Première partie

Théorie des ensembles

Chapitre 1

Ensembles ordonnés, ordinaux

Ce chapitre est consacré à l'étude des propriétés fondamentales des ordinaux. Les ordinaux ne sont en fait que des ensembles munis d'une certaine relation d'ordre. Pourtant cette structure si simple suffit pour leur accorder le statut d'unité de mesure de la taille d'un ensemble. Ainsi, les ordinaux offrent une généralisation naturelle des nombres naturels. Comme chaque généralisation leur étude nécessite plus de soin en raison des complications supplémentaires qui sont invisibles au niveau des nombres naturels.

1.1 Ordinaux

Définition 1.1.1 *Soit E un ensemble muni d'une relation binaire $\mathcal{R}$. Voici quelques propriétés bien connues dont une telle relation peut jouir :*

R *$\mathcal{R}$ est dite* réflexive *si $a\mathcal{R}a$ pour tout $a \in E$.*

AR *$\mathcal{R}$ est dite* antiréflexive *si pour tout $a \in E$, $a \not\mathcal{R}a$.*

S *$\mathcal{R}$ est dite* symétrique *si pour toute paire $(a, b) \in E^2$, $a\mathcal{R}b$ et $b\mathcal{R}a$.*

AntiS *$\mathcal{R}$ est* antisymétrique *si pour toute paire $(a, b) \in E^2$, $a\mathcal{R}b$ et $b\mathcal{R}a$ implique $a = b$.*

AS *$\mathcal{R}$ est* asymétrique *si pour toute paire $(a, b) \in E^2$, il n'est pas le cas que $a\mathcal{R}b$ et $b\mathcal{R}a$.*

T *$\mathcal{R}$ est* transitive *si pour tous $a, b, c \in E$, $a\mathcal{R}b$ et $b\mathcal{R}c$ implique que $a\mathcal{R}c$.*

Définition 1.1.2 *Soit E un ensemble muni d'une relation $\mathcal{R}$ binaire.*

1. *La relation binaire $\mathcal{R}$ sur E sera dite* relation d'ordre *si $\mathcal{R}$ est* **R**, **AntiS** *et* **T**.
2. *Si $\mathcal{R}$ est une relation d'ordre, elle est dite* totale *ou* linéaire *si pour toute paire $(a, b) \in E^2$, soit $a = b$, soit $a\mathcal{R}b$, soit $b\mathcal{R}a$.*
3. *Si $\mathcal{R}$ est une relation d'ordre, elle est dite* stricte *si elle est* **AR**, **AS** *et* **T**.
4. *Un ensemble ordonné est dit* bien ordonné *si chacune de ses parties non vides a un plus petit élément. L'ordre d'un ensemble bien ordonné est dit un* bon ordre.

La paire $(E, \mathcal{R})$ est, où E est un ensemble ordonné, est un exemple de l'une des notions les plus importantes de ce cours, celle d'une *structure*, une entité formée par un ensemble sous-jacent muni éventuellement des relations et des fonctions et muni toujours de la relation d'égalité.

Nous avons fait certaines définitions. Les exemples de relations d'ordre sont bien connus. Démontrons donc quelques lemmes pour compléter le paysage mathématique de nos premiers pas.

Lemme 1.1.3 *Soit $(E, <)$ un bon ordre. Alors l'application identité est le seul* automorphisme *de $(E, <)$. En d'autres termes, c'est une structure dont le groupe d'automorphismes est trivial, une structure* rigide.

Preuve. Un automorphisme f de la structure $(E, <)$ est une bijection telle que $x < y$ si et seulement si $f(x) < f(y)$. L'automorphisme f n'est pas l'identité alors si et seulement si $D =$

$\{x \in E | f(x) \neq x\} \neq \emptyset$. L'ensemble D est le support de f, et bien sûr celui de f^{-1} aussi. Soit alors $x \in D$ le plus petit élément de D. Si $f(x) < x$ alors $x = f^{-1}(f(x)) = f(x)$, absurde. Alors $f(x) > x$. Or ceci équivaut à $x = f^{-1}(f(x)) > f^{-1}(x)$. Donc $f(f^{-1}(x)) = f^{-1}(x)$, Or $x = f(f^{-1}(x))$ aussi, absurde. □

On définit un isomorphisme entre deux structures ordonnées de façon analogue à la notion d'automorphisme d'un ensemble ordonné, une bijection qui préserve la structure.

Lemme 1.1.4 *Soient $(E, <)$ et $(F, \prec)$ deux ensembles bien ordonnés. Si ces deux structures sont isomorphes, il existe un seul isomorphisme entre les deux.*

Preuve. Ce résultat découle du lemme 1.1.3 puisque si f et g sont deux isomorphismes de $(E, <)$ vers $(F, \prec)$ alors fg^{-1} est un automorphisme de $(E, <)$. □

Lemme 1.1.5 *Si $(E, <)$ est un ensemble strictement bien ordonné. Si $a \in E$ alors $(E, <)$ et $\{b \in E | b < a\}$ ne sont pas isomorphes.*

Preuve. Supposons par l'absurde qu'un isomorphisme f existe. Alors $f(a) \neq a$ puisque $a \notin \{b \in E | b < a\}$. Donc $\{x \in E | f(x) \neq x\} \neq \emptyset$. On peut raisonner comme dans la preuve du lemme 1.1.3. □

Exemple 1.1.1 Soit $(E, \in)$ un ensemble muni de la relation d'appartenance usuelle. Supposons que pour tout $\alpha \in E$, α soit un sous-ensemble de E. Voici un exemple :

$$\{\{\emptyset\}, \{\{\emptyset\}\}\} .$$

Notons que notre exemple n'existerait pas si on n'était pas assuré de l'existence de $\emptyset$. Rassurons-nous :

Axiome d'existence : Il existe un ensemble sans éléments. Dans le "langage" de notre structure ordonnée, on peut formellement exprimer cet énoncé : $\exists x \forall y (y \notin x)$.

La relation $\in$ définit dans l'ensemble E une relation d'ordre. Motivé par ce constat, nous introduisons une notion importante :

Définition 1.1.6 *Un ensemble E est dit* transitif *si la relation $\in$ y définit une relation d'ordre, en d'autres terms si pour tout $x \in E$, $x \subset E$ (tout élément de E en est une partie).*

Soulignons que cette relation d'ordre n'est pas nécessairement linéaire. Voici un ensemble transitif qui n'est pourtant pas totalement ordonné :

$$\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\} .$$

Définition 1.1.7 *Un ensemble est un* ordinal *s'il est strictement et bien ordonné par la relation d'appartenance.*

Exemple 1.1.2 La motivation pour la notion d'ordinal est de trouver des exemples canoniques de bons ordres qui généralisent nombres naturels :

$$0 \quad \emptyset$$

$$1 \quad \{\emptyset\}$$

$$2 \quad \{\emptyset, \{\emptyset\}\}$$

$$3 \quad \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$$

$$\vdots$$

Axiome de l'infini : Il existe un ensemble inductif. Plus formellement,

$$\exists x(0 \in x \wedge \forall y((y \in x) \rightarrow S(y) \in x)) \ .$$

La lettre S est utilisée pour noter l'opération *successeur* sur les ensembles :

$$\text{Si } x \text{ est un ensemble, alors } S(x) = x \cup \{x\} \ .$$

Le nouvel axiome est suffisant pour continuer :

$$\omega = \{0, 1, 2, 3, \ldots\}$$
$$S(\omega) = \omega + 1 = \omega \cup \{\omega\}$$
$$\omega + 2, \ldots$$
$$\omega.2 = \omega + \omega$$
$$\omega.3, \omega.4, \ldots$$
$$\omega^2 = \omega.\omega$$
$$\vdots$$

Nous pouvons procéder à démontrer certaines propriétés de base des ordinaux. Mais puisque nous avons déjà pris l'habitude des axiomes, pourquoi ne pas en introduire un autre. D'ailleurs ce nouvel axiome, certains raisonnements manqueraient de rigueur dans ce qui suit tant cet axiome est naturel. En fait, nous introduirons un *schéma d'axiomes* puisque pour toute propriété, l'axiome a un énoncé particulier.

Nous venons d'utiliser un autre mot dangeureusement naturel, à savoir "propriété". C'est une notion qu'on pourrait rendre rigoureuse dès maintenant en ayant recours aux notions de la théorie des modèles telles que les langages, les formules du premier ordre, etc. Nous préférons attendre avant la partie consacrée à la théorie des modèles de ce cours avant d'atteindre ce niveau de rigueur et nous nous contentons de dire qu'une propriété dans ce contexte est une propriété mathématique sur des ensembles qui s'exprime en quantifiant les variables et les joignant soit par des symboliques logiques, soit par $\in$, soit par $=$ ("une propriété définissable").

Schéma d'Axiome de Compréhension : Soit $P(x)$ une propriété. Pour tout ensemble A, il existe un ensemble B tel que $x \in B$ si et seulement si $x \in A$ et $P(x)$ est vraie.

Ce schéma d'axiome, si naturel soit-il, est crucial pour éviter les paradoxes du type "l'ensemble de tous les ensembles".

Lemme 1.1.8 *Si α et β sont deux ordinaux tels que $\alpha \subset \beta$ et que $\alpha \neq \beta$ alors $\alpha \in \beta$.*

Preuve. Soient α et β comme dans l'énoncé. On note γ le plus petit élément de l'ensemble $\beta - \alpha$ par rapport à l'ordre défini sur β par $\in$. Ceci est consistant puisque β est par hypothèse un ordinal. Il suffira de montrer que $\gamma = \alpha$.

Notons que β étant un ordinal, tout élément de γ appartient à β. Alors un élément δ tel que $\delta \in \gamma - \alpha$ contredirait le choix minimal de γ. Ainsi, $\gamma \subseteq \alpha$.

Si $\delta \in \alpha$, alors comme β est totalement ordonné par l'appartenance, soit $\delta \in \gamma$, soit $\delta = \gamma$, soit $\gamma \in \delta$. La deuxième possibilité est exclue puisque par hypothèse $\gamma \notin \alpha$. Il en est de même pour la troisième puisque par hypothèse α est un ordinal, donc transitivement ordonné par l'appartenance et que nous supposons $\gamma \notin \alpha$. Ainsi, $\alpha \subseteq \gamma$. □

Lemme 1.1.9 *Si α et β sont deux ordinaux, il en est de même pour leurs intersections.*

Preuve. La vérification des propriétés de bon ordre strict étant laissée aux bons soins des lecteurs, vérifions que $\in$ est transitif. Or, α et β sont transitivement ordonnés par $\in$. Donc, si $\gamma \in \alpha \cap \beta$, alors $\gamma \subset \alpha \cap \beta$. □

Notation 1.1.10 *Si α et β sont des ordinaux, alors sauf mention contraire, $\alpha < \beta$ et $\beta < \alpha$ auront le même sens q ue $\alpha \in \beta$ et $\beta \in \alpha$ respectivement.*

Proposition 1.1.11

1. *Si α et β sont deux ordinaux alors $\alpha < \beta$ ou $\alpha = \beta$ ou $\beta < \alpha$.*
2. *Soit E un ensemble d'ordinaux. Alors E est bien ordonné par $<$.*
3. *Si α est un ordinal alors il en est de même pour $S(\alpha)$. Il n'existe pas d'ordinal entre α et $S(\alpha)$.*
4. *Il n'existe pas d'ensemble de tous les ordinaux.*

Preuve. **(1)** D'après le lemme 1.1.9, $\alpha \cap \beta$ est un ordinal. Alors, le lemme 1.1.8 montre que si $\alpha \cap \beta \neq \alpha$ et que $\alpha \cap \beta \neq \beta$ alors $\alpha \cap \beta \in \alpha$ et que $\alpha \cap \beta \in \beta$. Alors $\alpha \cap \beta \in \alpha \cap \beta$. Or ceci est impossible en raison de l'asymétrie de $\in$.

(2) Soit E un ensemble d'ordinaux. Le point précédent montre que $\in$ y définit un ordre total. C'est un ordre strict. Il reste à vérifier que c'est un bon ordre. Soient $A \subseteq E$ et $\alpha \in A$. Si $\alpha \cap A = \emptyset$ alors α est le plus petit élément de A. En effet, E est ordonné par $\in$. Si $\alpha \cap A \neq \emptyset$, alors soit $\beta \in \alpha \cap A$ le plus petit élément de cette intersection. Un tel élément existe puisque α est un ordinal. Comme α est transitivement ordonné par $\in$, $\beta \cap A = \emptyset$. Ainsi β est le plus petit élément de A.

(3) Comme par hypothèse α est un ordinal, $S(\alpha)$ est transitivement ordonné par $\in$. L'ordre imposé par $\in$ est bon d'après le deuxième point. Finalement, si $\alpha < \gamma \leq S(\alpha)$, alors il n'y a qu'une seule possibilité : $\gamma = S(\alpha)$.

(4) Soit E un ensemble d'ordinaux. Alors $\bigcup E$, en d'autres termes l'union des éléments de E, est un ordinal (pourquoi ?). Si on pose $\alpha = S(\bigcup E)$, alors α est un ordinal d'après le point (3) et $\alpha \notin E$. □

Définition 1.1.12

1. *Un ordinal qui est de la forme $S(\alpha)$ où α est un ordinal aussi est dit* successeur.
2. *Un ordinal qui n'est pas successeur est un* ordinal limite.
3. *Si E est un ensemble d'ordinaux, alors $\bigcup E$ est noté* $\sup E$.

1.2 Induction et récurrence transfinies

Nous commençons par le principe d'induction transfinie. D'une façon similaire au cas des nombres naturels, il en existe deux formes :

Théorème 1.1 *Soit $P(x)$ une propriété exprimée éventuellement en utilisant des paramètres. On suppose que*

$$\text{si } P(\beta) \text{ est vraie pour tout } \beta < \alpha \text{ alors } P(\alpha) \text{ soit vraie.}$$

Alors $P(\alpha)$ est vraie pour tous les ordinaux.

Preuve. Notons dès le départ que l'hypothèse implique que $P(0)$ soit vraie puisqu'il n'existe pas d'ordinal strictement inférieur à 0. Supposons par l'absurde qu'il existe un ordinal α pour lequel $P(\alpha)$ est fausse. Alors, l'ensemble suivant (c'est un ensemble !) est non vide :

$$\{\beta \leq \alpha | P(\beta) \text{ est fausse}\} .$$

D'après la proposition 1.1.11 (2), cet ensemble a un plus petit élément α_0. Or $P(0)$ est vraie. Donc $P(\alpha_0)$ est vraie, une contradiction. □

Voici la deuxième forme dont la preuve se réduit facilement à la première.

Théorème 1.2 *Soit $P(x)$ une propriété éventuellement avec paramètres. On suppose que*

1. *$P(0)$ soit vraie ;*
2. *$P(\alpha)$ entraîne $P(\alpha+1)$ pour tous les ordinaux α ;*
3. *pour tout ordinal limite $\alpha \neq 0$, si $P(\beta)$ est vraie pour tout $\beta < \alpha$, alors $P(\alpha)$ soit vraie.*

Alors $P(x)$ est satisfaite par tous les ordinaux.

La discussion du théorème de récurrence sera plus longue. Comme dans le cas de la compréhension il s'agit plutôt d'un schéma d'axiomes qu'un seul :

Schéma d'Axiome de Remplacement : Soit $P(x,y)$ une propriété telle que pour tout x il existe un y et un seul tel que $P(x,y)$ soit vrai. Alors, pour tout ensemble A, il existe un ensemble B tel que pour tout $x \in A$, il existe $y \in B$ tel que $P(x,y)$ soit vrai.

Dans les énoncés du théorème de récurrence nous utiliserons le mot "opération". Nous aurions pu choisir l'appelation "fonction". Par contre les fonctions se définissent sur des ensembles tandis que les opérations sur des classes plus "larges". En l'occurrence, les opérations dont le théorème de récurrence transfinie parlera seront définies sur les ordinaux, et nous avons vu dans la proposition 1.1.11 (4) que les ordinaux forment plus qu'un ensemble.

Nous introduisons un outil technique qui sera nécessaire dans la preuve. Quand α est un ordinal et G une opération sur les ordinaux, nous dirons qu'une fonction t est une *α-opération définie par G* si t est une fonction de domaine α et que pour tout $\beta < \alpha$, $t(\beta) = G(t\lceil\beta)$.

Le contenu du théorème de récurrence transfinie est que si une opération permet de définir une α-opération pour tout ordinal α, alors il existe une opération dont la restriction à tout ordinal α est donnée par la α-opération correspondante. Intuitivement parlant, il s'agit de montrer que s'il existe une famille d''approximations" vers une opération dont l'existence n'est pas vérifiée, en l'occurrence les α-opérations, alors en effet ce point limite existe. Voici la première version du théorème de récurrence, celle que nous démontrerons :

Théorème 1.3 *Soit G une opération. Alors il existe une opération F et une seule telle que pour tout ordinal α*

$$F(\alpha) = G(F\lceil\alpha) \ .$$

Avant d'en donner la preuve essayons d'expliciter ce que ce théorème nous dit. En fait, il permet de définir l'opération F : si on connait les valeurs de F pour tous les éléments de α (tous les ordinaux strictement inférieurs à α), on en déduit la valeur de $F(\alpha)$ en appliquant G.

Preuve. On définit alors la propriété $P(x,y)$ par

$$P(x,y) \quad = \quad \begin{cases} y = t(x) \text{ avec } t \text{ une } \alpha\text{-opération définie par } G & \text{si } x \text{ est un ordinal} \\ y = \emptyset & \text{si } x \text{ n'est pas un ordinal} \end{cases}$$

L'objectif est de montrer, avec le recours du schéma d'axiome de remplacement et en utilisant le principe d'induction transfinie, que P définit une opération sur tous les ordinaux. Pour ce faire il suffira de montrer que pour tout x il existe un y et un seul tel que $P(x,y)$ soit vraie. En effet, si ceci est vérifié alors le schéma de remplacement permet de conclure.

Nous procéderons en utilisant le principe d'induction transfinie. Si $x = \emptyset$ ou x n'est pas un ordinal alors $t(x) = \emptyset$ et y est uniquement déterminé. Maintenant fixons α un ordinal strictement supérieur à 0. Supposons que pour tout $\beta < \alpha$ il existe une β-opération et une seule définie par G.

On pose

$$T = \{t | t \text{ est la } \beta\text{-opération définie par } G \text{ avec } \beta < \alpha\} \ .$$

Le schéma de remplacement assure qu'un tel ensembe existe. On considère l'union de toutes ces fonctions (vues comme des familles de paires d'antécédent et d'image) qu'on note $\bigcup T$. On définit alors le candidat pour une nouvelle α-opération :

$$\tau = \left(\bigcup T\right) \cup \left\{\left(\alpha, G\left(\bigcup T\right)\right)\right\} \ .$$

Il faut d'abord vérifier que τ est une fonction de domaine $\alpha + 1$. Si t_1 et t_2 sont deux membres de τ de domaines β_1 et β_2 respectivement, nous pouvons supposer en utilisant la proposition 1.1.11 (1) que $\beta_1 \leq \beta_2$, soit encore $\beta_1 \subseteq \beta_2$. Montrons en utilisant l'induction transfinie que $t_2 \lceil \beta_1 = t_1$. Soit alors $\gamma < \beta_1$. Supposons que pour tout $\delta < \gamma$, $t_1(\delta) = t_2(\delta)$. Alors $t_1 \lceil \gamma = t_2 \lceil \gamma$. On en conclut alors que

$$t_1(\gamma) = G(t_1 \lceil \gamma) = G(t_2 \lceil \gamma) = t_2(\gamma) \ .$$

Comme γ était arbitrairement choisi parmi les ordinaux strictement inférieurs à β_1, on conclut que $t_2 \lceil \beta_1 = t_1$. Donc τ est une fonction. Maintenant, essayez de vous convaincre que son domaine est $\alpha + 1$.

L'étape suivante est de montrer que τ est une $(\alpha + 1)$-opération définie par G. Il faut donc vérifier que si $\beta \leq \alpha$ alors $\tau(\beta) = G(\tau \lceil \beta)$. Si $\beta = \alpha$ alors $G(\beta) = G(\alpha) = G(\bigcup T) = G(\tau \lceil \alpha) = G(\tau \lceil \beta)$. Si $\beta < \alpha$, alors il existe $t \in T$ tel que $\beta \in \mathrm{Dom}(t)$. Alors, $\tau(\beta) = t(\beta) = G(t \lceil \beta) = G(\tau \lceil \beta)$.

Il reste à vérifier l'unicité de τ. Soit σ est une autre $(\alpha + 1)$-opération définie par G. Soit $\gamma \leq \alpha$. Vérifions que $\sigma(\gamma) = \tau(\gamma)$. Or $\sigma(\gamma) = G(\sigma \lceil \gamma) = G(\tau(\lceil \gamma) = \tau(\gamma)$.

De ce qui précède, on conclut en utilisant le principe d'induction transfinie que P définit une opération F qui est unique puisque $F \lceil \mathrm{Dom}(t) = t$ et que t est unique. Finalement, pour tout ordinal α, soit t l'unique α-opération définie par G. Alors $F(\alpha) = t(\alpha) = t(F \lceil \alpha) = G(F \lceil \alpha)$. □

L'énoncé suivant est la version *paramétrique* du théorème de récurrence transfinie :

Théorème 1.4 *Soit G une opération en deux variables. Alors la propriété suivante définit une opération F telle que $F(z, \alpha) = G(z, F(z, \ ) \lceil \alpha))$ pour tous les ordinaux α et z :*

$$P(x, y, z) \quad = \quad \begin{cases} y = t(z, x) \text{ avec } t \text{ une } \alpha\text{-opération définie par } G & \text{si } x \text{ est un ordinal} \\ y = \emptyset & \text{si } x \text{ n'est pas un ordinal} \end{cases}$$

1.3 Arithmétique des ordinaux

Dans cette section nous expliciterons comment le théorème de récursion, permet de mettre en place les fondations de l'arithmétique des ordinaux. Pour ce faire nous aurons besoin de la forme suivante du théorème de la récursion :

Théorème 1.5 *Soient G_1, G_2 et G_3 des opérations, éventuellement à paramètres. Alors il existe une opération F et une seule telle que*

$$\begin{aligned} F(0) &= G_1(0) \\ F(\alpha + 1) &= G_2(F(\alpha)) \text{ pour tout ordinal } \alpha \\ F(\alpha) &= G_3(F \lceil \alpha) \text{ pour tout ordinal limite } \alpha \neq 0 \end{aligned}$$

Nous commençons avec l'addition :

Définition 1.3.1 *Pour tout ordinal β*
(a) $\beta + 0 = \beta$;
(b) $\beta + S(\alpha) = S(\beta + \alpha)$ *pour tout ordinal* α ;
(c) $\beta + \alpha = \sup\{\beta + \gamma | \gamma < \alpha\}$ *pour tout ordinal limite* $\alpha \neq 0$.

Essayons de voir comment le théorème de récursion intervient dans la définition 1.3.1. Il suffit de considérer les opérations suivantes dans le cadre du théorème 1.5 :

$$\begin{aligned} G_1(x,y) &= x \\ G_2(x,y) &= S(y) \\ G_3(x,y) &= \sup(\text{Im } y) \end{aligned}$$

Notons que Im est utilisé pour noter l'image d'une fonction. Alors, on obtient

$$\begin{aligned} A(x,0) &= G_1(x,0) = x, \\ A(x,S(\alpha)) &= G_2(x,A(x,\alpha)) = S(A(x,\alpha)) \text{ pour tout } \alpha, \\ A(x,\alpha) &= G_3(x,A(x,\ )\lceil\alpha) = \sup(\text{Im}(A(x,\ )\lceil\alpha)) = \sup\{A(x,\delta)|\delta<\alpha\} \text{ pour tout ordinal limite } \alpha\neq 0. \end{aligned}$$

Alors, $\beta+\alpha = A(\beta,\alpha)$.

Faisons quelques remarques rapides. Dans ce qui suit nous utiliserons la notation $\alpha+1$ au lieu de $S(\alpha)$. Notons que l'addition des ordinaux n'est pas une opération commutative. En effet, $1+\omega = \omega \neq \omega+1$.

Procédons de la même manière pour définir les produit et l'exponentiation des ordinaux.

Définition 1.3.2 *Pour tout ordinal β*
(*a*) $\beta.0 = 0$ *;*
(*b*) $\beta.(\alpha+1) = \beta.\alpha+\beta$ *pour tout ordinal α ;*
(*c*) $\beta.\alpha = \sup\{\beta.\gamma|\gamma<\alpha\}$ *pour tout ordinal limite $\alpha\neq 0$.*

On peut tout exprimer encore une fois en utilisant le formalisme pas très convivial du théorème de récurrence.

$$\begin{aligned} G_1(x,y) &= 0 \\ G_2(x,y) &= y+x \\ G_3(x,y) &= \sup(\text{Im } x). \end{aligned}$$

Alors, on obtient

$$\begin{aligned} M(x,0) &= G_1(x,0) = 0, \\ M(x,S(\alpha)) &= G_2(x,M(x,\alpha)) = M(x,\alpha)+x \text{ pour tout } \alpha, \\ M(x,\alpha) &= G_3(x,M(x,\ )\lceil\alpha) = \sup(\text{Im}(M(x,\ )\lceil\alpha)) = \sup\{M(x,\delta)|\delta<\alpha\} \text{ pour tout ordinal limite } \alpha\neq 0. \end{aligned}$$

Finalement,

$$\beta.\alpha = M(\beta,\alpha)\ .$$

Comme la somme, le produit n'est pas commutatif non plus. En effet, $2.\omega = \omega \neq \omega+\omega = \omega.2$.

La dernière opération arithmétique est l'exponentiation :

Définition 1.3.3 *Pour tout ordinal β*
(*a*) $\beta^0 = \beta$ *;*
(*b*) $\beta^(S(\alpha)) = \beta^\alpha.\beta$ *pour tout ordinal α ;*
(*c*) $\beta^\alpha = \sup\{\beta^\gamma|\gamma<\alpha\}$ *pour tout ordinal limite $\alpha\neq 0$.*

Voici le formalisme dans le cadre du théorème de récurrence :

$$\begin{aligned} G_1(x,y) &= 1 \\ G_2(x,y) &= y.x \\ G_3(x,y) &= \sup(\text{Im } y). \end{aligned}$$

Alors,

$$\begin{aligned} E(x,0) &= G_1(x,0) = 1, \\ E(x,S(\alpha)) &= G_2(x,E(x,\alpha)) = E(x,\alpha).x \text{ pour tout } \alpha, \\ E(x,\alpha) &= G_3(x,E(x,\ )\lceil\alpha) = \sup(\mathrm{Im}(E(x,\ )\lceil\alpha)) = \sup\{E(x,\delta)|\delta<\alpha\} \text{ pour tout ordinal limite } \alpha\neq 0. \end{aligned}$$

Donc,

$$\beta^\alpha = E(\beta,\alpha)\ .$$

1.4 Propriétés des opérations sur les ordinaux

Dans cette section nous démontrerons quelques propriétés des trois opérations introduites dans la section précédente. L'objectif est aussi de présenter le fonctionnement général des raisonnements qui utilisent l'induction transfinie.

Lemme 1.4.1 *Soient α,β et δ trois ordinaux.*
(a) Si $\alpha=\beta$ alors $\delta+\alpha=\delta+\beta$.
(b) $\alpha<\beta$ si et seulement si $\delta+\alpha<\delta+\beta$.
(c) $\alpha=\beta$ si et seulement si $\delta+\alpha=\delta+\beta$.
(d) $(\alpha+\beta)+\delta=\alpha+(\beta+\delta)$.

Preuve. (a) Exercice.

(b) Pour vérifier l'implication de gauche à droite, on appliquera le théorème 1.2 à la propriété $P(\delta,\alpha,x)$ qui énonce que $\delta+\alpha<\delta+x$ si $\alpha<x$. Si $P(\delta,\alpha,\beta)$ alors

$$\delta+\alpha<\delta+\beta<(\delta+\beta)+1=\delta+(\beta+1)\ .$$

Si β est un ordinal limite alors $\delta+\alpha<\delta+\gamma$ pour tout $\gamma<\beta$ tel que $\alpha<\gamma$. Alors, $\delta+\beta=\sup_{\gamma<\beta}(\delta+\gamma)$. Or $\sup_{\gamma<\beta}(\delta+\gamma)$ est supérieur à tous les $\delta+\gamma$. Donc, $\delta+\alpha<\delta+\beta$.

Pour vérifier l'implication inverse, on commence par $\delta+\alpha<\delta+\beta$. Si $\beta\leq\alpha$, alors la première moitié de la preuve montre que $\alpha=\beta$. Dans ce cas, le point (a) permet d'aboutir à une contradiction.

(c) Si $\alpha=\beta$, alors le point (a) montre que $\delta+\alpha=\delta+\beta$. Si $\alpha\neq\beta$, alors la proposition 1.1.11 (1) montre que $\alpha<\beta$ ou $\beta<\alpha$. Alors le point (b) permet de conclure.

(d) Dans ce point la propriété $P(\alpha,\beta,x)$ qui énonce que $(\alpha+\beta)+x=\alpha+(\beta+x)$ sera vérifiée en utilisant le théorème 1.2. Clairement, $P(\alpha,\beta,0)$. Supposons maintenant $P(\alpha,\beta,\delta)$. Alors,

$$(\alpha+\beta)+(\delta+1)=((\alpha+\beta)+\delta)+1=(\alpha+(\beta+\delta))+1=\alpha+(\beta+(\delta+1))\ .$$

Finalement, si δ est un ordinal limite, alors

$$(\alpha+\beta)+\delta=\sup\{(\alpha+\beta)+\gamma|\gamma<\delta\}=\sup\{\alpha+(\beta+\gamma)|\gamma<\delta\}\ .$$

Il suffit donc de vérifier que la dernière expression est égale à $\alpha+(\beta+\delta)$. Or on peut vérifier en comparant les éléments des deux membres de l'égalite que

$$\sup\{\alpha+(\beta+\gamma)|\gamma<\delta\}=\sup\{\alpha+\xi|\xi<\beta+\delta\}\ .$$

La conclusion découle de la définition de l'addition des ordinaux. □

Le lemme suivant décrit les propriétés analogues du produit des ordinaux. Sa preuve est un bon exercice d'entraînement :

Lemme 1.4.2 *Soient α,β et δ trois ordinaux.*
(a) Si $\alpha=\beta$ alors $\delta.\alpha=\delta.\beta$.
(b) $\alpha<\beta$ si et seulement si $\delta.\alpha<\delta.\beta$ ($\delta\neq 0$).
(c) $\alpha=\beta$ si et seulement si $\delta.\alpha=\delta.\beta$ ($\delta\neq 0$).
(d) $(\alpha.\beta).\delta=\alpha.(\beta.\delta)$.

1.5 Les ensembles bien ordonnés et les ordinaux

Les ordinaux en général, contrairement aux nombres naturels, ne caractérisent pas la taille d'un ensemble. Il y a beaucoup d'ordinaux qui sont des ensembles dénombrables. Par contre, l'ordre imposé par $\in$ caractérise à isomorphisme près un nombre ordinal fixé. L'objectif de cette section est la preuve d'un théorème qui justifie cette conclusion :

Théorème 1.6 *Soit $(E, \prec)$ un ensemble bien ordonné. Alors il existe un ordinal α et un seul tel que $(\alpha, <)$ soit isomorphe à $(E, \prec)$. En plus, il existe exactement un isomorphisme entre les deux ensembles.*

Preuve. L'unicité de l'ordinal et celle de l'isomorphisme découlent des lemmes 1.1.5 et 1.1.4 respectivement. Il suffit donc de démontrer l'existence d'un ordinal ayant les propriétés exigées par l'énoncé.

On pose $E_x = \{y \in E | y \prec x\}$. Intuitivement, E_x est un "segment initial" de E. L'idée qui mène les raisonnements ci-dessous est qu'un segment initial d'un bon ordre (resp. ordinal) est un bon ordre (resp. ordinal). Nous définirons A comme l'ensemble des éléments $x \in E$ tels que l'ensemble bien ordonné $(E_x, \prec)$ soit isomorphe à un ordinal. Comme un tel ordinal est unique, le schéma de remplacement permet de définir une fonction f sur l'ensemble A :

f(x)=le seul ordinal isomorphe à $(E_x, \prec)$.

Soit maintenant α le plus petit ordinal qui n'appartient pas à $\mathrm{Im}(f)$. L'objectif est de montrer que f est un isomorphisme de $(E, \prec)$ sur α. Comme pour tout isomorphisme, les conditions à vérifier sont la préservation de la structure, la bijectivité de l'application en question, ses ensembles de départ et d'arrivée.

f préserve la structure : Pour ce faire, il suffit de vérifier que pour tous $x, y \in A$, si $x \prec y$ alors $f(x) \in f(y)$. Pour un tel choix de x et de y, soit h l'isomorphisme de $(E_y, \prec)$ sur $f(y)$. Le choix de x implique que $x \in E_y$, en d'autres termes $E_x \subset E_y$. La restriction de h à E_x est un isomorphisme de $(E_x, \prec)$ vers $f(y)$. Le lemme 1.1.5 montre que l'image de E_x sous l'action de h ne peut être que $f(x)$.

f est injective : Si $f(x) = f(y)$ alors le lemme 1.1.5 montre que $E_x = E_y$ et en conséquence $x = y$.

$\mathrm{Im}(f) = \alpha$: Soit $\beta \in \alpha$. Grâce au choix minimal de α, $\beta \in \mathrm{Im}(f)$. Inversement, soit $\beta \in \mathrm{Im}(f)$. Alors β est le seul ordinal isomorphe à $(E_x, \prec)$ pour un certain $x \in A$.

Montrons que $\beta \in \alpha$. Sinon, la proposition 1.1.11 (1) montre que $\alpha \in \beta$ ou que $\alpha = \beta$. Comme $\alpha \notin \mathrm{Im}(f)$, $\alpha \neq \beta$. Alors la seule possibilité est que $\alpha \in \beta$. Or, si $\alpha \in \beta$, l'isomorphisme entre $(E_x, \prec)$ et β associe à α un certain $y \prec x$. En plus, α est la possibilité unique dans cette association. Alors, $f(y) = \alpha$, ce qui contredit le choix de $\alpha \notin \mathrm{Im}(f)$.

$\mathrm{Dom}(f) = E$: On procède par l'absurde. Soit donc $x \in E - A$ minimal par rapport à $\prec$. Alors $f\restriction(E_x, \prec)$ est un isomorphisme vers α et l'image est un ordinal $\beta \in \alpha$. Il en découle que $f(x) = \beta$, absurde puisque x était choisi en dehors de A. □

Notons qu'une conséquence du théorème 1.6 est la généralisation immédiate des théorèmes d'induction et de récurrence transifinies aux ensembles bien ordonnés.

Un point qui mérite d'être souligné est que le théorème 1.6 ne dit absolument rien sur quels ensembles sont bien ordonnables. Son point de départ est un tel ensemble et sa motivation est de lui associer un représentant canonique qui s'avère être un ordinal.

1.6 Cardinaux, une introduction courte

Comment mesure-t-on la taille d'un ensemble ? Pour un ensemble fini, il existe une réponse "naturelle" : on compte le nombre de ses éléments. On *compare* donc l'ensemble en question à un nombre naturel, en d'autres termes à un ordinal fini. Cette idée de comparaison est plus générale :

Définition 1.6.1 *Deux ensembles A et B sont dits* équipotents *s'il existe une bijection entre les deux. Dans ce cas on écrit $|A| = |B|$.*

A une première vue, la notion d'équipotence semble généraliser le comptage fini à un ensemble quelconque. Or ce n'est pas tout à fait le cas, et ce pour plusieurs raisons. Tout d'abord, contrairement aux ordinaux finis, dans le cas général des ordinaux infinis, il en existe beaucoup qui sont équipotents. En effet $|\omega| = |\omega + 1| = |\omega + 2| = \ldots = |\omega + \omega| = \ldots = |\omega^\omega| = \ldots$. Il faut donc choisir d'une façon canonique un ordinal qui puisse présenter la "taille" d'un ensemble qui lui est équipotent. Si cette tentative réussit, elle fournira la bonne notion de "comptage infinie". La définition suivante met en oeuvre cette idée :

Définition 1.6.2 *Un nombre ordinal α est dit un* cardinal *s'il n'est équipotent à aucun β tel que $\beta \in \alpha$.*

Tout ordinal fini est un cardinal. Le plus petit cardinal infini (par rapport à l'ordre des ordinaux) est ω. En existe-t-il d'autres ? La réponse est affirmative mais nécessite plus de travail que nous effectuerons à la fin de cette section.

Avant de continuer, il nous faut faire face à une autre question qui est susceptible de ne pas être vue, tant sa "réponse" est naturelle. Si A et B sont deux ensembles, nous avons donné une définition de l'égalité de leur taille qui est conforme aux intuitions et à l'usage mathématique. On peut même généraliser :

Notation 1.6.3 *Soient A et B deux ensembles. On note*

1. *$|A| \leq |B|$ s'il existe une injection de A vers B ;*
2. *$|A| < |B|$ s'il existe une injection de A vers B mais il n'en existe pas de B vers A.*

Comme maintenant nous avons, grâce à la définition 1.6.2, une notion robuste de taille aussi, nous sommes convaincus que nous pouvons comparer et donc mesurer la taille de tout ensemble. Il s'agit après tout de trouver des injections entre des ensembles. Or, comment sait-on qu'il en existe une entre deux ensembles arbitrairement choisis quand ceux-ci sont infinis ? En fait, la réponse n'est affirmative qu'en présence d'un nouvel axiome, notamment l'Axiome du Choix qui sera le sujet du chapitre suivant.

La situation n'est pourtant pas si désespérée même en l'absence de l'Axiome du Choix. En effet, nous connaissons une classe d'ensembles où la comparaison des tailles est possible : celle des ordinaux, plus généralement (le théorème 1.6) celle des ensembles bien ordonnables. Le lemme suivant est un corollaire du théorème 1.6.

Lemme 1.6.4 *Un ensemble E est bien ordonnable si et seulement s'il est équipotent à un ordinal.*

Ce lemme qui peut paraître innocent énonce quand-même la forte conclusion qu'il suffit d'être bien ordonné pour être isomorphe à un ordinal, donc pour avoir une taille qui peut être mesurée par un ordinal. En conséquence, si l'on savait que tout ensemble est bien ordonnable, la proposition 1.1.11 permettrait de comparer tous les ensembles.

Ce qui précède se lie rapidement aux cardinaux :

Lemme 1.6.5 *Tout ensemble E bien ordonnable est équipotent à un cardinal et un seul, qu'on note $|E|$. En particulier, si E est dénombrable est l'ordinal $|E| = \omega$.*

Preuve. Le théorème 1.6 montre l'existence d'un ordinal α auquel E est équipotent. Parmi les éléments de α équipotents à E, soit α_0 le plus petit. Alors α_0 est un cardinal car sinon il existerait $\beta < \alpha_0$ tel que $|\alpha_0| = |\beta|$. Alors, $|\beta| = |E|$ et ceci contredirait le choix minimal de α_0. □

Définition 1.6.6 *Soit E un ensemble. Alors on note $h(E)$ est le plus petit ordinal qui ne soit équipotent à aucune partie de E.*

Lemme 1.6.7 *Pour tout ensemble E, $h(E)$ existe, et c'est un cardinal.*

Preuve. Commençons par l'existence. Tout ensemble E possède des parties qui sont des ordinaux, par exemples les parties finies de E. D'après le théorème 1.6 pour chaque partie bien ordonnée de E, il existe un ordinal et un seul qui lui soit isomorphe. Alors, le schéma de remplacement nous assure de l'existence d'un ensemble H d'ordinaux défini par la propriété de contenir exactement les ordinaux isomorphes aux parties bien ordonnées de E. Comme il s'agit d'un ensemble d'ordinaux, $\sup(H)$ est un ordinal. L'ordinal $\sup(H)$ ne peut pas être équipotent à une partie bien ordonnée de E puisque sinon $\sup(H) \in H$, absurde. Inversement, si $\beta \in \sup(H)$ alors β est par définition de H équipotent à une partie bien ordonnée de E. Ainsi, $h(E) = \sup(H)$.

Pour montrer que $h(E)$ est un cardinal, on peut procéder par l'absurde et supposer l'existence d'un ordinal η tel que $\eta \in h(E)$ et que $|\eta| = |h(E)|$. Or il en découle immédiatement que $h(E)$ est équipotent à une partie bien ordonnée de E, ce qui contredit la définition de $h(E)$. □

Le corollaire suivant doit éliminer tous les doutes sur l'existence des ordinaux infinis non dénombrables.

Corollaire 1.6.8 *Si E est un ensemble bien ordonnable alors $h(E)$ est un cardinal tel que à la fois $\alpha \in h(E)$ avec α l'ordinal isomorphe à E (le lemme 1.6.4), et que $|\alpha| < |h(E)|$, en d'autres termes, il existe une injection de α vers $h(E)$ mais il n'en existe pas de $h(E)$ vers α.*

A ce stade, nous sommes prêts à définir une hiérarchie ou échelle des cardinaux mais nous préférons patienter un peu.

Chapitre 2

Axiome du choix

Ce chapitre est consacré à l'Axiome du Choix. C'est un axiome un peu particulier. Comme le montreront des exemples en cours et aux travaux dirigés, l'axiome du choix très fréquemment utilisé en mathématiques. La plupart du temps, il s'agit des applications très "naturelles", intuitivement justifiées. Par contre, cet axiome entraîne parfois des conclusions étonnantes, ce qui fait douter si c'est judicieux de lui accorder le statut d'axiome. Les débats autour de ce statut ont fait couler beaucoup d'encre, sinon de sang, au début du 20e siècle.

Il faut bien souligner qu'un bon nombre des usages de l'axiome du choix, avec des conclusions étonnantes ou naturelles, sont indispensables. En utilisant des méthodes parfois très avancées, il est possible de montrer que beaucoup d'usages sont en fait nécessaires. Sans l'axiome du choix, l'arithmétique des cardinaux serait très réduite, il existerait des espaces vectoriels sans base, il serait difficile de faire de l'analyse fonctionnelle voire aborder la topologie. Par contre des théorèmes d'existence surprenants seraient évités.

Ce qui rend l'axiome du choix particulier n'est pas ses conéquences naturelles ou étranges. En science il est des résultats naturels, et d'autres inattendus. C'est un axiome qui énonce l'existence de certaines fonctions d'une façon très "non définissable" dans le cadre que nous nous sommes fixé, celui des ensembles avec les seules relations $\in$ et $=$. Admettre l'axiome du choix ajoute une propriété externe aux structures que nous nous sommes données. Mais, cela se fait beaucoup en théorie des modèles, et plus généralement en mathématiques. Et, nous ferons des mathématiques "avec choix".

2.1 Fonctions de choix

Dans cette section, nous introduirons l'Axiome du choix et en discuterons certaines importantes pour la théorie des ensembles. Ces énoncés équivalents permettent une étude cohérente des ensembles en utilisant les notions d'ordinal et de cardinal. Cela deviendra de plus en plus visible dans ce chapitre et le suivant.

Définition 2.1.1 *Soit $\mathcal{F}$ une famille d'ensembles. Une fonction f définie sur $\mathcal{F}$ est dite une* fonction de choix *si pour tout $E \in \mathcal{F}$ non vide, $f(E) \in E$.*

Lemme 2.1.2 *Soit $\mathcal{F}$ un système* fini *d'ensembles. Alors $\mathcal{F}$ admet une fonction de choix.*

Preuve. C'est une récurrence immédiate. Plus important est de constater que si $\mathcal{F} = \{E_1, \ldots, E_n\}$ et que l'on fixe un élément c_i de chaque E_i, alors l'ensemble $\{(E_i, c_i) | 1 \leq i \leq n\}$ décrit une propriété. Plus tard, nous dirons qu'un ensemble fini est définissable si on permet des constantes. □

Question : Pourquoi est-ce que cette preuve ne marche pas si on remplace "fini" par "dénombrable" ?

Axiome du choix : Pour toute famille d'ensembles il existe une fonction de choix.

Théorème 2.1 *Les énoncés suivants sont équivalents :*

(Axiome du Choix) *Pour toute famille d'ensembles il existe une fonction de choix.*

(Principe du Bon Ordre) *Tout ensemble peut être bien ordonné.*

(Lemme de Zorn) *Si toute partie totalement ordonnée d'un ensemble E partiellement ordonné admet un majorant dans E, alors E a un élément maximal.*

Preuve. Commençons par montrer que l'Axiome du Choix (AC) implique le Principe du Bon Ordre. Soit donc A un ensemble. (AC) assure qu'il existe une fonction de choix G dont l'ensemble de départ est $\mathcal{P}(A)$. On fixe un ordinal ξ qui n'est pas dans A et on définit alors par récurrence transfinie l'opération suivante :

$$F(\alpha) \quad = \quad \begin{cases} G(A \setminus \mathrm{Im}(F\lceil\alpha)) & \text{si } A \setminus \mathrm{Im}(F\lceil\alpha) \neq \emptyset \\ \xi & \text{si } A \setminus \mathrm{Im}(F\lceil\alpha) = \emptyset \end{cases}$$

Pour conclure que A est bien ordonnable, il suffira de montrer que F permet de définir une bijection. Constatez que pour le moment nous ne savons pas si cette construction par récurrence transfinie s'arrête. Nous utiliserons le cardinal $h(A)$ introduit dans la définition 1.6.6.

Soient $\alpha \in \beta \in h(A)$ deux ordinaux. Si $F(\beta) \neq \xi$ alors $F(\beta) = G(A \setminus \mathrm{Im}(F\lceil\beta))$. Or $G(A \setminus \mathrm{Im}(F\lceil\beta)) \in A \setminus \mathrm{Im}(F\lceil\beta)$ par la définition de G tandis que $F(\alpha) \in \mathrm{Im}(F\lceil\beta)$. Ainsi $F(\alpha) \neq F(\beta)$. Si $F(\alpha) \neq \xi$ pour tout $\alpha \in h(A)$ alors le raisonnement précédent montre qu'il existe une injection de $h(A)$ vers A, ce qui contredit la définition de $h(A)$. Cette dernière conclusion nous permet de définir $\lambda = \inf\{\alpha \in h(A) | F(\alpha) = \xi\}$. La preuve sera finie si on peut montrer que l'injection $F : \lambda \longrightarrow A$ est aussi surjective. Ceci est immédiat puisque sinon $A \setminus \mathrm{Im}(F\lceil\lambda) \neq \emptyset$, et alors par définition $F(\lambda) \neq \xi$.

Montrons maintenant que le (BO) implique le Lemme de Zorn (Z). Soit donc E un ensemble ordonné qui vérifie l'hypothèse de la condition (Z). Notons l'ordre partiel sur E par $\prec$. Par hypothèse il existe un bon ordre $<$ sur E. Ces deux ordres n'ont a priori aucun lien. Sans perte de généralité, nous les supposerons stricts. Notre objectif est de définir une fonction de domaine E en utilisant le théorème de récurrence transfinie. Pour ce faire nous définissons d'abord la fonction suivante sur E :

$$G(x) \quad = \quad \begin{cases} \sup_{\prec}(\inf_{<}(E - \mathrm{Im}(x)) \cup \mathrm{Im}(x)) & \text{si } \inf_{<}(E - \mathrm{Im}(x)) \text{ et } \mathrm{Im}(x) \text{ sont comparables} \\ \mathrm{Im}(x) & \text{sinon} \end{cases}$$

Soulignons que le sup et l'inf sont calculés par rapport aux deux ordres différents $\prec$ et $<$ respectivement. Alors, le théorème de récurrence transfinie montre qu'il existe une fonction F telle que

$$\text{pour tout } a \in E, \quad F(a) = G(F\lceil\{b \in E | b < a\}) \ .$$

La fonction F n'est en fait qu'une construction formelle d'une "suite croissante par rapport à $\prec$ éventuellement avec répétitions" en utilisant le bon ordre $<$. Par l'hypothèse du lemme de Zorn, $\mathrm{Im}(F)$ possède un majorant qui appartient à E que nous noterons a_∞.

Montrons maintenant que a_∞ est un élément maximal de E par rapport à $\prec$. S'il existe $x \in E$ tel que $a_\infty \prec x$ alors, comme E est bien ordonné par $<$, soit $a_\infty < x$, soit $x < a_\infty$. Dans le premier cas, $F(a_\infty) = a_\infty$ mais $F(x) = x$ ce qui contredit que a_∞ majore l'image de F. Dans le deuxième cas, $F(x) = x$ et $F(a_\infty) = x$, et on contredit encore une fois la propriété de a_∞.

Dernièrement montrons que (Z) implique (AC). Nous "construirons" une fonction. Soit donc $\mathcal{E}$ une famille d'ensembles non vides. Sur $\mathcal{E}$, il existe des familles de fonctions de choix "partielles". Si par exemple $E \in \mathcal{E}$ alors pour tout $x \in E$, le singleton $\{(E, x)\}$ est une fonction de choix pour la sous-famille $\{E\}$. L'ensemble vide est un autre exemple.

Soit $\mathcal{F}$ la famille de toutes les fonctions de choix partielles sur E. Les remarques du paragraphe précédent montrent que c'est une famille non vide. La famille $\mathcal{F}$ est partiellement ordonnée par $\subset$, et si $\mathcal{F}_0$ en est une partie totalement ordonnée, alors $\bigcup \mathcal{F}_0$ définit une nouvelle fonction de choix. Ainsi, $\bigcup \mathcal{F}_0 \in \mathcal{F}$, et il en découle que $(\mathcal{F}, \subset)$ vérifie les conditions de (Z). Alors, $\mathcal{F}$ a un élément maximal par rapport à $\subset$ que nous noterons f_∞. Montrons que f_∞ est une fonction de choix pour $\mathcal{F}$. Comme c'est une fonction de choix partielle, il suffit de vérifier que tout ensemble

E dans $\mathcal{F}$ appartient au domaine de f_∞. Si $E \in \mathcal{F} \setminus \mathrm{Dom}(f)$ alors on peut fixer un élément x arbitrairement choisi de E et $f_\infty \cup \{(E, x)\}$ est un élément de $\mathcal{F}$ qui contredit la maximalité de f dans $\mathcal{F}$ par rapport à $\subset$. La preuve est terminée. □

La preuve que nous avons donnée du théorème 2.1 n'est pas la seule puisque toutes les équivalences peuvent être démontrées sans passer par un troisième énoncé. Nous laissons cela comme un exercice utile aux lecteurs.

2.2 Axiome du choix et les cardinaux

Dans la section 1.6, nous avons rigoureusement défini la notion de cardinal. C'est une notion dont la richesse augmente substantiellement en présence de l'Axiome du Choix parce que la plupart des résultats sur les cardinaux en dépendent. Dans cette section nous illustrerons quelques exemples de ce phénomène.

Proposition 2.2.1

1. **(AC)** *Tout ensemble est équipotent à un cardinal.*
2. **(AC)** *Pour toute paire d'ensembles A et B, soit $|A| \leq |B|$, soit $|B| \leq |A|$.*
3. **(AC)** *Si f est une application et A est un ensemble, alors $|f(A)| \leq |A|$.*

Preuve. Le premier énoncé découle du lemme 1.6.4 et du théorème 2.1.

Le deuxième point découle de la proposition 1.1.11 (a) et du premier point.

La preuve du troisième point est une illustration de comment certains raisonnements "naturels" utilisent de façon cruciale l'Axiome du Choix. On peut supposer f non vide. On définit $A_x = \{y \in A | f(y) = f(x)\}$. D'après (AC), il existe une fonction de choix $g_{\mathcal{A}}$ sur la famille $\mathcal{A} = \{A_x | x \in A\}$. Pour conclure, il suffira de montrer que l'application suivante est injective :

$$\begin{array}{rccc} \iota & : & f(A) & \longrightarrow & A \\ & & f(x) & \longmapsto & g_{\mathcal{A}}(A_x) \end{array} .$$

Or, les ensembles A_x forment une partition de A. Ainsi, si $\iota(f(x)) = \iota(f(y))$ alors $g_{\mathcal{A}}(A_x) = g_{\mathcal{A}}(A_y)$, et donc, $A_x = A_y$. En d'autres termes, $f(x) = f(y)$. □

La proposition 2.2.1 a des conséquences importantes pour les cardinaux. Son deuxième point montre que l'ordre usuel des cardinaux (voir par exemple la notation 1.6.3) est consistant avec celui des ordinaux. En d'autres termes, pour deux ensembles arbitraires A et B, $|A| < |B|$ (il existe une bijection de A vers B mais il n'en existe aucune surjective) si et seulement si $|A| \in |B|$ en tant qu'ordinaux.

Avant de finir ce chapitre, donnons une dernière application de l'Axiome du Choix.

Proposition 2.2.2 (AC) *L'union d'un nombre dénombrable d'ensembles dénombrables est dénombrable.*

Preuve. Soit $\{A_i | i < \omega\}$ la famille d'ensembles en question. La preuve est très "intuitive" donc on peut même ne pas s'apercevoir de l'usage de (AC). Il s'agit de construire une application de $\mathbb{N} \times \mathbb{N}$ sur l'union en question. On définit donc

$$\begin{array}{rccc} \Phi & : & \mathbb{N} \times \mathbb{N} & \longrightarrow & \bigcup_{i<\omega} A_i \\ & & (n, k) & \longmapsto & g_n(k) \end{array}$$

où $g_n : \mathbb{N} \longrightarrow A_n$ est *une* énumération de A_n. □

Notons que la proposition 2.2.2 peut être montrée en supposant une version plus faible de l'Axiome du Choix mais qu'on ne peut pas complètement se passer d'un certain axiome du choix, et que ceci est un théorème.

Chapitre 3

Cardinaux, proprement dits

Nous avons déjà défini rigoureusement dans la section 1.6 la notion de cardinal en utilisant les ordinaux et montré comment cette définition se lie à la conception intuitive de la notion de cardinal. Au deuxième chapitre, la proposition 2.2.1 a montré que l'Axiome du Choix justifie ce lien et que la comparaison des cardinaux des ensembles en utilisant des injections entre ceux-ci est cohérente avec l'ordre défini par $\in$ entre les ordinaux.

Dans ce chapitre, nous développerons les bases de l'arithmétique des cardinaux. Sauf au début, l'axiome du choix sera omniprésent.

Un mot sur la notation et un autre sur les bases... Si E est un ensemble, nous noterons $\mathcal{P}(E)$ l'ensemble des sous-ensembles de E. Il faut souligner que l'existence de $\mathcal{P}(E)$ est un axiome :

Axiome de l'ensemble des parties : Pour tout ensemble E, il existe un ensemble $\mathcal{P}(E)$ tel que $X \in \mathcal{P}(E)$ si et seulement si $X \subset E$.

3.1 Résultats de base et la hiérarchie des cardinaux infinis

Théorème 3.1 (Cantor) *Pour tout ensemble E, $|E| < |\mathcal{P}(E)|$.*

Preuve. Clairement, $|E| \leq |\mathcal{P}(E)|$ puisque l'application qui associe à chaque $x \in E$, le singleton $\{x\}$ est une injection de E vers $\mathcal{P}(E)$. Si les deux ensembles étaient équipotents alors il existerait une bijection f de E vers $\mathcal{P}(E)$. Or, $\Delta = \{x \in E : x \notin f(x)\}$ est une partie de E qui n'a pas d'antécédent, une contradiction. □

La preuve précédente cache en fait ce qu'on appelle parfois le raisonnement diagonal. En effet, l'existence d'une bijection peut être interprétée comme l'existence d'une table dont les lignes, indexées par les éléments de E, correspondent aux éléments de $\mathcal{P}(E)$. Alors, Δ se construit à partir de la diagonale.

Le résultat suivant a une preuve plus compliquée mais son énoncé, très intuitif d'ailleurs, est bien connu.

Théorème 3.2 (Cantor-Bernstein) *Si X et Y sont deux ensembles tels que $|X| \leq |Y|$ et $|Y| \leq |X|$ alors $|X| = |Y|$.*

Preuve. On utilise la méthode de l'enseignant paresseux. La preuve, étant plus compliquée que celle du théorème 3.1, est reléguée aux travaux dirigés. □

Maintenant, en utilisant nos outils disponibles, nous introduirons une hiérarchie qui contient tous les cardinaux infinis. Cette définition, faite en utilisant la récurrence transfinie, ne nécessite pas l'axiome du choix. Dans la définition, pour tout ordinal α, nous utiliserons le cardinal $h(\alpha)$ introduit dans la définition 1.6.6.

Définition 3.1.1

$$\begin{aligned} \omega_0 &= \omega \\ \omega_{\alpha+1} &= h(\omega_\alpha) \\ \omega_\alpha &= \sup\{\omega_\beta | \beta < \alpha\} \text{ } \textit{si } \alpha \textit{ est un ordinal limite non nul.} \end{aligned}$$

Remarquons qu'il découle immédiatement de cette définition que les ω_α forment une suite strictement croissante d'ordinaux. S'ils sont des cardinaux et si tout cardinal peut être présenté sous la forme ω_α pour un certain ordinal α sont des questions naturelles auxquelles nous chercherons maintenant des réponses.

Lemme 3.1.2 *Pour toute paire d'ordinaux α et β, si $\alpha < \beta$ alors $|\omega_\alpha| < |\omega_\beta|$.*

Preuve. On peut procéder par récurrence sur β. Si $\beta = \gamma + 1$ pour un certain γ, alors $\alpha \leq \gamma$. Par récurrence, $|\omega_\alpha| \leq |\omega_\gamma|$. Or, la clause de la définition 3.1.1 pour les successeurs montre que $|\omega_\gamma| < |\omega_\beta|$. Alors, $|\omega_\alpha| < |\omega_\beta|$.

Si β est un ordinal limite et non nul, alors il existe γ tel que $\alpha < \gamma < \beta$. Par récurrence $|\omega_\alpha| < |\omega_\gamma|$. Comme, la troisième clause de la définition 3.1.1 implique que $\omega_\gamma \subset \omega_\beta$, $|\omega_\gamma| \leq |\omega_\beta|$. La conclusion en découle. □

Lemme 3.1.3 *Pour tout ordinal α, ω_α est un cardinal.*

Preuve. La preuve se fait en appliquant le principe d'induction transfinie aux indices des ω_α. La définition 3.1.1 et le lemme 1.6.7 montrent qu'il n'y a que le cas des α limites non nuls à étudier. Supposons par l'absurde qu'il existe ω_α avec α limite et non nul et un ordinal γ tels que $|\omega_\alpha| = |\gamma|$ mais que $\gamma < \omega_\alpha$. D'après la troisième clause de la définition 3.1.1, il existe $\beta < \alpha$ tel que $\gamma \leq \omega_\beta$. Par l'hypothèse d'induction, ω_β est un cardinal, et on en conclut en utilisant la deuxième clause de la définition 3.1.1 que

$$|\gamma| \leq |\omega_\beta| < |\omega_{\beta+1}| \ .$$

Or α étant limite, $\beta + 1 < \alpha$ et on conclut que $|\gamma| < |\omega_\alpha|$, absurde. □

Notons que non seulement les ω_α sont des cardinaux, mais ils forment une suite strictement croissante au sens de l'équipotence.

Lemme 3.1.4 *Pour tout ordinal α, $\alpha \leq \omega_\alpha$.*

Preuve. La preuve se fait en utilisant le principe d'induction transfinie. Soit donc α un ordinal. Si $\alpha = 0$, la réponse est donnée par la définition.

Si $\alpha = \beta + 1$, alors par l'hypothèse de récurrence, $\beta \leq \omega_\beta$. Par ailleurs, le lemme 3.1.2 montre que $|\omega_\beta| < |\omega_\alpha|$. Or, d'après le lemme 3.1.3, ω_β et ω_α sont des cardinaux. En conséquence $|\omega_\beta| = \omega_\beta$ et $|\omega_\alpha| = \omega_\alpha$, et $\omega_\beta < \omega_\alpha$. On déduit de ce qui précède que $\beta < \omega_\alpha$ et que, finalement, $\alpha = \beta + 1 \leq \omega_\alpha$.

Si α un ordinal limite, alors par définition, $\omega_\alpha = \sup\{\beta < \alpha | \omega_\beta\}$. Si $\omega_\alpha < \alpha$, alors il existe un ordinal β tel que $\omega_\alpha < \beta < \alpha$. Comme, par récurrence, $\beta \leq \omega_\beta$, on conclut les inégalités suivantes : $\omega_\alpha < \beta \leq \omega_\beta$. Or, d'après le lemme 3.1.2, $\omega_\beta < \omega_\alpha$. □

Un mot de prudence à propos du lemme 3.1.4 : toute tentation d'aboutir à une inégalité stricte est vaine. En effet, il suffit de considérer la borne supérieure de la suite

$$\begin{aligned} \alpha_0 &= \omega_\gamma \\ \alpha_{i+1} &= \omega_{\alpha_i} \quad (i \in \mathbb{N}) \end{aligned}$$

avec γ un ordinal quelconque.

Proposition 3.1.5 *Tout cardinal infini s'écrit sous la forme ω_α pour un certain ordinal α.*

Preuve. Soit κ un cardinal. D'après le lemme 3.1.4, $\kappa \leq \omega_\kappa$. Alors $\kappa < \omega_{\kappa+1}$ d'après la définition 3.1.1. Il suffit alors de montrer que pour tout ordinal α et cardinal infini κ, si $\kappa < \omega_\alpha$ alors il existe un ordinal $\beta < \alpha$ tel que $\kappa = \omega_\beta$. Il n'y a rien à faire si $\alpha = 0$ (si ce raisonnement est trop inquiétant, on peut toujours commencer par $\alpha = 1$.). Si $\alpha = \beta + 1$ et que $\kappa < \omega_\alpha$, alors $\kappa \leq \omega_\beta$ parce que $\omega_\alpha = h(\omega_\beta)$. Si $\kappa = \omega_\beta$, il n'y a rien à faire ; sinon, alors l'hypothèse d'induction permet de conclure. Si α est un ordinal limite, alors il existe $\beta < \alpha$ tel que $\kappa < \omega_\beta$. On applique l'hypothèse d'induction à ω_β. □

Il existe une autre notation très répandue qui sera utilisée aussi. Pour tout ordinal α on utilise $\aleph_\alpha$ pour noter ω_α. Ce n'est pas une complication inutile. En effet, nous utiliserons exclusivement les $\aleph$ pour souligner le caractère "cardinal" de ces objets tandis que la notation ω sera restreinte à l'usage "ordinal". En arithmétique des cardinaux, cette différenciation devient très utile car les opérations arithmétiques, quoique notées par les mêmes symboles, ont des comportements différents des opérations arithmétiques pour les ordinaux.

Il convient d'introduire à ce point un nouvel usage des mots "successeur" et "limite". Avant de donner la définition correspondante, nous voulons souligner qu'un cardinal est toujours un ordinal limite.

Définition 3.1.6 *Un cardinal de la forme $\aleph_\alpha$ est dit un* cardinal successeur *(resp. un* cardinal limite *si α est un ordinal successeur (resp. ordinal limite).*

3.2 Opérations sur les cardinaux

Dans cette section nous introduirons les trois opérations arithmétiques sur les cardinaux :la somme, le prouit et l'exponentiation. Dans un premier temps la somme et le produit ne feront intervenir qu'un nombre finis de cardinaux. Plus tard, nous considérerons les sommes et les produits infinis.

Définition 3.2.1 *Soient $n \in \mathbb{N}^*$, $\{\kappa_i | 1 \leq i \leq n\}$ et $\{X_i | 1 \leq i \leq n\}$ n ensembles deux à deux disjoints tels que $|X_i| = \kappa_i$ pour tout $i \in \{1, \ldots, n\}$. Alors la somme de ces cardinaux est définie et notée comme suit :*

$$\kappa_1 + \ldots + \kappa_n = \sum_{i=1}^{n} \kappa_i = |X_1 \sqcup \ldots \sqcup X_n| \ .$$

Avec les mêmes hypothèses, on définit et on note le produit de ces cardinaux comme suit :

$$\kappa_1 . \ldots . \kappa_n = \prod_{i=1}^{n} \kappa_i = |X_1 \times \ldots \times X_n| \ .$$

Soient κ_1 et κ_2 deux cardinaux. Si X_1 et X_2 sont deux ensembles de cardinaux κ_1 et κ_2 respectivement, alors $\kappa_1^{\kappa_2} = |X_1^{X_2}|$ où $X_1^{X_2}$ est l'ensemble de toutes les fonctions de X_2 vers X_1.

Il faut vérifier que ces opérations sont bien définies, en d'autres termes, qu'elles ne dépendent pas du choix des ensembles X_i.

Lemme 3.2.2 *Les trois opérations introduites dans la définition 3.2.1 sont indépendantes du choix d'ensembles représentatifs.*

Preuve. Pour la somme et le produit, il suffit d'étudier les opérations sur deux cardinaux. Soient donc $X_1,\ X_2, X_1', X_2'$ des ensembles tels que $|X_1| = |X_1'|$ et que $|X_2| = |X_2'|$. Si les équipotences s'expriment par les bijections $f_1 : X_1 \longrightarrow X_1'$ et $f_2 : X_2 \longrightarrow X_2'$, alors

$$\begin{array}{rccc} f & : & X_1 \sqcup X_2 & \longrightarrow & X_1' \sqcup X_2' \\ & & x & \longmapsto & \left\{ \begin{array}{ll} f_1(x) & \text{si} \quad x \in X_1 \\ f_2(x) & \text{si} \quad x \in X_2 \end{array} \right. \end{array}$$

est une bijection entre $X_1 \sqcup X_2$ et $X_1' \sqcup X_2'$.

Pour le produit, on utilise la correspondance $(x_1, x_2) \longmapsto (f_1(x_1), f_2(x_2))$.

Traitons maintenant le cas de l'exponentiation. A chaque fonction h de X_1 vers X_2, on associe l'application $h' : X_1' \longrightarrow X_2'$ qui fait correspondre à chaque paire $(x, h(x))$ la paire $(f_1(x), f_2(h(x)))$. □

Voici quelques propriétés élémentaires dont les preuves sont des exercices d'entrainement :

Lemme 3.2.3

1. *L'addition et le produit sont commutatifs, associatifs.*
2. *Le produit est distributif sur la somme.*
3. *Pour tout* $n \in \mathbb{N}$, $\aleph_0 + n = \aleph_0$ *;*
4. $\aleph_0 + \aleph_0 = \aleph_0$.
5. $\aleph_0.\aleph_0 = \aleph_0$.

Le résultat suivant que nous montrerons est un théorème fondamental qui généralise le point(4) du lemme précédent et dont un corollaire est une généralisation du point (6) du même lemme.

Théorème 3.3 $\aleph_\alpha.\aleph_\alpha = \aleph_\alpha$ *pour tout ordinal* α.

Preuve. La stratégie de la preuve est la construction d'un bon ordre $\prec$ sur l'ensemble $\omega_\alpha \times \omega_\alpha$ pour tout ordinal α, et utiliser le théorème 1.6 pour montrer que $|\omega_\alpha \times \omega_\alpha| \leq \omega_\alpha$. Ceci impliquera que $\aleph_\alpha.\aleph_\alpha \leq \aleph_\alpha$. L'autre inégalité étant claire (pourquoi ?), le théorème 3.2 permettra de conclure.

La première étape est la définition d'une relation binaire $\prec$ sur l'ensemble $\omega_\alpha \times \omega_\alpha$ et la vérification qu'il s'agit d'un bon ordre. L'ordre est réminiscent de l'ordre lexicographique mais il faut introduire un nouvel ingrédient dont vous pouvez motiver la définition en utilisant l'analogie au premier quart du plan cartésien.

Soient (α_1, α_2) et (β_1, β_2) deux paires d'ordinaux. On écrit

$$(\alpha_1, \alpha_2) \prec (\beta_1, \beta_2)$$

si et seulement si

$$\max\{\alpha_1, \alpha_2\} < \max(\beta_1, \beta_2)$$

ou

$$\max\{\alpha_1, \alpha_2\} = \max\{\beta_1, \beta_2\} \text{ et } \alpha_1 < \beta_1$$

ou

$$\max\{\alpha_1, \alpha_2\} = \max\{\beta_1, \beta_2\} \text{ et } \alpha_1 = \alpha_2 \text{ et } \beta_1 < \beta_2 \ .$$

Le soin de vérifier que $\prec$ est une relation d'ordre est laissé à nos lecteurs. Vérifions qu'il s'agit en fait d'un bon ordre. Soit alors X un ensemble non vide de paires d'ordinaux. Soit δ le plus petit ordinal de l'ensemble

$$\{\max\{\alpha, \beta\} | (\alpha, \beta) \in X\} \ .$$

Notons que c'est un ensemble (l'axiome de compréhension) et qu'il est non vide. Donc δ existe. Ensuite on définit l'ensemble

$$D = \{(\alpha, \beta) \in X | \max\{\alpha, \beta\} = \delta\} \ .$$

C'est un ensemble non vide. Pour toutes paires $(\alpha, \beta) \in D$ et $(\alpha_1, \beta_1) \in X \setminus D$, $(\alpha, \beta) \prec (\alpha_1, \beta_1)$. Donc il suffit de vérifier que D a un plus petit élément par rapport à $\prec$. Pour ce faire, on suit l'approche la plus naturelle. On "minimise" d'abord par rapport à la première coordonnée, ensuite par rapport à la deuxième. A chaque étape les opérations fournissent des ensembles non vides. La mise en place de ce procédé est un bon exercice.

Pour finir la preuve, il suffira de montrer que $|\omega_\alpha \times \omega_\alpha| \leq \aleph_\alpha$. Le raisonnement se fait en utilisant le Principe d'Induction Transfinie. Le lemme 3.2.3 (5) permet d'amorcer le raisonnement. Soit alors $0 < \alpha$ et supposons que $\aleph_\beta.\aleph_\beta \leq \aleph_\beta$ pour tout $\beta < \alpha$. L'idée principale est de comparer les segments initiaux de la structure bien ordonnée $(\omega_\alpha \times \omega_\alpha, \prec)$ à ceux du bon ordre $(\omega_\alpha, <)$. Fixons une paire arbitraire $(\alpha_1, \alpha_2) \in \omega_\alpha \times \omega_\alpha$. Nous définissons

$$S = \{(\xi_1, \xi_2) \in \omega_\alpha \times \omega_\alpha | (\xi_1, \xi_2) \prec (\alpha_1, \alpha_2)\} \ .$$

C'est le segment initial dont la borne supérieure est (α_1, α_2). Posons $\beta = \max\{\alpha_1, \alpha_2\}+1$. Comme ω_α est un ordinal limite, $\beta < \omega_\alpha$. Pour toute paire $(\xi_1, \xi_2) \in S$, $\max\{\xi_1, \xi_2\} \leq \max\{\alpha_1, \alpha_2\} < \beta$. En conséquence, $\xi_1, \xi_2 \in \beta$. Il en découle que $S \subseteq \beta \times \beta$.

Comme ω_α est un cardinal et que $\beta \in \omega_\alpha$, la proposition 3.1.5 montre qu'il existe un ordinal $\gamma < \alpha$ tel que $|\beta| = \aleph_\gamma$. D'après l'hypothèse d'induction, $\aleph_\gamma.\aleph_\gamma \leq \aleph_\gamma$. On en déduit que

$$|S| \leq |\beta \times \beta| = \aleph_\gamma.\aleph_\gamma \leq \aleph_\gamma \ .$$

Si $\aleph_\alpha < |\omega_\alpha \times \omega_\alpha|$ alors $\aleph_\alpha$ serat isomorphe à un segment initial de $\omega_\alpha \times \omega_\alpha$ (pourquoi ?). Or, comme S est un segment initial arbitraire de $\omega_\alpha \times \omega_\alpha$ par rapport à l'ordre $\prec$, le raisonnement précdent s'appliquerait alors $\aleph_\alpha$ aussi. De ceci découlerait alors que $\aleph_\alpha < \aleph_\alpha$, une contradiction. □

Corollaire 3.2.4 *1. Pour toute paire d'ordinaux α et β, $\aleph_\alpha.\aleph_\beta = \aleph_\beta$ si $\alpha \leq \beta$.*

2. Pour tout nombre naturel n non nul et pour tout ordinal α, $n.\aleph_\alpha = \aleph_\alpha$.

Preuve. Pour le premier point, la suite suivante d'inégalités donne la réponse :

$$\aleph_\beta \leq \aleph_\alpha.\aleph_\beta \leq \aleph_\beta.\aleph_\beta = \aleph_\beta \ .$$

Peut-être y a-t-il quelques points mineurs à vérifier. Remercions au moins Cantor et Bernstein.

La preuve du deuxième point suit un développement analogue. □

Corollaire 3.2.5 *1. Pour toute paire d'ordinaux α et β, $\aleph_\alpha + \aleph_\beta = \aleph_\beta$ si $\alpha \leq \beta$.*

2. Pour tout nombre naturel n et pour tout ordinal α, $n + \aleph_\alpha = \aleph_\alpha$.

Preuve. D'abord le premier point :

$$\aleph_\beta \leq \aleph_\alpha + \aleph_\beta \leq \aleph_\beta + \aleph_\beta = 2.\aleph_\beta = \aleph_\beta \ .$$

La preuve est similaire pour le deuxième point mais c'est plus distrayant de trouver une preuve directe en utilisant des bijections. □

Dans le reste de cette section, nous considérerons la somme et le produit d'une infinité de cardinaux. Les mêmes idées que celles du cas fini motivent les définitions ci-dessous, mais leurs mises en place nécessitent l'Axiome du choix. En effet, jusqu'à maintenant aucun raisonnement dans cette section n'a fait usage de l'Axiome du choix. Or, dans le cas infini et aussi dans la grande partie des propriétés de l'exponentiation des cardinaux que nous n'avons pas encore abordé, l'Axiome du choix fait son apparition dès les premiers pas.

Définition 3.2.6 *Soit $\{X_i : i \in I\}$ une famille d'ensembles deux à deux disjoints tels que $|X_i| = \kappa_i$. Alors la somme et le produit des κ_i sont respectivement définis de la façon suivante :*

$$\sum_{i \in I} \kappa_i = |\bigcup_{i \in I} X_i|$$

$$\prod_{i \in I} \kappa_i = |\prod_{i \in I} X_i| \ .$$

Comme dans le cas fini, la première chose à faire est de vérifier que ces deux définitions ne dépendent pas du choix des ensembles de cardinaux κ_i. Le lemme correspondant au lemme 3.2.2 nécessite l'Axiome du choix parce que les familles d'ensembles peuvent être infinies et qu'en conséquences il sera impossible de les "énumérer".

Lemme 3.2.7 *Soient $\{X_i : i \in I\}$ et $\{X'_i : i \in I\}$ deux familles d'ensembles tels que $|X_i| = |X'_i|$. Dans le cas de l'addition, on suppose aussi que les deux familles ne contiennent que des ensembles deux à deux disjoints. Alors*

$$|\bigcup_{i\in I} X_i| = |\bigcup_{i\in I} X'_i|$$

$$|\prod_{i\in I} X_i| = |\prod_{i\in I} X'_i|$$

Preuve. Pour chaque $i \in I$ il faut *choisir* une bijection $f_i : X_i \longrightarrow X'_i$. Une fois que c'est fait la première équipotence est assurée par la correspondance $x \longmapsto f_i(x)$ si et seulement si $x \in X_i$ tandis que pour la deuxième équipotence à $(x_i)_{i\in I}$ est associée la "suite" $(f_i(x_i))_{i\in I}$. □

Le comportement général de la somme est décrit dans les lemmes et la proposition suivants.

Lemme 3.2.8 *Si λ et κ sont deux cardinaux alors*

$$\sum_{i<\lambda} \kappa = \lambda.\kappa \ .$$

Preuve. Il suffit de construire une bijection entre $\bigcup_{i<\lambda} X_i$, où chaque X_i est un ensemble de cardinal κ_i, et $\lambda \times \kappa$. □

Lemme 3.2.9 *Soit $\{\kappa_\alpha | \alpha < \lambda\}$ une famille de cardinaux. Si $\kappa = \sup\{\kappa_\alpha | \alpha < \lambda\}$ alors*

$$\kappa \leq \sum_{\alpha<\lambda} \kappa_\alpha \ .$$

Preuve. Si $\delta < \kappa$, alors il existe $\beta < \lambda$ tel que $\delta < \kappa_\beta$. Or $\kappa_\beta \leq \sum_{\alpha<\lambda} \kappa_\alpha$. Alors $\delta \in \sum_{\alpha<\lambda} \kappa_\alpha$. □

Proposition 3.2.10 *Soient λ un cardinal infini et $\{\kappa_\alpha | \alpha < \lambda\}$ un ensemble de cardinaux non nuls. On pose $\kappa = \sup\{\kappa_\alpha | \alpha < \lambda\}$. Alors*

$$\sum_{\alpha<\lambda} \kappa_\alpha = \lambda.\kappa \ .$$

Preuve. Remarquons que $\sum_{\alpha<\lambda} \kappa_\alpha \leq \sum_{\alpha<\lambda} \kappa$. Le lemme 3.2.8 alors montre que $\sum_{\alpha<\lambda} \kappa_\alpha \leq \lambda.\kappa$. On applique le lemme 3.2.8 une deuxième fois pour conclure que

$$\lambda = \sum_{i<\lambda} 1 \leq \sum_{\alpha<\lambda} \kappa_\alpha \ .$$

Le lemme 3.2.9 montre alors que $\kappa \leq \sum_{\alpha<\lambda} \kappa_\alpha$. Il en découle que

$$\lambda.\kappa \leq \sum_{\alpha<\lambda} \kappa_\alpha . \sum_{\alpha<\lambda} \kappa_\alpha \ .$$

Or λ est un cardinal infini et les κ_α sont tous non nuls. En conséquence, le cardinal $\sum_{\alpha<\lambda} \kappa_\alpha$ est infini, et on peut appliquer le théorème 3.3 pour conclure que

$$\lambda.\kappa \leq \sum_{\alpha<\lambda} \kappa_\alpha . \sum_{\alpha<\lambda} \kappa_\alpha = \sum_{\alpha<\lambda} \kappa_\alpha \ .$$

La conclusion suit du théorème 3.2 (Cantor-Bernstein). □

Etudions certaines propriétés de base des produits et de l'exponentiation. Il existe un lemme analogue au lemme 3.2.8 qui lie le produit à l'exponentiation :

Lemme 3.2.11 *Si λ et κ sont deux cardinaux alors*

$$\prod_{i<\lambda} \kappa = \kappa^\lambda \ .$$

Preuve. La preuve découle de la définition générale de la notion de produit cartésien. En effet, l'ensemble $\prod_{i<\lambda} \kappa$ n'est que l'ensemble de toutes les fonctions de λ vers κ. □

Le lemme suivant montre que certaines propriétés bien connues de l'exponentiation des nombres naturels s'étendent au contexte des cardinaux infinis :

Lemme 3.2.12 *Soient κ, λ et μ trois cardinaux. Alors*

1. $\kappa^{\lambda+\mu} = \kappa^\lambda.\kappa^\mu$;
2. $(\kappa^\lambda)^\mu = \kappa^{\lambda.\mu}$.

Preuve. Pour le premier point, nous définissons une fonction de $\kappa^\lambda \times \kappa^\mu$ vers $\kappa^{\lambda+\mu}$ en associant à chaque paire de fonctions (f, g), avec $f : \lambda \longrightarrow \kappa$ et $g : \mu \longrightarrow \kappa$, la fonction $f \sqcup g$ qui associe à un élément de λ son image sous f et à un élément de μ son image sous g.

Pour le deuxième point, nous définissons une fonction qui associe à chaque fonction $f : \lambda \times \mu \longrightarrow \kappa$ une fonction $g : \mu \longrightarrow \kappa^\lambda$ avec g définie de la façon suivante :

$$\begin{array}{rccccc} \text{Pour tout } \alpha \in \mu, & g(\alpha) & : & \lambda & \longrightarrow & \kappa \\ & & & \beta & \longmapsto & f(\beta, \alpha) \ . \end{array}$$

La vérification que ce sont des bijections est un exercice. □

Puisque nous avançons vers le théorème de König, il convient de commencer à faire le lien avec le théorème fondateur 3.1 de Cantor. Voici un lemme simple mais important :

Lemme 3.2.13 *Si E est un ensemble tel que $|E| = \kappa$, alors $|\mathcal{P}(E)| = 2^\kappa$.*

Preuve. A chaque partie X de E, il suffit d'associer la fonction

$$\begin{array}{rccccc} f_X & : & E & \longrightarrow & \{0,1\} \\ & & x & \longmapsto & \left\{ \begin{array}{ll} 1 & \text{si } x \in X \\ 0 & \text{si } x \in E \setminus X \end{array} \right. \end{array}$$

□

Pour illustrer ces lemmes dans un contexte concret, calculons le produit $\prod_{i<\omega,\ i\neq 0} i$. D'un côté,

$$\prod_{i<\omega,\ i\neq 0} i \leq \prod_{i<\omega} \aleph_0 = \aleph_0^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} \overset{3.2.12}{=} 2^{(\aleph_0.\aleph_0)} = 2^{\aleph_0} \ .$$

De l'autre côté,

$$2^{\aleph_0} = \prod_{i<\omega} 2 \leq \prod_{i<\omega,\ i\neq 0} i \ .$$

Or, $2^{\aleph_0}$ est le cardinal de toutes les parties des nombres naturels (pourquoi ?), et d'après le théorème 3.1 $\aleph_0 < 2^{\aleph_0}$. Vu que la somme $\sum_{i<\omega} i = \aleph_0$, nous venons de voir une illustration des différences considérables entre les produits infinis et les sommes infinies.

Le théorème suivant est fondamental en ce qui concerne les liens (et la distance) entre la somme et le produit des cardinaux.

Théorème 3.4 (König) *Soient $\{\kappa_i | i \in I\}$ et $\{\lambda_i | i \in I\}$ deux familles de cardinaux indexées par le même ensemble I tels que $\kappa_i < \lambda_i$ pour tout $i \in I$. Alors*

$$\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i \ .$$

Notons immédiatement que ce théorème généralise le théorème 3.1. En effet, pour tout cardinal κ,

$$\kappa = \sum_{i < \kappa} 1 < \prod_{i < \kappa} = 2^{\kappa} \ .$$

L'importance du théorème de König va bien au delà de cette illustration. Nous en énoncerons une conséquence dans la section suivante.

Bien sûr, il faut donner une preuve du théorème 3.4. Nous avons préparé le terrain. Le reste du travail sera fait par nos lecteurs pendant leurs travaux dirigés.

3.3 Cardinaux réguliers et singuliers ; cofinalité

Cette section sera brève. Nous nous contenterons de certaines définitions et caractérisations qui nous seront nécessaires au quatrième chapitre. Ensuite, nous énoncerons une conséquence du théorème 3.4 sur la "cofinalité du continu". Ce sera une occasion de faire le lien avec l'Hypothèse du Continu et la complexité de l'exponentiation des cardinaux.

Nous commençons avec les notions très importantes de *cardinal régulier* et de *cardinal singulier*. Pour motiver, faisons la remarque suivante qui est bien évidente : l'ordinal ω, en d'autres termes l'ensemble des nombres naturels, ne peut pas s'écrire comme une union finie de ses parties finies. Une autre manière équivalente d'exprimer ce même fait est de dire que ω n'est pas la borne supérieure d'une suite strictement croissante et finie de nombres naturels. Que de manières pour parler d'une évidence si banale... Pourtant ce n'est ni une évidence ni une banalité quand on monte dans la hiérarchie des cardinaux. Il suffit de penser à

$$\aleph_\omega = \bigcup_{i < \omega} \aleph_i \ .$$

Le lemme suivant est pertinent :

Lemme 3.3.1 *Soit κ un cardinal. Alors les énoncés suivants sont équivalents :*

1. *Pour toute partie X de κ de cardinal strictement inférieur à κ, $\sup(X) < \kappa$.*
2. *Pour tout ordinal $\lambda < \kappa$, et toute famille $\{X_\alpha | \alpha \in \lambda\}$ de parties de κ telles que $|X_\alpha| < \kappa$, l'union $\bigcup_{\alpha < \kappa} X_\alpha$ est de cardinal strictement inférieur à κ.*

Proof. C'est un bon exercice que nos lecteurs aborderont aux travaux dirigés. □

Définition 3.3.2 *Un cardinal κ est dit* régulier *s'il vérifie les deux conditions équivalentes du lemme 3.3.1. Un cardinal qui n'est pas régulier est dit* singulier.

Exemple 3.3.1 Les nombres finis, $\aleph_0$ sont des cardinaux réguliers comme nous l'avons déjà constaté. Plus généralement, tous les cardinaux successeurs sont réguliers.

Du côté des singuliers, il n'en manque pas. En fait, tous les cardinaux limites que vous pouvez imaginer sont singuliers. S'il en existe quelques'uns que vous ne pouvez pas imaginer est indépendant des axiomes usuels de la théorie des ensembles, c'est à dire ZFC.

Les notions de régularité et de singularité s'expriment d'une manière plus efficace à l'aide de la notion de *cofinalité* dont nous parlerons maintenant. Si α est un ordinal, il existe toujours un ordinal $\beta \leq \alpha$ et une fonction $f : \beta \longrightarrow \alpha$ strictement croissante dont l'image n'est pas majorée dans α. On pourrait poser $\beta = \alpha$, et f serait l'application identité. Il n'est pas nécessaire que ce soit la seule possibilité, il s'agit quand-même d'un point de départ pour trouver un ordinal "canonique" qui aurait cette propriété :

Définition 3.3.3 *Soit α un ordinal. La* cofinalité *de α est le plus petit ordinal β tel qu'il existe une fonction $f : \beta \longrightarrow \alpha$ strictement croissante dont l'image est non majorée dans α. On note $\beta = \mathrm{cof}(\alpha)$.*

Voici une propriété fondamentale qui fait de la cofinalité une notion robuste :

Lemme 3.3.4 *Pour tout ordinal α, $\mathrm{cof}(\alpha)$ est un cardinal.*

Preuve. Un bon exercice... □

La proposition suivante lie la cofinalité à l'étude de la regularité :

Proposition 3.3.5

1. *Pour tout ordinal α, $\mathrm{cof}(\mathrm{cof}(\alpha)) = \mathrm{cof}(\alpha)$. En particulier, $\mathrm{cof}(\alpha)$ est un cardinal régulier.*
2. *Un cardinal κ est régulier si et seulement si $\mathrm{cof}(\kappa) = \kappa$.*

Preuve. Un autre bon exercice... □

Les trois notions introduites dans cette section sont fondamentales dans l'étude des cardinaux et de leur arithmétique. En particulier, les résultats sur l'exponentiation des cardinaux dépendent souvent de la régularité des cardinaux en question. Il est difficile d'arriver à des conclusions valables pour tous les cardinaux sans faire des hypothèses fortes dont la principale est l'Hypothèse du Continu. Dans le reste de cette section, nous fournirons quelques détails sur ce sujet.

Le langage introduit dans ce chapitre nous permet de remplacer l'énoncé du théorème 3.1 par un énoncé arithmétique qui est le suivant :

$$\text{pour tout ordinal } \alpha,\ \ 2^{\aleph_\alpha} \geq \aleph_{\alpha+1}\ .$$

L'Hypothèse du Continu Généralisée énonce qu'on a égalité. L'Hypothèse du Continu tout court s'adresse au cas où $\alpha = 0$. C'est un énoncé indépendant des autres axiomes de la théorie des ensembles. En fait, des techniques qui vont bien au delà de ce cours permettent de montrer que pour un grand choix de $\aleph_\beta$, il existe un "univers" (un terme un peu vague qui deviendra clair à partir du chapitre 5) où $2^{\aleph_\alpha} = \aleph_\beta$. Les restrictions, assez rares, s'expriment en fonction de la cofinalité :

Proposition 3.3.6 *Pour tout cardinal κ, $\mathrm{cof}(2^\kappa) > \kappa$.*

Preuve. C'est l'une des applications principales du théorème 3.4. Encore un autre bon exercice... □

Pour finir, citons le corollaire suivant pour illustrer comment la proposition 3.3.6 limite les choix de valeur pour l'exponentiation :

Corollaire 3.3.7 $2^{\aleph_0} \neq \aleph_\omega$.

Chapitre 4

Des ensembles aux groupes

Dans ce chapitre, nous démontrerons d'abord un théorème classique de la combinatoire infinie, à savoir le théorème de Fodor. Les notions et les raisonnements de la preuve de ce résultat sont omniprésents en théorie des ensembles, et plus généralement en toute branche des mathématiques ayant recours à la combinatoire des ensembles infinis.

La seconde moitié du chapitre illustrera des liens avec l'algèbre, plus précisément avec la théorie des groupes. Nous démontrerons un théorème de Simon Thomas sur les automorphismes de certaines classes des groupes infinis.

4.1 Les ensembles fermés et non bornés ; les ensembles stationnaires

Dès le début, fixons une notation qui, sauf mention contraire, sera valable jusqu'à la fin de cette section : κ est un cardinal *non* dénombrable et régulier.

Définition 4.1.1 *Soit $C \subset \kappa$:*

1. *C est dit* fermé *si pour tout ordinal limite $\lambda < \kappa$ et toute suite d'éléments de C de la forme*

$$\alpha_0 < \alpha_1 < \ldots < \alpha_\xi < \ldots \quad (\xi < \lambda) \ ,$$

   $\sup\{\alpha_\xi | \xi < \lambda\} \in C$.

2. *C est dit* non borné *si pour tout élément $\alpha \in \kappa$, il existe $\beta \in C$ tel que $\alpha < \beta$.*

Nous étudierons certaines propriétés de ensembles fermés et non bornés de κ. Dans la langue de Shakespeare telle qu'elle est pratiquée par les théoriciens d'ensembles, les synonymes de "fermé" et de "non borné" sont "closed" et "unbounded" respectivement. Comme dans ce contexte, les ensembles qui jouissent des deux propriétés simultanément jouent un rôle important, le mot "club" est assez courant. Pour simplifier l'usage nous adopterons le même mot [1].

Exemple 4.1.1
(1) Le cardinal κ lui-même est un club.
(2) $\{\alpha < \kappa | \alpha$ est un ordinal limite$\}$ est un club.
(3) Pour tout $\alpha < \kappa$, l'ensemble $\{\beta < \kappa | \alpha \leq \beta\}$ est un club.

Le lemme suivant sera vite dépassé mais sa preuve donne des indications éclairantes pour la proposition qui vient après.

Lemme 4.1.2 *Si C et D sont des clubs, alors $C \cap D$ est un club.*

[1] Depuis, nous avons appris que le bon mot est "clos cofinal". Merci Julien.

Preuve. La propriété d'être fermé est claire. Montrons que l'ensemble $C \cap D$ est non borné dans κ. Pour ce faire, soit $\alpha \in \kappa$. On peut construire une suite

$$\alpha < \beta_0 < \beta_1 < \beta_2 < \ldots < \beta_{2k} < \beta_{2k+1} < \ldots \quad (k < \omega) \ ,$$

dans laquelle $\{\beta_{2k} | k < \omega\} \subset C$ et $\{\beta_{2k+1} | k < \omega\} \subset D$ en utilisant l'hypothèse C et D soient non bornés. Alors ses deux suites extraites ont même borne supérieure qui appartient à $C \cap D$. □

Notons que le lemme 4.1.2 est suffisant pour conclure que l'ensemble des clubs sur κ est un *filtre*. En effet, cet ensemble contient κ tandis qu'il ne contient pas $\emptyset$ et il est stable par rapport aux intersections finies d'après le lemme 4.1.2. En général, ce n'est pas un *ultrafiltre*.

La proposition suivante montre que l'ensemble des clubs sur κ vérifie une propriété de stabilité beaucoup plus forte par rapport aux intersections :

Proposition 4.1.3 *Soient λ un cardinal strictement inférieur à κ et $\{C_\xi | \xi < \lambda\}$ une famille de clubs sur κ. Alors $\bigcap_{\xi<\lambda} C_\xi$ est un club.*

Preuve. La propriété d'être fermé est évidente. Vérifions la propriété de ne pas avoir de bornes en utilisant un raisonnement qui généralise la construction de suite du lemme 4.1.2.

Soit $\alpha \in \kappa$. On construira par récurrence transfinie, une suite strictement croissante

$$\{\gamma_{i,j} | i < \lambda, \ j < \omega; \ \gamma_{i,j} \in C_i \text{ pour tout } i < \lambda\}$$

d'éléments des C_ξ dont la borne supérieure appartiendra à l'intersection $\bigcap_{\xi<\lambda} C_\xi$. L'ordre est en fait déterminé par la relation suivante entre deux termes de la suite :

$$\gamma_{i,j} < \gamma_{i',j'} \text{ si et seulement si } \begin{cases} j < j' \\ \text{ ou} \\ j = j' \text{ et } i < i' \end{cases}$$

Pour commencer on fixe $\gamma_{0,0} \in C_0$ tel que $\alpha < \gamma_{0,0}$ en utilisant l'hypothèse que C_0 ne soit pas borné. Ensuite, on forme la suite strictement croissante

$$\gamma_{0,0} < \gamma_{1,0} < \ldots < \gamma_{i,0} < \ldots \quad (i < \lambda) \ ,$$

telle que $\gamma_{i,0} \in C_i$ pour chaque $i < \gamma$.

On pose $\gamma_1 = \sup\{\gamma_{i,0} | i < \lambda\}$. Supposons maintenant que γ_n soit calculé (γ_0 étant α). Alors, la suite suivante strictement croissante est construite :

$$\gamma_{0,n} < \gamma_{1,n} < \ldots < \gamma_{i,n} < \ldots \quad (i < \lambda)$$

avec cette fois $\gamma_n < \gamma_{0,n}$ et $\gamma_{i,n} \in C_i$. Cette étape de la construction est possible pour deux raisons. D'un côté, κ est un cardinal régulier strictement supérieur à λ ce qui implique que $\gamma_n < \kappa$. De l'autre côté, les C_i sont des clubs, plus précisément, des ensembles non borné dans κ. On pose à la fin de cette étape $\gamma_{n+1} = \sup\{\gamma_{i,n} | i < \lambda\}$.

Par construction, pour tout $n < \omega$ et $i < \lambda$, $\gamma_n < \gamma_{i,n} < \gamma_{n+1}$. Il en découle que

$$\sup\{\gamma_n | n < \omega\} = \sup\{\gamma_{i,n} | n \in \omega\}$$

pour tout $i < \lambda$. La règle des gendarmes... Or, chaque C_i est un ensemble fermé, donc

$$\sup\{\gamma_{i,n} | n \in \omega\} \in C_i$$

pour tout i. De manière équivalente, $\sup\{\gamma_n | n < \omega\} \in \cap_{i<\lambda} C_i$. Cette borne finale majore α aussi. □

Pour arriver à une nouvelle conclusion sur les intersections, il faudra considérer les *intersections diagonales* :

Définition 4.1.4 *Soit $\{X_\alpha|\alpha<\kappa\}$ une famille de sous-ensembles de κ. L'intersection diagonale des X_α est définie de la manière suivante*

$$\Delta X_\alpha = \{\xi<\kappa|\xi\in\bigcap_{\alpha<\xi} X_\alpha\}\ .$$

Proposition 4.1.5 *Soit $\{C_\alpha|\alpha<\kappa\}$ une famille de clubs dans κ. Alors ΔC_α est un club.*

Preuve. Contrairement aux deux énoncés précédents sur les intersections, il faut raisonner plus pour vérifier que ΔC_α est un ensemble fermé. Soit $(\alpha_i)_{i<\lambda}$ une suite strictement croissante d'éléments de ΔC_α. Par définition, chaque α_i vérifie la condition suivante :

$$\alpha_i\in\cap_{\gamma<\alpha_i}C_\gamma\ .$$

Posons $\sigma=\sup\{\alpha_i|i<\lambda\}$. Il suffit de vérifier que pour tout $\beta<\sigma$, $\sigma\in C_\beta$. D'après la définition de sup, il existe $i_0<\lambda$ tel que $\beta<\alpha_{i_0}$. Comme chaque élément de notre suite croissante appartient à ΔC_α, d'après la définition de celle-ci, si $i_0<i$ alors $\alpha_i\in C_\beta$. Or $\sigma=\sup\{\alpha_i|i_0<i<\lambda\}$ aussi. Comme C_β est un ensemble fermé et qu'il contient tout α_i vérifiant $i_0<i$, il en découle que $\sigma\in C_\beta$.

Maintenant, vérifions que ΔC_α est non bornée. Soit $\gamma<\kappa$. Nous construirons par récurrence une suite de longueur ω. On pose $\alpha_0=\gamma$. Supposons α_i défini avec $i\in\omega$ et procédons à la définition de α_{i+1}. D'après le lemme 4.1.3, l'intersection $\bigcap_{\alpha<\alpha_i}C_\alpha$ est un club. Alors, il existe $\alpha_{i+1}\in\bigcap_{\alpha<\alpha_i}C_\alpha$ tel que $\alpha_i<\alpha_{i+1}$. La construction est terminée.

Soit maintenant $\sigma=\sup\{\alpha_i|i<\omega\}$. Montrons que $\sigma\in\Delta C_\alpha$. Le raisonnement est similaire à ce que nous avons utilisé pour vérifier la propriété de fermeture. Il suffira de vérifier que $\sigma\in C_\beta$ pour tout $\beta<\sigma$. Fixons $\beta\in\sigma$. Il existe alors $i_0\in\omega$ tel que $\beta<\alpha_{i_0}$. Pour tout i tel que $i_0<i$, $\alpha_i\in C_\beta$ d'après la construction de la suite $(\alpha_i)_{i<\omega}$. Par ailleurs, $\sigma=\sup\{\alpha_i|i_0<i\}$ aussi. Comme C_β est un ensemble fermé, on déduit que $\sigma\in C_\beta$. □

Définition 4.1.6 *Une partie S de κ est dite* stationnaire *si $S\cap C\neq\emptyset$ pour tout club C sur κ.*

Exemple 4.1.2

(1) Tout club est un ensemble stationnaire.

(2) Si S est stationnaire et que C est un club alors $S\cap C$ est stationnaire.

(3) $S=\{\alpha<\omega_2|\mathrm{cof}(\alpha)=\omega\}$ est stationnaire dans ω_2 mais n'est pas un club. Pour vérifier que S est un ensemble stationnaire, il suffit de vérifier que chaque club contient un ordinal de cofinalité ω. Or, pour ce faire il suffit de considérer les ω premiers éléments d'un club. Si ces éléments sont notés α_i $(i<\omega)$, alors $\sup\{\alpha_i|i<\omega\}$ appartient au club et est de cofinalité ω. Pour vérifier que S n'est pas un club il suffit de considérer une suite strictement croissante dans S de longueur ω_1. Comme S est défini sur ω_2, il y a suffisamment d'espace pour une telle suite. En effet, il existe $\aleph_1$ ordinaux limites dénombrables, donc de cofinalité ω. Leur borne supérieure est ω_1 qui n'est pas de cofinalité ω.

Définition 4.1.7 *Soit E un ensemble d'ordinaux et f une fonction à valeurs ordinales définie sur E. Alors f est dite régressive sur E si pour tout $\alpha\in E$ et $\alpha\neq 0$, $f(\alpha)<\alpha$.*

Exemple 4.1.3 Nous donnons un exemple très simple pour illustrer la définition 4.1.7. Des exemples plus compliqués mais pertinents seront rencontrés lors des théorèmes qui suivent. Posons $S=\{\alpha<\kappa|\alpha$ est un ordinal successeur.$\}$. Alors la fonction suivante est régressive :

$$\begin{array}{rccc} f\ : & S & \longrightarrow & \kappa \\ & \alpha+1 & \longmapsto & \alpha \end{array}\ .$$

Théorème 4.1 (Fodor) *Si $S \subset \kappa$ est stationnaire et que $f : S \longrightarrow \kappa$ est une fonction régressive sur S, alors il existe un sous-ensemble stationnaire $T \subset S$ et $\gamma < \kappa$ tels que $f(\alpha) = \gamma$ pour tout $\alpha \in T$.*

Preuve. La preuve est par l'absurde. Sous l'hypothèse contradictoire, pour tout γ l'ensemble $\{\alpha \in S | f(\alpha) = \gamma\}$ n'est pas stationnaire, en d'autres termes, il existe un club C_γ tel que

$$(*) \qquad f^{-1}(\gamma) \cap C_\gamma = \emptyset \ .$$

Considérons ΔC_γ qui est aussi un club. Alors $\Delta C_\gamma \cap S \neq \emptyset$. Soit alors $\alpha \in \Delta C_\gamma \cap S$. Alors $f(\alpha) < \alpha$. D'après la définition de ΔC_γ, $\alpha \in \bigcap_{\gamma<\alpha} C_\gamma$. En particulier, $\alpha \in C_{f(\alpha)}$ puisque par hypothèse $f(\alpha) < \alpha$. Cette conclusion contredit $(*)$ quand $\gamma = f(\alpha)$. □

4.2 Tours d'automorphismes

Cette section est consacrée à certains groupes et leurs groupes d'automorphismes. Commençons par le rappel de quelques connaissances générales en théorie des groupes. Dans tout groupe G, il existe des sous-groupes qui y sont naturellements définis quelles que soit les propriétés particulières du groupe ambiant. Le *centralisateur* d'un élément g de G en est un :

$$C_G(g) = \{x \in G | xg = gx\} \ .$$

Le *centre* du groupe G peut se définir à partir des centralisateurs :

$$Z(G) = \{g \in G | C_G(g) = G\} \ .$$

Une autre manière moins encombrée de définir le centre est le suivant :

$$Z(G) = \{x \in G | \forall y(xy = yx)\} \ .$$

Le groupe G est dit *commutatif* ou *abélien* si $G = Z(G)$, de manière équivalente, si G vérifie la condition $\forall x \forall y(xy = yx)$ (propriétés définissables dans le langage des groupes...). Finalement, on peut généraliser la définition des centralisateurs à des parties quelconques de G : soit $X \subseteq G$, alors

$$C_G(Y) = \cap_{y \in Y} C_G(y) \ .$$

On dira qu'une partie X de G centralise une autre partie de Y si $X \subseteq C_G(Y)$.

Comme toute structure, un groupe G a des automorphismes. Ceux-ci forment un groupe quand ils sont munis de la composition usuelle des fonctions. Nous noterons ce "nouveau" groupe $\mathrm{Aut}(G)$. Dépendant de sa structure particulière G peut avoir des automorphismes dont d'autres groupes ne posséderaient pas d'analogues mais il existe au moins un ensemble d'automorphismes de G qu'on peut définir pour tous les groupes, à savoir les *automorphismes intérieurs*. Ce sont les automorphismes induits par la *conjugaison* par un élément fixé du groupe G : si $g \in G$ alors on définit

$$\begin{array}{rccc} i_g & : & G & \longrightarrow & G \\ & & x & \longmapsto & gxg^{-1} \end{array} \ .$$

Il est facile de vérifier que ce sont en effet des automorphismes et qu'ils forment un sous-groupe de $\mathrm{Aut}(G)$. Ce sous-groupe sera noté $\mathrm{Int}(G)$. Le sous-groupe $\mathrm{Int}(G)$ a une autre propriété importante. Il est *distingué* dans $\mathrm{Aut}(G)$ ce qui sera noté

$$\mathrm{Int}(G) \lhd \mathrm{Aut}(G) \ .$$

En général, rappelons qu'un sous-groupe H d'un groupe G est dit distingué si pour tout $g \in G$, $gHg^{-1} = H$, en d'autres termes, si et seulement si $i_g(H) = H$.

Comme pour toute paire d'éléments $g_1, g_2 \in G$, $i_{g_1 g_2} = i_{g_1} i_{g_2}$, l'application suivante est un homomorphisme de G vers $\mathrm{Aut}(G)$:

$$\begin{array}{rccc} \iota & : & G & \longrightarrow & \mathrm{Aut}(G) \\ & & g & \longmapsto & i_g \end{array} .$$

L'image de G est exactement $\mathrm{Int}(G)$. Le noyau de cet homomorphisme, en d'autres termes, les éléments de G dont l'image est l'élément neutre de $\mathrm{Aut}(G)$, est exactement $Z(G)$. Ceci entraîne l'isomorphisme naturel suivant :

$$G/Z(G) \cong \mathrm{Int}(G) .$$

Dans cette section, nous étudierons les automorphismes d'un groupe "sans centre". Comme vous pouvez l'avoir deviné, il s'agit d'un groupe G tel que $Z(G) = 1$. Le résumé que nous avons donné de la théorie des groupes montre que sous l'hypothèse d'être sans centre, $G \cong \mathrm{Int}(G)$. En conséquence, G se plonge dans $\mathrm{Aut}(G)$ et on peut le considérer comme un sous-groupe de $\mathrm{Aut}(G)$. Voici un premier lemme qui est légèrement au delà des généralités que nous avons révisées jusqu'à maintenant :

Lemme 4.2.1 *Pour tout groupe G, si $Z(G) = 1$, alors $C_{\mathrm{Aut}(G)}(\mathrm{Int}(G)) = 1$. En particulier, $Z(\mathrm{Aut}(G)) = 1$.*

Un mot sur la notation : nous utiliserons 1 pour noter tout élément neutre aussi bien que pour tout groupe qui ne contient que l'élément neutre.

Preuve. Même, sans l'hypothèse de la trivialité du centre de G, si $\rho \in C_{\mathrm{Aut}(G)}(\mathrm{Int}(G))$, alors pour tout $g, x \in G$

$$\rho(x) = i_g \rho i_{g^{-1}}(x) = g\rho(g^{-1}xg)g^{-1} = g\rho(g)^{-1}\rho(x)\rho(g)g^{-1} .$$

Il en découle que

$$\rho(g)g^{-1}\rho(x) = \rho(x)\rho(g)g^{-1} .$$

Comme ρ est un automorphisme de G, cette égalité, valable pour tout x et tout g dans G, implique que $\rho(g)g^{-1} \in Z(G)$. A ce point, l'hypothèse $Z(G) = 1$ est fort utile pour conclure que $\rho = 1$. □

Le lemme 4.2.1 nous montre qu'on peut voir $\mathrm{Aut}(G)$ comme un sous-groupe de $\mathrm{Aut}(\mathrm{Aut}(G))$, et toujours selon le même lemme, ce dernier est sans centre, et on peut le plonger dans... Il s'agit donc d'une "tour" d'automorphismes dont le rez-de-chaussé est G : soit α un ordinal, alors

$$\begin{array}{rcl} G_0 & = & \mathrm{Int}(G) \\ G_{\alpha+1} & = & \mathrm{Aut}(G_\alpha) \\ G_\alpha & = & \bigcup_{\beta<\alpha} G_\beta \text{ si } \alpha \text{ est un ordinal limite.} \end{array}$$

La construction s'arrête s'il existe un ordinal α tel que $\mathrm{Int}(G_\alpha) = G_{\alpha+1}$. Pour éviter de compliquer la notation, nous écrirons plutôt $G_\alpha = G_{\alpha+1}$, et plus généralement, nous dirons que $G_\alpha \leq G_{\alpha+1}$ au lieu de $\mathrm{Int}(G_\alpha) \leq G_{\alpha+1}$.

Théorème 4.2 (Vielandt, 1939) *Si G est un groupe fini sans centre, la tour a un nombre fini d'étages. Plus précisément, Si G est un groupe fini, alors il existe un nombre fini α tel que $G_\alpha = G_{\alpha+1}$.*

Nous nous intéressons au cas où G est infini :

Théorème 4.3 (Simon Thomas, 1984) *Si G est un groupe infini sans centre, alors le nombre d'étages est strictement inférieur à $\left(2^{|G|}\right)^+$. Plus précisément, le plus petit ordinal λ tel que $G_\lambda = G_{\lambda+1}$ est strictement inférieur à $\left(2^{|G|}\right)^+$.*

En général, pour tout cardinal κ, κ^+ est le cardinal successeur.

Nous aurons besoin d'une construction de sous-groupe bien connue. C'est le *normalisateur* d'un sous groupe : si G est un groupe et que H est un sous-groupe de G, alors le normalisateur de H dans G est

$$N_G(H) = \{g \in G | gHg^{-1} = H\} \ .$$

On dira qu'une partie X de G normalise H si $X \subseteq N_G(H)$.

Nous aurons besoin du lemme suivant :

Lemme 4.2.2 *Soient G un groupe sans centre, et α et β deux ordinaux.*

1. *Si $\alpha \leq \beta$, alors $C_{G_\beta}(G_\alpha) = 1$.*
2. *Si $\alpha < \beta$, alors $N_{G_\beta}(G_\alpha) = G_{\alpha+1}$.*

Preuve. (1) Il suffit de vérifier la conclusion quand $\alpha = 0$ puisque pour tout α, $G_\alpha \geq G_0$ et qu'en conséquence, $C_{G_\beta}(G_\alpha) \leq C_{G_\beta}(G_0)$. La démonstration se fait en utilisant l'induction transfinie sur β. Comme par hypothèse $Z(G) = 1$, le raisonnement est amorcé. Si β est un ordinal limite et que l'énoncé a été vérifié pour les ordinaux strictement inférieurs à β, alors il est vérifié pour β puisque $C_{G_\beta}(G_0) = \bigcup_{\lambda<\beta} C_{G_\lambda}(G_0) = 1$. Considérons le passage au successeur.

On pose $C = C_{G_{\beta+1}}(G_0)$. Supposons par l'absurde que $C \neq 1$. Notons que $C \cap G_\beta = 1$ puisque par récurrence $C_{G_\beta}(G_0) = 1$. Une autre remarque importante est que C ne centralise pas G_β puisque par construction $C_{G_{\beta+1}}(G_\beta) = 1$. Il existe donc un plus petit ordinal γ tel que G ne centralise pas G_γ. L'ordinal γ n'est pas nul puisque C centralise G_0. La construction de la suite tour d'automorphismes exclut la possibilité que γ soit un ordinal limite non nul. Il est donc de la forme $\delta + 1$ avec G_δ qui centralise C. Il s'ensuit de cette conclusion que $C \leq C_{G_{\beta+1}}(G_\delta)$. Or, $C_{G_{\beta+1}}(G_\delta) \leq C_{G_{\beta+1}}(G_0) = C$. Donc, $C = C_{G_{\beta+1}}(G_\delta)$ et, comme G_δ est normalisé par G_γ, C est normalisé par G_γ aussi. En conséquence, le groupe de commutateurs $[G_\gamma, C] = \langle g^{-1}c^{-1}gc | g \in G_\gamma,\ c \in C\rangle$ est contenu dans C. Or, $[G_\gamma, C]$ est contenu dans G_β aussi. En effet, comme $G_\gamma \leq G_\beta$ et que $G_\beta \lhd G_{\beta+1}$, $[G_\gamma, C] \leq [G_\beta, C] \leq G_\beta$. Donc $[G_\gamma, C] \leq G_\gamma \cap C = 1$. C'est une contradiction au choix de G_γ. Ainsi, $C_{G_{\beta+1}}(G_0) = 1$, et par le principe d'induction transfinie, $C_{G_\beta}(G_0) = 1$ pour tout ordinal β.

(2) Ce point se déduit rapidement du premier. En effet, par construction, $N_{G_\beta}(G_\alpha) \geq G_{\alpha+1}$. Par ailleurs, si $x \in N_{G_\beta}(G_\alpha)$, comme (1) a montré que $C_{G_\beta}(G_\alpha) = 1$, x induit un automorphisme de G_β. Or, par construction $G_{\alpha+1} = \mathrm{Aut}(G_\alpha)$, et il en résulte qu'il existe $y \in G_{\alpha+1}$ tel que y induit la même action sur G_α que x. Ainsi, $xy^{-1} \in C_{G_\beta}(G_\alpha) = 1$, et $x = y$. □

Maintenant, nous procédons à la preuve du théorème 4.3.

Preuve du Théorème 4.3. La première étape de la preuve est de borner le cardinal de chaque G_α dans la tour. En général, si G est un groupe et que H est un sous-groupe de G, alors la famille

$$\{gHg^{-1} | g \in G\}$$

des conjugués de H dans G est en bijection avec les classes de H dans G. En particulier, $[G : H] = |\{gHg^{-1} | g \in G\}|$, où $[G : H]$ est l'indice de H dans G. Nous appliquerons cette connaissance élémentaire à l'indice $[G_{\alpha+1} : G_1]$.

D'après le lemme 4.2.2 et le rappel que nous venons de faire, l'indice $[G_{\alpha+1} : G_1]$ est le cardinal de l'ensemble $\{gG_0g^{-1} | g \in G_{\alpha+1}\}$. Par ailleurs, tous ces conjugués sont contenus dans G_α puisque $G_0 \leq G_\alpha \lhd G_{\alpha+1}$. Ainsi, chaque conjugué gG_0g^{-1} est l'image dans G_α d'une injection de domaine G_0, et il en découle que

$$[G_{\alpha+1} : G_1] \leq |G_\alpha|^{|G|} \ .$$

Alors, comme $|\mathrm{Aut}(G)| \leq |G|^{|G|}$,

$$\begin{aligned} |G_{\alpha+1}| &= |G_1||G_\alpha|^{|G|} \\ &\leq |G|^{|G|}|G_\alpha|^{|G|} \\ &= |G_\alpha|^{|G|} \ . \end{aligned}$$

Avec cette estimation, en appliquant le principe d'induction transfinie, on montre que pour tout $\alpha < (2^{|G|})^+$, $|G_\alpha| \leq 2^{|G|}$. C'est un bon exercice d'entraînement pour nos lecteurs.

On pose $\lambda = (2^{|G|})^+$. Comme λ est un ordinal limite, $G_\lambda = \bigcup_{\alpha<\lambda} G_\alpha$. Par ailleurs, étant un cardinal sucesseur, λ est un cardinal régulier. En conséquence, l'ensemble

$$S = \{\alpha < \lambda | \mathrm{cof}(\alpha) = \mathrm{cof}(2^{|G|})\}$$

est un ensemble stationnaire (vérifiez !).

Par l'absurde, on supposera que pour tout ordinal $\alpha < \lambda$, $G_\alpha < G_{\alpha+1}$. Cette hypothèse permet de choisir pour tout $\alpha \in S$, $\pi_\alpha \in G_{\alpha+1} \setminus G_\alpha$. Notons aussi que, comme les éléments de S sont de cofinalité infinie, ils sont tous des ordinaux limites. Ainsi, pour tout $\alpha \in S$, $G_\alpha = \bigcup_{\beta<\alpha} G_\beta$.

Comme $G_0 \leq G_\alpha \lhd G_{\alpha+1}$, la dernière conclusion du paragraphe précédent montre que

$$\pi_\alpha G_0 \pi_\alpha^{-1} \leq \pi_\alpha G_\alpha \pi_\alpha^{-1} = G_\alpha = \bigcup_{\beta<\alpha} G_\beta \ .$$

Comme $\mathrm{cof}(\alpha) = \mathrm{cof}(2^{|G|}) > |G|$ d'après la proposition 3.3.6, il existe un ordinal $f(\alpha)$ strictement inférieur à α tel que $\pi_\alpha G_0 \pi_\alpha^{-1} \leq G_{f(\alpha)}$. En d'autres termes, la correspondance $\alpha \longmapsto f(\alpha)$ induit une fonction régressive de S vers λ. D'après le théorème 4.1, il existe $\gamma < \lambda$ et une partie stationnaire S_0 dans S tels que $f(S_0) = \{\gamma\}$.

Le reste du raisonnement utilise le constat que G_γ n'est pas suffisamment "spatieux". En effet, S_0 étant un ensemble stationnaire dans λ, il est de même cardinal que ce dernier. Alors, comme $\gamma < \lambda$, $|S_0| = (2^{|G|})^+ > 2^{|G|} = (2^{|G|})^{|G|} \geq |G_\gamma|^{|G|}$. Exprimé de manière informelle, il existe plus d'éléments dans S_0 qu'il y en a dans l'ensemble des applications de G vers G_γ.

Il s'ensuit de la conclusion du paragraphe précédent qu'il existe $\alpha < \beta \in S_0$, qui induisent de G_0 vers G_γ la même injection par conjugaison. Ainsi

$$\pi_\alpha G_0 \pi_\alpha^{-1} = \pi_\beta G_0 \pi_\beta^{-1} \ ,$$

ce qui équivaut à $\pi_\alpha^{-1}\pi_\beta \in N_{G_{\beta+1}}(G_0) = G_1$. Or $G_1 \leq G_\beta$ d'après le choix de β. Comme $\pi_\alpha \leq G_\alpha \leq G_\beta$, on conclut qu'il en est de même pour π_β. Or, par définition $\pi_\beta \notin G_\beta$. Cette contradiction finit la preuve. □

En 2006, Itay Kaplan et Saharon Shelah ont donné une nouvelle preuve du théorème 4.3. Cette nouvelle preuve minimise l'usage de l'axiome du choix.

Deuxième partie

Théorie des modèles

Chapitre 5

Notions de base

Nos objets principaux d'étude seront les *structures* et leurs *ensembles définissables*. Bien que les restrictions syntaxiques soient primordiales pour la notion de définissabilité, nous ne ferons pas la théorie des langages. La théorie des modèles s'intéresse aux structures, plus généralement aux classes de structures qui partagent certains aspects permettant d'obtenir des informations sur ces structures. Dans ce chapitre seront introduites les notions fondamentales de la théorie des modèles pour l'étude des structures.

5.1 Structures et notions liées

La notion de structure est bien naturelle pour tout mathématicien. Sa vie est consacrée à l'étude de certaines d'entre eux suivant une certaine optique. Sans attendre, définissons ce que nous en entendons :

Définition 5.1.1 *Une structure $\mathcal{M}$ est la donnée d'un ensemble sous-jacent M, dit* univers *ou* ensemble de base *muni*

1. *d'une famille $\{c_i | i \in I_C\}$, éventuellement vide, des éléments distingués de l'univers de $\mathcal{M}$, les* constantes *;*
2. *d'une famille $\{f_i | i \in I_F\}$, éventuellement vide, de* fonctions *chacune de M^{n_i} vers M avec $n_i \in \mathbb{N}^*$ qui ne dépend que de f_i $(i \in I_F)$;*
3. *d'une famille $\{R_i | i \in I_R\}$ de* relations*, chacune définie dans M^{k_i} avec $k_i \in \mathbb{N}^*$ qui ne dépend que de R_i $(i \in I_R)$.*

Les données des trois familles deux à deux disjointes d'indices I_C, I_F et I_R et d'une fonction qui associe à chaque indice le membre correspondant des familles $\{c_i | i \in I_C\}$, $\{(f_i, n_i) | i \in I_F\}$, $\{(R_i, k_i) | i \in I_R\}$ est la signature *de $\mathcal{M}$.*

Chaque nombre n_i (resp. k_i) est dit la arité *de la fonction (resp. la relation) indexé par i.*

Remarques/avertissements : 1. Dans la définition 5.1.1, nous n'avons pas dit que la famille des relations est "éventuellement vide". En effet, elle ne le sera jamais. L'égalité fera partie de toutes les structures que nous étudierons. C'est pour cela que nous omettrons sa mention.

2. Par abus, nous ne ferons pas de distinction entre la signature et son ensemble d'arrivée, en d'autres termes la famille indexée des constantes, relations et fonctions.

3. Notons aussi que les constantes ne sont que des fonctions 0-aires.

Une structure n'est donc pas qu'un ensemble. Elle est associée à une signature à laquelle peuvent être associée d'autres structures. C'est bien naturel. Il n'y a pas qu'un ensemble ordonné par $\in$. Pourtant, la signature décrit le cadre d'étude d'une structure. Illustrons avec quelques exemples :

Exemple 5.1.1 1. La première partie de ces notes est consacrée aux ensembles munis d'une relation d'ordre $\in$ avec certaines propriétés. La signature était $\{\in\}$, une relation binaire. C'est un exemple de *structure relationnelle*, en d'autres termes, une structure dont la signature ne contient pas de fonctions sauf éventuellement les constantes.

Parfois, nous nous sommes permis, comme par exemple dans la preuve du lemme 2.1.2, de distinguer certains ensembles finis à l'aide de leurs éléments. Il s'agissait d'augmenter la signature sans changer l'univers, un cas particulier de ce qu'on appelle une *expansion*.

2. Les nombres rationnels munis de leur ordre usuel est une structure que nous avons vue dès les premières heures. Sa signature $\{<\}$ est incluse dans la signature de toutes les structures ordonnées mais peut jouir des propriétés différentes suivant la structure. Pourtant, la propriété d'être muni d'une relation d'ordre est commune à chacune de ces structures.

3. Le corps des nombres réels $\mathbb{R}$ muni des fonctions binaires et unaires et de ces deux constantes $\{.,+,-,\ ^{-1},1,0\}$ est une structure algébrique dans la signature des corps. On peut associer au corps des réels des signatures plus réduites. On peut omettre des éléments de $\{-,\ ^{-1},1,0\}$. C'est un cas particulier d'une *réduction*. Bien que l'univers reste intact, la réduction, comme l'expansion, est un procédé qui peut avoir des effets importants sur la structure.

Dans une autre registre, le même univers $\mathbb{R}$ peut servir d'univers à une structure qui n'a a priori aucun lien à un corps, par exemple l'ensemble $\mathbb{R}$ muni de la seule relation $=$. Existerait-il un miracle qui permettrait de retrouver la richesse mathématique du corps des nombres réels dans cette structure si pauvre ?

La recherche des réponses à de telles questions nécessite l'introduction de la notion d'un *langage du premier ordre*.

5.2 Langages du premier ordre ; expansions, réduits

Nous avons vu dans la section précédente que le cadre d'étude d'une structure est décrit par sa signature. Pour poursuivre cette étude, il est indispensable d'exprimer les propriétés mathématiques des structures, ou des classes de structures associées à la même signature, et ce, en suivant certaines règles. La notion centrale de ce formalisme syntaxique qui sera développé dans ce chapitre est celle d'un *langage du premier ordre*.

Nous commençons avec l'alphabet. Les symboles d'un langage $\mathcal{L}$ du premier ordre se divisent en trois catégories :

1. Les symboles qui nomment les éléments de la signature fixée :

   R Symboles de relation

   F Symboles de fonction

   C Symboles de constante

2. Symboles logiques

   Quant Quantificateurs : $\forall$, $\exists$

   Con Connecteurs : $\neg$ (*la négation*), $\wedge$ (*la conjonction, "et"*), $\vee$ (*la disjonction, "ou"*), $\rightarrow$ (*l'implication*), $\leftrightarrow$ (*l'équivalence*)

3. Variables et parenthèses

La première catégorie varie d'une signature à une autre tandis que les deux autres appartiennent à chaque langage. Comme l'égalité est toujours présente dans la signature, le symbole d'égalité $=$ fait partie de chaque langage aussi. En conséquence, la description des symboles d'un langage ne mentionnera jamais $=$ ni les catégories (2) et (3) de symboles communs à tous les langages.

Le lien avec les structures est naturel. Une structure $\mathcal{M}$ sera associée à un langage $\mathcal{L}$ dont les symboles sont en correspondance bijective avec les éléments de la signature de $\mathcal{M}$. Une telle structure sera dite une *$\mathcal{L}$-structure*. Dans ce cas, les symboles de $\mathcal{L}$ sont *interprétés* par les éléments de la signature dans la $\mathcal{L}$-structure $\mathcal{M}$ de la façon suivante :

1. $\mathcal{L}$ contient un symbole de constante c_i et un seul pour chaque constante $c_i^{\mathcal{M}}$ $(i \in I_C)$ de $\mathcal{M}$.
2. $\mathcal{L}$ contient un symbole de fonction f_i et un seul pour chaque fonction $f_i^{\mathcal{M}}$ $(i \in I_F)$ de $\mathcal{M}$.
3. $\mathcal{L}$ contient un symbole de relation R_i et un seul pour chaque relation $R_i^{\mathcal{M}}$ $(i \in I_R)$ de $\mathcal{M}$.

Le *cardinal* $|\mathcal{L}|$ *d'un langage* $\mathcal{L}$ est par définition $\max(\aleph_0, |I_C \cup I_F \cup I_R|)$.
Illustrons ces définitions en révisant l'exemple 5.1.1 :

Retour sur l'exemple 5.1.1 : 1. Considérons un langage consistant d'un symbole de relation binaire, donc $\mathcal{L} = \{R_1\}$. Tout ensemble muni d'une relation d'ordre est en particulier une $\mathcal{L}$-structure. Alors $\mathcal{A} = (\aleph_1; \in)$ est une $\mathcal{L}$-structure avec $R_1^{\mathcal{A}} = \in$.

Dans ce même exemple nous avons considéré les expansions. En voici un exemple : nous augmentons le langage $\mathcal{L}$ en y ajoutant un symbole de constante c_1 pour obtenir $\mathcal{L}^+ = \{R_1, c_1\}$. La structure $\mathcal{A}^+ = (\aleph_1; \in, \omega)$ est une $\mathcal{L}^+$-structure avec $R_1^{\mathcal{A}^+} = \in$ et $c_1^{\mathcal{A}^+} = \omega$. C'est une expansion de $\mathcal{A}$. Nous avons obtenu une structure qui connait ses membres finis. Peut-être était-ce déjà le cas pour la structure $\mathcal{A}$, mais la réponse n'est pas claire. D'ailleurs, nous n'avons même pas précisé ce que nous entendons par "Nous avons obtenu une structure qui connait ses membres finis.".

2. Le langage du point 1 peut être utilisé pour ce point aussi. Si $\mathcal{Q} = (\mathbb{Q}; <)$ alors $R_1^{\mathcal{Q}}$ est $<$. Cette fois-ci considérons le langage $\mathcal{L}^+ = \mathcal{L} \cup \{R_2\}$ où R_2 est un symbole de relation unaire. Alors $\mathcal{Q}^+ = (\mathbb{Q}; <, P)$ est une $\mathcal{L}^+$-structure avec les interprétations $R_1^{\mathcal{Q}^+} = <$ et $R_2^{\mathcal{Q}^+} = P$ où $P(x)$ si et seulement si $x \in \mathbb{Q}_+$. C'est une expansion de $\mathcal{Q}$. Cette expansion connait ses éléments positifs. A la fin de ce chapitre vous aurez les connaissances suffisantes pour montrer que cette connaissance est impossible pour $\mathcal{Q}$.

3. Le point 3 est concerné par le "langage des corps". Nous poserons $\mathcal{L} = \{f_1, f_2, f_3, f_4, c_1, c_2\}$. Les fonctions f_1 et f_2 sont binaires tandis que f_3 et f_4 sont unaires. Nous pouvons donc considérer la $\mathcal{L}$-structure $\mathcal{R} = (\mathbb{R}; ., +, {}^{-1}, -, 1, 0)$ et faire les interprétations suivantes :

$$f_1^{\mathcal{R}} = . \ , \ f_2^{\mathcal{R}} = + \ , \ f_3^{\mathcal{R}} = {}^{-1} \ , \ f_4^{\mathcal{R}} = - \ , \ c_1^{\mathcal{R}} = 1 \ , \ c_2^{\mathcal{R}} = 0 \ .$$

On peut réduire le langage à $\mathcal{L}^- = \{f_1, f_2\}$. Le *réduit* correspondant serait $\mathcal{R}^- = (\mathbb{R}; +, .)$. En fait rien n'a été perdu cette fois-ci puisque tout ce qui a disparu du langage peut être "défini" en fonction de ce qui reste. Néanmoins il y a eu certains changements.

Considérons maintenant l'expansion $\mathcal{R}^+ = (\mathbb{R}; ., +, {}^{-1}, -, 1, 0, R)$ où R interprète un symbole de relation unaire telle que $R(x)$ si et seulement si x est un réel positif. Cette expansion des corps des réels officiellement connait ses membres positifs. Mais il en était de même, quoique officieusement, pour $\mathcal{R}$ aussi.

Deux autres réduits algébriquement intéressants de $\mathcal{R}$ sont le groupe additif et le groupe multiplicatif du corps des nombres réels. Un langage pour le premier est $\mathcal{L}_1 = \{+, -, 0\}$. Pour le deuxième $\mathcal{L}_2 = \{., {}^{-1}, 1\}$ est une possibilité.

Notons finalement que la notation dans les paragraphes précédents tend à être lourde. Dans la suite, quitte à un peu de recours à l'abus de langage, nous réduirons la distinction entre la notation pour les langages et celle pour les signatures.

5.3 Extensions, sous-structures

Les expansions et les réduits d'une structure sont des variations induites sur la structure en changeant le langage mais laissant l'univers intact. Deux autres types de variations sont obtenues en variant l'univers sans changer le langage. Elles correspondent aux notions d'*extension* et de *sous-structure*.

Définition 5.3.1 *Soient $\mathcal{L}$ un langage et $\mathcal{M}$ une $\mathcal{L}$-structure. Notons M l'ensemble de base de $\mathcal{M}$. Si M contient un sous-ensemble non vide N clos par rapport aux fonctions de $\mathcal{M}$,*

et qui contient toutes les constantes de $\mathcal{M}$, alors $\mathcal{N}$, qui est la structure obtenue en munissant l'ensemble N des restrictions des fonctions et des relations de $\mathcal{M}$ à N, est dite une sous-structure *de M. Dans ce cas, on dit aussi que $\mathcal{M}$ est une* extension *de $\mathcal{N}$.*

Considérons la structure $\mathcal{A} = \{\aleph_1; \in\}$ dans le langage $\mathcal{L} = \{\in\}$ (déjà un peu d'abus de langage). $\mathcal{A}_0 = \{\aleph_0; \in\}$ est une sous-structure de $\mathcal{A}$. En fait, tout sous-ensemble de $\aleph_1$ fournirait une sous-structure puisqu'il n'y a pas de fonctions dans la structure. Par contre, il n'est pas nécessaire que toute sous-structure soit un ordinal.

Finissons cette courte section avec les corps. Considérons les langages $\mathcal{L}_1 = \{., +, \ ^{-1}, -, 1, 0, \}$ et $\mathcal{L}_2 = \{., +, 1, 0, \}$. Si on considère un corps comme une $\mathcal{L}_2$-structure avec les interprétations naturelles des symboles, alors les sous-structures sont tous des corps. Si on remplace $\mathcal{L}_1$ par $\mathcal{L}_2$ alors les sous-anneaux seront des sous-structures aussi.

Il est temps de préciser le contexte.

5.4 Syntaxe

Nous avons introduit la notion de langage du premier ordre et nous l'avons illustré dans des exemples concrets. Il est temps d'apprendre à nous exprimer correctement. Nous introduirons donc les règles pour écrire des *formules du premier ordre* à partir des symboles d'un langage $\mathcal{L}$ fixé. Malgré le formalisme, ces règles ne sont pas des inventions insolites. Elles sont basées l'activité quotidienne en mathématiques. Pourtant, il faudra faire attention à certaines restrictions, les plus importantes étant la taille des formules et le domaine d'application des quantificateurs.

Nous travaillerons avec un langage fixé, noté $\mathcal{L}$. Il n'y aura aucune restriction au nombre de variables utilisées dans une formule pourvu que ce nombre reste fini pour une même formule. Les règles de parenthésage sont les usuelles. Les définitions ci-dessous ont un caractère récursif. Ceci permet de faire des preuves en utilisant la récurrence sur leurs complexités.

Le point de départ est la notion de *terme* qui précise exactement quelles fonctions peuvent être définies à partir d'un alphabet fixé.

Définition 5.4.1 *Soit $\mathcal{L}$ un langage. Un $\mathcal{L}$-terme est défini de la façon suivante :*

1. *Une variable est un terme.*
2. *Un symbole de constante est un terme.*
3. *Si $t_1, \ldots, t_n$ sont des termes et que f est un symbole de fonction n-aire, alors $f(t_1, \ldots, t_n)$ est un terme.*
4. *Une expression est un $\mathcal{L}$-terme seulement si l'on peut démontrer cela en utilisant 1, 2 et 3.*

Exemple 5.4.1 Une illustration assez naturelle de la définition 5.4.1 est les polynômes dans le langage des anneaux : $\mathcal{L} = \{., +, 0, 1\}$. Vous pouvez vérifier en utilisant la récurrence sur la taille des termes qu'une suite de symboles dans ce langage est un terme si et seulement si son interprétation dans un anneau en caractéristique 0 correspond à un polynôme à coefficients naturels.

Les *formules atomiques* sont les exemples les plus simples de formule du premier ordre.

Définition 5.4.2 *Soit $\mathcal{L}$ un langage. Une* formule atomique *est définie de la façon suivante :*

1. *Si t_1 et t_2 sont deux termes, alors $t_1 = t_2$ est une formule atomique.*
2. *Si $t_1, \ldots, t_n$ sont des termes et que R est un symbole de relation n-aire, alors $R(t_1, \ldots, t_n)$ est une formule atomique.*

Exemple 5.4.2 Si on utilise le langage de l'exemple 5.4.1, les équations polynômiales sont exactement les formules atomiques. Soulignons que les seuls coefficients possibles sont les naturels. Si $(m_0, \ldots, m_k), (n_0, \ldots, n_l) \in \mathbb{N}^k$ alors

$$m_0 + m_1 x + \ldots + m_k x^k = n_0 + n_1 + \ldots + n_l x^l$$

est une formule atomique... ou presque. L'écriture officielle est la suivante :

$$\underbrace{(1+\ldots+1)}_{m_0 \text{ fois}}+\underbrace{(1+\ldots+1)}_{m_1 \text{ fois}}.x+\ldots+\underbrace{(1+\ldots+1)}_{m_k \text{ fois}}\underbrace{(x.\ldots.x)}_{k \text{ fois}}$$
$$=$$
$$\underbrace{(1+\ldots+1)}_{n_0 \text{ fois}}+\underbrace{(1+\ldots+1)}_{n_1 \text{ fois}}.x+\ldots+\underbrace{(1+\ldots+1)}_{n_l \text{ fois}}\underbrace{(x.\ldots.x)}_{l \text{ fois}} .$$

Nous devons admettre que même cette expression manque de formalisme puisqu'il faudrait remplacer les . et + par les symboles correspondants de l'alphabet. Nous ne le ferons pas.

Maintenant nous pouvons définir une formule du premier ordre :

Définition 5.4.3 *Soit $\mathcal{L}$ un langage. On définit une $\mathcal{L}$-formule de la façon suivante :*

1. *Une formule atomique est une formule.*
2. *Si ϕ et ψ sont des formules, alors $\phi \wedge \psi$ et $\neg\phi$ sont des formules.*
3. *Si ϕ est une formule, alors $\forall x\phi$ et $\exists x\phi$, où x est une variable, est une formule.*
4. *Une expression est une $\mathcal{L}$-formule seulement si l'on peut démontrer cela en utilisant 1, 2 et 3.*

Notons que les symboles logiques $\vee$, $\rightarrow$, $\leftrightarrow$ n'apparaissent pas dans les définitions que nous venons de donner. C'est parce qu'ils sont expressibles en fonction des autres : $\phi \vee \psi$ est $\neg(\neg\phi \wedge \neg\psi)$ $\phi \rightarrow \psi$ est $\neg\phi \vee \psi$. Finalement, $\phi \leftrightarrow \psi$ est $(\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$. Soulignons aussi que $\forall x\phi$ est equivalent à $\neg\exists x\neg\phi$.

Une conséquence importante des règles que nous venons de définir est que dans une formule du premier ordre les quantifications se font sur les variables, et les formules ne contiennent qu'un nombre fini de symboles.

Les notions suivantes seront fréquemment utilisées :

Définition 5.4.4 *Les variables qui sont quantifiées sont appelées* variables liées. *Les variables non-quantifiées sont des variables* libres. *Une formule sans variable libre est un* énoncé.

Exemple 5.4.3 1. Dans la première partie de ces notes, à plusieurs occasions nous avons utilisé la notion de propriété sans en donner une définition rigoureuse. En fait, il ne s'agissait que des formules du premier ordre dans le langage $\mathcal{L} = \{\in\}$. Illustrons la propriété d'être successeur. Appelons $\phi(x_1, x_2)$ la formule suivante :

$$(x_1 \in x_2) \wedge \forall x_3(\ x_1 \in x_3\ \rightarrow (x_2 = x_3\ \vee\ x_2 \in x_3))$$

Cette formule, à quelques parenthèses près qui sont omises exprès, exprime que x_2 est successeur de x_1. C'est une propriété que nous avons utilisée pour exprimer l'Axiome de l'Infini : il existe un ensemble inductif.

$$\exists x(0 \in x \wedge \forall y \forall z((y \in x\ \wedge \phi(y, z)) \rightarrow z \in x)) .$$

Bien sûr, ϕ n'est pas un symbole du langage $\mathcal{L}$. Nous l'utilisons pour abbrévier sa définition donnée ci-dessus.

2. Si $\mathcal{L} = \{<\}$ est un langage où $<$ est un symbole de relation binaire interprété comme une relation d'ordre strict,

$$\forall x \forall y \exists z(x < y \rightarrow (x < z \wedge z < y))$$

exprime la densité de l'ordre.

3. Soit $\mathcal{L} = \{., {}^{-1}, 1\}$ le langage des groupes. La formule suivante décrit les éléments centraux :

$$\forall y(x.y = y.x)$$

Une autre façon d'écrire cela, puisque $\mathcal{L}$ contient un symbole de fonction unaire pour l'inversion est

$$\forall y(y^{-1}.x.y = x) \quad .$$

Dans chacune de ces formules y est une variable liée tandis que x ne l'est pas.

La formule suivante est par contre un énoncé

$$\forall x \forall y(x.y = y.x)$$

Dans l'interprétation naturelle dans un groupe cet énoncé exprime que le groupe concerné est abélien.

4. Soit $\mathcal{L} = \{+, ., 0, 1\}$ le langage des anneaux. Si $m \in \mathbb{N}^*$, alors

$$\underbrace{1 + \ldots + 1}_{m \text{ fois}} = 0$$

est un énoncé dans le langage des corps qui énonce que la caractéristique est m.

5.5 Satisfaction d'une formule du premier ordre

La *satisfaction* d'une formule du premier ordre, que nous définirons dans cette section, est le lien entre les considérations syntaxiques et celles sémantiques. Intuitivement, il s'agit de résoudre un système d'équations fini en utilisant les éléments de l'ensemble sous-jacent d'une structure donnée.

Nous définissons d'abord la *valeur* d'un terme :

Définition 5.5.1 *Soient $\mathcal{L}$ un langage et $\mathcal{M}$ une $\mathcal{L}$-structure avec ensemble de base M. Si t est un terme à n variables (noté $t = t(x_1, \ldots, x_n)$) et $\overline{m} = (m_1, \ldots, m_n) \in M^n$, alors la valeur $t[\overline{m}]$ est définie de la façon suivante :*

1. *Si t est la variable x_i, alors $t[\overline{m}] = m_i$.*
2. *Si t est un symbole de constante c, alors $t[\overline{m}] = c^{\mathcal{M}}$.*
3. *Si $t = f(t_1, \ldots, t_k)$ où f est une fonction et les t_i sont des termes, $t[\overline{m}] = f^{\mathcal{M}}(t_1[\overline{m}], \ldots, t_k[\overline{m}])$.*

En particulier les valeurs d'un terme sont des éléments de M.

Maintenant nous procédons à la définition de la notion de satisfaction d'une formule du premier ordre dans une structure. Plus précisément, nous fixerons un langage $\mathcal{L}$ du premier ordre et une $\mathcal{L}$-structure $\mathcal{M}$. Ensuite, nous définirons pour toute $\mathcal{L}$-formule ϕ dont les variables libres sont parmi $\{x_1, \ldots, x_k\}$ et pour toute k-uple $(m_1, \ldots, m_k)$ extrait de l'ensemble de base de $\mathcal{M}$ la *satisfaction de ϕ par $(m_1, \ldots, m_k)$ dans $\mathcal{M}$*, notée

$$\mathcal{M} \models \phi[(m_1, \ldots, m_k)]$$

ou

$$\mathcal{M} \models \phi[\overline{m}]$$

si on utilise l'abbréviation $\overline{m} = (m_1, \ldots, m_k)$.

Définition 5.5.2 *Soient $\mathcal{L}$ un langage et $\mathcal{M}$ une $\mathcal{L}$-structure avec ensemble de base M. Si $\phi(\overline{x})$ est une formule dont les variables libres sont parmi $\overline{x} = (x_1, \ldots, x_n)$, on définit récursivement la satisfaction de ϕ dans $\mathcal{M}$ de la façon suivante :*

1. *Si ϕ est de la forme $t_1 = t_2$ où t_1 et t_2 sont deux termes, alors $\mathcal{M} \models (t_1 = t_2)[\overline{m}]$ si et seulement si $t_1[\overline{m}] = t_2[\overline{m}]$.*
2. *Si ϕ est de la forme $R(t_1, \ldots, t_n)$ où R est un symbole de relation n-aire, alors $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $(t_1[\overline{m}], \ldots, t_n[\overline{m}])$ sont dans la relation $R^{\mathcal{M}}$.*

3. *Si ϕ est de la forme $\neg\alpha$, alors $\mathcal{M} \models \phi[\overline{m}]$ si et seulement s'il n'est pas le cas que $\mathcal{M} \models \alpha[\overline{m}]$ ($\mathcal{M} \not\models \alpha[\overline{m}]$).*
4. *Si ϕ est de la forme $\alpha \wedge \beta$, alors $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $\mathcal{M} \models \alpha[\overline{m}]$ et $\mathcal{M} \models \beta[\overline{m}]$*
5. *Si ϕ est de la forme $\forall y\theta(y, \overline{x})$, alors $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $\mathcal{M} \models \theta[m', \overline{m}]$ pour tout $m' \in M$.*
6. *Cette dernière clause est redondante puisque le quantificateur existentiel s'exprime en fonction du quantificateur universel comme nous l'avons remarqué après la définition 5.4.3. Pour des raisons pratiques faisons quand-même la définition suivante : Si ϕ est de la forme $\exists y\theta(y, \overline{x})$, alors $\mathcal{M} \models \phi[\overline{m}]$ si et seulement s'il existe $m' \in M$ tel que $\mathcal{M} \models \theta[m', \overline{m}]$.*

Un énoncé ψ, n'ayant pas de variables libres, est vrai ($\mathcal{M} \models \psi$) *ou* faux ($\mathcal{M} \not\models \psi$). *Donc pour tout uple $\overline{m}$ extrait de M et tout énoncé ψ, $\mathcal{M} \models \psi[\overline{m}]$ si ψ est vrai et $\mathcal{M} \not\models \psi[\overline{m}]$ si ψ est faux.*

Notre premier lemme illustrera le caractère inductif des définitions syntaxiques et sémantiques de ce chapitre.

Lemme 5.5.3 *Soient $\mathcal{L}$ un langage du premier ordre, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures. On suppose que $M \subset N$. Alors on a les conclusions suivantes :*

1. *Soit $\overline{m} = (m_1, \ldots, m_k) \in M^k$. Si $\phi(x_1, \ldots, x_k)$ est une formule sans quantificateurs, alors $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $\mathcal{N} \models \phi[\overline{m}]$.*
2. *Si $\phi(x_1, \ldots, x_k)$ est une formule sans quantificateurs dont les variables libres sont parmi $(x_1, \ldots, x_k)$ et que $\mathcal{M} \models \exists x_1 \ldots \exists x_k \phi(x_1, \ldots, x_k)$, alors $\mathcal{N} \models \exists x_1 \ldots \exists x_k \phi(x_1, \ldots, x_k)$.*
3. *Si $\phi(x_1, \ldots, x_k)$ est une formule sans quantificateurs dont les variables libres sont parmi $(x_1, \ldots, x_k)$ et que $\mathcal{N} \models \forall x_1 \ldots \forall x_k \phi(x_1, \ldots, x_k)$, alors $\mathcal{M} \models \forall x_1 \ldots \forall x_k \phi(x_1, \ldots, x_k)$.*

Preuve. Le premier point est une illustration de la technique élémentaire de preuve par récurrence sur la complexité d'une formule du premier ordre. La notion de complexité d'une formule peut se définir d'une façon formelle mais, ceci n'étant pas notre premier souci dans ce cours, nous nous contenterons de la définir comme la longueur d'une formule, en d'autres termes le nombre de symboles utilisés.

Si ϕ est de la forme $t_1 = t_2$ où t_1 et t_2 sont deux termes alors $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $t_1[\overline{m}] = t_2[\overline{m}]$. Puisque $M \subset N$, $\overline{m}$ est aussi extrait de N. Alors l'égalité $t_1[\overline{m}] = t_2[\overline{m}]$ équivaut à ce que $\mathcal{N} \models (t_1 = t_2)[\overline{m}]$ d'après le point (1) de la définition 5.5.2. Si ϕ est de la forme $R(t_1, \ldots, t_k)$ avec R un symbole de relation k-aire, alors $\mathcal{M} \models R(t_1, \ldots, t_k)[\overline{m}]$ si et seulement si $(t_1[\overline{m}], \ldots, t_k[\overline{m}]) \in R^{\mathcal{M}}$. Comme $M \subset N$, la conclusion de la dernière phrase équivaut à $(t_1[\overline{m}], \ldots, t_k[\overline{m}]) \in R^{\mathcal{N}}$.

Pour finir la preuve, il suffira de considérer les cas où ϕ est de la forme $\neg\psi$ et $\psi \wedge \theta$. Dans le premier cas, $\mathcal{M} \models \phi$ si et seulement si $\mathcal{M} \not\models \psi$. Par récurrence, ceci équivaut à $\mathcal{N} \not\models \psi$. De manière équivalente, $\mathcal{N} \models \phi$. Dans le deuxième cas, $\mathcal{M} \models \phi$ si et seulement si $\mathcal{M} \models \psi$ et $\mathcal{M} \models \theta$. Par récurrence, ceci équivaut $\mathcal{N} \models \psi$ et $\mathcal{N} \models \theta$. Cette dernière condition équivaut à $\mathcal{N} \models \phi$.

Maintenant, démontrons le deuxième point. D'après la définition 5.5.2 et les remarques qui la suivent, $\mathcal{M} \models \exists x_1 \ldots \exists x_k \phi(x_1, \ldots, x_k)$ si et seulement s'il existe un k-uple $(m_1, \ldots, m_k)$ extrait de M tel que $\mathcal{M} \models \phi[(m_1, \ldots, m_k)]$. Comme ϕ est sans quantificateur, le premier point montre que $\mathcal{N} \models \phi[(m_1, \ldots, m_k)]$. Alors, $\mathcal{N} \models \exists x_1 \ldots \exists x_k \phi(x_1, \ldots, x_k)$.

Dernièrement, démontrons le troisième point. D'après la définition 5.5.2, l'hypothèse de ce point équivaut à $\mathcal{N} \models \phi(n_1, \ldots, n_k)$ pour tout $(n_1, \ldots, n_k)$ extrait de N. En particulier, tous les k-uples extraits de M^k satisfont la formule ϕ dans $\mathcal{N}$. Comme ϕ est sans quantificateurs, le premier point montre qu'ils la satisfont dans $\mathcal{M}$ aussi. En d'autres termes, $\mathcal{M} \models \forall x_1 \ldots \forall x_k \phi(x_1, \ldots, x_k)$. □

Remarques : 1. Des énoncés du type $\exists x_1 \ldots \exists x_k \phi(x_1, \ldots, x_k)$ comme dans le point (2) du lemme 5.5.3 sont dits *existentiels* tandis que ceux du type $\forall x_1 \ldots \forall x_k \phi(x_1, \ldots, x_k)$ comme dans le point (3) sont dits *universels*.

2. Les point (2) et (3) du lemme 5.5.3 seraient faux si on mélangeait les quantificateurs. Illustrons cela par un exemple algébrique. Considérons le groupe $\mathcal{Q} = (\mathbb{Q}; +, 0)$. Alors, $\mathcal{Z} =$

$(\mathbb{Z}; +, 0)$ en est une sous-structure. Alors, $\mathcal{Q} \models \forall x_1 \exists x_2 (x_2 + x_2 = x_1)$ tandis que $\mathcal{Z} \not\models \forall x_1 \exists x_2 (x_2 + x_2 = x_1)$.

3. Le lemme 5.5.3 montre la simplicité des formules sans quantificateurs et les effets de l'introduction des quantificateurs. En théorie des modèles, *l'élimination des quantificateurs*, le procédé qui, intuitivement, consiste à remplacer chaque formule par une équivalente sans quantificateurs, est d'importance primordiale. C'est un thème qui sera abordé dans les prochains chapitres.

Maintenant, nous pouvons donner la définition de la notion qui baptise notre domaine :

Définition 5.5.4 *Soient $\mathcal{L}$ un langage du premier ordre, $\mathcal{M}$ une $\mathcal{L}$-structure et Φ un ensemble d'énoncés écrits dans $\mathcal{L}$. Si pour tout $\phi \in \Phi$, $\mathcal{M} \models \phi$, alors on dit que $\mathcal{M}$ est un* modèle *de Φ. Nous écrirons dans ce cas $\mathcal{M} \models \Phi$.*

Si Φ est un ensemble d'énoncés et que σ est un énoncé tel que pour toute structure $\mathcal{M}$, $\mathcal{M} \models \sigma$ chaque fois que $\mathcal{M} \models \Phi$, alors on dira que σ est une conséquence *de Φ. Dans ce cas, on écrira $\Phi \vdash \sigma$.*

Exemples : 1. Si $\mathcal{L} = \{.\}$ est un langage avec un symbole de fonction binaire et que l'on considère l'ensemble des énoncés Γ qui disent "je suis un groupe" alors pour tout groupe abélien $\mathcal{A} = (A; .)$

$$\mathcal{A} \models \ \Gamma \cup \ \{\forall x \forall y (x.y = y.x)\}$$

La conséquence suivante est une propriété élémentaire et bien connue des groupes :

$$\Gamma \cup \{\forall x (x^2 = 1)\} \vdash \forall x \forall y (x.y = y.x)$$

2. Soit $\mathcal{L} = \{<\}$ un langage avec une seule relation binaire. Soit σ l'ensemble des énoncés qui disent "je suis une relation d'ordre linéaire stricte". L'ensemble σ est fini. Or, le suivant est vrai

$$\sigma \cup \{\forall x \exists y (x < y) \ , \ \forall x \exists y (y < x)\} \vdash \infty$$

où ∞ dit "je suis un ensemble infini". C'est un bon exercice de décrire l'ensemble d'énoncés ∞.

5.6 Morphismes ; liens avec la définissabilité

Cette section a pour but de définir la notion de *morphisme*. Nous avons déjà rencontré des cas particulier de cette notion au début du cours dans le contexte des ensembles ordonnés. L'idée sous-jacente est naturelle, il s'agit de développer une notion robuste d'application préservant les aspects principaux de la structure sur laquelle l'application agit.

Définition 5.6.1 *Soient $\mathcal{L}$ un langage du premier ordre, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures avec ensembles de base M et N respectivement. Une application ν de M vers N et dite un* homomorphisme *si elle satisfait les conditions suivantes :*

1. *Soit $n \in \mathbb{N}^*$. Si f est un symbole de fonction n-aire de $\mathcal{L}$, alors pour tout $(a_1, \ldots, a_n) \in M^n$.*

$$\nu(f^{\mathcal{M}}(a_1, \ldots, a_n)) = f^{\mathcal{N}}(\nu(a_1), \ldots, \nu(a_n)).$$

2. *Si c est un symbole de constante, alors $\nu(c^{\mathcal{M}}) = c^{\mathcal{N}}$.*
3. *Soit $n \in \mathbb{N}^*$. Si R est un symbole de relation n-aire dans $\mathcal{L}$, alors pour tout $(a_1, \ldots, a_n) \in M^n$,*

$$\mathcal{M} \models R(a_1, \ldots, a_n) \text{ implique } \mathcal{N} \models R(\nu(a_1), \ldots, \nu(a_n)) \ .$$

4. *Un morphisme est un* plongement *si pour tout $n \in \mathbb{N}^*$, pour toute relation n-aire dans la signature de $\mathcal{M}$ et de $\mathcal{N}$ et pour tout n-uple $(a_1, \ldots, a_n) \in M^n$,*

$$\mathcal{M} \models R(a_1, \ldots, a_n) \text{ si et seulement si } \mathcal{N} \models R(\nu(a_1), \ldots, \nu(a_n)) \ .$$

5. *Un* isomorphisme *est un plongement surjectif. Un isomorphisme d'une structure sur elle-même est dit un* automorphisme.

Nous soulignons qu'un homomorphisme est défini entre deux structures d'un même langage. Par contre il n'est pas nécessaire qu'il soit expressible dans ce langage.

Exemple 5.6.1 1. Si nous étudions les groupes comme des $\mathcal{L}$-structures où $\mathcal{L} = \{., {}^{-1}, 1\}$, un homomorphisme entre deux structures est un homomorphisme de groupes au sens usuel. Voici un exemple d'isomorphisme dans ce contexte :

$$\mathcal{R}_1 = (\mathbb{R}, +, 0) \text{ et } \mathcal{R}_2 = (\mathbb{R}_+^*, ., 1)$$

sont isomorphes par le biais du l'application exponentielle.

2. On fixe $\mathcal{L} = \{<\}$ où $<$ est un symbole de relation binaire. $\mathcal{Q} = (\mathbb{Q}, <)$ est une $\mathcal{L}$-structure. L'application ν de $\mathbb{Q}$ dans $\mathbb{Q}$ définie par $f(x) = x - 1$ est un automorphisme de $\mathcal{Q}$.

Maintenant nous ajoutons à $\mathcal{L}$ un symbole de relation unaire P dont l'interprétation sera l'ensemble des rationnels positifs. En d'autres termes, $\mathcal{L}^+ = \mathcal{L} \cup \{P\}$ et $\mathcal{Q}^+ = (\mathbb{Q}, <, \mathbb{Q}^+)$. L'application ν n'est plus un automorphisme de cette nouvelle structure. La raison est que la partie de $\mathbb{Q}$ nommée par les P n'est pas stable sous l'action de l'automorphisme f. C'est un aspect général des parties *définissables* des structures, en d'autres termes des parties décrites par des formules du premier ordre. La proposition 5.6.2 en donne la justification.

Une autre façon d'exprimer les conclusions du paragraphe précédent serait de dire que f n'est pas un automorphisme de la structure $\mathcal{Q}^+$. Il est immédiat que les automorphismes $\mathcal{Q}^+$ sont exactement ceux de $\mathcal{Q}$ qui fixent 0.

La proposition suivante fournit les ingrédients techniques pour faire le lien entre les automorphismes et les parties des structures définies par des formules du premier ordre. Sa preuve, comme celle du lemme 5.5.3, est par récurrence sur la complexité des formules. Pour une dernière fois nous donnerons des détails. Rappelons, avant d'énoncer la proposition, que si M est un ensemble et que ν est une application de M vers un autre ensemble, pour un k-uple $\overline{m} = (m_1, \ldots, m_k)$ extrait de M, $\nu(\overline{m}) = (\nu(m_1), \ldots, \nu(m_k))$.

Proposition 5.6.2 *Soient $\mathcal{L}$ un langage du premier ordre, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures. Si ν est un homomorphisme de $\mathcal{M}$ vers $\mathcal{N}$, alors nous avons les conclusions suivantes :*

1. *Si $t(x_1, \ldots, x_k)$ est un terme de $\mathcal{L}$, alors pour tout $\overline{m} = (m_1, \ldots, m_k) \in M^k$, $\nu(t[\overline{m}]) = t[\nu(\overline{m})]$.*
2. *Si ν est un homomorphisme de $\mathcal{M}$ vers $\mathcal{N}$, alors pour toute formule atomique $\phi(x_1, \ldots, x_k)$ et tout $\overline{m} = (m_1, \ldots, m_k)$, si $\mathcal{M} \models \phi[\overline{m}]$ alors $\mathcal{N} \models \phi[\nu(\overline{m})]$.*
3. *Si ν est un plongement de $\mathcal{M}$ vers $\mathcal{N}$, alors pour toute formule sans quantificateurs $\phi(x_1, \ldots, x_k)$ et tout $\overline{m} = (m_1, \ldots, m_k)$, $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $\mathcal{N} \models \phi[\nu(\overline{m})]$.*
4. *Si ν est un isomorphisme, alors pour toute formule $\phi(x_1, \ldots, x_k)$ du premier ordre et tout $\overline{m} = (m_1, \ldots, m_l) \in M^k$, $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $\mathcal{N} \models \phi[\nu(\overline{m})]$.*

En particulier, deux $\mathcal{L}$-structures qui sont isomorphes satisfont les mêmes énoncés.

Preuve. Commençons par le premier point. La récurrence sera sur la complexité de l'écriture d'un terme. Soit t un terme. D'après la définition 5.4.1 (1), t peut être une variable x. Sa valeur, d'après la définition 5.5.1 (1), est obtenue en remplaçant x par un élément m de M. Donc, $\nu(t[m]) = \nu(m) = t[\nu(m)]$. Si t est un symbole de constante c, sa valeur est son interprétation $c^{\mathcal{M}}$ dans M. D'après la définition 5.6.1 (2), $\nu(c^{\mathcal{M}}) = c^{\mathcal{N}} = t[\nu(c^{\mathcal{M}})]$. Si f est un symbole de fonction k-aire et que $t = f(t_1, \ldots, t_k)$, alors

$$\begin{aligned}
\nu(t[\overline{m}]) &= \nu(f^{\mathcal{M}}(t_1[\overline{m}], \ldots, t_k[\overline{m}])) \text{ la définition 5.5.1 (3)} \\
&= f^{\mathcal{N}}(\nu(t_1[\overline{m}]), \ldots, \nu(t_k[\overline{m}])) \text{ la définition 5.6.1 (1)} \\
&= f^{\mathcal{N}}(t_1[\nu(\overline{m})], \ldots, t_k[\nu(\overline{m})]) \text{ hypothèse de récurrence} \\
&= t[\nu(\overline{m})] \text{ la définition 5.5.1 (3)}
\end{aligned}$$

Cela finit la preuve du premier point.

Démontrons le deuxième point. La définition 5.4.2 fournit la feuille de route. Si ϕ est de la forme $t_1 = t_2$ où t_1 et t_2 sont des termes et $\mathcal{M} \models (t_1 = t_2)[\overline{m}]$, alors d'après la définition 5.5.2 (1), $t_1[\overline{m}] = t_2[\overline{m}]$. Comme ν est une application bien définie, $\nu(t_1[\overline{m}]) = \nu(t_2[\overline{m}])$. Le premier point montre alors que $t_1[\nu(\overline{m})] = t_2[\nu(\overline{m})]$. Ainsi, $\mathcal{N} \models (t_1 = t_2)[\nu(\overline{m})]$. Si ϕ est de la forme $R(t_1, \ldots, t_k)$ où R est un symbole de relation et $\mathcal{M} \models R(t_1, \ldots, t_k)[\overline{m}]$, alors d'après la définition 5.5.2 (1), $(t_1[\overline{m}], \ldots, t_k[\overline{m}]) \in R^{\mathcal{M}}$. Il découle de la définition 5.6.1 (3) que $(\nu(t_1[\overline{m}]), \ldots, \nu(t_k[\overline{m}])) \in R^{\mathcal{N}}$. Le premier point montre alors que ceci équivaut à $(t_1[\nu(\overline{m})], \ldots, t_k[\nu(\overline{m})]) \in R^{\mathcal{N}}$, et on conclut en utilisant la définition 5.5.2 (2).

L'inclusion étant un cas particulier de plongement, la preuve du troisième point est la même que celle du premier point du lemme 5.5.3. Il suffit d'insérer ν aux bons endroits et d'utiliser le premier point aux bons moments. Nous laissons la partie atomique de la tâche à nos lecteurs et montrons l'étape de récurrence. A cette étape, ϕ est une négation ou une conjonction. Si ϕ est de la forme $\neg\psi$, alors d'après la définition 5.5.2 (3), $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $\overline{m}$ ne satisfait pas ψ dans $\mathcal{M}$. Par récurrence, $\mathcal{M} \models \psi[\overline{m}]$ *équivaut* à $\mathcal{N} \models \psi[\nu(\overline{m})]$. En utilisant la définition 5.5.2 3 nous concluons que $\mathcal{N} \models \neg\psi[\nu(\overline{m})]$. La conjonction est vérifiée en suivant le même raisonnement.

Dans la preuve du quatrième point, l'endroit où il y a du nouveau est bien sûr l'étape de récurrence concernant les quantificateurs. C'est ce que nous illustrerons ici. Le reste est un exercice pour nos lecteurs. Dans l'étape de récurrence qui fait intervenir les quantificateurs, ϕ est de la forme $\forall y\theta(y, x_1, \ldots, x_l)$ ayant ses variables libres parmi les x_i. D'après la définition 5.5.2 (5), $\mathcal{M} \models \forall y\theta(y, x_1, \ldots, x_l)[\overline{m}]$ si et seulement si $\mathcal{M} \models \theta[m', \overline{m}]$ pour tout $m' \in M$. Par récurrence, cela *équivaut* à $\mathcal{N} \models \theta[\nu(m'), \nu(\overline{m})]$ pour tout $m' \in M$. Or ν est une surjection par hypothèse, donc $\mathcal{N} \models \theta[n', \nu(\overline{m})]$ pour tout $n' \in N$. Maintenant on applique la définition 5.5.2 (5) pour conclure. □

Remarque sur la définissabilité : La proposition 5.6.2 est élémentaire et très naturel. Néanmoins, surtout son point (4) est fort lié à l'étude des *ensembles définissables* dans les structures, en d'autres termes, les parties des puissances cartesiennes de l'univers d'une structure décrites par des formules du premier ordre.

La structure $\mathcal{Q}^+$ du deuxième point de l'exemple 5.6.1 est une bonne illustration de ce lien. Cette structure, qui est une expansion de l'ordre dense linéaire sans extrémités $\mathcal{Q}$, est en fait plus "riche" que ce réduit. Le sous-ensemble nommé par la relation unaire P n'est pas *définissable* dans $\mathcal{Q}$. Plus précisément, on ne peut pas trouver une formule $\phi(x)$ du premier ordre dont x est la seule variable libre telle que pour tout $q \in \mathbb{Q}$, $\mathcal{Q} \models \phi[q]$ si et seulement si $q \in \mathbb{Q}^+$. En effet, si une telle formule existait, d'après le point (4) de la proposition 5.6.2 l'ensemble $\mathbb{Q}^+$ serait stable sous l'action de tout automorphisme de $\mathcal{Q}$. L'automorphisme $q \longmapsto q - 1$ montre que ce n'est pas possible. On peut arriver à des conclusions beaucoup plus fortes sur la définissabilité dans $\mathcal{Q}$ puisque c'est une structure qui a "beaucoup" d'automorphismes.

5.7 Théories et leurs modèles

Dans cette section nous introduirons des notions clés pour le reste du cours. Nous commençons par la notion d'*équivalence élémentaire.*

Définition 5.7.1 *Soient $\mathcal{A}$ et $\mathcal{B}$ deux $\mathcal{L}$-structures où $\mathcal{L}$ est un langage du premier ordre. $\mathcal{A}$ et $\mathcal{B}$ sont dits* élémentairement équivalentes *si elles sont modèles d'exactement les mêmes énoncés ; en d'autres termes, pour tout énoncé σ de $\mathcal{L}$, $\mathcal{A} \models \sigma$ si et seulement si $\mathcal{B} \models \sigma$. On écrit $\mathcal{A} \equiv \mathcal{B}$.*

La proposition 5.6.2 montre que deux structures isomorphes sont élémentairement équivalentes. L'implication inverse est fausse en général, et le *théorème de compacité* du chapitre suivant montrera qu'il y a une raison fondamentale pour cela. On peut néanmoins démontrer le résultat suivant.

Proposition 5.7.2 *Soient $\mathcal{L}$ un langage du premier ordre, $\mathcal{A}$ et $\mathcal{B}$ deux $\mathcal{L}$-structures finies. Alors $\mathcal{A} \equiv \mathcal{B}$ si et seulement si $\mathcal{A} \cong \mathcal{B}$.*

Preuve. Une direction de l'équivalence découle de la proposition 5.6.2. Nous procédons à démontrer l'autre direction. Donc, $\mathcal{A}$ et $\mathcal{B}$ sont deux structures finies élémentairement équivalentes dont les ensembles de base seront notés A et B respectivement. Les hypothèses impliquent que $|A| = |B|$ (pourquoi ?). En particulier, il suffit de trouver un plongement de $\mathcal{A}$ dans $\mathcal{B}$. Nous construirons un tel plongement en faisant des "va" successifs.

Soit $A = \{a_1, \ldots, a_m\}$ une énumération des éléments de A. On commence avec a_1 et on considère toutes les formules du premier ordre à au plus une variable libre dans le langage $\mathcal{L}$ qui sont satisfaites par a_1. En d'autres termes ce sont les énoncés, et les formules à exactement une variable libre satisfaites par a_1. Comme A est un ensemble fini, on peut supposer que l'ensemble des formules satisfaites par a_1 qui ne sont pas des énoncés est fini. En effet, pour chaque partie de A définie par une formule satisfaite par a_1, on peut retenir une formule et une seule. Alors, comme $\mathcal{P}(A)$ est un ensemble fini, nous aurons une liste finie.

Les formules retenues dans le paragraphe précédent seront notées $\phi_1(x), \ldots, \phi_n(x)$. Alors $\mathcal{A} \models \bigwedge_{i=1}^n \phi_i[a_1]$. En particulier $\mathcal{A} \models \exists x \bigwedge_{i=1}^n \phi(x)$. Comme $\mathcal{A} \equiv \mathcal{B}$, $\mathcal{B} \models \exists x \bigwedge_{i=1}^n \phi(x)$. Alors on choisit un élément $b_1 \in B$ tel que $\mathcal{B} \models \exists x \bigwedge_{i=1}^n \phi[b_1]$. Notons avant de passer à l'étape de récurrence que par construction, il est impossible qu'il existe une formule θ telle que $\mathcal{B} \models \theta[b_1]$ tandis que $\mathcal{A} \not\models \theta[a_1]$. En effet, si une telle θ existait alors une formule représentant la partie de A définie par $\neg\theta$ serait parmi $\phi_1, \ldots, \phi_n$. Or cette formule serait satifaite par b_1 dans $\mathcal{B}$. Vu que b_1 satisfait θ dans $\mathcal{B}$, c'est absurde.

Pour l'étape de récurrence on procède d'une façon similaire au dernier paragraphe. Supposons qu'une bijection dont le graphe est $\{(a_1, b_1), \ldots, (a_k, b_k)\}$ soit établie. On considère toutes les formules $\phi(x_1, \ldots, x_k, x_{k+1})$ à au plus $k+1$ variables libres telles que $\mathcal{A} \models \phi[a_1, \ldots, a_k, a_{k+1}]$. On peut supposer qu'il en existe en nombre fini, $\phi_1, \ldots, \phi_n$. Alors $\mathcal{A} \models \bigwedge_{i=1}^n \phi_i[a_1, \ldots, a_k, a_{k+1}]$. Il en découle que

$$\mathcal{A} \models \exists x \left(\bigwedge_{i=1}^{n} \phi(x_1, \ldots, x_k, x) \wedge \bigwedge_{j=1}^{k} (x \neq x_j) \right) [a_1, \ldots, a_k].$$

Par récurrence,

$$\mathcal{B} \models \exists x \left(\bigwedge_{i=1}^{n} \phi(x_1, \ldots, x_k, x) \wedge \bigwedge_{j=1}^{k} (x \neq x_j) \right) [b_1, \ldots, b_k].$$

On choisit un b_{k+1} témoignant cette satisfaction dans B.

La finitude des ensembles force ce procédé à aboutir et cela donne le plongement. □

La notion suivante, fondamentale en théorie des modèles, généralise l'équivalence élémentaire :

Définition 5.7.3 *Soient $\mathcal{L}$ un langage du premier ordre, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures telles que $\mathcal{M}$ soit une sous-structure de $\mathcal{N}$. Nous dirons que $\mathcal{M}$ est une* sous-structure élémentaire *de $\mathcal{N}$, ou que $\mathcal{N}$ est une* extension élémentaire *de $\mathcal{M}$ si pour toute formule $\phi(x_1, \ldots, x_k)$ de $\mathcal{L}$ et $\overline{m} = (m_1, \ldots, m_k) \in M^k$, $\mathcal{M} \models \phi[\overline{m}]$ si et seulement si $\mathcal{N} \models \phi[\overline{m}]$. Dans ce cas, nous écrirons $\mathcal{M} \preceq \mathcal{N}$. Si $M \neq N$, on préférera $\mathcal{M} \prec \mathcal{N}$.*

Plus généralement, s'il existe un plongement ϕ de $\mathcal{M}$ vers $\mathcal{N}$ tel que $\phi(\mathcal{M}) \preceq \mathcal{N}$, alors ϕ est dit un plongement élémentaire*.*

Soulignons que $\mathcal{M} \preceq \mathcal{N}$ implique $\mathcal{M} \equiv \mathcal{N}$, une conséquence immédiate de la définition précédente que nous utiliserons sans mention. L'implication inverse est fausse mais avant de donner des exemples nous donnerons une caractérisation fréquemment utilisée de la notion d'équivalence élémentaire.

Proposition 5.7.4 (Test de Tarski) *Soient $\mathcal{L}$ un langage du premier ordre, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures. On suppose que $\mathcal{M}$ soit une sous-structure de $\mathcal{N}$. Alors $\mathcal{M} \preceq \mathcal{N}$ si et seulement si pour toute formule $\phi(x_1, \ldots, x_k, y)$ et $(m_1, \ldots, m_k) \in M^k$, quand $\mathcal{N} \models \exists y \phi(m_1, \ldots, m_k, y)$, alors il existe un élément $m \in M$ tel que $\mathcal{N} \models \phi(m_1, \ldots, m_k, m)$.*

Preuve. La nécessité de la condition est claire. La suffisance est un exercice (voir la preuve de la proposition 5.6.2). □

Maintenant nous introduisons des notions qui suivent celles de la définition 5.5.4.

Définition 5.7.5 *Soit $\mathcal{L}$ un langage du premier ordre.*

1. *Un ensemble Θ d'énoncés dans $\mathcal{L}$ est dit* consistant, *si Θ a un modèle. Remarquons que si σ est un énoncé, il n'est pas possible que $\Theta \vdash \sigma$ et $\Theta \vdash \neg\sigma$ à la fois quand Θ est consistant. En outre, dans le cas où $\Theta \vdash \sigma$ et $\Theta \vdash \neg\sigma$, nous dirons que Θ est* contradictoire *ou* inconsistant.
2. *Une* théorie du premier ordre *T dans $\mathcal{L}$ est un ensemble consistant d'énoncés qui contient toutes ses conséquences. La théorie T est dite* engendrée *par un sous-ensemble A de T si toutes les conséquences de T sont aussi celles de A. On dit que A est un* ensemble d'axiomes *ou que A est une* axiomatisation *de T.*
3. *Une théorie T du premier ordre est dite* complète *si elle est maximale, en d'autres termes, si pour tout énoncé σ dans $\mathcal{L}$, $\sigma \in T$ ou $\neg\sigma \in T$.*

Parfois nous parlerons du cardinal d'une théorie du premier ordre. Il ne s'agira que du cardinal de son langage.

Lemme 5.7.6 *Si $\mathcal{A}$ et $\mathcal{B}$ sont deux modèles d'une théorie complète T dans un langage $\mathcal{L}$, alors $\mathcal{A} \equiv \mathcal{B}$.*

Preuve. Exercice. □

Dans la théorie des modèles on s'intéresse plutôt à des théories complètes. En général, il n'est pas très facile de conclure si une axiomatisation engendre une théorie complète. Il a fallu parfois des théorèmes profonds pour décider de la complétude d'une théorie. Dans notre contexte, le *va-et-vient* s'avérera un outil efficace pour vérifier la complétude d'une théorie. Néanmoins, tout cela cesse d'être un problème dès que la recherche d'une axiomatisation précise n'est pas le souci principal. La définition et le lemme suivants montrent que le contexte que nous avons développé dans les dernières sections est naturel pour étudier les théories complètes. En effet, chaque structure est accompagnée d'une théorie complète :

Définition 5.7.7 *Soient $\mathcal{L}$ un langage du premier ordre et $\mathcal{M}$ une $\mathcal{L}$-structure.* $\mathrm{Th}(\mathcal{M})$ *est l'ensemble de tous les énoncés qui sont vrais dans $\mathcal{M}$.*

Lemme 5.7.8 *Soient $\mathcal{L}$ un langage du premier ordre et $\mathcal{M}$ une $\mathcal{L}$-structure. Alors* $\mathrm{Th}(\mathcal{M})$ *est une théorie complète. En plus, $\mathcal{M} \equiv \mathcal{N}$ si et seulement si* $\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{N})$.

Preuve. Exercice. □

Dans la notation du lemme et de la définition précédents, $\mathrm{Th}(\mathcal{M})$ est dite la *théorie du premier-ordre* ou *la théorie élémentaire* de $\mathcal{M}$.

La complétude impose des restrictions aux cardinaux des modèles d'une théorie :

Lemme 5.7.9 *Soit T une théorie complète. Si T a un modèle infini, alors tous ses modèles sont infinis.*

Preuve. Exercice. □

Exemples : 1. Nous considérons le langage $\mathcal{L}$ contenant un seul symbole, $<$, une relation binaire. Les énoncés

$$\forall x(\neg(x < x))$$

$$\forall x \forall y(x < y \to \neg(y < x))$$

$$\forall x \forall y \forall z((x < y \wedge y < z) \to x < z)$$

$$\forall x \forall y \exists z(x < y \to (x < z \wedge z < y))$$

$$\forall x \exists y(x < y)$$

$$\forall x \exists y(y < x)$$

disent que tout modèle de cet ensemble d'énoncés est une chaîne dense sans extrémités. En particulier, c'est un ensemble consistant d'énoncés. Définissons T l'ensemble des conséquences de ces énoncés. T est une théorie. Nous verrons que c'est une théorie complète : la théorie des chaînes denses sans extrémités.

2. En utilisant le langage $\{+, ., -, \ ^{-1}, 0, 1\}$ des corps nous pouvons écrire les énoncés du premier ordre qui expriment le fait d'être un corps. Donc c'est un ensemble consistant, et l'ensemble T de toutes ses conséquences est une théorie du premier ordre, celle des corps. Cette théorie est incomplète pour une multitude de raisons élémentaires dont le quatrième point de l'exemple 5.4.3 ou le lemme 5.7.9. On peut ajouter une infinité d'énoncés qui expriment que les modèles de T sont infinis et de caractéristique fixée ainsi que leurs conséquences. La nouvelle théorie ne serait toujours pas complète. Pouvez-vous voir pourquoi ?

3. Nous avons déjà remarqué que si $\mathcal{A} \preceq \mathcal{B}$ alors $\mathcal{A} \equiv \mathcal{B}$. L'implication réciproque est fausse en général. Nous donnons un exemple. On considère le langage $\mathcal{L}$ contenant un symbole de fonction binaire $+$. Alors, $(\mathbb{Z}, +)$ et$(2\mathbb{Z}, +)$ sont deux $\mathcal{L}$-structures isomorphes. Donc $(\mathbb{Z}, +) \equiv (2\mathbb{Z}, +)$ d'après la proposition 5.6.2 (4). Par contre, $(\mathbb{Z}, +) \models \exists x(x + x = y)[2]$ bien que ce ne soit pas le cas pour $(2\mathbb{Z}, +)$.

4. Illustrons le phénomène du point (3) dans une structure relationnelle. Cette fois-ci notre langage sera $\mathcal{L} = \{<\}$, $<$ étant un symbole de relation binaire que l'on interprétera comme une relation d'ordre. $(\mathbb{Z}, <)$ et $(2\mathbb{Z}, <)$ sont des $\mathcal{L}$-structures isomorphes, et $(2\mathbb{Z}, <)$ est une sous-structure de $(\mathbb{Z}, <)$. Nous considérons la formule $\phi(x, y)$ définie par $\exists z(x < z \wedge z < y)$.

$$(\mathbb{Z}, <) \models \phi(x, y)[(0, 2)]$$

mais

$$(2\mathbb{Z}, <) \models \neg\phi(x, y)[(0, 2)] \ .$$

Chapitre 6

Compacité

"Ce malheureux Théorème de Compacité est entré par la petite porte, et on dirait que cette modestie originelle lui cause encore du tort dans les manuels de logique. C'est à mon avis un résultat beaucoup plus essentiel, primordial (et donc moins sophistiqué), que le Théorème de Complétude de Gödel,..." Bruno Poizat, CTM, p.78

6.1 Théorème de compacité

Dans cette section nous énoncerons plusieurs versions équivalentes du théorème de compacité. Ce théorème est l'un des plus fondamentaux dans la théorie des modèles et il est surprenant de voir ses applications simples mais cruciales dans les contextes les moins attendus. Une telle application simple à la frontière de l'algèbre et de la théorie des modèles sera présentée dans cette section. Dans la section suivante, la première application majeure en théorie des modèles sera présentée, à savoir le théorème de Löwenheim-Skolem.

Théorème 6.1 (Théorème de compacité) *Soit $\mathcal{L}$ un langage. Un ensemble de $\mathcal{L}$-énoncés est consistant si et seulement si chacun de ses sous-ensembles finis est consistant.*

Remarquons que le lien entre ce théorème et la notion topologique de compacité n'est pas seulement restreint à une similarité de terminologies. Les outils de la preuve que nous en donnerons sont suffisants pour transférer la discussion entière dans le domaine de la topologie et démontrer qu'un certain espace topologique est compact et en déduire l'énoncé ci-dessus. Nous suivrons le chemin plus logicien dans la section 6.3 mais présenterons quand-même la topologie pertinente dans la section 6.4

Nous donnons des versions équivalentes du théorème de compacité qui seront aussi utiles :

Théorème 6.2 *Soient $\mathcal{L}$ un langage et Φ un ensemble de $\mathcal{L}$-énoncés.*

1. *Φ est contradictoire si et seulement si Φ a un sous-ensemble fini qui est contradictoire.*
2. *ϕ est une conséquence de Φ si et seulement s'il existe un sous-ensemble fini de Φ dont ϕ est conséquence.*

Preuve. L'équivalence de ces énoncés au théorème 6.1 est un simple exercice laissé aux soins du lecteur. □

Nous présentons une application simple du théorème de compacité. Comme cette section assez courte est déjà en train de toucher à sa fin sans avoir exigé un effort particulier, nous suggérons à nos lecteurs d'en profiter pour faire une lecture attentive de sa preuve.

Proposition 6.1.1 *Il n'existe pas de théorie T du premier ordre (dans le langage des groupes) telle que $\mathcal{G} \models T$ si et seulement si $\mathcal{G}$ est un groupe périodique, en d'autre termes un groupe dont tous les éléments sont d'ordre fini.*

Preuve. Nous faisons un raisonnement par l'absurde. Supposons qu'une telle théorie T existe. Soit $\mathcal{L}$ le langage des groupes que nous augmentons à $\mathcal{L}^+$ en y ajoutant un symbole de constante c. Nous augmentons T à $T^+ = T \cup \{c^n \neq 1 : n \in \mathbb{N}\}$. Un sous-ensemble fini T_0 de T^+ est l'union d'un nombre fini d'énoncés de T et d'un ensemble du type $\{c^{n_1} \neq 1, \ldots, c^{n_k} \neq 1\}$. Si p est un nombre premier, alors le groupe cyclique C_p d'ordre p est un modèle de T et en particulier de $T \cap T_0$. Si $p \geq \max(n_1, \ldots, n_k)$ (d'autres p marcheraient aussi mais ce choix, qui est disonible, nous suffit), alors C_p est un modèle de T_0 parce qu'on peut interpréter c par un quelconque élément nontrivial de C_p. En d'autres termes, T_0 est consistante. Donc, T^+ a un modèle $\mathcal{H}^+$ et l'élément $c^{\mathcal{H}^+}$ est d'ordre infini. Maintenant nous pouvons réduire notre langage à $\mathcal{L}$, et la $\mathcal{L}$-structure réduite $\mathcal{H}$ que l'on obtient est un modèle de T qui contient un élément d'ordre infini. C'est une contradiction. □

L'application suivante du théorème de compacité mérite une section séparée. Si vous avez suffisamment réflechi sur la preuve de la proposition 6.1.1, il est temps d'aborder un autre théorème de la théorie des modèles.

6.2 Théorème de Löwenheim-Skolem

Dans cette section nous étudierons des théorèmes qui permettent de démontrer l'existence des extensions ou des sous-structures élémentaires d'une structure infinie. Le premier résultat est le théorème de Löwenheim-Skolem qui a deux parties dont nous ne démontrerons que la moitié qui dépend de la compacité.

Théorème 6.3 (Théorème de Löwenheim-Skolem) *Soient $\mathcal{L}$ un langage du premier ordre et $\mathcal{M}$ une $\mathcal{L}$-structure infinie.*

1. *Si $A \subseteq M$, alors M a une sous-structure élémentaire $\mathcal{M}_0$ dont l'univers contient A et qui est de cardinal égal à $\max(|A|, |\mathcal{L}|)$.*
2. *Si $\kappa \geq |\mathcal{M}|$, alors il existe une extension élémentaire de $\mathcal{M}$ de cardinal exactement κ.*

Preuve. Nous ne démontrerons que le deuxième point dont la preuve utilise la compacité avant d'être réduit au premier point. La preuve du premier point est plutôt un exercice en arithmétique des cardinaux.

La preuve du point (2) fait usage d'une méthode fréquemment utilisée en théorie des modèles. En suivant nos coutumes, nous noterons M l'univers de $\mathcal{M}$. D'abord on ajoute au langage $\mathcal{L}$ un symbole de constante c_m pour chaque $m \in M$. L'expansion de $\mathcal{M}$ ainsi obtenue sera notée $(\mathcal{M}, m)_{m \in M}$). Sa théorie du premier ordre est $\mathrm{Th}((\mathcal{M}, m)_{m \in M})$. Ainsi, $\phi(c_{m_1}, \ldots, c_{m_k})$ est un énoncé vrai dans $(\mathcal{M}, m)_{m \in M}$ si et seulement si $\mathcal{M} \models \phi[(m_1, \ldots, m_k)]$.

On augmente le langage davantage en y ajoutant un ensemble de symboles de constantes $\{a_i : i \in I\}$ où I est de cardinal κ. On tache de séparer ces nouveaux symboles de constantes de ceux utilisés pour nommer les éléments de l'univers de $\mathcal{M}$. On considère l'ensemble $T^+ = \mathrm{Th}((\mathcal{M}, m)_{m \in M}) \cup \{a_i \neq a_j : i \neq j,\ i, j \in I\}$. Nous montrerons que c'est un ensemble consistant d'énoncés en utilisant la compacité. Une partie finie T_0 de T^+ contient un sous-ensemble fini d'énoncés de $\mathrm{Th}((\mathcal{M}, m)_{m \in M})$ et un ensemble fini d'énoncés de la forme $\{a_{i_1} \neq a_{j_1}, \ldots, a_{i_k} \neq a_{j_k}\}$. Comme les a_i n'interviennent pas dans $\mathrm{Th}((\mathcal{M}, m)_{m \in M})$ et que M est un ensemble infini, nous voyons que $\mathcal{M} \models T_0$. En effet, il suffit d'interpréter les $\{a_{i_1}, a_{j_1}, \ldots, a_{i_k}, a_{j_k}\}$ de façon à satisfaire T_0 et le reste des a_i arbitrairement, quitte à éviter les valeurs choisies pour $\{a_{i_1}, a_{j_1}, \ldots, a_{i_k}, a_{j_k}\}$. Par compacité, T^+ est consistante. Nous prenons un modèle de T^+ et nous le considérons dans le langage réduit $\mathcal{L}$. Appelons cette $\mathcal{L}$-structure $\mathcal{N}$. Son ensemble de base, noté N, contient tous les éléments nommés par les a_i. Donc, $|N| \geq \kappa$.

Les éléments de $\mathcal{N}$ qui interprètent les constantes c_m témoignent de l'existence d'une sous-structure de $\mathcal{N}$ qui est isomorphe à $\mathcal{M}$. Donc, nous pouvons supposer que $\mathcal{M}$ est une sous-structure de $\mathcal{N}$. En plus, comme $(\mathcal{M}, m)_{m \in M}$ et $(\mathcal{N}, m)_{m \in M}$ sont des modèles de la théorie complète $\mathrm{Th}((\mathcal{M}, m)_{m \in M})$, $(\mathcal{M}, m)_{m \in M} \equiv (\mathcal{N}, m)_{m \in M}$. Ces deux observations nous permettront de montrer que $\mathcal{M} \preceq \mathcal{N}$. Soient donc $\phi(x_1, \ldots, x_r)$ une formule et $(m_1, \ldots, m_r)$ un r-uple

extrait de M. Alors, les équivalences suivantes sont vraies :

$$\begin{aligned} \mathcal{M} &\models \phi[m_1, \ldots, m_r] \text{ si et seulement si} \\ (\mathcal{M}, m)_{m\in M} &\models \phi(c_{m_1}, \ldots, c_{m_r}) \text{ si et seulement si} \\ (\mathcal{N}, m)_{m\in M} &\models \phi(c_{m_1}, \ldots, c_{m_r}) \text{ si et seulement si} \\ \mathcal{N} &\models \phi[m_1, \ldots, m_r] \end{aligned}$$

Ainsi, $\mathcal{M} \preceq \mathcal{N}$.

Finalement, il reste à montrer l'existence d'une extension élémentaire de cardinal exactement κ. L'ensemble des éléments nommés par les a_i est de cardinal κ, et M est de cardinal au plus κ. Appelons A l'union de ces deux sous-ensembles de N. D'après le point (1), il existe une sous-structure élémentaire $\mathcal{N}_0$ de $\mathcal{N}$ dont l'univers N_0 contient A et qui est de cardinal exactement κ. Il reste à vérifier que $\mathcal{M} \preceq \mathcal{N}_0$. Soient alors $\phi(x_1, \ldots, x_k)$ une $\mathcal{L}$-formule du premier ordre dont les variables libres sont parmi $\{x_1, \ldots, x_k\}$, et $(m_1, \ldots, m_k)$ un $k$uplet extrait de M. Comme $\mathcal{M} \preceq \mathcal{N}$,

$$\mathcal{M} \models \phi[(m_1, \ldots, m_k)] \text{ si et seulement si } \mathcal{N} \models \phi[(m_1, \ldots, m_k)] \ .$$

Or $M \subset N_0$ et $\mathcal{N}_0 \preceq \mathcal{N}$. Alors,

$$\mathcal{N}_0 \models \phi[(m_1, \ldots, m_k)] \text{ si et seulement si } \mathcal{N} \models \phi[(m_1, \ldots, m_k)] \ .$$

Il en découle que

$$\mathcal{M} \models \phi[(m_1, \ldots, m_k)] \text{ si et seulement si } \mathcal{N}_0 \models \phi[(m_1, \ldots, m_k)] \ .$$

□

La méthode que nous avons utilisée dans la preuve précédente pour obtenir $\mathcal{M} \preceq \mathcal{N}$ est dite la *méthode des diagrammes* en théorie des modèles. Son idée principale est simple mais suffisamment importante pour être reprise dans un lemme séparé dont la preuve est dans celle du théorème de Löwenheim-Skolem :

Lemme 6.2.1 *Soient $\mathcal{L}$ un langage, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures. Alors, $\mathcal{M} \preceq \mathcal{N}$ si et seulement si $M \subseteq N$ et $(\mathcal{M}, m)_{m\in M} \equiv (\mathcal{N}, m)_{m\in M}$. En d'autres termes, une extension élémentaire de $\mathcal{M}$ n'est qu'un modèle de* $\mathrm{Th}((\mathcal{M}, m)_{m\in M})$.

Notons aussi que la théorie $\mathrm{Th}((\mathcal{M}, m)_{m\in M})$ de la preuve du théorème de Löwenheim-Skolem est dite le *diagramme* $\mathcal{M}$. Cette terminologie peut varier d'un ouvrage à l'autre.

Le corollaire suivant est une application du théorème de Löwenheim-Skolem qui permet de vérifier que certaines théories sont complètes.

Corollaire 6.2.2 *Soient $\mathcal{L}$ un langage et T une $\mathcal{L}$-théorie dont tous les modèles sont infinis. On suppose en plus qu'il existe un cardinal $\kappa \geq |\mathcal{L}|$ tel que tous les modèles de T de cardinal κ soient isomorphes (on dit que T est κ-catégorique). Alors T est complète.*

Preuve. Soient $\mathcal{A}$ et $\mathcal{B}$ deux modèles de T. Nous montrerons que $\mathcal{A} \equiv \mathcal{B}$. Par le théorème de Löwenheim-Skolem il existe un modèle $\mathcal{A}_1$ et un autre $\mathcal{B}_1$ tous les deux de cardinal κ tels que $\mathcal{A} \equiv \mathcal{A}_1$ et que $\mathcal{B} \equiv \mathcal{B}_1$. Par hypothèse $\mathcal{A}_1 \cong \mathcal{B}_1$, et en particulier, $\mathcal{A}_1 \equiv \mathcal{B}_1$. La conclusion suit. □

Maintenant, nous appliquerons ce corollaire à la théorie des chaînes denses sans extrémités. L'étape principale est le théorème suivant qui était résumé pendant la séance de propagande de notre cours.

Théorème 6.4 *Deux chaînes denses sans extrémités qui sont dénombrables sont isomorphes.*

Preuve. Le raisonnement de la preuve, auquel nous avons fait allusion à la fin du chapitre précédent, est dit la méthode de *va-et-vient*. En utilisant cette méthode nous construirons un isomorphisme entre deux chaînes denses sans extrémités.

Soient donc $(E_1, \leq_1)$ et $(E_2, \leq_2)$ deux chaînes denses sans extrémités dénombrables. Nous construirons un sous-ensemble de $E_1 \times E_2$ qui ne sera que le graphe d'un isomorphisme f de E_1 sur E_2. Le graphe de f (qui sera noté par la même lettre f) sera construit par récurrence comme une suite croissante des graphes des isomorphismes partiels sur des sous-ensembles finis.

Comme E_1 et E_2 sont supposées être dénombrables, il existe une énumération $\{x_i : i \in \mathbb{N}\}$ de E_1 et une énumération $\{y_i : i \in \mathbb{N}\}$ de E_2. Par contre ces énumérations n'ont rien à faire avec $\leq_1$ et $\leq_2$.

La récurrence est amorcée en posant (x_0, y_0) comme le premier isomorphisme partiel. Supposons maintenant qu'un isomorhisme partiel $\{(x_{i_1}, y_{j_1}), \ldots, (x_{i_n}, y_{j_n})\}$ soit construit. Soit x_r l'élément ayant le plus petit indice (par rapport à l'énumération fixée pour E_1) dans $E_1 \setminus \{x_{i_1}, \ldots, x_{i_n}\}$. Nous trouverons une image pour x_r. Comme E_1 est un ensemble totalement ordonné, il existe trois possibilités pour la place de x_r par rapport aux éléments de l'ensemble $\{x_{i_1}, \ldots, x_{i_n}\}$:

(i) x_r est plus petit que tous les éléments de $\{x_{i_1}, \ldots, x_{i_n}\}$;

(ii) x_r est plus grand que tous les éléments de $\{x_{i_1}, \ldots, x_{i_n}\}$;

(iii) ni (i) ni (ii).

Supposons (i). Comme E_2 est un ensemble totalement ordonné sans extrémités, il contient un élément y_s strictement plus petit que tous les éléments de $\{y_{i_1}, \ldots, y_{i_n}\}$. Alors nous ajoutons (x_r, y_s) au graphe de notre isomorphisme partiel déjà construit. On procède d'une façon similaire pour (ii), quitte à utiliser la nonexistence d'un plus grand élément dans E_2.

Dans le cas (iii), l'ensemble $I = \{i_1, \ldots, i_n\}$ s'écrit comme l'union disjointe de deux sous-ensembles nonvides I_0 et I_1, tels que si $i_k \in I_0$, alors $x_{i_k} < x_r$, et si $i_k \in I_1$, alors $x_r < x_{i_k}$. L'ensemble $J = \{j_1, \ldots, j_n\}$ s'écrit comme l'union disjoint de deux sous-ensembles J_0 et J_1 suivant la décomposition de I, en d'autres termes, $j_k \in J_i$ $(i = 0, 1)$ si et seulement si $i_k \in J_i$. Comme J_0 et J_1 sont des ensembles finis et que E_2 est une chaîne dense, nous pouvons trouver $y_s \in E_2$ tel que y_s soit strictement supérieur à tous les éléments dont les indices sont dans J_0 et strictement inférieur à tous ceux dont les indice sont dans J_1. Alors nous ajoutons (x_r, y_s) au graphe déjà construit. C'était le *va*.

Maintenant le *vient*. Nous avons obtenu à la fin de l'étape *va* un nouveau graphe

$$\{(x_{i_1}, y_{j_1}), \ldots, (x_{i_n}, y_{j_n}), (x_r, y_s)\}.$$

On applique le procédé de *va* dans le sens inverse pour obtenir le *vient*. En d'autres termes, on choisit un élément $y_t \in E_2 \setminus \{y_{j_1}, \ldots, y_{j_n}, y_s)\}$ tel que t soit le plus petit indice qui n'est pas dans $J \cup \{s\}$, et on essaye de lui attribuer un antécédent x_u dans $E_1 \setminus \{x_{i_1}, \ldots, x_{i_n}, x_m\}$ en considérant les divers cas qui sont des analogues de (i), (ii) et (iii).

Cette construction épuise les deux chaînes complètement puisqu'on trouve pour x_k une image au plus tard à la fin du k^e *va*, et y_k aura son antécédent pas plus tard qu'à la fin du k^e *vient*. La bonne définition de f et son injectivité découlent de la construction qui évite d'utiliser les éléments qui sont déjà apparus dans le graphe. Nous avons construit notre isomorphisme. □

Corollaire 6.2.3 *La théorie des chaînes denses sans extrémités est complète.*

Preuve. Elle est catégorique en cardinal dénombrable ou *ω-catégorique* ou *$\aleph_0$-catégorique*. □

Remarquons que le corollaire 6.2.2 n'est pas aussi utile en pratique qu'il peut paraître. En effet, la plupart des structures n'ont pas une théorie du premier ordre catégorique en quelque cardinal que ce soit. Par ailleurs, même quand les structures jouissent de cette propriété très fortes, le va-et-vient impliqué dans la vérification de la catégoricité prouve la complétude en même temps. C'est aussi vrai pour la preuve du théorème 6.4 ci-dessus. Nous reviendrons à ce point dans les chapitres suivants.

Malgré nos remarques légèrement négatives à propos du corollaire 6.2.2 dans le paragraphe précédent, il nous a fourni une axiomatisation naturelle qui a aussi la vertu d'engendrer une théorie complète. En effet, l'axiomatisation de la théorie des chaînes denses sans extrémités que nous avons présentée dans la section 5.7 et ses conséquences forment une théorie complète, à savoir $\mathrm{Th}((\mathbb{Q}, <))$.

Nous finissons cette section avec une application importante de la méthode des diagrammes... et bien sûr de la compacité.

Proposition 6.2.4 *Soient $\mathcal{M}$ et $\mathcal{N}$ deux structures élémentairement équivalentes. Alors elles ont une extension élémentaire commune.*

Preuve. Nous noterons M et N les univers de $\mathcal{M}$ et de $\mathcal{N}$ respectivement. Comme dans la preuve du théorème de Löwenheim-Skolem, nous ajoutons au langage des symboles de constantes c_m pour chaque $m \in M$ et c_n pour chaque $n \in N$.

Nous montrerons que l'ensemble $\mathrm{Th}((\mathcal{M}, m)_{m\in M}) \cup \mathrm{Th}((\mathcal{N}, n)_{n\in N})$ est consistant en utilisant la compacité. Une partie finie de cette union contient des énoncés de la forme $\phi(c_{m_1}, \ldots, c_{m_r})$ et $\psi(c_{n_1}, \ldots, c_{n_s})$. Comme $\mathrm{Th}((\mathcal{M}, m)_{m\in M})$ et $\mathrm{Th}((\mathcal{N}, n)_{n\in N})$ sont complètes, elles contiennent les conjonctions finies de leurs énoncés aussi. Par conséquent, nous pouvons supposer qu'un sous-ensemble fini T_0 de $\mathrm{Th}((\mathcal{M}, m)_{m\in M}) \cup \mathrm{Th}((\mathcal{N}, n)_{n\in N})$ que nous fixerons maintenant contient exactement un énoncé de chaque type susmentionné.

D'après la préparation du paragraphe précédent, il existe deux formules du premier ordre $\phi(x_1, \ldots, x_r)$ et $\psi(x_1, \ldots, x_s)$ telles que

$$(\mathcal{M}, m)_{m\in M} \models \phi(c_{m_1}, \ldots, c_{m_r})$$

et

$$(\mathcal{N}, n)_{n\in N} \models \psi(c_{n_1}, \ldots, c_{n_s}) \ .$$

Il en découle que

$$\mathcal{N} \models \exists x_1 \ldots x_s \psi(x_1, \ldots, x_s) \ .$$

Comme $\mathcal{M} \equiv \mathcal{N}$,

$$\mathcal{M} \models \exists x_1 \ldots x_s \psi(x_1, \ldots, x_s)$$

aussi. Il en découle qu'il existe un $s$uplet $(m'_1, \ldots, m'_s) \in M^s$ tel que

$$\mathcal{M} \models \psi[(m'_1, \ldots, m'_s)] \ .$$

Ainsi,

$$(\mathcal{M}, m)_{m\in M} \models \phi(c_{m_1}, \ldots, c_{m_r}) \wedge \psi(c_{m'_1}, \ldots, c_{m'_s}) \ .$$

En conséquence, $(\mathcal{M}, m)_{m\in M} \models T_0$. Par compacité, $\mathrm{Th}((\mathcal{M}, m)_{m\in M}) \cup \mathrm{Th}((\mathcal{N}, n)_{n\in N})$ est consistant. On peut conclure comme dans la preuve du théorème de Löwenheim-Skolem qu'un modèle de $\mathrm{Th}((\mathcal{M}, m)_{m\in M}) \cup \mathrm{Th}((\mathcal{N}, n)_{n\in N})$, suite à une réduction à $\mathcal{L}$ est une extension élémentaire de $\mathcal{M}$ et de $\mathcal{N}$, ou plus précisément, contient deux sous-structures élémentaires isomorphes à $\mathcal{M}$ et $\mathcal{N}$ respectivement. □

Corollaire 6.2.5 *Si $\{\mathcal{M}_i : i \in I\}$ est une famille de structures deux à deux élémentairement équivalentes, les structures de cette famille ont une extension élémentaire commune.*

Preuve. Exercice. □

6.3 Preuve du théorème de compacité

Dans cette section, nous donnerons une preuve du théorème de compacité. La preuve que nous avons choisie utilise une notion importante d'aspects à la fois topologique et algébrique et qui en fait transcende plusieurs domaines des mathématiques. Il s'agit de la notion d'*ultraproduit*. Dans beaucoup d'applications, les raisonnements ni les idées de la preuve que nous donnerons ne seront visibles. Néanmoins, les ultraproduits sont des outils fréquemment utilisés en théorie des modèles et des ensembles parce qu'elles fournissent des extensions élémentaires.

Conformément à ce que nous avons indiqué au début de ce chapitre, les aspects topologiques du théorème de compacité vont bien au delà de son nom. Nos outils principaux pour la preuve, les ultraproduits, sont cruciaux pour cette interprétation topologique et sa preuve. Bien que nous ayons opté de choisir une preuve non topologique, nous résumerons les liens avec la topologie dans la section suivante. Notons que ces liens seront omniprésents dans d'autres contextes tels que la discussion des *types* au chapitre suivant.

Sans trop attendre, nous pouvons commencer la préparation de la preuve. Les ultraproduits et leurs discussions sont basées sur une notion déjà introduite aux travaux dirigés, à savoir la notion d'*ultrafiltre*, dont nous résumerons rapidement les aspects principaux. Nos lecteurs sont invités à réviser leur quatrième fiche.

Définition 6.3.1 *Soit I un ensemble non vide. Une partie $\mathcal{F}$ de $\mathcal{P}(I)$ est dite un* filtre *si elle vérifie les conditions suivantes :*

1. $I \in \mathcal{F}$; $\emptyset \notin \mathcal{F}$.
2. *Si* $A,\ B \in \mathcal{F}$ *alors* $A \cap B \in \mathcal{F}$.
3. *Si* $A \in \mathcal{F}$ *et* $B \subset I$ *alors* $B \in \mathcal{F}$.

Dans la fiche 4 des travaux dirigés plusieurs exemples de cette notion importante ont été donnés. Parmi eux, il convient de rappeler les voisinages d'un point en topologie. Une façon intuitive d'aborder les filtres est de les voir comme des sous-ensembles "larges" d'un espace. L'espace lui-même est large, deux larges s'intersectent largement, tout ensemble contenant un ensemble large est large. La largeur des éléments d'un filtre corrrespond à diverses notions suivant les domaines d'application de la notion : ensembles ouverts denses, propriété vérifiées presque partout par rapport à une mesure, etc.

Deux notions sont liées à celle d'un filtre :

Définition 6.3.2

1. Une prébase *de filtre $\mathcal{B}$ est un ensemble de parties d'un ensemble I non vide qui est contenu dans un filtre. De manière équivalente, une partie $\mathcal{B}$ de $\mathcal{P}(I)$ est une prébase de filtre si et seulement si l'intersection d'un nombre fini de ses éléments n'est jamais vide. Le plus petit filtre contenant $\mathcal{B}$ est dit le* filtre engendré *par $\mathcal{B}$. Il s'agit du filtre des parties de I contenant chacune une intersection finie des éléments de $\mathcal{B}$.*

2. Une prébase de filtre B est dite une base de filtre *si l'intersection de deux éléments arbitraires de $\mathcal{B}$ est dans $\mathcal{B}$. Le filtre engendré par une base de filtre est le filtre dont chaque élément contient un élément de $\mathcal{B}$.*

Voici deux exemples de bases :

Exemple 6.3.1 Soit I un ensemble non vide.

1. Toute partie A non vide de I est une base de filtre.
2. Si J est un ensemble et que I est l'ensemble des parties finies de J, alors pour tout $i \in I$ on définit
$$I_i = \{j \in I | j \supset i\} \quad .$$
L'ensemble $\{I_i | i \in I\}$ est une base de filtre des parties de I.

Un *ultrafiltre* est un filtre maximal. L'existence des ultrafiltres est liée à l'axiome du choix :

Axiome de l'ultrafiltre : Tout filtre est contenu dans un ultrafiltre.

Le lemme suivant résume certaines propriétés simples mais fondamentales des ultrafiltres :

Lemme 6.3.3 *Soient I un ensemble non vide et $\mathcal{F}$ un filtre des parties non vides de I.*

1. $\mathcal{F}$ est un ultrafiltre si et seulement si pour chaque partie $A \subset I$, soit $A \in \mathcal{F}$ soit $I \setminus A \in \mathcal{F}$.

2. Soit $\mathcal{U}$ un ultrafiltre de parties de I. Si, un nombre fini de parties $\{A_1, \ldots, A_k\}$ recouvre I alors il existe A_i tel que $A_i \in \mathcal{U}$.

Le caractère non constructif de l'axiome de l'ultrafiltre est la raison principale pour laquelle il est impossible d'expliciter un ultrafiltre sauf dans un cas, à savoir un ultrafiltre *principal*, qui par définition est la famille des parties d'un ensemble non vide I contenant une partie finie fixée $\{a_1, \ldots, a_k\}$. Il découle du lemme 6.3.3 (2) que ce sont les parties de I contenant un a_i fixé.

Nous verrons dans quelques paragraphes que les ultrafiltres principaux n'ont rien de principal pour nos objectifs, ils sont inutiles. Malgré ce manque de "visibilité" des ultrafiltres utiles, dits *nonprincipaux*, il est légitime (le lemme 6.3.3 (1) et l'axiome de l'ultrafiltre) et fructueux de fixer une partie infinie de I et d'essayer de déterminer les propriétés d'un ultrafiltre qui évite ou contient cette partie.

Maintenant nous pouvons procéder à la notion d'ultraproduit. Les données sont les suivantes. Nous nous fixons un langage $\mathcal{L}$, donc une signature qui lui correspond et une famille de $\mathcal{L}$-structures $\{\mathcal{M}_i | i \in I\}$ chacune d'univers M_i. A ces données nous ajoutons celle d'un ultrafiltre $\mathcal{U}$ sur les parties de I. Nous construirons une nouvelle $\mathcal{L}$-structure $\mathcal{M}$ dite l'*ultraproduit des $\mathcal{M}_i$ par l'ultrafiltre $\mathcal{U}$* qui sera notée

$$\prod_{i \in I} \mathcal{M}_i / \mathcal{U} \quad .$$

Pour obtenir l'ensemble sous-jacent de $\mathcal{M}$, on procède la façon suivante. Sur le produit cartésien $\prod_{i \in I} M_i$ on définit une relation d'équivalence. Si $m : I \longrightarrow \bigcup_{i \in I} M_i$ et $m' : I \longrightarrow \bigcup_{i \in I} M_i$ sont deux éléments de $\prod_{i \in I} M_i$, alors m et m' sont équivalentes si et seulement si l'ensemble $\{i \in I | m(i) = m'(i)\} \in \mathcal{U}$. Une expression informelle serait de dire que deux I-uplets sont égaux si et seulement s'ils s'entendent sur un large ensemble d'indices. Cette relation est réflexive puisque $I \in \mathcal{U}$, trivialement symétrique, et transitive puisque $\mathcal{U}$ est clos par rapport aux intersections finies.

Le terrain pour introduire l'univers M de l'ultraproduit $\mathcal{M} = \prod_{i \in I} \mathcal{M}_i / \mathcal{U}$ a été préparé dans le paragraphe précédent. Ce sera les classes d'équivalence de la relation d'équivalence qui vient d'être introduite. Pour chaque $m \in \prod_{i \in I} M_i$, cette classe sera notée $[m]$. Sur cette base, nous construirons maintenant une nouvelle structure à partir des $\mathcal{M}_i$. La signature est définie comme suit.

Les constantes de $\mathcal{M}$

Chaque symbole de constante $c \in \mathcal{L}$ est interprété par un élément $c^{\mathcal{M}_i}$ de M_i. La classe d'équivalence du I-uplet formé par toutes ces constantes sera l'interprétation $c^{\mathcal{M}}$ de c. Formellement, si $c_{\mathcal{M}}$ est l'application de I dans $\bigcup_{i \in I} M_i$ qui associe l'élément $c^{\mathcal{M}_i}$ à chaque i alors $c^{\mathcal{M}} = [c_{\mathcal{M}}]$. Cette définition est sans ambiguïté vu qu'il n'y a pas de choix de représentant pour les valeurs.

Les fonctions de $\mathcal{M}$

Le principe sera le même que pour les constantes mais sa mise en place sera plus compliquée non parce qu'elle nécessite des notions profondes mais que la notation est inévitablement compliquée. Soit donc f un symbole de fonction de arité $k \in \mathbb{N}^*$. Alors,

$$\begin{array}{rccc} f^{\mathcal{M}} & : & (\prod_{i \in I} M_i / \mathcal{U})^k & \longrightarrow & \prod_{i \in I} M_i / \mathcal{U} \\ & & ([m_1], \ldots, [m_k]) & \longmapsto & [f_{\mathcal{M}}(m_1, \ldots, m_k)] \end{array}$$

où $f_{\mathcal{M}}(m_1, \ldots, m_k)$ est l'élément suivant de $\prod_{i \in I} M_i$:

$$\begin{array}{rccc} f_{\mathcal{M}}(m_1, \ldots, m_k) & : & I & \longrightarrow & \bigcup_{i \in I} M_i \\ & & i & \longrightarrow & f^{\mathcal{M}_i}(m_1(i), \ldots, m_k(i)) \end{array} \quad .$$

Vérifions que la définition n'est pas ambiguë. En effet, pour tout $j \in \{1,\ldots,k\}$, m'_j est équivalent à m_j si et seulement si

$$U_j = \{i \in I | m_j(i) = m'_j(i)\} \in \mathcal{U}$$

pour tout $j \in \{1,\ldots,k\}$. Or $\bigcap_{j=1}^k U_j \in \mathcal{U}$, et si $i \in \bigcap_{j=1}^k U_j$, alors pour tout $j \in \{1,\ldots,k\}$, $m_j(i) = m'_j(i)$. Ceci équivaut à dire que les I-uplets $f_{\mathcal{M}}(m_1,\ldots,m_k)$ et $f_{\mathcal{M}}(m'_1,\ldots,m'_k)$ sont équivalents.

Les relations de $\mathcal{M}$

Soit R un symbole de relation k-aire ($k \in \mathbb{N}^*$) dans $\mathcal{L}$. Pour tout $([m_1],\ldots,[m_k]) \in (\prod_{i\in I} M_i/\mathcal{U})^k$,

$$([m_1],\ldots,[m_k]) \in R^{\mathcal{M}} \text{ si et seulement si } \{i \in I | (m_1(i),\ldots,m_k(i)) \in R^{\mathcal{M}_i}\} \in \mathcal{U} \ .$$

Encore une fois il faut vérifier que la définition ne dépend pas des représentants des classes d'équivalences. En effet, si pour tout $j \in \{1,\ldots,k\}$, m'_j est équivalent à m_j, alors le même raisonnement que celui nous avons détaillé dans le cas des fonctions montre que

$$\{i \in I \mid \mathcal{M}_i \models (R(x_1,\ldots,x_k) \leftrightarrow R(y_1,\ldots,y_k))[(m_1(i),\ldots,m_k(i),m'_1(i),\ldots,m'_k(i)] \ \} \in \mathcal{U} \ .$$

Nous venons de compléter la définition d'un ultraproduit de structures en toute généralité. Si les structures $\mathcal{M}_i$ dans la définition sont toutes la même, alors nous parlerons d'une *ultrapuissance*.

L'apparente complexité dans la définition d'un ultraproduit n'est causée que par la notation. La notion elle-même est suffisamment naturelle pour faire partie du patrimoine mathématique comme nous l'avons déjà indiqué. Avant de démontrer le lien fodamental entre la notion d'ultraproduit et la théorie des modèles, vérifions une remarque que nous avons faite sur les ultrafiltres principaux.

Proposition 6.3.4 *Soit $\{\mathcal{M}_i | i \in I\}$ une famille de structures de même signature. Si, pour un indice fixé $j \in I$, $\mathcal{U}_j$ est le filtre principal des parties de I contenant j, alors $\prod_{i\in I} \mathcal{M}_i/\mathcal{U}_j$ est isomorphe à $\mathcal{M}_j$.*

Preuve. Conformément au choix de notation dans cette section, nous posons dès le début

$$\mathcal{M} = \prod_{i\in I} \mathcal{M}_i/\mathcal{U}_j \ .$$

L'application proposée pour arriver à un isomorphisme est la suivante :

$$\begin{array}{rccc} \nu & : & \prod_{i\in I} M_i/\mathcal{U}_j & \longmapsto & M_j \\ & & [m] & \longrightarrow & m(j) \end{array} \ .$$

L'application ν est définie sans ambiguïté parce que si m est équivalent à m', alors $m(j) = m'(j)$ en raison de la définition de l'ultrafiltre $\mathcal{U}_j$. Clairement, l'application est surjective. Elle est aussi injective parce que si pour deux éléments $[m]$, $[m'] \in \mathcal{M}$, $m(j) = m'(j)$ alors m et m' s'entendent sur un ensemble large d'indices, en d'autres termes sur une partie de I contenant j. Tant il est facile de s'entendre dans le cas de ultrafiltres principaux...

En ce qui concerne les propriétés homomorphiques, le cas des constantes est clair suivant la définition que nous avons donnée des ultraproduits. Soit f un symbole de fonction $k$aire avec $k \in \mathbb{N}^*$. Nous utiliserons le symbolisme que nous avons introduit en définissant l'interprétation des des symboles de fonction dans $\mathcal{M}$. Alors, pour $([m_1],\ldots,[m_k]) \in (\prod_{i\in I} M_i/\mathcal{U}_j)^k$,

$$\begin{array}{rcl} \nu(f^{\mathcal{M}}([m_1],\ldots,[m_k])) & = & \nu([f_{\mathcal{M}}(m_1,\ldots,m_k)]) \\ & = & f_{\mathcal{M}}(m_1,\ldots,m_k)(j) \\ & = & f^{\mathcal{M}_j}(m_1(j),\ldots,m_k(j)) \\ & = & f^{\mathcal{M}_j}(\nu([m_1]),\ldots,\nu([m_k])) \ . \end{array}$$

Finalement, soit R un symbole de relation $k$aire avec $k \in \mathbb{N}^*$. Alors,

$$\begin{array}{lll} ([m_1],\ldots,[m_k]) \in R^{\mathcal{M}} & \text{si et seulement si} & j \in \{i \in I | (m_1(i),\ldots,m_k(i)) \in R^{\mathcal{M}_i}\} \\ & \text{si et seulement si} & (m_1(j),\ldots,m_k(j)) \in R^{\mathcal{M}_j} \\ & \text{si et seulement si} & (\nu([m_1]),\ldots,\nu([m_k])) \in R^{\mathcal{M}_j} \ . \end{array}$$

Cette dernière équivalence finit la preuve. □

Corollaire 6.3.5 *L'ultraproduit d'un nombre fini de structures par un ultrafiltre est isomorphe à l'une de ces structures.*

Le théorème suivant est fondamental pour faire le lien entre les ultraproduits et la théorie des modèles. Sa preuve est un simple raisonnement par récurrence sur la complexité des formules du premier ordre. Son apparence compliquée est due aux indices. Nous suivrons les notations de cette section.

Théorème 6.5 (Théorème de Los) *Soient $\{\mathcal{M}_i | i \in I\}$ une famille de structures de même signature, le langage correspondant étant $\mathcal{L}$, et $\mathcal{U}$ un ultrafiltre des parties de I. Nous posons $\mathcal{M} = \prod_{i\in I} \mathcal{M}_i/\mathcal{U}$. Pour toute formule $\phi(x_1,\ldots,x_k)$ dans le langage $\mathcal{L}$,*

$$\mathcal{M} \models \phi[([m_1],\ldots,[m_k])] \quad \textit{si et seulement si} \quad \{i \in I | \mathcal{M}_i \models \phi[(m_1(i),\ldots,m_k(i))]\} \in \mathcal{U} \ .$$

Preuve. La preuve est par récurrence sur la complexité des formules du premier ordre. La définition de l'ultraproduit montre qu'il n'y a presque rien à faire pour les formules atomiques. Le seul point qui manque est de généraliser aux termes l'étude que nous avons faite de l'évaluation des symboles de fonctions. Ceci utilisant une autre récurrence sur la complexité de l'écriture, nous ne la laissons même pas comme exercice à nos lecteurs. Alors commence l'étape inductive.

Si $\phi(x_1,\ldots,x_k)$ est de la forme $\neg\psi(x_1,\ldots,x_k)$, alors pour tout $([m_1],\ldots,[m_k]) \in M^k$, $\mathcal{M} \models \phi[([m_1],\ldots,[m_k])]$ si et seulement si $([m_1],\ldots,[m_k])$ ne satisfait pas $\psi(x_1,\ldots,x_k)$ dans $\mathcal{M}$. Par récurrence, ceci équivaut à $I \setminus \{i \in I \mid \mathcal{M}_i \models \psi[m_1(i),\ldots,m_k(i)] \} \in \mathcal{U}$. De manière équivalente, $\{i \in I \mid \mathcal{M}_i \models \phi[m_1(i),\ldots,m_k(i)] \} \in \mathcal{U}$.

Si ϕ est de la forme $\theta \wedge \psi$ alors on répète le même style de raisonnement que dans le paragraphe précédent mais cette fois-ci en utilisant le fait qu'un filtre soit clos par rapport aux intersections finies. Il reste donc à étudier les formules avec quantificateurs.

Si ϕ est de la forme $\exists x\psi$, alors $\mathcal{M} \models \phi[([m_1],\ldots,[m_k])]$ si et seulement s'il existe $[\beta] \in M$ tel que $\mathcal{M} \models \psi[([m_1],\ldots,[m_k],[\beta])]$ si et seulement si, par récurrence,

$$\{i \in I \mid \mathcal{M}_i \models \psi[(m_1(i),\ldots,m_k(i),\beta(i))]\} \in \mathcal{U} \ .$$

Or, cette partie de I est contenue dans

$$\{i \in I \mid \mathcal{M}_i \models \exists x\psi[(m_1(i),\ldots,m_k(i))] \} \ .$$

En conséquence,

$$\{i \in I \mid \mathcal{M}_i \models \phi[(m_1(i),\ldots,m_k(i))] \} \in \mathcal{U} \ .$$

Inversement, si

$$A = \{i \in I \mid \mathcal{M}_i \models \exists x\psi[(m_1(i),\ldots,m_k(i))] \} \in \mathcal{U} \ ,$$

alors on définit un élément $\beta \in \prod_{i\in I} M_i$. Si $i \in A$, on pose $\beta(i)$ égal à un élément de M_i tel que $\mathcal{M}_i \models \psi[(m_1(i),\ldots,m_k(i),\beta(i))]$; sinon, $\beta(i)$ est arbitrairement choisi. Comme $A \in \mathcal{U}$, cette définition donne un unique élément dans $\mathcal{M}$ tel que $\mathcal{M} \models \psi[([m_1],\ldots,[m_k],[\beta])]$. Le reste découle des équivalences déjà établies. □

Corollaire 6.3.6

1. On retient les hypothèses et la notation du théorème 6.5. Un énoncé σ est vrai dans l'ultraproduit $\prod_{i\in I}\mathcal{M}_i/\mathcal{U}$ si et seulement si $\{i \in I | \mathcal{M}_i \models \sigma\} \in \mathcal{U}$.

2. Soient I un ensemble non vide, $\mathcal{U}$ un ultrafiltre des parties de I et $\mathcal{M}$ une structure d'univers M. Le plongement diagonal *de $\mathcal{M}$ dans l'ultrapuissance $\prod_{i\in I}\mathcal{M}/\mathcal{U}$ qui associe à chaque élément de $m \in M$ la classe d'équivalence du I-uplet dont toutes les "coordonnées" sont égales à m est élémentaire.*

Nous n'avons toujours pas démontré le théorème de compacité. Pourtant toute la préparation était destinée vers cet objectif :

Preuve du Théorème 6.1. Soit E un ensemble de $\mathcal{L}$-énoncés dont chaque partie finie est consistante. On pose alors I l'ensemble des parties finies de E. Pour tout $i \in I$, il existe alors une structure $\mathcal{M}_i$ telle que $\mathcal{M}_i \models i$. Nous avons vu dans l'exemple 6.3.1 (2) que les ensembles $I_i = \{j \in I | i \subset j\}$ forment une base de filtre des parties de I. Soit $\mathcal{U}$ un ultrafiltre contenant cette base. Notre candidat de modèle pour l'ensemble E est alors

$$\prod_{i\in I}\mathcal{M}_i/\mathcal{U} \ .$$

En effet, si σ est un énoncé de E, alors pour tout $j \in I_{\{\sigma\}}$, $\mathcal{M}_j \models \sigma$. Alors, le point (1) du corollaire 6.3.6 montre que

$$\prod_{i\in I}\mathcal{M}_i/\mathcal{U} \models \sigma \ .$$

□

Finissons cette section avec un exemple algébrique :

Exemple 6.3.2 Soient G un groupe et $\mathcal{K}$ une famille infinie de sous-groupes distingués de G. Supposons que pour toute partie finie de G, il existe un membre de $\mathcal{K}$ qui sépare les éléments de cette partie finie. En d'autres termes, si X est une partie finie de G, il existe $N_X \lhd G$ tel que si g et h sont deux éléments distincts de X alors $gN_X \neq hN_X$. Sous cette hypothèse, on peut montrer que G se plonge dans un ultraproduit des quotients de la forme G/N avec $N \in \mathcal{K}$. Pour ce faire, il suffit de considérer comme ensemble d'indices I, les parties finies de la famille $\mathcal{K}$. Alors les sous-ensembles de I de la forme $I_{\{N_{i_1},\ldots,N_{i_r}\}}$, en d'autres termes les parties finies $\mathcal{K}$ contenant $\{N_{i_1},\ldots,N_{i_r}\}$ forment une base de filtre des parties de I. Si $\mathcal{U}$ est un filtre contenant cette base, alors G se plonge dans l'ultraproduit

$$\prod_{\{i_1,\ldots,i_r\}\in I} \left(G/N_{i_1} \times \ldots \times G/N_{i_r}\right)/\mathcal{U} \ .$$

Pour concrétiser la recette générale du premier paragraphe, nous conseillons de poser $G = (\mathbb{Z}, +)$, et de choisir $\mathcal{K} = \{p\mathbb{Z} \mid p \text{ premier}\}$.

6.4 Un espace compact

Qu'est-ce que le théorème de compacité sinon de vérifier qu'un espace topologique jouit de cette propriété ? Dans cette section nous établirons le cadre topologique pour ce faire et y appliquerons le théorème de compacité. Plus tard, cette approche topologique réapparaîtra dans le contexte des *types*.

Nous fixons un langage $\mathcal{L}$ et considérons l'ensemble $\mathcal{T}$ de toutes les théories complètes dans ce langage. Cet ensemble $\mathcal{T}$ peut être muni d'une topologie en suivant une approche dont des analogues sont souvent rencontrées dans diverses disciplines mathématiques telles que les algèbres de Boole, les spectres des anneaux, la topologie de Zariski, etc.

Pour chaque énoncé τ dans $\mathcal{L}$, nous définissons $\langle\tau\rangle$ comme l'ensemble de toutes les théories complètes dans $\mathcal{L}$ qui contiennent τ. La notation $[\tau]$ est souvent utilisée aussi mais nous suivons

Poizat parce que [ ] a été déjà mis en service pous plusieurs objectifs. Il convient de souligner une équivalence évidente : pour tout $T \in \mathcal{T}$, $T \in \langle\tau\rangle$ si et seulement si $\tau \in T$ si et seulement si $T \vdash \tau$. Les propriétés booléennes des énoncés décrivent les bases du cadre topologique. En effet, comme $T \vdash \sigma \wedge \tau$ équivaut à $T \in \langle\sigma\rangle \cap \langle\tau\rangle$ pour toute théorie T et tous énoncés σ, τ dans le langage $\mathcal{L}$, la famille des parties de $\mathcal{T}$ de la forme $\langle\sigma\rangle$ forme une base d'ouverts pour une topologie.

Certaines propriétés de la topologie dont $\mathcal{T}$ a été muni dans le paragraphe précédent sont rapidement vérifiées. L'espace $\mathcal{T}$ est *séparé*. En effet, deux théories complètes dans T et T' dans le langage $\mathcal{L}$ sont distincts distincts si et seulement s'il existe un énoncé σ tel que $T \vdash \sigma$ et $T' \vdash \neg\sigma$ dans lequel cas les ouverts disjoints $\langle\sigma\rangle$ et $\langle\neg\sigma\rangle$ séparent les deux points. Soulignons que la nécessité de l'existence d'un tel énoncé σ utilise la complétude des théories T et T'. Un raisonnement dans le même esprit montre que $\mathcal{T}$ est *totalement discontinu* : il admet une base dont les membres sont ouverts et fermés. Pour ce faire, il suffit de voir que si σ est un énoncé dans $\mathcal{L}$, alors $\mathcal{T}$ est l'union disjointe de $\langle\sigma\rangle$ et $\langle\neg\sigma\rangle$. Il en découle, en utilisant un raisonnement élémentaire de la topologie générale que toute partie connexe non vide de $\mathcal{T}$ est un singleton. Une autre propriété importante qui est vérifiée par le biais d'un raisonnement élémentaire est que les *singletons de $\mathcal{T}$ sont fermés*. Il suffit de constater que si $T \in \mathcal{T}$ alors

$$\{T\} = \bigcap_{\tau \in T} \langle\tau\rangle \ .$$

En général, ils ne sont pas ouverts ; il existe des théories qui ne sont pas finiment axiomatisables. Néanmoins, l'étude des points isolés fournit une grande quantité d'informations surtout dans le contexte des types.

Quels sont les ouverts et les fermés de $\mathcal{T}$ en général ? Par définition, tout ouvert est l'union des ensembles de type $\langle\tau\rangle$ avec τ un énoncé dans $\mathcal{L}$. Comme un ensemble est fermé si et seulement si son complémentaire est ouvert, les fermés sont l'intersection des ensembles de type $\langle\tau\rangle$. Cette caractérisation anodine a une interprétation plus parlante du point de vue de la théorie des modèles. Les fermés non vides correspondent aux théories, non nécessairement complètes, dans $\mathcal{L}$. En effet, $F = \cap_{i\in I}\langle\sigma_i\rangle$ est un fermé non vide de $\mathcal{T}$ si et seulement si toute théorie complète dans F contient chaque σ_i et les conséquences de celui-ci. En conséquence une théorie complète appartient à F si et seulement si elle contient la théorie T_F axiomatisée par les σ_i. En d'autres termes, F est l'ensemble des théories complètes dans $\mathcal{L}$ qui contiennent T_F.

Maintenant nous pouvons démontrer le fait non trivial à propos de la topologie de $\mathcal{T}$:

Théorème 6.6 *L'espace $\mathcal{T}$ est un espace topologique compact.*

Preuve. Nous utiliserons le mot compact pour les espaces dont les singletons sont fermés et chaque recouvrement par des ouverts a un sous-recouvrement fini. Souvent, le mot compact est utilisé pour la deuxième propriété.

Nous avons déjà vérifié que les singletons sont fermés. Une forme équivalente de la deuxième propriété énonce que si $\mathcal{F}$ est une famille de parties fermés de $\mathcal{T}$ tels que pour tous $F_1, \ldots, F_m \in \mathcal{F}$ l'intersection $F_1 \cap \ldots \cap F_m \neq \emptyset$, alors l'intersection de tous les membres de $\mathcal{F}$ est non vide. C'est ce que nous vérifierons en utilisant le théorème de compacité.

Nous commençons avec une famille $\mathcal{F}$ de fermés qui vérifient la condition sur les intersections finies. D'après la description que nous avons donnée des fermés de $\mathcal{T}$, chaque fermé dans $\mathcal{F}$ s'écrit comme l'intersection des fermés basiques $\langle\sigma\rangle$. Nous définissons $\tilde{\mathcal{F}}$ comme la famille formée par tous ces fermés basiques. Si $\{\langle\sigma_{i_1}\rangle, \ldots, \langle\sigma_{i_k}\rangle\}$ est une partie finie $\{F_1, \ldots, F_k\}$ de $\tilde{\mathcal{F}}$ telle que chaque $F_j \subset \langle\sigma_{i_j}\rangle$. Il s'ensuit que l'intersection

$$\langle\sigma_{i_1}\rangle \cap \ldots \cap \langle\sigma_{i_k}\rangle \supset F_1 \cap \ldots \cap F_k \neq \emptyset \ .$$

Ainsi, $\tilde{\mathcal{F}}$ jouit de la même propriété d'intersections finies non vides.

Pour préciser la notation, posons $\tilde{\mathcal{F}} = \{\langle\sigma_i\rangle | i \in I\}$. Alors, pour tous $i_1, \ldots, i_k \in I$, l'intersection fermée $\langle\sigma_{i_1}\rangle \cap \ldots \cap \langle\sigma_{i_k}\rangle$ est non vide. Or, la discussion générale que nous avons faite des fermés de l'espace $\mathcal{T}$ montre que cette intersection est l'ensemble des théories complètes

dans $\mathcal{L}$ qui contiennent la théorie aximoatisée par $\{\sigma_{i_1}, \ldots, \sigma_{i_k}\}$. L'intersection étant non vide, l'ensemble $\{\sigma_{i_1}, \ldots, \sigma_{i_k}\}$ est consistant. L'ensemble $\{\langle\sigma_{i_1}\rangle, \ldots, \langle\sigma_{i_k}\rangle\}$ étant arbitrairement choisi parmi les parties finies de $\tilde{\mathcal{F}}$, d'après le théorème de compacité, l'ensemble $\{\sigma_i | i \in I\}$ est consistant. Il en découle que la théorie axiomatisée par cet ensemble est consistante. En conséquence, l'intersection $\bigcap_{i\in I}\langle\sigma_i\rangle$ est non vide. Or, d'après notre définition de $\tilde{\mathcal{F}}$,

$$\bigcap_{F\in\mathcal{F}} F = \bigcap_{i\in I}\langle\sigma_i\rangle \ .$$

Le théorème est démontré. □

Un corollaire du théorème caractérise les fermés basiques de $\mathcal{T}$.

Corollaire 6.4.1 *Les ouverts-fermés de $\mathcal{T}$ sont les ensembles de la forme $\langle\tau\rangle$ avec τ un énoncé dans le langage $\mathcal{L}$.*

Cette section présente le début d'une approche topologique à la théorie des modèles. Les notions fondamentales de la topologie y trouvent leurs places pour offrir des synonymes topologiques à des notions fondamentales de la théorie des modèles. Il n'est pas nécessaire de suivre l'approche topologique mais certains mathématiciens la trouvent plus intuitive et parfois leurs poursuites fournissent des interactions surprenantes. Les travaux de Ludomir Newelski en sont un très bon exemple.

Chapitre 7

Types et saturation

La notion de *type* et celle de *modèle κ-saturé* sont parmi les plus importantes de la théorie des modèles. Les deux notions fournissent les outils primordiaux de l'étude d'une théorie du premier ordre. Un type est, d'une certaine optique, une formule de taille infinie, et d'une autre, une théorie complète du premier ordre quitte à ajouter au langage ambiant des symboles de constantes pour remplacer les formules par des énoncés. Un modèle κ-saturé est un modèle riche en certains types de types. Il est donc naturel que ces notions soient indispensables pour comprendre les théories élémentaires des structures. Dans les chapitres suivants, par le biais des exemples et d'autres résultats fondamentaux de la théorie des modèles, nous ne cesserons de rencontrer ces deux notions.

7.1 Types

Dans cette section et la prochaine, nous introduirons la notion fondamentale de *type*. Dans un premier temps nous étudierons ce qui est dit "un type sur l'ensemble vide" et la définition générale viendra après la discussion de la notion de *paramètre*. Depuis longtemps nous utilisons des paramètres sans les nommer ainsi mais l'introduction de la notion de type en deux étapes peut en faciliter la digestion.

Dans toute cette section, nous aurons un langage $\mathcal{L}$ fixé et les structures étudiées seront des $\mathcal{L}$-structures. La notion de type peut être vue comme une généralisation d'une formule du premier ordre à un contexte où il existe une infinité de formules satisfaites par un même k-uplet d'une $\mathcal{L}$-structure. Pour maîtriser des ensembles infinis de formules du premier ordre satisfaites par un même k-uplet, le théorème de compacité sera l'outil principal d'une façon similaire à son usage pour vérifier la consistance des ensembles infinis d'énoncés. En effet, un type à k-variables n'est qu'une généralisation de la notion de théorie complète à un langage élargi en y ajoutant k symboles pour nommer la "solution commune".

Commençons par introduire la notion de *réalisation/omission* d'une famille de formules du premier ordre.

Définition 7.1.1 *Soient $\Phi(x_1, \ldots, x_k)$ une famille de $\mathcal{L}$-formules du premier ordre dont les variables libres sont parmi $\{x_1, \ldots, x_k\}$ et $\mathcal{M}$ une $\mathcal{L}$ structure d'univers M. L'ensemble $\Phi(x_1, \ldots, x_k)$ est dit* réalisé *par un k-uplet $(m_1, \ldots, m_k) \in M^k$ dans $\mathcal{M}$ si pour toute formule $\phi(x_1, \ldots, x_k) \in \Phi(x_1, \ldots, x_k)$,*

$$\mathcal{M} \models \phi[(m_1, \ldots, m_k)] \ .$$

On dit aussi que $\mathcal{M}$ réalise $\Phi(x_1, \ldots, x_k)$ si un tel k-uplet existe dans M^k. On dit que $\mathcal{M}$ omet *$\Phi(x_1, \ldots, x_k)$ si M^k ne contient pas un tel k-uplet.*

Il convient de souligner que, dans la notation de la définition 7.1.1, le fait que $(m_1, \ldots, m_k)$ réalise $\Phi(x_1, \ldots, x_k)$ dans $\mathcal{M}$ équivaut à ce que la structure $(\mathcal{M}, m_1, \ldots, m_k)$ soit un modèle de l'ensemble d'*énoncés* $\Phi(x_1, \ldots, x_k)$ dans le langage $\mathcal{L}^+ = \mathcal{L} \cup \{x_1, \ldots, x_k\}$ où les symboles

$x_1, \ldots, x_k$ nomment $a_1, \ldots, a_k$ respectivement. La proposition suivante renforce ce lien avec les chapitres précédents en utilisant notre outil indispensable à savoir le théorème de compacité :

Proposition 7.1.2 *Soient T une théorie et $\Phi(x_1, \ldots, x_k)$ un ensemble de formules dont les variables libres sont parmi $\{x_1, \ldots, x_k\}$. Alors, les énoncés suivants sont équivalents :*

1. *T a un modèle qui réalise $\Phi(x_1, \ldots, x_k)$.*
2. *Toute partie finie de $\Phi(x_1, \ldots, x_k)$ est réalisée dans un certain modèle de T.*
3. *$T \cup \{\exists x_1 \ldots x_k(\phi_1 \wedge \ldots \wedge \phi_m) : \phi_1, \ldots, \phi_m \in \Phi(x_1, \ldots, x_k)\}$ est un ensemble consistant d'énoncés.*

Preuve. Clairement le premier point implique le deuxième. Le troisième se montre à partir du deuxième en appliquant le théorème de compacité. En effet, si T_0 est une partie finie de $T \cup \{\exists x_1 \ldots x_k(\phi_1 \wedge \ldots \wedge \phi_m) : \phi_1, \ldots, \phi_m \in \Phi(x_1, \ldots, x_k)\}$, alors on peut supposer que $T_0 = (T_0 \cap T) \cup \{\exists x_1 \ldots x_k(\phi_1 \wedge \ldots \wedge \phi_m)\}$ pour un certain $m \in \mathbb{N}$. Par hypothèse, il existe un modèle de T qui réalise $\{\phi_1(x_1, \ldots, x_k), \ldots, \phi_m(x_1, \ldots, x_k)\}$. Ce modèle est un modèle de T_0. Par conséquent, T_0 est consistant.

Quant à l'implication du point (1) par le point (3), il s'agit d'une autre application du théorème de compacité suivant une ligne conforme à celle du paragraphe qui précède la présente proposition. On pose $\mathcal{L}^+ = \mathcal{L} \cup \{c_1, \ldots, c_k\}$. Ensuite, on considère l'ensemble d'énoncés $T^+ = T \cup \{\phi(c_1, \ldots, c_k) | \phi(x_1, \ldots, x_k) \in \Phi(x_1, \ldots, x_k)\}$. Par hypothèse, toute partie finie de cet ensemble a un modèle, en d'autres termes elle est consistante. Par compacité, T^+ est consistant. Or, tout modèle de T^+ réalise $\Phi(x_1, \ldots, x_k)$. Le réduit d'un tel modèle au langage $\mathcal{L}$ est un modèle de T qui réalise $\Phi(x_1, \ldots, x_k)$. □

Définition 7.1.3 *1. Soit T une théorie. Un ensemble de formules tel que $\Phi(x_1, \ldots, x_k)$ de la proposition 7.1.2 est dit* consistant *avec T.*

2. *Soit T une théorie complète. Un ensemble de $\mathcal{L}$-formules $p(x_1, \ldots, x_k)$ dont les variables libres sont parmi $\{x_1, \ldots, x_k\}$ est dit un* k-type de T *($k \in \mathbb{N}$) si $p(x_1, \ldots, x_k)$ est un ensemble de formules consistant avec T et qui est maximal par rapport à cette propriété. L'ensemble des k-types de T est noté $S_k(T)$. On note $S(T) = \bigcup_{k \in \mathbb{N}} S_k(T)$.*

Remarques : La maximalité d'un type p implique que $T \subset p$. Par ailleurs, un 0-type n'est qu'une théorie complète. Dans le même esprit, voire de manière équivalente, un k-type est une théorie complète dans le langage $\mathcal{L} \cup \{x_1, \ldots, x_k\}$.

Exemples 7.1.1 1. Soient $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M, et $(m_1, \ldots, m_k) \in M^k$

$$\mathrm{tp}_{\mathcal{M}}(m_1, \ldots, m_k) = \{\phi(x_1, \ldots, x_k) | \mathcal{M} \models \phi[(m_1, \ldots, m_k)]\} .$$

Notons que cet exemple est d'une certaine manière le seul exemple de type à ceci près que la définition d'un type de la forme $\mathrm{tp}_{\mathcal{M}}(m_1, \ldots, m_k)$ se base sur une réalisation qui existe déjà dans le modèle ambiant, ce qui n'est pas nécessairement le cas. Or, comme l'a indiqué le lemme 7.1.2 et le précisera le lemme 7.1.4, une réalisation existe dans une extension élémentaire qui, une fois trouvée, peut être prise comme $\mathcal{M}$.

2. Un exemple bien connu de type non réalisé est le suivant. Soit $\mathcal{L} = \{+, -, ., 0, 1\}$ le langage des anneaux. On considère les corps algébriquement clos d'une caractéristique fixée en tant que $\mathcal{L}$-structures. Soit $K = \mathbb{Q}$ ou $K = \mathbb{F}_p$ pour un nombre premier p et $\overline{K}$ la clôture algébrique de K. L'ensemble des inéquations de la forme

$$a_0 + a_1 x + \ldots + a_n x^n \neq 0$$

où les polynômes ont des coefficients dans $\mathbb{Z}$ ou dans $\mathbb{F}_p$ est consistant. En utilisant le lemme de Zorn, on considère un 1-type $p(x)$ qui le contient. Un tel type n'est pas réalisé dans $\overline{K}$. Par

contre tout élément transcendant sur K est une réalisation de ce type. Si α est un tel élément et $\overline{K(\alpha)}$ est la clôture algébrique du corps engendré par K et α, alors

$$p(x) = \mathrm{tp}_{\overline{K(\alpha)}}(\alpha) \ .$$

3. Soit $\mathcal{L} = \{E\}$ où E est un symbole de relation binaire. La théorie T qui énonce que E est une relation d'équivalence avec une classe de cardinal n et une seule pour chaque $n \in \mathbb{N}^*$ est complète. Dans ce langage, il est possible d'écrire une formule du premier ordre à une seule variable $C_n(x)$ qui énonce que x appartient à la classe à n éléments. Alors, par compacité, l'ensemble

$$I = \{\neg C_n(x) | n \in \mathbb{N}^*\}$$

est consistant avec T. Si $\mathcal{M}$ est un modèle qui n'a que des classes finies d'équivalence, alors un 1-type $p(x)$ contenant l'ensemble I n'est pas réalisé dans $\mathcal{M}$. De l'autre côté, si $\mathcal{N}$ est une extension élémentaire de $\mathcal{M}$ qui a au moins une classe infinie, alors tout élément de l'univers de $\mathcal{N}$ qui n'est pas dans $\mathcal{M}$ réalise l'ensemble I. En fait, si α est un tel élément "transcendant", alors $\mathcal{M}$ a une extension élémentaire dénombrable $\mathcal{N}$ qui contient α. En utilisant les automorphismes de $\mathcal{N}$, on conclut que

$$p(x) = \mathrm{tp}_{\mathcal{N}}(\alpha) \ .$$

Comme c'est indiqué dans le point (1) de l'exemple 7.1.1 $\mathrm{tp}_{\mathcal{M}}(m_1, \ldots, m_k)$ est le seul "type" de type qui existe, quitte à se mettre dans un modèle contenant une réalisation de l'ensemble initial donné. Ceci était déjà visible dans la proposition 7.1.2 mais il est possible de préciser davantage :

Lemme 7.1.4 *Soit $\mathcal{M}$ une $\mathcal{L}$-structure et $p(x_1, \ldots, x_k)$ un k-type de $\mathrm{Th}(\mathcal{M})$. Alors, $\mathcal{M}$ a une extension élémentaire dont l'univers est de même cardinal que celui de $\mathcal{M}$ et qui contient une réalisation de p.*

Preuve. Il suffit d'adapter la preuve de la proposition 7.1.2 pour assurer que la réalisation recherchée apparaîtra dans une extension élémentaire. Donc, on ajoute au langage $\mathcal{L}$ un symbole de constante pour chaque élément de M et k symboles de constante pour nommer les éventuelles réalisations de p, et on vérifie que

$$T^+ = \mathrm{Th}((\mathcal{M}, m)_{m \in M}) \cup \{\phi(c_1, \ldots, c_k) | \phi \in p\} \ .$$

est consistant par un raisonnement de compacité. Clairement, la structure qui servira de modèle pour cette vérification est $\mathcal{M}$. Finalement, tout modèle de T^+ est une extension élémentaire de $\mathcal{M}$ qui contient une réalisation de p. Le théorème de Löwenheim-Skolem assure qu'il en existe un dont le cardinal de l'univers est égal à celui de $\mathcal{M}$. □

7.2 Paramètres

Dans cette section, nous introduirons d'une façon plus solide une méthode que nous avons déjà utilisée beaucoup de fois. Il s'agit de l'usage des *paramètres* dans les formules du premier ordre et en conséquence dans les structures où ces formules définissent des ensembles.

Comme d'habitude, nous travaillerons avec un langage $\mathcal{L}$ fixé et des $\mathcal{L}$-structures. L'une des expansions que nous avons le plus fréquemment utilisées, l'exemple le plus notable étant la méthode des diagrammes, a consisté à ajouter à $\mathcal{L}$ des symboles de constantes pour nommer des éléments de l'univers d'une structure fixée. Dans la méthode des diagrammes, nous avons en fait nommé tous les membres de l'univers. Nous commençons par mettre cette forme d'expansion dans un cadre particulier en introduisant une notation et une terminologie que nous utiliserons dans la suite :

Définition 7.2.1 *Soient $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M et $A \subset M$. Alors*

$$\mathcal{L}(A) = \mathcal{L} \cup \{c_a | a \in A\} \ .$$

La notation a déjà été utilisée dans le cas particulier où A est l'univers d'un modèle : $(\mathcal{M}, a)_{a\in A}$ est la $\mathcal{L}(A)$-structure qui correspond au langage élargi.

Les symboles de constantes ajoutés au langage $\mathcal{L}$ aussi bien que les éléments qu'ils sont censés nommer sont dits paramètres.

L'ajout des paramètres peut avoir des effets considérables. La théorie des ordres denses linéaires sans extrémités permet une illustration pertinente de ces effets.

Exemple 7.2.1 (Les ordres denses linéaires sans extrémités avec paramètres) Le langage $\mathcal{L}$ ne contient qu'un seul symbole $<$ de relation binaire. Nous considérons dans ce langage la théorie T des ordres denses linéaires sans extrémités. Nous avons vu dans le corollaire 6.2.3 que T est une théorie complète. En d'autres termes, $T = \mathrm{Th}((\mathbb{Q}, <))$. Considérons maintenant le langage $\mathcal{L}(\{0\})$ obtenu en ajoutant un symbole de constante $\{c_0\}$ pour nommer 0. La structure $((\mathbb{Q}, <), 0)$ a des propriétés de définissabilité qui n'étaient pas valables dans son réduit au langage $\mathcal{L}$. En effet, $((\mathbb{Q}, <), 0)$ connaît ses éléments positifs (resp. négatifs) :

$$\mathbb{Q}_+^* = \{q \in \mathbb{Q} \mid ((\mathbb{Q}, <), 0) \models \ (c_0 < x)[q] \ \} \ .$$

Nous avons déjà illustré dans les remarques suivant la proposition 5.6.2 une méthode pour montrer qu'il n'existe pas une formule miracle du premier ordre dans le langage $\mathcal{L}$ qui suffirait à définir l'ensemble des nombres rationnels positifs, ou de manière équivalente, qu'il n'existe pas de $\mathcal{L}$-formule $\psi(x)$ telle que

$$\mathrm{Th}((\mathbb{Q}, <), 0) \vdash \forall x(\psi(x) \leftrightarrow (c_0 < x)) \ .$$

La méthode consiste à trouver un automorphisme de $(\mathbb{Q}, <)$ qui ne stabilise pas $\mathbb{Q}_+^*$ et d'appliquer la proposition 5.6.2 (4). De la même proposition découle le constat important suivant : les automorphismes de la structure $((\mathbb{Q}, <), 0)$ sont les automorphismes de $(\mathbb{Q}, <)$ qui fixent 0.

La discussion du paragraphe précédent a des conséquences sur les types aussi. En effet, pour toute paire (q_1, q_2) de nombres rationnels $\mathrm{tp}_{(\mathbb{Q},<)}(q_1) = \mathrm{tp}_{(\mathbb{Q},<)}(q_2)$. Cette conclusion découle de l'étude des automorphismes de la structure $(\mathbb{Q}, <)$. Nous demandons à nos lecteurs de s'exercer à la vérifier en se basant sur le théorème 6.4. Par contre, suite à l'expansion à $((\mathbb{Q}, <), 0)$, chaque nombre rationnel sera la réalisation d'un des trois 1-types déterminés par les formules

$$x < c_0 \quad , \quad x = c_0 \quad , \quad c_0 < x \ .$$

Avant de considérer une nouvelle expansion, il convient de remarquer que cette description des 1-types reste inchangée si on considère des extensions élémentaires de $(\mathbb{Q}, <)$. Le lemme 7.2.3 mettra cette remarque dans le bon cadre.

Maintenant considérons le langage $\mathcal{L}(\{0, 1\})$, en d'autres termes, ajoutons un symbole de constante c_1 pour nommer 1. Le nouvel ajout augmente le nombre de parties qu'on peut définir comme le montre par exemple la formule

$$c_0 < x \wedge x < c_1$$

ou, la formule

$$c_1 < x$$

ou encore,

$$x = c_1 \ .$$

Aucune de ces formules n'a une équivalente dans le langage $\mathcal{L}$, ni dans le langage $\mathcal{L}(0)$. Nous laissons le soin de vérifier ceci à nos lecteurs.

Dans ce langage élargi, le nombre de types aussi augmente. En effet, les 1-types seront déterminés par l'appartenance aux intervalles ouverts

$$]\infty, 0[\ \ ,\ \ ]0, 1[\ \ ,\ \ ]1, +\infty[\ ,$$

et par être égal à 0 ou 1. La vérification de cette description exacte est un bon exercice pour lequel les automorphismes de $(\mathbb{Q}, <)$ qui fixent 0 et 1 sont des outils efficaces.

Maintenant, considérons le langage $\mathcal{L}(\mathbb{Z})$ et la structure $((\mathbb{Q}, c_k)_{k\in\mathbb{Z}})$. Avec l'expérience que nous avons accumulée, il n'est pas difficile de vérifier que beaucoup de nouveaux intervalles deviennent *définissables* dans la structure $((\mathbb{Q}, c_k)_{k\in\mathbb{Z}})$. Nous pouvons aussi parler des éléments "infinis". Considérons les deux ensembles de formules à une seule variable suivantes :

$$p_{+\infty}(x; c_k(k \in \mathbb{Z})) = \{c_k < x | k \in \mathbb{Z}\}$$

et

$$p_{-\infty}(x; c_k(k \in \mathbb{Z})) = \{x < c_k | k \in \mathbb{Z}\}$$

Le théorème de compacité montre que ces deux ensembles sont consistants avec

$$\mathrm{Th}(((\mathbb{Q}, <), c_k)_{k\in\mathbb{Z}})\ .$$

En appliquant la compacité soigneusement (le lemme 7.2.5), il est possible de vérifier que $((\mathbb{Q}, <), c_k)_{k\in\mathbb{Z}}$ a une extension élémentaire qui contient des réalisations de $p_{-\infty}$ et de $p_{+\infty}$. De même, la structure $(\mathbb{Q}, <)$ a une extension élémentaire contenant des éléments infinis. Il suffira d'oublier les paramètres. Toutes ces conclusions sont de bons exercices d'entraînement en compacité.

Quand on complète $p_{+\infty}$ (resp. $p_{-\infty}$) à des ensembles de formules à au plus une seule variable qui sont aussi consistants et maximaux par rapport à cette dernière propriété, on obtient deux types omis par $((\mathbb{Q}, <), c_k)_{k\in\mathbb{Z}}$. Ce sont des 1-*types à paramètres dans* $\mathbb{Z}$. C'est un bon exercice de vérifier que les complétés de $p_{+\infty}$ et de $p_{-\infty}$ sont uniques, et c'en est un autre de compter le cardinal de l'ensemble des types à paramètres dans $\mathbb{Z}$.

Maintenant, faisons un ajout spectaculaire, et considérons le langage $\mathcal{L}(\mathbb{Q})$. Il y a de nouveaux ensembles définissables et beaucoup de nouveaux types. En effet, en considérant les fractions rationnelles on conclut que chaque nombre réel α réalise dans $((\mathbb{R}, <), q)_{q\in\mathbb{Q}}$ un type avec paramètres dans $\mathbb{Q}$ qu'on peut appeler $p_\alpha(x; c_q(q \in \mathbb{Q}))$. Par ailleurs, en utilisant la densité des nombres rationnels dans $\mathbb{R}$, on conclut que pour deux nombres réels r, r', $r = r'$ si et seulement si $p_r = p_{r'}$. Ainsi, il existe au moins $2^{\aleph_0}$ types de $\mathrm{Th}(((\mathbb{R}, <), q)_{q\in\mathbb{Q}})$ avec des paramètres dans $\mathbb{Q}$. L'arithmétique cardinale montre alors qu'en fait c'est le cardinal exact.

L'exemple précédent non seulement illustre l'impact de l'usage des paramètres sur les théories mais aussi fournit des exemples de formules et de types à paramètres. Ceci motive la notion de *type sur un ensemble de paramètres*. C'est la définition générale de la notion de type qui inclut comme cas particulier la définition 7.1.3.

Définition 7.2.2 *Soient $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M et $A \subset M$. Un* type à paramètres dans A, *ou encore un* type sur A, *est un type de* $\mathrm{Th}((\mathcal{M}, a)_{a\in A})$. *Pour chaque $k \in \mathbb{N}$, on définit l'espace des k-types sur A :*

$$S_k^{\mathcal{M}}(A) = \{p(x_1, \ldots, x_k) \mid p \text{ est un } k\text{-type de } \mathrm{Th}((\mathcal{M}, a)_{a\in A})\}\ .$$

On pose alors $S^{\mathcal{M}}(A) = \bigcup_{k\in\mathbb{N}} S_k^{\mathcal{M}}(A)$.

Pour tout k-uplet $(m_1, \ldots, m_k) \in M^k$, le type de $(m_1, \ldots, m_k)$ sur A *est*

$$\mathrm{tp}_{\mathcal{M}}((m_1, \ldots, m_k)/A) = \{\phi(x_1, \ldots, x_k\ ;\ a_{i_1}, \ldots, a_{i_l}) | (\mathcal{M}, a)_{a\in A} \models \phi(x_1, \ldots, x_k; a_{i_1}, \ldots, a_{i_l})\}\ .$$

Le cas où $A = \emptyset$ correspond à la définition 7.1.3 et à l'exemple 7.1.1 (1).

Nous démontrons quelques lemmes simples mais cruciaux qui illustrent la démarche avec les paramètres. Les preuves, quoique simples, illustrent des techniques fréquentes que nous utiliserons plus tard avec moins de détails. Elles ont par ailleurs la vertu d'illustrer l'importance de la compacité et de la notion d'extension élémentaire.

Lemme 7.2.3 *Soient $\mathcal{L}$ un langage du premier ordre et $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M. Si $A \subset M$ et $\mathcal{M} \preceq \mathcal{N}$ alors pour tout $k \in \mathbb{N}$, $S_k^{\mathcal{M}}(A) = S_k^{\mathcal{N}}(A)$.*

Preuve. C'est clair puisque, A étant une partie de M et $\mathcal{N}$ une extension élémentaire de $\mathcal{M}$, $(\mathcal{M}, a)_{a\in A} \preceq (\mathcal{N}, a)_{a\in A}$. Alors $\mathrm{Th}((\mathcal{M}, a)_{a\in A}) = \mathrm{Th}((\mathcal{N}, a)_{a\in A})$ □

Le théorème de compacité permet d'étendre un type :

Lemme 7.2.4 *Soient $\mathcal{L}$ un langage du premier ordre et $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M. Si $A \subset B \subset M$ alors pour tout k-type $p \in S_k^{\mathcal{M}}(A)$, il existe un type $q \in S_k^{\mathcal{M}}(B)$ tel que $p \subset q$.*

Preuve. C'est une application du théorème de compacité. Il suffira de montrer que p est consistant avec $\mathrm{Th}((\mathcal{M}, b)_{b\in B})$ ce qui permettra en utilisant le lemme de Zorn de conclure que p est contenu dans un ensemble de formules consistant avec $\mathrm{Th}((\mathcal{M}, b)_{b\in B})$ et maximal par rapport à cette propriété. Il suffit de vérifier (pourquoi ?) que l'ensemble

$$\mathrm{Th}((\mathcal{M}, b)_{b\in B}) \cup \{\exists x_1 \dots x_k \phi(x_1, \dots, x_k; c_{a_1}, \dots, c_{a_l})\} \ ,$$

où $\phi(x_1, \dots, x_k; c_{a_1}, \dots, c_{a_l}) \in p$, est consistant. Par hypothèse,

$$\mathrm{Th}((\mathcal{M}, a)_{a\in A}) \cup \{\exists x_1 \dots x_k \phi(x_1, \dots, x_k; c_{a_1}, \dots, c_{a_l})\}$$

est consistant. Alors, il existe $(m_1, \dots, m_k) \in M^k$ qui satisfait $\phi(x_1, \dots, x_k; c_{a_1}, \dots, c_{a_l})$ dans $(\mathcal{M}, a)_{a\in A}$. De manière équivalente,

$$\mathcal{M} \models \phi[(m_1, \dots, m_k; a_1, \dots, a_l)] \ .$$

Ceci implique que $(m_1, \dots, m_k)$ satisfait $\phi(x_1, \dots, x_k; c_{a_1}, \dots, c_{a_l})$ dans $(\mathcal{M}, b)_{b\in B}$. Ainsi,

$$\mathrm{Th}((\mathcal{M}, b)_{b\in B}) \cup \{\exists x_1 \dots x_k \phi(x_1, \dots, x_k; c_{a_1}, \dots, c_{a_l})\}$$

est consistant. □

Un type comme q dans le lemme précédent est dit une *extension* (inflation de l'usage de ce mot ?) de p à l'ensemble de paramètres B. Il faut bien souligner qu'un type peut avoir plusieurs extensions à un même ensemble de paramètres. Illustrons ceci dans le cas des ordres denses linéaires sans extrémités en suivant la notation de l'exemple 7.2.1. Soient $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ et $a \in \mathbb{Q}$. Alors

$$\mathrm{tp}_{(\mathbb{R},<)}(a) = \mathrm{tp}_{(\mathbb{R},<)}(\alpha)$$

tandis que

$$\mathrm{tp}_{(\mathbb{R},<)}(a/\mathbb{Q}) \neq \mathrm{tp}_{(\mathbb{R},<)}(\alpha/\mathbb{Q}) \ .$$

Clairement, la formule $x = a$ appartient à $\mathrm{tp}_{(\mathbb{R},<)}(a/\mathbb{Q})$ mais pas à $\mathrm{tp}_{(\mathbb{R},<)}(\alpha/\mathbb{Q})$. D'une façon intuitive, l'extension $\mathrm{tp}_{(\mathbb{R},<)}(\alpha/\mathbb{Q})$ est plus informative à propos du type initial sur l'ensemble vide de paramètres parce qu'elle en garde plus de propriétés tandis que $\mathrm{tp}_{(\mathbb{R},<)}(a/\mathbb{Q})$ "dévie" de $\mathrm{tp}_{(\mathbb{R},<)}(a)$ en introduisant de nouvelles "formes" de formules, et en particulier, une forme faisant intervernir l'égalité.

Nous finissons cette section avec une autre application de la compacité. Il s'agit d'une version générale paramétrisée du lemme 7.1.4 que nous avons déjà evoquée dans l'exemple 7.2.1.

Lemme 7.2.5 *Soient $\mathcal{L}$ un langage du premier ordre et $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M. Si $A \subset M$ et $p \in S_k^{\mathcal{M}}(A)$, alors il existe une extension élémentaire $\mathcal{N}$ de $\mathcal{M}$ qui contient une réalisation de p. En d'autres termes, il existe un k-uplet $(n_1, \dots, n_k)$ extrait de l'univers de $\mathcal{N}$ qui satisfait toutes les formules de p dans $(\mathcal{N}, a)_{a\in A}$. Comme dans le lemme 7.1.4, $\mathcal{N}$ peut être choisi de façon à ce que son univers ait le même cardinal que M.*

Preuve. Exercice. □

7.3 Modèles saturés

Dans cette section, nous introduirons une notion clé pour l'étude des théories complètes, à savoir la notion de *modèle saturé*. Un modèle saturé d'une théorie complète, ou une extension élémentaire saturée d'une structure est une structure "riche" en réalisations des types vérifiant certaines hypothèses. Intuitivement, ceci revient à dire qu'une structure saturée présente concrètement les propriétés de sa théorie élémentaire.

La richesse en réalisation exigée pour l'AOC "saturée" est une condition suffisamment forte pour que les modèles saturés d'une théorie du premier ordre soient rares. Par conséquent, nous pourrons démontrer des théorèmes d'isomorphisme et d'homogénéité des structures saturées. Ce qui est remarquable est que les preuves de ces résultats ont des aspects communs faisant intervenir le va-et-vient que nos lecteurs n'auront pas de difficultés à reconnaître (*Indication : les modèles riches des travaux dirigés, les chaînes denses linéaires sans extrémités...*).

Comme d'habitude nous travaillerons dans un langage fixé $\mathcal{L}$. La nouveauté sera l'usage fréquent des paramètres. A ce paysage général s'ajoutera la présence d'une théorie complète. S'il n'est pas précisé de quelle théorie il s'agit, il suffira de considérer la théorie du premier ordre de la structure en question.

Définition 7.3.1 *Soient κ un cardinal infini, $\mathcal{L}$ un langage et $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M. La structure $\mathcal{M}$ est dite* κ-saturée *si l'une des trois conditions équivalentes est satisfaite :*

1. *pour tout $A \subset M$ de cardinal strictement inférieur à κ, tout 1-type dans $S_1^{\mathcal{M}}(A)$ est réalisé dans $(\mathcal{M}, a)_{a\in A}$;*
2. *pour tout $A \subset M$ de cardinal strictement inférieur à κ, tout 1-type de la théorie $\mathrm{Th}((\mathcal{M}, a)_{a\in A})$ est réalisé dans $(\mathcal{M}, a)_{a\in A}$;*
3. *pour tout $A \subset M$ de cardinal strictement inférieur à κ, tout 1-type de la théorie $\mathrm{Th}((\mathcal{M}, a)_{a\in A})$ est de la forme $\mathrm{tp}_{\mathcal{M}}(\alpha/A)$ pour un certain $\alpha \in M$.*

Une structure est dite saturée *si elle est finie ou elle est de cardinal κ et κ-saturée.*

La question d'existence des extensions élémentaires saturées ou κ-saturées est subtile à tel point qu'elle a des liens avec les axiomes de la théorie des ensembles. En effet, certaines conclusions d'existence nécessite l'hypothèse généralisée du continu. Heureusement, ces aspects trop ensemblistes peuvent être évités et par ailleurs, la plupart du temps les modèles ω-saturés seront suffisants pour arriver à notre objectif principal, à savoir la compréhension d'une théorie élémentaire et de ses modèles.

Commençons en donnant quelques exemples ω-saturés afin de motiver la notion. Nous ne vérifierons pas les énoncés. Fréquemment, un travail considérable sera nécessaire pour comprendre la structure d'un modèle ω-saturé.

Exemples 7.3.1 1. Tout ordre dense, linéaire, sans extrémités dans le langage usuel consistant en un seul symbole de relation binaire est une structure ω-saturée.

2. Toute extension saturée de la structure $\mathcal{N} = (\mathbb{N}, <)$ dans le langage $\mathcal{L} = \{<\}$ d'une relation d'ordre est de la forme $(\mathbb{N} \sqcup D \times \mathbb{Z}, \prec)$ avec D un ordre dense linéaire sans extrémités et $\prec$ qui donne priorité aux éléments de $\mathbb{N}$ et qui y induit l'ordre usuel des naturels tandis qu'il induit l'ordre lexicographique sur $D \times \mathbb{Z}$. En particulier, $\mathcal{N}$ a des extensions élémentaires ω-saturés et dénombrables.

3. Soit $\mathcal{L} = \{., {}^{-1}, 1\}$ le langage des groupes. Tout modèle ω-saturé de $\mathrm{Th}((\mathbb{Z}, +, 0))$ est de cardinal au moins $2^{\aleph_0}$. Pour vérifier cette conlusion il suffit d'étudier les conditions deux à deux disjointes de divisibilité par les parties des nombres premiers.

4. Un corps algébriquement clos de caractéristique p dans le langage des corps est ω-saturé si et seulement s'il est de degré de transcendance infini sur son corps premier.

Nous abordons notre étude par un lemme simple :

Lemme 7.3.2 *Soit $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M. $\mathcal{M}$ est κ-saturée si et seulement si pour tout $A \subset M$ de cardinal strictement inférieur à κ et pour tout $k \in \mathbb{N}^*$, tout type $p \in S_k^{\mathcal{M}}(A)$ est réalisé par un élément de M^k dans $(\mathcal{M}, a)_{a\in A}$.*

Preuve. La suffisance de la condition est claire. Démontrons-en la nécessité par récurrence sur k. La récurrence est amorcée par l'hypothèse de saturation. Ainsi, nous pouvons supposer $k > 1$. Soient A une partie de M et p un k-type dans $S_k^{\mathcal{M}}(A)$. Alors, l'ensemble

$$\overline{p} = \{\exists x_1 \phi(x_1, \ldots, x_k) | \phi(x_1, \ldots, x_k) \in p\}$$

est consistant avec $\mathrm{Th}((\mathcal{M}, a)_{a\in A})$. Comme $\overline{p}$ est un ensemble de formules à paramètres dans A, il existe par récurrence $(m_2, \ldots, m_k) \in M^{k-1}$ qui réalise $\overline{p}$ dans $(\mathcal{M}, a)_{a\in A}$. Il en découle que l'ensemble

$$p_1 = \{\phi(x, c_{m_2}, \ldots, c_{m_k}) | \phi(x_1, \ldots, x_k) \in p\}$$

est un 1-type dans $S_1^{\mathcal{M}}(A \cup \{m_2, \ldots, m_k\})$. Par la κ-saturation (ou par récurrence), p_1 est réalisé dans M par m_1. Le k-uplet $(m_1, \ldots, m_k)$ est une réalisation de p dans $(\mathcal{M}, a)_{a\in A}$. □

Dans ce qui suit, sauf mention contraire, les cardinaux des structures seront infinis. Une structure finie comme l'indique la définition 7.3.1 est κ-saturée pour tout cardinal infini κ. Le premier résultat conséquent concerne l'existence :

Théorème 7.1 *Soit $\mathcal{L}$ un langage. Si κ est un cardinal régulier supérieur ou égal à $|\mathcal{L}|$ et que $\mathcal{M}$ est une structure de cardinal κ alors $\mathcal{M}$ a une extension élémentaire κ-saturée.*

Le lemme suivant nous permettra de construire par récurrence une chaîne d'extensions élementaires.

Lemme 7.3.3 *Soient $\mathcal{L}$ un langage et $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M et de cardinal $\kappa \geq |\mathcal{L}|$. Alors, $\mathcal{M}$ a une extension élémentaire $\mathcal{N}$ telle que pour tout $A \subset M$ de cardinal strictement inférieur à M et $p \in S_1^{\mathcal{M}}(A)$ $(\mathcal{N}, a)_{a\in A}$ réalise p.*

Preuve. La preuve a deux étapes. Dans la première on construit un modèle qui réalise tous les 1-types sur un seul ensemble A, et dans la deuxieme on met ensemble tous ces modèles. Chaque étape, en outre de l'application du théorème de compacité, fait intervenir le corollaire 6.2.5.

Fixons d'abord une partie A de M comme dans l'énoncé et considérons séparément chaque $p \in S_1^{\mathcal{M}}(A)$. D'après le lemme 7.2.5, il existe une extension élémentaire $\mathcal{M}_p$ de $\mathcal{M}$ qui réalise p. Le corollaire 6.2.5 montre alors qu'il existe une extension élémentaire $\mathcal{M}_A$ de $\mathcal{M}$ qui étend élémentairement tous ces $\mathcal{M}_p$. C'est la fin de la première étape. La deuxième étape consiste à mettre ensemble les $\mathcal{M}_A$ où A est comme dans l'énoncé. Ceci se fait en utilisant une deuxième fois le corollaire 6.2.5. □

Preuve du Théorème 7.1. Nous construirons une chaîne croissante d'extensions élémentaires de longueur κ. Le lemme 7.3.3 servira pour les passages aux successeurs tandis que les limites seront l'union de ce qui a été déjà construit. La récurrence s'amorce sans peine : nous posons $\mathcal{M}_0 = \mathcal{M}$. Si pour $i < \kappa$, $\mathcal{M}_i$ est construit, alors $\mathcal{M}_{i+1}$ est obtenu en appliquant le lemme 7.3.3 à $\mathcal{M}_i$. Si $j \neq 0$ est un ordinal limite, alors $\mathcal{M}_j$ a pour base l'union de celles des $\mathcal{M}_i$ pour $i < j$. Ainsi s'obtient

$$\mathcal{M}_0 \preceq \mathcal{M}_1 \preceq \ldots \preceq \mathcal{M}_i \preceq \ldots \quad (i < \kappa)$$

La structure dont nous sommes à la recherche aura pour ensemble de base l'union des univers des $\mathcal{M}_i$. Appelons cette structure limite $\mathcal{M}_\kappa$. Par construction, $\mathcal{M} \preceq \mathcal{M}_\omega$ (pourquoi ?).

Il reste à vérifier la κ-saturation. Or κ est un cardinal régulier et c'est aussi la lonngueur de la chaîne que nous avons construite. En conséquence, si on fixe un ensemble de paramètres $A \subset M_\kappa$ de cardinal strictement inférieur à κ, alors A est nécessairement contenu dans l'une des M_i. La construction du début de la preuve montre alors que tout type sur A sera réalisé dans $\mathcal{M}_{i+1}$. □

Avant d'aborder les questions d'isomorphisme, nous énoncerons un théorème dont la preuve est plus subtile et qui fait le lien avec l'hypothèse généralisée du continu. En effet, si l'hypothèse généralisée du continu est acceptée alors il découle de ce théorème que chaque structure a une extension élémentaire saturée.

Théorème 7.2 (Theórème 9.15 du CTM de Poizat, p. 252) *Soient $\mathcal{L}$ un langage, κ un cardinal supérieur où égal à $|\mathcal{L}$ et T une théorie complète dans $\mathcal{L}$. Alors, tout modèle de T de cardinal inférieur ou égal à 2^κ se plonge élémentairement dans un modèle κ^+-saturé de cardinal 2^κ.*

Ce résultat n'a été donné qu'à titre de culture générale pour ceux qui s'intéresseraient aux liens entre la théorie des ensembles et la théorie des modèles.

Après ce court détour, nous continuons vers nos objectifs. En général, deux structures κ-saturées et élémentairement équivalentes ne sont pas nécessairement de même cardinal ce qui exclut toute possibilité d'un théorème d'isomorphisme. Néanmoins, comme deux structures élémentairement équivalentes satisfaisant la même condition de saturation ont en commun les propriétés de leur théorie du premier ordre, on s'attend à de fortes similarités entre les deux. Avant de préciser ce phénomène qui a été déjà étudié aux travaux dirigés, nous donnons un théorème d'isomorphisme. Notons que tout cardinal dans l'énoncé est dénombrable. Ceci n'est pas nécessaire mais nécessiterait d'introduire des I-types faisant intervenir des I-uplets éventuellement de longueur infinie. Nous préférons d'éviter cela.

Théorème 7.3 *Soient $\mathcal{L}$ un langage dénombrable, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures dénombrables élémentairement équivalentes, et ω-saturées, donc saturées. Alors $\mathcal{M} \cong \mathcal{N}$.*

Preuve. Il faut construire un isomorphisme en suivant la recette de la proposition 5.7.2 en appliquant le va-et-vient. Les détails, similaires à celles des études déjà effectuées pendant les travaux dirigés, sont laissés aux soins de nos lecteurs. □

La version générale du théorème précédent est la suivante. C'est un bon exercice, quoique non prioritaire, d'essayer de le faire pour voir où il faut introduire une extension de la notion de type telle que nous l'avons définie.

Théorème 7.4 *Soient $\mathcal{L}$ un langage, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures élémentairement équivalentes, de même cardinal $\kappa \geq |\mathcal{L}|$ et saturées. Alors $\mathcal{M} \cong \mathcal{N}$.*

La structure $(\mathbb{Z}, +, 0)$ (l'exemple 7.3.1 (3)), et le fait que l'équivalence élémentaire soit une condition plus faible que l'isomorphisme (c'est en fait une force de la théorie des modèles) font penser que le théorème 7.3 (ou 7.4) n'est pas celui qu'il faut rechercher dans l'étude des théories du premier ordre. En fait, dans les travaux dirigés cette étude était abordée en utilisant des applications qui ne sont que "localement" des isomorphismes. La particularité du théorème 7.3 (ou de 7.4) est la force de son hypothèse. Néanmoins, quand celle-ci est affaiblie, le va-et-vient est toujours très efficace. Nous mettrons cette aproche dans un cadre général et en tirerons quelques conséquences.

Précisons d'abord ce que nous entendons par "isomorphisme local".

Définition 7.3.4 *Soient $\mathcal{L}$ un langage du premier ordre, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures d'univers M et N respectivement. Les deux structures $\mathcal{M}$ et $\mathcal{N}$ sont dites* ∞-équivalentes *si pour tout $(a_1, \ldots, a_k) \in M^k$, $(b_1, \ldots, b_k) \in N^k$ qui satisfont les mêmes formules à au plus k variables libres dans $\mathcal{M}$ et $\mathcal{N}$ respectivement, les deux conditions suivantes sont vérifiées :*

1. *pour tout $a \in M$, il existe $b \in N$ tel que $(a_1, \ldots, a_k, a)$ et $(b_1, \ldots, b_k, b)$ satisfassent les mêmes formules ;*
2. *pour tout $b' \in N$, il existe $a' \in M$ tel que $(b_1, \ldots, b_k, b')$ et $(a_1, \ldots, a_k, a')$ satisfassent les mêmes formules.*

Notons que cette définition implique l'équivalence élémentaire des deux structures et qu'elle fournit le bon cadre pour les "similarités" de deux $\mathcal{L}$-structures En effet, le même va-et-vient que celui utilisé dans la preuve du théorème 7.3 montre que deux structures élémentairement équivalentes et ω-saturées sont ∞-équivalentes.

Lemme 7.3.5 *Soient $\mathcal{L}$ un langage, $\mathcal{M}$ et $\mathcal{N}$ deux $\mathcal{L}$-structures ∞-équivalentes.*

1. *Si $\mathcal{M}$ est ω-saturée, alors il en est de même pour $\mathcal{N}$.*
2. *Si les deux structures sont dénombrables, alors elles sont isomorphes.*

Preuve. Pour chacun des deux points, il suffit de suivre la démarche du théorème 7.3. □

Proposition 7.3.6 *Deux chaînes denses sans extrémités sont ∞-équivalentes.*

Preuve. Voir la preuve du théorème 6.4. □

Corollaire 7.3.7 *Une chaîne dense sans extrémités est une structure ω-saturée.*

Preuve. Le théorème 7.1 et le fait que la propriété d'être une chaîne dense sans extrémités soit exprimable par des énoncés du premier ordre montrent l'existence des chaînes denses sans extrémités ω-saturées. Alors on conclut en utilisant le théorème 7.1 et la proposition 7.3.6. □

Suivant les idées et raisonnements du lemme 7.3.5, la propriété suivante d'*homogénéité* peut être démontrée :

Proposition 7.3.8 *Soient $\mathcal{L}$ un langage, et $\mathcal{M}$ une $\mathcal{L}$-structure dénombrable et ω-saturée. Si $(a_1, \ldots, a_k)$ et $(b_1, \ldots, b_k)$ sont deux k-uplets extraits de l'univers de $\mathcal{M}$ tels que $\mathrm{tp}_{\mathcal{M}}(a_1, \ldots, a_k) = \mathrm{tp}_{\mathcal{M}}(b_1, \ldots, b_k)$, alors il existe un automorphisme ν de $\mathcal{M}$ tel que $\nu(a_i) = b_i$ pour $i \in \{1, \ldots, k\}$.*

Nous finissons cette section par un résultat d'existence plus concret, dans le contexte des modèles ω-saturés. Bien que le cadre de travail puisse paraître limité, le résultat s'avère utile dans beaucoup de contextes.

Théorème 7.5 *Soit T une théorie complète dans un langage $\mathcal{L}$ dénombrable. La théorie T a un modèle ω-saturé et dénombrable si et seulement si $S(T)$ est dénombrable.*

Lemme 7.3.9 *Soient T une théorie complète dans un langage $\mathcal{L}$ dénombrable, et $\mathcal{M}$ un modèle de T. Si $S(T)$ est dénombrable, alors pour tout sous-ensemble fini X de M, $S^{\mathcal{M}}(X)$ est dénombrable aussi.*

Preuve. Posons $X = \{a_1, \ldots, a_n\}$. Soient $p \in S_k(X)$ et $\bar{b} = (b_1, \ldots, b_k)$ une réalisation de p dans une extension élémentaire $\tilde{\mathcal{M}}$ de $\mathcal{M}$. Alors, $\mathrm{tp}_{\tilde{\mathcal{M}}}(\bar{b}, a_1, \ldots, a_n) \in S_{k+n}(T)$. L'application de $S_k(X)$ vers $S_{k+n}(T)$ qui associe $\mathrm{tp}_{\tilde{\mathcal{M}}}(\bar{b}/X)$ au type $\mathrm{tp}_{\tilde{\mathcal{M}}}(\bar{b}, a_1, \ldots a_n)$ est injective (pourquoi ?). La conclusion s'ensuit. □

Lemme 7.3.10 *Soient T une théorie complète dans un langage $\mathcal{L}$ dénombrable, et $\mathcal{M}$ un modèle dénombrable de T. Si $S(T)$ est dénombrable, alors il existe une extension élémentaire $\mathcal{N}$ de $\mathcal{M}$ telle que $|N| = |M|$ et que pour tout sous-ensemble fini $X \subseteq M$ et tout $p \in S_1^{\mathcal{M}}(X)$, p soit réalisé dans N.*

Preuve. Le lemme se démontre par un raisonnement de chaîne similaire à celui du théorème 7.1. Comme M est un ensemble dénombrable, le lemme 7.3.9 montre que l'ensemble $\bigcup_{X \subseteq M, X \text{ fini}} S^{\mathcal{M}}(X)$ est un ensemble dénombrable, soit $\{p_i : i \in \mathbb{N}\}$. On pose $\mathcal{M}_0 = \mathcal{M}$. Si $\mathcal{M}_i$ est déjà défini, alors pour obtenir $\mathcal{M}_{i+1}$ on applique le lemme 7.2.5 à $\mathcal{M}_i$ et p_i. La structure recherchée $\mathcal{N}$ aura pour univers l'union des univers des $\mathcal{M}_i$. Finalement, l'arithmétique des cardinaux montre que $|N| = |M|$. □

Preuve du Théorème 7.5. Dans un modèle dénombrable on ne peut réaliser qu'un nombre dénombrable de types, ainsi la condition est suffisante. Il reste alors à étudier sa nécessité.

D'après le théorème de Löwenheim-Skolem, T a un modèle dénombrable $\mathcal{M}_0$. La preuve consiste à définir par récurrence une chaîne croissante d'extensions élémentaires de longeur dénombrable. Nous venons de définir le premier élément de cette chaîne. Supposons maintenant que $\mathcal{M}_i$ soit définie. Nous appliquons le lemme 7.3.10 à $\mathcal{M}_i$ pour obtenir $\mathcal{M}_{i+1}$. Le modèle dénombrable recherché a pour base l'union des univers des $\mathcal{M}_i$. Tous les $\mathcal{M}_i$ ayant des univers dénombrables, il en est de même pour $\mathcal{N}$. □

7.4 Définissabilité, Théorème de Svenonius

Nous avons déjà abordé à maintes occasions la notion d'*ensemble définissable*. Il est temps d'étudier cette notion aussi naturelle que fondamentale dans tous ses aspects. Comme d'habitude nous fixons un langage $\mathcal{L}$, et une théorie, complète sauf mention contraire, dans ce langage nous accompagnera. Quand nous ne parlerons que d'une structure $\mathcal{M}$ la théorie sera $\mathrm{Th}(\mathcal{M})$. Notons que dans cette section les expansions innterviendront assez souvent.

Commençons avec une $\mathcal{L}$-structure $\mathcal{M}$ dont la base sera notée M. Un *ensemble définissable dans la structure* $\mathcal{M}$ est une partie d'une puissance cartesienne M^k ($k \in \mathbb{N}^*$) de M dont les éléments sont décrits par une formule du premier ordre. Plus précisément, une partie $E \subset M^k$ est définissable s'il existe une $\mathcal{L}$-formule du premier ordre $\phi(x_1, \ldots, x_k)$ telle que

$$E = \{(m_1, \ldots, m_k) \in M^k \mid \mathcal{M} \models \phi[(m_1, \ldots, m_k)]\} .$$

Les exemples sont abondants puisque toute formule du premier ordre permet de définir un ensemble définissable. Si elle n'est pas consistante, disons $x \neq x$, elle définit l'ensemble vide ; si elle est consistante, alors il suffit de déterminer les k-uplets extraits de M qui satisfont la formule en question. L'autre exemple extrême aux antipodes de l'ensemble vide est M^k qui est défini par la formule $\bigwedge_{i=1}^{k} x_i = x_i$. Une question plus pertinente et parfois difficile est si une partie arbitrairement choisie de M^k est définissable. Nous avons déjà vu que les automorphismes de la structure $\mathcal{M}$ fournissent un outil efficace si on espère arriver à des réponses négatives, et ceci grâce à la proposition 5.6.2 (4).

A la détermination de la définissabilité d'une partie arbitrairement choisie de M^k est lié un deuxième problème aussi important qui est celui de la détermination des *propriétés définissables*. Toute structure a des propriétés provenant de sa nature particulière, et souvent il n'est pas clair s'il existe une façon du premier ordre pour exprimer chacune de ces propriétés. Plus précisément, pour une propriété P, il n'est pas toujours clair s'il existe un énoncé σ du premier ordre telle que la structure $\mathcal{M}$ vérifie P si et seulement si $\mathcal{M} \models \sigma$. La proposition 6.1.1 était un exemple d'une propriété algébrique de ce genre. Un autre exemple est la théorie de la relation équivalence avec une classe finie et une seule pour chaque nombre naturel. Il n'existe pas d'énoncé σ qui est vrai dans un modèle (donc dans tous les modèles, la théorie étant complète) de cette théorie si et seulement si tout élément appartient à une classe finie. Par compacité, la théorie a des modèles avec des classes infinies. Par ailleurs, on peut montrer, toujours en utilisant la compacité qu'il n'existe pas de formule $\phi(x)$ à une variable libre qui soit satisfaite exactement par les éléments appartenant à une classe finie sauf si le modèle est celui qui ne contient que des classes finies. Dans ce cas particulier, la formule $x = x$ définit aussi l'ensemble des éléments appartenant à une classe finie. Or ce n'est qu'une coïncidence. Dans une extension élémentaire ce ne sera plus le cas, et encore une fois le théorème de compacité permet de vérifier qu'on ne peut trouver une formule du premier ordre qui survive aux passages aux extensions élémentaires (voyez-vous comment ?). Avant d'aborder certaines de ces questions plus en détail, complétons d'abord l'étude de la notion d'ensemble définissable.

Nous avons déjà observé que les expansions en utilisant les paramètres ont des effets considérables sur la nature des ensembles définissables (l'exemple 7.2.1). Ces effets motivent la notion d'un *ensemble définissable avec paramètres*. Si $A \subset M$ est fixé comme ensemble de paramètres, alors une partie E de M^k est dit *définissable dans $\mathcal{M}$ avec paramètres dans A* ou *A-définissable dans $\mathcal{M}$* s'il existe une formule $\phi(x_1, \ldots, x_k; c_{a_1}, \ldots, c_{a_l})$ avec $\{a_1, \ldots, a_k\} \subset A$ telle que

$$E = \{(m_1, \ldots, m_k) \in M^k \mid (\mathcal{M}, a)_{a \in A} \models \phi(x_1, \ldots, x_k; c_{a_1}, \ldots, c_{a_l})[(m_1, \ldots, m_k)]\}$$

soit encore,

$$E = \{(m_1, \ldots, m_k) \in M^k \mid \mathcal{M} \models \phi[(m_1, \ldots, m_k; a_1, \ldots, a_l)]\} .$$

La notation $\phi(\mathcal{M}; c_{a_1}, \ldots, c_{a_k})$ ou $\phi(\mathcal{M}; a_1, \ldots, a_k)$ est souvent utilisée.

La définissabilité avec paramètres est exactement la définissabilité dans l'expansion $(\mathcal{M}, a)_{a \in A}$. Alors, les automorphismes de l'expansion $(\mathcal{M}, a)_{a \in A}$, de manière équivalente, les automorphismes de $\mathcal{M}$ qui fixent l'ensemble A point par point, soit les $\mathcal{A}$-automorphismes de $\mathcal{M}$, stabilisent tout

ensemble $\mathcal{A}$-définissable (la proposition 5.6.2 (4)). Plus généralement, on peut considérer l'action des automorphismes de $\mathcal{M}$ sur les ensembles définissables dans $\mathcal{M}$ avec paramètres. Si ν est un tel automorphisme et que E est comme ci-dessus, alors

$$\begin{aligned}\nu(E) &= \{(\nu(m_1),\ldots,\nu(m_k)) \mid (m_1,\ldots,m_k)\in E\}\\ &= \{(m_1,\ldots,m_k)\in M^k \mid \mathcal{M}\models\phi[(m_1,\ldots,m_k,\nu(a_1),\ldots,\nu(a_l))]\}\end{aligned}$$

Quitte à ajouter au langage ($\mathcal{A}$), les symboles de constantes $c_{\nu(a_1)},\ldots,c_{\nu(a_l)}$ une autre façon de présenter $\nu(E)$ est la suivante :

$$\nu(E)=\{(m_1,\ldots,m_k)\in M^k \mid (\mathcal{M},a)_{a\in A}\models\phi(x_1,\ldots,x_k;c_{\nu(a_1)},\ldots,c_{\nu(a_l)})[(m_1,\ldots,m_k)]\}$$

Les automorphismes fournissent donc un outil efficace pour étudier les définissables. L'autre outil important est la compacité. En effet, nous invitons nos lecteurs à réflechir sur la remarque suivante : dans un modèle de la théorie de la relation d'équivalence à une classe finie et une seule pour chaque nombre naturel, qui a au moins une classe infinie, les éléments appartenant à une classe finie forme un ensemble qui n'est pas définissable avec des paramètres non plus.

Le théorème suivant fait le lien avec les types, autre ingrédient important de l'étude des aspects définissables d'une structure et de sa téorie. Nous garderons la notation introduite au début de cette section.

Théorème 7.6 *Soient $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M, a et b deux éléments de M. Alors $\mathrm{tp}_{\mathcal{M}}(a)=\mathrm{tp}_{\mathcal{M}}(b)$ si et seulement s'il existe une extension élémentaire $\mathcal{N}$ de $\mathcal{M}$ et un automorphisme ν de $\mathcal{N}$ tels que $\nu(a)=b$.*

Preuve. Le type $\mathrm{tp}_{\mathcal{M}}(a)$ est sans paramètres si bien que $\mathrm{tp}_{\mathcal{N}}(a)=\mathrm{tp}_{\mathcal{M}}(a)$ et que tout automorphisme de $\mathcal{N}$ (en fait de tout modèle de $\mathrm{Th}(\mathcal{M})$ contenant a) stabilise chaque ensemble dans ce type, et par conséquent fixe le type. Ainsi, l'existence d'un automorphisme tel que celui de l'énoncé engendre l'égalité des deux types.

La direction inverse nécessite un plus grand travail. On commence par poser $\mathcal{L}^+=\mathcal{L}(M)\cup\{\nu\}$ où ν est un symbole de fonction unaire qu'on aurait pu remplacer par un symbole de relation binaire suivant les goûts et les situations. Pour alléger la notation nous utiliserons a et b pour les symboles de constantes qui nomment a et b ainsi que pour les éléments eux-mêmes.

Nous posons ensuite

$$\begin{aligned}T^+ &= \mathrm{Th}((\mathcal{M},m)_{m\in M})\\ &\cup \{\nu(a)=b\}\cup\{\text{“}\nu\text{ est une bijection”}\}\\ &\cup \{\nu(c)=c \mid c \text{ est un symbole de constante de } \mathcal{L}\}\\ &\cup \{\nu(f(x_1,\ldots,x_k))=f(\nu(x_1),\ldots,\nu(x_k)) \mid f \text{ est un symbole de fonction de } \mathcal{L}\}\\ &\cup \{\forall x_1\ldots x_k(R(x_1,\ldots,x_k)\leftrightarrow R(\nu(x_1),\ldots,\nu(x_k))) \mid R \text{ est un symbole de relation de } \mathcal{L}\}\ .\end{aligned}$$

L'objectif de tout le reste est évident : vérifier que T^+ est un ensemble consistant d'énoncés.

Remarquons que tel qu'il est défini, il est possible que $|T^+|>\aleph_0$. Si le langage initial $\mathcal{L}$ n'est pas dénombrable, alors il n'est pas possible d'appliquer le théorème de Löwenheim-Skolem pour remplacer $\mathcal{M}$ par une sous-structure élémentaire dénombrable contenant a et b. Par contre, nous pourrons surmonter cet obstacle parce que, par compacité, le problème se réduit à l'étude des parties finies de T.

Soit T_0 une partie finie de T^+. Alors, T_0 fait intervenir un nombre fini de symboles de $\mathcal{L}$ nommant les éléments de la signature de $\mathcal{M}$. Ce langage réduit sera noté $\mathcal{L}_0$. Nous pouvons supposer que T_0 exprime aussi la bijectivité de ν, que a et b en tant que symboles de constantes interviennent au moins dans l'énoncé $\nu(a)=b$. Soient $m_1,\ldots,m_k$ les éléments de M nommés pas les symboles de constantes de $\mathcal{L}_0$. Nous utiliserons la même notation pour les symboles de constantes qui les nomment.

L'étape suivante est de trouver un cadre où on peut travailler dans une structure dénombrable qui permettra finalement d'utiliser la proposition 7.3.5 (2). Réduisons $\mathcal{M}$ à une $\mathcal{L}_0$-structure $\mathcal{M}_0$.

Alors, d'après le théorème de Löwenheim-Skolem $\mathcal{M}_0$ possède une $\mathcal{L}_0$-sous-structure élémentaire $\mathcal{N}_0$ qui est dénombrable et qui contient $\{m_1, \ldots, m_k, a, b\}$. En utilisant la structure $\mathcal{N}_0$ et le théorème de compacité, nous montrerons l'existence d'une $\mathcal{L}_0$-structure $\mathcal{N}_1$ dans laquelle ν sera interprété par un automorphisme. Une telle conclusion vérifiera que T_0 est consistante, et le reste découlera de la compacité cette fois-ci appliquée à T.

On introduit alors des symboles de relations qui seront les témoins d'un va-et-vient partiel. Plus précisément, pour chaque $i \in \mathbb{N}^*$ on ajoute $E_i(x_1, \ldots, x_i; y_1, \ldots, y_i)$ au langage $\mathcal{L}_0$. Ensuite, les énoncés de T_0 sont étendus à un ensemble d'énoncés qui expriment que chaque E_i est une relation d'équivalence et que les propriétés suivantes sont vérifiées :

$$\forall x \exists y E_1(x, y)$$

$$\forall x_1 y_1 x_2 \exists y_2 (E_1(x_1, y_1) \to E_2(x_1, x_2; y_1, y_2))$$

$$\vdots$$

$$\forall x_1 y_1 \ldots x_n y_n x_{n+1} \exists y_{n+1} (E_n(x_1, \ldots, x_n; y_1, \ldots, y_n) \to E_{n+1}(x_1, \ldots, x_n, x_{n+1}; y_1, \ldots, y_n, y_{n+1}))$$

$$\vdots$$

Pour toute formule atomique $\phi(x_1, \ldots, x_n, y)$ dans $\mathcal{L}_0$,

$$\forall x_1 y_1 \ldots x_n y_n (E_n(x_1, \ldots, x_n; y_1, \ldots, y_n) \to (\phi(x_1, \ldots, x_n, a) \leftrightarrow \phi(y_1, \ldots, y_n, b))) \quad .$$

En utilisant la compacité nous vérifierons que cette famille d'énoncés est consistante. Une partie finie de ces énoncés ne fait intervenir que les relations d'équivalence $E_1, \ldots, E_k$ pour un certain $k \in \mathbb{N}$. Pour ce faire, il suffit de faire un va-et-vient entre les uplets satisfaisant ces relations. D'ailleurs, on ne peut pas espérer de faire un va-et-vient infini puisque $\mathcal{N}_0$ n'est pas nécessairement un modèle ω-saturé, dans lequel il ne faudrait d'ailleurs rien faire grâce à la proposition 7.3.8 puisque $\mathrm{tp}(a) = \mathrm{tp}(b)$.

Nous utiliserons la base N_0 de $\mathcal{N}_0$ pour l'interprétation des symboles : pour $i \in \{1, \ldots, k\}$, deux i-uplets $(x_1, \ldots, x_i)$, $(y_1, \ldots, y_i)$ sont en relation E_i si et seulement s'ils satisfont les mêmes formules de rang de quantificateur au plus $k - i$. Le rang de quantificateur est la notion de complexité définie comme suit : les formules atomiques sont de rang 0 ; une formule $\neg\psi$ est de même rang de que ψ ; le rang de $\phi \wedge \psi$ est le maximum des rangs de ϕ et de ψ ; le rang de $\exists x \phi$ s'obtient en augmentant celui de ϕ par 1. Avec cette définition, nous pouvons procéder par récurrence sur k.

Si $k = 1$, alors pour tout $\alpha \in N_0$, il faut trouver $\beta \in N_0$ tel que (a, α) et (b, β) satisfassent les mêmes formules sans quantificateur. Or, le langage $\mathcal{L}_0$ ne contient qu'un nombre fini de symboles et nous considérons les formules sans quantificateur à deux variables libres (x_1, x_2). Il en existe en nombre fini, en conséquence de quoi, considérant la combinaison convenable il s'agit d'une seule formule. Comme a et b ont même type, il existe $\beta \in N_0$ satisfaisant la condition. Ce raisonnement amorce la récurrence. L'étape inductive pour passer de E_i à E_{i+1} se fait en suivant la même idée et en se rappelant que dans l'étape précédente il y avait un nombre fini de formules, donc une seule, et que a et b ont même type.

Le paragraphe précédent et la compacité montrent que la famille d'énoncés à propos des E_i est consistante. Alors, ceux-ci ont un modèle $\mathcal{N}_1$ dénombrable. Dans ce modèle, les E_i sont les témoins d'une ∞-equivalence qui transforme a en b. La structure $\mathcal{N}_1$ étant dénombrable, la proposition 7.3.5 (2) montre que cette ∞-équivalence est un automorphisme. L'interprétation de ν dans $\mathcal{N}_1$ sera cet automorphisme. Ainsi, $\mathcal{N}_1$ vérifie que T_0 est une famille consistante d'énoncés, et de cette conclusion découle, par compacité, la consistance de T^+. □

Corollaire 7.4.1 *Soient $\mathcal{M}$ une $\mathcal{L}$-structure d'univers M, a et b deux éléments de M, et $A \subset M$. Alors $\mathrm{tp}_{\mathcal{M}}(a/A) = \mathrm{tp}_{\mathcal{M}}(b/A)$ si et seulement s'il existe une extension élémentaire $\mathcal{N}$ de $\mathcal{M}$ et un A-automorphisme ν de $\mathcal{N}$ tels que $\nu(a) = b$.*

Remarque : Tous les raisonnements ci-dessus restent valables si on remplace les éléments de M par des k-uplets.

Le théorème 7.4.1 concerne les types. D'une certaine manière, il explique ce que c'est que d'être un type. Le théorème de Svenonius est dans le même esprit mais concerne plutôt les ensembles définissables. En effet, toute partie P de M^k où M est la base d'une $\mathcal{L}$-structure $\mathcal{M}$ peut devenir une partie définissable quitte à considérer l'expansion $(\mathcal{M}, P)$ obtenue en ajoutant un symbole de relation k-aire au langage $\mathcal{L}$ pour nommer l'ensemble P. La question est si cette nouvelle structure est vraiment nouvelle où s'il existe une $\mathcal{L}$-formule $\phi(x_1, \ldots, x_k)$ telle que

$$(\mathcal{M}, P) \models \forall x_1 \ldots x_k (P(x_1, \ldots, x_k) \leftrightarrow \phi(x_1, \ldots, x_k)) \ .$$

Le théorème de Svenonius décrit ce contexte et montre ses liens aux automorphismes de la structure. Notons que ce contexte a été l'un des viviers des problèmes les plus riches et difficiles de la théorie des modèles.

Théorème 7.7 (Svenonius) *Soient $\mathcal{L}$ un langage, $\mathcal{M}$ une $\mathcal{L}$-structure de base M, et P une partie de M^k pour un certain $k \in \mathbb{N}^*$. Si l'appartenance à l'ensemble P n'est pas définissable dans la structure $\mathcal{M}$, en d'autres termes, s'il n'existe pas de $\mathcal{L}$-formule $\phi(x_1, \ldots, x_k)$ telle que $(\mathcal{M}, P) \models \forall x_1 \ldots x_k (P(x_1, \ldots, x_k) \leftrightarrow \phi(x_1, \ldots, x_k))$ alors il existe une extension élémentaire $(\mathcal{M}', P')$ avec un automorphisme s qui ne préserve pas P'.*

Preuve. Dans la preuve, la partie P est traitée comme une relation définissable dans une expansion de $\mathcal{M}$. L'hypothèse dit que cette addition à la structure fournit une nouvelle structure. Cet ajout propre trouvera parmi ses témoins les automorphismes d'une extension élémentaire.

La preuve contient deux étapes dont les détails sont laissés à nos lecteurs. Dans la première étape, il est montré qu'il existe une extension élémentaire $(\mathcal{M}_1, P_1)$ et deux k-uplets $(a_1, \ldots, a_k)$ et $(b_1, \ldots, b_k)$ qui ont même type par rapport à la structure $\mathcal{M}$ tels que

$$(\mathcal{M}_1, P_1) \models P_1[(a_1, \ldots, a_k)] \wedge \neg P_1[(b_1, \ldots, b_k)] \ .$$

Le raisonnement commence par montrer que sous l'hypothèse contradictoire, l'appartenance de P à un type de $S_n(T, P)$ est complètement déterminée par la restriction de ce type à ses $\mathcal{L}$-formules. Dans une telle situation, la compacité montre qu'il existe une $\mathcal{L}$-formule qui définit P.

La deuxième étape suit la preuve du théorème 7.6 pour montrer l'existence d'une extension élémentaire $(\mathcal{M}', P')$ et un automorphisme ν de $\mathcal{M}'$ tels que $\nu(a_i) = b_i$ pour chaque $i \in \{1, \ldots, k\}$. En raison de la propriété des deux k-uples vérifiés dans la première étape, ν n'est pas un automorphisme de l'expansion $(\mathcal{M}', P')$. Ainsi il ne préserve pas P'. □

Chapitre 8

Corps algébriquement clos

Une théorie T est dite *κ-catégorique* ou *catégorique en κ* si elle a à isomorphisme près un seul modèle de cardinal κ. Deux types de catégoricité ont marqué l'histoire de la théorie des modèles : l'ω-catégoricité et la catégoricité non dénombrable. Le théorème de Ryll-Nardzewski qui caractérise les théories dénombrables et ω-catégoriques, et le théorème de Morley qui montre qu'une théorie dénombrable est $\aleph_1$-catégorique si et seulement si elle est catégorique en tous les cardinaux non dénombrables sont à l'origine de la théorie des modèles contemporaine.

La théorie du premier ordre d'un corps algébriquement clos de caractéristique fixée que nous étudierons dans ce chapitre est un exemple de théorie $\aleph_1$-catégorique. La propriété de catégoricité qui peut être vérifiée en utilisant le cardinal des bases de transcendance sur le corps premier, peut aussi être vérifiée à partir d'une propriété plus forte que nous montrerons dans ce chapitre : la théorie des corps algébriquement clos d'une caractéristique donnée est *fortement minimale*. Un aspect très important de ces notions de la théorie des modèles est leur nature géométrique. La minimalité forte décrit le comportement des parties fermées de la droite affine dans la *topologie de Zariski*. Le progrès dans la compréhension de cette nature géométrique est parmi les avancées les plus profondes de la théorie des modèles.

Notons que la théorie des corps algébriquement clos d'une caractéristique donnée n'est pas ω-catégorique. En fait, ces deux notions de catégoricité sont bien loin l'une de l'autre comme nous l'avons déjà observé dans le contexte des chaînes denses sans extrémités.

Dans ce chapitre, à des moments, nous serons moins formels en ce qui concerne l'usage des symboles. En parlant des corps, nous ne ferons plus de distinction entre la lettre représentant la structure et celle repésentant la base de cette structure. Nous suivrons néanmoins notre choix usuel de notation dans les discussions générales. Par ailleurs, nous utiliserons la même lettre pour un symbole de constante et son interprétation dans une structure.

8.1 Axiomes

Dans cette section, nous établirons notre cadre de d'étude des corps algébriquement clos. Nous nous fixerons un langage, et ensuite une théorie dont la vérification de la complétude sera une étape importante du travail que nous sommes en train d'entreprendre.

Fixons d'abord notre langage : $\mathcal{L}_C = \{0, 1, +, ., -, \ ^{-1}\}$. Notre point de départ est la liste infinie d'axiomes qui décrivent les propriétés bien connues des corps algébriquement clos de caractéristique fixée.

1. Axiomes décrivant un corps
2. $1 \neq 0$
3. (A_n) $(n \in \mathbb{N}^*)$ $\forall y_0 \dots y_{n-1} \exists x (y_0 + y_1.x + \dots + y_{n-1}.x^{n-1} + x^n = 0)$
4$_p$ $\underbrace{1 + \dots + 1}_{p \text{ fois}} = 0$ si la caractéristique est un nombre p premier

4_0 (C_n) $\underbrace{1+\ldots+1}_{n \text{ fois}} \neq 0$ pour tout $n \in \mathbb{N}^*$ si la caractéristique est nulle

Cette liste d'énoncés est consistante puisque tout corps algébriquement clos de caractéristique p, avec p éventuellement 0, en est modèle. Dans tout ce chapitre, pour p premier ou 0 fixé, nous noterons CAC_p la théorie formée par la liste ci-dessus et ses conséquences. Le résultat principal de cette section montrera que quelle que soit la valeur de p, CAC_p n'est pas finiment axiomatisable. Notre raisonnement fera usage des connaissances élémentaires en théorie de Galois et du théorème de la compacité.

Avant d'aborder notre étude en détail, il convient de souligner un point pratique : ces axiomes montrent que le fait d'être un corps algébriquement clos d'une certaine caractéristique se préserve quand on passe à des structures élémentairement équivalentes.

Rappelons aussi un lemme bien connu :

Lemme 8.1.1 *Tout corps algébriquement clos est infini.*

Afin d'établir un cadre cohérent de travail pour cette section nous commençons par un lemme général qui montre que la finitude d'une axiomatisation ne dépend pas du système d'axiomes choisis.

Lemme 8.1.2 *Soit T une théorie du premier ordre dans un langage $\mathcal{L}$. On suppose que T soit finiment axiomatisable. Alors de toute axiomatisation de T on peut extraire une axiomatisation finie.*

Preuve. En prenant la conjonction des énoncés dans un ensemble fini d'axiomes de T, nous pouvons supposer que d'un côté $\{\phi\}$ et de l'autre $\Psi = \{\psi_i : i \in I\}$ sont deux ensembles d'axiomes dont le deuxième est de cardinal arbitraire. La définition d'une axiomatisation montre que $\Psi \vdash \phi$. Par compacité, Ψ a une partie finie Ψ_0 tel que $\Psi_0 \vdash \phi$. □

Avant d'aborder la preuve du résultat principal de cette section, nous citons deux faits bien connus sur les corps finis :

Fait 8.1.3

1. *Le nombre d'éléments dans un corps fini de caractéristique p est une puissance de p. Pour tout $d \in \mathbb{N}^*$, il existe, à isomorphisme près, une extension de $\mathbb{F}_p$ de degré d et une seule. C'est l'ensemble des racines du polynôme $x^{p^d} - x$ et il contient p^d éléments.*
2. *Soient $d \in \mathbb{N}^*$ et E/F une extension de degré d'un corps fini F. Il existe une extension et une seule K/F, avec K sous-corps de E, de degré n si et seulement si $n|d$.*

Théorème 8.1 *Quelle que soit la caractéristique p, la théorie CAC_p n'est pas finiment axiomatisable.*

Preuve. Nous montrons d'abord que si p est un nombre premier et que $\{A_{i_1}, \ldots, A_{i_m}\}$ est une partie finie des axiomes A_n ci-dessus alors il existe un corps de caractéristique p qui n'est pas algébriquement clos mais qui satisfait les axiomes $\{A_{i_1}, \ldots, A_{i_m}\}$. Pour ce faire, posons $N = \max(i_1, \ldots, i_m)$.

En utilisant le fait 8.1.3, nous construirons un sous-corps propre de la clôture algébrique de $\mathbb{F}_p$ qui satisfera les axiomes A_n $(n \leq N)$. Nous considérons $K = \bigcup_{e=1}^{\infty} K_e$ où $K_1 = \mathbb{F}_p$ et $[K_{e+1} : K_e] = N!$. Nous montrons d'abord que K satisfait les A_n $(n \leq N)$. Soient $n \leq N$ et $P(X) = a_0 + a_1X + \ldots + a_{n-1}X^{n-1} + X^n$ un polynôme dans $K[X]$. Il existe donc $e \in \mathbb{N}^*$ tel que $(a_0, \ldots, a_{n-1}) \in K_e^n$. Si α est une racine de $P(X)$, alors, $K_e[X]$ étant un anneau principal, le polynôme minimal de α sur K_e divise $P(X)$. Donc, $[K_e(\alpha) : K_e] \leq n$, et en particulier, $[K_e(\alpha) : K_e]$ divise $N!$. En utilisant le fait 8.1.3 (2) ci-dessus, nous pouvons conclure que $\alpha \in K_{e+1}$. Cela montre que K satisfait tous les A_n pour $n \leq N$, et en particulier $\{A_{i_1}, \ldots, A_{i_m}\}$.

Par contre K n'est pas algébriquement clos. En effet, nous montrerons que pour tout $r \in \mathbb{N}$ $(r > 1)$, premier avec $N!$, K ne satisfait pas A_r. Soit donc L l'extension de $\mathbb{F}_p$ de degré r. Le fait

2 et la construction de K montrent que $L \cap K = \mathbb{F}_p$. Si $\alpha \in L \setminus \mathbb{F}_p$, alors son polynôme minimal Q sur $\mathbb{F}_p$ divise $x^{p^r} - x$ d'après le fait 8.1.3 (1). Comme Q est irréductible sur $\mathbb{F}_p$, il n'a pas de racine dans $\mathbb{F}_p$. Or toutes ces racines sont dans L d'après le fait 8.1.3 (1). Donc, Q n'a pas de racine dans K.

Nous étendons maintenant ce qui précède à la caractéristique 0. Le raisonnement est par compacité. Nous considérons l'ensemble T qui consiste des axiomes des corps, $1 \neq 0$ et

$$\{A_{i_1}, \ldots, A_{i_m}\} \cup \{\neg A_r\} \cup \{C_n : n \in \mathbb{N}^*\}$$

où r est un nombre naturel premier avec $N!$. Cet ensemble est consistant. En effet, soit T_0 un sous-ensemble fini.

$$T_0 = (T_0 \cap \{A_{i_1}, \ldots, A_{i_m}\}) \cup \{\neg A_r\} \cup \{C_{j_1}, \ldots, C_{j_k}\}$$

Le corps K construit dans la première partie de la preuve est un modèle de cet ensemble fini à condition que sa caractéristique soit strictement supérieure à $\max(j_1, \ldots, j_k)$. Ainsi, l'ensemble T est consistant. Or un modèle de T est un corps de caractéristique 0 qui satisfait $\{A_{i_1}, \ldots, A_{i_m}\}$, mais qui n'est pas algébriquement clos.

Maintenant nous pouvons conclure que la théorie des corps algébriquement clos d'une caractéristique donnée n'est pas finiment axiomatisable. Supposons par l'absurde qu'il existe une axiomatisation finie. D'après le lemme 8.1.2, l'axiomatisation que nous avons donnée contient une qui est finie. Or, nous avons vu que pour toute caractéristique p, toute partie finie des axiomes de CAC_p admet des modèles qui sont des corps non algébriquement clos. □

8.2 Complétude ; élimination des quantificateurs ; géométrie algébrique

Dans le reste de ce chapitre, la théorie CAC_p, avec p premier ou 0, sera notre objet d'étude. Le langage, quitte à ajouter des paramètres, sera $\mathcal{L}_C$. Dans cette section, nous établirons la complétude de CAC_p. La méthode sera le va-et-vient, et permettra aussi de caractériser les types dans $S(T)$. De cette caractérisation découlera l'élimination des quantificateurs dans le langage $\mathcal{L}_C$.

Les deux lemmes suivants établissent le cadre pour le va-et-vient. Notons que, comme le langage $\mathcal{L}_C$ ne fait intervenir que des symboles de fonction, les termes jouent un rôle important dans les raisonnements qui utilisent la récurrence sur la complexité des formules. Comme le langage $\mathcal{L}_C$ inclut les inversions par rapport à l'addition et à la multiplication, on montre par récurrence sur la complexité de l'écriture (exercice) que les termes sont des polynômes à coefficients dans le corps premier.

Lemme 8.2.1 *Soient K et L deux corps de même caractéristique, et k et l les corps premiers dans K et L respectivement. Si $(a_1, \ldots, a_m) \in K^m$ et $(b_1, \ldots, b_m) \in L^m$, alors $(a_1, \ldots, a_m)$ et $(b_1, \ldots, b_m)$ satisfont les mêmes formules sans quantificateurs dans le langage $\mathcal{L}_C$ si et seulement si $k(a_1, \ldots, a_m) \cong l(b_1, \ldots, b_m)$.*

Preuve. Il suffit de faire le raisonnement pour $m = 1$ puisque pour m plus large, on procède par récurrence en répétant le même raisonnement pour les corps isomorphes $k(a_1, \ldots, a_{m-1})$, $l(b_1, \ldots, b_{m-1})$ et les éléments a_m et b_m. Donc on remplace $(a_1, \ldots, a_m)$ par a et $(b_1, \ldots, b_m)$ par b. Notons aussi que l'hypothèse sur la caractéristique de K et L implique que $k \cong l$.

Si a et b satisfont les mêmes formules sans quantificateurs alors ils sont les racines des mêmes polynômes à coefficients dans k et l respectivement, par récurrence sur la complexité des formules atomiques (voir la définition 5.4.2 et l'exemple 5.4.2). Alors, $k(a)$ est soit isomorphe à $k(X)$ parce que a est transcendant sur k, soit isomorphe à $k[X]/\langle P(X)\rangle$ avec $P(X)$ le polynôme minimal de a (c'est un bon moment pour réviser la théorie de Galois élémentaire si cette dernière conclusion

n'est pas claire). Comme $k \cong l$, notre hypothèse sur a et b montre que les mêmes conditions sont satisfaites pour $l(b)$, et par conséquent $k(a) \cong l(b)$.

Dans l'autre direction supposons $k(a) \cong l(b)$. Nous procédons par récurrence sur la complexité des formules sans quantificateurs. Une formule atomique dans le langage $\mathcal{L}_C$ correspond à une équation polynômiale de la forme $t(X) = 0$ où les coefficients de t sont dans le corps premier (la définition 5.4.2 et l'exemple 5.4.2). Puisque $k(a) \cong l(b)$, $t(a) = 0$ si et seulement si $t(b) = 0$. Les formules sans quantificateurs plus complexes étant des conjonctions d'équations et d'inéquations à coefficients dans le corps premier, la conclusion découle du cas atomique. □

Lemme 8.2.2 *Soient K et L deux modèles ω-saturés de CAC_p. Si $= (a_1, \ldots, a_m) \in K^m$ et $(b_1, \ldots, b_m) \in L^m$ satisfont les mêmes formules sans quantificateurs et que $\alpha \in K$, alors il existe $\beta \in K$ tel que $\langle a_1, \ldots, a_m, \alpha\rangle \cong \langle b_1, \ldots, b_m, \beta\rangle$ où $\langle a_1, \ldots, a_m, \alpha\rangle$ (resp. $\langle b_1, \ldots, b_m, \beta\rangle$) est le sous-corps engendré par les éléments cités.*

Preuve. Le lemme 8.2.1 montre que les sous-corps engendrés par $\{a_1, \ldots, a_m\}$ et $\{b_1, \ldots, b_m\}$ sont isomorphes. On les note $k_{\overline{a}}$ et $k_{\overline{b}}$ respectivement.

Le raisonnement se divise en deux cas suivant si α est algébrique sur $k_{\overline{a}}$ ou non. Si α est algébrique sur $k_{\overline{a}}$, on considère son polynôme minimal sur $k_{\overline{a}}$ qui est $P(X)$ et que l'on transforme en $Q(X)$ en utilisant l'isomorphisme entre $k_{\overline{a}}$ et $k_{\overline{b}}$. Comme L est algébriquement clos, il contient une racine β de $Q(X)$. On associe β à α. Les corps $k_{\overline{a}}(\alpha)$ et $k_{\overline{b}}(\beta)$ sont isomorphes.

La deuxième possibilité est que α soit transcendant sur $k_{\overline{a}}$. Considérons la famille suivante de formules à une seule variable libre x :

$$p(x) = \{Q(x) \neq 0 \mid Q(X) \in k_{\overline{b}}[X]\} \ .$$

Chaque partie finie de cette famille est satisfaite par un élément de L puisque L est infini et qu'il existe un nombre fini de racines d'un polynôme à une variable. Donc p est consistant avec $\mathrm{Th}(L)$. Or, L est ω-saturé et l'ensemble $p(x)$ ne fait intervenir qu'un nombre fini de paramètres de L, notamment $\{b_1, \ldots, b_m\}$. Par conséquent, L contient une réalisation β de $p(x)$. Cette réalisation est transcendante sur $k_{\overline{b}}$, et il en découle que $k_{\overline{a}}(\alpha) \cong k_{\overline{b}}(\beta)$ dans ce cas aussi. □

Proposition 8.2.3 *Soient K et L deux modèles de CAC_p, $(a_1, \ldots, a_m) \in K^m$ et $(b_1, \ldots, b_m) \in L^m$. Alors $\mathrm{tp}_K(a_1, \ldots, a_m) = \mathrm{tp}_L(b_1, \ldots, b_m)$ si et seulement si ces deux m-uples satisfont les même formules sans quantificateurs dans K et L respectivement.*

Preuve. Nous pouvons supposer K et L ω-saturés. En effet, d'après le théorème 7.1, chacune de ces deux structures a une extension élémentaire ω-saturée qui est aussi modèle de CAC_p.

La nécessité de la condition étant claire, il reste à vérifier sa suffisance. Ayant établi dans le lemme 8.2.2 les conditions du va-et-vient suffisant pour ce faire, nous procédons à une récurrence sur la complexité des formules.

Soit alors $\phi(x_1, \ldots, x_m)$ une $\mathcal{L}_C$-formule dont les variables libres sont parmi $\{x_1, \ldots, x_m\}$ telle que $K \models \phi[(a_1, \ldots, a_m)]$. Si ϕ est sans quantificateur, alors $L \models \phi[(b_1, \ldots, b_m)]$ par l'hypothèse de la proposition. Il reste à étudier le cas où $\phi(x_1, \ldots, x_m)$ est de la forme $\exists y \theta(x_1, \ldots, x_m, y)$. Alors il existe $\alpha \in K$ tel que $K \models \theta[(a_1, \ldots, a_m, \alpha)]$. Le lemme 8.2.2 montre alors qu'il existe $\beta \in L$ tel que $\langle a_1 \ldots, a_m, \alpha\rangle \cong \langle b_1, \ldots, b_m, \beta\rangle$. Le lemme 8.2.1 montre alors que $(a_1 \ldots, a_m, \alpha)$ et $(b_1, \ldots, b_m, \beta)$ satisfont les mêmes formules libres. Par récurrence sur la complexité, $L \models \theta[(b_1, \ldots, b_m, \beta)]$, et $L \models \exists y \theta([(b_1, \ldots, b_m)], y)$; en d'autres termes $L \models \phi[(b_1, \ldots, b_m)]$. □

Théorème 8.2 *Pour tout p premier ou 0, CAC_p est complète.*

Preuve. Tout a été fait dans la preuve de la proposition 8.2.3. □

La proposition 8.2.3 caractérise les types de CAC_p pour p fixé. La proposition suivante est pertinente dans ce contexte. Bien qu'elle ait été abordée pendant les td, nous incluons cette version ici dans l'attente qu'elle puisse apporter une approche légèrement différente :

Proposition 8.2.4 *Soient $\mathcal{L}$ un langage, T une $\mathcal{L}$-théorie non nécessairement complète, $\mathcal{F}$ une famille de formules à variables libres $x_1, \ldots, x_k$ ($k \in \mathbb{N}^*$) exactement tels que le type d'un k-uple extrait d'un modèle de T soit déterminé par $\mathcal{F}$. En d'autres termes, deux tels k-uples ont même type s'ils satisfont les mêmes formules de $\mathcal{F}$. Alors pour toute formule $\phi(x_1, \ldots, x_k)$ à exactement k variables libres de $\mathcal{L}$, il existe $\psi(x_1, \ldots, x_k) \in \mathcal{F}$ telle que*

$$T \vdash \forall x_1 \ldots x_k (\phi(x_1, \ldots, x_k) \leftrightarrow \psi(x_1, \ldots, x_k)) \ .$$

Preuve. Soit $\phi(x_1, \ldots, x_k)$ une formule comme dans l'énoncé. Nous pouvons supposer que

$$T \vdash \exists x_1 \ldots x_k y_1 \ldots y_k (\phi(x_1, \ldots, x_k) \wedge \neg\phi(y_1, \ldots, y_k)) \ .$$

Soit $\mathcal{M}$ un modèle de T d'univers M. Alors il existe $(d_1, \ldots, d_k) \in M^k$ tel que

$$\mathcal{M} \models \phi[(d_1, \ldots, d_k)] \ .$$

Pour $(e_1, \ldots, e_k) \in M^k$ tel que $\mathcal{M} \not\models \phi[(e_1, \ldots, e_k)]$, il existe, d'après l'hypothèse sur la détermination des types, $\psi \in \mathrm{tp}_{\mathcal{M}}(d_1, \ldots, d_k) \cap \mathcal{F}$ qui vérifie les conditions suivantes :

$$\mathcal{M} \models \psi[(d_1, \ldots, d_k)] \quad \text{et} \quad \mathcal{M} \models \neg\psi[(e_1, \ldots, e_k)] \ .$$

Nous montrerons maintenant que toute réalisation de $\mathrm{tp}_{\mathcal{M}}(d_1, \ldots, d_k)$ dans un modèle de T satisfait $\phi(x_1, \ldots, x_k)$ dans ce modèle, ce qui peut s'abbrévier

$$T \cup (\mathrm{tp}_{\mathcal{M}}(d_1, \ldots, d_k) \cap \mathcal{F}) \ \vdash \ \phi(x_1, \ldots, x_k) \ .$$

Sinon, on pose $\mathcal{L}^+ = \mathcal{L} \cup \{c_1, \ldots, c_k\}$ et on considère l'ensemble des formules

$$\tilde{T} = T \cup \left\{ \neg\phi(c_1, \ldots, c_k) \wedge \bigwedge_{i=1}^{k} \psi_i(c_1, \ldots, c_k) \mid \psi_i \in \mathrm{tp}_{\mathcal{M}}(d_1, \ldots, d_k) \cap \mathcal{F} \right\} \ .$$

D'après l'hypothèse contradictoire, il existe un modèle $\mathcal{N}$ de T tel que

$$\mathcal{N} \models T \cup (\mathrm{tp}_{\mathcal{M}}(d_1, \ldots, d_k) \cap \mathcal{F}) \cup \neg\phi(x_1, \ldots, x_k) \ .$$

Cette structure $\mathcal{N}$ sert de base pour réaliser les parties finies de $\tilde{T}$. Par compacité, on trouve un modèle de $\tilde{T}$, donc en particulier un modèle $\mathcal{S}$ de T quand on réduit le langage à $\mathcal{L}$, contenant un k-uple qui interprète $(c_1, \ldots, c_k)$. Or $\mathrm{tp}_{\mathcal{S}}(c_1, \ldots, c_k) = \mathrm{tp}_{\mathcal{M}}(d_1, \ldots, d_k)$ puisque les deux k-uplets satisfont les mêmes formules de $\mathcal{F}$. C'est une contradiction.

Par compacité, il existe une partie finie de $T \cup (\mathrm{tp}_{\mathcal{M}}(d_1, \ldots, d_k) \cap \mathcal{F})$ dont $\phi(x_1, \ldots, x_k)$ est conséquence. En d'autres termes, il existe $\psi_1, \ldots, \psi_n \in \mathrm{tp}_{\mathcal{M}}(d_1, \ldots, d_k) \cap \mathcal{F}$ telles que

$$T \vdash \forall x_1, \ldots, x_k \left(\bigwedge_{i=1}^{n} \psi_i(x_1, \ldots, x_k) \rightarrow \phi(x_1, \ldots, x_k) \right) \ .$$

On répète maintenant le même raisonnement avec chaque k-uple de M^k qui satisfait ϕ dans $\mathcal{M}$. Ceci donne un recouvrement de ϕ par des formules de $\mathcal{F}$ qu'on peut remplacer par un recouvrement fini en appliquant la compacité une dernière fois. □

Clairement, tout a été préparé pour conclure le théorème suivant :

Théorème 8.3 *Dans le langage $\mathcal{L}_C$, la théorie CAC_p élimine les quantificateurs. En d'autres termes, pour toute $\mathcal{L}_C$-formule $\phi(x_1, \ldots, x_k)$ ($k \in \mathbb{N}^*$) à exactement k variables libres, il existe une $\mathcal{L}_C$-formule $\psi(x_1, \ldots, x_k)$ à k variables libres telle que*

$$\mathrm{CAC}_p \vdash \forall x_1 \ldots x_k (\phi(x_1, \ldots, x_k) \leftrightarrow \psi(x_1, \ldots, x_k)) \ .$$

Nos lecteurs reconnaîtront le lemme général suivant :

Lemme 8.2.5 *Soit T une théorie qui élimine les quantificateurs. Si $\mathcal{M}$ et $\mathcal{N}$ sont deux modèles de T d'univers M et N respectivement. Si $M \subset N$, alors $\mathcal{M} \preceq \mathcal{N}$.*

Preuve. C'est une conséquence du lemme 5.5.3 (1). □

Une théorie qui a la propriété du lemme 8.2.5 est dite *modèle-complète*. Avant d'élucider les liens de ce que nous avons accompli à la géométrie, il convient de souligner l'importance du choix du langage dans lequel nous travaillons. Nous avons déjà remarqué que le langage des anneaux, voire celui sans l'inversion additive est convenable pour étudier les coprs. Néanmoins, un tel contexte est limité pour l'étude de l'élimination des quantificateurs. En effet, dans un langage réduit par rapport à $\mathcal{L}_C$, l'élimination des quantificateurs ne serait pas possible puisque les sous-structures ne seraient pas des corps ce qui est nécessaire pour une théorie des corps qui élimine les quantificateurs d'après le lemme 8.2.5.

Jusqu'à maintenant nous avons mené une étude des corps algébriquement clos en utilisant les méthodes et les outils de la théorie des modèles. Cette approche nous a permis de démontrer un résultat typique de la théorie des modèles, en l'occurrence l'élimination des quantificateurs dans un certain langage. Les corps algébriquement clos offrent un cadre où cette approche permet une interaction avec un autre domaine des mathématiques, la géométrie algébrique. Le théorème suivant, bien connu en algèbre commutative, appartient à ce paysage :

Théorème 8.4 (Théorème des zéros de Hilbert) *Si un système* fini S *d'équations et d'inéquations exprimées dans le langage $\mathcal{L}_C$, aux inconnues $x_1, \ldots, x_k$ et à coefficients dans un corps k de caractéristique p éventuellement* 0*, a une solution dans une extension K de k, alors il a une solution dans toute extension algébriquement close de k.*

Preuve. Nous pouvons considérer S comme une formule sans quantificateurs du premier ordre $S(x_1, \ldots, x_k; a_1, \ldots, a_l)$ où $\{a_1, \ldots, a_l\}$ sont des paramètres provenant de k qui nomment les coefficients dans le système S.

D'après le théorème 8.3 et le lemme 8.2.5, la théorie CAC_p est modèle-complète. Soit L un corps algébriquement clos contenant K. Comme

$$K \models \exists x_1 \ldots x_k S(x_1, \ldots, x_k; [(a_1, \ldots, a_l)])$$

et que $S(x_1, \ldots, x_k; a_1, \ldots, a_l)$ est sans quantificateurs, le lemme 5.5.3 (2) montre que

$$L \models \exists x_1 \ldots x_k S(x_1, \ldots, x_k; [(a_1, \ldots, a_l)]) \ .$$

Si maintenant L' est un corps algébriquement clos qui contient k, alors L' contient la clôture algébrique $\tilde{k}$ de k. Comme T est modèle-complète, $\tilde{k} \preceq L$ et

$$\tilde{k} \models \exists x_1 \ldots x_k S(x_1, \ldots, x_k; [(a_1, \ldots, a_l)]) \ .$$

Par conséquent,

$$L' \models \exists x_1 \ldots x_k S(x_1, \ldots, x_k; [(a_1, \ldots, a_l)]) \ .$$

□

Notre préparation méticuleuse a permis une preuve simple donc conceptuellement claire d'un résultat en géométrie algébrique. Avant de finir cette section, nous aborderons une autre conséquence fondamentale de l'élimination des quantificateurs qui, comme la modèle-complétude de CAC_p a une interprétation géométrique. Commençons par rappeler une notion déjà introduite dans les dm :

Définition 8.2.6 *Soient $\mathcal{L}$ un langage et $\mathcal{M}$ une $\mathcal{L}$-structure infinie et d'univers M.*

1. *La structure $\mathcal{M}$ est dite* minimale *si pour toute $\mathcal{L}$-formule $\phi(x, y_1, \ldots, y_k)$ à $k+1$ variables libres avec $k \in \mathbb{N}$et $(m_1, \ldots, m_k) \in M^k$ la partie définissable $\phi(\mathcal{M}; m_1, \ldots, m_k)$ de M est soit finie soit cofinie.*
2. *La structure $\mathcal{M}$ est dite* fortement minimale *si toute extension élémentaire de $\mathcal{M}$ est minimale.*
3. *Une $\mathcal{L}$-théorie complète est dite* fortement minimale *si tous ses modèles sont minimaux.*

Théorème 8.5 *La théorie CAC_p est fortement minimale.*

Preuve. C'est une conséquence de l'élimination des quantificateurs. Soient K un corps algébriquement clos de caractéristique p, $\phi(x, y_1, \ldots, y_m)$ une formule à $m+1$ variables libres ($m \in \mathbb{N}$) et $(a_1, \ldots, a_m) \in K^m$. D'après le théorème 8.3, il existe une formule $\psi(x, y_1, \ldots, y_m)$ à $m+1$ variables libres et sans quantificateurs telle que

$$\mathrm{CAC}_p \vdash \forall x y_1 \ldots y_m (\phi(x, y_1, \ldots, y_m) \leftrightarrow \psi(x, y_1, \ldots, y_m)) \ .$$

En particulier,

$$K \models \forall x (\phi(x, a_1, \ldots, a_m) \leftrightarrow \psi(x, a_1, \ldots, a_m)) \ .$$

Or, $\psi(x, y_1, \ldots, y_m)$ est sans quantificateurs. Si elle est en outre atomique, alors il s'agit d'une équation polynomiale du type

$$P(X, Y_1, \ldots, Y_m) = 0 \ .$$

Or, l'équation

$$P(X, a_1, \ldots, a_m) = 0$$

a un nombre fini de solutions, et cette conclusion amorce une récurrence sur la compléxité des formules dans le langage $\mathcal{L}_C$. Si ψ est de la forme $\neg\theta$ alors $\psi(K, a_1, \ldots, a_m)$ est fini si et seulement si $\theta(K, a_1, \ldots, a_m)$ est cofini, et la conclusion pour le cardinal de $\psi(K, a_1 \ldots, a_m)$ découle de l'hypothèse de récurrence sur le cardinal de $\theta(K, a_1, \ldots, a_m)$. Une discussion légèrement plus compliquée règle le sort du cas où ψ est de la forme $\theta_1 \wedge \theta_2$.

Par ailleurs, étant algébriquement clos, K est un corps infini. Le corps K étant un modèle arbitrairement choisi de CAC_p, on conclut que CAC_p est fortement minimale. □

Ce dernier théorème aussi est de nature géométrique. En effet, la minimalité forte est une interprétation modèle-théorique d'une propriété de la topologie de Zariski sur l'ensemble K : les fermés sont petits, dans ce cas finis, les ouverts sont grands, dans ce cas, cofini. Notons que la nature géométrique des structures fortement minimales ne nécessite pas la présence d'un corps algébriquement clos. Leur cadre abstrait est suffisamment riche pour permettre de définir une notion d'indépendance et une notion de dimension qui en particulier montrent que leurs théories du premier ordre sont κ-catégoriques pour tout $\kappa > \aleph_0$.

8.3 Etude des modèles ω-saturés

Dans notre étude de la théorie CAC_p, nous avons utilisé les modèles ω-saturés pour établir le va-et-vient. Néanmoins, nous n'avons pas précisé leurs propriétés structurelles au delà d'être une ressource éventuellement infinie d'éléments transcendants indépendants entre eux aussi. En fait, cette propriété les caractérise.

Théorème 8.6 *Un modèle de* CAC_p *est ω-saturé si et seulement s'il est de degré de transcendance infini sur son corps premier.*

Preuve. Commençons par démontrer la nécessité de la condition. Soit K un modèle ω-saturé de CAC_p. Nous considérons le type $p_0(x)$ contenant les formules

$$\left\{ \sum_{i=0}^{n} a_i x^i \neq 0 : n \in \mathbb{N},\ a_i \in \mathbb{Z}_p \text{ ou } \mathbb{Z} \text{ selon } \mathrm{car}(K) \right\}.$$

$p_0(x)$ est réalisé par un élément transcendant sur le corps premier que nous notons α_0 que nous pouvons supposer dans K comme $p_0(x) \in S_1(\mathrm{CAC}_p)$. Nous considérons maintenant le type $p_1(x) \in S_1^K(\{\alpha_0\})$ qui contient l'union de

$$\left\{ \sum_{i=0}^{n} a_i x^i \neq 0 : n \in \mathbb{N},\ a_i \in \mathbb{Z}_p \text{ ou } \mathbb{Z} \text{ selon } \mathrm{car}(K) \right\}$$

et de

$$\left\{ \neg\exists x_{(i_0,i_1,j)} y_{(i_0,i_1)} \left(\left(\sum_{0\leq i_0+i_1\leq m} y_{(i_0,i_1)} \alpha_0^{i_0} x^{i_1} = 0 \right) \wedge \left(\sum_{j=0}^{n} x_{(i_0,i_1,j)} y_{(i_0,i_1)}^{n} = 0 \right) \right) : m, n \in \mathbb{N} \right\}.$$

Ces énoncés expriment le fait que la réalisation soit un élément transcendant indépendant de α_0. Une réalisation α_1 de p_1 appartient à K puisque K est ω-saturé. On continue de cette manière en définissant $p_{i+1}(x)$ qui exprime que sa réalisation est un élément transcendant indépendant des $\alpha_0, \ldots, \alpha_i$. Tous ces types sont réalisés dans K grâce à l'ω-saturation.

Pour ce qui est de la suffisance de la condition, nous montrons d'abord que si K est un modèle ω-saturé de CAC_p et que L est algébriquement clos et de degré de transcendance infini sur sur son corps premier alors ils sont ∞-équivalents. Le raisonnement est le même que ceux du lemme 8.2.2 et du théorème 8.2 sauf à un endroit. Si, dans la notation de la preuve du lemme 8.2.2, α est transcendant, il n'y a que l'hypothèse sur le degré de transcendance qui est disponible. Alors nous raisonnons de la façon suivante. Comme $(b_1, \ldots, b_m)$ est un m-uple, donc une suite finie d'éléments de L, le degré de transcendance du corps engendré par $(b_1, \ldots, b_m)$ ne peut être infini. Donc il existe un élément transcendant β qui est dans $L \setminus \langle b_1, \ldots, b_m \rangle$. Nous associons β à α. Pour conclure que L est ω-saturé il suffit d'appliquer le lemme 7.3.5 (1). □

La construction que nous avons faite dans la première moitié en obtenant à chaque étape un nouvel élément transcendant est bien connu en théorie des modèles. Sans le nommer, nous avons construit une *suite de Morley du type transcendant.*

Chapitre 9

Théories ω-catégoriques

Dans ce chapitre, nous étudierons les théories ω-catégoriques. Ce sont les théories qui, à isomorphisme près, ont un seul modèle dénombrable. Cette condition très restrictive offre néanmoins une classe de structures très variées dont l'étude a des liens avec plusieurs domaines des mathématiques dont les groupes de permutations, la topologie générale, la théorie des graphes.

Nous démontrerons la caractérisation fondamentale des théories ω-catégoriques en fonction de leurs types, à savoir le théorème de Ryll-Nardzewski. La preuve utilisera des méthodes et des idées qui sont utilisées dans d'autres parties de la théorie des modèles aussi. En outre, un théorème fondamental de la théorie des modèles sera crucial : le *théorème d'omission des types*. Nous omettrons sa preuve en raison de manque de temps. Les lecteurs intéressés sont encouragés à l'étudier.

Soulignons que certains théorèmes principaux de ce chapitre ne sont valables que pour les langages dénombrables. Certains d'entre eux sont aussi valables pour des théories non nécessairement complètes. Ces conditions seront indiquées aussi clairement que possible. Ce qui ne sera pas indiqué est que les structures que nous considérons sont toutes infinies. La théorie des ordres linéaires denses et sans extrémités offre une bonne illustration des idées générales qui jouent un role dans ce chapitre. Nous conseillons d'en faire usage pour mieux voir comment chaque idée abstraite se concrétise dans un exemple.

Comme dans le chapitre précédent, nous allégerons notre exposition. En particulier, nous utiliserons les mêmes symboles pour les symboles de constantes et les éléments qui les interprènt dans une structure. La notation $\bar{}$ sera plus fréquente et parfois, au lieu d'écrire $\mathcal{M} \models \phi[(a_1, \ldots, a_k)]$ nous écrirons pour la satisfaction dans $\mathcal{M} \models \phi(a_1, \ldots, a_k)$.

9.1 Théorème d'omission des types

Jusqu'à présent, nous avons étudié les réalisations des types et les conditions qui les rendent possible dans un modèle d'une théorie complète. Or, leur omission n'en est pas moins important. En effet, l'étude de l'extension d'un type à un ensemble de paramètres plus large montre qu'une extension qui ressemble le plus au type initial est un type qui n'est pas "plus" réalisé que le type initial.

La notion d'isolation d'un type est d'une importance majeure.

Définition 9.1.1

1. Soient $\mathcal{L}$ un langage, T une $\mathcal{L}$-théorie, $\mathcal{M}$ un modèle de T avec univers M et $A \subset M$. Une famille $\Phi(x_1, \ldots, x_k)$ de $\mathcal{L}$-formules à au plus k variables libres, est dite isolée *par une $\mathcal{L}$-formule $\sigma(x_1, \ldots, x_k)$ si*

(i) $T \vdash \exists x_1 \ldots x_k \sigma(x_1, \ldots, x_k)$, et

(ii) pour toute $\phi(x_1, \ldots, x_k) \in \Phi(x_1, \ldots, x_k)$,

$$T \vdash \forall x_1 \ldots x_k (\sigma(x_1 \ldots x_k) \to \phi(x_1 \ldots x_k)) \ .$$

2. Un type $p \in S_k^{\mathcal{M}}(A)$ est isolé si p est isolé par rapport à la théorie $\mathrm{Th}((\mathcal{M}, a)_{a\in A})$. *Par maximalité d'un type, une formule isolant p appartient à p.*

Comme la terminologie l'indique, la notion de type isolé a un caractère topologique. En effet, les types isolé correspondent à des points isolés d'un espace compact. Les voisinages qui les isolent sont déterminés par les formules qui les isolent.

Une autre appellation pour un type isolé est *principal*. Bien sûr, l'appellation souligner un lien avec les filtres principaux.

Nous illustrerons la notion de type isolé par quelques exemples :

Exemples 9.1.1 1. Soient $\mathcal{L}$, $\mathcal{M}$ et A comme dans la définition 9.1.1. Un 1-type p à paramètres dans A est dit *algébrique sur A* s'il contient une $\mathcal{L}(A)$-formule ϕ à au plus k variables libres qui est satisfaite par un nombre fini d'éléments dans $\mathcal{M}$, en d'autres termes $\phi(\mathcal{M})$ est une partie finie de M. Comme le nombre d'éléments dans $\phi(\mathcal{M})$ s'exprime par une formule du premier ordre de la forme $\exists^{=m} x\phi(x)$, la propriété de finitude de ϕ reste valable pour tout modèle de $\mathrm{Th}((\mathcal{M}, a)_{a\in A})$. Le type p est isolé par une formule algébrique décrivant l'ensemble fini du plus petit cardinal, en d'autres termes, par une formule correspondant au plus petit m, m étant comme dans la phrase précédente.

Pour concrétiser les notions abstraites du paragraphe précédent, les corps algébriquement clos fournissent un bon exemple. Comme dans le chapitre précédent, nous nous mettons dans le langage $\mathcal{L}_C = \{+, ., -, {}^{-1}, 0, 1\}$ et étudions la théorie CAC_p pour p premier ou 0. Considérons $S_1(\mathrm{CAC}_p)$. Si $p \in S_1(\mathrm{CAC}_p)$, alors il découle de l'élimination des quantificateurs dans le langage $\mathcal{L}_C$ qu'il y a deux possibilités pour p : soit p contient une formule de la forme $P(x) = 0$ où $P(X) \in \mathbb{Q}[X]$ ou $P(X) \in \mathbb{F}_p[X]$ suivant la valeur de p ; soit p ne contient aucune équation polynômiale de ce type, en d'autres termes, il est réalisé par un (tous les) élément(s) transcendants sur la clôture algébrique de $\mathbb{Q}$ ou $\mathbb{F}_p$ dans une extension élémentaire de celui-ci. Le premier cas correspond à un type algébrique, le deuxième à un type transcendant/omis/non isolé. La clôture algébrique du corps premier est un modèle *dénombrable* qui omet le type transcendant. Tout extension élémentaire plus grande, en d'autres termes de degré de transcendance non nul, de la clôture algébrique du corps premier contient inévitablement une réalisation du type transcendant. En particulier, ceci concerne les extensions élémentaires non dénombrables.

2. Pour motiver la discussion qui aura lieu dans ce chapitre, considérons le langage $\mathcal{L} = \{<\}$ avec $<$ un symbole de relation binaire $\mathcal{Q} = (\mathbb{Q}, <)$ en tant que $\mathcal{L}$-structure. Pour toute partie finie, éventuellement vide, A de $\mathbb{Q}$, tout élément de $S_k^{\mathcal{M}}(A)$ est isolé. En effet, comme la théorie des ordres denses linéaires sans extrémités éliminent les quantificateurs dans le langage $\mathcal{L}$, tout k-type sur A est déterminé par le positionnement des coordonnées d'une réalisation entre elles et par rapport aux éléments de A. En d'autres termes, si $(x_1, \ldots, x_k) \in \mathbb{Q}^k$ et que $A = \{a_1, \ldots, a_m\}$, alors le type de $(x_1, \ldots, x_k)$ sur A, est déterminé par la conjonction des formules de la forme $x_i < a_j$, $x_i = a_j$, $a_j < x_i$ et de leurs négations. Cette conjonction isole le type en question. Par contre, aucun 1-type sur A n'est algébrique. Cette dernière conlusion se vérifie en utilisant les automorphismes de $(\mathbb{Q}, <)$ qui fixent les éléments de A. En fait, $(\mathbb{Q}, <)$ est une structure qui a beaucoup d'automorphismes et très peu de k-types (en nombre fini) sur un ensemble fini de paramètres.

3. Maintenant considérons le langage $\mathcal{L} = \{R\}$ avec R un symbole de relation binaire. La théorie T d'une relation d'équivalence à une classe finie et une seule pour tout nombre naturel a un comportement fort différent de la théorie des ordres denses, linéaires et sans extrémités. Si p est un type dans $S_1(T)$ contenant $\{\neg C_i(x) | i \in \mathbb{N}^*\}$ où les C_i sont les formules qui expriment que x appartient à la classe à i éléments, alors p n'est pas isolé. En effet, s'il existait une formule $\sigma(x)$ isolant p, alors $\neg\sigma(x)$ serait une définition pour les éléments appartenant à une classe finie dans tout modèle de T. Or une telle formule n'existe pas. C'est un bon entraînement de vérifier les détails de ce raisonnement.

Le lemme suivant donne un premier aperçu du comportement des types isolés.

Lemme 9.1.2 *Soient $\mathcal{L}$ un langage, T une $\mathcal{L}$-théorie complète, $\mathcal{M}$ un modèle de T d'univers M, et $A \subset M$. Si $p \in S_k^{\mathcal{M}}(A)$ est un type isolé, alors p est réalisé dans $\mathcal{M}$, et plus généralement dans tout modèle de T contenant A.*

Preuve. Soit σ une $\mathcal{L}(A)$-formule qui isole p. Comme T est une théorie complète, $T \vdash \exists x_1 \dots x_k \sigma(x_1, \dots, x_k)$. En particulier, tout modèle contenant de T contenant A, sera satisfaite par un k-uple qui sera aussi une réalisation de p. □

Voici le théorème fondamental, dans le contexte des langages dénombrables, qui lie l'omission d'un type à sa (non) réalisation :

Théorème 9.1 *Soit T une théorie non nécessairement complète dans un langage $\mathcal{L}$ dénombrable. Si $\{\Phi_i(x_1, \dots, x_k) \mid i \in \mathbb{N}\}$ est une famille d'ensembles de formules à variables libres parmi $\{x_1, \dots, x_k\}$ chacune consistante avec T. On suppose qu'aucune des $\Phi_i(x_1, \dots, x_k)$ ne soit isolée. En d'autres termes, pour tout $i \in \mathbb{N}$, pour toute $\mathcal{L}$-formule $\sigma(x_1, \dots, x_k)$ à au plus k variables libres, il existe $\phi(x_1, \dots, x_k) \in \Phi_i(x_1, \dots, x_k)$ qui vérifie la condition suivante :*

$$T \vdash \exists \overline{x}(\sigma(\overline{x}) \wedge \neg\phi(\overline{x})) \ .$$

Alors T a un modèle dénombrable qui omet chaque $\Phi_i(x_1, \dots, x_k)$.

Remarques : 1. Le langage est nécessairement dénombrable pour que le théorème 9.1 soit valable. L'exemple suivant illustre cette nécessité. Nous nous plaçons dans le langage qui contient une infinité non dénombrable de symboles de constantes et l'égalité comme la seule relation, soit $\mathcal{L} = \{c_i \mid i < \omega_1\}$. La théorie T sera formée des conséquences de l'ensemble $\{c_i \neq c_j \mid i, j < \omega_1\}$.

Ensuite, nous augmentons le langage en y ajoutant une infinité dénombrable de symboles de constantes : $\{d_n | n \in \mathbb{N}\}$. La famille $\{x \neq d_n \mid n \in \mathbb{N}\}$ vérifie les hypothèses du théorème d'omission des types. Pourtant, elle est réalisée dans chaque modèle de T puisque chaque modèle de T est non dénombrable.

2. Le modèle fourni par le théorème d'omission des types est dénombrable aussi. Ceci est inévitable comme l'indique la discussion des corps algébriquement clos dans l'exemple 9.1.1 : un modèle non dénombrable de CAC_p ne peut pas omettre le type (sur $\emptyset$) transcendant. Pour l'omettre il faudra introduire de nouveaux paramètres.

9.2 Types isolés ; modèles atomiques ; modèles premiers

Dans cette section, nous étudierons les liens entre les types isolés et les modèles "les plus petits" d'une théorie complète. La définition suivante propose une notion de "petit" :

Définition 9.2.1 *Soient $\mathcal{L}$ un langage du premier ordre, T une $\mathcal{L}$-théorie complète, $\mathcal{M}$ un modèle de T d'univers M, et $A \subset M$.*

1. *Si $A \subset B \subset M$, alors B est dit* atomique *sur A si pour tout $\overline{b} \in B^k$ ($k \in \mathbb{N}$) $\mathrm{tp}_{\mathcal{M}}(\overline{b}/A)$ est isolé.*
2. *Un modèle $\mathcal{M}$ de T est dit* premier *si $\mathcal{M}$ se plonge élémentairement dans chaque modèle de T.*
3. *On dit que $\mathcal{M}$ est premier sur A si $(\mathcal{M}, a)_{a \in A}$ est un modèle premier de $\mathrm{Th}((\mathcal{M}, a)_{a \in A})$.*

Nous aurons besoin de certaines propriétés élémentaires mais cruciales des types isolés. L'idée motrice de la première provient de la nature topologique de la projection sur une coordonnée : c'est une application à la fois continue et ouverte.

Lemme 9.2.2 *Soient $\mathcal{L}$ un langage du premier ordre, T une $\mathcal{L}$-théorie complète et $\mathcal{M}$ un modèle de T d'univers M. Si $\overline{a} \in M^k$ et $\overline{b} \in M^l$, alors $\mathrm{tp}_{\mathcal{M}}(\overline{a}, \overline{b})$ est isolé si et seulement si $\mathrm{tp}_{\mathcal{M}}(\overline{a})$ (resp. $\mathrm{tp}_{\mathcal{M}}(\overline{b})$) est isolé et $\mathrm{tp}_{\mathcal{M}}(\overline{b}/\overline{a})$ (resp. $\mathrm{tp}_{\mathcal{M}}(\overline{a}\overline{b})$) est isolé.*

Preuve. Si $\phi(\overline{x}, \overline{y})$ isole $\mathrm{tp}_{\mathcal{M}}(\overline{a}, \overline{b})$ alors $\exists \overline{y} \phi(\overline{x}, \overline{y})$ isole $\mathrm{tp}_{\mathcal{M}}(\overline{a})$ et $\phi(\overline{a}, \overline{y})$ isole $\mathrm{tp}_{\mathcal{M}}(\overline{b}/\overline{a})$.

Si $\psi(\overline{x})$ isole $\mathrm{tp}_{\mathcal{M}}(\overline{a})$ et que $\theta(\overline{a}, \overline{y})$ isole $\mathrm{tp}_{\mathcal{M}}(\overline{b}/\overline{a})$, $\psi(\overline{x}) \wedge \theta(\overline{x}, \overline{y})$ isole $\mathrm{tp}_{\mathcal{M}}(\overline{a}, \overline{b})$. □

Lemme 9.2.3 *Soient $\mathcal{L}$ un langage du premier ordre, T une $\mathcal{L}$-théorie complète et $\mathcal{M}$ un modèle de T d'univers M. Soit $A \subset M$. Si A est atomique sur $\emptyset$ alors pour toute partie finie $\{a_1, \ldots, a_k\}$ de A, A est atomique sur $\{a_1, \ldots, a_k\}$.*

Preuve. C'est une application du lemme 9.2.2. □

Le lemme suivant illustre la transitivité de l'atomicité.

Lemme 9.2.4 *Soient $\mathcal{L}$ un langage du premier ordre, T une $\mathcal{L}$-théorie complète et $\mathcal{M}$ un modèle de T d'univers M. Si $A \subset B \subset C \subset M$, B est atomique sur A et C atomique sur B, alors C est atomique sur A.*

Preuve. C'est une occasion pour s'entraîner. □

Les types isolés permettent de faire le va-et-vient.

Proposition 9.2.5 *Soient $\mathcal{L}$ un langage, T une $\mathcal{L}$-théorie complète, $\mathcal{M}$ et $\mathcal{N}$ deux modèles de T d'univers M et N respectivement. Si $\mathcal{M}$ et $\mathcal{N}$ sont atomiques alors $\mathcal{M}$ et $\mathcal{N}$ sont ∞-équivalents.*

Preuve. Soient $\overline{a}$ et $\overline{b}$ deux k-uples extraits de M et de N tels que $\mathrm{tp}_{\mathcal{M}}(\overline{a}) = \mathrm{tp}_{\mathcal{N}}(\overline{b})$, et α un élément quelconque de M. Par hypothèse, $\mathrm{tp}_{\mathcal{M}}(\overline{a})$ et $\mathrm{tp}_{\mathcal{M}}(\overline{a}, \alpha)$ sont isolés par σ et θ respectivement. Alors

$$\mathcal{M} \models \sigma(\overline{a}) \rightarrow \exists x_{k+1} \theta(\overline{a}, x_{k+1})$$

puisque

$$\mathcal{M} \models \theta(\overline{a}, \alpha) \ .$$

Alors

$$\mathcal{N} \models \sigma(\overline{b}) \rightarrow \exists x_{k+1} \theta(\overline{b}, x_{k+1}) \ .$$

Soit donc $\beta \in N$ tel que

$$\mathcal{N} \models \theta(\overline{b}, \beta) \ .$$

Comme θ isole $\mathrm{tp}_{\mathcal{M}}(\overline{a}, \alpha)$, pour toute $\mathcal{L}$-formule $\phi(\overline{x}, y)$ dans $\mathrm{tp}_{\mathcal{M}}(\overline{a}, \alpha)$

$$\mathcal{M} \models \forall \overline{x} y (\theta(\overline{x}, y) \rightarrow \phi(\overline{x}, y)) \ .$$

Or $\mathcal{M} \equiv \mathcal{N}$. Ainsi,

$$\mathcal{N} \models \forall \overline{x} y (\theta(\overline{x}, y) \rightarrow \phi(\overline{x}, y)) \ ,$$

et $\phi(\overline{x}, y) \in \mathrm{tp}(\overline{b}, \beta)$. En particulier, $\mathrm{tp}_{\mathcal{M}}(\overline{a}, \alpha) \subseteq \mathrm{tp}_{\mathcal{N}}(\overline{b}, \beta)$. Par maximalité des types, on conclut que c'est en fait une égalité. □

Corollaire 9.2.6 *Soient $\mathcal{L}$ un langage, T une $\mathcal{L}$-théorie complète, $\mathcal{M}$ et $\mathcal{N}$ deux modèles de T d'univers M et N respectivement. Si $\mathcal{M}$ et $\mathcal{N}$ sont atomiques et dénombrables, alors $\mathcal{M}$ et $\mathcal{N}$ sont isomorphes.*

Corollaire 9.2.7 *Soient $\mathcal{L}$ un langage, T une $\mathcal{L}$-théorie complète, et $\mathcal{M}$ un modèle atomique dénombrable de T d'univers M. Soient $(a_1, \ldots, a_k)$, $(b_1, \ldots, b_k) \in M^k$ tels que $\mathrm{tp}_{\mathcal{M}}(a_1, \ldots, a_k) = \mathrm{tp}_{\mathcal{M}}(b_1, \ldots, b_k)$. Alors il existe un automorphisme ν de $\mathcal{M}$ tel que $\nu(a_i) = b_i$ pour $i \in \{1, \ldots, k\}$. Dans ce cas, on dit parfois que $\mathcal{M}$ est une structure* homogène.

Quand le langage est dénombrable, on peut caractériser les modèles atomiques et premiers.

Théorème 9.2 *Soient $\mathcal{L}$ un langage dénombrable, et T une $\mathcal{L}$-théorie complète. Alors, T a un modèle premier si et seulement s'il a un modèle atomique et dénombrable. Le modèle premier est à isomorphisme près le seul modèle atomique dénombrable.*

Preuve. Supposons d'abord $\mathcal{M}$ premier. Comme, d'après le théorème de Löwenheim-Skolem, T a un modèle dénombrable, et que $\mathcal{M}$, étant premier se plonge élémentairement dans celui-ci, $\mathcal{M}$ est nécessairement dénombrable. Nous montrerons que $\mathcal{M}$ est atomique. Supposons que ce ne soit pas le cas. Alors, il existe $k \in \mathbb{N}^*$ et $\overline{a} = (a_1, \ldots, a_k) \in M^k$ tels que $\mathrm{tp}_{\mathcal{M}}(\overline{a})$ ne soit pas isolé. D'après le théorème d'omission des types, il existe un modèle dénombrable $\mathcal{N}$ de T qui omet $\mathrm{tp}_{\mathcal{M}}(\overline{a})$. Or $\mathcal{M}$, étant premier par hypothèse, se plonge élémentairement dans $\mathcal{N}$. D'après le test de Tarski, l'image de $\overline{a}$ sous l'action de ce plongement est une réalisation dans $\mathcal{N}$ de $\mathrm{tp}_{\mathcal{M}}(\overline{a})$. C'est une contradiction.

Supposons maintenant $\mathcal{M}$ atomique. Alors, il découle du théorème de Löwenheim-Skolem que T a aussi un modèle dénombrable et atomique. On peut donc supposer $\mathcal{M}$ dénombrable. Soit $\mathcal{N}$ un modèle quelconque de T d'univers N. Nous effectuerons un "va" pour plonger $\mathcal{M}$ dans $\mathcal{N}$ élémentairement, et de cela découlera que $\mathcal{M}$ est premier. Soit $M = \{m_i \mid i < \omega\}$ une énumération de M. Alors $\mathrm{tp}_{\mathcal{M}}(m_0)$ est isolé par une formule σ_0. Un élément quelconque de N qui satisfait σ_0 dans $\mathcal{N}$ sera fixé comme l'image de m_0. Un tel élément existe puisque $\mathcal{M} \models \exists x \sigma_0(x)$ et que $\mathcal{M} \equiv \mathcal{N}$. Appelons-le n_0. Comme $\mathrm{tp}_{\mathcal{M}}(m_0)$ est isolé par σ_0 et que $\mathcal{M} \equiv \mathcal{N}$, on déduit que σ_0 isole $\mathrm{tp}_{\mathcal{N}}(n_0)$. En fait, $\mathrm{tp}_{\mathcal{M}}(m_0) = \mathrm{tp}_{\mathcal{N}}(n_0)$.

Supposons maintenant qu'une injection transformant $m_0, \ldots, m_t$ en $n_0, \ldots, n_t$ respectivement soit formée de façon à ce que $\mathrm{tp}_{\mathcal{M}}(m_0, \ldots, m_t) = \mathrm{tp}_{\mathcal{N}}(n_0, \ldots, n_t)$. Soit $m_{t+1} \in M \setminus \{m_0, \ldots, m_t\}$. Comme $\mathcal{M}$ est atomique, $\mathrm{tp}_{\mathcal{M}}(m_0, \ldots, m_t, m_{t+1})$ est isolé par une $\mathcal{L}$-formule σ_{t+1}. Comme

$$(\mathcal{M}, m_0, \ldots, m_t) \models \exists x_{t+1} \left(\bigwedge_{i=0}^{t} x_{t+1} \neq m_i \ \wedge \ \sigma_{t+1}(x_{t+1}, m_0, \ldots, m_t) \right) ,$$

on déduit de l'égalité $\mathrm{tp}_{\mathcal{M}}(m_0, \ldots, m_t) = \mathrm{tp}_{\mathcal{N}}(n_0, \ldots, n_t)$ qu'il existe $n_{t+1} \in N$ tel que

$$(\mathcal{N}, n_0, \ldots, n_t) \models \bigwedge_{i=0}^{t} n_{t+1} \neq n_i \ \wedge \ \sigma_{t+1}(n_{t+1}, n_0, \ldots, n_t) .$$

Alors, il découle de $\mathcal{M} \equiv \mathcal{N}$ et du fait que σ_{t+1} isole $\mathrm{tp}_{\mathcal{M}}(m_0, \ldots, m_t, m_{t+1})$ que

$$\mathrm{tp}_{\mathcal{M}}(m_0, \ldots, m_t, m_{t+1}) = \mathrm{tp}_{\mathcal{N}}(n_0, \ldots, n_t, n_{t+1}) .$$

Le va épuise M en assurant à chaque étape l'égalité des types, et par conséquent l'élémentarité du plongement. Soulignons qu'il n'y a aucune raison pour que ce plongement soit surjectif. □

Deux théorèmes puissants, en l'occurrence le théorème d'omission des types et le théorème 9.2, mis en collaboration, fournissent une caractérisation concise dans le cadre des langages dénombrables :

Théorème 9.3 *Soit T une théorie complète et dénombrable. Alors T a un modèle premier si et seulement si pour tout $k \in \mathbb{N}$, les types isolés sont "denses" dans $S_k(T)$, en d'autres termes toute formule ϕ à variables libres parmi $\{x_1, \ldots, x_k\}$ appartient à un k-type isolé.*

Preuve. Si T a un modèle premier, alors ce modèle est aussi atomique d'après le théorème 9.2. Par définition, tous les types de $S(T)$ réalisés dans $\mathcal{M}$ sont isolés. Comme T est complète, ils sont denses.

La nécessité de l'existence d'un modèle premier découle du théorème d'omission des types. En effet, si les types isolés satisfont la propriété "densité" de l'énoncé, alors chaque type non isolé vérifie la condition du théorème d'omission des types. Alors, il existe un modèle de T qui omet tous ces types. D'après le théorème 9.2, c'est le modèle premier de T. □

Nous invitons nos lecteurs à comparer le théorème 9.3 à l'exercice VII de la feuille 8.

9.3 Théories ω-catégoriques

Dans cette section nous donnerons la caractérisation des théories ω-catégoriques. Sa preuve incorpore toutes les idées introduites dans ce chapitre.

Théorème 9.4 (Le théorème de Ryll-Nardzewski) *Soit $\mathcal{L}$ un langage dénombrable. Une $\mathcal{L}$-théorie T complète est ω-catégorique si et seulement si $|S_k(T)| < \aleph_0$ pour tout $k \in \mathbb{N}$.*

Preuve. Montrons d'abord la nécessité de la condition. On montre d'abord tous les types de dans $S(T)$ sont isolés. En effet, sous l'hypothèse contradictoire, si $p \in S(T)$ est un type isolé, alors d'après le théorème d'omission des types, il existe un modèle dénombrable de T qui omet p. Or, d'après le lemme 7.1.4, T a un modèle dénombrable dans lequel p est réalisé. Or T est ω-catégorique.

La conclusion du paragraphe précédent nous permet d'établir une bijection entre les types dans $S(T)$ et les $\mathcal{L}$-formule qui consiste à associer à chaque $p \in S(T)$ une formule σ_p qui l'isole. Fixons alors $k \in \mathbb{N}^*$ et montrons en utilisant la compacité qu'on peut en déduire que $|S_k(T)| < \aleph_0$. En effet, pour tout $p \in S_k(T)$ il existe σ_p tel que σ_p isole p, en particulier que $\sigma_p \in p$. En suivant l'intuition indiquée au début du chapitre, on a tendance à dire que les σ_p recouvrent $S_k(T)$. Par compacité, il existe $p_1, \ldots, p_r \in S_k(T)$ tels que $T \vdash \forall x_1 \ldots x_k \left(\bigvee_{i=1}^{r} \sigma_{p_i}(x_1, \ldots, x_k)\right)$ (il est fort conseillé de vérifier les détails du raisonnement de compacité). Ainsi $S_k(T) = \{p_1, \ldots, p_r\}$.

Maintenant supposons que $|S_k(T)| < \aleph_0$ pour tout $k \in \mathbb{N}^*$, et posons pour k fixé $S_k(T) = \{p_1, \ldots, p_t\}$ les p_i étant deux à deux distincts. Pour chaque $i \in \{1, \ldots, t\}$, il existe alors une formule $\sigma_{i,j}$ $(j \neq i)$ telle que $\sigma_{i,j} \in p_i$ et $\neg\sigma_{i,j} \notin p_j$. La formule $\bigwedge_{j \neq i} \sigma_{i,j}$ isole p_i. Ceci montre que tout modèle de T est atomique. D'après le corollaire 9.2.6, il en existe un seul dénombrabre à isomorphisme près. □

Avant d'élaborer les caractérisations de la catégoricité dénombrable, donnons un corollaire immédiat du théorème 9.4 :

Corollaire 9.3.1 *Soient $\mathcal{L}$ un langage dénombrable et T une $\mathcal{L}$-théorie T complète est ω-catégorique. Alors tous les modèles de T sont ω-saturés.*

Preuve. Il est possible de déduire ce corollaire directement. Mais nous évoquerons le théorème 7.5 pour mieux illustrer le paysage général dans lequel nous sommes situés et l'importance de la notion de type dans l'étude de diverses structures. Le langage $\mathcal{L}$ est dénombrable, et $S(T)$ est dénombrable d'après le théorème 9.4. Alors T a un modèle ω-saturé et dénombrable d'après le théorème 7.5. Or T n'a qu'un seul modèle dénombrable à isomorphisme près, ainsi toute copie isomorphe de cette structure est ω-saturée aussi.

Maintenant soient $\mathcal{M}$ un modèle quelconque de T et $\{m_1, \ldots, m_n\}$ une partie finie extraite de l'univers de $\mathcal{M}$. D'après le théorème de Löwenheim-Skolem, cette partie finie est contenue dans une sous-structure élémentaire $\mathcal{M}_0$ dénombrable de $\mathcal{M}$. Dans le paragraphe précédent nous avons vu que cette sous-structure est ω-saturée. Comme $\mathcal{M}_0 \preceq \mathcal{M}$, la réalisation dans $\mathcal{M}_0$ d'un type à paramètres dans $\{m_1, \ldots, m_n\}$ en est aussi une réalisation dans $\mathcal{M}$. Ce raisonnement étant valable pour toute partie finie extraite de l'univers de $\mathcal{M}$, on conclut que $\mathcal{M}$ est ω-saturé. □

Il est possible d'augmenter le nombre de caractérisations de la catégoricité dénombrable. C'est ce que nous faisons dans la version élargie du théorème 9.4 que nous donnons ci-dessous. La preuve, qui s'obtient en appliquant les mêmes idées que celles utilisées dans la preuve du théorème 9.4, et le corollaire 9.2.7, est un bon exercice pour améliorer la compréhension de ce chapitre.

Chacune des caractérisations dans le théorème 9.5 est utile et importante. Par exemple, le point 2 fournit un lien important avec la théorie des groupes de permutation. C'est une caractérisation complètement algébrique d'une notion de la théorie des modèles qui a suscité une forte interaction entre les théoriciens de groupes, les spécialistes des géométries finies et les théoriciens des modèles.

Théorème 9.5 (Le théorème de Ryll-Nardzewski élargi) *Soient $\mathcal{L}$ un langage dénombrable et une $\mathcal{L}$-théorie T complète qui a des modèles infinis. Alors, les énoncés suivants sont équivalents :*

1. *T est ω-catégorique ;*
2. *pour tout modèle dénombrable $\mathcal{M}$ de T, le groupe d'automorphismes de $\mathcal{M}$ est* oligomorphique*, en d'autres termes, pour tout $n \in \mathbb{N}$, $\mathrm{Aut}(\mathcal{M})$ a un nombre fini d'orbites sur M ;*
3. *il existe un modèle dénombrable de T dont le groupe d'automorphismes est oligomorphique ;*
4. *pour tout $k \in \mathbb{N}$, il existe un nombre fini de $\mathcal{L}$-formules à k variables libres à équivalence par rapport à T près ;*
5. *pour tout $k \in \mathbb{N}$, tout k-type est isolé ;*
6. *pour tout modèle dénombrable $\mathcal{M}$ d'univers M, pour toute partie* finie *A de M, et pour tout $k \in \mathbb{N}$, $|S_k^{\mathcal{M}}(A)| < \aleph_0$.*

Nous finirons ce chapitre avec un résultat bien modeste sur les groupes dont la théorie du premier ordre est ω-catégorique. Par abus de langage, ils sont souvent dits "groupes ω-catégoriques". Le résultat qui restreint considérablement leurs structures algébriques est un corollaire général du théorème 9.5.

Pour l'énoncer précisons d'abord ce que nous entendons par *une sous-structure engendrée par une partie d'une structure* : si $\mathcal{M}$ est une $\mathcal{L}$-structure d'univers M et que $A \subset M$, alors la sous-structure engendrée par A est la plus petite sous-structure de $\mathcal{M}$ contenant A. On la note $\langle A \rangle$, ou encore $\langle A \rangle_{\mathcal{M}}$ si l'on souhaite préciser la structure ambiante. En raison de la définition de la notion de sous-structure $\langle A \rangle$ contient toutes les constantes de $\mathcal{M}$, et les fonctions de la signature y induisent des opérations internes. Par conséquent, si $\mathcal{L}$ ne contient pas de symboles de fonctions (y inclus les symboles de constantes), alors $\langle A \rangle_{\mathcal{M}} = A$. Plus généralement $b \in \langle A \rangle_{\mathcal{M}}$ s'il existe un terme k-aire t ($k \in \mathbb{N}$) et $\{a_1, \ldots, a_k\} \subset A$ tels que $t(a_1, \ldots, a_k) = b$.

Corollaire 9.3.2 *Soient $\mathcal{L}$ un langage dénombrable, T une théorie ω-catégorique et $\mathcal{M}$ un modèle de T d'univers M. Alors il existe une fonction $f : \mathbb{N} \to \mathbb{N}$, et une seule, qui ne dépend que de T, telle que pour tout $n \in \mathbb{N}$, $f(n)$ soit la borne supérieure des cardinaux des sous-structures engendrées par n éléments. En particulier, $\mathcal{M}$ est une structure* uniformément localement finie*.*

Preuve. Avant d'aborder la preuve proprement dite, il semble nécessaire de vérifier la cohérence de notre cadre de travail. Nous voulons déduire une conclusion de finitude locale. Or, la définition d'une sous-structure nécessite que toutes les constantes appartiennent à l'univers de la sous-structure. Or, ces constantes, sont-elles en nombre fini ? La réponse est oui. En effet, si $\{c_i | i \in \mathbb{N}\}$ sont des constantes de la structure $\mathcal{M}$ dont la théorie est ω-catégorique, alors d'après le théorème 9.5 (2), un nombre cofini d'entre eux doit être dans la même orbite du groupe d'automorphismes de $\mathcal{M}$. Or les automorphismes de $\mathcal{M}$ fixent les constantes.

Soit $A = \{a_1, \ldots, a_n\}$. En appliquant le théorème de Löwenheim-Skolem à A et $\mathcal{M}$, nous pouvons supposer que $\mathcal{M}$ est dénombrable. Si $b \in \langle A \rangle_{\mathcal{M}}$, alors tout automorphisme de $\mathcal{M}$ qui fixe A fixe b aussi. Or, d'après le théorème 9.5 (6) il existe un nombre fini de 1-types sur A. Or, si b et c dans $\langle A \rangle_{\mathcal{M}}$ sont distincts, alors il découle du théorème 9.5 (2) que $\mathrm{tp}_{\mathcal{M}}(b/A) \neq \mathrm{tp}_{\mathcal{M}}(c/A)$. Ainsi, on déduit que $\langle A \rangle_{\mathcal{M}}$ est fini.

Le paragraphe précédent règle le problème de finitude locale puisque toute partie finie de M est contenue dans une sous-structure finie. A ceci, la finitude des types ajoute une uniformité. En effet, pour tout $n \in \mathbb{N}$, $|S_n(T)| < \aleph_0$ d'après le théorème 9.5. En raison du point (2) du théorème 9.5, sous l'action du groupe d'automorphismes de $\mathcal{M}$, M^n ne contient qu'un nombre fini d'orbites. En d'autres termes, il existe $k \in \mathbb{N}^*$ et un ensemble $\{(a_{i,1}, \ldots, a_{i,n}) \in M^k \mid 1 \leq i \leq k\}$ tels que les éléments de cet ensemble représentent les orbites complètement. Alors on pose

$$f(n) = \max\{|\langle a_{i,1}, \ldots, a_{i,n} \rangle_{\mathcal{M}}| \mid 1 \leq i \leq n\} .$$

En particulier, aucune sous-structure de $\mathcal{M}$ engendrée par n éléments ne peut contenir plus de $f(n)$ éléments.

Finalement, vérifions que la fonction f ne dépend que de T. En fait il n'y a rien à faire puisque c'est une conséquence du corollaire 9.3.1. Tout type de $S(T)$ est réalisé dans un modèle dénombrable de T, et à isomorphisme près il n'en existe qu'un seul. En conséquence, deux modèles de T réalisent les mêmes types de $S(T)$, en d'autres termes ils les réalisent tous. Ainsi, f ne dépend pas du choix de modèle de T. □

Corollaire 9.3.3 *Soit $\mathcal{L} = \{., \ ^{-1}, 1\}$ le langage des groupes. Un groupe G dont la $\mathcal{L}$-théorie $Th(G)$ est ω-catégorique est d'exposant borné. S'il est en plus abélien, il est la somme directe des groupes cycliques.*

Preuve. La première conclusion découle du corollaire 9.3.2. La deuxième nécessite un peu la théorie des groupes abéliens. □

Ce corollaire montre qu'un exemple des travaux dirigés, à savoir l'exercice 11 de la feuille 7, n'est pas une coïncidence mais un prototype. Quant aux groupes non abéliens et ω-catégoriques, leur étude est bien plus compliquée et pleine de problèmes ouverts à la frontière de la théorie des modèles et de la théorie des groupes.