# Éléments de mathématiques pour le XXI$^{e}$ siècle, volume 3

## 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)*. 2019.
  - *Éléments de mathématiques pour le XXI^e siècle, volume 2 : Fondements des mathématiques 2 (théorie des ensembles, mathématiques discrètes, structures algébriques de base)*. 2019.
  - *Citations mathématiques : Plus de 200 citations sourcées et vérifiées*. 2021.
- À paraître : *Éléments de mathématiques pour le XXI^e siècle, volume 4 : Fondements des mathématiques 4*.

Étienne Bonheur

# Éléments de mathématiques pour le XXI$^{e}$ siècle, volume 3

Fondements des mathématiques 3
(théorie des ensembles, théorie des nombres, algèbre, théorie des modèles, théorie de la calculabilité, théorie des catégories et des topos)

Paysages Mathématiques

https://www.paysmaths.net

ISBN : 978-2-9569666-3-0
Dépôt légal : février 2022

# Table des matières

# Introduction

Ce livre est le troisième 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 mon lectorat 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 mon lectorat francophone 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 Uses of Some Words of Mathematics*. URL : `https://mathshistory.st-andrews.ac.uk/Miller/mathword/`
- Jeff MILLER. *Earliest Uses of Various Mathematical Symbols*. URL : `https://mathshistory.st-andrews.ac.uk/Miller/mathsym/`
- 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 précise non plus à chaque fois.

On notera aussi qu'aucun paragraphe ne commence par « exercice », ce qui ne veut pas dire que les lecteurs ne disposent 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 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… ».

Quelques algorithmes, implémentés en langage *python*, sont inclus dans cet ouvrage. Non seulement ce langage a un certain nombre d'avantages intrinsèques (simple, libre et gratuit, multiplateforme,…), mais il est aussi privilégié dans l'enseignement en France (que ce soit dans le secondaire ou le supérieur). Je précise néanmoins que ces scripts python sont de simples illustrations permettant de faire fonctionner les algorithmes dans des cas simples, et en aucun cas ne sont des exemples de code optimisé pour tel ou tel usage. En particulier, je ne tiens pas compte de contraintes purement informatiques (temps de calcul, mémoire utilisée,…), je ne vérifie pas la cohérence des données en entrée (la vérification doit se faire en amont), et je ne ferai appel qu'à des types de données basiques (et sans définir de nouvelles classes) : entiers, chaînes de caractères, booléens, et listes.

Les quatre premiers volumes de cette série 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 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 traite essentiellement de la notion de logique mathématique, et le deuxième de la théorie des ensembles de Zermelo-Fraenkel. Ce troisième volume est consacré à

- des compléments de théorie des ensembles : quelques généralités dans le chapitre 1, puis étude de la notion d'ordinal (qui étend le principe permettant d'*ordonner* un ensemble fini) et de celle de cardinal (qui étend le principe permettant de *dénombrer* un ensemble fini), dans les chapitres 5 et 6 (les cardinaux ont déjà été introduits de façon axiomatique dans le volume 2, ils sont ici construits, de façon classique, comme des ordinaux particuliers) ; enfin le chapitre 7 donne deux exemples de théories alternatives des ensembles (autres que la théorie standard de Zermelo-Fraenkel exposée dans le volume 2) : les théories des classes (très semblables) de von Neumann-Bernays-Gödel (NBG) et de Morse-Kelley (MK), et la théorie *New Foundations with Urelements* [Nouveaux Fondements avec Uréléments] (ou NFU) ;
- des compléments d'algèbre et de mathématiques discrètes, qui concernent les groupes et les anneaux dans les chapitre 2 et 4 (notamment la notion d'anneau quotient, permettant la construction des anneaux $\mathbb{Z}/n\mathbb{Z}$ des entiers modulo $n$), et la théorie élémentaire des nombres (divisibilité dans $\mathbb{N}$ et $\mathbb{Z}$, pgcd et ppcm…) dans le chapitre 3 ;
- l'introduction de différentes théories mathématiques plus avancées : la théorie des modèles (chapitre 8), qui étudie les relations pouvant exister entre des théories formelles et certaines structures algébriques, la théorie de la calculabilité (chapitre 9), qui s'intéresse à la formalisation du concept d'algorithme et permet

de démontrer certains théorèmes dits *de limitation* (comme les théorèmes d'incomplétude de Gödel), et la théorie des catégories et plus précisément celle des topos (chapitres 10 et 11), qui sera le cadre dans lequel je présenterai une autre théorie alternative des ensembles, la *théorie élémentaire de la catégorie des ensembles* (ou ETCS) de William Lawvere.

Le quatrième volume sera consacré à des compléments de mathématiques discrètes (théorie des nombres, analyse combinatoire, introduction à la théorie des graphes), à des compléments sur les différentes structures algébriques, à la construction des autres ensembles classiques de nombres (ensemble $\mathbb{Q}$ des rationnels, ensemble $\mathbb{R}$ des réels,...), à la présentation de structures linéaires et topologiques, et à l'introduction de la théorie homotopique des types (qui permet un autre fondement formel alternatif des mathématiques).

# Vocabulaire, notations et rappels

Je renvoie mon lectorat au premier volume pour quelques remarques introductives concernant le vocabulaire et les notations. Je ne rappellerai ici que quelques usages qui peuvent être peu répandus, voire personnels (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), ainsi que quelques principes vus dans les volumes précédents.

1. En ce qui concerne l'égalité, qui a en mathématiques un sens parfois subtil, je fais 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 :
   $$A \stackrel{\text{def}}{\equiv} B$$
   (« 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.
2. Dans une formule, $\vec{x}$ représente une liste de variables $x_1, \dots, x_n$, pour un entier $n$ indéterminé. De plus, 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 ($\vee$), ce qui donne, appliqué à deux propositions $P$ et $Q$ :
   $$\begin{array}{lll} P \wedge Q & \text{ou} & P \text{ et } Q \\ P \vee Q & \text{ou} & P \text{ ou } Q \end{array}$$
   Le connecteur pour la négation est noté $\neg$, celui pour l'implication $\Longrightarrow$ et celui pour l'équivalence $\Longleftrightarrow$, ce qui donne pour ces trois autres connecteurs logiques classiques :
   $$\begin{array}{ll} \neg P & (\text{« non-}P\text{ »}) \\ P \Longrightarrow Q & (\text{« } P \text{ implique } Q\text{ »}) \\ P \Longleftrightarrow Q & (\text{« } P \text{ est équivalent à } Q\text{ »}) \end{array}$$
3. Le symbole $\equiv$ désigne l'équivalence (sémantique ou syntaxique) de deux formules, dans le sens suivant : si $\Gamma$ 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

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

4. Je note $\subseteq$ la relation d'inclusion et $\subset$ la relation d'inclusion stricte : pour tous les ensembles $A$ et $B$

$$A \subseteq B \stackrel{\text{def}}{\equiv} \forall x \in A, x \in B$$
$$A \subset B \stackrel{\text{def}}{\equiv} A \subseteq B \text{ et } A \neq B$$

Si $A \subseteq B$ (respectivement $A \subset B$), on dit que $A$ est un *sous-ensemble* (respectivement un *sous-ensemble strict*), ou une *partie* (respectivement une *partie stricte*), de $B$.

5. Pour tout ensemble $A$, j'appellerai *relation* d'arité $n$, ou *prédicat* d'arité $n$, toute partie $P$ de $A^n$, et je pourrai noter indifféremment

$$(x_1, \dots, x_n) \in P \qquad \text{ou} \qquad P(x_1, \dots, x_n)$$

La donnée d'une relation équivaut à celle de sa *fonction indicatrice* (ou *fonction caractéristique*)

$$\chi_P : \begin{cases} A^n \longrightarrow \{0,1\} \\ (x_1, \dots, x_n) \longmapsto \begin{cases} 1 & \text{si } (x_1, \dots, x_n) \in P \\ 0 & \text{sinon} \end{cases} \end{cases}$$

Ainsi pour tous les éléments $x_1, \dots, x_n$ de $A$, les trois expressions suivantes sont équivalentes

$$(x_1, \dots, x_n) \in P \qquad P(x_1, \dots, x_n) \qquad \chi_P(x_1, \dots, x_n) = 1$$

de même que les trois expressions suivantes :

$$(x_1, \dots, x_n) \notin P \qquad \neg P(x_1, \dots, x_n) \qquad \chi_P(x_1, \dots, x_n) = 0$$

6. Une fonction $A \xrightarrow{f} B$ est définie entièrement sur $A$ (*fonction* est synonyme de *application*, terme que je n'utilise pas). Si le domaine de ce que j'ai appelé une *relation fonctionnelle* peut être strictement inclus dans $A$, il s'agit alors d'une *fonction partielle* de $A$ dans $B$.

7. Pour toute fonction $A \xrightarrow{f} B$ je note respectivement $\underset{\rightarrow}{f}$ et $\underset{\leftarrow}{f}$ les fonctions suivantes :

$$\underset{\rightarrow}{f} : \begin{cases} \mathscr{P}(A) \longrightarrow \mathscr{P}(B) \\ X \longmapsto \{f(x) \mid x \in X\} \end{cases} \qquad \text{et} \qquad \underset{\leftarrow}{f} : \begin{cases} \mathscr{P}(B) \longrightarrow \mathscr{P}(A) \\ Y \longmapsto \{x \in A \mid f(x) \in Y\} \end{cases}$$

$\underset{\rightarrow}{f}(X)$ représente donc l'image directe de l'ensemble $X$ par la fonction $f$, autrement dit ce qui est noté plus classiquement $f(X)$, et $\underset{\leftarrow}{f}(Y)$ représente donc l'image réciproque de l'ensemble $Y$ par la fonction $f$, autrement dit ce qui est noté plus classiquement $f^{-1}(Y)$.

8. Pour tous les ensembles $A$ et $B$ je note
   - $A \longrightarrow B$ ou $B^A$ l'ensemble des fonctions de $A$ dans $B$.
   - $\mathrm{Inj}(A, B)$ l'ensemble des injections de $A$ dans $B$.
   - $\mathrm{Surj}(A, B)$ l'ensemble des surjections de $A$ dans $B$.
   - $\mathrm{Bij}(A, B)$ l'ensemble des bijections de $A$ dans $B$.
   - $\mathfrak{S}_A$ l'ensemble des permutations de $A$ (c'est-à-dire des bijections de $A$ dans $A$).
9. Je désigne les intervalles de tout ensemble ordonné $A$ par la même notation que pour les intervalles classiques de $\mathbb{R}$ :

$$\begin{aligned} [a, b]_A &\overset{\text{def}}{=} \{x \in A \mid a \leqslant x \text{ et } x \leqslant b\} \\ [a, b[_A &\overset{\text{def}}{=} \{x \in A \mid a \leqslant x \text{ et } x < b\} \\ ]-\infty, b]_A &\overset{\text{def}}{=} \{x \in A \mid x \leqslant b\} \\ &\dots \end{aligned}$$

10. Pour tout ensemble $A$, $|A|$ représente le *cardinal* de $A$. Les cardinaux ont pour l'instant été définis de façon axiomatique (leur construction formelle se fera au chapitre 6), de telle sorte que pour tous les ensembles $A$ et $B$
    - $|A| = |B|$ si et seulement si $A$ et $B$ sont en bijection ;
    - $|A| \leqslant |B|$ si et seulement si il existe une injection de $A$ dans $B$.

    Le cardinal d'un ensemble fini est son nombre d'éléments. En particulier pour tout entier naturel $n$

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

11. L'expression $A \simeq B$ signifie, selon le contexte, que les ensembles $A$ et $B$ sont en bijection, ou que l'ensemble $A$, muni d'une certaine structure, et l'ensemble $B$, muni de la même structure, sont isomorphes. Par exemple si $(G, *, e)$ et $(G', \star, e')$ sont deux groupes,

$$(G, *, e) \simeq (G', \star, e') \qquad \text{ou juste} \qquad G \simeq G'$$

    signifie que les deux groupes sont isomorphes.

# Chapitre 1

# Compléments sur les ensembles

## 1.1 Généralités sur la notion de clôture

Nous avons vu dans le volume 2 un principe général qui s'applique à différents domaines des mathématiques : on considère une partie $A$ d'un ensemble $E$, un prédicat $\mathscr{P}$ sur $\mathscr{P}(E)$ (autrement dit une propriété que peuvent avoir les parties de $E$), et l'ensemble $\mathscr{E}$ des parties de $E$ incluant $A$ et vérifiant le prédicat $\mathscr{P}$, autrement dit

$$\mathscr{E} :\equiv \{X \subseteq E \mid A \subseteq X \text{ et } \mathscr{P}(X)\}$$

Si $\mathscr{E} \neq \emptyset$, on peut définir l'intersection $\bar{A}$ de cet ensemble :

$$\bar{A} :\equiv \bigcap_{X \in \mathscr{E}} X$$

Si de plus $\bar{A}$ vérifie le prédicat $\mathscr{P}$ (ce qui est notamment le cas si $\mathscr{P}$ est stable par intersection), alors $\bar{A}$ est le plus petit élément de $(\mathscr{E}, \subseteq)$, c'est-à-dire le plus petit sous-ensemble de $E$ incluant $A$ et vérifiant le prédicat $\mathscr{P}$.

C'est ainsi qu'ont été définis :

- le sous-groupe engendré par un sous-ensemble $A$ d'un groupe $G$, avec le prédicat « être un sous-groupe » : comme l'intersection de sous-groupes est un sous-groupe, l'intersection de tous les sous-groupes de $G$ incluant $A$ est le plus petit sous-groupe de $G$ incluant $A$.
- le sous-anneau engendré par un sous-ensemble $B$ d'un anneau $A$, avec le prédicat « être un sous-anneau » : comme l'intersection de sous-anneaux est un sous-anneau, l'intersection de tous les sous-anneaux de $A$ incluant $B$ est le plus petit sous-anneau de $A$ incluant $B$.
- l'idéal engendré par un sous-ensemble $B$ d'un anneau commutatif $A$, avec le prédicat « être un idéal » : comme l'intersection d'idéaux est un idéal, l'intersection de tous les idéaux de $A$ incluant $B$ est le plus petit idéal de $A$ incluant $B$.

Nous allons voir dans la suite de ce chapitre deux autres applications de ce principe, dit de *clôture*, la *clôture de relations binaires*, qui correspond au cas de relations binaires (c'est-à-dire de parties d'un ensemble produit $E \times E$) vérifiant certaines propriétés, et la *clôture d'opérations*, qui correspond au cas de parties d'un ensemble $E$ stables par certaines opérations (d'où l'on déduit le principe d'induction structurelle, que nous avons vu de façon informelle dans le volume 1).

## 1.2 Clôture de relations binaires

**Prérequis**
Les relations binaires (notamment les sections 1.5, 1.11, et 2.2 du volume 2).

**Théorème 1.2.1 (Intersection de relations binaires)**
Si $(\mathscr{R}_i)_{i\in I}$ est une famille non vide de relations binaires sur un ensemble $E$, alors l'intersection des relations

$$\mathscr{R} :\equiv \bigcap_{i\in I} \mathscr{R}_i$$

est une relation binaire sur $E$, telle que pour tout $x$ et $y$ dans $E$

$$x\mathscr{R}y \iff \forall i \in I, x\mathscr{R}_i y$$

De plus la relation $\mathscr{R}$ est plus forte que toutes les relations $\mathscr{R}_i$ (autrement dit $\mathscr{R}$ est incluse dans toutes les $\mathscr{R}_i$) :

$$\forall i \in I, \forall (x,y) \in E\times E, (x\mathscr{R}y \implies x\mathscr{R}_i y)$$

et $\mathscr{R}$ est la plus faible de toutes les relations sur $E$ plus fortes que toutes les $\mathscr{R}_i$ (autrement dit toute relation sur $E$ incluse dans toutes les $\mathscr{R}_i$ est incluse dans $\mathscr{R}$) : si $\mathscr{R}'$ est une relation binaire sur $E$ telle que

$$\forall i \in I, \forall (x,y) \in E\times E, (x\mathscr{R}'y \implies x\mathscr{R}_i y)$$

alors

$$\forall (x,y) \in E\times E, (x\mathscr{R}'y \implies x\mathscr{R}y)$$

**Preuve**
Il s'agit juste d'expliciter les propriétés de l'intersection d'une famille d'ensembles. Puisque tous les $\mathscr{R}_i$ sont des sous-ensembles de $E\times E$, leur intersection est aussi un sous-ensemble de $E\times E$, c'est-à-dire une relation binaire. Par définition de l'intersection

$$(x,y) \in \mathscr{R} \iff \forall i \in I, (x,y) \in \mathscr{R}_i$$

autrement dit

$$x\mathscr{R}y \iff \forall i \in I, x\mathscr{R}_i y$$

Par ailleurs, la relation $\mathscr{R}$ est la borne inférieure de $\{\mathscr{R}_i\}_{i\in I}$ dans $\mathscr{P}(E\times E)$ (propriété de la borne inférieure d'une famille), donc $\mathscr{R}$ est incluse dans tous les $\mathscr{R}_i$ (autrement dit $\mathscr{R}$ est plus forte que toutes les relations $\mathscr{R}_i$), et si $\mathscr{R}'$ est incluse dans toutes les $\mathscr{R}_i$, alors $\mathscr{R}' \subseteq \mathscr{R}$ (autrement dit si $\mathscr{R}'$ est plus forte que tous les $\mathscr{R}_i$, alors $\mathscr{R}'$ est plus forte que $\mathscr{R}$).

**Exemple 1.2.2 (Exemples d'intersections de relations binaires)**

1. Dans $\mathbb{Z}$, l'intersection des relations de congruence modulo $n$ (pour $n \in \mathbb{N}^*$) est la relation d'égalité, car le seul entier divisible par tous les entiers naturels non nuls est 0, donc

$$x\mathscr{R}y \quad\equiv\quad \forall n \in \mathbb{N}^*, n \mid (x-y) \quad\equiv\quad x-y=0 \quad\equiv\quad x=y$$

2. Dans $\mathbb{Z}$, l'intersection des relations de congruence modulo $n$, pour $n \in \{n_1,\dots,n_p\}$, est la relation de congruence modulo $\text{ppcm}(n_1,\dots,n_p)$. En effet, les lignes suivantes sont équivalentes :

$$x\mathscr{R}y$$

$$\forall k \in [1,p], x \equiv y \mod n_k$$
$$\forall k \in [1,p], n_k \mid (x-y)$$
$$\text{ppcm}(n_1,\dots,n_p) \mid (x-y)$$
$$x \equiv y \mod \text{ppcm}(n_1,\dots,n_p)$$

**Théorème 1.2.3 (Conservation de propriétés de relations binaires par intersection)**

1. L'intersection de relations binaires réflexives (respectivement antiréflexives, symétriques, antisymétriques, asymétriques, transitives) est réflexive (respectivement antiréflexive, symétrique, antisymétrique, asymétrique, transitive).
2. L'intersection de relations d'ordre (respectivement de préordre, d'ordre strict, d'équivalence) est une relation d'ordre (respectivement de préordre, d'ordre strict, d'équivalence).

**Preuve**

1.
   - Si $(\mathscr{R}_i)_{i\in I}$ est une famille de relations binaires réflexives sur $E$, alors pour tout $x \in E$ et pour tout $i \in I$, $(x,x) \in \mathscr{R}_i$, donc
   $$(x,x) \in \bigcap_{i\in I} \mathscr{R}_i = \mathscr{R}$$
   ce qui signifie par définition que $\mathscr{R}$ est réflexive.
   - Si $(\mathscr{R}_i)_{i\in I}$ est une famille de relations binaires antiréflexives, alors pour tout $x \in E$, on a $(x,x) \notin \mathscr{R}$ car sinon $(x,x)$ serait un élément de tous les $\mathscr{R}_i$ (on peut noter qu'il suffit qu'une des relations $\mathscr{R}_i$ soit antiréflexive pour que $\mathscr{R}$ le soit).
   - Enfin, faisons l'hypothèse que $(\mathscr{R}_i)_{i\in I}$ est une famille de relations binaires symétriques (le raisonnement est semblable pour des relations antisymétriques, asymétriques, transitives). Pour tout $(x,y) \in E \times E$, si $(x,y) \in \mathscr{R}$ alors pour tout $i \in I$, $(x,y) \in \mathscr{R}_i$ donc $(y,x) \in \mathscr{R}_i$, et par conséquent $(y,x) \in \mathscr{R}$.
2. Comme l'intersection conserve les propriétés caractérisant les relations d'ordre (respectivement de préordre, d'ordre strict, d'équivalence), on en déduit que l'intersection de relations d'ordre (respectivement de préordre, d'ordre strict, d'équivalence) est une relation d'ordre (respectivement de préordre, d'ordre strict, d'équivalence).

**Remarque 1.2.4 :** Si $(\mathscr{R}_i)_{i\in I}$ est une famille de relations binaires totales, leur intersection ne l'est pas nécessairement. Par exemple, sur $\{0,1\}$, les relations binaires

$$\{(0,0),(1,1),(0,1)\} \qquad \text{et} \qquad \{(0,0),(1,1),(1,0)\}$$

sont totales, mais leur intersection

$$\mathscr{R} :\equiv \{(0,0),(1,1)\}$$

ne l'est pas, car $(0,1) \notin \mathscr{R}$ et $(1,0) \notin \mathscr{R}$.

**Définition 1.2.5 (Clôture transitive, clôture réflexive transitive, relation d'équivalence engendrée)**
On considère une relation binaire $\mathscr{R}$ sur un ensemble $E$.

1. On appelle *clôture transitive* de $\mathscr{R}$ la plus petite (au sens de l'inclusion) des relations transitives incluant $\mathscr{R}$. C'est l'ensemble $\bar{\mathscr{R}}$ que l'on peut définir des deux façons équivalentes suivantes :
   - Définition de haut en bas : $\bar{\mathscr{R}}$ est l'intersection de toutes les relations transitives incluant $\mathscr{R}$.

- Définition de bas en haut :

$$\bar{\mathscr{R}} = \bigcup_{n\in\mathbb{N}} \mathscr{R}_n$$

où $(\mathscr{R}_n)_{n\in\mathbb{N}}$ est définie par récurrence :

$$\begin{cases} \mathscr{R}_0 = \mathscr{R} \\ \mathscr{R}_{n+1} = \left\{(x,y) \in E\times E \mid \exists z \in E, x\mathscr{R}z \text{ et } z\mathscr{R}_n y\right\} \end{cases}$$

Ce qui équivaut aussi à

$$x\bar{\mathscr{R}}y \iff \exists n\in\mathbb{N}, \exists (x_0,\dots,x_{n+1}) \in E^{n+2} \begin{cases} x_0 = x \\ x_{n+1} = y \\ \forall i \in [0,n], x_i \mathscr{R} x_{i+1} \end{cases}$$

2. On appelle *clôture réflexive transitive* de $\mathscr{R}$ la plus petite (au sens de l'inclusion) des relations réflexives et transitives incluant $\mathscr{R}$. C'est l'ensemble $\bar{\mathscr{R}}$ que l'on peut définir des deux façons équivalentes suivantes :
   - Définition de haut en bas : $\bar{\mathscr{R}}$ est l'intersection de toutes les relations réflexives et transitives incluant $\mathscr{R}$.
   - Définition de bas en haut :

$$\bar{\mathscr{R}} = \left(\bigcup_{n\in\mathbb{N}} \mathscr{R}_n\right) \cup \{(x,x) \mid x\in E\}$$

   (où $(\mathscr{R}_n)_{n\in\mathbb{N}}$ est définie comme précédemment), ce qui équivaut donc aussi à

$$x\bar{\mathscr{R}}y \iff (x=y) \text{ ou } \left( \exists n\in\mathbb{N}, \exists (x_0,\dots,x_{n+1}) \in E^{n+2} \begin{cases} x_0 = x \\ x_{n+1} = y \\ \forall i \in [0,n], x_i \mathscr{R} x_{i+1} \end{cases} \right)$$

   ou encore à

$$x\bar{\mathscr{R}}y \iff \exists n\in\mathbb{N}, \exists (x_0,\dots,x_n) \in E^{n+1} \begin{cases} x_0 = x \\ x_n = y \\ \forall i \in [0,n[, x_i \mathscr{R} x_{i+1} \end{cases}$$

3. On appelle *relation d'équivalence engendrée par* $\mathscr{R}$ la plus petite (au sens de l'inclusion) des relations d'équivalence incluant $\mathscr{R}$. C'est l'intersection de toutes les relations d'équivalence incluant $\mathscr{R}$.

**Preuve**

1. Définitions de haut en bas : toute relation $\mathscr{R}$ sur $E$ est incluse dans la relation $E\times E$ (l'ensemble de tous les couples de $E\times E$ est un cas particulier de relation binaire), qui est transitive, réflexive, symétrique (donc c'est aussi une relation d'équivalence), ce qui montre que l'ensemble des relations transitives (respectivement réflexives et transitives, d'équivalence) incluant $\mathscr{R}$ n'est pas vide, et permet de définir son intersection $\bar{\mathscr{R}}$. D'après le théorème précédent, $\bar{\mathscr{R}}$ est une relation transitive (respectivement réflexive et transitive, d'équivalence) et par conséquent c'est la plus petite des relations transitives (respectivement réflexives et transitives, d'équivalence) incluant $\mathscr{R}$ (principe général de définition par clôture).

**Ces pages ne sont pas incluses dans l'aperçu.**

**Le volume 3 des**
***Éléments de mathématiques pour le XXI^e siècle***
**(ISBN : 978-2-9569666-3-0) est disponible**
**en version papier et numérique.**
**Détails sur le site *Paysages Mathématiques* :**
**https://www.paysmaths.net/boutique**

**Définition 1.6.10 (Ordre lexicographique et antilexicographique sur des ensembles de fonctions)**
On considère un ensemble ordonné $A$ et un ensemble totalement ordonné $B$.

1. On appelle ordre *lexicographique strict* sur $B \longrightarrow A$, la relation d'ordre strict $<_{\mathscr{L}}$ suivante :

$$f <_{\mathscr{L}} g \overset{\text{def}}{\equiv} \exists b \in B, \begin{cases} f(b) < g(b) \\ \forall x < b, f(x) = g(x) \end{cases}$$

2. On appelle ordre *antilexicographique strict* sur $B \longrightarrow A$, la relation d'ordre strict $<_{\mathscr{A}}$ suivante :

$$f <_{\mathscr{A}} g \overset{\text{def}}{\equiv} \exists b \in B, \begin{cases} f(b) < g(b) \\ \forall x > b, f(x) = g(x) \end{cases}$$

**Remarque 1.6.11 :** Dans le cas lexicographique, l'élément $b$ est tel que

$$b = \min\{x \in B \mid f(x) \neq g(x)\}$$

et dans le cas antilexicographique, il est tel que

$$b = \max\{x \in B \mid f(x) \neq g(x)\}$$

Même si les $A_i$ sont des ensembles totalement ordonnés, les ordres lexicographique et antilexicographique définis sur $\prod_{i \in I} A_i$ ne sont pas nécessairement totaux. En effet, on déduit de la définition que deux éléments distincts $(x_i)_{i \in I}$ et $(y_i)_{i \in I}$ sont comparables pour l'ordre lexicographique (respectivement antilexicographique) si et seulement si $\{i \in I \mid x_i \neq y_i\}$ admet un plus petit élément (respectivement un plus grand élément) $m$. L'ordre sur les $A_i$ étant total, on a alors $x_m < y_m$ ou $y_m < x_m$.

Par exemple, si on se place dans l'ensemble $\mathbb{Z}_- \longrightarrow \{0,1\}$, alors pour l'ordre lexicographique les fonctions

$$f : \begin{cases} \mathbb{Z}_- \longrightarrow \{0,1\} \\ n \longmapsto \begin{cases} 0 & \text{si } n \text{ est pair} \\ 1 & \text{si } n \text{ est impair} \end{cases} \end{cases} \quad \text{et} \quad g : \begin{cases} \mathbb{Z}_- \longrightarrow \{0,1\} \\ n \longmapsto \begin{cases} 1 & \text{si } n \text{ est pair} \\ 0 & \text{si } n \text{ est impair} \end{cases} \end{cases}$$

ne sont pas comparables (car l'ensemble $\{x \in \mathbb{Z}_- \mid f(x) \neq g(x)\}$ n'a pas de plus petit élément), ce que l'on peut aussi visualiser ainsi :

$$\begin{aligned} f &: \quad \ldots 1,0,1,0,1,0 \\ g &: \quad \ldots 0,1,0,1,0,1 \end{aligned}$$

Si on se place dans l'ensemble $\mathbb{N} \longrightarrow \{0,1\}$, alors pour l'ordre antilexicographique les fonctions

$$f : \begin{cases} \mathbb{N} \longrightarrow \{0,1\} \\ n \longmapsto \begin{cases} 0 & \text{si } n \text{ est pair} \\ 1 & \text{si } n \text{ est impair} \end{cases} \end{cases} \quad \text{et} \quad g : \begin{cases} \mathbb{N} \longrightarrow \{0,1\} \\ n \longmapsto \begin{cases} 1 & \text{si } n \text{ est pair} \\ 0 & \text{si } n \text{ est impair} \end{cases} \end{cases}$$

ne sont pas comparables (car l'ensemble $\{x \in \mathbb{N} \mid f(x) \neq g(x)\}$ n'a pas de plus grand élément), ce que l'on peut aussi visualiser ainsi :

$$f : \quad 0,1,0,1,0,1,\ldots$$

$$g : \quad 1, 0, 1, 0, 1, 0, \ldots$$

Mais il suffit d'ajouter quelques conditions simples pour obtenir un ordre total. Une première façon de faire est de munir le domaine d'un bon ordre :

**Théorème 1.6.12 (Ordre (anti)lexicographique total sur un produit)**
On considère un ensemble totalement ordonné $I$ et une famille d'ensembles totalement ordonnés $(A_i)_{i\in I}$.

1. Si toute partie non vide de $I$ admet un plus petit élément (autrement dit si $(I, \leqslant_I)$ est bien ordonné), alors l'ordre lexicographique sur $\prod_{i\in I} A_i$ est total.
2. Si toute partie non vide de $I$ admet un plus grand élément (autrement dit si $(I, \geqslant_I)$ est bien ordonné), alors l'ordre antilexicographique sur $\prod_{i\in I} A_i$ est total.

**Preuve**

1. Pour toutes familles distinctes $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ de $\prod_{i\in I} A_i$, l'ensemble $\{i \in I \mid x_i \neq y_i\}$ est non vide, donc admet un plus petit élément $m$ (puisque $I$ est bien ordonné). On en déduit que $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ sont comparables (si $x_m < y_m$ alors $(x_i)_{i\in I} <_{\mathscr{L}} (y_i)_{i\in I}$, et si $x_m > y_m$ alors $(x_i)_{i\in I} >_{\mathscr{L}} (y_i)_{i\in I}$), et par conséquent l'ordre lexicographique sur $\prod_{i\in I} A_i$ est bien total.
2. Si toute partie non vide de $I$ admet un plus grand élément (pour $\leqslant_I$), cela signifie que $I$ muni de l'ordre inverse $\geqslant_I$ est bien ordonné. Il suffit alors d'appliquer le résultat précédent, car la définition de l'ordre antilexicographique sur $\prod_{i\in I} A_i$ lorsque $I$ est muni de $\leqslant_I$, correspond à la définition de l'ordre lexicographique sur $\prod_{i\in I} A_i$ lorsque $I$ est muni de l'ordre inverse $\geqslant_I$.

**Théorème 1.6.13 (Corollaire 1)**
Pour tout ensemble fini totalement ordonné $I$ et toute famille d'ensembles totalement ordonnés $(A_i)_{i\in I}$, les relations $\leqslant_{\mathscr{L}}$ et $\leqslant_{\mathscr{A}}$ définies sur $\prod_{i\in I} A_i$ sont des relations d'ordre total.

**Preuve**
Il suffit d'appliquer le théorème précédent, puisque dans un ensemble fini totalement ordonné, toute partie non vide admet un plus petit et un plus grand élément.

**Théorème 1.6.14 (Corollaire 2 : ordre (anti)lexicographique total sur $B \longrightarrow A$)**
On considère deux ensembles totalement ordonnés $A$ et $B$.

1. Si toute partie non vide de $B$ admet un plus petit élément (autrement dit si $(B, \leqslant)$ est bien ordonné), alors l'ordre lexicographique sur $B \longrightarrow A$ est un ordre total.
2. Si toute partie non vide de $B$ admet un plus grand élément (autrement dit si $(B, \geqslant)$ est bien ordonné), alors l'ordre antilexicographique sur $B \longrightarrow A$ est total.

**Preuve**
Il s'agit d'un cas particulier du cas général (où la famille $(A_i)_{i\in I}$ est constante).

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 2

# Compléments sur les anneaux

## 2.1 Rappels et propriétés diverses

Je rappelle d'abord brièvement quelques points en rapport avec les anneaux : à tout anneau $(A, +, \times, 0_A, 1_A)$ correspond le groupe additif sous-jacent $(A, +, 0_A)$ et le groupe multiplicatif $(A^\times, \times, 1_A)$ des éléments inversibles. Un *anneau trivial* est un anneau n'ayant qu'un élément (ou de manière équivalente, un anneau dans lequel $0_A = 1_A$). Un *anneau ordonné* est un anneau muni d'une relation d'ordre compatible avec l'addition, et avec la multiplication par un élément positif (ce qui équivaut à la stabilité des éléments positifs par multiplication). On a dans un anneau ordonné la règle des signes usuelles (le produit de deux éléments positifs, ou de deux éléments négatifs, est positif, et le produit d'un élément positif et d'un élément négatif est négatif). Un anneau *totalement ordonné* est un anneau ordonné pour lequel la relation d'ordre est totale ; on a alors notamment $a^2 \geqslant 0_A$ pour tout $a$, et $1_A > 0_A$ (si l'anneau n'est pas trivial). Dans un anneau totalement ordonné, on peut définir la valeur absolue de tout élément $a$ par

$$|a| \overset{\text{def}}{=} \max(a, -a) \qquad \text{ou de façon équivalente} \qquad |a| \overset{\text{def}}{=} \begin{cases} a & \text{si } a \geqslant 0 \\ -a & \text{si } a < 0 \end{cases}$$

Un *corps gauche* est un anneau non trivial dans lequel tout élément non nul est inversible.

Les *corps* et les *anneaux intègres* ont tous les deux été définis de façon assez semblable : ce sont des anneaux commutatifs non triviaux, avec une propriété supplémentaire :

- dans le cas des corps : tous les éléments non nuls sont inversibles (un corps est donc un corps gauche commutatif) ;
- dans le cas des anneaux intègres : il n'y a aucun diviseur de zéro, ce qui équivaut à dire que pour tout $a$ et $b$

$$ab = 0_A \iff (a = 0_A \text{ ou } b = 0_A)$$

De plus dans les deux cas les éléments non nuls sont simplifiables pour la multiplication. Nous savons que tout corps est un anneau intègre (car un élément inversible ne peut pas être un diviseur de zéro). La réciproque n'est pas vraie dans le cas général (par exemple $\mathbb{Z}$ est un anneau intègre mais pas un corps), mais elle l'est pour des anneaux finis :

**Théorème 2.1.1 (Anneaux intègres finis)**
Tout anneau intègre fini est un corps.

**Ces pages ne sont pas incluses dans l’aperçu.**

et c'est un élément de $A$ qui commute avec $a$ et $b$. On notera en particulier que pour tout entier $n$, $n_A$ commute avec tout $a \in A$, donc *si* l'élément $n_A$ est inversible, alors on peut définir pour tout $a \in A$

$$\frac{a}{n_A} \overset{\text{def}}{\equiv} a(n_A)^{-1} = (n_A)^{-1}a$$

Attention, $n_A$ n'est pas nécessairement inversible, même dans un corps, car il est possible d'avoir $n_A = 0$. Par exemple, dans le corps $\mathbb{Z}/3\mathbb{Z}$ (voir la section 4.3), on a $1 + 1 + 1 = 0$.

**Théorème 2.1.2**
Si $n$ est un entier non nul qui divise l'entier $k$, et si l'élément $n_A$ de l'anneau $A$ est inversible, alors pour tout $a \in A$

$$\frac{k}{n}a = k\frac{a}{n_A}$$

**Preuve**
On a

$$\begin{cases} \left(\frac{k}{n}a\right) \cdot n_A = n\left(\frac{k}{n}a\right) = \left(n\frac{k}{n}\right)a = ka \\ \left(k\frac{a}{n_A}\right) \cdot n_A = k\left(\frac{a}{n_A} \cdot n_A\right) = ka \end{cases}$$

Or $n_A$ est inversible, donc simplifiable, et par conséquent

$$\frac{k}{n}a = k\frac{a}{n_A}$$

## 2.2 Exemples de sommes classiques dans un anneau

**Prérequis**
Les calculs sur les sommes (section 4.8 du volume 2).

Commençons par deux théorèmes dont je pourrai utiliser les résultats dans certains calculs. Le premier nous assure que si, dans un anneau, deux sommes sont telles que les termes de l'une commutent avec ceux de l'autre, alors les deux sommes commutent :

**Théorème 2.2.1**
On considère des entiers $m \leqslant n$ et $p \leqslant q$, et deux familles $(a_i)_{m\leqslant i\leqslant n}$ et $(b_j)_{p\leqslant j\leqslant q}$ d'éléments d'un anneau $A$, telles que pour tout $i \in [m, n]$ et tout $j \in [p, q]$, $a_i$ et $b_j$ commutent. Alors

$$\left(\sum_{i=m}^{n} a_i\right)\left(\sum_{j=p}^{q} b_j\right) = \left(\sum_{j=p}^{q} b_j\right)\left(\sum_{i=m}^{n} a_i\right)$$

**Preuve**
On a

$$\left(\sum_{i=m}^{n} a_i\right)\left(\sum_{j=p}^{q} b_j\right) = \sum_{i=m}^{n}\sum_{j=p}^{q} a_i b_j = \sum_{i=m}^{n}\sum_{j=p}^{q} b_j a_i = \sum_{j=p}^{q}\sum_{i=m}^{n} b_j a_i = \left(\sum_{j=p}^{q} b_j\right)\left(\sum_{i=m}^{n} a_i\right)$$

Le deuxième théorème nous permet d'étendre certains résultats sur les sommes, que nous avons démontrés pour des produits d'éléments d'un anneau, à des termes de la forme $na$, où $n$ est un entier et $a$ un élément d'un anneau :

**Théorème 2.2.2**

1. On considère deux entiers $m \leqslant n$, un entier $k$ et une famille $(a_i)_{m\leqslant i\leqslant n}$ d'éléments d'un anneau. Alors

$$\sum_{i=m}^{n} k a_i = k \sum_{i=m}^{n} a_i$$

2. On considère deux entiers $m \leqslant n$, un élément $a$ d'un anneau et une famille $(k_i)_{m\leqslant i\leqslant n}$ d'entiers. Alors

$$\sum_{i=m}^{n} k_i a = \left(\sum_{i=m}^{n} k_i\right) a$$

**Preuve**
Par récurrence sur $n \geqslant m$ : le résultat est immédiat pour $n = m$ ($ka_m = ka_m$ et $k_m a = k_m a$), et si on fait l'hypothèse de récurrence au rang $n$, alors, en faisant appel aux propriétés rappelées dans la section précédente, on a

$$\sum_{i=m}^{n+1} k a_i = \sum_{i=m}^{n} k a_i + k a_{n+1} = k \sum_{i=m}^{n} a_i + k a_{n+1} = k\left(\sum_{i=m}^{n} a_i + a_{n+1}\right) = k\left(\sum_{i=m}^{n+1} a_i\right)$$

et

$$\sum_{i=m}^{n+1} k_i a = \sum_{i=m}^{n} k_i a + k_{n+1} a = \left(\sum_{i=m}^{n} k_i\right) a + k_{n+1} a = \left(\sum_{i=m}^{n} k_i + k_{n+1}\right) a = \left(\sum_{i=m}^{n+1} k_i\right) a$$

**Théorème 2.2.3 (Factorisation de $a^n - b^n$)**
Pour tous les éléments $a$ et $b$ d'un anneau $A$ tels que $ab = ba$, et pour tout entier naturel non nul $n$

$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^k \cdot b^{n-1-k} = (a-b)\sum_{k=0}^{n-1} a^{n-1-k} \cdot b^k$$

autrement dit

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})$$

**Preuve**
Comme $a$ et $b$ commutent, on a

$$(a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k} = \sum_{k=0}^{n-1} a^{k+1} b^{n-1-k} - \sum_{k=0}^{n-1} a^k b^{n-k} = \sum_{k=1}^{n} a^k b^{n-k} - \sum_{k=0}^{n-1} a^k b^{n-k} = a^n - b^n$$

On obtient l'autre égalité en effectuant le changement d'indice $k \longmapsto n-1-k$ dans la somme.

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 3

# Compléments de théorie des nombres

## 3.1 Compléments sur la divisibilité dans $\mathbb{N}$ et $\mathbb{Z}$

**Prérequis**
La divisibilité dans un anneau (section 2.4).

### Généralités

Nous allons d'abord préciser dans cette section quelques propriétés générales de la divisibilité dans $\mathbb{Z}$ (et dans son sous-ensemble $\mathbb{N}$), en reprenant notamment l'étude de la divisibilité dans un anneau intègre de la section 2.4. Les notions de pgcd et ppcm seront étudiées en détail dans la section 3.4.

Notons que si $a$ et $b$ sont des entiers naturels, $a$ divise $b$ dans $\mathbb{Z}$ si et seulement si $a$ divise $b$ dans $\mathbb{N}$ (au sens de la relation de divisibilité définie dans $\mathbb{N}$) :

- Si $a$ divise $b$ dans $\mathbb{N}$, cela signifie par définition qu'il existe $c \in \mathbb{N}$ tel que $b = ac$, donc a fortiori $a$ divise $b$ dans $\mathbb{Z}$ ($c \in \mathbb{Z}$).
- Réciproquement, si $a$ divise $b$ dans $\mathbb{Z}$, cela signifie par définition qu'il existe $c \in \mathbb{Z}$ tel que $b = ac$. Si $b = 0$ alors $a$ divise $b$ dans $\mathbb{N}$ ($a \times 0 = 0$). Sinon, $b > 0$ et $a > 0$, donc $c > 0$ d'après la règle des signes, et par conséquent $a$ divise $b$ dans $\mathbb{N}$.

On en déduit que l'ensemble des diviseurs de $a$ dans $\mathbb{N}$ est égal à l'ensemble des diviseurs positifs de $a$ dans $\mathbb{Z}$, autrement dit l'intersection de $\mathbb{N}$ et de l'ensemble des diviseurs de $a$ dans $\mathbb{Z}$ :

$$D_a^{\mathbb{N}} \overset{\text{def}}{=} \left\{d \in \mathbb{N} \mid d \mid_{\mathbb{N}} a\right\} = \left\{d \in \mathbb{Z}_+ \mid d \mid_{\mathbb{Z}} a\right\} = \mathbb{N} \cap \left\{d \in \mathbb{Z} \mid d \mid_{\mathbb{Z}} a\right\} \overset{\text{def}}{=} \mathbb{N} \cap D_a^{\mathbb{Z}}$$

De même, l'ensemble des multiples de $a$ dans $\mathbb{N}$ est égal à l'ensemble des multiples positifs de $a$ dans $\mathbb{Z}$, autrement dit l'intersection de $\mathbb{N}$ et de l'ensemble des multiples de $a$ dans $\mathbb{Z}$ :

$$a\mathbb{N} \overset{\text{def}}{=} \left\{m \in \mathbb{N} \mid a \mid_{\mathbb{N}} m\right\} = \left\{m \in \mathbb{Z}_+ \mid a \mid_{\mathbb{Z}} m\right\} = \mathbb{N} \cap \left\{m \in \mathbb{Z} \mid a \mid_{\mathbb{Z}} m\right\} \overset{\text{def}}{=} \mathbb{N} \cap a\mathbb{Z}$$

Je rappelle aussi quelques propriétés de la divisibilité dans $\mathbb{Z}$ et $\mathbb{N}$ :

- Si $a \in \mathbb{N}$ alors $a$ divise 1 si et seulement si $a = 1$. De même, puisque le produit de deux entiers relatifs est égal à 1 si et seulement si ils sont tous les deux égaux à 1, ou à $-1$, on en déduit qu'un entier relatif divise 1 si et seulement si il est égal à 1 ou $-1$.
- Si $a$ et $b$ sont des entiers naturels tels que $a \mid b$ et $b \neq 0$, alors $a \leqslant b$ (donc par contraposition si $a \mid b$ et $b < a$, alors $b = 0$, autrement dit le seul multiple de $a$ strictement inférieur à $a$ est 0).

## Éléments associés

Reprenons le concept d'éléments associés dans un anneau intègre. Puisque les seuls éléments inversibles de $\mathbb{Z}$ sont 1 et $-1$, si $a$ et $b$ sont des entiers relatifs

$$a \sim b \quad \equiv \quad a = b \text{ ou } a = -b \quad \equiv \quad |a| = |b|$$

On notera en particulier

$$a \sim |a|$$

On a aussi, d'après la définition des éléments associés,

$$|a| = |b| \quad \equiv \quad a\mathbb{Z} = b\mathbb{Z} \quad \equiv \quad D_a^{\mathbb{Z}} = D_b^{\mathbb{Z}}$$

Dans le cas où $a$ et $b$ sont des entiers naturels, $a \sim b$ si et seulement si $a = b$, puisque dans ce cas $a \mid b$ et $b \mid a$ si et seulement si $a = b$ (la relation de divisibilité n'est pas une relation d'ordre sur $\mathbb{Z}$, mais elle l'est sur $\mathbb{N}$), et on a aussi

$$a = b \quad \equiv \quad a\mathbb{N} = b\mathbb{N} \quad \equiv \quad D_a^{\mathbb{N}} = D_b^{\mathbb{N}}$$

Nous savons aussi que si $a \sim a'$ et $b \sim b'$, alors

$$a \mid b \iff a' \mid b'$$

En particulier, $a$ divise $b$ si et seulement si $|a|$ divise $|b|$.

## Division euclidienne

Je redonne pour finir le principe de la division euclidienne dans $\mathbb{N}$ ou $\mathbb{Z}$, auquel j'ajoute une variante que je n'avais pas mise dans le volume 2 :

**Théorème 3.1.1 (Division euclidienne)**

1. Pour tout $(a, b) \in \mathbb{N} \times \mathbb{N}^*$, il existe un unique couple $(q, r) \in \mathbb{N}^2$ tel que

$$\begin{cases} a = bq + r \\ r < b \end{cases}$$

2. Pour tout $(a, b) \in \mathbb{Z} \times \mathbb{N}^*$, il existe un unique couple $(q, r) \in \mathbb{Z}^2$ tel que

$$\begin{cases} a = bq + r \\ 0 \leqslant r < b \end{cases}$$

3. Pour tout $(a, b) \in \mathbb{Z} \times \mathbb{Z}^*$, il existe un unique couple $(q, r) \in \mathbb{Z}^2$ tel que

$$\begin{cases} a = bq + r \\ 0 \leqslant r < |b| \end{cases}$$

Les nombres $q$ et $r$ s'appellent respectivement le *quotient* et le *reste* de la division euclidienne de $a$ par $b$, ce que je note $\text{quo}(a, b)$ et $\text{res}(a, b)$.

**Preuve**
Les preuves des deux premières variantes ont été données dans le volume 2. On en déduit la preuve de la troisième (où $b$ peut être un entier négatif) : si $b < 0$ alors $-b > 0$, et on déduit de la deuxième variante deux entiers relatifs $q$ et $r$ tels que

$$\begin{cases} a = (-b)q + r = b(-q) + r \\ 0 \leqslant r < -b \end{cases}$$

Dans les deux cas ($b > 0$ ou $b < 0$), il existe deux entiers relatifs $q$ et $r$ tels que

$$\begin{cases} a = bq + r \\ 0 \leqslant r < |b| \end{cases}$$

(si $b > 0$ alors $|b| = b$, et si $b < 0$ alors $|b| = -b$). Et l'unicité se déduit de celle de la deuxième variante.

**Remarque 3.1.2 :**

1. Si $0 \leqslant a < b$ alors le reste de la division euclidienne de $a$ par $b$ est, par unicité, le nombre $a$ lui-même puisque $a = b \times 0 + a$ :
$$0 \leqslant a < b \implies \text{res}(a, b) = a$$

2. Le reste de la division euclidienne de $a$ par $b$ est nul si et seulement si $b$ divise $a$ :
$$\text{res}(a, b) = 0 \iff b \mid a$$
En effet, si $\text{res}(a, b) = 0$ alors $a = bq$, donc $b$ divise $a$, et réciproquement, si $b$ divise $a$ alors il existe $q \in \mathbb{Z}$ tel que $a = bq = bq + 0$, donc, par unicité du reste, $\text{res}(a, b) = 0$.

**Remarque 3.1.3 :** On trouve aussi la variante suivante de la division euclidienne, dans laquelle le reste peut être négatif : pour tout $(a, b) \in \mathbb{Z} \times \mathbb{Z}^*$, il existe un couple $(q, r) \in \mathbb{Z}^2$ tel que

$$\begin{cases} a = bq + r \\ |r| < |b| \end{cases}$$

Cette variante, qui se déduit immédiatement du théorème précédent (s'il existe $r$ tel que $0 \leqslant r < |b|$, alors a fortiori $|r| < b$), est un cas particulier d'une définition plus générale de division euclidienne valable dans d'autres anneaux (les anneaux dits *euclidiens*, qui partagent avec $\mathbb{Z}$ un certain nombre de propriétés liées à la divisibilité). Mais dans ce cas il n'y a plus unicité du couple $(q, r)$. Par exemple, la division euclidienne de 7 par 3 peut alors donner
$$3 \times 2 + 1 = 7 = 3 \times 3 - 2$$

## 3.2 Compléments sur les nombres premiers

**Prérequis**
Les nombres premiers (section 4.10 du volume 2) et les sommes et produits à support fini (section 4.8 du volume 2).

Commençons par une petite propriété simple des nombres premiers que je serai amené à employer dans un chapitre ultérieur.

**Ces pages ne sont pas incluses dans l’aperçu.**

**Preuve**
On considère un entier $n$ et son développement décimal

$$n = \sum_{k=0}^{p} a_k 10^k$$

On a

$$n = a_0 + \sum_{k=1}^{p} a_k 10^k = a_0 + \sum_{k=0}^{p-1} a_{k+1} 10^{k+1} = a_0 + 10 \sum_{k=0}^{p-1} a_{k+1} 10^k$$

Or les nombres 2, 5, 10 divisent 10, donc divisent aussi $10 \sum_{k=0}^{p-1} a_{k+1} 10^k$. On en déduit :

- $n$ est divisible par 2 si et seulement si $a_0$ est divisible par 2, c'est-à-dire (puisque $a_0 \in [0, 9]$) si et seulement si $a_0$ est égal à 0, 2, 4, 6, ou 8.
- $n$ est divisible par 5 si et seulement si $a_0$ est divisible par 5, c'est-à-dire si et seulement si $a_0$ est égal à 0 ou 5.
- $n$ est divisible par 10 si et seulement si $a_0$ est divisible par 10, c'est-à-dire si et seulement si $a_0$ est égal à 0.

De même

$$n = a_0 + 10a_1 + \sum_{k=2}^{p} a_k 10^k = a_0 + 10a_1 + \sum_{k=0}^{p-2} a_{k+2} 10^{k+2} = a_0 + 10a_1 + 100 \sum_{k=0}^{p-2} a_{k+2} 10^k$$

Or 4 divise 100, donc 4 divise $n$ si et seulement si 4 divise $a_0 + 10a_1$ (le nombre formé en décimal par les deux derniers chiffres de $\overline{n}$). Enfin, en notant $S$ la somme des chiffres de $\overline{n}$ on a

$$n - S = \sum_{k=0}^{p} a_k 10^k - \sum_{k=0}^{p} a_k = \sum_{k=0}^{p} a_k (10^k - 1)$$

Or pour tout entier $k$, $10 - 1$ divise $10^k - 1$, autrement dit 9 divise $10^k - 1$ (donc 3 divise aussi $10^k - 1$). Par conséquent 3 et 9 divisent $\sum_{k=0}^{p} a_k (10^k - 1)$. On en déduit que 3 (respectivement 9) divise $n$ si et seulement si 3 (respectivement 9) divise $S$.

**Remarque 3.3.7 :** Nous verrons d'autres exemples de critères de divisibilité dans la section 4.3 (exemple 4.3.18, p. 133).

## 3.4 Plus grand commun diviseur et plus petit commun multiple dans $\mathbb{N}$ et $\mathbb{Z}$

**Prérequis**
La divisibilité dans $\mathbb{N}$ et $\mathbb{Z}$ (section 3.1), les treillis (sections 8.7 du volume 1 et 2.6 du volume 2), et les idéaux d'un anneau (section 2.9 du volume 2).

### Définitions et premières propriétés

Nous allons étudier dans cette section les notions de plus grand commun diviseur (pgcd) et de plus petit commun multiple (ppcm), à la fois dans $\mathbb{N}$ et dans $\mathbb{Z}$. On pourrait penser que l'étude dans $\mathbb{Z}$ uniquement est suffisante, puisque $\mathbb{N} \subseteq \mathbb{Z}$, mais il peut être intéressant de définir ces notions sans faire appel à la structure d'anneau de $\mathbb{Z}$. Il y a par ailleurs une spécificité intéressante dans $\mathbb{N}$ : la relation de divisibilité est alors une relation d'ordre (ce qui n'est pas le cas dans $\mathbb{Z}$), et le pgcd et ppcm de deux éléments $a$ et $b$ sont respectivement la borne inférieure et la borne supérieure de $\{a, b\}$. Ainsi, $\mathbb{N}$, muni de la relation de divisibilité, est un treillis (il en est aussi de même pour $\mathbb{N}^*$).

**Théorème 3.4.1 (Plus grand commun diviseur et plus petit commun multiple dans $\mathbb{N}$)**
$(\mathbb{N}, |)$ et $(\mathbb{N}^*, |)$ sont des treillis, autrement dit toute paire d'entiers admet une borne inférieure et une borne supérieure pour la relation de divisibilité (et ces bornes sont différentes de 0 lorsque les deux entiers sont non nuls). De plus, pour tous les entiers $a$ et $b$ :

1. On appelle *plus grand commun diviseur* (ou en abrégé *pgcd*) de $a$ et $b$ la borne inférieure de $\{a, b\}$ (le plus *grand* des *minorants* de $\{a, b\}$ pour la relation de divisibilité, c'est-à-dire un diviseur de $a$ et de $b$ multiple de tous les diviseurs communs de $a$ et $b$), que l'on notera

$$a \wedge b \qquad \text{ou} \qquad \operatorname{pgcd}(a, b)$$

autrement dit par définition, $a \wedge b$ est l'unique entier tel que

- $a \wedge b$ divise $a$ et $b$ ;
- tout diviseur commun à $a$ et $b$ divise $a \wedge b$ ;

ce qui équivaut à

$$\forall d \in \mathbb{N}, ((d \mid a \text{ et } d \mid b) \iff d \mid a \wedge b)$$

ou encore

$$D_a \cap D_b = D_{a \wedge b}$$

De plus si $a$ et $b$ sont non nuls

$$a \wedge b = \prod_{p \in \mathbb{P}} p^{\min(v_p(a), v_p(b))}$$

2. On appelle *plus petit commun multiple* (ou en abrégé *ppcm*) de $a$ et $b$ la borne supérieure de $\{a, b\}$ (le plus *petit* des *majorants* de $\{a, b\}$ pour la relation de divisibilité, c'est-à-dire un multiple de $a$ et $b$ qui divise tous les multiples communs de $a$ et $b$), que l'on notera

$$a \vee b \qquad \text{ou} \qquad \operatorname{ppcm}(a, b)$$

autrement dit par définition, $a \vee b$ est l'unique entier tel que

- $a \vee b$ est un multiple de $a$ et de $b$ ;
- tout multiple commun à $a$ et $b$ est un multiple de $a \vee b$ ;

ce qui équivaut à

$$\forall m \in \mathbb{N}, ((a \mid m \text{ et } b \mid m) \iff (a \vee b) \mid m)$$

ou encore

$$a\mathbb{N} \cap b\mathbb{N} = (a \vee b)\mathbb{N}$$

De plus si $a$ et $b$ sont non nuls

$$a \vee b = \prod_{p \in \mathbb{P}} p^{\max(v_p(a), v_p(b))}$$

**Preuve**

1. Borne inférieure : si $a$ ou $b$ est nul : on peut choisir par exemple $a = 0$ (le raisonnement est semblable si $b = 0$). L'ensemble des diviseurs de $a$ est alors $\mathbb{N}$ et un diviseur commun à $a$ et $b$ est un diviseur de $b$. On en déduit que $b$ est la borne inférieure (pour $|$) de $a$ et $b$ ($0 \wedge b = b$). Si $a$ et $b$ sont différents de 0, notons

$$n :\equiv \prod_{p \in \mathbb{P}} p^{\min(v_p(a), v_p(b))}$$

et vérifions que $n$ est le pgcd de $a$ et $b$ :

**Ces pages ne sont pas incluses dans l’aperçu.**

**Théorème 3.4.17 (Théorème de Bachet-Bézout)**
Les entiers $a$ et $b$ sont premiers entre eux si et seulement si il existe des entiers $u$ et $v$ tels que

$$au + bv = 1$$

**Preuve**
Par définition, si $a$ et $b$ sont premiers entre eux, alors on a une relation de Bézout avec 1 comme pgcd. Réciproquement, si $au + bv = 1$, alors $a \wedge b \mid 1$, et par conséquent $a \wedge b = 1$.

**Remarque 3.4.18 (Remarque historique) :** Ce théorème porte traditionnellement, mais de façon impropre, le nom de Bézout, mais il est plus juste de l'attribuer au mathématicien et poète français Claude-Gaspard Bachet de Méziriac (1581–1638) [2] qui justifie, dans la seconde édition de ses *Problèmes plaisans et délectables, qui se font par les nombres* (1624), que si deux nombres $a$ et $b$ sont premiers entre eux, alors il existe des entiers $u$ et $v$ tels que $au + bv = 1$. Plusieurs problèmes sont consacrés à ce sujet, comme sa « Proposition XVIII » :

> « Deux nombres premiers entre eux étant donnés, trouver le moindre multiple de chacun d'iceux, surpassant de l'unité un multiple de l'autre. » [3]

Bézout, quant à lui, a étudié des problèmes connexes (on peut trouver par exemple dans son *Cours de mathématiques à l'usage des gardes du pavillon et de la marine* un exemple de résolution d'une équation en $x$ et $y$ de la forme $ax + by = c$), mais on lui doit surtout la généralisation du théorème précédent aux polynômes : parmi les conséquences d'une étude de Bézout exposée dans un mémoire de 1764 [4], il montre que le pgcd de deux polynômes peut s'exprimer comme combinaison linéaire de ces polynômes. L'attribution de ce principe à Bézout se fera au début du XX$^{e}$ siècle, et sera essentiellement popularisée par Bourbaki [5], dans le cadre plus général des anneaux principaux, dans un ouvrage publié en 1952 [6]. [7]

**Remarque 3.4.19 :** Nous verrons plus loin (algorithme 3.4.53, p. 100) comment trouver des coefficients $u$ et $v$ tels que

$$au + bv = a \wedge b$$

et de façon plus générale (théorème 3.4.55, p. 101) comment résoudre une équation d'inconnues $x$ et $y$ de la forme

$$ax + by = c$$

**Théorème 3.4.20 (Lemme de Gauss)**
Pour tous les entiers relatifs $a, b, c$

$$\begin{cases} a \mid bc \\ a \wedge b = 1 \end{cases} \implies a \mid c$$

2. D'où l'appellation *Théorème de Bachet-Bézout* que l'on peut trouver récemment dans le monde francophone.
3. Claude-Gaspard Bachet de MÉZIRIAC. *Problèmes plaisans et délectables, qui se font par les nombres. Partie recueillis de divers autheurs, partie inventez de nouveau avec leur démonstration.* 2$^{e}$ éd. 1624.
4. *Recherches sur le degré des équations résultantes de l'évanouissement des inconnues et sur les moyens qu'on doit employer pour trouver ces équations.*
5. Nicolas Bourbaki est le pseudonyme collectif d'un groupe de mathématiciens francophones, formé en 1935.
6. *Modules sur les anneaux principaux*.
7. Source du dernier paragraphe : Liliane ALFONSI. *De l'oubli à la reconnaissance : l'exemple des résultats mathématiques d'Étienne Bézout (1730-1783)*. avril 2009. URL : `https://hal.archives-ouvertes.fr/hal-00425064`. 134$^{e}$ Congrès national des Sociétés historiques et scientifiques à Bordeaux du 20 au 26 avril 2009 session "Savants célèbres ou obscurs".

**Preuve**
Puisque $a$ et $b$ sont premiers entre eux, il existe $u$ et $v$ dans $\mathbb{Z}$ tels que

$$au + bv = 1$$

donc en multipliant par $c$

$$auc + bvc = c$$

Or $a$ divise $auc$ et $bvc$ (car $a$ divise $bc$), donc $a$ divise $c$.

**Remarque 3.4.21 (Remarque historique) :** On trouve ce résultat dans les *Disquisitiones arithmeticae* [Recherches arithmétiques] (1801) du mathématicien, astronome et physicien allemand Carl Friedrich Gauss (1777–1855). Il s'agit en quelque sorte d'une généralisation du lemme d'Euclide (voir le théorème 3.4.27), qui dit que si un nombre premier $p$ divise le produit $ab$, alors $p$ divise $a$ ou $p$ divise $b$.

**Théorème 3.4.22 (Corollaire 1 : généralisation du lemme de Gauss)**
Pour tous les entiers relatifs $a, b_1, \dots, b_n, c$

$$\begin{cases} a \mid b_1 \dots b_n c \\ \forall i \in [1, n], a \wedge b_i = 1 \end{cases} \implies a \mid c$$

**Preuve**
On prouve par récurrence sur $n$ que pour tous les entiers $b_1, \dots, b_n, c$, on a la propriété indiquée : le cas où $n = 1$ correspond au lemme de Gauss, et si on fait l'hypothèse de récurrence au rang $n$, alors si $a$ divise $b_1 \dots b_n(b_{n+1}c)$, $a$ divise $b_{n+1}c$ par hypothèse de récurrence, donc $a$ divise $c$ (d'après le lemme de Gauss).

**Théorème 3.4.23 (Corollaire 2)**
1. Pour tous les entiers $a, b, c$ tels que $a$ et $c$ sont premiers entre eux
$$a \wedge b = a \wedge (bc)$$
2. Plus généralement, pour tous les entiers $a, b, c_1, \dots, c_n$ tels que pour tout $i \in [1, n]$, $a$ et $c_i$ sont premiers entre eux
$$a \wedge b = a \wedge (bc_1 \dots c_n)$$
3. En particulier, pour tous les entiers $a, c_1, \dots, c_n$ tels que pour tout $i \in [1, n]$, $a$ et $c_i$ sont premiers entre eux
$$a \wedge (c_1 \dots c_n) = 1$$

**Preuve**
1. Prouvons que $a \wedge b$ est le pgcd de $a$ et $bc$. D'une part, $a \wedge b$ divise $a$ et $b$ (donc aussi $bc$). D'autre part, si $d$ divise $a$ alors $d$ et $c$ sont premiers entre eux puisque $a$ et $c$ le sont. Par conséquent si $d$ divise $a$ et $bc$, alors $d$ divise $b$ d'après le lemme de Gauss, et par conséquent $d$ divise $a \wedge b$.
2. On en déduit la généralisation par une récurrence immédiate : le cas $n = 1$ correspond au point précédent, et si on fait l'hypothèse de récurrence au rang $n$, on a
$$a \wedge ((bc_1 \dots c_n)c_{n+1)}) = a \wedge (bc_1 \dots c_n) = a \wedge b$$
3. On obtient le dernier point en prenant pour $b$ l'un des $c_i$ : si pour tout $i \in [1, n]$, $a$ et $c_i$ sont premiers entre eux, alors
$$a \wedge (c_1 c_2 \dots c_n) = a \wedge c_1 = 1$$

**Ces pages ne sont pas incluses dans l’aperçu.**

**Théorème 3.4.45 (Algorithme d'Euclide)**
On considère deux entiers naturels $a$ et $b$ et la suite $(a_n, b_n)_{n\in\mathbb{N}^*}$ définie par récurrence par

$$\begin{cases}(a_1, b_1) = (a, b)\\ (a_{n+1}, b_{n+1}) = \begin{cases}(b_n, \text{res}(a_n, b_n)) & \text{si } b_n \neq 0\\ (a_n, b_n) & \text{si } b_n = 0\end{cases}\end{cases}$$

Alors il existe un entier $m$ tel que

$$\begin{cases}m = \min\{n \in \mathbb{N}^* \mid b_n = 0\}\\ a \wedge b = a_m\end{cases}$$

et si $m > 1$

$$a \wedge b = a_m = b_{m-1}$$

**Preuve**
Il existe $k \in \mathbb{N}^*$ tel que $b_k = 0$, car sinon pour tout entier $n \geqslant 1$, $b_{n+1}$ serait le reste de la division euclidienne de $a_n$ par $b_n$ donc on aurait $b_{n+1} < b_n$ et la suite $(b_n)_{n\in\mathbb{N}^*}$ serait strictement décroissante, ce qui est impossible. Notons alors

$$m :\equiv \min\{n \in \mathbb{N}^* \mid b_n = 0\}$$

Si $m = 1$ alors $b = b_1 = 0$ et on a bien

$$a \wedge b = a = a_1$$

Nous supposons dans la suite que $m > 1$ (donc $b \neq 0$). Prouvons par récurrence sur $n$ que si $n < m$ alors $a \wedge b = a_n \wedge b_n$.

- Pour $n = 1$, on a $1 < m$ et
$$a \wedge b = a_1 \wedge b_1$$
- Faisons l'hypothèse de récurrence au rang $n$. Si $n + 1 < m$ alors $n < m$ donc $b_n \neq 0$, et par conséquent
$$(a_{n+1}, b_{n+1}) = (b_n, \text{res}(a_n, b_n))$$
De plus $b_{n+1} \neq 0$ (car $n + 1 < m$) et
$$a_{n+1} \wedge b_{n+1} = b_n \wedge \text{res}(a_n, b_n) = a_n \wedge b_n = a \wedge b$$

On en déduit par récurrence que pour tout $n < m$, $a \wedge b = a_n \wedge b_n$. Enfin, puisque $b_{m-1} \neq 0$, on a

$$(a_m, 0) = (a_m, b_m) = (b_{m-1}, \text{res}(a_{m-1}, b_{m-1}))$$

donc $b_{m-1}$ divise $a_{m-1}$ (car $0 = b_m = \text{res}(a_{m-1}, b_{m-1})$), et par conséquent

$$a_m = b_{m-1} = a_{m-1} \wedge b_{m-1} = a \wedge b$$

**Remarque 3.4.46 :** Dans le cas où $a < b$, la première étape de l'algorithme revient à permuter les deux nombres (ce qu'il peut aussi être préférable de faire avant de l'effectuer, pour éviter une étape inutile), car dans ce cas le reste de la division euclidienne de $a$ par $b$ est $a$ lui-même, et les deux premières divisions euclidiennes effectuées sont

$$a = b \times 0 + a$$

puis

$$b = a \times q + r$$

Voici l'algorithme sous la forme d'un script python définissant la fonction qui à deux entiers $a$ et $b$ associe $a \wedge b$ :

**Algorithme 3.4.47 (Algorithme d'Euclide)**
Entrée : deux entiers naturels $a$ et $b$.
Sortie : $a \wedge b$.

```
def pgcd_euclide(a,b):
    while (b != 0): # tant que b est différent de 0
        a,b = b,a%b     # a prend la valeur b
                        # et b prend la valeur res(a,b)
    return a
```

**Exemple 3.4.48 (Exemple de calcul de pgcd)**
Calculons le pgcd de 341 et 36 en effectuant les divisions euclidiennes successives :

$$\begin{aligned} 341 &= \mathbf{36} \times 9 + \mathbf{17} \\ 36 &= \mathbf{17} \times 2 + \mathbf{2} \\ 17 &= \mathbf{2} \times 8 + \mathbf{1} \\ 2 &= \mathbf{1} \times 2 + \mathbf{0} \end{aligned}$$

donc $341 \wedge 36 = 1$.

Dans l'algorithme originel d'Euclide, qui apparaît dans le livre VII de ses *Éléments*, les divisions euclidiennes sont réalisées à l'aide de soustractions successives, ce qui est essentiellement le même algorithme, car la division euclidienne de $a$ par $b$

$$a = bq + r$$

revient, à partir de $a$, à effectuer $q$ soustractions de $b$ successives jusqu'à obtenir un entier $r < b$. Une variante de l'algorithme utilise aussi exclusivement des soustractions : on pose

$$\begin{cases} a_0 :\equiv \max(a, b) \\ b_0 :\equiv \min(a, b) \end{cases}$$

et on construit la suite $(a_n, b_n)_{n\in\mathbb{N}}$ telle que

$$(a_{n+1}, b_{n+1}) = (\max(b_n, a_n - b_n), \min(b_n, a_n - b_n))$$

Ainsi pour tout entier $n$, on a

$$a_{n+1} \wedge b_{n+1} = b_n \wedge (a_n - b_n) = a_n \wedge b_n$$

On en déduit par une récurrence immédiate, que pour tout entier $n$

$$\begin{cases} a_n \wedge b_n = a \wedge b \\ b_n \leqslant a_n \end{cases}$$

donc la suite $(a_n)_{n\in\mathbb{N}}$ est décroissante : en effet, soit $a_{n+1} = b_n \leqslant a_n$, soit $a_{n+1} = a_n - b_n \leqslant a_n$. Comme cette suite ne peut pas être strictement décroissante, il existe un entier $m$ tel que $a_m = a_{m+1}$. Si $0 < b_m < a_m$ alors $a_{m+1} < a_m$ (car $a_{m+1} = b_m < a_m$ ou $a_{m+1} = a_m - b_m < a_m$), donc

- soit $b_m = 0$ ;
- soit $b_m = a_m$, et alors $a_{m+1} = b_m = a_m$ et $b_{m+1} = 0$.

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 4

# Groupes et anneaux quotients

**Prérequis**
Les groupes (sections 8.2 du volume 1 et 2.7 du volume 2) et les anneaux (sections 8.4 du volume 1 et 2.9 du volume 2), les relations d'équivalence et ensembles quotients (section 2.2 du volume 2), les morphismes (sections 2.10 à 2.12 du volume 2) et les lois quotients (section 2.13 du volume 2).

Nous reprenons ici l'étude des lois quotients, c'est-à-dire des lois que l'on peut mettre sur un ensemble quotient $E/\sim$, induites par une loi sur $E$ compatible avec la relation d'équivalence $\sim$. Nous verrons quels types de relations d'équivalence permettent de munir d'une structure de groupe l'ensemble quotient obtenu à partir d'un groupe, et de munir d'une structure d'anneau l'ensemble quotient obtenu à partir d'un anneau.

## 4.1 Groupes quotients

Dans le cas de groupes, avoir une relation d'équivalence compatible avec la loi impose certaines conditions à cette relation :

**Théorème 4.1.1 (Relations d'équivalence compatibles avec une loi de groupe)**
On considère un groupe $G$.

1. Pour tout sous-groupe $H$ de $G$, et pour tous les éléments $a$ et $b$ de $G$, on a les équivalences suivantes :

$$a^{-1}b \in H \quad \equiv \quad b^{-1}a \in H \quad \equiv \quad b \in aH \quad \equiv \quad a \in bH \quad \equiv \quad aH = bH$$

   De plus, la relation

$$a \sim b :\equiv \ a^{-1}b \in H$$

   est une relation d'équivalence compatible à gauche avec la loi interne. La classe d'équivalence de $a$ est $aH$ (en particulier $H$ est la classe de l'élément neutre). On l'appelle la *classe à gauche* de $a$.
2. Réciproquement, si $\sim$ est une relation d'équivalence sur $G$ compatible à gauche avec la loi interne, alors la classe de l'élément neutre est un sous-groupe $H$ de $G$ tel que

$$a \sim b \iff a^{-1}b \in H$$

3. Pour tout sous-groupe $H$ de $G$, et pour tous les éléments $a$ et $b$ de $G$, on a les équivalences suivantes :

$$ab^{-1} \in H \quad \equiv \quad ba^{-1} \in H \quad \equiv \quad b \in Ha \quad \equiv \quad a \in Hb \quad \equiv \quad Ha = Hb$$

De plus, la relation

$$a \sim b \stackrel{\text{def}}{\equiv} ab^{-1} \in H$$

est une relation d'équivalence compatible à droite avec la loi interne. La classe d'équivalence de $a$ est $Ha$ (en particulier $H$ est la classe de l'élément neutre). On l'appelle la *classe à droite* de $a$.

4. Réciproquement, si $\sim$ est une relation d'équivalence sur $G$ compatible à droite avec la loi interne, alors la classe de l'élément neutre est un sous-groupe $H$ de $G$ tel que

$$a \sim b \iff ab^{-1} \in H$$

**Preuve**
Prouvons les deux premiers points du théorème (les démonstrations sont semblables pour les relations d'équivalence compatibles à droite).

1. La relation

$$a \sim b \stackrel{\text{def}}{\equiv} a^{-1}b \in H$$

est

- réflexive puisque

$$a^{-1}a = e \in H$$

- symétrique, puisque un élément appartient à un sous-groupe si et seulement si son symétrique lui appartient aussi :

$$a^{-1}b \in H \quad \equiv \quad (a^{-1}b)^{-1} \in H \quad \equiv \quad b^{-1}a \in H$$

- transitive, puisque si $a \sim b$ et $b \sim c$, alors

$$a^{-1}c = (a^{-1}b)(b^{-1}c) \in H$$

par stabilité d'un sous-groupe pour la loi de groupe.

C'est donc une relation d'équivalence, qui est compatible à gauche avec la loi car si $a^{-1}b \in H$, alors pour tout $c$ dans $G$

$$(ca)^{-1}(cb) = a^{-1}c^{-1}cb = a^{-1}b \in H$$

donc $ca \sim cb$. Par ailleurs, pour tout $a$ et $b$ dans $G$

$$a^{-1}b \in H \quad \equiv \quad aa^{-1}b \in aH \quad \equiv \quad b \in aH$$

ce qui montre que la classe d'équivalence de $a$ est $aH$. On en déduit aussi, du fait de la symétrie de la relation,

$$a^{-1}b \in H \quad \equiv \quad b^{-1}a \in H \quad \equiv \quad a \in bH$$

Et puisque $a \sim b$ si et seulement si leurs classes d'équivalence sont identiques

$$a \sim b \quad \equiv \quad aH = bH$$

2. Réciproquement, considérons une relation d'équivalence $\sim$ compatible à gauche avec la loi, et notons $H$ la classe de l'élément neutre $e$. Vérifions que $H$ est un sous-groupe :
   - $e \in H$ par définition d'une classe d'équivalence.
   - Si $a$ et $b$ sont des éléments de $H$, alors par définition $a \sim e$ et $b \sim e$ donc par compatibilité à gauche avec la loi

$$a^{-1}b \sim a^{-1}e \sim a^{-1}a = e$$

et par conséquent $a^{-1}b \in H$.

De plus, comme la relation $\sim$ est compatible à gauche avec la loi,

$$a \sim b \quad \equiv \quad a^{-1}a \sim a^{-1}b \quad \equiv \quad e \sim a^{-1}b \quad \equiv \quad a^{-1}b \in H$$

**Remarque 4.1.2 :** Si la notation de la loi de groupe est additive :

**Ces pages ne sont pas incluses dans l’aperçu.**

De plus $\varphi$ est surjectif puisque $\mathrm{pr}_{N'}$ l'est, donc son image est $G/N'$, et son noyau est

$$\mathrm{Ker}(\varphi) = \left\{xN \mid x \in G \text{ et } xN' = N'\right\} = \left\{xN \mid x \in N'\right\} = N'/N$$

Par conséquent $N'/N$ est un sous-groupe normal de $G/N$, et d'après le premier théorème d'isomorphisme

$$(G/N)/\mathrm{Ker}(\varphi) \simeq \mathrm{Im}(\varphi)$$

autrement dit

$$(G/N)/(N'/N) \simeq G/N'$$

**Remarque 4.1.21 (Vocabulaire) :** Les dénominations « premier théorème d'isomorphisme,... » sont usuelles, mais avec parfois des différences dans la numérotation (en particulier les deuxième et troisième théorèmes peuvent être inversés).

## 4.2 Anneaux quotients

Dans le cas des anneaux, les relations d'équivalence compatibles avec l'addition et la multiplication ne peuvent être que des congruences modulo un sous-groupe (pour l'addition), en raison de la structure de groupe additif des anneaux. Ces groupes étant commutatifs, tous les sous-groupes sont normaux, mais tous ne conviennent pas. Pour que la relation d'équivalence soit aussi compatible avec la multiplication, nous allons voir qu'on ne peut prendre comme sous-groupes que les idéaux bilatères :

**Théorème 4.2.1 (Anneau quotient)**

1. Pour tout idéal (bilatère) $I$ d'un anneau $(A, +, \times, 0, 1)$, la congruence modulo $I$ (pour l'addition), autrement dit la relation d'équivalence

$$a \equiv b \mod I \overset{\text{def}}{\iff} a - b \in I$$

est compatible avec la multiplication de $A$.

2. Réciproquement, si $\sim$ est une relation d'équivalence compatible avec l'addition et la multiplication, alors la classe de 0 est un idéal (bilatère) $I$ tel que

$$a \sim b \iff a - b \in I$$

3. Et alors le groupe quotient $(A/I, +, I)$ muni de la loi quotient de la multiplication, est un anneau, nommé *anneau quotient*. De plus
   - La projection canonique $A \xrightarrow{\mathrm{pr}} A/I$ est un morphisme d'anneaux, de noyau $I$.
   - Si $A$ est commutatif, alors $A/I$ est commutatif.

**Preuve**

1. La congruence modulo $I$ est définie par

$$a \equiv b \mod I \overset{\text{def}}{\iff} a - b \in I$$

Elle est compatible avec la multiplication par définition d'un idéal, car si $a \equiv b \mod I$ alors pour tout $c \in A$

$$\begin{cases} ca - cb = c(a-b) \in I \\ ac - bc = (a-b)c \in I \end{cases}$$

donc

$$\begin{cases} ca \equiv cb \mod I \\ ac \equiv bc \mod I \end{cases}$$

2. Si $\sim$ est une relation d'équivalence compatible avec l'addition et la multiplication, nous savons déjà que la classe de 0 est un sous-groupe $I$ de $(A, +, 0)$ tel que
$$a \sim b \iff a - b \in I$$
Il reste à prouver que $I$ est un idéal : si $i \in I$ et $a \in A$, alors $i \sim 0$ donc par compatibilité avec la multiplication
$$\begin{cases} a \cdot i \sim a \cdot 0 = 0 \\ i \cdot a \sim 0 \cdot a = 0 \end{cases}$$
et par conséquent $ai \in I$ et $ia \in I$.
3. La loi quotient de la multiplication dans $A/I$ est associative car la multiplication dans $A$ l'est (et elle est commutative si la multiplication dans $A$ l'est), elle admet pour élément neutre la classe de 1, et elle est distributive par rapport à la loi quotient de l'addition. On en déduit que $A/I$ muni des deux lois quotients est un anneau, et que la projection canonique est un morphisme d'anneaux (c'est un morphisme pour $+$ et $\times$ tel que l'image de l'unité soit l'unité) dont le noyau est $I$ (c'est le noyau du morphisme entre les deux groupes additifs sous-jacents).

**Remarque 4.2.2 :** Dans le cas particulier où $I$ est l'anneau trivial $\{0\}$, on a pour tout $a, b \in A$ :

$$a \equiv b \mod \{0\} \iff a = b$$

et la classe de $a$ est $\{a\}$. Dans le cas particulier où $I$ est l'anneau $A$ lui-même, on a $a \equiv b \mod A$ pour tout $a, b \in A$, et la classe de $a$ est $A$.

**Théorème 4.2.3 (Sous-groupe multiplicatif des éléments congrus à 1)**
Si $I$ est un idéal d'un anneau $A$, alors l'ensemble

$$\left\{a \in A^\times \mid a \equiv 1 \mod I\right\}$$

est un sous-groupe normal du groupe $(A^\times, \times, 1)$ des éléments inversibles.

**Preuve**
Notons

$$N :\equiv \left\{a \in A^\times \mid a \equiv 1 \mod I\right\}$$

$N$ est un sous-groupe de $A^\times$ car $1 \in N$, et le fait que la congruence soit compatible avec la multiplication implique que $N$ est stable pour la multiplication (si $a \equiv 1 \mod I$ et $b \equiv 1 \mod I$, alors $ab \equiv 1 \mod I$) et pour l'inverse (si $a \equiv 1 \mod I$, alors $1 = a^{-1}a \equiv a^{-1} \mod I$). De plus $N$ est un sous-groupe normal, car si $x \in N$ et $a \in A^\times$, alors

$$axa^{-1} - 1 = axa^{-1} - aa^{-1} = a(xa^{-1} - a^{-1}) = a(x-1)a^{-1}$$

Or $I$ est un idéal et $x - 1 \in I$ par définition de $N$, donc $a(x-1)a^{-1} \in I$, et par conséquent $axa^{-1} \in N$.

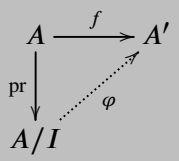
**Théorème 4.2.4 (Théorème de factorisation pour les anneaux)**
On considère un anneau $A$, un idéal $I$ de $A$, la projection canonique $A \xrightarrow{\text{pr}} A/I$ , et un morphisme d'anneaux $A \xrightarrow{f} A'$ .

1. Les propriétés suivantes sont équivalentes :
   - $f$ est compatible avec la congruence modulo $I$ (c'est-à-dire que $f$ est constant sur les classes

d'équivalence) : pour tout $x$ et $y$ dans $A$

$$x \equiv y \mod I \implies f(x) = f(y)$$

- $I$ est inclus dans le noyau de $f$ :
$$I \subseteq \mathrm{Ker}(f)$$
- Il existe un morphisme d'anneaux $A/I \xrightarrow{\varphi} A'$ tel que pour tout $x \in A$
$$f(x) = \varphi([x])$$
ce que l'on peut visualiser sur le diagramme commutatif suivant (où les flèches représentent des morphismes d'anneaux) :




2. Les propriétés suivantes sont alors vérifiées :
   - Le morphisme $\varphi$ est unique.
   - L'image de $\varphi$ est égale à l'image de $f$ (en particulier $\varphi$ est surjective si et seulement si $f$ est surjective).
   - $\varphi$ est injective si et seulement si les conditions équivalentes suivantes sont vérifiées
$$\forall x\ y \in A, (f(x) = f(y) \implies x \equiv y \mod I)$$
$$\forall x\ y \in A, (f(x) = f(y) \iff x \equiv y \mod I)$$
$$\mathrm{Ker}(f) \subseteq I$$
$$\mathrm{Ker}(f) = I$$

**Preuve**
Presque tout a déjà été démontré dans le théorème de factorisation pour les groupes. Il reste à justifier que le morphisme de groupes $\varphi$ est aussi un morphisme d'anneaux. On a

$$\varphi([1]) = f(1) = 1$$

et pour tout $x$ et $y$ dans $A$

$$\varphi([x][y]) = \varphi([xy]) = f(xy) = f(x)f(y) = \varphi([x])\varphi([y])$$

**Théorème 4.2.5 (Premier théorème d'isomorphisme pour les anneaux)**
On considère un morphisme d'anneaux $A \xrightarrow{f} A'$ et la projection canonique $A \xrightarrow{\mathrm{pr}} A/\mathrm{Ker}(f)$.
Il existe alors un unique morphisme d'anneaux $A/\mathrm{Ker}(f) \xrightarrow{\varphi} A'$ tel que pour tout $x \in A$

$$f(x) = \varphi([x])$$

De plus, ce morphisme est injectif et induit un isomorphisme

$$A/\mathrm{Ker}(f) \simeq \mathrm{Im}(f).$$

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 5

# Ordinaux

## 5.1 Définition et premières propriétés

**Prérequis**
Les ensembles transitifs, la définition des entiers naturels et de l'ensemble $\omega$ des entiers naturels (sections 3.1 à 3.3 du volume 2), et les relations de bon ordre (sections 1.13 et 2.11 du volume 2).

Les résultats de ce chapitre sont indépendants de l'axiome de fondation, et on ne supposera pas (sauf avis contraire) que cet axiome est nécessairement vérifié.

**Définition 5.1.1 (Ordinal)**
On appelle *ordinal* tout ensemble $\alpha$ vérifiant les trois propriétés équivalentes suivantes :

1. $\alpha$ est un ensemble transitif tel que la restriction de la relation $\in$ à $\alpha$ soit une relation de bon ordre strict, autrement dit :
   - $\alpha$ est un ensemble transitif : tout élément de $\alpha$ est un sous-ensemble de $\alpha$ :
   $$\forall x \in \alpha, x \subseteq \alpha$$
   - Antiréflexivité de la relation d'ordre strict : aucun élément de $\alpha$ n'appartient à lui-même :
   $$\forall x \in \alpha, x \notin x$$
   - Transitivité de la relation d'ordre strict : pour tous les éléments $x, y, z$ dans $\alpha$, si $x \in y$ et $y \in z$, alors $x \in z$ :
   $$\forall x\ y\ z \in \alpha, \left( \begin{cases} x \in y \\ y \in z \end{cases} \implies x \in z \right)$$
   - Relation de bon ordre : tout sous-ensemble non vide $b$ de $\alpha$ contient un plus petit élément (pour la relation d'ordre $\in$), c'est-à-dire un élément $x$ qui appartient à tous les autres éléments de $b$ :
   $$\forall b \subseteq \alpha, (b \neq \varnothing \implies (\exists x \in b, \forall y \in b, x \in y \text{ ou } x = y))$$
2. $\alpha$ est un ensemble bien ordonné tel que pour tout $\beta \in \alpha$
   $$\beta = ]-\infty, \beta[_\alpha$$
3. [en présence de l'axiome de fondation] $\alpha$ est un ensemble transitif dont les éléments sont des ensembles transitifs.

**Preuve (de l'équivalence des définitions)**

1. Faisons l'hypothèse qu'un ensemble $\alpha$ vérifie la première propriété, et démontrons qu'il vérifie la deuxième : si $\beta \in \alpha$ alors
$$]-\infty, \beta[_\alpha \overset{\text{def}}{\equiv} \{\gamma \in \alpha \mid \gamma \in \beta\}$$
donc $]-\infty, \beta[_\alpha \subseteq \beta$, et réciproquement, $\beta \subseteq ]-\infty, \beta[_\alpha$ car pour tout $\gamma \in \beta$, on a aussi $\gamma \in \alpha$ (puisque $\alpha$ est un ensemble transitif et $\beta \in \gamma$).
2. Faisons l'hypothèse qu'un ensemble $\alpha$ (muni d'une relation d'ordre $\leqslant$) vérifie la deuxième propriété, et démontrons qu'il vérifie la première : pour tout $\beta \in \alpha$
$$\beta = ]-\infty, \beta[_\alpha$$
et par conséquent :
   - Pour tout $\beta \in \alpha$ on a $\beta \subseteq \alpha$, donc $\alpha$ est un ensemble transitif.
   - Pour tout $\beta$ et $\gamma$ dans $\alpha$
   $$\gamma \in \beta \iff \gamma < \beta$$
   On en déduit que la restriction de $\in$ à $\alpha$ est une relation de bon ordre strict (puisque par hypothèse $<$ est une relation de bon ordre strict sur $\alpha$).
3. Faisons l'hypothèse qu'un ensemble $\alpha$ vérifie les propriétés 1 et 2 (équivalentes d'après ce qui précède), et démontrons qu'il vérifie aussi la propriété 3. Nous savons déjà que $\alpha$ est un ensemble transitif. Il reste à prouver que ses éléments le sont aussi. On considère $\beta \in \alpha$, $y \in \beta$, et $x \in y$. Comme $\beta \subseteq \alpha$, on a $y \in \alpha$ donc $x \in \alpha$ car $\alpha$ est transitif. Ainsi, $x, y, \beta$, sont des éléments de $\alpha$ tels que $x \in y$ et $y \in \beta$, donc par transitivité de la relation $\in$ sur $\alpha$, on en déduit $x \in \beta$. Par conséquent $\beta$ est un ensemble transitif.
4. Pour cette dernière implication, nous supposons que l'axiome de fondation est vérifié, et nous faisons l'hypothèse qu'un ensemble $\alpha$ vérifie la propriété 3. Démontrons qu'il vérifie aussi la propriété 1. Il reste à justifier que la restriction de $\in$ à $\alpha$ est une relation de bon ordre strict :
   - L'antiréflexivité de $\in$ est une conséquence de l'axiome de fondation ($\forall x \in \alpha, x \notin x$).
   - La transitivité de $\in$ est une conséquence de la transitivité des éléments de $\alpha$ (si $x, y, z$ dans $\alpha$ sont tels que $x \in y$ et $y \in z$, alors $x \in z$ par transitivité de $z$).
   - Relation de bon ordre : puisque la relation $\in$ est bien fondée d'après l'axiome de fondation, la seule chose restant à justifier est que la restriction de $\in$ à $\alpha$ est un ordre total. Cela revient à démontrer que si $x$ et $y$ sont des éléments de $\alpha$ tels que $x \notin y$ et $y \notin x$, alors $x = y$. Il suffit de prouver que tous les éléments $x$ et $y$ de $\alpha$ sont tels que $x \in y$ ou $y \subseteq x$ (d'où l'on déduit que si $x \notin y$ et $y \notin x$, alors $y \subseteq x$ et $x \subseteq y$, donc $x = y$). Considérons pour cela l'ensemble
   $$E :\equiv \left\{y \in \alpha \mid \exists x \in \alpha, y \nsubseteq x \text{ et } x \notin y\right\}$$
   dont nous voulons démontrer qu'il est vide. Faisons l'hypothèse contraire : d'après l'axiome de fondation, il existe $y \in E$ tel que $y \cap E = \varnothing$.

   — Par définition de $E$, il existe $x \in \alpha$ tel que
   $$y \nsubseteq x \text{ et } x \notin y$$
   En particulier $y \setminus x \neq \varnothing$ (car $y \nsubseteq x$), donc d'après l'axiome de fondation il existe $a \in y \setminus x$ tel que
   $$a \cap (y \setminus x) = \varnothing$$
   Par ailleurs $a$ est un élément de l'ensemble transitif $y$, donc $a \subseteq y$. Et comme $a \cap (y \setminus x) = \varnothing$, on en déduit $a \subseteq x$. De plus $a \neq x$ car $a \in y$ et $x \notin y$, donc $x \setminus a \neq \varnothing$. D'après l'axiome de fondation, il existe $b \in x \setminus a$ tel que
   $$b \cap (x \setminus a) = \varnothing$$
   Par ailleurs $b$ est un élément de l'ensemble transitif $x$, donc $b \subseteq x$. Et comme $b \cap (x \setminus a) = \varnothing$ on en déduit $b \subseteq a$.

   — On a $y \cap E = \varnothing$ par définition de $y$, donc $a \notin E$ (car $a \in y$). Comme par ailleurs $a$ et $b$ sont des éléments de $\alpha$ (car $a \in y$ et $y \in \alpha$, et de même $b \in x$ et $x \in \alpha$), on en déduit en particulier, par définition de $E$,
   $$a \subseteq b \text{ ou } b \in a$$
   — $b \in a$ est en contradiction avec $b \in x \setminus a$.

   — $a \subseteq b$ est aussi contradictoire, car si $a \subseteq b$ alors $a = b$ (puisque $b \subseteq a$), en contradiction avec $a \notin x$ et $b \in x$.

   On obtient donc une contradiction et par conséquent $E = \varnothing$ (négation de notre hypothèse de départ).

**Remarque 5.1.2 :** En présence de l'axiome de fondation, la formule $x \notin x$ est automatiquement vérifiée.

**Remarque 5.1.3 :** L'axiome de fondation a été utilisé pour démontrer que la troisième propriété implique les deux autres. Mais, même sans cet axiome, la preuve du théorème montre que si on définit un ordinal par l'une des deux premières propriétés (équivalentes), alors les éléments d'un ordinal sont aussi des ensembles transitifs.

**Remarque 5.1.4 :** On peut noter que le prédicat « être un ordinal » correspond à une formule du premier ordre dans le langage de la théorie des ensembles, la conjonction des formules de la définition ci-dessus :

$$\begin{array}{l}\alpha \text{ est un ordinal} \\ \overset{\text{def}}{\equiv} \\ \begin{cases} \forall x \in \alpha, x \subseteq \alpha \\ \forall x \in \alpha, x \notin x \\ \forall x, y, z \in \alpha, (x \in y \text{ et } y \in z) \implies x \in z \\ \forall b \subseteq \alpha, (b \neq \varnothing \implies (\exists x \in b, \forall y \in b, x \in y \text{ ou } x = y)) \end{cases}\end{array}$$

Ce qui permet de définir la classe des ordinaux, que je noterai $\boldsymbol{\mathcal{Ord}}$. Nous verrons plus loin que c'est une classe propre, c'est-à-dire que cette classe ne forme pas un ensemble, et j'utiliserai donc, comme indiqué dans le volume 2 à propos des classes, le symbole $\in$ pour les raccourcis suivants :

$$\begin{aligned} \alpha \in \boldsymbol{\mathcal{Ord}} &\overset{\text{def}}{\equiv} \alpha \text{ est un ordinal} \\ \forall \alpha \in \boldsymbol{\mathcal{Ord}}, \mathcal{F} &\overset{\text{def}}{\equiv} \forall \alpha, (\alpha \text{ est un ordinal} \implies \mathcal{F}) \\ \exists \alpha \in \boldsymbol{\mathcal{Ord}}, \mathcal{F} &\overset{\text{def}}{\equiv} \exists \alpha, (\alpha \text{ est un ordinal et } \mathcal{F}) \end{aligned}$$

Je rappelle que ce symbole $\in$ n'est pas formellement le même que le symbole $\in$ entre ensembles puisque $\boldsymbol{\mathcal{Ord}}$ n'est pas un ensemble.

**Théorème 5.1.5 (Propriétés élémentaires des ordinaux)**

1. L'ensemble vide est un ordinal.
2. Si $\alpha$ est un ordinal, alors $\alpha \notin \alpha$.
3. Si $\alpha$ est un ordinal, alors son successeur $\mathrm{S}\,\alpha$ (autrement dit $\alpha \cup \{\alpha\}$) est un ordinal.

**Preuve**

1. L'ensemble vide est un ordinal de façon triviale.
2. Pour tout ordinal $\alpha$, si $\alpha \in \alpha$ alors $\alpha \notin \alpha$ (comme tout élément d'un ordinal), ce qui est contradictoire, donc $\alpha \notin \alpha$.
3. On considère un ordinal $\alpha$. Nous savons déjà que le successeur d'un ensemble transitif est un ensemble transitif. Il reste à prouver les trois autres points de la définition :
   - Pour tout $x \in \mathrm{S}\,\alpha$, on a $x \in \alpha$ ou $x = \alpha$. Si $x \in \alpha$ alors $x \notin x$ (puisque $\alpha$ est un ordinal), et si $x = \alpha$ alors $x \notin x$ d'après la propriété précédente.
   - On considère des éléments $x, y, z$ de $\mathrm{S}\,\alpha$ tels que $x \in y$ et $y \in z$. Nous voulons démontrer que $x \in z$. Si $x, y, z$ sont dans $\alpha$, alors $x \in z$ puisque $\alpha$ est un ordinal. Sinon, au moins un de ces trois ensembles est égal à $\alpha$.
     — Si $x = \alpha$ alors $y \neq \alpha$ car $x \in y$ et $\alpha \notin \alpha$, donc $y \in \alpha$. On en déduit $\alpha \in y$ et $y \in \alpha$, donc $\alpha \in \alpha$ (car $\alpha$ est transitif), ce qui est impossible.
     — Si $y = \alpha$ alors $z \neq \alpha$ (car $\alpha \notin \alpha$), donc $z \in \alpha$. On en déduit comme précédemment que $\alpha \in z$ et $z \in \alpha$, donc $\alpha \in \alpha$, ce qui est impossible.
     — Si $z = \alpha$ alors $x \in y$ et $y \in \alpha$, donc $x \in \alpha$ puisque $\alpha$ est transitif, autrement dit $x \in z$.

**Ces pages ne sont pas incluses dans l’aperçu.**

On en déduit $\alpha \leqslant \sup A$, et par conséquent

$$\sup A = \alpha$$

## 5.4 Ordinaux et ensembles bien ordonnés

**Prérequis**
Les relations de bon ordre et les isomorphismes d'ensembles ordonnés (sections 1.13, 2.11, 5.2 et 6.3 du volume 2).

La construction précédente des ordinaux est due au mathématicien et physicien américano-hongrois John von Neumann (1903–1957). Une autre définition équivalente des ordinaux est basée sur les classes d'équivalence des ensembles bien ordonnés, c'est-à-dire qu'un ordinal représente une relation de bon ordre donnée (une collection d'ensembles isomorphes pour l'ordre). Cette approche est assez délicate pour la théorie des ensembles ZFC, puisqu'elle fait intervenir des collections qui ne sont pas des ensembles (la classe des ordinaux n'est pas un ensemble). On peut néanmoins retrouver cette approche des ordinaux en appliquant les résultats obtenus sur les ensembles ordonnés.

**Théorème 5.4.1**
Pour toutes classes d'ordinaux $\mathscr{A}$ et $\mathscr{B}$, pour toute fonction strictement croissante $\mathscr{A} \xrightarrow{f} \mathscr{B}$ , et tout $\alpha \in \mathscr{A}$

$$\alpha \leqslant f(\alpha)$$

**Preuve**
Faisons l'hypothèse que la classe

$$\mathscr{C} :\equiv \{\alpha \in \mathscr{A} \mid \alpha > f(\alpha)\}$$

n'est pas vide. $\mathscr{C}$ est une classe non vide d'ordinaux, donc contient un plus petit élément $\alpha$. Tout ordinal $\beta < \alpha$ n'appartient pas à $\mathscr{C}$, donc est tel que

$$\beta \leqslant f(\beta) < f(\alpha)$$

(car $f$ est strictement croissante), et par conséquent

$$\forall \beta \in \alpha, \beta \in f(\alpha)$$

autrement dit $\alpha \subseteq f(\alpha)$, ou encore $\alpha \leqslant f(\alpha)$, en contradiction avec $\alpha \in \mathscr{C}$. On en déduit $\mathscr{C} = \varnothing$, et par conséquent pour tout $\alpha \in \mathscr{A}$

$$\alpha \leqslant f(\alpha)$$

**Théorème 5.4.2 (Corollaire)**
Pour tous les ordinaux $\alpha$ et $\beta$, et toute fonction $\alpha \xrightarrow{f} \beta$ ,

1. si $f$ est strictement croissante, alors
$$\alpha \leqslant \beta$$
2. si $f$ est un isomorphisme pour l'ordre, alors
$$\begin{cases} \alpha = \beta \\ f = \mathrm{id}_\alpha \end{cases}$$

**Preuve**

1. D'après le théorème précédent, si $\beta \in \alpha$ (autrement dit $\beta < \alpha$), alors
$$\beta \leqslant f(\beta) < \beta$$
(puisque $f(\beta) \in \beta$ par définition de $f$), ce qui est contradictoire. On en déduit $\beta \geqslant \alpha$.
2. Si $\alpha \xrightarrow{f} \beta$ est un isomorphisme, alors $f$ et $f^{-1}$ sont strictement croissantes, donc d'après ce qui précède $\alpha \leqslant \beta$ et $\beta \leqslant \alpha$, et par conséquent $\alpha = \beta$. Et comme le seul automorphisme d'un ensemble bien ordonné est l'identité, on en déduit $f = \mathrm{id}_\alpha$. On pouvait aussi, sans faire appel à cette propriété, utiliser à nouveau le théorème précédent pour justifier $f = \mathrm{id}_\alpha$ : pour tout $\gamma \in \alpha$
$$\gamma \leqslant f(\gamma) \leqslant f^{-1}(f(\gamma)) = \gamma$$
donc $f(\gamma) = \gamma$.

**Théorème 5.4.3 (Isomorphisme entre ensemble bien ordonné et ordinal)**
Tout ensemble bien ordonné est isomorphe à un unique ordinal, et cet isomorphisme est unique.

**Preuve 1 (faisant appel au théorème de comparaison des bons ordres du volume 2)**
Notons tout d'abord que si un ensemble bien ordonné est isomorphe à un ordinal, cet ordinal est unique d'après le théorème précédent (puisque deux ordinaux isomorphes sont égaux), et l'isomorphisme est unique puisqu'il n'existe qu'un isomorphisme possible entre deux ensembles bien ordonnés (je ne referai pas cette remarque préliminaire pour les deux autres preuves de ce théorème). Il reste à justifier l'existence de cet ordinal. Il suffit pour cela d'appliquer le théorème de comparaison des bons ordres (théorème 2.11.19 du volume 2), au cas d'un ensemble bien ordonné $(E, \leqslant)$ et de la classe localement bien ordonnée $(\mathbf{Ord}, \in)$ :

- Soit $E$ est isomorphe à la classe des ordinaux, ce qui est impossible car $E$ est un ensemble et $\mathbf{Ord}$ une classe propre, et d'après le schéma de remplacement l'image d'un ensemble par une fonction doit être un ensemble.
- Soit la classe des ordinaux est isomorphe à un segment initial strict de $E$ ce qui est impossible pour la même raison.
- Soit $E$ est isomorphe à un segment initial strict de la classe des ordinaux, c'est-à-dire un ensemble de la forme $]-\infty, \alpha[_{\mathbf{Ord}}$, autrement dit $E$ est isomorphe à un ordinal $\alpha$.

**Preuve 2**
On considère un ensemble bien ordonné $E$. Construisons par récursion (métathéorème 6.3.1 du volume 2) la fonction $f$ de $E$ dans la classe de tous les ensembles, telle que pour tout $x \in E$
$$f(x) = \mathrm{Im}(f|_{]-\infty,x[}) = \{f(y) \mid y < x\}$$
et notons $\alpha$ l'image de cette fonction. D'après le schéma de remplacement, $\alpha$ est un ensemble. Nous allons démontrer que c'est un ordinal isomorphe à $E$. Voyons d'abord sur un exemple la construction de $\alpha$ : si $E$ est un ensemble bien ordonné ayant trois éléments tels que
$$a < b < c$$
alors
$$\begin{cases} f(a) = \{f(y) \mid y < a\} = \varnothing = 0 \\ f(b) = \{f(y) \mid y < b\} = \{f(a)\} = \{0\} = 1 \\ f(c) = \{f(y) \mid y < c\} = \{f(a), f(b)\} = \{0, 1\} = 2 \end{cases}$$
Dans cet exemple $\alpha = \{0, 1, 2\} = 3$.

- Prouvons que pour tout $x \in E$, $f(x) \notin f(x)$, en faisant l'hypothèse contraire que l'ensemble
$$A :\equiv \{x \in E \mid f(x) \in f(x)\}$$
est non vide. $A$ admet alors un plus petit élément $m$ tel que $f(m) \in f(m)$. Par définition de $f$, il existe $y < m$ tel que $f(m) = f(y)$. On en déduit $f(y) \in f(y)$, en contradiction avec le fait que $m$ soit le plus petit élément de $A$, et par conséquent $A = \varnothing$.
- Prouvons que $E \xrightarrow{f} \alpha$ est une bijection. Cette fonction étant surjective par définition ($\alpha$ est l'image de $f$), il reste à prouver l'injectivité. On considère $x$ et $y$ dans $E$ tels que $x \neq y$. L'ordre étant total, on a par exemple $y < x$ (le raisonnement est semblable si $x < y$), donc $f(y) \in f(x)$ par définition de $f$. Or d'après ce qui précède $f(x) \notin f(x)$, donc $f(x) \neq f(y)$.

**Ces pages ne sont pas incluses dans l’aperçu.**

**Théorème 5.6.5 (Caractérisation des fonctions normales)**
Si une fonction $\mathbf{Ord} \xrightarrow{f} \mathbf{Ord}$ est telle que pour tout ordinal limite $\lambda$

$$f(\lambda) = \sup_{\alpha<\lambda}\{f(\alpha)\}$$

alors $f$ est normale si et seulement si

$$\forall \alpha \in \mathbf{Ord}, f(\alpha) < f(\mathrm{S}\,\alpha)$$

**Preuve**
L'un des sens de l'équivalence est immédiat, puisqu'une fonction strictement croissante est en particulier telle que

$$\forall \alpha \in \mathbf{Ord}, f(\alpha) < f(\mathrm{S}\,\alpha)$$

Réciproquement, faisons l'hypothèse que tout ordinal $\alpha$ est tel que $f(\alpha) < f(\mathrm{S}\,\alpha)$, et considérons un ordinal $\alpha$ quelconque. Prouvons par récurrence transfinie sur $\beta$, que si $\alpha < \beta$ alors $f(\alpha) < f(\beta)$ :

- Le résultat est trivial pour $\beta = 0$ ($\alpha \not< 0$).
- Faisons l'hypothèse de récurrence pour $\beta$. Si $\alpha < \mathrm{S}\,\beta$ alors $\alpha \leqslant \beta$ donc
  - soit $\alpha = \beta$, et alors $f(\alpha) = f(\beta)$ ;
  - soit $\alpha < \beta$, et alors par hypothèse de récurrence $f(\alpha) < f(\beta)$.

  On a donc dans tous les cas
  $$f(\alpha) \leqslant f(\beta) < f(\mathrm{S}\,\beta)$$
- Faisons l'hypothèse de récurrence pour tout $\gamma < \beta$ avec $\beta$ ordinal limite. Si $\alpha < \beta$ alors $\mathrm{S}\,\alpha < \beta$ (puisque $\beta$ est limite), donc
  $$f(\alpha) < f(\mathrm{S}\,\alpha) \leqslant \sup_{\gamma<\beta}\{f(\gamma)\} = f(\beta)$$

**Théorème 5.6.6 (Stabilité des ordinaux limites par les fonctions normales)**
Si $\mathbf{Ord} \xrightarrow{f} \mathbf{Ord}$ est une fonction normale, et $\lambda$ un ordinal limite, alors $f(\lambda)$ est aussi un ordinal limite.

**Preuve**
On considère une fonction normale $\mathbf{Ord} \xrightarrow{f} \mathbf{Ord}$ et un ordinal limite $\lambda$. Pour justifier que $f(\lambda)$, qui est non nul (car $f(\lambda) \geqslant \lambda > 0$), est un ordinal limite, il suffit de prouver que tout $\beta < f(\lambda)$ est tel que $\mathrm{S}\,\beta < f(\lambda)$. Comme $f(\lambda) = \sup_{\alpha<\lambda}\{f(\alpha)\}$, on déduit de la définition d'une borne supérieure que si $\beta < f(\lambda)$ alors il existe un ordinal $\alpha < \lambda$ tel que $f(\alpha) > \beta$. Puisque $f$ est strictement croissante on a $f(\alpha) < f(\lambda)$, donc

$$\mathrm{S}\,\beta \leqslant f(\alpha) < f(\lambda)$$

Par conséquent $f(\lambda)$ est un ordinal limite.

**Théorème 5.6.7**
Pour toute fonction normale $\mathbf{Ord} \xrightarrow{f} \mathbf{Ord}$ et tout ordinal $\gamma \geqslant f(0)$, l'*ensemble*

$$\{\alpha \in \mathbf{Ord} \mid f(\alpha) \leqslant \gamma\}$$

admet un plus grand élément.

**Preuve**
Notons

$$A :\equiv \{\alpha \in \mathbf{Ord} \mid f(\alpha) \leqslant \gamma\} = \left\{\alpha \in [0,\gamma]_{\mathbf{Ord}} \mid f(\alpha) \leqslant \gamma\right\}$$

(voir la remarque ci-dessous). L'ensemble $A$ est transitif car si $\alpha \in \beta$ et $\beta \in A$, alors par définition $\alpha < \beta$ et $f(\beta) \leqslant \gamma$ donc par stricte croissance de $f$

$$f(\alpha) < f(\beta) \leqslant \gamma$$

et par conséquent $\alpha \in A$. On en déduit que $A$ est un ensemble transitif d'ordinaux, donc c'est un ordinal. Si $A$ est un ordinal limite, alors

$$f(A) = \sup_{\alpha\in A}\{f(\alpha)\} \leqslant \gamma$$

donc $A \in A$, ce qui est contradictoire. On en déduit que $A$ n'est pas un ordinal limite. De plus $A \neq 0$ (autrement dit $A \neq \varnothing$) car $0 \in A$. Par conséquent $A$ est un ordinal successeur, donc il a un plus grand élément (son prédécesseur).

**Remarque 5.6.8 :** La classe

$$\mathscr{A} :\equiv \{\alpha \in \mathbf{Ord} \mid f(\alpha) \leqslant \gamma\}$$

est bien un ensemble, puisque un ordinal appartenant à $\mathscr{A}$ est tel que

$$\alpha \leqslant f(\alpha) \leqslant \gamma$$

donc $\mathscr{A}$ est inclus dans l'ensemble des ordinaux inférieurs ou égaux à $\gamma$, c'est-à-dire que $\mathscr{A}$ est l'ensemble

$$\left\{\alpha \in [0,\gamma]_{\mathbf{Ord}} \mid f(\alpha) \leqslant \gamma\right\}$$

**Théorème 5.6.9**
Pour toute fonction normale $\mathbf{Ord} \xrightarrow{f} \mathbf{Ord}$ et tout ensemble non vide $A$ d'ordinaux,

$$f(\sup A) = \sup_{\alpha\in A}\{f(\alpha)\}$$

**Preuve**

- Tout $\alpha \in A$ est tel que $\alpha \leqslant \sup A$, donc puisque $f$ est (strictement) croissante

  $$f(\alpha) \leqslant f(\sup A)$$

  et par conséquent

  $$\sup_{\alpha\in A}\{f(\alpha)\} \leqslant f(\sup A)$$

- Inversement, prouvons $f(\sup A) \leqslant \sup\limits_{\alpha\in A}\{f(\alpha)\}$ :
  - Si $\sup A = 0$ alors $A = \{0\}$, donc $\vec{f}(A) = \{f(0)\}$ et $f(\sup A) = \sup\limits_{\alpha\in A}\{f(\alpha)\}$.
  - Si $\sup A$ est le successeur d'un ordinal $\mu$, alors puisque $\mu < \sup A$ il existe $\beta \in A$ tel que $\mu < \beta$, donc $\sup A = \mathrm{S}\,\mu \leqslant \beta$. On en déduit

    $$f(\sup A) \leqslant f(\beta) \leqslant \sup_{\alpha\in A}\{f(\alpha)\}$$

  - Si $\sup A$ est un ordinal limite, alors pour tout $\mu < \sup A$ il existe $\beta \in A$ tel que $\mu < \beta$, donc

    $$f(\mu) < f(\beta) \leqslant \sup_{\alpha\in A}\{f(\alpha)\}$$

    On en déduit, puisque $f$ est une fonction normale et $\sup A$ un ordinal limite,

    $$f(\sup A) = \sup_{\mu<\sup A}\{f(\mu)\} \leqslant \sup_{\alpha\in A}\{f(\alpha)\}$$

**Théorème 5.6.10 (Corollaire)**
Pour toute fonction normale $\mathbf{Ord} \xrightarrow{f} \mathbf{Ord}$, pour tout ordinal $\beta$ et toute fonction $\beta \xrightarrow{g} \mathbf{Ord}$,

$$f(\sup_{\gamma<\beta}\{g(\gamma)\}) = \sup_{\gamma<\beta}\{f(g(\gamma))\}$$

**Preuve**
Appliquons le théorème précédent à l'image de $g$, qui est un ensemble d'ordinaux :

$$f(\sup \mathrm{Im}(g)) = \sup_{\alpha\in\mathrm{Im}(g)}\{f(\alpha)\}$$

Or

$$\{f(\alpha) \mid \alpha \in \mathrm{Im}(g)\} = \{f(g(\gamma)) \mid \gamma \in \beta\}$$

donc

$$f(\sup_{\gamma<\beta}\{g(\gamma)\}) = f(\sup \mathrm{Im}(g)) = \sup_{\gamma<\beta}\{f(g(\gamma))\}$$

## 5.7 Arithmétique des ordinaux

**Prérequis**
Les différentes variantes de l'ordre antilexicographique (section 1.14 du volume 2 et section 1.6 de ce volume), et le principe de récurrence transfinie (section 5.5).

### Introduction

Nous pouvons définir des opérations arithmétiques sur la classe des ordinaux, qui généralisent les opérations arithmétiques sur $\mathbb{N}$ (addition, multiplication,...), de façon équivalente :

1. Par la propriété des ensembles bien ordonnés isomorphes à un unique ordinal : le résultat d'une opération entre deux ordinaux peut être défini comme l'unique ordinal isomorphe à un ensemble bien ordonné construit à partir de ces deux ordinaux.
2. Par récurrence transfinie en généralisant, à l'aide du théorème 5.5.2, les définitions de l'arithmétique de Peano.

On notera que ces méthodes font toutes les deux appel au schéma de remplacement (via l'existence d'un ordinal isomorphe à un ensemble bien ordonné, ou via le principe de définition par récurrence transfinie).

Nous avons aussi défini des opérations entre cardinaux (section 4.11 du volume 2) qui généralisent les opérations arithmétiques sur $\mathbb{N}$. Attention, ces opérations ne sont pas équivalentes. Si deux ordinaux $\alpha$ et $\beta$ sont aussi des cardinaux (nous verrons dans le chapitre 6 la construction usuelle des cardinaux dans ZFC, qui sont choisis parmi la classe des ordinaux), l'addition, par exemple, des cardinaux $\alpha$ et $\beta$, ne donnera pas le même résultat (en général) que l'addition des ordinaux $\alpha$ et $\beta$. C'est le contexte qui permettra de déterminer s'il s'agit d'une addition ordinale, ou d'une addition cardinale.

## Addition

**Définition 5.7.1 (Addition ordinale)**
Pour tous les ordinaux $\alpha$ et $\beta$ on définit $\alpha + \beta$, de manière équivalente

1. Comme l'unique ordinal isomorphe à l'ensemble (bien ordonné) $\alpha \sqcup \beta$, muni de l'ordre antilexicographique.
2. Par récurrence transfinie sur $\beta$ :

$$\begin{cases} \alpha + 0 = \alpha \\ \alpha + \mathrm{S}\,\beta = \mathrm{S}(\alpha + \beta) \\ \alpha + \beta = \sup\limits_{\gamma<\beta}\{\alpha + \gamma\} & \text{si } \beta \in \mathscr{Lim} \end{cases}$$

**Preuve**

1. La définition par l'isomorphisme a un sens : puisque les ordinaux $\alpha$ et $\beta$ sont bien ordonnés, l'ensemble $\alpha \sqcup \beta$, muni de l'ordre antilexicographique, est un ensemble bien ordonné, donc il est isomorphe à un unique ordinal.
2. Justifions en détail que la définition récursive a un sens : pour tout ordinal $\alpha$, appliquons le principe général de définition par récurrence transfinie avec

$$A :\equiv \mathscr{Ord} \qquad a :\equiv \alpha \qquad F_1 : \begin{cases} \mathscr{Ord} \longrightarrow \mathscr{Ord} \\ \beta \longmapsto \mathrm{S}\,\beta \end{cases} \qquad F_2 : \begin{cases} \mathscr{L} \longrightarrow \mathscr{Ord} \\ f \longmapsto \sup \mathrm{Im}(f) \end{cases}$$

On a :

$$\alpha + \beta \overset{\text{def}}{\equiv} G(\beta)$$

où $\mathscr{Ord} \xrightarrow{G} \mathscr{Ord}$ est l'unique fonction telle que pour tout ordinal $\beta$

$$\begin{cases} G(0) = \alpha \\ G(\mathrm{S}\,\beta) = F_1(G(\beta)) = \mathrm{S}(G(\beta)) \\ G(\beta) = F_2(G|_\beta) = \sup \mathrm{Im}(G|_\beta) = \sup\limits_{\gamma<\beta}\{G(\gamma)\} & \text{si } \beta \in \mathscr{Lim} \end{cases}$$

La justification de l'équivalence des définitions sera donnée plus loin.

**Remarque 5.7.2 :** Il est possible de prouver ici que les deux définitions sont équivalentes, mais je préfère donner d'abord différentes propriétés de l'addition ordinale, d'où découlera de façon immédiate cette équivalence.

**Remarque 5.7.3 :** Quelques conséquences immédiates de la définition par récurrence transfinie :

1. La restriction de l'addition ordinale aux entiers coïncide avec l'addition qui a été définie sur l'ensemble $\omega$ (c'est la même définition). En particulier, l'addition de deux ordinaux finis est un ordinal fini.
2. Pour tout ordinal $\alpha$
$$\alpha + 1 = \alpha + \mathrm{S}\,0 = \mathrm{S}(\alpha + 0) = \mathrm{S}\,\alpha$$
3. Pour tout ordinal $\alpha$, la fonction $x \longmapsto \alpha + x$ est une fonction normale, puisque par définition si $\beta$ est un ordinal limite
$$\alpha + \beta = \sup_{\gamma<\beta}\{\alpha + \gamma\}$$
et cette fonction est strictement croissante car pour tout ordinal $\beta$
$$\alpha + \mathrm{S}\,\beta = \mathrm{S}(\alpha + \beta) > \alpha + \beta$$

**Ces pages ne sont pas incluses dans l’aperçu.**

Dans le cas où $\beta$ est un ordinal limite, le théorème précédent montre que

$$\sup_{\gamma<\beta}\{\alpha^\gamma\} = \alpha^\beta$$

mais $\sup_{\gamma<\beta}\{\gamma^\alpha\}$ peut être strictement inférieur à $\beta^\alpha$. Par exemple

$$\sup_{\gamma<\omega}\{\gamma^\omega\} = \omega < \omega^\omega$$

(car $0^\omega = 0$, $1^\omega = 1$, et si $1 < \gamma < \omega$ alors $\gamma^\omega = \omega$).

**Théorème 5.7.48 (Équivalence des définitions de l'exponentiation ordinale)**
Les deux définitions de l'exponentiation ordinale sont équivalentes.

**Preuve**
Si on définit $\alpha^\beta$ comme l'unique ordinal isomorphe à l'ensemble (bien ordonné) $\alpha^{(\beta)}$, muni de l'ordre antilexicographique, alors les théorèmes précédents prouvent que pour tout $\alpha \neq 0$ et tout $\beta$

$$\begin{cases} \alpha^0 = 1 \\ \alpha^{S\beta} = \alpha^\beta \cdot \alpha \\ \alpha^\beta = \sup_{\gamma<\beta}\{\alpha^\gamma\} & \text{si } \beta \in \mathscr{Lim} \end{cases}$$

donc la fonction $x \longmapsto \alpha^x$ est l'unique fonction vérifiant cette récurrence. De plus, si $\alpha = 0$ on a bien

$$0^\beta = \begin{cases} 1 & \text{si } \beta = 0 \\ 0 & \text{sinon} \end{cases}$$

Nous pouvons encore une fois compléter la liste des ordinaux :

$$\begin{array}{ccccccccccccccc}
 & 0 & \longrightarrow & 1 & \longrightarrow & 2 & \dots & \omega & \longrightarrow & \omega+1 & \longrightarrow & \omega+2 & \dots \\
\dots & \omega\cdot 2 & \longrightarrow & \omega\cdot 2+1 & \longrightarrow & \omega\cdot 2+2 & \dots & \omega\cdot 3 & \longrightarrow & \omega\cdot 3+1 & \longrightarrow & \omega\cdot 3+2 & \dots \\
\dots & \omega^2 & \longrightarrow & \omega^2+1 & \dots & \omega^2+\omega & \dots & \omega^2+\omega\cdot 2 & \dots & \omega^3 & \dots & \omega^\omega & \dots \\
\dots & \omega^\omega\cdot 2 & \dots & \omega^{\omega\cdot 2} & \dots & \omega^{\omega^2} & \dots & \omega^{(\omega^\omega)} & \dots & & & &
\end{array}$$

## 5.8 Théorème de Goodstein

Nous allons voir dans cette section une application classique des ordinaux : la démonstration du théorème de Goodstein (évoqué dans le volume 1), à l'aide de calculs sur les ordinaux, ce qui peut d'ailleurs sembler un peu étrange au premier abord, car ce théorème ne porte que sur les nombres entiers.

**Théorème 5.8.1 (Lemme)**
On considère un ordinal $\gamma$, des ordinaux $\alpha_1, \dots, \alpha_n$ ($n \neq 0$) tels que

$$\alpha_n > \dots > \alpha_1$$

et des ordinaux $\delta_1, \dots, \delta_n$ strictement inférieurs à $\gamma$. Alors pour tout ordinal $\beta > \alpha_n$

$$\gamma^{\alpha_n}\cdot\delta_n + \dots + \gamma^{\alpha_1}\cdot\delta_1 < \gamma^\beta$$

**Ces pages ne sont pas incluses dans l’aperçu.**

prouvable dans l'arithmétique de Peano (du premier ordre). Comme par ailleurs il n'est pas non plus réfutable (car si l'arithmétique de Peano prouvait la négation du théorème, ZFC prouverait aussi cette négation, ce qui n'est pas le cas sauf si ZFC est contradictoire), cela signifie que cet énoncé est indépendant dans l'arithmétique de Peano.

## 5.9 Univers de von Neumann et axiome de fondation

**Prérequis**
La clôture transitive d'un ensemble (section 6.4 du volume 2) et l'axiome de fondation (section 6.5 du volume 2).

Nous allons dans cette section définir ce que l'on appelle l'univers de von Neumann, et revenir sur l'axiome de fondation (je rappelle que nous ne supposons pas, dans ce chapitre, que cet axiome est automatiquement vérifié).

**Définition 5.9.1 (Hiérarchie cumulative, univers $\mathscr{V}$ de von Neumann [a])**

1. On définit par récurrence transfinie la classe des ensembles $V_\alpha$ (où $\alpha$ est un ordinal), que l'on appelle la *hiérarchie cumulative* de von Neumann) :

$$\begin{cases} V_0 = 0 \quad (= \varnothing) \\ V_{\alpha+1} = \mathscr{P}\big(V_\alpha\big) \\ V_\alpha = \bigcup\limits_{\beta<\alpha} V_\beta \qquad \text{si } \alpha \in \mathscr{Lim} \end{cases}$$

2. On appelle *univers de von Neumann*, que l'on note $\mathscr{V}$, la classe des ensembles qui appartiennent à un ensemble $V_\alpha$ :

$$\mathscr{V} \overset{\text{def}}{=\!=} \bigcup_{\alpha\in\mathscr{Ord}} V_\alpha \overset{\text{def}}{=\!=} \big\{x \mid \exists \alpha \in \mathscr{Ord}, x \in V_\alpha\big\}$$

*a*. John von Neumann (1903–1957), mathématicien et physicien américano-hongrois.

**Remarque 5.9.2 :** $\mathscr{V}$ est par définition la classe des ensembles qui *appartiennent* à un $V_\alpha$, mais si un ensemble $A$ est *inclus* dans un $V_\alpha$, alors $A$ appartient aussi à $\mathscr{V}$, car $A \in \mathscr{P}\big(V_\alpha\big) = V_{\alpha+1}$.

**Exemple 5.9.3**
On a

$$\begin{aligned} V_0 &= 0 \\ V_1 &= \mathscr{P}\big(V_0\big) = \{0\} = 1 \\ V_2 &= \mathscr{P}\big(V_1\big) = \{0, \{0\}\} = \{0, 1\} = 2 \end{aligned}$$

*matical Society* 14 (1982), p. 285-293.

$$V_3 = \mathscr{P}\big(V_2\big) = \{0, \{0\}, \{1\}, \{0,1\}\} = \{0, 1, 2, \{1\}\}$$

$V_4$ est un ensemble dont les 16 éléments sont

$$\begin{aligned}
&0\\
&1, \{1\}, \{2\}, \{\{1\}\}\\
&2, \{0,2\}, \{0,\{1\}\}, \{1,2\}, \{1,\{1\}\}, \{2,\{1\}\}\\
&3, \{0,1,\{1\}\}, \{0,2,\{1\}\}, \{1,2,\{1\}\}\\
&\{0,1,2,\{1\}\}
\end{aligned}$$

$V_5$ est l'ensemble des parties de $V_4$, contenant donc $2^{16}$ éléments ; et ainsi de suite. Pour le plus petit ordinal limite, $\omega$, $V_\omega$ est la réunion de tous les ensembles $V_n$ (avec $n$ entier naturel) :

$$V_\omega = \bigcup_{n<\omega} V_n = \bigcup_{n\in\omega} V_n$$

**Théorème 5.9.4**
Pour tout ordinal $\alpha$, $V_\alpha$ est un ensemble transitif.

**Preuve**
C'est immédiat par récurrence transfinie, car d'une part si un ensemble est transitif il en est de même pour l'ensemble de ses parties, et d'autre part la réunion d'une famille d'ensembles transitifs est transitive :

- $V_0$, c'est-à-dire l'ordinal 0, est (trivialement) transitif.
- Si $V_\alpha$ est un ensemble transitif, alors $\mathscr{P}\big(V_\alpha\big)$, c'est-à-dire $V_{\alpha+1}$, est transitif.
- Si pour tout $\beta$ strictement inférieur à l'ordinal limite $\alpha$, $V_\beta$ est transitif, alors $V_\alpha$, qui est une réunion (indexée par $\alpha$) d'ensembles transitifs, est transitif.

**Théorème 5.9.5**
La famille $(V_\alpha)_{\alpha\in\mathbb{Ord}}$ est croissante : si $\alpha$ et $\beta$ sont deux ordinaux tels que $\alpha \leqslant \beta$, alors

$$V_\alpha \subseteq V_\beta$$

**Preuve**
On considère un ordinal $\alpha$. Démontrons $V_\alpha \subseteq V_\beta$ par récurrence transfinie sur $\beta \geqslant \alpha$ :

- Le résultat est trivial pour $\beta = \alpha$.
- Faisons l'hypothèse $V_\alpha \subseteq V_\beta$. Puisque $V_\beta$ est un ensemble transitif

$$V_\alpha \subseteq V_\beta \subseteq \mathscr{P}\big(V_\beta\big) = V_{\beta+1}$$

- On considère un ordinal limite $\beta$ tel que $\beta \geqslant \alpha$. Si $\alpha = \beta$ le résultat est établi, et si $\alpha < \beta$ alors on a par définition

$$V_\alpha \subseteq \bigcup_{\gamma<\beta} V_\gamma = V_\beta$$

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 6

# Cardinaux

## 6.1 Définition et premières propriétés

J'ai donné dans le volume 2 une définition provisoire des cardinaux consistant à affirmer de manière axiomatique l'existence, pour tout ensemble $A$, d'un ensemble $|A|$ tel que $|A| = |B|$ si et seulement si $A$ et $B$ sont en bijection. Nous allons maintenant voir comment nous pouvons construire, à l'aide des ordinaux et de l'axiome du choix, une telle fonction $(A \longmapsto |A|)$ définie sur la classe de tous les ensembles.

**Définition 6.1.1 (Cardinal)**

1. On dit qu'un ordinal $\alpha$ est un *cardinal* lorsqu'il est le plus petit des ordinaux en bijection avec lui (c'est-à-dire que si l'ordinal $\beta$ est en bijection avec $\alpha$, alors $\alpha \leqslant \beta$) :

$$\alpha \text{ est un cardinal } \stackrel{\text{def}}{\equiv} \begin{cases} \alpha \in \mathbf{Ord} \\ \forall \beta \in \mathbf{Ord}, (\beta \simeq \alpha \implies \alpha \leqslant \beta) \end{cases}$$

   ce qui équivaut aussi à dire, par contraposition, que tout ordinal strictement inférieur à $\alpha$ n'est pas en bijection avec $\alpha$ :

$$\alpha \text{ est un cardinal } \stackrel{\text{def}}{\equiv} \begin{cases} \alpha \in \mathbf{Ord} \\ \forall \beta \in \mathbf{Ord}, (\beta < \alpha \implies \beta \not\simeq \alpha) \end{cases}$$

2. Pour tout ensemble $A$, on appelle *cardinal de $A$* le plus petit ordinal en bijection avec $A$, autrement dit l'unique cardinal en bijection avec $A$. On le note $|A|$ :

$$|A| \stackrel{\text{def}}{\equiv} \min\{\alpha \in \mathbf{Ord} \mid \alpha \simeq A\}$$

**Preuve**

La définition du cardinal de $A$ a bien un sens. En effet, nous savons qu'il existe au moins un ordinal en bijection avec $A$ (tout ensemble peut être bien ordonné d'après le théorème de Zermelo, et tout ensemble bien ordonné est isomorphe à un ordinal, donc est a fortiori en bijection avec cet ordinal). Par conséquent la classe des ordinaux en bijection avec $A$ est non vide, et admet un plus petit élément, qui est par définition le cardinal de $A$. Ce plus petit élément $\alpha$ est par ailleurs un cardinal (tout ordinal en bijection avec lui est en bijection avec $A$ donc est supérieur ou égal à $\alpha$), et il est unique puisque deux cardinaux distincts ne peuvent pas être en bijection (si deux ordinaux en bijection sont tels que $\alpha < \beta$, alors par définition $\beta$ ne peut pas être un cardinal).

**Remarque 6.1.2 (Notations) :** Pour désigner le cardinal de $A$, on trouve aussi d'autres notations comme

$\|A\|$, $\#A$, ou $\mathrm{card}(A)$.

**Remarque 6.1.3 :** Un ensemble $A$ est un cardinal si et seulement si $A = |A|$.

**Remarque 6.1.4 :** Pour tout ordinal $\alpha$, soit $\alpha$ est un cardinal, et alors $|\alpha| = \alpha$, soit $\alpha$ n'est pas un cardinal, et alors $|\alpha| < \alpha$.

**Théorème 6.1.5 (Caractérisation des ensembles en bijection par leur cardinal)**
Deux ensembles $A$ et $B$ sont en bijection si et seulement si $|A| = |B|$.

**Preuve**
Si $|A| = |B|$ alors $A$ et $B$ sont en bijection avec un même ordinal, donc sont en bijection. Réciproquement, si $A$ et $B$ sont en bijection, $|B|$ est un ordinal en bijection avec $B$, donc avec $A$, et par conséquent $|A| \leqslant |B|$. De la même manière $|B| \leqslant |A|$, donc $|A| = |B|$.

**Remarque 6.1.6 :** En particulier deux cardinaux sont en bijection si et seulement si ils sont égaux.

**Remarque 6.1.7 :** Nous retrouvons la propriété prise comme définition provisoire des cardinaux dans le volume 2. La définition précédente est donc bien compatible avec cette définition provisoire. En particulier, les trois propriétés suivantes sont équivalentes (et signifient de manière informelle que les ensembles $A$ et $B$ ont la même taille) :

- $A$ et $B$ sont en bijection.
- Il existe une injection de $A$ dans $B$ et une injection de $B$ dans $A$.
- $|A| = |B|$.

**Remarque 6.1.8 :** On peut noter que c'est l'axiome du choix (via le théorème de Zermelo) qui a permis d'associer un cardinal à tout ensemble. Sans cet axiome, on peut toujours associer à tout ensemble qui peut être bien ordonné, le plus petit des ordinaux qui sont en bijection avec lui, mais ce n'est plus possible si l'ensemble ne peut pas être bien ordonné (ce qui peut arriver si l'axiome du choix, équivalent au théorème du bon ordre de Zermelo, n'est pas vérifié). Dans ce cas une possibilité pour définir le cardinal de tout ensemble est d'utiliser l'astuce de Scott, c'est-à-dire de définir le cardinal d'un ensemble $A$ comme l'ensemble de tous les ensembles en bijection avec $A$, ayant un rang minimum (dans l'univers de von Neumann) parmi les ensembles en bijection avec $A$ :

$$|A| \overset{\text{def}}{\equiv} \left\{ X \mid \begin{cases} X \simeq A \\ \forall Y, (Y \simeq A \implies \mathrm{rang}(Y) \geqslant \mathrm{rang}(X)) \end{cases} \right\}$$

Il s'agit bien d'un ensemble, et pas d'une classe propre, inclus dans un $V_\alpha$ : en notant

$$\alpha :\equiv \min\left\{\beta \in \mathbf{Ord} \mid \exists X \in V_\beta, X \simeq A\right\}$$

la classe précédente est l'intersection de $V_\alpha$ et de la classe des ensembles en bijection avec $A$, donc est incluse dans $V_\alpha$ :

$$|A| = \left\{ X \in V_\alpha \mid \begin{cases} X \simeq A \\ \forall Y, (Y \simeq A \implies \mathrm{rang}(Y) \geqslant \mathrm{rang}(X)) \end{cases} \right\}$$

Comme $\alpha \leqslant \mathrm{rang}(A) + 1$ (car $A \in V_{\mathrm{rang}(A)+1}$ et $A \simeq A$), on a aussi

$$|A| = \left\{ X \in V_{\mathrm{rang}(A)+1} \mid \begin{cases} X \simeq A \\ \forall Y, (Y \simeq A \implies \mathrm{rang}(Y) \geqslant \mathrm{rang}(X)) \end{cases} \right\}$$

**Ces pages ne sont pas incluses dans l’aperçu.**

**Remarque 6.2.2 :** Il était possible de procéder en sens contraire, c'est-à-dire de définir, pour tous les cardinaux $\alpha$ et $\beta$,

$$\begin{aligned}\alpha + \beta &\overset{\text{def}}{\equiv} |\alpha \sqcup \beta| \\ \alpha \cdot \beta &\overset{\text{def}}{\equiv} |\alpha \times \beta| \\ \alpha^\beta &\overset{\text{def}}{\equiv} |\beta \longrightarrow \alpha|\end{aligned}$$

et d'en déduire, pour tous les ensembles $A$ et $B$,

$$\begin{aligned}|A| + |B| &= |A \sqcup B| \\ |A| \cdot |B| &= |A \times B| \\ |A|^{|B|} &= |B \longrightarrow A|\end{aligned}$$

**Remarque 6.2.3 :** Les cardinaux étant des ordinaux particuliers, nous pouvons aussi effectuer des opérations arithmétiques ordinales entre cardinaux. Attention, ces opérations sont en général différentes (pour des nombres transfinis). Par exemple pour les opérations ordinales

$$\begin{cases}\omega + 1 = \mathrm{S}\,\omega > \omega \\ \omega \cdot 2 = \omega + \omega > \omega \\ 2^\omega = \omega\end{cases}$$

mais pour les opérations cardinales

$$\begin{cases}\omega + 1 = \omega \\ \omega \cdot 2 = \omega \\ 2^\omega > \omega\end{cases}$$

Je rappelle que pour tout ensemble $A$, on a $|\mathscr{P}(A)| = 2^{|A|}$, et que les propriétés des opérations cardinales sont très proches de celles des opérations usuelles entre entiers (commutativité et associativité de l'addition et de la multiplication, distributivité, $\alpha^{\beta\cdot\gamma} = (\alpha^\beta)^\gamma$,...). De plus, pour tous les cardinaux $\alpha, \beta, \gamma$, si $\alpha \leqslant \beta$ alors

$$\begin{aligned}\alpha + \gamma &\leqslant \beta + \gamma \\ \alpha \cdot \gamma &\leqslant \beta \cdot \gamma \\ \alpha^\gamma &\leqslant \beta^\gamma\end{aligned}$$

et si $\gamma \neq 0$

$$\gamma^\alpha \leqslant \gamma^\beta$$

Les opérations impliquant au moins un cardinal transfini sont même plus simples que leurs équivalents entre entiers, grâce à diverses propriétés que l'on peut déduire du théorème suivant, démontré dans le volume 2 à l'aide du lemme de Zorn (théorème 5.6.1) : tout cardinal transfini $\kappa$ 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 $A \times A$.

On en déduit par une récurrence immédiate, que si $A$ est un ensemble infini alors pour tout entier non nul $n$, $|A^n| = |A|$. Nous avons vu dans le volume 2 d'autres corollaires *intéressants* :

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 7

# Exemples de théories alternatives des ensembles

## 7.1 Théories des classes de von Neumann-Bernays-Gödel (NBG) et de Morse-Kelley (MK)

La théorie des classes de von Neumann-Bernays-Gödel (NBG) [1] et celle de Morse-Kelley (MK) [2] sont très proches de la théorie usuelle des ensembles de Zermelo-Fraenkel (ZF) étudiée dans le volume 2. Ces théories s'expriment dans un langage égalitaire du premier ordre, avec comme signature le seul symbole d'appartenance $\in$, mais contrairement à ZF les objets étudiés ne sont pas des ensembles, mais des classes, dans le même sens informel que dans ZF, c'est-à-dire des collections d'ensembles vérifiant une même formule.

Comme pour les ensembles, l'axiome d'extensionnalité dit que des classes sont égales si elles ont les mêmes éléments :

**Axiome 7.1.1 (Axiome d'extensionnalité)**
Deux classes contenant les mêmes éléments sont égales : pour tout $\mathscr{A}, \mathscr{B}$

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

**Remarque 7.1.2 :** Comme dans ZF, la réciproque de l'implication, c'est-à-dire

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

est aussi vraie, mais découle de la définition même de l'égalité.

**Remarque 7.1.3 (Notations) :** Comme je l'ai déjà fait, il peut m'arriver de varier la forme des variables (majuscule ou minuscule, police scripte,…), pour faciliter la compréhension intuitive des formules, mais d'un point de vue formel il n'y a aucune différence : les variables représentent toutes des objets de la théorie.

Parmi toutes les classes, on distingue les ensembles comme les classes appartenant à une autre classe :

1. John von Neumann (1903–1957), mathématicien et physicien américano-hongrois, Paul Bernays (1888–1977), mathématicien suisse, et Kurt Gödel (1906–1978), logicien et mathématicien austro-américain.
2. Anthony Morse (1911–1984) et John L. Kelley (1916–1999), mathématiciens américains.

**Définition 7.1.4 (Ensemble, classe propre)**

1. On appelle *ensemble* toute classe $E$ appartenant à une autre classe, c'est-à-dire vérifiant la formule $\text{Ens}(E)$ suivante :
$$\text{Ens}(E) \overset{\text{def}}{\equiv} \exists \mathscr{A}, E \in \mathscr{A}$$
2. On appelle *classe propre* toute classe qui n'est pas un ensemble, c'est-à-dire que $\mathscr{C}$ est une classe propre lorsque
$$\forall \mathscr{A}, \mathscr{C} \notin \mathscr{A}$$

**Remarque 7.1.5 :** D'une certaine façon, les ensembles sont des classes « assez petites » pour appartenir à une autre classe, et les classes propres sont des « grandes classes ».

Je dirai que les quantificateurs d'une formule sont relativisés aux ensembles lorsque tous les quantificateurs qui apparaissent sont de l'une des formes suivantes :

$$\forall x, (\text{Ens}(x) \implies \mathscr{F})$$
$$\exists x, (\text{Ens}(x) \text{ et } \mathscr{F})$$

L'axiome d'extensionnalité peut s'exprimer en quantifiant uniquement sur les ensembles :

**Théorème 7.1.6 (Corollaire de l'axiome d'extensionnalité)**
Pour tout $\mathscr{A}, \mathscr{B}$
$$\forall x, (\text{Ens}(x) \implies (x \in \mathscr{A} \iff x \in \mathscr{B})) \implies \mathscr{A} = \mathscr{B}$$

**Preuve**
Faisons l'hypothèse que pour tout *ensemble* $x$
$$x \in \mathscr{A} \iff x \in \mathscr{B}$$
On considère une *classe* $x$. Si $x \in \mathscr{A}$ alors $x$ est un ensemble (par définition d'un ensemble) donc par hypothèse $x \in \mathscr{B}$. De même, si $x \in \mathscr{B}$ alors $x \in \mathscr{A}$. On en déduit
$$x \in \mathscr{A} \iff x \in \mathscr{B}$$
donc $\mathscr{A} = \mathscr{B}$ d'après l'axiome d'extensionnalité.

La différence principale entre les théories NBG et MK est dans le schéma d'axiomes de compréhension, qui dit que pour toute formule il existe une classe dont les éléments sont les objets vérifiant cette formule. Dans MK la formule est quelconque, alors que dans NBG la formule ne peut être quantifiée que sur des ensembles.

**Axiome 7.1.7 (Schéma de compréhension dans la théorie NBG)**
Pour toute formule $\mathscr{F}[x, \vec{a}]$ dont tous les quantificateurs sont relativisés aux ensembles, il existe une classe $\mathscr{A}$ dont les éléments sont les classes $x$ (qui sont donc des ensembles) vérifiant la formule $\mathscr{F}$ : pour tout $\vec{a}$
$$\exists \mathscr{A} \forall x, (x \in \mathscr{A} \iff (\text{Ens}(x) \text{ et } \mathscr{F}(x)))$$
On note cette classe (qui est unique par extensionnalité)
$$\{x \mid \mathscr{F}(x)\}$$

**Ces pages ne sont pas incluses dans l’aperçu.**

**Remarque 7.1.28 :** Si $A$ et $B$ sont des ensembles, alors $A \times B$ est un ensemble, car $A \times B$ est inclus dans l'ensemble $\mathscr{P}(\mathscr{P}(A \cup B))$ (même raisonnement que dans ZF).

On peut définir, comme dans ZF, les relations et fonctions, et l'ensemble $\mathbb{N}$ à l'aide de l'axiome de l'infini :

**Axiome 7.1.29 (Axiome de l'infini)**
Il existe un ensemble récurrent :

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

On trouve aussi une variante de l'axiome du choix de ZFC :

**Axiome 7.1.30 (Axiome global du choix)**
Toute classe $\mathscr{C}$ d'ensembles non vides admet une fonction de choix, c'est-à-dire une fonction qui à tout élément $A$ de $\mathscr{C}$ associe un élément de $A$.

Comme pour ZF, les axiomes précédents sont suffisants pour formaliser une grande partie des mathématiques usuelles, mais on trouve aussi d'autres axiomes déjà présents dans ZF, et que je signale aussi sans entrer dans les détails (nous les avons étudiés dans le cadre de ZF) :

**Axiome 7.1.31 (Axiome de fondation)**
Pour tout $\mathscr{C}$

$$\mathscr{C} \neq \varnothing \implies \exists x \in \mathscr{C}, x \cap \mathscr{C} = \varnothing$$

**Axiome 7.1.32 (Axiome de remplacement)**
L'image directe d'un ensemble par une relation fonctionnelle est un ensemble.

## 7.2 Théorie *Nouveaux Fondements avec Uréléments* (NFU)

### Introduction

La théorie *New Foundations with Urelements* [Nouveaux Fondements avec Uréléments], ou NFU, proposée par le mathématicien américain Ronald Jensen (1936–) en 1969 et popularisée par le mathématicien Randall Holmes [3], est une variante de la théorie *New Foundations* [Nouveaux Fondements], ou NF, élaborée en 1937 par le philosophe et logicien américain Willard Van Orman Quine (1908–2000). Ces théories s'ins-

3. Ce qui est dans cette section est essentiellement adapté de
- Randall HOLMES. « Elementary Set Theory with a Universal Set ». Dans : *Cahiers du Centre de logique* 10 (1998). URL : `http://www.cahiersdelogique.be/Copies/CCL10.pdf`.
- Randall HOLMES. *Proof, Sets, and Logic*. 2017. URL : `https://randall-holmes.github.io/proofsetslogic.pdf`.
J'ai par ailleurs repris l'axiomatisation de NFU que l'on trouve dans
Paul K. GORBOW. « Categorical New Foundations ». Dans : (2018). URL : `https://arxiv.org/abs/1705.05021`.

pirent de la théorie des types, mais peuvent s'exprimer en logique des prédicats, à l'aide d'un petit nombre d'axiomes et d'un schéma d'axiomes (schéma de compréhension). Ce schéma d'axiomes peut même être remplacé par un nombre fini d'axiomes, ce qui fait que les théories NF et NFU sont finiment axiomatisables.

La particularité de NFU par rapport à NF ou ZF, c'est que les objets de la théorie ne sont pas tous des ensembles. En plus des ensembles, on peut trouver d'autres éléments, qu'on appelle *atomes* ou *uréléments*. Comme dans toute théorie du premier ordre, il n'y a pas de différence de nature entre les objets représentés par les variables, et la distinction entre atome et ensemble doit se faire par un prédicat : je prendrai comme langage celui de la théorie ZF (langage égalitaire avec le prédicat binaire $\in$), auquel j'ajoute le prédicat unaire « Ens » ; ainsi « Ens($A$) » signifie de façon informelle « $A$ est un ensemble » (et les objets de la théorie qui ne sont pas des ensembles seront donc appelés *atomes*). J'ajoute aussi au langage l'opérateur binaire $(-,-)$ pour représenter les couples ; ainsi, $(a, b)$ représentera le couple obtenu à partir des objets $a$ et $b$. La raison pour laquelle les couples ne sont pas définis directement dans la théorie (comme pour ZF) sera donnée plus loin.

### Axiome des ensembles et axiome d'extensionnalité

**Axiome 7.2.1 (Axiome des ensembles[a])**
Tout objet de la théorie qui a des éléments est un ensemble : pour tout $x$ et $E$

$$x \in E \implies \text{Ens}(E)$$

*a*. Je désigne ainsi cet axiome qui n'a pas de nom standard.

L'axiome d'extensionnalité dans NFU ne porte que sur les ensembles :

**Axiome 7.2.2 (Axiome d'extensionnalité)**
Pour tout $A$ et $B$

$$(\text{Ens}(A) \text{ et } \text{Ens}(B) \text{ et } \forall x, (x \in A \iff x \in B)) \implies A = B$$

On peut définir l'inclusion entre ensembles, dont on prouve comme dans ZF qu'elle détermine une relation d'ordre sur la classe des ensembles :

**Définition 7.2.3 (Inclusion entre ensembles)**
On a :

$$A \subseteq B \stackrel{\text{def}}{\equiv} \text{Ens}(A) \text{ et } \text{Ens}(B) \text{ et } \forall x, (x \in A \implies x \in B)$$

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 8

# Introduction à la théorie des modèles

**Prérequis**
La logique des prédicats (chapitres 4 et 7 du volume 1) et les morphismes (section 2.10 du volume 2).

## 8.1 Structures et modèles

### Structures

La théorie des modèles est essentiellement l'étude des rapports entre la logique et différentes structures mathématiques. Quand on parle de théorie des modèles, il est souvent implicite que la logique étudiée est celle qui sous-tend les mathématiques usuelles, c'est-à-dire la logique du premier ordre classique (non intuitionniste). C'est le cas de ce chapitre, dont les trois premières sections vont en grande partie reprendre, résumer et parfois compléter des notions vues dans les deux premiers volumes des *Éléments de mathématiques pour le XXI^e siècle* (les résultats déjà démontrés seront le plus souvent redonnés sans preuve).

Notre point de départ est la notion de structure : une structure consiste en l'interprétation, dans un ensemble $M$, de données formées d'un ensemble $O$ (que l'on peut considérer, dans le cadre de la logique, comme un ensemble d'opérateurs, c'est-à-dire de symboles d'opérations), d'un ensemble $C$ (que l'on peut considérer, dans le cadre de la logique, comme un ensemble de symboles de constantes), et d'un ensemble $P$ (que l'on peut considérer, dans le cadre de la logique, comme un ensemble de symboles de prédicats), chaque opérateur ou symbole de prédicat ayant par ailleurs une arité (nombre entier non nul). Formellement :

**Définition 8.1.1 (Signature, langage, structure)**

1. On appelle *signature* (ou *type*) d'une structure tout quadruplet

$$L :\equiv (O, C, P, ar)$$

où $O, C, P$ sont des ensembles deux à deux disjoints, et $ar$ une fonction de $O \cup P$ dans $\mathbb{N}^*$ ; pour tout élément $x$ de $O$ ou de $P$, $ar(x)$ est l'arité de $x$.

2. On appelle *$L$-structure* (ou juste *structure*), ou *interprétation* (ou *réalisation*) de $L$, tout quadruplet formé d'un ensemble non vide $M$ (l'*univers*, ou le *domaine*, de la structure) et de trois familles indexées respectivement par $O, C, P$

$$\mathscr{M} :\equiv (M, (\mathscr{I}_\varphi)_{\varphi\in O}, (\mathscr{I}_c)_{c\in C}, (\mathscr{I}_R)_{R\in P})$$

ce que je pourrai aussi noter

$$\mathscr{M} :\equiv (M, (\varphi_{\mathscr{M}})_{\varphi\in O}, (c_{\mathscr{M}})_{c\in C}, (R_{\mathscr{M}})_{R\in P})$$

et tel que

- Pour tout $\varphi \in O$ (pour tout symbole d'opération $\varphi$), $\varphi_{\mathscr{M}}$ (ou $\mathscr{I}_{\varphi}$) est une opération d'arité $ar(\varphi)$ sur $M$, autrement dit une fonction $M^{ar(\varphi)} \xrightarrow{\varphi_{\mathscr{M}}} M$ .
- Pour tout $c \in C$ (pour tout symbole de constante $c$), $c_{\mathscr{M}}$ (ou $\mathscr{I}_{c}$) est un élément de $M$.
- Pour tout $R \in P$ (pour tout symbole de prédicat $R$), $R_{\mathscr{M}}$ (ou $\mathscr{I}_{R}$) est une relation d'arité $ar(R)$ sur $M$, autrement dit un sous-ensemble de $M^{ar(R)}$.

3. On appelle *cardinal* d'une signature le cardinal de l'ensemble $O \cup C \cup P$, autrement dit

$$|O| + |C| + |P|$$

4. On appelle *cardinal* d'une $L$-structure le cardinal de son univers, c'est-à-dire

$$|\mathscr{M}| = |M|$$

5. On appelle *langage* de signature $L$ l'ensemble des formules (bien formées) du premier ordre construites à partir de $L$. Je noterai $|L|$ le cardinal de cet ensemble.

**Remarque 8.1.2 (Vocabulaire) :** Je pourrai fréquemment faire la confusion entre langage et signature, notamment en parlant de « langage $L$ » plutôt que de « langage de signature $L$ », ou de « langage d'une structure » plutôt que de « signature d'une structure ».

**Remarque 8.1.3 (Vocabulaire et notations) :** Dans la suite, je pourrai

- désigner par « $L$-formule » toute formule s'exprimant dans le langage $L$, « $L$-théorie » toute théorie s'exprimant dans le langage $L$, etc ;
- noter le nom de la structure en exposant plutôt qu'en indice, pour désigner les interprétations des différents symboles, c'est-à-dire noter $\varphi^{\mathscr{M}}$, $c^{\mathscr{M}}$, $R^{\mathscr{M}}$, plutôt que $\varphi_{\mathscr{M}}$, $c_{\mathscr{M}}$, $R_{\mathscr{M}}$ ;
- parler d'*opérateur* ou d'*opération*, de *constante* et de *prédicat* (plutôt que de *symbole d'opération*, de *symbole de constante* et de *symbole de prédicat*).

**Remarque 8.1.4 :** Sauf avis contraire, les langages que nous considérerons seront égalitaires, c'est-à-dire qu'il y aura toujours dans leur signature le symbole de prédicat d'égalité $=$. Par ailleurs, l'interprétation de ce symbole est imposée (c'est la même dans toute $L$-structure), pour correspondre à l'égalité (de la métathéorie) dans l'ensemble $M$ :

$$=_{\mathscr{M}} \overset{\text{def}}{\equiv} \{(x, x) \mid x \in M\}$$

**Théorème 8.1.5 (Cardinal d'un langage)**
Le cardinal du langage de signature $L :\equiv (O, C, P, ar)$ (autrement dit le cardinal de l'ensemble des formules sur cette signature) est

$$|L| = \max(\aleph_0, |O|, |C|, |P|)$$

**Ces pages ne sont pas incluses dans l’aperçu.**

**Définition 8.2.8 (Plongement élémentaire)**
Pour toutes $L$-structures $\mathscr{M}$ et $\mathscr{N}$, on appelle *plongement élémentaire* toute fonction $M \xrightarrow{f} N$ telle que pour toute formule $\mathscr{F}[x_1, \dots, x_n]$ et pour tout $(a_1, \dots, a_n) \in M^n$

$$\mathscr{M} \vDash \mathscr{F}[a_1, \dots, a_n] \quad \text{si et seulement si} \quad \mathscr{N} \vDash \mathscr{F}[f(a_1), \dots, f(a_n)]$$

Je pourrai noter

$$\mathscr{M} < \mathscr{N}$$

pour signifier qu'il existe un plongement élémentaire de $\mathscr{M}$ dans $\mathscr{N}$.

**Remarque 8.2.9 :** D'après la caractérisation des plongements, un plongement élémentaire est aussi un plongement.

**Remarque 8.2.10 :** La composée de deux plongements élémentaires en est un : en effet, si les fonctions $M_1 \xrightarrow{f} M_2$ et $M_2 \xrightarrow{g} M_3$ sont des plongements élémentaires entre les $L$-structures $\mathscr{M}_1$, $\mathscr{M}_2$ et $\mathscr{M}_3$, alors la fonction $M_1 \xrightarrow{g \circ f} M_3$ est un plongement élémentaire, car pour toute formule $\mathscr{F}[x_1, \dots, x_n]$ et pour tout $(a_1, \dots, a_n) \in M_1^n$, on a les équivalences suivantes :

$$\begin{aligned}
&\mathscr{M}_1 \vDash \mathscr{F}[a_1, \dots, a_n] \\
&\mathscr{M}_2 \vDash \mathscr{F}[f(a_1), \dots, f(a_n)] \\
&\mathscr{M}_3 \vDash \mathscr{F}[g(f(a_1)), \dots, g(f(a_n))]
\end{aligned}$$

**Remarque 8.2.11 :** La relation $<$ est une relation de préordre sur toute classe de structures de même signature, car elle est réflexive (l'identité est un plongement élémentaire) et transitive (la composée de deux plongements élémentaires en est un).

**Définition 8.2.12 (Équivalence élémentaire de modèles)**
Pour toutes $L$-structures $\mathscr{M}$ et $\mathscr{N}$, on dit que $\mathscr{M}$ et $\mathscr{N}$ sont *élémentairement équivalentes*, ce que je noterai

$$\mathscr{M} \equiv \mathscr{N}$$

lorsqu'elles sont modèles des mêmes énoncés, c'est-à-dire lorsque pour toute formule close $\mathscr{F}$,

$$\mathscr{M} \vDash \mathscr{F} \quad \text{si et seulement si} \quad \mathscr{N} \vDash \mathscr{F}$$

autrement dit lorsque

$$\mathrm{Th}(\mathscr{M}) = \mathrm{Th}(\mathscr{N})$$

**Remarque 8.2.13 :** Je rappelle que $\mathrm{Th}(\mathscr{M})$ désigne la théorie de la $L$-structure $\mathscr{M}$, c'est-à-dire l'ensemble des formules closes dont $\mathscr{M}$ est un modèle.

**Remarque 8.2.14 :** L'équivalence élémentaire est une relation d'équivalence sur toute classe de structures de même signature.

**Remarque 8.2.15 :** En appliquant la définition d'un plongement élémentaire au cas particulier des formules closes, on voit que s'il existe un plongement élémentaire d'une structure dans une autre, alors elles sont élémentairement équivalentes.

**Ces pages ne sont pas incluses dans l’aperçu.**

2. Si $\mathscr{F}$ est une formule existentielle, alors sa négation équivaut à une formule universelle. Donc d'après ce qui précède, si $\mathscr{N} \vDash \neg\mathscr{F}[a_1, \dots, a_p]$ alors $\mathscr{M} \vDash \neg\mathscr{F}[a_1, \dots, a_p]$. On en déduit par contraposition que si $\mathscr{M} \nvDash \neg\mathscr{F}[a_1, \dots, a_p]$ alors $\mathscr{N} \nvDash \neg\mathscr{F}[a_1, \dots, a_p]$, autrement dit si $\mathscr{M} \vDash \mathscr{F}[a_1, \dots, a_p]$ alors $\mathscr{N} \vDash \mathscr{F}[a_1, \dots, a_p]$.

**Théorème 8.2.31**
Pour toutes $L$-structures $\mathscr{M}$ et $\mathscr{N}$, les trois propriétés suivantes sont équivalentes :

- Il existe un plongement de $\mathscr{M}$ dans $\mathscr{N}$.
- $\mathscr{M}$ est isomorphe à une sous-structure de $\mathscr{N}$.
- $\mathscr{N}$ est isomorphe à une extension de $\mathscr{M}$.

**Preuve**

- Faisons l'hypothèse que $f$ est un plongement de $\mathscr{M}$ dans $\mathscr{N}$. Alors l'image $M'$ de $f$ est un sous-univers de $\mathscr{N}$, car c'est un sous-ensemble de $N$ non vide (puisque $M$ est non vide), tel que
  - Pour toute constante $c$, $c_{\mathscr{N}} = f(c_{\mathscr{M}}) \in M'$.
  - Pour tout opérateur $n$-aire $\varphi$, et tout $(y_1, \dots, y_n) \in M'^n$, il existe $(x_1, \dots, x_n) \in M^n$ tel que pour tout $i \in [1, n]$, $y_i = f(x_i)$, donc
  $$\varphi_{\mathscr{N}}(y_1, \dots, y_n) = \varphi_{\mathscr{N}}(f(x_1), \dots, f(x_n)) = f(\varphi_{\mathscr{M}}(x_1, \dots, x_n)) \in M'$$

  La $L$-structure induite $\mathscr{M}'$ est isomorphe à $\mathscr{M}$ car la bijection $M \xrightarrow{f} M'$ est un plongement de $\mathscr{M}$ dans $\mathscr{M}'$ :
  - Pour tout opérateur $n$-aire $\varphi$ et tout $(x_1, \dots, x_n) \in M^n$
  $$f(\varphi_{\mathscr{M}}(x_1, \dots, x_n)) = \varphi_{\mathscr{N}}(f(x_1), \dots, f(x_n)) = \varphi_{\mathscr{M}'}(f(x_1), \dots, f(x_n))$$
  - Pour toute constante $c$, $f(c_{\mathscr{M}}) = c_{\mathscr{N}} = c_{\mathscr{M}'}$.
  - Pour tout prédicat $n$-aire $R$, et tout $(x_1, \dots, x_n) \in M^n$
  $$(x_1, \dots, x_n) \in R_{\mathscr{M}} \equiv (f(x_1), \dots, f(x_n)) \in R_{\mathscr{N}} \equiv (f(x_1), \dots, f(x_n)) \in R_{\mathscr{M}'}$$
- Réciproquement, faisons l'hypothèse qu'il existe un isomorphisme $f$ de $\mathscr{M}$ dans une sous-structure $\mathscr{M}'$ de $\mathscr{N}$. Alors $f$ est un plongement de $\mathscr{M}$ dans $\mathscr{N}$, car
  - Pour tout opérateur $n$-aire $\varphi$ et tout $(x_1, \dots, x_n) \in M^n$
  $$f(\varphi_{\mathscr{M}}(x_1, \dots, x_n)) = \varphi_{\mathscr{M}'}(f(x_1), \dots, f(x_n)) = \varphi_{\mathscr{N}}(f(x_1), \dots, f(x_n))$$
  - Pour toute constante $c$, $f(c_{\mathscr{M}}) = c_{\mathscr{M}'} = c_{\mathscr{N}}$.
  - Pour tout prédicat $n$-aire $R$, et tout $(x_1, \dots, x_n) \in M^n$
  $$(x_1, \dots, x_n) \in R_{\mathscr{M}} \equiv (f(x_1), \dots, f(x_n)) \in R_{\mathscr{M}'} \equiv (f(x_1), \dots, f(x_n)) \in R_{\mathscr{N}}$$
- Enfin, si $\mathscr{M} \sqsubseteq \mathscr{M}'$, et si $f$ est un isomorphisme de $\mathscr{M}'$ dans $\mathscr{N}$, alors la restriction de $f$ à $\mathscr{M}$ est un plongement de $\mathscr{M}$ dans $\mathscr{N}$. Réciproquement, faisons l'hypothèse qu'il existe un plongement $f$ de $\mathscr{M}$ dans $\mathscr{N}$. Nous pouvons construire une extension $\mathscr{M}'$ de $\mathscr{M}$, isomorphe à $\mathscr{N}$, de la façon suivante :
  - On considère un ensemble $A$ de cardinal $|N \setminus \mathrm{Im}(f)|$, tel que $A$ et $M$ soient disjoints. Prenons comme univers $M'$ de $\mathscr{M}'$ l'ensemble $M \cup A$. Ainsi $M'$ est tel que
  $$\begin{cases} M \subseteq M' \\ |M' \setminus M| = |N \setminus \mathrm{Im}(f)| \end{cases}$$

  Comme $M \xrightarrow{f} \mathrm{Im}(f)$ est une bijection, et comme il existe une bijection de $M' \setminus M$ dans $N \setminus \mathrm{Im}(f)$, la réunion de ces deux bijections nous permet de prolonger $f$ en une bijection $M' \xrightarrow{g} N$ .
  - Pour toute constante $c$
  $$c_{\mathscr{M}'} :\equiv g^{-1}(c_{\mathscr{N}})$$

— Pour tout opérateur $\varphi$ d'arité $n$ et tout $(x_1, \dots, x_n) \in M'^n$

$$\varphi_{\mathscr{M}'}(x_1, \dots, x_n) :\equiv g^{-1}(\varphi_{\mathscr{N}}(g(x_1), \dots, g(x_n)))$$

— Pour tout prédicat $R$ d'arité $n$ et tout $(x_1, \dots, x_n) \in M'^n$

$$(x_1, \dots, x_n) \in R_{\mathscr{M}'} :\equiv (g(x_1), \dots, g(x_n)) \in R_{\mathscr{N}}$$

Par construction $g$ est un isomorphisme de $\mathscr{M}'$ dans $\mathscr{N}$. De plus, $\mathscr{M}'$ est une extension de $\mathscr{M}$ car l'injection canonique $M \xrightarrow{id} M'$ est un plongement de $\mathscr{M}$ dans $\mathscr{M}'$, composée des plongements $\mathscr{M} \xrightarrow{f} \mathscr{N}$ et $\mathscr{N} \xrightarrow{g^{-1}} \mathscr{M}'$.

**Exemple 8.2.32**
Dans la construction classique de $\mathbb{Z}$ à partir de $\mathbb{N}$, on a un plongement de $(\mathbb{N}, +, \times, 0, 1, \leqslant)$ dans $(\mathbb{Z}, +, \times, 0, 1, \leqslant)$, et on obtient ainsi une sous-structure de $(\mathbb{Z}, +, \times, 0, 1, \leqslant)$, isomorphe à $(\mathbb{N}, +, \times, 0, 1, \leqslant)$, que l'on identifie à $(\mathbb{N}, +, \times, 0, 1, \leqslant)$. Autrement dit on obtient un sous-ensemble de $\mathbb{Z}$ que l'on identifie à l'ensemble $\mathbb{N}$, deux éléments particuliers que l'on identifie aux entiers naturels 0 et 1, une addition et une multiplication que l'on identifie à l'addition et la multiplication sur $\mathbb{N}$, et une relation d'ordre que l'on identifie à la relation d'ordre sur $\mathbb{N}$.

**Définition 8.2.33 (Extension élémentaire, sous-structure élémentaire)**
Pour toutes $L$-structures $\mathscr{M}$ et $\mathscr{N}$, on dit que $\mathscr{N}$ est une *extension élémentaire* de $\mathscr{M}$, ou que $\mathscr{M}$ est une *sous-structure élémentaire* de $\mathscr{N}$, ce que je noterai

$$\mathscr{M} \sqsubset \mathscr{N}$$

lorsque l'univers de $\mathscr{M}$ est un sous-ensemble de celui de $\mathscr{N}$, et que l'injection canonique $M \xrightarrow{\text{id}} N$ est un plongement élémentaire, autrement dit lorsque pour toute formule $\mathscr{F}[x_1, \dots, x_n]$ et pour tout $(a_1, \dots, a_n) \in M^n$

$$\mathscr{M} \vDash \mathscr{F}[a_1, \dots, a_n] \quad \text{si et seulement si} \quad \mathscr{N} \vDash \mathscr{F}[a_1, \dots, a_n]$$

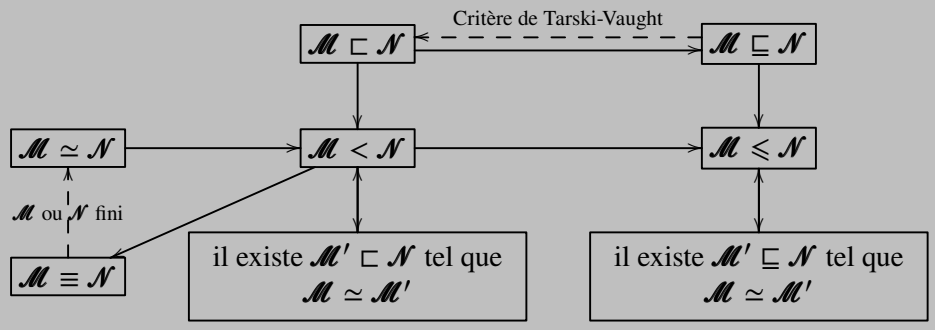
**Remarque 8.2.34 :** Si $\mathscr{M}$ est une sous-structure élémentaire de $\mathscr{N}$, alors $\mathscr{M}$ est une sous-structure de $\mathscr{N}$ (puisqu'un plongement élémentaire est un plongement), et il existe un plongement élémentaire de $\mathscr{M}$ dans $\mathscr{N}$ donc les structures $\mathscr{M}$ et $\mathscr{N}$ sont élémentairement équivalentes (si $\mathscr{M} \sqsubset \mathscr{N}$ alors $\mathscr{M} \sqsubseteq \mathscr{N}$, $\mathscr{M} < \mathscr{N}$ et $\mathscr{M} \equiv \mathscr{N}$).

**Remarque 8.2.35 :** La relation $\sqsubset$ est une relation d'ordre sur toute classe de structures de même signature, car elle est réflexive (l'identité est un plongement élémentaire), transitive (la composée de deux plongements élémentaires en est un, donc si $\mathscr{M}_1 \sqsubset \mathscr{M}_2$ et $\mathscr{M}_2 \sqsubset \mathscr{M}_3$ alors l'injection canonique $M_1 \xrightarrow{\text{id}} M_3$ est un plongement élémentaire), et antisymétrique (si $\mathscr{M} \sqsubset \mathscr{N}$ et $\mathscr{N} \sqsubset \mathscr{M}$, alors on a aussi $\mathscr{M} \sqsubseteq \mathscr{N}$ et $\mathscr{N} \sqsubseteq \mathscr{M}$, donc $\mathscr{M} = \mathscr{N}$).

**Théorème 8.2.36**
On considère trois $L$-structures $\mathscr{M}_1, \mathscr{M}_2, \mathscr{M}_3$. Si $\mathscr{M}_1$ et $\mathscr{M}_2$ sont des sous-structures élémentaires de $\mathscr{M}_3$ telles que $M_1 \subseteq M_2$, alors $\mathscr{M}_1$ est une sous-structure élémentaire de $\mathscr{M}_2$.

**Ces pages ne sont pas incluses dans l’aperçu.**

**Synthèse 8.2.47 (Plongement, sous-structure, équivalence élémentaire, isomorphisme)**
Pour toutes $L$-structures $\mathscr{M}$ et $\mathscr{N}$ :




Légende : une flèche indique une conséquence, une flèche en pointillés une conséquence dépendant d'une condition supplémentaire.

## 8.3 Structures de signature différente : expansions de structures, extensions de théories, diagrammes et diagrammes élémentaires

**Définition 8.3.1 (Enrichissement, restriction, d'un langage)**
1. On dit que le langage $L'$ *enrichit* le langage $L$, ou que $L$ est une *restriction* de $L'$, lorsque tous les symboles de $L$ (de constantes, opérations et prédicats) appartiennent aussi à $L'$.
2. On considère un langage $L'$ qui enrichit un langage $L$, $\mathscr{M}$ une $L$-structure et $\mathscr{M}'$ une $L'$-structure. On dit que $\mathscr{M}'$ est une *expansion* de $\mathscr{M}$, ou que $\mathscr{M}$ est une *restriction* de $\mathscr{M}'$, lorsque les univers de base sont identiques, et que chaque symbole de $L$ a la même interprétation dans $\mathscr{M}$ et $\mathscr{M}'$.

La véracité d'une formule dans une structure et une assignation donnée ne dépend que de la restriction de l'interprétation aux symboles du langage apparaissant dans la formule : si le langage $L'$ enrichit le langage $L$, si $\mathscr{M}$ est une $L$-structure et $\mathscr{M}'$ une $L'$-structure qui enrichit $\mathscr{M}$ (donc de même univers $M$), alors pour tout terme $t$ et toute formule $\mathscr{F}$ du langage $L$, et tout $(a_1, \dots, a_n) \in M^n$

$$t_{\mathscr{M}}[a_1, \dots, a_n] = t_{\mathscr{M}'}[a_1, \dots, a_n]$$
$$\mathscr{M} \vDash \mathscr{F}[a_1, \dots, a_n] \quad \text{si et seulement si} \quad \mathscr{M}' \vDash \mathscr{F}[a_1, \dots, a_n]$$

**Définition 8.3.2 (Extension conservative)**
On considère une théorie $\mathscr{T}$ dans un langage $L$.
1. On appelle *extension* de $\mathscr{T}$ toute théorie $\mathscr{T}'$, dans un langage $L'$ qui enrichit $L$, telle que toute formule de $\mathscr{T}$ est aussi une formule de $\mathscr{T}'$.
2. On appelle *extension conservative* (ou *extension conservatrice*) de $\mathscr{T}$ toute extension telle que tout théorème de $\mathscr{T}'$ exprimé dans le langage $L$ est aussi un théorème de $\mathscr{T}$, autrement dit pour toute $L$-formule $\mathscr{F}$

$$\text{si} \quad \mathscr{T}' \vdash \mathscr{F} \quad \text{alors} \quad \mathscr{T} \vdash \mathscr{F}$$

**Ces pages ne sont pas incluses dans l’aperçu.**

## 8.4 Filtres et ultrafiltres

Nous abordons maintenant la notion de filtre, qui a diverses applications en mathématiques, et qui sera notamment utilisée dans la section 8.6 pour une preuve du théorème de compacité (différente de celle donnée dans le volume 1).

**Définition 8.4.1 (Filtre)**
On appelle *filtre*, sur un ensemble $E$, tout sous-ensemble non vide de $\mathscr{P}(E) \setminus \{\varnothing\}$, stable par intersection et tel que toute partie de $E$ incluant un élément de $\mathscr{F}$ appartient à $\mathscr{F}$, autrement dit on appelle filtre tout ensemble non vide $\mathscr{F}$ de parties de $E$ vérifiant :

- $\varnothing \notin \mathscr{F}$.
- Si $X \in \mathscr{F}$ et $Y \in \mathscr{F}$, alors $X \cap Y \in \mathscr{F}$.
- Si $X \in \mathscr{F}$ et $X \subseteq Y \subseteq E$, alors $Y \in \mathscr{F}$.

**Remarque 8.4.2 :** La condition $\varnothing \notin \mathscr{F}$ n'est parfois pas incluse dans la définition, et un filtre $\mathscr{F}$ tel que $\varnothing \notin \mathscr{F}$ s'appelle alors un *filtre propre*. Avec cette définition $\mathscr{P}(E)$ est un filtre, et c'est le seul qui ne soit pas un filtre propre, car tout sous-ensemble $X$ de $E$ inclut l'ensemble vide, donc si $\varnothing \in \mathscr{F}$ alors $X \in \mathscr{F}$.

**Remarque 8.4.3 :** On considère un filtre $\mathscr{F}$ sur $E$.

1. $E \in \mathscr{F}$ car tout élément $X$ de $\mathscr{F}$ (qui est non vide par définition) est tel que $X \subseteq E$. Ainsi, on obtient une définition équivalente d'un filtre en remplaçant la condition $\mathscr{F} \neq \varnothing$ par $E \in \mathscr{F}$.
2. Puisque un filtre est stable par intersection, on en déduit par récurrence que si $X_1, \dots, X_n$ sont des éléments de $\mathscr{F}$, alors
$$X_1 \cap \cdots \cap X_n \in \mathscr{F}$$
Ainsi, on obtient une définition équivalente d'un filtre en remplaçant la condition « $\mathscr{F}$ est stable par intersection », par la condition « $\mathscr{F}$ est stable par intersection finie ».
3. On en déduit que si $X_1, \dots, X_n$ sont des éléments de $\mathscr{F}$, et si $Y$ est une partie de $E$ telle que
$$X_1 \cap \cdots \cap X_n \subseteq Y$$
alors $Y \in \mathscr{F}$ (puisque $Y$ inclut un élément de $\mathscr{F}$).

**Remarque 8.4.4 :** D'une certaine manière, un filtre distingue les parties « assez grosses » d'un ensemble (celles qui appartiennent au filtre), comme l'ensemble $E$ en entier, de celles qui ne le sont pas (celles qui n'appartiennent pas au filtre), comme l'ensemble vide, dans la mesure où tout ce qui est « plus gros » qu'un élément du filtre reste dans le filtre (si $X \in \mathscr{F}$ et si $Y \supseteq X$ alors $Y \in \mathscr{F}$). Le sens précis que l'on donne à l'expression « assez grosses » dépendra du filtre considéré (il existe plusieurs types de filtres, comme on pourra le constater dans les exemples qui suivent).

**Exemple 8.4.5 (Exemples de filtres)**
1. Si $A$ est un sous-ensemble non vide de $E$, alors l'ensemble des parties de $E$ incluant $A$ est un filtre sur $E$, appelé *filtre principal engendré par A* (voir la définition 8.4.7 ci-dessous).
2. Si $E$ est un ensemble infini, alors l'ensemble des parties de $E$ dont le complémentaire est un ensemble fini, est un filtre sur $E$ appelé *filtre de Fréchet* (voir la définition 8.4.13 ci-dessous).

3. Si $E$ est un espace topologique et $a$ un élément de $E$, alors l'ensemble des voisinages de $a$ est un filtre sur $E$.

**Théorème 8.4.6 (Propriétés élémentaires d'un filtre)**
On considère un filtre $\mathcal{F}$ sur $E$.

1. Pour toutes parties $X_1, \dots, X_n$ de $E$
$$X_1 \cap \cdots \cap X_n \in \mathcal{F} \iff X_1 \in \mathcal{F} \text{ et } \dots \text{ et } X_n \in \mathcal{F}$$
2. Si $X \in \mathcal{F}$ alors $\complement_E X \notin \mathcal{F}$.
3. Pour tout $X \in \mathcal{F}$ et $Y \subseteq E$, si $X \cap Y = \varnothing$ alors $\complement_E Y \in \mathcal{F}$ (et par contraposition, si $\complement_E Y \notin \mathcal{F}$ alors $X \cap Y \neq \varnothing$).

**Preuve**
1. Un filtre est stable par intersection finie, et la réciproque est vraie car $X_1 \cap \cdots \cap X_n$ est inclus dans chaque $X_i$ donc si cette intersection est un élément de $\mathcal{F}$, alors $X_1, \dots, X_n$ sont des parties de $E$ qui incluent un élément de $\mathcal{F}$, donc sont aussi des éléments de $\mathcal{F}$.
2. Si $X \in \mathcal{F}$ alors $\complement_E X \notin \mathcal{F}$, car sinon l'intersection de $X$ et $\complement_E X$ (l'ensemble vide) serait dans $\mathcal{F}$, en contradiction avec $\varnothing \notin \mathcal{F}$.
3. On considère $X \in \mathcal{F}$ et $Y \subseteq E$. Si $X \cap Y = \varnothing$ alors $X \subseteq \complement_E Y$, donc $\complement_E Y \in \mathcal{F}$.

**Définition 8.4.7 (Filtre principal)**
Si $A$ est un sous-ensemble non vide de $E$, alors l'ensemble des parties de $E$ incluant $A$ est un filtre sur $E$, appelé *filtre principal engendré par $A$*, que je noterai $\mathcal{F}_A$ :
$$\mathcal{F}_A \overset{\text{def}}{\equiv} \{X \subseteq E \mid A \subseteq X\}$$

**Preuve**
Vérifions que $\mathcal{F}_A$ est un filtre :
- $\mathcal{F}_A \neq \varnothing$ puisque $A \in \mathcal{F}_A$, et $\varnothing \notin \mathcal{F}$ puisque $A \neq \varnothing$.
- Si $X \in \mathcal{F}_A$ et $Y \in \mathcal{F}_A$, alors $A \subseteq X \cap Y$, donc $X \cap Y \in \mathcal{F}_A$.
- Si $X \in \mathcal{F}_A$ et $X \subseteq Y$, alors $A \subseteq X \subseteq Y$, donc $Y \in \mathcal{F}_A$.

**Remarque 8.4.8 :** $A \in \mathcal{F}_A$ car $A \subseteq A$.

**Remarque 8.4.9 (Notations) :** Je pourrai noter $\mathcal{F}_a$, plutôt que $\mathcal{F}_{\{a\}}$, le filtre principal engendré par le singleton $\{a\}$. On notera par ailleurs, puisque $\{a\} \subseteq X$ si et seulement si $a \in X$, que l'on a
$$\mathcal{F}_a = \{X \subseteq E \mid a \in X\}$$

**Définition 8.4.10 (Sous-ensemble cofini)**
On dit qu'un sous-ensemble est *cofini* lorsque son complémentaire est fini.

**Ces pages ne sont pas incluses dans l’aperçu.**

**Remarque 8.5.9 :** On peut noter que dans tout ce qui précède, la spécificité des ultrafiltres (par rapport aux filtres) n'a été utilisée que dans l'étape inductive du théorème précédent faisant intervenir la négation.

**Théorème 8.5.10 (Corollaire)**
Pour toute $L$-structure $\mathscr{M}$ et tout ultrafiltre $\mathscr{U}$ sur un ensemble $I$, la fonction

$$f : \begin{cases} M \longrightarrow M^I/\mathscr{U} \\ m \longmapsto \langle x \longmapsto m \rangle \end{cases}$$

est un plongement élémentaire de $\mathscr{M}$ dans $\mathscr{M}^I/\mathscr{U}$.

**Preuve**
Notons $\mathscr{N}$ la structure $\mathscr{M}^I/\mathscr{U}$. D'après le théorème de Łoś, pour toute formule $\mathscr{F}[x_1, \dots, x_n]$ et tout $(a_1, \dots, a_n) \in M^n$

$$\mathscr{N} \vDash \mathscr{F}[\langle x \longmapsto a_1 \rangle, \dots, \langle x \longmapsto a_n \rangle] \qquad \text{si et seulement si} \qquad \{x \in I \mid \mathscr{M} \vDash \mathscr{F}[a_1, \dots, a_n]\} \in \mathscr{U}$$

Or

$$\{x \in I \mid \mathscr{M} \vDash \mathscr{F}[a_1, \dots, a_n]\} = \begin{cases} I \in \mathscr{U} & \text{si } \mathscr{M} \vDash \mathscr{F}[a_1, \dots, a_n] \\ \varnothing \notin \mathscr{U} & \text{si } \mathscr{M} \nvDash \mathscr{F}[a_1, \dots, a_n] \end{cases}$$

Donc

$$\mathscr{M} \vDash \mathscr{F}[a_1, \dots, a_n] \qquad \text{si et seulement si} \qquad \mathscr{N} \vDash \mathscr{F}[\langle x \longmapsto a_1 \rangle, \dots, \langle x \longmapsto a_n \rangle]$$

et par conséquent $f$ est un plongement élémentaire.

## 8.6 Théorème de compacité et applications

Nous avons vu dans le volume 1 le théorème de compacité de la logique des prédicats, qui affirme qu'une théorie admet un modèle si et seulement si tout sous-ensemble fini de cette théorie en admet un, ou de façon équivalente qu'une formule est conséquence sémantique d'une théorie si et seulement si il existe un sous-ensemble fini de cette théorie dont elle est une conséquence sémantique. Ce théorème a été prouvé à l'aide des théorèmes de complétude et de correction, qui font intervenir à la fois la sémantique et un système de déduction formel. Mais l'énoncé du théorème de compacité ne faisant pas appel à un tel système de déduction, on s'attend à ce qu'on puisse le démontrer par des méthodes purement sémantiques. C'est en effet le cas, ce théorème étant une conséquence du théorème de Łoś :

**Métathéorème 8.6.1 (Théorème de compacité de la logique des prédicats)**
Toute théorie $\Gamma$ vérifie les deux propriétés (équivalentes) suivantes :

1. $\Gamma$ admet un modèle si et seulement si tout sous-ensemble fini de $\Gamma$ admet un modèle.
2. Pour toute formule $\mathscr{F}$, $\Gamma \vDash \mathscr{F}$ si et seulement si il existe un sous-ensemble fini $\Gamma'$ de $\Gamma$ tel que $\Gamma' \vDash \mathscr{F}$.

**Preuve 1 (à l'aide des théorèmes de correction et de complétude)**
La preuve a été donnée dans la section 7.4 du volume 1. Je la rappelle brièvement :

- L'une des implications du « si et seulement si » de chacune des deux versions est immédiate : si $\Gamma$ admet un modèle, alors tout sous-ensemble fini admet ce même modèle, et s'il existe un sous-ensemble fini $\Gamma'$ de $\Gamma$ tel que $\Gamma' \vDash \mathscr{F}$, alors $\Gamma \vDash \mathscr{F}$ car tout modèle de $\Gamma$ est a fortiori un modèle de $\Gamma'$.
- L'équivalence des deux autres implications est la suivante :
  — Faisons l'hypothèse du premier point, et considérons $\Gamma \vDash \mathscr{F}$. Alors la clôture universelle $\mathscr{F}'$ de $\mathscr{F}$ est aussi telle que

**Ces pages ne sont pas incluses dans l’aperçu.**

et la théorie

$$\Delta' :\equiv \Delta \cup \left\{\mathscr{G}_k \mid k \in \mathbb{N}\right\}$$

Tout sous-ensemble fini de $\Delta'$ admet $(\mathbb{N}, \leqslant)$ comme modèle : cette structure est en effet un modèle de $\Delta$ (il s'agit d'un ensemble bien ordonné), et si $n$ est le plus grand des indices $k$ des formules $\mathscr{G}_k$ apparaissant dans $\Delta'$, alors on peut par exemple interpréter dans $\mathbb{N}$ chaque $c_k$ par $n-k$. On en déduit que $\Delta'$ admet un modèle $\mathscr{M}$, pour lequel l'ensemble $\left\{c_k^{\mathscr{M}} \mid k \in \mathbb{N}\right\}$ est une partie non vide de $M$ sans plus petit élément (tout élément $c_k^{\mathscr{M}}$ de cet ensemble est tel que $c_k^{\mathscr{M}} > c_{k+1}^{\mathscr{M}}$), en contradiction avec le fait que $\mathscr{M}$ est un ensemble bien ordonné.

## 8.7 Théorèmes de Löwenheim-Skolem

**Théorème 8.7.1 (Théorèmes de Löwenheim-Skolem)**

1. *Théorème de Löwenheim-Skolem descendant* : Pour toute $L$-structure infinie $\mathscr{N}$, pour tout cardinal $\kappa$ et tout sous-ensemble $A$ de $N$ tels que

   $$\max(|L|, |A|) \leqslant \kappa$$

   il existe une sous-structure élémentaire $\mathscr{M}$ de $\mathscr{N}$ telle que

   $$\begin{cases} A \subseteq M \\ |\mathscr{M}| \leqslant \kappa \end{cases}$$

2. *Théorème de Löwenheim-Skolem ascendant* : Pour toute $L$-structure infinie $\mathscr{M}$ et tout cardinal $\kappa$, il existe une extension élémentaire $\mathscr{M}'$ de $\mathscr{M}$ telle que $\left|\mathscr{M}'\right| \geqslant \kappa$.

**Preuve**

1. Preuve du Théorème de Löwenheim-Skolem descendant : construisons une famille croissante (pour l'inclusion) $(M_n)_{n\in\mathbb{N}}$ de parties de $N$ incluant $A$, dont la réunion sera un sous-univers de $\mathscr{N}$ répondant à la question. L'idée est de partir de $A$ et d'ajouter assez d'éléments pour que le sous-ensemble de $N$ obtenu vérifie le critère de Tarski-Vaught.

   - On choisit donc

     $$M_0 :\equiv A$$

     et on construit $M_{n+1}$ à partir de $M_n$ de la façon suivante : pour tout $(a_1, \dots, a_p) \in M_n^p$ et pour toute formule $\mathscr{F}[x, x_1, \dots, x_p]$ telle que

     $$\mathscr{N} \vDash \exists x \mathscr{F}[a_1, \dots, a_p]$$

     il existe $a \in N$ tel que

     $$\mathscr{N} \vDash \mathscr{F}[a, a_1, \dots, a_p]$$

     On construit $M_{n+1}$ en ajoutant à $M_n$ tous les éléments $a$ ainsi obtenus. Formellement, notons $\Gamma$ l'ensemble de toutes les formules et

     $$X :\equiv \left\{((a_1, \dots, a_p), \mathscr{F}) \in \left(\bigcup_{k\in\mathbb{N}^*} M_n^k\right) \times \Gamma \mid \mathscr{F}[x, x_1, \dots, x_p] \text{ et } \mathscr{N} \vDash \exists x \mathscr{F}[a_1, \dots, a_p]\right\}$$

     D'après l'axiome du choix, il existe une fonction

     $$\varphi \in \prod_{((a_1,\dots,a_p),\mathscr{F})\in X} \left\{a \in N \mid \mathscr{N} \vDash \mathscr{F}[a, a_1, \dots, a_p]\right\}$$

     et l'ensemble $M_{n+1}$ est défini tel que

     $$M_{n+1} = M_n \cup \mathrm{Im}(\varphi)$$

     Notons

     $$M :\equiv \bigcup_{n\in\mathbb{N}} M_n$$

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 9

# Introduction à la calculabilité

## 9.1 Modélisation des algorithmes et thèse de Church-Turing

Si l'on veut décrire de manière informelle ce qu'est un algorithme, on peut par exemple donner la « définition » suivante : un *algorithme* est une procédure générale, sous la forme d'une suite finie d'instructions systématiques, permettant de résoudre par un calcul, en un nombre fini d'étapes, un problème donné.

Le terme *algorithme* vient du nom du mathématicien perse Al-Khwarizmi (v.780–v.850), dont le nom signifie littéralement « natif de Khwarezm » (ancien nom de la ville de Khiva, en Ouzbékistan), latinisé en *Algoritmi*. Il a écrit un traité sur le système de numération décimal de position (né en Inde probablement vers le Ier siècle), traité traduit en latin au XIIe siècle par Robert de Chester, sous le nom *Algoritmi de numero indorum* [Al-Khwarizmi sur les nombres indiens]. *Algoritmi* a donné en latin médiéval *algorismus*, puis en vieux français *algorisme* (XIIIe siècle), signifiant alors ce système de numération indien. C'est en grande partie pour cette raison que l'on désigne généralement nos chiffres actuels sous le nom de *chiffres arabes* (*chiffres indiens*, ou *chiffres indo-arabes*, serait probablement plus correct). *Algorisme* est ensuite devenu *algorithme*, sous l'influence de la racine grecque *arithmos* (nombre), que l'on trouve aussi dans *arithmétique*. Le sens évoluera plus tard pour désigner non plus seulement les règles des calculs arithmétiques liées à ce système de numération, mais n'importe quelle règle de calcul. Gottfried Wilhelm Leibniz écrit en 1684 :

> « De cette règle, connue sous le nom d'*algorithme*, pour ainsi dire, de ce calcul, que j'appelle *différentiel*, toutes les autres équations différentielles peuvent être trouvées au moyen d'un calcul général, et les maxima et minima, ainsi que les tangentes [peuvent être] obtenus [...] » [1]

L'usage du mot *algorithme* s'est répandu essentiellement au XXe siècle, en particulier avec l'essor de l'informatique. [2]

**Exemple 9.1.1 (Exemples d'algorithmes)**
1. L'addition, la soustraction, la multiplication, la division euclidienne de deux nombres à partir de

1. Gottfried Wilhelm LEIBNIZ. « Nova Methodvs pro maximis et minimis, itemque tangentibus, quae nec fractas, nec irrationales quantitates moratur, & singulare pro illis calculi genus *[Une nouvelle méthode pour les maxima et minima, ainsi que pour les tangentes, qui n'est pas entravée par des quantités fractionnaires ou irrationnelles]* ». Dans : *Acta Eruditorum* 3 (1684), p. 467-473 (p.469).

2. Sources du paragraphe :
Jeff MILLER. *Earliest Uses of Some Words of Mathematics*. URL : `https://mathshistory.st-andrews.ac.uk/Miller/mathword/`
Douglas HARPER. *Online etymology dictionary*. 2017. URL : `http://www.etymonline.com/`
*MacTutor History of Mathematics archive*. URL : `http://www-history.mcs.st-and.ac.uk/`

leur écriture décimale, par exemple

$$\begin{array}{cccc} & {\scriptstyle 1} & {\scriptstyle 1} & \\ & 5 & 3 & 4 \\ + & 7 & 9 & 6 \\ \hline 1 & 3 & 3 & 0 \end{array}$$

2. Les tests de primalité, pour savoir si un nombre entier $n$ est premier, comme la méthode simple consistant à tester la divisibilité de $n$ par tous les entiers compris entre 2 et $n-1$ (il existe des méthodes plus efficaces). Le crible d'Ératosthène (voir la section 4.10 du volume 2) est aussi une méthode algorithmique pour déterminer tous les nombres premiers inférieurs à un entier donné.
3. L'algorithme d'Euclide (voir la section 3.4), qui permet de calculer le pgcd de deux nombres.
4. Les algorithmes de résolution de casse-têtes comme le Rubik's Cube.
5. Les algorithmes de tri, pour classer un ensemble d'objets selon un ordre donné.

Un algorithme consiste donc essentiellement à effectuer un calcul, à partir de données initiales, permettant d'obtenir un certain résultat (par exemple calculer, à partir de $a$ et de $b$, la somme $a+b$ ou le pgcd de $a$ et $b$). Les situations qui ne semblent pas être des calculs directs peuvent néanmoins s'y ramener. On trouve par exemple, parmi les problèmes traités par des algorithmes, des problèmes dits *de décision* pour lesquels il faut répondre à une question par oui ou non, c'est-à-dire qu'il s'agit de savoir si les données initiales vérifient une certaine propriété ou pas (par exemple déterminer si un nombre entier est premier). On se ramène à un calcul à l'aide des fonctions indicatrices : si on se place dans un ensemble $U$ représentant l'univers des données possibles, déterminer si une de ces données vérifie une propriété $P$, c'est déterminer si un élément de $U$ appartient au sous-ensemble $S$ des éléments de $U$ vérifiant $P$, ce qui revient à calculer la fonction indicatrice de $S$ dans $U$ :

$$\chi_S : \begin{cases} U \longrightarrow \{0,1\} \\ x \longmapsto \begin{cases} 1 & \text{si } x \in S \\ 0 & \text{sinon} \end{cases} \end{cases}$$

Pendant longtemps, les mathématiciens n'ont pas formalisé la notion d'algorithme, tout le monde s'accordant sur ce qu'est, intuitivement, un algorithme. Si cela ne pose pas de problème quand il s'agit d'en construire un effectuant certaines opérations, c'est en revanche plus problématique quand on doit démontrer qu'il ne peut en exister aucun permettant de résoudre un certain problème. Pour cela, il est indispensable d'avoir une définition formelle de ce qu'est un algorithme. Voyons par exemple deux grands problèmes classiques :

1. Le problème de la décision de la logique du premier ordre, qu'on appelle aussi *Entscheidungsproblem* (terme allemand), a été posé par les mathématiciens allemands David Hilbert (1862–1943) et Wilhelm Ackermann (1896–1962) en 1928. Il consiste à déterminer si une formule en logique du premier ordre est valide (vraie dans tous les modèles) ou non.
2. Le dixième problème de Hilbert : il consiste à déterminer si une équation diophantienne [3], à une ou plusieurs inconnues, admet une solution. Ce problème fait partie d'une liste de 23, appelés *problèmes de Hilbert*, présentés par David Hilbert lors du congrès international des mathématiciens à Paris en 1900. Ces problèmes tenaient alors les mathématiciens en échec et devaient, selon lui, contribuer au développement futur des mathématiques.

Après la formalisation de la notion d'algorithme et de calcul, il a été démontré que ces deux problèmes étaient insolubles : il n'existe aucun algorithme permettant de répondre à la question posée. Ces résultats ont

3. Équation dont les coefficients sont des entiers relatifs.

été démontrés

1. pour le premier problème, en 1936 et de façon indépendante, par les mathématiciens et logiciens américains Alonzo Church (1903–1995) et Alan Turing (1912–1954) ;
2. pour le deuxième problème, par le mathématicien russe Youri Matiassevitch [4](1947–) en 1970, en s'appuyant sur les travaux des mathématiciens américains Martin Davis (1928–) et Julia Robinson (1919–1985), et du philosophe et mathématicien américain Hilary Putnam (1926–2016).

Pour modéliser les opérations que peut effectuer un algorithme, plusieurs modèles de calculs ont été proposés de façon indépendante au début des années 1930 :

1. les fonctions récursives, à partir des travaux du logicien et mathématicien austro-américain Kurt Gödel (1906–1978) et du mathématicien et logicien français Jacques Herbrand (1908–1931) ; les fonctions définies dans le système formel originel de Gödel et Herbrand sont aussi celles que l'on peut *représenter* (dans un sens qui sera précisé dans la section 9.13) dans l'arithmétique de Robinson (c'est-à-dire la théorie de l'arithmétique de Peano sans le schéma de récurrence) ou toute extension de cette théorie ; un peu plus tard, les travaux du mathématicien et logicien américain Stephen Cole Kleene (1909–1994), élève de Church, ont conduit à construire l'ensemble des fonctions récursives par induction à partir de quelques fonctions de base et de trois règles (composition, récurrence, et un opérateur de minimisation introduit par Kleene) ; c'est en général sous cette forme que l'on étudie de nos jours les fonctions récursives (voir les sections 9.2 à 9.11) ;
2. les machines de Turing, créées par Alan Turing (voir la section 9.16).
3. le lambda-calcul, créé par Alonzo Church, qui est un système formel que l'on pourrait résumer de façon informelle par l'expression « tout est fonction » (voir la section 9.17 pour une brève introduction au lambda-calcul) ;

Ces trois modèles sont fondés sur des principes très différents mais pourtant ils permettent d'effectuer exactement les mêmes calculs. Cette *coïncidence* est l'une des raisons qui fondent ce qu'on appelle la *thèse de Church* (ou plus correctement la *thèse de Church-Turing*), qui affirme que le concept d'algorithme est complètement caractérisé par ces systèmes, et que tout calcul effectué par un algorithme (dans le sens intuitif) peut l'être par chacun des systèmes précédents. Il s'agit d'une thèse dans le sens philosophique du terme, c'est-à-dire d'une affirmation à l'égard d'un sujet. Elle ne peut pas être démontrée, car sa formulation n'est pas formelle, mais elle est de nos jours largement acceptée. Depuis, d'autres modèles de calculs ont été proposés, qui sont équivalents aux précédents, et permettent tous de calculer les mêmes fonctions, et de réaliser théoriquement tous les algorithmes connus (on dit qu'ils sont Turing-complet). On peut citer par exemple

1. les machines à registres illimités ;
2. les automates cellulaires ;
3. les langages de programmation usuels (Python, JavaScript, Java, C#, C++, C, R…).

**Remarque 9.1.2 (Remarque historique) :** Au début des années 1930, Alonzo Church développe le lambda-calcul, assisté de deux de ses élèves, Stephen Cole Kleene et le mathématicien et logicien américain John Barkley Rosser (1907–1989). Church a conçu à l'origine le lambda-calcul dans l'espoir de développer un système logique adéquat pour les mathématiques, et dans lequel la notion de fonction est primordiale [5] ; il réalise qu'il peut représenter à l'intérieur de son système les nombres entiers, et de nombreuses fonctions définies sur les nombres entiers (on dit que ces fonctions sont $\lambda$-définissables).

À la même période, Gödel définit une classe de fonctions, en essayant d'expliciter la notion de récursivité (les fonctions définies de façon récursive, c'est-à-dire en faisant appel à elles-mêmes, généralisent d'une

4. Ou Yuri Matiyasevich.

5. Le lambda-calcul est en fait une partie du système originel envisagé par Church, Kleene et Rosser ayant prouvé que ce système complet était contradictoire.

certaine façon les définitions par récurrence classiques). Il s'appuie sur ses propres travaux (notamment son célèbre article de 1931, où il énonce ses deux théorèmes d'incomplétude, et définit ce qu'on appelle maintenant des fonctions primitives récursives), ainsi que sur une suggestion de Herbrand.

Début 1934, Church pense qu'une fonction est calculable, dans le sens intuitif du terme, si et seulement si elle peut être définie dans son système formel (il considère donc qu'il y a équivalence entre fonction calculable et fonction $\lambda$-définissable). Il propose à Gödel de définir ainsi les fonctions calculables, mais Gödel n'est pas du tout convaincu. En 1935, Church modifie sa proposition et suggère de définir les fonctions calculables comme étant les fonctions récursives d'Herbrand et Gödel. Cette proposition est annoncée lors d'une réunion de l'*American Mathematical Society* le 19 avril 1935, et apparaît dans le résumé d'un article en préparation, envoyé fin mars au bulletin de l'AMS [6]. Il écrit notamment que « la notion de fonction d'entiers positifs effectivement calculable, doit être identifiée avec celle de fonction récursive, car d'autres définitions plausibles de la calculabilité effective s'avèrent donner des notions qui sont soit équivalentes soit plus faibles que la récursivité ».

À cette date, Kleene et lui ont prouvé (indépendamment) que toutes les fonctions $\lambda$-définissables sont récursives. En juin 1935, Kleene prouve que toutes les fonctions récursives sont $\lambda$-définissables, justifiant ainsi l'équivalence de ces deux classes de fonctions. L'article définitif de Church paraît en 1936 [7]. Il y définit donc les fonctions calculables comme étant les fonctions récursives d'Herbrand et Gödel (ou de façon équivalente les fonctions $\lambda$-définissables). L'équivalence de ces deux classes de fonctions est d'ailleurs un point important de sa justification :

> « Le fait, cependant, que deux définitions de la calculabilité effective aussi différentes et (de l'avis de l'auteur) aussi naturelles, s'avèrent équivalentes, renforce les raisons invoquées ci-dessous pour croire qu'elles constituent une caractérisation générale de cette notion conforme à la compréhension intuitive habituelle de celle-ci. »

On peut noter qu'il s'agit bien pour lui de *définir* la calculabilité ; ce n'est que plus tard qu'on parlera de la « thèse de Church », à la suite de Kleene [8]. Néanmoins Gödel n'est toujours pas convaincu par les arguments de Church, assez ironiquement, pourrait-on dire, puisque Church utilise la classe de fonctions que Gödel a lui-même définie ; à cette époque, ce dernier pense que les fonctions calculables par une procédure finie sont les fonctions qui peuvent être définies par des méthodes récursives, mais il ne croit pas que sa définition couvre toutes les fonctions pouvant être construites par de telles méthodes.

En 1935 et 1936, Turing travaille sur son propre système, indépendamment des travaux de Church et Gödel. Ce n'est qu'au moment de publier ses résultats qu'il prend connaissance du texte de Church. Il ajoute à son article la preuve de l'équivalence de son système avec le lambda-calcul. Dans cet article, Turing justifie que tout algorithme, dans le sens informel du terme, peut être réalisé par une machine de Turing. Ses arguments sont extrêmement convaincants et emporteront immédiatement l'adhésion, notamment de Gödel. [9]

6. Alonzo CHURCH. « An Unsolvable Problem of Elementary Number Theory. Preliminary report ». Dans : *Bulletin of the American Mathematical Society* 41.5 (1935), p. 332-333.

7. Alonzo CHURCH. « An Unsolvable Problem of Elementary Number Theory ». Dans : *American Journal of Mathematics* 58.2 (avril 1936), p. 345-363.

8. Stephen Cole KLEENE. « Recursive predicates and quantifiers ». Dans : *Transactions of the American Mathematical Society* 53 (1943), p. 41-73.

9. Source principale : Martin DAVIS. « Why Gödel Didn't Have Church's Thesis ». Dans : *Information and Control* 54 (juillet-août 1982), p. 3-24.

## 9.2 Fonctions primitives récursives

Les fonctions étudiées dans cette section seront des fonctions de $\mathbb{N}^p$ dans $\mathbb{N}$, avec $p$ entier naturel non nul (ce qui ne sera pas précisé à chaque fois). Comme je l'ai déjà fait, je pourrai noter $\vec{x}^p$ ou $\vec{x}$ la liste de variables $x_1, \dots, x_p$.

**Définition 9.2.1 (Composée)**
Pour toutes les fonctions $\mathbb{N}^p \xrightarrow{f_1} \mathbb{N}$ , …, $\mathbb{N}^p \xrightarrow{f_n} \mathbb{N}$ , $\mathbb{N}^n \xrightarrow{g} \mathbb{N}$ , on appelle *fonction composée* de $\overrightarrow{f}$ (ou $f_1, \dots, f_n$) par $g$ la composée de $\langle f_1, \dots, f_n \rangle$ par $g$, c'est-à-dire la fonction

$$g \circ \langle \overrightarrow{f} \rangle : \begin{cases} \mathbb{N}^p \longrightarrow \mathbb{N} \\ \vec{x}^p \longmapsto g(f_1(\vec{x}^p), \dots, f_n(\vec{x}^p)) \end{cases}$$

**Définition 9.2.2 (Récursion primitive)**
Pour toutes fonctions $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ et $\mathbb{N}^{p+2} \xrightarrow{g} \mathbb{N}$ , on dit que la fonction $\mathbb{N}^{p+1} \xrightarrow{h} \mathbb{N}$ est définie par *récursion primitive* à partir de $f$ et de $g$, lorsque pour tout $\vec{x}^p$ et tout entier $n$ :

$$\begin{cases} h(\vec{x}^p, 0) = f(\vec{x}^p) \\ h(\vec{x}^p, n+1) = g(\vec{x}^p, n, h(\vec{x}^p, n)) \end{cases}$$

Je noterai

$$\mathrm{rec}(f, g) \overset{\text{def}}{\equiv} h$$

**Remarque 9.2.3 :** Nous avons vu dans le théorème 3.8.4 du volume 2, et en reprenant ses notations, que pour toutes fonctions $P \xrightarrow{a} A$ et $P \times \mathbb{N} \times A \xrightarrow{f} A$ , il existe une unique fonction $P \times \mathbb{N} \xrightarrow{g} A$ telle que pour tout $p \in P$ et tout $n \in \mathbb{N}$

$$\begin{cases} g(p, 0) = a(p) \\ g(p, n+1) = f(p, n, g(p, n)) \end{cases}$$

La fonction $\mathrm{rec}(f, g)$ est ainsi définie par une telle récurrence avec paramètres, en appliquant ce théorème en prenant $\mathbb{N}^p$ pour $P$ et $\mathbb{N}$ pour $A$. On notera par ailleurs qu'il y a unicité d'une telle fonction.

**Définition 9.2.4 (Fonctions primitives récursives)**
L'ensemble des fonctions *primitives récursives* (en abrégé PR), est défini par induction :

1. À partir de la base composée des fonctions suivantes :
   - la fonction constante égale à 0 :
   $$\begin{cases} \mathbb{N} \longrightarrow \mathbb{N} \\ x \longmapsto 0 \end{cases}$$

- la fonction successeur
$$\mathrm{S} : \begin{cases} \mathbb{N} \longrightarrow \mathbb{N} \\ x \longmapsto x+1 \end{cases}$$
- les $i$-èmes projections des $p$-uplets $(x_1, \dots, x_p)$, pour tout $1 \leqslant i \leqslant p$ :
$$\mathrm{pr}_i^p : \begin{cases} \mathbb{N}^p \longrightarrow \mathbb{N} \\ (x_1, \dots, x_p) \longmapsto x_i \end{cases}$$

2. Avec les deux règles suivantes :
   - Si $\mathbb{N}^p \xrightarrow{f_1} \mathbb{N}$ , …, $\mathbb{N}^p \xrightarrow{f_n} \mathbb{N}$ sont des fonctions PR, et si $\mathbb{N}^n \xrightarrow{g} \mathbb{N}$ est une fonction PR, alors la fonction composée $g \circ \langle \overrightarrow{f} \rangle$ est PR.
   - Si les fonctions $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ $f$ et $\mathbb{N}^{p+2} \xrightarrow{g} \mathbb{N}$ sont PR, alors la fonction $\mathrm{rec}(f, g)$, définie par récursion primitive à partir de $f$ et de $g$, est PR.

**Remarque 9.2.5 (Notations) :** J'ai utilisé l'expression éloquente « $x+1$ » pour noter la fonction successeur, mais je rappelle que cette fonction est une caractéristique intrinsèque de $\mathbb{N}$, qui ne dépend pas de l'addition (opération que l'on peut justement définir par récurrence à l'aide de S) : si l'ensemble $\mathbb{N}$ est défini comme une structure de Peano-Dedekind, la fonction successeur fait partie de la définition, et si $\mathbb{N}$ est défini comme un ensemble bien ordonné avec certaines propriétés, le successeur de $n$ est le plus petit des éléments strictement supérieurs à $n$.

**Remarque 9.2.6 :** Comme d'habitude, quand je dis qu'un certain ensemble est défini par induction, le détail de la construction est implicite (et je ne préciserai plus à chaque fois) : ici, l'ensemble des fonctions primitives récursives est donc défini de façon équivalente, comme

- le plus petit ensemble de fonctions incluant les fonctions de base, et stable par les deux règles (définition inductive de haut en bas) ;
- l'ensemble des fonctions obtenues par application successive (un nombre fini de fois) des deux règles, à partir des fonctions de base (définition inductive de bas en haut).

**Remarque 9.2.7 :** La règle de composition et les projections permettent notamment de permuter les arguments de fonctions PR, ou de modifier le nombre d'arguments. Par exemple si la fonction $\mathbb{N}^2 \xrightarrow{f} \mathbb{N}$ est PR, il en est de même des fonctions

$$g : \begin{cases} \mathbb{N}^2 \longrightarrow \mathbb{N} \\ (x, y) \longmapsto f(y, x) \end{cases} \qquad \text{et} \qquad h : \begin{cases} \mathbb{N} \longrightarrow \mathbb{N} \\ x \longmapsto f(x, x) \end{cases}$$

En effet, $g$ est la composée de $\langle \mathrm{pr}_2^2, \mathrm{pr}_1^2 \rangle$ par $f$, et $h$ est la composée de $\langle \mathrm{pr}_1^1, \mathrm{pr}_1^1 \rangle$ par $f$.

**Exemple 9.2.8 (Identité)**
L'identité
$$\mathrm{id} : \begin{cases} \mathbb{N} \longrightarrow \mathbb{N} \\ x \longmapsto x \end{cases}$$

**Ces pages ne sont pas incluses dans l’aperçu.**

**Exemple 9.3.9 (Relation de divisibilité)**
La relation de divisibilité est PR, car

$$x \mid y \iff \exists z \leqslant y, y = z \times x$$

En effet par définition

$$x \mid y \stackrel{\text{def}}{\equiv} \exists z \in \mathbb{N}, y = z \times x$$

mais cette formulation équivaut à celle où le quantificateur est borné : si $y = zx$, avec $x \geqslant 1$, on a $y = zx \geqslant z$, et si $x = 0$ alors $y = 0$ et on peut choisir $z :\equiv 0 \leqslant y$.

**Exemple 9.3.10 (Nombres premiers)**
Le prédicat « $p$ est un nombre premier » est PR, puisqu'on peut le définir à l'aide d'une quantification bornée et de divers connecteurs logiques, à partir des prédicats PR $p > 1$, $n \mid p$, $n = 1$, $n = p$ :

$$p \in \mathbb{P} \quad \iff \quad (p > 1) \text{ et } (\forall n \leqslant p, (n \mid p \implies (n = 1 \text{ ou } n = p))$$

## 9.4 Définition par cas, minimisation bornée

**Théorème 9.4.1 (Définition par cas)**
Pour toutes les fonctions PR $\mathbb{N}^p \xrightarrow{f_1} \mathbb{N}$, …, $\mathbb{N}^p \xrightarrow{f_n} \mathbb{N}$, et toutes les relations PR d'arité $p$ $P_1, \ldots, P_n$, formant une partition de $\mathbb{N}^p$, la fonction

$$f : \begin{cases} \mathbb{N}^p \longrightarrow \mathbb{N} \\ \vec{x} \longmapsto \begin{cases} f_1(\vec{x}) & \text{si } P_1(\vec{x}) \\ \ldots \\ f_n(\vec{x}) & \text{si } P_n(\vec{x}) \end{cases} \end{cases}$$

est PR.

**Preuve**
Le fait que les relations forment une partition de $\mathbb{N}^p$ permet d'avoir une définition correcte de $f$. Cette fonction est PR, car elle est la somme de fonctions PR : pour tout $\vec{x} \in N^p$

$$f(\vec{x}) = \sum_{i=1}^{n} f_i(\vec{x}) \times \chi_{P_i}(\vec{x})$$

**Remarque 9.4.2 :** En particulier, pour tout prédicat PR $P$ d'arité $p$, et pour toutes les fonctions PR $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ et $\mathbb{N}^p \xrightarrow{g} \mathbb{N}$, la fonction

$$h : \begin{cases} \mathbb{N}^p \longrightarrow \mathbb{N} \\ \vec{x}^p \longmapsto \begin{cases} f(\vec{x}^p) & \text{si } P(\vec{x}^p) \\ g(\vec{x}^p) & \text{sinon} \end{cases} \end{cases}$$

est PR.

**Ces pages ne sont pas incluses dans l’aperçu.**

on voit qu'il suffit de prendre la fonction $g$ suivante, qui est distincte de chaque $f_i$ sur la diagonale :

$$g : \begin{cases} \mathbb{N} \longrightarrow \mathbb{N} \\ n \longmapsto f_n(n) + 1 \end{cases}$$

Cette fonction est calculable (dans le sens intuitif), mais n'est pas PR car pour toute fonction PR $h$, il existe $p \in \mathbb{N}$ tel que $h = f_p$, et

$$g(p) = f_p(p) + 1 \neq f_p(p) = h(p)$$

donc $g \neq h$.

La deuxième justification de l'existence de fonctions calculables non PR consiste à donner un exemple concret d'une telle fonction. L'un des plus classiques est le suivant :

**Définition 9.7.1 (Fonction d'Ackermann-Péter)**
On considère la suite de fonctions $( \mathbb{N} \xrightarrow{A_m} \mathbb{N} )_{m\in\mathbb{N}}$, définie par récurrence de la façon suivante :

- $A_0$ est la fonction successeur ($n \longmapsto n + 1$).
- $A_{m+1}$ est la fonction définie par récurrence, par

$$\begin{cases} A_{m+1}(0) = A_m(1) \\ A_{m+1}(n + 1) = A_m(A_{m+1}(n)) \end{cases}$$

On appelle *fonction d'Ackermann-Péter* la fonction

$$A : \begin{cases} \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N} \\ (m, n) \longmapsto A_m(n) \end{cases}$$

On a donc, pour tout $(m, n) \in \mathbb{N}^2$

$$\begin{cases} A(0, n) = n + 1 \\ A(m + 1, 0) = A(m, 1) \\ A(m + 1, n + 1) = A(m, A(m + 1, n)) \end{cases}$$

**Remarque 9.7.2 (Remarque historique) :** C'est en 1928, dans un article du mathématicien allemand Wilhelm Ackermann (1896–1962), élève de David Hilbert, que l'on trouve le premier exemple publié de fonction calculable qui ne soit pas primitive récursive [11]. Il s'agit de la fonction $\mathbb{N}^3 \xrightarrow{f} \mathbb{N}$ que l'on peut définir par la récurrence suivante :

$$\begin{cases} f(m, n, 0) = m + n \\ f(m, 0, 1) = 0 \\ f(m, 0, 2) = 1 \\ f(m, 0, p) = m \qquad (\text{pour } p > 2) \\ f(m, n + 1, p + 1) = f(m, f(m, n, p + 1), p) \end{cases}$$

11. Wilhelm ACKERMANN. « Zum Hilbertschen Aufbau der reellen Zahlen *[Sur la construction par Hilbert des nombres réels]* ». Dans : *Mathematische Annalen* 99 (1928), p. 118-133.

Des versions à deux arguments de la fonction d'Ackermann ont ensuite été développées. La plus classique est celle que donne la mathématicienne hongroise Rózsa Péter (1905–1977) en 1935, définie par la récurrence

$$\begin{cases} g(0,n) = 2n+1 \\ g(m+1,0) = g(m,1) \\ g(m+1,n+1) = g(m,g(m+1,n)) \end{cases}$$

simplifiée ensuite en 1948 par le mathématicien américain Raphael M. Robinson (1911–1995) pour aboutir à la version de la définition précédente. Notons aussi que dans un article publié en 1927 [12], le mathématicien roumain Gabriel Sudan (1899–1977), autre élève de Hilbert, avait étudié une fonction définie sur des ordinaux (et non pas sur des entiers), dont on peut déduire une fonction calculable non primitive récursive. Mais cela n'a été mis en évidence que vers le milieu des années 1970 [13] par le mathématicien roumain Solomon Marcus (1925–2016), avec l'aide de ses élèves Cristian S. Calude (1952–) et Ionel Tevy, suivant une suggestion de Grigore Moisil (1906–1973) [14]. Ce sont eux qui nomment *fonction de Sudan* cette fonction $\mathbb{N}^3 \xrightarrow{h} \mathbb{N}$, définie par la récurrence suivante :

$$\begin{cases} h(m,n,0) = m+n \\ h(m,0,p+1) = m \\ h(m,n+1,p+1) = h(h(m,n,p+1),h(m,n,p+1)+n+1,p) \end{cases}$$

**Exemple 9.7.3 (Exemple de calcul)**
Supposons que l'on veuille calculer par exemple $A(2,1)$. On a

$$\begin{cases} A(2,1) = A(1,A(2,0)) \\ A(2,0) = A(1,1) = A(0,A(1,0)) = A(0,A(0,1)) = A(0,2) = 3 \end{cases}$$

donc

$$A(2,1) = A(1,A(2,0)) = A(1,3)$$

Il reste à calculer $A(1,3)$ :

$$\begin{cases} A(1,3) = A(0,A(1,2)) \\ A(1,2) = A(0,A(1,1)) \\ A(1,1) = A(0,A(1,0)) = A(0,A(0,1)) = A(0,2) = 3 \end{cases}$$

12. Gabriel SUDAN. « Sur le nombre transfini $\omega^\omega$ ». Dans : *Bulletin Mathématique de la Sociéte Roumaine des Sciences* 30.1 (jan. 1927), p. 11-30.

13. Même Sudan n'en avait pas conscience.

14. Voir :

- Cristian CALUDE et Brânduşa FÂNTÂNEANU. « On Recursive, Non-Primitive Recursive Functions ». Dans : *Bulletin mathématique de la Société des Sciences Mathématiques de la République Socialiste de Roumanie* 22(70).4 (1978), p. 355-358.
- Cristian CALUDE, Solomon MARCUS et Ionel TEVY. « The First Example of a Recursive Function Which is not Primitive Recursive ». Dans : *Historia Mathematica* 6 (1979), p. 380-384.
- Solomon MARCUS. « Grigore C. Moisil: A Life Becoming a Myth ». Dans : *International Journal of Computers, Communications & Control* 1.1 (2006), p. 73-79.

donc

$$\begin{cases} A(1,2) = A(0, A(1,1)) = A(0,3) = 4 \\ A(1,3) = A(0, A(1,2)) = A(0,4) = 5 \end{cases}$$

**Exemple 9.7.4 (Expression de $A_m$ pour des petites valeurs de $m$)**
$A_0$ est une fonction simple (la fonction successeur). Voyons ce qu'il en est des fonctions suivantes :

1. Étude de $A_1$ :
$$\begin{cases} A_1(0) = A_0(1) = 2 \\ A_1(n+1) = A_0(A_1(n)) = A_1(n) + 1 \end{cases}$$
Par conséquent $A_1$ est une suite arithmétique de premier terme 2 et de raison 1 :
$$A_1(n) = n + 2$$

2. Étude de $A_2$ :
$$\begin{cases} A_2(0) = A_1(1) = 3 \\ A_2(n+1) = A_1(A_2(n)) = A_2(n) + 2 \end{cases}$$
Par conséquent $A_2$ est une suite arithmétique de premier terme 3 et de raison 2 :
$$A_2(n) = 2n + 3$$

3. Étude de $A_3$ :
$$\begin{cases} A_3(0) = A_2(1) = 5 \\ A_3(n+1) = A_2(A_3(n)) = 2A_3(n) + 3 \end{cases}$$
Par conséquent $A_3$ est une suite arithmético-géométrique : on obtient après calcul (voir la section 2.3)
$$A_3(n) = 2^n(5+3) - 3 = 2^{n+3} - 3$$

4. Étude de $A_4$ :
$$\begin{cases} A_4(0) = A_3(1) = 13 \\ A_4(n+1) = A_3(A_4(n)) = 2^{A_4(n)+3} - 3 \end{cases}$$
Vérifions que
$$A_4(n) = 2 \uparrow\uparrow (n+3) - 3$$
Par unicité, il suffit de justifier que la suite définie pour tout $n$ par
$$u_n = 2 \uparrow\uparrow (n+3) - 3$$
vérifie la relation de récurrence : on a bien
$$u_0 = 2 \uparrow\uparrow 3 - 3 = 2^{2^2} - 3 = 16 - 3 = 13$$
et pour tout $n \in \mathbb{N}$
$$u_{n+1} = 2 \uparrow\uparrow (n+4) - 3 = 2^{2\uparrow\uparrow(n+3)} - 3 = 2^{u_n+3} - 3$$

**Ces pages ne sont pas incluses dans l’aperçu.**

## 9.10 Théorème de la forme normale

Je rappelle que nous avons effectué un codage des fonctions primitives récursives, en associant

- à la fonction constante égale à 0 le nombre $[0, 1]$, à la fonction successeur le nombre $[1, 1]$, à la projection $\mathrm{pr}_j^p$ le nombre $[2, p, j]$,
- à la composée de $\langle f_1, \dots, f_n \rangle$ par $g$, où
  - $f_1, \dots, f_n$ sont des fonctions d'arité $p$ et $g$ une fonction d'arité $n$,
  - $a_1, \dots, a_n$ codent respectivement $f_1, \dots, f_n$, et $b$ code $g$,

  le nombre
$$[3, p, a_1, \dots, a_n, b]$$
- à la fonction $\mathrm{rec}(f, g)$, où
  - $f$ est une fonction d'arité $p$ et $g$ une fonction d'arité $p + 2$,
  - $a$ et $b$ codent respectivement $f$ et $g$,

  le nombre
$$[4, p + 1, a, b]$$

Nous allons étendre ce codage à l'ensemble de toutes les fonctions récursives, et à l'ensemble de toutes les fonctions partielles récursives, en codant la fonction (ou fonction partielle) d'arité $p$ obtenue par minimisation à partir de la fonction codée par $a$, par le nombre

$$[5, p, a]$$

On dit d'un nombre codant la construction d'une fonction (ou fonction partielle) récursive $f$ que c'est un *indice* de $f$ (il n'y a pas unicité puisqu'une même fonction peut être construite de différentes manières).

Ce codage des fonctions va nous permettre, d'une certaine façon, de coder aussi tous les calculs possibles effectués par les fonctions récursives (ou par les fonctions partielles récursives). C'est ce que l'on utilise dans le théorème suivant :

**Théorème 9.10.1 (Théorème de la forme normale pour les fonctions récursives)**
Il existe une fonction primitive récursive $U$, et pour tout entier $p \geqslant 1$ un prédicat primitif récursif $T_p$, tels que pour toute fonction récursive $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$, il existe un entier $e$ tel que pour tout $(x_1, \dots, x_p) \in \mathbb{N}^p$
$$\begin{cases} \exists z \in \mathbb{N}, T_p(e, x_1, \dots, x_p, z) \\ f(x_1, \dots, x_p) = U(\mu_z T_p(e, x_1, \dots, x_p, z)) \end{cases}$$

**Preuve**

1. L'idée est que pour toute fonction récursive $f$,
   - $e$ est un indice de $f$, c'est-à-dire un nombre qui code une construction possible de $f$ ;
   - le prédicat $T_p(e, x_1, \dots, x_p, z)$ exprime le fait que le nombre $z$ représente un procédé qui calcule $f(x_1, \dots, x_p)$ ;
   - $\mu_z T_p(e, x_1, \dots, x_p, z)$ est donc l'un de ces nombres $z$ (en l'occurrence le plus petit) ;
   - la fonction $U$ calcule $f(x_1, \dots, x_p)$ à partir du nombre $z$ précédent.

   Le principe général (détaillé plus loin) est le suivant : nous allons construire par induction
   - un ensemble $Z$ de nombres codant le fait qu'on peut obtenir par un calcul une expression de la forme
$$f(x_1, \dots, x_p) = y$$

**Ces pages ne sont pas incluses dans l’aperçu.**

## 9.12 Arithmétique de Robinson

Je rappelle la définition de la théorie de l'arithmétique de Robinson (notée Q), que nous avons vue brièvement dans le chapitre 9 du volume 1. Il s'agit d'une théorie du premier ordre formée essentiellement des mêmes axiomes que ceux de l'arithmétique de Peano (notée PA), sans le schéma de récurrence mais avec un axiome indiquant que tout terme non nul est un successeur (formule aussi prouvable dans PA). La théorie PA est par conséquent une extension de la théorie Q, et les deux admettent comme modèle l'ensemble $\mathbb{N}$ des entiers naturels (avec les interprétations usuelles des opérations) ; la théorie $\mathrm{Th}(\mathbb{N})$ de cette structure, c'est-à-dire l'ensemble des formules closes (dans le langage de Q) vraies dans $\mathbb{N}$, est donc aussi une extension de Q.

**Définition 9.12.1 (Arithmétique de Robinson)**
On appelle *arithmétique de Robinson*, que l'on note Q, la théorie dans le langage de signature $\{0, \mathrm{S}, +, \times\}$, formée des axiomes suivants :

1. 0 n'est pas successeur : pour tout $x$
$$\mathrm{S}\, x \neq 0$$
2. S est injective : pour tout $x$ et $y$
$$\mathrm{S}\, x = \mathrm{S}\, y \implies x = y$$
ce qui équivaut, par contraposition, à
$$x \neq y \implies \mathrm{S}\, x \neq \mathrm{S}\, y$$
3. Tout terme non nul est successeur : pour tout $x$
$$x \neq 0 \implies \exists y, x = \mathrm{S}\, y$$
ce qui équivaut aussi à :
$$x = 0 \text{ ou } \exists y, x = \mathrm{S}\, y$$
4. Définition de l'addition (1) : pour tout $x$
$$x + 0 = x$$
5. Définition de l'addition (2) : pour tout $x$ et $y$
$$x + \mathrm{S}\, y = \mathrm{S}(x + y)$$
6. Définition de la multiplication (1) : pour tout $x$
$$x \times 0 = 0$$
7. Définition de la multiplication (2) : pour tout $x$ et $y$
$$x \times \mathrm{S}\, y = (x \times y) + x$$

Comme pour l'arithmétique de Peano, on définit des termes du langage qui représenteront les entiers naturels de la métathéorie :

**Ces pages ne sont pas incluses dans l’aperçu.**

## 9.13 Récursivité des fonctions représentables dans l'arithmétique de Robinson

Après ces généralités sur l'arithmétique de Robinson, nous pouvons maintenant voir ce que signifie *représenter* une fonction dans cette théorie :

**Définition 9.13.1 (Fonction représentable)**
On dit qu'une fonction $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est *représentée* dans Q par la formule $\mathscr{F}[x_1, \dots, x_p, y]$, lorsque pour tout $m_1, \dots, m_p, n$,

$$\text{si } f(m_1, \dots, m_p) = n \quad \text{alors} \quad \begin{cases} Q \vdash \mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{n}) \\ Q \vdash \forall y, (\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, y) \implies y = \underline{n}) \end{cases}$$

On dit alors aussi que $f$ est *représentable* dans Q.

**Théorème 9.13.2**
Si la fonction $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est représentée dans Q par $\mathscr{F}$, alors pour tous les entiers $m_1, \dots, m_p, n, k$

$$\text{si } f(m_1, \dots, m_p) = n \text{ et } k \neq n \quad \text{alors} \quad Q \vdash \neg\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{k})$$

**Preuve**
Faisons l'hypothèse $f(m_1, \dots, m_p) = n$ et $k \neq n$. On en déduit $\underline{k} \neq \underline{n}$, et comme $\mathscr{F}$ représente $f$, si $\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{k})$ alors $\underline{k} = \underline{n}$. Donc par contraposition

$$\neg\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{k})$$

**Remarque 9.13.3 :** Puisque la théorie Q est cohérente (car elle admet un modèle),

$$\text{si} \quad Q \vdash \neg\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{k}) \quad \text{alors} \quad Q \nvdash \mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{k})$$

Par conséquent, si $f(m_1, \dots, m_p) = n$ alors pour tout entier $k$

$$Q \vdash \mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{k}) \quad \text{si et seulement si} \quad k = n$$

**Remarque 9.13.4 :** Ce qui précède s'étend à n'importe quelle théorie $\mathscr{T}$ qui est une extension de Q (ce qui est par exemple le cas de l'arithmétique de Peano ou de Th($\mathbb{N}$)), puisque toute formule prouvée par Q est aussi prouvée par $\mathscr{T}$ : si la fonction $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est représentée dans Q par $\mathscr{F}[x_1, \dots, x_p, y]$, alors

- $f$ est aussi représentée par $\mathscr{F}$ dans $\mathscr{T}$, c'est-à-dire que si $f(m_1, \dots, m_p) = n$ alors

$$\begin{cases} \mathscr{T} \vdash \mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{n}) \\ \mathscr{T} \vdash \forall y, (\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, y) \implies y = \underline{n}) \end{cases}$$

- si $f(m_1, \dots, m_p) = n$ et $k \neq n$, alors

$$\mathscr{T} \vdash \neg\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{k})$$

- et si $\mathscr{T}$ est cohérente : si $f(m_1, \dots, m_p) = n$ alors pour tout entier $k$

$$\mathscr{T} \vdash \mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{k}) \quad \text{si et seulement si} \quad k = n$$

Ce résultat s'étend même à n'importe quelle théorie dans laquelle on peut « représenter » l'arithmétique de Robinson, par exemple la théorie des ensembles ZFC, ce que l'on peut justifier de manière informelle : nous avons vu qu'on peut construire dans ZFC un modèle $(\omega, \mathrm{S}, 0, +, \times)$ des axiomes de l'arithmétique de Peano (donc aussi a fortiori des axiomes de Q), avec pour tout entier $n$ du métalangage un terme $\underline{n}$ de ZFC tel que $\underline{n} \in \omega$. On en déduit que pour toute formule $\mathscr{F}[x_1, \dots, x_p, y]$ dans le langage de Q,

$$\text{si } Q \vdash \mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{n}) \quad \text{alors} \quad ZFC \vdash \mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{n})$$

et

$$\text{si } Q \vdash \mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, y) \implies y = \underline{n} \quad \text{alors} \quad ZFC \vdash (\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, y) \text{ et } y \in \omega) \implies y = \underline{n}$$

On en déduit que si la fonction $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est représentée dans Q par $\mathscr{F}[x_1, \dots, x_p, y]$, alors elle est représentée dans ZFC par

$$\mathscr{F}' :\equiv \mathscr{F}(x_1, \dots, x_p, y) \text{ et } y \in \omega$$

et

- si $f(m_1, \dots, m_p) = n$ et $k \neq n$, alors

$$ZFC \vdash \neg\mathscr{F}'(\underline{m_1}, \dots, \underline{m_p}, \underline{k})$$

- et si ZFC est cohérente : si $f(m_1, \dots, m_p) = n$ alors pour tout entier $k$

$$ZFC \vdash \mathscr{F}'(\underline{m_1}, \dots, \underline{m_p}, \underline{k}) \quad \text{si et seulement si} \quad k = n$$

Nous allons prouver qu'une fonction $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est récursive si et seulement si elle est représentable dans Q, en commençant par prouver que toute fonction représentable dans l'arithmétique de Robinson est récursive. Pour cela, l'idée est de coder par un entier les termes du langage de Q, puis les formules, puis les preuves formelles dans Q, de telle sorte que si $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est représentée par $\mathscr{F}$, on puisse exprimer par un prédicat décidable l'idée que pour tout $(m_1, \dots, m_p)$, un certain couple $(d, n)$ est tel que $d$ est le code d'une démonstration de la formule $\mathscr{F}(\underline{m_1}, \dots, \underline{m_p}, \underline{n})$. On pourra alors par minimisation (du nombre $[d, n]$) retrouver la valeur $n = f(m_1, \dots, m_p)$, qui s'exprimera alors comme une fonction récursive des variables $m_1, \dots, m_p$.

On commence donc par coder les termes de la théorie Q, en définissant par récursion une fonction de l'ensemble de ces termes dans $\mathbb{N}$. Nous avons vu, de façon générale, que les termes d'une théorie du premier ordre sont construits à partir des variables et constantes, avec une règle pour chaque opérateur (si $\varphi$ est un opérateur d'arité $p$ et si $t_1, \dots, t_p$ sont des termes, alors $\varphi(t_1, \dots, t_p)$ est un terme). Ici, la seule constante est 0, les variables seront notées $x_i$ ($i \geqslant 0$), et les opérateurs sont la fonction successeur, l'addition et la multiplication. Je noterai $[t]$ le nombre associé au terme $t$, défini par récursion ainsi (dans ce qui suit, $x_i$ est une variable, $t$ et $u$ sont des termes) :

$$[0] = [0] \qquad [x_i] = [i + 1]$$

et

$$[\mathrm{S}\, t] = [1, [t]] \qquad [t + u] = [2, [t], [u]] \qquad [t \times u] = [3, [t], [u]]$$

Comme je note de la même manière le codage des termes et le codage des listes d'entiers, il y a une légère ambiguïté, qui ne devrait pas être trop grave car le contexte permet une interprétation correcte : si $t$ est un terme du langage, alors $[t]$ est le code de ce terme, et si $n$ est un entier, alors $[n]$ est le code de la liste dont le seul élément est $n$ ; dans l'expression $[0] = [0]$, le membre de gauche est le code du terme 0 du langage de la théorie Q, égal par définition au membre de droite, qui est le code de la liste [0], c'est-à-dire $p_0^1 \times p_1^0$ (= 2).

**Remarque 9.13.5 :** On constate (par induction sur les termes) que si la variable $x_i$ apparaît dans un terme $t$, alors $i < [t]$ (car $i < i+1 = \mathrm{pr}_1[i+1] < [i+1] = [x_i]$, si $i < [t]$ alors $i < [t] = \mathrm{pr}_2[1,[t]] < [1,[t]] = [St]$, etc.).

**Définition 9.13.6**
On définit le prédicat PR $\mathrm{term}(n)$ signifiant « $n$ est le code d'un terme de Q », par récurrence sur $n$ :

$$\mathrm{term}(n) \overset{\text{def}}{\equiv} (\mathrm{lon}(n)=1) \quad \text{ou} \quad \begin{cases} \mathrm{lon}(n)=2 \\ \mathrm{pr}_1(n)=1 \\ \mathrm{term}(\mathrm{pr}_2(n)) \end{cases} \quad \text{ou} \quad \begin{cases} \mathrm{lon}(n)=3 \\ \mathrm{pr}_1(n)=2 \\ \mathrm{term}(\mathrm{pr}_2(n)) \\ \mathrm{term}(\mathrm{pr}_3(n)) \end{cases} \quad \text{ou} \quad \begin{cases} \mathrm{lon}(n)=3 \\ \mathrm{pr}_1(n)=3 \\ \mathrm{term}(\mathrm{pr}_2(n)) \\ \mathrm{term}(\mathrm{pr}_3(n)) \end{cases}$$

**Remarque 9.13.7 :** L'expression est bien définie puisque $\mathrm{pr}_i(n) < n$.

**Définition 9.13.8**
On définit le prédicat PR $\mathrm{var}(i,n)$ signifiant « $n$ est le code d'un terme $t$ et la variable $x_i$ apparaît dans $t$ », par récurrence sur $n$ :

$$\mathrm{var}(i,n) \overset{\text{def}}{\equiv} \mathrm{term}(n) \quad \text{et} \quad \left( (n=[i+1]) \quad \text{ou} \quad \begin{cases} \mathrm{lon}(n)=2 \\ \mathrm{var}(i,\mathrm{pr}_2(n)) \end{cases} \quad \text{ou} \quad \begin{cases} \mathrm{lon}(n)=3 \\ \mathrm{var}(i,\mathrm{pr}_2(n)) \text{ ou } \mathrm{var}(i,\mathrm{pr}_3(n)) \end{cases} \right)$$

**Définition 9.13.9**
On définit la fonction PR

$$\mathrm{num} : \begin{cases} \mathbb{N} \longrightarrow \mathbb{N} \\ n \longmapsto [\underline{n}] \end{cases}$$

qui à un entier $n$ (du métalangage) associe le code du terme $\underline{n}$ associé, par récurrence sur $n$ :

$$\begin{cases} \mathrm{num}(0) = [\underline{0}] = [0] \\ \mathrm{num}(n+1) = [\underline{n+1}] = [\mathrm{S}\,\underline{n}] = [1, \mathrm{num}(n)] \end{cases}$$

**Remarque 9.13.10 :** On voit (je ne détaillerai pas) que ce qui précède peut être adapté à n'importe quelle théorie du premier ordre, dans un langage dont la signature est dénombrable, en effectuant une variante du codage que j'ai choisi. Par exemple, on peut coder

$$[x_i] = [0,i]$$

pour toute constante $c_i$

$$[c_i] = [1,i]$$

**Ces pages ne sont pas incluses dans l’aperçu.**

**Théorème 9.14.8 (Corollaire)**
1. Toute théorie indécidable et semi-décidable est incomplète.
2. Toute théorie indécidable et complète n'est pas semi-décidable.

**Preuve**
Par contraposition du théorème précédent.

**Exemple 9.14.9**
Les théories de l'arithmétique de Robinson et de Peano sont semi-décidables car elles sont axiomatisables. Nous verrons dans la section 9.19 qu'elles sont indécidables (donc incomplètes). Nous verrons aussi que la théorie Th($\mathbb{N}$) est également indécidable, et par conséquent ne peut pas être axiomatisable (car elle est complète).

## 9.15 Représentabilité dans l'arithmétique de Robinson des fonctions récursives

Nous nous intéressons maintenant à la réciproque du résultat de la section 9.13 : nous allons démontrer par induction structurelle que les fonctions récursives sont représentables dans Q. Il faut pour cela justifier d'une part que les fonctions récursives de base sont représentables, et d'autre part que l'ensemble des fonctions représentables est stable pour les différentes règles de construction des fonctions récursives. Il est cependant difficile de justifier de la stabilité par récursion primitive. Pour traiter ce problème, nous allons donner une définition équivalente de l'ensemble des fonctions récursives, toujours par induction, mais sans la récursion primitive. L'idée est que nous allons pouvoir retrouver cette règle en codant par un entier la liste des valeurs $h(0), \dots, h(n)$ de la fonction $h$ que l'on cherche à définir. Mais la technique que nous avons utilisée jusqu'à présent, permettant de coder $(x_1, \dots, x_n)$ par un entier $[x_1, \dots, x_n]$, a été construite à l'aide de la récursion primitive (via la fonction exponentiation, la minimisation bornée, la suite des nombres premiers,...). Il va donc falloir employer un codage différent, plus sophistiqué que le précédent, mais ne faisant appel qu'à l'addition, la multiplication, et la division euclidienne : la fonction $\beta$ de Gödel.

**Théorème 9.15.1 (Fonction $\beta$ de Gödel)**
Il existe une fonction $\mathbb{N}^2 \xrightarrow{\beta} \mathbb{N}$ telle que pour tous les entiers $a_0, \dots, a_n$, il existe un entier $a$ tel que

$$\forall i \leqslant n, \beta(a, i) = a_i$$

**Preuve**
L'idée est, dans un premier temps, d'utiliser le théorème des restes chinois : si on trouve des entiers non nuls $d_0, \dots, d_n$, premiers entre eux deux à deux, tels que pour tout $i \leqslant n$, $a_i < d_i$, on en déduira l'existence d'un entier $c$ tel que pour tout $i \leqslant n$, $a_i$ est le reste de la division euclidienne de $c$ par $d_i$. On considère des entiers $a_0, \dots, a_n$, et on note

$$\begin{cases} m :\equiv \max(n, a_0, \dots, a_n) \\ d :\equiv m! \end{cases}$$

Vérifions qu'il suffit de choisir les $d_i$ tels que

$$d_i :\equiv 1 + (i+1)d$$

- Faisons l'hypothèse qu'il existe un nombre premier $p$ qui divise $d_i$ et $d_j$, avec $i > j$. Alors $p$ divise

$$d_i - d_j = (i-j)d$$

**Ces pages ne sont pas incluses dans l’aperçu.**

On déduit des théorèmes 9.13.22 et 9.15.12 la synthèse suivante :

**Synthèse 9.15.14 (Équivalence entre fonctions récursives et fonctions représentables dans Q)**
Une fonction $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est récursive si et seulement si elle est représentable dans l'arithmétique de Robinson.

Le résultat s'étend aux extensions de Q :

**Théorème 9.15.15**
Si $\mathscr{T}$ est une extension axiomatisable cohérente de l'arithmétique de Robinson, alors une fonction $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est récursive si et seulement si elle est représentable dans $\mathscr{T}$.

**Preuve**
Si une fonction est récursive, alors elle est représentable dans toute extension de Q, et réciproquement, nous avons vu qu'une fonction représentable dans $\mathscr{T}$ est récursive.

**Remarque 9.15.16 :** Comme nous avons pu le constater ci-dessus, une des implications nécessite des hypothèses moins fortes : si une fonction est récursive, alors elle est représentable dans toute extension de Q (même non axiomatisable).

**Remarque 9.15.17 :** Le résultat s'étend aussi à des théories comme ZFC, puisqu'une fonction représentable dans ZFC est récursive (sous réserve de cohérence), et qu'une fonction représentable dans Q est aussi représentable dans ZFC.

## 9.16 Introduction aux machines de Turing

Une machine de Turing est la représentation symbolique d'une machine pouvant réaliser un algorithme donné. L'idée de Turing était de reproduire, en les ramenant à leur plus simple expression, les opérations que pourrait effectuer un individu réalisant un calcul sur une feuille de papier : d'abord, cette feuille est découpée en plusieurs cases élémentaires, sur chacune desquelles peut être inscrit un certain symbole, choisi parmi un nombre fini (représentant par exemple une lettre, un mot, un chiffre, etc). L'opérateur, à un moment donné, ne peut observer qu'un petit nombre fini de cases en même temps. En fonction des symboles qu'il voit dans ces cases, et de l'« état d'esprit » dans lequel il se trouve (gardant en particulier en mémoire les opérations déjà effectuées), il peut choisir d'effacer ou de modifier certains de ces symboles, ou d'écrire un nouveau symbole dans une case vide. Il déplace ensuite son regard pour observer d'autres cases (à une distance finie proche de sa position actuelle), et ainsi de suite jusqu'à la fin de son calcul.

Ce principe général peut être adapté en de multiples variantes. Il existe en effet de nombreuses versions de machines de Turing, mais elles se sont révélées être toutes équivalentes. Je vais en détailler une.

Une machine de Turing est composée

1. d'un ensemble fini $E$ d'états, que je noterai $e_0, e_1, \ldots$ ;
2. d'un ruban infini vers la droite composé de cases numérotées (la case la plus à gauche est la case 0, la suivante la case 1,…)

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | … |
|---|---|---|---|---|---|---|---|---|---|

et dans lesquelles on peut écrire divers symboles [16]. L'ensemble de ces symboles sera formalisé par un alphabet $A$, contenant au moins un symbole correspondant à une case vide, que je noterai $\varnothing$ ou $0$ ;

3. d'une tête de lecture qui se déplace et indique une case sur le ruban [17] :

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | … |
|---|---|---|---|---|---|---|---|---|---|
| | | ↑ | | | | | | | |

On peut ainsi formaliser une configuration donnée de la machine par un triplet $(e, m, k)$ composé

1. de l'état $e \in E$ ;
2. du mot $m$ écrit sur le ruban (dans l'alphabet $A$), composé d'un nombre fini de symboles différents de $\varnothing$ ;
3. de la position $k \in \mathbb{N}$ de la tête de lecture.

Par exemple le triplet $(e, m, k)$, avec un état $e$, un mot $m = \circ \star \star \circ \varnothing \square$, et une position $k = 2$, représente la situation :

| $e$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ∘ | ★ | ★ | ∘ | | □ | | | | … |
| | | ↑ | | | | | | | … |

À tout ceci s'ajoute

1. une série d'instructions (ce qui correspond à la programmation d'un algorithme), permettant, à partir d'un symbole lu sur le ruban (à l'endroit de la tête de lecture), et de l'état dans lequel se trouve la machine,
   - de choisir un nouvel état (qui peut être le même) ;
   - puis
     — soit d'écrire un nouveau symbole sur le ruban à l'endroit de la tête de lecture (qui peut être le même symbole) ;
     — soit de déplacer la tête de lecture d'une case vers la droite ou vers la gauche.

   Ce que l'on peut représenter par une fonction partielle

   $$\varphi : E \times A \longrightarrow (A \cup \{d, g\}) \times E$$

   qui à un état (un élément de $E$) et un symbole lu sur le ruban (un élément de $A$), associe soit un nouveau symbole à écrire sur le ruban, soit une direction représentée par les symboles $d$ (déplacement d'une case vers la droite), et $g$ (déplacement d'une case vers la gauche), ainsi qu'un nouvel état [18] ; cette fonction partielle peut donc être décrite par un ensemble de couples $((e, x), (y, e'))$ [19], que je noterai pour simplifier sous la forme de quadruplets $(e, x, y, e')$, ou même plus simplement

   $$exye'$$

2. un état initial $e_0 \in E$ ;
3. un ensemble d'états finaux, ou accepteurs, $F \subseteq E$.

16. Parmi toutes les variantes possibles, on trouve des machines de Turing à plusieurs rubans, ou avec des rubans infinis dans les deux directions (à gauche et à droite).

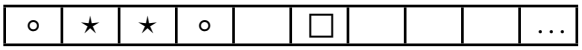
17. Parmi toutes les variantes possibles, on trouve des machines de Turing avec plusieurs têtes de lectures.

18. Parmi les variantes possibles, on trouve des machines de Turing qui effectuent dans la même instruction les deux actions (écriture sur le ruban et déplacement), d'autres qui séparent le fait d'effacer le symbole sur le ruban (ce qui revient à écrire le symbole qui correspond à une case vide) de l'écriture d'un autre symbole, d'autres qui autorisent un déplacement de plus d'une case.

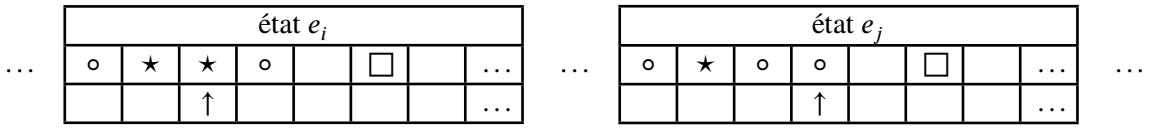
19. La série d'instructions ne devant pas être contradictoire, cet ensemble doit représenter une relation fonctionnelle, c'est-à-dire qu'à un couple $(e, x)$ donné ne correspond qu'un couple $(y, e')$ possible.

La marche à suivre pour effectuer « l'algorithme programmé » est alors le suivant :

1. On écrit sur le ruban la série de symboles correspondant aux données initiales.



   et on positionne la tête de lecture sur une case.
2. On fait fonctionner la machine, qui utilise ses instructions pour construire une succession de différentes configurations :




3. On lit le résultat sur le ruban.

Trois cas de figures sont possibles :

1. La machine s'arrête sur un des états accepteurs.
2. La machine s'arrête, mais son état ne fait pas partie des états accepteurs.
3. La machine ne s'arrête jamais.

Voyons maintenant plus précisément comment calculer des fonctions avec une machine de Turing, et vérifions l'équivalence de ce modèle de calcul avec les fonctions récursives. Ce qui suit est adapté de *Classical Recursion Theory* [20].

Le codage des entiers sur une machine se fera avec l'alphabet $\{0, 1\}$, dans un système unaire, c'est-à-dire par la répétition du symbole 1. Pour ne pas confondre le nombre 0 avec une case vide, qui servira à séparer différents nombres, il y aura un décalage de 1, autrement dit je coderai un nombre à l'aide de la bijection $\mathbb{N} \xrightarrow{n \longmapsto n+1} \mathbb{N}^*$ . Ainsi, l'entier $n$ sera codé par une succession de $n+1$ symboles 1, ce que je noterai $[n]$ :

$$[0] \stackrel{\text{def}}{=} 1 \quad [1] \stackrel{\text{def}}{=} 11 \quad [2] \stackrel{\text{def}}{=} 111 \quad [3] \stackrel{\text{def}}{=} 1111 \quad \dots$$

Je prendrai comme convention que si $n = 0$ alors $[n-1]$ représente une absence de case. Par exemple

$$1 \quad [n-1] \quad 0$$

représente

$$\begin{cases} 1 \quad \underbrace{1 \dots 1}_{n \text{ fois}} \quad 0 & \text{si } n > 0 \\ 1 \quad 0 & \text{si } n = 0 \end{cases}$$

Enfin, je noterai la position de la tête de lecture en gras :

- **0** La tête de lecture est sur cette case qui contient un 0
- **1** La tête de lecture est sur cette case qui contient un 1
- ∗ La tête de lecture est sur cette case qui contient une valeur indéterminée (0 ou 1)
- **[***n***]** La tête de lecture est sur le 1 le plus à droite de [n]
- **[***n*] La tête de lecture est sur le 1 le plus à gauche de [n]

20. Piergiorgio ODIFREDDI. *Classical Recursion Theory. The Theory of Functions and Sets of Natural Numbers*. 2e éd. Elsevier, 1999.

**Définition 9.16.1 (Fonction calculable par une machine de Turing)**
Je dirai qu'une fonction $\mathbb{N}^p \xrightarrow{f} \mathbb{N}$ est *calculable par une machine de Turing*, ou *T-calculable*, lorsqu'il existe une machine de Turing $T_f$ qui, pour tout $(x_1, \dots, x_p) \in \mathbb{N}^p$, à partir de la configuration initiale

$$0 \quad [x_1] \quad 0 \quad [x_2] \quad 0 \quad \dots \quad 0 \quad [x_p] \quad \mathbf{0}$$

(avec des cases vides à droite du dernier 0, mais éventuellement des cases non vides à gauche du premier 0), s'arrête après calcul sur la configuration suivante

$$0 \quad [x_1] \quad 0 \quad [x_2] \quad 0 \quad \dots \quad 0 \quad [x_p] \quad 0 \quad [f(x_1, \dots, x_p)] \quad \mathbf{0}$$

(et sans avoir modifié d'éventuelles cases à gauche).

**Remarque 9.16.2 :** On peut définir de la même façon la calculabilité par une machine de Turing pour une fonction partielle. Dans ce cas, la machine s'arrête si $(x_1, \dots, x_p) \in \mathrm{dom}(f)$, et ne s'arrête jamais sinon.

**Remarque 9.16.3 :** On peut coder de différentes façons ce type de calcul. Ici, les données initiales sont reproduites en sortie, ce qui n'est évidemment pas indispensable pour effectuer un calcul donné, et d'éventuelles cases à gauche des données initiales ne sont pas modifiées. Toutes ces conditions restrictives seront pratiques pour l'élaboration des différentes machines de Turing de cette section, dont une grande partie sera construite à l'aide d'une succession d'autres machines. Ces machines ne sont ni les seules, ni les plus simples, ni les plus rapides, pour effectuer un calcul donné, mais seront suffisantes pour justifier que l'on peut représenter toutes les fonctions récursives.

**Remarque 9.16.4 :** Par défaut, dans la définition d'une machine de Turing de cette section, l'état initial sera $e_0$ et l'état final $e_k$ sera tel que $k$ soit le plus grand des indices des états possibles et tel qu'aucune instruction ne soit associée à $e_k$ (une machine dans l'état $e_k$ est donc à l'arrêt). Ainsi la définition d'une machine de Turing construite en programmant successivement une machine T puis une machine T' sera facilitée : le renommage des états de T' (nécessaire pour qu'il n'y ait pas de confusion avec ceux de T) consistera en un simple décalage d'indice. Par exemple, si l'état final de T est $e_k$, alors les indices de tous les états de T' sont décalés de $k$ (l'état $e_0$ devient $e_k$, l'état $e_1$ devient $e_{k+1}$,...).

Nous allons démontrer par induction que toute fonction récursive est T-calculable, en justifiant que les fonctions de base (fonction constante 0, fonction successeur, fonctions projections) sont T-calculables, et que les trois règles (composition, récursion primitive, minimisation) transforment des fonctions T-calculables en une fonction T-calculable.

**Théorème 9.16.5**
La fonction

$$\begin{cases} \mathbb{N} \longrightarrow \mathbb{N} \\ x \longmapsto 0 \end{cases}$$

est T-calculable.

**Preuve**
On veut à partir de la configuration

$$0 \quad [x] \quad \mathbf{0}$$

obtenir la configuration

$$0 \quad [x] \quad 0 \quad 1 \quad \mathbf{0}$$

Il suffit pour cela de se déplacer d'une case vers la droite, de marquer le symbole 1, et de se redéplacer d'une case vers la droite :

$$e_0 0 d e_1 \qquad e_1 0 1 e_1 \qquad e_1 1 d e_2$$

**Théorème 9.16.6**
La fonction identité

$$\begin{cases} \mathbb{N} \longrightarrow \mathbb{N} \\ x \longmapsto x \end{cases}$$

est T-calculable.

**Preuve**
On veut à partir de la configuration

$$0 \quad [x] \quad \mathbf{0}$$

obtenir la configuration

$$0 \quad [x] \quad 0 \quad [x] \quad \mathbf{0}$$

La machine suivante convient :

1. On se déplace d'une case vers la droite :

   $$e_0 0 d e_1$$

2. On marque 1 :

   $$e_1 0 1 e_2$$

   On se retrouve donc dans la configuration suivante :

   $$0 \quad [x] \quad 0 \quad \mathbf{1}$$

   L'idée est de progressivement supprimer tous les 1 de $[x]$ pour les ajouter à droite. On obtiendra ainsi une succession de configurations de la forme

   $$0 \quad [x-j] \quad \underbrace{0 \dots 0}_{j \text{ fois}} \quad 0 \quad [j]$$

   (le cas initial ci-dessus correspond à $j = 0$).

3. On se place sur le 1 le plus à droite de $[x-j]$ et on le remplace par 0 :

   $$e_2 1 g e_2 \qquad e_2 0 g e_3 \qquad e_3 0 g e_3 \qquad e_3 1 0 e_4$$

   On se retrouve alors dans la configuration

   $$0 \quad [x-j-1] \quad \underbrace{\mathbf{0} \dots 0}_{(j+1) \text{ fois}} \quad 0 \quad [j]$$

4. On regarde s'il reste des 1 dans $[x-j-1]$, en se déplaçant d'une case à gauche pour voir si le symbole est 0 ou 1 :

   $$e_4 0 g e_5$$

   - S'il reste des 1, on se place à droite du dernier 1 de $[j]$, et on revient à l'étape 2 (état $e_1$) :

     $$e_5 1 d e_6 \qquad e_6 0 d e_6 \qquad e_6 1 d e_7 \qquad e_7 1 d e_7 \qquad e_7 0 0 e_1$$

   - S'il ne reste plus de 1, on est dans la configuration suivante

     $$\mathbf{0} \quad \underbrace{0 \dots 0}_{(x+1) \text{ fois}} \quad 0 \quad [x]$$

**Ces pages ne sont pas incluses dans l’aperçu.**

D'après les propriétés de Q

$$\underline{j} \leqslant y \vee y < \underline{j}$$

Si $\underline{j} \leqslant y$ alors $\neg \mathscr{F}(\underline{j}, \underline{m_1}, \dots, \underline{m_p}, \underline{i})$, ce qui est contradictoire. Donc $y < \underline{j}$, et par conséquent il existe un entier $k < j$ tel que $y = \underline{k}$. On en déduit $\mathscr{T} \vdash \mathscr{P}(\underline{m_1}, \dots, \underline{m_p}, \underline{k})$, donc $P(m_1, \dots, m_p, k)$ et par conséquent $R(m_1, \dots, m_p)$.

On déduit des théorèmes 9.18.5, 9.18.7 et 9.18.11, la synthèse suivante :

**Synthèse 9.18.12 (Caractérisation de la décidabilité et de la semi-décidabilité dans Q)**
On considère une relation $R$.

- Si la relation $R$ est décidable, alors elle est représentable dans toute extension de l'arithmétique de Robinson.
- Si la relation $R$ est faiblement représentable dans une extension axiomatisable de l'arithmétique de Robinson, alors elle est semi-décidable.
- Pour toute extension axiomatisable cohérente $\mathscr{T}$ de l'arithmétique de Robinson :
  - La relation $R$ est décidable si et seulement si elle est représentable dans $\mathscr{T}$.
  - La relation $R$ est semi-décidable si et seulement si elle est faiblement représentable dans $\mathscr{T}$.

## 9.19 Théorèmes de limitation : indécidabilité de l'arithmétique, non définissabilité de la vérité, et théorèmes d'incomplétude de Gödel

**Prérequis**
Semi-décidabilité (section 9.11), théories axiomatisables, décidables et semi-décidables (section 9.14), représentation des relations dans l'arithmétique de Robinson (section 9.18).

> « Le développement des mathématiques vers une plus grande précision a conduit, comme on le sait, à la formalisation de larges pans de celles-ci, de sorte que les démonstrations peuvent être effectuées en n'utilisant que quelques règles mécaniques. Les systèmes formels les plus complets mis en place à ce jour sont d'une part le système des *Principia Mathematica*, et d'autre part le système d'axiomes de la théorie des ensembles de Zermelo-Fraenkel [...]. Ces deux systèmes sont si vastes que toutes les méthodes de démonstration utilisées aujourd'hui en mathématiques peuvent y être formalisées, c'est-à-dire qu'elles peuvent se réduire à quelques axiomes et règles d'inférence. On pourrait donc conjecturer que ces axiomes et ces règles d'inférence suffisent à résoudre toutes les questions mathématiques pouvant être formellement exprimées dans ces systèmes. Dans ce qui suit, nous montrerons que ce n'est pas le cas, et qu'il y a au contraire, dans les deux systèmes cités, des problèmes relativement simples issus de la théorie des nombres entiers qui ne peuvent pas être résolus à partir des axiomes. »
>
> Kurt Gödel (1906–1978) [21]

Nous allons, pour finir ce chapitre, voir plusieurs théorèmes classiques et importants de la logique mathé-

21. Kurt GÖDEL. « Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I *[Sur les propositions formellement indécidables des Principia Mathematica et des systèmes apparentés]* ». Dans : *Monatshefte für Mathematik und Physik* 38 (1931), p. 173-198.

matique, faisant partie de ce qu'on appelle parfois des *théorèmes de limitation* (dans le sens où ils apportent des limites à ce que l'on peut faire dans un système formel, mettant un terme (au moins partiellement) aux espoirs du *programme de Hilbert*[22]) : l'indécidabilité de l'arithmétique, la non définissabilité de la vérité de l'arithmétique, et les théorèmes d'incomplétude de Gödel.

Dans la suite, nous allons être amenés à faire appel à des techniques utilisables dans l'arithmétique de Robinson, ou ses extensions, et qui reposent sur l'idée suivante : à toute formule $\mathscr{F}$ (dans le langage de Q) on peut associer son code $[\mathscr{F}]$, qui est un nombre entier. Mais puisqu'à tout entier $n$ correspond un terme $\underline{n}$ de la théorie, on peut aussi associer au nombre $[\mathscr{F}]$ le terme $\underline{[\mathscr{F}]}$ de la théorie, que je noterai $\langle\mathscr{F}\rangle$ :

**Définition 9.19.1**
Pour toute extension $\mathscr{T}$ de l'arithmétique de Robinson, et toute formule $\mathscr{F}$, je noterai

$$\langle\mathscr{F}\rangle \overset{\text{def}}{\equiv} \underline{[\mathscr{F}]}$$

le terme de $\mathscr{T}$ associé au code de $\mathscr{F}$.

Autrement dit, à toute formule $\mathscr{F}$ de Q on associe un terme $\langle\mathscr{F}\rangle$ de Q, que l'on peut ensuite utiliser dans les formules, pour avoir d'une certaine façon des formules qui s'appliquent à elles-mêmes. Nous allons par exemple justifier dans le théorème suivant que pour toute formule $\mathscr{P}$ à une variable, on peut trouver une formule $\mathscr{F}$ qui équivaut à la formule $\mathscr{P}$ appliquée à $\langle\mathscr{F}\rangle$, autrement dit $\mathscr{P}(\langle\mathscr{F}\rangle)$, ce que je noterai ainsi pour simplifier :

$$\mathscr{P}\langle\mathscr{F}\rangle \overset{\text{def}}{\equiv} \mathscr{P}(\langle\mathscr{F}\rangle)$$

Nous pourrons alors utiliser ce théorème pour justifier de l'existence de formules qui font référence à elles-mêmes, comme dans le paradoxe du menteur.

**Théorème 9.19.2 (Lemme diagonal)**
Si $\mathscr{T}$ est une extension de l'arithmétique de Robinson, alors pour toute formule $\mathscr{P}[x]$ à une variable libre, il existe une formule close $\mathscr{F}$ telle que

$$\mathscr{T} \vdash \mathscr{F} \iff \mathscr{P}\langle\mathscr{F}\rangle$$

**Preuve**
On considère la fonction, de l'ensemble des formules à une variable libre dans l'ensemble des formules closes, telle que

$$\mathscr{G}(x) \longmapsto \mathscr{G}\langle\mathscr{G}(x)\rangle$$

Nous exprimons cette fonction sur les codes des formules, à l'aide d'une fonction $\mathbb{N} \xrightarrow{d} \mathbb{N}$ telle que

$$d([\mathscr{G}(x)]) = [\mathscr{G}\langle\mathscr{G}(x)\rangle]$$

Cette fonction est PR, car on peut la définir par

$$n \longmapsto \begin{cases} \text{RpF}(\text{num}(n), i, n) & \text{si } \exists i < n, \begin{cases} \text{lib}(i,n) \\ \forall j < n, (j \neq i \implies \neg\,\text{lib}(j,n)) \end{cases} \\ 0 & \text{sinon} \end{cases}$$

22. Voir le premier chapitre du volume 1, pour une brève présentation des paradoxes et discussions qui ont accompagné les débuts de la formalisation des mathématiques.

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 10

# Introduction à la théorie des catégories

## 10.1 Introduction

La *théorie des catégories* permet d'unifier et de généraliser des notions mathématiques, notamment en étudiant certaines structures et les relations entre elles (morphismes de groupes, fonctions continues entre espaces topologiques…). Il existe deux approches des catégories :

1. Une approche unificatrice à partir de structures définies dans une théorie donnée (par exemple la théorie usuelle des ensembles ZFC). On peut alors vérifier par exemple que les ensembles, les groupes, les espaces topologiques…, forment des catégories particulières dont on peut étudier les propriétés, et mettre en évidence des phénomènes similaires apparaissant dans des champs mathématiques différents.
2. Une approche fondatrice où les structures sont définies en termes catégoriques. On peut par exemple définir les ensembles par des axiomes utilisant uniquement les concepts des catégories.

Pour m'adapter à ces deux approches, je donnerai plusieurs présentations de la notion de catégorie :

1. Deux présentations, dans un métalangage pouvant s'adapter à plusieurs formalismes. La théorie des catégories peut alors en particulier être partiellement formalisée dans la théorie des ensembles ; partiellement dans la mesure où certains concepts nécessitent des collections d'objets plus grandes que des ensembles ou des classes (comme des collections de classes), ce qui peut être difficile à traduire dans le langage de ZFC.
2. Une présentation formalisée dans un langage du premier ordre. Dans ce cas, la théorie des catégories ne s'appuie pas sur la théorie des ensembles ZFC, mais elle en partage les mêmes bases (la logique du premier ordre). Il est alors possible de définir une théorie alternative des ensembles dans un langage du premier ordre. C'est ce que je ferai dans le chapitre 11 où je présenterai la *théorie élémentaire de la catégorie des ensembles*, ou ETCS (*Elementary Theory of the Category of Sets*), du mathématicien américain William Lawvere (1937–).

**Remarque 10.1.1 (Remarque historique) :** Le concept de catégorie a été introduit par le mathématicien américain d'origine polonaise Samuel Eilenberg (1913–1998) et le mathématicien américain Saunders Mac Lane (1909–2005) dans un article de 1945 [1]. Il a été popularisé et développé un peu plus tard, notamment à partir de la fin des années 1950 par le mathématicien français Alexandre Grothendieck (1928–2014), et dans les années 1960 par William Lawvere.

1. Samuel EILENBERG et Saunders MACLANE. « General Theory of Natural Equivalences ». Dans : *Transactions of the American Mathematical Society* 58.2 (sept. 1945).

## 10.2 Description dans un métalangage informel

### Première présentation

Une catégorie est la donnée de deux collections [2] (une dont les éléments s'appellent des *objets* et une dont les éléments s'appellent des *flèches*, ou des *morphismes*) et de deux fonctions (de la collection des flèches dans la collection des objets), vérifiant certains axiomes. La théorie est donc composée des données suivantes, pour toute catégorie $\mathscr{C}$ :

1. Une collection $\mathrm{Ob}(\mathscr{C})$ d'éléments appelés *objets*, que l'on note souvent par une lettre majuscule ($A, B, C, \dots$).
2. Une collection $\mathrm{Fl}(\mathscr{C})$ d'éléments appelés *flèches*, ou *morphismes*, que l'on note souvent par une lettre minuscule ($f, g, h, \dots$).
3. À chaque flèche on associe un objet qu'on appelle sa *source*, ou son *domaine*, et un objet qu'on appelle son *but*, ou son *codomaine*. Cela revient à définir deux fonctions, de $\mathrm{Fl}(\mathscr{C})$ dans $\mathrm{Ob}(\mathscr{C})$ :
   - une fonction qui indique la source de la flèche, que je noterai *sou* ;
   - une fonction qui indique le but de la flèche, que je noterai *but*.

   Lorsque
   $$\mathrm{sou}(f) = A \text{ et } \mathrm{but}(f) = B$$
   on note
   $$f : A \longrightarrow B \quad \text{ou} \quad A \xrightarrow{f} B$$

Les axiomes sont les suivants :

**Axiomes 10.2.1 (Axiomes des catégories)**

1. Composition : pour tous les objets $A, B, C$, et toutes les flèches $f$ et $g$, tels que $A \xrightarrow{f} B$ et $B \xrightarrow{g} C$, il existe une flèche appelée composée de $f$ et de $g$, notée $g \circ f$, telle que $A \xrightarrow{g \circ f} C$.
2. Identités : pour tout objet $A$, il existe une flèche $A \xrightarrow{\mathrm{id}_A} A$, appelée identité de $A$, telle que pour tout objet $B$ et toutes les flèches $A \xrightarrow{f} B$ et $B \xrightarrow{g} A$ :
   $$\begin{cases} f \circ \mathrm{id}_A = f \\ \mathrm{id}_A \circ g = g \end{cases}$$
3. Associativité de $\circ$ : pour toutes les flèches $f, g, h$ telles que $f \circ (g \circ h)$ et $(f \circ g) \circ h$ existent :
   $$f \circ (g \circ h) = (f \circ g) \circ h$$

**Remarque 10.2.2 :** Pour tout objet $A$, son identité est unique puisque deux flèches $A \xrightarrow{\mathrm{id}_A} A$ et $A \xrightarrow{\mathrm{id}'_A} A$ vérifiant la propriété des identités sont telles que
$$\mathrm{id}'_A = \mathrm{id}'_A \circ \mathrm{id}_A = \mathrm{id}_A$$

2. C'est-à-dire de deux *ensembles*, dans le sens intuitif du terme, mais je n'utilise pas ce vocabulaire pour éviter la confusion avec les *ensembles* de la théorie ZFC que j'ai exposée dans le volume 2. Il y a aussi, dans la description des catégories, des problèmes de « taille » d'ensembles sur lesquels je reviendrai dans la section 10.7.

### Deuxième présentation

Une catégorie $\mathscr{C}$ est définie par les données suivantes :

1. Une collection $\mathrm{Ob}(\mathscr{C})$ d'*objets*.
2. Pour tous les objets $A$ et $B$, une collection $\mathrm{Fl}(A, B)$ de *flèches* (ou *morphismes*) ; on note aussi
$$A \xrightarrow{f} B$$
pour signifier que $f$ est une flèche de $\mathrm{Fl}(A, B)$. On dit alors que
   - $A$ est la *source*, ou le *domaine*, de $f$, ce que je noterai aussi $\mathrm{sou}(f)$.
   - $B$ est le *but*, ou le *codomaine*, de $f$, ce que je noterai aussi $\mathrm{but}(f)$.
3. Pour tous les objets $A, B, C$, une fonction
$$\begin{cases} \mathrm{Fl}(B, C) \times \mathrm{Fl}(A, B) \longrightarrow \mathrm{Fl}(A, C) \\ (f, g) \longmapsto f \circ g \end{cases}$$

avec les axiomes suivants :

**Axiomes 10.2.3 (Axiomes des catégories)**

1. Identités : pour tout objet $A$, il existe une flèche $A \xrightarrow{\mathrm{id}_A} A$ telle que pour tout objet $B$ et pour toutes les flèches $A \xrightarrow{f} B$ et $B \xrightarrow{g} A$
$$\begin{cases} f \circ \mathrm{id}_A = f \\ \mathrm{id}_A \circ g = g \end{cases}$$
2. Associativité de $\circ$ : pour tous les objets $A, B, C, D$, pour toutes les flèches $C \xrightarrow{f} D$, $B \xrightarrow{g} C$, $A \xrightarrow{h} B$,
$$f \circ (g \circ h) = (f \circ g) \circ h$$

### Quelques remarques sur les notations et le vocabulaire

**Remarque 10.2.4 (Notations) :** Je noterai $\mathrm{Fl}_{\mathscr{C}}(A, B)$ la collection de toutes les flèches de $\mathscr{C}$ de source $A$ et de but $B$, $\mathrm{Fl}_{\mathscr{C}}(A, \cdot)$ la collection de toutes les flèches de $\mathscr{C}$ de source $A$, et $\mathrm{Fl}_{\mathscr{C}}(\cdot, B)$ la collection de toutes les flèches de $\mathscr{C}$ de but $B$ (l'indice $\mathscr{C}$ pouvant être omis s'il n'y a pas d'ambiguïté sur la catégorie considérée). Pour désigner la collection des flèches d'un objet $A$ vers un objet $B$, on trouve aussi les notations $\mathscr{C}(A, B)$ ou $\hom(A, B)$.

**Remarque 10.2.5 (Notations) :** Pour désigner l'identité de $A$, on trouve aussi la notation $1_A$.

**Remarque 10.2.6 (Notations) :** Je pourrai employer fréquemment la convention classique qui consiste à omettre le symbole $\circ$ entre deux flèches (de la même façon qu'on peut omettre $\times$ dans une multiplication). Par exemple, si on considère deux flèches $A \xrightarrow{f} B$ et $B \xrightarrow{g} C$, alors $gf$ représente $g \circ f$. Il n'y a normalement pas d'ambiguïté, car la composition est la seule opération possible entre deux flèches. Attention néanmoins, il peut arriver que certaines personnes prennent la convention contraire en notant les flèches dans l'ordre de composition (par rapport au diagramme). Dans ce cas, la flèche $g \circ f$, correspondant au diagramme
$$A \xrightarrow{f} B \xrightarrow{g} C$$

serait notée $fg$.

**Remarque 10.2.7 (Notations) :** Je pourrai écrire

$$\text{les flèches } A \underset{g}{\overset{f}{\rightrightarrows}} B$$

pour deux flèches ayant la même source et le même but, comme raccourci pour

$$\text{les flèches } f \text{ et } g \text{ telles que } A \xrightarrow{f} B \text{ et } A \xrightarrow{g} B$$

Et je pourrai de même écrire

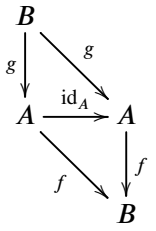
$$\text{les flèches } A \underset{g}{\overset{f}{\rightleftarrows}} B \qquad \text{ou} \qquad \text{les flèches } B \xrightarrow{g} A \xrightarrow{f} B$$

comme raccourci pour

$$\text{les flèches } f \text{ et } g \text{ telles que } A \xrightarrow{f} B \text{ et } B \xrightarrow{g} A$$

Du fait des axiomes, les compositions de flèches peuvent être représentées par des *diagrammes commutatifs*, pour lesquels tout chemin reliant deux objets correspond à la composée des flèches formant le chemin, quel que soit le chemin suivi, comme pour la composition de fonctions vue dans le volume 2 dans le cadre de la théorie des ensembles.

**Exemple 10.2.8**
On peut représenter la propriété des identités par le diagramme commutatif suivant :




Ce diagramme exprime le fait que les flèches $A \xrightarrow{f} B$ et $B \xrightarrow{g} A$ sont telles que

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

## Exemples de catégories

**Exemple 10.2.9 (Catégorie des ensembles)**
On définit la catégorie des ensembles $\mathcal{Ens}$, dont les objets sont les ensembles (au sens de la théorie des ensembles), et les flèches les fonctions. La composition est la composition des fonctions et l'identité $\mathrm{id}_A$ est la fonction identité usuelle ($x \longmapsto x$).

La catégorie des ensembles est un cas particulier de catégories dites *concrètes*, dont les objets sont les ensembles munis d'une structure donnée (comme par exemple une structure de groupe), et les flèches des fonctions qui préservent cette structure (dans l'exemple de groupes, il s'agit de morphismes de groupes).

**Exemple 10.2.10 (Exemples de catégories concrètes)**

1. La catégorie des ensembles ordonnés $\mathcal{Ord}$, dont les objets sont les ensembles ordonnés et les flèches les fonctions croissantes, puisque l'identité est une fonction croissante, et que la composée de deux fonctions croissantes est une fonction croissante.
2. La catégorie des groupes $\mathcal{Grp}$, dont les objets sont les groupes et les flèches les morphismes de groupes.
3. La catégorie des anneaux $\mathcal{Ann}$, dont les objets sont les anneaux et les flèches les morphismes d'anneaux.
4. La catégorie des $\mathbb{K}$-espaces vectoriels $\mathcal{Kev}$, dont les objets sont les $\mathbb{K}$-espaces vectoriels et les flèches les $\mathbb{K}$-fonctions linéaires (pour un corps $\mathbb{K}$ donné).
5. La catégorie des espaces topologiques $\mathcal{Top}$, dont les objets sont les espaces topologiques et les flèches les fonctions continues.

**Remarque 10.2.11 (Notations) :** La notation des différentes catégories n'est pas normalisée. Certains francophones peuvent utiliser les dénominations anglaises et noter par exemple $\mathcal{Set}$ la catégorie des ensembles et $\mathcal{Rng}$ (pour *ring*) la catégorie des anneaux.

Mais toutes les catégories ne sont pas de ce type, comme le montrent les exemples suivants :

**Exemple 10.2.12 (Catégorie associée à une relation d'ordre, de préordre, ou d'équivalence)**

1. On considère un ensemble $A$ muni d'un préordre, c'est-à-dire d'une relation binaire $\leqslant$ réflexive et transitive. On lui associe une catégorie, ayant au plus une flèche entre deux objets, en prenant comme objets les éléments de $A$, les flèches étant caractérisées par la propriété suivante : il existe une flèche $a \longrightarrow b$ si et seulement si $a \leqslant b$. L'identité correspond ici à la réflexivité (puisque $a \leqslant a$, il existe toujours une flèche de source $a$ et de but $a$), et la transitivité permet la composition des flèches (s'il existe deux flèches $a \longrightarrow b$ et $b \longrightarrow c$ , alors $a \leqslant b$ et $b \leqslant c$, donc $a \leqslant c$, et par conséquent il existe une flèche $a \longrightarrow c$ ). Réciproquement, à une catégorie ayant au plus une flèche entre deux objets, on peut associer un préordre.
2. Il en est de même pour un ensemble muni d'une relation d'ordre, ce qui est un cas particulier du précédent (il s'agit d'une relation de préordre à laquelle on ajoute la propriété d'antisymétrie). Dans ce cas s'il existe une flèche de $a$ vers $b$ et une de $b$ vers $a$, alors $a = b$.
3. Il en est aussi de même pour un ensemble muni d'une relation d'équivalence, ce qui est encore un cas particulier du premier point (il s'agit d'une relation de préordre à laquelle on ajoute la propriété de symétrie). Dans ce cas s'il existe une flèche de $a$ vers $b$, alors il existe une flèche de $b$ vers $a$.

**Exemple 10.2.13 (Catégorie associée à un monoïde)**

À tout monoïde $M$ on peut associer une catégorie ayant un seul objet, et dont les flèches sont les éléments de $M$. La composition de flèches correspond au produit de deux éléments (avec juste une

**Ces pages ne sont pas incluses dans l’aperçu.**

**Définition 10.4.3 (Split mono, split épi, rétraction, section)**
Lorsque les flèches $B \xrightarrow{g} A \xrightarrow{f} B$ sont telles que

$$fg = \mathrm{id}_B$$

on dit que

1. $g$ est un *split mono* (ou *split monomorphisme*), ou $g$ est *inversible à gauche*, ou $g$ est une *section* de $f$ (ou un *inverse à droite* de $f$).
2. $f$ est un *split épi* (ou *split épimorphisme*), ou $f$ est *inversible à droite*, ou $f$ est une *rétraction* de $g$ (ou un *inverse à gauche* de $g$) .

**Remarque 10.4.4 (Vocabulaire) :** Je pourrai utiliser le terme *section* comme synonyme de *split mono* (une section est une flèche telle qu'il existe une autre flèche dont elle est la section), et le terme *rétraction* comme synonyme de *split épi* (une rétraction est une flèche telle qu'il existe une autre flèche dont elle est la rétraction). Autrement dit :

- Un split mono, ou une section, est une flèche qui est inversible à gauche (et tout inverse à gauche est une rétraction de cette flèche).
- Un split épi, ou une rétraction, est une flèche qui est inversible à droite (et tout inverse à droite est une section de cette flèche).

**Remarque 10.4.5 (Vocabulaire) :** Les mathématiciens qui travaillent dans le domaine des catégories semblent préférer en général les termes *split mono*, *split épi*, *rétraction*, *section*, aux termes équivalents avec *inversible* ou *inverse*, à cause des ambiguïtés liées au vocabulaire droite/gauche : $f$ est inversible à droite quand elle est en même temps un inverse à gauche d'une autre flèche, et sur le schéma

$$B \xrightarrow{g} A \xrightarrow{f} B$$

qui représente la composée $fg$, la flèche $g$ se situe à gauche de $f$ (cela dépend de la position du diagramme, mais les flèches se tracent souvent de la gauche vers la droite) ; de plus certaines personnes prennent la convention contraire pour noter la composition des flèches, ce qui inverse droite et gauche (le schéma ci-dessus représente alors $gf$).

**Remarque 10.4.6 :** Les notions de split mono et split épi sont duales : $f$ est un split mono dans une catégorie si et seulement si son dual $f^*$ est un split épi.

**Exemple 10.4.7**
Nous avons vu que dans la théorie ZFC

- une fonction inversible à gauche est injective, et réciproquement, une fonction injective de domaine non vide est inversible à gauche, donc dans la catégorie des ensembles les split monos de domaine non vide sont exactement les fonctions injectives de domaine non vide ;
- une fonction est surjective si et seulement si elle est inversible à droite, donc dans la catégorie des ensembles les split épis sont exactement les fonctions surjectives.

En utilisant la notion de factorisation d'une flèche, les définitions précédentes s'expriment ainsi :

**Ces pages ne sont pas incluses dans l’aperçu.**

2. C'est la propriété duale : on considère un split épi $A \xrightarrow{f} B$, avec une section $B \xrightarrow{s} A$, et deux flèches $B \underset{h}{\overset{g}{\rightrightarrows}} C$ telles que

$$gf = hf$$

Alors

$$\begin{cases} (gf)s = g(fs) = g\,\mathrm{id}_B = g \\ (hf)s = h(fs) = h\,\mathrm{id}_B = h \end{cases}$$

Donc $g = h$.

**Remarque 10.4.26 :** En particulier, tout iso est à la fois un mono et un épi. La réciproque est fausse dans le cas général. Par exemple, dans la catégorie des espaces topologiques et fonctions continues, un iso est un homéomorphisme, c'est-à-dire une fonction bijective continue dont la réciproque est continue. Dans cette catégorie, une flèche qui est un mono et un épi est une fonction continue bijective, mais sa réciproque n'est pas nécessairement continue, autrement dit cette flèche n'est pas nécessairement un iso.

**Synthèse 10.4.27 (Synthèse des définitions précédentes)**

- La flèche $f$ est un *mono* lorsqu'elle est simplifiable *à gauche*

$$fg = fh \implies g = h$$

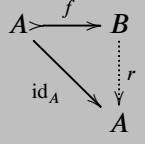
et un *split mono* lorsqu'elle est inversible *à gauche* : il existe une flèche $r$ telle que

$$rf = \mathrm{id}$$

Graphiquement, $A \rightarrowtail B$ est un *mono* lorsqu'une flèche de but $B$ ne peut se factoriser que d'une seule façon à travers $f$, autrement dit lorsque le diagramme commutatif suivant :

$$C \underset{h}{\overset{g}{\rightrightarrows}} A \rightarrowtail B$$

implique $g = h$, et est un *split mono* lorsque $\mathrm{id}_A$ (l'identité de sa source) se factorise à travers $f$, autrement dit lorsqu'on a le diagramme commutatif suivant :




- La flèche $f$ est un *épi* lorsqu'elle est simplifiable *à droite*

$$gf = hf \implies g = h$$

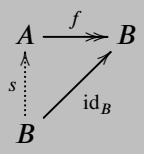
et un *split épi* lorsqu'elle est inversible *à droite* : il existe une flèche $s$ telle que

$$fs = \mathrm{id}$$

Graphiquement, $A \twoheadrightarrow B$ est un *épi* lorsqu'une flèche de source $A$ ne peut se factoriser que d'une seule façon à travers $f$, autrement dit lorsque le diagramme commutatif suivant :

$$A \xrightarrow{f} B \underset{h}{\overset{g}{\rightrightarrows}} C$$

implique $g = h$, et est un *split épi* lorsque $\mathrm{id}_B$ (l'identité de son but) se factorise à travers $f$, autrement dit lorsqu'on a le diagramme commutatif suivant :




**Théorème 10.4.28**
On considère des flèches $A \xrightarrow{f} B$ et $B \xrightarrow{g} C$ .

1. Si $f$ et $g$ sont des isos (respectivement monos, épis, split monos, split épis), alors $gf$ est un iso (respectivement mono, épi, split mono, split épi).
2. Si $gf$ est un mono (respectivement split mono), alors $f$ est un mono (respectivement split mono).
3. Si $gf$ est un épi (respectivement split épi), alors $g$ est un épi (respectivement split épi).
4. Si $gf$ est un iso, alors $f$ est un split mono et un mono, et $g$ est un split épi et un épi.

**Preuve**

1. Si $f$ et $g$ sont des monos et si $gfh = gfh'$, alors $fh = fh'$ (car $g$ est un mono), donc $h = h'$ (car $f$ est un mono), et par conséquent $gf$ est un mono. Si $f$ et $g$ sont des split monos, alors $f$ (respectivement $g$) admet un inverse à gauche $f'$ (respectivement $g'$), donc $gf$ est un split mono car $f'g'$ est un inverse à gauche de $gf$ :

$$f'g'gf = f' \,\mathrm{id}_B\, f = f'f = \mathrm{id}_A$$

   Par dualité, le raisonnement est semblable pour des épis et des split épis, et si $f$ et $g$ sont des isos, ce sont aussi des split monos et des split épis, donc $gf$ aussi, et par conséquent $gf$ est un iso.
2. Si $gf$ est un mono et si $fh = fh'$, alors $gfh = gfh'$, donc $h = h'$, et par conséquent $f$ est un mono. Si $gf$ est un split mono, alors il admet un inverse à gauche $r$ donc

$$(rg)f = r(gf) = \mathrm{id}_A$$

   On en déduit que $rg$ est un inverse à gauche de $f$, et par conséquent $f$ est un mono.
3. Par dualité, le raisonnement est semblable pour les épis et split épis.
4. Si $gf$ est un iso c'est un split mono, donc d'après ce qui précède $f$ est un split mono (donc aussi un mono), et c'est aussi un split épi donc d'après ce qui précède $g$ est un split épi (donc aussi un épi).

**Remarque 10.4.29 :** En d'autres termes (avec le vocabulaire droite/gauche) :

1. Si $f$ et $g$ sont inversibles (respectivement simplifiables à gauche, simplifiables à droite, inversibles à gauche, inversibles à droite), alors $gf$ est inversible (respectivement simplifiable à gauche, simplifiable à droite, inversible à gauche, inversible à droite).
2. Si $gf$ est simplifiable à gauche (respectivement inversible à gauche), alors $f$ est simplifiable à gauche (respectivement inversible à gauche).
3. Si $gf$ est simplifiable à droite (respectivement inversible à droite), alors $g$ est simplifiable à droite (respectivement inversible à droite).
4. Si $gf$ est inversible, alors $f$ est inversible à gauche et simplifiable à gauche, et $g$ est inversible à droite et simplifiable à droite.

**Remarque 10.4.30 :** En d'autres termes (avec le vocabulaire des factorisations) :

**Ces pages ne sont pas incluses dans l’aperçu.**

commutatif suivant :

$$\begin{array}{ccccc} A & \xrightarrow{\varphi} & C & \xrightarrow{\varphi'} & E \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} \\ B & \xrightarrow[\psi]{} & D & \xrightarrow[\psi']{} & F \end{array}$$

(le diagramme est bien commutatif car $\psi'\psi f = \psi' g\varphi = h\varphi'\varphi$), c'est-à-dire que l'on a

$$(\varphi', \psi')\circ(\varphi, \psi) \overset{\text{def}}{\equiv} (\varphi'\circ\varphi, \psi'\circ\psi)$$

## 10.7 Foncteurs

**Définition 10.7.1 (Foncteur)**
On appelle *foncteur* (ou *foncteur covariant*) $F$ entre les catégories $\mathscr{A}$ et $\mathscr{B}$, que je noterai

$$F : \mathscr{A} \longrightarrow \mathscr{B} \qquad \text{ou} \qquad \mathscr{A} \xrightarrow{F} \mathscr{B}$$

toute fonction qui à tout objet de $\mathscr{A}$ associe un objet de $\mathscr{B}$ et à toute flèche de $\mathscr{A}$ associe une flèche de $\mathscr{B}$, en préservant les sources, les buts, les identités, et la composition, autrement dit

1. Si $A \xrightarrow{f} A'$ est une flèche de $\mathscr{A}$, alors
$$F(A) \xrightarrow{F_f} F(A')$$
2. Pour tout objet $A$ de $\mathscr{A}$
$$F_{\mathrm{id}_A} = \mathrm{id}_{F(A)}$$
3. Pour toutes les flèches $f$ et $g$ de $\mathscr{A}$ dont la composée $fg$ existe
$$F_{fg} = F_f F_g$$

**Remarque 10.7.2 (Notations) :** Je pourrai indifféremment exprimer l'image d'une flèche $f$ par un foncteur $F$ sous la forme fonctionnelle classique $F(f)$, mais aussi, comme dans la définition ci-dessus, sous la forme indicielle $F_f$, pour faire une légère distinction avec l'image d'un objet $X$ noté alors $F(X)$. Il ne s'agit que de variantes dans la notation.

Un autre genre de foncteur, que l'on appelle *foncteur contravariant*, inverse le sens des flèches, ce qui revient à considérer un foncteur de la catégorie duale $\mathscr{A}^*$ dans $\mathscr{B}$ :

**Définition 10.7.3 (Foncteur contravariant)**
On appelle *foncteur contravariant* $F$ entre les catégories $\mathscr{A}$ et $\mathscr{B}$, que je noterai aussi

$$F : \mathscr{A} \longrightarrow \mathscr{B} \qquad \text{ou} \qquad \mathscr{A} \xrightarrow{F} \mathscr{B}$$

toute fonction qui à tout objet de $\mathscr{A}$ associe un objet de $\mathscr{B}$ et à toute flèche de $\mathscr{A}$ associe une flèche

**Ces pages ne sont pas incluses dans l’aperçu.**

**Définition 10.7.25 (Transformation naturelle)**
On considère deux catégories $\mathscr{A}$ et $\mathscr{B}$ et deux foncteurs $\mathscr{A} \underset{G}{\overset{F}{\rightrightarrows}} \mathscr{B}$ . On appelle *transformation naturelle* toute famille de flèches de $\mathscr{B}$ indexée par la collection des objets de $\mathscr{A}$, ce que je noterai $(\varphi_X)_{X:\mathscr{A}}$, telle que

1. Pour tout objet $X$ de $\mathscr{A}$, $F(X) \xrightarrow{\varphi_X} G(X)$ .
2. Pour toute flèche $X \xrightarrow{f} Y$ de $\mathscr{A}$
$$\varphi_Y F_f = G_f \varphi_X$$
ce qui peut se traduire par le diagramme commutatif suivant :
$$\begin{array}{ccc} F(X) & \xrightarrow{F_f} & F(Y) \\ {\scriptstyle \varphi_X}\downarrow & & \downarrow{\scriptstyle \varphi_Y} \\ G(X) & \xrightarrow[G_f]{} & G(Y) \end{array}$$

**Remarque 10.7.26 :** Je note $(\varphi_X)_{X:\mathscr{A}}$ par analogie avec les familles d'ensemble, mais ce n'est pas une notation standard. Par ailleurs, si $\mathscr{A}$ et $\mathscr{B}$ sont des petites catégories, alors la *collection* des $\varphi_X$ est bien un *ensemble*, et la famille $(\varphi_X)_{X:\mathscr{A}}$ est la famille $(\varphi_X)_{X\in \mathrm{Ob}(\mathscr{A})}$, qui est aussi elle-même un ensemble ; mais ce n'est pas nécessairement le cas en général.

**Exemple 10.7.27 (Exemple de transformation naturelle)**
On considère la catégorie des ensembles $\mathscr{Ens}$, les foncteurs

- $\mathscr{Ens} \xrightarrow{\mathrm{id}} \mathscr{Ens}$ tel que
$$\begin{cases} X \longmapsto X \\ X \xrightarrow{f} Y \longmapsto X \xrightarrow{f} Y \end{cases}$$
- $\mathscr{Ens} \xrightarrow{\mathscr{P}} \mathscr{Ens}$ tel que
$$\begin{cases} X \longmapsto \mathscr{P}(X) \\ X \xrightarrow{f} Y \longmapsto \mathscr{P}(X) \xrightarrow{\vec{f}} \mathscr{P}(Y) \end{cases}$$

et pour tout ensemble $X$ la fonction
$$\varphi_X : \begin{cases} X \longrightarrow \mathscr{P}(X) \\ x \longmapsto \{x\} \end{cases}$$

**Ces pages ne sont pas incluses dans l’aperçu.**

# Chapitre 11

# Introduction aux topos et théorie élémentaire de la catégorie des ensembles (ETCS)

## 11.1 Introduction

La théorie élémentaire de la catégorie des ensembles (ou ETCS) est due au mathématicien américain William Lawvere (1937–) [1]. Le terme *élémentaire* n'est pas à prendre ici dans le sens de *facile*, mais comme un synonyme de *du premier ordre* [2]. Il s'agit donc d'une théorie qui formalise dans la logique du premier ordre la notion intuitive d'*ensemble*, comme les théories ZFC, NBG ou NFU. Nous allons nous placer dans la théorie des catégories (c'est-à-dire en reprenant les axiomes du premier ordre des catégories de la section 10.3), mais en ajoutant d'autres axiomes pour définir une catégorie particulière, dont les objets s'appellent des ensembles, et les flèches s'appellent des fonctions. Je pourrai employer indifféremment les couples de termes *objet*/*flèche* ou *ensemble*/*fonction*. J'utiliserai en général *objets* et *flèches* dans les définitions (qui sont généralisables à d'autres catégories), mais *ensembles* et *fonctions* dans les axiomes (qui caractérisent la catégorie particulière des ensembles). Je pourrai de même employer indifféremment les couples de termes *source*/*but* ou *domaine*/*codomaine*. Les axiomes supplémentaires que nous allons voir peuvent s'exprimer de manière informelle (et approximative) ainsi :

1. Axiome 1 (Ensemble terminal) : il existe un ensemble terminal.
2. Axiome 2 (Produit cartésien) : pour tous les ensembles $A$ et $B$, on peut former le produit cartésien $A \times B$, ensemble de tous les couples $(a, b)$.
3. Axiome 3 (Ensemble de fonctions) : pour tous les ensembles $A$ et $B$, on peut former l'ensemble $A \longrightarrow B$ de toutes les fonctions de $A$ dans $B$.
4. Axiome 4 (Égaliseurs) : pour toutes les fonctions $A \underset{g}{\overset{f}{\rightrightarrows}} B$ , on peut former l'ensemble des solutions de l'équation $f(x) = g(x)$.
5. Axiome 5 (Classificateur de sous-ensembles) : il existe un ensemble $\Omega$ tel que toute partie $A$ d'un ensemble $E$ puisse être définie par sa fonction indicatrice $E \xrightarrow{\chi_A} \Omega$ , qui, à tout élément $x$ de $E$, associe la valeur *vrai* si et seulement si $x$ appartient à $A$.
6. Axiome 6 (Ensemble initial) : il existe un ensemble initial.

1. William LAWVERE. « An elementary theory on the category of sets (long version) with commentary ». Dans : *Reprints in Theory and Applications of Categories* 11 (2005), p. 1-35. URL : `http://www.tac.mta.ca/tac/reprints/articles/11/tr11abs.html`. Première publication en 1965 dans *National Academy of Science of the USA* 52, 1506-1511.

2. L'expression *théorie élémentaire* est un synonyme, de nos jours obsolète, de *théorie du premier ordre*.

7. Axiome 7 (Union disjointe) : pour tous les ensembles $A$ et $B$, on peut former la somme disjointe (ou union disjointe) $A + B$.
8. Axiome 8 (Coégaliseur), qui a pour conséquence : pour tout ensemble $A$ et toute relation d'équivalence $R$ sur $A$, on peut former l'ensemble quotient $A/R$.
9. Axiome 9 (Ensemble vide) : il existe un ensemble vide.
10. Axiome 10 (Générateur) : toute fonction est entièrement déterminée par les images des éléments de son domaine (si pour tout $x$, $f(x) = g(x)$, alors $f = g$).
11. Axiome 11 (Axiome du choix) : toute fonction $A \xrightarrow{f} B$ surjective est inversible à droite, c'est-à-dire qu'il existe une fonction $B \xrightarrow{g} A$ telle que $fg = \mathrm{id}_B$.
12. Axiome 12 (Ensemble des entiers naturels) : il existe un ensemble des entiers naturels.

Ces différents axiomes définissent plus généralement différentes classes de catégories :

- Les axiomes 1 et 2 définissent une *catégorie cartésienne*.
- L'ajout de l'axiome 3 définit une *catégorie cartésienne fermée*.
- L'ajout des axiomes 4 à 8 définit un *topos élémentaire*.
- L'ajout de l'axiome 9 définit un *topos élémentaire non dégénéré*.
- L'ajout de l'axiome 10 définit un *topos élémentaire bien pointé*.
- L'ajout des axiomes 11 et 12 définit la *catégorie des ensembles*.

Je précise enfin qu'en plus de ces axiomes du premier ordre, je donnerai un certain nombre de définitions et théorèmes généraux, portant sur ces différentes classes de catégories, certains de ces résultats n'étant pas exprimés dans la *théorie élémentaire de la catégorie des ensembles* (la théorie du premier ordre qui est le fil conducteur de ce chapitre), mais dans la métathéorie.

## 11.2 Éléments globaux, éléments généralisés

Pour que les objets de la théorie se comportent comme des ensembles (dans le sens de la théorie naïve des ensembles), nous devons définir la notion d'élément et d'appartenance. Pour cela, nous allons définir d'abord le concept de singleton (ensemble à un seul élément). L'idée est la suivante : pour tout ensemble $X$ il n'existe qu'une seule fonction de $X$ dans un singleton $S$ (à tout élément de $X$ on associe l'unique élément de $S$). Autrement dit $S$ est un objet terminal :

**Axiome 11.2.1 (Ensemble terminal)**
Il existe un ensemble terminal 1.

**Remarque 11.2.2 :** Par définition, pour tout ensemble $X$ il existe une unique fonction de $X$ dans 1, que je noterai $X \xrightarrow{1_X} 1$ .

**Remarque 11.2.3 :** On a choisi un ensemble terminal que l'on note 1, et qui servira d'une certaine manière de référence, mais ce choix n'a pas d'importance. Il peut y avoir d'autres éléments terminaux ou pas, ce n'est pas important car tous les objets terminaux sont isomorphes, et toutes les propriétés de l'ensemble 1 seront valables pour n'importe quel autre ensemble terminal.

De manière informelle, se donner un élément $x$ d'un ensemble $A$ équivaut à se donner une fonction du singleton 1 dans $A$, qui à l'unique élément de 1 associe $x$. C'est ce qui nous permet de retrouver les notions d'appartenance à un ensemble et d'élément :

**Définition 11.2.4 (Élément global, image)**

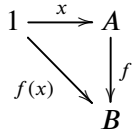
1. On appelle *élément global* (ou *point*) d'un objet $A$, toute flèche $1 \xrightarrow{x} A$ , et on note $x \in A$ :

$$x \in A \overset{\text{def}}{\equiv} 1 \xrightarrow{x} A$$

2. Pour toute flèche $A \xrightarrow{f} B$ et tout élément global $x$ de $A$, $fx$ est un élément global de $B$, que l'on appelle *image* de $x$ par $f$, et que l'on note aussi $f(x)$ :

$$f(x) \overset{\text{def}}{\equiv} fx \overset{\text{def}}{\equiv} f \circ x$$

Ce qui se traduit par le diagramme commutatif suivant :




**Remarque 11.2.5 :** On retrouve ainsi des propriétés usuelles des ensembles et des fonctions : pour toute fonction $A \xrightarrow{f} B$

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

et pour tout $x \in A$

$$\mathrm{id}_A(x) = x$$

**Remarque 11.2.6 (Vocabulaire) :** Dans le cadre des ensembles, je pourrai employer indifféremment les termes *élément global* ou *élément*.

**Remarque 11.2.7 :** L'objet 1 a exactement un élément, puisque l'identité $1 \xrightarrow{\mathrm{id}_1} 1$ est par définition un élément de 1, et par unicité (puisque 1 est un ensemble terminal) c'est le seul. Autrement dit, dans la catégorie des ensembles que nous sommes en train de définir, 1 est bien un singleton (un ensemble ayant exactement 1 élément). Nous verrons plus loin, à l'aide d'un autre axiome, que la réciproque est vraie, c'est-à-dire que tout singleton est un ensemble terminal (et par conséquent un ensemble est un singleton si et seulement si c'est un ensemble terminal, autrement dit si et seulement si il est isomorphe à 1).

Se donner une fonction $1 \xrightarrow{x} A$ revient, d'une certaine manière, à sélectionner un élément de $A$. De la même manière on peut considérer qu'une fonction $X \xrightarrow{f} A$ sélectionne un ensemble d'éléments de $A$ (ceux de son image). Cela revient aussi à se représenter $f$ comme un ensemble d'éléments de $A$ indexés par $X$, ou encore comme un élément variable de $A$ paramétré par $X$.

Ainsi toute flèche $X \xrightarrow{f} A$ généralise la notion d'élément associée aux flèches $1 \xrightarrow{x} A$ :

**Ces pages ne sont pas incluses dans l’aperçu.**

## 11.4 Catégories cartésiennes

Nous allons maintenant définir le produit de deux ensembles :

**Définition 11.4.1 (Produit)**
On appelle *produit* (ou *produit binaire*) des objets $A$ et $B$, tout objet, que l'on note $A \times B$, muni de deux flèches

$$A \xleftarrow{\mathrm{pr}_1} A \times B \xrightarrow{\mathrm{pr}_2} B$$

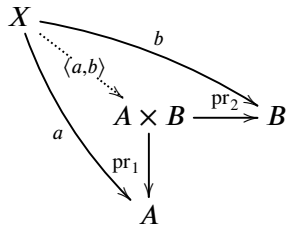
vérifiant la propriété universelle suivante : pour toutes les flèches $A \xleftarrow{a} X \xrightarrow{b} B$ , il existe une unique flèche

$$X \xrightarrow{\langle a,b \rangle} A \times B$$

telle que

$$\begin{cases} \mathrm{pr}_1 \langle a, b \rangle = a \\ \mathrm{pr}_2 \langle a, b \rangle = b \end{cases}$$

Ce qui se traduit par le diagramme commutatif suivant :




Les flèches $\mathrm{pr}_1$ et $\mathrm{pr}_2$ s'appellent respectivement *première* et *deuxième projection*.

**Remarque 11.4.2 (Vocabulaire) :** Dans la catégorie des ensembles que l'on est en train de définir, on peut aussi utiliser le terme *produit cartésien*.

**Remarque 11.4.3 (Notations) :** Je rappelle que la notation d'une flèche en pointillés dans un diagramme commutatif est une convention usuelle qui indique que le reste du diagramme implique l'existence d'une telle flèche.

**Remarque 11.4.4 :** Un produit $A \times B$ est défini comme limite du diagramme formé uniquement de deux objets $A$ et $B$ (sans flèches), puisqu'un cône d'un tel diagramme est de la forme

$$\begin{array}{ccc} X & \longrightarrow & B \\ \downarrow & & \\ A & & \end{array}$$

et la définition du produit équivaut à celle de la limite (c'est un objet terminal dans la catégorie de ces cônes). Un produit est donc bien défini par une propriété universelle (puisque c'est un cas particulier de limite), comme on peut aussi le voir en prenant le foncteur contravariant $\mathscr{C} \xrightarrow{F} \mathscr{E}ns$ (sous réserve d'une catégorie localement petite) tel que

$$\begin{cases} F(X) = \mathrm{Fl}(X, A) \times \mathrm{Fl}(X, B) \\ F(\ X \xrightarrow{f} Y\ ) = (y_A, y_B) \longmapsto (y_A f, y_B f) \end{cases}$$

**Ces pages ne sont pas incluses dans l’aperçu.**

**Définition 11.4.14 (Diagonale)**
Pour tout objet $A$, on appelle *diagonale* de $A$ la flèche

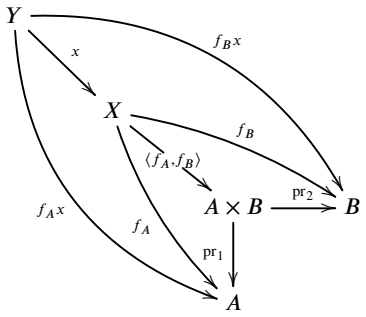
$$\delta \stackrel{\text{def}}{\equiv} \langle \mathrm{id}_A, \mathrm{id}_A \rangle$$

De plus $\delta$ est un split mono (car $\mathrm{pr}_1\, \delta = \mathrm{id}_A$), donc un mono, et par conséquent $A \xrightarrowtail{\delta} A \times A$ est une partie de $A \times A$.

**Théorème 11.4.15**
Pour toutes les flèches $X \xrightarrow{f_A} A$ , $X \xrightarrow{f_B} B$ , et $Y \xrightarrow{x} X$

$$\langle f_A, f_B \rangle x = \langle f_A x, f_B x \rangle$$

**Preuve**




On a

$$\begin{cases} \mathrm{pr}_1 \langle f_A, f_B \rangle x = f_A x = \mathrm{pr}_1 \langle f_A x, f_B x \rangle \\ \mathrm{pr}_2 \langle f_A, f_B \rangle x = f_B x = \mathrm{pr}_2 \langle f_A x, f_B x \rangle \end{cases}$$

donc

$$\langle f_A, f_B \rangle x = \langle f_A x, f_B x \rangle$$

**Remarque 11.4.16 :** En particulier

- si on prend 1 pour $Y$ : pour toutes les flèches $X \xrightarrow{f_A} A$ et $X \xrightarrow{f_B} B$ , et pour tout $x \in X$

$$\langle f_A, f_B \rangle (x) = \langle f_A(x), f_B(x) \rangle$$

- si on prend $X = B = A$ et l'identité pour $f_A$ et $f_B$ : pour tout $Y \xrightarrow{x} A$

$$\delta_A x = \langle \mathrm{id}_A, \mathrm{id}_A \rangle x = \langle \mathrm{id}_A\, x, \mathrm{id}_A\, x \rangle = \langle x, x \rangle$$

**Théorème 11.4.17**
Pour tous les objets $A$ et $B$, on a les isomorphismes suivants :

$$A \times B \simeq B \times A$$
$$A \times 1 \simeq 1 \times A \simeq A$$

**Preuve**
1. On considère les produits $A \xleftarrow{\mathrm{pr}_1} A \times B \xrightarrow{\mathrm{pr}_2} B$ et $B \xleftarrow{\mathrm{pr}'_1} B \times A \xrightarrow{\mathrm{pr}'_2} A$ .

**Ces pages ne sont pas incluses dans l’aperçu.**

entre 1 et $n+1$

$$\mathrm{pr}'_j\ \varphi = x_j$$

Il reste à vérifier l'unicité de $\varphi$. Si $\varphi'$ est une autre flèche vérifiant la même propriété, alors pr $\varphi'$ est telle que pour tout $j$ entre 1 et $n$

$$\mathrm{pr}_j(\mathrm{pr}\ \varphi') = x_j$$

donc par unicité (pour le produit $\prod_{i=1}^{n} A_i$)

$$\mathrm{pr}\ \varphi' = \langle x_1, \ldots, x_n \rangle$$

Comme de plus $\mathrm{pr}_{n+1}\ \varphi' = x_{n+1}$, on en déduit par unicité (pour le produit binaire) $\varphi' = \varphi$.

## 11.5 Catégories cartésiennes fermées

Comme c'était le cas dans la théorie ZFC on souhaite pouvoir construire, à partir de deux ensembles $A$ et $B$, l'ensemble de toutes les fonctions de $A$ dans $B$. De manière informelle, cela implique que si on prend un élément $\varphi$ de cet ensemble, que l'on notera $B^A$, et un élément $y$ de $A$, on en déduit un élément $\varphi(y)$ de $B$ : on doit donc avoir la fonction, dite *évaluation*, suivante :

$$e : \begin{cases} (B^A) \times A \longrightarrow B \\ (\varphi, y) \longmapsto \varphi(y) \end{cases}$$

On souhaite aussi que se donner une fonction à deux arguments

$$f : \begin{cases} X \times A \longrightarrow B \\ (x, y) \longmapsto f(x, y) \end{cases}$$

soit équivalent à se donner une fonction

$$\hat{f} : \begin{cases} X \longrightarrow (B^A) \\ x \longmapsto \hat{f}(x) \end{cases}$$

En effet, toujours de manière informelle, cela revient dans les deux cas à se donner deux éléments, un dans $X$ et un dans $A$, et en déduire un élément de $B$ (il s'agit du principe de curryfication que nous avons vu dans le volume 2). On veut donc obtenir $f(x, y)$ en appliquant l'image par $\hat{f}$ d'un élément $x$ de $X$, à un élément $y$ de $A$ :

$$\hat{f}(x)(y) = f(x, y)$$

autrement dit

$$e(\hat{f}(x), y) = f(x, y)$$

Cela conduit à la définition suivante :

**Définition 11.5.1 (Objet exponentiel)**
Pour tous les objets $A$ et $B$, on appelle *objet exponentiel* tout objet $B^A$ avec une flèche dite d'*évaluation*

$$(B^A) \times A \xrightarrow{e} B$$

vérifiant la propriété universelle suivante : pour tout objet $X$ et toute flèche $X \times A \xrightarrow{f} B$ , il existe

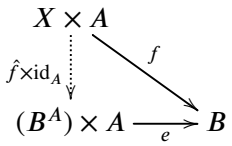
une unique flèche

$$X \xrightarrow{\hat{f}} (B^A)$$

que l'on appelle la *transposée exponentielle* de $f$, telle que

$$f = e(\hat{f} \times \mathrm{id}_A)$$

Ce qui se traduit par le diagramme commutatif suivant :




**Remarque 11.5.2 (Notations) :** J'utilise pour l'ensemble des fonctions de $A$ dans $B$ la notation usuelle $B^A$, et pas $A \longrightarrow B$ que j'ai employée dans l'exposé de la théorie des ensembles ZFC, car $B^A$ est un *objet* de la théorie, et pas une *flèche*, et la notation $A \longrightarrow B$ serait assez ambiguë.

**Remarque 11.5.3 (Notations) :** Il n'y a pas de notation standard pour désigner la transposée exponentielle d'une flèche $f$ (je noterai comme ci-dessus $\hat{f}$).

**Remarque 11.5.4 :** On trouve aussi une variante de la définition dans laquelle le produit $(B^A) \times A$ est remplacé par $A \times (B^A)$.

**Remarque 11.5.5 :** Un objet exponentiel est défini par une propriété universelle, comme on peut le voir en prenant le foncteur contravariant $\mathscr{C} \xrightarrow{F} \mathscr{E}ns$ (sous réserve d'une catégorie localement petite) tel que

$$\begin{cases} F(X) = \mathrm{Fl}(X \times A, B) \\ F(\ X \xrightarrow{f} Y\ ) = g \longmapsto g(f \times \mathrm{id}_A) \end{cases}$$

Le couple $(B^A, e)$ est un élément universel défini par le foncteur contravariant $F$, autrement dit c'est un objet terminal de la catégorie des $(X, x)$, avec $X \times A \xrightarrow{x} B$ .

**Remarque 11.5.6 :** On retrouve notamment la formule indiquée dans le préambule : pour tout $x \in X$ et $y \in A$

$$f\langle x, y\rangle = e(\hat{f} \times \mathrm{id}_A)\langle x, y\rangle = e\langle \hat{f}(x), y\rangle$$

**Remarque 11.5.7 :** Si on prend pour $f$ la flèche évaluation $(B^A) \times A \xrightarrow{e} B$ , alors par unicité la transposée est

$$\hat{e} = \mathrm{id}_{B^A}$$

**Remarque 11.5.8 :** La définition d'un objet exponentiel signifie que dans la métathéorie, la fonction

$$\begin{cases} \mathrm{Fl}(X, B^A) \longrightarrow \mathrm{Fl}(X \times A, B) \\ g \longmapsto e(g \times \mathrm{id}_A) \end{cases}$$

est une bijection, de bijection réciproque

$$\begin{cases} \mathrm{Fl}(X \times A, B) \longrightarrow \mathrm{Fl}(X, B^A) \\ f \longmapsto \hat{f} \end{cases}$$

**Ces pages ne sont pas incluses dans l’aperçu.**

**Théorème 11.6.19**
On considère une flèche $A \xrightarrow{f} B$ et le produit fibré

$$\begin{array}{ccc} A \underset{B}{\times} A & \xrightarrow{\pi_2} & A \\ \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle f} \\ A & \xrightarrow[f]{} & B \end{array}$$

Alors les trois propriétés suivantes sont équivalentes :

- $f$ est un mono.
- $\pi_1$ et $\pi_2$ sont des isos.
- $\pi_1 = \pi_2$.

**Preuve**
Notons d'abord que puisque l'identité de $A$ vérifie trivialement

$$f \,\mathrm{id}_A = f \,\mathrm{id}_A$$

on déduit de la propriété universelle du produit fibré qu'il existe une flèche $A \xrightarrow{\pi} A \underset{B}{\times} A$ telle que

$$\pi_1 \pi = \pi_2 \pi = \mathrm{id}_A$$

donc $\pi_1$ et $\pi_2$ sont des split épis. Démontrons l'équivalence des trois propriétés par implications circulaires :

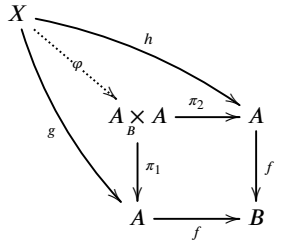
- Si $f$ est un mono, alors $\pi_1$ et $\pi_2$ sont des monos (d'après le théorème précédent), et des split épis, donc sont des isos.
- Si $\pi_1$ et $\pi_2$ sont des isos, alors $\pi_1^{-1} = \pi = \pi_2^{-1}$, donc $\pi_1 = \pi_2$.
- Si $\pi_1 = \pi_2$, et si $X \underset{h}{\overset{g}{\rightrightarrows}} A$ sont deux flèches telles que

  $$fg = fh$$

  alors, d'après la propriété universelle du produit fibré, il existe une flèche $X \xrightarrow{\varphi} A \underset{C}{\times} B$ telle que

  $$g = \pi_1 \varphi = \pi_2 \varphi = h$$

  donc $f$ est un mono.




On en déduit l'équivalence des trois propriétés. On peut aussi noter que puisque $f\pi_1 = f\pi_2$, on avait aussi directement une autre implication : si $f$ est un mono, alors $\pi_1 = \pi_2$.

**Métathéorème 11.6.20**

1. Dans une catégorie avec des produits binaires et des égaliseurs, il existe des produits fibrés.
2. Dans une catégorie avec un objet terminal et des produits fibrés, il existe des produits binaires (donc aussi tous les produits finis).
3. Dans une catégorie avec des produits binaires et des produits fibrés, il existe des égaliseurs.

**Preuve**

1. Nous avons déjà démontré le premier point, en construisant n'importe quel produit fibré à partir d'un produit binaire et d'un égaliseur.

**Ces pages ne sont pas incluses dans l’aperçu.**

c'est la composée de deux monos). On déduit alors de la propriété universelle du produit fibré, que pour tout $X \xrightarrow{x} E$

$$x \in_E A \cap B \iff (x \in_E A \text{ et } x \in_E B)$$

**Remarque 11.8.7 :** Par unicité d'un produit fibré à un iso près, si $C \rightarrowtail^{c} E$ et $C' \rightarrowtail^{c'} E$ sont deux intersections des deux mêmes parties, $C$ et $C'$ sont isomorphes et les intersections sont équivalentes (en tant que parties de $E$). En particulier, par symétrie de la définition, $A \cap B$ et $B \cap A$ représentent tous les deux l'intersection des parties $A \rightarrowtail^{a} E$ et $B \rightarrowtail^{b} E$, et par conséquent $a \cap b \sim b \cap a$.

## 11.9 Topos élémentaires

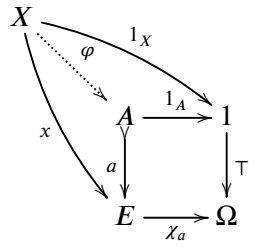
**Définition 11.9.1 (Classificateur de sous-objets)**
On appelle *classificateur de sous-objets*, tout objet $\Omega$ avec un élément $1 \xrightarrow{\top} \Omega$ tel que pour toute partie $A \rightarrowtail^{a} E$ de $E$, il existe une unique flèche $E \xrightarrow{\chi_a} \Omega$ (que j'appellerai *flèche indicatrice*) telle que le diagramme suivant

$$\begin{array}{ccc} A & \xrightarrow{1_A} & 1 \\ {\scriptstyle a}\downarrow & & \downarrow{\scriptstyle \top} \\ E & \xrightarrow[\chi_a]{} & \Omega \end{array}$$

soit un diagramme de produit fibré, c'est-à dire que les propriétés équivalentes suivantes sont vérifiées :

1. $E \xleftarrow{a} A \xrightarrow{1_A} 1$ est le produit fibré de $E \xrightarrow{\chi_a} \Omega \xleftarrow{\top} 1$ .
2. $A \rightarrowtail^{a} E$ est l'image réciproque de $1 \rightarrowtail^{\top} \Omega$ par $E \xrightarrow{\chi_a} \Omega$ .
3. $A \rightarrowtail^{a} E$ est l'égaliseur de $E \underset{\top 1_E}{\overset{\chi_a}{\rightrightarrows}} \Omega$ .

Donc pour toute flèche $X \xrightarrow{x} E$ telle que $\chi_a x = \top 1_X$, il existe une unique flèche $X \xrightarrow{\varphi} A$ telle que $x = a\varphi$, ce qui peut se traduire par le diagramme commutatif suivant




En particulier pour tout $X \xrightarrow{x} E$

$$x \in_E A \iff \chi_a x = \top 1_X$$

**Preuve (de l'équivalence des propriétés)**
Puisque $1 \xrightarrow{\top} \Omega$ est un mono, l'image réciproque correspond par définition au produit fibré. Démontrons l'équivalence

avec l'égaliseur. Notons d'abord que pour tout $X$ il existe toujours une flèche $X \xrightarrow{1_X} 1$, et s'il existe $X \xrightarrow{\varphi} A$, alors $1_A\varphi$ est l'unique flèche de $X$ vers 1, donc on a automatiquement $1_X = 1_A\varphi$. Par conséquent : $E \xleftarrow{a} A \xrightarrow{1_A} 1$ est le produit fibré de $E \xrightarrow{\chi_a} \Omega \xleftarrow{\top} 1$ si et seulement si

$$\begin{cases} \chi_a a = \top 1_A \\ \forall\ X \xrightarrow{x} E\ , (\chi_a x = \top 1_X \implies \exists!\ X \xrightarrow{\varphi} A\ , x = a\varphi) \end{cases}$$

Or $1_E a$ est l'unique flèche de $A$ vers 1 donc $1_A = 1_E a$, et de même $1_E x$ est l'unique flèche de $X$ vers 1 donc $1_X = 1_E x$. Par conséquent la propriété du produit fibré s'écrit

$$\begin{cases} \chi_a a = (\top 1_E)a \\ \forall\ X \xrightarrow{x} E\ , (\chi_a x = (\top 1_E)x \implies \exists!\ X \xrightarrow{\varphi} A\ , x = a\varphi) \end{cases}$$

ce qui équivaut à : $A \rightarrowtail^{a} E$ est l'égaliseur de $E \underset{\top 1_E}{\overset{\chi_a}{\rightrightarrows}} \Omega$.

**Remarque 11.9.2 :** En particulier, si on prend 1 pour $X$, alors $1_X = \mathrm{id}_1$, et on a

$$\forall x \in E, (x \in_E A \iff \chi_a(x) = \top)$$

Ainsi, pour toute partie $A \rightarrowtail^{a} E$ de $E$, $E \xrightarrow{\chi_a} \Omega$ représente la fonction indicatrice de $A$, c'est-à-dire que pour tout élément $x$ de $E$, $\chi_a(x)$ donne la valeur *vrai* ($\top$) si et seulement si $x$ appartient à $A$.

**Théorème 11.9.3 (Unicité d'un classificateur de sous-objets)**
Un classificateur de sous-objets est unique à un iso près.

**Preuve**
Si $1 \xrightarrow{\top} \Omega$ et $1 \xrightarrow{\top'} \Omega$ sont deux classificateurs de sous-objets, alors il existe deux flèches $\Omega \xrightarrow{f} \Omega'$ et $\Omega' \xrightarrow{g} \Omega$ telles que les deux diagrammes suivants soient des diagrammes de produits fibrés

$$\begin{array}{ccc} 1 & \xrightarrow{\mathrm{id}_1} & 1 \\ {\scriptstyle\top}\downarrow & & \downarrow{\scriptstyle\top'} \\ \Omega & \xrightarrow{f} & \Omega' \end{array} \qquad \begin{array}{ccc} 1 & \xrightarrow{\mathrm{id}_1} & 1 \\ {\scriptstyle\top'}\downarrow & & \downarrow{\scriptstyle\top} \\ \Omega' & \xrightarrow{g} & \Omega \end{array}$$

On en déduit en les joignant que le diagramme suivant est un diagramme de produit fibré :

$$\begin{array}{ccc} 1 & \xrightarrow{\mathrm{id}_1} & 1 \\ {\scriptstyle\top}\downarrow & & \downarrow{\scriptstyle\top} \\ \Omega & \xrightarrow{gf} & \Omega \end{array}$$

Or le diagramme suivant est un diagramme de produit fibré

$$\begin{array}{ccc} 1 & \xrightarrow{\mathrm{id}_1} & 1 \\ {\scriptstyle\top}\downarrow & & \downarrow{\scriptstyle\top} \\ \Omega & \xrightarrow{\mathrm{id}_\Omega} & \Omega \end{array}$$

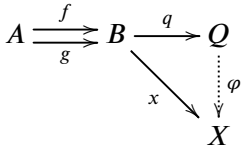
(il est commutatif et si la flèche $X \xrightarrow{x} \Omega$ est telle que $\mathrm{id}_\Omega\, x = \top 1_X$, alors $1_X$ est la seule flèche $X \xrightarrow{\varphi} 1$ telle que $x = \top\varphi$ et $1_X = \mathrm{id}_1\, \varphi$). On en déduit par unicité, d'après la définition d'un classificateur de sous-objets, que $gf = \mathrm{id}_\Omega$. On prouve de même que $fg = \mathrm{id}_{\Omega'}$, et par conséquent $f$ et $g$ sont des isos, donc $\Omega$ et $\Omega'$ sont isomorphes.

**Ces pages ne sont pas incluses dans l’aperçu.**

**Définition 11.10.13 (Coégaliseur)**
On appelle *coégaliseur* des flèches $A \underset{g}{\overset{f}{\rightrightarrows}} B$, tout objet $Q$ muni d'une flèche $B \xrightarrow{q} Q$ telle que $qf = qg$, vérifiant la propriété universelle suivante : pour toute flèche $B \xrightarrow{x} X$ telle que $xf = xg$, il existe une unique flèche $Q \xrightarrow{\varphi} X$ telle que

$$x = \varphi q$$

Ce qui se traduit par le diagramme commutatif suivant :




**Théorème 11.10.14 (Unicité d'un coégaliseur)**
Un coégaliseur est unique à un iso près. Plus précisément, si $B \xrightarrow{q} Q$ et $B \xrightarrow{q'} Q'$ sont deux coégaliseurs de $A \underset{g}{\overset{f}{\rightrightarrows}} B$, alors il existe un unique iso $Q \xrightarrow{\varphi} Q'$ tel que

$$q' = \varphi q$$

**Preuve**
C'est une conséquence du fait qu'un coégaliseur est défini par une propriété universelle.

**Axiome 11.10.15 (Coégaliseurs)**
Pour toutes les fonctions $A \underset{g}{\overset{f}{\rightrightarrows}} B$, il existe un coégaliseur $B \xrightarrow{q} Q$.

**Théorème 11.10.16**
Tout coégaliseur est un épi.

**Preuve**
Par dualité du théorème correspondant pour les égaliseurs (théorème 11.6.5, p. 497).

**Définition 11.10.17 (Épi régulier)**
On dit d'un coégaliseur que c'est un *épi régulier*, autrement dit : on appelle *épi régulier* toute flèche $B \xrightarrow{x} X$ telle qu'il existe deux flèches $A \underset{g}{\overset{f}{\rightrightarrows}} B$ dont $x$ soit un coégaliseur.

**Ces pages ne sont pas incluses dans l’aperçu.**

( $A \xleftarrow{\pi_1} A \cap B \xrightarrow{\pi_2} B$ étant par définition le produit fibré de $A \xrightarrow{a} E \xleftarrow{b} B$ , qui détermine l'intersection de $a$ et $b$).

**Preuve**
La fonction $i_1$ (respectivement $i_2$) est un mono car $ui_1$ (respectivement $ui_2$) l'est ($ui_1 = a$ et $ui_2 = b$). Le diagramme de gauche est commutatif par définition de $A \cup B$. Celui de droite est commutatif car $a\pi_1 = b\pi_2$ (par définition de $A \cap B$), $a = ui_1$ et $b = ui_2$ (par définition de $A \cup B$), et $i_1\pi_1 = i_2\pi_2$ car

$$ui_1\pi_1 = ue \text{ ijg } \pi_1 = a\pi_1 = b\pi_2 = ue \text{ ijd } \pi_2 = ui_2\pi_2$$

et $u$ est un mono.

## 11.14 Topos élémentaires non dégénérés

Le fait que l'objet initial $\varnothing$ est un ensemble vide n'est pas une conséquence des autres axiomes. Il est donc nécessaire d'en ajouter un nouveau. Mais nous avons déjà vu une condition nécessaire et suffisante pour qu'un objet initial soit vide, autrement dit pour qu'il n'ait aucun élément : si 0 est un objet initial, il est vide, autrement dit il n'existe aucune fonction $1 \xrightarrow{x} 0$ , si et seulement si 0 et 1 ne sont pas isomorphes, condition aussi équivalente au fait que 1 n'est pas un ensemble initial, ou que 0 n'est pas un ensemble terminal. Nous savons d'autre part que dans une catégorie cartésienne fermée, ces conditions sont équivalentes au fait qu'il existe des objets non isomorphes. C'est en ce sens que l'on dira d'un tel topos qu'il est *non dégénéré* (le cas *dégénéré* correspondant à la situation où tous les objets sont isomorphes). Ce sont donc ces conditions équivalentes que nous prenons comme axiome supplémentaire :

**Axiome 11.14.1 (Ensemble vide)**
Les cinq versions suivantes de cet axiome sont équivalentes :

1. $\varnothing$ n'a aucun élément, autrement dit il n'existe aucune fonction $1 \xrightarrow{x} \varnothing$ .
2. $\varnothing$ et 1 ne sont pas isomorphes.
3. 1 n'est pas un ensemble initial.
4. $\varnothing$ n'est pas un ensemble terminal.
5. Il existe des objets qui ne sont pas isomorphes.

**Remarque 11.14.2 :** D'après cet axiome, l'ensemble $\varnothing$ est vide (ce qui justifie la notation). La réciproque, c'est-à-dire le fait qu'un ensemble qui n'a pas d'élément soit un ensemble initial (isomorphe à $\varnothing$), est vraie dans la catégorie des ensembles, mais sa justification nécessite un autre axiome que nous verrons plus loin.

Les axiomes vus jusqu'à présent définissent ce qu'on appelle un *topos élémentaire non dégénéré* :

**Définition 11.14.3 (Topos élémentaire non dégénéré)**
On dit qu'un topos élémentaire est *non dégénéré* lorsqu'il existe des objets qui ne sont pas isomorphes (autrement dit lorsque 1 et $\varnothing$ ne sont pas isomorphes).

La catégorie des ensembles (que l'on est en train de définir) étant en particulier une catégorie cartésienne fermée, $\varnothing$ est un objet initial strict. On en déduit pour tout objet $X$ les propriétés suivantes :

**Ces pages ne sont pas incluses dans l’aperçu.**

**Théorème 11.15.22**
On a

$$\neg\neg = \mathrm{id}_\Omega$$

**Preuve**
$\neg\top$ est un élément de $\Omega$, qui ne peut donc être que $\top$ ou $\bot$. Si

$$\neg\top = \top = \top\,\mathrm{id}_1$$

on déduit de la propriété de produit fibré définissant $\neg$ qu'il existe une fonction de 1 dans 1, qui ne peut donc être que l'identité, telle que

$$\top = \bot\,\mathrm{id}_1 = \bot$$

en contradiction avec $\top \neq \bot$. Par conséquent $\neg\top = \bot$, et comme de plus $\neg\bot = \top$ par définition de $\neg$, on a

$$\begin{cases} \neg\neg\bot = \neg\top = \bot \\ \neg\neg\top = \neg\bot = \top \end{cases}$$

Puisque les deux seuls éléments de $\Omega$ sont $\bot$ et $\top$, on en déduit que pour tout $x \in \Omega$

$$\neg\neg(x) = \mathrm{id}_\Omega(x)$$

donc $\neg\neg = \mathrm{id}_\Omega$.

**Remarque 11.15.23 :** Dans n'importe quel topos élémentaire, les deux propriétés que l'on vient de voir ($1 + 1 \xrightarrow{\{\top,\bot\}} \Omega$ est un iso, et $\neg\neg = \mathrm{id}_\Omega$) sont en fait équivalentes, et un topos qui les vérifie est alors dit *booléen*.

## 11.16 Catégorie des ensembles

Il nous reste deux axiomes pour définir la catégorie des ensembles. Le premier est cet axiome classique déjà présent dans ZFC : l'axiome du choix.

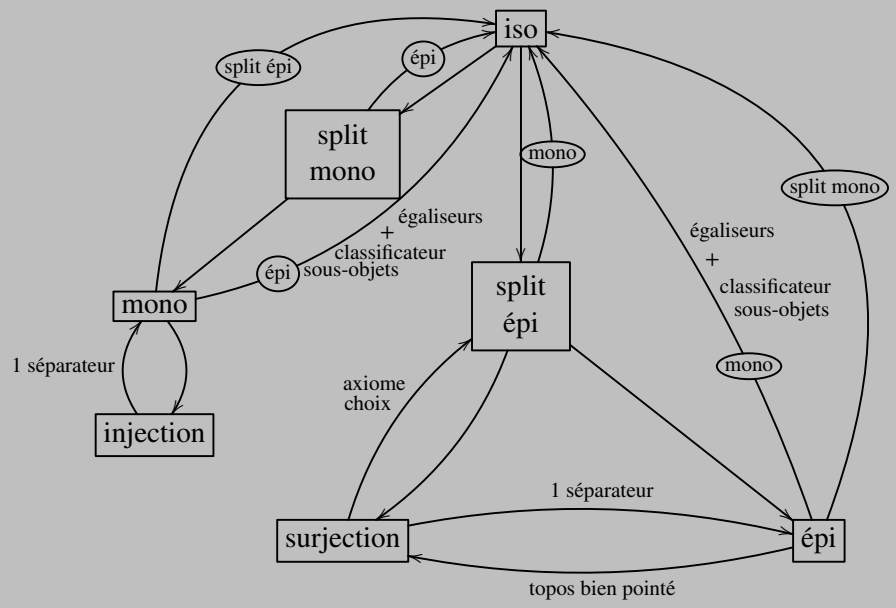
**Axiome 11.16.1 (Axiome du choix)**
Toute surjection est un split épi (autrement dit toute surjection est inversible à droite) : si $A \xrightarrow{f} B$ est une surjection, alors il existe une fonction $B \xrightarrow{g} A$ telle que

$$fg = \mathrm{id}_B$$

**Remarque 11.16.2 :** On retrouve l'une des formulations classiques de l'axiome du choix dans ZFC. Puisque dans un topos bien pointé, une fonction est surjective si et seulement si c'est un épi, on pouvait aussi prendre l'axiome équivalent (dans un topos bien pointé) suivant : tout épi est un split épi (autrement dit tout épi admet une section).

Nous pouvons maintenant faire une dernière synthèse présentant différentes relations entre monos, épis, split monos, split épis, isos, injections et surjections :

**Synthèse 11.16.3**



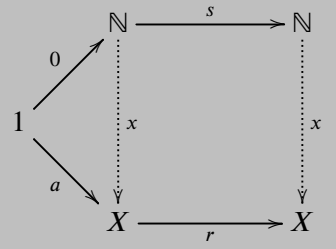

Légende : Une flèche indique une implication, soit toujours vraie (s'il n'y a pas de label), soit avec une ou plusieurs conditions supplémentaires.

Le dernier axiome nous assure de l'existence de l'ensemble $\mathbb{N}$ des entiers naturels :

**Axiome 11.16.4 (Ensemble des entiers naturels)**
Il existe un ensemble $\mathbb{N}$, un élément $0 \in \mathbb{N}$, et une fonction $\mathbb{N} \xrightarrow{s} \mathbb{N}$, tels que : pour tout ensemble $X$, pour tout élément $a \in X$, et pour toute fonction $X \xrightarrow{r} X$, il existe une unique fonction $\mathbb{N} \xrightarrow{x} X$ telle que

$$\begin{cases} x(0) = a \\ xs = rx \end{cases}$$

Ce qui peut se traduire par le diagramme commutatif suivant :




On dit de l'ensemble $\mathbb{N}$, muni des deux fonctions $1 \xrightarrow{0} \mathbb{N} \xrightarrow{s} \mathbb{N}$, que c'est un *objet entier naturel*.

**Ces pages ne sont pas incluses dans l’aperçu.**

# 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*) |
| « ou », « $\vee$ » | Connecteur logique pour la disjonction (*ou*) |
| $\Longrightarrow$ | Connecteur logique pour l'implication |
| $\Longleftrightarrow$ | Connecteur logique pour l'équivalence |
| $\vec{x}$ | Liste de variables $x_1, \ldots, x_n$, pour un entier $n$ indéterminé |
| $\mathscr{F}[x_1, \ldots, x_n]$ | Formule $\mathscr{F}$ dont les variables libres sont à prendre parmi $x_1, \ldots, x_n$ |
| $\mathscr{F}(t/x)$ | Formule $\mathscr{F}$ dans laquelle le terme $t$ remplace la variable $x$ |
| $\forall$ | Quantificateur universel : quel que soit |
| $\exists$ | Quantificateur existentiel : il existe |
| $\exists!x$ | Il existe un unique $x$ tel que … |
| $\equiv$ | Équivalence logique (sémantique ou syntaxique), page 5 |
| $\mathscr{M} \vDash \Gamma$ | La $L$-structure $\mathscr{M}$ est un modèle de $\Gamma$, page 258 |
| $\Gamma \vDash \mathscr{F}$ | $\mathscr{F}$ est une conséquence sémantique de $\Gamma$, page 258 |
| $\Gamma \vdash \mathscr{F}$ | $\mathscr{F}$ est une conséquence syntaxique de $\Gamma$, $\Gamma$ prouve $\mathscr{F}$ |
| $\mathrm{Th}(\mathscr{T})$ | Ensemble des théorèmes de la théorie $\mathscr{T}$, page 259 |
| $\mathrm{Th}(\mathscr{M})$ | Théorie de la $L$-structure $\mathscr{M}$ (ensemble des formules closes dont $\mathscr{M}$ est un modèle), page 265 |
| PA | Théorie de l'arithmétique de Peano |
| Q | Théorie de l'arithmétique de Robinson, page 361 |
| ZF | Théorie des ensembles de Zermelo-Fraenkel |
| ZFC | Théorie des ensembles de Zermelo-Fraenkel avec axiome du choix |
| NBG | Théorie des classes de von Neumann-Bernays-Gödel, page 235 |
| MK | Théorie des classes de Morse-Kelley, page 235 |
| NFU | Théorie *New Foundations with Urelements* [Nouveaux Fondements avec Uréléments], page 241 |

| | |
|---|---|
| $\subseteq$ | Symbole d'inclusion entre ensembles |
| $\subset$ | Symbole d'inclusion stricte entre ensembles |
| $\bigcup \mathscr{E}$ ou $\bigcup_{A\in\mathscr{E}} A$ | Réunion de l'ensemble $\mathscr{E}$ |
| $\bigcup_{i\in I} A_i$ | Réunion de la famille $(A_i)_{i\in I}$ |
| $A \cup B$ | Réunion des ensembles $A$ et $B$ |
| $\bigcap \mathscr{E}$ ou $\bigcap_{A\in\mathscr{E}} A$ | Intersection de l'ensemble $\mathscr{E}$ |
| $\bigcap_{i\in I} A_i$ | Intersection de la famille $(A_i)_{i\in I}$ |
| $A \cap B$ | Intersection des ensembles $A$ et $B$ |
| $A \setminus B$ | Différence des ensembles $A$ et $B$ |
| $\complement_E A$ | Complémentaire de l'ensemble $A$ dans $E$ |
| $A \Delta B$ | Différence symétrique des ensembles $A$ et $B$ |
| $\mathscr{P}(E)$ | Ensemble des sous-ensembles (ou parties) de l'ensemble $E$ |
| $A \times B$ | Produit cartésien des ensembles $A$ et $B$ |
| $\prod_{i\in I} A_i$ | Produit de la famille $(A_i)_{i\in I}$ |
| $A \sqcup B$ | Somme (ou union) disjointe des ensembles $A$ et $B$ : $A \sqcup B \stackrel{\text{def}}{=} (A \times \{0\}) \cup (B \times \{1\})$ |
| $E/\sim$ | Ensemble quotient de $E$ par la relation d'équivalence $\sim$ |
| $\lvert A \rvert$ | Cardinal de l'ensemble $A$, page 213 |
| $A^*$ | Ensemble des éléments non nuls de l'anneau $A$ |
| $A^\times$ | Ensemble des éléments inversibles de l'anneau $A$ (pour la multiplication) |
| $D_a$ | Ensemble des diviseurs de $a$, page 59 |
| $\varnothing$ | Ensemble vide |
| $\mathscr{U}$ | Classe de tous les ensembles |
| $\mathscr{V}$ | Univers de von Neumann (classe des ensembles $\mathscr{V}_\alpha$), page 202 |
| $\mathbb{B}$ | Ensemble $\{0, 1\}$ |
| $\mathbb{N}$ | Ensemble des entiers naturels |
| $\mathbb{N}^*$ | Ensemble des entiers naturels différents de 0 |
| $\mathbb{P}$ | Ensemble des nombres premiers |
| $\omega$ | Ensemble des ordinaux finis (entiers naturels), page 153 |
| $\mathbb{Z}$ | Ensemble des entiers relatifs |
| $\mathbb{Z}/n\mathbb{Z}$ | Anneau des entiers modulo $n$, page 127 |
| $a \equiv b \mod H$ | Congruence modulo $H$ (relation d'équivalence) : $a \equiv b \mod H \stackrel{\text{def}}{=} a^{-1}b \in H$, page 118 |

$a \equiv b \mod n$ Congruence modulo $n$ (relation d'équivalence) : $a \equiv b \mod n \stackrel{\text{def}}{\iff} a - b \in n\mathbb{Z}$, page 128

$\aleph_0$ Cardinal de $\mathbb{N}$ (plus petit cardinal transfini)

$\aleph_\alpha$ $\alpha$-ième aleph (cardinal transfini), page 230

$\mathcal{Ord}$ Classe des ordinaux, page 145

$\mathcal{Lim}$ Classe des ordinaux limites, page 155

$\mathcal{Card}$ Classe des cardinaux, page 216

$f : A \longrightarrow B$ ou $A \xrightarrow{f} B$ $f$ est une fonction de $A$ dans $B$

$f : A \longrightarrow B$ ou $A \xrightarrow{f} B$ $f$ est une flèche de $A$ vers $B$ (théorie des catégories), page 426

$A \rightarrowtail B$ $f$ est un mono (théorie des catégories), page 440

$A \twoheadrightarrow B$ $f$ est un épi (théorie des catégories), page 440

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

$f \circ g$ ou $fg$ Composée de la flèche $g$ par la flèche $f$ (théorie des catégories), page 426

$(A_i)_{i \in I}$ Famille indexée par $I$ (fonction $i \longmapsto A_i$)

$\mathrm{dom}(f)$ Domaine de la fonction $f$

$\mathrm{cod}(f)$ Codomaine (ensemble d'arrivée) de la fonction $f$

$\mathrm{Im}(f)$ Image de la fonction $f$

$\vec{f}(A)$ Image directe de l'ensemble $A$ par la fonction $f$, page 6

$\overleftarrow{f}(A)$ Image réciproque de l'ensemble $A$ par la fonction $f$, page 6

$\mathrm{supp}(f)$ Support de la fonction $f$, page 37

$\mathrm{sou}(f)$ Source de la flèche $f$ (théorie des catégories), page 426

$\mathrm{but}(f)$ But de la flèche $f$ (théorie des catégories), page 426

$A \longrightarrow B$ ou $B^A$ Ensemble des fonctions de $A$ dans $B$

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

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

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

$\mathcal{S}_E$ Ensemble des permutations de $E$ (les bijections de $E$ dans $E$)

$B^{(A)}$ Ensemble des fonctions de $A$ dans $B$ à support fini, page 37

$A \simeq B$ Les ensembles $A$ et $B$ sont en bijection, page 7

$A \simeq B$ ou $(A, \dots) \simeq (B, \dots)$ Les structures $(A, \dots)$ et $(B, \dots)$ sont isomorphes, page 7

$\mathrm{Ob}(\mathcal{C})$ Collection des objets de la catégorie $\mathcal{C}$, page 426

$\mathrm{Fl}(\mathcal{C})$ Collection des flèches de la catégorie $\mathcal{C}$, page 426

$\mathrm{Fl}_{\mathcal{C}}(A, B)$ Collection des flèches de la catégorie $\mathcal{C}$ de source $A$ et de but $B$, page 427

$\mathrm{Fl}_{\mathcal{C}}(A, \cdot)$ Collection des flèches de la catégorie $\mathcal{C}$ de source $A$, page 427

$\mathrm{Fl}_{\mathscr{C}}(\cdot, B)$ Collection des flèches de la catégorie $\mathscr{C}$ de but $B$, page 427

$\mathrm{id}_A$ Fonction identité de l'ensemble $A$ ($x \longmapsto x$)

$\chi_A$ Fonction indicatrice de l'ensemble $A : x \longmapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$

$\langle f_1, \dots, f_n \rangle$ Fonction $x \longmapsto (f_1(x), \dots, f_n(x))$

$f_1 \times \cdots \times f_n$ Fonction $(x_1, \dots, x_n) \longmapsto (f_1(x_1), \dots, f_n(x_n))$

$\mathrm{rec}(f, g)$ Fonction définie par récursion primitive à partir de $f$ et $g$, page 309

$g \circ \langle \overrightarrow{f} \rangle$ Composée de $\langle \overrightarrow{f} \rangle$ par $g$, page 309

$\mu f$ Fonction définie à partir de $f$ par minimisation, page 340

$\mathrm{pr}_n^p$ $n$-ième projection de $A_1 \times \cdots \times A_p : (x_1, \dots, x_p) \longmapsto x_n$, page 310

$\uparrow^p$ Notation des puissances itérées de Knuth, page 316

$[x_1, \dots, x_n]$ Codage de la liste $(x_1, \dots, x_n)$ par un entier, page 326

$\mathrm{pr}_i(x)$ Si $x = [x_1, \dots, x_n]$ alors $\mathrm{pr}_i(x) = x_i$, page 326

$\eta_e^p$ Fonction partielle récursive d'arité $p$ et d'indice $e$, page 354

$H \leqslant G$ $H$ est un sous-groupe de $G$

$H \trianglelefteq G$ $H$ est un sous-groupe normal de $G$ ($H$ est un sous-groupe de $G$ stable par les automorphismes intérieurs), page 116

$\mathrm{SGr}(G)$ Ensemble des sous-groupes de $G$

$\mathrm{SGr}_{\trianglelefteq}(G)$ Ensemble des sous-groupes normaux de $G$, page 116

$Z(G)$ Centre du groupe $G$ (ensemble des éléments qui commutent avec tous les autres)

$\mathrm{Aut}(G)$ Groupe des automorphismes du groupe $G$ (isomorphismes de $G$ dans lui-même)

$\mathrm{Int}(G)$ Groupe des automorphismes intérieurs du groupe $G$ (automorphismes de la forme $x \longmapsto axa^{-1}$)

$\mathrm{Ker}(f)$ Noyau du morphisme de groupes $f$ (ensemble des éléments dont l'image est l'élément neutre)

$[a, b]_A, [a, b[_A, \dots$ Intervalles de l'ensemble ordonné $A$, page 7

$a \wedge b$ Borne inférieure de $\{a, b\}$

$a \wedge b$ ou $\min(a, b)$ Minimum de $a$ et $b$

$a \wedge b$ ou $\mathrm{pgcd}(a, b)$ pgcd (plus grand commun diviseur) de $a$ et $b$, page 75

$a \vee b$ Borne supérieure de $\{a, b\}$

$a \vee b$ ou $\max(a, b)$ Maximum de $a$ et $b$

$a \vee b$ ou $\mathrm{ppcm}(a, b)$ ppcm (plus petit commun multiple) de $a$ et $b$, page 75

$a \mid b$ $a$ divise $b$

$\mathrm{quo}(n, p)$ Quotient de la division euclidienne de $n$ par $p$, page 64

$\mathrm{res}(n, p)$ Reste de la division euclidienne de $n$ par $p$, page 64

| | |
|---|---|
| $\sum_{k=p}^{n} a_k$ | Somme : $a_p + a_{p+1} + \cdots + a_{n-1} + a_n$ |
| $\sum_{i \in I} \kappa_i$ | Somme d'une famille de cardinaux, page 224 |
| $\prod_{k=p}^{n} a_k$ | Produit : $a_p \times a_{p+1} \times \cdots \times a_{n-1} \times a_n$ |
| $\prod_{i \in I} \kappa_i$ | Produit d'une famille de cardinaux, page 224 |

# Index des notions

# Index des noms propres