Cours de logique mathématique

Martin Hils

15 janvier 2013

Table des matières

Chapitre 1

Compter à l'infini

Dans ce premier chapitre, les notions d'*ensemble* et d'*entier* sont prises dans leur sens naïf. La théorie des ordinaux et cardinaux, due à Cantor (fin du 19^{e} siècle) sera développée du point de vue naïf.

Voici les deux principes de la Théorie des ensembles naïve :

- *Extensionalité* : Deux ensembles contenant les mêmes éléments sont égaux.
- *Compréhension (globale)* : Pour toute propriété (raisonnable) P, la collection $\{a \mid a \text{ vérifie } P\}$ est donnée par un ensemble.

C'est le second principe qui est problématique.

Antinomie de Russell. *Soit* $a = \{x \mid x \notin x\}$. *Alors* $a \in a$ *ssi* $a \notin a$.

Grâce à cette antinomie montrant que la Théorie des ensembles naïve est contradictoire, des fondements rigoureuses de la Théorie des ensembles (et par conséquent de la mathématique en générale) étaient devenues nécessaires.

Dans le dernier chapitre de ce cours nous traitons le système d'axiomes ZFC (*axiomes de Zermelo-Fraenkel* plus l'*axiome du choix*) ; la plupart de ces axiomes consisteront en une forme restreinte du principe de compréhension. Nous y verrons que les notions et résultats de ce premier chapitre restent valides en *Théorie axiomatique des ensembles*, en n'utilisant que le système d'axiomes ZFC.

Notation. Si A, B sont des ensembles, $A \cup B$, $A \cap B$ et $A \setminus B$ désignent la *réunion*, l'*intersection* et la *différence* d'ensembles, respectivement. Si $(A_i)_{i \in I}$ est une famille d'ensembles, on note sa réunion par $\bigcup_{i \in I} A_i = \{x \mid x \in A_i \text{ pour un } i \in I\}$, et son intersection par $\bigcap_{i \in I} A_i$.

Nous écrivons $A \subseteq B$ si A est une *partie* de B, et $A \subset B$ si A est une *partie propre* de B. L'*ensemble des parties* de A est noté $\mathcal{P}(A)$.

1.1 Le Théorème de Cantor et le Théorème de Cantor-Bernstein

L'existence d'une fonction injective / surjective / bijective entre deux ensembles sera utilisée comme moyen pour comparer leur « taille ». Nous commençons avec deux premiers résultats qui vont dans ce sens.

Théorème 1.1.1 (Cantor)**.** *Soit A un ensemble. Il n'existe pas de surjection de A sur $\mathcal{P}(A)$.*

Démonstration. Soit $f : A \to \mathcal{P}(A)$ une appliction. On considère l'ensemble

$$B = \{x \in A \mid x \notin f(x)\}.$$

Pour tout $x \in A$ avec $f(x) = B$, on a $x \in B$ si et seulement si $x \notin B$. Donc B n'est pas dans l'image de f. □

Théorème 1.1.2 (Cantor-Bernstein)**.** *Soient A et B deux ensembles, $f : A \to B$ et $g : B \to A$ deux injections. Alors il existe une bijection $h : B \to A$.*

Démonstration. On peut supposer que A est une partie de B et que l'application f est donnée par l'inclusion. (Il suffit de remplacer g par $f \circ g$ et A par $f(A)$.) Soit $C = \{g^n(x) \mid n \in \mathbb{N}, x \in B \setminus A\}$. On pose $h(c) = g(c)$ pour $c \in C$ et $h(x) = x$ pour $x \in B \setminus C$. □

Définition. Soient X et Y deux ensembles. On dit que X et Y sont *équipotents*, noté $X \sim Y$, s'il existe une bijection entre X et Y ; on dit que X est *subpotent* à Y, noté $X \preccurlyeq Y$, s'il existe une injection de X dans Y.

Dans cette terminologie, le théorème de Cantor-Bernstein dit que si $X \preccurlyeq Y$ et $Y \preccurlyeq X$, alors $X \sim Y$.

Exercice 1.1.3. Montrer qu'il existe une bijection entre $\mathbb{R}$ et $\mathcal{P}(\mathbb{N})$.

1.2 Notions d'ordre

Définition. Un *ordre partiel* $<$ sur un ensemble X est une relation binaire (c.-à-d. donnée par une partie de $X \times X$) qui est *transitive* (si $x < y$ et $y < z$ alors $x < z$) et *antiréflexive* ($x \not< x$). Si de plus $<$ est une relation *connexe* (pour tout $x, y \in X$ on a $x < y$, $x = y$ ou $y < x$) on dit que $<$ est un *ordre total*.

On note $x \leq y$ pour $x < y$ ou $x = y$, puis $x > y$ pour $y < x$, et $x \geq y$ pour $y \leq x$.

Si $Y \subseteq X$, $y \in Y$ est un *plus petit élément* si pour tout y' dans Y, $y \leq y'$. C'est un *élément minimal* si pour tout y' dans Y, $y' \not< y$. On définit de même un *plus grand élément* et un *élément maximal*. Un *minorant* de Y est un élément de X qui est $\leq$ à tous les éléments de Y. Une *borne inférieure* de Y est un plus

grand élément parmi les minorants. On définit de même la notion de *majorant* et de *borne supérieure*.

Noter que, dans un order partiel, $a \leq b$ et $b \leq a$ entraîne $a = b$. Un plus petit (grand) élément est donc nécessairement unique, ce qui n'est pas le cas pour un élément minimal (maximal) en général.

Remarque 1.2.1. *Si $<$ est un ordre partiel alors $\leq$ est une relation* réflexive *($x \leq x$ pour tout $x \in X$), transitive et* faiblement antisymétrique *($x \leq y$ et $y \leq x$ entraîne $x = y$).*

Réciproquement, si $\leq$ est une relation réflexive, transitive et faiblement antisymétrique, la relation $<$ définie par $x < y :\Leftrightarrow (x \leq y$ et $x \neq y)$ est un ordre partiel.

Démonstration. Exercice. □

Définition.
- Soit $<$ un ordre partiel sur X. On dit que $<$ est *bien-fondé* si toute partie non vide de X contient un élément minimal.
- Un *bon ordre* est un ordre total bien-fondé.

Exemples 1.2.2.
1. L'ordre usuel $<$ sur $\mathbb{N}$ est un bon ordre, tandis qu'il définit un ordre total non bien-fondé sur $\mathbb{Z}$.
2. Pour tout ensemble X, la relation $\subset$ est un ordre partiel sur $\mathcal{P}(X)$ qui est bien-fondé si et seulement si X est fini[1].
3. La restriction d'un ordre partiel (total, bien-fondé) sur X à une partie $Y \subseteq X$ est un ordre partiel (total, bien-fondé) sur Y.

Remarque 1.2.3. *Soit $<$ un ordre partiel sur X.*

1. *L'application $a \mapsto X_{\leq a} = \{x \in X \mid x \leq a\}$ identifie $(X, <)$ à une partie Y de $\mathcal{P}(X)$ avec l'ordre partiel sur Y induit par $\subset$.*
2. *$<$ est bien-fondé si et seulement s'il n'existe pas de suite infinie strictement décroissante dans X.*
3. *$<$ est un bon ordre si et seulement si toute partie non vide de X contient un plus petit élément.*

Démonstration. Les preuves de (1) et (3) sont laissées en exercice.

Montrons (2). Si $(X, <)$ n'est pas bien-fondé, il existe une partie $\emptyset \neq Y \subseteq X$ sans élément minimal. Par induction on peut donc choisir des éléments $y_n \in Y$ tels que $y_{n+1} < y_n$. Réciproquement, si $(y_n)_{n\in\mathbb{N}}$ est une suite strictement décroissante, il est clair que $Y = \{y_n \mid n \in \mathbb{N}\}$ ne contient pas d'élément minimal. □

1.3 Opérations sur les ordres

Définition. Soient X et Y deux ensembles partiellement ordonnés.

[1] Par un ensemble *fini* nous entendons un ensemble dans lequel on ne peut pas injecter $\mathbb{N}$.

– La *somme ordonnée* de X et Y, notée $X+Y$, est l'ensemble ordonné formé des paires $(x,0)$ avec $x \in X$ et $(y,1)$ avec $y \in Y$ et où l'ordre est défini ainsi : $(a,i) < (b,j)$ si $i < j$ ou si $i = j$ et $a < b$.
– Le *produit lexicographique* de X et Y est obtenu en munissant le produit cartésien $X \times Y$ de l'ordre suivant : $(x,y) < (x',y')$ si $y < y'$ ou si $y = y'$ et $x < x'$. On le note également $X \times Y$.

Lemme 1.3.1. *1. La somme ordonnée d'ordres totaux (resp. bien-fondés) est un ordre total (resp. bien-fondé).*

2. Le produit lexicographique d'ordres totaux (resp. bien-fondés) est un ordre total (resp. bien-fondé).

3. Soient X, Y et Z des ensembles partiellement ordonnés. On a les isomorphismes suivants d'ensembles ordonnés.

(a) $(X+Y)+Z \cong X+(Y+Z)$

(b) $(X \times Y) \times Z \cong X \times (Y \times Z)$

(c) $X \times (Y+Z) \cong (X \times Y)+(X \times Z)$

Démonstration. Le seul point non trivial est de montrer que si deux ordres sont bien-fondés, alors leur produit lexicographique l'est aussi. Soient donc X et Y des ensembles ordonnés bien-fondés. Soit Z un sous-ensemble non vide de $X \times Y$. On dénote $\pi : X \times Y \to Y$ la projection sur la deuxième coordonnée. Comme l'ordre sur Y est bien-fondé, il existe un élément minimal y_0 dans $\pi(Z) \subseteq Y$. Comme l'ordre sur X est bien-fondé, il existe x_0 minimal dans $Z_{y_0} = \{x \in X \mid (x,y_0) \in Z\}$. Il est clair que (x_0,y_0) est minimal dans Z. □

Définition. Soient X et Y des ensembles totalement ordonnés. On suppose que X possède un plus petit élément 0. On définit l'*exponentiation*, notée $X^{(Y)}$, comme suit. Comme ensemble, il s'agit des suites à support fini, c'est-à-dire le sous-ensemble de X^Y formé des applications $f : Y \to X$ avec $\mathrm{supp}(f) = \{y \in Y \mid f(y) \neq 0\}$ fini. On pose $f < g$ s'il existe $y \in Y$ tel que $f(y) < g(y)$ et $f(y') = g(y')$ pour tout $y' > y$.

Proposition 1.3.2. *1. La relation $<$ définit un ordre total sur $X^{(Y)}$ qui est bien-fondé si les ordres sur X et Y le sont.*

2. On a les isomorphismes $X^{(Y+Z)} \cong X^{(Y)} \times X^{(Z)}$ et $X^{(Y \times Z)} \cong \left(X^{(Y)}\right)^{(Z)}$.

Démonstration. L'énoncé sur les isomorphismes est une conséquence directe des définitions. Il est également facile à voir qu'il s'agit d'un ordre total.

Supposons que les ordres sur X et Y soient bien-fondés (ce sont donc des bons ordres). Soit Z une partie non vide de $X^{(Y)}$. Il faut montrer que Z contient un plus petit élément. Si la fonction constante à 0 est dans Z, il n'y a rien à montrer. On peut donc supposer que $\mathrm{supp}(f) \neq \emptyset$ pour tout $f \in Z$. Soit $Y_1 = \{s_1(f) \mid f \in Z\}$, où $s_1(f) = \max(\mathrm{supp}(f))$. Soit y_1 le plus petit élément de Y_1, et $Z_1' = \{f \in Z \mid s_1(f) = y_1\}$. L'ensemble Z_1' est un *ségment initial* de Z, c'est-à-dire $f < g$ pour tout $f \in Z_1'$ et $g \in Z \setminus Z_1'$.

Soit x_1 le plus petit élément de $\{f(y_1) \mid f \in Z_1'\}$. On pose

$$Z_1 = \{f \in Z_1' \mid f(y_1) = x_1\}.$$

L'ensemble Z_1 est un ségment initial de Z_1'. Si Z_1 contient la fonction qui est constante 0 en dehors de y_1, on a terminé. Sinon, $\mathrm{supp}(f) \setminus \{y_1\} \neq \emptyset$ pour tout $f \in Z_1$. Soit $Y_2 = \{s_2(f) \mid f \in Z_1\}$, où $s_2(f) = \max(\mathrm{supp}(f) \setminus \{y_1\})$. Soit y_2 le plus petit élément de Y_2, puis x_2 le plus petit élément de $\{f(y_2) \mid f \in Z_1 \text{ et } y_2 = s_2(f)\}$. On pose $Z_2 = \{f \in Z_1 \mid s_2(f) = y_2 \text{ et } f(y_2) = x_2\}$. C'est un ségment initial de Z_1. Si Z_2 contient la fonction qui est constante 0 en dehors de $\{y_1, y_2\}$, on a terminé, sinon on continue de la même façon, construisant Y_3, y_3, Z_3', x_3, Z_3 et ainsi de suite. Comme la suite des y_i est strictement décroissante dans Y, ce procédé s'arrête après un nombre fini d'étapes. □

1.4 Ordinaux

Un ensemble X est *transitif* si pour tout $x \in X$ et $y \in x$ on a $y \in X$. (C'est équivalent à $x \in X \Rightarrow x \subseteq X$.

Définition. Un ensemble X est un *ordinal* s'il est transitif et si $\in\restriction_{X \times X}$ définit un bon ordre sur X.

Proposition 1.4.1. *Soient α et β des ordinaux.*

1. *$\emptyset$ est un ordinal.*
2. *Si $\alpha \neq \emptyset$, alors $\emptyset \in \alpha$.*
3. *$\alpha \notin \alpha$*
4. *Si $x \in \alpha$, alors $x = S_{<x} := \{y \in \alpha \mid y < x\}$.*
5. *Si $x \in \alpha$, alors x est un ordinal.*
6. *$\beta \subseteq \alpha$ si et seulement si $\beta \in \alpha$ ou $\beta = \alpha$.*
7. *$x := \alpha \cup \{\alpha\}$ est un ordinal. On le note α^+.*

Démonstration. (1) est clair. Quant à (2), on considère $x \in \alpha$ minimal. S'il existait $y \in x$, alors $y \in \alpha$ par transitivité de α, et x ne serait donc pas minimal. Dans (3), par antiréflexivité, on a $x \notin x$ pour tout $x \in \alpha$. Donc $\alpha \in \alpha$ entraîne $\alpha \notin \alpha$. (4) suit du fait que $<$ est donné par $\in$. Pour montrer (5), notons que $\in$ se restreint en un bon ordre sur x, car $x \subseteq \alpha$. De plus, $x = S_{<x}$ est un ensemble transitif, car $z \in y \in x \Rightarrow z < x \Rightarrow z \in S_{<x}$.

Pour montrer le sens direct dans (6), supposons que $\beta \subset \alpha$. Soit x minimal dans $\alpha \setminus \beta$. Clairement $\beta \supseteq S_{<x}$ par minimalité. Réciproquement, si $y \in \beta$, alors $y \in x$ car sinon $x \in y$ et $x \in \beta$. Donc $\beta = S_{<x} = x \in \alpha$. L'autre sens dans (6) est trivial, et (7) se vérifie immédiatement. □

Proposition 1.4.2. *Soit X un ensemble non vide d'ordinaux. Alors $\bigcap_{\alpha \in X} \alpha$ est un plus petit élément de X.*

Démonstration. L'intersection d'ensembles transitifs est transitive, et l'ordre induit sur un sous-ensemble d'un bon ordre est un bon ordre. Donc $\beta = \bigcap_{\alpha \in X} \alpha$ est un ordinal. On a $\beta \subseteq \alpha$ pour tout $\alpha \in X$. Si $\beta \notin X$, alors $\beta \in \alpha$ pour tout $\alpha \in X$, par Proposition 1.4.1(6). D'où $\beta \in \beta$, absurde. □

Théorème 1.4.3. *Soient α et β deux ordinaux. Une et une seule des trois propriétés suivantes est vérifiée :*

(1) $\alpha \in \beta$, (2) $\alpha = \beta$, (3) $\beta \in \alpha$.

Démonstration. On pose $X = \{\alpha, \beta\}$, et on applique Proposition 1.4.2. Si $\alpha \cap \beta = \alpha$, alors $\alpha \subseteq \beta$, d'où $\alpha = \beta$ ou $\alpha \in \beta$ par Proposition 1.4.1(6). De même, si $\alpha \cap \beta = \beta$, alors $\alpha = \beta$ ou $\beta \in \alpha$. L'exclusivité de ces propriétés est claire. □

Notation. Dans la suite, nous écrivons $\alpha < \beta$ pour $\alpha \in \beta$, et $\alpha \leq \beta$ pour $\alpha \subseteq \beta$, quand α et β sont des ordinaux.

Proposition 1.4.4. *Soit X un ensemble d'ordinaux. Alors $b = \bigcup_{\alpha \in X} \alpha$ est un ordinal. De plus, si $\gamma < b$, il existe $\alpha \in X$ tel que $\gamma \in \alpha$. On écrira aussi $b = \sup_{\alpha \in X} \alpha$.*

Démonstration. Comme b est réunion d'ensembles transitifs, il est transitif. De plus, b ne contient que des ordinaux. Par le Théorème 1.4.3, $\in$ induit un ordre total sur b. Si $\emptyset \neq Z \subseteq b$, alors $\bigcap_{\alpha \in Z} \alpha$ est un plus petit élément de Z par 1.4.2. Cela montre que cet ordre est bien-fondé. □

Un ordinal de la forme α^+ est appelé *ordinal successeur*. Il est clair que α^+ est le plus petit ordinal $> \alpha$, autrement dit le successeur de α.

Définition. Un *ordinal limite* est un ordinal non vide qui n'est pas successeur.

Proposition 1.4.5. *Pour un ordinal $\lambda \neq \emptyset$, sont équivalents :*

1. *λ est limite ;*
2. *$\lambda = \bigcup_{\alpha < \lambda} \alpha$.*

Démonstration. (1)⇒(2). Soit $\beta = \bigcup_{\alpha < \lambda} \alpha$ et λ limite. Il est clair que $\beta \subseteq \lambda$. Réciproquement, supposons $\alpha < \lambda$. Donc $\alpha^+ \leq \lambda$ et alors $\alpha^+ < \lambda$ car λ est limite. On conclut, puisque $\alpha \in \alpha^+ \subseteq \beta$.

(2)⇒(1). Si $\lambda = \gamma^+$, alors $\bigcup_{\alpha < \lambda} \alpha = \bigcup_{\alpha \leq \gamma} \alpha = \gamma < \lambda$. □

Exemples 1.4.6.

1. On peut retrouver les *entiers comme ordinaux* de la façon suivante. On pose $\underline{0} := \emptyset$, puis inductivement $\underline{n+1} := \underline{n}^+$ pour $n \in \mathbb{N}$. On a par exemple $\underline{1} = \{\emptyset\}$, $\underline{2} = \{\underline{0}, \underline{1}\} = \{\emptyset, \{\emptyset\}\}$, $\underline{3} = \{\underline{0}, \underline{1}, \underline{2}\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$.
   On montre par induction que $\underline{n}$ est un ordinal pour tout entier n. Dans la suite, on identifiera souvent $\underline{n}$ et n.
2. On note $\omega := \bigcup_{n \in \mathbb{N}} \underline{n}$. C'est un ordinal par 1.4.4.

Définition. On dit qu'un ordinal est *fini* si ni lui-même ni aucun de ses éléments n'est limite.

Proposition 1.4.7. *1. ω est l'ensemble des ordinaux finis.*

2. ω est le plus petit ordinal limite.

Démonstration. On montre d'abord, par induction sur $n \in \mathbb{N}$, que tous les éléments de ω sont des ordinaux finis. De plus, $\alpha < \omega$ entraîne $\alpha^+ < \omega$. Cela montre (2). Si $\alpha \notin \omega$, alors $\omega \leq \alpha$, donc ou bien $\alpha = \omega$ ou bien $\omega \in \alpha$. Dans les deux cas, α n'est pas fini. Cela montre (1). □

Théorème 1.4.8 (Classification des bons ordres par les ordinaux)**.**
Tout ensemble bien-ordonné X est isomorphe, comme ensemble ordonné, à un ordinal. De plus, l'ordinal et l'isomorphisme en question sont uniques.

Pour démontrer le théorème, nous avons besoin d'un lemme.

Lemme 1.4.9. *Soit $f : \alpha \to \alpha'$ une application strictement croissante entre deux ordinaux. Alors $f(\beta) \geq \beta$ pour tout $\beta \in \alpha$. En particulier, on a $\alpha \leq \alpha'$, et si f est un isomorphisme, alors $\alpha = \alpha'$ et f est égale à l'identité.*

Démonstration. S'il existe $\beta \in \alpha$ avec $f(\beta) < \beta$, on choisit β_0 minimal avec cette propriété. Comme f est strictement croissante, on a $f(f(\beta_0)) < f(\beta_0)$, ce qui contredit la minimalité. □

Démonstration du Théorème 1.4.8. L'unicité est une conséquence du lemme. Quant à l'existence, notons d'abord que pour tout $x \in X$, tout isomorphisme entre $S_{<x}$ et un ordinal α s'étend en un isomorphisme entre $S_{\leq x} = S_{<x} \cup \{x\}$ et α^+. Soit

$$Y = \{y \in X \mid \text{ il existe } f : S_{\leq y} \cong \alpha \text{ pour un ordinal } \alpha\}.$$

Par unicité, pour $y \in Y$, l'ordinal $\alpha = \alpha(y)$ et l'isomorphisme $f = f_y$ sont uniques. Nous allons montrer que $Y = X$. Sinon, il existe $x \in X$ minimal dans $X \setminus Y$. Pour $y < x$ on a $f_y : S_{\leq y} \cong \alpha(y)$. De plus, il s'agit d'un système cohérent d'isomorphismes, c'est-à-dire pour tout $y' < y < x$ on a $f_y \restriction_{S_{\leq y'}} = f_{y'}$. (Remarquons pour cela qu'un segment initial d'un ordinal est un ordinal.) On pose

$$\alpha = \sup_{y<x} \alpha(y) \text{ et } f : S_{<x} \to \alpha,\ f(y) := f_y(y).$$

Il est clair que f est bien définie et induit un isomorphisme d'ensembles ordonnés entre $S_{<x}$ et α. Par ce qui était dit au début, f s'étend en un ismorphisme entre $S_{\leq x}$ et α^+. Contradiction. On a donc bien $Y = X$. Pour conclure, il suffit d'employer un argument du même type, en posant $\alpha(X) := \sup_{x \in X} \alpha(x)$ et $f : X \cong \alpha(X)$, $x \mapsto f_x(x)$. □

Remarque 1.4.10 (Induction transfinie)**.**
Soit P une propriété qui porte sur les ordinaux. On suppose :

- *$\emptyset$ vérifie P ;*
- *pour tout ordinal α : si α vérifie P, alors α^+ vérifie P ;*
- *pour tout ordinal limite λ : si tout $\alpha < \lambda$ vérifie P, alors λ vérifie P.*

Alors tout ordinal vérifie P.

1.5 Arithmétique ordinale

Si α et β sont des ordinaux, on note $\alpha + \beta$ l'unique ordinal isomorphe à la somme ordonnée de α et β (qui existe par le Théorème 1.4.8). On définit de même $\alpha\beta$ comme l'unique ordinal isomorphe au produit lexicographique $\alpha \times \beta$ et α^β comme l'unique ordinal isomorphe à l'ensemble ordonné $\alpha^{(\beta)}$. Notons que 0^β n'est pas encore défini. On pose $0^0 := 1$, puis $0^\beta := 0$ pour tout $\beta > 0$.

Proposition 1.5.1 (Addition ordinale). *1.* $\alpha + 0 = 0 + \alpha = \alpha$

2. $\alpha + 1 = \alpha^+$
3. $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$, *en particulier* $\alpha + \beta^+ = (\alpha + \beta)^+$
4. $\alpha < \beta$ *si et seulement s'il existe* $\gamma > 0$ *tel que* $\beta = \alpha + \gamma$
5. *Si* $\beta < \beta'$, *alors* $\alpha + \beta < \alpha + \beta'$ *pour tout* α. *En particulier, on peut simplifier à gauche :* $\alpha + \beta = \alpha + \beta' \Rightarrow \beta = \beta'$.
6. *Si* λ *est limite, alors* $\alpha + \lambda = \sup_{\beta<\lambda}(\alpha + \beta)$ *(continuité).*
7. $1 + \alpha = \alpha + 1$ *si* α *est fini,* $1 + \alpha = \alpha$ *sinon.*

Démonstration. (1)–(3) sont clairs. Dans le sens non trivial de (4), on vérifie sans peine que l'ordinal γ isomorphe à l'ensemble bien-ordonné $\beta \setminus \alpha$ marche.

(5) Si $\beta < \beta'$, on utilise (2) et (4) pour obtenir $\beta' = \beta + \gamma$ et donc $\alpha + \beta' = (\alpha + \beta) + \gamma$ pour un $\gamma > 0$.

(6) $\alpha + \lambda \geq \sup_{\beta<\lambda}(\alpha + \beta)$ suit de (5). Réciproquement, si $\alpha \leq \mu < \alpha + \lambda$, alors $\mu = \alpha + \delta$ pour un δ avec $0 \leq \delta < \lambda$. Comme λ est limite, on a $\delta^+ < \lambda$, d'où $\mu < \alpha + \delta^+ \leq \sup_{\beta<\lambda}(\alpha + \beta)$.

(7) On montre par induction sur $n \in \mathbb{N}$ que $\underline{1}+\underline{n} = \underline{n}+\underline{1}$. Puis, on a $1+\omega = \omega$ par (6). Enfin, $\alpha \geq \omega$ s'écrit $\alpha = \omega+\beta$, d'où $1+\alpha = 1+\omega+\beta = \omega+\beta = \alpha$. □

Dans ce qui suit, nous omettons quelques parenthèses, en suivant la convention que l'exponentiation lie plus fortement que la multiplication et que la multiplication lie plus fortement que l'addition. Ainsi, il faudra par exemple lire $\alpha\beta + \gamma$ comme $(\alpha\beta) + \gamma$, et $\gamma\alpha^\beta$ comme $\gamma\left(\alpha^\beta\right)$.

Proposition 1.5.2 (Multiplication ordinale). *1.* $\alpha 0 = 0\alpha = 0$

2. $\alpha 1 = 1\alpha = \alpha$
3. $\alpha(\beta\gamma) = (\alpha\beta)\gamma$
4. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$, *en particulier* $\alpha\beta^+ = \alpha\beta + \alpha$
5. $2\omega = \omega < \omega 2 = \omega + \omega$
6. *On suppose* $\alpha \neq 0$. *Si* $\beta < \beta'$, *alors* $\alpha\beta < \alpha\beta'$. *En particulier, on peut simplifier à gauche :* $\alpha\beta = \alpha\beta' \Rightarrow \beta = \beta'$.
7. *Si* λ *est limite, alors* $\alpha\lambda = \sup_{\beta<\lambda} \alpha\beta$ *(continuité).*

Démonstration. (1) et (2) sont clairs, (3) et (4) sont conséquences du Lemme 1.3.1. Dans (5), $2\omega = \omega$ suit de (7), les autres parties de (5) sont simples. Quant à (6), il suffit de noter que si $\beta < \beta'$ alors $\beta' = \beta + \gamma$ pour un $\gamma > 0$, d'où $\alpha\beta' = \alpha\beta + \alpha\gamma$ par (4) et alors $\alpha\beta' > \alpha\beta$.

(7) On peut supposer $\alpha \neq 0$. Soit λ un ordinal limite. L'inégalité $\alpha\lambda \geq \sup_{\beta<\lambda} \alpha\beta =: \delta$ est clair par (6). Réciproquement, on se donne $\gamma < \alpha\lambda$. La division euclidienne, démontrée dans le lemme suivant, fournit une paire d'ordinaux (ρ, μ) telle que $\gamma = \alpha\mu + \rho$, avec $\rho < \alpha$. Comme $\mu < \lambda$ (par (6)), on a $\mu^+ < \lambda$ car λ est limite, d'où $\gamma = \alpha\mu + \rho < \alpha\mu + \alpha = \alpha\mu^+ \leq \delta$. □

Lemme 1.5.3 (Division euclidienne)**.** *Soient α et β des ordinaux, avec $\alpha \neq 0$. Alors il existe une unique paire d'ordinaux (ρ, μ) telle que $\rho < \alpha$ et $\beta = \alpha\mu + \rho$.*

Démonstration. Unicité : Supposons que $\alpha\mu + \rho = \alpha\mu' + \rho'$ avec $\rho, \rho' < \alpha$. Si $\mu < \mu'$, alors $\alpha\mu + \rho < \alpha\mu^+ \leq \alpha\mu' \leq \alpha\mu' + \rho'$, ce qui est absurde. Donc $\mu = \mu'$ par symétrie, et on obtient également $\rho = \rho'$ en simplifiant.

Existence : Si $\beta = 0$ il n'y a rien a montrer. Soit donc $\beta \neq 0$. L'application $f_0 : \beta \to \alpha \times \beta$, $x \mapsto (0, x)$ est strictement croissante, d'où $\beta \leq \alpha\beta$ par 1.4.9. Si $\beta = \alpha\beta$, on pose $\mu = \beta$ et $\rho = 0$. Sinon, on a $\beta \in \alpha\beta$. Soit f l'isomorphisme d'ensembles ordonnés entre $\alpha\beta$ et $\alpha \times \beta$. On pose $(\rho, \mu) = f(\beta)$. On vérfie que $\beta = \alpha\mu + \rho$ (exercice). □

Exercice 1.5.4.
1. Montrer que α est limite si et seulement s'il existe $\beta \neq 0$ tel que $\alpha = \omega\beta$.
2. Montrer que $\omega^2 = \omega\omega$ n'est pas de la forme $\delta + \omega$.

Proposition 1.5.5 (Exponentiation ordinale)**.**

1. *Pour tout α on a $\alpha^0 = 1$, $\alpha^1 = \alpha$ et $1^\alpha = 1$. Si $\alpha \neq 0$, alors $0^\alpha = 0$.*
2. *$\alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$, en particulier $\alpha^{\beta^+} = \alpha^\beta \alpha$.*
3. *$\left(\alpha^\beta\right)^\gamma = \alpha^{\beta\gamma}$*
4. *Si $\alpha > 1$, alors $\beta < \beta' \Rightarrow \alpha^\beta < \alpha^{\beta'}$.*
5. *Si λ est limite et $\alpha \neq 0$, alors $\alpha^\lambda = \sup_{\beta<\lambda} \alpha^\beta$ (continuité).*

Démonstration. (1) se vérifie directement, et (2) ainsi que (3) suivent de 1.3.2.

(4) $\beta < \beta' \Rightarrow \beta' = \beta + \rho$ pour un $\rho > 0$. Donc $\alpha^{\beta'} = \alpha^{\beta+\rho} = \alpha^\beta \alpha^\rho$. Or $\alpha^\rho > 1$ est clair (l'ensemble $\alpha^{(\rho)}$ contient au moins 2 éléments), d'où $\alpha^{\beta'} > \alpha^\beta$ par 1.5.2(6).

Montrons l'inégalité non triviale dans (5). On se donne $f \in \alpha^{(\lambda)}$. On peut supposer que f n'est pas la fonction constante 0. Alors $s_1(f) < \lambda$, d'où $\beta = s_1(f)^+ < \lambda$, ce qui montre que $S_{\leq f}$ admet une application strictement croissante dans $\alpha^{(\beta)}$. On conclut par le Lemme 1.4.9. □

Exercice 1.5.6. Soit $\alpha > 1$.

1. Montrer que $\alpha^\gamma \geq \gamma$ pour tout γ. (Est-ce qu'il y a des exemples avec $\alpha^\gamma = \gamma$?)
2. Soit $\beta > 0$. Montrer qu'il existe γ tel que $\alpha^\gamma \leq \beta < \alpha^{\gamma^+}$.

3. En déduire que tout ordinal β admet un développement en base α : il existe une suite finie d'ordinaux $\beta_1 > \ldots > \beta_n \geq 0$ et des ordinaux k_i tels avec $0 < k_i <$ *alpha* tels que

$$\beta = \alpha^{\beta_1} k_1 + \ldots + \alpha^{\beta_n} k_n.$$

De plus, l'entier n ainsi que les suites (β_i) et (k_i) sont uniques.

On appelle le développement en base ω la *forme normale de Cantor*.

Remarque 1.5.7. *Les formules suivantes permettraient de donner une autre définition (par récurrence transfinie) de l'addition, de la multiplication ainsi que de l'exponentiation ordinale :*

- $\alpha + 0 = \alpha$, $\alpha + \beta^+ = (\alpha + \beta)^+$, *puis* $\alpha + \lambda = \sup_{\beta<\lambda}(\alpha + \beta)$ *pour* λ *limite.*
- $\alpha 0 = 0$, $\alpha\beta^+ = \alpha\beta + \alpha$, *puis* $\alpha\lambda = \sup_{\beta<\lambda}(\alpha\beta)$ *pour* λ *limite.*
- *(*$\alpha \neq 0$*.)* $\alpha^0 = 1$, $\alpha^{\beta^+} = \alpha^\beta \alpha$, *puis* $\alpha^\lambda = \sup_{\beta<\lambda} \left(\alpha^\beta\right)$ *pour* λ *limite.*

Exercice 1.5.8 (Topologie de l'ordre)**.** Dans tout ordre total X, on peut définir une topologie, appelé *topologie de l'ordre*, dont une base d'ouverts est donnée par les intervalles ouverts, c'est-à-dire par les ensembles de la forme $(-\infty, b) = \{x \in X \mid x < b\}$, $(a, b) = \{x \in X \mid a < x < b\}$ ou (a, ∞), pour $a, b \in X$.

1. Montrer que pour X quelconque, cette topologie est séparée.
2. Soit maintenant α un ordinal muni de la topologie de l'ordre. Montrer :
   (a) La topologie sur α est discrète si et seulement si $\alpha \leq \omega$.
   (b) α est compact si et seulement s'il n'est pas limite.
   (c) Une application $f : \alpha \to \beta$ entre deux ordinaux et qui est faiblement croissante ($x \leq y \Rightarrow f(x) \leq f(y)$) est continue si et seulement si, pour tout $\lambda \in \alpha$ limite, on a $f(\lambda) = \sup_{\gamma<\lambda} f(\gamma)$.

1.6 Axiome du choix

Pour une famille d'ensembles $(X_i)_{i\in I}$ on pose

$$\prod_{i\in I} X_i = \{f : I \to \bigcup_{i\in I} X_i \mid f(i) \in X_i \text{ pour tout } i \in I\},$$

appelé le *produit* de la famille.

Définition. L'*Axiome du choix (AC)* dit que le produit d'une famille d'ensembles non vides est non vide, c'est-à-dire que si $X_i \neq \emptyset$ pour tout $i \in I$, alors $\prod_{i\in I} X_i \neq \emptyset$.

Dans le système d'axiomes de Zermelo-Fraenkel (ZF), (AC) est équivalent au *Lemme de Zorn* ainsi qu'au *Théorème de Zermelo*. Nous le démontrons au dernier chapitre du cours et l'admettons comme résultat pour l'instant.

Définition. Un ensemble partiellement ordonné X est *inductif* si toute partie $Y \subseteq X$ totalement ordonné possède un majorant (dans X). En particulier, un tel X n'est pas vide.

Lemme de Zorn. Tout ensemble partiellement ordonné inductif possède un élément maximal.

Théorème de Zermelo (Wohlordnungssatz)**.** Tout ensemble X admet un bon ordre.

1.7 Cardinaux

On suppose jusqu'à la fin de l'avant-dernier chapitre que l'axiome du choix est vérifié.

Définition. On appelle *cardinal* tout ordinal qui n'est pas équipotent à un ordinal plus petit.

Exemples 1.7.1.
1. Tout ordinal fini et un cardinal.
2. L'ordinal ω est un cardinal. Il sera noté $\aleph_0$.
3. Si α est infini, alors α^+ n'est pas un cardinal. (α^+ et α sont équipotents.)

Proposition 1.7.2. *Tout ensemble X est équipotent à un unique cardinal, noté* $\mathrm{card}(X)$.

Démonstration. Par le Théorème de Zermelo et 1.4.8, X est équipotent à un ordinal α. Soit $\beta \leq \alpha$ minimal tel que β soit équipotent à α. Alors β est un cardinal et en bijection avec X. L'unicité est claire. □

Proposition 1.7.3. *Soient X et Y deux ensembles avec X non vide. Sont équivalents :*

1. $\mathrm{card}(X) \leq \mathrm{card}(Y)$
2. *Il existe une injection de X dans Y.*
3. *Il existe une surjection de Y sur X.*

Démonstration. (1)⇒(2) est facile.

(2)⇒(3) : Soit $f : X \to Y$ une injection. On choisit $x_0 \in X$ et on obtient une surjection $g : Y \to X$, en posant $g(y) := x_0$ si $y \notin \mathrm{im}(f) = \{f(x) \mid x \in X\}$, et $g(y) := f^{-1}(y)$ sinon.

(3)⇒(1) : S'il existe une surjection de Y sur X, alors il existe une surjection $g : \lambda = \mathrm{card}(Y) \to \kappa = \mathrm{card}(X)$. L'application f qui à $\alpha \in \kappa$ associe l'élément minimal β dans λ avec $g(\beta) = \alpha$ définit une injection de κ dans λ. En particulier, κ est en bijection avec un ordinal $\gamma \leq \lambda$. (On prend γ l'unique ordinal qui est isomorphe au bon ordre induit sur $\mathrm{im}(f)$.) □

Définition. On appelle un ensemble X *dénombrable* si $\mathrm{card}(X) \leq \aleph_0$, et *fini* si $\mathrm{card}(X) < \aleph_0$.

Dans la suite, κ, λ etc. dénoteront des cardinaux.

Proposition 1.7.4. *Soit X un ensemble de cardinaux. Alors $\lambda = \sup_{\kappa \in X} \kappa$ est un cardinal.*

Démonstration. Si $\alpha < \lambda$, alors $\alpha < \kappa$ pour un $\kappa \in X$. Comme κ est un cardinal, on a $\kappa = \mathrm{card}(\kappa) \leq \mathrm{card}(\lambda)$, d'où $\alpha < \mathrm{card}(\lambda)$. Cela montre que λ n'est pas équipotent avec un ordinal plus petit. □

Il n'y a pas de plus grand cardinal. En effet, si κ est un cardinal, alors $\lambda := \mathrm{card}(\mathcal{P}(\kappa)) > \kappa$ par le théorème de Cantor. L'ensemble des cardinaux $\leq \lambda$ qui sont $> \kappa$ est donc non vide. On note κ^+ son plus petit élément, appelé *successeur cardinal* de κ. Pour éviter des confusions, le successeur ordinal de α sera noté $\alpha + 1$ dans la suite.

Définition. La *hiérarchie des* $\aleph$ est l'application des ordinaux dans les cardinaux qui est définie comme suit :

- $\aleph_0 := \omega$
- $\aleph_{\alpha+1} := \aleph_\alpha^+$
- $\aleph_\alpha := \sup_{\beta<\alpha} \aleph_\beta$, si α est un ordinal limite.

Par induction transfinie, on montre que $\alpha < \beta \Rightarrow \aleph_\alpha < \aleph_\beta$. En combinaison avec le résultat suivant, on en déduit que la hiérarchie des $\aleph$ établit une énumération strictement croissante des cardinaux infinis par les ordinaux.

Proposition 1.7.5. *Tout cardinal infini est de la forme $\aleph_\alpha$.*

Démonstration. Soit κ un cardinal infini. La fonction $\beta \mapsto \aleph_\beta$ est strictement croissante sur $\kappa + 1$ (et à valeurs dans $\aleph_{\kappa+1}$). Donc $\aleph_\kappa \geq \kappa$ par 1.4.9, d'où $\aleph_{\kappa+1} > \kappa$. Soit $\alpha \leq \kappa + 1$ minimal avec $\aleph_\alpha > \kappa$. Comme $\kappa \geq \aleph_0$, on a $\alpha > 0$. Si α était limite, par définition on aurait $\kappa \in \bigcup_{\beta<\alpha} \aleph_\beta$ et alors $\kappa \in \aleph_\beta$ pour un $\beta < \alpha$, ce qui contredirait la minimalité de α. Donc $\alpha = \beta + 1$ et alors $\aleph_\beta \leq \kappa < \aleph_{\beta+1} = \aleph_\beta^+$. Comme $\aleph_\beta^+$ est le successeur cardinal de $\aleph_\beta$, nécessairement $\aleph_\beta = \kappa$. □

1.8 Opérations sur les cardinaux

Si X et Y sont des ensembles, on note $X + Y$ leur réunion disjointe, $X \times Y$ leur produit cartésien et X^Y l'ensemble des applications de Y dans X. Si κ et λ sont deux cardinaux, on note $\kappa + \lambda$ le cardinal de leur réunion disjointe, $\kappa\lambda$ le cardinal de leur produit cartésien et κ^λ le cardinal de l'ensemble des applications de λ dans κ. Ces opérations sont appelées *addition (resp. multiplication, exponentiation) cardinale*. Il ne faut pas confondre ces opérations avec les opérations ordinales correspondantes. Ainsi, on a par exemple $2^\omega = \omega = \aleph_0 < 2^{\aleph_0}$, où encore $\aleph_0 2 = \aleph_0$, mais $\omega < \omega 2$. Il est clair que ces opérations coïncident avec les opérations arithmétiques usuelles sur les cardinaux finis.

On notera aussi que $\mathrm{card}(X + Y) = \mathrm{card}(X) + \mathrm{card}(Y)$, $\mathrm{card}(X \times Y) = \mathrm{card}(X)\mathrm{card}(Y)$ et $\mathrm{card}(X^Y) = \mathrm{card}(X)^{\mathrm{card}(Y)}$.

La preuve des résultats suivants est immédiate.

Proposition 1.8.1. *1. L'addition et la multiplication cardinale sont commutatives et associatives, la multiplication est distributive par rapport à l'addition, $\kappa^{\lambda+\mu} = \kappa^\lambda \kappa^\mu$, $(\kappa^\lambda)^\mu = \kappa^{\lambda\mu}$ et $(\kappa\lambda)^\mu = \kappa^\mu \lambda^\mu$.*

2. Si $\kappa \leq \kappa'$, alors $\kappa + \lambda \leq \kappa' + \lambda$, $\kappa\lambda \leq \kappa'\lambda$ et $\kappa^\lambda \leq \kappa'^\lambda$ (si $\kappa > 0$) ainsi que $\lambda^\kappa \leq \lambda^{\kappa'}$ (si $\lambda > 0$). □

Proposition 1.8.2. *On a* $\mathrm{card}(\mathbb{R}) = 2^{\aleph_0}$.

Démonstration. Cela reprend l'exercice 1.1.3, compte tenu de la bijection canonique entre $\mathcal{P}(\mathbb{N})$ et $2^{\aleph_0}$.

On a une injection $h : 2^{\aleph_0} \to \mathbb{R}$ qui à une suite $(a_i)_{i\in\mathbb{N}}$ de 0 et de 1 associe $\sum_i a_i 2^{-i}$ si le support de la suite est infini, et $2 + \sum_i a_i 2^{-i}$ sinon. Cela montre $2^{\aleph_0} \leq \mathrm{card}(\mathbb{R})$. D'autre part l'image de h contient l'intervalle $]0, 1[$ qui est équipotent à $\mathbb{R}$ (par exemple via $x \mapsto 1/\pi \arctan(x) + 1/2$), d'où $\mathrm{card}(\mathbb{R}) \leq 2^{\aleph_0}$. □

Proposition 1.8.3 (Théorème de Hesseberg)**.**
Pour tout cardinal infini κ on a $\kappa\kappa = \kappa$.

Démonstration. Par induction sur α, on montrera que $\aleph_\alpha \aleph_\alpha = \aleph_\alpha$. Pour $\alpha = 0$ c'est clair. (L'application $\alpha_2 : \mathbb{N}^2 \to \mathbb{N}$, $\alpha_2(m, n) := 1/2(m + n + 1)(m + n) + n$ est bijective.) On suppose que $\aleph_\beta \aleph_\beta = \aleph_\beta$ pour tout $\beta < \alpha$, et on munit $\aleph_\alpha \times \aleph_\alpha$ de l'ordre suivant :

$(\beta, \gamma) < (\beta', \gamma')$ si $\max(\beta, \gamma) < \max(\beta', \gamma')$ ou si $\max(\beta, \gamma) = \max(\beta', \gamma')$ et $\beta < \beta'$ ou si $\max(\beta, \gamma) = \max(\beta', \gamma')$ et $\beta = \beta'$ et $\gamma < \gamma'$.

On vérifie qu'il s'agit d'un bon ordre.

De plus, pour tout $\delta < \aleph_\alpha$, la partie $\delta \times \delta$ est un ségment initial pour $<$. Soit $f : \varepsilon \to \aleph_\alpha \times \aleph_\alpha$ l'unique isomorphisme d'ensembles ordonnés avec ε un ordinal. Si $\varepsilon > \aleph_\alpha$, alors $\aleph_\alpha \in \varepsilon$ et $f(\aleph_\alpha) = (\beta_0, \gamma_0) \in \aleph_\alpha \times \aleph_\alpha$. Posons $\delta_0 := \max(\beta_0, \gamma_0) + 1$. Comme aucun ordinal successeur infini n'est un cardinal (par 1.7.1), on a $\delta_0 < \aleph_\alpha$ et f restreinte à $\aleph_\alpha$ définit une injection de $\aleph_\alpha$ dans $\delta_0 \times \delta_0$, un ensemble de cardinal $\mathrm{card}(\delta_0 \times \delta_0) = \mathrm{card}(\delta_0) \leq \delta_0 < \aleph_\alpha$ par hypothèse d'induction. Contradiction. Donc $\aleph_\alpha \aleph_\alpha \leq \aleph_\alpha$. L'inégalité dans l'autre sens est claire. □

Exemple 1.8.4. Soit $\mathcal{T} = \{X \subseteq \mathbb{R} \mid X \text{ est un ouvert}\}$. Alors $\mathrm{card}(\mathcal{T}) = 2^{\aleph_0}$.

Démonstration. L'application qui à $r \in \mathbb{R}$ associe l'intervalle $(r, +\infty)$ définit une injection de $\mathbb{R}$ dans $\mathcal{T}$, ce qui montre que $\mathrm{card}(\mathcal{T}) \geq 2^{\aleph_0}$.

Réciproquement, notons que tout ouvert de $\mathbb{R}$ s'écrit comme réunion sur des intervalles de la forme $=(q, q + q')$, avec $q \in \mathbb{Q}$ et $q' \in \mathbb{Q}_{>0}$. L'application qui à $Y \subseteq \mathbb{Q} \times \mathbb{Q}_{>0}$ associe $\bigcup_{(q,q')\in Y}(q, q + q')$ définit donc une surjection de $\mathcal{P}(\mathbb{Q} \times \mathbb{Q}_{>0})$ sur $\mathcal{T}$. Comme $\mathbb{Q} \times \mathbb{Q}_{>0}$ est dénombrable, on obtient $2^{\aleph_0} \geq \mathrm{card}(\mathcal{T})$. □

Proposition 1.8.5. *1. Si X et Y sont des ensembles non vides dont l'un au moins est infini, alors*

$$\mathrm{card}(X \cup Y) = \mathrm{card}(X \times Y) = \max(\mathrm{card}(X), \mathrm{card}(Y)).$$

2. *Soit $\kappa \geq \aleph_0$ et $\lambda > 0$ des cardinaux. Alors $\kappa + \lambda = \kappa\lambda = \max(\kappa, \lambda)$.*
3. *Soit $(X_i)_{i\in I}$ une famille d'ensembles avec au moins un X_i infini. Alors*

$$\operatorname{card}\left(\bigcup_{i\in I} X_i\right) \leq \sup\left(\{\operatorname{card}(X_i) \mid i \in I\} \cup \{\operatorname{card}(I)\}\right). \tag{1.1}$$

(En particulier, une réunion dénombrable d'ensembles dénombrables est dénombrable.) Si de plus les X_i sont tous non vides et 2 à 2 disjoints, alors on a égalité dans (1.1).

Démonstration. (1) Soit $\kappa = \max(\operatorname{card}(X), \operatorname{card}(Y))$. On a $\kappa \leq \operatorname{card}(X \cup Y) \leq \kappa + \kappa = 2\kappa \leq \kappa\kappa$ ainsi que $\kappa \leq \operatorname{card}(X \times Y) \leq \kappa\kappa$. On conclut par le théorème de Hesseberg. (2) est un cas particulier de (1).

(3) Soit $X = \{(x_i, i) \mid x_i \in X_i \text{ pour un } i \in I\}$ la réunion disjointe des X_i. On a une surjection canonique de X sur $\bigcup_{i\in I} X_i$ et il suffit donc de montrer que $\operatorname{card}(X) \leq \sup\left(\{\operatorname{card}(X_i) \mid i \in I\} \cup \{\operatorname{card}(I)\}\right)$. Soit $\kappa = \sup(\operatorname{card}(X_i))$, et soit Y_i l'ensemble des applications injectives de X_i dans κ. Comme les Y_i sont tous non vides, par l'axiome du choix il existe $f = (f_i)_{i\in I} \in \prod_{i\in I} Y_i$. La fonction $g : X \to \kappa \times I$, $g((x_i, i)) := (f_i(x_i), i)$ est injective, d'où $\operatorname{card}(X) \leq \kappa \operatorname{card}(I) = \max(\kappa, \operatorname{card}(I))$. L'énoncé sur l'égalité est clair. □

La proposition précédente montre que l'addition cardinale et la multiplication cardinale deviennent triviaux pour des cardinaux infinis. C'est tout à fait le contraire pour l'exponentiation cardinale. Notons que le système d'axiomes ZFC ne détermine pas complètement l'exponentiation cardinale. Ainsi, il ne permet par exemple pas de trancher sur l'hypothèse du continu.

Définition.
- L'*Hypothèse du continu (CH)* est l'énoncé $2^{\aleph_0} = \aleph_1$.
- L'*Hypothèse du continu généralisée (GCH)* dit que $2^\kappa = \kappa^+$ pour tout cardinal infini κ.

Si $(\kappa_i)_{i\in I}$ est une famille de cardinaux, on note $\sum_{i\in I} \kappa_i$ le cardinal de la réunion disjointe des κ_i, et $\prod_{i\in I} \kappa_i$ le cardinal du produit de la famille.

Théorème 1.8.6 (Théorème de König)**.** *Soient $(\kappa_i)_{i\in I}$ et $(\lambda_i)_{i\in I}$ deux familles de cardinaux dont on suppose que $\kappa_i < \lambda_i$ pour tout i. Alors $\sum_{i\in I} \kappa_i < \prod_{i\in I} \lambda_i$.*

Démonstration. Soit $f : \sum_{i\in I} \kappa_i \to \prod_{i\in I} \lambda_i$. Pour chaque i, f induit une application $f_i : \kappa_i \to \lambda_i$ donnée par la i-ème composante de la restriction de f à κ_i. Comme $\kappa_i < \lambda_i$, l'ensemble $B_i := \lambda_i \setminus \operatorname{im}(f_i)$ est non vide pour tout i. Par l'axiome du choix il existe $b \in \prod_{i\in I} B_i \subseteq \prod_{i\in I} \lambda_i$. Clairement $b \notin \operatorname{im}(f)$. □

1.9 Cofinalité

Afin de pouvoir démontrer des énoncés du type $2^{\aleph_0} \neq \aleph_\omega$, nous aurons besoin de la notion de cofinalité.

Définition.
- Soit X un ensemble totalement ordonné. On dit qu'une partie $Y \subseteq X$ est *cofinale* dans X si Y n'est pas borné dans X, c'est-à-dire si pour tout $x \in X$ il existe $y \in Y$ tel que $x \leq y$.
- Soit $f : \beta \to \alpha$ une application entre ordinaux. On dit que f est *cofinale* si $\mathrm{im}(f)$ est cofinale dans α.
- La *cofinalité* de α, notée $\mathrm{cof}(\alpha)$, est le plus petit ordinal β tel qu'il existe une fonction $f : \beta \to \alpha$ qui est cofinale.

Exemples 1.9.1.
1. $\mathrm{cof}(0) = 0$
2. $\mathrm{cof}(\alpha + 1) = 1$ pour tout α
3. $\mathrm{cof}(\omega) = \omega$

Proposition 1.9.2. *1.* $\mathrm{cof}(\alpha) \leq \alpha$

2. $\mathrm{cof}(\alpha)$ *est un cardinal.*

3. $\mathrm{cof}(\alpha)$ *est égal au plus petit ordinal* β *tel qu'il existe une application cofinale et (strictement) croissante* $f : \beta \to \alpha$.

4. $\mathrm{cof}(\mathrm{cof}(\alpha)) = \mathrm{cof}(\alpha)$

Démonstration. (1) est clair, et (2) suit du fait que tout ordinal β est en bijection avec $\mathrm{card}(\beta) \leq \beta$.

(3) Il suffit de construire $\beta \leq \mathrm{cof}(\alpha)$ et une application $f : \beta \to \alpha$ qui soit strictement croissante et cofinale. Par hypothèse il existe $h : \mathrm{cof}(\alpha) \to \alpha$ cofinale. On définit

$$X = \{x \in \mathrm{cof}(\alpha) \mid h(y) < h(x) \text{ pour tout } y < x\}.$$

L'ensemble $h(X) = \{h(x) \mid x \in X\}$ est une partie cofinale de α. En effet, soit $\gamma < \alpha$ donné. Par cofinalité de h, il existe $y \in \mathrm{cof}(\alpha)$ tel que $h(y) \geq \gamma$. Si y est minimal avec cette propriété, on a $y \in X$.

Comme $(X, <) \cong (\beta, \in)$ pour un $\beta \leq \mathrm{cof}(\alpha)$, on a donc terminé, car h restreinte à X est strictement croissante et cofinale.

(4) $\mathrm{cof}(\mathrm{cof}(\alpha)) \leq \mathrm{cof}(\alpha)$ suit de (1). Pour l'autre sens, on choisit deux fonctions croissantes et cofinales $f : \mathrm{cof}(\mathrm{cof}(\alpha)) \to \mathrm{cof}(\alpha)$ et $g : \mathrm{cof}(\alpha) \to \alpha$ ce qui est possible par (3). Alors $g \circ f : \mathrm{cof}(\mathrm{cof}(\alpha)) \to \alpha$ est cofinale, d'où $\mathrm{cof}(\alpha) \leq \mathrm{cof}(\mathrm{cof}(\alpha))$. □

On dit qu'un cardinal infini κ est *régulier* si $\mathrm{cof}(\kappa) = \kappa$, et *singulier* si $\mathrm{cof}(\kappa) < \kappa$.

Proposition 1.9.3. *Tout cardinal infini successeur est régulier. En particulier,* $\aleph_1$ *est régulier.*

Démonstration. Soit $\kappa = \aleph_{\beta+1} = \aleph_\beta^+$. Notons que pour un ordinal limite α, une partie $X \subseteq \alpha$ est cofinale si et seulement si $\alpha = \bigcup_{\gamma \in X} \gamma$. (C'est une conséquence de 1.4.5.) On considère $f : \lambda \to \kappa$ pour un cardinal $\lambda < \kappa$. Alors $\lambda \leq \aleph_\beta$ et 1.8.5(3) entraîne que $\mathrm{card}\left(\bigcup_{\beta<\lambda} f(\beta)\right) \leq \sup\left(\{f(\beta) \mid \beta < \lambda\} \cup \{\lambda\}\right) \leq \aleph_\beta$. Donc f n'est pas cofinale. □

Proposition 1.9.4. *Si λ est un ordinal limite, alors* $\mathrm{cof}(\aleph_\lambda) = \mathrm{cof}(\lambda)$.

Démonstration. Si $f : \alpha \to \lambda$ est cofinale, alors il est clair que $\tilde{f} : \alpha \to \aleph_\lambda$, $\beta \mapsto \aleph_{f(\beta)}$, est cofinale aussi, car $\aleph_\lambda = \bigcup_{\gamma<\lambda} \aleph_\gamma$ par définition. Cela montre $\mathrm{cof}(\aleph_\lambda) \leq \mathrm{cof}(\lambda)$. Réciproquement, soit $g : \alpha \to \aleph_\lambda$ cofinale. On définit $\tilde{g} : \alpha \to \lambda$, $\tilde{g}(\beta) = 0$ si $g(\beta)$ est fini, et $\tilde{g}(\beta) = \gamma$ si $\mathrm{card}(g(\beta)) = \aleph_\gamma$. Clairement, $\tilde{g}$ est cofinale. □

Proposition 1.9.5. *Si $\kappa \geq 2$ et $\lambda \geq \aleph_0$ sont deux cardinaux, alors* $\mathrm{cof}\left(\kappa^\lambda\right) > \lambda$.

Démonstration. Soit $f : \alpha \to \kappa^\lambda$ une application, avec α un ordinal $\leq \lambda$. Comme $f(\beta) < \kappa^\lambda$ pour tout $\beta < \alpha$, le théorème de König donne

$$\mathrm{card}\left(\bigcup_{\beta<\alpha} f(\beta)\right) \leq \sum_{\beta<\alpha} \mathrm{card}(f(\beta)) < \prod_{\beta<\alpha} \left(\kappa^\lambda\right) = \left(\kappa^\lambda\right)^{\mathrm{card}(\alpha)} = \kappa^{\lambda\cdot\mathrm{card}(\alpha)} \leq \kappa^\lambda.$$

Donc f n'est pas cofinale. □

Corollaire 1.9.6. *On a $2^{\aleph_0} \neq \aleph_\omega$.*

Démonstration. On a $\mathrm{cof}(\aleph_\omega) = \mathrm{cof}(\omega) = \omega = \aleph_0 < \mathrm{cof}\left(2^{\aleph_0}\right)$. □

Chapitre 2

Calcul des prédicats

2.1 Langages et structures

Les énoncés du calcul des prédicats sont des chaînes de symboles qui décrivent des propriétés de structures. L'énoncé

$$\varphi = \forall x(x > 0 \to \exists y\, y \cdot y \doteq x)$$

par exemple est satisfait dans un corps ordonné si et seulement si tout élément positif est un carré. Il est vrai dans le corps ordonné des réels $\mathfrak{R} = \langle \mathbb{R}; 0, 1, +, -, < \rangle$ et faux dans le corps ordonné des rationnels.

On verra que l'énoncé φ s'exprime dans le cadre que nous allons étudier, c'est-à-dire dans la logique du premier ordre. Par contre, nous verrons que d'autres propriétés de corps ordonnés comme le fait d'être *archimédien* (pour tout $x > 0$ il existe $n \in \mathbb{N}$ tel que $nx > 1$) ou *complet* (toute partie bornée non vide admet un supremum) ne s'expriment pas par des énoncés du premier ordre.

Définition. Un *langage (du premier ordre)* est un ensemble de symboles $\mathcal{L}$ qui se compose de deux parties :

1. La première partie (commune à tous les langages) consiste en les symboles auxiliaires « (» et «) » ainsi qu'en les *symboles logiques* suivants :

l'ensemble des *variables*	$\mathcal{V} = \{v_n \mid n \in \mathbb{N}\}$
le *symbole de l'égalité*	$\doteq$ (« égal »)
les *connecteurs*	$\neg$(négation, « non ») , $\wedge$(conjonction, « et »)
le *quanteur existentiel*	$\exists$ (« il existe »)

2. La deuxième partie, appelée la *signature* de $\mathcal{L}$ et notée $\sigma^{\mathcal{L}}$, consiste en les *symboles non logiques* de $\mathcal{L}$. Elle est formée
   - d'un ensemble de *symboles de constante* $\mathcal{C}^{\mathcal{L}}$;
   - d'une suite d'ensembles $\mathcal{F}_n^{\mathcal{L}}$, $n \in \mathbb{N}^*$, où les éléments de $\mathcal{F}_n^{\mathcal{L}}$ sont appelés *symboles de fonction n-aires* ;
   - d'une suite d'ensembles $\mathcal{R}_n^{\mathcal{L}}$, $n \in \mathbb{N}^*$, où les éléments de $\mathcal{R}_n^{\mathcal{L}}$ sont appelés *symboles de relation* (ou *prédicat*) *n-aires*.

Le langage $\mathcal{L}$ est donné par la réunion (disjointe) de ces ensembles de symboles.

On observe que tout langage est infini. Comme la première partie est commune à tous les langages, nous identifions souvent $\mathcal{L}$ et $\sigma^{\mathcal{L}}$, par un abus de notation.

Exemples 2.1.1.

$\mathcal{L}_\emptyset = \emptyset$	Le langage vide.
$\mathcal{L}_{an} = \{\underline{0}, \underline{1}, +, -, \cdot\}$	Le langage des anneaux (avec 1).
$\mathcal{L}_{ord} = \{<\}$	Le langage des ordres.
$\mathcal{L}_{c.ord} = \mathcal{L}_{an} \cup \mathcal{L}_{ord}$	Le langage des corps ordonnés.
$\mathcal{L}_{ar} = \{\underline{0}, S, +, \cdot, <\}$	Le langage de l'arithmétique.
$\mathcal{L}_{ens} = \{\in\}$	Le langage de la théorie des ensembles.

Dans ces exemples, $\underline{0}$ et $\underline{1}$ sont des symboles de constantes, $-$ et S des symboles de fonction unaires, $+$ et $\cdot$ des symboles de fonction binaires, et $<$ ainsi que $\in$ des symboles de relation binaires.

Définition. Soit $\mathcal{L}$ un langage. Une *$\mathcal{L}$-structure* $\mathfrak{A}$ est la donnée d'un ensemble non vide A (appelé l'*ensemble de base* de $\mathfrak{A}$) muni d'un élément $c^{\mathfrak{A}} \in A$ pour chaque $c \in \mathcal{C}^{\mathcal{L}}$, d'une fonction $f^{\mathfrak{A}} : A^n \to A$ pour chaque $f \in \mathcal{F}_n^{\mathcal{L}}$ et d'une partie $R^{\mathfrak{A}} \subseteq A^n$ pour chaque $R \in \mathcal{R}_n^{\mathcal{L}}$. On l'écrit $\mathfrak{A} = \langle A; (Z^{\mathfrak{A}})_{Z \in \sigma^{\mathcal{L}}} \rangle$.

$Z^{\mathfrak{A}}$ est une *interprétation* du symbole $Z \in \sigma^{\mathcal{L}}$.

Exemples 2.1.2.
1. $\mathfrak{N} = \langle \mathbb{N}; 0, S, +, \cdot, < \rangle$ est une $\mathcal{L}_{ar}$-structure, avec S l'application successeur qui à x associe $x + 1$.
2. $\mathfrak{C} = \langle \mathbb{C}; 0, 1, +, -, \cdot \rangle$, le corps des complexes, est une $\mathcal{L}_{an}$-structure.
3. $\mathfrak{R} = \langle \mathbb{R}; 0, 1, +, -, \cdot, < \rangle$, le corps ordonné des réels, est une $\mathcal{L}_{c.ord}$-structure.

Définition. On dit que deux $\mathcal{L}$-structures $\mathfrak{A}$ et $\mathfrak{B}$ sont *isomorphes*, $\mathfrak{A} \cong \mathfrak{B}$, s'il existe un *isomorphisme* $F : \mathfrak{A} \cong \mathfrak{B}$, une bijection $F : A \to B$ qui commute avec les interprétations des symboles dans $\sigma^{\mathcal{L}}$, c'est-à-dire
- $F(c^{\mathfrak{A}}) = c^{\mathfrak{B}}$ pour tout $c \in \mathcal{C}^{\mathcal{L}}$,
- $F(f^{\mathfrak{A}}(a_1, \ldots, a_n)) = f^{\mathfrak{B}}(F(a_1), \ldots, F(a_n))$ pour tout $f \in \mathcal{F}_n^{\mathcal{L}}$ et tout uple $(a_1, \ldots, a_n) \in A^n$,
- $(a_1, \ldots, a_n) \in R^{\mathfrak{A}} \iff (F(a_1), \ldots, F(a_n)) \in R^{\mathfrak{B}}$ pour tout $R \in \mathcal{R}_n^{\mathcal{L}}$ et tout uple $(a_1, \ldots, a_n) \in A^n$.

2.2 Termes et formules

Un *mot* sur un ensemble (alphabet) E est une chaîne finie $m = a_0 a_1 \cdots a_{k-1}$ avec $a_i \in E$ pour tout i. On appelle k la *longueur* de m, et on note E^* l'ensemble des mots sur E.

Définition. Soit $\mathcal{L}$ un langage. L'ensemble $\mathcal{T}^{\mathcal{L}}$ des $\mathcal{L}$-termes est le plus petit sous-ensemble de $\mathcal{L}^*$ qui contient les variables et les symboles de constante et tel que si $f \in \mathcal{F}_n^{\mathcal{L}}$ et $t_1, \ldots, t_n \in \mathcal{T}^{\mathcal{L}}$, alors $f t_1 \cdots t_n \in \mathcal{T}^{\mathcal{L}}$.

On a $\mathcal{T}^{\mathcal{L}} = \bigcup_{n\in\mathbb{N}} \mathcal{T}_n^{\mathcal{L}}$, où on définit $\mathcal{T}_0^{\mathcal{L}} = \mathcal{C}^{\mathcal{L}} \cup \mathcal{V}^{\mathcal{L}}$, puis inductivement $\mathcal{T}_{n+1}^{\mathcal{L}} = \mathcal{T}_n^{\mathcal{L}} \cup \{ft_1\cdots t_k \mid f \in \mathcal{F}_k^{\mathcal{L}} \text{ et } t_1,\ldots,t_k \in \mathcal{T}_n^{\mathcal{L}}\}$.

Proposition 2.2.1 (Lecture unique des termes)**.** *Tout terme $t \in \mathcal{T}^{\mathcal{L}}$ vérifie une et une seule des trois possibilités suivantes :*

1. *t est une variable,*
2. *t est un symbole de constante,*
3. *il existe un unique entier $n \geq 1$, un unique symbole de fonction n-aire f et une unique suite $(t_1,\ldots,t_n)$ de termes tels que $t = ft_1\cdots t_n$.*

Démonstration. Exercice. (On montrera d'abord, par induction sur la longueur des termes, qu'aucun ségment initial propre d'un terme n'est un terme.) □

Notation. Pour faciliter la lecture des termes, on écrira dorénavant $f(t_1,\ldots,t_n)$ pour $ft_1\cdots t_n$. Si f est binaire, on écrira parfois aussi t_1ft_2 au lieu de ft_1t_2. Par exemple, $(x+y)\cdot z$ signifie $\cdot + xyz$.

Si t est un terme, son *hauteur* $\mathrm{ht}(t)$ est défini comme le plus petit entier k tel que $t \in \mathcal{T}_k^{\mathcal{L}}$. Il suit de la propriété de lecture unique pour les termes que $\mathrm{ht}(ft_1\cdots t_n) = 1 + \max(\mathrm{ht}(t_i))$.

Définition. 1. Une *$\mathcal{L}$-formule atomique* est
   - soit un mot de la forme $t_1 \doteq t_2$, où t_1 et t_2 sont des $\mathcal{L}$-termes,
   - soit un mot de la forme $Rt_1\cdots t_n$, où $R \in \mathcal{R}_n^{\mathcal{L}}$ et tous les t_i sont des $\mathcal{L}$-termes.

2. L'ensemble $\mathcal{F}ml^{\mathcal{L}}$ des $\mathcal{L}$-formules est le plus petit sous-ensemble de $\mathcal{L}^*$ qui contient les formules atomiques et tel que si $\varphi,\psi \in \mathcal{F}ml^{\mathcal{L}}$ et $x \in \mathcal{V}$, alors $\neg\varphi$, $(\varphi\wedge\psi)$ et $\exists x\varphi$ sont des éléments de $\mathcal{F}ml^{\mathcal{L}}$.

On a $\mathcal{F}ml^{\mathcal{L}} = \bigcup_{n\in\mathbb{N}} \mathcal{F}ml_n^{\mathcal{L}}$, où $\mathcal{F}ml_0^{\mathcal{L}}$ est définit comme l'ensemble des formules atomiques, puis inductivement

$$\begin{aligned}\mathcal{F}ml_{n+1}^{\mathcal{L}} \quad = \quad & \mathcal{F}ml_n^{\mathcal{L}} \cup \{\neg\varphi \mid \varphi \in \mathcal{F}ml_n^{\mathcal{L}}\} \cup \{(\varphi\wedge\psi) \mid \varphi,\psi \in \mathcal{F}ml_n^{\mathcal{L}}\} \\ & \cup \{\exists x\varphi \mid x \in \mathcal{V} \text{ et } \varphi \in \mathcal{F}ml_n^{\mathcal{L}}\}.\end{aligned}$$

Proposition 2.2.2 (Lecture unique des formules)**.**
Toute formule $\varphi \in \mathcal{F}ml^{\mathcal{L}}$ vérifie une et une seule des possibilités suivantes :

1. *φ est une formule atomique, soit de la forme $Rt_1\cdots t_n$ pour R et $t_1,\ldots,t_n$ uniques, soit de la forme $t_1 \doteq t_2$ pour t_1, t_2 uniques ;*
2. *$\varphi = \neg\psi$ pour une unique $\mathcal{L}$-formule ψ ;*
3. *$\varphi = (\psi\wedge\chi)$ pour des uniques $\mathcal{L}$-formules ψ et χ ;*
4. *$\varphi = \exists x\psi$ pour une unique variable x et une unique $\mathcal{L}$-formule ψ.*

Démonstration. Exercice. (On montrera d'abord qu'aucun ségment initial propre d'une formule n'est une formule.) □

Si φ est une formule, son *hauteur* $\mathrm{ht}(\varphi)$ est définie comme le plus petit entier k tel que $\varphi \in \mathcal{F}ml_k^{\mathcal{L}}$. Il suit de la propriété de lecture unique pour les formules que $\mathrm{ht}(\exists x\varphi) = \mathrm{ht}(\neg\varphi) = 1 + \mathrm{ht}(\varphi)$ et $\mathrm{ht}((\varphi \wedge \psi)) = 1 + \max(\mathrm{ht}(\varphi), \mathrm{ht}(\psi))$. Cela nous permet de donner des définitions par induction sur la hauteur des formules.

Définition. 1. Soit v_k une variable. On définit, par induction sur la hauteur de φ, la notion d'*occurrence libre* de v_k dans φ :
   - Si φ est atomique, toutes les occurrences de v_k dans φ sont libres.
   - Les occurrences libres de v_k dans $\varphi = \neg\psi$ sont celles dans ψ.
   - Les occurrences libres de v_k dans $\varphi = (\psi \wedge \chi)$ sont celles dans ψ et celles dans χ.
   - Si $l \neq k$, les occurrences libres de v_k dans $\varphi = \exists v_l\psi$ sont celles de ψ.
   - Aucune occurrence de v_k dans $\varphi = \exists v_k\psi$ n'est libre.

   Les occurrences non libres d'une variable sont dites *liées*.

2. Les *variables libres de* φ sont celles qui admettent au moins une occurrence libre dans φ. On note $\mathrm{Lib}(\varphi)$ l'ensemble des variables libres de φ.

3. Un *énoncé* est une fomule sans variable libre.

Exemple. Dans $\varphi = (\exists v_0\, v_0 < v_1 \wedge v_0 \dot{=} v_1)$, les deux premières occurrences de v_0 sont liées, tandis que la troisième est libre. Toutes les occurrences de v_1 dans φ sont libres, d'où $\mathrm{Lib}(\varphi) = \{v_0, v_1\}$.

Notation. On utilisera les abbréviations suivantes :

$(\varphi \vee \psi) = \neg(\neg\varphi \wedge \neg\psi)$	(*disjonction*, « ou »)
$(\varphi \to \psi) = \neg(\varphi \wedge \neg\psi)$	(*implication*, « implique »)
$(\varphi \leftrightarrow \psi) = ((\varphi \to \psi) \wedge (\psi \to \varphi))$	(*équivalence*, « si et seulement si »)
$\forall x\varphi = \neg\exists x\neg\varphi$	(*quanteur universel*, « pour tout »)

On écrira $\exists x_1, \ldots, x_n$ au lieu de $\exists x_1 \cdots \exists x_n$ (de même pour le quanteur universel), $R(t_1, \ldots, t_n)$ au lieu de $Rt_1 \cdots t_n$, et parfois $t_1 R t_2$ au lieu de Rt_1t_2.

On notera $(\varphi_0 \wedge \cdots \wedge \varphi_n)$ au lieu de $(\cdots((\varphi_0 \wedge \varphi_1) \wedge \varphi_2) \wedge \ldots \wedge \varphi_n)$, de même pour $\vee$ à la place de $\wedge$.

Enfin, pour faciliter la lecture des formules, on ajoutera parfois des parenthèses, ou on les omettra. Dans ce cas, on lira les formules selon l'*affinité* des symboles logiques :

Affinité maximale :	$\neg\ \exists\ \forall$
	$\wedge$
	$\vee$
Affinité minimale :	$\to\ \leftrightarrow$

Ainsi, $\forall x\varphi \wedge \psi \to \chi$ voudra dire $((\forall x\varphi \wedge \psi) \to \chi) = \neg((\forall x\varphi \wedge \psi) \wedge \neg\chi) = \neg((\neg\exists x\neg\varphi \wedge \psi) \wedge \neg\chi)$.

Exemple. Voici la liste des axiomes de corps dans $\mathcal{L}_{an}$.

1. $\forall x\, x + \underline{0} \dot{=} x$
2. $\forall x, y\, x + y \dot{=} y + x$

3. $\forall x\,(-x) + x = \underline{0}$
4. $\forall x, y, z\,(x + y) + z \doteq x + (y + z)$
5. $\forall x\, x \cdot \underline{1} \doteq x$
6. $\forall x, y\, x \cdot y \doteq y \cdot x$
7. $\forall x, y, z\,(x \cdot y) \cdot z \doteq x \cdot (y \cdot z)$
8. $\forall x, y, z\, x \cdot (y + z) \doteq (x \cdot y) + (x \cdot z)$
9. $\forall x\,(\neg x \doteq \underline{0} \to \exists y\, x \cdot y \doteq \underline{1})$
10. $\neg \underline{0} \doteq \underline{1}$

2.3 Sémantique

Dans ce paragraphe, nous donnons un sens aux formules.

Définition. 1. Soit $\mathfrak{A}$ une $\mathcal{L}$-structure. Une *affectation* est une fonction $\alpha : \mathcal{V} \to A$ de l'ensemble des variables dans l'ensemble de base de $\mathfrak{A}$.

2. Si α est une affectation, $\mathfrak{A}$ une $\mathcal{L}$-structure et t un $\mathcal{L}$-terme, on définit $t^{\mathfrak{A}}[\alpha]$, par induction sur $\mathrm{ht}(t)$, de la manière suivante :
   - $v_i^{\mathfrak{A}}[\alpha] = \alpha(v_i)$ (pour $v_i \in \mathcal{V}$) et $c^{\mathfrak{A}}[\alpha] = c^{\mathfrak{A}}$ (pour $c \in \mathcal{C}^{\mathcal{L}}$),
   - $f(t_1, \ldots, t_n)^{\mathfrak{A}}[\alpha] = f^{\mathfrak{A}}(t_1^{\mathfrak{A}}[\alpha], \ldots, t_n^{\mathfrak{A}}[\alpha])$.

La preuve du lemme suivant est claire.

Lemme 2.3.1. *Si les affectations α et β coïncident sur toutes les variables qui ont une occurrence dans t, alors $t^{\mathfrak{A}}[\alpha] = t^{\mathfrak{A}}[\beta]$.* □

Notation. Si t est un terme, on écrira $t = t(x_1, \ldots, x_n)$ si les variables x_i sont distinctes et si toutes les variables ayant au moins une occurrence dans t figurent parmi les x_i.

Si $t = t(x_1, \ldots, x_n)$ et $a_1, \ldots, a_n \in A$, on définit $t^{\mathfrak{A}}[a_1, \ldots, a_n]$ par $t^{\mathfrak{A}}[\alpha]$ où α est une affectation avec $\alpha(x_i) = a_i$ pour tout i. (C'est bien défini par le lemme.)

Définition (Satisfaction d'une fomule)**.** Soit $\mathfrak{A}$ une $\mathcal{L}$-structure. On définit pour toutes les affectations α et toutes les formules la relation $\mathfrak{A} \models \varphi[\alpha]$ qu'on lit « φ est satisfait dans $\mathfrak{A}$ par α » :

$$
\begin{array}{lll}
\mathfrak{A} \models t_1 \doteq t_2[\alpha] & :\Longleftrightarrow & t_1^{\mathfrak{A}}[\alpha] = t_2^{\mathfrak{A}}[\alpha] \\
\mathfrak{A} \models R t_1 \cdots t_n[\alpha] & :\Longleftrightarrow & (t_1^{\mathfrak{A}}[\alpha], \ldots, t_n^{\mathfrak{A}}[\alpha]) \in R^{\mathfrak{A}} \\
\mathfrak{A} \models \neg\psi[\alpha] & :\Longleftrightarrow & \mathfrak{A} \not\models \psi[\alpha] \\
\mathfrak{A} \models (\psi \wedge \chi)[\alpha] & :\Longleftrightarrow & \mathfrak{A} \models \psi[\alpha] \text{ et } \mathfrak{A} \models \chi[\alpha] \\
\mathfrak{A} \models \exists x \psi[\alpha] & :\Longleftrightarrow & \text{il existe } a \in A \text{ tel que } \mathfrak{A} \models \psi[\alpha_{a/x}]
\end{array}
$$

Ici, $\alpha_{a/x}(y) := \begin{cases} \alpha(y), \text{ si } y \neq x, \\ a, \text{ si } y = x. \end{cases}$

Proposition 2.3.2. *Si les affectations α et β coïncident sur* $\mathrm{Lib}(\varphi)$*, alors on a* $\mathfrak{A} \models \varphi[\alpha] \iff \mathfrak{A} \models \varphi[\beta]$.

Démonstration. Par induction sur $\mathrm{ht}(\varphi)$. Le cas des formules atomiques suit du lemme. Dans l'étape d'induction, seul le cas $\varphi = \exists x\psi$ mérite un argument. Si $\mathfrak{A} \models \varphi[\alpha]$, il existe $a \in A$ tel que $\mathfrak{A} \models \psi[\alpha_{a/x}]$. Toute variable $y \neq x$ qui est libre dans ψ est aussi libre dans φ. Donc $\mathfrak{A} \models \psi[\beta_{a/x}]$ par l'hypothèse d'induction, d'où $\mathfrak{A} \models \varphi[\beta]$. □

Notation. Si φ est une formule, on écrira $\varphi = \varphi(x_1, \ldots, x_n)$ si les variables x_i sont distinctes et si toutes les variables libres dans φ figurent parmi les x_i.

Si $\varphi = \varphi(x_1, \ldots, x_n)$ et $a_1, \ldots, a_n \in A$, on définit $\mathfrak{A} \models \varphi[a_1, \ldots, a_n]$ par $\mathfrak{A} \models \varphi[\alpha]$, où α est une affectation avec $\alpha(x_i) = a_i$. (C'est bien défini par la proposition.)

Ainsi, la formule $\varphi(x_1, \ldots, x_n)$ définit une relation n-aire dans la structure $\mathfrak{A}$, donnée par $\{(a_1, \ldots, a_n) \in A^n \mid \mathfrak{A} \models \varphi[a_1, \ldots, a_n]\}$.

En particulier, si φ est un énoncé, on peut définir le symbole

$$\mathfrak{A} \models \varphi,$$

lu comme « φ est *satisfait* (ou *vrai*) dans $\mathfrak{A}$ » ou bien « $\mathfrak{A}$ est *modèle* de φ ».

Exemple. Une $\mathcal{L}_{an}$-structure $\langle K; 0, 1, +, -, \cdot\rangle$ est un corps si et seulement si elle satisfait aux axiomes de corps (1)–(10) donnés auparavant.

Exercice 2.3.3.

1. Soit $\mathfrak{A}$ une $\mathcal{L}$-structure et B une partie non vide de A qui contient les interprétations $c^{\mathfrak{A}}$ de toutes les constantes de $\mathcal{L}$ et qui est close sous toutes les opérations $f^{\mathfrak{A}}$. En restreignant les interprétations des symboles de A à B on obtient alors une $\mathcal{L}$-structure $\mathfrak{B}$, appelée *sous-structure* de $\mathfrak{A}$. Nous écrivons $\mathfrak{B} \subseteq \mathfrak{A}$ si $\mathfrak{B}$ est une sous-structure de $\mathfrak{A}$.
   Montrer que l'intersection d'une famille de sous-structures de $\mathfrak{A}$ est une sous-structure ou vide. En déduire que toute partie non vide X de A est contenue dans une plus petite sous-structure de $\mathfrak{A}$, appelée la *sous-structure engendrée par* X et notée $\langle X\rangle_{\mathfrak{A}}$. Montrer que l'ensemble de base de $\langle X\rangle_{\mathfrak{A}}$ est donné par $\{t^{\mathfrak{A}}[\alpha] \mid t \in \mathcal{T}^{\mathcal{L}} \text{ et } \alpha : \mathcal{V} \to X\}$.
2. Soit $\mathfrak{A} = \langle A; 0, 1, +, -, \cdot\rangle$ une $\mathcal{L}_{an}$-structure. On suppose que $\mathfrak{A}$ est un anneau. Montrer que la sous-structure engendrée par une partie X de A est donnée par le sous-anneau engendré par X.

2.4 Substitution

Le but de ce paragraphe est de donner une bonne définition de la substitution (dans une formule φ) d'une variable x par un terme s, d'une telle manière que les propriétés sémantiques qu'on attend (voir 2.4.2) soient satisfaites.

On pourrait tout simplement remplacer toute occurrence libre de x dans φ par s, mais alors il se peut qu'une variable de s est liée, dans la formule ainsi obtenue, par un quanteur. Cela peut avoir des effets sémantiques indésirables, comme par exemple pour $\varphi(v_0) = \exists v_1 \neg v_1 \doteq v_0$, $x = v_0$ et $s = v_1$. On obtiendrait

l'énoncé $\psi = \exists v_1 \neg v_1 \doteq v_1$ satisfait par aucune structure, tandis que si $\mathfrak{A}$ contient au moins deux éléments, alors $\mathfrak{A} \models \varphi[a]$ pour tout $a \in A$.

La définition suivante, un peu plus compliquée, rémédie à ce problème.

Définition. Soient $x_0, \ldots, x_r$ des variables distinctes et $s_0, \ldots, s_r$ des termes. On définit la *substitution simultanée* des x_i par les s_i comme suit.

1. Soit t un terme. Alors $t_{s_0/x_0,\ldots,s_r/x_r} = t_{\overline{s}/\overline{x}}$ est le mot obtenu en « remplaçant simultanément toutes les occurrences de x_i dans t par s_i », c'est-à-dire (par induction sur la hauteur de t)
   - $x_{\overline{s}/\overline{x}} = \begin{cases} x, & \text{si } x \neq x_0, \ldots, x \neq x_r \\ s_i, & \text{si } x = x_i \end{cases}$
   - $c_{\overline{s}/\overline{x}} = c$
   - $[ft^1 \cdots t^n]_{\overline{s}/\overline{x}} = ft^1_{\overline{s}/\overline{x}} \cdots t^n_{\overline{s}/\overline{x}}$.
2. Par induction sur la hauteur d'une formule, on définit
   - $[t \doteq t']_{\overline{s}/\overline{x}} = t_{\overline{s}/\overline{x}} \doteq t'_{\overline{s}/\overline{x}}$ ainsi que $[Rt^1 \cdots t^n]_{\overline{s}/\overline{x}} = Rt^1_{\overline{s}/\overline{x}} \cdots t^n_{\overline{s}/\overline{x}}$.
   - $[\neg\psi]_{\overline{s}/\overline{x}} = \neg\psi_{\overline{s}/\overline{x}}$.
   - $(\psi \wedge \chi)_{\overline{s}/\overline{x}} = (\psi_{\overline{s}/\overline{x}} \wedge \chi_{\overline{s}/\overline{x}})$.
   - Soient $x_{i_1}, \ldots x_{i_k}$ $(i_1 < \ldots < i_k)$ les variables parmi $x_0, \ldots, x_r$ qui sont libres dans $\exists x\psi$. (En particulier on a $x \neq x_{i_1}, \ldots, x \neq x_{i_k}$.)
     - Si x n'a pas d'occurrence dans $s_{i_1}, \ldots, s_{i_k}$, on pose
     $$[\exists x\psi]_{\overline{s}/\overline{x}} = \exists x\psi_{s_{i_1}/x_{i_1},\ldots,s_{i_k}/x_{i_k}};$$
     - si x a une d'occurrence dans $s_{i_1}, \ldots, s_{i_k}$, on pose
     $$[\exists x\psi]_{\overline{s}/\overline{x}} = \exists u\psi_{s_{i_1}/x_{i_1},\ldots,s_{i_k}/x_{i_k},u/x}, \tag{2.1}$$
     où u est la première variable dans l'énumération $v_0, v_1, v_2, \ldots$ qui n'a pas d'occurrence dans les mots $\exists x\psi, s_{i_1}, \ldots, s_{i_k}$.

Exercice 2.4.1. Montrer (par induction sur la hauteur) :

1. Si t est un terme, alors $t_{\overline{s}/\overline{x}}$ est un terme.
2. Si φ est une formule, alors $\varphi_{\overline{s}/\overline{x}}$ est une formule et on a $\mathrm{ht}(\varphi_{\overline{s}/\overline{x}}) = \mathrm{ht}(\varphi)$.

Soient $x_0, \ldots, x_r$ des variables distinctes, α une affection à valeurs dans $\mathfrak{A}$ et $a_0, \ldots, a_r$ des éléments de A. On définit l'affectation $\alpha_{a_0/x_0,\ldots,a_r/x_r} = \alpha_{\overline{a}/\overline{x}}$ comme $\alpha_{\overline{a}/\overline{x}}(x_i) = a_i$ et $\alpha_{\overline{a}/\overline{x}}(y) = \alpha(y)$ si $y \neq x_i$ pour tout i.

Proposition 2.4.2 (Lemme de substitution)**.** *Soient $x_0, \ldots, x_r$ des variables distinctes, $s_0, \ldots, s_r$ des termes et α une affectation à valeurs dans $\mathfrak{A}$.*

1. *Pour tout terme t on a $t^{\mathfrak{A}}_{\overline{s}/\overline{x}}[\alpha] = t^{\mathfrak{A}}\left[\alpha_{s_0^{\mathfrak{A}}[\alpha]/x_0,\ldots,s_r^{\mathfrak{A}}[\alpha]/x_r}\right]$.*
2. *Pour toute formule φ on a*
$$\mathfrak{A} \models \varphi_{\overline{s}/\overline{x}}[\alpha] \iff \mathfrak{A} \models \varphi\left[\alpha_{s_0^{\mathfrak{A}}[\alpha]/x_0,\ldots,s_r^{\mathfrak{A}}[\alpha]/x_r}\right].$$

Démonstration. (1) est facile. On montre (2) par induction sur la hauteur de φ, le seul cas non évident étant $\varphi = \exists x\psi$. Soient, comme dans la définition de $[\exists x\psi]_{\overline{s}/\overline{x}}$, les variables $x_{i_1}, \ldots, x_{i_k}$ exactement celles parmi les x_i qui sont libres dans $\exists x\psi$. Le cas où x n'a pas d'occurrence dans $s_{i_1}, \ldots, s_{i_k}$ est facile et laissé en exercice. Supposons donc que x a une occurrence dans l'un des termes $s_{i_1}, \ldots, s_{i_k}$, et soit u la variable choisie comme dans (2.1). On a alors

$$\begin{aligned}
&\mathfrak{A} \models [\exists x\psi]_{\overline{s}/\overline{x}}[\alpha] \iff \mathfrak{A} \models \exists u \psi_{s_{i_1}/x_{i_1},\ldots,s_{i_k}/x_{i_k},u/x}[\alpha] \\
\iff\ & \text{il existe } a \in A \text{ tel que } \mathfrak{A} \models \psi_{s_{i_1}/x_{i_1},\ldots,s_{i_k}/x_{i_k},u/x}[\alpha_{a/u}] \\
\iff\ & \text{il existe } a \in A \text{ tel que } \mathfrak{A} \models \psi[\alpha_{s^{\mathfrak{A}}_{i_1}[\alpha_{a/u}]/x_{i_1},\ldots,s^{\mathfrak{A}}_{i_k}[\alpha_{a/u}]/x_{i_k},u^{\mathfrak{A}}[\alpha_{a/u}]/x}] \\
& \quad \text{(par hypothèse d'induction)} \\
\iff\ & \text{il existe } a \in A \text{ tel que } \mathfrak{A} \models \psi[\alpha_{s^{\mathfrak{A}}_{i_1}[\alpha]/x_{i_1},\ldots,s^{\mathfrak{A}}_{i_k}[\alpha]/x_{i_k},a/x}] \\
& \quad \text{(par 2.3.1, comme } u \text{ est différent de tous les } x_{i_j}) \\
\iff\ & \mathfrak{A} \models \exists x\psi[\alpha_{s^{\mathfrak{A}}_{i_1}[\alpha]/x_{i_1},\ldots,s^{\mathfrak{A}}_{i_k}[\alpha]/x_{i_k}}] \\
\iff\ & \mathfrak{A} \models \exists x\psi[\alpha_{s^{\mathfrak{A}}_0[\alpha]/x_0,\ldots,s^{\mathfrak{A}}_r[\alpha]/x_r}]
\end{aligned}$$

La dernière équivalence suit de 2.3.2, car les affectations $\alpha_{s^{\mathfrak{A}}_{i_1}[\alpha]/x_{i_1},\ldots,s^{\mathfrak{A}}_{i_k}[\alpha]/x_{i_k}}$ et $\alpha_{s^{\mathfrak{A}}_0[\alpha]/x_0,\ldots,s^{\mathfrak{A}}_r[\alpha]/x_r}$ coïncident sur l'ensemble $\mathrm{Lib}(\exists x\psi)$. □

Exemples 2.4.3.

1. Soit $\varphi = \varphi(v_0) = \exists v_1 \neg v_1 \doteq v_0$ et $s = v_1$. Alors $\varphi_{s/v_0} = \exists v_2[\neg v_1 \doteq v_0]_{v_1/v_0, v_2/v_1} = \exists v_2 \neg v_2 \doteq v_1$.
2. Soient φ une formule et $x_1, \ldots, x_n$ des variables distinctes. On suppose que $s_1, \ldots, s_n$ ne contiennent pas de variables ayant une occurrence dans φ. Alors $\varphi_{\overline{s}/\overline{x}}$ est le mot obtenu en remplaçant toute occurrence libre de x_i par s_i. (Preuve par induction sur $\mathrm{ht}(\varphi)$, laissée en exercice.)

Lemme 2.4.4. *Soit y une variable sans occurrence dans φ. Alors $\left(\varphi_{y/x}\right)_{x/y} = \varphi$.*

Démonstration. On fait une induction sur la hauteur de φ, le seul cas non trivial étant $\varphi = \exists z\psi$. Si $z = x$, alors ni x ni y n'est libre dans $\varphi = \exists x\psi$, d'où $(\varphi_{y/x})_{x/y} = \varphi_{x/y} = \varphi$. Sinon, on a $z \neq x$ et $z \neq y$, et alors $([\exists z\psi]_{y/x})_{x/y} = (\exists z \psi_{y/x})_{x/y} = \exists z (\psi_{y/x})_{x/y}$. Or $(\psi_{y/x})_{x/y} = \psi$ par hypothèse d'induction. □

Notation. Soient $s_1, \ldots, s_n$ des termes. Si $t(x_1, \ldots, x_n)$ est un terme, nous écrivons souvent $t(s_1, \ldots, s_n)$ pour $t_{s_1/x_1,\ldots,s_n/x_n}$, et $\varphi(s_1, \ldots, s_n)$ au lieu de $\varphi_{s_1/x_1,\ldots,s_n/x_n}$ si $\varphi = \varphi(x_1, \ldots, x_n)$ est une formule.

2.5 Formules universellement valides

Définition. Une $\mathcal{L}$-formule φ est dite *universellement valide*, noté $\models \varphi$, si elle est satisfaite par toute $\mathcal{L}$-structure $\mathfrak{A}$ et toute affectation α à valeurs dans $\mathfrak{A}$, c'est-à-dire si $\mathfrak{A} \models \varphi[\alpha]$ pour tout $\mathfrak{A}$ et tout α.

Remarque 2.5.1. *La formule $\varphi = \varphi(x_1, \ldots, x_n)$ est universellement valide si et seulement si l'énoncé $\forall x_1, \ldots, x_n \varphi$ est universellement valide.*

On appelle $\forall x_1, \ldots, x_n \varphi$ une *clôture universelle de* φ.

Exemples. 1. $\exists x\, x \doteq x$ est universellement valide. (Par définition, une structure est non vide !)

2. Soient φ et ψ des $\mathcal{L}$-formules quelconques. Alors $(\varphi \to (\psi \to \varphi))$ est universellement valide.

Le lemme suivant justifie que nous ne faisons pas référence au langage $\mathcal{L}$ dans la notation $\models \varphi$.

Lemme 2.5.2. *Soit φ une $\mathcal{L}$-formule et $\mathcal{L}' \supseteq \mathcal{L}$. Alors φ est universellement valide en tant que $\mathcal{L}$-formule si et seulement si elle l'est en tant que $\mathcal{L}'$-formule.*

Démonstration. Si $\mathfrak{A}' = \langle A; (Z^{\mathfrak{A}'})_{Z \in \sigma^{\mathcal{L}'}} \rangle$ est une $\mathcal{L}'$-structure, on pose $\mathfrak{A} := \mathfrak{A}' \restriction_{\mathcal{L}} := \langle A; (Z^{\mathfrak{A}'})_{Z \in \sigma^{\mathcal{L}}} \rangle$, le *réduit de $\mathfrak{A}'$ à $\mathcal{L} \subseteq \mathcal{L}'$* (la structure $\mathfrak{A}'$ est appelée *expansion de $\mathfrak{A}$ à $\mathcal{L}'$*).

On peut identifier les affectations à valeurs dans $\mathfrak{A}$ et celles à valeurs dans $\mathfrak{A}'$, et on a $\mathfrak{A}' \models \varphi[\alpha] \iff \mathfrak{A} \models \varphi[\alpha]$ pour toute affectation α.

Il suffit donc de montrer que toute $\mathcal{L}$-structure $\mathfrak{A}$ admet une expansion à une $\mathcal{L}'$-structure. Ceci est clair. □

Afin de préciser un cadre pour les formules de la forme $(\varphi \to (\psi \to \varphi))$ qui sont toujours vérifiées indépendemment des valeurs de vérité de φ et ψ, on introduira les formules du *calcul propositionnel*.

On fixe un ensemble $\mathcal{P} = \{p_i \mid i \in \mathbb{N}\}$, où les p_i sont des *variables propositionnelles* (elles prennent comme valeurs « vrai » ou « faux »).

Les formules du calcul propositionnel sont des mots sur l'alphabet $\mathcal{P} \cup \{\neg, \wedge, (,)\}$, formés selon les règles suivantes :

– p_i est une formule pour tout i ;
– si F et G sont des formules, alors $\neg F$ et $(F \wedge G)$ aussi.

Soit $\mathcal{F}ml_{\mathcal{P}}$ l'ensemble des formules du calcul propositionnel. Comme avant on introduit $\vee, \to$ et $\leftrightarrow$ comme abbréviations. On identifie $\mathbb{Z}/2 = \{0, 1\}$ avec $\{$ « faux », « vrai »$\}$, en associant 0 à « faux » et 1 à « vrai ». Une *affectation* (ou *distribution de valeurs de vérité*) est une application $\delta : \mathcal{P} \to \{0, 1\}$.

Toute affectation δ induit une fonction $\delta^* : \mathcal{F}ml_{\mathcal{P}} \to \{0, 1\}$ donnée (inductivement) par $\delta^*(p_i) = \delta(p_i)$, $\delta^*(\neg F) = 1 - \delta^*(F)$ et $\delta^*((F \wedge G)) = \delta^*(F) \cdot \delta^*(G)$.

On écrit $F = F(q_1, \ldots, q_n)$ si les q_i sont des variables distinctes et si toutes les variables ayant une occurrence dans F figurent parmi les q_i.

Exercice 2.5.3. Pour toute fonction $g : \{0, 1\}^n \to \{0, 1\}$ il existe une formule $F(p_1, \ldots, p_n)$ telle que pour toute affectation δ on ait $\delta^*(F) = g(\delta(p_1), \ldots, \delta(p_n))$.

Définition. 1. Une *tautologie du calcul propositionnel* est une formule $F \in \mathcal{F}ml_{\mathcal{P}}$ telle que $\delta^*(F) = 1$ pour toute distribution de valeurs de vérité δ. (Autrement dit, si $F = F(q_1, \ldots, q_n)$, alors F est vraie quelle que soient les valeurs de vérité des q_i.)

2. La $\mathcal{L}$-formule φ est une *tautologie du calcul des prédicats* s'il existe une tautologie $F = F(q_1, \ldots, q_n)$ du calcul propositionnel et des $\mathcal{L}$-formules $\psi_1, \ldots, \psi_n$ telles que $\varphi = F_{\psi_1/q_1, \ldots, \psi_n/q_n}$ (le mot obtenu en substituant les ψ_i aux variables q_i).

Lemme 2.5.4. *Les tautologies (du calcul des prédicats) sont des formules universellement valides.*

Démonstration. Clair. □

Lemme 2.5.5 (Axiomes de l'égalité)**.**
Les énoncés suivants sont universellement valides :

- $\forall x\, x \doteq x$ *(réflexivité)*
- $\forall x, y\, (x \doteq y \to y \doteq x)$ *(symétrie)*
- $\forall x, y, z\, ((x \doteq y \wedge y \doteq z) \to x \doteq z)$ *(transitivité)*
- $\forall x_1, \ldots, x_n, y_1, \ldots, y_n\, (x_1 \doteq y_1 \wedge \ldots \wedge x_n \doteq y_n \to f x_1 \cdots x_n \doteq f y_1 \cdots y_n)$ *pour tout* $f \in \mathcal{F}_n^{\mathcal{L}}$ *(congruence 1)*
- $\forall x_1, \ldots, x_n, y_1, \ldots, y_n\, (x_1 \doteq y_1 \wedge \ldots \wedge x_n \doteq y_n \wedge R x_1 \cdots x_n \to R y_1 \cdots y_n)$ *pour tout* $R \in \mathcal{R}_n^{\mathcal{L}}$ *(congruence 2)*

Remarque. *Ces énoncés expriment que* $\doteq$ *est une* relation de congruence, *c'est-à-dire une relation d'équivalence qui est compatible avec les fonctions et les relations du langage.*

Lemme 2.5.6 (Axiomes du quanteur existentiel)**.** *Soient* φ *une* $\mathcal{L}$*-formule,* t *un* $\mathcal{L}$*-terme et* x *une variable. Alors la formule* $\varphi_{t/x} \to \exists x \varphi$ *est universellement valide.*

Démonstration. On pose $\psi = \varphi_{t/x} \to \exists x \varphi$. Soit α une affectation à valeurs dans $\mathfrak{A}$. Par le lemme de substitution, on a

$$\mathfrak{A} \models \varphi_{t/x}[\alpha] \Rightarrow \mathfrak{A} \models \varphi\left[\alpha_{t^{\mathfrak{A}}[\alpha]/x}\right] \Rightarrow \mathfrak{A} \models \exists x \varphi[\alpha].$$

On en déduit que $\mathfrak{A} \models \psi[\alpha]$. □

Lemme 2.5.7 (Modus ponens)**.** *Si* φ *et* $(\varphi \to \psi)$ *sont universellement valides, alors* ψ *aussi.* □

Lemme 2.5.8 ($\exists$-introduction)**.** *On suppose que* x *n'est pas libre dans* ψ*. Alors si* $\varphi \to \psi$ *est universellement valide,* $\exists x \varphi \to \psi$ *l'est aussi.*

Démonstration. Si $\mathfrak{A} \models \exists x \varphi[\alpha]$, il existe $a \in A$ tel que $\mathfrak{A} \models \varphi[\alpha_{a/x}]$.

Si $\models \varphi \to \psi$, on a donc $\mathfrak{A} \models \psi[\alpha_{a/x}]$. On en déduit $\mathfrak{A} \models \psi[\alpha]$ par Proposition 2.3.2, puisque x n'est pas libre dans ψ. Cela établit $\mathfrak{A} \models (\exists x \varphi \to \psi)[\alpha]$. □

Exemple. La restriction sur la variable x dans le lemme est nécessaire, come le montre l'exemple suivant. Soient $\varphi = \psi = x \doteq c$. On a $\models \varphi \to \psi$, mais $\not\models (\exists x\, x \doteq c \to x \doteq c)$.

Définition. Une $\mathcal{L}$*-théorie* est un ensemble de $\mathcal{L}$-énoncés. Soit T une $\mathcal{L}$-théorie.

- Une $\mathcal{L}$-structure $\mathfrak{A}$ est *modèle de* T, notée $\mathfrak{A} \models T$, si $\mathfrak{A} \models \varphi$ pour tout $\varphi \in T$.
- L'énoncé φ est *conséquence logique* de T si pour toute $\mathcal{L}$-structure $\mathfrak{A}$ qui est modèle de T on a $\mathfrak{A} \models \varphi$. On le note $T \models \varphi$.

La preuve du lemme 2.5.2 montre que $T \models \varphi$ ne dépend pas du langage $\mathcal{L}$.

2.6 Démonstrations formelles et théorème de complétude de Gödel

On va formaliser la notion de preuve.

Axiomes logiques :
- les tautologies ;
- les axiomes de l'égalité ;
- ($\exists$-ax) les axiomes du quanteur existentiel.

Règles de déduction :
- (MP) À partir de φ et $\varphi \to \psi$, on peut déduire ψ.
- ($\exists$-intro) Supposons x non libre dans ψ. À partir de $\varphi \to \psi$, on peut déduire $\exists x\varphi \to \psi$.

Définition. Soit φ une $\mathcal{L}$-formule et T une $\mathcal{L}$-théorie. Une *preuve formelle de φ dans T* est une suite finie de $\mathcal{L}$-formules $(\varphi_0, \ldots, \varphi_n)$ avec $\varphi_n = \varphi$ et telle que, pour chaque $i \leq n$, soit $\varphi_i \in T$, soit φ_i est un axiome logique, soit il s'obtient par (MP) à partir de φ_j et φ_k avec $j, k < i$, soit il s'obtient par ($\exists$-intro) à partir d'une formule φ_j avec $j < i$.

On dit que φ est *prouvable dans T*, notée $T \vdash_{\mathcal{L}} \varphi$, s'il existe une preuve formelle de φ dans T.

A priori, $\vdash_{\mathcal{L}}$ dépend du langage $\mathcal{L}$ dans lequel la formule φ et la théorie T sont considérées, car si $\mathcal{L}' \supseteq \mathcal{L}$, il y a plus de $\mathcal{L}'$-preuves que de $\mathcal{L}$-preuves. Nous verrons plus tard que $T \models \varphi$ si et seulement si $T \vdash_{\mathcal{L}} \varphi$, ce qui montrera en particulier que la prouvabilité d'une formule est une notion qui est indépendante du langage.

Dans la suite, on donnera aussi des preuves $(\varphi_0, \ldots, \varphi_n)$ dans lesquelles des formules φ_i dont la prouvabilité a déjà été établie antérieurement seront utilisées comme axiomes. On pourra toujours transformer une telle preuve en une preuve formelle — il suffit de remplacer φ_i par une preuve formelle de φ_i.

On utilisera parfois l'axiome (dérivé) et les règles (dérivées) qui sont donnés dans le lemme suivant.

Lemme 2.6.1.

1. *(MPGen) Si $T \vdash_{\mathcal{L}} \varphi_i$ pour $i = 1, \ldots, n$ et $T \vdash_{\mathcal{L}} \varphi_1 \wedge \ldots \wedge \varphi_n \to \psi$, alors $T \vdash_{\mathcal{L}} \psi$.*
2. *($\forall$-ax) On a $\vdash_{\mathcal{L}} \forall x\, \varphi \to \varphi_{t/x}$ pour toute $\mathcal{L}$-formule φ, tout $\mathcal{L}$-terme t et toute variable x.*
3. *($\forall$-intro) Si x n'est pas libre dans φ et si $T \vdash_{\mathcal{L}} \varphi \to \psi$, alors $T \vdash_{\mathcal{L}} \varphi \to \forall x\, \psi$.*
4. *(Généralisation) Si $T \vdash_{\mathcal{L}} \psi$, alors $T \vdash_{\mathcal{L}} \forall x\psi$.*

Démonstration. (1) La formule suivante est une tautologie :

$$((\varphi_1 \wedge \ldots \wedge \varphi_n \to \psi) \to (\varphi_1 \to (\varphi_2 \to \ldots \to (\varphi_n \to \psi) \cdots).$$

On applique (MP) $n+1$ fois pour obtenir $T \vdash_{\mathcal{L}} \psi$.

(2) Soient φ, t et x donnés. Voici une preuve formelle de $\forall x\, \varphi \to \varphi_{t/x}$ (rappelons que $\forall x\varphi = \neg\exists x \neg\varphi$ par définition) :

$\varphi_0 = (\neg\varphi_{t/x} \to \exists x\, \neg\varphi)$	($\exists$-ax)
$\varphi_1 = (\neg\varphi_{t/x} \to \exists x\, \neg\varphi) \to (\neg\exists x \neg\varphi \to \varphi_{t/x})$	(tautologie)
$\varphi_2 = \forall x\, \varphi \to \varphi_{t/x}$	((MP) sur φ_0 et φ_1)

(3) La preuve de ($\forall$-intro) est similaire que celle de ($\forall$-ax). On l'obtient par contraposition à partir de ($\exists$-intro).

(4) On suppose $T \vdash_{\mathcal{L}} \psi$. On choisit un énoncé prouvable φ.

$\varphi_0 = \varphi$	(prouvable par notre choix)
$\varphi_1 = \psi$	(prouvable dans T par hypothèse)
$\varphi_2 = \psi \to (\varphi \to \psi)$	(tautologie)
$\varphi_3 = \varphi \to \psi$	((MP) sur φ_1 et φ_2)
$\varphi_4 = \varphi \to \forall x\psi$	(($\forall$-intro) appliquée à φ_3)
$\varphi_5 = \forall x\, \psi$	((MP) sur φ_0 et φ_4)

□

Lemme 2.6.2 (Lemme de déduction)**.** *Soit χ un $\mathcal{L}$-énoncé, T une $\mathcal{L}$-théorie et φ une $\mathcal{L}$-formule.*

1. *$T \cup \{\chi\} \vdash_{\mathcal{L}} \varphi$ ssi $T \vdash_{\mathcal{L}} \chi \to \varphi$.*
2. *$T \vdash_{\mathcal{L}} \varphi$ ssi il existe $\psi_1, \ldots, \psi_n \in T$ tels que $\vdash_{\mathcal{L}} (\psi_1 \wedge \cdots \wedge \psi_n) \to \varphi$.*

Démonstration. (1) Il est clair que $T \vdash_{\mathcal{L}} \chi \to \varphi$ entraîne $T \cup \{\chi\} \vdash_{\mathcal{L}} \varphi$. Réciproquement, soit $(\varphi_0, \ldots, \varphi_n)$ une preuve formelle de $\varphi = \varphi_n$ dans $T \cup \{\chi\}$. Par induction sur i, on montrera que $T \vdash_{\mathcal{L}} (\chi \to \varphi_i)$. Si $\varphi_i = \chi$, c'est clair, et si $T \vdash_{\mathcal{L}} \varphi_i$, cela suit de (MP) et du fait que $(\varphi_i \to (\chi \to \varphi_i))$ est une tautologie. Cela montre $T \vdash_{\mathcal{L}} (\chi \to \varphi_i)$ dans les cas où φ_i est un axiome logique ou un élement de T.

Si φ_i s'obtient par (MP) à partir de φ_j et $\varphi_k = (\varphi_j \to \varphi_i)$ pour $j, k < i$, il suffit d'utiliser (MP) et le fait que $((\chi \to \varphi_j) \wedge (\chi \to (\varphi_j \to \varphi_i)) \to (\chi \to \varphi_i))$ est une tautologie.

Enfin, si $\varphi_i = (\exists x\psi \to \varphi)$ s'obtient par ($\exists$-intro) à partir de $\varphi_j = (\psi \to \varphi)$ pour $j < i$ (donc x n'est pas libre dans φ), on utilise la tautologie $(F \to (G \to H)) \to (G \to (F \to H))$ deux fois et le fait que x n'est pas libre dans $(\chi \to \varphi)$ pour conclure :

$$\begin{aligned} T \vdash_{\mathcal{L}} \chi \to (\psi \to \varphi) &\Rightarrow T \vdash_{\mathcal{L}} \psi \to (\chi \to \varphi) \\ (\text{par } \exists\text{-intro}) &\Rightarrow T \vdash_{\mathcal{L}} \exists x\psi \to (\chi \to \varphi) \\ &\Rightarrow T \vdash_{\mathcal{L}} \chi \to (\exists x\psi \to \varphi) \end{aligned}$$

La partie (2) se déduit facilement de (1). □

Lemme 2.6.3 (Simulation des constantes par des variables)**.** *Soit ψ une $\mathcal{L}$-formule, T une $\mathcal{L}$-théorie, et C un ensemble de constantes tel que $\mathcal{L} \cap C = \emptyset$.*

(a) Soit x une variable et $c \in C$. Les propriétés suivantes sont équivalentes :

(1) $T \vdash_{\mathcal{L}} \psi$

(2) $T \vdash_{\mathcal{L}\cup\{c\}} \psi_{c/x}$

(3) $T \vdash_{\mathcal{L}\cup\{c\}} \psi$

(b) On a $T \vdash_{\mathcal{L}} \psi$ si et seulement si $T \vdash_{\mathcal{L}\cup C} \psi$.

Démonstration. On supposera dans la preuve que $T = \emptyset$. La preuve du cas général est la même. Montrons d'abord la partie (a).

(1)⇒(3) est triviale. (Toute $\mathcal{L}$-preuve est une $\mathcal{L} \cup \{c\}$-preuve.)

(3)⇒(2). Si $\vdash_{\mathcal{L}\cup\{c\}} \psi$, alors $\vdash_{\mathcal{L}\cup\{c\}} \forall x\psi$ par généralisation. En utilisant (∀-ax) et (MP), on obtient $\vdash_{\mathcal{L}\cup\{c\}} \psi_{c/x}$.

(2)⇒(1). Pour $\tilde{\varphi}$ une $\mathcal{L} \cup \{c\}$-formule et y une variable sans occurrence dans $\tilde{\varphi}$, on notera $\varphi = \tilde{\varphi}_{y/c}$ le mot obtenu en substituant toute occurrence de c dans $\tilde{\varphi}$ par y.

Soient $\tilde{\varphi}, \tilde{\psi}$ et $\tilde{\chi}$ des $\mathcal{L} \cup \{c\}$-formules sans occurrences de y. Alors on a les propriétés suivantes :

(i) φ est une $\mathcal{L}$-formule.
(ii) $\varphi_{c/y} = \tilde{\varphi}$.
(iii) Si $\tilde{\varphi}$ est un axiome d'égalité (respectivement une tautologie ou un ∃-axiome), alors φ aussi.
(iv) Si $\tilde{\varphi}$ s'obtient par (MP) à partir de $\tilde{\psi}$ et $\tilde{\chi}$, alors φ s'obtient par (MP) à partir de ψ et χ.
(v) Si $\tilde{\varphi}$ s'obtient par (∃-intro) à partir de $\tilde{\psi}$, alors φ s'obtient par (∃-intro) à partir de ψ.

On peut montrer (i) par induction sur la hauteur de φ, et (ii) est une conséquence de l'exemple 2.4.3(2). Les propriétés (iv) et (v) sont immédiates, de même les cas d'un axiome d'égalite ou d'une tautologie dans (iii). Reste à vérifier le cas d'un (∃-ax) dans (iii) : Si $\tilde{\varphi} = (\tilde{\delta}_{t/x} \to \exists x\tilde{\delta})$, il faut monter que $\varphi = ([\tilde{\delta}_{t/x}]_{y/c} \to \exists x[\tilde{\delta}]_{y/c})$ est un axiome de quanteur existentiel. Ceci est une conséquence de l'identité

$$[\tilde{\delta}_{t/x}]_{y/c} = [\tilde{\delta}_{y/c}]_{t_{y/c}/x}$$

dont la preuve se fait par induction et est laissée en exercice.

Maintenant, soit $(\tilde{\varphi}^0, \ldots, \tilde{\varphi}^m)$ une $\mathcal{L}\cup\{c\}$-preuve formelle de $\psi_{c/x}$. On choisit une variable y sans occurrence dans $\tilde{\varphi}^0, \ldots, \tilde{\varphi}^m$ et telle que $y \neq x$.

On déduit des propriétés (i)-(v) que $(\varphi^0, \ldots, \varphi^m)$ est une $\mathcal{L}$-preuve formelle de $\varphi^m = (\psi_{c/x})_{y/c} = \psi_{y/x}$. (On utilise 2.4.3(2) pour établir cette égalité.) En généralisant on obtient $\vdash_{\mathcal{L}} \forall y\psi_{y/x}$.

On utilise (∀-ax) et (MP) pour montrer $\vdash_{\mathcal{L}} (\psi_{y/x})_{x/y}$. Or, le lemme 2.4.4 montre que $\left(\psi_{y/x}\right)_{x/y} = \psi$, ce qui termine la preuve de (2)⇒(1).

La partie (b) est une conséquence immédiate de (a). En effet, comme toute $\mathcal{L} \cup C$-preuve n'utilise qu'un nombre fini de constantes de C, on peut supposer que C est fini. On termine donc par une induction sur le cardinal de C, en utilisant (1)⇔(3). □

Définition. Soit T une $\mathcal{L}$-théorie.
- T est *contradictoire* s'il existe un énoncé φ tel que $T \vdash_{\mathcal{L}} \varphi$ et $T \vdash_{\mathcal{L}} \neg\varphi$.
- T est *cohérente* si elle est non contradictoire.
- T est *complète* si elle est cohérente et pour tout $\mathcal{L}$-énoncé φ on a soit $T \vdash_{\mathcal{L}} \varphi$, soit $T \vdash_{\mathcal{L}} \neg\varphi$.

Remarque 2.6.4. *Pour une $\mathcal{L}$-théorie T, sont équivalents :*

1. *T est contradictoire ;*
2. *$T \vdash_{\mathcal{L}} \varphi$ pour tout $\mathcal{L}$-énoncé φ ;*
3. *il existe $\psi_1, \ldots, \psi_n \in T$ tels que $\vdash_{\mathcal{L}} \neg(\psi_1 \wedge \cdots \wedge \psi_n)$.*

Démonstration. Exercice. □

Lemme 2.6.5. *Soit T une $\mathcal{L}$-théorie.*

1. *Soit ψ un $\mathcal{L}$-énoncé. Alors $T \vdash_{\mathcal{L}} \psi$ ssi $T \cup \{\neg\psi\}$ est contradictoire.*
2. *Soit $\varphi(x)$ une $\mathcal{L}$-formule et $c \in \mathcal{L}$ une constante qui n'a pas d'occurrence dans $T \cup \{\varphi(x)\}$. On suppose que T est cohérente. Alors la théorie $T \cup \{\exists x\varphi \to \varphi_{c/x}\}$ est une $\mathcal{L}$-théorie cohérente.*

Démonstration. (1) Le sens direct est clair. Réciproquement, si $T \cup \{\neg\psi\}$ est contradictoire, par le lemme de déduction, il existe un $\mathcal{L}$-énoncé φ tel que $T \vdash_{\mathcal{L}} \psi \to \varphi$ et $T \vdash_{\mathcal{L}} \psi \to \neg\varphi$. On utilise la tautologie $((\psi \to \varphi) \to (\psi \to \neg\varphi)) \to \psi$ et (MP) pour conclure.

(2) Si la conclusion était fausse, on aurait $T \vdash_{\mathcal{L}} \exists x\varphi \wedge \neg\varphi_{c/x}$ par la première partie du lemme. Par le lemme de déduction il existerait donc $\psi_1, \ldots, \psi_n \in T$ tels que $\vdash_{\mathcal{L}} (\psi_1 \wedge \ldots \wedge \psi_n) \to (\exists x\varphi \wedge \neg\varphi_{c/x})$.

Nous allons montrer que si $\vdash_{\mathcal{L}} \psi \to (\exists x\varphi \wedge \neg\varphi_{c/x})$ pour un $\mathcal{L}$-énoncé ψ qui ne contient pas c, alors $\vdash_{\mathcal{L}} \neg\psi$. En effet, on a clairement $\vdash_{\mathcal{L}} \neg\exists x\varphi \to \neg\psi$ et $\vdash_{\mathcal{L}} \varphi_{c/x} \to \neg\psi$. Le dernier donne $\vdash_{\mathcal{L}} \varphi \to \neg\psi$ par le Lemme 2.6.3(a), puis $\vdash_{\mathcal{L}} \exists x\varphi \to \neg\psi$ par ($\exists$-intro). En utilisant des tautologies convenables et (MP), on obtient $\vdash_{\mathcal{L}} \neg\psi$. □

Exemples 2.6.6.
1. Soit $\mathfrak{A}$ une $\mathcal{L}$-structure. Alors $\mathrm{Th}(\mathfrak{A}) := \{\varphi\ \mathcal{L}\text{-énoncé} \mid \mathfrak{A} \models \varphi\}$ est une théorie, appelée la *théorie de* $\mathfrak{A}$. C'est une théorie complète. (Sa cohérence sera une conséquence de la direction facile du théoréme de complétude.)
2. La théorie des corps algébriquement clos est la $\mathcal{L}_{an}$-théorie CAC qui outre les axiomes de corps contient un énoncé χ_n pour tout $n \geq 1$ exprimant que tout polynôme de degré n a une racine.
   [$\chi_n = \forall z_0, \ldots, z_{n-1} \exists x(x^n + z_{n-1}x^{n-1} + \ldots + z_0 \doteq 0)$ marche ; cette formule peut être réécrite dans $\mathcal{L}_{an}$.]

Théorème 2.6.7 (Théorème de complétude de Gödel)**.** *Soit T une $\mathcal{L}$-théorie et φ un $\mathcal{L}$-énoncé. Alors $T \models \varphi$ si et seulement si $T \vdash_{\mathcal{L}} \varphi$.*

Remarque. *Comme $\models$ ne dépend pas du langage considéré, a postiori $\vdash_{\mathcal{L}}$ n'en dépend pas non plus. Une fois le théorème de complétude établi, nous écrivons donc $\vdash$ au lieu de $\vdash_{\mathcal{L}}$.*

Démonstration. $T \vdash_{\mathcal{L}} \varphi \Rightarrow T \models \varphi$. Ce sens facile exprime que notre calcul de preuves formelles est correct. En effet, $\vdash_{\mathcal{L}} \varphi \Rightarrow \models \varphi$ est une conséquence des lemmes 2.5.4–2.5.8. Si $T \vdash_{\mathcal{L}} \varphi$, il existe $\psi_1, \ldots, \psi_n \in T$ tels que $\vdash_{\mathcal{L}} (\psi_1 \wedge \ldots \wedge \psi_n) \to \varphi$. Donc $\models (\psi_1 \wedge \ldots \wedge \psi_n) \to \varphi$ et alors pour tout $\mathfrak{A} \models T$ on a $\mathfrak{A} \models \varphi$ (car $\mathfrak{A} \models \psi_1 \wedge \ldots \wedge \psi_n$).

$T \models \varphi \Rightarrow T \vdash_{\mathcal{L}} \varphi$. C'est le sens non trivial. Il sera conséquence du théorème d'existence de modèle suivant. En fait, comme

$$\begin{aligned} T \not\models \varphi &\iff T \cup \{\neg\varphi\} \text{ a un modèle} \quad \text{(par définition), et} \\ T \not\vdash_{\mathcal{L}} \varphi &\iff T \cup \{\neg\varphi\} \text{ est cohérente} \quad \text{(Lemme 2.6.5(1)),} \end{aligned}$$

ce théorème est équivalent au théorème de complétude. □

Théorème 2.6.8. *Une théorie a un modèle si et seulement si elle est cohérente.*

Démonstration. La direction facile a déja été montrée. Pour montrer que toute théorie cohérente T a un modèle, l'idée est de construire une expansion T^+ de T dans un langage $\mathcal{L}^+ \supseteq \mathcal{L}$ qui ressemble au « diagramme complet d'une structure ». Nous avons besoin d'une définition.

Définition. Soit $\mathcal{L}$ un langage et C un ensemble de constantes avec $\mathcal{L} \cap C = \emptyset$. On dit qu'une $\mathcal{L} \cup C$-théorie T^+ *admet des témoins de Henkin dans C* si pour toute $\mathcal{L} \cup C$-formule $\varphi = \varphi(x)$ il existe $c \in C$ tel que $\exists x\varphi \to \varphi_{c/x} \in T^+$.

Si $\mathfrak{A}$ est une $\mathcal{L}$-structure et $A = \{a_c \mid c \in C\}$ une énumération (non nécessairement injective) de son ensemble de base par C, on note $\mathfrak{A}^+$ la $\mathcal{L} \cup C$-structure obtenue à partir de $\mathfrak{A}$ en interprétant c par a_c. Alors $\mathrm{Th}(\mathfrak{A}^+)$ est une théorie complète qui admet des témoins de Henkin dans C. En fait, toute théorie complète qui admet des témoins de Henkin dans C est de cette forme :

Proposition 2.6.9. *Toute $\mathcal{L} \cup C$-théorie complète T^+ qui admet des témoins de Henkin dans C a un modèle $\mathfrak{A}^+$ formé de constantes de C, c'est-à-dire dont l'ensemble de base est de la forme $A^+ = \{c^{\mathfrak{A}^+} \mid c \in C\}$.*

Démonstration de 2.6.9. Quitte à remplacer T^+ par $\{\varphi \ \mathcal{L} \cup C\text{-énoncé} \mid T^+ \vdash_{\mathcal{L}\cup C} \varphi\}$, on peut supposer que T^+ est *déductivement close* ($T \vdash_{\mathcal{L}\cup C} \varphi$ ssi $\varphi \in T^+$).

Sur C, la relation $c \sim d :\iff c \doteq d \in T^+$ est une relation d'équivalence. En effet, cela découle des axiomes de l'égalité. On pose $a_c := c/\sim$ et $A^+ := \{a_c \mid c \in C\}$. Évidemment, on a $A^+ \neq \emptyset$. On définit une $\mathcal{L} \cup C$-structure $\mathfrak{A}^+$ sur A^+ comme suit :

– Pour d une constante, on pose $d^{\mathfrak{A}^+} := a_c$ si $c \dot{=} d \in T^+$. Pour montrer qu'un tel c existe, on raisonne ainsi : $d \dot{=} d \to \exists x x \dot{=} d \in T^+$ (par ($\exists$-ax)), donc $\exists x x \dot{=} d \in T^+$ par (MP) ; comme T^+ admet des témoins de Henkin, il existe $c \in C$ tel que $\exists x x \dot{=} d \to c \dot{=} d \in T^+$, d'où $c \dot{=} d \in T^+$ par (MP). C'est bien défini par les axiomes de l'égalité.

– $R \in \mathcal{R}_n^{\mathcal{L}}$. On définit $(a_{c_1}, \ldots, a_{c_n}) \in R^{\mathfrak{A}^+} :\iff Rc_1 \cdots c_n \in T^+$. C'est bien défini par les axiomes de l'égalité (congruence 2).

– $f \in \mathcal{F}_n^{\mathcal{L}}$. On définit $f^{\mathfrak{A}^+}(a_{c_1}, \ldots, a_{c_n}) = a_{c_0} :\iff fc_1 \cdots c_n \dot{=} c_0 \in T^+$. On montre sans problème en utilisant les axiomes de l'égalité (congruence 1) que c'est bien défini. Un argument similaire à celui donné pour les constantes montre que $f^{\mathfrak{A}^+}$ est une fonction qui est définie partout.

(I) Si t est un $\mathcal{L} \cup C$-terme sans variable, alors $t^{\mathfrak{A}^+} = a_c \iff t \dot{=} c \in T^+$.

(II) Soit ψ un $\mathcal{L} \cup C$-énoncé. Alors $\mathfrak{A}^+ \models \psi \iff \psi \in T^+$.

On démontre (I) par induction sur $\mathrm{ht}(t)$, en utilisant les axiomes de l'égalité.

Quant à (II), on raisonne par induction sur $\mathrm{ht}(\psi)$, où ψ est un $\mathcal{L} \cup C$-*énoncé*. Rappelons que $\mathrm{ht}(\varphi) = \mathrm{ht}(\varphi_{\overline{s}/\overline{x}})$ (voir l'exercice 2.4.1).

Si $\psi = t_1 \dot{=} t_2$, alors (II) suit de (I).

Si $\psi = Rt_1 \cdots t_n$, on choisit $c_i \in C$ tel que $t_i^{\mathfrak{A}^+} = a_{c_i}$. Comme $t_i \dot{=} c_i \in T^+$, on a $\mathfrak{A}^+ \models \psi \iff Rc_1 \cdots c_n \in T^+ \iff \psi \in T^+$. La première équivalence suit de la définition de $\mathfrak{A}^+$, la seconde suit de la congruence 2.

Le cas $\psi = (\varphi_1 \wedge \varphi_2)$ est facile.

Si $\psi = \neg\varphi$, on a $\mathfrak{A}^+ \models \psi \iff \mathfrak{A}^+ \not\models \varphi \iff \varphi \notin T^+ \iff \psi \in T^+$, puisque T^+ est complète par hypothèse.

Finalement, soit $\psi = \exists x \varphi$. On montre d'abord que $\psi \in T^+ \iff$ il existe $c \in C$ tel que $\varphi_{c/x} \in T^+$. En effet, comme T^+ admet des témoins de Henkin dans C, on a le sens direct. Pour le sens indirect, il suffit d'appliquer ($\exists$-ax), (MP) et utiliser que T^+ est déductivement close.

$$\begin{aligned} \text{On obtient les équivalences} \quad \mathfrak{A}^+ \models \psi &\iff \mathfrak{A}^+ \models \varphi[a_c] \text{ pour un } c \in C \\ \text{(Lemme de substitution)} &\iff \mathfrak{A}^+ \models \varphi_{c/x} \text{ pour un } c \in C \\ \text{(Hypothèse d'induction)} &\iff \varphi_{c/x} \in T^+ \text{ pour un } c \in C \\ &\iff \psi \in T^+ \end{aligned}$$

$\mathfrak{A}^+$ est donc un modèle de T^+ qui est formé de constantes de C. □

Le lemme suivant terminera la preuve du théorème 2.6.8, car si $\mathfrak{A}^+$ est un modèle de la $\mathcal{L}^+ = \mathcal{L} \cup C$-théorie T^+ avec $T^+ \supseteq T$, alors le réduit de $\mathfrak{A}^+$ au langage $\mathcal{L}$ est un modèle de T. □

Lemme 2.6.10. *Soit T une $\mathcal{L}$-théorie cohérente. Alors il existe une $\mathcal{L} \cup C$-théorie complète T^+ contenant T qui admet des témoins de Henkin dans C.*

Démonstration. Pour toute $\mathcal{L}$-formule à une variable libre $\varphi = \varphi(x)$ on introduit une nouvelle constante c_φ. Soit C_1 l'ensemble des c_φ. On pose $\mathcal{L}_1 = \mathcal{L} \cup C_1$ et $T_1 := \tilde{T} := T \cup \{\exists x \varphi \to \varphi_{c_\varphi/x} \mid \varphi = \varphi(x)\}$. Montrons que T_1 est une $\mathcal{L}_1$-théorie cohérente. Comme une théorie est cohérente si et seulement si toute partie finie

est cohérente, il suffit évidemment de montrer que la théorie $\tilde{T}' = T \cup \{\exists x_i \varphi^i \to \varphi^i_{c_{\varphi^i}/x_i}\}$ est cohérente pour tout ensemble fini $\{\varphi^1, \ldots, \varphi^n\}$ de formules. Comme T est cohérente en tant que $\mathcal{L}_1$-théorie par 2.6.3, c'est une conséquence du lemme 2.6.5(2), quitte à raisonner par induction sur n.

On répète cette construction, avec C_2 un ensemble de nouvelles constantes, $\mathcal{L}_2 := \mathcal{L}_1 \cup C_2$ et la $\mathcal{L}_2$-théorie $T_2 := \tilde{T}_1$. On montre qu'il s'agit d'une théorie cohérente. Si T_n est construite, on trouve $T_{n+1} := \tilde{T}_n$, une théorie cohérente dans le langage $\mathcal{L}_{n+1} = \mathcal{L}_n \cup C_{n+1}$. Posons $C := \bigcup_{n\in\mathbb{N}} C_n$ et $\mathcal{L}^+ := \mathcal{L} \cup C$. Alors $(T_n)_{n\in\mathbb{N}}$ est une suite croissante de $\mathcal{L}^+$-théories cohérentes. La $\mathcal{L}^+$-théorie $T' := \bigcup_{n\in\mathbb{N}} T_n$ est alors cohérente. Par construction, elle admet des témoins de Henkin dans C.

Comme toute $\mathcal{L}^+$-théorie S' contenant T' admet des témoins de Henkin dans C, il reste à montrer que toute $\mathcal{L}^+$-théorie cohérente est contenue dans une $\mathcal{L}^+$-théorie complète. C'est une conséquence du Lemme de Zorn. En effet, l'ensemble $\mathcal{S} = \{S' \supseteq T' \mid S' \text{ et une } \mathcal{L}^+\text{-théorie cohérente}\}$ est non vide et ordonné partiellement par inclusion. Comme la réunion d'une chaîne de théories cohérentes est cohérente, cet ordre est inductif. Par le lemme de Zorn, il existe un élément maximal dans $\mathcal{S}$, autrement dit il existe une $\mathcal{L}^+$-théorie $T^+ \supseteq T'$ qui est maximale cohérente. Soit φ un $\mathcal{L}^+$-énoncé. Si $T \nvdash_{\mathcal{L}^+} \varphi$, la $\mathcal{L}^+$-théorie $T^+ \cup \{\neg\varphi\}$ est cohérente par 2.6.5(1). On en déduit que $\neg\varphi \in T^+$ par maximalité, ce qui montre que T^+ est complète et termine la preuve.

Si $\mathcal{L}$ est dénombrable, le lemme admet une preuve par une construction plus directe. Nous allons l'esquisser maintenant. On choisit un ensemble $C = \{c_n \mid n \in \mathbb{N}\}$ de nouvelles constantes et on énumère l'ensemble des $\mathcal{L} \cup C$-énoncés via $\varphi_1, \varphi_2, \ldots$. Par induction, on construit une suite croissante de $\mathcal{L} \cup C$-théories cohérentes ayant les propriétés suivantes :

- $T_0 = T$, et $T_n \setminus T$ est fini pour tout $n \in \mathbb{N}$;
- $\varphi_n \in T_{n+1}$ ou $\neg\varphi_n \in T_{n+1}$;
- si $\varphi_n = \exists x\psi$, alors il existe $c \in C$ tel que $\psi_{c/x} \in T_{n+1}$.

Une fois la suite des T_n construite, il suffit de poser $T^+ := \bigcup_{n\in\mathbb{N}} T_n$. C'est une théorie complète par construction, et elle admet des témoins de Henkin dans C. En effet, soit $\psi(x)$ donnée. Alors il existe $n \in \mathbb{N}$ tel que $\exists x\psi = \varphi_n$. Si $\varphi_n \in T^+$, alors par construction $\psi_{c/x} \in T^+$ pour un $c \in C$, donc $\exists x\psi \to \psi_{c/x} \in T^+$. Sinon, on a $\neg\exists x\psi \in T^+$, d'où trivialement $\exists x\psi \to \psi_{c/x} \in T^+$ pour tout c.

Supposons T_n construite. Si $T_n \cup \{\varphi_n\}$ est cohérente, on pose $\psi_n = \varphi_n$. Sinon, $\psi_n = \neg\varphi_n$. Dans les deux cas, $T_n \cup \{\psi_n\}$ est cohérente. Si ψ_n n'est pas de la forme $\exists x\chi$, on pose $T_{n+1} := T_n \cup \{\psi_n\}$. Sinon, soit m minimal tel que c_m n'a pas d'occurrence dans T_n (il existe car $T_n \setminus T$ est fini). On pose $T_{n+1} := T_n \cup \{\psi_n, \chi_{c_m/x}\}$. La théorie T_{n+1} est cohérente par le lemme 2.6.5(2). □

Chapitre 3

Premiers pas en théorie des modèles

3.1 Quelques théorèmes fondamentaux

Théorème 3.1.1 (Théorème de compacité)**.** *Soit T une théorie dont toute partie finie a un modèle. Alors T a un modèle.*

Démonstration. La théorie T est cohérente si et seulement si toute partie finie de T est cohérente. Le résultat suit du théorème 2.6.8. □

Exercice. Soit X l'ensemble des $\mathcal{L}$-théories T complètes et closes par déduction. Sur X, on définit une topologie de la manière suivante. Pour φ un $\mathcal{L}$-énoncé, on note $\langle\varphi\rangle := \{T \in X \mid \varphi \in T\}$, et on prend la collection des $\langle\varphi\rangle$ comme base d'ouverts d'une topologie. Montrer :

1. $\langle\varphi \wedge \psi\rangle = \langle\varphi\rangle \cap \langle\psi\rangle$ et $\langle\neg\varphi\rangle = X \setminus \langle\varphi\rangle$. En particulier, les $\langle\varphi\rangle$ forment une base d'ouverts-fermés.
2. Cette topologie est compacte et séparée.

Définition. Soient $\mathfrak{M}$ et $\mathfrak{N}$ des $\mathcal{L}$-structures.

1. On dit que $\mathfrak{M}$ et $\mathfrak{N}$ sont *élémentairement équivalentes*, si $\mathrm{Th}(\mathfrak{M}) = \mathrm{Th}(\mathfrak{N})$, c'est-à-dire si elles satisfont les mêmes énoncés. On le note $\mathfrak{M} \equiv \mathfrak{N}$.
2. Soit $\mathfrak{M}$ une sous-structure de $\mathfrak{N}$ (ce que l'on note $\mathfrak{M} \subseteq \mathfrak{N}$). On dit que c'est une *sous-structure élémentaire* (et $\mathfrak{N}$ est appelée *extension élémentaire* de $\mathfrak{M}$) si pour toute $\mathcal{L}$-formule $\varphi = \varphi(x_1, \ldots, x_n)$ et tout uplet $\overline{a} = (a_1, \ldots, a_n) \in M^n$ on a $\mathfrak{M} \models \varphi[\overline{a}]$ si et seulement si $\mathfrak{N} \models \varphi[\overline{a}]$. On le note $\mathfrak{M} \preccurlyeq \mathfrak{N}$.

Remarque. *1. Si $\mathfrak{M} \preccurlyeq \mathfrak{N}$, alors $\mathfrak{M} \equiv \mathfrak{N}$.*

2. Si $\mathfrak{M}$ et $\mathfrak{N}$ sont isomorphes, alors $\mathfrak{M} \equiv \mathfrak{N}$. □

L'exemple suivant montre que $\mathfrak{M} \subseteq \mathfrak{N}$ et $\mathfrak{M} \equiv \mathfrak{N}$ n'entraîne pas toujours $\mathfrak{M} \preccurlyeq \mathfrak{N}$.

Exemple. Dans le langage $\mathcal{L}_<$, on considère $\mathfrak{N} := \langle \mathbb{N}; < \rangle$ et $\mathfrak{M} := \langle \mathbb{N}^*; < \rangle$, où $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$.

On a $\mathfrak{M} \subseteq \mathfrak{N}$ et $\mathfrak{M} \cong \mathfrak{N}$, en particulier $\mathfrak{M} \equiv \mathfrak{N}$. Or, l'extension $\mathfrak{M} \subseteq \mathfrak{N}$ n'est pas élémentaire, puisque $\mathfrak{M} \models \neg \exists x\, x < 1$ et $\mathfrak{N} \models \exists x\, x < 1$.

Théorème 3.1.2 (Test de Tarski-Vaught)**.** *Soient $\mathfrak{M} \subseteq \mathfrak{N}$ deux $\mathcal{L}$-structures. On suppose que pour toute $\mathcal{L}$-formule $\varphi(x_0, \ldots, x_n)$ et tout uplet $(a_1, \ldots, a_n) \in M^n$, s'il existe $b_0 \in N$ tel que $\mathfrak{N} \models \varphi[b_0, a_1, \ldots, a_n]$, alors il existe $a_0 \in M$ tel que $\mathfrak{N} \models \varphi[a_0, a_1, \ldots, a_n]$. Alors $\mathfrak{M} \preccurlyeq \mathfrak{N}$.*

Démonstration. Par induction sur $\mathrm{ht}(\varphi(x_1, \ldots, x_n))$ on montre : pour tout $\overline{a} \in M^n$ on a $\mathfrak{M} \models \varphi[\overline{a}] \iff \mathfrak{N} \models \varphi[\overline{a}]$.

Si φ est une formule atomique, c'est vrai car $\mathfrak{M}$ est une sous-structure de $\mathfrak{N}$.

Le cas des connecteurs logiques est clair.

Soit donc $\varphi = \varphi(x_1, \ldots, x_n) = \exists x_0 \psi(x_0, \ldots, x_n)$. (Comme x_0 n'est pas libre dans φ, on peut supposer que $x_0 \neq x_i$ pour $i = 1, \ldots, n$.) Soit $(a_1, \ldots, a_n) \in M^n$.

$$\begin{aligned}
\text{On a } \mathfrak{M} \models \varphi[a_1, \ldots, a_n] &\iff \text{il existe } a_0 \in M \text{ tel que } \mathfrak{M} \models \psi[a_0, \ldots, a_n] \\
\text{(hypothèse d'induction)} &\iff \text{il existe } a_0 \in M \text{ tel que } \mathfrak{N} \models \psi[a_0, \ldots, a_n] \\
\text{(hypothèse du théorème)} &\iff \text{il existe } b_0 \in N \text{ tel que } \mathfrak{N} \models \psi[b_0, \ldots, a_n] \\
&\iff \mathfrak{N} \models \varphi[a_1, \ldots, a_n].
\end{aligned}$$

□

Notons que pour tout langage $\mathcal{L}$ on a $\mathrm{card}(\mathcal{L}) = \mathrm{card}(\mathcal{F}ml^{\mathcal{L}}) \geq \mathrm{card}(\mathcal{T}^{\mathcal{L}})$. (Exercice.)

Théorème 3.1.3 (Théorème de Löwenheim-Skolem descendant)**.** *Soit $\mathfrak{M}$ une $\mathcal{L}$-structure et $A \subseteq M$ une partie de l'ensemble de base de $\mathfrak{M}$. On suppose que $\mathrm{card}(M) \geq \mathrm{card}(\mathcal{L})$. Alors il existe une sous-structure élémentaire $\mathfrak{M}_0$ de $\mathfrak{M}$ contenant A et de cardinalité $\sup(\mathrm{card}(A), \mathrm{card}(\mathcal{L}))$.*

Démonstration. Quitte à agrandir A on peut supposer que $\mathrm{card}(A) \geq \mathrm{card}(\mathcal{L})$.

Dans cette preuve, si $\emptyset \neq B \subseteq M$, nous notons $\tilde{B} \subseteq M$ l'ensemble de base de $\langle B \rangle_{\mathfrak{M}}$ (la sous-structure engendrée par B). On a $\tilde{B} = \{t^{\mathfrak{M}}[b_1, \ldots, b_n] \mid t = t(x_1, \ldots, x_n) \in \mathcal{T}^{\mathcal{L}}, b_1, \ldots, b_n \in B\}$ (Exercice 2.3.3). En particulier, si $\mathrm{card}(B) \geq \mathrm{card}(\mathcal{L})$, alors $\mathrm{card}(B) = \mathrm{card}(\tilde{B})$, puisqu'il existe une surjection de $\mathcal{T}^{\mathcal{L}} \times \bigcup_{n \in \mathbb{N}} B^n$ sur $\tilde{B}$.

Soit $\mathfrak{A}_0 := \langle A \rangle_{\mathfrak{M}}$ et $A_0 := \tilde{A}$. Étant défini $A_i \subseteq M$, on construit A_{i+1} comme suit. Pour toute $\mathcal{L}$-formule $\varphi(x_0, \ldots, x_n)$ et tout n-uplet $\overline{a} \in A_i^n$, si $\mathfrak{M} \models \varphi[b_0, \overline{a}]$ pour un $b_0 \in M$, on choisit $c(\varphi, \overline{a}) \in M$ tel que $\mathfrak{M} \models \varphi[c(\varphi, \overline{a}), \overline{a}]$. On pose $B_{i+1} := A_i \cup \{c(\varphi, \overline{a}) \mid \varphi = \varphi(x_0, \ldots, x_n) \in \mathcal{F}ml^{\mathcal{L}}, a_1, \ldots, a_n \in A_i\}$, puis $A_{i+1} := \widetilde{B_{i+1}}$, l'ensemble de base de $\mathfrak{A}_{i+1} = \langle B_{i+1} \rangle_{\mathfrak{M}}$.

Soit $M_0 := \bigcup_{i \in \mathbb{N}} A_i$. Alors M_0 contient $c^{\mathfrak{M}}$ pour tout $c \in \mathcal{C}^{\mathcal{L}}$ et est clos par $f^{\mathfrak{M}}$ pour tout $f \in \mathcal{F}^{\mathcal{L}}$. C'est donc l'ensemble de base d'une sous-structure $\mathfrak{M}_0$ de $\mathfrak{M}$.

Soit $\varphi = \varphi(x_0, \ldots, x_n)$ une formule, $\overline{a} \in M_0^n$ et $b_0 \in M$ tel que $\mathfrak{M} \models \varphi[b_0, \overline{a}]$. Il existe $N \in \mathbb{N}$ tel que $a_1, \ldots, a_n \in A_N$. Par construction de A_{N+1} il existe $c_0 \in A_{N+1} \subseteq M_0$ tel que $\mathfrak{M} \models \varphi[c_0, \overline{a}]$. Donc $\mathfrak{M}_0 \preccurlyeq \mathfrak{M}$ par le test de Tarski-Vaught. Le fait que $\mathrm{card}(M_0) = \mathrm{card}(A)$ est clair. □

Exemples 3.1.4. Soit $\mathfrak{R} = \langle \mathbb{R}; +, -, 0, 1, \cdot, < \rangle$ le corps ordonné des réels.

1. Il existe $\mathfrak{R}' \equiv \mathfrak{R}$ avec $\mathfrak{R}'$ non archimédien, c'est-à-dire contenant un élément $\varepsilon > 0$ tel que $n \cdot \varepsilon = \underbrace{\varepsilon + \ldots + \varepsilon}_{n \text{ fois}} < 1$ pour tout $n \in \mathbb{N}$. Un tel ε est appelé *infinitésimal*.
2. Il existe $\mathfrak{R}' \preccurlyeq \mathfrak{R}$ tel que $\mathfrak{R}'$ ne soit pas complet.

Démonstration. (1) Soit c une nouvelle constante et $\mathcal{L} := \mathcal{L}_{c.ord} \cup \{c\}$. On considère la $\mathcal{L}$-théorie $T := \mathrm{Th}(\mathfrak{R}) \cup \{0 < c\} \cup \{\varphi_n = \underbrace{c + \ldots + c}_{n \text{ fois}} < 1 \mid n \in \mathbb{N}\}$. Soit $T_0 \subseteq T$ une partie finie. Alors il existe $N \in \mathbb{N}$ tel que $\varphi_m \notin T_0$ pour tout $m \geq N$. L'expansion de $\mathfrak{R}$ à $\mathcal{L}$ obtenue en interprétant c par $1/N \in \mathbb{R}$ est donc un modèle de T_0. Par compacité, T a un modèle $\mathfrak{R}'$. L'élément $\varepsilon = c^{\mathfrak{R}'}$ montre que $\mathfrak{R}'$ n'est pas archimédien.

(2) Par le théorème de Löwenheim-Skolem descendant (pour $A = \emptyset$), il existe $\mathfrak{R}' \preccurlyeq \mathfrak{R}$ avec $\mathbb{R}'$ dénombrable. On a $\mathbb{Q} \subseteq \mathbb{R}'$, car $\mathfrak{R}'$ est un corps. On choisit $r \in \mathbb{R} \setminus \mathbb{R}'$. Alors la partie $\{r' \in \mathbb{R}' \mid r' < r\}$ n'a pas de supremum dans $\mathbb{R}'$, car $\mathbb{R}'$ est dense dans $\mathbb{R}$. □

Remarque. *On peut montrer que* $\mathfrak{R}_{alg} \preccurlyeq \mathfrak{R}$*, où* $\mathfrak{R}_{alg} = \langle \mathbb{R}_{alg}; +, -, 0, 1, \cdot, < \rangle$ *et* $\mathbb{R}_{alg} = \{r \in \mathbb{R} \mid$ *il existe* $0 \neq p(X) \in \mathbb{Q}[X]$ *tel que* $p(r) = 0\}$.

Définition. Soit $\mathcal{P}$ une propriété que chaque $\mathcal{L}$-structure est susceptible de vérifier ou non.

On dit que $\mathcal{P}$ est *axiomatisable* (resp. *finiment axiomatisable*) s'il existe une $\mathcal{L}$-théorie T (resp. un $\mathcal{L}$-énoncé φ) telle que pour toute $\mathcal{L}$-structure $\mathfrak{M}$, la propriété $\mathcal{P}$ est satisfaite par $\mathfrak{M}$ si et seulement si $\mathfrak{M} \models T$ (resp. $\mathfrak{M} \models \varphi$).

Dans ce cas, on dit que T (resp. φ) *axiomatise* la propriété $\mathcal{P}$.

Proposition 3.1.5. *Soit* $\mathcal{P}$ *une propriété que chaque* $\mathcal{L}$*-structure est susceptible de vérifier ou non. Alors* $\mathcal{P}$ *est finiment axiomatisable si et seulement si et* $\mathcal{P}$ *et la négation de* $\mathcal{P}$ *sont axiomatisables.*

Démonstration. Si φ axiomatise $\mathcal{P}$, alors $\neg\varphi$ axiomatise la négation de $\mathcal{P}$. Réciproquement, on suppose que T axiomatise $\mathcal{P}$ et que T' axiomatise la négation de $\mathcal{P}$. La théorie $T \cup T'$ est contradictoire. Il existe donc $\varphi_1, \ldots, \varphi_n \in T$, $\varphi'_1, \ldots, \varphi'_m \in T'$ tels que $\{\varphi_1, \ldots, \varphi_n, \varphi'_1, \ldots, \varphi'_m\}$ soit contradictoire. On a

$$\mathfrak{M} \models T \Rightarrow \mathfrak{M} \models \bigwedge_{i=1}^{n} \varphi_i \Rightarrow \mathfrak{M} \not\models \bigwedge_{i=1}^{m} \varphi'_i \Rightarrow \mathfrak{M} \not\models T' \Rightarrow \mathfrak{M} \models T.$$

Cela montre que $\varphi = \bigwedge_{i=1}^{n} \varphi_i$ est une axiomatisation finie de $\mathcal{P}$. □

Exemples 3.1.6. 1. Soit p un nombre premier. Alors les corps de caractéristique p sont finiment axiomatisables (dans le langage $\mathcal{L}_{an}$).

2. Les corps de caractéristique 0 sont axiomatisables dans $\mathcal{L}_{an}$, mais pas finiment. (En effet, si l'énoncé φ_0 était une axiomatisation finie, il serait conséquence de $\varphi_{corps} \cup \{\neg \underbrace{1+\ldots+1}_{p \text{ fois}} \doteq 0 \mid p \text{ premier}\}$, et on aurait alors $\varphi_{corps} \cup \{\neg \underbrace{1+\ldots+1}_{p \text{ fois}} \doteq 0 \mid p \text{ premier et } p < N\} \vdash \varphi_0$ pour un $N \in \mathbb{N}$. Contradiction, puisqu'il existe des corps de caractéristique $p > N$.)

3. Les corps ordonnés archimédiens ne sont pas axiomatisables (cf. 3.1.4(1)).

4. Les corps ordonnés complets ne sont pas axiomatisables (cf. 3.1.4(2)).

On observe que les deux derniers exemples précisent ce que nous avons dit au début du chapitre 2.

3.2 La méthode des diagrammes

Soit $\mathfrak{M}$ une $\mathcal{L}$-structure. On considère le langage $\mathcal{L}_M$ obtenu en rajoutant à $\mathcal{L}$ une constante nouvelle c_m pour chaque élément m de M. La structure $\mathfrak{M}$ admet une expansion naturelle au langage $\mathcal{L}_M$, en interprétant c_m par m. On note $\mathfrak{M}^*$ la $\mathcal{L}_M$-structure ainsi obtenue.

Par abus de notation, on identifiera parfois m et c_m dans la suite. Cela ne pose pas de vrai problème, à cause du lemme de substitution.

Définition. 1. Le *diagramme complet* de $\mathfrak{M}$, noté $D(\mathfrak{M})$, est défini comme $\mathrm{Th}(\mathfrak{M}^*)$. Il est donc formé des $\varphi(c_{m_1}, \ldots, c_{m_n})$, où $\varphi = \varphi(x_1, \ldots, x_n)$ est une $\mathcal{L}$-formule et $\overline{m} \in M^n$ tel que $\mathfrak{M} \models \varphi[m_1, \ldots, m_n]$.

2. Le *diagramme simple* de $\mathfrak{M}$, noté $\Delta(\mathfrak{M})$, est formé des $\varphi(c_{m_1}, \ldots, c_{m_n})$, où φ est une $\mathcal{L}$-formule sans quanteur et $\overline{m} \in M^n$ tel que $\mathfrak{M} \models \varphi[m_1, \ldots, m_n]$.

Proposition 3.2.1. *Les réduits au langage $\mathcal{L}$ des modèles de $D(\mathfrak{M})$ correspondent, à $\mathcal{L}$-isomorphisme près, aux extensions élémentaires de $\mathfrak{M}$.*

Démonstration. Soit $\mathfrak{N}^* \models D(\mathfrak{M})$ et $\mathfrak{N} := \mathfrak{N}^* \restriction_{\mathcal{L}}$ son $\mathcal{L}$-réduit. L'application $m \mapsto c_m^{\mathfrak{N}}$ définit un isomorphisme entre $\mathfrak{M}$ et une sous-structure élémentaire de $\mathfrak{N}$ (exercice facile).

Réciproquement, si $\mathfrak{M} \preccurlyeq \mathfrak{N}$, il suffit de considérer la $\mathcal{L}_M$-expansion $\mathfrak{N}^*$ de $\mathfrak{N}$ obtenue en interprétant c_m par m. On a alors $\mathfrak{N}^* \models D(\mathfrak{M})$. En effet, soit $\varphi(c_{m_1}, \ldots, c_{m_n}) \in D(\mathfrak{M})$. Alors $\mathfrak{M} \models \varphi[m_1, \ldots, m_n] \Rightarrow \mathfrak{N} \models \varphi[m_1, \ldots, m_n] \Rightarrow \mathfrak{N}^* \models \varphi(c_{m_1}, \ldots, c_{m_n})$, par élémentarité et le lemme de substitution. □

Proposition 3.2.2. *Les réduits au langage $\mathcal{L}$ des modèles de $\Delta(\mathfrak{M})$ correspondent, à $\mathcal{L}$-isomorphisme près, aux extensions de $\mathfrak{M}$.*

Démonstration. Similaire à la preuve précédente. □

Exercice 3.2.3.

1. Soient $\mathfrak{M}_1$ et $\mathfrak{M}_2$ deux $\mathcal{L}$-structures. Montrer que $\mathfrak{M}_1 \equiv \mathfrak{M}_2$ si et seulement s'il existe $\mathfrak{N}_1, \mathfrak{N}_2$ et $\mathfrak{P}$ telles que $\mathfrak{M}_i \cong \mathfrak{N}_i$ et $\mathfrak{N}_i \preccurlyeq \mathfrak{P}$ pour $i = 1, 2$.
   (Indication : On pourra montrer que $D(\mathfrak{M}_1) \cup D(\mathfrak{M}_2)$ est une théorie cohérente, où les deux ensembles de constantes sont choisis disjoints.)
2. Soit $T = \mathrm{Th}(\mathfrak{N})$, où $\mathfrak{N}$ est la $\mathcal{L}_{ar}$-structure $\langle \mathbb{N}; S, 0, +, \cdot, < \rangle$. Montrer qu'il existe $\mathfrak{N}' \models T$ tel que N' contienne un nombre premier non standard p' (par cela, on entend que $\mathfrak{N}' \models \forall x(\exists y\, x \cdot y \doteq p' \to (x \doteq 1 \vee x \doteq p'))$ et $\mathfrak{N}' \models (\underbrace{S \cdots S}_{n \text{ fois}} 0) < p'$ pour tout $n \in \mathbb{N}$).

Théorème 3.2.4 (Théorème de Löwenheim-Skolem ascendant)**.** *Soit $\mathfrak{M}$ une $\mathcal{L}$-structure infinie et κ un cardinal tel que $\kappa \geq \sup(\mathrm{card}(M), \mathrm{card}(\mathcal{L}))$. Alors il existe une extension élémentaire de $\mathfrak{M}$ qui est de cardinalité κ.*

Corollaire 3.2.5. *Soit T une théorie ayant un modèle infini. Alors T a un modèle en tout cardinal $\geq \mathrm{card}(\mathcal{L})$.* □

Démonstration du théorème 3.2.4. Il suffit de construire une extension élémentaire $\mathfrak{N}$ de $\mathfrak{M}$ qui est de cardinalité $\geq \kappa$, puis d'appliquer le théorème de Löwenheim-Skolem descendant à un ensemble $A \subseteq N$ de cardinalité κ et contenant M.

Pour chaque $i \in \kappa$, soit c_i une nouvelle constante. Soit $\tilde{\mathcal{L}} := \mathcal{L}_M \cup \{c_i \mid i \in \kappa\}$, et soit $\tilde{T} := D(\mathfrak{M}) \cup \{\neg c_i \doteq c_j \mid i < j < \kappa\}$. C'est une $\tilde{L}$-théorie qui est finiment satisfaisable. En effet, si $\tilde{T}_0$ est une partie finie de $\tilde{T}$, il suffit d'interpréter les symboles de constantes $c_{i_1}, \ldots, c_{i_n}$ qui apparaissent dans $\tilde{T}_0$ par des éléments distincts de M pour construire une expansion de $\mathfrak{M}^*$ qui est modèle de $\tilde{T}_0$. Comme M est infini, c'est possible. Par compacité, on trouve alors $\tilde{\mathfrak{N}} \models \tilde{T}$ dont le $\mathcal{L}$-réduit $\mathfrak{N}$ est une extension élémentaire de $\mathfrak{M}$ (par 3.2.1) qui est de cardinalité $\geq \kappa$, puisqu'il contient les κ éléments 2 à 2 distincts $c_i^{\tilde{\mathfrak{N}}}$, $i \in \kappa$. □

Exercice 3.2.6. Soit $\mathcal{L}$ un langage fini (c'est-à-dire dont la signature $\sigma^{\mathcal{L}}$ est finie), et soit $\mathfrak{M}$ une $\mathcal{L}$-structure finie.

1. Montrer qu'il existe un $\mathcal{L}_M$-énoncé $\varphi \in \Delta(\mathfrak{M})$ tel que $\vdash \varphi \to \varphi'$ pour tout $\varphi' \in \Delta(\mathfrak{M})$.
2. En déduire que si $\mathfrak{N} \equiv \mathfrak{M}$, alors $\mathfrak{N} \cong \mathfrak{M}$.
3. (Plus difficile.) Soit $\mathfrak{M}'$ une $\mathcal{L}'$-structure finie, avec $\mathcal{L}'$ non nécessairement fini. Montrer que $\mathfrak{N}' \equiv \mathfrak{M}' \Rightarrow \mathfrak{N} \cong \mathfrak{M}$.

3.3 Expansions par définitions

Soit T une $\mathcal{L}$-théorie et $\varphi = \varphi(x_1, \ldots, x_n)$ une $\mathcal{L}$-formule. On peut introduire un nouveau symbole de relation (n-aire) R et définir la $\mathcal{L}' = \mathcal{L} \cup \{R\}$-théorie $T \cup \{\forall x_1, \ldots, x_n(\varphi(x_1, \ldots, x_n) \leftrightarrow Rx_1 \cdots x_n)\}$. Alors T' est une expansion *conservatrice* de T, c'est-à-dire : pour tout $\mathcal{L}$-énoncé ψ on a $T \models \psi$ si et seulement si $T' \models \psi$. C'est clair, car tout modèle de T admet une $\mathcal{L}'$-expansion (même unique) en un modèle de T'.

On dit que deux formules $\varphi_1(x_1, \ldots, x_n)$ et $\varphi_2(x_1, \ldots, x_n)$ sont *équivalentes dans la théorie* $\tilde{T}$ si $\tilde{T} \models \forall x_1, \ldots, x_n(\varphi_1 \leftrightarrow \varphi_2)$.

Par induction, on montre que toute $\mathcal{L}'$-formule est équivalente, dans T', à une $\mathcal{L}$-formule.

Notation. $\exists! x\psi$ (« il existe un et un seul x tel que ψ ») est une abbréviation pour $\exists x(\psi \wedge \forall x'(\psi_{x'/x} \to x = x'))$, où x' est une variable distincte de x.

Si l'on rajoute un symbole de fonction (resp. symbole de constante) pour une fonction (resp. constante) définissable, l'expansion T' ainsi construite satisfait toujours les mêmes propriétés (Proposition 3.3.2).

Lemme 3.3.1. *Toute formule est logiquement équivalente à une formule dans laquelle ne figurent que des termes de hauteur* ≤ 1.

Démonstration. Pour φ une formule, soit $M(\varphi)$ la hauteur maximale d'un terme dans φ. On fait une preuve par induction sur $(M(\varphi), \mathrm{ht}(\varphi))$ ordonné lexicographiquement. Le seul cas non trivial est le cas d'une formule atomique φ, disons $\varphi = Rt_1 \cdots t_n$. Pour simplifier les notations, nous supposons que $\mathrm{ht}(t_i) > 1$ pour $i = 1, \ldots, m$ et $\mathrm{ht}(t_i) \leq 1$ pour $i > m$. Pour $i = 1, \ldots, m$, soit $t_i = f_i(s_{i,1}, \ldots, s_{i,k_i})$. Pour $i = 1, \ldots, m$ on choisit des nouvelles variables $y_{i,j}$ pour $1 \leq j \leq k_i$. Si $\overline{y}_i := (y_{i,1}, \ldots, y_{i,k_i})$, alors φ est équivalente à

$$\psi = \exists \overline{y}_1, \ldots, \overline{y}_m \left(Rf_1(\overline{y}_1) \cdots f_m(\overline{y}_m) t_{m+1} \cdots t_n \wedge \bigwedge_{i,j} y_{i,j} \doteq s_{i,j} \right)$$

et on a $M(\psi) < M(\varphi)$. Le cas $\varphi = t_1 \doteq t_2$ est similaire. □

Proposition 3.3.2 (Fonctions définissables)**.** *Soit* $\varphi(x_0, \ldots, x_n)$ *une* $\mathcal{L}$*-formule. On suppose que* $T \models \forall x_1, \ldots, x_n \exists! x_0\, \varphi$*, c'est-à-dire que dans tout modèle de* T*,* φ *définit le graphe d'une fonction* n*-aire.*

- *Si* $n > 0$*, soit* f *un nouveau symbole de fonction* n*-aire,* $\mathcal{L}' = \mathcal{L} \cup \{f\}$ *et soit* $T' = T \cup \{\forall x_1, \ldots, x_n\, \varphi(f(x_1, \ldots, x_n), x_1, \ldots, x_n)\}$.
- *Si* $n = 0$*, soit* c *un nouveau symbole de constante,* $\mathcal{L}' = \mathcal{L} \cup \{c\}$ *et soit* $T' = T \cup \{\varphi(c)\}$.

Alors T' *est une expansion conservatrice de* T*. De plus, toute* $\mathcal{L}'$*-formule* ψ *est équivalente dans* T' *à une* $\mathcal{L}$*-formule* ψ^*.

Démonstration. Nous donnons l'argument dans le cas $n > 0$, le cas $n = 0$ étant similaire.

Tout modèle $\mathfrak{M}$ de T admet une $\mathcal{L}'$-expansion (même unique) en un modèle $\mathfrak{M}'$ de T'. En effet, il faut et il suffit de définir $f^{\mathfrak{M}'}(a_1, \ldots, a_n) = a_0$ si $\mathfrak{M} \models \varphi[a_0, a_1, \ldots, a_n]$. Donc $T' \supseteq T$ est conservatrice.

Soit d'abord ψ une $\mathcal{L}'$-formule atomique ne contenant que des termes de hauteur ≤ 1, disons $\psi = Rt_1 \cdots t_m$. On suppose que pour $i = 1, \ldots, k$ on ait $t_i = f(s_1^i, \ldots, s_n^i)$ et que f n'apparaisse pas dans t_i pour $i > k$. Les s_j^i sont

des termes de hauteur 0, donc en particulier des $\mathcal{L}$-termes. Soient $z_1, \ldots, z_k$ des nouvelles variables. Alors la $\mathcal{L}$-formule

$$\psi^* := \exists z_1, \ldots, z_k \left(Rz_1 \cdots z_k t_{k+1} \cdots t_m \wedge \bigwedge_{i=1}^{m} \varphi(z_i, s_1^i, \ldots, s_n^i) \right)$$

est équivalente (modulo T') à ψ. Si $\psi = t_1 \doteq t_2$, l'argument est le même.

Soit maintenant ψ une $\mathcal{L}'$-formule arbitraire. Par le lemme, on peut supposer que tous les termes figurant dans ψ sont de hauteurs au plus 1. Pour toute sous-formule atomique φ de ψ, on choisit une $\mathcal{L}$-formule équivalente φ^*. (On vient de construire une telle formule.) On définit ψ^* comme la $\mathcal{L}$-formule obtenue en remplaçant toute sous-formule atomique φ de ψ par la formule φ^* correspondante.

C'est un fait général que si on remplace des sous-formules par des sous-formules équivalentes, on obtient une formule équivalente. (Exercice.) □

Une expansion par des relations, fonctions et constantes définissables est appelée une *expansion par définition*.

Exemples 3.3.3.
1. Soit $T' = \mathrm{Th}(\mathfrak{R})$ la théorie du corps ordonné des réels, et soit T la théorie du (pur) corps des réels. Comme dans $\mathbb{R}$ on a $r < s$ si et seulement s'il existe $t \neq 0$ tel que $t^2 = s - r$, T' est une expansion par définition de T.
2. Soit $T = \mathrm{Th}(\langle \aleph_1; < \rangle)$. Alors ω est une constante définissable dans T, c'est-à-dire il existe une formule $\varphi(x_0)$ telle que $T \models \exists! x_0 \varphi$ et $\langle \aleph_1; < \rangle \models \varphi[\omega]$. En effet, l'ensemble des ordinaux limites est défini par la formule $\mathrm{Lim}(x) = \exists y\, y < x \wedge \forall y \exists z (y < x \to y < z \wedge z < x)$, et il suffit de poser $\varphi(x_0) = \mathrm{Lim}(x_0) \wedge \forall y (\mathrm{Lim}(y) \to \neg y < x_0)$

3.4 Élimination des quanteurs

Par induction sur la hauteur, il est facile de montrer que toute formule φ est logiquement équivalente à une formule ψ qui est sous *forme prénexe*, c'est-à-dire de la forme $Q_1 x_1 \cdots Q_n x_n \chi$, avec χ sans quanteur et $Q_i \in \{\exists, \forall\}$ pour tout i. En général, le nombre d'alternances des quanteurs $\exists$ et $\forall$ fournit une bonne indication pour la complexité de la formule φ (ainsi que de l'ensemble défini par φ dans une structure donnée). Souvent, on comprend assez bien les ensembles qui sont définissables à l'aide d'une formule sans quanteur.

Nous commençons par un résultat technique important et dont la preuve est un peu longue.

Théorème 3.4.1. *Soit T une $\mathcal{L}$-théorie, $n \geq 1$ un entier et $\varphi = \varphi(x_1, \ldots, x_n)$ une $\mathcal{L}$-formule. Sont équivalents :*

1. *Il existe une $\mathcal{L}$-formule sans quanteur $\psi(x_1, \ldots, x_n)$ qui est équivalente à φ dans T.*

2. *Soient $\mathfrak{M}$ et $\mathfrak{N}$ deux modèles de T et $\mathfrak{A}$ une sous-structure commune de $\mathfrak{M}$ et de $\mathfrak{N}$. Alors pour tout $\overline{a} \in A^n$ on a $\mathfrak{M} \models \varphi[\overline{a}] \iff \mathfrak{N} \models \varphi[\overline{a}]$.*

Remarque 3.4.2. *Le théorème s'applique aussi dans le cas où φ est un énoncé. Il suffit de considérer φ comme $\varphi(x)$ pour trouver une formule $\psi(x)$ sans quanteurs et équivalente à φ dans T. Ainsi, un théorème de T, par exemple $\exists y y \doteq y$ est équivalent dans T à la formule $x \doteq x$. Si le langage $\mathcal{L}$ ne contient pas de symbole de constante, il n'y a pas de $\mathcal{L}$-énoncé sans quanteur et on ne peut donc pas faire mieux.*

Une autre manière de traiter ce problème serait d'introduire une constante propositionnelle $\top$ pour un énoncé toujours vrai.

Démonstration du théorème 3.4.1. (1)⇒(2). On note d'abord que si $\mathfrak{A} \subseteq \mathfrak{B}$, la formule $\psi(x_1, \ldots, x_n)$ est sans quanteur et $\overline{a} \in A^n$, alors $\mathfrak{A} \models \psi[\overline{a}] \iff \mathfrak{B} \models \psi[\overline{a}]$.

Si $\mathfrak{M}$ et $\mathfrak{N}$ sont des modèles de T ayant $\mathfrak{A}$ comme sous-structure et si $\varphi = \varphi(x_1, \ldots, x_n)$ est équivalente dans T à la formule $\psi(x_1, \ldots, x_n)$ sans quanteur, alors pour $\overline{a} \in A^n$ on a donc

$$\mathfrak{M} \models \varphi[\overline{a}] \iff \mathfrak{M} \models \psi[\overline{a}] \iff \mathfrak{A} \models \psi[\overline{a}] \iff \mathfrak{N} \models \psi[\overline{a}] \iff \mathfrak{N} \models \varphi[\overline{a}].$$

(2)⇒(1). On considère l'ensemble de formules

$$\Gamma(\overline{x}) := \{\chi(x_1, \ldots, x_n) \text{ sans quanteur } \mid T \models \forall x_1, \ldots, x_n(\varphi \to \chi)\}.$$

On choisit $c_1, \ldots, c_n$ des nouvelles constantes 2 à 2 distinctes et on considère la théorie $\Gamma(\overline{c}) := \{\chi(c_1, \ldots, c_n) \mid \chi \in \Gamma(\overline{x})\}$. Soit $\mathcal{L}' = \mathcal{L} \cup \{c_1, \ldots, c_n\}$. Nous allons montrer :

$$T \cup \Gamma(\overline{c}) \models \varphi(\overline{c}) \tag{3.1}$$

Si (3.1) était faux, on pourrait trouver $\mathfrak{M}' \models T \cup \Gamma(\overline{c}) \cup \neg\varphi(\overline{c})$. Soit $\mathfrak{A}' := \langle c_1^{\mathfrak{M}'}, \ldots, c_n^{\mathfrak{M}'} \rangle_{\mathfrak{M}'} = \langle A; \ldots \rangle$ la sous-structure engendrée par les $c_i^{\mathfrak{M}'}$ dans $\mathfrak{M}'$. On observe que $\Gamma(\overline{c}) \subseteq \Delta(\mathfrak{A}')$. Montrons que

$$\Sigma := T \cup \Delta(\mathfrak{A}') \cup \{\varphi(\overline{c})\}$$

a un modèle.

Sinon, on aurait $T \cup \Delta(\mathfrak{A}') \models \neg\varphi(\overline{c})$, d'où $T \cup \Delta(\mathfrak{A}') \vdash \neg\varphi(\overline{c})$. Comme tout élément de A s'écrit comme un $\mathcal{L}'$-terme, si on note $\Delta_{\overline{c}}(\mathfrak{A}')$ l'ensemble des $\mathcal{L}'$-énoncés sans quanteurs dans $\Delta(\mathfrak{A}')$, alors $T \cup \Delta(\mathfrak{A}')$ est une expansion conservatrice de $T \cup \Delta_{\overline{c}}(\mathfrak{A}')$ par la proposition 3.3.2. En particulier, on a

$$T \cup \Delta_{\overline{c}}(\mathfrak{A}') \vdash \neg\varphi(\overline{c}).$$

Il existe alors des $\mathcal{L}$-formules sans quanteur $\xi_1(\overline{x}), \ldots, \xi_k(\overline{x})$ telles que

$$T \vdash \bigwedge_{i=1}^{k} \xi_i(\overline{c}) \to \neg\varphi(\overline{c}) \quad \text{et} \quad \Delta(\mathfrak{A}') \vdash \bigwedge_{i=1}^{k} \xi_i(\overline{c}) =: \xi(\overline{c}).$$

Comme les c_i n'apparaissent ni dans T ni dans $\varphi(\overline{x}), \xi(\overline{x})$, on en déduit (par exemple par 2.6.3) que $T \vdash \forall \overline{x}(\xi(\overline{x}) \to \neg\varphi(\overline{x}))$ et alors $T \vdash \forall \overline{x}(\varphi(\overline{x}) \to \neg\xi(\overline{x}))$. Mais alors $\neg\xi(\overline{x}) \in \Gamma(\overline{x})$ par définition, et $\neg\xi(\overline{c}) \in \Gamma(\overline{c})$, d'où $\neg\xi(\overline{c}) \in \Delta(\mathfrak{A}')$. Contradiction.

Donc Σ a un modèle $\mathfrak{N}^*$, et le $\mathcal{L}$-réduit $\mathfrak{N}$ de $\mathfrak{N}^*$ contient une copie isomorphe $\mathfrak{B}'$ de $\mathfrak{A}'$ comme sous-structure par la proposition 3.2.2. Quitte à identifier $\mathfrak{B}'$ et $\mathfrak{A}'$, on a donc construit deux modèles $\mathfrak{M} = \mathfrak{M}' \restriction_{\mathcal{L}}$ et $\mathfrak{N}$ de T contenant une sous-structure commune $\mathfrak{A} = \mathfrak{A}' \restriction_{\mathcal{L}}$ tels que si l'on pose $a_i = c_i^{\mathfrak{M}'}$ alors $\mathfrak{N} \models \varphi[\overline{a}]$ et $\mathfrak{M} \models \neg\varphi[\overline{a}]$, ce qui contredit la condition (2). On a donc démontré (3.1).

Par compacité il existe $\zeta_1(\overline{c}), \ldots, \zeta_m(\overline{c}) \in \Gamma(\overline{c})$ tels que $T \models \bigwedge_{i=1}^m \zeta_i(\overline{c}) \to \varphi(\overline{c})$, ce qui entraîne comme avant $T \models \forall \overline{x}\,(\bigwedge_{i=1}^m \zeta_i(\overline{x}) \to \varphi(\overline{x}))$. Mais alors, comme $T \models \forall \overline{x}(\varphi \to \zeta_i)$ pour tout i, on obtient $T \models \forall \overline{x}\,(\bigwedge_{i=1}^m \zeta_i(\overline{x}) \leftrightarrow \varphi(\overline{x}))$, avec $\bigwedge_{i=1}^m \zeta_i(\overline{x})$ sans quanteur. □

Définition. Soit T une $\mathcal{L}$-théorie. On dit que T admet *l'élimination des quanteurs* (dans le langage $\mathcal{L}$) si toute $\mathcal{L}$-formule φ est équivalente dans T à une $\mathcal{L}$-formule sans quanteur.

Lemme 3.4.3. *Si pour toute formule sans quanteur φ et toute variable x il existe une formule sans quanteur ψ telle que $\exists x\varphi$ et ψ soient équivalentes dans T, alors T admet l'élimination des quanteurs.*

Démonstration. Soient ψ et ψ' deux formules qui sont équivalentes dans T, noté $\psi \sim_T \psi'$. Alors $\neg\psi \sim_T \neg\psi'$, $\exists x\psi \sim_T \exists x\psi'$ et $\chi \wedge \psi \sim_T \chi \wedge \psi'$ pour toute formule χ. On peut donc raisonner par induction sur la hauteur de la formule, et la preuve est claire. □

Théorème 3.4.4. *Soit T une $\mathcal{L}$-théorie. On suppose que pour toute paire de modèles $\mathfrak{M}$ et $\mathfrak{N}$ de T, pour toute sous-structure commune $\mathfrak{A}$ de $\mathfrak{M}$ et de $\mathfrak{N}$ et toute formule sans quanteur $\varphi(x_0, \ldots, x_n)$, s'il existe $\overline{a} \in A^n$ et $b_0 \in M$ tels que $\mathfrak{M} \models \varphi[b_0, \overline{a}]$, alors il existe $c_0 \in N$ tel que $\mathfrak{N} \models \varphi[c_0, \overline{a}]$.*

Alors T admet l'élimination des quanteurs.

Remarque. *La réciproque de ce résultat est claire : toute théorie avec l'élimination des quanteurs satisfait à l'hypothèse du théorème.*

Démonstration. Soient $\mathfrak{A} \subseteq \mathfrak{M}, \mathfrak{N}$ données, avec $\mathfrak{M}, \mathfrak{N} \models T$. Quitte à symétriser l'hypothèse du théorème, elle exprime que si φ est sans quanteur et $\chi = \exists x_0\varphi$, alors $\mathfrak{M} \models \chi[\overline{a}] \iff \mathfrak{N} \models \chi[\overline{a}]$ pour tout $\overline{a} \in A^n$. Par le théorème 3.4.1, χ est équivalente dans T à une formule sans quanteur. Par le lemme, cela suffit. □

Proposition 3.4.5. *Soit T une théorie qui admet l'élimination des quanteurs.*

1. *Soient $\mathfrak{M}$ et $\mathfrak{N}$ deux modèles de T ayant une sous-structure commune. Alors $\mathfrak{M} \equiv \mathfrak{N}$.*
2. *Si $\mathfrak{M} \subseteq \mathfrak{N}$, où $\mathfrak{M}$ et $\mathfrak{N}$ sont deux modèles de T, alors $\mathfrak{M} \preccurlyeq \mathfrak{N}$.*

Démonstration. (1) Il s'agit d'un cas particulier de la direction facile du théorème 3.4.1. En effet, tout énoncé φ est équivalent dans T à une formule $\psi(x)$ sans quanteur. Pour $a \in A$ arbitraire, où $\mathfrak{A}$ est une sous-structure commune de $\mathfrak{M}$ et de $\mathfrak{N}$, on a donc

$$\mathfrak{M} \models \varphi \iff \mathfrak{M} \models \psi[a] \iff \mathfrak{A} \models \psi[a] \iff \mathfrak{N} \models \psi[a] \iff \mathfrak{N} \models \varphi.$$

(2) Clair. □

Exercice 3.4.6. Toute théorie T admet une expansion par définition qui admet l'élimination des quanteurs.

3.5 Corps algébriquement clos

On considère la $\mathcal{L}_{an}$-théorie CAC des corps algébriquement clos (voir 2.6.6). On utilisera des notions et résultats qui seront vus en Algèbre 1.

Soit A un sous-anneau d'un corps K. Un élément de K est appelé *algébrique sur* A s'il est racine d'un polynôme non nul à coefficients dans A.

Soit A un anneau intègre. Une *clôture algébrique de* A est un corps algébriquement clos contenant A et dont tout élément est algébrique sur A.

Fait 3.5.1. *Soit A un anneau intègre.*

1. *A admet une clôture algébrique.*
2. *Si K et K' sont deux clôtures algébriques de A, alors il existe un isomorphisme $f : K \cong K'$ tel que $f \restriction_A = \mathrm{id}_A$.*
3. *Soit $A \subseteq L$, où L est un corps algébriquement clos. Alors le sous-corps $A_L^{alg} = \{b \in L \mid b \text{ est algébrique sur } A\}$ est une clôture algébrique de A.*
4. *Soit $\mathbb{F}_p^{alg}$ une clôture algébrique du corps à p éléments $\mathbb{F}_p$. Alors $\mathbb{F}_p^{alg}$ est une réunion croissante de sous-corps finis F_N, $N \in \mathbb{N}$. Plus précisément, pour tout entier $k \geq 1$, l'ensemble des racines du polynôme $X^{p^k} - X$ est un sous-corps $\mathbb{F}_{p^k}$ (à p^k éléments), et $\bigcup_{k\in\mathbb{N}} \mathbb{F}_{p^k} = \mathbb{F}_p^{alg}$. Comme de plus $\mathbb{F}_{p^k} \subseteq \mathbb{F}_{p^l}$ si $k \mid l$ (car alors $X^{p^l} - X$ est divisible par $X^{p^k} - X$), il suffit de poser $F_N := \mathbb{F}_{p^{N!}}$.*
5. *Tout corps algébriquement clos est infini.*
6. *Soit $K \subseteq L$ une extension de corps avec K algébriquement clos, et soit $b \in L \setminus K$. Alors b n'est pas algébrique sur K.*

Théorème 3.5.2. *La théorie* CAC *admet l'élimination des quanteurs.*

Démonstration. On observe qu'une sous-structure d'un corps dans $\mathcal{L}_{an}$ n'est rien d'autre qu'un sous-anneau. Par le théorème 3.4.4 il suffit donc de montrer que si K et L sont algébriquement clos, A est un sous-anneau commun et $\varphi = \varphi(x_0, \ldots, x_n)$ une formule sans quanteur, alors pour tout $\overline{a} \in A^n$ il existe $b \in L$ tel que $L \models \varphi[b, \overline{a}]$ si et seulement s'il existe $c \in K$ tel que $K \models \varphi[c, \overline{a}]$.

Par le fait, K et L contiennent des clôtures algébriques F_K et F_L de A qui sont isomorphes au-dessus de A. Quitte à agrandir A, on peut donc supposer que A est algébriquement clos. On peut alors même supposer que $A = K \subseteq L$. Dans cette situation, il faut montrer que s'il existe $b \in L$ tel que $L \models \varphi[b, \overline{a}]$, alors il existe $c \in K$ tel que $K \models \varphi[c, \overline{a}]$.

La formule φ est logiquement équivalente à une formule de la forme $\bigvee_i \bigwedge \chi_{i,j}$, avec $\chi_{i,j} = \chi_{i,j}(x_0, \ldots, x_n)$ atomique ou négation d'atomique. Si $L \models \varphi[b, \overline{a}]$, il existe i tel que $L \models \bigwedge_j \chi_{i,j}[b, \overline{a}]$. Il suffit donc de traiter le cas où φ est une conjonction de formules atomiques ou négations d'atomiques. Dans la théorie des corps, toute formule atomique est équivalente à $P(\overline{x}) \doteq 0$ pour un polynôme $P(\overline{x})$ à coefficients entiers. Alors on peut supposer que

$$\varphi(\overline{x}) = \bigwedge_{i=1}^{n} P_i(\overline{x}) \doteq 0 \wedge \bigwedge_{i=1}^{m} \neg Q_i(\overline{x}) \doteq 0.$$

Si l'un des $P_i(x_0, a_1, \ldots, a_n) \in K[x_0]$ est un polynôme non nul, alors b est algébrique sur K, d'où $b \in K$ et on a terminé.

On peut donc supposer que $\varphi = \bigwedge_{i=1}^{m} \neg Q_i(\overline{x}) \doteq 0$. Comme chaque $Q_i(x_0, \overline{a}) \in K[x_0]$ est différent du polynôme nul (sinon, b ne pourrait pas exister), il n'a qu'un nombre fini de racines. On en déduit qu'il existe $c \in K$ tel que $K \models \varphi[c, \overline{a}]$, car K est un corps infini (il est algébriquement clos). □

Pour p un nombre premier ou $p = 0$ on note CAC_p la théorie des corps algébriquement clos de caractéristique p.

Théorème 3.5.3. *Les théories* CAC_p *sont complètes (p premier ou* $p = 0$*).*

Démonstration. Tout corps de caractéristique $p > 0$ contient $\mathbb{F}_p$ comme sous-corps. Si K et L sont algébriquement clos de caractéristique p, alors $K \equiv L$ par le théorème 3.5.2 et la proposition 3.4.5(1), ce qui montre que CAC_p est complète.

Pour CAC_0, on fait le même argument, en remplaçant $\mathbb{F}_p$ par $\mathbb{Q}$. □

Théorème 3.5.4 (Principe de Lefschetz)**.** *Soit* φ *un* $\mathcal{L}_{an}$*-énoncé. Les propriétés suivantes sont équivalentes :*

1. $\mathbb{C} \models \varphi$.
2. *Il existe un corps algébriquement clos de caractéristique 0 qui vérifie* φ.
3. *Tout corps algébriquement clos de caractéristique 0 vérifie* φ.
4. *Il existe* $N \in \mathbb{N}$ *tel que* φ *est vérifé dans tout corps algébriquement clos de caractéristique* $p > N$.
5. *Il existe un ensemble infini de nombres premiers* $\mathcal{P}$ *tel que pour tout* $p \in \mathcal{P}$ *il existe* $K_p \models \mathrm{CAC}_p$ *vérifiant* φ.

Démonstration. (1) $\iff$ (2) $\iff$ (3) suit de 3.5.3.

(3)$\Rightarrow$(4). Notons que CAC_0 est donnée par $\mathrm{CAC} \cup \{\chi_p \mid p \text{ premier}\}$, où χ_p exprime que $p = 1 + \cdots + 1$ est différent de 0. Si $\mathrm{CAC} \models \varphi$, par compacité il

existe une partie finie Δ de CAC_0 telle que $\Delta \models \varphi$. Or Δ ne contient qu'un nombre fini d'axiomes de la formes χ_p. En particulier il existe $N \in \mathbb{N}$ tel que $K \models \Delta$ pour tout corps algébriquement clos de caractéristique $p > N$.

(4)⇒(5) est trivial.

(5)⇒(3). Pour $p \in \mathcal{P}$, soit $K_p \models \mathrm{CAC}_p$ tel que $K_p \models \varphi$. Si $\mathrm{CAC}_0 \not\models \varphi$, alors $\mathrm{CAC}_0 \models \neg\varphi$ par complétude. Par (3)⇒(4) il existe $N \in \mathbb{N}$ tel que $\neg\varphi$ soit vérifié dans tout corps algébriquement clos de caractéristique $p > N$. En particulier $\mathcal{P}$ est fini. □

Théorème 3.5.5 (Théorème des zéros (Nullstellensatz) de Hilbert)**.**
Soient K un corps algébriquement clos et $P_1(\overline{X}), \ldots, P_m(\overline{X}) \in K[X_1, \ldots, X_n]$. Si le système d'équations polynômiales $P_1(\overline{X}) = P_2(\overline{X}) = \ldots = P_m(\overline{X}) = 0$ a une solution dans un corps $L \supseteq K$, alors il a une solution dans K.

Démonstration. Soient $L \supseteq K$ et $\overline{a} \in L^n$ tel que $P_1(\overline{a}) = \ldots = P_m(\overline{a}) = 0$. Quitte à agrandir L, on peut supposer que L est algébriquement clos. Comme CAC admet l'élimination des quanteurs, on a $K \preccurlyeq L$ par 3.4.5.

On choisit des $\mathcal{L}_{an}$-termes $F_i(\overline{X}, \overline{Z_i})$ et des uplets $\overline{b}_i$ extraits de K tels que $P_i(\overline{X}) = F_i(\overline{X}, \overline{b}_i)$.

Alors $L \models \exists\overline{x} \bigwedge F_i(\overline{x}, \overline{b}_i) \doteq 0$, d'où $K \models \exists\overline{x} \bigwedge F_i(\overline{x}, \overline{b}_i) \doteq 0$, comme $K \preccurlyeq L$. □

3.6 Le théorème d'Ax

Une chaîne de $\mathcal{L}$-structures est une suite $(\mathfrak{M}_i)_{i\in\mathbb{N}}$ de $\mathcal{L}$-structures telle que $\mathfrak{M}_i \subseteq \mathfrak{M}_{i+1}$ pour tout i.

Si $(\mathfrak{M}_i)_{i\in\mathbb{N}}$ est une telle chaîne, il existe une et une seule $\mathcal{L}$-structure $\mathfrak{M}$ avec ensemble de base $M = \bigcup_{i\in\mathbb{N}} M_i$ telle que $\mathfrak{M}_i \subseteq \mathfrak{M}$ pour tout i. En effet, il faut et il suffit d'interpréter les symboles du langage comme suit : $c^{\mathfrak{M}} = c^{\mathfrak{M}_0}$, $f^{\mathfrak{M}} = \bigcup_{i\in\mathbb{N}} f^{\mathfrak{M}_i}$ et $R^{\mathfrak{M}} = \bigcup_{i\in\mathbb{N}} R^{\mathfrak{M}_i}$. Il est facile à voir que c'est bien défini.

La $\mathcal{L}$-structure $\mathfrak{M}$ ainsi obtenue est notée $\bigcup_{i\in\mathbb{N}} \mathfrak{M}_i$.

Définition. Une *formule* $\forall\exists$ est une formule de la forme $\forall x_1, \ldots, x_n \exists y_1, \ldots, y_m \varphi$, où φ est sans quanteur et $m, n \geq 0$.

Lemme 3.6.1 (Préservation des énoncés $\forall\exists$ par union de chaîne)**.**
Soit ψ un $\mathcal{L}$-énoncé $\forall\exists$ et $(\mathfrak{M}_i)_{i\in\mathbb{N}}$ une chaîne de $\mathcal{L}$-structures avec $\mathfrak{M}_i \models \psi$ pour tout i. Alors $\mathfrak{M} = \bigcup_{i\in\mathbb{N}} \mathfrak{M}_i \models \psi$.

Démonstration. Soit $\psi = \forall x_1, \ldots, x_n \exists y_1, \ldots, y_m \varphi(\overline{x}, \overline{y})$ avec φ sans quanteur. Il faut montrer que $\mathfrak{M} \models \exists y_1, \ldots, y_m \varphi[\overline{a}, \overline{y}]$ pour tout $\overline{a} \in M^n$. Comme M_i est une suite croissante, il existe $k \in \mathbb{N}$ tel que $\overline{a} \in M_k^n$. Il existe donc $b_1, \ldots, b_m \in M_k$ tels que $\mathfrak{M}_k \models \varphi[\overline{a}, \overline{b}]$, car $\mathfrak{M} \models \psi$. On en déduit que $\mathfrak{M} \models \varphi[\overline{a}, \overline{b}]$, car φ est sans quanteur et $\mathfrak{M}_k$ est une sous-structure de $\mathfrak{M}$. □

Remarque. *On peut montrer qu'un énoncé est préservé par union de chaîne si et seulement s'il est logiquement équivalent à un énoncé $\forall\exists$.*

Proposition 3.6.2. *Soit φ un $\mathcal{L}_{an}$-énoncé $\forall\exists$ vrai dans tout corps fini. Alors* $\mathrm{CAC} \models \varphi$. *En particulier il est vrai dans $\mathbb{C}$.*

Démonstration. Par le fait 3.5.1(4), $\mathbb{F}_p^{alg}$ est réunion croissante de sous-corps finis. Donc $\mathbb{F}_p^{alg} \models \varphi$ par le lemme, pour tout p premier. On en déduit le résultat par le principe de Lefschetz. □

Théorème 3.6.3 (Théorème d'Ax)**.** *Soit $f : \mathbb{C}^n \to \mathbb{C}^n$ une application polynômiale, c'est-à-dire de la forme $f = (f_1, \ldots, f_n)$ pour des polynômes $f_i \in \mathbb{C}[x_1, \ldots, x_n]$. Si f est injective, alors f est surjective.*

Démonstration. Il existe des $\mathcal{L}_{an}$-termes — ce sont en particulier des polynômes à coefficients entiers — $P_{n,d}(\overline{z}, \overline{x})$ tels que pour tout corps K, tout polynôme $f(\overline{x}) \in K[x_1, \ldots, x_n]$ de degré $\leq d$ s'écrit comme $P_{n,d}(\overline{a}, \overline{x})$ pour un uplet $\overline{a}$ d'éléments de K. L'énoncé suivant $\psi_{n,d}$ est $\forall\exists$ et il exprime que toute fonction polynômiale et injective $f : K^n \to K^n$, définie à l'aide de polynômes f_i de degré au plus d, est surjective :

$$\psi_{n,d} = \forall \overline{z}_1, \ldots, \overline{z}_n \forall \overline{u} \exists \overline{x} \exists \overline{x}' \exists \overline{y}$$
$$[(\textstyle\bigwedge_{i=1}^n P_{n,d}(\overline{z}_i, \overline{x}) \dot{=} P_{n,d}(\overline{z}_i, \overline{x}') \wedge \neg \bigwedge_{i=1}^n x_i \dot{=} x_i') \vee (\bigwedge_{i=1}^n P_{n,d}(\overline{z}_i, \overline{y}) \dot{=} u_i)]$$

Évidemment, $\psi_{n,d}$ est vrai dans tout corps fini, donc $\mathrm{CAC} \models \psi_{d,n}$ par la proposition 3.6.2, et en particulier $\mathbb{C} \models \psi_{d,n}$. □

Chapitre 4

Récursivité

Dans ce chapitre, nous développons la théorie des fonctions et ensembles d'entiers qui sont « calculables ». Nous donnons deux précisions de la notion intuitive d'une fonction calculable — la notion de fonction récursive, et la notion de fonction calculable par une machine de Turing — et nous montrons que ces deux notions coïncident.

Pour $n \in \mathbb{N}$, on note $\mathcal{F}_n := \{f : \mathbb{N}^n \to \mathbb{N}\}$ et $\mathcal{F} := \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$.

Notation. Soit f la fonction qui à $(x_1, \ldots, x_n)$ associe $f(x_1, \ldots, x_n) \in \mathbb{N}$. On notera parfois $f = \lambda x_1 \cdots x_n.f(x_1, \ldots, x_n)$.

4.1 Fonctions primitives récursives

Définition. L'ensemble des fonctions *primitives récursives* est le plus petit sous-ensemble E de $\mathcal{F}$ vérifiant les conditions suivantes :

(R0) E contient les *fonctions de base* suivantes :
- $S = \lambda x.x + 1$ (la fonction successeur) ;
- la fonction 0-aire constante à 0, notée C_0^0 ;
- pour tout $n \geq 1$ et tout $1 \leq i \leq n$ la projection P_i^n sur la i$^{\text{e}}$coordonnée, c'est-à-dire $P_i^n = \lambda x_1 \cdots x_n.x_i$.

(R1) E est clos par *composition* : si $f_1, \ldots, f_n \in \mathcal{F}_p \cap E$ et si $h \in \mathcal{F}_n \cap E$, alors $g = h(f_1, \ldots, f_n) \in E$.

(R2) E est clos par *récurrence* : pour tout entier n, si $g \in \mathcal{F}_n \cap E$ et $h \in \mathcal{F}_{n+2} \cap E$, alors la fonction f définie par

$$\begin{aligned} f(x_1, \ldots, x_n, 0) &:= g(x_1, \ldots, x_n), \\ f(x_1, \ldots, x_n, y+1) &:= h(x_1, \ldots, x_n, y, f(x_1, \ldots, x_n, y)) \end{aligned}$$

est dans E aussi.

Observons que toutes les fonctions constantes $C_k^n = \lambda x_1 \cdots x_n.k$ sont primitives récursives pour tout $k, n \in \mathbb{N}$. (La fonction C_0^1 est obtenue par récurrence à partir de $g = C_0^0$ et $h = P_2^2$.)

Si $X \subseteq \mathbb{N}^n$, on notera $\mathbb{1}_X$ la *fonction caractéristique* de X, c'est-à-dire $\mathbb{1}_X(x_1,\ldots,x_n) = \begin{cases} 1, \text{ si } (x_1,\ldots,x_n) \in X; \\ 0, \text{ sinon.} \end{cases}$

Lemme 4.1.1. *Les fonctions $\lambda xy.x+y$, $\lambda xy.x\cdot y$, $\lambda xy.x^y$, $\lambda x.x!$, $\lambda xy.x\dot{-}y$ et $\mathrm{sg} = \mathbb{1}_{\mathbb{N}^*}$ sont primitives récursives, où $x\dot{-}y := \begin{cases} x-y, \text{ si } x \geq y\,; \\ 0, \text{ sinon.} \end{cases}$.*

Démonstration. L'addition $f = \lambda xy.x+y$ est obtenue par récurrence à partir de $g = P^1_1$ et $h = S \circ P^3_3$, comme $x+0 = x$ et $x+(y+1) = S(x+y)$.

Il est également facile d'obtenir les autres fonctions. Donnons juste l'argument pour $\lambda xy.x\dot{-}y$ et sg. On construit d'abord $\lambda y.y\dot{-}1$ par récurrence, via $0\dot{-}1 = 0$, $(y+1)\dot{-}1 = y$. Puis, on peut définir $x\dot{-}0 = x$, $x\dot{-}(y+1) = (x\dot{-}y)\dot{-}1$. Enfin, on a $\mathrm{sg}(x) = 1\dot{-}(1\dot{-}x)$. □

Définition. Une partie $X \subseteq \mathbb{N}^n$ est *primitive récursive* si sa fonction caractéristique $\mathbb{1}_X$ l'est.

Lemme 4.1.2.

1. *L'ensemble des fonctions primitives récursives est clos par permutation des variables.*
   [Clair]
2. *Si $X \subseteq \mathbb{N}^n$ est primitif récursif et si $f_1,\ldots,f_n \in \mathcal{F}_p$ sont primitives récursives, alors $Y = \{(x_1,\ldots,x_p) \in \mathbb{N}^p \mid (f_1(\overline{x}),\ldots,f_n(\overline{x})) \in X\}$ est primitif récursif.*
   [On a $\mathbb{1}_Y = \mathbb{1}_X(f_1,\ldots,f_n)$.]
3. *L'ensemble des parties primitives récursives de $\mathbb{N}^n$ contient $\emptyset$ et $\mathbb{N}^n$, et il est stable par $\cup$, $\cap$ et par passage au complément (donc par combinaisons booléennes).*
   [On a $\mathbb{1}_{\mathbb{N}^n\setminus X} = 1\dot{-}\mathbb{1}_X$ et $\mathbb{1}_{X\cap Y} = \mathbb{1}_X \cdot \mathbb{1}_Y$.]
4. *L'ensemble $\{(x,y) \mid x < y\} \subseteq \mathbb{N}^2$ est primitif récursif.*
   [On a $\mathbb{1}_<(x,y) = \mathrm{sg}(y\dot{-}x)$.]
5. *(Définition par cas.) Soit $\mathbb{N}^n = A_1 \dot\cup \ldots \dot\cup A_k$ une partition de $\mathbb{N}^n$ en un nombre fini de parties primitives récursives A_i, et soient $f_1,\ldots,f_k \in \mathcal{F}_n$ primitives récursives. Alors la fonction f, définie par $f(\overline{x}) = f_i(\overline{x})$ si $\overline{x} \in A_i$, est primitive récursive aussi. En particulier, les fonctions $\lambda x_1\cdots x_n.\max(x_i)$ et $\lambda x_1\cdots x_n.\min(x_i)$ sont primitives récursives.*
   [En effet, on a $f = \mathbb{1}_{A_1}f_1 + \ldots + \mathbb{1}_{A_k}f_k$.]
6. *(Sommes et produits limités.) Si $f \in \mathcal{F}_{n+1}$ est primitive récursive, alors les fonctions suivantes aussi :*
$$\lambda x_1\cdots x_n y.\sum_{t=0}^{y} f(\overline{x},t) \;\text{ et }\; \lambda x_1\cdots x_n y.\prod_{t=0}^{y} f(\overline{x},t)$$
   [Par récurrence.]

7. *(Schéma μ borné.) Soit $X \subseteq \mathbb{N}^{n+1}$ primitif récursif. Alors la fonction $f \in \mathcal{F}_{n+1}$, $f(\overline{x}, z) = (\mu t \leq z)\,((\overline{x}, t) \in X)$, définie par*

$$f(\overline{x}, z) := \begin{cases} 0, \text{ s'il n'existe pas de } t \leq z \text{ tel que } (\overline{x}, t) \in X; \\ t_0, \text{ si } t_0 \text{ est le plus petit entier } t \leq z \text{ tel que } (\overline{x}, t) \in X \end{cases}$$

est primitive récursive.

[On a $f(\overline{x}, 0) = 0$, $f(\overline{x}, z+1) = \begin{cases} f(\overline{x}, z), \text{ si } \sum_{t=0}^{z} \mathbb{1}_X(\overline{x}, t) \geq 1\,; \\ z+1, \text{ sinon et si } (\overline{x}, z+1) \in X; \\ 0, \text{ sinon.} \end{cases}$

Donc f est primitive récursive (récurrence et définition par cas).]

8. *(Quantification bornée.) Si $X \subseteq \mathbb{N}^{n+1}$ est primitif récursif, alors aussi*

$$X_e = \{(x_1, \ldots, x_n, z) \in \mathbb{N}^{n+1} \mid (\exists t \leq z)\,(\overline{x}, t) \in X\} \text{ ainsi que}$$
$$X_a = \{(x_1, \ldots, x_n, z) \in \mathbb{N}^{n+1} \mid (\forall t \leq z)\,(\overline{x}, t) \in X\}.$$

[Comme on peut passer au complémentaire, il suffit de traiter le premier cas. On a $\mathbb{1}_{X_e}(\overline{x}, z) = 1$ si $\sum_{t=0}^{z} \mathbb{1}_X(\overline{x}, t) \geq 1$, et $\mathbb{1}_{X_e}(\overline{x}, z) = 0$ sinon.]

Exemples 4.1.3.

1. Soit $q : \mathbb{N}^2 \to \mathbb{N}$ la fonction qui à (x, y) associe la partie entière de x/y si $y \neq 0$, et 0 sinon. Alors q est primitive récursive.
   [On a $q(x, y) = (\mu t \leq x)\,((t+1) \cdot y > x)$.]
2. $\{(x, y) \in \mathbb{N}^2 : y \mid x\}$ est un ensemble primitif récurisf.
   [On a $y \mid x \iff \exists z\, z \cdot y = x \iff x = q(x, y) \cdot y$.]
3. L'ensemble P des nombres premiers est primitif récursif.
   [On a $x \in P \iff x \geq 2 \wedge (\forall y \leq x)\ (y \mid x \to (y = 1 \vee y = x))$. Par les propriétés de clôture du lemme précédent, P est donc primitif récursif.]
4. Soit π la fonction qui à x associe le $(x+1)^{\text{e}}$ nombre premier. Alors π est primitive récursive.
   [On a $\pi(0) = 2$, $\pi(x+1) = (\mu z \leq \pi(x)! + 1)\,(z > \pi(n) \text{ et } z \in P)$.]
5. Une bijection primitive récursive entre $\mathbb{N}^2$ et $\mathbb{N}$ est donnée par la fonction $\alpha_2 = \lambda xy.\frac{1}{2}(x+y+1)(x+y)+x$. On a $\alpha_2(\beta_1^2, \beta_2^2) = \mathrm{id}_{\mathbb{N}}$ pour des fonctions primitives récursives $\beta_1^2, \beta_2^2 \in \mathcal{F}_1$. En effet, comme $\alpha_2(x, y) \geq \min(x, y)$, on a $\beta_1^2(x) = (\mu z \leq x)(\exists t \leq x(\alpha_2(z, t) = x))$, de même pour β_2^2.
   Par induction sur $p \geq 2$, on définit une bijection primitive récursive $\alpha_p : \mathbb{N}^p \to \mathbb{N}$ avec composantes de l'inverse $\beta_1^p, \ldots, \beta_p^p$ primitives récursives. Il suffit de poser $\alpha_{p+1}(x_1, \ldots, x_{p+1}) = \alpha_p(x_1, \ldots, x_{p-1}, \alpha_2(x_p, x_{p+1}))$.

Lemme 4.1.4. $\langle x_0, \ldots, x_{n-1} \rangle := \pi(0)^{x_0} \cdot \ldots \cdot \pi(n-2)^{x_{n-2}} \cdot \pi(n-1)^{x_{n-1}+1} - 1$ *définit une bijection $\langle\ \rangle : \mathbb{N}^* \to \mathbb{N}$, où $\mathbb{N}^*$ dénote l'ensemble des suites finies d'entiers. Elle satisfait aux propriétés suivantes :*

1. *La fonction* composante *à deux places $(x)_i$, définie par*

$$(\langle x_0, \ldots, x_{n-1} \rangle)_i = \begin{cases} x_i, \text{ si } i < n\,; \\ 0, \text{ sinon} \end{cases}$$

est primitve récursive.

2. *La fonction* longueur, $\lg(\langle x_0, \ldots, x_{n-1}\rangle) = n$, *est primitive récursive.*
3. *Pour tout* $n \in \mathbb{N}$, $\langle\ \rangle \restriction_{\mathbb{N}^n} : \mathbb{N}^n \to \mathbb{N}$ *est primitive récursive.*
4. *On a* $\lg(x) \leq x$ *pour tout* x, *et si* $x > 0$, *alors* $(x)_i < x$ *pour tout* i.

On appelle $\langle x_0, \ldots, x_{n-1}\rangle$ *le* nombre de Gödel *de la suite d'entiers* $(x_0, \ldots, x_{n-1})$.

Démonstration. (4) est clair, et (3) suit de 4.1.3.

(2) On a $\lg(x) = \begin{cases} 0, \text{ si } x = 0\,; \\ (\mu y \leq x)\,[(\forall z \leq x)\,(y \leq z \to \pi(z) \dot{-} (x+1))]\,, \text{ sinon.} \end{cases}$

(1) On a $(x)_i = \begin{cases} 0, \text{ si } i \geq \lg(x)\,; \\ (\mu y \leq x)\,\big(\pi(i)^{y+2} \dot{-} (x+1)\big)\,, \text{ si } i+1 = \lg(x)\,; \\ (\mu y \leq x)\,\big(\pi(i)^{y+1} \dot{-} (x+1)\big)\,, \text{ si } i+1 < \lg(x). \end{cases}$ □

4.2 La fonction d'Ackermann

Voici une fonction qui est calculable au sens intuitif, la *fonction d'Ackermann* $\xi \in \mathcal{F}_2$ définie ainsi :

- $\xi(0, x) := 2^x$
- $\xi(y, 0) := 1$
- $\xi(y+1, x+1) := \xi(y, \xi(y+1, x))$

Cette fonction est bien définie et se calcule en un nombre fini d'étapes (exercice ; il suffit de considérer l'ordre lexicographique sur $\mathbb{N}^2$).

On écrit aussi $\xi_n(x)$ pour $\xi(n, x)$, et ξ_n^k pour $\overbrace{\xi_n \circ \ldots \circ \xi_n}^{k \text{ fois}}$.

Lemme 4.2.1. 1. $\xi_n(x) > x$ *pour tout* $n, x \in \mathbb{N}$.

2. ξ_n *est strictement croissante pour tout* $n \in \mathbb{N}$.
3. $\xi_{n+1}(x) \geq \xi_n(x)$ *pour tout* $n, x \in \mathbb{N}$.
4. ξ_n^k *est strictement croissante pour tout* $k, n \in \mathbb{N}$.
5. $\xi_n^k(x) < \xi_n^{k+1}(x)$ *pour tout* $k, n, x \in \mathbb{N}$.
6. $\xi_m^k(x) \leq \xi_n^k(x)$ *pour tout* $k, m, n, x \in \mathbb{N}$ *avec* $m \leq n$.
7. $\xi_n^k(x) \leq \xi_{n+1}(x+k)$ *pour tout* $k, n, x \in \mathbb{N}$.

Démonstration. (1) On fait une induction sur n, le cas $n = 0$ étant clair. On a $\xi_{n+1}(x+1) = \xi_n(\xi_{n+1}(x)) > \xi_{n+1}(x) > x$.

(2) C'est clair pour $n = 0$, et on vient de voir que $\xi_{n+1}(x+1) > \xi_{n+1}(x)$.

(3) $\xi_{n+1}(0) = \xi_n(0)$, $\xi_{n+1}(x+1) = \xi_n(\underbrace{\xi_{n+1}(x)}_{\geq x+1}) \geq \xi_n(x+1)$ par (2).

(4), (5) et (6) sont clairs.

(7) On fait une récurrence sur k, le cas $k = 0$ étant clair. On a $\xi_n^{k+1}(x) = \xi_n(\xi_n^k(x)) \leq \xi_n(\xi_{n+1}(x+k)) = \xi_{n+1}(x+k+1)$. □

Définition. On dit que la fonction $f \in \mathcal{F}_1$ *domine* la fonction $g \in \mathcal{F}_n$ s'il existe un entier N tel que pour tout $\overline{x} \in \mathbb{N}^n$ on ait $g(\overline{x}) \leq f(\max(x_i, N))$.

Notons que si f est strictement croissante, alors f domine g si et seulement si $g(\overline{x}) \leq f(\max(x_i))$ sauf pour un ensemble fini de n-uplets.

On note C_n l'ensemble des fonctions dans $\mathcal{F}$ qui sont dominées par au moins une des fonctions ξ_n^k, pour $k \in \mathbb{N}$.

Lemme 4.2.2. *1. $C_n \subseteq C_m$ pour tout $n \leq m$.*

2. C_0 contient les fonctions de base ainsi que les fonctions $\lambda xy.x+y$, $\lambda x.k \cdot x$ (pour k fixé) et $\lambda x_1 \ldots x_n.\max(x_i)$.

3. C_n est stable par composition.

4. Si $g \in \mathcal{F}_p \cap C_n$ et $h \in \mathcal{F}_{p+2} \cap C_n$, alors f définie par récurrence à partir de g et h est dans C_{n+1}.

En conclusion, $C = \bigcup_{n\in\mathbb{N}} C_n$ contient l'ensemble des fonctions primitives récursives.

Démonstration. (1) et (2) sont clairs. Quant à (3), on se donne $f_1, \ldots, f_m \in \mathcal{F}_p \cap C_n$ et $g \in \mathcal{F}_m \cap C_n$. Il existe des entiers $N, N_1, \ldots, N_m, k, k_1, \ldots, k_m$ tels que $g(\overline{y}) \leq \xi_n^k(\max(y_j, N))$ pour tout $\overline{y} \in \mathbb{N}^m$ et $f_i(\overline{x}) \leq \xi_n^{k_i}(\max(x_j, N_i))$ pour tout $\overline{x} \in \mathbb{N}^p$. Posons $M = \max(N, N_1, \ldots, N_m)$ et $l = \max(k_1, \ldots, k_m)$. Pour tout $\overline{x} \in \mathbb{N}^p$ on a $g(f_1(\overline{x}), \ldots, f_m(\overline{x})) \leq \xi_n^k(\xi_n^l(\max(x_j, M))) = \xi_n^{k+l}(\max(x_j, M))$.

(4) On a $f(\overline{x}, 0) = g(\overline{x})$ et $f(\overline{x}, y+1) = h(\overline{x}, y, f(\overline{x}, y))$. Il existe k_1, N_1, k_2, N_2 tels que $g(\overline{x}) \leq \xi_n^{k_1}(\max(x_j, N_1))$ et $h(\overline{x}, y, t) \leq \xi_n^{k_2}(\max(x_j, y, t, N_2))$. Par récurrence sur y, on montre sans problème que

$$f(\overline{x}, y) \leq \xi_n^{k_1 + yk_2}(\max(x_j, y, N_1, N_2)). \tag{4.1}$$

Par le lemme 4.2.1(7), on déduit de (4.1) que

$$f(\overline{x}, y) \leq \underbrace{\xi_{n+1}(\max(x_j, y, N_1, N_2) + k_1 + k_2 y)}_{\text{composée de fonctions dans } C_{n+1}},$$

d'où le résultat par la partie (3). □

Théorème 4.2.3. *La fonction d'Ackermann n'est pas primitive récursive.*

Remarque. *Par contre, une induction sur n montre que toutes les fonctions ξ_n sont primitives récursives.*

Démonstration du théorème 4.2.3. Par l'absurde. On suppose donc que ξ est primitive récursive. Alors $\lambda x.\xi(x, 2x)$ l'est aussi, et il existe $k, n, N \in \mathbb{N}$ tels que $\xi(x, 2x) \leq \xi_n^k(x)$ pour tout $x > N$. Par 4.2.1(7) on a donc $\xi_x(2x) \leq \xi_{n+1}(x + k)$ pour tout $x > N$. Pour $x > \max(k, n+1)$, c'est absurde. □

4.3 Fonctions partielles récursives

Compte tenu de l'existence de la fonction d'Ackermann — calculable au sens intuitif et non primitive récursive — nous allons élargir la classe des fonctions que nous considérons. Pour des raisons techniques, il convient également de passer aux fonctions partielles.

Une *fonction partielle* de $\mathbb{N}^n$ dans $\mathbb{N}$ est la donnée d'un couple (A, f) avec $A \subseteq \mathbb{N}^n$ et $f : A \to \mathbb{N}$. L'ensemble A est le domaine de définition de f, noté $\mathrm{dom}(f)$. On note $\mathcal{F}_n^*$ l'ensemble de ces couples et $\mathcal{F}^* = \bigcup_{n\in\mathbb{N}} \mathcal{F}_n^*$.

Définition. L'ensemble des *fonctions partielles récursives* est le plus petit sous-ensemble E de $\mathcal{F}^*$ vérifiant les propriétés $(R0)^* - (R3)^*$ suivantes :

$(R0)^*$ E contient les fonctions de base (C_0^0, S et P_i^n pour $n \geq 1$ et $1 \leq i \leq n$).

$(R1)^*$ E est clos par composition (de fonctions partielles) : si $f_1, \ldots, f_n \in \mathcal{F}_m^* \cap E$ et $h \in \mathcal{F}_n^* \cap E$, alors $g = h(f_1, \ldots, f_n) \in E$, où g est la fonction partielle qui à $\overline{x}$ associe $h(f_1(\overline{x}), \ldots, f_n(\overline{x}))$ si $\overline{x} \in \mathrm{dom}(f_i)$ pour tout i et $(f_1(\overline{x}), \ldots, f_n(\overline{x})) \in \mathrm{dom}(h)$. Sinon, g n'est pas définie en $\overline{x}$.

$(R2)^*$ E est clos par récurrence (de fonctions partielles) : si $g \in \mathcal{F}_p^* \cap E$ et $h \in \mathcal{F}_{p+2}^* \cap E$, alors $f \in E$, où
- $f(\overline{x}, 0) = g(\overline{x})$ si $\overline{x} \in \mathrm{dom}(g)$, et sinon f n'est pas définie en $(\overline{x}, 0)$;
- $f(\overline{x}, y+1) = h(\overline{x}, y, f(\overline{x}, y))$ si f est définie en $(\overline{x}, y)$ (relation définie en même temps par récurrence sur y) et si $(\overline{x}, y, f(\overline{x}, y)) \in \mathrm{dom}(h)$, et sinon f n'est pas définie en $(\overline{x}, y+1)$.

$(R3)^*$ E est clos par *schéma* μ : soit $f \in \mathcal{F}_{n+1}^* \cap E$. Alors $g \in E$, où $g \in \mathcal{F}_n^*$, $g(\overline{x}) = \mu y\,(f(\overline{x}, y) = 0)$ est définie ainsi :
- s'il existe z tel que $f(\overline{x}, z) = 0$ et $(\overline{x}, z') \in \mathrm{dom}(f)$ pour tout $z' \leq z$, alors $g(\overline{x})$ est le plus petit tel z ;
- sinon, g n'est pas définie en $\overline{x}$.

Une fonction partielle $f \in \mathcal{F}_n^*$ est *totale* si $\mathrm{dom}(f) = \mathbb{N}^n$. On appelle *fonction récursive totale* toute fonction totale qui est récursive partielle.

L'opération suivante est réservée pour les fonctions totales.

Soit f une fonction totale à $n+1$ arguments telle que $\forall x_1, \ldots, x_n \exists y\, f(\overline{x}, y) = 0$. Alors on dit que la fonction (totale) g, $g(\overline{x}) = \mu y\,(f(\overline{x}, y) = 0)$ est obtenue par *schéma* μ *total* à partir de f.

La propriété $(R3)$ exprime qu'un ensemble de fonctions totales est clos par schéma μ total.

Exercice 4.3.1. Montrer que la fonction d'Ackermann est récursive totale. (Voir Proposition 4.5.4.)

4.4 Fonctions calculables par machine de Turing

Les machines de Turing fournissent une précision de la notion de « calculabilité par une machine ».

Définition. Une *machine de Turing* $\mathcal{M}$ est la donnée
- d'un nombre fini (≥ 1) de *bandes* $B_1, B_2, \ldots$ placées horizontalement, toutes bornées à gauche et non bornées à droite ; chaque bande est divisée en cases, numérotées par $\mathbb{N}^*$ de gauche à droite. Les bandes sont arrangées de sorte que les cases du même numéro se trouvent sur une même verticale.
- d'une *tête de lecture* qui peut lire, effacer et écrire des symboles sur les cases (un symbole par case). Cette tête se déplace horizontalement et est placée à chaque instant sur une verticale et manipule toutes les cases de cette verticale d'un coup.
  L'ensemble de symboles est donné par $S = \{d, |, b\}$, d comme « début de bande », | comme « bâton » et b comme « blanc ».

Voici les données qui sont spécifiques à $\mathcal{M}$:
- $n = n(\mathcal{M}) \in \mathbb{N}^*$ (le nombre de bandes)
- Un ensemble fini d'*états* E. Il y a deux états spécifiques appartenant à E : e_i (l'état initial) et e_f (l'état final).
- Une fonction $M : S^n \times E \to S^n \times E \times \{-1, 0, 1\}$, appelée *table de transition* de $\mathcal{M}$.

Fonctionnement de la machine :
- À chaque instant $t \in \mathbb{N}$, $\mathcal{M}$ se trouve dans un état donné (de E).
- $\mathcal{M}$ fonctionne en changeant d'état, en effaçant et écrivant des symboles sur les bandes et en déplaçant sa tête à chaque instant.

Le fonctionnement est sujet aux règles suivantes :
1. À l'instant $t = 0$, $\mathcal{M}$ se trouve dans l'état initial et sa tête se trouve devant les cases n^{o} 1.
2. À chaque instant t, $\mathcal{M}$ lit les symboles $(s_1, \ldots, s_n) \in S^n$ écrits devant sa tête, et la table M décrit alors ce que $\mathcal{M}$ est censée faire. Si $\mathcal{M}$ est dans l'état e et lit $(s_1, \ldots, s_n)$, alors, soit $M(\overline{s}, e) = (\overline{s}', e', \varepsilon)$. La tête efface $\overline{s}$, écrit $\overline{s}'$, se déplace par ε horizontalement et $\mathcal{M}$ se met dans l'état e', puis passe à l'instant $t + 1$.
3. $\mathcal{M}$ s'arrête quand elle atteint l'état final.

Entrée correcte :
- À l'instant $t = 0$, d est inscrit dans les cases n° 1 et uniquement là.
- Toutes les cases sont remplies, et il y a un nombre fini de bâtons.

(Cela reste vrai tout au long du calcul, vu les contraintes qui suivront).

Contraintes :
- Pour tout $\overline{s} \in S^n$ on a $M(\overline{s}, e_f) = (\overline{s}, e_f, 0)$.
- La tête ne peut ni écrire ni effacer le symbole de début de bande :
  - Pour tout $e \in E$, on a $M((d, \ldots, d), e) = ((d, \ldots, d), e', \varepsilon)$ pour un $\varepsilon \in \{0, 1\}$;
  - si $\overline{s} \neq (d, \ldots, d)$, alors $M(\overline{s}, e) = (\overline{s}', e', \varepsilon)$ avec $s'_i \neq d$ pour tout i.

Remarque. *Il s'agit d'une machine* déterministe*, c'est-à-dire on peut prévoir pour tout instant t la position de la tête, l'état de la machine et le remplissage des bandes.*

Définition. 1. Une bande *représente* (à un instant donné) l'entier m si la bande remplie est égale à $(d, \underbrace{|, \ldots, |}_{m \text{ fois}}, b, \ldots, b, \ldots)$.

2. Une machine de Turing $\mathcal{M}$ *calcule* $f \in \mathcal{F}_p^*$ si $n(\mathcal{M}) \geq p+1$ et si pour tout $\overline{m} \in \mathbb{N}^p$, quand on fait fonctionner $\mathcal{M}$ sur l'entrée où, pour $i = 1, \ldots, p$, la bande B_i représente m_i et, pour $i > p$, la bande B_i représente l'entier 0, alors
   - si $\overline{m} \in \mathrm{dom}(f)$, $\mathcal{M}$ s'arrête après un temps fini et sur les bandes sont représentés successivement les entiers $(m_1, \ldots, m_p, f(\overline{m}), 0, \ldots, 0)$;
   - si $\overline{m} \notin \mathrm{dom}(f)$, $\mathcal{M}$ ne s'arrête jamais (c'est-à-dire e_f n'est jamais atteint).
3. La fonction partielle f est dite *T-calculable* s'il existe une machine de Turing $\mathcal{M}$ qui calcule f.

Lemme 4.4.1. *Les fonctions de base (C_0^0, S et les P_i^n) sont T-calculables.*

Démonstration. C_0^0 est calculable par $\mathcal{M}$ à une bande et avec $E = \{e_i, e_f\}$, où la table de transition est donnée par $M(d, e_i) = (d, e_f, 0)$. (Ici comme dans la suite, nous nous contentons de donner la partie essentielle de M.)

S est calculable par $\mathcal{M}$ à deux bandes, avec $E = \{e_i, e_f\}$ et où $M(d, d, e_i) = (d, d, e_i, +1), M(|, b, e_i) = (|, |, e_i, +1)$ et $M(b, b, e_i) = (b, |, e_f, 0)$.

P_i^n est calculable par $\mathcal{M}$ à $n+1$ bandes, avec $E = \{e_i, e_f\}$ et où la table de transition est donnée par $M(d, \ldots, d, e_i) = (d, \ldots, d, e_i, +1)$ ainsi que

$$M(s_1, \ldots, s_n, b, e_i) = \begin{cases} (s_1, \ldots, s_n, |, e_i, +1), \text{ si } s_i = |, \\ (s_1, \ldots, s_n, b, e_f, 0), \text{ si } s_i = b. \end{cases} \quad \square$$

Lemme 4.4.2. *L'ensemble des fonctions partielles T-calculables est clos par composition.*

Démonstration. Soient $f_1, \ldots, f_n \in \mathcal{F}_p^*$ et $g \in \mathcal{F}_n^*$ T-calculables, et soit $h = g(f_1, \ldots, f_n)$.

Par hypothèse il existe $\mathcal{M}_i$, $1 \leq i \leq n$, à $p_i \geq p+1$ bandes et ensemble d'états E_i calculant f_i, ainsi que $\mathcal{M}'$ à $n' \geq n+1$ bandes et ensemble d'états E' calculant g.

Soit $\mathcal{M}$ une machine de Turing à $p + (n' - n) + \sum_{i=1}^n (p_i - p)$ bandes et ensemble d'états $E = \{e_d, e_n\} \dot\cup E' \dot\cup \dot\bigcup_i E_i$. On déclare que l'état initial de $\mathcal{M}$ est celui de $\mathcal{M}_1$, l'état final celui de $\mathcal{M}'$.

Voici le fonctionnement de $\mathcal{M}$ (la description exacte de la table de transition de $\mathcal{M}$ est laissée en exercice) :

Étant représentés $(m_1, \ldots, m_p, 0, \ldots, 0)$ sur les bandes, $\mathcal{M}$ commence à calculer $f_1(m_1, \ldots, m_p)$ comme $\mathcal{M}_1$ le ferait, en se servant des états dans $E_1 \setminus \{e_f^1\}$ et de $(p_1 - p)$ bandes supplémentaires (disjointes de B_{p+1}). En l'état e_f^1, elle revient en début des bandes, puis se met dans l'état $e_i^2 \in E_2$ et calcule $f_2(\overline{m})$ en se servant des états dans $E_2 \setminus \{e_f^2\}$ et de $(p_2 - p)$ bandes supplémentaires (disjointes

de B_{p+1}). Ainsi, $\mathcal{M}$ calcule successivement $f_1(\overline{m}), \ldots, f_n(\overline{m})$. Une fois $f_n(\overline{m})$ calculée, $\mathcal{M}$ se met dans l'état initial de $\mathcal{M}'$ et calcule $h(f_1(\overline{m}), \ldots, f_n(\overline{m}))$ comme $\mathcal{M}'$ le ferait, en se servant de E' et de $n' - n$ bandes supplémentaires (disjointes de B_{p+1}). Elle utilise pour ce calcul les bandes sur lesquelles sont inscrites les $f_i(\overline{m})$ comme bandes d'entrée et B_{p+1} comme bande de sortie. Il faudra donc renuméroter les bandes pour pouvoir se servir des tables de transition de $\mathcal{M}_1, \ldots, \mathcal{M}_n$ et $\mathcal{M}'$ en établissant la table de $\mathcal{M}$.

Une fois $h(f_1(\overline{m}), \ldots, f_n(\overline{m}))$ calculée, au lieu de passer à l'état final de $\mathcal{M}'$, $\mathcal{M}$ se met dans l'état e_d qui sert pour retrouver le début de bande, puis nettoie les bandes sur lesquelles sont représentés les $f_i(\overline{m})$, en se servant de l'état e_n, pour enfin se mettre dans l'état final. □

Lemme 4.4.3. *L'ensemble des fonctions partielles T-calculables est clos par schéma μ.*

Démonstration. Exercice. (C'est facile.) □

Lemme 4.4.4. *L'ensemble des fonctions partielles T-calculables est clos par récurrence.*

Démonstration. Soit $g \in \mathcal{F}_p^*$ calculée par $\mathcal{M}$ à $p + 1 + k$ bandes et ensemble d'états E, et soit $h \in \mathcal{F}_{p+2}^*$ calculée par $\mathcal{M}'$ à $p + 3 + k'$ bandes et ensemble d'états E'. On considère la fonction partielle $f \in \mathcal{F}_{p+1}^*$ définie par $f(\overline{x}, 0) = g(\overline{x})$ et $f(\overline{x}, y + 1) = h(\overline{x}, y, f(\overline{x}, y))$.

Voici la description informelle d'une machine de Turing $\mathcal{N}$ qui calcule f : $\mathcal{N}$ a $p + 4 + k + k'$ bandes, et son ensemble d'états est composé de $E \dot{\cup} E'$ plus certains états auxiliaires, avec $e_i^{\mathcal{N}} = e_i^{\mathcal{M}}$ et $e_f^{\mathcal{N}} = e_f^{\mathcal{M}'}$.

Fonctionnement de $\mathcal{N}$ sur l'entrée $(m_1, \ldots, m_{p+1})$:

(1) Calcul de $g(m_1, \ldots, m_p)$ avec bandes d'entrée $B_1, \ldots, B_p$, bande de sortie B_{p+2} et bandes auxiliaires $B_{p+5}, \ldots, , B_{p+4+k}$, en travaillant comme $\mathcal{M}$, à renumérotation des bandes près.

(2) Comparer m_{p+1} et le contenu y de B_{p+3} (tout au long du calcul, la bande B_{p+3} représentera un entier $\leq m_{p+1}$) :
   – si $m_{p+1} = y$, aller à l'étape (5) ;
   – si $m_{p+1} > y$, aller à l'étape (3).

(3) Calculer $h(m_1, \ldots, m_p, y, f(\overline{x}, y)) = f(\overline{x}, y + 1)$ comme le ferait $\mathcal{M}'$, avec bandes d'entrée $B_1, \ldots, B_p, B_{p+3}, B_{p+2}$, bande de sortie B_{p+4} et les k' dernières bandes comme bandes auxiliaires.

(4) Copier le contenue de B_{p+4} sur la bande B_{p+2}, puis nettoyer B_{p+4} et incrémenter le contenu de B_{p+3} par un bâton, c'est-à-dire passer de y à $y + 1$. Revenir à l'étape (2).

(5) Nettoyer B_{p+3} et s'arrêter. □

Théorème 4.4.5. *Les fonctions partielles récursives sont T-calculables.*

Démonstration. On combine les quatre lemmes précédents. □

Exercice. Décrire une machine de Turing qui calcule $\lambda xy.x + y$.

Pour pouvoir établir la réciproque du théorème 4.4.5, nous allons coder les machines de Turing ainsi que leur fonctionnement.

Codage des machines de Turing :
On identifie $S = \{b, d, |\}$ à $\{0, 1, 2\}$, via $0 \leftrightarrow b$, $1 \leftrightarrow d$ et $2 \leftrightarrow |$. Une suite $(s_i)_{i \leq n}$, ou une suite $(s_i)_{i \in \mathbb{N}}$ avec $s_i = 0$ pour presque tout i, est codée par $\Gamma((s_i)) = \sum_{i \geq 0} s_i 3^i \in \mathbb{N}$.

Soit $\mathcal{M}$ une machine de Turing. Elle est donnée par
- $n = n(\mathcal{M}) \geq 1$, le nombre de bandes de $\mathcal{M}$;
- l'ensemble fini d'états E : on supposera que $E = \{0, \ldots, m\}$ ($m \geq 1$ est donc égal à $\mathrm{card}(E) - 1$), avec $e_i = 0$ et $e_f = 1$;
- la table de transition $M : S^n \times E \to S^n \times E \times \{-1, 0, 1\}$: pour coder M, si $\rho = (s_1, \ldots, s_n, e) \in S^n \times E$ et $M(\rho) = (t_1, \ldots, t_n, e', \varepsilon)$, on pose $r_1(\rho) = \alpha_2(\Gamma(s_1, \ldots, s_n), e)$ et $r_2(\rho) = \alpha_3(\Gamma(t_1, \ldots, t_n), e', \varepsilon + 1)$, puis

$$\ulcorner M \urcorner = \prod_{\rho \in S^n \times E} \pi(r_1(\rho))^{r_2(\rho)}.$$

Pour décoder, on utilisera la fonction δ, avec $\delta(i, x) := \mu z \leq x \left(\pi(i)^{z+1} \nmid x\right)$. On a alors $\delta(\alpha_2(\Gamma(s_1, \ldots, s_n), e), \ulcorner M \urcorner) = \alpha_3(\Gamma(t_1, \ldots, t_n), e', \varepsilon + 1)$.

Enfin, l'*indice de* $\mathcal{M}$ est donné par $\ulcorner \mathcal{M} \urcorner = \alpha_3(n, m, \ulcorner M \urcorner)$.

Lemme 4.4.6. *Pour $p \geq 0$, l'ensemble*

$$I_p = \{\ulcorner \mathcal{M} \urcorner \mid \mathcal{M} \textit{ est une machine de Turing à } \geq p + 1 \textit{ bandes}\}$$

est primitif récursif.

Démonstration. On voit facilement que l'on peut reconnaître de manière primitive récursive si les contraintes sur la table de transition sont satisfaites. Les détails sont laissés en exercice. □

Une *configuration* $C = C(t)$ de $\mathcal{M}$ (à un instant t) est un élément $(s_i)_i \in S^{\mathbb{N}}$ tel que s_{nv+w} soit l'élément écrit sur la $(v+1)^{\mathrm{e}}$ case de la bande B_{w+1}). (On suppose qu'il n'y a qu'un nombre fini de symboles non blancs et que le symbole d est écrit au début de bande et uniquement là.)

On note $\Gamma(C) := \Gamma((s_i))$ le code de la configuration C. Pour décoder, on utilisera la fonction $\eta(\Gamma(C), u, v, n) = r\left(q(\Gamma(C), 3^{n(u \dot{-} 1) + (v \dot{-} 1)}), 3\right)$, ($q$ étant le quotient de la division avec reste, r étant le reste) qui donne le symbole s écrit sur la u^{e} case de la bande B_v. Si $\sigma = (s_1, \ldots, s_n)$ est la suite des symboles écrits sur les cases n$^{\mathrm{o}}$ u, alors $\Gamma(\sigma) = \varepsilon(\Gamma(C), u, n)$, où $\varepsilon(x, y, z) = r\left(q(x, 3^{z(y \dot{-} 1)}), 3^z\right)$.

La *situation de* $\mathcal{M}$ (à un instant t) est donnée par $\mathrm{Sit} = \mathrm{Sit}(t) = (e, k, C(t))$, où e désigne l'état de $\mathcal{M}$ à l'instant t, k est le n$^{\mathrm{o}}$ de case devant laquelle la tête se trouve à l'instant t et $C(t)$ la configuration à l'instant t. On la code via $\Gamma(\mathrm{Sit}(t)) = \alpha_3(e, k, \Gamma(C(t))$.

Lemme 4.4.7. *Soit $p \geq 0$. Il existe une fonction primitive récursive $g^p \in \mathcal{F}_2$ telle que*
- *$g^p(i,x) = 0$ si $i \notin I_p$;*
- *si $i = \ulcorner \mathcal{M} \urcorner$ et x est le code de la situation de $\mathcal{M}$ à l'instant t, alors $g^p(i,x)$ est le code de la situation de $\mathcal{M}$ à l'instant $t+1$.*

Démonstration. On suppose $i = \ulcorner \mathcal{M} \urcorner \in I_p$. Alors
- l'état de la machine est $\beta_1^3(x) = e$;
- le n° des cases observées est $\beta_2^3(x) = k$;
- le code de la configuration $C(t)$ est $\beta_3^3(x) = \Gamma(C(t))$; le nombre de bandes est donné par $\beta_1^3(i) = n$;
- le nombre d'états de $\mathcal{M}$ moins 1 est égal à $\beta_2^3(i) = m$;
- le code de la table de transition est donné par $\beta_3^3(i) = \ulcorner M \urcorner$;
- le code de ce que lit la tête à l'instant t est donné par

$$\varepsilon\left(\Gamma(C(t)), k, n\right) = \varepsilon\left(\beta_3^3(x), \beta_2^3(x), \beta_1^3(i)\right) =: c.$$

Si x n'est pas le code d'une situation, on pose $g^p(i,x) = 0$. C'est le cas si $\beta_1^3(x) > m$, si $\beta_2^3(x) = 0$ ou si $\beta_3^3(x)$ n'est pas le code d'une configuration avec d en début de bande et uniquement là (la dernière condition s'exprime de manière primitive récursive, en utilisant les fonctions ε et η).

Sinon, soit $\delta := \delta\left(\alpha_2(c,e), \ulcorner M \urcorner\right)$. Alors $c' = \beta_1^3(\delta)$ est le code de la suite écrite à la place de celle codée par c. On pourra poser $g^p(i,x) = \alpha_3\left(e', k', \Gamma(C')\right)$, où $\Gamma(C') = \left(\Gamma(C) + 3^{n(k \dot{-} 1)} \cdot c'\right) \dot{-} \left(3^{n(k \dot{-} 1)} \cdot c\right)$ est le code de la configuration, $e' = \beta_2^3(\delta)$ l'état de la machine, et enfin $k' = \left(\beta_2^3(x) + \beta_3^3(\delta)\right) \dot{-} 1$ le n° des cases observées à l'instant $t+1$. □

On définit une fonction $\mathrm{ST}^p \in \mathcal{F}_{p+2}$:
- $\mathrm{ST}^p(i,t,\overline{x}) = 0$, si $i \notin I_p$;
- sinon, $\mathrm{ST}^p(i,t,\overline{x})$ est le code $\Gamma(\mathrm{Sit}(t))$ de la situation à l'instant t de la machine $\mathcal{M}$ d'indice i qui a commencé à fonctionner en $t = 0$ sur la configuration suivante : sur $B_1, \ldots, B_p$ sont représentés les entiers $x_1, \ldots, x_p$, et les autres bandes représentent 0.

Théorème 4.4.8. *Pour tout $p \geq 0$, la fonction ST^p est primitive récursive.*

Démonstration. La fonction $\lambda i \overline{x}.\mathrm{ST}^p(i, 0, \overline{x})$ est primitive récursive. (C'est facile et laissé en exercice.) Puis, on a $\mathrm{ST}^p(i, t+1, \overline{x}) = g^p\left(i, \mathrm{ST}^p(i,t,\overline{x})\right)$. □

On appelle *configuration de sortie pour l'entrée* $\overline{x} = (x_1, \ldots, x_p)$ une configuration où les bandes $B_1, \ldots, B_p$ représentent $x_1, \ldots, x_p$, la bande B_{p+1} représente un entier, et où les autres bandes repésentent 0.

Pour $p \geq 0$, on définit un prédicat $E^p \subseteq \mathbb{N}^{p+2}$ de la manière suivante : $(i, c, \overline{x}) \in E^p) \iff i \in I_p$ et c est le code d'une configuration de sortie pour l'entrée $\overline{x}$.

Il est facile à voir que E^p est un ensemble primitif récursif. [Exercice. Pour exprimer par exemple que B_{p+1} représente un entier, notons que dans la formule $\forall z \left(\eta(c, z, p+1, n) = 2 \wedge z \geq 3 \rightarrow \eta(c, z \dot{-} 1, p+1, n) = 2\right)$, on peut borner le quanteur universel par c.]

On définit aussi les ensembles suivants :

$$B^p = \left\{(i, t, \overline{x}) \in \mathbb{N}^{p+2} \mid \beta_1^3\left(\mathrm{ST}^p(i, t, \overline{x})\right) = 1 \text{ et } (i, \beta_3^3(\mathrm{ST}^p(i, t, \overline{x})), \overline{x}) \in E^p\right\}$$

(à l'instant t, la machine d'indice i se trouve dans l'état final, et sa configuration est une configuration de sortie pour l'entrée $\overline{x}$)

$$C^p = \left\{(i, y, t, \overline{x}) \in \mathbb{N}^{p+3} \mid (i, t, \overline{x}) \in B^p \text{ et } y \text{ est représenté sur } B_{p+1}\right\}$$

Les ensembles B^p et C^p sont primitifs récursifs.

Soit maintenant $f \in \mathcal{F}_p^*$ une fonction partielle T-calculable. Choisissons une machine de Turing $\mathcal{M}$ qui calcule f, et posons $i = \ulcorner \mathcal{M} \urcorner$. On peut définir la fonction partielle $T_{\mathcal{M}}$ (qui donne le temps de calcul) comme

$$T_{\mathcal{M}}(\overline{x}) := \mu t \left((i, t, \overline{x}) \in B^p\right).$$

Alors on a

$$f(\overline{x}) = \big(\mu y \leq T_{\mathcal{M}}(\overline{x})\big)\big((i, y, T_{\mathcal{M}}(\overline{x}), \overline{x}) \in C^p\big). \tag{4.2}$$

On appelle (4.2) la *forme normale de Kleene* de f.

En particulier, nous avons montré les deux résultats suivants.

Théorème 4.4.9. *Les fonctions (partielles) T-calculables sont récursives.* □

Proposition 4.4.10.
1. *Si f est totale et T-calculable en un temps primitif récursif, alors f est primitive récursive.*
2. *L'ensemble des fonctions partielles récursives est le plus petit sous-ensemble de $\mathcal{F}^*$ contenant les fonctions primitives récursives et clos par composition et schéma μ.*
3. *L'ensemble des fonctions totales récursives est le plus petit sous-ensemble de $\mathcal{F}$ contenant les fonctions primitives récursives et clos par composition et schéma μ total. Autrement dit, toute fonction récursive totale s'obtient par un nombre fini d'applications des règles $(R0) - (R3)$.* □

On observera que nos arguments pour montrer que toute fonction T-calculable est récursive ne sont pas spécifiques aux machines de Turing. Cela justifie la

Thèse de Church.
Toute fonction calculable (au sens intuitif du terme) est récursive.

4.5 Fonctions universelles

Il suffit de varier l'indice de la machine de Turing pour obtenir une fonction universelle. On définit d'abord la fonction $T^p \in \mathcal{F}^*_{p+1}$ comme suit :

– si $i \notin I_p$, $T^p(i, \overline{x})$ n'est pas définie ;

– si $i \in I_p$, on pose $T^p(i, \overline{x}) = \mu t\,((i, t, \overline{x}) \in B^p)$.

Ensuite, on définit la fonction partielle récursive $\varphi^p \in \mathcal{F}^*_{p+1}$,

$$\varphi^p(i, \overline{x}) = \mu y\,[(i, y, T^p(i, \overline{x}), \overline{x}) \in C^p]\,.$$

On note $\varphi^p_i = \lambda \overline{x}.\varphi^p(i, \overline{x})$.

Proposition 4.5.1. *φ^p est une fonction partielle récursive* universelle *: toute fonction partielle récursive en p variables est de la forme φ^p_i pour un entier i convenable.* □

Théorème 4.5.2 (Théorème SMN)**.** *Pour tout couple d'entiers m et n il existe une fonction $s^m_n \in \mathcal{F}_{n+1}$ primitive récursive telle que pour tout $i \in \mathbb{N}, \overline{x} \in \mathbb{N}^n$ et $\overline{y} \in \mathbb{N}^m$ on ait*

$$\varphi^{n+m}_i(\overline{x}, \overline{y}) = \varphi^m_{s^m_n(i,\overline{x})}(\overline{y}).$$

Démonstration. Soit $\mathcal{M}$ une machine de Turing à $\geq m + n + 1$ bandes. Étant donné $(a_1, \ldots, a_n) \in \mathbb{N}^n$, on considère la machine $\mathcal{M}'$ qui fonctionne ainsi :

(1) Elle écrit $a_1, \ldots, a_n$ sur les bandes $B_{m+2}, \ldots, B_{m+n+1}$.

(2) Puis, elle travaille comme $\mathcal{M}$, à permutation des bandes près, et écrit le résultat sur la bande B_{m+1}.

(3) Elle nettoie les bandes $B_{m+2}, \ldots, B_{m+n+1}$, puis s'arrête.

Il est clair qu'il existe une fonction primitive récursive $g \in \mathcal{F}_{n+1}$ qui calcule $\ulcorner \mathcal{M}' \urcorner \in I_m$ à partir de $\ulcorner \mathcal{M} \urcorner$ et $a_1, \ldots, a_n$, c'est-à-dire $g(\ulcorner \mathcal{M} \urcorner, a_1, \ldots, a_n) = \ulcorner \mathcal{M}' \urcorner$ et $g(i, \overline{a}) = 0$ si $i \notin I_{m+n}$. Il suffit donc de poser $s^m_n = g$. □

Théorème 4.5.3 (Théorème du point fixe)**.** *Soit $m > 0$ et $\alpha \in \mathcal{F}_1$ récursive totale. Alors il existe $i \in \mathbb{N}$ tel que $\varphi^m_i = \varphi^m_{\alpha(i)}$.*

Démonstration. On considère $g = \lambda y x_1 \ldots x_m.\varphi^m(\alpha(s^m_1(y, y)), x_1, \ldots, x_m)$. Par universalité (Proposition 4.5.1), on a $g = \varphi^{m+1}_a$ pour un indice $a \in \mathbb{N}$. On calcule

$$\varphi^m_{\alpha(s^m_1(y,y))}(\overline{x}) = \varphi^{m+1}_a(y, \overline{x}) = \varphi^m_{s^m_1(a,y)}(\overline{x}).$$

Il suffit de poser $i := s^m_1(a, a)$ pour obtenir $\varphi^m_{\alpha(i)} = \varphi^m_i$. □

Le théorème du point fixe fournit un argument élégant pour montrer le résultat suivant (cf. Exercice 4.3.1).

Proposition 4.5.4. *La fonction d'Ackermann ξ est récursive.*

Démonstration. On définit $\theta \in \mathcal{F}^*_3$ récursive partielle comme suit :

– $\theta(i, y, x) = 2^x$ si $y = 0$;

– $\theta(i,y,x) = 1$ si $x = 0$;
– $\theta(i,y,x) = \varphi^2(i, y\dot{-}1, \varphi^2(i,y,x\dot{-}1))$ sinon.

Par universalité, il existe $a \in \mathbb{N}$ tel que $\theta = \varphi^3_a$. Par le théorème SMN, on a donc

$$\theta(i,y,x) = \varphi^2_{s^2_1(a,i)}(y,x).$$

Posons $\alpha = \lambda i.s^2_1(a,i)$. Par le théorème du point fixe il existe $i_0 \in \mathbb{N}$ tel que $\varphi^2_{i_0} = \varphi^2_{\alpha(i_0)}$. La fonction $\varphi^2_{i_0}$ est récursive et satisfait aux relations qui définissent la fonction d'Ackermann : pour $y, x \geq 1$, on a

$$\varphi^2_{i_0}(y,x) = \theta(i_0,y,x) = \varphi^2(i_0, y-1, \varphi^2(i_0,y,x-1)) = \varphi^2_{i_0}(y-1, \varphi^2_{i_0}(y,x-1)).$$

(Notons que la totalité de $\varphi^2_{i_0}$ s'obtient également par récurrence.) □

4.6 Ensembles récursivement énumérables

Définition. Un ensemble $R \subseteq \mathbb{N}^n$ est appelé *récursivement énumérable* (r.e.) s'il existe $S \subseteq \mathbb{N}^{n+1}$ récursif tel que $R = \pi(S)$, où π est la projection sur les n premières coordonnées.

Une relation $R(x_1, \ldots, x_n)$ sur $\mathbb{N}$ est donc récursivement énumérable s'il existe une relation récursive $S(\overline{x}, y)$ telle que (dans $\mathbb{N}$) $R(\overline{x}) \iff \exists y S(\overline{x}, y)$.

Remarque. *Tout ensemble récursif est r.e.* □

Proposition 4.6.1. *L'ensemble des relations r.e. est clos par conjonction, disjonction, quantification universelle bornée et par quantification existentielle. De plus, si $f_1, \ldots, f_n \in \mathcal{F}_m$ sont récursives et si $R \subseteq \mathbb{N}^n$ est r.e., alors $\{\overline{x} \in \mathbb{N}^m \mid (f_1(\overline{x}), \ldots, f_m(\overline{x})) \in R\}$ est r.e.*

Démonstration. Soient P, Q et R des relations telles que $P(\overline{x}) \iff \exists y \overline{P}(\overline{x}, y)$, $Q(\overline{x}) \iff \exists y \overline{Q}(\overline{x}, y)$ et $R(\overline{x}, z) \iff \exists y \overline{R}(\overline{x}, z, y)$ pour $\overline{P}, \overline{Q}$ et $\overline{R}$ récursives. Alors on a les équivalences suivantes :

$$(P(\overline{x}) \vee Q(\overline{x})) \iff \exists y \left(\overline{P}(\overline{x}, y) \vee \overline{Q}(\overline{x}, y)\right)$$

$$(P(\overline{x}) \wedge Q(\overline{x})) \iff \exists y \left(\overline{P}(\overline{x}, \beta^2_1(y)) \wedge \overline{Q}(\overline{x}, \beta^2_2(y))\right)$$

$$\exists z R(\overline{x}, z) \iff \exists z \exists y \overline{R}(\overline{x}, z, y) \iff \exists s \overline{R}(\overline{x}, \beta^2_1(s), \beta^2_2(s))$$

$$R(f_1(\overline{x}), \ldots, f_n(\overline{x})) \iff \exists y \overline{R}(f_1(\overline{x}), \ldots, f_n(\overline{x}), y)$$

$$\forall z \leq w R(\overline{x}, z) \iff \forall z \leq w \exists y \overline{R}(\overline{x}, z, y) \iff \exists s \forall z \leq w \overline{R}(\overline{x}, z, (s)_z)$$

Ici, $(s)_z$ désigne la fonction composante du lemme 4.1.4. □

Théorème 4.6.2. *Soit $X \subseteq \mathbb{N}^n$. Sont équivalents :*

1. *X est r.e.*
2. *X est vide ou l'image d'une fonction $f = (f_1, \ldots, f_n) : \mathbb{N} \to \mathbb{N}^n$ avec $f_1, \ldots, f_n \in \mathcal{F}_1$ récursives (totales).*

3. $X = \mathrm{dom}(f)$ *pour une fonction partielle récursive* $f \in \mathcal{F}_n^*$.

4. *Il existe* $Y \subseteq \mathbb{N}^{n+1}$ *primitif récursif tel que* $X = \pi(Y)$.

Démonstration. Quitte à identifier $\mathbb{N}^n$ et $\mathbb{N}$ par α_n (une fonction primitive récursive), on peut supposer $n = 1$.
(1)⇒(2) : Soit $r \in X$ et $X = P_1^2(Y)$ pour Y récursif. Alors $X = \mathrm{im}(f)$, pour $f \in \mathcal{F}_1$ définie ainsi :

$$f(z) := \begin{cases} \beta_1^2(z), \text{ si } (\beta_1^2(z), \beta_2^2(z)) \in Y\,; \\ r, \text{ sinon.} \end{cases}$$

(2)⇒(3) : On a $\emptyset = \mathrm{dom}(g)$ pour $g = \lambda x.\mu t(x > x)$. Si $X = \mathrm{im}(f)$, on définit $g = \lambda x.\mu t(x = f(t))$. Alors $\mathrm{dom}(g) = \mathrm{im}(f) = X$.
(3)⇒(4) : Soit $f \in \mathcal{F}_1^*$ récursive partielle et $X = \mathrm{dom}(f)$. Alors par universalité il existe $i_0 \in \mathbb{N}$ tel que $f = \varphi_{i_0}^1$, c'est-à-dire

$$X = \{x \in \mathbb{N} \mid \exists t \left[(i_0, t, x) \in B^1\right]\}.$$

Cela établit (4), puisque B^1 est primitif récursif.
(4)⇒(1) : Trivial. □

Remarque. *Notons que la preuve de (1)⇒(2), compte tenu de l'équivalence entre (1) et (4), montre que l'on pourra même supposer dans (2) que les fonctions* $f_1, \ldots, f_n$ *sont primitives récursives.*

Corollaire 4.6.3. *L'ensemble* $U = \mathrm{dom}(\varphi^n) \subseteq \mathbb{N}^{n+1}$ *est* universel r.e. *: il est r.e. et pour tout* $R \subseteq \mathbb{N}^n$ *r.e. il existe* $e \in \mathbb{N}$ *tel que* $R = \{\overline{x} \mid (e, \overline{x}) \in U\}$. □

Théorème 4.6.4. *Un ensemble* $X \subseteq \mathbb{N}^n$ *est récursif ssi* X *et* $\mathbb{N}^n \setminus X$ *sont r.e.*

Démonstration. "⇒" est clair.

Quant à "⇐", soit $X \subseteq \mathbb{N}^n$ tel que $X = \pi(Y)$ et $\mathbb{N}^n \setminus X = \pi(Y')$ avec $Y, Y' \subseteq \mathbb{N}^{n+1}$ récursifs. Alors $\mathbb{1}_X(\overline{z}) = \mathbb{1}_Y(\overline{z}, \mu t[(\overline{z}, t) \in Y \cup Y'])$). □

Théorème 4.6.5. *L'ensemble* $\mathrm{dom}(\varphi^1)$ *n'est pas récursif. En particulier, il existe un ensemble r.e. qui n'est pas récursif.*

Démonstration. On montre que $g = \lambda x.\varphi^1(x, x)$ n'est pas de domaine récursif. (On ne peut pas décider si une machine s'arrête quand on la fait fonctionner sur son propre code.)

Soit $D = \mathbb{N} \setminus \mathrm{dom}(g)$. Si D était r.e., alors $D = \mathrm{dom}(\varphi_{i_0}^1)$ pour un $i_0 \in \mathbb{N}$ par le corollaire 4.6.3. En particulier, on aurait $i_0 \in D$ si et seulement si $\varphi^1(i_0, i_0)$ est définie. Or, par définition de g, $i_0 \in D$ si et seulement si $\varphi^1(i_0, i_0)$ n'est pas définie. Absurde. □

Une petite variation montre que le *problème de l'arrêt* n'est pas décidable :

Théorème 4.6.6 (Undécidabilité du problème de l'arrêt)**.** *L'ensemble des indices des machines de Turing qui s'arrêtent lorsqu'elles commencent à fonctionner sur l'entrée vide n'est pas récursif.*

Démonstration. Par l'absurde. Si l'ensemble

$$I_a = \{i \in I_1 \mid \mathcal{M} \text{ d'indice } i \text{ s'arrête si elle est lancée sur l'entrée vide}\}$$

est récursif, il est facile à voir que l'ensemble $I_s = \operatorname{dom}\left(\lambda x.\varphi^1(x,0)\right)$ est récursif aussi, où I_s est donc l'ensemble des indices des machines de Turing qui s'arrêtent sur une configuration de sortie pour 0 lorsqu'elles sont lancées sur l'entrée vide : si $i \in I_a$, alors $i \in I_s$ ssi $\left(i, \beta_3^3(\mathrm{ST}^1(i, \mu t[\beta_1^3(\mathrm{ST}^1(i,t,0)) = 1], 0)), 0\right) \in E^1$.

On considère $A \subseteq \mathbb{N}$ r.e. non récursif (un tel A existe par 4.6.5). Soit $\mathcal{M}$ une machine de Turing qui calcule une fonction $f \in \mathcal{F}_2^*$ avec $\operatorname{dom}(f) = A \times \{0\}$. Soit i_0 l'indice de $\mathcal{M}$. Alors $A = \{n \in \mathbb{N} \mid s_1^1(i_0, n) \in I_s\}$, ce qui montre que A est récursif. Contradiction. □

Théorème 4.6.7 (Théorème de Rice)**.** *Soit $\mathcal{X}$ un ensemble de fonctions partielles récursives à une variable. On suppose que $\mathcal{X}$ est non vide et distinct de l'ensemble de toutes les fonctions partielles récursives à une variable. Alors $I = \{i \in \mathbb{N} \mid \varphi_i^1 \in \mathcal{X}\}$ n'est pas récursif.*

Démonstration. Quitte à passer au complément, on peut supposer que la fonction de domaine vide est dans $\mathcal{X}$. On choisit $j \in \mathbb{N} \setminus I$, et on note $\psi(x,y,z) = \left(\varphi^1(j,z) + \varphi^1(x,y)\right) \dot{-} \varphi^1(x,y)$, puis $\psi_{x,y} = \lambda z.\psi(x,y,z)$. Alors

$$\psi_{x,y} \in \mathcal{X} \iff (x,y) \notin \operatorname{dom}(\varphi^1) = U.$$

On a $\psi = \varphi_{i_0}^3$ pour un entier i_0, donc $\psi_{x,y} = \varphi^1_{s_2^1(i_0,x,y)}$ par le théorème SMN. Posons $h = \lambda xy.s_2^1(i_0,x,y)$. C'est une fonction primitive récursive. On a $(x,y) \in \mathbb{N}^2 \setminus U \iff h(x,y) \in I$. Comme $\mathbb{N}^2 \setminus U$ n'est pas récursif par le théorème 4.6.4, I n'est pas récursif non plus. □

Exemples 4.6.8.

1. Soit $f \in \mathcal{F}_1^*$ une fonction partielle récursive. Alors l'ensemble $\{i \in \mathbb{N} \mid \psi_i^1 = f\}$ n'est pas récursif.
2. $\{(i,j) \in \mathbb{N}^2 \mid \varphi_i^1 = \varphi_j^1\} \subseteq \mathbb{N}^2$ n'est pas récursif.
3. L'ensemble des indices des fonctions totales (récursives) n'est pas récursif.

4.7 Élimination de la récurrence

Théorème 4.7.1 (Élimination de la récurrence)**.** *Toute fonction récursive totale s'obtient à partir des fonctions de base $C_0^0, S, P_i^n, +, \cdot$ et $\mathbb{1}_<$, en appliquant uniquement les règles $(R1)$ et $(R3)$.*

Démonstration. Une fonction totale est appelée #-récursive si elle satisfait à la conclusion du théorème 4.7.1. Par la proposition 4.4.10(3), il suffira de montrer que la classe des fonctions #-récursives est close par $(R2)$ (récurrence).

Lemme 4.7.2. *(1) La fonction $\lambda xy.x\dot{-}y$ est #-récursive.*

(2) La classe des ensembles #-récursifs est close par combinaisons booléennes et quantification bornée.

(3) $\mathbb{1}_{=}$ est #-récursive.

(4) L'ensemble $\{(x,y,z) \mid x \equiv y \mod z\}$ est #-récursif.

(5) La classe des fonctions #-récursives est close par définition par cas (voir Lemme 4.1.2(5)), si les fonctions et les ensembles de la partition sont #-récursifs.

Preuve du lemme 4.7.2. (1) $x\dot{-}y = \mu z\,(x < (y+z)+1)$

On a $\mathbb{1}_{X^c} = 1\dot{-}\mathbb{1}_X$ et $\mathbb{1}_{X\cap Y} = \mathbb{1}_X \cdot \mathbb{1}_Y$, ce qui établit la clôture par combinaisons booléennes.

(3) On a $x = y \iff (\neg x < y \wedge \neg y < x)$.

Maintenant, soit $X \subseteq \mathbb{N}^{n+1}$ un ensemble #-récursif. On définit $g \in \mathcal{F}_{n+1}$,

$$g(\overline{x}, z) := \mu y\,((\overline{x}, y) \in X \text{ ou } y = z+1)\,.$$

Comme la condition entre parenthèses s'exprime facilement par $f(\overline{x}, y, z) = 0$ pour une fonction #-récursive f, on voit que g est #-récursive aussi. On a $\exists y \leq z : (\overline{x}, y) \in X \iff g(\overline{x}, z) < z+1$, d'où la clôture par quantification existentielle bornée. Comme $\forall y \leq z\,\varphi$ est équivalent à $\neg\exists y \leq z\,\neg\varphi$, la preuve de (2) est terminée.

(4) On a $x \equiv y \mod z \iff \exists w \leq x + y\big(x = y + w \cdot z \vee y = x + w \cdot z\big)$.

(5) Soit $\mathbb{N}^n = A_1 \dot\cup \ldots \dot\cup A_k$ une partition avec A_i #-récursif pour tout i, et soient $f_1, \ldots, f_k \in \mathcal{F}_n$ des fonctions #-récursives. Par induction sur k, on montre que $\sum_{i=1}^{k} \mathbb{1}_{A_i} \cdot f_i$ est #-récursive. □

Lemme 4.7.3 (La fonction β de Gödel)**.** *Il existe une fonction #-récursive $\beta \in \mathcal{F}_3$ telle que pour toute suite finie d'entiers $(c_0, \ldots, c_{n-1})$ il existe $a, b \in \mathbb{N}$ avec $\beta(a,b,i) = c_i$ pour $i = 0, \ldots, n-1$.*

Démonstration. On pose $\beta(a,b,i) := \mu z\,[z \equiv a \mod (b(i+1)+1)]$. La fonction β est #-récursive par le lemme 4.7.2. Soient $c_0, \ldots, c_{n-1}$ donnés, et soit $b \in \mathbb{N}$, b divisible par $n!$ et $b > c_i$ pour tout i. Alors $b+1, 2b+1, 3b+1, \ldots, nb+1$ sont 2 à 2 premiers entre eux. En effet, si p premier divise $ib+1$ et $jb+1$ pour $1 \leq i < j \leq n$, alors $p \nmid b$ et $p \mid b(i-j)$. Donc $p \mid (i-j)$. Comme $(i-j) \mid n!$, c'est absurde.

Par le théorème chinois il existe a tel que pour tout $i = 0, \ldots, n-1$ on ait $a \equiv c_i \mod ((i+1)b+1)$. Comme $c_i < (i+1)b+1$ pour tout i, c_i est le plus petit entier z tel que $z \equiv a \mod ((i+1)b+1)$. □

Revenons à la preuve du théorème 4.7.1. Soient donc $g \in \mathcal{F}_n$ et $h \in \mathcal{F}_{n+2}$ #-récursives, et soit $f \in \mathcal{F}_{n+1}$ définie par récurrence à partir de g et h, c'est-à-dire $f(\overline{x}, 0) = g(\overline{x})$ et $f(\overline{x}, y+1) = h(\overline{x}, y, f(\overline{x}, y))$. La relation

$$R(\overline{x}, y, a, b) :\iff [\beta(a,b,0) = g(\overline{x}) \wedge \forall i \leq y\,(\beta(a,b,i+1) = h(\overline{x}, i, \beta(a,b,i)))]$$

est #-récursive par les lemmes 4.7.2 et 4.7.3.

Par le choix de β, on a $\forall \overline{x}, y \exists a, b\, R(\overline{x}, y, a, b)$. Alors la fonction e qui à $(\overline{x}, y)$ associe $\mu s\, (\exists a \leq s \exists b \leq s R(\overline{x}, y, a, b))$ est #-récursive.

Enfin, $f(\overline{x}, y) = \mu z\, [\exists a \leq e(\overline{x}, y) \exists b \leq e(\overline{x}, y)\, (R(\overline{x}, y, a, b) \wedge z = \beta(a, b, y))]$, ce qui montre que f est #-récursive. □

Chapitre 5

Modèles de l'arithmétique et théorèmes de limitation

5.1 Codage des formules et des preuves

Soit $\sigma^{\mathcal{L}} = \{\lambda_1, \ldots, \lambda_l\}$ une signature finie. On associe aux symboles s de $\mathcal{L}$ leur *nombre de Gödel* $\ulcorner s \urcorner$ comme suit : à $\doteq, \wedge, \neg, (,), \exists$ on associe successivement $\langle 0,0\rangle, \langle 0,1\rangle, \ldots, \langle 0,5\rangle$, à λ_i on associe $\langle 0, i+5\rangle$, et à v_i ($i \in \mathbb{N}$) on associe $\langle 1, i\rangle$.

À un mot $m = s_1 \cdots s_n$ sur l'alphabet $\mathcal{L}$ on associe le *nombre de Gödel*

$$\#m = \langle \ulcorner s_1 \urcorner, \ldots, \ulcorner s_n \urcorner \rangle.$$

Ce codage est évidemment injectif.

Lemme 5.1.1. *Les ensembles suivants sont primitifs récursifs :*

1. Term $= \{\#t \mid t$ *est un* $\mathcal{L}$*-terme*$\}$
2. Form $= \{\#\varphi \mid \varphi$ *est une* $\mathcal{L}$*-formule*$\}$
3. $\{(\#t, n) \mid t$ *est un* $\mathcal{L}$*-terme dans lequel* v_n *a une occurrence*$\}$ *et* $\{(\#t, n) \mid t$ *est un* $\mathcal{L}$*-terme dans lequel* v_n *n'a pas d'occurrence*$\}$
4. $\{(\#\varphi, n) \mid \varphi$ *est une formule dans laquelle* v_n *a une occurrence*$\}$, *de même pour* 'n'a pas d'occurrence', 'a au moins une occurrence libre', 'n'a pas d'occurrence libre', 'a au moins une occurrence liée' *et* 'n'a pas d'occurrence liée' *à la place de* 'a une occurrence'.
5. $\{\#\varphi \mid \varphi$ *est un* $\mathcal{L}$*-énoncé*$\}$.

Démonstration. On utilise les propriétés du codage $\langle \ldots \rangle$ des suites finies (Lemme 4.1.4), en particulier $n = \lg(\langle s_0, \ldots, s_{n-1}\rangle) \leq \langle s_0, \ldots, s_{n-1}\rangle$ et la fonction primitive récursive de décodage à deux places $\lambda xi.(x)_i$.

Par les propriétés de lecture unique, on pourra non seulement argumenter par induction sur la longueur, mais aussi par induction sur la hauteur (des termes et des formules). Par exemple si x est le code d'un mot m qui commence avec

'(', alors x est le code d'une formule si et seulement s'il existe $y_1, y_2 < x$ tels que $y_i = \#\varphi_i$ pour des formules φ_1, φ_2 et tels que $m = (\varphi_1 \wedge \varphi_2)$. Les détails sont laissés en exercice. □

Rappelons que nous avons défini la substitution dans les termes et les formules. Si $x_1, \ldots, x_n$ sont n variables 2-à-2 distinctes, et $s_1, \ldots, s_n$ des termes, nous avons défini $t_{\overline{s}/\overline{x}}$ et $\varphi_{\overline{s}/\overline{x}}$. On montre sans problème que les fonctions suivantes $S_t : \mathbb{N}^3 \to \mathbb{N}$ et $S_f : \mathbb{N}^3 \to \mathbb{N}$ sont primitives récursives : S_t est la fonction qui à (a, b, c) associe $\#t_{\overline{s}/v_{i_1},\ldots,v_{i_n}}$ si $a = \#t \in \text{Term}$, $b = \langle i_1, \ldots, i_n\rangle$ est le code d'une suite d'entiers 2-à-2 distincts et $c = \langle \#s_1, \ldots, \#s_n\rangle$, où $\#s_i \in \text{Term}$ pour tout i, et 0 sinon.

La fonction S_f est définie de la même manière, c.-à-d. pour $(a, b, c) = (\#\varphi, \langle i_1, \ldots, i_n\rangle, \langle \#s_1, \ldots, \#s_n\rangle)$ convenable on a $S_f(a, b, c) = \#\varphi_{\overline{s}/v_{i_1},\ldots,v_{i_n}}$.

En particulier, nous obtenons donc le résultat suivant.

Lemme 5.1.2. *Il existe des fonctions primitives récursives* $\text{Subst}_t, \text{Subst}_f \in \mathcal{F}_3$ *telles que pour tout* n*, si* s *et* t *sont des termes et* φ *une formule, alors* $\text{Subst}_t(n, \#s, \#t) = \#t_{s/v_n}$ *et* $\text{Subst}_f(n, \#s, \#\varphi) = \#\varphi_{s/v_n}$. □

De la même manière, on code les formules du calcul propositionnel, via $F \mapsto \#F \in \mathbb{N}$, où $F \in \mathcal{F}ml_{\mathcal{P}}$.

Lemme 5.1.3. *L'ensemble* Taut *des codes des tautologies du calcul des prédicat est primitif récursif.*

Démonstration. On montre :

- $\text{Form}_{\mathcal{P}} = \{\#F \mid F \in \mathcal{F}ml_{\mathcal{P}}\}$ est un ensemble primitif récursif ;
- la fonction qui à (d, x) associe la valeur de vérité (0 ou 1) de la formule $F = F[p_0, \ldots, p_{n-1}]$ pour la distribution des valeurs de vérité $d = \langle d_0, \ldots, d_{n-1}\rangle$ si d est le code d'une suite de 0 et 1 de longueur n et $x = \#F$ est le code de $F \in \mathcal{F}ml_{\mathcal{P}}$ ne contenant que des variables propositionnelles p_i avec $i \leq n-1$, et 0 sinon, est primitive récursive ;
- l'ensemble $\text{Taut}_{\mathcal{P}} = \{\#F \mid F \text{ est une tautologie}\}$ est primitif récursif ;
- il existe une fonction primitive récursive qui à $(\langle \#\varphi_0, \ldots \#\varphi_{n-1}\rangle, \#F)$ associe $\#F_{\overline{\varphi}/\overline{p}}$ si $F = F[p_0, \ldots, p_{n-1}]$.

Ceci permet de conclure. □

On peut également reconnaître de manière primitive récursive les codes des axiomes de l'égalité ainsi que les codes des axiomes du quanteur existentiel (compte tenu du lemme 5.1.2). Nous avons donc montré le théorème suivant.

Théorème 5.1.4. *L'ensemble* Ax *des codes* $\#\varphi$ *des axiomes logiques* φ *de* $\mathcal{L}$ *est primitif récurisf.* □

Codage des preuves formelles

Si $d = (\varphi_0, \ldots, \varphi_{n-1})$ est une suite finie de $\mathcal{L}$-formules, on la code via $\#\#d = \langle \#\varphi_0, \ldots, \#\varphi_{n-1}\rangle$.

Lemme 5.1.5. $\mathrm{Dem} = \{(x, \#\varphi) \mid x = \#\#d$ *et* d *est une preuve formelle de* $\varphi\}$ *est un ensemble primitif récursif.*

Démonstration. Il suffit de décoder et de tester si toutes les composantes de x sont des codes de $\mathcal{L}$-formules φ_i, la dernière étant égale à φ, et si pour tout i, soit φ_i est un axiome logique (ce qui se voit de manière primitive récursive par le théorème 5.1.4), soit elle peut s'obtenir par une règle de déduction à partir de formules déjà obtenues avant. □

Proposition 5.1.6. $U = \{\#\varphi \mid \vdash \varphi\}$ *est récursivement énumérable.*

Démonstration. On a $n \in U \Leftrightarrow \exists d\,(d, n) \in \mathrm{Dem}$. □

5.2 Théories décidables

Pour T une théorie, on note $\mathrm{Thm}(T) = \{\varphi\ \mathcal{L}\text{-énoncé} \mid T \vdash \varphi\}$, la clôture déductive de T.

Définition. Soit T une $\mathcal{L}$-théorie.

1. T est dite *récursive* si $\{\#\varphi \mid \varphi \in T\}$ est un ensemble récursif.
2. T est dite *récursivement* (ou *effectivement*) *axiomatisable* s'il existe une $\mathcal{L}$-théorie récursive T_0 telle que $\mathrm{Thm}(T) = \mathrm{Thm}(T_0)$.
3. T est dite *décidable* si $\mathrm{Thm}(T)$ est une théorie récursive, c.-à-d. si l'ensemble $\#\mathrm{Thm}(T) = \{\#\varphi \mid T \vdash \varphi\}$ est récursif.

Si T est une $\mathcal{L}$-théorie, on définit l'ensemble

$$\mathrm{Dem}(T) = \{(\#\#d, \#\varphi) \mid \ d \text{ est une preuve de } \varphi \text{ dans } T\}.$$

Lemme 5.2.1. *Si* T *est récursive,* $\mathrm{Dem}(T)$ *est un ensemble récursif.*

Démonstration. Exercice. □

Théorème 5.2.2. *Si* T *est effectivement axiomatisable, alors* $\#\mathrm{Thm}(T)$ *est récursivement énumérable.*

Démonstration. Soit T_0 une théorie récursive telle que $\mathrm{Thm}(T) = \mathrm{Thm}(T_0)$.

On a $\varphi \in \mathrm{Thm}(T)$ si et seulement si φ est un $\mathcal{L}$-énoncé et il existe une preuve de φ dans T_0. On conclut par les lemmes 5.1.1(5) et 5.2.1. □

Remarque 5.2.3. *La réciproque du théorème 5.2.2 est vraie aussi.*

Démonstration. Supposons que l'ensemble $\#\mathrm{Thm}(T)$ est récursivement énumérable. Il existe donc $f \in \mathcal{F}_1$ récursive telle que $\#\mathrm{Thm}(T) = \{f(i) = \#\varphi_i \mid i \in \mathbb{N}\}$. La fonction g qui à n associe $\#\bigwedge_{i=0}^{n} \varphi_i$ est récursive et on a $g(n) \geq n$ pour tout n. Son image est donc un ensemble récursif. Par ailleurs, $\{\bigwedge_{i=0}^{n} \varphi_i \mid n \in \mathbb{N}\}$ est une axiomatisation de $\mathrm{Thm}(T)$. □

Théorème 5.2.4. *Soit T effectivement axiomatisable et complète. Alors T est décidable.*

Démonstration. Par le théorème 5.2.2, on sait que $\#\mathrm{Thm}(T)$ est récursivement énumérable. Or, son complémentaire $\mathbb{N} \setminus \#\mathrm{Thm}(T)$ est donné par la rénuion de $\{x \mid x \text{ n'est pas le code d'un } \mathcal{L}\text{-énoncé}\}$ et de $\{\#\varphi \mid \neg\varphi \in \mathrm{Thm}(T)\}$.

Le premier ensemble est primitif récursif, et le deuxième est récursivement énumérable, car $\#\varphi \mapsto \#\neg\varphi$ est donné par une fonction (primitive) récursive. On conclut par le théorème 4.6.4. □

Exemples 5.2.5.

1. Toute théorie finiment axiomatisable est effectivement axiomatisable. [Clair.]
2. CAC_p, p premier ou 0, est décidable.
   [En effet, l'axiomatisation que nous avons donnée est certainement récursive, et on conclut donc en combinant le théorème 3.5.3 (CAC_p est compléte) avec le théorème 5.2.4.]
3. CAC est une théorie décidable.
   [En effet, comme CAC est effectivement axiomatisable, il suffit de montrer que l'ensemble $X = \{\#\varphi \mid \varphi \text{ est un énoncé et } \mathrm{CAC} \not\models \varphi\}$ est récursivement énumérable. Or, $\mathrm{CAC} \not\models \varphi$ si et seulement s'il existe p premier tel que $\mathrm{CAC}_p \models \neg\varphi$. (C'est une conséquence du Principe de Lefschetz.) Cela est équivalent à $\mathrm{CAC} \models \chi_p \to \neg\varphi$, où $\chi_p = \underbrace{1 + \cdots + 1 \doteq 0}_{p \text{ fois}}$.
   Il s'en suit que $\#\varphi \in X$ si et seulement s'il existe p premier et y tels que $(y, \#(\chi_p \to \neg\varphi)) \in \mathrm{Dem}(\mathrm{CAC})$. Comme l'application $(p, \#\varphi) \mapsto \#(\chi_p \to \neg\varphi)$ est (primitive) récursive, on conclut que X est récursivement énumérable.]
4. La théorie du corps ordonné des réels $\mathcal{R}$ est décidable (c'est un résultat de Tarski).
   [Comme elle est complète, il suffit de donner une axiomatisation effective. On peut montrer que tout corps ordonné $\langle K; +, -, 0, 1, \cdot, < \rangle$ satisfaisant aux deux propriétés suivantes est élémentairement équivalent à $\mathcal{R}$:
   – tout élément positif est un carré $(\forall x(x \geq 0 \to \exists y\, y \cdot y \doteq x))$;
   – K a la propriété de la valeur intermédiaire pour les fonctions polynômiales, c.-à-d. si $P(X) \in K[X]$ est un polynôme et $a < b$ avec $P(a) \cdot P(b) < 0$, alors il existe $c \in K$ tel que $a < c < b$ et $P(c) = 0$.
   Pour le prouver, Tarski a montré que $\mathrm{Th}(\mathcal{R})$ admet l'élimination des quanteurs dans le langage des corps ordonnés, un résultat que nous ne pouvons pas présenter dans ce cours, par manque de temps.]

Exercice 5.2.6. Si T est décidable et $\varphi_0, \ldots, \varphi_{n-1}$ sont des $\mathcal{L}$-énoncés, alors $T \cup \{\varphi_0, \ldots, \varphi_{n-1}\}$ est décidable.

5.3 Arithmétique de Peano

On considère le langage de l'arithmétique $\mathcal{L}_{ar} = \{0, S, +, \cdot, <\}$.

Définition. L'ensemble des *axiomes de Peano faibles* est l'ensemble fini $\mathcal{P}_0$ des 8 axiomes suivants :

(A1) $\forall v_0\, \neg Sv_0 \doteq 0$

(A2) $\forall v_0\, \exists v_1 (\neg v_0 \doteq 0 \to Sv_1 \doteq v_0)$

(A3) $\forall v_0 \forall v_1\, (Sv_0 \doteq Sv_1 \to v_0 \doteq v_1)$

(A4) $\forall v_0\, v_0 + 0 \doteq v_0$

(A5) $\forall v_0 \forall v_1\, v_0 + Sv_1 \doteq S(v_0 + v_1)$

(A6) $\forall v_0\, v_0 \cdot 0 \doteq 0$

(A7) $\forall v_0 \forall v_1\, v_0 \cdot Sv_1 \doteq (v_0 \cdot v_1) + v_0$

(A8) $\forall v_0 \forall v_1\, (v_0 < v_1 \leftrightarrow (\exists v_2 v_2 + v_0 \doteq v_1 \wedge \neg v_0 \doteq v_1))$

L'ensemble des *axiomes de Peano* est l'ensemble (infini) $\mathcal{P}$ formé de $\mathcal{P}_0$, ainsi que pour chaque $\mathcal{L}_{ar}$-formule $\varphi = \varphi[v_0, \ldots, v_n]$ de l'axiome d'induction suivant (on pose $\overline{v} = (v_1, \ldots, v_n)$) :

$$\forall v_1 \forall v_2 \cdots \forall v_n ((\varphi(0, \overline{v}) \wedge \forall v_0 (\varphi(v_0, \overline{v}) \to \varphi(Sv_0, \overline{v})) \to \forall v_0 \varphi(v_0, \overline{v})).$$

Remarque 5.3.1. *1. On a $\mathcal{N} = \langle \mathbb{N}; 0, succ, +, \cdot, < \rangle \models \mathcal{P}$.*

2. *$\mathcal{P}_0$ est une expansion par définition de la $\mathcal{L}_{ar} \setminus \{<\}$-théorie (A1)-(A7). (C'est une conséquence de (A8).)*
3. *$\mathcal{P}_0$ est une théorie très faible. On peut construire des modèles dans lesquels l'addition (ou la multiplication) n'est pas commutative / associative ; de même, il y a des modèles dans lesquels $\leq$ ne définit pas une relation d'ordre.*
4. *Pour $n \in \mathbb{N}$, dans tout modèle de $\mathcal{P}_0$ on trouve l'élément qui interprète $\underline{n} = \underbrace{S \cdots S}_{n \text{ fois}} 0$. Un élément est dit* non standard *s'il est différent de $\underline{n}$ pour tout $n \in \mathbb{N}$. Par compacité, il existe des modèles de $\mathcal{P}$ contenant des éléments non standards.*

Définition. Soit $\mathfrak{M} \subseteq \mathfrak{M}'$ deux $\mathcal{L}_{ar}$-structures. On dit que $\mathfrak{M}$ est un *segment initial* de $\mathfrak{M}'$ si

- pour tout $a' \in M'$ et tout $a \in M$ tel que $\mathfrak{M}' \models a' < a$ on a $a' \in M$, et
- pour tout $a' \in M' \setminus M$ et tout $a \in M$ on a $\mathfrak{M}' \models a < a'$.

Lemme 5.3.2. *Soit $\mathfrak{M} \models \mathcal{P}_0$. Alors $N = \{\underline{n}^{\mathfrak{M}} \mid n \in \mathbb{N}\}$ est l'ensemble de base d'une sous-structure de $\mathfrak{M}$ isomorphe à $\mathcal{N}$ qui forme un segment initial de $\mathfrak{M}$.*

Démonstration. Soit $\mathfrak{M} \models \mathcal{P}_0$. On montre :
(i) Pour tout $m, n \in \mathbb{N}$ on a $\mathcal{P}_0 \models \underline{m} + \underline{n} \doteq \underline{m+n}$.

[Preuve par induction sur n, en utilisant (A4) et (A5).]

(ii) Pour tout $m, n \in \mathbb{N}$ on a $\mathcal{P}_0 \models \underline{m} \cdot \underline{n} \doteq \underline{m \cdot n}$.

[Preuve par induction sur n, en utilisant (A6), (A7) et (i).]

(iii) $\mathcal{P}_0 \models \forall x \forall y (x < y \iff Sx < Sy)$.

[En effet, si $c \in M$, alors $\mathfrak{M} \models (c + x \doteq y \wedge \neg x \doteq y) \overset{(A3)}{\Leftrightarrow} \mathfrak{M} \models (S(c+x) \doteq Sy \wedge \neg Sx \doteq Sy) \overset{(A5)}{\Leftrightarrow} \mathfrak{M} \models (c + Sx \doteq Sy \wedge \neg Sx \doteq Sy)$. Le résultat suit de (A8).]

(iv) Pour tout $n \in \mathbb{N}$ on a $\mathcal{P}_0 \models \forall x(x < \underline{n} \leftrightarrow \bigvee_{i=0}^{n-1} x \doteq \underline{i})$.

[Preuve par induction sur n. Si $\mathfrak{M} \models c < 0$ pour un $c \in M$, alors (dans $\mathfrak{M}$) $a + c = 0$ pour un $a \in M$ et $c \neq 0$ (par (A8)), donc $c = Sd$ pour un d (par (A2)) et enfin (par (A5)) $0 = a + Sd = S(a + d)$, ce qui contredit (A1). Cela montre le cas $n = 0$.

Pour l'étape d'induction, on utlise (iii) ainsi que le fait que $0 < c$ pour tout $c \neq 0$ (conséquence de (A4) et (A8)).]

(v) Soit $c \in M \setminus N$ et $n \in \mathbb{N}$. Alors $\mathfrak{M} \models \underline{n} < c$.

[Preuve par induction sur n, le cas $n = 0$ étant clair. Pour l'étape d'induction $n \mapsto n + 1$, notons que $c = Sd$ pour un $d \notin N$. On a donc $\mathfrak{M} \models \underline{n} < d$, ce qui donne $\mathfrak{M} \models \underline{n+1} < c$, en utilisant (iii).]

Il est clair que cela termine la preuve. □

Quitte à nommer les éléments de $\mathbb{N}$ par des constantes (ce qui est inutile puisque tout élément de $\mathbb{N}$ est donné par l'interprétation d'un terme), nous avons donc en particulier montré que si $\mathfrak{M} \models \mathcal{P}_0$, alors $\mathfrak{M} \models \Delta(\mathcal{N})$.

Définition.
- L'ensemble des *formules* Σ_1 est le plus petit sous-ensemble des $\mathcal{L}_{ar}$-formules contenant les formules sans quanteurs et qui est clos par $\wedge$, $\vee$, quantification existentielle $\exists x$ et *quantification universelle bornée* $\forall x < t$, où t est un terme. [On définit $(\forall x < t)\, \varphi := \forall x(x < t \to \varphi)$.]
- L'ensemble des *formules* Σ_1 *strictes* est le plus petit sous-ensemble des $\mathcal{L}_{ar}$-formules contenant les formules $0 \doteq x$, $sx \doteq y$, $x + y \doteq z$, $x \cdot y \doteq z$, $x \doteq y$, $\neg x \doteq y$, $x < y$, $\neg x < y$ et qui est clos par $\wedge$, $\vee$, $\exists x$ et $\forall x < y$.

Lemme 5.3.3. *1. Toute formule Σ_1 est équivalente à une formule Σ_1 stricte.*

2. Si φ est une formule Σ_1, alors toute formule $\varphi_{\overline{s}/\overline{x}}$ obtenue par substitution est Σ_1 aussi.

Démonstration. On peut éliminer des termes complexes à l'aide de quanteurs existentiels, voir la preuve du lemme 3.3.1. Par exemple, la formule $x \cdot Sy \doteq z$ est équivalente à $\exists u \exists v(Sy \doteq u \wedge x \cdot u \doteq v \wedge Sz \doteq v)$. La seconde partie est claire. □

Proposition 5.3.4. *Tout énoncé Σ_1 vrai dans $\mathcal{N}$ est conséquence de $\mathcal{P}_0$.*

Démonstration. On montrera que pour toute formule Σ_1 stricte $\varphi(x_1, \ldots, x_n)$ et tout $m_1, \ldots, m_n \in \mathbb{N}$ on a

$$\mathcal{N} \models \varphi[m_1, \ldots, m_n] \Rightarrow \mathcal{P}_0 \models \varphi(\underline{m_1}, \ldots, \underline{m_n}).$$

Cela suffira par le lemme 5.3.3.

Si φ est une formule de base, c'est une conséquence du lemme 5.3.2.

Supposons que le résultat a été démontré pour $\varphi(x_0, \ldots, x_n)$ et $\psi(x_0, \ldots, x_n)$. Les cas $(\varphi \wedge \psi)$ et $(\varphi \vee \psi)$ sont alors clairs. Quant au quanteur existentiel, si

$\mathcal{N} \models \exists x_0 \varphi[m_1, \ldots, m_n]$, alors il existe $m_0 \in \mathbb{N}$ tel que $\mathcal{N} \models \varphi[m_0, \ldots, m_n]$, ce qui implique $\mathcal{P}_0 \models \varphi(\underline{m_0}, \ldots, \underline{m_n})$ par hypothèse d'induction, d'où $\mathcal{P}_0 \models \exists x_0 \varphi(\underline{m_1}, \ldots, \underline{m_n})$.

Finalement, si $\mathcal{N} \models (\forall x_0 < x_1)\, \varphi[m_1, \ldots, m_n]$, alors $\mathcal{N} \models \varphi[m_0, m_1, \ldots, m_n]$ pour tout $m_0 < m_1$ par définition de $\forall x_0 < x_1$. Par hypothèse d'induction, on a donc $\mathcal{P}_0 \models \varphi(\underline{m_0}, \ldots, \underline{m_n})$ pour tout $m_0 < m_1$. Par (iv) de la preuve du lemme 5.3.2, on obtient alors $\mathcal{P}_0 \models (\forall x_0 < x_1)\varphi(\underline{m_1}, \ldots, \underline{m_n})$. □

Définition. Soit $f \in \mathcal{F}_p$. On dit que la formule $\varphi = \varphi(x_1, \ldots, x_{p+1})$ *représente* f si pour tout $n_1, \ldots, n_p \in \mathbb{N}$ on a

$$\mathcal{P}_0 \models \forall y \left(\varphi(\underline{n_1}, \ldots, \underline{n_p}, y) \leftrightarrow y \doteq \underline{f(n_1, \ldots, n_p)} \right).$$

On dit que $\varphi(x_1, \ldots, x_p)$ *représente* $A \subseteq \mathbb{N}^p$ si pour tout $n_1, \ldots, n_p \in \mathbb{N}$ on a

$$\overline{n} \in A \Rightarrow \mathcal{P}_0 \models \varphi(\underline{n_1}, \ldots, \underline{n_p}) \quad \text{et} \quad \overline{n} \notin A \Rightarrow \mathcal{P}_0 \models \neg\varphi(\underline{n_1}, \ldots, \underline{n_p}).$$

Remarque. *Si $\varphi(x_1, \ldots, x_p)$ représente A, alors $\mathbb{1}_A$ est représentée par la formule $(\varphi(\overline{x}) \wedge x_{p+1} \doteq \underline{1}) \vee (\neg\varphi(\overline{x}) \wedge x_{p+1} \doteq 0)$.*

Réciproquement, si $\psi(\overline{x}, x_{p+1})$ représente $\mathbb{1}_A$, alors $\psi(\overline{x}, \underline{1})$ représente A. □

Théorème 5.3.5 (Théorème de représentabilité)**.**
Toute fonction récursive totale est représentée par une formule Σ_1.

Démonstration.
- Les fonctions de base $S, P_i^n, C_0^0, +, \cdot$ ainsi que $\mathbb{1}_<$ sont représentables par des formules sans quanteurs (qui sont Σ_1 par définition).
- Si $f_1, \ldots, f_p \in \mathcal{F}_m$ sont représentées par $\varphi_1, \ldots, \varphi_p$ et si $g \in \mathcal{F}_p$ est représentée par ψ, alors $h = g(f_1, \ldots, f_p) \in \mathcal{F}_m$ est représentée par

$$\exists y_1, \ldots, \exists y_p \bigwedge_{i=1}^{p} \varphi_i(\overline{x}, y_i) \wedge \psi(y_1, \ldots, y_p, x_{m+1}).$$

  En particulier, l'ensemble des fonctions qui sont représentables par une formule Σ_1 est clos par composition.
- Soit $f \in \mathcal{F}_{p+1}$ et $g \in \mathcal{F}_p$ définie à partir de f par schéma μ total, c.-à-d. $g(\overline{x}) = \mu y\, (f(\overline{x}, y) = 0)$. Soit $A = \{(\overline{x}, y) \mid f(\overline{x}, y) = 0\} \subseteq \mathbb{N}^{p+1}$. On suppose que f est représentée par une formule Σ_1. Alors $\mathbb{1}_A = \mathbb{1}_=(f(\overline{x}, y), 0)$ est également représentée par une formule Σ_1 (c'est une composition de fonction qui le sont), disons par $\varphi(\overline{x}, y, z)$. La fonction g est donc représentée par la formule Σ_1 $\psi(\overline{x}, y) = \varphi(\overline{x}, y, \underline{1}) \wedge (\forall y' < y)\, \varphi(\overline{x}, y', 0)$.
  En particulier, l'ensemble des fonctions qui sont représentables par une formule Σ_1 est clos par schéma μ total.

On conclut par le théorème 4.7.1. □

Corollaire 5.3.6. *Un ensemble $A \subseteq \mathbb{N}^n$ est récursivement énumérable si et seulement s'il existe une fomule Σ_1 qui définit A (dans la structure $\mathcal{N}$).*

Démonstration. Tout ensemble définissable sans quanteur est (primitif) récursif, donc un ensemble défini par une formule Σ_1 est récursivement énumérable par clôture des ensembles récursivement énumérables par projection.

Réciproquement, comme un ensemble récursivement énumérable est la projection d'un ensemble récursif, il suffit de montrer que tout ensemble récursif est définit par une formule Σ_1, ce qui est une conséquence du théorème de représentabilité. □

Proposition 5.3.7. *Soit $\mathfrak{M}$ un modèle de $\mathcal{P}$. Alors, les propriétés suivantes sont satisfaites dans $\mathfrak{M}$:*

1. *$+$ et $\cdot$ sont commutatives et associatives.*
2. *$\cdot$ est distributive par rapport à $+$.*
3. *$<$ définit un ordre total, compatible avec $+$ (si $a < b$, alors $a + c < b + c$) et avec la multiplication avec un élément non nul (si $a < b$ et $c \neq 0$, alors $a \cdot c < b \cdot c$) ; 0 est l'élément minimal, et pour tout a, Sa est le successeur immédiat de a.*
4. *Tout élément est régulier pour l'addition ($a + c = b + c \Rightarrow a = b$), et tout élément non nul est régulier pour la multiplication (si $a \cdot c = b \cdot c$ et $c \neq 0$, alors $a = b$).*

Démonstration. Toutes ces propriétés s'obtiennent facilement par les axiomes d'induction de $\mathcal{P}$. À titre d'exemple, nous donnons les arguments pour l'associativité et la commutativité de l'addition. Le reste est similaire.

Associativité de l'addition. Soit $\varphi(x, y, z) := x + (y + z) \doteq (x + y) + z$.

Dans $\mathfrak{M}$, on a $a + (b + 0) = a + b = (a + b) + 0$ par (A4), ce qui montre que $\mathfrak{M} \models \forall x, y\, \varphi(x, y, 0)$.

Si $a, b, c \in M$ avec $(a + b) + c = a + (b + c)$, alors on a $a + (b + Sc) = a + S(b + c) = S(a + (b + c)) = S((a + b) + c) = (a + b) + Sc$ par (A5), d'où $\mathfrak{M} \models \forall x, y\, (\varphi(x, y, z) \to \varphi(x, y, Sz))$.

L'axiome d'induction associé à φ donne $\mathfrak{M} \models \forall x, y, z\, \varphi(x, y, z)$.

Commutativité de l'addition. On montre les propriétés suivantes :
(i) $\mathfrak{M} \models \forall x\, 0 + x \doteq x$ (par induction, en utilisant (A4) et (A5))
(ii) $\mathfrak{M} \models \forall x\, \underline{1} + x \doteq Sx$ (par induction, en utilisant (i) et (A5))
(iii) $\mathfrak{M} \models \forall x, y\, x + y \doteq y + x$ (par induction, en utilisant (i), (ii) et l'associativité de $+$) □

Remarque. *Cela montre que $\mathcal{P}$ suffit pour faire la théorie des nombres élémentaire. $\mathcal{P}$ est une théorie incomplète (voir plus loin), mais il n'est pas facile de trouver un énoncé mathématique indépendant de $\mathcal{P}$. Le théorème de Goodstein (TD n° 2, exercice 6) en est un exemple.*

Lemme 5.3.8 (Overspill)**.** *Soit $\mathfrak{M} \models \mathcal{P}$ non standard et $\varphi = \varphi(x)$ une $\mathcal{L}_{ar}$-formule. Si $\mathfrak{M} \models \varphi(\underline{n})$ pour tout $n \in \mathbb{N}$, alors il existe $c \in M$ non standard tel que $\mathfrak{M} \models \varphi[c]$.*

Démonstration. Sinon, on aurait $\mathfrak{M} \models \varphi(0)$ et $\mathfrak{M} \models \forall x(\varphi(x) \to \varphi(Sx))$, car φ n'est satisfaite que par les éléments standards. Mais alors $\mathfrak{M} \models \forall x\, \varphi(x)$, ce qui est absurde car $\mathfrak{M}$ contient des éléments non standards par hypothèse. □

5.4 Les théorèmes de Tarski et de Church

On considère une fonction primitive récursive $\text{subst} \in \mathcal{F}_2$ telle que si $\varphi(v)$ est une $\mathcal{L}_{ar}$-formule à une variable libre v et $n \in \mathbb{N}$, alors $\text{subst}(\#\varphi, n) = \#\varphi(\underline{n})$.

Par le théorème de représentabilité, il existe une formule $G(x, y, z)$ représentant subst et qui est Σ_1. Pour $\varphi = \varphi(v)$, on considère la formule

$$H(x) = \exists z\, (G(x, x, z) \wedge \varphi(z))\,.$$

On pose $n := \#(\neg H)$, puis $\Delta_\varphi := \neg H(\underline{n})$.

Proposition 5.4.1 (L'argument diagonal)**.** *Soit $\varphi = \varphi(v)$ une formule à une variable libre. Alors*

$$\mathcal{P}_0 \vdash \varphi(\underline{\#\Delta_\varphi}) \leftrightarrow \neg\Delta_\varphi.$$

Démonstration. Comme $\text{subst}(n, n) = \text{subst}(\#(\neg H), n) = \#\Delta_\varphi$, pour tout modèle $\mathfrak{M} \models \mathcal{P}_0$, on a

$$\mathfrak{M} \models \forall z \left(G(\underline{n}, \underline{n}, z) \leftrightarrow z \doteq \underline{\#\Delta_\varphi} \right). \tag{5.1}$$

Si $\mathfrak{M} \models \varphi(\underline{\#\Delta_\varphi})$, alors $\mathfrak{M} \models \underbrace{\exists z\, (G(\underline{n}, \underline{n}, z) \wedge \varphi(z))}_{H(\underline{n})}$ par (5.1), et on a donc $\mathfrak{M} \models \neg\Delta_\varphi$ par définition de Δ_φ.

Réciproquement, si $\mathfrak{M} \models \neg\Delta_\varphi$, alors $\mathfrak{M} \models \exists z\, (G(\underline{n}, \underline{n}, z) \wedge \varphi(z))$, et on a donc $\mathfrak{M} \models \varphi(\underline{\#\Delta_\varphi})$ par (5.1). □

Théorème 5.4.2 (Théorème de Tarski sur la non définissabilité de la vérité)**.** *Soit $\mathfrak{M} \models \mathcal{P}_0$. Alors $\#\,\text{Th}(\mathfrak{M})$ n'est pas définissable dans $\mathfrak{M}$, c'est-à-dire il n'existe pas de formule $S_{\mathfrak{M}}(x)$ à une variable libre telle que pour tout $\mathcal{L}_{ar}$-énoncé φ on ait $\mathfrak{M} \models \varphi$ si et seulement si $\mathfrak{M} \models S_{\mathfrak{M}}(\underline{\#\varphi})$.*

Démonstration. Si $S = S_{\mathfrak{M}}$ est une telle formule, par l'argument diagonal on obtient $\mathcal{P}_0 \vdash S(\underline{\#\Delta_S}) \leftrightarrow \neg\Delta_S$. En particulier, on a $\mathfrak{M} \models S(\underline{\#\Delta_S}) \leftrightarrow \neg\Delta_S$, ce qui est absurde, vu que S définit la satisfaction dans $\mathfrak{M}$. □

Corollaire 5.4.3. *Il n'existe pas de $\mathcal{L}_{ar}$-formule $S_{\mathcal{N}}(x)$ telle que pour tout $\mathcal{L}_{ar}$-énoncé φ on ait $\mathcal{N} \models \varphi$ si et seulement si $\mathcal{N} \models S_{\mathcal{N}}(\underline{\#\varphi})$.*

En particulier, $\text{Th}(\mathcal{N})$ *n'est pas décidable.* □

Théorème 5.4.4 (Théorème de Church)**.** *Soit $T \supseteq \mathcal{P}_0$ une $\mathcal{L}_{ar}$-théorie consistante. Alors T n'est pas décidable.*

Démonstration. Sinon, $\#\mathrm{Thm}(T)$ serait un ensemble récursif, et il existerait, par le théorème de représentabilité, une formule $\tau(x)$ représentant $\#\mathrm{Thm}(T)$.

Si $T \vdash \Delta_\tau$, c'est-à-dire si $\#\Delta_\tau \in \#\mathrm{Thm}(T)$, on a $\mathcal{P}_0 \vdash \tau(\underline{\#\Delta_\tau})$ et donc $\mathcal{P}_0 \vdash \neg\Delta_\tau$ par l'argument diagonal. Absurde.

Si $T \nvdash \Delta_\tau$, c'est-à-dire si $\#\Delta_\tau \notin \#\mathrm{Thm}(T)$, on a $\mathcal{P}_0 \vdash \neg\tau(\underline{\#\Delta_\tau})$ et donc $\mathcal{P}_0 \vdash \Delta_\tau$ par l'argument diagonal. Également absurde. □

Corollaire 5.4.5 (Indécidabilité du calcul des prédicats — Church)**.** *Il existe un langage fini $\mathcal{L}$ tel que $U = \{\#\varphi \mid \varphi$ est une $\mathcal{L}$-formule universellement valide$\}$ n'est pas récursif.*

Démonstration. Soit $\mathcal{L} = \mathcal{L}_{ar}$. Supposons U récursif. Alors, comme φ est universellement valide si et seulement si une clôture universelle de φ l'est, la $\mathcal{L}_{ar}$-théorie vide est décidable. Comme $\mathcal{P}_0$ est une théorie finie, $\mathcal{P}_0$ serait donc également décidable par l'exercice 5.2.6, contredisant le théorème de Church.

[Voici l'argument pour le transfer de la décidabilité de la théorie vide à $\mathcal{P}_0$:

On a $\varphi \in \mathrm{Thm}(\mathcal{P}_0)$ ssi $\mathcal{P}_0 \vdash \varphi$ ssi $\vdash \underbrace{\bigwedge_{i=1}^{8} A_i \to \varphi}_{\psi_\varphi}$ ssi $\psi_\varphi \in \mathrm{Thm}(\emptyset)$.] □

Remarque. *Il suffit de considérer un langage $\mathcal{L}$ qui contient un seul symbole de relation binaire, c'est-à-dire $\mathcal{L}_{ens} = \{\in\}$, pour obtenir le résultat précédent de Church.*

Par contre, les langages plus simples ne suffisent pas, comme le montre l'exercice suivant.

Exercice 5.4.6. Soit $\mathcal{L} = \{P, c\}$, où P est un prédicat unaire et c une constante.

1. Déterminer les $\mathcal{L}$-structures finies et dénombrables à isomorphisme près.
2. En déduire : deux $\mathcal{L}$-structures $\mathfrak{M}$ et $\mathfrak{N}$ sont élémentairement équivalentes si et seulement si
   - $\mathfrak{M} \models Pc$ ssi $\mathfrak{N} \models Pc$, et
   - $\mathfrak{M} \models \exists^{\geq k} x\, Qx$ ssi $\mathfrak{N} \models \exists^{\geq k} x\, Qx$ pour tout $k \in \mathbb{N}$ et tout $Q \in \{P, \neg P\}$.
3. Montrer qu'un $\mathcal{L}$-énoncé φ est universellement valide ssi $\mathfrak{M} \models \varphi$ pour toute $\mathcal{L}$-structure finie. En déduire que la $\mathcal{L}$-théorie vide est décidable.

5.5 Les théorèmes d'incomplétude de Gödel

Théorème 5.5.1 (Premier théorème d'incomplétude de Gödel)**.** *Soit T une $\mathcal{L}_{ar}$-théorie consistante, récursive et contenant $\mathcal{P}_0$. Alors T n'est pas complète.*

Démonstration. Si T était complète, elle serait décidable par 5.2.4, contredisant le théorème de Church. □

Dans la suite, nous gardons les hypothèses du théorème 5.5.1, c'est-à-dire $T \supseteq \mathcal{P}_0$ consistante et récursive. Il existe donc un énoncé ψ tel que ni $T \vdash \psi$ ni $T \vdash \neg\psi$. Nous allons présenter une amélioration, due à Rosser, qui en fournit un exemple.

Rappel : $\mathrm{Dem}(T) = \{(\#\varphi, \#\#d) \mid d \text{ est une preuve formelle de } \varphi \text{ dans } T\}$ est récursif. Par le théorème de représentabilité il existe une formule Σ_1 $P_T(x,y)$ qui représente $\mathrm{Dem}(T) \subseteq \mathbb{N}^2$.

Soit $\mathrm{Neg} : \mathbb{N} \to \mathbb{N}$ une fonction primitive récursive telle que $\mathrm{Neg}(\#\varphi) = \#(\neg\varphi)$ et $\mathrm{Neg}(x) = 0$ si x n'est pas le code d'une formule. On choisit une fomule Σ_1 $\nu(x,y)$ représentant Neg, et on modifie P_T en

$$P_T^R(x,y) := P_T(x,y) \wedge \neg\exists z \leq y \exists u (P_T(u,z) \wedge \nu(x,u)).$$

Considérons $h_T^R(x) := \exists y P_T^R(x,y)$, puis $\Delta_T^R := \Delta_{h_T^R}$.

Théorème 5.5.2 (Variante de Rosser du permier théorème d'incomplétude)**.** *Soit $T \supseteq \mathcal{P}_0$ consistante et récursive. Alors $T \nvdash \Delta_T^R$ et $T \nvdash \neg\Delta_T^R$.*

Démonstration. On pose $\Delta := \Delta_T^R$ et $m := \#\Delta$. Par l'argument diagonal, on a donc

$$\mathcal{P}_0 \vdash \left(\exists y P_T^R(\underline{m}, y) \leftrightarrow \neg\Delta\right). \tag{5.2}$$

Supposons $T \vdash \Delta$. Alors il existe une preuve de Δ dans T, autrement dit $(m,p) \in \mathrm{Dem}(T)$ pour un $p \in \mathbb{N}$, d'où $\mathcal{P}_0 \vdash P_T(\underline{m}, \underline{p})$. Comme T est consistante, pour tout $p' \in \mathbb{N}$ on a $\mathcal{P}_0 \vdash \neg\exists u \left(\nu(\underline{m}, u) \wedge P_T(u, \underline{p'})\right)$. Il suffit de noter que $\mathcal{P}_0 \vdash (x \leq \underline{p} \leftrightarrow \bigvee_{i=0}^{p} x \doteq \underline{i})$ pour déduire que $\mathcal{P}_0 \vdash P_T^R(\underline{m}, \underline{p})$. Par (5.2), on a $\mathcal{P}_0 \vdash \neg\Delta$, d'où $T \vdash \neg\Delta$ (car $T \supseteq \mathcal{P}_0$). Contradiction.

Supposons $T \vdash \neg\Delta$. Alors il existe $p \in \mathbb{N}$ tel que $\mathcal{P}_0 \vdash P_T(\underline{\#\neg\Delta}, \underline{p})$. Comme T est consistante, pour tout $p' \in \mathbb{N}$ on a $\mathcal{P}_0 \vdash \neg P_T(\underline{m}, \underline{p'})$. Il s'en suit que $\mathcal{P}_0 \vdash \neg\exists y P_T^R(\underline{m}, y)$, puisque $\mathcal{N}$ est un segment initial de tout modèle de $\mathcal{P}_0$. Par (5.2), on a $\mathcal{P}_0 \vdash \Delta$, donc $T \vdash \Delta$. Contradiction. □

Soient $h_T(x) := \exists y P_T(x,y)$ et $\Delta_T := \Delta_{h_T}$.

Lemme 5.5.3. *Soit $T \supseteq \mathcal{P}_0$ consistante et récursive. Alors $T \nvdash \Delta_T$.*

Démonstration. La preuve est similaire à celle de 5.5.2 et laissée en exercice. □

Soit $a := \#(\neg 0 \doteq 0) \in \mathbb{N}$. Pour T une théorie récursive (quelconque), on pose $\mathrm{Coh}(T) := \neg\exists y P_T(\underline{a}, y)$.

Remarque. *On a $\mathcal{N} \models \mathrm{Coh}(T)$ si et seulement si T est une théorie consistante.* □

Théorème 5.5.4 (Second théorème d'incomplétude de Gödel — sans preuve)**.** *Soit T une théorie consistante récursive contenant l'arithmétique de Peano $\mathcal{P}$. Alors $T \cup \{\mathrm{Coh}(T)\} \vdash \Delta_T$. En particulier, $T \nvdash \mathrm{Coh}(T)$, c'est-à-dire T ne démontre pas sa propre consistance.*

Remarque 5.5.5. *1. Il suffit de demander que T contienne $\mathcal{P}_0$ plus tous les axiomes d'induction associés aux formules Σ_1, pour obtenir la conclusion du second théorème d'incomplétude.*

2. Il s'en suit qu'il n'est pas exclus que $T \vdash \neg\mathrm{Coh}(T)$. Ce sera le cas par exemple pour la théorie $T' = T \cup \{\neg\mathrm{Coh}(T)\}$ qui est consistante par le théorème et qui 'démontre' sa propre inconsistance. Or, ce n'est pas choquant puisque $\mathcal{N} \not\models T'$. En effet, la 'preuve' de $\neg 0 \doteq 0$ en question provient d'un entier non standard.)

Ce n'est pas le cas de $\mathcal{P}$, puisque $\mathcal{N} \models \mathcal{P} \cup \{\mathrm{Coh}(\mathcal{P})\}$, plus généralement pour toute théorie récursive T dont $\mathcal{N}$ est un modèle.

Idée de la preuve du second théorème d'incomplétude.
La preuve complète est très longue et compliquée. Nous nous contentons de donner les étapes principales.

- On représente $\mathrm{Dem}(T)$ et $\mathrm{Dem}(\mathcal{P}_0)$ par des formules $P_T(x,y)$ et $P_{\mathcal{P}_0}$ de manière naturelle, en demandant
  - $\vdash P_{\mathcal{P}_0} \to P_T$;
  - $\vdash \Big(\exists y P_T(\underline{\#\varphi}, y) \wedge \exists y P_T(\underline{\#(\varphi \to \psi)}, y)\Big) \to \exists y P_T(\underline{\#\psi}, y)$, de même pour $\exists$-intro au lieu du modus ponens et pour $P_{\mathcal{P}_0}$ au lieu de P_T.
- Nous avons vu dans la proposition 5.3.4 que pour tout énoncé Σ_1 φ, on a $\mathcal{N} \models \varphi \to \exists y P_{\mathcal{P}_0}(\underline{\#\varphi}, y)$. On considère alors la théorie

$$\mathcal{P}_1 := \mathcal{P}_0 \cup \{\varphi \to \exists y P_{\mathcal{P}_0}(\underline{\#\varphi}, y) \mid \varphi \text{ énoncé } \Sigma_1\}.$$

  La théorie $\mathcal{P}_1$ est récursive, avec $\mathcal{N} \models \mathcal{P}_1$.
- On montre (cette étape n'est pas très dificile) : Soit T consistante et récursive avec $T \vdash \mathcal{P}_1$, c'est-à-dire $T \vdash \psi$ pour tout $\psi \in \mathcal{P}_1$. Alors on a $T \cup \{\mathrm{Coh}(T)\} \vdash \Delta_T$, et donc en particulier, $T \nvdash \mathrm{Coh}(T)$.
- On montre : $\mathcal{P} \vdash \mathcal{P}_1$.
  C'est dans la preuve de cet énoncé où se trouve toute la difficulté. À la base, l'idée est très simple. On montre $\mathcal{P} \vdash \mathcal{P}_1$ 'comme on a montré $\mathcal{N} \models \mathcal{P}_1$', en vérifiant à chaque étape que l'on peut tout faire en ne raisonnant que dans l'arithmétique de Peano. Entre autre, il faut montrer une sorte de théorème de complétude dans Peano. Cela demande un travail considérable de codage. □

Chapitre 6

Théorie axiomatique des ensembles

Dans ce dernier chapitre du cours, nous formalisons la théorie des ensembles à l'aide des outils que nous avons développés, c'est-à-dire au premier ordre. Nous travaillons dans le langage $\mathcal{L} = \mathcal{L}_{ens} = \{\in\}$, et nous considérons des $\mathcal{L}_{ens}$-structures $\mathcal{U} = \langle U; \in \rangle$ ($\mathcal{U}$ comme *univers d'ensembles*). Nous appelons *ensemble* tout élément de U. L'univers U, quant à lui, est un ensemble au sens naïf. L'appartenance entre deux ensembles est donnée par le symbole de relation $\in$.

On peut alors former le langage $\mathcal{L}_U$ (un ensemble au sens naïf).

Définition. Une partie naïve D de U est une *classe* s'il existe une $\mathcal{L}_U$-formule $\varphi = \varphi(x)$ telle que $D = \varphi[\mathcal{U}]$.

Notons que tout ensemble a définit une classe C_a, donnée par $\varphi(x) := x \in a$. Par l'axiome d'extensionalité, deux ensembles a et b sont égaux si $C_a = C_b$.

Par abus de notation, nous utilisons parfois $\in$ pour désigner l'appartenance (au sens naïf) à une classe, par exemple quand nous écrivons $c \in D$ au lieu de $\mathcal{U} \models \varphi[c]$, si D est définie par φ. Cela ne devrait pas créer de confusion.

6.1 Les axiomes de Zermelo-Fraenkel

Définition.

1. Le *système d'axiomes de Zermelo-Fraenkel* ZF est donné par les axiomes suivants : Axiome d'extensionalité, schéma d'axiomes de compréhension, axiome de la paire, axiome de la réunion, axiome des parties, schéma d'axiomes de remplacement, axiome de fondation, axiome de l'infini.
2. Si on ajoute l'axiome du choix (AC) au système ZF, on on obtient le *système d'axiomes* ZFC.

Remarque. *L'axiome de fondation n'est pas toujours inclus dans* ZF.

Dans ce qui suit, nous discutons les axiomes en détail.

Axiome d'extensionalité (Ext) $\forall x, y\,(\forall z(z \in x \leftrightarrow z \in y) \rightarrow x \doteq y)$

Il exprime que deux ensembles ayant les mêmes éléments sont égaux.

Notation. Nous écrivons $x \subseteq y$ comme abbréviation de $\forall z(z \in x \rightarrow z \in y)$.

Ainsi, (Ext) équivaut à $\forall x, y\,((x \subseteq y \wedge y \subseteq x) \rightarrow x \doteq y)$.

On observera que les cinq (schémas d') axiomes suivants sont des cas particuliers de la compréhension globale.

Schéma d'axiomes de compréhension (Com)

Pour toute $\mathcal{L}_{ens}$-formule $\varphi = \varphi(v_0, \ldots, v_n)$, un axiome de la forme

$$\forall v_1, \ldots, v_{n+1} \exists v_{n+2} \forall v_0\,(v_0 \in v_{n+2} \leftrightarrow (v_0 \in v_{n+1} \wedge \varphi(v_0, \ldots, v_n)))\,.$$

Il exprime : si a est un ensemble et D une classe, alors la classe $\{b \in a \mid b \in D\}$ est un ensemble.

Voici quelques conséquences de ces axiomes :

- Si a et b sont deux ensembles, on peut former leur intersection $a \cap b = \{c \in a \mid c \in b\}$ ainsi que leur différence $a \setminus b = \{c \in a \mid c \notin b\}$.
- On obtient l'existence de l'ensemble vide, noté $\emptyset$. En effet, soit a un ensemble arbitraire. Alors $\emptyset = \{c \in a \mid \neg c \doteq c\}$. L'unicité de $\emptyset$ suit de (Ext).
- Si $a \neq \emptyset$ est un ensemble, on obtient l'existence de $\bigcap_{b \in a} b$. Il suffit de choisir $b_0 \in a$, puis de remarquer que $\bigcap_{b \in a} b = \{c \in b_0 \mid \forall x(x \in a \rightarrow c \in x)\}$.

Remarque 6.1.1. *L'antinomie de Russel devient un théorème dans* ZFC. *En effet, on a* (Com) $\vdash \neg\exists x \forall z\, z \in x$. *Sinon,* $\{z \mid \neg z \in z\}$ *serait également un ensemble, ce qui est absurde.*

L'axiome de la paire (Paire) $\forall y_1, y_2 \exists x \forall z\,(z \in x \leftrightarrow (z \doteq y_1 \vee z \doteq y_2))$

Il exprime que si a et b sont deux ensembles, alors $\{a, b\}$ est un ensemble.

Définition. La *paire ordonnée* (appelée aussi *paire de Kuratowski*) de deux ensembles a et b est l'ensemble $(a, b) := \{\{a\}, \{a, b\}\}$.

Lemme 6.1.2. *Avec les axiomes que nous avons énoncés jusqu'à maintenant, on a :*

1. *Il y a une formule qui définit la classe des paires ordonnées. De plus, il existe une* $\mathcal{L}_{ens}$*-formule* $\varphi(x, y)$ *telle que l'on ait* $\mathcal{U} \models \varphi[a, b]$ *ssi* a *est une paire de Kuratowski de la forme* $a = (b, c)$. *Même chose pour la deuxième composante.*
2. *On a* $(b, c) = (b', c')$ *si et seulement si* $b = b'$ *et* $c = c'$.

Démonstration. Exercice. □

L'axiome de la réunion ($\bigcup$) $\forall y \exists x \forall z\, (z \in x \leftrightarrow \exists w(z \in w \wedge w \in y))$

Il exprime que pour tout ensemble a, la classe $\bigcup a := \{z \mid \exists w(z \in w \wedge w \in a\}$ est un ensemble.

Notons qu'en combinant (Paire) et ($\bigcup$), on peut montrer que la réunion $a \cup b$ de deux ensembles a et b est un ensemble.

L'axiome des parties (Parties) $\forall y \exists x \forall z (z \in x \leftrightarrow z \subseteq y)$

Il postule l'existence de l'ensemble des parties de a : $\mathcal{P}(a) = \{b \mid b \subseteq a\}$.

Lemme 6.1.3. *Les axiomes énoncés jusqu'à maintenant entraînent l'existence du produit cartésien de deux ensembles a et b : $a \times b = \{(x, y) \mid x \in a \wedge y \in b\}$ est un ensemble.*

Démonstration. Si $x \in a$ et $y \in b$, on a $\{x\}, \{x, y\} \in \mathcal{P}(a \cup b)$, d'où $(x, y) \in \mathcal{P}(\mathcal{P}(a \cup b))$. On conclut par (Com), en utilisant le lemme 6.1.2. □

On peut également définir des triplets $(x, y, z) := ((x, y), z)$, puis inductivement des n-uplets, via $(x_1, \ldots, x_{n+1}) := ((x_1, \ldots, x_n), x_{n+1})$. On obtient alors $a \times b \times c$, et plus généralement $a_1 \times \cdots \times a_n$.

Définition.
- Une *relation (binaire)* R est un ensemble de paires ordonnées. On pose $\mathrm{dom}(R) := \{x \mid \exists y(x, y) \in R\}$ et $\mathrm{im}(R) := \{y \mid \exists x(x, y) \in R\}$.
- Une *fonction* f est une relation qui est unique à droite, c'est-à-dire on a $\forall x, y_1, y_2\, ((x, y_1) \in f \wedge (x, y_2) \in f \to y_1 \doteq y_2)$.

Remarque.
- *On note que* $\mathrm{dom}(R)$ *et* $\mathrm{im}(R)$ *sont bien des ensembles, car si* $(x, y) \in R$*, alors* $x, y \in \bigcup\bigcup R$.
- *Par définition, nous identifions une fonction avec son graphe.*

On écrira $f(x) = y$ au lieu de $(x, y) \in f$. Parfois, pour $x \notin \mathrm{dom}(f)$, nous posons $f(x) := \emptyset$. Si $a = \mathrm{dom}(f)$ et $\mathrm{im}(f) \subseteq b$, on écrit $f : a \to b$.

Lemme 6.1.4. *Soit a et b deux ensembles. Alors, avec les axiomes que nous avons énoncés, on obtient :*

1. $\{R \mid R$ *est une relation avec* $\mathrm{dom}(R) \subseteq a$ *et* $\mathrm{im}(f) \subseteq b\}$ *est un ensemble.*
2. $\{f \mid f : a \to b\}$ *est un ensemble.*
3. *Les fonctions forment une classe. On note* $\mathrm{Fn}(x)$ *une formule qui définit cette classe.*

Démonstration. Exercice. □

Remarque. *Si $I \neq \emptyset$ est un ensemble et $(a_i)_{i \in I}$ est une* famille d'ensembles*, c'est-à-dire donnée par une fonction f avec* $\mathrm{dom}(f) = I$*, alors la classe* $\prod_{i \in I} a_i = \{g : I \to \bigcup_{i \in I} a_i \mid \forall z(z \in I \to g(z) \in z)\}$ *est un ensemble.* □

Définition. On dit que $F \subseteq U^2$ est une *classe fonctionnelle* s'il existe une $\mathcal{L}_U$-formule $\varphi(x, y)$ qui définit F telle que

$$\mathcal{U} \models \forall x, y_1, y_2\, (\varphi(x, y_1) \wedge \varphi(x, y_2) \to y_1 \doteq y_2)\,.$$

La classe $\mathrm{Dom}(F)$ définie par $\exists y \varphi$ est appelée le *domaine de F*.

Notons qu'à une classe fonctionnelle correspond une fonction naïve (dont la classe fonctionnelle est le graphe et avec domaine la classe $\mathrm{dom}(F)$).

Schéma d'axiomes de remplacement (Rem)

Pour toute $\mathcal{L}_{ens}$-formule $\varphi = \varphi(x, y, v_1, \ldots, v_n)$, un axiome de la forme

$$\forall v_0, v_1, \ldots, v_n[\forall x, y_1, y_2(\varphi(x, y_1, \overline{v}) \wedge \varphi(x, y_2, \overline{v}) \rightarrow y_1 \dot{=} y_2) \\ \rightarrow \exists v_{n+1} \forall y(y \in v_{n+1} \leftrightarrow \exists x(x \in v_0 \wedge \varphi(x, y, \overline{v}))].$$

Il exprime que si $F \subseteq U^2$ est une classe fonctionnelle et a un ensemble, alors $F[a] := \{z \mid \exists u(u \in a \wedge (u, z) \in F)\}$ est un ensemble. (Il est obtenu en 'remplaçant' tout élément de a par son image sous F.)

Ces axiomes ne sont pas indépendants. En effet :

Lemme 6.1.5. *1. Le schéma (Rem) implique le schéma (Com).*

2. (Paire) est une conséquence des autres axiomes.

Démonstration. Soit $\varphi(x)$ donnée. On pose $F(x, y) := (x \dot{=} y \wedge \varphi(x))$. Alors $\{x \in a \mid \varphi(x)\} = F[a]$, d'où (1).

Quant à (2), on se donne deux ensembles a et b. On remarque que $\emptyset \in \mathcal{P}(\emptyset) = \{\emptyset\}$, en particulier $\emptyset \neq \mathcal{P}(\emptyset)$. On pose $F(x, y) := ((x \dot{=} \emptyset \wedge y \dot{=} a) \vee (x \dot{=} \mathcal{P}(\emptyset) \wedge y \dot{=} b))$. C'est bien une relation fonctionnelle. Pour $c = \mathcal{P}(\mathcal{P}(\emptyset)) = \{\emptyset, \mathcal{P}(\emptyset)\}$, on a alors $F[c] = \{a, b\}$. □

Remarque. *Nous pouvons travailler dans une expansion par définition et utiliser par exemple des symboles de relation $x \subseteq y$ (binaire) et $f : x \rightarrow y$ (ternaire), des symboles de fonction $x \cup y$, $x \cap y$, $x \setminus y$, $\bigcup y$, $\mathcal{P}(y)$, $\{x, y\}$, (x, y), $x \times y$, $\mathrm{dom}(R)$, $\mathrm{im}(R)$, la constante $\emptyset$ etc. dans les schémas d'axiomes (Com) et (Rem). Par la proposition 3.3.2, cela ne change rien.*

Rappelons que x est un *ordinal* si x est un ensemble transitif tel que la relation d'appartenance $\in\restriction_{x \times x}$ définit un bon ordre sur x, c'est-à-dire un ordre total bien-fondé. (Cette dernière propriété se dit ainsi : $\forall y(y \in \mathcal{P}(x) \wedge \neg y \dot{=} \emptyset \rightarrow \exists z(z \in y \wedge z \cap y \dot{=} \emptyset))$.) On note $\mathrm{Ord}(x)$ une formule qui définit la classe des ordinaux. De même, on peut définir la classe des cardinaux : $\mathrm{Card}(x) := \mathrm{Ord}(x) \wedge \forall y(y \in x \rightarrow \neg \exists f : y \rightarrow x \text{ surjective})$

Notation. Dans la suite, nous utilisons α, β, γ etc. uniquement pour des ordinaux, et nous écrivons par exemple $\forall \gamma \varphi$ au lieu de $\forall \gamma(\mathrm{Ord}(\gamma) \rightarrow \varphi)$.

L'axiome de fondation (AF) $\forall x(\neg x \dot{=} \emptyset \rightarrow \exists z(z \in x \wedge z \cap x \dot{=} \emptyset))$

Notons que (AF), avec l'axiome de la paire, entraîne que pour tout ensemble x on a $x \notin x$. En effet, sinon $\{x\}$ contredirait (AF).

L'axiome de l'infini (AI) $\exists x(\emptyset \in x \wedge \forall z(z \in x \rightarrow z \cup \{z\} \in x))$

Il postule l'existence d'un ensemble contenant $\emptyset$ et clos par 'successeur'. De manière équivalente, on pourrait postuler l'existence d'un ordinal infini (ou de ω).

On note ω le plus petit ordinal limite, et $\mathrm{Lim}(x)$ une formule qui définit la classe des ordinaux limites.

Remarque 6.1.6. *1. Sous* ZF*, les ordinaux vérifient les mêmes propriétés que celles vues au premier chapitre du cours. De même les cardinaux, sous* ZFC*.*

2. La classe des ordinaux n'est pas un ensemble (pareil pour les cardinaux). [Si $a = \mathrm{Ord}$, alors a serait transitif et bien-ordonné par $\in$, donc un ordinal et alors $a \in a$. Or $a \notin a$ pour tout ordinal.]

Lemme 6.1.7 (Induction transfinie)**.** *Soit $\mathcal{U} \models$ ZF et soit $\varphi(x)$ une $\mathcal{L}_U$-formule. Alors $\mathcal{U}$ satisfait la propriété d'induction suivante :*

$$\varphi(\emptyset) \wedge \forall\gamma(\varphi(\gamma) \to \varphi(\gamma \cup \{\gamma\}) \wedge \forall\gamma([\mathrm{Lim}(\gamma) \wedge \forall\delta(\delta \in \gamma \to \varphi(\delta))] \to \varphi(\gamma)) \to \forall\gamma\varphi(\gamma).$$

Démonstration. Si $\mathcal{U} \models \varphi(\emptyset) \wedge \forall\gamma(\varphi(\gamma) \to \varphi(\gamma \cup \{\gamma\}) \wedge \forall\gamma([\mathrm{Lim}(\gamma) \wedge \forall\delta(\delta \in \gamma \to \varphi(\delta))] \to \varphi(\gamma))$ et $\mathcal{U} \models \neg\varphi[\alpha]$ pour un ordinal α, il suffit de choisir un tel α minimal pour arriver à une contradiction. □

On peut également considérer des classes fonctionnelles F qui correspondent à des fonctions naïves entre une classe $D \subseteq U^n$ et U.

Théorème 6.1.8 (Définition par récurrence transfinie)**.** *Soit G une classe fonctionnelle de domaine U^{n+1}. Alors il existe une unique classe fonctionnelle F de domaine $U^n \times \mathrm{Ord}$ telle que pour tout uplet d'ensembles $\overline{w}$ et tout ordinal α on ait $F(\overline{w}, \alpha) = G(\overline{w}, F \restriction_{\{\overline{w}\} \times \alpha})$.*

Démonstration. L'idée est d'approximer F par des fonctions. On montre d'abord que pour tout $\overline{w}$ et tout ordinal β il existe une unique fonction $f_{\overline{w},\beta}$ de domaine β telle que

$$f_{\overline{w},\beta}(\alpha) = G(\overline{w}, \{\overline{w}\} \times f_{\overline{w},\beta} \restriction_\alpha) \text{ pour tout } \alpha < \beta.$$

L'unicité est claire par induction (transfinie).

Quant à l'existence, on montrera par induction sur β :

$$\forall\overline{w}\exists f_{\overline{w},\beta} \text{ fonction de domaine } \beta : \underbrace{\forall\alpha\,(\alpha < \beta \to f_{\overline{w},\beta}(\alpha) \doteq G(\overline{w}, \{\overline{w}\} \times f_{\overline{w},\beta} \restriction_\alpha))}_{(*)_{\overline{w},\beta}}$$

Si $\beta = 0$, on pose $f_{\overline{w},\beta} = \emptyset$.

Soit $\beta = \beta' + 1$ et $f_{\overline{w},\beta'}$ une fonction satisfaisant à $(*)_{\overline{w},\beta'}$.
On pose $f_{\overline{w},\beta} = f_{\overline{w},\beta'} \cup \{(\beta', G(\overline{w}, \{\overline{w}\} \times f_{\overline{w},\beta'})\}$.

Soit β un ordinal limite. On considère l'ensemble suivant :
$X_{\overline{w},\beta} := \{f \mid \exists\beta' < \beta : f \text{ est une fonction satisfaisant } (*)_{\overline{w},\beta'} \text{ t.q. } \mathrm{dom}(f) = \beta'\}$.
Par (Rem) et unicité, $X_{\overline{w},\beta}$ est un ensemble. Également par unicité, $\bigcup_{f' \in X_{\overline{w},\beta}} f'$ est donc une fonction qui convient.

On pose $(\overline{w}, \beta, y) \in F :\Leftrightarrow$ pour toute fonction f de domaine $\beta+1$ satisfaisant à $(*)_{\overline{w},\beta+1}$ on a $f(\beta) = y$. □

Exemples 6.1.9 (Applications de la définition par récurrence transfinie)**.**

1. Les fonctions de l'arithmétique ordinale (addition, multiplication et exponentiation ordinale) peuvent se définir par récurrence (voir la remarque 1.5.7).
2. La hiérarchie des $\aleph$ est donnée par une classe fonctionnelle $\aleph : \mathrm{Ord} \to \mathrm{Card}$. En effet, il suffit d'appliquer la récurrence transfinie à la classe fonctionnelle $G : U \to U$ définie comme suit :
   – $G(0) = \omega$;
   – si $f : \alpha \to \beta$ pour deux ordinaux α, β, alors
   $$G(f) = \begin{cases} \text{le plus petit cardinal} > f(\alpha'), \text{ si } \alpha = \alpha' + 1; \\ \bigcup \mathrm{im}(f), \text{ si } \mathrm{dom}(f) \text{ est un ordinal limite.} \end{cases}$$
   – si x n'est pas une fonction entre deux ordinaux, alors $G(x) = \emptyset$.
3. La *hiérarchie de von Neumann* : par récurrence sur $\alpha \in \mathrm{Ord}$, on définit une classe fonctionnelle $\alpha \mapsto V_\alpha$ via
   – $V_0 = \emptyset$;
   – $V_{\alpha+1} = \mathcal{P}(V_\alpha)$;
   – $V_\lambda = \bigcup_{\alpha<\lambda} V_\alpha$ pour λ limite.

Proposition 6.1.10 (ZF − (AF))**.** *L'ordinal ω muni des opérations ordinales (addition, multiplication et successeur), avec $0 = \emptyset$ et $<$ donné par $\in$ est un modèle de $\mathcal{P}$.*

Démonstration. Exercice. □

6.2 Axiome du choix

L'axiome du choix (AC)

$$\forall f[(\mathrm{Fn}(f) \wedge \emptyset \notin \mathrm{im}(f)) \\ \to \exists g\,(\mathrm{Fn}(g) \wedge \mathrm{dom}(g) \doteq \mathrm{dom}(f) \wedge \forall x(x \in \mathrm{dom}(g) \to g(x) \in f(x)))]$$

Il exprime que le produit d'une famille d'ensembles non vides est non vide.

Définition. Soit a un ensemble. On appelle *fonction de choix* sur a une fonction $h : \mathcal{P}(a)' := \mathcal{P}(a) \setminus \{\emptyset\} \to a$ telle que $h(A) \in A$ pour tout $A \in \mathcal{P}(a)'$.

Proposition 6.2.1. *(AC) est équivalent à l'existence d'une fonction de choix pour tout ensemble a.*

Démonstration. Supposons (AC) et soit a un ensemble. Par l'axiome du choix, $\prod_{A \in \mathcal{P}(a)'} A \neq \emptyset$. Tout élément de ce produit est une fonction de choix sur a.

Réciproquement, soit $(a_i)_{i \in I}$ une famille d'ensembles avec $a_i \neq \emptyset$ pour tout $i \in I$. Soit $a = \bigcup_{i \in I} a_i$ et $h : \mathcal{P}(a)' \to a$ une fonction de choix sur a. Alors $h \circ f \in \prod_{i \in I} a_i$, où $f : I \to \mathcal{P}(a)'$, $f(i) := a_i$. □

Théorème 6.2.2. *Dans ZF, son équivalents :*

1. *(AC) ;*
2. *le lemme de Zorn ;*
3. *le théorème de Zermelo (Wohlordnungssatz).*

Démonstration. (1)⇒(2) : Par l'absurde. Soit donc $\langle X, <\rangle$ un ordre partiel inductif sans élément maximal. On considère l'ensemble

$$\mathcal{T} = \{T \in \mathcal{P}(X) \mid T \text{ est totalement ordonné par } <\}.$$

Comme il n'existe pas d'élément maximal dans X, pour tout $T \in \mathcal{T}$, l'ensemble $B(T) := \{x \in X \mid x > t \ \forall t \in T\}$ est non vide. Par (AC), il existe une fonction $b : \mathcal{T} \to X$ telle que $b(T) \in B(T)$ pour tout $T \in \mathcal{T}$.

Soit G la classe fonctionnelle suivante :
- si g est une fonction avec $\mathrm{dom}(g) \in \mathrm{Ord}$ et $\mathrm{im}(g) \in \mathcal{T}$, alors $G(g) = b(\mathrm{im}(g))$;
- $G(g) = \emptyset$ sinon.

Par le théorème 6.1.8 il existe une classe fonctionnelle F de domanie la classe des ordinaux telle que $F(\alpha) = G(F \restriction_\alpha)$ pour tout ordinal α. On vérifie par induction que $F(\alpha) \in X$ pour tout α et

$$\alpha < \beta \Rightarrow F(\alpha) < F(\beta). \tag{6.1}$$

Par compréhension, $O_X := \{x \in X \mid$ il existe un ordinal α tel que $(\alpha, x) \in F\}$ est un ensemble.

En tant que fonction naïve, F est injective par (6.1). La classe F^{-1} définie par $(x, \alpha) \in F^{-1} :\Leftrightarrow (\alpha, x) \in F$ est donc une classe fonctionnelle. On a $F^{-1}[O_X] = \mathrm{Ord}$, et Ord est un ensemble par remplacement. Ceci contredit la remarque 6.1.6(2).

(2)⇒(3) : Soit a un ensemble. Il faut montrer que a admet un bon ordre. On considère $X := \{(b, R) \mid b \in \mathcal{P}(a)$ et R est un bon ordre sur $b\}$. Il est facile à voir que X est un ensemble (exercice).

Sur X, on définit un ordre partiel comme suit :

$$(b, R) \leq (b', R') \ :\Leftrightarrow \ b \text{ est un segment initial de } (b', R') \text{ et } R' \restriction_b = R\}.$$

Cet ordre est inductif. En effet, si $(b_i, R_i)_{i \in I}$ est une partie totalement ordonnée de X, alors $(b, R) := (\bigcup_{i \in I} b_i, \bigcup_{i \in I} R_i)$ est un majorant de cette partie. Par le lemme de Zorn, il existe un élément $(\tilde{b}, \tilde{R}) \in X$ qui est maximal pour $\leq$. Si $\tilde{b} \neq a$, il existe $y \in a \setminus \tilde{b}$. On pose $b' = \tilde{b} \cup \{y\}$ et $R' := \tilde{R} \cup \{(x, y) \mid x \in \tilde{b}\}$. Il est alors clair que R' est un bon ordre sur b' qui prolonge celui sur $\tilde{b}$, ce qui contredit la maximalité de $(\tilde{b}, \tilde{R})$.

(3)⇒(1) : Soit a un ensemble. Par hypothèse, a peut être bien-ordonné, disons par $<$. La fonction $f : \mathcal{P}(a)' \to a$ qui à une partie non vide de a associe son plus petit élément est une fonction de choix sur a. On conclut par la proposition 6.2.1. □

Remarque 6.2.3 (Paradoxe de Skolem)**.** *Si ZFC est consistante, elle a un modèle (nécesairement infini, vu les axiomes). Comme $\mathcal{L}_{ens}$ est dénombrable, ZFC a donc un modèle dénombrable $\mathfrak{M}$ par le théorème de Löwenheim-Skolem descendant. Mais il existe des ensembles non-dénombrables dans $\mathfrak{M}$, par exemple $\aleph_1$ ou $\mathbb{R}$.*

Explication : *La notion de 'dénombrabilité' dépend du modèle de ZFC dans lequel on travaille. Elle est donc relative. Être dénombrable au sens de $\mathfrak{M}$ et l'être au sens du modèle de la théorie des ensembles (naïf) de fond n'est pas la même chose. Qui plus est, l'ensemble de base M de $\mathfrak{M}$ est un un ensemble du point de vue naïf, et une classe propre du point de vue de $\mathfrak{M}$.*

Plus généralement, la notion de cardinalité dépend du modèle. Ainsi, l'ensemble des entiers $\mathbb{N}$ (au sens de $\mathfrak{M}$) et l'ensemble des réels $\mathbb{R}$ (au sens de $\mathfrak{M}$) sont tous les deux dénombrables d'un point de vue du modèle de fond. Or, la bijection entre les deux ensembles qui existe comme ensemble naïf n'est même pas représenté par une classe dans $\mathfrak{M}$. En effet, sinon, par remplacement, elle serait donnée par un ensemble au sens de $\mathfrak{M}$, c'est-à-dire par un élément de M.

6.3 La hiérarchie de von Neumann et l'axiome de fondation

On note $\mathrm{ZF}^- := \mathrm{ZF} - (\mathrm{AF})$ et $\mathrm{ZFC}^- := \mathrm{ZFC} - (\mathrm{AF})$. Dans cette partie, nous supposons que $\mathcal{U} \models \mathrm{ZF}^-$.

Définition. Soit a un ensemble. Par récurrence sur $n \in \omega$, on définit $a_0 := a$, $a_{n+1} := a_n \cup \bigcup a_n$. Puis, on pose $\mathrm{ct}(a) := \bigcup_{n\in\omega} a_n$, appelée la *clôture transitive* de a.

Lemme 6.3.1 (ZF^-)**.** *L'ensemble $\mathrm{ct}(a)$ est le plus petit ensemble transitif contenant a comme partie. Plus précisément, on a les propriétés suivantes :*

1. $a \subseteq \mathrm{ct}(a)$
2. *$\mathrm{ct}(a)$ est un ensemble transitif.*
3. *Si $a \subseteq t$ avec t un ensemble transitif, alors $\mathrm{ct}(a) \subseteq t$. En particulier, $a \subseteq b \Rightarrow \mathrm{ct}(a) \subseteq \mathrm{ct}(b)$.*
4. *Si a est transitif, alors $\mathrm{ct}(a) = a$.*
5. *Si $b \in a$, alors $\mathrm{ct}(b) \subseteq \mathrm{ct}(a)$.*
6. $\mathrm{ct}(a) = a \cup \bigcup_{b\in a} \mathrm{ct}(b)$.

Démonstration. (1) et (2) sont clairs. Quant à (3), on montre par induction que $a_n \subseteq t$ pour tout $n \in \omega$. (4) est une conséquence de (1-3)

Preuve de (5) : On a $b \in a \Rightarrow b \in \mathrm{ct}(a) \Rightarrow b \subseteq \mathrm{ct}(a) \Rightarrow \mathrm{ct}(b) \subseteq \mathrm{ct}(a)$. (La première implication suit de (1), la seconde de (2), et la dernière de (3).)

Preuve de (6) : On a $\mathrm{ct}(a) \supseteq a \cup \bigcup_{b\in a} \mathrm{ct}(b)$ par (1) et (5). Pour l'autre inclusion, par (3), il suffit de montrer que $a \cup \bigcup_{b\in a} \mathrm{ct}(b)$ est transitif, ce qui est clair. □

Rappelons la hiérarchie de von Neumann : $V_0 := \emptyset$, $V_{\alpha+1} := \mathcal{P}(V_\alpha)$, $V_\lambda := \bigcup_{\alpha<\lambda} V_\alpha$ (pour λ limite).

On définit la classe V par la formule $\exists \alpha\, x \in V_\alpha$. Informellement, on a donc "$V = \bigcup_{\alpha \in \mathrm{Ord}} V_\alpha$".

Définition. Soit a un ensemble. On pose

$$\mathrm{Rg}(a) := \begin{cases} \text{le plus petit ordinal } \gamma \text{ tel que } a \in V_{\gamma+1}, \text{ si un tel } \gamma \text{ existe;} \\ \infty, \text{ sinon.} \end{cases}$$

On l'appelle le *rang* de a.

Lemme 6.3.2. *1. V_α est un ensemble transitif pour tout ordinal α.*

2. $\beta \leq \alpha \Rightarrow V_\beta \subseteq V_\alpha$

3. $V_\alpha = \{x \in V \mid \mathrm{Rg}(x) < \alpha\}$

4. Si $x \in V$ et $y \in x$, alors $y \in V$ et $\mathrm{Rg}(y) < \mathrm{Rg}(x)$.

5. Si $x \in V$, alors $\mathrm{Rg}(x) = \sup\{\mathrm{Rg}(y) + 1 \mid y \in x\}$.

6. Si $x \in V$ est transitif, alors $\{\mathrm{Rg}(y) \mid y \in x\}$ est un ordinal α.

7. On a $\mathrm{Rg}(\alpha) = \alpha$ pour tout α. En particulier, $\alpha \in V$ et $V_\alpha \cap \mathrm{Ord} = \alpha$.

8. Soit x un ensemble. Alors $x \in V$ si et seulement si $x \subseteq V$.

9. On suppose (AC). Si $x \in V$ est transitif, on a $x \in V_{\mathrm{card}(x)^+}$.

Démonstration. (1) & (2) Par induction transfinie sur α, on montre que V_α est transtif et que $V_\beta \subseteq V_\alpha$ pour tout $\beta \leq \alpha$. Les cas $\alpha = 0$ et α limite sont clairs. Soit donc $\alpha = \gamma + 1$. Comme V_γ est transitif par hypothèse d'induction, $\mathcal{P}(V_\gamma) = V_\alpha$ est transitif aussi. Si $\beta < \alpha$, alors $\beta \leq \gamma$ et inductiviement $V_\beta \subseteq V_\gamma$, d'où $V_\beta \in V_\alpha$ et enfin $V_\beta \subseteq V_\alpha$ par transitivité.

(3) Si $x \in V$, on a $\mathrm{Rg}(x) < \alpha$ ssi $\exists \beta < \alpha : x \in V_{\beta+1}$ ssi $x \in V_\alpha$.

(4) Soit $\alpha = \mathrm{Rg}(x)$. Alors $x \in V_{\alpha+1} = \mathcal{P}(V_\alpha)$. Pour $y \in x$, on obtient $y \in V_\alpha$, d'où $y \in V$ et $\mathrm{Rg}(y) < \alpha$ (par (3)).

(5) Si $x \in V$, on pose $\alpha = \sup\{\mathrm{Rg}(y) + 1 \mid y \in x\}$. Par (4), on a $\alpha \leq \mathrm{Rg}(x)$. Comme $\mathrm{Rg}(y) < \alpha$ pour tout $y \in x$, (3) entraîne que $x \subseteq V_\alpha$, d'où $x \in V_{\alpha+1}$ et alors $\alpha \geq \mathrm{Rg}(x)$ par définition.

(6) Soit $x \in V$ transitif et $\beta < \mathrm{Rg}(x)$ donné. Par (5) il existe $y \in x$ avec $\beta \leq \mathrm{Rg}(y)$. On choisit un tel y avec $\mathrm{Rg}(y)$ minimal. Si $z \in y$, on a $z \in x$ (par transitivité de x) et $\mathrm{Rg}(z) < \mathrm{Rg}(y)$ par (4), d'où $\mathrm{Rg}(z) < \beta$ par minimalité de $\mathrm{Rg}(y)$. Cela montre que $y \subseteq V_\beta$ et alors $y \in V_{\beta+1}$, d'où $\mathrm{Rg}(y) \leq \beta$.

(7) On démontre $\alpha \in V$ et $\mathrm{Rg}(\alpha) = \alpha$ par induction transfinie, le cas $\alpha = 0$ étant clair. Supposons le résultat vrai pour tout $\beta < \alpha$. Alors $\beta \in V_{\beta+1} \subseteq V_\alpha$ pour tout $\beta < \alpha$, d'où $\alpha \subseteq V_\alpha$ et donc $\alpha \in V_{\alpha+1}$. Par (5), on a $\mathrm{Rg}(\alpha) = \sup\{\beta + 1 \mid \beta < \alpha\} = \alpha$.

(8) $x \in V \Rightarrow x \subseteq V$ est une conséquence de (4). Réciproquement, si x est un ensemble tel que $x \subseteq V$, alors $\{\mathrm{Rg}(y) + 1 \mid y \in x\}$ est un ensemble par remplacement. Pour $\alpha = \sup\{\mathrm{Rg}(y) + 1 \mid y \in x\}$ on a donc $x \subseteq V_\alpha$ et enfin $x \in V_{\alpha+1}$.

(9) est une conséquence de (6). □

Théorème 6.3.3. *Soit $\mathcal{U} \models \mathrm{ZF}^- = \mathrm{ZF} - (\mathrm{AF})$. Sont équivalents :*

1. $\mathcal{U} \models (\mathrm{AF})$
2. $\mathcal{U} \models \forall x V(x)$

Démonstration. (2)⇒(1) : Soit x un ensemble non vide. Il faut trouver un élément y de x tel que $y \cap x = \emptyset$. Par hypothèse, tout ensemble est dans V. On choisit $y \in x$ avec $\mathrm{Rg}(y)$ minimal. Alors $y \cap x = \emptyset$ par le lemme 6.3.2(4).

(1)⇒(2) : Soit x un ensemble donné. On pose $y := \mathrm{ct}(x)$, et on considère l'ensemble $z := \{t \in y \mid t \notin V\}$. On a $x \subseteq y$. Par 6.3.2(8), pour montrer que $x \in V$, il suffit donc de montrer que tout élément de y est dans V, autrement dit que $z = \emptyset$. Si z n'était pas vide, par (AF) il existerait $t \in z$ avec $t \cap z = \emptyset$. Considérons un tel t. Pour $u \in t$, on a $u \in y$ (car y est transitif) et donc $u \in V$ (car $t \cap z = \emptyset$). Mais alors $t \in V$ par 6.3.2(8). Contradiction. □

Remarque. *Comme les éléments de V sont construits à partir de l'ensemble vide (par un procédé transfini), le théorème 6.3.3 exprime que (AF) est équivalent au fait que tout ensemble est construit à partir de l'ensemble vide.*

Remarque 6.3.4. *Dans ZFC^-, sont équivalents :*

1. *(AF)*
2. *Il n'existe pas de suite $(a_i)_{i\in\omega}$ telle que $a_{i+1} \in a_i$ pour tout $i \in \omega$.*

Démonstration. (1)⇒(2) : (Cette implication n'utilise pas (AC).) Si $(a_i)_{i\in\omega}$ est une suite d'ensembles, on considère $a = \{a_i \mid i \in \omega\}$. Par (AF), il existe $b \in a$ tel que $b \cap a = \emptyset$. Donc pour un $n \in \omega$ on a $a_n \cap a = \emptyset$, et en particulier $a_{n+1} \notin a_n$.

(2)⇒(1) : Soit $a \neq \emptyset$ un ensemble qui contredit (AF). Pour tout $b \in a$ on a donc $b \cap a \neq \emptyset$. Par (AC) il existe une fonction $f : a \to a$ telle que $f(b) \in b$ pour tout $b \in a$. On choisit $a_0 \in a$, puis on définit par récurrence sur ω une suite $(a_i)_{i\in\omega}$, en posant $a_{i+1} = f(a_i)$. □

Lemme 6.3.5. *1. Si $x \in V$, alors $\bigcup x, \mathcal{P}(x)$ et $\{x\}$ sont dans V. Ce sont des ensembles de rang plus petit que $\mathrm{Rg}(x) + \omega$.*

2. Si $x, y \in V$, alors $x \times y, x \cup y, x \cap y, \{x, y\}, (x, y)$ et x^y sont dans V aussi. De plus, le rang de ces ensembles est plus petit que $\max\{\mathrm{Rg}(x), \mathrm{Rg}(y)\} + \omega$.

Démonstration. Exercice. □

Exercice 6.3.6. On travaille dans $\mathcal{U} \models \mathrm{ZF}^-$, et on suppose que les ensembles $\mathbb{N}, \mathbb{Z}, \ldots$ sont définis de manière usuelle.

1. Montrer que les ensembles $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}, C^0([0,1], \mathbb{C})$ sont dans $V_{\omega \cdot 2}$.
2. Calculer les rangs des ensembles mentionnés dans (1).

Remarque. *Comme l'axiome d'extensionalité (qui dit en quelque sorte qu'il n'y a que des ensembles), l'axiome de fondation restreint l'univers des ensembles à l'endroit où les mathématiques usuelles ont lieu :*

- *Nous avons déjà vu dans l'exercice précédent que des objets mathématiques usuels sont dans V.*
- *Si on travaille dans* ZFC^-*, toute structure (du premier ordre) a une copie isomorphe dans V. (Supposons $\sigma^{\mathcal{L}}$ fini pour simplifier la présentation.) En effet, soit $\mathfrak{M} = \langle M; R_i, c_j, f_k \rangle$ une $\mathcal{L}$-structure donnée et $g : M \cong \mathrm{card}(M) = \kappa$ une bijection. Il existe une unique $\mathcal{L}$-structure $\mathfrak{N}$ avec ensemble de base κ telle que g soit un $\mathcal{L}$-isomorphisme entre $\mathfrak{M}$ et $\mathfrak{N}$. On a $\kappa \in V_{\kappa+1}$, puis $\mathfrak{N} \in V_{\kappa+\omega}$ par le lemme 6.3.5.*
- *En dehors des structures du premier ordre, cela reste vrai. Prenons par exemple le cas d'un espace topologique $(X, \mathcal{T})$ avec X un ensemble et $\mathcal{T} \subseteq \mathcal{P}(X)$ l'ensemble des ouverts de X. Alors si $g : X \cong \kappa$ est une bijection, on peut raisonner comme avant.*

6.4 Quelques résultats d'incomplétude, d'indépendance et de consistance relative

Dans cette dernière partie du chapitre sur la théorie des ensembles, nous supposons codées les formules ainsi que les preuves dans la théorie ZF^-. Pour cela, on pourra par exemple se servir du codage que nous avons donné dans l'arithmétique et utiliser que si $\mathcal{U} \models \mathrm{ZF}^-$ et $\omega \in U$, alors $\langle \omega; 0, +, \cdot, \mathrm{succ}, \in\restriction_\omega \rangle \models \mathcal{P}$ (voir 6.1.10). Nous continuons à écrire $\#\varphi$ pour le code de la formule φ.

Les résultats d'incomplétude de Gödel ont alors leurs analogues en théorie des ensembles (avec une preuve un peu plus simple). Notons que ZF^-, ZF, ZFC^- et ZFC sont des théories récursives.

Théorème 6.4.1 (Premier théorème d'incomplétude de Gödel)**.** *Soit $T \supseteq \mathrm{ZF}^-$ une $\mathcal{L}_{ens}$-théorie récursive et consistante. Alors T est incomplète.*

Remarque. *Nous allons voir (sans preuve) que si* ZF *est consistante, alors* $\mathrm{ZF} \nvdash (\mathrm{AC})$ *et* $\mathrm{ZF} \nvdash \neg(\mathrm{AC})$*. De même,* ZFC $\nvdash (\mathrm{CH})$ *et* $\mathrm{ZFC} \nvdash \neg(\mathrm{CH})$*. Dans les deux cas, il s'agit de propriétés avec un sens mathématique. Contrairement à cela, dans le cas de l'arithmétique, il fallait chercher loin pour trouver un énoncé indépendant de $\mathcal{P}$.*

Théorème 6.4.2 (Second théorème d'incomplétude de Gödel)**.** *Si $T \supseteq \mathrm{ZF}^-$ est une $\mathcal{L}_{ens}$-théorie récursive et consistante, alors $T \nvdash \mathrm{Coh}(T)$.*

On se place dans un modèle $\mathcal{U}$ de ZF^-. On peut alors coder la satisfaction dans $\mathcal{U}$. Soit $\mathcal{L}$ un langage fini et $\mathfrak{A} = \langle A; (Z^{\mathfrak{A}})_{Z \in \sigma^{\mathcal{L}}} \rangle$ une $\mathcal{L}$-structure avec $\mathfrak{A} \in U$. On identifie l'ensemble des affectations à valeur dans $\mathfrak{A}$ à A^ω, un élément de U.

Le lemme suivant s'obtient par induction sur la hauteur d'une formule, ce qui revient à une induction sur ω. Les détails sont laissés en exercice.

Lemme 6.4.3. *Il existe une fonction dans $\mathcal{U}$ qui à $(\#\varphi, \alpha)$ associe 1 si $\mathfrak{A} \models \varphi[\alpha]$, et 0 sinon.*

En particulier, $\#\,\mathrm{Th}(\mathfrak{A}) = \{\#\varphi \mid \varphi$ est un $\mathcal{L}$-énoncé tel que $\mathfrak{A} \models \varphi\}$ est donné par un élément de $\mathcal{U}$. □

Un théorème de *consistance relative* a la forme suivante : étant données deux théories T_1 et T_2, alors, si T_1 est consistante, T_2 est consistante aussi. Le principe de preuve sera de construire un modèle de T_2 à partir d'un modèle de T_1.

Commençons par un résultat qui établit une relation entre la théorie des ensembles et l'arithmétique. On vérifie sans problème que la preuve de la proposition 6.1.10 se fait dans ZF^-. Vu que ω est un ensemble, on peut déduire du lemme 6.4.3 le résultat suivant.

Proposition 6.4.4. *On a $\mathrm{ZF}^- \vdash \mathrm{Coh}(\mathcal{P})$. En particulier, si ZF^- est consistante, l'arithmétique de Peano $\mathcal{P}$ est consistante aussi.* □

Soit $\mathcal{U} \models \mathrm{ZF}^-$. Si X est une classe non vide contenue dans U, on peut considérer la $\mathcal{L}_{ens}$-structure $\langle X; \in\restriction_X \rangle$.

Dans les preuves des résultats de consistance relative en théorie des ensembles que nous allons donner, nous construisons des classes X telles que si $\mathcal{U} \models T_1$, alors $\langle X; \in\restriction_X \rangle \models T_2$, d'où le résultat de consistance relative voulu.

Notation. On écrira parfois $X \models \varphi$ au lieu de $\langle X; \in\restriction_X \rangle \models \varphi$.

Définition (Relativisation)**.** Soit $F(v_0)$ une $\mathcal{L}_{ens}$-formule. Pour toute formule φ on définit, par induction sur $\mathrm{ht}(\varphi)$, une formule φ^F, la *relativisée de φ à F* :

- $\varphi^F = \varphi$ si φ est atomique ;
- $(\varphi \wedge \psi)^F = (\varphi^F \wedge \psi^F)$ et $(\neg\varphi)^F = \neg(\varphi^F)$;
- $(\exists x \varphi)^F = \exists x (F(x) \wedge \varphi^F)$

Proposition 6.4.5. *1. Soit $\mathcal{U} \models \mathrm{ZF}^-$ et $X \subseteq U$ une classe. On suppose que $X = F[\mathcal{U}] = G[\mathcal{U}]$, c'est-à-dire que $F(v_0)$ et $G(v_0)$ définissent X. Alors pour toute formule φ, les formules φ^F et φ^G sont équivalentes dans $\mathcal{U}$. On pourra donc écrire φ^X au lieu de φ^F, si on ne s'intéresse qu'à la formule à équivalence près.*

2. Pour tout $a_1, \ldots, a_n \in X$ et toute formule $\varphi = \varphi(x_1, \ldots, x_n)$, on a l'équivalence suivante : $\mathcal{U} \models \varphi^X[\overline{a}]$ ssi $\langle X; \in\restriction_X \rangle \models \varphi[\overline{a}]$

Démonstration. (1) Preuve par induction sur $\mathrm{ht}(\varphi)$ (exercice).

(2) Preuve par induction sur $\mathrm{ht}(\varphi)$. Seul le cas $\varphi = \exists x_0 \psi$ est non trivial. On a $\mathcal{U} \models \varphi^X[a_1, \ldots, a_n]$ ssi $\mathcal{U} \models \exists x_0 (F(x_0) \wedge \psi^X)[a_1, \ldots, a_n]$ ssi il existe $b \in U$ tel que $\mathcal{U} \models F[b]$ et $\mathcal{U} \models \psi^X[b, \overline{a}]$ ssi il existe $b \in X$ tel que $\mathcal{U} \models \psi^X[b, \overline{a}]$ ssi il existe $b \in X$ tel que $X \models \psi[b, \overline{a}]$ ssi $X \models \varphi[\overline{a}]$. □

Définition. Soit X une classe définie par $F(v_0)$. Une formule $\varphi(x_1, \ldots, x_n)$ est dite *absolue pour X* si $\mathcal{U} \models \forall x_1, \ldots, x_n \left(\bigwedge_{i=1}^n F(x_i) \to (\varphi \leftrightarrow \varphi^X) \right)$.

Définition. L'ensemble des $\mathcal{L}_{ens}$*-formules* Δ_0 est le plus petit ensemble de formules qui contient les formules atomiques et qui est clos par combinaisons booléennes et *quantification bornée* (si φ est Δ_0, alors $\exists x(x \in y \wedge \varphi)$ et $\forall x(x \in y \to \varphi)$ sont des formules Δ_0 aussi).

Lemme 6.4.6. *1. Si X est une classe transitive (c'est-à-dire $z \in y \in X \Rightarrow z \in X$), alors toute formule Δ_0 est absolue pour X.*

2. Les propriétés suivantes s'expriment par des formules Δ_0 : $x \subseteq y$, $x = \emptyset$, $x = y \cup \{y\}$, 'x est transitif', $z = \{x, y\}$, $y = \bigcup x$

Démonstration. (1) Les formules atomiques sont absolues pour toute classe non vide. De plus, pour toute classe X, l'ensemble des formules absolues est clos par combinaisons booléennes. Montrons que si X est une classe transitive non vide et si $\varphi(x, y, v_1, \ldots, v_n)$ est absolue pour X, alors la formule $\exists x(x \in y \wedge \varphi)$ est absolue pour X aussi. Pour $b, c_1, \ldots, c_n \in X$, on a les équivalences suivantes :

$$\begin{aligned}
& \mathcal{U} \models (\exists x(x \in y \wedge \varphi))^X [b, \overline{c}] \\
\text{(par définition)} \quad & \iff \mathcal{U} \models \left(\exists x(F(x) \wedge x \in y \wedge \varphi^X)\right) [b, \overline{c}] \\
& \iff \text{il existe } a \in X \cap b \text{ tel que } \mathcal{U} \models \varphi^X [a, b, \overline{c}] \\
\text{(hypothèse d'induction)} \quad & \iff \text{il existe } a \in X \cap b \text{ tel que } \mathcal{U} \models \varphi [a, b, \overline{c}] \\
\text{(par transitivité de } X) \quad & \iff \text{il existe } a \in b \text{ tel que } \mathcal{U} \models \varphi [a, b, \overline{c}] \\
& \iff \mathcal{U} \models (\exists x(x \in y \wedge \varphi)) [b, \overline{c}]
\end{aligned}$$

La preuve de (2) est facile. Par exemple, la transitivité de x est exprimée par la formule Δ_0 suivante : $\forall y \, (y \in x \to \forall z(z \in y \to z \in x))$. □

Lemme 6.4.7. *On se place dans $\mathcal{U} \models \mathrm{ZF}^-$.*

1. *Si X est une classe transitive non vide, elle satisfait (Ext).*
2. *$V_\alpha \models (\bigcup)$ pour tout ordinal $\alpha > 0$. De même, $V \models (\bigcup)$.*
3. *Si λ est un ordinal limite, alors $V_\lambda \models (Paire)$. De même, $V \models (Paire)$.*
4. *Si $\emptyset \neq X$ est transitive telle que $\mathcal{U} \models \forall x \in X \exists y \in X(\mathcal{P}(x) \cap X = y)$, alors $X \models (Parties)$. En particulier, $V_\lambda \models (Parties)$ pour tout ordinal limite λ, et $V \models (Parties)$. De plus, $y = \mathcal{P}(x)$ est absolue dans ces deux cas.*

Démonstration. (1) L'enoncé $\forall y \in X \forall z \in X \, ([\forall x \in X(x \in y \leftrightarrow x \in z] \to y \doteq z)$ exprime $(\mathrm{Ext})^X$. Comme X est transitive, si $y, z \in X$, alors $y, z \subseteq X$, donc $(\mathrm{Ext})^X$ est satisfait dans $\mathcal{U}$.

(2) Commençons par une remarque générale : Si $G(x, y)$ définit une classe fonctionnelle (dans $\mathcal{U}$) de domaine $\mathcal{U}$ telle que $G(x, y)$ soit une formule absolue pour X, alors, pour montrer $X \models \forall x \exists! y G(x, y)$, il suffit de montrer que si $a \in X$, alors l'unique $b \in U$ tel que $\mathcal{U} \models G(a, b)$ est également dans X.

Comme $y = \bigcup x$ est une formule Δ_0 par le lemme 6.4.6, il suffit donc de montrer que si $x \in V_\alpha$, alors $\bigcup x \in V_\alpha$, ce qui est clair. En effet, on a les implications suivantes :

$$x \in V_\alpha \;\Rightarrow\; \mathrm{Rg}(x) = \beta < \alpha \;\Rightarrow\; \mathrm{Rg}(y) < \beta \text{ pour tout } y \in x' \in x$$
$$\Rightarrow\; \bigcup x \subseteq V_\beta \;\Rightarrow\; \bigcup x \in V_{\beta+1} \subseteq V_\alpha$$

Le même argument donne le résultat pour V.

(3) $x, y \in V_\alpha \Rightarrow \{x, y\} \in V_{\alpha+1}$, en particulier $x, y \in V_\lambda \Rightarrow \{x, y\} \in V_\lambda$ pour λ limite. Comme la formule $z = \{x, y\}$ est Δ_0, par la remarque générale on a $V_\lambda \models$ (Paire). De même pour V au lieu de V_λ.

(4) On a $\mathcal{U} \models \forall x \in X\, \exists y \in X(\mathcal{P}(x) \cap X \doteq y)$ si et seulement si $\mathcal{U} \models \forall x \in X\, \exists y \in X\, \forall z \in X(z \subseteq x \leftrightarrow z \in y)$. Ce dernier énoncé est juste $(\text{Parties})^X$, compte tenu du fait que $z \subseteq x$ est absolue pour X par transitivité de X.

Soit maintenant $x \in V_\alpha$. On a alors $\mathcal{P}(x) \subseteq V_\alpha$ et donc $\mathcal{P}(x) \in V_{\alpha+1}$. On obtient donc le résultat pour V_λ et pour V. L'absoluité de $y = \mathcal{P}(x)$ est claire. □

Lemme 6.4.8. *Les formules* $\mathrm{Ord}(x)$ *et* $\mathrm{Card}(x)$ *sont absolues pour* V *ainsi que pour* V_λ *si* λ *est un ordinal limite.*

Démonstration. Commençons par $\mathrm{Ord}(x)$. La transitivité de x ainsi que le fait que $\in\restriction_x$ définisse un ordre total s'expriment par des formules Δ_0 et sont donc absolues. Quant à la bonne fondation, on utilise l'absoluité de $y = \mathcal{P}(x)$ pour montrer que la formule suivante est absolue aussi :

$$\forall z\,(z \in \mathcal{P}(x) \wedge z \neq \emptyset \rightarrow \exists u(u \in z \wedge \forall w(w \in z \rightarrow w \notin u)))$$

La formule $\mathrm{Ord}(x) \wedge \forall y(y \in x \rightarrow \neg\exists f \in \mathcal{P}(x \times y)\, f : x \cong y)$ est équivalente à $\mathrm{Card}(x)$. On montre que les formules $f : x \cong y$ (en les trois variables f, x, y) ainsi que $z = x \times y$ sont absolues pour V ainsi que pour V_λ si λ est un ordinal limite (exercice). Cela établit l'absoluité de $\mathrm{Card}(x)$ dans les deux cas. □

Définition. On se place dans ZF^-.
- Un cardinal λ est appelé *fortement limite* si pour tout $\mu < \lambda$ on a $2^\mu < \lambda$.
- On dit que λ est (fortement) *inaccessible* s'il est fortement limite, régulier et $> \aleph_0$.

Exemple. On définit, par récurrence sur $\alpha \in \mathrm{Ord}$, une hiérarche cardinale comme suit : $\beth_0 := \aleph_0$, $\beth_{\alpha+1} := 2^{\beth_\alpha}$, et $\beth_\lambda := \sup_{\alpha<\lambda} \beth_\alpha$ pour λ limite.

Pour tout ordinal α limite, le cardinal $\beth_\alpha$ est alors fortement limite. Par contre, comme $\mathrm{cof}(\beth_\alpha) = \mathrm{cof}(\alpha)$ dans ce cas, $\beth_\alpha$ est en général singulier.

Lemme 6.4.9. *Soit* $\mathcal{U} \models \mathrm{ZF}^-$*, et soit ,* $X = V$ *ou* $X = V_\omega$*, ou* $X = V_\kappa$ *pour* κ *inaccessible. Dans le dernier cas, on suppose de plus que* $\mathcal{U} \models$ (AC).

Alors X *satisfait au schéma d'axiomes de remplacement ainsi qu'à l'axiome de fondation.*

Démonstration. On choisit une formule $F(x)$ qui définit X. Soit $G(v_0, v_1)$ une formule qui définit une classe fonctionnelle dans $\langle X, \in\restriction_X\rangle$. Alors la formule $H(v_0, v_1) = F(v_0) \wedge F(v_1) \wedge G^F$ définit une classe fonctionnelle dans $\mathcal{U}$. Soit $b := H[a] = \{H(c) \mid c \in a\}$, où a est un élément de X.

– Si $X = V$, alors $b \in V$ car $b \subseteq V$ et b est un ensemble.
– Si $X = V_\omega$, alors, comme a est fini, b est fini et une partie de V_ω, d'où $b \subseteq V_n$ pour un $n \in \omega$ et enfin $b \in V_\omega$.
– Si $X = V_\kappa$ pour κ inaccessible, essentiellement le même argument marche. On montre par induction sur α que $\mathrm{card}(V_\alpha) < \kappa$ pour tout $\alpha < \kappa$. (Comme on suppose $\mathcal{U} \models (\mathrm{AC})$, on peut se servir de la notion de cardinal.) Pour α successeur, on utilise que κ est fortement limite ; pour α limite, on utilise la régularité de κ.
Maintenant, si $a \in V_\kappa$, on a $a \in V_\alpha$ pour un $\alpha < \kappa$ et alors $a \subseteq V_{\alpha+1}$. Il s'en suit que $\mathrm{card}(b) \leq \mathrm{card}(a) < \kappa$. Par ailleurs, on a $b \subseteq V_\kappa$. Par régularité de κ, on obtient $\sup\{\mathrm{Rg}(c) \mid c \in b\} < \kappa$ et enfin $b \in V_\kappa$.

Cela montre le schéma d'axiomes de remplacement dans les trois cas.

Quant à (AF), on se donne $\emptyset \neq a \in X$. On a $a \subseteq X$ dans les trois cas, car X est transitive. Pour tout $b \in a$ de rang minimal on a $a \cap b = \emptyset$, et un tel b est dans X. □

Lemme 6.4.10. *On a $V \models (\mathrm{AI})$, $V_\kappa \models (\mathrm{AI})$ pour κ inaccessible et $V_\omega \models \neg(\mathrm{AI})$.*

Démonstration. On a $\omega \in V$, $\omega \in V_\kappa$ et $\omega \notin V_\omega$. Les détails sont laissés en exercice. □

Lemme 6.4.11. *Si $\mathcal{U} \models \mathrm{ZFC}^-$, alors $V \models (\mathrm{AC})$ et $V_\kappa \models (\mathrm{AC})$, pour κ inaccessible.*

Démonstration. Soit $(X_i)_{i \in I}$ un élément de V_κ. Alors $I \in V_\kappa$ et tout élément $g \in \prod_{i \in I} X_i$ est dans V_κ. Pour V, on conclut par le même argument. □

Théorème 6.4.12 (Consistance relative de (AF))**.**

1. *Si $\mathcal{U} \models \mathrm{ZF}^-$, alors $V \models \mathrm{ZF}$. En particulier, on a $\mathrm{Coh}(\mathrm{ZF}^-) \Rightarrow \mathrm{Coh}(\mathrm{ZF})$.*
2. *Si $\mathcal{U} \models \mathrm{ZFC}^-$, alors $V \models \mathrm{ZFC}$. En particulier, on a $\mathrm{Coh}(\mathrm{ZFC}^-) \Rightarrow \mathrm{Coh}(\mathrm{ZFC})$.*

Démonstration. Il suffit de combiner les lemmes précédents. (Rappelons que le schéma (Rem) implique le schéma (Com), voir 6.1.5.) □

Théorème 6.4.13. *Si $\mathcal{U} \models \mathrm{ZF}^-$, alors $V_\omega \models \mathrm{ZF} - (\mathrm{AI}) + \neg(\mathrm{AI})$. En particulier, $\mathrm{ZF}^- \vdash \mathrm{Coh}(\mathrm{ZFC} - (\mathrm{AI}) + \neg(\mathrm{AI}))$.*

Démonstration. Nous avons tout démontré dans les lemmes sauf $V_\omega \models (\mathrm{AC})$. Il suffit de remarquer que tout élément de V_ω est un ensemble fini et admet donc un bon ordre, c'est-à-dire qu'il existe une bijection avec un élément de ω. Une telle bijection est dans V_ω. □

Axiome des cardinaux inaccessibles (CI) : Il existe un cardinal inaccessible.

Théorème 6.4.14. *Soit $\mathcal{U} \models \mathrm{ZFC}^-$ et $\kappa \in U$ un cardinal inaccessible. Alors $V_\kappa \models \mathrm{ZFC}$. En particulier, $\mathrm{ZFC}^- + (\mathrm{CI}) \models \mathrm{Coh}(\mathrm{ZFC})$.*

Démonstration. Nous avons montré tous les axiomes dans les lemmes précédents. □

Par le second théorème d'incomplétude, le théorème 6.4.14 a comme corollaire que ZFC $\nvdash$ (CI). Voici le résultat de consistance relative correspondant que nous montrons plus directement, sans faire appel au second théorème d'incomplétude.

Théorème 6.4.15. *On a* $\mathrm{Coh}(\mathrm{ZFC}) \Rightarrow \mathrm{Coh}(\mathrm{ZFC} + \neg(\mathrm{CI}))$.

Démonstration. Soit $\mathcal{U} \models \mathrm{ZFC}$. On peut supposer que $\mathcal{U} \models (\mathrm{CI})$. Soit κ le plus petit cardinal inaccessible dans U. Alors $V_\kappa \models \mathrm{ZFC}$ par le théorème précédent. Il suffit de noter que $V_\kappa \models \neg(\mathrm{CI})$, ce qui suit des observations suivaintes :

- $y = \mathcal{P}(x)$, $\mathrm{Ord}(x)$ et $\mathrm{Card}(x)$ sont des formules absolues pour V_κ (Lemmes 6.4.7 et 6.4.8) ;
- "λ est un cardinal régulier" est absolu pour V_κ (cette propriété s'exprime par la formule $\neg\exists\alpha < \lambda(\exists f : \alpha \to \lambda$ cofinale), exercice) ;
- "λ est un cardinal fortement limite" est absolu pour V_κ (cette propriété s'exprime par la formule $\forall\alpha < \lambda(\neg\exists f : \mathcal{P}(\alpha) \to \lambda$ surjective), exercice).

□

À la fin de ce cours, nous mentionnons quelques résultats importants que nous citons sans preuve. Le livre de Kunen [4] est une référence excellente.

Le théorème suivant peut s'obtenir de manière assez élémentaire par la *méthode de Fraenkel-Mostowski* (voir la feuille de TD 14).

Théorème 6.4.16. *1. Si* ZF *est consistante, alors* $\mathrm{ZFC}^- + \neg(\mathrm{AF})$ *aussi.*

2. Si ZF *est consistante, alors* $\mathrm{ZF}^- + \neg(\mathrm{AC})$ *aussi.*

Il est possible de remplacer $\mathrm{ZF}^- + \neg(\mathrm{AC})$ par $\mathrm{ZF} + \neg(\mathrm{AC})$ dans la seconde partie du théorème précédent, mais la preuve est alors beaucoup plus difficile.

Rappelons que l'hypothèse généralisée du continue GCH est donnée par l'énoncé $\forall\alpha(2^{\aleph_\alpha} = \aleph_{\alpha+1})$.

Théorème 6.4.17 (Gödel)**.** *Si* ZF *est consistante, alors* $\mathrm{ZFC} + (\mathrm{GCH})$ *aussi.*

Gödel obtient ce résultat par la *méthode des constructibles*. Voici l'idée de la construction qui est similaire à la celle de V.

- On suppose formalisées dans $\mathcal{U} \models \mathrm{ZF}$ la syntaxe ainsi que la satisfaction pour des structures $\langle a; \in\restriction_a \rangle$, où $a \in U$.
- On montre alors qu'il existe une classe fonctionnelle $\mathcal{D}$ qui donne l'ensemble des parties définissables, c'est-à-dire qu'on a $b \in \mathcal{D}(a) \Leftrightarrow b \subseteq a$ et il existe $n \in \omega$, une formule $\varphi[v_0, \ldots, v_n]$ et $c_1, \ldots, c_n \in a$ tels que $b = \{c_0 \in a \mid \langle a; \in\restriction_a \rangle \models \varphi[c_0, c_1, \ldots, c_n]\}$.
- On définit $L_0 := \emptyset$, $\mathcal{L}_{\alpha+1} := \mathcal{D}(L_\alpha)$, et $L_\lambda := \bigcup_{\alpha<\lambda} L_\alpha$ pour λ limite. Puis, on définit "$L = \bigcup_\alpha L_\alpha$" (une classe propre).

– Comme pour la hiérarchie de von Neumann, on établit certaines propriétés de base, par exemple $\alpha \leq \beta \Rightarrow L_\alpha \subseteq L_\beta$ et la transitivité des L_α. On a donc une hiérarchie croissante et continue d'ensembles transitifs $(L_\alpha)_{\alpha \in \mathrm{Ord}}$.
– On définit un rang Rg_L comme pour V, et on montre que $\mathrm{Rg}_L(\alpha) = \alpha$, autrement dit $L_\alpha \cap \mathrm{Ord} = \alpha$.
– On montre que $L^L = L$ (dans ZF), c'est-à-dire que L satisfait à l'*axiome de constructibilité* $\forall x\, L(x)$. Pour cela, on utilise que la classe fonctionnelle $\mathcal{D}$ est absolue pour une classe X qui est modèle de $\mathrm{ZF} - (\text{Parties})$, et donc la classe fonctionnelle $\alpha \mapsto L_\alpha$ aussi.
– La preuve que $L \models \mathrm{ZF}$ est similaire à celle pour V.
– Afin d'établir $L \models (\mathrm{AC}) + (\mathrm{GCH})$, on construit, par récurrence sur α, un bon ordre sur L_α tel que si $\alpha \leq \beta$, alors L_α est un segment initial de L_β. Dans l'étape successeur, à l'aide du bon ordre sur L_α, on construit un bon ordre sur $L_\alpha^{<\omega}$ (l'ensemble des suites finies d'éléments de L_α), puis sur $L_{\alpha+1}$, en énumérant les (codes des) formules également.

Théorème 6.4.18 (Cohen)**.** *1. Si* ZF^- *est consistante, alors* $\mathrm{ZFC} + \neg(\mathrm{HC})$ *aussi.*

2. Si ZF^- *est consistante, alors* $\mathrm{ZF} + \neg(\mathrm{AC})$ *aussi.*

Contrairement aux constructions de modèles de ZF^- que nous avons vues ou esquissées jusqu'ici, les constructions utilisées pour montrer ce résultat *ne se font pas à l'intérieur du modèle de départ* $\mathcal{U}$. En revanche, la méthode du forcing (introduite par Cohen en 1964 pour montrer son théorème et utilisée en théorie des ensembles jusqu'à nos jours) permet de construire des modèles de ZF qui sont des *extensions* du modèle de départ.

Affaiblissements de l'axiome du choix

L'axiome du choix dépendant (ACD)

$$\forall r \forall a \forall x_0 \big[x_0 \in a \wedge r \subseteq a^2 \wedge \forall x \in a\, \exists y \in a\, (x, y) \in r \\ \rightarrow \exists f (f : \omega \rightarrow a \wedge f(0) = x_0 \wedge \forall n \in \omega\, ((f(n), f(n+1)) \in r) \big]$$

L'axiome du choix dénombrable (ACC)

$$\forall f [(\mathrm{Fn}(f) \wedge \mathrm{dom}(f) \doteq \omega \wedge \emptyset \notin \mathrm{im}(f)) \\ \rightarrow \exists g\, (\mathrm{Fn}(g) \wedge \mathrm{dom}(g) \doteq \omega \wedge \forall x (x \in \mathrm{dom}(g) \rightarrow g(x) \in f(x)))]$$

L'axiome (ACC) exprime que le produit d'une famille dénombrable d'ensembles non vides est non vide.

Lemme 6.4.19. *On a (AC)⇒(ACD)⇒(ACC).*

Démonstration. (AC)⇒(ACD) : Soit $h : \mathcal{P}'(a) \rightarrow a$ une fonction de choix sur a. Par induction sur ω, on définit une fonction f, via $f(0) := x_0$, $f(n+1) := h(\{y \in a \mid (f(n), y) \in r\})$. Il est clair que f convient.

(ACD)⇒(ACC) : Soit $(X_n)_{n\in\omega}$ une famille dénombrable d'ensembles non vides. On pose $Y_n := \{n\} \times X_n$, $a := \bigcup_{n\in\omega} Y_n$ et

$$r := \{(x, y) \in a^2 \mid \text{ il existe } n \in \omega \text{ tel que } x \in Y_n \text{ et } y \in Y_{n+1}\}.$$

Par (ACD) il existe une fonction $f : \omega \to a$ telle que $f(n) \in Y_n$ pour tout $n \in \omega$. Alors $g \in \prod_{n\in\omega} X_n$, où $g(n) := \pi((f(n))$, avec π la projection sur la deuxième coordonnée. □

La construction d'un ensemble non mesurable de réels qu'on voit au cours d'intégration se fait dans ZFC. Cela donne le résultat suivant.

Proposition 6.4.20. ZFC $\models$ *"il existe une partie de* $\mathbb{R}$ *non mesurable"*

Le théorème que nous mentionnons maintenant sans preuve est plus difficile. Pour le montrer, on utilise la technique du forcing.

Théorème 6.4.21 (Solovay, 1970)**.** *Si* ZFC + (CI) *est consistante, alors la théorie* ZF + (ACD) + *"toute partie de* $\mathbb{R}$ *est mesurable" est consistante aussi.*

Bibliographie

[1] René Cori et Daniel Lascar, *Logique mathématique*, Tome 1, Dunod, 2003.

[2] René Cori et Daniel Lascar, *Logique mathématique*, Tome 2, Dunod, 2003.

[3] Heinz-Dieter Ebbinghaus, Jörg Flum et Wolfgang Thomas, *Mathematical Logic*, Second Edition, Undergraduate Texts in Mathematics, Springer, 1994.

[4] Kenneth Kunen, *Set Theory. An Introduction to Independence Proofs*, Studies in Logic and the Foundations of Mathematics vol. **102**, Elsevier, 1980.

[5] François Loeser, *Un premier cours de logique*, Notes de cours, 2010 (disponible sur la page `http://www.math.jussieu.fr/~loeser/notes.php`).

[6] Martin Ziegler, *Mathematische Logik*, Mathematik Kompakt, Birkhäuser, 2010.