UN PREMIER COURS DE LOGIQUE

par

François Loeser

Table des matières

1. Apprendre à compter

Dans cette section on donne aux termes "ensembles", "entiers", etc., leur sens usuel en mathématiques. Dans un chapitre ultérieur on verra que l'essentiel des notions et résultats de ce chapitre sont valides en n'utilisant que les axiomes de Zermelo-Fraenkel et l'axiome du choix.

1.1. Le Théorème de Cantor-Bernstein et le Théorème de Cantor. — Ce sont deux résultats classiques fondamentaux concernant les ensembles:

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

Démonstration. — On pose $A_{\geq 0} = A$ et $B_{\geq 0} = B$ et on définit par récurrence $A_{\geq n+1} = g(B_{\geq n})$ et $B_{\geq n+1} = f(A_{\geq n})$. On pose $A_n = A_{\geq n} \setminus A_{\geq n+1}$ et on note A_∞ le complémentaire de la réunion des A_n dans A. On définit de même B_n et B_∞. On note A_{pair} la réunion des A_{2n}, A_{impair} la réunion des A_{2n+1}, de même pour B. L'ensemble A est réunion disjointe de A_{pair}, A_{impair}, et A_∞ et de même pour B. On remarque que f établit une bijection entre A_{pair} et B_{impair}, g^{-1} établit une bijection entre A_{impair} et B_{pair} et f entre A_∞ et B_∞. On définit donc une bijection h entre A et B en posant $h(x) = g^{-1}(x)$ si $x \in A_{\text{impair}}$ et $h(x) = f(x)$ sinon. □

***1.1.2. Théorème* (Cantor).** — *Soit X un ensemble et $\mathfrak{P}(X)$ l'ensemble des parties de X. Il n'existe pas de surjection $X \to \mathfrak{P}(X)$.*

Démonstration. — Soit $f : X \to \mathfrak{P}(X)$. Considérons

$$Y = \{a \in X; a \notin f(a)\}.$$

Supposons que $Y = f(b)$, pour $b \in X$. Si $b \in Y$, $b \notin f(b)$, absurde. Si $b \notin Y$, $b \in f(b)$, également absurde. □

1.2. Rappel sur les ensembles ordonnés. — Soit X un ensemble. Une relation d'ordre $<$ sur X est une relation transitive (si $x < y$ et $y < z$ alors $x < z$) et antiréflexive ($x \not< x$). Si pour tout $x \neq y$, $x < y$ ou $y < x$ on dit que $<$ est un ordre total. On note $x \leq y$ pour $x < y$ ou $x = y$. Si $Y \subset X$, $y \in Y$ est un plus petit élément si pour tout x dans Y, $y \leq x$. C'est un élément minimal si pour tout x dans Y, $y \not> x$. 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 élément maximal parmi les minorants. On définit de même la notion de majorant et de borne supérieure.

Noter qu'un plus petit (grand) élément est nécessairement unique, ce qui n'est pas le cas pour un élément minimal (maximal) en général.

1.2.1. Définition. — Un bon ordre sur un ensemble X est un ordre tel que toute partie non vide de X admet un plus petit élément.

1.2.2. Remarques. — (1) Un bon ordre est toujours total.

(2) Dans un ensemble bien ordonné il n'existe pas de suites strictement décroissantes.

1.3. Ordinaux. —

1.3.1. Définition. — Un ensemble X est transitif si quand $y \in X$ et $x \in y$, alors $x \in X$. Autrement dit, $y \in X \implies y \subset X$. Un ensemble X est un ordinal s'il est transitif et s'il est bien ordonné par $\in$.

1.3.2. Proposition. — *(1) L'ensemble vide $\emptyset$ est un ordinal.*
(2) Si X est un ordinal non vide alors $\emptyset \in X$.
(3) Si α est un ordinal, $\alpha \notin \alpha$.
(4) Si α est un ordinal et $\beta \in \alpha$, β est un ordinal.
(5) Si α est un ordinal et $\beta \in \alpha$, alors $\beta = S_{<\beta} := \{x \in \alpha, x < \beta\}$.
(6) Si α et β sont des ordinaux, $\beta \subset \alpha$ ssi $\beta \in \alpha$ ou $\beta = \alpha$.
(7) Si α est un ordinal, $\beta := \alpha \cup \{\alpha\}$ est un ordinal. On le note α^+.

Démonstration. — (1) est clair.

(2): En effet X admet un élément minimal y qui ne peut être que $\emptyset$. En effet, sinon il existerait $x \in y$ et y ne serait pas minimal.

(3): si $\alpha \in \alpha$, $\alpha \notin \alpha$ par antiréflexivité.

(4): comme $\beta \subset \alpha$, $\in$ se restreint en un bon ordre sur β (qui est transitif car $\beta \in \alpha$).

(5) est clair.

(6): En effet si $\beta \subset \alpha$ et $\beta \neq \alpha$, on considère x élément minimal de $\alpha \setminus \beta$. Comme aucun élément de $\alpha \setminus \beta$ ne peut être $< x$, on a $S_{<x} \subset \beta$. D'autre part, si $z \in \beta$, on a $z < x$, car sinon $x < z$ et $x \in \beta$. On a donc $\beta = S_{<x}$, d'où par (5) $x = \beta$, et $\beta \in \alpha$. La réciproque est claire. □

1.3.3. Théorème. — *Soient α et β deux ordinaux. L'une des trois propriétés suivantes est vérifiée*

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

Démonstration. — Soit $\gamma = \alpha \cap \beta$. C'est un ensemble transitif bien ordonné par $\in$, c'est donc un ordinal. Si $\gamma = \alpha$, on a $\alpha \subset \beta$ et donc $\alpha \in \beta$ ou $\alpha = \beta$. De même si $\gamma = \beta$ on a terminé. Sinon $\gamma \in \alpha$ et $\gamma \in \beta$, donc $\gamma \in \gamma$, absurde. □

1.3.4. Proposition. — *Si A est un ensemble non vide d'ordinaux il possède un plus petit élément, à savoir $\cap_{\alpha \in A} \alpha$.*

Démonstration. — En effet, $\beta = \cap_{\alpha \in A} \alpha$ est clairement un ordinal et, pour tout $\alpha \in A$, $\beta \leq \alpha$. De plus, si pour tout $\alpha \in A$, $\beta < \alpha$, on aurait $\beta \in \beta$, absurde. Donc $\beta \in A$. □

On tire de la proposition précédente une autre preuve que si α et β sont des ordinaux, alors $\alpha > \beta$ ou $\beta \leq \alpha$.

1.3.5. Proposition. — *Soit A un ensemble d'ordinaux. Alors $a = \cup_{\alpha \in A} \alpha$ est un ordinal et si $\beta < a$, il existe $\alpha \in A$ tel que $\beta \in \alpha$. On écrira aussi $a = \sup_{\alpha \in A} \alpha$.*

Démonstration. — En effet il est clair que a est ordonné. Si X est une partie non vide de a, l'intersection des éléments de X est le plus petit élément de X. Ceci prouve que a est un ordinal. Si $\alpha \in A$, $\alpha \subset a$, donc $\alpha \leq a$. Si $\beta < a$, il existe α dans A avec $\beta \in \alpha$, donc $\beta < \alpha$. □

1.3.6. Définition. — Un ordinal est limite s'il n'est pas la forme α^+ et s'il est non vide.

1.3.7. Proposition. — *Soit λ un ordinal non vide. Les deux conditions suivantes sont équivalentes :*

(1) λ est limite
(2) $\lambda = \cup_{\alpha \in \lambda} \alpha$.

Démonstration. — Soit λ un ordinal limite. Soit $\beta = \cup_{\alpha \in \lambda} \alpha$. Supposons $\alpha \in \lambda$. Comme $\lambda \neq \alpha^+$, on a $\alpha^+ \in \lambda$ ou $\lambda \in \alpha^+$. Si $\lambda \in \alpha^+$, $\lambda \in \alpha$ ou $\lambda = \alpha$, absurde. En particulier, si $\alpha \in \lambda$, alors $\alpha^+ \subset \beta$ et donc $\alpha \in \beta$, d'où $\lambda \subset \beta$. L'inclusion réciproque est claire. Si $\lambda = \gamma^+$, alors $\cup_{\alpha \in \lambda} \alpha = \gamma$, d'où le résultat. □

Pour tout entier naturel n, on définit un ordinal $\underline{n}$ par $\underline{0} = \emptyset$ et $\underline{n} = \underline{n} - 1^+$, pour $n \geq 1$. On a $\underline{1} = \{\emptyset\}$, $\underline{2} = \{\emptyset, \{\emptyset\}\}$, etc. On note ω l'ordinal $\cup_{n\in\mathbb{N}}\underline{n}$. C'est un ordinal limite et en fait le plus petit ordinal limite. On dit qu'un ordinal est fini si ni lui-même, ni aucun de ses éléments n'est limite. Clairement un ordinal est fini si et seulement si il est de la forme $\underline{n}$ avec $n \in \mathbb{N}$. Dorénavant on identifiera $\underline{n}$ et n.

1.3.8. Théorème. — *Tout ensemble bien ordonné X est isomorphe comme ensemble ordonné à un ordinal. De plus un tel isomorphisme est unique.*

Démonstration. — Commençons par établir l'unicité. Il suffit de montrer que si $f : \alpha \to \alpha'$ est un isomorphisme d'ensembles ordonnés entre deux ordinaux alors $\alpha = \alpha'$ et f est l'identité. Supposons que l'ensemble des x de α avec $f(x) \neq x$ est non vide et considérons un tel x minimal. Alors la restriction de f à $S_{<x}$ est l'identité. Mais alors $y \in x \iff f(y) \in f(x)$, autrement dit x et $f(x)$ ont les mêmes éléments, donc $f(x) = x$, absurde.

Pour l'existence, remarquons que tout isomorphisme entre $S_{<x}$ et un ordinal α, s'étend en un isomorphisme entre $S_{\leq x} = S_{<x} \cup \{x\}$ et α^+. Il s'ensuit que pour tout x dans X il existe un isomorphisme $f_x : S_{\leq x} \to \alpha(x)$ avec $\alpha(x)$ un ordinal. En effet, considérons sinon x minimal ne vérifiant pas cette propriété. Pour $y < x$ on dispose alors d'un isomorphisme $f_y : S_{\leq y} \to \alpha(y)$. Si $y' < y$, f_y se restreint en $f_{y'}$ par unicité. Posons $\alpha = \cup_{y<x}\alpha(y)$. C'est un ordinal et $f : S_{<x} \to \alpha$ défini par $f(y) = f_y(y)$ est un isomorphisme car la restriction de f à $S_{\leq y}$ coïncide avec f_y. Un argument similaire permet de conclure : on pose $\alpha = \cup_{x\in X}\alpha(x)$. C'est un ordinal et $f : X \to \alpha$ défini par $f(x) = f_x(x)$ est un isomorphisme car la restriction de f à $S_{\leq x}$ coïncide avec f_x. □

1.4. Opérations sur les ordinaux. —

1.4.1. Somme de deux ensembles ordonnés. — Soient A et B deux ensembles ordonnés. On note $A + B$ l'ensemble formé des paires $(a, 1)$ avec $A \in A$ et $(b, 2)$ avec $b \in B$. On définit un ordre sur $A+B$ en posant $(x, i) < (y, j)$ si $i = j$ et $w < y$ ou si $i < j$. Si A et B sont totalement (resp. bien) ordonnés, alors $A + B$ est totalement (resp. bien) ordonné. Notons que $(A + B) + C$ est isomorphe à $A + (B + C)$.

1.4.2. Produit de deux ensembles ordonnés. — Soient A et B deux ensembles ordonnés. On munit le produit $A \times B$ de l'ordre $(a, b) < (a', b')$ si $b < b'$ ou si $b = b'$ et $a < a'$. Si A et B sont totalement (resp. bien) ordonnés, alors $A \times B$ est totalement (resp. bien) ordonné. A isomorphisme près cette opération est associative ($(A \times B) \times C$ est isomorphe à $A \times (B \times C)$) et distributive à gauche par rapport à l'addition ($A \times (B + C)$ est isomorphe à $(A \times B) + (A \times C)$).

1.4.3. Exponentiation. — Soient A et B deux ensembles ordonnés. On suppose que A possède un plus petit élément 0. On considère l'ensemble $A^{(B)}$ des suites à support fini, c'est à dire le sous-ensemble de A^B formé des applications $f : B \to A$ avec $\{i \in B; f(i) \neq 0\}$ fini. On pose $f < g$ s'il existe i tel que $f(i) < g(i)$ et $f(j) = g(j)$ pour $j > i$. Si A et B sont totalement (resp. bien) ordonnés, alors $A^{(B)}$

est totalement (resp. bien) ordonné. On a des isomorphismes d'ensembles ordonnés $A^{(B+C)} \simeq A^{(B)} \times A^{(C)}$ et $A^{(B\times C)} \simeq (A^{(B)})^{(C)}$.

Le seul point non trivial est le suivant :

1.4.4. Proposition. — *Si A et B sont bien ordonnés, alors $A^{(B)}$ est bien ordonné.*

Démonstration. — Soit X une partie non vide de $A^{(B)}$. Si la fonction constante f_0 de valeur 0 est dans X, on a terminé. Sinon chaque fonction f dans X a un support non vide et on note $s_1(f)$ le plus grand élément du support de f. Soit alors b_1 la valeur minimale de $s_1(f)$ et a_1 la valeur minimale de $f(b_1)$ pour les f avec $s_1(f) = b_1$. L'ensemble X_1 formé des f de X avec $s_1(f) = b_1$ et $f(b_1) = a_1$ est une partie non vide de X, en fait un segment initial de X. En effet si $g \in X_0$ et $f \in X \setminus X_0$, ou bien $s_1(f) > b_1$ et alors $g < f$, ou bien $s_1(f) = b_1$ et $f(b_1) > a_1$ et alors $g < f$ également. Si X_1 contient la fonction dont la seule valeur non nulle est $f(b_1) = a_1$, on a terminé. Sinon, toutes les fonctions dans X_1 ont un support de cardinal ≥ 2 et, pour f dans X_1, on note $s_2(f)$ le plus grand élément du support de f après $s_1(f)$. Soit alors b_2 la valeur minimale de $s_2(f)$ pour f dans X_1 et a_2 la valeur minimale de $f(b_2)$ pour les f avec $s_2(f) = b_2$. On considère l'ensemble X_2 formé des f dans X_2 avec $s_2(f) = b_2$ et $f(b_2) = a_2$ qui est un segment initial de X_2. On continue si nécessaire en construisant b_i, a_i, X_i. Ce processus s'arrête nécessairement, car la suite des b_i étant strictement décroissante, elle ne peut prendre qu'un nombre fini de valeurs. □

1.4.5. Arithmétique des ordinaux. — Si α et α' sont deux ordinaux, on note $\alpha + \alpha'$ l'unique ordinal isomorphe à la somme des ensembles ordonnés α et α'. On définit de même $\alpha\alpha'$ comme l'unique ordinal isomorphe à $\alpha \times \alpha'$ et $\alpha^{\alpha'}$ comme l'unique ordinal isomorphe à $\alpha^{(\alpha')}$.

***1.4.6. Proposition* (Addition).** — *(1)* $\alpha + 0 = 0 + \alpha = \alpha$
(2) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
(3) $\alpha + 1 = \alpha^+$
(4) On a $\alpha < \beta$ ssi il existe γ non nul tel que $\beta = \alpha + \gamma$.
(5) Si $\beta < \beta'$ alors pour tout α $\alpha + \beta < \alpha + \beta'$. En particulier si $\alpha + \beta = \alpha + \beta'$, $\beta = \beta'$.
(6) Si λ est limite, alors $\alpha + \lambda = \sup_{\beta<\lambda} \alpha + \beta$.
(7) $1 + \alpha = \alpha + 1$ si α est fini et $1 + \alpha = \alpha$ sinon.

Démonstration. — (1), (2), (3) et (4) sont clairs. Pour (5) on écrit $\beta' = \beta + \gamma$ avec γ non nul et on utilise l'associativité. Pour (6), certainement $\alpha + \lambda$ est au moins égal au sup. Réciproquement, soit $\alpha \leq \theta < \alpha + \lambda$. Alors $\theta = \alpha + \delta$ avec $\delta < \lambda$. Pour le deuxième point de (7), d'après (5) il suffit de démontrer que $1 + \omega = \omega$, ce qui est clair. □

***1.4.7. Proposition* (Multiplication).** — *(1)* $\alpha 0 = 0\alpha = 0$.
(2) $\alpha 1 = 1\alpha = \alpha$.
(3) $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ *et* $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

(4) $2\omega = \omega \neq \omega 2 = \omega + \omega$.
(5) Supposons $\alpha \neq 0$. *Si* $\beta < \beta'$, $\alpha\beta < \alpha\beta'$. *En particulier si* $\alpha\beta = \alpha\beta'$ *alors* $\beta = \beta'$.
(6) Si λ *est limite, alors* $\alpha\lambda = \sup_{\beta<\lambda} \alpha\beta$.

Démonstration. — Le seul point non évident est (6). Clairement $\alpha\lambda$ est au moins égal au sup. Pour l'inégalité réciproque, soit $\theta < \alpha\lambda$. Par le lemme qui suit $\theta = \alpha\sigma + \rho$ avec $\sigma < \lambda$ et $\rho < \alpha$. On a $\theta < \alpha(\sigma + 1)$, donc θ est strictement inférieur au sup. □

1.4.8. Lemme. — *Si* $\gamma < \alpha\beta$, *il existe* $\rho < \alpha$ *et* $\sigma < \beta$ *tels que* $\gamma = \alpha\sigma + \rho$.

Démonstration. — On pose $(\rho, \sigma) = f^{-1}(\gamma)$ pour $f : \alpha \times \beta \to \alpha\beta$ l'unique isomorphisme entre ensembles bien ordonnés. □

1.4.9. Exercice. — Montrer que pour tout ordinal β et tout ordinal non nul α et il existe un unique couple d'ordinaux (σ, ρ) avec $\beta = \alpha\sigma + \rho$ et $\rho < \alpha$.

***1.4.10. Proposition* (Exponentiation).** — *(1) On a* $\alpha^0 = 1$, $\alpha^1 = \alpha$, $1^\beta = 1$. *Pour* $\beta \neq 0$, $0^\beta = 0$.
(2) $\alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$ *et* $\alpha^{\beta\gamma} = (\alpha^\beta)^\gamma$.
(3) Si $\alpha > 1$, $\alpha^\beta < \alpha^{\beta'}$ *si* $\beta < \beta'$.
(4) Si $\alpha > 1$, *et* λ *est limite,* $\alpha^\lambda = \sup_{\beta<\lambda} \alpha^\beta$.

Démonstration. — Montrons la direction non triviale de (4). Soit f une fonction à support fini de λ dans α. Il existe $\beta < \lambda$ strictement plus grand que tous les éléments du support de f. Il suit que le segment initial de $\alpha^{(\lambda)}$ déterminé par f peut être plongé dans un segment initial de $\alpha^{(\beta)}$. □

1.5. Suites de Goodstein. — On écrit un entier n en base p:

$$n = p^{n_1} c_1 + \cdots + p^{n_k} c_k$$

avec les c_i des entiers $< p$. On peut alors exprimer les exposants n_i en base p et itérer le processus. On appelle représentation de n en base p itérée la décomposition ainsi obtenue.

Pour $q \geq p \geq 2$, on définit une fonction $f_{p,q} : \mathbf{N} \to \mathbf{N}$ comme suit. L'entier $f_{p,q}(n)$ est obtenu en remplaçant tous les p par des q dans la décomposition de n en base p itérée.

Exemple : $f_{3,4}(26) = f_{3,4}(3^2 \cdot 2 + 3 \cdot 2 + 2) = 4^2 \cdot 2 + 4 \cdot 2 + 2 = 42$.

Pour a un entier, on pose $g_2(a) = a$ et $g_{p+1}(a) = f_{p,p+1}(g_p(a)) - 1$ si $g_p(a)$ est non nul, $g_{p+1}(a) = 0$ sinon.

Ainsi $g_2(5) = 5 = 2^{2^1} + 1$, $g_3(5) = f_{2,3}(5) - 1 = (3^{3^1} + 1) - 1 = 27 = 3^{3^1}$, $g_4(5) = f_{3,4}(27) - 1 = (4^{4^1}) - 1 = 255 = 4^3 3 + 4^2 3 + 4^1 3 + 3$, $g_5(5) = f_{4,5}(255) - 1 = (5^3 3 + 5^2 3 + 5^1 3 + 3) - 1 = 447 = 5^3 3 + 5^2 3 + 5^1 3 + 2$, $g_6(5) = f_{5,6}(447) - 1 = (6^3 3 + 6^2 3 + 6^1 3 + 2) - 1 = 775$, etc.

***1.5.1. Théorème* (Goodstein).** — *Pour chaque entier a il existe un entier p tel que $g_p(a) = 0$.*

Démonstration. — On introduit la fonction $f_{p,\omega}$ de $\mathbf{N}$ à valeurs dans les ordinaux définie sur le modèle précédent : $f_{p,\omega}(n)$ est obtenu en remplaçant tous les p par des ω dans la décomposition de n en base p itérée et les entiers $< p$ par les ordinaux finis correspondants.

Démontrons par récurrence que $f_{p,\omega}(n) < f_{p,\omega}(n+1)$ pour tout $n \in \mathbf{N}$ et $p \geq 2$. Pour $n = 0$ et $n = 1$ le résultat est clair. Comme $p^n - 1 > n$ pour $p \geq 2$ et $n \geq 2$, on peut écrire

$$n = p^{n-1}c_{n-1} + \cdots + p^0 c_0$$

avec les c_i des entiers $< p$ qui ne sont pas tous égaux à $p-1$. Soit m le plus petit entier tel que c_m ne soit pas égal à $p-1$. On a

$$n+1 = p^{n-1}c_{n-1} + \cdots + p^m(c_m + 1).$$

Pour montrer que $f_{p,\omega}(n+1) > f_{p,\omega}(n)$ il suffit de vérifier que

$$\omega^{f_{p,\omega}(m)} > \omega^{f_{p,\omega}(m-1)}c_{m-1} + \cdots + \omega^{f_{p,\omega}(0)}c_0.$$

Or le terme de droite est majoré par $\omega^{f_{p,\omega}(m-1)}(c_{m-1}+1)$ qui est lui même strictement majoré par $\omega^{f_{p,\omega}(m-1)+1}$.

Pour conclure on considère la suite $\tilde{g}_p(a) := f_{p,\omega}(g_p(a))$. Si $g_p(a)$ est non nul, on a

$$\tilde{g}_{p+1}(a) = f_{p+1,\omega}(g_{p+1}(a)) = f_{p+1,\omega}(f_{p,p+1}(g_p(a)) - 1) < f_{p+1,\omega}(f_{p,p+1}(g_p(a))),$$

or

$$f_{p+1,\omega}(f_{p,p+1}(g_p(a))) = f_{p,\omega}(g_p(a)) = \tilde{g}_p(a).$$

Comme il n'existe pas de suite infinie strictement décroissante d'ordinaux, on en déduit le résultat. □

***1.5.2. Remarque*.** — C'est un théorème de Kirby et Paris que l'énoncé du théorème de Goodstein n'est pas conséquence des axiomes de Peano. Par contre il se déduit des axiomes de Zermelo-Fraenkel.

1.6. L'axiome du choix. — Soit $(X_i)_{i\in I}$ une famille d'ensembles indexée par un ensemble I. On note $\prod_{i\in I} X_i$ l'ensemble des applications $f : I \to \cup_{i\in I} X_i$ telles que $f(i) \in X_i$ pour tout i.

L'axiome du choix est l'énoncé suivant :

***1.6.1. Axiome du choix* (AC).** — *Si tous les ensembles X_i sont non vides, alors $\prod_{i\in I} X_i$ n'est pas vide.*

Dans le système de Zermelo-Fraenkel l'axiome du choix est équivalent au Lemme de Zorn, ainsi qu'au théorème de Zermelo.

***1.6.2. Définition*.** — Un ensemble ordonné X est inductif si toute partie Y de X totalement ordonnée par $<$ admet un majorant dans X.

Un tel ensemble n'est pas vide.

***1.6.3. Théorème* (Lemme de Zorn).** — *Tout ensemble ordonné inductif admet un élément maximal.*

***1.6.4. Théorème* (Théorème de Zermelo).** — *Pour tout ensemble X il existe un bon ordre sur X.*

1.7. Cardinaux. — A partir de maintenant on supposera que l'axiome du choix est vérifié. Ceci est nécessaire pour développer une théorie des cardinaux raisonnable. Notons que la section sur les ordinaux ne requiérait par contre pas l'usage de l'axiome du choix.

Soient A et B deux ensembles. On dit qu'ils sont équipotents s'ils sont en bijection et on dit que A est subpotent à B s'il existe une injection de A dans B.

On appelle cardinal un ordinal qui n'est équipotent à aucun ordinal strictement plus petit. Ainsi les ordinaux finis sont des cardinaux ainsi que ω (ces assertions claires en théorie naïve des ensembles, par contre en théorie des ensembles, les énoncés analogues requièrent une démonstration). Par contre l'ordinal $\omega + 1$ n'est pas un cardinal.

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

Démonstration. — Soit α un ordinal. L'ensemble des $x \in \alpha + 1$ en bijection avec α est non vide et son plus petit élément est un cardinal équipotent à α. Par le théorème de Zermelo tout ensemble est équipotent à un ordinal. Si deux ordinaux distincts sont en bijection, l'un des deux appartient à l'autre, donc au plus un est un cardinal. □

1.7.2. Proposition. — *Soient A et B deux ensembles. On suppose A non vide. Les conditions suivantes sont équivalentes:*

(1) $\mathrm{card}\,(A) \leq \mathrm{card}\,(B)$
(2) il existe une injection de A dans B
(3) il existe une surjection de B sur A.

Démonstration. — Pour (1) $\implies$ (2) on peut supposer que A et B sont des cardinaux, le résultat est alors clair. S'il existe une injection de A dans B et $\mathrm{card}\,(A) \not\leq \mathrm{card}\,(B)$ alors $\mathrm{card}\,(B) \leq \mathrm{card}\,(A)$ et il existe une injection de $B \to A$. Les ensembles A et B sont alors en bijection et $\mathrm{card}\,(A) = \mathrm{card}\,(B)$, d'où (2) $\implies$ (1). Pour (3) $\implies$ (2), si $g : B \to A$ est surjective, on munit B d'un bon ordre et on définit $f : A \to B$ par $f(x)$ est le plus petit élément de $g^{-1}(x)$. Pour la réciproque, on fixe x_0 dans A, et si $f : A \to B$ est injective, on envoie $x \in B$ sur $f^{-1}(x)$ si x est dans l'image de A, sur x_0 sinon. □

1.7.3. Lemme. — *Si A est un ensemble de cardinaux, alors $\beta = \sup_{\alpha \in A} \alpha$ est un cardinal.*

Démonstration. — Si λ est un ordinal $< \beta$, il existe $\gamma > \lambda$ dans A, donc $\mathrm{card}\,(\beta) \geq \mathrm{card}\,(\gamma) = \gamma > \lambda$. On en déduit que β est un cardinal. □

1.7.4. Lemme. — *Si $f : \beta \to \alpha$ est une application strictement croissante entre ordinaux, pour tout γ dans β, $f(\gamma) \geq \gamma$.*

Démonstration. — Par l'absurde. Sinon, soit γ le plus petit ordinal avec $f(\gamma) < \gamma$. Pour tout $\delta < \gamma$, $\delta \leq f(\delta) < f(\gamma)$. On a donc $\gamma \subset f(\gamma)$, soit $\gamma \leq f(\gamma)$. □

Notons que par le théorème de Cantor, pour tout cardinal α il existe un ensemble X tel que $\mathrm{card}\,(X) > \alpha$. L'ensemble des cardinaux $\beta \leq \mathrm{card}\,(X)$ tels que $\beta > \alpha$ est donc non vide et possède un plus petit élément. Il suit qu'il existe un plus petit cardinal $> \alpha$, on le note α^+ (pour éviter tout risque de confusion on note désormais $\gamma + 1$ le successeur d'un ordinal γ).

On définit une application strictement croissante $\aleph$ de la classe des ordinaux dans celle des cardinaux de la façon suivante. On pose $\aleph_0 = \omega$, $\aleph_{\alpha+1} = \aleph_\alpha^+$, si α est un ordinal limite on pose $\aleph_\alpha = \sup_{\beta<\alpha} \aleph_\beta$ (qui est un cardinal par le lemme 1.7.3).

1.7.5. Proposition. — *Tout cardinal infini est un $\aleph_\alpha$.*

Démonstration. — Du lemme on tire que pour tout cardinal infini λ, $\aleph_{\lambda+1} > \lambda$, donc il existe un plus petit ordinal α tel que $\aleph_\alpha > \lambda$. Ce ne peut être un ordinal limite ni 0, donc $\alpha = \beta + 1$ et on a $\aleph_\beta \leq \lambda < \aleph_{\beta+1}$. On en tire que $\lambda = \aleph_\beta$. □

1.8. Opérations sur les cardinaux. — Si X et Y sont des ensembles on note $X+Y$ la somme disjointe de X et Y, $X \times Y$ 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 la réunion disjointe de κ et λ, $\kappa\lambda$ le cardinal de $\kappa \times \lambda$, et κ^λ le cardinal de l'ensemble des applications de λ dans κ. Ces notations sont cohérentes car $\mathrm{card}\,(X + Y) = \mathrm{card}\,X + \mathrm{card}\,Y$, $\mathrm{card}\,(X \times Y) = \mathrm{card}\,X \mathrm{card}\,Y$ et $\mathrm{card}\,(X^Y) = \mathrm{card}\,X^{\mathrm{card}\,Y}$. Par contre, ces opérations ne doivent pas être confondues avec les opérations correspondantes sur les ordinaux. Si κ et λ sont finis, ces opérations coïncident avec les opérations correspondantes sur les entiers.

Quelques propriétés de vérification directe:

1.8.1. Proposition. — *(1) Si $\lambda \leq \mu$ et $\nu \leq \kappa$, alors $\lambda + \nu \leq \mu + \kappa$, $\lambda\nu \leq \mu\kappa$ et, si $\mu \neq 0$, $\lambda^\nu \leq \mu^\kappa$.*

(2) L'addition et le produit des cardinaux sont commutatifs et associatifs, le produit est distributif par rapport à l'addition, $\lambda^{\mu+\nu} = \lambda^\mu\lambda^\nu$, $\lambda^\nu\mu^\nu = (\lambda\mu)^\nu$, $(\lambda^\mu)^\nu = \lambda^{\mu\nu}$.

1.8.2. Proposition. — $\mathrm{card}\,(\mathbf{R}) = 2^{\aleph_0}$.

Démonstration. — On a une injection $h : 2^{\aleph_0} \to \mathbf{R}$ qui à une suite a_i de 0 et de 1 associe $\sum_i a_i 2^{-i}$ si le support de la suite est infini, et $2 + \sum_i a_i 2^{-i}$ si le support est fini. On en tire $\mathrm{card}\,(\mathbf{R}) \geq 2^{\aleph_0}$. D'autre part la fonction $1/\pi \mathrm{arctg}(x) + 1/2$ établit une bijection entre $\mathbf{R}$ et $]0,1[$, et l'image de h contient $]0,1[$. On en tire $\mathrm{card}\,(\mathbf{R}) \leq \mathrm{card}\,(h(2^{\aleph_0})) \leq 2^{\aleph_0}$. □

Un ensemble A est dénombrable si $\mathrm{card}\,(A) = \omega$.

1.8.3. Lemme. — *Si α est un ordinal infini,* $\mathrm{card}\,(\alpha) = \mathrm{card}\,(\alpha + 1)$.

Démonstration. — Si $\alpha = \omega$, on considère $f : \mathbf{N} \to \mathbf{N} \cup \{\infty\}$ envoyant 0 sur ∞ et $i > 0$ sur $i - 1$. C'est une bijection, d'où le résultat. Le cas général s'en déduit en écrivant $\alpha = \omega + \alpha'$. □

Maintenant un résultat non trivial.

1.8.4. Proposition. — *Pour tout cardinal infini λ on a $\lambda \cdot \lambda = \lambda$.*

Démonstration. — Par induction : on suppose que pour tout cardinal infini $\mu < \lambda$ on a $\mu \cdot \mu = \mu$. On munit $\lambda \times \lambda$ de l'ordre suivant $(\beta, \gamma) < (\beta_1, \gamma_1)$ si $\sup(\beta, \gamma) < \sup(\beta_1, \gamma_1)$ ou $\sup(\beta, \gamma) = \sup(\beta_1, \gamma_1)$ et $\beta < \beta_1$ ou $\sup(\beta, \gamma) = \sup(\beta_1, \gamma_1)$, $\beta = \beta_1$ et $\gamma < \gamma_1$. On vérifie que c'est un bon ordre. Il existe donc un ordinal α et un isomorphisme d'ensembles ordonnés $f : \alpha \to \lambda \times \lambda$. Si $\alpha > \lambda$, alors $\lambda \in \alpha$. Dans ce cas, on écrit $f(\lambda) = (\beta_0, \gamma_0)$ et la restriction de f à λ induit une bijection entre λ et $Y = \{(\beta, \gamma) < (\beta_0, \gamma_0)\}$. Soit $\delta_0 = \sup(\beta_0, \gamma_0) + 1$. On a $\mathrm{card}\,(\delta_0) < \lambda$ (par le lemme) et $Y \subset \delta_0 \cdot \delta_0$. En appliquant l'hypothèse on en tire que $\mathrm{card}\,(Y) < \lambda$, contradiction. On a donc $\lambda \cdot \lambda \leq \lambda$, l'inégalité dans l'autre sens est claire. □

La bijection $\alpha_2 : \mathbf{N}^2 \to \mathbf{N}$ donnée par $\alpha_2(p, n) = 1/2(n+p+1)(n+p)+n$ fournit une preuve directe de l'égalité $\omega \cdot \omega = \omega$.

1.8.5. Proposition. — *(1) Si λ est infini alors $\lambda + \lambda = \lambda$.*
(2) Si X et Y sont des ensembles non vides dont l'un au moins est infini

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

(3) Si $(X_i)_{i \in I}$ est une famille d'ensembles dont l'un au moins est infini, alors $\mathrm{card}\,(\cup_{i \in I} X_i) \leq \sup(\mathrm{card}\,(X_i), \mathrm{card}\,(I))$. *Si, de plus les X_i sont tous non vides, alors on a égalité. En particulier, une réunion dénombrable d'ensembles dénombrables est dénombrable.*

Démonstration. — Pour (1) on remarque que $\lambda \leq \lambda + \lambda \leq \lambda \cdot \lambda = \lambda$.

Démontrons (2). Soit $\lambda = \sup(\mathrm{card}\,(X), \mathrm{card}\,(Y))$. Certainement, $\lambda \leq \mathrm{card}\,(X \cup Y)$ et $\lambda \leq \mathrm{card}\,(X \times Y)$. Dans l'autre sens $\mathrm{card}\,(X \times Y) \leq \lambda \cdot \lambda = \lambda$ et $\mathrm{card}\,(X \cup Y) \leq \lambda + \lambda = \lambda$.

Démontrons (3). Par l'axiome du choix il existe une application $f : X = \cup X_i \to I$ telle que $x \in X_{f(x)}$ pour tout x (munir I d'un bon ordre et poser $f(x) = \inf\{i \in I; x \in X_i\}$). Pour chaque i l'ensemble Y_i des applications injectives $X_i \to \lambda := \sup(\mathrm{card}\,(X_i), \mathrm{card}\,(I))$ est non vide. Par l'axiome du choix il existe donc une famille $(g_i)_{i \in I}$ d'applications injectives $g_i : X_i \to \lambda$. On en déduit une application injective $X \to I \times \lambda$ donnée par $x \mapsto (f(x), g_{f(x)}(x))$. Il suit que $\mathrm{card}\,(X) \leq \lambda \cdot \lambda = \lambda$. L'égalité quand les X_i sont tous non vides est claire. □

***1.8.6. Théorème* (Théorème de König).** — *Soient* $(X_i)_{i\in I}$ *et* $(Y_i)_{i\in I}$ *deux familles d'ensembles dont on suppose que* $\mathrm{card}\,(X_i) < \mathrm{card}\,(Y_i)$ *pour tout* i. *Alors*

$$\mathrm{card}\,(\cup_{i\in I} X_i) < \mathrm{card}\,(\prod_{i\in I} Y_i).$$

Démonstration. — On pose $X = \cup_{i\in I} X_i$ et $Y = \prod_{i\in I} Y_i$. Soit $f : X \to Y$. Pour chaque i, f induit une application $f_i : X_i \to Y_i$ donnée par la i-ème composante de la restriction de f à X_i. Par hypothèse le complémentaire B_i de l'image de f_i dans Y_i n'est pas vide. Par l'axiome du choix il existe b dans $\prod_{i\in I} B_i$. Certainement b n'est pas dans l'image de f. □

1.9. Cofinalité: $\aleph_\omega \neq 2^{\aleph_0}$. — Pour démontrer que $\aleph_\omega \neq 2^{\aleph_0}$ on va introduire la notion de cofinalité.

Soit A un ensemble muni d'un ordre total $<$. On dit qu'une partie $B \subset A$ est cofinale dans A si B est non borné dans A, autrement dit, si pour tout $a \in A$ il existe b dans B avec $b \geq a$.

Soit $f : \beta \to \alpha$ une application entre ordinaux. On dit que f est cofinale si l'image de f est cofinale dans α.

Soit α un ordinal. La cofinalité de α est le plus petit ordinal β tel qu'il existe une application cofinale strictement croissante $\beta \to \alpha$. On note $\mathrm{cf}(\alpha)$ la cofinalité de α.

1.9.1. Exemples. — (0) $\mathrm{cf}(\alpha) \leq \alpha$.
(1) $\mathrm{cf}(0) = 0$.
(2) Si α est un ordinal $\mathrm{cf}(\alpha + 1) = 1$.
(3) Notons que $\mathrm{cf}(\omega) = \omega$.

D'après le lemme suivant, on peut supprimer la condition "strictement croissante" dans la définition :

1.9.2. Lemme. — *Pour tout ordinal* α, *l'ordinal* $\mathrm{cf}(\alpha)$ *est le plus petit ordinal* θ *tel qu'il existe une application cofinale* $\theta \to \alpha$.

Démonstration. — Il suffit de démontrer que s'il existe $f : \beta \to \alpha$ une application entre ordinaux cofinale dans α, il existe une application strictement croissante cofinale $g : \beta' \to \alpha$ avec β' un ordinal $\leq \beta$. Si $\alpha = 0$ ou est un ordinal successeur le résultat est clair. Supposons donc que α est un ordinal limite.

Pour cela on construit une fonction $\varphi : \beta \to \alpha + 1$ par induction de la façon suivante. (On peut supposer β non vide.) On pose $\varphi(0) = f(0)$. Maintenant on pose $\kappa(\lambda) = \sup_{\lambda'<\lambda}(\varphi(\lambda'))$ pour $\lambda < \beta$. Si $\kappa(\lambda) \geq \alpha$ on pose $\varphi(\lambda) = \alpha$ tandis que si $\kappa(\lambda) < \alpha$ on définit $\varphi(\lambda)$ comme le plus petit élément de l'ensemble non vide $f(\beta) \cap \{\alpha' \in \alpha; \alpha' \geq \kappa(\lambda) + 1\}$ (c'est ici que l'on utilise que α est un ordinal limite). On pose $\beta' = \{\lambda \in \beta; \varphi(\lambda) < \alpha\}$ et on note g la restriction de φ à β'. Par construction g est strictement croissante et cofinale dans α. □

1.9.3. Proposition. — *Pour tout ordinal* α, *on a* $\mathrm{cf}(\alpha) \leq \alpha$, $\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha)$ *et* $\mathrm{cf}(\alpha)$ *est un cardinal.*

Démonstration. — Le premier énoncé est clair. On en tire que $\mathrm{cf}(\mathrm{cf}(\alpha)) \leq \mathrm{cf}(\alpha)$. Notons que si C est cofinal dans B et B est cofinal dans A, alors C est cofinal dans A. Il suit que $\mathrm{cf}(\mathrm{cf}(\alpha))$ est cofinal dans α et donc $\mathrm{cf}(\mathrm{cf}(\alpha)) \geq \mathrm{cf}(\alpha)$.

Soit β un ordinal qui n'est pas un cardinal. Il existe alors une application surjective $\theta \to \beta$ avec θ un ordinal $< \beta$. Par le lemme $\mathrm{cf}(\beta) \leq \theta$, donc $\mathrm{cf}(\beta) < \beta$. Il en résulte que $\mathrm{cf}(\alpha)$ est un cardinal. □

1.9.4. Proposition. — *Si λ est un ordinal limite,*

$$\mathrm{cf}(\aleph_\lambda) = \mathrm{cf}(\lambda).$$

Démonstration. — On a $\aleph_\lambda = \sup\{\aleph_\alpha; \alpha < \lambda\}$. Si $(a_\gamma)_{\gamma<\theta}$ est une suite strictement croissante cofinale dans λ, la suite $(\aleph_{a_\gamma})_{\gamma<\theta}$ est strictement croissante cofinale dans $\aleph_\lambda$ donc $\mathrm{cf}(\aleph_\lambda) \leq \mathrm{cf}(\lambda)$. Réciproquement, si $(b_\gamma)_{\gamma<\theta}$ est une suite strictement croissante cofinale dans $\aleph_\lambda$, on définit une suite cofinale non nécessairement strictement croissante par $\aleph_{a_\gamma} = \mathrm{card}\,(b_\gamma)_{\gamma<\theta}$ et on en déduit $\mathrm{cf}(\aleph_\lambda) \geq \mathrm{cf}(\lambda)$ par le lemme. □

Un cardinal infini κ est dit régulier si $\mathrm{cf}(\kappa) = \kappa$, singulier sinon.

1.9.5. Proposition. — *Tout cardinal infini successeur est régulier.*

Démonstration. — Si $\kappa = \lambda^+$, et $f : \alpha \to \kappa$ est cofinale avec $\alpha < \kappa$, alors, par le lemme 1.9.6

$$\kappa = \bigcup_{\beta\in\alpha} f(\beta)$$

serait la réunion de $\leq \lambda$ ensembles chacun de cardinalité $\leq \lambda$ et donc serait de cardinal $\leq \lambda$, absurde. □

1.9.6. Lemme. — *Soit $f : \alpha \to \kappa$ une application cofinale entre ordinaux. On suppose que κ est un ordinal limite. Alors $\kappa = \cup_{\beta\in\alpha} f(\beta)$.*

Démonstration. — En effet, $\delta = \cup_{\beta\in\alpha} f(\beta)$ est un ordinal $\geq \kappa$. Si $\delta \neq \kappa$, alors $\delta \in \kappa$ et il existe $\beta \in \alpha$ avec $f(\beta) \geq \theta + 1 > \theta$ par cofinalité et le fait que κ soit un ordinal limite, contradiction. □

On ne connait pas de cardinal limite régulier autre que $\aleph_0$. Plus généralement on dit qu'un cardinal λ est fortement limite si $2^\mu < \lambda$ pour tout $\mu < \lambda$ et on dit que λ est inaccessible s'il est régulier, fortement limite et $> \aleph_0$. L'existence de cardinaux inaccessibles est indépendante des axiomes ZFC mais est néanmoins parfois très utile en mathématiques. On peut la voir comme une généralisation de l'axiome de l'infini.

1.9.7. Proposition. — *Pour tout cardinal $\kappa \geq 2$ et tout cardinal infini λ on a*

$$\mathrm{cf}(\kappa^\lambda) > \lambda.$$

Démonstration. — Soit $f : \theta \to \kappa^\lambda$ une application avec $\mathrm{card}\,(\theta) \leq \lambda$ et $\theta \neq 0$. D'après le théorème de König,

$$\mathrm{card}\,(\cup_{\alpha<\theta} f(\alpha)) < \mathrm{card}\,(\prod_{\alpha<\theta}(\kappa^\lambda)) = (\kappa^\lambda)^{\mathrm{card}\,(\theta)} = \kappa^{\lambda\cdot\mathrm{card}\,(\theta)} = \kappa^\lambda,$$

en particulier, f ne peut pas être cofinale d'après le lemme 1.9.6. □

1.9.8. Corollaire. — *On a*

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

Démonstration. — En effet, $\mathrm{cf}(\aleph_\omega) = \omega$ tandis que $\mathrm{cf}(2^{\aleph_0}) > \omega$. □

2. Théorie des modèles

2.1. Langages. — Un langage est composé

(a) d'un ensemble infini dénombrable de variables $\mathcal{V} = \{v_0, \cdots, v_n, \cdots\}$,
(b) des symboles logiques $\neg$ (négation), $\wedge$ (et), $\vee$ (ou), $\Longrightarrow$, $\Longleftrightarrow$, $\forall$ et $\exists$,
(c) d'une suite d'ensembles $\mathcal{F}_n$, $n \in \mathbf{N}$,
(d) d'une suite d'ensembles $\mathcal{R}_n$, $n \in \mathbf{N}$.

Les éléments de $\mathcal{F}_n$ sont appelés symboles de fonctions n-aires, ceux de $\mathcal{R}_n$ sont appelés symboles de relations n-aires. Un symbole fonction 0-aire est une constante. De plus, on suppose que $\mathcal{R}_0$ contient un symbole $\top$ (qui sera interprété comme l'énoncé toujours vrai).

Le langage L est formé de la réunion de ces ensembles de symboles. Un mot du langage L est une suite finie $m = (a_0, \cdots, a_k)$ d'éléments de L. On écrira aussi $m = a_0 a_1 \cdots a_k$. On note L^* l'ensemble des mots de L.

2.2. Termes. — L'ensemble $\mathcal{T}(L)$ des termes du langage L est le plus petit ensemble de L^* qui contient les variables et les symboles de constante et tel que si f est dans $\mathcal{F}_n$, t_1, …, t_n sont dans $\mathcal{T}(L)$, alors $ft_1 \cdots t_n$ est un terme.

Notons $\mathcal{T}_0(L)$ l'ensemble des variables et des symboles de constante, $\mathcal{T}_{k+1}(L)$ la réunion de $\mathcal{T}_k(L)$ et des $ft_1 \cdots t_n$ avec f dans $\mathcal{F}_n$ et les t_i dans $\mathcal{T}_k(L)$. Alors $\mathcal{T}(L) = \cup_{n\in\mathbf{N}} \mathcal{T}_k(L)$. On appelle hauteur d'un terme t le plus petit entier k tel que t appartienne à $\mathcal{T}_k(L)$.

On a la propriété suivante de lecture unique (de vérification facile, laissée en exercice) :

2.2.1. Proposition. — *Tout terme t de $\mathcal{T}(L)$ vérifie une et une seulement des trois possibilités suivantes :*

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

2.2.2. Notation. — On écrira dorénavant $f(t_1, \cdots, t_n)$ pour $ft_1 \cdots t_n$.

2.2.3. Notation. — Si t est un terme, on écrira $t = t[v_{i_1}, \ldots, v_{i_n}]$ si toutes les variables ayant au moins une occurence dans t figurent parmi les v_{i_j}.

2.2.4. Substitution dans les termes. — Si w_1, …, w_n sont des variables distinctes et u_1, …, u_n des termes, t un terme, on définit, par récurrence sur la hauteur de t, un terme $t_{u_1/w_1,\ldots,u_n/w_n}$ en substituant les u_i aux w_i.

2.3. Formules. — Une formule atomique est un mot de la forme $Rt_1 \cdots t_n$ avec R une relation n-aire et les t_i des termes. On écrira $Rt_1 \cdots t_n = R(t_1, \ldots, t_n)$.

On définit par récurrence les formules de hauteur m de la façon suivante : les formules atomiques sont de hauteur 0, si F est une formule de hauteur m, alors $\neg F$, $\forall v_n F$ et $\exists v_n F$ sont des formules de hauteur $m+1$, si de plus G est une formule de hauteur m', alors αFG - notée $(F\alpha G)$ - est une formule de hauteur $\sup(m, m') + 1$, pour $\alpha \in \{\wedge, \implies, \iff\}$. L'ensemble des formules dans le langage L est constitué des formules de hauteur m pour $m \in \mathbf{N}$. C'est le plus petit sous-ensemble de L^* contenant les formules atomiques et tels que si F et G sont des formules $\neg F$, $\forall v_n F$ et $\exists v_n F$, $(F \wedge G)$, $(F \vee G)$, $(F \implies G)$, et $(F \iff G)$ sont des formules.

On a un énoncé d'unicité similaire à celui pour les termes (de vérification facile, laissée en exercice) :

2.3.1. Proposition. — *Tout formule F de $\mathcal{T}(L)$ vérifie une et une seulement des cinq possibilités suivantes :*

(a) F est une formule atomique et s'écrit de façon unique comme $R(t_1, \ldots, t_n)$.
(b) F est de la forme $\neg G$ avec G une unique formule.
(c) F est de la forme αGH avec G et H d'uniques formules et α unique dans $\{\wedge, \vee, \implies, \iff\}$.
(d) F est de la forme $\forall v_n G$ avec n et G uniques.
(e) F est de la forme $\exists v_n G$ avec n et G uniques.

2.3.2. Variables libres, liées. — Soit v_k une variable. Si F est atomique toutes les occurences de v_k dans F sont libres. Si $F = \neg G$ les occurences libres de v_k dans F sont celles de v_k dans G. Les occurences libres de v_k dans $(F\alpha G)$ sont celles dans F et celles dans G. Si $F = \forall v_h G$ ou $\exists v_h G$ avec $h \neq k$, les occurences libres de v_k dans F sont celles de v_k dans G. Si $h = k$, aucune des occurences de v_k dans F n'est libre. Les occurences non libres d'une variable sont dites liées.

Les variables libres dans F sont celles admettant au moins une occurence libre. Un énoncé est une formule sans variable libre.

On écrit $F[v_{i_1}, \ldots, v_{i_n}]$ si les variables libres de F figurent parmi les v_{i_j} (supposés distincts).

Soit F une formule et v une variable ayant une occurence liée dans F. Alors considérons la sous-formule QvG de F correspondante. Les occurences libres de v dans G ainsi que l'occurence de v suivant le quantificateur Q sont appelées occurences de v figurant dans le champ du quantificateur Qv.

2.4. Substitution dans les formules. — Soit F une formule, $w_1, \ldots, w_n$ des variables distinctes et $u_1, \ldots, u_n$ des termes, on définit une expression $F_{u_1/w_1,\ldots,u_n/w_n}$ "en remplaçant w_i par u_i dans les variable libres". Ainsi si $F = R(t_1, \ldots, t_n)$ est atomique, $F_{u_1/w_1,\ldots,u_n/w_n}$ est obtenue par substitution dans chaque terme t_i, si $F = \neg G$, on effectue la substitution dans G, idem pour $(F \alpha G)$ et $\forall v G$, $\exists v G$, si v ne figure pas parmi les w_i. Par contre dans $\forall w_i G$ et $\exists w_i G$ on se contente de substituer les u_j aux w_j pour $j \neq i$.

2.5. Structures. —

2.5.1. Structures. — Soit L un langage. Une L-structure $\mathfrak{M}$ est un ensemble non vide M muni, pour chaque symbole de fonction n-aire f de L d'une fonction $f^{\mathfrak{M}} : M^n \to M$ et pour chaque symbole de relation n-aire R d'une partie $R^{\mathfrak{M}}$ de M^n. On demande que $\top^{\mathfrak{M}} = M^0$.

2.5.2. Interprétation d'un terme dans une structure. — Soit $t[w_0, \ldots, w_n]$ un terme. Si $a_0, \ldots, a_n$ sont dans M, on note $t[a_0, \ldots, a_n]$ l'interprétation de t dans M avec w_i interprété par a_i: si $t = w_i$, c'est a_i. Si t est la constante c, c'est $c^{\mathfrak{M}}$. Si $t = f(t_1, \ldots t_r)$, c'est $f^{\mathfrak{M}}(t_1[a_0, \ldots, a_n], \ldots, t_r[a_0, \ldots, a_n])$.

2.5.3. Satisfaction d'une formule dans une structure. — Soit $F[w_1, \ldots, w_n]$ une formule. Si $a_0, \ldots, a_n$ sont dans M, on dit que $(a_1, \ldots, a_n)$ satisfait F dans $\mathfrak{M}$, et on note $\mathfrak{M} \models F[a_1, \ldots, a_n]$, si la formule obtenue en interprétant w_i par a_i est satisfaite dans M:

(1) si $F = R(t_1, \ldots, t_r)$, $\mathfrak{M} \models F[a_1, \ldots, a_n]$ si $R^{\mathfrak{M}}(t_1[a_0, \ldots, a_n], \ldots, t_r[a_0, \ldots, a_n])$ est vérifiée.
(2) $\mathfrak{M} \models \neg F[a_1, \ldots, a_n]$ ssi $\mathfrak{M} \not\models F[a_1, \ldots, a_n]$.
(3) $\mathfrak{M} \models (F \wedge G)[a_1, \ldots, a_n]$ ssi $\mathfrak{M} \models F[a_1, \ldots, a_n]$ et $\mathfrak{M} \models G[a_1, \ldots, a_n]$.
(4) $\mathfrak{M} \models (F \vee G)[a_1, \ldots, a_n]$ ssi $\mathfrak{M} \models F[a_1, \ldots, a_n]$ ou $\mathfrak{M} \models G[a_1, \ldots, a_n]$.
(5) Idem pour $\Longrightarrow$ et $\Longleftrightarrow$.
(6) Si $F = (\forall v G)[w_1, \ldots, w_n]$ et v différent des w_i, alors $G = G[v, w_1, \ldots, w_n]$ et $\mathfrak{M} \models (\forall v G)[a_1, \ldots, a_n]$, si pour tout a dans M, $\mathfrak{M} \models G[a, a_1, \ldots, a_n]$, idem pour $\exists$.
(7) Si $F = (\forall w_i G)[w_1, \ldots, w_n]$, alors $\mathfrak{M} \models (\forall w_i G)[a_1, \ldots, a_n]$, si pour tout a dans M, $\mathfrak{M} \models F[a_1, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_n]$. Idem pour $\exists$.

Si F est un énoncé, alors soit $M \models F$, soit $M \not\models F$. Si $F[v_1, \cdots, v_n]$ est une formule on écrit $M \models F[v_1, \cdots, v_n]$ si, pour tout $(a_1, \cdots, a_n) \in M^n$, $M \models F[a_1, \cdots, a_n]$. Si F est une formule de la forme $F[v_1, \cdots, v_n]$ on appelle clôture universelle de F l'énoncé $F' = \forall v_1 \cdots \forall v_n F[v_1, \cdots, v_n]$. Noter la non unicité de la clôture universelle (il y'en a une pour chaque ensemble fini ordonné contenant les variables libres de F). On a $M \models F[v_1, \cdots, v_n]$ si et seulement si $M \models F'$ (noter que cette équivalence ne vaudrait pas pour des structures vides).

2.5.4. Vocabulaire. — Un énoncé de L est dit universellement valide s'il est vérifié dans toute L-structure. Une formule $F[v_1, \cdots, v_n]$ de L est universellement valide si

une (et donc toute) clôture universelle l'est. Deux formules F et G sont logiquement équivalentes si $F \iff G$ est universellement valide.

Une théorie est un ensemble d'énoncés de L. Une L-structure $\mathfrak{M}$ est un modèle d'une théorie T si tous les énoncés de T sont satisfaits dans $\mathfrak{M}$. Si T admet au moins un modèle on dit que T est consistante.

2.6. Tautologie. — Pour définir les tautologies il est commode d'introduire le langage L_P du calcul propositionnel. Il est constitué d'un ensemble dénombrable de symboles P_i, $i \in \mathbb{N}$, ainsi que des symboles logiques $\neg, \wedge, \vee, \implies$ et $\iff$ et du symbole $\top$.

L'ensemble des formules du calcul propositionnel est le plus petit sous-ensemble de L_P^* contenant les P_i et $\top$ et tel que si M et N sont des formules du calcul propositionnel alors $\neg M$, $M \wedge N$, $M \vee N$, $M \implies N$, et $M \iff N$ aussi. Notons $\mathcal{P}$ l'ensemble des formules du calcul propositionnel. Considérons

$$\theta : \mathcal{P} \longrightarrow \mathbf{Z}/2\mathbf{Z}[X_0, \ldots, X_n, \ldots]$$

envoyant $\top$ sur 1, P_i sur X_i, $\neg M$ sur $1 + \theta(M)$, $M \wedge N$ sur $\theta(M)\theta(N)$, $M \vee N$ sur $\theta(M) + \theta(N) + \theta(M)\theta(N)$, $M \implies N$ sur $1 + \theta(M) + \theta(M)\theta(N)$ et $M \iff N$ sur $1 + \theta(M) + \theta(N)$. Une tautologie du calcul propositonnel est une formule J de $\mathcal{P}$ tel que le polynôme $\theta(J)$ ne prend que la valeur 1 sur $\{0,1\}^{\mathbf{N}}$. Autrement dit J est toujours vrai quelle que soit la valeur de vérité des P_i.

Maintenant si $J[P_0, \ldots, P_n]$ est une tautologie du calcul propositionnel formée à partir des variables P_i, et si $F_0, \ldots, F_n$ sont des formules du langage L, alors la formule $J[F_0, \ldots, F_n]$ obtenue par substitution des F_i aux P_i est appelé tautologie dans L. Une tautologie est clairement une formule universellement valide.

2.7. Modèles égalitaires. — On suppose ici que L contient un symbole $=$. Une L-structure $\mathfrak{M}$ est dite égalitaire si $=^{\mathfrak{M}}$ correspond à la vraie égalité dans M. Un modèle égalitaire d'une théorie dans L est un modèle qui est une L-structure égalitaire.

On considère la théorie E formée des énoncés suivants :

(a) $\forall v(v = v)$
(b) $\forall v_0 \forall v_1 (v_0 = v_1 \implies v_1 = v_0)$
(c) $\forall v_0 \forall v_1 \forall v_2 ((v_0 = v_1 \wedge v_1 = v_2) \implies v_0 = v_2)$
(d) pour chaque symbole de fonction n-aire f de l'énoncé exprimant que si $v_i = v_{i+n}$, pour tout i, $f(v_1, \ldots, v_n) = f(v_{n+1}, \ldots, v_{2n})$
(e) pour chaque symbole de relation n-aire R de l'énoncé exprimant que si $v_i = v_{i+n}$, pour tout i, $R(v_1, \ldots, v_n)$ ssi $R(v_{n+1}, \ldots, v_{2n})$.

2.7.1. Proposition. — *T admet un modèle égalitaire ssi $T \cup E$ admet un modèle.*

Démonstration. — Le premier sens est trivial. Si $T \cup E$ admet un modèle $\mathcal{M}$ on obtient un modèle égalitaire en quotientant $\mathcal{M}$ par la relation d'équivalence $=^{\mathfrak{M}}$. □

2.8. Démonstrations formelles. — On veut donner un sens précis à "l'énoncé F est démontrable à partir de la théorie T".

Règles formelles. — Il y en a deux :

(1) Le modus ponens : à partir de F et de $F \implies G$, le modus ponens permet de déduire G.
(2) La règle de généralisation : à partir de F où F est une formule, on déduit $\forall v F$. (Si on sait démontrer $F(v)$, on sait démontrer $\forall v F$).

Axiomes logiques. — Ils sont de deux types :

(1) Les tautologies
(2) Les axiomes des quantificateurs
    (a) $\exists v F \iff \neg \forall v \neg F$ où F est une formule quelconque
    (b) $\forall v(F \implies G) \implies (F \implies \forall v G)$ pour F et G des formules et v une variable n'ayant pas d'occurence libre dans F.
    (c) $\forall v F \implies F_{t/v}$, pour toute formule F et tout terme t tels qu'aucune occurence libre de v dans F n'apparait dans le champ d'un quantificateur liant une variable de t.

2.8.1. Remarque. — La restriction sur les occurences libres de v dans (c) est nécessaire. En effet, prenons $F = \exists v_1 \neg R(v, v_1)$ et $t = v_1$ Alors $F_{v_1/v} = \exists v_1 \neg R(v_1, v_1)$ tandis que $\forall v F = \forall v \exists v_1 \neg R(v, v_1)$. La formule $\forall v F \implies F_{t/v}$ n'est pas satisfaite pour R interprété par l'égalité dans une structure de cardinal > 1.

2.8.2. Remarque. — La formule $\forall v(F \implies \forall v F)$ n'est pas universellement valide en général (notons que cela ne contredit pas le fait que la règle de généralisation soit raisonnable). En effet si $F = P(v)$ avec P une relation 1-aire, $\forall v(F \implies \forall v F)$ est logiquement équivalente à $\exists v P(v) \implies \forall v P(v)$ qui n'est pas universellement valide.

2.8.3. Définition. — Soit T une théorie et F une formule de L. Une preuve formelle de F dans T est une suite finie $(F_0, \ldots, F_n)$ avec $F_n = F$, telle que pour chaque i, soit $F_i \in T$, soit F_i est un axiome logique, soit F_i se déduit à partir d'une ou deux formules précédentes dans la suite par l'une des règles de déduction.

S'il existe une démonstration de F dans T on dit que F est une conséquence syntaxique de T et on écrit $T \vdash F$. Si F est une formule telle que pour tout modèle $\mathfrak{M}$ de T, on a $\mathfrak{M} \models F$, on dit que F est conséquence sémantique de T et on écrit $T \models F$.

Les règles de déduction et les axiomes logiques ont été choisis de façon à ce que l'énoncé suivant soit vérifié :

2.8.4. Lemme. — *Soit T une théorie et soit F une formule de L. Si $T \vdash F$, alors $T \models F$.*

Démonstration. — Vérification directe laissée au lecteur. □

Toutefois on aurait pu choisir d'autres règles et obtenir peut-être une autre notion de conséquence syntaxique. Il résultera du théorème de complétude de Gödel que $T \vdash F$ ssi $T \models F$, ce qui justifiera a posteriori nos choix.

Si $T \vdash F$ avec T la théorie vide, on écrit $\vdash F$.

2.8.5. Exemples. — (1) Si F et G sont des énoncés, alors $\{F, G\} \vdash F \wedge G$.
(2) Soit F une formule et t un terme tels qu'aucune occurence libre de v dans F n'apparaît dans le champ d'un quantificateur liant une variable de t. Alors $\vdash F_{t/v} \implies \exists v F$.
(3) Si la variable w n'a aucune occurence dans F, alors $\vdash \forall w F_{w/v} \implies \forall v F$.
(4) $\vdash \forall v F \implies F$.

Démonstration. — (1) A partir de F et de la tautologie $F \implies (G \implies (F \wedge G))$ on démontre $G \implies (F \wedge G)$ par modus ponens, d'où l'on déduit $F \wedge G$ par modus ponens.

(2) Utiliser l'axiome $\forall v \neg F \implies \neg F_{t/v}$ et la tautologie $(\forall v \neg F \implies \neg F_{t/v}) \implies (F_{t/v} \implies \neg \forall v \neg F)$.

(3) De l'axiome des quantificateurs (c) on déduit $\forall w F_{w/v} \implies (F_{w/v})_{v/w}$. Mais $(F_{w/v})_{v/w} = F$, d'où par généralisation $\forall v(\forall w F_{w/v} \implies F)$ et on conclut par l'axiome des quantificateurs (b) et modus ponens.

(4) C'est une conséquence de l'axiome des quantificateurs (c) car $F_{v/v} = F$. □

2.9. Cohérence. — On dit que T est cohérente s'il n'existe pas de formule F telle que $T \vdash F$ et $T \vdash \neg F$. Si T n'est pas cohérente alors $T \vdash G$ pour toute formule G (utiliser la tautologie $F \implies (\neg F \implies G)$).

Notons que si $T \vdash F$, alors il existe $T_0 \subset T$ finie telle que $T_0 \vdash F$.

2.9.1. Corollaire. — *Si toutes les parties finies de T sont cohérentes, T aussi.*

2.9.2. Corollaire. — *Soient T_i des théories, $i \in I$, avec $T_i \subset T_j$ ou $T_j \subset T_i$ pour tout i et j. Si les T_i sont cohérentes, $T = \cup_{i \in I} T_i$ aussi.*

2.9.3. Lemme **(Lemme de déduction)**. — *Soit F un énoncé. Si $T \cup \{F\} \vdash G$, alors $T \vdash F \implies G$.*

Démonstration. — Soit $G_0, \ldots, G_n$ une démonstration de G dans $T \cup \{F\}$. On va obtenir une démonstration de $F \implies G$ dans T en faisant des insertions dans la suite $F \implies G_0, \ldots, F \implies G_n$ et en effectuant une récurrence sur n. Notons que si G_i est une tautologie, $F \implies G_i$ aussi. Si G_i est un axiome de quantificateurs ou est dans T, on démontre $F \implies G_i$ à partir de G_i et de la tautologie $G_i \implies (F \implies G_i)$. Supposons que G_i se déduise par modus ponens de G_j et de $G_k = (G_j \implies G_i)$. On considère alors la tautologie

$$(F \implies G_j) \implies ((F \implies (G_j \implies G_i)) \implies (F \implies G_i)).$$

Par modus ponens on en déduit $(F \implies (G_j \implies G_i)) \implies (F \implies G_i)$. Comme $F \implies (G_j \implies G_i)$ par hypothèse de récurrence, on conclut par modus ponens. Supposons enfin que $G_i = \forall v G_j$. Par généralisation on déduit $\forall v(F \implies G_j)$ à

partir de $F \implies G_j$. Comme $\forall v(F \implies G_j) \implies (F \implies \forall v G_j)$ est un axiome de quantificateurs (F est un énoncé), on conclut par modus ponens. □

2.9.4. Corollaire. — *$T \vdash F$ ssi $T \cup \{\neg F\}$ n'est pas cohérente.*

Démonstration. — En effet, si $T \vdash F$, $T \cup \{\neg F\}$ n'est pas cohérente. Réciproquement, si $T \cup \{\neg F\}$ n'est pas cohérente, elle démontre F. Par le lemme 2.9.3, $T \vdash \neg F \implies F$. Comme $(\neg F \implies F) \implies F$ est une tautologie, on en tire $T \vdash F$. □

Autre lemme utile :

2.9.5. Lemme. — *Soit T une théorie, soit v une variable et soit $F[v]$ une formule ayant au plus une variable libre. Soit c un symbole de constante n'apparaissant ni dans T ni dans F. Si $T \vdash F_{c/v}$, alors $T \vdash \forall v F$.*

Démonstration. — Soit $G_0, \ldots, G_n$ une démonstration de $F_{c/v}$ dans T. Soit w une variable n'apparaissant dans aucune des formules G_i et soit K_i obtenue en remplaçant c par w dans G_i. Si G_i est un axiome logique, K_i aussi. Si G_i se déduit des précédentes par modus ponens ou généralisation, K_i aussi. Enfin, si G_i est dans T, $K_i = G_i$ aussi. Il en résulte que $K_0, \ldots, K_n$ est une démonstration de $F_{w/v}$ dans T. Par généralisation on en déduit $T \vdash \forall w F_{w/v}$ et donc $T \vdash \forall v F$ par l'exemple (3). □

2.9.6. Corollaire. — *Soit T une théorie, soit v une variable et soit $F[v]$ une formule ayant au plus une variable libre. Soit c un symbole de constante n'apparaissant ni dans T ni dans F. Si T est cohérente, alors $T \cup \{\exists v F[v] \implies F_{c/v}\}$ est cohérente.*

Démonstration. — En effet, si $T \cup \{\exists v F[v] \implies F_{c/v}\}$ n'est pas cohérente $T \vdash \neg(\exists v F[v] \implies F_{c/v})$, autrement dit, $T \vdash \exists v F[v] \wedge \neg F_{c/v}$. Mais alors par le lemme 2.9.5 et les axiomes de quantificateurs, $T \vdash \exists v F[v] \wedge \forall v \neg F[v]$, ce qui contredit la cohérence de T. □

2.9.7. Lemme. — *Soit T une théorie et soit F un énoncé dans un langage L. Soit L' un langage contenant L et tel que $L' \setminus L$ soit uniquement composé de symboles constantes. Alors T démontre F dans L' si et seulement si T démontre F dans L. En particulier T est cohérente dans L si et seulement si elle est cohérente dans L'.*

Démonstration. — On procède comme dans la preuve du lemme 2.9.5. Soit $G_0, \ldots, G_n$ une démonstration de F à partir de T dans L'. Pour chaque constante c_j de $L' \setminus L$ apparaissant dans un G_i on introduit une nouvelle variable w_j n'apparaissant dans aucune des formules G_i et on considère K_i obtenue en remplaçant les c_j par les w_j dans G_i. Pour les mêmes raisons que dans la preuve du lemme 2.9.5 $K_0, \ldots, K_n$ est une démonstration de F à partir de T dans le langage L. □

2.10. Construction de modèles. — On dit qu'une théorie T est complète dans un langage L si elle est cohérente et si pour tout énoncé F dans L, $T \vdash F$ ou $T \vdash \neg F$.

On dit que T a des témoins de Henkin dans L si, pour toute variable v et toute formule $F = F[v]$, il existe un symbole de constante c dans L tel que $(\exists v F[v] \implies F[c]) \in T$.

2.10.1. Théorème. — *Toute théorie T complète et admettant des témoins de Henkin admet un modèle.*

Démonstration. — On prend comme ensemble M l'ensemble des termes clos (c'est-à-dire dans lesquels aucune variable n'apparaît) de L. Cet ensemble est non vide car il existe au moins un symbole de constante (par exemple c dans L tel que $(\exists v \top[v] \implies \top[c]) \in T$). Si c est un symbole de constante on l'interprète par lui-même. Si R est un symbole de relation n-aire on l'interprète par $(t_1, \ldots, t_n) \in R^{\mathfrak{M}}$ ssi $T \vdash R(t_1, \ldots, t_n)$. Si f est un symbole de fonctions n-aire, par définition $f^{\mathfrak{M}}(t_1, \ldots, t_n)$ est le terme $f(t_1, \ldots, t_n)$.

Il suit des définitions que si F est un énoncé atomique, on a $T \vdash F$ ssi $\mathfrak{M} \models F$. On va faire une induction sur la hauteur de F pour établir le même énoncé pour F arbitraire.

1) Supposons que $F = \neg G$ et que le résultat est connu pour G. On a $\mathfrak{M} \models F$ ssi $\mathfrak{M} \not\models G$. Or $\mathfrak{M} \not\models G$ ssi $T \not\vdash G$. Comme T est complète, le dernier énoncé équivaut à $T \vdash F$.

2) si $F = G \wedge H$ et que le résultat est connu pour G et H, il l'est pour F. Vérification immédiate.

3) si $F = G \vee H$, et le résultat est connu pour G et H, il l'est pour F : résulte de ce qui précède. Idem pour $F = (G \implies H)$ et $F = ((G \iff H)$.

4) $F = \forall v G$. Si $T \vdash \forall v G$, et si t est dans M, il résulte des axiomes logiques que $T \vdash G_{t/v}$ et donc que $\mathfrak{M} \models G(t)$, ceci pour tout t, donc $\mathfrak{M} \models \forall v G$.

Dans l'autre sens, si T ne démontre pas $\forall v G$, alors T démontre $\neg \forall v G$ et par conséquent aussi $\exists v \neg G$. Comme il existe des témoins de Henkin, il existe un symbole de constante c tel que $T \vdash \neg G(c)$ et donc $\mathfrak{M} \models \neg G(c)$, $\forall v G$ est donc faux dans $\mathfrak{M}$.

5) Le cas $F = \exists v G$ résulte du précédent et de 1). □

2.10.2. Proposition. — *Soit T une théorie cohérente dans un langage L. Alors il existe un langage L' contenant L et une théorie T' contenant T, complète dans le langage L' et admettant des témoins de Henkin dans L'.*

Démonstration. — Supposons tout d'abord le langage L fini ou dénombrable. On considère L' obtenu en rajoutant à L des nouveaux symboles de constantes $c_0, \ldots, c_n, \ldots$. Le langage L' est dénombrable et on peut énumérer les énoncés de L' en $F_0, \ldots, F_n, \ldots$.

On pose $T_0 = T$. Par le lemme 2.9.7, T_0 est cohérente dans L'. On va construire pas récurrence sur n une suite croissante T_n de théories cohérentes dans L' avec $T_n - T$ fini, F_n ou $\neg F_n$ dans T_{n+1}, et si $F_n = \exists v H$ est dans T_{n+1}, il existe un symbole de constante c tel que $H_{c/v}$ est dans T_{n+1}.

Supposons T_n construite. Si $T_n \cup \{F_n\}$ est cohérente, on pose $G_n = F_n$. Sinon, $T_n \vdash \neg F_n$, et on pose $G_n = \neg F_n$. Dans les deux cas $T_n \cup \{G_n\}$ est cohérente.

Si G_n n'est pas de la forme $\exists v H$, on pose $T_{n+1} = T_n \cup \{G_n\}$. Si $G_n = \exists v H$, on considère un symbole c_i avec i minimal n'apparaissant ni dans T_n ni dans G_n. On pose $T_{n+1} = T_n \cup \{G_n\} \cup \{H[c_i])\}$. Montrons que T_{n+1} est cohérente. Sinon, $T_n \cup \{\exists v H\} \vdash \neg H_{c/v}$. Mais par le lemme 2.9.5, il suit que $T_n \cup \{\exists v H\} \vdash \forall v \neg H$, ce qui contredit la cohérence de $T_n \cup \{\exists v H\}$.

On considère la réunion T' des T_n. C'est une théorie cohérente par le corollaire 2.9.2 et complète par construction. Montrons qu'elle admet des témoins de Henkin. Soit $H = H[v]$. Il existe n tel que $F_n = \exists v H$ et soit $\neg F_n \in T'$ soit il existe c tel que $H_{c/v} \in T'$. Dans les deux cas $T' \vdash \exists v H \implies H_{c/v}$, ce qui force $\exists v H \implies H_{c/v}$ à appartenir à T' car sinon $\neg \exists v H \implies H_{c/v}$ appartiendrait à T' ce qui contredirait la cohérence.

Dans le cas général, on introduit un nouveau symbole de constante c_F pour chaque formule $F[v]$ avec v une variable. On note $L_1 = \tilde{L}$ le langage ainsi obtenu. Par le corollaire 2.9.6, pour chaque famille finie F_i, $1 \leq i \leq n$, la théorie $T \cup \{\exists w_p F_p[w_p] \implies F_p[c_{F_p}]\}_{1 \leq p \leq n}$ est cohérente, donc la théorie $T_1 = \tilde{T} = T \cup \{\exists v F[v] \implies F[c_F]\}$ est cohérente dans L_1 par le corollaire 2.9.1 et le lemme 2.9.7.

On pose $L_i = \tilde{L}_{i-1}$ et $T_i = \tilde{T}_{i-1}$. On note L' la réunion des L_i et T' celle des T_i. Par application répétée du corollaire 2.9.1 et du lemme 2.9.7 chaque T_i est cohérent. Par le corollaire 2.9.2 on en tire que la théorie T' est cohérente. Par construction elle admet des témoins de Henkin dans le langage L'.

Il reste à démontrer que toute théorie T' dans un langage L' qui est cohérente est contenue dans une théorie T'' cohérente et complète dans le même langage (car si T' admet des témoins de Henkin, T'' aussi) .

Cela résulte du lemme de Zorn. On considère l'ensemble des théories cohérentes de L' contenant T' ordonné par l'inclusion. C'est un ensemble inductif, par le lemme de Zorn il existe un élément maximal T''. Soit F un énoncé n'appartenant pas à T''. Par maximalité $T'' \cup \{F\}$ n'est pas cohérente. Il suit que $T'' \vdash \neg F$. La théorie T'' est donc complète. □

On en tire le théorème de complétude de Gödel (1930) :

2.10.3. Théorème. — *Toute théorie cohérente a un modèle.*

Démonstration. — Soit T une théorie cohérente dans un langage L. Il existe un langage L' contenant L et une théorie T' contenant T, complète dans le langage L' et admettant des témoins de Henkin dans L'. La théorie T' admet un modèle $\mathfrak{M}$. Ce modèle peut aussi être vu comme une L-structure (cette L-structure est appelée réduite de $\mathfrak{M}$). C'est un modèle de T. □

On en tire

2.10.4. Proposition. — *Soit T une théorie et soit F un énoncé vrai dans tout modèle de T. Alors $T \vdash F$.*

Démonstration. — Si $T \nvdash F$, alors on a vu que $T \cup \{\neg F\}$ est cohérente, et donc admet un modèle. □

Il n'y a donc pas de différence entre conséquence sémantique et conséquence syntaxique.

Une autre conséquence est le théorème de compacité :

2.10.5. Théorème. — *Si T est une théorie dont toute partie finie a un modèle, alors T a un modèle.*

Démonstration. — Toute partie finie de T est cohérente, donc T est cohérente. Par le théorème de complétude T a un modèle. □

2.11. Sous-structures élémentaires. — Soit $\mathfrak{M}$ une L-structure. Soit N un sous-ensemble de M, contenant les $c^{\mathfrak{M}}$, c symbole de constante et les $f^{\mathfrak{M}}(N^k)$ pour f symboles de fonctions k-aires. On munit alors N d'une structure de L-structure $\mathfrak{N}$ en interprétant les symboles de relation k-aires par $R^{\mathfrak{N}} = R^{\mathfrak{M}} \cap N^k$, les constantes par $c^{\mathfrak{M}}$ et les fonctions f par les restrictions des $f^{\mathfrak{M}}$ à N^k. On dit que $\mathfrak{N}$ est une sous-structure de $\mathfrak{M}$.

On dit que $\mathfrak{N}$ est une sous-structure élémentaire de $\mathfrak{M}$ (ou que $\mathfrak{M}$ est une extension élémentaire de $\mathfrak{N}$) si pour toute formule $F[v_1, \ldots, v_n]$ de L et tous les $(a_1, \ldots, a_n)$ dans N^k, $\mathfrak{M} \models F[a_1, \ldots, a_n]$ si et seulement si $\mathfrak{N} \models F[a_1, \ldots, a_n]$.

Soit $\mathfrak{M}$ une L-structure. On note $Th(\mathfrak{M})$ l'ensemble des énoncés de L vérifiés dans $\mathfrak{M}$. Si $\mathfrak{N}$ est une autre L-structure et que $Th(\mathfrak{N}) = Th(\mathfrak{M})$ on dit que $\mathfrak{N}$ et $\mathfrak{M}$ sont élémentairement équivalentes. En particulier si $\mathfrak{N}$ est une sous-structure élémentaire de $\mathfrak{M}$, $\mathfrak{N}$ et $\mathfrak{M}$ sont élémentairement équivalentes.

Voici un critère permettant de vérifier qu'une sous-structure est élémentaire.

***2.11.1. Théorème* (Test de Tarski-Vaught).** — *Soit $\mathfrak{M}$ une structure et $\mathfrak{N}$ une sous-structure. On suppose que pour toute formule $F[v_0, \ldots, v_n]$ de L et tous les $(a_1, \ldots, a_n)$ dans N^k, si*

$$\mathfrak{M} \models \exists v_0 F[v_0, a_1, \ldots, a_n]$$

alors il existe a_0 dans N tel que

$$\mathfrak{M} \models F[a_0, \ldots, a_n].$$

Alors $\mathfrak{N}$ est une sous-structure élémentaire de $\mathfrak{M}$.

Démonstration. — On va montrer que pour toute formule $G[v_1, \ldots, v_n]$ de L et tous les $(a_1, \ldots, a_n)$ dans N^k, $\mathfrak{M} \models G[a_1, \ldots, a_n]$ si et seulement si $\mathfrak{N} \models G[a_1, \ldots, a_n]$ par induction sur G. Pour G atomique, c'est clair et le cas des connecteurs propositionnels ne pose pas de difficulté. En remplaçant G par une formule logiquement équivalente on se ramène au cas où $G = \exists v_0 F[v_0, v_1, \ldots, v_n]$.

Si $\mathfrak{N} \models \exists v_0 F[v_0, a_1, \ldots, a_n]$, il existe a_0 dans N tel que $\mathfrak{N} \models F[a_0, \ldots, a_n]$ et $\mathfrak{M} \models F[a_0, \ldots, a_n]$ par hypothèse, donc $\mathfrak{M} \models \exists v_0 F[v_0, a_1, \ldots, a_n]$.

Réciproquement, si $\mathfrak{M} \models \exists v_0 F[v_0, a_1, \ldots, a_n]$, $\mathfrak{M} \models F[a_0, \ldots, a_n]$ par hypothèse, $\mathfrak{N} \models F[a_0, \ldots, a_n]$ et $\mathfrak{N} \models \exists v_0 F[v_0, a_1, \ldots, a_n]$. □

On note $\operatorname{card} L$ le cardinal de l'ensemble des formules de L.
On va en déduire le théorème de Löwenheim-Skolem descendant:

2.11.2. Théorème. — *Soit $\mathfrak{M}$ une L-structure, A une partie de M. On suppose que $\operatorname{card} M \geq \operatorname{card} L$. Alors il existe une sous-structure élémentaire $\mathfrak{M}_0$ de $\mathfrak{M}$ contenant A de cardinalité $\sup(\operatorname{card} A, \operatorname{card} L)$.*

Démonstration. — Quitte à agrandir A on peut supposer que $\operatorname{card} A \geq \operatorname{card} L$. Notons que si B est une partie de M de cardinal $\geq L$, la sous-structure engendrée par B a même cardinal que B. On considère A_0 la structure engendrée par A. Etant définie A_i on construit A_{i+1} comme suit. Pour chaque formule $F[v_0, \ldots, v_n]$ de L et chaque suite $(a_1, \ldots, a_n)$ dans A_i, si $\mathfrak{M} \models \exists v_0 F[v_0, a_1, \ldots, a_n]$, on choisit $c(F, a_1, \ldots, a_n)$ dans M tel que $\mathfrak{M} \models F[c(F, a_1, \ldots, a_n), a_1, \ldots, a_n]$. On considère B_i obtenu en rajoutant tous les $c(F, a_1, \ldots, a_n)$ à A_i et on définit A_{i+1} comme la sous-structure engendrée par B_i. On définit $\mathfrak{M}_0$ comme la réunion des A_i. C'est une sous-structure de cardinal $\operatorname{card} A$. Elle est élémentaire par le test de Tarski-Vaught. □

2.12. Diagrammes. — Soit $\mathfrak{M}$ une L-structure égalitaire. On considère le langage L_M obtenu en rajoutant à L une constante pour chaque élément de M. La structure $\mathfrak{M}$ s'enrichit naturellement en une L_M-structure notée $\mathfrak{M}^*$.

Le diagramme complet $D(\mathfrak{M})$ est la théorie complète de $\mathfrak{M}^*$. Il est formé des $F[a_1, \ldots, a_n]$ avec $F[v_1, \ldots, v_n]$ formule de L, a_1, …, a_n dans M et $\mathfrak{M} \models F[a_1, \ldots, a_n]$.

2.12.1. Proposition. — *Le réduit au langage L d'un modèle de $D(\mathfrak{M})$ est - à isomorphisme près - une extension élémentaire de $\mathfrak{M}$. Réciproquement toute extension élémentaire de $\mathfrak{M}$ est le réduit au langage L d'un modèle de $D(\mathfrak{M})$.*

Démonstration. — Soit $\mathfrak{N}^*$ un modèle de $D(\mathfrak{M})$ de réduit au langage $\mathfrak{N}$. L'application $M \to N$ qui à un élément de M associe son interprétation dans N est injective (si m et m' sont distincts alors $m \neq m'$ appartient à $D(\mathfrak{M})$. $\mathfrak{N}$ est donc isomorphe à une extension de $\mathfrak{M}$ qui est clairement élémentaire. Réciproquement, si $\mathfrak{N}$ est une extension élémentaire de $\mathfrak{N}$, l'inclusion $M \subset N$ permet d'étendre $\mathfrak{N}$ en une L_M-structure $\mathfrak{N}^*$ qui est un modèle de $D(\mathfrak{M})$. □

De façon similaire, on considère le diagramme simple $\Delta(\mathfrak{M})$ formé des $F[a_1, \ldots, a_n]$ avec $F[v_1, \ldots, v_n]$ formule sans quantificateur de L, a_1, …, a_n dans M et $\mathfrak{M} \models F[a_1, \ldots, a_n]$.

2.12.2. Proposition. — *Les réduits au langage L des modèle de $\Delta(\mathfrak{M})$ correspondent à isomorphisme près aux extensions de $\mathfrak{M}$.*

Démonstration. — Preuve similaire à celle de la Proposition 2.12.1. □

On peut maintenant démontrer le théorème de Löwenheim-Skolem ascendant:

2.12.3. Théorème. — *Soit $\mathfrak{M}$ une L-structure infinie et κ un cardinal tel que $\kappa \geq \sup(\operatorname{card} M, \operatorname{card} L)$. Alors il existe une extension élémentaire de $\mathfrak{M}$ de cardinal κ.*

Démonstration. — Quitte à agrandir le langage on peut supposer que $\mathfrak{M}$ est égalitaire. Il suffit de construire une extension élémentaire de $\mathfrak{M}$ de cardinal $\geq \kappa$. et d'appliquer le théorème de Löwenheim-Skolem descendant. Pour chaque $i \in \kappa$ on considère un nouveau symbole de constante c_i et on considère le langage L^* obtenu en rajoutant les c_i à L. On introduit la théorie T formée de $D(\mathfrak{M})$ et des énoncés $\neg c_i = c_j$ pour $i \neq j$. Cette théorie est consistante car toute partie finie de T a un modèle (à savoir $\mathfrak{M}$, les c_i étant interprétés par des points distincts de $\mathfrak{M}$). Il existe donc un modèle de T' dont le réduit est (isomorphe à) une extension élémentaire de $\mathfrak{M}$ et dont le cardinal est $\geq \kappa$. □

2.12.4. Corollaire. — *Soit T une théorie. Si T a un modèle infini, elle a un modèle en tout cardinal $\geq \operatorname{card} L$.*

2.13. Elimination des quantificateurs. — Le résultat suivant donne une caractérisation sémantique des formules équivalentes à une formule sans quantificateurs dans une théorie.

2.13.1. Théorème. — *Soit T une théorie égalitaire et soit $F[\underline{v}]$ une formule à m variables libres. Les deux conditions suivantes sont équivalentes :*

1) Il existe une formule sans quantificateur $G[\underline{v}]$ telle que $T \models \forall \underline{v}(F[\underline{v}] \iff G[\underline{v}])$.

2) Si $\mathcal{A}$ est une L-structure, $\mathfrak{M}$ et $\mathfrak{N}$ sont des modèles de T contenant chacun $\mathcal{A}$, pour tout uplet $\underline{a}$ de A, $\mathcal{M} \models F[\underline{a}]$ si et seulement si $\mathcal{N} \models F[\underline{a}]$.

Démonstration. — 1) $\implies$ 2) est clair car une formule sans quantificateur est vérifiée dans un modèle si et seulement si elle est vérifiée dans une sous-structure.

Pour la réciproque, on considère l'ensemble $\Gamma(\underline{v})$ formé des formules sans quantificateur $G[\underline{v}]$ telles que $T \models \forall \underline{v}(F[\underline{v}] \implies G[\underline{v}])$. On rajoute des nouvelles constantes de symbole $d_1, \ldots, d_m$.

Montrons que $T \cup \Gamma(\underline{d}) \models F[\underline{d}]$. Sinon, il existerait $\mathfrak{M}$ tel que $\mathfrak{M} \models T \cup \Gamma(\underline{d}) \cup \{\neg F[\underline{d}]\}$. Considérons la sous-structure $\mathcal{A}$ de $\mathfrak{M}$ engendrée par $\underline{d}$. On considère $\Sigma = T \cup \Delta(\mathcal{A}) \cup \{F[\underline{d}]\}$. Montrons que Σ a un modèle. En effet, si Σ n'a pas de modèle, il existe par compacité des formules sans quantificateurs $g_1[\underline{d}], \ldots, g_n[\underline{d}]$ dans $\Delta(\mathcal{A})$ telles que

$$T \models \wedge_i g_i[\underline{d}] \implies \neg F[\underline{d}].$$

Comme les d_i n'apparaissent pas dans T ni dans L, il suit que

$$T \models \forall \underline{v}(\wedge_i g_i[\underline{v}] \implies \neg F[\underline{v}]).$$

Mais alors

$$T \models \forall \underline{v}(F[\underline{v}] \implies \vee_i \neg g_i[\underline{v}])$$

et $\vee_i \neg g_i[\underline{v}]$ appartient à $\Gamma(\underline{v})$ et donc $\mathcal{A} \models \vee_i \neg g_i[\underline{d}]$, contradiction. Il existe donc un modèle $\mathfrak{N}$ de Σ contenant $\mathcal{A}$. On a $\mathcal{M} \models \neg F[\underline{d}]$ tandis que $\mathcal{N} \models F[\underline{d}]$, ce qui contredit 2).

On a donc montré que $T \cup \Gamma(\underline{d}) \models F[\underline{d}]$. il existe par compacité des formules sans quantificateurs $g_1[\underline{d}], \ldots, g_n[\underline{d}]$ dans $\Gamma(\underline{d})$ telles que

$$T \models \wedge_i g_i[\underline{d}] \implies F[\underline{d}].$$

il suit que

$$T \models \forall \underline{v}(\wedge_i g_i[\underline{v}] \implies F[\underline{v}]).$$

Mais alors

$$T \models \forall \underline{v}(\wedge_i g_i[\underline{v}] \iff F[\underline{v}])$$

et $\wedge_i g_i[\underline{v}]$ est sans quantificateurs. □

On dit que la théorie T admet l'élimination des quantificateurs dans le langage L si toute formule de L est équivalente à une formule sans quantificateurs dans T.

On a le critère très simple suivant :

2.13.2. Lemme. — *Si pour toute formule sans quantificateur $\theta(v, \underline{w})$ il existe une formule sans quantificateur $\psi(\underline{w})$ telle que $T \models \exists v \theta(v, \underline{w}) \iff \psi(\underline{w})$, alors T admet l'élimination des quantificateurs.*

Démonstration. — Comme $\forall v\theta$ est logiquement équivalente à $\neg\exists v \neg\theta$, l'hypothèse du lemme vaut aussi quand on remplace $\exists$ par $\forall$. Comme toute formule est équivalente à une formule sous forme prénexe, c'est à dire de la forme $Q_1 x_1 \cdots Q_n x_n \varphi$ avec $Q_1, \ldots, Q_n$ des quantificateurs $\exists$ ou $\forall$, on conclut par récurrence sur n. □

Compte-tenu du théorème précédent on en tire:

2.13.3. Théorème. — *Soit T une théorie égalitaire. On suppose que pour toute paire de modèles $\mathfrak{M}$ et $\mathfrak{N}$ de T, pour toute sous-structure $\mathcal{A}$ commune à $\mathfrak{M}$ et $\mathfrak{N}$ et toute formule sans quantificateur $\theta(v, \underline{w})$, s'il existe $\underline{a}$ un uplet de A et b dans $\mathfrak{M}$ tels que $\mathcal{M} \models \theta(b, \underline{a})$, alors il existe c dans $\mathfrak{N}$ tel que $\mathcal{N} \models \theta(c, \underline{a})$. Alors T admet l'élimination des quantificateurs.*

2.14. Corps algébriquement clos. — On se place dans le langage des anneaux $L = \{+, -, \cdot, 0, 1, =\}$. Dans ce langage on considère la théorie T_{ac} formée des axiomes des corps et pour chaque entier $n \geq 2$ de l'énoncé exprimant que tout polynôme de degré n a une racine. Un corps algébriquement clos est un modèle de T_{ac}.

On utilisera les énoncés et notions suivants qui seront vus en Algèbre 1. Soit A un sous-anneau d'un corps K. Un élément de K est dit algébrique sur A s'il est racine d'un polynôme non nul à coefficients dans A. Une clôture algébrique d'un anneau intègre A est un corps algébriquement clos contenant A et dont tous les éléments sont algébriques sur A. Si K est un corps algébriquement clos contenant A, l'ensemble des éléments de K algébriques sur A est une clôture algébrique de A. Enfin, si K et K' sont deux clôtures algébriques de A, il existe un isomorphisme de

corps entre K et K' dont la restriction à A est l'identité de A. Enfin tout anneau intègre admet une clôture algébrique.

2.14.1. Théorème. — *La théorie T_{ac} admet l'élimination des quantificateurs dans le langage des anneaux.*

Démonstration. — Il suffit de démontrer que si K et F sont des corps algébriquement clos avec $F \subset K$ et $\theta(v, \underline{w})$ est une formule sans quantificateur, pour tout $\underline{a}$ uplet de F, s'il existe b dans K tel que $K \models \theta(b, \underline{a})$, alors il existe c dans F tel que $F \models \theta(c, \underline{a})$. En effet si A est une sous-structure de K_1 et K_2 algébriquement clos, A est un anneau intègre. Soient F_1 et F_2 la clôture algébrique de A dans K_1 et K_2 respectivement. Soit a un uplet dans A. S'il existe b dans K_1 tel que $K_1 \models \theta(b, \underline{a})$ alors il existe c_1 dans F_1 tel que $F_1 \models \theta(c_1, \underline{a})$. Par isomorphisme, il existe c_2 dans F_2 tel que $F_2 \models \theta(c_2, \underline{a})$ et nécessairement $K_2 \models \theta(c_2, \underline{a})$.

Remarquons que θ est équivalente logiquement à une formule de la forme $\vee_i \wedge_j \theta_{i,j}(v, \underline{w})$ avec les $\theta_{i,j}$ atomiques ou négations d'atomiques. Maintenant, si $K \models \theta(b, \underline{a})$, il existe i tel que $K \models \wedge_j \theta_{i,j}(b, \underline{a})$, donc on peut supposer que θ est une conjonction de formules atomiques ou négations d'atomiques. Dans le langage des anneaux toute formule atomique est de la forme $p(v_1, \cdots, v_n) = 0$ avec p un polynôme à coefficients entiers. On peut donc supposer que $\theta(v, \underline{a})$ est de la forme

$$\wedge_{i=1}^{n} p_i(v) = 0 \wedge \wedge_{i=1}^{m} q_i(v) \neq 0$$

avec p_i et q_i des polynômes à coefficients dans F. Si l'un des p_i est non nul, b est algébrique sur A, donc appartient à F. On peut donc supposer que $\theta(v, \underline{a})$ est de la forme $\wedge_{i=1}^{m} q_i(v) \neq 0$. Comme chaque polynôme q_i n'a qu'un nombre fini de racines et que F est infini, il existe c dans F tel que $\theta(c, \underline{a})$. □

Pour p un nombre premier on note $T_{\mathrm{ac},p}$ la théorie T_{ac} augmentée de l'énoncé $p = 0$ et on note $T_{\mathrm{ac},0}$ la théorie T_{ac} augmentée des énoncés $p \neq 0$, p parcourant l'ensemble des nombres premiers.

2.14.2. Corollaire. — *Les théories $T_{\mathrm{ac},p}$ et $T_{\mathrm{ac},0}$ sont complètes.*

Démonstration. — Notons $\bar{\mathbf{F}}_p$ une clôture algébrique du corps fini à p éléments $\mathbf{F}_p$. La théorie $\mathrm{Th}(\bar{\mathbf{F}}_p)$ formée des énoncés vrais dans $\bar{\mathbf{F}}_p$ est complète par définition. Tout corps algébriquement clos de caractéristique p K contient $\mathbf{F}_p$ donc également un sous-corps isomorphe à $\bar{\mathbf{F}}_p$. Par élimination des quantificateurs $\mathrm{Th}(\bar{\mathbf{F}}_p) = \mathrm{Th}(K)$. Il s'ensuit que tous les énoncés de $\mathrm{Th}(\bar{\mathbf{F}}_p)$ sont des conséquences de $T_{\mathrm{ac},p}$ qui est donc complète. En raisonnant de même en remplaçant $\mathbf{F}_p$ par $\mathbf{Q}$ on obtient l'énoncé pour $T_{\mathrm{ac},0}$. □

***2.14.3. Théorème* (Principe de Lefschetz).** — *Soit φ un énoncé dans le langage des anneaux. les énoncés suivant sont équivalents*

1. *L'énoncé φ est vérifié dans $\mathbf{C}$.*
2. *L'énoncé φ est vérifié dans un corps algébriquement clos de caractéristique 0.*
3. *L'énoncé φ est vérifié dans tout corps algébriquement clos de caractéristique 0.*

4. *L'énoncé φ est vérifié dans tout corps algébriquement clos de caractéristique p avec p suffisamment grand.*
5. *Il existe un ensemble infini de nombres premiers $\mathcal{P}$, tel que pour chaque p dans $\mathcal{P}$ φ est vérifié pour au moins un corps algébriquement clos de caractéristique p.*

Démonstration. — L'équivalence de 1), 2) et 3) résulte de la complétude. Montrons 3) $\implies$ 4). Dans ce cas $T_{\mathrm{ac},0} \models \varphi$. Par compacité il existe une partie finie Δ de $T_{\mathrm{ac},0}$ telle que $\Delta \models \varphi$. Comme pour p suffisament grand $T_{\mathrm{ac},p} \models \Delta$, on déduit 4). Comme 4) $\implies$ 5) il reste à montrer que 5) $\implies$ 3). Supposons qu'il existe un ensemble infini de nombres premiers $\mathcal{P}$, tel que φ est vérifié pour p dans $\mathcal{P}$ pour au moins un corps algébriquement clos de caractéristique p. Si $T_{\mathrm{ac},0} \not\models \varphi$, $T_{\mathrm{ac},0} \models \neg\varphi$ par complétude. Mais alors $\neg\varphi$ est vérifiée dans tout corps algébriquement clos de caractéristique p avec p suffisament grand, absurde. □

Donnons une application remarquable, le théorème d'Ax.

***2.14.4. Théorème* (Ax).** — *Soit $f : \mathbf{C}^n \to \mathbf{C}^n$ une application polynomiale, c'est à dire de la forme $f = (f_1, \cdots, f_n)$, $f_i \in \mathbf{C}[X_1, \cdots, X_n]$. Si f est injective, alors f est surjective.*

Démonstration. — Il suffit de démontrer l'énoncé correspondant pour le corps $\bar{\mathbf{F}}_p$ pour p premier. En effet pour toute paire d'entiers n et d il existe un énoncé du premier ordre $\Phi_{n,d}$ exprimant dans un corps K que pour toute famille $f_1, \cdots, f_n$ de polynômes en n variables de degré $\leq d$ si l'application associée $K^n \to K^n$ est injective alors elle est surjective. Si $\Phi_{n,d}$ est une conséquence de $T_{\mathrm{ac},p}$ pour tout p, alors c'est une conséquence de $T_{\mathrm{ac},0}$.

Pour tout entier $k \geq 1$, l'ensemble des racines du polynôme $X^{p^k} - X$ dans $\bar{\mathbf{F}}_p$ est un sous-corps $\mathbf{F}_{p^k}$ de cardinal p^k de $\bar{\mathbf{F}}_p$, $\mathbf{F}_{p^k}$ est un sous-corps de $\mathbf{F}_{(p^k)^r}$ et $\bar{\mathbf{F}}_p$ est la réunion de ses sous-corps $\mathbf{F}_{p^k}$. En particulier toute partie finie de $\bar{\mathbf{F}}_p$ est contenue dans un sous-corps fini de $\bar{\mathbf{F}}_p$.

Soit $f = (f_1, \cdots, f_n) : \bar{\mathbf{F}}_p^n \to \bar{\mathbf{F}}_p^n$ une application polynomiale injective. Supposons que f n'est pas surjective. Soit b dans $\bar{\mathbf{F}}_p^n$ qui n'est pas dans l'image de f. Soit k un sous-corps fini de $\bar{\mathbf{F}}_p$ contenant les composantes de b ainsi que les coefficients des polynômes f_i. L'application f induit une application $k^n \to k^n$ qui est injective et non surjective, ce qui est absurde vu que k^n est un ensemble fini. □

3. Récursivité

3.1. Fonctions primitives récursives. — Pour chaque entier p, on note $\mathcal{F}_p$ l'ensemble des fonctions $\mathbf{N}^p \to \mathbf{N}$. On note $\mathcal{F}$ la réunion des ensembles $\mathcal{F}_p$.

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

1. E contient pour tout entier p les fonctions constantes $\mathbf{N}^p \to \mathbf{N}$.

2. E contient pour tout entier p et tout $1 \leq i \leq p$ la projection $P_i^p : \mathbf{N}^p \to \mathbf{N}$ sur le i-ème facteur
3. E contient la fonction successeur $S : x \mapsto x+1$ de $\mathcal{F}_1$.
4. E est clos par composition : si $f_1, \cdots, f_n$ sont n fonctions de $\mathcal{F}_p$ appartenant à E et si h est une fonction de $\mathcal{F}_n$ appartenant à E, alors la fonction $h(f_1, \cdots, f_n)$ appartient à E.
5. E est clos par récurrence : pour tout entier p, si g est une fonction de $\mathcal{F}_p$ appartenant à E et si h est une fonction de $\mathcal{F}_{p+2}$ appartenant à E, alors la fonction f de $\mathcal{F}_{p+1}$ définie par $f(x_1, \cdots, x_p, 0) = g(x_1, \cdots, x_p)$ et $f(x_1, \cdots, x_p, y+1) = h(x_1, \cdots, x_p, y, f(x_1, \cdots, x_p, y))$ appartient à E.

Une partie de $\mathbf{N}^p$ est primitive récursive si sa fonction caractéristique l'est.

3.1.2. Exemples. — Les fonctions $(x,y) \mapsto x+y$, $(x,y) \mapsto x \times y$, $(x,y) \mapsto x^y$, la fonction $x \dot{-} 1$ définie par $0 \dot{-} 1 = 0$ et $x+1 \dot{-} 1 = x$ est primitive récursive, ainsi que la fonction $x \dot{-} y$ définie par $x \dot{-} 0 = x$ et $x \dot{-} (y+1) = (x \dot{-} y) \dot{-} 1$. La fonction sg valant 0 en 0 et 1 pour $x > 0$ est primitive récursive (elle est égale à $x \mapsto 1 \dot{-} (1 \dot{-} x)$). Le prédicat $x > y$ est primitif récursif car sa fonction caractéristique est $\mathrm{sg}(x \dot{-} y)$.

3.1.3. Remarques. — (1) L'ensemble des fonctions primitives récursives est clos par permutation des variables.

(2) Si $A \subset \mathbf{N}^n$ est primitif récursif et $f_1, \ldots, f_n$ sont dans $\mathcal{F}_p$ et primitives récursives alors l'ensemble des x tels que $(f_1(x), \cdots, f_n(x)) \in A$ est primitif récursif car sa fonction caractéristique est $\mathbf{1}_A(f_1, \cdots, f_n)$.

(3) L'ensemble des parties primitives récursives de $\mathbf{N}^n$ est stable par réunion, intersection et complémentaire.

(4) Si f et g sont dans $\mathcal{F}_p$ primitives récursives et A est primitif récursif, alors la fonction $h = f\mathbf{1}_A + g\mathbf{1}_{\mathbf{N}^p \setminus A}$ aussi. Plus généralement pour une partition de $\mathbf{N}^p$ en un nombre fini de parties primitives récursives. En particulier les fonctions $\sup(x_1, \cdots, x_p)$ et $\inf(x_1, \cdots, x_p)$ sont primitives récursives.

(5) Sommes et produits limités : si f est dans $\mathcal{F}_{p+1}$ est primitive récursive les fonctions $(x_1, \cdots, x_p, y) \mapsto \sum_{0 \leq t \leq y} f(x_1, \cdots, x_p, t)$ et $(x_1, \cdots, x_p, y) \mapsto \prod_{0 \leq t \leq y} f(x_1, \cdots, x_p, t)$ également.

3.1.4. Le schéma μ borné. — Soit A primitif récursif dans $\mathbf{N}^{p+1}$. Alors la fonction $f = \mu t \leq z((x_1, \cdots, x_p, t) \in A)$ définie par $f(x_1, \cdots, x_p, z) = 0$ s'il n'existe pas de $t \leq z$ tel que $(x_1, \cdots, x_p, t) \in A$ et par $f(x_1, \cdots, x_p, z) = t$ si t est le plus petit entier $\leq z$ tel que $(x_1, \cdots, x_p, t) \in A$ est primitive récursive. En effet f est définie par $f(x_1, \cdots, x_p, 0) = 0$, $f(x_1, \cdots, x_p, z+1) = f(x_1, \cdots, x_p, z)$ si $\sum_{0 \leq t \leq z} \mathbf{1}_A(x_1, \cdots, x_p, t) \geq 1$, $f(x_1, \cdots, x_p, z+1) = z+1$ sinon et si $(x_1, \cdots, x_p, z+1) \in A$ et $f(x_1, \cdots, x_p, z+1) = 0$ sinon.

3.1.5. Quantification bornée. — Si A est primitif récursif dans $\mathbf{N}^{p+1}$ il en est de même de $\{(x_1, \cdots, x_p, z), \exists t \leq z(x_1, \cdots, x_p, t) \in A\}$ et de $\{(x_1, \cdots, x_p, z), \forall t \leq z(x_1, \cdots, x_p, t) \in A\}$.

3.1.6. Exemples. — (1) La fonction $q(x,y)$ qui est égale à partie entière de x/y si y non nul et à 0 sinon est primitive récursive (c'est $\mu t \leq x((t+1)y > x)$.

(2) L'ensemble $\{(x,y); y \,\text{divise}\, x\}$ est primitif récursif car y divise x ssi $x = yq(x,y)$

(3) L'ensemble des nombres premiers est primitif récursif

(4) La fonction π qui à un entier n associe le $n+1$-ème nombre premier est primitive récursive. Elle est définie par $\pi(0) = 2$ et $\pi(n+1) = \mu z \leq (\pi(n)! + 1)(z > \pi(n) \,\text{et}\, z \,\text{est premier})$.

(5) Une bijection récursive entre $\mathbf{N}^2$ et $\mathbf{N}$: $\alpha_2((p,n)) = 1/2(n+p+1)(n+p) + n$. Son inverse est aussi donnée par des fonctions récursives. En effet notons que $\alpha_2((p,n))$ est au moins égal à n et p. On peut donc considérer la fonction primitive récursive $\beta_2^1(x) = \mu z \leq x(\exists t \leq x \alpha_2(z,t) = x)$ et de même $\beta_2^2(x)$. On considère $\alpha_p(x_1, \cdots, x_{p+1}) = \alpha_p(x_1, \cdots, x_{p-1}, \alpha_2(x_p, x_{p+1}))$ qui établit un isomorphisme entre $\mathbf{N}^{p+1}$ et $\mathbf{N}$ la i-ème composante de l'inverse étant une fonction primitive récursive β_{p+1}^i.

3.2. La fonction d'Ackermann. — On considère la fonction ξ à deux variables définie par $\xi(0,x) = 2^x$, $\xi(y,0) = 1$ et $\xi(y+1, x+1) = \xi(y, \xi(y+1,x))$.

On écrit aussi $\xi_n(x)$ pour $\xi(n,x)$.

Notons que $\xi_n(x) > x$ par récurrence sur n. Pour $n = 0$ c'est clair. Pour $n > 0$ on va faire une récurrence sur x. On a $\xi_n(x+1) = \xi_{n-1}(\xi_n(x)) > \xi_n(x) > x$.

Notons que $\xi_n(x)$ est strictement croissante. Enfin $\xi_n(x) \geq \xi_{n-1}(x)$ (car $\xi_n(x+1) = \xi_{n-1}(\xi_n(x)) \geq \xi_{n-1}(x+1)$.

On note ξ_n^k la fonction ξ_n itérée k fois. Il est clair que les fonctions ξ_n^k sont strictement croissantes, $\xi_n^k < \xi_n^{k+1}$, $\xi_n^k(x) \geq x$, et $\xi_m^k \leq \xi_n^k$ si $m \leq n$.

On dit que la fonction f dans $\mathcal{F}_1$ domine g de $\mathcal{F}_p$ s'il existe A entier tel que pour tout uplet $(x_1, \cdot, x_p)$, $g(x_1, \cdot, x_p) \leq f(\sup(x_i, A))$. Si f est strictement croissante f domine g ssi $g(x_1, \cdot, x_p) \leq f(\sup(x_i))$ sauf pour un ensemble fini de uplets.

On note C_n l'ensemble des fonctions dominées par au moins une fonction ξ_n^k. Notons que C_0 contient les fonctions constantes, successeur et les projections. Il est facile de vérifier que C_n est clos par composition.

Montrons que si g et h sont dans C_n, la fonction f définie par récurrence à partir de g et h est dans C_{n+1}. On a $f(x_1, \cdots, x_p, 0) = g(x_1, \cdots, x_p)$ et $f(x_1, \cdots, x_p, y+1) = h(x_1, \cdots, x_p, y, f(x_1, \cdots, x_p, y))$. On a $g(x) \leq \xi_n^{k_1}(\sup(x_i, A_1))$ et $h(x,y,t) \leq \xi_n^{k_2}(\sup(x_i, y, t, A_2))$. Par récurrence sur y, on vérifie que $f(x_1, \cdots, x_p, y) \leq \xi_n^{k_1 + yk_2}(\sup(x_i, y, t, A_1, A_2))$. On remarque (récurrence sur $k \geq 1$) que $\xi_n^k(x) \leq \xi_{n+1}(x+k)$, d'où

$$f(x_1, \cdots, x_p, y) \leq \xi_{n+1}(\sup(x_i, y, A_1, A_2) + k_1 + k_2 y)$$

or la fonction majorante est composée de fonctions de C_{n+1}.

On en déduit que $C = \cup_n C_n$ contient l'ensemble des fonctions primitives récursives.

Montrons maintenant que ξ n'est pas primitive. Sinon la fonction $x \mapsto \xi(x, 2x)$ le serait également. Il existerait alors des entiers A, n et k tels que $\xi(x, 2x) \leq \xi_n^k(x)$

pour $x > A$. On aurait alors $\xi_x(2x) \leq \xi_{n+1}(x+k)$ pour $x > A$, ce qui est impossible dès que $x > \sup(1, k, n+1)$.

3.3. Fonctions partielles récursives. — Une application partielle de $\mathbf{N}^p$ dans $\mathbf{N}$ est un couple (A, f) avec $A \subset \mathbf{N}^p$ et $f : A \to \mathbf{N}$. L'ensemble A est le domaine de définition de f. On note $\mathcal{F}_p^*$ l'ensemble de ces couples et $\mathcal{F}^*$ la réunion des $\mathcal{F}_p^*$.

Composition de fonctions partielles : Soient $f_1, \ldots, f_n$ dans $\mathcal{F}_n^*$ et g dans $\mathcal{F}_n^*$. La fonction composée $h = g(f_1, \cdots, f_n)$ a pour domaine de définition l'ensemble des uplets en lesquels les f_i et $g(f_1, \cdots, f_n)$ est définie.

Soit g une fonction partielle dans $\mathcal{F}_p^*$ et h une fonction partielle de $\mathcal{F}_{p+2}^*$, alors on définit la fonction partielle f de $\mathcal{F}_{p+1}^*$ par $f(x_1, \cdots, x_p, 0) = g(x_1, \cdots, x_p)$ si g est définie et $f(x_1, \cdots, x_p, y+1) = h(x_1, \cdots, x_p, y, f(x_1, \cdots, x_p, y)$ si $h(x_1, \cdots, x_p, y, f(x_1, \cdots, x_p, y)$ est définie.

Le schéma μ: soit $f \in \mathcal{F}_{p+1}^*$ Alors on définit la fonction partielle $g(x_1, \cdots, x_p) = \mu y(f(x_1, \cdots, x_p, y) = 0)$ par s'il existe un entier z tel que $f(x_1, \cdots, x_p, z) = 0$ et que $f(x_1, \cdots, x_p, z')$ est définie pour $z' \leq z$, alors $g(x_1, \cdots, x_p)$ est le plus petit de ces entiers z. Sinon g n'est pas définie.

L'ensemble des fonctions partielles récursives est le plus petit sous-ensemble de $\mathcal{F}^*$ contenant les fonctions totales constantes, les projections, la fonction successeur et clos par composition, définition par récurrence et schéma μ.

Un sous-ensemble de $\mathbf{N}^p$ est récursif si sa fonction caractéristique l'est.

3.4. Machines de Turing. — Une machine de Turing est composée d'un nombre fini de bandes horizontales et d'une tête de lecture. Chaque bande est bornée à gauche, illimitée à droite et divisée en cases successives numérotés de la façon suivante : 1 pour la case la plus à gauche, 2 pour la suivante à droite, etc. La tête peut lire, écrire ou effacer des symboles et se déplacer horizontalement. Elle peut écrire 3 symboles : le symbole de début de bande d, le symbole blanc b, et le bâton $|$. On pose $S := \{d, b, |\}$. Une machine de Turing est déterminée par le nombre n de ses bandes, un ensemble fini E d'états contenant l'état initial e_i et l'état final e_f et une table M qui est une application de $S^n \times E$ dans $S^n \times E \times \{-1, 0, +1\}$.

La machine fonctionne de la façon suivante. A l'instant $t = 0$, la tête se trouve devant les cases numéro 1 sur lesquelles est écrit le symbole d. Un symbole est écrit sur chaque case. La machine est dans l'état e_i. A l'instant t la machine lit les symboles s_1, ..., s_n écrits devant sa tête. Si elle est dans l'état e et si $M(s_1, \cdots, s_n, e) = (s_1', \cdots, s_n', e', \varepsilon)$, alors la machine efface les symboles s_i qu'elle remplace par les s_i', et elle déplace sa tête de ε cases vers la droite et elle passe à l'état e'. On est alors à l'instant $t+1$. La machine s'arrête quand elle est en l'état e_f.

Les contraintes sont les suivantes :

(1) La machine s'arrête en e_f, ie $M(s_1, \cdots, s_n, e_f) = (s_1, \cdots, s_n, e_f, 0)$.

(2) On ne peut effacer ou écrire le symbole d qui se trouve uniquement sur les cases numéro 1. Il n'est pas possible à la tête de se déplacer sur la gauche quand elle lit le symbole d.

(3) A l'instant $t = 0$ il n y a qu'un nombre fini de cases remplies par autre chose que b.

On dit qu'une bande représente un entier x si elle commence par un d suivi de x bâtons puis uniquement de symboles b.

3.5. Fonctions T-calculables. — Soit f une fonction dans $\mathcal{F}_p^*$ et $\mathcal{M}$ une machine de Turing possédant au moins $p+1$ bandes. On dit que $\mathcal{M}$ calcule f si pour tout suite d'entiers $x_1, \ldots, x_p$, si on fait fonctionner la machine $\mathcal{M}$ à partir de la configuration initiale où les p premières bandes représentent respectivement $x_1, \ldots, x_p$, les autres étant blanches (ie sans bâton), alors si $f(x_1, \cdots, x_n)$ n'est pas définie la machine ne s'arrête jamais (ie e_f n'est jamais atteint), si $f(x_1, \cdots, x_n)$ est définie, la machine s'arrête en un temps fini et son état final est le suivant : les p premières bandes représentent respectivement $x_1, \ldots, x_p$, la $p+1$-ème représente $f(x_1, \cdots, x_n)$ les autres étant blanches.

3.5.1. Lemme. — *La fonction successeur est T-calculable par une machine à deux bandes avec états $\{e_i, e_f\}$.*

Démonstration. — On considère une machine avec une table vérifiant $M(d, d, e_i) = (d, d, e_i, +1)$, $M(|, b, e_i) = (|, |, e_i, +1)$ et $M(b, b, e_i) = (b, |, e_f, 0)$. □

3.5.2. Lemme. — *La projection P_p^i est T-calculable par une machine à $p+1$ bandes avec états $\{e_i, e_f\}$.*

Démonstration. — On considère une machine avec une table vérifiant $M(d, \cdots, d, e_i) = (d, \cdots, d, e_i, +1)$, $M(s_1, \cdots, s_p, b, e_i) = (s_1, \cdots, s_p, |, e_i, +1)$ si $s_i = |$, $M(s_1, \cdots, s_p, b, e_i) = (s_1, \cdots, s_p, b, e_f, 0)$ si $s_i = b$. □

3.5.3. Lemme. — *La fonction constante égale à k en p variables est T-calculable par une machine à $p+1$ bandes avec états $\{e_i, e_f, e_1, \cdots, e_k\}$.*

Démonstration. — On considère une machine avec une table vérifiant $M(d, \cdots, d, e_i) = (d, \cdots, d, e_1, +1)$, $M(s_1, \cdots, s_p, b, e_n) = (s_1, \cdots, s_p, |, e_{n+1}, +1)$ pour n entre 1 et $k-1$, $M(s_1, \cdots, s_p, b, e_k) = (s_1, \cdots, s_p, |, e_f, 0)$. □

3.5.4. Lemme. — *L'ensemble des fonctions T-calculables est clos par composition.*

Démonstration. — Etant données des machines de Turing $\mathcal{M}_i$ à p_i bandes calculant f_i et une machine de Turing $\mathcal{N}$ à n' bandes calculant g, il s'agit de décrire une machine de Turing $\mathcal{M}$ à $p + \sum(p_i - p) + n' - n$ bandes calculant $g(f_1, \cdots, f_n)$. A l'instant initial on représente les x_i sur les p premières bandes les autres étant blanches. On calcule $f_1(x_1, \cdots, x_p)$ à l'aide de $\mathcal{M}_1$ en utilisant les p premières bandes et $p_1 - p$ autres bandes mais pas la $p+1$-ème. On calcule de même les $f_i(x_1, \cdots, x_p)$ en utilisant $p_i - p$ nouvelles bandes à chaque fois mais pas la $p+1$-ème. Une fois cela fait on calcule $g(f_1, \cdots, f_n)$ à partir des $f_i(x_1, \cdots, x_p)$ en utilisant $\mathcal{N}$ (et en renumérotant les bandes) et en utilisant la $p+1$-ème bande pour inscrire le résultat. □

3.5.5. Lemme. — *L'ensemble des fonctions T-calculables est clos par récurrence.*

Démonstration. — Soit g une fonction partielle dans $\mathcal{F}_p^*$ calculée par $\mathcal{M}$ avec $p+1+k$bandes et h une fonction partielle de $\mathcal{F}_{p+2}^*$ calculée par $\mathcal{M}'$ avec $p+3+k'$ bandes. On suppose les ensembles d'états disjoints. Alors on peut calculer la fonction partielle f de $\mathcal{F}_{p+1}^*$ définie par $f(x_1,\cdots,x_p,0) = g(x_1,\cdots,x_p)$ si g est définie et $f(x_1,\cdots,x_p,y+1) = h(x_1,\cdots,x_p,y,f(x_1,\cdots,x_p,y)$ si $h(x_1,\cdots,x_p,y,f(x_1,\cdots,x_p,y)$ est définie de la façon suivante. La machine $\mathcal{N}$ a $p+4+k+k'$ bandes et ses états sont formés de E, E' et de 8 nouveaux états e_0, ..., e_7. Pour commencer on calcule $g(x_1,\cdots,x_p)$ avec $\mathcal{M}$ en utilisant les bandes 1, ..., p, $p+4$ et k bandes calcul. Ensuite on passe à l'état e_0 et on calcule succesivement $f(x_1,\cdots,x_p,1)$, ..., $f(x_1,\cdots,x_p,x_{p+1})$. Quand on calcule $f(x_1,\cdots,x_p,y+1)$, y est codé sur la bande $p+2$, $f(x_1,\cdots,x_p,y)$ sur la bande $p+3$. En état e_0 elle transfère le contenu de la bande $p+4$ sur la bande $p+3$ en effaçant celui de la bande $p+4$ et compare le contenu des bandes $p+2$ et $p+1$. S'ils sont égaux elle efface la bande $p+2$ et s'arrête. Sinon elle revient en début de bande et travaille avec $\mathcal{M}'$ en se servant des bandes $1,\cdots,p$, $p+2$ et $p+3$ comme bandes de données et de la bande $p+4$ pour écrire le résultat. Quand cela est terminé elle rajoute un bâton sur la bande $p+2$, replace sa tête en début de bande et se met en état e_0. □

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

Démonstration. — C'est clair. □

On a donc démontré

3.5.7. Théorème. — *Les fonctions partielles récursives sont T-calculables.*

3.6. Les fonctions T-calculables sont récursives. — On fixe une machine de Turing $\mathcal{M}$ et on suppose que $\mathcal{M}$ calcule la fonction partielle f.

On suppose que $\mathcal{M}$ possède n bandes. On note $\{0,1,\cdots,m\}$ l'ensemble des états avec 0 initial et 1 final. On identifie le symbole blanc avec 0, d avec 1 et | avec 2. On appelle configuration de $\mathcal{M}$ à l'instant t la suite $C(t)=(s_0,s_1,\cdots)$ avec s_{nu+v} le symbole écrit sur la case $u+1$ de la bande $v+1$ $(0\leq v<n)$. La situation de de $\mathcal{M}$ à l'instant t est le triplet $S(t)=(e,k,C(t))$ avec e l'état et k le numéro des cases devant la tête à cet instant.

On code la configuration C par $\Gamma(C)=\sum_i s_i 3^i$ et la situation $S=(e,k,C)$ par $\Gamma(S)=\alpha_3(e,k,\Gamma(C))$. (Plus généralement, on code une suite finie $C=(s_i)$ de symboles par $\Gamma(C)=\sum_i s_i 3^i$.) A partir de $\Gamma(C)$ on retrouve le symbole écrit sur la case numéro u de la bande v par une fonction primitive récursive $\eta(\Gamma(C),u,v) = r(q(\Gamma(C),3^{n(u-1)+v-1}),3)$. Ici $q(x,y)$ et $r(x,y)$ désignent le quotient et le reste de la division par y. On retrouve de même la suite σ des n symboles sur les cases u: $\Gamma(\sigma)=\varepsilon(\Gamma(C),u,n)$, avec $\varepsilon(x,y,z)=r(q(x,3^{z(y-1)}),3^z)$.

3.6.1. Lemme. — *Il existe une fonction primitive récursive g telle que si x est le code de la situation de la machine à l'instant t, $g(x)$ est le code de la situation de la machine à l'instant $t+1$.*

Démonstration. — L'état de la machine est $\beta_3^1(x)$, le numéro des cases observées $\beta_3^2(x)$, et le code de la configuration est $\beta_3^3(x)$. Le code de la suite que la tête lit est $c = \varepsilon(\beta_3^3(x), \beta_3^2(x), n)$. Si $\beta_3^1(x) = j$ et si $\varepsilon(\beta_3^3(x), \beta_3^2(x), n) = c$, avec $c = \Gamma(s_0, \cdots, s_{n-1})$, $M(s_0, \cdots, s_{n-1}, j) = (t_0, \cdots, t_{n-1}, h, \omega)$ et

$$g(x) = \alpha_3(h, \beta_3^2(x) + \omega, \beta_3^3(x) + 3^{n(\beta_3^2(x)-1)}(c' - c))$$

si $c' = \Gamma(t_0, \cdots, t_{n-1})$. Si $\beta_3^1(x) > m$ ou si $\varepsilon(\beta_3^3(x), \beta_3^2(x), n) > 3^n - 1$ (dans ce cas x n'est pas code d'une situation), on pose $g(x) = 0$. □

3.6.2. Lemme. — *La fonction* $\mathrm{Sit}(t, x_1, \cdots, x_p)$ *donnant le code de la situation de la machine à l'instant t est une fonction primitive récursive.*

Démonstration. — Comme $\mathrm{Sit}(t+1, x_1, \cdots, x_p) = g(\mathrm{Sit}(t, x_1, \cdots, x_p))$, il suffit de montrer que la fonction $\mathrm{Sit}(0, x_1, \cdots, x_p)$ est primitive récursive. Or comme $\mathrm{Sit}(0, x_1, \cdots, x_p) = \alpha_3(0, 1, \Gamma(C))$ avec C la configuration des bandes où x_i est représenté sur la i-ème bande, les autres étant blanches, ceci résulte du fait que $\Gamma(C)$ est primitive récursive comme fonction des x_i. □

Montrons maintenant que la fonction f calculée par $\mathcal{M}$ est partielle récursive. Le temps de calcul est la fonction $T(x_1, \cdots, x_p) = \mu t(\beta_3^1(\mathrm{Sit}(t, x_1, \cdots, x_p)) = 1)$. La fonction f est alors égale au nombre de bâtons sur la $p+1$-ème ligne à l'instant T. Si x est le code de situation de la machine $\eta(\beta_3^3(x), y+2, p+1) = 0$ signifie que le symbole sur la case $y+2$ de la bande $p+1$ est un blanc. Considérons la fonction $\alpha(x) = \mu y(\eta(\beta_3^3(x), y+2, p+1) = 0)$. On a donc

$$f(x_1, \cdots, x_p) = \alpha(\mathrm{Sit}(T(x_1, \cdots, x_p), x_1, \cdots, x_p).$$

Remarquons que c'est seulement dans la définition de $T(x_1, \cdots, x_p)$ que l'on utilise le schéma μ non borné (pour $\alpha(x)$ on peut utiliser $\mu y \leq x$).

On en déduit de cette remarque :

3.6.3. Proposition. — *(1) Si f est une fonction totale calculable par une machine de Turing en un temps $T(x_1, \cdots, x_p)$ 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.

3.7. Machines et fonctions universelles. —

3.7.1. Codage des machines de Turing. — Une machine de Turing est définie par son nombre de bandes n, son ensemble d'états $E = \{0, \cdots, m\}$ et sa table de transition M qui est une application de $S^n \times E \to S^n \times E \times \{-1, 0, +1\}$. Pour $\rho = (s_1, \cdots, s_n, e)$ de $S^n \times E$ avec $M(\rho) = (t_1, \cdots, t_n, e', \varepsilon)$, on pose $r_1 = \alpha_2(\Gamma(s_1, \cdots, s_n), e)$, $r_2 = \alpha_3(\Gamma(t_1, \cdots, t_n), e', \varepsilon + 1)$ et $n(\rho) = \pi(r_1)^{r_2}$, avec $\pi(i)$ le i-ème nombre premier. Le code de la table M est l'entier

$$u = \prod_{\rho \in S^n \times E} n(\rho).$$

Pour décoder on utilise la fonction primitive récursive

$$\delta(i, x) = \mu z \leq x(x \quad \text{n'est pas divisible par} \quad \pi(i)^{z+1}).$$

Si c est le code de $(s_1, \cdots, s_n)$, c' celui de $(t_1, \cdots, t_n)$, on a $\delta(\alpha_2(c, e), u) = \alpha_3(c', e', \varepsilon + 1)$.

L'indice de la machine de Turing $\mathcal{M}$ est l'entier $\alpha_3(n, m, u)$.

Notons I_p l'ensemble des codes des machines de Turing ayant au moins $p + 1$ bandes. C'est un ensemble primitif récursif.

On définit alors une fonction $G^p(i, t, x_1, \cdots, x_p)$ par $G^p(i, t, x_1, \cdots, x_p) = 0$ si $i \notin I_p$ et $G^p(i, t, x_1, \cdots, x_p) = \Gamma(S(t))$, avec $\Gamma(S(t))$ le code à l'instant t de la situation de la machine d'indice i avec configuration initiale les x_i représentés sur les p premières bandes, les autres blanches, si $i \in I_p$.

3.7.2. Théorème. — *Pour tout entier p la fonction $G^p(i, t, x_1, \cdots, x_p)$ est primitive récursive.*

Démonstration. — La fonction $G^p(i, 0, x_1, \cdots, x_p)$ est certainement primitive récursive (exercice). Il suffit donc de montrer que la fonction $g(i, x)$ qui donne le code de situation de la machine d'indice i à l'instant $t+1$ en fonction de i et du code x de situation de la machine d'indice i à l'instant t est primitive récursive. C'est clair. Explicitement, $c = \varepsilon(\beta_3^3(x), \beta_3^2(x), \beta_3^1(i))$ est le code de ce que la tête lit à l'instant t. Si on pose $\delta = \delta(\alpha_2(c, \beta_3^1(x)), \beta_3^3(i))$, le code c' de la suite écrite à la place est égal à $\beta_3^1(\delta)$ et $g(i, x) = \alpha_3(e', k', \Gamma(C'))$ avec $e' = \beta_3^2(\delta)$, $k' = \beta_3^2(x) + \beta_3^3(\delta) - 1$ et $\Gamma(C') = \beta_3^3(x) + 3^{\beta_3^1(i)(\beta_3^2(x)-1)}(c' - c)$. □

3.7.3. Fonction universelle. — Pour chaque entier p, on note $T^p(i, x_1, \cdots, x_p)$ la fonction récursive partielle définie ainsi. Si $i \notin I_p$ elle n'est pas définie. Si $i \in I_p$, on pose

$$T^p(i, x_1, \cdots, x_p) = \mu t(\beta_3^1(G^p(i, t, x_1, \cdots, x_p)) = 1)).$$

C'est le temps de calcul de la machine de Turing d'indice i avec données initiales $(x_1, \cdots, x_p)$.

On note B^p l'ensemble des $(i, t, x_1, \cdots, x_p)$ tels que $\beta_3^1(G^p(i, t, x_1, \cdots, x_p)) = 1$ et C^p l'ensemble des $(i, y, t, x_1, \cdots, x_p)$ tels que $(i, t, x_1, \cdots, x_p)$ appartienne à B^p et y soit le nombre de bâtons sur la $p + 1$-ème bande de la machine d'indice i à l'instant t avec données initiales $x_1, \cdots, x_p$. Ce sont des ensembles primitifs récursifs.

On définit la fonction partielle récursive

$$\varphi^p(i, x_1, \cdots, x_p) = \mu y((i, y, T^p(i, x_1, \cdots, x_p), x_1, \cdots, x_p) \in C^p).$$

Quand elle est définie $\varphi^p(i, x_1, \cdots, x_p)$ donne le résultat du calcul par la machine de Turing d'indice i et de données initiales $x_1, \cdots, x_p$. Elle n'est définie que si $i \in I_p$ et si le calcul s'arrête.

Pour tout entier i, on note φ_i^p la fonction

$$\varphi_i^p : (x_1, \cdots, x_p) \mapsto \varphi^p(i, x_1, \cdots, x_p).$$

La fonction φ^p est une fonction partielle récursive universelle : toute fonction partielle récursive f en p variables $x_1, \cdots, x_p$ est de la forme φ_i^p pour i convenable. On dit que i est un indice de f. Il n'est en général pas unique.

3.8. Ensembles récursivement énumérables. —

3.8.1. Définition. — Une partie A de $\mathbf{N}^p$ est dite récursive si sa fonction caractéristique l'est.

L'ensemble des parties récursives est clos par intersection, réunion et passage au complémentaire.

3.8.2. Définition. — Une partie A de $\mathbf{N}^p$ est dite récursivement énumérable si elle est le domaine de définition d'une fonction partielle récursive.

Si l'on note W_i^p le domaine de définition de φ_i^p, les parties récursivement énumérables de $\mathbf{N}^p$ sont exactement les W_i^p.

3.8.3. Lemme. — *Tout ensemble récursif est récursivement énumérable.*

Démonstration. — Il existe une fonction partielle récursive f de domaine $\mathbf{N} \setminus \{0\}$, par exemple $\mu y(y + 1 = x)$. Si A est récursif, A est le domaine de $f \circ \mathbf{1}_A$. □

3.8.4. Théorème. — *Tout ensemble récursivement énumérable de $\mathbf{N}^p$ est la projection d'un ensemble primitif récursif de $\mathbf{N}^{p+1}$.*

Démonstration. — On écrit $A = W_i^p$. L'ensemble A est alors la projection de l'ensemble primitif récursif $B^p(i)$ formé de l'ensemble des points de B^p de première composante i. □

3.8.5. Théorème. — *La projection d'un ensemble récursivement énumérable est récursivement énumérable.*

Démonstration. — Montrons que si $A \subset \mathbf{N}^{p+1}$ est récursivement énumérable, sa projection B sur $\mathbf{N}^p$ l'est aussi. On écrit $A = W_i^{p+1}$. Alors B est la projection de $B^{p+1}(i)$, ie $(x_1, \cdots, x_p) \in B$ ssi il existe t et x_0 tels que $(i, t, x_0, x_1, \cdots, x_p) \in B^{p+1}$. L'ensemble B est domaine de définition de la fonction $\mu z((i, \beta_2^1(z), \beta_2^2(z), x_1, \cdots, x_p) \in B^{p+1})$. □

Notons $B^p(i)$ et $C^p(i)$ l'ensemble des points de B^p et C^p de première composante i. Ce sont des ensembles primitifs récursifs.

L'intersection de deux ensembles récursivement énumérables est récursivement énumérable car le domaine de $f_1 + f_2$ est l'intersection des domaines de f_1 et f_2. La réunion de deux ensembles récursivement énumérables est récursivement énumérable car si A_1 et A_2 sont respectivement projection de B_1 et B_2 primitifs récursifs, $A_1 \cup A_2$ est la projection de $B_1 \cup B_2$.

3.8.6. Théorème. — *Soit $A \subset \mathbf{N}^p$. L'ensemble A est récursif si et seulement si A et son complémentaire sont récursivement énumérables.*

Démonstration. — Démontrons l'implication non triviale. Soit i l'indice d'une fonction de domaine A et i' celui d'une fonction de domaine $\mathbf{N}^p \setminus A$. La fonction $h(x_1, \cdots, x_p)$ égale à $\mu t((t, x_1, \cdots, x_p) \in B^p(i) \cup B^p(i')$ est récursive totale. Soit g la fonction caractéristique de $B^p(i)$. La fonction $g(h(x_1, \cdots, x_p), x_1, \cdots, x_p)$ est égale à la fonction caractéristique de A. □

Soit f une fonction partielle récursive de domaine W_i^p. Le graphe de f est la projection de l'ensemble primitif récursif $C^p(i)$ il est donc récursivement énumérable. En particulier, l'image de f qui est projection de son graphe est récursivement énumérable.

Réciproquement, toute partie récursivement énumérable non vide A de $\mathbf{N}$ est l'image d'une fonction primitive récursive à une variable. En effet on écrit $A = W_i^1$ et on choisit $n \in A$. Alors A est l'image de la fonction g définie par $g(z) = \beta_2^2(z)$ si $(\beta_2^1(z), \beta_2^2(z)) \in B^1(i)$ et $g(z) = n$ sinon.

3.9. Le problème de l'arrêt. — Soit $g(x) = \varphi^1(x, x)$ et soit A le domaine de définition de g. Montrons que le complémentaire de A n'est pas récursivement énumérable. Sinon, il existerait un entier n tel que $\mathbf{N} \setminus A = W_n^1$. Notons que $x \in A$ ssi $A \in W_x^1$. On en déduit que $n \in \mathbf{N} \setminus A$ ssi $n \in A$, absurde.

3.9.1. Corollaire. — *Le domaine de définition de φ^1 n'est pas récursif.*

3.10. Compléments sur la récursivité. — Soit $\mathcal{M}$ une machine de Turing à au moins $p+p'$ bandes. Etant donnés des entiers $a_1, \cdots, a_{p'}$, on considère la machine de Turing $\mathcal{M}'$ qui fonctionne de la façon suivante. Elle écrit $a_1, \cdots, a_{p'}$ sur les bandes $p+2, \ldots, p+p'+1$. Ensuite elle travaille comme $\mathcal{M}$ à permutation des bandes près : étant donnée une donnée initiale $x_1, \cdots, x_p$ qu'elle écrit sur les p premières bandes elle effectue le même calcul que $\mathcal{M}$ (à permutation des bandes près) pour écrire le résultat sur la bande $p+1$. Enfin elle efface les bandes $p+2, \ldots, p+p'+1$. Il est clair qu'il existe une fonction primitive récursive g en $p'+1$ variables telle que l'indice de $\mathcal{M}'$ dans I_p soit égal à $g(i, a_1, \cdots, a_{p'})$ avec i l'indice de $\mathcal{M}$ dans $I_{p+p'}$.

On déduit de cette observation

***3.10.1. Théorème* (SMN).** — *Pour tout couple d'entiers m et n il existe une fonction primitive récursive s_n^m à $n+1$ variables telle que*

$$\varphi^{m+n}(i, x_1, \cdots, x_n, y_1, \cdots, y_m) = \varphi^m(s_n^m(i, x_1, \cdots, x_n), y_1, \cdots, y_m)$$

pour tous i, x_j et y_k.

***3.10.2. Théorème* (Rice).** — *Soit $\mathcal{X}$ un ensemble de fonctions partielles récursives à une variable non vide et distinct de l'ensemble de toutes les fonctions partielles récursives à une variable. Alors $A = \{i; \varphi_i^1 \in \mathcal{X}\}$ n'est pas récursif.*

Démonstration. — Quitte à remplacer $\mathcal{X}$ par son complémentaire on peut suppose que la fonction $\emptyset$ de domaine vide est dans $\mathcal{X}$. Soit b un entier qui n'est pas dans A. On considère la fonction

$$\psi(x, y, z) = \varphi^1(b, z) + \varphi^1(x, y) - \varphi^1(x, y)$$

et on note $\psi_{x,y}$ la fonction vue comme fonction de z seulement. Notons que $\psi_{x,y}$ appartient à $\mathcal{X}$ ssi $\varphi^1(x, y)$ n'est pas défini. On écrit $\psi(x, y, z) = \varphi^3(i, x, y, z) = \varphi^1(h(x, y), z)$ avec h primitive récursive. Notons que $h(x, y)$ est un indice pour $\psi_{x,y}$. Si W est l'ensemble des (x, y) tels que $\varphi^1(x, y)$ n'est pas défini, on a $(x, y) \in W$ ssi $h(x, y) \in A$. Si A était récursif, W le serait, absurde. □

Exemples d'applications :
- l'ensemble des indices d'une fonction partielle récursif n'est pas récursif
- l'ensemble des (i, j) avec $\varphi_i^1 = \varphi^j$ n'est pas récursif
- l'ensemble des indices des fonctions totales n'est pas récursif.

***3.10.3. Théorème* (Point fixe).** — *Soit p un entier positif et α une fonction récursive totale à une variable. Alors il existe un entier i tel que $\varphi_i^p = \varphi_{\alpha(i)}^p$.*

Démonstration. — On considère la fonction $\varphi^p(\alpha(s_1^p(y, y)), x_1, \cdots, x_p)$. Elle admet un indice a. On a donc

$$\varphi^p_{\alpha(s_1^p(y,y))}(x_1, \cdots, x_p) = \varphi_a^{p+1}(y, x_1, \cdots, x_p) = \varphi^p_{s_1^p(a,y)}(x_1, \cdots, x_p).$$

L'entier $i = s_1^p(a, a)$ convient. □

Montrons en application que la fonction d'Ackermann est récursive. On définit une fonction partielle récursive θ de la façon suivante :
- $\theta(i, y, x) = 2^x$ si $y = 0$.
-$\theta(i, y, x) = 1$ si $x = 0$.
- $\theta(i, y, x) = \varphi^2(i, y - 1, \varphi^2(i, y, x - 1))$ sinon.

On a

$$\theta(i, y, x) = \varphi^3(a, i, y, x) = \varphi^2(s_1^2(a, i), x, y)$$

pour a convenable. Il existe donc une fonction primitive récursive α_i telle que $\theta(i, y, x) = \varphi^2_{\alpha(i)}(y, x)$. Par le théorème du point fixe il existe un entier j tel que $\varphi_j^2 = \varphi^2_{\alpha(j)}$. La fonction φ_j^2 est récursive et satisfait les relations déterminant la fonction d'Ackermann. Celle-ci est donc égale à φ_j^2 et est donc récursive.

4. Modèles de l'Arithmétique et théorèmes de limitation

4.1. Axiomes de Peano. — On considère le langage $\mathcal{L}_0$ de l'arithmétique comprenant, outre l'égalité, 4 symboles : un symbole de constante 0, un symbole de fonction unaire S, deux symboles de fonctions binaires $+$ et $\times$.

L'ensemble des axiomes de Peano faibles est l'ensemble fini $\mathcal{P}_0$ formé des 7 axiomes suivants :

$(A1)$ $\forall v_0 \neg Sv_0 = 0$
$(A2)$ $\forall v_0 \exists v_1 (\neg v_0 = 0 \implies Sv_1 = v_0)$
$(A3)$ $\forall v_0 \forall v_1 (Sv_0 = Sv_1 \implies v_0 = v_1)$
$(A4)$ $\forall v_0 v_0 + 0 = v_0$
$(A5)$ $\forall v_0 \forall v_1 v_0 + Sv_1 = S(v_0 + v_1)$
$(A6)$ $\forall v_0 v_0 \times 0 = 0$
$(A7)$ $\forall v_0 \forall v_1 v_0 \times Sv_1 = (v_0 \times v_1) + v_0$

L'ensemble des axiomes de Peano est l'ensemble infini $\mathcal{P}$ formé de $\mathcal{P}_0$, et pour chaque formule $F[v_0, \cdots, v_n]$ de $\mathcal{L}_0$ sans autre variable libre que v_0, $\ldots$, v_n, de l'axiome d'induction

$$\forall v_1 \ldots \forall v_n ((F[0, v_1, \cdots, v_n] \wedge (F[v_0, \cdots, v_n] \implies F[Sv_0, \cdots, v_n])) \implies \forall v_0 F[v_0, \cdots, v_n]).$$

4.2. Modèles de $\mathcal{P}$ et de $\mathcal{P}_0$. — Notons tout d'abord que $\mathbf{N}$ est muni d'une $\mathcal{L}_0$-structure naturelle qui en fait un modèle de $\mathcal{P}$.

Dans une $\mathcal{L}_0$ structure, pour chaque entier n on notera $\underline{n}$ ou n le terme composé de n occurences du symbole S suivi du symbole 0 Ainsi $\underline{2}$ désigne $SS0$. Un tel élément de la structure est appelé standard.

Démontrons qu'il existe des modèles de $\mathcal{P}$ contenant des éléments non standard, en particulier non isomorphes à $\mathbf{N}$. En effet on considère le langage $\mathcal{L} = \mathcal{L}_0 \cup \{c\}$ et la théorie T formée de $\mathcal{P}$ et des énoncés $\neg c = \underline{n}$ pour tout entier n. Par le théorème de compacité T admet un modèle, dont le réduit à $\mathcal{L}_0$ est un modèle de $\mathcal{P}$ admettant un élément non standard.

On note $v_0 \leq v_1$ la formule $\exists v_2 (v_2 + v_0 = v_1)$.

La théorie $\mathcal{P}_0$ est très faible. On peut montrer qu'il existe des modèles de $\mathcal{P}_0$ pour lesquels l'addition n'est pas commutative.

Toutefois

4.2.1. Proposition. — *Soit $\mathfrak{M}$ un modèle de $\mathcal{P}_0$. Alors le sous-ensemble de $\mathfrak{M}$ formé des entiers standard est une sous-structure de $\mathfrak{M}$ qui en est un segment initial et est isomorphe à $\mathbf{N}$.*

On dit qu'une sous-structure $\mathfrak{N}$ de $\mathfrak{M}$ en est un segment initial si pour tout a dans $\mathfrak{N}$ et tout b dans $\mathfrak{M}$, si $\mathfrak{M} \models b \leq a$ alors $b \in \mathfrak{N}$ et si $b \notin \mathfrak{N}$, $\mathfrak{M} \models a \leq b$.

Démonstration. — Le seul point non immédiat est que le sous-ensemble de $\mathfrak{M}$ formé des entiers standard en est un segment initial. Notons que si $c \leq 0$ alors $c = 0$ et que $Sa \leq Sb$ si et seulement si $a \leq b$ (en effet $b = c + a$ ssi $Sb = S(c + a)$ ssi $Sb = c + Sa$). Montrons par récurrence sur $n \in \mathbf{N}$ que si $c \leq \underline{n}$, alors c est standard. Pour $n = 0$

c'est clair. Si $c \leq \underline{n+1}$ et que $c \neq 0$, $c = Sc'$ et $c' \leq \underline{n}$, c' est alors standard, ainsi que c. Montrons par récurrence sur $n \in \mathbf{N}$ que si c est non standard, alors $c \geq \underline{n}$. Pour $n = 0$, c'est clair. Si c est non standard, il s'écrit $c = Sc'$ avec c' non standard. Par hypothèse $c' \geq \underline{n}$, d'où $c \geq \underline{n+1}$. □

4.2.2. Proposition. — *Soit $\mathfrak{M}$ un modèle de $\mathcal{P}$. L'addition et la multiplication sont associatives et commutatives et la multiplication est distributive par rapport à l'addition. De plus tout élément (resp. tout élément non nul) de $\mathcal{M}$ est régulier pour l'addition (resp. la multiplication). Enfin la formule $v_0 \leq v_1$ définit un ordre total sur $\mathcal{M}$ compatible avec l'addition et la multiplication.*

Démonstration. — Exercice facile sur la récurrence. □

Le lemme suivant est très utile

4.2.3. Lemme. — *Soit $\mathfrak{M}$ un modèle non standard de $\mathcal{P}$. Soit $F(x)$ une formule avec une variable libre. Si pour tout n dans $\mathbf{N}$, $\mathfrak{M} \models F(n)$, alors il existe c non standard dans $\mathfrak{M}$ tel que $\mathfrak{M} \models F(c)$.*

Démonstration. — Sinon, seuls les éléments standard vérifieraient F dans $\mathfrak{M}$. Mais alors on aurait $F(0)$ et $\forall v F(v) \implies F(Sv)$. Absurde. □

4.3. Représentabilité des fonctions récursives. — On rappelle que $\mathcal{F}_p$ désigne l'ensemble des fonctions totales $\mathbf{N}^p \to \mathbf{N}$.

On dit qu'une fonction f dans $\mathcal{F}_p$ est représentée par une formule $F[v_1, \cdots, v_p, v_{p+1}]$ si pour tout p-uplet d'entiers $n_1, \ldots, n_p$ on a

$$\mathcal{P}_0 \models \forall v (F[\underline{n_1}, \cdots, \underline{n_p}, v] \iff v = \underline{f(n_1, \cdots, n_p)}).$$

Notons qu'en particulier, si m est un entier, $f(n_1, \cdots, n_p) = m$ si et seulement si $\mathcal{P}_0 \models F[\underline{n_1}, \cdots, \underline{n_p}, \underline{m}]$.

On dit qu'une partie A de $\mathbf{N}^p$ est représentable si sa fonction caractéristique $\mathbf{1}_A$ l'est. Dans ce cas, si F est une formule représentant $\mathbf{1}_A$, on a dit que la formule $G = F[v_1, \cdots, v_p, v_p, \underline{1}]$ représente A. Pour tout p-uplet d'entiers $n_1, \ldots, n_p$ on a $(n_1, \ldots, n_p) \in A$ si et seulement si $\mathcal{P}_0 \models G[\underline{n_1}, \cdots, \underline{n_p}]$.

On note Σ_1 le plus petit ensemble de formules de $\mathcal{L}_0$ contenant les formules sans quantificateurs, stable par $\wedge$ et $\vee$, clos par quantification existentielle, clos par quantification universelle bornée (i.e. si F est dans Σ_1, $\forall v_0 (v_0 < v_1 \implies F)$ aussi.

4.3.1. Théorème. — *Toute fonction récursive totale est représentable par une formule Σ_1.*

Démonstration. — Notons que les fonctions successeur, addition, multiplication, projection et constante sont représentables par des formules Σ_1.

Etape 1 : L'ensemble des fonctions représentables par une formule Σ_1 est stable par composition.

En effet si $f_1, \ldots, f_n$ dans $\mathcal{F}_p$ et g dans $\mathcal{F}_n$ sont représentées par F_i et G, la fonction $g(f_1, \cdots, f_n)$ est représentée par

$$\exists w_1 \cdots \exists w_n (G[w_1, \ldots, w_n, v] \wedge \bigwedge_{1 \leq i \leq n} F_i[v_1, \cdots, v_p, w_i]).$$

Etape 2 : L'ensemble des fonctions représentables par une formule Σ_1 est stable par l'opérateur μ total.

En effet si la fonction caractéristique de $A \subset \mathbf{N}^{p+1}$ est représentée par F et la fonction $f = \mu y((y, x_1, \ldots, x_p) \in A)$ est totale, alors f est représentée par

$$F(v_0, v_1, \ldots, v_p, 1) \wedge \forall w < v_0 F(w, v_1, \ldots, v_p, 0).$$

Etape 3 : Il existe une fonction β (la fonction de Gödel) à trois variables, récursive primitive et représentable par une formule Σ_1, telle que pour tout entier p et toute suite $(n_1, \cdots, n_p)$ dans $\mathbf{N}^p$, il existe des entiers a et b tels que pour tout i entre 1 et p on ait $\beta(i, a, b) = n_i$.

On définit $\beta(i, a, b)$ comme le reste de la division euclidienne de b par $ai + 1$. Notons que β est représentée par la formule

$$B[v_1, \cdots, v_3, v] = \exists w v_3 = (w \times S(v_2 \times v_1)) + v \wedge v < S(v_2 \times v_1).$$

Montrons que β possède la propriété voulue. Soit $\alpha_1, \cdots, \alpha_p$ une suite d'entiers. On choisit m supérieur à $p + 1$ tel que $a = m!$ soit supérieur à tous les α_i. Notons que les $ai + 1$ sont premiers entre eux deux à deux pour i entre 1 et p. En effet si c est un diviseur premier de $ai + 1$ et $aj + 1$ pour $j < i$, il divise $a(i - j) = m!(i - j)$ et donc $c \leq m$, ce qui contredit le fait que c divise $m!i + 1$.

Par le théorème chinois il existe un entier b tel que pour tout i entre 1 et p b soit congru à α_i modulo $ai + 1$. Comme $\alpha_i \leq a < ai + 1$, on a bien $\alpha_i = \beta(i, a, b)$.

Etape 4 : Soient g dans $\mathcal{F}_p$ et h dans $\mathcal{F}_{p+2}$ représentables par une formule Σ_1. Alors la fonction f définie par récurrence à partir de g et h: $f(x_1, \cdots, x_p, 0) = g(x_1, \cdots, x_p)$, $f(x_1, \cdots, x_p, x_{p+1} + 1) = h(x_1, \cdots, x_p, x_{p+1}, f(x_1, \cdots, x_p, x_{p+1}))$ est représentable par par une formule Σ_1.

Pour exprimer que $y = f(x_1, \cdots, x_{p+1})$ on écrit qu'il existe une suite d'entiers $(z(0), \cdots, z(x_{p+1}))$ telle que $z(0) = g(x_1, \cdots, x_p)$, $z(x_{p+1}) = y$ et $z(i + 1) = h(x_1, x_2, \cdots, x_p, i, z(i))$ au moyen de la fonction β. Soit G et H représentant g et h, B représentant β. On remplace B par $B'[v_1, v_2, v_3, v] = B[v_1, v_2, v_3, v] \wedge \forall v_4 < v \neg B[v_1, v_2, v_3, v_4]$ qui présente l'avantage que si $\mathfrak{M}$ est un modèle de $\mathcal{P}_0$ et n est standard avec $\mathfrak{M} \models B'[a, b, c, n]$, alors aucun autre point x de $\mathfrak{M}$ ne vérifie $B'[a, b, c, x]$. Notons que B' est Σ_1 car le quantificateur existentiel dans B peut être borné. Alors la formule

$$\exists w_1 \exists w_2 (\exists w (B'[1, w_1, w_2, w] \wedge G[v_1, \cdots, v_p, w]) \wedge B'[v_{p+1} + 1, w_1, w_2, v] \wedge$$
$$\forall w_3 < v_{p+1} \exists w_4 \exists w_5 (B'[S w_3, w_1, w_2, w_4] \wedge B'[SS w_3, w_1, w_2, w_5] \wedge H[v_1, \cdots, v_p, w_3, w_4, w_5])).$$

représente f. □

4.4. Codage des formules. — A toute formule F dans $\mathcal{L}_0$ on va associer un code $\#F$ de la façon suivante.

Par induction sur la hauteur du terme t on définit $\#t$ par:

si $t = 0$, $\#t = \alpha_3(0,0,0)$, si $t = v_n$, $\#t = \alpha_3(n+1,0,0)$, si $t = St_1$, $\#t = \alpha_3(\#t_1, 0, 1)$, si $t = t_1 + t_2$, $\#t = \alpha_3(\#t_1, \#t_2, 2)$, si $t = t_1 \times t_2$, $\#t = \alpha_3(\#t_1, \#t_2, 3)$.

4.4.1. Lemme. — *L'ensemble* Term *des $\#t$ pour t un terme de $\mathcal{L}_0$ est primitif récursif. Le codage $t \mapsto \#t$ est injectif.*

Si maintenant F est une formule on définit son code (ou nombre de Gödel) $\#F$ par induction sur la hauteur :

si $F = (t_1 = t_2)$, $\#F = \alpha_3(\#t_1, \#t_2, 0)$, si $F = \neg F_1$, $\#F = \alpha_3(\#F_1, 0, 1)$, si $F = F_1 \wedge F_2$, $\#F = \alpha_3(\#F_1, \#F_2, 2)$, si $F = F_1 \vee F_2$, $\#F = \alpha_3(\#F_1, \#F_2, 3)$, si $F = F_1 \implies F_2$, $\#F = \alpha_3(\#F_1, \#F_2, 4)$, si $F = F_1 \iff F_2$, $\#F = \alpha_3(\#F_1, \#F_2, 5)$, si $F = \forall v_n F_1$, $\#F = \alpha_3(\#F_1, n, 6)$, si $F = \exists v_n F_1$, $\#F = \alpha_3(\#F_1, n, 7)$, si $F = \top$, $\#F = \alpha_3(0,0,8)$.

4.4.2. Lemme. — *L'ensemble* Form *des $\#F$ pour T une formule de $\mathcal{L}_0$ est primitif récursif. Le codage $F \mapsto \#F$ est injectif.*

4.4.3. Lemme. — *Les ensembles suivants sont primitifs récursifs :*

(1) L'ensemble des $(\#t, n)$ avec t un terme et v_n n'a pas d'occurence (resp. a au moins une occurence) dans t.
(2) L'ensemble des $(\#F, n)$ avec F une formule et v_n n'a pas d'occurence (resp. n'a pas d'occurence libre, resp. liée) dans F.
(3) L'ensemble des $(\#F, n)$ avec F une formule et v_n a au moins une occurence libre (resp. liée) dans F.
(4) L'ensemble des $\#F$ avec F énoncé.

Notons qu'on aurait pu utiliser d'autres codages. Par exemple à toute formule F on pourrait associer le code ASCII de son expression en LaTeX.

Substituer dans une formule ou un terme est primitif récursif : il existe deux fonctions primitives récursives Subst_t et Subst_f telles que pour tout entier n, si t et u sont des termes et F une formule, $\mathrm{Subst}_t(n, \#t, \#u) = \#u_{t/v_n}$ et $\mathrm{Subst}_f(n, \#t, \#F) = \#F_{t/v_n}$.

Notons qu'on peut de même coder de façon injective les formules du calcul propositionnel:

si $P = \top$, $\#P = \alpha_3(0,0,0)$, si $P = P_n$, $\#P = \alpha_3(n+2,0,0)$, si $P = \neg P_1$, $\#P = \alpha_3(\#P_1, 0, 1)$, si $P = P_1 \wedge P_2$, $\#P = \alpha_3(\#P_1, \#P_2, 2)$, si $P = P_1 \vee F_2$, $\#P = \alpha_3(\#P_1, \#P_2, 3)$, si $P = P_1 \implies P_2$, $\#P = \alpha_3(\#P_1, \#P_2, 4)$, si $P = P_1 \iff P_2$, $\#P = \alpha_3(\#P_1, \#P_2, 5)$.

4.4.4. Lemme. — *L'ensemble des codes des tautologies du calcul propositionnel est primitif récursif.*

Démonstration. — Pour chaque entier k, on pose $\lambda_k(P_n) = 1$ si $\pi(n)$ divise k et $\lambda_k(P_n) = 0$ sinon, et on étend λ_k en une valeur de vérité sur toutes les formules du calcul propositionnel. On vérifie que la fonction E définie par $E(k, x) = 0$ si x n'est pas le code d'une proposition et par $E(k, \#P) = \lambda_k(P)$ est primitive récursive. Enfin P est une tautologie si et seulement si $\forall k \leq \pi(\#P)! E(k, \#P) = 1$. □

4.4.5. Lemme. — *L'ensemble des codes des tautologies du calcul des prédicats est primitif récursif.*

Démonstration. — Si F est une formule on écrit $F = P[F_1, \cdots, F_k]$ avec P une proposition et les F_i ne peuvent être décomposées à l'aide de connecteurs propositionnels. On pose $P_F = P[P_{\#F_i}, \cdots, P_{\#F_k}]$ et $P_\top = \top$. Alors F est une tautologie si et seulement si P_F en est une. Pour terminer il suffit de construire une fonction primitive récursive telle que $\gamma(\#F) = \#P_F$ pour tout formule F. □

On en déduit assez facilement

4.4.6. Théorème. — *L'ensemble* Ax *des* $\#F$ *avec* F *axiome logique de* $\mathcal{L}_0$ *est primitif récursif.*

On dit qu'une théorie T est récursive si $\#T = \{\#F, F \in T\}$ est récursif. Par exemple $\mathcal{P}_0$ et $\mathcal{P}$ sont récursives. On note $\#\mathrm{Th}(T)$ l'ensemble $\{\#F, T \vdash F\}$ pour F un énoncé. On dit que T est décidable si $\#\mathrm{Th}(T)$ est récursif.

Codage des démonstrations : Si $d = (F_0, \cdots, F_n)$ est une suite de formules on note $\#\#d = \pi(0)^{\#F_0} \cdot \pi(1)^{\#F_1} \ldots \pi(n)^{\#F_n}$.

4.4.7. Proposition. — *Si* T *est une théorie récursive alors l'ensemble* $\mathrm{Dem}(T) = \{(\#F, \#\#d)\}$*avec* F *une formule et* d *une démonstration de* F *à partir de* T *est récursif.*

4.4.8. Corollaire. — *Si* T *est récursive alors l'ensemble* $\#\mathrm{Th}(T)$ *est récursivement énumérable. En particulier l'ensemble des* $\#F$ *telles que* $\vdash F$*, resp.* $\mathcal{P}_0 \vdash F$*, resp.* $\mathcal{P} \vdash F$*, est récursivement énumérable.*

4.4.9. Corollaire. — *Si* T *est complète et récursive elle est décidable.*

4.5. L'argument diagonal. — Il résulte de ce qu'on a vu qu'il existe une fonction primitive récursive subst telle que pour toute formule F à une variable libre et tout entier n, $\mathrm{subst}[\#F, n] = \#F(n)$. On choisit une formule Σ_1 $G(x, y, z)$ représentant subst. Soit $F[x]$ une formule à une variable libre x. On considère la formule $H(x) = \exists z G(x, x, z) \wedge F(z)$. On pose $n = \#(\neg H)$ et $\Delta_F = \neg H(n)$.

4.5.1. Proposition. — *Pour toute* $F[x]$ *une formule à une variable libre* x*, on a*

$$\mathcal{P}_0 \vdash F(\#\Delta_F) \iff \neg\Delta_F.$$

Démonstration. — On a $\text{subst}[n,n] = \text{subst}[\#(\neg H), n] = \#\neg H(n) = \#\Delta_F$. On en déduit que dans tout modèle $\mathfrak{M}$ de $\mathcal{P}_0$ on a $G(n,n,\#\Delta_F)$ et $\forall z(G(n,n,z) \implies z = \#\Delta_F)$.

D'autre part

$$\Delta_F = \neg\exists z(G(n,n,z) \wedge F(z)).$$

Soit $\mathfrak{M}$ un modèle de $\mathcal{P}_0$. Si $\mathfrak{M} \models F(\#\Delta_F)$, $\mathfrak{M} \models \exists z(G(n,n,z) \wedge F(z))$ et donc $\mathfrak{M} \models \neg\Delta_F$. Réciproquement, si $\mathfrak{M} \models \neg\Delta_F$, $\mathfrak{M} \models \exists z(G(n,n,z) \wedge F(z))$. Il existe donc un élément c de $\mathcal{M}$ qui ne peut être que $\#\Delta_F$ tel que $G(n,n,c) \wedge F(c)$. On a alors $\mathfrak{M} \models F(\#\Delta_F)$. □

4.6. Le théorème de Tarski. —

***4.6.1. Théorème* (Tarski).** — *Soit $\mathfrak{M}$ un modèle de $\mathcal{P}_0$. Alors l'ensemble des nombres de Gödel des énoncés de $\mathcal{L}_0$ vrais dans $\mathfrak{M}$ n'est pas définissable dans $\mathfrak{M}$: il n'existe pas de formule $S_{\mathfrak{M}}(x)$ telle que pour tout énoncé F de $\mathcal{L}_0$ on ait $\mathfrak{M} \models F$ si et seulement si $\mathfrak{M} \models S_{\mathfrak{M}}(\#F)$.*

Démonstration. — Soit S une telle formule et posons $\Delta = \Delta_S$. Par construction, on a $\mathfrak{M} \models S(\#\Delta) \iff \neg\Delta$ ce qui contredit l'hypothèse faite sur S. □

4.6.2. Corollaire. — *Il n'existe pas de formule $S_{\mathbf{N}}(x)$ telle que pour tout énoncé F de $\mathcal{L}_0$ on ait $\mathbf{N} \models F$ si et seulement si $\mathbf{N} \models S_{\mathbf{N}}(\#F)$.*

4.6.3. Corollaire. — *L'ensemble $\#\text{Th}(\mathbf{N})$ des codes des énoncés vrais dans $\mathbf{N}$ n'est pas récursivement énumérable.*

Démonstration. — Comme F est vérifié dans $\mathbf{N}$ si et seulement si $\neg F$ ne l'est pas, si $\#\text{Th}(\mathbf{N})$ était récursivement énumérable son complémentaire dans $\mathbf{N}$ le serait aussi, donc $\#\text{Th}(\mathbf{N})$ serait récursif. Il serait alors représentable par une formule $G(x)$ qui vérifierait la propriété énoncée dans le corollaire. □

4.6.4. Corollaire. — *Soit T une théorie récursive telle que $\mathbf{N} \models T$. Alors il existe un énoncé de $\mathcal{L}_0$ vrai dans $\mathbf{N}$ mais non démontrable à partir de T.*

Démonstration. — En effet l'ensemble des codes des énoncés démontrables à partir de T est récursivement énumérable. □

4.7. Le théorème de Church. —

***4.7.1. Théorème* (Church).** — *Soit T un ensemble consistant d'énoncés de $\mathcal{L}_0$ contenant $\mathcal{P}_0$. Alors T n'est pas décidable.*

Démonstration. — Sinon $\#\text{Th}(T)$ serait récursif, donc représentable par une formule $F(x)$. Posons $\Delta = \Delta_F$. On sait que $\mathcal{P}_0 \vdash F(\#\Delta) \iff \neg\Delta$.

Si $T \vdash \Delta$, on a $\#\Delta \in \#\text{Th}(T)$ et donc $\mathcal{P}_0 \vdash F(\#\Delta)$, absurde. Si $T \nvdash \Delta$, on a $\#\Delta \notin \#\text{Th}(T)$ et donc $\mathcal{P}_0 \vdash \neg F(\#\Delta)$, absurde. □

4.7.2. Corollaire. — *Si T est une théorie récursive et consistante contenant $\mathcal{P}_0$, alors $\#\text{Th}(T)$ est récursivement énumérable non récursif.*

4.7.3. Corollaire. — *L'ensemble T_0 des énoncés de $\mathcal{L}_0$ universellement valides n'est pas récursif.*

Démonstration. — Soit H la conjonction des axiomes de $\mathcal{P}_0$. Pour tout énoncé F de $\mathcal{L}_0$, $\mathcal{P}_0 \vdash F$ si et seulement si $(H \implies F) \in T_0$. Si T_0 était récursif, $\mathcal{P}_0$ serait décidable, ce qui contredit le théorème de Church. □

4.8. Les théorèmes d'incomplétude de Gödel. — Soit T une théorie récursive. Alors $\mathrm{Dem}(T) = \{(\#F, \#\#d)\}$ est récursif. Soit $P_T(x,y)$ une formule Σ_1 représentant $\mathrm{Dem}(T)$ et notons $h_T(x)$ la formule $\exists y P_T(x,y)$. On note $\Delta_T = \Delta_{h_T}$.

4.8.1. Théorème **(Premier théorème d'incomplétude de Gödel)**
Soit T une théorie récursive consistante contenant $\mathcal{P}_0$. Alors $\mathbf{N} \models \Delta_T$, mais Δ_T n'est pas démontrable à partir de T.

Démonstration. — Soit $\Delta = \Delta_T$ et $n = \#\Delta$. On a $\mathcal{P}_0 \vdash \exists y P_T(n,y) \iff \neg\Delta$.

Supposons que $T \vdash \Delta$. Alors il existe une preuve de Δ dans T, autrement dit il existe un entier p dans $\mathbf{N}$ tel que (n,p) appartienne à $\mathrm{Dem}(T)$ et donc, par représentabilité, $\mathcal{P}_0 \vdash P_T(n,p)$. On a donc $\mathcal{P}_0 \vdash \exists y P_T(n,y)$ et donc $\mathcal{P}_0 \vdash \neg\Delta$, ce qui contredit l'hypothèse de consistance de T. Si $\mathbf{N} \not\models \Delta$, alors $\mathbf{N} \models \exists y P_T(n,y)$. Il existe alors un entier p tel que $P_T(n,p)$ soit vérifié dans $\mathbf{N}$. Autrement dit, il existe une preuve dans T de Δ et donc $T \vdash \Delta$, ce qui est absurde. □

Présentons maintenant l'amélioration technique, due à Rosser, de ce résultat. On définit une fonction primitive récursive $N : \mathbf{N} \to \mathbf{N}$ de la façon suivante : si n est le nombre de Gödel d'un énoncé F, $N(n)$ est le nombre de Gödel de $\neg F$ sinon $N(n) = 0$. On choisit une formule Σ_1 $\mathcal{N}$ représentant cette fonction. Si T est une théorie récursive, on modifie $P_T(x,y)$ en

$$P_T^R(x,y) = P_T(x,y) \wedge \neg\exists z \leq y \, \exists u (P_T(u,z) \wedge \mathcal{N}(x,u)).$$

On note $h_T^R(x)$ la formule $\exists y P_T^R(x,y)$ et $\Delta_T^R = \Delta_{h_T^R}$.

4.8.2. Théorème **(Variante de Rosser)**. — *Soit T une théorie récursive consistante contenant $\mathcal{P}_0$. Alors ni Δ_T^R ni sa négation ne sont démontrables à partir de T.*

Démonstration. — La preuve que $T \not\vdash \Delta_T^R$ est similaire à celle du théorème précédent. Supposons que $T \vdash \neg\Delta_T^R$. Il résulte qu'il existe un entier naturel e code d'une démonstration de $\neg\Delta_T^R$, et donc $\mathcal{P}_0 \vdash P_T(\#\neg\Delta_T^R, e)$. De plus, T étant consistante, pour tout entier naturel e' on a $\mathcal{P}_0 \not\vdash P_T(\#\Delta_T^R, e')$. Il résulte que

$$\mathcal{P}_0 \vdash \forall y (P_T(\#\Delta_T^R, y) \implies \exists z \leq y \, \exists u (P_T(u,z) \wedge \mathcal{N}(\#\Delta_T^R, u))).$$

Comme ce dernier énoncé est logiquement équivalent à $\neg h_T^R(\#\Delta_T^R)$, on en déduit que $\mathcal{P}_0 \vdash \Delta_T^R$, contradiction. □

Soit T une théorie récursive dans le langage $\mathcal{L}_0$. Soit a le code de l'énoncé $\neg 0 = 0$. On note $\mathrm{Coh}(T)$ l'énoncé

$$\neg \exists y P_T(S^a(0), y).$$

On a $\mathbf{N} \models \mathrm{Coh}(T)$ si et seulement si T est cohérente.

4.8.3. Théorème **(Second théorème d'incomplétude de Gödel)**
Soit T une théorie cohérente récursive et contenant $\mathcal{P}$ (ou au moins $\mathcal{P}_0$ et les axiomes d'induction pour les formules Σ_1). Alors $T \cup \{\mathrm{Coh}(T)\} \vdash \Delta_T$. En particulier $\mathrm{Coh}(T)$ n'est pas démontrable à partir de T.

Idée informelle de la preuve. — La preuve longue et technique est rarement présentée complètement en détail. Indiquons en l'idée. Si $T \cup \{\mathrm{Coh}(T)\} \nvdash \Delta_T$, alors il existe un modèle $\mathcal{N}$ de $T \cup \{\mathrm{Coh}(T)\}$ avec $\mathcal{N} \models \neg\Delta_T$, autrement dit $\mathcal{N} \models \exists y P_T(\#\Delta_T, y)$. Le fait que $\mathcal{N} \models \mathrm{Coh}(T)$ permet à $\mathcal{N}$ de "mimer" la construction d'un modèle de Henkin en faisant comme si les vrais entiers étaient les éléments de la structure $\mathcal{N}$. Ainsi un terme au sens de $\mathcal{N}$ sera un élément de N vérifiant une formule Σ_1 fixée représentant l'ensemble des codes de termes. La construction d'un modèle de Henkin peut ainsi être mimée (car $\mathcal{N} \models \mathrm{Coh}(T)$) par des constructions définissables par des formules à paramètres dans N et permet d'obtenir un modèle $\mathcal{M}$ de T au sens de $\mathcal{N}$, qui vérifie $\mathcal{M} \models T$. Comme $\mathcal{N} \models \exists y P_T(\#\Delta_T, y)$, il existe une preuve "au sens de $\mathcal{N}$ de Δ_T, ce qui permet d'obtenir que Δ_T est satisfaite dans $\mathcal{M}$. De même que tout modèle de $\mathcal{P}_0$ contient une copie de $\mathbf{N}$ comme segment initial, $\mathcal{M}$ contient comme segment initial une sous-structure isomorphe à $\mathcal{N}$. Mais, comme P_T est Σ_1, $\mathcal{N} \models \exists y P_T(\#\Delta_T, y)$ et $\mathcal{M} \models \neg\exists y P_T(\#\Delta_T, y)$ sont alors contradictoires. L'hypothèse que T doit contenir au moins les axiomes d'induction pour les formules Σ_1 permet de garantir que tout se passe comme indiqué. □

5. Théorie des ensembles

Dans cette section on considère le langage $\mathcal{L}$ formé outre de l'égalité d'un prédicat binaire noté $\in$. On considère une $\mathcal{L}$-structure $\mathcal{U}$. L'ensemble naïf sous-jacent U sera appelé univers et les éléments de U seront appelés ensembles. On ne parlera plus d'ensemble ou d'appartenance etc. dans le sens naïf (sauf exception).

5.1. Axiomes de Zermelo-Fraenkel. — Les axiomes de Zermelo-Fraenkel comprennent les axiomes d'extensionnalité, de la réunion, des parties, de remplacement et de l'infini.

5.1.1. Axiome d'extensionnalité. — Il exprime que deux ensembles ayant les mêmes éléments sont égaux :

$$\forall v_0 \forall v_1 (\forall v_2 (v_2 \in v_0 \iff v_2 \in v_1) \implies v_0 = v_1).$$

On notera $a \subset b$ si $\forall v_0 (v_0 \in a \implies v_0 \in b)$.

5.1.2. Axiome de la réunion. — Il exprime que étant donné un ensemble a, il existe un ensemble dont les éléments sont exactement les éléments des éléments de a :

$$\forall v_0 \exists v_1 \forall v_2 (v_2 \in v_1 \iff \exists v_3 (v_3 \in v_0 \wedge v_2 \in v_3)).$$

5.1.3. Axiome des parties. — Il exprime que étant donné un ensemble a, il existe un ensemble (unique par extensionnalité) que l'on notera $\mathfrak{P}(a)$ dont les éléments sont les parties de a :

$$\forall v_0 \exists v_1 \forall v_2 (v_2 \in v_1 \iff \forall v_3 (v_3 \in v_2 \implies v_3 \in v_0)).$$

5.1.4. Axiome de compréhension. — Pour tout formule $F[v_0, \ldots, v_n]$ de $\mathcal{L}$ on considère l'axiome

$$\forall v_1 \cdots \forall v_{n+1} \exists v_{n+2} \forall v_0 (v_0 \in v_{n+2} \iff (v_0 \in v_{n+1} \wedge F[v_0, \ldots, v_n])).$$

Une autre formulation est que pour toute formule $H[x]$ à paramètres dans U et tout ensemble a il existe un ensemble b tel que $y \in b$ si et seulement si $y \in a$ et $H[y]$ est vérifié. Une formule à paramètres dans $\mathcal{U}$ est une formule dans le langage $\mathcal{L}(\mathcal{U})$ (on rajoute à un $\mathcal{L}$ un symbole de constante pour chaque élément naïf de U). Notons qu'il est primordial d'imposer la restriction $y \in a$. En effet il n'existe pas d'ensemble a tel que $x \in a \iff x \notin x$.

Soit $H[x]$ une formule à paramètres dans U. On appelera classe définie par $H[x]$ l'ensemble naïf formés des ensembles satisfaisant H. En général une classe n'est pas un ensemble. Ainsi U est une classe (définie par $x = x$). Tout ensemble a est une classe (définie par $x = a$).

5.1.5. Quelques conséquences de l'axiome de compréhension. — Comme U est non vide, soit a un ensemble. Il existe un unique ensemble $\emptyset = \{x \in a, \neg x = x\}$. De plus cet ensemble ne dépend pas de a.

Si a et b sont deux ensembles il existe un unique ensemble $a \cap b$ dont les éléments sont les éléments de a appartenant à b. De même si a est non vide, il existe un unique ensemble $\cap_{x \in a} x$ formé des ensembles appartenant à tous les éléments de a.

5.1.6. Axiome de remplacement. — Il entraîne l'axiome de compréhension. Pour toute formule $F[w_0, w_1, v_1, \ldots, v_n]$ de $\mathcal{L}$ on demande

$$\forall v_0 \forall v_1 \cdots \forall v_n (\forall w_0 \forall w_1 \forall w_2 ((F[w_0, w_1, v_1, \ldots, v_n] \wedge F[w_0, w_2, v_1, \ldots, v_n]) \implies w_1 = w_2)$$
$$\implies \exists v_{n+1} \forall v_{n+2} (v_{n+2} \in v_{n+1} \iff \exists w_0 (w_0 \in v_0 \wedge F[w_0, v_{n+2}, v_1, \ldots, v_n]))).$$

Une relation fonctionnelle en w_0 est une formule $F[w_0, w_1, a_1, \ldots, a_n]$ à deux variables libres et à paramètres dans $\mathcal{U}$ telle que

$$\forall w_0 \forall w_1 \forall w_2 ((F[w_0, w_1, a_1, \ldots, a_n] \wedge F[w_0, w_2, a_1, \ldots, a_n] \implies w_1 = w_2).$$

A une telle relation correspond une fonction partielle naïve φ_F qui associe à b l'unique c tel que $F[b, c]$ s'il existe et qui n'est pas définie sinon. Le schéma de remplacement exprime que si $F[w_0, w_1]$ est une relation fonctionnelle en w_0 et si a est un ensemble, alors la classe formée des images par φ_F des éléments de a est un ensemble, que l'on notera $\{x; \exists v_0 \in a F[v_0, x]\}$. Le schéma de remplacement implique celui de

compréhension. En effet si $F[v_0]$ est une formule à une variable libre à paramètres dans $\mathcal{U}$, l'axiome de compréhension pour F est équivalent à celui de remplacement pour la relation fonctionnelle $w_0 = w_1 \wedge F[w_0]$.

5.1.7. Quelques conséquences de l'axiome de remplacement. — Montrons l'existence pour deux ensembles a et b de l'ensemble $\{a, b\}$ dont les éléments sont exactement a et b. On remarque que $\mathfrak{P}(\emptyset)$ possède un unique élément à savoir $\emptyset$. L'ensemble $\mathfrak{P}(\mathfrak{P}(\emptyset))$ possède donc exactement deux éléments à savoir $\emptyset$ et $\{\emptyset\}$. On applique l'axiome de remplacement à la relation fonctionnelle $(x = \emptyset \wedge y = a) \vee (x = \{\emptyset\} \wedge y = b)$.

On pose $\{a\} = \{a, a\}$ et $(a, b) = \{\{a\}, \{a, b\}\}$. Pour n un entier ≥ 3, on pose $(a_1, \ldots, a_n) = (a_1, (a_2, \ldots, a_n))$. On a $(a_1, \ldots, a_n) = (a'_1, \ldots, a'_n)$ si et seulement si $a_i = a'_i$ pour tout i.

Le produit de deux ensembles a et b est la classe définie par

$$\exists x \exists y (z = (x, y) \wedge x \in a \wedge y \in b)$$

C'est un ensemble car si $x \in a$ et $y \in b$, $(x, y) \in \mathfrak{P}(\mathfrak{P}(a \cup b))$.

Soit $R(x, y)$ une relation fonctionnelle à un argument. Son domaine est la classe définie par $\exists y R(x, y)$. Si ce domaine est un ensemble, alors l'image de R est un ensemble c et il existe un ensemble $f \subset a \times c$ dont les éléments sont les couples (x, y) tels que $R(x, y)$. Un tel ensemble f est appelé fonction de domaine a ou famille d'ensembles indexées par a. Si b est un ensemble, une fonction (ou application) de a dans b est une relation fonctionnelle de domaine a et d'image contenue dans b.

Exercice : écrire l'énoncé "f est une fonction de a dans b" dans le langage $\mathcal{L}$. Une fonction de a dans b étant un élément de $\mathfrak{P}(a \times b)$, il résulte de l'axiome de compréhension que la classe des fonctions des a dans b est un ensemble.

Soit a une famille d'ensemble indexée par un ensemble I, c'est à dire une fonction de domaine I. On note $\bigcup_{i \in I} a_i$ la réunion des éléments de l'image de la fonction a.

L'intersection $\bigcap_{i \in I} a_i$ de la famille est la classe définie par $\forall i (i \in I \implies x \in a_i)$ Si I est non vide c'est un ensemble.

Considérons la classe des applications f de I dans $\bigcup_{i \in I} a_i$ telles que $f(i) \in a_i$ pour tout $i \in I$. Une telle application est un élément de $(\bigcup_{i \in I} a_i)^I$, c'est donc un ensemble par compréhension. On le note $\prod_{i \in I} a_i$.

5.1.8. Ordinaux et axiome de l'infini. — On définit les ordinaux comme en théorie naïve. Il existe une formule à une variable libre $\mathrm{Ord}(x)$ exprimant que x est un ordinal. Les ordinaux forment donc une classe. Ils vérifient les mêmes propriétés que celles vues dans le premier chapitre. Notons que la classe des ordinaux n'est pas un ensemble. En effet soit X un tel ensemble. Ce serait un ordinal, on aurait donc $X \in X$ et $X \notin X$, absurde.

On dit qu'un ordinal est fini si ni lui-même, ni aucun de ses éléments n'est un ordinal limite. Sinon il est dit infini.

L'axiome de l'infini suppose l'existence d'un ordinal infini. Il équivaut à

$$\exists v_0 (\mathrm{Ord}(v_0) \wedge \neg v_0 = \emptyset \wedge \forall v_1 \neg v_0 = v_1 \cup \{v_1\}).$$

On note ω le plus petit ordinal infini (pour tout ordinal infini a, ω est le plus petit ordinal infini appartenant à $a+1$). Notons que ω est l'ensemble des ordinaux finis.

5.1.9. Proposition. — *L'ordinal ω muni de l'addition et de la multiplication ordinale, de l'application successeur et de $0=\emptyset$ est un modèle des axiomes de Peano, en général non standard.*

Démonstration. — Exercice. □

5.2. L'axiome du choix. —

5.2.1. Induction transfinie. — Soit F une formule à une variable libre et à paramètres dans $\mathcal{U}$. Pour montrer que $F[\alpha]$ est vérifiée pour tout ordinal α il suffit de montrer que si F est satisfaite pour tout $\gamma<\alpha$ alors elle est satisfaite pour α. En effet sinon, soit α le plus petit ordinal tel que $F(\alpha)$ ne soit pas satisfaite. Alors F est satisfaite pour tout $\gamma<\alpha$, absurde.

Le résultat suivant permet de construire des fonctions par induction transfinie.

5.2.2. Théorème. — *Soit $F[v_0,v_1,v_2]$ une formule avec paramètres dans $\mathcal{U}$ telle que*

$$\forall v_0 \forall v_1((\mathrm{Ord}(v_0)\wedge v_1 \textit{ est une application de domaine } v_0) \implies \exists! v_2 F[v_0,v_1,v_2]).$$

Alors, il existe une unique relation fonctionnelle φ de domaine la classe des ordinaux telle que pour tout ordinal α, $\varphi(\alpha)$ est l'unique ensemble x tel que $F(\alpha,\varphi_{|\alpha},x)$.

Démonstration. — On considère la condition $(*)$ suivante sur une fonction f : le domaine de f est un ordinal α et pour tout $\beta\in\alpha$ on a $F(\beta,f_{|\beta},f(\beta))$. Certainement si f vérifie $(*)$, la restriction de f à tout élément de α aussi. D'autre part pour tout ordinal α il existe au plus une telle fonction f (si f' est une autre telle fonction considérer le plus petit β tel que $f(\beta)\neq f'(\beta)$). On en déduit l'unicité.

On considère la formule $G[v_0,v_1]$ exprimant que v_0 est un ordinal et qu'il existe une application f de domaine v_0+1 satisfaisant $(*)$ et telle que $v_1=f(v_0)$. Cette formule est fonctionnelle en v_0. Il est reste à démontrer qu'elle est de domaine la classe des ordinaux.

Montrons par induction que pour tout ordinal α il existe un ensemble x satisfaisant $G[\alpha,x]$. On suppose que c'est vrai pour tout $\beta<\alpha$. Dans ce cas la restriction de G à α définit une fonction g de domaine α. Cette application satisfait $(*)$ (raisonner par l'absurde et considérer le plus petit β tel $F(\beta,g_{|\beta},g(\beta))$ ne soit pas vérifié). Par hypothèse il existe un ensemble x tel que $F[\alpha,g,x]$ soit vérifié. On étend g en une fonction f de domaine $\alpha+1$ en posant $f(\alpha)=x$) et on a bien $G[\alpha,x]$. □

5.2.3. L'axiome du choix. — On rappelle que l'axiome du choix AC est l'énoncé suivant : le produit d'une famille d'ensembles non vides est non vide. Pour tout ensemble a on appelle fonction de choix sur a une fonction h de l'ensemble $\mathfrak{P}(a)'$ des parties non vides de a dans a, telle que, pour tout $x\in\mathfrak{P}(a)'$, $h(x)\in x$

5.2.4. Proposition. — *L'axiome du choix est équivalent à l'existence d'une fonction de choix pour tout ensemble a.*

Démonstration. — Supposons que l'axiome du choix est vérifié. Soit a un ensemble. Alors le produit $\prod_{x\in\mathfrak{P}(a)'} x$ est non vide. Un élément de ce produit n'est autre qu'une fonction de choix sur a. Réciproquement si $(a_i)_{i\in I}$ est une famille d'ensemble non vides, considérons $a = \cup_{i\in I} a_i$ et une fonction de choix h sur a. En composant la fonction $I \to \mathfrak{P}(a)'$ correspondant à la famille des a_i avec la fonction de choix h on obtient une fonction $h : I \to a$ qui est un élément de $\prod_{i\in I} a_i$. □

5.2.5. Théorème. — *Les énoncés suivant sont équivalents dans* ZF *:*

(1) L'axiome du choix
(2) Le lemme de Zorn
(3) Le théorème de Zermelo.

Démonstration. — (1) $\implies$ (2). Supposons qu'il existe un ensemble $(X, <)$ qui est inductif et n'admet pas d'élément maximal. On considère l'ensemble T des parties totalement ordonnées par $<$. Pour chaque Y dans T l'ensemble $\{x \in X; \forall y \in Y\, y < x\}$ n'est pas vide. Par l'axiome du choix, il existe donc une fonction $k : T \to X$ telle que pour tout Y dans T, $\forall y \in Y$, $y < k(Y)$.

On considère la formule $F[v_0, v_1, v_2]$ exprimant que v_0 est un ordinal, que v_1 est une fonction de domaine v_0 et que si l'image de v_1 appartient à T, alors $v_2 = k(\mathrm{Im}(v_1))$ et sinon, $v_2 = \emptyset$. Cette formule vérifie les conditions du théorème 5.2.2. On en déduit une relation fonctionnelle h de domaine la classe des ordinaux telle que pour tout ordinal α, $F[\alpha, h_{|\alpha}, h(\alpha)]$ soit vérifiée. On vérifie par induction que pour tout ordinal α, $h(\alpha) \in X$ et que $h(\beta) < h(\alpha)$ si $\beta < \alpha$. Pour conclure on considère la formule $H[v_0, v_1]$ exprimant que v_0 est un ordinal, que $v_1 \in T$ et qu'il existe un isomorphisme de v_0 sur v_1. Comme un ensemble ordonné ne peut être isomorphe à deux ordinaux différents, la formule H est fonctionnelle en v_1. Par l'axiome de remplacement l'image de T par la fonction définie par H est un ensemble. Mais cet ensemble serait l'ensemble de tous les ordinaux !

(2) $\implies$ (3). Soit a un ensemble. On considère l'ensemble X des couples (b, R) avec $b \subset a$ et R un bon ordre sur b. On ordonne X par $(b, R) \leq (b', R')$ si b est un segment initial de (b', R') et R est la restriction de R'. Cette relation est inductive : en effet si $(b_i, R_i)_{i\in I}$ est une famille totalement ordonnée d'éléments de X, la réunion b des b_i munie de l'unique relation d'ordre induisant R_i sur chaque b_i est un majorant des b_i. Soit (b, R) maximal dans X. Si $b \neq a$, il existe un ensemble $b' = b \cup \{x\}$ avec $x \in a \setminus b$, et on peut prolonger R en un bon ordre sur b' en imposant que x est strictement supérieur à tout élément de a. Contradiction.

(3) $\implies$ (1). Soit a un ensemble. Par hypothèse il peut être bien ordonné. La fonction $\mathfrak{P}(a)' \to a$ qui à une partie non vide de a associe son plus petit élément pour ce bon ordre est une fonction de choix sur a. □

On note ZF les axiomes de Zermelo-Fraenkel et ZFC les axiomes de Zermelo-Fraenkel augmentés de l'axiome du choix.

Dans ZFC on peut développer la notion de cardinal et obtenir les mêmes énoncés que ceux obtenus dans le premier chapitre.

5.3. Paradoxe de Skolem. — Si la théorie ZFC est consistante, elle a un modèle, nécessairement infini à cause des axiomes, et, d'après le théorème de Löwenheim-Skolem, ce modèle possède une sous-structure dénombrable qui satisfait exactement les mêmes énoncés, en particulier les axiomes de ZFC. Les points de cette structure, qui sont des ensembles, sont, en tant qu'ensembles, inclus dans la classe de tous les ensembles, le support de la structure.

Or, l'un des premiers résultats de la théorie des ensembles est la démonstration de l'existence d'un ensemble non dénombrable. C'est une conséquence immédiate du théorème de Cantor. Comment un ensemble non dénombrable pourrait-il être inclus dans une sous-structure dénombrable ?

Comme souligné par Skolem, le problème réside dans la relativité de ce qu'on appelle ici dénombrable. En théorie des ensembles, un ensemble est dénombrable s'il est en bijection avec $\mathbf{N}$, l'ensemble des entiers naturels. Mais nous avons utilisé cette notion en deux sens différents : les ensemble dénombrables au sens du modèle de ZFC, et les ensemble dénombrables au sens de la théorie intuitive dans laquelle nous avons énoncé le théorème de Löwenheim-Skolem. On peut tout à fait formaliser ce théorème en théorie des ensembles, mais on ne peut faire coïncider le modèle de ZFC dans lequel on a effectué cette formalisation, et celui auquel on applique le théorème. Dans le modèle dénombrable de ZFC obtenu par Löwenheim-Skolem, il existe bien une collection (un ensemble de l'univers de la formalisation) de couples qui établit une bijection entre les ensembles $\mathbf{N}$ et $\mathbf{R}$ du modèle, mais comme $\mathbf{R}$, l'ensemble des réels, n'est pas dénombrable, cette collection n'est pas représentée par un ensemble de ce modèle. Ce n'est même pas une classe. Il n'y a aucun moyen d'en parler dans ce modèle. L'ensemble $\mathbf{R}$ est bien non dénombrable au sens du modèle.

5.4. Autres axiomes, incomplétude et énoncés de consistance relative. —

5.4.1. Axiome de fondation. — L'axiome de fondation AF est l'énoncé suivant : pour tout ensemble a non vide, il existe un élément de a dont l'intersection avec a est vide.

Sous l'axiome du choix il est équivalent à l'énoncé suivant : il n'existe pas de famille $(a_i)_{i\in\omega}$ telle que pour tout i dans ω, $a_{i+1} \in a_i$.

En effet supposons que l'axiome de fondation soit vérifié et montrons le second énoncé (sans axiome du choix). On considère l'ensemble a dont les éléments sont les a_i. Par hypothèse il existe a_n tel que $a \cap a_n$ soit vide, ce qui contredit $a_{n+1} \in a_n$. Réciproquement, soit x un ensemble non vide tel que pour tout élément y de x, $y \cap x$ soit non vide. Par l'axiome du choix il existe $f : x \to x$ telle que pour tout $y \in x$, $f(y) \in y$. Soit a_0 dans x. On définit une suite par récurrence sur i : $a_{i+1} = f(a_i)$.

En particulier sous l'axiome de fondation tout ensemble x vérifie $x \notin x$.

Dans l'univers $\mathcal{U}$, on définit par induction une relation fonctionnelle V_α de domaine la classe des ordinaux par $V_\alpha = \cup_{\beta<\alpha}\mathcal{P}(V_\beta)$. On a $V_0 = \emptyset$, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$. On considère la classe V formée de la réunion des V_α. Si x est dans V, on définit $r(x)$ comme le plus petit α tel que $x \in V_\alpha$.

5.4.2. Lemme. — *Un ensemble x est dans V si et seulement si tous ses éléments y sont. Si x est de rang r, les éléments de x sont de rang $< r$.*

Démonstration. — Soit x un ensemble dans V. Il est de rang $r = \alpha + 1$. On a donc $x \subset V_\alpha$ et donc tous les éléments de x sont de rang $< r$. Réciproquement, si tous les éléments de x sont dans V, l'ensemble des rangs de ces éléments est un ensemble d'ordinaux, il est donc contenu dans un ordinal α. On a alors $a \subset V_\alpha$, d'où le résultat. □

5.4.3. Lemme. — *Tout ordinal α est dans V et le rang de α est $\alpha + 1$.*

Démonstration. — Soit α le plus petit ordinal tel que $\alpha \notin V_{\alpha+1}$ s'il en existe. Alors si $\beta \in \alpha$, $\beta \in V_{\beta+1}$. Il suit que $\alpha \subset \cup_{\beta<\alpha} V_{\beta+1} = V_\alpha$, et $\alpha \in V_{\alpha+1}$. Soit α le plus petit ordinal tel que $\alpha \in V_\alpha$ s'il en existe. Il existe alors $\beta < \alpha$ tel que $\alpha \in \mathcal{P}(V_\beta)$. On a alors $\alpha \subset V_\beta$ et donc $\beta \in V_\beta$, absurde. □

5.4.4. Proposition. — *L'axiome de fondation équivaut à*

$$\forall x V(x).$$

Démonstration. — Supposons $\forall x V(x)$ satisfait et soit a un ensemble non vide. Soit $b \in a$ de rang minimal. Alors $a \cap b = \emptyset$.

Réciproquement, soit a un ensemble qui n'est pas dans V. Soit b un ensemble transitif contenant a et soit c l'ensemble des $x \in b$ tel que x ne soit pas dans V. C'est un ensemble non vide : considérons $x \in a$ qui n'est pas dans V, alors $x \in b$. Maintenant soit $x \in c$. Il existe $y \in x$ qui n'est pas dans V. Comme b est transitif, $y \in b$, donc $y \in c$. Il suit que $c \cap x \neq \emptyset$ et l'axiome de fondation n'est pas vérifié.

Il reste à vérifier que tout ensemble a est contenu dans un ensemble transitif b (c'est à dire tel que si $x \in b$ et $y \in x$ alors $y \in b$). Pour cela on considère la fonction f définie par induction sur ω par $f(0) = a$, $f(n+1) = \cup_{x \in f(n)} x$ et on pose $b = \cup_{n \in \omega} f(n)$. L'ensemble b est la clôture transitive de a : c'est un ensemble transitif contenant a et tout tel ensemble contient b. □

5.4.5. Formes faibles de l'axiome du choix. — L'axiome du choix dépendant, noté ACD, est l'énoncé suivant : pour tout ensemble a, pour tout élément x_0 de a et tout $r \subset a^2$ tel que $\forall x \in a \exists y \in a((x,y) \in r)$, il existe une suite $f : \omega \to a$ telle que $f(0) = x_0$ et $(f(n), f(n+1)) \in r$ pour tout $n \in \omega$. L'axiome du choix entraîne ACD dans ZF. En effet si on dispose d'une fonction de choix h sur a on peut définir par récurrence par $f(0) = x_0$ et $f(n+1) = h(\{y \in a; (f(n), y) \in r\})$.

L'axiome du choix dépendant entraîne l'axiome du choix dénombrable qui énonce que si $(X_n)_{n \in \omega}$ est une suite d'ensembles non vides alors le produit des X_n est non vide. En effet il suffit de poser $Y_n = \{n\} \times X_n$, $a = \cup_{n \in \omega} Y_n$ et de définir r par $(x,y) \in r$ si et seulement si il existe n tel que $x \in Y_n$ et $y \in Y_{n+1}$. Par l'axiome de choix dépendant on a alors une suite $f : \omega \to a$ avec $f(n) \in Y_n$ pour tout n, autrement dit $f(n) = (n, g(n))$ avec g une fonction de domaine ω définissant un élément du produit des X_n.

5.4.6. Axiome des cardinaux inaccessibles. — On se place dans ZFC. Un cardinal λ est appelé fortement limite si pour tout cardinal μ strictement inférieur à λ, 2^μ est aussi strictement inférieur à λ. On dit que λ est inaccessible s'il est fortement limite, strictement supérieur à $\aleph_0$ et régulier. L'axiome CI des cardinaux inaccessibles énonce l'existence d'un cardinal inaccessible.

5.4.7. Théorème de Gödel. — Le codage de Gödel est également possible dans le cadre du langage de la théorie des ensembles, car les nombre entiers naturels s'interpètent dans tout modèle de ZF. Ainsi il existe un énoncé $\mathrm{Coh}_{\mathrm{ZF}}$ exprimant la cohérence de ZF. De plus cet énoncé est arithmétique : les quantificateurs portent sur des éléments de ω et l'énoncé peut se réécrire dans le langage de l'arithmétique.

Le deuxième théorème d'incomplétude de Gödel s'étend à ce cadre:

***5.4.8. Théorème* (Gödel).** — *Si* ZF *est cohérente, alors* $\mathrm{ZF} \nvdash \mathrm{Coh}_{\mathrm{ZF}}$.

Par contre, on peut montrer assez facilement que

$$\mathrm{ZFC} + \mathrm{CI} \vdash \mathrm{Coh}_{\mathrm{ZF}}.$$

5.4.9. Quelques énoncés de consistance relative. — (1) Si ZF est consistant alors ZF + AF est consistant.

En effet on vérifie que si $\mathcal{U}$ est un modèle de ZF, alors la classe V est un modèle de ZF + AF.

(2) Si ZF est consistant alors $\mathrm{ZF}^- + \neg$Infini est consistant. Ici ZF^- désigne ZF sans l'axiome de l'infini.

En effet on vérifie que si $\mathcal{U}$ est un modèle de ZF, alors $(V_\omega, \in)$ est un modèle de $\mathrm{ZF}^- + \neg$Infini.

(3) Si ZFC est consistant alors ZFC + ¬CI est consistant.

Soit $\mathcal{U}$ un modèle de ZFC. S'il n'admet pas de cardinal inaccessible c'est terminé. Sinon, soit λ le plus petit cardinal inaccessible. Alors $(V_\lambda, \in)$ est un modèle de ZFC + ¬CI.

(4) Si ZF est consistant alors ZF + ¬AC est consistant (Mostowski 1939, méthode de Fraenkel-Mostowski).

(5) Si ZF est consistant alors ZFC + AF + HGC est consistant (Gödel 1940, ensembles constructibles).

Rappelons que l'hypothèse du continu est l'énoncé

$$2^{\aleph_0} = \aleph_1,$$

tandis que l'hypothèse du continu généralisée HGC est l'énoncé

$$2^{\aleph_\alpha} = \aleph_{\alpha+1}$$

pour tout ordinal α.

Soit $\mathcal{U}$ un modèle de ZF. Il existe une relation fonctionnelle de domaine $\mathcal{U}$ qui à un ensemble a associe l'ensemble $\Pi(a)$ des parties de a définissables par une formule $F(x)$ à paramètres dans a, ie de la forme $\{x \in a | F(x)\}$. [Noter la présence de deux

difficultés dans cette définition : comme on quantifie sur les formules, il faut définir une formule ZF définissant l'ensemble des formules dans le langage ZF, cela se fait par codage de Gödel ; d'autre part il faut définir la notion de satisfaction, pour cela on se restreint à ne considérer que la structure $(a \in)$.] On définit par induction une relation fonctionnelle L_α de domaine la classe des ordinaux par $L_\alpha = \cup_{\beta<\alpha}\Pi(L_\beta)$. L'axiome de constructibilité est l'énoncé $U = L$.

Il n'est pas très difficile de vérifier que L satisfait ZF + AF. Par contre il est moins facile (Gödel) de montrer que L satisfait l'axiome de constructibilité. Pour conclure Gödel démontre que l'axiome de constructibilité (joint à ZF) entraîne AC et HGC

Les énoncés précédents se démontrent en construisant un modèle à l'intérieur d'un modèle donné. La méthode de forcing, introduite par Cohen en 1964, permet de construire des modèles plus grands que des modèles donnés. Elle permet de démontrer les résultats suivants :

(6) Si ZF est consistant alors ZFC + AF + ¬HC est consistant (Cohen 1964).

(7) Si ZF est consistant alors ZF + AF + ¬AC est consistant (Cohen 1964).

(8) Si ZFC + CI est consistant alors ZF + ACD + "toute partie de $\mathbf{R}$ est mesurable" est consistant (Solovay 1970).

François Loeser, École Normale Supérieure, Département de mathématiques et applications, 45 rue d'Ulm, 75230 Paris Cedex 05, France (UMR 8553 du CNRS)
E-mail : Francois.Loeser@ens.fr