Éléments de mathématiques pour le XXI^{e} siècle, volume 2

Du même auteur

- Déjà paru :
  - Éléments de mathématiques pour le XXI^e siècle, volume 1 : Fondements des mathématiques 1 (logique des propositions et des prédicats, systèmes déductifs formels, arithmétique de Peano, structures algébriques de base)
- À paraître :
  - Éléments de mathématiques pour le XXI^e siècle, volume 3 : Fondements des mathématiques 3
  - Éléments de mathématiques pour le XXI^e siècle, volume 4 : Fondements des mathématiques 4

Éléments de mathématiques pour le XXI^{e} siècle, volume 2

Fondements des mathématiques 2
(théorie des ensembles, mathématiques discrètes, structures algébriques de base)

Étienne Bonheur

https://www.paysmaths.net

ISBN : 978-2-9569666-1-6
Dépôt légal : Juillet 2019

Table des matières

Introduction

Ce livre est le deuxième volume d'une série qui doit, à terme, couvrir l'ensemble des notions du premier cycle universitaire en mathématiques, tout en débordant largement sur le deuxième cycle. Il sera donc utile aux étudiants en licence ou en classes préparatoires scientifiques, ainsi qu'aux étudiants en master, y compris ceux préparant le CAPES ou l'agrégation (dont les programmes sont également très largement couverts par cette série d'ouvrages) [1]. Les enseignants y trouveront aussi de nombreux éléments leur permettant de préparer leurs cours, ou de compléter leurs connaissances dans des domaines qui ne leur sont pas familiers.

De manière plus générale, cette série d'ouvrages pourra être utile à toute personne s'intéressant aux mathématiques actuelles (les *mathématiques du XXI^e siècle* auxquelles fait référence le titre [2]). Elle devrait, *en théorie*, être accessible même sans connaissance préalable. En effet, les mathématiques sont prises à leur début et les différents concepts progressivement construits, chaque définition, théorème et démonstration ne faisant appel qu'à ce qui a été défini précédemment. Ce principe général aura cependant quelques exceptions : je pourrai, pour des raisons didactiques (notamment dans les remarques et exemples), ou par volonté de synthèse, être parfois amené à faire référence à des notions postérieures. À noter aussi que je suivrai un ordre me permettant d'enchaîner logiquement les différentes notions, mais qui n'est pas nécessairement l'ordre que l'on pourrait trouver dans un cursus universitaire, c'est-à-dire, par exemple, que certains éléments apparaissant dans les premiers volumes peuvent être enseignés traditionnellement dans des classes de troisième année de licence, voire au-delà. Néanmoins les chapitres peuvent être largement indépendants, et la compréhension d'un chapitre donné n'est pas toujours nécessaire à la compréhension de ceux qui suivent. Par ailleurs, lorsque cela peut être utile, les prérequis principaux seront indiqués au début d'une section [3].

Chaque ouvrage se veut à la fois

- didactique, avec des preuves très détaillées, des explications informelles, et de nombreux exemples et contre-exemples ;
- complet, voire encyclopédique, avec un exposé de nombreuses notions, des théorèmes tous démontrés, et de nombreux détails historiques (notamment sur l'origine des notations et du vocabulaire mathématique) ;
- synthétique, avec en particulier la volonté de multiplier les points de vue ; par exemple, les sujets pourront être abordés de façon à la fois formelle et informelle, et il pourra arriver que je donne plusieurs définitions équivalentes d'un même concept, ou plusieurs preuves d'un même théorème.

J'ai décidé de ne pas inclure de bibliographie, qui ne serait qu'une très longue liste de documents, et dont l'intérêt serait limité, sachant que dans cet ouvrage, tous les termes sont définis, tous les théorèmes sont prouvés, et le lecteur peut ainsi vérifier par lui-même tous les résultats. Les affirmations non justifiées (par exemple les remarques historiques) et certaines démonstrations sont directement sourcées dans les notes de bas de page. Cependant, pour les remarques portant sur l'origine du vocabulaire et des notations, je n'indiquerai pas à chaque fois mes sources principales, qui sont

1. Ou des cursus équivalents, pour les lecteurs francophones non français.

2. Le début du titre faisant par ailleurs référence aux *Éléments* d'Euclide, et aux *Éléments de Mathématique* de Bourbaki, deux œuvres partageant avec la présente série la volonté d'exposition des savoirs selon un ordre logique précis, à partir d'axiomes donnés.

3. Les différents prérequis indiqués ne correspondent ni à un minimum, ni à un maximum à connaître pour comprendre la section en cours, mais doivent être pris comme une aide pour identifier, parmi les notions abordées précédemment, celles pouvant être utiles.

- Jeff MILLER. *Earliest Known Uses of Some of the Words of Mathematics*. URL : `http://jeff560.tripod.com/mathword.html`.
- Jeff MILLER. *Earliest use of various mathematical symbols*. URL : `http://jeff560.tripod.com/mathsym.html`.
- Florian CAJORI. *A history of mathematical notations*. The Open Court Publishing Co., 1928-1929.

Je précise que les sources indiquées ne sont pas nécessairement exhaustives (je peux par exemple donner uniquement une source simple d'accès, ce qui est le cas des précédentes), et que dans la mesure du possible, je vérifie et recoupe toutes les informations, y compris les références données par ces différentes sources. Par ailleurs, les citations issues de textes non francophones feront automatiquement l'objet d'une traduction personnelle, sans que je le signale non plus à chaque fois.

Je précise aussi que les remarques historiques sont nécessairement succinctes, et ne rendent pas forcément compte de la complexité de l'évolution des concepts étudiés. Il en est de même pour la présentation de ces concepts, sous la forme de définitions, axiomes ou théorèmes. Celle-ci peut parfois donner l'impression que les notions s'articulent entre elles de façon naturelle, mais encore une fois cela cache les multiples variantes et points de vue, les contributions des différents mathématiciens, ainsi que leurs interrogations et tâtonnements, qui ont permis de façonner les mathématiques contemporaines.

On notera enfin qu'aucun paragraphe ne commence par « exercice », ce qui ne veut pas dire que le lecteur ne dispose d'aucun matériel pour s'exercer : les exemples ainsi que les nombreux théorèmes peuvent être considérés comme autant d'exercices corrigés (beaucoup d'énoncés que l'on trouve fréquemment dans la littérature mathématique sous l'intitulé *exercice* se trouvent ici sous l'intitulé *théorème*). Ainsi, chaque théorème étant suivi d'une preuve complète, il n'y aura pas dans cette série d'ouvrages d'expressions comme « la preuve est laissée en exercice », « le lecteur prouvera lui-même que … », et autres « on démontre facilement que … ».

Les quatre premiers volumes traitent des fondements modernes des mathématiques. Je prends cette expression dans un sens un peu général : au-delà de son acception la plus usuelle (comprenant, pour faire simple, la logique [4] mathématique et la théorie des ensembles), j'inclus d'autres sujets comme la construction des ensembles classiques de nombres (ensemble $\mathbb{N}$ des entiers naturels, ensemble $\mathbb{R}$ des nombres réels, …) ou l'étude de certaines structures algébriques de base (comme les groupes ou les anneaux).

Le premier volume est essentiellement consacré à la notion de logique mathématique et à la présentation, dans le cadre de la logique du premier ordre, de quelques structures algébriques, et de l'arithmétique de Peano qui formalise les propriétés des nombres entiers et des opérations associées (addition, multiplication).

Ce deuxième volume expose la théorie des ensembles de Zermelo-Fraenkel (ou ZF), qui est le fondement formel des mathématiques le plus classique :

- Le premier chapitre présente les axiomes principaux de la théorie, et quelques conséquences (en particulier, introduction des notions de *fonction* et de *relation*).
- Le deuxième chapitre reprend et généralise les structures algébriques de base vues dans le volume 1 (mais en les étudiant maintenant dans le cadre de la théorie des ensembles), et introduit le concept de *morphisme* entre structures et celui de *cardinal* (qui généralise le principe permettant de *dénombrer* un ensemble fini, c'est-à-dire de compter ses éléments).
- Le troisième chapitre présente l'axiome dit *de l'infini* et la construction de l'ensemble $\mathbb{N}$ des entiers naturels.
- Le quatrième chapitre est consacré à la construction de l'ensemble $\mathbb{Z}$ des entiers relatifs, et à la présentation d'applications diverses dans le domaine de ce qu'on appelle les *mathématiques discrètes* : éléments de théorie des nombres, c'est-à-dire l'étude des propriétés des nombres entiers (en reprenant et généralisant la présentation de l'arithmétique de Peano faite dans le volume 1), et introduction à l'analyse combinatoire (c'est-à-dire l'étude du dénombrement d'ensembles finis).

4. Logique : du grec *logikê*, dérivé de *logos*, signifiant à la fois raison, langage, et raisonnement.

- Le cinquième chapitre traite de l'axiome dit *du choix* et de ses conséquences.
- Le sixième chapitre donne les autres axiomes de la théorie (schéma d'axiomes de remplacement, axiome de fondation), dont les conséquences sont principalement utilisées dans la théorie elle-même, mais assez peu en dehors (dans d'autres domaines des mathématiques).

Les troisième et quatrième volumes seront consacrés

- à des compléments de théorie des ensembles : étude du concept d'*ordinal*, qui généralise le principe permettant d'*ordonner* un ensemble fini (en numérotant ses éléments), et compléments sur les cardinaux ;
- à des exemples de théories alternatives des ensembles : théorie des classes de von Neumann-Bernays-Gödel (NBG) et de Morse-Kelley (MK), théorie NFU (*New Foundations with Urelements*) ;
- à l'*introduction* de différentes théories mathématiques plus avancées en rapport avec la logique et les fondements des mathématiques : théorie des modèles, théorie de la calculabilité, théorie des catégories et des topos (et théorie élémentaire de la catégorie des ensembles, autre théorie alternative des ensembles), théorie homotopique des types (permettant aussi un autre fondement formel alternatif des mathématiques) ;
- à l'étude d'autres éléments de mathématiques discrètes (compléments de théorie des nombres et d'analyse combinatoire, introduction à la théorie des graphes) ;
- à des compléments sur les différentes structures mathématiques, et à la construction des autres ensembles classiques de nombres (ensemble $\mathbb{Q}$ des rationnels, ensemble $\mathbb{R}$ des réels, …).

Vocabulaire et notations

Je renvoie le lecteur au premier volume pour quelques remarques introductives concernant le vocabulaire et les notations. Je ne rappellerai ici que quelques notations qui peuvent être peu répandues, voire personnelles. Les notations classiques ne feront pas l'objet d'un rappel systématique, mais peuvent être trouvées dans la liste des symboles à la fin de cet ouvrage.

En ce qui concerne l'égalité, qui a en mathématiques un sens parfois subtil, je ferai la distinction entre *égalité*, *égalité par définition*, et *affectation* :

- égalité :
$$A = B$$
(« A est égal à B »)
signifie : les objets A et B sont identiques.
- égalité par définition :
$$A \stackrel{\text{def}}{=} B$$
(« A est égal, par définition, à B »)
signifie : on donne par définition, à l'objet B, le nom A.
- affectation :
$$A :\equiv B$$
(« A prend la valeur B »)
signifie : la variable A prend la valeur B. Il s'agit en quelque sorte d'une affectation, dans le sens informatique du terme.

Dans une formule, je peux noter simplement « et » le connecteur logique pour la conjonction, à la place de la notation usuelle ($\wedge$), et « ou » le connecteur logique pour la disjonction, à la place de la notation usuelle ($\vee$). De plus

- $\mathscr{F}(t/x)$ désigne la formule $\mathscr{F}$ dans laquelle la variable x a été remplacée par le terme t ;
- $\mathscr{F}[x_1, \dots, x_n]$ signifie que les variables libres de la formule $\mathscr{F}$ sont à prendre parmi $x_1, \dots, x_n$.

Enfin, le symbole $\equiv$ désigne l'équivalence (sémantique ou syntaxique) de deux formules, dans le sens suivant : si Γ est une théorie donnée (en général implicite)

$$\mathscr{F} \equiv_\Gamma \mathscr{G} \qquad \text{ou juste} \qquad \mathscr{F} \equiv \mathscr{G}$$

signifie

$$\Gamma \vdash \mathscr{F} \iff \mathscr{G}$$

ce qui équivaut aussi à

$$\begin{cases} \Gamma, \mathscr{F} \vdash \mathscr{G} \\ \Gamma, \mathscr{G} \vdash \mathscr{F} \end{cases}$$

le symbole $\vdash$ (conséquence syntaxique, c'est-à-dire inférence du système de déduction), pouvant être remplacé par le symbole $\models$ (conséquence sémantique). Cette formulation me permet en particulier de noter

$$\mathscr{F} \equiv \mathscr{G} \equiv \mathscr{H}$$

pour signifier

$$\begin{cases} \Gamma \vdash \mathscr{F} \iff \mathscr{G} \\ \Gamma \vdash \mathscr{G} \iff \mathscr{H} \end{cases}$$

Chapitre 1

Théorie des ensembles de Zermelo-Fraenkel (ZF), 1re partie : axiomes de base et conséquences

1.1 Introduction

La théorie des ensembles a été fondée par le mathématicien allemand Georg Cantor (1845-1918) dans un article publié en 1874. Sa théorie originale, que l'on qualifie aussi de *théorie naïve des ensembles* (car ce n'était pas une théorie formelle s'appuyant sur des axiomes), était sujette à différents paradoxes (voir le chapitre 1 du volume 1). Mais cette théorie s'étant avérée être un outil très utile pour les mathématiciens, le besoin d'avoir des bases solides et non contradictoires se faisait sentir. C'est le mathématicien allemand Ernst Zermelo (1871-1953) qui a été le premier à formaliser la théorie des ensembles en 1908 [1], de façon à éliminer les paradoxes connus. La théorie a ensuite été modifiée dans les années 1920 par le mathématicien allemand Abraham Fraenkel (1891-1965), le mathématicien et logicien norvégien Thoralf Skolem (1887-1963) et le mathématicien et physicien américano-hongrois John von Neumann (1903-1957), pour aboutir à la théorie des ensembles dite de Zermelo-Fraenkel (ou ZF), qui est le fondement formel des mathématiques le plus classique [2].

La théorie ZF utilise un langage logique du premier ordre égalitaire (voir le chapitre 4 du volume 1), dont la signature ne comprend que le prédicat binaire d'appartenance que l'on note $\in$. Si

$$x \in y$$

on dit que x appartient à y, ou que x est un élément de y, ou que y contient x. On note aussi $\notin$ la négation de $\in$:

$$x \notin y \stackrel{\text{def}}{\equiv} \neg(x \in y)$$

Ainsi, l'expression « $x \in y$ » formalise l'idée d'un élément x qui appartient à l'ensemble y, mais il faut se méfier d'une interprétation trop intuitive de cette notion. On notera par exemple que, comme dans toute théorie logique du premier ordre, il n'existe qu'un seul type d'objets (ici, des ensembles). Ainsi, dans cette théorie, il n'y a pas de différence de nature entre un ensemble et ses éléments : tous les éléments d'un ensemble sont eux-mêmes des ensembles ; c'est ce qu'on appelle parfois des *ensembles purs*, ou des *ensembles héréditaires*. Il serait possible d'envisager deux types de données : des ensembles, et des objets pouvant appartenir à un ensemble, mais qui n'en sont pas eux-mêmes, que l'on appelle en général *atomes*, ou *uréléments*. C'est

1. Ernst ZERMELO. « Untersuchungen über die Grundlagen der Mengenlehre *[Études sur les fondements de la théorie des ensembles]* ». Dans : *Mathematische Annalen* 65 (1908), p. 261-281.

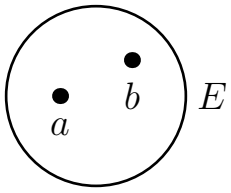
2. Il existe néanmoins différentes variantes, que je pourrai évoquer dans la suite.

le cas de certaines variantes de la théorie [3]. Mais cette approche nécessite alors l'astuce usuelle pour se ramener à une logique du premier ordre (à une seule sorte d'objets) : l'ajout d'un prédicat signifiant « être un ensemble » ou « être un urélément », ce qui complique l'expression de toutes les formules. Cependant, comme nous le verrons, tous les objets mathématiques classiques (les nombres entiers, les fonctions, ...) peuvent être *représentés* (on pourrait dire *codés*) par les *ensembles purs*. Il est donc « inutile » de considérer dans la théorie d'autres types d'objets.

Les différents axiomes permettront essentiellement de justifier l'existence d'ensembles vérifiant certaines formules, de telle sorte que l'on puisse les construire et les manipuler de façon intuitive et en accord avec l'usage (construction d'un ensemble à partir de la liste de ses éléments, réunion et intersection d'ensembles, ensemble des parties d'un ensemble, définition de fonctions entre deux ensembles ...), mais que les paradoxes évoqués précédemment (dans le volume 1) soient impossibles. Ainsi, un regroupement quelconque d'objets ne sera pas automatiquement un ensemble. Dans la suite, je réserverai le mot *ensemble* à un objet de la théorie, dont l'existence (éventuelle) est assurée de manière très précise par les axiomes. J'utiliserai le mot *collection* pour décrire un ensemble dans le sens intuitif et courant du terme (un regroupement de différents objets), et le mot *classe* pour parler d'une collection d'objets vérifiant une formule, mais qui ne forment pas obligatoirement un ensemble. Par exemple la collection des ensembles qui n'appartiennent pas à eux-mêmes, que l'on peut décrire comme la collection des *ensembles* x vérifiant la formule $x \notin x$, ou la collection des ensembles égaux à eux-mêmes, que l'on peut décrire comme la collection des *ensembles* x vérifiant la formule $x = x$, sont des *classes*, mais ces collections ne peuvent pas être des *ensembles* à cause de paradoxes classiques : la collection des x tels que $x \notin x$ conduit au paradoxe de Russell (un ensemble qui appartient à lui-même si et seulement si il n'appartient pas à lui-même), la collection des x tels que $x = x$ conduit au paradoxe de Cantor (l'ensemble de tous les ensembles, qui ne peut pas exister à cause du théorème de Cantor).

Remarque 1.1.1 (Origine des notations) : Le symbole $\in$ a été introduit par le mathématicien et linguiste italien Giuseppe Peano (1858-1932) [4]. Il l'écrit sous la forme de la lettre epsilon (ϵ), comme abréviation du latin *est*, qui signifie « il est, il existe ». L'idée étant par exemple que si P représente l'ensemble des nombres pairs, $x \in P$ signifie x *est* pair. Le symbole de non appartenance $\notin$ aurait été utilisé pour la première fois par le groupe Bourbaki [5] (Nicolas Bourbaki étant le pseudonyme collectif d'un groupe de mathématiciens francophones, formé en 1935).

Nous pouvons aussi illustrer l'appartenance à un ensemble à l'aide des diagrammes d'Euler et de Venn. Chaque cercle représente un ensemble, dont les éléments sont à l'intérieur (et ce qui n'appartient pas à l'ensemble est à l'extérieur du cercle). Par exemple, l'ensemble E contenant les éléments a et b peut être représenté par



En ce qui concerne les notations, puisque dans la théorie ZF tout objet est un ensemble, il n'est pas nécessaire de distinguer un ensemble de ses éléments par une lettre majuscule ou minuscule, comme il est courant de le faire en mathématiques. Néanmoins, je pourrai faire cette distinction dans certaines formules, pour faciliter la compréhension. Si on écrit par exemple $x \in E$, x est un ensemble, mais dans cette formulation on insiste sur la relation d'appartenance entre l'*élément* x et l'*ensemble* E. Je pourrai aussi parfois utiliser une

3. C'est notamment le cas de la théorie originale de Zermelo.

4. Giuseppe PEANO. *Arithmetices principia, nova methodo exposita* [Les principes de l'arithmétique, nouvelle méthode d'exposition]. 1889, p.vi, x.

5. Nicolas BOURBAKI. *Théorie des ensembles*. 1939, p. 4.

Ces pages ne sont pas incluses dans l'aperçu.

Le volume 2 des
Éléments de mathématiques pour le XXI^e siècle
(ISBN : 978-2-9569666-1-6) est disponible
en version papier et numérique.
Détails sur le site *Paysages Mathématiques* :
https://www.paysmaths.net/boutique

1.2 Axiomes de la théorie Z_{fini}

Le premier axiome affirme qu'un ensemble est complètement déterminé par ses éléments :

Axiome 1.2.1 (Axiome d'extensionnalité)
Deux ensembles contenant les mêmes éléments sont égaux : pour tout A et B

$$\forall x, (x \in A \iff x \in B) \implies A = B$$

autrement dit

$$(A \subseteq B \text{ et } B \subseteq A) \implies A = B$$

ce que l'on peut aussi écrire

$$(A \subseteq B \subseteq A) \implies A = B$$

Remarque 1.2.2 : Je rappelle que j'utilise une expression comme « pour tout A et B » pour signifier que l'on prend la clôture universelle des formules, qui consiste à ajouter devant chacune d'elles un quantificateur universel par variable libre. Autrement dit, l'énoncé complet, sous la forme d'une formule close, s'écrit

$$\forall A\ B, (\forall x, (x \in A \iff x \in B) \implies A = B)$$

Remarque 1.2.3 : La réciproque de l'implication, c'est-à-dire

$$A = B \implies \forall x, (x \in A \iff x \in B)$$

est aussi vraie, mais découle de la définition même de l'égalité : si $A = B$, on peut remplacer A par B dans une formule (et réciproquement), donc

$$x \in A \iff x \in B$$

Remarque 1.2.4 : Ainsi, pour justifier que deux ensembles A et B sont égaux, on peut :

- Soit procéder directement par équivalence : on montre que pour tout x

$$x \in A \iff x \in B$$

- Soit procéder par double inclusion : on montre

$$\begin{cases} \forall x, (x \in A \implies x \in B) \\ \forall x, (x \in B \implies x \in A) \end{cases}$$

c'est-à-dire

$$\begin{cases} A \subseteq B \\ B \subseteq A \end{cases}$$

Remarque 1.2.5 : Il est possible d'exprimer la théorie dans un langage non égalitaire (sans le symbole $=$), en définissant l'égalité entre deux ensembles comme le raccourci suivant

$$A = B \overset{\text{def}}{\equiv} \forall x, (x \in A \iff x \in B)$$

signifiant que deux ensembles sont égaux lorsqu'ils ont les mêmes éléments. Dans ce cas, l'axiome d'extensionnalité doit être remplacé par exemple par : pour tout x et y

$$x = y \implies \forall A, (x \in A \iff y \in A)$$

c'est-à-dire

$$\forall z, (z \in x \iff z \in y) \implies \forall A, (x \in A \iff y \in A)$$

afin que des ensembles égaux vérifient les mêmes formules. Sans cela, si deux ensembles x et y étaient égaux, on n'aurait pas de moyen de justifier que si $x \in A$, alors $y \in A$.

Le prochain axiome va justifier l'existence d'un ensemble particulier, l'ensemble vide. L'idée d'un ensemble qui soit vide peut sembler étrange, mais c'est un concept mathématique important. On peut aussi faire l'analogie suivante pour « justifier » l'existence d'un ensemble vide : un sac peut contenir différents objets, mais même vide, ça reste toujours un sac.

Axiome 1.2.6 (Axiome de l'ensemble vide)
Il existe un ensemble qui ne contient aucun élément :

$$\exists E \forall x, x \notin E$$

Définition 1.2.7 (Ensemble vide)
Il existe un unique ensemble E tel que

$$\forall x, x \notin E$$

On l'appelle l'*ensemble vide*, et on le note $\varnothing$.

Preuve
L'existence d'un tel ensemble est l'axiome de l'ensemble vide, et son unicité est une conséquence de l'axiome d'extensionnalité. En effet, soient A et B deux ensembles vérifiant l'axiome de l'ensemble vide, c'est-à-dire que pour tout x, on a $x \notin A$ et $x \notin B$. On en déduit

$$\begin{cases} \forall x, (x \in A \implies x \in B) \\ \forall x, (x \in B \implies x \in A) \end{cases}$$

puisqu'une formule de la forme $\mathscr{F} \implies \mathscr{G}$ est toujours vraie lorsque $\mathscr{F}$ est faux. Donc

$$\forall x, (x \in A \iff x \in B)$$

et par conséquent, d'après l'axiome d'extensionnalité, $A = B$.

Remarque 1.2.8 (Notations) : On trouve aussi la notation $\{\}$ pour désigner l'ensemble vide.

Remarque 1.2.9 (Origine des notations) : Le symbole $\varnothing$ apparaît en 1939 dans *Théorie des ensembles* de Bourbaki :

> « Certaines propriétés […] ne sont vraies pour *aucun* élément de E […] la partie qu'elles définissent est appelée la *partie vide* de E, et désignée par la notation $\varnothing$. » [9]

Le mathématicien français André Weil (1906-1998), l'un des membres fondateurs du groupe Bourbaki, raconte dans son autobiographie qu'il est à l'origine de l'adoption de ce symbole, issu de l'alphabet norvégien, et ressemblant à un zéro :

> « Bien plus tard la part que j'avais prise à ces débats me valut le respect de ma fille Nicolette quand je lui dis que j'étais personnellement responsable de l'adoption du symbole $\varnothing$ pour l'ensemble vide, symbole dont elle venait d'apprendre l'usage à l'école. Le $\varnothing$ appartenait à l'alphabet norvégien et j'étais seul dans Bourbaki à le connaître. » [10]

9. Nicolas BOURBAKI. *Théorie des ensembles (fascicule de résultats)*. 1939, p. 4.
10. André WEIL. *Souvenirs d'apprentissage*. Springer, 1991, p. 119.

Ces pages ne sont pas incluses dans l’aperçu.

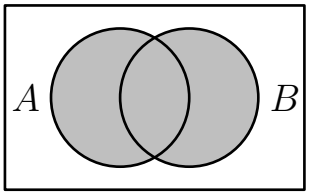


FIGURE 1.2 – Diagramme de Venn : $A \cup B$ (la zone correspondante est grisée).

Remarque 1.2.58 : La notation $\bigcup_{i=1}^{n} A_i$ peut être adaptée selon la numérotation des ensembles. Par exemple la réunion des ensembles $A_p, \dots, A_n$ $(p \leqslant n)$ est

$$\bigcup_{i=p}^{n} A_i \overset{\text{def}}{\equiv} A_p \cup \cdots \cup A_n$$

Par ailleurs cette expression pourra être redéfinie plus loin à l'aide du concept de *familles indexées* (section 1.7), et avec une variable n qui ne sera plus un entier du métalangage (comme c'est le cas pour l'instant) mais un objet de la théorie interprétant un tel entier. Après la construction de l'ensemble $\mathbb{N}$ des entiers naturels, la variable n représentera même directement un objet de la théorie (un élément de $\mathbb{N}$) sans aucune référence à la métathéorie (section 3.3).

Théorème 1.2.59 (Propriétés de la réunion)

1. Commutativité : pour tout A, B
$$A \cup B = B \cup A$$
2. Associativité : pour tout A, B, C
$$A \cup (B \cup C) = (A \cup B) \cup C$$
3. Idempotence : pour tout A
$$A \cup A = A$$
4. $\varnothing$ est élément neutre : pour tout A
$$A \cup \varnothing = \varnothing \cup A = A$$
5. Pour tout A et B, $A \cup B$ est la borne supérieure de $\{A, B\}$, c'est-à-dire le plus petit des ensembles (pour l'inclusion) qui sont supérieurs à la fois à A et à B :
$$A \subseteq A \cup B$$
$$B \subseteq A \cup B$$
$$\forall C, \left(\begin{cases} A \subseteq C \\ B \subseteq C \end{cases} \implies A \cup B \subseteq C \right)$$
la conjonction de ces trois formules étant aussi équivalente à l'unique formule
$$\forall C, \left(\begin{cases} A \subseteq C \\ B \subseteq C \end{cases} \iff A \cup B \subseteq C \right)$$
6. Caractérisation de l'inclusion en fonction de $\cup$: pour tout A et B
$$A \subseteq B \iff A \cup B = B$$

Preuve

1. La commutativité est une conséquence de celle de la disjonction (le connecteur « ou ») : pour tout x

$$x \in A \cup B \equiv x \in A \text{ ou } x \in B \equiv x \in B \text{ ou } x \in A \equiv x \in B \cup A$$

On en déduit d'après l'axiome d'extensionnalité (je ne le préciserai plus à chaque fois)

$$A \cup B = B \cup A$$

2. L'associativité est une conséquence de celle de la disjonction : pour tout x

$$\begin{aligned} x \in A \cup (B \cup C) &\equiv x \in A \text{ ou } x \in B \cup C \equiv x \in A \text{ ou } (x \in B \text{ ou } x \in C) \\ &\equiv (x \in A \text{ ou } x \in B) \text{ ou } x \in C \equiv x \in A \cup B \text{ ou } x \in C \\ &\equiv x \in (A \cup B) \cup C \end{aligned}$$

3. L'idempotence est une conséquence de celle de la disjonction : pour tout x

$$x \in A \cup A \equiv x \in A \text{ ou } x \in A \equiv x \in A$$

4. $\varnothing$ est élément neutre car

$$x \in A \cup \varnothing \equiv x \in A \text{ ou } x \in \varnothing \equiv x \in A$$

puisque par définition de l'ensemble vide

$$\forall x, x \notin \varnothing$$

5. Si $x \in A$, alors par définition $x \in A \cup B$ et de même si $x \in B$, alors $x \in A \cup B$, ce qui prouve

$$\begin{cases} A \subseteq A \cup B \\ B \subseteq A \cup B \end{cases}$$

On fait maintenant l'hypothèse $A \subseteq C$ et $B \subseteq C$.

- si $x \in A$, alors $x \in C$;
- si $x \in B$, alors $x \in C$.

Donc si $x \in A \cup B$, autrement dit si $x \in A$ ou $x \in B$, on en déduit $x \in C$ par disjonction des cas. Enfin, l'équivalence avec l'unique formule a déjà été démontrée dans la section 8.7 du volume 1 qui traite des treillis (nous aborderons aussi ce sujet dans la section 2.1). Je la redonne brièvement :

- Si $A \subseteq A \cup B$ et $B \subseteq A \cup B$, alors on déduit de $A \cup B \subseteq C$

$$\begin{cases} A \subseteq A \cup B \subseteq C \\ B \subseteq A \cup B \subseteq C \end{cases}$$

- De

$$A \cup B \subseteq C \implies \begin{cases} A \subseteq C \\ B \subseteq C \end{cases}$$

on déduit, puisque $A \cup B \subseteq A \cup B$, $A \subseteq A \cup B$ et $B \subseteq A \cup B$.

6. Caractérisation de l'inclusion en fonction de $\cup$:
   - On fait l'hypothèse $A \subseteq B$. Si $x \in A \cup B$, alors par définition $x \in B$ ou $x \in A$ (et alors $x \in B$). On en déduit par disjonction des cas $x \in B$. Par conséquent $A \cup B \subseteq B$, et comme l'inclusion réciproque est toujours vraie, $A \cup B = B$.
   - Réciproquement, on fait l'hypothèse $A \cup B = B$. Alors

$$A \subseteq A \cup B = B$$

On notera que l'on a étendu progressivement la signature de la théorie en ajoutant de nouveaux symboles, en plus du symbole d'appartenance $\in$. J'ai par exemple défini pour l'instant les symboles $\subseteq$, $\varnothing$, $\bigcup$, $\cup$. Cela revient à ajouter à chaque fois un nouveau symbole et un nouvel axiome correspondant à la définition du symbole en question. Par exemple

Ces pages ne sont pas incluses dans l’aperçu.

pour

$$\forall x_1 \forall x_2 \dots \forall x_n, ((x_1, x_2, \dots x_n) \in A_1 \times A_2 \times \cdots \times A_n \implies \mathscr{F})$$

et

$$\exists (x_1, x_2, \dots, x_n) \in A_1 \times A_2 \times \cdots \times A_n, \mathscr{F}$$

pour

$$\exists x_1 \exists x_2 \dots \exists x_n, ((x_1, x_2, \dots x_n) \in A_1 \times A_2 \times \cdots \times A_n \text{ et } \mathscr{F})$$

Remarque 1.3.23 : Si on a défini par récurrence $(a_1, \dots, a_{n+1})$ comme étant $(a_1, (a_2, \dots, a_{n+1}))$, alors on peut définir

$$A_1 \times \cdots \times A_{n+1} \overset{\text{def}}{=\!=} A_1 \times (A_2 \times \cdots \times A_{n+1})$$

Comme indiqué précédemment à propos des n-uplets, il peut arriver que l'on fasse la confusion, par exemple entre $(A \times B) \times C$ et $A \times (B \times C)$ que l'on note dans tous les cas $A \times B \times C$.

Remarque 1.3.24 : À cause de la façon dont ont été construits les n-uplets, deux produits cartésiens faisant intervenir un nombre différent d'ensembles peuvent ne pas être disjoints. Par exemple, il est possible que $A \times A$ et $A \times A \times A$ ne soient pas disjoints, si certains éléments de A sont des couples d'autres éléments : par exemple, si A contient les éléments a, b, c et (a, b), alors puisque $(a, b, c) \overset{\text{def}}{=\!=} ((a, b), c)$, (a, b, c) est à la fois un élément de $A \times A \times A$ et de $A \times A$.

Remarque 1.3.25 : Si l'un des A_i (i variant entre 1 et n) est vide alors le produit cartésien des A_i est vide, et réciproquement si tous les A_i sont non vides, il existe un élément dans le produit cartésien, donc

$$A_1 \times A_2 \times \cdots \times A_n = \varnothing \quad \text{si et seulement si} \quad \text{il existe un entier } i \text{ entre 1 et } n \text{ tel que } A_i = \varnothing$$

Remarque 1.3.26 : La notation $\prod_{i=1}^{n} A_i$ peut être adaptée selon la numérotation des ensembles. Par exemple le produit des ensembles $A_p, \dots, A_n$ $(p \leqslant n)$ est

$$\prod_{i=p}^{n} A_i \overset{\text{def}}{=\!=} A_p \times \cdots \times A_n$$

Par ailleurs cette expression pourra être redéfinie après la construction de l'ensemble $\mathbb{N}$ des entiers naturels, avec une variable n qui ne sera plus un entier du métalangage mais un objet de la théorie (un élément de $\mathbb{N}$) sans aucune référence à la métathéorie (section 3.5).

Théorème 1.3.27 (Propriétés du produit cartésien)

1. Produit cartésien et inclusion : pour tout A, B, A', B'

$$\begin{cases} A \subseteq A' \\ B \subseteq B' \end{cases} \implies A \times B \subseteq A' \times B'$$

et si A et B sont non vides

$$A \times B \subseteq A' \times B' \implies \begin{cases} A \subseteq A' \\ B \subseteq B' \end{cases}$$

Par conséquent, si A et B sont non vides, ou si A' et B' sont non vides

$$A \times B = A' \times B' \implies \begin{cases} A = A' \\ B = B' \end{cases}$$

2. Produit cartésien et intersection :
   - Le produit cartésien d'intersections est l'intersection de produits cartésiens : pour tout A, B, C, D

$$(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D) = (B \times C) \cap (A \times D)$$

   - Distributivité du produit cartésien sur l'intersection : pour tout A, B, C

$$A \times (B \cap C) = (A \times B) \cap (A \times C)$$
$$(B \cap C) \times A = (B \times A) \cap (C \times A)$$

3. Produit cartésien et réunion :
   - Réunion de produits cartésiens : pour tout A, B, C, D

$$(A \times C) \cup (B \times D) \subseteq (A \cup B) \times (C \cup D)$$
$$(A \times C) \cup (B \times D) = (A \setminus B) \times C \quad \cup \quad (A \cap B) \times (C \cup D) \quad \cup \quad (B \setminus A) \times D$$

   - Produit cartésien de réunions : pour tout A, B, C, D

$$(A \cup B) \times (C \cup D) = (A \times C) \cup (B \times D) \cup (A \times D) \cup (B \times C)$$

   - Distributivité du produit cartésien sur la réunion : pour tout A, B, C

$$A \times (B \cup C) = (A \times B) \cup (A \times C)$$
$$(B \cup C) \times A = (B \times A) \cup (C \times A)$$

4. Produit cartésien et différence :
   - Différence de produits cartésiens : pour tout A, B, C, D

$$(A \times C) \setminus (B \times D) = (A \times (C \setminus D)) \cup ((A \setminus B) \times C)$$

   - Produit cartésien de différences : pour tout A, B, C, D

$$(A \setminus B) \times (C \setminus D) = (A \times C) \setminus ((A \times D) \cup (B \times C))$$

   - Distributivité du produit cartésien sur la différence : pour tout A, B, C

$$A \times (B \setminus C) = (A \times B) \setminus (A \times C)$$
$$(B \setminus C) \times A = (B \times A) \setminus (C \times A)$$

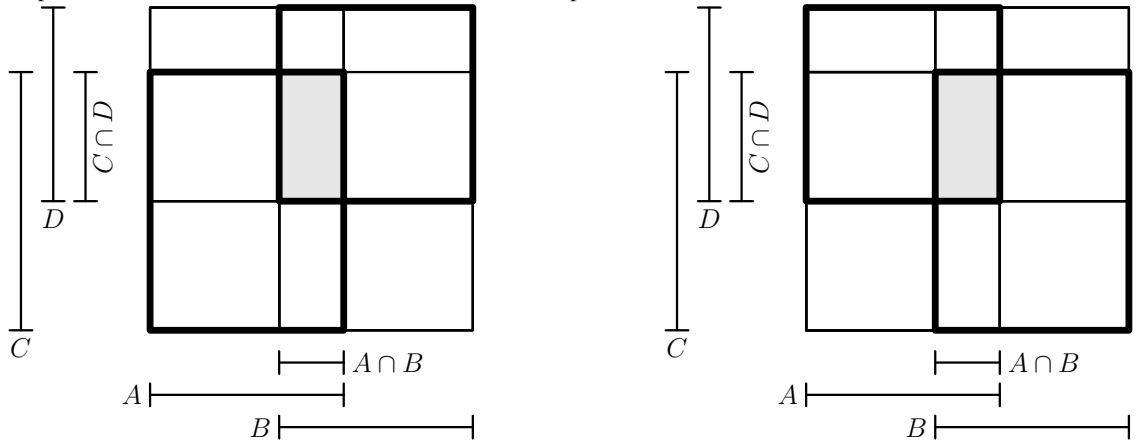
5. Produit cartésien et complémentaire : pour tout $A \subseteq A'$ et $B \subseteq B'$

$$\begin{aligned} \complement(A \times B) &= (\complement A \times B') \cup (A' \times \complement B) \\ &= (\complement A \times \complement B) \cup (\complement A \times B) \cup (A \times \complement B) \end{aligned}$$

avec $\complement(A \times B) \overset{\text{def}}{=} \complement_{A' \times B'}(A \times B)$ et $\complement A \overset{\text{def}}{=} \complement_{A'} A$ et $\complement B \overset{\text{def}}{=} \complement_{B'} B$

Preuve

1. Produit cartésien et inclusion :
   - La première implication est immédiate : si $A \subseteq A'$ et $B \subseteq B'$ alors $A \times B \subseteq A' \times B'$ puisque tout élément de $A \times B$ est de la forme (a, b), avec $a \in A$ (donc $a \in A'$), et $b \in B$ (donc $b \in B'$).
   - Réciproquement, on fait l'hypothèse $A \times B \subseteq A' \times B'$ (avec A et B non vides). Puisque $B \neq \emptyset$, il existe $b \in B$. On en déduit que pour tout $x \in A$
   $$(x, b) \in A \times B \subseteq A' \times B'$$
   donc $x \in A'$. Par conséquent $A \subseteq A'$. On prouve de manière symétrique (en permutant A avec B et A' avec B') que $B \subseteq B'$.
   - La dernière implication s'en déduit. Notons tout d'abord que si $A \times B = A' \times B'$, il suffit que A et B, ou que A' et B', soient non vides, pour que tous les ensembles A, B, A', B' le soient (puisque le produit cartésien de deux ensembles est vide si et seulement si l'un des ensembles l'est). Et alors on a
   $$\begin{cases} A \times B & \subseteq A' \times B' \\ A' \times B' & \subseteq A \times B \end{cases}$$
   donc d'après le résultat précédent
   $$A \subseteq A' \text{ et } B \subseteq B' \text{ et } A' \subseteq A \text{ et } B' \subseteq B$$
   et par conséquent $A = A'$ et $B = B'$.
2. Produit cartésien et intersection : pour ces deux propriétés, on utilise essentiellement l'associativité et la commutativité de la conjonction (le connecteur « et »).
   - Le produit cartésien d'intersections est une intersection de produits cartésiens :



$$\begin{aligned}(x, y) \in (A \cap B) \times (C \cap D) &\equiv (x \in A \cap B) \text{ et } (y \in C \cap D) \equiv x \in A \text{ et } x \in B \text{ et } y \in C \text{ et } y \in D \\ &\equiv (x \in A \text{ et } y \in C) \text{ et } (x \in B \text{ et } y \in D) \equiv (x, y) \in A \times C \text{ et } (x, y) \in B \times D \\ &\equiv (x, y) \in (A \times C) \cap (B \times D)\end{aligned}$$

   Par ailleurs, puisque l'intersection est commutative, on ne change pas le résultat en permutant A avec B ou C avec D, donc
   $$(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D) = (B \times C) \cap (A \times D)$$
   - Distributivité du produit cartésien : en appliquant ce qui précède en prenant $A = B$ on obtient (puisque $A \cap A = A$)
   $$A \times (C \cap D) = (A \times C) \cap (A \times D)$$
   et de même en prenant $C = D$ on obtient
   $$(A \cap B) \times C = (A \times C) \cap (B \times C)$$
   On pouvait aussi démontrer directement le résultat, par exemple pour la première égalité :
   $$\begin{aligned}(x, y) \in A \times (B \cap C) &\equiv x \in A \text{ et } y \in B \cap C \equiv x \in A \text{ et } y \in B \text{ et } y \in C \\ &\equiv (x \in A \text{ et } y \in B) \text{ et } (x \in A \text{ et } y \in C) \equiv (x, y) \in (A \times B) \text{ et } (x, y) \in (A \times C) \\ &\equiv (x, y) \in (A \times B) \cap (A \times C)\end{aligned}$$

Ces pages ne sont pas incluses dans l’aperçu.

Théorème 1.5.24 (Stabilité des propriétés pour une relation induite)
Toute relation induite par une relation réflexive (respectivement antiréflexive, symétrique, antisymétrique, asymétrique, transitive, totale) est réflexive (respectivement antiréflexive, symétrique, antisymétrique, asymétrique, transitive, totale).

Preuve
C'est immédiat, car si $\mathscr{C}$ est une sous-classe de $\mathscr{A}$, alors par définition tout élément de $\mathscr{C}$ est aussi un élément de $\mathscr{A}$. Par exemple, si $\mathscr{R}$ est une relation binaire réflexive sur $\mathscr{A}$, alors tout x dans la classe $\mathscr{C}$ est aussi un élément de $\mathscr{A}$ donc est tel que $x\mathscr{R}x$, c'est-à-dire que la relation induite sur $\mathscr{C}$ est aussi réflexive.

J'ai défini les relations sur des classes (pouvant être propres), et pas sur des ensembles uniquement (ce qui est plus classique), afin de me placer dans un cadre un peu large me permettant par exemple de parler de relation d'ordre ou de relation d'équivalence sur la classe de tous les ensembles (comme respectivement la relation d'inclusion, ou la relation d'équipotence), de relation d'ordre sur la classe des ordinaux (voir le volume 3), ou encore de fonctions entre classes propres.

Néanmoins, dans la suite, la plupart des énoncés en rapport avec les relations seront donnés pour des ensembles, afin d'avoir des formulations classiques (l'usage de classes propres n'est pas très courant en mathématiques, en dehors de certains domaines comme la théorie des ensembles). Cependant la grande majorité des définitions et théorèmes resteront valables pour des classes (essentiellement tout ce qui ne fait pas intervenir les deux grandes limitations, c'est-à-dire le fait qu'on ne puisse pas quantifier sur les classes, et le fait qu'une classe ne puisse pas appartenir à un ensemble ou à une classe). Je m'autoriserai, quand cela ne pose pas de problème, et sans plus de précision, à appliquer à des classes un énoncé écrit pour des ensembles. Les cas problématiques pourront faire l'objet de remarques spécifiques.

1.6 Relations fonctionnelles, fonctions

Définition 1.6.1 (Formule fonctionnelle)
On dit que la formule $\mathscr{F}[y,\vec{a}]$ est une *formule fonctionnelle* en y lorsque pour tout $\vec{a}$

$$\forall y\ z, \left(\begin{cases} \mathscr{F}(y) \\ \mathscr{F}(z) \end{cases} \implies y = z \right)$$

Définition 1.6.2 (Relation fonctionnelle, fonction partielle, fonction, image, antécédent)
On considère deux ensembles A et B.

1. On appelle *relation fonctionnelle* toute relation f telle que la formule

   $$(x,y) \in f$$

   soit une formule fonctionnelle en y, autrement dit telle que pour tout x, y, z

   $$\begin{cases} (x,y) \in f \\ (x,z) \in f \end{cases} \implies y = z$$

   Lorsque $(x,y) \in f$, ce que l'on peut aussi noter

   $$f : x \mapsto y$$

on dit que

- y est *l'image* de x par f, et on la note $f(x)$, ou f_x, ou fx.
- x est *un antécédent* de y par f.

Ainsi, pour tout x appartenant au domaine de f, les expressions suivantes ont toutes, par définition, le même sens :

$$\begin{aligned} &y \text{ est l'image de } x \text{ par } f \\ &f : x \mapsto y \\ &y = f(x) \\ &(x, y) \in f \end{aligned}$$

2. On appelle *fonction partielle* de A dans B toute relation fonctionnelle f dont le domaine est inclus dans A et l'image incluse dans B, c'est-à-dire toute relation binaire $f \subseteq A \times B$ telle que

$$\forall x \in A, \forall y\, z \in B, \left(\begin{cases} (x, y) \in f \\ (x, z) \in f \end{cases} \implies y = z \right)$$

3. On appelle *fonction* de A dans B toute relation fonctionnelle f de domaine A et dont l'image est incluse dans B, c'est-à-dire toute relation binaire $f \subseteq A \times B$ telle que

$$\forall x \in A, \exists! y \in B, (x, y) \in f$$

ce que l'on peut aussi écrire, avec les notations précédentes

$$\forall x \in A, \exists! y \in B, y = f(x)$$

4. Pour signifier que f est une fonction de A dans B, on note

$$f : A \longrightarrow B \quad \text{ou} \quad A \xrightarrow{f} B$$

et je pourrai écrire directement

$$\text{la fonction } A \xrightarrow{f} B$$

comme raccourci pour

$$\text{la fonction } f \text{ de } A \text{ dans } B$$

Remarque 1.6.3 : Comme indiqué en fin de section précédente, même si j'ai défini les relations fonctionnelles pour des ensembles, la même définition s'applique aux classes (pouvant être propres). Les différences entre les deux cas seront abordées dans diverses remarques.

Remarque 1.6.4 : L'image de x, si elle existe, est unique par définition. Par contre tout élément peut avoir plusieurs antécédents (ou aucun).

Remarque 1.6.5 (Origine des notations [19]) : La notation $f(x)$ est due au mathématicien suisse Leonhard Euler (1707-1783), dans *Commentarii Academiae Scientiarum Petropolitanae* (1734). On trouve la flèche

19. Source complémentaire : site *History of Science and Mathematics* du réseau *Stack Exchange* : `https://hsm.stackexchange.com/questions/5772/when-was-the-function-arrow-notation-x-mapsto-y-first-used`.

$\longrightarrow$ pour indiquer une correspondance entre un élément et son image dans *L'Algèbre Abstraite* (1936) du mathématicien norvégien Øystein Ore (1899-1968) :

> « Nous dirons que deux systèmes algébriques S et S' sont homomorphes (par rapport à l'addition et à la multiplication) s'il existe une correspondance a $\longrightarrow$ a' entre les éléments de S et S' donnant à chaque élément a de S une image unique a' dans S' telle que chaque élément de S' soit l'image d'au moins un élément de S et en outre telle que de a $\longrightarrow$ a', b $\longrightarrow$ b' on puisse conclure
>
> $$a + b \longrightarrow a' + b', ab \longrightarrow a'b' \text{ »}$$

Cette même notation est utilisée dès 1939 dans les premières éditions des *Éléments de mathématique* du groupe Bourbaki [20] (« l'application $x \longrightarrow f(x)$ »).

Le fait de représenter une fonction à l'aide d'une flèche entre son domaine et son codomaine serait apparu dans les années 1940. Le mathématicien américain Saunders Mac Lane (1909-2005) écrit dans *Categories for the Working Mathematician* :

> « L'idée fondamentale de représenter une fonction par une flèche est apparue pour la première fois en topologie vers 1940, probablement dans des notes ou conférences de W.Hurewicz sur les groupes d'homotopie relatifs [...] La flèche $f : X \longrightarrow Y$ a rapidement remplacé la notation occasionnelle $f(X) \subset Y$ pour une fonction. » [21]

Le symbole $\longrightarrow$ avait donc alors deux sens distincts, indiquant soit la correspondance entre un élément et son image, soit la fonction entre un domaine et son codomaine. L'apparition du symbole $\mapsto$ pour le premier sens aurait eu lieu vers 1963-64 dans le groupe Bourbaki. Il est possible de voir la transition entre les deux symboles en regardant le détail des éditions des *Éléments de mathématique* : en 1963, $\longrightarrow$ est utilisé (pour la correspondance entre un élément et son image) dans *Intégration, Chapitres 7 et 8* ; en 1965, dans la deuxième édition de *Intégration, Chapitres 1 à 4*, c'est $\mapsto$ qui est utilisé. Et on trouve à ce sujet, dans *Théorie des ensembles, Fascicule de résultats* (la quatrième édition de 1964) :

> « Lorsqu'une relation de la forme $y = \langle x \rangle$ (où $\langle x \rangle$ désigne une combinaison de signes dans laquelle peut figurer x) est une relation fonctionnelle en y, on désignera parfois la fonction qu'elle détermine par la notation $x \longrightarrow \langle x \rangle$, ou même simplement par $\langle x \rangle$, ce qui est un abus de langage très fréquent (par exemple, on parlera de la fonction $\sin x$ dans $\boldsymbol{R}$). On notera que le sens de la flèche $\longrightarrow$ est alors tout à fait différent de celui qui lui a été attribué ci-dessus [où une fonction est notée $f : E \longrightarrow F$] ; pour cette raison, on remplace souvent la notation $x \longrightarrow \langle x \rangle$ par la notation légèrement différente $x \mapsto \langle x \rangle$ pour éviter toute ambiguïté. Par exemple, si X, Y sont deux parties d'un ensemble E, la relation $Y = \complement X$ est fonctionnelle en Y, et il y a lieu de désigner par $X \mapsto \complement X$ l'application de $\mathscr{P}(E)$ dans lui-même qu'elle détermine, pour éviter de la confondre avec une application de X dans $\complement X$. »

Remarque 1.6.6 (Vocabulaire et notations) : Quand le domaine d'une fonction est de la forme $A_1 \times A_2 \cdots \times A_n$ on peut dire aussi que c'est une fonction d'*arité n*, ou une fonction à n arguments. Dans ce cas, pour éviter de multiplier les parenthèses, l'image du n-uplet $(x_1, \ldots, x_n)$ par f s'écrit en général

$$f(x_1, \ldots, x_n) \quad \text{plutôt que} \quad f((x_1, \ldots, x_n))$$

Remarque 1.6.7 (Vocabulaire) : Si une fonction (respectivement fonction partielle) a pour codomaine B, on peut aussi dire que c'est une fonction (respectivement fonction partielle) *à valeurs dans B*.

Remarque 1.6.8 (Vocabulaire) : On trouve très fréquemment le terme *application*, à la place de *fonction*. Mais l'usage veut aussi que l'on utilise *fonction* lorsque le codomaine est inclus dans l'ensemble $\mathbb{R}$ des

20. Nicolas Bourbaki est le pseudonyme collectif d'un groupe de mathématiciens francophones, formé en 1935.
21. Saunders Mac LANE. *Categories for the Working Mathematician*. Springer, 1971, p. 29.

nombres réels ou l'ensemble $\mathbb{C}$ des nombres complexes (on parle de la *fonction exponentielle*, de *fonctions trigonométriques*, de la *fonction gamma*, …), même pour les auteurs qui utilisent le terme *application*. L'usage le plus courant semble être d'employer *application* en algèbre (quand on travaille sur des ensembles quelconques) et *fonction* en analyse (quand on travaille sur des ensembles de nombres), mais comme dans la pratique, ces termes sont synonymes, je préfère n'en utiliser qu'un seul, *fonction* (qui a aussi l'avantage de correspondre au terme anglophone équivalent, *function*), et je n'emploierai donc pas *application*.

Par ailleurs, certains auteurs définissent le terme *fonction* comme ce que j'ai appelé une *fonction partielle*, c'est-à-dire qu'une fonction n'est alors définie que sur une partie de son ensemble de départ. Cette définition (qui semble trouver son origine dans la réforme de l'enseignement de la fin des années 1960, dite des *maths modernes* [22]), en contradiction avec l'emploi usuel du mot *fonction* en mathématiques, est à proscrire.

Remarque 1.6.9 (Notations) : Si Y est un terme ne faisant pas apparaître la variable y, la formule $y = Y$ est une formule fonctionnelle en y, et on note de la même manière

$$f : x \mapsto Y$$

pour signifier que pour tout x, l'unique y tel que $y = Y$ est l'image de x par f. À la place de

$$x \mapsto Y$$

on peut aussi trouver la notation suivante, issue du lambda-calcul [23] :

$$\lambda x.Y$$

Pour désigner la fonction f de A dans B telle que l'image de x soit Y, je pourrai noter

$$f : \begin{cases} A & \longrightarrow B \\ x & \longmapsto Y \end{cases}$$

Dans toutes ces expressions, la variable x est muette, et on peut la remplacer par n'importe quelle variable (n'apparaissant pas dans Y). Par ailleurs la flèche $\mapsto$ (ou le symbole λ) est un moyen rapide de désigner une fonction sans lui donner un nom. Par exemple :

- Si on se place dans l'ensemble $\mathbb{N}$, on peut désigner la fonction qui à un nombre entier x associe $x + 2$, comme la fonction $x \mapsto x+2$ (ou $y \mapsto y+2$, …), ou, de manière équivalente, comme la fonction $\lambda x.x+2$.
- Toujours dans $\mathbb{N}$, on peut désigner l'addition comme la fonction $(x, y) \mapsto x+y$ (ou, de manière équivalente, comme la fonction $\lambda(x, y).x + y$).
- Si on se place dans la classe de tous les ensembles, on peut désigner la fonction qui à un ensemble associe le singleton le contenant, comme la fonction $A \mapsto \{A\}$ (ou, de manière équivalente, comme la fonction $\lambda A.\{A\}$).

Remarque 1.6.10 : Le prédicat « être une relation fonctionnelle » peut s'exprimer par une formule du premier ordre :

$$f \text{ est une relation fonctionnelle} \overset{\text{def}}{\iff} \forall z, (z \in f \implies \exists x\ y, z = (x, y)) \quad \text{et} \quad \forall x\ y\ z, (((x, y) \in f \text{ et } (x, z) \in f) \implies y = z)$$

22. Mais qui n'est pas, contrairement à ce qu'on peut lire parfois, dans les *Éléments de Mathématique* de Bourbaki.
23. Voir le volume 3 pour une introduction au lambda-calcul.

Ces pages ne sont pas incluses dans l’aperçu.

donc $f = g$. Par conséquent si les fonctions $A \xrightarrow{f} B$ et $A \xrightarrow{g} B$ sont telles que

$$\forall x \in A, f(x) = g(x)$$

cela signifie aussi, puisque les domaines et codomaines sont égaux, que les *fonctions* $A \xrightarrow{f} B$ et $A \xrightarrow{g} B$ sont égales.

Remarque 1.6.17 : La réciproque est vraie par définition de l'égalité, donc on a aussi

$$(\forall x \in A, f(x) = g(x)) \iff f = g$$

Définition 1.6.18 (Image directe, image réciproque, fibre)
On considère une relation fonctionnelle f et une classe $\mathscr{A}$ (qui peut être un ensemble).

1. On appelle *image directe* (ou *image*) de $\mathscr{A}$ par f, que je noterai $\underset{\rightarrow}{f}(\mathscr{A})$, la classe de toutes les images des éléments de $\mathscr{A}$:
$$\underset{\rightarrow}{f}(\mathscr{A}) \overset{\text{def}}{\equiv} \{y \mid \exists x \in \mathscr{A}, (x, y) \in f\}$$
ce qui équivaut à
$$\underset{\rightarrow}{f}(\mathscr{A}) \overset{\text{def}}{\equiv} \{y \mid \exists x \in \mathscr{A} \cap \text{dom}(f), y = f(x)\}$$
et si $\mathscr{A}$ est inclus dans le domaine de f
$$\underset{\rightarrow}{f}(\mathscr{A}) \overset{\text{def}}{\equiv} \{y \mid \exists x \in \mathscr{A}, y = f(x)\}$$
ce que l'on peut aussi noter plus simplement
$$\{f(x) \mid x \in \mathscr{A}\} \qquad \text{ou} \qquad \{f_x\}_{x \in \mathscr{A}}$$
2. On appelle *image réciproque* de $\mathscr{A}$, que je noterai $\underset{\leftarrow}{f}(\mathscr{A})$, la classe des éléments dont l'image appartient à $\mathscr{A}$:
$$\underset{\leftarrow}{f}(\mathscr{A}) \overset{\text{def}}{\equiv} \{x \mid \exists y \in \mathscr{A}, (x, y) \in f\}$$
ce qui équivaut à
$$\underset{\leftarrow}{f}(\mathscr{A}) \overset{\text{def}}{\equiv} \{x \in \text{dom}(f) \mid f(x) \in \mathscr{A}\}$$
3. On appelle *fibre* de y par f l'image réciproque de $\{y\}$, autrement dit l'ensemble de tous les antécédents de y :
$$\underset{\leftarrow}{f}(\{y\}) = \{x \in \text{dom}(f) \mid f(x) = y\}$$

Remarque 1.6.19 : Ces définitions s'appliquent aux fonctions : l'image directe (respectivement image réciproque) de $\mathscr{A}$ par une fonction $\mathscr{E} \xrightarrow{f} \mathscr{F}$, est l'image directe (respectivement image réciproque) de $\mathscr{A}$ par la relation fonctionnelle f.

Remarque 1.6.20 : On notera que l'image directe de $\mathscr{A}$ est définie même si $\mathscr{A}$ n'est pas inclus dans le domaine de f, et l'image réciproque de $\mathscr{A}$ est définie même si $\mathscr{A}$ n'est pas inclus dans l'image de f.

Remarque 1.6.21 : Pour toute relation fonctionnelle f

$$\underset{\rightarrow}{f}(\varnothing) = \varnothing \qquad \text{et} \qquad \underset{\leftarrow}{f}(\varnothing) = \varnothing$$

Remarque 1.6.22 :

1. L'image de f est égale, par définition, à l'image directe de son domaine :
$$\mathrm{Im}(f) \overset{\text{def}}{\equiv} \{y \mid \exists x \in \mathrm{dom}(f), y = f(x)\} \overset{\text{def}}{\equiv} \underset{\rightarrow}{f}(\mathrm{dom}(f))$$
2. Si f est définie sur $A \times B$ (en particulier pour toute fonction $A \xrightarrow{f} B$), l'image réciproque du codomaine B de f est son domaine :
$$\underset{\leftarrow}{f}(B) = \mathrm{dom}(f)$$

Remarque 1.6.23 (Notations) : La notation la plus usuelle pour l'image directe de A est $f(A)$, et celle pour l'image réciproque de A est $f^{-1}(A)$. Mais

- Il y a des risques de confusion entre l'image d'un élément par f, et l'image directe. Par exemple, si on définit une relation fonctionnelle dont le domaine est l'ensemble
$$A :\equiv \{a, \{a\}\}$$
alors $f(\{a\})$ pourrait signifier l'image de l'élément $\{a\}$ par f, ou l'image directe de l'ensemble $\{a\}$, puisque $\{a\}$ est un sous-ensemble de A (a, seul élément de $\{a\}$, appartient à A).
- Il y a aussi un problème avec la notation $f^{-1}(A)$ qui pourrait aussi signifier, lorsque f est une fonction bijective (voir la section 5.3), l'image directe de A par la fonction réciproque f^{-1}, même si cette ambiguïté n'est pas trop grave car dans ce cas les deux ensembles possibles sont identiques.
- Si f est une fonction de l'ensemble A dans l'ensemble B, l'image directe induit une nouvelle fonction de $\mathscr{P}(A)$ dans $\mathscr{P}(B)$, et l'image réciproque induit une nouvelle fonction de $\mathscr{P}(B)$ dans $\mathscr{P}(A)$, et il est pratique de disposer d'une notation pour ces nouvelles fonctions (respectivement $\underset{\rightarrow}{f}$ et $\underset{\leftarrow}{f}$).

Certains auteurs utilisent des notations particulières pour lever les ambiguïtés, sans qu'il y ait de consensus sur la question. Les notations que je propose sont personnelles ; on peut aussi trouver par exemple, pour noter l'image directe de A, $f^{\rightarrow}(A)$, $f[A]$ ou $f_{\star}(A)$, et pour noter l'image réciproque de A, $f^{\leftarrow}(A)$, $f^{-1}[A]$ ou $f^{\star}(A)$.

Remarque 1.6.24 (Notations) : Je pourrai néanmoins noter, s'il n'y a pas d'ambiguïté
$$\underset{\leftarrow}{f}(y) \quad \text{plutôt que} \quad \underset{\leftarrow}{f}(\{y\})$$
la fibre de y.

Remarque 1.6.25 : Si E et F sont deux ensembles, une image directe et une image réciproque par une fonction $E \xrightarrow{f} F$ est toujours un ensemble d'après le schéma de compréhension, puisque
$$\underset{\rightarrow}{f}(\mathscr{A}) \overset{\text{def}}{\equiv} \{y \in F \mid \exists x \in \mathscr{A}, (x, y) \in f\}$$
$$\underset{\leftarrow}{f}(\mathscr{A}) \overset{\text{def}}{\equiv} \{x \in E \mid f(x) \in \mathscr{A}\}$$
donc $\underset{\rightarrow}{f}(\mathscr{A}) \in \mathscr{P}(F)$ et $\underset{\leftarrow}{f}(\mathscr{A}) \in \mathscr{P}(E)$. Ainsi, toute fonction f de l'ensemble E dans l'ensemble F induit deux nouvelles fonctions, une fonction $\underset{\rightarrow}{f}$ de $\mathscr{P}(E)$ dans $\mathscr{P}(F)$, qui à tout sous-ensemble de E associe son image directe :
$$\underset{\rightarrow}{f} : \begin{cases} \mathscr{P}(E) & \longrightarrow \mathscr{P}(F) \\ A & \longmapsto \underset{\rightarrow}{f}(A) \end{cases}$$
et une fonction $\underset{\leftarrow}{f}$ de $\mathscr{P}(F)$ dans $\mathscr{P}(E)$, qui à tout sous-ensemble de F associe son image réciproque :
$$\underset{\leftarrow}{f} : \begin{cases} \mathscr{P}(F) & \longrightarrow \mathscr{P}(E) \\ A & \longmapsto \underset{\leftarrow}{f}(A) \end{cases}$$

Ces pages ne sont pas incluses dans l’aperçu.

Remarque 1.6.37 : En particulier, si F est un système compatible de relations fonctionnelles de codomaine B, alors $\left(\bigcup_{f\in F} \mathrm{dom}(f)\right) \xrightarrow{\bigcup_{f\in F} f} B$ est une fonction qui prolonge chacune des relations fonctionnelles de F, ce qui signifie que pour tout $g \in F$

$$\left(\bigcup_{f\in F} f\right)|_{\mathrm{dom}(g)} = g$$

Exemple 1.6.38
L'ensemble des deux fonctions $A_1 \xrightarrow{f_1} B$ et $A_2 \xrightarrow{f_2} B$, avec $A_1 \cap A_2 = \emptyset$, est un système compatible (les domaines sont deux à deux disjoints), ce qui permet de définir la fonction

$$f_1 \cup f_2 : \begin{cases} A_1 \cup A_2 & \longrightarrow B \\ x & \longmapsto \begin{cases} f_1(x) & \text{si } x \in A_1 \\ f_2(x) & \text{si } x \in A_2 \end{cases} \end{cases}$$

Définition 1.6.39 (Ensemble de fonctions)
On considère deux *ensembles* A et B. Je noterai

$$A \longrightarrow B \qquad \text{ou} \qquad B^A$$

l'ensemble de toutes les fonctions de domaine A et de codomaine B

$$\begin{aligned} A \longrightarrow B &\overset{\text{def}}{\equiv} \{f \in \mathscr{P}(A \times B) \mid f \text{ est une relation fonctionnelle de domaine } A\} \\ &\overset{\text{def}}{\equiv} \{f \in \mathscr{P}(A \times B) \mid \forall x \in A, \exists! y \in B, (x, y) \in f\} \end{aligned}$$

Remarque 1.6.40 : Si on définit une fonction comme un triplet (A, B, f), l'ensemble des fonctions de A dans B devrait être l'ensemble des triplets, ce qui complique la manipulation de tels ensembles.

Remarque 1.6.41 : Si A et B sont des classes (pouvant être propres), une relation fonctionnelle $f \subseteq A \times B$ est une classe, et on ne peut pas définir l'ensemble (ni même la classe) de toutes ces relations fonctionnelles (car cela n'aurait pas de sens).

Remarque 1.6.42 (Notations) : Les notations les plus usuelles pour l'ensemble des fonctions de A dans B sont $\mathscr{F}(A, B)$ et B^A (on trouve aussi la variante ${}^A B$). La raison de cette notation *exponentielle* est qu'un ensemble de fonctions généralise, d'une certaine manière, le produit cartésien fini. En effet, se donner par exemple un élément de B^n, c'est-à-dire un n-uplet $(b_0, b_1, \ldots, b_{n-1})$ d'éléments de B, équivaut à se donner une fonction b d'un ensemble à n éléments dans B :

$$\begin{array}{ccccccc} 0 & 1 & 2 & \ldots & n-1 \\ \downarrow & \downarrow & \downarrow & \ldots & \downarrow \\ b_0 & b_1 & b_2 & \ldots & b_{n-1} \end{array} \qquad \text{équivaut à} \qquad b : \begin{cases} \{0, 1, 2, \ldots n-1\} & \longrightarrow B \\ k & \longmapsto b(k) \end{cases}$$

Ayant noté $\underline{n}$ l'ensemble $\{0, 1, 2, \ldots n-1\}$, il est équivalent de se donner un élément de B^n et un élément de l'ensemble des fonctions de $\underline{n}$ dans B.

J'utiliserai le plus souvent la notation $A \longrightarrow B$, qui a l'avantage de ne pas prêter à confusion avec l'exponentiation de deux ordinaux (voir le volume 3), ou de deux cardinaux (voir la section 4.11), d'être intuitivement compréhensible, et d'être cohérente avec la notation

$$f : A \longrightarrow B$$

qui signifie : f est une fonction de A dans B, et que l'on peut aussi comprendre comme

$$f \in A \longrightarrow B$$

c'est-à-dire : f appartient à l'ensemble des fonctions de A dans B.

Néanmoins, je pourrai parfois utiliser la notation B^A, en particulier dans le cadre des familles d'ensembles (voir la section 1.7), qui est une autre façon de manipuler les fonctions, ou pour certaines formules, cette notation pouvant alors être appropriée. Par exemple nous verrons qu'il existe une bijection de $A \longrightarrow (B \longrightarrow C)$ dans $(A \times B) \longrightarrow C$ ce qui peut s'écrire en notation exponentielle

$$(C^B)^A \simeq C^{A \times B}$$

Remarque 1.6.43 (Notations) : Je pourrai utiliser le raccourci

$$A \xrightarrow{f} B$$

pour

$$f \in A \longrightarrow B$$

par exemple avec des quantificateurs :

$$\forall\, A \xrightarrow{f} B \stackrel{\text{def}}{\equiv} \forall f \in A \longrightarrow B$$
$$\exists\, A \xrightarrow{f} B \stackrel{\text{def}}{\equiv} \exists f \in A \longrightarrow B$$

Remarque 1.6.44 :

1. Pour tout ensemble B, il existe une unique fonction de $\varnothing$ dans B. C'est l'ensemble vide $\varnothing$.
2. Par contre, si $A \neq \varnothing$, il n'existe aucune fonction de A dans $\varnothing$.
3. L'ensemble des fonctions de A dans le singleton $\{a\}$ est le singleton dont l'unique élément est la fonction constante égale à a.

On a donc, pour tous les ensembles A, B et a :

$$\varnothing \longrightarrow B = \{\varnothing\}$$
$$A \longrightarrow \varnothing = \begin{cases} \varnothing & \text{si } A \neq \varnothing \\ \{\varnothing\} & \text{si } A = \varnothing \end{cases}$$
$$A \longrightarrow \{a\} = \{f\} \text{ avec } f \text{ fonction constante} \quad f : \begin{cases} A & \longrightarrow \{a\} \\ x & \longmapsto a \end{cases}$$

Ces pages ne sont pas incluses dans l’aperçu.

5. La bijection réciproque de la fonction

$$f : \begin{cases} \mathbb{R}^+ & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto x^2 \end{cases}$$

est

$$f^{-1} : \begin{cases} \mathbb{R}^+ & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto \sqrt{x} \end{cases}$$

Théorème 1.9.20 (Inversibilité à droite et à gauche de fonctions)
On considère une fonction $A \xrightarrow{f} B$, avec $A \neq \emptyset$.

1. f est injective si et seulement si elle est inversible à gauche, autrement dit si et seulement si il existe une fonction $B \xrightarrow{r} A$ telle que
$$r \circ f = \mathrm{id}_A$$
Et alors r est une fonction surjective, qu'on appelle une *rétraction* de f.
2. Si f est inversible à droite, autrement dit s'il existe une fonction $B \xrightarrow{s} A$ telle que
$$f \circ s = \mathrm{id}_B$$
alors f est surjective, et s est une fonction injective, qu'on appelle une *section* de f.

Preuve
1. Injectivité :
   - On fait l'hypothèse qu'il existe une fonction $B \xrightarrow{r} A$ telle que $r \circ f = \mathrm{id}_A$. Si $f(x) = f(y)$, alors $r(f(x)) = r(f(y))$, donc $x = y$. On en déduit que f est injective.
   - Réciproquement, si f est injective et A non vide, on considère un élément quelconque $a \in A$, et la fonction r définie de B dans A de la façon suivante :
     — Si $y \in \mathrm{Im}(f)$, $r(y)$ est l'unique antécédent de y (y admet un antécédent puisque $y \in \mathrm{Im}(f)$ et cet antécédent est unique par injectivité de f).
     — Sinon, $r(y) :\equiv a$.

     Pour tout x dans A, $r \circ f(x)$ est l'unique antécédent de $f(x)$, c'est-à-dire x. Par conséquent $r \circ f = \mathrm{id}_A$.

   Le fait que r est surjective est immédiat, puisque pour tout élément x de A, $x = r(f(x))$, autrement dit $f(x)$ est un antécédent de x par r.
2. Surjectivité : Si $f \circ s = \mathrm{id}_B$ alors, comme démontré ci-dessus, f est surjective (tout élément $x \in B$ admet $s(x)$ comme antécédent par f), et s est injective d'après le point précédent (car s est inversible à gauche).

Remarque 1.9.21 : Si $A = \emptyset$, alors f est (trivialement) injective, mais il n'existe en général pas de fonction de B dans A (sauf si $B = \emptyset$), donc f n'admet pas de rétraction. Par contre les autres propriétés restent valables même si $A = \emptyset$: si f est inversible à gauche, alors f est injective, et si f est inversible à droite, alors f est surjective, mais de façon triviale (on a nécessairement aussi $B = \emptyset$, et les fonctions sont toutes la fonction vide).

Remarque 1.9.22 : Nous verrons dans la section 5.3, qu'avec l'ajout d'un autre axiome, l'axiome du choix, la réciproque de la deuxième propriété est vraie : si f est surjective, alors elle est inversible à droite (et par conséquent : f est surjective si et seulement si elle est inversible à droite). De manière informelle, si f est surjective, on peut prendre pour s une fonction qui, à tout élément de B, associe l'un de ses antécédents par

f. Mais l'axiome du choix est nécessaire pour justifier formellement de l'existence d'une telle fonction.

Remarque 1.9.23 : Dans la théorie des catégories (voir le volume 3), les flèches (généralisation des fonctions) qui sont inversibles à gauche s'appellent des split monomorphismes, et celles qui sont inversibles à droite des split épimorphismes. Donc d'après le théorème précédent, dans la catégorie des ensembles, les split monomorphismes sont des injections, les split épimorphismes sont des surjections, et réciproquement les injections de domaine non vide sont des split monomorphismes, et (avec l'axiome du choix) les surjections sont des split épimorphismes.

Théorème 1.9.24 (Corollaire 1)
Si deux fonctions $A \xrightarrow{f} B$ et $B \xrightarrow{g} A$ sont telles que

$$f \circ g = \mathrm{id}_B$$

alors f est surjective et g est injective.

Preuve
C'est l'application directe du théorème précédent (dans le cas où l'un des ensembles est vide, on a $A = B = f = g = \varnothing$ et les fonctions f et g sont trivialement injectives et surjectives).

Théorème 1.9.25 (Corollaire 2)
S'il existe une injection de A (non vide) dans B, alors il existe une surjection de B dans A.

Preuve
S'il existe une injection f de A (non vide) dans B, alors f admet une rétraction r qui est une surjection de B dans A.

Remarque 1.9.26 : Avec l'ajout de l'axiome du choix, la réciproque sera aussi vérifiée ; par conséquent il y aura alors équivalence entre l'existence d'une injection de A dans B et l'existence d'une surjection de B dans A.

Théorème 1.9.27 (Caractérisation des fonctions bijectives comme fonctions inversibles)
Une fonction $A \xrightarrow{f} B$ est bijective si et seulement si elle est inversible à droite et à gauche, autrement dit si et seulement si il existe deux fonctions $B \xrightarrow{g_1} A$ et $B \xrightarrow{g_2} A$ telles que

$$\begin{cases} g_1 \circ f = \mathrm{id}_A \\ f \circ g_2 = \mathrm{id}_B \end{cases}$$

et alors

$$g_1 = g_2 = f^{-1}$$

En particulier, f est bijective si et seulement si elle est inversible, autrement dit si et seulement si il existe une fonction $B \xrightarrow{g} A$ telle que

$$\begin{cases} g \circ f = \mathrm{id}_A \\ f \circ g = \mathrm{id}_B \end{cases}$$

(et alors $g = f^{-1}$).

Ces pages ne sont pas incluses dans l’aperçu.

On dit alors que A est *minoré* par m, ou que m *minore* A.

2. On appelle *majorant* de A dans E tout élément $M \in E$ tel que

$$\forall x \in A, M \geqslant x$$

On dit alors que A est *majoré* par M, ou que M *majore* A.

3. On dit que l'ensemble A est *borné* lorsqu'il est minoré et majoré, autrement dit lorsqu'il existe m et M dans E tels que

$$\forall x \in A, m \leqslant x \leqslant M$$

Remarque 1.11.22 : Un minorant de l'ensemble A n'appartient pas obligatoirement à A. Mais si c'est le cas, alors c'est le plus petit élément de A. Et réciproquement, si A admet un plus petit élément, alors c'est un minorant. Cette remarque reste bien entendu valable en remplaçant dans ce qui précède *minorant* par *majorant* et *plus petit élément* par *plus grand élément*.

Exemple 1.11.23

1. Dans $\mathbb{N}$, les éléments 0, 1, 2 sont des minorants de $\{3, 4, 5, 6, 7\}$. 3 est aussi un autre minorant de cet ensemble, et en est le plus petit élément. De même, 7 est le plus grand élément de cet ensemble, et tous les entiers supérieurs ou égaux à 7 en sont des majorants.
2. Pour tout ensemble $\mathscr{A}$, puisque

$$\forall A \in \mathscr{A}, A \subseteq \bigcup_{X \in \mathscr{A}} X$$

on en déduit que $\bigcup_{X \in \mathscr{A}} X$ est un majorant de $\mathscr{A}$ (pour la relation d'inclusion) dans la classe des ensembles.

Définition 1.11.24 (Borne inférieure, borne supérieure)
On considère un ensemble ordonné E et un sous-ensemble A de E.

1. S'il existe un plus grand minorant de A dans E (qui est donc unique), on l'appelle la *borne inférieure* de A, et on le note

$$\inf(A) \quad \text{ou} \quad \inf A \quad \text{ou} \quad \bigwedge A$$

Autrement dit m est la borne inférieure de A lorsque m est un minorant de A tel que tout minorant de A soit inférieur à m :

$$m = \inf(A) \overset{\text{def}}{\iff} \begin{cases} \forall x \in A, m \leqslant x \\ \forall y \in E, ((\forall x \in A, y \leqslant x) \implies y \leqslant m) \end{cases}$$

ce qui équivaut aussi à la définition suivante :

$$m = \inf(A) \overset{\text{def}}{\iff} \forall y \in E, ((\forall x \in A, y \leqslant x) \iff y \leqslant m)$$

Si de plus l'ordre est total, on en déduit par contraposition que m est la borne inférieure de A si et seulement si m est un minorant de A tel que tout élément strictement supérieur à m n'est pas un

minorant de A, autrement dit quand l'ordre est total

$$m = \inf(A) \iff \begin{cases} \forall x \in A, m \leqslant x \\ \forall y \in E, (y > m \implies (\exists x \in A, x < y)) \end{cases}$$

ce qui équivaut aussi à

$$m = \inf(A) \iff \forall y \in E, ((\exists x \in A, x < y) \iff m < y)$$

2. S'il existe un plus petit majorant de A dans E (qui est donc unique), on l'appelle la *borne supérieure* de A, et on le note

$$\sup(A) \quad \text{ou} \quad \sup A \quad \text{ou} \quad \bigvee A$$

Autrement dit M est la borne supérieure de A lorsque M est un majorant de A tel que tout majorant de A soit supérieur à M :

$$M = \sup(A) \stackrel{\text{def}}{\equiv} \begin{cases} \forall x \in A, M \geqslant x \\ \forall y \in E, ((\forall x \in A, y \geqslant x) \implies y \geqslant M) \end{cases}$$

ce qui équivaut aussi à la définition suivante :

$$M = \sup(A) \stackrel{\text{def}}{\equiv} \forall y \in E, ((\forall x \in A, y \geqslant x) \iff y \geqslant M)$$

Si de plus l'ordre est total, on en déduit par contraposition que M est la borne supérieure de A si et seulement si M est un majorant de A tel que tout élément strictement inférieur à M n'est pas un majorant de A, autrement dit quand l'ordre est total

$$M = \sup(A) \iff \begin{cases} \forall x \in A, M \geqslant x \\ \forall y \in E, (y < M \implies (\exists x \in A, x > y)) \end{cases}$$

ce qui équivaut aussi à

$$M = \sup(A) \iff \forall y \in E, ((\exists x \in A, x > y) \iff y < M)$$

Preuve (de l'équivalence des définitions)
Vérifions l'équivalence des deux définitions de la borne inférieure (le raisonnement est semblable pour la borne supérieure) :

- On fait l'hypothèse que m est le plus grand des minorants de A. On en déduit d'une part que tout $y \leqslant m$ est aussi un minorant (si $y \leqslant m$ alors pour tout $x \in A$, $y \leqslant m \leqslant x$), ce qui se traduit formellement par

$$\forall y \in E, (y \leqslant m \implies (\forall x \in A, y \leqslant x))$$

et d'autre part que tout minorant de A est inférieur à m, ce qui se traduit par l'implication réciproque, et par conséquent

$$\forall y \in E, ((\forall x \in A, y \leqslant x) \iff y \leqslant m)$$

- Réciproquement, si cette formule est vérifiée, on a d'une part l'implication

$$\forall y \in E, ((\forall x \in A, y \leqslant x) \implies y \leqslant m)$$

qui exprimer formellement l'idée que tout minorant de A est inférieur à m, et d'autre part, en appliquant l'implication réciproque avec $y :\equiv m$ (puisque $m \leqslant m$), on a

$$\forall x \in A, m \leqslant x$$

ce qui traduit formellement l'idée que m est un minorant de A. Par conséquent m est le plus grand des minorants de A.

Ces pages ne sont pas incluses dans l’aperçu.

Ce qui précède s'applique aussi à toute relation de bon ordre avec la particularité que, pour tout ensemble E non vide, il n'existe alors qu'un seul élément minimal (le plus petit élément de E).

Définition 1.13.24 (Segment initial)
On considère une classe ordonnée $(\mathscr{C}, \leqslant)$. On appelle *segment initial* toute classe $\mathscr{A} \subseteq \mathscr{C}$ telle que

$$\forall x\ y \in \mathscr{C}, \left(\begin{cases} x \leqslant y \\ y \in \mathscr{A} \end{cases} \implies x \in \mathscr{A} \right)$$

Si la classe $\mathscr{C}$ est totalement ordonnée, la définition équivaut aussi, par contraposition, à

$$\forall x \in \mathscr{C}, \forall y \in \mathscr{A}, x \notin \mathscr{A} \implies y < x$$

autrement dit

$$\forall x \in \mathscr{C} \setminus \mathscr{A}, \forall y \in \mathscr{A}, y < x$$

Un segment initial différent de $\mathscr{C}$ s'appelle un *segment initial strict* (ou *segment initial propre*).

Remarque 1.13.25 (Vocabulaire) : On trouve aussi l'expression *section commençante* pour désigner un segment initial.

Remarque 1.13.26 : La classe $\mathscr{C}$ elle-même est un segment initial.

Théorème 1.13.27 (Réunion de segments initiaux)
Si $\mathscr{A}$ est une classe de segments initiaux de $\mathscr{C}$, alors $\bigcup_{X \in \mathscr{A}} X$ est un segment initial de $\mathscr{C}$.

Preuve
Si x et y sont des éléments de $\mathscr{C}$ tels que $x \leqslant y$ et $y \in \bigcup_{X \in \mathscr{A}} X$, alors il existe un segment initial $A \in \mathscr{A}$ tel que $y \in A$, donc $x \in A$ et par conséquent $x \in \bigcup_{X \in \mathscr{A}} X$.

Remarque 1.13.28 : Les éléments de la classe $\mathscr{A}$ sont nécessairement des ensembles (par définition), et pas des classes propres.

Théorème 1.13.29 (Propriétés élémentaires d'un segment initial)
On considère un segment initial $\mathscr{A}$ d'une classe $\mathscr{C}$.

1. Pour tout $a \in \mathscr{A}$
$$]-\infty, a[_{\mathscr{A}} =]-\infty, a[_{\mathscr{C}}$$
2. Si la classe $\mathscr{C}$ est totalement ordonnée, et si $\mathscr{A}$ admet un minimum a, alors a est aussi le minimum de $\mathscr{C}$.
3. Si $\mathscr{B}$ est un segment initial de $\mathscr{A}$ (pour l'ordre induit par l'ordre sur $\mathscr{C}$), alors $\mathscr{B}$ est un segment initial de $\mathscr{C}$.

Preuve

1. Puisque $\mathscr{A}$ est une sous-classe de $\mathscr{C}$, on a
$$]-\infty, a[_{\mathscr{A}} \subseteq]-\infty, a[_{\mathscr{C}}$$
et pour tout élément $x \in \mathscr{C}$ tel que $x < a$, on a $x \in \mathscr{A}$ (puisque $\mathscr{A}$ est un segment initial de $\mathscr{C}$) et $x < a$, donc $x \in]-\infty, a[_{\mathscr{A}}$. D'où l'inclusion réciproque, et par conséquent
$$]-\infty, a[_{\mathscr{C}} =]-\infty, a[_{\mathscr{A}}$$
2. Si l'ordre est total, tout élément de $\mathscr{A}$ est strictement inférieur à tout élément de $\mathscr{C} \setminus \mathscr{A}$, donc en particulier pour tout $x \in \mathscr{C} \setminus \mathscr{A}$, $a < x$, et par conséquent a est le minimum de $\mathscr{C}$.
3. On considère un segment initial $\mathscr{B}$ de $\mathscr{A}$ et deux éléments $x \in \mathscr{C}$ et $y \in \mathscr{B}$ tels que $x \leqslant y$. Puisque $\mathscr{B} \subseteq \mathscr{A}$, on a $y \in \mathscr{A}$. Or $\mathscr{A}$ est un segment initial de $\mathscr{C}$, donc $x \in \mathscr{A}$, et puisque $\mathscr{B}$ est un segment initial de $\mathscr{A}$, on en déduit $x \in \mathscr{B}$.

La classe des éléments de $\mathscr{C}$ inférieurs (strictement ou pas) à un élément donné a, c'est-à-dire les intervalles

$$]-\infty, a] \stackrel{\text{def}}{=} \{x \in \mathscr{C} \mid x \leqslant a\}$$
$$]-\infty, a[\stackrel{\text{def}}{=} \{x \in \mathscr{C} \mid x < a\}$$

sont des segments initiaux. Le deuxième exemple est particulièrement important, car c'est toujours un segment initial strict, et dans le cas d'un bon ordre *défini sur un ensemble*, tout segment initial strict est de cette forme. D'où les deux théorèmes suivants :

Théorème 1.13.30
On considère une classe ordonnée $\mathscr{C}$. Pour tout $a \in \mathscr{C}$

- La classe $]-\infty, a]$ est un segment initial.
- La classe $]-\infty, a[$ est un segment initial strict.

Preuve
Soit $a \in \mathscr{C}$. Par définition, $]-\infty, a]$ et $]-\infty, a[$ sont bien inclus dans $\mathscr{C}$. On considère x et y dans $\mathscr{C}$ tels que $x \leqslant y$.

- Si $y \in]-\infty, a]$, alors $x \leqslant y \leqslant a$, et par transitivité $x \leqslant a$.
- Si $y \in]-\infty, a[$, alors $x \leqslant y < a$, donc $x < a$.

On en déduit que $]-\infty, a]$ et $]-\infty, a[$ sont des segments initiaux. Enfin, $]-\infty, a[\neq \mathscr{C}$ puisque $a \notin]-\infty, a[$.

Théorème 1.13.31 (Caractérisation des segments initiaux stricts dans un ensemble bien ordonné)
On considère un *ensemble E bien ordonné*, et un sous-ensemble A de E. A est un segment initial strict si et seulement si il existe $a \in E$ tel que $A =]-\infty, a[$.

Preuve
Nous venons de voir que $]-\infty, a[$ est un segment initial strict. Démontrons la réciproque. Soit un segment initial A distinct de E. Alors $\complement_E A$, complémentaire de A dans E, est un sous-ensemble non vide de E (puisque $A \neq E$), donc contient un plus petit élément a (car E est bien ordonné). Justifions que $A =]-\infty, a[$:

- Si $x < a$, alors $x \in A$ puisque a est le plus petit élément de $\complement_E A$ (aucun élément de $\complement_E A$ n'est strictement inférieur à a).
- Réciproquement, on fait l'hypothèse $x \in A$. Si $a \leqslant x$, alors $a \in A$ (car A est un segment initial). Or $a \notin A$, donc $a \not\leqslant x$, c'est-à-dire $x < a$ (l'ordre est total puisque c'est un bon ordre).

La preuve du théorème précédent n'est plus valable dans une classe bien ordonnée, puisque le complémentaire d'une classe est une classe, et on ne peut pas conclure qu'elle admet un plus petit élément. Mais on

Ces pages ne sont pas incluses dans l’aperçu.

toutes les parties A et B de E

$$A \subseteq B \equiv A \cap B = A \equiv A \cup B = B$$

Il serait aussi possible de justifier d'abord que $(\mathscr{P}(E), \cap, \cup, \varnothing, E, \complement)$ est une algèbre de Boole, car nous savons que les six propriétés caractérisant cette structure (commutativité, associativité, lois d'absorption, élément neutre, distributivité, complément) sont vérifiées, puis d'en déduire la structure de treillis (voir la remarque 2.6.11, p. 233). Enfin, les propriétés liées à la borne supérieure et la borne inférieure d'une partie quelconque de $\mathscr{P}(E)$ sont aussi déjà connues, et la compatibilité de la réunion et l'intersection avec l'inclusion est une conséquence immédiate de la définition des bornes supérieure et inférieure d'une paire, que nous avons prouvée dans le volume 1, mais que l'on peut aussi redémontrer. Pour la première formule par exemple : si $A \subseteq B$ alors $A \cap C \subseteq A \subseteq B$ et $A \cap C \subseteq C$, donc $A \cap C \subseteq B \cap C$.

Remarque 2.1.6 : Nous reviendrons dans la section 2.6 sur $\mathscr{P}(E)$, qui est non seulement un treillis, mais aussi une algèbre de Boole.

Théorème 2.1.7 (Existence d'un plus petit et d'un plus grand élément dans un treillis complet)
Tout treillis complet est borné (autrement dit tout treillis complet admet un plus petit élément et un plus grand élément).

Preuve
Puisque dans un treillis complet E toute partie admet une borne supérieure et une borne inférieure, c'est en particulier le cas de l'ensemble vide (qui est inclus dans tous les ensembles). Donc

- $\varnothing$ admet une borne supérieure, ce qui signifie que E admet un plus petit élément (tout élément de E est trivialement un majorant de $\varnothing$, et la borne supérieure de $\varnothing$ est donc le plus petit élément de E).
- De même, $\varnothing$ admet une borne inférieure, ce qui signifie que E admet un plus grand élément (tout élément de E est trivialement un minorant de $\varnothing$, et la borne inférieure de $\varnothing$ est donc le plus grand élément de E).

Ainsi, tout treillis complet admet un plus petit élément et un plus grand élément.

Remarque 2.1.8 : Si $(E, \leqslant)$ est un ensemble totalement ordonné, c'est un treillis, mais pas nécessairement un treillis complet. Par exemple $(\mathbb{N}, \leqslant)$ et $(\mathbb{R}, \leqslant)$ sont des ensembles totalement ordonnés qui ne sont pas des treillis complets (ils n'ont pas de plus grand élément). Par contre si $a < b$, $[a, b]_{\mathbb{N}}$ et $[a, b]_{\mathbb{R}}$ sont des treillis complets (dans le cas de $\mathbb{N}$, la borne inférieure et la borne supérieure d'un sous-ensemble non vide A de $[a, b]_{\mathbb{N}}$ sont même respectivement le plus petit et le plus grand élément de A).

Théorème 2.1.9 (Intervalle $[a, b]$ dans un treillis complet)
Pour tout treillis complet $(E, \leqslant)$ et tous les éléments $a \in E$ et $b \in E$ tels que $a \leqslant b$, l'intervalle $[a, b]_E$, muni de la relation d'ordre induite, est un treillis complet.

Preuve
Notons d'abord que si $a \leqslant b$, l'intervalle $[a, b]$ n'est pas vide (il contient au moins a et b). On considère un sous-ensemble A de $[a, b]$. Si $A = \varnothing$, alors a est trivialement la borne supérieure de A. On suppose dans la suite que $A \neq \varnothing$.

- Puisque $A \subseteq [a, b]$, a est un minorant de A et b est un majorant de A. De plus, comme A est aussi un sous-ensemble de E, treillis complet, A admet (dans E) une borne inférieure m et une borne supérieure M, qui sont telles que $a \leqslant m$ et $M \leqslant b$.
- Pour prouver que m et M sont respectivement la borne inférieure et la borne supérieure de A dans $[a, b]$, il reste à justifier que ce sont des éléments de cet intervalle : m étant un minorant de A et M un majorant (et A étant non vide), on a $m \leqslant M$, donc

  $$a \leqslant m \leqslant M \leqslant b$$

  ce qui signifie que m et M sont des éléments de $[a, b]$, et par conséquent ce sont respectivement la borne inférieure et la borne supérieure de A dans $[a, b]$.

Théorème 2.1.10 (Lemme de Knaster-Tarski)
Si $(E, \leqslant)$ est un treillis complet, alors toute fonction croissante $E \xrightarrow{f} E$ admet un point fixe. Plus précisément, f admet un plus petit point fixe, qui est

$$\inf\{x \in E \mid f(x) \leqslant x\}$$

et un plus grand point fixe, qui est

$$\sup\{x \in E \mid x \leqslant f(x)\}$$

Preuve
On considère l'ensemble

$$A :\equiv \{x \in E \mid f(x) \leqslant x\}$$

Notons d'abord que tout point fixe de f est un élément de A. Puisque E est un treillis complet, A admet une borne inférieure, que l'on note m. Vérifions que m est le plus petit des points fixes de f :

- Prouvons d'abord que $m \in A$ (ce qui montrera que m est le plus petit élément de A, et donc que m est inférieur à tous les points fixes de f). Pour tout $x \in A$, on a $m \leqslant x$, donc $f(m) \leqslant f(x)$, puisque f est croissante, et $f(x) \leqslant x$, puisque x est un élément de A, et par conséquent $f(m) \leqslant x$. On en déduit que $f(m)$ est un minorant de A, donc par définition de la borne inférieure, $f(m) \leqslant m$. Cela signifie par définition de A que $m \in A$, donc m est le plus petit élément de A.
- Prouvons que m est un point fixe de f. Puisque f est croissante et $f(m) \leqslant m$, on en déduit

  $$f(f(m)) \leqslant f(m)$$

  ce qui signifie que $f(m) \in A$. Or m est le plus petit élément de A, donc $m \leqslant f(m)$, et par conséquent $f(m) = m$, autrement dit m est un point fixe de f.
- Puisque tout point fixe de f appartient à A, que m est un point fixe de f et que c'est le plus petit élément de A, m est donc le plus petit des points fixes de f.

On démontre de la même façon que f admet un plus grand point fixe, qui est la borne supérieure de

$$\{x \in E \mid x \leqslant f(x)\}$$

Remarque 2.1.11 : Le preuve du théorème montre que le plus petit point fixe (respectivement le plus grand point fixe), qui est la borne inférieure de l'ensemble $\{x \in E \mid f(x) \leqslant x\}$ (respectivement la borne supérieure de l'ensemble $\{x \in E \mid x \leqslant f(x)\}$), appartient à cet ensemble, et en est donc le plus petit élément (respectivement le plus grand élément).

Remarque 2.1.12 (Vocabulaire) : La démonstration du théorème précédent est la première partie de la preuve du théorème dit de Knaster-Tarski (voir ci-dessous). D'où la dénomination de *lemme de Knaster-Tarski*, qui n'est pas usuelle, mais qui est appropriée.

Théorème 2.1.13 (Théorème de Knaster-Tarski, ou théorème de point fixe de Tarski)
On considère un treillis complet $(E, \leqslant)$, et une fonction croissante $E \xrightarrow{f} E$. Alors l'ensemble des points fixes de f, muni de la relation induite par $\leqslant$, est un treillis complet.

Preuve
On note F l'ensemble des points fixes de f, M le plus grand élément de E, et on considère un sous-ensemble A de F dont on veut démontrer qu'il admet une borne supérieure dans F. Notons S la borne supérieure de A dans E. L'intervalle $[S, M]$ est donc l'ensemble des majorants de A dans E, et l'ensemble des points fixes de f appartenant à cet intervalle est l'ensemble des majorants de A dans F. Nous allons utiliser le lemme pour démontrer que ce dernier ensemble admet un

Ces pages ne sont pas incluses dans l’aperçu.

Théorème 2.3.26 (Théorème de Cantor-Bernstein)
On considère deux ensembles non vides A et B. S'il existe une injection de A dans B, et une injection de B dans A, alors A et B sont en bijection.

Preuve 1 (faisant appel au lemme de Knaster-Tarski)
On considère deux injections $A \xrightarrow{f} B$ et $B \xrightarrow{g} A$, et la fonction

$$h : \begin{cases} \mathscr{P}(A) & \longrightarrow \mathscr{P}(A) \\ X & \longmapsto \complement_A \underset{\rightarrow}{g}(\complement_B \underset{\rightarrow}{f}(X)) \end{cases}$$

Démontrons que h est une fonction croissante (pour l'inclusion) : soient X et Y des éléments de $\mathscr{P}(A)$, tels que $X \subseteq Y$. On a donc

$$\begin{gathered} \underset{\rightarrow}{f}(X) \subseteq \underset{\rightarrow}{f}(Y) \\ \complement_B \underset{\rightarrow}{f}(Y) \subseteq \complement_B \underset{\rightarrow}{f}(X) \\ \underset{\rightarrow}{g}(\complement_B \underset{\rightarrow}{f}(Y)) \subseteq \underset{\rightarrow}{g}(\complement_B \underset{\rightarrow}{f}(X)) \\ \complement_A \underset{\rightarrow}{g}(\complement_B \underset{\rightarrow}{f}(X)) \subseteq \complement_A \underset{\rightarrow}{g}(\complement_B \underset{\rightarrow}{f}(Y)) \\ h(X) \subseteq h(Y) \end{gathered}$$

D'après le lemme de Knaster-Tarski appliqué aux ensembles de parties, h admet un point fixe, autrement dit il existe un ensemble $F \in \mathscr{P}(A)$ tel que $F = h(F)$. Donc

$$\complement_A F = \complement_A h(F) = \complement_A \complement_A \underset{\rightarrow}{g}(\complement_B \underset{\rightarrow}{f}(F)) = \underset{\rightarrow}{g}(\complement_B \underset{\rightarrow}{f}(F))$$

On en déduit que la restriction de g à $\complement_B \underset{\rightarrow}{f}(F)$, qui est une fonction injective, a pour image $\complement_A F$, donc induit une bijection $\complement_B \underset{\rightarrow}{f}(F) \xrightarrow{g} \complement_A F$, dont on notera $\complement_A F \xrightarrow{g'} \complement_B \underset{\rightarrow}{f}(F)$ la réciproque. De plus, la restriction de f à F, qui est une fonction injective d'image $\underset{\rightarrow}{f}(F)$, induit une bijection $F \xrightarrow{f} \underset{\rightarrow}{f}(F)$. Or

$$F \cap \complement_A F = \underset{\rightarrow}{f}(F) \cap \complement_B \underset{\rightarrow}{f}(F) = \varnothing$$

Par conséquent la fonction $F \cup \complement_A F \xrightarrow{f \cup g'} \underset{\rightarrow}{f}(F) \cup \complement_B \underset{\rightarrow}{f}(F)$, c'est-à-dire

$$f \cup g' : \begin{cases} A & \longrightarrow B \\ x & \longmapsto \begin{cases} f(x) \text{ si } x \in F \\ g'(x) \text{ si } x \in \complement_A F \end{cases} \end{cases}$$

est une bijection de A dans B.

Preuve 2 (faisant appel à l'ensemble $\mathbb{N}$ des entiers naturels qui sera défini plus loin)
On considère deux injections $A \xrightarrow{f} B$ et $B \xrightarrow{g} A$. On définit par récurrence la suite $(A_n)_{n\in\mathbb{N}}$ de parties de A :

$$\begin{cases} A_0 = \complement_A \operatorname{Im}(g) \\ A_{n+1} = (\underset{\rightarrow}{g \circ f})(A_n) \end{cases}$$

Notons E la réunion de la suite $(A_n)_{n\in\mathbb{N}}$:

$$E :\equiv \bigcup_{n\in\mathbb{N}} A_n = \left\{(g\circ f)^n(x) \mid x \in \complement_A \operatorname{Im}(g), n \in \mathbb{N}\right\}$$

C'est un sous-ensemble de A incluant $\complement_A \operatorname{Im}(g)$, stable par $g\circ f$, puisque si $x \in E$, alors il existe $n \in \mathbb{N}$ tel que $x \in A_n$, donc

$$g\circ f(x) \in A_{n+1} \subseteq E$$

(c'est même le plus petit sous-ensemble de A contenant $\complement_A \operatorname{Im}(g)$, stable par $g\circ f$, puisque par une récurrence immédiate,

Ces pages ne sont pas incluses dans l’aperçu.

- Prouvons que pour tout $x \in A$

$$x \in A_B \iff f(x) \in B_B$$

L'une des implications est immédiate : si $x \in A_B$, alors il existe un entier n et $b \in B \setminus \mathrm{Im}(f)$ tels que $x = g\circ(f\circ g)^n(b)$, donc $f(x) = (f\circ g)^{n+1}(b) \in B_B$. Réciproquement, on considère $x \in A$ tel que $f(x) \in B_B$. Il existe donc un entier n et $b \in B \setminus \mathrm{Im}(f)$ tels que

$$f(x) = (f\circ g)^n(b)$$

Par ailleurs $n \neq 0$ car sinon $f(x) = b$ en contradiction avec $b \notin \mathrm{Im}(f)$. On a donc

$$f(x) = f\circ(g\circ f)^{n-1}\circ g(b)$$

donc par injectivité de f

$$x = (g\circ f)^{n-1}\circ g(b)$$

Si $n = 1$, alors $x = g(b)$, donc $x \in A_B$, et si $n \geqslant 2$, alors

$$x = g\circ(f\circ g)^{n-2}\circ f\circ g(b) = g\circ(f\circ g)^{n-1}(b) \in A_B$$

On déduit de ce qui précède que $x \in A_A \cup A_B$ si et seulement si $f(x) \in B_A \cup B_B$, et par conséquent $\underset{\rightarrow}{f}(A_\infty) = B_\infty$, donc f induit une bijection de A_∞ dans B_∞.

- On peut donc construire une bijection de A dans B en prenant la réunion des bijections $A_A \xrightarrow{f|_{A_A}} B_A$, $A_\infty \xrightarrow{f|_{A_\infty}} B_\infty$, et $A_B \xrightarrow{g|^{-1}_{B_B}} B_B$, c'est-à-dire la fonction $A \xrightarrow{\varphi} B$ telle que pour tout $x \in A$

$$\begin{cases} x \in A_A & \mapsto f(x) \\ x \in A_B & \mapsto g^{-1}(x) \\ x \in A_\infty & \mapsto f(x) \end{cases}$$

Remarque 2.3.27 : Les preuves 2 à 5 utilisent l'ensemble des entiers naturels $\mathbb{N}$ qui n'a pas encore été introduit (voir la section 3.1) ; je les donne pour être complet, car ce sont des preuves classiques. Mais ce théorème peut être démontré uniquement à partir de notions déjà vues comme le montre la preuve 1. Par ailleurs, il n'y a pas de cercle vicieux dans l'exposé des preuves 2 à 5, car la définition de $\mathbb{N}$ et le principe de construction de suites par récurrence, utilisés dans ces preuves, et que nous verrons plus loin, ne font pas appel au théorème de Cantor-Bernstein. Je pourrai dans d'autres théorèmes procéder de la même façon (en indiquant une ou plusieurs preuves s'appuyant sur des notions pas encore vues) sans faire nécessairement cette remarque à chaque fois.

Remarque 2.3.28 (Vocabulaire) : On trouve aussi ce théorème sous le nom de théorème de Schröder-Bernstein ou théorème de Cantor-Schröder-Bernstein.

Remarque 2.3.29 : On déduit du théorème une autre justification du fait que pour tout ensemble E, il n'existe aucune injection de $\mathscr{P}(E)$ dans E : s'il existait une telle injection, alors puisque la fonction $x \mapsto \{x\}$ est une injection de E dans $\mathscr{P}(E)$, les ensembles E et $\mathscr{P}(E)$ seraient en bijection, ce qui est impossible d'après le théorème de Cantor.

Remarque 2.3.30 (Remarque historique [5]) : Le mathématicien allemand Richard Dedekind (1831-1916) a été le premier à prouver ce théorème en 1887 (avec un principe semblable à celui de la preuve 3 ci-dessus), mais sans communiquer son résultat, la démonstration n'étant publiée qu'en 1932 dans le tome 3 du recueil posthume de *Gesammelte mathematische Werke* [6].

5. Sources :
Lorenz HALBEISEN. « Comparing cardinalities in Zermelo's system ». Dans : *Cahiers du Centre de Logique* 17 (2010), p. 9-19.
Encyclopédie Wikipedia : `https://en.wikipedia.org/wiki/Schröder-Bernstein_theorem`.

6. Richard DEDEKIND. *Gesammelte mathematische Werke III* [Œuvres mathématiques rassemblées III]. Sous la dir. de Robert FRICKE, Emmy NOETHER et Öystein ORE. Friedrich Vieweg & Sohn, Braunschweig, 1932.

Ces pages ne sont pas incluses dans l’aperçu.

4. On fait l'hypothèse $A \subseteq B$. Tout élément de aA est de la forme ay, avec $y \in A$, donc aussi $y \in B$. On en déduit $aA \subseteq aB$. De plus, si a admet un symétrique et si $aA \subseteq aB$, alors d'après les résultats précédents

$$a^{-1}(aA) \subseteq a^{-1}(aB)$$
$$(a^{-1}a)A \subseteq (a^{-1}a)B$$
$$eA \subseteq eB$$
$$A \subseteq B$$

Les raisonnements sont semblables pour Aa et Ba.

Définition 2.7.8 (Groupe)
On appelle *groupe* tout monoïde dans lequel tout élément est symétrisable, autrement dit tout ensemble G muni d'une loi de composition interne et d'un élément e, vérifiant les formules suivantes :

1. Associativité : pour tout x, y, z dans G

$$(xy)z = x(yz)$$

2. Élément neutre : pour tout x dans G

$$xe = ex = x$$

3. Symétrique :

$$\forall x \in G, \exists y \in G, yx = xy = e$$

Remarque 2.7.9 : Ce qui a été vu dans le volume 1 (section 8.2) s'applique. On notera en particulier que dans un groupe, tous les éléments x admettent un unique symétrique que l'on note x^{-1}, donc sont simplifiables pour la loi. Je renvoie aussi à cette section pour des exemples de monoïdes et de groupes.

Remarque 2.7.10 : Dans un groupe, les translations à droite et à gauche sont des bijections, car

$$ax = y \iff x = a^{-1}y \qquad \text{et} \qquad xa = y \iff x = ya^{-1}$$

Remarque 2.7.11 : On notera qu'en toute rigueur, un groupe est défini par un triplet $(G, *, e)$, où G est un ensemble, $*$ une loi de composition interne sur G et e un élément de G. Il est aussi possible d'ajouter la fonction $x \mapsto x^{-1}$ (de G dans G) dans la définition, donc de définir un groupe par le quadruplet $(G, *, e, ^{-1})$, la propriété du symétrique, dans la définition, étant remplacée par

$$\forall x \in G, x^{-1}x = xx^{-1} = e$$

Définition 2.7.12
On considère un sous-ensemble H d'un groupe.

$$H^{-1} \overset{\text{def}}{=} \left\{x^{-1} \mid x \in H\right\}$$

et si la notation de la loi est additive (+), on note

$$-H \overset{\text{def}}{=} \{-x \mid x \in H\}$$

Remarque 2.7.13 : On a donc

$$x \in H \iff x^{-1} \in H^{-1} \qquad \text{et} \qquad x \in H^{-1} \iff x^{-1} \in H$$

Définition 2.7.14 (Groupe symétrique)
L'ensemble $\mathcal{S}_E$ des permutations d'un ensemble E (les bijections de E dans E), muni de la composition des fonctions $\circ$, est un groupe, que l'on appelle le *groupe symétrique* de E. L'élément neutre est la fonction identité id_E ($x \mapsto x$), et le symétrique d'une bijection f est sa réciproque f^{-1}.

Preuve
La composition de fonctions est une loi interne sur $\mathcal{S}_E$ (la composée de deux bijections est une bijection), associative, l'identité est une bijection qui est élément neutre pour $\circ$, et pour toute permutation f, la réciproque f^{-1} est aussi une permutation telle que

$$f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_E$$

Il est possible de construire un groupe à partir d'autres groupes de différentes façons. Tout d'abord, si G est un groupe et E un ensemble, on peut munir de façon naturelle l'ensemble $E \longrightarrow G$ des fonctions de E dans G, d'une structure de groupe :

Théorème 2.7.15 (Groupe de l'ensemble des fonctions à valeurs dans un groupe)
On considère un ensemble E et un groupe $(G, \cdot, e)$. Alors $E \longrightarrow G$, muni de la loi de composition $*$ suivante, est un groupe :

$$f * g : \begin{cases} E & \longrightarrow G \\ x & \longmapsto f(x) \cdot g(x) \end{cases}$$

L'élément neutre pour cette opération est la fonction constante égale à e :

$$\begin{cases} E & \longrightarrow G \\ x & \longmapsto e \end{cases}$$

Le symétrique de la fonction f est la fonction :

$$\begin{cases} E & \longrightarrow G \\ x & \longmapsto f(x)^{-1} \end{cases}$$

De plus si G est commutatif, alors $E \longrightarrow G$ l'est aussi.

Preuve
Tout a déjà été démontré précédemment. Si $(G, \cdot, e)$ est un groupe, la loi $\cdot$ est associative donc la loi $*$ l'est aussi, la fonction constante égale à e est élément neutre pour $*$, et puisque G est un groupe, tout élément de G admet un symétrique et toute fonction $E \xrightarrow{f} G$ admet pour symétrique

$$\varphi : \begin{cases} E & \longrightarrow G \\ x & \longmapsto f(x)^{-1} \end{cases}$$

Enfin, nous avons déjà vu que si la loi sur G est commutative, celle sur $E \longrightarrow G$ aussi.

Si G_1 et G_2 sont des groupes, on peut munir le produit cartésien $G_1 \times G_2$ d'une structure de groupe :

Ces pages ne sont pas incluses dans l’aperçu.

2.10 Morphismes et isomorphismes

La notion d'isomorphisme généralise celle d'équipotence, et apparaît dans plusieurs domaines des mathématiques. L'idée informelle est de dire que deux ensembles peuvent être différents, mais se comporter, du point de vue de certaines opérations ou relations, de la même façon. Par exemple, on munit l'ensemble $\{0,1\}$ de l'addition suivante :

$$0+1=1+0=1 \qquad 0+0=0 \qquad 1+1=0$$

Et on munit l'ensemble formé de deux objets distincts quelconques que je noterai a et b, donc l'ensemble $\{a,b\}$, de l'opération $*$ suivante :

$$a*b=b*a=b \qquad a*a=a \qquad b*b=a$$

Même si ces ensembles ne sont pas les mêmes (ils sont formés d'objets différents), ils sont en bijection (à chaque élément de l'un correspond un unique élément de l'autre) et se comportent, vis à vis de leur opération respective $(+,*)$, de la même façon (on peut remplacer 0 par a, 1 par b, $+$ par $*$). On dit que ces ensembles sont isomorphes pour l'opération en question et on peut d'une certaine façon considérer que la structure construite est la même. Dans ce qui suit, je donne une définition très générale des notions de morphisme et d'isomorphisme, mais ces concepts pourront être redéfinis pour certaines structures classiques (morphismes de groupes, morphismes d'anneaux, …).

Définition 2.10.1 (Morphisme, isomorphisme, plongement, pour des opérations)
On considère un entier non nul n, un ensemble E muni d'une opération n-aire φ et un ensemble E' muni d'une opération n-aire φ'. On appelle

1. *morphisme* entre (ou pour) φ et φ' toute fonction $E \xrightarrow{f} E'$ telle que
$$\forall (x_1,\dots,x_n)\in E^n, f(\varphi(x_1,\dots,x_n))=\varphi'(f(x_1),\dots,f(x_n))$$
En particulier on appelle
- morphisme entre les opérations unaires $\intercal$ et $\intercal'$ toute fonction $E \xrightarrow{f} E'$ telle que
$$\forall x\in E, f(\intercal x)=\intercal'(f(x))$$
autrement dit telle que
$$f\circ\intercal=\intercal'\circ f$$
ce qui peut se représenter par le diagramme commutatif suivant :
$$\begin{array}{ccc} E & \xrightarrow{\intercal} & E \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle f} \\ E' & \xrightarrow{\intercal'} & E' \end{array}$$
- morphisme entre les lois $*$ et $\star$ toute fonction $E \xrightarrow{f} E'$ telle que
$$\forall (x,y)\in E\times E, f(x*y)=f(x)\star f(y)$$
ce qui peut se représenter par le diagramme commutatif suivant :
$$\begin{array}{ccc} E\times E & \xrightarrow{*} & E \\ \downarrow{\scriptstyle f\times f} & & \downarrow{\scriptstyle f} \\ E'\times E' & \xrightarrow{\star} & E' \end{array}$$

Ces pages ne sont pas incluses dans l’aperçu.

Preuve
- La fonction identité est un isomorphisme, ce qui prouve la réflexivité.
- La fonction réciproque d'un isomorphisme est un isomorphisme, ce qui prouve la symétrie.
- La composée de deux isomorphismes est un isomorphisme, ce qui prouve la transitivité.

Théorème 2.10.37 (Isomorphisme entre $\mathscr{P}(E)$ et $E \longrightarrow \mathbb{B}$)

1. Pour tout ensemble E, la bijection

$$\chi : \begin{cases} \mathscr{P}(E) & \longrightarrow (E \longrightarrow \mathbb{B}) \\ A & \longmapsto \chi_A \end{cases}$$

définit un isomorphisme

$$\text{de } (\mathscr{P}(E), \subseteq, \cup, \cap, \varnothing, E, \Delta, \complement) \text{ dans } (E \longrightarrow \mathbb{B}, \leqslant, +, \times, 0, 1, \oplus, \bar{}\,)$$

les lois $+, \times, \oplus$ sur l'ensemble $E \longrightarrow \mathbb{B}$ étant les lois induites par les lois internes sur $\mathbb{B}$, le complément $\bar{f}$ étant la fonction $x \mapsto \bar{x}$, et les éléments 0 et 1 étant respectivement les fonctions constantes égales aux éléments 0 et 1 de $\mathbb{B}$.

2. On en déduit que les algèbres de Boole $(\mathscr{P}(E), \subseteq, \cup, \cap, \varnothing, E, \complement)$ et $(E \longrightarrow \mathbb{B}, \leqslant, +, \times, 0, 1, \bar{}\,)$ sont isomorphes, et que les anneaux de Boole $(\mathscr{P}(E), \Delta, \cap, \varnothing, E)$ et $(E \longrightarrow \mathbb{B}, \oplus, \times, 0, 1)$ sont isomorphes :

$$(\mathscr{P}(E), \subseteq, \cup, \cap, \varnothing, E, \complement) \simeq (E \longrightarrow \mathbb{B}, \leqslant, +, \times, 0, 1, \bar{}\,)$$
$$(\mathscr{P}(E), \Delta, \cap, \varnothing, E) \simeq (E \longrightarrow \mathbb{B}, \oplus, \times, 0, 1)$$

Preuve

1. Nous savons déjà que χ est une bijection, et que

$$\begin{cases} \chi_\varnothing = 0 \\ \chi_E = 1 \end{cases}$$

Vérifions les autres propriétés :

- Prouvons que

$$A \subseteq B \iff \chi_A \leqslant \chi_B$$

— Si pour tout $x \in E$, $\chi_A(x) \leqslant \chi_B(x)$, alors si $x \in A$, $\chi_A(x) = 1$, donc $\chi_B(x) = 1$, et par conséquent $x \in B$. On en déduit $A \subseteq B$.

— Réciproquement, on fait l'hypothèse $A \subseteq B$, et on considère $x \in E$.

— Si $x \in A$, alors $x \in B$, donc $1 = \chi_A(x) = \chi_B(x)$, et a fortiori $\chi_A(x) \leqslant \chi_B(x)$.

— Si $x \notin A$, alors $\chi_A(x) = 0$, donc $\chi_A(x) \leqslant \chi_B(x)$.

On en déduit que $\chi_A \leqslant \chi_B$.

- Pour tout $x \in E$

$$\chi_{A\cup B}(x) = 0 \equiv x \notin A \text{ et } x \notin B \equiv \chi_A(x) = 0 \text{ et } \chi_B(x) = 0 \equiv \chi_A(x) + \chi_B(x) = 0$$

donc

$$\chi_{A\cup B} = \chi_A + \chi_B$$

- Pour tout $x \in E$

$$\chi_{A\cap B}(x) = 1 \equiv x \in A \text{ et } x \in B \equiv \chi_A(x) = 1 \text{ et } \chi_B(x) = 1 \equiv \chi_A(x) \times \chi_B(x) = 1$$

Donc

$$\chi_{A\cap B} = \chi_A \times \chi_B$$

Ces pages ne sont pas incluses dans l’aperçu.

Chapitre 3

Théorie des ensembles de Zermelo-Fraenkel (ZF), 3e partie : axiome de l'infini et construction de $\mathbb{N}$

3.1 Axiome de l'infini, ensemble ω des entiers naturels, principe de récurrence

Le prochain axiome va justifier l'existence d'un ensemble infini, et permettre de construire l'ensemble des entiers naturels, à partir de la notion d'*ensemble récurrent*. Nous avons déjà défini $\underline{0}$ (comme étant l'ensemble vide) et le successeur d'un ensemble quelconque x comme

$$Sx \overset{\text{def}}{\equiv} x \cup \{x\}$$

à partir de quoi nous pouvons construire les ensembles

$$\begin{aligned} \underline{1} &\overset{\text{def}}{\equiv} S\underline{0} \\ \underline{2} &\overset{\text{def}}{\equiv} S\underline{1} \\ &\dots \end{aligned}$$

Nous allons maintenant définir un ensemble récurrent comme un ensemble contenant l'ensemble vide ainsi que le successeur de tous ses éléments, ce qui aura en particulier pour conséquence que les ensembles précédents ($\underline{0}, \underline{1}, \dots$) seront automatiquement des éléments d'un ensemble récurrent. Mais aucun des axiomes vus pour l'instant ne permet de justifier l'existence d'un tel ensemble. C'est pourquoi un nouvel axiome est nécessaire.

Définition 3.1.1 (Ensemble récurrent)
On appelle *ensemble récurrent* tout ensemble E contenant l'ensemble vide et stable pour la fonction successeur, c'est-à-dire tel que

$$\begin{cases} \varnothing \in E \\ \forall x \in E, x \cup \{x\} \in E \end{cases}$$

Remarque 3.1.2 (Vocabulaire) : On trouve aussi *ensemble inductif* à la place de *ensemble récurrent*, mais cette expression a aussi d'autres sens en mathématiques, en particulier dans le cadre des ensembles ordonnés, en rapport avec le lemme de Zorn (voir la remarque 5.4.8, p. 515).

Axiome 3.1.3 (Axiome de l'infini)
Il existe un ensemble récurrent :

$$\exists E, \begin{cases} \varnothing \in E \\ \forall x \in E, x \cup \{x\} \in E \end{cases}$$

Définition 3.1.4 (Entier naturel, ensemble des entiers naturels)
L'intersection de tous les ensembles récurrents est un ensemble, que l'on note ω (autrement dit ω est l'ensemble des ensembles n qui appartiennent à tous les ensembles récurrents), dont les éléments s'appellent des *entiers naturels* (ou juste *entiers*) :

$$n \text{ est un entier naturel} \overset{\text{def}}{\iff} \forall E, \left(\begin{cases} \varnothing \in E \\ \forall x \in E, x \cup \{x\} \in E \end{cases} \implies n \in E \right)$$

$$\omega \overset{\text{def}}{=} \bigcap \{E \mid \varnothing \in E \text{ et } (\forall x \in E, x \cup \{x\} \in E)\} \overset{\text{def}}{=} \{n \mid n \text{ est un entier naturel}\}$$

De plus ω est le plus petit (au sens de l'inclusion) des ensembles récurrents, autrement dit c'est un ensemble récurrent tel que pour tout ensemble récurrent E

$$\omega \subseteq E$$

Preuve

- Vérifions que cette définition est cohérente, c'est-à-dire qu'elle permet bien de définir un ensemble. Soit E un ensemble récurrent (il en existe au moins un d'après l'axiome de l'infini). On peut définir un ensemble ω, d'après le schéma de compréhension, par
$$\{n \in E \mid n \text{ est un entier}\}$$
(l'expression « n est un entier » étant un raccourci pour la formule indiquée dans la définition ci-dessus). Cette définition ne dépend pas de E, car si F est un autre ensemble récurrent,
$$\{n \in E \mid n \text{ est un entier}\} = \{n \in F \mid n \text{ est un entier}\}$$
car tout élément de $\{n \in E \mid n \text{ est un entier}\}$ est un entier, et par conséquent il appartient à l'ensemble récurrent F, donc à $\{n \in F \mid n \text{ est un entier}\}$, et réciproquement.
- De plus, $\varnothing$ appartient à ω, et si n est un élément de ω (autrement dit, si n est un entier), alors Sn est aussi un entier (car le successeur d'un élément d'un ensemble récurrent appartient aussi par définition à cet ensemble), donc Sn appartient à ω. Par conséquent ω est un ensemble récurrent.
- Enfin, si E est un ensemble récurrent et si $n \in \omega$, alors n est un entier naturel donc appartient à E, et par conséquent $\omega \subseteq E$.

Remarque 3.1.5 (Vocabulaire) : Le terme *entier* seul a plusieurs sens, car il peut aussi s'appliquer aux entiers relatifs (que nous définirons formellement plus loin), mais le contexte permet souvent de lever l'ambiguïté.

Remarque 3.1.6 : Je pourrai utiliser l'expression « n est un entier » comme raccourci pour la formule du premier ordre associée.

Remarque 3.1.7 : On peut noter que le successeur d'un entier naturel est un entier naturel, puisque si n est un entier naturel, il appartient à tous les ensembles récurrents, donc par définition, son successeur Sn appartient aussi à tous les ensembles récurrents, ce qui signifie que Sn est un entier naturel. Autrement dit ω est stable pour la fonction successeur (comme tous les ensembles récurrents).

Ces pages ne sont pas incluses dans l’aperçu.

autrement dit

$$k \in n \implies Sk \in Sn$$

Preuve
Prouvons par récurrence sur n la formule

$$\forall k \in \omega, (k \in n \implies (Sk \in n \text{ ou } Sk = n))$$

- Au rang $n = 0$, le résultat est trivial puisque $k \notin 0$.
- On fait l'hypothèse de récurrence au rang n, et on considère $k \in \omega$ tel que $k \in Sn \overset{\text{def}}{\equiv} n \cup \{n\}$.
  - Soit $k = n$, et alors $Sk = Sn$;
  - soit $k \in n$, et alors par hypothèse de récurrence $Sk \in n$ ou $Sk = n$, autrement dit $Sk \in Sn$.

  On en déduit $Sk \in Sn$ ou $Sk = Sn$.

Théorème 3.2.16 (Totalité de la relation d'ordre, et caractérisation par la relation d'inclusion) La restriction de la relation d'appartenance ($\in$) à ω est une relation d'ordre strict total, qui coïncide avec la restriction de la relation d'inclusion stricte ($\subset$). On a donc, pour tous les entiers n et p

$$n < p \equiv n \subset p \equiv n \in p$$
$$n \leqslant p \equiv n \subseteq p \equiv (n \in p \text{ ou } n = p)$$

Preuve 1 (faisant appel au lemme précédent)
1. Prouvons d'abord le principe de trichotomie pour la restriction de la relation $\in$ à ω. On considère un entier p. Démontrons par récurrence sur n la formule

   $$p \in n \text{ ou } n \in p \text{ ou } p = n$$

   - Pour $n = 0$: le résultat est établi par la formule que nous avons démontrée précédemment

     $$0 = p \text{ ou } 0 \in p$$

   - On fait l'hypothèse de récurrence au rang n.
     - Si $p \in n$, alors puisque $n \in Sn$, on en déduit par transitivité $p \in Sn$.
     - Si $p = n$, alors on a de même : $p \in Sn$.
     - Si $n \in p$, alors $Sn \in Sp$ d'après le lemme, autrement dit par définition de Sp

       $$Sn \in p \text{ ou } Sn = p$$

2. Nous savons déjà que pour tous les entiers n et p, si $n \in p$ alors $n \subset p$. Réciproquement, si $n \subset p$, alors puisque $p \neq n$ et $p \notin n$ (car sinon $p \subset n$), on en déduit par trichotomie $n \in p$. Par conséquent

   $$n \in p \iff n \subset p$$

3. On en déduit que sur ω, l'inclusion stricte ($\subset$) coïncide avec $\in$, et que c'est une relation d'ordre strict vérifiant le principe de trichotomie, donc l'ordre associé $\subseteq$ est total (le fait que $\subseteq$ soit une relation d'ordre est déjà connu, mais sur la classe des ensembles, cet ordre n'est pas total).

Preuve 2 (ne faisant pas appel au lemme précédent)
1. Prouvons d'abord

   $$n \in p \iff n \subset p$$

   On sait déjà que si $n \in p$, alors $n \subset p$. Réciproquement, on note

   $$A :\equiv \{p \in \omega \mid \forall n \in \omega, (n \subset p \implies n \in p)\}$$

   Démontrons par récurrence que $A = \omega$:

Ces pages ne sont pas incluses dans l’aperçu.

Exemple 3.7.14
Démontrons par récurrence (forte) sur n que tout nombre entier supérieur ou égal à 2 admet un diviseur premier.

- Pour $n = 2$, le résultat est immédiat puisque 2 admet comme diviseur premier 2.
- On fait l'hypothèse de récurrence que la propriété est vérifiée pour tout $k \leqslant n$. Si $n+1$ est un nombre premier, il admet un diviseur premier (lui-même). Sinon, $n + 1$ admet un diviseur d strictement compris entre 1 et $n + 1$, donc compris entre 2 et n. Par hypothèse de récurrence, d admet un diviseur premier, qui est aussi un diviseur premier de $n + 1$.

Enfin, voyons quelques exemples de *fausses* démonstrations par récurrence, c'est-à-dire quelques exemples d'erreurs de raisonnement en rapport avec le principe de récurrence. Les deux erreurs principales consistent

1. à ne pas vérifier que la propriété est correctement initialisée, c'est-à-dire qu'elle est vérifiée pour un certain entier k. Par exemple si
$$\forall n \geqslant k, (\mathscr{F}(n) \implies \mathscr{F}(n+1))$$
mais si la propriété n'est pas vérifiée pour k, on ne peut pas en conclure qu'elle est vérifiée pour tout entier.
2. à prouver la propriété dite *d'hérédité*, le passage de n à $n+1$, à partir d'un entier différent de l'entier initial k. Par exemple si
$$\begin{cases} \mathscr{F}(k) \\ \forall n > k, (\mathscr{F}(n) \implies \mathscr{F}(n+1)) \end{cases}$$
on ne peut pas en conclure que la propriété est vérifiée pour tout entier.

Exemple 3.7.15 (Fausse récurrence)
On va faussement démontrer par récurrence qu'un ensemble de n éléments est formé d'éléments identiques.

- Le rang $n = 1$ est trivial.
- On fait l'hypothèse de récurrence au rang n, et on considère un ensemble de $n + 1$ éléments, que l'on numérote de 1 à $n + 1$. Les n premiers (de 1 à n) forment un ensemble de n éléments, donc ils vérifient l'hypothèse de récurrence, et par conséquent ils sont identiques. Les n derniers (de 2 à $n+1$) vérifient aussi l'hypothèse de récurrence, donc ils sont aussi identiques. Mais les éléments de 2 à n sont communs aux deux ensembles, donc tous les éléments de 1 à $n + 1$ sont identiques.

On en déduit le résultat par récurrence. Où est l'erreur ?

L'erreur vient ici du passage de n à $n + 1$ qui n'est vérifié que pour $n \geqslant 2$, au lieu de $n \geqslant 1$ pour pouvoir appliquer le théorème ; en effet si $n = 1$, l'ensemble des n premiers éléments et des n derniers (de l'ensemble à $n + 1$ éléments) sont disjoints.

Remarque 3.7.16 (Remarque historique) : Ce principe (de raisonnement incorrect par récurrence) apparaît en 1954 dans *Mathematics and Plausible Reasoning*[5] du mathématicien américain (d'origine hongroise) George Pólya (1887-1985), qui l'applique à la couleur des yeux d'un groupe de filles (dans un groupe de n filles, toutes ont les yeux de la même couleur). De manière générale, on peut utiliser n'importe quelle propriété (par exemple, dans un groupe de personnes, toutes ont le même âge, le même nom, le même sexe...). On trouve

5. George PÓLYA. *Mathematics and Plausible Reasoning. Induction and analogy in mathematics*. Princeton University Press, 1954, p. 120.

souvent ce raisonnement en français appliqué à des crayons de couleurs (n crayons de couleurs sont toujours de la même couleur), et en anglais appliqué à des chevaux (tous les chevaux sont de la même couleur). Ce dernier exemple est présent dans un article de Joel E. Cohen paru en 1961 (repris en 1973 dans un livre de Robert L. Weber [6]), et aussi dans *Concrete Mathematics : A Foundation for Computer Science* (1989) [7], où les auteurs citent George Pólya comme leur source, ce qui fait que ce dernier est souvent crédité (à tort) de l'exemple des chevaux [8].

Exemple 3.7.17 (Fausse récurrence)
Un autre exemple de fausse récurrence est donné par l'auteur et mathématicien Richard Johnsonbaugh [9]. On va faussement démontrer par récurrence sur n, que pour tous les entiers a et b, si $\max(a, b) \leqslant n$ alors $a = b$.

- Au rang $n = 0$: si $\max(a, b) \leqslant 0$, alors $a = b = 0$.
- On fait l'hypothèse de récurrence au rang n, et on considère deux entiers a et b tels que $\max(a, b) \leqslant n + 1$. Alors $\max(a - 1, b - 1) \leqslant n$, donc par hypothèse de récurrence, $a - 1 = b - 1$ et par conséquent $a = b$.

On en déduit le résultat par récurrence, et on l'applique avec $a = 0$ et $n = b$: $\max(0, b) = b$, donc $\max(0, b) \leqslant b$, et on en déduit $b = 0$. Par conséquent tous les entiers sont nuls, ce qui est absurde.

L'erreur est dans le passage de n à $n + 1$. Deux entiers a et b quelconques n'ont pas nécessairement un prédécesseur (s'ils sont nuls), c'est-à-dire que $a - 1$ ou $b - 1$ n'est pas nécessairement défini, et on ne peut pas appliquer l'hypothèse de récurrence.

3.8 Construction d'une suite par récurrence

Théorème 3.8.1 (Définition par récurrence, ou définition récursive)
Soit A un ensemble, a un élément de A et une fonction $\mathbb{N} \times A \xrightarrow{f} A$. Alors il existe une unique fonction $\mathbb{N} \xrightarrow{g} A$ vérifiant :

$$\begin{cases} g(0) & = a \\ \forall n \in \mathbb{N}, g(n+1) & = f(n, g(n)) \end{cases}$$

autrement dit, en notation indicielle, il existe une unique suite $(g_n)_{n\in\mathbb{N}}$ d'éléments de A vérifiant

$$\begin{cases} g_0 & = a \\ \forall n \in \mathbb{N}, g_{n+1} & = f(n, g_n) \end{cases}$$

6. Joel E. COHEN. « On the nature of mathematical proofs ». Dans : Robert L. WEBER. *A random walk in science*. Sous la dir. d'Eric MENDOZA. Institute of Physics, 1973, p. 34-36. Article original paru dans *Opus* (1961).

7. Ronald GRAHAM, Donald KNUTH et Oren PATASHNIK. *Concrete Mathematics : A Foundation for Computer Science*. Addison-Wesley Publishing Company, 1989, p. 17.

8. L'erreur apparaît le 14 mars 2006 sur la page Wikipedia anglophone consacrée au raisonnement par récurrence (`https://en.wikipedia.org/wiki/Mathematical_induction`), et le 13 décembre 2006 sur la page (créée le 26 novembre 2002) consacrée à ce faux raisonnement (`https://en.wikipedia.org/wiki/All_horses_are_the_same_color`), ce qui a probablement contribué à sa propagation.

9. Cité dans : Jean-Paul DELAHAYE. *Les inattendus mathématiques*. Pour la science, 2004.

Ces pages ne sont pas incluses dans l’aperçu.

Chapitre 4

Théorie des ensembles de Zermelo-Fraenkel (ZF), 4e partie : éléments de mathématiques discrètes

Prérequis
L'étude de l'ensemble $\mathbb{N}$ du volume 1 (le chapitre 9 sur arithmétique de Peano) et du chapitre 3.

4.1 Ensembles finis, cardinaux d'ensembles finis

Commençons par trois théorèmes en rapport avec les surjections, qui nous seront utiles dans la suite.

Théorème 4.1.1 (Inversibilité à droite des surjections dont le domaine est inclus dans $\mathbb{N}$)
On considère une partie A de $\mathbb{N}$, un ensemble B et une fonction $A \xrightarrow{f} B$. Alors f est surjective si et seulement si f est inversible à droite, autrement dit si et seulement si il existe une fonction $B \xrightarrow{s} A$ (nécessairement injective) telle que

$$f \circ s = \mathrm{id}_B$$

Preuve
Nous avons déjà vu que si f est inversible à droite, alors f est surjective (et son inverse est une injection). Il reste à prouver la réciproque. On considère une fonction surjective $A \xrightarrow{f} B$. Si $B = \emptyset$, alors f ne peut être que la fonction vide (et $A = \emptyset$), et f est trivialement inversible à droite ($\emptyset \circ \emptyset = \emptyset$). Sinon, on peut définir la fonction

$$s : \begin{cases} B & \longrightarrow A \\ x & \longmapsto \min \underset{\leftarrow}{f}(\{x\}) \end{cases}$$

En effet, puisque f est surjective, alors pour tout $x \in B$, $\underset{\leftarrow}{f}(\{x\})$ est une partie non vide de $\mathbb{N}$, donc admet un plus petit élément. Et s est telle que

$$f \circ s = \mathrm{id}_B$$

Théorème 4.1.2 (Corollaire 1)
On considère un ensemble A en bijection avec une partie de $\mathbb{N}$, un ensemble B et une fonction

$A \xrightarrow{f} B$. Alors f est surjective si et seulement si f est inversible à droite, autrement dit si et seulement si il existe une fonction $B \xrightarrow{s} A$ (nécessairement injective) telle que

$$f \circ s = \mathrm{id}_B$$

Preuve
Si f est inversible à droite, alors f est surjective. Réciproquement, on fait l'hypothèse que f est surjective, et que A est en bijection avec une partie de $\mathbb{N}$. Il existe donc un sous-ensemble N de $\mathbb{N}$ et une bijection $N \xrightarrow{g} A$. On en déduit que $f \circ g$ est une fonction surjective de N dans B donc d'après le théorème précédent il existe une fonction $B \xrightarrow{h} N$ telle que

$$f \circ g \circ h = \mathrm{id}_B$$

Par conséquent f est inversible à droite.

Théorème 4.1.3 (Corollaire 2)
On considère un ensemble A en bijection avec une partie de $\mathbb{N}$, et un ensemble B non vide. Alors il existe une surjection de A dans B si et seulement si il existe une injection de B dans A.

Preuve
Nous avons déjà vu qu'une fonction est injective si et seulement si elle est inversible à gauche (lorsque son domaine est non vide), et que son inverse est une surjection, donc s'il existe une injection de B dans A (avec $B \neq \varnothing$), alors il existe une surjection de A dans B. Réciproquement, puisque A est en bijection avec une partie de $\mathbb{N}$, s'il existe une surjection de A dans B alors d'après le théorème précédent il existe une injection de B dans A.

Remarque 4.1.4 : S'il existe une surjection de A dans B, alors il existe une injection de B dans A, même si $B = \varnothing$, puisque dans ce cas, la surjection de A dans B ne peut être que la fonction vide de $\varnothing$ dans $\varnothing$, et la fonction vide est une injection (triviale) de $\varnothing$ dans $\varnothing$.

Remarque 4.1.5 : Avec l'ajout d'un autre axiome, l'axiome du choix, les deux théorèmes précédents seront valables pour tous les ensembles (que A soit en bijection avec une partie de $\mathbb{N}$ ou pas).

Remarque 4.1.6 : Les ensembles en bijection avec une partie de $\mathbb{N}$ sont appelés des ensembles dénombrables. Nous les étudierons plus en détail dans la section 4.12.

Pour définir les ensembles finis, nous allons formaliser l'idée de nombre d'éléments d'un ensemble. Par exemple, l'ensemble $\{a, b, c\}$ (avec a, b, c distincts) a (intuitivement) trois éléments, ce que l'on va traduire en disant qu'il est en bijection avec un autre ensemble à trois éléments de référence, qui peut être par exemple l'ensemble $\{0, 1, 2\}$, c'est-à-dire l'entier naturel 3, ou encore l'ensemble $\{1, 2, 3\}$.

Plus généralement, nous avons vu que nous pouvons formaliser les ensembles $\{1, \ldots, n\}$ et $\{0, \ldots, n-1\}$ par

$$\{1, 2, 3, \ldots, n-1, n\} \overset{\text{def}}{\equiv} [1, n] \overset{\text{def}}{\equiv} \{k \in \mathbb{N} \mid k \geqslant 1 \text{ et } k \leqslant n\}$$
$$\{0, 1, 2, \ldots, n-1\} \overset{\text{def}}{\equiv} [0, n[\overset{\text{def}}{\equiv} \{k \in \mathbb{N} \mid k \geqslant 0 \text{ et } k < n\}$$

avec n entier naturel quelconque ($n \in \mathbb{N}$). Je rappelle aussi que pour tout entier n, ces deux intervalles sont en bijection, et que dans le cas où $n = 0$, on a

$$[1, 0] = [0, 0[= \varnothing$$

et dans le cas où $n \neq 0$

$$[0, n[= [0, n-1]$$

(ce que l'on peut prolonger à $n = 0$ avec la convention $[0, 0-1] \stackrel{\text{def}}{\equiv} \varnothing$).

Il est alors naturel de définir un ensemble fini comme un ensemble en bijection avec un intervalle de la forme $[1, n]$, ou de la forme $[0, n[$, ou directement comme un ensemble en bijection avec un entier (puisque $[0, n[= n$). Et ces trois définitions sont équivalentes puisque $[0, n[$ et $[1, n]$ sont en bijection. Le choix d'un de ces deux intervalles, plutôt que directement l'entier n, a l'avantage de ne pas faire appel à la construction particulière de la relation d'ordre (comme restriction de $\in$ sur ω), pour laquelle $n = [0, n[$, mais uniquement aux propriétés de l'ensemble ordonné $(\mathbb{N}, \leqslant)$. Je choisis les intervalles de la forme $[1, n]$, mais dans ce qui suit, les définitions et théorèmes restent valables quand on remplace $[1, n]$ par $[0, n[$ (ou n).

Définition 4.1.7 (Fini, infini)

1. On appelle *ensemble fini* tout ensemble A tel qu'il existe un entier naturel n tel que A et $[1, n]$ soit en bijection.
2. On appelle *ensemble infini* tout ensemble qui n'est pas fini.
3. On appelle *famille finie* (respectivement *famille infinie*) toute famille indexée par un ensemble fini (respectivement infini).
4. On appelle réunion (respectivement intersection, produit) *finie* (respectivement *infinie*) toute réunion (respectivement intersection, produit) d'une famille finie (respectivement infinie).

Remarque 4.1.8 : On notera que par définition, les entiers naturels (en particulier l'ensemble vide) sont des ensembles finis.

Remarque 4.1.9 : Nous verrons dans cette section quelques propriétés des ensembles finis. Celles des ensembles infinis seront abordées dans la section 4.3.

Pour définir correctement le cardinal d'un ensemble fini (correspondant à l'idée informelle de nombre d'éléments), nous allons devoir justifier que la relation donnée dans la définition précédente (qui à un ensemble fini associe un entier naturel n) est fonctionnelle, c'est-à-dire qu'un ensemble fini ne peut pas être en bijection avec deux entiers naturels différents.

Voyons d'abord dans un lemme une propriété importante des ensembles finis, qui est de ne pas pouvoir être en bijection avec une de leurs parties strictes. Le théorème qui suit s'applique aux intervalles de la forme $[1, n]$, mais sera généralisé à tous les ensembles finis un peu plus loin (théorème 4.1.20, p. 364).

Théorème 4.1.10 (Lemme)

Les trois énoncés suivants de ce théorème sont équivalents.

1. Pour tout entier n, $[1, n]$ n'est en bijection avec aucune de ses parties strictes.
2. Pour tout entier n, toute injection de $[1, n]$ dans $[1, n]$ est bijective.
3. Pour tout entier n et toute fonction $[1, n] \xrightarrow{f} [1, n]$

$$f \text{ est injective } \equiv f \text{ est surjective } \equiv f \text{ est bijective}$$

Ces pages ne sont pas incluses dans l’aperçu.

4.2 Arithmétique des entiers naturels, introduction à l'analyse combinatoire

Prérequis
La construction d'une suite par récurrence (section 3.8), le chapitre 9 du volume 1 sur l'arithmétique de Peano, les ensembles en bijection et les cardinaux (sections 2.3 et 2.4) et plus précisément les cardinaux des ensembles finis (section 4.1).

Il est possible de définir les opérations usuelles sur $\mathbb{N}$ (addition, multiplication, puissance) de façon équivalente :

- Comme application du principe de définition par récurrence, de telle sorte que la structure $(\mathbb{N}, 0, S, +, \times)$ soit un modèle de la théorie de l'arithmétique de Peano.
- À l'aide des cardinaux, en définissant chaque opération comme le cardinal d'un ensemble fini particulier.

Dans ce qui suit, j'utilise la notation Sn, plutôt que $n+1$, pour désigner le successeur de n, afin d'éviter le risque de confusion entre la notation $n+1$ et le symbole $+$ de l'addition que l'on est en train de définir :

Définition 4.2.1 (Addition)
On définit l'*addition*

$$+ : \begin{cases} \mathbb{N}\times\mathbb{N} & \longrightarrow \mathbb{N} \\ (p,n) & \longmapsto p+n \end{cases}$$

de manière équivalente,

1. par récurrence, comme l'unique fonction de $\mathbb{N}\times\mathbb{N}$ dans $\mathbb{N}$ telle que

$$\forall p\, n \in \mathbb{N}, \begin{cases} p+0 & = p \\ p+Sn & = S(p+n) \end{cases}$$

2. par les cardinaux : pour tous les entiers naturels p et n, $[1,p] \sqcup [1,n]$ est un ensemble fini dont on note $p+n$ le cardinal :

$$p+n \overset{\text{def}}{=\!=} |[1,p] \sqcup [1,n]|$$

Preuve

- Vérifions que la définition par récurrence est correcte : on applique le principe de définition par récurrence (avec paramètre) aux fonctions

$$a : \begin{cases} \mathbb{N} & \longrightarrow \mathbb{N} \\ p & \longmapsto p \end{cases} \qquad \text{et} \qquad f : \begin{cases} \mathbb{N}\times\mathbb{N}\times\mathbb{N} & \longrightarrow \mathbb{N} \\ (p,n,x) & \longmapsto Sx \end{cases}$$

Il existe donc une unique fonction $\mathbb{N}\times\mathbb{N} \xrightarrow{g} \mathbb{N}$ telle que

$$\forall p\, n \in \mathbb{N}, \begin{cases} g(p,0) & = p \\ g(p,Sn) & = S(g(p,n)) \end{cases}$$

et on note

$$p+n \overset{\text{def}}{=\!=} g(p,n)$$

- On considère un entier p. Prouvons par récurrence sur n que $[1,p] \sqcup [1,n]$ est un ensemble fini.

— Pour $n=0$,

$$[1,p] \sqcup [1,0] \simeq [1,p]$$

qui est un ensemble fini de cardinal p.

Ces pages ne sont pas incluses dans l’aperçu.

4.3 Ensembles infinis, cardinal $\aleph_0$

Commençons par quelques propriétés des ensembles infinis, qui se déduisent simplement de celles des ensembles finis.

Le théorème qui suit traduit l'idée que si un ensemble est « plus grand » qu'un ensemble infini, alors il est aussi infini.

Théorème 4.3.1
Si A est un ensemble infini et si $A \subseteq B$ (respectivement : s'il existe une injection de $A \longrightarrow B$, s'il existe une surjection de $B \longrightarrow A$), alors B est infini.

Preuve
On considère deux ensembles A et B avec $A \neq \emptyset$ tels que $A \subseteq B$ (respectivement : il existe une injection de $A \longrightarrow B$, il existe une surjection de $B \longrightarrow A$). Si B est un ensemble fini, alors A est un ensemble fini. Donc par contraposition si A est infini, alors B est infini.

Théorème 4.3.2 (Corollaire)
1. Si on ajoute ou si on enlève un nombre fini d'éléments à un ensemble infini A, on obtient un ensemble infini : pour tout ensemble fini B, $A \cup B$ et $A \setminus B$ sont des ensembles infinis.
2. Si A est un ensemble infini, alors $\mathscr{P}(A)$ est un ensemble infini.
3. Si $A_1, \dots, A_n$ sont des ensembles dont l'un (au moins) est infini, alors $\bigcup_{i=1}^{n} A_i$ est un ensemble infini.
4. Si $A_1, \dots, A_n$ sont des ensembles non vides dont l'un (au moins) est infini, alors $\prod_{i=1}^{n} A_i$ est un ensemble infini.

Preuve
1. On considère un ensemble infini A et un ensemble fini B. Puisque $A \subseteq A \cup B$, on déduit du théorème précédent que $A \cup B$ est infini. Par ailleurs
$$A = (A \setminus B) \cup (A \cap B)$$
avec $A \cap B$ ensemble fini (il est inclus dans B). Par conséquent $A \setminus B$ est infini, car s'il était fini A serait un ensemble fini.
2. On considère un ensemble infini A. La fonction
$$\begin{cases} A & \longrightarrow \mathscr{P}(A) \\ x & \longmapsto \{x\} \end{cases}$$
est injective, donc $\mathscr{P}(A)$ est infini.
3. On considère des ensembles $A_1, \dots, A_n$, dont l'un (au moins) est infini, par exemple A_j. Puisque $A_j \subseteq \bigcup_{i=1}^{n} A_i$, on en déduit que $\bigcup_{i=1}^{n} A_i$ est infini.
4. On considère des ensembles non vides $A_1, \dots, A_n$, dont l'un (au moins) est infini
   - Pour justifier que le produit cartésien est infini, prouvons d'abord le résultat dans le cas de deux ensembles : considérons deux ensembles non vides A et B avec A infini (le raisonnement est semblable si B est infini). Puisque $B \neq \emptyset$, il existe $b \in B$. La fonction
$$\begin{cases} A & \longrightarrow A \times B \\ x & \longmapsto (x, b) \end{cases}$$
est injective, donc $A \times B$ est infini.

- Prouvons par récurrence sur $n \geqslant 1$ que si $A_1, \dots, A_n$ sont des ensembles non vides dont l'un (au moins) est infini, alors leur produit cartésien est infini.
  - — Le résultat est immédiat pour $n = 1$.
  - — On fait l'hypothèse de récurrence au rang n, et on considère les ensembles non vides $A_1, \dots, A_{n+1}$ dont l'un au moins est infini. On a
$$\prod_{i=1}^{n+1} A_i \simeq \left(\prod_{i=1}^{n} A_i\right) \times A_{n+1}$$
(on notera que cette formule est valable pour la formalisation des produits cartésiens par des ensembles de listes). Soit l'un des A_i, avec $i \leqslant n$, est infini, et alors par hypothèse de récurrence, $\prod_{i=1}^{n} A_i$ est infini, soit A_{n+1} est infini. On déduit du point précédent que $\prod_{i=1}^{n+1} A_i$ est infini.

Théorème 4.3.3 (Caractérisation d'un ensemble infini)
Un ensemble A est infini si et seulement si pour tout entier n, il existe un sous-ensemble de A fini de cardinal n.

Preuve
- On considère un ensemble A infini. Démontrons par récurrence sur n qu'il existe un sous-ensemble fini B de A de cardinal n.
  - — C'est immédiat pour $n = 0$, car $\varnothing \subseteq A$.
  - — On fait l'hypothèse de récurrence au rang n. Il existe donc $B \subseteq A$ en bijection avec $[1, n]$. De plus $B \neq A$ car sinon A serait fini, donc $A \setminus B \neq \varnothing$. Il existe donc $x \in A \setminus B$ et on a
$$B \cup \{x\} \subseteq A$$
De plus, puisque $x \notin B$, $B \cup \{x\}$ est un ensemble fini tel que
$$|B \cup \{x\}| = |B| + 1 = n + 1$$
- Réciproquement, si pour tout entier n, il existe un sous-ensemble de A fini de cardinal n, alors A est infini, car si A est fini de cardinal p, tout sous-ensemble de A est fini de cardinal inférieur ou égal à p, ce qui est contradictoire.

Nous avons vu que toute injection d'un ensemble fini dans lui-même est surjective. Par contraposition, on en déduit que s'il existe une injection non surjective de E dans lui-même (ou, ce qui revient au même, si E est en bijection avec une de ses parties strictes), alors E est infini. Un tel ensemble est dit *infini au sens de Dedekind*. Cette propriété est parfois prise comme définition d'un ensemble infini. Elle a la particularité de définir les ensembles finis et infinis sans faire référence aux nombres entiers. Les deux notions ne sont pas équivalentes à partir des axiomes vus jusqu'à présent, mais le seront avec l'ajout d'un autre axiome, l'axiome du choix.

Théorème 4.3.4 (Ensemble infini au sens de Dedekind)
Pour tout ensemble A, les quatre propriétés suivantes sont équivalentes :

1. Il existe une injection de $\mathbb{N}$ dans A.
2. Il existe un sous-ensemble de A en bijection avec $\mathbb{N}$.
3. Il existe une injection non surjective de A dans A.
4. A est en bijection avec une de ses parties strictes.

Ces pages ne sont pas incluses dans l’aperçu.

4.8 Sommes et produits dans un monoïde : notations Σ et Π, méthodes de calculs

Prérequis
La construction d'une suite par récurrence (section 3.8), les règles de calculs dans un monoïde (section 2.7) et dans un anneau (section 2.9).

Définitions

Le principe de définition par récurrence va nous permettre de formaliser l'écriture de sommes ou produits de nombres entiers, de la forme

$$a_1 + a_2 + \cdots + a_n \qquad \text{ou} \qquad a_1 \times a_2 \cdots \times a_n$$

De façon plus générale, on peut définir de façon formelle (sans les points de suspension), dans un ensemble muni d'une loi de composition interne associative *, l'expression

$$a_1 * a_2 * \cdots * a_n$$

Lorsque la loi est additive (notée +), on utilise le symbole Σ

$$\sum_{k=1}^{n} a_k \overset{\text{def}}{=\!=} a_1 + a_2 + \cdots + a_n$$

et sinon (en particulier pour une loi multiplicative, notée $\times$ ou $\cdot$), j'utiliserai le symbole Π (d'autres symboles peuvent être parfois utilisés pour des lois ni additives ni multiplicatives, mais il ne me semble pas utile de multiplier les notations)

$$\prod_{k=1}^{n} a_k \overset{\text{def}}{=\!=} a_1 * a_2 * \cdots * a_n$$

Plus précisément :

Définition 4.8.1 (Notations Σ et Π)
On considère un ensemble A muni d'une loi associative *, un entier naturel p, et une suite $(a_k)_{k\in\mathbb{N}}$ d'éléments de A.

1. Pour tout entier $n \geqslant p$, on définit $\displaystyle\prod_{k=p}^{n} a_k$ par récurrence sur n :

$$\begin{cases} \displaystyle\prod_{k=p}^{p} a_k = a_p \\ \displaystyle\prod_{k=p}^{n+1} a_k = \prod_{k=p}^{n} a_k * a_{n+1} \end{cases}$$

Autrement dit

$$\prod_{k=p}^{n} a_k = a_p * a_{p+1} * \cdots * a_n$$

Ces pages ne sont pas incluses dans l’aperçu.

$$\sum_{x\in\{a,b,c\}} f_x = f_a + f_b + f_c$$

Théorème 4.8.28 (Propriétés élémentaires de sommes et produits indexés par un ensemble fini)

1. On considère un monoïde A, deux ensembles finis et disjoints I et J, et une famille $(a_i)_{i\in I\cup J}$ d'éléments de A. Alors
$$\sum_{i\in I\cup J} a_i = \sum_{i\in I} a_i + \sum_{j\in J} a_j$$
2. On considère un monoïde A, un ensemble fini I et une famille $(a_i)_{i\in I}$ d'éléments de A. Si x est un élément de A tel que
$$\forall i \in I, x + a_i = a_i + x$$
(ce qui est en particulier le cas si la loi est commutative), alors
$$x + \sum_{i\in I} a_i = \sum_{i\in I} a_i + x$$
3. On considère un monoïde commutatif A, un ensemble fini I et deux familles $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ d'éléments de A. Alors
$$\sum_{i\in I}(a_i + b_i) = \sum_{i\in I} a_i + \sum_{i\in I} b_i$$

Preuve

1. Prouvons que si I et J sont des ensembles finis tels que $I \cap J = \emptyset$, alors
$$\sum_{i\in I\cup J} a_i = \sum_{i\in I} a_i + \sum_{j\in J} a_j$$
On note n et p les cardinaux respectifs de I et J et on considère deux bijections $[1,n] \xrightarrow{f} I$ et $[1,p] \xrightarrow{g} J$, d'où on déduit la bijection
$$h : \begin{cases} [1,n+p] & \longrightarrow I \cup J \\ k & \longmapsto \begin{cases} f(k) \text{ si } k \leqslant n \\ g(k-n) \text{ sinon} \end{cases} \end{cases}$$
On a
$$\sum_{i\in I\cup J} a_i = \sum_{k=1}^{n+p} a_{h(k)} = \sum_{k=1}^{n} a_{h(k)} + \sum_{k=n+1}^{n+p} a_{h(k)} = \sum_{k=1}^{n} a_{f(k)} + \sum_{k=n+1}^{n+p} a_{g(k-n)} = \sum_{k=1}^{n} a_{f(k)} + \sum_{k=1}^{p} a_{g(k)} = \sum_{i\in I} a_i + \sum_{j\in J} a_j$$
2. On considère une bijection $[1,n] \xrightarrow{f} I$. Si x commute avec les a_i
$$x + \sum_{i\in I} a_i = x + \sum_{k=1}^{n} a_{f(k)} = \sum_{k=1}^{n} a_{f(k)} + x = \sum_{i\in I} a_i + x$$
3. On considère une bijection $[1,n] \xrightarrow{f} I$.
$$\sum_{i\in I}(a_i + b_i) = \sum_{k=1}^{n}(a_{f(k)} + b_{f(k)}) = \sum_{k=1}^{n} a_{f(k)} + \sum_{k=1}^{n} b_{f(k)} = \sum_{i\in I} a_i + \sum_{i\in I} b_i$$

Ces pages ne sont pas incluses dans l’aperçu.

Remarque 4.9.11 (Vocabulaire) : On trouve aussi ce théorème sous le nom *formule du crible*, *formule de Poincaré*, ou *principe d'inclusion-exclusion*.

Remarque 4.9.12 : La formule peut s'écrire sous forme plus détaillée

$$\left|\bigcup_{k=1}^{n} A_k\right| = \sum_{1\leqslant i_1\leqslant n}\left|A_{i_1}\right| - \sum_{1\leqslant i_1<i_2\leqslant n}\left|A_{i_1}\cap A_{i_2}\right| + \sum_{1\leqslant i_1<i_2<i_3\leqslant n}\left|A_{i_1}\cap A_{i_2}\cap A_{i_3}\right| + \cdots + (-1)^{n-1}\sum_{1\leqslant i_1<\cdots<i_n\leqslant n}\left|A_{i_1}\cap\cdots\cap A_{i_n}\right|$$

Exemple 4.9.13 (Formule du crible)

1. Formule du crible pour deux ensembles :

$$|A\cup B| = |A| + |B| - |A\cap B|$$

2. Formule du crible pour trois ensembles :

$$|A\cup B\cup C| = |A| + |B| + |C| - |A\cap B| - |A\cap C| - |B\cap C| + |A\cap B\cap C|$$

3. Formule du crible pour quatre ensembles :

$$\begin{aligned}\left|A_1\cup A_2\cup A_3\cup A_4\right| = {}& \left|A_1\right| + \left|A_2\right| + \left|A_3\right| + \left|A_4\right| \\ & - \left|A_1\cap A_2\right| - \left|A_1\cap A_3\right| - \left|A_1\cap A_4\right| - \left|A_2\cap A_3\right| - \left|A_2\cap A_4\right| - \left|A_3\cap A_4\right| \\ & + \left|A_1\cap A_2\cap A_3\right| + \left|A_1\cap A_2\cap A_4\right| + \left|A_1\cap A_3\cap A_4\right| + \left|A_2\cap A_3\cap A_4\right| \\ & - \left|A_1\cap A_2\cap A_3\cap A_4\right|\end{aligned}$$

4.10 Nombres premiers

Définition 4.10.1 (Nombre premier)

Les définitions suivantes d'un nombre premier sont équivalentes :

1. On dit qu'un nombre entier p est *premier* lorsqu'il admet exactement deux diviseurs (distincts) :

$$p \text{ est premier} \overset{\text{def}}{\iff} |\{n\in\mathbb{N} \mid n \mid p\}| = 2$$

2. On dit qu'un nombre entier p est *premier* lorsque $p > 1$ et que l'ensemble de ses diviseurs est $\{1, p\}$:

$$p \text{ est premier} \overset{\text{def}}{\iff} \begin{cases} p > 1 \\ \{n\in\mathbb{N} \mid n \mid p\} = \{1, p\} \end{cases}$$

autrement dit, puisque tout entier est divisible par 1 et par lui-même, p est premier lorsque $p > 1$ et que les seuls diviseurs de p sont 1 et p :

$$p \text{ est premier} \overset{\text{def}}{\iff} \begin{cases} p > 1 \\ \forall n\in\mathbb{N}, (n \mid p \implies n = 1 \text{ ou } n = p) \end{cases}$$

3. On dit qu'un nombre entier p est *premier* lorsque $p > 1$ et que p ne peut pas être le produit de deux entiers différents de 1 (c'est-à-dire que si p est le produit de deux nombres, au moins l'un des deux est 1) :

$$p \text{ est premier } \stackrel{\text{def}}{\equiv} \begin{cases} p > 1 \\ \forall a\ b \in \mathbb{N}, (p = ab \implies a = 1 \text{ ou } b = 1) \end{cases}$$

On notera $\mathbb{P}$ l'ensemble des nombres premiers.

Preuve (de l'équivalence des définitions)
Il s'agit essentiellement de la même définition présentée de façon différente. Prouvons en détail l'équivalence par implications circulaires.

- On fait l'hypothèse que p admet deux diviseurs distincts. On en déduit que $p > 1$ (car 0 a une infinité de diviseurs et 1 a un seul diviseur), et que ces diviseurs sont nécessairement 1 et p (car tout entier est divisible par 1 et par lui-même). Par conséquent, si n divise p, alors $n = 1$ ou $n = p$.
- On fait l'hypothèse que $p > 1$ et que ses seuls diviseurs sont 1 et p. Alors si $p = ab$, cela signifie par définition que b est un diviseur de p donc soit $b = 1$, soit $b = p$ et alors $a = 1$.
- On fait l'hypothèse que $p > 1$ et que p ne peut pas être le produit de deux entiers différents de 1. Donc si n est un diviseur de p, alors il existe par définition un entier k tel que $p = kn$ et par conséquent soit $n = 1$, soit $k = 1$ et alors $n = p$. Donc

$$\{n \in \mathbb{N} \mid n \mid p\} = \{1, p\}$$

et puisque $p > 1$, le cardinal de cet ensemble est 2.

Remarque 4.10.2 : On aurait pu, dans la définition d'un nombre premier, remplacer la condition $p > 1$ par $p \neq 1$, puisque tous les entiers divisent 0 et donc 0 ne vérifie ni la formule

$$\forall n \in \mathbb{N}, (n \mid p \implies n = 1 \text{ ou } n = p)$$

ni la formule

$$\forall a\ b \in \mathbb{N}, (p = ab \implies a = 1 \text{ ou } b = 1)$$

Remarque 4.10.3 : Si p et q sont deux nombres premiers tels que $p \mid q$, alors $p = q$ (puisque $p \neq 1$).

Remarque 4.10.4 : En prenant la négation de la troisième définition, on voit qu'un entier $n > 1$ n'est pas premier si et seulement si il existe deux entiers a et b tels que

$$n = ab \text{ et } a \neq 1 \text{ et } b \neq 1$$

et puisque a et b ne peuvent pas être nuls (car $n \neq 0$), on obtient la caractérisation suivante des nombres non premiers : si $n > 1$, n n'est pas un nombre premier si et seulement si il existe deux entiers a et b strictement supérieurs à 1 tels que $n = ab$.

Exemple 4.10.5
Les nombres premiers inférieurs à 100 sont

2 3 5 7 11 13 17 19 23 29 31 37 41
43 47 53 59 61 67 71 73 79 83 89 97

Pour le justifier, il suffit, pour tout entier n entre 2 et 100, de chercher les diviseurs de n (en vérifiant pour chaque entier inférieur à n s'il divise n ou pas). Mais il y a des méthodes plus rapides, comme

Ces pages ne sont pas incluses dans l’aperçu.

Théorème 4.11.12 (Corollaire)
Pour tout cardinal α

$$\alpha \geqslant \aleph_0 \iff \alpha + \aleph_0 = \alpha$$

Preuve
D'après la caractérisation par l'addition de la relation d'ordre sur les cardinaux, si $\alpha = \aleph_0 + \alpha$, alors $\aleph_0 \leqslant \alpha$. Réciproquement, si $\aleph_0 \leqslant \alpha$, alors il existe un cardinal γ tel que $\alpha = \aleph_0 + \gamma$, donc

$$\aleph_0 + \alpha = \aleph_0 + \aleph_0 + \gamma = \aleph_0 + \gamma = \alpha$$

Théorème 4.11.13 (Propriétés de la multiplication de cardinaux)
1. Associativité : pour tous les cardinaux α, β, γ

$$\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$$

2. Commutativité : pour tous les cardinaux α et β

$$\alpha \cdot \beta = \beta \cdot \alpha$$

3. 1 est élément neutre : pour tout cardinal α

$$\alpha \cdot 1 = 1 \cdot \alpha = \alpha$$

4. 0 est élément absorbant : pour tout cardinal α

$$\alpha \cdot 0 = 0 \cdot \alpha = 0$$

5. Distributivité de la multiplication sur l'addition : pour tous les cardinaux α, β, γ

$$\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$$
$$(\beta + \gamma) \cdot \alpha = \beta \cdot \alpha + \gamma \cdot \alpha$$

Preuve
Nous avons déjà vu que pour tous les ensembles A, B, C

$$A \times (B \times C) \simeq (A \times B) \times C$$

d'où l'on déduit l'associativité de la multiplication. De même

$$A \times B \simeq B \times A$$

d'où l'on déduit la commutativité,

$$A \times \{\varnothing\} \simeq A$$

d'où l'on déduit que 1 est élément neutre,

$$A \times \varnothing = \varnothing$$

d'où l'on déduit que 0 est élément absorbant,

$$A \times (B \sqcup C) \simeq (A \times B) \sqcup (A \times C)$$

d'où l'on déduit la distributivité de la multiplication sur l'addition.

Ces pages ne sont pas incluses dans l’aperçu.

Remarque 5.1.7 : Je rappelle qu'un ensemble A peut toujours être muni d'un bon ordre s'il existe une surjection d'un ensemble bien ordonné dans A, ou une bijection de A dans un ensemble bien ordonné. En particulier, si $\bigcup_{X\in\mathscr{E}} X$ est un ensemble dénombrable non vide, alors il existe une fonction de choix sur $\mathscr{E}\setminus\{\varnothing\}$, et si E est un ensemble dénombrable non vide, alors il existe une fonction de choix sur $\mathscr{P}(E)\setminus\{\varnothing\}$, puisqu'un ensemble non vide A est dénombrable si et seulement si il existe une surjection de $\mathbb{N}$ (ensemble bien ordonné) dans A.

Mais l'existence d'une fonction de choix pour tout ensemble ne peut pas être justifiée à partir des axiomes existants [1]. Il est pour cela nécessaire d'ajouter un nouvel axiome :

Axiome 5.1.8 (Axiome du choix)
Les huit énoncés suivants de l'axiome du choix sont équivalents :

1. Pour tout ensemble $\mathscr{E}$ d'ensembles non vides, il existe une fonction de choix sur $\mathscr{E}$: pour tout $\mathscr{E}$
$$\varnothing\notin\mathscr{E} \implies \exists\, \mathscr{E}\xrightarrow{f}\bigcup\mathscr{E}\ ,\forall A\in\mathscr{E}, f(A)\in A$$
2. Pour tout ensemble E, il existe une fonction de choix sur $\mathscr{P}(E)\setminus\{\varnothing\}$: pour tout E
$$\exists\, \mathscr{P}(E)\setminus\{\varnothing\}\xrightarrow{f}E\ ,\forall A\in\mathscr{P}(E)\setminus\{\varnothing\}, f(A)\in A$$
3. Le produit cartésien d'une famille d'ensembles non vides, est non vide, c'est-à-dire que pour toute famille d'ensembles $(A_i)_{i\in I}$:
$$(\forall i\in I, A_i\neq\varnothing) \implies \prod_{i\in I} A_i\neq\varnothing$$
4. Pour tout ensemble $\mathscr{E}$ d'ensembles non vides et deux à deux disjoints, il existe un sous-ensemble C de $\bigcup\mathscr{E}$ ayant exactement un élément en commun avec chaque élément de $\mathscr{E}$: pour tout $\mathscr{E}$
$$(\varnothing\notin\mathscr{E} \text{ et } \forall X\ Y\in\mathscr{E}, (X\neq Y \implies X\cap Y=\varnothing)) \implies \exists C\subseteq\bigcup\mathscr{E}, \forall X\in\mathscr{E}, \exists x, C\cap X=\{x\}$$
5. Pour toute partition $\mathscr{P}$ d'un ensemble E, il existe un sous-ensemble C de E ayant exactement un élément en commun avec chaque élément de $\mathscr{P}$, autrement dit si $\mathscr{P}$ est une partition de E :
$$\exists C\subseteq E, \forall X\in\mathscr{P}, \exists x, C\cap X=\{x\}$$
6. Toute relation d'équivalence $\sim$ sur un ensemble E admet un ensemble de représentants, c'est-à-dire qu'il existe un sous-ensemble C de E ayant exactement un élément en commun avec chaque classe d'équivalence, autrement dit si $\sim$ est une relation d'équivalence sur E :
$$\exists C\subseteq E, \forall X\in E/\!\sim, \exists x, C\cap X=\{x\}$$
7. Toute fonction surjective est inversible à droite, autrement dit si $A\xrightarrow{f}B$ est surjective, il existe une fonction $B\xrightarrow{g}A$ telle que
$$f\circ g=\mathrm{id}_B$$
8. Pour tous les ensembles A et B et toute relation
$$\mathscr{R}\subseteq A\times B$$
il existe une fonction $f\subseteq\mathscr{R}$ de domaine $\mathrm{dom}(\mathscr{R})$, autrement dit une fonction $\mathrm{dom}(\mathscr{R})\xrightarrow{f}B$ telle que
$$\forall x\in\mathrm{dom}(\mathscr{R}), x\mathscr{R}f(x)$$

1. Voir la remarque 5.1.12, p. 501.

Ces pages ne sont pas incluses dans l’aperçu.

De plus

$$i = f \circ g(i) = j$$

donc

$$g(i) = (i, x)$$

avec $x \in A_i$, et par conséquent $pr_2 \circ g(i) \in A_i$. Notons, même si ce n'est pas nécessaire pour la suite de la démonstration, que la réciproque (l'énoncé 3 implique l'énoncé 7) est immédiate : si $A \xrightarrow{f} B$ est une surjection,

$$\prod_{y \in B} \underset{\leftarrow}{f}(\{y\}) \neq \emptyset$$

et toute fonction $g \in \prod_{y \in B} \underset{\leftarrow}{f}(\{y\})$ est inverse à droite de f ($f \circ g = \mathrm{id}_B$).

7. On démontre enfin que l'énoncé 3 implique l'énoncé 1 comme dans la preuve précédente.

Remarque 5.1.9 : Dans le cas où $I = \emptyset$, la famille $(A_i)_{i \in I}$ est la fonction vide (l'ensemble vide) et

$$\prod_{i \in I} A_i = \{\emptyset\}$$

Remarque 5.1.10 : L'une des variantes de cet axiome est que si une fonction est surjective, alors elle est inversible à droite. Nous avons déjà vu que la réciproque est aussi vraie (indépendamment de l'axiome du choix). En effet, s'il existe une fonction $B \xrightarrow{g} A$ telle que $f \circ g = \mathrm{id}_B$, alors pour tout y dans B, $y = f(g(y))$, donc f est surjective. Par conséquent une fonction est surjective si et seulement si elle est inversible à droite (voir aussi la section 5.3 pour d'autres propriétés en rapport avec la composition de fonctions).

Remarque 5.1.11 : L'axiome du choix signifie, de façon informelle, que si on se donne une collection d'ensembles, on peut choisir simultanément un élément dans chacun d'entre eux. Comme nous l'avons vu précédemment, pour justifier l'existence d'une telle fonction de choix, l'axiome n'est pas nécessaire s'il y a un nombre fini d'ensembles, ou si un procédé permet de *choisir* un élément (par exemple en utilisant les propriétés des ensembles bien ordonnés). Par contre, si le nombre d'ensembles est infini et qu'on ne connaît pas de procédé pour choisir, l'axiome du choix est nécessaire. Dans *Introduction to mathematical philosophy* (1919), le mathématicien, logicien et philosophe britannique Bertrand Russell (1872-1970) propose une analogie qui est souvent résumée de la manière suivante :

> Pour choisir une chaussette pour chaque paire d'une collection infinie, on a besoin de l'axiome du choix. Mais pour les chaussures, ce n'est pas la peine.

En effet, il existe un procédé pour choisir une chaussure dans une paire (par exemple, choisir la chaussure gauche), donc l'axiome du choix n'est pas nécessaire. Mais ce n'est pas le cas pour les deux chaussettes d'une paire, a priori identiques.

Remarque 5.1.12 (Remarque historique) : L'axiome du choix apparaît explicitement pour la première fois en 1890 dans un article sur les équations différentielles du mathématicien et linguiste italien Giuseppe Peano (1858-1932), il est mentionné par le mathématicien italien Beppo Levi (1875-1961) en 1902, mais c'est en 1904 que le mathématicien allemand Ernst Zermelo (1871-1953) introduit formellement cet axiome (sans le nommer), pour prouver, dans le théorème qui porte son nom, que tout ensemble peut être muni d'un bon ordre :

> « La présente preuve est basée sur l'hypothèse que les affectations γ existent, c'est-à-dire sur le principe que même pour un ensemble infini d'ensembles, il y a toujours des correspondances dans lesquelles à chaque ensemble correspond l'un de ses éléments, ou, exprimé de façon formelle, que le produit d'un ensemble infini d'ensembles, dont chacun contient au moins un élé-

Ces pages ne sont pas incluses dans l’aperçu.

- Si f est surjective et $g \circ f = h \circ f$, alors $g = h$: en effet, il existe une fonction $B \xrightarrow{f'} A$ telle que $f \circ f' = \mathrm{id}_B$, donc

$$\begin{aligned} g \circ f &= h \circ f \\ g \circ f \circ f' &= h \circ f \circ f' \\ g &= h \end{aligned}$$

5.4 Lemme de Zorn, théorème de Zermelo, principe de maximalité de Hausdorff, lemme de Teichmüller-Tukey

Prérequis
Les relations de bon ordre (section 1.13).

Nous allons étudier plusieurs théorèmes classiques équivalents à l'axiome du choix, en commençant par ce qu'on appelle le *lemme de Zorn*. Prouvons d'abord un théorème qui sera utile à sa démonstration, ainsi qu'à celle du théorème de Zermelo, et que je désignerai par l'expression *lemme fondamental*.

Théorème 5.4.1 (Lemme fondamental)
On considère deux familles $(A_i)_{i\in I}$ et $(\leqslant_i)_{i\in I}$ telles que

- pour tout $i \in I$, $\leqslant_i$ est une relation de bon ordre sur A_i ;
- pour tout i et j dans I, A_i est un segment initial de $(A_j, \leqslant_j)$ et $\leqslant_i$ est la relation d'ordre induite par $\leqslant_j$ sur A_i, ou A_j est un segment initial de $(A_i, \leqslant_i)$ et $\leqslant_j$ est la relation d'ordre induite par $\leqslant_i$ sur A_j.

On note

$$A :\equiv \bigcup_{i\in I} A_i \qquad \leqslant :\equiv \bigcup_{i\in I} \leqslant_i$$

Alors $\leqslant$ est une relation de bon ordre sur A, telle que pour tout $j \in I$, A_j est un segment initial de $(A, \leqslant)$, et $\leqslant_j$ est la relation induite par $\leqslant$ sur A_j.

Preuve
- Prouvons que $\leqslant$ est une relation d'ordre sur A :
  - — Réflexivité : si $x \in A$, il existe $i \in I$ tel que $x \in A_i$, donc $x \leqslant_i x$ et par conséquent $x \leqslant x$.
  - — Antisymétrie : si $x \in A$ et $y \in A$ sont tels que $x \leqslant y$ et $y \leqslant x$, il existe i et j dans I tels que $x \leqslant_i y$ et $y \leqslant_j x$. Par hypothèse, $\leqslant_i$ est la relation induite par $\leqslant_j$ sur A_i, ou $\leqslant_j$ est la relation induite par $\leqslant_i$ sur A_j. On fait par exemple l'hypothèse que $\leqslant_i$ est la relation induite par $\leqslant_j$ (le raisonnement est semblable dans l'autre cas). On en déduit $(x, y) \in A_j \times A_j$, $x \leqslant_j y$ et $y \leqslant_j x$, donc $x = y$.
  - — Transitivité : si $x \leqslant y$ et $y \leqslant z$, on en déduit de même qu'il existe i et j dans I tels que par exemple $x \leqslant_j y$ et $y \leqslant_j z$, donc $x \leqslant_j z$ et par conséquent $x \leqslant z$.
- Prouvons que $(A, \leqslant)$ est bien ordonné. On considère une partie non vide P de A. Il existe donc $i \in I$ tel que $P \cap A_i \neq \varnothing$. Par conséquent $P \cap A_i$ est une partie non vide de l'ensemble bien ordonné $(A_i, \leqslant_i)$, qui admet donc un plus petit élément que l'on notera m. Prouvons que c'est aussi le plus petit élément de P. On considère $x \in P$.
  - — Si $x \in A_i$, alors $m \leqslant_i x$ (par définition de m) donc $m \leqslant x$.
  - — Sinon, il existe $j \in I$ tel que $x \in A_j$. Puisque A_j n'est pas un segment initial de A_i (car il n'est pas inclus dans A_i), on en déduit que A_i est un segment initial de A_j. Donc $x \not<_j m$ (car $x \notin A_i$), et par conséquent $m \leqslant_j x$ (puisque l'ordre sur

A_j est total), donc $m \leqslant x$.

- On considère $j \in I$. Vérifions que $\leqslant_j$ est la relation induite par $\leqslant$ sur A_j. En effet, si $(x, y) \in A_j \times A_j$ et $x \leqslant_j y$, on a $x \leqslant y$ (par définition de $\leqslant$) et réciproquement, si $x \leqslant y$, il existe $i \in I$ tel que $x \leqslant_i y$, et
  - soit $\leqslant_i$ est la relation induite par $\leqslant_j$ sur A_i, et alors comme $(x, y) \in A_i \times A_i$ et $x \leqslant_i y$, on en déduit $x \leqslant_j y$;
  - soit $\leqslant_j$ est la relation induite par $\leqslant_i$ sur A_j, et alors comme $(x, y) \in A_j \times A_j$ et $x \leqslant_i y$, on en déduit $x \leqslant_j y$.
- Prouvons enfin que A_j est un segment initial de $(A, \leqslant)$. Soit $x \in A$ et $y \in A_j$ tels que $x \leqslant y$. Il existe donc $i \in I$ tel que $x \leqslant_i y$ (en particulier $x \in A_i$). Soit A_i est un segment initial de $(A_j, \leqslant_j)$, et alors $A_i \subseteq A_j$ donc $x \in A_j$, soit A_j est un segment initial de $(A_i, \leqslant_i)$, donc $x \in A_j$. On en déduit que A_j est un segment initial de A.

Théorème 5.4.2 (Corollaire)
On considère un ensemble $\mathscr{E}$ de sous-ensembles bien ordonnés d'un ensemble ordonné $(E, \leqslant)$, tel que pour tout A et B dans $\mathscr{E}$, A est un segment initial de B ou B est un segment initial de A. Alors $\bigcup_{X \in \mathscr{E}} X$ est un sous-ensemble bien ordonné de E tel que tout élément de $\mathscr{E}$ est un segment initial de $\bigcup_{X \in \mathscr{E}} X$.

Preuve
On applique le lemme aux familles $(X)_{X \in \mathscr{E}}$ et $(\leqslant_X)_{X \in \mathscr{E}}$, où $\leqslant_X$ est la relation (de bon ordre) induite par $\leqslant$ sur X. On en déduit que $\bigcup_{X \in \mathscr{E}} X$ est un sous-ensemble bien ordonné de E tel que tout élément de $\mathscr{E}$ est un segment initial de $\bigcup_{X \in \mathscr{E}} X$.

Théorème 5.4.3 (Lemme de Zorn, ou lemme de Kuratowski-Zorn)
Si E est un ensemble ordonné dans lequel tout sous-ensemble *bien ordonné* admet un *majorant* dans E, alors E admet un *élément maximal*.

Preuve
On considère un ensemble E dans lequel tout sous-ensemble bien ordonné admet un majorant. Notons d'abord que puisque $\varnothing$ est (trivialement) un sous-ensemble bien ordonné de E, $\varnothing$ doit avoir un majorant dans E, ce qui implique que E n'est pas vide. Je désigne dans la suite par le terme *bonne chaîne* tout sous-ensemble bien ordonné de E (un tel sous-ensemble est en particulier une chaîne, autrement dit un sous-ensemble totalement ordonné par l'ordre induit), et je note $\mathscr{C}$ l'ensemble de toutes les bonnes chaînes. Pour prouver que E a un élément maximal, on raisonne par l'absurde en faisant l'hypothèse contraire qu'il n'existe aucun élément maximal, avec l'objectif d'obtenir une contradiction.

- Toute bonne chaîne admet alors un majorant strict, puisque tout majorant m d'une bonne chaîne C n'étant pas un élément maximal (on a fait l'hypothèse que E n'a pas d'élément maximal), il existe dans E un élément $m' > m$ qui est donc un majorant strict de C. On en déduit que pour toute bonne chaîne C, l'ensemble M_C de ses majorants stricts n'est pas vide, donc d'après l'axiome du choix, on peut définir une fonction $f \in \prod_{C \in \mathscr{C}} M_C$, c'est-à-dire une fonction qui à toute bonne chaîne associe un de ses majorants stricts.
- Notons $\mathscr{A}$ l'ensemble des bonnes chaînes C telles que

$$\forall x \in C, x = f(]-\infty, x[_C)$$

On peut remarquer que l'ensemble vide est une bonne chaîne vérifiant (trivialement) cette formule, donc $\varnothing \in \mathscr{A}$. Par ailleurs, si C est un élément de $\mathscr{A}$, alors $C \cup \{f(C)\}$ est une bonne chaîne, car si X est une partie non vide de $C \cup \{f(C)\}$, soit $X = \{f(C)\}$ (et alors $f(C)$ est le plus petit élément de X), soit $X \setminus \{f(C)\}$ est une partie non vide de C dont le plus petit élément est aussi le plus petit élément de X. De plus par définition de f

$$C =]-\infty, f(C)[_{C \cup \{f(C)\}}$$

donc

$$f(C) = f\left(]-\infty, f(C)[_{C \cup \{f(C)\}}\right)$$

Ces pages ne sont pas incluses dans l’aperçu.

avec $\mathbb{R}$ ensemble non dénombrable et $\mathbb{Q}$ ensemble dénombrable.

5.6 Compléments sur l'arithmétique des cardinaux

Prérequis
L'arithmétique des cardinaux (section 4.11).

Théorème 5.6.1
Tout cardinal transfini κ est tel que

$$\kappa \cdot \kappa = \kappa$$

ce qui revient à dire que si A est un ensemble infini, alors

$$|A \times A| = |A|$$

autrement dit tout ensemble infini A est en bijection avec $A \times A$.

Preuve (faisant appel au lemme de Zorn)
On considère un ensemble A infini. Prouvons

$$|A \times A| = |A|$$

1. On considère l'ensemble des bijections de $B \times B$ dans B, où B est un sous-ensemble infini de A :

$$F :\equiv \left\{ f \mid \exists B \subseteq A, |B| \geqslant \aleph_0 \text{ et } f \in \mathrm{Bij}(B \times B, B) \right\}$$

Puisque A est infini, il a un sous-ensemble ∞-dénombrable B, qui est tel que $B \times B \simeq B$, donc $F \neq \varnothing$. Nous allons appliquer le lemme de Zorn à F. Considérons pour cela une $\subseteq$-chaîne non vide C de F, et prouvons que $\bigcup_{f \in C} f \in F$. Puisque C est une chaîne, c'est un ensemble compatible de relations fonctionnelles, donc $\bigcup_{f \in C} f$ est une relation fonctionnelle injective, que l'on notera φ. Notons aussi B son image :

$$B :\equiv \mathrm{Im}\,\varphi = \bigcup_{f \in C} \mathrm{Im}(f)$$

Puisque les images des fonctions de C sont des ensembles infinis, on en déduit que B est infini. Nous allons vérifier que φ est une bijection de $B \times B$ dans B, ce qui prouvera que $\varphi \in F$. Il reste à démontrer que le domaine de φ est $B \times B$.

- Soit $(x, y) \in B \times B$. Il existe f et g dans C telles que $x \in \mathrm{Im}(f)$ et $y \in \mathrm{Im}(g)$. De plus, $f \subseteq g$ ou $g \subseteq f$. Supposons par exemple $f \subseteq g$ (le raisonnement est semblable dans l'autre cas). Alors

$$(x, y) \in \mathrm{Im}(g) \times \mathrm{Im}(g) = \mathrm{dom}(g) \subseteq \mathrm{dom}(\varphi)$$

- Réciproquement, si $z \in \mathrm{dom}(\varphi)$, il existe $f \in C$ telle que

$$z \in \mathrm{dom}(f) = \mathrm{Im}(f) \times \mathrm{Im}(f) \subseteq B \times B$$

Par conséquent φ est un élément de F. Donc, d'après le lemme de Zorn, F admet un élément maximal f.

2. Il existe donc un sous-ensemble infini B de A tel que $B \times B \xrightarrow{f} B$ soit une bijection, maximale dans F. Notons κ le cardinal de B. Par définition, κ est un cardinal transfini tel que

$$\kappa \cdot \kappa = \kappa$$

Nous allons justifier que $|A \setminus B| < \kappa$. On fait l'hypothèse contraire

$$\kappa \leqslant |A \setminus B|$$

Ces pages ne sont pas incluses dans l’aperçu.

Chapitre 6

Théorie des ensembles de Zermelo-Fraenkel (ZF), 6e partie : autres axiomes

6.1 Introduction

Les axiomes précédents sont suffisants pour construire la presque totalité des notions mathématiques actuelles, mais il reste encore deux axiomes, dont l'intérêt est essentiellement lié au développement de la théorie des ensembles elle-même : le schéma d'axiomes de remplacement, et l'axiome de fondation.

6.2 Schéma d'axiomes de remplacement

Pour toute relation fonctionnelle f (qui peut être définie sur un produit de classes) et toute classe $\mathscr{A}$, nous avons défini respectivement l'image directe et l'image réciproque de $\mathscr{A}$ par f comme

$$\underset{\rightarrow}{f}(\mathscr{A}) = \{y \in \mathrm{Im}(f) \mid \exists x \in \mathscr{A}, (x,y) \in f\}$$
$$\underset{\leftarrow}{f}(\mathscr{A}) = \{x \in \mathrm{dom}(f) \mid \exists y \in \mathscr{A}, (x,y) \in f\}$$

Donc d'après le schéma de compréhension, si l'image de f est incluse dans un ensemble, alors l'image directe de $\mathscr{A}$ est un ensemble, et si le domaine de f est inclus dans un ensemble, alors l'image réciproque de $\mathscr{A}$ est un ensemble (même si $\mathscr{A}$ est une classe propre). En particulier si la fonction $A \xrightarrow{f} B$ est définie sur les ensembles A et B, alors l'image directe et l'image réciproque d'une classe par f est un ensemble.

Mais si on ne dispose pas d'information sur l'image de f, rien ne permet de conclure que l'image directe d'un ensemble par une relation fonctionnelle est un ensemble, alors que cela peut sembler intuitivement vrai puisque cela revient à dire, d'une certaine manière, que si on remplace chaque élément d'un ensemble on obtient encore un ensemble. Il faut pour cela ajouter un nouveau schéma d'axiomes : le schéma d'axiomes de remplacement.

Axiome 6.2.1 (Schéma de remplacement, ou schéma de substitution)
Les formulations suivantes de cet axiome sont équivalentes :

1. L'image directe d'un ensemble par une relation fonctionnelle est un ensemble : pour toute formule

Ces pages ne sont pas incluses dans l’aperçu.

6.5 Axiome de fondation

Prérequis
Les relations bien fondées (sections 1.13, 3.6 et 5.2).

Nous allons voir un dernier axiome, dit de fondation, qui va limiter les ensembles possibles uniquement à ce que l'on appelle des ensembles bien fondés, une classe d'ensembles suffisants pour construire toutes les mathématiques usuelles. L'idée générale (informelle) est que l'on souhaite que les ensembles soient organisés par « étages », de telle sorte que les ensembles appartenant à un étage donné ne soient construits qu'à partir de ceux apparaissant dans les étages inférieurs, et que ces étages soient inclus les uns dans les autres, pour qu'un ensemble appartenant à un étage appartienne aussi automatiquement aux étages supérieurs. Par exemple, si l'ensemble A apparaît pour la première fois à un étage donné, l'ensemble $\mathscr{P}(A)$ ne peut être construit qu'à partir de l'étage suivant.

Justifions de manière informelle que la volonté d'avoir un tel système d'ensembles organisés en « étages » impose une condition, que l'on prendra comme axiome. On considère un ensemble E non vide, appartenant à un étage donné, et dont les éléments doivent donc appartenir à des étages inférieurs. Notons p le plus petit des étages où apparaissent pour la première fois les éléments de E [14], et on considère un élément x de E apparaissant pour la première fois à l'étage p. Alors $x \cap E = \varnothing$, car si $y \in E$ et $y \in x$, alors y doit apparaître à un étage strictement inférieur à p, ce qui contredit le fait que p est le plus petit des étages où apparaissent pour la première fois les éléments de E. En résumé, si E est un ensemble non vide, il existe $x \in E$ tel que $x \cap E = \varnothing$, c'est-à-dire tel que pour tout $y \in E$, $y \notin x$. Donc dans tout ensemble non vide il existe un élément minimal pour la relation $\in$, autrement dit $\in$ est une relation bien fondée. C'est cette propriété que l'on appelle axiome de fondation :

Axiome 6.5.1 (Axiome de fondation, ou axiome de régularité)
La relation d'appartenance $\in$ est une relation bien fondée sur la classe de tous les ensembles, ce qui signifie que tout ensemble non vide possède au moins un élément minimal pour $\in$: pour tout E

$$E \neq \varnothing \implies \exists x \in E, \forall y \in E, y \notin x$$

autrement dit

$$E \neq \varnothing \implies \exists x \in E, x \cap E = \varnothing$$

Remarque 6.5.2 : L'axiome de fondation équivaut aussi à : pour tout ensemble E, la restriction de $\in$ à E est une relation bien fondée. En effet :

- Si la relation $\in$ est bien fondée sur la classe de tous les ensembles, a fortiori elle est bien fondée sur la restriction à n'importe quel ensemble.
- Réciproquement, si pour tout ensemble E, la restriction de $\in$ à E est une relation bien fondée, cela signifie que toute partie non vide de E admet un élément $\in$-minimal. En particulier, si E est non vide, alors E admet un élément $\in$-minimal. Donc $\in$ est une relation bien fondée sur la classe de tous les ensembles.

Remarque 6.5.3 : Il est possible d'obtenir une classe d'ensembles organisés en « étages » (voir la remarque précédant l'axiome) en construisant progressivement des ensembles, à partir de l'ensemble vide, par application successive (éventuellement infinie) de la réunion et de l'ensemble des parties. On obtient ainsi une

14. Je considère pour simplifier que les étages sont indexés par des entiers naturels, ce qui ne sera pas formellement le cas (ils sont indexés par des ordinaux), mais le principe général reste valable.

Ces pages ne sont pas incluses dans l’aperçu.

Liste des figures

Liste des tableaux

Liste des symboles

$=$	Égalité logique entre deux objets identiques, page 5
$\overset{\text{def}}{=}$	Égalité par définition, page 5
$:\equiv$	Égalité par affectation, page 5
$\neg$	Connecteur logique pour la négation
« et », « $\wedge$ »	Connecteur logique pour la conjonction (*et*), page 5
« ou », « $\vee$ »	Connecteur logique pour la disjonction (*ou*), page 5
$\Longrightarrow$	Connecteur logique pour l'implication
$\Longleftrightarrow$	Connecteur logique pour l'équivalence
$\oplus$	Connecteur logique ou exclusif (OUX, ou XOR), négation du connecteur logique équivalence
$\equiv$	Équivalence logique (sémantique ou syntaxique), page 5
$\mathscr{M} \vDash \Gamma$	La L-structure $\mathscr{M}$ est un modèle de Γ
$\Gamma \vDash \mathscr{F}$	$\mathscr{F}$ est une conséquence sémantique de Γ
$\Gamma \vdash \mathscr{F}$	$\mathscr{F}$ est une conséquence syntaxique de Γ, Γ prouve $\mathscr{F}$
$\mathscr{F}(t/x)$	Formule $\mathscr{F}$ dans laquelle le terme t remplace la variable x, page 5
$\mathscr{F}[x_1, \ldots, x_n]$	Formule $\mathscr{F}$ dont les variables libres sont à prendre parmi $x_1, \ldots, x_n$, page 5
$\vec{x}$	Liste de variables $x_1, \ldots, x_n$, pour un entier n indéterminé
$A \times B$	Produit cartésien des ensembles A et B, page 49
$E/\sim$	Ensemble quotient de E par la relation d'équivalence $\sim$, page 178
$\in$	Prédicat binaire d'appartenance, page 7
$\mathscr{P}(E)$	Ensemble des sous-ensembles de l'ensemble E, page 41
$\varnothing$	Ensemble vide, page 14
$\subseteq$	Symbole d'inclusion entre ensembles, page 11
$\subset$	Symbole d'inclusion stricte entre ensembles, page 11
$\forall$	Quantificateur universel : quel que soit
$\forall x \in E, \mathscr{F}$	Tous les éléments x de E sont tels que $\mathscr{F}$, page 9
$\exists$	Quantificateur existentiel : il existe

$\exists x \in E, \mathscr{F}$ Il existe un élément x de E tel que $\mathscr{F}$, page 9

$\exists! x$ Il existe un unique x tel que …

$\exists! x \in E, \mathscr{F}$ Il existe un unique élément x de E tel que $\mathscr{F}$, page 10

$\bigcup \mathscr{E}, \bigcup_{A \in \mathscr{E}} A$ Réunion de l'ensemble $\mathscr{E}$, page 23

$A \cup B$ Réunion des ensembles A et B, page 26

$\bigcap \mathscr{E}, \bigcap_{A \in \mathscr{E}} A$ Intersection de l'ensemble $\mathscr{E}$, page 30

$A \cap B$ Intersection des ensembles A et B, page 31

$A \setminus B$ Différence des ensembles A et B, page 34

$A \sqcup B$ Somme (ou union) disjointe des ensembles A et B, page 159

$\complement_E A$ Complémentaire de l'ensemble A dans E, page 35

$A \Delta B$ Différence symétrique des ensembles A et B, page 38

$[a, b]_E, [a, b[_E, \ldots$ Intervalles de l'ensemble ordonné E, page 134

ω Ensemble des entiers naturels, page 312

$\mathbb{N}$ Ensemble des entiers naturels, page 325

$\mathbb{N}^*$ Ensemble des entiers naturels différents de 0, page 325

$\mathbb{Z}$ Ensemble des entiers relatifs, page 406

$\mathbb{B}$ Ensemble $\{0, 1\}$, page 83

$|A|$ Cardinal de l'ensemble A, page 210

$\aleph_0$ Cardinal de $\mathbb{N}$ (plus petit cardinal infini), page 396

$\mathscr{Card}$ Classe des cardinaux, page 210

$\mathscr{U}$ Classe de tous les ensembles, page 57

id_A Fonction identité de l'ensemble A ($x \mapsto x$), page 82

$\langle f_1, \ldots, f_n \rangle$ Fonction $x \mapsto (f_1(x), \ldots, f_n(x))$, page 85

χ_A Fonction indicatrice de l'ensemble A, page 83

$f : A \longrightarrow B$, $A \xrightarrow{f} B$ f est une fonction de A dans B, page 67

$f \circ g$ Composée de la fonction g par la fonction f : $f \circ g(x) = f(g(x))$, page 91

$f_1 \times \cdots \times f_n$ Fonction $(x_1, \ldots, x_n) \mapsto (f(x_1), \ldots, f(x_n))$, page 85

$A \longrightarrow B, B^A$ Ensemble des fonctions de A dans B, page 80

$\mathrm{Inj}(A, B)$ Ensemble des injections de A dans B, page 106

$\mathrm{Surj}(A, B)$ Ensemble des surjections de A dans B, page 106

$\mathrm{Bij}(A, B)$ Ensemble des bijections de A dans B, page 106

$\mathscr{S}_E$ Ensemble des permutations de E (les bijections de E dans E), page 106

$A^{(E)}$ Ensemble des fonctions de E dans A à support fini, page 370

$A \simeq B$	Les ensembles A et B sont en bijection, page 107
$A \simeq B, (A, \dots) \simeq (B, \dots)$	Les structures $(A, \dots)$ et $(B, \dots)$ sont isomorphes, page 279
$(A_i)_{i \in I}$	Famille indexée par I (fonction $i \mapsto A_i$), page 86
$\bigcup_{i \in I} A_i$	Réunion de la famille $(A_i)_{i \in I}$, page 87
$\bigcap_{i \in I} A_i$	Intersection de la famille $(A_i)_{i \in I}$, page 87
$\prod_{i \in I} A_i$	Produit de la famille $(A_i)_{i \in I}$, page 90
$\mathrm{dom}(\mathscr{R})$	Domaine de la relation $\mathscr{R}$ (classe des premières composantes des éléments de $\mathscr{R}$), page 63
$\mathrm{cod}(\mathscr{R})$	Codomaine (classe d'arrivée) de la relation $\mathscr{R}$, page 59
$\mathrm{Im}(\mathscr{R})$	Image de la relation $\mathscr{R}$ (classe des deuxièmes composantes des éléments de $\mathscr{R}$), page 63
$\mathrm{dom}(f)$	Domaine de la fonction f, page 70
$\mathrm{cod}(f)$	Codomaine (ensemble d'arrivée) de la fonction f, page 70
$\mathrm{Im}(f)$	Image de la fonction f, page 70
$\underset{\rightarrow}{f}(A)$	Image directe de l'ensemble A par la fonction f, page 74
$\underset{\leftarrow}{f}(A)$	Image réciproque de l'ensemble A par la fonction f, page 74
$\mathrm{supp}(f)$	Support de la fonction f, page 370
$\sum_{k=p}^{n} a_k$	Somme : $a_p + a_{p+1} + \cdots + a_{n-1} + a_n$, page 430
$\prod_{k=p}^{n} a_k$	Produit : $a_p \times a_{p+1} \times \cdots \times a_{n-1} \times a_n$, page 430
$H \leqslant G$	H est un sous-groupe de G, page 239
$\mathrm{SGr}(G)$	Ensemble des sous-groupes de G, page 239
$Z(G)$	Centre du groupe G (éléments commutant avec tous les autres), page 245
$\mathrm{Aut}(G)$	Groupe des automorphismes du groupe G, page 294
$\mathrm{Int}(G)$	Groupe des automorphismes intérieurs du groupe G, page 299
$\mathrm{Ker}(f)$	Noyau du morphisme de groupes f (ensemble des éléments dont l'image est l'élément neutre), page 296
$\mathrm{quo}(n, p)$	Quotient de la division euclidienne de n par p, page 405
$\mathrm{res}(n, p)$	Reste de la division euclidienne de n par p, page 405

Index des notions

Index des noms propres