LOGIQUE DU PREMIER ORDRE

Antoine Chambert-Loir

Antoine Chambert-Loir
Université Paris-Diderot.
E-mail : Antoine.Chambert-Loir@math.univ-paris-diderot.fr

Version du 5 décembre 2016, 1h21
La version la plus récente de ce texte devrait être accessible en ligne, à l'adresse
http://webusers.imj-prg.fr/~antoine.chambert-loir/enseignement/2016-17/logique/logique.pdf

TABLE DES MATIÈRES

CHAPITRE 1

ENSEMBLES ET ARITHMÉTIQUE DES CARDINAUX — *COMPTER*

« Oui ; cela n'est pas mal ; mais ce compte-là n'est rien de réel. »
— Molière, *L'Avare*, acte II, scène 5

La logique mathématique est la branche des mathématiques dont l'objet d'étude est l'ensemble des théories mathématiques elles-mêmes. Cette mise en abîme en inquiète certains, en indiffère beaucoup, et attise la curiosité d'autres. Si ce cours s'adresse par nature à ces derniers, j'espère qu'il rassurera les premiers et suscitera quelque intérêt chez les seconds.

Comme l'expliquent Cori et Lascar dans l'introduction de leur ouvrage (2003*a* ; 2003*b*), la logique mathématique utilise tout l'attirail des mathématiques : raisonnement, structures, etc. Ce premier chapitre développe quelques aspects de théorie « naïve » des ensembles que nous utiliserons par la suite.

1.1. Rappels de théorie des ensembles

1.1.1. — La notion naïve d'ensemble semble claire. Un ensemble possède des éléments, qui peuvent d'ailleurs eux-mêmes être des ensembles ; on écrit $x \in \mathrm{A}$ pour dire qu'un élément x appartient à l'ensemble A, et $x \notin \mathrm{A}$ pour dire qu'il n'y appartient pas. La théorie des ensembles est guidée par deux principes :

a) Deux ensembles qui ont les mêmes éléments sont égaux (*extension*) ;
b) Si $\boldsymbol{P}(x)$ est une propriété d'un objet, on peut former l'ensemble $\{x\,;\,\mathbf{P}(x)\}$ des objets qui la vérifient (*compréhension*).

Comme l'a observé G. Russell, les choses ne peuvent pas être aussi simples.

Théorème **(1.1.2)** (Paradoxe de Russell). — *Il n'existe pas d'ensemble dont les éléments soient les ensembles vérifiant la propriété $x \notin x$.*

Je renvoie à l'excellent Doxiadis & Papadimitriou (2009) (pages 162 et suivantes) pour un récit savoureux de la découverte de ce paradoxe.

Démonstration. — Raisonnons par l'absurde en supposant qu'un tel ensemble existe et notons-le R. La contradiction apparaît lorsqu'on cherche à savoir si R appartient à R ou pas.

Si R appartient à R, on a $R \in R$, donc $R \notin R$ par définition de l'ensemble R ; c'est une contradiction. Donc R n'appartient pas à R, c'est-à-dire que $R \notin R$, donc $R \in R$ par définition de R ; apparaît de nouveau une contradiction. □

1.1.3. — Ce paradoxe indique qu'une formalisation précise de la théorie des ensembles est nécessaire. Parmi les formalisations possibles, l'axiomatique (ZF) de Zermelo–Fraenkel, complémentée (ZFC) par l'axiome du choix, est la plus couramment utilisée. Dans ce cours, nous nous contenterons d'une compréhension intuitive, telle qu'elle est développée en licence, et renvoyons au cours du second semestre pour sa description détaillée. Nous utiliserons librement l'axiome du choix ; il sera parfois invoqué sous la forme du théorème de Zorn (théorème 1.1.8), mais parfois aussi dans des raisonnements d'apparence anodine dont la formalisation le nécessite.

1.1.4. — Soit A et B des ensembles. On dit que B est contenu dans A, ou que B est une partie de A, et on note $B \subseteq A$, si tout élément de B appartient à A.

L'ensemble vide, l'ensemble A lui-même sont des parties de A.

Si A est un ensemble, on note $\mathscr{P}(A)$ l'ensemble des parties de A.

1.1.5. — Soit $f : A \to B$ une application. On dit que f est injective (resp. surjective, resp. bijective) si tout élément de B a au plus (resp. au moins, resp. exactement un) antécédent par f.

Si f est bijective, on note $f^{-1} : B \to A$ sa bijection réciproque.

Si f est injective et A n'est pas vide, il existe une application $g : B \to A$ telle que $g \circ f = \mathrm{Id}_A$ (rétraction). Choisissons un élément $a \in A$ et définissons g comme suit : si $b \in B$ appartient à l'image de f, on note $g(b)$ son unique antécédent ; sinon, on pose $g(b) = a$. Par construction, on a $g(f(x)) = x$ pour tout $x \in A$.

Si f est surjective, il existe une application $g : B \to A$ telle que $f \circ g = \mathrm{Id}_B$ (section). Cela se déduit de l'« axiome du choix » en choisissant, pour tout $b \in B$, un antécédent $g(b)$ de b par f. Le lemme suivant affirme que l'application g est injective.

Lemme **(1.1.6).** — *Soit $f : A \to B$ et $g : B \to C$ des applications.*

a) *Si f et g sont injectives, alors $g \circ f$ est injective ;*

b) *Si f et g sont surjectives, alors $g \circ f$ est surjective ;*
c) *Si $g \circ f$ est injective, alors f est injective ;*
d) *Si $g \circ f$ est surjective, alors g est surjective.*
e) *En particulier, si f et g sont bijectives, alors $g \circ f$ est bijective.*

1.1.7. — On suppose connu l'ensemble $\mathbf{N} = \{0, 1, 2, \ldots\}$ des entiers naturels. On note $s : \mathbf{N} \to \mathbf{N}$ l'application « successeur », définie par $s(n) = n+1$. Le couple $(\mathbf{N}, s)$ sera caractérisé par les propriétés suivantes :

a) L'élément 0 n'est le successeur d'aucun élément ;
b) L'application s est injective ;
c) Principe de récurrence : Soit A une partie de $\mathbf{N}$; si $0 \in A$ (initialisation) et si $s(A) \subset A$ (hérédité), alors $A = \mathbf{N}$.

Théorème **(1.1.8)** (Zorn). — *Soit* A *un ensemble muni d'une relation d'ordre $\leq$. On suppose que* A *est* inductif, *c'est-à-dire que toute partie totalement ordonnée de* A *est majorée. Alors,* A *possède un élément maximal.*

Plus généralement, pour tout élément a de A, *il existe un élément maximal b tel que $a \leq b$.*

1.2. Ensembles finis

Lemme **(1.2.1)**. — *Soit m et n des entiers naturels.*

a) *Pour qu'il existe une surjection $f : \{0, 1, \ldots, m-1\} \to \{0, 1, \ldots, n-1\}$, il faut et il suffit que $m \geqslant n$.*
b) *Pour qu'il existe une injection $f : \{0, 1, \ldots, m-1\} \to \{0, 1, \ldots, n-1\}$, il faut et il suffit que $m \leqslant n$.*
c) *Pour qu'il existe une bijection $f : \{0, 1, \ldots, m-1\} \to \{0, 1, \ldots, n-1\}$, il faut et il suffit que $m = n$.*

Démonstration. — a) Si $m \geqslant n$, l'application f de $\{0, 1, \ldots, m-1\}$ dans $\{0, 1, \ldots, n-1\}$ donnée par $f(i) = \min(i, n-1)$ est surjective, ce qui établit la suffisance de la condition indiquée. Démontrons sa nécessité par récurrence sur n. Il n'y a rien a démontrer si $n = 0$, car l'ensemble $\{0, \ldots, n-1\}$ est alors vide. Supposons l'assertion vérifiée pour $n-1$ et soit $f : \{0, \ldots, m-1\} \to \{0, \ldots, n\}$ une application surjective. Nécessairement, on a $m \geqslant 1$. Posons $b = \min(f(m-1), n-1)$ et soit $g : \{0, \ldots, m-2\} \to \{0, \ldots, n-1\}$ l'application donnée par $g(a) = f(a)$ si $f(a) \leqslant n-1$ et $g(a) = b$ sinon ; démontrons qu'elle est surjective.

Soit $i \in \{0, \dots, n-1\}$ et soit a un antécédent de i par f. Si $i \neq f(m-1)$, on a $a \neq m-1$, donc $g(a) = f(a) = i$. Sinon, on a $i = f(m-1) = b$; soit a un antécédent de n par f ; comme $f(a) = n \neq i$, on a $a \neq m-1$, d'où $g(a) = i$. Ainsi, g est surjective, donc $m - 2 \geqslant n - 1$, puis $m - 1 \geqslant n$, comme il fallait démontrer.

b) Si $m \leqslant n$, l'application f de $\{0, 1, \dots, m-1\}$ dans $\{0, 1, \dots, n-1\}$ donnée par $f(i) = i$ est injective, ce qui prouve la suffisance de la condition indiquée. Démontrons sa nécessité : soit $f : \{0, \dots, m-1\} \to \{0, \dots, n-1\}$ une application injective ; démontrons que $m \leqslant n$. L'assertion étant évidente si $m = 0$, on suppose donc $m \geqslant 1$; alors $0 \leqslant f(0) \leqslant n - 1$, donc $n \geqslant 1$. On définit alors une application g de $\{0, \dots, n-1\}$ dans $\{0, \dots, m-1\}$ de la façon suivante. Soit $a \in \{0, \dots, n-1\}$; si a possède un antécédent par f, on note $g(a)$ cet antécédent ; sinon, on pose $g(a) = 0$. L'application g ainsi définie vérifie $g(f(a)) = a$ pour tout $a \in \{0, \dots, m-1\}$, donc g est surjective. Par suite, $m \leqslant n$.

c) Si $m = n$, l'application identique est une bijection de $\{0, \dots, m-1\}$ dans lui-même, d'où la suffisance de la condition donnée. Sa nécessité résulte des deux cas précédents.

□

Définition **(1.2.2)**. — *On dit qu'un ensemble* A *est* fini *s'il existe un entier naturel n et une bijection* $f : \{0, 1, \dots, n-1\} \to \mathrm{A}$. *Il existe alors un seul tel entier ; on dit que c'est le* cardinal *de l'ensemble* A *et on le note* $\mathrm{Card}(\mathrm{A})$, *ou* $|\mathrm{A}|$.

L'ensemble vide est fini, de cardinal 0.

Lemme **(1.2.3)**. — *Soit n un entier et soit* A *une partie de* $\{0, \dots, n-1\}$. *Il existe une bijection* $f : \{0, \dots, n-1\} \to \{0, \dots, n-1\}$ *et un entier* $m \in \{0, \dots, n\}$ *tels que* $f(\{m, \dots, n-1\}) = \mathrm{A}$.

Démonstration. — Si $\mathrm{A} = \{0, \dots, n-1\}$, on pose $f = \mathrm{Id}$ et $m = 0$. Sinon, on raisonne par récurrence descendante sur le plus grand élément p de $\{0, 1, \dots, n-1\}$ qui n'appartient pas à A ; de sorte que $\{p+1, \dots, n-1\} \subset \mathrm{A}$ et $p \notin \mathrm{A}$. Si $\mathrm{A} \cap \{1, \dots, p\} = \varnothing$, on pose $f = \mathrm{Id}$ et $m = p + 1$. Sinon, soit a un élément de $\mathrm{A} \cap \{1, \dots, p\}$ et soit σ la transposition (a, p). Alors, $\sigma(\mathrm{A})$ est une partie de $\{0, \dots, n-1\}$ qui contient $\{p, \dots, n-1\}$ mais n'est pas égale à $\{0, \dots, n-1\}$. Par récurrence, il existe une bijection g de $\{0, \dots, n-1\}$ dans lui-même et un entier $m \in \{0, \dots, n\}$ tels que $g(\{m, \dots, n-1\}) = \sigma(\mathrm{A})$. Alors, $f = \sigma \circ g$ est une bijection de $\{0, \dots, n-1\}$ dans lui-même et l'on a $g(\{m, \dots, n-1\}) = \mathrm{A}$. □

Proposition **(1.2.4)**. — *Soit* A *un ensemble fini. Pour toute partie* B *de* A*, l'ensemble* B *est fini et l'on a* $\mathrm{Card}(B) \leqslant \mathrm{Card}(A)$. *Plus précisément,* $A - B$ *est également fini et l'on a* $\mathrm{Card}(A) = \mathrm{Card}(B) + \mathrm{Card}(A - B)$.

Démonstration. — Soit $n = \mathrm{Card}(A)$ et soit $f : \{0, \ldots, n-1\} \to A$ une bijection. Soit $g : \{0, \ldots, n-1\} \to \{0, \ldots, n-1\}$ une bijection et soit m un entier tels que $g(\{m, \ldots, n-1\}) = f^{-1}(B)$. Alors, $f^{-1}(A - B) = \{0, \ldots, m-1\}$; par suite, la restriction de $f \circ g$ définit une bijection de $\{0, \ldots, m-1\}$ sur $A - B$; cela prouve que $A - B$ est fini et que $\mathrm{Card}(A - B) = m$. De plus, l'application $\{0, \ldots, n-m\} \to A$ donnée par $p \mapsto f(g(m+p))$ définit une bijection de $\{0, \ldots, n-m-1\}$ sur B ; cela prouve que B est fini et que $\mathrm{Card}(B) = n-m$. Ainsi, $\mathrm{Card}(A) = (n-m)+m = \mathrm{Card}(B) + \mathrm{Card}(A - B)$. □

Corollaire **(1.2.5)**. — *Soit* A *et* B *des ensembles finis tels que* $\mathrm{Card}(A) = \mathrm{Card}(B)$ *et soit* $f : A \to B$ *une application. Les conditions suivantes sont équivalentes :*

a) *f est injective ;*
b) *f est surjective ;*
c) *f est bijective.*

Démonstration. — Supposons que f est injective. Elle induit une bijection de A sur son image ; on a alors $\mathrm{Card}(f(A)) = \mathrm{Card}(A) = \mathrm{Card}(B)$, donc $\mathrm{Card}(B - f(A)) = 0$ et $B - f(A) = \varnothing$, c'est-à-dire $f(A) = B$. Par suite, f est surjective.

Supposons que f est surjective et soit $g : B \to A$ une section de f. Comme g est injective, le cas déjà traité entraîne que g est surjective, donc bijective. Puisque $f \circ g = \mathrm{Id}_B$, l'application $f = g^{-1}$ est bijective.

Enfin, si f est bijective, elle est également injective. □

Corollaire **(1.2.6)**. — *L'ensemble* **N** *n'est pas fini.*

Démonstration. — Si **N** était fini, on aurait $n \leqslant \mathrm{Card}(\mathbf{N})$ pour tout entier n. □

Corollaire **(1.2.7)**. — *Soit* A *un ensemble fini et soit* $(A_1, \ldots, A_p)$ *une partition finie de* A*. Pour tout entier* $i \in \{1, \ldots, p\}$*, l'ensemble* A_i *est fini et l'on a* $\mathrm{Card}(A) = \sum_{i=1}^{p} \mathrm{Card}(A_i)$.

Démonstration. — Cela se déduit de la proposition par récurrence sur p. □

Corollaire **(1.2.8)**. — *Soit* $f : A \to B$ *une application. On suppose que* B *est fini et que pour tout* $b \in B$*, l'ensemble* $f^{-1}(b)$ *est fini ; alors* A *est fini et l'on a* $\mathrm{Card}(A) = \sum_{b \in B} \mathrm{Card}(f^{-1}(b))$.

Démonstration. — Soit $g : \{0, \dots, p-1\} \to \mathrm{B}$ une bijection ; pour tout $i \in \{0, \dots, p-1\}$, on pose $\mathrm{A}_i = f^{-1}(g(i))$. La suite $(\mathrm{A}_0, \dots, \mathrm{A}_{p-1})$ est une partition de A. Par suite, $\mathrm{Card}(\mathrm{A}) = \sum_{i=0}^{p-1} \mathrm{Card}(\mathrm{A}_i) = \sum_{b \in \mathrm{B}} \mathrm{Card}(f^{-1}(b))$. □

Corollaire **(1.2.9)**. — *Soit* A *et* B *des ensembles finis. L'ensemble* $\mathrm{A} \times \mathrm{B}$ *est fini, de cardinal* $\mathrm{Card}(\mathrm{A})\,\mathrm{Card}(\mathrm{B})$.

Démonstration. — Considérons la seconde projection $p(a, b) \mapsto b$ de $\mathrm{A} \times \mathrm{B}$ dans B. Pour tout $b \in \mathrm{B}$, l'application $a \mapsto (a, b)$ est une bijection de A sur l'ensemble $p^{-1}(b)$. Le corollaire résulte donc du corollaire précédent, appliqué à l'application p. □

Proposition **(1.2.10)**. — *Soit* A *et* B *des ensembles finis. L'ensemble* $\mathscr{F}(\mathrm{A}; \mathrm{B})$ *des applications de* A *dans* B *est fini, de cardinal* $\mathrm{Card}(\mathrm{B})^{\mathrm{Card}(\mathrm{A})}$.

Démonstration. — Soit n un entier et soit $f : \{0, \dots, n-1\} \to \mathrm{A}$ une bijection. On démontre la proposition par récurrence sur n. Si $n = 0$, alors $\mathrm{A} = \varnothing$ et il n'y a qu'une application de A dans B (dont le graphe est vide). Alors, $\mathrm{Card}(\mathscr{F}(\mathrm{A}; \mathrm{B})) = 1 = \mathrm{Card}(\mathrm{B})^0$. Posons sinon $a = f(n-1)$ et $\mathrm{A}' = \mathrm{A} - \{a\}$. L'application $u \mapsto (u|_{\mathrm{A}'}, u(a))$ est une bijection de $\mathscr{F}(\mathrm{A}; \mathrm{B})$ sur $\mathscr{F}(\mathrm{A}'; \mathrm{B}) \times \mathrm{B}$. Par récurrence, l'ensemble $\mathscr{F}(\mathrm{A}'; \mathrm{B})$ est fini, de cardinal $\mathrm{Card}(\mathrm{B})^{\mathrm{Card}(\mathrm{A}')}$. Ainsi, l'ensemble $\mathscr{F}(\mathrm{A}; \mathrm{B})$ est fini et son cardinal est égal à

$$\begin{aligned} \mathrm{Card}(\mathscr{F}(\mathrm{A}'; \mathrm{B}))\,\mathrm{Card}(\mathrm{B}) &= \mathrm{Card}(\mathrm{B})^{\mathrm{Card}(\mathrm{A}')}\,\mathrm{Card}(\mathrm{B}) \\ &= \mathrm{Card}(\mathrm{B})^{\mathrm{Card}(\mathrm{A}')+1} \\ &= \mathrm{Card}(\mathrm{B})^{\mathrm{Card}(\mathrm{A})}. \end{aligned}$$

□

Corollaire **(1.2.11)**. — *Pour tout ensemble fini* A*, l'ensemble* $\mathscr{P}(\mathrm{A})$ *des parties de* A *est fini, de cardinal* $2^{\mathrm{Card}(\mathrm{A})}$.

Démonstration. — Pour toute partie B de A, notons χ_{B} sa fonction indicatrice, définie par $\chi_{\mathrm{B}}(a) = 1$ si $a \in \mathrm{B}$ et $\chi_{\mathrm{B}}(a) = 0$ sinon. L'application $\mathrm{B} \mapsto \chi_{\mathrm{B}}$ est une bijection de $\mathscr{P}(\mathrm{A})$ dans $\mathscr{F}(\mathrm{A}; \{0, 1\})$. D'après la proposition précédente, l'ensemble $\mathscr{F}(\mathrm{A}; \{0, 1\})$ est fini, de cardinal $2^{\mathrm{Card}(\mathrm{A})}$. Le corollaire en résulte. □

1.3. Ensembles dénombrables

Proposition **(1.3.1)**. — *Soit* A *un ensemble. Si* A *n'est pas fini, il existe une application injective* $f : \mathbf{N} \to \mathrm{A}$.

Démonstration. — Puisque A n'est pas fini, il n'est pas vide ; soit donc a_0 un élément de A et posons $A_1 = A - \{a_0\}$. Si l'ensemble A_1 était fini, l'ensemble A serait fini, comme réunion de sa partie finie A_0 et de son complémentaire $\{a_0\}$; en particulier, A_1 n'est pas vide ; soit a_1 un élément de A_1 et posons $A_2 = A_1 - \{a_1\}$. En répétant cet argument, on construit une suite $(a_0, a_1, \dots)$ d'éléments de A, deux à deux distincts. L'application $n \mapsto a_n$ est une application injective de **N** dans A.[1] □

Définition **(1.3.2)**. — *Soit* A *un ensemble. On dit que* A *est dénombrable s'il existe une application injective de* A *dans* **N**.

Un ensemble fini est dénombrable.
Une partie d'un ensemble dénombrable est dénombrable.

Lemme **(1.3.3)**. — *Soit* $f : A \to B$ *une application surjective. Si* A *est dénombrable, alors* B *est dénombrable.*

Démonstration. — Soit $g : B \to A$ une section de l'application f. C'est une bijection de B sur une partie de A. Par suite, B est dénombrable. □

Proposition **(1.3.4)**. — *Soit* A *une partie de* **N**. *Si* A *n'est pas finie, il existe une unique bijection croissante* $f : \mathbf{N} \to A$.

Démonstration. — Pour tout entier $m \in \mathbf{N}$ et toute application $f : \{0, \dots, m-1\} \to A$, l'ensemble $A - \{f(0), \dots, f(m-1)\}$ n'est pas fini, donc n'est pas vide ; il possède donc un plus petit élément. On définit alors une application $f : \mathbf{N} \to A$ par récurrence de la façon suivante : pour tout entier m, $f(m)$ est le plus petit élément de $A - \{f(0), \dots, f(m-1)\}$. L'application f ainsi définie est strictement croissante. Soit en effet un élément $m \in \mathbf{N}$ tel que $m \geqslant 1$; par hypothèse, $f(m-1)$ est le plus petit élément de $A - \{f(0), \dots, f(m-2)\}$ et $f(m)$ est le plus petit élément de $A - \{f(0), \dots, f(m-1)\}$. Cela entraîne $f(m) \neq f(m-1)$; si l'on avait $f(m) \leqslant f(m-1)$, on aurait $f(m) < f(m-1)$ ce qui contredirait la définition de $f(m-1)$. En particulier, l'application f est injective. Démontrons qu'elle est surjective. Raisonnons par l'absurde et supposons qu'un élément $a \in A$ n'appartient pas à l'image de f. Comme f est injective, on a $f(m) \geqslant m$ pour tout entier m ; en particulier, $f(a) \geqslant a$, donc $f(a) > a$. Par suite,

[1] La démonstration précédente est informelle, en ce qu'elle requiert une succession infinie de choix ; une présentation rigoureuse mettrait en évidence l'utilisation de l'axiome du choix, par exemple par l'utilisation du théorème de Zorn.

a est un élément de $A - \{f(0), \dots, f(a-1)\}$ strictement inférieur à $f(a)$, ce qui contredit la définition de $f(a)$. Ainsi, $f : \mathbf{N} \to A$ est une bijection croissante.

Soit $g : \mathbf{N} \to A$ une bijection croissante et posons $h = g^{-1} \circ f$. L'application h est une bijection croissante de $\mathbf{N}$ dans $\mathbf{N}$; démontrons par récurrence sur n que $h(n) = n$ pour tout $n \in \mathbf{N}$. Soit a l'antécédent de 0 par h. Comme $a \geqslant 0$, on a $0 = h(a) \geqslant h(0) \geqslant 0$, d'où $h(0) = 0$. Supposons que $h(m) = m$ pour tout entier m tel que $m < n$ et démontrons que $h(n) = n$; soit a l'antécédent de n par h. Comme $h(m) < n$ pour $m < n$, on a $a \geqslant n$. Par suite, $n = h(a) \geqslant h(n) > h(n-1) = n - 1$, donc $n = h(n)$. On a donc $h = \mathrm{Id}$, d'où $f = g$. □

Corollaire **(1.3.5)**. — *Soit* A *un ensemble dénombrable. Si* A *n'est pas fini, il existe une bijection de* $\mathbf{N}$ *dans* A.

Démonstration. — Soit $f : A \to \mathbf{N}$ une application injective et soit $f' : f(A) \to A$ l'application qui associe à un élément de $f(A)$ son unique antécédent par f. Alors $f(A)$ est une partie de $\mathbf{N}$ qui n'est pas finie. Il existe donc une bijection $g : \mathbf{N} \to f(A)$ et l'application $f' \circ g$ est une bijection de $\mathbf{N}$ sur A. □

Proposition **(1.3.6)**. — *L'ensemble* $\mathbf{N} \times \mathbf{N}$ *est dénombrable.*

Démonstration. — Voici deux preuves.

a) Soit $f : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ l'application définie par $f(m,n) = 2^m 3^n$; elle est injective, en vertu de l'unicité de la décomposition en facteurs premiers. Cela démontre que $\mathbf{N} \times \mathbf{N}$ est dénombrable.

b) On construit une application de $\mathbf{N} \times \mathbf{N}$ dans $\mathbf{N}$ en parcourant le quadrant $\mathbf{N}^2$ diagonale descendante par diagonale descendante, par la formule suivante :

$$f(m,n) = \sum_{p=0}^{m+n-1} (p+1) + m = \frac{1}{2}(m+n)(m+n+1) + m.$$

On vérifie que f est bijective, d'où la proposition.

□

Corollaire **(1.3.7)**. — *Soit* $f : A \to B$ *une application. On suppose que l'ensemble* B *est dénombrable et que, pour tout élément* $b \in B$, *l'ensemble* $f^{-1}(b)$ *est dénombrable. Alors l'ensemble* A *est dénombrable.*

Démonstration. — Soit $g : B \to \mathbf{N}$ une application injective. Pour tout $b \in B$, soit $h_b : f^{-1}(b) \to \mathbf{N}$ une application injective. Soit alors $h : A \to \mathbf{N} \times \mathbf{N}$ l'application définie par $h(a) = (g(f(a)), h_{f(a)}(a))$. Démontrons que h est injective. Soit

$a, a' \in \mathrm{A}$ tels que $h(a) = h(a')$ et démontrons que $a = a'$. On a d'abord $g(f(a)) = g(f(a'))$; comme g est injective, cela entraîne $f(a) = f(a')$. Comme $h_{f(a)}$ est injective, on a donc $a = a'$. Par suite, h est injective, donc A est dénombrable. □

Corollaire **(1.3.8)**. — *La réunion* $\bigcup_{n\in\mathbf{N}} \mathrm{A}_n$ *d'une suite* $(\mathrm{A}_n)_{n\in\mathbf{N}}$ *d'ensembles dénombrables est dénombrable.*

Démonstration. — Soit $\mathrm{A} = \bigcup_{n\in\mathbf{N}} \mathrm{A}_n$ la réunion de la suite (A_n) et soit A' sa somme ; rappelons que A' est la partie de $\mathbf{N}\times\mathrm{A}$ formée des couples (n, a) tels que $a \in \mathrm{A}_n$. Soit $f : \mathrm{A}' \to \mathbf{N}$ et $g : \mathrm{A}' \to \mathrm{A}$ les deux projections. D'après le corollaire précédent appliqué à l'application f, l'ensemble A' est dénombrable. Comme g est surjective, l'ensemble A est dénombrable. □

Corollaire **(1.3.9)**. — *Soit* n *un entier et soit* $\mathrm{A}_1, \ldots, \mathrm{A}_n$ *des ensembles dénombrables ; l'ensemble* $\mathrm{A}_1 \times \cdots \times \mathrm{A}_n$ *est dénombrable.*

Démonstration. — Le résultat est évident si $n = 1$. Démontrons le cas général par récurrence sur n. Posons $\mathrm{B} = \mathrm{A}_1 \times \cdots \times \mathrm{A}_{n-1}$ et soit $f : \mathrm{A}_1 \times \cdots \times \mathrm{A}_n \to \mathrm{B}$ l'application définie par $f(a_1, \ldots, a_n) = (a_1, \ldots, a_{n-1})$. Par hypothèse de récurrence, B est dénombrable ; de plus, pour tout $b = (a_1, \ldots, a_{n-1})$, l'application $a \mapsto (a_1, \ldots, a_{n-1}, a)$ induit une bijection de A_n sur $f^{-1}(b)$, si bien que $f^{-1}(b)$ est dénombrable. Par suite, $\mathrm{A}_1 \times \cdots \times \mathrm{A}_n$ est dénombrable. □

1.4. Cardinaux infinis

Définition **(1.4.1)**. — *On dit que deux ensembles* A *et* B *sont équipotents s'il existe une bijection* $f : \mathrm{A} \to \mathrm{B}$.

C'est une relation d'équivalence. La théorie des ensembles (ZFC) permet de construire, pour tout ensemble A, un ensemble « canonique » $\mathrm{Card}(\mathrm{A})$, appelé *cardinal de* A, qui lui est équipotent et qui vérifie la propriété : pour que deux ensembles A et B soient équipotents, il faut et il suffit que $\mathrm{Card}(\mathrm{A}) = \mathrm{Card}(\mathrm{B})$. D'autres auteurs préfèrerent supposer choisi un ensemble dans cette classe, n'ayant d'autre propriété que l'assertion précédente ; c'est par exemple le cas de Bourbaki (1970) où $\mathrm{Card}(\mathrm{A})$ est défini à l'aide du méta-axiome du choix que permet le symbole « τ » de Hilbert. Dans tous les cas, les nombres entiers sont définis de sorte que si A est un ensemble fini de cardinal n, alors $\mathrm{Card}(\mathrm{A}) = n$.

On appelle *cardinal* un ensemble de la forme $\mathrm{Card}(\mathrm{A})$. Les cardinaux donnent lieu à une arithmétique intéressante mais délicate, et parfois surprenante. Dans ce dernier paragraphe, nous nous limiterons à ses aspects les plus élémentaires.

1.4.2. — Si A et B sont des ensembles, notons $\mathrm{A} \preceq \mathrm{B}$ s'il existe une injection de A dans B, et $\mathrm{A} \equiv \mathrm{B}$ s'il existe une bijection de A sur B.

S'il existe une application surjective f de A dans B, toute section g de f est une application injective de B dans B, de sorte que $\mathrm{B} \preceq \mathrm{A}$.

Théorème **(1.4.3)**. — *Soit* A *et* B *des ensembles.*

a) *Ou bien* $\mathrm{A} \preceq \mathrm{B}$*, ou bien* $\mathrm{B} \preceq \mathrm{A}$ (Théorème de Cantor)*;*
b) *Si* $\mathrm{A} \preceq \mathrm{B}$ *et* $\mathrm{B} \preceq \mathrm{A}$*, alors* $\mathrm{A} \equiv \mathrm{B}$ (Théorème de Cantor-Bernstein).

Démonstration. — a) Soit I l'ensemble des couples (V, f), où V est une partie de A et f une injection de V dans B. On munit I de la relation suivante : $(\mathrm{V}, f) \leqslant (\mathrm{W}, g)$ si $\mathrm{V} \subset \mathrm{W}$ et si $f = g|_{\mathrm{V}}$. Démontrons que c'est une relation d'ordre :

– On a $(\mathrm{V}, f) \leqslant (\mathrm{V}, f)$ pour tout $(\mathrm{V}, f) \in \mathscr{J}$.
– Supposons $(\mathrm{V}, f) \leqslant (\mathrm{W}, g)$ et $(\mathrm{W}, g) \subset (\mathrm{V}, f)$; alors $\mathrm{V} \subset \mathrm{W} \subset \mathrm{V}$, donc $\mathrm{V} = \mathrm{W}$, et $g = g|_{\mathrm{V}} = f$.
– Supposons $(\mathrm{V}, f) \leqslant (\mathrm{W}, g)$ et $(\mathrm{W}, g) \subset (\mathrm{X}, h)$; alors $\mathrm{V} \subset \mathrm{W} \subset \mathrm{X}$ et $f = g|_{\mathrm{V}} = h|_{\mathrm{W}}|_{\mathrm{V}} = h|_{\mathrm{V}}$, de sorte que $(\mathrm{V}, f) \leqslant (\mathrm{X}, h)$.

Pour $i \in \mathrm{I}$, notons (V_i, f_i) l'élement i.

Soit J une partie totalement ordonnée de I. Soit V la réunion des V_j, pour $j \in \mathrm{J}$. Soit $a \in \mathrm{V}$. Soit $j, k \in \mathrm{J}$ tels que $a \in \mathrm{V}_j \cap \mathrm{V}_j$; par hypothèse, on a $j \leqslant k$ ou $k \leqslant j$, ce qui entraîne $f_k(a) = f_j(a)$; notons $f(a)$ cette valeur commune.

Soit $a, b \in \mathrm{V}$ tels que $f(a) = f(b)$. Par hypothèse il existe $j, k \in \mathscr{J}$ tels que $a \in \mathrm{V}_j$ et $b \in \mathrm{V}_k$. Si $j \leqslant k$, alors $a \in \mathrm{V}_k$ et $f_k(a) = f(a) = f(b) = f_k(b)$; comme f_k est injective, cela entraîne $a = b$. On raisonne de même si $k \leqslant j$. Ainsi, l'application $f : \mathrm{V} \to \mathrm{B}$ est injective, et le couple (V, f) est un élément de I qui majore J. Cela démontre que l'ensemble ordonné I est inductif. D'après le théorème de Zorn, il possède un élément maximal (V, f).

Si $\mathrm{V} = \mathrm{A}$, alors f est une injection de A dans B, ce qui démontre que $\mathrm{A} \preceq \mathrm{B}$.

Supposons $\mathrm{V} \neq \mathrm{A}$ et démontrons que f est surjective. Dans le cas contraire, considérons un élément $a \in \mathrm{A} - \mathrm{V}$ et un élément $b \in \mathrm{B} - f(\mathrm{V})$, posons $\mathrm{V}' = \mathrm{V} \cup \{a\}$ et soit $f' : \mathrm{V}' \to \mathrm{B}$ l'application définie par $f'(x) = f(x)$ pour $x \in \mathrm{V}$ et $f'(a) = b$. Elle est injective, par choix de b, et l'on a $(\mathrm{V}, f) < (\mathrm{V}', f')$, ce

qui contredit l'hypothèse que (V, f) est un élément maximal de I. Ainsi, f est surjective. L'application $g : B \to A$ qui associe à tout élément b de B son unique antécédent par f est injective, ce qui démontre que $B \preceq A$.

b) Soit $f : A \to B$ et $g : B \to A$ des applications injectives et construisons une application bijective h de A dans B. On définit des ensembles (A_n) et (B_n) par récurrence sur n en posant $A_0 = A$, $B_0 = B$, et, pour $n \geqslant 0$, $A_{n+1} = g(B_n)$ et $B_{n+1} = f(A_n)$; on pose aussi $A_n^* = A_n - A_{n+1}$ et $B_n^* = B_n - B_{n+1}$; soit enfin $A_\infty = \bigcap_{n\in\mathbf{N}} A_n$ et $B_\infty = \bigcap_{n\in\mathbf{N}} B_n$.

Par construction, la suite $(A_0^*, A_1^*, \ldots, A_\infty)$ est une partition de A, et la suite $(B_0^*, B_1^*, \ldots, B_\infty)$ est une partition de B. Par restriction, l'application f induit une bijection de A_n^* sur B_{n+1}^*, pour tout $n \in \mathbf{N}$, et une bijection de A_∞ sur B_∞ ; de même, l'application g induit une bijection de B_n^* sur A_{n+1}^*, pour tout entier n.

On définit une application $h : A \to B$ de la façon suivante. Soit $a \in A$. Si $a \in A_\infty$, on pose $h(a) = f(a)$. Sinon, soit n l'unique entier tel que $a \in A_n^*$; si n est pair, on pose $h(a) = f(a)$; si n est impair, alors a appartient à $A_1 = g(B)$, donc il possède un unique antécédent par g que l'on note $h(a)$. L'application h est bijective. □

1.4.4. — D'après le théorème précédent, la relation $A \preceq B$ sur les ensembles induit par restriction une relation d'ordre total sur les cardinaux, notée $\preceq$ ou, plus simplement, $\leqslant$.

On note $\aleph_0$ le cardinal de l'ensemble $\mathbf{N}$; c'est le plus petit cardinal infini.[2]

1.5. Arithmétique des cardinaux

1.5.1. — Soit α, β des cardinaux. On note respectivement $\alpha + \beta$, $\alpha\beta$ et α^β les cardinaux des ensembles somme des ensembles α et β, produit des ensembles α et β, et de l'ensemble $\mathscr{F}(\beta, \alpha)$ des fonctions de β dans α.

Lemme **(1.5.2)**. — *Soit* A, A', B, B' *des ensembles tels que* $A' \preceq A$ *et* $B' \preceq B$.

On a les relations $A' + B' \preceq A + B$ *et* $A' \times B' \preceq A \times B$. *Sauf si* $A = B' = \varnothing$ *et* $B \neq \varnothing$, *on a de plus* $\mathscr{F}(B', A') \preceq \mathscr{F}(B, A)$.

[2] Le symbole $\aleph$, *aleph*, est la première lettre de l'alphabet hébreu. Elle a été introduite vers 1893 par Georg Cantor pour noter les cardinaux infinis, parce que « les autres alphabets étaient déjà trop largement utilisés. » La notation a été conservée depuis.

En particulier, $\mathrm{Card}(A + B) = \mathrm{Card}(A) + \mathrm{Card}(B)$, $\mathrm{Card}(A \times B) = \mathrm{Card}(A)\,\mathrm{Card}(B)$ *et* $\mathrm{Card}(\mathscr{F}(B, A)) = \mathrm{Card}(A^B) = \mathrm{Card}(A)^{\mathrm{Card}(B)}$.

Démonstration. — Choisissons des injections $f : A' \to A$ et $g : B' \to B$; pour traiter la deuxième assertion, on suppose que f et g sont bijectives et que $A' = \mathrm{Card}(A)$ et $B' = \mathrm{Card}(B)$.

L'unique application h de $A'+B'$ dans $A+B$ dont la restriction à A' est égale à f et la restriction à B' est égale à g est injective; elle est bijective lorsque f et g sont bijectives. Cela prouve la première inégalité, ainsi que l'égalité $\mathrm{Card}(A + B) = \mathrm{Card}(A) + \mathrm{Card}(B)$.

De même, l'application de $A' \times B'$ dans $A \times B$ définie par $(a, b) \mapsto (f(a), g(b))$ est injective, et bijective si f et g sont bijectives. Par suite, on a $A' \times B' \preceq A \times B$, ainsi que $\mathrm{Card}(A \times B) = \mathrm{Card}(A)\,\mathrm{Card}(B)$.

Pour démontrer la dernière inégalité, supposons d'abord que l'injection g possède une rétraction $g' : B \to B'$; c'est par exemple le cas lorsque $B' \neq \varnothing$ ou lorsque g est bijective. L'application k de $\mathscr{F}(B', A')$ dans $\mathscr{F}(B, A)$ donnée par $u \mapsto f \circ u \circ g'$ est alors injective; on a donc $\mathscr{F}(B', A')) \preceq \mathscr{F}(B, A)$. Si f et g sont bijectives, l'application k est bijective, d'où l'égalité $\mathrm{Card}(A^B) = \mathrm{Card}(\mathscr{F}(B, A)) = \mathrm{Card}(A)^{\mathrm{Card}(B)}$. Enfin, dans le cas où $B' = \varnothing$, l'ensemble des fonctions de B' dans A est réduit à un élément; comme $A \neq \varnothing$, l'ensemble $\mathscr{F}(B; A)$ n'est pas vide, d'où l'inégalité $\mathscr{F}(B'; A') \preceq \mathscr{F}(B; A)$. □

Théorème **(1.5.3)** (Hessenberg). — *Soit* A *un ensemble infini. L'ensemble* $A \times A$ *est équipotent à* A.

Démonstration. — Soit I l'ensemble des couples (V, f), où V est une partie de A et $f : V \to V \times V$ est bijective; on note aussi (V_i, f_i) l'élement i de I. Munissons I de la relation définie par $(V, f) \preceq (W, g)$ si $V \subset W$ et si $g|_V = f$; démontrons que c'est une relation d'ordre. Pour tout $(V, f) \in I$, on a $(V, f) \preceq (W, f)$. Soit (V, f) et (W, g) des éléments de I tels que $(V, f) \leqslant (W, g)$ et $(W, g) \leqslant (V, f)$; alors $V \subset W \subset V$, de sorte que $V = W$; puis $f = g|_V = g$, donc $(V, f) = (W, g)$. Soit enfin (V, f), (W, g) et (X, h) des élements de I tels que $(V, f) \leqslant (W, g)$ et $(W, g) \leqslant (X, h)$; on a $V \subset W \subset X$, donc $V \subset X$; de plus, $f = g|_V = h|_W|_V = h|_V$, ce qui prouve que $(V, f) \leqslant (X, h)$.

Démontrons que l'ensemble ordonné $(I, \leqslant)$ est inductif. Soit J une partie totalement ordonnée de I; posons $V = \bigcup_{j \in J} V_j$. Pour tout $a \in V$, l'ensemble des éléments $j \in J$ tels que $a \in V_j$ n'est pas vide. soit $a \in V$ et soit $j, k \in J$ tels que

$a \in V_j \cap V_k$. Si $j \leqslant k$, alors $V_j \subset V_k$ et $f_k|_{V_j} = f_j$, de sorte que $f_j(a) = f_k(a)$; par symétrie, cette égalité vaut aussi lorsque $k \leqslant j$. Il existe alors une unique application $f : V \to A \times A$ telle que $f(a) = f_j(a)$ pour tout $j \in J$ tel que $a \in V_j$. Démontrons que f est injective et que son image est égale à $V \times V$. Soit $a, b \in V$ tels que $f(a) = f(b)$. Soit $j, k \in J$ tels que $a \in V_j$ et $b \in V_k$; si $j \leqslant k$, alors $V_j \subset V_k$, de sorte que $f_k(a) = f(a) = f(b) = f_k(b)$, d'où $a = b$ car f_k est injective ; on raisonne de même si $k \leqslant j$. Cela démontre que f est injective. Soit $a \in V$ et soit $j \in J$ tel que $a \in V_j$; alors, $f(a) = f_j(a) \in V_j \times V_j \subset V \times V$. L'image de f est donc contenue dans $V \times V$. Soit enfin $(a, b) \in V$; soit $j, k \in J$ tels que $a \in V_j$ et $b \in V_k$. Si $j \leqslant k$, alors $a, b \in V_k$; puisque l'image de f_k est égale à $V_k \times V_k$, il existe $c \in V_k$ tel que $f_k(c) = (a, b)$; on a alors $c \in V$ et $f(c) = f_k(c) = (a, b)$. On raisonne de même si $k \leqslant j$. Par suite, le couple (V, f) est un élément de I ; il majore J par construction. Cela prouve que l'ensemble I est inductif.

Comme A est infini, il contient une partie B_0 équipotente à $\mathbf{N}$, laquelle est équipotente à $\mathbf{N} \times \mathbf{N}$; soit $f_0 : B_0 \to B_0 \times B_0$ une bijection. Soit (B, f) un élément de I qui est maximal et majore (B_0, f_0) (il en existe, en vertu du théorème de Zorn) ; posons $C = A - B$. Par construction, l'ensemble B est infini. Par suite, on a les inégalités

$$\operatorname{Card}(B) \leqslant 2\operatorname{Card}(B) \leqslant 3\operatorname{Card}(B) \leqslant \operatorname{Card}(B)^2 = \operatorname{Card}(B),$$

d'où les égalités

$$\operatorname{Card}(B)^2 = 3\operatorname{Card}(B) = 2\operatorname{Card}(B) = \operatorname{Card}(B).$$

Pour démontrer que B est équipotent à A, nous allons prouver que $\operatorname{Card}(C) < \operatorname{Card}(B)$. Raisonnons par l'absurde et supposons que $\operatorname{Card}(C) \geqslant \operatorname{Card}(B)$. Soit alors B_1 une partie de C qui est équipotente à B. Posons $B' = B \cup B_1$ et définissons une application $f' : B' \to B' \times B'$ qui prolonge f. L'ensemble $B' \times B'$ est la réunion des parties $B \times B$, $B_1 \times B$, $B \times B_1$ et $B_1 \times B_1$, deux à deux disjointes et toutes équipotentes à $B \times B$, donc à B. Comme $\operatorname{Card}(B) = 3\operatorname{Card}(B)$, il existe une bijection f_1 de B_1 sur $(B_1 \times B) \cup (B \times B_1) \cup (B_1 \times B_1)$. L'application $f' : B' \to B' \times B'$ définie par $f'(a) = f(a)$ pour $a \in B$ et par $f'(a) = f_1(a)$ pour $a \in B_1$ est bijective. Le couple (B', f') appartient à I, majore (B, f) et est distinct de (B, f), ce qui contredit l'hypothèse que (B, f) est un élément maximal de I.

Par conséquent, $\operatorname{Card}(C) < \operatorname{Card}(B)$, d'où $\operatorname{Card}(A) = \operatorname{Card}(B) + \operatorname{Card}(C) = \operatorname{Card}(B)$ puisque B est infini. Ainsi, $\operatorname{Card}(A \times A) = \operatorname{Card}(B \times B) = \operatorname{Card}(B) = \operatorname{Card}(A)$, ce qu'il fallait démontrer. □

Corollaire **(1.5.4)**. — *Soit α et β des cardinaux tels que $1 \leqslant \beta \leqslant \alpha$; Supposons que α est infini. On a alors $\alpha + \beta = \alpha$ et $\alpha\beta = \alpha$. En particulier, $\alpha^2 = 2\alpha = \alpha$.*

Démonstration. — Comme α est infini, le théorème précédent implique l'égalité $\alpha\alpha = \alpha$. Puisque $2 \leqslant \alpha$, on a l'inégalité $\alpha \leqslant \alpha + \beta \leqslant \alpha + \alpha \leqslant 2\alpha \leqslant \alpha\alpha = \alpha$, si bien que $\alpha + \beta = \alpha$. Comme $\beta \neq 0$, on a aussi $\alpha \leqslant \alpha\beta \leqslant \alpha\alpha = \alpha$, d'où l'égalité $\alpha\beta = \alpha$. □

Corollaire **(1.5.5)**. — *Soit* A, B *des ensembles et soit* $f : A \to B$ *une application. Soit α un cardinal infini tel que* $\mathrm{Card}(B) \leqslant \alpha$ *et* $\mathrm{Card}(f^{-1}(b)) \leqslant \alpha$ *pour tout* $b \in B$*; alors,* $\mathrm{Card}(A) \leqslant \alpha$.

Démonstration. — Pour tout $b \in B$, soit $g_b : f^{-1}(b) \to \alpha$ une application injective. L'application $g : A \to B \times \alpha$ définie par $g(a) = (f(a), g_{f(a)}(b))$ est injective. On a donc $\mathrm{Card}(A) \leqslant \mathrm{Card}(B \times \alpha) \leqslant \alpha\alpha = \alpha$. □

Corollaire **(1.5.6)**. — *Soit α un cardinal infini. Soit* I *un ensemble et soit* $(A_i)_{i\in I}$ *une famille d'ensembles; on note* $A = \bigcup_{i\in I} A_i$. *Si* $\mathrm{Card}(I) \leqslant \alpha$ *et* $\mathrm{Card}(A_i) \leqslant \alpha$ *pour tout* $i \in I$*, alors* $\mathrm{Card}(A) \leqslant \alpha$.

Démonstration. — Soit A' la somme de la famille A_i ; c'est l'ensemble des couples $(i, a) \in I \times A$ tels que $a \in A_i$. Soit $f : A' \to I$ et $g : A' \to A$ les deux projections. Pour tout $i \in I$, l'application de A_i dans $f^{-1}(i)$ définie par $a \mapsto (i, a)$ est bijective, donc $\mathrm{Card}(f^{-1}(i)) \leqslant \alpha$; d'après le corollaire précédent, on a donc $\mathrm{Card}(A') \leqslant \alpha$. L'application g est surjective, par hypothèse, donc $\mathrm{Card}(A) \leqslant \mathrm{Card}(A') \leqslant \alpha$. □

Théorème **(1.5.7)** (Cantor). — *Soit* A *un ensemble. Il n'existe pas d'application surjective de* A *dans* $\mathscr{P}(A)$.

Démonstration. — Raisonnons par l'absurde et considérons une application surjective $f : A \to \mathscr{P}(A)$. Soit B l'ensemble des éléments $a \in A$ tels que $a \notin f(a)$. Par hypothèse, il existe un élément $b \in A$ tel que $B = f(b)$. Si $b \in B$, alors $b \notin f(b) = B$, par définition de B ; donc $b \notin B$, c'est-à-dire $b \in f(b) = B$, contradiction. □

Corollaire **(1.5.8)**. — *Pour tout cardinal α, on a $2^\alpha > \alpha$.*

Proposition **(1.5.9)**. — *Les ensembles* $\mathscr{P}(\mathbf{N})$ *et* $\mathbf{R}$ *sont équipotents.*

On dit qu'ils ont la *puissance du continu.*

Démonstration. — En identifiant une partie A de $\mathbf{N}$ à sa fonction caractéristique $\chi_A : \mathbf{N} \to \{0,1\}$, on remplace $\mathscr{P}(\mathbf{N})$ par l'ensemble $\{0,1\}^{\mathbf{N}}$ des fonctions de $\mathbf{N}$ dans $\{0,1\}$. L'idée de la construction est de représenter un nombre réel par son développement en base 2. Cependant, certains nombres réels ont plusieurs développements, par exemple $1{,}000\cdots = 0{,}1111\ldots$, ce qui nous oblige à quelque précaution. Pour $a \in A$, on pose

$$f(a) = 0{,}a_0a_1a_2\cdots = \sum_{n=0}^{\infty} a_n 2^{-n-1}.$$

L'image de f est l'intervalle $[0,1]$, d'où l'inégalité

$$\mathrm{Card}([0,1]) \leqslant \mathrm{Card}(\mathscr{F}(\mathbf{N},\{0,1\})).$$

Soit $a, b \in A$ tels que $f(a) = f(b)$ et $a \neq b$; soit m le plus petit entier tel que $a_m \neq b_m$; quitte à échanger a et b, on suppose que $a_m = 0$ et $b_m = 1$. Démontrons que $a = (a_0, \ldots, a_{m-1}, 0, 1, 1, \ldots)$ et $b = (a_0, \ldots, a_{m-1}, 1, 0, 0, \ldots)$. Posons $a' = (a_m, a_{m+1}, \ldots)$ et $b' = (b_m, b_{m+1}, \ldots)$; on a donc

$$f(a') = 2^{m+1}\left(f(a) - \sum_{n=0}^{m-1} a_n 2^{-n-1}\right) = 2^{m+1}\left(f(b) - \sum_{n=0}^{m-1} b_n 2^{-n-1}\right) = f(b'),$$

ce qui permet de supposer que $m = 0$ et ramène à prouver $a = (0, 1, 1, \ldots)$ et $b = (1, 0, 0, \ldots)$. On déduit de la définition de f que $f(a) \leqslant 1 \leqslant f(b)$, donc $f(a) = f(b) = 1$. Supposons qu'il existe un entier $m \geqslant 1$ tel que $a_m \neq 1$, on a

$$f(a) = \sum_{n=0}^{\infty} a_n 2^{-n-1} \leqslant \sum_{\substack{n=0 \\ n\neq m}} a_n 2^{-n-1} \leqslant \sum_{\substack{n=0 \\ n\neq m}} 2^{-n-1} = 1 - 2^{-m-1} < 1.$$

Cette contradiction prouve que $a = (0, 1, 1, \ldots)$. De même, s'il existe un entier $m \geqslant 1$ tel que $b_m \neq 0$, on a

$$f(b) = \sum_{n=0}^{\infty} b_n 2^{-n-1} \geqslant 1 + 2^{-m-1} > 1,$$

si bien que $b = (1, 0, 0, \ldots)$. Par conséquent, chaque élément de $[0,1]$ a au plus deux antécédents. Appliquons le corollaire 1.5.5 à l'application f et au cardinal $\mathrm{Card}([0,1])$; on en déduit

$$\mathrm{Card}(\mathscr{F}(\mathbf{N},\{0,1\})) \leqslant 2 \cdot \mathrm{Card}([0,1]) = \mathrm{Card}([0,1]).$$

On a donc l'égalité $\mathrm{Card}([0,1]) = \mathrm{Card}(\mathscr{P}(\mathbf{N}))$.

Pour conclure la démonstration de la proposition, il reste à vérifier que $\mathrm{Card}([0,1]) = \mathrm{Card}(\mathbf{R})$. L'application $t \mapsto \cot(\pi t)$ définit une bijection de $]0,1[$ sur $\mathbf{R}$; on a donc

$$\mathrm{Card}([0,1]) \leqslant \mathrm{Card}(\mathbf{R}) = \mathrm{Card}(]0,1[) \leqslant \mathrm{Card}([0,1]),$$

d'où l'égalité voulue. □

Remarque **(1.5.10)**. — L'*hypothèse du continu* postule qu'il n'existe aucun cardinal α tel que $\aleph_0 < \alpha < 2^{\aleph_0}$; autrement dit, si une partie de $\mathbf{R}$ n'est pas dénombrable, son cardinal est la puissance du continu. D'après des théorèmes remarquables de Gödel et Cohen, elle ne peut en fait ni être infirmée, ni démontrée, à partir des simples axiomes de la théorie des ensembles (ZFC).

CHAPITRE 2

LANGAGES DU PREMIER ORDRE — *PARLER*

« (...) J'enrage
Lorsque j'entends tenir ces sortes de langage. »
— Molière, *Le Tartuffe*, acte II, scène 3

2.1. Alphabets, mots, langages

2.1.1. — Soit A un ensemble non vide. On note A^* l'ensemble des suites finies d'éléments de A, c'est-à-dire la réunion des ensembles A^n, lorsque n parcourt **N**. L'ensemble A est appelé *alphabet* ; un élément de A^* est appelé *mot* sur l'alphabet A. Si $m = (a_0, \ldots, a_{n-1})$ est un élément de A^*, l'entier n est appelé *longueur*du mot m et noté $\ell(m)$. La suite de longueur 0 est le mot vide ; on la notera ε. On identifiera aussi l'alphabet A à la partie de A^* des mots de longueur 1.

Lemme **(2.1.2)**. — *Soit* A *un alphabet. L'ensemble* A^* *des mots sur l'alphabet* A *est équipotent à* **N** *si* A *est dénombrable, et à* A *sinon.*

Démonstration. — Supposons d'abord que A est fini. Pour tout entier n, l'ensemble A^n est fini. Comme une réunion dénombrable d'ensembles finis est dénombrable, l'ensemble A^* est dénombrable. Par ailleurs, comme l'alphabet A n'est pas vide, A^* contient des mots de toute longueur. Par suite, A^* est infini. On a donc $A^* \equiv \mathbf{N}$.

Supposons maintenant que A est infini. Pour tout entier $n \geqslant 1$, on a donc $\mathrm{Card}(A^n) = \mathrm{Card}(A)$. Puisque $\aleph_0 \leqslant \mathrm{Card}(A)$, on en déduit que $\mathrm{Card}(A^*) \leqslant \mathrm{Card}(A)$. D'autre part, l'ensemble des mots de longueur 1 est une partie de A^* équipotente à A, donc $\mathrm{Card}(A) \leqslant \mathrm{Card}(A^*)$. Il en résulte que $\mathrm{Card}(A^*) = \mathrm{Card}(A)$. □

2.1.3. — Soit A un alphabet. Soit $m = (a_0, \dots, a_{n-1})$ et $m' = (a'_0, \dots, a'_{n'-1})$ des mots dans l'alphabet A. On note $mm' = (a_0, \dots, a_{n-1}, a'_0, \dots, a'_{n'-1})$ le mot obtenu par concaténation de m et m'. On a $\ell(mm') = \ell(m + m')$.

L'opération $(m, m') \mapsto mm'$ dans A^* est associative, le mot vide est un élément neutre. Par ailleurs, tout mot est simplifiable à droite et à gauche.

Compte tenu de l'identification des éléments de A avec les mots de longueur 1, on a $m = a_0 \dots a_{n-1}$.

Définition **(2.1.4)**. — *Soit* A *un alphabet et soit* $m = (a_0, \dots, a_{n-1})$ *un mot sur* A.

a) *Soit* $a \in$ A. *On dit que* a *apparaît dans* m *s'il existe un entier* k *tel que* $0 \leqslant k \leqslant n-1$ *et* $a = a_k$ *; un tel entier est appelé* occurrence *de la lettre* a.

b) *On dit qu'un mot* m' *est un* segment initial *(resp.* terminal*) de* m *s'il existe un mot* m'' *tel que* $m = m'm''$ *(resp.* $m = m''m'$*); cela revient à dire qu'il existe un entier* k *tel que* $0 \leqslant k \leqslant n$ *et* $m' = (a_0, \dots, a_{k-1})$ *(resp.* $m' = (a_k, \dots, a_{n-1})$*).*

2.1.5. — Soit V un ensemble. Ses éléments sont appelés symboles de *variables*. Conformément aux usages, nous supposons que V est infini et dénombrable et notons $x_0, x_1, x_2, \dots$ ses éléments. Cependant, d'autres notations $(x, y, y_0, y_1, \dots)$ sont possibles, et parfois plus pratiques. Cet ensemble sera supposé fixé dans toute la suite.

2.1.6. — Un *langage du premier ordre* est la donnée de deux suites $(\mathrm{R}_n)_{n\geqslant 0}$ et $(\mathrm{F}_n)_{n\geqslant 0}$ d'ensembles, deux à deux disjoints, et disjoints de l'ensemble V et de l'ensemble formé des symboles), (, $\vee$, $\wedge$, $\neg$, $\Rightarrow$, $\Leftrightarrow$, $\forall$, $\exists$.

Les éléments de R_n sont appelés symboles de *relations n-aires*, ceux de F_n symboles de *fonctions n-aires*, ou *constantes* lorsque $n = 0$. On suppose aussi que l'ensemble R_0 contient deux éléments $\top$ (« vrai ») et $\bot$ (« faux ») et que l'ensemble R_2 contient le symbole = (« égalité »).

Si ces symboles n'ont aucune signification à ce stade de la théorie, le rôle ultime d'un langage est d'exprimer les propriétés utilisées dans une théorie mathématique donnée, données par des *mots* de l'alphabet

$$\mathrm{V} \cup \bigcup_{n=0}^{\infty} \mathrm{R}_n \cup \bigcup_{n=0}^{\infty} \mathrm{F}_n \cup \{\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, (,)\}.$$

Exemple **(2.1.7)**. — Voici ainsi quelques exemples de langages.

a) *Ensembles.* L'ensemble F_0 possède un élément $\varnothing$; l'ensemble R_2 possède deux éléments, $\in$ et = ; les autres ensembles de relations et de fonctions sont

vides. Dans ce langage, on peut écrire les axiomes de la théorie des ensembles et développer cette théorie, les relations et fonctions supplémentaires sont codées via leur graphe.

b) *Groupes.* L'ensemble F_0 possède un élément e ; l'ensemble F_1 possède un élément i ; l'ensemble R_2 possède deux éléments $\cdot$ et $=$; les autres ensembles de relations et de fonctions sont vides. Dans ce langage, on peut en effet écrire les axiomes de la théorie des groupes et la développer ; le symbole de constante e correspond à l'élément neutre, le symbole i à la fonction inverse et le symbole $\cdot$ à la loi de groupe. La propriété, pour un groupe, d'être commutatif s'écrit, par exemple, $\forall x \forall y\, x \cdot y = y \cdot x$.

c) *Anneaux.* L'ensemble F_0 possède deux éléments 0 et 1 ; l'ensemble R_2 possède quatre éléments $+$, $-$, $\cdot$ et $=$.

d) *Géométrie plane.* L'ensemble R_1 possède deux éléments D et P, correspondant respectivement à la propriété d'être une droite et à la propriété d'être un point ; l'ensemble R_2 possède un élément $\in$, correspondant à la relation d'incidence, c'est-à-dire, pour un point d'appartenir à une droite. L'axiome des parallèles s'énonce par exemple

$$\forall x \forall y (\mathrm{P}(x) \wedge \mathrm{D}(y) \wedge \neg x \in y) \Rightarrow (\exists z (\mathrm{D}(z) \wedge \forall t (\mathrm{P}(t) \wedge t \in z \Rightarrow \neg t \in y))),$$

c'est-à-dire pour tout point x et toute droite y ne contenant pas x, il existe une droite z passant par x et ne contenant aucun point de y.

2.2. Termes

2.2.1. — On fixe un langage du premier ordre $\mathscr{L} = ((R_n), (F_n))$. On notera A_L l'alphabet correspondant, réunion de V, des R_n, des F_n, et de l'ensemble des symboles logiques $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, (,)$.

Définition **(2.2.2)**. — *On note $\mathscr{T}_L$ l'intersection de toutes les parties* T *de A_L^* qui vérifient les propriétés suivantes :*

– T *contient les symboles de variables* V *;*

– *Pour tout entier $n \in \mathbf{N}$, tout symbole $f \in F_n$ de fonction n-aire et toute suite $(t_1, \ldots, t_n)$ d'éléments de* T*, le mot $f t_1 \ldots t_n$ appartient à* T.

On voit immédiatement que $\mathscr{T}_L$ vérifie les deux propriétés précédentes ; c'est donc la plus petite partie de A_L^* qui les vérifie. Un élément de $\mathscr{T}_L$ est appelé un *terme* dans le langage L.

Soit T une partie de A_L^*. Pour prouver que T contient $\mathscr{T}_L$, il suffit d'établir les propriétés de la définition. Cela donne lieu à une méthode de « raisonnement par récurrence sur la définition d'un terme ». Par exemple, on vérifie ainsi que les seuls symboles apparaissant dans un terme sont les symboles de fonctions et de variables. De même, l'ensemble $\mathscr{T}_L - \{\varepsilon\}$ vérifie les propriétés de la définition, donc le mot vide n'est pas un terme.

2.2.3. — On définit par récurrence une suite croissante $(\mathscr{T}_L^{(h)})$ de mots de la façon suivante : on pose $\mathscr{T}_L^{(0)} = V \cup F_0$ et, si $\mathscr{T}_L^{(h)}$ est défini, on définit $\mathscr{T}_L^{(h+1)}$ comme la réunion de $\mathscr{T}_L^{(h)}$ et de l'ensemble des mots de la forme $f t_1 \dots t_n$, où $n \in \mathbf{N}^*$, $f \in F_n$ et $t_1, \dots, t_n \in \mathscr{T}_L^{(h)}$. Par récurrence, on a $\mathscr{T}_L^{(h)} \subset \mathscr{T}_L$ pour tout entier h. On vérifie immédiatement que l'ensemble $\bigcup_{h\in\mathbf{N}} \mathscr{T}_L^{(h)}$ vérifie les propriétés de la définition précédente. Par suite, $\mathscr{T}_L = \bigcup_{h\in\mathbf{N}} \mathscr{T}_L^{(h)}$. La *hauteur* d'un terme t est le plus petit entier h tel que $t \in \mathscr{T}_L^{(h)}$.

Remarque **(2.2.4)**. — La définition des termes n'est pas tout à fait conforme à l'habitude mathématique. Par exemple, si f est un symbole de fonction binaire et x, y sont des variables, on écrit fxy au lieu de $f(x, y)$. De même, on écrit $=xy$ au lieu de $x = y$. Cette définition formelle s'avère très pratique, parce qu'elle est concise et inambigüe. Le plus souvent, nous écrirons cependant les termes sous leur forme classique.

2.2.5. — Pour $v \in V$, on dit qu'un terme t fait intervenir le symbole v s'il existe une occurrence de ce symbole dans v. Si W est une partie de l'ensemble V des symboles de variables, on note $\mathscr{T}_{L,W}$ l'ensemble des termes qui ne font intervenir que les symboles de variables appartenant à W. On vérifie immédiatement que l'ensemble $\mathscr{T}_{L,W}$ est l'intersection des parties de A_L^* qui vérifient la définition 2.2.2 où l'on remplace V par W dans la première propriété.

2.2.6. — On définit le *poids* $p(x)$ d'un symbole x de A_L de la façon suivante : si $x \in V$, on pose $p(x) = -1$; si $x \in F_n$, on pose $p(x) = n - 1$; sinon, on pose $p(x) = 0$.

Soit $m \in A_L^*$; on appelle *poids* du mot m la somme des poids des symboles qui le constituent : si $m = x_1 \dots x_n$, alors $p(m) = p(x_1) + \cdots + p(x_n)$.

Proposition **(2.2.7)**. — *Soit $t \in \mathrm{A}_{\mathrm{L}}^*$ un mot ne faisant intervenir que des symboles de variables et de fonctions. Pour que t soit un terme il faut et il suffit que les deux propriétés suivantes soient satisfaites :*

(i) *On a $p(t) = -1$;*
(ii) *Pour tout segment initial t' de t tel que $t' \neq t$, on a $p(t') \geqslant 0$.*

Démonstration. — Soit $\mathscr{T}_{\mathrm{L}}'$ l'ensemble T des mots qui vérifient ces propriétés; démontrons que $\mathscr{T}_{\mathrm{L}}'$ satisfait les conditions énoncés dans la définition de l'ensemble des termes.

– Soit t un symbole de variable ou de constante. Alors $p(t) = -1$ par définition. De plus, si t' est un segment initial distinct de t, on a $t' = \varepsilon$, car t est de longueur 1, donc $p(t') = 0$.

– Soit $f \in \mathrm{F}_n$, soit $t_1, \dots, t_n$ des termes de $\mathscr{T}_{\mathrm{L}}'$ et soit $t = f t_1 \dots t_n$. On a $p(t) = p(f) + p(t_1) + \cdots + p(t_n) = (n-1) + (-n) = -1$. Soit d'autre part t' un segment initial de t. Si $t' = \varepsilon$, alors $p(t') = 0$. Sinon, il existe un entier m tel que $1 \leqslant m \leqslant n$ et un segment initial t'_m de t_m tel que $t' = f t_1 \dots t_{m-1} t'_m$; par récurrence, on a alors

$$p(t') = p(f) + \sum_{k=1}^{m-1} p(t_i) + p(t'_m) = (n-1) - (m-1) + p(t'_m) \geqslant n - m \geqslant 0.$$

Par définition de l'ensemble des termes, on a donc $\mathscr{T}_{\mathrm{L}} \subset \mathscr{T}_{\mathrm{L}}'$.

Inversement, soit t un mot vérifiant ces propriétés et démontrons par récurrence sur la longueur de t que t est un terme. Comme $p(\varepsilon) = 0$ et $p(t) = -1$, on a $\ell(t) > 0$; soit donc f le premier symbole de t. Si f est un symbole de variable ou de constante, on a $p(f) = -1$; par suite, f n'est pas un segment initial strict de t, donc $t = f$ et t est un terme. Sinon, f est un symbole de fonction dont l'arité est $\geqslant 1$; notons n cette arité. Puisque $p(f) = n - 1 \geqslant 0$, la longueur de f est au moins égale à 2.

Écrivons $t = x_1 \dots x_m$, de sorte que $m = \ell(t)$ et $f = x_1$. Pour $i \in \{1, \dots, m\}$, posons $p_i = 1 + p(x_1 \dots x_i)$. Par hypothèse, on a $p_1 = n$, $p_i \geqslant 1$ pour $i \in \{1, \dots, m-1\}$, et $p_m = 0$. Comme $p_i - p_{i-1} = p(x_i) \geqslant -1$ pour tout i, p_i atteint au moins une fois chaque valeur entre 0 et n. Pour $k \in \{1, \dots, n\}$, soit j_k le plus petit entier $j \geqslant 1$ tel que $p_j = n - k$. Par le même argument, p_i atteint $n - k$ avant d'avoir atteint $n - k - 1$, de sorte que l'on a $1 \leqslant j_1 < \cdots < j_n \leqslant m$. Posons alors $t_1 = x_2 \dots x_{j_1}$, $t_2 = x_{j_1+1} \dots x_{j_2}$,..., $t_n = x_{j_{n-1}+1} \dots x_{j_n}$. Alors, $f t_1 \dots t_n$ est

un segment initial de t tel que $p(f\,t_1 \dots t_n) = p(x_1 \dots x_{j_n}) = p_{j_n} - 1 = -1$; par hypothèse, on a donc $t = f\,t_1 \dots t_n$ (et $j_n = m$).

Démontrons que les mots $t_1, \dots, t_n$ sont des termes. Soit $i \in \{1, \dots, n\}$. On a $p(t_i) = p_{j_i} - p_{j_{i-1}} = -1$ par construction des entiers j_i. Soit t' un segment initial de t_i distinct de t_i, et soit $j \in \{j_{i-1}, \dots, j_i - 1\}$ l'unique entier tel que $t' = x_{j_{i-1}+1} \dots x_j$; par définition de l'entier j_i, on a $p(t') = p_j - p_{j_{i-1}} > p_{j_i} - p_{j_{i-1}} = -1$, de sorte que $p(t') \geqslant 0$. Comme la longueur de t_i est strictement inférieure à $m - 1$, l'hypothèse de récurrence entraîne que t_i est un terme.

Comme f est un symbole de fonction d'arité n, il en résulte que $t = f\,t_1 \dots t_n$ est un terme, ce qu'il fallait démontrer. □

Corollaire **(2.2.8)**. — *Soit $t \in \mathscr{T}_{\mathrm{L}}$. Aucun segment initial de t, distinct de t, n'est un terme.*

Démonstration. — Considérons en effet un segment initial t' de t tel que $t' \neq t$. D'après la proposition, on a $p(t') \geqslant 0$. Comme le poids d'un terme est égal à -1, donc t' n'est pas un terme. □

Théorème **(2.2.9)**. — *Soit $t \in \mathscr{T}_{\mathrm{L}}$ un terme. Alors une, et une seule, des assertions suivantes est vérifiée :*

i) *t est un symbole de variable de* L*;*

ii) *Il existe un unique entier $n \geqslant 0$, un unique symbole de fonction n-aire f, et une unique suite $(t_1, \dots, t_n)$ de termes tels que $t = f\,t_1 \dots t_n$.*

Démonstration. — Le fait qu'une des deux assertions soit vraie se vérifie immédiatement par récurrence sur la définition d'un terme. D'autre part, ces deux cas s'excluent mutuellement puisque dans le second, le premier symbole de t est un élément de F_n, donc n'appartient pas à V.

Supposons donc que l'on ait $t = f\,t_1 \dots t_n = g\,u_1 \dots u_m$, pour des symboles de fonction $f \in \mathrm{F}_n$, $g \in \mathrm{F}_m$, et des termes $t_1, \dots, t_n, u_1, \dots, u_m$. Démontrons que $f = g$, $n = m$, et que $t_i = u_i$ pour tout i. En considérant la première lettre de t, on voit que $f = g$; en particulier, $n = m$. Démontrons par récurrence que $t_i = u_i$ pour tout i. Raisonnons par l'aburde : supposons $t_1 = u_1, \dots, t_{i-1} = u_{i-1}$ et $t_i \neq u_i$. Puisque $f\,t_1 \dots t_{i-1} = g\,u_1 \dots u_{i-1}$ et $f\,t_1 \dots t_n = g\,u_1 \dots u_n$, on a donc $t_i \dots t_n = u_i \dots u_n$. Supposons que la longueur de t_i est inférieure ou égale à celle de u_i; alors t_i est un segment initial de u_i; comme t_i est un terme, on a $p(t_i) = -1$; cela entraîne que $t_i = u_i$. □

Corollaire **(2.2.10)**. — *Soit* $s : \mathrm{V} \to \mathscr{T}_\mathrm{L}$ *une application. Il existe une unique application* $\sigma : \mathscr{T}_\mathrm{L} \to \mathscr{T}_\mathrm{L}$ *vérifiant les propriétés suivantes :*

a) *Si* t *est un symbole de variable, alors* $\sigma(t) = s(t)$;
b) *Si* $f \in \mathrm{F}_n$, $t_1, \dots, t_n \in \mathscr{T}_\mathrm{L}$, *alors* $\sigma(f t_1 \dots t_n) = f\sigma(t_1) \dots \sigma(t_n)$.

Le terme $\sigma(t)$ est appelé terme déduit de t par *substitution* des termes $s(x)$ aux variables x.

En pratique, on dispose d'une partie finie $\{x_1, \dots, x_k\}$ de variables (deux à deux distinctes) et de termes $u_1, \dots, u_k$, l'application s étant définie par $s(x_i) = u_i$ pour tout i, et $s(x) = x$ pour $x \in \mathrm{V} - \{x_1, \dots, x_k\}$. Dans ce cas, le terme $\sigma(t)$ est parfois noté $t_{u_1/x_1,\dots,u_k/x_k}$. Il arrive aussi que le terme t soit noté $t(x_1, \dots, x_k)$ pour mettre en valeur les variables ; dans ce cas, le terme $\sigma(t)$ peut être noté $t(u_1, \dots, u_k)$.

Démonstration. — Il découle du théorème précédent que ces conditions définissent une unique application $\sigma : \mathscr{T}_\mathrm{L} \to \mathrm{A}_\mathrm{L}^*$. On démontre alors par récurrence sur la définition d'un terme que $\sigma(t)$ est un terme, pour tout terme t. □

2.2.11. Sous-termes. — Soit t un terme dans le langage L. On définit par récurrence l'ensemble $\mathrm{st}(t)$ des sous-termes de L :

– Si t est un symbole de variable v, alors $\mathrm{st}(t) = \{t\}$;
– Si $t = f t_1 \dots t_n$, où $f \in \mathrm{F}_n$ et $t_1, \dots, t_n \in \mathscr{T}_\mathrm{L}$, alors $\mathrm{st}(t)$ est la réunion de $\{t\}$ et de l'ensemble somme des ensembles $\mathrm{st}(t_i)$, pour $i \in \{1, \dots, n\}$.

Cet ensemble est naturellement muni d'une relation d'ordre, elle-même définie par récurrence : si t' et t'' sont des sous-termes de t, $t' \preceq t''$ si t' est un sous-terme de t''. On voit que $\mathrm{st}(t)$ s'organise en un arbre (inversé) qu'on appelle l'arbre d'évaluation du terme t : Le plus grand élément t est au sommet ; si $t = f t_1 \dots t_n$, les fils de t sont les termes $t_1, \dots, t_n$, et ainsi de suite par récurrence.

Donnons l'exemple du terme $t = -+\cdot xxy1$ du langage des anneaux, où 1 est un symbole de constante et x, y des symboles de variables. correspondant au polynôme $x^2 + y - 1$. Appliquons la définition :

$$\begin{aligned}\mathrm{st}(t) &= \{t\} \cup \mathrm{st}(+\cdot xxy) \cup \mathrm{st}(1)\\ &= \{-+\cdot xxy1, +\cdot xxy, 1\} \cup \mathrm{st}(\cdot xx) \cup \mathrm{st}(y)\\ &= \{-+\cdot xxy1, +\cdot xxy, 1, \cdot xx, y, x, x\}.\end{aligned}$$

Prendre garde que les deux éléments x figurant dans la formule précédente sont distincts : ils correspondent aux deux arguments du symbole de fonction $\cdot$. La représentation en arbre est indiquée dans la figure 1; on a écrit dans chaque nœud un symbole de variable ou de fonction, le sous-terme correspondant s'en déduit en lisant le sous-arbre qui en est issu.

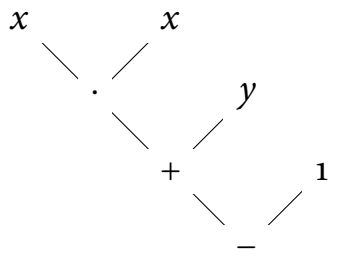

FIGURE 1. Arbre d'évaluation du terme $-+\cdot xxy1$

2.3. Formules

On fixe un langage L.

Définition **(2.3.1)**. — *Un mot m dans l'alphabet* $\mathrm{A_L}$ *est appelé* formule atomique *s'il existe un entier naturel n, un symbole de relation $r \in \mathrm{R}_n$ et une suite $(t_1, \ldots, t_n)$ de termes du langage* L *tels que $m = rt_1 \ldots t_n$.*

Proposition **(2.3.2)**. — *Soit m une formule atomique dans le langage* L. *Il existe un unique entier n, un unique symbole de relation $r \in \mathrm{R}_n$ et une unique suite $(t_1, \ldots, t_n)$ de termes du langage* L *tels que $m = rt_1 \ldots t_n$.*

Démonstration. — Soit k un entier, $s \in \mathrm{R}_k$ un symbole de relation k-aire et $(u_1, \ldots, u_k)$ une suite de termes dans le langage L tels que $m = su_1 \ldots u_k$. Démontrons que $k = n$, $s = r$ et $(u_1, \ldots, u_k) = (t_1, \ldots, t_n)$. Considérons la variante L′ du langage L dont l'ensemble de symboles de fonctions la réunion des ensembles de symboles de fonctions et de symboles de relations de L. Si t est un terme dans le langage L, c'est encore un terme dans le langage L′ ; comme r est symbole de fonction du langage L, le mot m est un terme du langage L′. D'après le théorème de lecture unique des termes, il existe on a $k = n$, $s = r$ et $(u_1, \ldots, u_k) = (t_1, \ldots, t_n)$, ce qu'il fallait démontrer. □

Définition **(2.3.3)**. — *On note* $\mathscr{F}_{\mathrm{L}}$ *l'intersection de toutes les parties* F *de* $\mathrm{A}_{\mathrm{L}}^{*}$ *vérifiant les propriétés suivantes :*

– F *contient toutes les formules atomiques ;*

– *Pour tout mot* $m \in \mathrm{F}$*, le mot* $\neg m$ *appartient à* F*;*

– *Pour tous mots* $m, m' \in \mathrm{F}$*, les mots* $(m \vee m')$*,* $(m \wedge m')$*,* $(m \Rightarrow m')$ *et* $(m \Leftrightarrow m')$ *appartiennent à* F*;*

– *Pour tout mot* $m \in \mathrm{F}$ *et tout symbole de variable* $v \in \mathrm{V}$*, les mots* $\exists v m$ *et* $\forall v m$ *appartiennent à* F.

On voit immédiatement que $\mathscr{F}_{\mathrm{L}}$ vérifie les quatre propriétés précédentes ; c'est donc la plus petite partie de $\mathrm{A}_{\mathrm{L}}^{*}$ qui les vérifie. On dit qu'un élément de $\mathscr{F}_{\mathrm{L}}$ est une *formule* dans le langage L. Si m et m' sont des formules, on dit que $\neg m$ est la *négation* de la formule m, que $(m \vee m')$ et $(m \wedge m')$ sont la disjonction et la conjonction des formules m et m'. Soit F une partie de $\mathrm{A}_{\mathrm{L}}^{*}$. Pour prouver que F contient $\mathscr{F}_{\mathrm{L}}$, il suffit d'établir les propriétés de la définition. Cela donne lieu à une méthode de « raisonnement par récurrence sur la définition d'un terme ». On vérifie par exmple ainsi que le mot vide n'est pas une formule.

2.3.4. — On définit par récurrence une suite croissante $(\mathscr{F}_{\mathrm{L}}^{(h)})$ d'ensembles de mots de la façon suivante : on note $\mathscr{F}_{\mathrm{L}}^{(0)}$ l'ensemble des formules atomiques et, si $\mathscr{F}_{\mathrm{L}}^{(h)}$ est défini, on note $\mathscr{F}_{\mathrm{L}}^{(h+1)}$ la réunion de $\mathscr{F}_{\mathrm{L}}^{(h)}$ et de l'ensemble des mots de la forme $\neg m$, $(m \vee m')$, $(m \wedge m')$, $(m \Rightarrow m')$, $(m \Leftrightarrow m')$, $\exists v m$ ou $\forall v m$, où m, m' sont des éléments de $\mathscr{F}_{\mathrm{L}}^{(h)}$ et $v \in \mathrm{V}$ est un symbole de variable. On voit par récurrence que pour tout entier h, $\mathscr{F}_{\mathrm{L}}^{(h)}$ est contenu dans $\mathscr{F}_{\mathrm{L}}$. De plus, la réunion $\bigcup_{h \in \mathbf{N}} \mathscr{F}_{\mathrm{L}}^{(h)}$ vérifie les propriétés de la définition de l'ensemble des termes ; cela entraîne l'égalité $\mathscr{F}_{\mathrm{L}} = \bigcup_{h \in \mathbf{N}} \mathscr{F}_{\mathrm{L}}^{(h)}$. La *hauteur*d'une formule m est le plus petit entier h tel que $m \in \mathscr{F}_{\mathrm{L}}^{(h)}$.

2.3.5. — Si m est un mot, on note $p'(m)$ la différence du nombre de parenthèses ouvrantes et du nombre de parenthèses fermantes qui apparaissent dans m.

Proposition **(2.3.6)**. — *Soit* $m \in \mathscr{F}_{\mathrm{L}}$ *une formule dans le langage* L.

a) *On a* $p'(m) = 0$*;*

b) *Pour tout segment initial* m' *de* m*, on a* $p'(m') \geqslant 0$*;*

c) *Si le premier symbole de* m *est* (*, alors* $p'(m') \geqslant 1$ *pour tout segment initial* m' *de* m *qui est distinct de* ε *et de* m*.*

Démonstration. — a) Raisonnons par récurrence sur la définition d'une formule. Si m est une formule atomique, alors m n'a pas de parenthèse, donc $p'(m) = 0$; si m est de la forme $\neg m_1$, où m_1 est une formule, alors $p'(m) = p'(m_1) = 0$; si m est de la forme $(m_1 \bullet m_2)$, alors $p'(m) = 1 + p'(m_1) + p'(m_2) - 1 = 0$; si m est de la forme $\exists v m_1$ ou $\forall v m_1$, alors $p'(m) = p'(m_1) = 0$. L'assertion en résulte.

b) et c) Raisonnons encore par récurrence sur la définition d'une formule. Soit m' un segment initial de m, que l'on peut supposer distinct de m. Si m est une formule atomique, alors m n'a pas de parenthèses, donc m' non plus et $p'(m') = 0$. Supposons m de la forme $\neg m_1$; alors il existe un segment initial m'_1 de m tel que $m' = \neg m'_1$; on a donc $p'(m') = p'(m'_1) \geqslant 0$. Supposons m de la forme $(m_1 \bullet m_2)$, où m_1 et m_2 sont des formules. Si $\ell(m') = 0$, alors $p'(m') = 0$; si $1 \leqslant \ell(m') \leqslant 1 + \ell(m_1)$, alors il existe un segment initial m'_1 de m_1 tel que $m' = (m'_1$, donc $p'(m') = 1 + p'(m'_1) \geqslant 1$. Si $2 + \ell(m_1) \leqslant \ell(m') < \ell(m)$, il existe un segment initial m'_2 de m_2 tel que $m' = (m_1 \bullet m'_2$; on a donc $p'(m') = 1 + p'(m_1) + p'(m'_2) \geqslant 1$. Les deux assertions b) et c) en résultent.

□

Corollaire **(2.3.7)**. — *Soit m une formule et soit m' un segment initial de m, distinct de m. Alors, m' n'est pas une formule.*

Démonstration. — Raisonnons par récurrence sur la définition d'une formule. Le mot vide n'est pas une formule ; on suppose dans la suite que $m' \neq \varepsilon$. On raisonne par l'absurde. Il y alors quatre cas à traiter :

a) Supposons que m soit une formule atomique. Alors, m' commence par le premier symbole de m qui n'est ni une parenthèse, ni le symbole $\neg$, ni un symbole de quantificateur. Si m' est une formule, c'est donc une formule atomique. Mais un segment initial strict d'une formule atomique n'est pas une formule atomique, contradiction.

b) Supposons que m commence par le symbole $\neg$. Il en est de même de m' ; puisque m' est une formule, il existe une formule m'_1 telle que $m' = \neg m'_1$; alors, m'_1 est un segment initial de m_1, distinct de m_1 ; contradiction.

c) Supposons que m commence par une parenthèse ouvrante. D'après le lemme précédent, on a alors $p'(m') \geqslant 1$, ce qui contredit l'hypothèse que m' est une formule.

d) Supposons que le premier symbole de m soit un quantificateur, disons $\exists$ pour fixer les idées. Il en est de même du premier symbole de m' ; par suite, $\ell(m') \geqslant 2$ et il existe une formule m'_1 telle que $m' = \exists v m'_1$, où v est le second symbole de m. On observe que m'_1 est un segment initial de la formule m_1, distinct de m_1, ce qui est absurde.

□

Théorème **(2.3.8)**. — *Soit $m \in \mathscr{F}_{\mathrm{L}}$ une formule dans le langage* L. *Alors, une, et une seule, des propositions suivantes est satisfaite :*

a) *m est une formule atomique ;*

b) *Il existe une unique formule m' telle que $m = \neg m'$;*

c) *Il existe un unique couple (m', m'') de formules et un unique élément $\bullet \in \{\wedge, \vee, \Rightarrow, \Leftrightarrow\}$ tel que $m = (m' \bullet m'')$;*

d) *Il existe un unique élément $v \in \mathrm{V}$ et une unique formule m' tels que $m = \exists v m'$;*

e) *Il existe un unique élément $v \in \mathrm{V}$ et une unique formule m' tels que $m = \forall v m'$.*

Démonstration. — Le mot vide n'est pas une formule ; par suite, les cinq cas cités s'excluent mutuellement par la nature du premier symbole : de relation dans le cas a), $\neg$ dans le cas b), (dans le cas c), $\exists$ dans le cas d), $\forall$ dans le cas e). Il résulte aussi de la définition d'une formule que l'assertion d'existence d'un des cinq cas cités est vérifiée, il s'agit donc d'établir l'assertion d'unicité dans chacun des cas.

a) L'assertion est vide dans ce cas.

b) Supposons qu'il existe deux formules m' et m'' telles que $m = \neg m' = \neg m''$. Alors, $m' = m''$.

c) Supposons que l'on ait $m = (m' \bullet m'') = (m'_1 \square m''_1)$, et prouvons que $\bullet = \square$, $m' = m'_1$ et $m'' = m''_1$. Supposons, pour fixer les idées, que la longueur de m'_1 soit inférieure ou égale à celle de m'. Alors, m'_1 est un segment initial de m' et est une formule ; on a donc $m'_1 = m'$. Considérant le symbole suivant, on a donc $\bullet = \square$, puis $m'') = m''_1)$, d'où $m'' = m''_1$ en simplifiant le dernier symbole.

d) Supposons que l'on ait $m' = \exists v m' = \exists v_1 m'_1$. En comparant les deuxièmes symboles, on a $v = v_1$. En simplifiant les deux premiers symboles, il vient alors $m' = m'_1$.

e) Ce cas est analogue au précédent. □

2.3.9. — Soit m une formule dans le langage L. On définit l'ensemble $\mathrm{sf}(m)$ des *sous-formules* de m par récurrence sur la définition de m :

– Si m est une formule atomique, alors $\mathrm{sf}(m) = \{m\}$;
– Si $m = \neg m'$, alors $\mathrm{sf}(m) = \mathrm{sf}(m') \cup \{m\}$;
– Si $m = (m' \bullet m'')$, alors $\mathrm{sf}(m) = \mathrm{sf}(m') \cup \mathrm{sf}(m'') \cup \{m\}$;
– Si $m = \forall v m'$ ou $m = \exists v m'$, alors $\mathrm{sf}(m) = \mathrm{sf}(m') \cup \{m\}$.

L'ensemble $\mathrm{sf}(m)$ est fini ; ses éléments de $\mathrm{sf}(m)$ sont des formules et sont aussi des sous-mots du mot m.

2.4. Variables libres, variables liées, substitution

2.4.1. — Soit m une formule dans le langage L et soit $v \in \mathrm{V}$ un symbole de variable. Les occurrences de v dans le mot m vont être de nature très différentes suivant qu'elles sont liées par un symbole de quantificateur ou *libres*. La définition procède par récurrence sur la définition d'une formule :

– Si m est une formule atomique, toutes les occurrences de v sont libres ;
– Si $m = \neg m'$, les occurrences libres de v dans m sont celles qui sont libres dans m' ;
– Si $m = (m' \bullet m'')$, les occurrences libres de v dans m sont les occurrences libres de v dans m' et les occurrences libres de v dans m'' ;
– Si $m = \exists w m'$ ou $m = \forall w m'$, où $w \in \mathrm{V}$ est un symbole de variable *distinct de v,* les occurrences libres de v dans m sont les occurrences libres de v dans m' ;
– Si $m = \exists v m'$, aucune occurrence de v dans m n'est libre ; les occurrences de v qui sont libres dans m', ainsi que celle en seconde position de m, sont dites *liées par le quantificateur existentiel* $\exists$*,* ou *dans le champ* de ce quantificateur ;
– Si $m = \forall v m'$, aucune occurrence de v dans m n'est libre ; les occurrences de v qui sont libres dans m', ainsi que celle en seconde position de m, sont dites *liées par le quantificateur universel* $\forall$*,* ou *dans le champ* de ce quantificateur.

On vérifie par récurrence que toute occurrence d'un symbole de variable est soit libre, soit liée par une, et une seule, occurrence de quantificateur.

Définition **(2.4.2)**. — *Soit m une formule dans le langage* L *et soit $v \in \mathrm{V}$ un symbole de variable. On dit que v est* libre *dans m si ce symbole admet au moins une occurrence libre.*

Soit W une partie de V ; on note $\mathscr{F}_{\mathrm{L,W}}$ l'ensemble des formules dans le langage L dont les symboles de variables libres sont contenus dans W.

Remarque **(2.4.3)**. — Soit $m = m_1 \dots m_n$ un mot et soit $v \in \mathrm{V}$ un symbole de variable. La définition formelle d'une occurrence de v dans m est un indice i tel que $v = m_i$. Si l'on se tient à cette définition, par exemple si l'on veut programmer il est nécessaire de bien interpréter la définition précédente. Expliquons par exemple le second cas, lorsque $m = \neg m'$. On a $m_1 = \neg$ et $m' = m_2 \dots m_n = m'_1 \dots m'_{n-1}$. Soit j une occurrence de v dans m' ; on a donc $v = m'_j = m_{j+1}$. Ainsi, c'est l'entier $j + 1$ qui est une occurrence de v dans m, et non j.

Définition **(2.4.4)**. — *On dit qu'une formule est* close *si aucune variable n'y est libre.*

Proposition **(2.4.5)**. — *Soit* $s : \mathrm{V} \to \mathscr{T}_{\mathrm{L}}$ *une application. Pour tout symbole de variable* $v \in \mathrm{V}$*, on note* s_v *l'application de* V *dans* $\mathscr{T}_{\mathrm{L}}$ *qui coïncide avec* s *hors de* v *et telle que* $s_v(v) = v$.

Il existe une unique application $\sigma : \mathscr{F}_{\mathrm{L}} \to \mathscr{F}_{\mathrm{L}}$ *vérifiant les propriétés suivantes :*

– *L'application* σ *coïncide avec celle déjà définie sur les formules atomiques ;*

– *Si* $m = \neg m'$*, alors* $\sigma(m) = \neg\sigma(m')$ *;*

– *Si* $m = (m' \bullet m'')$*, alors* $\sigma(m) = (\sigma(m') \bullet \sigma(m'')$ *;*

– *Si* $m = \exists v m'$*, alors* $\sigma(m) = \exists v \sigma_v(m')$*, où* $\sigma_v : \mathscr{F}_{\mathrm{L}} \to \mathscr{F}_{\mathrm{L}}$ *est l'application définie à partir de l'application* s_v *;*

– *Si* $m = \forall v m'$*, alors* $\sigma(m) = \forall v \sigma_v(m')$*, où* $\sigma_v : \mathscr{F}_{\mathrm{L}} \to \mathscr{F}_{\mathrm{L}}$ *est l'application définie à partir de l'application* s_v

On l'appelle l'application de *substitution des occurrences libres* des symboles de variables v par les termes $s(v)$. En effet, si une occurrence de variable v se trouve dans le champ d'un quantificateur, elle figure dans une sous-formule de la forme $\exists v m'$, ou $\forall v m'$, qui est transformée en $\exists v \sigma_v(m')$ ou $\forall v \sigma_v(m')$. Cependant, comme $s_v(v) = v$, σ_v ne modifie jamais le symbole de variable v.

Démonstration. — Il résulte de la définition des formules par récurrence et du théorème de lecture unique, qu'il existe une unique application $\sigma : \mathscr{F}_{\mathrm{L}} \to \mathrm{A}_{\mathrm{L}}^*$ vérifiant ces propriétés. On vérifie ensuite par récurrence que $\sigma(m)$ est une formule, pour toute formule m. □

CHAPITRE 3

RÉALISATIONS, MODÈLES — *INTERPRÉTER*

« D'un fort beau caractère on voit là le modèle, Madame (...) »
— Molière, *Le Misanthrope*, acte v, scène 4

3.1. Structures

Soit $L = ((R_n), (F_n))$ un langage du premier ordre.

Définition **(3.1.1)**. — *On appelle* réalisation du langage L*, ou* L-structure*, un ensemble* A *muni des données suivantes :*

– *Pour tout entier n et tout symbole de relation* $r \in R_n$*, une relation n-aire* r^A *dans le produit cartésien* A^n *;*

– *Pour tout entier n et tout symbole de fonction* $f \in F_n$*, une fonction* $f^A : A^n \to A$.

On exige aussi que les graphes des relations $\top^A$ *et* $\bot^A$ *dans* A^o *soient respectivement* A^o *et* $\varnothing$*, et que le graphe de la relation* $=^A$ *dans* A^2 *soit la diagonale.*

On dit que A est l'*univers* de la réalisation ; pour tout symbole de relation r on dit que la relation r^A est l'interprétation du symbole r dans A ; pour tout symbole de fonction f, on dit que la fonction f^A est l'interprétation du symbole f dans A. On dira souvent « soit A une réalisation... » en sous-entendant les interprétations.

Remarque **(3.1.2)**. — Rappelons que A^o désigne l'ensemble des applications de l'ensemble à o éléments, l'ensemble vide, dans A. Il y a donc exactement un élément dans l'ensemble A^o ; notons-le ε.

Puisque se donner une relation dans A^o revient à se donner une partie de l'ensemble $\{\varepsilon\}$, il y a deux telles relations : $\{\varepsilon\}$, et $\varnothing$. La première est toujours vérifiée et possède la valeur de vérité *vrai*, ou $\top$; la seconde n'est jamais vérifiée et possède la valeur de vérité *faux*, ou $\bot$.

De même, les symboles de constantes sont les éléments de F_o. Se donner une application de A^o dans A revient à se donner l'image dans A de l'élément ε,

autrement dit, à choisir, pour tout symbole de constante $c \in \mathrm{F}_0$, un élément $c^{\mathrm{A}} \in \mathrm{A}$.

3.1.3. — Supposons que le langage L soit celui des anneaux, les symboles de variables étant représentés par des lettres de l'alphabet. Une réalisation de ce langage est un ensemble, disons A, muni de deux lois binaires + et ·, $\mathrm{A} \times \mathrm{A} \to \mathrm{A}$, et de deux éléments 0 et 1. En revanche, on ne fait encore aucune hypothèse sur les relations entre ces lois et les deux éléments 0, 1. En effet, si les axiomes des anneaux s'expriment comme des formules du langage L, l'associativité de l'addition s'écrit $\forall x \forall y \forall z = ++xyz+x+yz$, ou, avec une notation plus courante, $\forall x \forall y \forall z\,((x+y)+z = x+(y+z))$. Mais pour pouvoir dire si l'addition de A est associative, il faut pouvoir *interpréter* les formules et les termes dans A.

3.1.4. — Les deux propositions ci-dessous affirment qu'il y a une unique façon raisonnable d'interpréter termes et formules dans une réalisation donnée A du langage L. Elles sont énoncées de façon un peu exotique. En effet, l'interprétation d'un terme sera une application du produit cartésien A^{V} de la structure dans la structure. Rappelons que A^{V} est l'ensemble des familles $(a_v)_{v\in\mathrm{V}}$ d'éléments de A, indexées par l'ensemble V des symboles de variables ; comme on l'imagine, seuls les indices v apparaissant dans un terme seront pris en compte dans l'interprétation de ce terme. Il est cependant plus simple L'interprétation d'une formule est une application du produit cartésien A^{V} dans l'ensemble $\{\bot, \top\}$ des valeurs de vérité, même si seuls les indices v correspondant aux variables libres d'une formule donnée interviendront pour l'interprétation de cette formule.

Les symboles logiques $\neg, \vee, \wedge, \Rightarrow, \Leftrightarrow$ vont avoir l'interprétation usuelle, donnée par les tables de vérité bien connues suivantes :

$\neg$	$\top$	$\bot$
	$\bot$	$\top$

$\wedge$	$\top$	$\bot$
$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\bot$

$\vee$	$\top$	$\bot$
$\top$	$\top$	$\top$
$\bot$	$\top$	$\bot$

$\Rightarrow$	$\top$	$\bot$
$\top$	$\top$	$\bot$
$\bot$	$\top$	$\top$

$\Leftrightarrow$	$\top$	$\bot$
$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\top$

De même, le quantificateur existentiel $\exists v$ va être interprété comme l'existence d'un élément *dans* A, et le quantificateur universel $\forall v$ va être interprété comme une propriété *pour tout élément de* A.

Proposition **(3.1.5)**. — *Soit* W *une partie de* V. *Il existe une unique application* $\tau_{\mathrm{W}} : \mathscr{T}_{\mathrm{L,W}} \to \mathscr{F}(\mathrm{A}^{\mathrm{W}}, \mathrm{A})$ *de l'ensemble* $\mathscr{T}_{\mathrm{L,W}}$ *des termes dont les symboles de*

variables appartiennent à W *dans l'ensemble des fonctions de* A^{W} *dans* A *vérifiant les propriétés suivantes :*

– *Pour* $v \in \mathrm{W}$, $\tau_{\mathrm{W}}(v)$ *est l'application de* A^{W} *dans* A *donnée par* $(a_v)_{v\in\mathrm{W}} \mapsto a_w$ *;*

– *Pour* $n \in \mathbf{N}$, $f \in \mathrm{F}_n$ *et* $(t_1, \dots, t_n) \in \mathscr{T}_{\mathrm{L,W}}$, $\tau_{\mathrm{W}}(f\,t_1 \dots t_n)$ *est l'application* $a \mapsto f^{\mathrm{A}}(\tau_{\mathrm{W}}(t_1)(a), \dots, \tau_{\mathrm{W}}(t_n)(a))$.

En pratique, l'application $\tau_{\mathrm{W}}(t) : \mathrm{A}^{\mathrm{W}} \to \mathrm{A}$ est notée t^{A}. La formule de récurrence se récrit ainsi $t^{\mathrm{A}}(a) = f^{\mathrm{A}}(t_1^{\mathrm{A}}(a), \dots, t_n^{\mathrm{A}}(a))$, ou encore $t^{\mathrm{A}} = f^{\mathrm{A}} \circ (t_1^{\mathrm{A}}, \dots, t_n^{\mathrm{A}})$. L'ensemble des variables apparaissant dans le terme t est donc implicite et on verra ci-dessous que les ambiguités permises par cet abus de langage sons sans danger.

Démonstration. — Cela découle du théorème de lecture unique des termes. □

Lemme **(3.1.6)**. — *Soit* W *une partie de* V.

a) *Soit* W' *une partie de* W. *Pour tout terme* $t \in \mathscr{T}_{\mathrm{L,W}'}$ *et tout* $a \in \mathrm{A}^{\mathrm{W}}$, *on a* $\tau_{\mathrm{W}}(t)(a) = \tau_{\mathrm{W}'}(t)(a')$, *où* $a' = (a_w)_{w\in\mathrm{W}'}$. *Autrement dit, l'interprétation d'un terme ne dépend que des variables qui y apparaissent effectivement.*

b) *Soit* W' *une partie de* V, *soit* $s : \mathrm{W} \to \mathscr{T}_{\mathrm{L,W}'}$ *une application et soit* $\sigma : \mathscr{T}_{\mathrm{L,W}} \to \mathscr{T}_{\mathrm{L,W}'}$ *l'application de substitution correspondante. Pour tout terme* t *et tout* $a \in \mathrm{A}^{\mathrm{W}'}$, *on a* $\sigma(t)^{\mathrm{A}}(a) = t^{\mathrm{A}}((s(w)^{\mathrm{A}}(a))_{w\in\mathrm{W}})$.

Démonstration. — Cela se prouve par récurrence sur la définition d'un terme. □

Proposition **(3.1.7)**. — *Il existe une unique famille* $(\varphi_{\mathrm{W}})_{\mathrm{W}\in\mathscr{P}(\mathrm{V})}$ *où, pour toute partie* W *de* V, $\varphi_{\mathrm{W}} : \mathscr{F}_{\mathrm{L,W}} \to \mathscr{R}(\mathrm{A}^{\mathrm{W}})$ *est une application de l'ensemble* $\mathscr{F}_{\mathrm{L,W}}$ *des formules du langage* L *à variables libres dans* W *dans l'ensemble* $\mathscr{R}(\mathrm{A}^{\mathrm{W}})$ *des relations sur* A^{W}, *vérifiant les propriétés suivantes :*

a) *Si* $m = r t_1 \dots t_n$, *où* $r \in \mathrm{R}_n$ *et* $t_1, \dots, t_n \in \mathscr{T}_{\mathrm{L,W}}$, *alors* $\varphi_{\mathrm{W}}(m)(a) = r^{\mathrm{A}}(t_1^{\mathrm{A}}(a), \dots, t_n^{\mathrm{A}}(a))$ *pour tout* $a \in \mathrm{A}^{\mathrm{W}}$ *;*

b) *Si* $m = \neg m'$, *où* $m' \in \mathscr{F}_{\mathrm{L,W}}$, *alors* $\varphi_{\mathrm{W}}(m)(a) = \neg\varphi_{\mathrm{W}}(m')(a)$ *pour tout* $a \in \mathrm{A}^{\mathrm{W}}$ *;*

c) *Si* $m = (m' \bullet m'')$, *où* $m', m'' \in \mathscr{F}_{\mathrm{L,W}}$ *et* $\bullet \in \{\vee, \wedge, \Rightarrow, \Leftrightarrow\}$, *alors* $\varphi_{\mathrm{W}}(m)(a) = \varphi_{\mathrm{W}}(m')(a) \bullet \varphi_{\mathrm{W}}(m'')(a)$ *pour tout* $a \in \mathrm{A}^{\mathrm{L,W}}$ *;*

d) *Si* $m = \exists v m'$, *avec* $v \in \mathrm{V}$ *et* $m' \in \mathscr{F}_{\mathrm{L,W}'}$, *où* $\mathrm{W}' = \mathrm{W} \cup \{v\}$, *alors on a* $\varphi_{\mathrm{W}}(m)(a)$ *si et seulement s'il existe un élément* $a' \in \mathrm{A}^{\mathrm{W}'}$ *tel que* $a'_w = a_w$ *pour* $w \neq v$ *tel que* $\varphi_{\mathrm{W}'}(m')(a')$ *;*

e) *Supposons* $m = \forall v m'$, *avec* $v \in \mathrm{V}$ *et* $m' \in \mathscr{F}_{\mathrm{L},\mathrm{W}'}$, *où* $\mathrm{W}' = \mathrm{W} \cup \{v\}$, *alors* $\varphi_{\mathrm{W}}(m)(a)$ *si et seulement si, pour tout élément* $a' \in \mathrm{A}^{\mathrm{W}'}$ *tel que* $a'_w = a_w$ *pour* $w \neq v$, *on a* $\varphi(m')(a')$.

En pratique, la relation $\varphi_{\mathrm{W}}(m)$ dans A^{W} est notée m^{A}, l'ensemble W de variables libres étant sous-entendu. On verra plus loin que cela ne crée pas d'ambiguité.

Démonstration. — Cela se démontre par récurrence sur la définition d'une formule. D'après le théorème de lecture unique, un et un seul des cinq cas indiqué se produit, mais il faut vérifier que les ensembles de variables libres qui apparaissent sont bien comme indiqué, ce qui résulte de la définition par récurrence de l'ensemble des occurrences libres d'un symbole de variables. Supposons $m \in \mathscr{F}_{\mathrm{L},\mathrm{W}}$.

a) Si $m = r t_1 \dots, t_n$ est une formule atomique, ses variables libres sont les symboles variables libres apparaissant dans $t_1, \dots, t_n$; par suite, alors $t_1, \dots, t_n \in \mathscr{T}_{\mathrm{L},\mathrm{W}}$;

b) Si $m = \neg m'$, les variables libres de m sont celles de m' ;

c) Si $m = (m' \bullet m'')$, les variables libres de m sont celles de m' et de m''. Dans ces deux derniers cas, on a donc $m', m'' \in \mathscr{F}_{\mathrm{L},\mathrm{W}}$;

d) Si $m = \exists v m'$ ou $m = \forall v m'$, les variables libres de m sont celles de m', privées de v ; alors, $m' \in \mathscr{F}_{\mathrm{L},\mathrm{W}'}$, où $\mathrm{W}' = \mathrm{W} \cup \{v\}$. □

Lemme **(3.1.8)**. — *Soit* W, W′ *des parties de* V *telles que* $\mathrm{W}' \subset \mathrm{W}$ *et soit* $m \in \mathscr{F}_{\mathrm{L},\mathrm{W}'}$, *de sorte que* $m \in \mathscr{F}_{\mathrm{L},\mathrm{W}}$. *Pour tout* $a \in \mathrm{A}^{\mathrm{W}}$, *on a* $\varphi_{\mathrm{W}}(m)(a) = \varphi_{\mathrm{W}'}(a')$, *où* $a' = (a_w)_{w \in \mathrm{W}'}$. *Autrement dit, la relation* m^{A} *dans* A^{W} *ne dépend pas des composantes dont l'indice n'a pas d'occurrence libre dans* W.

Démonstration. — Cela se prouve par récurrence sur la définition d'une formule. □

3.1.8.1. — La substitution d'une variable libre par un terme n'est pas toujours compatible à l'interprétation d'une formule. Prenons pour exemple la formule $\exists x(x^2 = y)$. Substituons d'abord à y le terme $t = y + z$; on obtient la formule $\exists x(x^2 = y + z)$ dont l'interprétation en un couple (b, c) coïncide avec l'interprétation de la formule initiale en $b + c$. Substituons maintenant à y le terme $u = x + y + z$; on obtient la formule $\exists x(x^2 = x + y + z)$ dont l'interprétation

en un couple (b, c) est vraie si et seulement si l'équation $x^2 - x = b + c$ a une solution, et non l'équation $x^2 = b + c$.

On dira qu'une formule m est *admissible* pour une substitution $s : \mathrm{W} \to \mathscr{T}_{\mathrm{L}}$ si pour toute sous-formule de m de la forme $\mathrm{Q}vm'$ où $\mathrm{Q} \in \{\forall, \exists\}$ et tout symbole de variable $w \in \mathrm{W} - \{v\}$ qui a une occurrence libre dans m', le terme $s(w)$ ne fait pas intervenir le symbole v.

Lemme **(3.1.9)**. — *Soit* W *une partie de* V. *Soit* $s : \mathrm{W} \to \mathscr{T}_{\mathrm{L},\mathrm{W}'}$ *une application et soit* $\sigma : \mathscr{F}_{\mathrm{L},\mathrm{W}} \to \mathscr{F}_{\mathrm{L},\mathrm{W}'}$ *l'application de substitution correspondante. Pour tout* $a \in \mathrm{A}^{\mathrm{W}'}$, *notons* $s^{\mathrm{A}}(a) = (s(w)^{\mathrm{A}}(a))_{w\in\mathrm{W}}$; *pour tout* $a \in \mathrm{A}^{\mathrm{W}'}$ *et toute formule* $m \in \mathscr{F}_{\mathrm{L},\mathrm{W}}$ *qui est admissible pour la substitution* s, *on a* $\sigma(m)^{\mathrm{A}}(a) = m^{\mathrm{A}}(s^{\mathrm{A}}(a)))$.

Démonstration. — Raisonnons par récurrence sur la définition d'une formule. Pour simplifier, nous poserons $b = s^{\mathrm{A}}(a)$.

a) Supposons que m soit une formule atomique, $m = rt_1 \dots, t_n$. Alors $\sigma(m) = ru_1 \dots u_n$, où $u_i = \sigma(t_i) \in \mathscr{T}_{\mathrm{L},\mathrm{W}}$ pour tout i. On a donc $\sigma(m)^{\mathrm{A}}(a) = r^{\mathrm{A}}(u_1^{\mathrm{A}}(a), \dots, u_n^{\mathrm{A}}(a))$. Pour tout $i \in \{1, \dots, n\}$, on a $u_i^{\mathrm{A}}(a) = \sigma(t_i)^{\mathrm{A}}(a) = t_i^{\mathrm{A}}(s^{\mathrm{A}}(a)) = t_i^{\mathrm{A}}(b)$, d'après le lemme 3.1.6. Par suite, on a

$$\sigma(m)^{\mathrm{A}}(a) = r^{\mathrm{A}}(t_1^{\mathrm{A}}(b), \dots, t_n^{\mathrm{A}}(b)) = m^{\mathrm{A}}(b) = m^{\mathrm{A}}(s^{\mathrm{A}}(a)),$$

ce qu'il fallait démontrer.

b) Supposons que $m = \neg m'$, où $m' \in \mathscr{F}_{\mathrm{L},\mathrm{W}}$. Alors, $\sigma(m) = \neg\sigma(m')$, donc $\sigma(m)^{\mathrm{A}}(a) = \neg\sigma(m')^{\mathrm{A}}(a)$. Par récurrence, $\sigma(m')^{\mathrm{A}}(a) = (m')^{\mathrm{A}}(b)$, d'où $\sigma(m)^{\mathrm{A}}(a) = \neg(m')^{\mathrm{A}}(b) = m^{\mathrm{A}}(b)$.

c) On raisonne de façon analogue lorsque m est de la forme $(m' \bullet m'')$, avec $\bullet \in \{\vee, \wedge, \Rightarrow, \Leftrightarrow\}$.

d) Supposons que $m = \exists v m'$; posons $\mathrm{W}_0 = \mathrm{W} - \{v\}$, $\mathrm{W}_1 = \mathrm{W} \cup \{v\} = \mathrm{W}_0 \cup \{v\}$, de sorte que $m' \in \mathscr{F}_{\mathrm{L},\mathrm{W}_1}$ et $m \in \mathscr{F}_{\mathrm{L},\mathrm{W}_0}$.

Posons $\mathrm{W}_1' = \mathrm{W}' \cup \{v\}$ et soit $s_1 : \mathrm{W}_1 \to \mathscr{T}_{\mathrm{L},\mathrm{W}_1'}$ l'application qui coïncide avec s sur $\mathrm{W}_0 - \{v\}$ et telle que $s_1(v) = v$; soit σ_1 l'application de substitution correspondante. Par définition, on a $\sigma(m) = \exists v \sigma_1(m')$,

Considérons une sous-formule de m' de la forme $\mathrm{Q}xp$, où $\mathrm{Q} \in \{\forall, \exists\}$. Comme c'est une sous-formule de m, pour tout symbole de variable $w \in \mathrm{W}$ qui a une occurrence libre dans p, le terme $s(w)$ ne fait pas intervenir le symbole x. Soit $w \in \mathrm{W}_1$ un symbole de variable qui a une occurrence libre dans p et tel que $w \neq x$; on a $s_1(w) = s(w)$, donc $s_1(w)$ ne fait pas intervenir x. Cela prouve que la formule m' est admissible pour la substitution s_1.

L'interprétation $\sigma_1(m')^{\mathrm{A}}$ est une relation dans $\mathrm{A}^{\mathrm{W}'_1}$; l'interprétation $(m')^{\mathrm{A}}$ est une relation dans $\mathrm{A}^{\mathrm{W}_1}$. Par récurrence, on a donc $\sigma_1(m')^{\mathrm{A}}(a_1) = (m')^{\mathrm{A}}((s_1(w)^{\mathrm{A}}(a_1))_{w\in \mathrm{W}_1})$ pour tout $a_1 \in \mathrm{A}^{\mathrm{W}'_1}$.

Soit $a \in \mathrm{A}^{\mathrm{W}'}$. Puisque $\sigma(m) = \exists v \sigma(m')$, on a $\sigma(m)^{\mathrm{A}}(a) = \top$ si et seulement s'il existe $a_1 \in \mathrm{A}^{\mathrm{W}'_1}$ tel que $\sigma(m')^{\mathrm{A}}(a_1) = \top$ et tel que toutes les composantes de a_1 autres que celle d'indice v coïncident avec celles de a. (On dira qu'un tel a_1 est v-proche de a.)

Soit $w \in \mathrm{W}_1$ un symbole de variable qui a une occurrence libre dans m' ; par hypothèse, le symbole de variable v n'apparaît pas dans le terme $s_1(w) = s(w)$; on a donc $s_1(w)^{\mathrm{A}}(a_1) = b_w$ pour tout élément convenable a_1. D'autre part, $s_1(v) = v$, donc $s_1(v)^{\mathrm{A}}(a_1)$ est la composante d'indice v de a_1. Ainsi, l'application $a_1 \mapsto (s_1(w)^{\mathrm{A}}(a_1))_{w\in \mathrm{W}_1}$ induit une bijection de l'ensemble des éléments $a_1 \in \mathrm{A}^{\mathrm{W}'_1}$ qui sont v-proches de a sur l'ensemble des éléments $b_1 \in \mathrm{A}^{\mathrm{W}_1}$ qui sont v-proches de b.

Cela démontre que $\sigma(m)^{\mathrm{A}}(a) = \top$ si et seulement s'il existe un élément b_1 qui est v-proche de b tel que $(m')^{\mathrm{A}}(b) = \top$, ce qui signifie exactement $m^{\mathrm{A}}(b) = \top$.

e) Le cas du quantificateur $\forall$ est analogue.

□

3.1.10. — On a l'habitude, lorsque P est un polynôme à coefficients dans un anneau commutatif A en des indéterminées $\mathrm{T}_1, \ldots, \mathrm{T}_n$, de noter $\mathrm{P}(\mathrm{T}_1, \ldots, \mathrm{T}_n)$ le polynôme P. On note aussi couramment $\mathrm{P}(a_1, \ldots, a_n)$ l'évaluation du polynôme P en un n-uplet $(a_1, \ldots, a_n) \in \mathrm{A}^n$.

Nous allons mettre en place cette notation dans le contexte de l'évaluation des termes et des formules dans un langage L. Soit A une réalisation de L, soit $t \in \mathscr{T}_{\mathrm{L}}$ un terme et soit $m \in \mathscr{T}_{\mathrm{L}}$ une formule dans le langage L.

Par analogie avec les polynômes, on notera aussi $t(x_1, \ldots, x_n)$ le terme t lorsque $(x_1, \ldots, x_n)$ une suite de symboles de variables *deux à deux distincts* contenant les symboles de variables qui apparaissent dans t. Si $(a_1, \ldots, a_n) \in \mathrm{A}^n$, on notera $t^{\mathrm{A}}(a_1, \ldots, a_n)$ l'évaluation du terme t en remplaçant x_i par a_i.

De même, si m est une formule dans le langage L, on écrira $m(x_1, \ldots, x_n)$ lorsque $(x_1, \ldots, x_n)$ est une suite de symboles de variables *deux à deux distincts* contenant les symboles de *variables libres* qui apparaissent dans t. Si $(a_1, \ldots, a_n) \in \mathrm{A}^n$, on notera $m^{\mathrm{A}}(a_1, \ldots, a_n)$ l'évaluation de la formule m en remplaçant x_i par a_i.

On notera enfin $\mathrm{A} \vDash m(a_1, \ldots, a_n)$ au lieu de $m^{\mathrm{A}}(a_1, \ldots, a_n) = \top$.

En particulier, lorsque m est une formule close, son interprétation m^{A} est une relation dans l'ensemble à un élément $\mathrm{A}^{\varnothing}$, c'est-à-dire tout ou rien. Lorsque $m^{\mathrm{A}} = \top$, on dira que m est *satisfaite* dans A et on écrira $\mathrm{A} \vDash m$.

3.2. Équivalence sémantique

Définition **(3.2.1)**. — *Soit* L *un langage. On dit que deux formules dans le langage* L *sont* sémantiquement équivalentes *si leurs interprétations sont égales dans toute réalisation du langage* L.

Plus explicitement, deux formules m et m' dans le langage L sont équivalentes si pour toute réalisation A de L, les fonctions m^{A} et $(m')^{\mathrm{A}}$ de A^{V} dans $\{\bot, \top\}$ sont égales. C'est bien sûr une relation d'équivalence.

On définit aussi la relation de quasi-équivalence en n'exigeant l'égalité des interprétations que dans toute réalisation *non vide*.

Exemple **(3.2.2)**. — Soit m, m', m'' des formules et soit x, y des symboles de variables. Les formules suivantes sont sémantiquement équivalentes :

a) m et $\neg\neg m$;
b) $\neg(m \vee m')$ et $(\neg m \wedge \neg m')$;
c) $(m \vee (m' \vee m''))$ et $((m \vee m') \vee m'')$;
d) $(m \wedge (m' \wedge m''))$ et $((m \wedge m') \wedge m'')$;
e) m et $(m \wedge m)$;
f) m et $(m \vee m)$;
g) $m \Rightarrow m'$ et $(\neg m \vee m')$;
h) $(m \vee \neg m)$ et $\top$;
i) $\neg\exists x m$ et $\forall x \neg m$;
j) $\forall x(m \wedge m')$ et $(\forall x m \wedge \forall x m')$.

Lemme **(3.2.3)**. — *Soit m, m', p, p' des formules ; soit x un symbole de variable. On suppose que m et m' d'une part, p et p' d'autre part, sont sémantiquement équivalentes. Alors les formules suivantes sont sémantiquement équivalentes :*

a) $\neg m$ *et* $\neg m'$;
b) $(m \bullet p)$ *et* $(m' \bullet p')$, *pour tout symbole* $\bullet \in \{\wedge, \vee, \Rightarrow, \Leftrightarrow\}$;
c) $\exists x m$ *et* $\exists x m'$;
d) $\forall x m$ *et* $\forall x m'$.

Exemple **(3.2.4)**. — Soit $\bullet$ un élément de $\{\vee, \wedge\}$, soit Q un élément de $\{\forall, \exists\}$. Toute formule est sémantiquement équivalente à une formule où les seuls symboles logiques apparaissant sont $\neg$, $\bullet$, Q.

Cela se démontre par récurrence sur la définition d'une formule. Traitons par exemple le cas où $\bullet = \vee$ et $Q = \exists$. On remplace alors $(m \wedge m')$ par $\neg(\neg m \vee \neg m')$, $(m \Rightarrow m')$ par $(\neg m \vee m')$, $(m \Leftrightarrow m')$ par $\neg(\neg(\neg m \vee m') \vee \neg(\neg m' \vee m))$, et $\forall x m$ par $\neg\exists x \neg m$.

3.2.5. — On dit qu'une formule m est sans quantificateurs si elle n'admet aucune occurrence de quantificateur. On dit qu'elle est en *forme prenex* s'il existe une formule sans quantificateurs m', un entier n, des symboles de variables $v_1, \ldots, v_n$ et des quantificateurs $Q_1, \ldots, Q_n$ tels que $m = Q_1 v_1 \ldots Q_n v_n m'$. Une formule prenex est dite *universelle* si elle n'admet pas d'occurrence de quantificateur existentiel $(\exists)$; elle est dite *existentielle* si elle n'admet pas d'occurrence de quantificateur universel $(\forall)$.

Proposition **(3.2.6)**. — *Toute formule m est quasi-équivalente à une formule prenex.*

Démonstration. — On commence par se ramener au cas où les seuls symboles logiques qu'elle contient sont $\neg$ et $\vee$.

On suppose aussi que chaque symbole de variable liée n'est lié que par un seul symbole quantificateur, et qu'aucun symbole de variable liée n'a d'occurrence libre. Que cela soit possible se vérifie par récurrence. Par exemple, si $m = Qxp$, où Q est un symbole de quantificateur et p une formule dans laquelle x possède une occurence liée, on substitue les occurences libres de x dans m par un symbole de variable y qui n'apparaît pas dans p. De même, si $m = (p \bullet q)$, on s'est assuré par récurrence que les symboles de variables qui sont liés dans p n'apparaissent pas dans q, et inversement.

La démonstration procède ensuite par récurrence sur la définition d'une formule.

Si Q est un symbole de quantificateur, on notera Q^* le quantificateur dual, de sorte que $\forall^* = \exists$ et $\exists^* = \forall$.

– Si $m = \neg p$, où p est une formule prenex, on écrit $p = Q_1 v_1 \ldots Q_n v_n p'$, où p' est sans quantificateur ; alors, $\neg p$ est équivalente à $Q_1^* v_1 \ldots Q_n p * v_n \neg p'$.

– Supposons $m = (p \vee q)$, où p et q sont des formules prenex; on écrit alors $p = Q_1 v_1 \dots Q_n v_n p'$ et $q = R_1 w_1 \dots R_k w_k q'$, où p' et q' sont sans quantificateurs; en outre, $v_1, \dots, v_n$ n'apparaissent pas dans q, et $w_1, \dots, w_k$ n'apparaissent pas dans p. Alors, $m = (p \vee q)$ est quasi-équivalente à $Q_1 v_1 \dots Q_n v_n R_1 w_1 \dots R_k w_k (p' \vee q')$. Cela se démontre par récurrence sur $n + k$, variable après variable, grâce aux deux remarques :

a) Les formules $(\forall x p_1 \vee q_1)$ et $\forall x (p_1 \vee q_1)$ sont sémantiquement équivalentes;

b) Les formules $(\exists x p_1 \vee q_1)$ et $\exists x (p_1 \vee q_1)$ sont quasi-équivalentes. Elles ne sont par contre pas forcément équivalentes : dans la structure vide, l'interprétation de $\exists x (p_1 \vee q_1)$ ne prend que la valeur $\bot$, tandis que celle de $(\exists x p_1 \vee q_1)$ coïncide avec celle de q_1. □

3.3. Morphismes de structures

Définition **(3.3.1)**. — *Soit* L *un langage, soit* A *et* A′ *des réalisations de* L *et soit* $\varphi : A' \to A$ *une application.*

On dit que φ *est un* morphisme *de structures si les conditions suivantes sont satisfaites :*

(i) *Pour tout symbole* $f \in F_n$ *de fonction n-aire et tout élément* $(a_1, \dots, a_n) \in (A')^n$*, on a*

$$\varphi\big(f^{A'}(a_1, \dots, a_n)\big) = f^{A}(\varphi(a_1), \dots, \varphi(a_n));$$

(ii) *Pour tout symbole* $r \in R_n$ *de relation n-aire et tout élément* $(a_1, \dots, a_n) \in (A')^n$ *tel que* $r^{A'}(a_1, \dots, a_n)$*, on a* $r^{A}(\varphi(a_1), \dots, \varphi(a_n))$.

On dit que φ *est un* monomorphisme *de structures, ou un* plongement*, si l'on a de plus*

(ii′) *Pour tout symbole* $r \in R_n$ *de relation n-aire et tout élément* $(a_1, \dots, a_n) \in (A')^n$*, alors*

$$r^{A'}(a_1, \dots, a_n) \quad \Leftrightarrow \quad r^{A}(\varphi(a_1), \dots, \varphi(a_n)).$$

Comme le langage L comporte un symbole de relation binaire qui est interprété comme l'égalité dans A, *un monomorphisme de structures est injectif.*

Exemple **(3.3.2)**. — Le langage des groupes possède deux symboles de fonctions, $e \in \mathscr{F}_0$ représente l'élément neutre, et $\cdot \in \mathscr{F}_2$ la loi de composition, et n'a comme

symbole de relation que l'égalité. Une réalisation de ce langage est un magma « pointé », c'est-à-dire muni d'un élément privilégié. Un homomorphisme de structures est un morphisme de magmas pointés ; un monomomorphisme est un morphisme injectif.

3.3.3. — Le composé de deux morphismes (resp. monomorphismes) est un morphisme (resp. un monomorphisme) ; l'identité Id_{A} est un monomorphisme.

On dit qu'un morphisme $\varphi : \mathrm{A}' \to \mathrm{A}$ est un isomorphisme s'il existe un morphisme $\psi : \mathrm{A} \to \mathrm{A}'$ tel que $\psi \circ \varphi = \mathrm{Id}_{\mathrm{A}'}$ et $\varphi \circ \psi = \mathrm{Id}_{\mathrm{A}}$. Cela revient à dire que φ est un monomorphisme bijectif.

3.3.4. — Soit A et A′ des réalisations du langage L. Lorsque A′ est une partie de A et que l'inclusion de A′ dans A est un monomorphisme de structures, on dit que A′ *est une sous-structure de* A.

Inversement, soit A une réalisation du langage L et soit A′ une partie de A vérifiant la propriété suivante : *Pour tout symbole* $f \in \mathrm{F}_n$ *et tout* $(a_1, \dots, a_n) \in (\mathrm{A}')^n$, *on a* $f^{\mathrm{A}}(a_1, \dots, a_n) \in \mathrm{A}'$. Alors, il existe une interprétation du langage L dans A′ de sorte que l'inclusion de A′ dans A soit un plongement de structures.

En effet, il suffit de poser $f^{\mathrm{A}'}(a_1, \dots, a_n) = f^{\mathrm{A}}(a_1, \dots, a_n)$ et $r^{\mathrm{A}'}(a_1, \dots, a_n) = r^{\mathrm{A}}(a_1, \dots, a_n)$ pour tout $f \in \mathrm{F}_n$, tout $r \in \mathrm{R}_n$ et tout $(a_1, \dots, a_n) \in (\mathrm{A}')^n$.

3.3.5. — Soit A une réalisation du langage L. L'intersection d'une famille (A_i) de sous-structures de A est une sous-structure de A.

Pour toute partie S de A, il existe donc une plus petite sous-structure de A qui contient S, à savoir l'intersection de toutes d'entre elles ; on dit que c'est la sous-structure engendrée par A.

Lemme **(3.3.6)**. — *Soit* L *un langage et soit* A *une réalisation de* L. *Soit* S *une partie de* A. *La sous-structure de* A *engendrée par* S *est l'ensemble des éléments de* A *de la forme* $t^{\mathrm{A}}(a_1, \dots, a_n)$, *où* $n \in \mathbf{N}$, $t \in \mathscr{T}_{\mathrm{L}}$ *est un terme en les symboles de variables* $x_1, \dots, x_n$, *et* $(a_1, \dots, a_n) \in \mathrm{S}^n$.

Démonstration. — Soit A′ l'ensemble des éléments indiqués et soit B la sous-structure de A engendrée par S. Pour tout $n \in \mathbf{N}$, tout terme $t(x_1, \dots, x_n)$ et tout $(a_1, \dots, a_n) \in \mathrm{S}^n$, on a donc $t^{\mathrm{A}}(a_1, \dots, a_n) = t^{\mathrm{B}}(a_1, \dots, a_n) \in \mathrm{B}$. Cela prouve que $\mathrm{A}' \subset \mathrm{B}$. Inversement, nous allons démontrer que A′ est une sous-structure de A qui contient S. Soit x un symbole de variable et soit t le terme x ; on a $t^{\mathrm{A}}(a) = a$ pour tout $a \in \mathrm{S}$, donc A′ contient S. Soit $n \in \mathbf{N}$, soit $f \in \mathrm{F}_n$ un

symbole de fonction n-aire et soit $a_1, \dots, a_n$ des éléments de A' ; démontrons que $f^{\mathrm{A}}(a_1, \dots, a_n)$ appartient à A'. Pour $i \in \{1, \dots, n\}$, soit t_i un terme en des symboles de variables $x_{i,j}$ (pour $1 \leqslant j \leqslant m_i$) et soit $(a_{i,j})_{1 \leqslant j \leqslant m_i}$ un élément de A^{m_i} tel que $a_i = t_i^{\mathrm{A}}(a_{i,1}, \dots, a_{i,m_i})$. Comme l'ensemble de symboles de variables est infini, on se ramène d'abord au cas où les symboles $x_{i,j}$ sont deux à deux distincts : choisir une famille $(y_{i,j})$ de symboles deux à deux distincts, puis, pour tout i, et substituer $y_{i,j}$ à $x_{i,j}$ dans t_i. Soit aussi $x_1, \dots, x_n$ des symboles de variables deux à deux distincts et soit t le terme obtenu en substituant dans le terme $f x_1 \dots x_n$ le symbole de variable x_i par le terme t_i. On a

$$f^{\mathrm{A}}(a_1, \dots, a_n) = f^{\mathrm{A}}(t_1^{\mathrm{A}}(a_{1,1}, \dots, a_{1,m_1}), \dots, t_n^{\mathrm{A}}(a_{n,1}, \dots)) = t^{\mathrm{A}}((a_{i,j})),$$

ce qui prouve que $f^{\mathrm{A}}(a_1, \dots, a_n)$ appartient à A'. Cela conclut la démonstration du lemme. □

3.4. Extension de langages

3.4.1. — Soit $\mathrm{L} = ((\mathrm{R}_n), (\mathrm{F}_n))$ et $\mathrm{L}' = ((\mathrm{R}'_n), (\mathrm{F}'_n))$ des langages. On dit que le langage L' étend le langage L, ou que le langage L réduit le langage L' si pour tout entier n, on a $\mathrm{R}_n \subset \mathrm{R}'_n$ et $\mathrm{F}_n \subset \mathrm{F}'_n$.

Dans ce cas, toute réalisation A' du langage L' donne lieu à une réalisation A du langage L : on conserve l'univers en posant $\mathrm{A} = \mathrm{A}'$, et on choisit pour interprétations f^{A} et r^{A} des symboles de fonctions $f \in \mathrm{F}_n$ et de relations $r \in \mathrm{R}_n$ les interprétations de ces symboles dans la réalisation A'.

Exemple **(3.4.2)**. — Soit L un langage, soit A une réalisation de L. Soit S une partie de A. On définit un langage L_{S} en ajoutant aux symboles de constantes de L un symbole c_s pour chaque élément $s \in \mathrm{S}$. Un terme ou une formule du langage L_{S} est aussi appelé un terme ou une formule du langage L *à paramètres dans* S.

Le langage L_{S} possède une réalisation naturelle dans l'univers A, où, pour tout $s \in \mathrm{S}$, l'interprétation c_s^{A} du symbole de constante c_s est prise égale à s.

Soit $m(x_1, \dots, x_n) \in \mathscr{F}_{\mathrm{L}}$ une formule dans le langage L en des variables libres $x_1, \dots, x_n$. Supposons qu'il existe $(a_1, \dots, a_n) \in \mathrm{S}^n$ tel que $\mathrm{A} \vDash m(a_1, \dots, a_n)$, c'est-à-dire $m^{\mathrm{A}}(a_1, \dots, a_n) = \top$. Dans la formule m, substituons simultanément le symbole de variable libre x_i par le symbole de constante c_{a_i}, considéré comme

un terme du langage $\mathrm{L_S}$; la formule obtenue, naturellement notée $m(c_{a_1}, \ldots, c_{a_n})$, est une formule close de $\mathrm{L_S}$. Elle est satisfaite dans la réalisation A.

Exemple **(3.4.3)**. — Soit L un langage et soit A une réalisation de L. Soit x, y deux symboles de variables (distincts) et soit m une formule dont les symboles de variables libres appartiennent à $\{x, y\}$. Soit aussi L′ le langage obtenu en adjoignant à L un (nouveau) symbole de fonction 1-aire que l'on note f_m.

Supposons que la formule close $\forall x \exists y m(x, y)$ soit satisfaite dans A. Cela signifie que pour tout $a \in \mathrm{A}$, il existe un élément $b \in \mathrm{A}$ tel que $m^{\mathrm{A}}(a, b)$. Pour tout $a \in \mathrm{A}$, *choisissons* un tel élément et notons-le $f_m^{\mathrm{A}}(a)$.

On a ainsi promu l'ensemble A de réalisation de L en une réalisation de L′. Pour tout ensemble de symboles de variables W, les formules de L ne changent pas d'interprétation en tant que formule de L′. De plus, le mot $\forall x m(x, f_m(x))$ est une formule close du langage L′ qui est satisfaite dans A.

Inversement, soit A′ une réalisation de L′ dans laquelle la formule $\forall x m(x, f_m(x))$ est satisfaite. En oubliant l'interprétation du symbole de fonction f_m, l'ensemble A est une réalisation de L dans laquelle la formule close $\forall x \exists y m(x, y)$ est satisfaite.

Ce procédé est appelé *skolemisation*. Il se généralise évidemment d'une formule à un ensemble arbitraire de formules.

3.5. Théories, modèles

3.5.1. — Soit L un langage. On appelle *théorie* un ensemble de formules closes dans le langage L.

Soit T une théorie dans le langage L. On dit qu'une réalisation A du langage L est un *modèle* de la théorie T si toute formule appartenant à T est satisfaite dans A. On dit que la théorie T est *consistante* si elle possède un modèle.

Exemple **(3.5.2)**. — Soit L un langage et soit A une réalisation de L. La théorie de A dans le langage L, notée $\mathrm{Th_L(A)}$, est l'ensemble de toutes les formules closes de L qui sont réalisées dans L. Une théorie de cette forme est consistante ; on dit qu'elle est *complète*.

Si A est un modèle d'une théorie T, on a $\mathrm{T} \subset \mathrm{Th_L(A)}$. Toute théorie consistante T est donc contenue dans une théorie complète.

Remarque **(3.5.3)**. — Pour qu'une théorie consistante T soit complète, il faut et il suffit qu'elle soit maximale.

Supposons d'abord que T est complète et soit A un modèle de T telle que $T = \mathrm{Th}_L(A)$. Soit m une formule de L qui n'appartient pas à T, c'est-à-dire $m^A = \bot$; alors, $(\neg m)^A = \top$, donc $\neg m \in T$. Si B est un modèle de T, on a $(\neg m)^B = \top$, donc $m^B = \bot$, d'où $m \notin \mathrm{Th}_L(B)$. Cela prouve que $T \cup \{m\}$ n'est pas consistant, si bien que T est une théorie maximale.

Soit T une théorie consistante maximale et soit A un modèle de T. On a $T \subset \mathrm{Th}_L(A)$. Puisque T est maximale, on a $T = \mathrm{Th}_L(A)$: la théorie T est complète.

3.5.4. — Soit T une théorie dans le langage L. On dit qu'une formule close m est *conséquence sémantique* de T si m satsisfaite dans tout modèle de T. On dit qu'une théorie T′ est conséquence sémantique de T si tout modèle de T est modèle de T′.

On dit qu'une théorie est *finiment axiomatisable* si elle est conséquence sémantique d'un ensemble fini de formules closes.

3.6. Extensions élémentaires, théorème de Löwenheim–Skolem

Définition **(3.6.1)**. — *Soit* L *un langage, soit* A *une réalisation de* L *et soit* B *une sous-structure de* A. *On dit que* B *est une* sous-structure élémentaire *de* A, *ou que* A *est une* extension élémentaire *de* B, *si pour tout entier n, toute formule* $m(x_1, \ldots, x_n)$ *dans le langage* L *et tout élément* $(b_1, \ldots, b_n) \in B^n$, *on a* $m^A(b_1, \ldots, b_n) = m^B(b_1, \ldots, b_n)$.

On note $B \prec A$.

Remarque **(3.6.2)**. — En appliquant la définition à une formule close, on observe que $\mathrm{Th}_L(A) = \mathrm{Th}_L(B)$ si B est une sous-structure élémentaire de A.

Plus généralement, soit L_B le langage du premier ordre obtenu en adjoignant à L un symbole de constante c_b pour chaque élément b de B. Considérons A comme une réalisation de L_B, chaque symbole c_b étant interprété comme l'élément b de B, de sorte que B est une sous-structure. Les formules closes de L_B sont de la forme $m(c_{b_1}, \ldots, c_{b_n})$, où $m(x_1, \ldots, x_n)$ est une formule de L en les variables libres $x_1, \ldots, x_n$.

Cela montre que A est extension élémentaire de B si et seulement si $\mathrm{Th}_{L_B}(A) = \mathrm{Th}_{L_B}(B)$.

Théorème **(3.6.3)** (Test de Tarski-Vaught). — *Soit* A *une réalisation de* L, *soit* B *une partie de* A. *On suppose que pour tout entier n, toute formule* $m(x_1, \ldots, x_n, y)$ *de* L *et tout élément* $(a_1, \ldots, a_n) \in \mathrm{B}^n$ *tel que* $\mathrm{A} \vDash \exists y m(a_1, \ldots, a_n, y)$, *il existe un élément* $b \in \mathrm{B}$ *tel que* $\mathrm{A} \vDash m(a_1, \ldots, a_n, b)$. *Alors,* B *est une sous-structure élémentaire de* A.

Le point important de l'énoncé est de ne faire intervenir la satisfaction que dans la structure A.

Démonstration. — Démontrons que B est une sous-structure de A. Soit $n \in \mathbf{N}$ et soit $f \in \mathrm{F}_n$ un symbole de fonction n-aire, soit $(b_1, \ldots, b_n) \in \mathrm{B}^n$. En appliquant la définition à la formule $\exists y(y = f(x_1, \ldots, x_n))$, on constate que $f^{\mathrm{A}}(b_1, \ldots, b_n) \in \mathrm{B}$. Cela prouve que B est une sous-structure de A.

Prouvons que c'est une sous-structure *élémentaire*. Il s'agit de démontrer que pour toute formule $m(x_1, \ldots, x_n)$ et tout élément $(a_1, \ldots, a_n) \in \mathrm{B}^n$, alors $\mathrm{A} \vDash m(a_1, \ldots, a_n)$ si et seulement si $\mathrm{B} \vDash m(a_1, \ldots, a_n)$.

Si un symbole de variable x_i n'a pas d'occurrence libre dans m, les fonctions $m^{\mathrm{A}} : \mathrm{A}^n \to \{\top, \bot\}$ et $m^{\mathrm{B}} : \mathrm{B}^n \to \{\top, \bot\}$ ne dépendent pas de la coordonnée a_i ; en raisonnant par récurrence sur n, on suppose que tous les symboles de variables $x_1, \ldots, x_n$ ont une occurrence libre dans m.

On raisonne alors par récurrence sur la définition d'une formule :

a) Si m est une formule atomique, cela résulte de la définition d'une sous-structure ;

b) Supposons qu'il existe une formule m' telle que $m = \neg m'$. Par récurrence, on a $(m')^{\mathrm{B}}(a_1, \ldots, a_n) = (m')^{\mathrm{A}}(a_1, \ldots, a_n)$; par suite,

$$m^{\mathrm{B}}(a_1, \ldots, a_n) = \neg(m')^{\mathrm{B}}(a_1, \ldots, a_n) = \neg(m')^{\mathrm{A}}(a_1, \ldots, a_n) = m^{\mathrm{A}}(a_1, \ldots, a_n).$$

c) Supposons qu'il existe des formules m' et m'' et un élément $\bullet \in \{\vee, \wedge, \Rightarrow, \Leftrightarrow\}$ tels que $m = (m' \bullet m'')$. Par récurrence, on a $(m')^{\mathrm{B}}(a_1, \ldots, a_n) = (m')^{\mathrm{A}}(a_1, \ldots, a_n)$ et $(m'')^{\mathrm{B}}(a_1, \ldots, a_n) = (m'')^{\mathrm{A}}(a_1, \ldots, a_n)$; par suite,

$$\begin{aligned} m^{\mathrm{B}}(a_1, \ldots, a_n) &= ((m')^{\mathrm{B}}(a_1, \ldots, a_n) \bullet (m'')^{\mathrm{B}}(a_1, \ldots, a_n)) \\ &= ((m')^{\mathrm{A}}(a_1, \ldots, a_n) \bullet (m'')^{\mathrm{A}}(a_1, \ldots, a_n)) \\ &= m^{\mathrm{A}}(a_1, \ldots, a_n). \end{aligned}$$

d) Supposons qu'il existe une formule m' et un symbole de variable y tel que $m = \exists y m'$. Par définition d'une occurrence libre, y n'est pas une variable

libre de m, si bien que $y \notin \{x_1, \ldots, x_n\}$, de sorte que m' est une formule en les variables libres $x_1, \ldots, x_n, y$.

Supposons $m^{\mathrm{A}}(a_1, \ldots, a_n) = \top$. Par définition, il existe $b \in \mathrm{A}$ tel que $(m')^{\mathrm{A}}(a_1, \ldots, a_n, b)$. Par hypothèse, il existe alors un tel élément b qui appartient à B. Par récurrence, on a alors $(m')^{\mathrm{B}}(a_1, \ldots, a_n, b) = (m')^{\mathrm{A}}(a_1, \ldots, a_n, b) = \top$. Il en résulte que $m^{\mathrm{B}}(a_1, \ldots, a_n) = \top$.

Supposons inversement que $m^{\mathrm{B}}(a_1, \ldots, a_n) = \top$. Il existe donc $b \in \mathrm{B}$ tel que $(m')^{\mathrm{B}}(a_1, \ldots, a_n, b) = \top$. Par récurrence, on a donc $(m')^{\mathrm{A}}(a_1, \ldots, a_n, b) = \top$, ce qui entraîne que $m^{\mathrm{A}}(a_1, \ldots, a_n) = \top$.

e) Supposons enfin qu'il existe une formule m' et un symbole de variable y tel que $m = \forall y m'$. Par définition d'une occurrence libre, y n'est pas une variable libre de m, si bien que $y \notin \{x_1, \ldots, x_n\}$, de sorte que m' est une formule en les variables libres $x_1, \ldots, x_n, y$.

Supposons $m^{\mathrm{A}}(a_1, \ldots, a_n) = \top$. Par définition, pour tout $b \in \mathrm{A}$, on a $(m')^{\mathrm{A}}(a_1, \ldots, a_n, b) = \top$. Pour tout $b \in \mathrm{B}$, on a on a alors $(m')^{\mathrm{B}}(a_1, \ldots, a_n, b) = (m')^{\mathrm{A}}(a_1, \ldots, a_n, b) = \top$, par récurrence. Il en résulte que $m^{\mathrm{B}}(a_1, \ldots, a_n) = \top$.

Supposons que $m^{\mathrm{A}}(a_1, \ldots, a_n) = \bot$. Par définition, il existe $b \in \mathrm{A}$ tel que $(m')^{\mathrm{A}}(a_1, \ldots, a_n, b) = \top$. L'hypothèse du théorème, appliquée à la formule $\neg m'$ entraîne qu'il existe un élément $b \in \mathrm{B}$ tel que $(m')^{\mathrm{A}}(a_1, \ldots, a_n, b) = \bot$. Par récurrence, on a alors $(m')^{\mathrm{B}}(a_1, \ldots, a_n, b) = \bot$, ce qui entraîne que $m^{\mathrm{B}}(a_1, \ldots, a_n) = \bot$.

□

3.6.4. — Soit T une théorie (consistante) dans un langage L, soit A un modèle de T et soit S une partie de A.

Nous allons construire par récurrence une suite croissante (B_p) de sous-structures de A contenant S dont la réunion $\mathrm{B} = \bigcup_p \mathrm{B}_p$ sera un modèle de T.

On définit B_0 comme la sous-structure de A engendrée par S. Supposons B_p définie.

Pour tout entier n, toute formule $m(x_1, \ldots, x_n, y)$ et tout élément $(a_1, \ldots, a_n) \in (\mathrm{B}_p)^n$ tel que $\exists y m^{\mathrm{A}}(a_1, \ldots, a_n, y)$, choisissons un élément $b \in \mathrm{A}$ tel que $m^{\mathrm{A}}(a_1, \ldots, a_n, b)$. Soit B_{p+1} la sous-structure de A engendrée par B_p et tous ces éléments b.

Soit $\mathrm{B} = \bigcup_{p \in \mathbf{N}} \mathrm{B}_p$. C'est une sous-structure de A. Soit en effet $n \in \mathbf{N}$, $f \in \mathrm{F}_n$ et $(a_1, \ldots, a_n) \in \mathrm{B}^n$; soit $p \in \mathbf{N}$ tel que $a_1, \ldots, a_n$ appartiennent à B_p ; par définition d'une sous-structure, alors $f^{\mathrm{A}}(a_1, \ldots, a_n) \in \mathrm{B}_p$, si bien que $f^{\mathrm{A}}(a_1, \ldots, a_n) \in \mathrm{B}$.

Démontrons alors que B est un modèle de T en appliquant le test de Tarski–Vaught (théorème 3.6.3). Soit m une formule en les variables libres $x_1, \dots, x_n, y$ et soit $(a_1, \dots, a_n) \in \mathrm{B}^n$; supposons qu'il existe $a \in \mathrm{A}$ tel que $\mathrm{A} \vDash m(a_1, \dots, a_n, a)$. Soit $p \in \mathbf{N}$ tel que $a_1, \dots, a_n \in \mathrm{B}_p$. Par construction de B_{p+1}, il existe donc un élément $b \in \mathrm{B}_{p+1}$ tel que $m(a_1, \dots, a_n, b)$; en particulier, $b \in \mathrm{B}$. Cela prouve que B est un modèle de T.

Proposition **(3.6.5)**. — *Si* $\mathrm{Card}(\mathrm{S}) \geqslant \mathrm{Card}(\mathrm{L})$, *alors* $\mathrm{Card}(\mathrm{B}) = \mathrm{Card}(\mathrm{S})$.

Démonstration. — Posons $\alpha = \mathrm{Card}(\mathrm{S})$; par hypothèse, α est un cardinal infini. On peut supposer que l'ensemble V des symboles de variables est égal à $\{y, x_1, x_2, \dots, \}$, en particulier, dénombrable.

L'ensemble des termes de L est alors de cardinal $\mathrm{Card}(\mathrm{L})$; pour tout entier n, on a $\mathrm{Card}(\mathrm{S}^n) = \mathrm{Card}(\mathrm{S}) = \alpha$; le cardinal de la sous-structure B_0 engendrée par S est alors égal à α.

Supposons par récurrence que $\mathrm{Card}(\mathrm{B}_p) = \alpha$. L'ensemble des formules en les variables libres $x_1, \dots, x_n, y$ est de cardinal $\mathrm{Card}(\mathrm{L})$, et $\mathrm{Card}(\mathrm{B}_p^n) = \alpha$. Par construction, B_{p+1} est la sous-structure engendrée par B_p et un ensemble d'éléments b en cardinal au plus $\aleph_0 \alpha = \alpha$. On a donc $\mathrm{Card}(\mathrm{B}_{p+1}) = \alpha$.

Par récurrence, $\mathrm{Card}(\mathrm{B}_p) = \alpha$ pour tout entier p ; par suite, $\mathrm{Card}(\mathrm{B}) = \alpha$. □

Théorème **(3.6.6)** (Löwenheim–Skolem). — *Soit* A *une réalisation de* L *telle que* $\mathrm{Card}(\mathrm{A}) \geqslant \mathrm{Card}(\mathrm{L})$. *Pour toute partie* S *de* A, *il existe une sous-structure élémentaire* B *de* A *telle que* $\mathrm{Card}(\mathrm{B}) = \sup(\mathrm{Card}(\mathrm{L}), \mathrm{Card}(\mathrm{S}))$.

En particulier, si une théorie possède un modèle de cardinal $\alpha \geqslant \mathrm{Card}(\mathrm{L})$, elle possède un modèle de tout cardinal β tel que $\mathrm{Card}(\mathrm{L}) \leqslant \beta \leqslant \alpha$.

Démonstration. — Quitte à agrandir l'ensemble S, on suppose que $\mathrm{Card}(\mathrm{S}) \geqslant \mathrm{Card}(\mathrm{L})$. D'après la proposition précédente, la sous-structure B construite est de cardinal $\mathrm{Card}(\mathrm{S})$. □

3.7. Filtres et ultrafiltres

Définition **(3.7.1)**. — *Soit* X *un ensemble et soit* $\mathscr{F}$ *un ensemble de parties de* X. *On dit que* $\mathscr{F}$ *est un* filtre *si les propriétés suivantes sont satisfaites :*

(i) *L'ensemble* $\mathscr{F}$ *ne contient pas la partie vide ;*

(ii) *L'ensemble* X *appartient à* $\mathscr{F}$ *;*

(iii) *Si* A *et* B *sont des parties de* X *appartenant à* $\mathscr{F}$*, alors* $A \cap B$ *appartient à* $\mathscr{F}$ *;*

(iv) *Si* A *est une partie de* X *appartenant à* $\mathscr{F}$*, alors toute partie de* X *contenant* A *appartient à* $\mathscr{F}$*.*

On dit qu'un filtre $\mathscr{F}$ *est un* ultrafiltre *si tout filtre* $\mathscr{F}'$ *contenant* $\mathscr{F}$ *est égal à* $\mathscr{F}$*.*

Autrement dit, un ultrafiltre est un filtre maximal pour la relation d'ordre définie par l'inclusion.

Il résulte des propriétés (i) et (iii) que si A et B sont deux parties de X appartenant à un filtre, alors $A \cap B$ n'est pas vide.

Exemple **(3.7.2)**. — Soit X un ensemble infini. L'ensemble $\mathscr{F}$ des parties A de X telles que $X - A$ est fini est un filtre sur X ; on l'appelle le filtre de Fréchet.

Exemple **(3.7.3)**. — La topologie générale fournit de nombreux exemples de filtres.

a) Soit X un espace topologique et soit A une partie non vide de X. L'ensemble des voisinages de A est un filtre sur X.

b) Supposons $X = \mathbf{R}$ et soit $a \in X$. Une partie de $\mathbf{R}$ est un « voisinage à droite » de a si elle contient un intervalle de la forme $]a, b[$, où $b > a$. L'ensemble des voisinages à droite de a est un filtre sur X

c) Supposons encore $X = \mathbf{R}$ et prenons $a = +\infty$. Une partie de $\mathbf{R}$ est un « voisinage de $+\infty$ » si elle contient un intervalle de la forme $]a, +\infty[$. L'ensemble des voisinages de $+\infty$ est un filtre sur X.

Exemple **(3.7.4)**. — Soit X un ensemble ordonné. On suppose que X est *filtrant*, c'est-à-dire que toute partie finie de X est majorée ; autrement dit, X n'est pas vide et pour tout couple (x, y) d'éléments de X, il existe $z \in X$ tel que $x \leq z$ et $y \leq z$. Pour tout $x \in X$, on note A_x l'ensemble des $y \in X$ tels que $x \leq y$. Soit $\mathscr{F}$ l'ensemble des parties A de X contenant un ensemble de la forme A_x.

Comme $x \in A_x$, aucun élément de $\mathscr{F}$ n'est vide. Soit $x \in X$; alors X contient A_x, donc $X \in \mathscr{F}$. Soit A et B deux éléments de $\mathscr{F}$; soit $x, y \in X$ tels que $A \supset A_x$ et $B \supset A_y$; soit $z \in X$ un majorant de x et y ; alors, $A_z \subset A_x \cap A_y \subset A \cap B$, donc $A \cap B$ appartient à $\mathscr{F}$. Enfin, toute partie contenant une partie de $\mathscr{F}$ appartient à $\mathscr{F}$, par construction. Cela prouve que $\mathscr{F}$ est un filtre sur X.

Lorsque $X = \mathbf{R}$, on retrouve le filtre des voisinages de $+\infty$.

Nous utiliserons cette construction lorsque X est l'ensemble des parties finies d'un ensemble T, muni de la relation d'inclusion. Il est bien filtrant puique si x et y sont deux parties finies de T, $x \cup y$ est une partie finie de T contenant à la fois x et y.

Théorème **(3.7.5)**. — *Soit* X *un ensemble. Tout filtre sur* X *est contenu dans un ultrafiltre.*

Démonstration. — On munit l'ensemble des filtres sur X de la relation d'ordre définie par l'inclusion. Démontrons que cet ensemble ordonné est inductif. Soit Φ une partie totalement ordonnée de l'ensemble des filtres sur X; soit $\mathscr{F}$ la réunion de $\{X\}$ et des éléments de Φ. Démontrons que $\mathscr{F}$ est un filtre sur X.

L'ensemble vide n'appartenant à aucun filtre, il n'appartient pas à $\mathscr{F}$; l'ensemble X appartient à $\mathscr{F}$ par construction. Soit A et B des éléments de $\mathscr{F}$; soit $\mathscr{F}'$ et $\mathscr{F}''$ des filtres appartenant à Φ tels que $A \in \mathscr{F}'$ et $B \in \mathscr{F}''$. Puisque Φ est totalement ordonné, on a $\mathscr{F}' \subset \mathscr{F}''$, ou $\mathscr{F}'' \subset \mathscr{F}'$. Supposons, pour fixer les idées, que $\mathscr{F}' \subset \mathscr{F}''$; alors, $A, B \in \mathscr{F}''$, donc $A \cap B \in \mathscr{F}''$ puisque $\mathscr{F}''$ est un filtre; ainsi, $A \cap B \in \mathscr{F}$. Soit A un élément de $\mathscr{F}$ et soit B une partie de X qui contient A; soit $\mathscr{F}'$ un filtre appartenant à Φ tel que $A \in \mathscr{F}'$; alors $B \in \mathscr{F}'$, car $\mathscr{F}'$ est un filtre, donc $B \in \mathscr{F}$.

D'après le théorème de Zorn, tout filtre est contenu dans un filtre maximal, c'est-à-dire dans un ultrafiltre. □

Lemme **(3.7.6)**. — *Soit* X *un ensemble et soit* $\mathscr{F}$ *un filtre sur* X. *Pour que le filtre* $\mathscr{F}$ *soit un ultrafiltre, il faut et il suffit que pour toute partie* A *de* X, *soit* $A \in \mathscr{F}$, *soit* $X - A \in \mathscr{F}$.

On remarque que les conditions $A \in \mathscr{F}$ et $X - A \in \mathscr{F}$ s'excluent, car $A \cap (X - A) = \varnothing$ n'appartient pas à $\mathscr{F}$.

Démonstration. — *a*) Supposons que $\mathscr{F}$ est un ultrafiltre sur X. Soit A une partie de X; supposons que $X - A$ n'appartient pas à $\mathscr{F}$.

Soit B une partie de X appartenant à $\mathscr{F}$. Supposons que $A \cap B = \varnothing$; alors, $B \subset X - A$, donc $X - A \in \mathscr{F}$, contradiction.

Soit alors $\mathscr{F}'$ l'ensemble des parties de X qui contiennent une partie de X de la forme $A \cap F$, où $F \in \mathscr{F}$. D'après ce qui précède, $\varnothing \notin \mathscr{F}'$. Soit B, B' des parties de X appartenant à $\mathscr{F}'$. Il existe donc des ensembles $F, F' \in \mathscr{F}$ tels que $B = A \cap F$ et $B' = A \cap F'$; alors, $F \cap F' \in \mathscr{F}$ et $B \cap B' = A \cap (F \cap F') \in \mathscr{F}'$. Enfin,

soit B une partie de X appartenant à $\mathscr{F}'$ et soit B′ une partie de X contenant B ; soit F une partie appartenant à $\mathscr{F}$ tel que $B = A \cap F$; on a alors $F \cup B' \in \mathscr{F}$ et $B' = A \cap (F \cup B')$, donc $B' \in \mathscr{F}'$.

Cela démontre que $\mathscr{F}'$ est un filtre sur X. Par construction, $\mathscr{F}'$ contient $\mathscr{F}$. Puisque $\mathscr{F}$ est un ultrafiltre, on a $\mathscr{F}' = \mathscr{F}$. Comme $A \in \mathscr{F}'$, on a $A \in \mathscr{F}$, ce qu'il fallait démontrer.

b) Supposons maintenant que pour toute partie A de X, le filtre $\mathscr{F}$ contient A ou $X - A$. Soit $\mathscr{F}'$ un filtre contenant $\mathscr{F}$. Soit A une partie de X appartenant à $\mathscr{F}'$. Si A n'appartient pas à $\mathscr{F}$, alors $X - A \in \mathscr{F}$, donc $X - A \in \mathscr{F}'$; alors, $\varnothing = A \cap (X - A) \in \mathscr{F}'$, contradiction. Cela prouve que $A \in \mathscr{F}$, d'où $\mathscr{F} = \mathscr{F}'$. Ainsi, $\mathscr{F}$ est un ultrafiltre. □

Remarque **(3.7.7)**. — Le langage des filtres est très utile en topologie générale. Soit X un espace topologique. On dit qu'un filtre $\mathscr{F}$ sur X converge vers un point $x \in X$, ou que x est point limite de $\mathscr{F}$, si $\mathscr{F}$ contient le filtre des voisinages de x.

Supposons X séparé. Alors, deux points distincts possèdent des voisinages disjoints, si bien qu'un filtre ne peut avoir qu'au plus un point limite.

Supposons que X est compact et démontrons que tout ultrafiltre sur X converge. Raisonnons par l'absurde et considérons un ultrafiltre $\mathscr{F}$ qui n'a pas de point limite. Soit $x \in X$; puisque x n'est pas point limite de $\mathscr{F}$, il existe un voisinage ouvert V_x de x qui n'appartient pas à $\mathscr{F}$; son complémentaire appartient donc à $\mathscr{F}$, car $\mathscr{F}$ est un ultrafiltre. Puisque X est compact, il existe une partie finie S de X telle que $X = \bigcup_{x \in S} V_x$; alors, $\varnothing = \bigcap_{x \in S} (X - V_x)$, ce qui contredit la définition d'un filtre.

Inversement, supposons que tout ultrafiltre sur X converge et démontrons que X est compact. Considérons une famille (F_i) de parties fermées de X dont aucune intersection finie n'est vide ; démontrons que l'intersection des F_i n'est pas vide. L'ensemble $\mathscr{F}$ des parties de X qui contiennent une intersection finie $F_{i_1} \cap \cdots \cap F_{i_n}$ est un filtre sur X. En effet, on a $\varnothing \notin \mathscr{F}$, $X \in \mathscr{F}$; si A et B sont deux ensembles de $\mathscr{F}$, contenant respectivement des intersections finies $\bigcap_{i \in J} F_i$ et $\bigcap_{i \in K} F_i$, alors $A \cap B$ contient l'intersection finie $\bigcap_{i \in J \cup K} F_i$, donc appartient à $\mathscr{F}$; enfin, toute partie contenant un élément de $\mathscr{F}$ appartient à $\mathscr{F}$, par construction de $\mathscr{F}$.

Soit $\mathscr{U}$ un ultrafiltre contenant $\mathscr{F}$ et soit x Un point limite de $\mathscr{U}$. Soit $i \in I$. Soit V un voisinage ouvert de x ; l'ensemble V appartient à $\mathscr{U}$, de même que F_i,

puisque $F_i \in \mathscr{F}$, de sorte que $V \cap F_i$ appartient à $\mathscr{U}$; en particulier, $V \cap F_i$ n'est pas vide. Cela prouve que x est adhérent à F_i, donc $x \in F_i$ puisque F_i est fermé. Par suite, x appartient à tous les V_i, ce qui prouve que l'intersection des F_i n'est pas vide.

3.8. Ultraproduits, théorème de Łos

3.8.1. — Soit I un ensemble et soit $(A_i)_{i\in I}$ une famille d'ensembles indexée par I. Soit $\mathscr{F}$ un ultrafiltre sur I ; nous allons définir l'*ultraproduit* $\prod_{i\in I}^{\mathscr{F}} A_i$ des ensembles A_i le long de $\mathscr{F}$.

Soit J l'ensemble des $i \in I$ tels que $A_i \neq \varnothing$. Si $J \notin \mathscr{F}$, on pose $\prod_{i\in I}^{\mathscr{F}} A_i = \varnothing$.

Supposons maintenant que $J \in \mathscr{F}$. Soit $A = \prod_{i\in J} A_i$. Soit $\sim$ la relation dans A définie par $(a_i)_{i\in J} \sim (a'_i)_{i\in J}$ si l'ensemble des $i \in J$ tels que $a_i = a'_i$ appartient à $\mathscr{F}$. Démontrons que c'est une relation d'équivalence. Pour tout $(a_i)_{i\in J} \in A$, on a $(a_i) \sim (a_i)$ car l'ensemble J appartient à $\mathscr{F}$; cela démontre que la relation est réflexive. Elle est évidemment symétrique. Démontrons enfin qu'elle est transitive ; soit (a_i), (a'_i) et (a''_i) des éléments de A tels que $(a_i) \sim (a'_i)$ et $(a'_i) \sim (a''_i)$. Soit J' l'ensemble des $i \in J$ tels que $a_i = a'_i$; soit J'' l'ensemble des $i \in J$ tels que $a'_i = a''_i$; pour tout $i \in J' \cap J''$, on a $a_i = a'_i = a''_i$; de plus, $J', J'' \in \mathscr{F}$, donc $J' \cap J'' \in \mathscr{F}$; par suite, $(a_i) \sim (a''_i)$.

On définit alors $\prod_{i\in I}^{\mathscr{F}} A_i$ comme l'ensemble quotient $A/\sim$.

3.8.2. — Soit I un ensemble, soit $\mathscr{F}$ un ultrafiltre sur I. Soit $(A_i)_{i\in I}, (B_i)_{i\in I}$ des familles d'ensembles indexées par I ; soit $A^{\mathscr{F}}, B^{\mathscr{F}}$ leurs ultraproduits le long de l'ultrafiltre $\mathscr{F}$. Pour tout i, posons $C_i = A_i \times B_i$. Nous allons identifier l'ultraproduit $C^{\mathscr{F}}$ de la famille (P_i) avec le produit $A^{\mathscr{F}} \times B^{\mathscr{F}}$.

Soit J, resp. K l'ensemble des $i \in I$ tels que $A_i \neq \varnothing$, resp. $B_j \neq \varnothing$. Posons $A = \prod_{i\in J} A_i$ et $B = \prod_{i\in K} B_i$.

Supposons $A^{\mathscr{F}} = \varnothing$, de sorte que $J \notin \mathscr{F}$. Alors $C_i = A_i \times B_i = \varnothing$ pour tout $i \notin J$, donc $C^{\mathscr{F}} = \varnothing$.

On prouve de même que $C^{\mathscr{F}} = \varnothing$ si $A^{\mathscr{F}} = \varnothing$.

Supposons maintenant que $A^{\mathscr{F}}$ et $B^{\mathscr{F}}$ ne sont pas vides, de sorte que J et L appartiennent à $\mathscr{F}$. Soit L l'ensemble des $i \in I$ tels que $C_i \neq \varnothing$; on a $L = J \cap K$, donc $L \in \mathscr{F}$. Soit $C = \prod_{i\in L} P_i$ et soit $f : A \times B \to C$ l'application donnée par

$$f((a_i)_{i\in J}, (b_i)_{i\in K}) = ((a_i, b_i))_{i\in L}.$$

Soit $(a_i, b_i)_{i\in L}$ un élément de C. Pour tout $i \in J - L$, choisissons un élément a_i arbitraire de A_i ; pour tout $i \in K - L$, choisissons un élément b_i arbitraire de B_i. Alors, $f((a_i)_{i\in J}, (b_i)_{i\in K}) = (a_i, b_i)_{i\in L}$. Cela démontre que l'application f est surjective.

Soit $(a_i), (a'_i) \in A$, $(b_i), (b'_i) \in B$ tels que $(a_i) \sim (a'_i)$ et $(b_i) \sim (b'_i)$. Soit J' l'ensemble des $i \in J$ tels que $a_i = a'_i$, soit K' l'ensemble des $i \in K$ tels que $b_i = b'_i$; ce sont des éléments de $\mathscr{F}$ par hypothèse. Comme l'ensemble des $i \in L$ tels que $(a_i, b_i) = (a'_i, b'_i)$ contient $J' \cap K'$, on a $(a_i, b_i)_{i\in L} \sim (a'_i, b'_i)_{i\in L}$. Cela démontre que l'application f est compatible aux relations d'équivalence définies par l'ultrafiltre $\mathscr{F}$ dans les ensembles A, B, C. Elle induit donc, par passage aux quotients, une application φ de $A^{\mathscr{F}} \times A^{\mathscr{F}}$ sur $C^{\mathscr{F}}$.

Comme l'application f est surjective, il en est de même de l'application φ.

Soit $(a_i), (a'_i) \in A$, $(b_i), (b'_i) \in B$ tels que $f((a_i), (b_i)) \sim f((a'_i), (b'_i))$. Soit J' l'ensemble des $i \in J$ tels que $a_i = a'_i$, soit K' l'ensemble des $i \in K$ tels que $b_i = b'_i$. Alors, $J' \cap K'$ est l'ensemble des $i \in L$ tels que $(a_i, b_i) = (a'_i, b'_i)$, donc $J' \cap K'$ appartient à $\mathscr{F}$. Par suite, $J' \in \mathscr{F}$ et $K' \in \mathscr{F}$, ce qui entraîne que $(a_i) \sim (a'_i)$ et $(b_i) \sim (b'_i)$. Cela prouve que l'application φ est injective.

3.8.3. — Soit I un ensemble, soit $\mathscr{F}$ un ultrafiltre sur I. Soit $(A_i)_{i\in I}, (B_i)_{i\in I}$ des familles d'ensembles indexées par I ; soit $A^{\mathscr{F}}, B^{\mathscr{F}}$ leurs ultraproduits le long de l'ultrafiltre $\mathscr{F}$. Pour tout $i \in I$, soit f_i une application de A_i dans B_i.

On déduit de la famille (f_i) une application $f^{\mathscr{F}}$ de $A^{\mathscr{F}}$ dans $B^{\mathscr{F}}$. Elle associe à la classe d'une famille (a_i) la classe de la famille $(f_i(a_i))$.

3.8.4. — Soit L un langage, soit $(A_i)_{i\in I}$ une famille de réalisations de L indexée par un ensemble I et soit $\mathscr{F}$ un ultrafiltre sur I. Soit $A^{\mathscr{F}}$ l'ultraproduit de la famille $(A_i)_{i\in I}$ le long de l'ultrafiltre $\mathscr{F}$.

Soit n un entier et soit $f \in F_n$ un symbole de fonction n-aire. Par un raisonnement similaire à celui du paragraphe précédent, on déduit de la famille (f^{A_i}) une fonction $f^{A^{\mathscr{F}}}$. de $(A^{\mathscr{F}})^n$ dans $A^{\mathscr{F}}$.

Soit n un entier et soit $r \in R_n$ un symbole de relation n-aire. On déduit de la famille (r^{A_i}) une relation n-aire $r^{A^{\mathscr{F}}}$ dans l'ensemble $A^{\mathscr{F}}$: la classe d'une famille $(a_i)_{i\in J} \in \prod_{i\in J} A_i^n$ appartient au graphe de cette relation si et seulement si l'ensemble des $i \in J$ tels que $r_i(a_i) = \top$ appartient à $\mathscr{F}$.

Ainsi, l'ultraproduit $A^{\mathscr{F}}$ est naturellement muni d'une structure de réalisation du langage L.

3.8.5. — Soit $t \in \mathscr{T}_{\mathrm{L}}$ un terme et soit V l'ensemble des symboles de variables qui apparaissent dans t. On démontre par récurrence sur la définition d'un terme que l'interprétation $t^{\mathrm{A}^{\mathscr{F}}}$ de t dans la structure $\mathrm{A}^{\mathscr{F}}$ est l'application de $(\mathrm{A}^{\mathscr{F}})^{\mathrm{V}}$ dans $\mathrm{A}^{\mathscr{F}}$ déduite des applications $t^{\mathrm{A}_i} : \mathrm{A}_i^{\mathrm{V}} \to \mathrm{A}_i$.

Soit $m \in \mathscr{F}_{\mathrm{L}}$ une formule et soit V l'ensemble des symboles de variables libres de m. On démontre par récurrence sur la définition d'un terme que l'interprétation $m^{\mathrm{A}^{\mathscr{F}}}$ de m dans la structure $\mathrm{A}^{\mathscr{F}}$ est l'application de $(\mathrm{A}^{\mathscr{F}})^{\mathrm{V}}$ dans $\{\bot, \top\}$ déduite des applications $m^{\mathrm{A}_i} : \mathrm{A}_i^{\mathrm{V}} \to \{\bot, \top\}$.

Théorème **(3.8.6)** (Łos). — *Soit* L *un langage, soit* $(\mathrm{A}_i)_{i\in\mathrm{I}}$ *une famille de réalisations de* L *et soit* $\mathscr{F}$ *un ultrafiltre sur l'ensemble* I. *Soit* $\mathrm{A}^{\mathscr{F}} = \prod_{i\in\mathrm{I}}^{\mathscr{F}} \mathrm{A}_i$ *l'ultraproduit de la famille* (A_i) *selon l'ultrafiltre* $\mathscr{F}$.

Soit $m \in \mathscr{F}_{\mathrm{L}}$ *une formule close dans le langage* L. *Pour que la formule* m *soit réalisée dans l'ultraproduit* $\mathrm{A}^{\mathscr{F}}$, *il faut et il suffit que l'ensemble des* $i \in \mathrm{I}$ *tels que* m *est réalisée dans* A_i *appartienne à* $\mathscr{F}$.

Démonstration. — C'est exactement la définition de l'interprétation de la formule m dans la réalisation $\mathrm{A}^{\mathscr{F}}$. □

3.9. Théorème de compacité et théorème de Tarski

Théorème **(3.9.1)** (Théorème de compacité, Gödel). — *Soit* L *un langage et soit* T *une théorie. Pour que* T *soit consistante, il faut et il suffit que toute partie finie de* T *soit consistante.*

Démonstration. — Toute partie d'une théorie consistante est consistante; il faut donc démontrer que si toute partie finie de T est consistante, alors T est consistant. L'assertion est évidente si T est finie; supposons donc que T soit infinie. Soit I l'ensemble (infini) des parties finies de T; pour $i \in \mathrm{I}$, on notera T_i la théorie i. D'après l'exemple 3.7.4, il existe un filtre $\mathscr{F}$ sur I tel que pour tout $i \in \mathrm{I}$, l'ensemble des parties finies de T contenant i appartient à $\mathscr{F}$. D'après le théorème 3.7.5, il existe un tel filtre qui est un ultrafiltre.

Pour tout $i \in \mathrm{I}$, soit A_i un modèle de la théorie T_i. Soit $\mathrm{A} = \prod_{i\in\mathrm{I}}^{\mathscr{F}} \mathrm{A}_i$ l'ultraproduit de la famille (A_i) le long de l'ultrafiltre $\mathscr{F}$. C'est une réalisation du langage L. Démontrons que c'est un modèle de T.

Soit $m \in \mathrm{T}$. Démontrons que m est réalisée dans A. Par définition des structures A_i, la formule m est réalisée dans A_i dès que $m \in \mathrm{T}_i$. Par construction de

l'ultrafiltre $\mathscr{F}$, l'ensemble des parties finies de T contenant m appartient à $\mathscr{F}$. D'après le théorème de Łos, la formule m est donc réalisée dans A. □

3.9.2. — Soit $\mathscr{T}_{\mathrm{L}}$ l'ensemble des théories complètes dans le langage L ; c'est une partie de $\mathscr{P}(\mathscr{F}_{\mathrm{L},\varnothing})$.

Pour tout formule close m, notons V_m l'ensemble des théories complètes qui contiennent m. Pour toute ensemble S de formules closes, notons $\mathrm{V}_{\mathrm{S}} = \bigcap_{m\in\mathrm{S}} \mathrm{V}_m$; c'est l'ensemble des théories complètes qui contiennent S.

On a $\mathrm{V}_{\top} = \mathscr{T}_{\mathrm{L}}$ et $\mathrm{V}_{\bot} = \varnothing$.

Si m et n sont des formules closes, alors $\mathrm{V}_m \cap \mathrm{V}_n = \mathrm{V}_{(m\wedge n)}$; en effet, si m, n appartiennent à une théorie complète T et si A est un modèle de T, alors $m^{\mathrm{A}} = n^{\mathrm{A}} = \top$, donc $(m \wedge n)^{\mathrm{A}} = \top$, ce qui prouve que $(m \wedge n)$ appartient à $\mathrm{Th}_{\mathrm{L}}(\mathrm{A}) = \mathrm{T}$.

Par suite, les ensembles de la forme V_m forment une base d'ouverts d'une unique topologie sur $\mathscr{T}_{\mathrm{L}}$: une partie de $\mathscr{T}_{\mathrm{L}}$ est ouverte si et seulement si c'est une réunion de parties de la forme V_m.

Observons que $\mathrm{V}_m \cap \mathrm{V}_{\neg m} = \varnothing$ et que $\mathrm{V}_m \cup \mathrm{V}_{\neg m} = \mathscr{T}_{\mathrm{L}}$, pour toute formule close m. En effet, si A est une réalisation de L, alors soit m est satisfaite dans A, soit $\neg m$ est satisfaite dans A, mais pas les deux. Par suite, les ensembles V_m sont ouverts et fermés dans $\mathscr{T}_{\mathrm{L}}$.

Comme une intersection de parties fermées est fermée, V_{S} est fermé, pour tout ensemble S de formules closes. Inversement, soit F une partie fermée de $\mathscr{T}_{\mathrm{L}}$; soit U son complémentaire et soit S un ensemble de formules closes tel que $\mathrm{U} = \bigcup_{m\in\mathrm{S}} \mathrm{V}_m$; alors,

$$\mathrm{F} = \bigcap_{m\in\mathrm{S}} (\mathscr{T}_{\mathrm{L}} - \mathrm{V}_m) = \bigcap_{m\in\mathrm{S}} \mathrm{V}_{\neg m} = \mathrm{V}_{\mathrm{S}'},$$

où $\mathrm{S}' = \{\neg m\,;\, m \in \mathrm{S}\}$.

Corollaire **(3.9.3)**. — *L'espace topologique $\mathscr{T}_{\mathrm{L}}$ est compact et totalement discontinu.*

Démonstration. — Considérons enfin une famille $(\mathrm{F}_i)_{i\in\mathrm{I}}$ de parties fermées de $\mathscr{T}_{\mathrm{L}}$ dont l'intersection est vide et démontrons qu'il existe une sous-famille finie d'intersection vide. Pour tout i, soit S_i un ensemble de formules closes tel que $\mathrm{F}_i = \mathrm{V}_{\mathrm{S}_i}$; posons $\mathrm{S} = \bigcup_i \mathrm{S}_i$. On a $\mathrm{V}_{\mathrm{S}} = \bigcap_i \mathrm{V}_{\mathrm{S}_i} = \bigcap_i \mathrm{F}_i = \varnothing$, ce qui signifie que l'ensemble S n'est pas consistant. D'après le théorème de compacité, il existe une partie finie S' de S qui est inconsistante. Pour toute formule $m \in \mathrm{S}'$, choisissons

un élément i_m tel que $m \in \mathrm{S}_i$; soit I′ l'ensemble, fini, de ces i_m. Si une théorie consistante T appartient à $\bigcap_{i\in \mathrm{I}'} \mathrm{F}_i$, alors T appartient à F_i pour tout i, donc T contient S′, ce qui contredit la définition pour une théorie d'être consistante. Par suite, $\bigcap_{i\in \mathrm{I}'} \mathrm{F}_i = \varnothing$.

Soit T et T′ des théories complètes telles que $\mathrm{T} \neq \mathrm{T}'$. Il existe donc $m \in \mathrm{T}$ tel que $m \notin \mathrm{T}'$; on a donc $\neg m \in \mathrm{T}'$. Alors, V_m et $\mathrm{V}_{\neg m}$ sont des voisinages ouverts de T et T′ respectivement, disjoints. Cela entraîne que $\mathscr{T}_{\mathrm{L}}$ est séparé.

L'espace topologique $\mathscr{T}_{\mathrm{L}}$ vérifie l'axiome de Borel–Lebesgue et est séparé ; il est donc compact.

Démontrons enfin que $\mathscr{T}_{\mathrm{L}}$ est totalement discontinu, c'est-à-dire que ses composantes connexes sont réduites à des points. Soit T une théorie complète soit U sa composante connexe dans $\mathscr{T}_{\mathrm{L}}$; démontrons que $\mathrm{U} = \{\mathrm{T}\}$. Pour toute formule close m appartenant à T, l'ensemble $\mathrm{U} \cap \mathrm{V}_m$ est ouvert et fermé dans U, et contient T, donc est égal à U ; on a donc $\mathrm{U} \subset \mathrm{V}_m$. L'intersection, pour $m \in \mathrm{T}$, des ensembles V_m est l'ensemble des théories complètes contenant T, donc est égale à $\{\mathrm{T}\}$. Cela prouve que $\mathrm{U} \subset \{\mathrm{T}\}$. □

Théorème **(3.9.4)** (Tarski). — *Soit* A *une réalisation infinie de* L. *Pour tout cardinal* $\alpha \geqslant \sup(\mathrm{Card}(\mathrm{A}), \mathrm{Card}(\mathrm{L}))$, *il existe une extension élémentaire* B *de* A *de cardinal* α.

Démonstration. — Soit L′ le langage obtenu en adjoignant à L une famille de symboles de constantes (c_i) indexée par un ensemble I de cardinal α.

Soit T′ l'ensemble de formules closes du langage L′ obtenu en adjoignant à T les formules $c_i \neq c_j$ pour tout couple (i, j) d'éléments de I tels que $i \neq j$.

Démontrons que T′ est consistant. D'après le théorème de compacité, il suffit de prouver que toute partie finie de T′ est consistante ; une telle partie finie est contenue dans la réunion T'_{J} de T et de l'ensemble de formules $c_i \neq c_j$, pour $i \neq j$ dans un ensemble fini J. Comme A est infini, il existe une famille $(a_i)_{i\in \mathrm{J}}$ telle que $a_i \neq a_j$ pour $i \neq j$. On observe alors que A, muni des interprétations $c_i^{\mathrm{A}} = a_i$, pour $i \in \mathrm{J}$, est un modèle de T'_{J}. Par suite, T′ est consistant, donc possède un modèle.

Soit B un modèle de T′. C'est une réalisation B de L munie d'une famille $(b_i)_{i\in \mathrm{I}}$ dans laquelle les formules de T d'une part, les formules $b_i \neq b_j$ pour $i \neq j$ d'autre part, sont satisfaites. Par suite, B est un modèle de T et $\mathrm{Card}(\mathrm{B}) \geqslant \mathrm{Card}(\mathrm{I}) = \alpha$.

En choisissant une partie S de B de cardinal α, on déduit du théorème de Löwenheim–Skolem qu'il existe un modèle B′ de B contenant S et dont le cardinal est égal à α. □

CHAPITRE 4

THÉORIE DE LA DÉMONSTRATION — *RAISONNER*

« Pour de l'esprit, j'en ai, sans doute ; et du bon goût,
À juger sans étude et raisonner de tout. »
— Molière, *Le Misanthrope*, acte III, scène 1

Ce chapitre est consacré aux aspects *syntaxiques* de la logique du premier ordre, c'est-à-dire à la définition d'une démonstration formelle et à l'étude des formules qui en possèdent une.

Soit L un langage du premier ordre et soit T une théorie de L. Par définition, on dira qu'une formule close m se déduit de T s'il existe une suite finie $(m_1, \ldots, m_n)$ de formules closes telles que pour tout i, soit $m_i \in$ T, soit m_i se déduit de $(m_1, \ldots, m_{i-1})$ par application d'une règle de déduction, et si $m_n = m$. Une telle suite est appelée démonstration de m dans T. Ces règles de déduction sont soumises à deux tensions opposées :

– Être « raisonnables », au sens où une démonstration formelle doit conduire, une fois convenablement interprétée, à un énoncé vrai ;
– Être « complètes », c'est-à-dire prouver tout ce qui peut l'être.

La première contrainte sera vérifiée dans le paragraphe 4.2 : on démontre que si une formule m possède une démonstration formelle dans une théorie T, alors m est vraie dans tout modèle de T. La seconde contrainte est Le théorème de complétude de Gödel (théorème 4.5.3). vérifie la seconde contrainte : si m est vraie dans tout modèle de T, alors m possède une démonstration dans T.

Les règles de déduction sont de deux types : deux d'entre elles, modus ponens et généralisation, disent comment écrire de nouvelles formules à partir de précédentes ; les autres sont des axiomes, parfois appelés *axiomes logiques* : tautologies, syntaxe des quantificateurs, égalité. Comme nous avons autorisé la réalisation vide, les règles que nous donnons ci-dessous diffèrent légèrement de celles proposées dans la plupart des ouvrages, notamment par Cori & Lascar (2003*a*). Ces auteurs introduisent par exemple la règle de particularisation (4.1.4.3) ($\forall v\, m \Rightarrow m(t)$, où m est une formule et t un terme) même lorsque

le symbole de variable v n'a pas d'occurrence libre dans m. L'exemple de la formule $\phi = \forall x \bot$, qui est satisfaite uniquement dans la structure vide, nous force à quelques précautions supplémentaires.

4.1. Démonstrations formelles

4.1.1. Modus ponens. — Soit m, n deux formules. On fait l'hypothèse que toute variable libre de m admet au moins une occurrence libre dans n. Alors, la règle *modus ponens*[1] déduit n des deux formules m et $(m \Rightarrow n)$.

4.1.2. Règle de généralisation. — Soit m une formule et v un symbole de variable. La *règle de généralisation* permet de déduire $\forall v m$ de m.

4.1.3. Tautologies. — Soit $t(q_1, \ldots, q_n)$ un terme du calcul des prédicats en n variables $x_1, \ldots, x_n$. Supposons que t soit une *tautologie*, c'est-à-dire que l'on a $t(a_1, \ldots, a_n) = \top$ pour tout $(a_1, \ldots, a_n) \in \{\bot, \top\}^n$.

Alors, pour toute suite $(m_1, \ldots, m_n)$ de formules de L, la règle

(4.1.3.1) $$t(m_1, \ldots, m_n)$$

affirme la formule (tautologique) déduire de t dans laquelle la formule m_i est substituée à la variable x_i.

4.1.4. Axiomes des quantificateurs. — Soit v un symbole de variable et soit m, n des formules. On dispose des trois règles de déduction suivantes, aussi appelées *axiomes des quantificateurs*.

La règle

(4.1.4.1) $$(\exists v m \Leftrightarrow \neg \forall v \neg m)$$

exprime la relation entre quantificateur existentiel et quantificateur universel.

La règle suivante

(4.1.4.2) $$(\forall v (m \Rightarrow n) \Rightarrow (\forall v m \Rightarrow \forall v n))$$

permet de distribuer le quantificateur universel dans une implication.

Soit $t(x_1, \ldots, x_n)$ est un terme en les variables libres $x_1, \ldots, x_n$ et soit $\tau : \mathscr{F}_{\mathrm{L},\{v\}} \to \mathscr{F}_{\mathrm{L}}$ l'application de substitution correspondante. Supposons que le

[1] Dérivée de l'expression latin *modus ponendo ponens* qui signifie « manière d'affirmer en affirmant ».

symbole v possède une occurrence libre dans m et que m soit admissible pour la substitution $\{v\} \to \mathscr{T}_{\mathrm{L}}$ d'image t. Alors, la *règle de particularisation*

(4.1.4.3) $$(\forall v m \Rightarrow \tau(m))$$

permet de déduire de $\forall v m$ la formule $\tau(m)$ dans laquelle les occurrences libres de la variable v ont été remplacées par le terme t.

4.1.5. Axiomes de l'égalité. — Enfin, on ajoute les axiomes suivants qui régissent l'égalité :

(4.1.5.1) $$x = x$$

(4.1.5.2) $$((x = y) \Rightarrow (y = x))$$

(4.1.5.3) $$((x = y) \wedge (y = z)) \Rightarrow (x = z)$$

(4.1.5.4) $$((\bigwedge_{1 \leqslant i \leqslant n} x_i = y_i) \Rightarrow (f x_1 \dots x_n = f y_1 \dots y_n))$$

(4.1.5.5) $$((\bigwedge_{1 \leqslant i \leqslant n} x_i = y_i) \Rightarrow (r x_1 \dots x_n \Rightarrow r y_1 \dots y_n)),$$

où $x, y, z, x_1, \dots, x_n, y_1, \dots, y_n$ sont des symboles de variables deux à deux distincts, f un symbole de fonction n-aire et r un symbole de relation n-aire.

Définition **(4.1.6)**. — *Soit* T *une théorie et soit m une formule de* L. *Une* démonstration formelle *de m dans* T *est une suite finie $(m_1, \dots, m_n)$ de formules de* L *telle que $m_n = m$ et telle que pour tout $i \in \{1, \dots, n\}$, l'une des propriétés suivantes soit satisfaite :*

a) *On a $m_i \in$ T;*

b) *Il existe un symbole de variable v et un entier j tel que $1 \leqslant j < i$ de sorte que $m_i = \forall v m_j$;*

c) *La formule m_i est une tautologie ou un axiome de quantificateur;*

d) *Il existe deux entiers j, k tels que $1 \leqslant j, k < i$ et $m_k = (m_j \Rightarrow m_i)$, où toute variable libre dans m_j est libre dans m_i, de sorte que la formule m_i se déduise des formules m_j et m_k par application de* modus ponens.

Si m possède une démonstration formelle dans T, on dit aussi qu'elle est *démontrable* dans T, qu'elle est conséquence syntaxique de T, ou que c'est un *théorème* de T ; on écrit $\mathrm{T} \vdash m$. Lorsque T est vide, on dit que m est *démontrable* et l'on écrit $\vdash m$.

Exemple **(4.1.7)**. — Soit m, n deux formules closes et soit $\mathrm{T} = \{m, n\}$. La suite de formules suivante constitue une démonstration formelle de $m \wedge n$ dans T :

(1) m — on a $m \in \mathrm{T}$;
(2) n — on a $n \in \mathrm{T}$;
(3) $(m \Rightarrow (n \Rightarrow (m \wedge n)))$ — c'est une tautologie ;
(4) $(n \Rightarrow (m \wedge n))$ — *modus ponens,* (1) et (3) ;
(5) $(m \wedge n)$ — *modus ponens,* (2) et (4).

Exemple **(4.1.8)**. — Soit T, T′ des théories et soit m une formule telle que $\mathrm{T} \cup \mathrm{T}' \vdash m$. Supposons que $\mathrm{T} \vdash n$ pour toute formule $n \in \mathrm{T}'$; alors, $\mathrm{T} \vdash m$.

En effet, considérons une démonstration $(m_1, \ldots, m_k)$ de m dans $\mathrm{T} \cup \mathrm{T}'$. Pour tout $i \in \{1, \ldots, k\}$ tel que $m_i \in \mathrm{T}'$, on remplace m_i par une démonstration $(m_i^1, \ldots, m_i^{k_i})$ de m_i dans T. On obtient ainsi une démonstration de m dans T.

Exemple **(4.1.9)**. — Soit m et n deux formules. Supposons que le symbole de variable v n'ait pas d'occurence libre dans m et démontrons la variante suivante de l'axiome de quantificateur (4.1.4.2) :

(4.1.9.1) $$(\forall v(m \Rightarrow n) \Rightarrow (m \Rightarrow \forall v n)).$$

En voici une démonstration formelle :

(1) $(m \Rightarrow \forall v m)$ — règle de généralisation
(2) $(\forall v(m \Rightarrow n) \Rightarrow (\forall v m \Rightarrow \forall v n))$ — axiome (4.1.4.2) ;
(3) $((a \Rightarrow (b \Rightarrow c)) \Rightarrow ((b' \Rightarrow b) \Rightarrow (a \Rightarrow (b' \Rightarrow c))))$ — tautologie, appliquée avec $a = \forall v(m \Rightarrow n)$, $b' = m$, $b = \forall v m$, $c = (m \Rightarrow \forall v n)$;
(4) $((m \Rightarrow \forall v m) \Rightarrow (\forall v(m \Rightarrow n) \Rightarrow (m \Rightarrow \forall v n)))$ — *modus ponens*, (2) et (3) toute variable libre de (3) ayant au moins une occcurence libre dans (4) ;
(5) $(\forall v(m \Rightarrow n) \Rightarrow (m \Rightarrow \forall v n)))$ — *modus ponens*, (1) et (4), toute variable libre de (1) a bien une occurrence libre dans (5).

Exemple **(4.1.10)**. — Soit m, n, p des formules telles que $(m \Rightarrow n)$ et $(n \Rightarrow p)$ soient démontrables dans une théorie T. Prouvons que si toute variable libre de n possède au moins une occurrence libre dans m ou dans p, alors $(m \Rightarrow p)$ est démontrable dans T. En effet, on concatène des démonstrations formelles de $(m \Rightarrow n)$ et de $(n \Rightarrow p)$ et on ajoute les formules suivantes :

(1) $(m \Rightarrow n)$;
(2) $(n \Rightarrow p)$;
(3) $((m \Rightarrow n) \Rightarrow ((n \Rightarrow p) \Rightarrow (m \Rightarrow p)))$ — tautologie ;

(4) $((n \Rightarrow p) \Rightarrow (m \Rightarrow p))$ — *modus ponens* à partir de (3), toute variable libre de (1) ayant une occurence libre dans (4) ;

(5) $(m \Rightarrow p)$ — *modus ponens* à partir de (4), toute variable libre de (2) ayant au moins une occurrence libre dans (5), par hypothèse.

Remarque **(4.1.11)**. — Soit m une formule, soit v un symbole de variable et soit t un terme. La règle de particularisation (4.1.4.3), $(\forall v m \Rightarrow m(t))$, n'est admise que lorsque v admit une occurrence libre dans m. Lorsque v n'a pas d'occurrence libre dans m, $m(t) = m$ et cette règle s'écrirait $(\forall v m \Rightarrow m)$. Alors, la formule $\phi = \forall x \bot$, vraie dans la réalisation vide, et fausse sinon, nous permettrait de déduire $\bot$ de ϕ, ce qui est déraisonnable : si l'ensemble vide est d'un intérêt limité en mathématiques, il ne s'agit tout de même pas qu'il conduise à des contradictions !

Bien sûr, un langage qui possède des symboles de constantes n'a pas de réalisation vide. Démontrons ainsi que si le langage contient une symbole de constante c, alors la formule $(\forall v m \Rightarrow m)$ est démontrable pour toute formule m où le symbole de variable v n'a pas d'occurrence libre. Les formules suivantes sont en effet démontrables :

(1) $(v = v) \Leftrightarrow \top$ — aviome de l'égalité ;

(2) $((p \Leftrightarrow \top) \Rightarrow (m \Rightarrow (m \wedge p)))$ — tautologie, qu'on va appliquer avec $p = (v = v)$;

(3) $(m \Rightarrow (m \wedge (v = v))$, *modus ponens*, car v est la seule varible libre de (1) et a une occurrence libre dans (3) ;

(4) $\forall v(m \Rightarrow (m \wedge (v = v)))$ — généralisation de (3) ;

(5) $(\forall v(m \Rightarrow (m \wedge (v = v))) \Rightarrow (\forall v m \Rightarrow \forall v(m \wedge (v = v)))$ — aviome logique (4.1.4.2) ;

(6) $(\forall v m \Rightarrow \forall v(m \wedge (v = v)))$;

(7) $(\forall v(m \wedge (v = v)) \Rightarrow (m \wedge (c = c)))$ — particularisation ;

(8) $((m \wedge (c = c)) \Rightarrow m)$ — tautologie.

On est dans la situation $p \Rightarrow q$, $q \Rightarrow r$, $r \Rightarrow s$, où les variables libres de $p = \forall v m$, $q = \forall v(m \wedge (v = v))$, $r = (m \wedge (c = c))$, $s = m$, sont celles de m. D'après l'exemple 4.1.10, la formule $(\forall v m \Rightarrow m)$ est donc démontrable dans T.

4.2. Démonstration et satisfaction

Définition **(4.2.1)**. — *Soit m une formule d'un langage* L. *Une* quantification universelle *de m est une formule close de la forme $\forall x_1 \dots \forall x_n m$, où $x_1, \dots, x_n$ sont des symboles de variables.*

La condition que m est une formule close revient à dire que l'ensemble des variables libres de m est contenu dans l'ensemble $\{x_1, \dots, x_n\}$. En effet, si un symbole de variable ayant une occurrence libre dans m n'appartenait pas à $\{x_1, \dots, x_n\}$, il serait encore libre dans la formule $\forall x_1 \dots \forall x_n m$.

Soit W un ensemble de symboles de variables. On dira qu'une formule $m \in \mathscr{F}_{\mathrm{L,W}}$ est satisfaite dans une structure A si la fonction $m^{\mathrm{A}} : \mathrm{A}^{\mathrm{W}} \to \{\top, \bot\}$ ne prend que la valeur $\top$. Cela revient à dire que toute quantification universelle de m est satisfaite dans A. On notera cela par abus $m^{\mathrm{A}} = \top$.

C'est un abus à cause des possibles structures vides. En effet, si A est vide et que W n'est pas vide, l'ensemble A^{W} est vide, donc il est vrai qu'une fonction de A^{W} dans $\{\top, \bot\}$ ne prend que la valeur $\top$ — mais il est aussi vrai qu'elle ne prend que la valeur $\bot$.

D'autre part, soit m une formule close telle que $m^{\mathrm{A}} = \bot$ dans la structure vide A. Pour tout symbole de variable x, $m' = \forall x m$ est une formule close, mais on a $(m')^{\mathrm{A}} = \top$.

Théorème **(4.2.2)**. — *Soit* L *un langage, soit* T *une théorie et soit m une formule de* L *qui est démontrable dans* T. *Alors m est satisfaite dans tout modèle de* T.

Ce théorème garantit donc que la définition d'une démonstration formelle est « saine » : dans un modèle donné, à partir d'énoncés vrais, elle ne peut pas conduire à des énoncés faux. Modulo un argument de récurrence élémentaire, la preuve se déduit des lemmes suivants.

Lemme **(4.2.3)**. — *Soit* A *une réalisation de* L *et soit m, n des formules dans le langage* L *telle que toute variable libre de m possède au moins une occurrence libre dans n. Si $m^{\mathrm{A}} = \top$ et $(m \Rightarrow n)^{\mathrm{A}} = \top$, alors $n^{\mathrm{A}} = \top$.*

Démonstration. — Soit W un ensemble de symboles de variables contenant les variables libres de n. Démontrons que la fonction $n^{\mathrm{A}} : \mathrm{A}^{\mathrm{W}} \to \{\top, \bot\}$ ne prend que la valeur $\top$.

Soit $a \in \mathrm{A}^{\mathrm{W}}$. Par hypothèse, tout symbole de variable qui a une occurrence libre dans m a aussi une occurrence libre dans n. La formule m appartient ainsi

à $\mathscr{F}_{\mathrm{L,W}}$, et les variables libres de la formule $(m \Rightarrow n)$ sont aussi contenues dans W. Par hypothèse, on a $m^{\mathrm{A}}(a) = \top$ et $(m \Rightarrow n)^{\mathrm{A}}(a) = \top$. Par définition de l'évaluation des formules, on a donc $n^{\mathrm{A}}(a) = \top$. □

Lemme **(4.2.4)**. — *Les tautologies (formules (4.1.3.1)) et les axiomes logiques (formules (4.1.4.1), (4.1.4.2) et (4.1.4.3)) sont satisfaites dans toute réalisation.*

Démonstration. — Il s'agit de démontrer que l'évaluation de ces formules dans toute réalisation de L ne prend que la valeur $\top$. Soit donc A une réalisation de L.

Traitons d'abord le cas des tautologies. Posons $p = t(m_1, \ldots, m_n)$ et soit W un ensemble de symboles de variables contenant les variables libres de p. Alors, W contient les variables libres de m_i, pour tout i. L'évaluation des formules étant compatible aux règles logiques, on a $p^{\mathrm{A}}(a) = t^{\mathrm{A}}(m_1^{\mathrm{A}}(a), \ldots, m_n^{\mathrm{A}}(a)) = \top$ pour tout $a \in \mathrm{A}^{\mathrm{W}}$, ce qu'il fallait démontrer.

Traitons maintenant les axiomes logiques.

Commençons par la règle (4.1.4.1) : $(\exists v m \Leftrightarrow \neg\forall v \neg m)$. Notons p_1 cette formule, soit W un ensemble de symboles de variables contenant ses variables libres et posons $\mathrm{W}' = \mathrm{W} \cup \{v\}$. Soit $a \in \mathrm{A}^{\mathrm{W}}$. Par définition de l'évaluation, on a $p_1^{\mathrm{A}}(a) = \top$ si et seulement si $(\exists v m)^{\mathrm{A}}(a) = (\neg\forall v \neg m)^{\mathrm{A}}(a)$, c'est-à-dire si et seulement si $(\exists v m)^{\mathrm{A}}(a) \neq (\forall v \neg m)^{\mathrm{A}}(a)$, La première formule vaut $\top$ si et seulement s'il existe $a' \in \mathrm{A}^{\mathrm{W}'}$ ne différant de a qu'en v tel que $m^{\mathrm{A}}(a') = \top$; la seconde vaut $\top$ si et seulement si pour tout $a' \in \mathrm{A}^{\mathrm{W}'}$ ne différant de a qu'en v, on a $\neg m^{\mathrm{A}}(a') = \top$, c'est-à-dire $m^{\mathrm{A}}(a') = \bot$. On a donc $p_1^{\mathrm{A}}(a) = \top$, comme il fallait démontrer.

Posons $p_2 = (\forall v(m \Rightarrow n)) \Rightarrow (\forall v m \Rightarrow \forall v n))$ (formule (4.1.4.2)). Soit W un ensemble de symboles de variables contenant les variables libres de p_2 et soit $\mathrm{W}' = \mathrm{W} \cup \{v\}$. Soit $a \in \mathrm{A}^{\mathrm{W}}$. On a $(p_2)^{\mathrm{A}}(a) = \top$ si et seulement si $(\forall v(m \Rightarrow n))^{\mathrm{A}}(a) = \bot$ ou $(\forall v m \Rightarrow \forall v n)^{\mathrm{A}}(a) = \top$, c'est-à-dire si et seulement si $(\forall v(m \Rightarrow n))^{\mathrm{A}}(a) = \bot$, $(\forall v m)^{\mathrm{A}}(a) = \bot$ ou $(\forall v n)^{\mathrm{A}}(a) = \top$. SUpposons donc que $(\forall v(m \Rightarrow n))^{\mathrm{A}}(a) = (\forall v m)^{\mathrm{A}}(a) = \top$ et prouvons que $(\forall v n)^{\mathrm{A}}(a) = \top$. Soit $a' \in \mathrm{A}^{\mathrm{W}'}$ ne différant de a qu'en v ; nous devons démontrer que $n^{\mathrm{A}}(a') = \top$. Par hypothèse, on a $(m \Rightarrow n)^{\mathrm{A}}(a') = m^{\mathrm{A}}(a') = \top$, d'où $n^{\mathrm{A}}(a') = \top$, comme il fallait démontrer.

Reprenons les notations de la formule (4.1.4.3), posons $p_3 = (\forall v m \Rightarrow \tau(m))$. Les variables libres de p_3 sont celles de m privées de v, ainsi que celles apparaissant dans le terme t, car v a une occurrence libre dans m. Soit W un ensemble de symboles de variables contenant les variables libres de p_3. Posons aussi $\mathrm{W}' = \mathrm{W} \cup \{v\}$ et notons $\varphi : \mathrm{A}^{\mathrm{W}} \to \mathrm{A}^{\mathrm{W}'}$ l'application telle que $\varphi_v(a) = t^{\mathrm{A}}(a)$, et $\varphi_w(a) = a_w$ pour $w \in \mathrm{W} - \{v\}$ et tout $a \in \mathrm{A}^{\mathrm{W}}$.

Soit $a \in \mathrm{A}^{\mathrm{W}}$ et démontrons que $p_3^{\mathrm{A}}(a) = \top$. Par définition, on peut supposer que $(\forall v m)^{\mathrm{A}}(a) = \top$ et il s'agit de prouver que $(\tau(m))^{\mathrm{A}}(a) = \top$. Puisque $\varphi(a)$ ne diffère de a qu'en v, la formule $(\forall v m)^{\mathrm{A}}(a) = \top$ entraîne que $m^{\mathrm{A}}(\varphi(a)) = \top$. Par compatibilité de l'évaluation à la substitution (lemme 3.1.9), on a donc $(\tau(m))^{\mathrm{A}}(a) = (m^{\mathrm{A}})(\varphi(a)) = \top$, comme il fallait démontrer. □

Démonstration du théorème 4.2.2. — Soit A un modèle de T. Soit m une formule dans le langage L qui est démontrable dans T. Soit $(m_1, \ldots, m_n)$ une démonstration formelle de m. Prouvons par récurrence sur i que $m_i^{\mathrm{A}} = \top$, c'est-à-dire que pour tout ensemble de symboles de variables W_i contenant les variables libres de m_i, on a $m_i^{\mathrm{A}}(a) = \top$ pour tout $a \in \mathrm{A}^{\mathrm{W}_i}$. Il y a quatre cas :

a) Si $m_i \in \mathrm{T}$, alors $m_i^{\mathrm{A}} = \top$, par définition d'un modèle ;

b) Supposons que m_i se déduise d'une formule m_j (où $j < i$) par la règle de généralisation, c'est-à-dire que $m_i = \forall v m_j$, où v est un symbole de variable. Soit W_i un ensemble de symboles de variables contenant les variables libres de m_i et soit $a \in \mathrm{A}^{\mathrm{W}_i}$; posons $\mathrm{W}_i' = \mathrm{W}_i \cup \{v\}$. Par définition, $m_i^{\mathrm{A}}(a) = \top$ si et seulement si pour tout élément $a' \in \mathrm{A}^{\mathrm{W}_i'}$ qui ne diffère de a qu'en v, $m_j^{\mathrm{A}}(a') = \top$. Puisque W_i' contient les symboles de variables qui sont libres dans m_j, cela résuite de l'hypothèse de récurrence $m_j^{\mathrm{A}} = \top$;

c) Si m_i est une tautologie ou un axiome de quantificateur, alors $m_i^{\mathrm{A}} = \top$ en vertu du lemme 4.2.4 ;

d) Si m_i se déduit de formules m_j et m_k (où $1 \leqslant j, k < i$) par la règle *modus ponens*, on a $m_i^{\mathrm{A}} = \top$ d'après le lemme 4.2.3.

Par récurrence, on a donc $m^{\mathrm{A}} = m_n^{\mathrm{A}} = \top$, ainsi qu'il fallait démontrer. □

4.3. Théories cohérentes

Définition **(4.3.1)**. — *On dit qu'une théorie* T *est* cohérente *si la formule close* $\bot$ *n'est pas un théorème de* T.

Toute partie d'une théorie cohérente est cohérente.

D'après le théorème 4.2.2, toute théorie consistante est cohérente. Le théorème de complétude affirmera la réciproque.

Remarque **(4.3.2)**. — Soit T une théorie. Supposons qu'il existe une formule close m telle que $T \vdash m$ et $T \vdash \neg m$; alors, T n'est pas cohérente.

En effet, puisque T prouve m et $\neg m$, la tautologie

$$(m \Rightarrow (\neg m \Rightarrow \bot))$$

et la règle *modus ponens* entraînent (m est close) qu'elle démontre $(\neg m \Rightarrow \bot)$, puis $\bot$.

Inversement, supposons que T ne soit pas cohérente et soit m une formule close de L. Comme $(\bot \Rightarrow m)$ est une tautologie, la théorie T démontre m.

En revanche, si m n'est pas close, il est possible qu'une théorie cohérente T prouve m et $\neg m$. Ce n'est pas absurde car la satisfaction de m signifie la satisfaction de ses quantifications universelles. Supposant par exemple que m possède un seul symbole de variable libre, x, la satisfaction de m équivaut à celle de $\forall x m$, et la satisfaction de $\neg m$ équivaut à celle de $\forall x \neg m$. Ces deux formules $\forall x m$ et $\forall x \neg m$ ne sont pas contradictoires, mais elles impliquent la formule $\phi = \forall x \bot$. En particulier, la théorie T n'a pas de modèle non vide.

Exemple **(4.3.3)**. — La théorie vide est consistante, donc cohérente.

Proposition **(4.3.4)** (Finitude). — *Soit* T *une théorie. Pour toute formule* m *telle que* $T \vdash m$, *il existe une partie finie* T_0 *de* T *telle que* $T_0 \vdash m$.

Démonstration. — Soit $(m_1, \ldots, m_n)$ une démonstration formelle de m dans T. Soit T_0 l'ensemble des formules closes de T qui sont égales à l'une des m_i. Alors, $(m_1, \ldots, m_n)$ est une démonstration formelle de m dans T_0, ce qui prouve que $T_0 \vdash m$. □

Corollaire **(4.3.5)**. — *Soit* T *une théorie. Pour que* T *soit cohérente, il faut et il suffit que toute partie finie de* T *soit cohérente.*

Corollaire **(4.3.6)**. — *Soit* $(T_i)_{i \in I}$ *une famille de théories cohérentes, filtrante pour l'inclusion. Alors,* $T = \bigcup_{i \in I} T_i$ *est une théorie cohérente.*

Démonstration. — L'assertion est évidente si $I = \varnothing$, car la théorie vide est cohérente. Supposons donc $I \neq \varnothing$ et considérons une partie finie S de T. Il existe

un élément $i \in I$ tel que $S \subset T_i$. Comme T_i est cohérente, il en est de même de S. D'après la proposition 4.3.4, cela démontre que T est cohérente. □

Corollaire **(4.3.7)**. — *Pour toute théorie cohérente* T, *il existe une théorie cohérente maximale* T' *telle que* $T \subset T'$.

Démonstration. — Soit $\mathscr{C}$ l'ensemble des théories cohérentes, ordonné par l'inclusion. Démontrons que $\mathscr{C}$ est inductif. Soit $(T_i)_{i\in I}$ une famille totalement ordonnée de théories cohérentes; d'après le corollaire précédent, $T = \bigcup_{i\in I} T_i$ est une théorie cohérente qui contient chaque T_i. Cela prouve que toute partie totalement ordonnée de $\mathscr{C}$ est majorée; l'ensemble ordonné $\mathscr{C}$ est donc ordonné. Le corollaire résulte ainsi du théorème de Zorn (théorème 1.1.8). □

Proposition **(4.3.8)** (Déduction). — *Soit* T *une théorie, soit m une formule close et soit n une formule close. Les deux assertions suivantes sont équivalentes :*

(i) *La formule n possède une démonstration formelle dans* $T \cup \{m\}$;
(ii) *La formule* $(m \Rightarrow n)$ *possède une démonstration formelle dans* T.

En particulier, pour démontrer la formule $(m \Rightarrow n)$, on peut démontrer n en supposant m.

Démonstration. — L'implication (ii)⇒(i) est banale. Considérons une démonstration formelle $(m_1, \dots, m_s)$ de $(m \Rightarrow n)$ dans T et adjoignons-lui les deux formules m et n. La suite $(m_1, \dots, m_s, m, n)$ de formules de $T \cup \{m\}$ est bien une démonstration formelle de n dans $T \cup \{n\}$ puisque $m \in T \cup \{m\}$ et que n se déduit de $(m \Rightarrow n)$ et m par modus ponens.

Supposons maintenant que $T \cup \{m\} \vdash n$ et considérons une démonstration formelle $(m_1, \dots, m_s)$ de n dans $T \cup \{m\}$. Démontrons que $T \vdash (m \Rightarrow n)$. Soit $i \in \{1, \dots, s\}$ et analysons les divers cas possibles.

a) Supposons que $m_i \in T \cup \{m\}$. Si $m_i = m$, la formule $(m \Rightarrow m)$ est une tautologie, Sinon, on a $m_i \in T$ et on insère avant $(m \Rightarrow m_i)$ la formule m_i de T et la tautologie $(m_i \Rightarrow (m \Rightarrow m_i))$, dont la formule $(m \Rightarrow m_i)$ se déduit par *modus ponens*;

b) Supposons que m_i soit un axiome logique du langage L. On insère alors avant $(m \Rightarrow m_i)$ l'axiome logique m_i et la tautologie $(m_i \Rightarrow (m \Rightarrow m_i))$, de sorte que $(m \Rightarrow m_i)$ est démontré par *modus ponens*.

c) Supposons enfin que m_i se déduise par *modus ponens* des formules m_j et m_k, où $1 \leqslant j, k < i$. Quitte à échanger j et k, on a donc $m_k = (m_j \Rightarrow m_i)$; on insère alors avant la formule $(m \Rightarrow m_i)$ la tautologie

$$((m \Rightarrow m_j) \Rightarrow ((m \Rightarrow (m_j \Rightarrow m_i)) \Rightarrow (m \Rightarrow m_i))),$$

la formule $((m \Rightarrow (m_j \Rightarrow m_i)) \Rightarrow (m \Rightarrow m_i))$ qui s'en déduit par *modus ponens* avec $(m \Rightarrow m_j)$, de sorte que la formule $(m \Rightarrow m_i)$ se déduit par *modus ponens* de la précédente et de $(m \Rightarrow (m_j \Rightarrow m_i)) = (m \Rightarrow m_k)$.

La suite de formules ainsi obtenue est une démonstration formelle de $(m \Rightarrow m_s) = (m \Rightarrow n)$, d'où la proposition. □

Corollaire **(4.3.9)**. — *Soit* T *une théorie et soit* m *une formule close de* T. *Pour que* m *soit un théorème de* T, *il faut et il suffit que la théorie* $\mathrm{T} \cup \{\neg m\}$ *ne soit pas cohérente.*

Ce corollaire justifie le principe du *raisonnement par l'absurde :* pour démontrer m, on suppose $\neg m$ et on obtient une contradiction.

Démonstration. — Supposons que m est un théorème de T. Alors, m est un théorème de $\mathrm{T} \cup \{\neg m\}$, de même que $\neg m$. Cela prouve que m n'est pas cohérente.

Supposons maintenant que $\mathrm{T} \cup \{\neg m\}$ ne soit pas cohérente, de sorte que $\bot$ est un théorème de $\mathrm{T} \cup \{\neg m\}$. D'après la proposition 4.3.8, $(\neg m \Rightarrow \bot)$ est un théorème de T. En appliquant la règle *modus ponens* à ce théorème et à la tautologie $((\neg m \Rightarrow \bot) \Rightarrow m)$, on en déduit que m est un théorème de T. □

Exemple **(4.3.10)**. — Soit x un symbole de variable, soit p une formule n'ayant que x comme variable libre et soit q une formule close. Démontrons que si $\mathrm{T} \vdash \exists x p$ et $\mathrm{T} \vdash \forall x (p \Rightarrow q)$, alors $\mathrm{T} \vdash q$. La suite de formules ci-dessous démontre que les formules closes $\forall x \neg p$ et sa négation sont toutes deux démontrables dans la théorie

$$\mathrm{T}' = \mathrm{T} \cup \{\exists x p, \forall x (p \Rightarrow q), \neg q\} :$$

(1) $\forall x (p \Rightarrow q)$;
(2) $\neg q$;
(3) $(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$ — tautologie ;
(4) $\forall x (p \Rightarrow q) \Rightarrow \forall x (\neg q \Rightarrow \neg p)$ — application de l'axiome de quantificateur (4.1.4.2) ;
(5) $\forall x (\neg q \Rightarrow \neg p)$;

(6) $\neg q \Rightarrow \forall x \neg p$ — application de la variante (4.1.9.1) de l'axiome de quantificateur (4.1.4.2) ;

(7) $\forall x \neg p$ — application de *modus ponens* ;

(8) $\neg \forall x \neg p$ — équivalent de $\exists x p$.

Si une théorie démontre une formule close et sa négation, elle n'est pas cohérente, donc la théorie T′ n'est pas cohérente, ce qu'il fallait démontrer.

4.4. Théories henkiniennes

Définition **(4.4.1)**. — *On dit qu'une théorie* T *est* syntaxiquement complète *si elle est cohérente et si pour toute formule close* m*, on a ou bien* $\mathrm{T} \vdash m$*, ou bien* $\mathrm{T} \vdash \neg m$.

Une théorie complète est syntaxiquement complète.

Une théorie cohérente maximale est syntaxiquement complète. En effet, si T est cohérente et maximale, et si m est une formule close telle que $m \notin \mathrm{T}$, alors $\mathrm{T} \cup \{m\}$ n'est pas cohérente, donc $\mathrm{T} \vdash \neg m$. Par conséquent, il découle du corollaire 4.3.7 que toute théorie cohérente est contenue dans une théorie syntaxiquement complète.

Définition **(4.4.2)**. — *Soit* T *une théorie dans un langage* L*. On dit que* T *est* henkinienne *si les deux propriétés suivantes sont satisfaites :*

a) *La théorie* T *est syntaxiquement complète.*

b) *Pour toute formule* $m(x)$ *en une seule variable libre* x *telle que* $\mathrm{T} \vdash \exists x m$*, il existe un terme clos* t *dans* L *tel que* $\mathrm{T} \vdash m(t)$.

En pratique, l'existence de tels termes clos sera garantie par l'existence, pour toute formule m telle que $\mathrm{T} \vdash \exists x m$, d'un symbole de constante c tel que $m(c)$ appartienne à T.

4.4.3. — Soit T une théorie henkinienne. On considère la relation suivante dans l'ensemble $\mathscr{T}_{\mathrm{L},\varnothing}$ des termes clos dans L : pour $a, b \in \mathscr{T}_{\mathrm{L},\varnothing}$, on pose $a \equiv b$ si et seulement si $\mathrm{T} \vdash (a = b)$. Les axiomes de l'égalité entraînent que c'est une relation d'équivalence. On pose alors $\mathrm{A} = \mathscr{T}_{\mathrm{L},\varnothing} / \equiv$; on note $[a]$ la classe dans A d'un élément $a \in \mathscr{T}_{\mathrm{L},\varnothing}$.

Lemme **(4.4.4)**. — a) *Soit n un entier, soit f un symbole de fonction n-aire, soit $t_1, \dots, t_n, u_1, \dots, u_n$ des termes clos. Supposons que* $\mathrm{T} \vdash t_i = u_i$ *pour tout* $i \in \{1, \dots, n\}$*; alors* $\mathrm{T} \vdash f t_1 \dots t_n = f u_1 \dots u_n$.

b) *Soit n un entier, soit r un symbole de relation n-aire, soit $t_1, \dots, t_n, u_1, \dots, u_n$ des termes clos. Supposons que* $\mathrm{T} \vdash t_i = u_i$ *pour tout* $i \in \{1, \dots, n\}$. *Alors* $\mathrm{T} \vdash (r t_1 \dots t_n \Leftrightarrow r u_1 \dots u_n)$.

Démonstration. — Cela découle des axiomes de l'égalité. Traitons par exemple le cas des fonctions unaire. Soit t, u des termes clos. Remarquons que la suite de formules suivante est une démonstration de $ft = fu$ dans la théorie $\mathrm{T} \cup \{t = u\}$:

a) $t = u$;
b) $(x = y) \Rightarrow (fx = fy)$ — axiome de l'égalité;
c) $(t = y) \Rightarrow (ft = fy)$ — substitution de x par t dans (2);
d) $(t = u) \Rightarrow (ft = fu)$ — substitution de y par u dans (3);
e) $ft = fu$ — modus ponens, (1) et (4).

D'après le théorème de déduction si l'on a $\mathrm{T} \vdash t = u$, alors on a $\mathrm{T} \vdash ft = fu$. Les autres cas sont analogues. □

4.4.4.1. — Soit n un entier et soit f un symbole de fonction n-aire. D'après le lemme précédent, il existe une unique application $f^{\mathrm{A}} : \mathrm{A}^n \to \mathrm{A}$ telle que

$$f^{\mathrm{A}}([t_1], \dots, [t_n]) = [f t_1 \dots t_n]^{\mathrm{A}}$$

pour tous $t_1, \dots, t_n \in \mathscr{T}_{\mathrm{L},\varnothing}$.

De même, soit n un entier et soit r un symbole de relation n-aire. D'après le lemme précédent, il existe une unique relation n-aire $r^{\mathrm{A}} : \mathrm{A}^n \to \{\top, \bot\}$ dans A^n telle que

$$r^{\mathrm{A}}([t_1], \dots, [t_n]) = \top \quad \text{si et seulement si} \mathrm{T} \vdash r t_1 \dots t_n.$$

Nous avons ainsi fait l'ensemble A une réalisation du langage L. On dit que c'est la *réalisation canonique* associée à la théorie henkinienne T.

Lemme **(4.4.5)**. — *Pour tout terme $t \in \mathscr{T}_{\mathrm{L},\varnothing}$, on a $t^{\mathrm{A}} = [t]$.*

Démonstration. — On raisonne par récurrence sur la définition d'un terme. Puisque t est clos, il existe un entier $n \geqslant 0$, un symbole de fonction n-aire, f, et des termes $t_1, \dots, t_n$ tels que $t = f t_1 \dots t_n$. Comme t est clos, les termes t_i sont clos. Par définition de l'évaluation, on a $t^{\mathrm{A}} = f^{\mathrm{A}}(t_1^{\mathrm{A}}, \dots, t_n^{\mathrm{A}})$. Par récurrence, on

a $t_i^{\mathrm{A}} = [t_i]$ pour tout i. Par suite, on a $t^{\mathrm{A}} = f^{\mathrm{A}}([t_1], \ldots, [t_n]) = [f t_1 \ldots t_n] = [t]$, ce qu'il fallait démontrer. □

Théorème **(4.4.6)** (Henkin). — *La réalisation canonique associée à une théorie henkinienne est un modèle de cette théorie.*

Démonstration. — Soit T une théorie henkinienne et soit A la réalisation canonique qui lui est associée. Soit $m \in \mathscr{F}_{\mathrm{L}}$ une formule close de L ; démontrons que l'on a $\mathrm{T} \vdash m$ si et seulement si $\mathrm{A} \vDash m$.

On raisonne par récurrence sur la définition de m. Précisément, on suppose que la propriété est vraie pour toute formule close qui, soit possède strictement moins de quantificateurs que m, soit soit strictement plus courte que m.

a) Supposons d'abord que m soit une formule atomique. Il existe donc un entier n, un symbole de relation n-aire, r, et des termes $t_1, \ldots, t_n$ tels que $m = r t_1 \ldots t_n$. Puisque m est une formule close, les termes t_i sont clos. Par définition de l'évaluation de m dans la structure A, on a $m^{\mathrm{A}} = r^{\mathrm{A}}(t_1^{\mathrm{A}}, \ldots, t_n^{\mathrm{A}})$. D'après le lemme précédent, on a $t_i^{\mathrm{A}} = [t_i]$ pour tout i ; par suite, $m^{\mathrm{A}} = r^{\mathrm{A}}([t_1], \ldots, [t_n])$. Cela prouve que $m^{\mathrm{A}} = \top$ si et seulement si $\mathrm{T} \vdash r t_1 \ldots t_n$, c'est-à-dire si et seulement si $\mathrm{T} \vdash m$.

b) Supposons maintenant que m soit de la forme $\neg m'$, où m' est une formule. Puisque T est cohérente, $\mathrm{T} \vdash m$ équivaut à $\mathrm{T} \nvdash m'$. Par récurrence, $\mathrm{T} \vdash m'$ si et seulement si $(m')^{\mathrm{A}} = \top$, donc $\mathrm{T} \nvdash m'$ équivaut à $(m')^{\mathrm{A}} = \bot$, donc à $m^{\mathrm{A}} = \top$ par définition de l'évaluation. ce que $m = \neg m'$ ne soit pas démontrable

c) Supposons maintenant que m soit de la forme $(m' \wedge m'')$, où m' et m'' sont des formules ; comme m est une formule close, m' et m'' sont closes. Par définition de l'évaluation, $m^{\mathrm{A}} = \top$ si et seulement si $(m')^{\mathrm{A}} = (m'')^{\mathrm{A}} = \top$; par récurrence, $m^{\mathrm{A}} = \top$ équivaut donc à ce que m' et m'' soient démontrables dans T.

Pour conclure, on observe que $(m' \wedge m'')$ est démontrable dans T si et seulement si m' et m'' le sont. En effet, la suite

(1) $(m' \wedge m'')$;
(2) $((m' \wedge m'') \Rightarrow m')$ — tautologie ;
(3) m' — *modus ponens*

est une démonstration de m' dans la théorie $\mathrm{T} \cup \{(m' \wedge m'')\}$. D'après l'exemple 4.1.8, m' est un théorème de T si $(m' \wedge m'')$ l'est ; on raisonne de même pour m''. Inversement, la suite

(1) m' ;

(2) m'' ;
(3) $(m' \Rightarrow (m'' \Rightarrow (m' \wedge m'')))$ — tautologie ;
(4) $(m'' \Rightarrow (m' \wedge m''))$ — modus ponens (1) et (3) ;
(5) $(m' \wedge m'')$ — modus ponens (2) et (4)

est une démonstration de $(m' \wedge m'')$ dans $\mathrm{T} \cup \{m', m''\}$. Ainsi, $(m' \wedge m'')$ est démontrable dans T si m' et m'' le sont.

d) Supposons que $m = (m' \vee m'')$, où m' et m'' sont des formules ; elles sont closes, car m est une formule close. Par définition de l'évaluation et par récurrence, $m^{\mathrm{A}} = \top$ si et seulement si au moins une des deux formules m' et m'' est démontrable dans T.

Pour conclure, on observe que m est démontrable dans T si et seulement si au moins l'une des deux formules m' et m'' l'est. La suite

(1) m' ;
(2) $(m' \Rightarrow (m' \vee m''))$ — tautologie ;
(3) $(m' \vee m'')$ — modus ponens (1) et (2)

est une démonstration de m dans $\mathrm{T} \cup \{m'\}$, si bien que $\mathrm{T} \vdash m$ si $\mathrm{T} \vdash m'$. On démontre de même que $\mathrm{T} \vdash m$ si $\mathrm{T} \vdash m''$.

Supposons inversement que ni m', ni m'' ne soient démontrables dans T. Comme T est syntaxiquement complète, les formules $\neg m'$ et $\neg m''$ sont des théorèmes de T. Comme la suite

(1) $\neg m'$;
(2) $\neg m''$;
(3) $(\neg m' \Rightarrow (\neg m'' \Rightarrow \neg(m' \vee m')))$ — tautologie ;
(4) $(\neg m'' \Rightarrow \neg(m' \vee m''))$ — modus ponens (1) et (3) ;
(5) $\neg(m' \vee m'')$ — modus ponens (2) et (4)

est une démonstration de $\neg m$ dans $\mathrm{T} \cup \{\neg m', \neg m''\}$, la formule $\neg m$ est un théorème de T.

e) Les deux cas $m = (m' \Rightarrow m'')$ et $m = (m' \Leftrightarrow m'')$ sont laissés au lecteur.

f) Supposons maintenant qu'il existe une formule m' et un symbole de variable x tels que $m = \exists x m'$. Comme m est une formule close, l'ensemble des variables libres de m' est contenu dans $\{x\}$.

Supposons que $\mathrm{T} \vdash m$. Par définition d'une théorie henkinienne, il existe un terme clos $t \in \mathscr{T}_{\mathrm{L},\varnothing}$ tel que $\mathrm{T} \vdash m'(t)$. La formule $m'(t)$ ayant un quantificateur de moins que la formule m, l'hypothèse de récurrence affirme que $(m'(t))^{\mathrm{A}} = \top$.

Puisque le terme t est clos, la substitution de x par t est admissible dans m', et l'on a $(m')^{\mathrm{A}}(t^{\mathrm{A}}) = \top$. Cela prouve que $m^{\mathrm{A}} = (\exists x m)^{\mathrm{A}} = \top$.

Supposons inversement que $\mathrm{T} \nvdash m$. Puisque T est syntaxiquement complète, la formule $\neg m = \neg\exists x m'$ est un théorème de T. La suite

(1) $(\exists x m' \Leftrightarrow \neg\forall x \neg m')$ — axiome de quantificateur (4.1.4.1) ;

(2) $(\neg\exists x m')$;

(3) $(\neg p \Rightarrow ((p \Leftrightarrow \neg q) \Rightarrow q))$ — tautologie, appliquée avec $p = \exists x m'$ et $q = \forall x \neg m'$;

(4) $((p \Leftrightarrow \neg q) \Rightarrow q)$, où $p = \exists x m'$ et $q = \forall x \neg m'$ — modus ponens, (2) et (3) ;

(5) $\forall x \neg m'$ — modus ponens, (1) et (4)

est une démonstration formelle de $\forall x \neg m'$ dans $\mathrm{T} \cup \{\neg m\}$, ce qui démontre que $\forall x \neg m'$ est un théorème de T. Démontrons maintenant que $m^{\mathrm{A}} = \bot$; il s'agit de prouver que $(m')^{\mathrm{A}}(a) = \bot$ pour tout $a \in \mathrm{A}$. Soit donc $a \in \mathrm{A}$; par définition de A, il existe un terme clos t tel que $a = [t] = t^{\mathrm{A}}$. Par application de l'axiome de quantificateur (4.1.4.3), la formule $\neg m'(t)$ est encore un théorème de T, donc $(\neg m'(t))^{\mathrm{A}} = \top$ par récurrence, et donc $(m'(t))^{\mathrm{A}} = \bot$. Par suite, $(m')^{\mathrm{A}}(a) = (m')^{\mathrm{A}}(t^{\mathrm{A}}) = (m'(t))^{\mathrm{A}} = \neg(\neg m'(t))^{\mathrm{A}} = \bot$, comme il fallait démontrer.

g) Supposons enfin qu'il existe une formule m' et un symbole de variable x tels que $m = \forall x m'$. Comme m est une formule close, l'ensemble des variables libres de m' est contenu dans $\{x\}$.

Supposons que m est un théorème de T. Pour tout terme clos t, la formule $m'(t)$ déduite de m' par substitution de x par t est donc un théorème de T ; par récurrence, $(m'(t))^{\mathrm{A}} = \top$. Puisque le terme t est clos, la formule m' est admissible pour la substitution qui applique x sur t et l'on a $(m')^{\mathrm{A}}(t^{\mathrm{A}}) = (m'(t))^{\mathrm{A}} = \top$. Comme tout élément de A est la classe d'un tel terme, on a $(m')^{\mathrm{A}}(a) = \top$ pour tout $a \in \mathrm{A}$, ce qui prouve que $m^{\mathrm{A}} = \top$.

Supposons inversement que $m = \forall x m'$ n'est pas un théorème de T. Puisque $(\neg\neg m' \Rightarrow m')$ est une tautologie et que T est cohérente, la formule $\forall x \neg\neg m'$ n'est pas un théorème de T (axiome (4.1.9.1)). Puisque T est cohérente, la formule $\neg\forall x \neg\neg m'$ est donc un théorème de T. Par application de l'axiome de quantificateur (4.1.4.1), il en est de même de la formule $\exists x \neg m'$. On applique alors à la formule $\neg m'$ l'argument développé ci-dessus. Comme T est henkinienne, il existe un terme clos t tel que $\neg m'(t)$ soit un théorème de T. Par récurrence, on a

donc $(\neg m'(t))^{\mathrm{A}} = \top$, d'où $(m')^{\mathrm{A}}(t^{\mathrm{A}}) = (m'(t))^{\mathrm{A}} = \neg(\neg m'(t))^2 = \bot$. L'élément $a = t^{\mathrm{A}}$ de A vérifie donc $(m')^{\mathrm{A}}(a) = \bot$, si bien que $m^{\mathrm{A}} = (\forall x m')^{\mathrm{A}} = \bot$. □

Corollaire **(4.4.7)**. — *Toute théorie henkinienne a un modèle.*

4.5. Le théorème de complétude de Gödel

Lemme **(4.5.1)**. — *Soit* T *une théorie dans un langage* L*, soit x un symbole de variable, soit m une formule de* L *dont la seule variable libre est x. Soit c un symbole de constante n'appartenant pas à* L*. On suppose que* $\mathrm{T} \vdash m(c)$ *dans le langage* $\mathrm{L} \cup \{c\}$*. Alors,* $\mathrm{T} \vdash \forall x m$ *dans le langage* L*.*

Autrement dit, si l'on sait prouver $m(c)$ sans rien savoir de c, c'est qu'on sait prouver $m(x)$ pour tout x, sans faire intervenir c.

Démonstration. — Prouvons $\mathrm{T} \vdash \forall x m$. Soit $(m_1, \ldots, m_n)$ une démonstration formelle de m. Choisissons un symbole de variable auxiliaire z qui n'apparaît dans aucune des formules $m_1, \ldots, m_n$. Quitte à changer les variables, on se ramène au cas où $z = x$.

Si p est une formule de L, on notera p' la formule déduite de p en remplaçant chaque occurrence du symbole de constante c par x.

Si p est un axiome de quantificateur, une tautologie, ou un des axiomes de l'égalité, il en est de même de p'. Par généralisation, $\forall x p'$ est démontrable dans T.

Si p est un axiome de généralisation $m \Rightarrow \forall v m$, alors $p' = (m' \Rightarrow \forall v m')$ en est un également.

Supposons que p se déduise par *modus ponens* de formules q et $(q \Rightarrow p)$ telles que $\forall x q'$ et $\forall x(q \Rightarrow p)' = \forall x(q' \Rightarrow p')$ soient démontrables dans T. Alors $\forall x q' \Rightarrow \forall x p'$ est démontrable dans T. Par hypothèse, toute variable libre de q a au moins une occurrence libre dans p. Les variables libres de p' sont celles de p, ainsi que x si c apparaît dans p, donc celles de $\forall x p'$ sont celles de p ; de même pour $\forall x q'$. On peut donc appliquer la règle *modus ponens* et on en déduit que $\forall x p'$ est démontrable dans T.

En appliquant le raisonnement précédent par récurrence à $m_1, \ldots, m_n$, on en déduit que $\forall x m_i'$ est démontrable dans T, pour tout T. Par suite, $\mathrm{T} \vdash \forall x m$. □

Proposition **(4.5.2)**. — *Soit* T *une théorie dans un langage* L*. Soit x un symbole de variable et soit* E *un ensemble de formules de m n'ayant que x comme variable libre ; supposons que pour tout $p \in$* E*, la formule close $\exists x p$ soit un théorème de* T*.*

Soit L' *le langage obtenu en adjoignant à* L, *pour toute formule* $p \in E$, *un symbole de constante* c_p. *Soit* T' *la théorie dans le langage* L' *obtenue en adjoignant à* T *les formules* $p(c_p)$, *pour* p *parcourant* E.

La théorie T *est cohérente si et seulement si la théorie* T' *est cohérente.*

Démonstration. — Si la théorie T' est cohérente, alors la théorie T l'est également puisque $T \subset T'$. Supposons donc que T soit cohérente et que T' soit incohérente. Il existe donc une démonstration formelle de $\bot$ dans T'. Comme une démonstration formelle de T' ne fait intervenir qu'un nombre fini de formules, on peut supposer que l'ensemble E est fini, puis raisonner par récurrence sur $\mathrm{Card}(E)$. On est ainsi ramené au cas où $E = \{p\}$. Par hypothèse, $T \cup \{p(c)\} \vdash \bot$ (dans le langage L'). D'après le théorème de déduction, $T \vdash (p(c) \Rightarrow \bot)$. D'après le lemme 4.5.1, il existe une démonstration formelle de $\forall x(p \Rightarrow \bot)$ dans la théorie T et le langage L.

Il reste à observer que si $T \vdash \exists x\, p$ et $T \vdash \forall x(p \Rightarrow \bot)$, alors $T \vdash \bot$, ce qui est un cas particulier de l'exemple 4.3.10. □

Théorème **(4.5.3)** (Gödel). — *Toute théorie cohérente possède un modèle.*

C'est le *théorème de complétude* de Gödel (1930). Il signifie en effet que les règles de déduction qui interviennent dans une démonstration formelle sont complètes, au sens où elles permettent de déduire tout ce qui peut l'être.

Démonstration. — Soit T une théorie cohérente. On définit de la façon suivante une suite (L_n) de langages et une suite (T_n) de théories syntaxiquement complètes où, pour tout $n \geqslant 1$, L_{n+1} est une extension du langage L_n et T_{n+1} est une théorie dans le langage L_{n+1} qui contient T_n. On choisit pour T_0 une complétion syntaxique de T. Supposons T_n définie. Notons alors E_n l'ensemble des formules p de L_n en une seule variable libre x telles que $T_n \vdash \exists x\, p$. Soit L_{n+1} le langage obtenu en adjoignant à L_n des symboles de constantes c_p, pour $p \in E_n$. Soit T'_{n+1} la théorie obtenue en adjoignant à T_n les formules $p(c_p)$, pour $p \in E_n$. D'après la proposition 4.5.2, la théorie T'_{n+1} est cohérente. Soit T_{n+1} une complétion syntaxique de T'_{n+1}.

Soit L' le langage réunion des langages L_n et soit T' la théorie réunion des théories T dans le langage L'. Démontrons que T' est une théorie henkinienne :

– C'est une réunion croissante de théories cohérentes, donc elle est cohérente.

– Prouvons qu'elle est syntaxiquement complète. Soit m une formule close dans le langage L'. Il existe un entier n tel que m soit une formule du langage L_n.

Puisque T_n est syntaxiquement clos, on a $T_n \vdash m$ ou $T_n \vdash \neg m$. Par suite, $T' \vdash m$ ou $T' \vdash \neg m$.

– Démontrons enfin que pour toute formule p du langage L' en une seule variable libre x telle que $T' \vdash \exists x\, p$, il existe un terme clos t de L' tel que $p(t)$. Il existe un entier n tel que p soit une formule du langage L_n. Alors, c_p est un symbole de constante de L_{n+1} et la formule close $p(c_p)$ appartient à T'_{n+1}, donc à T_{n+1}, donc à T'.

D'après le théorème de Henkin (théorème 4.4.6), la théorie T' possède un modèle.

Soit A un modèle de T'. Comme le langage L' contient L, on déduit de A une réalisation du langage L en oubliant les symboles de constantes ajoutés. Soit m une formule close du langage L qui appartient à T. C'est une formule close de L' appartenant à T', donc elle est satisfaite dans A. Par suite, A est un modèle de T. □

Corollaire **(4.5.4)**. — *Soit* T *une théorie et soit* m *une formule close. Si* m *est satisfaite dans tout modèle de* T*, alors* m *est démontrable dans* T.

Démonstration. — Supposons que m ne soit pas démontrable dans T. D'après le corollaire 4.3.9, la théorie $T \cup \{\neg m\}$ est cohérente. D'après le théorème de complétude, elle possède un modèle. Dans ce modèle, la formule m n'est pas satisfaite. □

Corollaire **(4.5.5)** (Compacité). — *Soit* T *une théorie. Pour que* T *soit consistante, il faut et il suffit que toute partie finie de* T *soit consistante.*

Démonstration. — En effet, consistance et cohérence coïncident, en vertu du théorème de complétude. Le corollaire découle donc du théorème de finitude. □

BIBLIOGRAPHIE

N. Bourbaki (1970), *Éléments de mathématique. Théorie des ensembles*, Hermann, Paris.

R. Cori & D. Lascar (2003*a*), *Logique mathématique, tome 1. Calcul propositionnel, algèbre de Boole, calcul des prédicats*, Dunod.

R. Cori & D. Lascar (2003*b*), *Logique mathématique, tome 2. Fonctions récursives, théorie des modèles*, Dunod.

A. Doxiadis & C. H. Papadimitriou (2009), *Logicomix—An epic search for truth*, Bloomsbury Press, New York. Character design and drawings by Alecos Papadatos, color by Annie Di Donna.

B. Poizat (2000), *A course in model theory. An introduction to contemporary mathematical logic*, Universitext, Springer-Verlag, New York. Translated from the French by Moses Klein and revised by the author.

INDEX