Avant-propos

Une théorie magnifique, mais objet de bien des malentendus

La théorie des ensembles est l'exploration, à l'aide des outils des mathématiques, de la notion d'infini. L'idée de l'infini, partagée par la plupart des humains bien que non évidemment inscrite dans l'univers physique qui nous entoure, est objet de fascination. Développée depuis un siècle et demi, la théorie des ensembles a permis de découvrir de nombreux aspects de l'infini, en particulier l'existence d'une multitude d'infinis différents les uns des autres mais entretenant entre eux des liens subtils. C'est une magnifique théorie, qui fait appel à des ressources variées et à des méthodes extrêmement sophistiquées, une théorie dont les mots eux-mêmes font rêver : hypothèse du continu, argument diagonal, paradoxe de Russell, théorème de non-définissabilité de la vérité, méthode du forcing, ultrapuissances itérées, cardinaux inaccessibles, cardinaux géants, cardinaux indescriptibles, réel $0^\sharp$, modèle $\boldsymbol{L}^{\mathrm{ultime}}$, ...

Pour autant, cette théorie est méconnue et objet de plusieurs malentendus. Le premier, et le plus grave, trouve son origine dans les succès mêmes de l'entreprise. Lors de la crise des fondements qui a secoué les mathématiques au début du XX$^{\mathrm{e}}$ siècle, il est apparu que la théorie des ensembles, une fois installée sur une base axiomatique ferme, pouvait également, et même si ce n'était pas sa vocation première, servir de base fondationnelle pour la totalité de l'édifice mathématique : (presque) tout objet mathématique peut être représenté par un ensemble. De ce fait, on a un temps imaginé de faire jouer à la théorie des ensembles le rôle d'une sorte de théorie universelle englobant toutes les autres et à l'intérieur de laquelle toutes les mathématiques, et même peut-être l'enseignement des mathématiques, devraient être pensés. Cette conception réductrice a fait long feu, mais elle a durablement brouillé l'image de la théorie des ensembles : une fois dissipée la fascination pour un formalisme obscur et constatée l'évidence que la théorie des ensembles restait impuissante dans la plupart des domaines, cette vision biaisée a détourné l'attention de ce qui est et reste la vocation première de la théorie, à savoir explorer l'infini.

Le second malentendu, lui aussi né des réussites de la théorie des ensembles, est la croyance que celle-ci s'est achevée en 1963, lorsque Paul Cohen a établi que l'hypothèse du continu n'est pas prouvable à partir des axiomes du système de Zermelo–Fraenkel. Certains mathématiciens ont pensé, à tort, que ce résultat magnifique marque la fin de la théorie, et que les questions laissées ouvertes le resteront à jamais, dans un mystérieux statut ni vrai-ni faux indiquant peut-être qu'il s'agit de problèmes vides de sens. La théorie des ensembles serait ainsi comme une étoile un temps brillante mais désormais éteinte. Quelle erreur ! Ce qu'a vérifié Paul Cohen est que le système de Zermelo–Fraenkel est incomplet, contribuant ainsi non pas à fermer, mais à ouvrir la théorie, mise en demeure de compléter une axiomatisation que personne n'avait jamais prétendue exhaustive. Un demi-siècle plus tard, les problèmes ne sont pas tous réglés, mais il est indéniable que l'on en sait beaucoup plus qu'aux temps de Cohen, et *a fortiori* de Zermelo et Fraenkel.

Les objectifs de ce texte

Ce texte est une introduction à la théorie des ensembles d'aujourd'hui. Son but est d'*expliquer* les bases de cette théorie — tout au moins ce que l'auteur du texte en a compris ! Il n'est évidemment pas question d'offrir un tableau complet : les deux mille pages du Handbook [34] n'y suffisent pas, qui pourtant sont probablement inaccessibles à la plupart des débutants. En pratique, notre ambition sera d'arriver à la compréhension des trois résultats que l'on peut estimer être les plus marquants de la théorie au XX[e] siècle :

- la consistance de l'axiome du choix et de l'hypothèse du continu par rapport au système de Zermelo-Fraenkel, établie par K. Gödel en 1938,
- la consistance de la négation de l'hypothèse du continu par rapport au système de Zermelo-Fraenkel, établie par P. Cohen en 1963,
- l'analyse exhaustive des ensembles de cardinalité dénombrable telle qu'elle résulte des théorèmes sur la détermination projective, établis par D. A. Martin, J. Steel, et H. Woodin en 1985 et 1987.

Pour cela, on essaiera de prendre le lecteur par la main depuis le début, pour le mener aux résultats ci-dessus et à un point d'où, ensuite, il pourra facilement continuer dans les textes pour spécialistes. On aimerait combler ici le fossé séparant, d'une part, les nombreux textes existants exclusivement élémentaires et destinés aux non-spécialistes, qui souvent reflètent une vision datée et depuis longtemps dépassée, et, d'autre part, des ouvrages spécialisés d'un abord plus difficile comme [58] ou [59].

Chemin faisant, on aura à découvrir bien des outils et des méthodes : théorèmes de complétude et d'incomplétude, modèles intérieurs, forcing et extensions génériques, axiomes de grands cardinaux, propriété de détermination... À chaque fois, on essaiera de ne pas rester à un niveau exclusivement technique et de mettre en perspective les résultats avec les objectifs de la théorie, en s'interrogeant sur leur signification et leur portée. La théorie des ensembles est souvent remise en question aujourd'hui en tant que théorie fondationnelle, et c'est une question intéressante à discuter à la lueur de résultats précis. Sans prétendre à une quelconque compétence en philosophie, on a néanmoins inclus quelques remarques naïves dans cette direction.

Notre espoir est que des publics variés puissent trouver leur pitance à travers le copieux menu qui est proposé. Le lecteur découvrira vite qu'il existe plusieurs niveaux de lecture possibles, suggérés par la typographie : les (abondants) textes en petits caractères sont soit des démonstrations, soit des commentaires, discussions, compléments, non nécessaires à la lecture du texte principal, mais supposés l'éclairer. On espère que la diversité des formats facilitera la navigation.

Organisation du texte

Le texte comporte trois parties distinctes tant par leur contenu que par leur style. Dans la partie A (chapitres I à V), on expose la théorie élémentaire des ensembles sans supposer aucun prérequis, en particulier de logique. L'objectif est de présenter une approche aussi peu dogmatique que possible de la théorie, en s'efforçant au fur et à mesure de justifier les options adoptées et en évoquant à l'occasion certaines des pistes alternatives non retenues. Comme dans toutes les mathématiques, quelques bribes de logique apparaissent ici ou là, mais uniquement comme une sténographie. Les arguments de cette partie résolument « pré-logique » sont donc principalement combinatoires. Malgré ces limitations, on verra que des résultats hautement non triviaux peuvent être établis au prix d'arguments délicats, comme à la fin du chapitre V.

Le but de la partie B (chapitres VI à IX) est de fournir, là encore sans supposer de connaissance préalable, les bases de logique mathématique nécessaires pour le développement ultérieur d'une théorie des ensembles plus avancée. Il s'agit donc d'un cours de logique pour débutants, partant d'une introduction à la logique propositionnelle et aux logiques du premier ordre et cheminant pour donner un accès aussi rapide que possible aux résultats de logique indispensables pour la partie C, à savoir principalement le théorème de complétude des logiques du premier ordre et les théorèmes d'incomplétude de Gödel. Des démonstrations complètes sont données, mais l'aperçu présent ne peut prétendre constituer un cours complet de logique mathématique, en particulier parce que le traitement de la théorie des modèles et de la théorie de la démonstration y est indigent.

Dans la partie C (chapitres X à XVI), disposant désormais des résultats fondamentaux de la logique mathématique et d'un cadre précis pour la théorie descriptive, on revient à la théorie des ensembles pour en poursuivre l'élaboration dans un cadre conceptuel centré sur la notion de modèle de ZFC. Cette partie correspond à ce qui est souvent appelé la théorie axiomatique des ensembles, par opposition à la théorie élémentaire de la partie A. De la même façon que l'on peut étudier des modèles abstraits de l'arithmétique qui ne sont pas nécessairement le modèle standard constitué des vrais entiers et de leurs vraies opérations, on peut considérer des modèles abstraits de la théorie des ensembles qui ne sont pas nécessairement constitués des vrais ensembles et de la vraie appartenance. En permettant de considérer simultanément plusieurs modèles et de développer des opérations adaptées, cette approche, dite sémantique, a révolutionné la théorie, et elle en est le cœur depuis des décennies. On trouvera ici une présentation de notions devenues classiques comme les ensembles constructibles et la méthode du forcing, mais aussi et surtout un aperçu de développements plus récents autour de la hiérarchie des grands cardinaux et des propriétés de détermination, ainsi que quelques ouvertures vers des points de recherche en cours. Globalement, le texte est orienté vers la découverte de nouveaux axiomes susceptibles de compléter le système ZFC, lequel, à l'évidence, n'épuise pas l'intuition de la notion d'infini telle qu'elle se dégage des résultats accumulés depuis un siècle et demi. L'idée importante qui, on l'espère, devrait émerger, est que la théorie des ensembles ne se limite pas à des résultats négatifs de non-prouvabilité et d'indépendance, mais qu'elle est principalement une théorie structurelle visant à élaborer une conception positive de l'infini mathématique. On espère qu'au-delà des détails techniques, parfois délicats, le lecteur pourra s'imprégner de la philosophie générale de cette démarche et ressentir l'harmonie de la conception qui s'en dégage. Comme on l'a déjà dit, notre objectif ici n'est pas d'offrir un exposé exhaustif ni même toujours des démonstrations complètes, mais plutôt de mettre l'accent sur les notions et les idées fondamentales, de façon à permettre ensuite au lecteur d'accéder, s'il le souhaite, à des textes spécialisés.

Les chapitres I à VIII sont suivis de quelques exercices. Par ailleurs, on trouvera à la fin de chaque chapitre quelques repères chronologiques : il doit être clair que ceci ne prétend pas fournir une vision complète d'un sujet par ailleurs très riche, mais simplement de donner quelques dates pour fixer les idées. Sur les aspects historiques du développement de la théorie des ensembles, on pourra consulter [32, 44, 52] pour les origines et, spécialement, [60] pour la période plus récente. Pour les aspects philosophiques, on pourra consulter [92].

Le texte a été revu bien des fois, mais il n'est guère douteux que des coquilles, voire des erreurs, sont encore présentes. Les lecteurs sont invités à les signaler à `patrick.dehornoy@unicaen.fr`. Une liste des corrections sera tenue à jour à l'adresse

`"https://dehornoy.users.lmno.cnrs.fr/ens_erratum.html"`.

Les chapitres

La partie A regroupe cinq chapitres. Le chapitre I dessine un cadre axiomatique adapté à l'exploration des ensembles. Partant de l'existence de plusieurs infinis distincts, on est rapidement confronté à des problèmes non évidents, comme le problème du continu de Cantor, et ce sont eux qui rendent nécessaire l'élaboration d'une théorie des ensembles et la motivent. Comme il semble difficile de définir les ensembles de façon satisfaisante, on recourt à une approche axiomatique qui, au terme de plusieurs ajustements, aboutit au système de Zermelo, qui est la base axiomatique sur laquelle on propose de développer une théorie — donc d'explorer l'infini.

Le chapitre II contient une introduction aux ordinaux, qui sont un prolongement transfini de la suite des entiers naturels, et qui jouent le rôle d'une véritable épine dorsale du monde des ensembles. Après quelques résultats généraux sur les bons ordres, on présente la construction des ordinaux de von Neumann, qui fait de ceux-ci des ensembles transitifs particuliers. Cette approche permet d'établir rapidement les propriétés de base des ordinaux, en particulier de développer leur arithmétique. Comme applications, on établit le théorème de Cantor–Bendixson sur la décomposition des fermés de $\mathbb{R}$ et le théorème de Goodstein sur la convergence des suites éponymes.

Dans le chapitre III, utilisant en particulier les ordinaux finis comme contreparties des nombres entiers, on montre comment la plupart des objets mathématiques usuels peuvent être représentés par des ensembles dits purs, puis on vérifie que, pourvu qu'il soit amendé d'axiomes d'infini et de remplacement qui en font le système standard ZF de Zermelo-Fraenkel, le système axiomatique obtenu permet de légitimer tous les résultats de représentation par des ensembles purs, ainsi que toutes les propriétés usuelles des ordinaux, dont les définitions récursives. Il apparaît donc raisonnable à ce point de poursuivre le développement sur la base du système ZF.

Le chapitre IV est consacré à l'axiome du choix AC. On y établit l'équivalence de AC avec ses variantes classiques, théorème de Zermelo et lemme de Zorn, et on introduit deux formes faibles, l'axiome des choix dépendants et l'axiome du choix dénombrable. Ensuite, on passe en revue quelques applications de AC dans des domaines variés, allant de l'algèbre (existence de bases dans les espaces vectoriels, d'idéaux maximaux dans les anneaux) à l'analyse (existence d'ultrafiltres, théorème de Tikhonov) et à la géométrie (théorème de Hahn–Banach, décompositions paradoxales de Banach–Tarski). Enfin, anticipant sur les résulats d'indépendance démontrés dans la partie C, on discute brièvement du statut de AC pour conclure qu'il est cohérent avec le point de vue adopté d'inclure celui-ci dans le système de base, qui devient donc ZFC.

Dernier de la partie A, le chapitre V regroupe quelques résultats élémentaires sur les cardinalités (infinies) qui constituent le cœur de la théorie. On y définit la notion de cardinal, avec l'énumération des cardinaux infinis par les $\aleph_\alpha$ de Cantor et le développement d'une arithmétique cardinale où l'exponentiation pose des problèmes ardus. Comme outil structurant et permettant de mettre un peu de clarté, on introduit la cofinalité, avec les notions dérivées de cardinal successeur et de cardinal limite, et on conclut avec le théorème de Fodor, petit aperçu de combinatoire sur ω_1 et une de ses applications, le théorème de Silver sur la continuité de l'hypothèse du continu en cofinalité ω_1, bon exemple d'un résultat purement élémentaire et néanmoins de démonstration délicate.

Viennent ensuite les quatre chapitres de la partie B. Le chapitre VI est consacré au calcul propositionnel, ou booléen. Il n'est pas strictement indispensable à la suite, mais il offre l'avantage de présenter sur un exemple simple les notions de cohérence et de complétude pour un système de preuve. Ayant introduit les principales notions dans le cadre d'une logique formelle générale, on les illustre dans celui la logique booléenne, en établissant notamment le théorème de complétude qui montre que l'unique règle de coupure adossée à quelques schémas de preuve simples épuise les possibilités du raisonnement booléen.

Le chapitre VII est consacré aux logiques du premier ordre. Après un passage en revue de la syntaxe (inévitablement rébarbatif, comme il se doit ...), on passe rapidement à la notion de preuve, avec le théorème de complétude, ici soigneusement démontré et qui ouvre la voie à la notion de modèle d'une théorie, fondement du développement de la théorie des ensembles contemporaine, et à la méthode sémantique de preuve qui en dérive. En vue des développements du chapitre VIII, il est utile d'appliquer ce qui précède aux modèles de l'arithmétique de Peano faible.

Ensuite, le chapitre VIII est consacré aux célèbres résultats de limitation établis dans les années 1930, qui tracent un cadre indépassable pour le pouvoir déductif des systèmes axiomatiques. Le point-clé est un certain résultat technique, dit lemme diagonal. Une fois celui-ci établi, il est rapide de déduire à la fois le théorème d'indécidabilité de Church, le théorème de non-définissabilité de la vérité de Tarski, et les deux théorèmes d'incomplétude de Gödel, le second ici pour les systèmes au moins aussi forts que le système de Zermelo (ce qui est suffisant pour la théorie des ensembles). Mais, pour arriver au lemme diagonal, plusieurs résultats préliminaires sont nécessaires, et on commence par expliquer tout cela : bases sur les fonctions récursives, existence d'une arithmétisation récursive de la prouvabilité en logique du premier ordre, absoluité des formules d'arithmétique de complexité Σ_1 vis-à-vis des modèles de l'arithmétique de Robinson.

Le chapitre IX contient des bases de théorie descriptive des ensembles, qui est l'étude spécifique des sous-ensembles de la droite réelle $\mathbb{R}$, à la frontière de la topologie et de la théorie des ensembles. Cette étude, nécessaire en vue du chapitre XV, n'est pas directement partie de la logique, mais ses méthodes et son esprit sont directement réminiscents de ceux du chapitre VIII, et il est donc naturel qu'elle soit placée là. Outre les définitions des hiérarchies borélienne, projective, arithmétique, et analytique, on y établit l'universalité de l'espace de Baire parmi les espaces polonais, et la représentatibilité des ensembles $\mathbf{\Sigma}^1_1$ comme projections des branches d'un arbre sur $\omega \times \omega_1$.

Enfin viennent les sept chapitres de la partie C. Le chapitre X est introductif et rassemble des résultats élémentaires sur les modèles de ZF, avec en particulier les notions de modèle intérieur et d'absoluité d'une formule pour une famille de structures. Les résultats sont rudimentaires mais permettent de se familiariser avec la notion de modèle de ZF. Comme application facile, on établit l'indépendance de la plupart des axiomes du système ZF les uns par rapport aux autres.

Le chapitre XI décrit le modèles des ensembles constructibles qui, à l'intérieur de tout modèle de référence, est un modèle intérieur minimal, que l'on peut imaginer comme analogue au sous-corps premier d'un corps. On montre en particulier comment l'existence de ce modèle et l'étude de ses propriétés permettent d'établir la consistance relative de l'axiome du choix et de l'hypothèse généralisée du continu, ainsi que celle de la négation de l'hypothèse de Souslin.

Le chapitre XII est consacré à la méthode du forcing, qui permet, à partir d'un modèle convenable, d'en construire une extension, c'est-à-dire un nouveau modèle dont le modèle initial est modèle intérieur. On peut cette fois penser à une extension algébrique de corps. On décrit ici le mécanisme des extensions par forcing, et on développe quelques exemples classiques, permettant notamment d'établir la consistance de la négation de l'axiome du choix et de l'hypothèse du continu, ainsi que la consistance de l'hypothèse de Souslin via l'axiome de Martin.

Les chapitres suivants sont une introduction aux axiomes de grands cardinaux, qui occupent une place fondamentale dans la plupart des développements de la théorie des ensembles depuis les années 1970. Le chapitre XIII est centré sur les « petits » grands cardinaux, avec notamment ici les cardinaux inaccessibles et les cardinaux faiblement compacts, et les propriétés d'arbre et de partition associées. Ensuite, un développement approfondi est consacré aux cardinaux mesurables, à la charnière des « petits » et des « grands » grands cardinaux, qui sont fondamentaux à plusieurs titres. On étudie ici la notion-clé d'ultrapuissance d'un modèle de ZF associée à un cardinal mesurable, ainsi que la trace que laisse un tel cardinal sur le modèle intérieur $\boldsymbol{L}$ via une famille d'ordinaux indiscernables.

Le chapitre XIV présente quelques « grands » grands cardinaux, cardinaux forts, cardinaux de Woodin, cardinaux supercompacts, tous définis en termes de l'existence de plongements élémentaires convenables du modèle de référence dans un modèle intérieur. Une place spéciale est accordée aux cardinaux de Laver et à la structure des itérés des plongements élémentaires associés, qui mène à des applications étonnantes (et, pour le moment, peu connues) en combinatoire finie, avec les propriétés asymptotiques des périodes dans les tables de Laver.

Le chapitre XV est consacré à la propriété de détermination d'un sous-ensemble de la droite réelle. On y montre que l'hypothèse que tous les ensembles projectifs sont déterminés (« axiome de détermination projective DP ») fournit une description optimale de ces ensembles, et qu'elle résulte d'un certain axiome de grand cardinal. On explique alors comment un tel résultat, étendant au monde $H_{\aleph_1}$ des ensembles dénombrables la compréhension précédemment acquise pour le monde $H_{\aleph_0}$ des ensembles finis, légitime un nouveau consensus dans lequel ZFC amendé en $\mathsf{ZFC}+\mathsf{DP}$ devient la base axiomatique standard de la théorie des ensembles.

Enfin, dans le chapitre XVI, en guise de conclusion, on commence par revenir brièvement sur le rôle fondationnel de la théorie des ensembles pour souligner ce qui devrait apparaître à l'évidence comme des malentendus au lecteur parvenu à ce point, puis on tente une sorte d'évaluation de la théorie au vu des résultats acquis, pour revenir finalement aux mathématiques pour un aperçu de quelques travaux en cours, principalement autour des axiomes de forcing et des modèles canoniques, et des perspectives qu'ils ouvrent pour l'analyse de la structure $H_{\aleph_2}$ et pour une solution du problème du continu, toujours ouvert à ce jour, mais dont le lecteur devrait être désormais d'accord pour penser avec la plupart des spécialistes qu'il n'a nulle vocation à le rester à jamais.

Remerciements

Les dix premiers chapitres de ce texte sont issus de cours enseignés pendant des années à l'université de Caen et à l'École normale supérieure de Paris. Je remercie particulièrement les collègues qui m'ont épaulé à cette occasion et ont contribué à enrichir ces notes : Olivier Laurent, Julien Lévy, Claude Sureson, et Philippe Toffin. Je remercie également collectivement les étudiants qui ont suivi ces cours et les lecteurs connus ou inconnus des diverses versions qui m'ont depuis des années adressé commentaires, questions, corrections, suggestions, et en particulier Pierre Ageron, Maxime Bourrigan, Lorenzo Carlucci, Cédric Cessio, Ikram Cherigui, René Cori, Pierre Dehornoy, Rémi Goblot, Marc Hoyois, Gérard Lang, Serge Leblanc, Marc Mezzaroba, Simon Pépin-Lehalleur, Vincent Robin, Marc Sage, Pierre Simon, Pierre Simonnet, Lorenzo Tortora, Frédéric Wang.

Je remercie mes collègues de la communauté de théorie des ensembles, notamment Steve Jackson, Aki Kanamori, Stevo Todorcevic, Friedrich Wehrung, et Hugh Woodin, pour les nombreuses explications et précisions qu'ils m'ont apportées. Je remercie également Henri Lombardi qui, en jouant le rôle d'un opposant acharné à l'approche ensembliste, m'a aidé à dépasser les *a priori* dogmatiques.

Merci également à Alberto Arabia et Rached Mneimné qui, par leur expertise technique et leur exigence, m'ont aidé à préparer une mise en page soignée et, j'espère, agréable pour le lecteur, malgré sa densité. Et à Arlette Dehornoy et Isabella Bembo pour leur saisissante représentation des grands cardinaux.

Enfin et surtout, je remercie Serge Grigorieff qui, par une relecture tout à la fois minutieuse et compétente de l'intégralité de ce texte, a grandement contribué à en améliorer la qualité.

J'ajoute mes remerciements aux lecteurs qui, en débusquant fautes et imprécisions du premier tirage, ont permis d'améliorer le deuxième, que voici. Ma gratitude va spécialement à Alexandre Bailleul et à Martial Leroy, deux redoutables et souvent complémentaires chasseurs de coquilles.

Caen, Paris, Arnières-sur-Iton, juin 2017, et mai 2018

Table des matières

Partie B. Un peu de logique mathématique

VI. Logique propositionnelle

VII. Logique du premier ordre

VIII. Théorèmes de limitation