}^{\infty} \left\{ x \in E : \quad |f_p^m(x) - f^m(x)| < \frac{1}{m} \right\}. \tag{5.3}$$

Soit x un nombre réel tel que

$$x \in \prod_{m=1}^{\infty} \left(\bigcup_{j=1}^{m} A_{n_j}^m \right). \tag{5.4}$$

Il s'ensuit, en vertu de l'équation (5.3) que

$$\left| f_p^m(x) - f^m(x) \right| \; < \; \frac{1}{m}$$

pour $p \geq n_m$ et $m \in \mathbb{N}$, d'où en particulier

$$\left| f^m_{n_m}(x) - f^m(x) \right| \; < \; \frac{1}{m}$$

pour $m \in \mathbb{N}$, ce qui donne d'après la condition 2

$$\lim_{m \to \infty} f^m_{n_m}(x) \; = \; 0.$$

Or, en vertu du théorème 5.9, cette égalité ne peut être remplie que pour les nombres réels x formant un ensemble au plus dénombrable. La formule (5.4) ne peut donc se présenter que tout au plus pour une infinité dénombrable de x, de sorte que la suite double $\{A^i_k\}_{k,l \in \mathbb{N}}$ satisfait également à la condition 3 du théorème 5.10.

■

5.6 Application du principe de dualité de Sierpiński–Erdős aux ensembles de Lusin

Voici une définition correspondant par dualité à celle de la propriété de Lusin et des ensembles de Lusin.

Définition 5.2.

1. Un ensemble linéaire jouit de la *propriété de Sierpiński* s'il intersecte tout ensemble de mesure nulle en un nombre au plus dénombrable de points ;
2. Un ensemble linéaire S ayant la puissance du continu et jouissant de la propriété de Sierpiński est appelé *ensemble de Sierpiński*.

Théorème 5.11. (Sierpiński, 1924 (55)) Il découle de l'hypothèse du continu qu'il existe un ensemble de Sierpiński.

Par une preuve similaire à celle de la proposition 5.2, on montre que tout ensemble de Sierpiński est de première catégorie.

Le mathématicien autrichien et néo-écossais d'adoption Fritz Rothberger a établi un seconde quasi-équivalence entre l'existence d'un ensemble de Lusin et l'hypothèse du continu en montrant le théorème suivant :

Théorème 5.12. (Rothberger, 1938 (49)) S'il existe un ensemble de Lusin et un ensemble de Sierpiński, alors l'hypothèse du continu est vérifiée.

Les deux prochains résultats donnent des conditions nécessaires et suffisantes à l'existence - dans un ensemble linéaire - de sous-ensembles jouissant de la propriété de Lusin et de la propriété de Sierpiński respectivement. Il va sans dire que ces deux résultats se correspondent mutuellement par dualité.

Théorème 5.13. Il résulte de l'hypothèse du continu qu'il faut et qu'il suffit pour qu'un ensemble linéaire contienne un sous-ensemble indénombrable jouissant de la propriété de Lusin qu'il soit de deuxième catégorie de Baire.

Démonstration.

$\Rightarrow$) Soit A un ensemble linéaire de première catégorie et L un ensemble jouissant de la propriété de Baire. Alors, l'ensemble $A \cap L$ est au plus dénombrable. Par conséquent, A ne peut contenir L puisque ce dernier a la puissance du continu.

$\Leftarrow$) Soit A un ensemble linéaire de deuxième catégorie et soit $\{X_\alpha\}_{\alpha \in \omega_1}$ une décomposition de $\mathbb{R}$ vue dans la preuve du théorème 4.18 où l'on pose $K := \mathcal{M}$. Soit L l'ensemble obtenu en sélectionnant un et un seul point dans chaque ensemble non vide de la forme $A \cap X_\alpha$.

Comme ensemble de deuxième catégorie, A n'est pas l'union de $\aleph_0$ ensembles de première catégorie. Ainsi, une quantité indénombrable d'ensembles $A \cap X_\alpha$ est non-vide et, conséquemment, L est indénombrable. Par l'hypothèse du continu, la puissance de L est $2^{\aleph_0}$.

Puisque les X_α sont disjoints, aucun sous-ensemble indénombrable de L ne peut être couvert par un nombre au plus dénombrable d'ensembles X_α. Il découle donc du théorème 4.18 qu'iaucun sous-ensemble de L ayant la puissance du continu n'est de première catégorie ; $L \subseteq A$ jouit de la propriété de Lusin.

■

Corollaire 5.14. Il résulte de l'hypothèse du continu qu'il faut et qu'il suffit pour qu'un ensemble linéaire contienne un sous-ensemble indénombrable jouissant de la propriété de Sierpiński qu'il soit de mesure extérieure positive.

Il est à noter qu'aucune des propositions duales étudiées jusqu'ici n'a fait intervenir simultanément les notions de mesure et de catégorie. C'est donc dire que nous n'avons fait utilisation que du seul principe de dualité de Sierpiński. Voici un exemple nécessitant le principe de dualité de Sierpiński–Erdős dans sa pleine mesure.

Théorème 5.15. Sous l'hypothèse du continu, pour toute classe indénombrable $\mathcal{K}$ de transformations bijectives de la droite réelle préservant la mesure nulle, il existe un ensemble linéaire indénombrable E de première catégorie que chaque T dans $\mathcal{K}$ transforme en lui-même, abstraction faite tout au plus d'une infinité dénombrable de points. Autrement dit, $T(E) \triangle E$ est au plus dénombrable pour tout $T \in \mathcal{K}$.

Démonstration. Indiçons les éléments de $\mathcal{K}$ et $\mathbb{R}$ de sorte que :

$$\mathcal{K} \;=\; \{T_\alpha\}_{\alpha<\omega_1} \qquad \text{et} \qquad \mathbb{R} \;=\; \{r_\alpha\}_{\alpha<\omega_1}.$$

Par le théorème 4.10, il existe un ensemble de mesure nulle A tel que $\mathbb{R}\backslash A$ soit de première catégorie. Fixons $0<\alpha<\omega_1$. Soit G_α la collection formée de tous les produits de la forme

$$T_{\beta_1}^{k_1}T_{\beta_2}^{k_2}...T_{\beta_n}^{k_n},$$

où $\beta_i<\alpha$ et $k_i=\pm 1$ avec $i\in\{1,...,n\}$ et $n\in\mathbb{N}$. Alors, G_α est au plus dénombrable et chaque T dans G_α préserve la mesure nulle. Pour tout T dans G_α, l'ensemble $T(A)$ est de mesure nulle. Par conséquent,

$$A_\alpha := \bigcup_{T\in G_\alpha} T(A)$$

est de mesure nulle. Soit $x_0 := r_0$ et supposons que, pour $\beta<\alpha$, les points $x_\beta\in\mathbb{R}$ ont tous été définis. Posons

$$B_\alpha := \{T(x_\beta): \quad \beta<\alpha,\ T\in G_\alpha\}.$$

Alors, B_α est un ensemble au plus dénombrable et $A_\alpha\cup B_\alpha$ est un ensemble de mesure nulle. Posons x_α comme étant le premier élément de l'énumération des nombres réels donnée plus haut qui n'est pas dans $A_\alpha\cup B_\alpha$. Posons

$$E_\alpha := \{T(x_\alpha): \quad T\in G_\alpha\}$$

et définissons

$$E := \bigcup_{0<\alpha<\omega_1} E_\alpha.$$

Alors, E_α est au plus dénombrable et E est indénombrable. De plus, E est un sous-ensemble de $\mathbb{R}\backslash A$. Ainsi, E est de première catégorie. Pour tout $\beta<\alpha<\omega_1$, on a que $T_\beta(E_\alpha)=E_\alpha$. Conséquemment,

$$T_\beta(E)\triangle E \;\subseteq\; \bigcup_{\alpha\leq\beta}\Big(E_\alpha\cup T_\beta(E_\alpha)\Big).$$

Cela montre que $T(E)\triangle E$ est au plus dénombrable quelque soit T dans $\mathcal{K}$.

■

Corollaire 5.16. En supposant l'hypothèse du continu, pour toute classe $\mathcal{K}$ indénombrable de transformations bijectives de la droite réelle préservant les catégories de Baire telle que $\operatorname{card}(\mathcal{K})=2^{\aleph_0}$, il existe un ensemble linéaire indénombrable E de mesure nulle que chaque T dans $\mathcal{K}$ transforme en lui-même, abstraction faite tout au plus d'une infinité dénombrable de points. Autrement dit, $T(E)\triangle E$ est au plus dénombrable pour tout $T\in\mathcal{K}$.

Conclusion

Au cours de ce mémoire, nous aurons saisi que ni la théorie axiomatique des ensembles de Zermelo–Fraenkel, ni ses extensions ne peuvent prétendre formaliser l'*intégralité des mathématiques*, à supposer que cette expression ait un sens. Il existe des énoncés que l'on ne pourra jamais déterminer en restant dans le cadre de cette théorie. En ce début du 21$^{\text{e}}$ siècle, le problème du continu jouit d'un regain de popularité. Certains spécialistes de la théorie des ensembles estiment que telle ou telle question suggère l'introduction de tel ou tel nouvel axiome *plausible* ou *intuitivement évident* et que cela permet élucider le problème du continu. C'est notamment le cas du mathématicien américain William Hugh Woodin qui soutient que ses travaux - tout en permettant d'harnacher de nouveaux fleuves de connaissances et d'explorer des champs mathématiques nouveaux - s'inscrivent en faux contre l'hypothèse du continu. Soulignons cependant que les résultats de Woodin sont encore loin de faire l'unanimité chez les spécialistes (20).

Dans le cas inverse, c'est-à-dire s'il s'avérait que la quête d'un axiome qui soit en mesure de dénouer le noeud gordien achoppe définitivement et qu'il faille se résoudre à faire un choix cornélien entre l'hypothèse du continu et sa négation basé sur le seul critère de leur utilité à la science, les 82 conséquences identifiées par Sierpiński pourraient incidemment faire pencher la balance en faveur de l'hypothèse du continu.

En situation de *statu quo*, les questions en suspens demeurent nombreuses. D'un côté, le principe de dualité, les applications de Sierpiński–Erdős ainsi que les ensembles de Lusin et de Sierpiński n'ont pas fini de livrer leurs secrets. À l'autre bout du spectre, le rejet de l'hypothèse du continu soulève de nombreuses questions et nous confronte à notre ignorance. À quel $\aleph$ correspond le continu ? Qu'est-ce qui caractérise l'ensemble $\mathscr{P}(A)$ pour un ensemble A quelconque ? Quelles sont les propriétés des ensembles linéaires de cardinalité strictement comprise entre $\aleph_0$ et $2^{\aleph_0}$? La science actuelle n'a guère de réponses à offrir sur l'arithmétique cardinale sans l'hypothèse du continu, à plus forte raison si l'on s'interdit de recourir à l'axiome du choix.

Bibliographie

[1] Banach, Stefan. (1923). Sur le problème de la mesure. Fundamenta Mathematicae, 4, 7-33.

[2] Banach, Stefan et Kuratowski, Kazimierz. (1929). Sur une généralisation du problème de la mesure. Fundamenta Mathematicae, 14, 127-131.

[3] Bartoszyński, Tomek. (2000). A Note on Duality Between Measure and Category. Proceedings of the American Mathematical Society, 2745-2748.

[4] Belna, Jean-Pierre. (2000). Cantor (Vol. 20). Belles Lettres.

[5] Boolos, George. (1971). The Iterative Conception of Set. The Journal of Philosophy, 68(8), 215-231.

[6] Borel, Émile. (1898). Leçons sur la théorie des fonctions. Gauthier–Villars.

[7] Borel, Émile. (1905). Leçons sur les fonctions de variables réelles. Gauthier–Villars.

[8] Borel, Émile. (1914). Leçons sur les fonctions. Gauthier–Villars.

[9] Brieskorn, Egbert et Chatterji, Srishti D. (2002). Felix Hausdorff Grundzüge der Mengenlehre. Springer.

[10] Bukovský, Lev. (2011). The Structure of the Real Line, 71. Springer.

[11] Burali-Forti, Cesare. (1897). Una questione sui numeri transfiniti. Rendiconti del Ciroclo Matematico di Palermo, 11, 154-164.

[12] Cantor, Georg. (1878). Ein Beitrag zur Mannigfaltigkeitslehre Journal de Crelle, 84, 242-258.

[13] Cantor, Georg. (1880). Über unendliche, lineare Punktmannigfaltigkeiten, II. Mathematische Annalen, 17, 355-358.

[14] Cantor, Georg. (1882). Über unendliche, lineare Punktmannigfaltigkeiten, III. Mathematische Annalen, 20, 113-121.

[15] Cantor, Georg. (1883). Über unendliche, lineare Punktmannigfaltigkeiten, IV. Mathematische Annalen, 21, 51-58.

[16] Cantor, Georg. (1883a). Über unendliche, lineare Punktmannigfaltigkeiten, V. Mathematische Annalen, 21 545-591.

[17] Cantor, Georg. (1887). Mitteilungen zur Lehre vom Transfiniten. Zeitschrift für Philosophie und philosophische Kritik, 91, 81-125.

[18] Cantor, Georg. (1891). Über eine elementare Frage zur Mannigfaltigkeitslehre. Jahresbericht der Deutschen Mathematker-Vereinigung, 1, 75-78.

[19] Cantor, Georg. (1895). Beiträge zur Begründung der transfiniten Mengenlehre. Mathematische Annalen, 46(4), 481-512.

[20] Clark, Pete L. *Order and Arithmetic of Cardinalities.* [En ligne]. p. 8. http ://math.uga.edu/ pete/settheorypart2.pdf (Page consultée le 8 août 2013)

[21] Cohen, Paul. (1963). The Independence of the Continuum Hypothesis. Proceedings of the U.S. National Academy of Sciences 50(6), 1143-48.

[22] Cohen, Paul J. (1966). Set theory and the continuum hypothesis. Dover Publications.

[23] Egoroff, Dmitri. (1911). Sur les suites des fonctions mesurables. Comptes Rendus des Séances de l'Académie des sciences, 152, 244-246.

[24] Enderton, Herbert B. (1977). Elements of Set Theory. Academic Press.

[25] Erdős, Paul. (1943). Some remarks on set theory. The Annals of Mathematics, 44(4), 643-646.

[26] Fraenkel, Adolf. (1922). Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre. Mathematische annalen, 86(3), 230-237.

[27] Gödel, Kurt et Brown, George W. (1951). The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory. Princeton University Press.

[28] Halbeisen, Lorenz J. (2011). Combinatorial Set Theory. Springer.

[29] Halbeisen, Lorenz J. et Shelah, Shaharon. (2001). Relations between some cardinals in the absence of the axiom of choice. Bulletin of Symbolic Logic, 237-261.

[30] Halmos, Paul R. (1960). Naive Set Theory. Springer.

[31] Hartogs, Friedrich. (1915). Über das Problem der Wohlordnung. Mathematische Annalen, 76(4), 438-443.

[32] Heine, Eduard. (1872). Die Elemente der Functionenlehre. Journal für die reine und angewandte Mathematik, 74, 172-188.

[33] Herrlich, Horst. (2006). Axiom of Choice. (No. 1876). Springer.

[34] Hrbacek, Karel et Jech, Thomas. (1999). Introduction to Set Theory, Revised and Expanded. Crc Press.

[35] Jech, Thomas J. (2008). The axiom of choice. Dover Publications.

[36] Koellner, Peter. (2011). The Continuum Hypothesis. Stanford Encyclopedia of Philosophy.

[37] Korselt, Alwin R. (1911). Über einen Beweis des Äquivalenzsatzes. Mathematische Annalen, 70(2), 294-296.

[38] Kunen, Kenneth. (1980). Set Theory. North-Holland Publishing Company.

[39] Kuratowski, Kazimierz. (1921). Sur la notion de l'ordre dans la Théorie des Ensembles. Fundamenta Mathematicae, 2(1), 161-171.

[40] Kysiak, Marcin. (2003). Another Note on Duality Between Measure and Category. Bull. Polish Acad. Sci. Math, 51, 269-281.

[41] Lebesgue, Léon-Henri. (1904). Leçons sur l'intégration et la recherche des fonctions primitives. Gauthier–Villars.

[42] Lusin, Nikolaï. (1914). Sur un problème de M. Baire. Comptes Rendus des Séances de l'Académie des Sciences, 158. 1268-1261.

[43] Maddy, Penelope. (1988). Believing the Axioms. I. The Journal of Symbolic Logic, 53(2), 481-511.

[44] Mahlo, Paul. (1913). Über Teilmengen des Kontinuums von dessen Mächtigkeit. Jahresbericht der Deutschen Mathematiker-Vereinigung, 25, 283-315.

[45] Martin, Donald A. et Solovay, Robert M. (1970). Internal Cohen Extensions. Annals of Mathematical Logic, 2(2), 143-178.

[46] Moore, Gregory H. (2013). Zermelo's Axiom of Choice : Its Origins, Development and Influence. Dover Publications.

[47] Oxtoby, John. (1971). Measure and Category : A Survey of the Analogies between Topological and Measure Spaces. Springer-Verlag.

[48] Panasawatwong, Supakun et Vejjajiva, Pimpen. (2013). Some remarks on cardinal arithmetic without choice. Science Asia, 29, 91-94.

[49] Rothberger, Fritz. (1938). Eine Äquivalenz zwischen der Kontinuumhypothese und der Existenz der Lusinschen und Sierpinskischen Mengen. Fundamenta Mathematicae, 30, 215-217.

[50] Rothberger, Fritz. (1938). Une remarque concernant l'hypothèse du continu. Fundamenta Mathematicae, 31, 224-226.

[51] Russell, Bertrand. (1906). On some difficulties in the theory of transfinite numbers and order types. Proceedings of the London Mathematical Society, 2(4), 29-53.

[52] Schröder, Ernst. (1898). Über zwei Definitionen der Endlichkeit und G. Cantor'sche Sästze. Kaiserliche Leopoldino–Carolinische Deutsche Akademie der Naturforsher, Nova Acta, 71(6), 303-376.

[53] Severini, Carlo. (1910). Sulle successioni di funzioni ortogonali. Atti dell'Accademia Gioenia, 3(5), 1-7.

[54] Shinoda, Juichi. (1973). Some Consequences of Martin's Axiom and the Negation of the Continuum Hypothesis. Nagoya Mathematcial Journal, 49, 117-125.

[55] Sierpiński, Wacław. (1924). Sur l'hypothèse du continu ($2^{\aleph_0} = \aleph_1$). Fondamenta Mathematicae, 5(1), 177-187.

[56] Sierpiński, Wacław. (1934). Hypothèse du Continu. Monografje Matematyczne. 4. Warszawa-Lwow : Subwencji Funduszu Kultur. Narodowej.

[57] Sierpiński, Wacław. (1934). Sur la dualité entre la première catégorie et la mesure nulle. Fundamenta Mathematicae, 22(1), 276-280.

[58] Sierpiński, Wacław. (1947). L'hypothèse généralisée du continu et l'axiome du choix. Fundamenta Mathematicae, 34(1), 1-5.

[59] Sierpiński, Wacław. (1965). Cardinal and Ordinal Numbers. Warszawa : Panstwowe Wydawnictwo Naukowe.

[60] Skolem, Thoralf. (1923). Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre. Den femte skandivaniska matematikerkongressen, 290-301.

[61] Solovay, Robert M. (1970). A Model of Set-Theory in Which Every Set of Reals is Lebesgue Measurable. Annals of Mathematics, 92(1), 1-56.

[62] Szpilrajn, Edward. (1934). Remarques sur les fonctions complètement additives d'ensembles et sur les ensembles jouissant de la propriété de Baire. Fundamenta Mathematicae, 22(1), 303-311.

[63] Tarski, Alfred. (1924). Sur les ensembles finis. Fundamenta Mathematicae, 6(1), 45-95.

[64] Taylor, Angus E. et Dugac, Pierre. (1981). Quatre lettres de Lebesgue à Fréchet. Revue d'Histoire des sciences, 34(2), 149-169.

[65] Van Heijenoort, Jean. (1977). From Freege to Gödel : A Source Book in Mathematical Logic, 1879-1931. Harvard University Press.

[66] Vitali, Giuseppe. (1905). Sul Problema della Mesura dei Gruppi di Punti di una Retta. Bologna, Tip. Gamberini e Parmeggiana.

[67] von Neumann, John. (1928). Die Axiomatisierung der Mengenlehre. Mathematische Zeitschrift, 27(1), 669-752.

[68] Woodin, W. Hugh. (1999). The continuum hypothesis. I. Notices of the American Mathematical Society, 48(6), 567-576.

[69] Woodin, W. Hugh. (2001). The continuum hypothesis. II. Notices of the American Mathematical Society, 48(7), 681-690.

[70] Zermelo, Ernst. (1904). Beweis, dass jede Menge wohlgeordnet werden kann. Mathematische Annalen, 59(4), 514-516.

[71] Zermelo, Ernst. (1908). Untersuchungen über die Grundlagen der Mengenlehre. I. Mathematische Annalen, 65(2), 261-281.