Filière MP - ENS de Paris-Saclay, Lyon, Rennes et Paris - Session 2022

Page de garde du rapport de TIPE

NOM : LASCAUD	Prénoms : Louis Maxime
Classe : MP	
Lycée : Gay-Lussac	Numéro de candidat : 14849
Ville : Limoges	

Concours auxquels vous êtes admissible, dans la banque MP inter-ENS (les indiquer par une croix) :

ENS Cachan	MP - Option MP		MP - Option MPI	
	Informatique			
ENS Lyon	MP - Option MP		MP - Option MPI	
	Informatique - Option M		Informatique - Option P	
ENS Rennes	MP - Option MP	X	MP - Option MPI	
	Informatique			
ENS Paris	MP - Option MP		MP - Option MPI	
	Informatique			

Matière dominante du TIPE (la sélectionner d'une croix inscrite dans la case correspondante) :

Informatique		Mathématiques	X	Physique	

Titre du TIPE : Généralisations du théorème de Cantor-Bernstein aux autres catégories que les ensembles

Nombre de pages (à indiquer dans les cases ci-dessous) :

Texte	37	Illustration	3	Bibliographie	1

Attention, les illustrations doivent figurer dans le corps du texte et non en fin du document !

Résumé ou descriptif succinct du TIPE (6 lignes, maximum) :
Après avoir exposé le théorème de Cantor-Bernstein ensembliste et l'avoir démontré, on cherche à l'étendre aux catégories concrètes usuelles. Typiquement, s'il existe entre deux structures deux morphismes injectifs en sens inverses, on se demande si celles-ci sont nécessairement isomorphes. On verra que la réponse dépend de la catégorie, et qu'un théorème peut être parfois établi moyennant des hypothèses supplémentaires.

À Limoges,

Le 9 juin 2022

Signature du professeur responsable de la classe préparatoire dans la discipline

Cachet de l'établissement

LYCÉE GAY-LUSSAC
12, Bd Georges Périn - 87000 LIMOGES
Tél. 05 55 79 70 01
ce.0870015u@ac-limoges.fr

Signature du (de la) candidat(e)

LASCAUD

La signature du professeur responsable et le tampon de l'établissement ne sont pas indispensables pour les candidats libres (hors CPGE).

Généralisations du théorème de Cantor-Bernstein aux autres catégories que les ensembles

Résumé

Après avoir exposé le théorème de Cantor-Bernstein ensembliste sur les cardinaux et l'avoir démontré, on cherche à le généraliser à d'autres catégories que celle des ensembles. Typiquement, s'il existe entre deux structures, par exemple, pour fixer les idées, des groupes, un morphisme injectif dans un sens, et un autre en sens inverse, ces deux groupes sont-ils nécessairement isomorphes, en plus d'être en bijection ? Par l'étude de propriétés de structure, on verra que la réponse varie selon les catégories considérées, et qu'un théorème de Cantor-Bernstein modulo une catégorie $\mathcal{C}$ peut-être établi au moyen d'hypothèses supplémentaires sur les morphismes ou les objets.

Abstract

Having shown and proven the Schröder-Bernstein theorem about cardinalities from set theory, we aim to generalise it to other categories than sets. Specifically, if an injective morphism exists between two groups, for instance, to set these ideas, and another backwards, are these two groups necessarily isomorphic, besides being equinumerous? Studying structure properties, we will see that the answers depend on the category considered, and that a Schröder-Bernstein theorem modulo a category $\mathcal{C}$ can be established by means of additional hypotheses both on morphisms or objects.

Introduction

Le théorème de Cantor-Bernstein, démontré à l'aube du XX$^{\text{e}}$ siècle et de la théorie des ensembles, généralise à tous les ensembles une propriété simple en milieu fini : si un ensemble a moins d'éléments qu'un deuxième, et si ce deuxième a moins d'éléments que le premier, alors ces deux ensembles sont équipotents, ce qui donne une méthode générique pour établir l'équipotence en permettant notamment de vérifier que le préordre cardinal est un ordre sur la classe de tous les ensembles, et prépare l'étude de la hiérarchie cardinale.

La généralisation de ce théorème prend forme spontanément au sein de la théorie des catégories : étant donné deux structures, nous nous demandons si l'existence de deux morphismes injectifs, en sens contraires l'un de l'autre, implique nécessairement leur isomorphie ; nous allons voir que la réponse dépend de la catégorie envisagée. Dans une première approche, on peut espérer que la vérification d'un théorème de Cantor-Bernstein spécifique à une catégorie donnée dépend de la malléabilité de la structure de ses objets (*i. e.*, un espace topologique a un « degré de structuration » moins fort qu'un espace vectoriel normé), intuition qui, soit dit en passant, reste en désaccord avec le théorème classique pour les ensembles pour lesquels aucune structure n'est donnée. Cette incohérence apparente nous poussera entre autres à considérer d'autres paramètres dans le principe général que nous venons d'énoncer.

Dans une première partie, nous voyons l'énoncé, la preuve et les conséquences du théorème de Cantor-Bernstein tel qu'il est pour les ensembles, avant de le généraliser ou d'en infirmer la généralisation dans un tour d'horizon méthodique des catégories usuelles. Avant cela, nous précisons également quelques résultats, parfois classiques, qui nous seront utiles dans la suite de notre composition.

1 Théorème de Cantor-Bernstein classique

1.1 Quelques notions en théorie des ensembles

1.1.1 Rappels axiomatiques

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.

Définition. *(Application)*
Soient E, F deux ensembles. On appelle *application* de E dans F un triplet (E, F, Γ) où Γ est une partie de $E \times F$ telle que pour tout $x \in E$, il existe un unique $y \in F$ tel que $(x, y) \in \Gamma$.

Il importe de distinguer cette définition de celle de *fonction*, plus générale, pour laquelle la locution « il existe un unique y » est remplacée par « il existe au plus un y » . Les applications peuvent vérifier plusieurs propriétés, telles l'injectivité (deux éléments distincts ont des images distinctes) et la surjectivité (tout élément d'arrivée est image d'au moins un élément). Ce sont elles, et leur conjonction, la *bijectivité*, qui nous intéressent dans la partie suivante afin de définir la notion pratique de nombre d'élément d'un ensemble, car, s'il existe toujours une application d'un ensemble dans un autre, il n'est pas certain que telle application puisse toujours vérifier les propriétés précédentes.

Avant d'y venir, nous devons rappeler un axiome de la théorie des ensembles, parfois omis parmi les précédents, que nous aurons besoin par la suite de postuler, en ayant garde de signaler toujours chacun de ses usages.

Définition. *(Axiome du choix)*
Soient I un ensemble et $(A_i)_{i \in I}$ une famille d'ensembles. Alors il existe une application σ de I dans $\bigcup\limits_{i \in I} A_i$ telle que, pour tout $i \in I$, $\sigma(i) \in A_i$. On note cet axiome AC et on dit que σ est une *fonction de choix*.

Avec la formulation des familles d'ensembles, il est équivalent de dire que le produit cartésien est intègre pour l'ensemble vide.

1.1.2 Cardinalité

Définition. *(Relation d'équipotence)*
On dit que deux ensembles A et B sont *équipotents*, ou *en bijection*, ou *isomorphes (au sens ensembliste)*, ou encore *qu'ils ont le même cardinal*, s'il existe une bijection $A \longrightarrow B$, ou ce qui est équivalent, s'il existe une bijection $B \longrightarrow A$. On note : $A \simeq B$.

Justification. Ces deux formulations sont bien équivalentes, à cause de la symétrie de la relation d'équipotence, dont nous montrons plus fortement ci-dessous que c'est une relation d'équivalence.

Propriété. La relation d'équipotence est une relation d'équivalence sur la classe des ensembles.

▷ La relation est réflexive, puisque $A \simeq A$ par l'application $id_A : x \longmapsto x$. La transitivité nous vient de la composition de deux bijections : si $f : A \longrightarrow B$ et $g : B \longrightarrow C$ sont bijections, alors $g \circ f : A \longrightarrow C$ est une bijection

[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 $F = \{E\}$.

donc $A \simeq B \wedge B \simeq C \Rightarrow A \simeq C$. Enfin, si $A \simeq B$ par l'application bijective f, alors $B \simeq A$ par l'application $h = f^{-1}$, dite *bijection réciproque* de f. En effet, si $y \in B$, alors $f^{-1}(y)$ est non vide par surjectivité de f et a au plus un élément d'après son injectivité, donc c'est un singleton, et l'on pose $h(y)$ l'unique élément de ce singleton, de sorte qu'on ait : $f \circ h = id_B$ et $h \circ f = id_A$. ■

Axiome. *(Cardinaux)*

On appelle *cardinal* ou *taille* tout représentant d'une classe d'équivalence de la relation d'équipotence (ce qui, foncièrement, signifie que tout ensemble est un cardinal et réciproquement...). C'est pourquoi on préfère fixer une *transversale* pour cette relation et fixer le cardinal commun à toute une classe d'équipotence, transversale que l'on appellera *classe des cardinaux*. Tous les éléments d'une classe d'équivalence de la relation d'équipotence sont donc isomorphes au sens ensembliste.

La transversale choisie par habitude contient notamment : $\emptyset$, les $[\![1, n]\!]$ pour $n \in \mathbb{N}^*$, l'ensemble $\mathbb{N}$ lui-même, $\mathcal{P}(\mathbb{N}) \simeq \mathbb{R}$, puis $\mathcal{P}(\mathbb{R})$, $\mathcal{P}(\mathcal{P}(\mathbb{R}))$, etc.

Dans le cas des ensembles finis, la construction de la cardinalité est plus simple. Nous en donnerons les rudiments plus tard dans notre composition.

Propriété. *(Comparaison cardinale)*

La relation définie sur la classe des cardinaux par $C \hookrightarrow C'$ si et seulement s'il existe une injection de C dans C', est un ordre sur la classe des cardinaux, ou ${}^{\mathrm{Set}}\!/_{\simeq}$.

▷ L'identité de tout ensemble id_E, qui est toujours injective, montre la réflexivité de cette relation. La transitivité provient de la composition de deux injections, qui reste en tous cas une injection. Enfin, l'antisymétrie de cette relation est l'énoncé du théorème de Cantor-Bernstein. ■

Remarques.

1. On dit parfois que C est *subpotent* à C'. Ce terme est tombé en désuétude.
2. Si l'on se restreint à la classe des cardinaux non nuls (l'ensemble vide est l'unique élément de sa classe d'équivalence d'après l'axiome d'extensionnalité), la relation $C \hookrightarrow C'$ est symétrisable et équivaut à la relation $C' \twoheadrightarrow C$ définie par : il existe une surjection de C' dans C.

   ▷ La démonstration de ce résultat fait intervenir l'axiome du choix noté AC, que nous postulons (notre système d'axiomes est ZFC, pour Zermelo-Fraenkel + Choix). Nous y reviendrons dans les sections suivantes.
3. On peut se demander, assez naturellement, si cet ordre est total. Cette affirmation est en fait équivalente à l'axiome du choix, ce qui constitue le *théorème de comparabilité cardinale*. La démonstration repose sur le lemme de Zorn, qui équivaut lui-même à AC.
4. Si $E \subseteq F$, l'injection canonique $x \longmapsto x$ permet d'écrire $E \hookrightarrow F$. Ainsi $\subseteq$ est plus forte que $\hookrightarrow$, ce qui signifie que son graphe, au sens des classes, est inclus dans $\Gamma_{\hookrightarrow}$.

Avant d'en faire davantage sur les cardinaux, nous nous permettons une (fausse) digression.

1.1.3 Considérations sur l'axiome du choix

Nous montrons ici une formulation équivalente classiques de l'axiome du choix, le lemme de Zorn. Ce théorème n'est pas fondamentalement lié au théorème de Cantor-Bernstein classique, mais nous l'utiliserons lorsqu'il s'agira de constructions dans les autres catégories.

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

Exemple. On considère l'ensemble A des parties strictes d'un ensemble E. Si a est un élément de E, alors l'élément $E - \{a\}$ de A est un élément maximal pour l'inclusion (mais a priori ce n'est pas un maximum).

Précisions. Tout maximum (aussi appelé *plus grand élément*) est un élément maximal, et c'est même alors l'unique élément maximal de E. Par contre, un élément maximal n'est pas nécessairement un maximum. Les deux notions coïncident lorsque l'ordre $\leqslant$ est total.

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).

Propriété. *(Lemme de Zorn)*
Tout ensemble inductif a un élément maximal.

$\triangleright$ 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 $\forall x \in X, x \leqslant a$ 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 $K' = \{C \in K \mid \forall D \in K,\ C \subseteq D \vee D \subseteq C\}$. 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 $K' = K$. ■

1.1.4 Arithmétique cardinale

Dans la suite, nous aurons l'occasion d'utiliser un résultat de mathématiques supérieures à la limite du programme. Nous le rappelons ci-dessous, avec les propriétés qui lui sont directement liées.

Propriété. Soient f, g des applications quelconques. On suppose que $f : E \longrightarrow F$ et $g : F \longrightarrow E$.
(i) si $g \circ f$ est injective, alors f est injective ;
(ii) si E est non vide, f est injective si et seulement si elle est inversible à gauche. Cet inverse, appelé *rétraction*, est toujours surjectif ;
(iii) si $g \circ f$ est surjective, alors g est surjective ;
(iv) g est surjective si et seulement si elle est inversible à droite. Cet inverse, appelé *section*, est toujours injectif.

▷ Montrons ces quatre résultats.

1. Soient f de E dans F et g de F dans G deux applications. On suppose que la composée $g \circ f$ est injective. Montrons que f l'est également. Soient x, y deux éléments de E. Alors si $f(x) = f(y)$, il suffit de composer par g pour obtenir $g \circ f(x) = g \circ f(y)$. Or $g \circ f$ est injective par hypothèse, donc $x = y$, ce qu'il fallait prouver.

2. Soit f une application de E dans F inversible à gauche, c'est-à-dire qu'il existe une application g de F dans E telle que $g \circ f = id_E$. En particulier, $g \circ f$ est injective, donc f est injective. De plus $g \circ f$ est surjective donc la propriété *(iii)* garantit indépendamment que g est surjective. Réciproquement, supposons que f soit injective. On va construire une rétraction g telle que $g \circ f = id_E$. oit $x_0 \in E$ fixé (c'est là qu'on utilise le fait que E est non vide). On pose la fonction

$$\begin{aligned} g : F &\longrightarrow E \\ x &\longmapsto \begin{cases} f^{-1}(x) \text{ si } x \in \mathrm{Im}(f) \\ x_0 \text{ sinon,} \end{cases} \end{aligned}$$

$f^{-1}(x)$ étant dans le premier cas un singleton, car il est non vide par définition de l'image et a au plus un élément par injectivité de f. Alors, pour tout $t \in E$, $f(t) \in \mathrm{Im}(f)$ donc $g \circ f(t) = t$ par construction, ce qui montre que $g \circ f = id_E$.

3. Soient f de E dans F et g de F dans G deux applications. On suppose que la composée $g \circ f$ est surjective. Montrons que g l'est également. Soit x un élément de E. Soit $x' \in E$ tel que $g \circ f(x') = x$, qui existe par hypothèse. Alors $f(x') = y$ est un élément de F tel que $g(y) = x$, ce qui montre que g est surjective.

4. Soit g une application de F dans E inversible à droite, c'est-à-dire qu'il existe une application f de E dans F telle que $g \circ f = id_E$. En particulier, $g \circ f$ est surjective, donc g est surjective. De plus $g \circ f$ est injective donc la propriété *(i)* garantit indépendamment que f est injective. Réciproquement, supposons que g soit surjective. On va construire une section f telle que $g \circ f = id_E$. Soit x un élément de E. g étant surjective, $g^{-1}(x)$ est non vide. On prend f une fonction de choix sur $(g^{-1}(x))_{x \in E}$. Alors par construction, on a $g[f(t)] \in \{t\}$, car si $t \in E$, $f(t) \in g^{-1}(t)$, d'où $g \circ f = id$. ■

Remarque. L'énoncé *(iv)* est en fait équivalent à l'axiome du choix.

On montre également ce résultat que nous avons évoqué précédemment :

Théorème. *(Symétrisation de l'ordre cardinal)*
S'il existe une injection de E dans F, et E est non vide, alors il existe une surjection de F dans E. S'il existe une surjection de A dans B, alors il existe toujours une injection de B dans A.

$\triangleright$ C'est une conséquence directe des propriétés *(ii)* et *(iv)* précédentes. ■

Remarques. Il faut donc considérer le cas de l'ensemble vide à part : $\emptyset$ s'injecte dans tout ensemble, mais seul l'ensemble vide s'injecte dans $\emptyset$, car il n'y a que l'ensemble vide qui admette une application partant de lui à valeurs dans l'ensemble vide (ce qui correspond à l'intuition que $\emptyset$ est le plus petit ensemble). D'autre part, tout ensemble se surjecte dans l'ensemble vide, mais l'ensemble vide se surjecte seulement dans l'ensemble vide. Toutes ces propriétés sont démontrées à l'aide des règles sur les quantificateurs.

1.2 Théorème de Cantor-Bernstein

1.2.1 Preuve

On démontre maintenant le théorème dont il est question dans notre composition, énoncé en 1887 par Cantor et démontré en 1895 par lui au moyen de l'axiome du choix. En 1896, Felix Bernstein le démontre sans l'axiome du choix. Ernst Schröder croit le démontrer en 1898, mais il doit corriger sa preuve en 1902. Dedekind avait également démontré le théorème en 1887.

Propriété. *(Théorème de Cantor-Schröder-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.

Remarque. Ce théorème est plutôt évident dans le cas des ensembles finis, comme nous le développons au paragraphe suivant.

Il existe un grand nombre de preuves du théorème de Cantor-Bernstein : pour référence, nous en donnons une constructive et qui ne fait pas recours à l'axiome du choix. Nous l'appelons *preuve 0*. D'autres démonstrations classiques existent ; il n'est pas possible pourtant de donner ici un inventaire exhaustif, la multitude des preuves ayant été rédigées par les mathématiciens pour le théorème de Cantor-Bernstein dépassant largement le cadre de notre exposé.

Preuve 0. On se propose d'expliciter ici en ligne la démonstration du théorème, puisqu'elle est d'une importance majeure pour notre composition.

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 $\tilde{g} : B \longrightarrow \mathcal{I}m(g) \subseteq A$, 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 $h : A \longrightarrow \mathcal{I}m(g)$, c'est-à-dire un bijection de A sur son sous-ensemble $\mathcal{I}m(g)$ 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 $A_0 = \complement_A \mathcal{I}m(g)$, 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 :

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

Vérifions que h est une application bijective.

- ▷ Elle est bien définie partout sur son ensemble de définition.
- ▷ Elle est aussi à valeurs dans $\mathcal{I}m(g)$, puisque par construction, $g \circ f$ envoie $\bigcup_{n=0}^{+\infty} A_n$ sur $\bigcup_{n=1}^{+\infty} A_n \subseteq \mathcal{I}m(g)$, et que si $a \notin \bigcup_{n=0}^{+\infty} A_n$, alors en particulier $a \notin A_0 = \complement_A \mathcal{I}m(g)$ donc $a = h(a) \in \mathcal{I}m(g)$.
- ▷ 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 $y \in \mathcal{I}m(g)$, 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 $\bigcup_{n=0}^{+\infty} A_n \subseteq A$.

  2e cas. $y \notin \bigcup_{n=1}^{+\infty} A_n$. 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.

Heuristique. Nous pouvons interpréter la preuve précédente par une métaphore qui n'est pas sans rappeler celle de l'hôtel de Hilbert. Des gens (*a priori* une infinité de personnes) ont des places de cinéma, un cinéma *a priori* infini. Correctement, deux personnes différentes reçoivent des places différentes : c'est l'injection f des gens vers les places du cinéma. Ce n'est pas forcément une surjection, car toutes les places du cinéma n'ont pas forcément été réservées, mais f étant bien une application, c'est-à-dire définie sur tout A, tous les gens considérés ont réservé une place. Le jour de la séance arrive, mais au lieu de s'assoir à leurs places, la foule arrive et certaines personnes (pas toutes, car g n'est pas forcément surjective) s'assoient au hasard de manière à remplir toute la salle de cinéma : c'est l'autre injection, g, des places vers les gens. (Bien sûr, l'intuition donnée par le cas fini nous laisse déjà penser qu'il faut qu'il y ait autant de gens que de places pour qu'une telle disposition soit possible.) Le patron du cinéma parvient sur les lieux et veut mettre un peu d'ordre : il ordonne d'abord aux gens restés dehors, qui constituent $A_0 = \complement_A \mathrm{Im}(g)$, d'entrer et d'aller à leur place réservées : autrement dit, on cherche son image par f. Chaque fois qu'une personne vient légitimement réclamer sa place, l'occupant indu, qu'on trouve en composant par g (on a donc un nouvel élément de A donné par $g \circ f$) rejoint celle qu'il avait lui-même réservée, et ainsi de suite. Notons que cette nouvelle personne n'est bien sûr par dans A_0, et peut être

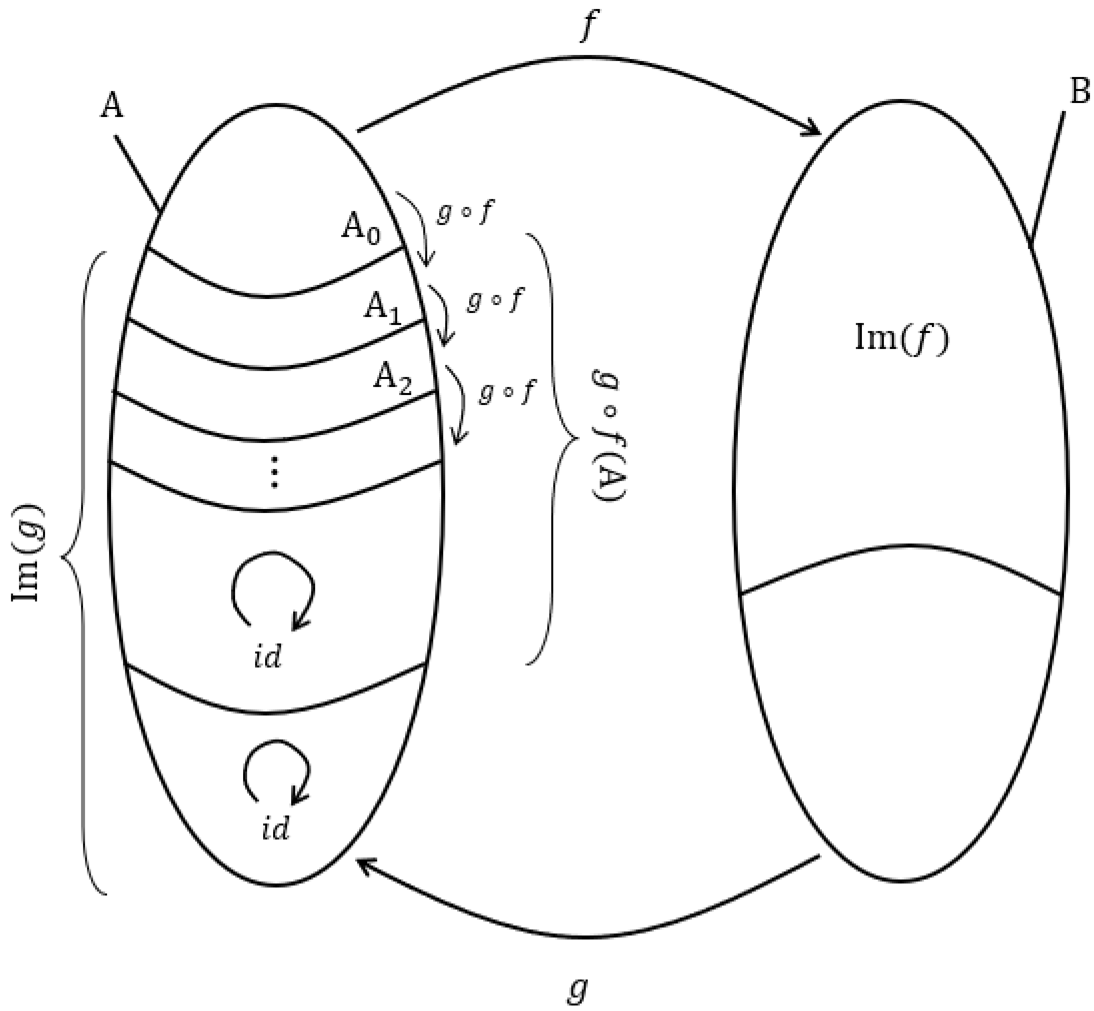

Figure 1. — *Illustration de la preuve 0 du théorème de Cantor-Bernstein.* Les petites flèches (les trois marquées $g \circ f$ et les deux arrondies marquées id) situées sur A représentent l'application h qui envoie A sur Im(g). Les éléments qui sont dans l'un des A_i, y compris $A_0 = \complement_A \mathrm{Im}(g)$, sont envoyés dans le suivant A_{i+1} et s'accumulent en restant dans $\bigcup\limits_{n=1}^{+\infty} A_n$. S'ils sont dans ce qui reste de $g \circ f(A)$, ils sont envoyés sur eux-mêmes. De même, enfin, s'ils sont dans la partie de Im(g) restante.

soit mal placée, et c'est le cas où l'on itère infiniment (d'où l'importance des tailles infinies des réservations ou du cinéma), soit bien placée. Il reste encore à placer les gens qui ne sont pas restés dehors : soit ces personnes se sont initialement bien placées (c'est la part la plus basse du diagramme précédent), soit elles sont bien placées à force de réarrangements orchestrés par le patron du cinéma (c'est la part immédiatement supérieure) : on fait en sorte, évidemment, que ces personnes restent à leur place ; soit encore elle sont mal placées et alors il est garanti qu'à un moment ou à un autre une personne arrivera réclamer son siège. Voilà une correspondance biunivoque entre les gens ayant réservé des places et les gens étant rentrés sans permission. Le théorème assure qu'à la fin, tout le monde est assis et aucune place n'est vide : c'est la bijection $g^{-1} \circ h$ entre les places et les gens.

Comme évoqué précédemment, d'autres preuves existent.

1.2.2 Le cas des ensembles finis

Nous adoptons la définition la plus courante des ensembles finis, qui pré-existe à une théorie des ensembles axiomatique.

Définition. *(Ensemble fini)*
Un ensemble E est fini s'il existe un entier $n \in \mathbb{N}$ tel que E soit en bijection avec $[\![1, n]\!]$.

Remarque. L'ensemble vide est donc fini, puisqu'on peut l'écrire comme $[\![1, 0]\!]$.

Partant de cette définition, il est facile de légitimer les propriétés usuelles des ensembles finis en utilisant le résultat technique suivant.

Théorème. Toute injection d'un ensemble fini dans lui-même est bijective.

▷ Il suffit de montrer le résultat pour les intervalles $[\![1, n]\!]$ d'après la définition donné ci-dessus. On raisonne par récurrence sur $n \geqslant 1$, le résultant étant trivial pour l'ensemble vide, la seule application à considérer étant alors la flèche vide $\emptyset_\emptyset$ qui est à la fois injective et bijective. Pour $n = 1$, l'intervalle $\{1, ..., n\}$ est le singleton $\{1\}$ et la seule application de cet ensemble dans lui-même est l'identité, qui est une bijection, ce qui constitue notre initialisation. Supposons maintenant $n \geqslant 1$, et soit f une injection de $[\![1, n]\!]$ dans lui-même. Posons $m = f(n)$. On définit une application g de $[\![1, n-1]\!]$ dans lui-même en posant $g(i) = f(i)$ si $f(i) < m$ et $g(i) = f(i) - 1$ si $f(i) > m$. Alors g est injective puisque m n'appartient pas à l'image de $[\![1, n]\!]$ par f. Par hypothèse de récurrence, g est donc surjective. Par construction, $\mathrm{Im}(f) = \mathrm{Im}(g) \cup \{n\}$, d'où $\mathrm{Im}(f) = [\![1, n]\!]$, ce qui signifie que f est elle-même surjective, donc f est une bijection, ce qu'il fallait montrer. ■

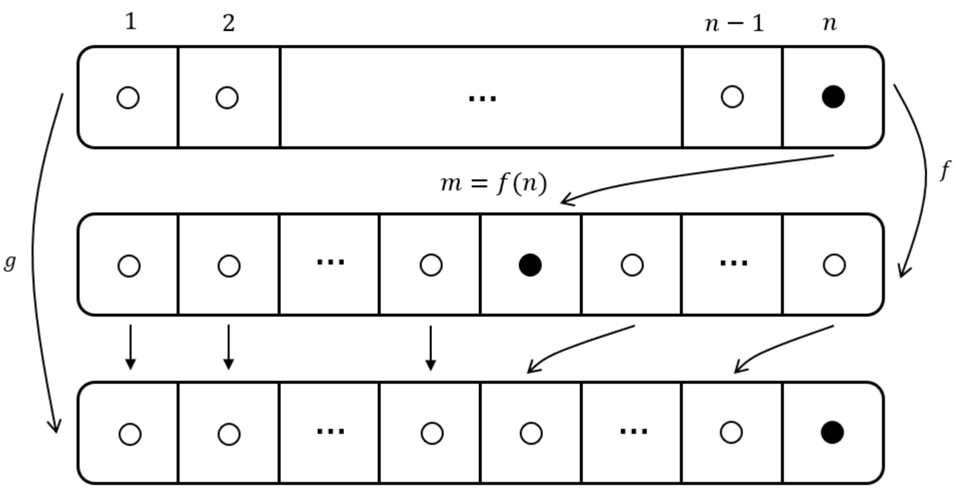

FIGURE 2. — *Illustration de la preuve du théorème fondamental sur les ensembles finis.* Les ensembles représentés contiennent des cercles blancs (si un cercle est noir, c'est que l'élément n'appartient pas à l'ensemble considéré).

À partir de là, il y a deux façons, plus ou moins élégantes, de conclure.

Première méthode.

Propriété. Tout ensemble fini est en bijection avec un unique intervalle $[\![1,n]\!]$ de $\mathbb{N}$, n étant un entier naturel quelconque. L'entier n est alors appelé *cardinal* de l'ensemble.

▷ Par définition, si E est un ensemble fini, il est en bijection avec au moins un intervalle, disons $[\![1,p]\!]$, $p \in \mathbb{N}$. Il suffit de montrer que cet entier p est unique. C'est bien le cas, car pour $q \neq p$, on a une bijection de E sur $[\![1,q]\!]$ d'où une bijection g de $[\![1,p]\!]$ sur $[\![1,q]\!]$. Sans perte de généralité, on peut prendre $q < p$. Dans ce cas, $[\![1,q]\!]$ s'injecte canoniquement dans $[\![1,p]\!]$ par f. Dans ce cas, $f \circ g$ est une injection de $[\![1,p]\!]$ dans lui-même. Cependant, elle ne peut être bijective, directement parce que f n'est pas surjective. Cela contredit le théorème précédent. ■

On peut maintenant démontrer le cas particulier du théorème de Cantor-Bernstein pour les ensembles finis. En effet, si A s'injecte dans B, alors $\mathrm{card}(A) \leqslant \mathrm{card}(B)$, et de même, si B s'injecte dans A, $\mathrm{card}(B) \leqslant \mathrm{card}(A)$. Or dans le cas des ensembles finis, $\mathrm{card}(A)$ et $\mathrm{card}(B)$ sont des entiers naturels, donc par antisymétrie de la relation d'ordre sur les entiers, on a $\mathrm{card}(A) = \mathrm{card}(B)$, ce qui signifie que A et B sont en bijection, dans ce cas, tous deux avec $[\![1,\mathrm{card}(A)]\!]$, et l'on conclut par transitivité.

Deuxième méthode.

Elle est plus rapide et plus élégante, mais demande d'utiliser le résultat démontré en **1.1. 4.**. Nous pouvons alors conclure.

Théorème. *(Théorème de Cantor-Bernstein pour les ensembles finis)*
Soient A, B deux ensembles finis. Si A s'injecte dans B et B s'injecte dans A, alors A et B sont en bijection.

▷ Bien que cet énoncé soit une conséquence du théorème de Cantor-Bernstein général dont nous avons déjà dressé la preuve, nous en donnons deux preuves plus élémentaires : la première a été fournie précédemment. D'autre part, il suffit de remarquer que, si $f : A \hookrightarrow B$ et $g : B \hookrightarrow A$, alors $g \circ f$ est une injection de A dans A. D'après le théorème fondamental, c'est une bijection. En particulier, elle est inversible, disons par h, et donc $g \circ f \circ h = id_A$. D'après *(iii)*, g est surjective, puisqu'elle est inversible à droite par l'application $f \circ h$. C'est donc une bijection, de B sur A, et c'est terminé. ■

1.2.3 Remarques

Compte-tenu des remarques liées à la relation de subpotence et notamment à sa relation duale, on peut énoncer quelques formulations duales du théorème de Cantor-Bernstein.

Théorème. *(Théorème de Cantor-Bernstein dual)*
Si E se surjecte dans F et F et se surjecte dans E, E et F sont en bijection.

Théorème. *(Théorème de Cantor-Bernstein univoque)*
Si E s'injecte et se surjecte dans F, E et F sont en bijection.

Remarque. Ce n'est pas tautologique du tout. Ce le serait si l'injection et la surjection considérées étaient la même application.

Théorème. *(Équivalence)*
Le théorème de Cantor-Bernstein, sa formulation duale et sa version univoque sont logiquement équivalentes.

▷ On utilise à l'envi le deuxième théorème de la partie **1. 1. 4.**. Supposons le théorème de Cantor-Bernstein vrai. Si E se surjecte dans F, alors F s'injecte dans E d'après cette propriété. De même E s'injecte dans F, et en appliquant le théorème de Cantor-Bernstein classique, on en déduit que E et F sont en bijection, ce qui suffit pour montrer le théorème de Cantor-Bernstein dual. Supposons maintenant le théorème dual vérifié et montrons le théorème. Dans le cas où E est non vide, si E s'injecte dans F, alors F se surjecte dans E. Comme E se surjecte dans F par hypothèse, on peut appliquer le théorème dual, et E et F sont équipotents. Dans le cas où E est vide, comme E se surjecte dans F, d'après la remarque sur l'ensemble vide faite à la suite de la propriété de symétrisation de l'ordre cardinal, F est vide, donc E et F sont tous les deux égaux à l'ensemble vide ; *a fortiori* ils sont en bijection. Ceci montre le théorème de Cantor-Bernstein univoque. Supposons enfin ce dernier théorème vérifié. Supposons donc que E s'injecte dans F et que F s'injecte dans E. Si F est non vide, alors E se surjecte dans F, et en appliquant le théorème univoque, on en déduit CB. Si F est vide, alors E s'injecte dans F donc E est vide, et E et F sont équipotents. Ceci montre le théorème de Cantor-Bernstein à partir de sa version univoque. Par conséquence, les trois énoncés sont deux à deux équivalents. ■

1.3 Comment généraliser le théorème de Cantor-Bernstein

1.3.1 La théorie des catégories

On rappelle le formalisme essentiel des catégories, issu de la théorie des catégories qui prolonge celle des ensembles et se place dans le cadre 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.

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

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).

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).

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 $\mathcal{C}$, $F(g \circ f) = F(g) \circ F(f)$.

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 $f = g$.

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 $F(f)$.

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.

Nous introduisons enfin la notion capitale pour généraliser le théorème de Cantor-Bernstein aux catégories concrètes.

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 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.

D'autres catégories, parmi les plus usuelles, sont parfois notées de la façon suivante :

Catégorie	Objets	Morphismes
Set ou Ens	Ensembles	Applications
Ord	Ensembles ordonnées	Applications croissantes
Tot-Ord	Ensembles totalement ordonnées	Idem
$\mathbb{K}$-Vect	$\mathbb{K}$-espaces vectoriels	Applications linéaires
Met	Espaces métriques	Morphismes pour la topologie métrique
Top	Espaces topologiques	Applications continues
Mon	Monoïdes	Morphismes de monoïdes
Grp	Groupes	Morphismes de groupes
Ab	Groupes abéliens	Idem
Ring	Anneaux (unitaires)	Morphismes d'anneaux unitaires
ACU	Anneaux commutatifs unitaires	Idem
Krp	Corps	Morphismes de corps
$\mathbb{K}$-Alg	$\mathbb{K}$-algèbres	Morphismes d'algèbres

Tableau 2. — *Catégories usuelles*

1.3.2 Un petit peu de vocabulaire

Définition. *(Théorème de Cantor-Bernstein modulo une catégorie concrète)*

On appelle *$\mathcal{C}$-théorème de Cantor-Bernstein* ou *théorème de Cantor-Bernstein modulo une catégorie concrète* $\mathcal{C}$, la proposition : « pour tous objets A et B de $\mathcal{C}$, s'il existe un morphisme injectif de A vers B et s'il existe un morphisme injectif de B vers A, alors A et B sont isomorphes ». On note : $\mathcal{C}$-CB, en hommage à l'abréviation courante du théorème de Cantor-Bernstein en anglais : CBT.

Remarques.

1. Cette définition fait donc varier les définitions d'objets mais surtout de morphismes, qui définissent également la notion d'isomorphie dans une catégorie donnée. Deux objets sont isomorphes si et seulement s'il existe un *isomorphisme* de l'un vers l'autre, mais un isomorphisme n'est pas toujours seulement un morphisme bijectif ! Par définition, un isomorphisme est une application
2. De ce que nous savons, le Set-théorème de Cantor-Bernstein est vrai (c'est le théorème de Cantor-Bernstein classique). D'après la remarque qui le précède, on sait aussi que le $\mathcal{C}$-théorème de Cantor-Bernstein, avec $\mathcal{C}$ la catégorie des ensembles finis, est vrai. Plus généralement, **si $\mathcal{C}$-CB et $\mathcal{C}'$ est une sous-catégorie *pleine* de $\mathcal{C}$, alors le $\mathcal{C}$-théorème de Cantor-Bernstein implique le $\mathcal{C}'$-théorème de Cantor-Bernstein**.
3. Il s'agit bien sûr de ne pas se perdre. Le morphisme injectif dans le sens inverse n'est pas a priori la surjection réciproque du premier associée par l'axiome du choix, car celle-ci c'est pas a priori un morphisme (il suffit de se référer à la preuve dans **1. 1. 4.** pour s'en convaincre).

Définition. *(Catégorie bernsteinienne)*
Une catégorie concrète $\mathcal{C}$ est dite *bernsteinienne*, si $\mathcal{C}$-CB. La *bernsteinté*, ou *Cantor-Bernsteinté* de la catégorie désigne son caractère bernsteinien.

Remarque. Cette terminologie est inspirée d'un terme dont nous avons trouvé une occurrence hapactique dans la littérature mais tout à fait intéressante : « Cantor-Bernsteinness ».

Nous pouvons dès à présent nous rendre compte de l'intérêt direct des catégories bernsteiniennes, c'est-à-dire dans lesquelles le théorème CB est vérifié : ces catégories sont naturellement munies d'une relation d'ordre, au sens large des classes, donnée par $A \leqslant B$ s'il existe un morphisme injectif de A dans B, dès lors que l'on quotiente cette catégorie par sa relation d'isomorphie, toujours au sens des classes. En effet, pour un objet donné, l'identité, qui existe toujours par axiome, fournit un morphisme injectif de A dans A, d'où la réflexivité. D'autre part, l'axiome de composition des morphismes ainsi que la propriété de composition des injections en théorie des ensembles donne la propriété de transitivité. Ainsi, modulo l'isomorphie, $\mathcal{C}$-CB consiste exactement en l'antisymétrie de ce préordre et donc en l'existence ou non d'un ordre cohérent sur $\mathcal{C}/\simeq$.

Cette cohérence est toute à fait intuitive et a forcé, dans la définition générale des théorèmes de Cantor-Bernstein, l'introduction de la notion de morphisme injectif. En effet, l'existence d'une injection ne suffit plus pour caractériser les structures au sein d'une catégorie. Considérons par exemple toutes les $\mathbb{R}$-espaces vectoriels de dimension finie ; modulo la relation d'isomorphie, la catégorie obtenue se ramène à une catégorie petite entièrement décrite par la famille des $(\mathbb{R}^n)_{n\in\mathbb{N}}$, en abusant de la notation élidant les lois d'espace vectoriel. Nous remarquons alors que, ensemblistement, mis à part le cas pathologique $n = 0$, tous les ensembles décrits par cette famille sont en bijection (voir dans la section suivante pour une preuve). Pour l'ordre cardinal, chacun est donc plus petit que tous les autres, ce qui rend l'utilisation de cet outil totalement caduque. En revanche, si nous imposons maintenant, pour l'ordre, qu'il existe une application linéaire injective, alors nous savons, d'après le cours, que la dimension de l'espace de départ doit être inférieure ou égale à la dimension de l'espace d'arrivée. Autrement dit, nous avons ordonné les espaces de cette catégorie simplifiée selon leurs dimensions, et nous savons également que la dimension, en dimension finie, caractérise les classes d'isomorphie des espaces vectoriel (on le rappelle, en passant par l'existence d'une base de cardinal unique n). Cet exemple nous fournit l'intuition complète pour généraliser l'ordre cardinal aux autres catégories comme nous l'avons fait. Nous avons, au demeurant, seulement généralisé mot à mot en théorie des catégories le lexique spécifique des ensembles.

Par surcroît, la généralisation du théorème de Cantor-Bernstein se voit d'ores et déjà monter des barrières directes. En observant la démonstration fournie ci-dessus (preuve 0), on se convainc aisément que l'application exhibée, mettant en bijection A et B, n'est pas un morphisme dans la plupart des catégories concrètes, par exemple algébriques : d'une part, parce que h est définie par morceaux sur les A_i et le complémentaire de leur réunion, ce qui entame l'espoir que l'application ainsi définie ait la régularité propre aux morphismes, d'autre part, parce que l'application g^{-1} n'est pas forcément un morphisme. Par exemple, dans certaines catégories, la réciproque d'un morphisme bijectif (ici, on aurait bien un morphisme injectif corestreint à son image) n'est pas nécessairement un morphisme comme nous l'avons vu et donc ce morphisme bijectif n'est pas un isomorphisme au sens des catégories. Ces considérations permettent de penser que la généralisation de Ens-CB n'est pas du tout immédiate, et nous allons voir, dans les catégories usuelles, dans quels cas celui-ci subsiste et tenterons d'en extraire des conditions générales pour que le théorème puisse s'étendre.

2 Vers un théorème de Cantor-Bernstein topologique

2.1 Cas des espaces vectoriels : rupture de symétrie avec l'analogue ensembliste ?

2.1.1 Un $\mathbb{K}$-Vect-théorème de Cantor-Bernstein

Soit $\mathbb{K}$ un corps commutatif fixé dont on passe les lois sous silence. On remarque que le théorème de Cantor-Bernstein modulo la catégorie des espaces vectoriels sur $\mathbb{K}$ de dimension finie sur $\mathbb{K}$, pour la même raison que dans le cas des ensembles finis, est trivialement vrai, ce qui est dû à l'analogie entre finitude pour les ensembles et dimension finie pour les espaces vectoriels, et parallèlement applications et applications linéaires. Précisons cette analogie : entre ensembles finis de même cardinal, toute application injective et bijective (nous l'avons démontré dans la section sur-précédente). Entre espaces vectoriels de dimensions finies égales, toute application linéaire, c'est-à-dire morphisme d'espaces vectoriels, injective et en réalité bijective, c'est-à-dire un isomorphisme. Ceci est conséquence directe du théorème du rang : si $f : E \longrightarrow F$ est une application linéaire injective, alors $\mathrm{Ker}(f) = \{0\}$ par caractérisation. Or $\dim(E) = n = \dim \mathrm{Ker}(f) + \dim \mathrm{Im}(f) = \dim \mathrm{Im}(f)$, mais le seul sous-espace vectoriel de dimension n d'un espace de dimension n est l'espace lui-même, ce qui signifie que f est surjective et termine la preuve. Nous le reformulons ci-dessous.

Propriété. *(Théorème de Cantor-Bernstein modulo la catégorie des espaces vectoriels de dimension finie)*
Si deux espaces vectoriels de dimension finie sur $\mathbb{K}$ s'injectent réciproquement l'un dans l'autre aux moyens d'applications linéaires, alors ces deux espaces sont isomorphes.

▷ On sait que leurs dimensions sont réciproquement inférieures l'une à l'autre, donc sont égales puisque ces grandeurs sont entières. On en déduit que les espaces E et F sont isomorphes, par transitivité en passant par $\mathbb{K}^{\dim(E)}$. ■

Naturellement, on se demande s'il y a rupture de symétrie dans le cas de la dimension infinie, autrement dit, si le $\mathbb{K}$-Vect-théorème de Cantor-Bernstein est infirmé. On se rend compte que ce n'est pas le cas, grâce à la théorie de la dimension.

Commençons par établir le résultat fondamental de la dimension infinie quelconque que nous connaissons déjà en dimension finie.

Théorème. *(Théorème de la base incomplète)*
Dans un espace vectoriel E quelconque sur $\mathbb{K}$, de toute famille génératrice $\mathcal{G}$, pour toute famille libre $\mathcal{L}$ incluse dans $\mathcal{G}$, on peut trouver une famille de vecteurs contenue dans $\mathcal{G}$ et contenant tous les vecteurs de $\mathcal{L}$ qui soit une base de E.

▷ Cet énoncé est plus précis que ce dont nous avons besoin, mais nous le montrons sous cette forme cependant. Soit E un espace vectoriel sur $\mathbb{K}$, $\mathcal{G}$ une famille génératrice de vecteurs de E, c'est-à-dire telle que tout vecteur de E se décompose comme combinaison linéaire finie de vecteurs de $\mathcal{G}$, et $\mathcal{L}$ une famille libre de vecteurs de E, c'est-à-dire telle que toute combinaison linéaire d'un nombre fini de ses vecteurs et nulle soit triviale.

Soit $\mathcal{F}$ l'ensemble des familles libres telles que $\mathcal{L} \subseteq \mathcal{L}' \subseteq \mathcal{G}$, en identifiant sans problème une famille à l'image de son support. L'ensemble $\mathcal{F}$ est alors partiellement ordonné par l'inclusion en tant que sous-ensemble de l'ensemble des parties de E. Montrons que $\mathcal{F}$ est un ensemble inductif. Il est non vide, car par hypothèse, il contient $\mathcal{L}$. Soit C une chaîne de $\mathcal{F}$, c'est-à-dire une partie de $\mathcal{F}$ totalement ordonnée, et montrons qu'elle est majorée dans $\mathcal{F}$. Cette chaîne, on peut l'écrire : $C = \{\mathcal{L}_i \mid i \in I\}$ où I est totalement ordonné, avec donc pour tous $i, j \in I$, $i \leqslant j \Longrightarrow \mathcal{L}_i \subseteq \mathcal{L}_j$. Dans ce cas, la famille $\mathcal{L}' = \bigcup_{i \in I} \mathcal{L}_i$ est libre. En effet, si $x_1, ..., x_n$ appartiennent à $\mathcal{L}'$, si l'on a des scalaires vérifiant $\lambda_1 x_1 + ... \lambda_n x_n = 0$, pour tout $j \in [\![1, n]\!]$, il existe $i_j \in I$ tel que $x_j \in \mathcal{L}_{i_j}$. Soit $i_0 = \max_{1 \leqslant j \leqslant n} i_j$. Puisque la famille $(\mathcal{L}_i)_{i \in I}$ est totalement ordonnée, on a $x_1, ..., x_n \in \mathcal{L}_{i_0}$. La famille $\mathcal{L}_{i_0}$ étant libre, on en déduit $\lambda_1 = ... = \lambda_n = 0$. Ainsi, la famille $\mathcal{L}$ est libre et vérifiant évidemment $\mathcal{L} \subseteq \mathcal{L}' \subseteq \mathcal{G}$, on a $\mathcal{L}' \in \mathcal{F}$. Comme $\mathcal{L}_i \subseteq \mathcal{L}'$ pour tout i, $\mathcal{L}'$ est un majorant de C.

D'après le lemme de Zorn, l'ensemble $\mathcal{F}$ possède un élément maximal $\mathcal{B}$. On va montrer que cette famille est une base de E, telle que $\mathcal{L} \subseteq \mathcal{B} \subseteq \mathcal{G}$. Par définition, c'est une famille libre et par construction elle vérifie les inclusions précédentes. Il ne reste donc qu'à voir qu'elle est génératrice. Soit $x \in E$ et montrons que x s'écrit comme combinaison linéaire d'éléments de $\mathcal{B}$. Il suffit de traiter le cas où $x \in \mathcal{G}$, puisque $\mathcal{G}$ est-elle même génératrice de E. Si $x \in \mathcal{B}$, il n'y a rien à faire. Supposons donc $x \notin \mathcal{B}$. Si la famille $\mathcal{B} \cup \{x\}$ était libre, alorsce serait un élément de $\mathcal{F}$ contenant strictement $\mathcal{B}$, ce qui contredit la maximalité de $\mathcal{B}$. Ainsi, la famille $\mathcal{B} \cup \{x\}$ est liée, et c'est terminé, en effet : il existe des scalaires $\mu_1, ..., mu_n, \mu$ tels que $\mu_1 x_1 + ... + \mu_n x_n + \mu x = 0$ où $\mu_1, ..., \mu_n$ sont non tous nuls et $x_1, ..., x_n \in \mathcal{B}$. Si $\mu = 0$, alors puisque $\mathcal{B}$ est libre, tous les scalaires jusqu'à n sont nuls, donc tous les μ_i, μ sont nuls ce qui est exclu. Ainsi, on peut écrire $x = -\mu^{-1}(\mu_1 x_1 + ... + \mu_n x_n)$. x s'écrit donc comme combinaison linéaire d'éléments de $\mathcal{B}$, ce qu'il fallait montrer. ■

Notons au passage que ce théorème dépend du lemme de Zorn, et donc de l'axiome du choix. Pour montrer $\mathbb{K}$-Vect-CB, on justifie d'abord des lemmes de construction qui généralisent les propriétés classiques de la dimension finie.

Propriété. *(Existence de bases)*
Tout espace vectoriel admet des bases.

▷ Rappelons pour commencer que l'espace nul a pour base la famille vide. Ce cas pathologique ayant été exclu, on peut considérer un vecteur non nul $x_0 \in E$, et dans ce cas, $\{x_0\}$ forme une famille libre. Elle est incluse dans la famille trivialement génératrice $(x)_{x \in E}$. On applique alors le théorème de la base incomplète, et c'est terminé. ■

Théorème. *(Théorème de la dimension)*
Dans tout espace vectoriel, le cardinal de toute partie libre est inférieur au sens de l'ordre cardinal au cardinal de toute partie génératrice de E.

▷ Soient $\mathcal{L}$ une partie libre de E et $\mathcal{G}$ une partie génératrice. Dans le cas où $\mathcal{G}$ est finie, par définition E est de dimension finie et le théorème d'échange du programme garantit que $\mathcal{L}$ est elle-même finie de cardinal inférieur à celui de $\mathcal{G}$. Dans le cas général, pour tout vecteur $l \in \mathcal{L}$, avec l'axiome du choix, choisissons une partie finie $f(l)$ de G telle que l appartienne à $\mathrm{Vect}(f(l))$. Pour tout $K \in \mathrm{Fin}(\mathcal{G})$ l'ensemble des parties finies de $\mathcal{G}$, d'après le cas fini, on a $\mathrm{card}(f^{-1}(\{K\})) \leqslant \mathrm{card}(K) \leqslant \aleph_0$, dont on déduit que $\mathrm{card}(\mathcal{L}) = \sum_{K \in \mathrm{Fin}(\mathcal{G})} \mathrm{card}(f^{-1}(\{K\})) \leqslant \sum_{K \in \mathrm{Fin}(\mathcal{G})} \aleph_0 \leqslant \mathrm{card}(\mathrm{Fin}(\mathcal{G}))\aleph_0 = \mathrm{card}(\mathrm{Fin}(\mathcal{G})) = \mathrm{card}(\mathcal{G})$ d'après les propriétés de l'arithmétique cardinale, soit $\mathrm{card}(\mathcal{G}) \geqslant \mathrm{card}(\mathcal{L})$. ■

Corollaire. *(Unicité de la dimension)*
Sur un espace vectoriel quelconque, toutes les bases ont le même cardinal. Celui-ci définit alors la *dimension* de l'espace vectoriel considéré.

▷ Soient $\mathcal{B}, \mathcal{B}'$ deux bases d'un espace vectoriel E. Alors $\mathcal{B}$ est libre et $\mathcal{B}'$ est génératrice, donc d'après ce qui précède, $\mathrm{card}(\mathcal{B}) \leqslant \mathrm{card}(\mathcal{B}')$. Mais ces deux bases ayant un rôle rigoureusement symétrique, on a $\mathrm{card}(\mathcal{B}') \leqslant \mathrm{card}(\mathcal{B})$. **En utilisant le théorème de Cantor-Bernstein**, par antisymétrie de l'ordre cardinal, on a $\mathrm{card}(\mathcal{B}) = \mathrm{card}(\mathcal{B}')$, ce qu'il fallait montrer. Notons que la définition de la dimension est ainsi univoque grâce aux résultats conjoints du corollaire du théorème de la base incomplète, qui garantit l'existence, et du théorème de la dimension pour les espaces vectoriels, qui garantit l'unicité. ■

On termine la théorie élémentaire de la dimension infinie par un résultat très intéressant dont nous n'aurons vraiment besoin que de la réciproque.

Corollaire. *(Caractérisation des classes d'isomorphie)*
Deux espaces vectoriels sont isomorphes si et seulement s'ils ont la même dimension.

▷ Rappelons que pour que deux espaces vectoriels soient isomorphes, il suffit qu'il existe une application linéaire bijective de l'un vers l'autre, la réciproque d'une telle application étant nécessairement un morphisme d'après le cours. Supposons que E, F soient deux espaces vectoriels sur $\mathbb{K}$ qui soient isomorphes, par f de E dans F. On vérifie aisément (voir ci-dessous la preuve du $\mathbb{K}$-Vect-théorème de Cantor-Bernstein) qu'un isomorphisme f transforme un famille libre en famille libre et une famille génératrice en famille génératrice. Il transforme donc une base en une base. Puisqu'une bijection préserve le cardinal, par définition, on a donc une base de F de même cardinal qu'une base de E. Par unicité de la dimension, E et F sont de même dimension. Réciproquement, si E est de dimension I, où I est un ensemble quelconque, alors E est isomorphe à $\mathbb{K}^{(I)}$ l'ensemble des applications de I dans $\mathbb{K}$ à support fini. Par transitivité sur cet ensemble, deux espaces de même dimension sont isomorphes. Justifions enfin que $E \simeq \mathbb{K}^{(I)}$. L'isomorphisme à considérer est celui qui à tout vecteur de E envoie la famille de ses composantes dans une base de cardinal I fixée, cette famille étant à support fini par définition du caractère générateur. On vérifie facilement qu'elle est linéaire, surjective par les axiomes de stabilité des espaces vectoriels, et pour l'injectivité, celle-ci vient de la liberté des bases. ■

On peut maintenant conclure.

Lemme. *(Majoration du cardinal des familles libres)*
Dans un espace vectoriel quelconque, toute famille libre est de cardinal inférieur à cette dimension.

▷ C'est déjà fait. ■

Théorème. *($\mathbb{K}$-Vect-théorème de Cantor-Bernstein)*
Si deux espaces vectoriels quelconques sur $\mathbb{K}$ s'injectent réciproquement l'un dans l'autre aux moyens d'applications linéaires, alors ces deux espaces sont isomorphes.

▷ Soient $(\mathbb{E}, +, .)$ et $(\mathbb{F}, +, .)$ deux espaces vectoriels sur le corps $\mathbb{K}$; pour simplifier, on note identiquement leurs lois respectives, mais ce n'est pas nécessaire, et l'on ne perd pas de généralité. Supposons qu'il existe une application linéaire injective $f : \mathbb{E} \longrightarrow \mathbb{F}$ et une autre application linéaire injective $g : \mathbb{F} \longrightarrow \mathbb{E}$. Soient maintenant $\mathcal{B}$ une base de $\mathbb{E}$ et $\mathcal{B}'$ une base de $\mathbb{F}$, l'existence de ces bases étant garantie dans tout espace vectoriel quelconque par le lemme de Zorn et donc l'axiome du choix. $\mathcal{B}$ étant une base, c'est en particulier une famille $\mathbb{K}$-libre de l'espace $\mathbb{E}$, et l'application f étant un morphisme (une application linéaire), $f(\mathcal{B})$ est une famille libre de l'espace $\mathbb{F}$. En effet : si $n \in \mathbb{N}, \lambda_1, ..., \lambda_n \in \mathbb{K}$, si $y_1, ..., y_n \in f(\mathcal{B})$, il existe des $x_1, ...x_n \in \mathcal{B}$ tels que $f(x_1) = y_1, ..., f(x_n) = y_n$, et :

$$\begin{aligned}\lambda_1 y_1 + ... + \lambda_n y_n = 0 &\Rightarrow \lambda_1 f(x_1) + ... + \lambda_n f(x_n) = 0 \\ &\Rightarrow f(\lambda_1 x_1 + ... + \lambda_n x_n) = 0 \text{ car } f \in \mathcal{L}(\mathbb{E}) \\ &\Rightarrow \lambda_1 x_1 + ... + \lambda_n x_n = 0 \text{ car } f \text{ est injective} \\ &\Rightarrow \lambda_1 = ... = \lambda_n = 0,\end{aligned}$$

puisque $\mathcal{B}$ est libre. Si l'on identifie sans problème une famille à l'image de son support, l'application f étant injective, on peut écrire canoniquement : $\mathcal{B} \hookrightarrow f(\mathcal{B})$. Or par liberté, d'après le lemme, on a $f(\mathcal{B}) \hookrightarrow \mathcal{B}'$, d'où, par transitivité, $\mathcal{B} \hookrightarrow \mathcal{B}'$. Symétriquement, on obtient $\mathcal{B}' \hookrightarrow \mathcal{B}$. D'après le théorème de Cantor-Bernstein classique pour les ensembles, on peut en déduire que $\mathcal{B} \simeq \mathcal{B}'$, ce qui signifie que $\mathbb{E}$ et $\mathbb{F}$ ont la même dimension sur $\mathbb{K}$, autrement dit qu'ils sont isomorphes. ■

2.1.2 Applications

Intéressons-nous au cas particulier de $\mathbb{R}$ et $\mathbb{R}^2$ munis de la structure d'espace vectoriel sur $\mathbb{R}$. Ces deux ensembles sont en bijection, car on sait, d'après l'axiome du choix, que tout ensemble infini est équipotent à son carré cartésien. On peut néanmoins le démontrer avec des moyens élémentaires.

Théorème. $\mathbb{R}$ et $\mathbb{R}^2$ sont en bijection.

$\triangleright$ Montrons ce résultat à l'aide du théorème de Cantor-Bernstein. $\mathbb{R}$ s'injecte dans $\mathbb{R}^2$ par l'application $x \longmapsto (x, 0)$. 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 :

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

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 $(0,a_1a_2a_3...;0,b_1b_2b_3...)$, $(0,a_1'a_2'a_3'...;0,b_1'b_2'b_3'...) \in [0,1[$, les représentations ici étant des développements illimités propres. Supposons que $\varphi(0,a_1a_2a_3...;0,b_1b_2b_3...) = \varphi(0,a_1'a_2'a_3'...;0,b_1'b_2'b_3'...)$, c'est-à-dire $0,a_1b_1a_2b_2a_3b_3... = 0,a_1'b_1'a_2'b_2'a_3'b_3'...$, 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 $0,a_1a_2a_3... = 0,a_1'a_2'a_3'...$ d'une part et $0,b_1b_2b_3... = 0,b_1'b_2'b_3'...$ d'autre part, ce qui montre l'injectivité de φ. ■

Remarquons par là que ce théorème donne que tous les $\mathbb{R}^n$ sont en bijection, ce que nous avons évoqué dans nos développements précédents et postulé seulement. En effet, on établit que $\mathbb{R}$ est en bijection avec $\mathbb{R}^n$ par récurrence sur $n \in \mathbb{N}^*$, l'initialisation étant la réflexivité de la relation d'ordre cardinale et l'hérédité s'écrit en remarquant aisément que $\mathbb{R}^{n+1} \simeq \mathbb{R} \times \mathbb{R}^n$, et puisque $\mathbb{R}^n \simeq \mathbb{R}$, on a $\mathbb{R}^{n+1} \simeq \mathbb{R} \times \mathbb{R} = \mathbb{R}^2 \simeq \mathbb{R}$, et tous ces ensembles ont donc la puissance du continu.

On sait donc que $\mathbb{R}$ et $\mathbb{R}^2$ sont en bijection. Cependant, si ces deux structures sont ainsi isomorphes du point de vue ensembliste, on sait qu'elle ne le sont pas dans la catégorie $\mathbb{R}$-Vect, car $\dim_{\mathbb{R}}(\mathbb{R}) = 1$ et $\dim_{\mathbb{R}}(\mathbb{R}^2) = 2$, et deux espaces vectoriels sont isomorphes si et seulement s'ils ont la même dimension. Or $x \longmapsto (x, 0)$ est un morphisme de $\mathbb{R}$ dans $\mathbb{R}^2$, tandis que l'autre application n'en est pas. On peut se demander s'il existe un morphisme injectif à la place de l'application que nous avons posée. Le théorème de Cantor-Bernstein pour les espaces vectoriels sur un même corps donne qu'il n'existe pas de telle application.

Cette application du théorème est d'autant plus intéressante en dimension finie, où elle permet d'établir directement l'inexistence de certains morphismes injectifs. Pour donner un exemple, on peut considérer $\mathbb{R}[X]$ et $\mathbb{R}(\mathbb{X})$. $\mathbb{R}[X]$ s'injecte canoniquement dans $\mathbb{R}(\mathbb{X})$ par un morphisme, mais ces deux espaces vectoriels sur $\mathbb{R}$ ne sont pas isomorphes, car ils n'ont pas la même dimension : en effet, $\mathbb{R}[X]$ est de dimension $\aleph_0$ (soit $\mathbb{N}$) en exhibant la famille $(X_n)_{n\in\mathbb{N}}$, mais $\mathbb{R}(X)$ est de dimension au moins $2^{\aleph_0}$ (soit $\mathbb{R} \simeq \mathcal{P}(\mathbb{N})$) d'après le théorème de décomposition en éléments simples, en exhibant la famille des $(\frac{1}{X-\alpha})_{\alpha\in\mathbb{R}}$. On en déduit qu'il n'existe pas d'application linéaire injective du corps des fractions rationnelles dans les polynômes à une indéterminée.

2.1.3 Commentaire

Remarquons simplement l'élément central de la preuve du $\mathbb{K}$-Vect théorème de Cantor-Bernstein. En effet, parmi les preuves précédentes, nous avons abondamment utilisé le fait que les classes d'isomorphie dans la catégorie des espaces vectoriels sur un même corps commutatif étaient caractérisés de façon exacte par leur dimension, c'est-à-dire le cardinal commun à toutes ses bases. On dispose donc, dans la catégorie $\mathbb{K}$-Vect, d'un *invariant* d'isomorphie, comme nous en donnons une définition plus précise ci-dessous pour notre étude.

Définition. *(Invariant d'isomorphie)*
Soit $\mathcal{C}$ une catégorie. On appelle *invariant d'isomorphie*, tout foncteur de $\mathcal{C}$ dans Set qui passe au quotient pour la relation d'isomorphie $\simeq$ de $\mathcal{C}$ et la relation d'équipotence sur Set en un foncteur tel que l'application quotient sur les objets soit injective.

Dans notre exemple, le foncteur f considéré est celui qui à un espace vectoriel, associe sa dimension (en ayant fixé au préalable une transversale de cardinaux). En effet, deux espaces vectoriels isomorphes ont la même dimension, c'est-à-dire que f est constant sur les classes d'isomorphie, autrement dit, il passe au quotient pour la relation d'équivalence $\simeq$: « être isomorphe à ». L'application quotient (où nous avons généralisé le théorème de factorisation aux applications entre classes propres) sur les objets est alors injective, car à deux espaces non isomorphes, on associe deux dimensions et donc cardinaux différents. On s'aperçoit, au travers de cet exemple, qu'une catégorie sur laquelle il existe un invariant d'isomorphie vérifiera généralement un théorème de Cantor-Bernstein, en appliquant justement celui-ci aux invariants d'isomorphie.

Nous aurons l'occasion, dans la suite, de préciser cette notion et d'en voir les limites.

2.2 Approche topologique

2.2.1 Négation de Top-CB

On rappelle la structure d'espace topologique et le lien avec celle d'espace métrique.

Définition. *(Espace topologique)*
Un espace topologique $(E, \mathcal{O})$ est la donnée d'un ensemble E et de $\mathcal{O}$ une famille de parties de E non vide, stable par réunion et par intersection finie, appelée famille des *ouverts* de E.

Théorème. *(Ouverts d'un métrisable)*
Les ouverts d'un espace métrique (E, d) sont les réunions quelconques de boules ouvertes pour d. En particulier, tout espace métrique est un espace topologique.

$\triangleright$ Soit (E, d) un espace métrique et pour tout $x \in E$, pour tout $r \in \mathbb{R}$, on note $B(x, r)$ la boule centrée en x de rayon r.

Vérifions que cette définition correspond bien à celle des espaces vectoriels normés, classiquement, une partie O de E est ouverte si $\forall x \in O \ \exists r > 0 \quad B(x, r) \subseteq O$. Soit O un ouvert pour cette définition. Pour tout $x \in O$, il existe $r > 0$ tel que $B(x, r) \subseteq O$. Ainsi le réel strictement positif $r_x = \sup\{r > 0, B(x, r) \subseteq O\}$ est bien défini pour tout x et on a $O = \bigcup_{x \in O} B(x, r_x)$. Réciproquement, on vérifie aisément que toute réunion quelconque de boules ouvertes est un ouvert pour la définition ci-dessus.

On note $\mathcal{O}$ l'ensemble des réunions de boules ouvertes. Montrons que $\mathcal{O}$ définit une topologie sur E comme famille

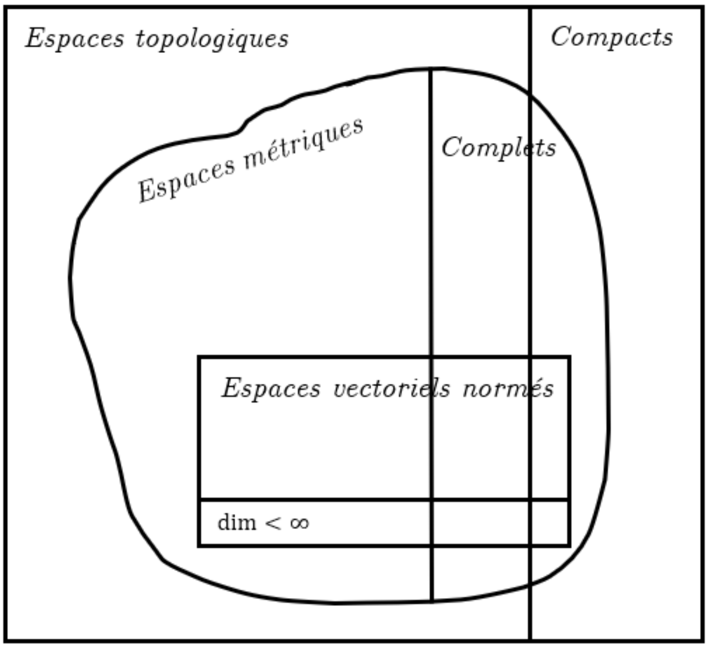

FIGURE 3. — *Structures topologiques usuelles et leur hiérarchisation.* La complétude est une notion purement métrique (elle est couplée à celle de suite de Cauchy). Tout espace vectoriel normé est un espace métrique pour la distance issue de la norme $d(x, y) = \|x - y\|$; tout espace métrique est un espace topologique pour la topologie métrique décrite ci-dessus.

d'ouverts. Cette famille est non vide ; elle contient, par exemple, l'ensemble vide en prenant $r < 0$ ou encore E en l'écrivant $\bigcup_{x \in E} B(x, 0)$. Elle est stable par réunion quelconque par associativité de la réunion. La stabilité par intersection finie se vérifie facilement avec l'autre définition dont on vient de montrer l'équivalence. ■

Ces bases ayant été posées, on peut désormais donner un contre-exemple très simple avant d'infirmer le théorème de Cantor-Bernstein sur la catégorie des espaces topologiques.

Théorème. *(Infirmation de Top-CB)*
$\neg$ Top-CB.

▷ Prenons les espaces métriques, munis de la distance usuelle, $E = \mathbb{R}$ et $F = \mathbb{R}_+$. Ce sont des espaces topologiques. L'injection canonique de F dans E est évidemment continue ; d'autre part, la fonction exponentielle fournit une injection continue de E dans F. Cependant, E et F ne sont pas homéomorphes, car F privé du point $\{0\}$ est encore connexe par arcs tandis que E privé de n'importe quel point ne l'est plus. ■

Ainsi, il n'y a pas de théorème de Cantor-Bernstein topologique. L'intuition de ce résultat provient de ce que, si les espaces vectoriels de l'algèbre linéaire sont une structure très « rigide », du fait de ce que tout espace vectoriel admet des bases qui en réduisent considérablement la difficulté de l'étude, les espaces métriques et les espaces topologiques *a fortiori* sont des structures relativement « molles ». Pour exemple, il est possible de munir tout ensemble d'une structure topologique au moyen de la distance discrète : $d(x, y) = \delta_x^y$, où δ est le symbole de Kronecker. En revanche, on ne peut pas munir n'importe quel ensemble d'une structure d'espace vectoriel.

2.2.2 Le cas des espaces métriques

Le plus généralement possible, Met-CB est faux.

Théorème. *(Infirmation de Met-CB)*
$\neg$ Met-CB.

▷ C'est déjà fait. ■

Notons en même temps que le contre-exemple dans le cas de la catégorie $\mathcal{V}$ des espaces vectoriels normés de dimensions infinies donnait déjà un contre-exemple pour Met-Cb ainsi que Top-CB : en effet, d'après la remarque dans **1. 3. 2.**, $\mathcal{V}$ étant une sous-catégorie pleine de Met et ensuite de Top, par contraposée, ni Met ni Top ne sont bernsteiniennes. Le contre-exemple donné pour Top est beaucoup plus élémentaire néanmoins.

2.3 Renforcements des propriétés des espaces pour que leur catégorie soit bernsteinienne

Il n'y a pas de théorème de Cantor-Bernstein pour les espaces topologiques connexes, c'est-à-dire qui ne s'écrivent pas comme réunion de deux ouverts disjoints non triviaux. En effet, le contre-exemple $E = \mathbb{R}$ et $F = \mathbb{R}_+$ donne deux espaces topologiques connexes (ce sont des intervalles, ils sont donc connexes par arcs donc connexes).

Ce n'est pas du tout surprenant : les espaces connexes ne sont pas munis de rigidité particulière, selon l'intuition, qui permettrait de démontrer un théorème CB ; en particulier, la connexité n'est pas stable pour les opérations usuelles. En revanche, les notions de complétude et de compacité fournissent cette même rigidité, respectivement grâce au théorème de l'application ouverte et à la notion de recouvrement fini par des ouverts.

2.3.1 Espaces topologiques compacts

De façon plutôt inattendue, on n'établit pas de théorème de Cantor-Bernstein pour les espaces compacts. On peut donner un contre-exemple assez rudimentaire :

Théorème. La catégorie des espaces topologiques compacts est non bernsteinienne.

▷ On prend deux espaces topologiques définis comme sous-espaces métriques des réels pour la distance usuelle, avec $E = [0, 1]$ et $F = [0, 1] \cup [2, 3]$. Un segment de $\mathbb{R}$ est compact en tant que fermé borné en dimension finie, donc E est compact et F est compact comme réunion de deux compacts. Ces deux espaces ne sont pas homéomorphes, car l'un est connexe en tant qu'intervalle des réels et l'autre n'est pas connexe en tant que réunion de deux fermés disjoints non triviaux. Pourtant, E s'injecte continûment dans F par l'injection canonique découlant de l'inclusion, et F s'injecte dans E en envoyant $[0, 1]$ dans $[0, 1/4]$ (par une application affine bien choisie) et en envoyant $[2, 3]$ dans $[3/4, 1]$. Cette injection, en tant que recollement de deux injections à images disjointes, est continue sur $[0, 1]$ et continue sur $[2, 3]$, donc continue sur F. ■

On aurait pu s'attendre à un théorème vérifié pour les espaces compacts, car dans une certaine mesure, les espaces compacts jouent, parmi les espaces topologiques, le rôle des ensembles finis parmi les ensembles. Le théorème suivant justifie pourquoi.

Théorème. *(Bijection continue sur un compact)*
Tout morphisme bijectif sur un compact restreint à son image séparée est un isomorphisme au sens topologique, c'est-à-dire, toute bijection continue sur un compact est un homéomorphisme.

$\triangleright$ Soient E, F deux espaces topologiques. Soit f une application de E dans F. On suppose que E est compact et on considère le sous-espace topologique $F' = \text{Im}(f)$, muni de la topologie induite. On suppose également que f est bijective de E dans F et continue. Montrons que f est un homéomorphisme : il suffit de vérifier qu'elle est bicontinue, soit ici que l'application réciproque $g = f^{-1}$ est continue. On procède par caractérisation par images réciproques de fermés. Soit I un fermé de $E = \text{Im}(g)$. E étant compact, I est un fermé dans un compact ; on sait alors que I est un compact. Alors $g^{-1}(I) = f(I)$, qui est compact en tant qu'image continue d'un compact. En particulier, c'est un fermé de F, car F est séparé. Ainsi, g est continue, ce qu'il fallait montrer. ■

On donne enfin la définition topologique des espaces compacts, afin de donner toute sa généralité à notre paragraphe.

Théorème. *(Propriété de Borel-Lebesgue)*
Un espace topologique est compact si et seulement s'il est séparé et que de tout recouvrement de l'espace par des ouverts, on peut extraire un sous-recouvrement fini.

2.3.2 Espaces vectoriels normés complets

Nous donnons ici une construction rapide de la notion de complétude. On fixe (E, d) un espace métrique quelconque.

Définition. *(Suite de Cauchy)*
Une suite $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ est dite *de Cauchy* si $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n, p \geqslant N \quad d(u_n, u_p) \leqslant \varepsilon$, ou de façon équivalente, $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geqslant N\ \forall k \geqslant 0 \quad d(u_{n+k}, u_n) \leqslant \varepsilon$.

Remarque. Dans les espaces vectoriels normés, on écrit parfois $\|u_p - u_q\| \longrightarrow_{p,q\to+\infty} 0$.

Propriété. Toute suite convergente est de Cauchy.

$\triangleright$ Soit (u_n) une suite convergente ; on note l sa limite. Soit $\varepsilon > 0$. Alors il existe N_1 tel que pour tout $n \geqslant N_1$, $d(u_n, l) \leqslant \frac{\varepsilon}{2}$. De même, il existe N_2 tel que pour tout $p \geqslant N_2$, $d(u_p, l) \leqslant \frac{\varepsilon}{2}$. Dans ce cas, si l'on note $N = \max(N_1, N_2)$, alors pour tous $n, p \leqslant N$, les inégalités précédentes sont vérifiées, et par inégalité triangulaire, $d(u_n, u_p) \leqslant d(u_n, l) + d(l, u_p) = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, ce qui signifie que la suite (u_n) est une suite de Cauchy. ■

Propriété. Toute suite de Cauchy est bornée.

$\triangleright$ On copie ici la preuve pour les suites convergentes. Soit (u_n) une suite de Cauchy à valeurs dans E. Pour $\varepsilon = 1$, il existe N tel que pour tout $n, p \geqslant N$, $d(u_n, u_p) \leqslant 1$. En particulier, $d(u_n, u_N) \leqslant 1$. On pose $M = \max(d(u_0, u_N), d(u_1, u_N), ..., d(u_{N-1}, u_N), 1)$, les N premiers termes étant en nombre fini. (u_n) est alors bornée par M. ■

Propriété. Toute suite de Cauchy admettant une sous-suite convergente converge.

$\triangleright$ Soit (u_n) une suite de Cauchy, et φ une extractrice donnant une sous-suite convergente. On note l la valeur d'adhérence associée. Soit $\varepsilon > 0$. Il existe alors N_1 tel que pour tout $n \geqslant N_1$, $d(u_{\varphi(n)}, l) \leqslant \frac{\varepsilon}{2}$ par définition de la limite. D'autre part, il existe N_2 tel que tous $n, p \geqslant N_2$, $d(u_n, u_p) \leqslant \frac{\varepsilon}{2}$ puisque (u_n) est de Cauchy. Or pour toute extractrice

φ, pour tout $n \in \mathbb{N}$, $\varphi(n) \geqslant n$, donc en appliquant la relation précédente, pour tout $n \geqslant N_1$, $d(u_{\varphi(n)}, u_n) \leqslant \frac{\varepsilon}{2}$. Ainsi, en notant $N = \max(N_1, N_2)$, pour tout $n \geqslant N$, on a $d(u_n, l) \leqslant d(u_n, u_{\varphi(n)}) + d(u_{\varphi(n)}, l) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, ce qui signifie que (u_n) a pour limite l. En particulier elle converge. ■

Définition. *(Complétude)*
Un espace métrique dans lequel toute suite de Cauchy converge est dit *complet.*

Définition. *(Espaces de Banach)*
Un espace vectoriel normé complet est appelé *espace de Banach.*

Remarques importante. Tout espace vectoriel normé de dimension finie sur $\mathbb{R}$ ou $\mathbb{C}$ est un espace de Banach. En effet, soit (u_n) une suite de Cauchy à valeurs dans $\mathbb{R}^n$ pour fixer les idées. Cette suite est bornée comme on l'a vu ; en particulier, on peut l'inclure dans un pavé de la forme $[-M, M]^n$. Or un tel pavé est un compact, car en dimension finie, les compacts sont exactement les fermés bornés. Nous pouvons appliquer le théorème de Bolzano-Weierstrass et extraire ainsi de (u_n) une sous-suite convergente ; d'après ce qui précède, la suite (u_n) est donc convergente, ce qu'il fallait montrer.

Il existe des espaces vectoriels normés de dimensions infinies qui sont également des espaces de Banach, mais tous les espaces vectoriels normés ne sont pas de Banach.

Nous énonçons une propriété importante pour la suite.

Propriété. *(Fermé dans un complet)*
Toute partie fermée d'un espace complet est complète.

$\triangleright$ Soit (E, d) un espace métrique complet et A une partie fermée de E. Montrons que, munie de la distance induite, A est un espace métrique complet. Soit donc (u_n) un suite de Cauchy de A. Cette notion étant extrinsèque, c'est aussi une suite de Cauchy de E, donc elle converge dans E. Mais A est fermée, et comme (u_n) est à valeurs dans A qui est fermée, sa limite qui existe est à valeurs dans A. Autrement, dit (u_n) converge dans A. On a montré que A était complète. ■

On en déduit le résultat suivant.

Propriété. *(Sous-espaces de Banach)*
Les sous-espaces de Banach d'un espace de Banach sont ses sous-espaces vectoriels fermés.

$\triangleright$ Soit E un espace de Banach, muni de lois et d'une norme. Soit F un sous-espace vectoriel de E qui soit une partie fermée. Alors E est complet d'après la propriété précédente, c'est donc un espace de Banach, c'est-à-dire un sous-espace de Banach de E. Réciproquement, soit F un sous-espace de Banach de E. C'est un sous-espace vectoriel de E par définition. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans F qui converge vers une limite dans E. C'est une suite convergente de F donc une suite de Cauchy de F. Comme F est complet, c'est une suite convergente de F autrement dit sa limite est dans F. Ainsi F est fermé. ■

Remarque. En particulier, tout sous-espace vectoriel de dimension finie d'un espace de Banach en est un sous-espace de Banach, car tout sous-espace vectoriel de dimension finie d'un espace vectoriel normé de dimension quelconque est fermé, d'après le cours.

Nous concluons notre construction des espaces de Banach par une caractérisation que nous utiliserons par la suite de la complétude des espaces vectoriels normés.

Propriété. *(Absolue convergence et convergence des séries dans les Banach)*
Un espace vectoriel normé est complet si et seulement si toute série absolument convergente est convergente.

▷ Seul le sens direct nous intéresse véritablement. Soit E un espace de Banach. Soit (x_n) une suite d'éléments de E telle que $\sum \|x_n\|$ converge. Alors les sommes $S_N = \sum\limits_{n \leqslant N} x_n$ vérifient pour tout $M \geqslant N$:

$$\|S_M - S_N\| = \|\sum_{n=N+1}^{M} x_n\| \leqslant \sum_{n=N+1}^{M} \|x_n\|.$$

Or, la série $\sum \|x_n\|$ étant convergente, les tranches de Cauchy du dernier membre tendent vers zéro lorsque M, N tendent vers $+\infty$. Ainsi, la suite $(S_N)_{N\in\mathbb{N}}$ est de Cauchy. Comme E est complet, elle converge, c'est-à-dire que la série est convergente.

On donne la réciproque par souci d'exhaustivité. Supposons que toute série absolument convergente converge dans E. Soit (x_n) une suite de Cauchy à valeurs dans E. On peut extraire une sous-suite $(x_{\varphi(n)})$ telle que $\|x_{\varphi(n+1)} - x_{\varphi(n)}\| \leqslant 1/2^n$, en choisissant $\varphi(n) = N$ le module associé à $\varepsilon = 1/2^n$. On pose alors $u_n = x_{\varphi(n+1)} - x_{\varphi(n)}$ et la série $\sum u_n$ est absolument convergente, donc convergente par hypothèse. Or $x_{\varphi(n+1)} - x_{\varphi(0)} = \sum\limits_{k=0}^{n} u_k$, donc on déduit que la sous-suite $(x_{\varphi(n)})$ converge. Comme toute suite de Cauchy admettant une sous-suite convergente converge, on en déduit la convergence de (x_n), et E est un espace de Banach. ■

On mentionne d'abord un résultant facilitant la notion d'isomorphisme dans les espaces de Banach, le théorème de l'application ouverte. En effet, la notion intuitive de morphismes dans les espaces de Banach correspond aux applications linéaires continues (on rappelle qu'en dimension finie, la continuité est une conséquence de la linéarité). Celui-ci est une conséquence du théorème de Baire, dont nous donnons également une démonstration, lui-même déduit du théorème des fermés emboîtés.

Théorème. *(Théorème des fermés emboîtés)*
Dans tout espace métrique complet, toute intersection décroissante de fermés non vides dont les diamètres tendent vers zéro est réduite à un singleton.

▷ Soit (E, d) un espace métrique complet, et (F_n) une suite décroissante pour l'inclusion de fermés non vides telle que $\lim_{n\to+\infty} \delta(F_n) = 0$ où $\delta(F_n) = \sup_{(x,y\in F_n^2} d(x, y)$. Montrons qu'il existe $x \in E$ tel que $\{x\} = \bigcap\limits_{n\in\mathbb{N}} F_n$. Pour tout $n \in \mathbb{N}$, F_n est non vide donc il contient un élément x_n d'après l'axiome du choix dénombrable. Soit $\varepsilon > 0$. Les diamètres tendant vers zéro, on peut choisir $N \in \mathbb{N}$ tel que $\delta(F_N) < \varepsilon$, et alors pour tous $p, q > N$, $d(x_p, x_q) < \varepsilon$, car par décroissance de (F_n), $x_p, x_q \in F_N$. La suite des (x_n) est donc une suite de Cauchy. E étant complet, elle converge vers un élément x de E. Or pour tout $p \in \mathbb{N}$, F_p est fermé et $x_n \in F_p$ pour tout $n \geqslant \mathbb{N}$, donc x appartient à F_p. On en déduit que $x \in \bigcap\limits_{p\in\mathbb{N}} F_p$. Ainsi $\bigcap\limits_{n\in\mathbb{N}} F_n$ est non vide. Supposons enfin que $x, y \in \bigcap\limits_{n\in\mathbb{N}} F_n$. Alors pour tout $n \in \mathbb{N}$, $x, y \in F_n$ donc $0 \leqslant d(x, y) \leqslant \delta(F_n)$. En passant à la limite, $d(x, y) = 0$ d'où par séparation $x = y$. On en déduit que $\bigcap\limits_{n\in\mathbb{N}} F_n$ est le singleton $\{x\}$. ■

Théorème. *(Théorème de Baire)*
Dans un espace complet, toute intersection dénombrable d'ouverts denses est dense.

$\triangleright$ Soit (E, d) un espace métrique complet, et $(O_n)_{n\in\mathbb{N}}$ une suite d'ouverts denses dans E. Montrons que $\bigcap\limits_{n\in\mathbb{N}} O_n$ est dense dans E, c'est-à-dire que pour tout ouvert non vide V de E, $V \cap \bigcap\limits_{n\in\mathbb{N}} O_n \neq \emptyset$. Soit donc V un ouvert de E. On construit par récurrence une suite (B_n) de boules fermées de E telles que pour tout $n \in \mathbb{N}$, B_n est de rayon non nul inférieur à $1/2^n$, et d'autre part, $B_0 \subseteq O_0 \cap V$ et $B_{n+1} \subseteq O_{n+1} \cap \mathring{B}_n$. On notera que ces boules sont emboîtées.

L'ouvert O_0 est dense dans E donc $O_0 \cap V \neq \emptyset$. Or cet ensemble est ouvert par intersection de deux ouverts, donc il existe une boule ouverte $B(x_0, r) \subseteq O_0 \cap V$. Si B_0 est la boule fermé de centre x_0 et de rayon $r/2$ (ou 1 si $r/2 > 1$), on a donc $B_0 \subseteq O_0 \cap V$. Supposons les boules $B_0, ..., B_n$ construites et vérifiant les propriétés voulues. L'ouvert O_{n+1} étant dense dans E, $O_{n+1} \cap \mathring{B}_n$ est un ouvert non vide. Il existe donc une boule ouverte $B(x, r)$ incluse dans $O_{n+1} \cap \mathring{B}_n$, et si l'on prend pour B_{n+1} la boule fermé de centre x et de rayon $\min(r/2, 1/2^{n+1})$, on a $B_{n+1} \subseteq O_{n+1} \cap \mathring{B}_n$, et B_{n+1} vérifiant bien les propriétés voulues.

Par construction, $(B_n)_{n\in\mathbb{N}}$ est une suite décroissante de fermés non vides de E dont le diamètre tend vers 0. De plus, E est complet, donc d'après le théorème des fermés emboîtés, il existe $x \in E$ tel que $\bigcap\limits_{n\in\mathbb{N}} B_n = \{x\}$. Comme $B_0 \subseteq V$, on a en particulier $x \in V$. D'autre part, par construction, $B_n \subseteq O_n$ pour tout n, donc $x \in O_n$ pour tout $n \in \mathbb{N}$. Par conséquent, $x \in \bigcap\limits_{n\in\mathbb{N}}$ donc $x \in V \cap \bigcap\limits_{n\in\mathbb{N}}$, ce qu'il fallait montrer. ■

Théorème. *(Théorème de Banach-Schauder, théorème de l'application ouverte)*

Toute application linéaire continue surjective entre deux espaces de Banach est ouverte (c'est-à-dire, l'image de tout ouvert est ouvert).

$\triangleright$ Soit $f : E \longrightarrow F$ une application linéaire continue et surjective. Puisque f est surjective, f s'écrit comme la réunion des fermés $F_n = \overline{f(B(0, n))}$, boule de l'espace E. D'après le théorème de Baire, l'un de ces fermés, notons-le F_N, est d'intérieur non vide ; il contient donc une boule $B(y, \eta)$ de l'espace F. Le fermé F_{2N} contient donc la boule $B(0, \eta)$. On dispose ainsi, par linéarité, d'un réel M tel que $B_F(0, 1) \subseteq \overline{B_E(0, M)}$. On en déduit que, pour tout $n \in \mathbb{N}$, $B_F(0, \frac{1}{2^n}) \subseteq \overline{f(B_E(0, \frac{M}{2^n})}$. Montrons que $B_F(0, 1) \subseteq B_E(0, 2M)$. Soit $z \in B_F(0, 1)$. D'une part, il existe x_0 de norme strictement inférieure à M tel que $z_1 = z - f(x_0)$ soit de norme inférieure à $1/2$, et d'autre part, il existe x_1 de norme inférieure à $M/2$ tel que $z_2 = z_1 - f(x_1)$ soit de norme inférieure à $1/4$. On construit alors, par récurrence, une suite (x_n) de vecteurs de E telle que $\|x_n\| \leqslant M/2^n$ et $z_{n+1} = z - f(x_0 + ... + x_n)$ soit de norme inférieure à $1/2^{n+1}$. La série $\sum x_n$ est en particulier absolument convergente, donc comme E est un espace de Banach, elle est convergente. De plus, $\|\sum\limits_{n=0}^{+\infty} x_n\| \leqslant \sum\limits_{n=0}^{+\infty} \|x_n\| < M \sum\limits_{n=0}^{+\infty} \frac{1}{2^n} = 2M$. Par passage à la limite, $z = f(\sum\limits_{n=0}^{+\infty} x_n) \in f(B_E(0, 2M))$, ce qu'il fallait démontrer. ■

On en déduit directement que les morphismes bijectifs, dans la catégorie des espaces de Banach, sont exactement leurs isomorphismes. En effet, si l'on a une application linéaire continue bijective, la surjectivité donne en particularité qu'elle est ouverte, d'après le théorème précédente. Par caractérisation de la continuité par images réciproques d'ouverts, sa bijection réciproque est continue, c'est-à-dire que la bijection linéaire est nécessairement bicontinue. C'est donc en particulier un homéomorphisme, c'est-à-dire un isomorphisme topologique.

On termine par un résultat que l'on peut trouver dans la littérature. Il existe un théorème également, ressemblant au théorème de Cantor-Bernstein, dans les espaces de Banach, avec des hypothèses supplémentaires. Pour l'énoncer, on aura besoin de la notion suivante.

Définition. *(Sous-espace complémenté)*

Un sous-espace de Banach d'un espace de Banach est dit *complémenté* s'il admet un supplémentaire qui soit un sous-espace de Banach, c'est-à-dire un supplémentaire fermé.

Théorème. *(Théorème de Pelczynski)*

Soient X, Y deux espaces de Banach. On suppose X est isomorphe à un sous-espace complémenté de Y, que Y est isomorphe à un sous-espace complémenté de X, et que X et Y sont isomorphes à leur carré. Alors X et Y sont isomorphes.

$\triangleright$ X est isomorphe à un sous-espace complémenté de Y, c'est-à-dire[2] qu'il existe un espace de Banach R tel que $X \times R \sim Y$, où $X \times R$ est muni de la topologie produit. Ainsi $X \times Y \sim X \times (X \times R) \sim (X \times X) \times R \sim X \times R \sim Y$, car $X \sim X \times X$ par hypothèse. Puisque X et Y jouent des rôles symétriques, on a aussi $Y \times X \sim X$. D'autre part, $X \times Y$ et $Y \times X$ sont isomorphes par l'application $(x, y) \longrightarrow (y, x)$. Par transitivité, on a que $X \sim Y$, ce qu'il fallait démontrer. ■

3 Problème en algèbre générale

Définissons avant tout quelques notions de base.

Définition. *(Magma)*

Un *magma* $(E, \star)$ est la donnée d'un ensemble E et d'une loi de composition interne sur E, c'est à dire une application $\star : E \times E \longrightarrow E$.

Définition. *(Magma associatif)*

Un magma $(E, \star)$ est dit *associatif* si pour tous $x, y, z \in E$, $x \star (y \star z) = (x \star y) \star z$.

Définition. *(Morphisme de magmas)*

Un *morphisme de magmas* f de $(E, \star)$ dans $(F, \diamond)$ est une application de E dans F telle que pour tous $x, y \in E$, $f(x \star y) = f(x) \diamond f(y)$.

Définition. *(Monoïde)*

Un *monoïde* est un magma associatif contenant un élément neutre à la fois à droite et à gauche.

Définition. *(Morphisme de monoïdes)*

Un *morphisme de monoïdes* est un morphisme entre les magmas correspondant qui envoie le neutre d'un magma sur l'autre.

Remarque. Il existe des morphismes de magmas entre monoïdes qui ne sont pas des morphismes de monoïdes. Par exemple, en considérant $E = F = \mathbb{N}$ muni de la loi $n \star m = \max(n, m)$, l'application $n \longrightarrow n + 1$ est un morphisme de magma (d'un magma dans lui-même, ici un monoïde) qui n'est pas un morphisme de monoïdes.

[2] En effet, la somme directe externe d'espaces vectoriels sur un même corps est défini comme la sous-structure $\bigoplus_{i \in I} E_i$ des familles à supports finis du produit direct $\Pi_{i \in I}$, ce qui, lorsque I est fini, légitime l'identification $E \oplus F \cong E \times F$.

3.1 Le théorème de Cantor-Bernstein pour les groupes

3.1.1 Aparté : le cas fini

Plus généralement, pour toute catégorie dont les objets sont tous finis et tout morphisme bijectif est un isomorphisme, c'est-à-dire que sa bijection réciproque est également un morphisme, propriété que nous notons $(*)$, le théorème de Cantor-Bernstein est trivialement vrai. On peut donc énoncer :

Théorème. *($\mathcal{C}$-théorème de Cantor-Bernstein pour $\mathcal{C}$ à objets finis dans les catégories assez rigides)*
Si $\mathcal{C}$ est une catégorie concrète dont les objets sont tous finis, et $(*)$ y est vérifiée, alors $\mathcal{C}$-CB.

Ceci est par exemple vrai pour les groupes finis. En effet, la catégorie des groupes finis... est à objets finis, et l'on sait d'après que le cours, que la réciproque de tout isomorphisme de groupes et encore un isomorphisme. Rappelons pourquoi.

Théorème. La réciproque de tout isomorphisme de groupes $f : G \longrightarrow G'$ est un isomorphisme de groupes.

$\triangleright$ Soit $f : G \longrightarrow G'$ un isomorphisme de groupes. Sa bijection réciproque est bien sûr une bijection, et il reste à montrer que c'est un morphisme. Soient $u, v \in G'$. Il existe, par surjectivité de f, deux éléments $x, y \in G$ tels que $f(x) = u$ et $f(y) = v$. Alors $f^{-1}(uv) = f^{-1}(f(x)f(y)) = f^{-1}(f(xy)) = xy$. Or $x = f^{-1}(u)$ et $y = f^{-1}(v)$, c'est-à-dire que l'on peut récrire $f^{-1}(uv) = f^{-1}(u)f^{-1}(v)$. Ainsi f^{-1} est un morphisme, ce qui termine la preuve. ■

3.1.2 Le cas général des structures « magma-like »

Nous nous intéressons directement au cas des groupes. Si nous l'infirmons, alors nous pourrons l'infirmer pour les catégories contenues dans celles des groupes dont les hypothèses pour définir le morphisme ne sont pas plus fortes que celles de groupe, ce qui est évidemment le cas pour les magmas. Nous l'infirmons en effet dans le cas des groupes quelconques.

Heuristique. Nous cherchons deux groupes G, G' (leurs lois ne sont pas forcément les mêmes), tels que $G \hookrightarrow G'$ par un morphisme de groupes (*i. e.* une application f telle que pour tous $x, y \in G$, $f(xy) = f(x)f(y)$), également $G' \hookrightarrow G$ par un morphisme de groupes, et tels que $G \not\simeq G'$ au sens de l'isomorphie entre groupes ; d'après le paragraphe précédent, les ensembles G, G' sont infinis et d'après le théorème de Cantor-Bernstein pour les ensembles, $\mathrm{card}(G) = \mathrm{card}(G')$ ce qui oriente la recherche aux groupes infinis de même cardinaux. Intuitivement, rien ne sert de chercher trop loin et l'on peut commencer avec les groupes dénombrables. Les groupes additifs $\mathbb{Z}^n$ sont pourtant une suite strictement croissante pour l'injection par morphismes[3], ce qui force à complexifier nos attentes et à nous intéressons à des groupes de cardinaux légèrement plus grands. Le vrai vice de structure apparaît lorsqu'on mêle les groupes additifs $\mathbb{Z}$ et celui des rationnels $\mathbb{Q}$, et surtout lorsqu'on étudie un produit infini de groupes de rationnels, c'est-à-dire un groupe de suites de rationnels. On voit ci-dessous que la contradiction est, comme souvent dans de telles études, de nature arithmétique.

Théorème. *(Infirmation de Grp-CB)*
La catégorie des groupes est non bernsteinienne.

[3] Par composition, il suffit de montrer que $\mathbb{Z}^n$ s'injecte dans $\mathbb{Z}^{n+1}$ par un morphisme mais pas le contraire. Le sens direct est évident. Pour montrer qu'il n'existe pas de morphisme injectif de $\mathbb{Z}^{n+1}$ dans $\mathbb{Z}^n$, il s'agit d'utiliser la monogénéité de $\mathbb{Z}$, ici sur chaque composante. La contradiction vient en effet du point de vue algébrique, car tous ces groupes ont le cardinal de $\mathbb{Z}$ par produit fini de dénombrables.

▷ On introduit le contre exemple : $G = \mathbb{Q}^{\mathbb{N}}$ l'ensemble des suites de rationnels muni de la loi $+$; cet ensemble en est naturellement muni de la structure de groupe, $(\mathbb{Q}, +)$ étant un groupe. On pose également $G' = \mathbb{Z} \times \mathbb{Q}^{\mathbb{N}^*}$. Cet ensemble est un groupe en tant que groupe produit de $(\mathbb{Z}, +)$ avec $\mathbb{Q}^{\mathbb{N}^*}$. Ces groupes s'injectent l'un dans l'autre par des morphismes : en effet, il existe un morphisme injectif de G dans G' avec $f \longmapsto (0, \tilde{f})$ où $\tilde{f}$ est décalée 1 ($\tilde{f}(i) = f(i-1)$), l'injectivité se déduisant directement de la propriété fondamentale des couples, et il existe un morphisme injectif de G' dans G, en posant l'application qui à (k, f) associe la suite g telle que $g(0) = k$ et $g(n) = f(n)$ sinon ; elle est injective par ce que g détermine complètement k et f. Montrons maintenant que G et G' ne sont pas isomorphes. Supposons le contraire : il existe $\phi : G \longrightarrow G'$ un isomorphisme. Alors il existe un antécédent $f \in \mathbb{Q}^{\mathbb{N}}$ de $\phi(f) = (2, 0)$ où 0 désigne la suite nulle de rationnels définie à partir de 1. f s'écrit $f = (u_1, ..., u_i, ...)$ en tant que suite ; on peut définir le morphisme coordonnée par rapport à chaque composante. Alors $\phi(3\frac{u_1}{3}, ..., 3\frac{u_i}{3}, ...) = 3\phi(\frac{u_1}{3}, ..., \frac{u_i}{3}, ...) = (2, 0)$. Considérons p le morphisme renvoyant la première coordonnée, à valeurs dans $\mathbb{Z}$ puisque ϕ est à valeurs dans G'. Dans ce cas, on peut écrire $2 = 3(p \circ f)(\frac{u_1}{3}, ..., \frac{u_i}{3}, ...)$. $(p \circ f)(\frac{u_1}{3}, ..., \frac{u_i}{3}, ...)$ étant un entier, on a $3 \mid 2$, ce qui est absurde. ■

Ce contre-exemple, pas tout à fait trivial, est très puissant en ce qu'il permet d'infirmer la majorité des théorèmes de Cantor-Bernstein pour les catégories de l'algèbre abstraite ayant une seule loi, ce que nous appelons, comme le vocabulaire anglais le suggère, des structures « magma-like ». Remarquons par exemple que les groupes donnés en exemple juste précédemment sont abéliens, ce qui permet de ne pas disjoindre le cas commutatif. En particulier, il n'y aura pas de théorème de Cantor-Bernstein du type : « si G est isomorphe à un sous-groupe distingué de G', et réciproquement, alors... »comme on l'a vu dans les espaces de Banach. Par suite, successivement :

- le théorème de Cantor-Bernstein pour les groupes quelconques, pour les groupes abéliens, est faux (on vient de le voir) ;
- le théorème de Cantor-Bernstein pour les magmas quelconques, qu'ils soient tous commutatifs ou non, est également faux, car tout groupe est magma et tout morphisme de groupes est morphisme de magmas pour les mêmes lois ;
- de même pour les magmas associatifs, car l'associativité de la loi apparaît dans les axiomes du groupe ;
- le théorème de Cantor-Bernstein pour les monoïdes est également faux pour les mêmes raisons. En effet, un morphisme de groupes est toujours un morphisme de monoïdes, car on sait que $f(e_G) = e_{G'}$ avec les notations usuelles. C'est très simple à redémontrer : $f(e_G \times e_G) = f(e_G) = f(e_G) \times f(e_G)$. Par simplifiabilité de $f(e_G)$ dans G', on a $e_{G'} = f(e_G)$, ce qu'il fallait démontrer.

Intuitivement, l'absence d'un théorème de Cantor-Bernstein pour, disons, les groupes, comme nous l'avons pris pour exemple dans l'introduction de notre étude, donne pour image que la catégorie des groupes quelconques n'est pas assez rigide dans l'acception floue que nous nous sommes permise d'employer jusqu'ici. À l'instar des espaces topologiques, la structure algébrique de groupe est trop faible pour permettre de caractériser l'isomorphie au moyen d'une relation d'ordre, l'existence d'un morphisme injectif entre deux groupes. *A fortiori*, c'est encore le cas pour les structures encore moins élaborées que sont les magmas et les monoïdes. Dans le même mouvement que dans la section précédente, nous pouvons nous demander si, dans le cas de structure plus élaborées telles que les anneaux ou les corps, le théorème de Cantor-Bernstein peut subsister.

3.2 Cantor-Bernstein et les anneaux et les corps

3.2.1 Préliminaires

Quelques définitions nous donnent un peu plus de liberté.

Définition. *(Pseudo-anneau)*
On appelle *pseudo-anneau*, un anneau sans l'axiome d'existence d'élément unité, ce que la littérature appelle parfois *anneau* alors que nous réservons ce terme à ce que ces mêmes auteurs appellent *anneaux unitaires.*

Exemple. Avec cette définition, $(2\mathbb{N}, +, \times)$ est un pseudo-anneau.

Définition. *(Morphisme d'anneaux, morphisme de pseudo-anneaux)*
Un *morphisme d'anneaux* est une application qui soit un morphisme de magmas additifs et un morphisme de monoïdes multiplicatifs. Pour ce qui est des morphismes de pseudo-anneaux, $(A, \times)$ n'est justement pas un monoïde, d'où l'on définit un *morphisme de pseudo-anneaux* comme un morphisme de magmas additifs et un morphisme de magmas multiplicatifs.

De la même manière que dans le paragraphe précédent, nous avons intérêt à nous intéresser sans détour au cas des corps (commutatifs). De plus, nous aurons grandement besoin des notions définies ci-dessous.

Définition. *(Clôture algébrique)*
Un corps K est dit *algébriquement clos* si tout polynôme non constant à coefficients dans K admet au moins une racine qui soit un élément de K. (Il revient au même de dire que tout polynôme non constant a autant de racines que son degré, comptées avec leur multiplicité, ou que tout polynôme non constant est scindé.)

Petite propriété. Si deux corps K_1 et K_2 sont isomorphes, et si l'un est algébriquement clos, alors l'autre est aussi algébriquement clos.

$\triangleright$ Ce n'est pas tout à fait trivial, nous nous donnons la peine de l'écrire. Supposons, par symétrie, que K_1 est algébriquement clos. Soit $P = a_0 + a_1X + ... + a_nX^n$ un polynôme à coefficients dans K_2. Par surjectivité de l'isomorphisme de corps $\varphi : K_1 \longrightarrow K_2$ que nous pouvons invoquer, il existe des éléments de K_2, notés $b_0, ..., b_1$ dont les images par φ sont respectivement les $a_1, ..., a_n$, de sorte que l'on puisse écrire $P = \varphi(a_0) + \varphi(a_1)X + ... + \varphi(a_n)X^n$. D'autre part, le polynôme $b_0 + b_1X + ... + b_nX^n$ est à coefficients dans K_1. Comme ce corps est algébriquement clos, il admet une racine α, soit $b_0 + b_1\alpha + ... + b_n\alpha^n = 0_{K_1}$. Composons par φ : $\varphi(b_0 + b_1\alpha + ... + b_n\alpha^n) = \varphi(0_{K_1}) = 0_{K_2}$. Or φ est un morphisme de corps, ce qui permet, en utilisant les trois axiomes de morphisme d'anneaux, d'écrire exactement $\varphi(b_0 + b_1\alpha + ... + b_n\alpha^n) = \varphi(b_0) + \varphi(b_1)\varphi(\alpha) + ... + \varphi(b_n)\varphi(\alpha)^n = a_0 + a_1\varphi(\alpha) + ... + a_n\varphi(\alpha)^n = P(\varphi(\alpha))$, soit $P(\varphi(\alpha)) = 0_{K_2}$. En notant $\beta = \varphi(\alpha) \in K_2$, $P(\beta) = 0$, donc P a une racine dans K_2. ■

Il nous faudra surtout utiliser le théorème suivant, dont la démonstration, remarquons-le, nécessite (encore) l'axiome du choix au moyen du lemme de Zorn démontré en début de composition.

Théorème. *(Théorème de Steinitz)*
Tout corps commutatif admet une clôture algébrique, c'est-à-dire une extension de corps (aussi appelée parfois *sur-corps*) algébriquement close.

$\triangleright$ Soit K un corps (commutatif, non trivial). On choisit un ensemble Ω de cardinal $\max(2^{\aleph_0}, 2^{\mathrm{card}(K)})$. C'est possible ; il suffit de considérer l'ensemble des réels ou bien l'ensemble des parties de K. On considère l'ensemble des triplets $(L, +, \times)$ avec L un sous-ensemble de Ω contenant K, et des lois prolongeant celles de K faisant de L une extension algébrique, c'est-à-dire telle que tous les éléments de L sont algébriques sur K, c'est-à-dire racines d'un polynôme à coefficients dans K. On définit une relation d'ordre sur l'ensemble de ces triplets par $(L,+,\times) \leqslant (L',+,\times)$ si $L \subseteq L'$ et si L est un sous-corps de L' (on entend, pour ces lois). On vérifie aisément que l'ensemble des triplets considéré est alors un ensemble inductif. Il est non vide, car il contient K. Par le lemme de Zorn, on en déduit qu'il admet un élément maximal, que nous notons F. Montrons que F est une clôture algébrique de K.

Soit E une extension algébrique de F. F étant algébrique sur K par construction, il est de même cardinal que K ou, lorsque K est fini, est au plus dénombrable. Justifions le : $F = \bigcup_{n\in\mathbb{N}} \bigcup_{P\in K_n[X], P\neq 0} \mathrm{Zeros}(P)$, d'après la définition d'algébricité, où l'on note $\mathrm{Zeros}(P)$ l'ensemble des racines de P. Or cet ensemble est fini puisque P est non nul, $K_n[X]$ est un K-espace de dimension $n+1$, donc isomorphe à K^{n+1}. Si K est infini, cet ensemble est équipotent à K, car tout ensemble infini est équipotent à son carré cartésien, et comme il est encore infini, en passant à une réunion dénombrable, F est de cardinal K. Si K est fini, cet ensemble est encore fini, puis en passant à la réunion dénombrable, il est au plus dénombrable. E étant une extension algébrique de F, il est, de même cardinal que F, donc de même cardinal que K. Par suite, le complémentaire de F dans E est de cardinal inférieur à celui de $\complement_\Omega F$, qui a le même cardinal que Ω. Ainsi, il existe une application injective de E dans Ω qui soit l'identité sur F. Si l'on munit son image de la structure de corps induite par celle de E, on obtient une extension algébrique de F. Par maximalité de F, cette image est égale à F. Donc $E = F$, et F est algébriquement clos, ce qui termine la preuve. ■

Remarque. Cette clôture algébrique est unique à isomorphisme près. En effet, si F_1, F_2 sont deux clôtures algébriques de K, on considère les couples (L,ρ) où L est une sous-K-extension de F_1 et où $\rho : L \longrightarrow F_2$ est un morphisme de corps. L'ensemble de ces couples est non vide et est naturellement ordonné. On vérifie également qu'il est inductif. Soit (L,ρ) un élément maximal donné par le lemme de Zorn. Si a est un élément de F_1, on considère son polynôme minimal $P(x) \in L[x]$ sur L, qui existe. Le polynôme $\rho(P(x)) \in \rho(L)[x]$ admet une racine b dans F_2. Il existe alors un morphisme de corps $L[a] \longrightarrow F_2$ qui vaut ρ sur L, envoyant a sur b. Par maximalité de (L,ρ), on a $L[a] = L$, donc $L = F_1$. Comme $\rho(F_1) \subseteq F_2$ est algébriquement clos, on a $\rho(F_1) = F_2$, puis ρ est un isomorphisme de F_1 sur F_2.

Enfin, nous aurons besoin de la notion de base de transcendance.

Définition. *(Base de transcendance)*

Pour tout extension de corps L de K, une famille d'élements de L est une *base de transcendance* si elle est algébriquement indépendante sur K et si elle n'est strictement contenue dans aucune famille algébriquement indépendante de L. Une famille $\mathcal{F}$ d'éléments de L est dite *algébriquement indépendante* sur K si pour tous $x_1, ..., x_n \in L$, il n'existe pas de polynôme non nul $P \in K[X_1, ..., X_n]$ tel que $P(x_1, ..., x_n) = 0$.

Théorème. *(Existence de bases de transcendance)*

Pour tout corps K, pour tout extension transcendante L de K (c'est-à-dire, qui n'est pas algébrique), il existe des bases de transcendance de L sur K.

▷ Il s'agit encore une fois d'appliquer le lemme de Zorn. On considère l'ensemble $\mathcal{F}$ des familles d'éléments de L algébriquement indépendantes sur K. Cet ensemble est inductif. En effet, il est non vide par définition d'une extension transcendante : il existe un élément α de L non algébrique sur K, et sa famille $\{\alpha\}$ est algébriquement indépendante sur K. Soit maintenant C une chaîne de $\mathcal{F}$, c'est-à-dire un ensemble de familles algébriquement indépendantes sur K totalement ordonnées pour l'inclusion. Leur réunion est encore algébriquement indépendante sur K, et c'en est trivialement un majorant. Montrons-le. Cette chaîne, on peut l'écrire $C = \{\mathcal{A}_i \mid i \in I\}$ où I est totalement ordonné, avec donc pour tous $i,j \in I$, $i \leqslant j \Longrightarrow \mathcal{A}_i \subseteq \mathcal{A}_j$. Notons $\mathcal{A} = \bigcup_{i\in I} \mathcal{A}_i$. Pour tous $x_1, ..., x_n \in \mathcal{A}$, pour tout $j \in [\![1,n]\!]$, il existe $i_j \in I$ tel que $x_j \in \mathcal{A}_{i_j}$. Soit $i_0 = \max_{1\leqslant j\leqslant n} i_j$. Puisque la famille $(\mathcal{A}_i)_{i\in I}$ est totalement ordonnée, on a $x_1, ..., x_n \in \mathcal{A}_{i_0}$. La famille $\mathcal{A}_{i_0}$ étant algébriquement indépendante, on en déduit qu'il n'existe pas de polynôme non nul

$P \in K[X_1, ..., X_n]$ tel que $P(x_1, ..., x_n) = 0$. Ainsi $\mathcal{A}$ est algébriquement indépendante.

Ainsi, il existe un élément maximal de $\mathcal{F}$. Par définition, c'est une base de transcendance de L sur K. ■

Théorème. *(Caractérisation des bases de transcendance)*

Pour tout corps K, pour tout extension transcendante L de K, $(\xi_i)_{i\in I}$ est une base de transcendance de L sur K si et seulement si les ξ_i sont algébriquement indépendants et L est algébrique sur $K(\xi_i)_{i\in I}$.

▷ C'est immédiat : si l'extension n'était pas algébrique, il y aurait un $\eta \in L$ transcendant sur $K(\xi_i)_{i\in I}$; mais alors la famille obtenue en rajoutant η aux $(\xi_i)_{i\in I}$ est encore algébriquement indépendante, contredisant la maximalité de la famille $(\xi_i)_{i\in I}$. ■

3.2.2 Conséquences et conclusions

Nous pouvons maintenant fournir un contre-exemple, quoique simple, au théorème de Cantor-Bernstein pour les corps.

Rappelons pourquoi la recherche d'idées pour un tel contre-exemple est beaucoup facilitée dans ce cas.

Propriété. Tout morphisme de corps est injectif.

▷ Soit $f : K \longrightarrow K'$ un morphisme de corps. C'est en particulier un morphisme d'anneaux, donc son noyau $\mathrm{Ker}(f)$ est un idéal de K. Or les idéaux d'un corps sont triviaux : c'est soit $\{0\}$, soit K lui-même. Or cette dernière possibilité est exclue, car $f(1_K) = 1_{K'} \neq 0_{K'}$ un corps n'étant jamais trivial, donc $\mathrm{Ker}(f) = \{0_K\}$, ce qui caractérise l'injectivité des morphismes. ■

Nous pouvons conclure.

Théorème. La catégorie des corps est non bernsteinienne.

▷ On prend deux corps $K_1 = \mathbb{C}$ le corps des complexes, et $K_2 = \mathbb{C}(X)$ le corps des fractions rationnelles sur $\mathbb{C}$. On va exhiber deux morphismes injectifs en sens inverses et justifier que ces corps ne sont pas isomorphes au sens des morphismes de corps, qui sont exactement les morphismes d'anneaux entre deux corps.

$\mathbb{C}$ s'injecte trivialement dans $\mathbb{C}(X)$ par l'injection canonique $z \longrightarrow zX^0$. Le plus dur est d'exhiber un morphisme de corps en sens inverse. Soit $(x_i)_{i\in I}$ une base de transcendance de $\mathbb{C}$ sur $\mathbb{Q}$, qui existe d'après le théorème précédent, car bien sûr[4] $\mathbb{C}$ est une extension transcendante de $\mathbb{Q}$. I est infini, car $\mathbb{C}$ n'est pas dénombrable (il est plus grand que $\mathbb{R}$). On note $\overline{K}$ la clôture algébrique de $K = \mathbb{Q}(x_i)_{i\in I}$. Soit $L = K(X)$ et $\overline{L}$ la clôture algébrique de L. Alors $\overline{K}$ et $\overline{L}$ sont isomorphes, puisque K et L le sont, et sont tous deux isomorphes à $\mathbb{C}$ par définition d'une base de transcendance et le fait que $\mathbb{C}$ est algébriquement clos. En effet, K est isomorphe au corps $\mathbb{Q}(X_i)_{i\in I}$ les X_i étant des indéterminées. $L = K(X)$ donc L est isomorphe à $\mathbb{Q}(X_j)_{j\in J}$ où $J = I \cup i_0$ où i_0 est en dehors de I. Comme I et J ont en fait le même

[4] Il suffit d'exhiber un nombre transcendant. Le plus simple reste de démontrer que l'ensemble des réels transcendants est indénombrable, ce qui constitue le *théorème de Liouville*. Il suffit alors de montrer que son complémentaire l'ensemble des réels algébriques est dénombrable, car $\mathbb{R}$ est indénombrable. Or $\mathbb{Q}$ étant dénombrable, $\mathbb{Q}[X]$ est dénombrable : en effet, c'est la réunion dénombrable des $\mathbb{Q}_n[X]$, qui par dimension, sont isomorphes à $\mathbb{Q}^{n+1}$ qui est dénombrable par produit cartésien fini de dénombrables. Dans ce cas, l'ensemble des nombres algébriques s'écrit $\bigcup_{P\in\mathbb{Q}[X], P\neq 0} \mathrm{Zeros}(P)$ où l'ensemble des zéros d'un polynôme non nul $\mathrm{Zeros}(P)$ est fini. Par réunion dénombrable de finis, l'ensemble des nombres algébriques est au plus dénombrable. Ainsi, l'ensemble des réels transcendants est indénombrable, en particulier, il est non vide.

cardinal, K et L sont isomorphes. D'autre part, par définition de K, tout élément de $\mathbb{C}$ est algébrique sur K, et comme $\mathbb{C}$ est algébriquement clos, on a $\overline{K} \simeq \mathbb{C}$.

Le corps $\overline{K}$ s'identifie à l'ensemble des éléments de $\overline{L}$ qui sont algébriques sur K. En effet : soit K_1 l'ensemble des éléments de $\overline{L}$ qui sont algébriques sur K. Par définition K_1 est une extension algébrique de K. Soit P un polynôme à coefficients dans K_1. Il a une racine x dans $\overline{L}$, car $\overline{L}$ est algébriquement clos, et x est algébrique sur K_1, donc sur K, par conséquent $x \in K_1$. Ceci montre que K_1 est algébriquement clos, donc isomorphe à $\overline{K}$. De plus l'élément $X \in L$ est transcendant sur $\overline{K}$ puisqu'il l'est sur K, donc on a un morphisme de corps $\mathbb{C}(X) \simeq \overline{K}(X) \to \overline{L} \simeq \mathbb{C}$. Ce morphisme de corps est automatiquement injectif comme nous l'avons rappelé ci-dessous.

Ainsi $K_1 \hookrightarrow K_2$ et $K_1 \hookrightarrow K_1$. Cependant, ces deux corps ne sont pas isomorphes. On utilise le lemme justifié précédemment : $\mathbb{C}$ est algébriquement clos d'après le théorème de d'Alembert-Gauss. Par contre, $\mathbb{C}(X)$ n'est pas algébriquement clos : le polynôme X n'a pas de racine carrée ; en effet, s'il existait une fraction rationnelle $(\frac{A}{B})^2 = X$, alors $A^2 = XB^2$ où A, B sont des polynômes, on aurait $2\deg(A) = 2\deg(B) + 1$. Or B est non nul, et A est non nul puisque X est non nul, dont ces degrés sont des grandeurs entières, l'une paire, l'autre impaire, d'où la contradiction. Ainsi le polynôme de $\mathbb{C}(Y)$ défini par $Y^2 - X$ n'est pas scindé. Ainsi K_1 et K_2 ne sont pas isomorphes, ce qui fournit un contre-exemple. ■

De la même manière que pour les structures topologiques, on en déduit une négation des théorèmes de Cantor-Bernstein sur les sous-catégories pleines usuelles de la catégorie des corps, comme explicité ci-dessous :

Théorème. La catégorie des pseudo-anneaux est non bernsteinienne.

Théorème. La catégorie des anneaux est non bernsteinienne.

Théorème. La catégorie ACU est non bernsteinienne.

Théorème. La catégorie des anneaux intègres est non bernsteinienne.

Théorème. La catégorie des corps gauches est non bernsteinienne.

▷ Dans tous les cas, l'exemple précédent convient. ■

Remarquons que ce contre-exemple du théorème de Cantor-Bernstein pour la catégorie des corps en fournit automatiquement un pour la catégorie des groupes, mais nous voyons aussi que la construction de celui-là est nettement plus laborieuse que celui donné pour les groupes dont on avait justifié même qu'il était le plus élémentaire possible.

Pour résumer, le théorème de Cantor-Bernstein n'est pratiquement jamais vérifié en algèbre générale, c'est-à-dire, pour les catégories concrètes pleines de l'algèbre générale (groupes, anneaux, corps), le théorème est systématiquement infirmé. En particulier, dans le cas des corps, il est surprenant d'invoquer la notion de base de transcendance lorsque c'est justement, dans le cas de l'algèbre linéaire traité plus haut, la notion de base algébrique qui a permis de généraliser le théorème de Cantor-Bernstein, et les bases de transcendance vérifient des propriétés analogues : par exemple, elles ont toutes le même cardinal. Nous devons nous apercevoir que la notion de base de transcendance n'a pas pourtant l'universalité de celle de base algébrique, et le problème vient directement de ce que l'on parle toujours d'une base de transcendance de L sur K, ce qui rend ce concept dépend d'un couple de corps, dont par surcroît, l'un doit être une extension transcendance de l'autre. En définitive, nous observons que les structures algébriques n'ont pas la « rigidité »attendue pour qu'un théorème de Cantor-Bernstein soit vérifié non trivialement sur leurs catégories pleines, notamment à cause de la facilité avec laquelle sont définis leurs morphismes.

4 Pistes pour un TCB universel

Des considérations des sections précédentes, nous pouvons tirer des pistes pour généraliser TCB en théorie des catégories.

Catégorie $\mathcal{C}$	$\mathcal{C}$-CB
Structures ensemblistes	
Ensembles finis	✔
Ensembles	✔
Structures algébriques	
Magmas	✘
Monoïdes	✘
Groupes	✘
Groupes abéliens	✘
Groupes finis	✔
Anneaux (unitaires)	✘
Corps	✘
Structures topologiques	
Espaces vectoriels de dimension finie	✔
Espaces vectoriels	✔
Espaces vectoriels normés de dimension finie	✔
Espaces de Banach	*Théorème de Pelczynski*
Espaces métriques	✘
Espaces topologiques	✘
Espaces compacts	✘

TABLEAU 3. — *Résumé de notre visite des catégories concrètes.* Le choix des isomorphismes pour une catégorie donnée, nous l'avons vu, importe grandement pour savoir si $\mathcal{C}$-CB y est vérifié.

4.1 Le cas des catégories rigides

Nous l'avons vu à plusieurs reprises dans notre composition : il ne suffit pas que les morphismes bijectifs soient automatiquement des isomorphismes pour que la catégorie considérée soit bernsteinienne, même si elle est pleine, c'est-à-dire, si elle contient bien tous les morphismes possibles (cf. le cas des groupes). Si, en revanche, on impose que les objets sont tous finis, alors le théorème devient vrai, c'est pourquoi nous introduisons la terminologie suivante.

Définition. *(Catégorie très petite)*
Une catégorie $\mathcal{C}$ sera dite *très petite* si tous les objets de $\mathcal{C}$ sont des ensembles finis.

Définition. *(Catégorie rigide)*
Une catégorie concrète est dite *rigide* si tout morphisme bijectif est un isomorphisme.

Dans ce cas, nous pouvons énoncer le

Théorème. Pour toute catégorie très petite et rigide $\mathcal{C}$, $\mathcal{C}$-CB est trivialement vérifié.

▷ Soit $\mathcal{C}$ une catégorie très petite rigide. Soient A, B deux objets de cette catégorie. On suppose qu'il existe un morphisme injectif de A dans B, noté f, et un morphisme injectif de B dans A. D'après le théorème de Cantor-Bernstein pour les ensembles finis, A et B sont de même cardinal (fini). L'application f est donc une injection entre ensembles finis de même cardinal, donc c'est une bijection. On dispose donc d'un morphisme bijectif de A dans B. Par rigidité de $\mathcal{C}$, c'est un isomorphisme, soit $A \simeq B$, ce qu'il fallait montrer. ■

4.2 Les invariants d'isomorphie

Nous précisons ici les considérations apparues à la suite de l'établissement du théorème de Cantor-Bernstein pour les $\mathbb{K}$-espaces vectoriels de dimensions quelconques.

Pour formaliser cette idée naturelle, nous avons besoin de la notion d'ensembles quotients.

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 et 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 $\mathcal{R}$. ■

$$\begin{tikzcd} E \arrow[d, two heads, "\pi"'] \\ E/\mathcal{R} \end{tikzcd}$$

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.* $\forall x, y \in E \quad x \sim y \Rightarrow f(x) = f(y)$) 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 :

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, two heads, "\pi"'] & F \\ E/\mathcal{R} \arrow[ur, "\tilde{f}"'] & \end{tikzcd}$$

Le théorème de factorisation établit donc la commutation de ce diagramme. On énonce enfin un corollaire d'intérêt principalement formel, mais que nous utiliserons en priorité.

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 $\widehat{\cdot}$ et f une application de E dans F. Alors f est compatible avec $\mathcal{R}$ modulo $\mathcal{S}$ (*i. e.* $\forall x, y \in E \quad x \sim y \Rightarrow f(x) \equiv f(y)$) si et seulement s'il existe une unique application $\tilde{f}$ telle que $\chi \circ f = \tilde{f} \circ \pi$ (se qui se réécrit $\widehat{f(x)} = \tilde{f}(\overline{x})$ 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 :

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, two heads, "\pi"'] & F \arrow[d, "\chi"] \\ E/\mathcal{R} \arrow[r, "\tilde{f}"'] & F/\mathcal{S} \end{tikzcd}$$

Dans chacun des deux théorèmes précédents, on vérifie facilement que :

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

On remarque aussi que, pour passer au quotient dans une application entre deux ensembles quotients, la relation d'équivalence de l'ensemble d'arrivée n'intervient pas.

On s'aperçoit que la construction des ensembles quotients ci-dessous peut être généralisée sans problème à la théorie des classes. De même, on peut le généraliser aux foncteurs entre catégories en appliquant simplement le théorème de factorisation d'une part à l'application sur les objets, d'autre part à l'application sur les morphismes ; il s'agit ensuite de vérifier que le foncteur quotient vérifie bien les axiomes du foncteur. Nous définissons ainsi :

Définition. *(Invariant d'isomorphie)*
Soit $\mathcal{C}$ une catégorie. On appelle *invariant d'isomorphie*, tout foncteur de $\mathcal{C}$ dans Set qui passe au quotient pour la relation d'isomorphie $\simeq$ de $\mathcal{C}$ et la relation d'équipotence sur Set en un foncteur tel que l'application quotient sur les objets soit injective.

Nous pouvons énoncer :

Théorème. *(Théorème de Cantor-Bernstein sur une catégorie possédant un invariant d'isomorphie)*
Soit $\mathcal{C}$ une catégorie concrète qui admette un invariant d'isomorphie f **qui transforme les morphismes injectifs en injections**. Alors $\mathcal{C}$-CB.

$\triangleright$ Soient A, B deux objets de $\mathcal{C}$ tels que A s'injecte dans B par un morphisme u et B s'injecte dans A par un morphisme v. En appliquant le foncteur f à u et à v, on en déduit qu'il existe une injection de $f(A)$ dans $f(B)$ et une injection de $f(B)$ dans $f(A)$, car les morphismes de Set sont les applications, et on a supposé, ce qui est très raisonnable, que f transformait les morphismes injectifs en injections.

$f(A), f(B)$ étant des ensembles, on peut appliquer le théorème de Cantor-Bernstein classique qui donne que $f(A) \simeq f(B)$ pour l'équipotence. Ceci se récrit : $\tilde{f}(A) = \tilde{f}(B)$. Or $\tilde{f}$ étant injectif sur les objets, on a $A \sim B$ pour l'isomorphie de $\mathcal{C}$, ce qu'il fallait démontrer. ■

La situation des invariants d'isomorphie peut se représenter comme ci-dessous :

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "f"] \arrow[d, "\pi_1"'] & \mathrm{Set} \arrow[d, "\pi_2"] \\
\mathcal{C}/\sim \arrow[r, hook, "\tilde{f}"'] & \mathrm{Set}/\simeq
\end{tikzcd}$$

5 Conclusion

Le phénomène général est le suivant : bien que les ensembles n'aient aucune structure et que leur collection soit pourtant bernsteinienne, plus les objets d'une catégorie ont une structure « rigide », plus on observe systématiquement l'existence d'un théorème de Cantor-Bernstein au sein de leur catégorie pleine.

Cependant, ce mouvement d'ensemble est accompagné d'un mouvement contraire, dû à la rigidité de la structure des catégories mêmes (finitude des éléments, isomorphie des morphismes bijectifs, etc) : si la structure des objets peut en effet rendre plus aisée l'introduction d'un isomorphisme entre deux structures qui se ressemblent, par exemple, grâce aux bases algébriques dans $\mathbb{K}$-Vect, il faut également que l'existence de morphismes injectifs en sens réciproques soit une condition assez forte pour obtenir cette similitude. En particulier, la vérification d'un théorème de Cantor-Bernstein modulo une certaine catégorie est largement dépendante, non seulement des objets bien sûr, mais également des morphismes choisis pour la catégorie, dont le choix n'est pas nécessairement univoque.

Une formalisation complète en théorie des catégories, faisant intervenir seulement les objets de cette théorie, et ne se limitant plus uniquement aux catégories concrètes, dont on connaît les propriétés, permet peut-être d'y répondre de façon fermée.

A Bibliographie

Références

[1] *La théorie des ensembles : introduction à une théorie de l'infini et des grands cardinaux*, Patrick Dehornoy, 2017, Calvage et Mounet

[2] *Analyse mathématique : grands théorèmes du vingtième siècle*, G. H. Hardy et J. E. Littlewood, 2009, Calvage et Mounet

[3] *Topoi : the categorial analysis of logic*, Robert Goldblatt, 1979, Dover Publications

[4] *Algèbre : le grand combat. Cours et exercices*, Grégory Berhuy, 2018, Calvage et Mounet, 2e éd.

[5] *Raisonnements divins. Quelques démonstrations mathématiques particulièrement élégantes*, titre en anglais : "Proofs from the book", Martin Aigner, Günter M. Ziegler, André Warusfel, 2013, Springer, 3e éd.

[6] *Mathématiques. Tout-en-un pour la Licence*, sous la direction de Jean-Pierre Ramis, André Warusfel, 2006, Dunod, 2e éd., vol. 1

[7] *Mathématiques. Tout-en-un pour la Licence*, sous la direction de Jean-Pierre Ramis, André Warusfel, 2006, Dunod, 2e éd., vol. 2

[8] *Les maths en tête : Analyse*, Xavier Gourdon, 2021, Ellipses, 3e éd., vol. 1

[9] *Les maths en tête : Algèbre, probabilités*, Xavier Gourdon, 2021, Ellipses, 3e éd., vol. 2

[10] *Oraux x-ens mathématiques*, Serge Nicolas, Serge Francinou, Hervé Gianella, 2022, Cassini, nouvelle série, vol. 6

[11] Devoir surveillé de MPSI présentant un problème en trois parties sur le théorème de Cantor-Bernstein, Alain Troesch, polycopié du lycée Louis le Grand

[12] *Introduction à la topologie*, Francis Nier, Dragos Iftimie, polycopié de l'université de Rennes 1 pour la licence de mathématiques

B Iconographie

Table des figures

Liste des tableaux

Table des matières