Cours de mathématiques

TOME I

LOGIQUE & THÉORIE DES ENSEMBLES

Mathématiques générales

France $\sim$ 2024

Écrit et réalisé par Louis Lascaud

Chapitre 1

Logique mathématique

Résumé

Les mathématiques sont formées sur les bases de la théorie des ensembles, construction étroitement intriquée avec la logique théorique. Pour appréhender la première de façon satisfaisante, il n'est pas nécessaire pourtant de comprendre les mécanismes très savants de la seconde ; c'est pourquoi nous fournissons aux débutants, comme tous les manuels de mathématique l'ont toujours proposé, un préambule de logique dite pratique, sans axiomatisation dure. Au cours de cette démarche, nous tirerons deux conclusions : d'une part, de la beauté et de la rigueur des fondations sur lesquelles s'étendent les règles de la logique « habituelle » ; d'autre part, de la difficulté de se convaincre de son bien-fondé... deux constats fort paradoxaux. En tout cas, nous ne pouvons qu'exhorter le lecteur a susciter toute son attention dès les premières lignes de ce cours, même si seulement les parties prochaines auront l'air de la praticité.

1.1 Calcul des propositions

La logique est la science qui permet d'établir la vérité, mais la vérité de quoi ? Celle des propositions. C'est la brique de base, partout reprise, de la logique.

« Définition ». ***(Proposition, assertion)***

Une *proposition* est une phrase assertive ou parfois *assertion* constituée de mots (*métalangage*) ou de symboles mathématiques (*langage mathématique*) qui soit bien formée.
En pratique, on confond souvent les termes *proposition* et *assertion*

Ce n'est pas une bonne définition, à cause du syntagme *bien formé* qui n'a jamais été défini. On précise dans la remarque suivante dans quelle mesure cette notion pourrait être axiomatisée (en légitimant ainsi notre construction[1]).

À cause de la mauvaise définition des propositions avec laquelle nous sommes forcés de composer, nous voyons qu'il n'est pas possible de déterminer définitivement du fait qu'une

proposition donnée en soit une ou non. Nous comptons sur la liste d'exemples donnée ci-dessous pour forger dans l'esprit du lecteur une conception empirique de ce qu'est une proposition, qui peut suffire longtemps encore.

Exemples

1. « $2+2=4$ » est une proposition.
   Il se trouve qu'elle est vraie ; on l'apprend en classe de maternelle.
2. « $2+2=5$ » est une proposition.
   Puisque la précédente est vraie, on démontre avec les règles de l'arithmétique des entiers naturels que celle-ci est fausse.
3. « $2+=4$ » n'est pas une proposition.
   En effet, elle est *mal formée* au niveau du langage mathématique.
4. « $x+1=4$ » n'est pas une proposition.
   En effet, le symbole x n'est pas défini dans aucun des langages, courant ou mathématique. Cependant, « $\forall x \in \mathbb{N} \quad x+1=4$ » est une assertion bien formée, même si elle est bien évidemment fausse.
5. « César a franchi le Rubicon » est une proposition.
   Sa véracité revient en détermination à l'historien. Encore faut-il définir précisément tous les termes : César, le Rubicon, l'action de franchir...
6. « Tout entier pair supérieur à 3 peut s'écrire comme la somme de deux nombres premiers » est une proposition.
   Les mathématiciens ignorent encore aujourd'hui si elle est vraie ou fausse, ce qui est, j'en conviens, un peu honteux.
7. « J'ai faim » est une proposition.
   Il est nécessaire de sous-entendre : que *je* est Louis Lascaud, que l'action se déroule à 16 h 24 le 17 janvier 2023, que la faim est une notion biologiquement déterminée sans ambiguïté.
8. « As-tu faim ? » n'est pas une proposition.
   En effet, ce n'est pas une phrase assertive, mais interrogative. Même problème pour une phrase du type : « Viens ! ».
9. « Par une demain » n'est pas une proposition.
   En effet, elle est mal formée au niveau du métalangage.

Si, normalement, cette définition est largement suffisante pour l'instant, ça n'est pas toujours évident, même pour des étudiants ayant poursuivi plusieurs années d'études supérieures de mathématiques. Nous citons cet exemple véridique. On demande à un élève de démontrer que tout groupe fini dont tout élément est d'ordre 2 a pour ordre une puissance de 2.
L'élève propose de raisonner par récurrence sur le cardinal du groupe, ce qui est très possible. Néanmoins, énoncer l'hypothèse de récurrence et la propriété à montrer en hérédité relève une vraie difficulté chez lui : on voit écrire pour ne citer qu'une absurdité : « on suppose que tout groupe d'ordre n dont tout élément est d'ordre 2 a pour cardinal une puissance de 2 ; montrons que le plus petit groupe de cardinal supérieur a pour cardinal une puissance de 2 supérieure ». Cette dernière phrase n'est pas une proposition valide et fausse la résolution du problème ! Ce qui semble grossier sorti du contexte, n'est pas si ridicule lorsque plongé dans l'exercice en question.

Remarque. Il est aisé de voir comment la formule *bien formé* peut être formalisée. Dans la définition précédente, nous nous sommes permis d'inclure des *mots* dans les propositions, ce qui fait partie du *métalangage* mathématique, en opposition au *langage* mathématique composé exclusivement de symboles. En effet, toute proposition mathématique peut être écrite seulement avec des symboles : par exemple, « $\forall x\ \forall y \quad (x = y) \iff ((\forall t \in x, t \in y) \wedge (\forall t \in y, t \in x))$ » n'est autre que le principe de double inclusion que connaissent bien les étudiants de première année lorsqu'il veulent montrer l'égalité de deux ensembles. À partir de là, on peut supposer que, dans la définition des propositions, nous nous restreignons aux propositions contenant un métalangage que l'on peut convertir en langage mathématique. Pour exemple, « x appartient à E » n'est autre qu'une phrase en français pour dire mathématiquement « $x \in E$ ». Ainsi, en toute rigueur, « César a franchi le Rubicon », comme on l'a indiqué déjà, n'est pas une proposition... À cause de César, du Rubicon et de l'action de franchir qui ne font pas partie de la théorie des ensembles, tout au moins sans définition préalable.

Puisque, comme toute autre science, les résultats mathématiques doivent pouvoir être formalisés au-delà de la multiplicité des langages, la précision de la science a mathématique a forcé à éliminer ainsi le risque de malentendu lié au langage oral ou courant en faisant le choix de traduire les propositions en un langage artificiel formel ne contenant que des termes logiques. Dès la deuxième section de ce cours, nous commencerons d'en donner les principes.

En mathématiques, une proposition bien formée du métalangage ne comporte pas forcément de point final.

Comme dans toutes les mathématiques, nous travaillons de surcroît dans le cadre d'un certain calcul *littéral*, c'est-à-dire avec des lettres, afin de pouvoir généraliser nos constructions, ici, sur les propositions. Habituellement, nous notons par les lettres P,Q,R,S, etc., des propositions quelconques, non déterminées par une phrase du même type que celles ci-dessous (constituée de métalangage ou de langage mathématique, ou des deux), mais qui, sans condition autre, le pourraient être à n'importe quel moment.

⟶ **Notation.** Pour donner un nom littéral à une proposition déterminée, par exemple, « César n'a jamais franchi le Rubicon », nous notons de la manière suivante :

$$P : \text{ « César n'a jamais franchi le Rubicon ».}$$

Les lecteurs avancés reconnaîtront à juste titre la rédaction propre à l'en-tête d'un raisonnement par récurrence (Montrons pour tout $n \in \mathbb{N}$, $\mathcal{P}_n$: « $n \leqslant 2^n$ » par exemple).

Astuce !

Pour les lecteurs n'ayant pas commencé la lecture de ce cours par le présent chapitre, nous déconseillons vivement de jamais quantifier P,Q... dans des formules impliquant de telles propositions. En effet, non seulement l'usage des quantificateurs $\forall, \exists$ est par définition réservé aux ensembles, mais en outre, la théorie dite *logique du premier ordre*, dont nous utilisons les outils, interdit formellement la quantification des prédicats (une proposition n'étant autre qu'un prédicat à aucun argument). Ainsi, en logique théorique, cet usage est incorrect ; dans la logique naïve que nous présentons, il est plutôt mal placé. Rien n'empêche, par contre, d'écrire : « Pour toute proposition P, [...] ».

1.1.1 Axiomatique primaire d'une logique naïve

En guise d'activité introductive, on propose au lecteur l'exercice suivant à résoudre avec plus de jugeotte que d'austérité d'esprit.

Exercice 1

(Quelques valeurs de vérité à attribuer) Pour chacun des assertions suivantes, donner sa valeur de vérité, à savoir le vrai ou le faux.

1. « Le rosier est une plante »
2. « La baleine est un poisson »
3. « $15684 + 749412 = 795096$ »
4. « Le nombre 4 est un nombre premier »
5. « Il existe des fonctions continues non dérivables »
6. « 5 est un nombre premier si et seulement si deux droites parallèles et disjointes n'ont aucun point commun »
7. « Si 5 est impair, alors la somme des angles d'un triangle plan vaut 180ř. »
8. « Tout quadrilatère avec deux côtés opposés égaux et deux angles opposés égaux est un parallélogramme »

Pour formaliser proprement les fondements du calcul propositionnel, par souci purement cosmogonique[1], nous introduisons une fonction multivaluée sur la classe des propositions lui attribuant une valeur de vérité. On rappelle qu'une fonction multivaluée f de $\mathcal{C}$ dans $\mathcal{D}$ est une fonction définie sur une partie de $\mathcal{C}$ et qui à tout élément x de cette partie associe un ou plusieurs éléments de $\mathcal{D}$, par exemple y, et que dans ce cas, on note $f(x) \ni y$.

Axiome. *(Notion de booléen)*

Une proposition peut être vraie ou fausse. Plus précisément, on suppose qu'il existe une fonction définie partout sur la collection de toutes les propositions à valeurs dans $\{0,1\}$ (a priori multivaluée).

⟶ **Notation.** Notons valver cette fonction multivaluée. On a donc :

$$\text{valver} : \mathcal{P} \multimap \{0,1\}$$

avec le symbole des multifonctions, en notant $\mathcal{P}$ la classe de toutes les propositions.

On appelle *vrai* le nombre 1 et *faux* le nombre 0.

Avec ce premier axiome, nous y allons calmement. Nous énonçons qu'à toute proposition, c'est-à-dire à toute assertion bien formée, on peut faire correspondre au moins une *valeur de vérité*, ce qui, dans le langage courant, correspond au vrai et au faux. Cependant, sans les précisions suivantes, on ignore si une proposition peut être ni vraie ni fausse, ou s'il existe des propositions vraies et fausses simultanément. Les deux axiomes qui arrivent énoncent que ces deux cas de figure sont exclus.

[1] Et qui risque d'alourdir le discours gravement. On invite donc les lecteurs fragiles à se focaliser sur les points importants.

Axiome. ***(Principe du tiers exclu, principe de la double valeur)***

La fonction valver est définie partout.

Le plus important des deux axiomes précédents est le mot **partout**.

On constate que cet axiome énonce exactement qu'une proposition est soit vraie, soit fausse, mais au moins l'un des deux.

⟶ **Notation.** On note simplement

$$P$$

pour $= \text{valver}(P) \ni 1$, mais parfois, ce qui est regrettable, on note « P est vraie » par souci de clarté, avec ou sans guillemets. On peut aussi d'ores et déjà noter Non P, ou encore $\neg P$ pour $\text{valver}(P) \ni 0$, ou, ce qui est encore plus regrettable, « P est fausse », toujours pour clarté.

Exercice 2

Que dire de la proposition « Tout entier pair supérieur à 3 peut s'écrire comme la somme de deux nombres premiers » ?

▷ **Éléments de réponse.**

Cette assertion (*conjecture de Goldbach*) est, d'après le principe précédent, soit vraie, soit fausse. Nous ignorons en fait si elle est vraie ou fausse, mais qu'importe : ce n'est pas parce que nous ne savons pas la valeur de vérité de cette proposition, qu'elle n'en a pas.

Axiome. ***(Principe de non-contradiction)***

L'application valver est en fait monovaluée.

On constate que cet axiome énonce exactement qu'une proposition ne peut être à la fois vraie et fausse ; c'est-à-dire, conjoint à l'axiome précédent, qu'une proposition est soit vraie, soit fausse, l'un des deux, et un seul des deux.

Exercice 3

Se convaincre que l'on ne peut trouver de proposition vraie et fausse.

▷ **Éléments de réponse.**

Par exemple, en déduire que tout est vrai, et tout est faux. Montrer ensuite que la proposition « tout est vrai et tout est faux » est fausse, puis tourner en rond. On admet alors qu'un diallèle est un illogisme.

Des logiques non binaires

La propriété (ou, pour nous, axiome) précédente n'est autre que la **bivalence** de la logique classique, unique logique utilisée depuis Aristote jusqu'à son développement formel à la fin du XIX^{e} siècle. Son principe très ancien rend ainsi compte d'une intuition profonde : les choses sont vraies ou fausses ; elles doivent absolument être l'une des deux, et rien n'est vrai et faux à la fois, faute de paradoxe. Cette logique est celle sur laquelle devait se fonder tous les écrits des logiciens pendant deux millénaires. Elle n'est pas aussi définitive pour tout le monde.

Certains mathématiciens étudient des logiques *polyvalentes* (on dit aussi *multivalentes*) c'est-à-dire où les propositions peuvent prendre plus de deux valeurs de vérité. Les premières logiques polyvalentes sont développées dans les années 1920 à la suite des travaux du mathématicien polonais Jan Łukasiewicz. Par exemple, la *logique tertiaire* entend répondre aux besoins de la mécanique quantique en prenant en compte trois états de vérité, VRAI, FAUX et INCONNU (ou, comme il conviendrait mieux de dire, INDÉTERMINÉ). Il est important de voire que ce troisième état est complètement absent de notre modèle de logique théorique, comme on l'a déjà mentionné : une proposition, même de valeur de vérité inconnue (par quelqu'un), est soit vraie, soit fausse.

Il existe même des logiques dont le nombre de valeurs de vérité possible est infini, et même non dénombrable ! La *logique floue* inventée en 1965 par Lofti Zadeh permet de créer des propositions dont la valeur de vérité prend ses valeurs dans l'intervalle de réels [0,1], en fonction du degré de sûreté de la proposition : 0 si elle est résolument fausse, 1 si elle est résolument vraie.

Remarque. Les deux axiomes précédent permettent d'énoncer le fait suivant : valver est une multifonction monovaluée définie partout... Autrement dit, c'est une simple application de la classe de toutes les propositions dans $\{0,1\}$. Nous admettons qu'il n'y a eu qu'un intérêt pédagogique à amener successivement que, d'une part, valver est définie partout, ce qui en fait une multi-application, et d'autre part, qu'elle est monovaluée, ce qui en fait une simple fonction. On peut donc noter en toute rigueur, dans le formalisme des classes :

$$\text{valver} \in \{0,1\}^{\mathcal{C}}.$$

Certains logiciens adoptent un point de vue très ensembliste de la logique naïve, en remarquant que, après ce qui a été vu précédemment, $\mathcal{C}$ s'écrit comme la réunion disjointe de la sous-classe des propositions vraies $\mathcal{V}$ et la sous-classe des propositions fausses qui est son complémentaire $\mathcal{F}$. On a même exactement $\mathcal{V} = \text{valver}^{-1}(\{1\})$ et $\mathcal{F} = \text{valver}^{-1}(\{1\})$. Pour eux, le calcul propositionnel s'effectue non plus sur $\mathcal{C}$ grâce à valver mais sur $\mathcal{C}/\mathcal{R}$ où $\mathcal{R}$ est la relation définie sur $\mathcal{C}$ par $P\mathcal{R}Q$ si et seulement si $\text{valver}(P) = \text{valver}(Q)$ et l'on travaille avec l'application quotient $\tilde{\text{valver}}$.

Dans cette conception, il n'y a que deux propositions : une proposition vraie, appelée *le vrai*, notée $\top$, et une proposition fausse, appelée *le faux*, notée $\bot$. Comme les conceptions $\mathcal{C}$ et $\tilde{\mathcal{C}}$ de la logique naïve se chevauchent, on n'écrit jamais les propositions $\tilde{P}$, et même lorsque la première conception prévaut, on s'autorise à écrire, pour « P est vraie » successivement :

$$P = \top \text{ ou } P \iff \top \text{ ou } P : \top,$$

de même pour « P est fausse ».

Remarquons d'ores et déjà que $\mathcal{R}$ n'est autre que la relation d'équivalence définie plus bas dans ce document. La conception $\tilde{\mathcal{C}}$ a l'intérêt de ne considérer que le vrai et le faux indépendamment de la réalité pratique des propositions ; quant aux constructions de la section suivante, elle les légitime totalement *(voir à ce moment)*.

Cette identification des propositions à leurs valeurs de vérité n'aura plus vraiment de sens pour les prédicats qui, eux, *prennent* véritablement des valeurs de vérité de façon variante.

Reformulation. ***(Axiomes des propositions)***

On peut résumer les trois faits précédents par les points suivants :

(i) Une proposition, si elle existe, peut prendre la valeur de vérité vrai ou faux.

(ii) Il n'existe pas d'autres valeurs de vérité.

(iii) Toute proposition prend au moins l'une des deux valeurs de vérité vrai ou faux.

(iv) Une proposition ne prend pas simultanément les deux valeurs de vérité vrai et faux.

Quel est l'intérêt de cette construction ? Nous allons définir, pour développer dans le bon sens le calcul propositionnel, des opérations pour créer des formules, qui ne seront autres que des propositions compositions de propositions au moyen d'opérations : $f(P,Q,R)$. Selon la conception précédente, la création $f(P)$ n'a que peu d'intérêt logique : P ne peut toujours prendre que deux valeurs et bien sûr $f(P)$ en tant que proposition (si tant est que, bien sûr, elle est bien formée), mais une proposition de la forme $f(P_1,...,P_n)$ peut prendre toujours seulement deux valeurs de vérité en tant que proposition bien formée, mais cette valeur variera dans 2^n cas selon les valeurs de vérité prises par $P_1,...,P_n$!

Afin de *calculer* avec les propositions, on distinguera chacun de ses 2^n cas de figure. Afin de tous les représenter, on aura profit, dès que $n \geqslant 2$, de les écrire dans un tableau, appelé *table de vérité* comme représentée ci-dessous, ayant pour buts exhaustivité et clarté, et qui facilite beaucoup en pratique le calcul propositionnel comme on pourra l'éprouver très rapidement.

P
V
F

TABLE 1.1 : *Table de vérité triviale à une variable.* —
Lorsqu'on a qu'une proposition, elle ne peut prendre que la variable de vérité *vrai* ou *faux*. On sent déjà qu'il n'y aura donc que deux opérateurs unaires : l'identité et la négation, c'est-à-dire, conserver les deux valeurs de vérité ou les inverser.

P	Q
V	V
V	F
F	V
F	F

TABLE 1.2 : *Table de vérité à deux variables (sans proposition composée à calculer.* —
Voilà la forme générale d'une table de vérité à deux variables. **Important.** Pour ne pas s'y perdre, on commence toujours par la gauche où l'on inscrit les valeurs de vérité vraies et fausses à la suite (on divise par deux le tableau, le haut et le bas). Dans chacune de ces divisions, on divise le tableau en deux, le haut et le bas, qui seront les places respectivement du vrai et du faux de la deuxième proposition, et ainsi de suite, de sorte que la dernière colonne sera constituée par l'alternance une à une des valeurs vrai et faux de la n-ième proposition.

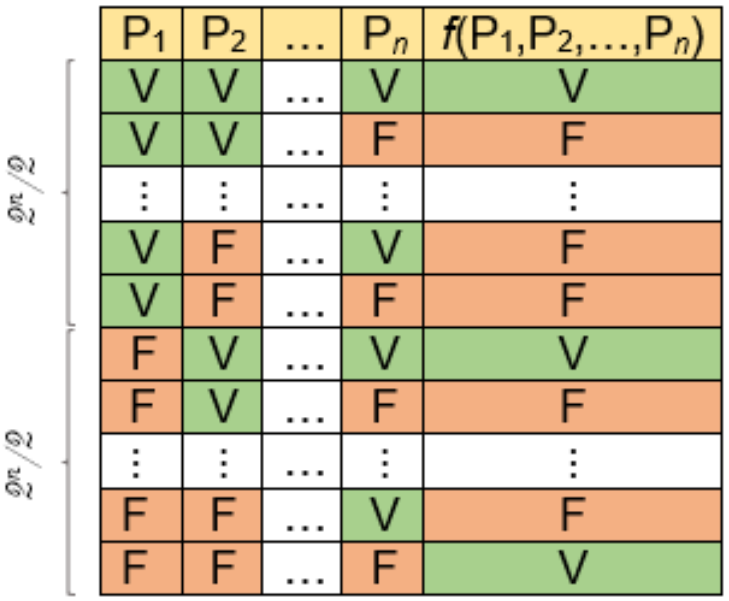


TABLE 1.3 : *Table de vérité à n variables.* — Généralisation des faits précédents.

Ce qu'il faut retenir

→ La logique naïve travaille sur les *propositions*. On suppose qu'elles existent, elles ne sont pas définies très proprement mais on essaie d'identifier leur *bonne formation*.

→ Comme au fil de toute démarche axiomatique, ces propositions, quoique non définies, sont régies par des *axiomes*. Ceux-ci portent sur la notion de *valeur de vérité*, qu'une proposition peut prendre selon des règles précises.

→ Une certaine façon de voire les choses (conception quotient, point de vue ensembliste) alors est d'*identifier* les propositions à leurs valeurs de vérité, ce qui peut paraître cruel mais est très cohérent pour l'œil pratique. Plus couramment, on adopte une conception classique (*i.e.* sur les classes, point de vue prédicatif).

→ Pour écrire les calculs sur un ensemble fini de propositions, on utilisera le formalisme des *tableaux de vérité* permettant de représenter clairement tous les cas de figure selon que des propositions quelconques prennent l'une ou l'autre des deux valeurs de vérité possibles.

→ Ce qui est intéressant, en logique puis en mathématique, c'est de former des formules complexes qui soient vraies pour les 2^n valeurs de vérité considérées : une telle formule sera appelée *théorème du calcul propositionnel*. Plus généralement, mais peut-être aussi de façon plus biaisée, on peut considérer que toute l'activité mathématique consiste à établir de telles formules, appelées alors *théorèmes*.

1.1.2 Opérations sur les propositions

1.1.2.1 Négation d'une proposition

La définition suivante n'en est pas vraiment une dans la première conception de la logique naïve, mais elle en est une dans la deuxième. Ainsi, dans la première conception, il faudra remplacer le mot « Définition » par « Axiome » et commencer l'énoncé par *On admet qu'il existe une unique proposition telle que...*

La négation transforme une proposition en une autre proposition : ce sera un opérateur unaire (d'ailleurs, le seul, avec l'identité bien sûr qui ne change rien à la proposition, ce qui nous fera intéresser très vite aux opérateurs binaires entre propositions).

Définition. ***(Négation)***

Soit P une proposition. On appelle *négation* de P la proposition Q telle que Q est vraie si P est fausse et Q est fausse si P est vraie. On note : $\neg P$ ou $\text{Non}(P)$ ou capiNon P ou $\text{Neg}(P)$.

Exercice 4

Se convaincre que l'on ne peut trouver de proposition ni vraie, ni fausse.

▷ **Éléments de réponse.**

Soit une proposition ni vraie ni fausse. Que dire de sa négation ?

Une autre façon de définir la négation. L'opération de négation d'une proposition peut être entièrement définie par sa table de vérité, comme suit :

P	Non P
V	F
F	V

Table 1.4 : *Table de vérité de la négation.* — Cette table a pour nous valeur de définition.

Intuitivement, la négation d'une proposition est celle qui dit le « contraire » de la proposition de départ. Comme on l'a déjà évoqué (encore ! qu'est-ce qu'on évoque dans ce cours), elle échange les valeurs de vérité selon celles de la proposition initiale, tandis que l'identité (que nous ne prenons pas la peine de définir, car ce n'est même pas un opérateur dans la conception quotient) les conserve.

On admet pour l'instant que deux propositions sont équivalentes si et seulement si elles ont les mêmes valeurs de vérité, vraies ou fausses.

Propriété. ***(Involutivité de la négation)***

Pour toute proposition P, $\neg\neg P \iff P$. Autrement dit, l'opérateur de négation est involutif.

▷ La méthode incontournable pour montrer ce genre de formules, même ici très élémentaire, est la formalisation en tables de vérité. On rédigé de la manière suivante : ■

Logique intuitionniste : *first encounter*

La double négation, selon le théorème précédent, est traité directement dans notre logique. Ce n'est plus le cas en logique intuitionniste.

1.1.2.2 Conjonction

1.1.2.3 Disjonction

1.1.2.4 Implication

1.1.2.5 Équivalence

1.1.2.6 Opérateurs binaires (en général)

1.1.3 Conséquences du calcul propositionnel

Théorème. ***(Reformulation du principe de non-contradiction)***

capiNon P capiou P.

1.2 Prédicats

1.2.1 Définition

1.2.2 Principe de la preuve

1.3 Théorèmes de la logique classique

1.3.1 Lois usuelles

1.3.2 Principes démonstratifs

Chapitre 2

Les raisonnements mathématiques

Résumé

On classifie les façons de démontrer les plus classiques en mathématiques, et qui sont les seules qui serviront chez nous.

2.1 Méthodes générales de démonstration

2.1.1 Méthodes de démonstration directes

2.1.2 Méthodes de démonstration indirectes

2.1.2.1 Contraposée versus absurde

2.1.3 Autres méthodes générales de démonstration

2.2 Pratique de la démonstration

2.2.1 Principes de démonstration

2.2.2 Paradigmes de preuve

2.2.2.1 Paradigmes analytiques

2.3 Quelques pièges dans les démonstrations mathématiques

Chapitre 3

Théorie naïve des ensembles

Résumé

On présente une théorie des ensembles munie de l'axiomative naïve de la fin du XIX^{e} siècle. Cet absence d'axiomatique n'empêche pas l'arithmétique cardinale et ordinale toutefois. Toute considération touchant aux modèles de théorie est hors de propos.

3.1 La démarche axiomatique en philosophie des sciences

Nous avons déjà rencontré l'exemple trébuchant de la logique naïve. Essayons de généraliser ce processus.

3.2 Axiomes de la théorie des ensembles

3.2.1 L'axiome de choix

3.2.1.1 Théorème de Zorn

Définition. ***(Maximal)***

Soit $(E, \leqslant)$ un ensemble ordonné. Un élément a de E est dit *maximal* si pour tout $x \in E$, $a \leqslant x \implies a = x$. On définit de même la notion d'*élément minimal*.

Définition. ***(Chaîne)***

Une *chaîne* d'un ensemble ordonné $(E, \leqslant)$ est une partie A de cet ensemble E sur laquelle la restriction $\leqslant_{|A\times A}$ de l'ordre est totale.

Définition. ***(Ensemble inductif)***

Un ensemble inductif E est un ensemble ordonné dont toute chaîne est majorée (par un élément a priori dans E).

Théorème. ***(Lemme de Zorn)***

Tout ensemble inductif a un élément maximal.

▷ Nous donnons une preuve plutôt laborieuse de ce résultat, dite « par au-dessus » (*top-down* en anglais). Celle-ci est beaucoup moins judicieuse qu'une preuve « par en dessous » (*bottom-up* en anglais), qui est l'autre preuve classique du lemme de Zorn, mais a l'avantage de ne pas recourir à la théorie des ordinaux.

Soit $(E, \leqslant)$ un ensemble ordonné inductif. Remarquons que l'ensemble vide n'est pas inductif. Soit σ une fonction de choix sur la famille de toutes les parties non vides de E. Pour $X \subseteq E$ et $a \in E$, on note $X \preccurlyeq a$ pour $\forall x \in X \quad x \leqslant a$ et de même pour une inégalité stricte. Pour toute chaîne C, on définit $C^+ = C \cup \{\sigma(\{a \mid C \prec a\})\}$ si $\{a \mid C \prec a\}$ est non vide, et $C^+ = C$ sinon.

Soit C une chaîne quelconque de E (il en existe toujours une, par exemple la partie vide). Parce que E est inductif, il existe a vérifiant $C \preccurlyeq a$ par définition. Si a est maximal dans E, il n'y a rien à faire. Sinon, par définition, il existe un certain b dans E vérifiant $a < b$, et donc par transitivité $C \prec b$. Par conséquent, si C est une chaîne telle que $\{a \mid C \prec a\}$ soit vide, c'est-à-dire vérifiant $C^+ = C$, alors il existe une élément maximal dans E. Nous allons construire une telle chaîne.

On appelle *close* toute famille K de chaînes de E telle que $C \in K$ entraîne $C^+ \in K$, et que, si J est une partie de K formées de chaînes deux à deux comparables pour l'inclusion, alors leur réunion appartienne encore à K. La famille de toutes les chaînes de A est évidemment close, et l'on vérifie que toute intersection de familles closes est close. Il existe donc une plus petite famille close K (l'intersection de toutes les familles closes, qui ne pose pas de problème de définition puisque l'ensemble des familles closes est non vide). Posons enfin $K' = \{C \in K \mid \forall D \in K \quad C \subseteq D \vee D \subseteq C\}$. On va montrer que $K' = K$, c'est-à-dire que K est composée de chaînes deux à deux comparables pour l'inclusion. Supposons cela démontré. On pose C la réunion des éléments de K. Par définition, on a $C \in K$ et donc $C^+ \in K$. Or, par construction, on a $D \subseteq \bigcup K = C$ pour toute chaîne D dans K. En particulier, on a donc $C^+ \subseteq C$, d'où $C^+ = C$, comme souhaité.

Revenons sur notre postulat. Puisque K est la plus petite famille close, et que l'on a $K' \subseteq K$, il suffit, pour montrer $K' = K$, de montrer que K' est close. Soit $C \in K'$. Posons $K_C = \{D \in K \mid D \subseteq C \vee C^+ \subseteq K'\}$. Supposons que $D \in K_C$. Si $C^+ \subseteq D$, on a *a fortiori* $C^+ \subseteq D^+$. Pour $C = D$, on a trivialement $C^+ = D^+$. Supposons alors $D \subsetneq C$. Par hypothèse, D^+ est dans K, et C est dans K', donc on a $D^+ \subseteq C$ ou $C \subsetneq D^+$. Le second cas est incompatible avec $D \subsetneq C$ puisque D^+ privé de D est un singleton. Dans tous les cas, $D \in K_C$ entraîne donc $D^+ \in K_C$. Supposons maintenant que J soit un sous-ensemble de K_C formé de chaînes deux à deux comparables pour l'inclusion. Ou bien on a $D \subseteq C$ pour tout D dans J, et l'on a alors $\bigcup J \subseteq C$, ou bien il existe D dans J vérifiant $C^+ \subseteq D$, et l'on a alors $C^+ \subseteq \bigcup J$: dans les deux cas, $\bigcup J$ est dans K_C. Ainsi, K_C est une famille close, donc $K_C = K$, ce qui montre que C^+ est dans K' dès que C s'y trouve.

Finalement, supposons que J est un sous-ensemble de K' formé de chaînes deux à deux comparables pour l'inclusion. Soit D une chaîne quelconque dans K. Ou bien on a $C \subseteq D$ pour toute chaîne C dans J, dont on déduit que $\bigcup J \subseteq D$, ou bien il existe C dans J vérifiant $D \subseteq C$, dont on déduit que $D \subseteq \bigcup J$. Donc, dans tous les cas, $\bigcup J \in K'$. Il en résulte que K' est close, et on a donc $K' = K$. ■

3.2.2 L'axiome de fondation

Un petit développement sur l'axiome de fondation, axiome supplémentaire de la théorie des ensembles classiques (au même titre que l'axiome du choix) et qui n'est pas du tout utile. C'est pourquoi les discussions à propos de son adoption ne sont pas véhémentes, et l'on peut considérer des théories tout à fait semblables en termes des mathématiques que nous connaissons munies soit d'un axiome de fondation, soit d'un axiome d'anti-fondation. Le principe général de l'axiome de fondation est d'interdire la construction d'ensembles qui s'appartiennent eux-mêmes.

3.2.2.1 Retour sur la relation d'appartenance

En théorie naïve des ensembles, on pose qu'il existe des objets, appelés *ensembles*, liés par une relation dite d'appartenance, et notée $\in$, et dont les règles de construction sont regroupées en une liste d'axiomes. Tout ce que nous appelons *élément* est ensemble, et réciproquement tout ensemble peut-être vu comme un élément[1]. Ce que sont les ensembles n'est pas précisé. Plus généralement, on regroupe le concept intuitif de collection d'objets sous le terme de *classe*, de sorte que tout ensemble soit une classe, mais ce n'est pas réciproque : par exemple, la classe regroupant tous les ensembles n'est pas un ensemble (paradoxe de Russell) ; elle est dite impropre.

La liste des axiomes choisis constitue les fondements de la théorie ; si quelques axiomes semblent essentiels à la construction d'une théorie des ensembles pertinente, pour d'autres, la communauté n'est pas décidée, tant qu'ils sont indépendants (aucun axiome n'est conséquence logique des autres), c'est en particulier le cas pour l'axiome du choix. Ceci mène à l'élaboration de différents *modèles* d'une théorie, à laquelle nous associons les noms de leur créateur, avec plus ou moins de précision : Z pour Zermelo, ZF pour Zermelo et Fraenkel, ZFC pour ce système d'axiomes additionné de l'axiome du choix.

Ainsi les axiomes, qui sont admis obtusément, ont pour but primaire de régir les règles de construction d'ensembles et relatives à la relation d'appartenance notée $\in$, en hommage au ϵ grec, pour le verbe « est » en latin. Le but de cette section est de donner de la relation d'appartenance quelques propriétés importances, avec pour conséquence notable la construction véritable des entiers naturels.

3.2.2.1.1 Notion intuitive d'appartenance

Axiome. ***(Relation d'appartenance)***

Sur la classe des ensembles, il existe une relation notée $\in$, c'est-à-dire une partie (impropre) de Ens $\times$ Ens.

[1] En effet, si E : Ens, *i.e.* « E est un ensemble », alors d'après l'axiome de la paire, $E \in F$ en posant $F = \{E\}$.

Exemples

1. $1,2,3 \in \mathbb{N}$;
2. $-1 \notin \mathbb{N}$;
3. $-1,1 \in \mathbb{Z}$;
4. $(-1,1) \notin \mathbb{N}$;
5. $(-1,1) \in \mathbb{Z} \times \mathbb{N}$;
6. $(-1,1) \in \{-1\{\times\{1\}, \{-1,1\}^2$;
7. $\mathbb{N} \in \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}\}$;
8. $\mathbb{R} \notin \mathbb{R}^2$;
9. $\mathbb{N} \notin \mathbb{R}$;
10. $\emptyset \in \{\emptyset\}$;
11. $\emptyset \notin \emptyset$;
12. $\{1,2,3\} \subseteq \{1,2,3,4\}$;
13. $\{1,2,3\} \notin \{1,2,3,4\}$;
14. $\left\{^{\{1,2,3\} \in \{\{1,2,3\},1,2,3,4}\right\}$;
15. $\mathbb{N} \subseteq \mathbb{N} \cup \{-1\}$;
16. $\{\emptyset\} \notin \emptyset$.

Quelques propriétés très intuitives.

Propriété. ***(Non-transitivité de l'appartenance)***

La relation d'appartenance n'est pas transitive (mais pas intransitive).

▷ Il suffit de prendre pour exemples : $1 \in \{1,2\} \in \{\{1,2\}, 2\}$, mais $1 \notin \{\{1,2\}, 2\}$. La construction et structure des ensembles entiers naturels sera justifiée plus tard. ■

Propriété. ***(Non-totalité de l'appartenance)***

La relation d'appartenance est partielle.

▷ On a $\emptyset \notin \{\{\emptyset\}\}$ et $\{\{\emptyset\}\} \notin \emptyset$ (le vérifier soi-même). ■

Remarques.

1. Par essence de la théorie des ensembles, tout objet est ensemble. Un problème émerge : il n'y a pas de distinction absolue entre les ensembles et leurs éléments. On comprend que $\mathbb{N}$, $\mathbb{Z}$,..., $\mathbb{R}$, soient des ensembles, mais on conçoit mal que $1,2, \pi, x \longmapsto x^2, \mathbb{P}, \int_1^2 e^{it}\mathrm{d}t, \int_1^{+\infty} e^t \mathrm{d}t$ le soient également. C'est pourtant le cas.
2. L'axiome de fondation, noté AF dans la suite, répond aux questions suivantes : $\in$ est-elle réflexive ? et $\in$ est-elle symétrique, antisymétrique ? On propose au lecteur de se poser la question lui-même avant de trouver la réponse dans la suite.

Exercice 5

Expliciter les ensembles 1,2,π, $x \longmapsto x^2$, $\mathbb{P}$, $\displaystyle\int_1^2 e^{it}\mathrm{d}t$.

3.2.2.1.2 Axiomes déjà connus quant à $\in$

Axiome. ***(Principe d'extensionnalité)***

Pour tous A,B : Ens, $(A = B) \iff (\forall x \quad (x \in A \iff x \in B))$.

Remarques.

1. C'est le premier axiome.
2. C'est en fait une définition, celle de « $=$ ». Cette relation d'égalité est définie entre les ensembles, et l'on peut garder en mémoire que toutes les relations d'égalités utilisées quotidiennement sont des *restrictions* de cette relation définie sur une classe impropre.
3. Le syntagme « $\forall x$ » seul apparaissant dans la formulation de l'axiome n'est pas un abus, mais la notation la plus correcte pour : « pour tout ensemble x... ». Quand on note, $\forall x \in \mathbb{R} \quad \mathfrak{Re}(x) = x$, on raccourcit la plus correcte : $(\forall x, x \in \mathbb{R} \implies \mathfrak{Re}(x) = x)$.

Un petit rappel qui ne fait pas de mal.

Définition. ***(Inclusion)***

Pour tous A,B : Ens, on dit que $A \subseteq B$ si $\forall x \quad x \in A \implies x \in B$.

Remarque. L'idée fondamentale de l'inclusion est son rapport avec l'appartenance : l'appartenance $\in$ est un notion locale, alors que l'inclusion $\subseteq$ est une notion globale.

On donne quelques applications du principe d'extensionnalité, sachant qu'un peu de raisonnement axiomatique n'est pas luxueux. Dans le théorème suivant, l'existence d'un ensemble des parties en conséquence de l'axiome d'existence d'ensemble des parties. Nous ne donnons pas tous les axiomes, au risque de faire inventaire. Le lecteur intéressé peut les trouver à l'adresse : http://math.univ-lyon1.fr/~melleray/AnnexeA.pdf.

Théorème. ***(Identité d'ensembles par les ensembles des parties)***

Pour tous A,B : Ens, $A = B \iff \mathcal{P}(A) = \mathcal{P}(B)$.

▷ Si $A = B$, $\forall x \quad x \in A \iff x \in B$. Or $\forall X \quad X \subseteq A \iff \forall x \in X \quad x \in A$, donc on vérifie : $\forall X \quad X \subseteq A \iff X \subseteq B$, soit par définition $\mathcal{P}(A) = \mathcal{P}(B)$: c'est l'extensionnalité, car $\mathcal{P}(A) = \{X \mid X \subseteq A\}$. Réciproquement, si $A \neq B$, il existe $x \in B$, $x \notin A$ (ou $x \in A, x \notin B$, cas qui se traite de la même manière). Dans le premier cas, $\{x\} \subseteq B$ mais $\{x\} \not\subseteq A$, car sinon

$x \in A$. Ainsi $\{x\} \in \mathcal{P}(B)$ mais $\{x\} \notin \mathcal{P}(A)$, donc par extensionnalité $\mathcal{P}(A) \neq \mathcal{P}(B)$. Par contraposée, $\mathcal{P}(A) = \mathcal{P}(B) \implies A = B$, d'où l'équivalence. ■

Remarque. Une chose remarquable du théorème, et que l'égalité des ensembles des parties n'est qu'une égalité d'ensembles, et pas une correspondance deux à deux des parties.

Théorème. ***(Singletons associés)***

$\forall A,B \quad A = B \iff \{A\} = \{B\}$.

▷ D'une part, on suppose $A = B$. Soit $x \in \{A\}$. Alors puisque c'est un singleton, $x = A$. Or $A = B$, donc $x = B$, et $B \in \{B\}$ donc $x \in \{B\}$. Ainsi $\{A\} \subseteq \{B\}$. Semblablement, $\{B\} \subseteq \{A\}$, donc par double inclusion $\{A\} = \{B\}$.

Réciproquement, supposons $\{A\} = \{B\}$. $A \in \{A\}$ et $\{A\} = \{B\} \iff (\forall x \in \{A\} \quad x \in \{B\} \wedge \forall x \in \{B\} \quad x \in \{A\})$. Le premier point de la conjonction, avec la première remarque faite donne $A \in \{B\}$, soit $A = B$, car $\{B\}$ est un singleton. ■

Remarques.

1. De même que pour le théorème précédement, l'existence du singleton contenant A est axiomatique. (Elle vient de l'axiome... de la paire. Il n'y pas d'axiome du singleton : pour l'en déduire, il suffit de prendre, dans la paire, les deux éléments identiques, et l'on se rend compte qu'un axiome de singleton serait superflu, car il est déjà constructible à partir de l'axiome de la paire.)
2. La contraposée du théorème donne : $A \neq B \iff \{A\} \neq \{B\}$.

Propriété. ***(Partition triviale par événements atomiques, partition discrète)***

Soit Ω un ensemble. Alors $(\{x\})_{x\in\Omega}$ partitionne Ω.

▷ Il suffit de reprendre point par point la définition de partition.

⋆ **Habitations.** Soit $x \in \Omega$, c'est-à-dire $\{x\}$ dans la partition. $\text{card}(\{x\}) = 1 \neq 0$, donc les parties de la partition ne sont pas vides.

⋆ **Disjonction deux à deux.** C'est la contraposée du théorème des singletons associés.

⋆ **Réunion.** $\bigcup_{x\in\Omega} x = \Omega$. En effet : $X \in \bigcup_{x\in\Omega} x \iff \exists x \in \Omega \quad x = X$, soit $\bigcup_{x\in\Omega} x \iff x \in \Omega$. Ainsi, on a l'égalité par extensionnalité,

ce qui termine la preuve. ■

Remarque. On aurait pu le démontrer beaucoup plus rapidement : l'égalité sur E est une relation d'équivalence, dont les classes sont les $(\{x\})_{x\in\Omega}$ et d'après le théorème fondamental des relations d'équivalence, c'est une partition de Ω.

3.2.2.1.3 Clarification de l'ambivalence entre ensemble et élément

On a vu que tout élément est en fait un ensemble complétement, et que la distinction entre les deux n'est réellement qu'un agrément de langage. Dans l'assertion :

$$E \in F,$$

E,F sont des ensembles, mais on dit plutôt que E est un élément, à savoir un élément de F. D'autre part, tout ensemble peut être vu comme un élément, on l'a déjà remarqué, car $\forall x \quad x \in \{x\}$, ce qui justifie de confondre le concept d'élément avec celui d'ensembles. Nous voulons, dans ce qui suit, corriger les imprécisions mentales issues de l'ambivalence entre les deux termes.

Propriété. ***(Inclusions générales)***

(i) Tout ensemble inclut un ensemble ;
(ii) tout ensemble est inclus dans un ensemble ;
(iii) si $\mathrm{card}(E) \geqslant 1$, un ensemble E inclut un autre ensemble ;
(iv) dans le cas général, tout ensemble est inclus dans un autre ensemble.

▷ Successivement :

1. $E \subseteq E$;
2. $E \subseteq E$;
3. $\emptyset \subseteq E$, et $E \neq \emptyset$, car, par hypothèse, il est de cardinal non nul, et l'ensemble vide existe d'après un axiome ;
4. Soit F un ensemble, $F \notin E$. Il existe d'après le paradoxe de Cantor. Alors $E \subseteq E \cup \{F\} = G$, mais $G \neq E$, car $F \in G$ mais $F \notin E$. ■

En général, un élément d'un ensemble n'en est pas une partie. Par exemple $\{1\} \in \{\{1\}\}$, mais $1 \notin \{\{1\}\}$ donc $\{1\} \not\subseteq \{\{1\}\}$. Ceci n'est pas universel, c'est même faux dès que $E \ni \emptyset$ (pourquoi ?). Réciproquement, une partie d'un ensemble, en général, ne lui appartient pas : $\{1,2\} \notin \mathbb{N}$ alors que $\{1,2\} \subseteq \mathbb{N}$.

Propriété. ***(Remarques supplémentaires)***

(i) Un ensemble n'est pas nécessairement disjoint de l'ensemble de ses parties ;
(ii) on peut avoir $E \subseteq F$ et $E \in F$ même si $E \neq \emptyset$;
(iii) pour tout E, il existe F tel que $E \subseteq F$ et $E \in F$.

▷ (i) et (ii) ont déjà été traités ci-dessus. Pour (iii), il suffit de choisir $F = E \cup \{E\}$. ■

Remarquons que $E \cup \{E\} \neq E$: cette propriété universelle découle de l'axiome de fondation ; avant de l'introduire, on rend compte des implications d'une trop grande liberté dans les constructions relatives relation d'appartenance.

3.2.2.2 Bizarreries de la relation d'appartenance

Formalisons les conséquences des propriétés juste précédentes, pour en montrer les limites. Le premier objet bizarre engendré par de telles constructions est celui d'*ensemble transitif*.

3.2.2.2.1 Ensembles transitifs

On a vu que $\in$ n'était pas a priori transitive, mais également, qu'elle n'était pas pour autant intransitive. Les ensembles transitifs sont tels que la transitivité est toujours vraie, lorsqu'on ne regarde qu'eux, un par un.

Définition. ***(Ensemble transitif)***

Un ensemble est dit *transitif* si les éléments de ses éléments en sont tous des éléments, autrement dit, A est transitif si et seulement si $\forall x \in A \; \forall a \in X \quad x \in A$.

Exemples

1. $\emptyset$ est transitif. En effet, $\forall x \in \emptyset \; \forall a \in x \quad x \in \emptyset$, car toute propriété commençant par $\forall x \in \emptyset$ est vraie par principe d'explosion ;
2. $\emptyset \cup \{\emptyset\}$ est transitif (le vérifier) ; ainsi l'ensemble vide n'est pas le seul transitif.

La notion d'ensemble transitif n'est pas du tout intuitive : elle montre les pathologies de la relation d'appartenance. AF n'interdit pas les ensembles transitifs, mais il les limite[1], en pratique, à ceux que nous voyons maintenant, c'est-à-dire, l'ensemble vide et ses composés selon le théorème suivant.

L'ensemble des parties d'un ensemble A est noté indifféremment $\mathcal{P}(A)$ ou $\mathfrak{b}(A)$.

Propriété. ***(Caractérisation de la transitivité par l'ensemble des parties)***

Un ensemble A est transitif si et seulement si $A \subseteq \mathfrak{b}(A)$.

▷ C'est une simple reformulation de la définition. Le lecteur un peu perdu aura intérêt à rédiger l'équivalence. ■

Théorème. ***(Construction des ensembles transitif)***

Soit A un ensemble transitif. Alors :

- $A \cup \{A\}$ est transitif ;
- $\mathcal{P}(A)$ est transitif.

▷ Successivement :

[1] En effet, si $E \in E$, alors E est automatiquement transitif.

1. On fait une disjonction des cas : si $x \in A \cup \{A\}$, soit $x \in A$, dans ce cas, on applique la transitivité de A, donc pour tout $a \in x$, $x \in A$ donc $x \in A \cup \{A\}$. Si d'autre part $x \in \{A\}$, alors $x = A$, donc si $a \in x = A$, $a \in A$ donc $a \in A \cup \{A\}$ de même.
2. Si $x \in a \in \mathcal{P}(A)$, $x \in a \subseteq A$, soit $x \in A$ par définition de l'inclusion, donc par définition $x \subseteq A$, soit $x \in \mathcal{P}(A)$ et $\mathcal{P}(A)$ est transitif. ■

On espère avoir assez brouillé les esprits croyant l'apparente commodité de la relation d'appartenance. Avant de passer à l'énoncé de AF, on rappelle le paradoxe suivant, beaucoup utile.

3.2.2.2.2 Paradoxe de Russell

Soit la collection des objets : $\{X \mid X \notin X\} = \mathcal{C}$.

Paradoxe. ***(Paradoxe de Russell)***

La construction de $\mathcal{C}$ est paradoxale.

▷ Par principe logique de tiers-exclu, on a, soit $\mathcal{C} \in \mathcal{C}$, soit $\mathcal{C} \notin \mathcal{C}$. Supposons pour commencer $\mathcal{C} \in \mathcal{C}$. Alors par définition de $\mathcal{C}$, $\mathcal{C} \notin \mathcal{C}$, car $\mathcal{C}$ est un X tel que $X \in \mathcal{C}$. Inversement, supposons que $\mathcal{C} \notin \mathcal{C}$. Par définition, $\mathcal{C}$ contient tous les X tels que $X \notin X$, et $\mathcal{C}$ vérifie ce prédicat logique, donc $\mathcal{C} \in \mathcal{C}$. Ainsi, dans les deux cas logiques possibles, on a ($\mathcal{C} \in \mathcal{C}$ ET $\mathcal{C} \notin \mathcal{C}$), donc cette proposition est universellement vraie. Or elle est universellement fausse selon le principe de non-contradiction, donc la propriété « la propriété $\mathcal{C} \in \mathcal{C}$ capiet $\mathcal{C} \notin \mathcal{C}$ est vraie » est la fois vraie et fausse, ce qui est contradictoire par non-contradiction. ■

Le constat du paradoxe de Russell a conduit à la création d'un des premiers axiomes de la théorie des ensembles, les schémas de séparation : on ne peut pas construire des ensembles à partir de rien, mais il y a une règle : si E existe, et $\mathcal{P}$ est un prédicat logiquement bien formé sur E, alors on peut considérer l'ensemble $\{x \in E \mid \mathcal{P}(x)\}$, d'où le nom de définition par séparation et compréhension. C'est la raison pour laquelle il faut écrire toujours : $\forall x \in \mathbb{R}...$, $\forall f \in \mathbb{C}^{\mathbb{Z}}...$ (Notons qu'il faut également postuler l'existence d'au moins un ensemble, ce que l'on fait avec l'axiome de l'ensemble vide.)

La faute commise dans le paradoxe de Russell est d'avoir postulé l'existence de l'ensemble $\mathcal{C}$: il n'est pas défini par séparation, donc a priori, il n'existe pas. En effet, l'on sait que $\mathcal{C}$, d'après l'axiome de fondation, est Ens la classe de tous les ensembles, et d'après le théorème de Cantor, ce n'est pas un ensemble, autrement dit une classe impropre.

3.2.2.3 L'axiome de fondation proprement dit

⟶ **Notation.** On notera $ZF_\bullet$ l'ensemble des axiomes de Zermelo-Fraenkel sans l'axiome de fondation, et $ZF = ZF_\bullet + AF$.

On rappelle les quelques interrogations initiales de cette section :

⋆ $\in$ est-elle réflexive ?

⋆ $\in$ est-elle symétrique, antisymétrique, ou sous quelles conditions ?

c'est-à-dire, existe-t-il, et lesquels, des ensembles E,F tels que :

⋆ $E \in E$?

⋆ $E \in F$ et $F \in E$?

3.2.2.3.1 Énoncé et premières propriétés

Axiome. ***(Axiome de fondation)***

Pour tout ensemble A, $A \neq \emptyset \implies \exists B \in A \; B \cap A = \emptyset$.

Remarques.

1. C'est chelou.
2. On verra que ceci exprime que la relation $\in$ est *bien fondée* sur la classe des ensembles, c'est-à-dire, qu'il existe toujours sur tout ensemble et l'ensemble des éléments qu'il contient, et l'ensemble des éléments qu'ils contiennent, etc., un élément minimal pour l'appartenance, et donc qu'il n'existe pas de suite infinie du type $E \ni F \ni G \ni ...$ Nous verrons même que, modulo un axiome du choix spécial, cette assertion est équivalente à AF.

Théorème. ***(Irréflexivité de $\in$)***

Pour tout ensemble E, $E \notin E$.

▷ Si $E \in E$ pour un certain ensemble E, notons $A = \{E\}$. Dans ce cas, $A \neq \emptyset$, car A est un singleton donc de cardinal $1 \neq 0$. De plus, si $x \in A$, $x \in \{E\}$ soit $x = E$ et $x \cap A \neq \emptyset$, car $x \cap A = E \cap \{E\}$, et puisque $E \in \{E\}$ et $E \in E$ par hytpohèse, on aurait $E \in E \cap \{E\}$, ce qui contredirait AF. Par contraposition, $AF \implies \forall E,\ E \notin E$, c'est-à-dire $\nexists E,\ E \in E$. ■

Remarques.

1. Ce théorème d'irréflexivité se réécrit en : il n'existe pas d'ensemble E tel que $E = \{E\}$.
2. On en déduit immédiatement ce que l'on a évoqué tout à l'heure : pour tout ensemble E, $E \cup \{E\} \neq E$. En effet, cela voudrait dire que, comme $E \in E \cup \{E\}$, $E \in E$.
3. On constate que l'irréflexivité est encore vérifiée pour l'ensemble vide. En effet, $\emptyset$ ne peut contenir aucun élément, y compris $\emptyset$.

On arrive désormais aux équivalences centrales de cette partie, qui donnent tout son sens à la formulation initiale un peu obscure de l'axiome de fondation. Avant cela, on rappelle l'énoncé de l'axiome du choix dépendant qui servira ensuite pour établir une formulation équivalente de AF.

Axiome. ***(Axiome du choix dépendant)***

Pour tout $X \neq \emptyset$, pour toute relation binaire $\mathcal{R}$ sur X, si le domaine de définition de $\mathcal{R}$ est bien X (autrement dit si tout élément $x \in X$ est bien tel qu'il existe un $y \in X$ tel que $(x,y) \in \mathcal{R}$), alors il existe une suite $(x_n) \in X^{\mathbb{N}}$ telle que pour tout entier naturel n, $x_n \mathcal{R} x_{n+1}$. On note cet axiome DC.

Propriété. ***(Formulations diverses de l'axiome de fondation)***

On considère les propriétés suivantes :

AF l'axiome de fondation ;

(I) l'irréflexivité de la relation d'appartenance ;

(II) « Il n'existe pas x_1,...,x_n n ensembles, $n \in \mathbb{N}^*$ tels que $x_1 = x_n$ et $x_1 \in ... \in x_n$ » ;

(III) « $\in$ est bien fondée, c'est-à-dire qu'il n'existe pas de suite $(x_n)_{n\in\mathbb{N}}$ d'ensembles décroissante pour $\in$ ».

Alors on a les implications relatives comme représentées sur la figure 3.2.1. En particulier, elles sont toutes vraies dans un système ayant pour axiome AF. Les autres sont de simples conséquences logiques les unes des autres, mais $(III) \iff AF$ en présence de l'axiome du choix dépendant.

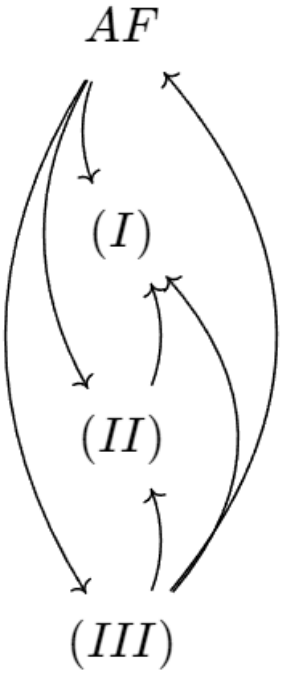


FIGURE 3.2.1 : *Formulations faibles de l'axiome de fondation.* — Illustration des implications réciproques.

▷ Montrons chacune des flèches précédentes, qui représentent des implications.

▶ $AF \implies (I)$. On l'a déjà montré, c'est l'objet de la propriété précédente.

- $AF \implies (II)$. De même, par contraposée. Supposons $x_1 \in ... \in x_n = x_1$. On pose $E = \{x_1,...,x_n\}$, qui n'est pas vide puisque $n \neq 0$. On construit ainsi un contre-exemple de l'axiome de fondation. En effet, soit $x \in E$. $x = x_i$ où $i \in [\![1,n]\!]$, car $x_1 = x_n$. $x \cap E = x_i \cap \{x_1,...,x_{n-1}\}$. Dans le premier cas, $i \neq 1$. On a $x_{i-1} \in x_i$ par hypothèse et $x_{i_1} \in \{x_1,...,x_{n-1}\}$, donc $x_{i-1} \in x \cap E$ qui est donc non vide. Dans le deuxième cas, $i = 1$. Alors $x_{n-1} \in x_i = x_n$ et $x_{n-1} \in \{x_1,...,x_{n-1}\}$ donc $x_{n-1} \in x \cap E$ qui n'est pas vide encore une fois. Ainsi $\neg AF$.
- $AF \implies (III)$. Remarquons en passant que c'est faux pour une suite qui serait croissante, il suffit de prendre $x_0 = \emptyset$ et $x_{i+1} = \{x_i\}$ pour tout $i \in \mathbb{N}$. Soit donc une suite infinie décroissante $x_0 \ni ... \ni x_n \ni ...$; l'axiome de la réunion permet de former $E = \bigcup_{n\in\mathbb{N}} \{x_n\} = \mathrm{Im}(x)$: l'existence de la suite (x_n) est hypothétique donc certaine. $E \neq \emptyset$, car $E \ni x_{47}$. Soit $x \in E$. $x \cap E = x_i \cap \bigcup_{n\in\mathbb{N}} \{x_n\}$, $i \in \mathbb{N}$ fixé choisi. On a à la fois $x_{i+1} \in x_i$ par hypothèse de chaîne et $x_{i+1} \in \{x_0,...,x_n,...\} = \bigcup_{n\in\mathbb{N}} \{x_n\}$ donc $x_{i+1} \in x \cap E$ qui est non vide. Ainsi, on nie l'axiome de fondation et l'on conclut par contraposition.
- $(III) \implies (I)$. On raisonne encore par contraposée : si $A \in A$, alors la suite constante $(A)_{n\in\mathbb{N}}$ convient pour nier (III).
- $(III) \implies (II)$. On raisonne par contraposée, en concaténant : $x_n \ni ... \ni x_1 \ni x_{n-1} \ni ... \ni x_1 \ni x_{n-1} \ni ....$ Plus précisément, on définit comme contre-exemple de (III) la suite infinie décroissante : $(u_i)_{i\in\mathbb{N}}$ telle que $u_0 = x_n$, et pour tout $i \in \mathbb{N}$, $u_i = x_{n-k-1}$ où k est le reste dans la division euclidienne de i par n.
- $(II) \implies (I)$. Il suffit de prendre $n = 1 \in \mathbb{N}^*$.
- $(III), DC \implies AF$. Toujours par contraposition. Le principe de démonstration est le suivant : nions AF. On suppose qu'il existe $A_0 \neq \emptyset$ tel que $\forall B \in A_0, B \cap A_0 \neq \emptyset$ (c'est exactement $negAF$). A_0 étant non vide, prenons $B_0 \in A_0$. $B_0 \cap A_0$ d'après ce qui précède, donc on peut prendre $A_1 \in B_0 \cap A_0$. Mais en particulier $A_1 \in A_0$, donc $A_1 \cap A_0 \neq \emptyset$ par hypothèse. Alors on peut prendre $A_2 \in A_1 \cap A_0$. Mais $A_2 \in A_0$, donc $A_2 \cap A_0 \neq \emptyset$ toujours par hypothèse, et l'on prend $A_3 \in A_2 \cap A_0$, etc. Cette intuition que l'on va pouvoir creuser à l'infini dans les éléments de B, d'où le terme de fondation, se formalise exactement avec DC. On prend, dans la définition de DC, $X = A_0 \neq \emptyset$ et pour relation $\ni$ la relation symétrique de $\in$. $\ni$ est bien définie partout sur X, en effet c'est l'hypothèse : $\forall B \in A_0 \quad B \cap A_0 \neq \emptyset$, c'est-à-dire, $\forall B \in A_0 \quad \exists x \in A_0\ x \in B$ soit $B \ni x$. Ainsi, DC s'applique et il existe une suite $(x_i)_{i\in\mathbb{N}}$ telle que $\forall i \in \mathbb{N} \quad x_i \ni x_{i+1}$, soit $\neg(III)$,

et tout est démontré. ■

Enfin, on règle le compte de la symétrie de la relation d'appartenance. On voit qu'elle n'est pas symétrique, et même asymétrique. De plus, elle n'est asymétrique que si l'un des deux ensembles est vide afin d'appliquer le principe d'explosion.

Corollaire. ***(Asymétrie de la relation d'appartenance)***

Il n'existe pas d'ensembles E,F tels que $E \in F$ et $F \in E$.

▷ C'est une conséquence de (II) pour $n = 2$. ■

Remarque. Une autre façon de le dire, est qu'aucun ensemble n'est élément d'un de ses éléments.

Questions ouvertes. Il est naturel de se demander si l'on peut clore le diagramme ci-dessous, de sorte que les quatre propositions soient en fait équivalentes : (II), et *a fortiori* (I), impliquent-ils AF ? La question n'a peut-être pas de réponse, car trouver un exemple de théorie (on dit : un *modèle*) dans laquelle les axiomes de ZF sont vérifiés est en fait impossible, c'est le théorème d'incomplétude de Gödel, et il serait problématique de montrer alors que AF et (I) ne sont pas équivalentes, d'où la difficulté de traiter $\neg AF \implies \neg(I)$.

3.2.2.3.2 Conséquences pour la construction d'objets mathématiques

Exercice 6

Montrer que $\{\{\emptyset\}\} \neq \{\emptyset\}$.

Exercice 7

Montrer que, pour tout ensemble x, $\{\{x\}\} \neq x$.

Exercice 8

Montrer que, pour tout ensemble x, pour tout $p \in \mathbb{N}^*$, $\underbrace{\{\{...\{}_{p \text{ fois}} x \underbrace{\}...\}\}}_{p \text{ fois}} \neq x$.

Exercice 9

Montrer que, pour tout ensemble x, pour tous n,$p \in \mathbb{N}^*$, $\underbrace{\{\{...\{}_{n \text{ fois}} x \underbrace{\}...\}\}}_{n \text{ fois}} \neq \underbrace{\{\{...\{}_{p \text{ fois}} x \underbrace{\}...\}\}}_{p \text{ fois}}$ si et seulement si $n \neq p$.

On va, sommairement, construire de façon ensembliste quelques-uns des objets mathématiques les plus utilisés, notamment les entiers naturels de l'ensemble $\mathbb{N}$. L'axiome de fondation permet, non de créer les entiers naturels (c'est l'axiome de l'infini qui le permet), mais de montrer que les constructions obtenues sont deux à deux distinctes, autrement, de justifier que $1 \neq 2$. Il est important de comprendre que la « représentation » ci-dessous est bel et bien une construction, c'est-à-dire une façon tout au moins de justifier l'existence de tels objets, même si, l'on en convient, ce qu'ils sont n'a pratiquement aucun intérêt à côté de leurs propriétés ; elles sont l'objet de l'arithmétique.

Définition. ***(Représentation des entiers naturels de Von Neumann)***

Nous posons : $0 = \emptyset$, puis : $1 = \{\emptyset\} = \{0\}$, puis $2 = \{0,1\} = 1 \cup \{1\} = \{\emptyset, \{\emptyset\}\}$, etc., c'est-à-dire, à l'infini (ce qui est justifié par l'axiome de l'infini) $n+1 = n \cup \{n\}$. L'existence de $\emptyset$ est garantie par l'axiome de l'ensemble vide.

John Von Neumann

D'origine hongroise, fils d'un banquier réputé, János Lajos Neumann, dit Von Neumann commence à étudier à Budapest. Enfant surdoué, il lit et mémorise tout ce qui lui tombe sous la main, parle grec et latin à l'âge de six ans. Calculateur prodige, il stupéfie ses instituteurs et les amis de la famille, dont Lipót Fejér qui dirigera sa thèse, par sa mémoire prodigieuse et ses capacités en calcul mental.

Malgré une situation politique instable en Hongrie, Neumann entreprend des études supérieures de mathématiques à Budapest en 1919 qu'il complète par trois années d'études de chimie à Berlin et Zurich. Il rencontrera ainsi Erhard Schmidt, Herman Weyl et Polya. Il s'intéresse en fait plus aux ensembles et aux nombres transfinis de Cantor qu'à la chimie... C'est à Budapest qu'il soutiendra finalement sa thèse dirigée par Fejér portant sur les ensembles transfinis, fin 1926.

Professeur à Göttingen puis à l'université de Berlin, la réputation de Neumann s'instaure outre-Atlantique : il se rend aux États-Unis à Princeton à l'invitation de Veblen en 1930 à l'occasion de la mise en place du tout nouveau Institute for Advanced Study.

Juif, afin d'échapper à la répression du pouvoir hitlérien soutenu par le régime hongrois, von Neumann s'installe définitivement aux États-Unis en 1933 et fit toute sa carrière au célèbre institut. Il meurt prématurément, en 1954, à 54 ans, d'un cancer des os sans doute causé par ses nombreuses expositions aux radiations lors des expérimentations pour la mise au point de la première bombe atomique.

L'ensemble formé par cette infinité d'ensembles est noté $\mathbb{N}$; on peut lui définir une addition, une multiplication, et vérifier qu'elles vérifient toutes les propriétés habituelles qui lui sont associés ; également un ordre qui permet d'énoncer la propriété fondamentale de $\mathbb{N}$: toute partie non vide a un minimum. C'est peu intéressant et sans révolution conceptuelle non plus ; le lecteur intéressé trouvera une construction plutôt complète dans *Épistémologie mathématique* de Henri Lombardi. De plus, on vérifie qu'il vérifie les cinq propriétés axiomatiques de l'arithmétique de Peano, que nous énonçons à titre informatif ci-dessous :

Définition. ***(Arithmétique de Peano)***

On appelle *entiers naturels de Peano*, un ensemble $\mathbb{N}$ vérifiant les propriétés suivantes, dits *axiomes de Peano* :

(i) il contient au moins un élément, notons le 0 ;
(ii) il existe une fonction σ de $\mathbb{N}$ dans $\mathbb{N}$, appelée *successeur* ;
(iii) aucun entier naturel n'est suivi par 0 ($0 \notin \mathrm{Im}(\sigma)$) ;
(iv) deux entiers naturels ayant le même successeur sont égaux (σ est injective) ;
(v) un principe de récurrence : si un ensemble contient 0 et le successeur de chacun de ses éléments, cet ensemble est $\mathbb{N}$.

Habituellement, la fonction successeur est donnée par $\sigma(n) = n + 1$.

Concluons par l'intérêt principal de cette partie.

Théorème. ***(Distinction deux à deux des entiers naturels)***

En présence de l'axiome de fondation, les entiers naturels de Von Neumann sont deux à deux distincts.

▷ On l'a déjà vu : n ne peut appartenir à n, pour tout n, donc $n + 1 \neq n = n \cup \{n\}$. Ceci montre que deux entiers successifs sont distincts. Pour montrer que les entiers naturels sont deux à deux distincts, il s'agit simplement le principe du quatrième exercice présenté ci-dessus, qui en découle. ■

Nous espérons que, par ces considérations, le lecteur sera convaincu que la totalité des objets mathématiques qu'il manipule est une construction ensembliste : un ordre, par exemple, est une relation sur, disons, $\mathbb{N}$, c'est-à-dire une partie du produit $\mathbb{N} \times \mathbb{N}$: la notation $n \leqslant p$ traduit simplement $(n,p) \in \leqslant$. Une fonction f est un triplet $f = (E,F,\Gamma)$, où E est l'ensemble de départ, F l'ensemble d'arrivée et Γ une partie de $E \times F$ vérifiant la propriété fondamentale des fonctions : tout élément a au plus une image. Les nombres réels sont identifiés, par exemple, aux coupures de Dedekind : ce sont alors des couples (A,B) tels que $A,B \subseteq \mathbb{Q}$, $AcupB = \mathbb{Q}$, $A \cap B = \emptyset$ et $\forall a \in A\ \forall b \in B \quad a < b$. Et ainsi de suite.

3.2.2.3.3 Considérations logiques

On peut montrer que si $ZF_\bullet$ est consistant, *i. e.* s'il n'y a pas d'incohérence dans ses axiomes et qu'un modèle est envisageable, alors il ne prouve ni AF, ni sa négation : on dit que AF est indépendant des axiomes de $ZF_\bullet$. Cela s'exprime :

$$ZF_\bullet \text{ consistant} \implies ZF \text{ consistant} .$$

Méthode. ***(Étudier un ensemble)***

- S'il est partie ou sur-ensemble d'un autre ensemble ?
- Sous-parties remarquables
- Déterminer son cardinal
- Dénombrer certaines parties spéciales (c'est-à-dire déterminer leur cardinal)
- Est-il ensemble des parties d'un certain ensemble ? Se réalise-t-il naturellement comme tel ou partie ?

3.3 Cardinalité

3.3.1 Théorème de Cantor-Bernstein

Théorème. ***(Théorème de Cantor-Bernstein)***

Soient A et B deux ensembles. S'il existe une injection de A dans B, et s'il existe une injection de B dans A, alors A et B sont en bijection. Autrement dit, la relation $\hookrightarrow$ est antisymétrique.

▷ Il existe un grand nombre de preuves du théorème de Cantor-Bernstein ; nous en donnons une constructive et qui ne fait pas recours à l'axiome du choix. Soient A et B deux ensembles, dont on suppose qu'il existe une application injective $f : A \longrightarrow B$, et d'autre part qu'il existe une application $g : B \longrightarrow A$ injective. Ce sont des applications, c'est-à-dire que tout élément de leur départ admet une image.

Remarquons que la corestriction $\tilde{g} : B \longrightarrow \mathrm{Im}(g) \subseteq A$, qui à un élément de B fait correspondre son image par g, est toujours injective, et surjective par construction. C'est donc une bijection. Si l'on exhibe une bijection $h : A \longrightarrow \mathrm{Im}(g)$, c'est-à-dire un bijection de A sur son sous-ensemble $\mathrm{Im}(g)$ a priori strict, alors la fonction $h^{-1} \circ \tilde{g} : B \longrightarrow A$ est une bijection de B dans A et A et B sont en bijection.

Pour construire h, on introduit la suite $(A_n)_{n\in\mathbb{N}}$ de parties de A en posant $A_0 = \complement_A \mathrm{Im}(g)$, et pour tout $n \in \mathbb{N}^*$, $A_{n+1} = g \circ f(A_n)$. De façon immédiate, on a, pour tout entier naturel n, $A_n = (g \circ f)^n(A_0)$. Démontrons d'abord que les A_n sont toutes deux à deux disjointes. Si i est un entier naturel non nul, alors $A_i = (g \circ f)^i(A_0) = g(f \circ (g \circ f)^{i-1})(A_0)$. Par suite, $A_i \subseteq \complement_A A_0$, ce qui signifie exactement que A_i et A_0 sont disjoints. Soit maintenant un entier naturel n. Par composition, $g \circ f$ est une injection, puis encore $(g \circ f)^n$ est injective. En composant l'intersection $A_0 \cap A_i = \emptyset$ par une injection, on obtient l'inclusion $(g \circ f)^n(A_0) \cap (g \circ f)^n(A_i) \subseteq g \circ f(\emptyset) = \emptyset$, l'image de l'ensemble vide par une application étant toujours vide. Ainsi $(g \circ f)^n(A_0) \cap (g \circ f)^n(A_i) = \emptyset$. Or $(g \circ f)^n(A_0) = A_n$ et $(g \circ f)^n(A_i) = A_{n+i}$, donc A_n est disjoint de A_{n+i} pour tous $n \geq 0, i > 0$. Par conséquent, les $A_n, n \in \mathbb{N}$, sont deux à deux disjoints.

Cette construction permet d'écrire que : $g \circ f(\bigcup_{n=0}^{+\infty} A_n) = \bigcup_{n=0}^{+\infty} g \circ f(A_n) = \bigcup_{n=1}^{+\infty} A_n$. (Le caractère injectif n'intervient pas dans cette égalité.)

Définissons la fonction h par :

$$h : \begin{array}{ccc} A & \longrightarrow & \mathrm{Im}(g) \\ a & \longmapsto & \begin{cases} g \circ f(a) \text{ si } a \in \bigcup_{n=0}^{+\infty} A_n \\ a \text{ sinon.} \end{cases} \end{array}$$

Vérifions que h est une application bijective.

- ⋆ Elle est bien définie partout sur son ensemble de définition.
- ⋆ Elle est aussi à valeurs dans $\mathrm{Im}(g)$, puisque par construction, $g \circ f$ envoie $\bigcup_{n=0}^{+\infty} A_n$ sur $\bigcup_{n=1}^{+\infty} A_n \subseteq \mathrm{Im}(g)$, et que si $a \notin \bigcup_{n=0}^{+\infty} A_n$, alors en particulier $a \notin A_0 = \complement_A \mathrm{Im}(g)$ donc $a = h(a) \in \mathrm{Im}(g)$.
- ⋆ L'injectivité provient de ce que d'abord $g \circ f$ est une injection. Par suite, $id_{|\complement_A \bigcup_{n=0}^{+\infty} A_n}$ et les $(g \circ f)_{|A_n}$ sont des injections, par restriction. De plus, ces injections sont à images disjointes d'après ce que nous avons montré précédemment, car les A_0,...,A_n,... sont deux à deux disjointes et toutes dans $\bigcup_{n=0}^{+\infty} A_n$ qui est disjoint de $\complement_A \bigcup_{n=0}^{+\infty} A_n$.
- ⋆ D'autre part, on a dit que $g \circ f$ envoie $\bigcup_{n=0}^{+\infty} A_n$ sur $\bigcup_{n=1}^{+\infty} A_n$, ce qui garantit la surjectivité. En effet, si $y \in \mathrm{Im}(g)$, alors $y \notin A_0$, et l'on a :

  1er cas. $y \in \bigcup_{n=1}^{+\infty} A_n$. Alors la remarque précédente donne l'existence d'un antécédent dans $\bigcup_{n=0}^{+\infty} A_n \subseteq A$.

  2nd cas. $y \notin \bigcup_{n=1}^{+\infty} A_n A_n$. Alors $y \notin \bigcup_{n=0}^{+\infty} A_n$, donc $h(y) = y$ et l'antécédent y convient.

Ainsi h est une bijection, ce qui permet de conclure. ■

Exercice 10

1. Montrer que $\sum \frac{\sin(n)}{n}$ est semi-convergente.
2. En admettant le théorème de Cantor-Bernstein, montrer que $\mathfrak{S}(\mathbb{N})$ a exactement la puissance du continu.

3.3.2 Arithmétique cardinale

Exercice 11

Montrer qu'un ensemble E non vide est infini si et seulement si l'on a conjointement $E^E \simeq \mathcal{P}(E)$ et $\mathrm{card}(E) \neq 2$.

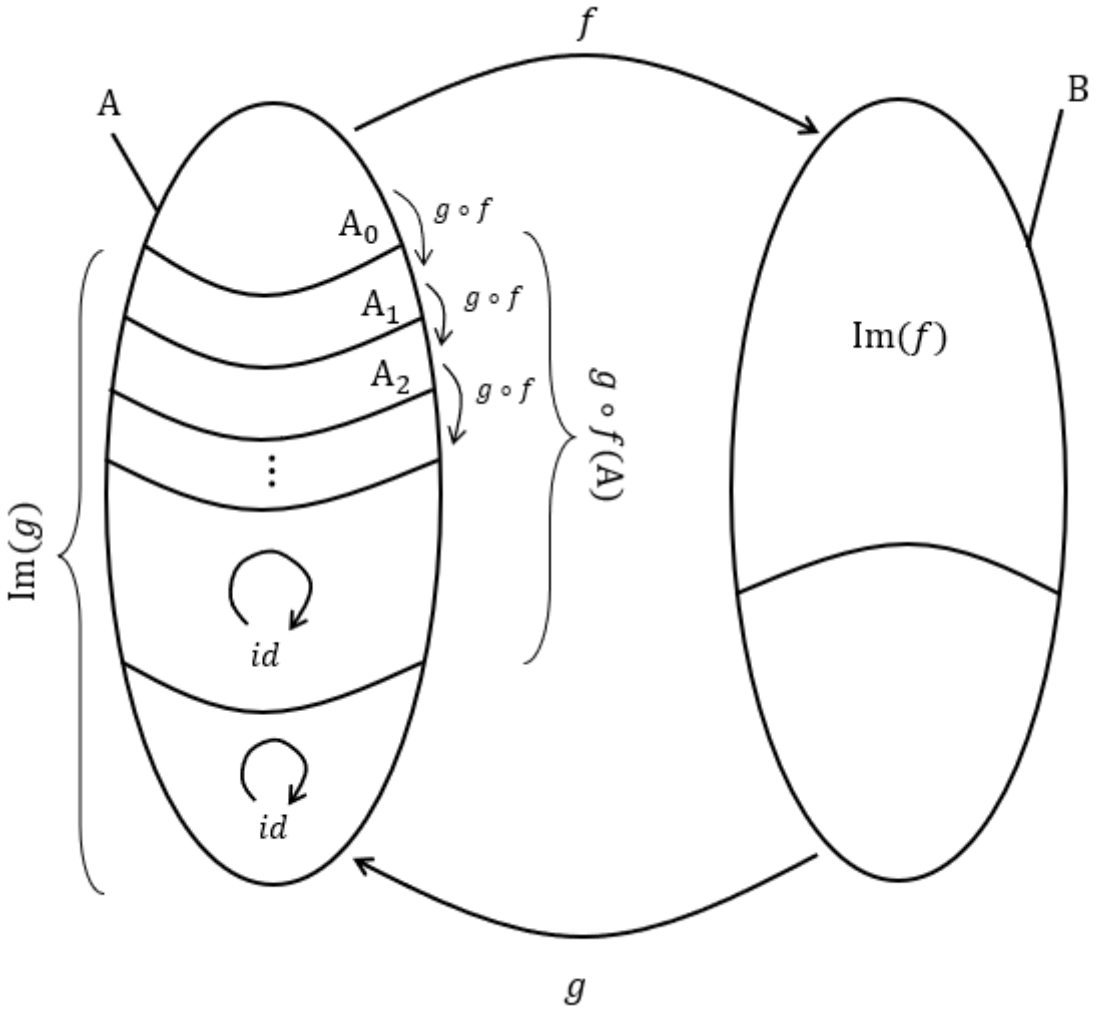


FIGURE 3.3.1 : *Illustration de la preuve du théorème de Cantor-Bernstein.* —

3.3.3 Dénombrabilité

On suppose connus les fondements sur les ensembles de la première année de classe préparatoire. Quelques références sûres pour ces notions sont données en bibliographie.

Exercice 12

1. Montrer que $\{0,1\}^{\mathbb{N}}$ est bijection avec $\mathcal{P}(\mathbb{N})$.
2. Montrer que tout intervalle non trivial est en bijection avec $\mathbb{R}$.
3. Montrer que $\mathcal{P}(\mathbb{N})$ et $\mathbb{R}$ sont en bijection.

Nous ne nous donnons pas en objet de former ici un cours complet sur les cardinaux dénombrables, mais seulement un complément de cours rudimentaire sur lequel les habitudes des classes préparatoires sont lacunaires : deux démonstrations non seulement au programme, mais dont les méthodes sont réutilisables. On rappelle d'abord une convention fluctuante.

Définition. ***(Dénombrabilité)***

Deux définitions coexistent pour la dénombrabilité :

- La convention faible : un ensemble est *dénombrable* par définition si et seulement s'il est en bijection avec une partie de $\mathbb{N}$, ou, ce qui est équivalent, s'il s'injecte dans $\mathbb{N}$. Dans ce cas, *infini dénombrable* signifie exactement « en bijection avec $\mathbb{N}$ ». Un ensemble dénombrable est donc soit fini (un ensemble est fini si et seulement s'il est en bijection avec un certain $[\![1,n]\!]$, $n \in \mathbb{N}$) soit infini dénombrable.
- La convention forte : *dénombrable* signifie maintenant « en bijection avec $\mathbb{N}$ » et le terme *au plus dénombrable* couvre les ensembles finis ou dénombrables au sens fort. *Infini dénombrable* est alors un pléonasme. **C'est la convention du programme donc nous nous efforcerons de nous y soumettre.**

Dans les deux cas, on réserve le mot *indénombrable* pour les cardinaux plus grands que $\mathbb{N}$ (on dit qu'un cardinal A est plus grand que B si l'ensemble B s'injecte dans A).

Exercice 13

Montrer que $\mathbb{Q}$ est dénombrable.

Remarque. Un ensemble est infini si et seulement s'il n'est pas fini. On dit qu'un ensemble est *infini au sens de Dedekind* s'il est équipotent à l'une de ses parties strictes, ou, ce qui est équivalent, s'il est plus grand que $\mathbb{N}$, c'est-à-dire s'il contient une partie infinie dénombrable. Dans la théorie des ensembles classiques, tout ensemble infini au sens de Dedekind est infini. La réciproque n'est pas démontrable dans le système ZF où ne figure pas l'axiome du choix, mais dès lors qu'on l'y rajoute (ou du moins sa forme dite dénombrable), les deux notions sont équivalentes. Dans ce cas, un ensemble est infini si et seulement s'il en existe une suite d'éléments deux à deux distincts (c'est alors une injection de $\mathbb{N}$ dans cet ensemble), et $\mathbb{N}$ est le plus petit infini (*i.e.* $\mathbb{N}$ s'injecte dans tout ensemble infini).

Exercice 14

Montrer que le complémentaire d'un ensemble fini dans un ensemble infini est infini.

Exercice 15

1. *(Théorème de Cantor)* Soit E un ensemble. Montrer que E et $\mathcal{P}(E)$ ne peuvent être en bijection.
2. Quel est le cardinal de l'ensemble des parties d'un ensemble fini ?
3. Montrer qu'il n'existe pas d'ensemble des parties d'aucun ensemble qui soit infini dénombrable.

Exercice 16

On admet le résultat de cours de la partie suivante.

1. Montrer qu'un ensemble E est infini au sens de Dedekind, si et seulement si $\mathbb{N} \hookrightarrow E$ (que l'on formule : $\mathbb{N}$ s'injecte dans E, ou encore E contient une copie de $\mathbb{N}$).
2. Établir que cette condition est équivalente à ce qu'il existe une suite d'éléments de E deux à deux distincts.
3. Montrer que tout ensemble infini au sens de Dedekind est infini au sens classique.
4. L'axiome du choix dénombrable annonce que si $(E_n)_{n\in\mathbb{N}}$ sont des ensembles non vides, il existe une suite x telle que pour tout n, $x_n \in E_n$. Montrer que modulo cet axiome, tout ensemble infini est infini au sens de Dedekind.
5. En déduire que $\mathbb{N}$ est le plus petit ensemble infini.

3.3.3.1 Parties de $\mathbb{N}$

Le théorème suivant, explicitement au programme, permet de décrire le cardinal de toutes les parties de $\mathbb{N}$, dont on a vu qu'il était le premier infini : elles sont soient finies, soit automatiquement dénombrables dans la convention que nous avons fixé. On verra brièvement qu'une telle taxonomie n'est généralement plus possible entre les parties des cardinaux infinis.

Exercice 17

Montrer qu'une partie de $\mathbb{N}$ est infinie si et seulement si elle est non majorée.

Propriété. ***(Axiome du bon ordre de $\mathbb{N}$)***

Toute partie non vide de $\mathbb{N}$ a un minimum.

▷ Tout dépend de ce que l'on pose comme axiome de $\mathbb{N}$. ■

Propriété. ***(Cardinal des parties de $\mathbb{N}$)***

Toute partie de $\mathbb{N}$ est au plus dénombrable, autrement dit, toute partie de $\mathbb{N}$ est soit finie, soit infinie dénombrable, ou encore, tout partie infinie de $\mathbb{N}$ est en bijection avec $\mathbb{N}$.

▷ On choisit cette dernière formulation. Soit A une partie infinie de $\mathbb{N}$. On veut former une bijection de $\mathbb{N}$ sur A, c'est-à-dire une suite bijective. On la définit par récurrence de la manière suivante : on pose $u_0 = \min(A)$ et $u_0,\ldots,u_n$ étant déjà construits, on pose $u_{n+1} = \min\{x \in A \mid x > u_n\}$. Montrons que cette suite est bien définie, à valeurs dans A, injective et surjective.

⋆ La suite est bien définie, car u_n n'est pas un majorant de A : en effet, si c'était le cas, A serait majorée et une partie de $\mathbb{N}$ est finie si et seulement si elle est majorée. Ainsi $\{x \in A \mid x > u_n\}$ est une partie non vide de A, partie de $\mathbb{N}$, donc par axiome, elle admet un plus petit élément. Remarquons que par construction $u_{n+1} > u_n$ $\quad (*)$.

⋆ Le minimum d'une partie lui appartient, donc u_n appartient à la partie de A définie ci-dessus donc à A. De même u_0 est dans A, donc la suite est bien définie en image également.

⋆ D'après la remarque $(*)$, u est strictement croissante donc en particulier injective.

⋆ Il reste à montrer que u est surjective. Soit $a \in A$. On considère $A_0 = \{a_n \mid n \in \mathbb{N}^*\}$. Par injectivité, cette partie de $\mathbb{N}$ est infinie donc non majorée, en particulier non majroée par a. L'ensemble A_1 des majorants de A_0 est donc une partie non vide de $\mathbb{N}$ et admet donc par axiome un plus petit élément que nous noterons $a_n = \min(A_1)$. Alors par définition du minimum, $a_n > a$ et $a_{n-1} \leqslant a$. Supposons un instant que cette dernière inégalité soit stricte. Alors par construction, $a_n = \min\{x \in A \mid x > a_{n-1}\}$ et par cette dernière hypothèse, $a_n \leqslant a$ par définition du minimum et $a \in A$ appartenant à ce dernier ensemble. C'est absurde avec $a_n > a$, donc il y a égalité : $a = a_{n-1}$ ce qui donne un antécédent par la suite $a : n - 1$.

Ceci conclut la démonstration. ■

Remarques.

1. Cette méthode est exactement la même que celle qui permet de montrer ce résultat souvent passé sous silence : étant donnés $x_1, ..., x_n$ n réels deux à deux distincts, il existe une unique permutation $\sigma \in \mathfrak{S}_n$ telle que $x_{\sigma(1)}, ..., x_{\sigma(n)}$ les ordonne dans l'ordre croissant.
2. Des arguments semblables permettent de montrer ce que nous appelons personnellement *lemme de recouvrement croissant*[1] : étant donné I un ensemble et $(J_n)_{n\in\mathbb{N}}$ un recouvrement ($\bigcup_{n\in\mathbb{N}} J_n \supseteq I$) de parties de I (il y a donc égalité dans le recouvrement) croissant ($J_n \subseteq J_{n+1}$ pour tout $n \in \mathbb{N}$), la suite des couronnes définie par $I_0 = J_0$ et pour tout $n \geqslant 1$, $I_n = J_n \setminus J_{n-1}$ forme une partition à parties éventuellement vides (*i.e.* un partage) de I.

Exercice 18

1. Quel est le cardinal de l'ensemble des parties finies de $\mathbb{N}$?
2. Quel est le cardinal de l'ensemble des parties infinies de $\mathbb{N}$?

Exemple fondamental. *(Suite des nombres premiers)*

L'ensemble des nombres premiers $\mathcal{P}$ est infini : en effet, si $p_1, ..., p_n$ sont les seuls nombres premiers, alors $n = p_1...p_n + 1$ est encore premier, mais distinct de tous les autres... Puisque $\mathcal{P}$ est une partie de $\mathbb{N}$, elle est infinie dénombrable et l'on peut, d'après la preuve précédente, former la suite croissante des nombres premiers $(p_n)_{n\in\mathbb{N}}$. On peut par exemple montrer que $\sum \frac{1}{p_n}$ diverge, et que plus généralement $\sum \frac{1}{p_n^\alpha}$ a le même critère de convergence

[1] Ceci est utilisé dans le cadre du programme comme lemme des théorèmes de limite monotone en probabilités discrètes. Il sert aussi à démontrer un corollaire du théorème de sommation par paquets, que l'on utilise notamment dans une preuve par récurrence de l'identité d'Euler.

que celui de Riemann.

Exercice 19

Soit (u_n) une suite réelle et A une partie de $\mathbb{N}$. À quelle condition la quantité $\sum_{n\in A} u_n$ est-elle définie ?

Les cardinaux

Le cardinal est la notion intuitive de nombre d'éléments, et notamment dans le cas des ensembles finis où la notion est plus élémentaire. Cependant, dans le cas des ensembles finis, elle mène à de nombreux paradoxes et notamment à celui donnant que des ensembles peuvent avoir exactement le même nombre d'éléments (et donc, le même cardinal) que leurs parties strictes : par exemple, l'ensemble des entiers naturels et l'ensemble des entiers naturels pairs, propriété qui vient d'après ce que l'on a dit précédemment caractériser justement les ensembles infinis.

On dit que deux ensembles ont le même cardinal s'ils sont en bijection. Ceci définit une relation d'équivalence sur la classe de tous les ensembles (qui n'est pas un ensemble !) dont les classes ont pour représentants des ensembles typiques, qui sont habituellement : $\emptyset$, les $[\![1,n]\!]$, $n \in \mathbb{N}$, $\mathbb{N}$, $\mathcal{P}(\mathbb{N}) \simeq \mathbb{R}$, puis $\mathcal{P}(\mathcal{P}(\mathbb{N}))$, etc. Et encore, tous les cardinaux ne sont pas représentés par cette suite, strictement croissante pour l'ordre cardinal.

C'est Gottlob Frege et Georg Cantor qui définissent ces notions, posant les fondements de la théorie des ensembles à partir des années 1880, que ce dernier décrit comme l'étude de l'infini. Celle-ci est profondément liée à la logique théorique, deux branches mathématiques tout à fait co-dépendantes.

3.3.3.2 Réunion dénombrable de dénombrables

Lorsqu'on sait que pour tout $k \in \mathbb{N}$, $\mathbb{N}^k$ est dénombrable (voir ci-après), il est facile d'établir qu'une réunion finie d'ensembles finis est dénombrable. En effet, si les $A_1, ..., A_k$ sont des ensembles, leur réunion s'injecte trivialement dans $\prod_{i=1} kA_i$. La question se pose différemment dans le cas d'une réunion au plus dénombrable. Dans la suite, on utilisera le fait dû à l'axiome du choix qu'un ensemble *non vide* s'injecte dans un autre si et seulement s'il existe une surjection en sens inverse.

Exercice 20

(Grille de Cantor) Montrer que $(p,q) \longmapsto \frac{(p+q)(p+q+1)}{2} + q$ est une bijection de $\mathbb{N}^2$ sur $\mathbb{N}$, sans nécessairement expliciter la bijection réciproque.

Théorème. ***(Dénombrabilité des réunions dénombrables de dénombrables)***

Si I est au plus dénombrable et $(A_i)_{i\in I}$ une famille d'ensembles au plus dénombrables, alors leur réunion est encore au plus dénombrable. Le théorème est encore vrai en remplaçant par *dénombrable* à chaque occurrence de *au plus dénombrable*.

▷ On a vu le cas I fini ; ne considérons plus que I infini dénombrable. On peut donc prendre $(A_n)_{n\in\mathbb{N}}$ une famille d'ensembles tous au plus dénombrables. Ainsi, pour tout $n \in \mathbb{N}$, il existe une surjection $f_n : N \twoheadrightarrow A_n$. Considérons l'application $f : \mathbb{N}\times\mathbb{N} \longrightarrow \bigcup_{i\in\mathbb{N}} A_i$ qui à un couple (n,k) fait correspondre $f_n(k)$: l'existence d'une telle fonction est garantie par l'**axiome du choix**. Cette application est une surjection par définition de la réunion puis surjectivité des f_n : pour tout $x \in \bigcup_{i\in\mathbb{N}} A_i$, il existe k tel que $x \in A_k$ puis $n \in \mathbb{N}$ tel que $x = f_n(k)$. Or $\mathbb{N}^2$ est au plus dénombrable *(voir ci-après)*, donc $\bigcup_{i\in\mathbb{N}} A_i$ (son cardinal est plus petit que $\mathbb{N}$). Enfin, dans le cas de la dénombrabilité forte, il suffit de remarquer que A_0 s'injecte dans $\cup_{i\in\mathbb{N}} A_i$ pour avoir l'*infinie* dénombrabilité. ■

Remarque. La dernière phrase est inutile si l'on convient de la dénombrabilité faible, comme expliqué ci-haut, ce qui lui donne tout son intérêt. Mais en pratique, il est trivial de vérifier qu'une opération sur ensembles infinis est infinie.

Exercice 21

1. Un nombre réel est *algébrique*, s'il est racine d'un polynôme non nul à coefficients rationnels. Partitionner intelligemment $\mathbb{Q}[X]$ pour trouver le cardinal de l'ensemble des nombres algébriques.
2. *(Théorème de Liouville)* En déduire qu'il existe au moins un réel transcendant, c'est-à-dire non algébrique.

Il ne faut pas confondre ce dernier résultat théorématique avec celui sur les produits cartésiens, qui présente une dissymétrie. En effet, un produit cartésien fini d'ensembles au plus dénombrables est au plus dénombrable, mais c'est faux pour un produit cartésien infini, même dénombrable : on a vu que $\{0,1\}^{\mathbb{N}}$ était indénombrable dans le premier exercice.

Justifions cette première affirmation. Soit k un entier naturel (non nul, le produit cartésien vide étant vide), et $p_1,...,p_k$ k nombres premiers. Alors l'application de $\mathbb{N}^k$ dans $\mathbb{N}$ qui à $(n_1,...,n_k)$ fait correspondre $p_1^{n_1} \times ... \times p_k^{n_k}$ est injective d'après le théorème fondamental de d'Alembert, ce qui donne la dénombrabilité de $\mathbb{N}^k$, et à une bijection près celle d'un produit cartésien fini d'au plus dénombrables (une manière beaucoup plus rapide que celle de la grille de Cantor).

Exercice 22

Montrer que $\mathbb{R}^n$, en tant que $\mathbb{Q}$-espace vectoriel, est de dimension infinie. *Une base de* $\mathbb{R}$ *vu comme un* $\mathbb{Q}$*-espace vectoriel est appelée base de Hamel.*

Un ensemble a la puissance du continu, par définition, s'il est en bijection avec $\mathbb{R}$, qui est en bijection avec $\mathcal{P}(\mathbb{N})$. La fameuse *hypothèse du continu*, dont il a été démontré (Kurt Gödel en 1938 puis Paul Cohen en 1963 avec sa célèbre méthode du *forcing*) qu'elle était indécidable, à savoir démontrable et de négation démontrable, dans le cadre usuel de la théorie des ensembles, stipule qu'il n'existe pas de cardinal strictement compris entre $\mathbb{N}$ et $\mathcal{P}(\mathbb{N})$, autrement dit, que toute partie de $\mathbb{R}$ est au plus dénombrable ou a la puissance du continu. Cette affirmation rompt la continuité du théorème sur les parties de $\mathbb{N}$ établi à la partie précédente à laquelle on pouvait s'attendre.

Propriété. ***(Réunions dénombrables de puissances du continu)*** **HP**

Si $(A_n)_{n\in\mathbb{N}}$ est une famille d'ensembles ayant la puissance du continu, alors leur réunion est encore dénombrable.

▷ À faire en exercice, de façon tout à fait similaire à la démonstration précédente. ■

Si l'on s'aperçoit que $\mathbb{R}^2$ et $\mathbb{R}$ sont équipotents[1] (*i.e.* en bijection), on peut démontrer qu'une réunion de puissances de continu sur un ensemble ayant la puissance de continu a la puissance du continu. Plus généralement, si l'on sait montrer que tout ensemble infini est équipotent à son carré, ce qui n'est pas évident, mais vrai, alors on peut démontrer pareillement que si $(A_i)_{i\in E}$ est une famille dont tous les éléments sont de cardinal E, leur réunion a pour cardinal E au plus.

[1] Montrons ce résultat à l'aide du théorème de Cantor-Bernstein, théorème dont la preuve, quoique difficile, permet d'établir que si deux ensembles s'injectent l'un dans l'autre réciproquement, ils sont en bijection. $\mathbb{R}$ s'injecte dans $\mathbb{R}^2$ par l'application canonique $x \longmapsto (x,0)$. Réciproquement, exhibons une injection de $\mathbb{R}^2$ dans $\mathbb{R}$. Pour ça, il suffit d'exhiber une bijection de $[0,1[^2$ dans $[0,1[$, puisque comme c'est un intervalle non trivial, $[0,1[$ est en bijection avec $\mathbb{R}^2$. Posons l'application qui à deux éléments de $[0,1[$ associe l'entrelacement des décimales de leurs développements illimités propres, définie par :

$$\begin{array}{rccc} \varphi : & [0,1[^2 & \longrightarrow & [0,1[\\ & (0{,}a_1a_2a_3\ldots; 0{,}b_1b_2b_3\ldots) & \longmapsto & 0{,}a_1b_1a_2b_2a_3b_3\ldots. \end{array}$$

Cette fonction est bien définie d'après l'existence et l'unicité du développement illimité propre d'un réel, c'est-à-dire qu'on ne peut choisir qu'une unique suite de décimales non stationnaire à 9 qui représente un réel donné. Montrons son injectivité. Soient $(0{,}a_1a_2a_3\ldots; 0{,}b_1b_2b_3\ldots)$, $(0{,}a'_1a'_2a'_3\ldots; 0{,}b'_1b'_2b'_3\ldots) \in [0,1[$, les représentations ici étant des développements illimités propres. Supposons que $\varphi(0{,}a_1a_2a_3\ldots; 0{,}b_1b_2b_3\ldots) = \varphi(0{,}a'_1a'_2a'_3\ldots; 0{,}b'_1b'_2b'_3\ldots)$, c'est-à-dire $0{,}a_1b_1a_2b_2a_3b_3\ldots = 0{,}a'_1b'_1a'_2b'_2a'_3b'_3\ldots$, alors ces écritures ne sont pas des développements impropres. En effet, si $a_1b_1a_2b_2a_3b_3$ est stationnaire à 9, alors à partir d'un certain rang N, $(a_n, b_n) = (9{,}9)$ pour tous $n \geqslant N$. En particulier, (a_n) est stationnaire à 9, ce qui est exclu. De même pour l'écriture de droite. Or il y a unicité du développement décimal propre, ce qui impose : $a_1 = a'_1$ sur la première décimale, puis $b_1 = b'_1$, puis $a_2 = a'_2$, etc., de sorte que $0{,}a_1a_2a_3\ldots = 0{,}a'_1a'_2a'_3\ldots$ d'une part et $0{,}b_1b_2b_3\ldots = 0{,}b'_1b'_2b'_3\ldots$ d'autre part, ce qui montre l'injectivité de φ.

Exercice 23

En informatique théorique, on considère des *alphabets* finis, c'est-à-dire ensembles finis, prenons-en un $\mathscr{A}$, dont les éléments sont appelés *lettres*. Un *mot fini* sur $\mathcal{A}$ est un n-uplet d'éléments de $\mathscr{A}$. L'ensemble des *langages* sur $\mathscr{A}$ est défini comme $\mathcal{P}(\mathscr{A}^*)$. On dit qu'un langage $\mathcal{L} \subseteq \mathscr{A}^*$ est *reconnu par un algorithme* s'il existe un programme informatique ayant pour entrées les $u \in \mathscr{A}^*$ et qui calcule la fonction caractéristique de $\mathcal{L}$. Montrer qu'il existe des langages indécidables, c'est-à-dire qui ne sont pas reconnus par des algorithmes.

3.4 Quotients d'ensemble : ensembles sans structure donnée (préambule aux quotients de groupe)

Ces notions sont complètement hors programme. Toutefois, elles permettent de s'éclaircir grandement les idées sur certaines notions en théorie des groupes mais également en algèbre linéaire comme nous l'allons voir. De plus, elle permet de clore la théorie élémentaire des opérations ensemblistes : la somme (réunion), la différence (différence ensembliste, différence symétrique), le produit (cartésien ou intersection), sont complétés par le quotient, qu'il faut introduire au moyen d'un objet à première vue retors : la relation d'équivalence.

Ne nous leurrons pas pourtant : lorsqu'il s'agit d'ensembles simplement, il n'y a pas de notion de quotient de deux ensembles, mais seulement d'*ensemble quotient par une relation d'équivalence*. On ne pourra définir la notion de quotient d'ensembles par l'une de ses parties, et encore il les faudra structurées (sous-groupe pour un groupe, sous-espace vectoriel, etc., ce qui correspondra à la notion de diviseur) que dans de telles structures.
On rappelle un résultat de mathématiques supérieures dans sa totalité.

Théorème. ***(Théorème fondamental des relations d'équivalence)***

Soit E un ensemble. L'application $\mathcal{R} \longmapsto \{\overline{x}_{\mathcal{R}} \mid x \in E\}$ est une bijection de l'ensemble des relations d'équivalence sur E dans l'ensemble des partitions de E.

▷ Il s'agit démontrer d'abord que cette application est bien définie, c'est-à-dire que les classes d'une relation d'équivalence partitionnent l'ensemble sur lequel elle est définie, ce qui est un résultat de sup (le refaire !). On vérifie facilement qu'à toute partition de E, on peut faire correspondre une relation d'équivalence définie par l'appartenance de deux éléments à la même partie de la partition. Ces deux applications sont alors bijections réciproques. ■

Exercice 24

Le n-ième nombre de Bell, noté B_n, est le nombre de relations d'équivalence sur un ensemble à n éléments.

1. *(Relation d'Aitken)* Montrer que $B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k$.
2. *(Formule de Dobinski)* En déduire que $B_n = \frac{1}{e} \sum_{k=0}^{+\infty} \frac{k^n}{k!}$ le n-ième moment de la Poisson de moyenne $\lambda = 1$.
3. *(Congruence de Touchard)* Montrer que si p est premier, alors $B_{n+p} \equiv B_n + B_{n+1}$ $[p]$.

Exemple. ***(Relation « avoir le même signe »)***

Si l'on partitionne $\mathbb{R}^*$ en ses deux composantes connexes, on obtient la relation « avoir le même signe ». Remarquer que c'est la relation engendrée par la partition donnée par les fibres de la fonction signe σ, une fibre par une fonction f n'étant autre que l'image réciproque d'un point.

Propriété

Soit $\mathcal{R}$ une relation d'équivalence sur un ensemble E. Alors pour tous x,$y \in E$, $x\mathcal{R}y \iff \overline{x} = \overline{y}$.

▷ En exercice. ■

Définition. ***(Ensemble quotient par une relation d'équivalence)***

Soit E un ensemble et $\mathcal{R}$ une relation d'équivalence sur E. Alors on note $E/\mathcal{R}$ l'*ensemble quotient par* $\mathcal{R}$ défini par $E/\mathcal{R} = \{\overline{x}_{\mathcal{R}} \mid x \in E\}$.

⟶ **Convention.** On aurait aussi pu définir, de manière équivalente, l'ensemble quotient par un système de représentants de $\mathcal{R}$, c'est-à-dire un élément de la classe de x et un seul dans l'ensemble quotient pour tout x de E. Cependant, nous fixons la construction de la preuve précédente.

Fixons maintenant E un ensemble, $\mathcal{R}$ une relation d'équivalence sur E. On se permet de la noter $\sim$. Pour rappel, on peut noter indifféremment $\overline{x}_{\mathcal{R}} = \overline{x}_{\sim} = \overline{x} = cl(x)$ quand il n'y a pas d'ambiguïté la classe d'équivalence de l'élément $x \in E$.

Définition-propriété. ***(Projection canonique)***

L'application $\pi : E \longrightarrow E/\mathcal{R}$ qui à x fait correspondre la classe de x par $\mathcal{R}$, notée $\overline{x}_{\mathcal{R}}$, est une surjection appelée *projection canonique.*

▷ La surjectivité vient de la réflexivité de $\mathcal{R}$. ■

Propriété

Pour tout $a \in E/\mathcal{R}$, $\pi^{-1}(a) = cl(x)$ où x est un antécédent de a par π (il en existe au moins un d'après la proposition précédente).

▷ Évident. ■

A priori, bien sûr, l'ensemble et l'ensemble des classes d'équivalence par $\mathcal{R}$ n'ont pas le même cardinal.

Exercice 25

Montrer que la projection canonique est une bijection si et seulement si $\mathcal{R}$ égale la relation d'égalité sur E, notée $=_E$.

▷ **Éléments de réponse.**

Elle est bijective si et seulement si elle est injective.

Heuristique

L'introduction de la notion d'ensemble quotient débrouille la perplexité initiale devant l'idée d'appeler quotient un ensemble formée des classes d'une relation : la surjectivité permet tout d'abord d'avoir que, dans tous les cas, même en milieu infini, $\text{card}(E) \geqslant \text{card}(E/\mathcal{R})$, ce qui permet d'assimiler $\mathcal{R}$ à, par exemple, un nombre supérieur à 1, dans le cas E non vide, par l'artifice frauduleux ci-dessous :

$$\text{card}(E) \geqslant \text{card}(E/\mathcal{R}) = \frac{\text{card}(E)}{\text{card}(\mathcal{R})} \text{ donc } \text{card}(\mathcal{R})\text{card}(E) \geqslant \text{card}(E) \text{ donc } \text{card}(\mathcal{R}) \geqslant 1,$$

nombre qui diviserait E en autant parties disjointes. L'important dans la suite sera double : d'une part que la structure quotientée soit préservée (ce qui sera à peu près toujours le cas, et constituera à chaque section la première partie sur la compatibilité de la structure avec la relation d'équivalence : sous-groupe distingué, linéarité de la projection, continuité en topologie...), d'autre part, l'établissement de théorèmes de factorisation et d'isomorphisme à propos des morphismes partant de structures quotientées dont on verra que, par cet effet, ils sont, sous certaines hypothèses, « simplifiables » en applications quotients.

Exercice 26

Une relation d'équivalence sur $\mathbb{R}$ permet de définir l'argument principal d'un complexe : quel est le cardinal de l'ensemble quotient par elle ?

Les théorèmes sur les ensembles quotients sont très bien résumés par les diagrammes. Nous introduisons d'ores et déjà celui qui sert de base à tous les autres avec les propriétés précédentes. On rappelle que parmi les applications qui sont des flèches $\longrightarrow$, les injections sont notées $\hookrightarrow$,

les surjections sont notées $\twoheadrightarrow$, et les bijections sont notées $\simeq$ ou comme la conjonction d'une injection et d'une surjection.

$$\begin{array}{c} E \\ {\scriptstyle\pi}\big\downarrow \\ E/\mathcal{R} \end{array}$$

Exercice 27

Pour se débarbouiller l'esprit sur le raisonnement par analyse-synthèse, montrer que :

1. Toute fonction de $\mathbb{R}$ dans $\mathbb{R}$ se décompose de façon unique comme somme d'une fonction paire et d'une fonction impaire.
2. Toute matrice à coefficients dans un corps de caractéristique différente de 2 ($2 \neq 0$) se décompose de façon unique comme somme d'une matrice symétrique et d'une matrice antisymétrique.
3. Pour toute base $(e_1,...,e_n)$ d'un espace vectoriel de dimension finie, il existe une unique base de son dual, base duale, notée $(e_i^*)_{i\in[\![1,n]\!]}$ telle que pour tous $i,j \in [\![1,n]\!]$, $e_i^*(e_j) = \delta_i^j$.

Théorème. ***(Théorème de factorisation pour les applications)***

Soit F un ensemble quelconque et f une application de E dans F. Alors f est compatible avec $\mathcal{R}$ (*i.e.* $\forall x,y \in E \quad x \sim y \implies f(x) = f(y)$) si et seulement s'il existe une unique application $\tilde{f}$ telle que $f = \tilde{f} \circ \pi$ (se qui se réécrit $f(x) = \tilde{f}(\overline{x})$ pour tout $x \in E$). Dans ce cas de compatibilité, on dit qu'on *passe au quotient* dans l'application f.

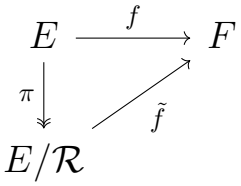
▷ La preuve est très facile si l'on se focalise sur le choix des bons arguments.

▶ Dans le sens direct, on suppose que f est compatible avec la relation $\mathcal{R}$. On montre l'unicité et l'existence de $\tilde{f}$ par analyse-synthèse. Supposons que pour tout $x \in E$, puisque par définition $\pi(x) = \overline{x}$, $\tilde{f}(\overline{x}) = f(x)$: ceci définit de manière explicite et donc unique l'application $\tilde{f}$, ce qui termine l'analyse. Le point crucial est que **cette écriture a un sens**, vu que $f(x)$ ne dépend pas du représentant choisi de $\overline{x}$ *ce qui est exactement ce que signifie la condition de compatibilité*. Pour la synthèse, c'est encore plus immédiat : pour tout $x \in E$, $\overline{x} = \pi(x)$ donc $\tilde{f}(\overline{x}) = \tilde{f} \circ \pi(x)$. Mais l'hypothèse de synthèse définit ce que l'analyse conclut, soit $\tilde{f}(\overline{x}) = f(x)$, d'où $f(x) = \tilde{f} \circ \pi(x)$ ce qui signifie par définition de l'égalité des fonctions que $f = \tilde{f} \circ \pi$.

▶ Réciproquement, si l'on suppose que f se factorise en $\tilde{f} \circ \pi$ (de façon unique, mais on ne s'en sert pas), soient $x \sim y$ deux éléments de E. Alors $\pi(x) = \pi(y)$ par définition de π. Ainsi $\tilde{f} \circ \pi(x) = \tilde{f} \circ \pi(y)$ puisque $\tilde{f}$ est une application, et par hypothèse ces deux quantités égalent $f(x) = f(y)$, et donc f est compatible avec $\mathcal{R}$.

Ainsi f est compatible si et seulement elle se factorise, ce qu'il fallait démontrer. ■

La situation se présente comme suit :



Le théorème de factorisation établit donc la commutation de ce diagramme.

Exemple fondamental. ***(Anneaux quotients)***

Soit $(\mathbb{A}, +, \times)$ un anneau et

$$\begin{aligned} f : \ \mathbb{Z} &\longrightarrow \mathbb{A} \\ k &\longmapsto k.1_{\mathbb{A}} \end{aligned}$$

(l'unique) morphisme de l'anneau $\mathbb{Z}$ dans $\mathbb{A}$. Si $\mathbb{A}$ est fini, $\mathrm{Ker}(f)$ étant un idéal de $\mathbb{Z}$, il est de la forme $p\mathbb{Z}$ où p est pris minimal, $p \neq 0$ puisque f ne peut être injective par cardinalité. Dans ce cas, f est compatible avec la relation de congruence modulo p sur les entiers, et l'on peut définir

$$\begin{aligned} \tilde{f} : \ \mathbb{Z}/p\mathbb{Z} &\longrightarrow \mathbb{A} \\ \overline{k} &\longmapsto k.1_A. \end{aligned}$$

Il faut remarquer que c'est bien la compatibilité et elle seule (équivalence dans le théorème de factorisation) qui permet de définir ce nouveau morphisme, même il faut vérifier indépendamment que l'application est un morphisme : ce n'est pas dur mais ce sera systématique dans la partie suivante. Si $\mathbb{A}$ est intègre, et l'on peut montrer qu'un anneau fini est intègre si et seulement si c'est un corps, alors on vérifie que p est premier. Dans ce cas, $\tilde{f}$ est injectif car c'est un morphisme de corps, ce qui nous donne notamment que tout corps fini de cardinal premier p est isomorphe à $\mathbb{Z}/p\mathbb{Z}$ et que tout corps fini a pour cardinal p^d où d est un certain entier : la dimension sur $\mathrm{Im}(\tilde{f})$ de $\mathbb{K}$ vu comme $\mathbb{Z}/p\mathbb{Z}$-espace vectoriel.

Exercice 28

Fournir un exemple où une application ne passe pas au quotient par une relation d'équivalence donnée.

Méthode. ***(Recette pour passer au quotient dans les applications)***

J'ai une application φ d'un ensemble quotient Q dans F un ensemble quelconque.

1. J'identifie la relation d'équivalence qui quotiente : $Q = E/\mathcal{R}$ et E l'ensemble initial.
2. Je vérifie que $\mathcal{R}$ est une relation d'équivalence pour justifier mon propos.
3. Je pose une application f de E dans F définie sans aucun problème et qui devra, une fois passée au quotient, retomber sur φ.
4. Je montre que pour tous éléments x,y de E, si $x\mathcal{R}y$, alors $f(x)$ et $f(y)$ sont égaux.
5. Je peux maintenant définir une application $\varphi = \tilde{f} : Q \longrightarrow F$ sans trouble, telle que pour tout $\overline{x} \in Q$, $\tilde{f}(\overline{x}) = f(x)$, et j'insiste bien sur ce que cette construction n'est possible que grâce à la compatibilité vérifiée précédemment.

Si je veux une propriété d'injectivité, de surjectivité, voire de bijectivité pour mon application, je me réfère au résultat de l'exercice suivant.

Un corollaire d'intérêt principalement formel.

Théorème. ***(Théorème de factorisation carré pour les applications)***

Soit F un ensemble muni d'une relation d'équivalence $\mathcal{S}$ que l'on se permet de noter $\equiv$, ses classes $\hat{\cdot}$ et f une application de E dans F. Alors f est compatible avec $\mathcal{R}$ modulo $\mathcal{S}$ (*i.e.* $\forall x,y \in E \quad x \sim y \implies f(x) \equiv f(y)$) si et seulement s'il existe une unique application $\tilde{f}$ telle que $\chi \circ f = \tilde{f} \circ \pi$ (se qui se réécrit $f\hat{(}x) = \tilde{f}(\overline{x})$ pour tout $x \in E$). Dans le cas de compatibilité, on dit encore qu'on passe au quotient dans f.

▷ On applique le théorème de factorisation à l'application $\chi \circ f$, l'ensemble F étant maintenant $F/\mathcal{S}$. Il suffit de vérifier alors que la compatibilité de $\chi \circ f$ avec $\mathcal{R}$ est équivalente à la compatibilité de f avec $\mathcal{R}$ modulo $\mathcal{S}$, ce qui est immédiat par définition de cette dernière notion. ■

Le diagramme correspondant est le suivant :

$$\begin{array}{ccc} E & \xrightarrow{\quad f \quad} & F \\ {\scriptstyle\pi}\big\downarrow & & \big\downarrow{\scriptstyle\chi} \\ E/\mathcal{R} & \xrightarrow[\quad \tilde{f} \quad]{} & F/\mathcal{S} \end{array}$$

Exercice 29

Dans chacun des deux théorèmes précédents, montrer que :

1. $\tilde{f}$ est injective si et seulement si $\forall x,y \in E \quad x \sim y \iff f(x) = f(y)$ (respectivement $\forall x,y \in E \quad x \sim y \iff f(x) \equiv f(y)$) ;
2. $\tilde{f}$ est surjective si et seulement si f est surjective ;
3. $\tilde{f}$ est bijective si et seulement si f est surjective et $\forall x,y \in E \quad x \sim y \iff f(x) = f(y)$ (respectivement f est surjective et $\forall x,y \in E \quad x \sim y \iff f(x) \equiv f(y)$).

Méthode. *(Recette pour passer au quotient dans les applications entre deux espaces quotients)*

C'est exactement la même que précédemment : **la structure quotient de l'espace d'arrivée n'intervient pas**. Pour une illustration, voir dans la section suivante sur les magmas.

Exercice 30

(Difficile) Soit n un entier naturel. On note $\mathbb{P}^n$ le quotient de $\mathbb{R}^{n+1} \setminus \{0\}$ où ici $0 \in \mathbb{R}^n$ par la relation : pour $x,y \in \mathbb{R}^{n+1}$, $x \sim y \iff \exists \lambda \in \mathbb{R} \quad x = \lambda y$. Montrer que pour naturels n,m, l'application : $\begin{array}{ccc} \mathbb{P}^n \times \mathbb{P}^m & \longrightarrow & \mathbb{P}^{n+m-1} \\ (x_0,y_0),(x_1,y_1) & \longmapsto & (x_0y_0,x_1y_0,x_0y_1,x_1y_1) \end{array}$ (*plongement de Segre*) est bien définie. (En GÉOMÉTRIE ALGÉBRIQUE, il permet de justifier que le produit de deux variétés projectives est encore une variété projective.)

On examine enfin un cas particulier important, celle de la relation $x \sim y \iff f(x) = f(y)$ qui est définie sur tout ensemble étant donné une application f partant de cet ensemble, dont les classes sont parfois appelées *fibres*. Le théorème de factorisation appliqué à f donne un résultat intéressant.

Théorème. ***(Théorème d'isomorphisme ensembliste, théorème de bijection quotient)***

Soit f une application de E dans F deux ensembles quelconques. On considère la relation d'équivalence $\mathcal{R}$ sur E définie par $x \sim y \iff f(x) = f(y)$. Dans ce cas, f est compatible avec $\mathcal{R}$ et l'application quotient de f par cette relation réalise une bijection de $E/\mathcal{R}$ sur $\mathrm{Im}(f)$.

▷ Encore une fois, le théorème ne consiste qu'en des astuces de langage. Procédons par étapes claires et distinctes pour rasséréner les esprits malades. D'abord, on vérifie aisément que la relation des fibres $\mathcal{R}$ est bien une relation d'équivalence : la réflexivité est tautologique, la transitivité vient de ce que deux choses égales à un même sont égales, et la symétrie de celle de la relation d'égalité même. Dans ce cas, $E/\mathcal{R} = \{f^{-1}(y) \mid y \in \mathrm{Im}(f)\}$.

Pour pouvoir avoir seulement l'audace de rêver du théorème de factorisation, il nous faut vérifier la compatibilité de f avec $\mathcal{R}$. Nous laissons aux cerveaux fatigués le soin de découvrir par eux-mêmes,

pourquoi c'est évident. La relation d'équivalence des fibres est également la relation d'équivalence associée à l'application f, définissable sur tout ensemble dont elle part par « avoir la même image par f ». Cette compatibilité exprime que cette relation est moins fine que la relation définie dans le théorème.

La remarque du paragraphe surprécédent donne en particulier, ayant introduit l'application quotient $\tilde{f}$ par $\mathcal{R}$, que $\boxed{\tilde{f}[f^{-1}(y)] = y}$. En effet

Soient deux éléments de $E/\mathcal{R}$; d'après l'expression de l'ensemble quotient ci-dessus, il existe y,$y' \in \mathrm{Im}(f)$ tels que ces éléments s'écrivent $f^{-1}(y)$, $f^{-1}(y')$. Si $\tilde{f}(f^{-1}(y)) = \tilde{f}(f^{-1}(y'))$, alors d'après ce qui précède $y = y'$ (c'est tout simplement une réécriture des deux termes), d'où $f^{-1}(y) = f^{-1}(y')$, et donc $\tilde{f}$ est injective.

De surcroît, $\mathrm{Im}(f) = \mathrm{Im}(\tilde{f})$. En effet, $\mathrm{Im}(\tilde{f}) = \{\tilde{f}(f^{-1}(y)) \mid f^{-1}(y) \in E/\mathcal{R}\} = \{\tilde{f}(f^{-1}(y)) \mid y \in \mathrm{Im}(f)\}$ d'après l'égalité encadrée et encore d'après l'égalité encadrée ceci égale $\{y \mid y \in \mathrm{Im}(f)\} = \mathrm{Im}(f)$.

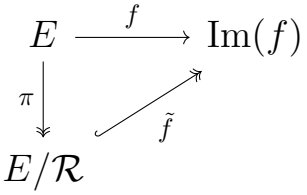
Ainsi $\tilde{f}$ est une bijection de $E/\mathcal{R}$ sur $\mathrm{Im}(\tilde{f}) = \mathrm{Im}(f)$, ce qui termine la démonstration. ■

Ceci constitue un résultat théorique : l'utilité de la relation d'équivalence associée à f est factice ; il permet seulement d'établir que l'image d'une application est isomorphe, au sens ensembliste (c'est-à-dire en bijection) avec l'espace des fibres par f, ce qui se récrit : $\mathrm{Im}(f) \simeq \{f^{-1}(y) \mid y \in \mathrm{Im}(f)\}$. Remarquons qu'on aurait pu l'établir plus élémentairement.

Exercice 31

1. Démontrer que, pour dénombrer son troupeau, un berger peut se « contenter » de compter les pattes puis diviser par quatre.
2. *(Lemme des bergers)* Soit f une surjection de A dans B deux ensembles finis, telle que pour tout y dans B, y ait exactement k antécédents par f. Montrer qu'alors $\mathrm{card}(E) = k.\mathrm{card}(F)$ et retrouver ce qui précède[1].
3. Que dire si f n'est plus surjective ?

A picture is worth a thousand words.



À ce stade, nous recommandons très chaudement d'étudier la notion algébrique de quotient dans l'ordre suivant :

- sous-groupes distingués et groupes quotients,
- anneaux quotients,
- espaces vectoriels quotients et co-dimension,
- topologie quotient.

L'étude des structures quotients est un tout : il est très envisageable de ne s'intéresser à ce concept qu'à partir de la troisième année universitaire, mais dans ce cas, mieux vaut enchaîner les cinq chapitres cités ci-haut (y compris le préambule à propos des ensembles quotients).

On dispose d'une vue d'ensemble des structures quotients (ainsi que du parallélisme de construction) grâce au **tableau récapitulatif** de la figure 3.1.

Catégorie	Quotient par...	Théorèmes de factorisation	Théorèmes d'isomorphisme
Ensembles	Relation d'équivalence $\mathcal{R}$	f est compatible avec $\mathcal{R}$ ssi $\exists! \tilde{f} \quad f = \tilde{f} \circ \pi$	Pour $\mathcal{R} : x \sim y \Leftrightarrow f(x) = f(y)$, $E/\mathcal{R} \simeq \mathrm{Im}\, f$ (par $\tilde{f}$)
Magmas	Relation d'équivalence $\mathcal{R}$ avec laquelle la loi de magma est compatible	φ morph. est compatible avec $\mathcal{R}$ ssi $\exists! \tilde{\varphi}$ morph. $\quad \varphi = \tilde{\varphi} \circ \pi$	Pour $\mathcal{R} : x \sim y \Leftrightarrow \varphi(x) = \varphi(y)$, $E/\mathcal{R} \simeq \mathrm{Im}\, \varphi$ (par $\tilde{\varphi}$)
Monoïdes, groupes (naïvement)	Idem	Idem	Idem
Groupes	Sous-groupe distingué H	$H \subseteq \mathrm{Ker}(f)$ ssi $\exists! \tilde{f}$ morph. $\quad f = \tilde{f} \circ \pi$	$G/\mathrm{Ker}(f) \simeq \mathrm{Im}\, f$
Anneaux	Idéal I	$I \subseteq \mathrm{Ker}(f)$ ssi $\exists! \tilde{f}$ morph. $\quad f = \tilde{f} \circ \pi$	$A/\mathrm{Ker}(f) \simeq \mathrm{Im}\, f$
Espaces vectoriels	Sous-espace vectoriel V	$V \subseteq \mathrm{Ker}(f)$ ssi $\exists! \tilde{f}$ lin. $\quad f = \tilde{f} \circ \pi$	$E/\mathrm{Ker}(f) \simeq \mathrm{Im}\, f$
Espaces topologiques	Relation d'équivalence $\mathcal{R}$, quotient muni de la topologie quotient	f C^0 est compatible avec $\mathcal{R}$ ssi $\exists! \tilde{f}$ C^0 $\quad f = \tilde{f} \circ \pi$	Pour $\mathcal{R} : x \sim y \Leftrightarrow f(x) = f(y)$, f continue et ouverte (*quotient map*), $E/\mathcal{R} \simeq \mathrm{Im}\, f$ (par $\tilde{f}$)

TABLE 3.1 : *Tableau récapitulatif des phénomènes de factorisation et d'isomorphisme dans les structures quotients selon les catégories.* —
Il est très beau.

Signalons avant de partir le fait suivant, qui se transpose aux autres catégories :

Fait. ***(Universalité des projections)***

Soient X,Y deux ensembles et $f : X \longrightarrow Y$ surjective. Alors il existe une relation d'équivalence R sur X, compatible avec f, telle que $Y \simeq X/R$.

▷ Conséquence essentielle du théorème d'isomorphisme. En fait, il suffit de prendre $xRy \iff f(x) = f(y)$. ■

Chapitre 4

Théorie de la logique

Résumé

On expose la théorie de la logique formalisée telle que présentée classiquement en maîtrise de mathématiques pures.

4.1 Logique du premier ordre

Chapitre 5

Théorie des catégories

Résumé

Les catégories sont la version « plus-plus » des ensembles. S'il est assez faux de dire que la théorie des catégories trivialise des résultats mathématiques, il est très vrai de constater qu'elle uniformise un certain nombre de concepts d'apparence transverses, et c'est par là d'ailleurs que sourdent les premiers résultats substantiels. En conséquence, on s'attache à décrire dans un premier temps le vocabulaire, avec force exemple, pour que le lecteur se familiarise avec ; il reste difficile de n'être pas rebuté par l'aridité des premiers théorèmes (caractérisation des équivalences, lemme de Yoneda)... courage !

5.1 Exposé des premières définitions de la théorie des catégories

5.1.1 Introduction

5.1.2 Définition

La théorie des catégories prolonge celle des ensembles et doit se placer dans le cadre axiomatique (semblable) de la théorie des classes.

Définition. ***(Catégorie)***

La *catégorie* est l'élément de base de la théorie des catégories. On est en présence d'une catégorie $\mathcal{C}$, si :

- une collection d'*objets*, représentée par une classe, propre ou impropre, identifiée à $\mathcal{C}$;
- la collection des *flèches* ou *morphismes* entre ces objets ;
- l'association, à toute flèche f de A dans B, de son départ (ou *domaine*) A et de son arrivée B (ou *co-domaine*) ;

- une loi de composition interne entre les flèches, appelée *composition*, dès que le co-domaine de celle de droite correspond au domaine de celle de gauche ;
- l'associativité de la composition ;
- un élément neutre pour la composition, associé à tout objet A de la catégorie $\mathcal{C}$, noté id_A.

Remarques.

1. On note parfois $\mathrm{Ob}(\mathcal{C})$, $\mathrm{ob}(\mathcal{C})$ ou $\mathrm{Obj}(\mathcal{C})$ la classe des objets de $\mathcal{C}$.
2. En tout cas, pour tous $X,Y \in \mathrm{Ob}(\mathcal{C})$, on note $\mathrm{Hom}_{\mathcal{C}}(X,Y)$ la classe des morphismes entre X et Y de X à Y.
3. Avec ces notations, la composition est définie $\circ : \mathrm{Hom}(Y,Z) \times \mathrm{Hom}(X,Y) \longrightarrow \mathrm{Hom}(X,Z)$, et $\forall x\ id_X \in \mathrm{Hom}(X,X)$.
4. **(Important)** On écrit simplement $f : X \longrightarrow Y$ pour, $f \in \mathrm{Hom}(X,Y)$. Cette notation doit être employée savamment, car elle ne fait pas mention de la catégorie considérée (dans un espace de Banach par exemple, on s'y perd facilement).

Autres remarques.

1. La théorie de catégorie a pour motivation de démontrer les théorèmes observés de façon analogue dans diverses catégories, une fois pour toutes.
2. Une catégorie est totalement encodée par ses morphismes : la connaissance d'un objet conditionnée par la connaissance de son identité. C'est pourquoi les objets, en théorie des catégories, ont une importance moindre face aux flèches.
3. Un monoïde est exactement la donnée d'une catégorie dont la classe des objets est réduire à un point. En effet, si $\star$ est un point, alors cette catégorie est déterminée par $\mathrm{Hom}(X,X)$, qui, d'après les axiomes de catégorie, doit et doit seulement être un monoïde (non nécessairement commutatif).

L'exemple le plus rudimentaire où la catégorie est celle des ensembles, on la note Set ou Ens en anglais ou en français.

5.1.3 Catégorie petite, localement petite

Définition. ***(Catégorie petite, catégorie localement petite)***

Une catégorie est *petite* si la classe de tous ses morphismes est impropre (autrement dit, c'est un ensemble). Une catégorie est *localement petite* si pour tous objets A,B, la classe des morphismes de A dans B, notée $\mathrm{Hom}(A,B)$, est impropre.

Remarque. Une catégorie petite est localement petite (car toute partie d'un ensemble est un ensemble), mais la réciproque est fausse. De plus, la classe des objets d'une catégorie petite n'est pas forcément impropre, car on n'est pas forcé, pour définir une certaine catégorie de

« prendre toutes les flèches possibles ». C'est le cas cependant pour les catégories concrètes dont le foncteur d'oubli vers Set est plein *(voir ci-dessous)*.

5.1.4 Sous-catégorie, sous-catégorie pleine

Définition. ***(Sous-catégorie pleine)***

Une sous-catégorie $\mathcal{C}'$ d'une catégorie $\mathcal{C}$ est la donnée de certains objets de $\mathcal{C}$, mais pas forcément tous, et de certaines flèches de $\mathcal{C}$, mais pas forcément toutes. Une sous-catégorie est dite *pleine* si pour tous objets A,B de $\mathcal{C}'$, $\mathrm{Hom}_{\mathcal{C}'}(A,B) = \mathrm{Hom}_{\mathcal{C}}(A,B)$ (il revient au-même de dire que le foncteur d'oubli de $\mathcal{C}'$ dans $\mathcal{C}$ est plein).

5.1.5 Foncteurs

Afin de représenter les liens entre les différentes catégories, on introduit la notion de foncteur. L'injectivité intuitive d'un foncteur est traduite par la notion de fidélité, et la surjectivité intuitive d'un foncteur est traduite par la notion de plénitude.

Définition. ***(Foncteur)***

Un *foncteur* ou *foncteur covariant* d'une catégorie $\mathcal{C}$ dans une catégorie $\mathcal{D}$ est la donnée d'une fonction qui à tout objet X de $\mathcal{C}$, associe un objet $F(X)$ de $\mathcal{D}$ et d'une fonction qui à tout morphisme $f : X \longrightarrow Y$ de $\mathcal{C}$, associe un morphisme $F(f) : F(X) \longrightarrow F(Y)$ de $\mathcal{D}$, vérifiant les deux propriétés supplémentaires : $F(id_X) = id_{F(X)}$ pour tout objet X de $\mathcal{C}$, et pour tous objets X,Y,Z et morphismes $f : X \longrightarrow Y$ et $g : Y \longrightarrow Z$ de $\mathcal{C}$, $F(g \circ f) = F(g) \circ F(f)$.

Exemples

1. On peut considérer le foncteur d'oubli Top $\longrightarrow$ Ens. On peut alors également considérer le foncteur de Top dans Top_h qui à $x \mapsto x$ et à $f \mapsto [f]$.

Définition. ***(Composition entre foncteurs)***

On peut égaler définir la composition entre foncteurs pour $F : C \to D$ et $G : D \to E$, on définit GF de façon évidente. De plus pour toute catégorie C, il existe un foncteur identité Id_C qui ne touche à rien. Ainsi, un *isomorphisme de catégories* est un foncteur bi-inversible.

Notons que l'on ne définira pas l'équivalence de catégories avec l'existence d'un foncteur bi-inversible.

On a (en anticipant sur la suite) pour tout $(x,y) \in \mathcal{C}$, une application de $\mathrm{Hom}_{\mathcal{C}}(x,y) \longrightarrow \mathrm{Hom}_{\mathcal{D}}(F(x),F(y))$. En particulier, $\mathrm{Aut}_{\mathcal{C}}(x) \longrightarrow \mathrm{Aut}_{\mathcal{D}}(F(x))$ est un morphisme de groupes.

5.1.5.1 Foncteur fidèle

Définition. ***(Foncteur fidèle)***

Un foncteur d'une catégorie $\mathcal{C}$ dans une catégorie $\mathcal{D}$ est *fidèle* si pour tous morphismes $f,g : X \longrightarrow Y$, si $F(f) = F(g)$, alors $f = g$.

5.1.5.2 Foncteur plein

Définition. ***(Foncteur plein)***

Un foncteur d'une catégorie $\mathcal{C}$ dans une catégorie $\mathcal{D}$ est *plein* si tout morphisme $F(X) \longrightarrow F(Y)$ est égal à un $F(f)$.

5.1.5.3 Catégorie concrète

Définition. ***(Catégorie concrète)***

Une catégorie est *concrète* s'il existe un foncteur fidèle, dit *foncteur d'oubli* de cette catégorie vers la catégorie des ensembles ; on peut donc la voir comme une sous-catégorie de Set.

Toutes les catégories évoquées dans les prochains développement de notre composition sont des catégories concrètes. Par exemple, la catégorie des groupes est concrète, car à tout groupe, on peut faire correspondre un unique ensemble et à tout morphisme de groupes une unique application par $(G,\star) \longrightarrow G$. Ce n'est qu'en conclusion que nous étudierons le cas de la théorie abstraite des catégories.

5.1.6 Isomorphisme

Nous introduisons enfin la notion capitale des mathématiques.

Définition. ***(Isomorphisme)***

Un isomorphisme entre deux objets X, Y d'une catégorie $\mathcal{C}$ est un morphisme f de X dans Y tel qu'il existe un morphisme g de Y dans X tel que $g \circ f = f \circ g = id_X$.

Remarque. Il revient au même, dans une catégorie concrète, de se donner un morphisme bijectif donc l'inverse est encore un morphisme. Dans beaucoup de catégories, cette dernière condition est automatiquement vérifiée : groupes, espaces vectoriels, etc. Ce n'est pas général toutefois : par exemple, dans la catégorie des ensembles partiellement ordonnés, l'inverse d'une bijection croissante n'est pas nécessairement croissante.

5.1.6.1 Automorphisme

Définition. ***(Automorphisme)***

Soit $\mathcal{C}$ une catégorie, $x \in \mathcal{C}$. On note $\mathrm{Aut}_{\mathcal{C}}(x)$ l'ensemble des isomorphismes de $x \to x$.

Propriété. ***(Description des isomorphismes)***

Soit C une catégorie et X,Y deux objets de C. Si X et Y sont isomorphes (dans C), par disons $\varphi : X \simeq Y$, alors l'ensemble des isomorphismes de X dans Y est $\varphi \circ \alpha$ pour α parcourant $\mathrm{Aut}(X)$.

▷ Il est clair qu'un tel objet est un isomorphisme de X dans Y, d'inverse $\alpha^{-1} \circ \varphi^{-1}$. Réciproquement, soit $f : X \longrightarrow Y$ un isomorphisme. Alors $f^{-1} \circ \varphi$ est un isomorphisme (par composition) de X dans X, donc $\alpha = f^{-1} \circ \varphi \in \mathrm{Aut}(X)$. Ainsi $\varphi \circ \alpha^{-1} = f$. Puisque $\mathrm{Aut}(X) = \mathrm{Aut}(X)^{-1}$, le théorème est montré. ■

5.1.7 Produit de catégories

Définition. ***(Produit de catégories)***

Le produit $C \times D$ de deux catégories, est défini par $\mathrm{Ob}(C \times D) = \mathrm{Ob}(C) \times \mathrm{Ob}(D)$ et $\mathrm{Hom}((x,x'),(y,y')) = \mathrm{Hom}(x,y) \times \mathrm{Hom}(x',y')$.

5.2 Transformations naturelles et lemme de Yoneda

5.2.1 Définition

Définition. ***(Transformation naturelle)***

Soient F,G : $C \longrightarrow D$ deux foncteurs. Une transformation naturelle $\eta : F \longrightarrow G$ est la donnée pour tout $x \in C$ d'un morphisme de D, $\eta_x : F(x) \longrightarrow G(x)$, tel que $\forall f : x \longrightarrow y$ ($\in \mathrm{Hom}_C(x,y)$),

$$\begin{array}{ccc} F(x) & \xrightarrow{F(f)} & F(y) \\ {\scriptstyle \eta_x}\downarrow & & \downarrow{\scriptstyle \eta_y} \\ G(x) & \xrightarrow{G(f)} & G(y) \end{array}$$

commute, soit : $\eta_y \circ F(f) = G(f) \circ \eta_x$.

Exemple fondamental : la bidualité est une transformation naturelle. La naturalité se traduit par le fait suivant : si l'on enferme deux mathématiciens dans des pièces séparées et leur demande de construire un morphisme, il y a de grandes chances qu'ils construisent le même.

Si C,D sont petites, alors il existe une catégorie $\mathrm{Fun}(C,D)$ dont les objets sont les foncteurs $C \longrightarrow D$, et les morphismes sont les transformations. De même qu'avec les foncteurs, on peut parler d'isomorphismes naturels, etc.

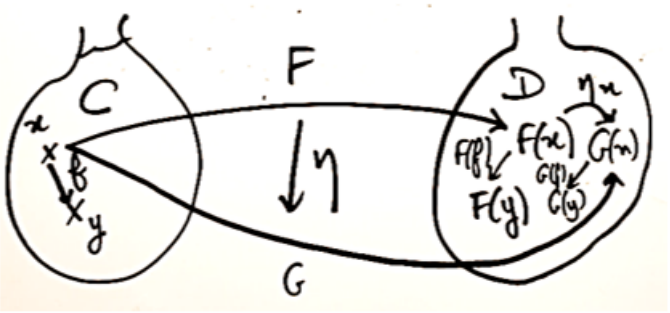


FIGURE 5.2.1 : *Transformation naturelle entre deux foncteurs entre deux mêmes catégories.* — La commutation se passe dans $\mathcal{D}$.

Propriété

Une transformation naturelle $\eta : F \longrightarrow G$ est inversible si et seulement si pour tout $x \in C$, $\eta_x \in \mathrm{Hom}(F(x),G(x))$ est un isomorphisme.
Dans ce cas, on parle d'*équivalence naturelle* ou d'*isomorphisme naturel*.

▷ C'est la définition. ■

On peut maintenant définir l'équivalence de catégories, qui est nettement moins forte que l'existence d'un foncteur inversible.

Définition. *(Équivalence de catégories)*

Soient C,D deux catégories. Une *équivalence de catégorie* est un foncteur $F : C \longrightarrow D$ tel qu'il existe $G : D \longrightarrow C$ tel que $FG \simeq Id_D$ et $DG \simeq Id_C$ par des isomorphismes naturels. On dit que G est un *quasi-inverse*.

Remarque. Le quasi-inverse est unique à équivalence naturelle près mais pas à unique équivalence naturelle près près (il est donc essentiellement unique, mais pas canonique). Plus précisément, deux quasi-inverses ne sont pas les mêmes, mais toujours isomorphes, et même le quasi-inverse fixé, l'équivalence naturelle de la composition à l'identité n'est pas unique (mais sont-elles naturellement isomorphes dans Fun ? Vous avez trois heures[1]).

[1] Dans la première heure, vous pouvez remarquer : naturellement équivalent dans Fun = naturellement équivalent naturellement équivalent

Définition-propriété. ***(Pleine fidélité, surjectivité essentielle)***

Un foncteur $F : C \longrightarrow D$ est une équivalence si et seulement si F est *pleinement fidèle* : $\forall x,y \in C \quad \mathrm{Hom}_C(x,y) \xrightarrow{f} \mathrm{Hom}_D(F(x),F(y))$ est bijectif ; et *essentiellement surjectif* : $\forall y \in D \; \exists c \in C \quad F(x) \simeq y$.

$\triangleright$ « $\Longrightarrow$ » Soit (G,η,μ) un quasi-inverse de F. Alors $GF \underset{y}{\simeq} Id_C$ et $FG \underset{\mu}{\simeq} Id_D$. Alors $\mathrm{Hom}_D(F(x),F(y)) \xrightarrow{G} \mathrm{Hom}_C(GF(x),GF(y)) \xrightarrow{y} \mathrm{Hom}_C(x,y)$ est l'inverse de $\mathrm{Hom}_C(x,y) \xrightarrow{F} \mathrm{Hom}((F(x),F(y))$. Donc F est pleinement fidèle. Si $y \in D$, $y \underset{\mu}{\simeq} F(G(y)) \in F(C)$, donc il est essentiellement surjectif.

Réciproquement, si F est essentiellement surjectif, pour tout $y \in D$, on choisit (si l'on croît à l'axiome du choix sur les classes) un objet $x \in C$, $\psi_y : F(x) \simeq y$. On pose $G(y) = x$. Pour tout $x \in C$, on a un isomorphisme $\psi_x : FG(y) \simeq y$. Pour $g : y_1 \longrightarrow y_2$ dans D, on vérifie que le diagramme suivant commute, c'est le seul truc qui pouvait poser problème.

$$\begin{array}{ccc} FGy_1 & \xrightarrow{FG(g)} & FGy_2 \\ {\scriptstyle \psi_{y_1}}\downarrow & & \downarrow{\scriptstyle \psi_{y_2}} \\ y_1 & \xrightarrow{g} & y_2 \end{array}$$

Ainsi $G(g) = \psi_{y_2}^{-1} g \psi_{y_1}$. Donc par construction, G est un foncteur et $\psi : FG \simeq Id_Y$ est naturel. Ceci définit $G(g)$.

Si maintenant $x \in C$,

$$\begin{array}{ccc} F(x) & \xrightarrow{F(\psi_x)} & FG(x) \\ & {\scriptstyle id_F}\searrow & \downarrow{\scriptstyle \varphi_{F(x)}} \\ & & F(x) \end{array}$$

donc G est un foncteur et ψ, φ donnent des isomorphismes $FG \simeq Id$, $GF \simeq Id$. ■

5.2.2 Premières propriétés

5.2.3 Le lemme de Yoneda

Chapitre 6

Exercices

Difficulté des exercices :

- ● ○ ○ ○ ○ Question de cours, application directe, exercice purement calculatoire sans réelle difficulté technique
- ● ● ○ ○ ○ Exercice faisable, soit intuitivement, soit en employant des moyens rudimentaires ou des techniques déjà vues
- ● ● ● ○ ○ Exercice relativement difficile et dont la résolution appelle à une réflexion plus importante à cause d'obstacles techniques ou conceptuels, qui cependant devraient être à la portée de la plupart des étudiants bien entraînés
- ● ● ● ● ○ Exercice très exigeant, destiné aux élèves prétendant aux concours les plus difficiles, exercice « classique ».
- ● ● ● ● ● La résolution de l'exercice requiert un raisonnement et des connaissances extrêmement avancés, dépassant les attentes du prérequis. Il est presque impossible de le mener à terme sans indication. Bien qu'exigibles à très peu d'endroits, ces exercices sont très intéressants et présentent souvent des résultats forts.

Appendice

Table des matières

Bibliographie

[1] *Titre du livre*, Auteur du livre, date, maison d'édition

Table des figures

Liste des tableaux