Théorie des ensembles pour les philosophes

Mirna Džamonja

Théorie des Ensembles, *pour les philosophes*

Mirna Džamonja

Pour Ada et Jean-Marc, ma famille .

Table des matières

Préface

Ce livre s'est articulé autour d'un cours que l'auteure enseigne à l'Université Paris 1-Panthéon-Sorbonne au sujet de la Théorie des Ensembles au niveau M1 dans l'UFR de la Philosophie depuis 2014. Cette intervention est survenue après de nombreuses années d'enseignement de la théorie des ensembles aux étudiants en mathématiques, mais le défi du cours était d'enseigner aux philosophes. Je me suis vite rendu compte que la littérature que je connaissais n'était pas adaptée, car trop centrée sur les questions mathématiques et trop peu sur les questions philosophiques. De plus, il n'y a qu'un seul livre relativement récent en théorie des ensembles écrit en français, ce qui me semble un peu dommage pour cette belle thématique encore à explorer du point de vue pas seulement ma-

thématique, où on a appris à tout faire en anglais, mais aussi en philosophie où j'insiste qu'il ne faut pas 'apprendre' à tout faire dans une seule langue.

La théorie des ensembles fait partie des fondements des mathématiques et elle est indispensable à ceux qui sont se posés la question des fondements mathématiques. Elle n'est pas la seule façon de faire les fondements, mais elle est probablement la plus ancienne et elle donne des bons fondements à une grande partie des mathématiques. Elle a des limitations, connues et moins connues, mais même dans ces limitations, elle a été utile car elle a nous permis de se poser de bonnes questions. Ici, on va expliquer quelques unes de ces questions et donner des points de réflexion pour permettre au lecteur d'aborder la thématique, armé de suffisamment de prouesse pour ne pas tomber devant les premières difficultés. Bien sûr que la théorie moderne des ensembles est un sujet vaste et difficile en mathématiques, le sujet qui est le domaine de mes recherches mathématiques. Ce livre ne prétend pas faire de vous un adeptd des méthodes mathématiques, mais simplement vous présente les idées.

Si vous trouvez après la lecture de cet ouvrage que vous avez aimé les mathématiques exposées, vous êtes invités à consulter des livres plus avancés dont j'ai mis plusieurs dans la bibliographie.

Pour conclure, je voudrais par ce livre faire comprendre que les fondements des mathématiques et la philosophie sont en relation étroite, de plus que de très bonnes questions mathématiques ont souvent été découvertes avec une motivation philosophique et, pour les philosophes intéressés, qu'il est possible d'apprendre les méthodes mathématiques sans avoir eu une formation strictement mathématique au préalable.

Je tiens à remercier mes collègues à l'Institut d'Histoire et de Philosophie des Sciences et des Techniques IHPST de l'université Paris 1-Panthéone-Sorbonne, qui m'ont fait confiance en m'élisant membre associée et en me confiant l'enseignement de la théorie des ensembles. Je remercie Mme Catherine Schuhl et M. Jean-Marc Vanden-Broeck pour leur soutien et les corrections apportées à ce texte. Les fautes qui restent sont de la seule responsabilité de l'auteure.

En espérant que vous apprécierez votre lecture.

Première partie

Les notions de base et les axiomes de ZFC

Chapitre 1

Les premières notions et la dénombrabilité

1.1 La genèse et les premières définitions

Une série trigonométrique est une suite de la forme $\langle S_n(x) : n = 0, 1, 2, \ldots \rangle$ où chaque S_n est périodique avec la période T et la forme $S_n(x) = \Sigma_{k=0}^{n}[a_k \cos(\frac{2k\pi}{T}x) + b_k \sin(\frac{2k\pi}{T}x)]$, où l'on considère chaque $S_n(x)$ comme une fonction d'une variable réelle x. L'exemple le plus classique de série trigonométrique est la série de

Fourier, inventée en 1822. L'intérêt de ce genre de série est que pour une classe assez large de fonctions réelles $f(x)$, on peut trouver une série trigonométrique $\langle S_n(x) : n = 0, 1, 2; \ldots \rangle$ telle que pour chaque x on a $f(x) = \lim_{n\to\infty} S_n(x)$, comme cela a été déjà démontré par Fourier. Donc, même si f est une fonction assez compliquée et si on veut simplement avoir une approximation de f, comme c'est souvent suffisant dans les applications pratiques, il nous suffit de calculer S_n pour n assez large, sachant que les séries trigonométriques sont faciles à calculer. Il est donc intéressant de savoir exactement pour quelles fonctions on peut utiliser l'approximation par une série trigonométrique.

En 1869 Eduard Heine avait posé à Cantor, alors âgé de 24 ans, la question de l'unicité de l'écriture d'une fonction périodique d'une variable réelle comme la limite d'une série de fonctions trigonométriques. Intéressé par ce problème, Cantor obtint l'unicité pour les fonctions continues, dans une série d'articles entre 1870 [Can70a], [Can70b] et 1872 [Can72]. Il étendit aussi son résultat, le rendant valable lorsqu'il y a convergence en tout point sauf en un nombre

fini. En 1872, il s'attacha à définir l'ensemble des points de discontinuité des fonctions non-continues qui ont l'unicité de l'écriture, ce qui présuppose donc de manipuler des ensembles infinis. On appelle *ensemble d'unicité* un sous-ensemble de la droite réelle, défini modulo 2π, sur lequel le résultat d'unicité de Cantor s'applique.

Cantor fut donc le fondateur de la théorie des ensembles, à partir de 1874. Il s'est basé sur les idées développées par Bernard Bolzano [Bol51] en 1847. Avant cela le concept d'ensemble était plutôt basique, et avait été utilisé implicitement depuis les débuts des mathématiques, depuis Aristote. Personne n'avait compris que cette théorie avait des éléments non implicites. Avant Cantor, il n'y avait en fait que les ensembles finis (qui sont aisés à comprendre) et les ensembles infinis (qui étaient plutôt sujets à une discussion philosophique). Une des grandes contributions de Cantor est d'avoir commencé à penser à la *taille* ou à la *puissance* d'un ensemble infini. Pour ce faire, il a utilisé la notion d'une fonction *bijective*, déjà connue par Bolzano qui l'a utilisé pour démontrer qu'un ensemble est infini si est seulement s'il est bijectif avec une propre partie de

lui-même.

Définition 1.1.1. *Soit f une application de A à B. Elle est* injective *si et seulement si (ssi) elle conserve la différence, c'est-à-dire pour tout a, a' en A,*

$$f(a) = f(a') \implies a = a'.$$

Elle est surjective *ssi son image couvre l'ensemble B, c'est-à-dire que pour tout $b \in B$, il existe $a \in A$ tel que $f(a) = b$. f est une* bijection *ssi elle est injective et surjective.*

Cette notion correspond à la façon dont nous apprenons à compter « sur les doigts » : pour énumérer 5 objets, il suffit d'associer chacun de ses objets à un de nos doigts. Tandis que, pour les ensembles finis, il existe bien d'autres méthodes pour compter, la méthode introduite par Cantor pour les ensembles infinis est exactement que deux ensembles infinis sont de *la même taille* ss'il y a une bijection entre eux. Cette définition s'applique aussi aux ensembles finis et, en fait, on peut définir les nombres *naturels* $0, 1, 2, \ldots$ comme les

étalons des bijections : un ensemble a puissance 5 s'il est en correspondance bijective avec l'étalon 5. Un ensemble est fini s'il est en bijection avec un nombre naturel. Un ensemble qui n'est pas fini est infini. On reviendra sur cela avec plus de précision, mais pour l'instant regardons le premier étalon infini, l'ensemble $\mathbb{N} = \{0, 1, 2, \ldots\}$ des nombres naturels.

Définition 1.1.2. *Un ensemble est* dénombrable *ss'il est fini ou bijectif avec* $\mathbb{N}$.

On peut aussi introduire un ordre entre les tailles des ensembles : on dit que la puissance $|A|$ de l'ensemble A est $\leq$ à $|B|$ s'il existe une injection $f : A \to B$. Comme une composition des injections est aussi une injection, on a la transitivité de cette relation, c'est-à-dire si $|A| \leq |B|$ et $|B| \leq |C|$, alors $|A| \leq |C|$. Tout de suite aussi, on a la proposition suivante :

Théorème 1.1.3. *La puissance de* $\mathbb{N}$ *est la plus petite puissance infinie.*

Preuve Il s'agit de démontrer que pour chaque ensemble infini,

disons A, nous avons $|\mathbb{N}| \leq |A|$. Donc, étant donné A, nous allons définir une injection $f : \mathbb{N} \to A$. Nous allons utiliser *le principe d'induction mathématique* , défini de la façon suivante :

Définition 1.1.4. *Si nous avons une liste d'énoncés*

$$P_0, P_1, P_2, P_3, ..., P_n, ...$$

et que nous avons démontré de façon irréfutable que

1 P_0 *est vrai (ce critère est parfois appelé* la condition de base*)*

2 *si* P_n *est vrai, il faut que* P_{n+1} *soit vrai, (appelé* l'énoncé d'induction*),*

on doit conclure que P_n *est vrai, peu importe la valeur de* n*.*

Pour commencer la construction de f, choisissons n'importe quel élément de A pour $f(0)$. Supposons que $f(0), f(1), \ldots f(n)$ ont été choisis. Comme A est infini, il n'est pas égal à l'ensemble $\{f(0), f(1), \ldots f(n)\}$, qui est évidemment un ensemble fini. Donc il y a un élément de A qui n'est pas en $\{f(0), f(1), \ldots f(n)\}$, et nous choisissons un tel élément pour $f(n+1)$.

Notre construction est terminée par le principe d'induction. ★1.1.3

Après avoir découvert les ensembles, Cantor a démontré que certains ensembles infinis dénombrables sont les ensembles d'unicité. Après 1872 il a centré ses recherches autour de la théorie des ensembles et n'a plus travaillé sur les ensembles d'unicité. Son travail a été néanmoins validé par le résultat de W.H. Young qui a démontré en 1909 (voir [KL87]) que les ensembles dénombrables sont des ensembles d'unicité, donc ce « petit infini » a été la bonne généralisation du fini. Les ensembles infinis qui ne sont pas dénombrables (on démontrera qu'il en y a) ne sont pas en général les ensembles d'unicité. La référence standard pour des ensembles d'unicité est le livre [KL87].

Cantor a conçu un ensemble comme une notion déterminée par une description d'appartenance, comme $\{x : \varphi(x)\}$. On verra un peu plus tard que cette description est un peu trop simple, mais que le sens que lui a donné Cantor, qui s'est intéressé la plupart du temps aux sous-ensembles de la droite réelle, était correct. Pour

l'instant nous allons rester avec cette intuition vague, mais déjà en §3.2 nous allons adopter quelques axiomes qui font partie de la théorie axiomatique de ZF(C).

On voit, en consultant la preuve de Théorème 1.1.3, que les ensembles admettent des opérations simples comme la différence et il convient de les nommer. On nomme aussi quelques relations entre les ensembles.

Définition 1.1.5. *Supposons que A et B sont deux ensembles quelconques.*

— La réunion *de A et B est l'ensemble*

$$A \cup B = \{x : x \in A \text{ ou } x \in B\}.$$

— L'intersection *de A et B est l'ensemble*

$$A \cap B = \{x : x \in A \text{ et } x \in B\}.$$

— La différence *entre* A *et* B *est l'ensemble*

$$A \setminus B = \{x : x \in A \text{ et } x \notin B\}.$$

On dit que A *est* un sous-ensemble *de* B *ssi chaque élément de* A *appartient à* B*.*

Dans §3.2 nous allons voir des généralisations de ces notions. Revenons maintenant à la notion d'infini dénombrable.

1.2 Les infinis dénombrables et ceux qui ne sont pas

On voit tout de suite que les ensembles infinis sont très différents des ensembles finis : un ensemble infini peut avoir la même puissance qu'un de ses propres sous-ensembles. Par exemple, on peut voir facilement que les entiers $\mathbb{Z}$ ont la même puissance que les entiers positifs $\mathbb{N}$, en regardant la fonction $2n \mapsto n$ et $2n-1 \mapsto -n$. Il est intéressant de constater qu'un ensemble encore plus grand,

les rationnels $\mathbb{Q}$, a encore la même puissance que $\mathbb{N}$, comme on va le démontrer.

Pour commencer il nous faut prouver un lemme :

Lemme 1.2.1. *(Schroeder-Bernstein) Supposons qu'il y a une injection f entre A et B et qu'il y a aussi une injection g entre B et A. Alors il y a une bijection entre A et B.*

En fait ce lemme était déjà connu de Cantor, mais sa démonstration posait des problèmes. La démonstration que nous allons présenter provient de Denis Koenig dans les années 1930. Ce lemme montre que $|A| \leq |B|$ et $|B| \leq |A|$ implique $|A| = |B|$. La démonstration de Koenig n'est pas optimale, car elle utilise un axiome qui n'est pas nécessaire pour le théorème (l'axiome du choix), mais elle a l'avantage d'être facile à visualiser et à comprendre.

Preuve Notons que la difficulté est qu'on ne sait pas si f ou g sont des surjections, autrement le lemme serait facile.

Pour chaque $a \in A$, définissons *l'orbite* de a, $o(a)$, c'est-à-dire

la suite

$$\ldots, g^{-1}(f^{-1}(g^{-1}(a))), f^{-1}(g^{-1}(a)), g^{-1}(a), a, f(a), g(f(a)), f(g(f(a))), \ldots$$

Cette suite va toujours aller jusqu'à l'infini à droite, mais il peut se passer qu'elle s'arrête à gauche car e.g. $g^{-1}(a)$ n'est pas défini. Si cette suite ne s'arrête jamais, alors f ou g forment une bijection entre $o(a) \cap A$ et $o(a) \cap B$. Si elle s'arrête sur un élément de A, c'est que g^{-1} n'est pas défini sur cet élément. Dans ce cas, f est une bijection entre $o(a) \cap A$ et $o(a) \cap B$. Si la suite s'arrête sur un élément de B, alors c'est g qui est une bijection entre $o(a) \cap A$ et $o(a) \cap B$. Dans tous les cas, il y a une bijection entre $o(a) \cap A$ et $o(a) \cap B$.

Maintenant il faut observer que les orbites des éléments a et a' peuvent être soit le même ensemble, soit disjointes. Cela veut dire que ces orbites tranchent la réunion $A \cup B$ dans des ensembles disjoints tels que leurs parties dans A et dans B sont respectivement bijectives. En plus, la réunion est couverte par ces tranches. En mettant ces bijections ensemble, nous arrivons à une bijection

entre A et B. $\bigstar_{1.2.1}$

Maintenant on peut démontrer :

Théorème 1.2.2. *L'ensemble* $\mathbb{Q}$ *est dénombrable.*

Preuve En utilisant le lemme de Schroeder-Bernstein, nous voyons qu'il nous suffit de démontrer l'existence d'une injection f de $\mathbb{N}$ à $\mathbb{Q}$, et d'une autre injection g de $\mathbb{Q}$ à $\mathbb{N}$. Il est facile de définir f, on prend simplement $f(x) = x$. Pour g on peut utiliser

$$g(p/q) = \begin{cases} 2^p \cdot 3^q \cdot 5 & \text{si } p \geq 0 \\ 2^p \cdot 3^q & \text{si } p < 0 \end{cases}$$

On utilise ici une représentation des rationnels par p/q où p et q sont premiers entre eux et $q > 0$. $\bigstar_{1.2.2}$

Dans les exercices nous allons montrer qu'il y a une quantité d'ensembles infinis dénombrables. On peut se demander si tout ensemble infini est dénombrable. Cela n'est pas le cas, car Cantor a aussi démontré que les réels ne sont pas dénombrables :

Théorème 1.2.3. *L'ensemble* $\mathbb{R}$ *n'est pas dénombrable.*

Preuve Observons que si un ensemble est dénombrable, alors chaque sous-ensemble l'est aussi. Il nous suffit donc de démontrer que l'intervalle $[0,1)$ n'est pas dénombrable. Supposons, pour contradiction, qu'il y a une bijection $f : \mathbb{N} \longrightarrow [0,1)$. On peut alors écrire $[0,1) = \{f(0), f(1), f(2), \ldots\}$. Comme chaque élément de $[0,1)$ a une représentation en terme de $0, a_0 a_1 \ldots$ où $a_0, a_1, \ldots \in \{0, 1, 2, \ldots 9\}$, on va utiliser cela pour écrire $f(n) = 0, a_0^n, a_1^n, a_2^n, \ldots$ pour chaque n.

Maintenant on va définir un nombre b qui est dans $[0,1)$ mais qui n'est pas parmi les $f(n)$, d'où une contradiction. On choisit pour chaque n un chiffre $b_n \notin \{a_n^n, 9\}$ et on définit $b = 0, b_0 b_1 \ldots$.

★1.2.3

Cantor a beaucoup essayé de trouver un sous-ensemble infini de réels qui n'est pas dénombrable mais également qui n'a pas la même puissance que $\mathbb{R}$. Il n'a pas réussi et alors il a formulé :

HC-l'hypothèse du continu. *Chaque sous-ensemble infini de* $\mathbb{R}$ *qui n'est pas dénombrable a la même puissance que* $\mathbb{R}$.

Ce problème a été formulé en 1874. Il est alors devenu clair que

c'était un problème suffisamment difficile pour constituer le premier de la célèbre liste des 23 problèmes établie par David Hilbert (1862-1943) pour le congrès international des mathématiciens de 1900 à Paris, afin de guider la recherche en mathématiques du siècle alors naissant.

Dans ce chapitre nous avons développé certaines notions d'une façon naïve, c'est-à-dire sans formalité axiomatique. C'est la façon adoptée par Cantor et aujourd'hui considérée comme une bonne façon d'aborder le sujet. Les livres [Hal97] et [End77] sont de bonnes références pour cette approche. Dans le chapitre 3 nous allons développer la présentation axiomatique qui peut s'utiliser pour justifier les notions abordées ici. Pour l'instant nous continuons notre investigation des infinis et de leurs tailles.

Outre l'ensemble des réels, d'autres ensembles faciles à définir ne sont pas dénombrables. C'est le cas pour l'ensemble de parties de $\mathbb{N}$.

Définition 1.2.4. *Pour un ensemble B nous écrivons $\mathcal{P}(B)$ pour* l'ensemble de parties *de B, c'est-à-dire $\{B : B \subseteq A\}$.*

Théorème 1.2.5. *Soit A un ensemble quelconque. Alors il n'existe pas une injection de $\mathcal{P}(A)$ à A.*

Preuve Si A est un ensemble vide, l'ensemble $\mathcal{P}(\emptyset)$ est l'ensemble $\{\emptyset\}$, donc A a zéro élément et $\mathcal{P}(A)$ a un élément, alors il ne peut pas y avoir une injection de $\mathcal{P}(A)$ à A. Soit A un ensemble qui n'est pas vide et supposons, au contraire, que $f : \mathcal{P}(A) \to A$ est une injection. Considérons l'ensemble

$$B = \{a \in A : \ f^{-1}(a) \text{ est un ensemble qui ne contient pas } a\}.$$

Donc B est un sous-ensemble de A, c'est-à-dire $B \in \mathcal{P}(A)$. La valeur $f(B)$ est définie et est égale à un élément b de A. Est-que b appartient à B ? Par la définition de B, nous voyons que les deux possibilités $b \in B$ et $b \notin B$ nous donnent une contradiction. Donc f ne peut pas exister. $\bigstar_{1.2.5}$

Corollaire 1.2.6. *L'ensemble $\mathcal{P}(\mathbb{N})$ n'est pas dénombrable.*

Preuve Autrement, il y aurait une bijection, donc une injection, de $\mathcal{P}(\mathbb{N})$ à $\mathbb{N}$, contradiction. (Il faut aussi noter que $\mathcal{P}(\mathbb{N})$ n'est

pas fini, ce qui est facile à voir en utilisant les ensembles $\{n\}$ pour $n \in \mathbb{N}$). $\bigstar_{1.2.6}$

Le lemme de Schroeder-Bernstein nous donne un ordre linéaire entre les puissances infinies, car étant donné que $|A| \leq |B|$ quand il existe une injection de A à B, le lemme peut se traduire en disant que $|A| \leq |B|$ et $|B| \leq |A|$ impliquent $|A| = |B|$. Comme la notion d'injection est fermée pour la composition, nous avons aussi la transitivité : $|A| \leq |B|$ et $|B| \leq |C|$ impliquent $|A| \leq |B|$. Bien sûr, nous avons aussi $|A| \leq |A|$ pour chaque A, par l'identité. Donc nous pouvons imaginer les tailles infinies sur une ligne, en commençant par le plus petit élément $|\mathbb{N}|$. Théorème 1.2.5 dit que cette ligne n'a pas l'élément le plus grand : car si un certain $|A|$ est le plus grand, néanmoins $|\mathcal{P}(A)|$ est encore plus grand !

1.3 Exercices et points de réflexion

Dans les exercices nous allons utiliser quelques notions qui n'ont pas encore été définies, dont :

Définition 1.3.1. *Le* produit cartésien *ou simplement le* produit *de deux ensembles A et B est donné par*

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

Si $A = B$, on écrit A^2 à la place de $A \times A$. Cette définition peut se prolonger par induction en définissant $A^{n+1} = A^n \times A$, pour $n \in \mathbb{N} \setminus \{0\}$. On déclare $A^1 = A$.

On peut de la même façon définir le produit d'un nombre fini d'ensembles quelconques $A_0 \times A_1 \times \ldots A_n$.

Définition 1.3.2. *Un nombre réel est* algébrique *s'il est la solution d'une équation polynomiale de forme*

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0,$$

où $a_0, a_1, \ldots a_n$ sont rationnels et au moins un parmi eux est non-zéro. Un nombre qui n'est pas algébrique est appelé transcendant.

1.3.1 À vous de jouer

Démontrez les énoncés suivants et répondez aux questions posées.

1 Tout sous-ensemble d'un ensemble dénombrable est lui-même dénombrable.

2 L'ensemble $\mathbb{Z}$ d'entiers est dénombrable.

3 $\mathbb{N} \times \mathbb{N}$ est dénombrable.

4 Le produit d'un nombre fini d'ensembles dénombrables est dénombrable.

5 Chaque nombre transcendant est irrationnel. Donnez un exemple d'un nombre irrationnel algébrique. Essayez de démontrez l'existence d'un nombre transcendant sans utiliser la théorie des ensembles.

6

Théorème 1.3.3 (Cantor 1874)**.** *L'ensemble des nombres transcendantaux n'est pas dénombrable.*

7 Qui était Nicolas Bourbaki ? Que pensez-vous de son modèle de travail ? À quoi sert de toujours nommer l'auteur d'un texte ?

1.4 Éléments historiques et bibliographiques

Pour une approche philosophique du fini et de l'infini, on peut consulter une conférence de vulgarisation d'Alain Badiou [Bad10], dans laquelle est également exposée une différence philosophique entre « l'infini virtuel » et « l'infini actuel ». En gros, l'infini virtuel est le fait de savoir qu'il n'y a pas un nombre fini qui soit le plus grand et qu'on approche l'infini par une suite croissante des nombres finis, sans pourtant jamais toucher à l'infini *actuel*. C'était l'idée d'infini telle que la concevait les philosophes et les mathématiciens avant Cantor. L'idée de Cantor de nommer cet infini, de l'actualiser, de le traiter comme un objet existant, l'infini actuel, était très révolutionnaire et elle a causé beaucoup de chagrin à Cantor lui-même et à d'autres. Étant un homme très religieux, Cantor se demandait s'il avait le *droit* de discuter de l'infini actuel. Badiou cite même une lettre que Cantor a écrite au Pape pour lui demander la permission d'étudier cet infini. Cantor était très admiré par des mathématiciens tels que Dedekind, Mittag-Lefler et Hilbert,

mais il a rencontré aussi beaucoup de résistance, en particulier de la part de Kroenecker qui a mené une vraie guerre contre lui. Une biographie passionnante de Cantor a été écrite par George Dauben [Dau79]. Une connexion entre la théorie des ensembles de Cantor et une histoire vraie du mysticisme religieux est exposée dans [KG10].

L'existence d'un nombre transcendant a été démontrée par Liouville en 1844, qui a construit un nombre nommé **c** qui est transcendant par construction. Les exemples les plus connus de transcendantaux sont les deux constantes favorites des mathématiciens : e et π. C'est Hermite qui a démontré en 1873 que e est transcendant et il a fallu attendre 1882 et von Lindemann pour savoir que π aussi est transcendant. Cela a résolu une question très ancienne : savoir s'il est possible, en n'utilisant que les opérations de la géométrie classique, de trouver un carré de la même surface qu'un cercle donné. La réponse est négative, car π n'est pas algébrique et, en gros, les opérations géométriques ne produisent que les nombres algébriques. Donc le théorème de Cantor de 1874 montre, alors qu'on ne connaissait que quelques nombres transcendantaux, que la plu-

part des réels sont transcendantaux ! C'est là la différence entre une preuve d'existence d'un objet et la possibilité de le construire, qui va devenir l'objet des études profondes en mathématiques et en informatique au vingtième siècle. Les nombres transcendantaux, quant à eux, sont toujours très mystérieux et c'est même une question ouverte en ce moment de savoir si $\pi + e$ est transcendant.

Malheureusement la littérature sur la théorie des ensembles en français n'est pas abondante. Il y a bien sûr l'ouvrage de Bourbaki [Bou70], très bien écrit, mais un peu daté. Un très bel ouvrage est celui de Jean-Louis Krivine [Kri09] (voir aussi [Kri07]), qui s'adresse surtout aux étudiants en mathématiques.

Pour ce qui est de la question de la terminologie, nous notons qu'il y a des auteurs, par exemple Krivine [Kri09], qui prennent la notion de dénombrable pour signifier seulement les ensembles infinis bijectifs avec $\mathbb{N}$, pas des ensembles finis, comme il y a en d'autres qui utilisent la terminologie utilisée ici. Un autre point de différence est que parfois le nombre 0 n'est pas inclus dans l'ensemble $\mathbb{N}$. Ces petites différences sont purement une question de présentation et

ne changent pas l'essentiel.

Chapitre 2

Les cardinaux et les ordinaux

Il est possible de découvrir les cardinaux et les ordinaux tout en restant dans le monde intuitif de Cantor, même si nous allons revisiter ces notions une fois que l'axiomatique est à notre disposition.

2.1 Cardinalités

La définition de cardinalité selon notre approche jusqu'ici est plutôt celle de Gotlob Frege, le grand logicien allemand, presque contemporain de Cantor (1848-1925). Son œuvre est énorme et il est

considéré comme l'un des plus grands logiciens de tous les temps. La partie de son œuvre la plus pertinente ici est son livre *Fondements de l'arithmétique*, où il tente de dériver l'arithmétique de la logique. Il se trouve que cette tentative a connu le même problème que la théorie des ensembles naïve, ce qu'on va discuter dans §3.1. Restons pour le moment dans le paradis de l'ignorance !

Donc, comme on a déjà commencé à le faire, une première théorie de la cardinalité peut se construire comme une théorie de la relation d'équipotence, sans définir ce qu'est vraiment un nombre cardinal en toute généralité. La relation d'équipotence étant réflexive, symétrique et transitive sur la classe des ensembles [1], chaque classe d'équivalence est appelée nombre cardinal ou simplement cardinal. Par exemple, tout entier positif est un cardinal, représentant sa taille, et zéro aussi est un cardinal représentant la taille de l'ensemble sans éléments.

En partant de cette définition, on va définir $\aleph_0$ comme la cardinalité de $\mathbb{N}$. Cela veut dire que $\aleph_0$ est la cardinalité de chaque

1. C'est exactement l'idée de la classe d'ensembles qu'il faudrait revoir pour éviter les paradoxes.

ensemble infini dénombrable. Théorème 1.2.3 ou Corollaire 1.2.6 montrent qu'il y a d'autres cardinalités infinies à part $\aleph_0$. Théorème 1.2.5 montre que pour chaque cardinalité il y en a une plus grande. La cardinalité de $\mathbb{R}$ est appelée *le continu* et est notée $\mathfrak{c}$. Il est très important de remarquer que les cardinaux eux-mêmes sont des ensembles, car ils sont des représentants des classes d'équivalence des ensembles.

2.2 Opérations avec les cardinalités

Si on veut penser des cardinaux en terme de nombres, il faut chercher à voir s'il y a une arithmétique de ces nombres. Un des défis est que, comme les entiers sont des cardinaux, il faut que cette arithmétique, quand elle est restreinte sur les entiers, retrouve les opérations habituelles. Il se trouve qu'on peut développer une telle arithmétique, mais il y a des surprises. Commençons avec l'addition.

Si on pense à l'addition des entiers, on peut la définir en disant

que pour deux entiers m et n représentant les tailles de deux ensembles A et B, on a $m+n$ comme taille de l'ensemble $A \cup B$, sauf que pour cela il faut bien éviter qu'il y ait une intersection entre A et B, car autrement le calcul ne serait pas correct. Cela ne pose pas de problème car on peut toujours choisir deux ensembles disjoints pour représenter m et n, comme un ensemble A de m poires et un ensemble B de n pommes, donc $m+n$ représente le nombre de fruits dans l'ensemble $A \cup B$. On va utiliser exactement cette définition pour la somme $\kappa + \lambda$ en général. Pour réussir avec la demande d'utilisation des ensembles disjoints, on va prendre pour des analogues des pommes et des poires les ensembles $\kappa \times \{0\}$ et $\lambda \times \{1\}$, qui sont disjoints car le premier ne contient que des paires de la forme $(x, 0)$, et le deuxième que des paires de la forme $(y, 1)$.

Définition 2.2.1. *Soit κ et λ des cardinaux. On définit $\kappa + \lambda$ comme étant la taille de $\kappa \times \{0\} \cup \lambda \times \{1\}$.*

L'addition de cardinaux est commutative car $\lambda + \kappa$ par définition est la taille de $\lambda \times \{0\} \cup \kappa \times \{1\}$, qui est en bijection avec $\kappa \times \{0\} \cup \lambda \times \{1\}$, défini comme étant $\kappa + \lambda$. On peut également

voir que cette addition est associative. Et voici qu'arrive la première surprise :

Théorème 2.2.2. *Pour κ et λ cardinaux tels qu'au moins l'un d'eux est infini, on a* $\kappa + \lambda = \max\{\kappa, \lambda\}$.

Pour démontrer ce théorème, il nous serait plus facile d'utiliser la représentation correcte de cardinaux comme « ordinaux initiaux ». Comme pour l'instant nous n'avons pas la notion d'ordinal, nous allons réserver la preuve pour le moment où nous aurons étudié les ordinaux. Nous allons maintenant aborder cette question.

2.3 Les ensembles bien ordonnés

Une des notions les plus importantes dans la théorie des ensembles est celle du *bon ordre*, et c'est grâce à cette notion que nous pourrons éventuellement définir les ordinaux.

Définition 2.3.1. *Un ensemble ordonné $(A, \leq)$ est* bien ordonné par $\leq$ *et la relation $\leq$ est un* bon ordre *si la condition suivante est satisfaite :*

Toute partie non vide de A possède un plus petit élément par rapport à $\leq$. Nous écrivons $<$ pour l'ordre strict induit par $\leq$.

Il est facile de voir que chaque ordre fini est un bon ordre et, aussi, que l'ensemble des entiers naturels muni de l'ordre habituel est bien ordonné. On note cet ensemble avec cet ordre ω, qui est à peu près la même chose que $\mathbb{N}$, mais la notation surligne l'importance d'ordre et pas l'importance d'autres propriétés de $\mathbb{N}$, tel que l'addition, par exemple. Chaque sous-ensemble d'un ensemble bien ordonné est bien ordonné par la relation héritée. On peut également voir que les réels ou les rationnels munis de l'ordre habituel ne sont pas bien ordonnés, car même ces ensembles eux-mêmes n'ont pas d'éléments minimaux. La proposition suivante est très utile pour caractériser les bons ordres. Même si nous n'avons pas encore abordé l'axiomatique, nous pouvons remarquer que la preuve de Proposition 2.3.2 nécessite une forme faible de l'axiome du choix (l'axiome du choix dépendant).

Proposition 2.3.2. *Un ensemble ordonné $(A, \leq)$ est* bien ordonné par $\leq$ *ss'il n'existe pas une suite infinie strictement décroissante*

en $(A, \leq)$.

Preuve Si $x_0 > x_1 > \ldots$ est une suite infinie strictement décroissante en $(A, \leq)$, alors l'ensemble $\{x_n : n < \omega\}$ n'a pas un plus petit élément par rapport à $\leq$. Dans l'autre direction, si un ensemble ordonné $(A, \leq)$ n'est pas bien ordonné, il existe alors un sous-ensemble non vide A_0 de A qui n'a pas de plus petit élément. Choisissons un élément $x_0 \in A_0$. Comme il n'est pas le plus petit, nous pouvons trouver $x_1 \in A_0$ tel que $x_0 > x_1$. Continuant de la même façon, nous pouvons produire une suite infinie strictement décroissante $x_0 > x_1 > \ldots$. $\bigstar_{2.3.2}$

La notion par laquelle nous comparons les ordres est celle d'isomorphisme d'ordres.

Définition 2.3.3. *Deux ensembles ordonnés* $(A, \leq_A)$ *et* $(B, \leq_B)$ *sont* isomorphes *s'il existe une bijection* $f : A \to B$ *telle que*

$$\forall x, y \in A\,(x \leq_A y \iff f(x) \leq_B f(y)).$$

La propriété la plus importante des bons ordres est que, quels

que soient les deux que l'on choisisse, ils sont comparables par l'isomorphisme. Pour comprendre ce que cela veut dire, il nous faut la notion de l'ensemble de *précédents*, défini par

$$\mathrm{pre}_A(x) = \{y \in A : y <_A x\},$$

pour un élément x d'un ensemble ordonné $(A, \leq_A)$. Le théorème monumental de Cantor sur ce sujet est le « Théorème de la Trichotomie » :

Théorème 2.3.4. *Étant donnés deux ensembles bien ordonnés,* $(A, <_A)$ *et* $(B, <_B)$ *:*

- *soit* $(A, <_A)$ *est isomorphe à* $(B, <_B)$*,*
- *soit* $(A, < A)$ *est isomorphe à une section commençante propre de* $(B, <_B)$*, donc un ensemble de la forme* $\mathrm{pre}_B(x)$*,*
- *soit* $(B, <_B)$ *est isomorphe à une section commençante propre de* $(A, <_A)$*, donc un ensemble de la forme* $\mathrm{pre}_A(x)$*,*

et ces trois cas sont exclusifs.

Avant de commencer la preuve du Théorème 2.3.4, notons d'abord

que la supposition que $(A, <_A)$ et $(B, <_B)$ sont tous les deux bien ordonnés est essentielle. Car, en prenant pour $(A, <_A)$ l'ensemble $\mathbb{N}$ avec l'ordre habituel et pour $(B, <_B)$ l'ensemble $\mathbb{Z}$ avec l'ordre habituel, nous voyons qu'aucun cas de la trichotomie n'est satisfait. Supposer qu'au moins entre $(A, <_A)$ et $(B, <_B)$ est un bon ordre n'est pas suffisant pour la trichotomie. Pour la preuve de la trichotomie, nous avons besoin de quelques lemmes.

Lemme 2.3.5. *Si $(A, <_A)$ est un bon ordre, alors il n'y a aucun $x \in A$ tel que $(A, <_A)$ est isomorphe à $(\mathrm{pre}_A(x), <_A)$.*

Preuve Supposons que $f : A \to \mathrm{pre}_A(x)$ est un isomorphisme, et soit $y = f(x)$. Donc $y \in \mathrm{pre}_A(x)$ et alors $y <_A x$. En appliquant f, nous avons $f(y) <_A f(x) = y$, et en continuant nous obtiendrons que $y >_A f(y) >_A f(f(y)) \ldots$ est une suite infinie décroissante. C'est une contradiction car $(A, <_A)$ est un bon ordre. $\bigstar_{2.3.5}$

Lemme 2.3.6. *Si $(A, <_A)$ et $(B, <_B)$ sont deux bons ordres isomorphes, alors l'isomorphisme entre eux est unique.*

Preuve Supposons que $f : A \to B$ et $g : A \to B$ sont deux

isomorphismes entre $(A, <_A)$ et $(B, <_B)$, mais que $f \neq g$. Donc $\{x \in A : f(x) \neq g(x)\} \neq \emptyset$ et il existe alors a l'élément le plus petit de cet ensemble, disons x^*. Nous avons $f(x^*) \neq g(x^*)$ et donc, ou bien $f(x^*) <_B g(x^*)$, ou bien $g(x^*) <_B f(x^*)$. Disons $f(x^*) <_B g(x^*)$, l'autre cas étant symétrique. Soit $y^* = f(x^*)$. Comme g est un isomorphisme, il y a $x' \in A$ tel que $g(x') = y^*$. Donc $g(x') = y^* <_B g(x^*)$ et alors $x' <_A x^*$. Par le choix de x^*, nous avons que $f(x') = g(x') = f(x^*)$, contradiction. $\bigstar_{2.3.6}$

Lemme 2.3.7. *Si $(A, <_A)$ est un bon ordre et $y, y' \in A$ sont tels que les ensembles $(\mathrm{pre}_A(y), <_A)$ et $(\mathrm{pre}_A(y'), <_A)$ sont isomorphes, alors $y = y'$.*

Preuve Autrement, supposons que $y \neq y'$, donc soit $y <_A y'$, soit $y' <_A y$. Disons $y' <_A y$, car l'autre cas est symétrique. Alors $y' \in \mathrm{pre}_A(y)$. Remarquons que, en étant un sous-ensemble d'un bon ordre, $\mathrm{pre}_A(y)$ est bien ordonné par $<_A$. Aussi, $\mathrm{pre}_{\mathrm{pre}_A(y)}(y') = \mathrm{pre}_A(y')$ et alors, $\mathrm{pre}_{\mathrm{pre}_A(y)}(y')$ est isomorphe avec $\mathrm{pre}_A(y)$, en contradiction avec Lemme 2.3.6. $\bigstar_{2.3.7}$

Pour la preuve suivante nous allons devoir définir ce qu'on com-

prend par une fonction dans la théorie des ensembles. On utilise l'identification des fonctions avec leurs graphes, alors une fonction entre A et B est un ensemble des couples $(x, y) \in A \times B$ tels que pour chaque $x \in A$ il y a exactement un $y \in B$ tel que $(x, y) \in f$.

Preuve (de la Trichotomie) Nous allons supposer que les ensembles A et B sont tous les deux non vides, car autrement la conclusion est évidente. Soit

$$f = \{(x, y) : x \in A, y \in B \text{ et } \mathrm{pre}_A(x) \text{ et } \mathrm{pre}_B(y) \text{ sont isomorphes}\}.$$

Définissons $\mathrm{dom}(f) = \{x \in A : (\exists y \in B)(x, y) \in f\}$, $\mathrm{ran}(f) = \{y \in B : (\exists x \in A)(x, y) \in f\}$. Nous allons démontrer que f est une fonction entre $\mathrm{dom}(f) \subseteq A$ et $\mathrm{ran}(f) \subseteq B$. Pour ce faire, supposons qu'il y a un $x \in A$ tel que pour $y \neq y' \in B$ nous avons $(x, y), (x, y') \in f$. Donc, $\mathrm{pre}_B(y)$ et $\mathrm{pre}_B(y')$ sont tous les deux isomorphes à $\mathrm{pre}_A(x)$ et alors isomorphes entre eux. En utilisant Lemme 2.3.7, nous avons $y = y'$, contradiction.

Maintenant nous allons démontrer que f est un isomorphisme

entre $\mathrm{dom}(f)$ et $\mathrm{ran}(f)$. C'est une surjection par la définition de $\mathrm{ran}(f)$. Pour voir qu'il s'agit d'une injection, supposons que pour un $x \in \mathrm{dom}(f)$ il existe $y, y' \in B$ tels que $(x, y), (x, y') \in f$. Donc $\mathrm{pre}_A(x)$ et $\mathrm{pre}_B(y)$ sont isomorphes et aussi $\mathrm{pre}_A(x')$ et $\mathrm{pre}_B(y')$ sont isomorphes. En composant ces isomorphismes nous avons que $\mathrm{pre}_B(y)$ et $\mathrm{pre}_B(y')$ sont isomorphes, donc $y = y'$ par Lemme 2.3.7. Finalement, pour voir que f est un isomorphisme, supposons que $x <_A x'$ sont deux éléments de $\mathrm{dom}(f)$ et montrons que $y = f(x) <_B y' = f(x')$. Soit g l'isomorphisme entre $\mathrm{pre}_A(x')$ et $\mathrm{pre}_B(y')$ et soit $h = g \restriction \mathrm{pre}_A(x)$. On peut facilement vérifier que h est un isomorphisme entre $\mathrm{pre}_A(x)$ et $\mathrm{pre}_B(g(x))$, et donc $(x, g(x)) \in f$. Par l'argument utilisé pour montrer que f est une fonction, nous avons que $g(x) = f(x) = y$ et donc, $y \in \mathrm{pre}_B(y')$, ce qu'il fallait démontrer.

Si $\mathrm{dom}(f) = A$ et $\mathrm{ran}(f) = B$, nous avons trouvé un isomorphisme entre A et B. Supposons que $\mathrm{dom}(f) \neq A$ et soit x le $<_A$-plus petit élément de $A \setminus \mathrm{dom}(f)$. Si $\mathrm{ran}(f) \neq B$, nous pouvons choisir le $<_B$-plus petit élément y de $A \setminus \mathrm{ran}(f)$ et dans ce

cas nous voyons que f montre que $(x, y) \in f$, contradiction. Donc $\text{ran}(f) = B$ et f montre qu'il y a un isomorphisme entre $\text{pre}_A(x)$ et B. Nous pouvons voir de la même façon que si $\text{ran}(f) \neq B$, f démontre que nous sommes dans le troisième cas de la trichotomie. Il est également facile de voir que les trois cas de la trichotomie sont exclusifs. $\bigstar_{2.3.4}$

2.4 Les ordinaux

Les ordinaux sont la vraie colonne vertébrale de la théorie des ensembles. Ils font la base de la hiérarchie cumulative des ensembles, qui est un concept que nous allons aborder plus tard. Pour l'instant, présentons les ordinaux.

Définition 2.4.1. *(1) Un ensemble x est* transitif *si pour chaque $y \in x$, nous avons $y \subseteq x$. Autrement dit, si $y \in x$ et $z \in y$, alors $z \in x$.*

(2) Un ordinal *est un ensemble transitif bien ordonné par $\in$.*

Quelques exemples d'ordinaux sont $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$. Nous

pouvons continuer à construire de plus et plus d'ordinaux en utilisant la méthode par laquelle nous avons construit les trois exemples précédents. En théorie des ensembles les nombres naturels sont définis à partir de cette idée, en utilisant la récurrence.

Définition 2.4.2. *Le nombre 0 est identifié avec* $\emptyset$*. Pour chaque* n*, déjà défini comme nombre naturel, nous définissons* $n+1 = n \cup \{n\}$*.*

Définition 2.4.2 implique que pour chaque nombre naturel $n \geq 1$, nous avons que $n = \{0, 1, 2, \ldots n-1\}$. On définit ω comme l'ensemble des nombres naturels et on peut facilement vérifier que ω est un ordinal. Cette définition de ω concorde avec la définition donnée précédemment. On peut continuer cette définition (et le lecteur est invité à réfléchir sur les axiomes qui nous permettent de faire ces définitions), de la façon suivante :

Définition 2.4.3. *Pour un ordinal* α*, définissons* $\alpha+1 = \alpha \cup \{\alpha\}$*. Un ordinal* β *est appelé* successeur *s'il y a* α *tel que* $\beta = \alpha + 1$*. Les ordinaux qui ne sont pas successeurs sont appelés* limites*.*

Pour un ensemble A d'ordinaux, définissons $\sup(A) = \bigcup A$. *Si α et β sont des ordinaux, écrivons $\alpha < \beta$ si $\alpha \in \beta$.*

Théorème 2.4.4. *(1) Si x est un ordinal et $y \in x$, alors y est un ordinal.*

(2) Si deux ordinaux x et y sont isomorphes, alors ils sont égaux, c'est-à-dire qu'ils ont les mêmes éléments.

(3) Pour deux ordinaux x et y, soit $x < y$, soit $x = y$ ou soit $x > y$.

(4) Si x, y, z sont des ordinaux, $x \leq y$ et $y \leq z$, alors $x \leq z$.

(5) Si C est une famille quelconque d'ordinaux, si C n'est pas vide, alors cette famille a un plus petit élément.

La preuve de Théorème 2.4.4 est proposée comme un exercice. Ce théorème montre qu'entre eux les ordinaux sont bien ordonnés, même si cette phrase invite à une réflexion plus profonde reliée au paradoxe de Buralli-Forti. Nous aborderons ce point dans la section d'axiomatique.

Les ordinaux sont en fait un code pour tous les bons ordres, comme le montre Théorème 2.4.5. Nous allons réserver la preuve

de ce théorème pour le moment où nous aurons présenté la récursion sur les bons ordres, ce qui nous sera plus facile avec une meilleure compréhension des axiomes au Chapitre 3.

Théorème 2.4.5. *Soit $(A, <_A)$ un bon ordre. Il y a exactement un ordinal α tel que $(A, <_A)$ est isomorphe avec α.*

Définition 2.4.6. *Si $(A, <_A)$ est un bon ordre,* le type d'ordre de *$(A, <_A)$ est défini comme l'unique ordinal α tel que $(A, <_A)$ est isomorphe avec α. On écrit $\alpha = \mathrm{otp}((A, <_A))$.*

2.5 Les cardinaux

La définition moderne des cardinaux est basée sur les ordinaux.

Définition 2.5.1. *Un ordinal α est un* cardinal *s'il n'est bijectif avec aucun $\beta < \alpha$.*

Les nombres finis sont des cardinaux et ω est un cardinal. C'est le plus petit cardinal qui soit infini, c'est la cardinalité des ensembles dénombrables infinis. C'est $\aleph_0$! En définissant $\aleph_0$ de cette

façon, nous avons évité les difficultés de la définition des cardinaux comme représentants des classes d'equivalance des ensembles, qui apparaissait dans l'approche de Cantor.

La majorité des ordinaux ne sont pas des cardinaux, par exemple $\omega + 1$ ou $\omega + 1 + 1$. Comme pour chaque α infini il y a toujours une bijection entre α et $\alpha + 1$, un cardinal ne peut pas être un ordinal successif.

Maintenant, après avoir donné la définition correcte des cardinaux, nous allons démontrer Théorème 2.2.2.

Preuve (du Théorème 2.2.2) Supposons que κ et λ sont deux cardinaux, disons $\kappa \leq \lambda$. En supposant que λ est infini, nous devons démontrer que $\kappa + \lambda = \lambda$. Pour ce faire, nous allons utiliser un lemme.

Lemme 2.5.2. *Soit θ un cardinal infini. Alors, θ peut être décomposé en une réunion de deux ensembles disjoints, chacun de puissance θ.*

Preuve (du Lemme 2.5.2) Cette preuve se déroule par récurrence sur les cardinaux et ordinaux, qui est une procédure de légi-

timité dont nous allons discuter au Chapitre 3.

La décomposition de ω dans l'ensemble de nombres pairs et l'ensemble de nombres impairs nous montre que le théorème est vrai pour $\theta = \aleph_0$. Supposons qu'il est vrai pour tous les cardinaux infinis $\rho < \theta$ et démontrons qu'il est vrai pour θ. Nous allons définir deux sous-ensembles, A_0 et A_1 de θ. Soit $0 \in A_0$. Pour chaque α qui est en A_0, soit $\alpha+1 \in A_1$. Pour chaque α qui est en A_1, soit $\alpha + 1 \in A_0$. Si un ordinal $\delta < \theta$ est limite, nous mettons $\delta \in A_0$. Il est clair que $A_0, A_1 \subseteq \theta$, (sachant que θ ne peut pas s'écrire sous la forme $\alpha + 1$ pour $\alpha < \theta$), $A_0 \cup A_1 = \theta$ et que A_0 et A_1 sont disjoints. En définissant $\alpha \mapsto \alpha + 1$ nous obtiendrons une injection de A_0 à A_1, et la même formule nous donne une injection de A_1 à A_0. Donc, par le Lemme de Schroeder-Bernstein, nous avons que $|A_0| = |A_1|$. Soit cette cardinalité égale à ρ. Donc $\rho \leq \theta$ et il est clair que ρ doit être infini. Si $\rho < \theta$, nous avons supposé que $\rho + \rho = \rho$ et donc la cardinalité de $\theta = A_0 \cup A_1$ est aussi $\rho < \theta$, contradiction. Alors $\rho = \theta$ et le lemme est démontré.

$\bigstar_{2.5.2}$

Utilisons maintenant Lemme 2.5.2 pour décomposer λ en réunion de deux ensembles disjoints A_0 et A_1, chacun de puissance λ. Représentons $\kappa + \lambda$ comme la puissance de $\kappa \times \{0\} \cup \lambda \times \{1\}$. Il est évident que l'application $\alpha \mapsto (\alpha, 1)$ est une injection de λ à $\kappa + \lambda$. Dans l'autre direction, nous pouvons facilement injecter $\kappa \times \{0\}$ dans A_0 et $\lambda \times \{1\}$ dans A_1. Donc le Lemme de Schroeder-Bernstein nous donne que $\kappa + \lambda = \lambda$. $\bigstar_{2.2.2}$

Il y a d'autres opérations qu'on peut faire avec des ordinaux et des cardinaux. Nous allons les aborder au Chapitre 3, après avoir adopté l'approche axiomatique.

2.5.1 À vous de jouer

Dans ces exercices, sans vous en rendre compte, vous allez utiliser plusieurs fois les axiomes de ZFC. Une fois que vous aurez appris le sujet du Chapitre 3, il sera utile de revoir ces exercices et de vérifier où chaque axiome a été utilisé.

1 Trouvez un bon ordre sur l'ensemble de nombres naturels qui le rend isomorphe avec $\omega + \omega$.

2 Démontrez que pour chaque ordinal α qui est dénombrable et infini, il y a un bon ordre sur $(\omega, <_\alpha)$ qui est isomorphe à α.

3 Démontrez Théorème 2.4.4.

4 Démontrez l'associativité de l'addition de cardinaux, $(\kappa + \lambda) + \theta = \kappa + (\lambda + \theta)$ pour chaque κ, λ et θ.

5 Supposez qu'un ensemble A peut être bien ordonné. Démontrez que le plus petit α tel qu'il existe un bon ordre de A avec le type d'ordre égal à α, est un cardinal.

2.6 Éléments historiques et bibliographiques

La définition des ordinaux par Cantor est basée sur les classes d'équivalence d'ensembles ordonnés. Elle ne se formalise pas dans la théorie de ZFC, car elle utilise la *classe propre* des bons ordres, ce qu'on peut voir par le paradoxe de Buralli-Forti 3.4. La définition moderne des ordinaux est due à John von Neumann. Avant lui, Zermelo dans son article fondamental [Zer08] a donné une autre définition des ordinaux qui est formalisable dans la théorie ZFC, mais qui présente le désavantage par rapport à la définition de von Neumann de ne pas pouvoir définir les cardinaux comme les ordinaux initiaux.

Pour ce qui est de la terminologie, les bons ordres sont souvent nommés *beaux ordres*.

Chapitre 3

Les axiomes

3.1 Intuitive versus axiomatique

Du seul fait que Cantor ait réussi à résoudre de grands problèmes en mathématiques en utilisant sa méthode des ensembles, on se doit une réflexion sur la notion d'ensemble. Qu'est-ce que c'est exactement, un ensemble ? Pour Cantor la réponse était simple : un ensemble est constitué par tous les objets qui vérifient une certaine propriété, $\{x : \varphi(x)\}$. Par exemple, l'ensemble des entiers positifs divisibles par 3, ou l'ensemble de tous les nombres transcendantaux.

Cette représentation, appelée *la compréhension sans restriction*, est assez intuitive, mais elle n'est pas aussi précise qu'on l'aimerait. Notamment, ce qu'on veut dire par une « propriété » n'est pas clair. Cette approche a eu un grand revers par la découverte d'une antinomie par Bertrand Russell, en 1901.

On peut formuler le paradoxe ainsi : soit $U = \{x : x = x\}$, donc U est la collection de tous les ensembles (ou toutes les choses !), car chaque ensemble x vérifie que $x = x$. En utilisant $\varphi(x) \equiv x = x$, on voit que U se présente sur la forme $\{x : \varphi(x)\}$, donc, selon Cantor, U est un ensemble. Maintenant regardons l'ensemble $B = \{x \in U : x \notin x\}$. Là aussi, on peut facilement montrer que B est un ensemble. Le paradoxe viendra lorsqu'on se demandera si $B \in B$. Si on répond oui, alors, comme par définition les membres de B n'appartiennent pas à eux-mêmes, il n'appartient pas à lui-même : contradiction. Mais si on répond non, alors il a la propriété requise pour appartenir à lui-même : contradiction de nouveau. On a donc une contradiction dans les deux cas, ce qui rend l'existence d'un tel ensemble paradoxale.

On a déjà évoqué le paradoxe de Russell et son effet catastrophique sur l'œuvre monumentale de Frege, *Fondements de l'arithmétique*. En effet, l'ouvrage venait d'être achevé en 1902 quand Frege reçut une lettre de Russell où le paradoxe était expliqué, ce qui lui montrait que l'une des règles introduites, la compréhension non restreinte, rendait sa théorie contradictoire. En 1903 Frege adjoignit au second volume de *Fondements de l'arithmétique* un appendice où il exposait le paradoxe en faisant précéder l'analyse de cet aveu d'une honnêteté confondante : « Pour un écrivain scientifique, il est peu d'infortunes pires que de voir l'une des fondations de son travail s'effondrer alors que celui-ci s'achève. C'est dans cette situation inconfortable que m'a mis une lettre de M. Bertrand Russell, alors que le présent volume allait paraître ».

Comment peut-on résoudre le paradoxe ? Évidemment la supposition qui nous a amené au paradoxe est que U est un ensemble. Donc, pour résoudre le paradoxe, il faut avoir une approche plus soignée de la façon dont nous formons les ensembles. Il faut avoir une axiomatique où la compréhension sans restriction n'est pas

permise. Ce fut le travail de logiciens au début du 20e siècle, notamment Zermelo, de trouver une telle axiomatique.

3.2 La théorie axiomatique

Maintenant nous allons (re)-développer la théorie des ensembles en utilisant une approche formelle, que nous appelons la théorie axiomatique. Dans notre théorie axiomatique, nous allons nous exprimer en logique du premier ordre en utilisant le langage $\mathcal{L} = \{\in\}$, où $\in$ est une relation binaire, interprétée par l'appartenance. Nous allons également utiliser l'égalité qui est toujours censée représenter la vraie égalité. Dans cette approche les variables sont donc censées représenter les ensembles et *on ne demande pas* la définition d'un ensemble, juste comme, par exemple, en géométrie on ne demande pas la définition d'un point. Un ensemble pour nous est une notion de base, prise comme étant donnée. [1]

Nous commençons par deux axiomes qui sont évidents.

1. Ce n'est pas du tout l'approche de Cantor qui a essayé (sans succès évidemment) de donner des définitions de plus et plus complexes d'un ensemble.

L'axiome d'Existence d'Ensembles Il existe un ensemble.

Cet axiome est parfois pris comme une conséquence de l'utilisation d'une formalisation logique car il veut simplement dire que le domaine du discours n'est pas vide. L'axiome suivant désigne la relation entre les ensembles et l'appartenance. Il implique qu'un ensemble est complètement déterminé par ces éléments.

L'axiome d' Extensionalité Chaque ensemble est déterminé par ses éléments, ou

$$\forall A \forall B[\forall x(x \in A \iff x \in B) \implies A = B].$$

Cela revient à dire que : si tout élément de l'ensemble A est aussi un élément de B, et inversement, tout élément de l'ensemble B appartient à l'ensemble A, alors les deux ensembles A et B sont égaux.

Une fois qu'un ensemble est donné, nous allons adopter une série d'axiomes qui vont nous permettre de construire de nouveaux ensembles à partir des ensembles déjà connus. C'est ici qu'intervient

l'idée de Zermelo qui a résolu le paradoxe de Russell. Ce qu'on appelle la Compréhension Étendue veut dire que pour n'importe quelle propriété φ des ensembles, on peut former l'ensemble de tous les ensembles qui vérifient φ, c'est-à-dire $\{x : \varphi(x)\}$. Comme nous l'avons déjà vu, cela donne une contradiction, montrée par le paradoxe de Russell. L'idée de Zermelo introduit une Compréhension Restreinte : pour chaque φ *et* chaque ensemble A, on peut former l'ensemble de tous les éléments de A qui vérifient φ.

Schéma d'axiomes de Compréhension Pour chaque énoncé φ de $\mathcal{L}$ et chaque ensemble A, la collection $\{x \in A : \varphi(x)\}$ est un ensemble, ou, formellement, pour chaque énoncé φ avec les variables libres parmi $x, z, w_1, \ldots, w_n$,

$$(\forall z)(\forall w_1) \ldots (\forall w_n)(\exists y)(\forall x)\,(x \in y \iff x \in z \wedge \varphi).$$

Remarquons que le Schéma de Compréhension représente un nombre infini d'axiomes, un pour chaque énoncé φ. Il n'est pas

possible de donner une axiomatique finie équivalente à ZFC, comme cela a été démontré par Richard Montague dans sa thèse doctorale [Mon57].

À la base des axiomes introduits jusqu'ici, nous pouvons définir un ensemble spécifique, l'ensemble vide :

Définition 3.2.1. *Soit x^* un ensemble. Alors* $\emptyset = \{x \in x^* : x \neq x\}$.

Dans Définition 3.2.1 nous avons utilisé le symbole $\neq$, et même avant, dans la formulation du Schéma de Compréhension, nous avons utilisé la notation $x \in A$. Cette notation n'appartient pas normalement à notre langage, mais, comme d'habitude en logique, nous allons nous permettre d'utiliser des abréviations qui peuvent être traduites pzr une expression écrite formellement en $\mathcal{L}$.

Nous pouvons vérifier que l'ensemble vide est unique : si un autre ensemble z vérifie les mêmes propriétés que $\emptyset$, alors z et $\emptyset$ ont les mêmes éléments (nil) et donc ils sont égaux par l'extensionalité. Par contre, basés sur les axiomes nous avons introduits jusqu'ici, nous ne pouvons pas démontrer qu'il y a d'autres en-

sembles que l'ensemble vide. C'est une conséquence de la Solidité et de la Complétude de la logique du Premier Ordre, démontrées par Kurt Gödel, [Göd30]. Ce théorème dit qu'une théorie T (ensemble des phrases) dans la logique du premier ordre démontre une phrase φ ssi dans chaque modèle de T la phrase φ est satisfaite. En prenant pour T les axiomes exposés jusqu'ici, nous pouvons facilement vérifier que $M = \{\emptyset\}$ est un modèle de T et que la phrase $(\exists x)\, x \neq \emptyset$ n'est pas vérifiée dans ce modèle.

Nous pouvons aussi vérifier que le paradoxe de Russell est vraiment résolu :

Théorème 3.2.2. *L'ensemble de tous les ensembles n'existe pas.*

Preuve Supposons qu'il existe un x tel que pour tout y, $y \in x$. En appliquant Compréhension, nous avons que $B = \{y \in x : x \notin x\}$ est un ensemble, ce qui est une contradiction car la question de savoir si $B \in B$ n'a pas de réponse. $\bigstar_{3.2.2}$

Les deux axiomes suivants correspondent à deux opérations bien connues en théorie des ensembles, la formation d'une paire et la formation de la réunion d'une famille d'ensembles.

L'axiome de la Paire Dans la version formelle :

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

où, pour chaque choix de deux ensembles x, y, il y en a un troisième qui contient les deux.

Nous pourrions postuler l'Axiome de la Paire en demandant que z ne contienne que x et y. Cela n'est pas important, car si nous avons un z qui contient x, y, nous pouvons utiliser l'Axiome de Compréhension pour construire $w = \{o \in z : o = x \vee o = y\}$, qui contient exactement x et y. Dans les axiomes suivants nous allons utiliser la même logique pour ne pas préciser exactement les éléments d'ensembles construits.

Définition 3.2.3. *Si x, y sont deux ensembles quelconques, nous notons par $\{x, y\}$ l'ensemble qui contient exactement x et y. C'est* la paire *de x, y.*

Définition 3.2.3 ne demande pas expressément que x et y soient différents, donc on peut parfaitement bien définir le singleton $\{x\}$

comme la paire $\{x, x\}$. La définition ensembliste d'un *couple*, également appelé une paire *ordonnée*, est $(x, y) = \{\{x\}, \{x, y\}\}$.

L'axiome de la Réunion Pour tout ensemble $\mathcal{F}$, il existe un ensemble qui contient tous les éléments des ensembles éléments de l'ensemble $\mathcal{F}$, ou formellement

$$(\forall \mathcal{F})(\exists Z)(\forall a)(a \in Z \iff (\exists A \in \mathcal{F})(a \in A))).$$

Nous notons par $\bigcup \mathcal{F} = \{a : (\exists A \in \mathcal{F})(a \in A)\}$ et nous appelons cet ensemble *la réunion* de $\mathcal{F}$.

Nous pouvons maintenant définir plusieurs opérations habituelles de la théorie des ensembles, comme ceci :

Définition 3.2.4. *Pour deux ensembles A et B, nous écrivons $A \cup B$ pour $\bigcup\{A, B\}$ et $A \setminus B$ pour $\{x \in A : x \notin B\}$. Si $\mathcal{F}$ est une famille non vide d'ensembles, nous pouvons définir $\bigcap \mathcal{F} = \{z : (\exists A \in \mathcal{F})(a \in Z)\}$.*

Les axiomes que nous avons introduits ne nous donnent pas le pouvoir de définir le produit cartésien. Pour cela, il nous faut un

nouvel axiome, en fait encore une fois un schéma d'axiomes.

Schéma d'axiomes de Remplacement Pour chaque énoncé φ avec les variables libres parmi $x, y, A, w_1, \ldots, w_n$,

$$(\forall A)(\forall w_1)\ldots(\forall w_n)[(\forall x \in A)(\exists! y)\varphi \implies (\exists Y)(\forall x \in A)(\exists y \in Y)\varphi].$$

Informellement, cela veut dir que, un ensemble A étant donné, son image par une relation fonctionnelle est un ensemble.

On entend ici, par une relation fonctionnelle, une relation qui choisit pour chaque élément de son domaine un seul élément dans le range. Cela est en accord avec la définition d'une fonction dans la théorie des ensembles, que nous avons déjà rencontrée. Maintenant, nous sommes en mesure de justifier cette définition à partir des axiomes.

Définition 3.2.5. *Pour deux ensembles x et B, en observant que pour chaque $y \in B$ il y a exactement un ensemble (x, y), nous définissons $\{x\} \times B = \{(x, y) : y \in B\}$. Pour un ensemble quelconque A, nous définissons $A \times B = \bigcup\{\{x\} \times B : x \in A\}$.*

Une fois que nous avons défini le produit cartésien, nous pouvons définir une relation entre A et B comme un sous-ensemble quelconque de $A \times B$. Nous définissons une fonction de A à B comme une relation f entre A et B telle que pour chaque $x \in A$ il y a seulement un $y \in B$ tel que $(x, y) \in f$. Le lecteur peut vérifier que les notions d'ordre, de bon ordre et d'ordinal ou de cardinal sont toutes bien justifiées par les axiomes introduits jusqu'ici. Nous pouvons aussi facilement définir les nombres naturels $0 = \emptyset, 1 = \{\emptyset\}, \ldots n + 1 = \{0, 1, \ldots n\}$, etc. Mais, pour justifier l'existence d'un ensemble infini, il nous faut un autre axiome. Pour l'expliquer, nous remarquons que l'opération du successeur $x \mapsto x \cup \{x\}$ est bien définie à partir des axiomes que nous avons déjà introduits.

L'Axiome de l'Infini Il existe un ensemble auquel appartient l'ensemble vide et qui est clos par application du successeur, c'est-à-dire dans le langage formel :

$$\exists A(\emptyset \in A \wedge \forall x(x \in A \implies x \cup \{x\} \in A)).$$

L'axiome suivant parle d'une autre opération habituelle sur les ensembles, à savoir former l'ensemble de parties, c'est-à-dire l'ensemble des sous-ensembles d'un ensemble donné. À partir de cet axiome, nous allons présenter seulement les formulations informelles, sachant qu'une traduction formelle peut se faire de la même façon qu'auparavant.

L'Axiome de l'Ensemble des Parties Pour tout ensemble A, il existe un ensemble auquel appartiennent tous les sous-ensembles de A, et seulement ceux-ci. Cet ensemble s'appelle *l'ensemble des parties de* A et il est noté $\mathcal{P}(A)$.

Comme nous l'avons déjà vu dans le Théorème 1.2.5, en utilisant un ensemble infini quelconque, l'Axiome de l'Ensemble des Parties permet de prouver l'existence d'ensembles infinis non dénombrables. En raison de ce théorème de Cantor, en prenant de la façon répétitive l'ensemble des parties sur des ensembles infinis, on obtient des ensembles de plus en plus grands en cardinalité ; pour cette raison l'axiome renforce considérablement la théorie. Nous

discuterons cela en connexion avec la hiérarchie cumulative, qui est une façon d'organiser tous les ensembles. Pour ce faire, il nous faut un axiome très utile en théorie des ensembles et presque jamais invoqué ailleurs !

L'axiome de Fondation Tout ensemble x non vide possède soit un élément minimal pour l'appartenance sur x, soit un élément y n'ayant aucun élément en commun avec x.

Les axiomes évoqués jusqu'ici forment la théorie de ZF, nommée par Zermelo et Fraenkel. Finalement, parlons de l'Axiome le plus connu de la théorie des ensembles, l'axiome qui fait la différence entre ZF et ZFC : l'Axiome du Choix. Il est facile à formuler, mais nous allons voir que cette simplicité est décevante, car l'Axiome du Choix a des formulations surprenantes et inattendues.

L'axiome du Choix Tout ensemble x admet un bon ordre, c'est-à-dire une relation R sur x telle que (x, R) forment un bon ordre.

Par exemple, l'Axiome du Choix implique que l'ensemble $\mathbb{R}$ des réels admet un bon ordre. Évidemment, cet ordre n'est pas l'ordre

habituel sur $\mathbb{R}$, qui n'est pas un bon ordre car il produit des suites infinies décroissantes.

3.3 Les équivalents de l'Axiome du Choix

L'Axiome du Choix a beaucoup de formulations équivalentes. Les livres [RR70] et [RR85] en donnent plusieurs centaines. Nous allons évoquer, sans preuve, quelques uns des équivalents les plus connus.

Lemme de Zorn Un ensemble non vide partiellement ordonné $(A, \leq)$ est dit *inductif* quand toute partie totalement ordonnée (ou chaîne) de A admet un majorant, qui est un élément m de A vérifiant $a \leq m$ pour chaque $a \in A$. Le lemme de Zorn dit que tout ensemble inductif non vide admet un élément maximal.

La fonction de choix Étant donné un ensemble $\mathcal{F}$ d'ensembles non vides, il existe une fonction f définie sur $\mathcal{F}$, appelée *fonction de choix*, qui à chaque $F \in \mathcal{F}$ associe un de ses éléments $f(F) \in F$.

Le choix de représentants d'une relation d'équivalence Pour toute relation d'équivalence R sur un ensemble A, il existe un choix de représentants de R, autrement dit un sous-ensemble B de A tel que tout élément de A est R-équivalent à un unique élément de B.

3.3.1 Récursion Ordinale

Définition 3.3.1. *Soit A un ensemble et α un ordinal. On note par ${}^{\alpha}A$ l'ensemble des fonctions de la forme $f : \alpha \to A$ et on appelle ces fonctions des* suites de longueur α. *Pour un ordinal α^* on note par ${}^{<\alpha^*}A$ l'ensemble $\bigcup_{\alpha<\alpha^*} {}^{\alpha}A$.*

Pour une suite g de longueur β et $\alpha < \beta$ nous notons par $g \restriction \alpha$ la fonction g restreinte sur α.

Théorème 3.3.2. *Soit A un ensemble, α^* un ordinal et F une fonction de $\alpha^* \times {}^{<\alpha^*}A$. Alors il existe une unique fonction $g : \alpha^* \to A$ telle que pour tout ordinal $\alpha < \alpha^*$,*

$$g(\alpha) = F(\alpha, g \restriction \alpha). \qquad (*)$$

Preuve La preuve est par l'induction sur α^* appartenant à la classe des ordinaux. Si $\alpha^* = 0$, l'unique fonction g qui satisfait $(*)$ est $g = \emptyset$. Si $\alpha^* = \alpha+1$ pour un ordinal α, il nous suffit d'utiliser pour $g \restriction \alpha$ l'unique fonction $g_\alpha : \alpha \to A$ qui satisfait $(*)$, déjà défini par l'hypothèse de l'induction et de rajouter $g(\alpha) = F(\alpha, g_\alpha)$. Il est évident que cette définition produit une unique fonction sur α^* qui satisfait $(*)$.

Le cas de α^* un ordinal limite non-zéro est le plus intéressant. Par l'hypothèse de l'induction, pour chaque $\beta < \alpha^*$ il existe une unique fonction $g_\beta : \beta \to A$ qui satisfait $(*)$ pour chaque $\alpha < \beta$. Cela implique que pour $\beta < \gamma < \alpha^*$, nous avons $g_\beta = g_\gamma \restriction \beta$. Donc, les fonctions $g_\beta\,(\beta < \alpha^*)$ sont deux à deux compatibles et leur réunion $g = \bigcup_{\beta<\alpha^*} g_\beta$ est la fonction recherchée. Il est facile de voir qu'une telle fonction doit être unique. $\bigstar_{3.3.2}$

3.3.2 À vous de jouer

1 Démontrez que la définition de (x, y) satisfait la propriété de base des paires ordonnées : $(x, y) = (x', y')$ ssi $x = x'$ et $y = y'$.

2 Écrivez une formule pour exprimer $z = ((x, y), (v, w))$ en utilisant seulement $\in$ et $=$.

3 Démontrez que la définition de $\bigcap \mathcal{F}$ pour $\mathcal{F} \neq \emptyset$ est justifiée par les axiomes. Pourquoi demande-t-elle que $\mathcal{F} \neq \emptyset$?

4 Démontrez que les 'équivalents' de l'Axiome du Choix cités ici sont vraiment équivalents à l'Axiome du Choix.

5 Démontrez, en utilisant l'Axiome du Choix, que chaque espace vectoriel a une base.

3.4 Éléments historiques et bibliographiques

Le paradoxe de Russell utilise très peu de propriétés de l'appartenance, une relation binaire suffit. Cela a permis à Bertrand Russell d'illustrer le même paradoxe sous la forme plus imagée du paradoxe du barbier. Un barbier se propose de raser tous les hommes qui ne se rasent pas eux-mêmes, et seulement ceux-là. Le barbier doit-il se raser lui-même ? L'étude des deux possibilités conduit de nouveau à une contradiction. On résout le problème en affirmant qu'un tel barbier ne peut exister. [2] Le paradoxe de Russell n'était pas le premier paradoxe à apparaître dans la théorie des ensembles : il y avait le paradoxe de Buralli-Forti, découvert par ce dernier en 1897, et très clairement interprété par Georg Cantor dans une lettre de 1899 à Richard Dedekind où il montre que l'« ensemble » paradoxal en jeu, que nous appelons aujourd'hui la classe de tous les ordinaux, n'est pas un ensemble (), ou plus exactement est de nature différente. Donc, à cette époque, Cantor ne pense pas que tout

2. Jean-Marc Vanden-Broeck a remarqué qu'il suffit de proposer que le barbier est une femme.

prédicat définisse un ensemble, même s'il ne donne pas de définition précise de la différence entre ce que nous appelons aujourd'hui « ensemble » et « classe propre », et qu'il évoque sous les termes de « multiplicité [vielheit] consistante et inconsistante ». La solution de Cantor aux paradoxes ensemblistes, trop peu formelle, ne réussit pas vraiment à convaincre Richard Dedekind, l'un des premiers à utiliser la notion d'ensemble, et qui resta très ébranlé par la découverte des paradoxes.

La notion d'un couple qu'on utilise dans la théorie des ensembles a été découverte par Kazimierz Kuratowski, en 1921. Le Schéma d'axiomes de Remplacement a été introduit en 1922 indépendamment par Abraham Fraenkel et Thoralf Skolem. C'est grâce à ce schéma qu'on parle de la théorie des ensembles selon Zermelo et Fraenkel. Dans le même contexte, Abraham Fraenkel et Thoralf Skolem ont introduit l'Axiome de Fondation, aussi formulé par John von Neumann en 1925.

L'Axiome du Choix a une réputation volatile, car autant qu'il est nécessaire pour des énoncés les plus naturels des mathéma-

tiques, tel que chaque espace vectoriel a une base, il est aussi responsable pour un nombre de constructions d'ensembles assez paradoxales. La plus connue est probablement la construction de Stefan Banach et Alfred Tarski en 1924, qui affirme qu'il est possible de couper une boule de l'espace $\mathbb{R}^3$ en un nombre fini de morceaux et de ré-assembler ces morceaux pour former deux boules identiques à la première, à un déplacement près. Ce résultat paradoxal implique que ces morceaux sont non mesurables, sans quoi on obtiendrait une contradiction (le volume étant un exemple de mesure, cela veut plus simplement dire que ces morceaux n'ont pas de volume). L'Axiome du Choix est nécessaire pour construire des ensembles non mesurables. Ce paradoxe est le sujet du livre de Stanley Wagon, [Wag93]

Au sujet de la terminologie, le Schéma d'axiomes de Compréhension est parfois nommé Schéma d'axiomes de Séparation ou, simplement, Axiome de Compréhension ou Axiome de Séparation. Le Schéma d'axiomes de Remplacement est parfois appelé Schéma d'axiomes de Substitution. L'Axiome de Fondation est encore appelé l' Axiome de Régularité.

Au sujet de la grammaire française, l'auteure et sa co-rédactrice, toutes féministes qu'elles soient, ont hésité, dans le paragraphe précédent de §3.2, à utiliser « nous a amenée », sachant que ce « nous »est un « nous de modestie »....

Deuxième partie

Modèles de ZFC

Dans cette partie du livre notre style deviendra un peu plus mathématique, ce qui est nécessité par la thématique. On peut toujours en faire une première lecture en se concentrant que sur les énoncés, pour revenir aux preuves plus tard.

Chapitre 4

L'axiome de Fondation, hiérarchies **V** et **L**

4.1 La hiérarchie cumulative

L'axiome de Fondation (AF) n'entre pas dans le discours quotidien des mathématiciens, mais il joue un rôle considérable dans la théorie des ensembles elle-même, comme nous allons le voir dans ce chapitre. L'axiome dit que tout ensemble x non vide possède soit un élément minimal pour l'appartenance sur x, soit un élément

y n'ayant aucun élément en commun avec x. Une autre façon de comprendre cet axiome est de considérer la relation d'appartenance sur les ensembles (donc, la classe d'ensembles) et de dire qu'elle est *bien fondée*, c'est-à-dire que chaque ensemble non vide possède un élément minimal pour x. En particulier, sur chaque ensemble individuel, la relation $\in$ est bien fondée. Tout de suite, nous pouvons tirer des conséquences agréables de AF.

Corollaire 4.1.1. **1** *Un ordinal est un ensemble transitif sur lequel* $\in$ *représente un ordre linéaire.*

2 *Il n'y a pas une suite infinie* $\langle x_n : n < \omega \rangle$ *descendante en* $\in$*, c'est-à-dire vérifiant que* $x_{n+1} \in x_n$ *pour chaque* n*.*

3 *Il n'existe pas d'ensemble* x *tel que* $x \in x$*.*

4 *Il n'existe pas de cycle* $\langle x_n : n \leq k \rangle$ *tel que* $x_0 \in x_1 \in x_2 \in \ldots x_k \in x_0$*.*

Preuve

1 Comme $\in$ est bien fondée sur chaque ensemble, si nous avons un ensemble α qui est transitif et sur lequel $\in$ représente un ordre

linéaire, nous obtenons tout de suite que α est bien ordonné par $\in$ et, donc, α est un ordinal.

2 En discutant les bons ordres, nous avons effectivement déjà remarqué qu'une relation qui est bien fondée n'admet pas une suite infinie décroissante.

3 Si $x \in x$, alors l'ensemble $\{x\}$ est non vide et il n'a pas d'éléments minimaux en $\in$.

4 Un cycle $\langle x_n : n \leq k \rangle$ tel que $x_0 \in x_1 \in x_2 \in \ldots x_k \in x_0$ ne possède pas un élément minimal en $\in$.

$\bigstar_{4.1.1}$

L'axiome de Fondation nous permet d'organiser la classe des ensembles dans une hiérarchie stratifiée, comme nous allons l'expliquer maintenant. En utilisant le reste des axiomes, sans AF, nous définissons la classe V :

Définition 4.1.2. *Par induction sur α, un ordinal, nous définissons :*

— $V_0 = \emptyset$,

— $V_{\alpha+1} = \mathcal{P}(V_\alpha)$ *et*

— *pour* $\alpha > 0$ *un ordinal limite,* $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$.

Notons par V *la classe* $\bigcup_{\alpha\in\ Ord} V_\alpha$.

L'idée est que les ensembles V_α forment une stratification croissante de V dans les couches transitives et que l'Axiome de Fondation équivaut au fait que tous les ensembles se trouvent quelque part en V. Pour démontrer cela, nous allons développer quelques propriétés basiques de V.

Lemme 4.1.3. **1** *Pour tout ordinal* α, *nous avons* $V_\alpha = \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta)$.
Si $\beta \leq \alpha$, *alors* $V_\beta \subseteq V_\alpha$.

2 *Chaque* V_α *est un ensemble transitif.*

3 *Pour chaque* $x \in V$ *il existe un premier* α *tel que* $x \in V_\alpha$ *et un tel* α *est toujours un ordinal successif.*
Pour $x \in V$ *soit* $\rho(x) = \alpha$ *ssi* $\alpha + 1$ *est le premier ordinal* β *tel que* $x \in V_\beta$. *On l'appelle le* rang *de* x.

4 *Un ensemble* x *est dans* V *ssi tous* $z \in x$ *y sont, et dans ce cas nous avons* $\rho(x) = \sup\{\rho(z) + 1 : z \in x\}$.

5 *Chaque ordinal* α *est dans* V *et* $\rho(\alpha) = \alpha$. *Donc* $V_\alpha \cap Ord = \alpha$.

Preuve

1 La démonstration de ces deux énoncés est par induction sur α. Les énoncés sont vrais trivialement pour $\alpha = 0$. Pour $\alpha = \gamma + 1$ nous avons que $\beta < \alpha \iff \beta \leq \gamma$, et dans ce cas, par l'hypothèse de l'induction appliquée sur γ, nous avons que $V_\beta \subseteq V_\gamma$. Donc $\mathcal{P}(V_\beta) \subseteq \mathcal{P}(V_\gamma)$ et alors $\bigcup_{\beta<\alpha} \mathcal{P}(V_\beta) \subseteq \mathcal{P}(V_\gamma) = V_\alpha$. Il est évident que $V_\alpha = \mathcal{P}(V_\gamma) \subseteq \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta)$ et, donc, nous avons démontré que $V_\alpha = \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta)$. Si $\beta \leq \alpha$ alors $V_\beta = \bigcup_{\delta<\beta} \mathcal{P}(V_\delta) \subseteq \bigcup_{\delta<\alpha} \mathcal{P}(V_\delta) = V_\alpha$.

Si $\alpha > 0$ est un ordinal limite, nous avons que $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$, qui est, par l'hypothèse de l'induction, égal à $\bigcup_{\beta<\alpha} \bigcup_{\gamma<\beta} \mathcal{P}(V_\gamma) = \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta)$. Aussi, si $\beta < \alpha$, il est clair que $V_\beta \subseteq V_\alpha$ par la définition de V_α.

2 En utilisant les axiomes de la réunion et de l'ensemble des parties, nous voyons que chaque V_α est un ensemble. Pour démontrer que cet ensemble est transitif, nous utilisions l'induction sur α.

3 La définition de V implique que pour chaque $x \in V$ il existe

un ordinal α tel que $x \in V_\alpha$ et donc, comme les ordinaux sont bien ordonnés, il existe un premier tel ordinal α. Il est clair, en sachant que $V_0 = \emptyset$ et que pour $\alpha > 0$ limite nous avons $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$, qu'un tel premier α ne peut pas être une limite.

4 Si $x \in V$, soit $\alpha = \rho(x)$. Donc pour chaque $y \in x$ nous avons $y \in V_{\alpha+1}$, car $V_{\alpha+1}$ est transitif. En particulier, $y \in V$. Une ramification de cet argument donne le calcul exact des rangs.

5 Nous pouvons démontrer le premier énoncé par induction sur α, le deuxième énoncé est alors un conséquence.

$\bigstar_{4.1.3}$

L'observation cruciale est donnée par Théorème 4.1.4, dont nous allons donner la preuve dans le reste de cette section.

Théorème 4.1.4. *L'axiome de Fondation est équivalent à l'énoncé que la hiérarchie V contient tous les ensembles.*

Pour démontrer Théorème 4.1.4 nous allons introduire l'opération de la clôture transitive. Remarquons que cette définition n'a pas besoin de l'Axiome de Fondation.

Définition 4.1.5. *Pour un ensemble x, nous définissions par induction sur $n < \omega$,*

1 $\bigcup^0 x = x$,

2 $\bigcup^{n+1} x = x \cup \bigcup^n x$.

Alors on défini $\mathrm{trcl}(x) = \bigcup_{n<\omega} \bigcup^n x$.

Quelques propriétés de base de cette nouvelle notion sont faciles à démontrer, pour chaque ensemble A.

Lemme 4.1.6. *(a)* $A \subseteq \mathrm{trcl}(A)$,

(b) $\mathrm{trcl}(A)$ *est transitive,*

(c)Si $A \subseteq T$ *et* T *est un ensemble transitif, alors* $\mathrm{trcl}(A) \subseteq T$,

(d) Si A *est déjà transitif, alors* $\mathrm{trcl}(A) = A$,

(e) Si $x \in A$ *alors* $\mathrm{trcl}(x) \subseteq \mathrm{trcl}(A)$,

(f) $\mathrm{trcl}(A) = A \cup \bigcup_{x\in A} \mathrm{trcl}(x)$.

Preuve (b) Si $y \in \bigcup^n A$, alors $y \subseteq \bigcup^{n+1} A$. (c) Par induction sur n nous démontrons que $\bigcup^n A \subseteq T$. (d) Appliqons (c) avec $A = T$. (e) Si $x \in A$ alors $x \in \mathrm{trcl}(A)$ et donc $x \subseteq \mathrm{trcl}(A)$, et maintenant appliquons (c). Pour (f), l'inclusion $\supseteq$ est impliquée

par (a) et (e). L'inclusion $\subseteq$ est démontrée grâce à la transitivité d'ensemble $A \cup \bigcup_{x \in A} \mathrm{trcl}(x)$, qu'on peut vérifier facilement. $\bigstar_{4.1.6}$

Lemme 4.1.7. *Si A est transitif et $\in$ est bien-fondé sur A, alors $A \in V$.*

Preuve La preuve est par contradiction. Nous avons déjà remarqué que V contient comme élémént chaque ensemble x tel que $x \subseteq V$. Si l'énoncé n'est pas correct, nous ne pouvons donc pas avoir que $A \subseteq V$ et donc l'ensemble $A \setminus V$ n'est pas vide. Par l'hypothèse, nous pouvons choisir an x qui le $\in$-plus petit élément d'$A \setminus V$. A est transitif, donc $x \subseteq A$. Si $y \in x$, par le choix de x nous avons que $y \notin A \setminus V$, donc $y \in A \cap V$. Cela montre que $x \subseteq V$, et donc $x \in V$, contradiction. $\bigstar_{4.1.7}$

Preuve (du Théorème 4.1.4) Supposons que l'Axiome de Fondation est vrai et démontrons que V contient tous les ensembles. Pour un ensemble A quelconque, considérons $\mathrm{trcl}(A)$. C'est un ensemble transitif sue lequel $\in$ est bien fondé, donc par Lemme 4.1.7, nous avons $\mathrm{trcl}(A) \in V$. Donc $A \subseteq \mathrm{trcl}(A) \subseteq V$, et alors $A \in V$ car V contient tous ses sous-ensembles.

L'autre direction est claire car $\in$ est bien fondé sur V, comme le montre l'existence du rank. $\bigstar_{4.1.4}$

4.2 Relativisation ou les énoncés restreints

Définition 4.2.1. *On écrit* $(\exists x \in z)\varphi(x)$ *pour* $(\exists x)(\varphi(x) \wedge x \in z)$ *et* $(\forall x \in z)\varphi(x)$ *pour* $(\forall x)(\varphi(x) \wedge x \in z)$.

Définition 4.2.2. *Considérons une classe* M *et un énoncé* φ *du langage de la théorie des ensembles, dont tous les paramètres sont des objets de* M. *La* relativisation φ^M *de* φ *à* M *est définie par l'induction sur la complexité de* φ, *de la façon suivante :*

— *Si* φ *est atomique,* $\varphi^M = \varphi$,

— $(\neg\varphi)^M = \neg\varphi^M$, $(\varphi \wedge \psi)^M = \varphi^M \wedge \psi^M$,

— $(\exists x\varphi(x))^M$ *est* $(\exists x \in M)(\varphi^M(x))$.

Nous pouvons en déduire la forme de la relativisation pour les connectives $\vee$, $\Longrightarrow$ et $\Longleftrightarrow$, car elles sont définissables en terme de $\neg$ et $\wedge$. Également, rappelons que le quantificateur $\forall$ peut se définir en terme de $\neg$ et $\exists$, donc $(\forall x\varphi(x))^M$ est $(\forall x \in M)(\varphi^M(x))$.

De cette même manière, la plupart du temps dans nos arguments nous allons seulement vérifier les cas de $\neg, \wedge$ et $\exists$.

L'idée de la relativisation est que la relativisation d'un énoncé φ est vraie dans un modèle M ssi φ^M est vrai dans l'univers $\mathbf{V}$ où se passe la discussion (donc le méta-univers). Par exemple par $\varphi^M \iff \varphi^N$ nous entendons que nous pouvons vérifier cette équivalence en $\mathbf{V}$. Il s'agit ici d'un point philosophique important, car il implique que la notion de vérité dans un modèle et dans le méta-univers ne sont pas toujours les mêmes ! Il est facile de vérifier que la définition de la relativisation est faite d'une façon analogue à la définition de la satisfaction d'un énoncé dans un modèle dans la théorie des modèles. Donc, dans le cas où M est un ensemble, la relativisation est déjà une notion connue. Nous allons l'utiliser aussi dans un contexte des classes propres.

4.3 Absoluité

Nous avons déjà discuté la relativisation et nous avons défini la notion de φ^M pour un énoncé φ et un modèle M. Nous avons eu l'occasion de voir que le fait de vérifier qu'un M est un modèle d'une théorie revient à vérifier plusieurs φ^M, et qu'il est souvent très utile de savoir que φ^M est de la forme φ^N pour un modèle N que nous connaissons déjà. Ici nous allons développer une méthode générale pour la comparaison de φ^M et φ^N pour des paires (M, N) de modèles. Dans cette section, les lettres φ, ψ et θ sont utilisées pour les énoncés du premier ordre du langage de la théorie des ensembles (même si quelques énoncés s'appliquent dans un contexte plus large) et M, N pour des classes possédant une interprétation du symbole $\in$. La notation $\bar{x}, \bar{y}, \bar{z}$ est réservée pour les n-tuplets (pour un entier n) de variables et en écrivant $\varphi(\bar{x})$ nous sous-entendons que φ est un énoncé avec le nombre de variables correspondant à la longueur de $\bar{x}$. Nous allons souvent oublier cette pédanterie et écrire que $\varphi(x)$, même en sachant qu'il fallait vrai-

ment parler de $\varphi(\bar{x})$.

La notion principale est donnée par la définition suivante :

Définition 4.3.1. *Supposons que $M \subseteq N$. On dit que la formule* $\varphi(\bar{x})$ est absolue pour M, N *ssi pour chaque $\bar{a} \in M$, nous avons $\varphi^M[\bar{a}] \iff \varphi^N[\bar{a}]$. Si N est la hiérarchie cumulative V, on dit que* φ est absolu pour M *au lieu de dire que φ est absolu pour M, V.*

Un lemme, qui peut se vérifier comme exercice, est que la notion d'absoluité est fermée par les opérations logiques :

Lemme 4.3.2. *Si φ et ψ sont absolus pour M, N, alors $\neg\varphi$ et $\varphi \wedge \psi$ le sont aussi.*

Donc, dans le contexte de Lemme 4.3.2, les énoncés $\varphi \vee \psi$, $\varphi \implies \psi$ et $\varphi \iff \psi$ sont absolus pour M, N aussi, car on peut les définir en terme de négations et de conjonctions. Une autre conséquence facile de Lemme 4.3.2, mais qui mérite d'être énoncée, est que *tout* énoncé sans quantificateurs est absolu, pour M, N quelconques. Donc le jeu de relativisation se passe au niveau

des quantificateurs. Il est important de noter que la transitivité, si présente, a une influence sur l'absoluité, dans le cas des quantificateurs.

Lemme 4.3.3. *Supposons que $M \subseteq N$, que M est transitive et que $\in^M = \in^N = \in$, tandis que φ est absolu pour M, N. Alors $(\exists x \in z)\varphi(x)$ est aussi absolu pour M, N.*

Preuve Supposons que $z \in M$ et que $(\exists x \in z)\varphi^M(x)$. Évidemment, ce même x est aussi en N car $M \subseteq N$ et, comme φ est absolu pour M, N, nous avons que $(\exists x \in z)\varphi^N(x)$. Dans l'autre direction, s'il y a un $x \in N$ tel que $x \in z$ et $\varphi^N(x)$, en utilisant la transitivité de M et que $z \in M$, nous avons $z \subseteq M$ et donc le même $x \in M$ montre que $(\exists x \in z)\varphi^M(x)$. $\bigstar_{4.3.3}$

On peut se demander si la supposition que $\in^M = \in$ est difficile à satisfaire. En fait, un théorème de Mostowski [Mos49] montre que cette restriction n'a pas d'importance.

Théorème 4.3.4. *(Lemme de Contraction de Mostowski) Soit M un modèle du language $\mathfrak{L} = \{\in\}$ vérifiant l'Axiome d'Extensionalité et l'Axiome de Fondation. Alors il existe un unique modèle N*

de $\{\in\}$ *tel que* $\in^N = \in$ *est il y a un isomorphisme* $\pi : (M, \in^M) \to N$. *En plus,* N *est transitif et l'isomorphisme est unique.*

Preuve On définit l'isomorphisme π par réccurence sur $\in^M$ par $\pi(x) \stackrel{\text{def}}{=} \{\pi(y) : y \in^M x\}$. $\bigstar_{4.3.3}$

Le N obtenu dans Théorème 4.3.3 s'appelle *contracté ou le collapse de Mostowski* de N. Grâce à ce théorème, sauf avis contraire, nous allons dès maintenant discuter seulement des modèles M tel que $\in^M = \in$.

La notion de la quantification bornée donne lieu à la classe des énoncés Δ_0, aussi connus comme q.u.b. (cf. [Kri09]).

Définition 4.3.5. *Nous définissons la classe* Δ_0 *comme étant la plus petite classe des énoncés de la théorie des ensembles telle que :*

— *tout énoncé atomique est* Δ_0,

— *si* φ *et* ψ *sont* Δ_0, *alors* $\neg\varphi$ *et* $\varphi \wedge \psi$ *sont* Δ_0,

— *si* ψ *est* Δ_0, *alors* $(\exists x \in z)\psi$ *est* Δ_0.

Donc, les lemmes précédents peuvent se résumer en disant que les énoncés Δ_0 sont absolus pour les classes M, N tels que M est

transitive. Il est évident que les énoncés qui sont *équivalents* à un énoncé Δ_0 sont absolus aussi, c'est-à-dire que, si à partir d'axiomes on peut démontrer que $(\forall x)(\varphi(x) \iff \psi(x))$ et que φ est Δ_0, alors ψ aussi est absolu pour M, N tels que M est transitive. Nous allons utiliser cette observation pour obtenir quelques instances du Théorème 4.3.6. Soit ZF_0 la théorie ZF moins les Axiomes de l'Infini et des Parties. Pour une sous-théorie S de ZFC nous notons par S^- la théorie S privée de l'Axiome de Fondation.

Théorème 4.3.6. *On peut vérifier dans la théorie ZF_0^- que les énoncés et les notions suivantes sont équivalents à un énoncé Δ_0, et donc ils sont absolus pour les modèles $M \subseteq N$ qui sont transitifs et vérifient ZF_0^- :*

— $x \in y,\ x = y,$

— $\{x, y\},\ \{x\},\ (x, y),$

— $\emptyset, x \cup y, x \cap y, x \setminus y,\ x \cup \{x\},\ \bigcup x,\ \bigcap x$ *(si $x \neq \emptyset$),*

— *« x est transitif ».*

Preuve Nous allons donner quelques instances. Les énoncés atomiques sont Δ_0 par définition. L'ensemble $\{x, y\}$ est l'unique z tel

que

$$x \in z, y \in z \wedge (\forall w \in z)(w = x \vee w = y),$$

qui est évidemment un énoncé Δ_0. La notion « x est transitif »est équivalente à

$$(\forall y \in x)(\forall z \in y)(z \in x).$$

★4.3.6

En fait, nous n'avons utilisé que la transitivité de M. Comme notre situation en pratique va toujours se présenter avec deux modèles transitifs, nous allons nous permettre de manquer de précision et de considérer seulement des paires de modèles transitifs. Il y a des exemples d'énoncés qui sont absolus et qui ne sont pas Δ_0. Des exemples importants de notions qui ne sont ni Δ_0 ni absolus sont $x \subseteq y$ et $\mathcal{P}(x)$. Le fait que $\mathcal{P}(x)$ n'est presque jamais absolu est la cause de plusieurs résultats d'indépendance en théorie des ensembles, notamment l'indépendance de HC.

Pour continuer notre analyse des quantificateurs, tout énoncé peut se classer par le nombre des quantificateurs non bornés qui

sont nécessaires pour écrire l'énoncé, comme on peut le voir dans Définition 4.3.7.

Définition 4.3.7. *Par induction sur $n \geq 0$ nous définissons les classes Σ_n, Π_n des formules :*

- *Les classes Σ_0 et Π_0 sont simplement égales à Δ_0,*
- *Supposons que $\theta \in \Pi_n$ et soit $\varphi = (\exists \bar{x})(\theta(\bar{x}))$. Alors φ est Σ_{n+1}.*
- *Supposons que $\theta \in \Sigma_n$ et soit $\varphi = (\forall \bar{x})(\theta(\bar{x}))$. Alors φ est Π_{n+1}.*

Nous définissons pour une théorie S les classes Σ_n^S et Π_n^S en disant qu'un énoncé φ est Σ_n^S si on peut démontrer en S que φ est équivalent à un énoncé Σ_n, et de même pour Π_n^S. Nous définissons aussi la classe Δ_n^S comme étant la classe des énoncés qui sont Σ_n^S et Π_n^S.

Nous avons déjà tacitement utilisé dans la preuve du Théorème 4.3.6 une observation importante, qui généralise nos observations dans les lemmes précédents, la définition de la classe Δ_0 : si M, N sont des modèles transitifs d'une théorie S et si l'énoncé S est dans

la classe Δ_0^S, alors φ est absolu pour M, N. Au niveau des énoncés Σ_1 et Π_1, nous n'avons plus l'absoluité, mais nous observons un transfert parmi les modèles transitifs.

Lemme 4.3.8. *Supposons que $M \subseteq N$ sont tout les deux transitifs, que φ est Σ_1 et ψ est que Π_1. Alors $\varphi^M \implies \varphi^N$ et $\psi^N \implies \psi^M$. En conséquence, les énoncés dans la classe Δ_1^S sont absolus pour M, N, pour toute S.*

Preuve Supposons que θ est Δ_0 et que $\varphi = (\exists x)\theta(x)$. Alors φ^M est $(\exists x \in M)\theta^M(x)$ et donc $\varphi^M \implies \varphi^N$. La preuve pour ψ est similaire. $\bigstar_{4.3.8}$

Nous pouvons utiliser ces observations pour discuter l'absoluité des énoncés pour les modèles transitifs de ZF_0^-.

Théorème 4.3.9. *Les notions suivantes sont absolues pour les modèles transitifs de ZF_0^- :*

- *« x est un couple ordonné »,*
- *$A \times B$,*
- *« R est une relation », « R est une fonction », « R est injective »*

— $\mathrm{dom(R)}, \mathrm{ran(R)}$, $R(x)$.

4.4 Quelques modèles des sous-théories de ZFC

Maintenant nous sommes prêts à discuter des modèles des axiomes de ZFC. Nous allons utiliser la théorie ZF^- dans notre méta-univers.

Lemme 4.4.1. *Si M est une classe transitive, alors M vérifie l'Axiome d'Extensionalité.*

Preuve Nous pouvons écrire l'Axiome d'Extensionalité en disant $\forall x, y,\ \forall z(z \in x \iff z \in y) \implies x = y$. Notons la dernière expression par $\varphi(x, y)$. Donc $\varphi^M(x, y)$ est $\forall z \in M(z \in x \iff z \in y) \implies x = y$, qui est, par la transitivité de M, équivalent à $\varphi(x, y)$, qui est vrai par l'Axiome d'Extensionalité dans le méta-univers. $\bigstar_{4.4.1}$

Lemme 4.4.2. *Si pour tout $z \in M$ nous avons $\mathcal{P}(z) \subseteq M$, alors*

l'Axiome de Compréhension est vrai en M.

Preuve Étant donnés des ensembles $z, z_1, \ldots, z_n$ en M et une formule $\varphi(x, z_1, \ldots, z_n)$, nous avons que

$$A = \{x | x \in z \wedge \varphi^M(x, z_1, \ldots, z_n)\}$$

est un ensemble dans le méta-univers, et donc un élément de $\mathcal{P}(z)$. Par l'hypothése, $A \in M$. $\bigstar_{4.4.2}$

Théorème 4.4.3. *Soit $M = \{\emptyset\}$. Alors M vérifie l'Axiome d'Extensionalité, l'Axiome de Fondation et la phrase $\forall y(y = \emptyset)$.*

La preuve est évidente, en utilisant Lemme 4.4.2 pour la Compréhension. Nous avons donc démontré, en utilisant l'abréviation *Con*, qui est la notation habituelle, pour la cohérence :

Corollaire 4.4.4. *La cohérence de* (ZF$^-$) *implique la cohérence de Extensionalité+ Fondation + Compréhension+ $\forall y(y = \emptyset)$).*

En utilisant l'absoluité de $\emptyset$ et de la formation des ordinaux finis, nous pouvons voir que les nombres finis sont absolus. Il se

trouve que cela est vrai aussi pour ω dans le contexte des modèles transitifs de ZF_0^-.

Lemme 4.4.5. *Soit M un modèle transitif de ZF_0^-. Si $\omega \in M$, alors $\omega^M = \omega$ et M vérifie l'Axiome de l'Infini. Aussi, si M vérifie l'Axiome de l'Infini, alors $\omega \in M$.*

Preuve L'Axiome de l'Infini dit que

$$(\exists x)(\emptyset \in x \wedge (\forall y \in x)S(y) \in x),$$

donc on peut tirer la première conclusion en utilisant l'absoluité. Pour la deuxième conclusion, soit A un ensemble quelconque vérifiant l'Axiome de l'Infini en M. On peut définir ω comme le plus petit sous-ensemble de A tel que $\emptyset \in z$ et $(\forall y \in z)S(y) \in z$, autrement dit, la clôture en A de $\emptyset$ par l'opération S. $\bigstar_{4.4.5}$

Maintenant nous allons aborder l'Axiome du Choix. Il dit que chaque ensemble peut être bien ordonné, donc discutons d'abord l'absoluité des notions impliquées. L'énoncé « R est un ordre linéaire »est Δ_0. La notion d'un bon ordre est plus complexe à dé-

crire, mais nous pouvons démontrer Lemme 4.4.6, qui va suffire pour nos exemples. Notons déjà que le fait d'être un élément minimal d'une relation sur un ensemble est Δ_0 et utilisons cela pour démontrer le lemme.

Lemme 4.4.6. *Supposons que M est une classe transitive telle que pour chaque $A \in M$, $A \times A \in M$ et $\mathcal{P}(A) \subseteq M$. Alors pour une relation R sur un élément $A \in M$, nous avons « R est un bon ordre sur A »ssi « (R est un bon ordre sur A)^M ».*

Preuve Si $A \in M$, alors $A \times A \in M$ et donc chaque relation R sur A appartient à $\mathcal{P}(A \times A) \subseteq M$. Si R est un bon ordre sur $A \in M$ et $\emptyset \neq X \subseteq A$, alors $X \in M$ a un élément qui est R-minimal, disons x. Par la transitivité de M nous avons que $x \in M$, et donc (R est un bon ordre sur A)^M. L'autre direction est similaire. $\bigstar_{4.4.6}$

Les résultats de cette section impliquent le théorème suivant.

Théorème 4.4.7. *En utilisant les axiomes de ZF^-, nous pouvons démontrer que V_ω est un modèle de ZF sans l'Axiome de l'Infini et vérifiant la négation de l'Axiome de l'Infini. Si nous avons aussi*

recours à l' Axiome de Choix, nous pouvons démontrer que V_ω est un modèle de l'Axiome de Choix.

En résumé,

— La cohérence de ZF^- implique la cohérence de $ZF\backslash$ Infini + $\neg$ Infini.

— La cohérence de ZFC^- implique la cohérence de $ZFC\backslash$ Infini + $\neg$ Infini.

La discussion des modèles de la forme V_α nous emmènera dans le domaine des *grands cardinaux*, comme nous allons le voir de suite.

Définition 4.4.8. *Un cardinal régulier $\kappa > \aleph_0$ est* faiblement inaccessible *ssi il est un cardinal limite. Un tel κ qui satisfait également*

$$(\forall \lambda < \kappa)(2^\lambda < \kappa)$$

est dit d'être fortement inaccessible.

En utilisant les techniques des sections précédentes, nous pouvons facilement démontrer :

Théorème 4.4.9. *Si κ est fortement inaccessible, alors V_κ est un modèle de ZFC.*

Cela implique, par l'Incomplétude de Gödel que :

Corollaire 4.4.10. *On ne peut pas démontrer l'existence d'un cardinal qui est fortement inaccessible par une preuve en ZFC.*

En fait, le même est vrai pour les cardinaux qui sont faiblement inaccessibles car Solovay a démontré par la méthode du forcing (et avec une preuve qui est facile aujourd'hui) que si on a un modèle de ZFC avec un cardinal qui est faiblement inaccessible, alors nous pouvons changer le modèle pour avoir un cardinal qui est fortement inaccessible. On peut voir cela aussi en regardant l'univers constructible, qui le sujet de §4.5. Donc, nous ne pouvons pas démontrer en ZFC l'existence d'un cardinal qui est faiblement inaccessible. Les *grands cardinaux* sont en fait définis comme les propriétés que peut avoir un cardinal, disons $\varphi(\kappa)$, telles que l'existence d'un cardinal qui satisfait $\varphi(\kappa)$ ne peut pas être établie en ZFC.

4.5 L, l'univers constructible

Notre construction de la hiérarchie cumulative $\mathbf{V}$ nous a permis une vérification assez facile des axiomes de ZFC dans l'univers $\mathbf{V}$. Mais, quand à HCG, cette construction ne nous permet pas de déterminer l'arithmétique des cardinaux ni de savoir si HCG est satisfaite en $\mathbf{V}$. Dans ses considérations de HCG, Kurt Gödel a développé une modification de la hiérarchie cumulative qui permet de vérifier HCG, l'univers constructible $\mathbf{L}$. L'idée de la construction de $\mathbf{L}$ est d'avoir encore une hiérarchie comme $\mathbf{V}$, mais de ne prendre que certains sous-ensembles dans les étapes de la forme $\alpha + 1$. La construction est telle qu'on peut compter le nombre de sous-ensembles pris.

Définition 4.5.1. *En commençant avec un modèle* $(M, \in)$ *du langage* $\mathcal{L} = \{\in\}$*, nous disons q'un sous-ensemble* $A \subseteq M$ *est* définissable en M *ssi il y a une formule* φ *de* $\mathcal{L}$ *et des paramètres* $a_i\,(i < n)$ *dans* M *tels que* $A = \{a \in M : \varphi[a; a_0, \ldots a_{n-1}]\}$*. Nous écrivons* $\mathrm{Def}(M) = \{A : A \text{ est définissable en } M\}$*.*

Le point important de cette définition est qu'on peut toute suite voir que la cardinalité de $\mathrm{Def}(M)$ pour M infinie est $|M|$. C'est la cas car nous avons que pour $M^{<\omega} = \bigcup_{n<\omega} M^n$ la cardinalité est $|M^{<\omega}| = |M|$, qui est le nombre des paramètres possibles. Il n'y a qu'un nombre dénombrable des formules φ, donc il y a au plus $|M|$ sous-ensembles définissables de M.

Des exemples d'ensembles définissables sont des intervalles (a, b) dans $\mathbb{Q}$ et $\emptyset$ pour un M quelconque.

Définition 4.5.2. *Par récurrence sur un ordinal α définissons :*

— $L_0 = \emptyset$,

— $L_{\alpha+1} = \mathrm{Def}(L_\alpha)$,

— $L_\delta = \bigcup_{\alpha<\delta} L_\alpha$ *pour δ limite > 0.*

Nous définisons $\mathbf{L} = \bigcup_{\alpha\in Ord} L_\alpha$.

Lemme 4.5.3. *(1) Pour $(M, \in)$ quelconque, $\emptyset, M \in \mathrm{Def}(M)$ et $M \subseteq \mathrm{Def}(M) \subseteq \mathcal{P}(M)$. Tout ensemble fini appartenant à $\mathcal{P}(M)$ est en $\mathrm{Def}(M)$.*

(2) Quelques propriétés simples de **L** *sont :*

— *Chaque L_α est un ensemble transitif et* **L** *est une classe*

transitive,

— $\alpha < \beta \implies L_\alpha \subseteq L_\beta$.

— $\alpha \in L_{\alpha+1}$.

Preuve (1) $M = \{x \in M : x = x\} \in \mathrm{Def}(M)$, $\emptyset = \{x \in M : x \neq x\} \in \mathrm{Def}(M)$ et pour tout $F = \{a_0, \ldots a_{n-1}\} \subseteq M$, $F = \{x \in M : x = a_0 \vee \ldots x = a_{n-1}\}$.

(2) Par l'induction sur α.

(3) Par l'induction sur α. Nous avons $\alpha \notin V_\alpha$, donc certainement $\alpha \notin L_\alpha$. Mais, en sachant que $\alpha \subseteq L_\alpha$ et que α est définissable comme l'ensemble de tous les ordinaux en L_α, nous obtenons que $\alpha \in L_{\alpha+1}$. $\bigstar_{4.5.3}$

De la même façon que la définition de V a donné lieu a un rank, on peut définir le rank sur $\mathbf{L}$:

Définition 4.5.4. *Pour tout $x \in \mathbf{L}$, on défini le rank $\rho_L(x)$ comme l'ordinal minimal α tel que $x \in L_{\alpha+1}$.*

4.5.1 Les axioms de ZF en $\mathbf{L}$

Ici nous vérifions un par un les axiomes de ZF en $\mathbf{L}$.

L'axiome de fondation La relation $\in$ est bien fondée en $\mathbf{L}$.

$(\mathbf{L}, \in)$ est une sous-structure de $(V, \in)$ qui est bien fondée, elle l'est donc également.

L'axiome d'extensionalité Deux ensembles ssi ils ont les mêmes éléments.

Par transitivité de $\mathbf{L}$.

L'axiome de la paire Si x, y sont des ensembles, alors $\{x, y\}$ est un ensemble.

Si $x, y \in \mathbf{L}$, alors il existe un ordinal α tel que $x, y \in L_\alpha$. Par conséquent $\{x, y\} = \{s | s \in L_\alpha \wedge (s = x \vee s = y)\} \in L_{\alpha+1}$.

L'axiome de l'union Pour tout ensemble x, il existe un ensemble y dont les éléments sont précisément les éléments des éléments de x.

Si $x \in L_\alpha$, alors ses éléments sont dans L_α et leurs éléments appartiennent à L_α. La réunion de x, disons y, est donc un sous-ensemble de L_α, Maintenant, $y = \{s | s \in L_\alpha \wedge (\exists z \in x) s \in z\} \in L_{\alpha+1}$. Donc $y \in \mathbf{L}$.

L'axiome de l'infini Il existe un ensemble x tel que $\emptyset \in x$ et

tel que si $y \in x$, l'union $y \cup \{y\}$ lui appartient également.

Car $\omega \in \mathbf{L}$.

L'axiome de compréhension Étant donnés des ensembles $z, z_1, \ldots, z_n$ et une proposition $P(x, z_1, \ldots, z_n)$, alors $\{x | x \in z \wedge P(x, z_1, \ldots, z_n)\}$ est un ensemble.

Par induction sur les sous-formules de P, on peut montrer qu'il existe un α pour lequel L_α contient z et $z_1, \ldots, z_n$ et P est vraie dans L_α si et seulement si P est vraie dans $\mathbf{L}$ — c'est le principe de réflexion traité en §4.5.4. Donc

$$\{x | x \in z \wedge P^{\mathbf{L}}(x, z_1, \ldots, z_n)\} = \{x | x \in L_\alpha \cap z \wedge P^{L_\alpha}(x, z_1, \ldots, z_n)\},$$

qui est un ensemble en $L_{\alpha+1}$. Par conséquent ce sous-ensemble est dans $\mathbf{L}$.

L'axiome de remplacement Étant donné n'importe quel ensemble S et n'importe quelle classe fonctionnelle, définie formellement comme une proposition $P(x, y)$ pour laquelle $P(x, y)$ et $P(x, z)$ implique $y = z$, la collection $\{y : (\exists x \in S) P(x, y)\}$ est un

ensemble.

Soit $Q = P^{\mathbf{L}}$. En général, Q est une formule bien plus complexe que P, mais comme P était fonctionnelle sur $\mathbf{L}$, Q est fonctionnelle sur $\mathbf{V}$ (vérifiez cela) ; par conséquent, on peut appliquer le remplacement dans $\mathbf{V}$ à Q. Donc

$$\{y \in \mathbf{L} | (\exists x \in S) P^{\mathbf{L}}(x, y)\} = \{y \in L | \, (\exists x \in S) Q(x, y)\}$$

est un ensemble dans $\mathbf{V}$ et est une sous-classe de $\mathbf{L}$. On peut montrer qu'il existe un α tel que cet ensemble est inclus dans $L_\alpha \in L_{\alpha+1}$.

L'axiome de l'ensemble des parties Pour tout ensemble x, il existe un ensemble y dont les éléments sont précisément les sous-ensembles de x.

En général, certaines parties d'un ensemble x de $\mathbf{L}$ ne sont pas dans $\mathbf{L}$, donc l'ensemble de toutes les parties de x non plus. Ce qu'on doit montrer ici, c'est que l'intersection avec $\mathbf{L}$ de cet ensemble des parties appartient à $\mathbf{L}$. Nous utilisons l'axiome de remplacement dans V pour montrer qu'il existe un α pour lequel

cette intersection est un sous-ensemble de L_α. L'intersection s'exprime alors $\{z \in L_\alpha | z \subseteq x\} \in L_{\alpha+1}$. L'ensemble requis est donc dans **L**.

4.5.2 L'axiome du choix en L

Nous allons montrer qu'il existe un bon ordre sur **L** dont la définition fonctionne avec la même définition à l'intérieur de **L** lui-même. La preuve ici est due à Jensen. Supposons que $x \neq y$ soient deux ensembles en **L**, et que nous souhaitons déterminer si $x < y$ ou si $x > y$. Si $\rho_L(x) = \alpha$, $\rho_L(y) = \alpha$ et $\alpha \neq \beta$, alors définissons $x < y$ si et seulement si $\alpha < \beta$. À partir de ce point nous supposerons que $\alpha = \beta$.

Rappelons que $L_{\alpha+1}$ est défini grâce à des formules paramétrisées par des ensembles en L_α. Si on oublie temporairement les paramètres, les formules peuvent être énumérées par une énumération standard de Gödel . Si φ et ψ sont les formules avec des plus petits nombres de Gödel permettant de définir respectivement x et y, et si ces formules sont différentes, nous définissons $x < y$ si

et seulement si le nombre de φ est strictement inférieur au nombre de ψ. Maintenant nous supposerons $\varphi = \psi$.

Continuons la définition en supposant que φ possède n paramètres de L_α. Supposons que la séquence des paramètres $z_1, ..., z_n$ soit la séquence de paramètres de φ utilisée pour définir x, et que de même $w_1, ..., w_n$ soit celle de y. Choisissons alors $x < y$ si et seulement si $z_n < w_n$ ou $z_n = w_n$ et $z_{n-1} < w_{n-1}$ etc, c'est-à-dire l'ordre lexicographique inverse. S'il existe plusieurs séquences définissant le même ensemble, nous choisissons la plus petite par cette définition. Nous devons remarquer que les valeurs possibles pour un paramètre sont ordonnées d'après l'ordre sur **L** restreint à L_α, cette définition est donc une définition par récurrence transfinie.

Notons que ce bon ordre peut être défini dans **L** lui-même par une formule sans paramètre de la théorie des ensembles, seulement avec les variables libres x et y.

Tout l'intérêt de **L** est que la preuve que c'est un modèle de ZFC nécessite uniquement que V soit un modèle de ZF, c'est-à-dire qu'elle ne dépend pas de la validité de l'axiome du choix dans

V. Donc, on obtient :

Théorème 4.5.5 (Gödel)**.** *La cohérence de ZF implique la cohérence de ZFC.*

4.5.3 La minimalité et l'unicité de L

On appelle une classe M *un modèle standard* si M est une classe transitive dont la relation d'appartenance est celle de **V**. Par conséquent, toute formule Δ_0 est absolue. Un *modèle intérieur* est un modèle standard qui est une sous-classe de V et qui contient tous les ordinaux de **V**. **L** est donc un modèle intérieur. En fait, **L** est un modèle intérieur canonique, dans le sens suivant.

Soit W un modèle standard de ZF qui possède les mêmes ordinaux que V. L'univers **L** défini dans W est alors le même que celui défini dans **V**, par l'absoluité. En particulier, L_α est le même dans **V** et W pour n'importe quel ordinal α, et les mêmes formules et paramètres dans $\mathrm{Def}(L_\alpha)$ produisent les mêmes ensembles constructibles dans $L_{\alpha+1}$.

De plus, comme **L** est une sous-classe de **V** et de W, L est la

plus petite classe qui à la fois contient tous les ordinaux et qui est un modèle standard de ZF. En effet $\mathbf{L}$ est l'intersection de toutes ces classes.

4.5.4 Le principe de la réflexion

Une propriété importante des hiérarchies des ensembles, $\mathbf{V}$ et $\mathbf{L}$, est la *réflexion*.

Théorème 4.5.6. *(1) Pour n'importe quels énoncés $\varphi_0, \varphi_1, \ldots \varphi_n$ de la théorie des ensembles, il existe un ordinal α tel que pour chaque $i \leq n$, φ_i est vrai en $\mathbf{V}$ ssi $V_\alpha \models \varphi_i$.*
(2) Pour quelconques énoncés $\varphi_0, \varphi_1, \ldots \varphi_n$ de la théorie des ensembles, il existe un ordinal α tel que pour chaque $i \leq n$, φ_i est vrai en $\mathbf{V}$ ssi $V_\alpha \models \varphi_i$.

Pour démontrer Théorème 4.5.6 il nous sera utile d'établir le lemme suivant.

Lemme 4.5.7. *Soit M, N deux classes telles que $M \subseteq N$ et $\varphi_0, \varphi_1, \ldots \varphi_n$ des formules telles que la liste contient toutes les sous-*

formules. Alors, les énoncés suivants sont equivalents :

(i) $\varphi_0, \varphi_1, \ldots \varphi_n$ *sont absolus pour* M *et* N,

(i) *Pour chaque* φ_j *de la forme* $\exists x\, \varphi_i(x, y_0, \ldots, y_m)$ *et pour chaque* $y_0, \ldots, y_m$ *dans* M, *il existe* $x \in M$ *tel que* $\varphi_i^M(x, y_0, \ldots, y_m)$ *ssi il existe* $x \in N$ *tel que* $\varphi_i^M(x, y_0, \ldots, y_m)$.

Preuve La direction (i) $\Longrightarrow$ (ii). Nous avons que $\varphi_j^M(y_0, \ldots, y_m)$ ssi $(\exists x \varphi_i(x, y_0, \ldots, y_m))^M$ ss'il existe $x \in M$ tel que $\varphi_i^M(x, y_0, \ldots, y_m)$. Aussi, $\varphi_j^N(y_0, \ldots, y_m)$ ss'il existe $x \in N$ tel que $\varphi_i^N(x, y_0, \ldots, y_m)$. Par l'absoluité de φ^j et φ^i, nous avons qu'il existe $x \in M$ tel que $\varphi_i^M(x, y_0, \ldots, y_m)$ ss'il existe $x \in N$ tel que $\varphi_i^M(x, y_0, \ldots, y_m)$.

Pour l'autre direction, nous pouvons supposer que si φ_i est une propre sous-formule de φ_j, alors $i < j$. La preuve est par induction sur j. En arrivant sur φ_j, il nous suffit de vérifier le cas que $\varphi_j = \varphi_i \wedge \varphi_k$, $\varphi_j = \neg\varphi_i$ et le cas que $\varphi_j(y_0, \ldots, y_m)$ est $\exists x \varphi_i(x, y_0, \ldots, y_m)$ pour des $i, k < j$, le cas atomique étant évident. Les deux premiers cas sont aussi évidents, et le dernier cas est couvert par la supposition (ii). $\bigstar_{4.5.7}$

Continuons la preuve du théorème :

Preuve du Théorème 4.5.6 Nous n'allons démontrer que (1), car (2) a la même preuve. Étant données $\varphi_0, \varphi_1, \ldots \varphi_n$, nous pouvons supposer que la liste est fermée par les sous-formules, en ajoutant des formules si nécessaire. Remarquons que nous cherchons un α tel que V_α reflète la vérité (c'est à dire, la situation en $\mathbf{V}$) sur chaque φ_i qui est un énoncé (car en ajoutant les sous-formules dans la liste, nous avons maintenant des φ qui ne sont pas des énoncés). Le lemme nous montre qu'il nous suffit de trouver des témoins pour les formules existentielles. Pour trouver un tel α nous allons définir des fonctions qui donnent les témoins.

Pour chaque $i \leq n$ nous définissions $F_i : \mathrm{Ord} \to \mathrm{Ord}$. Si la formule φ_i n'est pas de la forme $\exists x \varphi_i(x, y_0, \ldots, y_m)$, laissons F_i être idéntiquement 0. Autrement, pour chaque $y_0, \ldots, y_m$ pour lequel il existe x tel que $\varphi_i(x, y_0, \ldots, y_m)$, définissons $G_i(y_0, \ldots, y_m)$ comme le minimum η tel que pour un $x \in V_\eta$, nous avons $\varphi_i(x, y_0, \ldots, y_m)$. S'il n'y a pas un tel x, définissons $G_i(y_0, \ldots, y_m) = 0$. Maintenant définissons $F_i(\xi) = \sup\{G_i(y_0, \ldots, y_m) : y_0, \ldots, y_m \in V_\xi\}$. Disons que un ordinal α est fermé pour F_i si pour chaque $\beta < \alpha$,

$F(\beta) < \alpha$. Il convient de vérifier que si α est fermé pour toutes F_i, alors α satisfait les conditions du théorème.

Pour trouver un tel α nous allons discuter une notion qui est très utile dans le contexte d'applications sur les ordinaux.

Définition 4.5.8. *Un ensemble ou une classe C d'ordinaux est* fermé *si pour chaque $\alpha < \sup(C)$, si $\sup(\alpha \cap C) = \alpha$, alors $\alpha \in C$. Pour un ordinal α et un ensemble C, on dit que C est* borné *en α si $\sup(C \cap \alpha) < \alpha$. On définit de façon similaire ce que veut dire pour une classe d'ordinaux d'être bornée dans la classe* Ord *des ordinaux.*

Lemme 4.5.9. *Supposons que $F : \text{Ord} \to \text{Ord}$ est une fonction, alors la classe $\mathcal{F}$ d'ordinaux fermés pour $\mathcal{F}$ est fermée et n'est pas bornée.*

Preuve Supposons que $\sup(\mathcal{F} \cap \alpha) = \alpha$ et démontrons que $\alpha \in \mathcal{F}$. Si $\beta < \alpha$, alors il existe $\gamma < \alpha$ tel que $\beta < \gamma$ et $\gamma \in \mathcal{F}$. Donc $F(\beta) < \gamma < \alpha$. Cela montre que $\mathcal{F}$ est fermée.

Pour voir que $\mathcal{F}$ n'est pas bornée, prenons un ordinal α et trouvons $\alpha^* > \alpha$ tel que $\alpha^* \in \mathcal{F}$. Définissons $\alpha_0 = \alpha + 1$ et

continuons par l'induction sur n. Étant donné α_n, soit $\alpha_{n+1} = \sup\{F(\beta) + 1 : \beta < \alpha\}$. Finalement, soit $\alpha^* = \sup_{n<\omega} \alpha_n$. Si $\beta < \alpha^*$, nous avons que $\beta < \alpha_n$ pour un $n < \omega$, et donc $F(\beta) < \alpha_{n+1} \leq \alpha^*$. $\bigstar_{4.5.9}$

Pour finir la preuve du théorème, définissions $F = \max\{F_i : i \leq n\}$. Il nous suffit de trouver un α qui est fermé par F. L'existence de toute une classe des tels α est impliquée par Lemme 4.5.9. $\bigstar_{4.5.6}$

4.6 Éléments historiques et bibliographiques

L'univers constructible était découvert par Kurt Gödel en [Göd38] pour résoudre le problème du continu. Dans l'article, Gödel montrait que **L** est un modèle intérieur de la théorie ZF et que l'axiome du choix et l'hypothèse généralisée du continuu sont vérifiés dans ce modèle. Ceci prouve que ces deux propositions sont cohérentes avec les axiomes de ZF, à condition que ZF soit déjà cohérente. De nombreux autres théorèmes n'étant applicables que

si l'une ou les deux hypothèses sont vraies, leur cohérence est un résultat très important.

De la façon plus complexe qu'on ne le pense, on peut aussi démontrer (voir [Kun80]) que l'univers constructible satisfait l'axiome de constructibilité, $\mathbb{V} = L$. Cela veut dire que tous les ensembles sont constructibles. Pour citer un blog très pertinent par David Madore [Mad] : « L'axiome de constructibilité rend explicites et 'naturels' des objets dont l'axiome du choix ne fait que postuler l'existence, et il apporte suffisamment d'information sur les parties de ω pour démontrer et éclairer l'hypothèse du continu (et plus généralement sur les parties de n'importe quel ensemble pour l'hypothèse généralisée du continu). L'axiome de constructibilité ordonne (et « déroule ») tout l'univers mathématiques selon les ordinaux et permet donc de répondre à un nombre étonnant de questions combinatoires autrement indécidables, dont la véracité de l'axiome du choix et de l'hypothèse du continu ne sont que deux exemples : ce qui ne veut pas dire, cependant, que nous devions forcément accepter ces réponses comme correctes. »

Chapitre 5

Indépendance

L'idée d'indépendance en mathématiques n'est pas seulement très intéressante philosophiquement, mais elle appartient à une tradition assez ancienne. Qui n'a pas entendu parler de la question de l'indépendance du Postulat des Parallèles d'Euclid, étudié déjà par Gauss ? Nous avons déjà abordé l'indépendance dans notre travail sur les modèles de sous-théories de ZF. Un énoncé φ est indépendant d'une théorie T si T ne peut ni démontrer ni réfuter φ. Au moins dans le contexte de la logique du premier ordre, où nous avons recours à la Solidité et à la Complétude de Gödel, qui dit en

gros que les notions de prouvabilité et de satisfaction sont logiquement équivalentes, pour montrer qu'un énoncé φ est indépendant d'une théorie T, il suffit de trouver un modèle de T où φ est satisfaite et un autre modèle de T où $\neg\varphi$ est satisfaite. Dans le cas des modèles de ZFC il y a une difficulté supplémentaire : nous cherchons une preuve en ZFC, donc par l'Incomplétude de Gödel, nous ne pouvons pas (en ZFC) démontrer qu'il y a même un modèle de ZFC. Alors nous ne pouvons pas démontrer que, par exemple, $ZFC + \neg HC$ est cohérente. Pour cette raison, le mieux que nous pouvons espérer trouver est une cohérence *relative*. Cela veut dire qu'en supposant qu'il existe un modèle de ZFC, nous pouvons en produire un autre qui est un modèle de $ZFC + \neg HC$.

La méthode pour démontrer des résultats de cohérence relative est le plus souvent la méthode de forcing. Elle est un peu trop complexe pour l'aborder en détail dans ce texte. Notre but sera de donner une introduction intuitive. Cette introduction peut ensuite être développée en consultant un de nombreux livres de référence, dont [Kun80], [Jec03] et [Hal12].

Nous allons aborder le thème par un exemple, l'exemple d'arbres.

Définition 5.0.1. *Un* arbre $(T, \leq_T)$ *est un ordre partiel tel que chaque élément* $t \in T$ *a un ensemble bien-ordonné par* $\leq_T$ *des précédents. On dit que le type d'ordre de* $(\mathrm{pre}(t), \leq_T)$ *est la* hauteur *de* t, $\mathrm{ht}(t)$. *Le niveau* α *de* T, $\mathrm{lev}_\alpha(T)$ *est l'ensemble d'éléments de* T *de hauteur* α. *La* hauteur *de* T *est le premier* α *tel que* $\mathrm{lev}_\alpha(T) = \emptyset$.

Une branche *de* T *est une chaine maximale en* T.

Nous allons nous contenter, sans grande perte de généralité, de discuter les arbres avec une *racine* ou *origine*, qui est l'élément unique du niveau 0.

Théorème 5.0.2. *Tout arbre infini à branchement fini a une branche infinie.*

Le branchement d'un arbre est fini si chaque élément a un nombre fini de successeurs immédiats.

Preuve Un nœud est dit *infini* s'il a un nombre infini de successeurs. Supposons qu'on ait un arbre infini T d'origine O à bran-

chement fini. T est infini, donc O est infini. Comme O n'a qu'un nombre fini de successeurs immédiats, l'un d'entre eux au moins, noté $N(1)$, est infini ; sinon O ne serait pas infini. De même l'un au moins, noté $N(2)$, des successeurs immédiats de $N(1)$ est infini. On définit ainsi un nombre infini de nœuds $N(i)$, pour tout entier positif i, qui ensemble forment une branche infinie. $\bigstar_{5.0.2}$

Définition 5.0.3. *Si κ est un cardinal, un* κ-arbre d'Aronszajn *est un arbre T de hauteur κ dont aucune branche n'est de hauteur κ, et tel que tous les niveaux de T sont de cardinalité $< \kappa$; les arbres d'Aronszajn sont les $\aleph_1$-arbres d'Aronszajn, ce qui équivaut à des arbres non dénombrables dont toutes les branches sont dénombrables, et dont chaque élément a au plus un nombre dénombrable de successeurs.*

Le lemme de König dit qu'il n'existe pas de $\aleph_0$-arbre d'Aronszajn. Des arbres d'Aronszajn (c'est-à-dire des $\aleph_1$-arbres d'Aronszajn) furent construits en 1934 par Nachman Aronszajn, ce qui montre que l'analogue du lemme de König ne s'applique pas aux arbres non dénombrables.

L'existence de $\aleph_2$-arbres d'Aronszajn est indécidable (en admettant un certain axiome de grand cardinal) ; plus précisément, l'hypothèse du continu implique l'existence d'un -$\aleph_2$-arbre d'Aronszajn, mais Mitchell et Silver ont montré ([Mit73]), en admettant l'existence d'un cardinal faiblement compact, qu'il est cohérent de supposer qu'aucun $\aleph_2$-arbre d'Aronszajn n'existe.

Théorème 5.0.4. *Il existe un arbre d'Aronszajn.*

Définition 5.0.5. *Un arbre d'Aronszajn T est Souslin ssi chaque antichaine en T est dénombrable.*

Définition 5.0.6. *Un ensemble $S \subseteq \omega_1$ est* stationnaire *ssi $S \cap C \neq \emptyset$ pour chaque club $C \subseteq \omega_1$.*

Le principe 'diamant' dit que

$\diamondsuit$: Il existe une suite $\langle A_\alpha : \alpha < \omega_1 \rangle$ telle que chaque $A_\alpha \subseteq \alpha$ et pour chaque $A \subseteq \omega_1$ l'ensemble $\{\alpha < \omega_1 : A_\alpha = A \cap \alpha\}$ est stationnaire.

Théorème 5.0.7 (Jensen)**.** *(1) $\diamondsuit$ est vrai en* $\mathbf{L}$*.*
(2) $\diamondsuit$ implique l'existence d'un arbre de Souslin.

Preuve Nous ne démontrons que (2), car (1) est un peu dehors des limites de ce livre.

(2) Soit $\langle A_\alpha : \alpha < \omega_1 \rangle$ une suite diamant. Nous allons construire un arbre T avec l'ensemble d'éléments equal à ω_1. La construction va donc établir un ordre $\leq_T$. Cet ordre sera tel que $\alpha <_T \beta \implies \alpha < \beta$, donc nous avons que l'ensemble des précédents de chaque élément β est bien ordonné par $<_T$, car chaque suite infinie descendante en $<_T$ est aussi descendante en $<$. En fait, nous allons nous assurer que les niveaux de T correspondent, comme ensembles aux multiples de ω, donc $\mathrm{lev}_\beta(T) = \{\omega \cdot \beta + n : n < \omega\}$. Alors, chaque niveau de T est dénombrable. Donc par induction sur β nous construisons $T_\beta = <_T \restriction \cup_{\alpha<\beta} \mathrm{lev}_\beta(T)$, d'une façon cohérente, c'est-à-dire que pour $\alpha < \beta$ l'ordre sur $T_\beta \restriction T_\alpha$ est celui que nous avons construit sur T_α. Nous nous assurons aussi que pour chaque $\alpha < \beta$ et t de hauteur α, il y a au moins un t' de hauteur β tel que $t <_T t'$. De cette façon nous sommes sûres que la hauteur de T est ω_1.

A l'étape $\beta + 1$, nous décidons que $\omega \cdot \beta + n <_T \omega \cdot \beta + 2n, \omega \cdot$

$\beta + 2n + 1$.

L'idée principale de la construction se passe aux niveaux limites de l'arbre. Là, nous allons demander

$$T_\beta = \beta \wedge A_\beta \text{ est une antichaine maximale en } T_\beta \implies (\forall t \in \mathrm{lev}_\beta(T))(\exists t' \in A_\beta)(t' <_T t); \tag{5.1}$$

Il est facile de s'assurer dans les étapes limites que l'équation 5.1 est satisfaite. Pour finir la preuve

Lemme 5.0.8. *Si T satisfait l'equation 5.1, alors T est un arbre de Souslin.*

Preuve(du lemme) Montrons d'abord que chaque antichaine de T est dénombrable. Soit A une antichaine de T, supposons non dénombrable, et supposons aussi que A est maximale. Il est facile de démontrer que l'ensemble $C = \{\alpha \text{ limite} : T_\alpha = \alpha \text{ et } A \cap \alpha \text{ est une antichaine maximale en } T_\alpha\}$ est un club de ω_1, donc il existe $\beta \in C$ tel que $A \cap \beta = A_\beta$. Soit t^+ an élément de A de hauteur au moins aussi grande que β est soit t' l'unique élément de T de

hauteur β qui est $\leq_T t^+$. L'équation 5.1 nous assure qu'il y a un t dans A_β tel que $t <_T t'$ et donc, $t \in A \cap \beta$ est comparable avec t^+. C'est une contradiction.

Maintenant, démontrons que T n'a pas de chaines qui ne sont pas dénombrables. Sinon, soit C une chaine qui n'est pas dénombrable et prenons une chaine maximale qui contient C, disons C^*. Alors pour chaque β, C^* a un élément sur le niveau β de T, disons $c_\beta = \omega\beta + n_\beta$. Nous savons qu'au moins un parmi $\omega\beta + 1 + 2n_\beta$ et $\omega\beta + 1 + 2n_\beta + 1$ n'est pas sur C^*, choissisons pour a_β un tel élément. Maintenant, il nous suffit d'observer que $\{a_\beta : \beta < \omega_1\}$ est une antichaine, une contradiction. $\bigstar_{5.0.8}$

$\bigstar_{5.0.7}$

On peut bien remarquer que le principe $\diamondsuit$ implique HC.

Théorème 5.0.9. *Le principe $\diamondsuit$ implique que $|\mathcal{P}(\omega)| = \aleph_1$.*

Preuve Supposons que $\langle A_\alpha : \alpha < \omega_1 \rangle$ est une suite diamant. Pour chaque $A \subseteq \omega$, en remarquant que c'est également le cas que $A \subseteq \omega_1$, nous pouvons trouver un ensemble stationnaire S_A tel que pour chaque $\alpha < \omega_1$ nous avons $A \cap \alpha = A_\alpha$. En étant stationnaire,

S_A est cofinal en ω_1 et donc il contient un élément $\alpha > \omega$. Alors $A = A \cap \alpha = A_\alpha$. Cela implique que tous les sous-ensembles de ω se trouvent parmi les éléments de $\{A_\alpha : \alpha < \omega_1\}$. $\bigstar_{5.0.9}$

Par contre, par la méthode de forcing nous pouvons trouver un modèle de ZFC où il n'y a pas d'arbres de Souslin.

5.1 Le Forcing

En écrivant cette section, nous avons privilégié une description peu technique, mais courte et intuitive. Les détails sont beaucoup trop complexes pour cette présentation et, de plus, très bien expliqués dans les livres mathématiques, tel que [Kun80].

Quand on construit un objet par l'induction ou par l'induction transfinie, on peut considérer la construction comme une suite totalement ordonnée d'approximations de l'objet qu'on est en train de construire. Par exemple, dans la construction de l'arbre de Souslin dans le Théorème 5.0.7, l'arbre T est construit comme la réunion d'approximations $\langle T_\alpha : \alpha < \omega_1 \rangle$. Il est évident qu'une telle réunion

est bien définie et qu'elle forme un arbre.

Imaginons maintenant une construction où nous construisons un objet par l'induction sur un ordre qui est plutôt partiel. Comme example, considérons l'arbre T construit dans le Théorème 5.0.7, qui est en particulier un ordre partiel. Considérons les chaines en T. Parmi ces chaines on en trouve certaines qui ont une majorante, mais ce n'est pas le cas que chaque chaine en T a une majorante. Par exemple, si nous regardons le niveau ω de T, il contient un ensemble dénombrable d'éléments. Pour chaque tel élément t nous pouvons associer à t, la chaine des précédents de t en T et cette chaine a alors une majorante, par exemple t. Mais on peut former 2^ω chaines en T en utilisant les éléments de T_ω et seulement un nombre dénombrable parmi eux va avoir une majorante. Nous pouvons également observer qu'une chaine de longueur ω_1 en T ne peut pas exister.

Mais, soyons absolument déterminés à construire une chaine de longueur ω_1 en T ! Appelons la G. Nous pouvons imaginer que l'arbre T présente un ensemble d'approximations d'une telle chaine

G : chaque élément p de T détermine la chaine qui est formée par les précédents de T et si nous avons que $p \in G$, alors G contient tous les précédents de p. Si $p \leq q$ et $q \in G$, alors q détermine encore plus de G que p. Donc notre essai consiste à construire G à partir de T en choisissant les approximations, c'est-à-dire, les éléments de T qui vont former G d'une façon judicieuse, de telle sorte qu'à la fin nous obtiendrons une branche de T qui n'est pas dénombrable. Mais T n'a pas de telle branche ! Si, pourvu qu'on change l'univers de la théorie des ensembles. Le prix à payer ...

Pour expliquer cela dans des termes moins dramatiques, imaginons que nous sommes dans un univers V de la théorie des ensembles, autrement dit, nous supposons que la théorie des ensembles est cohérente et que nous sommes dans un modèle de ZFC, qui s'appelle V. Nous pouvons même, par exemple en nous concentrant sur le $\mathbf{L}$ d'un modèle de ZFC, supposer que notre modèle V possède un arbre de Souslin, disons T. Comme V est un ensemble, il n'est pas égal à la classe de tous les ensembles. La méthode de forcing nous permet de construire un ensemble G qui n'est

pas en V et tel qu'à partir de V et G nous pouvons construire un autre modèle $V[G]$ de ZFC de telle façon que dans $V[G]$, l'arbre T est toujours un ω_1-arbre, mais cette fois avec une branche qui n'est pas dénombrable.

Maintenant, si nous pouvons itérer cette procédure de façon à nous occuper des tous les arbres de Souslin, y compris ceux qui ont apparu pendant l'itération, nous allons avoir un modèle de ZFC où il n'y a pas d'arbres de Souslin. Il se trouve qu'une telle itération est, effectivement, possible, ce qui montre que la cohérence de ZFC implique la cohérence de 'ZFC + il n'y a pas d'arbres de Souslin'.

Le théorème à l'origine de la méthode de forcing appartient à Paul Cohen, qui a démontré que l'hypothèse de continu n'est pas impliquée par les axiomes ZFC. En faisant cela, il a supposé la cohérence de ZFC pour commencer avec un modèle V de ZFC. Il a utilisé la méthode de forcing et l'ordre partiel

$$P = \{f : \ f : \omega \to 2, f \text{ fonction partielle}, |f| < \aleph_0\},$$

ordonné par $\subseteq$. Un filtre générique de P, dont nous ne citons pas

la définition formelle, produit une fonction de ω à 2 qui n'est pas en V. Si on utilise un ordre un peu plus complexe

$$P = \{f : f : (2^{\aleph_0{}^+})^V \times \omega \to 2, f \text{ fonction partielle}, |f| < \aleph_0\},$$

on obtient $(2^{\aleph_0{}^+})^V$ nouvelles fonctions de ω à 2, ce qui implique que $(2^{\aleph_0})^{V[G]} \geq (2^{\aleph_0{}^+})^V$ et le modèle $V[G]$ ne peut pas alors satisfaire l'hypothèse du continu.

5.2 Éléments historiques et bibliographiques

Le principe $\diamondsuit$ fait partie de la structure fine de l'univers constructible, découverte par Ronald Jensen [Jen72], justement pour résoudre le problème d'existence d'un arbre de Souslin. Cette structure a développé sa propre vie et elle est à la base de toute la théorie des modèles intérieurs.

La méthode de forcing a été découverte par Paul Cohen in 1963, publié en [Coh66]. Il a utilisé cette méthode pour démontrer l'indépendance de HC, c'est-à-dire pour montrer que la cohérence de

ZFC implique la cohérence de $ZFC + \neg HC$. Pour ce résultat sur l'hypothèse du continu, Cohen a gagné la médaille Fields en 1966, ainsi que la National Medal of Science en 1967. La découverte était très surprenante et originale, même si dans sa mise en œuvre Cohen a beaucoup bénéficié de ses conversations avec Solomon Feferman et Kurt Gödel. La méthode est beaucoup plus générale que le résultat de l'indépendance de l'hypothèse du continu et ses vraies limites étaitent avancées par Robert Solovay dans les années 1960 et 1970. Parmi ses travaux les plus importants, on trouve la preuve que, sous l'hypothèse de l'existence d'un cardinal inaccessible, la mesurabilité (au sens de Lebesgue) de toutes les parties de $\mathbb{R}$ est cohérente avec la théorie ZF sans l'axiome du choix [Sol70]. Le travail de Solovay est toujours la base de la théorie des ensembles modernes. La méthode de l'itération de forcing, qui est nécessaire pour le résultat de l'indépendance de l'existence d'un arbre de Souslin, était développée par Solovay et Tennenbaum [ST71]. Des découvertes importantes dans la théorie de forcing ont été faites par James Baumgartner, Saharon Shelah (l'invention du proper forcing

[She98]), Hugh Woodin et d'autres.

Essayant d'être Souslin

Troisième partie

Quelques Corrigés

5.3 Chapitre 1

1 Soit A un ensemble dénombrable. S'il est fini, chacun de ses sous-ensembles est fini et donc dénombrable. Supposons donc que A est infini dénombrable et que B est un sous-ensemble de A, lui-même infini. Comme A est dénombrable, il y a une bijection f de A à $\mathbb{N}$ et, en réduisant f à B, nous obtenons une injection de B à $\mathbb{N}$. Maintenant nous allons démontrer l'existence d'une injection de $\mathbb{N}$ à A et notre démonstration s'achèvera grâce au lemme de Schroeder-Bernstein. Soit $g(0)$ un élément quelconque de B, qui existe car B n'est pas fini, donc il n'est pas vide. Par l'induction sur $n \in \mathbb{N}$, étant donnés $g(0), g(1), \ldots, g(n)$, nous observons qu'ils n'épuisent pas la totalité de B, car B est infini. Donc nous pouvons trouver un élément $g(n+1)$ de B qui n'est pas parmi $g(0), g(1), \ldots, g(n)$ et l'induction continue.

2 Nous avons déjà démontré que $\mathbb{Q}$ est dénombrable, donc son sous-ensemble $\mathbb{Z}$ est dénombrable lui aussi.

3 Il nous suffit de produire deux injections $f : \mathbb{N} \to (\mathbb{N} \times \mathbb{N})$ et $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. Nous pouvons utiliser $f : n \to (n, 0)$ et $g : (a, b) \to 2^a \cdot 3^b$.

4 Nous allons présenter le cas principal qui est le cas du produit $A_0 \times A_1 \times \ldots A_n$ d'ensembles infinis dénombrables, où $n \geq 1$ est un entier naturel. En utilisant une bijection $f_i : A_i \to \mathbb{N}$, pour chaque $i \leq n$ nous pouvons facilement construire une bijection entre $A_0 \times A_1 \times \ldots A_n$ et $\mathbb{N}^{n+1}$, donc il nous suffit de démontrer par l'induction sur $n \geq 1$ que $\mathbb{N}^{n+1}$ est dénombrable. Nous avons déjà démontré le cas de base $n = 1$. Pour le cas de $n+1$, représentons $\mathbb{N}^{n+1}$ comme le produit $\mathbb{N}^n \times \mathbb{N}$. Par l'hypothèse de l'induction, nous avons une bijection entre $\mathbb{N}^n$ et $\mathbb{N}$, que nous pouvons utiliser pour produire une bijection entre $\mathbb{N}^{n+1}$ et $\mathbb{N}^2$. Il nous suffit donc de démontrer que $\mathbb{N}^2$ est dénombrable. Il est évident qu'il y a une injection de $\mathbb{N}$ à $\mathbb{N}^2$, par exemple la fonction $f : n \to (n, 0)$. La fonction $f(m, n) = 2^m \cdot 3^n$ est une injection de $\mathbb{N}^2$ à $\mathbb{N}$. L'existence de ces deux injections montre qu'il y a une bijection entre $\mathbb{N}^2$ et $\mathbb{N}$, par le lemme de

Schroeder-Bernstein.

5 $\sqrt{2}$ est irrationnel mais algébrique, car $x = \sqrt{2}$ est une solution d'équation $x^2 - 2 = 0$.

6 Le nombre des équations algébriques de dégré n pour chaque n est dénombrable, car le nombre des suites de longueur n de rationnelles est dénombrable, et chaque équation est uniquement représentée par la suite de ses coefficients. Une équation de degré n peut avoir au plus n solutions, donc l'ensemble A_n des réels qui sont les solutions des équations de degré n, pour chaque n, est dénombrable. L'ensemble des nombres algébriques $A = \bigcup_{n\in\mathbb{N}} A_n$ est donc lui-même dénombrable. L'ensemble des nombres transcendantaux est $T = \mathbb{R} \setminus A$, donc il ne peut pas être dénombrable car $\mathbb{R}$ n'est pas dénombrable.

5.4 Chapitre 2

1 Soit $m \leq^* n$ pour $m, n < \omega$ si $m = n \,(\mathrm{mod}\, 2)$, ou bien si m est pair et n est impair.

2 Soit $f : \alpha \to \omega$ une bijection. Pour $m, n < \omega$ définissons $m <_\alpha n$ ssi $f^{-1}(m) < f^{-1}(n)$.

3 (1) Comme x est transitif et $y \in x$, nous avons $y \subseteq x$. Donc, en étant un sous-ensemble d'un ensemble bien ordonné par $\in$, y lui-même est bien ordonné par $\in$. Si $z \in y$ et $w \in z$ nous avons $z \in x$ (car $y \subseteq x$) et donc $z \subseteq x$ (car x est transitif), ce qui implique $w \in x$. L'ensemble x est bien ordonné par $\in$ et nous avons $w, y \in x$. Donc soit $w = y$, soit $y \in w$, soit $w \in y$. Si $y \in w$, nous avons $y \in w, w \in z$ et donc $y \in z$ car $\in$ est transitive sur x. Mais dans ce cas, $\{y, z\}$ n'a pas le plus petit élément, contradiction. En utilisant la même logique, nous pouvons voir que la supposition $w = y$ donne une contradiction, qui nous laisse avec $w \in y$. Cela montre que $z \subseteq y$ et donc y est transitif.

(2) Si $x = \emptyset$, la conclusion est évidente. Autrement, soit f l'unique (par la trichotomie) isomorphisme entre x et y. Nous allons démonter que f est l'identité. Supposons que non et soit z le $\in$-plus petit élément de x tel que $f(z) \neq z$. Pour chaque

$w \in z$ nous avons $w \in x$ (car x est un ordinal) et, par le choix de z, nous avons $f(w) = w$. Donc $z = \{w : w \in z\} = \{f(w) : w \in z\} = f(z)$, car f est un isomorphisme. C'est une contradiction avec le choix de z.

(3) Par la trichotomie, si $x \neq y$, nous avons que x est isomorphe avec une partie initiale de y ou que y est isomorphe avec une partie initiale de z. Disons que x est isomorphe avec une partie initiale de y. Nous pouvons vérifier que la transitivité de y implique que chaque partie initiale de y est un ordinal de la forme $z \in y$. Donc x est isomorphe avec z, qui implique par (2) que $x = z$ et donc $x < y$. La possibilité que y soit isomorphe avec une partie initiale de z nous donne, de la façon similaire, que $y < x$.

(4) Supposons, pour la simplicité, que $x < y$ et $y < z$, les possibilités restantes étant très similaires. Donc nous avons $y \in z$ par la définition de $<$ et $y \subseteq z$ par la transitivité. Alors $x < y$ implique que $x \in y$ et donc $x \in z$, ce qui implique $x < z$.

(5) Choisissons $x \in C$. Si x est le plus petit élément de C, nous

avons trouvé un tel élément. Autrement, $\{y : y \in x \cap C\}$ est un sous-ensemble non vide de x et donc il a un plus petit élément, disons z. Nous allons démontrer que z est le plus petit élément de C. Si ce n'est pas le cas, il y a $w \in C$ tel que $w \in z$, ce qui implique par la transitivité de x que $w \in x$. Donc $w \in x \cap C$, ce qui est une contradiction avec le choix de z.

4 Il s'agit d'une lecture attentive de la définition : $(\kappa + \lambda) + \theta$ est bijectif avec $(\kappa + \lambda) \times \{1\} \cup \theta \times \{2\}$, qui est, à son tour, bijectif avec $\kappa \times \{0\} \cup \lambda \times \{1\}$, donc $(\kappa + \lambda) + \theta$ est bijectif avec $\kappa \times \{0\} \cup \lambda \times \{1\} \cup \theta \times \{2\}$. Une autre application de la même logique nous assure que $\kappa + (\lambda + \theta)$ est bijectif avec le même ensemble.

5 Supposons que α n'est pas un cardinal, ce qui implique qu'il y a un $\beta < \alpha$ tel que β est bijectif avec α et donc aussi avec A. Une telle bijection f induit un bon ordre $<_A$ de type β sur A, et définissant $a <_A b$ ssi $f^{-1}(a) < f^{-1}(b)$. C'est une contradiction avec le choix de α.

Quatrième partie

La suite

Chapitre 6

À lire

Pour approfondir l'étude philosophique de la théorie des ensembles, ainsi que pour comprendre d'autres directions de recherche, voici quelques suggestions incontournables.

6.1 Alain Badiou

Ce philosophe français, né en 1937, à fait appel à la logique mathématique, seule capable, selon lui, de déployer l'ontologie. Il soutient que l'ontologie, théorie de l'être, est identique aux mathématiques et, spécifiquement, à la théorie des ensembles. Il associe

également la phénoménologie, étude des degrés de l'apparaître et de l'événement, à la théorie des topoi. Il a beaucoup écrit sur ces sujets, notamment les deux livres [Bad88] et [Bad06]. Une des thèses qui émerge de ces textes est que, du point de vue de l'être, rien n'appartient à soi, ce qui est une conséquence de l'Axiome de Fondation et que l'événement n'est possible que s'il y a précisément une telle auto-appartenance. En conséquence, l'événement n'est pas. Il exploite la théorie de forcing pour une conception de la vérité.

6.2 Fraenkel, Bar-Hillel et Lévy

Ces trois auteurs ont entrepris un livre de mise au point des fondements des mathé-matiques dans l'année 1973 [FBHL73]. Ils expliquent la crise des fondements au début du vingtième siècle et sa résolution par Zermelo et Fraenkel. En suite, ils expliquent des approches de la théorie des ensembles ZFC et les nouveaux fondements de Quine, mais également l'approche de la théorie des types et des conceptions intuitionistiques. Le livre est très détaillé, sans

être trop technique. Dans le monde d'aujourd'hui il nous manque un livre de cet ampleur sur notre propre état dans les fondements.

6.3 Thomas Jech et l'Encyclopédie

S'il nous manque un livre de mise au point dans les fondements, dans la théorie des ensembles il y a le livre de Jech [Jec03] et celui, encore plus récent, [MF10]. Armé de ces deux livres, on voit un panorama de la théorie des ensembles moderne et plusieurs directions de recherche. La théorie des ensembles y est présentée dans toute sa complexité et toute sa gloire. Un autre livre, plus adapté à l'apprentissage, très bien écrit et amusant, est le livre de Lorenz Halbeisen [Hal12].

6.4 La philosophie des ensembles

Le rapprochement entre la philosophie et la théorie des ensembles est très en vogue dans les cercles mathématiques et nous sommes plusieurs à croire que pour pouvoir poser des bonnes ques-

tions en théorie des ensembles nous devrons intégrer la motivation philosophique. Ce genre de recherche est représenté à Harvard par Peter Koellner et Hugh Woodin, à Berkeley par John Steel, à CUNY par Joel Hamkins, à Helsinki par Juliette Kennedy et Jouko Väänänen, à Amsterdam et Hamburg par Benedikt Loewe, pour nommer que quelques personnes très actives. À Paris il y a un institut qui s'est depuis l'années 1930 consacré au rapprochement entre les sciences exactes et les sciences humaines, l'Institut d'histoire des sciences et des techniques à l'Université Paris 1 Panthéon-Sorbonne. La mission de l'Institut, comme noté dans les textes de 1932 est « Créer un institut où collaboreront des professeurs de la Faculté des sciences et du Collège de France » , « Faire travailler ensemble étudiants de laboratoires et étudiants d'histoire ou de philosophie », « Refuser la division des études entre lettres et sciences ». Le travail en fondements des mathématiques, dont le nôtre, s'intègre très naturellement dans ce contexte, toujours représenté par l'IHPST.

6.5 Les fondements non-ensemblistes

Il n'est pas question de discuter des fondements des mathématiques sans noter qu'il n'y a pas une seule façon de fonder les mathématiques et, en fait, qu'il n'est pas possible en ce moment d'avoir une vision totalement ontologique de ce qu'on veut dire par les fondements. La théorie des ensembles est une bonne façon de fonder une grande partie des mathématiques, mais pas toutes les mathématiques. Par exemple, dans la géométrie algébrique il est beaucoup plus naturel d'utiliser les catégories, comme le montre le travail de Grothendieck et d'autres après lui. Dans le contexte de la vérification automatique de preuves, il s'agit très naturellement de travailler avec les fondements constructivistes, notamment avec les fondements univalents de Voevodsky, qui sont assez récents. Comment définir les fondements aujourd'hui revient à une question profondément philosophique. Mon opinion à ce sujet est expliquée dans l'article [Dža17], où on peut également trouver d'autres références.

Chapitre 7

Quelques problèmes pour s'entrainer

Les problèmes proposés dans les livres de la théorie des ensembles pour les mathématiciens ont la tendance d'être trop techniques pour les philosophes, car ils demandant des connaissances mathématiques trop détaillées et il ne discutent pas des connaissances philosophiques. Nous proposons ici une petite sélection de problèmes plus abordables, mais assez utiles pour approfondir ses connaissances de la matière traitée. Il s'agit d'un vrai examen passé

par les étudiants à Paris 1.

2016–2017

LoPhiSC M1

Théorie des Ensembles,

par Mirna Džamonja

Durée: 120 Minutes.

1.(i) Donnez la définition complète de ce que on veut dire par un ensemble dénombrable.

(ii) Démontrez que l'ensemble des sous-ensembles finis d'un ensemble dénombrable est dénombrable.

(iii) Donnez la définition complète d'un ordinal.

(iv) Quelle est la cardinalité de l'ensemble des ordinaux dénombrables ?

2.(i) Donnez la définition complète d'un cardinal.

(ii) Si κ est un cardinal, démontrez que κ est un ordinal limite.

(iii) Définissons la *cofinalité* $\mathrm{cf}(\alpha)$ d'un ordinal α comme le plus petit ordinal $\beta \leq \alpha$ tel qu'il existe une fonction $f : \beta \to \alpha$ avec $\sup(\mathrm{ran}(f)) = \alpha$. Disons qu'un cardinal κ est *régulier* si $\mathrm{cf}(\kappa) = \kappa$, sinon disons que κ est singulier.

— Démontrez que $\aleph_0$ est régulier.

— Démontrez que $\aleph_1$ est régulier.

— Donnez un exemple d'un cardinal singulier et justifiez votre réponse.

3.(i) Donnez une définition complète d'un *ordre partiel* et d'un *ordre linéaire (ou total)* sur un ensemble A.

(ii) Notons par ${}^{\omega}\omega$ la famille des fonctions de ω à ω. Définissons la relation $\leq^*$ sur ${}^{\omega}\omega$ par $f \leq^* g$ ss'il existe $n \in \omega$ tel que pour tout $m \geq n$ nous avons $f(m) \leq g(m)$. Montrez que $\leq^*$ est réflexive et transitive.

(iii) Donnez un exemple qui montre que $\leq^*$ n'est pas un ordre partiel sur ${}^{\omega}\omega$.

NOUVELLE PAGE

(iv) Supposez que $f_n (n < \omega)$ est une suite de fonctions en ${}^{\omega}\omega$ telle que pour tout n nous avons $f_n \leq^* f_{n+1}$. Montrez qu'il existe une fonction $g \in {}^{\omega}\omega$ qui vérifie $f_n \leq g^*$ pour tout n.

4. Un ensemble ordonné tel que tout sous-ensemble totalement ordonné possède un majorant est appelé ensemble inductif. Le lemme de Zorn énonce que tout ensemble inductif admet au moins un élément maximal.

(Ce lemme est un des équivalents de l'Axiome du Choix).

(i) Un *filtre* sur ω est une famille $\mathcal{F} \subseteq \mathcal{P}(\omega)$ non vide qui vérifie :

— $\emptyset \notin \mathcal{F}$,

— Si $A \subseteq B \subseteq \omega$ et $A \in \mathcal{F}$, alors $B \in \mathcal{F}$.

— Si $A, B \in \mathcal{F}$, alors $A \cap B \in \mathcal{F}$.

Un filtre $\mathcal{U}$ sur ω est un *ultrafiltre* sur ω ssi pour chaque $A \subseteq \omega$, soit $A \in \mathcal{U}$, soit $\omega \setminus A \in \mathcal{U}$.

— Donnez un exemple d'un filtre sur ω.

(ii) Démontrez qu'un filtre $\mathcal{F} \subseteq \mathcal{P}(\omega)$ est un ultrafiltre ssi il est maximal, c'est-à-dire qu'il n'existe pas un filtre $\mathcal{F}^* \subseteq \mathcal{P}(\omega)$ tel que $\mathcal{F}$ est un sous-ensemble de $\mathcal{F}^*$ et $\mathcal{F} \neq \mathcal{F}^*$.

(iii) En utilisant le lemme de Zorn, démontrez que pour chaque filtre $\mathcal{F}$ sur ω il existe un ultrafiltre $\mathcal{U}$ sur ω tel que $\mathcal{F} \subseteq \mathcal{U}$.

5.(i) Expliquez le paradoxe de Russell.

(ii) Expliquez le paradoxe de Burali-Forti.

(iii) Énoncez le *Schéma d'axiomes de Compréhension.*

(iv) Expliquez comment le Schéma d'axiomes de Compréhension a résolu les paradoxes en (i) and (ii).

La FIN

Bibliographie

[Bad88] Alain Badiou, *L'être et l'événement*, Seuil, 1988.

[Bad06] ———, *Logiques des mondes. L'Etre et l'Evénement, 2*, Seuil, 2006.

[Bad10] Alain Badiou, *Le fini et l'infini*, Les Petites Conférences, Bayard, 2010.

[Bol51] Bernard Bolzano, *Paradoxien des Unendlichen*, Leipzig : C. H. Re-clamsen, 1851.

[Bou70] Nicolas Bourbaki, *Théorie des ensembles*, Eléments de mathématique, Diffusion CCLS Paris, 1970.

[Can70a] George Cantor, *Beweis, dass eine für jeden reellen Werth von x durch eine trigonometrische Reihe gegebene Function $f(x)$ sich nur auf eine einzige Weise in dieser Form darstellen lässt*, J. Reine Angew. Math. **72** (1870), 139–142. MR 1579489

[Can70b] ———, *Ueber einen die trigonometrischen Reihen betreffenden Lehrsatz*, J. Reine Angew. Math. **72** (1870), 130–138. MR 1579488

[Can72] ———, *Ueber die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen*, Math. Ann. **5** (1872), no. 1, 123–132. MR 1509769

[Coh66] Paul Cohen, *Set theory and the continuum hypothesis*, Benjamin, New York, 1966.

[Dau79] Joseph Warren Dauben, *George Cantor, his mathematics and philosophy of the infinite*, paperback edition 1990 ed., Princeton University Press, 1979.

[Dža17] Mirna Džamonja, *Set theory and its place in the foundations of mathematics-a new look at an old question*, Journal of the Indian Council of Philosophical Research (2017).

[End77] Herbert B. Enderton, *Elements of set theory*, Academic Press, New York, 1977.

[FBHL73] Abraham A. Fraenkel, Yehoshoua Bar-Hillel, and Azriel Lévy, *Foundations of set theory (2nd revised edition)*, North-Holland, Amsterdam, 1973.

[Göd30] Kurt Gödel, *Die Vollständigkeit der Axiome des logischen Funktionenkalküls*, Monatsh. Math. Phys. **37** (1930), no. 1, 349–360. MR 1549799

[Göd38] Kurt Gödel, *The consistency of the axiom of choice and of the generalized continuum hypothesis*, Proc. Nat. Acad. Sci. **24** (1938), 556–557.

[Hal97] Paul R. Halmos, *Introduction à la théorie des ensembles*, Les Grands Classiques Gauthier-Villars, Jacques Gabay, 1997.

[Hal12] Lorenz J. Halbeisen, *Combinatorial Set Theory, with a gentle introduction to forcing*, Springer Monographs in Mathematics, Springer, London, 2012.

[Jec03] Thomas Jech, *Set theory*, 3rd millenium ed., Springer-Verlag, Berlin Heidelberg, 2003.

[Jen72] Ronald Björn Jensen, *The fine structure of the constructible hierarchy*, Ann. Math. Logic **4** (1972), 229–308 ; erratum, ibid. 4 (1972), 443, With a section by Jack Silver.

[KG10] Jean-Michel Kantor and Loren Graham, *Au nom de l'infini. une histoire vraie de mysticisme religieux et de création mathématique*, Éditions Belin, 2010.

[KL87] Alexander S. Kechris and Alain Louveau, *Descriptive set theory and the structure of sets of uniqueness*, London Mathematical Society Lecture Notes Series, vol. 128, Cambridge University Press, Cambridge, 1987.

[Kri07] Jean-Louis Krivine, *Théorie des ensembles*, 2nd ed., Nouvelle bibliothèque mathématique, vol. 5, Cassini, Paris, 2007.

[Kri09] Jean-Louis Krivine, *Logique et théories axiomatiques*, Disponible sur le site web d'auteur, version mai 2009.

[Kun80] Kenneth Kunen, *Set theory*, Studies in Logic and the Foundations of Mathematics, vol. 102, North-Holland Publishing Co., 1980.

[Mad] David Madore, *Pourquoi l'univers constructible de Gödel est important mathématiquement et philosophiquement*, a blog entry.

[MF10] Akihiro Kanamori (eds.) Matthew Foreman (ed.), *Handbook of set theory*, 1 ed., Springer Netherlands, 2010.

[Mit73] William Mitchell, *Aronszajn trees and the independence of the transfer property*, Ann. Math. Logic **5** (1972/73), 21–46. MR 0313057

[Mon57] Richard Montague, *Contributions to the axiomatic foundations of set theory*, Ph.D. thesis, Berkeley University, 1957.

[Mos49] Andrzej Mostowski, *An undecidable arithmetical statement*, Fund. Math. **36** (1949), 143–164. MR 0035721

[RR70] Herman Rubin and Jean E. Rubin, *Equivalents of the axiom of choice*, North-Holland Publishing Co., Amsterdam-London, 1970, Studies in Logic and the Foundations of Mathematics. MR 0434812

[RR85] ———, *Equivalents of the axiom of choice. II*, Studies in Logic and the Foundations of Mathematics, vol. 116, North-Holland Publishing Co., Amsterdam, 1985. MR 798475

[She98] Saharon Shelah, *Proper and improper forcing*, second ed., Perspectives in Mathematical Logic, Springer-Verlag, Berlin, 1998.

[Sol70] Robert M. Solovay, *A model of set-theory in which every set of reals is Lebesgue measurable*, Annals of Mathematics, Second Series **92** (1970), 1–56.

[ST71] Robert M. Solovay and Stanley Tennenbaum, *Iterated Cohen extensions and Souslin's problem*, Ann. of Math. **94** (1971), no. 2, 201–245.

[Wag93] Stanley Wagon, *The Banach-Tarski paradox*, Cambridge University Press, 1993.

[Zer08] E. Zermelo, *Untersuchungen über die Grundlagen der Mengenlehre. I*, Math. Ann. **65** (1908), no. 2, 261–281. MR 1511466

Index