Suites de Goodstein et incomplétude de l'arithmétique de Peano

Gabriel Scherer et Silvain Rideau

17 Juin 2009

Nous remercions François Loeser de nous avoir proposé ce sujet, Laurence Rideau et Catherine Kikuchi pour les corrections.
Le premier théorème d'incomplétude de Gödel donne l'existence de formules exprimables dans $\mathcal{L}_{\mathrm{P}}$, le langage de P l'arithmétique de Peano, qui ne soient pas démontrables dans P. Cependant, la formule construite par Gödel pour démontrer son théorème n'a aucun contenu arithmétique, il faut attendre 1982 pour que Laurie Kirby et Jeff Paris (voir [KP82]) démontrent que le théorème de Goodstein, qui est un résultat purement arithmétique, n'est pas démontrable dans l'arithmétique de Peano. On démontrera ici ce résultat en utilisant le théorème de Wainer qui donne une majoration des fonctions prouvablement totales dans P en s'inspirant de la démonstration de [AZ97] et [Cic83].
On s'intéressera, dans la deuxième partie, au lien entre P et l'ordinal[1] ϵ_0 défini comme suit :

Définition 0.0.1 (ϵ_0)**.** Soit la suite d'ordinaux $(\omega_n)_{n\in\mathbb{N}}$ telle que :

$$\begin{cases} \omega_0 &= 1 \\ \omega_{n+1} &= \omega^{\omega_n} \quad \forall n \in \mathbb{N} \end{cases}$$

On note $\epsilon_0 = \bigcup_{n\in\mathbb{N}} \omega_n$.

Ce lien a été mis en avant par Gentzen en démontrant que la bonne fondation de ϵ_0 implique la cohérence de P. On montrera ici que le théorème de Goodstein est équivalent à l'induction jusqu'à ϵ_0 et donc par le deuxième théorème d'incomplétude de Gödel, cela fournit une deuxième preuve du fait que le théorème de Goodstein est indémontrable dans P. Cela montre de plus à quel point le théorème de Goodstein est lié à l'incomplétude de P.

[1] voir appendice A pour des précisions sur les ordinaux.

Chapitre 1

Le théorème de Kirby et Paris

1.1 Les suites de Goodstein

Commençons par définir les suites de Goodstein et démontrer qu'elles sont stationnaires en 0 à partir d'un certain rang. Pour cela on définit une variante de la notion de base, la base itérée ou pure. Elle consiste à écrire les exposants aussi dans cette base, puis les exposants des exposants et ainsi de suite jusqu'à ce que la notation ne contienne plus que des coefficients inférieurs à la base. Plus précisément, on définit une fonction de remplacement de q par x dans cette notation en base q itérée.

Définition 1.1.1 (Remplacement dans la base q itérée)**.** Soient $p, q \in \mathbb{N}$ et $p = \sum_{i=0}^{k} q^i a_i$ sa notation en base q. On définit la fonction suivante par récurrence sur p :

$$f_{q,x}(p) = \sum_{i=0}^{k} x^{f_{q,x}(i)} a_i$$

On peut l'étendre et définir $f_{q,\omega}(p) = \sum_{i=0}^{k} \omega^{f_{q,\omega}(i)} a_i \ < \epsilon_0$ car ϵ_0 est stable par addition, multiplication et exponentiation.

On peut alors définir les suites suivantes :

Définition 1.1.2 (Suite de Goodstein de p commençant en base q)**.** Soient $p, q \in \mathbb{N}$ on définit la suite $g_n^{p,q}$ (souvent notée simplement g_n) par :

$$\begin{cases} g_0 &= p & \\ g_{n+1} &= 0 & \text{si } g_n = 0 \\ g_{n+1} &= f_{q+n,q+n+1}(g_n) - 1 & \text{sinon} \end{cases}$$

A chaque élément de la suite g_n on associe $\alpha_n = f_{q+n,\omega}(g_n)$

.
La proposition suivante montre que, bien que la suite semble exploser avec le remplacement dans la base itérée, la présence du -1 rend la suite d'ordinaux associés décroissante.

Proposition 1.1.3. *Si $g_n \neq 0$ alors $\alpha_n = f_{q+n+1,\omega}(g_{n+1} + 1)$.*

Démonstration. $f_{q+n+1,\omega}(g_{n+1}+1) = f_{q+n+1,\omega}(f_{q+n,q+n+1}(g_n)) = f_{q+n,\omega}(g_n)$ car en écrivant en base $q+n$ itérée puis en remplaçant les $q+n$ par des $q+n+1$ on obtient l'écriture en base $q+n+1$ itérée. □

On en déduit donc le théorème suivant :

Théorème 1.1.4 (Théorème de Goodstein)**.** *Toutes les suites de Goodstein sont stationnaires en* 0 *à partir d'un certain rang.*

Démonstration. On commence par démontrer que pour tout $n \in \mathbb{N}$, $f_{n,\omega}(x)$ est une fonction strictement croissante de x.
En effet par récurrence sur x, si $x = 0$ alors $f_{n,\omega}(1) = 1 > 0 = f_{n,\omega}(0)$. Sinon soient k maximal tel que $n^k \leqslant x+1$, $a > 0$ maximal tel que $an^k \leqslant x+1$ et $j = x+1-an^k < n^k$. Si $j \neq 0$ alors

$$f_{n,\omega}(x+1) = \omega^{f_{n,\omega}(k)} \cdot a + f_{n,\omega}(j) > \omega^{f_{n,\omega}(k)} \cdot a + f_{n,\omega}(j-1) = f_{n,\omega}(x)$$

Sinon $x+1 = an^k$ et donc $x = (a-1)n^k + \sum_{i=0}^{k-1}(n-1)n^i$.
Par récurrence, $f_{n,\omega}(k) \geqslant f_{n,\omega}(k-1)+1$ donc

$$\omega^{f_{n,\omega}(k)} \geqslant \omega^{f_{n,\omega}(k-1)+1} > \omega^{f_{n,\omega}(k-1)} \cdot n > \sum_{i=0}^{k-1} \omega^{f_{n,\omega}(i)} \cdot (n-1)$$

et donc

$$\omega^{f_{n,\omega}(k)} \cdot a > \omega^{f_{n,\omega}(k)} \cdot (a-1) + \sum_{i=0}^{k-1} \omega^{f_{n,\omega}(i)} \cdot (n-1)$$

On en conclut que pour tout n, $f_{n,\omega}(x)$ est strictement croissante.

D'après la proposition 1.1.3 et la croissance de $f_{q+n+1,\omega}(x)$, si $g_n \neq 0$ alors :

$$\alpha_n = f_{q+n+1,\omega}(g_{n+1}+1) > f_{q+n+1,\omega}(g_{n+1}) = \alpha_{n+1}$$

Donc comme les ordinaux sont bien fondés, $\exists k\ g_k = 0$ et donc $\forall n > k\ g_n = 0$. □

On peut donc définir la fonction suivante. Elle est récursive car le remplacement en base q itérée est récursif, ie $f_{n,m}(x)$ vue comme une fonction de n,m et x est récursive, ce qui implique que $g_n^{p,q}$ est une fonction récursive de n, p et q. De plus, elle est totale par le théorème de Goodstein.

Définition 1.1.5 (Fonction de Goodstein)**.**

$$h : \begin{array}{rcl} \mathbb{N} \times \mathbb{N} \backslash \{0,1\} & \to & \mathbb{N} \\ (p,q) & \mapsto & \min\{t \mid g_t^{p,q} = 0\} \end{array}$$

1.2 Quelques résultats sur les ordinaux

Pour démontrer le résultat de Kirby et Paris, on aura besoin de quelques outils sur les ordinaux dus entre autre à Ketonen et Solovay (voir [KS81]). Ces résultats nous permettront de définir la hiérarchie de Hardy qui joue, pour les fonctions récursives prouvablement totales dans P, le rôle que joue la fonction de Ackerman pour les fonctions primitives récursives. Ils permettrons aussi de prouver le théorème de Cichon qui donne une valeur de $h(p, q)$.

1.2.1 La forme normale de Cantor et la base ω itérée

Pour commencer, on montre le résultat suivant qui permet l'écriture en base ω des ordinaux $< \epsilon_0$

Proposition 1.2.1 (Forme normale de Cantor)**.** *Tout $\lambda < \epsilon_0$ peut être écrit sous sa forme normale de Cantor :*

$$\lambda = \omega^{\lambda_1} + \cdots + \omega^{\lambda_n} = \beta + \omega^{\lambda_n} \text{ où } \lambda > \lambda_1 \geqslant \cdots \geqslant \lambda_n \geqslant 0$$

Démonstration. Montrons l'existence et l'unicité de cette forme normale par induction structurelle sur $\lambda < \epsilon_0$.
Si $\lambda = 1$ alors $n = 1$ et $1 = \omega^0$.
Si λ est successeur, $\lambda = \gamma + 1$. Par induction $\gamma = \sum_{i=1}^{n} \omega^{\lambda_i}$ donc $\lambda = \sum_{i=1}^{n+1} \omega^{\lambda_i}$ avec $\lambda_{n+1} = 0$ et on a bien $\lambda > \gamma > \lambda_1 \geqslant \cdots \geqslant \lambda_n \geqslant \lambda_{n+1} \geqslant 0$.
Pour l'unicité, remarquons d'abord que $\lambda = \sum_{i=1}^{n} \omega^{\lambda_i}$ alors $n \neq 0$ et $\lambda_n = 0$ sinon λ est limite. Donc si $\lambda = \sum_{i=1}^{n} \omega^{\lambda_i} = \sum_{i=1}^{m} \omega^{\mu_i}$ deux formes normales de Cantor, alors $\lambda_n = \mu_m = 0$ et $\gamma = \sum_{i=1}^{n-1} \omega^{\lambda_i} = \sum_{i=1}^{m-1} \omega^{\mu_i}$. Par induction la forme normale de Cantor de γ est unique donc $n = m$ et $\forall i < n\ \lambda_i = \mu_i$.
Si λ est limite, soit $\lambda_1 = \max\{\kappa \mid \omega^\kappa \leqslant \lambda\}$. Cet ensemble est non vide car $\omega^0 = 1 \leqslant \lambda$ et borné car $\omega^\lambda > \lambda$.
En effet, supposons $\omega^\lambda \leqslant \lambda$ alors $\lambda \geqslant \epsilon_0$ car $\omega_0 = 1 < \lambda$ et par récurrence $\forall n\ \omega_{n+1} = \omega^{\omega_n} \leqslant \omega^\lambda \leqslant \lambda$. Donc $\epsilon_0 = \bigcup_{n\in\mathbb{N}} \omega_n \leqslant \lambda$. Mais comme $\lambda < \epsilon_0$ par hypothèse, on a $\epsilon_0 < \epsilon_0$ ce qui est absurde.
Soit $\gamma < \lambda$ tel que $\lambda = \omega^{\lambda_1} + \gamma$ alors par induction $\gamma = \sum_{i=2}^{n} \omega^{\lambda_i}$ avec $\gamma > \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0$ il ne reste donc plus qu'à montrer que $\lambda > \lambda_1 \geqslant \lambda_2$.
Si $\lambda_2 > \lambda_1$ alors $\exists \delta > 0\ \lambda_2 = \lambda_1 + \delta$ et :

$$\begin{aligned} \lambda &= \omega^{\lambda_1} + \omega^{\lambda_2} + \ldots \\ &= \omega^{\lambda_1}(1 + \omega^\delta) + \ldots \\ &= \omega^{\lambda_1}\omega^\delta + \ldots \end{aligned}$$

Donc $\omega^{\lambda_2} \leqslant \lambda$ avec $\lambda_2 > \lambda_1$ ce qui contredit la définition de λ_1.
De plus si $\lambda_1 \geqslant \lambda$ alors $\lambda \geqslant \omega^{\lambda_1} \geqslant \omega^\lambda$ ce qui est impossible comme on l'a montré précédemment.
Montrons alors l'unicité, supposons $\lambda = \sum_{i=1}^{n} \omega^{\lambda_i} = \sum_{j=1}^{m} \omega^{\mu_j}$. Comme $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ et $\mu_1 \geqslant \cdots \geqslant \mu_m$, on a, pour tout $i \leqslant n,\ \lambda_i = \lambda_n + \kappa_i$, pour tout $j \leqslant m,\ \mu_j = \mu_m + \nu_i$ et

$$\lambda = (\sum_{i=1}^{m} \omega^{\kappa_i})\omega^{\lambda_m} = \alpha\omega^{\lambda_m} = (\sum_{i=1}^{n} \omega^{\nu_i})\omega^{\mu_n} = \beta\omega^{\mu_n}$$

où α et β sont successeurs et comme $\lambda_m, \mu_n \neq 0$ (sinon λ serait successeur) $\alpha, \beta < \lambda$. Si $\lambda_m \geqslant \mu_n$ alors $\lambda_m = \mu_m + \kappa$ et $\alpha\omega^\kappa = \beta$. Comme β est successeur, $\kappa = 0$ donc $\lambda_m = \mu_n$ et $\sum_{i=1}^m \omega^{\kappa_i} = \alpha = \beta = \sum_{i=1}^n \omega^{\nu_i}$. Par unicité de la forme normale de Cantor, $n = m$ et $\forall i \leqslant n\ \kappa_i = \nu_i$ et donc $\lambda_i = \mu_i$. □

En rassemblant les λ_i identiques on obtient pour tout $\alpha < \epsilon_0$ son écriture en base ω :

$$\alpha = \sum_{i=0}^{k} \omega^{\lambda_i} \cdot a_i$$

où

$$\forall i \leqslant k\ a_i \in \mathbb{N} \text{ et } \alpha > \lambda_0 > \cdots > \lambda_k \geqslant 0$$

La somme d'ordinaux n'étant pas commutative, $\sum_{i=0}^k$ doit se comprendre comme la somme des termes en commençant à gauche par le terme $i = 0$ et en finissant à droite par le terme $i = k$.
En réitérant le processus pour les exposants comme on l'a fait pour les entiers, on obtient la représentation en base ω itérée de tous les ordinaux appartenant à ϵ_0. On peut alors définir,comme on l'a fait pour les entiers, pour tout $n \in \mathbb{N}$ et $\alpha < \epsilon_0$, $f_{\omega,n}(\alpha)$.
Cette écriture en base ω itérée permet de coder les ordinaux appartenant à ϵ_0 dans l'arithmétique de Peano, par exemple par un n-uplet qui contient le code des λ_i. On peut vérifier que toutes les opérations sur les ordinaux, ainsi que les suites ordinales et les fonctions de la hiérarchie de Hardy (vues comme des fonctions de deux arguments, le code de l'ordinal et l'entier) sont récursives grâce à cet encodage des ordinaux.
De plus pour tout $\alpha < \epsilon_0$, si tous les coefficients de α en base ω itérée sont strictement inférieurs à un certain $n \geqslant 2$, on dit que α est n-représentable et on a $f_{n,\omega}(f_{\omega,n}(\alpha)) = \alpha$.
On définit de plus la relation suivante sur les ordinaux :

Définition 1.2.2. Soient $\alpha, \beta < \epsilon_0$, soient $\alpha = \sum_{i=0}^k \omega^{\alpha_i}$ et $\beta = \sum_{j=0}^l \omega^{\beta_j}$ leurs formes normales de Cantor. On dit que $\alpha \ll \beta$ si $\beta_0 \geqslant \alpha_k$.

Pour finir on remarque que :

Proposition 1.2.3 (Majoration par un ω_n)**.** *Pour tout $\alpha < \epsilon_0$ il existe $n \in \mathbb{N}$ tel que $\alpha < \omega_n$*

Démonstration. Cette proposition est un cas particulier de la proposition A.2.5, appliquée au cas de ϵ_0. On peut aussi la démontrer directement à partir de ce qui précède.
Montrons cette propriété par induction sur $\alpha < \epsilon_0$.
En effet $\alpha = \sum_{i=0}^k \omega^{\lambda_i} \cdot a_i < \omega^{\lambda_0} \cdot (a_0 + 1) < \omega^{\lambda_0 + 1}$. Par induction, il existe $n \in \mathbb{N}$, $\lambda_0 + 1 < \omega_n$. On en déduit que $\alpha < \omega^{\omega_n} = \omega_{n+1}$. □

1.2.2 Suites ordinales

On définit alors une première suite d'ordinaux associée à chaque ordinal appartenant à ϵ_0.

Définition 1.2.4. Pour tout $\alpha < \epsilon_0$, on définit la suite $(\langle\alpha\rangle(n))_{n\in\mathbb{N}}$ par induction sur α :

$$\begin{cases} \langle 0\rangle(n) &= 0 & \\ \langle \alpha\rangle(n) &= \alpha - 1 & \text{si } \alpha \text{ est successeur} \\ \langle \beta + \omega^\delta\rangle(n) &= \beta + \omega^{\langle\delta\rangle(n)} \cdot (n-1) + \langle \omega^{\langle\delta\rangle(n)}\rangle(n) & \text{si } \alpha \text{ est limite} \end{cases}$$

Pour tout $k > 1$, on note $\langle\alpha\rangle(n_1, \ldots, n_k) = \langle\langle\alpha\rangle(n_1, \ldots, n_{k-1})\rangle(n_k) = \langle\langle\alpha\rangle(n_1)\rangle(n_2, \ldots, n_k)$.

Cette suite est bien définie car on vérifie facilement par induction que si $\alpha \neq 0$ alors pour tout $n \in \mathbb{N}$, $\langle\alpha\rangle(n) < \alpha$.
Les trois propositions suivantes décrivent les propriétés de ces suites, leur lien avec le remplacement par ω dans les bases itérées et donc avec la suite ordinale associée à une suite de Goodstein.

Proposition 1.2.5. *Soient $n \in \mathbb{N}$ et $\alpha, \beta < \epsilon_0$ tels que $\beta \neq 0$ et $\beta \ll \alpha$, alors*

$$\langle\alpha + \beta\rangle(n) = \alpha + \langle\beta\rangle(n)$$

Démonstration. Si $\beta = \gamma + 1$ alors $\langle\alpha + \gamma + 1\rangle(n) = \alpha + \gamma = \alpha + \langle\beta\rangle(n)$
Sinon par unicité de la forme normale de Cantor, $\alpha + \gamma + \omega^\delta$ est en forme normale de Cantor et $\langle\alpha + \gamma + \omega^\delta\rangle(n) = \alpha + \gamma + \omega^{\langle\delta\rangle(n)} \cdot (n-1) + \langle\omega^{\langle\delta\rangle(n)}\rangle(n) = \alpha + \langle\beta\rangle(n)$ □

Proposition 1.2.6. *Soient $n > 1$ et $m \in \mathbb{N}$,*

$$\langle f_{n,\omega}(m+1)\rangle(n) = f_{n,\omega}(m)$$

Démonstration. Soient k maximal tel que $n^k \leqslant m+1$, $a > 0$ maximal tel que $an^k \leqslant m+1$ et $j = m + 1 - an^k < n^k$. On a $f_{n,\omega}(m+1) = \omega^{f_{n,\omega}(k)} \cdot a + f_{n,\omega}(j)$.
Si $k = 0$, $m + 1 < n$ et

$$\begin{aligned}
\langle f_{n,\omega}(m+1)\rangle(n) &= \langle m+1\rangle(n) \\
&= m \\
&= f_{n,\omega}(m)
\end{aligned}$$

Si $j \neq 0$ alors $f_{n,\omega}(j) \neq 0$ et

$$\begin{aligned}
\langle f_{n,\omega}(m+1)\rangle(n) &= \langle\omega^{f_{n,\omega}(k)} \cdot a + f_{n,\omega}(j)\rangle(n) \\
&= \omega^{f_{n,\omega}(k)} \cdot a + \langle f_{n,\omega}(j)\rangle(n) \\
&= \omega^{f_{n,\omega}(k)} \cdot a + f_{n,\omega}(j-1) \text{ par récurrence} \\
&= \omega^{f_{n,\omega}(k)} \cdot a + f_{n,\omega}(m - an^k) \\
&= f_{n,\omega}(m)
\end{aligned}$$

Si $j = 0$, $m + 1 = an^k$ et

$$\begin{aligned}
\langle f_{n,\omega}(m+1)\rangle(n) &= \langle\omega^{f_{n,\omega}(k)} \cdot (a-1) + \omega^{f_{n,\omega}(k)}\rangle(n) \\
&= \omega^{f_{n,\omega}(k)} \cdot (a-1) + \omega^{\langle f_{n,\omega}(k)\rangle} \cdot (n-1) + \langle\omega^{\langle f_{n,\omega}(k)\rangle}\rangle(n) \\
&= \omega^{f_{n,\omega}(k)} \cdot (a-1) + \omega^{f_{n,\omega}(k-1)} \cdot (n-1) + \langle\omega^{f_{n,\omega}(k-1)}\rangle(n) \\
&= \omega^{f_{n,\omega}(k)} \cdot (a-1) + \omega^{f_{n,\omega}(k-1)} \cdot (n-1) + \langle f_{n,\omega}(n^{k-1})\rangle(n) \\
&= \omega^{f_{n,\omega}(k)} \cdot (a-1) + \omega^{f_{n,\omega}(k-1)} \cdot (n-1) + f_{n,\omega}(n^{k-1} - 1) \\
&= f_{n,\omega}((a-1)n^k + (n-1)n^{k-1} + n^{k-1} - 1) \\
&= f_{n,\omega}(m) \qquad \square
\end{aligned}$$

On en déduit le corollaire suivant qui sera utile dans la partie 2.4.

Corollaire 1.2.7. *Soit $\alpha < \epsilon_0$, si $\alpha \neq 0$ et si α est n-représentable, alors $\langle\alpha\rangle(n)$ est le plus grand ordinal β tel que $\beta < \alpha$ et β est n-représentable.*

Démonstration. Soit $p = f_{\omega,n}(\alpha)$. Comme α est n-représentable, $f_{n,\omega}(p) = \alpha$, et comme $\alpha \neq 0$, $p \neq 0$. La proportion précédente implique que $\langle\alpha\rangle(n) = f_{n,\omega}(p-1)$ donc $\langle\alpha\rangle(n)$ est n-représentable.
Supposons que l'on ait β n-représentable tel que $\beta < \alpha$. Comme $f_{n,\omega}$ est strictement croissante (cf. démonstration de théorème 1.1.4) et que $f_{n,\omega}(f_{\omega,n}(\beta)) = \beta < \alpha = f_{n,\omega}(f_{\omega,n}(\alpha))$, on a $f_{\omega,n}(\beta) < f_{\omega,n}(\alpha) = p$. On a alors $f_{\omega,n}(\beta) \leqslant p-1$ et donc,

$$\beta = f_{n,\omega}(f_{\omega,n}(\beta)) \leqslant f_{n,\omega}(p-1) = \langle\alpha\rangle(n)$$

donc $\langle\alpha\rangle(n)$ est bien le plus grand ordinal n-représentable strictement inférieur à α. □

Proposition 1.2.8. *Soit α_n la suite d'ordinaux associée à une suite de Goodstein commençant en base q,*

$$\forall n \in \mathbb{N},\ \alpha_{n+1} = \langle\alpha_n\rangle(q+n+1)$$

et donc

$$\forall n \in \mathbb{N},\ \alpha_n = \langle\alpha\rangle(q+1,\ldots,q+n)$$

Démonstration. Par les propositions 1.1.3 et 1.2.6, on a, si $g_n \neq 0$:

$$\langle\alpha_n\rangle(q+n+1) = \langle f_{q+n+1,\omega}(g_{n+1}+1)\rangle(q+n+1) = f_{q+n+1,\omega}(g_{n+1}) = \alpha_{n+1}$$

Si $g_n = 0$ alors $g_{n+1} = 0$ et donc $\langle\alpha_n\rangle(q+n+1) = \langle 0\rangle(q+n+1) = 0 = \alpha_{n+1}$. □

On définit aussi une deuxième suite d'ordinaux associée à tout ordinal limite :

Définition 1.2.9 (Suite canonique)**.** On définit pour tout $\alpha < \epsilon_0$ la suite $(\{\alpha\}(n))_{n\in\mathbb{N}}$ par induction sur α.

$$\begin{cases} \{0\}(n) &= 0 \\ \{\beta+1\}(n) &= \beta \\ \{\beta+\omega^\delta\}(n) &= \beta+\omega^{\{\delta\}(n)} & \text{si } \delta \text{ est limite} \\ \{\beta+\omega^\delta\}(n) &= \beta+\omega^{\delta-1}\cdot n & \text{sinon} \end{cases}$$

Cette suite vérifie les propriétés suivantes :

Proposition 1.2.10. *Si α est limite alors $\{\alpha\}(n)$ est strictement croissante et $\bigcup_{n\in\mathbb{N}}\{\alpha\}(n)$.*

Démonstration. Par induction sur α.
Supposons la proposition vérifiée pour tout $\beta < \alpha$. Soient $n < m \in \mathbb{N}$ Soit $\alpha = \beta + \omega^\delta$ la forme normale de Cantor, si δ est successeur,

$$\{\alpha\}(n) = \beta + \omega^{\delta-1}\cdot n < \beta + \omega^{\delta-1}\cdot m = \{\alpha\}(m)$$

et

$$\bigcup_{n\in\mathbb{N}}\{\alpha\}(n) = \beta + \omega^{\delta-1}\cdot\bigcup_{n\in\mathbb{N}} n = \beta+\omega^\delta = \alpha$$

Si δ est limite $\{\delta\}(n) < \{\delta\}(m)$ et donc :

$$\{\alpha\}(n) = \beta + \omega^{\{\delta\}(n)} < \beta + \omega^{\{\delta\}(m)} = \{\alpha\}(m)$$

et

$$\bigcup_{n\in\mathbb{N}} \{\alpha\}(n) = \beta + \omega^{\bigcup_{n\in\mathbb{N}}\{\delta\}(n)} = \beta + \omega^{\delta} = \alpha$$

□

Proposition 1.2.11. *Soient $\alpha, \beta < \epsilon_0$ tels que $\beta \neq 0$ et $\beta \ll \alpha$, alors :*

$$\forall n\ \{\alpha + \beta\}(n) = \alpha + \{\beta\}(n)$$

Démonstration. Si β est successeur, $\beta = \gamma+1$ et $\{\alpha+\gamma+1\}(n) = \alpha+\gamma = \alpha+\{\beta\}(n)$. Si β est limite, soit $\beta = \gamma + \omega^{\delta}$ sa forme normale de Cantor, par unicité de la forme normale de Cantor $\alpha + \gamma + \omega^{\delta}$ est la forme normale de $\alpha + \beta$ et donc

$$\{\alpha + \beta\}(n) = \alpha + \gamma + \omega^{\{\delta\}(n)} = \alpha + \{\beta\}(n)$$

ou

$$\{\alpha + \beta\}(n) = \alpha + \gamma + \omega^{\delta-1} \cdot n = \alpha + \{\beta\}(n)$$

suivant si δ est limite ou non. □

Pour finir on montre la proposition suivante qui relie les deux suites définies jusqu'ici :

Proposition 1.2.12. *Soit $\alpha < \epsilon_0$ un ordinal limite,*

$$\forall n \in \mathbb{N}\ \langle\{\alpha\}(n)\rangle(n) = \langle\alpha\rangle(n)$$

Démonstration. Par induction sur α.
Si $\alpha = 0$, $\forall n \in \mathbb{N}\ \langle\{0\}(n)\rangle(n) = \langle 0\rangle(n)$.
Sinon, soit $\alpha = \beta + \omega^{\delta}$ la forme normale de Cantor de α, on a pour tout $n \in \mathbb{N}$, si δ est limite :

$$\begin{aligned}
\langle\{\beta + \omega^{\delta}\}(n)\rangle(n) &= \langle\beta + \omega^{\{\delta\}(n)}\rangle(n) \\
&= \beta + \omega^{\langle\{\delta\}(n)\rangle(n)} \cdot (n-1) + \langle\omega^{\langle\{\delta\}(n)\rangle(n)}\rangle(n) \\
&= \beta + \omega^{\langle\delta\rangle(n)} \cdot (n-1) + \langle\omega^{\langle\delta\rangle(n)}\rangle(n) \\
&= \langle\beta + \omega^{\delta}\rangle(n)
\end{aligned}$$

si δ est successeur :

$$\begin{aligned}
\langle\{\beta + \omega^{\delta}\}(n)\rangle(n) &= \langle\beta + \omega^{\delta-1} \cdot n\rangle(n) \\
&= \langle\beta + \omega^{\delta-1} \cdot (n-1) + \omega^{\delta-1}\rangle(n) \\
&= \beta + \omega^{\langle\delta\rangle(n)} \cdot (n-1) + \langle\omega^{\langle\delta\rangle(n)}\rangle(n) \\
&= \langle\beta + \omega^{\delta}\rangle(n) \qquad \square
\end{aligned}$$

1.2.3 La relation de Ketonen et Solovay

Ketonen et Solovay introduisent dans [KS81] une relation sur les ordinaux qui s'appuie sur la suite canonique et qui sera utile pour démontrer la croissance de la hiérarchie de Hardy par rapport aux ordinaux en 1.2.17.

Définition 1.2.13 (Relation de Ketonen et Solovay)**.** Soient $\alpha, \beta < \epsilon_0$ et $k \in \mathbb{N}$, on note $\alpha \to_k \beta$ si il existe une suite d'ordinaux $(\gamma)_{i=0\dots n}$ telle que $\alpha = \gamma_0$, $\beta = \gamma_n$ et $\gamma_{i+1} = \{\gamma_i\}(k)$.

Proposition 1.2.14 (propriétés de la relation de Ketonen et Solovay)**.** *Soient $\alpha, \beta < \epsilon_0$ la relation de Ketonen et Solovay vérifie :*

(i) Si $\alpha = \gamma + 1$ alors pour tout $k \in \mathbb{N}$, $\alpha \to_k \gamma$. De plus si $\alpha \to_k \beta$ et $\alpha > \beta$ alors $\gamma \to_k \beta$.

(ii) Soient $x > y \in \mathbb{N}$, pour tout $k \in \mathbb{N}$, $\{\alpha\}(x) \to_k \{\alpha\}(y)$

(iii) Si $\alpha \to_k \beta$ alors pour tout $n \geqslant k$, $\alpha \to_n \beta$.

(iv) Si $\alpha \geqslant \beta$ alors il existe $k \in \mathbb{N}$ tel que $\alpha \to_k \beta$.

Démonstration. (i) Pour tout $k \in \mathbb{N}$ $\{\gamma + 1\}(k) = \gamma$ donc $\alpha \to_k \gamma$.
Si $\alpha \to_k \beta$ et $\alpha > \beta$, on a $\{\alpha\}(k) \to_k \beta$ par définition de $\to_k$.

(ii) Par induction sur α, $\{0\}(x) = 0 = \{0\}(y)$ et $\{\gamma + 1\}(x) = \gamma = \{\gamma + 1\}(y)$ donc si $\alpha = 0$ ou est successeur c'est évident.
Sinon, soit $\alpha = \beta + \omega^\delta$ sa forme normale de Cantor. Si δ est successeur,

$$\{\alpha\}(x) = \beta + \omega^{\delta-1} \cdot x = \beta + \omega^{\delta-1} \cdot y + \omega^{\delta-1} \cdot (x - y) = \{\alpha\}(y) + \omega^{\delta-1} \cdot (x - y)$$

Mais pour tout $k \in \mathbb{N}$, $\omega^{\delta-1} \cdot (x - y) \to_k 0$ car la suite définie par $\lambda_0 = \omega^{\delta-1} \cdot (x - y)$ et pour tout $n \in \mathbb{N}$, $\lambda_{n+1} = \{\lambda_n\}(k)$ est strictement décroissante tant que $\lambda_n \neq 0$. Comme les ordinaux sont bien fondés il existe $n \in \mathbb{N}$, $\lambda_n = 0$. C'est exactement la définition de $\omega^{\delta-1} \cdot (x - y) \to_k 0$.
Une conséquence immédiate de la proposition 1.2.11 est que si $\alpha \to_k \beta$ alors $\gamma + \alpha \to_k \gamma + \beta$ pour tout $\gamma \gg \alpha$. Donc $\{\alpha\}(x) \to_k \{\alpha\}(y)$.
Si δ est limite,

$$\{\alpha\}(x) = \beta + \omega^{\{\delta\}(x)} \to_k \beta + \omega^{\{\delta\}(y)} = \{\alpha\}(y)$$

car si $\lambda \to_k \mu$ alors il existe $(\gamma_i)_{i=0\dots n}$ tels que $\gamma_0 = \lambda$,$\gamma_n = \mu$ et $\gamma_{i+1} = \{\gamma_i\}(k)$ donc en posant $\kappa_i = \omega^{\lambda_i}$ on a $\{\kappa_i\}(k) = \omega^{\{\lambda_i\}(k)} = \omega^{\lambda_{i+1}} = \kappa_{i+1}$ et donc $\omega^\lambda \to_k \omega^\mu$. Comme par induction $\{\delta\}(x) \to_k \{\delta\}(y)$ on conclut.

(iii) Soit $\alpha \to_k \beta$ si $\alpha = \beta$ alors pour tout $n \in \mathbb{N}$ $\alpha \to_n \beta$ sinon $\{\alpha\}(k) \to_k \beta$. Soit dans ce cas $n > k$. Par le (ii) on a $\{\alpha\}(n) \to_n \{\alpha\}(k)$ et comme par définition $\alpha \to_n \{\alpha\}(n)$, par transitivité $\alpha \to_n \beta$.

(iv) Fixons β et montrons par induction sur $\alpha \geqslant \beta$ que la proposition est vérifiée.
Si $\alpha = \beta$ alors pour tout $k \in \mathbb{N}$, $\alpha \to_k \beta$.
Si $\alpha = \gamma + 1$, on a $\gamma \geqslant \beta$ et donc par induction $\exists k \in \mathbb{N}$ $\gamma \to_k \beta$. Or par le (i), $\alpha \to_k \gamma$ et donc par transitivité $\alpha \to_k \beta$.

Si α est limite, comme $\bigcup_{n\in\mathbb{N}}\{\alpha\}(n) = \alpha > \beta$, il existe $n_0 \in \mathbb{N}$, $\{\alpha\}(n_0) > \beta$. Par induction et en utilisant le (iii), il existe $n_1 \in \mathbb{N}$ tel que $\forall n \geqslant n_1$ $\{\alpha\}(n_0) \to_n \beta$ et $\forall n \geqslant n_0$ $\alpha \to_n \{\alpha\}(n_0)$. En prenant $k \geqslant \max(n_0, n_1)$ on a $\alpha \to_k \{\alpha\}(n_0)$ et $\{\alpha\}(n_0) \to_k \beta$ et on conclut par transitivité.

□

1.2.4 Hiérarchie de Hardy

On aura besoin par la suite d'une famille de fonctions, la hiérarchie de Hardy, indicée par les ordinaux appartenant à ϵ_0. Sa principale caractéristique est de permettre d'énoncer une condition nécessaire pour qu'une fonction récursive soit prouvablement totale dans P.

Définition 1.2.15 (Hiérarchie de Hardy)**.** $(H_\alpha)_{\alpha<\epsilon_0}$ est la famille de fonctions $\mathbb{N} \to \mathbb{N}$ définie par induction sur $\alpha < \epsilon_0$ en posant :

$$\begin{cases} H_0(x) &= x \\ H_\alpha(x) &= H_{\alpha-1}(x+1) & \text{si } \alpha \text{ est successeur} \\ H_\alpha(x) &= H_{\{\alpha\}(x)}(x) & \text{si } \alpha \text{ est limite} \end{cases}$$

Proposition 1.2.16. *Pour tout $\alpha, \beta < \epsilon_0$ tels que $\alpha \gg \beta$, $H_{\alpha+\beta}(x) = H_\alpha \circ H_\beta(x)$. En particulier, si on note $f^{(0)} = \mathrm{Id}$ et $\forall n \in \mathbb{N}$ $f^{(n)} = f \circ f^{(n+1)}$, alors :*

$$H_{\omega^\alpha \cdot n}(x) = H^{(n)}_{\omega^\alpha}(x)$$

Démonstration. Par induction sur β :
Si $\beta = 0$, $H_{\alpha+0}(x) = H_\alpha \circ \mathrm{Id}(x) = H_\alpha \circ H_\beta(x)$
Si β est successeur, $\beta = \gamma + 1$ et

$$H_{\alpha+\gamma+1}(x) = H_{\alpha+\gamma}(x+1) = H_\alpha \circ H_\gamma(x+1) = H_\alpha \circ H_\beta(x)$$

Si β est limite alors $\alpha + \beta$ est limite et

$$H_{\alpha+\beta}(x) = H_{\{\alpha+\beta\}(x)}(x) = H_{\alpha+\{\beta\}(x)} = H_\alpha \circ H_{\{\beta\}(x)}(x) = H_\alpha \circ H_\beta(x)$$

Montrons alors $H_{\omega^\alpha \cdot n}(x) = H^{(n)}_{\omega^\alpha}(x)$ par récurrence sur n.
Si $n = 0$, $H_0(x) = x = H^{(0)}_{\omega^\alpha}(x)$.
Sinon en supposant $H_{\omega^\alpha \cdot n}(x) = H^{(n)}_{\omega^\alpha}(x)$, on a :

$$\begin{aligned} H_{\omega^\alpha \cdot (n+1)}(x) &= H_{\omega^\alpha \cdot n + \omega^\alpha}(x) \\ &= H_{\omega^\alpha \cdot n} \circ H_{\omega^\alpha}(x) \\ &= H^{(n)}_{\omega^\alpha} \circ H_{\omega^\alpha}(x) \\ &= H^{(n+1)}_{\omega^\alpha}(x) \end{aligned}$$

□

La proposition suivante donne divers résultats de croissance de la Hiérarchie de Hardy en fonction de ses deux paramètres, l'ordinal et l'entier.

Proposition 1.2.17 (Croissance de la hiérarchie de Hardy)**.** *Soient $f, g : \mathbb{N} \to \mathbb{N}$, on note $f <_* g$ s'il existe $x_0 \in \mathbb{N}$ tel que $\forall x > x_0\ f(x) < g(x)$. On vérifie aisément que cette relation est transitive.*

(i) Pour tout $\alpha < \epsilon_0$, H_α est une fonction strictement croissante.

(ii) Soient $\beta < \alpha < \epsilon_0$, si $\alpha \to_k \beta$ pour un certain k alors

$$\forall x \geqslant k,\ H_\alpha(x) > H_\beta(x)$$

(iii) Soient $\alpha < \beta < \epsilon_0$, alors $H_\alpha <_* H_\beta$.

Démonstration. On montre (i) et (ii) simultanément par induction sur α.
Si $\alpha = 0$, $H_\alpha = \mathrm{Id}$ est strictement croissante. Pour (ii), il n'y a pas de $\beta < \alpha$ donc le cas de base est démontré.
Si $\alpha = \gamma + 1$ alors $H_\alpha(x) = H_\gamma(x+1)$ et comme, par induction, H_γ est strictement croissante, H_α aussi.
Soit $\beta < \alpha$ et $k \in \mathbb{N}$, on a soit $\beta = \gamma$, car pour tout $k \in \mathbb{N}$, $\alpha \to_k \gamma$ d'après la proposition 1.2.14.(i), soit $\beta < \gamma$. Dans le premier cas, comme H_γ est strictement croissante, pour tout $x > k$ on a

$$H_\alpha(x) = H_\gamma(x+1) > H_\gamma(x)$$

Si $\beta < \gamma$ et $\alpha \to_k \beta$ alors par la proposition 1.2.14.(i) $\gamma \to_k \beta$ et donc par induction

$$\forall x \geqslant k,\ H_\beta(x) < H_\gamma(x) < H_\alpha(x)$$

Si α est limite alors $H_\alpha(x) = H_{\{\alpha\}(x)}$. Soient $x > y$, par la proposition 1.2.14.(ii) on a $\{\alpha\}(x) \to_y \{\alpha\}(y)$, et donc par le (ii) appliqué à $\{\alpha\}(x)$ et la croissance de $H_{\{\alpha\}(y)}$,

$$H_\alpha(x) = H_{\{\alpha\}(x)}(x) > H_{\{\alpha\}(y)}(x) > H_{\{\alpha\}(y)}(y) = H_\alpha(y)$$

Pour le (ii), soit $k \in \mathbb{N}$ et $\beta < \alpha$ tel que $\alpha \to_k \beta$ alors pour tout $x \geqslant k$, $\alpha \to_x \beta$. Mais comme $\beta < \alpha$, on a $\{\alpha\}(x) \to_x \beta$. Donc par induction

$$H_\alpha(x) = H_{\{\alpha\}(x)}(x) > H_\beta(x)$$

Le (i) et le (ii) sont donc démontrés.
Le (iii) découle du (ii). En effet soient $\beta < \alpha < \epsilon_0$, par la proposition 1.2.14.(iv) il existe k tel que $\alpha \to_k \beta$ et donc $\forall x \geqslant k\ H_\alpha(x) > H_\beta(x)$ □

1.2.5 Théorème de Cichon

Le résultat qui suit donne une formule de la fonction de Goodstein en fonction de la hiérarchie de Hardy. On remarque que les arguments de h apparaissent non seulement comme arguments de H_α mais aussi pour définir α. Ainsi h croît plus vite que n'importe quelle fonction de la hiérarchie de Hardy (en un sens qu'il faudra formaliser vu que h prend 2 variables). Elle croît en quelque sorte trop vite pour être prouvablement totale comme l'indique le théorème de Wainer (théorème 1.4.1 de ce mémoire).

Théorème 1.2.18 (Cichon, 1983)**.** *Soit $\alpha = f_{q,\omega}(p)$. Alors*

$$h(p,q) = H_\alpha(q+1) - (q+1)$$

Démonstration. Par la proposition 1.2.8, on a :

$$h(p,q) = \min\{n \mid \langle\alpha\rangle(q+1,\ldots,q+n) = 0\}$$

Il suffit donc de montrer par induction sur α que pour tout $q \in \mathbb{N}$ on a

$$\min\{n \mid \langle\alpha\rangle(q+1,\ldots,q+n) = 0\} = H_\alpha(q+1) - (q+1)$$

Si $\alpha = 0$, on a bien $0 = H_0(q+1) - (q+1)$.
Si α est successeur, $\alpha = \beta + 1$, $H_\alpha(q+1) = H_\beta(q+2)$ et $\langle\alpha\rangle(q+1) = \beta$ donc

$$\begin{aligned}
\min\{n \mid \langle\alpha\rangle(q+1,\ldots,q+n) = 0\} &= \min\{n \mid \langle\beta\rangle(q+2,\ldots,q+n) = 0\} \\
&= \min\{n \mid \langle\beta\rangle(q+2,\ldots,q+1+n) = 0\} + 1 \\
&= H_\beta(q+2) - (q+2) + 1 \\
&= H_\alpha(q+1) - (q+1)
\end{aligned}$$

Si α est limite, par la proposition 1.2.12, $\langle\alpha\rangle(q+1) = \langle\{\alpha\}(q+1)\rangle(q+1)$ donc

$$\begin{aligned}
\min\{n \mid \langle\alpha\rangle(q+1,\ldots,q+n) = 0\} &= \min\{n \mid \langle\{\alpha\}(q+1)\rangle(q+1,\ldots,q+n) = 0\} \\
&= H_{\{\alpha\}(q+1)}(q+1) - (q+1) \\
&= H_\alpha(q+1) - (q+1)
\end{aligned}$$

□

1.3 Un indicateur des segments initiaux qui vérifient P

Par la suite $\mathbb{M}$ est un modèle de P et on l'identifie à M, la structure sous-jacente. On note $\mathbb{N}$ le modèle standard de P et on identifie $\mathbb{N}$ et le segment initial de $\mathbb{M}$ constitué des entiers standards. On notera de plus $a > \mathbb{I}$ si $\forall b \in \mathbb{I}$, $a > b$. Enfin, pour des raisons de concision, on notera $\iff$ et $\Rightarrow$ l'équivalence et l'implication dans les démonstrations, qu'il ne faudra pas confondre avec $\rightarrow$ et $\leftrightarrow$ les symboles de $\mathcal{L}_{\mathrm{P}}$.
Pour ce qui est du codage, les ensembles seront codés comme le produit des nombres premiers dont le numéro est dans l'ensemble, si ce produit existe (c'est à dire dans le cas des ensembles finis dans $\mathbb{N}$), et les fonctions seront codées par le code de la formule qui les représente.

1.3.1 Segments initiaux forts

La définition suivante est une caractérisation de certains segments initiaux des modèles de P, pour plus de détails sur ces derniers se référer à [KP76].

Définition 1.3.1 (Segments initiaux forts)**.** Soit $\mathbb{I}$ un segment initial propre de $\mathbb{M}$, ie $\mathbb{I} \neq \emptyset$, $\mathbb{I} \neq \mathbb{M}$ et si $a \in \mathbb{I}$ alors pour tout $b < a$, $b \in \mathbb{I}$. On dit qu'il est fort s'il est clos par successeur et si pour tout $f : \mathbb{I} \rightarrow \mathbb{M}$ que l'on peut coder dans $\mathbb{M}$, il existe $e > \mathbb{I}$ tel que pour tout $a \in \mathbb{I}$,

$$f(a) \in \mathbb{I} \vee f(a) > e$$

La proposition qui suit indique qu'être un segment initial fort est une propriété très forte car elle est suffisante pour être un modèle de P.

Proposition 1.3.2 (Segments initiaux forts et modèles de P)**.** *Si* $\mathbb{I}$ *est un segment initial fort de* $\mathbb{M}$ *alors* $\mathbb{I} \models \mathrm{P}$*.*

Démonstration. Montrons d'abord que $\mathbb{I}$ est clos par addition et multiplication.
Supposons qu'il existe $u, v \in \mathbb{I}$ tels que $u + v \notin \mathbb{I}$. Soit $f(x) = u + x$ codable dans $\mathbb{M}$, il existe, par définition de I, $e > \mathbb{I}$ tel que pour tout $x \in \mathbb{I}$, $f(x) \in \mathbb{I}$ ou $f(x) \geqslant e$. En particulier $u + v \geqslant e$.
Soit $x = e - 1 - u$, alors $x \geqslant v - 1$ donc $x \in \mathbb{I}$ mais $f(x) = e - 1 < e$. Comme $\mathbb{I}$ est clos par successeur $e - 1 \notin \mathbb{I}$ sinon e le serait aussi et on a $x \in \mathbb{I}$ tel que $\mathbb{I} < f(x) < e$ ce qui est absurde.
Pour ce qui est de la multiplication, si $u, v \in \mathbb{I}$ et $u * v \notin \mathbb{I}$ on considère $f(x) = u * x$ et il existe $e > I$ tel que $f(x) \in \mathbb{I}$ ou $f(x) > e$. Comme $f(v) > e$, $M \models \exists x\ u * x > e$ et par le principe du minimum (qui est équivalent au principe de récurrence) appliqué dans $\mathbb{M}$ il existe v_0 minimal tel que $u * v_0 > e$, donc $f(v_0 - 1) \leqslant e$ mais comme $v_0 \leqslant v \in \mathbb{I}$, on a $v_0 \in \mathbb{I}$ et donc $f(v_0) \in \mathbb{I}$. Mais $u * v_0 = u * (v_0 - 1) + u$ cela contredit donc la clôture par addition de $\mathbb{I}$.
De plus $\mathbb{I}$ vérifie les axiomes de P_0, Peano faible, car toutes les opérations dans $\mathbb{I}$ peuvent être vues comme des opérations dans $\mathbb{M}$.
Il reste à démontrer que $\mathbb{I}$ vérifie le schéma de récurrence. On démontre plutôt, ce qui est strictement équivalent, que $\mathbb{I}$ vérifie le schéma du minimum : soit F une formule de $\mathcal{L}_{\mathrm{P}}$ et $\bar{a}^n = (a_i)_{i=0\ldots n} \in \mathbb{I}^n$ alors

$$(\mathbb{I} \models \exists x\ F[\bar{a}^n, x]) \Rightarrow (\mathbb{I} \models \exists x\ (F[\bar{a}^n, x] \wedge \forall y < x\ \neg F[\bar{a}^n, y]))$$

Pour cela on définit la transformation suivante sur les formules. Soit $F[\bar{x}^n]$ une formule, il existe $F^*[\bar{x}^n, \bar{y}^m]$ dont tous les quantificateurs sont bornés et $\bar{b}^m \in \mathbb{M}^m$ tels que pour tout $\bar{a}^n \in \mathbb{I}^n$

$$(\mathbb{I} \models F[\bar{a}^n]) \iff (\mathbb{M} \models F^*[\bar{a}^n, \bar{b}^m])$$

On construit F^* et $\bar{b}^m$ par induction sur la formule F.
Si F est atomique $F^* = F$ convient. De même $(\neg F)^* = \neg F^*$ et $(F \to G)^* = F^* \to G^*$ en gardant les même $\bar{b}^m$.
Reste le cas $F[\bar{x}^n] = \forall z\ G[\bar{x}^n, z]$. Supposons qu'on ait déjà G^* et le $\bar{b}^m$ qui lui correspond. Posons alors $f : \mathbb{I} \to \mathbb{M}$ tel que, si on note $\lfloor \bar{a}^n \rfloor$ le code dans $\mathbb{I}$ du n-uplet $\bar{a}^n$:

$$f(\lfloor \bar{a}^n \rfloor) = \min\{z \mid \neg G^*[\bar{a}^n, z, \bar{b}^m]\} \text{ s'il existe}, b > \mathbb{I} \text{ quelconque sinon}$$

Cette fonction est codable dans $\mathbb{M}$ donc par définition d'un segment initial fort, $\exists e > \mathbb{I}$ tel que pour tout $\bar{a}^n \in \mathbb{I}^n$, $f(\lfloor \bar{a}^n \rfloor) \in \mathbb{I}$ ou $f(\lfloor \bar{a}^n \rfloor) > e$. On pose alors $F^* = \forall z < e\ G^*$ et on a bien pour tout $\bar{a}^n \in \mathbb{I}^n$:

$$\begin{aligned} \mathbb{I} \models \forall z\ G[\bar{a}^n, z] &\iff f(\lfloor \bar{a}^n \rfloor) > e \\ &\iff M \models \forall z < e\ G^*[\bar{a}^n, z] \end{aligned}$$

Montrons maintenant que $\mathbb{I}$ vérifie le schéma du minimum. Soit une formule $F[x, \bar{y}^n]$ et $\bar{a}^n \in \mathbb{I}^n$ tels que $\mathbb{I} \models \exists x\ F[x, \bar{a}^n]$. Soit $a \in \mathbb{I}$ tel que $\mathbb{I} \models F[a, \bar{a}^n]$ alors $\mathbb{M} \models F^*[a, \bar{a}^n, \bar{b}^m]$

et donc comme $\mathbb{M} \models \mathrm{P}$, il vérifie le schéma du minimum et donc il existe un a_0 minimum tel que $\mathbb{M} \models F^*[a_0, \bar{a}^n, \bar{b}^m]$. Comme $a_0 < a \in \mathbb{I}$ et que $\mathbb{I}$ est un segment initial on a $a_0 \in \mathbb{I}$ et donc $\mathbb{I} \models F[a_0, \bar{a}^n]$ et a_0 est bien le plus petit $a \in \mathbb{I}$ vérifiant cette propriété. Ce qui montre que $\mathbb{I}$ vérifie le schéma d'induction et donc $\mathbb{I} \models \mathrm{P}$. □

1.3.2 Un indicateur de segments initiaux forts

Pour une définition plus combinatoire de la notion suivante, on pourra consulter [Par79]. Les résultats de [KS81] et [Par79] montrent qu'elles sont équivalentes pour les intervalles.

Définition 1.3.3 (Ensembles α-larges)**.** Soit $S \subset \mathbb{M}$, soit $x^+ = \min\{y \in S \mid y > x\}$, pour tout $x \in S$. On définit de nouvelles fonctions de Hardy pour tout $\alpha < \epsilon_0$ définies partiellement sur S par :

$$\begin{cases} H_0^S(x) & = & x \\ H_\alpha^S(x) & = & H_{\alpha-1}^S(x^+) \quad \text{si } \alpha \text{ est successeur} \\ H_\alpha^S(x) & = & H_{\{\alpha\}(x)}^S(x) \quad \text{si } \alpha \text{ est limite} \end{cases}$$

Comme pour les fonctions de Hardy normales, $H_{\omega^\alpha.n}^S(x) = (H_{\omega^\alpha}^S)^{(n)}(x)$ à condition que le terme de gauche soit bien défini. La preuve est identique.
Si S est un ensemble borné et $\alpha < \epsilon_0$, soit $a = \min S$ et $b = \max S$, on dit que S est α-large si

$$H_\alpha^S(a) \leqslant b$$

Soient $a, b \in \mathbb{M}$, pour tout $\alpha < \epsilon_0$ $H_\alpha^{[\![a;b]\!]}(a) = H_\alpha(a)$ et donc $[\![a;b]\!]$ est α-large si

$$H_\alpha(a) \leqslant b$$

Soit S un ensemble codable dans $\mathbb{M}$. Notons $H^0[x, y, z, t]$ la formule qui représente $H_\alpha^S(x)$ (ie en notant $\lfloor \alpha \rfloor \in \mathbb{N}$ le code de l'ordinal α et $\lfloor S \rfloor \in \mathbb{N}$ le code de S, on a pour tout $x \in \mathbb{M}$, $\mathbb{M} \models H^0[x, \lfloor \alpha \rfloor, \lfloor S \rfloor, z] \iff H_\alpha^S(x) = z$) et notons $\omega_n^k = \{\omega_{n+1}\}(k)$, en notant Min et Max les formules représentant le min et le max de l'ensemble codé par x, la formule :

$$A[x, n, k] = \exists t, u, v \ (H^0[u, \lfloor \omega_n^k \rfloor, x, t] \wedge z \geqslant v \wedge Min[x, u] \wedge Max[x, v])$$

exprime exactement que l'ensemble codé par x est ω_n^k-large.
On a aussi, en notant H[x,y,z] la formule qui représente $H_\alpha(x)$:

$$L[a, b, n, k] = \exists z \ (H[a, \lfloor \omega_n^k \rfloor, z] \wedge z \leqslant b)$$

qui exprime exactement que $[\![a;b]\!]$ est ω_n^k-large.

La proposition suivante peut prendre une forme plus générale en la formulant en terme d'indicateur, c'est à dire que si on note $Y(a, b) = \max\{c \mid [\![a;b]\!] \text{ est } \omega_c\text{-large } \}$ alors c'est un indicateur des modèles de P, c'est à dire que si $Y(a, b) > \mathbb{N}$, il existe un segment initial qui est un modèle de P tel que $a \in \mathbb{I} < b$ (voir [Par79] et [Par78]). On la démontre ici formulée un peu différemment, dans le cas d'un modèle dénombrable tel que $\mathbb{M} \models \mathrm{Th}(\mathbb{N})$ (où $\mathrm{Th}(\mathbb{N})$ est l'ensemble des énoncés vrais dans $\mathbb{N}$) et sans démontrer que l'indicateur Y est bien une fonction.

Proposition 1.3.4. *Soit* $\mathbb{M} \models \mathrm{Th}(\mathbb{N})$ *dénombrable et* $a, b \in \mathbb{M}$ *tel que pour tout* $n, k \in \mathbb{N}$ *on ait* $\mathbb{M} \models L[a, b, n, k]$ *alors il existe* $\mathbb{I}$ *segment initial fort de* $\mathbb{M}$ *tel que* $a \in \mathbb{I} < b$

Pour démontrer ce résultat, on a adapté la démonstration de [AZ97] relativement besogneuse mais ne nécessitant pas de théorie supplémentaire. On va donc d'abord définir la notion d'ensemble approximant une fonction avant de démontrer trois lemmes sur ces approximations. On démontre les trois lemmes dans $\mathbb{N}$ mais comme on considère $\mathbb{M} \models \mathrm{Th}(\mathbb{N})$, ils seront aussi vrais dans $\mathbb{M}$.

Définition 1.3.5 (Approximation de fonction)**.** Soit $f : \mathbb{M} \to \mathbb{M}$ une fonction partielle, et $S \subset \mathbb{M}$ borné. On dit que S est une approximation de f si pour tout $x \in S \backslash \{\max S\}$ et tout $y < x - 2$:

$$y \in \mathrm{dom}(f) \Rightarrow (f(y) < x^+ \text{ ou } f(y) \geqslant \max S)$$

Lemme 1.3.6. *Soit* $S \subset \mathbb{N}$ *codable* $\omega^{\alpha+1}$*-large et* $f : \mathbb{N} \to \mathbb{N}$ *une fonction partielle. Supposons que* $\min S = s_0 > 0$*. Alors il existe* $a > s_0$ *et* $S' \subset S$ *codable tels que* S' *soit* ω^α*-large,* $\min S' = a$ *et pour tout* $x < s_0 - 2$ *dans* $\mathrm{dom}(f)$ *:*

$$f(x) < a \text{ ou } f(x) \geqslant \max S'$$

Démonstration. Notons $\min S = a_0$ et $\max S = b_0$. Par définition $H_{\omega^{\alpha+1}}(a_0) \leqslant b_0$ or $\{\omega^{\alpha+1}\}(a_0) = \omega^\alpha \cdot a_0$. Donc $(H^S_{\omega^\alpha})^{(a_0)} = b_0$.
Soit $(a_i)_{i=0\ldots s_0}$ tels que pour tout $i \geqslant 0$ $a_{i+1} = H^S_{\omega^\alpha \cdot a_0}(a_i)$. On a alors $a_{a_0} \leqslant b$.
Comme $f([\![0; a_0 - 3]\!])$ contient au plus $a_0 - 2$ valeurs et qu'il y a $a_0 - 1$ intervalles $[\![a_i; a_{i+1}[\![$ pour $i > 0$, l'un d'eux ne contient aucun $f(x)$ pour $x < a_0 - 2$. Nommons le i correspondant i_0. Alors $a = a_{i_0} > a_0$ et $S' = [\![a_{i_0}; a_{i_0+1}[\![$ conviennent (il est codable car S' et $[\![a_{i_0}; a_{i_0+1}[\![$ le sont).
En effet pour tout $\alpha < \epsilon_0$ et $a, b \in \mathbb{N}$ on vérifie par induction que $H^{S \cap [\![a;b[\![}_\alpha = H^S_\alpha|_{[\![a;b[\![}$. S' est bien ω^α-large. □

Lemme 1.3.7. *Soit* S *codable et* ω^α*-large,* $\min S > 0$ *et* $f : \mathbb{N} \to \mathbb{N}$ *une fonction partielle. Il existe alors* $S' \subset S$ *codable tel que* S' *soit une approximation de* f*,* S' α*-large et* $\min S' = \min S$*.*

Démonstration. Par induction sur α.
Si $\alpha = 0$ alors S est 1-large. Soit $a_0 = \min S$ alors $S' = \{a_0\}$ est 0-large et S' est une approximation de f vu que l'on quantifie sur le vide dans la définition.
Si α successeur, alors $\alpha = \gamma + 1$. Soient S_1 et $a = \min S_1 > \min S$ tels que dans le lemme 1.3.6. S_1 est alors ω^γ-large et par induction, il existe $S_2 \subset S_1$ codable γ-large tel que $\min S_2 = \min S_1 = a$ et S_2 est une approximation de f. Vérifions que $S_3 = \{\min S\} \cup S_2 \subset S$ convient (il est codable car S_2 l'est).
$\min S_3 = \min S$ par définition. De plus si on note $a_0 = \min S_3$

$$H^{S_3}_{\gamma+1}(a_0) = H^{S_3}_\gamma(a_0^+) = H^{S_3}_\gamma(a) \leqslant \max S_2 = \max S_3$$

donc S_3 est $(\gamma + 1)$-large.
Enfin, soit $a \in S_3 \backslash \{s_n\}$ si $a \neq s_0$ alors comme S_2 est une approximation de f, on a bien $\forall x < a - 2 f(x) < a^+$ ou $f(x) \geqslant s_n = \max S_2$ sinon $a = s_0$ et comme S_1 vérifie que pour

tout $x < \min S - 2 = s_0 - 2$ dans $\mathrm{dom}(f)$, $f(x) < a = s_1$ ou $f(x) \geqslant \max S'$ on a bien que S_3 est une approximation de f.
Si α est limite, notons $\min S = a_0$. Comme $\{\omega^\alpha\}(a_0) = \omega^{\{\alpha\}(a_0)}$,

$$H^S_{\omega^\alpha}(a_0) = H^S_{\{\omega^\alpha\}(a_0)}(a_0) = H_{\omega^{\{\alpha\}(a_0)}}(a_0)$$

donc S est $\omega^{\{\alpha\}(s_0)}$-large.
Par induction il existe $S' \subset S$ codable et $\{\alpha\}(a_0)$-large tel que $\min S' = a_0$ et tel que S' soit une approximation de f. Mais S' est aussi α-large car

$$H^{S'}_\alpha(a_0) = H^{S'}_{\{\alpha\}(a_0)}(a_0) \leqslant \max S' \qquad \square$$

Lemme 1.3.8. *Soit S_0 codable dans $\mathbb{N}$. Alors pour toutes $f_1, \ldots, f_n$ fonctions partielles codables dans $\mathbb{N}$ il existe $S_1 \ldots S_n$ codables tels que pour tout $i > 0$, $S_i \subset S_{i-1}$, S_i est une approximation de f_i et $\mathrm{card}\, S_i > k$*
En particulier pour tous $n, k \in \mathbb{N}$ et x code d'ensemble,

$$\begin{aligned}\mathbb{N} \models A[x, n, k] \rightarrow \quad & \forall y \text{ code de fonction } \exists z \text{ code d'ensemble} \\ & (z \subset x \wedge z \text{ approxime } y \wedge A[z, n-1, k] \wedge \mathrm{card}\, z > k)\end{aligned}$$

De plus comme les intervalles sont finis dans $\mathbb{N}$, leurs sous-ensembles sont codables et on a donc pour tous $a, b \in \mathbb{N}$:

$$\begin{aligned}\mathbb{N} \models L[a, b, n, k] \rightarrow \quad & \forall y \text{ code de fonction } \exists z \text{ code d'ensemble} \\ & (z \subset \lfloor [\![a; b]\!] \rfloor \wedge z \text{ approxime } y \wedge A[z, n-1, k] \wedge \mathrm{card}\, z > k])\end{aligned}$$

Démonstration. Remarquons d'abord que pour tout $n > 0$,

$$\omega^k_n = \{\omega^{\omega_n}\}(k) = \omega^{\{\omega_n\}(k)} = \omega^{\omega^k_{n-1}}$$

et $\omega^k_0 = k$.
Montrons ce lemme par récurrence sur n.
Si $n = 1$, par le lemme 1.3.7, il existe $S_1 \subset S_0$ ω^k_0-large tel que S_1 approximation de f_1, et en notant $S = \{s_1, \ldots, s_n\}$

$$H_k(s_1) = H_{k-1}(s_2) = \cdots = H_0(s_{k+1}) = s_{k+1} \leqslant s_n$$

donc $\mathrm{card}\, S_1 = n \geqslant k + 1 > k$.
Sinon, toujours par le lemme 1.3.7, il existe $S_1 \subset S_0$ ω^k_{n-1}-large tel que S_1 approximation de f_1. Par récurrence il existe $S_2, \ldots, S_n$ vérifiant les bonnes propriétés. Comme $S_2 \subset S_1$ et $\mathrm{card}\, S_2 > k$ on a encore $\mathrm{card}\, S_1 > k$. $\square$

On peut maintenant construire le segment initial fort dont on a besoin.

Démonstration de la proposition 1.3.4. Pour tous $n, k \in \mathbb{N}$, $\mathbb{M} \models L[a, b, n, k]$. En particulier pour tout $n \in \mathbb{N}$, $\mathbb{M} \models L[a, b, n, n]$. Par le lemme du débordement il existe $c > \mathbb{N}$ tel que $\mathbb{M} \models L[a, b, c, c]$.
Soit $(f_i)_{i \geqslant 1}$ une énumération des fonctions codables dans $\mathbb{M}$ dénombrable. On construit par récurrence $(S_i)_{i \geqslant 1}$ telle que S_i approxime f_i, $\mathbb{M} \models A[\lfloor S_i \rfloor, c - i, c]$ et $\mathrm{card}\, S_i > c$ à

l'aide du lemme 1.3.8. Comme c est non standard $\operatorname{card} S_i > c$ implique que S_i est infini. Soit d_i le i$^{\text{e}}$élément de S_i, on définit $\mathbb{I} = \{x \in \mathbb{M} \mid \exists i \in \mathbb{N} x \leqslant d_i\}$.
Comme $S_{i+1} \subset S_i$, $d_i + 1 \leqslant d_{i+1}$, on en déduit que $\mathbb{I}$ est clos par successeur. De plus si $a \in \mathbb{I}$ on a clairement pour tout $b \leqslant a$, $b \leqslant a \leqslant d_i$ pour un certain i donc $b \in \mathbb{I}$.
Montrons que $\mathbb{I}$ est fort. Soit f une fonction codable dans $\mathbb{M}$, alors il existe i_0 tel que $f = f_{i_0}$. Posons $e = \max S_i$. Soit $x \in \mathbb{I}$, il existe j tel que $x \leqslant d_j$. Par le même argument que pour le successeur, $x < d_{j+3} - 2$. Il existe donc $k \geqslant i$ tel que $x < d_k - 2$. Comme $S_k \subset S_i$, on a $d_k \in S_i$. Or S_i approxime f donc $f(x) < d_k^{+,S_i}$ ou $f(x) > \max S_i = e$. Il suffit donc de démontrer que d_k^{+,S_i} est dans $\mathbb{I}$.
Mais comme $S_{k+1} \subset S_k \subset S_i$, on a $d_k^{+,S_i} \leqslant d_k^{+,S_{k+1}} = d_{k+1}$, donc $d_k^{+,S_i} \in \mathbb{I}$. □

1.4 Théorème de Wainer

On peut maintenant démontrer le résultat probablement le plus important de ce mémoire, le théorème de Wainer qui caractérise les fonctions récursives prouvablement totales dans P (la réciproque du théorème se trouve dans [BW87]). Le théorème de Kirby et Paris en est un corollaire.

Théorème 1.4.1 (Wainer, 1970)**.** *Soit $f : \mathbb{N} \to \mathbb{N}$ une fonction récursive. Soit $F[x, y]$ une formule Σ_1 qui représente f. Si $\mathrm{P} \vdash \forall x \exists! y\ F(x, y)$ (ie. f est prouvablement totale dans P), alors*

$$\exists \alpha < \epsilon_0\ \forall x \in \mathbb{N}\ f(x) \leqslant H_\alpha(x)$$

On commence par démontrer le lemme suivant :

Lemme 1.4.2. *Soient $f : \mathbb{N} \to \mathbb{N}$ récursive et $F[x, y]$ qui représente f. Si pour tout $n, k \in \mathbb{N}$ on a*

$$\mathbb{N} \models \exists x \exists y\ (L[x, y, n, k] \wedge F[x, y])$$

alors il existe $\mathbb{M}$ modèle dénombrable de P et $a, b \in \mathbb{M}$ tels que :

$$\mathbb{M} \models \mathrm{Th}(\mathbb{N}) \text{ et } \mathbb{M} \models F[a, b] \text{ et } \forall n, k \in \mathbb{N}\ \mathbb{M} \models L[a, b, n, k]$$

Démonstration. Considérons la théorie $T = \mathrm{Th}(\mathbb{N}) \cup \{F[a, b]\} \cup \{L[a, b, n, k]\}_{n,k \in \mathbb{N}}$ dans le langage de P augmenté de deux constantes a et b.
Pour tout $n, k \in \mathbb{N}$, $\mathbb{N} \models \exists x \exists y\ \bigwedge_{k' \leqslant k, n' \leqslant n} L[x, y, n', k']$. En effet en notant $a_{n,k}$ tel que $\mathbb{N} \models \exists y\ F[a_{n,k}, y] \wedge L[a_{n,k}, y, n, k]$, $\{H_{\omega_{n'}^{k'}}(a_{n',k'}) \mid n' \leqslant n \text{ et } k' \leqslant k\}$ est fini. Il admet donc un plus grand élément atteint en n_0 et k_0. On en déduit que $[\![a_{n_0,k_0}; f(a_{n_0,k_0})]\!]$ est ω_n^k-large pour tout $n' \leqslant n$ et $k' \leqslant k$.
La théorie T admet donc $\mathbb{N}$ comme modèle pour tout fragment fini. Par le théorème de compacité, il existe donc $\mathbb{M} \models T$. Par le théorème de Löwenheim-Skolem descendant, comme le cardinal de $\mathcal{L}_\mathrm{P} \cup \{a, b\}$ est fini, on peut prendre $\mathbb{M}$ dénombrable. □

Démonstration du théorème de Wainer. Démontrons ce résultat par l'absurde.
Soit $f : \mathbb{N} \to \mathbb{N}$ récursive prouvablement totale et F qui la représente telle que pour tout $\alpha < \epsilon_0$ on ait $x \in \mathbb{N}$ tel que $f(x) > H_\alpha(x)$, en particulier

$$\forall\ n, k \in \mathbb{N}\ \exists x \in \mathbb{N}\ f(x) > H_{\omega_n^k}(x)$$

Par définition (voir 1.3.3), $[\![x; f(x)]\!]$ est ω_n^k-large. On en déduit que

$$\mathbb{N} \models \forall n \forall k \exists x \exists y \ (L[x, y, n, k] \wedge F[x, y])$$

donc par le lemme 1.4.2 il existe $\mathbb{M}$ un modèle dénombrable de P et $a, b \in \mathbb{M}$ tel que

$$\mathbb{M} \models \mathrm{Th}(\mathbb{N}) \text{ et } \mathbb{M} \models F[a, b] \text{ et } \forall n, k \in \mathbb{N}, \ \mathbb{M} \models L[a, b, n, k]$$

Par la proposition 1.3.4 il existe $\mathbb{I}$ un segment initial fort de $\mathbb{M}$ tel que $a \in \mathbb{I} < b$ et comme l'indique la proposition 1.3.2, $\mathbb{I} \models \mathrm{P}$. D'après les hypothèses $\mathrm{P} \vdash \forall x \exists! y \ F(x, y)$ donc $\exists c \in \mathbb{I} \ c = f(a)$. Mais alors $\mathbb{M} \models F(a, b)$ et $\mathbb{M} \models F(a, c)$ et comme $c \in \mathbb{I} < b$, $c \neq b$ ce qui est impossible vu que $\mathbb{M} \models \mathrm{P}$ et que $\mathrm{P} \vdash \exists! y \ F(a, y)$. □

La première démonstration de ce théorème se trouve dans [Wai70] et [Wai72]. On trouvera aussi une preuve ultérieure de ce résultat dans [BW87] passant par la théorie de la démonstration.

1.5 Indépendance du théorème Goodstein dans P

On a maintenant défini tous les outils nécessaires à la preuve du résultat qui nous intéresse principalement dans cette partie. Comme indiqué précédemment, le théorème de Goodstein n'est pas dans P car la fonction de Goodstein (que l'on voit comme une fonction de q seulement en fixant p en fonction de q) majore (au sens de $<_*$) toutes les fonctions de la hiérarchie de Hardy.

Théorème 1.5.1 (Kirby-Paris, 1982)**.**

$$\mathrm{P} \nvdash \forall p \forall q (q > 1 \Rightarrow \exists m F(m, p, q, 0))$$

En d'autres termes, le théorème de Goodstein n'est pas démontrable dans l'arithmétique de Peano.

Démonstration. Par l'absurde, supposons que le théorème de Goodstein soit démontrable dans P alors h (étendue en $(p, 0)$ et $(p, 1)$ pour tout p par 0 par exemple) est prouvablement totale dans P.
Soit $puis : \mathbb{N}^2 \to \mathbb{N}$ primitive récursive définie par $puis(x, 0) = 1$ et pour tout $n > 0$ $puis(x, n) = x^{puis(x,n-1)}$. On vérifie par une récurrence simple que $f_{x,\omega}(p(x, n)) = \omega_n$
On a alors

$$g : \left\{ \begin{array}{ccc} \mathbb{N} & \to & \mathbb{N} \\ n & \mapsto & h(puis(n, n), n) + n + 1 \end{array} \right.$$

récursive prouvablement totale dans P.
Par le théorème de Wainer (voir 1.4.1) il existe $\alpha < \epsilon_0$ tel que, pour tout $n \in \mathbb{N}$, $g(n) < H_\alpha(n)$.
Or, par le théorème de Cichon (voir 1.2.18), pour tout $n \in \mathbb{N}$:

$$g(n) = H_{f_{n,\omega}(puis(n,n))}(n + 1) = H_{\omega_n}(n)$$

De plus par la propriété 1.2.3 $\exists n_0$ tel que $\alpha < \omega_{n_0}$ donc par la propriété 1.2.17 de monotonie de la hiérarchie de Hardy par rapport aux ordinaux :

$$\exists n_1 \ \forall n > n_1 \ H_{\omega_{n_0}}(n) > H_\alpha(n)$$

et comme $\forall p > q \ \omega_p \to_1 \omega_q$ voir 1.2.14, toujours par la propriété 1.2.17 on a pour tout $p > q$:

$$\forall n > 1 \ H_{\omega_p}(n) > H_{\omega_q}(n)$$

Donc en prenant $n > \max(2, n_1, n_2)$ on a

$$H_{\omega_n}(n) > H_{\omega_{n_0}}(n) > H_\alpha(n) > H_{\omega_n}(n)$$

ce qui est absurde.
On a ainsi démontré que le théorème de Goodstein n'est pas démontrable dans l'arithmétique de Peano. □

Chapitre 2

Le théorème de Gentzen et ϵ_0

On présentera, tout d'abord, dans cette partie le théorème de Gentzen dont la conséquence est la suivante : la cohérence de l'arithmétique peut être démontrée de façon finitiste à une induction jusqu'à ϵ_0 près. On finira en montrant que le théorème de Goodstein implique la bonne fondation de ϵ_0 (en suivant [KH89]). Ceci fournira une deuxième preuve du théorème de Kirby-Paris par le deuxième théorème d'incomplétude de Gödel.
Cette preuve utilise des outils de la théorie de la démonstration. Au contraire de la première partie, qui reposait sur des arguments sémantiques et la théorie des modèles, ce sera donc une présentation de méthodes syntaxiques pour répondre à des questions sur l'arithmétique de Peano.
L'idée générale développée ici est de travailler sur la *structure* des preuves formelles. Les systèmes de preuves à la Hilbert[1] sont caractérisés par un très petit nombre de règles, en général seulement le *modus ponens* et la généralisation. En contrepartie, ils ont un grand nombre d'axiomes comme "toute les tautologies du calcul propositionnel sans quantificateurs". Pour manipuler plus aisément la structure des preuves, on présentera ici un autre système de preuve formelle, possédant plus de règles et moins d'axiomes : le calcul des séquents.

2.1 Calcul des séquents pour la logique propositionnelle

2.1.1 Langage de premier ordre

On reprend les définitions habituelles d'un langage de premier ordre.

Définition 2.1.1. Un langage de premier ordre L est la donnée :
– d'un ensemble de symboles de constantes ;
– d'un ensemble de symboles de variables ;
– d'un ensemble de symboles de fonctions f_i d'arités $|f_i|$;
– d'un ensemble de symboles de relations R_i d'arités $|R_i|$.

Définition 2.1.2. On définit inductivement les *termes* de ce langage :
– une constante est un terme ;
– une variable est un terme ;

[1] Par exemple celui utilisé cette année dans le cours de Logique du premier semestre

– si f_i est un symbole de fonction et $t_1, t_2, ..., t_{|f_i|}$ des termes, alors $f_i(t_1, ..., t_{|f_i|})$ est un terme.

Définition 2.1.3. On définit inductivement les *formules* sur ce langage :
– Si R_i est un symbole de relation et $t_1, ..., t_{|R_i|}$ sont des termes de L, $R_i(t_1, ..., t_{|R_i|})$ est une formule. On dit de plus que c'est une *formule atomique*.
– Si A et B sont formules, alors $\neg A$, $A \wedge B$, $A \vee B$ sont des formules.
– Si A est une formule et x une variable, $\forall x A$ et $\exists x A$ sont des formules.

Définition 2.1.4. Pour mesurer la "complexité" d'une formule, on définit inductivement une fonction *comp* :
– Si A est atomique, $\text{comp}(A) = 1$
– $\text{comp}(\neg A) = \text{comp}(\exists x A) = \text{comp}(\forall x A) = 1 + \text{comp}(A)$
– $\text{comp}(A \wedge B) = \text{comp}(A \vee B) = 1 + \max(\text{comp}(A), \text{comp}(B))$

2.1.2 Séquents et inférences

Définition 2.1.5. Un *séquent* $\Gamma \vdash \Delta$ est la donnée de deux ensembles de formules Γ et Δ. On dit que les éléments de Γ sont les *hypothèses* du séquent, et ceux de Δ sont ses *résultats*. On le note sous la forme $\Gamma \vdash \Delta$.

Informellement, on peut interpréter un séquent $\Gamma \vdash \Delta$ comme l'affirmation que la conjonction des hypothèses entraîne la disjonction de ses résultats : "si toutes les formules de Γ sont vraies, au moins une des formules de Δ est vraie".
$\emptyset \vdash \Delta$ représente une tautologie (la disjonction des formules de Δ est vraie), et $\Gamma \vdash \emptyset$ une négation/contradiction (la conjonction des formules de Γ est fausse). En particulier, le séquent $\emptyset \vdash \emptyset$, souvent noté simplement $\vdash$, représente le faux.

Définition 2.1.6. Une *inférence* est la donnée d'un groupe (fini) de séquents, les *prémisses*, et d'un séquent, la *conclusion*. Elle indique la validité du raisonnement qui passe des prémisses à la conclusion. On la note sous la forme

$$\frac{\Gamma_1 \vdash \Delta_1 \qquad \Gamma_2 \vdash \Delta_2}{\Gamma_3 \vdash \Delta_3}$$

Les prémisses sont au dessus de la barre, et la conclusion en dessous.

2.1.3 Règles et démonstrations

Définition 2.1.7. Un *système de preuve* est la donnée d'un ensemble d'inférences, ses *règles*.

On considère dans cette partie le système **LK**, correspondant aux preuves de la simple logique du premier ordre habituelle, déterminé par l'ensemble de règles qui suivront.
Les premières règles de notre système formel, les *règles logiques*, décrivent la manière dont on peut créer des formules propositionnelles à partir d'un connecteur logique et de formules propositionnelles plus simples, par exemple $A \wedge B$ à partir de A et de B. Pour chaque connecteur il y a deux types de règles : les "règles gauches" qui permettent son apparition parmi les hypothèses du séquent, et les "règles droites" qui concernent les conclusions du séquent.

$$\frac{\Gamma \vdash A, \Delta}{\Gamma, \neg A \vdash \Delta}\ (\neg \vdash) \qquad \frac{\Gamma, A \vdash \Delta}{\Gamma \vdash \neg A, \Delta}\ (\vdash \neg)$$

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}\ (\wedge_1 \vdash) \qquad \frac{\Gamma, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}\ (\wedge_2 \vdash) \qquad \frac{\Gamma \vdash A, \Delta \qquad \Gamma \vdash B, \Delta}{\Gamma \vdash A \wedge B, \Delta}\ (\vdash \wedge)$$

$$\frac{\Gamma, A \vdash \Delta \qquad \Gamma, B \vdash \Delta}{\Gamma, A \vee B \vdash \Delta}\ (\vee \vdash) \qquad \frac{\Gamma \vdash A, \Delta}{\Gamma \vdash A \vee B, \Delta}\ (\vdash \vee_1) \qquad \frac{\Gamma \vdash B, \Delta}{\Gamma \vdash A \vee B, \Delta}\ (\vdash \vee_2)$$

$$\frac{\Gamma, A[x := t] \vdash \Delta}{\Gamma, \forall x A \vdash \Delta}\ (\forall \vdash) \qquad \frac{\Gamma \vdash A, \Delta}{\Gamma \vdash \forall x A, \Delta}\ (\vdash \forall)$$

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, \exists x A \Delta}\ (\exists \vdash) \qquad \frac{\Gamma \vdash A[x := t] \Delta}{\Gamma \vdash \exists A, \Delta}\ (\vdash \exists)$$

Les notations Γ, Γ' et Γ, A décrivent les ensembles $\Gamma \cup \Gamma'$ et $\Gamma \cup \{A\}$. $A[x := t]$ désigne la formule A dans laquelle on a remplacé toutes les occurrences libres de la variable x par le terme t (à renommage de variables près, pour éviter les accidents). Dans les deux règles qui l'utilisent ($\forall \vdash$, $\vdash \exists$) on demande de plus à ce que la variable x ne fasse pas partie des variables libres de Γ et Δ : sinon l'inférence n'est pas une règle.
Pour pouvoir manipuler les formules d'un séquent, on a envie de pouvoir les réordonner selon ses besoins, supprimer les doublons, ajouter des hypothèses ou des résultats possibles. C'est ce que permettent les *règles structurelles* :

$$\frac{\Gamma, A, B, \Gamma' \vdash \Delta}{\Gamma, B, A, \Gamma' \vdash \Delta}\ (\text{échange} \vdash)) \qquad \frac{\Gamma \vdash \Delta, A, B, \Delta\text{'}}{\Gamma \vdash \Delta, B, A, \Delta'}\ (\vdash \text{échange})$$

$$\frac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta}\ (\text{contraction} \vdash) \qquad \frac{\Gamma \vdash A, A, \Delta}{\Gamma \vdash A, \Delta}\ (\vdash \text{contraction})$$

$$\frac{\Gamma \vdash \Delta}{\Gamma, A \vdash \Delta}\ (\text{affaiblissement} \vdash) \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash A, \Delta}\ (\vdash \text{affaiblissement})$$

Contrairement à ce qu'on pourrait penser, ces règles structurelles ont de l'intérêt en soi : si on les restreint, on obtient des logiques spécifiques qui peuvent avoir des avantages théoriques et/ou des applications pratiques. Par exemple si on enlève la règle d'échange, on ne peut plus utiliser que l'hypothèse la plus à droite (puisque toutes les règles logiques sont de la forme $\Gamma, A \vdash ...$) ; cela permet entre autres de modéliser la logique de langages de programmations utilisant une pile pour stocker des valeurs, et qui ne peuvent utiliser que la valeur placée en haut de la pile.
Enfin, il reste deux *règles d'identité*, l'axiome et la coupure :

$$\frac{}{A \vdash A}\ (\text{axiome}) \qquad \frac{\Gamma \vdash A, \Delta \qquad A, \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\ (A\text{-coupure})$$

L'intérêt de la coupure est qu'il permet la réutilisation des preuves : on peut combiner

deux preuves séparées pour obtenir un nouveau résultat sans devoir repartir de zéro. Cela permet une certaine modularité.
Par exemple, si l'on sait prouver pour toutes les formules A et B les séquents $A, B \vdash A \wedge B$ et $A, A \Rightarrow B \vdash B$, on peut obtenir le séquent $C, D, (C \wedge D) \Rightarrow E \vdash E$ par coupure, sans devoir récrire les preuves pour tenir compte des variations des hypothèses.

2.1.4 Preuves

Définition 2.1.8. Une *preuve* ϕ d'un séquent $\Gamma \vdash \Delta$ est la donnée d'une règle dont la conclusion est ce séquent, et d'une preuve ϕ_i de chacune des prémisses de cette règle (appelée *dernière règle* de la preuve). On note $\phi \Vdash \Gamma \vdash \Delta$, ou bien

$$\begin{array}{c} \phi \\ \vdots \\ \Gamma \vdash \Delta \end{array}$$

Les preuves sont donc des arbres d'inférences, qui se terminent par des règles sans prémisses, donc des règles axiomes.
Voici par exemple une preuve du séquent $A \wedge \neg A \vdash B$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}\,(\text{axiome})}{A \vdash B, A}\,(\vdash \text{affaiblissement})}{A \vdash A, B}\,(\vdash \text{échange})}{A, \neg A \vdash B}\,(\vdash \neg)}{A \wedge \neg A \vdash B}\,(\wedge \vdash)$$

Les preuves se lisent usuellement de bas en haut. Celle-ci est linéaire, puisque chaque inférence utilise une seule prémisse, mais elles sont en général arborescentes :

$$\dfrac{\dfrac{\dfrac{\dfrac{}{A, B \vdash B}\,(\text{axiome}) \quad \dfrac{}{A, B \vdash A}\,(\text{axiome})}{A, B \vdash B \wedge A}\,(\vdash \wedge)}{A \wedge B \vdash B \wedge A}\,(\wedge \vdash)$$

On utilise $=$ pour signaler la présence d'une ou plusieurs inférences structurelles passées sous silence.

2.2 Élimination des coupures

2.2.1 Propriété de la sous-formule

Si l'on suit les deux preuves ci-dessus de la conclusion jusqu'aux prémisses, on peut se rendre compte qu'elles sont quasiment "automatiques" : à chaque étape, selon la structure des formules du séquent à prouver, une ou plusieurs règles sont utilisables, et l'on peut essentiellement choisir arbitrairement l'une d'entre elles, obtenir une expression plus simple et aboutir (modulo quelques règles structurelles ça et là) à des règles axiomes.
Cette méthode n'est pas utilisable pour une règle coupure :

$$\dfrac{\Gamma \vdash A, \Delta \qquad A, \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,(A\text{-coupure})$$

Pour établir l'inférence en partant de $\Gamma, \Gamma' \vdash \Delta, \Delta'$, il faudrait "deviner" le terme A. On peut décrire plus précisément ce problème par la propriété de la sous-formule.

Définition 2.2.1. Les *sous-formules directes* d'une formule sont les termes utilisés pour sa construction :
– les sous-formules directes de $A \wedge B$ et $A \vee B$ sont A et B
– la sous-formule directe de $\neg A$ est A
– les sous-formules directes de $\forall x A$ et $\exists x A$ sont les $A[x := t]$ pour chaque terme t.
Une *sous-formule* d'une formule donnée est sa sous-formule directe, ou la sous-formule directe d'une de ses sous-formules. Cette définition est bien fondée car la complexité d'une formule est toujours strictement supérieure aux complexités de ses sous-formules (directes).

Proposition 2.2.2. *Toutes les règles d'inférences, sauf la coupure, ont la propriété de la sous-formule : les formules des prémisses des formules de la conclusion, ou des sous-formules de celles-ci.*

La propriété de la sous-formule est une propriété désirable, en particulier dans l'optique d'une preuve de cohérence : si on démontre le faux, alors une sous-formule du faux est dans les prémisses, autrement dit "seul le faux implique le faux". C'est exactement la preuve que l'on va faire, mais pour cela il faut éliminer les coupures.

Théorème 2.2.3 (Élimination des coupures)**.** *Toute preuve du calcul des séquents peut être récrite en une preuve du même séquent n'utilisant pas la règle de coupure.*

La démonstration de ce résultat est l'objet principal de cette section. Elle est inspirée de [GTL89] et [Tak75].

Remarque 1. Même avec l'élimination des coupures, la procédure de démonstration automatique esquissée plus haut ne marche pas : parmi les règles structurelles à l'air innocent se trouve la règle de contraction qui a la propriété (si on regarde de bas en haut, des conclusions aux prémisses) de dupliquer des formules, augmentant alors le nombre de possibilités à l'étape suivante. Au sein du calcul propositionnel sans quantificateurs, on pourrait en étant précautionneux retrouver une procédure de décision, mais au premier ordre cela ne fonctionne plus.

2.2.2 Cas clés

2.2.2.1 Exemple : cas $\wedge$

L'idée est de faire "remonter" les coupures dans les preuves, en transformant une coupure sur une formule en une ou plusieurs coupures sur ses sous-formules, situées plus haut dans l'arbre de preuve.
Les cas clés sont les cas où les coupures font correspondre une règle gauche avec la règle droite qui lui est associée. Par exemple, si l'on a la preuve suivante :

$$\dfrac{\dfrac{\dfrac{\overset{\phi_1}{\cdots}}{\Gamma \vdash A, \Delta} \quad \dfrac{\overset{\phi_2}{\cdots}}{\Gamma \vdash B, \Delta}}{\Gamma \vdash A \wedge B, \Delta}\,(\vdash \wedge) \quad \dfrac{\overset{\phi'}{\cdots}}{\Gamma', A, \vdash \Delta'}\,(\wedge_1 \vdash)}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,((A \wedge B)\text{-coupure})$$

Par les preuves ϕ_1 et ϕ_2 on prouve les résultat A et B respectivement, donc le résultat $A \wedge B$ utilisé par la prémisse de droite. Le point essentiel pour l'élimination de cette coupure est de remarquer qu'on effectue ici un *détour* : on prouve A et B alors que la preuve ϕ' n'utilise en fait que l'hypothèse A. On peut donc se contenter de couper sur A :

$$\dfrac{\overset{\phi_1}{\overset{\cdots\cdots}{\Gamma \vdash A, \Delta}} \qquad \overset{\phi'}{\overset{\cdots\cdots}{\Gamma', A, \vdash \Delta'}}}{\Gamma, \Gamma' \vdash \Delta, \Delta'} (A\text{-coupure})$$

On a transformé une coupure sur $A \wedge B$ en une coupure sur sa sous-formule A : la complexité de la formule coupée décroît strictement.
Le cas de la règle $(\wedge_2 \vdash)$ est clairement symétrique. On peut énumérer tous les autres cas clés, où une coupure fait correspondre une règle gauche et une règle droite de la même formule.

Remarque 2. Les informaticiens savent depuis longtemps qu'il existe des liens forts entre la logique et le typage d'un langage de programmation, éclairant certains concepts logiques sous un angle opérationnel. Les coupures présentées ici correspondent effectivement aux réductions d'un lambda-calcul, par exemple ici, avec des paires, $\pi_1(< s, t >) \to s$. On peut alors interpréter la réduction des coupures comme un processus de normalisation des preuves.

2.2.2.2 Cas $\vee$

$$\dfrac{\dfrac{\overset{\phi}{\overset{\cdots\cdots}{\Gamma \vdash A, \Delta}}}{\Gamma \vdash A \vee B, \Delta} (\vdash \vee_1) \qquad \dfrac{\overset{\phi'_1}{\overset{\cdots\cdots}{\Gamma', A \vdash \Delta'}} \qquad \overset{\phi'_2}{\overset{\cdots\cdots}{\Gamma', B \vdash \Delta'}}}{\Gamma', A \vee B \vdash \Delta'} (\vee \vdash)}{\Gamma, \Gamma' \vdash \Delta, \Delta'} ((A \vee B)\text{-coupure})$$

se réduit en

$$\dfrac{\overset{\phi}{\overset{\cdots\cdots}{\Gamma \vdash A, \Delta}} \qquad \overset{\phi'_1}{\overset{\cdots\cdots}{\Gamma', A \vdash \Delta'}}}{\Gamma, \Gamma' \vdash \Delta, \Delta'} (A\text{-coupure})$$

Le cas $(\vdash \vee_2)$ est symétrique.

2.2.2.3 Cas $\neg$

$$\dfrac{\dfrac{\overset{\phi}{\overset{\cdots\cdots}{\Gamma A \vdash \Delta}}}{\Gamma \vdash \neg A, \Delta} (\vdash \neg) \qquad \dfrac{\overset{\phi'}{\overset{\cdots\cdots}{\Gamma' \vdash A, \Delta'}}}{\Gamma' \neg A \vdash \Delta'} (\neg \vdash)}{\Gamma, \Gamma' \vdash \Delta, \Delta'} ((\neg A)\text{-coupure})$$

se réduit en

$$\dfrac{\overset{\phi}{\overset{\cdots\cdots}{\Gamma \vdash A, \Delta}} \qquad \overset{\phi'}{\overset{\cdots\cdots}{\Gamma', A \vdash \Delta'}}}{\Gamma, \Gamma' \vdash \Delta, \Delta'} (A\text{-coupure})$$

2.2.2.4 Cas $\forall$

$$\dfrac{\dfrac{\begin{array}{c}\phi\\ \vdots\\ \Gamma \vdash A, \Delta\end{array}}{\Gamma \vdash \forall x A, \Delta}\,(\vdash \forall) \qquad \dfrac{\begin{array}{c}\phi'\\ \vdots\\ \Gamma', A[x := t] \vdash \Delta'\end{array}}{\Gamma, \forall x A \vdash \Delta}\,(\forall \vdash)}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,((\forall x A)\text{-coupure})$$

se réduit en

$$\dfrac{\begin{array}{c}\phi[x := t]\\ \vdots\\ \Gamma \vdash A[x := t], \Delta\end{array} \qquad \begin{array}{c}\phi'\\ \vdots\\ \Gamma', A[x := t] \vdash \Delta'\end{array}}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,(A\text{-coupure})$$

Ici $\phi[x := t]$ désigne la preuve ϕ dans laquelle on a remplacé toutes les occurrences libres de x issues de A par le terme t.

2.2.2.5 Cas $\exists$

$$\dfrac{\dfrac{\begin{array}{c}\phi\\ \vdots\\ \Gamma \vdash A[x := t], \Delta\end{array}}{\Gamma \vdash \exists x A, \Delta}\,(\vdash \exists) \qquad \dfrac{\begin{array}{c}\phi'\\ \vdots\\ \Gamma', A \vdash \Delta'\end{array}}{\Gamma', \exists x A \vdash \Delta'}\,(\exists \vdash)}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,((\forall x A)\text{-coupure})$$

se réduit en

$$\dfrac{\begin{array}{c}\phi\\ \vdots\\ \Gamma \vdash A[x := t], \Delta\end{array} \qquad \begin{array}{c}\phi'[x := t]\\ \vdots\\ \Gamma', A[x := t] \vdash \Delta'\end{array}}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,(A\text{-coupure})$$

2.2.2.6 Résumé

Les cas les plus importants, où il "se passe quelque chose", sont traités. Il reste à traiter tous les autres cas, où les règles utilisées juste avant la coupure ne correspondent pas. Il faut alors faire remonter la coupure correctement pour que les règles concordent à nouveau. La propriété essentielle est cependant déjà visible : c'est la réduction de la complexité des formules coupées (mais pas en général de la taille des preuves ou du nombre de coupures).

Proposition 2.2.4. *Dans chacun des cas clés, la complexité de la formule coupée décroît strictement.*

Définition 2.2.5. Le *degré* d'une coupure est la complexité de la formule coupée.

2.2.3 Preuve complète

Définition 2.2.6. Le *degré* d'une preuve est la borne supérieure du degré de ses coupures. Le degré d'une preuve sans coupure est 0, ce qui est cohérent puisque la complexité d'une formule (donc le degré d'une coupure) est au moins 1.
Pour dire qu'une preuve $\phi \Vdash \Gamma \vdash \Delta$ est de degré strictement inférieur à d, on note $\phi \Vdash_{<d} \Gamma \vdash \Delta$.

Définition 2.2.7. La *hauteur* d'une preuve ϕ est la hauteur de l'arbre associé : $H(\phi) := 1 + \sup_i(H(\phi_i))$.

Lemme 2.2.8. *Soit A une formule de complexité d. Si $\phi \Vdash_{<d} \Gamma \vdash A, \Delta$ et $\phi' \Vdash_{<d} \Gamma' \vdash A, \Delta'$, alors il existe une preuve $\psi \Vdash_{<d} \Gamma, \Gamma' \vdash \Delta, \Delta'$.*

Démonstration. Sans perdre de généralité, on peut supposer que A n'apparaît pas dans $\Gamma, \Gamma', \Delta, \Delta'$ (sinon, on peut récupérer les formes exactes en ajoutant des inférences structurelles)[2].
On procède alors par induction sur $H(\phi) + H(\phi')$. On fait des distinctions de cas selon les dernières règles r et r' des preuves ϕ et ϕ' :

1. r est un axiome. Il y a deux sous-cas :
   - r prouve $A \vdash A$ (donc $\Gamma = \{A\}, \Delta = \emptyset$), on récupère une preuve de $(\Gamma, \Gamma' \vdash \Delta, \Delta') = (A, \Gamma' \vdash \Delta')$ en réordonnant, par échange, les hypothèses de $\phi' \Vdash \Gamma', A \vdash \Delta'$.
   - r prouve $B \vdash B$ avec $B \neq A$: alors $(\Gamma, \Gamma' \vdash \Delta, \Delta') = (B, \Gamma' \vdash B, \Delta')$ est obtenu en ajoutant, par affaiblissement, Γ' et Δ' à $\phi \Vdash B \vdash B$.

   Si r' est un axiome, la situation est similaire. Dans les deux cas, il n'y a pas de coupure donc le degré est $0 < d$.
2. r ou r' sont des règles structurelles (supposons qu'il s'agit de r, le cas r' est symétrique) : la prémisse $\Gamma_1 \vdash \Delta_1$ est prouvée par ϕ_1 avec $H(\phi_1) < H(\phi)$, donc on peut obtenir par hypothèse d'induction sur ϕ_1 et ϕ' une preuve $\psi \Vdash_{<d} \Gamma_1, \Gamma' \vdash \Delta_1, \Delta'$. Puisque r est une règle structurelle, le passage de Γ_1 et Δ_1 à Γ et Δ ne demande que l'application d'une ou plusieurs règles structurelles, que l'on applique à ψ, sans modifier son degré, pour obtenir notre résultat.
3. r est une règle logique qui n'est pas une règle de création (à gauche ou à droite) de la formule A. L'application de l'hypothèse d'induction sur les couples (ϕ_i, ϕ') (où ϕ_i parcourt les preuves associées aux prémisses de r) donne des preuves $\psi_i \Vdash \Gamma_i, \Gamma' \vdash \Delta_i, \Delta'$, que l'on peut utiliser comme prémisses de la règle r. Celle-ci ne crée pas d'occurrence de A, donc produit bien une preuve du séquent $\Gamma, \Gamma' \vdash \Delta, \Delta'$ avec A n'apparaissant pas dans $\Gamma, \Gamma', \Delta, \Delta'$. La situation est similaire pour r'.
4. r et r' sont des règles sur A, l'une gauche et l'autre droite. C'est le cas le plus important.

   On commence par appliquer l'hypothèse d'induction sur les (ϕ_i, ϕ') et les (ϕ, ϕ'_j) pour obtenir des preuves $\psi_i \Vdash_{<d} \Gamma_i, \Gamma' \vdash \Delta_i, \Delta'$ et $\psi'_j \Vdash_{<d} \Gamma, \Gamma'_j \vdash \Delta, \Delta'_j$.

   Ensuite on applique r aux ψ_i et r' aux ψ_j, pour obtenir (modulo quelques règles structurelles) une preuve $\psi \Vdash_{<d} \Gamma, \Gamma', \vdash A, \Delta, \Delta'$ et une preuve $\psi' \Vdash_{<d} \Gamma, \Gamma', A \vdash \Delta, \Delta'$.

[2] Pour un traitement détaillé de ce point technique, consulter [Tak75]

En utilisant la règle de coupure sur ψ et ψ', on obtient alors une preuve de $\Gamma, \Gamma' \vdash \Delta, \Delta'$ de degré $\text{comp}(A) = d$. Mais c'est précisément un des cas clés : on peut donc le transformer comme vu précédemment en une preuve de degré strictement inférieur à d.

□

Proposition 2.2.9. *Si ϕ est une preuve de degré $d > 0$, il existe une preuve du même séquent de degré strictement inférieur à d.*

Démonstration. Par induction sur $H(\phi)$. On utilise l'hypothèse d'induction sur les prémisses (ϕ_i) de ϕ pour en obtenir des preuves (ψ_i) de degré inférieur à d.
– si la dernière règle r de ϕ n'est pas une règle de coupure, on l'applique simplement aux (ψ_i) et on obtient le résultat désiré.
– Si r est une coupure, on utilise le lemme précédent.

□

Théorème 2.2.10 (Élimination des coupures dans LK)**.** *Si ϕ est une preuve du séquent s dans* **LK***, il existe une preuve de s de degré nul.*

Démonstration. Par itération de la proposition précédente. □

Il est important de remarquer que ce résultat d'existence est en fait constructif : on a littéralement construit une preuve sans coupure en réduisant ϕ. En particulier, dans un système formel convenable, on pourrait exprimer cette réduction avec des fonctions récursives primitives.

2.3 Extension du résultat à l'arithmétique

La difficulté est maintenant d'étendre cette preuve à l'arithmétique. On a procédé dans le cas de **LK** par induction sur les hauteurs et degrés des preuves. On a besoin pour l'arithmétique d'induction sur des ordinaux bien plus gros, en fait jusqu'à ϵ_0 : c'est là que se situera le lien avec les suites de Goodstein.

2.3.1 Formalisation de P en calcul des séquents

On utilise pour P le langage habituel de l'arithmétique : constante 0, fonctions succ, $+$, $\times$ et relation $=$.
Le système de preuve P est obtenu en ajoutant des axiomes à **LK**.
Les *axiomes mathématiques* (pour s, t, r termes arbitraires) :
– $\text{succ}(s) = \text{succ}(t) \vdash s = t$
– $\text{succ}(s) = 0 \vdash$ ($\emptyset$ implicite à droite)
– $s = t, t = r \vdash s = r$
– $s = t \vdash \text{succ}(s) = \text{succ}(t)$
– $\vdash s + 0 = s$
– $\vdash s \times 0 = 0$
– $\vdash s \times \text{succ}(t) = s \times t + s$
Pour $F(x)$ une formule à un paramètre (pointant toutes les occurrences libres de x) et t un terme du langage, on ajoute la règle d'induction sur F

$$\frac{\Gamma, F(a) \to F(\mathrm{succ}(a)), \Delta}{\Gamma, F(0) \vdash F(t), \Delta}\ (Inf_F)$$

Les inductions compliquent en fait significativement la preuve d'élimination des coupures : elles permettent une forme de "raisonnement infini" qui rend inapplicables les simples concepts de hauteur et de degré.
On peut cependant adapter cette idée, en utilisant des ordinaux pour mesurer la "hauteur" des preuves. Une numérotation astucieuse des preuves par des ordinaux permet de vérifier la terminaison de la réduction des coupures de la même manière que pour **LK** ; c'est l'idée de la preuve de Gentzen, reprise par Takeuti ([Gen38], [Tak75]).
Cependant, les choix de numérotation ne sont pas forcément naturels et rendent la preuve un peu absconse. Pour rendre les notations plus explicites, on peut passer à un système *infinitaire*, où les preuves sont des objets infinis, et se laissent alors naturellement numéroter par des ordinaux. Cette idée, due à Schütte ([Sch56]), rend la preuve nettement plus accessible.

2.3.2 Système infinitaire

2.3.2.1 Le système P^∞

On définit ici un nouveau système de preuve, P^∞, qui permet de manipuler des preuves de taille infinie : le nombre de prémisses d'une inférence n'est plus forcément fini.
On conserve comme règle de P^∞ les axiomes et les règles identités de P. On utilise par contre les règles logiques suivantes :

$$\frac{\Gamma \vdash A, \Delta}{\Gamma, \neg A \vdash \Delta}\ (\neg \vdash) \qquad \frac{\Gamma, A \vdash \Delta}{\Gamma \vdash \neg A, \Delta}\ (\vdash \neg)$$

$$\frac{\Gamma, A_k \vdash \Delta}{\Gamma, A_1 \wedge A_2 \vdash \Delta}\ (\wedge_k \vdash) \quad k \in 1, 2 \qquad \frac{\Gamma \vdash A_1, \Delta \qquad \Gamma \vdash A_2, \Delta}{\Gamma \vdash A_1 \wedge A_2, \Delta}\ (\vdash \wedge)$$

$$\frac{\Gamma, A_1 \vdash \Delta \qquad \Gamma, B_2 \vdash \Delta}{\Gamma, A_1 \vee A_2 \vdash \Delta}\ (\vee \vdash) \qquad \frac{\Gamma \vdash A_k, \Delta}{\Gamma \vdash A_1 \vee A_2, \Delta}\ (\vdash \vee_k) \quad k \in 1, 2$$

$$\frac{\Gamma, A[x := k] \vdash \Delta}{\Gamma, \forall x A \vdash \Delta}\ (\forall_k \vdash) \quad k \in \mathbb{N} \qquad \frac{(\Gamma \vdash A[x := i], \Delta)_{i \in \mathbb{N}}}{\Gamma \vdash \forall x A, \Delta}\ (\vdash \forall)$$

$$\frac{(\Gamma, A[x := i] \vdash \Delta)_{i \in \mathbb{N}}}{\Gamma, \exists x A \vdash \Delta}\ (\exists \vdash) \qquad \frac{\Gamma \vdash A[x := k]\Delta}{\Gamma \vdash \exists A, \Delta}\ (\vdash \exists) \quad k \in \mathbb{N}$$

La notation $(\Gamma_i \vdash \Delta_i)_{i \in \mathbb{N}}$ indique que l'inférence utilise une infinité de prémisses, tous les $\Gamma_i \vdash \Delta_i$ pour i entier naturel.
On peut de plus voir les $\wedge$ et $\vee$ comme des cas particuliers, finis et limités à deux éléments, des $\forall$ et $\exists$.

2.3.2.2 L'induction dispensable

Tout l'intérêt de ce système infinitaire est qu'il ne nécessite pas d'ajouter une règle supplémentaire pour l'induction : on peut déjà écrire des preuves par induction dans le système donné.
En effet, pour une formule $F(k)$ et un $n \in \mathbb{N}$ donné, il est facile de démontrer le séquent $F(0), \forall k(F(k) \Rightarrow F(\mathrm{succ}(k)) \vdash F(n)$. Par exemple pour $n = 2$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))) \vdash F(0)}\,(\text{axiome})}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))), \neg F(0) \vdash F(1)}\,(\neg \vdash) \quad \dfrac{}{F(1) \vdash F(1)}\,(\text{axiome})}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))), \neg F(0) \vee F(1) \vdash F(1)}\,(\vee \vdash)}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))) \vdash F(1)}\,(\forall \vdash)}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))), \neg F(1) \vdash F(2)}\,(\neg \vdash) \quad \dfrac{}{F(2) \vdash F(2)}\,(\text{axiome})}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))), \neg F(1) \vee F(2) \vdash F(2)}\,(\vee \vdash)}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))) \vdash F(2)}\,(\forall \vdash)$$

On peut en fait construire par récurrence une suite ϕ^n de preuves de $F(0), \forall k(F(k) \Rightarrow F(\mathrm{succ}(k)) \vdash F(n)$: ci-dessous, ϕ^0 puis $\phi^{\mathrm{succ}(n)}$ à partir de ϕ^n.

$$\dfrac{}{F(0), \forall (k)(\neg F(k) \vee F(\mathrm{succ}(k))) \vdash F(0)}\,(\text{axiome})$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\phi^n}{F(0), \forall (k)(\neg F(k) \vee F(\mathrm{succ}(k))) \vdash F(n)}}{F(0), \forall (k)(\neg F(k) \vee F(\mathrm{succ}(k))), \neg F(n) \vdash F(\mathrm{succ}(n))}\,(\neg \vdash) \quad \dfrac{}{F(\mathrm{succ}(n)) \vdash F(\mathrm{succ}(n))}\,(\text{axiome})}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))), \neg F(n) \vee F(\mathrm{succ}(n)) \vdash F(\mathrm{succ}(n))}\,(\vee \vdash)}{F(0), \forall k(\neg F(k) \vee F(\mathrm{succ}(k))) \vdash F(\mathrm{succ}(n))}\,(\forall \vdash)$$

On peut alors, en utilisant la nature infinitaire de $\forall$, construire une preuve de l'induction selon F :

Définition 2.3.1. $$\dfrac{\left(\phi^n \Vdash F(0), \forall k(F(k) \Rightarrow F(\mathrm{succ}(k))) \vdash F(n)\right)_{n \in \mathbb{N}}}{F(0), \forall k(F(k) \Rightarrow F(\mathrm{succ}(k))) \vdash \forall n F(n)}\,(\vdash \forall)$$

2.3.2.3 Réduction des coupures

On conserve la définition du degré d'une preuve de **LK**. Il faut par contre adapter un peu la notion de hauteur pour travailler sur des ordinaux. Au lieu de la définir intrinsèquement à partir des preuves, on annote les preuves par leur hauteur.

Définition 2.3.2. Une *preuve ordonnée* ϕ d'un séquent est la donnée d'un ordinal $H(\phi)$ et d'une preuve de ϕ dont les preuves des prémisses ϕ_i sont ordonnées, et qui vérifient $\forall i H(\phi_i) < H(\phi)$.
On note $\phi \Vdash^{\alpha}_{<d} \Gamma \vdash \Delta$ quand ϕ est de hauteur α et de degré strictement inférieur à d, ou encore

$$\alpha \ \dfrac{\overset{\phi}{\cdots\cdots}}{\Gamma \vdash \Delta}$$

Par la suite on utilisera généralement des preuves ordonnées, et on ne le précisera plus.

Lemme 2.3.3. *Soit A une formule de complexité d. Il existe un opérateur R_A sur les preuves, tel que pour $\phi \Vdash^{\alpha}_{<d} \Gamma \vdash A, \Delta$ et $\phi' \Vdash^{\beta}_{<d} \Gamma' \vdash A, \Delta'$, on ait $R_A(\phi, \phi') \Vdash^{\alpha+\beta}_{<d} \Gamma, \Gamma' \vdash \Delta, \Delta'$.*

Démonstration. C'est exactement la même preuve que pour le lemme 2.2.8. La gestion des hauteurs, qui n'est pas vraiment nouvelle puisqu'on faisait déjà l'induction sur $H(\phi) + H(\phi')$, ne pose absolument aucune difficulté.
Il faut cependant reformuler les cas clés correspondant à $\forall$ et $\exists$.

$$\dfrac{\dfrac{\overset{\phi}{\cdots\cdots}}{(\Gamma \vdash A[x := i], \Delta)_{i\in\mathbb{N}}}\,(\vdash \forall) \qquad \dfrac{\overset{\phi'}{\cdots\cdots}}{\Gamma', A[x := k] \vdash \Delta'}\,(\forall_k \vdash) \quad k \in \mathbb{N}}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,((\forall xA)\text{-coupure})$$

se réduit en

$$\dfrac{\overset{\phi_k}{\cdots\cdots} \qquad \overset{\phi'}{\cdots\cdots}}{\Gamma \vdash A[x := k], \Delta \qquad \Gamma', A[x := k] \vdash \Delta'}\,(A[x := k]\text{-coupure})$$

$$\dfrac{\dfrac{\overset{\phi}{\cdots\cdots}}{\Gamma \vdash A[x := k], \Delta}\,(\vdash \exists_k) \quad k \in \mathbb{N} \qquad \dfrac{\overset{\phi'}{\cdots\cdots}}{(\Gamma', A[x := i] \vdash \Delta')_{i\in\mathbb{N}}}\,(\exists \vdash)}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\,((\exists xA)\text{-coupure})$$

se réduit en

$$\dfrac{\overset{\phi}{\cdots\cdots} \qquad \overset{\phi'_k}{\cdots\cdots}}{\Gamma \vdash A[x := k], \Delta \qquad \Gamma', A[x := k] \vdash \Delta'}\,(A[x := k]\text{-coupure})$$

□

Proposition 2.3.4. *On peut se donner un opérateur $\mathcal{E}$ tel que si $\phi \Vdash^{\alpha}_{<d+1} \Gamma \vdash \Delta$, alors $\mathcal{E}(\phi) \Vdash^{\omega^\alpha}_{<d} \Gamma \vdash \Delta$.*

Démonstration. Cette fois c'est l'analogue de la proposition 2.2.9 :
– si la dernière règle r de *phi* n'est pas une règle de coupure, on utilise l'hypothèse d'induction sur ses prémisses, puis r sur les sous-preuves obtenues, conservant leur degré strictement inférieur à d, en attribuant la hauteur ω^α à la preuve obtenue. Cette hauteur est correcte parce que les hauteurs ω^{α_i} des prémisses obtenues par induction à partir des $\phi_i \Vdash^{\alpha_i}$ ... vérifient bien $\omega^{\alpha_i} < \omega^\alpha$ car nécessairement $\alpha_i < \alpha$.

– Si r est une coupure sur A, on a par induction des preuves ϕ_1 et ϕ_2 de ses prémisses, de degrés inférieurs à d. On utilise alors le lemme précédent pour obtenir $R_A(\phi_1, \phi_2) \Vdash^{\alpha+\beta}_{<d}$. Pour obtenir enfin une preuve de hauteur $\omega^{\alpha+\beta}$, il faut pouvoir donner une hauteur supplémentaire à la preuve.
Pour cela on rajoute à P^∞ une règle structurelle n'ayant aucun intérêt de démonstration mais permettant de ré-étiqueter une preuve pour lui donner une nouvelle hauteur :

$$\frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta}\,(\epsilon)$$

On peut alors donner la preuve ordonnée

$$\omega^{\alpha+\beta}\,\frac{R_A(\phi_1, \phi_2)}{\Gamma \vdash \Delta}\,(\epsilon)$$

□

Corollaire 2.3.5 (Élimination des coupures dans P^∞)**.** *Par itération de l'opération précédente, on peut transformer une preuve du séquent s en une preuve de s de degré nul.*

Proposition 2.3.6. *La hauteur d'une preuve de* (P^∞) *est déterminée par la hauteur de ses prémisses de façon calculable (primitive récursive) et elle appartient à ϵ_0.*

Démonstration. La hauteur d'un axiome est 1. Les autres hauteurs sont obtenues par addition ($\alpha, \beta \mapsto \alpha + \beta$) et exponentiation ($\alpha \mapsto \omega^\alpha$), opérations par lesquelles ϵ_0 est stable. □

Théorème 2.3.7. *Cohérence de* P^∞ *Par bon ordre de ϵ_0, il n'y a pas de preuve du séquent $\bot := \emptyset \vdash \emptyset$ dans* P^∞.

Démonstration. S'il existe une preuve de $\bot$, il existe une preuve ϕ de $\bot$ sans coupure. Par la propriété de la sous-formule, la dernière règle de ϕ admet nécessairement des hypothèses vides. Ça ne peut pas être un axiome (les conclusions ne sont pas vides) donc c'est la règle (ϵ), de prémisse $\bot$. La preuve ϕ_1 attachée aux prémisses de la règle (ϵ) est donc elle aussi une preuve de $\bot$, qui vérifie $H(\phi_1) < H(\phi)$. On peut par cette méthode produire une suite infinie de preuves de $\bot$ de hauteurs strictement décroissante, ce qui contredit l'hypothèse de bon ordre de ϵ_0. □

2.3.3 Système finitaire

L'inconvénient du système infinitaire est qu'il n'est pas satisfaisant pour les mathématiciens appréciant les méthodes finitistes. Pour contenter tout le monde, on peut écrire une traduction d'un système finitiste vers P^∞, qui permette de réutiliser la méthode de réduction des coupures de P^∞ par des moyens purement finitistes. Cette traduction, due à Buchholz ([Buc97]), permet même de justifier à posteriori les numérotations surprenantes de la preuve de Gentzen.
Comme système finitaire, on prend P comme défini précédemment, en lui rajoutant une règle (E) correspondant à la règle (ϵ) de P^∞ :

$$\frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta}\,(\mathrm{E})$$

Définition 2.3.8. On définit simultanément la traduction ϕ^∞ dans P^∞ d'une preuve $\phi \Vdash \Gamma \vdash \Delta$ de P, et la hauteur et le degré d'une preuve de P, de façon à vérifier

$$\phi^\infty \Vdash^{H(\phi)}_{<\deg(\phi)} \Gamma \vdash \Delta$$

Selon la dernière règle r de ϕ :
- Si r est une A-coupure :
  - $\deg(\phi) := \max(\deg(A), \deg(\phi_1), \deg(\phi_2))$
  - $H(\phi) := H(\phi_1) + H(\phi_2)$
  - $\phi^\infty := R_A(\phi_1^\infty, \phi_2^\infty)$
- Si r est (E) :
  - $deg(\phi) := deg(\phi_1) - 1$
  - $H(\phi) := \omega^{H(\phi_1)}$
  - $\phi^\infty := \mathcal{E}(\phi_1^\infty)$
- Si r est une induction sur F :
  - $\deg(\phi) := max(\deg(F), \deg(\phi_1)$
  - $H(\phi) := H(\phi_1) \times \omega$
  - ϕ^∞ construit dans la définition 2.3.1
- Sinon :
  - $\deg(\phi) = sup_i(\deg(\phi_i))$
  - $H(\phi) = 1 + sup_i(H(\phi_i))$
  -
  - Si ϕ est un axiome, $\phi^\infty := \phi$
  - $\left(\dfrac{\phi_0[x := t]}{\Gamma \vdash \Delta} (\forall \vdash) \right)^\infty := \dfrac{(\phi_0[x := i])_{i\in\mathbb{N}}}{\Gamma \vdash \Delta} (\forall \vdash)$
  - $\left(\dfrac{\phi_0[x := t]}{\Gamma \vdash \Delta} (\exists \vdash) \right)^\infty := \dfrac{(\phi_0[x := i])_{i\in\mathbb{N}}}{\Gamma \vdash \Delta} (\exists \vdash)$
  - sinon $\left(\dfrac{(\phi_i)_{i\in I}}{\Gamma \vdash \Delta} r \right)^\infty := \dfrac{(\phi_i)_{i\in I}}{\Gamma \vdash \Delta} r$, avec $I = \{1, 2\}$ $(\vee \vdash, \vdash \wedge)$ ou $I = \{k\}$ $(\vdash \vee_k, \wedge_k \vdash)$ ou $I = \{1\}$ $(\vdash \exists, \forall \vdash, \neg \vdash, \vdash \neg$, règles structurelles).

Remarque 3. Les hauteurs ordinales des preuves Gentzen-Takeuti prennent en fait tout leur sens : la hauteur $h(\phi_0) \times \omega$ de l'induction correspond aux ω prémisses de l'inférence infinitaire correspondante, et les tours d'exponentielles correspondent à plusieurs inférences (ϵ) implicitement utilisées pour réduire le degré de la preuve.

Proposition 2.3.9. *La hauteur d'une preuve de* (P) *est déterminée par la hauteur de ses prémisses de façon calculable (primitive récursive) et elle appartient à* ϵ_0.

Démonstration. La hauteur d'un axiome est 1. Les autres hauteurs sont obtenues par addition $(\alpha, \beta \mapsto \alpha+\beta)$, multiplication $(\alpha \mapsto \alpha\times\omega)$ et exponentiation $(\alpha \mapsto \omega^\alpha)$, opérations par lesquelles ϵ_0 est stable. □

Remarque 4. Avec cette traduction, on a en fait déjà prouvé la cohérence de P : s'il existait une preuve de $\emptyset \vdash \emptyset$ dans P, on pourrait la traduire dans P^∞ et y obtenir une preuve de $\emptyset \vdash \emptyset$ dans P, ce qui est contradictoire.
Cependant, cette preuve de consistance utilise les méthodes non finitistes de P^∞. Pour obtenir une preuve finitiste, il faut se donner dans P seul les outils permettant de faire la

preuve de cohérence. C'est ce qu'on va faire maintenant : à partir d'une preuve ϕ de P, on va construire des sous-preuves $\phi[i]$, qui ne sont pas les prémisses de ϕ mais permettent de construire une preuve du même séquent, avec en plus de bonnes propriétés (héritées de P^∞) permettant une preuve de cohérence.
La définition des $\phi[i]$ est assez fastidieuse. Leur principal intérêt est l'obtention d'une méthode purement finitiste, induction sur ϵ_0 exceptée.

On notera $R(\phi)$ la dernière règle de la preuve ϕ et $S(\phi)$ son séquent de conclusion.
À partir d'une preuve ϕ de P, on définit ses sous-preuves $\phi[i]$ dans P, qui vérifient $(\phi[i])^\infty = \phi_i^\infty$. On aura aussi besoin de définir $\mathrm{concl}(\phi)$, une règle de P^∞ telle que $\mathrm{concl}(\phi) = R(\phi^\infty)$.
On vérifiera par induction rapide, au moment de leur définition, que ces $\phi[i]$ satisfont les propriétés suivantes, dont l'intérêt est de vérifier que la traduction de P vers P^∞ préserve l'élimination des coupures :

Proposition 2.3.10. *1.* $\dfrac{(S(\phi[i]))_i}{S(\phi)}$ *est une règle de* P^∞

2. Si $\mathrm{concl}(\phi) = (A\text{-}$*coupure*$)$*, alors* $\mathrm{comp}(A) < \deg(\phi)$

3. $\forall i, \deg(\phi[i]) \leqslant \deg(\phi)$

4. $\forall i, H(\phi[i]) < H(\phi)$

Définition 2.3.11. Définition des $\phi[i]$ et de $\mathrm{concl}(\phi)$ par induction sur ϕ et distinction de sa dernière règle r :
– Ind_F : $\phi[n] := S(\phi^n)$ (les ϕ^n sont définies en 2.3.1), et $\mathrm{concl}(\phi) := (\epsilon)$
– (E) :
  – Si $\mathrm{concl}(\phi_1) = A$-coupure, $\mathrm{concl}(\phi) := (\epsilon)$ et
  $$\phi[1] := \left(\dfrac{\dfrac{\phi_1[1]}{\phi_1[1]}\,(\mathrm{E}) \quad \dfrac{\phi_1[2]}{\phi_1[2]}\,(\mathrm{E})}{S(\phi)}\,(A\text{-coupure}) \right)$$
  – sinon $\mathrm{concl}(\phi) := \mathrm{concl}(\phi_1)$ et $\phi[i] := \dfrac{\phi_1[i]}{\phi_1[i]}\,(\mathrm{E})$
– Si r est une A-coupure, on utilise la procédure de réduction des coupures du lemme 2.2.8 : on obtient une règle r' et deux prémisses ϕ'_1, ϕ'_2, et alors on pose $\mathrm{concl}(\phi) := r'$, $\phi[1] = \phi'_1$, $\phi[2] = \phi'_2$.
– dans les autres cas, la règle r existe encore dans P^∞ donc on a $\mathrm{concl}(\phi) = r$ et :
  – (axiome) : pas de prémisses $\phi[i]$
  – $(\vdash \wedge), (\vee \vdash) : \phi[i] := \phi_i \quad i \in \{1, 2\}$
  – $(\wedge_i \vdash), (\vdash \vee_i) : \phi[1] := \phi_1$
  – $(\vdash \forall), (\exists \vdash) : \phi[i] := \phi_1[x := i] \quad i \in \mathbb{N}$
  – $(\forall_k \vdash), (\vdash \exists_k) : \phi[k] := \phi_k$

L'intérêt de ces fastidieuses définitions et des fastidieuses vérifications qui y sont associées est le suivant : on a défini les prémisses et la dernière règle de la preuve qu'on obtiendrait dans P après élimination des coupures. On ne manipule plus que des objets de (P) : $\mathrm{concl}(\phi)$ est une règle de P^∞ mais elle ne sert qu'à guider la définition des $\phi[i]$. On a donc obtenu une méthode de réduction des coupures dans P qui serait définissable par des moyens purement finitistes, sans devoir passer par P^∞.
On peut maintenant passer à l'objet de cette partie, le résultat de cohérence de l'arithmétique.

Proposition 2.3.12. *Soit* $\mathrm{P}_\perp$ *l'ensemble des preuves* ϕ *du séquent* $\emptyset \vdash \emptyset$ *de degré nul. Si* $\phi \in \mathrm{P}_\perp$*, alors* $\phi[1] \in \mathrm{P}_\perp$ *et* $H(\phi[1]) < H(\phi)$*.*

Démonstration. D'après le premier fait de la proposition 2.3.10, $\dfrac{(S(\phi[i]))_i}{S(\phi)}$ est une règle de P^∞. La seule règle de P^∞ pouvant produire un séquent vide est la règle (ϵ), donc nécessairement $\mathrm{concl}(\phi) = (\epsilon)$, donc $\phi[1]$ existe et vérifie $H(\phi[1]) < H(\phi)$. □

Théorème 2.3.13 (Cohérence de P)**.** *Par bonne fondation de l'ordinal* ϵ_0*, il n'existe pas de preuve du faux (*$\emptyset \vdash \emptyset$*) dans* P*.*

Démonstration. D'après la proposition précédente, l'existence d'un élément dans $\mathrm{P}_\perp$ implique l'existence d'une suite infinie de hauteurs ordinales décroissantes. Or d'après la proposition 2.3.9, ces hauteurs appartiennent à ϵ_0, qui est supposé bien fondé : $\mathrm{P}_\perp$ est nécessairement vide.
Supposons maintenant qu'il existe une preuve ϕ de $\emptyset \vdash \emptyset$ de degré d non nul. Alors par définition

$$\dfrac{\dfrac{\dfrac{\phi}{\vdots}}{\dfrac{S(\phi)}{\vdots}}\,(\mathrm{E})}{\begin{array}{c}d \text{ applications} \\ \text{de la règle (E)}\end{array}}\,\dfrac{\vdots}{\dfrac{\vdots}{S(\phi)}}\,(\mathrm{E})$$

est une preuve de ϕ de degré nul (chaque application de la règle (E) fait descendre de 1 le degré), ce qui est contradictoire. □

Remarque 5. Étant donné que l'on est passé de LK à P en ajoutant essentiellement la règle d'induction sur les entiers naturels, on a essentiellement montré que l'induction sur les entiers naturels "était cohérente", en utilisant l'induction... jusqu'à ϵ_0. À première vue, il y a ici une arnaque : on aurait admis plus que ce que l'on veut obtenir !
En réalité, la preuve utilise un principe formel bien moins puissant que l'induction transfinie sur ϵ_0 : alors que l'induction sur P s'applique à n'importe quelle formule F, on n'a en fait utilisé ici qu'une induction sur une formule *sans quantificateurs*. C'est ce qui fait tout l'intérêt de la preuve : on demande une induction sur des ordinaux plus gros, mais sans quantificateurs, ce qui correspond à des opérations toujours calculables. On peut en particulier obtenir ce principe de démonstration par le biais des suites de Goodstein, comme nous allons le montrer dans la section suivante.

2.4 Suites de Goodstein et bonne fondation de ϵ_0

Pour finir démontrons la proposition suivante.

Proposition 2.4.1. *Le théorème de Goodstein implique dans* P *la bonne fondation de* ϵ_0*. Plus précisément,* P *et le théorème de Goodstein permettent de démontrer qu'il n'existe pas de suite strictement décroissante d'ordinaux strictement inférieurs à* ϵ_0 *qui soit récursive prouvablement totale (ie la suite des codes des ordinaux récursive prouvablement totale)*

Pour commencer, on définit une généralisation des suites de Goodstein (ou plutôt de la suite ordinale associée), les suites de Goodstein lentes. On démontrera tout d'abord que l'on peut à partir de toute suite d'ordinaux décroissants définir une suite de Goodstein lente qui la majore. Puis on démontrera que leur terminaison (le fait qu'elles finissent stationnaires en zéro) est équivalent au théorème de Goodstein (terminaison des suites de Goodstein standards).

Définition 2.4.2 (Suites de Goodstein lentes de α de transition ρ)**.** Soit ρ strictement croissante telle que $\rho(0) \geqslant 2$ et $\alpha < \epsilon_0$ $\rho(0)$-représentable. On définit $\gamma_n^{\alpha,\rho}$ (de nouveau, souvent notée γ_n) suite d'ordinaux par :

$$\begin{cases} \gamma_0 &= \alpha \\ \gamma_{n+1} &= \langle \gamma_n \rangle(\rho(n+1)) \quad \text{pour tout } n \in \mathbb{N} \end{cases}$$

Lemme 2.4.3. *Soit $(\lambda_i)_{i\in\mathbb{N}}$ une suite d'ordinaux strictement décroissante récursive prouvablement totale dans Peano. Il existe alors $\alpha < \epsilon_0$ et ρ récursive prouvablement totale dans* P *telle que pour tout $i \in \mathbb{N}$,*

$$\gamma_i^{\alpha,\rho} \geqslant \lambda_i$$

Démonstration. Soit $\rho(0) = 1 + \max$ des coefficients apparaissant dans l'écriture en base ω itérée de λ_0. Posons de plus pour tout n, $\rho(n+1) = 1 + \max$ de $\rho(n)$ et des coefficients apparaissant dans l'écriture en base ω itérée de λ_{n+1}. ρ est récursive prouvablement totale dans P car la suite $(\lambda_i)_{i\in\mathbb{N}}$ l'est et il est récursif prouvablement total de récupérer les coefficient et en faire le max. De plus ρ est strictement croissante et pour tout $i \in \mathbb{N}$, λ_i est $\rho(i)$-représentable.
Soit $(\gamma_n)_{n\in\mathbb{N}}$ la suite de Goodstein lente de λ_0 de transition ρ. Montrons par récurrence que pour tout $i \in \mathbb{N}$ on a $\gamma_i \geqslant \lambda_i$.
Si $i = 0$, comme $\gamma_0 = \lambda_0$, c'est vérifié.
Supposons $\gamma_i \geqslant \lambda_i$, $\gamma_{i+1} = \langle \gamma_i \rangle(\rho(i+1)) \geqslant \langle \lambda_i \rangle(\rho(i+1))$. $\lambda_i = 0$ est impossible vu que par définition $0 \leqslant \gamma_{i+1} < \gamma_i$, donc $\lambda_i \neq 0$ et $\langle \lambda_i \rangle(\rho(i+1))$ est le plus grand ordinal qui vérifie $\beta < \lambda_i$ et β est $\rho(i+1)$-représentable (cf. proposition 1.2.7). Or λ_{i+1} vérifie aussi ces deux propriétés donc $\langle \lambda_i \rangle(\rho(i+1)) \geqslant \lambda_{i+1}$. On a donc bien la propriété voulue. □

Lemme 2.4.4. *Soient $\alpha, \beta < \epsilon_0$ avec $\alpha \gg \beta$ et $(q_n)_{n\in\mathbb{N}}$ une suite d'entiers, s'il existe n_0 tel que $\langle \alpha + \beta \rangle(q_1, \ldots, q_{n_0}) \leqslant \alpha$, alors il existe $n \leqslant n_0$ tel que $\langle \alpha + \beta \rangle(q_1, \ldots, q_{n_0}) = \alpha$*

Démonstration. Si $\langle \beta \rangle(q_1, \ldots, q_k) > 0$ alors $\langle \alpha + \beta \rangle(q_1, \ldots, q_{k+1}) = \alpha + \langle \beta \rangle(q_1, \ldots, q_{k+1}$ (cela se démontre par récurrence sur k en appliquant la proposition 1.2.5). Donc si $\forall k \in \mathbb{N}$, $\langle \beta \rangle(q_1, \ldots, q_k) > 0$ alors $\forall k \in \mathbb{N}$, $\langle \alpha + \beta \rangle(q_1, \ldots, q_k) > \alpha$ ce qui rentre en contradiction avec les hypothèses. il existe donc un n minimal tel que $\langle \beta \rangle(q_1, \ldots, q_n) = 0$ et alors comme $\langle \beta \rangle(q_1, \ldots, q_{n-1}) > 0$, on a $\langle \alpha + \beta \rangle(q_1, \ldots, q_{n_0}) = \alpha$. □

Lemme 2.4.5. *Le théorème de Goodstein implique la terminaison de toutes les suites de Goodstein lentes à transition récursive prouvablement totale dans* P*.*

Démonstration. Soit $(\gamma_i)_{i\in\mathbb{N}}$ la suite de Goodstein lente de α et de transition ρ récursive prouvablement totale dans P.
Par le théorème de Wainer (théorème 1.4.1), il existe $\gamma < \epsilon_0$ tel que pour tout $x \in \mathbb{N}$, $\rho(x) < H_\gamma(x)$. Par la proposition 1.2.3, il existe $n_0 \in \mathbb{N}$ tel que $\alpha, \gamma < \omega_{n_0}$ et donc

$\rho <_* H_{\omega_{n_0}}$. Comme on cherche à montrer que la suite devient stationnaire en zéro, ce qui est une propriété qui ne concerne pas les premiers termes, on peut supposer que $\rho < H_{\omega_{n_0}}$. Par le théorème de Cichon (voir 1.2.18), en notant $h(\alpha, p) = \min\{z \mid \langle\alpha\rangle(p, \ldots, z) = 0\}$, pour tout $x \in \mathbb{N}$,

$$\rho(x) < H_{\omega_{n_0}}(x) = h(\omega_{n_0}, x)$$

Soit $\beta = \omega^{\omega_{n_0}+\alpha+1}$ où on a fait attention à avoir $\omega_{n_0} \gg \alpha$. Soit $(\beta_n)_{n\in\mathbb{N}}$ la suite ordinale associée à la suite de Goodstein standard commençant à $f_{\omega,\rho(0)+1}(\beta)$ et en base $\rho(0)+1$. On a alors bien $\beta_0 = \beta$ car α est par définition $\rho(0)$-représentable et ω_{n_0} est 2-représentable donc $\rho(0)+1$-représentable.
On pose $\phi(1) = \rho(0) + 2$.
Comme

$$\langle\beta\rangle(\rho(0)+2) = \omega^{\omega_{n_0}+\langle\alpha+1\rangle(\rho(0)+2)} \cdot (\rho(0)+1) + \langle\omega^{\omega_{n_0}+\langle\alpha+1\rangle(\rho(0)+2)}\rangle(\rho(0)+2)$$

et que la suite Goodstein finit stationnaire en zéro, par le lemme 2.4.4, il existe $\phi(1)$ tel que $\beta_{\phi(1)-1} = \omega^{\omega_{n_0}+\langle\alpha+1\rangle(\phi(0))}$.
Par récurrence, tant que $\langle\alpha+1\rangle(\phi(0), \ldots, \phi(i)) > 0$, on construit $\phi(i+1)$ comme le plus petit terme j supérieur à $\phi(i)$ tel que $\beta_{j-1} = \omega^{\omega_{n_0}+\langle\alpha+1\rangle(\phi(0),\ldots,\phi(i))}$. Si $\langle\alpha+1\rangle(\phi(0), \ldots, \phi(i)) = 0$ on prend $\phi(i+1) = \phi(i) + 1$.
Le théorème de Goodstein implique que β_n s'annule à partir d'un certain rang. Donc $\langle\alpha+1\rangle(\phi(0), \ldots, \phi(i))$ s'annule aussi à partir d'un certain rang car $\omega^\delta > 0$ pour tout $\delta < \epsilon_0$.
Comme $\omega_{n_0} > \alpha$, et $\rho(0) \geqslant 2$, par le lemme 2.4.4, pour tout $i \geqslant 1$ tel que $\langle\alpha+1\rangle(\phi(0), \ldots, \phi(i)) > 0$, avant d'atteindre $\phi(i+1)$ il faut annuler un terme de la forme $\omega^{\omega_{n_0}+\langle\alpha+1\rangle(\phi(0),\ldots,\phi(i))}$ et donc :

$$\begin{aligned}
\phi(i+1) &\geqslant \min\{j \mid \langle\omega^{\omega_{n_0}+\langle\alpha+1\rangle(\phi(0),\ldots,\phi(i))}\rangle(\phi(i)+1, \ldots, j) = 0\} \\
&\geqslant \min\{j \mid \langle\omega_{n_0}\rangle(\phi(i)+1, \ldots, j) = 0\} \\
&\geqslant h(\omega_{n_0}, \phi(i)+1) \\
&> \rho(\phi(i)+1) \\
&\geqslant \rho(i+2)
\end{aligned}$$

car ρ est strictement croissante et $\phi(i) \geqslant i+1$ car $\phi(0) \geqslant 3$ et ϕ est strictement croissante. On a démontré que la suite de Goodstein lente commençant en α et de transition ϕ s'annule à partir d'un certain rang. Comme $\phi(i) > \rho(i)$ pour tout i, par récurrence et en appliquant le fait que $\langle\alpha\rangle(n)$ soit croissante en fonction de n, on a

$$\langle\alpha+1\rangle(\phi(0), \ldots, \phi(i)) \geqslant \langle\alpha+1\rangle(\rho(0), \ldots, \rho(i)) = \langle\alpha\rangle(\rho(1), \ldots, \rho(i))$$

La suite de Goodstein lente commençant en α et de transition ρ termine donc aussi. □

Démonstration de la proposition 2.4.1. Soit $(\lambda_i)_{i\in\mathbb{N}}$ une suite d'ordinaux strictement décroissante récursive prouvablement totale dans P. Par le lemme 2.4.3, il existe une suite de Goodstein lente à transition récursive prouvablement totale dans P qui la majore. Le lemme 2.4.5 implique, comme on suppose le théorème de Goodstein, que la suite lente termine. Il existe donc $n \in \mathbb{N}$ tel que pour tout $i \geqslant n$, $\lambda_i = 0$ ce qui contredit la décroissance stricte. Une suite d'ordinaux appartenant à ϵ_0 strictement décroissante n'existe pas si on suppose le théorème de Goodstein. □

On remarquera cependant que la preuve du lemme 2.4.5 n'est pas finitiste telle qu'elle est donnée ici. En effet notre démonstration du théorème de Wainer ne l'est pas, mais la preuve donnée dans [BW87] passant par la théorie de la démonstration l'est.
Dans son article de 1944 (voir [Goo44]), Goodstein introduit ses suites en étant à la recherche d'une preuve finitiste la bonne fondation de ϵ_0 (c'est à dire démontrable dans P). Il aurait alors pu démontrer que la "reine Zahlentheorie" de Gentzen (et donc P) démontre sa propre cohérence et est donc incohérente, d'après le deuxième théorème d'incomplétude de Gödel.
On remarquera la grande ironie, ou la grande régularité de tout cela : si ses suites permettent bien de montrer la cohérence de P, son théorème par contre n'admet pas de preuve finitiste...

Annexe A

Tout ce que vous avez toujours voulu savoir sur les ordinaux sans jamais oser le demander

On introduira dans cet appendice quelques notions sur les ordinaux qui sont utiles à la compréhension de ce mémoire, il s'inspire du cours donné par François Loeser à la FIMFA.

A.1 Ordres et bons ordres

La notion d'ordinal étant essentiellement liée à celle de bon ordre, on commencera par introduire ces notions.

Définition A.1.1 (Ordre)**.** Soit X un ensemble, la relation binaire $<$ est un ordre, si elle est transitive ($\forall x, y, z,\ x < y \wedge y < z \Rightarrow x < z$) et antiréflexive ($\forall x,\ \neg\ x < x$)).
On note $x \leqslant y$ si $x < y$ ou $x = y$.

Définition A.1.2 (Bon ordre)**.** Soit $Y \subset X$, Y admet un plus petit élément x si $\forall y \in Y,\ x \leqslant y$.
Un ordre $<$ est un bon ordre si toute partie non vide admet un plus petit élément.
Un bon ordre est forcément total, c'est à dire que pour tout $x, y \in X$, on a soit $x = y$, soit $x < y$, soit $y < x$. En effet si $x \neq y$, alors $\{x, y\}$ admet un plus petit élément.

Lemme A.1.3 (Bonne fondation des bons ordres)**.** *Soit X un ensemble bien ordonné. Alors il n'existe pas dans X de suite infinie strictement décroissante*

On remarque d'ailleurs que c'est plutôt la bonne fondation que le bon ordre qui est utilisée dans ce mémoire.

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite strictement décroissante dans X. $\{x_n\}_{n\in\mathbb{N}}$ admet un plus petit élément car X est bien ordonné, soit n_0 l'indice de ce plus petit élément dans la suite. On a alors $x_{n_0+1} < x_{n_0}$ par stricte décroissance, ce qui contredit la minimalité de x_{n_0}. □

A.2 Ordinaux

A.2.1 Définition et première propriétés

Définition A.2.1 (Ordinal)**.** Un ensemble X est un ordinal s'il est bien ordonné par $\in$ et s'il est transitif ($\forall x \in X,\ \forall y \in x,\ y \in X$).

Les propriétés suivantes des ordinaux, qui sont des conséquences directes de la définition, donnent une première idée de ce que sont ces objets peut être un peu bizarres à première vue.

Proposition A.2.2 (Premières propriétés des ordinaux)**.** *Soit α un ordinal :*

(i) $\emptyset$ *est un ordinal.*

(ii) si $\alpha \neq \emptyset$*, alors* $\emptyset \in \alpha$*.*

(iii) $\alpha \notin \alpha$

(iv) Soit $\beta \in \alpha$ *alors* β *est un ordinal.*

(v) Soit $\beta \in \alpha$ *alors* $\beta = S_{<\beta} = \{\gamma \in \alpha \mid \gamma < \beta\}$*.*
On appelle segment initial tout $Y \subset X$ *ensemble ordonné tel que pour tout* $y \in Y$ *et tout* $x \in X$*, si* $x < y$ *alors* $x \in Y$*.*
On déduit du résultat précédent que tout segment initial d'un ordinal est un ordinal. Et plus précisément, si c'est un segment initial strict, c'est l'élément minimal des ordinaux qui ne sont pas dans le segment initial.

(vi) Soit β *un ordinal,* $\beta \subset \alpha \iff \beta = \alpha$ *ou* $\beta \in \alpha$*.*

(vii) $\alpha \cup \{\alpha\}$ *est un ordinal noté* α^+*. De plus si on a un ordinal* β *tel que* $\alpha < \beta$ *alors* $\alpha^+ \subset \beta$*.* α^+ *est donc le plus petit ordinal strictement supérieur à* α*.*

Démonstration. (i) La première propriété est vérifiée car $\emptyset$ ne contient aucun élément et la deuxième est vérifiée car l'appartenance est bien un ordre sur $\emptyset$ et qu'il n'a pas de sous partie non vide.

(ii) Comme $\in$ est un bon ordre pour α, α admet un plus petit élément β. Supposons que $\beta \neq \emptyset$, alors il existe $\gamma \in \beta$, mais on a alors $\gamma \in \alpha$ par transitivité et $\gamma < \beta$ ce qui contredit la minimalité de β dans α.

(iii) Si $\alpha = \emptyset$, par définition, $\alpha \notin \alpha$.
Sinon supposons $\alpha \in \alpha$. Soit $\beta \in \alpha$ comme $\in$ est un ordre dans α, on a, par antiréflexivité, $\beta \notin \beta$, en particulier $\alpha \notin \alpha$ ce qui est absurde.

(iv) La transitivité de α traduit exactement que $\beta \subset \alpha$. Il est donc bien ordonné. Pour ce qui est de la transitivité de β, soit $\delta \in \gamma \in \beta \subset \alpha$. Par transitivité de α, $\delta \in \alpha$. Comme α est totalement ordonné on a $\delta < \beta$ ou $\delta = \beta$ ou $\beta < \delta$, mais comme $\delta < \gamma < \beta$ dans les deux derniers cas, on a $\beta < \beta$ ce qui contredit l'antiréflexivité.

(v) Soit $\gamma \in \beta$, par transitivité de α, $\gamma \in \alpha$ et donc $\gamma \in S_{<\beta}$. Pour ce qui est de la deuxième inclusion, soit $\gamma \in S_{<\beta}$, on a alors $\gamma < \beta$, ie $\gamma \in \beta$.
Pour ce qui est de la deuxième affirmation, soit X un segment initial strict d' un ordinal α (s'il n'est pas strict, c'est α qui est un ordinal). Soit γ l'élément minimal de $\alpha \backslash X \neq \emptyset$. Montrons que $X = S_{<\gamma}$.

Soit $\delta \in S_{<\gamma}$, si $\delta \notin X$, on a $\delta \in \alpha \backslash X$ et $\delta < \gamma$, ce qui contredit la minimalité de γ.

Soit $\delta \in X$, si $\delta > \gamma$, comme X est un segment initial, on devrait avoir $\gamma \in X$ ce qui contredit le fait que $\gamma \in \alpha \backslash X$.

On a donc bien $X = S_{<\gamma} = \gamma$ d'après la première affirmation.

(vi) Montrons d'abord ($\Leftarrow$). Si $\beta \in \alpha$ alors d'après le (v), $\beta = \{\gamma \in \alpha \mid \gamma \in \beta\} \subset \alpha$. Et évidemment $\alpha \subset \alpha$.

Pour ($\Rightarrow$), supposons $\beta \subset \alpha$ et $\beta \neq \alpha$. On a alors $\alpha \backslash \beta \neq \emptyset$ qui admet un élément minimal γ.

Montrons que $\beta = S_{<\gamma}$. Soit $\delta \in \beta \subset \alpha$, on a alors $\delta \in \alpha$. Supposons que $\delta \geqslant \gamma$, alors $\gamma \in \beta$ ce qui est impossible, donc $\delta < \gamma$ et $\delta \in S_{<\gamma}$. Soit $\delta \in S_{<\gamma}$ alors, si $\delta \notin \beta$, $\delta \in \alpha \backslash \beta$ et comme $\delta < \gamma$, cela contredit la minimalité de γ. Donc $\delta \in \beta$ et on a bien $\beta = S_{<\gamma}$. Or, d'après le (v), $S_{<\gamma} = \gamma$, on a donc $\beta = \gamma \in \alpha$.

(vii) $\in$ est bien un ordre sur α^+ car si on a $x \in y \in z$ dans α^+, soit $z \in \alpha$ et donc $y \in \alpha$ et $x \in \alpha$ et comme c'est un ordre sur α on a bien $x \in z$, soit $z = \alpha$ et par transitivité de α on a bien $x \in \alpha$. Pour ce qui est de l'antiréflexivité, pour tout $x \in \alpha^+$, soit $x \in \alpha$ et $x \notin x$ par antiréflexivité de $\in$ dans α, soit $x = \alpha$ et d'après le (iii), $\alpha \notin \alpha$.

De plus c'est un bon ordre car si on considère une sous-partie $\mathcal{P}$ de α^+ soit $\mathcal{P} \subset \alpha$ et dans quel cas comme l'ordre est bien fondé sur α, $\mathcal{P}$ admet un plus petit élément, soit cette partie est $\{\alpha\}$ qui admet α comme plus petit élément, soit c'est $\mathcal{P} \cup \{\alpha\}$ où $\mathcal{P}$ est une sous-partie non vide de α qui admet donc un plus petit élément $\beta \in \mathcal{P} \subset \alpha$. On a donc $\beta < \alpha$ et c'est donc aussi l'élément minimal de $\mathcal{P} \cup \{\alpha\}$.

Supposons alors que $\alpha < \beta$. Par la proposition (vi), $\alpha \subset \beta$ donc $\alpha^+ = \alpha \cup \{\alpha\} \subset \beta$ donc, toujours par la proposition (vi), $\alpha^+ = \beta$ ou $\alpha^+ < \beta$.

□

Le théorème suivant est fondamental pour décrire la classe des ordinaux :

Théorème A.2.3. *Soient α et β deux ordinaux, on a alors :*
- *soit $\alpha \in \beta$*
- *soit $\alpha = \beta$*
- *soit $\beta \in \alpha$*

Démonstration. Soit $\gamma = \alpha \cap \beta$. γ est transitif car soit $\lambda \in \delta \in \gamma$, on a $\lambda \in \delta \in \alpha$ et $\lambda \in \delta \in \beta$et donc, par transitivité de α et β, $\lambda \in \alpha$ et $\lambda \in \beta$.
De plus $\in$ est un ordre, et un bon ordre, dans γ comme sous-ensemble d'ensemble bien ordonné. C'est donc un ordinal.
Si $\gamma = \alpha$, alors $\alpha \subset \beta$ et par la proposition A.2.2.(vi), $\alpha \in \beta$ ou $\alpha = \beta$.
Si $\gamma = \beta$, on a de même $\beta \in \alpha$ ou $\beta = \alpha$.
Enfin le cas $\gamma \neq \alpha$ et $\gamma \neq \beta$ est impossible car on a alors, toujours par la proposition A.2.2.(vi), comme $\gamma \subset \alpha$ et $\gamma \subset \beta$, $\gamma \in \alpha$ et $\gamma \in \beta$. Et donc $\gamma \in \gamma$ ce qui contredit la proposition A.2.2.(iii). □

Proposition A.2.4. *Si $\mathcal{A}$ est un ensemble non vide d'ordinaux, il contient un plus petit élément qui est $\bigcap_{\alpha \in \mathcal{A}}$.*

Démonstration. Par le même argument que dans la preuve précédente, $\beta = \bigcap_{\alpha\in\mathcal{A}} \alpha$ est un ordinal. De plus pour tout $\alpha \in \mathcal{A}$ on a $\beta \leqslant \alpha$.
Montrons par l'absurde que $\beta \in \mathcal{A}$. Si quelque soit $\alpha \in \mathcal{A}$ $\beta < \alpha$, alors $\beta \in \bigcap_{\alpha\in\mathcal{A}} \alpha = \beta$ ce qui rentre en contradiction avec la proposition A.2.2.(iii). □

Proposition A.2.5. *Soit $\mathcal{A}$ un ensemble d'ordinaux. Notons $\beta = \bigcup_{\alpha\in\mathcal{A}} \alpha$.*
Alors β est un ordinal et si γ est un ordinal tel que $\gamma < \beta$, alors il existe $\alpha \in \mathcal{A}$ tel que $\gamma < \alpha$.

Démonstration. Montrons d'abord que β est transitif. Soient $\delta \in \gamma \in \beta$. Il existe $\alpha \in \mathcal{A}$ tel que $\delta \in \gamma \in \alpha$. Par transitivité de α, $\delta \in \alpha$ et donc $\delta \in \beta$.
On munit β de l'ordre définit comme suit : soient $\gamma, \delta \in \beta$ il existe $\lambda, \mu \in \mathcal{A}$ tels que $\gamma \in \lambda$ et $\delta \in \mu$. Par le théorème A.2.3 et la proposition A.2.2.(vi), on a soit $\lambda \subset \mu$, soit $\mu \subset \lambda$. Supposons $\mu \subset \lambda$, on a alors $\gamma, \delta \in \lambda$ et on prend l'ordre dans λ. On vérifie aisément que c'est bien un ordre par hérédité dans les éléments de $\mathcal{A}$.
Montrons que c'est un bon ordre. Soit $\mathcal{X} \subset \beta$ non vide. $\mathcal{X}$ est un ensemble d'ordinaux donc d'après la proposition A.2.4, il admet un plus petit élément, $\bigcap_{x\in\mathcal{X}} x$.
La deuxième affirmation est juste la définition de l'union où les $<$ remplacent les $\in$. □

A.2.2 Terminologie

Définition A.2.6 (Successeur, limite)**.** Soit α un ordinal :

1. On dit que α^+ est le successeur de α.
2. Si $\alpha \neq \emptyset$ et si α n'est pas un successeur, on dit que α est limite.

La proposition qui suit relie un ordinal limite aux ordinaux inférieurs de même que l'on avait une relation entre un ordinal successeur et son prédécesseur.

Proposition A.2.7. *Soit $\alpha \neq \emptyset$ un ordinal. On a équivalence entre :*
(1) α limite
(2) $\alpha = \bigcup_{\beta\in\alpha} \beta$
(3) Si $\gamma \in \alpha$, alors $\gamma^+ \in \alpha$

Démonstration. Commençons par (1) $\Rightarrow$ (2). On a toujours $\bigcup_{\beta\in\alpha} \subset \alpha$, c'est la définition de la transitivité de α.
Montrons l'inclusion réciproque. Supposons que l'on ait pas $\alpha = \bigcup_{\beta\in\alpha} \beta$. Soit γ, le plus petit élément de $\alpha \setminus \bigcup_{\beta\in\alpha} \beta$.
On a alors $\gamma^+ \notin \alpha$, sinon comme $\gamma \in \gamma^+$ on aurait $\gamma \in \bigcup_{\beta\in\alpha} \beta$. On a donc $\gamma < \alpha \leqslant \gamma^+$ donc $\alpha = \gamma^+$ d'après la proposition A.2.2.(vii), c'est à dire α est successeur.
Montrons maintenant (2) $\Rightarrow$ (3). Soit $\gamma \in \alpha$. Comme $\alpha = \bigcup_{\beta\in\alpha} \beta$, il existe $\beta \in \alpha$ tel que $\gamma < \beta$ mais comme γ^+ est le plus petit ordinal strictement supérieur à γ, $\gamma^+ \leqslant \beta < \alpha$.
Montrons enfin (3) $\Rightarrow$ (1). Supposons que α soit successeur. On a donc $\alpha = \gamma^+$. En particulier, $\gamma \in \alpha$ mais $\gamma^+ = \alpha$ n'appartient pas à α d'après la proposition A.2.2.(iii). □

Proposition A.2.8.

Définition A.2.9 (Ordinaux finis, ω)**.** Soit α un ordinal :

1. α est dit fini si ni α ni aucun éléments de α n'est limite.
2. On note ω le plus petit ordinal limite.
3. Pour tout $n \in \mathbb{N}$, on note $\underline{n}$ (ou plus simplement n s'il n'y a pas d'ambiguïté sur le fait que ce soit un ordinal) l'unique ordinal de cardinal n. On a donc :

$$\begin{aligned} \underline{0} &= \emptyset \\ \underline{1} &= \{\emptyset\} \\ \underline{n+1} &= \underline{n}^+ \end{aligned}$$

Cette dernière définition a un sens car il y a un seul ordinal de cardinal n. En effet, si α est de cardinal n alors comme $\alpha \notin \alpha$, α^+ est de cardinal $n+1$. Donc si on a $\beta > \alpha$, alors comme $\alpha^+ \subset \beta$ par la proposition A.2.2.(vii), β est de cardinal supérieur à $n+1$. De plus n^+ est bien de cardinal $n+1$ et donc, par une récurrence évidente, pour tout $n \in \mathbb{N}$ il existe un ordinal de cardinal n.

Proposition A.2.10. *Les ordinaux finis sont exactement les ordinaux de cardinal fini. De plus ce sont exactement les éléments de ω, d'où $\omega = \bigcup_{n\in\mathbb{N}} \underline{n}$.*

ω est isomorphe à $(\mathbb{N}, <)$.

Démonstration. Montrons d'abord que les éléments de ω sont les ordinaux finis. Soit α un ordinal fini. D'après le théorème A.2.3, on a soit $\alpha \in \omega$, soit $\alpha = \omega$ soit $\omega \in \alpha$, mais dans les deux derniers cas α est soit limite soit contient un ordinal limite et n'est donc pas fini. Donc tous les ordinaux finis appartiennent à ω. Soit maintenant $\alpha < \omega$ comme ω est le plus petit ordinal limite, α n'est pas limite. Il ne contient pas non plus d'ordinal β limite car sinon par transitivité de $\beta < \omega$ ce qui contredit la définition de ω. Donc tous les éléments de ω sont finis. On a donc montré que α fini $\iff \alpha \in \omega$.
Montrons ensuite, par récurrence que les ordinaux de cardinal fini sont finis.
Tout d'abord $\underline{0}$ est fini, il ne contient rien donc a priori pas d'ordinal limite et n'est pas limite lui même.
De plus supposons $\underline{n}$ fini, alors comme $\underline{n+1} = n^+$, il n'est pas limite. De plus $\underline{n+1} = \underline{n} \cup \{\underline{n}\}$, donc comme $\underline{n}$ n'est pas limite et ne contient pas d'ordinaux limites, alors $\underline{n+1}$ non plus.
Montrons maintenant que $\beta = \bigcup_{n\in\mathbb{N}} \underline{n}$ est limite. Tout d'abord pour tout $n \in \mathbb{N}$, $\beta > \underline{n}$, en effet $\underline{n}^+ = \underline{n+1} \subset \beta$ donc $\underline{n} < \underline{n}^+ \leqslant \beta$, ce qui implique en autre que β ne peut être de cardinal fini sinon on aurait $\underline{n} = \beta > \underline{n}$ pour un certain $n \in \mathbb{N}$.
Supposons au contraire que β soit successeur, c'est à dire que $\beta = \gamma^+$. γ ne peut pas être de cardinal fini sinon β le serait aussi. Donc pour tout $n \in \mathbb{N}$, par le théorème A.2.3, comme on ne peut pas avoir $\gamma \in \underline{n}$ (sinon par transitivité de n, γ serait aussi de cardinal fini car inclut dans $\underline{n}$) ni $\gamma = \underline{n}$, on a $\underline{n} < \gamma$. Donc $\beta = \bigcup_{n\in\mathbb{N}} \underline{n} \subset \gamma \in \gamma^+ = \beta$ ce qui est impossible par la proposition A.2.2.(iii).
Donc $\omega \leqslant \beta = \bigcup_{n\in\mathbb{N}} \underline{n}$, en particulier quelque soit α fini, comme $\alpha \in \omega \subset \beta$, il existe, par la proposition A.2.5, $n \in \mathbb{N}$ tel que $\alpha \in \underline{n}$. En particulier α est de cardinal fini. Donc ω est l'ensemble de tous les ordinaux de cardinal fini.

L'isomorphisme que l'on cherche est donc simplement :

$$\begin{cases} \mathbb{N} & \to & \omega \\ n & \mapsto & \underline{n} \end{cases}$$

□

A.2.3 Ordinaux et bons ordres

On conclura cette partie par le théorème suivant qui indique le lien qui existe entre ordinaux et bons ordres. Tout bon ordre peut être représenté par un et un seul ordinal.

Théorème A.2.11 (Ordinaux et bons ordres)**.** *Soit X un ensemble bien ordonné. Alors il existe un unique ordinal α et une unique bijection $f : X \to \alpha$ telle que pour tout $x \in X$, si $x < y$ alors $f(x) < f(y)$*
On dit que X et α sont isomorphes en tant qu'espaces ordonnés.

Démonstration. On commencera par montrer l'unicité de cet ordinal et de cette fonction si ils existent avant de montrer qu'ils existent.
Supposons que l'on a $f : X \to \alpha$ et $f' : X \to \alpha'$ deux tels isomorphismes. Posons $g = f^{-1} \circ f' : \alpha \to \alpha'$ qui est un isomorphisme. Supposons que $g \neq \mathrm{Id}$. Soit β_0 le plus petit élément de α tel que $g(\beta_0) \neq \beta_0$. Quelque soit $\gamma \in \beta_0$, $g(\gamma) = \gamma$. Or comme g est un isomorphisme, $\gamma \in \delta \iff g(\gamma) \in g(\delta)$ et particulier $\gamma \in \beta_0 \iff \gamma = g(\gamma) \in g(\beta_0)$, donc $g(\beta_0) = \beta_0$ ce qui est absurde.
Donc $g = \mathrm{Id}$ et donc $\alpha = \alpha'$ et $f = f'$.
Pour montrer l'existence, on remarque d'abord que, pour tous $x \in X$ et α un ordinal, s'il existe un isomorphisme entre $S_{<x}$ et *alpha* alors on étendant l'isomorphisme par α en x on obtient un isomorphisme entre $S_{\leqslant x}$ et α^+. Montrons maintenant que quelque soit $x \in X$ il existe un ordinal $\alpha(x)$ et un isomorphisme f_x entre $S_{\leqslant x}$ et α. Sinon, soit x_0 l'élément minimal tel qu'ils n'existent pas. Quelque soit $y < x$ on a $\alpha(y)$ isomorphe par f_y a $S_{<y}$. Par unicité pour tout $y' < y$, $f_y|_{S_{\leqslant y'}} = f_{y'}$. On pose $\alpha = \bigcup_{y<x_0} \alpha(y)$ et $f : S_{<x_0} \to \alpha$ par $f(y) = f_y(y)$. f est bien un isomorphisme car $f|_{S_{\leqslant y}} = f_y$. D'après la remarque liminaire on a un isomorphisme entre $S_{\leqslant x_0}$ et α^+, ce qui rentre en contradiction avec la définition de x_0.
On procède de même pour construire l'ordinal et l'isomorphisme pour X. Soit $\alpha = \bigcup_{x\in X} \alpha(x)$ et $f : X \to \alpha$ tel que $f(x) = f_x(x)$. Comme pour tout $x \in X$, $f|_{S_{<x}} = f_x$, f est bien un isomorphisme. □

A.3 Arithmétique ordinale

La troisième et dernière partie de cette introduction aux ordinaux traitera la question des opérations sur les ordinaux : l'addition, la multiplication et l'exponentiation ; leur définition et leurs propriétés.

A.3.1 Addition

Définition A.3.1. Addition de deux ensembles ordonnés Soient $(A, <)$ et $(B, <)$ deux ensembles ordonnés. On note $A + B$ l'ensemble $\{(a, 0) \mid a \in A\} \sqcup \{(b, 1) \mid b \in B\}$ (où $\sqcup$

indique l'union disjointe). On muni $A + B$ de l'ordre lexicographique, c'est à dire pour $c, d \in (A \cup B)$ et $i, j \in \{0, 1\}$, on a $(c, i) < (d, j))$ si $i < j$ ou si $i = j$ et $c < d$.

Proposition A.3.2. *Si $(A, <)$ et $(B, <)$ sont biens ordonnés, alors $(A + B, <_{lex})$ est bien ordonné.*

Démonstration. Soit $X \subset A + B$ une partie non vide. Notons $X_A = X \cap (A, 0)$ et $X_B = X \cap (B, 1)$ où $(A, 0) = \{(a, 0) \mid a \in A\}$ et $(B, 1) = \{(b, 1) \mid b \in B\}$. Comme $A + B = (A, 0) \sqcup (B, 1)$, $X = X_A \sqcup X_B$. Supposons $X_A \neq \emptyset$. Comme $((A, 0), <_{lex})$ est canoniquement isomorphe à $(A, <)$, il est aussi bien ordonné. X_A admet donc un plus petit élément $m = (a, 0)$. Par définition de l'ordre lexicographique, pour tout $b \in B$, $(a, 0) < (b, 1)$ et donc m est bien le plus petit élément de X. Si $X_A = \emptyset$ alors $X = X_B$ et vu que, comme précédemment, $(B, 1)$ est bien ordonné, alors X_B admet un plus petit élément qui est donc aussi le plus petit élément de X. □

On peut alors définir la somme de deux ordinaux

Définition A.3.3. Somme d'ordinaux Soient deux ordinaux α et β. On note $\alpha+\beta$ l'ordinal isomorphe à l'ensemble bien ordonné $\alpha + \beta$.

la somme que l'on vient de définir a les propriétés suivantes :

Proposition A.3.4 (Propriétés de la somme d'ordinaux)**.** *Soient α, β, γ trois ordinaux,*

(i) *La somme d'ordinaux est associative, ie $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.*

(ii) *0 est neutre à gauche et à droite, ie $\alpha + 0 = 0 + \alpha = \alpha$.*

(iii) *$\alpha + 1 = \alpha^+$.*

(iv) *$\alpha < \beta$, si et seulement s'il existe $\gamma > 0$ tel que $\beta = \alpha + \gamma$.*

(v) *Si $\beta < \gamma$ alors $\alpha + \beta < \alpha + \gamma$. En particulier si $\alpha + \beta = \alpha + \gamma$ alors $\beta = \gamma$ (la somme d'ordinaux est intègre à gauche).*

(vi) *Si β est limite alors $\alpha + \beta = \bigcup_{\gamma \in \beta} \alpha + \gamma$.*

(vii) *Si α est fini, $\alpha + 1 = 1 + \alpha$. Si α est infini, $1 + \alpha = \alpha$.*

Démonstration. (i) On a en fait la propriété plus générale que si A, B, C sont des ensembles ordonnés, alors $(A+B)+C$ est isomorphe à $A+(B+C)$ en tant qu'ensembles ordonnés. On pourra vérifier que

$$\begin{cases} (A + B) + C & \to & A + (B + C) \\ ((a, 0), 0) & \mapsto & (a, 0) \\ ((b, 1), 0) & \mapsto & ((b, 0), 1) \\ (c, 1) & \mapsto & ((c, 1), 1) \end{cases}$$

est l'isomorphisme qui convient. L'unicité de l'ordinal isomorphe à un ensemble bien ordonné (théorème $A.2.11$) permet de conclure dans le cas des ordinaux. $(\alpha + \beta) + \gamma$ et $\alpha + (\beta + \gamma)$ sont deux ordinaux isomorphes à deux ensembles ordonnés isomorphes. Ils sont donc isomorphes et donc égaux.

(ii) $\alpha + 0 = (\alpha, 0)$ muni de l'ordre lexicographique, qui est canoniquement isomorphe à $(\alpha, \in)$. De même $0 + \alpha = (\alpha, 1)$ muni de l'ordre lexicographique, qui est canoniquement isomorphe à $(\alpha, \in)$.

(iii) Comme $\alpha+1$ est isomorphe à $(\alpha,0)\sqcup\{(1,1)\}$, et qu'on a l'isomorphisme d'ensembles ordonnés suivant :

$$\begin{cases} \alpha+1 & \to & \alpha^+ \\ (\beta,0)\in(\alpha,0) & \mapsto & \beta \\ (1,1) & \mapsto & \alpha \end{cases}$$

on a bien $\alpha+1$ isomorphe à α^+. Comme ce sont tous deux des ordinaux, ils sont égaux par le théorème A.2.11.

(iv) Montrons d'abord ($\Rightarrow$). D'après la proposition A.2.2.(v), comme $\alpha\in\beta$, $\alpha=S_{<\alpha}\subsetneq\beta$. Notons $X=\beta\backslash\alpha$. X est bien ordonné, non vide et est donc isomorphe à un ordinal $\gamma\neq 0$. De plus quelque soit $\delta\in X$, $\delta\geqslant\alpha$ donc quelque soit $\lambda\in\alpha$, $\lambda<\delta$. On en déduit que $\alpha+\gamma$ est isomorphe à β.

Montrons maintenant ($\Leftarrow$). Considérons le morphisme injectif suivant :

$$\phi:\begin{cases} \alpha & \to & (\alpha,0)\sqcup(\gamma,1) \\ \delta & \mapsto & (\delta,0) \end{cases}$$

qui envoie α sur un segments initial. En notant ψ l'isomorphisme entre $(\alpha,0)\sqcup(\gamma,1)$ et $\alpha+\gamma$, on a un morphisme injectif $\psi\circ\phi$ qui envoie α sur un segment initial de $\alpha+\gamma$. Par la proposition A.2.2.(v), $\mathrm{Im}(\alpha)$ est un ordinal. Il est isomorphe à α, par le théorème A.2.11, c'est α lui même. On a donc $\alpha\subset\alpha+\gamma$.

Si $\gamma\neq 0$, ϕ n'est pas surjectif donc $\psi\circ\phi$ ne l'est pas non plus. On en déduit donc que $\alpha\subsetneq\alpha+\gamma$. Par la proposition A.2.2.(vi), on a $\alpha\in\alpha+\gamma$.

(v) D'après la proposition (iv), il existe $\delta>0$ tel que $\gamma=\beta+\delta$. On a donc $\alpha+\gamma=\alpha+\beta+\delta=(\alpha+\beta)+\delta>\alpha+\beta$ toujours par la proposition (iv).

Supposons $\alpha+\beta=\alpha+\gamma$, si $\beta<\gamma$, $\alpha+\beta<\alpha+\gamma$ ce qui est absurde, de même si $\gamma<\beta$ alors $\alpha+\gamma<\alpha+\beta$ ce qui est absurde. Donc $\beta=\gamma$.

(vi) Notons $\delta=\bigcup_{\gamma\in\beta}\alpha+\gamma$. D'après la proposition (v), pour tout $\gamma\in\beta$, $\alpha+\gamma\in\alpha+\beta$. Donc, d'après la proposition A.2.2.(vi), $\alpha+\gamma\subset\alpha+\beta$ donc $\delta\subset\alpha+\beta$.

Réciproquement, soit $\lambda\in\alpha+\beta$. Supposons d'abord que $\lambda\geqslant\alpha$. D'après la proposition (iv), il existe $\gamma\in\beta$ tel que $\lambda=\alpha+\gamma$ donc $\lambda\in\delta$.

Si $\lambda<\alpha$, comme β est limite, $0\in\beta$. Donc $\lambda\in\alpha+0\subset\delta$ ce qui permet de conclure.

(vii) Soit $\alpha=\underline{n}$ un ordinal fini. Comme $\underline{1}=\{\emptyset\}$, $\underline{1}+\underline{n}$ est isomorphe à $\{(\emptyset,0)\}\sqcup(\underline{n},1)$ qui est de cardinal $n+1$. C'est donc l'ordinal de cardinal $n+1$, ie $\underline{n+1}=\underline{n}^+=\alpha^+=\alpha+\underline{1}$ d'après la proposition (iii).

Montrons d'abord que $\underline{1}+\omega=\omega$. Par définition $\underline{1}+\omega$ est isomorphe à $\{(\emptyset,0)\}\sqcup(\omega,1)$. On dispose de plus de l'isomorphisme suivant :

$$\begin{cases} \{(\emptyset,0)\}\sqcup(\omega,1) & \to & \omega \\ (\emptyset,0) & \mapsto & \underline{0} \\ (\underline{n},1) & \mapsto & \underline{n+1} \end{cases}$$

$\underline{1}+\omega$ est donc isomorphe à ω. Ils sont donc égaux.

Soit α un ordinal qui ne soit pas fini. $\alpha\geqslant\omega$ d'après la proposition A.2.10. D'après la proposition (iv), il existe $\beta\geqslant 0$ tel que $\alpha=\omega+\beta$. On en déduit que $1+\alpha=1+\omega+\beta=\omega+\beta=\alpha$. □

A.3.2 Multiplication

Définition A.3.5 (Produit de deux ensembles ordonnés)**.** Soient $(A, <)$ et $(B, <)$ deux ensembles ordonnés. On note $A \cdot B$ le produit cartésien $A \times B$ muni de l'ordre lexicographique (soient $a, c \in A$ et $b, d \in B$, $(a, b) <_{lex} (c, d)$ si $b < d$ ou si $b = d$ et $a < c$).

Proposition A.3.6. *Si $(A, <)$ et $(B, <)$ sont biens ordonnés, alors $(A \cdot B, <_{lex})$ est bien ordonné.*

Démonstration. Soit $X \subset A \times B$ non vide.
On note $X_B = \{b \in B \mid \exists a \in A, (a, b) \in X\} \subset B$. Comme $X \neq \emptyset$, il existe $(a, b) \in A \times B$ tel que $(a, b) \in X$. On a alors $b \in X_B$ qui est donc non vide. Comme B est bien ordonné, X_B a un élément minimal b_0. Notons $X_A^{b_0} = \{a \in A \mid (a, b_0) \in X\} \subset A$. Comme $b_0 \in X_B$, il existe $b \in B$ tel que $(a, b_0) \in X$. On a alors $a \in X_A^{b_0}$ qui est donc non vide. Comme A est bien ordonné, notons a_0 l'élément minimal de $X_A^{b_0}$.
Montrons que (a_0, b_0) est l'élément minimal de X. Soit $(a, b) \in X$, par définition $b \in X_B$ et donc $b_0 \leqslant b$. Si $b_0 < b$, on a bien $(a_0, b_0) < (a, b)$.
Sinon $b = b_0$ et donc $a \in X_A^{b_0}$. On en déduit que $a_0 \leqslant a$ et donc $(a_0, b_0) \leqslant (a, b)$. □

On définit alors le produit de deux ordinaux de la manière suivante :

Définition A.3.7. produit de deux ordinaux] Soient α, β deux ordinaux, on note $\alpha \cdot \beta$ l'ordinal isomorphe à l'ensemble bien ordonné $\alpha \cdot \beta$.

Le produit que l'on vient de définir vérifie les propriétés suivantes :

Proposition A.3.8 (Propriétés du produit d'ordinaux)**.** *Soient α, β, γ trois ordinaux,*

(i) *Le produit d'ordinaux est associatif, ie $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$,*

(ii) $0 \cdot \alpha = \alpha \cdot 0 = 0$,

(iii) *1 est neutre à gauche et à droite, ie $1 \cdot \alpha = \alpha \cdot 1 = \alpha$,*

(iv) *La multiplication est distributive à gauche sur l'addition, ie $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$,*

(v) $\omega \cdot 2 \neq 2 \cdot \omega = \omega$,

(vi) *Si $\alpha \neq 0$ et $\beta < \gamma$ alors $\alpha \cdot \beta < \alpha \cdot \gamma$. En particulier, si $\alpha \cdot \beta = \alpha \cdot \gamma$ alors $\beta = \gamma$, ie la multiplication est intègre à gauche,*

(vii) *Si β est limite, $\alpha \cdot \beta = \bigcup_{\gamma \in \beta} \alpha \cdot \gamma$,*

(viii) *si $\beta \neq 0$, il existe λ, ρ tels que $\alpha = \beta \cdot \lambda + \rho$ avec $\rho < \beta$.*

Démonstration. (i) De même que dans le cas de l'addition, on a la propriété plus générale suivante : si A, B, C sont des ensembles ordonnés alors $(A \cdot B) \cdot C$ est isomorphe à $A \cdot (B \cdot C)$. On pourra vérifier que

$$\begin{cases} (A \times B) \times C & \to & A \times (B \times C) \\ ((a, b), c) & \mapsto & (a, (b, c)) \end{cases}$$

est l'isomorphisme qui convient. Dans le cas d'ordinaux, comme deux ordinaux isomorphes sont égaux (théorème A.2.11), on a bien $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$.

(ii) $\emptyset \times \alpha = \alpha \times \emptyset = \emptyset$ ce qui permet de conclure.

(iii) $\{\emptyset\} \times \alpha$ est canoniquement isomorphe à α par l'application qui a $(\emptyset, \beta)$ associe β, (de même que $\alpha \times \emptyset$).

(iv) on vérifiera que l'application

$$\begin{cases} A \times ((B,0) \sqcup (C,1)) & \to & (A \times B, 0) \sqcup (A \times C, 1) \\ (a,(b,0)) & \mapsto & ((a,b),0) \\ (a,(c,1)) & \mapsto & ((a,c),1) \end{cases}$$

est l'isomorphisme qui convient.

(v) On vérifiera aisément que l'application suivante est un isomorphisme :

$$\begin{cases} \underline{2} \times \omega & \to & \omega \\ (\emptyset, \underline{n}) & \mapsto & \underline{2n} \\ (\{\emptyset\}, \underline{n}) & \mapsto & \underline{2n+1} \end{cases}$$

ω et $2 \cdot \omega$ sont donc isomorphes. Ils sont donc égaux.

De plus $\omega \cdot 2 = \omega + \omega$ par distributivité. On ne peut donc pas avoir $2 \cdot \omega = \omega \cdot 2$ sinon on aurait $\omega = \omega + \omega$ et donc $\omega = 0$ ce qui est absurde car $1 \in \omega$.

(vi) Comme $\gamma > \beta$, il existe $\delta > 0$ tel que $\gamma = \beta + \delta$. On a alors $\alpha \cdot \gamma = \alpha \cdot \beta + \alpha \cdot \delta$.

De plus $\alpha \cdot \delta = 0$ si et seulement si $\alpha = 0$ ou $\delta = 0$. En effet si $\alpha \neq \emptyset$ et $delta \neq \emptyset$ alors $\alpha \times \delta \neq \emptyset$ et ne peut donc pas être en bijection avec l'ensemble vide.

Comme ici $\alpha \neq 0$ et $\delta \neq 0$ par hypothèse, $\alpha \cdot \delta > 0$ et donc $\alpha \cdot \gamma > \alpha \cdot \beta$.

Supposons $\beta \cdot \alpha = \gamma \cdot \alpha$. Si $\beta > \gamma$, alors $\alpha \cdot \beta > \alpha \cdot \gamma$ ce qui est absurde, et si $\beta < \gamma$, alors $\alpha \cdot \beta < \alpha \cdot \gamma$ ce qui est aussi absurde. Donc $\beta = \gamma$.

(vii) D'après la proposition (vi), pour tout $\gamma \in \beta$, $\alpha \cdot \gamma < \alpha \cdot \beta$ donc, d'après la proposition A.2.2.(vi), $\alpha \cdot \gamma \subset \alpha \cdot \beta$, donc $\bigcup_{\gamma \in \beta} \alpha \cdot \gamma \subset \alpha \cdot \beta$.

Montrons maintenant l'inclusion réciproque. Montrons d'abord que si on a un ordinal δ tel que pour tout $\gamma \in \beta$, $\delta > \alpha \cdot \gamma$ alors $\delta \geqslant \alpha \cdot \beta$. Montrons plus précisément que $\alpha \cdot \beta$ est isomorphe à un segment initial de δ.

En effet, comme β est limite, $\alpha \cdot \beta = \alpha \cdot \bigcup_{\gamma \in \beta} \gamma \simeq \alpha \times \bigcup_{\gamma \in \beta} \gamma$. Pour tout $\gamma \in \beta$, on note f_γ l'isomorphisme entre $\alpha \times \gamma$ et $\alpha \cdot \gamma$. Soit $(\lambda, \mu) \in \alpha \times \bigcup_{\gamma \in \beta} \gamma$, il existe donc $\gamma \in \beta$ tel que $\mu \in \gamma$. Posons $\phi(\lambda, \mu) = f_\gamma(\lambda, \mu)$. Cette application est bien définie car si $\mu \in \gamma' \in \beta$, on peut supposer $\gamma \leqslant \gamma'$. on remarque qu'alors $\alpha \times \gamma$ est un segment initial de $\alpha \times \gamma'$ donc comme $f_{\gamma'}$ est un morphisme $f_{\gamma'}(\alpha \times \gamma)$ est un segment initial de $\alpha \cdot \gamma'$, c'est donc un ordinal d'après la proposition A.2.2.(v). Cet ordinal est isomorphe à $\alpha \times \gamma$ par $f_{\gamma'}|_{\alpha \times \gamma}$ qui est lui même isomorphe à $\alpha \cdot \gamma$ par f_γ. D'après le théorème A.2.11, $f_{\gamma'}|_{\alpha \times \gamma} = f_\gamma$. En particulier comme $(\lambda, \mu) \in \alpha \times \gamma$, $f_{\gamma'}(\lambda, \mu) = f_\gamma(\lambda, \mu)$. ϕ est donc bien définie.

C'est un morphisme injectif car si on a $\mu_1, \mu_2 \in \bigcup_{\gamma \in \beta} \gamma$, alors il existe $\gamma \in \beta$ tel que $\mu_1, \mu_2 \in \gamma$, et, pour tout $\gamma \in \beta$, $\phi|_{\alpha \times \gamma} = f_\gamma$ qui est un morphisme injectif.

De plus $\phi(\alpha \times \bigcup_{\gamma \in \beta} \gamma)$ est un segment initial. En effet, soient $\mu \in \phi(\alpha \times \bigcup_{\gamma \in \beta} \gamma)$ et $\lambda \in \delta \backslash \phi(\alpha \times \bigcup_{\gamma \in \beta} \gamma)$. Il existe $\gamma \in \beta$, $\nu \in \gamma$ et $\kappa \in \alpha$ tels que $\mu = \phi(\kappa, \nu) = f_\gamma(\kappa, \nu) \in \alpha \cdot \gamma$. Comme $\alpha \cdot \gamma \subset \phi(\alpha \times \bigcup_{\gamma \in \beta} \gamma)$ et que c'est un segment initial de δ, on a bien $\lambda \in \delta \backslash (\alpha \cdot \gamma)$ et donc $\mu < \lambda$.

On a donc un isomorphisme entre $\alpha \cdot \beta \simeq \alpha \times \bigcup_{\gamma \in \beta} \gamma$ et un segment initial de δ. On en conclut donc que $\alpha \cdot \beta \leqslant \delta$.

On peut maintenant conclure. Notons $\delta = \bigcup_{\gamma \in \beta} \alpha \cdot \gamma$. Comme β est limite, pour tout $\beta \in \gamma$, $\gamma^+ \in \beta$ d'après la proposition A.2.7 et donc $\alpha \cdot \gamma < \alpha \cdot \gamma^+ \leqslant \delta$. On en déduit d'après le résultat précédent que $\delta \geqslant \alpha \cdot \beta$.

(viii) On démontrera cette propriété en supposant que l'on sait que $\alpha \in \beta \cdot \gamma$ pour un certain γ. On peut démontrer par exemple que $\alpha \leqslant \beta \cdot \alpha$ (et donc $\alpha < \beta \cdot \alpha^+$) en montrant qu'il existe un morphisme injectif de α dans $\beta \cdot \alpha$ puis en montrant que si un ordinal s'injecte dans un autre alors il en est un segment initial. Mais cette dernière proposition nécessiterai un certain travail.

Soit $\delta = \min\{\delta \in \gamma^+ \mid \beta \cdot \delta > \alpha\}$. Ce plus petit élément existe car cet ensemble est non vide, il contient γ. Montrons que δ est forcément successeur. Si δ était limite, on aurait alors $\alpha \cdot \delta = \bigcup_{\gamma \in \delta} \alpha \cdot \gamma$ d'après la proposition (vii). Comme $\alpha \in \beta \cdot \delta$, il existerai $\gamma \in \delta$ tel que $\alpha \in \beta \cdot \gamma$ ce qui rentre en contradiction avec la minimalité de δ. On en déduit que δ est successeur, ie $\delta = \lambda^+$. Par minimalité de δ, $\beta \cdot \lambda \leqslant \alpha$. Il existe donc ρ tel que $\alpha = \beta \cdot \lambda + \rho$.

Il reste à montrer que $\rho < \beta$. Si ce n'était pas le cas, on aurait

$$\begin{aligned} \alpha &= \beta \cdot \lambda + \rho \\ &\geqslant \beta \cdot \lambda + \beta \\ &= \beta \cdot \lambda^+ \\ &> \alpha \end{aligned}$$

ce qui est impossible. □

A.3.3 Exponentiation

Définition A.3.9 (Puissance d'un ensemble ordonné par un autre)**.** Soient $(A, <)$ et $(B, <)$ deux ensembles ordonnés, et on suppose que A possède un plus petit élément noté 0. On note $A^{(B)}$ l'ensemble des fonctions de A dans B à support fini (c'est à dire, en notant $\mathrm{supp}(f) = \{b \in B \mid f(b) \neq 0\}$ le support de f, $\{f : B \to A \mid |\mathrm{supp}(f)| < \infty\}$). On le muni de l'ordre lexicographique : soient $f, g \in A^{(B)}$, $f <_{lex} g$ s'il existe $b_0 \in B$ tel que $f(b_0) < g(b_0)$ et pour tout $b \in B$ tel que $b > b_0$, on ait $f(b) = g(b)$.

Proposition A.3.10. *Si $(A, <)$ et $(B, <)$ sont biens ordonnés, alors $(A^{(B)}, <_{lex})$ est bien ordonné.*

Démonstration. Soit $X_0 \subset A^{(B)}$ non vide. Si X_0 contient la fonction identiquement nulle, comme c'est l'élément minimal de $A^{(B)}$, c'est aussi celui de X_0. On peut donc supposer que toutes les fonctions dans X_0 sont à support non vide.

Soit $f \in X_0$, on note $s(f) = \max(\mathrm{supp}(f))$ (qui existe car $\mathrm{supp}(f)$ est non vide de cardinal fini). Soit $b_1 = \min\{s(f) \mid f \in X_0\}$, soit $a_1 = \min\{f(b_1) \mid f \in X_0 \text{ et } s(f) = b_1\}$, soit

$$X_1 = \{f \in X_0 \mid s(f) = b_1 \text{ et } f(b_1) = a_1\}$$

et enfin soit

$$X_1' = \{f \in A^{(B)} \mid f(b_1) = 0 \text{ et } \exists g \in X_1,\ \forall b \neq b_1,\ f(b) = g(b)\}$$

X_1 est un segment initial de X_0. En effet, soit $f \in X_0 \backslash X_1$ et $g \in X_1$. On a $s(f) \geqslant b_1$
Supposons $s(f) > b_1$, $g(s(f)) = 0 < f(s(f))$ de plus quelque soit $b > s(f) > s(g) = b_1$, $f(b) = 0 = g(b)$. On a donc bien $g < f$.
Sinon $s(f) = b_1$, on a alors $f(b_1) > a_1 = g(b_1)$ sinon f serait dans X_1. De plus, quelque soit $b > b_1$, $f(b) = 0 = g(b)$. On a donc bien $g < f$.
De plus on peut vérifier que l'application suivante est un isomorphisme d'ensembles ordonné :

$$\begin{array}{ccc} X_1' & \stackrel{\phi}{\to} & X_1 \\ f & \mapsto & \begin{pmatrix} B & \to & A \\ b_1 & \mapsto & a_1 \\ b \neq b_1 & \mapsto & f(b) \end{pmatrix} \end{array}$$

De plus, si on a le plus petit élément de X_1', on a celui de X_0. En effet soit f_1 le plus petit élément de X_1', comme ϕ est un morphisme, $f_0 = \phi(f_1)$ est le plus petit élément de X_1. Comme X_1 est un segment initial de X_0, f_0 est aussi le plus petit élément de X_0.
Si X_1' contient la fonction identiquement nulle, on s'arrête là sinon on construit b_2, a_2, X_2, X_2' à partir de X_1' comme on a construit b_1, a_1, X_1, X_1' à partir de X_0, et ainsi de suite.
La suite des X_i' que l'on construit est forcément finie car $b_i > b_{i+1}$. En effet quelque soit $f \in X_i'$, en notant $g = \phi(f) \in X_i$, on a $\mathrm{supp}(f) = \mathrm{supp}(g) \backslash b_i$ or $b_i = \max(\mathrm{supp}(g))$ donc $s(f) = \max(\mathrm{supp}(f)) < b_i$. On en déduit que $b_{i+1} = \min\{s(f) \mid f \in X_i'\} < b_i$.
Si cette suite n'était pas finie on aurait une suite infinie strictement décroissante dans un ensemble bien ordonné, ce qui est impossible d'après la propriété A.1.3.
Il existe donc i_0 tel que X_{i_0} contient la fonction identiquement nulle, c'est donc son élément minimal. Comme on l'a remarqué précédemment, si on connaît l'élément minimal de X_{i+1} on sait construire l'élément minimal de X_i. On en déduit donc l'élément minimal de X_0. □

On peut alors définir l'exponentiation d'ordinaux.

Définition A.3.11 (Puissance d'un ordinal par un autre)**.** Soient α, β deux ordinaux. On note α^β l'ordinal isomorphe à l'ensemble bien ordonné $\alpha^{(\beta)}$.

L'exponentiation que l'on vient de définir vérifie les propriétés suivantes :

Proposition A.3.12. *Soient α, β, γ trois ordinaux,*

(i) $\alpha^0 = 1$, $\alpha^1 = \alpha$, $1^\alpha = 1$ *et si* $\alpha \neq 0$, $0^\alpha = 0$,

(ii) $\alpha^{\beta+\gamma} = \alpha^\beta \cdot \alpha^\gamma$ *et* $\alpha^{\beta \cdot \gamma} = (\alpha^\beta)^\gamma$,

(iii) Si $\alpha > 1$ *et* $\beta < \gamma$ *alors* $\alpha^\beta < \alpha^\gamma$,

(iv) si β *est limite,* $\alpha^\beta = \bigcup_{\gamma \in \beta} \alpha^\gamma$.

Démonstration. (i) Comme $0 = \emptyset$, α^0 est isomorphe à l'ensemble des fonctions à support fini du vide dans α, il y en a une, celle à image vide. α^0 est donc l'ordinal à un élément.

α^1 est l'ensemble des fonction (à support forcément fini) de $\{\emptyset\}$ dans α. C'est $\{0 \mapsto \beta\}_{\beta \in \alpha}$ qui est canoniquement isomorphe à α.

1^α est l'ensemble des fonction à support fini de α dans $\{0\}$.Il y en a une seule, la fonction identiquement nulle (qui est bien à support finie) si α est non vide (le cas $\alpha = 0$ a déjà été traité). 1^α est donc l'ordinal à un élément.

Soit $\alpha \neq 0$, 0^α est inclus dans l'ensemble des fonctions d'un ensemble non vide dans l'ensemble vide, ie l'ensemble vide.

(ii) On vérifiera que les deux applications

$$\left\{ \begin{array}{ccc} & A^{((B,0)\sqcup(C,1))} & \to & A^{(B)} \times A^{(C)} \\ \left(\begin{array}{ccc} (b,0) & \mapsto & f_B(b) \\ (c,1) & \mapsto & f_C(c) \end{array} \right) & \mapsto & (f_B, f_C) \end{array} \right.$$

$$\left\{ \begin{array}{ccc} A^{(B\times C)} & \to & (A^{(B)})^{(C)} \\ f & \mapsto & (c \mapsto (b \mapsto f(b,c))) \end{array} \right.$$

sont définies et sont les isomorphismes qui conviennent.

(iii) Comme $\beta < \gamma$, il existe $\delta > 0$ tel que $\gamma = \beta + \delta$. On a donc $\alpha^\gamma = \alpha^{\beta+\delta} = \alpha^\beta \cdot \alpha^\delta$. Comme $\delta \neq 0$ et $\alpha > 1$, $\alpha^{(\delta)} > 1$, il contient au moins la fonction identiquement nulle et la fonction qui a 0 associe 1 et 0 au reste. D'après la proposition A.3.8.(vi), $\alpha^\beta \cdot \alpha^\delta < \alpha^\beta \cdot 1 = \alpha^\beta$.

(iv) Si $\gamma \in \beta$, d'après la proposition (iii), $\alpha^\gamma < \alpha^\beta$, donc $\bigcup_{\gamma\in\beta} \alpha^\gamma \subset \alpha^\beta$.

Montrons maintenant l'inclusion réciproque. Soit $\gamma \in \alpha^\beta$, on l'identifie à $f : \beta \to \alpha$ à support fini. Montrons que, comme β est limite, il existe $\delta \in \beta$ tel que $\mathrm{supp}(f) \subset \delta$. On peut par exemple prendre $\max\{\delta^+ \mid \delta \in \mathrm{supp}(f)\}$ (c'est le plus grand élément d'un ensemble fini totalement ordonné, il existe donc).

On note $g = f_{|\delta} \in \alpha^{(\delta)}$. Soit ϕ_1 l'isomorphisme entre $\alpha^{(\beta)}$ et α^β et ϕ_2 l'isomorphisme entre $\alpha^{(\delta)}$ et α^δ. On a, de plus, ϕ_3 le morphisme injectif suivant :

$$\left\{ \begin{array}{ccc} \alpha^{(\delta)} & \to & \alpha^{(\beta)} \\ f & \mapsto & \left(\begin{array}{ccc} \lambda \in \delta & \mapsto & f(\lambda) \\ \lambda \geqslant \delta & \mapsto & 0 \end{array} \right) \end{array} \right.$$

L'image de ce morphisme est un segment initial de $\alpha^{(\beta)}$, donc l'image de $\alpha^{(\delta)}$ par $\phi_1 \circ \phi_3$ est un ordinal. Comme $\alpha^{(\delta)}$ est isomorphe à α^{delta} par ϕ_2, d'après le théorème A.2.11, $\phi_1 \circ \phi_3 = \phi_2$.

Comme $\mathrm{supp}(f) \subset \gamma$, $\phi_3(g) = f$. Donc $\gamma = \phi_1 \circ \phi_3(g) = \phi_2(g) \in \alpha^\delta$. On en déduit que $\alpha^\beta \subset \bigcup_{\delta\in\beta} \alpha^\delta$. □

Bibliographie

[AZ97] Z. Adamowicz and P. Zbierski. *Logic of Mathematics*, chapter 22. Wiley-interscience, 1997.

[Buc97] Wilfried Buchholz. Explaining gentzen's consistency proof within infinitary proof theory. In *KGC '97 : Proceedings of the 5th Kurt Gödel Colloquium on Computational Logic and Proof Theory*, pages 4–17, London, UK, 1997. Springer-Verlag.

[BW87] W. Buchholz and S.S. Wainer. Provably computable functions and the fast growing hierarchy. In *Contemporary Mathematics, Logic and Combinatorics*, volume 65, pages 179–198. American Mathematical Society, 1987.

[Cic83] E.A. Cichon. A short proof of two recently discovered independence results using recursion theoretic methods. *Proodedings of the American Mathematical Society*, 87(4) :704–706, Avril 1983.

[Gen38] G. Gentzen. Neue fassung des widerspruchsfreiheitsbeweises fÃ$\frac{1}{4}$r die reine zahlentheorie. *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften*, 4 :1–44, 1938.

[Goo44] R.L. Goodstein. On the restricted ordinal theorem. *The Journal of Symbolic Logic*, 9(2) :33–41, Juin 1944.

[GTL89] Jean Y. Girard, Paul Taylor, and Yves Lafont. *Proofs and types*. CUP, Cambridge, 1989.

[KH89] C.F. Kent and B.R. Hogdson. Extensions of arithmetic for proving termination of computations. *The Journal of Symbolic Logic*, 54(3) :779–794, Septembre 1989.

[KP76] L.A.S Kirby and J.B. Paris. Initials segments of peano's axioms. In *Proceedings of the Bierutowice Conference*. Springer-Verlag, Berlin et New York, 1976.

[KP82] L.A.S. Kirby and J. B. Paris. Accessible independant results for peano's arithmetic. *Bulletin of the London Mathematical Society*, 14 :285–293, 1982.

[KS81] J. Ketonen and R. Solovay. Rapidly growing ramsey functions. *The Annals of Mathematics*, 113(2) :267–314, Mars 1981.

[Par78] J.B. Paris. Some independence results for peano arithmetic. *The Journal of Symbolic Logic*, 43(4) :725–731, Decembre 1978.

[Par79] J.B. Paris. A hierarchy of cuts in models of arithmetic. In *Proceedings of the Karpacz Conference*. Lecture Notes in Mathematics, Springer Berlin, 1979.

[Sch56] K. Schütte. Beweistheoretische erfassung der unendliche induktion in der zahlentheorie. *Mathemastiche Annalen*, 122 :369–389, 1956.

[Tak75] G. Takeuti. *Proof Theory*, volume 81. American Elsevier, 1975.

[Wai70] S.S. Wainer. A classification of the ordinal recursive functions. *Archiv für matematishe Logik und Grundlagenforschung*, 13 :136–153, 1970.

[Wai72] S.S. Wainer. Ordinal recursion and a refinement of the extended grzegorczyk hierarchy. *The Journal of Symbolic Logic*, 37(2) :281–292, Juin 1972.

Table des matières