Introduction à la Logique

Zoé Chatzidakis, ENS, automne 2015

Introduction

Ce cours présentera quelques résultats de base en logique mathématique. La logique mathématique est vaste, elle comporte trois sujets principaux, qui ont des connexions fortes avec d'autres domaines mathématiques ou scientifiques.

– La théorie des ensembles (avec l'analyse fonctionnelle, la topologie)
– La théorie des modèles (avec l'algèbre)
– La récursivité, maintenant appelée calculabilité (avec l'informatique).

1 Rudiments de théorie des ensembles : ordinaux, cardinaux, etc.

1.1. Qu'est-ce qu'un ensemble ? La réponse à cette question est plus compliquée qu'on ne le croit. En particulier, existe-t-il un ensemble X ayant pour éléments tous les ensembles ? Un tel X satisferait $X \in X$; et donc, définissant

$$Y = \{x \in X \mid x \notin x\},$$

et posant la question "Y appartient-il à Y?" on obtiendrait une contradiction :

$$Y \in Y \iff Y \notin Y.$$

Il faut donc faire un peu attention. Nous parlerons de la *collection* de tous les ensembles, ou bien de la *classe* de tous les ensembles.

1.2. Notations, conventions.
$\in$: appartient à, ou: est membre de;
$\mathbb{N}$, les entiers naturels $\{0, 1, 2, \dots\}$;
$\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, les nombres entiers, rationnels, réels, complexes.
$A \cap B$, $A \cup B$, $A \setminus B$, $A \times B$ les opérations ensemblistes de base: intersection, union, complémentaire

relatif, produit cartésien.
Etant donnée une famille (A_i) d'ensembles indexée par un ensemble I, nous avons

$$\begin{aligned}\bigcup_{i\in I} A_i &= \{x \mid \exists i \in I,\, x \in A_i\},\\ \bigcap_{i\in I} A_i &= \{x \mid \forall i \in I,\, x \in A_i\}.\end{aligned}$$

Etant donné un ensemble A, nous avons son *ensemble de parties*,

$$\mathcal{P}(A) = \{B \mid B \subseteq A\}.$$

Si $f : A \to B$ est une fonction, alors $f(A)$ est l'image de f: $\{f(a) \mid a \in A\}$.

1.1 Deux théorèmes

Théorème 1.3. (Cantor) *Soit A un ensemble. Il n'existe pas de surjection de A sur $\mathcal{P}(A)$.*

Démonstration. Soit $f : A \to \mathcal{P}(A)$, et définissons

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

Nous allons montrer que $B \notin f(A)$. En effet, s'il existait $a \in A$ tel que $f(a) = B$, alors nous aurions

$$a \in B \iff a \notin B,$$

ce qui nous donne la contradiction désirée.

Théorème 1.4. (Cantor-Bernstein) *Soient A et B deux ensembles, et supposons qu'il existe des injections $f : A \to B$ et $g : B \to A$. Alors il existe une bijection $h : B \to A$.*

Démonstration. Les ensembles A et $f(A)$ sont en bijection, nous pouvons donc supposer que A est inclus dans B (et que f est l'inclusion). Si $n \in \mathbb{N}$, nous notons g^n l'identité si $n = 0$, et la composée $g \circ g \cdots \circ g$ n fois de la fonction g pour $n > 0$. Nous définissons

$$C = \{g^n(x) \mid x \in B \setminus A,\ n \in \mathbb{N}\},$$

nous avons donc $C = \bigcup_{n\in\mathbb{N}} g^n(B \setminus A)$, et comme l'image de f est contenue dans A, on obtient que $C = (B \setminus A) \cup g(C)$, l'union étant une union disjointe. De plus, on a

$$A \setminus C = B \setminus C,$$

puisque $B \setminus A$ est contenu dans C. Nous définissons maintenant $h : B \to A$ en posant

$$h(x) = \begin{cases} g(x) & \text{si } x \in C,\\ x & \text{sinon.}\end{cases}$$

Pour montrer que h est injective, il suffit de montrer que $h(C) \cap B \setminus C = \emptyset$, puisque h est injective sur C et sur $B \setminus C$. C'est clair, puisque $h(C) = g(C) \subseteq C$, et $h(B \setminus C) = B \setminus C$.
Pour la surjectivité, on note que $A = (A \cap C) \cup (A \setminus C) = g(C) \cup (B \setminus C) = h(B)$.

Remarque 1.5. On dit que deux ensembles A et B sont *équipotents* (noté $A \sim B$) s'il existe une bijection entre A et B. On dit que A est *subpotent* à B (noté $A \preceq B$) s'il existe une injection de A dans B.
Le théorème précédent nous dit donc que si chacun de A, B est subpotent à l'autre, alors ils sont équipotents.

Définition 1.6. Un ensemble infini est *dénombrable* s'il est équipotent à $\mathbb{N}$.

Remarque 1.7. Le théorème de Cantor (1.3) nous dit donc que les ensembles infinis ne sont pas tous équipotents : il n'existe pas de bijection entre $\mathbb{N}$ et $\mathcal{P}(\mathbb{N})$. Celui de Cantor-Bernstein (1.4) permet de montrer facilement que $\mathcal{P}(\mathbb{N})$ et $\mathbb{R}$ sont équipotents.

1.2 Notions d'ordre

Définition 1.8. Soit A un ensemble. Un *ordre* (partiel, strict) sur A est une relation binaire $<$ (i.e., donnée par un sous-ensemble de $A \times A$) et satisfaisant aux conditions suivantes :

(i) (transitivité) si $x, y, z \in A$ sont tels que $x < y$ et $y < z$ alors $x < z$.
(ii) (anti-réflexivité) si $x \in A$, alors $x \not< x$. (Ici $\not<$ veut dire : n'est pas $<$)
Si pour tout $x, y \in A$, on a $x < y$ ou $x = y$ ou $x > y$, on dira que l'ordre est *total*, ou *linéaire*.

On dénote par $x \leq y$: $x < y$ ou $x = y$. Alors $\leq$ est toujours transitif, mais il n'est que *faiblement antisymmétrique* : $x \leq y$ et $y \leq x$ impliquent $x = y$.

Exemples 1.9. Voici quelques exemples bien connus :
$(\mathbb{N}, <)$, $(\mathbb{Z}, <)$ et $(\mathbb{R}, <)$, où $<$ est l'ordre usuel. Ce sont des ordres totaux.
$(\mathcal{P}(\mathbb{N}), \subset)$, où l'ordre est donné par l'inclusion (stricte). C'est un ordre partiel.

Définition 1.10. Soient $(A, <)$ un ensemble ordonné, $a, a' \in A$, et $B \subseteq A$.

(i) a est un *plus petit élément* de B si $a \in B$, et pour tout $b \in B$, si $b \neq a$, alors $b > a$.
(ii) a est un *élément minimal* de B si $a \in B$ et pour tout $b \in B$, on a $b \not< a$.
(iii) a est un *minorant* de B si pour tout $b \in B$ on a $a \leq b$.
(iv) a est une borne inférieure de B si a est le plus grand élément de l'ensemble des minorants de B.
(v) Les notions duales de *plus grand élément*, d'*élément maximal*, de *majorant*, de *borne supérieure* sont claires je pense.
(vi) a et a' sont *incomparables* si $a \neq a'$, $a \not< a'$ et $a' \not< a$.

Exemples 1.11. $(\mathbb{N}, <)$ a un plus petit élément, et en fait, tout sous-ensemble de $\mathbb{N}$ a un plus petit élément.
Par contre $(\mathbb{Z}, <)$, $(\mathbb{Q}, <)$ et $(\mathbb{R}, <)$ n'ont pas de plus petit élément.
$(\mathcal{P}(\mathbb{N}), \subset)$ a un plus petit élément: $\emptyset$, l'*ensemble vide*[1]. Mais $(\mathcal{P}(\mathbb{N}) \setminus \{\emptyset\}, \subset)$ n'a pas de plus petit élément. Ses éléments minimaux sont les singletons $\{n\}$, $n \in \mathbb{N}$, et ils sont deux à deux incomparables.

Définition 1.12. Un ordre est *bien fondé* si toute partie non vide de A a un élément minimal. Un *bon ordre* est un ordre total qui est bien fondé.

Remarque 1.13. Il est facile de montrer qu'un ordre est bien fondé si et seulement s'il ne contient pas de suite décroissante stricte infinie.

Exemples 1.14. Je dénote par $\mathcal{P}^f(\mathbb{N})$ l'ensemble des parties finies de $\mathbb{N}$. Alors $(\mathbb{N}, <)$ et $(\mathcal{P}^f(\mathbb{N}), \subset)$ sont bien fondés. Mais $(\mathbb{Z}, <)$ et $(\mathcal{P}(\mathbb{N}), \subset)$ ne le sont pas.

Notation 1.15. La notation $\simeq$ entre deux ensembles ordonnés veut dire qu'il existe une bijection entre ces deux ensembles qui préserve l'ordre. On dit alors que les deux ensembles ordonnés sont *isomorphes*.

1.3 Opérations sur les ordres

Soient X et Y des ensembles (partiellement) ordonnés.

1.16. La somme ordonnée de X et Y, notée $X + Y$. L'ensemble sous-jacent de $X + Y$ est la somme disjointe de X et Y (notée $X \coprod Y$, j'écris aussi parfois $X \bigcup Y$), qu'on peut décrire ensemblistement de la façon suivante : on identifie X et $X \times \{0\}$, Y et $Y \times \{1\}$ de la façon naturelle, ce qui permet de mettre un ordre sur ces deux ensembles, et aussi les rend disjoints. On prend alors comme ensemble sous-jacent de $X + Y$ l'ensemble $(X \times \{0\}) \cup (Y \times \{1\})$, sur lequel on définit

$$(a, i) < (b, j) \iff \begin{cases} i < j, \text{ ou} \\ i = j \text{ et } a < b \end{cases}$$

Cela revient donc à mettre une copie de l'ordre Y "après" l'ordre X.

1.17. Le produit ordonné de X et Y, noté $X \times Y$, ou bien parfois $X \overset{\leftarrow}{\times} Y$. L'ensemble sous-jacent est le produit cartésien $X \times Y$, muni de l'ordre *anti-lexicographique*, c'est à dire,

$$(a_1, b_1) < (a_2, b_2) \iff \begin{cases} a_2 < b_2 \text{ ou} \\ a_2 = b_2 \text{ et } a_1 < b_1. \end{cases}$$

C'est donc la deuxième coordonnée qui domine. (L'ordre lexicographique est celui où c'est la première coordonnée qui domine).

[1] qui n'a aucun élément

Lemme 1.18. *Soient X, Y, Z des ensembles ordonnés (non vides).*

(1) *La somme ordonnée de deux ordres totaux [resp., bien fondés] est un ordre total [resp., bien fondé]*
(2) *Même chose pour le produit ordonné.*
(3) *(Associativité de la somme et du produit)* $(X + Y) + Z \simeq X + (Y + Z)$; $(X \times Y) \times Z \simeq X \times (Y \times Z)$.
(Distributivité) $X \times (Y + Z) \simeq (X \times Y) + (X \times Z)$.

Démonstration. (1) et (3) sont faciles, ainsi que le fait que le produit de deux ordres totaux est total, et sont laissés en exercice. Nous montrons maintenant que le produit de deux ordres bien fondés est bien fondé.

Supposons X et Y bien fondés, et soit $Z \subseteq X \times Y$ un ensemble non vide. On considère $\pi : X \times Y \to Y$ la projection sur la 2ème coordonnée. Soit y_0 un élément minimal de $\pi(Z)$ $(\subseteq Y)$, et considérons la fibre de Z au-dessus de y_0,

$$Z_{y_0} = \{x \in X \mid (x, y_0) \in Z\}.$$

Si x_0 est un élément minimal de Z_{y_0}, alors (x_0, y_0) est minimal dans Z.

Exercice 1.19. (1) Donnez les preuves manquantes du lemme.
(2) A-t-on $(X + Y) \times Z \simeq (X \times Z) + (Y \times Z)$?
(3) Et $(X + Y) \times Z \simeq X \times Z + Y \times Z$?

1.20. L'exponentielle (faible) de X par Y, notée $X^{(Y)}$. On suppose maintenant que X et Y sont totalement ordonnés, et que X a un plus petit élément 0. Soit X^Y l'ensemble des fonctions $f : Y \to X$. Si $f \in X^Y$, alors le *support de* f, $\mathrm{Supp}(f)$, est l'ensemble des $y \in Y$ tels que $f(y) \neq 0$.

L'ensemble sous-jacent de $X^{(Y)}$ est l'ensemble des fonctions f de Y dans X telles que $\mathrm{Supp}(f)$ soit fini. On définit un ordre sur $X^{(Y)}$ de la façon suivante : si $f, g \in X^{(Y)}$ sont distincts, la réunion de leurs supports est finie ; l'ensemble $\{y \in Y \mid f(y) \neq g(y)\}$ est donc fini, et a un plus grand élément, y_0. Alors on a $f < g$ ssi $f(y_0) < g(y_0)$.

Remarque 1.21. Remarquez que pour pouvoir définir l'ordre sur $X^{(Y)}$ on a absolument besoin du fait que Y soit totalement ordonné, puisqu'il faut que tout sous-ensemble fini de Y ait un plus grand élément. De même il faut que X ait un plus petit élément pour qu'on puisse définir le support. Mais on peut supposer que l'ordre sur X ne soit pas total : la définition ci-dessus donne alors un ordre partiel sur $X^{(Y)}$.

Définition 1.22. Soit X un ensemble totalement ordonné. Un sous-ensemble J de X est un *segment initial* si pour tout $a, b \in X$, $a < b$ et $b \in J$ impliquent $a \in J$.
Si $J \subseteq Y \subseteq X$, on parlera de *segment initial de* Y, en considérant Y avec l'ordre induit par celui de X.

Proposition 1.23. *Soient X et Y deux ordres totaux, X ayant un plus petit élément 0, et considérons $X^{(Y)}$ avec l'ordre défini ci-dessus.*

(1) *$X^{(Y)}$ est un ordre total.*
(2) *Si X et Y sont bien fondés, alors aussi $X^{(Y)}$.*
(3) *$X^{(Y+Z)} \simeq X^{(Y)} \times X^{(Z)}$; $X^{(Y\times Z)} \simeq (X^{(Y)})^{(Z)}$.*

Démonstration. (1) Cela suit facilement de la définition.

(2) Supposons X, Y bien fondés, et soit $Z \subseteq X^{(Y)}$ un ensemble non vide.

Si $\bar{0}$, la fonction constante sur Y égale à 0, est dans Z, alors $\bar{0}$ est minimal dans Z (car elle est minimale dans $X^{(Y)}$). Supposons donc que $\bar{0} \notin Z$, c'est à dire, $\mathrm{Supp}(f) \neq \emptyset$ pour tout $f \in Z$. On considère

$$Y_1 = \{\max \mathrm{Supp}(f) \mid f \in Z\}.$$

(Chaque support étant fini et non vide, a bien un plus grand élément, et nous considérons l'ensemble de ces plus grands éléments.) Soit y_1 le plus petit élément de Y_1 ($\subseteq Y$), et regardons

$$Z_1' = \{f \in Z \mid \max \mathrm{Supp}(f) = y_1\}.$$

Par définition de l'ordre, on a que si $f \in Z_1'$ et $g \in Z \setminus Z_1'$, alors $f < g$, et donc : Z_1' est un segment initial de Z. Soit x_1 le plus petit élément de $\{f(y_1) \mid f \in Z_1'\}$ ($\subseteq X$), et soit

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

Alors Z_1 est un segment initial de Z_1' et donc de Z. On identifie Z_1 avec un sous-ensemble de $X^{(Y\setminus\{y_1\})}$ (en oubliant la valeur de f en y_1), et on recommence. On regarde d'abord si la fonction f_1 qui vaut x_1 en y_1 et 0 ailleurs, est dans Z_1 : si oui, ce serait notre plus petit élément. Si non, alors on définit :

$$\begin{aligned}
Y_2 &= \{\max(\mathrm{Supp}(f) \setminus \{y_1\}) \mid f \in Z_1\},\\
y_2 &= \text{plus petit élément de } Y_2,\\
Z_2' &= \{f \in Z_1 \mid \max(\mathrm{Supp}(f) \setminus \{y_1\}) = y_2\},\\
x_2 &= \min\{f(y_2) \mid f \in Z_2'\},\\
Z_2 &= \{f \in Z_2' \mid f(y_2) = x_2\},\\
Y_3 &= \max(\mathrm{Supp}(f) \setminus \{y_1, y_2\}),\\
&\ldots
\end{aligned}$$

Comme ci-dessus on vérifie que $y_2 < y_1$, et que Z_2' et Z_2 sont des segments initiaux de Z_1. On répète la procédure, et obtient ainsi une suite $y_1, x_1, Z_1, y_2, x_2, Z_2, \ldots, y_n, x_n, Z_n, \ldots$. Mais comme les y_n forment une suite strictement décroissante et que Y est bien fondé, cette suite s'arrête forcément pour un certain n. Pourquoi s'arrête-t-elle ? Parce que l'ensemble Z_n contient une fonction (nécéssairement unique) de support $\{y_1, \ldots, y_n\}$: elle sera plus petite que tous les autres membres de Z_n.

(3) C'est facile.

1.4 Ordinaux

On considère la relation d'appartenance ($\in$) entre des ensembles. Nous utiliserons l'

Axiome d'extensionnalité : *Deux ensembles ayant les même éléments sont égaux.*

Définition 1.24. (1) Un ensemble X est *transitif* si pour tout $x \in X$ et $y \in x$ on a $y \in X$.
(2) Un ensemble X est un ordinal s'il est transitif, et $\in$ définit un bon ordre sur X.
(3) Je noterai On la classe de tous les ordinaux. Et j'utiliserai indifféremment $<$ ou $\in$ pour l'ordre sur un ordinal.

Proposition 1.25. *Soient α et β des ordinaux.*

(1) *$\emptyset$ est un ordinal.*
(2) *Si $\alpha \neq \emptyset$ alors $\emptyset \in \alpha$.*
(3) *$\alpha \notin \alpha$.*
(4) *Si $x \in \alpha$ alors $x = S_{<x} := \{y \in \alpha \mid y < x\}$ $(= \{y \in \alpha \mid y \in x\})$*
(5) *Si $x \in \alpha$, alors x est un ordinal (ce que je noterai aussi parfois $x \in$ On – ce n'est pas tout à fait correct, mais c'est plus court).*
(6) *$\beta \subseteq \alpha$ ssi $\beta = \alpha$ ou $\beta \in \alpha$.*
(7) *$x = \alpha \cup \{\alpha\}$ est un ordinal, noté (temporairement[2]) α^+.*

Démonstration. (1) Les conditions sont vides, et donc trivialement satisfaites.

(2) Un élément minimal de α (pour $\in$) doit être vide : si $x \in \alpha$ contient un élément y, alors $y \in \alpha$ par transitivité de α, et donc x n'est pas minimal.

(3) Si $x \in \alpha$ alors $x \notin x$ (puisque $\in$ est anti-réflexif sur les éléments de α). Alors $\alpha \in \alpha$ impliquerait $\alpha \notin \alpha$, une contradiction. On a donc bien $\alpha \notin \alpha$.

(4) Vient de la définition de $<$ sur α ($< = \in$).

(5) Si $x \in \alpha$, alors x est un ordinal : $x \subset \alpha$ implique que l'ordre est total et bien fondé ; si $z \in y$ et $y \in x$ alors $z < x$, d'où $z \in S_{<x} = x$, et x est donc transitif.

(6) On suppose β strictement contenu dans α, et on veut montrer que $\beta \in \alpha$. Soit x minimal dans $\alpha \setminus \beta$. Alors $\beta \supseteq S_{<x}$ (par définition de x).

Soit $y \in \beta$. Comme $\in$ est un ordre total sur α, on a $y = x$, ou $x \in y$, ou $y \in x$.
$y = x$: non, car $x \notin \beta$.
$x \in y$: non, car $y \in \beta$ impliquerait $x \in \beta$.
Donc, $y \in x$. Tous les éléments de β sont des éléments de x, et comme aussi $\beta \supseteq S_{<x}$, et par (4), on a $\beta = x$, et donc $\beta \in \alpha$.

(7) La vérification est facile.

[2] plus tard, nous le noterons $\alpha + 1$.

Exemple 1.26. Nous avons vu que $\emptyset$ est un ordinal, et qu'il appartient à tous les autres ordinaux. Mais quels sont les autres ? Tout d'abord, l'item (7) permet d'en construire une infinité à partir de $\emptyset$. On aura un ordre discret commençant avec $\emptyset$, puis $\{\emptyset\}$ (l'ensemble dont le seul élément est l'ensemble vide), puis $\{\emptyset, \{\emptyset\}\}$, puis $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, etc.
Pour simplifier les notations, nous les noterons 0, 1, 2, 3, etc.
L'ensemble des ordinaux $0, 1, 2, \ldots$ est noté ω. On vérifie qu'il est bien ordonné et transitif, c'est donc un ordinal. On peut donc naturellement l'identifier à $\mathbb{N}$. Mais il y a d'autres ordinaux : $\omega^+ = \omega \cup \{\omega\}$, etc.

Exercice 1.27. Soit α un ordinal. Montrez que si β est un ordinal vérifiant $\alpha \leq \beta \leq \alpha^+$, alors $\beta = \alpha$ ou $\beta = \alpha^+$. L'ordinal α^+ est donc un successeur de α (pour l'ordre $\in$).

Proposition 1.28. *Soit X un ensemble non-vide d'ordinaux. Alors $\bigcap_{\alpha \in X} \alpha$ est le plus petit élément de X.*

Démonstration. On vérifie facilement que $\bigcap_{\alpha \in X} \alpha$ est transitif et bien ordonné (une intersection d'ensembles transitifs est transitive ; un sous-ensemble d'un ensemble bien ordonné est bien ordonné). C'est donc un ordinal β, et on a $\beta \subseteq \alpha$ pour tout $\alpha \in X$.

Si $\beta \notin X$, alors $\beta \subset \alpha$ pour tout $\alpha \in X$ (par 1.25(6)), i.e.: $\beta \in \alpha$ pour tout $\alpha \in X$, et donc $\beta \in \beta$, ce qui est absurde (par 1.25(3)). On a donc $\beta \in X$, et il est clairement minimal dans X.

Théorème 1.29. *Soient α et β des ordinaux. Alors une (et une seule) des propriétés suivantes est satisfaite :*
- $\alpha = \beta$
- $\alpha \in \beta$
- $\beta \in \alpha$.

Démonstration. Appliquant la Proposition précédente 1.28 à $X = \{\alpha, \beta\}$, nous obtenons $\alpha \cap \beta \in \{\alpha, \beta\}$. Si $\alpha \cap \beta = \alpha$, alors $\alpha \subseteq \beta$, i.e., $\alpha = \beta$ ou bien $\alpha \subset \beta$ ($\leftrightarrow \alpha \in \beta$ par 1.25(6)), les deux cas étant bien sûr exclusifs. De même $\alpha \cap \beta = \beta$ implique $\beta = \alpha$ ou $\beta \in \alpha$.

Proposition 1.30. *Soit X un ensemble d'ordinaux. Alors $b = \bigcup_{\alpha \in X} \alpha$ est un ordinal. De plus, si $\gamma < b$, il existe $\alpha \in X$ tel que $\gamma \in \alpha$.*
On écrit aussi $b = \sup_{\alpha \in X} \alpha$.

Démonstration. Les éléments de b sont tous des ordinaux, et b est transitif (facile). Par le théorème précédent, $\in$ est un ordre total sur b. Soit $Z \subseteq b$ un ensemble non vide. Alors $\bigcap_{\alpha \in Z} \alpha$ est le plus petit élément de Z, donc l'ordre est bien fondé. Cela montre que b est un ordinal.

Soit $\beta < b$. Alors $\beta \in b$, et par définition de b, il existe $\alpha \in X$ tel que $\beta \in \alpha$, i.e., $\beta < \alpha$.

Remarque 1.31. Soit X l'ensemble des ordinaux $0, 1, 2, \ldots$. Alors $\bigcup_{\alpha \in X} \alpha = \omega \notin X$.

Définition 1.32. (1) Si α est un ordinal, alors α^+ ($= \alpha \cup \{\alpha\}$) est appelé le *successeur* de α.
(2) Un ordinal β est un *ordinal successeur* s'il existe un ordinal α tel que $\beta = \alpha \cup \{\alpha\}$. Notons que cet α peut aussi être décrit comme le plus grand élément de β.

(3) Un ordinal **non vide** qui n'est pas un ordinal successeur est appelé un *ordinal limite*.

Proposition 1.33. *Soit $\lambda \neq \emptyset$ un ordinal. Les conditions suivantes sont équivalentes :*

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

Démonstration. Si $\lambda = \beta \cup \{\beta\}$, alors $\bigcup_{\alpha<\lambda} \alpha = \beta$, car β est le plus grand élément de λ. Cela montre (2) $\to$ (1).

Pour l'autre direction. Puisque λ est limite, il n'a pas de plus grand élément : s'il en avait un, disons β, alors nous aurions $\lambda = \beta^+$. Donc $\beta \in \lambda$ implique $\beta^+ \in \lambda$.

Nous savons que $\beta = \bigcup_{\alpha<\lambda} \alpha$ est un ordinal, et $\beta \subseteq \lambda$, puisque tous ses éléments sont dans λ. On ne peut avoir $\beta < \lambda$, car on aurait $\beta^+ \in \lambda$, et donc $\beta \in \beta$, ce qui est ridicule. Cela montre que $\lambda = \beta$, et donc l'implication (1) $\to$ (2).

1.34. Induction transfinie. Soit P une propriété des ordinaux. On suppose :
– $\emptyset$ satisfait P ;
– Si un ordinal α satisfait P, alors α^+ satisfait P ;
– (λ ordinal limite) Si tous les $\alpha < \lambda$ satisfont P, alors λ satisfait P.

Alors tous les ordinaux satisfont P.

En effet, si tous les ordinaux ne satisfaisaient pas P, il en existerait un (disons α), et donc un plus petit (puisque α^+ est bien ordonné). Ce plus petit ordinal, disons β, contredirait nos hypothèses.

Remarquez qu'on peut fusionner les trois conditions, et dire tout simplement : Si α est un ordinal, et tous les $\beta \in \alpha$ satisfont P alors α satisfait P. Cependant, quand on veut vérifier l'hypothèse d'induction, en général, on le fait pour chacun des trois cas séparément.

Définition 1.35. Un ordinal α est *fini* si $\alpha = \emptyset$, ou bien si α et tous ses éléments non vides sont des successeurs.

Proposition 1.36. *ω est l'ensemble des ordinaux finis, et c'est le plus petit ordinal limite.*

Démonstration. Par définition de ω, tous ses éléments non vides sont des successeurs, mais lui-même n'est pas un successeur. Il n'est donc pas fini, et il est le plus petit ordinal limite.

Nous allons montrer :

Théorème 1.37. *Tout ensemble bien ordonné est isomorphe, comme ensemble ordonné, à un ordinal. Cet ordinal, ainsi que l'isomorphisme, sont uniques.*

Lemme 1.38. *Soit $f : \alpha \to \alpha'$ une fonction strictement croissante entre deux ordinaux α et α'. Alors $f(\beta) \geq \beta$ pour tout $\beta \in \alpha$.*

De plus, $\alpha' \geq \alpha$, et si f est un isomorphisme, alors $\alpha = \alpha'$ et f est l'identité.

Démonstration. S'il existe $\beta \in \alpha$ tel que $f(\beta) < \beta$, on prend un tel β minimal, β_0. Comme f est strictement croissante, nous obtenons (en appliquant f)

$$f(f(\beta_0)) < f(\beta_0)$$

ce qui contredit la minimalité de β_0. On a donc $f(\beta) \geq \beta$ pour tout $\beta \in \alpha$. Donc, si $\beta \in \alpha$, alors $\beta \in \alpha'$, ce qui entraine que $\alpha \subseteq \alpha'$, i.e., $\alpha \leq \alpha'$.

Si f est un isomorphisme, alors f^{-1} est strictement croissante, on applique la première partie à f^{-1} pour obtenir le résultat.

1.39. *Démonstration du théorème 1.37.* L'unicité suit du lemme précédent. Soit X un ensemble bien ordonné. Si $x \in X$, tout isomorphisme f_x entre le segment initial $S_{<x} = \{y \in X \mid y < x\}$ de X et un ordinal α s'étend alors en un isomorphisme $f : S_{\leq x} := S_{<x} \cup \{x\} \to \alpha^+$ obtenu en envoyant x sur α.
On pose

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

Comme remarqué ci-dessus, si $y \in Y$, alors l'ordinal $\alpha = \alpha(y)$ et l'isomorphisme f_y sont uniques. Supposons $Y \neq X$, et soit x minimal dans $X \setminus Y$. Si $y < x$, on a $y \in Y$, et donc $f_y : S_{\leq y} \simeq \alpha(y)$.

L'unicité des f_y implique qu'ils sont compatibles : si $y < z < x$, alors la restriction de f_z à $S_{\leq y}$ coïncide avec f_y. Soit $\alpha = \sup_{y \in Y} \alpha(y)$, et définissons $f : S_{<x} \to \alpha$ par $f(y) = f_y(y)$. (C'est OK car les f_y sont tous compatibles).

Alors f est un isomorphisme, qui s'étend à un isomorphisme entre $S_{\leq x}$ et α^+, ce qui nous donne une contradiction. Nous avons donc $Y = X$. Posant $\alpha = \sup_{y \in X} \alpha(y)$, et $f(y) = f_y(y)$ pour $y \in X$, nous obtenons l'isomorphisme entre X et α.

1.5 Opérations sur les ordinaux

Il y a deux façons de définir les opérations sur les ordinaux, et elles donnent le même résultat. L'équivalence des deux sera laissée en exercice (de TD?). La façon classique est la suivante, qui utilise une induction transfinie. Dans ce qui suit, les lettres grecques ($\alpha, \beta, \ldots$) dénotent des ordinaux.

1.40. Somme. $\alpha + \beta$ est définie par induction sur β :
- $\alpha + 0 = \alpha$;
- $\alpha + (\beta^+) = (\alpha + \beta)^+$;
- si β est limite, $\alpha + \beta = \sup_{\gamma<\beta} \alpha + \gamma$.

Remarques 1.41. (1) Sur les ordinaux finis, l'addition coïncide avec l'addition sur les entiers. Donc pas de problème. Mais sur les ordinaux infinis, pas du tout. En particulier

L'addition n'est pas commutative!!!
$0^+ + \omega = \sup_{n\in\omega} 1 + n = \omega$,
$\omega + 0^+ = \omega^+ > \omega$.
Si l'on note $0^+ = 1$, alors α^+ devient $\alpha + 1$.

1.42. Produit. De même, le produit $\alpha\beta$ est défini par induction sur β :
- $\alpha 0 = 0$;
- $\alpha(\beta^+) = \alpha\beta + \alpha$;
- si β est limite, $\alpha\beta = \sup_{\gamma<\beta} \alpha\gamma$.

Remarque 1.43. $2\omega = \sup_{n\in\omega} 2n = \omega < \omega 2 = \omega + \omega$.

1.44. Exponentielle. L'exponentielle α^β, pour $\alpha \neq 0$, est définie par induction sur β :
- $\alpha^0 = 1$ $(= 0^+)$;
- $\alpha^{\beta^+} = \alpha^\beta \alpha$;
- si β est limite, $\alpha^\beta = \sup_{\gamma<\beta} \alpha^\gamma$.

Exemple 1.45. $2^\omega = \sup_{n\in\omega} 2^n = \omega$.

Exercice 1.46. Soient κ et λ des ordinaux infinis, tels que λ contienne tous les ordinaux α^β, $\alpha + \beta$ et $\alpha\beta$ pour $\alpha, \beta \in \kappa$, $\alpha \neq 0$. Nous considérons κ et λ comme des ensembles topologiques, la topologie étant celle de l'ordre. Fixons $\alpha \in \kappa$, $\alpha \neq 0$. Montrez que la fonction $\kappa \to \lambda$, qui à $\beta \in \kappa$ associe $\alpha + \beta$, est continue.

Faites de même avec les fonctions $\beta \mapsto \alpha\beta$ et $\beta \mapsto \alpha^\beta$.

Rappel/définition. Soit $(I, <)$ un ensemble totalement ordonné. La topologie de l'ordre sur I est la topologie dont une base d'ouverts est donnée par les intervalles ouverts

$$(a, b)^3 =]a, b[:= \{x \in I \mid a < x < b\}$$

[3] J'utilise souvent la notation (a, b), qui est celle utilisée internationalement.

pour des éléments $a < b$ de I, et

$$(-\infty, a) =]-\infty, a[= \{x \in I \mid x < a\},\ (a, +\infty) =]a, +\infty[= \{x \in I \mid x > a\},$$

pour $a \in I$.

1.47. L'autre façon de définir les opérations ordinales. Les résultats de la section 1.3 montrent que si α et β sont des ordinaux, alors les ensembles ordonnés "somme", "produit" et "exponentielle" sont aussi bien ordonnés, et donc, par le Théorème 1.37, sont isomorphes à des ordinaux. Nous posons donc :

– $\alpha + \beta$ est l'unique ordinal isomorphe à la somme ordonnée de α et β ;
– $\alpha\beta$ est l'unique ordinal isomorphe au produit ordonné de α et β ;
– et si $\alpha \neq 0$, α^β est l'unique ordinal isomorphe à l'exponentielle ordonnée $\alpha^{(\beta)}$.

On pose de plus $0^0 = 1$, et $0^\beta = 0$ si $\beta > 0$.

Les preuves que les deux définitions coïncident ne sont pas difficiles, et sont faites par induction sur β. Par exemple, voici l'étape successeur de la preuve pour le produit :
On suppose qu'on a $\alpha\beta \simeq \alpha \overset{\leftarrow}{\times} \beta$. Alors

$$\alpha(\beta^+) = \alpha\beta + \beta \simeq \alpha \overset{\leftarrow}{\times} \beta + \alpha.$$

Si $+'$ dénote la somme ordonnée, on a ensuite

$$\alpha \overset{\leftarrow}{\times} \beta +' \alpha \simeq \alpha \overset{\leftarrow}{\times} \beta +' \alpha \times \{\beta\} \simeq \alpha \overset{\leftarrow}{\times} (\beta + 1),$$

ce qui donne le résultat pour $\beta^+ = \beta + 1$.

Exercice 1.48. Montrez, par induction sur β, que les deux définitions coïncident, pour la somme, le produit, et l'exponentielle.

L'équivalence des définitions rend alors plus facile la preuve de certaines propriétés :

Proposition 1.49. (1) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.
(2) $\alpha \leq \beta$ *ssi il existe* γ *tel que* $\alpha + \gamma = \beta$.
(3) *Si* $\beta < \beta'$, *alors pour tout* α, $\alpha + \beta < \alpha + \beta'$.
(Et donc: $\alpha + \beta = \alpha + \beta'$ *implique* $\beta = \beta'$ *– simplification à gauche de l'addition).*
(4) $1 + \alpha = \alpha$ *si* $\alpha \geq \omega$.
(5) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.
(6) $\alpha(\beta\gamma) = (\alpha\beta)\gamma$.
(7) *Si* $\alpha \neq 0$ *et* $\beta < \beta'$, *alors* $\alpha\beta < \alpha\beta'$.
(Et donc si $\alpha \neq 0$, *et* $\alpha\beta = \alpha\beta'$, *alors* $\beta = \beta'$ *– simplification à gauche de la multiplication par un élément non nul).*
(8) $\alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$.
(9) $(\alpha^\beta)^\gamma = \alpha^{\beta\gamma}$.

(10) *Si $\alpha > 1$ et $\beta < \beta'$, alors $\alpha^\beta < \alpha^{\beta'}$.*

Démonstration. Pour (2), on prend pour γ l'ordinal isomorphe à l'ensemble bien ordonné $\beta \setminus \alpha$. Les autres sont faciles, prouvés souvent par induction ou bien en utilisant (2).

Proposition 1.50. (Division euclidienne) *Soient α, β des ordinaux, avec $\alpha \neq 0$. Alors il existe une unique paire (ρ, μ) d'ordinaux telle que $\rho < \alpha$ et $\beta = \alpha\mu + \rho$.*

Démonstration. Unicité : on suppose $\beta = \alpha\mu + \rho = \alpha\mu' + \rho'$. Si $\mu < \mu'$ alors

$$\alpha\mu + \rho < \alpha\mu + \alpha = \alpha\mu^+ \leq \alpha\mu' \leq \alpha\mu' + \rho$$

ce qui est absurde. De même, on ne peut avoir $\mu' < \mu$, et donc on a $\mu = \mu'$. Mais alors $\rho = \rho'$ car on peut simplifier à gauche.

Existence : Si $\beta = 0$, OK.

Assertion. $\beta \leq \alpha\beta$.
En effet, l'application $f_0 : \beta \to \alpha \overset{\leftarrow}{\times} \beta$, $x \mapsto (0, x)$, est strictement croissante. On applique le lemme 1.38.

Si $\beta = \alpha\beta$, on prend $\mu = \beta$, $\rho = 0$. Sinon, on a $\beta < \alpha\beta$. Soit $f : \alpha\beta \to \alpha \overset{\leftarrow}{\times} \beta$ l'isomorphisme d'ensembles ordonnés. On pose $(\rho, \mu) = f(\beta)$, et on vérifie que çà marche : On montre par induction sur $\nu < \beta$ que $f(\alpha\nu) = (0, \nu)$; puis que $f(\alpha\nu + \sigma) = f(\alpha\nu) + (\sigma, 0)$ pour $\sigma < \alpha$.

Proposition 1.51. *Les lettres grecques dénotent des ordinaux. On suppose $\alpha > 1$.*

(1) *$\alpha^\gamma \geq \gamma$ pour tout γ.*
(2) *Si $\beta > 0$, alors il existe γ tel que $\alpha^\gamma \leq \beta < \alpha^{\gamma+1}$.*
(3) *Tout ordinal $\beta > 0$ admet un développement en base α, i.e. : il existe $n \in \mathbb{N}$, des ordinaux $\beta_1 > \beta_2 > \cdots > \beta_n \geq 0$ et des ordinaux k_i avec $0 < k_i < \alpha$ tels que*

$$\beta = \alpha^{\beta_1} k_1 + \alpha^{\beta_2} k_2 + \cdots + \alpha^{\beta_n} k_n.$$

Quand $\alpha = \omega$, cette écriture est appelée la forme normale de Cantor.

Démonstration. (1) Induction transfinie. Si $\gamma = 0$, OK : $\alpha^0 = 1 > 0$. Supposons le vrai pour γ. Alors

$$\alpha^{\gamma+1} = \alpha^\gamma \alpha \geq \alpha^\gamma + \alpha^\gamma.$$

L'inégalité vient du fait que $\alpha \geq 2$. Si $\gamma = 0$, le terme de droite égale $2 \geq \gamma + 1 = 1$. Si $\gamma \geq 1$, alors $\gamma + \gamma \geq \gamma + 1$, et cela donne le résultat.

Supposons maintenant que γ soit limite, et le résultat vrai pour tout $\beta < \gamma$. Alors

$$\alpha^\gamma = \sup_{\beta<\gamma} \alpha^\beta \geq \sup_{\beta<\gamma} \beta = \gamma.$$

(2) Par (1), on sait que $\beta < \alpha^{\beta^+}$, et donc il existe un plus petit δ satisfaisant $\alpha^\delta > \beta$. Si $\delta = \gamma^+$, alors γ est l'élément désiré : par définition de δ, on a $\alpha^\gamma \leq \beta$. Il suffit donc de montrer que δ ne peut être limite.

Soit δ un ordinal limite tel que $\alpha^\delta > \beta$. Comme $\alpha^\delta = \bigcup_{\gamma<\delta} \alpha^\gamma$, il existe $\gamma < \delta$ tel que $\beta \in \alpha^\gamma$ (cf Proposition 1.30). Mais $\beta \in \alpha^\gamma$ implique alors que δ ne peut être minimal tel que $\beta < \alpha^\delta$, et cela nous donne la contradiction désirée : δ ne peut pas être un ordinal limite.

(3) On prend γ comme dans (2) et on l'appelle β_1, puis on fait la division euclidienne de β par α^{β_1}, pour trouver k_1 et $\beta' < \alpha^{\beta_1}$ tels que $\beta = \alpha^{\beta_1} k_1 + \beta'$. Si $\beta' = 0$, on a fini. Sinon, on répète la procédure et on trouve d'abord β_2 tel que $\alpha^{\beta_2} \leq \beta' < \alpha^{\beta_2+1}$, puis k_2 et $\beta'' < \alpha^{\beta_2}$ tels que $\beta' = \alpha^{\beta_2} k_2 + \beta''$. Comme $\beta' < \alpha^{\beta_1}$, on a nécéssairement $\beta_2 < \beta_2 + 1 \leq \beta_1$. On construit ainsi une suite strictement décroissante d'ordinaux, qui sera donc finie. Si à l'étape n on ne peut continuer, c'est parce que

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

1.6 Axiome(s) du choix

Je vais vous donner 5 versions de l'axiome du choix (abbrégé par **AC**). Nous montrerons plus tard qu'elles sont équivalentes, modulo les axiomes de Zermelo-Fraenkel (**ZF**).

Définition 1.52. Soit $(X_i)_{i\in I}$ une famille d'ensembles indexée par l'ensemble I. On pose $\prod_{i\in I} X_i = \{f : I \to \bigcup_{i\in I} X_i \mid f(i) \in X_i,\ \forall i \in I\}$. Cet ensemble est appelé le *produit cartésien de la famille* (X_i). Quand nous parlerons de famille, nous supposerons toujours que les indices appartiennent à un ensemble. J'utiliserai parfois la notation indicielle – f_i au lieu de $f(i)$.

AC - version 1. *Le produit d'une famille d'ensembles non vides est non vide.*

Définition 1.53. Un ensemble ordonné non vide X est *inductif* si pour tout sous-ensemble totalement ordonné Y de X, Y admet une borne supérieure dans X.

AC - version 2 - Lemme de Zorn. *Tout ensemble ordonné non vide et inductif admet un élément maximal.*

AC - version 3 - Théorème de Zermelo. *Tout ensemble X admet un bon ordre.*

AC - version 4. *Si X est un ensemble dont les éléments sont non vides et disjoints, alors il existe un ensemble Y tel que, si $x \in X$, alors $x \cap Y$ a un seul élément.*

AC - version 5. Soit X un ensemble non vide. Alors il existe une fonction $h : \mathcal{P}(X) \to X$ telle que si $\emptyset \neq Y \subseteq X$, alors $h(Y) \in Y$.

Les versions 4 et 5 parlent bien de choix : dans AC4, l'ensemble Y choisit un élément de chaque membre de X, et dans AC5, la fonction h choisit un élément de chaque sous-ensemble non vide de X. Remarquons qu'il est facile, à partir d'une famille (X_i) d'ensembles, d'en construire une autre qui consiste d'ensembles deux à deux disjoints : on prend $Y_i = \{i\} \times X_i$. Cette remarque montre que AC4 implique AC1.

1.7 Cardinaux

Dans tout ce chapitre, on supposera que l'axiome du choix (AC) est vrai, et on utilisera chacune des 5 versions équivalentes.

Définition 1.54. Un cardinal est un ordinal qui n'est pas équipotent à un ordinal plus petit.

Exemples 1.55. (1) Les ordinaux finis sont des cardinaux.
(2) ω est le premier cardinal infini. On le note $\aleph_0$.
(3) Si α est un ordinal infini, alors α et $\alpha + 1$ sont équipotents. La démonstration est laissée en exercice : si $\alpha \geq \omega$, on commencera par trouver une bijection entre ω et $\omega \cup \{\alpha\}$.

Un cardinal infini sera donc nécéssairement un ordinal limite.

A partir de maintenant, je n'utilise plus la notation $^+$ pour le successeur ordinal, mais la notation +1

Proposition 1.56. *Tout ensemble X est équipotent à un unique cardinal, noté card(X).*

Démonstration. Par Zermelo, il existe un ordinal α qui est équipotent à X. On prend le plus petit ordinal β qui soit équipotent à α. Ce sera nécéssairement un cardinal.

Proposition 1.57. *Soient X et Y des ensembles, $X \neq \emptyset$. Les propriétés suivantes sont équivalentes :*

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

Démonstration. (1) $\leftrightarrow$ (2) : presque par définition.
(2) $\rightarrow$ (3) : soit $f : X \to Y$ injective, et choisissons $x_0 \in X$. On définit une fonction $g : Y \to X$ par

$$g(y) = \begin{cases} f^{-1}(y) \text{ si } y \in f(X), \\ x_0 \text{ sinon .} \end{cases}$$

(3) $\rightarrow$ (2). Soient λ et κ des cardinaux équipotents à X et à Y respectivement. Alors une fonction surjective de Y sur X nous donne une fonction surjective $g : \lambda \to \kappa$. Pour $\alpha \in \kappa$, on pose $f(\alpha) = \min\{\beta \in \lambda \mid g(\beta) = \alpha\}$. Alors f est injective, de κ dans λ.

Exemples 1.58. Bien sûr 0, 1, 2, sont des cardinaux. Ainsi que ω : si $\alpha < \omega$, alors α est fini, donc ne peut être équipotent à ω.
Par contre : $\omega + 1$ et $\omega + \omega$ ne sont pas des cardinaux : On construit sans peine des bijections entre ces ensembles et ω.

Proposition 1.59. *Soit X un ensemble de cardinaux. Alors $\lambda = \bigcup_{\alpha \in X} \alpha$ est un cardinal.*

Démonstration. Si $\beta < \lambda$, alors il existe $\alpha \in X$ tel que $\beta < \alpha$. Comme α est un cardinal, on a $\mathrm{card}(\beta) < \alpha \leq \lambda$.

Notation 1.60. On sait que $\mathrm{card}(\mathcal{P}(\kappa)) > \kappa$ pour tout cardinal κ (exercice, cf. 1.3). Il n'y a donc pas de "plus grand cardinal". On notera κ^+ le cardinal successeur de κ, i.e., le plus petit cardinal strictement supérieur à κ. (D'où la nécéssité d'utiliser $+1$ pour le successeur ordinal).

1.61. La hiérarchie des $\aleph$. On a déjà vu que ω, le premier cardinal infini, était aussi noté $\aleph_0$. On définit par induction une suite de cardinaux indexée par des ordinaux, en posant $\aleph_{\alpha+1} = \aleph_\alpha^+$, et pour un ordinal limite λ, $\aleph_\lambda = \sup_{\beta<\lambda} \aleph_\beta$.

Il existe d'autres hiérarchies. Par exemple celle définie par $\beth_{\alpha+1} = \mathrm{card}(\mathcal{P}(\beth_\alpha))$.

Proposition 1.62. *Tout cardinal infini est un $\aleph_\alpha$, où α est un ordinal.*

Démonstration. Soit κ un cardinal infini. Aors la fonction définie sur l'ordinal $\kappa + 2$ et qui envoie β sur $\aleph_\beta$ est strictement croissante. Donc $\aleph_\kappa \geq \kappa$ (par 1.38) et $\aleph_{\kappa+1} > \kappa$. Soit $\alpha \leq \kappa + 1$ minimal tel que $\aleph_\alpha > \kappa$. Alors $\alpha > 0$ car $\kappa \geq \aleph_0$.

Asssertion. α n'est pas limite.
On sait que $\kappa \in \aleph_\alpha$. Si α était limite, alors on aurait $\aleph_\alpha = \bigcup_{\beta<\alpha} \aleph_\beta$, d'où il existerait $\beta < \alpha$ tel que $\kappa \in \aleph_\beta$, et on aurait $\aleph_\beta > \kappa$, ce qui contredirait la minimalité de α.

Donc $\alpha = \beta + 1$, et $\aleph_\beta \leq \kappa < \aleph_\alpha$, i.e., $\kappa = \aleph_\beta$.

Remarque 1.63. Notez que la preuve ci-dessus montre en particulier le résultat suivant : si κ est un cardinal infini, alors c'est un ordinal limite.

1.64. Hypothèse(s) du continu. Il est difficile de ne pas mentionner cette question, car elle est quand même fondamentale : *Où* $\mathrm{card}(\mathcal{P}(\mathbb{N}))$ *se place-t-il dans la hiérarchie des* $\aleph$ *?* En fait, cette question n'a pas de réponse dans ZF. Nous en discuterons plus tard.

Hypothèse du continu (CH). $2^{\aleph_0} = \aleph_1$.

Hypothèse généralisée du continu (GCH). $2^\kappa = \kappa^+$ pour tout cardinal infini κ.

1.8 Opérations sur les cardinaux

Elles sont en fait plus simples que celles sur les ordinaux, et surtout beaucoup plus simples à décrire. **Attention, elles ne coïncident pas avec les opérations sur les ordinaux décrites précédemment. (Sauf sur les entiers, bien sûr).**

Notation 1.65. Soient A et B des ensembles. Alors $A \coprod B$ dénote l'union disjointe de A et B, $A \times B$ leur produit cartésien, et si A est non vide, A^B est l'ensemble des fonctions de B dans A.

Remarque 1.66. L'union disjointe de deux ensembles A et B n'est pas uniquement définie. C'est par définition, à bijection près, un ensemble C, avec deux injections $f : A \to C$ et $g : B \to C$ telles que $C = f(A) \cup g(B)$ et $f(A) \cap g(B) = \emptyset$. Par exemple, on peut identifier $A \coprod A$ avec $A \times \{0\} \cup A \times \{1\}$. Ou bien avec $A \times \{a\} \cup A \times \{b\}$ si $a \neq b$. Au niveau de l'équipotence, cela ne fait pas de différence.

Définition 1.67. Soient κ et λ deux cardinaux. On note :
$\kappa + \lambda = \mathrm{card}(\kappa \coprod \lambda)$ (addition de deux cardinaux),
$\kappa\lambda = \mathrm{card}(\kappa \times \lambda)$ (multiplication), et enfin
si $\kappa \neq 0$, $\kappa^\lambda = \mathrm{card}(\{f : \lambda \to \kappa\})$ (exponentielle).

Enfin, si κ_i est une famille de cardinaux indexée par l'ensemble I, on pose $\sum_{i\in I} \kappa_i = \mathrm{card}(\coprod_{i\in I} \kappa_i)$ et $\prod_{i\in I} \kappa_i = \mathrm{card}(\prod_{i\in I} \kappa_i)$. (Il y a une petite ambiguité ici, puisque $\prod_i \kappa_i$ dénote un cardinal, et aussi un produit cartésien d'ensembles). Notez ici que $\coprod_{i\in I} \kappa_i$ peut naturellement être identifié à $\bigcup_{i\in I} \{i\} \times \kappa_i$.

Remarque 1.68. Il suit immédiatement de la définition que l'addition et la multiplication des cardinaux sont des opérations commutatives. On montre aussi très facilement, pour des cardinaux κ, κ', λ, μ :
Distributivité : $\kappa(\lambda + \mu) = \kappa\lambda + \kappa\mu$;
$\kappa^{\lambda+\mu} = \kappa^\lambda \kappa^\mu$;
$\kappa^{\lambda\mu} = (\kappa^\lambda)^\mu$;
Si $\kappa \leq \kappa'$, alors $\kappa + \lambda \leq \kappa' + \lambda$, $\kappa\lambda \leq \kappa'\lambda$, $\kappa^\lambda \leq \kappa'^\lambda$ (si $\kappa > 0$), et $\lambda^\kappa \leq \lambda^{\kappa'}$.

Proposition 1.69. (Hessenberg) *Soit κ un cardinal infini. Alors $\kappa\kappa = \kappa$.*

Démonstration. Par induction sur l'ordinal α tel que $\aleph_\alpha = \kappa$. Si $\alpha = 0$, c'est bien connu : il existe une bijection entre $\mathbb{N}^2$ et $\mathbb{N}$, définie par $(m, n) \mapsto (m + n + 1)(m + n)/2 + n$.

On suppose le résultat vrai pour tous les $\beta < \alpha$. On met un ordre sur l'ensemble (de paires ordonnées d'ordinaux) $\aleph_\alpha \times \aleph_\alpha$, de la façon suivante :

$$(\beta_1, \gamma_1) < (\beta_2, \gamma_2) \iff \begin{cases} \max\{\beta_1, \gamma_1\} < \max\{\beta_2, \gamma_2\}, \text{ ou} \\ \max\{\beta_1, \gamma_1\} = \max\{\beta_2, \gamma_2\} \text{ et } \beta_1 < \beta_2, \text{ ou} \\ \max\{\beta_1, \gamma_1\} = \max\{\beta_2, \gamma_2\}, \ \beta_1 = \beta_2 \text{ et } \gamma_1 < \gamma_2. \end{cases}$$

C'est un ordre total, et on montre que c'est un bon ordre. En effet, soit $Z \subseteq \aleph_\alpha \times \aleph_\alpha$ non vide. On pose δ le plus petit élément de $\{\max\{\beta, \gamma\} \mid (\beta, \gamma) \in Z\}$, puis on prend $Z_1 = \{(\beta, \gamma) \in Z \mid \max\{\beta, \gamma\} = \delta\}$. Alors tout élément de $Z \setminus Z_1$ est plus grand que tout élément de Z_1. On s'aperçoit que sur Z_1, on a en fait l'ordre lexicographique, ce qui nous donne un plus petit élément.

On remarque ensuite que si $\delta < \aleph_\alpha$, alors $\delta \times \delta$ est un segment initial de $\aleph_\alpha \times \aleph_\alpha$. En effet, si $(\beta, \gamma) \in \aleph_\alpha \times \aleph_\alpha \setminus \delta \times \delta$, cela veut dire exactement que $\max\{\beta, \gamma\} \geq \delta$, et donc que (β, γ) est $>$ à tous les éléments de $\delta \times \delta$. De plus on a

$$\aleph_\alpha \times \aleph_\alpha = \bigcup_{\delta \in \aleph_\alpha} \delta \times \delta.$$

En effet, si $\beta, \gamma \in \aleph_\alpha$, alors prenant $\delta = \max\{\beta, \gamma\} + 1$, on a que $\delta < \aleph_\alpha$ ($\aleph_\alpha$ est infini, donc limite), et $(\beta, \gamma) \in \delta \times \delta$. Par hypothèse d'induction, $\text{card}(\delta \times \delta) = \text{card}(\delta)^2 = \text{card}(\delta)$ si $\delta \geq \aleph_0$, car $\text{card}(\delta) < \aleph_\alpha$.

Nous avons donc montré les choses suivantes : $\aleph_\alpha \times \aleph_\alpha$, avec l'ordre défini ci-dessus, est bien ordonné. Il est donc isomorphe à un ordinal, disons β (par le Théorème 1.37). D'autre part, il est clair que sa cardinalité est $\geq \aleph_\alpha$. Nous venons aussi de montrer que tout segment initial de β est de cardinalité $< \aleph_\alpha \leq \text{card}(\beta)$. Cela a pour conséquence d'abord que β est un cardinal, ensuite qu'il est $\leq \aleph_\alpha$, et donc est égal à $\aleph_\alpha$.

Proposition 1.70. *(Les lettres grecques dénotent des cardinaux)*

(1) *Soient $\kappa \geq \aleph_0$ et $\lambda > 0$. Alors*

$$\kappa + \lambda = \kappa\lambda = \max\{\kappa, \lambda\}.$$

(2) *Si $(X_i)_{i \in I}$ est une famille d'ensembles avec au moins un X_i infini, alors*

$$\text{card}(\bigcup_{i \in I} X_i) \leq \sup\{\text{card}(X_i) \mid i \in I\} + \text{card}(I).$$

Démonstration. (1) Sans perte de généralité, nous supposerons que $\kappa \geq \lambda$. Alors $\text{card}(\kappa \coprod \lambda) \leq \kappa + \kappa \leq \kappa \times \kappa = \kappa$.

(2) Soit $\lambda = \sup\{\text{card}(X_i) \mid i \in I\}$. Alors

$$\text{card}(\bigcup_{i \in I} X_i) \leq \text{card}(\coprod_{i \in I} X_i) \leq \text{card}(\lambda \times I) \leq \sup\{\lambda, \text{card}(I)\}.$$

Exercice 1.71. Donnez des exemples des phénomènes suivants :
(Toutes les lettres grecques dénotent des cardinaux)
$\kappa < \kappa'$, mais $\lambda + \kappa = \lambda + \kappa'$ et $\lambda\kappa = \lambda\kappa'$.

Exercice 1.72. Montrez que si κ est un cardinal infini alors $2^\kappa = \kappa^\kappa$.

Théorème 1.73. (Koenig) *Soient κ_i, λ_i, $i \in I$, deux familles de cardinaux telles que $\kappa_i < \lambda_i$ pour tout $i \in I$ (I un ensemble non vide). Alors*

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

Démonstration. Soit f une fonction de $\sum_{i \in I} \kappa_i \to \prod_{i \in I} \lambda_i$. Nous allons montrer que f n'est pas surjective, ce qui nous donnera le résultat. Pour chaque i, on définit $f_i : \kappa_i \to \lambda_i$ en composant la restriction de f à κ_i avec l'évaluation en i, autrement dit :

$$\text{si } x \in \kappa_i, \text{ on pose } f_i(x) = f(x)(i) \in \lambda_i.$$

(Rappel : le produit cartésien $\prod_i \lambda_i$ est un ensemble de fonctions de I à valeurs dans $\bigcup \lambda_i$).

Comme $\kappa_i < \lambda_i$ sont des cardinaux, la fonction f_i ne peut être surjective, et l'ensemble $B_i = \lambda_i \setminus f_i(\kappa_i)$ est donc non vide. Soit $b \in \prod_{i \in I} B_i$ (un tel b existe par AC), alors b n'est pas dans l'image de f. En effet, soit $c \in \sum_i \kappa_i$, et soit i tel que $c \in \kappa_i$. Alors $f(c)(i) = f_i(c) \neq b_i$. Donc, $f(c) \neq b$.

1.9 Cofinalités

Définition 1.74. (1) Soit X un ensemble totalement ordonné. Un sous-ensemble Y de X est *cofinal* dans X ssi pour tout $x \in X$ il existe $y \in Y$ tel que $y \geq x$.

(2) Soit X un ensemble totalement ordonné, et $f : Y \to X$. Alors f est *cofinale* (dans X) ssi son image est cofinale dans X.

(3) Soient α, β des ordinaux. Une fonction $f : \beta \to \alpha$ est *cofinale* ssi l'image de f, $f(\beta)$, est cofinale dans α.

(4) Soit α un ordinal. Alors, $\mathrm{cof}(\alpha)$, la *cofinalité de* α, est le plus petit ordinal β tel qu'il existe une fonction $f : \beta \to \alpha$ cofinale dans α.

Remarques 1.75. (1) Soit α un ordinal. Alors $\mathrm{cof}(\alpha + 1) = 1$, car $\alpha + 1$ a un plus grand élément, α.
(2) $\mathrm{cof}(\omega) = \omega$.
(3) Si α est un ordinal limite, alors $\mathrm{cof}(\aleph_\alpha) = \mathrm{cof}(\alpha)$.
(4) Si α est un ordinal limite, alors $X \subseteq \alpha$ est cofinal si et seulement si $\alpha = \bigcup_{\gamma \in X} \gamma$.
(5) Soient α et β des ordinaux, avec α limite, et $f : \beta \to \alpha$ une fonction. Alors f est cofinale dans α si et seulement si $\bigcup_{\gamma \in \beta} f(\gamma) = \alpha$.

Démonstration. (1) et (2) sont clairs. Pour (3), soit $f : \beta \to \alpha$ une application cofinale. Alors $\tilde{f} : \beta \to \aleph_\alpha$ définie par $\tilde{f}(\gamma) = \aleph_{f(\gamma)}$ est cofinale dans $\aleph_\alpha$. Cela montre que $\mathrm{cof}(\aleph_\alpha) \leq \mathrm{cof}(\alpha)$. D'autre part, étant donnée $g : \beta \to \aleph_\alpha$ cofinale, on définit $\tilde{g} : \beta \to \alpha$ par $\tilde{g}(\gamma) =$ le plus petit δ tel que $g(\gamma) < \aleph_\delta$.

(4) On sait que $\alpha = \bigcup_{\beta \in \alpha} \beta$ (par 1.33). Si $\alpha = \bigcup_{\beta \in X} \beta$, prenons $\gamma \in \alpha$; par définition de l'union, il existe $\beta \in X$ tel que $\gamma \in \beta$, et donc en particulier $\gamma < \beta$, ce qui montre bien que X

est cofinal. Réciproquement, si X est cofinal dans α, nous savons que $\bigcup_{\beta\in X}\beta$ est un ordinal (par 1.30), disons γ, et on a $\gamma \leq \alpha$ puisque tous ses éléments sont dans α. Il faut montrer que $\gamma = \alpha$: mais sinon, on aurait $\gamma < \alpha$, et donc il existerait $\beta \in \alpha$ tel que $\beta > \gamma$ (c'est ici qu'on utilise le fait que α soit limite), puis un $\delta \in X$ tel que $\delta \geq \beta$, ce qui contredit le fait que $\gamma = \sup_{\beta\in X}\beta$.

(5) Suit immédiatement de (4).

Proposition 1.76. *Soit α un ordinal.*

(1) $\mathrm{cof}(\alpha) \leq \alpha$.
(2) $\mathrm{cof}(\alpha)$ *est un cardinal.*
(3) $\mathrm{cof}(\alpha)$ *est le plus petit ordinal β tel qu'il existe une application cofinale strictement croissante $f : \beta \to \alpha$.*
(4) $\mathrm{cof}(\mathrm{cof}(\alpha)) = \mathrm{cof}(\alpha)$.

Démonstration. (2) Soit f une bijection entre $\mathrm{card}(\mathrm{cof}(\alpha))$ et $\mathrm{cof}(\alpha)$, et $g : \mathrm{cof}(\alpha) \to \alpha$ une application cofinale. Alors $g \circ f : \mathrm{card}(\mathrm{cof}(\alpha)) \to \alpha$ est cofinale, et $\mathrm{card}(\mathrm{cof}(\alpha)) \leq \mathrm{cof}(\alpha)$, d'øù le résultat, par minimalité de $\mathrm{cof}(\alpha)$.

(3) Soit $f : \mathrm{cof}(\alpha) \to \alpha$ une application cofinale. Nous allons montrer qu'on peut en trouver une qui soit strictement croissante. Définissons

$$X = \{x \in \mathrm{cof}(\alpha) \mid f(y) < f(x)\ \forall y < x\}.$$

Alors $f(X)$ est cofinale dans α. En effet, si $\gamma \in \alpha$, il existe $y \in \mathrm{cof}(\alpha)$ tel que $f(y) \geq \gamma$, et on prend un tel y minimal, il sera forcément dans X. De plus, la restriction de f à X est strictement croissante. L'ensemble X est bien ordonné, donc isomorphe à un ordinal β. Mais comme $X \subseteq \mathrm{cof}(\alpha)$ et $\mathrm{card}(X) \geq \mathrm{cof}(\alpha)$, on a nécéssairement un isomorphisme $g : X \to \mathrm{cof}(\alpha)$, et $h = f \circ g$ est notre application strictement croissante cofinale dans α.

Définition 1.77. Un cardinal infini κ est *régulier* si $\mathrm{cof}(\kappa) = \kappa$.

Exemples 1.78. $\aleph_0$ est régulier. Mais $\aleph_\omega$ ne l'est pas (puisqu'il est de cofinalité $\aleph_0$).

Proposition 1.79. *Tout cardinal infini successeur est régulier.*

Démonstration. Soit β un ordinal tel que $\kappa = \aleph_{\beta+1}$. Soient $\lambda < \kappa$ et $f : \lambda \to \kappa$, λ un cardinal. Alors $\lambda \leq \aleph_\beta$, et par 1.70(2),

$$\mathrm{card}(\bigcup_{\gamma\in\lambda} f(\gamma)) \leq \sup\{\mathrm{card}(f(\gamma)) \mid \gamma \in \lambda\} + \lambda.$$

Mais $\mathrm{card}(f(\gamma)) \leq \aleph_\beta$ (car $f(\gamma) \in \kappa$, et κ est un cardinal), et $\lambda < \kappa$. Ce qui montre que f n'est pas cofinale. (Cf 1.75(5))

Proposition 1.80. *Soient $\kappa \geq 2$ et $\lambda \geq \aleph_0$ des cardinaux. Alors* $\mathrm{cof}(\kappa^\lambda) > \lambda$.

Démonstration. Soit $f : \alpha \to \kappa^\lambda$, où α est un ordinal $\leq \lambda$. Alors $f(\beta) < \kappa^\lambda$ pour tout $\beta < \alpha$, et le théorème de König 1.73 donne

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

Mais $\prod_{\beta<\alpha} \kappa^\lambda = (\kappa^\lambda)^{\mathrm{card}(\alpha)} = \kappa^{\lambda \mathrm{card}(\alpha)} = \kappa^\lambda$. Nous avons donc montré que f n'est pas cofinale.

Corollaire 1.81. $2^{\aleph_0} \neq \aleph_\omega$.

Démonstration. $\mathrm{cof}(\aleph_\omega) = \omega = \mathrm{cof}(\omega) = \aleph_0 < \mathrm{cof}(2^{\aleph_0})$.

Exercice 1.82. Soit X un ensemble de cardinaux. Montrez que $\bigcup_{\alpha\in X} \alpha$ est un cardinal.

Remarque 1.83. Attention : en général

$$\mathrm{card}(\bigcup_{\alpha\in X} \alpha) \neq \bigcup_{\alpha\in X} \mathrm{card}(\alpha).$$

Par exemple, si κ est un cardinal infini, alors $\bigcup_{\alpha<\kappa} \alpha = \kappa$, mais si κ est successeur, disons $\kappa = \lambda^+$, alors $\bigcup_{\alpha<\kappa} \mathrm{card}(\alpha) = \lambda$.

2 Langages, formules, satisfaction, etc.

2.1 Langages et structures

Définition 2.1. Un langage est un ensemble de relations, fonctions et constantes. Les relations et fonctions viennent avec leur *arité*, c'est à dire un entier > 0.

2.2. On ajoute à ce langage des symboles logiques qui permettront de construire des formules : des symboles de variables ($x, y, \ldots, x_1, \ldots, u, v, \ldots$), des connecteurs ($\wedge$ la conjonction, et $\neg$ la négation), des parenthèses, le symbole $=$, et enfin des quantificateurs $\exists$, $\forall$ (il existe, pour tout). On utilisera aussi souvent les abbréviations $\vee$ ($A \vee B$ ssi $\neg(\neg A \wedge \neg B)$), $\rightarrow$ ($A \rightarrow B$ ssi $\neg A \vee B$) et $\leftrightarrow$ ($A \leftrightarrow B$ ssi $(A \rightarrow B) \wedge (B \rightarrow A)$, ssi $(A \wedge B) \vee (\neg A \wedge \neg B)$).

Exemple 2.3. 1 - Le langage des groupes : $\{\cdot, ^{-1}, e\}$, où $\cdot$ est un symbole de fonction binaire, $^{-1}$ un symbole de fonction unaire, et e un symbole de constante.

2 - Le langage des anneaux ordonnés : $\{+, -, \cdot, 0, 1, <\}$. Ici, $+, -, \cdot$ sont des symboles de fonctions binaires, 0 et 1 des symboles de constantes, et $<$ un symbole de relation binaire.

Définition 2.4. Soit $\mathcal{L}$ un langage. Une $\mathcal{L}$-structure est donnée par un univers M (en général supposé **non-vide**), et une interprétation de chaque symbole du langage dans M. On fait parfois la distinction entre la structure et l'univers, je ne la ferai pas, ou rarement. L'interprétation se fait de la façon suivante :
Pour chaque symbole de relation $R \in \mathcal{L}$, d'arité n, R^M un sous-ensemble de M^n ;
Pour chaque symbole de fonction $f \in \mathcal{L}$, d'arité n, f^M est une fonction $M^n \rightarrow M$;
Pour chaque symbole de constante $c \in \mathcal{L}$, c^M est un élément de M.

Exemple 2.5. Soit $\mathbb{R}$ l'ensemble des nombres réels. On peut le considérer comme une structure du langage des anneaux ordonnés, notée $(\mathbb{R}, +, -, \cdot, 0, 1, <)$, en donnant leur interprétation habituelle aux symboles $+, -, \cdot$ (addition, soustraction et multiplication), $0, 1$ (les éléments 0 et 1) et $<$ (l'ordre). Mais c'est un choix. Il existe de multiples façons (non naturelles bien sûr) de mettre une $\{+, -, \cdot, 0, 1, <\}$-structure sur $\mathbb{R}$.

Définition 2.6. Soient $\mathcal{L}$ un langage, M une $\mathcal{L}$-structure. Une *sous-structure* N de M, notée $N \subseteq M$, est une $\mathcal{L}$-structure dont l'univers N est un sous-ensemble de M, et dont la structure est "la restriction à N de celle sur M", plus précisément :
Si $R \in \mathcal{L}$ est une relation n-aire, alors $R^N = R^M \cap N^n$;
Si $f \in \mathcal{L}$ est une fonction n-aire, alors f^N est la restriction de f^M à N^n;
Si $c \in \mathcal{L}$ est une constante, alors $c^M = c^N$.

Remarque 2.7. Soit M une $\mathcal{L}$-structure, et N un sous-ensemble de M. Si N contient les interprétations des constantes dans M, et est clos par les (interprétations dans M des) fonctions du langage, alors N peut être muni (de façon unique) d'une $\mathcal{L}$-structure qui en fait une sous-structure de M. Cette $\mathcal{L}$-structure sur N est appelée la $\mathcal{L}$-structure *induite*.

Exemple 2.8. 1 - $\mathcal{L} = \{+, -, \cdot, 0, 1, <\}$, $\mathbb{R}$ muni de la $\mathcal{L}$-structure usuelle. Alors une sous-structure de $\mathbb{R}$ sera un sous-anneau de $\mathbb{R}$. Par exemple $\mathbb{Q}$ avec la $\mathcal{L}$-structure induite. Ou bien $\mathbb{Z}$. Par contre $\mathbb{N}$ ne sera jamais une sous-structure de $\mathbb{R}$.

2 - $\mathcal{L}_1 = \{\cdot, e\}$, $\mathcal{L}_2 = \mathcal{L}_1 \cup \{^{-1}\}$. (Langage des monoïdes, des groupes). Soit $\mathbb{R}$ avec sa structure naturelle de groupe additif. Donc $\cdot$ est interprété par l'addition, et $^{-1}$ par $-$, e par 0.
Une sous-$\mathcal{L}_1$-structure de $\mathbb{R}$ sera alors un sous-monoïde contenant 0, donc clos par addition. Par exemple $\mathbb{N}$. Par contre, une sous-$\mathcal{L}_2$-structure de $\mathbb{R}$ sera un sous-groupe additif.

Définition 2.9. Soient M et N deux $\mathcal{L}$-structures, $\mathcal{L}$ un langage.

(1) Un *morphisme* de M dans N (ou bien, un $\mathcal{L}$-morphisme, ou bien un *homomorphisme*) est une application $F : M \to N$, qui respecte les $\mathcal{L}$-structures de M et de N, c'est à dire :
Si $c \in \mathcal{L}$ est une constante, alors $F(c^M) = c^N$;
si $R \in \mathcal{L}$ est une relation n-aire, et $\bar{a} \in M^n$, alors $\bar{a} \in R^M$ implique $F(\bar{a}) \in R^N$;
et si $f \in \mathcal{L}$ est une fonction n-aire, $\bar{a} \in M^n$ alors $F^M(f(\bar{a})) = f^N(F(\bar{a}))$.
[Ici, si $\bar{a} = (a_1, \ldots, a_n)$, $F(\bar{a})$ dénote le n-uplet $(F(a_1), \ldots, F(a_n))$.]

(2) Un *plongement* de M dans N est un morphisme F qui est injectif, et de plus satisfait, pour toute relation n-aire $R \in \mathcal{L}$ et $\bar{a} \in M^n$,

$$\bar{a} \in R^M \iff F^M(\bar{a}) \in R^N.$$

(3) Un *isomorphisme* entre M et N est un plongement de M dans N qui est surjectif.

Remarques 2.10. 1 - Si F est un morphisme, il se peut que $F(\bar{a})$ satisfasse plus de relations dans M que $\bar{a}$ n'en satisfaisait dans N. Un morphisme injectif n'est donc pas nécéssairement un plongement.

2 - Si $F : M \to N$ est un plongement, alors $F(M)$ est une sous-structure de N, et F définit un isomorphisme entre M et $F(M)$.

3 - La notion de morphisme dépend beaucoup du langage. Par exemple, si on ajoute la relation (binaire) $\neq$ au langage, alors tout morphisme sera forcément injectif, puisqu'il doit respecter l'inégalité.
De même, quand on étudie des structures ordonnées, il est en général souhaitable d'utiliser la relation $\leq$ au lieu de la relation $<$: un $\{<\}$-morphisme de $(\mathbb{N}, <)$ vers $(I, <)$ où $<$ est anti-réflexif, sera automatiquement *injectif*. Ce ne sera pa le cas si on prend comme relation de base $\leq$, où $x \leq y$ est défini par $(x = y) \vee (x < y)$.

4 - Soient $\mathcal{L}$ un langage, et M une $\mathcal{L}$-structure. On forme le langage $\mathcal{L}(M)$ en ajoutant à $\mathcal{L}$ un nouveau symbole de constante $\underline{a}$ pour tout élément a de M. (On obtient alors un langage de cardinalité $\mathrm{card}(\mathcal{L}) + \mathrm{card}(M)$.) La $\mathcal{L}$-structure M a une $\mathcal{L}(M)$-structure naturelle, obtenue en interprétant le symbole $\underline{a}$ par $\ldots$ a. *Ce langage sera constamment utilisé par la suite, il est très utile.*

Exemple 2.11. Soit $\mathcal{L} = \{\oplus, *, \underline{1}\}$, où $\oplus$ est une fonction binaire, $*$ est une fonction unaire, et $\underline{1}$ est une constante. Soit $\mathcal{R}_1$ la $\mathcal{L}$-structure d'univers $\mathbb{R}$ où $\oplus$ est interprété par l'addition, $*$ par

la fonction $x \mapsto -x$, et $\underline{1}$ par 0. Soit $\mathcal{R}_2$ la $\mathcal{L}$-structure d'univers $\mathbb{R} \setminus \{0\}$, où $\oplus$ est interprété par la mutiplication, $*$ par la fonction $x \mapsto x^{-1}$, et $\underline{1}$ par 1. On considère la fonction exponentielle exp. Cette fonction définit bien un morphisme de $\mathcal{R}_1$ (le groupe additif des réels) vers $\mathcal{R}_2$ (le groupe multiplicatif des réels). Ce morphisme est en fait un plongement, et il respecte l'ordre de $\mathbb{R}$ – il serait donc aussi un plongement de $\mathcal{L} \cup \{<\}$-structures, si on munissait $\mathcal{R}_1$ et $\mathcal{R}_2$ de l'ordre induit par l'ordre de $\mathbb{R}$. Ce n'est pas un isomorphisme, car son image est un sous-groupe propre de $\mathbb{R} \setminus \{0\}$. [Notez que $\mathcal{R}_2$ muni de cet ordre n'est pas un groupe ordonné, car il devrait satisfaire par exemple $\forall x\,(x < 1 \to x^2 < 1)$, et -1 en donne un contre-exemple. Nous discuterons des groupes ordonnés plus tard dans le cours].

2.2 Termes et formules

Soit $\mathcal{L}$ un langage. A part les symboles de $\mathcal{L}$, nous avons à notre disposition des symboles de variables, des connecteurs ($\wedge$, $\neg$; à partir desquels nous définissons $\vee$, $\to$ et $\leftrightarrow$), le symbole d'égalité, et enfin le quantificateur $\exists$, "il existe" (à partir duquel on définira le quantificateur $\forall$, "pour tout").

Définition 2.12. La collection des termes du langage est le plus petit ensemble $\mathcal{T}erm$ satisfaisant les conditions suivantes :
Toute variable est dans $\mathcal{T}erm$;
toute constante est dans $\mathcal{T}erm$;
si $t_1, \ldots, t_n$ sont dans $\mathcal{T}erm$ et $f \in \mathcal{L}$ est une fonction n-aire, alors $f(t_1, \ldots, t_n)$ est dans $\mathcal{T}erm$.

2.13. Commentaires. Un terme est donc construit par induction : on commence avec des symboles de variables ou de constantes, et on applique les fonctions du langage. Il existe une notion naturelle de *compléxité* d'un terme, qui associe aux constantes et variables la valeur 0, et à un terme $f(t_1, \ldots, t_n)$ la valeur 1+ le sup des compléxités des t_i.

On rajoute suffisamment de parenthèses pour que l'écriture soit unique.

Si t est un terme, la notation $t(x_1, \ldots, x_n)$ voudra dire que les seules variables apparaissant dans l'écriture du terme sont parmi $x_1, \ldots, x_n$. Mais elles n'y apparaissent pas nécéssairement, et donc un terme en $(x_1, \ldots, x_n)$ pourra aussi être vu comme un terme en $(x_1, \ldots, x_{n+1})$, ou en $(x_1, \ldots, x_n, y_{24}, z, u)$ si on veut.

Exemple 2.14. $\mathcal{L} = \{+, \cdot, -, 0, 1\}$ le langage des anneaux. D'après notre définition, un terme s'écrira par exemple

$$-(\cdot(+(x, y), x_1), +(1 + 1))$$

ce qui n'est pas extrêmement facile à lire. Une amélioration serait déja obtenue en utilisant la notation des opérations au lieu de celle des fonctions, i.e., $(x + y)$ au lieu de $+(x, y)$, ce qui donnerait

$$(((x + y) \cdot x_1) - (1 + 1)).$$

Il faut noter que les termes $(x + y) + z$ et $x + (y + z)$ sont différents, et c'est pour cela que nous sommes obligés d'utiliser des parenthèses. Quand on travaille dans une structure donnée (par exemple un anneau), on pourra simplifier la notation et écrire tout simplement $(x + y)x_1 - 2$.

2.15. Soient $\mathcal{L}$ un langage, et M une $\mathcal{L}$-structure. Un terme $t(x_1, \ldots, x_n)$ du langage s'interprète alors naturellement comme une fonction $t^M : M^n \to M$. Ceci est montré par induction sur la complexité du terme :

La variable x_i définit une fonction $M \to M$ (la fonction identité); la constante c définit une fonction $M^0 \to M$, qui prend la valeur c^M; et enfin supposons $t_1^M, \ldots, t_m^M$ définies (où $t_1, \ldots, t_m$ sont des termes ne faisant intervenir que les variables $x_1, \ldots, x_n$) et $f \in \mathcal{L}$ une fonction m-aire, alors si $t = f(t_1, \ldots, t_m)$, on aura

$$t^M(x_1, \ldots, x_n) = f^M(t_1^M(x_1, \ldots, x_n), \ldots, t_m^M(x_1, \ldots, x_n)).$$

Exemple 2.16. On considère $\mathbb{R}$ avec sa $\mathcal{L}$-structure naturelle, $\mathcal{L} = \{+, -, \cdot, 0, 1\}$. Alors les fonctions correspondant aux termes du langage sont tout simplement les polynômes à coefficients ... dans $\mathbb{Z}$. En effet, tous les éléments de $\mathbb{Z}$ sont des termes. On remarque aussi que la plus petite sous-$\mathcal{L}$-structure de $\mathbb{R}$ est $\mathbb{Z}$.

Si on avait considéré le langage $\mathcal{L}_0 = \{+, -, \cdot\}$, alors il y aurait moins de fonctions, puisque les éléments de $\mathbb{Z}$ ne sont pas représentés par des termes, et que $\emptyset$ est une sous-structure de $\mathbb{R}$ (si on admet les $\mathcal{L}$-structures vides). On obtiendrait alors tous les polynômes à coefficients dans $\mathbb{Z}$ et avec coefficient constant égal à 0.
On peut aussi avoir plus de fonctions : si maintenant on considère $\mathbb{R}$ avec sa $\mathcal{L}(\mathbb{R})$-structure naturelle, les fonctions correspondant aux termes du langage seront alors tous les polynômes à coefficients dans $\mathbb{R}$.

Définition 2.17. Une formule est une chaîne de caractères pris parmi les symboles du langage $\mathcal{L}$ et les symboles logiques mentionnés ci-dessus, et qui obéit à certaines règles. Une formule sera construite à partir de blocs de base, en utilisant certaines opérations. L'ensemble des formules est défini par induction, comme le plus petit ensemble contenant les formules atomiques, et clos par conjonction et quantification. La formulation précise est donnée en (3) ci-dessous, pour un langage $\mathcal{L}$ fixé.

(1) Une formule *atomique* est une formule qui est de la forme $t_1 = t_2$, ou bien $R(t_1, \ldots, t_n)$, où $t_1, t_2, \ldots, t_n$ sont des termes du langage, et $R \in \mathcal{L}$ une relation n-aire.
(2) On appelle formule *négatomique* une formule de la forme $\neg\varphi$, où φ est atomique. Ce n'est pas une terminologie très souvent utilisée, et elle ne sert pas pour la définition de l'ensemble des formules.
(3) L'ensemble des formules est le plus petit ensemble $\mathcal{F}orm$ contenant les formules atomiques, et satisfaisant les conditions suivantes :
   (i) Si $\varphi \in \mathcal{F}orm$ alors $(\neg\varphi) \in \mathcal{F}orm$;
   (ii) si $\varphi_1, \varphi_2 \in \mathcal{F}orm$ alors $(\varphi_1 \wedge \varphi_2) \in \mathcal{F}orm$;
   (iii) si $\varphi \in \mathcal{F}orm$ et v est une variable, alors $(\exists v\, \varphi) \in \mathcal{F}orm$.

2.18. Abbréviations J'utiliserai constamment les abbréviations suivantes (où ϕ, ψ dénotent des formules) :

$(\phi \vee \psi)$ pour $\neg((\neg\phi) \wedge (\neg\psi))$;
$(\phi \to \psi)$ pour $(\neg\phi) \vee \psi$ (c'est à dire : $\neg(\phi \wedge (\neg\psi))$) ;
$\forall v\ \phi$ pour $\neg(\exists v\ (\neg\phi))$.

De plus on omettra parfois des quantificateurs quand ils ne changent pas : c'est-à-dire $\exists x_1 \exists x_2\, \phi$ pourra être abbrévié par $\exists x_1, x_2\, \phi$. De même, $\forall x, y, z\, \phi$ sera une abbréviation pour $\forall x \forall y \forall z\, \phi$.

Remarques 2.19. 1 - L'ensemble des formules est donc aussi clos par disjonction et quantification universelle : si φ_1, φ_2 sont des formules, alors aussi $(\varphi_1 \vee \varphi_2)$ et $(\forall v\, \varphi_1)$.

2 - Pour que l'écriture d'une formule soit non ambigüe, j'utilise des parenthèses ; on peut bien sûr omettre sans problème les parenthèses extérieures. On peut aussi donner des ordres de préséance aux opérations. Par exemple, on considère souvent que les quantificateurs portent sur aussi peu que possible. Donc $\exists v\, \varphi \wedge \psi$ serait une abbréviation pour $((\exists v\, \varphi) \wedge \psi)$.

3 - L'écriture $\exists v\, \varphi$ n'implique pas que la variable v apparaît dans φ.

4 - La formule $\exists v\, (\exists v\, \phi)$ est permise (mais la quantification extérieure n'apporte rien).

5 - On aurait pu mettre comme conditions que $\mathcal{F}orm$ soit close par $\vee$, $\neg$ et quantification universelle, cela n'aurait pas changé l'ensemble de formules, seulement leur écriture. Cependant, il aurait fallu faire attention au moment de la définition de la satisfaction.

Définition 2.20. Soit φ une formule. Une variable peut apparaître de façon *liée* ou *libre* dans une formule, le "ou" n'étant pas exclusif. La définition est faite par induction sur la compléxité des formules, de la façon suivante :
Si φ est une formule atomique, alors toutes les variables apparaissant dans φ ont toutes leurs occurrences libres, et aucune variable n'a d'occurrence liée ;
si $\varphi = \neg\psi$, alors les occurrences libres de la variable v dans φ sont celles de ψ, et de même pour les occurrences liées ;
si $\varphi = \psi \wedge \theta$, alors les occurrences libres de la variable v dans φ est la réunion des occurrences libres de v dans ψ et dans θ ; de même pour les occurrences liées de la variable v ;
si $\varphi = \forall v\, \psi$, alors toutes les occurrences de v dans φ sont liées (et aucune n'est libre). La qualité (libre ou liée) des occurrences des autres variables apparaissant dans φ ne change pas.

Exemple 2.21. Soient A, B, C des relations (binaires) du langage. On considère les formules : $\varphi_1 := A(x, y)$; $\varphi_2 := \exists x\, (A(x, y) \wedge B(x, z))$; $\varphi_3 := \forall v\, A(x, y)$; $\varphi_4 := \exists x\, (A(x, y) \wedge \forall x\, (B(x, z) \wedge C(x, t))) \vee \neg(C(x, t))$.

Alors x, y sont libres dans φ_1 ; y et z sont libres dans φ_2, mais x est liée et n'a pas d'occurrence libre ; x et y sont libres dans φ_3 et v est liée (mais en fait cette quantification ne sert à rien) ; et pour φ_4, il faut vraiment la décomposer pour voir ce qui se passe : $\varphi_4 = (\exists x(\psi_1 \wedge \psi_2)) \vee \psi_3$. La variable x apparaît libre dans ψ_1 et liée dans ψ_2, donc de toute façon est liée dans $\exists x(\psi_1 \wedge \psi_2)$ donc dans φ_4. Elle apparaît libre dans ψ_3, donc aussi dans φ_4. Finalement toutes les autres variables apparaissent comme libres.

2.22. Commentaires. De même que pour les termes, il existe des notions de compléxité pour les formules, qu'on peut facilement formaliser. Pour montrer qu'une propriété P est vraie de

toutes les formules, on montrera d'abord qu'elle est vraie pour les formules atomiques (ce qui nécéssitera peut-être une induction sur la compléxité des termes) ; puis que si φ_1, φ_2 ont P alors aussi $(\varphi_1 \wedge \varphi_2)$ et $(\neg\varphi)$; et enfin, ce qui est souvent le plus difficile, que si φ a P alors aussi $\exists v\, \varphi$. (Ou bien, que si φ a P alors aussi $\forall v\, \varphi$.)

Définition 2.23. Un *énoncé* est une formule dans laquelle aucune variable n'apparaît de façon libre.

2.3 Satisfaction

Soient $\mathcal{L}$ un langage, M une $\mathcal{L}$-structure, et $\varphi(x_1, \ldots, x_n)$ une formule, $\bar{a} = (a_1, \ldots, a_n) \in M^n$. Nous allons définir par induction le fait que "M satisfait $\varphi(\bar{a})$", ou encore "$\bar{a}$ satisfait φ dans M", ou encore "M est un modèle de $\varphi(\bar{a})$", noté

$$M \models \varphi(\bar{a}).$$

Les énoncés ci-dessous vont vous sembler complètement triviaux, et c'est normal. La satisfaction d'une formule doit être celle que l'on attend si on lit la formule à voix haute. Le fait que les formules soient parfois écrites avec la même variable jouant des roles différents, comme dans la formule φ_4 ci-dessus, rend parfois les choses un peu compliquées, et c'est pourquoi on est obligé de définir la satisfaction de façon tout à fait formelle. L'induction est sur la complexité des formules, pas sur le nombre de variables libres.

2.24. Soient $\bar{x} = (x_1, \ldots, x_n)$ un uplet de variables, $\bar{a} \in M^n$, $t_1, t_2, \ldots$ des termes en $\bar{x}$, et R un symbole de relation m-aire.
Si $\varphi(\bar{x})$ est la formule atomique $t_1(\bar{x}) = t_2(\bar{x})$, alors $M \models t_1(\bar{a}) = t_2(\bar{a})$ si et seulement si $t_1^M(\bar{a}) = t_2^M(\bar{a})$.
Si $\varphi(\bar{x}) := R(t_1(\bar{x}), \ldots, t_m(\bar{x}))$ alors $M \models \varphi(\bar{a})$ si et seulement si le m-uplet $(t_1^M(\bar{a}), \ldots, t_m^M(\bar{a}))$ est dans R^M.

Cela règle le cas des formules atomiques. Celui des combinaisons booléennes est aussi très simple :

Si $\varphi(\bar{x}) = \varphi_1(\bar{x}) \wedge \varphi_2(\bar{x})$ alors $M \models \varphi(\bar{a})$ si et seulement si $M \models \varphi_1(\bar{a})$ et $M \models \varphi_2(\bar{a})$.

Si $\varphi(\bar{x}) = \neg\psi(\bar{x})$ alors $M \models \varphi(\bar{x})$ si et seulement si M ne satisfait pas la formule $\psi(\bar{a})$, noté $M \not\models \psi(\bar{a})$.

Si $\varphi(\bar{x}) = \exists y\, \psi(\bar{x}, y)$, alors $M \models \varphi(\bar{a})$ si et seulement il existe $b \in M$ tel que $M \models \psi(\bar{a}, b)$.

2.25. Commentaires. Il est clair que $M \models (\varphi_1 \vee \varphi_2)(\bar{a})$ ssi $M \models \varphi_1(\bar{a})$ ou $M \models \varphi_2(\bar{a})$. De même, $M \models \forall y\, \psi(\bar{a}, y)$ ssi pour tout $b \in M$, on a $M \models \psi(\bar{a}, b)$.

Puisque $M \models \neg\varphi(\bar{a})$ si et seulement si $M \not\models \varphi(\bar{a})$, on obtient

$$M \models \varphi(\bar{a}) \qquad \text{ou} \qquad M \models \neg\varphi(\bar{a}),$$

les deux possibilités étant bien sûr exclusives.

Exemple 2.26. On considère $\mathcal{L} = \{+, -, \cdot, 0, 1, <\}$, $\mathbb{Q} \subset \mathbb{R}$ avec la $\mathcal{L}$-structure naturelle. On regarde la formule $\varphi(x) = x > 0 \to (\exists y\, y^2 = x)$.

[Comme je travaille dans des anneaux, je m'autorise les abbréviations habituelles : y^2 est $y \cdot y$. Cette notation peut être un peu trompeuse car elle peut faire croire que tous les monômes ont la même compléxité, ce qui n'est pas vrai.]
Cette formule est toujours vraie dans $\mathbb{R}$: tout élément positif est en effet un carré, et donc

$\mathbb{R} \models \forall x\, \varphi(x)$. Elle est fausse dans $\mathbb{Q}$: en effet, 2 est positif, mais n'a pas de racine carrée dans $\mathbb{Q}$: $\mathbb{Q} \models \neg\varphi(1+1)$.

Définition 2.27. Une formule est *sans quantificateurs* ... si son écriture ne comporte pas de quantificateurs. Elle est donc obtenue à partir des formules atomiques en utilisant les opérations booléennes $\wedge$ et $\neg$ (et $\vee$).

Exercice 2.28. Soient $\mathcal{L}$ un langage, et $M \subseteq N$ des $\mathcal{L}$-structures. (La notation $\subseteq$ dans ce cas sous-entend que l'inclusion est une inclusion de $\mathcal{L}$-structures. Si ce n'est pas le cas, je le dirai explicitement).

(1) Soient $\varphi(x_1, \ldots, x_n)$ une formule sans quantificateurs, et $(a_1, \ldots, a_n)$ un n-uplet de M. Montrez que
$$M \models \varphi(a_1, \ldots, a_n) \qquad \text{ssi} \qquad N \models \varphi(a_1, \ldots, a_n).$$

(2) Soient $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ une formule sans quantificateurs et $(a_1, \ldots, a_n)$ un n-uplet de M. Montrez que si
$$M \models \exists y_1, \ldots, \exists y_m \varphi(a_1, \ldots, a_n, y_1, \ldots, y_m)$$
alors
$$N \models \exists y_1, \ldots, \exists y_m \varphi(a_1, \ldots, a_n, y_1, \ldots, y_m).$$
Donnez un exemple montrant que la réciproque est fausse.

2.29. Commentaire. Une formule comme dans le (2) de l'exercice s'appelle une *formule existentielle*. Sa négation (qui n'aura donc que des quantificateurs $\forall$) s'appelle une *formule universelle*. Attention aux pièges : la formule $(\forall x\, x^4 \neq 1) \to (\forall y\, y^2 \neq 1)$ n'est pas une formule universelle, car en fait le 1er quantificateur est vraiment un quantificateur existentiel.

Exercice 2.30. Soient $\mathcal{L}$ un langage, M une $\mathcal{L}$-structure, et $\mathcal{L}(M)$ le langage obtenu en ajoutant un nouveau symbole de constante pour chaque élément de M. On considère M avec sa $\mathcal{L}(M)$-structure naturelle. Soit $\mathrm{Diag}^+(M)$ l'ensemble des $\mathcal{L}(M)$-énoncés atomiques vrais dans M. On l'appelle aussi le *diagramme atomique de* M. Il s'agit donc de toutes les relations de la forme $t_1(\bar{a}) = t_2(\bar{a})$, ou bien $R(t_1(\bar{a}), \ldots, t_n(\bar{a}))$ (avec t_i des termes, $\bar{a}$ un uplet d'éléments de M de la bonne arité, $R \in \mathcal{L}$ une relation) qui sont vérifiées dans M.

Montrez qu'une $\mathcal{L}(M)$-structure N satisfait tous les énoncés de $\mathrm{Diag}^+(M)$ si et seulement s'il existe un $\mathcal{L}$-morphisme $h : M \to N$.

2.31. Notation plus formelles. Je vais introduire une notation plus formelle pour le fait de remplacer des variables par des éléments d'un modèle. On pense à une assignation (= traduction de l'anglais "assignement"), aussi appellé "affectation", comme à une fonction de l'ensemble Var de toutes les variables dans la $\mathcal{L}$-structure M. Donc, au lieu d'être un uplet fini, ce sera un uplet infini.

Si $\alpha : \mathrm{Var} \to M$ est une assignation, et φ une formule, alors la notation $\varphi(\alpha)$ voudra dire que nous avons remplacé dans φ, chaque occurrence d'une variable libre par la valeur de α en cette variable. Donc, quand je définissais $M \models \varphi(a_1, \ldots, a_n)$ pour la formule φ dont les variables libres étaient (parmi) $x_1, \ldots, x_n$, ce que j'écrivais $M \models \varphi(\bar{a})$ (avec $\bar{a} = (a_1, \ldots, a_n)$) aurait aussi pu être écrit $M \models \varphi(\alpha)$, avec α une assignation telle que $\alpha(x_i) = a_i$, sa valeur dans les autres variables n'ayant aucune importance.
Dans la pratique j'utiliserai peu cette terminologie, et me restreindrai à des assignations ayant domaine fini. Typiquement, si nous avons une formule dont nous écrivons les variables libres comme $(x_1, x_2, \ldots, x_n)$, alors une assignation sera un n-uplet d'éléments de M, dont le premier correspondra à x_1, le 2ème à x_2, etc. Si nous écrivons les variables libres $(x_1, x_2, y, z, \ldots)$ alors le 3ème élément du n-uplet correspondra à y, et le 4ème à z.

Notation. Si $\alpha : \mathrm{Var} \to M$ est une assignation, et $b \in M$, x une variable, alors $\alpha(b/x)$ est défini par

$$\alpha(b/x)(y) = \begin{cases} b & \text{si } y = x, \\ \alpha(y) & \text{sinon.} \end{cases}$$

2.32. Je vais donner la preuve de la 2ème partie de 2.25 : $M \models \forall y\, \varphi(y, \bar{a})$ ssi pour tout $b \in M$ on a $M \models \varphi(b, \bar{a})$. Ici, cette notation veut dire que toutes les occurrences **libres** de y sont remplacées par b. J'utiliserai la notation des assignations introduite ci-dessus. Rappelons que $\forall x \varphi$ est une abbréviation pour $\neg(\exists x(\neg\varphi))$.

Soient M une $\mathcal{L}$-structure, α une assignation. Alors $M \models \forall y \varphi(\alpha)$
ssi $M \not\models (\exists y\, \neg(\varphi))(\alpha)$,
ssi il n'existe pas de $b \in M$ tel que $M \models (\neg\varphi)(\alpha(b/y))$,
ssi pour tout $b \in M$ on a $M \not\models (\neg\varphi)(\alpha(b/y))$,
ssi pour tout $b \in M$ on a $M \models \varphi(\alpha(y/b))$, qui est ce que nous voulions montrer.

La preuve de la 1ère partie est similaire.

2.4 Théories et modèles, réduits, expansions

Nous introduisons ici un peu de terminologie. Soient $\mathcal{L}$ un langage, M une $\mathcal{L}$-structure, et Σ un ensemble d'énoncés (du langage $\mathcal{L}$), $\mathcal{K}$ une classe de $\mathcal{L}$-structures, et enfin $\varphi(x_1, \ldots, x_n)$ une formule.

A partir de maintenant, toutes les structures seront non vides

Définition 2.33. (1) M est un modèle de Σ, noté $M \models \Sigma$, ssi M satisfait tous les énoncés de Σ (donc, pour tout $\psi \in \Sigma$, on a $M \models \psi$). On note (temporairement) la classe des modèles de Σ par $\mathrm{Mod}(\Sigma)$.

(2) $\mathrm{Th}(M)$, la *théorie de* M, est l'ensemble des énoncés satisfaits par M. Si $N \models \mathrm{Th}(M)$, on notera $N \equiv M$ (se lit : N *est élémentairement équivalent à* M).

(3) La *théorie de* $\mathcal{K}$ est l'ensemble des énoncés vrais dans tous éléments de $\mathcal{K}$. Autrement dit, $\mathrm{Th}(\mathcal{K}) = \bigcap_{K \in \mathcal{K}} \mathrm{Th}(K)$.

(4) On a $\mathcal{K} \subseteq \mathrm{Mod}(\mathrm{Th}(\mathcal{K}))$ et $\Sigma \subseteq \mathrm{Th}(\mathrm{Mod}(\Sigma))$. De plus, on a $\mathrm{Th}(\mathcal{K}) = \mathrm{Th}(\mathrm{Mod}(\mathrm{Th}(\mathcal{K})))$ et $\mathrm{Mod}(\Sigma) = \mathrm{Mod}(\mathrm{Th}(\mathrm{Mod}(\Sigma)))$.

(5) Quel sens donner à $M \models \varphi(x_1, \ldots, x_n)$? Nous dirons $M \models \varphi(x_1, \ldots, x_n)$ si pour tous $a_1, \ldots, a_n$ dans M, on a $M \models \varphi(a_1, \ldots, a_n)$. Autrement dit, la satisfaction d'une formule ayant des variables libres auxquelles on n'a pas donné de valeur dans M, est obtenue en considérant que ces variables sont quantifiées universellement. $M \models \varphi(x_1, \ldots, x_n)$ ssi $M \models \forall x_1, \ldots, x_n \, \varphi(x_1, \ldots, x_n)$.

Définition 2.34. Soient $\mathcal{L} \subseteq \mathcal{L}'$ des langages. Si M est une $\mathcal{L}$-structure, on appelle *expansion* de M une $\mathcal{L}'$-structure $\mathcal{M}$ ayant même univers que M, et telle que tous les symboles de $\mathcal{L}$ ont la même interprétation dans $\mathcal{M}$ et dans M. On dira aussi que M est un *réduit de $\mathcal{M}$ au langage $\mathcal{L}$*, noté $\mathcal{M}_{|\mathcal{L}}$.

Exemple 2.35. Le groupe additif des réels $(\mathbb{R}, +, -, 0)$ est un réduit de l'anneau des réels $(\mathbb{R}, +, -, \cdot, 0, 1)$.

Définition 2.36. Soient $\mathcal{L}$ un langage, M une $\mathcal{L}$-structure, n un entier ≥ 1. Un sous-ensemble D de M^n est *définissable* s'il existe une formule $\varphi(\bar{x})$ ($\bar{x}$ un n-uplet de variables) telle que

$$D = \{\bar{a} \in M^n \mid M \models \varphi(\bar{a})\}.$$

Soit $A \subseteq M$. Un sous-ensemble D de M^n est *définissable sur A*, ou bien *avec paramètres dans A*, s'il existe une formule $\varphi(\bar{x}, \bar{y})$ ($\bar{x}$ un n-uplet de variables, $\bar{y}$ un uplet de variables), et un uplet $\bar{b}$ dans A (de même longueur que $\bar{y}$) telle que

$$D = \{\bar{a} \in M^n \mid M \models \varphi(\bar{a}, \bar{b})\}.$$

Remarque 2.37. Si $\mathrm{Def}(M)$ dénote la réunion, sur $n \in \mathbb{N}$, des sous-ensembles définissables de M^n, alors on observe que $\mathrm{Def}(M)$ est close par : combinaisons booléennes ($\cap$, $\cup$, complémentaire ; ces opérations correspondant aux connecteurs $\wedge$, $\vee$ et $\neg$). Mais aussi par *projection* : Si $D \subseteq M^{n+1}$ est définissable, et π est la projection sur les n premières coordonnées, alors $\pi(M)$ est aussi définissable. La projection correspond au quantificateur existentiel.

2.5 Substitutions et formules universellement valides

Définition 2.38. Soit $t = t(v_1, \ldots, v_n)$ un terme. Soit maintenant $\varphi(x, \ldots,)$ une formule. Nous allons définir par induction sur la compléxité de φ la substitution de x par t dans φ. Intuitivement, nous avons le droit de remplacer seulement les occurrences libres de x par t — nous ajouterons alors peut-être des variables libres à notre formule : en effet, x ne sera plus une variable libre de la nouvelle formule, mais les variables de t seront des variables libres. Très formellement, voici les règles :

(1) si φ est atomique, on remplace toutes les occurrences de x par t pour obtenir $\varphi(t/x)$;

(2) si $\varphi = \neg\psi$, alors $\varphi(t/x) = \neg(\psi(t/x))$;
(3) si $\varphi = \varphi_1 \wedge \varphi_2$, alors $\varphi(t/x) = \varphi_1(t/x) \wedge \varphi_2(t/x)$;
(4) si $\varphi = \exists v\, \psi$, alors il y a plusieurs cas à distinguer :
  (i) si $v = x$, alors on ne fait rien : $\varphi(t/x) = \varphi$;
  (ii) si $v \neq x$ et si v n'apparaît pas parmi les variables de t, alors on pose $\varphi(t/x) = \exists v\, \psi(t/x)$;
  (iii) si $v \neq x$ et si v est une des variables de t, alors on définit $\varphi(t/x) = \exists u\, \psi(u/v)(t/x)$, où u est une nouvelle variable qui n'apparaît pas dans ψ ni dans t.

  Dans le dernier sous-cas, on a d'abord changé la variable v en u, pour se ramener au sous-cas précédent.

Remarque 2.39. Le dernier sous-cas semble bien compliqué. On peut l'éviter, mais cela a un coût. Les lemmes suivants qui parlent des substitutions devraient alors faire des hypothèses supplémentaires sur la variable quantifiée.

Remarque 2.40. Soit φ une formule dans laquelle la variable y n'apparaît pas du tout. Alors si $\psi = \varphi(x/y)$, on aura aussi $\varphi = \psi(y/x)$.
Ceci est faux si y apparaît dans φ : $\exists y\, y = x$ devient $\exists x\, x = x$.

Proposition 2.41. *(Lemme de substitution). Soient M une $\mathcal{L}$-structure, x une variable, s, t des termes, φ une formule, et $\bar{a}$ un uplet de M, correspondant à un uplet de variables contenant les variables de s, t, φ et la variable x. Si $c \in M$, on note $\bar{a}(c/x)$ le uplet obtenu à partir de $\bar{a}$ en remplaçant l'élément correpondant à la variable x par c.*

(1) $t(s/x)(\bar{a}) = t(\bar{a}(s(\bar{a})/x))$.
(2) *Alors* $M \models \varphi(s/x)(\bar{a})$ *ssi* $M \models \varphi(\bar{a}(s(\bar{a})/x))$.

Démonstration. (1) Clair ?

(2) La preuve est par induction sur la compléxité de φ, le seul cas délicat étant celui de $\varphi = \exists v\, \psi$ avec $v \neq x$. Ecrivons les variables libres de ψ explicitement comme $(v, \bar{y})$. Si v n'apparaît pas dans le terme s, alors $\varphi(s/x) = \exists v\, \psi(s/x)$ par définition. De plus, dans ce cas, $s(\bar{a}) = s(b/v)(\bar{a})$ pour tout $b \in M$, puisque $s = s(\bar{y})$ ne dépend pas de la variable v. Alors, $M \models \varphi(s/x)(\bar{a})$ ssi il existe $b \in M$ tel que $M \models \psi(s/x)(b, \bar{a})$, ssi il existe $b \in M$ tel que $M \models \psi(b, \bar{a}(s(\bar{a})/x))$ (par induction), ssi il existe $b \in M$ tel que $M \models \psi(b, \bar{a}(s(\bar{a})/x))$, ssi $M \models \exists v\, \psi(v, \bar{a}(s(\bar{a})/x))$.

Supposons maintenant que v soit une des variables de s. Il faut d'abord noter que $M \models \exists v\, \varphi$ ssi $M \models \exists u\, \varphi(u/v)$, quand u est une variable qui n'apparait pas parmi les variables de φ. La définition de $\varphi(s/x)$ nous permet alors de conclure en utilisant le cas précédent.

Exercice 2.42. Soient $t = t(v_1, \ldots, v_n)$ un terme, φ une formule et x une variable. Nous dirons que t *est libre pour x dans φ* si aucune occurrence libre de x n'apparaît dans une sous-formule de φ dans laquelle une des variables v_i est quantifiée. Plus formellement :

Si φ est atomique, alors t est libre pour x dans φ ;

t est libre pour x dans $\neg\psi$ ssi il est libre pour x dans ψ; t est libre pour x dans $\varphi_1 \wedge \varphi_2$ ssi il est libre dans φ_1 et dans φ_2 ;

t est libre pour x dans $\forall v\,\psi$ si :

ou bien x n'a pas d'occurrence libre dans ψ

ou bien v est distinct des variables $v_1, \ldots, v_n$ de t, et t est libre pour x dans ψ.

Montrez que si φ, t et x sont comme ci-dessus, et si t est libre pour x dans φ, alors $\varphi(t/x)$ est exactement obtenue en remplaçant toutes les occurrences libres de x par t dans φ. C'est à dire, le sous-cas (iii) n'apparaît pas, quand on définit $\varphi(t/x)$ en utilisant l'induction sur les sous-formules de φ.

Exemple 2.43. Voici un exemple qui montre le problème. On considère une formule sans quantificateurs $\psi(x, y)$, et la formule $\varphi = \exists y\,\psi(x, y)$. La variable x est donc libre dans φ : on voudrait faire la substitution de x par y. La règle de substitution (cas (iii)) nous donne $\varphi(x)(y/x) = \exists u\,\psi(y, u)$ (si on appelle u la nouvelle variable).

Supposons qu'au lieu de faire cela nous ayions tout simplement remplacé l'occurrence de x par y. On aurait obtenu la formule (en fait un énoncé) $\theta = \exists y\,\psi(y, y)$.

Il est clair que θ n'est pas équivalente à $\varphi(x)(y/x)$, et on peut facilement construire un modèle M et une assignation α tels que $M \models \varphi(y/x)(\alpha)$, et $M \models \neg\theta$. (On remarque ici que la satisfaction de θ est indépendante de l'assignation α.)

2.44. Tautologies. En principe, vous avez déjà vu le calcul propositionnel (abbrévié CP). On regarde des expressions formées à partir de lettres (les "variables"), des connecteurs $\wedge$ et $\neg$ (à partir desquels on peut former $\vee$ et $\rightarrow$), et de parenthèses. Et on regarde les valeurs de vérité de ces expressions, en fonction des valeurs (0 ou 1; avec 0 correspondant à "Faux", et 1 à "Vrai") données aux lettres de l'expression. On construit alors des *tables de vérité* :

A	B	$\neg A$	$A \wedge B$
0	0	1	0
0	1	1	0
1	0	0	0
1	1	0	1

Les formules du calcul propositionnel sont définies par induction : si α et β sont des formules, alors aussi $(\neg\alpha)$ et $(\alpha \wedge \beta)$.

Etant donné une fonction f de l'ensemble des lettres (ou variables) et à valeurs dans $\{0, 1\}$, on peut définir la *valeur de vérité* $\tilde{f}$ des formules du CP, par induction sur la compléxité. Soit α une formule. Si α est une variable, alors $\tilde{f}(\alpha) = f(\alpha)$. Si $\alpha = (\beta \wedge \gamma)$ alors $\tilde{f}(\alpha) = \tilde{f}(\beta)\tilde{f}(\gamma)$; si $\alpha = \neg\beta$, alors $\tilde{f}(\alpha) = 1 - \tilde{f}(\beta)$. On en déduit facilement que $\tilde{f}(\alpha \vee \beta) = \sup\{\tilde{f}(\alpha), \tilde{f}(\beta)\}$, que $\tilde{f}(\alpha \rightarrow \beta) = \sup\{1 - \tilde{f}(\alpha), \tilde{f}(\beta)\}$, et que $\tilde{f}(\alpha \leftrightarrow \beta) = \sup\{\tilde{f}(\alpha)\tilde{f}(\beta), (1 - \tilde{f}(\alpha))(1 - \tilde{f}(\beta))\}$.

Une *tautologie* est une formule (du CP) qui est toujours vraie (c'est-à-dire, prend toujours la valeur 1), quelle que soit la valeur assignée à ses variables. En voici quelques exemples :

Exemple 2.45. $A \vee (\neg A)$;
$(A_1 \to (A_2 \to (\cdots (A_n \to B) \cdots))) \leftrightarrow ((A_1 \wedge A_2 \wedge \cdots \wedge A_n) \to B)$;
$A \to (B \to A)$;
$A \to A \vee B$.

Définition 2.46. Soient $\mathcal{L}$-un langage, Γ un ensemble de formules, et φ une formule.

(1) φ est *universellement valide*, ou bien *logiquement valide*, noté $\models \varphi$ si pour toute $\mathcal{L}$-structure M, on a $M \models \varphi$. [Rappel : si φ a des variables libres, alors on quantifie universellement sur ces variables libres.]
(2) $\Gamma \models \varphi$ ssi pour toute $\mathcal{L}$-structure M, si $M \models \Gamma$ alors aussi $M \models \varphi$.

2.47. Règles de déduction. Nous allons formaliser la notion de *preuve*, et de *théorème* de notre logique (qui est appellée *Calcul des prédicats*; à ne pas confondre avec le *Calcul propositionnel*). Nous commençons pas des définitions et des lemmes presque tous évidents.

Définition 2.48. Tout d'abord les *tautologies du langage* $\mathcal{L}$ sont obtenues à partir des formules du calcul propositionnel en remplaçant les lettres par des $\mathcal{L}$-formules. On aura donc, pour φ, ψ des $\mathcal{L}$-formules, que $\varphi \vee \neg\varphi$ et $\varphi \to (\psi \to \varphi)$ sont des tautologies.

Lemme 2.49. *Si la $\mathcal{L}$-formule φ est une tautologie, alors $\models \varphi$.*

Démonstration. Evident.

Lemme 2.50. *(Axiomes de l'égalité) Les énoncés suivants sont universellement valides :*
– $\forall x\, x = x$;
– $\exists x\, x = x$; *(d'où la nécéssité d'interdire les structure vides)*
– $\forall x, y\, (x = y \to y = x)$;
– $\forall x, y, z\, \big((x = y) \wedge (y = z)\big) \to (x = z)$;
– $\forall x_1, \ldots, x_n, y_1, \ldots, y_n \left(\bigwedge_{i=1}^n (x_i = y_i)\right) \to (f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n))$, *si $f \in \mathcal{L}$ est une fonction n-aire.*
– $\forall x_1, \ldots, x_n, y_1, \ldots, y_n \left(\bigwedge_{i=1}^n (x_i = y_i) \wedge R(x_1, \ldots, x_n)\right) \to R(y_1, \ldots, y_n)$, *si $R \in \mathcal{L}$ est une relation n-aire.*

[$\bigwedge_{i=1}^n$ se lit "conjonction de $i = 1$ à n"; $\bigwedge_{i=1}^n (x_i = y_i)$ est donc une abbréviation pour $(x_1 = y_1) \wedge (x_2 = y_2) \wedge \cdots \wedge (x_n = y_n)$. De même pour la disjonction de $i = 1$ à n, $\bigvee_{i=1}^n$.]

Remarque 2.51. On en déduit que si $\varphi(x_1, \ldots, x_n)$ est une $\mathcal{L}$-formule sans quantificateurs, alors
$\models \forall x_1, \ldots, x_n, y_1, \ldots, y_n\, (\bigwedge_{i=1}^n (x_i = y_i) \wedge \varphi \to \varphi(y_1/x_1, \ldots, y_n/x_n)$.

Lemme 2.52. *(Axiome du quantificateur existentiel) Soient φ une formule, x une variable et t un terme. Alors la formule $\varphi(t/x) \to \exists x\, \varphi$ est universellement valide.*

Démonstration. Soient $\bar{y}$ l'union des variables libres de φ et de t, et $\bar{a}$ un uplet de même longueur que $\bar{y}$, dans une $\mathcal{L}$-structure M. Si $M \models \varphi(t/x)(\bar{a})$, alors $M \models \varphi(\bar{a}(t(\bar{a})/x))$ par le lemme de substitution 2.41. Donc $M \models \exists x\, \varphi(\bar{a})$, le témoin de cette existence étant $b = t(\bar{a})$.

Lemme 2.53. *(Modus Ponens) Si φ et $(\varphi \to \psi)$ sont universellement valides, alors ψ l'est aussi.*

Démonstration. Evident.

Lemme 2.54. *($\exists$-introduction) Si la variable x n'a pas d'occurrence libre dans ψ, et si $\varphi \to \psi$ est universellement valide, alors $(\exists x\, \varphi) \to \psi$ l'est aussi.*

Démonstration. Soient M une $\mathcal{L}$-structure, $\bar{x}$ l'union de x et des variables libres de φ et ψ, et $\bar{a}$ un uplet dans M de même longueur que $\bar{x}$.

On suppose $M \models \exists x\, \varphi(\bar{a})$; donc il existe $b \in M$ tel que $M \models \varphi(\bar{a}(b/x))$, et cela entraine que $M \models \psi(\bar{a}(b/x))$. Comme x n'est pas libre dans ψ, on a $M \models \psi(\bar{a}(b/x))$ ssi $M \models \psi(\bar{a})$. Et donc $M \models \psi(\bar{a})$, ce qui montre que $M \models ((\exists x\, \varphi) \to \psi)(\bar{a})$.

2.6 Preuves formelles

Nous avons à notre disposition plusieurs outils :
Les *axiomes logiques* :

- Les tautologies ;
- les axiomes de l'égalité (2.50) ;
- l'axiome du quantificateur existentiel (2.52) ;

des règles de déduction :

- Le modus ponens (2.53) ;
- $\exists$-introduction (2.54) : si x n'est pas libre dans ψ alors de $\varphi \to \psi$ on peut déduire $(\exists x\, \varphi) \to \psi$.

Définition 2.55. Soient φ une formule, et T un ensemble d'énoncés. Une *preuve formelle de φ dans T* est une suite finie de formules $\{\varphi_1, \ldots, \varphi_n\}$ avec $\varphi_n = \varphi$, et telle que pour tout $i \leq n$,
ou bien $\varphi_i \in T$,
ou bien φ_i est un axiome logique,
ou bien φ_i s'obtient par Modus Ponens (MP) à partir de φ_j et φ_k, pour des entiers $j, k < i$,
ou bien φ_i s'obtient par $\exists$-introduction à partir d'un φ_j, avec $j < i$.

On dit que φ est *prouvable à partir de T*, noté $T \vdash \varphi$, s'il existe une preuve formelle de φ dans T. Si $T = \emptyset$, on note $\vdash \varphi$.

Remarque 2.56. Je ne précise pas le langage explicitement - il est entendu que nous travaillons dans des preuves faites dans le langage $\mathcal{L}$. On pourrait imaginer que si $\mathcal{L} \subset \mathcal{L}'$, alors une formule ait une preuve dans $\mathcal{L}'$ mais pas dans $\mathcal{L}$. Nous verrons que ce n'est pas le cas. Mais pour l'instant nous travaillons tout le temps dans un langage fixé $\mathcal{L}$, sauf mention du contraire.

2.57. Le but du restant de cette section est de montrer que

$$\boxed{T \models \varphi \text{ ssi } \ T \vdash \varphi}$$

pour une théorie T (= ensemble d'énoncés) et une formule φ. En fait, nous avons déjà montré la moitié de cette équivalence :

Théorème 2.58. *Soient $\mathcal{L}$ un langage, T une $\mathcal{L}$-théorie et φ une $\mathcal{L}$-formule. Si $T \vdash \varphi$, alors $T \models \varphi$.*

Démonstration. Par induction sur la longueur de la preuve, en utilisant les lemmes 2.49 à 2.54.

Maintenant voici quelques définitions et propriétés simples :

Définition 2.59. (1) Une théorie T est *contradictoire* s'il existe une formule φ telle que $T \vdash \varphi$ et $T \vdash \neg\varphi$. (Anglicisme : *inconsistante*).

(2) Une théorie est *cohérente* si elle n'est pas contradictoire. (Anglicisme : *consistante*).

(3) Une théorie T est *complète*, si pour tout énoncé θ, soit $T \vdash \theta$, soit $T \vdash \neg\theta$.

Proposition 2.60. *Soient $\varphi_1, \ldots, \varphi_n, \varphi, \psi \ldots$ des formules, T une théorie (qui peut être vide), t un terme, x une variable.*

(0) *Si $T \vdash \varphi_i$ pour $i = 1, \ldots, n$ alors $T \vdash (\varphi_1 \wedge \cdots \wedge \varphi_n)$.*

(1) *Si $T \vdash \varphi_i$ pour $i = 1, \ldots, n$ et $T \vdash (\varphi_1 \wedge \cdots \wedge \varphi_n) \to \psi$, alors $T \vdash \psi$.*

(2) *(axiome $\forall$) $\vdash (\forall x \varphi) \to \varphi(t/x)$. C'est la contraposée de l'axiome du quantificateur existentiel.*

(3) *($\forall$-introduction) Si x n'est pas libre dans φ et si $T \vdash \varphi \to \psi$, alors $T \vdash \varphi \to \forall x\, \psi$.*

(4) *Si $T \vdash \varphi$ alors $T \vdash \forall x\, \varphi$.*

Démonstration. (0) En utilisant une induction sur n, il suffit de le prouver pour $n = 2$. On utilise le fait que $(A \to (B \to (A \wedge B)))$ est une tautologie du calcul propositionnel, puis deux fois le modus ponens, pour obtenir la preuve suivante :
$(\varphi_1 \to (\varphi_2 \to (\varphi_1 \wedge \varphi_2)))$
preuve de φ_1 à partir de T
$(\varphi_2 \to (\varphi_1 \wedge \varphi_2))$ (MP)
preuve de φ_2 à partir de T
$\varphi_1 \wedge \varphi_2$ (MP).

(1) Evident en utilisant (0) et MP.

(3) On utilise l'équivalence de $(A \to B)$ et $(\neg B) \to (\neg A)$ pour obtenir la preuve (dans laquelle j'ai "raccourci" les étapes de passage à la contraposée) :
preuve de $\varphi \to \psi$
$\neg\psi \to \neg\varphi$ (contraposée)
$\exists x\, (\neg\psi) \to \neg\varphi$ ($\exists$-introduction)
$\varphi \to \neg(\exists x\, (\neg\psi))$ (contraposée)
$\varphi \to \forall x\, \psi$ (équivalence entre $\forall x\, \psi$ et $\neg(\exists x\, \neg\psi)$).

(2) Preuve similaire, en utilisant la contraposée, et l'axiome $\exists$.

(4) On sait introduire le quantificateur $\forall$ dans une implication, par (3). Soit ψ un énoncé prouvable à partir de T. Par exemple, $\psi = \forall z\, z = z$. Voici une preuve :
ψ
preuve de φ
$\varphi \to (\psi \to \varphi))$ (tautologie $A \to (B \to A)$)
$\psi \to \varphi$ (MP)
$\psi \to \forall x\, \varphi$ (on utilise (3))
$\forall x\, \varphi$ (MP).

Corollaire 2.61. (1) *Si T est contradictoire, il existe un sous-ensemble fini de T qui est contradictoire (puisque toute preuve ne fait intervenir qu'un nombre fini d'éléments de T).*

(2) *Il suit qu'une théorie T est cohérente ssi tout sous-ensemble fini de T est cohérent.*

(3) *Les conditions suivantes sont équivalentes :*

(a) T est contradictoire,

(b) Pour toute formule φ, $T \vdash \varphi$,

(c) Il existe $\psi_1, \ldots, \psi_n \in T$ tels que $T \vdash \neg(\psi_1 \wedge \cdots \psi_n)$.

Démonstration. (1) et (2) sont évidents. Pour (3), supposons (a), et soit ψ telle que $T \vdash \psi$ et $T \vdash \neg\psi$. On utilise la tautologie $(A \to (\neg A \to B))$ appliquée à $A = \psi$ et $B = \varphi$, et le Modus Ponens deux fois pour déduire $T \vdash \varphi$. Cela prouve (b). Le fait que (b) implique (c) est clair, puisque T prouve n'importe quoi. Supposons (c). Par (0) du lemme précédent, on sait que $T \vdash \psi_1 \wedge \cdots \wedge \psi_n$, ce qui montre (a).

Proposition 2.62. *Soient T une théorie, χ un énoncé et φ une formule. Alors*

$$T \cup \{\chi\} \vdash \varphi \qquad \textit{ssi} \qquad T \vdash (\chi \to \varphi).$$

Démonstration. Si on a une preuve de $(\chi \to \varphi)$ à partir de T, alors aussi à partir de $T \cup \{\chi\}$, et donc on obtient une preuve de φ en utilisant MP.

Pour l'autre direction, on va montrer que si $\alpha_0, \ldots, \alpha_m = \varphi$ est une preuve de φ à partir de $T \cup \{\chi\}$, alors $T \vdash (\chi \to \alpha_i)$. La démonstration est par induction sur i. Si α_i est un axiome logique, alors $\vdash \alpha_i$, et donc $\vdash (\chi \to \alpha_i)$. Si $\alpha_i \in T \cup \{\chi\}$, il est clair que $T \vdash (\chi \to \alpha_i)$.

Si α_i est obtenue à partir de α_j et $\alpha_k = \alpha_j \to \alpha_i$ pour des $j, k < i$, alors on utilise l'hypothèse d'induction et la tautologie $[(A \to B) \wedge (A \to (B \to C)] \to (A \to C)$.

Et si $\alpha_i = (\exists x\, \psi) \to \theta$ est obtenue à partir de $\alpha_j = (\psi \to \theta)$, pour un $j < i$, en utilisant $\exists$-introduction, alors par induction, on a
$T \vdash \chi \to (\psi \to \theta)$,
$T \vdash \psi \to (\chi \to \theta)$ (tautologie $[A \to (B \to C)] \to [B \to (A \to C)]$)
$T \vdash (\exists x\, \psi) \to (\chi \to \theta)$ ($\exists$-introduction, car x n'apparaît pas libre dans $\chi \to \theta$)
$T \vdash \chi \to ((\exists x\, \psi) \to \theta)$ (tautologie).

Remarque 2.63. (Très utile) Soient T une théorie, φ un énoncé. Si $T \cup \{\varphi\}$ est contradictoire alors $T \vdash \neg\varphi$.

Démonstration. On sait que $T \cup \{\varphi\} \vdash \neg\varphi$, ce qui entraîne que $T \vdash \varphi \to (\neg\varphi)$. Mais $(A \to \neg A) \to (\neg A)$ est une tautologie, d'où on déduit $T \vdash \neg\varphi$ par MP.

Corollaire 2.64. *Une $\mathcal{L}$-théorie qui est cohérente et maximale parmi les $\mathcal{L}$-théorie cohérentes, est complète. (**AC**) Tout théorie cohérente est contenue dans une théorie complète.*

Démonstration. Soit T une théorie $\mathcal{L}$-théorie cohérente, et maximale pour ces propriétés. Alors si φ est un énoncé du langage, ou bien $\varphi \in T$ et alors $T \vdash \varphi$, ou bien $T \cup \{\varphi\}$ est contradictoire. Dans ce dernier cas, on a $T \vdash \neg\varphi$ par la remarque précédente.

Pour la deuxième partie, on considère l'ensemble des théories cohérentes contenant T. Cet ensemble est inductif : l'union d'une chaîne de théories cohérentes est cohérente. Il a donc un élément maximal, et cet élément maximal sera une théorie complète.

Définition 2.65. Une théorie T est *close par déduction* si pour tout énoncé φ, on a $T \vdash \varphi$ si et seulement si $\varphi \in T$. Toute théorie est contenue dans une plus petite théorie déductivement clos, qui sera appelée sa *clôture déductive* ou *clôture par déduction.*

On remarque qu'une théorie complète n'est pas forcément maximale cohérente ; mais sa clôture par déduction le sera.

Définition 2.66. Soient $\mathcal{L}$ un langage, T une $\mathcal{L}$-théorie qui est cohérente, et soit C un ensemble de constantes de $\mathcal{L}$. Nous disons que C est un *ensemble de témoins pour* T si pour toute formule $\varphi(x)$ (au plus x comme variable libre) il existe une constante $c \in C$ telle que $T \vdash (\exists x\varphi) \to \varphi(c/x)$. On dira aussi que T *admet* C *comme ensemble de témoins.*

2.67. Nous allons montrer deux choses : comment fabriquer, à partir d'une $\mathcal{L}$-théorie cohérente T, une autre théorie cohérente la contenant, et admettant un ensemble de constantes témoins. Puis montrer qu'on peut, à partir de ces nouvelles constantes, fabriquer un modèle de T.

Nous fixons le langange $\mathcal{L}$, la $\mathcal{L}$-théorie cohérente T, et un ensemble C de nouvelles constantes, de cardinalité $\mathrm{card}(\mathcal{L}) + \aleph_0$ (donc le sup de $\mathrm{card}(\mathcal{L})$, $\aleph_0$), et enfin $\bar{\mathcal{L}} = \mathcal{L} \cup C$.

Nous commençons par un petit lemme, qui servira à montrer qu'ajouter des constantes au langage n'affecte pas la cohérence de T.

Lemme 2.68. *Soit φ une $\mathcal{L}$-formule dont la seule variable libre est v. Si $T \vdash \varphi(c/v)$ (dans $\mathcal{L} \cup \{c\}$), alors $T \vdash \forall v\, \varphi$ (dans $\mathcal{L}$).*

Démonstration. Soit $\psi_1, \ldots, \psi_m = \varphi(c/v)$ une $\mathcal{L} \cup \{c\}$-preuve. On choisit une variable w qui n'apparaît dans aucune des ψ_i, et pour chaque $i \leq m$, on pose θ_i la $\mathcal{L}$-formule obtenue en remplaçant dans ψ_i toutes les occurrences de c par w. On voit que :

– si ψ_i est un axiome logique, alors aussi θ_i ;

– si $\psi_i \in T$, alors $\psi_i = \theta_i \in T$;

– si ψ_i est obtenue par (MP) à partir de ψ_j et ψ_k, alors θ_i est obtenue par (MP) à partir de θ_j et θ_k ;

– si ψ_i est obtenue par $\exists$-introduction à partir de ψ_j, alors aussi θ_i à partir de θ_j.

Il suit que $\theta_1, \ldots, \theta_m$ est une preuve dans $\mathcal{L}$. Nous avons que $\theta_m = \varphi(w/v)$, et donc par 2.60(4), T prouve $\forall v\, \varphi$ (en fait, il faudrait donner un peu plus de détails : T prouve $\forall w\, \varphi(w/v)$, donc elle prouve $\varphi(v/w)(w/v)$ par l'axiome $\forall$, mais $\varphi(v/w)(w/v) = \varphi$ et donc par 2.60(4) elle prouve $\forall v\, \varphi$).

Corollaire 2.69. *Si T est cohérente en tant que $\mathcal{L}$-théorie, alors elle est aussi cohérente en tant que $\bar{\mathcal{L}}$-théorie.*

Démonstration. Il suffit d'ajouter les constantes une à la fois, puisque une preuve de contradiction n'utiliserait qu'un nombre fini de constantes. Le résultat suit du lemme précédent, car si T n'était pas $\mathcal{L} \cup \{c\}$-cohérente, alors elle prouverait $(v \neq v)(c/v)$ dans $\mathcal{L} \cup \{c\}$, et donc par le lemme, elle prouverait $\forall v\, v \neq v$ dans $\mathcal{L}$, i.e., serait déjà contradictoire dans $\mathcal{L}$.

Remarque 2.70. Le lemme 2.68 nous dit que si c est un symbole de constante qui n'est pas dans $\mathcal{L}$, et si T est une $\mathcal{L}$-théorie et $\varphi(x)$ une $\mathcal{L}$-formule, alors $T \vdash \varphi(c/x)$ implique $T \vdash \forall x\, \varphi$. On peut généraliser ce résultat à une formule $\varphi(x_1, \ldots, x_n)$ et à des constantes $c_1, \ldots, c_n$ n'appartenant pas à $\mathcal{L}$, par induction sur n. On obtient alors le résultat suivant, obtenu en prenant $T = \{\psi\}$ et utilisant l'équivalence de $\psi \vdash \theta$ et de $\vdash \psi \to \theta$:
Soient $\varphi(x_1, \ldots, x_n)$ une $\mathcal{L}$-formule, et ψ un $\mathcal{L}$-énoncé. Supposons que les symboles de constantes $c_1, \ldots, c_n$ n'apparaissent pas dans ψ. Si $\vdash \psi \to \varphi(c_1, \ldots, c_n)$, alors $\vdash \psi \to \forall x_1, \ldots, x_n\, \varphi$.
La démonstration se fait par induction sur n et est laissée **en exercice**.

Lemme 2.71. *(**AC**) Il existe une $\bar{\mathcal{L}}$-théorie $\bar{T}$, qui est cohérente, contient T, et admet C comme ensemble de témoins.*

Démonstration. Soit $\kappa = \text{card}(\mathcal{L}) + \aleph_0$, et soit $(c_\alpha)_{\alpha<\kappa}$ une liste (sans répétition) des éléments de C. Nous regardons l'ensemblede toutes les $\bar{\mathcal{L}}$-formules en au plus une variable libre, en prenons une liste φ_α, $\alpha < \kappa$, et réécrivons la formule φ_α de façon que la variable libre de φ_α (si elle existe) soit x_α. (C'est à dire, nous substituons la variable x_α pour la variable libre de φ_α.)

[En classe il y a eu des questions sur la cardinalité de cet ensemble, et il est clair qu'il faut faire des restrictions car sinon on risque d'avoir un ensemble trop grand, par exemple si on permet une quantité arbitrairement grande de variables. Une façon serait d'identifier des formules qui ne diffèrent que par les variables utilisées, et qui sont donc logiquement équivalentes : la formule φ, vue comme une liste de symboles, dans laquelle on remplace toute occurrence de la variable x par une autre variable qui n'apparaît pas dans l'écriture de φ ; alors on obtient une nouvelle formule qui lui est logiquement équivalente. Une autre façon, plus simple peut-être, est de fixer un ensemble de variables dénombrable. Comme une formule est une suite **finie** de symboles pris dans un ensemble de cardinalité $\kappa = \text{card}(\bar{\mathcal{L}}) = \sup\{\text{card}(\mathcal{L}), \aleph_0\}$, on aura donc κ formules, puisque pour tout entier $n > 0$, on a $\kappa^n = \kappa$. De plus, cet ensemble de cardinalité κ a une énumération indexée par les ordinaux $\alpha < \kappa$. Il faut se rappeler que tout ensemble de cardinalité κ est (par définition de la cardinalité) en bijection avec κ, c'est à dire avec l'ensemble des ordinaux $< \kappa$.]

Nous construisons par induction une suite croissante T_α, $\alpha < \kappa$, de $\bar{\mathcal{L}}$-théories, et une suite d_α, $\alpha < \kappa$ de constantes de C, satisfaisant les conditions suivantes :
– (i) Chaque T_α est cohérente (en tant que $\bar{\mathcal{L}}$-théorie) ;
– (ii) Si $\alpha = \beta + 1$, alors $T_\alpha = T_\beta \cup \{(\exists x_\beta\, \varphi_\beta) \to \varphi_\beta(d_\beta/x_\beta)\}$. (Si φ_β n'a pas de variable libre, cet énoncé ne dit donc rien du tout.)
– (iii) Si α est limite, alors $T_\alpha = \bigcup_{\beta<\alpha} T_\beta$.
Par le lemme précédent, nous savons que $T = T_0$ est $\bar{\mathcal{L}}$-cohérente. Supposons T_β construite et cohérente, et notons que comme $\beta < \kappa$, $\text{card}(T_\beta \setminus T) < \kappa$. Il existe donc un élément de

C qui n'apparaît pas dans T_β, et nous prenons pour d_β le premier tel élément de C (pour l'énumération c_α). Nous allons montrer que

$$T_{\beta+1} = T_\beta \cup \{(\exists x_\beta \varphi_\beta) \to \varphi_\beta(d_\beta/x_\beta)\}$$

est cohérente. Sinon, nous aurions

$$T_\beta \vdash \neg((\exists x_\beta \varphi_\beta) \to \varphi_\beta(d_\beta/x_\beta)),$$

i.e.,

$$T_\beta \vdash (\exists x_\beta \varphi_\beta) \wedge \neg(\varphi_\beta(d_\beta/x_\beta)),$$

et donc par le lemme 2.68, que

$$T_\beta \vdash (\exists x_\beta \, \varphi_\beta) \wedge (\forall x_\beta \, (\neg\varphi_\beta)),$$

ce qui nous donne la contradiction désirée, puisque T_β était supposée cohérente.

Si α est limite, alors $T_\alpha = \bigcup_{\beta<\alpha} T_\beta$ est cohérente. Posons maintenant $\bar{T} = \bigcup_{\alpha<\kappa} T_\alpha$. C'est une $\bar{\mathcal{L}}$-théorie, qui est cohérente, et admet (par construction) C comme ensemble de témoins.

Lemme 2.72. *(**AC**) Soit T une $\mathcal{L}$-théorie cohérente ayant C comme ensemble de témoins. Alors T a un modèle, dont l'univers consiste d'interprétations d'éléments de C.*

Démonstration. Il faut noter que si $T' \supset T$ est une $\mathcal{L}$-théorie cohérente, alors C est aussi un ensemble de témoins pour T'. En effet les conditions pour admettre un ensemble de témoins portent sur un ensemble d'énoncés qui doivent être prouvables par T, et cet ensemble dépend seulement du langage et de C, pas de la théorie T. Donc si la $\mathcal{L}$-théorie T admet C comme ensemble de témoins, alors toute $\mathcal{L}$-théorie la contenant l'admettra aussi, puisqu'elle prouvera tous les énoncés nécessaires.

En utilisant le lemme de Zorn, nous supposerons que T est maximale cohérente. (L'ensemble des $\bar{\mathcal{L}}$-théories cohérentes contenant T est inductif, donc contient un élément maximal). Deux conséquences de cette hypothèses sont, pour un énoncé φ :

– Si $T \vdash \varphi$, alors $\varphi \in T$. En effet, comme T est cohérente et prouve φ, alors $T \cup \{\varphi\}$ est cohérente, et par maximalité $\varphi \in T$.

– Si $T \nvdash \varphi$, alors $\neg\varphi \in T$. En effet, si $T \nvdash \varphi$, alors $T \cup \{\neg\varphi\}$ est cohérente. Par maximalité de T, nous avons que $\neg\varphi \in T$.

Ainsi,

$$T \vdash \varphi \qquad \text{ssi} \qquad \varphi \in T \qquad \text{ssi} \qquad (\neg\varphi) \notin T.$$

En particulier, T est complète.

Pour deux constantes $c, d \in C$, nous posons $c \sim d$ si $c = d \in T$. Comme T est maximale cohérente, si $(c = d) \notin T$, alors $(c \neq d) \in T$, et donc il suit que $\sim$ définit une relation d'équivalence sur C. Par les axiomes de l'égalité, cette relation d'équivalence est compatible avec les fonctions et relations du langage. D'autre part, si $d \in \mathcal{L} \setminus C$ est un symbole de constante, alors nous avons $T \vdash \exists v \, v = d$; donc, puisque C est un ensemble de témoins, nous

aurons $T \vdash c = d$ pour un $c \in C$, et en fait $T \vdash c' = d$ pour tout c' dans la $\sim$-classe d'équivalence de c. De la même façon, si f est une fonction n-aire, alors $(f(c_1, \ldots, c_n) = c) \in T$ pour un (ou plusieurs) $c \in C$.

Nous pouvons donc définir une $\mathcal{L}$-structure sur $A = C/\sim$ de façon naturelle. Notons $\tilde{c}$ la classe d'équivalence de $c \in C$. Si R est un symbole de relation n-aire, et $(c_1, \ldots, c_n)$ un n-uplet de C, nous posons $R(\tilde{c_1}, \ldots, \tilde{c_n})$ si et seulement si $R(c_1, \ldots, c_n) \in T$. De même, si f est une foncion n-aire de $\mathcal{L}$ et $(c_1, \ldots, c_n, c)$ un $(n+1)$-uplet de C tel que $(f(c_1, \ldots, c_n) = c) \in T$, nous posons $f(\tilde{c_1}, \ldots, \tilde{c_n}) = \tilde{c}$. Et enfin si d est un symnbole de constante de $\mathcal{L} \setminus C$, alors il existe $c \in C$ tel que $(c = d) \in T$, et nous interprétons la constante d par $\tilde{c}$. Les axiomes de l'égalité nous garantissent que cette définition de $\mathcal{L}$-structure sur A ne dépend pas du choix des éléments de C dans les $\sim$-classes d'équivalence.

Il nous faut maintenant montrer que A, avec cette $\mathcal{L}$-structure, est bien un modèle de T. Pour cela nous allons montrer que si φ est un $\mathcal{L}$-énoncé, alors

$$(*) \qquad A \models \varphi \qquad \text{ssi} \qquad \varphi \in T.$$

C'est montré par induction sur la compléxité du $\mathcal{L}$-énoncé φ.
φ atomique. Soient $t_1, \ldots, t_n$ des termes sans variables libres. (Ce sont donc des termes dans lesquels les éventuelles variables libres ont été remplacées par des constantes du langage). La façon dont la $\mathcal{L}$-structure de A a été définie garantit que $A \models t_1 = t_2$ ssi $(t_1 = t_2) \in T$. De même, si R est un symbole de relation n-aire, alors $A \models R(t_1, \ldots, t_n)$ ssi $R(t_1, \ldots, t_n) \in T$. $(*)$ est donc vrai pour les formules atomiques.

Si $(*)$ est vrai pour φ_1 et φ_2 alors il est certainement vrai pour $\neg\varphi_1$ et pour $\varphi_1 \wedge \varphi_2$.

Supposons maintenant que $\varphi = \exists x \psi$. Si $A \models \varphi$, alors il existe $\tilde{c} \in A$ tel que $A \models \psi(\tilde{c}/x)$. Par hypothèse d'induction, cela veut dire que $\psi(c/x) \in T$, et donc $\exists x \psi \in T$ (par maximalité de T, et en utilisant l'axiome existentiel et MP). Ceci montre une direction.

Si $\exists x \psi \in T$, alors, puisque C est un ensemble de témoins pour T, nous savons qu'il existe $c \in C$ tel que $T \vdash \exists x \psi \to \psi(c/x)$, et donc $\psi(c/x) \in T$ par (MP). D'où $A \models \psi(\tilde{c})$ et $A \models \varphi$. Cela termine la preuve.

Théorème 2.73. *(Théorème de complétude) (**AC**)). Soit T une $\mathcal{L}$-théorie. Alors T est cohérente ssi T a un modèle.*

Démonstration. Les deux (ou trois) résultats précédents montrent la direction difficile : si T est cohérente, alors T a un modèle. D'autre part, si T a un modèle, alors T n'est pas contradictoire, car sinon elle prouverait l'énoncé $\exists x\, x \neq x$ qui est faux dans toutes les $\mathcal{L}$-structures, et cela contredit 2.58.

2.74. Une formulation équivalent est la suivante : *pour tout énoncé φ,*

$$T \models \varphi \qquad \text{ssi} \qquad T \vdash \varphi.$$

En effet, $T \models \varphi$ ssi aucun modèle de T ne satisfait $\neg\varphi$, ssi $T \cup \{\neg\varphi\}$ est contradictoire, ssi $T \vdash \varphi$.

Remarques 2.75. Notons qu'une des conséquences de la preuve du théorème de complétude, est qu'une théorie cohérente aura un modèle de cardinalité au plus $\mathrm{card}(\mathcal{L}) + \aleph_0$.
Si le langage $\mathcal{L}$ est dénombrable, on peut se passer de l'axiome du choix, en modifiant la construction.

Théorème 2.76. *(Compacité) Une $\mathcal{L}$-théorie T a un modèle ssi tous ses sous-ensembles finis ont un modèle.*

2.77. Le théorème de compacité est un des outils fondamentaux de la logique. Il est constamment utilisé, mais attention – on fait facilement des erreurs. Voici deux exemples d'applications.

Théorème 2.78. *Une théorie qui a des modèles finis arbitrairement grands, a un modèle infini.*

Démonstration. On ajoute au langage $\mathcal{L}$ des nouveaux symboles de constantes $\{c_n \mid n \in \mathbb{N}\}$, et on considère la théorie $\bar{T} = T \cup \{c_n \neq c_m \mid n \neq m\}$ dans le langage $\bar{\mathcal{L}} = \mathcal{L} \cup \{c_n \mid n \in \mathbb{N}\}$. Alors tout sous-ensemble fini T_0 de $\bar{T}$ a un modèle : si n est tel que aucun c_m, $m \geq n$, n'apparaît dans les énoncés de T_0, alors par hypothèse, T_0 a un modèle M avec au moins n éléments. Ce modèle a une expansion $\bar{M}$ à une $\bar{\mathcal{L}}$-structure dans laquelles les interprétations des constantes $c_0, \ldots, c_{n-1}$ sont distinctes. Alors $\bar{M} \models T_0$.

Tous les sous-ensembles finis de $\bar{T}$ ont donc des modèles, ce qui entraine que $\bar{T}$ a un modèle. Ce modèle sera infini, et son réduit à $\mathcal{L}$ est un modèle de T.

Remarque 2.79. On peut se demander quelles sont les cardinalités possibles pour des modèles de T. La preuve du résultat précédent peut être facilement modifiée pour montrer que si T a des modèles finis arbitrairement grands, alors elle en a de cardinalité $\geq \kappa$ pour tout cardinal κ.

Définition 2.80. Un *groupe ordonné* est un groupe muni d'un ordre total $<$ qui est compatible avec la loi de groupe, i.e. qui vérifie

$$\forall x, y, z \, \big(x < y \to (xz < yz) \wedge (zx < zy)\big).$$

Un groupe est dit *ordonnable* s'il peut être muni d'un ordre qui en fait un groupe ordonné. [Je devrais vraiment écrire la multiplication avec $\cdot$, ici j'utilise la notation usuelle, qui l'omet.]

2.81. Si G est un groupe ordonné alors on a $x < y$ ssi $x^{-1}y > e$ (on multiplie à gauche par x^{-1}), et $x > e$ ssi $x^{-1} < e$, où e est l'identité du groupe. Notons aussi que G est sans torsion : si $g^n = e$ avec $n \in \mathbb{N}^{>0}$, $g \in G$, alors $g = e$. Cela suit du fait que l'un de g ou g^{-1} est $\geq e$, disons g ; et alors on aura $e \leq g \leq g^2 \leq \cdots \leq g^n = e$.

Définition 2.82. Un groupe est *de type fini* s'il est engendré par un nombre fini d'éléments.

2.83. Les diagrammes. Soient $\mathcal{L}$ un langage, M une $\mathcal{L}$-structure. On forme alors le langage $\mathcal{L}(M)$ en ajoutant à $\mathcal{L}$ un nouveau symbole de constante pour chaque élément de M. Par abus de notation je noterai de la même façon : l'élément $m \in M$, la constante qui lui est associée, et éventuellement l'interprétation de cette constante dans M.
Le *diagramme atomique de M*, noté $\mathrm{Diag}^+(M)$ ou $\mathrm{Diag}^{at}(M)$, est l'ensemble des $\mathcal{L}(M)$-énoncés

atomiques vrais dans M.
Le *diagramme sans quantificateurs de* M, noté ... Diagsansquant(M), ou bien ici tout simplement Diag(M), est l'ensemble des $\mathcal{L}(M)$-énoncés sans quantificateurs vrais dans M.
Et enfin, le *diagramme élémentaire de* M, noté DiagElem(M), est l'ensemble de tous les $\mathcal{L}(M)$-énoncés vrais dans M.

Exemple 2.84. Soit G un groupe, vu comme $\mathcal{L}$-structure, $\mathcal{L} = \{\cdot, ^{-1}, e\}$. Alors $\mathrm{Diag}^+(G)$ consistera de tous les $\mathcal{L}(G)$-énoncés de la forme $g_1 \cdot g_2 = g_3$ où g_1g_2 est égal à g_3 dans G. $\mathrm{Diag}(G)$ consistera des énoncés de $\mathrm{Diag}^+(G)$, des $\mathcal{L}(G)$-énoncés de la forme $g_1 \cdot g_2 \neq g_3$ où g_1g_2 n'est pas égal à g_3 dans G, et enfin des conjonctions et disjonctions finies de ces énoncés.

Remarque 2.85. Important. Soit N une $\mathcal{L}(M)$-structure. On peut alors définir une application $f : M \to N$ de façon naturelle, en envoyant un élément m de M sur l'interprétation m^N dans N de la constante associée. On peut voir f comme une application entre les $\mathcal{L}(M)$-structures M et N, ou bien entre les $\mathcal{L}$-structures M et le réduit de N à $\mathcal{L}$. On a alors que
(1) f est un morphisme si et seulement si $N \models \mathrm{Diag}^+(M)$. (Cf. 2.30)
(2) f est un plongement si et seulement si $N \models$ Diagsansquant(M).

La preuve a plus ou moins été faite en TD. Les détails manquants sont en **exercice**.

Le théorème de compacité s'applique aussi pour montrer

Théorème 2.86. *Soit G un groupe. Alors G est ordonnable si et seulement si tout sous-groupe de type fini de G est ordonnable.*

Démonstration. La preuve du résultat s'appuie sur la remarque suivante, et sur la *méthode des diagrammes* :
Un sous-groupe d'un groupe ordonné est ordonnable : on prend tout simplement l'ordre induit.
On considère le langage des groupes $\mathcal{L} = \{\cdot, ^{-1}, e\}$, et on pose $\mathcal{L}' = \mathcal{L} \cup \{<\}$. Soit T_0 la $\mathcal{L}'$-théorie des groupes ordonnés. Elle est donc axiomatisée par les axiomes de groupe $(\forall x, y, z \Big(\big((x \cdot y) \cdot z = x \cdot (y \cdot z) \big) \wedge (x \cdot x^{-1} = x^{-1} \cdot x = e) \wedge (x \cdot e = e \cdot x = x) \Big))$, par ceux d'ordre total $(\forall x, y, z \big((x < y) \wedge y < z) \to x < z \big) \wedge \big(\big(x < y) \vee (x = y) \vee (x > y) \big) \wedge \neg (x < x))$, et enfin par celui disant que l'ordre est compatible avec les opérations de groupe et donné ci-dessus. Nous prenons maintenant T_1 la $\mathcal{L}'(G)$-théorie obtenue en adjoignant à T_0 le diagramme sans quantificateurs de la $\mathcal{L}$-structure G.

Il nous suffit donc de montrer que la théorie T_1 a un modèle : si H est un modèle de T_1, alors H est ordonné, et contient une copie du groupe G (cf. 2.83). Cette copie de G est ordonnable et donc G aussi.

Pour montrer que T_1 a un modèle, il suffit de montrer (par le théorème de compacité 2.76) que tout fragment[4] fini de T_1 a un modèle.

Mais un fragment fini T_2 de T_1 ne mentionne qu'un nombre fini d'éléments de G, disons, $g_1, \ldots, g_m$. Par hypothèse, le sous-groupe de G engendré par $g_1, \ldots, g_m$ est ordonnable, i.e., a une $\mathcal{L}'$-structure qui en fait un groupe ordonné. Son expansion naturelle à $\mathcal{L}' \cup \{g_1, \ldots, g_m\}$ sera donc un modèle de T_2.

[4]fragment = morceau, sous-ensemble.

Définition 2.87. On dit qu'une formule $\varphi(x_1, \ldots, x_n)$ est *sous forme prénexe*, si elle s'écrit $Q_1 y_1 Q_2 y_2 \ldots Q_m y_m \, \psi(x_1, \ldots, x_n, y_1, \ldots, y_m)$, avec $Q_i \in \{\forall, \exists\}$, les y_i des variables distinctes des variables (libres) $x_1, \ldots, x_n$, et la formule ψ une formule sans quantificateurs. Le nom des variables n'a aucune importance, j'impose seulement qu'elles soient toutes distinctes.

L'intérêt des formules sous forme prénexe, est que les variables qui y apparaissent sont soit liées, soit libres, mais pas les deux. C'est parce que les quantificateurs sont au début. On montre alors facilement le résultat suivant :

Proposition 2.88. *Toute formule $\varphi(x_1, \ldots, x_n)$ est logiquement équivalente à une formule $\theta(x_1, \ldots, x_m)$ sous forme prénexe. (C'est à dire : $\vdash \varphi(x_1, \ldots, x_n) \leftrightarrow \psi(x_1, \ldots, x_n)$).*

Démonstration. On le montre par induction sur la compléxité des formules. Pour les formules sans quantificateurs il n'y a rien à montrer. Si $\psi(x_1, \ldots, x_n)$ est sous forme prénexe, alors aussi $\neg\psi(x_1, \ldots, x_n)$ et $\exists x_i \psi(x_1, \ldots, x_n)$. En effet, pour la première assertion, $\neg Q_1 y_1 Q_2 y_2 \ldots Q_m y_m$ $\psi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ est (presque) $\tilde{Q}_1 y_1 \tilde{Q}_2 y_2 \ldots \tilde{Q}_m y_m \, \neg\psi(x_1, \ldots, x_n, y_1, \ldots, y_m)$, avec $\tilde{Q}_i = \forall$ si $Q_i = \exists$, et $\tilde{Q}_i = \exists$ si $Q_i = \forall$. La deuxième est immédiate. L'ensemble des $\mathcal{L}$-formules logiquement équivalentes à une formule sous forme prénexe est donc clos par négation et par quantification. Il reste à montrer qu'il est clos par conjonction.
On suppose $\varphi_1(x_1, \ldots, x_n)$ et $\varphi_2(x_1, \ldots, x_n)$ sous forme prénexe. Soient $v_1, \ldots, v_r$ les variables liées apparaissant dans φ_2, et choisissons de nouvelles variables $w_1, \ldots, w_r$ qui n'apparaissent ni dans φ_1 ni dans φ_2, ni parmi $\{x_1, \ldots, x_n\}$. On vérifie, par induction sur r, que la formule φ_2' obtenue à partir de φ_2 en remplaçant chaque occurrence de v_i par w_i pour $i = 1, \ldots, r$, est logiquement équivalente à φ_2 (et est aussi sous forme prénexe). On peut donc supposer que les variables liées de φ_2 n'apparaissent pas dans φ_1 ni dans $\{x_1, \ldots, x_n\}$. Ecrivons $\varphi_2(x_1, \ldots, x_n) = Q_1 w_1 \ldots Q_r w_r \, \psi(x_1, \ldots, x_n, w_1, \ldots, w_r)$. Alors $\varphi_1(x_1, \ldots, x_n) \wedge \varphi_2(x_1, \ldots, x_n)$ est logiquement équivalente à $Q_1 w_1 \ldots Q_r w_r \, (\varphi_1(x_1, \ldots x_n) \wedge \psi(x_1, \ldots, x_n, w_1, \ldots, w_r))$.

Exemple 2.89. Considérons des formules de la forme $\exists x \psi_1(x)$ et $\exists x \psi_2(x)$ (les ψ_i étant sans quantificateurs). Alors la formule $(\exists x \psi_1(x)) \wedge (\exists x \psi_2(x))$ n'est pas équivalente à $\exists x \, (\psi_1(x) \wedge \psi_2(x))$, elle est bien plus faible. Cependant, elle est équivalente à $\exists x \exists y \, (\psi_1(x) \wedge \psi_2(y))$.

2.90. (NON FAIT EN CLASSE, A ÉTÉ MENTIONNÉ EN TD). Voici une troisième application du théorème de compacité. Soient $\mathcal{L}$ un langage, T une ($\mathcal{L}$)-théorie cohérente. On note $S(T)$ l'ensemble de toutes les $\mathcal{L}$-théories complètes et déductivement closes qui contiennent T. [Convention/rappel : une théorie T' est complète si elle est cohérente et pour tout énoncé φ, $T' \nvdash \varphi$ alors $T' \vdash (\neg\varphi)$. Une théorie T' est close par déduction si pour tout énoncé φ, si $T' \vdash \varphi$ alors $\varphi \in T'$] $S(T)$ est appellé l'*espace de Lindenbaum* de T. On met dessus une topologie, dont une base d'ouverts est donnée par les ensembles

$$\langle \varphi \rangle = \{T' \in S(T) \mid \varphi \in T'\},$$

où φ parcourt l'ensembles des énoncés.

Théorème 2.91. *L'espace topologique $S(T)$ est Hausdorff, totalement discontinu (i.e., la topologie a une base d'ouverts consistant d'ouverts-fermés) et compact.*

Démonstration. Soient $T_1 \neq T_2$ deux éléments de $S(T)$. Alors il existe un énoncé φ tel que $\varphi \in T_1$ et $\varphi \notin T_2$. Comme T_2 est complète et déductivement close, on a $\neg\varphi \in T_2$, autrement dit : $T_1 \in \langle\varphi\rangle$, $T_2 \in \langle\neg\varphi\rangle$. Comme les ouverts $\langle\varphi\rangle$ et $\langle\neg\varphi\rangle$ sont disjoints, cela montre que $S_1(T)$ est Hausdorff. Cela montre aussi que tous les ouverts de notre base sont des ouverts-fermés, puisque leur complémentaire est aussi ouvert. Il reste à montrer que si U_i, $i \in I$, est une famille d'ouverts de $S(T)$ telle que $\bigcup_{i\in I} U_i = S(T)$, alors il existe un sous-ensemble fini J de I tel que $\bigcup_{i\in J} U_i = S(T)$.

On remarque d'abord que la famille d'ouverts $\langle\varphi\rangle$ est bien close par intersection finie, car $\langle\varphi_1 \wedge \varphi_2\rangle = \langle\varphi_1\rangle \cap \langle\varphi_2\rangle$. Tout ouvert s'écrit comme une union d'ouverts de base, et on peut donc supposer que les U_i sont de la forme $\langle\varphi_i\rangle$, puisque chaque U_i sera une union d'ensembles de cette forme.

Notre hypothèse entraine alors que tout modèle de T satisfait l'un des φ_i. En effet, si $M \models T$, alors $\mathrm{Th}(M) \in S(T)$, et donc il existe $i \in I$ tel que $\mathrm{Th}(M) \in \langle\varphi_i\rangle$; autrement dit, il existe $i \in I$ tel que $M \models \varphi_i$.

Prenant la contraposée, aucun modèle de T n'est un modèle de $\{\neg\varphi_i \mid i \in I\}$, et la théorie $T \cup \{\neg\varphi_i \mid i \in I\}$ n'a pas de modèles. Par compacité, il existe $J \subset I$ fini, tel que $T \cup \{\neg\varphi_i \mid i \in J\}$ n'a pas de modèle. Redéroulant l'argument, cela montre que $\bigcup_{i\in J}\langle\varphi_i\rangle = S(T)$.

3 Un peu plus de théorie des modèles

Soit $\mathcal{L}$ un langage.

Définition 3.1. (Rappel) Soient M et N deux $\mathcal{L}$-structures. Alors M et N sont *élémentairement équivalents*, noté $M \equiv N$ ssi elles satisfont les même énoncés (de $\mathcal{L}$). $\equiv$ est une relation d'équivalence sur les $\mathcal{L}$-structures.

Supposons $M \subseteq N$. Alors M est une *sous-structure élémentaire de* N, ou bien N est une *extension élémentaire de* M, noté $M \prec N$, si pour toute $\mathcal{L}$-formule $\varphi(\bar{x})$ et tout uplet $\bar{a}$ dans M, on a

$$M \models \varphi(\bar{a}) \qquad \text{ssi} \qquad N \models \varphi(\bar{a}).$$

Proposition 3.2. *Soient M, N, U des $\mathcal{L}$-structures.*

(1) *Si $M \prec N$ alors $M \equiv N$.*
(2) *$M \prec M$.*
(3) *Si $M \prec N$ et $N \prec U$, alors $M \prec U$.*
(4) *Si $M \prec U$, $N \prec U$, et $M \subset N$ alors $M \prec N$.*

Démonstration. Exercice.

Exemples 3.3. Voici trois exemples montrant qu'on peut avoir $M \equiv N$, $M \subseteq N$, et $M \not\prec N$.
(1) Soit $\mathcal{L}$ le langage $\{S\}$ où S est une fonction unaire. On munit $\mathbb{N}$ d'une $\mathcal{L}$-structure en

définissant $S(x) = x + 1$ sur $\mathbb{N}$. C'est notre modèle N. Nous regardons maintenant sa sous-structure M, d'univers $\mathbb{N}^{>0}$. Alors $M \simeq N$, et donc $M \equiv N$. Cependant $M \not\prec N$. En effet, nous avons

$$M \models \forall x \neg(Sx = 1)$$

tandis que $N \models S0 = 1$ et donc en particulier $N \models \exists x\, Sx = 1$.

(2) On considére les structures $M = (\mathbb{N}^{>0}, <) \subset N = (\mathbb{N}, <)$. Alors $M \simeq N$ et $1 \in M$, mais $M \models \forall x \neg(x < 1)$ alors que $N \models \exists x\, (x < 1)$.

(3) Soit $\mathcal{L} = \{+, -, 0\}$, et prenons $N = \mathbb{Z}$ avec sa $\mathcal{L}$-structure habituelle. Alors $M = 2\mathbb{Z}$ est une sous-structure de N, qui lui est isomorphe. Donc on a $M \subset N$, et $M \equiv N$. Cependant, $M \models \forall x\, (x + x \neq 2)$, alors que $N \models \exists x\, (x + x = 2)$.

Exemples 3.4. Voici maintenant deux vrais exemples, ils seront montrés bientôt.

(1) $(\mathbb{Q}, <) \prec (\mathbb{R}, <)$ (l'ordre étant l'ordre habituel)

(2) $(\bar{\mathbb{Q}}, +, -, \cdot, 0, 1) \prec (\mathbb{C}, +, -, \cdot, 0, 1)$.

Ici, $\bar{\mathbb{Q}}$ dénote la clôture algébrique de $\mathbb{Q}$ dans $\mathbb{C}$. Tout ses éléments satisfont un polynôme (non nul) à coefficients dans $\mathbb{Q}$, et c'est le plus petit corps algébriquement clos contenant $\mathbb{Q}$ (et contenu dans $\mathbb{C}$).

Définition 3.5. Soient M et N des $\mathcal{L}$-structures.

(1) On appelle *diagramme élémentaire de M*, noté $\mathrm{DiagElem}(M)$, la théorie de la $\mathcal{L}(M)$-structure M.

(2) Une application $F : M \to N$ est un *plongement élémentaire* si c'est un plongement, et $F(M) \prec N$. Autrement dit, pour toute $\mathcal{L}$-formule $\varphi(\bar{x})$ et uplet $\bar{a}$ dans M,

$$M \models \varphi(\bar{a}) \qquad \text{ssi} \qquad N \models \varphi(F(\bar{a})).$$

Exercice 3.6. Exercice sur les diagrammes. Soient M et N des $\mathcal{L}$-structures, et enrichissons N en une $\mathcal{L}(M)$-structure en définissant pour chaque $a \in M$ une interprétation a^N de la constante associée à l'élément a de M. Nous mettons sur M la $\mathcal{L}(M)$-structure naturelle, et nous obtenons donc une application $F : M \to N$, définie par $a \mapsto a^N$ pour chaque $a \in M$. Je rappelle aussi que :
$\mathrm{Diag}^+(M)$ est l'ensemble de tous les $\mathcal{L}(M)$-énoncés atomiques vrais dans M ;
$\mathrm{Diag}(M)$ est l'ensemble de tous les $\mathcal{L}(M)$-énoncés sans quantificateurs vrais dans M. (Cet exercice a déjà été proposé – cf 2.85).

(1) Montrez que F est un homomorphisme (de $\mathcal{L}$-structures) si et seulement si la $\mathcal{L}(M)$-structure N est un modèle de $\mathrm{Diag}^+(M)$.

(2) Montrez que F est un plongement si et seulement si $N \models \mathrm{Diag}(M)$.

(3) Montrez que F est un plongement élémentaire si et seulement si $N \models \mathrm{DiagElem}(M)$.

Théorème 3.7. *(Théorème de Löwenheim-Skolem-Tarski ascendant) Soient M une $\mathcal{L}$-structure infinie, et κ un cardinal. Alors M a une extension élémentaire de cardinalité $\geq \kappa$.*

Démonstration. Cela n'a d'intérêt que si $\text{card}(M) < \kappa$. Dans ce cas on ajoute au langage $\mathcal{L}(M)$ un ensemble C de nouvelles constantes, avec $\text{card}(C) = \kappa$. On considère la théorie

$$\text{DiagElem}(M) \cup \{c \neq d \mid c, d \in C, c \neq d\}.$$

Tout fragment fini de cette théorie a un modèle, d'univers M. Elle a donc un modèle N. On aura $\text{card}(N) \geq \text{card}(C) = \kappa$. Et aussi $M \prec N$ par l'exercice 3.6.
Notons que si $\text{card}(\mathcal{L}) \leq \kappa$, alors par la preuve de la construction d'un modèle d'une théorie cohérente, ce modèle N peut être choisi de cardinalité exactement κ.

Proposition 3.8. *(Le test ou critère de Tarski-Vaught). Soient $M \subseteq N$ deux $\mathcal{L}$-structures. Supposons que pour toute $\mathcal{L}$-formule $\varphi(x, \bar{y})$ et uplet $\bar{a}$ de M, si*

$$N \models \exists x\, \varphi(x, \bar{a})$$

alors il existe b dans M tel que

$$N \models \varphi(b, \bar{a}).$$

Alors $M \prec N$.

Démonstration. Notez qu'on n'exige pas que le b trouvé dans M satisfasse la formule dans M, seulement dans N.

On montre par induction sur la compléxité des formules, que si $\varphi(x_1, \ldots, x_n)$ est une $\mathcal{L}$-formule alors

$$(*) : \text{pour tout } a_1, \ldots, a_n \in M,\ M \models \varphi(a_1, \ldots, a_n) \text{ ssi } N \models \varphi(a_1, \ldots, a_n).$$

$(*)$ est vrai pour les formules atomiques (cela utilise seulement le fait que $M \subseteq N$). Si φ et ψ satisfont $(*)$ alors aussi $\varphi \wedge \psi$ et $\neg\varphi$. Nous supposons maintenant que $\varphi(\bar{x}, y)$ satisfait $(*)$, et voulons montrer que $\exists y\, \varphi(\bar{x}, y)$ satisfait $(*)$.

Soit $\bar{a} \in M^n$, et supposons $M \models \exists y\, \varphi(\bar{a}, y)$. Alors il existe $b \in M$ tel que $M \models \varphi(\bar{a}, b)$ et par hypothèse d'induction, on a $N \models \varphi(\bar{a}, b)$, i.e., $N \models \exists y\, \varphi(\bar{a}, y)$.

Supposons maintenant que $N \models \exists y\, \varphi(\bar{a}, y)$. Par hypothèse (b), il existe $b \in M$ tel que $N \models \varphi(\bar{a}, b)$, et par hypothèse d'induction, on a $M \models \varphi(\bar{a}, b)$. Cela implique $M \models \exists y\, \varphi(\bar{a}, y)$ et finit la preuve.

Exercice 3.9. (1) Soit $(M_n)_{n\in\omega}$ une suite de $\mathcal{L}$-structures satisfaisant $M_n \prec M_{n+1}$ pour tout $n \in \omega$. Montrez que pour tout n, on a $M_n \prec \bigcup_{n\in\omega} M_n$.
(2) Généralisez le résultat à un ordinal κ quelconque : si $(M_\beta)_{\beta<\kappa}$ est une suite de $\mathcal{L}$-structures satisfaisant $M_\beta \prec M_\gamma$ pour tout $\beta < \gamma < \kappa$, montrez que $M_\beta \prec \bigcup_{\beta\in\kappa} M_\beta$ pour tout $\gamma < \kappa$.
On peut énoncer cet exercice de façon un peu différente et apparemment plus forte. Soit (M_α), $\alpha < \kappa$ une chaîne de $\mathcal{L}$-structures satisfaisant pour tout α : $M_\alpha \prec M_{\alpha+1}$; et si α est un ordinal limite, alors $M_\alpha = \bigcup_{\beta<\alpha} M_\beta$. Montrez que $M = \bigcup_{\alpha<\kappa} M_\alpha$ est une extension élémentaire de chaque M_α.

Théorème 3.10. *(Löwenheim-Skolem descendant). Soient M une $\mathcal{L}$-structure et $A \subseteq M$. On suppose* $\mathrm{card}(M) \geq \mathrm{card}(\mathcal{L}) + \aleph_0$. *Alors M a une sous-structure élémentaire N contenant A et de cardinalité* $\leq \mathrm{card}(A) + \mathrm{card}(\mathcal{L}) + \aleph_0 = \kappa$.

Démonstration. On construit une chaîne croissante $(A_n)_{n\in\omega}$ de sous-structures de M de cardinalité $\leq \kappa$ et qui satisfait la condition suivante, pour tout $n \in \mathbb{N}$:
Si $\varphi(x, \bar{y})$ est une formule, $\bar{a}$ un uplet de A_n et $M \models \exists x\, \varphi(x, \bar{a})$, alors il existe $b \in A_{n+1}$ tel que $M \models \varphi(b, \bar{a})$.

On posera ensuite $N = \bigcup_{n\in\omega} A_n$. C'est donc une $\mathcal{L}$-sous-structure de M. Le test de Tarski-Vaught 3.8 montre $N \prec M$: si $\varphi(x, \bar{y})$ est une $\mathcal{L}$-formule, $\bar{a}$ un uplet dans N, et $M \models \exists x\, \varphi(x, \bar{a})$, alors $\bar{a}$ est dans un A_n, et donc il existera $b \in A_{n+1}$ tel que $M \models \varphi(b, \bar{a})$. De plus, $\mathrm{card}(N) \leq \sup \mathrm{card}(A_n) + \aleph_0 \leq \kappa$.

Nous construisons la suite de la façon suivante. On prend pour A_0 la sous-structure de M engendrée par A ; elle est bien de cardinalité $\leq \kappa$, car A_0 consiste exactement des éléments de M de la forme $t(\bar{a})$, où t est un terme du langage, et $\bar{a}$ est un uplet d'éléments de A de la bonne longueur. On a donc $\mathrm{card}(A_0) \leq \mathrm{card}(\mathcal{L} \cup A) + \aleph_0 \leq \kappa$.

Supposons A_n construit, de cardinalité $\leq \kappa$. Soit $\Sigma(A_n)$ l'ensemble de toutes les formules $\varphi(x)$ de $\mathcal{L}(A_n)$ ayant exactement la variable libre x et telles que $M \models \exists x\, \varphi(x)$. Pour chaque $\varphi(x) \in \Sigma(A_n)$ on choisit $b_\varphi \in M$ tel que $M \models \varphi(b_\varphi)$ et on prend pour A_{n+1} la sous-structure de M engendrée par $A_n \cup \{b_\varphi \mid \varphi \in \Sigma(A_n)\}$. Comme $\mathrm{card}(\Sigma(A_n)) \leq \kappa$ on a bien $\mathrm{card}(A_{n+1}) \leq \kappa$.

Définition 3.11. Une théorie cohérente T *élimine les quantificateurs* si pour toute formule $\varphi(\bar{v})$, $\bar{v}$ un uplet de variables, il existe une formule $\psi(\bar{v})$ telle que $T \models \forall \bar{v}\, (\varphi(\bar{v}) \leftrightarrow \psi(\bar{v}))$.
Voici une terminologie utile : si T est une théorie, $\varphi(\bar{v}), \psi(\bar{v})$ des formules, et si $T \models \forall \bar{v}(\varphi(\bar{v}) \leftrightarrow \psi(\bar{v}))$, on dira aussi φ *et* ψ *sont équivalentes modulo* T.

Remarque 3.12. Quid d'un énoncé φ ?
Vous avez remarqué que dans ma définition, les formules φ et ψ ont les mêmes variables libres. Donc si φ n'a pas de variables libres, alors aussi ψ devrait ne pas en avoir. Mais il est possible que le langage ne contienne pas de constantes, et donc aucun énoncé sans quantificateurs.

Discussion du cas où le langage ne contient pas de symbole de constante. Il y a deux façons de remédier au problème d'absence dénoncé sans quantificateur. La première est de ne rien dire sur les variables libres de la formule ψ ; et la seconde est de dire que $\bot$ (toujours faux) et $\top$ (toujours vrai) sont des énoncés sans quantificateurs du langage.
J'opterai pour la deuxième solution : **considérer $\bot$ et $\top$ comme des énoncés sans quantificateurs**. Mais cela entraine que les seules théories (dans un langage sans constantes) qui éliminent les quantificateurs, sont les théories complètes. Une troisième possibilité serait de rajouter au langage un symbole de constante : après tout, comme nos structures sont toujours supposées non vides, ça ne change pas grand chose à la théorie T ni à ses modèles.

Remarque 3.13. Soient $\bar{a}$ et $\bar{b}$ des uplets dans des $\mathcal{L}$-structures M et N respectivement. Que veut dire la phrase

($*$) a et $\bar{b}$ satisfont les mêmes formules sans quantificateurs (dans M et N respectivement). Tout d'abord ils ont la même longueur, disons $\bar{a} = (a_1, \ldots, a_n)$ et $\bar{b} = (b_1, \ldots, b_n)$. On sait que si $t_1, \ldots, t_m$ sont des termes du langage (en $\bar{x} = (x_1, \ldots, x_n)$) et R est un symbole de relation m-aire du langage, ou $=$, alors

$$M \models R(t_1(\bar{a}), \ldots, t_m(\bar{a})) \qquad \text{ssi} \qquad N \models R(t_1(\bar{b}), \ldots, t_m(\bar{b})).$$

Si $\langle \bar{a} \rangle$ dénote la sous-structure de M engendrée par $\bar{a}$, ses éléments sont exactement les (interprétations des) termes du langage évalués en $\bar{a}$. La bijection qui envoie a_i sur b_i pour $1 \leq i \leq n$ s'étend donc de façon unique en un isomorphisme de $\mathcal{L}$-structures entre la sous-structure $\langle \bar{a} \rangle$ de M et la sous-structure $\langle \bar{b} \rangle$ de N. On a donc l'équivalence entre ($*$) et
L'application $\bar{a} \to \bar{b}$ s'étend en un isomorphisme entre la sous-structure $\langle \bar{a} \rangle$ de M engendrée par $\bar{a}$ et la sous-structure $\langle \bar{b} \rangle$ de N engendrée par $\bar{b}$.

Proposition 3.14. *Soient T une théorie qui élimine les quantificateurs, et M, N deux modèles de T.*

(1) *Soit A une $\mathcal{L}$-structure (non vide), qui est une sous-structure de M et de N. Alors les $\mathcal{L}(A)$-structures M et N sont élémentairement équivalentes.*
(2) *Si M et N contiennent des sous-structures isomorphes, alors $M \equiv N$.*
(3) *Suppposons $M \subseteq N$. Alors $M \prec N$.*

Démonstration. Ces trois remarques sont tout à fait évidentes, une fois qu'on décortique un peu les définitions. Notez l'équivalence : Si $\varphi(\bar{x})$ est une $\mathcal{L}$-formule et $\bar{a}$ un uplet de A, alors on peut aussi voir $\varphi(\bar{a})$ comme un $\mathcal{L}(A)$-énoncé.

(1) Le fait que A soit une sous-structure commune de M et de N dit que M et N satisfont les mêmes $\mathcal{L}(A)$-énoncés sans quantificateurs. Comme ils sont modèles de la théorie T qui élimine les quantificateurs, ils satisfont donc les mêmes $\mathcal{L}(A)$-énoncés, ce qui veut dire que les $\mathcal{L}(A)$-structures M et N sont élémentairement équivalentes.

(2) Suit immédiatement de (1), en notant que l'ensemble des $\mathcal{L}(A)$-énoncés contient l'ensemble des $\mathcal{L}$-énoncés.

(3) On prend $A = M$ et on applique (1).

Lemme 3.15. *Soit T une théorie telle que pour toute formule $\varphi(x, \bar{y})$ sans quantificateurs (x une variable, $\bar{y}$ un uplet de variables), il existe une formule $\psi(\bar{y})$ sans quantificateurs telle que*

$$T \models \forall \bar{y} \big((\exists x \, \varphi(x, \bar{y})) \leftrightarrow \psi(\bar{y}) \big).$$

Alors T élimine les quantificateurs.

Démonstration. On regarde l'ensemble S des $\mathcal{L}$-formules qui sont équivalentes modulo T à une formule sans quantificateurs. Il contient certainement les formules sans quantificateurs, et si l'on montre qu'il est clos par conjonction, négation et quantification existentielle, cela montrera qu'il coïncide avec l'ensemble de toutes les $\mathcal{L}$-formules. L'ensemble S est clairement clos par

conjonction et négation. Il suffit donc de montrer qu'il est clos par quantification existentielle : si $\varphi(x, \bar{y}) \in S$, alors $\varphi(x, \bar{y})$ est équivalente modulo T à une formule sans quantificateurs $\theta(x, \bar{y})$; donc $\exists x\, \varphi(x, \bar{y})$ est équivalente modulo T à $\exists x\, \theta(x, \bar{y})$, qui par hypothèse, est équivalente modulo T à une formule sans quantificateurs.

Corollaire 3.16. *Les conditions suivantes sont équivalentes, pour une théorie cohérente T, et une formule $\varphi(x, \bar{y})$ sans quantificateurs, avec $|\bar{y}| = n$, x une seule variable :*

(1) *$\exists x\, \varphi(x, \bar{y})$ est équivalente modulo T à une formule sans quantificateurs.*
(2) *Pour tous modèles M et N de T, et n-uplets $\bar{a}$ de M et $\bar{b}$ de N, si $\bar{a}$ et $\bar{b}$ satisfont les mêmes formules sans quantificateurs et s'il existe $c \in M$ tel que $M \models \varphi(c, \bar{a})$, alors il existe $d \in M$ tel que $N \models \varphi(d, \bar{b})$.*
(3) *Pour tous modèles M et N de T contenant un n-uplet $\bar{a}$ (dans une sous-structure commune), s'il existe $c \in M$ tel que $M \models \varphi(c, \bar{a})$, alors il existe $d \in N$ tel que $N \models \varphi(d, \bar{a})$.*

Démonstration. Il est clair que (1) implique (2) qui implique (3).
(3) implique (2). (La seule difficulté est de montrer que nous pouvons supposer que $\bar{b} = \bar{a}$.) Supposons donc que nous ayions des modèles M et N de T, des n-uplets $\bar{a} \in M$ et $\bar{b} \in N$ qui satisfont les mêmes formules sans quantificateurs. Alors l'application f qui envoie $\bar{a}$ sur $\bar{b}$ s'étend (de façon unique) à un isomorphisme $\tilde{f} : \langle \bar{a} \rangle \to \langle \bar{b} \rangle$ entre les sous-structures de M et N engendrées par $\bar{a}$ et $\bar{b}$ respectivement. On étend ensuite $\tilde{f}$ à une bijection de domaine M (en envoyant $M \setminus \langle \bar{a} \rangle$ sur n'importe quel ensemble disjoint de $\langle \bar{b} \rangle$), puis on met sur $\tilde{f}(M)$ l'unique $\mathcal{L}$-structure telle que $\tilde{f}$ soit un isomorphisme. On applique ensuite (3) à $(\tilde{f}(M), \bar{b})$ et à $(N, \bar{b})$, puis on revient via $\tilde{f}^{-1}$.

(2) implique (1). Soient $\bar{c}$ un n-uplet de nouvelles constantes, et posons $\mathcal{L}' = \mathcal{L} \cup \{\bar{c}\}$. On suppose $T \cup \{\exists x\, \varphi(x, \bar{c})\}$ cohérente. (Sinon on prend $\psi = \bot$). Nous allons montrer que $\exists x\, \varphi(x, \bar{c})$ est équivalente modulo T à un $\mathcal{L}'$-énoncé sans quantificateurs. Un tel énoncé s'écrit $\psi(\bar{c})$ avec ψ une $\mathcal{L}$-formule sans quantificateurs, et comme les éléments de $\bar{c}$ ne sont pas dans $\mathcal{L}$, le fait que $T \vdash \psi(\bar{c}) \leftrightarrow \exists x\, \varphi(\bar{c})$ entraine $T \vdash \forall \bar{y}[\psi(\bar{y}) \leftrightarrow (\exists x\, \varphi(x, \bar{y}))]$.

Soit Δ l'ensemble des $\mathcal{L}'$-énoncés sans quantificateurs, et posons

$$\Gamma = \{\psi \in \Delta \mid T \cup \{\exists x\, \varphi(x, \bar{c})\} \vdash \psi\}.$$

Nous allons montrer que $T \cup \Gamma \vdash \exists x\, \varphi(x, \bar{c})$. Par compacité, il existera alors $\psi \in \Gamma$ tel que $T \cup \{\psi\} \vdash \exists x\, \varphi(x, \bar{c})$, et ce sera le $\mathcal{L}'$-énoncé désiré. [Ici j'utilise que Γ est clos par conjonction finie].

Il faut donc montrer que tout modèle de $T \cup \Gamma$ satisfait $\exists x\, \varphi(x, \bar{c})$. Soit $N \models T \cup \Gamma$, et posons

$$\Sigma = \{\psi \in \Delta \mid N \models \psi\}.$$

Assertion. $T \cup \Sigma \cup \{\exists x\, \varphi(x, \bar{c})\}$ est cohérente.
Sinon, il existerait $\psi_1, \ldots, \psi_m \in \Sigma$ telles que

$$T \cup \{\exists x\, \varphi(x, \bar{c})\} \vdash \neg(\psi_1 \wedge \ldots \wedge \psi_m).$$

Mais alors $\neg(\psi_1 \wedge \ldots \wedge \psi_m) \in \Gamma$, ce qui contredit le fait que $N \models \Gamma$, et $N \models \psi_1 \wedge \ldots \wedge \psi_m$. Cela montre l'assertion.

Soit M un modèle de $T \cup \Sigma \cup \{\exists x\, \varphi(x, \bar{c})\}$. Alors M et N satisfont les mêmes $\mathcal{L}'$-énoncés sans quantificateurs (par définition de Σ). Donc par hypothèse de (2), puisque $M \models \exists x\, \varphi(x, \bar{c})$, aussi $N \models \exists x\, \varphi(x, \bar{c})$.

Voici un exercice facile, qui généralise ce résultat à d'autres ensembles de formules ou énoncés :

Exercice 3.17. Soient T_1 et T_2 des théories du langage $\mathcal{L}$, et Δ un ensemble d'énoncés de $\mathcal{L}$ qui est clos par disjonction finie. Supposons que $T_1 \cup T_2$ est cohérente. Les conditions suivantes sont équivalentes :

(1) Il existe $\Gamma \subseteq \Delta$ tel que $T_1 \cup \Gamma \models T_2$ et $T_1 \cup T_2 \models \Gamma$.
(2) Pour tous modèles M et N de T_1, si $M \models T_2$ et N satisfait tous les énoncés de Δ satisfaits par M, alors $N \models T_2$.

Pour (2) implique (1), on suit la preuve de 3.16, avec les changements suivants : $\Gamma = \{\psi \in \Delta \mid T_1 \cup T_2 \models \psi\}$; $\Sigma = \{\neg\psi \mid \psi \in \Delta, N \models \neg\psi\}$.

3.1 Les corps algébriquement clos

Dans cette section, nous étudierons la théorie des corps algébriquement clos, dans le langage des anneaux $\{+, -, \cdot, 0, 1\}$. Rappelons d'abord l'axiomatisation de la théorie ACF des corps algébriquement clos (parfois notée CAC en français) :
– Les axiomes de corps ;
– Pour chaque entier $n > 1$, un axiome disant que tout polynôme unitaire de degré n a une solution : $\forall x_1, \ldots, x_n \exists y\, y^n + x_1 y^{n-1} + \cdots + x_n = 0$.

Théorème 3.18. *ACF élimine les quantificateurs dans le langage des anneaux.*

Démonstration. Soient M et N deux corps algébriquement clos, et $\bar{a} \in M$, $\bar{b} \in N$ deux uplets satisfaisant les mêmes formules sans quantificateurs. Si A dénote le sous-corps de M engendré par $\bar{a}$ et B le sous-corps de N engendré par $\bar{b}$, il existe donc un isomorphisme f entre A et B qui envoie $\bar{a}$ sur $\bar{b}$. (On utilise le fait que les sous-structures $\langle\bar{a}\rangle$ et $\langle\bar{b}\rangle$ sont isomorphes, c'est à dire, que les sous-anneaux de M et N engendrés par $\bar{a}$ et $\bar{b}$ sont isomorphes). Prenons $c \in M$, et $\varphi(x, \bar{y})$ une formule telle que $M \models \varphi(c, \bar{a})$. Nous allons utiliser le critère donné par 3.16. Il y a deux cas à considérer.

Cas 1. c est algébrique sur A. (C'est à dire : il existe $P(X) \in A[X]$, $P \neq 0$, tel que $P(c) = 0$) Soit $P(X)$ le polynôme minimal (i.e., polynôme de degré minimal tel que $P(c) = 0$; et donc P est irréductible) unitaire de c sur A, et soit $P^f(X)$ le polynome obtenu en appliquant f aux coefficients de P. Comme f est un isomorphisme entre A et B, nous savons que P^f est unitaire,

de même degré que P, et irréductible sur B. Puisque N est algébriquement clos, il contient une racine d de $P^f(X) = 0$. Alors f s'étend en un isomrphisme qui envoie c sur d. En effet, on a

$$A(c) \simeq A[X]/(P(X)) \qquad \text{et} \qquad B(d) \simeq B[X]/(P^f(X)).$$

$(P(X))$ dénote l'ensemble des multiples de $P(X)$; c'est un idéal de $A[X]$, c'est-à-dire clos par $+$ et par multiplication par des éléments de $A[X]$. Nous aurons donc $N \models \varphi(d, \bar{b})$.

Cas 2. c est *transcendant* sur A (c'est à dire, c ne satisfait aucune équation polynomiale non triviale à coefficients dans A).

Si N contient un élément d transcendant aur B, alors l'isomorphisme f se prolonge en un isomorphisme entre $A(c)$ et $B(d)$ qui envoie c sur d, et on raisonne comme dans le cas précédent. Si tous les éléments de N sont algébriques sur B, alors prenons N^* une extension élémentaire propre de N. Tout élément $d \in N^* \setminus N$ sera transcendant sur N, et donc sur B. Alors nous aurons que $N^* \models \varphi(d, \bar{b})$, donc $N^* \models \exists x\, \varphi(x, \bar{b})$, et $N \models \exists x\, \varphi(x, \bar{b})$ (parce que $N \prec N^*$), d'où il existe $d' \in N$ tel que $N \models \varphi(d', \bar{b})$, ce qui termine la preuve.

3.19. Soit p un nombre premier. On dénote par ACF_p la théorie $\mathrm{ACF} \cup \{\text{“}p = 0\text{”}\}$, et par ACF_0 la théorie $\mathrm{ACF} \cup \{\text{“}p \neq 0\text{”} \mid p \text{ nombre premier}\}$.

Corollaire 3.20. *Les théories ACF_p (pour p premier) et ACF_0 sont complètes. De plus, tout corps algébriquement clos est modèle d'une de ces théories (suivant sa caractéristique) : les théories ACF_0 et ACF_p sont appelées les* complétions *de T.*

Démonstration. Tout corps algébriquement clos est modèle d'une de ces théories, et deux corps élémentairement équivalents ont certainement la même caractéristique.

Soient M et N des corps algébriquement clos. S'ils ont la même caractéristique p (p un premier ou 0), alors ils sont modèles de ACF_p. Ils contiennent des sous-structures isomorphes : le corps premier, qui est $\mathbb{Q}$ si $p = 0$, et $\mathbb{F}_p$ si $p > 0$. Par élimination des quantificateurs de la théorie ACF, il suit que $M \equiv N$. (Cf. 3.14).

3.21. Vous avez probablement tous entendu parler du principe de Lefschetz, qui dit qu'une propriété qui est vraie pour $\mathbb{C}$ est vraie dans presque tous les corps algébriquement clos. Il est clair que la notion de “propriété” demande à être précisée. Mais maintenant vous pouvez le faire, et ce principe devient un corollaire du théorème précédent.

Corollaire 3.22. *(Le principe de Lefschetz) Soit φ un énoncé du langage des anneaux. Si $\mathbb{C} \models \varphi$, alors il existe N tel que φ est vrai dans tout corps algébriquement clos de caractéristique $> N$. (Ce qui équivaut à : φ est vrai dans la clôture algébrique $\bar{\mathbb{F}}_p$ de $\mathbb{F}_p$ pour $p > N$).*

Démonstration. Puisque $\mathbb{C} \models \varphi$, et ACF_0 est complète, nous savons que $\mathrm{ACF}_0 \models \varphi$. Un fragment fini Γ de ACF_0 prouve donc φ, et ce fragment ne contient qu'un nombre fini d'énoncés de la forme “$p \neq 0$” : on choisit N plus grand que tous les p apparaissant dans Γ.

Remarques 3.23. (1) Notons aussi la contraposée : si φ est vraie dans une infinité de $\bar{\mathbb{F}}_p$, alors elle est vraie dans tous corps algébriquement clos de caractéristique 0.
(2) Le fait que $\bar{\mathbb{F}}_p$ soit une union de corps finis, est très utile. Il existe des énoncés qui sont trivialement vérifiés pour $\bar{\mathbb{F}}_p$, et grâce au fait que ACF_p est complète, le deviennent pour tout corps algébriquement clos de caractéristique p. Un exemple est le fameux résultat d'Ax, qui sera montré en TD:
(Ax) *Soit $f : \mathbb{C}^n \to \mathbb{C}^n$ une fonction injective, donnée par des polynômes. Alors f est une bijection.*

Ces deux exemples n'ont pas été faits en classe.

Exemple 3.24. Considérons la structure $(\mathbb{N}, S)$, où S est la fonction successeur, et T sa théorie. Alors T n'élimine pas les quantificateurs. En effet, on voit que 1 et 2 satisfont les mêmes formules sans quantificateurs, car les sous-structures qu'ils engendrent sont isomorphes : $\langle 1 \rangle = \mathbb{N}^{\geq 1}$, $\langle 2 \rangle = \mathbb{N}^{\geq 2}$. Cependant, nous avons $\mathbb{N} \models \exists x\, SSx = 2$, et $\mathbb{N} \models \forall x\, SSx \neq 1$.

Exemple 3.25. Considérons $\mathbb{R}$ dans le langage des anneaux $\{+, -, \cdot, 0, 1\}$, et soit T sa théorie. Alors T n'élimine pas les quantificateurs. En effet, on voit que $\sqrt{2}$ et $-\sqrt{2}$ satisfont les mêmes formules sans quantificateurs, puisque $\langle\sqrt{2}\rangle = \mathbb{Z}[\sqrt{2}]$ et $\langle-\sqrt{2}\rangle = \mathbb{Z}[\sqrt{2}]$ sont isomorphes par un isomorphisme qui envoie $\sqrt{2}$ sur $-\sqrt{2}$. Cependant

$$\mathbb{R} \models (\exists x\, x^2 = \sqrt{2}) \wedge (\forall x\, x^2 \neq -\sqrt{2})$$

On peut montrer qu'il suffit de rajouter l'ordre au langage pour obtenir que la théorie de $\mathbb{R}$ élimine les quantificateurs. Vous en connaissez déjà une instance : l'équation $aX^2 + bX + c$ a une solution dans $\mathbb{R}$ ssi $b^2 - 4ac$ est ≥ 0.

3.2 Ultrafiltres, ultraproduits

Définition 3.26. Soit I un ensemble.

(1) Un *filtre* sur I est un sous-ensemble $\mathcal{F}$ de $\mathcal{P}(I)$ (l'ensemble des parties de I) satisfaisant :
   (i) $\emptyset \notin \mathcal{F}$; $I \in \mathcal{F}$;
   (ii) Si $U \in \mathcal{F}$ et $U \subseteq V \subseteq I$ alors $V \in \mathcal{F}$;
   (iii) Si $U, V \in \mathcal{F}$ alors $U \cap V \in \mathcal{F}$.

(2) Un *ultrafiltre* sur I est un filtre sur I qui est maximal (pour l'inclusion).

(3) Un filtre $\mathcal{F}$ sur I est *principal* s'il existe $U \in \mathcal{F}$ tel que $\mathcal{F} = \{V \subseteq I \mid V \supseteq U\}$.

(4) Si I est infini, le *filtre de Fréchet* est l'ensemble des parties *cofinies* de I (c'est à dire, des $V \subseteq I$ tels que $I \setminus V$ soit fini).

Remarques 3.27. (1) En utilisant l'axiome du choix (ou plutôt sa conséquence le Lemme de Zorn), tout filtre est contenu dans un ultrafiltre.
(2) Un ultrafiltre principal est forcément de la forme $\{V \subseteq I \mid i \in V\}$ pour un élément $i \in I$.
(3) Si I est fini, alors tout ultrafiltre sur I est principal, car il a un plus petit élément

Exemple 3.28. Soit X un espace topologique, et $x \in X$. On appelle voisinage de x un sous-ensemble de X contenant un ouvert qui contient x. Alors l'ensemble des voisinages de x forme un filtre de parties de X.

Définition 3.29. Soit $S \subset \mathcal{P}(I)$. On dit que S *a la propriété de l'intersection finie* si pour tous n et $U_1, \ldots, U_n \in S$, $U_1 \cap \cdots \cap U_n \neq \emptyset$.

Lemme 3.30. *Soit $S \subset \mathcal{P}(I)$ ayant la propriété de l'intersection finie. Alors l'ensemble $\mathcal{F}$ consistant des sous-ensembles U de I tels qu'il existe $n \in \mathbb{N}$ et $U_1, \ldots, U_n \in S$ avec $U_1 \cap \cdots \cap U_n \subseteq U$, est un filtre. (On l'appelle le* filtre engendré par S*).*

Démonstration. Immédiat.

Lemme 3.31. *Soit $\mathcal{F}$ un filtre sur I. Alors $\mathcal{F}$ est un ultrafiltre si et seulement si pour tout sous-ensemble U de I, nous avons $U \in \mathcal{F}$ ssi[5] $I \setminus U \notin \mathcal{F}$.*

[5] ssi est une abbréviation pour si et seulement si

Démonstration. Supposons d'abord $U \in \mathcal{F}$ ssi $I \setminus U \notin \mathcal{F}$. Alors aussi $U \notin \mathcal{F}$ ssi $I \setminus U \in \mathcal{F}$. Notons que aucun filtre ne peut contenir à la fois U et $I \setminus U$: s'il les contient tous les deux, il contient $U \cap (I \setminus U) = \emptyset$ (par (iii)), mais cela contredit (i). Ces deux remarques montrent que $\mathcal{F}$ est maximal : s'il ne contient pas U, il contient $I \setminus U$, et aucun filtre le contenant ne contiendra U.

Supposons maintenant que $\mathcal{F}$ soit un ultrafiltre, et soit $U \subseteq I$ tel que $I \setminus U \notin \mathcal{F}$. Nous allons montrer que $U \in \mathcal{F}$. Soit $\mathcal{G} = \{V \cap U \mid V \in \mathcal{F}\}$. J'affirme que $\emptyset \notin \mathcal{G}$. Sinon, il existerait $V \in \mathcal{F}$ tel que $U \cap V = \emptyset$, c'est à dire, $V \subset I \setminus U$, et nous aurions donc que $I \setminus U \in \mathcal{F}$, contrairement à notre hypothèse. La famille $\mathcal{G}$ satisfait donc la propriété de l'intersection finie ; par 3.30(1), $\mathcal{G}$ engendre un filtre $\mathcal{F}'$, et ce filtre $\mathcal{F}'$ contient $\mathcal{F} \cup \{U\}$. Par maximalité de $\mathcal{F}$, nous avons $\mathcal{F}' = \mathcal{F}$, c'est à dire, $U \in \mathcal{F}$. On en déduit facilement l'implication $\rightarrow$.

Lemme 3.32. *Soit I un ensemble infini, et $\mathcal{F}$ un ultrafiltre sur I. Alors $\mathcal{F}$ est non-principal si et seulement s'il contient le filtre de Fréchet sur I.*

Démonstration. Un ultrafiltre est non-principal si et seulement s'il ne contient aucun singleton (par 3.27), si et seulement s'il contient les complémentaires des singletons (par 3.31), si et seulement s'il contient le filtre de Fréchet (par (iii) de la définition).

Définition 3.33. Soit $\mathcal{F}$ un filtre sur I, et $U \in \mathcal{F}$. On définit

$$\mathcal{F}|_U = \{V \cap U \mid V \in \mathcal{F}\}.$$

Exercice 3.34. Vérifiez que $\mathcal{F}|_U$ est un filtre sur U.

Définition 3.35. Soit $(A_i)_{i \in I}$, une famille de $\mathcal{L}$-structures. On munit le produit cartésien $\prod_{i \in I} A_i$ d'une $\mathcal{L}$-structure de la façon suivante :
Si $c \in \mathcal{L}$ est un symbole de constante, l'interprétation de c dans le produit est l'élément $(c_i)_i$, où chaque c_i est l'interprétation de c dans A_i.
Si R est une relation n-aire du langage, et $a^1 = (a^1_i)_i, \ldots, a^n = (a^n_i)_i$ sont des éléments du produit, alors

$$R(a^1, \ldots, a^n) \qquad \text{ssi} \qquad A_i \models R(a^1_i, \ldots, a^n_i) \text{ pour tout } i \in I.$$

Si f est une fonction n-aire du langage, et $a^1, \ldots, a^n$ sont comme ci-dessus, on définit

$$f(a^1, \ldots, a^n) = (f(a^1_i, \ldots, a^n_i))_i.$$

Exercice 3.36. Soient I, A_i comme ci-dessus, et $\varphi(\bar{x})$ une formule atomique, $\bar{a} = (\bar{a}_i)_i$ un uplet de $\prod A_i$.
(1) Alors

$$\prod_{i \in I} A_i \models \varphi(\bar{a}) \qquad \text{ssi} \qquad A_i \models \varphi(\bar{a}_i) \text{ pour tout } i.$$

(2) Pour chaque $j \in I$, la projection $\prod_{i \in I} A_i \rightarrow A_j$ est un homomorphisme.
(3) Soient A et B des $\mathcal{L}$-structures, et $A \times B$ leur produit cartésien, avec sa $\mathcal{L}$-structure comme

ci-dessus (On prend $|I| = 2$). Donnez-un exemple de uplets $c = (a_1, b_1)$ et $d = (a_2, b_2)$, et de formule $\varphi(x, y)$ tels que

$$A \times B \models \varphi(c, d), \qquad A \models \neg\varphi(a_1, b_1).$$

Définition 3.37. Soient I un ensemble, $\mathcal{F}$ un filtre sur I, et $(A_i)_{i\in I}$ une famille de $\mathcal{L}$-structures. Nous définissons le *produit réduit des A_i sur le filtre $\mathcal{F}$*, noté $\prod_{i\in I} A_i/\mathcal{F}$, comme étant la $\mathcal{L}$-structure suivante :

L'univers de $\prod_{i\in I} A_i/\mathcal{F}$ est le quotient de $\prod_{i\in I} A_i$ par la relation d'équivalence $\sim_{\mathcal{F}}$ définie par

$$(a_i) \sim_{\mathcal{F}} (b_i) \qquad \text{ssi} \qquad \{i \in I \mid a_i = b_i\} \in \mathcal{F}.$$

Notons que

$$\{i \in I \mid a_i = b_i\} \cap \{i \in I \mid b_i = c_i\} \subseteq \{i \in I \mid a_i = c_i\},$$

ce qui montre la transitivité de $\sim_{\mathcal{F}}$. La réfléxivité et la symétrie de $\sim_{\mathcal{F}}$ sont évidentes. Nous notons $(a_i)_{\mathcal{F}}$ la classe d'équivalence de (a_i) pour la relation $\sim_{\mathcal{F}}$.

Pour c un symbole de constante, l'interprétation de c dans $\prod_{i\in I} A_i/\mathcal{F}$ est tout simplement $(c_i)_{\mathcal{F}}$, où chaque c_i est l'interprétation de c dans A_i.

Si f est un symbole de fonction n-aire du langage, et $(a_i^1), \ldots, (a_i^n) \in \prod_{i\in I} A_i$, nous posons

$$f((a_i^1)_{\mathcal{F}}, \ldots, (a_i^n)_{\mathcal{F}}) = (f(a_i^1, \ldots, a_i^n))_{\mathcal{F}}.$$

Notons que si $(b_i^j)_{\mathcal{F}} = (a_i^j)_{\mathcal{F}}$ pour $j = 1, \ldots, n$, alors $\{i \in I \mid b_i^j = a_i^j \text{ pour } j = 1, \ldots, n\}$ est dans $\mathcal{F}$, ce qui montre que f est bien définie.

Si R est un symbole de relation n-aire du langage, et $(a_i^1), \ldots, (a_i^n) \in \prod_{i\in I} A_i$, alors nous posons

$$\prod_{i\in I} A_i/\mathcal{F} \models R((a_i^1)\mathcal{F}, \ldots, (a_i^n)_{\mathcal{F}}) \qquad \text{ssi} \qquad \{i \in I \mid A_i \models R(a_i^1, \ldots, a_i^n)\} \in \mathcal{F}.$$

Définition 3.38. S'il existe une $\mathcal{L}$-structure M telle que tous les A_i sont égaux à M, alors on écrit $M^I/\mathcal{F}$ au lieu de $\prod_{i\in I} A_i/\mathcal{F}$, et on parle de *puissance réduite de M*. Elle est aussi parfois notée $M^{\mathcal{F}}$.
Si le filtre $\mathcal{F}$ est un ultrafiltre, alors $\prod_{i\in I} A_i/\mathcal{F}$ est appelé un *ultraproduit des structures A_i*, et $M^I/\mathcal{F}$ est appelée une *ultrapuissance de M*.

Remarques 3.39. Soit I un ensemble infini.

(1) Soient $\mathcal{F} \subseteq \mathcal{F}'$ des filtres sur I. Alors l'application $\prod_{i\in I} A_i/\mathcal{F} \to \prod_{i\in I} A_i/\mathcal{F}'$ définie par $(a_i)_{\mathcal{F}} \mapsto (a_i)_{\mathcal{F}'}$, est un homomorphisme de $\mathcal{L}$-structures.

(2) Notons aussi que le produit cartésien est un cas particulier de produit réduit : on prend le filtre $\{I\}$.

(3) Soient $\mathcal{F}$ un filtre sur I, et $U \in \mathcal{F}$. Alors l'application naturelle $\prod_{i\in I} A_i \to \prod_{i\in U} A_i$, $(a_i)_{i\in I} \mapsto (a_i)_{i\in U}$, induit un isomorphisme de $\mathcal{L}$-structures

$$\prod_{i\in I} A_i/\mathcal{F} \simeq \prod_{i\in U} A_i/\mathcal{F}|_U.$$

En particulier, si $\{i\} \in \mathcal{F}$, alors

$$\prod_{i\in I} A_i/\mathcal{F} \simeq A_i.$$

Théorème 3.40. *(*Théorème de Łos.*) Soient I un ensemble, $\mathcal{F}$ un ultrafiltre sur I, et $(M_i)_{i\in I}$ une famille de $\mathcal{L}$-structures, $M^* = \prod_{i\in I} M_i/\mathcal{F}$.*

(1) *Si $t(x_1, \ldots, x_n)$ est un terme du langage, et $(a_i^1), \ldots, (a_i^n) \in \prod_{i\in I} M_i$, alors*

$$t((a_i^1)_\mathcal{F}, \ldots, (a_i^n)_\mathcal{F}) = (t(a_i^1, \ldots, a_i^n))_\mathcal{F}.$$

(En fait, ceci est vrai aussi dans un produit réduit.)

(2) *Soient $\varphi(x_1, \ldots, x_n)$ une formule, et $(a_i^1), \ldots, (a_i^n) \in \prod_{i\in I} M_i$. Alors*

$$M^* \models \varphi((a_i^1)_\mathcal{F}, \ldots, (a_i^n)_\mathcal{F}) \qquad ssi \qquad \{i \in I \mid M_i \models \varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{F}.$$

(3) *Soit φ un énoncé. Alors*

$$M^* \models \varphi \qquad ssi \qquad \{i \in I \mid M_i \models \varphi\} \in \mathcal{F}.$$

Démonstration. (1) est clair par la définition de la $\mathcal{L}$-structure de M^*, et (3) est un cas particulier de (2). Il suffit donc de montrer (2). C'est fait par induction sur la complexité des formules, pour tous les uplets $\bar{a}$ de $\prod_{i\in I} M_i/\mathcal{F}$: nous allons d'abord montrer que c'est vrai pour les formules atomiques.

Les formules atomiques sont de la forme $R(t_1(\bar{x}), \ldots, t_n(\bar{x}))$, où $\bar{x}$ est un uplet de variables, chaque $t_i(\bar{x})$ est un terme du langage, et R est une symbole de relation n-aire, ou bien le graphe de l'égalité. L'équivalence suit de la définition de la $\mathcal{L}$-structure M^*. Remarquons qu'ici nous utilisons seulement le fait que $\mathcal{F}$ est un filtre.

Supposons que l'équivalence soit vraie pour les formules $\varphi(\bar{x})$ et $\psi(\bar{x})$. Elle est alors vraie pour la formule $\varphi(\bar{x}) \wedge \psi(\bar{x})$. De même, si elle est vraie pour la formule $\theta(\bar{x}, y)$, alors elle est vraie pour la formule $\exists y\, \theta(\bar{x}, y)$. Encore une fois, nous utilisons seulement le fait que $\mathcal{F}$ est un filtre.

Il reste maintenant à montrer que si le résultat est vrai pour $\varphi(\bar{x})$, alors il est vrai pour $\neg\varphi(\bar{x})$, et c'est là enfin que nous utiliserons le fait que $\mathcal{F}$ est un ultrafiltre. Nous avons par hypothèse d'induction :

$$M^* \models \neg\varphi((a_i^1)_\mathcal{F}, \ldots, (a_i^n)_\mathcal{F}) \qquad \text{ssi} \qquad \{i \in I \mid M_i \models \varphi(a_i^1, \ldots, a_i^n)\} \notin \mathcal{F}.$$

Mais, comme $\mathcal{F}$ est un ultrafiltre,

$$\{i \in I \mid M_i \models \varphi(a_i^1, \ldots, a_i^n)\} \notin \mathcal{F} \qquad \text{ssi} \qquad \{i \in I \mid M_i \models \neg\varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{F},$$

ce qui nous donne que la formule $\neg\varphi(\bar{x})$ satisfait aussi l'équivalence.

Remarque 3.41. Attention, si $\mathcal{F}$ est seulement un filtre, le résultat n'est plus vrai. On trouve facilement des exemples avec un produit cartésien $A \times B$.

Comme corollaire du Théorème de Łos 3.40, nous obtenons une version directe du théorème de compacité :

Théorème 3.42. *(Compacité) Soient T une théorie. Alors T a un modèle si et seulement si tous ses fragments finis ont un modèle.*

Démonstration. Soit I l'ensemble des parties finies de T ; par hypothèse, si $\Sigma \in I$, alors il existe une $\mathcal{L}$-structure M_Σ qui est modèle de Σ. Pour chaque énoncé φ, nous définissons

$$U_\varphi = \{\Sigma \in I \mid \Sigma \models \varphi\}.$$

Prenons $\mathcal{S} = \{U_\varphi \mid \varphi \in T\}$. J'affirme que $\mathcal{S}$ a la propriété de l'intersection finie : cela suit du fait que si $\varphi_1, \ldots, \varphi_n \in T$, alors $\bigcap_{i=1}^n U_{\varphi_i} = \{\Sigma \in I \mid \varphi_1, \ldots, \varphi_n \in \Sigma\}$ n'est pas vide. Soit $\mathcal{F}$ un ultrafiltre contenant $\mathcal{S}$ (qui existe par 3.30 et 3.27), et $M^* = \prod_{\Sigma \in I} M_\Sigma$. Soit $\varphi \in T$. Alors nous savons que

$$\{\Sigma \in I \mid M_\Sigma \models \varphi\} \supseteq U_\varphi,$$

et appartient donc à $\mathcal{F}$. Par le Théorème de Łos, nous avons donc $M^* \models \varphi$.

Corollaire 3.43. *Soit $\mathcal{F}$ un ultrafiltre non-principal sur l'ensemble I des nombres premiers. Alors*

$$\prod_{p \in I} \bar{\mathbb{F}}_p / \mathcal{F} \simeq \mathbb{C}.$$

Démonstration. Le théorème de Łos ainsi que le fait que la clôture déductive de la théorie ACF_0 soit complète, donne facilement que $\prod_{p \in I} \bar{\mathbb{F}}_p / \mathcal{F} \equiv \mathbb{C}$. Ensuite, il faut utiliser le fait que deux corps algébriquement clos de même caractéristique et de même cardinalité non dénombrable, sont isomorphes. Et enfin, il faut utiliser le fait que $\mathrm{card}(\prod_{p \in I} \bar{\mathbb{F}}_p / \mathcal{F}) = 2^{\aleph_0}$. Je ne prouverai pas ces deux assertions, en tout cas pas maintenant. La première vient de l'algèbre, et est facile si vous connaissez le concept de base de transcendance. La deuxième vient en fait d'un phénomène plus général, que je vais maintenant discuter.

Remarque 3.44. On peut se demander quelles sont les cardinalités possibles des ultrapuissances. Elles sont certainement bornées par la cardinalité du produit cartésien, que en principe on sait calculer. En fait on peut montrer que si les structures A_i sont finies, et I est dénombrable, $\mathcal{F}$ un ultrafiltre sur I alors ou bien $\prod_{i \in I} A_i / \mathcal{F}$ est fini, ou bien il est de cardinalité $2^{\aleph_0}$.

4 Fonctions primitives récursives et fonctions récursives

L'ensemble des fonctions totales de domaine $\mathbb{N}^p$ à valeurs dans $\mathbb{N}$ est noté $\mathcal{F}_p$. Une *fonction partielle* de $\mathbb{N}^p$ dans $\mathbb{N}$ est donnée par un sous-ensemble A de $\mathbb{N}^p$ et une fonction $f : A \to \mathbb{N}$. L'ensemble A est le *domaine* de la fonction f, noté $\mathrm{Dom}(f)$, et l'ensemble des fonctions partielles de $\mathbb{N}^p$ dns $\mathbb{N}$ est noté $\mathcal{F}_p^*$. On pose $\mathcal{F} = \bigcup_p \mathcal{F}_p$ et $\mathcal{F}^* = \bigcup_p \mathcal{F}_p^*$.

4.1. Motivation - 1. Les fonctions calculables. On essaie de définir ce qu'est une fonction *calculable*, calculable voulant dire calculable par une machine, qui ne fait que ce qu'on lui dit. On ne regarde que des fonctions d'une puissance cartésienne de $\mathbb{N}$ dans $\mathbb{N}$ (ou bien dans une puissance de $\mathbb{N}$ – ça n'a pas beaucoup d'importance, grâce à des bijections que nous définirons), certaines étant sans doute seulement partielles. Voici une définition informelle :
Une fonction $f \in \mathcal{F}_p^*$ est *semi-calculable* s'il existe un programme (un algorithme) qui, à partir d'une donnée $\bar{x} \in \mathbb{N}^p$, nous donne $f(\bar{x})$ si $\bar{x} \in \mathrm{dom}(f)$, et ou bien nous donne la valeur -1 (c'est à dire nous dit que $\bar{x} \notin \mathrm{dom}(f)$), ou bien ne s'arrête pas. On dit que f est calculable si elle est semi-calculable et s'il existe de plus un programme qui nous dit quand elle ne s'arrête pas pour la donnée $\bar{x}$. Evidemment, la question suivante est : *Qu'est-ce qu'un algorithme ou un programme ?*

4.2. Motivation - 2. Algorithmes ou machines. Une définition informelle d'un calcul par machine :

- On a un ensemble fini P d'instructions.
- Une machine L qui applique les instructions aux données.
- La machine peut garder en mémoire les données, les calculs intermédiaires, les étapes du calcul, etc.
- La machine L réagit étape par étape.
- Tous les calculs sont déterminés : pas de tirage au sort, pas de choix au hasard.
- Quand la machine s'arrête, elle donne un résultat. Mais elle peut aussi ne pas s'arrêter ...

4.3. Motivation - 3. Les thèses de Church.

Thèse de Church - version faible. *Il existe un ensemble explicite de fonctions semi-calculables, et un ensemble fini de règles qui permettent de construire de nouvelles fonctions semi-calculables en un nombre fini d'étapes.*

Plusieurs personnes ont proposé de tels ensembles de fonctions semi-calculables et de règles - il se trouve que l'on obtient toujours les mêmes fonctions (en tout cas pour des propositions raisonnables). Cela conduit à la thèse plus forte :

Thèse de Church. *Les fonctions semi-calculables sont les fonctions récursives. Les fonctions calculables sont les fonctions récursives dont le domaine est récursif.*

Dans ce chapitre, nous allons définir les fonctions récursives. Nous commençons par un sous-ensemble, plus simple, mais qui contient déjà beaucoup de fonctions.

4.1 Fonctions primitives récursives

Définition 4.4. L'ensemble E des fonctions *primitives récursives* est le plus petit sous-ensemble de $\mathcal{F}$ satisfaisant les conditions suivantes :

- Pour tout entier p, E contient les fonctions constantes $\mathbb{N}^p \to \mathbb{N}$.
- Pour tout entiers $1 \leq i \leq p$, E contient la projection $\pi_i^p : \mathbb{N}^p \to \mathbb{N}$, $(x_1, \ldots, x_p) \mapsto x_i$. Donc en particulier, E contient $id_{\mathbb{N}}$ $(= \pi_1^1)$.
- E contient la fonction successeur S, $x \mapsto x + 1$.
- E est clos par composition : si $f_1, \ldots, f_n \in E \cap \mathcal{F}_p$, et $g \in \mathcal{F}_n \cap E$, alors $g \circ (f_1, \ldots, f_n) \in E$.
- E est clos par récurrence : pour tout entier p, si $g \in E \cap \mathcal{F}_p$, $h \in E \cap \mathcal{F}_{p+2}$, alors la fonction $f \in \mathcal{F}_{p+1}$ définie par

$$f(\bar{x}, y) = \begin{cases} g(\bar{x}) & \text{si } y = 0, \\ h(\bar{x}, y-1, f(\bar{x}, y-1)) & \text{sinon.} \end{cases}$$

  est dans E.

Un sous-ensemble A de $\mathbb{N}^p$ est *primitif récursif* si sa fonction caractéristique, $\mathbf{1}_A$, est primitive récursive.

Grâce à ces opérations, on peut définir la plupart des fonctions usuelles, notamment l'exponentielle.

Exemple 4.5. (1) L'addition est définie par récurrence en utilisant les fonctions $g(x) = x$ et $h(x, y, z) = S(z)$.
(2) La multiplication est définie par récurrence en utilisant les fonctions $g(x) = 0$, $h(x, y, z) = z + x$.
(3) L'exponentielle x^y ...

Remarque 4.6. (1) Les fonctions suivantes sont primitives récursives :
(i) l'exponentielle $(x, y) \mapsto x^y$;
(ii) la fonction $x \mapsto x \dot{-} 1 := \sup\{0, x - 1\}$ (par récurrence, en prenant $h(y, z) = y$) ;
(iii) la fonction $(x, y) \mapsto x \dot{-} y := \sup\{0, x - y\}$;
(iv) si $A \subset \mathbb{N}^p$ est primitif récursif, et $f, g \in \mathcal{F}_p$ sont primitives récursives, alors aussi la fonction h qui vaut f sur A et g sur $\mathbb{N}^p \setminus A$: $f = \mathbf{1}_A f + (1 \dot{-} \mathbf{1}_A)g$; le résultat se généralise sans peine à une partition de $\mathbb{N}^p$ en n ensembles disjoints primitifs récursifs $A_1, \ldots, A_n$ et des fonctions $f_1, \ldots, f_n \in E$: on considère $\sum_{i=1}^n \mathbf{1}_{A_i} f_i$;
(v) la fonction $\mathrm{sg} := \mathbf{1}_{\mathbb{N}^*}$, $x \mapsto 1 \dot{-} (1 \dot{-} x)$;
(vi) la fonction $(x, y) \mapsto |x - y|$, on utilise sa définition par cas ;
(vii) Si $f \in \mathcal{F}_{p+1} \cap E$, alors $g_1(\bar{x}, y) = \sum_{t=0}^y f(\bar{x}, t)$ et $g_2(\bar{x}, y) = \prod_{t=0}^y f(\bar{x}, t)$ sont dans E.
(2) (i) La relation $<$ sur $\mathbb{N}$ (vue comme un sous-ensemble de $\mathbb{N}^2$) est primitive récursive.
(ii) L'ensemble des ensembles primitifs récursifs est clos par les opérations booléennes ($\cap$, $\cup$ et complément).

(iii) Si $f_1, \ldots, f_n \in E \cap \mathcal{F}_p$ et $A \subset \mathbb{N}^n$ est primitif récursif, alors $\{\bar{x} \in \mathbb{N}^p \mid (f_1(\bar{x}), \ldots, f_n(\bar{x})) \in A\}$ est primitif récursif.
(iv) Si $f \in E$, alors le graphe de f, Γ_f, est primitif récursif : $\mathbf{1}_{\Gamma_f}(\bar{x}, y) = 1 \dot{-} |f(\bar{x}) - y|$. Attention, la réciproque est fausse.

Démonstration. Exercice. Faites-le, cela voua aidera à apprendre les définitions, techniques, et propriétés.

4.7. Notons en particulier le
Schéma μ borné : Si $A \subseteq \mathbb{N}^{p+1}$ est primitif récursif, on définit la fonction

$$f(\bar{x}, y) = (\mu z \leq y)((\bar{x}, z) \in A)$$

par

$$f(\bar{x}, y) = \begin{cases} \text{le plus petit élément } z \leq y \text{ tel que } (\bar{x}, z) \text{ soit dans } A & \text{s'il existe,} \\ 0 & \text{s'il n'existe pas de tel } z. \end{cases}$$

Alors la fonction $(\mu z \leq y)((\bar{x}, z) \in A)$ est primitive récursive. Il faut noter que cette fonction est définissable à partir de f en utilisant les deux opérations ci-dessus (et les fonctions de base, celles déjà construites, ainsi que la définition par cas) :

$f(\bar{x}, 0) = 0$
$f(\bar{x}, y+1) = f(\bar{x}, y)$ si $\sum_{z=0}^{y} \mathbf{1}_A(\bar{x}, z) \geq 1$
$f(\bar{x}, y+1) = y+1$ si $\sum_{z=0}^{y} \mathbf{1}_A(\bar{x}, z) = 0 \wedge (\bar{x}, y+1) \in A$
$f(\bar{x}, y+1) = 0$ si $\sum_{z=0}^{y+1} \mathbf{1}_A(\bar{x}, z) = 0$.

4.8. Quantification bornée: Si $A \subset \mathbb{N}^{p+1}$ est primitif récursif, alors aussi

$$B_1 = \{(\bar{x}, y) \mid \exists t(t \leq y \wedge (\bar{x}, t) \in A)\},$$

et

$$B_2 = \{(\bar{x}, y) \mid \forall t(t \leq y \to (\bar{x}, t) \in A\}.$$

En effet, la fonction caractéristique de B_1 est $\mathrm{sg}(\sum_{t \leq z} \mathbf{1}_{\mathbf{A}}(\bar{x}, t))$. Quelle est celle de B_2 ?

4.9. Nombres premiers. L'ensemble P des nombres premiers est primitif récursif : il est défini par $x \in P$ si et seulement si $x \geq 2$ et $\forall u, v \leq x\, (uv = x \to u = 1 \vee v = 1)$ (cf 4.8). Il suit que la fonction π qui à $n = 0$ associe 2 et à $n > 0$ associe le $n+1$-ième nombre premier est aussi primitive récursive : on a $\pi(n+1) = (\mu z \leq \pi(n)! + 1)(\pi(n) < z \wedge z \in P)$. Bien que ce ne soit pas la forme exacte du schéma μ-borné, elle peut s'en déduire (voir aussi l'exercice ci-dessous 4.16). On a $\pi(n+1) = h(\pi(n))$, où $h(y) = h_1(y, y! + 1)$, et $h_1(y, z) = (\mu t \leq z)(y < t \wedge t \in P)$.

On peut aussi, plus simplement, utiliser le fait que $\pi(n) \sim (n+1)\log(n+1)$, ce qui entraine que pour n suffisamment grand, on a $(*) : \pi(n) < n^2$ (puisque $\lim \pi(n)\log(n+1)/(n+1) = 1$).

On a donc un entier A tel que si $n > A$ alors $(*)$ est vrai. On définit donc

$$\pi(n) = \begin{cases} 2 & \text{si } n = 0, \\ 3 & \text{si } n = 1, \\ 5 & \text{si } n = 2, \\ & \vdots \\ \pi(A) & \text{si } n = A, \\ (\mu z \leq n^2)(\pi(n-1) < z \wedge z \in P) & \text{si } n > A. \end{cases}$$

4.2 Fonctions de codage

Dans cette sous-section je donne deux exemples de fonctions donnent des bijections entre $\mathbb{N}^p$ et $\mathbb{N}$, et entre les suites finies d'entiers et $\mathbb{N}$. Elles nous permettent, à partir d'un entier n, de récupérer un p-uplet, ou bien une suite finie. Je mentionne aussi une troisième façon de récupérer des suites finies.

4.10. Première fonction de codage. La fonction

$$\alpha : \mathbb{N}^2 \to \mathbb{N}, \qquad (x, y) \mapsto (x + y)(x + y + 1)/2 + x,$$

définit une bijection entre $\mathbb{N}^2$ et $\mathbb{N}$. Elle est primitive récursive, et les deux fonctions β_1 et β_2 définies par $\alpha^{-1} = (\beta_1, \beta_2)$ le sont aussi, en utilisant le schéma μ-borné, car on a $\alpha(m, n) \geq m, n$: on a

$$\beta_1(x) = \mu(z \leq x)(\exists y \leq x\, \alpha(z, y) = x).$$

Cette fonction peut alors nous servir à définir des bijections (primitives récursives) entre $\mathbb{N}^p$ et $\mathbb{N}$, pour tout $p > 2$. Elles seront notées α_p pour $p > 2$, et les fonctions "inverses" correspondantes seront notées β_i^p pour $1 \leq i \leq p$ (et définies par $\alpha_p(\beta_1^p, \ldots, \beta_p^p) = id_{\mathbb{N}}$). On peut, par exemple, définir $\alpha_{p+1}(\bar{x}, y) = \alpha(\alpha_p(\bar{x}), y)$, si $\bar{x} \in \mathbb{N}^p$ et $y \in \mathbb{N}$. Mais ce n'est pas la seule façon. Ce qui nous intéresse, c'est que ces fonctions soient toutes primitives récursives.

4.11. Deuxième fonction de codage. Nous pouvons aussi définir une fonction, qui à une suite finie $(a_0, \ldots, a_n)$ de $\mathbb{N}$ associe un élément de $\mathbb{N}$ de la façon suivante :

$$\langle a_0, \ldots, a_n \rangle = 2^{a_0} \cdots \pi(n)^{a_n+1} - 1.$$

On définit $\langle \emptyset \rangle = 0$. On vérifie que cela définit bien une bijection de l'ensemble des suites finies d'entiers ≥ 0 dans $\mathbb{N}$. De plus, les opérations "inverses" suivantes sont primitives récursives :
$\langle a_0, \ldots, a_n \rangle \mapsto n$;
$(\langle a_0, \ldots, a_n \rangle, i) \mapsto a_i$ si $0 \leq i \leq n$, 0 sinon ;
pour chaque $n > 0$, la fonction $(a_1, \ldots, a_n) \mapsto \langle a_1, \ldots, a_n \rangle$;
$(\langle a_0, \ldots, a_n \rangle, i) \mapsto \langle a_0, \ldots a_i \rangle$ si $i \leq n$, $\langle a_0, \ldots, a_n \rangle$ sinon.

4.12. Décodage ? Comment fait-on pour trouver l'inverse de cette fonction ? Soit N un entier. Si $N = 0$, alors $N = \langle\emptyset\rangle$. Si $N \geq 1$, on écrit $N + 1 = \prod_{i=0}^{n} \pi(i)^{e_i}$, cette écriture est unique si on impose $e_n \neq 0$. Alors $N = \langle e_0, e_1, \ldots, e_n - 1\rangle$.

Remarquons que n est le plus grand $z \leq N + 1$ tel que $\pi(z)$ (le $z + 1$-ième nombre premier) soit dans P et $\pi(z)$ divise $N + 1$; que si $i = n$, alors a_n est l'entier $\leq N + 1$ tel que $\pi(n)^{a_n+1}$ divise $N + 1$ mais $\pi(n)^{a_n+2}$ ne le divise pas ; et si $i < n$, alors a_i est l'entier $\leq N + 1$ tel que $\pi(i)^{a_i}$ divise $N + 1$ et $\pi(i)^{a_i+1}$ ne le divise pas. A partir de ces remarques faciles, on prouve facilement les assertions ci-dessus, ainsi que celle-ci :
La fonction $\mathbb{N}^2 \to \mathbb{N}$, $(\langle a_0, \ldots, a_n\rangle, \langle b_0, \ldots, b_m\rangle) \mapsto \langle a_0, \ldots, a_n, b_0, \ldots, b_m\rangle$ est primitive récursive.

4.13. La fonction de Gödel. C'est une fonction $\mathbb{N}^3 \to \mathbb{N}$, définie par $\beta(n, k, i) = \mathrm{Rem}(n, (1 + i)k + 1)$, le reste de la division de n par $(1 + i)k + 1$. Cette fonction est primitive récursive. En effet, pour $y > 0$,

$$\mathrm{Rem}(x, y) = (\mu z \leq y)(\exists t \leq x\, yt = x - z).$$

Proposition 4.14. *Pour toute suite $a_0, \ldots, a_m$ d'entiers, il existe n et k tels que pour $i \leq m$, on ait $\beta(n, k, i) = a_i$.*

Démonstration. On choisit d'abord N tel que $N \geq m + 1$, et $1 + (i + 1)N! > a_i$ pour tout $i \leq m$. On pose $k = N!$. On remarque que si $0 \leq i < j \leq m$, alors les nombres $1 + (i + 1)N!$ et $1 + (j + 1)N!$ sont relativement premiers : en effet, soit d un diviseur premier de ces deux nombres, alors il divisera aussi $(j - i)N!$, et comme $(j - i) < N$, il sera $\leq N$ et donc divisera $N!$. Cela entraîne que d divise 1, ce qui est impossible.
Par le théorème du reste chinois, le système de congruences

$$x \equiv a_i \mod ((1 + i)N! + 1)$$

pour $i = 0, \ldots, m$, a une solution, n.

4.15. Codage ? Ce n'est pas à strictement parler une fonction de codage ou décodage, pour les raisons suivantes : (1) étant donnée une suite finie $(a_0, \ldots, a_m)$, il existe bien des entiers n et k tels que pour tout $i \leq m$ on ait $\beta(n, k, i) = a_i$, mais il en existe une infinité ; (2) la fonction $\beta(n, k, i)$ est définie pour tout i, et donc sera, pour i suffisamment grand, constante égale à n. On peut pallier à ces deux problèmes de la façon suivante : on associe à $(a_1, \ldots, a_m)$ la suite $(m, a_1, \ldots, a_m)$; on prend $N = \sup\{m + 1, a_1, \ldots, a_m\}$ et $k = N!$, puis $n < 1 + (m + 1)k$ tel que $n \equiv m \mod (1 + k)$ et $n \equiv a_i \mod (1 + (i + 1)k)$ for $1 \leq i \leq m$. Ces nombres k et n sont donc uniques pour une suite finie donnée.

Au "décodage", comme $\beta(n, k, 0)$ nous dit la longueur de la suite, on sait quand s'arrêter.

Exercice 4.16. Soient $f \in \mathcal{F}_{p+1}$, $g \in \mathcal{F}_p$ et $h \in \mathcal{F}_p$ des fonctions primitives récursives, et soit A un ensemble primitif récursif.

(1) Montrez que la fonction $\sum_{z=g(\bar{x})}^{z=h(\bar{x})} f(\bar{x}, y)$ est primitive récursive. Ici j'impose que la somme soit nulle si $g(\bar{x}) > h(\bar{x})$.

(2) Montrez que la fonction qui à n associe le n-ième élément de A, est primitive récursive, si vous savez que cette fonction est dominée par une fonction primitive récursive.

(3) Supposons maintenant qu'il existe une fonction $g \in \mathcal{F}_1 \cap E$, telle que pour tout $n \in \mathbb{N}$, si a_n dénote le n-ième élément de A, alors $a_{n+1} \leq g(a_n)$. Montrez que la fonction $n \mapsto a_n$ est dans E. (Cela s'applique donc à l'ensemble P des nombres premiers).

4.3 Fonctions récursives partielles

Définition 4.17. On considère maintenant l'ensemble $\mathcal{F}^*$ des fonctions partielles. L'ensemble des *fonctions partielles récursives* est le plus petit sous-ensemble de $\mathcal{F}^*$ qui contient les fonctions de base (projections, constantes, et fonction successeur), et qui est clos par les trois opérations suivantes :

– Composition : si $g \in \mathcal{F}^*_n$ et $f_1, \ldots, f_n \in \mathcal{F}^*_p$ sont récursives, alors aussi $f(\bar{x}) = g(f_1(\bar{x}), \ldots, f_n(\bar{x}))$; son domaine sera $\bigcap \mathrm{Dom}(f_i) \cap \{\bar{x} \mid (f_1(\bar{x}), \ldots, f_n(\bar{x})) \in \mathrm{Dom}(g)\}$.

– Récurrence : si $h \in \mathcal{F}^*_{p+2}$, $g \in \mathcal{F}^*_p$ sont récursives, alors aussi la fonction $f \in \mathcal{F}^*_{p+1}$ définie par :
(i) $f(\bar{x}, 0) = g(\bar{x})$ si $\bar{x} \in \mathrm{Dom}(g)$, non définie sinon ;
(ii) $f(\bar{x}, y+1)$ n'est pas définie s'il existe $z \leq y$ tel que $f(\bar{x}, z)$ n'est pas définie, ou bien si $(\bar{x}, y, f(\bar{x}, y)) \notin \mathrm{Dom}(h)$;
(iii) et enfin si pour tout $z \leq y$, $(\bar{x}, z) \in \mathrm{Dom}(f)$, et si $(\bar{x}, y, f(\bar{x}, y)) \in \mathrm{Dom}(h)$, alors $f(\bar{x}, y+1) = h(\bar{x}, y, f(\bar{x}, y))$.

– Schéma μ : Soit $f \in \mathcal{F}^*_{p+1}$ une fonction récursive. Alors la fonction $g(\bar{x}) = (\mu y)(f(\bar{x}, y) = 0)$ est récursive. Elle est définie par : $g(\bar{x}) =$ le plus petit z s'il existe, tel que $f(\bar{x}, z) = 0$ et pour tout $t < z$, $(\bar{x}, t) \in \mathrm{Dom}(f)$. S'il n'existe pas de tel z alors la fonction $(\mu y)(f(\bar{x}, y) = 0)$ n'est pas définie.

Définition 4.18. (1) On appelle *fonction récursive totale* une fonction récursive qui est dans $\mathcal{F}$.
(2) Un sous-ensemble A de $\mathbb{N}^p$ est *récursif* si sa fonction caractéristique $\mathbf{1}_A$ est récursive totale.
(3) Soit $f \in \mathcal{F}_{p+1}$ une fonction récursive totale, et supposons que pour tout $\bar{x}$ il existe y tel que $f(\bar{x}, y) = 0$. La fonction $(\mu y)(f(\bar{x}, y) = 0)$ est alors totale, et nous dirons qu'elle est obtenue à partir de f en appliquant le *schéma μ total*.

Remarques 4.19. (1) Une fonction récursive partielle est donc construite par récurrence à partir des fonctions de base, en utilisant les trois opérations ci-dessus, et en obéissant à certaines règles.
(2) Les fonctions primitives récursives sont récursives (heureusement) et totales.
(3) On remarque que les deux premières opérations appliquées à des fonctions totales, produisent des fonctions totales. C'est seulement la troisième opération qui introduit des fonctions partielles : La fonction $(\mu y)(|x - 2y| = 0)$ n'est définie que si x est pair. Attention : j'ai utilisé la vraie fonction $-$, et pas $\dot{-}$.
(4) Le schéma μ-borné est une conséquence du schéma μ. En effet, on a $(\bar{x}, y) \in A$ si et seulement si $1 \dot{-} \mathbf{1}_A(\bar{x}, y) = 0$, et si on ajoute en plus la condition $z \leq y$, on aura

$$(\mu z \leq y)((\bar{x}, z) \in A) = (\mu z)(1 \dot{-} \mathbf{1}_A(\bar{x}, z)\mathbf{1}_{\leq}(z, y) = 0),$$

où $\mathbf{1}_{\leq}$ est la fonction caractéristique de la relation $\leq$ ($\subset \mathbb{N}^2$).
(5) Soit $g(\bar{x}) = (\mu z)(f(\bar{x}, z) = 0)$, où $f \in \mathcal{F}^*_{p+1}$. Alors $g(\bar{x}) > 0$ implique $f(\bar{x}, g(\bar{x})) = 0$ et pour tout $t < g(\bar{x})$, $(\bar{x}, t) \in \mathrm{dom}(f)$ (et $f(\bar{x}, t) \neq 0$).
$g(\bar{x}) = 0$ implique $f(\bar{x}, 0) = 0$.

Et enfin $\bar{x} \notin \mathrm{dom}(f)$ ne dit rien sur l'existence de y tels que $f(\bar{x}, y) = 0$, il peut en exister.
(6) Si $f(\bar{x}, y)$ et $g(\bar{x})$ sont récursives, alors aussi $g(\bar{x}) = (\mu z)(f(\bar{x}, y) = g(\bar{x}))$: appliquer le schéma μ à $|f(\bar{x}, y) - g(\bar{x})|$.

Définition 4.20. Un ensemble $E \subset \mathbb{N}^p$ est *récursivement énumérable* (abbrégé par RE) s'il est vide, ou bien si c'est le domaine d'une fonction récursive.

Remarque 4.21. La collection des sous-ensembles récursifs de $\mathbb{N}^p$ est close par intersection, par union, et par complémentaire. Par contre, celle des sous-ensembles récursivement énumérables de $\mathbb{N}^p$ est seulement close par intersection et union. Cela vient du fait qu'il existe des ensembles récursivement énumérables qui ne sont pas récursifs. Il est facile de montrer que la classe des ensembles RE est close par intersection, je ne connais pas de façon facile de montrer directement qu'elle est close par union.

Définition 4.22. Soient $f \in \mathcal{F}^*, m \in \mathbb{N}$. On appelle *niveau* de f en m l'ensemble $f^{-1}(m) := \{\bar{x} \in \mathrm{dom}(f) \mid f(\bar{x}) = m\}$.

Remarque 4.23. Les propriétés suivantes sont équivalentes, pour un $E \subset \mathbb{N}^p$:

(1) E est RE.
(2) E est le niveau d'une fonction récursive.
(3) $E = f^{-1}(0)$ pour une fonction récursive f.

Démonstration. Si E est RE, alors $E = \mathrm{Dom}(f)$, f réc., et donc E est le niveau en 0 de $\underline{0} \circ f$, où $\underline{0}$ est la fonction constante égale à 0. Cela montre que (1) implique (3) ; il est clair que (3) implique (2) ; pour (2) implique (1), soient m et f tels que $E = f^{-1}(m)$. On pose

$$g(\bar{x}) = (\mu y)(|f(\bar{x}) - m| = 0).$$

Cette fonction n'est définie que si $f(\bar{x}) = m$, et prend alors la valeur 0.

Théorème 4.24. *Les propriétés suivantes sont équivalentes, pour un $E \subset \mathbb{N}^p$:*

(1) *E est RE.*
(2) *E est la projection d'un niveau de fonction primitive récursive.*

Démonstration. Il faut d'abord montrer que si $E = \pi(S)$, où $S = f^{-1}(m)$ et f est primitive récursive, alors E est RE. Par la remarque précdente, on peut supposer que $m = 0$. On a $\pi : \mathbb{N}^{p+n} \to \mathbb{N}^p$ la projection sur les p premières coordonnées. En utilisant la fonction de codage primitive récursive $\alpha_n : \mathbb{N}^n \to \mathbb{N}$, dont les inverses sont aussi primitive récursives, on peut supposer que $n = 1$. En effet, nous avions

$$\bar{x} \in E \iff \exists \bar{y}\, f(\bar{x}, \bar{y}) = 0,$$

($\bar{x}$ un p-uplet, $\bar{y}$ un n-uplet). Si $\alpha_n : \mathbb{N}^n \to \mathbb{N}$ est la bijection primitive récursive donnée en 4.10, et β_i^n, $1 \leq i \leq n$ sont ses "inverses", alors on a

$$\bar{x} \in E \iff \exists z\, f(\bar{x}, \beta_1^n(z), \ldots, \beta_n^n(z)) = 0,$$

et la fonction $g(\bar{x}, z) = f(\bar{x}, \beta_1^n(z), \ldots, \beta_n^n(z))$ est primitive récursive si f l'est. On pose alors

$$g(\bar{x}) = (\mu y)(f(\bar{x}, y) = 0).$$

Ici nous avons utilisé le fait que la fonction f est totale.

Preuve de la direction difficile. Tout d'abord une définition : nous appellons ensemble *primitif RE*, abbrégé PRE, un ensemble qui est la projection d'un niveau de fonction primitive récursive. Clairement, tout ensemble primitif récursif est PRE, et si f est primitive récursive, alors son graphe Γ_f est PRE (c'est-à-dire, sa fonction caractéristique est primitive récursive). L'intérêt de PRE est qu'on voit tout de suite que

Etape 1/Assertion. *L'ensemble des ensembles PRE est clos par produit cartésien fini, intersection et union finies, et projection.*

Démonstration. Exercice.

De plus, par le raisonnement donné pour la direction facile du théorème (voir ci-dessus), nous pourrons toujours supposer qu'un ensemble $E \subset \mathbb{N}^p$ qui est PRE est défini par

$$E = \{\bar{x} \mid \exists y\, F(\bar{x}, y) = 0\},$$

où y est une seule variable, et F est primitive récursive.

Supposons que $\Gamma_f \subset \mathbb{N}^{p+1}$ soit PRE. Alors $f^{-1}(0)$ est PRE, puisque $f^{-1}(0)$ est la projection sur les p-premières coordonnées de $\Gamma_f \cap (\mathbb{N}^p \times \{0\})$. De plus, $\mathrm{Dom}(f)$ est aussi PRE : c'est la projection sur les p premières coodonnées de Γ_f.

Grâce à la dernière assertion, il suffit donc de montrer que l'ensemble des fonctions $f \in \mathcal{F}^*$ telles que Γ_f soit PRE, est clos par composition, récurrence et schéma μ. Réécrivons encore une fois ce que cela veut dire : il existe une fonction primitive récursive F telle que

$$f(\bar{x}) = y \iff \exists z\, F(\bar{x}, y, z) = 0.$$

De plus on peut remplacer z par un uplet de variables si on veut.

Etape 2. (Clos par composition)
Soient $f_1, \ldots, f_n \in \mathcal{F}_p^*$ et $g \in \mathcal{F}_n^*$, et $h = g \circ (f_1, \ldots, f_n)$. Si Γ_g et les Γ_{f_i} sont PRE, alors aussi Γ_h. En effet, Γ_h est la projection de l'ensemble

$$(\Gamma_{f_1} \times \Gamma_{f_2} \times \cdots \times \Gamma_{f_n} \times \mathbb{N}) \cap \Gamma'$$

sur $p+1$ coordonnées (les p premières et la dernière), où Γ' est l'ensemble des $((p+1)n+1)$-uplets $(x_1, \ldots, x_{(p+1)n+1})$ satisfaisant $(x_{p+1}, x_{2(p+1)}, \ldots, x_{n(p+1)}, x_{n(p+1)+1}) \in \Gamma_g$ et $x_i = x_{i+k(p+1)}$ pour $i = 1, \ldots, p, k = 1, \ldots, n-1\}$.

On peut aussi le voir de façon plus intuitive : $(\bar{x}, z) \in \Gamma_h$ si et seulement si il existe $y_1, \ldots, y_n$ $\bigwedge_{i=1}^n (\bar{x}, y_i) \in \Gamma_{f_i} \wedge (y_1, \ldots, y_n, z) \in \Gamma_h$. Si les fonctions primitives récursives permettant de définir les f_i et g sont F_i et G respectivement, alors on a donc :

$(\bar{x}, z) \in \Gamma_h$ si et seulement si $\exists y_1, \ldots, y_n, u, u_1, \ldots, u_n \, [G(\bar{y}, z, u) = 0 \wedge \bigwedge_{i=1}^{n} F_i(\bar{x}, y_i, u_i) = 0.]$ L'ensemble défini par la formule entre crochets est bien l'ensemble des zéros d'une fonction primitive récursive, et cela entraine que Γ_h est PRE.

Etape 3. Supposons que $g(\bar{x}) = (\mu y)(f(\bar{x}, y) = 0)$, où Γ_f est PRE, $f \in \mathcal{F}^*_{p+1}$. Alors Γ_g est PRE.

Sans perte de généralité, on peut supposer que $\Gamma_f = \pi(F^{-1}(0))$, où $F \in \mathcal{F}_{p+3}$. Posons $\mathrm{Gd}(u, v) = \beta(\beta_1(u), \beta_2(u), v)$ (les fonctions β_i sont les fonctions qui définissent la bijection de $\mathbb{N}^2$ avec $\mathbb{N}$, cf 4.10, la fonction β est celle donnée dans 4.13. C'est juste pour pouvoir utiliser μ qui nous permettra de quantifier existentiellement sur les variables $\beta_1(u)$ et $\beta_2(u)$). Il suit (4.14) donc que : étant donnée une suite finie $a_0, \ldots, a_n$ d'entiers, il existe m tel que que pour tout $i \leq n$, on a $\mathrm{Gd}(m, i) = a_i$; de plus la fonction Gd est primitive récursive. Et posons $s(0) = 1$, $s(x) = 0$ pour $x \geq 1$ (une fonction primitive récursive).

Nous voulons donc les choses suivantes :
(i) $f(\bar{x}, y) = 0 : \exists u \, F(\bar{x}, y, 0, u) = 0$.
(ii) Si $y > 0$ et $k < y$, alors $f(\bar{x}, k)$ est définie, mais non nulle : $\exists u_k \, \exists y_k F(\bar{x}, k, y_k, u_k) = 0 \wedge s(y_k) = 0$.

On remarque qu'une somme d'entiers est nulle si et seulement si chacun de ses termes est nul. Les conditions précédentes nous suggèrent alors la fonction $G \in \mathcal{F}_{p+4}$ définie par :

$$G(\bar{x}, y, u, t, t_1) = F(\bar{x}, y, 0, u) + \sum_{k=0}^{y-1} s(\mathrm{Gd}(t, k)) + F(\bar{x}, k, \mathrm{Gd}(t, k), \mathrm{Gd}(t_1, k)).$$

(Ici les entiers t et t_1 nous serviront à coder les deux suites finies $f(\bar{x}, 0), \ldots, f(\bar{x}, k-1)$ et $u_0, \ldots, u_{k-1}$.) Nous allons montrer que Γ_g est la projection du niveau 0 de G sur les coordonnées $(\bar{x}, y)$. I.e., $g(\bar{x}) = y$ si et seulement si $\exists u, t, t_1 \, G(\bar{x}, y, u, t, t_1) = 0$.

Soit $(\bar{x}, y)$ tel qu'il existe u, t, t_1 avec $G(\bar{x}, y, u, t, t_1) = 0$. Nous voulons montrer que $g(\bar{x}) = y$, c'est à dire que $f(\bar{x}, y) = 0$ et pour $k < y$, $f(\bar{x}, k)$ est définie et ≥ 1.
Tous les termes de G sont égaux à 0 : $F(\bar{x}, y, 0, u)$, $s(\mathrm{Gd}(t, k))$ et $F(\bar{x}, k, \mathrm{Gd}(t, k), \mathrm{Gd}(t_1, k))$ pour $k < y$.
– $F(\bar{x}, y, 0, u) = 0$, nous donne la condition (i).
Si $y = 0$, il n'y a donc rien à prouver : $(\bar{x}, y) \in \Gamma_g$. Supposons $y > 0$, et soit $k < y$. Alors
– $s(\mathrm{Gd}(t, k)) = 0$ implique $\mathrm{Gd}(t, k) \neq 0$, et $F(\bar{x}, k, \mathrm{Gd}(t, k), \mathrm{Gd}(t_1, k)) = 0$ implique $f(\bar{x}, k) = \mathrm{Gd}(t, k) \neq 0$. On a donc la condition (ii) pour tous les $k < y$.

Nous avons donc montré que s'il existe u, t, t_1 tels que $G(\bar{x}, y, u, t, t_1) = 0$, alors $f(\bar{x}, y) = 0$, et pour tout $k < y$, la fonction $f(\bar{x}, k)$ est définie, et non nulle. Ces deux propriétés définissent $g(\bar{x})$ (s'il existe) de façon unique.

Pour l'autre direction, étant donné $(\bar{x}, y) \in \Gamma_g$, nous devons choisir des valeurs pour u, t, t_1 de telle façon que tous les termes de la somme égalent 0. Nous savons que $(\bar{x}, y, 0) \in \Gamma_f$, et donc il existe u tel que $F(\bar{x}, y, 0, u) = 0$. Si $y = 0$, nous choisissons t et t_1 n'importe comment.
Supposons $y > 0$. Nous trouvons d'abord t tel que $\mathrm{Gd}(t, k) = f(\bar{x}, k)$ pour tout $k < y$ (cf

4.14). Pour chaque $k < y$, nous savons que $f(\bar{x}, k)$ est définie, et donc il existe u_k tel que $F(\bar{x}, k, f(\bar{x}, k), u_k) = 0$. Nous prenons t_1 tel que $\mathrm{Gd}(t_1, k) = u_k$. On vérifie que ça marche.

Résumons : l'entier t nous sert à coder la suite des valeurs de $f(\bar{x}, k)$ pour $k < y$, et l'entier t_1 la suite des témoins u_k nécessaires pour déterminer la valeur de $f(\bar{x}, k)$.

Etape 4. Stabilité par récursion.
$g \in \mathcal{F}_p^*$, $h \in \mathcal{F}_{p+2}^*$, permettant de définir $f \in \mathcal{F}_{p+1}^*$ par récurrrence. On veut montrer que si Γ_g et Γ_h sont PRE, alors aussi Γ_f. C'est le même genre d'astuce. On suppose
– $g(\bar{x}) = w$ ssi il existe u, $G(\bar{x}, w, u) = 0$,
– $h(\bar{x}, y, z) = w$ ssi il existe v, $H(\bar{x}, y, z, w, v) = 0$.
Voici l'intuition : $(\bar{x}, 0, w) \in \Gamma_f$ ssi $\exists u\, G(\bar{x}, w, u) = 0$; et si $y > 0$, alors $(\bar{x}, y, w) \in \Gamma_f$ ssi $\exists w_0, \ldots, w_y \exists u, v_1, \ldots, v_y [\bigwedge_{k=1}^{y} H(\bar{x}, k-1, w_{k-1}, w_k, v_k) = 0 \wedge G(\bar{x}, 0, w_0, u) = 0]$. La suite des w_i est la suite des valeurs $f(\bar{x}, i)$, $0 \leq i \leq y$, et la suite des v_i est celle des témoins requis pour obtenir h.

On pose $\tilde{s}(0) = 0$ et $\tilde{s}(x) = 1$ pour $x \geq 1$, et

$$F(\bar{x}, y, w, u, t, t_1) = G(\bar{x}, \mathrm{Gd}(t, 0), u) + \tilde{s}(|w - \mathrm{Gd}(t, y)|) + \sum_{k=1}^{y} H(\bar{x}, k-1, \mathrm{Gd}(t, k-1), \mathrm{Gd}(t, k), \mathrm{Gd}(t_1, k)).$$

La projection du niveau 0 de F sur les coordonnées $(\bar{x}, y, w)$ nous donne alors le graphe de f.

Supposons qu'il existe u, t, t_1 tels que $F(\bar{x}, y, w, u, t, t_1) = 0$. Alors tous ses termes égalent 0, et donc $\mathrm{Gd}(t, y) = w$. Si $y = 0$, alors $G(\bar{x}, \mathrm{Gd}(t, 0), u) = 0$ implique donc $g(\bar{x}) = w$. Si $y > 0$, alors pour tout $1 \leq k \leq y$, on a $H(\bar{x}, k-1, \mathrm{Gd}(t, k-1), \mathrm{Gd}(t, k), \mathrm{Gd}(t_1, k)) = 0$, ce qui implique $\mathrm{Gd}(t, k) = h(\bar{x}, k-1, \mathrm{Gd}(t, k-1))$. En raisonnant par induction, on a donc que $f(\bar{x}, k) = \mathrm{Gd}(t, k)$ pour $k \leq y$, ce qui montre que $w = f(\bar{x}, y)$.

Supposons maintenant que $f(\bar{x}, y) = w$. Pour trouver t et t_1 on raisonne comme dans l'étape précédente.

Remarque 4.25. En fait on peut montrer que l'ensemble des fonctions récursives est le plus petit ensemble de fonctions contenant les fonctions de base, ainsi que la multiplication, l'addition, et la fonction caractéristique de la relation $<$, et qui est close par composition et schéma μ. On n'a donc pas besoin de la récurrence, mais il faut avoir quelques fonctions supplémentaires au début.
Le résultat que nous venons de montrer a plusieurs conséquences importantes, plus ou moins faciles, et que nous allons voir maintenant.

Lemme 4.26. *Une union de deux ensembles RE est RE.*

Démonstration. Soient f et g des fonctions primitives récursives telles que $A = \{\bar{x} \mid \exists y\, f(\bar{x}, y) = 0\}$ et $B = \{\bar{x} \mid \exists y\, g(\bar{x}, y) = 0\}$. Alors $A \cup B = \{\bar{x} \mid \exists y\, f(\bar{x}, y) g(\bar{x}, y) = 0\}$, qui est RE.

Lemme 4.27. *$f \in \mathcal{F}_p^*$ est récursive si et seulement si Γ_f est RE.*

Démonstration. On connaissait déjà une direction : si f est récursive, alors son graphe est RE : on définit $g(\bar{x}, y) = (\mu z)(|f(\bar{x}) - y| = 0)$. Elle est définie en $(\bar{x}, y)$ si et seulement si $f(\bar{x}) = y$.

Pour l'autre direction, soit F primitive récursive, donnée par 4.24, telle que $(\bar{x}, y) \in \Gamma_f$ si et seulement s'il existe z, $F(\bar{x}, y, z) = 0$. On pose $H(\bar{x}, u) = F(\bar{x}, \beta_1(u), \beta_2(u))$ (primitive récursive), et $h(\bar{x}) = (\mu u)(H(\bar{x}, u) = 0)$. Alors h est une fonction partielle, de domaine $\mathrm{Dom}(f)$, et $f(\bar{x}) = \beta_1(h(\bar{x}))$.

Corollaire 4.28. *Une fonction récursive f a une description dans laquelle le schéma μ n'est appliqué qu'une fois. De plus, si f est totale, alors les fonctions intervenant dans la description de f sont toutes totales.*

Démonstration. Par le lemme précédent et le théorème 4.24, il existe une fonction primitive récursive F telle que $f(\bar{x}) = y$ si et seulement s'il existe z tel que $F(\bar{x}, y, z) = 0$. Comme dans le lemme précédent, on a

$$f(\bar{x}) = \beta_1((\mu u)(F(\bar{x}, \beta_1(u), \beta_2(u)) = 0)).$$

Corollaire 4.29. *Soit $A \subset \mathbb{N}^p$ un ensemble RE non vide. Alors il existe une fonction f primitive récursive ($\in \mathcal{F}_1$) telle que l'image de $(\beta_1^p, \ldots, \beta_p^p) \circ f$ soit A. Ici les fonction $\beta_1^p, \ldots, \beta_p^p$ sont les fonctions inverses de la bijection $\alpha_p : |natp^p \to \mathbb{N}$, cf. 4.10.*

Démonstration. On suppose d'abord $p = 1$. Soit F une fonction primitive récursive telle que $\bar{x} \in A$ si et seulement si il existe y, $F(\bar{x}, y) = 0$. On choisit $a \in A$, et on définit

$$f(u) = \begin{cases} \beta_1(u) & \text{si } F(\beta_1(u), \beta_2(u)) = 0, \\ a & \text{sinon.} \end{cases}$$

En utilisant la fonction primitive récursive $\alpha_p : \mathbb{N}^p \to \mathbb{N}$, on a que $\alpha_p(A)$ est un sous-ensemble RE de $\mathbb{N}$; si $\alpha(A) = \mathrm{Im}(f)$, alors $A = (\beta_1^p, \ldots, \beta_p^p)(\alpha(A)) = \mathrm{Im}((\beta_1 \ldots, \beta_p^p) \circ \alpha)$.

Remarque 4.30. Penser à ce résultat de la façon suivante : un ensemble récursivement énumérable peut être énuméré de façon ... récursive (effective, par une machine, ...). On peut donc décider si un élément est dans notre ensemble RE A, mais pas forcément s'il n'y est pas.

Théorème 4.31. *Un sous-ensemble A de $\mathbb{N}^p$ est récursif si et seulement si lui et son complémentaire sont RE.*

Démonstration. La nécéssité est claire : si A est récursif, alors aussi $\mathbb{N}^p \setminus A$, et donc ils sont tous deux RE. Montrons l'autre direction.

Il existe des fonctions récursives f et g telles que $A = \mathrm{Dom}(f)$ et $\mathbb{N}^p \setminus g = \mathrm{Dom}(g)$. En les composant avec des fonctions constantes, nous pouvons supposer que la seule valeur prise par f est 1, et la seule valeur prise par g est 0. Alors $\Gamma_f \cup \Gamma_g$ est le graphe de la fonction totale $\mathbf{1}_A$, et est RE.

Lemme 4.32. *Soit $A \subset \mathbb{N}$, non vide. Montrez que A est récursif si et seulement s'il existe une fonction croissante récursive totale $f \in \mathcal{F}_1$ telle que $A = \mathrm{Im}(f)$.*

Démonstration. Le cas fini étant évident, je supposerai A infini récursif, et vais définir une fonction strictement croissante. $f(0) = (\mu y)(\mathbf{1}_A(y) = 1)$; pour $x > 0$, $f(x) = (\mu y)(\mathbf{1}_A(y) = 1 \wedge y > f(x-1))$.

Dans l'autre direction, supposons f croissante totale récursive donnée, $A = \mathrm{Im}(f)$. Alors A est RE. Si A est fini, alors A est certainement récursif. Et si A est infini, alors $x \notin \mathrm{Im}(f)$ si et seulement si $\exists y\, f(y) < x < f(y+1)$. Cela montre que $\mathbb{N} \setminus A$ est aussi RE, et donc que A est récursif.

Exercice 4.33. Nous savons maintenant que la classe des ensembles RE est close par projection, union, intersection, produit direct. La preuve des étapes 3 et 4 du théorème 4.24 peut être généralisée pour montrer la choses suivante :
Soit F une fonction primitive récursive définie sur $\mathbb{N}^{p+3}$. Alors l'ensemble

$$E = \{(\bar{x}, y) \mid \forall z \leq y \exists u\, F(\bar{x}, y, z, u) = 0\}$$

est RE.
Le problème est évidemment que la quantification bornée universelle apparait **avant** le quantificateur existentiel.

Remarque 4.34. En faisant un peu attention, on peut donc en déduire : Si $E \subset \mathbb{N}^p$ est définissable par une formule dans laquelle toutes les fonctions et relations qui apparaissent sont récursives, et dans laquelle les quantificateurs sont existentiels, ou bien "bornés", alors E est RE. C'est parfois utile.

4.4 Fonctions universelles, etc.

4.35. Je note Alg_p l'ensemble des constructions de fonctions récursives de domaine $\subseteq \mathbb{N}^p$ faites en utilisant les opérations et règles énoncées ci-dessus, et Alg la réunions des Alg_p. Un tel "algorithme" peut donc être représenté par une suite de symboles, ou si vous préférez, "écrit" de la même façon que vous écrivez des formules.

On prend donc un alphabet Σ qui contient des symboles pour toutes les fonctions de base – donc par exemple un symbole pour chaque fonction constante définie sur $\mathbb{N}^p$ – des symboles pour les trois opérations - disons Comp, Rec, Mu, ainsi que des symboles auxiliaires (parenthèses, virgules, ...), qui nous permettront de récupérer notre algorithme à partir de son écriture. On construit une première bijection effective entre Σ et $\mathbb{N}$, qui nous sert ensuite à construire une injection F de Alg dans $\mathbb{N}$ en utilisant les fonctions de codage introduites précédemment. On exige quelques propriétés de F :

(1) L'image de F, $F(\mathrm{Alg})$, est primitive récursive ; pour chaque $p \geq 0$, $F(\mathrm{Alg}_p)$ est primitif récursif.

(2) La fonction qui à un élement $F(I)$, $I \in$Alg, associe l'arité de la fonction définie par I, est

primitive récursive.

(3) Si la chaîne de symboles s apparaît comme sous-suite de la chaîne de symboles t, alors $F(s) \leq F(t)$.

On supposera que les arguments de fonctions sont parmi $x_1, \ldots$ et que la variable sur laquelle μ porte est la dernière de l'énumération des variables de la fonction considérée.

Exemple 4.36. Par exemple, la fonction $\mathrm{Mu}(\mathrm{Rec}(h, \mathrm{Comp}(g, f_1, f_2)))$, où f_1, f_2, g, h sont des fonctions déjà codées, et f_1, f_2 sont p-aires, g est binaire, h est $p+2$-aire, sera codée par

$$\ulcorner \mathrm{Mu} \urcorner \ulcorner (\urcorner \ulcorner \mathrm{Rec} \urcorner \ulcorner (\urcorner \ulcorner h \urcorner \ulcorner , \urcorner \ulcorner \mathrm{Comp} \urcorner \ulcorner (\urcorner \ulcorner g \urcorner \ulcorner , \urcorner \ulcorner f_1 \urcorner \ulcorner , \urcorner \ulcorner f_2 \ulcorner) \urcorner \ulcorner) \urcorner \ulcorner) \urcorner.$$

Ici, $\ulcorner \ \urcorner$ dénote le code donné par la bijection entre Σ et $\mathbb{N}$ si le symbole est dans l'alphabet, et $\ulcorner f_1 \urcorner$, $\ulcorner f_2 \urcorner$, $\ulcorner g \urcorner$ et $\ulcorner h \urcorner$ représentent les codes déjà attribués à f_1, f_2, g, h. L'opération μ portera sur la dernière variable de la fonction $p+1$-aire $f(x_1, \ldots, x_p, x_{p+1})$ définie par $f(x_1, \ldots, x_p, 0) = g(f_1(x_1, \ldots, x_p), f_2(x_1, \ldots, x_p))$, et $f(x_1, \ldots, x_p, x_{p+1}) = h(x_1, \ldots, x_p, x_{p+1} - 1, f(x_1, \ldots, x_p, x_{p+1} - 1))$ si $x_{p+1} > 0$, c'est à dire sur x_{p+1}.

En principe, à partir de l'écriture donnée ci-dessus, on doit être capable d'identifier, dans la suite finie d'entiers $a_0, \ldots a_N$ codée par notre nombre, les morceaux correspondant aux sous-fonctions Rec, Comp, f_1, f_2, g, h. Donc d'abord identifier le code de $\mathrm{Rec}(h, \mathrm{Comp}(g, f_1, f_2))$: Apres $a_0 = \ulcorner \mathrm{Mu} \urcorner$, nous avons $a_1 = \ulcorner (\urcorner$ et $a_N = \ulcorner) \urcorner$; la fonction μ est donc appliquée à la fonction dont le code est $\langle a_2, \ldots, a_{N-1} \rangle$; Nous avons $a_2 = \ulcorner \mathrm{Rec} \urcorner$, puis a_3 et a_{N-1} des codes de parenthèses. Nous voulons identifier les sous-suites correspondant aux deux arguments de Rec. Nous regardons maintenant a_4 ; si c'est le code d'une fonction de base, ce sera le premier argument de notre fonction Rec, et $a_5 = \ulcorner , \urcorner$, et la suite $a_6, \ldots, a_{N-2}$ code notre second argument. Sinon : le premier argument sera donc une fonction plus compliquée, donc qui mettra en jeu des parenthèses : on commence à compter les parenthèses ouvrantes et fermantes qui apparaissent dans la suite $a_4, \ldots, a_{N-2}$, en principe le premier nombre est $\geq$ au deuxième, dès qu'ils sont égaux nous avons notre sous-suite correspondant au premier terme de la Rec. Et cette parenthèse fermante qui donne l'égalité du nombre de parenthèses ouvrantes et fermantes doit être suivie d'une virgule. [Vous voyez que finalement je n'avais pas vraiment besoin des virgules – c'est un dispositif de sécurité].
Etc.

Remarque 4.37. Tout ça est très ennuyeux à écrire de façon précise. Mais vous voyez bien que cela peut être fait. Et de façon tout à fait effective, récursive. Donc, étant donnée une fonction récursive, je choisis une façon de la présenter, de l'écrire, et puis j'associe à cette écriture un code. Il existe plusieurs façons de décrire la même fonction, chaque fonction récursive aura donc une infinité de codes.

4.38. Maintenant nous allons faire une hypothèse, très forte, qui est vraie mais que je ne prouverai pas. (En tout cas, pas dans l'immédiat). Nous supposons qu'il existe une fonction récursive $T : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ satisfaisant les conditions suivantes, pour toute paire $(n, m) \in \mathbb{N} \times \mathbb{N}$:

1 – S'il existe p tel que n est le code d'un élément de Alg_p permettant de définir la fonction récursive f, et si m est le code d'une suite de longueur p, $m = \langle a_1, \ldots, a_p \rangle$, alors

si $f(a_1, \ldots, a_p)$ est définie, alors $T(n, m) = f(a_1, \ldots, a_p)$;

si $f(a_1, \ldots, a_p)$ n'est pas définie, alors $T(n, m)$ n'est pas définie.

2 – Si pour tout p, (n n'est pas le code d'un élément de Alg_p ou m n'est pas le code d'une suite de p éléments) alors $T(n, m) = 0$. (Idéalement, on aimerait pouvoir définir $T(n, m) = -1$).

Définition 4.39. Une telle fonction est appellée *récursive universelle*.

Remarque 4.40. Comme je le mentionnais plus haut, il en existe, il est même facile de les "définir" à partir d'une définition convenable de F, mais la difficulté est de montrer qu'elles sont récursives, car leur définition met en jeu des doubles récurrences : la définition de $T(m+1, n+1)$ utilise la définition de $T(m+1, n)$ et celle de $T(m, j)$ pour des j dont on ne connait pas la taille.
L'existence d'une fonction universelle T entraine l'existence d'une famille de fonctions récursives universelles (ϕ^n) comme dans le TD : On fixe une bijection g_n (récursive) entre $F(\mathrm{Alg}_n)$ et $\mathbb{N}$, et on pose $\phi^n(i, a_0, \ldots, a_{n-1}) = T(g_n^{-1}(i), \langle a_0, \ldots, a_{n-1} \rangle$.

Proposition 4.41. *Soit T une fonction récursive universelle.*

(1) *T n'est pas totale.*
(2) *L'ensemble $\{n \in \mathbb{N} \mid T(n,-)$ est totale$\}$ n'est pas récursif.*
(3) *Dom(T) n'est pas récursif.*

Démonstration. (1) En effet, si n est le code d'une fonction qui n'est pas totale, il existera un m tel que $T(n,m)$ ne soit pas défini.

(2) (*partie corrigée*) Si cet ensemble était récursif, alors aussi son intersection A avec $F(\mathrm{Alg}_1)$. Si $i : A \to \mathbb{N}$ est l'unique bijection croissante de A avec $\mathbb{N}$, alors i^{-1} est récursive (par 4.32). La fonction T' définie par $T'(x,y) = T(i^{-1}(x), 2^{y+1}-1)$ est alors totale, puisque $2^{y+1}-1$ est le code du 1-uplet y. (Ici j'utilise la fonction $\langle a_0, \ldots, a_{n-1}\rangle = \pi(n-1)\prod_{i=0}^{n-1}\pi(i)^{a_n} - 1$; on pourrait en choisir une autre).
On considère la fonction $f(x) = T'(x,x)+1$. Elle est totale, récursive, donc (un de ses algorithmes) a un code $\ulcorner f \urcorner$ dans A ; si $n = i(\ulcorner f \urcorner)$, on aura alors
– $f(n) = T(\ulcorner f \urcorner, 2^{n+1}-1)$ par définition de $\ulcorner f \urcorner$ et de T, et donc $f(n) = T'(n,n)$, et d'autre part,
– $f(n) = T'(n,n)+1$ par définition de f,
ce qui nous donne la contradiction désirée.

(3) Si Dom(T) était récursif, la fonction

$$U(n,x) = \begin{cases} T(n,x) & \text{si } (n,x) \in \mathrm{Dom}(T), \\ 0 & \text{sinon.} \end{cases}$$

serait totale, récursive, et universelle pour les fonctions récursives totales. La même astuce que dans le (2) nous donne la contradiction désirée.

4.42. Notation. Nous avions fixé une fonction F de l'ensemble Alg des écritures de (définitions de) fonctions récursives à valeurs dans $\mathbb{N}$, satisfaisant quelques propriétés simples, cf 4.35. Parmi ces propriétés, il faut que la fonction qui à (m,n) associe l'image par F de la fonction de $\mathcal{F}_m$ constante égale à n (notée Cons(m,n)) soit primitive récursive ; et de même que la fonction qui à $0 \le n \le m$ associe l'image par F de la projection de $\mathbb{N}^m$ sur la n-ième coordonnée soit primitive récursive. Si $I \in \mathrm{Alg}$, je noterai $[I]$, ou $[F(I)]$, la fonction récursive définie par I.

Théorème 4.43. (Théorème smn) *Soient m et n des entiers. Il existe une fonction primitive récursive $s_n^m \in \mathcal{F}_{n+1}$ telle que pour tout $\ell \in \mathbb{N}$, et $a_1, \ldots, a_m \in \mathbb{N}$:*
– si $\ell \in F(\mathrm{Alg}_{m+n})$, alors $s_n^m(\ell, a_1, \ldots, a_n)$ code la fonction $(y_1, \ldots, y_m) \mapsto [\ell](a_1, \ldots, a_n, y_1, \ldots, y_m)$;
– $s_n^m(\ell, a_1, \ldots, a_n) = 0$ sinon.

Démonstration. Nous voulons montrer que la fonction qui à $(z, x_1, \ldots, x_n)$ associe $F(\mathrm{Comp}(F^{-1}(z), \mathrm{Cons}(n+m, x_1), \ldots, \mathrm{Cons}(n+m, x_n)))$ si $z \in F(\mathrm{Alg}_{n+m})$ et 0 sinon, est primitive récursive. C'est évident par notre définition de F.

Remarque 4.44. Le théorème smn nous donne une fonction totale récursive s_n^0, qui associe à une paire (a, b) avec $a \in F(\mathrm{Alg}_n)$ et b le code d'un n-uplet, un élément de Alg_0. Cet élément est seulement une écriture pour un résultat (qui n'existe peut-être pas si la fonction $[a]$ n'est pas définie en b). La fonction s_n^0 n'est donc pas une fonction universelle au sens ou je l'ai définie.

Corollaire 4.45. *Soit $f \in \mathcal{F}_2^*$. Alors il existe $g \in \mathcal{F}_1$, récursive, telle que pour tout x, y,*

$$g(x) \in F(\mathrm{Alg}_1) \text{ et } f(x, y) = [g(x)](y).$$

Démonstration. On applique le théorème smn avec $m = n = 1$ à la fonction $f(x, y)$ et à ℓ tel que $[\ell] = f$: $g(x) = s_1^1(\ell, x)$.

Théorème 4.46. *(Théorème du point fixe de Kleene) Pour toute fonction totale récursive $f \in \mathcal{F}_1$, il existe $n \in F\mathrm{Alg})$ tel que $[n] = [f(n)]$.*

Démonstration. On considère la fonction $H(x, y) = [f([x](x))](y)$. Par le Corollaire précédent, il existe une fonction totale récursive g telle que $H(x, y) = [g(x)](y)$ pour tout x, y, avec de plus $g(x) \in F(\mathrm{Alg}_1)$. Cela bien que la fonction $[x](x) = T(x, x)$ ne soit pas toujours définie en x. Notez que pour tout $a \in \mathbb{N}$, $g(a)$ est simplement l'image par F d'un algorithme qui peut-être ne s'arrête pas. On aura donc, pour tout $a \in \mathbb{N}$ $[f([a](a))] = [g(a)]$ $(*)$.
Si m est tel que $[m] = g$, on regarde $n = [m](m)$: alors

$$[f(n)] = [f([m](m))] = [g(m)] = [[m](m)] = [n].$$

La première égalité suit de la définition de n, la 2ème par $(*)$ ci-dessus, la 3ème parce que $[m]$ est un code pour g, et enfin la 4ème par définition de n.

Théorème 4.47. (Théorème de Rice) *Soit $E \subset F(\mathrm{Alg})$ un ensemble, qui n'est pas vide ni tout $F(\mathrm{Alg})$; on suppose de plus que pour tous m, n, si $[m] = [n]$ alors $m \in E \iff n \in E$. Alors E n'est pas récursif.*

Démonstration. Sinon, soit $a \in F(\mathrm{Alg}) \setminus E$, $b \in E$. Alors la fonction totale

$$T(x) = \begin{cases} a & \text{si } x \in E, \\ b & \text{sinon.} \end{cases}$$

serait récursive. Elle vérifierait

$$x \in E \iff T(x) \notin E.$$

Mais cela contredirait le théorème de Kleene 4.46 : si $n \in F(\mathrm{Alg})$ est tel que $[n] = [T(n)]$ alors on a $n \in E \iff n \notin E$, ce qui est absurde.
Je crois que cet exercice a aussi été fait en TD.

4.5 La fonction d'Ackermann

La fonction d'Ackermann est définie en utilisant une double récurrence. Nous verrons qu'elle n'est pas primitive récursive. Elle est différente de celle donnée en TD, mais a les mêmes propriétés : vous le verrez quand vous calculerez $A(2, x)$. Je donne ici un peu plus de détails qu'en cours.

Définition 4.48. Soit $\xi : \mathbb{N}^2 \to \mathbb{N}$, définie par

$$\xi(0, x) = 2^x, \qquad \xi(y, 0) = 1, \qquad \xi(y + 1, x + 1) = \xi(y, \xi(y + 1, x)).$$

Elle est appelée la *fonction d'Ackermann.*
Si $n, m \in \mathbb{N}$, nous noterons ξ_n la fonction définie par $\xi_n(x) = \xi(n, x)$, et ξ_n^m la fonction obtenue en itérant ξ_n m fois (et donc ξ_n^0 est l'identité). On a donc

$$\xi_{n+1}(x + 1) = \xi_n(\xi_{n+1}(x + 1)).$$

Exemple 4.49. La fonction ξ_0 est tout simplement la fonction $x \mapsto 2^x$.
On a $\xi_1(0) = 1$, $\xi_1(1) = \xi_0(\xi_1(0)) = \xi_0(1) = 2$, $\ldots$, $\xi_1(m) = \xi_0(\xi_1(m - 1)) = \ldots = \xi_0^m(1)$. Donc, $\xi_1(m)$ est une tour d'exponentielles de longueur m.
De même, on montre que $\xi_{n+1}(m) = \xi_n^m(1)$ et $\xi_{n+1}(x + m) = \xi_n^m(\xi_{n+1}(x))$.

Remarque 4.50. On montre les résultats suivants, pour tout n, x, k $(k > 0)$:

(0) Chaque ξ_n est primitive récursive.
(1) $\xi_n(x) > x$.
(2) $\xi_n(x)$ est strictement croissante.
(3) $\xi_{n+1}(x) \geq \xi_n(x)$.
(4) $\xi_n^k(x) \leq \xi_n^k(x)$; $\xi_n^k(x) \geq x$; $\xi_m^k(x) \leq \xi_n^k(x)$. Les inégalités sont strictes si $k > 0$.

Démonstration. Facile. On fait une induction sur n, puis sur x.

Définition 4.51. On dit que $f \in \mathcal{F}_1$ *domine* $g \in \mathcal{F}_p$ s'il existe $A \in \mathbb{N}$ tel que pour tout $\bar{x} = (x_1, \ldots, x_p) \in \mathbb{N}^p$, $g(\bar{x}) \leq f(\sup\{A, x_1, \ldots, x_p\})$.

On pose C_n l'ensemble des fonctions (de $\mathcal{F}$) dominées par au moins une fonction ξ_n^k, et $C = \bigcup_{n\in\omega} C_n$. Nous allons montrer que C contient toutes les fonctions primitives récursives, mais que ξ n'est pas dans C. Cela montrera donc que ξ n'est pas primitive récursive.

Lemme 4.52. *C_0 contient les fonctions constantes, les projections et la fonction successeur.*

Démonstration. Elles sont toutes dominées par ξ_0.

Lemme 4.53. *Chaque C_n est clos par composition : si $f \in \mathcal{F}_m \cap C_n$ et $g_1, \ldots, g_m \in \mathcal{F}_p \cap C_n$, alors $f \circ (g_1, \ldots, g_m) \in C_n$.*

Démonstration. On trouve d'abord $k \in \mathbb{N}$ tel que tous les g_i sont dominées par ξ_n^k, puis ℓ tel que f soit dominée par ξ_n^ℓ. Alors $f \circ (g_1, \ldots, g_m)$ est dominée par $\xi_n^{k+\ell}$.

Lemme 4.54. *Soient $g \in \mathcal{F}_p \cap C_n$, $h \in \mathcal{F}_{p+2} \cap C_n$, et $f \in \mathcal{F}_{p+1}$ définie par récurrence à partir de g et h. Alors $f \in C_{n+1}$.*

Démonstration. Soient k, ℓ, A, B tels que $g(\bar{x}) \leq \xi_n^k(\sup\{x_i, A\})$ et $h(\bar{x}, y, z) \leq \xi_n^\ell(\sup\{x_i, y, z, B\})$. On vérifie alors, par induction sur y, que

$$f(\bar{x}, y) \leq \xi_n^{k+y\ell}(\sup\{x_i, y, A, B\}).$$

Comme $\xi_n^k(x) \leq \xi_n^k(\xi_{n+1}(x)) = \xi_{n+1}(x+k)$, on obtient $f(\bar{x}, y) \leq \xi_{n+1}(\sup\{x_i, y, A_1, A_2\}+k+\ell y)$.

Corollaire 4.55. $\bigcup_n C_n$ *contient toutes les fonctions primitives récursives.*

Démonstration. Clair, puisqu'elle contient les fonctions de base, et est close par composition et définition par récurrence.

Corollaire 4.56. $\xi \notin \bigcup_n C_n$. *La fonction ξ n'est pas primitive récursive.*

Démonstration. La deuxième assertion suit de la première et du corollaire précédent. Supposons, pour obtenir une contradiction, que ξ soit primitive récursive ; alors aussi la fonction $\xi(x, 2x)$, et donc il existe A, n et k tels que

$$\xi(x, 2x) \leq \xi_n^k(x) \text{ pour tout } x > A.$$

En particulier

$$\xi_x(2x) = \xi(x, 2x) \leq \xi_{n+1}(x + k) \text{ si } x > A.$$

Mais c'est impossible si $x > 1, k, n + 1$.

5 L'arithmétique et le théorème d'incomplétude de Gödel

Dans ce chapitre, nous montrerons que la théorie des entiers avec l'addition et la mutiplication, est extrêmement indécidable. Nous commençons avec un essai d'axiomatisation, dû à Peano.

5.1 L'arithmétique de Peano

Nous travaillons dans le langage $\mathcal{L} = \{+, \times, S, 0\}$, où S est une fonction unaire (la fonction successeur), $+$ et $\times$ sont les fonctions binaires usuelles d'addition et de multiplication, et 0 la constante 0.

5.1. Les axiomes de $\mathcal{P}_0$. Nous commençons par une liste finie d'axiomes simples satisfaits par le *modèle standard* $(\mathbb{N}, +, \times, S, 0\}$:

A1 $\forall x\,(S(x) \neq 0)$

A2 $\forall x \exists y\, (x = 0 \vee Sy = x)$
A3 $\forall x, y\, (S(x) = S(y) \to x = y)$
A4 $\forall x\, (x + 0 = x)$
A5 $\forall x, y\ (x + S(y) = S(x + y))$
A6 $\forall x\ (x \times 0 = 0)$
A7 $\forall x, y\ (x \times S(y) = (x \times y) + x)$

5.2. Commentaires.
(1) Si S était omise, $\{0\}$ serait une sous-structure de $\mathbb{N}$, ce que nous ne voulons pas. On aurait pu aussi rajouter une constante pour 1 $(= S(0))$.

(2) Nous abrègerons $S^n(0)$ (l'itérée n fois de S appliquée à 0) par n ou bien par $\underline{n}$ si nous voulons insister sur le fait que nous parlons d'un terme du langage. Un tel n sera aussi appelé un *entier standard*.

(3) Notez que tout modèle de $\mathcal{P}_0$ contient une copie de $\mathbb{N}$: la sous-structure engendrée par $\emptyset$ (qui contient donc 0 puisque c'est une sous-structure, et est close par S). On vérifie facilement qu'elle est close par $+$ et par $\times$, en montrant (par induction) que $S^{m+n}(0) = S^m(0) + S^n(0)$ et que $S^{mn}(0) = S^m(0) \times S^n(0)$.

(4) La théorie $\mathcal{P}_0$ est très faible. On peut montrer qu'elle a des modèles dans lesquels l'addition n'est pas commutative. Je suppose que la raison de prendre une axiomatisation si faible est pour voir jusqu'où on peut descendre et encore avoir certains résultats de définissabilité. Si vous voulez, vous pouvez rajouter les axiomes de commutativité de $+$, $\times$, la distributivité, et tout fragment fini que vous avez envie d'avoir.

(5) On écrit $(x \leq y)$ pour $\exists z\ (z + x = y)$. [Le fait d'ajouter le z "à gauche" est important].

(6) Mon symbole $\times$ va souvent disparaitre – j'utiliserai la notation usuelle xy pour $x \times y$.

Proposition 5.3. *Soit M un modèle de $\mathcal{P}_0$. Alors le sous-ensemble de M formé par les éléments $S^n(0)$, $n \in \mathbb{N}$, est une sous-structure de M qui est isomorphe à $\mathbb{N}$, et qui est un segment initial de M.*

Démonstration. Nous avons déjà parlé de la première assertion. Par "abus de notation", je noterai toujours $\mathbb{N}$ la sous-structure engendrée par $\emptyset$.

Pour la dernière assertion, a priori on ne sait pas que $\leq$ définit un ordre sur M, donc par segment initial, on entend : si $a \leq b$ et $b \in \mathbb{N}$, alors $a \in \mathbb{N}$.

Il faut d'abord montrer que si $c \leq 0$, alors $c = 0$: en effet, soit $d \in M$ tel que $d + c = 0$. Si $c \neq 0$, alors il existe c_1 tel que $c = S(c_1)$, et donc $0 = d + c = d + S(c_1) = S(d + c_1)$ (par A5), ce qui contredit A1.

On montre par induction sur n que si $c \leq n$, alors $c \in \mathbb{N}$. Si $n = 0$ nous venons de le montrer. Supposons le montré pour n, et soit $c \leq n + 1$. Si $c \leq n$ ou si $c = 0$, nous n'avons rien à montrer. Si $c \neq 0$, alors $c = S(c_1)$, et on a $S(c_1) \leq S(n)$. Il existe donc d tel que $d + S(c_1) = S(n)$, d'où $S(d + c_1) = S(n)$ (A5), $d + c_1 = n$ (A3), $c_1 \leq n$ (définition de $\leq$), $c_1 \in \mathbb{N}$ par HI, et $c = S(c_1) \in \mathbb{N}$.

5.4. Peano - suite. A l'origine, Peano a énoncé des propriétés caractérisant $\mathbb{N}$: c'est un modèle de $\mathcal{P}_0$, qui de plus satisfait

Tout sous-ensemble non vide de $\mathbb{N}$ a un plus petit élément.

Cet dernier axiome est un *énoncé de la logique du second ordre*, car il quantifie sur les sous-ensembles du modèle. Il est clair qu'une $\mathcal{L}$-structure modèle de $\mathcal{P}_0$ ayant ces propriétés doit être $\mathbb{N}$, puisque on sait qu'elle en contient une copie comme segment initial, mais que cette copie n'a pas de plus grand élément (et donc le complémentaire n'a pas de plus petit élément, ce qui entraine qu'il est vide). Une façon équivalente d'énoncer cette propriété est de dire : un sous-ensemble P de $\mathbb{N}$ qui contient 0 et est clos par successeur, est tout $\mathbb{N}$. Un essai d'axiomatiser la théorie élémentaire de $(\mathbb{N}, +, \times, S, 0)$ conduit alors à l'ensemble d'axiomes suivant, où l'on restreint son attention aux sous-ensembles **définissables** de $\mathbb{N}$.

5.5. Le schéma $\mathcal{P}$ d'induction de Peano. L'ensemble des axiomes de Peano, $\mathcal{P}$, est formé de $\mathcal{P}_0$, et du schéma d'axiomes suivant : pour toute formule $\varphi(x, \bar{y})$ (x une variable, $\bar{y}$ un uplet de variables), on prend l'axiome

$$\forall \bar{y}[(\varphi(0, \bar{y}) \wedge \forall x\, (\varphi(x, \bar{y}) \to \varphi(S(x), \bar{y}))) \to (\forall x\, \varphi(x, \bar{y}))].$$

Il s'agit donc d'un schéma d'axiomes qui est RE, si on fixe une énumération récursive des variables du langage. Nous montrerons plus tard qu'il n'axiomatise pas $\mathrm{Th}(\mathbb{N})$. On étudie souvent des morceaux de $\mathcal{P}$ (appelés *fragments de Peano*, ou *fragments de l'arithmétique*), en restreignant le schéma d'induction à certaines formules. Les plus célèbres sont le schéma ouvert (*Open induction*) qui ne s'occupe que des formules sans quantificateurs, et le schéma $I\Delta_0$, qui ne permet de quantification que bornée (voir ci-dessous). Il existe en fait toute une hiérarchie de fragments de l'arithmétique, dépendant de la complexité des formules pour lesquelles on ajoute le schéma d'induction.

Proposition 5.6. *Soit M un modèle de $\mathcal{P}$. L'addition et la mutiplication y sont commutatives, la mutiplication est distributive par rapport à l'addition. Tout élément y est régulier pour l'addition, et pour la multiplication s'il est non nul. La formule $x \leq y$ définit un ordre total, compatible avec $+$ et $\times$.*

Démonstration. Régularité : $\forall x, y, z$, $x + y = x + z$ implique $y = z$, et $\forall x, y, z$, si $x \neq 0$, alors $x \times y = x \times z$ implique $y = z$.
Compatibilité : $\forall x, y, z$, si $x \leq y$ alors $x + z \leq y + z$ et $x \times z \leq y \times z$.

Exercice : on utilise l'induction. Quand vous essayerez de montrer la commutativité de $+$, vous vous rendrez compte qu'il faut d'abord montrer des choses plus simples, par exemple : $0 + x = x$, et $S(x) + y = S(x + y)$.

Définition 5.7. On appelle *modèle non-standard* (de $\mathcal{P}_0$ ou de $\mathcal{P}$) un modèle (de $\mathcal{P}_0$ ou de $\mathcal{P}$) qui contient strictement $\mathbb{N}$. De tels modèles existent par Löwenheim Skolem. La sous-structure engendrée par 0 sera appelée le *modèle standard* et sera toujours notée $\mathbb{N}$.

Proposition 5.8. *Soit M un modèle non-standard de $\mathcal{P}$. Soit $\varphi(x) \in \mathcal{L}(M)$ une formule avec une seule variable libre, x. Si pour tout $n \in \mathbb{N}$, on a $M \models \varphi(n)$, alors il existe $c \in M \setminus \mathbb{N}$ tel que $M \models \varphi(c)$.*

Démonstration. Nous avons $M \models \varphi(0)$. Si $M \models \forall x\, \varphi(x)$, alors on prend n'importe quel $c \in M \setminus \mathbb{N}$. Sinon, par le schéma d'induction, $M \models \neg(\forall x(\varphi(x) \to \varphi(S(x)))$, c'est à dire, il existe c tel que $M \models \varphi(c) \wedge \neg\varphi(S(c))$. Un tel c est nécéssairement dans $M \setminus \mathbb{N}$.

5.9. Commentaires. Cela dit que $\mathbb{N}$ n'est pas définissable dans M. Parfois on appelle ce phénomène *overspill* (débordement?). En fait, on a quelque chose d'un peu plus fort, et qui est équivalent au schéma d'induction :

Soit M un modèle de $\mathcal{P}$, et $S \subseteq M$ un sous-ensemble non vide définissable (dans $\mathcal{L}(M)$). Alors S a un plus petit élément.

Démonstration. Soit $\varphi(x)$ la formule (de $\mathcal{L}(M)$) définissant S ; si $M \models \varphi(0)$ il n'y a rien à faire. Sinon, on applique le schéma d'induction à la formule $\tilde{\varphi}(x) = \forall y\, (y \leq x \to \neg\varphi(y))$. On a $M \models \tilde{\varphi}(0)$. Comme $M \models \exists x\ \varphi(x)$, donc $M \models \exists x \neg\tilde{\varphi}(x)$. Par le schéma d'induction appliqué à la formule $\tilde{\varphi}$, il existe $c \in M$, tel que

$$M \models \tilde{\varphi}(c) \wedge \neg\tilde{\varphi}(S(c)).$$

Alors on a que $M \models \forall x\, (x \leq c \to \neg\varphi(x)) \wedge \varphi(c+1)$, et donc $c+1$ est l'élément désiré.

Définition 5.10. Soit $f \in \mathcal{F}_p$ (totale, $\mathbb{N}^p \to \mathbb{N}$). On dit que la $\mathcal{L}$-formule $\varphi(\bar{x}, y)$ *représente* f, si pour tout p-uplet $\bar{a}$, si $b = f(\bar{a})$, alors

$$\mathcal{P}_0 \models \forall y\, \varphi(\underline{\bar{a}}, y) \leftrightarrow y = \underline{b}.$$

Une fonction est *représentable* s'il existe une $\mathcal{L}$-formule qui la représente. Un sous-ensemble E de $\mathbb{N}^p$ est *représentable* si sa fonction caractéristique $\mathbf{1}_E$ est représentable. De façon équivalente, s'il existe une formule $\varphi(\bar{x})$ telle que pour tout uplet $\bar{a}$ dans $\mathbb{N}$ on a $\mathcal{P}_0 \models \varphi(\underline{\bar{a}})$ si $\bar{a} \in E$, et $\mathcal{P}_0 \models \neg\varphi(\underline{\bar{a}})$ si $\bar{a} \notin E$ (Exercice).

Notons qu'ici nous utilisons les termes $\underline{\bar{a}}$ et $\underline{b}$, et que la formule de droite est en fait un énoncé. Le fait qu'une fonction f soit représentable par φ implique que dans tout modèle M de $\mathcal{P}_0$, l'ensemble $\{(\bar{a}, b) \mid \bar{a}, b \in \mathbb{N}, M \models \varphi(\bar{a}, b)\}$ est le graphe de f. A priori, on ne sait pas que la formule $\varphi(\bar{x}, y)$ définit une fonction sur tout M^p.

Définition 5.11. On note Σ_1 le plus petit ensemble de formules qui contient toutes les formules sans quantificateurs, et est clos par $\wedge$, $\vee$, quantification existentielle, et quantification universelle bornée (i.e., si $\varphi(x, \ldots) \in \Sigma_1$, alors $\exists x\, \varphi$ et $\forall y(y < x \to \varphi(x, y, \ldots)$ y sont aussi ; pour la quantification bornée on impose que x et y apparaissent en variables libres).

5.12. Corollaire de la Proposition 5.3. *Soit φ un énoncé Σ_1. Alors*

$$\mathbb{N} \models \varphi \text{ si et seulement si } \mathcal{P}_0 \vdash \varphi.$$

Démonstration. La suffisance de la condition est claire, puisque $\mathrm{Th}(\mathbb{N}) \supset \mathcal{P}_0$. Pour l'autre direction, par le théorème de complétude, il suffit de montrer que si $\mathbb{N} \models \varphi$ alors tout modèle de $\mathcal{P}_0$ satisfait aussi φ. Cela suit de la proposition 5.3 : supposons $\mathbb{N} \models \varphi$, où φ est un énoncé Σ_1, et soit M un modèle de $\mathcal{P}_0$; alors la sous-structure $\mathbb{N}^M$ engendrée par $\emptyset$ satisfait φ puisqu'elle est isomorphe à $\mathbb{N}$. On en déduit que $M \models \varphi$ en se servant du petit lemme suivant qui est laissé en exercice (la démonstration est par induction sur la compléxité de la formule) :

Exercice 5.13. Soient $N \subset M$ des $\mathcal{L}_0$-structures, avec N totalement ordonné, et un *segment initial* de M (i.e., $b \in N$ et $a < b$ impliquent $a \in N$). Si φ est une formule Σ_1, et $\bar{a}$ un uplet de N, alors

$$N \models \varphi(\bar{a}) \text{ implique } M \models \varphi(\bar{a}).$$

Démonstration. L'ensemble des formules satisfaisant cette implication contient les formules sans quantificateurs, est close par conjonction et disjonction, et par quantification existentielle. Il suffit donc de montrer qu'il est clos par quantification universelle bornée. Je vous laisse terminer la preuve.

Remarque 5.14. Ce résultat a pour conséquence le fait suivant : soit $f \in \mathcal{F}_p$ dont le graphe Γ_f est définissable par une formule $\varphi(\bar{x}, y)$ qui est Σ_1. Alors f est représentable. En effet, pour tout $(\bar{a}, b) \in \mathbb{N}^{p+1}$, on a $f(\bar{a}) = b$ ssi $\mathbb{N} \models \varphi(\bar{a}, b)$ ssi $\mathcal{P}_0 \vdash \varphi(\bar{a}, b)$.

Proposition 5.15. *Toute fonction récursive totale est representée par une formule Σ_1.*

Démonstration. Par la remarque précédente, pour montrer qu'une fonction totale est représentable, il suffit de montrer que son graphe est définissable par une fonction Σ_1. La preuve est très similaire à celle donnée pour montrer 4.24. On montre les choses suivantes : l'ensemble des fonctions représentables par une formule Σ_1
– contient les fonctions de base S, $+$ et $\times$, représentées par les formules $y = S(x)$, $x + y = z$ et $x \times y = z$,
– est clos par composition (ici on utilise des quantificateurs existentiels),
– est clos par récurrence.

Pour cette dernière assertion, on utilise le fait que la fonction β de Gödel β est représentable (cf. 4.13 pour la définition), et donc aussi la fonction Gd définie dans la preuve de 4.24. En effet $\beta(x, y, z) = t$ ssi $t < y(z + 1) + 1 \wedge \exists u\, u(y(z + 1) + 1) = x$, et $\beta_1^2(x) = y$ ssi $\exists z(z \leq x \wedge (y + z)(y + z + 1) + 2 = 2x)$, $\beta_2^2(x) = \ldots$

On va exprimer qu'il existe un nombre qui code la suite $f(\bar{x}, 0), f(\bar{x}, 1), \ldots, f(\bar{x}, y)$, dont le dernier membre est z. Si f est définie par récurrence à partir de g et h, on a

$$f(\bar{x}, y) = z \text{ ssi } \exists u(Gd(u, 0) = g(\bar{x})) \wedge (\forall v < y\; Gd(u, S(v)) = h(\bar{x}, y, Gd(u, v))) \wedge Gd(u, y) = z.$$

On vérifie sans peine qu'en utilisant les formules Σ_1 représentant g et h, on obtient une formule Σ_1.

Cela montre donc que toutes les fonctions primitives récursives sont représentables par des formules Σ_1. Maintenant nous utilisons le théorème 4.24 : si f est une fonction récursive (totale ou non), alors il existe une fonction primitive récursive F telle que $f(\bar{x}) = y$ ssi $\exists z\, F(\bar{x}, y, z) = 0$. Il suit que f est représentable par une formule Σ_1 puisque F l'est.

5.16. Retour sur les codages de suites. On a vu deux façons de (dé)coder les suites finies d'entiers, l'une basée sur la fonction β et sa variante Gd, l'autre sur les facteurs premiers. Concernant le codage grâce aux facteurs premiers, il en existe plusieurs variantes, si on n'exige pas que tout entier code une suite. Nous avions vu le codage

$$(a_0, \ldots, a_n) \mapsto \pi(0)^{a_0} \cdots \pi(n-1)^{a_{n-1}} \pi(n)^{a_n+1} - 1$$

grâce auquel tout entier était un code de suite finie, et à partir du code N on pouvait retrouver la longueur de la suite (= 1+ l'indice du plus grand nombre premier divisant le nombre), ainsi que les termes de la suite, comme une fonction de N et de l'indice du terme désiré.

Une des choses qu'on veut, est que à partir du code, on retrouve la longueur de la suite par une fonction primitive récursive ; aussi, on veut que si on a deux suites, on puisse en fabriquer une autre en les mettant bout à bout de façon primitive récursive. Je rappelle que la fonction π qui à n associe le $(n+1)$-ième nombre premier est primitive récursive.

Bien sûr on veut aussi que deux suites différentes n'aient pas le même code, et la fonction $(a_0, \ldots, a_n) \mapsto \prod_{i=0}^{n} \pi(i)^{a_i}$ est donc proscrite, mais on peut aussi considérer les deux fonctions suivantes :

$$(a_0, \ldots a_n) \mapsto \prod_{i=0}^{n} \pi(i)^{a_i+1}, \text{ et}$$
$$(a_0, \ldots, a_n) \mapsto 2^{n+1} \times \prod_{i=0}^{n} \pi(i+1)^{a_i}.$$

Les deux ont leurs avantage. La première est plus facile à définir par récurrence ; la deuxième est plus explicite pour la longueur. J'utiliserai en fait la première, et la noterai Ω.

On note $\lg(N)$ la longueur de la suite codée par N. Faites l'exercice 5.17, qui montre que Ω a les propriétés que nous désirons.

5.2 Codage des formules, des preuves, etc.

On utilise la fonction Ω, qui à une suite finie $(a_0, \ldots, a_n)$ associe $\prod_{i=0}^{n} \pi(i)^{a_i+1}$. On définit $\Omega(\emptyset) = 1$ (Ω ne prendra donc pas la valeur 0). Je présente les idées de comment on arrive à coder des preuves, mais donnerai très peu de détails car ils sont absolument assommants. Je vous renvoie au livre de R. Cori et D. Lascar (Logique Mathématique, Vol. 2, Dunod. Pages 81 – 93 si vous voulez (presque) tous les détails. Il faut absolument que vous soyiez à l'aise avec les propriétés des fonctions primitives récursives, et comment on en construit de nouvelles à partir d'anciennes. Je vous conseille en particulier de faire l'exercice 4.16. Faites surtout l'exercice ci-dessous, j'utiliserai ses résultats par la suite.

Exercice 5.17. (Propriétés de Ω).

(1) Montrez que l'image de Ω, notée $\mathrm{Im}(\Omega)$, est primitive récursive. (C'est-à-dire, sa fonction caractéristique est primitive récursive. Rappel : l'ensemble des ensembles définissables primitifs récursifs est clos par $\cap$, complémentaire et quantification bornée – 4.8).

(2) Montrez que la fonction lg définie par

$$\lg(x) = \begin{cases} 0 & \text{si } x \notin \mathrm{Im}(\Omega), \\ \text{la longueur de la suite codée par } \Omega & \text{sinon.} \end{cases}$$

est primitive récursive. (Par définition, la suite $\emptyset$ est de longueur 0).

(3) Soit $(x, i) \in \mathbb{N}^2$. Si $x \in \mathrm{Im}(\Omega)$, et si $i \in \mathbb{N}$, je note $(x)_i$ le $(i+1)$-ième élément de la suite codée par x si $i < \lg(x)$, et 0 sinon. Si $x \notin \mathrm{Im}(\Omega)$, alors on pose $(x)_i = 0$. Montrez que la fonction $(x, i) \mapsto (x)_i$ est primitive récursive.

Je continue avec un petit lemme, qui nous sera utile pour les définitions par induction sur la compléxité (des termes ou des formules).

Lemme 5.18. *Soient p et n des entiers, $f_1, \ldots, f_n \in \mathcal{F}_1$, $g \in \mathcal{F}_p$ et $h \in \mathcal{F}_{p+n+1}$ des fonctions primitives récursives. On suppose de plus que pour tout $y \in \mathbb{N}$ et $1 \leq i \leq n$, $f_i(y) < y$. Alors l'unique fonction $f \in \mathcal{F}_{p+1}$ définie par*

$$f(\bar{x}, y) = \begin{cases} g(\bar{x}) & \textit{si } y = 0, \\ h(\bar{x}, y, f(\bar{x}, f_1(y)), \ldots, f(\bar{x}, f_n(y))) & \textit{sinon,} \end{cases}$$

est primitive récursive.

Démonstration. On va définir une fonction F telle que $F(\bar{x}, y)$ code la suite $f(\bar{x}, 0), \ldots, f(\bar{x}, y)$ en utilisant le codage Ω. On définit $F(\bar{x}, 0) = 2^{g(\bar{x})+1}$; et

$$F(\bar{x}, y+1) = F(\bar{x}, y)\pi(y+1)^{\gamma},$$

où

$$\gamma = 1 + h(\bar{x}, y, (F(\bar{x}, y))_{f_1(y+1)}, \ldots, (F(\bar{x}, y))_{f_n(y+1)}).$$

La fonction F est primitive récursive, et donc f aussi, puisque $f(\bar{x}, y) = (F(\bar{x}, y))_y$.

5.19. Retour sur la fonction α_3. Je rappelle que la fonction

$$\alpha_2 : (m, n) \mapsto \frac{(m+n)(m+n+1)}{2} + n$$

définit une bijection entre $\mathbb{N}^2$ et $\mathbb{N}$. On a $\alpha_2(0,0) = 0$, $\alpha_2(1,0) = 1$, $\alpha_2(0,1) = 2$, et donc si $\alpha_2(m, n) > 1$, alors $m, n < \alpha_2(m, n)$.

On pose $\alpha_3(k, m, n) = \alpha_2(k, \alpha_2(m, n))$. Cela définit une bijection entre $\mathbb{N}^3$ et $\mathbb{N}$, qui satisfait aussi que si $\alpha_3(k, m, n) > 1$ alors $k, m, n < \alpha_3(k, m, n)$. Je noterai[6] $\gamma_1, \gamma_2, \gamma_3$ les fonctions $\mathbb{N} \to \mathbb{N}$ telles que $\alpha_3^{-1} = (\gamma_1, \gamma_2, \gamma_3)$, et aussi $\gamma_i(\alpha_3(x_1, x_2, x_3)) = x_i$ pour $i = 1, 2, 3$. Les fonctions γ_i sont primitives récursives (grâce au schéma μ-borné).

[6]Dans le Cori-Lascar la notation est différente, β_3^i est notre γ_i.

Exercice 5.20. Calculez $\alpha_3^{-1}(0)$, $\alpha_3^{-1}(1)$, $\alpha_3^{-1}(2)$ et $\alpha_3^{-1}(3)$.

5.21. Stratégie. Nous allons coder les termes et formules en utilisant des fonctions α_3. Les termes ou formules sont tous construits par induction à partir de morceaux plus simples, en utilisant ou bien des fonctions pour les termes, ou bien des symboles logiques pour les fonctions ; le dernier argument de la fonction α_3 nous dira ce que nous faisons.

5.22. Codage des termes. Nous fixons une énumération (indexée par les entiers) $v_0, v_1, \ldots$ des variables du langage ; toute variable, même notée x, sera l'une des v_i. Nous assignons à chaque terme t du langage $\mathcal{L}$ un nombre $\#t$ ou $\#(t)$, le *numéro de Gödel* de t, de la façon suivante, par induction sur la compléxité du terme t :
– $\#0 = \alpha_3(0,0,0)$; $\quad \#v_i = \alpha_3(i+1,0,0)$;
– $\#S(t) = \alpha_3(\#t, 0, 1)$;
– $\#(t_1 + t_2) = \alpha_3(\#t_1, \#t_2, 2)$;
– $\#(t_1 \times t_2) = \alpha_3(\#t_1, \#t_2, 3)$;
De ce codage, on voit facilement qu'on peut savoir si un nombre entier est le numéro de Gödel d'un terme, en utilisant une induction et en regardant la valeur de γ_3. On a :

Lemme 5.23. *L'ensemble Term des $\#(t)$, t un terme du langage, est primitif récursif.*

Démonstration. Soit $x \in \mathbb{N}$. Si $x = 0$ ou $x = 1$, alors $x \in$ Term ($= \#0$, ou $= \#v_0$). Supposons $x > 1$. La définition de la fonction caractéristique sera une définition par cas, chaque cas étant clairement primitif récursif, puisqu'on regardera des valeurs de $\gamma_i(x)$, qui sont strictement plus petites que x, et on pourra donc utiliser le Lemme 5.18. J'écris f pour la fonction caractéristique $\mathbf{1}_{\text{Term}}$ de Term. On a $f(0) = f(1) = 1$. On suppose maintenant $x > 1$:
– Si $\gamma_3(x) = 0$: alors $f(x) = 1$ si $\gamma_2(x) = 0$, et $f(x) = 0$ sinon ;
– Si $\gamma_3(x) = 1$: alors $f(x) = 1$ si $\gamma_2(x) = 0$ et $f(\gamma_1(x)) = 1$, $f(x) = 0$ sinon ;
– Si $\gamma_3(x) = 2$ ou 3, alors $f(x) = f(\gamma_1(x))f(\gamma_2(x))$;
– Si $\gamma(x) > 3$ alors $f(x) = 0$.

5.24. Codage des formules. Nous allons maintenant définir les nombres de Gödel des formules, en commençant par les formules atomiques. Les valeurs de γ_3 des numéros de Gödel des formules prendront toutes les valeurs entre 4 et 11, et aucune autre. Les lettres t_1, t_2 dénotent des termes, φ_1, φ_2, φ des formules, et v une variable.
– $\#(t_1 = t_2) = \alpha_3(\#t_1, \#t_2, 4)$;
– $\#\neg\varphi = \alpha_3(\#\varphi, 0, 5)$;
– $\#(\varphi_1 \wedge \varphi_2) = \alpha_3(\#\varphi_1, \#\varphi_2, 6)$;
– $\#(\varphi_1 \vee \varphi_2) = \alpha_3(\#\varphi_1, \#\varphi_2, 7)$;
– $\#(\varphi_1 \to \varphi_2) = \alpha_3(\#\varphi_1, \#\varphi_2, 8)$;
– $\#(\varphi_1 \leftrightarrow \varphi_2) = \alpha_3(\#\varphi_1, \#\varphi_2, 9)$;
– $\#(\exists v\, \varphi) = \alpha_3(\#\varphi, \#v, 10)$;
– $\#(\forall v\, \varphi) = \alpha_3(\#\varphi, \#v, 11)$.

Commentaire. Dans les articles ou livres auxquels je renvoie pour les détails, les fonctions

sont définies un peu différemment, notamment les "3-ièmes coordonnées" varient de 0 à 7 au lieu de varier de 4 à 11 : j'ai choisi de bien séparer les numéros des termes de ceux des formules. En classe j'ai choisi d'omettre les cas $\gamma_3 = 7, 8, 9, 11$, puisque pour écrire une formule, on n'a pas besoin de $\vee$, $\to$, $\leftrightarrow$ et $\forall$, ils sont définissables à partir des autres en utilisant les opérations booléennes. On peut les mettre ou ne pas les mettre, ça n'a aucune importance, cela influe simplement sur le nombre de façons d'écrire une formule. En particulier, si vous les rajoutez, il faudra mettre parmi les tautologies (voir ci-dessous) celle qui dit que $\varphi \vee \psi$ est équivalente à $\neg((\neg\varphi) \wedge (\neg\psi))$.

Lemme 5.25. *L'ensemble Atom des numéros de Gödel des formules atomiques est primitif récursif, ainsi que l'ensemble Form des numéros de Gödel des formules.*

Démonstration. Exercice (faites-le, c'est facile). Remarquez que Atom est l'ensemble des $x \in$ Form tels que $\gamma_3(x) = 4$, donc il suffit de montrer la seconde assertion.

Lemme 5.26. *Les ensembles suivants sont primitifs récursifs :*

(1) *L'ensemble des $(\#t, n)$ avec t un terme, et v_n n'ayant pas d'occurrence dans t ; donc aussi l'ensemble des $(\#t, n)$ avec t un terme, et v_n ayant au moins une occurrence dans t.*
(2) *L'ensemble des $(\#\varphi, n)$ avec φ une formule, et v_n n'ayant pas d'occurrence libre dans φ ; respectivement, n'ayant pas d'occurrence liée dans φ ; respectivement, n'ayant pas d'occurrence dans φ.*
(3) *L'ensemble des $(\#\varphi, n)$ avec φ une formule, et v_n ayant au moins une occurrence libre dans φ ; respectivement, ayant au moins une occurrence liée dans φ ; respectivement, ayant au moins une occurrence dans φ.*
(4) *L'ensemble des $(\#\varphi, n)$ avec φ un énoncé.*

Démonstration. Je ne vais pas faire la preuve. Elle se fait par induction sur la compléxité des termes (en faisant attention quand $x = 0$ ou $x = 1$), puis sur la compléxité des formules. Par exemple, si f est la fonction caractéristique de l'ensemble des $(\#\varphi, n)$ avec φ une formule, et v_n ayant au moins une occurrence libre dans φ, on aura $f(\#(\neg\varphi), n) = f(\#\varphi, n)$, $f(\#(\varphi_1 \wedge \varphi_2), n) = \sup\{f(\#\varphi_1, n), f(\#\varphi_2, n), \ldots, f(\#(\exists v\varphi), n) = 0$ si $\#v = \#v_n$, et $= f(\#\varphi, n)$ si $\#v \neq \#v_n$.

5.27. De quoi avons-nous encore besoin pour coder des preuves? Nous avons une définition de ce que doit être une preuve. Tout d'abord il y a des énoncés valides qui peuvent toujours être utilisés, ce serait bien que leurs numéros de Gödel forment un ensemble primitif récursif. Ensuite il y a les deux règles de déduction (ou d'inférence) : le Modus Ponens (MP), et la règle de $\exists$-introduction (2.54). Pour cela nous avons besoin de substitutions de variables dans des termes et dans des formules. Ce qui justifie les quelques lemmes suivants, qui sont plus durs :

Lemme 5.28. *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 φ une formule, alors*

$$\mathrm{Subst}_t(n, \#t, \#u) = \#u(t/v_n), \qquad \mathrm{Subst}_f(n, \#t, \#\varphi) = \#\varphi(t/v_n).$$

Démonstration. La définition de la substitution est faite par induction sur la compléxité des termes ou formules. Pour les termes, il n'y a pas de problème ; pour les formules, les connecteurs logiques ne posent aucun problème non plus, mais il faut faire attention quand on regarde une formule de la forme $\exists v\,\varphi$, ou $\forall v\,\varphi$. Il y avait notamment un cas un peu bizarre, le sous-cas (iii) de 2.38, où on remplace v par une variable u n'apparaissant pas dans t : $(\exists v\,\psi)(t/v_n) = \exists u\,(\psi(u/v))(t/v_n)$ (on fait d'abord la substitution de v par u, puis celle de v_n par t). En fait ce cas pose vraiment des problèmes, si on le fait directement. On va faire en sorte qu'il n'apparaisse pas. Il faut faire la procédure en plusieurs étapes séparées :
La première cherche les variables de φ qui sont liées et apparaissent aussi dans t, donne une liste N de leurs nombres de Gödel (codée par Ω), puis donne une liste M de la même longueur de numéros de Gödel de variables n'apparaissant pas dans φ ou dans t.
La deuxième étape fait la substitution (simultanément) des variables données par N par les variables données par M : c'est défini par induction, il n'y a pas de problème puisque le sous-cas (iii) de 2.38 n'apparait pas. Ne pas oublier de substituer aussi les variables quantifiées. Nous obtenons donc le numéro de Gödel d'une formule φ'.
Et la troisième étape calcule $\mathrm{Subst}_f(n, \#t, \#\varphi')$, par induction sur la compléxité de φ' : le sous-cas problématique n'apparait pas.

Je vous laisse regarder les détails dans Cori-Lascar, si vous avez vraiment envie de les voir.

5.29. Les tautologies. Nous avons presque fini d'introduire les outils nécéssaires pour coder des preuves. Il nous reste encore à nous occuper des tautologies. Rappelons d'abord ce qu'est une tautologie. Chaque tautologie est obtenue à partir d'une formule tautologique P du calcul propositionnel (i.e., P est une combinaison Booléenne de variables que je noterai $A_0, A_1, \ldots,$ pour les distinguer des autres) en remplaçant les variables A_i de la formule P par des formules du langage.

Il nous faut donc d'abord étudier les tautologies du langage propositionnel.

5.30. Codage des formules du calcul propositionnel. Les variables sont notées $A_0, A_1, \ldots,$, les formules P, P_1, P_2. On pose :
– $\#A_n = \alpha_3(n+1, 0, 0)$ (il faut éviter que 0 soit un code) ;
– $\#(\neg P) = \alpha_3(\#P, 0, 5)$;
– $\#(P_1 \wedge P_2) = \alpha_3(\#P_1, \#P_2, 6)$;
– ...
– $\#(P_1 \leftrightarrow P_2) = \alpha_3(\#P_1, \#P_2, 9)$.

Lemme 5.31. *L'ensemble Prop des numéros de Gödel de formules du calcul propositionnel est primitif récursif. L'ensemble Taut-Pr des numéros de Gödel des tautologies du calcul propositionnel est primitif récursif.*

Démonstration. La première assertion est facile. Pour la seconde, nous avons besoin d'un lemme intermédiaire, mais d'abord quelques définitions et remarques.

Définition 5.32. On utilise la notation ci-dessus. Une *valeur de vérité* est une fonction $\lambda : \{A_i \mid i \in \mathbb{N}\} \to \{0, 1\}$. On l'étend à toutes les formules du calcul propositionnel, par induction

sur la compléxité de la formule, en posant : $\lambda(\neg P) = 1 - \lambda(P)$; $\lambda(P_1 \wedge P_2) = \lambda(P_1) \times \lambda(P_2)$; et donc aussi $\lambda(P_1 \vee P_2) = \sup\{\lambda(P_1), \lambda(P_2)\}$, $\lambda(P_1 \to P_2) = \sup\{1 - \lambda(P_1), \lambda(P_2)\}$ et $\lambda(P_1 \leftrightarrow P_2) = \sup\{\lambda(P_1) \times \lambda(P_2), (1 - \lambda(P_1)) \times (1 - \lambda(P_2))\}$.

Pour $k \in \mathbb{N}$, on définit une valeur de vérité λ_k sur les variables A_i par $\lambda_k(i) = 1$ si $\pi(i)$ (le $(i+1)$-ième nombre premier) divise k, et 0 sinon. On voit facilement que si λ est une valeur de vérité, et si $N \in \mathbb{N}$, alors il existe un k tel que pour tout $i < N$ on a $\lambda_k(i) = \lambda(i)$. On peut prendre $k = \prod_{i=0}^{N-1} \pi(i)^{\lambda(i)}$; alors $k \leq \pi(N)!$. On étend λ_k aux formules du calcul propositionnel comme ci-dessus.

Lemme 5.33. *La fonction E définie par*

$$E(k,x) = \begin{cases} 0 & \text{si } x \notin \text{Prop}, \\ \lambda_k(P) & \text{si } x \in \text{Prop},\ x = \#P, \end{cases}$$

est primitive récursive.

Démonstration. Par induction sur la compléxité de la formule P. On suppose $x \in$ Prop.
– Si $\gamma_3(x) = 4$: donc $x = \#A_n$ et $E(k,x) = 1$ ssi $\pi(\gamma_1(x) - 1)$ divise k ;
– Si $\gamma_3(x) = 5$: $E(k,x) = 1 \dot{-} E(k, \gamma_1(x))$, puisque $x = \#(\neg P)$;
– Si $\gamma_3(x) = 6$; $E(k,x) = E(k, \gamma_1(x)) \times E(k, \gamma_2(x))$;
– ...

Fin de la preuve de 5.31. On a $x \in$ Taut-Pr si et seulement si pour tout $k \leq \pi(x)!$, $E(k,x) = 1$.

Lemme 5.34. *Etant donnée une formule φ, il existe une formule P_φ du calcul propositionnel, et des formules $\psi_0, \ldots, \psi_n$ telles que $\varphi = P_\varphi(\psi_0, \ldots, \psi_n)$, et pour chaque i, ψ_i est ou bien une formule atomique, ou bien de la forme $\exists v\, \theta$, ou bien de la forme $\forall v\, \theta$.*

Démonstration. Evident par induction sur la compléxité des formules.

Lemme 5.35. *La fonction qui à $x = \#\varphi$ (φ une formule) associe $\#P_\varphi$, et qui prend la valeur 0 en dehors de Form (P_φ donnée par le lemme précédent), est primitive récursive.*

Démonstration. La fonction f est définie par cas, nulle sur le complémentaire de Form, et non nulle sur Form.
– Si $\gamma_3(x) = 4$, 10 ou 11 , $f(x) = \alpha_3(x+1, 0, 0)$;
– Si $\gamma_3(x) = 5$: $f(x) = \alpha_3(f(\gamma_1(x)), 0, 5)$;
– Si $6 \leq \gamma_3(x) \leq 9$: $f(x) = \alpha_3(f(\gamma_1(x)), f(\gamma_2(x)), \gamma_3(x))$.

Lemme 5.36. *L'ensemble Taut des numéros de Gödel des tautologies du calcul des prédicats est primitif récursif.*

Démonstration. On remarque que φ est une tautologie si et seulement si P_φ est une tautologie. Le résultat suit alors des lemmes précédents.

Théorème 5.37. *L'ensemble Ax des numéros de Gödel des axiomes logiques est primitif récursif.*

Démonstration. Facile, cf. sous-section 2.6.

5.38. Autres langages. Il est clair que la méthode que nous avons utilisée pour le langage de l'arithmétique peut être étendue à d'autres langages dénombrables, dès lors qu'ils ont une présentation effective. Par cela, je veux dire qu'il existe une injection (temporairement notée f comme toutes mes fonctions) de l'ensemble des symboles (non logiques) du langage dans $\mathbb{N}$, telle que la fonction décrite ci-dessous soit récursive :

$$G(x) = \begin{cases} 0 & \text{si } x \notin \mathrm{Im}(f), \\ 1 & \text{si } f^{-1}(x) \text{ est un symbole de constante,} \\ \alpha_2(\text{arité de } F, 0) + 2 & \text{si } f^{-1}(x) \text{ est le symbole de fonction } F, \\ \alpha_2(\text{arité de } R, 1) + 2 & \text{si } f^{-1}(x) \text{ est le symbole de relation } R. \end{cases}$$

On peut alors généraliser les résultats précédents à ce langage, en remarquant qu'il faut éventuellement remplacer la conclusion "primitif récursif" par "récursif total" si notre fonction G ci-dessus n'est pas primitive récursive mais est seulement récursive totale. J'ai choisi une façon de formaliser les chose que je demande à mon codage du langage : je veux que son image soit récursive, que je puisse décider de quel genre de symbole il s'agit (constante, relation, fonction), et éventuellement quelle est son arité. Il est clair que pour pouvoir avoir un codage effectif de formules, il faut d'abord un codage effectif des symboles.

Définition 5.39. Soit T une théorie (dans $\mathcal{L}_0$, ou bien dans un langage ayant une présentation effective, comme discuté ci-dessus).

(1) (Je l'utilise tout le temps sans le dire). Un sous-ensemble de $\mathbb{N}^p$ est *primitif récursif* si sa fonction caractéristique est primitive récursive.
(2) On dit que T est *récursive* si $\{\#\varphi \mid \varphi \in T\}$ est récursif.
(3) On dit que T est *récursivement énumérable* si $\{\#\varphi \mid \varphi \in T\}$ est RE.
(4) On dit que T est *décidable* si l'ensemble $\#\mathrm{Th}(T) = \{\#\varphi \mid T \vdash \varphi\}$ est récursif ; et sinon on dit que T est *indécidable*.

5.40. Codage des preuves. Nous allons maintenant utiliser la fonction Ω, (ou n'importe quelle autre fonction qui nous permette de coder les suites finies).
Si $d = (\varphi_1, \ldots, \varphi_n)$ est une suite de formules, on pose

$$\#\#d = \Omega(\#\varphi_1, \ldots, \#\varphi_n).$$

Proposition 5.41. *Soit T une théorie [primitive] récursive. Alors l'ensemble* $\mathrm{Dem}(T)$ *consistant des paires $(\#\varphi, \#\#d)$ où φ est une formule et $\#\#d$ est une démonstration de φ à partir de T, est [primitif] récursif.*

Démonstration. On regarde la définition de preuve : il faut montrer que chacune des formules φ_i de la suite codée par $\#\#d$ est ou bien un axiome logique, ou bien est dans T (ce qui est [primitif] récursif), ou bien est obtenue à partir des précédentes par MP, ou bien est une instance de $\exists$-introduction. Et aussi bien sûr que la dernière formule de la suite codée par $\#\#d$ est bien φ. Sonc, l'ensemble des paires $(\#\varphi, \#\#d)$ comme ci-dessus est défini par une formule ne mettant en jeu que des quantificateurs existentiels et des fonctions primitives récursives.

Corollaire 5.42. *Soit T une théorie RE. Alors $\#\mathrm{Th}(T)$ est RE.*

Démonstration. Soit f une fonction récursive définie sur $\mathbb{N}$ et dont l'image est $\{\#\varphi \mid \varphi \in T\}$ (cf 4.29). Alors $\#\varphi \in \mathrm{Th}(T)$ ssi il existe N, il existe D tel que D code une preuve d de φ à partir de la théorie (finie) $T_N = \{\varphi \mid \exists m < N \ \#\varphi = f(m)\}$. La projection d'un ensemble récursif est RE.

Corollaire 5.43. *Si T a une axiomatisation qui est RE, et si on sait que pour tout énoncé φ, $T \vdash \varphi$ ou bien $T \vdash \neg\varphi$[7], alors T est décidable.*

Démonstration. Par le corollaire précédent, nous savons que $\#\mathrm{Th}(T)$ est RE. Nous savons aussi que pour un énoncé φ,

$$T \nvdash \varphi \iff T \vdash \neg\varphi,$$

ce qui implique que le complémentaire de $\#\mathrm{Th}(T)$ est aussi RE. (Ici, j'utilise sans le dire, le fait que le complémentaire de $\{\#\theta \mid \theta \text{ un énoncé}\}$ est primitif récursif.)

Corollaire 5.44. *Les ensembles des numéros de Gödel de formules prouvables à partir de $\mathcal{P}_0$ ou de $\mathcal{P}$, sont RE.*

Démonstration. Immédiat à partir de 5.42, puisque $\mathcal{P}_0$ et $\mathcal{P}$ sont récursives.

Corollaire 5.45. *L'ensemble des numéros de Gödel de formules universellement valides est RE.*

5.46. Construction diagonale. La fonction Subn qui à une paire $(m, n) \in \mathbb{N}^2$ associe $\#\theta(S^n(0)/v_0)$ si $m = \#\theta(v_0)$ (θ une formule ayant au plus v_0 comme variable libre), et 0 si m n'est pas de cette forme, est primitive récursive. Je note la formule Σ_1 qui la représente subn. Donc, pour tous $(m, n) \in \mathbb{N}^2$, si $k = \mathrm{Subn}(m, n)$ alors

$$\mathcal{P}_0 \vdash \mathrm{subn}(\underline{m}, \underline{n}, \underline{k}) \wedge \forall z\, \mathrm{subn}(\underline{m}, \underline{n}, z) \to z = \underline{k}.$$

On fixe une formule $\varphi(v_0)$, et on définit $\psi(v_0) = \forall v_1[\neg\mathrm{subn}(v_0, v_0, v_1) \vee \neg\varphi(v_1)]$. Etant donné un entier n, nous aurons donc

$$\psi(n) \leftrightarrow (\neg\varphi(\mathrm{Subn}(n, n))), \qquad (*)$$

puisque la fonction Subn est totale sur $\mathbb{N}^2$. On pose $N = \#(\psi(v_0))$ et $\Delta_\varphi = \psi(N)$.

[7]Rappel : une théorie ayant ces propriétés est appelée *complète*.

Proposition 5.47. *Soient φ, ψ, N et Δ_φ comme ci-dessus. Alors*

$$\mathcal{P}_0 \vdash \varphi(\#\Delta_\varphi) \longleftrightarrow \neg\Delta_\varphi.$$

Démonstration. Il suffit de montrer que l'équivalence est vraie dans tout modèle M de $\mathcal{P}_0$. On sait que

$$\mathrm{Subn}(N, N) = \mathrm{Subn}(\#\psi, N) = \#\psi(N) = \#\Delta_\varphi \tag{1}$$

par définition de N et de $\Delta_\varphi = \psi(N)$.
Soit $M \models \mathcal{P}_0$. Si $M \models \varphi(\#\Delta_\varphi)$, alors $M \models \varphi(\mathrm{Subn}(N, N))$, et par $(*)$, nous avons donc $M \models \neg\psi(N)$, c'est à dire, $M \models \neg\Delta_\varphi$.
Réciproquement, si $M \models \neg\Delta_\varphi$, alors $M \models \neg\psi(N)$ et $M \models \varphi(\mathrm{Subn}(N, N))$ (ici nous utilisons le fait que Subn est définie sur $\mathbb{N}$ et $(*)$), i.e., $M \models \varphi(\#\Delta_\varphi)$.

Théorème 5.48. *(Tarski) Soit $M \models \mathcal{P}_0$. Alors l'ensemble*

$$\#\mathrm{Th}(M) = \{\#\varphi \mid \varphi \text{ un énoncé}, M \models \varphi\}$$

n'est pas relativement définissable *dans M, c'est-à-dire, il n'existe pas de formule Sat telle que pour tout énoncé φ, on ait*

$$M \models \mathrm{Sat}(\#\varphi) \text{ si et seulement si } M \models \varphi.$$

Démonstration. Remarque : Si M est non-standard, $\#\mathrm{Th}(M)$ ne peut pas être définissable dans M, car sinon on pourrait définir $\mathbb{N}$ à l'intérieur de M – très certainement l'ensemble $\#\mathrm{Th}(M)$ est cofinal dans $\mathbb{N}$. Le mieux qu'on puisse espérer est que $\#\mathrm{Th}(M)$ soit l'intersection avec $\mathbb{N}$ d'un ensemble définissable de M.

Soit σ une formule, et considérons Δ_σ. Alors, puisque $M \models \mathcal{P}_0$, nous avons, par 5.47,

$$M \models \sigma(\#\Delta_\sigma) \leftrightarrow \neg\Delta_\sigma.$$

Cela montre que σ ne peut être notre formule Sat.

Corollaire 5.49. *Il n'existe pas de formule $Sat_{\mathbb{N}}$ telle que pour tout énoncé φ, on ait $\mathbb{N} \models \varphi$ si et seulement si $\mathbb{N} \models \mathrm{Sat}_{\mathbb{N}}(\#\varphi)$.*

Corollaire 5.50. *L'ensemble $\#\mathrm{Th}(\mathbb{N})$ n'est pas RE.*

Démonstration. Si $\#\mathrm{Th}(\mathbb{N})$ était RE, il serait représenté par une formule Σ_1. Sa fonction caractéristique serait une fonction $\mathrm{Sat}_{\mathbb{N}}$.

Théorème 5.51. *(Church) Soit T une théorie cohérente contenant $\mathcal{P}_0$. Alors T n'est pas décidable.*

Démonstration. Sinon, $\#\mathrm{Th}(T)$ serait récursive, donc représentable par une formule $\varphi(v_0)$. Je rappelle que çà veut dire que la fonction caractéristique de $\#\mathrm{Th}(T)$ est représentable, et donc on a, pour tout énoncé θ, si $T \vdash \theta$, alors $\mathcal{P}_0 \vdash \varphi(\#\theta)$, et si $T \nvdash \theta$, alors $\mathcal{P}_0 \vdash \neg\varphi(\#\theta)$. En particulier, pour tout énoncé θ,

$$T \vdash \theta \leftrightarrow \varphi(\#\theta),$$

puisque T contient $\mathcal{P}_0$.

On considère $\theta = \Delta_\varphi$. Alors par 5.47 on a

$$\mathcal{P}_0 \vdash \varphi(\#\Delta_\varphi) \leftrightarrow \neg\Delta_\varphi.$$

Mettant les deux ensembles, on obtient une contradiction.

Corollaire 5.52. *Si T est une théorie récursive cohérente contenant $\mathcal{P}_0$, alors $\#\mathrm{Th}(T)$ est RE, non récursive.*

Démonstration. Nous avons déjà montré qu'il est RE. Le Théorème de Church nous dit qu'il n'est pas récursif.

Corollaire 5.53. *L'ensemble T_0 des (numéros de Gödel des) énoncés universellement valides n'est pas récursif.*

Démonstration. Soit ψ la conjonction des énoncés de $\mathcal{P}_0$. Alors pour tout énoncé φ, on a

$$\mathcal{P}_0 \vdash \varphi \iff \vdash (\psi \to \varphi).$$

Si $\#T_0$ était récursif, alors aussi $\#\mathrm{Th}(\mathcal{P}_0)$.

Corollaire 5.54. *Soit T une théorie ayant $\mathbb{N}$ pour modèle. Alors T n'est pas décidable.*

Démonstration. Soit ψ la conjonction des énoncés de $\mathcal{P}_0$. Puisque $\mathbb{N}$ est un modèle de T, il suit que $T \cup \{\psi\}$ est cohérente, et implique $\mathcal{P}_0$. Mais on a :

$$\{\varphi \mid T \vdash (\psi \to \varphi)\} = \{\varphi \mid T \cup \{\psi\} \vdash \varphi\},$$

et l'application $\#\varphi \mapsto \#(\psi \to \varphi)$ est clairement récursive. Par le théorème de Church, 5.51, $T \cup \{\psi\}$ est indécidable et donc la même chose est vraie de T.

Corollaire 5.55. *(Non fait en classe) Soit T une théorie (dans le langage $\mathcal{L}_0$, ou bien dans le langage des anneaux) ayant $\mathbb{Z}$ pour modèle (avec l'interprétation naturelle des symboles du langage). Alors T est indécidable.*

Démonstration. Si on travaille dans le langage des anneaux, il faut remarquer que l'ordre est définissable dans $\mathbb{Z}$ par la formule

$$x < y := \exists z_1, z_2, z_3, z_4\, (y - x) = 1 + z_1^2 + z_2^2 + z_3^2 + z_4^2.$$

En effet, dans $\mathbb{Z}$ tout élément positif ou nul est la somme de 4 carrés. Dans ce qui suit, je travaillerai dans $\mathcal{L}_0$, je laisse en exercice la preuve dans le langage des anneaux. Appelons Nat la formule unaire qui exprime $x \geq 0$ dans $\mathbb{Z}$. Supposons T décidable. A tout énoncé φ, associons φ^{Nat}, son relativisé à Nat (voir ci-dessous 5.56). Alors $T \cup \{\psi^{\mathrm{Nat}}\}$ est cohérente (elle a $\mathbb{Z}$ pour modèle), et l'ensemble $\{\#\varphi \mid \psi^{\mathrm{Nat}} \to \varphi^{\mathrm{Nat}}\}$ est décidable, c'est à dire, la théorie $T' = \{\varphi \mid \psi^{\mathrm{Nat}} \to \varphi^{\mathrm{Nat}}\}$ est décidable. Mais cette théorie est cohérente, contient $\mathcal{P}_0$, ce qui contredit le théorème de Church 5.51.

5.56. Relativisation des formules. Je rappelle que si $\theta(x)$ est une formule (unaire), et $\varphi(x_1, \ldots, x_n)$ une formule quelconque, la formule φ relativisée à θ est la formule φ^θ obtenue à partir de $\varphi(x_1, \ldots, x_n)$ en remplaçant chaque occurrence de $\exists y(\cdots)$ par $\exists y\,(\theta(y) \wedge \cdots)$, chaque occurrence de $\forall y(\cdots)$ par $\forall y\,(\theta(y) \to \cdots)$, et prenant la conjonction de la formule obtenue avec $\theta(x_1) \wedge \cdots \wedge \theta(x_n)$. La relativisée d'un énoncé sera un énoncé.

Corollaire 5.57. *Les théories suivantes sont indécidables : Th({anneaux principaux}), Th({anneaux Noethériens}).*

Je rappelle que la théorie d'une classe de structures est l'ensemble des énoncés vrais dans tous les membres de cette classe. Cela suit du résultat precédent puisque $\mathbb{Z}$ est dans les deux classes.

5.58. Soit T une théorie (primitive) récursive. Alors l'ensemble $\mathrm{Dem}(T)$ des paires $(\#\varphi, \#\#d)$ où φ est un énoncé, et d une preuve de φ à partir de T, est (primitif) récursif (cf 5.41), et nous prenons $\mathrm{Pr}_T(v_0, v_1)$ une formule (Σ_1) le représentant, et posons $h_T(v_0)$ la formule $\exists v_1 \, \mathrm{Pr}_T(v_0, v_1)$.

Théorème 5.59. *(Théorème d'incomplétude de Gödel) Soient T une théorie récursive et cohérente contenant $\mathcal{P}_0$, et h_T la formule définie ci-dessus. Alors*

$$\mathbb{N} \models \Delta_{h_T}, \ \textit{mais } T \not\vdash \Delta_{h_T}.$$

Théorème 5.59. (Théorème d'incomplétude de Gödel) *Soient T une théorie récursive et cohérente contenant $\mathcal{P}_0$, et h_T la formule définie ci-dessus. Alors*

$$\mathbb{N} \models \Delta_{h_T}, \ \textit{mais } T \not\vdash \Delta_{h_T}.$$

Démonstration. Soit $N = \#\Delta_{h_T}$. Nous avons donc, par 5.47,

$$\mathcal{P}_0 \models \exists v_1 \, \mathrm{Pr}_T(N, v_1) \leftrightarrow \neg \Delta_{h_T}. \qquad (*)$$

Supposons $T \vdash \Delta_{h_T}$. Il existe donc une preuve de Δ_{h_T} à partir de T, c'est-à-dire : il existe un entier n tel que $\mathbb{N} \models \mathrm{Pr_T}(N, n)$. Mais comme Pr_T est Σ_1, cela entraine (par 5.12) que $\mathcal{P}_0 \vdash \mathrm{Pr_T}(N, n)$; d'où $\mathcal{P}_0 \vdash \neg \Delta_{h_T}$, impossible car $T \supset \mathcal{P}_0$ et est cohérente. Donc $T \not\vdash \Delta_{h_T}$. Cela montre la deuxième assertion.

Pour la première, nous raisonnons aussi par contradiction, et supposons que $\mathbb{N} \models \neg \Delta_{h_T}$. Par $(*)$ nous avons $\mathbb{N} \models \exists v_1 \, \mathrm{Pr}_T(N, v_1)$. Il existe donc une preuve de Δ_{h_T} à partir de T, ce qui montre $T \vdash \Delta_{h_T}$ et contredit ce que nous venons de montrer.

5.60. Il existe une forme un peu plus forte du Théorème d'incomplétude, dûe à Rosser. La fonction qui à $\#\varphi$ associe $\#(\neg\varphi)$ est primitive récursive, soit $\mathrm{Neg}(x, y)$ la formule qui la représente. Nous définissons ensuite

$$\mathrm{Rosser}_T(v_0, v_1) = \mathrm{Pr}_T(v_0, v_1) \wedge \neg(\exists v_2 \leq v_1 \exists u \, (\mathrm{Pr}_T(u, v_2)) \wedge \mathrm{Neg}(v_0, u)).$$

Donc, grosso modo, cette formule dit qu'il existe une preuve de l'énoncé codé par v_0 mais il n'en existe pas de sa négation. On a alors

Théorème 5.61. *(Rosser) Soient T une théorie récursive et cohérente contenant $\mathcal{P}_0$, et h_T^R la formule $\exists v_1 \, \mathrm{Rosser}(v_0, v_1)$. Alors*

$$T \not\vdash \Delta_{h_T^R}, \ \textit{et } T \not\vdash \neg \Delta_{h_T^R}.$$

Démonstration. Soient $N = \#\Delta_{h_T^R}$, et $N_1 = \mathrm{Neg}(N)$. Si $T \vdash \Delta_{h_T^R}$, alors $T \vdash \neg(\exists v_1 \, \mathrm{Rosser}(N, v_1))$, par 5.47.

Mais d'autre part, il existe une preuve de $\Delta_{h_T^R}$ à partir de T, disons de numéro n, et comme T est cohérente, il n'existe pas de preuve de $\neg \Delta_{h_T^R}$ à partir de T. C'est à dire, $\mathbb{N} \models \mathrm{Rosser}(N, n)$,

mais il faut un peu argumenter avant de dire que $\mathcal{P}_0$ prouve $\mathrm{Rosser}(N,n)$ car cette formule contient un quantificateur universel. Nous avons

$$\mathbb{N} \models \mathrm{Pr}_T(N,n) \wedge \neg(\exists v_2 \leq v_1(\mathrm{Pr}_T(N_1,v_2)) \wedge \mathrm{Neg}(N,N_1)),$$

d'où

$$\mathcal{P}_0 \vdash \mathrm{Pr}_T(N,n) \wedge \neg(\exists v_2 \leq v_1(\mathrm{Pr}_T(N_1,v_2)) \wedge \mathrm{Neg}(N,N_1)),$$

et donc

$$\mathcal{P}_0 \vdash \mathrm{Rosser}(N,n).$$

Cela nous donne la contradiction désirée, et donc $T \nvdash \Delta_{h_T^R}$.
L'autre direction est similaire : supposons que $T \vdash \neg\Delta_{h_T^R}$, et soit $e \in \mathbb{N}$ tel que $\mathcal{P}_0 \vdash \mathrm{Pr}_T(N_1,e)$. Alors $\mathcal{P}_0 \vdash \forall x(e \leq x \to \exists z \leq x\, \mathrm{Pr}_T(N_1,z))$ (c'est trivial, puisque $z = e$ marche). De plus, comme T est cohérente, il n'existe pas de preuve de $\Delta_{h_T^R}$ à partir de T, et donc en particulier, on a

$$\mathcal{P}_0 \vdash \forall x\, \mathrm{Pr}_T(N,x) \to e \leq x.$$

Nous avons donc que

$$\mathcal{P}_0 \vdash \forall x\, (\mathrm{Pr}_T(N,x) \to \exists z \leq x\, \exists u\, \mathrm{Neg}(N,u) \wedge \mathrm{Pr}_T(u,z)). \qquad (\dagger)$$

D'un autre côté, comme $T \vdash \neg\Delta_{h_T^R}$, par 5.47, nous avons $T \vdash h_T^R(N)$, c'est-à-dire,

$$T \vdash \exists x\, \mathrm{Pr}_T(N,x) \wedge \neg(\exists z \leq x\, \exists u\, \mathrm{Neg}(N,u) \wedge \mathrm{Pr}_T(u,z)), \qquad (\ddagger)$$

ce qui nous donne la contradiction désirée.

5.62. Soit T une théorie récursive. On note $\mathrm{Coh}(T)$ l'énoncé

$$\neg(\exists v_1 \mathrm{Pr}_T(\#(0 \neq 0), v_1)).$$

Alors $\mathbb{N} \models \mathrm{Coh}(T)$ si et seulement si T est cohérente.

Théorème 5.63. *(Second théorème d'incomplétude de Gödel) Soit T une théorie récursive et cohérente contenant $\mathcal{P}$ (ou au moins $\mathcal{P}_0$ plus le schéma d'induction appliqué aux formules Σ_1). Alors $T \cup \{\mathrm{Coh}(T)\} \vdash \Delta_{h_T}$. En particulier, Coh$(T)$ n'est pas démontrable à partir de T.*

h_T est la formule du 1er théorème d'incomplétude. Pour les détails voir le livre de Cori et Lascar, Logique Mathématique, Tome 2. En fait je vous mens un peu : la formule $\mathrm{Coh}(T)$ est une formule qui exprime la propriété ci-dessus, mais elle est a priori beaucoup plus compliquée.

5.3 Le dixième problème de Hilbert

5.64. En 1900, lors du congrès international des mathématiciens, David Hilbert a proposé une liste de 23 problèmes, dont plusieurs restent ouverts à ce jour, et font encore l'objet de recherches poussées. Le 10ème problème de Hilbert a été résolu par la négative, mais a plusieurs variantes qui sont encore ouvertes.

H10 (Formulation d'origine) *Etant donné $n \geq 2$, y a-t-il un algorithme qui permette de décider si une équation $P(x_1, \ldots, x_n) = Q(x_1, \ldots, x_n)$ a une solution dans $\mathbb{N}^n$, où P et Q sont des polynômes à coefficients dans $\mathbb{N}$?*

Ceci nous pose une question sur $\mathbb{N}$. En fait, on peut la traduire facilement en une question portant sur $\mathbb{Z}$, en faisant les remarques suivantes :
$\mathbb{N}$ est le sous-ensemble de $\mathbb{Z}$ consistant de sommes de 4 carrés ;
Une somme de carrés est nulle si et seulement chacun de ses termes est nul : cela nous permet de considérer des systèmes finis d'équations.

H10 (version pour $\mathbb{Z}$). *($n \geq 2$) Y a-t-il un algorithme qui permette de décider si une équation $P(x_1, \ldots, x_n) = 0$ a une solution dans $\mathbb{Z}^n$, où P est un polynôme à coefficients dans $\mathbb{Z}$?*

Définition 5.65. Un sous-ensemble $S \subset \mathbb{Z}^n$ est *diophantien* s'il est de la forme

$$S = \{\bar{x} \in \mathbb{Z}^n \mid \mathbb{Z} \models \exists \bar{y}\, P(\bar{x}, \bar{y}) = 0\}$$

pour un polynôme P à coefficients dans $\mathbb{Z}$.

Théorème 5.66. *(Matyasevich) Les ensembles RE sont diophantiens.*

Ce théorème apporte la touche finale au travail de plusieurs autres personnes : Martin Davis, Hilary Putnam et Julia Robinson. Je vais vous indiquer quelques-uns des ingrédients.

On travaille dans la $\mathcal{L}_0$-structure $\mathbb{N}$ ou dans l'anneau $\mathbb{Z}$ indifféremment. On sait déjà que tout ensemble RE est définissable par une formule Σ_1. En fait, on peut affiner cette représentation, en montrant que si $S \subset \mathbb{Z}^n$ est RE, alors il existe un polynôme P tels que pour tout $\bar{a} \in \mathbb{Z}^n$, on a

$$\bar{a} \in S \iff \exists x \forall y \leq x \exists z_1, \ldots, z_m\, (P(\bar{a}, x, y, \bar{z}) = 0).$$

Il faut bien sûr éliminer le quantificateur $\forall y \leq x$, c'est ce qui pose problème.

On montre ensuite qu'on arrive à l'éliminer si on permet des équations polynomiales-exponentielles (= termes du langage des anneaux avec l'exponentielle x^y) de la forme $E(\bar{z}) = E'(\bar{z})$. Cela réduit le problème à montrer que le graphe de $(x, y) \mapsto x^y$ est diophantien. Et c'est ce que prouve Matyasevich.

Exercice 5.67. Pourquoi le résultat de Matyasevich implique-t-il la réponse négative à H10 ?

5.68. Variantes. Le problème peut être généralisé à un anneau R quelconque, à condition que cet anneau R soit récursif, i.e., qu'il existe un plongement $f : R \to \mathbb{N}$ tel que les images par f des ensembles suivants soient récursifs : R ; le graphe de l'addition ; le graphe de la multiplication.

H10(R) *Y a-t-il un algorithme qui permette de décider si une équation $P(x_1, \ldots, x_n) = 0$ a une solution dans R^n, où P est un polynôme à coefficients dans R ?*

Quelques noms de personnes travaillant ou ayant travaillé dans ce domaine : Denef, Mazur, Pheidas, Shlapentokh, Vidaux, Zahidi, et beaucoup d'autres. La plupart des réponses obtenues sont négatives. Deux problèmes extrêmement ouverts :

H10($\mathbb{Q}$)
H10($C(t)$), C un corps algébriquement clos dénombrable.

6 Théorie des ensembles – la suite

Nous avons déjà vu un peu de théorie naïve des ensembles, qui suppose très peu de choses, et qui est très informelle. Nous allons introduire le système d'axiomes ZF, introduit par Zermelo-Fraenkel. Il s'agit de donner des axiomes pour la théorie des ensembles, nous avons vu qu'il fallait faire attention pour ne pas arriver à des contradictions. Pour le moment, on n'a pas trouvé de contradiction de ZF.

Nous avons vu que l'ensemble de tous les ensembles ... n'existe pas. Mais comment peut-on en parler? Je parlerai de *collection*, ou bien de *classe* de tous les ensembles.

Nous travaillerons dans le langage avec un seul symbole binaire, $\in$; $a \in b$ se lit *a appartient à b*. Notre univers (la classe de tous les ensembles) sera notée $\mathcal{U}$, et nous allons imposer une série d'axiomes sur $\mathcal{U}$, dont nous exigeons que $(\mathcal{U}, \in)$ soit un modèle. Nous exigeons d'abord que $\mathcal{U}$ est **non vide**. Il faudra donc faire la différence entre deux sortes d'ensembles : ceux qui sont des éléments de $\mathcal{U}$, et les autres, dont on parle dans le langage courant, et qui seront en général définissables dans $\mathcal{U}$.

Exercice 6.1. Montrez que la collection On des ordinaux est définissable dans $\mathcal{U}$, et n'est pas un ensemble.

6.1 Les axiomes de Zermelo - Fraenkel

6.2. ZF1 - l'Axiome d'extensionnalité. Nous avons déjà vu cet axiome : il dit que deux ensembles ayant les mêmes éléments sont égaux. Vérifions que c'est bien une formule du premier ordre :

$\forall x \forall y \, [\forall z \, (z \in x \leftrightarrow z \in y) \rightarrow x = y]$.

6.3. ZF2 – Axiome de la réunion ou de la somme. Etant donné un ensemble a, il existe un ensemble b dont les éléments sont les éléments d'éléments de a ; cet ensemble est noté $\bigcup_{x \in a} x$, et est unique.

$\forall x \exists y \forall z \, [z \in y \leftrightarrow \exists t \, (t \in x \wedge x \in t)]$.
Pensez à a comme à une famille d'ensembles ; alors b est l'union des membres de cette famille.

6.4. ZF3 – Axiome des parties. On note $a \subset b$ pour $\forall x \, (x \in a \rightarrow x \in b)$, lu *a sous-ensemble de b*, ou bien *a inclus dans b*. Cet axiome dit que l'ensemble $\mathcal{P}(a)$ des sous-ensembles de l'ensemble a existe.

$\forall x \exists y \forall z\, [z \in y \leftrightarrow z \subset x]$.

Définition 6.5. Une *relation fonctionnelle en* w_0 est une formule $\varphi(w_0, w_1, a_1, \ldots, a_n)$, où les a_i sont dans $\mathcal{U}$, telle que

$$\mathcal{U} \models \forall w_0, w_1, w_2\, (\varphi(w_0, w_1, a_1, \ldots, a_n) \wedge \varphi(w_0, w_2, a_1, \ldots, a_n)) \rightarrow (w_1 = w_2).$$

C'est à dire que $\varphi(-, -, a_1, \ldots, a_n)$ définit le graphe d'une fonction partielle. Je noterai parfois cette relation fonctionnelle $f_{\bar{a}}$, mais il vaut mieux réserver l'appellation *fonction* à une "vraie" fonction, c'est à dire avec domaine et image qui soient des ensembles et pas des classes.

Plus généralement, je noterai souvent F une relation fonctionnelle (par abus de notation). Il sera sous-entendu qu'il existe une formule φ qui définit dans $\mathcal{U}$ la relation fonctionnelle F.

6.6. ZF4 – Axiome de substitution ou remplacement. Le schéma dit : pour tout n-uplet $\bar{v}$, si la formule $\varphi(w_0, w_1, \bar{v})$ définit une relation fonctionnelle $f_{\bar{v}}$ en w_0, et si a est un ensemble, alors l'ensemble des images par $f_{\bar{v}}$ des éléments de a est aussi un ensemble. Donc

$\forall v_1, v_1, \ldots, v_n \forall w_0, w_1, w_2 [(\varphi(w_0, w_1, v_1, \ldots, v_n) \wedge \varphi(w_0, w_2, v_1, \ldots, v_n) \rightarrow (w_1 = w_2)) \rightarrow$
$(\exists v_{n+1} \forall v_{n+2}\, (v_{n+2} \in v_{n+1} \leftrightarrow \exists w_0\, w_0 \in v_0 \wedge \varphi(w_0, w_2, v_1, \ldots, v_n)))]$.

6.7. Schéma de compréhension. Ce schéma dit que si $\varphi(x)$ est une formule (à paramètres dans $\mathcal{U}$) et si a est dans $\mathcal{U}$ alors la collection des éléments de a qui satisfont φ est un ensemble.

Pour toute formule $\varphi(x, v_1, \ldots, v_n)$, la formule suivante
$\forall v_1, \ldots, v_n, v_{n+1} \exists x \forall y\, [y \in x \leftrightarrow y \in v_{n+1} \wedge \varphi(y, v_1, \ldots, v_n)]$.
Preuve que ZF4 implique le schéma de compréhension : on considère la formule

$$\psi(z, y, v_1, \ldots, v_n) := (z = y) \wedge \varphi(y, v_1, \ldots, v_n).$$

Elle est fonctionnelle en z ; si $a_1, \ldots, a_n$ sont dans $\mathcal{U}$, alors le domaine de la fonction $f_{\bar{a}}$ définie par ψ est la classe des ensembles x qui satisfont $\varphi(x, \bar{a})$, et égale son image. On applique ZF4 à ψ et à l'ensemble v_{n+1}.

Théorème 6.8. *(Existence de $\emptyset$) Il existe un ensemble n'ayant aucun élément.*

Démonstration. Soit a un ensemble. On applique le schéma de compréhension à a et à la formule $x \neq x$. On obtient donc

$$b = \{x \in a \mid x \neq x\}.$$

Un tel b ne peut pas avoir d'éléments.

6.9. Axiome de la paire. Il est en fait une conséquence des axiomes ZF3 et ZF4, et dit que étant donnés deux ensembles a et b, il existe un ensemble c dont les éléments sont exactements a et b. On écrit $c = \{a, b\}$, et on parle de la *paire* $\{a, b\}$.

$\forall x, y\, \exists z\, \forall t\, [t \in z \leftrightarrow (t = x \vee t = y)]$.

On remarque que si $a = b$, alors $\{a, b\}$ est le *singleton* $\{a\}$. La *paire ordonnée* (a, b), aussi appelée le *couple* (a, b), est l'ensemble $\{\{a\}, \{a, b\}\}$. Alors $(a, a) = \{\{a\}\}$. Elle est définie par

$\forall x, y \exists z \forall t\, [t \in z \leftrightarrow (t = \{x\} \vee t = \{x, y\})]$, où la formule $t = \{x, y\}$ est donnée ci-dessus, et la formule exprimant $t = \{x\}$ est ... ?

6.10. *Démonstration de l'axiome de la paire.* On remarque que si $a \subset \emptyset$ alors $a = \emptyset$. Donc $\mathcal{P}(\emptyset) = \{\emptyset\}$, et $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$. Nous avons donc un ensemble $c = \mathcal{P}(\{\emptyset\})$ ayant deux éléments distincts. On considère la formule

$\varphi(x, y, \emptyset, \{\emptyset\}, a, b) := (x = \emptyset \wedge y = a) \vee (x = \{\emptyset\} \wedge y = b)$.

Cette formule est fonctionnelle en x, on applique le schéma de substitution à c, et on obtient un ensemble d, dont les éléments sont a et b.

Théorème 6.11. *Soient a, b, a', b' des ensembles.*

(1) $\{a, b\} = \{a', b'\}$ *si et seulement si* $(a = a'$ *et* $b = b')$ *ou* $(a = b'$ *et* $b = a')$.
(2) $(a, b) = (a', b')$ *si et seulement si* $a = a'$ *et* $b = b'$.

Démonstration. Exercice.

Définition 6.12. Soient n un entier et $a_1, \ldots, a_n$ des ensembles. On définit $(a_1, \ldots, a_n)$ par induction sur n en posant $(a_1, \ldots, a_n) = (a_1, (a_2, \ldots, a_n))$.

On vérifie de la même façon que $(a_1, \ldots, a_n) = (b_1, \ldots, b_n)$ ssi $a_1 = b_1$, …, $a_n = b_n$.

6.13. Conséquences. Soient a, b deux ensembles, et $c = \{a, b\}$. Alors $\bigcup_{x \in c} x = a \cup b$. De même, si $c = \{a_1, \ldots, a_n\}$, alors $\bigcup_{x \in c} x = a_1 \cup \cdots \cup a_n$.

6.14. Le produit de deux ensembles existe. *Soient a et b des ensembles ; on considère la collection X des paires ordonnées (x, y) où $x \in a, y \in b$, et nous allons montrer que c'est un ensemble.*
Si (x, y) est dans X, alors $(x, y) \in \mathcal{P}(\mathcal{P}(a \cup b))$: en effet, les éléments de (x, y) sont $\{x\}$ et $\{x, y\}$, qui sont dans $\mathcal{P}(a \cup b)$. Par compréhension (et définissabilité de X), X est bien un ensemble.

6.15. Ensembles de fonctions. *Soient a et b des ensembles. Alors la collection de toutes les fonctions de a dans b, notée b^a, est un ensemble.*
Une application de a dans b est donnée par son graphe, qui est inclus dans $a \times b$. On a donc que f est une fonction si et seulement si

$f \subset a \times b \wedge \forall x, y, y' \, [(x, y) \in f \wedge (x, y') \in f \to y = y'] \wedge \forall x \, [x \in a \to \exists y \, (y \in b \wedge (x, y) \in f)]$.

6.16. Réunions, intersections et produits de familles. Soit a une famille d'ensembles indexée par l'ensemble I, c'est à dire, a est une fonction de domaine I. Si $i \in I$, on notera (souvent) a_i pour $a(i)$. Si $b = \{a_i \mid i \in I\}$ (qui est un ensemble par ZF4), alors on définit

$$\bigcup_{i \in I} a_i = \bigcup_{x \in b} x$$

qui existe par ZF2. Cette union est aussi définie par la formule (en y) $\forall x[x \in y \leftrightarrow \exists i \, (i \in I \wedge x \in a_i)]$.

On définit alors l'intersection des a_i, notée $\bigcap_{i \in I} a_i$, de la façon suivante. Si $I = \emptyset$, ce n'est pas un ensemble car c'est tout $\mathcal{U}$. Supposons I non vide, et soit $c = \bigcup_{i \in I} a_i$ (un ensemble). Alors la collection des $x \in c$ tels que $\forall i \, (i \in I \to x \in a_i)$ est un ensemble.

Finalement, le produit des a_i, noté $\prod_{i\in I} a_i$, est l'ensemble des fonctions $f : I \to \bigcup_{i\in I} a_i$ telles que $f(i) \in a_i$ pour tout $i \in I$. Une telle fonction est dans $(\bigcup_{i\in I} a_i)^I$. Il suit que $\prod_{i\in I} a_i$ est bien un ensemble.

6.17. ZF5 - l'axiome de l'infini. C'est un axiome qui dit qu'il existe un ensemble ayant une infinité d'éléments. Si on prend tout simplement l'axiome $\exists x \forall y\, x \neq \emptyset \wedge (y \in x \to y \cup \{y\} \in x)$, ça ne suffit pas, car il se peut que x contienne un seul élément y, qui satisfait $y \in y$. La solution est d'imposer que $\emptyset \in x$. Donc l'axiome désiré est :

$\exists x\, [\emptyset \in x \wedge \forall y\, (y \in x \to y \cup \{y\} \in x)]$.

6.2 L'axiome du choix (AC) – quelques versions équivalentes

Voici cinq axiomes, et nous allons montrer qu'ils sont tous équivalents (modulo la théorie ZF). Par AC nous entendrons n'importe laquelle des 3 premières versions.Comme le nom l'indique, AC dit que étant donné une famille d'ensembles indexée par un ensemble, alors il existe une fonction qui choisit exactement un élément dans chaque ensemble de la famille.

AC1 – Le produit d'une famille d'ensembles non vides est non vide.

AC2 – Pour tout ensemble a non vide, il existe une fonction $f : \mathcal{P}(a) \to a$ telle que si $x \in a$ est non vide, alors $f(x) \in x$.

AC3 – Si a est un ensemble dont tous les éléments sont non vides et deux à deux disjoints (autrement dit, ils forment une partition de $\bigcup_{x\in a} x$), alors il existe un ensemble c tel que pour tout $x \in a$, $x \cap c$ a exactement un élément.

Zorn – Tout ensemble non vide partiellement ordonné et inductif, admet un élément maximal.

Zermelo – Tout ensemble non vide peut être muni d'un bon ordre.

Définition 6.18. Je rappelle qu'un ensemble $(X, \leq)$ partiellement ordonné est *inductif* si X est non vide, et pour tout sous-ensemble $Y \subset X$ qui est totalement ordonné, il existe un *majorant* de Y dans X, c'est à dire un $a \in X$ tel que $y \leq a$ pour tout $y \in Y$.

Avant de commence la preuve des équivalences, voici un lemme et une remarque qui nous seront utiles.

Lemme 6.19. *Soient α un ordinal, S une collection, $\mathcal{F}$ la collection des applications définies sur les ordinaux $\beta < \alpha$ et prenant leurs valeurs dans S, et F une relation fonctionnelle de domaine $\mathcal{F}$, à valeurs dans S. Alors il existe une fonction f (dans $\mathcal{F}$), et une seule, définie sur α, et telle que*

$$f(\beta) = F(f_{|\beta}) \qquad (*)$$

pour tout $\beta < \alpha$.

Démonstration. Unicité. Soient f et g deux fonctions satisfaisant $(*)$. Alors $f(\emptyset) = g(\emptyset)$, car ils sont égaux à $F(\emptyset)$ ($\emptyset^\emptyset = \{\emptyset\}$, donc $f_{|\emptyset} = \emptyset$). On montre par induction sur $\beta < \alpha$ que $f(\beta) = g(\beta)$. Supposons le résultat montré pour tout $\gamma < \beta$. Alors $f_{|\beta} = g_{|\beta}$, et donc

$f(\beta) = g(\beta)$ par $(*)$.
Existence. Soit τ l'ensemble des ordinaux $\beta \in \alpha$ tels qu'il existe $f_\beta \in \mathcal{F}$ définie sur β et telle que $f_\beta(\gamma) = F(f_{\beta|_\gamma})$ pour tout $\gamma < \beta$. (τ est bien un ensemble, car il est définissable et contenu dans α). Alors τ est un segment initial de α, et donc est un ordinal. Par unicité, si $\beta < \gamma < \tau$ alors $f_\beta = f_{\gamma|_\beta}$. On définit f_τ en posant $f_\tau(\beta) = F(f_\beta)$ pour tout $\beta < \tau$. Alors $f_\beta = f_{\tau|_\beta}$. Si $\tau < \alpha$, alors $\tau \in \tau$, ce qui est absurde. Donc $\tau = \alpha$. [Je me rends compte que je n'avais pas fini la preuve en classe, désolée.]

Remarque 6.20. La preuve du résultat précédent se généralise avec α remplacé par la classe On de tous les ordinaux, et $\mathcal{F}$ la collection des applications définies sur des ordinaux et prenant leurs valeurs dans S. On aura une relation fonctionnelle $y = f(\alpha)$ définie par : il existe $f_\alpha \in \mathcal{F}$ de domaine α, et satisfaisant $(*)$ pour tout $\beta < \alpha$.

Nous allons maintenant montrer les équivalences.

6.21. AC2 implique AC3. Soient $b = \bigcup_{x\in a} x$ et f donnée par AC2 appliquée à b. Par ZF4, $c = \{f(x) \mid x \in a\}$ est un ensemble. Puisque tous les éléments de a sont non vides, on sait que pour tout $x \in a$, $f(x) \in x$, et donc $c \cap x$ a au moins un élément. Mais si x, y sont des éléments distincts de a, nous savons que $x \cap y = \emptyset$, ce qui entraine que $f(y) \notin x$, et donc $c \cap x$ a exactement un élément.

6.22. AC3 implique AC1. Soit $(a_i)_{i\in I}$ une famille d'ensembles non vides, et posons $b_i = \{i\} \times a_i$. Les b_i sont des ensembles non vides, et deux à deux disjoints ; chaque b_i est en bijection avec a_i. Soit c un ensemble donné par AC3 appliqué à l'ensemble $\{b_i \mid i \in I\}$. Alors $c \in \prod_{i\in I} a_i$, car c'est le graphe d'une fonction $I \to \bigcup_{i\in I} a_i$ qui envoie chaque $i \in I$ sur un élément de a_i.

6.23. AC1 implique AC2. Soit a un ensemble non vide, et considérons $c = \prod_{x\subset a, \emptyset\neq x} x$, qui est non-vide par AC1. Si $f \in c$, alors f est une fonction de $\{x \subset a \mid x \neq \emptyset\}$ dans a telle que pour tout x dans son domaine, $f(x) \in x$. On étend f à $\mathcal{P}(a)$ tout entier en prenant n'importe quoi pour $f(\emptyset)$.

Nous avons donc montré l'équivalence des 3 premières versions de AC. Maintenant les deux autres.

6.24. Zermelo implique AC2. On prend un bon ordre sur a, et à une partie non vide x de a, on associe le plus petit élément de x. En associant à $\emptyset$ n'importe quel élément de a, on obtient la fonction désirée.

6.25. AC2 implique Zermelo. Soient a non vide, et $f : \mathcal{P}(a) \to a$ donnée par AC2. On définit $g : \mathcal{P}(a) \to a$ en posant $g(x) = f(a \setminus x)$. On a donc $g(x) \notin x$ pour tout $x \subset a$ tel que $x \neq a$. Je note $\mathcal{P}'(a) = \mathcal{P}(a) \setminus \{a\}$.

On prend θ dans $\mathcal{U}$, $\theta \notin a$. (Un tel θ existe puisque a est un ensemble, et $\mathcal{U}$ est une classe qui n'est pas un ensemble, donc contient strictement a). Nous allons construire une relation fonctionnelle F, de domaine On (la classe des ordinaux), et à valeurs dans $a \cup \{\theta\}$. Cette relation fonctionnelle est définie par induction sur les ordinaux de la façon suivante. On pose $F(0) = g(\emptyset) \in a$, et pour un ordinal $\alpha > 0$,

$$F(\alpha) = \begin{cases} g(\{F(\beta) \mid \beta < \alpha\}) & \text{si } \{F(\beta) \mid \beta < \alpha\} \in \mathcal{P}'(a) \\ \theta & \text{sinon.} \end{cases}$$

Nous allons d'abord supposer que pour tout ordinal α, nous avons $F(\alpha) \in a$, et nous obtiendrons une contradiction. Par définition de F, $F(\alpha) \in a$ implique $F(\beta) \in a$ pour tout $\beta < \alpha$, et $\{F(\beta) \mid \beta < \alpha\} \neq a$, d'où on a que $F(\alpha) \notin \{F(\beta) \mid \beta < \alpha\}$ (car $g(x) \notin x$). Cela montre que F définit une fonction injective de la classe On dans l'ensemble a, ce qui est absurde. En effet, comme F est injective sur la classe des ordinaux, la formule $G(x, y)$ qui décrit F^{-1}, c'est-à-dire, $G(x, y) := F(y, x) \wedge \mathrm{On}(x)$ est une relation fonctionnelle (en y), de domaine contenu dans a, et dont l'image est donc un ensemble par ZF4 : mais son image est la classe de tous les ordinaux,

qui n'est pas un ensemble.

Il existe donc un ordinal α tel que $F(\alpha) = \theta$, et on prend le plus petit - appelons-le α_0. Nous avons donc :

Pour tout $\beta < \gamma < \alpha_0$, $F(\beta)$ et $F(\gamma)$ sont distincts, et appartiennent à a. Donc la restriction h de F à α_0 est injective. De plus, nous savons que $\{F(\beta) \mid \beta < \alpha_0\}$, bien que contenu dans a, n'est pas dans $\mathcal{P}'(a)$: cet ensemble est donc égal à a. Nous avons montré qu'il existe une bijection h entre α_0 et a ; nous définissons un bon ordre $<$ sur a de la seule façon possible pour que h soit un isomorphisme d'ensembles ordonnés. (C'est-à-dire : si $x, y \in a$, $x < y$ ssi $h^{-1}(x) < h^{-1}(y)$).

Exercice 6.26. Donnez explicitement la formule qui définit la relation fonctionnelle F définie ci-dessus. Les paramètres qui y apparaissent sont (au moins) a et la fonction g.

6.27. AC2 implique Zorn. Soit $(a, \leq)$ un ensemble ordonné, non vide et inductif. Soit $f : \mathcal{P}(a) \to a$ donnée par AC2 appliqué à a. Si $x \subset a$, on appelle *majorant strict de* x un élément $y \in a$ tel que $z < y$ pour tout $z \in x$. Nous considérons

$$C = \{x \subset a \mid x \text{ a un majorant strict dans } a\}.$$

Alors $\emptyset \in C$ car $a \neq \emptyset$. On définit $m : C \to a$ par

$$m(x) = f(\{y \in a \mid y \text{ est un majorant strict de } x\}).$$

Donc m choisit un majorant strict de x si x en a un.

Soit θ dans $\mathcal{U}$, et $\theta \notin a$. On va définir une relation fonctionnelle $F : \mathrm{On} \to a \cup \{\theta\}$ en posant

$$F(\alpha) = \begin{cases} m(\{F(\beta) \mid \beta < \alpha\}) & \text{si } \{F(\beta) \mid \beta < \alpha\} \in C, \\ \theta & \text{sinon.} \end{cases}$$

On raisonne comme précédemment. Notons que si $\beta < \alpha$ et $F(\alpha) \in a$, alors $F(\beta) \in a$ et $F(\beta) < F(\alpha)$, par définition de F. Si pour tout ordinal α on avait $F(\alpha) \in a$, on aurait donc une injection de la classe On dans l'ensemble a, ce qui est absurde.

Soit α_0 le plus petit ordinal tel que $F(\alpha_0) = \theta$. La restriction g de F à α_0 définit donc un isomorphisme d'ensembles ordonnés entre α_0 et son image $g[\alpha_0] := \{F(\beta) \mid \beta < \alpha_0\}$. Comme $F(\alpha_0) \notin a$, nous savons que $g[\alpha_0] \notin C$, et donc n'a pas de majorant strict. Cependant, $g[\alpha_0]$ est un sous-ensemble de a qui est bien ordonné, et a donc un majorant (disons, d) puisque a est inductif. Ce majorant d n'est pas un majorant strict : cela montre que d est un élément maximal de a.

6.28. Zorn implique AC3. Soit a un ensemble non vide, dont tous les éléments sont non vides et disjoints deux à deux. Posons $b = \bigcup_{x \in a} x$, et

$$X = \{c \subset b \mid \forall x \in a \ |c \cap x| \leq 1\}.$$

(Ici, par $|c \cap x|$, je dénote la cardinalité de $c \cap x$, c'est à dire, combien il a d'éléments). Alors X est ordonné par inclusion. On montre facilement que X est inductif : si $Y \subset X$ est totalement

ordonné par l'inclusion, et consiste d'ensembles y qui intersectent chaque $x \in a$ en au plus un élément, alors aussi $z = \bigcup_{y \in Y} y$ est dans X. Par Zorn, X a donc un élément maximal, notons-le d. Alors pour tout $x \in a$, on a $|d \cap x| \leq 1$. Supposons qu'il existe $x \in a$ tel que $x \cap d = \emptyset$. Prenons alors un élément $u \in x$, et considérons $d_1 = d \cup \{u\}$. Alors $d_1 \in X$, ce qui contredit la maximalité de d. Donc, pour tout $x \in a$ nous avons $d \cap x \neq \emptyset$, et d est l'élément désiré.

Exercice 6.29. On note Z la théorie formée des axiomes ZF1, ZF2, ZF3, de l'axiome de la paire, du schéma de compréhension, et de ZF5. C'est la *théorie de Zermelo*, historiquement la première axiomatisation proposée par Zermelo pour la théorie des ensembles (le schéma de remplacement est dû à Fraenkel). Montrer, dans la théorie Z, que pour tout ensemble x, il existe un ensemble bien ordonné $(y, \prec)$ qui ne peut pas s'injecter dans x (indication : on pourra prendre pour y l'ensemble des bon ordres sur une partie de x, quotienté par la relation d'isomorphisme, que l'on munira de l'ordre $\prec$ défini par $(u, <_u) \prec (v, <_v)$ si et seulement si $(u, <_u)$ est isomorphe à un segment initial strict de $(v, <_v)$). En déduire que l'équivalence entre les différentes formes de l'axiome du choix (AC1, AC2, AC3, Zermelo et Zorn) est démontrable dans Z.

On notera alors ZC la théorie Z à laquelle on a ajouté AC.

Exercice 6.30. Voici une preuve directe de Zermelo à partir de Zorn. Soit a non vide. On considère l'ensemble X consistant de couples $(b, <_b)$, où $b \subset a$, et $<_b$ est un bon ordre sur b. On définit un ordre partiel $\sqsubset$ sur X en posant $(b, <_b) \sqsubset (c, <_c)$ si et seulement si c contient strictement b, l'ordre $<_c$ prolonge l'ordre $<_b$, et b est un segment initial de c.

(1) Montrez que X est bien un ensemble, en exhibant un ensemble qui le contient, et une formule qui le définit.
(2) Montrez que $(X, \sqsubset)$ est inductif.
(3) Montrez qu'un élément maximal de X donne un bon ordre sur a.

6.3 Axiome de fondation, construction de V

Comme toujours, on travaille dans un modèle $\mathcal{U}$ de ZF.

6.31. L'axiome de fondation, AF. C'est l'axiome disant

$$\forall x \, [x \neq \emptyset \to \exists y \, (y \in x \wedge y \cap x = \emptyset)].$$

6.32. Attention. Les français considèrent que l'axiome de fondation ne fait pas partie de ZF, alors que la plupart des autres mathématiciens mettent AF dans ZF.

Conséquences d'AF

(1) Il n'existe pas de suite $(u_n)_{n \in \omega}$ telle que $u_{n+1} \in u_n$ pour tout $n \in \omega$.
(2) Pour tout x, on a $x \notin x$.

(3) Pour tout entier $n > 0$ on ne peut trouver de $\in$-cycle de longueur n, i.e., des ensembles $u_1, \ldots, u_n$ tels que $u_1 \in u_2$, $u_2 \in u_3$, $\ldots$, $u_{n-1} \in u_n$ et $u_n \in u_1$.

Démonstration. (1) S'il existe une telle suite $(u_n)_{n\in\omega}$, on prend $a = \{u_n \mid n \in \omega\}$: on ne peut trouver de $b \in a$ tel que $a \cap b = \emptyset$. En effet $b \in a$ entraine $b = u_n$ pour un n, et donc que $u_{n+1} \in a \cap b$.
(2) et (3) Evidents par (1).

Remarque 6.33. On a une réciproque partielle : dans ZFC (= ZF + AC), la négation de AF entraine l'existence (dans $\mathcal{U}$) d'une suite infinie $(u_n)_{n\in\omega}$ telle que $u_{n+1} \in u_n$ pour tout $n \in \omega$.

Démonstration. Soit $a \in \mathcal{U}$ contredisant AF (et non vide). Alors, pour tout $y \in a$, $y \cap a \neq \emptyset$. Grâce à AC, il existe une fonction $f : a \to a$, qui associe à tout élément $b \in a$ un élément de $b \cap a$. On prend $u_0 \in a$, et on définit par induction (sur les éléments de ω), la suite $u_{n+1} = f(u_n)$, à valeurs dans a. Alors la suite $(u_n)_{n\in\omega}$ ($\subset \omega \times a$) est dans $\mathcal{U}$.

Théorème 6.34. *(AF) Pour que X soit un ordinal, il faut et il suffit que $\in$ définisse sur X un ordre total.*

Démonstration. Puisque $(X, \in)$ est un ordre total, X est donc un ensemble transitif. De plus, AF entraine que cet ordre est bien fondé.

Remarque 6.35. Au contraire des axiomes ZF2 à ZF5 et de AC, l'axiome de fondation n'a pas pour conséquence l'existence d'ensemble souhaîtés, mais plutôt l'inexistence d'ensembles pathologiques. De ce fait, hors théorie des ensembles, il n'est pas utilisé en mathématiques. Il est plutôt là pour donner à l'univers des ensembles une forme "convenable" pour les théoriciens des ensembles.

Ensembles transitifs

Avant d'aller plus loin dans l'étude d'AF, revenons dans ZF et donnons quelques résultats sur les ensembles transitifs. On rappelle qu'un ensemble x est transitif si la relation d'appartenance restreinte à x est transitive, autrement dit si pour tout $y \in x$, on a $y \subseteq x$. On a les propriétés suivantes :

6.36. Propriétés

(1) Une réunion d'ensembles transitifs est transitive.
(2) Si x est transitif, alors $\mathcal{P}(x)$ l'est également.

Démonstration. (1) Si les x_i sont transitifs et si $y \in x = \bigcup_{i\in I} x_i$, alors $y \in x_i$ pour un certain $i \in I$, donc $y \subseteq x_i$, donc $y \subseteq x$.

(2) Si $y \in \mathcal{P}(x)$, alors $y \subseteq x$; pour tout $z \in y$, on a alors $z \in x$, donc $z \subseteq x$ par transitivité, donc $z \in \mathcal{P}(x)$. Ceci montre que $y \subseteq \mathcal{P}(x)$.

Proposition 6.37. *Soit x un ensemble. Alors il existe un unique ensemble transitif y contenant x et qui est contenu dans tout ensemble transitif contenant x. On l'appelle la clôture transitive de x.*

Démonstration. Définissons par récurrence sur ω une suite $(x_n)_{n<\omega}$ par $x_0 = x$ et $x_{n+1} = \bigcup x_n$ (réunion des éléments de x_n). On pose $y = \bigcup_{n<\omega} x_n$ et on montre que y convient. C'est bien un ensemble transitif: si $z \in y$, alors $z \in x_n$ pour un certain n, donc $z \subseteq x_{n+1} \subseteq y$. De plus, si t est un ensemble transitif contenant x, alors on montre par récurrence sur n qu'il contient x_n : si $x_n \subseteq t$ et si $z \in x_{n+1}$, alors il existe $u \in x_n$ tel que $z \in u$; on a $u \in t$, donc par transitivité $u \subseteq t$, et donc $z \in t$. Ceci montre que $x_{n+1} \subseteq t$.

On rappelle également le résultat suivant, dont la preuve à été vue au TD 1 :

Théorème 6.38 (Collapse de Mostowski)**.** *Soit $(P, \lhd)$ un ensemble muni d'une relation bien-fondée et extenstionnelle, c'est-à-dire que $(P, \lhd)$ satisfait l'axiome d'extensionnalité. Alors il existe un unique couple (x, π) où x est un ensemble transitif et π un isomorphisme entre $(P, \lhd)$ et x muni de la restriction de la relation d'appartenance.*

Remarque 6.39. Si la relation $\lhd$ est un bon ordre, on retrouve immédiatement le théorème disant que tout bon-ordre est isomorphe à un ordinal. Mais ce théorème est égalemen très utile sous $ZF + AF$, lorsque P est un ensemble transitif muni de la restriction de la relation d'appartenance.

La hiérarchie de Von Neumann

6.40. Définition de V. On définit une relation fonctionnelle $y = V_\alpha$, pour α un ordinal, de la façon suivante :

$$V_0 = \emptyset; \qquad V_\alpha = \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta) \text{ si } \alpha > 0.$$

On a donc $V_{\alpha+1} = V_\alpha \cup \mathcal{P}(V_\alpha)$, et si α est un ordinal limite, alors $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$. Cela montre que si $\beta < \alpha$, alors V_β est un sous-ensemble (strict) de V_α. On prend V la collection qui est réunion des V_α, pour α dans On. V est donc défini par la formule (en x) $\exists \alpha, (\mathrm{On}(\alpha) \wedge x \in V_\alpha)$.

Exercice 6.41. Donnez explicitement la formule qui définit la relation fonctionnelle $y = V_\alpha$ dans $\mathcal{U}$.

Remarque 6.42. Les propriétés 6.36 permettent de montrer, par une récurrence immédiate sur α, que chaque V_α est transitif. Et donc $V_{\alpha+1} = \mathcal{P}(V_\alpha)$.

Définition 6.43. Soit a dans V. Le rang de a, $rg(a)$, est le plus petit ordinal α tel que $a \in V_\alpha$. Notez que $rg(a)$ est toujours un ordinal successeur.

Lemme 6.44. *Soit a un ensemble.*

(1) *a est dans V si et seulement si tous ses éléments sont dans V.*

(2) *Si a est dans V et $b \in a$, alors $rg(b) < rg(a)$.*
(3) $\mathrm{On} \subset V$ *et si α est un ordinal, alors $rg(\alpha) = \alpha + 1$.*

Démonstration. (1) Soit a dans V, de rang $\beta + 1$. Donc $a \subset V_\beta$, et ses éléments sont dans V. Réciproquement, supposons que tous les éléments de a soient dans V. Alors (par remplacement), il existe α tel que tous les éléments de a sont dans V_α (on prend $\alpha = \bigcup_{x \in a} rg(x)$). Cela entraine que $a \subset V_\alpha$, et $a \in \mathcal{P}(V_\alpha) \subset V_{\alpha+1}$.

(2) Si $\beta + 1 = rg(a)$, alors $a \subset V_\beta$, et donc pour tout $b \in a$, on aura $rg(b) \leq \beta < rg(a)$.

(3) Par induction sur α. On sait que $0 = \emptyset$ est dans V et son rang est 1. Soit α le plus petit ordinal (s'il existe) tel que $\alpha \notin V_{\alpha+1}$. Par définition de α, si $\beta \in \alpha$ alors $\beta \in V_{\beta+1}$, d'où

$$\alpha = \{\beta \mid \beta \in \alpha\} \subset \bigcup_{\beta<\alpha} V_{\beta+1} = \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta) = V_\alpha,$$

et donc $\alpha \in V_{\alpha+1}$.

Il faut aussi montrer qu'on ne peut avoir $rg(\alpha) \leq \alpha$: sinon, soit α le plus petit ordinal (s'il existe) tel que $\alpha \in V_\alpha$. Ce n'est pas 0. Donc $\alpha \in \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta)$, et il existe $\beta < \alpha$ tel que $\alpha \in \mathcal{P}(V_\beta)$, i.e., $\alpha \subset V_\beta$, et comme $\beta \in \alpha$, on a $\beta \in V_\beta$, ce qui contredit la minimalité de α.

Théorème 6.45. *Pour que l'axiome de fondation soit satisfait, il faut et il suffit que V soit l'univers tout entier, c'est à dire, $\mathcal{U} \models \forall x\, V(x)$.*

Démonstration. Rappel: la formule $V(x)$ est une abbréviation pour la formule $\exists \alpha\, \mathrm{On}(\alpha) \wedge x \in V_\alpha$). Supposons $\mathcal{U} \models \forall x\, V(x)$. Soit a non vide dans $\mathcal{U}$, et $b \in a$ de rang minimal. Alors $a \cap b = \emptyset$, car les éléments de b sont de rang strictement inférieur à celui de b.

Pour l'autre direction, on suppose que l'axiome de fondation est vérifié dans $\mathcal{U}$, mais qu'il existe un ensemble a dans $\mathcal{U}$ qui n'est pas dans V. Soit b transitif contenant a, et posons $b' = \{x \in b \mid \neg V(x)\}$. Alors $b' \neq \emptyset$: comme a n'est pas dans V, il existe $x \in a$ qui n'est pas dans V (par le petit Lemme 6.44). Un tel x sera dans b puisque $a \subset b$.

De plus, si $x \in b'$, alors $x \cap b' \neq \emptyset$. En effet, si $x \in b'$, alors il existe $y \in x$ qui n'est pas dans V, et donc ce y est dans b', i.e., dans $b' \cap x$. (Ici on a utilisé que b est transitif). L'ensemble b' contredit AF.

Introduction aux preuves de consistance relative

6.46. Le second théorème d'incomplétude en théorie des ensembles. En théorie des ensembles aussi, on a une version du second théorème d'incomplétude de Gödel. Introduire les notions nécessaires à son énonciation est presque aussi difficile que le démontrer, mais je vais tenter d'en donner une idée en quelques mots. À l'intérieur d'un univers $\mathcal{U}$ satisfaisant ZF, on peut définir, comme on l'a fait au chapitre 2, des notions de formules, preuves formelles, structures et modèles. La théorie ZF peut alors être vue comme un ensemble de ces formules "internes à l'univers $\mathcal{U}$". On dira qu'une théorie interne T (un ensemble de ces formules internes)

dans le langage des ensembles, est consistante si il n'existe pas de preuve formelle de $0 \neq 0$ dans cette théorie, ou encore si cette théorie possède un modèle (un modèle, ici, est un élément M de l'univers $\mathcal{U}$ muni d'une relation binaire $\varepsilon \subseteq M^2$). On notera alors $\mathrm{Coh}(T)$ l'énoncé interne qui, interprété dans une structure (M, ε), dit que la théorie T *écrite avec les formules internes à* M est consistante, c'est à dire qu'il existe des objets N et ε' de M tels que (N, ε') soit modèle de T *pour la théorie des modèles développée dans* (M, ε). (Attention, il y a une subtilité : même le fait de "transporter" la théorie T, qui est a priori interne à $\mathcal{U}$, dans (M, ε), pour en faire une théorie interne à ce modèle, n'est pas évident. Cela ne peut a priori être fait que si T est, en tant qu'objet de l'univers $\mathcal{U}$, définissable par une formule sans paramètres, ce qui est heureusement le cas pour les théories usuelles comme ZF et ZFC.) Le second théorème d'incomplétude dit alors que si T est une théorie (interne à $\mathcal{U}$) qui est récursivement énumerable (pour un bon codage des formules dans l'arithmétique), qui est consistante et qui contient ZF, alors il n'existe pas de preuve formelle de $\mathrm{Coh}(T)$ dans T.

Une conséquence importante de ce théorème est qu'il est impossible de prouver la consistance de ZFC : en effet, comme toutes les mathématiques telles qu'actuellement pratiquées sont formalisables dans tout univers satisfaisant ZFC, alors on pourrait "transporter" cette preuve à l'intérieur d'un tel univers, ce qui montrerait que dans tout univers, on a $ZFC \vdash \mathrm{Coh}(ZFC)$, contredisant ainsi le second théorème d'incomplétude.

En théorie des ensembles, on se contentera donc de prouver des résultats de *consistance relative*, qui sont des résultats du type "Si une certaine théorie T est consistante, alors une autre théorie T' l'est également". Le but principal du reste de cette section est par exemple de montrer que si ZFC est consistante, alors $ZFC + AF$ l'est également. La méthode utilisée pour montrer ce genre de résultats est la suivante : on considère un univers $\mathcal{U}$ qui est modèle de T, et à partir de $\mathcal{U}$, on construit un univers $\mathcal{M}$ modèle de T'. Dans la plupart des cas, $\mathcal{M}$ sera une classe transitive M de $\mathcal{U}$ munie de la restriction de la relation d'appartenance de $\mathcal{U}$. Dans la suite, quand je dirai qu'une classe transitive M satisfait un axiome ou une théorie sans préciser quelle est la relation d'appartenance associée, M sera toujours munie de la restriction de la relation d'appartenance de $\mathcal{U}$. La définition suivante permet de ramener la satisfaction d'une formule dans M à celle d'une autre formule dans $\mathcal{U}$.

Définition 6.47 (Relativisation des formules)**.** Rappel: si $S \subset \mathcal{U}$ est une collection, définie par la formule $S(x)$, alors on définit par induction sur la complexite des formules du langage, les relativisées à S. Si φ est atomique, alors $\varphi^S = \varphi$; $(\varphi \wedge \psi)^S = \varphi^S \wedge \psi^S$; $(\neg\varphi)^S = \neg\varphi^S$; et enfin $(\exists x\, \varphi^S) = \exists x\, (S(x) \wedge \varphi^S)$, $(\forall x\, \varphi)^S = \forall x\, (S(x) \to \varphi^S)$. Les deux dernières formules sont souvent abbrégées par $\exists x \in S\, \varphi^S$, et $\forall x \in S\, \varphi^S$. Autrement dit, on relativise tous les quantificateurs à S.

Exercice 6.48. Nous utiliserons le résultats de cet exercice. Il est facile à montrer, faites-le. ($\mathcal{L} = \{\in\}$,$\mathcal{U} \models$ ZF) Soient S définissable dans $\mathcal{U}$ et non vide, $\varphi(\bar{x})$ une formule, et $\bar{a}$ un uplet d'éléments de S. Montrez que

$$\mathcal{U} \models \varphi^S(\bar{a}) \text{ ssi } S \models \varphi(\bar{a}).$$

Cet exercice utilise le fait que le sous-ensemble défini par S est une sous-structure de $\mathcal{U}$, car le langage est relationnel (i.e., n'a aucun symbole de constante ou de fonction).

Définition 6.49. Soit S une collection défine par $S(x)$. Une formule $\varphi(\bar{y})$ est *absolue* pour S (ou bien φ *se reflète dans* S) si pour tout uplet $\bar{a}$ de S, on a $\mathcal{U} \models \varphi(\bar{a})$ ssi $S \models \varphi(\bar{a})$. Autrement dit, si

$$\mathcal{U} \models \forall \bar{y} \, [\bigwedge_{i<|\bar{y}|} S(y_i) \to (\varphi(\bar{y}) \leftrightarrow \varphi^S(\bar{y})).$$

Définition 6.50. (1) L'ensemble des formules Δ_0 est le plus petit ensemble de formules contenant les formules atomiques, qui est clos par combinaisons booléennes et par *quantification bornée* : si $\varphi(x, y, \bar{z})$ est une formule Δ_0, alors aussi $\exists x \, (x \in y \wedge \varphi(x, y, \bar{z}))$, et $\forall x \, (x \in y \to \varphi(x, y, \bar{z}))$.

(2) L'ensemble des formules Σ_1 est le plus petit ensemble de formules contenant les formules atomiques et négatomiques, et qui est clos par conjonction, disjonction, quantification universelle bornée et quantification existentielle.

(3) L'ensemble des formules Π_1 est le plus petit ensemble de formules contenant les formules atomiques et négatomiques, et qui est clos par conjonction, disjonction, quantification universelle et quantification existentielle bornée.

Le lemme suivant est alors un analogue des lemmes de préservation vus en cours de théorie des modèles.

Lemme 6.51. *Soit S une collection transitive.*

(1) *Toute formule Δ_0 est absolue pour S.*
(2) *Si φ est une formule Σ_1, alors pour tout uplet $\bar{a}$ de S, si $S \models \varphi(\bar{a})$, alors $\mathcal{U} \models \varphi(\bar{a})$.*
(3) *Si φ est une formule Π_1, alors pour tout uplet $\bar{a}$ de S, si $\mathcal{U} \models \varphi(\bar{a})$, alors $S \models \varphi(\bar{a})$.*

Démonstration. (1) Par induction sur la complexité des formules. C'est clair pour les formules atomiques (qui sont absolues pour n'importe quel ensemble définissable), et la propriété d'être absolu est préservée par combinaisons booléennes. Il faut vérifier que si $\varphi(x, y, \bar{z})$ est absolue pour S, alors aussi $\exists x \in y \, \varphi(x, y, \bar{z})$. Soient $b, \bar{c}$ dans S. On a : $\mathcal{U} \models \exists x \in b \, \varphi(x, b, \bar{c})$ ssi il existe a tel que $\mathcal{U} \models a \in b \wedge \varphi(a, b, \bar{c})$, ssi il existe $a \in S$ tel que $\mathcal{U} \models a \in b \wedge \varphi(a, b, \bar{c})$ (transitivité de S), ssi $S \models a \in b \wedge \varphi(a, b, \bar{c})$ (HI), ssi $S \models \exists x \in b \, \varphi(x, b, \bar{c})$.

(2) Il suffit de montrer que si le résultat est vrai pour φ, alors il l'est pour $\exists x \, \varphi$, et le reste a déjà été fait dans la preuve de (1). Mais si $S \models \exists x \, \varphi(\bar{a})$, alors on peut trouver $b \in S$ tel que $M \models \varphi(b, \bar{a})$, donc $\mathcal{U} \models \varphi(b, \bar{a})$ par hypothèse d'induction, donc $\mathcal{U} \models \exists x \, \varphi(\bar{a})$.

(3) Immédiat à partir de (2) en utilisant le fait que la négation d'une formule Π_1 est logiquement équivalente à une formule Σ_1.

Lemme 6.52. *Les propriétés suivantes sont exprimables par des formules Δ_0 : $x \subset y$; $x = \emptyset$; $x = y \cup \{y\}$; x est transitif ; $z = \{x, y\}$; $y = \bigcup_{z \in x} z$.*

Démonstration. Elles s'écrivent respectivement :
$\forall z \in x \, (z \in y)$

$\forall z \in x\,(z \neq z)$
$\forall z \in x\,(z = y \vee z \in y) \wedge y \subset x \wedge y \in x$
$\forall z \in x\,(\forall y \in z\,(y \in x))$
$\forall u \in z\,(u = x \vee u = y) \wedge x \in z \wedge y \in z$
$\forall u \in y\,(\exists z \in x\,(u \in z) \wedge \forall z \in x(z \subset y)$

Exercice 6.53. Montrer que la formule $\mathrm{On}(x)$ est Π_1. Montrer que modulo $ZF + AF$, elle est équivalente à une formule Δ_0. En déduire que si $\mathcal{U}$ est un univers modèle de $ZF + AF$ et si M est une classe transitive de $\mathcal{U}$ qui est également modèle de $ZF + AF$, alors on aura $\mathrm{On}^M = \mathrm{On} \cap M$.

Proposition 6.54. *Soient $\mathcal{U} \models ZF$ et S une classe transitive non vide de $\mathcal{U}$.*

(1) *S satisfait ZF1.*
(2) *Si pour tout $x \in S$, on a $\bigcup x \in S$, alors S satisfait ZF2.*
(3) *Si pour tous $x, y \in S$, on a $\{x, y\} \in S$, alors S satisfait l'axiome de la paire.*
(4) *On suppose que $\mathcal{U} \models \forall x \in S \exists y \in S\,(\mathcal{P}(x) \cap S = y)$. Alors S satisfait ZF3.*
(5) *Si pour tout $x \in S$, on a $\mathcal{P}(x) \subseteq S$, alors S satisfait le schéma de compréhension.*
(6) *Si pour toute relation fonctionnelle F de domaine et d'image contenus dans S, et pour tout $a \in S$, on a $\{F(x) \mid x \in a\} \in S$, alors S satisfait ZF4.*
(7) *Si $\omega \in S$, alors S satisfait ZF5.*
(8) *Si $\mathcal{U}$ satisfait AC et si pour tout $x \in S$, on a $\mathcal{P}\left(\bigcup x\right) \subseteq S$, alors S satisfait AC.*

Démonstration. (1) ZF1 est Π_1.

(2) La formule $y = \bigcup x$ étant absolue pour S par les deux lemmes précédents, on en déduit que la relativisation de ZF2 à S est équivalente à l'énoncé $\forall x \in S\, \exists y \in S\ y = \bigcup x$, d'où le résultat.

(3) Exactement pareil que (2).

(4) $\mathcal{U}$ satisfait donc : $\forall x \in S \exists y \in S \forall z \in S(z \subset x \leftrightarrow x \in y)$, ce qui est la version relativisée à S de l'axiome des parties.

(5) Soit φ une formule à une variable libre et à paramètres dans S. La formule $y \subseteq x$ étant absolue pour S, la relativisation à S de l'instance du schéma de compréhension associée à φ s'écrit $\forall x \in S\, \exists y \in S\ (y \subseteq x \wedge \forall z \in x\ (z \in y \Leftrightarrow \varphi^S(z)))$. L'ensemble y dont on a besoin existe par le schéma de compréhension appliqué à φ^S dans $\mathcal{U}$, il suffit alors de montrer que $y \in S$. Mais comme $y \subseteq x$, ceci est vrai si $\mathcal{P}(x) \subseteq S$.

(6) Soit $\varphi(x, y)$ une formule à deux variables libres et à paramètres dans S telle que $S \models$ "$\varphi(x, y)$ définit une relation fonctionnelle". Ceci s'écrit, dans $\mathcal{U}$, $\forall x, y, y' \in S\ (\varphi^S(x, y) \wedge \varphi^S(x, y') \Leftrightarrow y = y')$. En particulier, la formule $x \in S \wedge y \in S \wedge \varphi^S(x, y)$ définit une relation fonctionnelle de domaine et d'image inclus dans S, que l'on notera $y = F(x)$. Soit maintenant $a \in S$; on cherche $b \in S$ tel que $S \models \forall y\ (y \in b \Leftrightarrow \exists x \in a\ \varphi(x, y))$, c'est-à-dire tel que $\forall y \in S\ (y \in b \Leftrightarrow \exists x \in a\ y = F(x))$. Par ZF4 appliqué dans $\mathcal{U}$, $\{F(x) \mid x \in a\}$ est un ensemble

que l'on notera b, et par l'hypothèse, cet ensemble appartient à S. Il vérifie clairement la propriété voulue.

(7) L'ordinal ω satisfait la formule $\varnothing \in x \wedge \forall y \in x\ (y \cup \{y\} \in x)$, et par le lemme précédent cette formule est absolue pour S. ω la satisfait donc aussi dans S, ce qui montre que $ZF5$ est satisfait.

(8) On vérifie facilement que les formules "x est un ensemble d'ensembles non-vides deux-à-deux disjoints" et "l'intersection de y et de tout élément de x est un singleton" sont Δ_0, donc absolues pour S. L'axiome de choix relativisé à S (dans sa forme AC3) est donc équivalent à dire que pour tout $x \in S$ ensemble d'ensemble non-vides deux-à-deux disjoints, il existe $y \in S$ dont l'intersection avec tout élément de x est un singleton. Mais si $x \in S$, alors par l'axiome de choix dans $\mathcal{U}$, l'ensemble y recherché existe dans $\mathcal{U}$, il suffit donc de montrer qu'il est dans S. C'est le cas puisque c'est un élément de $\mathcal{P}(\bigcup x)$.

Proposition 6.55. *Soient $\mathcal{U} \models ZF$ et α un ordinal.*

(1) *Si $x, y \in V_\alpha$, alors $\bigcup x \in V_\alpha$, $\mathcal{P}(x) \in V_{\alpha+1}$ et $\{x, y\} \in V_{\alpha+1}$.*
(2) *En particulier, V_α satisfait ZF1 et ZF2, et si α est limite, alors V_α satisfait de plus l'axiome de la paire, ZF3, le schéma de compréhension, et l'axiome de choix s'il est satisfait par $\mathcal{U}$. De même, la collection V satisfait ZF1, ZF2, l'axiome de la paire, ZF3, le schéma de compréhension, et l'axiome de choix s'il est satisfait par $\mathcal{U}$.*

Démonstration. (1) Si $x \in V_\alpha$, alors $x \subseteq V_\beta$ pour un certain $\beta < \alpha$, et donc par transitivité de V_β, on a $\bigcup x \subseteq V_\beta$, donc $\bigcup x \in V_\alpha$. De même, pour tout $z \subseteq x$, on a $z \subseteq V_\beta$, donc $z \in V_\alpha$, et donc $\mathcal{P}(x) \subseteq V_\alpha$ et $\mathcal{P}(x) \in V_{\alpha+1}$. Le fait que $\{x, y\} \in V_{\alpha+1}$ est quant à lui immédiat par la définition de $V_{\alpha+1}$.

(2) Le fait que V_α et V satisfont ZF1 et ZF2 découle de la proposition 6.54 et du fait que si $x \in V_\alpha$ (resp. $x \in V$), alors $\bigcup x \in V_\alpha$ (resp. $\bigcup x \in V$). Le reste découle de la même proposition et du fait, conséquence de (1), que si α est limite et si $x, y \in V_\alpha$ (resp. $x, y \in V$), alors $\cup x$, $\mathcal{P}(x)$ et $\{x, y\}$ sont des éléments et des parties de V_α (resp. de V).

Théorème 6.56. *Soit $\mathcal{U} \models ZF$ (resp. $\mathcal{U} \models ZFC$). Alors la collection V de $\mathcal{U}$ est modèle de ZF + AF (resp. de ZFC + AF). En particulier, si ZF est consistant, alors ZF + AF l'est également, et si ZFC est consistant, alors ZFC + AF l'est également.*

Démonstration. L'essentiel a déjà été fait dans la proposition précédente, il ne reste plus qu'à montrer que V satisfait ZF4 et AF. Commençons par ZF4. Par la proposition 6.54, il suffit de montrer qu'étant donné F une relation fonctionnelle à domaine et image dans V, pour tout $a \in V$, on a $\{F(x) \mid x \in a\} \in V$. Mais ceci est immédiat par le lemme 6.44, qui dit qu'un ensemble dont tous les éléments sont dans V est lui-même dans V.

En ce qui concerne AF, le relativisé à V de cet axiome dit que pour tout $x \in V$, il existe $y \in x$ tel que $x \cap y = \varnothing$; il est satisfait, il suffit en effet de prendre $y \in x$ de rang minimal, comme dans la preuve du théorème 6.45.

Exercice 6.57. Rappel: la formule $\mathrm{On}(x)$ est la formule exprimant que $\in$ définit un ordre total sur x, et que cet ordre est bien fondé. Montrez que si $\alpha \in V$, alors $V \models \mathrm{On}(\alpha)$ si et seulement si $\mathcal{U} \models \mathrm{On}(\alpha)$.

Remarque 6.58. La preuve de la consistance relative de AF qu'on vient de donner est sémantique : elle fait appel à des modèles des théories considérées. Il peut être intéressant d'en donner une preuve syntaxique (ne se basant que sur la manipulation de preuves formelles), de façon à pouvoir l'exprimer dans une métathéorie la plus simple possible (ici, comme on fait référence aux modèles, il faut une théorie permettant de manipuler des ensembles, alors que pour faire des preuves syntaxiques, une arithmétique faible suffit). Je vais expliquer comment en obtenir une.

Si on analyse la preuve qu'on vient de faire, ce qu'on vient de montrer, c'est en fait que pour tout axiome φ de ZFC + AF, on a $\mathrm{ZFC} \vdash \varphi^V$ (et il n'y a pas besoin de référence aux modèles pour montrer cela). Or, si ZFC + AF était inconsistente, alors il existerait un nombre fini d'énoncés $\varphi_1, \ldots, \varphi_n$ de cette théorie à partir desquels on peut obtenir une preuve formelle de $0 \neq 0$. Cette preuve formelle est une suite $(\psi_1, \ldots, \psi_k)$ de formules. On vérifie alors (par induction sur la complexité de la formule définissant V) que $(\psi_1^V, \ldots, \psi_k^V)$ est une preuve formelle de $(0 \neq 0)^V$, c'est-à-dire de $0 \neq 0$, à partir des formules $\varphi_1^V, \ldots, \varphi_n^V$. En combinant cette preuve avec des preuves formelles des φ_i^V dans ZFC, on obtient une preuve, dans ZFC, de $0 = 0$, ce qui conclut.

De façon plus générale, cette méthode permet de transformer toutes les preuves sémantiques de résultats de consistance relative en preuves syntaxiques.

Théorème 6.59. *V_ω satisfait ZF1 – ZF4, ainsi que la négation de ZF5.*

Démonstration. On sait déjà qu'il satisfait ZF1, ZF2 et ZF3. Pour ZF4, nous remarquons d'abord que tout élément de V_ω est fini. En effet, $V_0 = \emptyset$, et il suit, par induction sur n, que chaque V_n est fini pour $n \in \mathbb{N}$. Si $a \in V_\omega$, alors $a \in V_n$ pour un entier n, et donc $a \subset V_n$ (par transitivité de V_n), ce qui implique que a est fini. Soient alors F une relation fonctionnelle à domaine et image dans V_ω et $a \in V_\omega$. Comme a est fini, l'ensemble $\{\mathrm{rg}(F(x)) \mid x \in a\}$ est un ensemble fini d'ordinaux finis, donc sa borne supérieure, que l'on notera n, est finie. On a alors $b = \{F(x) \mid x \in a\} \subseteq V_n$, donc $b \in V_{n+1} \subseteq V_\omega$.

D'autre part, puisque tout élément de V_ω est fini, V_ω ne satisfait pas ZF5 : tout ordinal non nul de V_ω est un successeur.

Exercice 6.60. Montrer que V_ω satisfait AC même si l'univers de départ ne le satisfait pas.

Définition 6.61. On travaille dans ZFC, avec des cardinaux, λ, μ etc.

(1) λ est *fortement limite* si pour tout $\mu < \lambda$, on a $2^\mu < \lambda$.
(2) (Rappel) λ est *régulier* si aucun ensemble $X \subset \lambda$ de cardinalité inférieure à λ n'est cofinal dans λ.
(3) λ est *inaccessible* si λ est fortement limite et régulier, et $\lambda > \aleph_0$.

Exemple 6.62. (1) $\aleph_0$ est fortement limite et régulier.
(2) On pose $\beth_0 = \aleph_0$, $\beth_{n+1} = 2^{\beth_n}$, et $\beth_\omega = \bigcup_{n\in\omega} \beth_n$. Alors $\beth_\omega$ est fortement limite. Mais il n'est pas régulier ... quelle est sa cofinalité ?

Théorème 6.63. *Si ZFC est consistant, alors ZFC + "Il n'existe pas de cardinal inaccessible" est consistant.*

Démonstration. Soit $\mathcal{U}$ un modèle de ZFC. S'il contient un cardinal inaccessible, on prend le plus petit, κ. (Et s'il n'en contient pas on n'a rien à faire.) On va montrer que V_κ est un modèle de ZFC, et qu'il ne contient pas de cardinal inaccessible. Nous savons déjà que V_κ satisfait ZF1, ZF2, ZF3 et AC.

Pour montrer ZF4, nous allons d'abord montrer que si $a \in V_\kappa$, alors $|a| < \kappa$ ($|a|$ dénote la cardinalité de a). Comme κ est limite, il existe $\alpha < \kappa$ tel que $a \in V_\alpha$, d'où $a \subset V_\alpha$, et il suffit donc de montrer que si $\alpha < \kappa$, alors $|V_\alpha| < \kappa$. Par induction sur α, et pour $\alpha = 0$ c'est clair. Si $\alpha = \beta + 1$ alors $V_\alpha = 2^{V_\beta}$ et donc $|V_\alpha| = 2^{|V_\beta|}$, est $< \kappa$ par IH et puisque κ est fortement limite. Si α est limite, alors V_α est une union indexée par un ordinal $< \kappa$ d'ensembles de cardinalité $< \kappa$, et est donc de cardinalité $< \kappa$ car κ est régulier.

Le reste de la preuve est faite comme pour V_ω, en remplaçant "fini" par "de cardinalité $< \kappa$". Soient F une relation fonctionnelle de domaine et d'image dans V_κ, et $a \in V_\kappa$. Il faut montrer que $b = \{F(x) \mid x \in a\}$ est dans V_κ. Nous savons que c'est un ensemble (par ZF4 pour $\mathcal{U}$). Nous savons que les éléments de Y sont dans V_κ, et que $|Y| < \kappa$ (puisque $|a| < \kappa$) ; par régularité de κ, il existe $\beta < \kappa$ tel que $Y \subset V_\beta$, et nous avons alors $Y \in V_{\beta+1}$.

Il est clair je pense que $\omega \in V_\kappa$, et donc V_κ satisfait ZF5.

Montrons enfin que V_κ ne contient pas de cardinal inaccessible. Nous savons que les ordinaux de V_κ sont $< \kappa$. S'il en contenait un, nous aurions donc un $\lambda < \kappa$ qui satisferait (dans V_κ) les formules suivantes :
$\forall \mu < \lambda,\ 2^\mu < \lambda$
$\lambda > \aleph_0$
$\mathrm{cf}(\lambda) = \lambda$
La première formule est clairement Δ_0, et sera donc satisfaite dans $\mathcal{U}$ aussi (puisque V_κ est transitif). Pour la deuxième, supposons λ dénombrable (au sens de $\mathcal{U}$) ; alors il existe une surjection $h : \omega \to \lambda$, $h \in \mathcal{U}$; mais comme précédemment, cet h est dans V_κ, et donc λ serait dénombrable au sens de V_κ, contradiction. Enfin, la troisième propriété s'exprime en disant : si $Y \subset \lambda$ est cofinal dans λ, alors $|Y| = \lambda$. Mais un tel Y est de rang $\lambda + 2$, et comme $\mathcal{P}(\lambda) \subset V_\kappa$, on obtient le résultat (comme ci-dessus, la bijection entre λ et Y est dans $V_{\lambda+3}$).

Théorème 6.64. *(Schéma de reflexion, en supposant ZF+AF). Pour toute formule $\varphi(\bar{v})$ ($\bar{v} = (v_1, \ldots, v_n)$), on a*

$$\mathcal{U} \models \forall \bar{v} \exists \alpha \, \mathrm{On}(\alpha) \wedge \bigwedge_i v_i \in V_\alpha \wedge (\varphi(\bar{v}) \leftrightarrow \varphi^{V_\alpha}(\bar{v})).$$

Nous aurons besoin de plusieurs petits résultats pour montrer cela. Tout d'abord

Lemme 6.65. *Soient $\varphi(\bar{v})$ une formule, et $(X_n \mid n \in \omega)$ une suite croissante d'ensembles (ou de classes). On suppose que pour tout $n \in \omega$, φ et toutes ses sous-formules sont absolues pour X_n. Alors φ est absolue pour $X = \bigcup_{n\in\omega} X_n$.*

Démonstration. La démonstration est par induction sur la complexité de φ. Si φ est atomique, alors c'est clair, puisqu'une formule sans quantificateurs est absolue pour n'importe quelle classe/ensemble. De même, si φ_1 et φ_2 sont absolues pour X, alors aussi $\varphi_1 \wedge \varphi_2$ et $\neg\varphi_1$.

Il reste le cas du quantificateur, supposons $\varphi(\bar{v}) = \exists x \, \psi(x, \bar{v})$. Notre hypothèse dit que $\psi(x, \bar{v})$ (et toutes ses sous-formules) est absolue pour chaque X_n. Par HI, cela entraine que $\psi(x, \bar{v})$ est absolue pour X. Soit $\bar{a}$ dans X, et supposons d'abord que $\mathcal{U} \models \exists x \, \psi(x, \bar{a})$; comme $\varphi(\bar{v})$ est absolue pour X_m, nous avons $X_m \models \exists x \, \psi(x, \bar{a})$; soit $b \in X_m$ tel que $X_m \models \psi(b, \bar{a})$; par HI, on a $X \models \psi(b, \bar{a})$, ce qui montre $X \models \varphi(\bar{a})$. L'autre direction est similaire, sauf qu'on prend m tel que b et $\bar{a}$ sont dans X_m.

6.66. *Démonstration du schéma de reflexion 6.64.* Nous allons montrer quelque chose d'un peu plus fort.
($*$) Soient φ une formule, β un ordinal. Il existe $\alpha > \beta$ tel que φ et toutes ses sous-formules sont absolues pour V_α.
Démonstration. La démonstration est par induction sur la complexité de φ. Si φ est atomique, il n'y a rien à faire ; si $\varphi = \neg\psi$, il n'y a rien à faire non plus, car φ est absolue pour un ensemble ssi ψ est absolue pour ce même ensemble.

Déjà le cas de $\varphi_1 \wedge \varphi_2$ est un peu délicat, à cause des sous-formules. On prend $\alpha_0 = \beta$, et on construit par induction sur $n \in \omega$ une suite strictement croissante α_n d'ordinaux. Supposons α_{n-1} construit. Si n est impair, on prend pour α_n le plus petit ordinal $> \alpha_{n-1}$ tel que φ_1 et toutes ses sous-formules sont absolues pour V_{α_n} (un tel α_n existe par HI). Et si n est pair, on prend pour α_n le plus petit ordinal $> \alpha_{n-1}$ tel que φ_2 et toutes ses sous-formules sont absolues pour V_{α_n}. On pose

$$\alpha = \bigcup_{n\in\omega} \alpha_{2n} = \bigcup_{n\in\omega} \alpha_{2n+1}.$$

Alors le lemme précédent nous donne le résultat désiré : φ_1 et toutes ses sous-formules, φ_2 et toutes ses sous-formules, sont absolues pour V_α, et $\alpha > \beta$. Donc aussi $\varphi_1 \wedge \varphi_2$ et toutes ses sous-formules sont absolues pour V_α.

Passons maintenant au cas où $\varphi = \exists x \, \psi(x, \bar{v})$. Nous allons d'abord montrer la chose suivante :

Assertion. Si γ est un ordinal, alors il existe un ordinal δ tel que pour tout $\bar{a}$ dans V_γ, s'il

existe b (dans $\mathcal{U} = V$) tel que $\mathcal{U} \models \psi(b, \bar{a})$, alors il existe un tel b dans V_δ.
On applique le schéma de remplacement à la fonction qui à un uplet $\bar{a}$ de V_γ associe le plus petit $\alpha \geq \gamma$ tel que V_α contienne un b satisfaisant $\varphi(b, \bar{a})$ s'il en existe un dans V, et γ sinon. Nous avons donc un (vrai) ensemble d'ordinaux, et son supremum sera le δ désiré.

Cette fonction est définissable, et nous donne donc une relation fonctionnelle sur les ordinaux, que je noterai F (étant donné un ordinal γ on prend le plus petit ordinal $\delta > \gamma$ tel que pour tout uplet $\bar{a}$ dans V_γ ...).

Nous définissons une suite strictement croissante α_n, en posant $\alpha_0 = \beta$; supposons α_{n-1} choisi. Si n est impair, alors α_n est le plus petit ordinal $> \alpha_{n-1}$ tel que ψ et toutes ses sous-formules sont absolues pour V_{α_n} (HI). Si n est pair, on prend $\alpha_n = F(\alpha_{n-1})$. Et finalement on prend

$$\alpha = \bigcup_{n \in \omega} \alpha_{2n} = \bigcup_{n \in \omega} \alpha_{2n+1}.$$

ψ et toutes ses sous-formules sont absolues pour V_α ; il reste à montrer que φ est absolue pour V_α. Soit $\bar{a}$ un uplet de V_α ; comme la suite α_n est strictement croissante, α est un ordinal limite, et donc il existe $m \in \omega$, m pair, tel que $\bar{a} \in V_{\alpha_m}$.

Si $\mathcal{U} \models \exists x\, \psi(x, \bar{a})$, alors il existe $b \in V_{\alpha_{m+1}}$ tel que $\mathcal{U} \models \psi(b, \bar{a})$. Comme $\alpha_{m+1} < \alpha$ et ψ est absolue pour V_α, nous obtenons $V_\alpha \models \psi(b, \bar{a})$, et donc $V_\alpha \models \varphi(\bar{a})$.

L'autre direction est encore plus simple.

Comme quelqu'un l'avait remarqué en classe, on peut se permettre de ne pas faire le cas de la conjonction, en mettant la formule sous forme prénexe : une formules sans quantificateurs est toujours absolue.

6.4 Les ensembles constructibles

Nous travaillons dans un modèle $\mathcal{U}$ de ZFC + AF. Le langage est $\{\in\}$. Je donnerai très peu de preuves, les personnes intéréssées peuvent les trouver dans le livre de J.-L. Krivine, *Théorie axiomatique des ensembles*.

Définition 6.67. Soit a un ensemble. On définit $\Pi(a) \subset \mathcal{P}(a)$ comme l'ensemble des sous-ensembles définissables (avec paramètres) de la structure $(a, \in)$.

Remarques 6.68. (1) Donc $b \in \Pi(a)$ si et seulement s'il existe une formule $\varphi(\bar{v}, w)$ et un uplet fini $\bar{c}$ de a tel que

$$b = \{x \in a \mid a \models \varphi(\bar{c}, x)\}.$$

(2) $\Pi(a)$ est une algèbre de Boole, qui contient a, $\emptyset$, et tous les sous-ensembles finis de a. C'est-à-dire, vue comme sous-ensemble de $\mathcal{P}(a)$, elle est close par $\bigcap$ et complémentaire (et union, contient $\emptyset$, a).

(3) Si a est fini, alors $\Pi(a) = \mathcal{P}(a)$. Cependant si a est infini (rappel, nous avons l'axiome du choix), alors $\mathrm{card}(\Pi(a)) = \mathrm{card}(a)$, puisqu'il y a exactement $\mathrm{card}(a)$ formules avec paramètres dans a. En particulier, $\Pi(a)$ est beaucoup plus petit que $\mathcal{P}(a)$.

(4) On peut avoir $a \subset b$ et $\Pi(a) \not\subset \Pi(b)$: en effet, c'est le cas si $a \subset b$ et $a \notin \Pi(b)$.

Lemme 6.69. *Si $a \subset b$ et $a \in b$, alors $\Pi(a) \subset \Pi(b)$.*

Démonstration. Soit $c \subset a$ défini par la formule $\varphi(x, \bar{d})$. Alors il sera défini dans b par la relativisée $\varphi^a(x, \bar{d})$ de φ à a. (Cf 5.56).

Définition 6.70. On définit par induction sur l'ordinal α une relation fonctionnelle $y = L_\alpha$, de domaine la classe des ordinaux. On pose

$$L_0 = \emptyset, \qquad L_\alpha = \bigcup_{\beta<\alpha} \Pi(L_\beta).$$

La collection L est la réunion des L_α : $x \in L$ si et seulement si $\mathcal{U} \models \exists \alpha\, \mathrm{On}(\alpha) \wedge x \in L_\alpha$. Les éléments de L sont appelés des *constructibles*.

Remarque 6.71. Il y tout de suite un petit problème : pour pouvoir définir la relation fonctionnelle $y = L_\alpha$, il faut pouvoir définir la relation $y = \Pi(x)$. Il faut donc exprimer dans notre langage la propriété suivante d'une paire (x, y) :
Il existe une formule $\varphi(w, \bar{v})$ et un uplet $\bar{a}$ d'éléments de x (de longueur $|\bar{v}|$) tels que $y = \{z \in x \mid x \models \varphi(z, \bar{a})\}$.
Nous admettrons que cette relation est définissable. La preuve formelle est longue, en voici l'idée. Dans $\mathcal{U}$, nous avons l'ordinal définissable ω, ensemble de tous les ordinaux finis. En utilisant la fonction successeur, et une définition par induction, on arrive à définir les opérations $+$ et $\times$ sur ω, identifié au semi-anneau $\mathbb{N}$ des entiers. Les résultats sur les codages de formules et leur complexité, nous permettent alors de reconnaître les numéros de Gödel des formules du langage, leur arité, et étant donnés $\#\varphi(\bar{v})$ et un uplet $\bar{a}$ d'éléments de l'ensemble b, si $b \models \varphi(\bar{a})$ ou non. La satisfaction étant définie par induction sur la complexité des formules. Je supposerai donc admis la chose suivante :

il existe une formule Δ_0 $\mathrm{Sat}(x, y, z)$, (et à paramètres dans $L_\omega = V_\omega$) telle que $\mathcal{U} \models \mathrm{Sat}(m, b, c)$ si et seulement si
(i) $m \in \omega$ est le numéro de Gödel d'une formule $\varphi(\bar{v})$;
(ii) $c \in b^{|\bar{v}|}$ (donc, un $|\bar{v}|$-uplet d'éléments de b ; comme ci-dessus, $|\bar{v}|$ dénote la longueur de $\bar{v}$) ;
(iii) $(b, \in) \models \varphi(c)$.

Notez que le fait que cette formule soit Δ_0 n'est pas surprenant, puisque tous les quantificateurs sont soumis aux conditions $\in \omega$ ou bien $\in y$.

Remarques 6.72. Voici quelques propriétés faciles :
(1) Si $\gamma < \alpha$ alors $L_\gamma \subset L_\alpha$. En effet, par définition, $L_\gamma = \bigcup_{\beta<\gamma} \Pi(L_\beta) \subset \bigcup_{\beta<\alpha} \Pi(L_\beta) = L_\alpha$. De plus $L_\gamma \in L_\alpha$, car $L_\gamma \in \Pi(L_\gamma)$.
(2) Donc $\beta < \alpha$ implique $\Pi(L_\alpha) \subset \Pi(L_\beta)$ (par 6.69), et $L_{\alpha+1} = \Pi(L_\alpha)$, $L_\beta = \bigcup_{\gamma<\beta} L_\gamma$ si β est un ordinal limite.
(3) $L_\omega = V_\omega$. En effet, $V_n = L_n$ pour tout $n \in \omega$ car V_n est fini.
(4) On peut voir L_α comme étant inclus dans V_α : cela suit des inclusions naturelles : si $a \subset b$, alors $\Pi(a) \subseteq \mathcal{P}(a) \subset \mathcal{P}(b)$.

Définition 6.73. Si a est constructible, on appelle *ordre de* a, noté $od(a)$, le plus petit ordinal α tel que $a \in L_\alpha$. C'est un ordinal successeur.

Lemme 6.74. *Si a est constructible et $b \in a$ alors b est constructible, et $od(b) < od(a)$.*

Démonstration. Si $\alpha = \beta+1 = od(a)$, alors $a \in L_\alpha = \Pi(L_\beta)$, i.e. $a \subset L_\beta$, $b \in L_\beta$, $od(b) \leq \beta < \alpha$.

Corollaire 6.75. *Chaque L_α est transitif.*

Lemme 6.76. *Tout ordinal α est constructible, et $od(\alpha) = \alpha + 1$.*

Démonstration. On sait déjà que $od(\alpha) \not\leq \alpha$, car $L_\alpha \subseteq V_\alpha$ et nous avons montré que $\alpha \notin V_\alpha$. Il suffit donc de montrer que $\alpha \in L_{\alpha+1}$. Supposons que ce ne soit pas le cas, et prenons un tel α minimal. Ce n'est pas 0 car $\emptyset \in L_1 = \{\emptyset\}$. Alors pour tout $\beta < \alpha$ on a $\beta \in L_{\beta+1}$, i.e., $\beta \in \Pi(L_\beta)$. Donc $\alpha \subset \bigcup_{\beta\in\alpha} \Pi(L_\beta)$, i.e., $\alpha \subset L_\alpha$. Il faut maintenant montrer que α est définissable dans L_α. Mais on sait que α est l'ensemble des ordinaux inférieurs à α, et le fait d'être un ordinal est définissable par une formule Δ_0 : x est un ordinal ssi $\in$ définit un ordre total sur x et x est transitif. Notez que comme nous avons AF, nous n'avons pas besoin de dire que l'ordre est bien fondé (ce qui ne serait pas Δ_0). Donc α est définissable dans L_α, et donc appartient à $L_{\alpha+1}$.

On peut alors montrer toute une série de résultats d'indépendance, disant que si ZFC a un modèle, alors d'autres théories ont un modèle. Nous avions déjà vu que si ZFC a un modèle, alors aussi ZFC + AF. (On passe à la sous-classe V).

Théorème 6.77. *Soit $\mathcal{U} \models ZFC + AF$, et L défini par $x \in L$ ssi $\exists\alpha\, \mathrm{On}(\alpha) \wedge x \in L_\alpha$. Alors $L \models ZFC + AF$, et $L \models \forall x \exists\alpha\, \mathrm{On}(\alpha) \wedge x \in L_\alpha$.*

6.78. Quelques commentaires. La première assertion n'est pas très difficile à montrer, on fait un peu comme pour V. La deuxième est plus difficile, car bien que V sache que tous les éléments de L sont constructibles, peut-être L ne le sait pas. Pour cela il faut montrer que l'appartenance à L_α est définissable par une formule qui, si elle est vraie dans L, est aussi vraie dans $\mathcal{U}$.
La dernière assertion est souvent abbrégée en disant "$V = L$". Je vais faire quelques esquisses de preuves que j'ai préparées, il y a plus de détails dans le livre de Krivine.

6.79. *Quelques éléments de preuve.* Comme la formule $\mathrm{On}(x)$ est Δ_0, elle définit dans L exactement les ordinaux de $\mathcal{U}$. De même, ω est défini par la formule $\emptyset \in x$ et $\mathrm{On}(x)$ et $\forall y \in x \exists z \in x\, y \in z$ et $\forall y \in x \exists z\, y = z \cup \{z\} \vee y = \emptyset$, qui est Δ_0. Ce qui veut dire que ce que $\mathcal{U}$ nous dit au sujet de la satisfaction de formules sera aussi vrai dans L_α pour $\alpha > \omega$. On sait déjà que chaque L_α satisfait ZF1, ZF2 (par la Proposition 6.54, et la preuve de 6.55). Pour ZF3, il faut montrer que si $a \in L_\alpha$, alors $\Pi(a)$ est dans L, ce qui est évident : car $\Pi(a) \subset \Pi(L_\alpha)$, et est définissable par la formule $x \in \Pi(a)$ si et seulement si $x \subset a$ (cette formule marche dans L, pas dans $\mathcal{U}$).
Pour ZF4: Soit $\bar{c}$ un uplet de L, $R(x, y, \bar{c})$ une relation fonctionnelle (pour L), et $a \in L$. On veut montrer que l'ensemble $\{b \mid \exists x \in a, L \models R(x, b, \bar{c})\}$ est dans L. Tout d'abord on prouve

un schéma de réflexion généralisé : la preuve de 6.64 se généralise à L : si α est tel que a et $\bar{c}$ sont dans L_α, alors il existe un $\beta > \alpha$ tel que pour tout $b, b' \in L_\beta$, on a $L \models R(b, b', \bar{c})$ si et seulement si $L_\beta \models R(b, b', \bar{c})$. Alors, comme l'image de a par $R(-, -, \bar{c})$ est définissable, elle sera en fait dans $L_{\beta+1}$. La preuve de ZF5 est facile ; et pour AF on fait comme on avait fait pour V : si $a \in L$ est non-vide, on prend un élément $b \in a$ ayant ordre minimal. Alors $a \cap b = \emptyset$. Il reste à montrer que L satisfait AC. Pour cela, nous allons mettre un ordre bien fondé sur tous les L_α, de façon que si $\omega \leq \alpha < \beta$ alors L_α est un segment initial de L_β. Si $a \in L$, alors a est dans un L_α, et l'ensemble des singletons d'élements de a est un sous-ensemble de $L_{\alpha+1}$ et donc héritera d'un bon ordre. On va supposer $\alpha \geq \omega$, (car $V_\omega = L_\omega$ – on doit pouvoir définir un bon ordre sur L_ω). Pour α limite, on prend la réunion des ordres définis sur les L_β, $\beta < \alpha$. Regardons maintenant le cas successeur, je note $<_\alpha$ le bon ordre sur L_α. Un élément de $\Pi(L_\alpha)$ est donc un sous-ensemble de L_α définissable dans L_α par une formule à paramètres. On choisit un bon ordre $\prec_\alpha$ sur les formules (du langage $\{\in\}$) à paramètres dans L_α et avec une seule variable libre. Ce bon ordre est bien sûr défini à partir du bon ordre de L_α. Puis, on définit un ordre sur $L_{\alpha+1}$ de la façon suivante : soient $a, b \in L_{\alpha+1}$. S'ils sont tous deux dans L_α, alors $a <_{\alpha+1} b$ ssi $a <_\alpha b$: si a est dans L_α et b dans $L_{\alpha+1} \setminus L_\alpha$, alors $a <_{\alpha+1} b$; et enfin, si a et b sont tous deux dans $L_{\alpha+1} \setminus L_\alpha$, alors $a <_{\alpha+1} b$ si et seulement si la "plus petite" formule définissant a est $\prec_\alpha$ à celle définissant b. Cela définit un bon ordre sur $L_{\alpha+1} \setminus L_\alpha$, et il est clair que l'ordre sur $L_{\alpha+1}$ satisfait les conditions voulues. Nous avons donc montré que L satisfait Zermelo. La dernière assertion utilise les propriétés suivantes, qu'il faut montrer, mais que j'espère sont crédibles :
La relation fonctionnelle $y = \Pi(x)$ est Δ_0. En effet, elle parle de On, de ω, qui sont tous deux Δ_0-définissables, elle parle de satisfaction, et tous les quantificateurs sont bornés. $[\forall z(z \in y \leftrightarrow \exists n \in \omega,\ n = \#\varphi(\bar{v}, w) \wedge \exists \bar{v} \in x^{|\bar{v}|}, (\forall t\, t \in z \iff t \in x \wedge \varphi^x(\bar{v}, t)$, ou quelque chose comme ça]. On utilise aussi une version améliorée du Lemme d'induction 6.19 : si on suppose que la relation fonctionnelle F est définissable par une formule Σ_1, et que S est tout l'univers, alors la relation (de domaine les ordinaux) définie par $f(\alpha) = F(f_{|_\alpha})$, est aussi Σ_1. En effet, elle est définissable par la formule $y = f(\alpha)$ ssi $\mathrm{On}(\alpha)$ et $\exists g$ (g est une fonction définie sur α et $\forall \beta \in \alpha, g(\beta) = F(g_{|_\beta})$ et $y = F(g)$). Un quantificateur existentiel appliqué à des formules Σ_1. De cela on déduit :
$y = L_\alpha$ est une relation fonctionnelle Σ_1. Donc si L croit que $y = L_\alpha$, alors c'est aussi vrai dans $\mathcal{U}$. On en déduit que

$$L \models \forall x \exists \alpha\, x \in L_\alpha.$$

En effet, soit a dans L, donc dans un L_α, avec α un ordinal. Mais l'ensemble L_α est le même dans $\mathcal{U}$ et dans L (puisque $y = L_\alpha$ est Σ_1.) Donc $L \models a \in L_\alpha$.

Théorème 6.80. *$\mathcal{U} \models ZFC + AF$. Alors $L \models GCH$ (l'hypothèse généralisée du continu).*

La preuve nécéssite un résultat que je ne montrerai pas :

Fait. *(ZFC + AF). Soit α un ordinal $\geq \omega$. Alors* $\mathrm{card}(L_\alpha) = \mathrm{card}(\alpha)$.
Soit a dans L. Alors $\mathrm{card}(od(a)) \leq \mathrm{card}(\mathrm{Cl}(a)) + \aleph_0$, *où* Cl *dénote la clôture transitive de a.*

Preuve de GCH pour L. On peut donc appliquer les résultats du fait précédent à L puisque nous

savons maintenant que L est un modèle de ZFC et de AF. Nous voulons montrer que $\mathcal{P}(\aleph_\rho)$ a cardinalité $\aleph_{\rho+1}$. Soit $a \subset \aleph_\rho$ constructible. Alors $\mathrm{card}(od(a)) \leq \mathrm{card}(\mathrm{Cl}(a)) + \aleph_0 \leq \aleph_\rho$: c'est parce que comme $a \subset \aleph_\rho$, alors aussi $\mathrm{Cl}(a)$. Donc $od(a) < \aleph_{\rho+1}$, i.e., $a \in L_{\aleph_{\rho+1}}$. Donc $\mathcal{P}(\aleph_\rho) \subset L_{\aleph_{\rho+1}}$; mais $\mathrm{card}(L_{\aleph_{\rho+1}}) = \aleph_{\rho+1}$, ce qui finit la preuve.

6.81. Suggestions de lecture, pour aller plus loin

Théorie des Modèles :
David Marker, *Model Theory: an introduction*, Graduate texts in Mathematics 217, Springer-Verlag 2002.
Katrin Tent, Martin Ziegler, *A course in Model theory*, Lecture notes in Logic 40, Cambridge University Press, 2012.
(plus classique) C. C. Chang, J.K. Keisler, *Model Theory*, 3rd edition, Studies in Logic and the Foundation of Mathematics 73, North-Holland, 1990.

Théorie des Ensembles :
Kenneth Kunen, *An introduction to independence proofs*. Studies in Logic and the Foundation of Mathematics 102, North Holland 1983.
Alexander S. Kechris, *Classical Descriptive Set theory*, Graduate texts in Mathematics, Springer-Verlag 1995.

Récursivité :
Hartley Rogers Jr, *Theory of recursive functions and effective computability*, 2nd edition, MIT Press, 1987.

Bibliographie

[1] R. Cori, D. Lascar, *Logique Mathématique*, Tome 2. Dunod, Paris 2003.

[2] M. Hils, *Cours de logique mathématique*, 2012.
(http://www.logique.jussieu.fr/∼hils/enseignement/Notes_Cours2012-13.pdf)

[3] J.-L. Krivine, *Théorie axiomatique des ensembles*, PUF, Paris 1972.

[4] F. Loeser, *Un premier cours de logique*, 2010.
(http://www.math.jussieu.fr/∼loeser/notes.php)

[5] Y.I. Manin, *A course in mathematical logic for mathematicians*, Springer, 2010.