Le problème de l'inclusion des automates de Parikh faiblement non ambigus

8.1 Introduction

8.1.1 Introduction au problème de l'inclusion

Dans ce chapitre, nous étudions une conséquence algorithmique de la propriété d'holonomie des séries de comptage des automates de Parikh faiblement non ambigus. Le problème de l'inclusion est le problème de décision suivant :

`entrée` : deux automates $\mathcal{A}$ et $\mathcal{B}$

`sortie` : est-ce que $\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B})$?

La décidabilité et la complexité de ce problème dépendent des classes d'automates dont sont issus $\mathcal{A}$ et $\mathcal{B}$. Si $\mathcal{A}$ et $\mathcal{B}$ sont des automates finis, le problème est classiquement décidable, et est PSPACE-complet [MS72]. Si les automates sont non ambigus, le problème est décidable en temps polynomial [SI85], par un argument de comptage que nous allons développer en détail dans cette introduction. Si $\mathcal{A}$ et $\mathcal{B}$ sont des grammaires hors-contextes, ce problème est indécidable, car l'universalité est déjà indécidable [BHPS61, Theorem 6.2]. Le problème reste indécidable si $\mathcal{A}$ et $\mathcal{B}$ sont des grammaires déterministes [GG66, Theorem 5.3], et donc *a fortiori* si les grammaires sont non ambiguës ([AN00] fournit une preuve directe dans le cas non ambigu, par réduction du problème de correspondance de Post). D'autres variantes ont été étudiées, en mélangeant les classes (par exemple si $\mathcal{A}$ est une grammaire hors-contexte quelconque, et $\mathcal{B}$ un automate fini, le problème est EXPTIME-complet [KI92]).

Une approche simple pour aborder le problème de l'inclusion, d'un point de vue purement automate, consiste à utiliser l'équivalence suivante :

$$\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B}) \Leftrightarrow \mathcal{L}(\mathcal{A}) \cap \overline{\mathcal{L}(\mathcal{B})} = \emptyset\,,$$

où $\overline{\mathcal{L}(\mathcal{B})}$ désigne le complémentaire de $\mathcal{L}(\mathcal{B})$. Il suffit pour décider l'inclusion de calculer ce complémentaire, lorsque c'est possible, puis de calculer l'intersection,

et enfin de tester le vide du langage obtenu. Cette approche souffre du recours à la complémentation : même dans le cas des automates finis qui sont clos par complémentaire, cette opération a un coût exponentiel ; par ailleurs pour des classes plus compliquées, le calcul du complémentaire n'est pas pas forcément faisable (les langages algébriques (non ambigus) ne sont pas clos par complémentaire [HU66], et nous ne savons pas si les automates de Parikh faiblement non ambigus sont clos par complémentaire). Pour contourner ce problème, dans le cadre des automates finis non ambigus, Stearns et Hunt [SI85] ont utilisé plutôt l'équivalence suivante :

$$\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B}) \Leftrightarrow \mathcal{L}(\mathcal{A}) \cap \overline{\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})} = \emptyset\,.$$

Mathématiquement, il s'agit exactement de la même égalité ; en effet en utilisant les lois de De Morgan, $\mathcal{L}(\mathcal{A}) \cap \overline{\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})} = \mathcal{L}(\mathcal{A}) \cap \overline{\mathcal{L}(\mathcal{B})}$. Cependant, considérer l'intersection $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$ est intéressant car l'inclusion $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B}) \subseteq \mathcal{L}(\mathcal{A})$ est toujours vraie, avec égalité si et seulement si $\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B})$. Par conséquent, pour décider le problème de l'inclusion, il suffit, pour tout $n \in \mathbb{N}$, de vérifier qu'il y a bien autant de mots de longueur n dans $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$ que de mots de longueur n dans $\mathcal{L}(\mathcal{A})$. Comme $\mathcal{A}$ et $\mathcal{B}$ sont des automates finis non ambigus, $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$ est reconnu par un automate fini non ambigu $\mathcal{C}$, obtenu en utilisant la construction produit. Notons $A(x) = \sum_n a_n x^n$ et $C(x) = \sum_n c_n x^n$ les séries génératrices de $\mathcal{L}(\mathcal{A})$ et $\mathcal{L}(\mathcal{C}) = \mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$. Le raisonnement par dénombrement de Stearns et Hunt s'écrit sous la forme :

$$\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B}) \Leftrightarrow \forall n \in \mathbb{N}, a_n - c_n = 0\,.$$

La non ambiguïté et l'astuce de l'intersection permettent ainsi de remplacer le calcul des mots d'un complémentaire, qui nécessite une construction exponentielle, par le dénombrement des mots de ce complémentaire, qui est bien plus simple. En effet, en utilisant la non ambiguïté des automates, il est possible de montrer que la suite (a_n) (resp. (c_n)) satisfait une récurrence linéaire à coefficients constants, d'ordre borné par $|Q_{\mathcal{A}}|$ (resp. $|Q_{\mathcal{A}}||Q_{\mathcal{B}}|$), où $|Q_{\mathcal{A}}|$ et $|Q_{\mathcal{B}}|$ désignent le nombre d'états de $\mathcal{A}$ et $\mathcal{B}$. Les suites de cette forme sont closes par soustraction, si bien que $d_n := a_n - c_n$ satisfait une équation de la forme :

$$\forall n \in \mathbb{N},\ d_{n+r} = \frac{1}{a_r}(a_{r-1}d_{n+r-1} + \ldots + a_0 d_n)\,,$$

avec $a_r \neq 0$, et $r \leqslant |Q_{\mathcal{A}}||Q_{\mathcal{B}}| + |Q_{\mathcal{A}}|$. Cette égalité implique que (d_n) est la suite nulle si et seulement si ses r premiers termes sont nuls. Stearns et Hunt ont donc démontré l'équivalence suivante :

$$\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B}) \Leftrightarrow \forall n \in \{0, \ldots, (|Q_{\mathcal{A}}|+1)|Q_{\mathcal{B}}| - 1\},\ a_n - c_n = 0\,.$$

Il suffit donc pour résoudre le problème de l'inclusion de comparer les nombres de mots acceptés par $\mathcal{A}$ et $\mathcal{C}$ pour toutes les tailles inférieures à $(|Q_{\mathcal{A}}|+1)|Q_{\mathcal{B}}| - 1$, ce qui se fait en temps polynomial en la taille de $\mathcal{A}$ et de $\mathcal{B}$. Stearns et Hunt déduisent aussi de cet argument que si $\mathcal{L}(\mathcal{A}) \not\subseteq \mathcal{L}(\mathcal{B})$, alors il existe un mot w de longueur plus petite que $|Q_{\mathcal{A}}||Q_{\mathcal{B}}| + |Q_{\mathcal{A}}|$, tel que $w \in \mathcal{L}(\mathcal{A})$ mais $w \notin \mathcal{L}(\mathcal{B})$.

Dans ce chapitre, nous cherchons à étendre la méthode de Stearns et Hunt, qui s'est avérée fructueuse dans le cas des automates finis, aux automates de Parikh non ambigus. Le problème est indécidable pour des automates de Parikh quelconques ; il

Rappelons que l'universalité des automates de Parikh est indécidable [KR02]

est co-NEXP-complet pour les automates de Parikh déterministes [FGM19], lorsque l'ensemble semilinéaire est représenté sous la forme d'une formule de Presburger existentielle. La décidabilité du problème de l'inclusion pour les automates de Parikh faiblement non ambigus peut se déduire des travaux de [CM17] sur la classe RCM, mais les auteurs ne fournissent pas de borne de complexité, et passent sous silence la complication due aux racines du polynôme de tête de la récurrence.

Notre but est donc de déterminer une borne $n_{\max}(\mathcal{A}, \mathcal{B})$ telle que :

$$\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B}) \Leftrightarrow \forall n \in \{0, \dots, n_{\max}(\mathcal{A}, \mathcal{B})\},\ a_n - c_n = 0\,.$$

Si la méthode de Stearns et Hunt s'adapte aux automates de Parikh faiblement non ambigus, elle se confronte cependant aux difficultés suivantes :

— les séries génératrices $A(x)$ et $C(x)$ ne s'obtiennent plus directement à partir des automates en résolvant un simple système linéaire. En effet, ces séries ne sont plus rationnelles, mais holonomes, et l'opération utilisée pour obtenir leurs équations différentielles, que ce soit par un produit d'Hadamard ou une diagonale de série, n'est pas triviale.

— la suite $d_n = a_n - c_n$ satisfait une équation de récurrence linéaire *à coefficients polynomiaux*, et non plus constants, de la forme :

$$p_r(n)d_{n+r} = \sum_{k=0}^{r-1} p_k(n)d_{n+k}$$

avec $p_r(n)$ qui n'est pas le polynôme nul. Les racines du polynôme $p_r(n)$ compliquent le comportement de la suite (d_n) : il ne suffit plus que les r premiers termes de d_n soient nuls pour que la suite soit nulle. Par exemple, le polynôme $D(x) = x^{1000}$ satisfait l'équation $1000D(x) - x\partial_x D(x) = 0$, et la récurrence $(n - 1000)d_n = 0$. Il est clair que vérifier que $d_0 = 0$ ne suffit pas à affirmer que (d_n) est la suite nulle. Pour pouvoir s'assurer que la suite (d_n) est nulle, il faut dépasser à la fois l'ordre de la récurrence et la plus grande racine entière du polynôme de tête. Pour borner cette racine, nous devons borner aussi la taille des coefficients des polynômes de la récurrence.

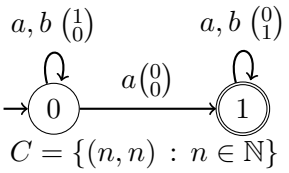
Ces deux difficultés compliquent considérablement l'analyse du problème de l'inclusion dans le cas des automates de Parikh faiblement non ambigus.

8.1.2 Introduction à l'algorithme de Lipshitz

Comme nous l'avons vu aux chapitres précédents, la propriété principale pour démontrer que la série génératrice d'un automate de Parikh non ambigu est holonome est la stabilité des séries holonomes par *diagonale* (ou produit d'Hadamard). Il s'agit d'une opération non triviale sur les fonctions holonomes, qui est toujours l'objet de recherche active en calcul formel. La clôture des fonctions holonomes par diagonale a été démontrée par Lipshitz en 1988 [Lip88] ; sa preuve repose sur des séries formelles, et consiste à interpréter la diagonale d'une série comme une extraction de coefficients. Cette extraction de coefficients, du point de vue analyse complexe, avec la formule de Cauchy, s'exprime sous la forme d'une intégrale (ou période), ce qui a donné lieu à de nouvelles méthodes de calcul de diagonale, plus récentes et efficaces, fondées sur le *télescopage créatif* ([Chy14, BCLS18]). Nous nous sommes limités dans cette

thèse à l'algorithme historique de Lipshitz [Lip88], qui a l'avantage d'être facile à comprendre et à étudier (mais s'avère peu efficace en pratique).

Pour donner une idée du principe de l'algorithme de Lipshitz, nous illustrons son fonctionnement dans cette introduction, sur un exemple simple. Soit $\mathcal{A}_{\text{ex}}$ l'automate de Parikh non ambigu suivant :



L'automate $\mathcal{A}_{\text{ex}}$ reconnaît les mots de longueur impaire qui ont un a au milieu. Ce langage est en fait algébrique non ambigu, et sa série génératrice $L(x) = \frac{x}{1-4x^2}$ est même rationnelle. Regardons comment l'algorithme de Lipshitz permet de trouver une équation différentielle satisfaite par $L(x)$.

Comme nous l'avons démontré au chapitre 7, en notant $A(x, y_1, y_2) = \frac{x}{(1-2xy_1)(1-2xy_2)}$ la série multivariée des calculs de l'automate, et $C(y_1, y_2) = \frac{1}{1-y_1y_2}$ la série du semilinéaire, la série $L(x)$ s'exprime comme le produit d'Hadamard $A(x, y_1, y_2) \odot \frac{1}{1-x}C(y_1, y_2)$, spécialisé en $y_1 = y_2 = 1$. L'idée de Lipshitz est d'exprimer le produit d'Hadamard sous la forme d'une extraction de coefficient :

$$A(x, y_1, y_2) \odot \frac{1}{1-x}C(y_1, y_2) = [u_1^{-1}u_2^{-1}]\frac{1}{u_1u_2}A(x, u_1, u_2)C(y_1/u_1, y_2/u_2)\,.$$

Même si nous sortons du monde des séries entières en écrivant $C(y_1/u_1, y_2/u_2)$, des arguments algébriques permettent d'étendre la notion d'holonomie à ce type de séries formelles. Plutôt que de calculer ce produit d'Hadamard en trois variables, pour évaluer deux variables à 1 ensuite, nous effectuons la substitution avant l'extraction de coefficient (nous justifierons dans la suite que nous avons le droit de le faire, et que les objets considérés sont bien définis). Ainsi :

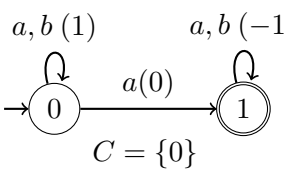
$$L(x) = [u_1^{-1}u_2^{-1}]\frac{1}{u_1u_2}A(x, u_1, u_2)C(u_1^{-1}, u_2^{-1})\,.$$

Pour que les calculs restent raisonnables dans cette introduction, nous simplifions encore cette formule. En remarquant que la contrainte du semilinéaire demande l'égalité des exposants des variables y_1 et y_2, qui s'exprime déjà par une extraction de coefficient bien choisie, nous pouvons vérifier simplement que :

$$L(x) = [y^{-1}]\tfrac{1}{y}A(x, y, y^{-1}) = [y^{-1}]\frac{x}{(1-2xy)(y-2x)}\,.$$

Remarque 8.1.

▶ Cette formule a une interprétation simple si on considère une généralisation équivalente aux automates de Parikh faiblement non ambigus, avec des vecteurs dans $\mathbb{Z}^d$ (sans contrainte de positivité sur les vecteurs étiquetant les calculs). Ainsi l'automate suivant reconnaît le même langage que $\mathcal{A}_{\text{ex}}$:



et la série des calculs finaux de cet automate est bien $A(x, y, y^{-1})$. ◀

Nous posons $F = \frac{P}{Q}$ avec $P = x$ et $Q = (1 - 2xy)(y - 2x)$. L'argument de Lipshitz consiste à remarquer que si on arrive à trouver une équation différentielle pour F de la forme :

$$\sum_{i=i_0}^{r} p_i(x, \partial_x)\partial_y^i F(x, y) = 0$$

avec les p_i des opérateurs différentiels, *qui ne dépendent pas de* y, et $p_{i_0}(x, \partial_x) \neq 0$, alors $p_{i_0}(x, \partial_x)L(x) = 0$.

La preuve de Lipshitz dit, par un argument de dimension, qu'il est toujours possible de trouver une telle équation pour F, en considérant la famille des polynômes $Q^{N+1}\partial_x^j\partial_y^i F$ avec $i + j < N$, pour des valeurs croissantes de N. Vérifions-le sur notre exemple : pour $N = 3$, nous demandons à Maple de résoudre le système suivant dans $\mathbb{Q}(x)$:

$$\lambda_1 Q^4 F + \lambda_2 Q^4 \partial_x F + \lambda_3 Q^4 \partial_y F + \lambda_4 Q^4 \partial_x^2 F + \lambda_5 Q^4 \partial_x\partial_y F + \lambda_6 Q^4 \partial_y^2 F = 0\,.$$

Nous trouvons alors la relation de dépendance suivante (le gras est juste utilisé pour rendre l'équation plus lisible) :

$$\begin{aligned} 0 = 16\ &\left(4\,x^2+3\right) x^4 \boldsymbol{Q^4\partial_x F} \\ &+ 2\,x^3\left(2\,x-1\right)\left(2\,x+1\right)\left(4\,x^2+1\right) \boldsymbol{Q^4\partial_x^2 F} \\ &+ \left(2\,x-1\right)\left(2\,x+1\right)\left(16\,x^4+16\,x^2+1\right) \boldsymbol{Q^4\partial_y F} \\ &+ x\left(4\,x^2+1\right)\left(2\,x-1\right)^2\left(2\,x+1\right)^2 \boldsymbol{Q^4\partial_x\partial_y F} \end{aligned}$$

En extrayant le coefficient de plus petit degré en ∂_y (les deux premières lignes), puis en divisant par $2x^3Q^4$, nous obtenons l'équation différentielle suivante satisfaite par $L(x)$:

$$\left(2\,x-1\right)\left(2\,x+1\right)\left(4\,x^2+1\right)\partial_x^2 L(x) + 8x\left(4\,x^2+3\right)\partial_x L(x) = 0\,.$$

On vérifie facilement que $L(x) = \frac{x}{1-4x^2}$ est bien solution de cette équation différentielle. On remarque que la méthode de Lipshitz ne donne pas une équation minimale satisfaite par $L(x)$; en effet, en tant que fraction rationnelle, $L(x)$ vérifie une équation du premier ordre :

$$x(1-2x)(1+2x)\partial_x L(x) - (1+4x^2)L(x) = 0\,.$$

Nous avons ainsi vu dans cet exemple introductif que la méthode de Lipshitz peut se résumer à trouver une relation de dépendance sur des polynômes bien choisis, obtenus en dérivant une fraction rationnelle issue d'un automate. Une partie centrale de ce chapitre consiste ainsi à analyser cet algorithme de Lipshitz, afin d'obtenir des bornes sur les degrés, l'ordre et surtout la taille des coefficients de l'équation différentielle satisfaite par la série génératrice d'un automate de Parikh non ambigu.

Nous rappelons enfin que l'idée de ce chapitre n'est pas d'implémenter l'algorithme de Lisphitz, ni d'autres algorithmes de calcul formel pour résoudre le problème de l'inclusion. Nous cherchons juste à obtenir des bornes sur les tailles des équations différentielles en jeu, pour répondre au problème de l'inclusion par une simple énumération de mots reconnus par les automates.

8.1.3 Cadre de travail, notations, et plan du chapitre

Dans la suite de ce chapitre, un automate de Parikh faiblement inambigu de dimension $d \geqslant 1$ est donné sous la forme d'un tuple $\mathcal{A} = (\Sigma, Q, q_I, F, C, \Delta)$ où on rappelle que Σ est l'alphabet, Q l'ensemble d'états, $q_I \in Q$ l'état initial, $F \subseteq Q$ l'ensemble des états finaux, $C \subseteq \mathbb{N}^d$ son ensemble semilinéaire d'acceptation, et $\Delta \subseteq Q \times (\Sigma \times \mathbb{N}^d) \times Q$ la relation de transition. On suppose que C est donné sous une forme inambiguë $C = \biguplus_{i=1}^{p} \boldsymbol{c_i} + P_i^*$.

En particulier il n'y a pas de ε-transitions

Définition 8.2 (Taille d'un automate de Parikh faiblement inambigu).
▶ Nous introduisons les notations suivantes pour quantifier la taille de $\mathcal{A}$:
- on désigne par $|\mathcal{A}| := |Q| + |\Delta| + p + \sum_i |P_i|$ la taille unaire de $\mathcal{A}$;
- pour $j \in [d]$, on note $\|\mathcal{A}\|_{j,\infty}$ la valeur maximale qui apparaît en coordonnée j dans les vecteurs de Δ, des P_i et des $\boldsymbol{c_i}$;
- on note $\|\mathcal{A}\|_\infty = \max_j \|\mathcal{A}\|_{j,\infty}$ la coordonnée maximale qui apparaît dans toutes les composantes des vecteurs de Δ, des P_i et des $\boldsymbol{c_i}$;
- on notera, pour $q \in Q$, $\deg_q^{out}$ le nombre de transitions sortantes de l'état q.

◀

Les sections 8.2 et 8.3 sont dédiées au calcul de la série génératrice d'un automate de Parikh non ambigu, en bornant notamment les coefficients de leur équation différentielle. Le résultat principal de ces deux sections est résumé dans la proposition suivante :

Proposition 8.3 (Taille de l'équation différentielle satisfaite par un langage de Parikh faiblement non ambigu).
▶ Soit $\mathcal{A}$ un automate de Parikh faiblement non ambigu. Alors la série $L(x)$ du langage de $\mathcal{A}$ satisfait une équation différentielle linéaire de la forme :

$$q_s(x)\partial_x^s L(x) + \cdots + q_0(x)L(x) = 0\,,$$

avec $s \leqslant (d|\mathcal{A}|\,\|\mathcal{A}\|_\infty)^{O(d)}$, et pour tout $i \in [0, s]$,

$$\deg(q_i) \leqslant (d|\mathcal{A}|\,\|\mathcal{A}\|_\infty)^{O(d^2)}\,,$$
$$\log \|q_i\|_1 \leqslant (d|\mathcal{A}|\,\|\mathcal{A}\|_\infty)^{O(d^2)}\,.$$

◀

Ensuite, la section 8.4 exploite les majorations obtenues précédemment pour borner la taille du mot minimal qui témoigne de la non inclusion des langages $\mathcal{L}(\mathcal{A})$ et $\mathcal{L}(\mathcal{B})$; les deux résultat principaux de cette section sont les suivants :

Théorème 8.4 (Taille du plus petit témoin de non inclusion).
▶ Si $\mathcal{L}(A)$ n'est pas inclus dans $\mathcal{L}(B)$, alors il existe un mot w qui soit dans $\mathcal{L}(\mathcal{A})$ mais pas dans $\mathcal{L}(\mathcal{B})$ tel que :

$$|w| \leqslant 2^{(dM)^{O(d^2)}}\,,$$

avec $d = d_\mathcal{A} + d_\mathcal{B}$, et $M = |\mathcal{A}||\mathcal{B}| \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)$. ◀

Corollaire 8.5 (Borne de complexité du problème de l'inclusion).
▶ Soient deux automates de Parikh faiblement non ambigus $\mathcal{A}$ and $\mathcal{B}$ de dimensions $d_\mathcal{A}$ et $d_\mathcal{B}$. On peut décider le problème de l'inclusion $L(\mathcal{A}) \subseteq L(\mathcal{B})$ en temps $2^{2^{O(d^2 \log(dM))}}$ où $d = d_\mathcal{A} + d_\mathcal{B}$ et $M = |\mathcal{A}|\,|\mathcal{B}|\,\max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)$. ◀

8.2 Séries génératrices rationnelles associées à un automate de Parikh

8.2.1 Notations, bornes sur les déterminants, règle de Cramer

Dans ce chapitre, nous allons travailler avec des polynômes multivariés à coefficients rationnels, c'est-à-dire des éléments de $\mathbb{Q}[x_1, \ldots, x_n]$. On note $\boldsymbol{\alpha} \in \mathbb{N}^n$ le *multi-indice* $(\alpha_1, \ldots, \alpha_n)$ et $x^{\boldsymbol{\alpha}}$ le monôme $x^{\boldsymbol{\alpha}} = \prod_{i=1}^n x_i^{\alpha_i}$. Le *degré total* de $x^{\boldsymbol{\alpha}}$ est défini par $\sum_{i=1}^n \alpha_i$, mais dans la suite nous allons principalement travailler avec les degrés partiels donnés par le multi-indice, et le *degré maximal* $\deg_m(x^{\boldsymbol{\alpha}}) = \max_{i=1}^n \alpha_i$.

On rappelle qu'un polynôme non nul P de $\mathbb{Q}[x_1, \ldots, x_n]$ est une combinaison linéaire finie de monômes sur $\mathbb{Q}$

$$P = \sum_{\boldsymbol{\alpha} \in \mathrm{Supp}(P)} \lambda_{\boldsymbol{\alpha}} x^{\boldsymbol{\alpha}},$$

où $\mathrm{Supp}(P)$ est un ensemble fini de multi-indices, vérifiant $\lambda_{\boldsymbol{\alpha}} \neq 0$ pour tout $\boldsymbol{\alpha} \in \mathrm{Supp}(P)$. L'ensemble $\mathrm{Supp}(P)$ est appelé le support monomial de P. On note

$$\mathfrak{s}(P) = \mathrm{card}(\mathrm{Supp}(P))\,.$$

Le *degré total* de P, $\deg(P)$, est le degré maximal de ses monômes :

$$\deg(P) = \max\{\deg(x^{\boldsymbol{\alpha}}) : \ \boldsymbol{\alpha} \in \mathrm{Supp}(P)\}\,.$$

Le degré partiel de P par rapport à x_i, noté $\deg_{x_i}(P)$, est la plus grande i-ième composante des tuples de $\mathrm{Supp}(P)$: autrement dit il s'agit du degré de P vu comme un polynôme en x_i.

Le degré maximal d'un polynôme est le maximum des degrés maximaux de ses monômes :

$$\deg_m(P) = \max\{\deg_m(x^{\boldsymbol{\alpha}}) : \ \boldsymbol{\alpha} \in \mathrm{Supp}(P)\} = \max_{i \in [n]}(\deg_{x_i}(P)).$$

On utilisera les normes classiques suivantes sur les polynômes de $\mathbb{Q}[x_1, \ldots, x_n]$:

$$\|P\|_1 = \sum_{\boldsymbol{\alpha} \in \mathrm{Supp}(P)} |\lambda_{\boldsymbol{\alpha}}|, \ \text{ et } \ \|P\|_\infty = \max_{\boldsymbol{\alpha} \in \mathrm{Supp}(P)} |\lambda_{\boldsymbol{\alpha}}|.$$

On obtient facilement la relation suivante :

$$\|P\|_1 \leqslant \mathfrak{s}(P)\|P\|_\infty \leqslant \|P\|_\infty \prod_{i=1}^n (\deg_{x_i}(P)+1) \leqslant (\deg_m(P)+1)^n \|P\|_\infty\,, \quad (8.1)$$

et ces inégalités sont des égalités pour le polynôme $P = \sum_{\boldsymbol{\alpha} \in [d]^n} x^{\boldsymbol{\alpha}}$, car alors $\|P\|_1 = (d+1)^n$.

Lemme 8.6 (Produit).
▶ Si P et Q sont des polynômes de $\mathbb{Q}[x_1, \ldots, x_n]$, alors

$$\|PQ\|_\infty \leqslant \|P\|_\infty \|Q\|_1 \leqslant \|P\|_1 \|Q\|_1 \quad \text{et} \quad \|PQ\|_1 \leqslant \|P\|_1 \|Q\|_1\,. \qquad \blacktriangleleft$$

Démonstration. On note $P = \sum \lambda_{\boldsymbol{\alpha}} x^{\boldsymbol{\alpha}}$ et $Q = \sum \mu_{\boldsymbol{\beta}} x^{\boldsymbol{\beta}}$. Alors $PQ = \sum c_{\boldsymbol{\gamma}} x^{\boldsymbol{\gamma}}$ avec $|c_{\boldsymbol{\gamma}}| = |\sum_{\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\gamma}} \lambda_{\boldsymbol{\alpha}}\mu_{\boldsymbol{\beta}}| \leqslant \|P\|_\infty \sum_{\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\gamma}} |\mu_{\boldsymbol{\beta}}| \leqslant \|P\|_\infty \|Q\|_1$. L'autre inégalité vient du fait que $\|P\|_\infty \leqslant \|P\|_1$.

Remarquons que la première inégalité est atteinte pour $P = Q = \sum_{\boldsymbol{\alpha}\in[d]^n} x^{\boldsymbol{\alpha}}$, car on a alors $\|P^2\|_\infty = (d+1)^n = \|P\|_1$.
Enfin, $PQ = \sum_{\boldsymbol{\alpha}} \sum_{\boldsymbol{\beta}} \lambda_{\boldsymbol{\alpha}}\mu_{\boldsymbol{\beta}} x^{\boldsymbol{\alpha}+\boldsymbol{\beta}}$, d'où $\|PQ\|_1 \leqslant \sum_{\boldsymbol{\alpha}} |\lambda_{\boldsymbol{\alpha}}| \sum_{\boldsymbol{\beta}} |\mu_{\boldsymbol{\beta}}| = \|P\|_1 \|Q\|_1$. ■

Lemme 8.7 (Borne polynomiale simple).
▶ Soit A une matrice $p \times p$ dont les coefficients sont des polynômes de $\mathbb{Q}[x_1, \ldots, x_n]$. Alors

$$\|\det(A)\|_1 \leqslant \prod_{i=1}^{p} \sum_{j=1}^{p} \|A_{i,j}\|_1 .$$

◀

Démonstration. Par définition du déterminant :

La dernière égalité s'obtient par distributivité.

$$\|\det(A)\|_1 \leqslant \sum_{\sigma\in\mathfrak{S}_p} \prod_{i=1}^{p} \|A_{i,\sigma(i)}\|_1 \leqslant \sum_{\sigma:[1,p]\to[1,p]} \prod_{i=1}^{p} \|A_{i,\sigma(i)}\|_1 = \prod_{i=1}^{p} \sum_{j=1}^{p} \|A_{i,j}\|_1.$$

■

Remarque 8.8 (Borne d'Hadamard).
▶ Nous pouvons prouver une borne plus fine sur les déterminants, à partir des bornes d'Hadamard, en utilisant la norme 2 du déterminant :

$$\|\det(A)\|_2 \leqslant \prod_{i=1}^{p} \sqrt{\sum_{j=1}^{p} \|A_{i,j}\|_1^2} .$$

Cette formule se démontre en adaptant la preuve de [GG74], qui porte sur $\mathbb{C}[x]$, aux polynômes multivariés ; une borne similaire apparaît dans [EP05, Lemma 3.1] et, dans un contexte différent, dans [Bro71, Eq. (21)] (sans preuve). Cependant, la norme 1 intervient plus naturellement dans l'analyse du problème, et nous ne sommes pas arrivés à tirer pleinement parti de cette borne plus fine : nous finissions par borner la somme des carrés par le carré de la somme, ce qui revient à utiliser la borne du Lemme 8.7. ◀

Un *mineur* d'ordre m d'une matrice $m \times n$ avec $m < n$ est un sous-déterminant de la matrice, obtenu en sélectionnant m colonnes de A. La proposition suivante permet d'exprimer les solutions d'un système linéaire à l'aide de déterminants :

Proposition 8.9 (Formules de Cramer).
▶ Soit $\mathbb{K}$ un corps quelconque.

a. Soit A une matrice $n \times n$ inversible à coefficients dans $\mathbb{K}$, $\boldsymbol{b}$ et $\boldsymbol{v}$ deux vecteurs de $\mathbb{K}^n$ tels que $A\boldsymbol{v} = \boldsymbol{b}$. Alors pour tout $i \in [1, n]$, $v_i = \frac{\det(A_i)}{\det(A)}$ où A_i désigne la matrice obtenue à partir de A en remplaçant sa i-ième colonne par le vecteur $\boldsymbol{b}$.

b. Soit A une matrice $m \times n$ à coefficients dans $\mathbb{K}$ où $m < n$. Il existe un vecteur $\boldsymbol{v} \in \mathbb{K}^m \setminus \{\mathbf{0}\}$ dont les coordonnées sont des mineurs ou des opposés de mineurs de A d'ordre m, tel que $A\boldsymbol{v} = \mathbf{0}$.

◀

Démonstration. Voir [AADM17] et [Kau14]. ■

8.2.2 Séries génératrices des calculs de l'automate et du semilinéaire

Définition 8.10 (Séries génératrices associées à l'automate).

▶ On définit dans un premier temps deux séries associées à un automate de Parikh :

— La série (multivariée) des calculs de $\mathcal{A}$, notée $A(x, y_1, \dots, y_d)$, est définie par :

$$A(x, y_1, \dots, y_d) := \sum_{n\in\mathbb{N}, \boldsymbol{v}\in\mathbb{N}^d} a(n, \boldsymbol{v}) x^n y_1^{v_1} \cdots y_d^{v_d},$$

où $a(n, \boldsymbol{v})$ compte le nombre de calculs de $\mathcal{A}$ de longueur n, partant de l'état initial, finissant dans un état final, et étiquetés par le vecteur $\boldsymbol{v}$. Comme on suppose qu'aucune transition n'est étiquetée par ε, $a(n, \boldsymbol{v})$ est bien fini pour tout $n \in \mathbb{N}$ et $\boldsymbol{v} \in \mathbb{N}^d$.

— La série génératrice du semilinéaire C est définie par

$$C(y_1, \dots, y_d) = \sum_{\boldsymbol{v}\in C} y_1^{v_1} \cdots y_d^{v_d}.$$

◀

Remarque 8.11 (Notation condensée).

▶ Nous utiliserons les notations $\boldsymbol{y} := (y_1, \dots, y_d)$ et $\boldsymbol{y}^{\boldsymbol{v}} := y_1^{v_1} \cdots y_d^{v_d}$. ◀

Lemme 8.12 (Série génératrice des calculs).

▶ La série des calculs $A(x, y_1, \dots, y_d)$ est égale à une fraction $\frac{R}{S}$, où R et S sont deux polynômes de $\mathbb{Z}[x, y_1, \dots, y_d]$ vérifiant :

$$\begin{array}{cc} \deg_x(R) \leqslant |Q_{\mathcal{A}}| - 1 & \deg_x(S) \leqslant |Q_{\mathcal{A}}| \\ \forall i \in [d],\ \deg_{y_i}(R) \leqslant (|Q_{\mathcal{A}}| - 1)\|\mathcal{A}\|_{i,\infty} & \deg_{y_i}(S) \leqslant |Q_{\mathcal{A}}|\|\mathcal{A}\|_{i,\infty} \\ \|R\|_\infty \leqslant \|R\|_1 \leqslant |F| \prod_{q\in Q_{\mathcal{A}}} (1 + \deg_q^{out}) & \|S\|_\infty \leqslant \|S\|_1 \leqslant \prod_{q\in Q_{\mathcal{A}}} (1 + \deg_q^{out}) \end{array}$$

◀

Démonstration. Pour tout état $q \in Q_{\mathcal{A}}$, on considère la série

$$q(x, y_1, \dots, y_d) = \sum_{n\in\mathbb{N}, \boldsymbol{v}\in\mathbb{N}^d} a_q(n, \boldsymbol{v}) x^n y_1^{v_1} \dots y_d^{v_d},$$

où $a_q(n, \boldsymbol{v})$ décompte le nombre de calculs de longueur n depuis l'état q jusqu'à un état final, étiquetés par le vecteur $\boldsymbol{v}$. En particulier, la série génératrice $A(x, y_1, \dots, y_d)$ des calculs de $\mathcal{A}$ est égale à $q_I(x, y_1, \dots, y_d)$.

Pour tout état q, nous avons classiquement l'égalité :

$$q(x, y_1, \dots, y_d) = [\![q \in F]\!] + \sum_{(q,(a,\boldsymbol{v}),q')\in\Delta} x y_1^{v_1} \cdots y_d^{v_d} \cdot q'(x, y_1, \dots, y_d),$$

où $[\![P]\!] = 1$ si la propriété P est vraie et $[\![P]\!] = 0$ sinon.

Par conséquent, le vecteur de séries génératrices $\boldsymbol{q} = (q(x, y_1, \dots, y_d))_{q\in Q_{\mathcal{A}}}$ vérifie l'équation :

$$\boldsymbol{q} = xM\boldsymbol{q} + \boldsymbol{f},$$

où $\boldsymbol{f} = (\llbracket q \in F \rrbracket)_{q \in Q_\mathcal{A}}$, et M est une matrice $|Q_\mathcal{A}| \times |Q_\mathcal{A}|$ de polynômes, définie par $M_{q,q'}(y_1, \ldots, y_d) = \sum_{(q,(a,\boldsymbol{v}),q') \in \Delta} y_1^{\boldsymbol{v}_1} \cdots y_d^{\boldsymbol{v}_d}$ pour tout q et q' dans $Q_\mathcal{A}$. En particulier, pour tout $q \in Q_\mathcal{A}$,

$$\sum_{q' \in Q_\mathcal{A}} \|M_{q,q'}\|_1 = d_q^{out}.$$

Le vecteur $\boldsymbol{q}$ satisfait l'équation $(\mathtt{Id} - xM)\boldsymbol{q} = \boldsymbol{f}$. La matrice $\mathtt{Id} - xM$ est inversible dans le corps des fractions rationnelles $\mathbb{Q}(x, y_1, \ldots, y_d)$. En effet, $\det(\mathtt{Id} - M)(x, y_1, \ldots, y_d)$ est un polynôme non nul car $\det(\mathtt{Id} - xM)(0, \ldots, 0) = \det(I) = 1$. Par les formules de Cramer (Proposition 8.9) :

$$A(x, y_1, \ldots, y_d) = q_I(x, y_1, \ldots, y_d) = \frac{R(x, y_1, \ldots, y_d)}{S(x, y_1, \ldots, y_d)},$$

avec $R(x, y_1, \ldots, y_d) = \det(M')$, où M' est la matrice obtenue en remplaçant la colonne numéro q_I de $\mathtt{Id} - xM$ par le vecteur $\boldsymbol{f}$, et $S(x, y_1, \ldots, y_d) = \det(\mathtt{Id} - xM)$.

Remarquons alors que pour $q \in Q_\mathcal{A}$:

$$\sum_{q' \in Q_\mathcal{A}} \|(\mathtt{Id} - xM)_{q,q'}\|_1 = 1 + \deg_q^{out}$$

avec le 1 qui vient de l'identité. Ainsi en utilisant la borne du Lemme 8.7 à la matrice $\mathtt{Id} - xM$, on obtient

$$\|S\|_\infty \leqslant \|S\|_1 \leqslant \prod_{q \in Q_\mathcal{A}} (1 + \deg_q^{out}).$$

Par définition du déterminant, chaque terme de S est une combinaison linéaire de produits de $|Q_\mathcal{A}|$ polynômes, de degré maximal 1 en x, et de degré maximal $\|A\|_{i,\infty}$ en y_i, donc $\deg_x(S) \leqslant |Q_\mathcal{A}|$ et $\deg_{y_i}(S) \leqslant |Q_\mathcal{A}| \|\mathcal{A}\|_{i,\infty}$.

En développant le déterminant de M' selon la première colonne $\boldsymbol{f}$, et en appliquant la borne du Lemme 8.7 aux sous-déterminants, nous obtenons la formule suivante :

$$\|R\|_1 \leqslant \sum_{q_f \in F} \prod_{q \neq q_f} \sum_{q' \neq q_I} \|(\mathtt{Id} - xM)_{q,q'}\|_1 .$$

Trivialement, $\sum_{q' \neq q_I} \|(\mathtt{Id} - xM)_{q,q'}\|_1 \leqslant \sum_{q' \in Q_\mathcal{A}} \|(\mathtt{Id} - xM)_{q,q'}\|_1 = 1 + \deg_q^{out}$.

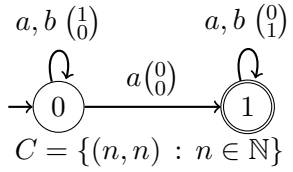
Et $\prod_{q \neq q_f} (1 + \deg_q^{out}) \leqslant \prod_{q \in Q_\mathcal{A}} (1 + \deg_q^{out})$, car $(1 + \deg_{q_f}^{out}) \geqslant 1$. Ainsi :

$$\|R\|_\infty \leqslant \|R\|_1 \leqslant |F| \prod_{q \in Q_\mathcal{A}} (1 + \deg_q^{out}).$$

■

Exemple 8.13 .

▶ Tout au long de cette section et de la suivante, nous illustrons les résultats des différentes propositions sur l'automate $\mathcal{A}_{\text{ex}}$ de l'introduction. Nous rappelons que $\mathcal{A}_{\text{ex}}$ est l'automate suivant, qui reconnaît les mots de longueur impaire qui ont un a au milieu :



Réécrivons le semilinéaire sous la forme voulue $C = \mathbf{0} + \{(1,1)\}^*$. Dans ces conditions :

- $d = 2$
- $|\mathcal{A}_{\text{ex}}| = 2 + 3 + 1 + 1 = 7$;
- $\|\mathcal{A}_{\text{ex}}\|_{1,\infty} = \|\mathcal{A}_{\text{ex}}\|_{2,\infty} = \|\mathcal{A}_{\text{ex}}\|_{\infty} = 1$;
- $\deg_{q_0}^{out} = 3$ et $\deg_{q_1}^{out} = 2$. ◀

Nous obtenons alors le système suivant pour les séries génératrices des calculs :

$$\begin{cases} q_0(x, y_1, y_2) = 2xy_1 q_0(x, y_1, y_2) + xq_1(x, y_1, y_2) \\ q_1(x, y_1, y_2) = 2xy_2 q_1(x, y_1, y_2) + 1 \end{cases}$$

qui s'écrit sous la forme matricielle :

$$\begin{pmatrix} 1 - 2xy_1 & -x \\ 0 & 1 - 2xy_2 \end{pmatrix} \cdot \begin{pmatrix} q_0(x, y_1, y_2) \\ q_1(x, y_1, y_2) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

et avec les formules de Cramer nous avons $A(x, y_1, y_2) = \frac{R}{S}$ avec

$$R = \begin{vmatrix} 0 & -x \\ 1 & 1 - 2xy_2 \end{vmatrix} = x \quad \text{et} \quad S = \begin{vmatrix} 1 - 2xy_1 & -x \\ 0 & 1 - 2xy_2 \end{vmatrix} = (1-2xy_1)(1-2xy_2)\,.$$

Et nous vérifons bien les inégalités :

$$\begin{aligned} \deg_x(R) = 1 &\leqslant |Q_{\mathcal{A}_{\text{ex}}}| - 1 = 1 & \deg_x(S) = 2 &\leqslant |Q_{\mathcal{A}_{\text{ex}}}| = 2 \\ \deg_{y_1}(R) = 0 &\leqslant (|Q_{\mathcal{A}_{\text{ex}}}| - 1)\|\mathcal{A}_{\text{ex}}\|_{1,\infty} = 1 & \deg_{y_1}(S) = 1 &\leqslant |Q_{\mathcal{A}_{\text{ex}}}|\|\mathcal{A}_{\text{ex}}\|_{1,\infty} = 2 \\ \deg_{y_2}(R) = 0 &\leqslant (|Q_{\mathcal{A}_{\text{ex}}}| - 1)\|\mathcal{A}_{\text{ex}}\|_{2,\infty} = 1 & \deg_{y_2}(S) = 1 &\leqslant |Q_{\mathcal{A}_{\text{ex}}}|\|\mathcal{A}_{\text{ex}}\|_{2,\infty} = 2 \\ \|R\|_1 = 1 &\leqslant |F| \prod_{q \in Q_{\mathcal{A}_{\text{ex}}}} (1 + \deg_q^{out}) = 12 & \|S\|_1 = 9 &\leqslant \prod_{q \in Q_{\mathcal{A}_{\text{ex}}}} (1 + \deg_q^{out}) = 12 \end{aligned}$$

Remarque 8.14 .

▶ En utilisant la définition du déterminant $\det(\mathtt{Id} - xM)$ comme une somme portant sur toutes les permutations s de $Q_{\mathcal{A}}$, nous pouvons interpréter la décomposition de s en produits de cycles disjoints en termes de cycles dans l'automate, pour borner plus finement le degré de R et de S. Soit $i \in [d]$, pour tout cycle σ élémentaire de $\mathcal{A}$, notons, dans cette remarque uniquement, $|\sigma|_i$ la somme des i-ièmes composantes des vecteurs des transitions empruntées dans le cycle σ.

En notant Ω l'ensemble des cycles élémentaires de $\mathcal{A}$ et $\mathcal{P}_{\neq}(\Omega)$ l'ensemble des parties de Ω composées de cycles qui ne passent par aucun état en commun, alors $\deg_{y_i}(S) \leqslant \max_{X \in \mathcal{P}_{\neq}(\Omega)} \sum_{\sigma \in X} |\sigma|_i$.

Cette interprétation du déterminant avec des cycles dans un graphe est un cas particulier de la formule plus générale de Lindström-Gessel-Viennot [GV89]. Nous n'aurons pas besoin d'être aussi précis cependant. Cette borne implique celle que

nous avons trouvée plus haut : pour tout cycle élémentaire σ, $|\sigma|_i \leqslant \|\mathcal{A}\|_{i,\infty}|\sigma|$ où $|\sigma|$ est le nombre d'états par lequel passe le cycle élémentaire. Comme les cycles sont élémentaires et ne passent par aucun état en commun, $\sum_{\sigma\in X}|\sigma|_i \leqslant |Q_{\mathcal{A}}|\|\mathcal{A}\|_{i,\infty}$. ◀

Lemme 8.15 (Série génératrice du semilinéaire).
▶ On rappelle que le semilinéaire est donné sous une forme non ambiguë $C = \biguplus_{i=1}^{p} \boldsymbol{c_i} + P_i^*$. Alors la série génératrice de C s'écrit sous la forme d'une fraction $C(\boldsymbol{y}) = \frac{P_C}{Q_C}$ avec :

$$\forall i \in [d],\ \deg_{y_i}(P_C) \leqslant (1+\sum_{j=1}^{p}|P_j|)\|\mathcal{A}\|_{i,\infty} \qquad \deg_{y_i}(Q_C) \leqslant (\sum_{j=1}^{p}|P_j|)\|\mathcal{A}\|_{i,\infty}$$
$$\|P_C\|_1 \leqslant p2^{\sum_{i=1}^{p}|P_i|} \qquad \|Q_C\|_1 \leqslant 2^{\sum_{i=1}^{p}|P_i|}$$

◀

Démonstration. Par non-ambiguïté, $C(\boldsymbol{y}) = \sum_{i=1}^{p} \frac{\boldsymbol{y}^{\boldsymbol{c_i}}}{\prod_{\boldsymbol{v}\in P_i}(1-\boldsymbol{y}^{\boldsymbol{v}})}$. En mettant au même dénominateur, on obtient $C(\boldsymbol{y}) = \frac{P_C}{Q_C}$ avec :

$$P_C = \sum_{i=1}^{p} \boldsymbol{y}^{\boldsymbol{c_i}} \prod_{j\neq i}\prod_{\boldsymbol{v}\in P_j}(1-\boldsymbol{y}^{\boldsymbol{v}}) \quad \text{et} \quad Q_C = \prod_{i=1}^{p}\prod_{\boldsymbol{v}\in P_i}(1-\boldsymbol{y}^{\boldsymbol{v}}).$$

Ainsi $\deg_{y_i}(P_C) \leqslant (1+\sum_{j=1}^{p}|P_j|)\|\mathcal{A}\|_{i,\infty}$, et $\deg_{y_i}(Q_C) \leqslant (\sum_{j=1}^{p}|P_j|)\|\mathcal{A}\|_{i,\infty}$ pour tout $i \in [d]$.

De plus par non-ambiguïté, $\|\prod_{\boldsymbol{v}\in P_i}(1-\boldsymbol{y}^{\boldsymbol{v}})\|_\infty = 1$, et $\|\prod_{\boldsymbol{v}\in P_i}(1-\boldsymbol{y}^{\boldsymbol{v}})\|_1 = \mathfrak{s}\big(\prod_{\boldsymbol{v}\in P_i}(1-\boldsymbol{y}^{\boldsymbol{v}})\big) = 2^{|P_i|}$.

Ainsi $\|Q_C\|_1 \leqslant 2^{\sum_{i=1}^{p}|P_i|}$, et $\|P_C\|_1 \leqslant p2^{\sum_{i=1}^{p}|P_i|}$. ■

8.3 Série génératrice d'un automate de Parikh faiblement non ambigu

8.3.1 Préparation de la série génératrice à l'algorithme de Lipshitz

Un des objectifs de ce chapitre est de fournir des bornes sur les degrés et les normes des coefficients des polynômes qui apparaissent dans l'équation différentielle satisfaite par la série génératrice du langage de $\mathcal{A}$. Pour cela nous décrivons rapidement dans un premier temps l'approche de Lipshitz [Lip88], qui consiste à introduire la fonction :

$$F(x,\boldsymbol{y},\boldsymbol{u}) = \frac{1}{u_1\cdots u_d}A\left(x,u_1,\ldots,u_d\right)\cdot C\left(\frac{y_1}{u_1},\ldots,\frac{y_d}{u_d}\right).$$

Si cette opération s'interprète bien du côté des fractions rationnelles, il faut vérifier qu'elle ait bien un sens dans celui des séries, car nous quittons l'anneau des séries formelles $\mathbb{Q}[x,\boldsymbol{y},\boldsymbol{u}]$: la substitution $y_1 \leftarrow \frac{y_1}{u_1}$ dans la série de C introduit une infinité de monômes à exposants négatifs. Lipshitz se place en fait dans l'ensemble $\mathcal{M}$ des

sommes infinies formelles de la forme :

$$G = \sum_{\substack{n\in\mathbb{N},\boldsymbol{i}\in\mathbb{N}^d,\boldsymbol{j}\in\mathbb{Z}^d \\ i_1+j_1\geqslant -k \\ \cdots \\ i_d+j_d\geqslant -k}} a(n,\boldsymbol{i},\boldsymbol{j}) x^n \boldsymbol{y}^{\boldsymbol{i}} \boldsymbol{u}^{\boldsymbol{j}}$$

où k est une constante qui dépend de G. L'algorithme de Lipshitz repose sur les propriétés suivantes de l'ensemble $\mathcal{M}$:

- $\mathcal{M}$ est un $\mathbb{Q}[x,\boldsymbol{y},\boldsymbol{u}]$-module (au sens où pour tous $G, G' \in \mathcal{M}$, et $p \in \mathbb{Q}[x,\boldsymbol{y},\boldsymbol{u}]$, $G+G' \in \mathcal{M}$ et $pG \in \mathcal{M}$)
- $\mathcal{M}$ est stable par dérivation (si $t \in \{x, y_1, \dots, y_d, u_1, \dots, u_d\}$ et $G \in \mathcal{M}$, alors $\partial_t G \in M$, où $\partial_t G$ désigne la dérivation formelle terme à terme de G).
- enfin, si $p \in \mathbb{Q}[x,\boldsymbol{y},\boldsymbol{u}]$ et $G \in \mathcal{M}$ sont tels que $p \neq 0$ et $pG = 0$, alors $G = 0$.

L'étude de $F(x,\boldsymbol{y},\boldsymbol{u})$ est alors justifiée par l'égalité suivante :

$$A(x,y_1,\dots,y_d) \odot \frac{1}{1-x} C(y_1,\dots,y_d) = [u_1^{-1}\cdots u_d^{-1}]\, F(x,\boldsymbol{y},\boldsymbol{u})\,,$$

l'extraction de coefficients ayant bien un sens par la définition de $\mathcal{M}$. Multiplier par $\frac{1}{y_1\dots y_d}$ dans la définition de F, pour ensuite extraire $[(y_1\dots y_d)^{-1}]F$, peut sembler inutile au premier abord. En fait, l'argument de Lipshitz repose fortement sur le fait que le coefficient extrait est bien $[(u_1\dots u_d)^{-1}]F$: comme $(u_1\dots u_d)^{-1}$ ne peut pas s'obtenir à partir de la dérivation formelle d'un terme de la forme $u_1^{\pm i_1}\dots u_d^{\pm i_d}$, cela permettra d'intervertir dérivation et extraction de coefficient dans une équation différentielle de F bien choisie, pour la transformer en une équation pour $[(u_1\dots u_d)^{-1}]F$.

L'algorithme de Lipshitz permet ainsi de trouver une équation différentielle partielle en x satisfaite par $A(x,y_1,\dots,y_d) \odot \frac{1}{1-x} C(y_1,\dots,y_d)$. Nous en déduisons ensuite en spécialisant $y_i = 1$ une équation différentielle pour $L(x)$.

C'est ce que nous avons fait dans les annexes de notre article publié à ICALP en 2020. Cette approche a le gros inconvénient d'introduire des variables auxiliaires $u_1,\dots,u_d$, et d'ainsi doubler le nombre de variables par rapport à la dimension de l'automate de départ. Cette augmentation de la dimension vouait à l'échec toute tentative d'appliquer l'algorithme de Lipshitz sur des exemples, même élémentaires. Cette méthode implique de calculer l'équation du produit d'Hadamard de deux séries en $x, y_1, \dots, y_d$, pour ensuite remplacer toutes les variables y_i par 1. Nous calculons donc un objet plus complexe, la série multivariée, pour ensuite "jeter" les variables.

Il nous faudra donc faire les mêmes vérifications que Lipshitz, en définissant un module analogue à $\mathcal{M}$ adapté aux séries que nous rencontrerons.

Dans ce chapitre, j'opte donc pour une méthode légèrement différente de celle de notre article, en remplaçant toutes les variables y_i par 1 *avant* d'appliquer l'algorithme de Lipshitz. Nous nous sommes donc intéressés à la somme infinie suivante :

$$\frac{1}{y_1\cdots y_d} \sum_{n\in\mathbb{N},\boldsymbol{v}\in\mathbb{N}^d,\boldsymbol{v'}\in C} a(n,\boldsymbol{v}) x^n \boldsymbol{y}^{\boldsymbol{v}-\boldsymbol{v'}}$$

$$\text{" = "} \frac{1}{y_1\cdots y_d} A\left(x,y_1,\dots,y_d\right)\cdot C\left(\frac{1}{y_1},\dots,\frac{1}{y_d}\right)$$

Il faut cependant faire attention, car cette somme infinie n'est pas automatiquement bien définie. Nous le voyons sur un exemple simple : si $A(x,y) = x\sum_{n\in\mathbb{N}} y^n$, et $C(y) = \sum_{n\in\mathbb{N}} y^n$, alors $A(x,y)C(1/y) = "x\sum_{n,m\in\mathbb{N}} y^{n-m}"$ n'est pas bien définie

en tant que somme infinie, même si la fraction rationnelle $\frac{x}{1-y}\frac{1}{1-1/y}$ existe. Le problème s'observe aussi en se plongeant dans les séries entières : si les rayons de convergence de $A(x,y)$ et $C(x,y)$ sont 1 par exemple, la série de $A(x,y)$ converge pour $|y|<1$, mais celle de $C(x,1/y)$ converge pour $|y|>1$.

Cependant, dans notre cas particulier, nous allons voir que la somme est bien définie, car $A(x,\boldsymbol{y})$ a une forme particulière. Dans un premier temps nous introduisons un ensemble $\mathcal{M}'$ analogue à celui de Lipshitz, adapté à notre problème :

Proposition 8.16 .
▶ L'ensemble $\mathcal{M}'$ des sommes formelles de la forme

$$G = \sum_{\substack{n\in\mathbb{N},\boldsymbol{v}\in\mathbb{Z}^d \\ v_1\leqslant nk \\ \cdots \\ v_d\leqslant nk}} a(n,\boldsymbol{v})x^n\boldsymbol{y}^{\boldsymbol{v}},$$

où k est une constante qui dépend de G, vérifie toutes les propriétés nécessaires à l'algorithme de Lipshitz :

— $\mathcal{M}'$ est un $\mathbb{Q}[x,\boldsymbol{y}]$-module

— $\mathcal{M}'$ est stable par dérivation

— enfin, si $p\in\mathbb{Q}[x,\boldsymbol{y}]$ est non nul, $G\in\mathcal{M}'$, et de plus $pG=0$, alors $G=0$.

◀

Démonstration. Seul le dernier point n'est pas trivial. Il se démontre par une adaptation de la preuve pour le module $\mathcal{M}$ de Lipshitz [Lip88] : si $p(x,\boldsymbol{y})G(x,\boldsymbol{y})=0$ avec $p(x,\boldsymbol{y})\neq 0$ et $G(x,\boldsymbol{y})\in\mathcal{M}'$, notons k la constante associée à G, et $r := \max(k,\deg_m(p))$. Alors $G(xy_1^r\dots y_d^r, y_1^{-1},\dots,y_d^{-1})\in\mathbb{Q}[[x,y_1,\dots,y_d]]$ est une série formelle, et $p(xy_1^r\dots y_d^r, y_1^{-1},\dots,y_d^{-1})$ est toujours un polynôme non nul. Ainsi $p(xy_1^r\dots y_d^r, y_1^{-1},\dots,y_d^{-1})G(xy_1^r\dots y_d^r, y_1^{-1},\dots,y_d^{-1})=0$, et comme l'anneau des séries formelles est intègre, $G(xy_1^r\dots y_d^r, y_1^{-1},\dots,y_d^{-1})=0$. En remarquant que la substitution qui remplace $(x,\boldsymbol{y})$ par $(xy_1^r\dots y_d^r, y_1^{-1},\dots,y_d^{-1})$ ne crée pas de collision, au sens où les deux séries ont exactement le même ensemble de coefficients, nous en déduisons que $G=0$. ■

La proposition suivante démontre que F est bien définie :

Proposition 8.17 (Définition de F).
▶ La somme infinie suivante :

$$F(x,\boldsymbol{y}) := \frac{1}{y_1\cdots y_d}\sum_{n\in\mathbb{N},\boldsymbol{v}\in\mathbb{N}^d,\boldsymbol{v'}\in C} a(n,\boldsymbol{v})x^n\boldsymbol{y}^{\boldsymbol{v}-\boldsymbol{v'}}$$

est bien définie et appartient à $\mathcal{M}'$. ◀

Démonstration. Rappelons que l'automate $\mathcal{A}$ ne possède pas d'ε-transitions, et qu'ainsi la longueur d'un calcul de $\mathcal{A}$ coïncide avec la longueur du mot qui l'étiquette. Comme tout vecteur $\boldsymbol{v}$ qui étiquette un calcul de longueur n vérifie $\|\boldsymbol{v}\|_\infty\leqslant n\|\mathcal{A}\|_\infty$, pour tout $n\in\mathbb{N}$, $[x^n]A(x,\boldsymbol{y}) = \sum_{\|\boldsymbol{v}\|_\infty\leqslant n\|\mathcal{A}\|_\infty} a(n,\boldsymbol{v})\boldsymbol{y}^{\boldsymbol{v}}$ est donc un polynôme, dont le degré maximal est borné par $n\|\mathcal{A}\|_\infty$.

Ainsi la somme infinie $\sum_{n\in\mathbb{N},\boldsymbol{v}\in\mathbb{N}^d,\boldsymbol{v'}\in C} a(n,\boldsymbol{v})x^n\boldsymbol{y}^{\boldsymbol{v}-\boldsymbol{v'}}$ est bien définie, et F appartient à l'ensemble $\mathcal{M}'$. ■

Remarque 8.18 .
▶ Avec un peu de recul, nous aurions sans doute dû tourner les choses un tout petit peu différemment. Si je suis parti de l'expression $A(x, \boldsymbol{y})C(\boldsymbol{y}^{-\mathbf{1}})$, c'est parce qu'elle s'interprète bien en imaginant un automate qui calcule d'abord un vecteur positif en incrémentant des compteurs, puis les décrémente avec une machine qui calcule le semilinéaire. D'un point de vue séries formelles cependant, il est plus naturel d'inverser l'endroit où nous inversons les exposants, en considérant plutôt :

$$F(x, \boldsymbol{y}) = \frac{1}{y_1 \cdots y_d} A\left(x, \frac{1}{y_1}, \ldots, \frac{1}{y_d}\right) \cdot C\left(y_1, \ldots, y_d\right).$$

Avec cette expression, pour tout $n \in \mathbb{N}$, $[x^n]F(x, \boldsymbol{y})$ n'a qu'un nombre fini de termes qui ont un exposant négatif, autrement dit $[x^n]F(x, \boldsymbol{y})$ est une série de Laurent formelle. Les deux approches sont analogues, mais celle de cette remarque a l'avantage d'introduire des classes de séries formelles plus standard.

Si les propositions qui suivent sont en fait valides pour les deux définitions de F, je n'ai pas changé la définition de F faute de temps, car il aurait fallu refaire tous les exemples ; ce n'était pas prioritaire dans le processus de rédaction de la thèse. ◀

Nous vérifions alors facilement que formellement, en notant m un entier plus grand que $\deg_m(P_C)$ et $\deg_m(Q_C)$, nous avons l'égalité :

$$S(x, \boldsymbol{y})(y_1 \ldots y_d)^{m+1} Q_C(\tfrac{1}{y_1}, \ldots, \tfrac{1}{y_d}) \cdot F(x, \boldsymbol{y}) = R(x, \boldsymbol{y})(y_1 \ldots y_d)^m P_C(\tfrac{1}{y_1}, \ldots, \tfrac{1}{y_d}).$$

Autrement dit :

$$F(x, \boldsymbol{y}) = \frac{1}{y_1 \cdots y_d} A\left(x, y_1, \ldots, y_d\right) \cdot C\left(\frac{1}{y_1}, \ldots, \frac{1}{y_d}\right)$$

où le membre droit est interprété comme une fraction rationnelle P/Q avec $P, Q \in \mathbb{Z}[x, \boldsymbol{y}]$. Enfin, une extraction de coefficient de $F(x, \boldsymbol{y})$ fait apparaître $L(x)$:

$$\begin{aligned}[y_1^{-1} \cdots y_d^{-1}]F(x, \boldsymbol{y}) &= [y_1^{-1} \cdots y_d^{-1}] \frac{1}{y_1 \cdots y_d} \sum_{n \in \mathbb{N}, \boldsymbol{v} \in \mathbb{N}^d, \boldsymbol{v'} \in C} a(n, \boldsymbol{v}) x^n \boldsymbol{y}^{\boldsymbol{v} - \boldsymbol{v'}} \\ &= \sum_{n \in \mathbb{N}, \boldsymbol{v} \in C} a(n, \boldsymbol{v}) x^n = L(x)\end{aligned} \tag{8.2}$$

où nous rappelons que la série génératrice des calculs acceptants de $\mathcal{A}$ vaut $L(x)$ par faible non ambiguïté.

Tout le terrain théorique préparatoire à la mise en place de l'algorithme de Lipshitz est donc prêt, et nous pouvons désormais dériver les premières propriétés de la fraction rationnelle $F(x, \boldsymbol{y})$:

Lemme 8.19 (Propriétés de la fraction rationnelle).
▶ La fraction rationnelle $F(x, \boldsymbol{y})$ peut s'écrire sous la forme $\frac{P}{Q}$, avec P et Q des polynômes de $\mathbb{Z}[x, \boldsymbol{y}]$ qui vérifient :

$$\begin{array}{cc} \deg_x(P) \leqslant |Q_{\mathcal{A}}| - 1 & \deg_x(Q) \leqslant |Q_{\mathcal{A}}| \\ \forall i \in [d],\ \deg_{y_i}(P) < |\mathcal{A}|\|\mathcal{A}\|_{i,\infty} & \deg_{y_i}(Q) < |\mathcal{A}|\|\mathcal{A}\|_{i,\infty} \\ \|P\|_1 \leqslant 2^{|\mathcal{A}|} & \|Q\|_1 \leqslant 2^{|\mathcal{A}|} \end{array}$$

◀

Démonstration. Nous utilisons les bornes des Lemmes 8.12 et 8.15.

Posons $m_i = \max(1, \sum_{j=1}^{p} |P_j|)\|\mathcal{A}\|_{i,\infty}$ pour tout $i \in [d]$. Alors $\frac{1}{y_1 \dots y_d} \frac{\boldsymbol{y}^{\boldsymbol{m}} P_C(\boldsymbol{y}^{-\mathbf{1}})}{\boldsymbol{y}^{\boldsymbol{m}} Q_C(\boldsymbol{y}^{-\mathbf{1}})}$ s'écrit sous la forme $\frac{\tilde{P}_C}{\tilde{Q}_C}$, avec $\tilde{P}_C$ et $\tilde{Q}_C$ deux polynômes de $\mathbb{Z}[x, \boldsymbol{y}]$ qui vérifient $\|\tilde{P}_C\|_1 = \|P_C\|_1$ et $\|\tilde{Q}_C\|_1 = \|Q_C\|_1$. De plus $\deg_{y_i}(\tilde{P}_C) \leqslant m_i$ et $\deg_{y_i}(\tilde{Q}_C) \leqslant m_i + 1$, pour tout $i \in [d]$.

Ainsi $P = R\tilde{P}_C$ et $Q = S\tilde{Q}_C$, si bien que :

$$\begin{aligned}
\log(\|P\|_1) &\leqslant \log(\|R\|_1 \|\tilde{P}_C\|_1) \\
&\leqslant \log(|F|) + \sum_{q \in Q_\mathcal{A}} \log(1 + \deg_q^{out}) + \log(p) + \sum_{i=1}^{p} |P_i| \\
&\leqslant \log(|F|) + \sum_{q \in Q_\mathcal{A}} \deg_q^{out} + \log(p) + \sum_{i=1}^{p} |P_i| \\
&\leqslant \log(|F|) + |\Delta| + \log(p) + \sum_{i=1}^{p} |P_i| \leqslant |\mathcal{A}|
\end{aligned}$$

Nous obtenons la même majoration pour Q : $\|Q\|_1 \leqslant 2^{|\mathcal{A}|}$.

En ce qui concerne la variable x, $\deg_x(P) = \deg_x(R)$ et $\deg_x(Q) = \deg_x(S)$.

Enfin, $\deg_{y_i}(P) \leqslant \deg_{y_i}(R) + m_i \leqslant (|Q_\mathcal{A}| - 1 + 1 + \sum_{j=1}^{p} |P_j|)\|\mathcal{A}\|_{i,\infty} < |\mathcal{A}|\|\mathcal{A}\|_{i,\infty}$.

Nous supposons que $\|\mathcal{A}\|_{i,\infty} \neq 0$, sinon cette coordonnée ne sert à rien.

De même, $\deg_{y_i}(Q) \leqslant \deg_{y_i}(S) + m_i + 1 \leqslant (|Q_\mathcal{A}| + 1 + \sum_{j=1}^{p} |P_j|)\|\mathcal{A}\|_{i,\infty} + 1 < |\mathcal{A}|\|\mathcal{A}\|_{i,\infty}$. ■

Exemple 8.20.

▶ Si nous reprenons l'exemple de l'automate $\mathcal{A}_{\text{ex}}$, nous avons

$$A(x, y_1, y_2) = \frac{x}{(1 - 2xy_1)(1 - 2xy_2)} \quad \text{et} \quad C(y_1, y_2) = \frac{1}{1 - y_1 y_2},$$

si bien que :

$$\begin{aligned}
F(x, y_1, y_2) &= \frac{1}{y_1 y_2} \frac{x}{(1 - 2xy_1)(1 - 2xy_2)} \frac{1}{1 - y_1^{-1} y_2^{-1}} \\
&= \frac{x}{(1 - 2xy_1)(1 - 2xy_2)(y_1 y_2 - 1)}
\end{aligned}$$

et ainsi $L(x) = [y_1^{-1} y_2^{-1}] F(x, y_1, y_2)$.

On vérifie bien que les bornes annoncées sont vérifiées sur cet exemple, même si sur cet exemple nous surestimons un peu, avec notamment $1 = \|P\|_1 \leqslant 2^{|\mathcal{A}_{\text{ex}}|} = 128$). ◀

Remarque 8.21.

▶ Nous n'utilisons plus l'astuce, propre à cet exemple particulier, utilisée dans l'introduction pour réduire le nombre de variables. En effet le but de cet exemple est désormais d'illustrer l'approche générale, adaptée à tout type de semilinéaire. ◀

8.3.2 Équation différentielle satisfaite par F

Dans cette section nous analysons le procédé utilisé par Lipshitz pour dériver l'équation différentielle satisfaite par F.

Nous nous intéressons à la fraction $F = \frac{P}{Q}$ en les variables $x, y_1, \ldots, y_d$.

Pour simplifier les notations dans un premier temps, supposons qu'il existe des constantes M_i telles que $\deg_{y_i}(P), \deg_{y_i}(Q) \leqslant M_i - 1$ pour tout $i \in [d]$, et M_0 telle que $\deg_x(P), \deg_x(Q) \leqslant M_0 - 1$. Par le Lemme 8.19, $M_0 = |Q_{\mathcal{A}}| + 1 \leqslant |\mathcal{A}|$, et $M_i = |\mathcal{A}|\|\mathcal{A}\|_{i,\infty}$.

Nous noterons $y_0 = x$ pour simplifier le traitement des différentes variables, lorsque x ne joue pas de rôle particulier par rapport aux variables $\boldsymbol{y}$.

Lemme 8.22.

▶ Soit $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_d) \in \mathbb{N}^{d+1}$. On note $\partial_{\boldsymbol{\alpha}}$ l'opérateur défini par $\partial_{\boldsymbol{\alpha}} = \partial_{y_1}^{\alpha_1} \ldots \partial_{y_d}^{\alpha_d} \partial_x^{\alpha_0}$. Alors il existe un polynôme $P_{\boldsymbol{\alpha}}$ de $\mathbb{Z}[x, \boldsymbol{y}]$ tel que $Q^{s+1}\partial_{\boldsymbol{\alpha}} F = P_{\boldsymbol{\alpha}}$, avec $s = \|\boldsymbol{\alpha}\|_1 = \sum_i \alpha_i$ et :

$$\deg_{y_i}(P_{\boldsymbol{\alpha}}) \leqslant (s+1)(M_i - 1)$$
$$\|P_{\boldsymbol{\alpha}}\|_1 \leqslant 2^s s! \|Q\|_1^s \|P\|_1 \prod_{i=0}^{d} (M_i - 1)^{\alpha_i}$$ ◀

Démonstration. Par récurrence sur $s = \|\boldsymbol{\alpha}\|_1$. Si $s = 0$, alors $QF = P$ et la proposition est vraie.

Soit $s > 0$ tel que la proposition soit vraie au rang $s - 1$. Soit $\boldsymbol{\alpha}$ tel que $\|\boldsymbol{\alpha}\|_1 = s$. Écrivons $\boldsymbol{\alpha} = \boldsymbol{\beta} + \boldsymbol{e_i}$, pour une coordonnée i telle que $\alpha_i > 0$; ainsi $\|\boldsymbol{\beta}\|_1 = s - 1$.

Par hypothèse de récurrence $Q^s \partial_{\boldsymbol{\beta}} F = P_{\boldsymbol{\beta}}$, où $P_{\boldsymbol{\beta}}$ satisfait les conclusions de la proposition au rang $s - 1$. Dérivons cette relation par rapport à y_i (en utilisant le temps de cette preuve la notation $y_0 = x$) :

$$(\partial_{y_i} Q) s Q^{s-1} \partial_{\boldsymbol{\beta}} F + Q^s \partial_{\boldsymbol{\alpha}} F = \partial_{y_i} P_{\boldsymbol{\beta}}.$$

Multiplions par Q :

$$Q^{s+1} \partial_{\boldsymbol{\alpha}} F = Q \partial_{y_i} P_{\boldsymbol{\beta}} - (\partial_{y_i} Q) s Q^s \partial_{\boldsymbol{\beta}} F = Q \partial_{y_i} P_{\boldsymbol{\beta}} - (\partial_{y_i} Q) s P_{\boldsymbol{\beta}}.$$

Posons $P_{\boldsymbol{\alpha}} = Q \partial_{y_i} P_{\boldsymbol{\beta}} - (\partial_{y_i} Q) s P_{\boldsymbol{\beta}}$. Le polynôme $P_{\boldsymbol{\alpha}}$ est bien à coefficients entiers.

De plus, pour $j \neq i$, $\deg_{y_j}(P_{\boldsymbol{\alpha}}) \leqslant \deg_{y_j}(Q) + \deg_{y_j}(P_{\boldsymbol{\beta}}) \leqslant_{H.R.} \deg_{y_j}(Q) + s(M_j - 1) \leqslant (s+1)(M_j - 1)$.

Et de même $\deg_{y_i}(P_{\boldsymbol{\alpha}}) \leqslant M_i - 1 + s(M_i - 1) - 1 \leqslant (s+1)(M_i - 1)$.

De plus :

$$\begin{aligned}
\|P_{\boldsymbol{\alpha}}\|_1 &\leqslant \|Q\|_1 \|\partial_{y_i} P_{\boldsymbol{\beta}}\|_1 + s \|\partial_{y_i} Q\|_1 \|P_{\boldsymbol{\beta}}\|_1 \\
&\leqslant \|Q\|_1 \deg_{y_i}(P_{\boldsymbol{\beta}}) \|P_{\boldsymbol{\beta}}\|_1 + s \deg_{y_i}(Q) \|Q\|_1 \|P_{\boldsymbol{\beta}}\|_1 \\
&\leqslant_{H.R.} 2s(M_i - 1) \|Q\|_1 \|P_{\boldsymbol{\beta}}\|_1 \\
&\leqslant_{H.R.} 2^s s! \|Q\|_1^s \|P\|_1 \prod_{i=0}^{d} (M_i - 1)^{\alpha_i}
\end{aligned}$$

■

Lemme 8.23.
▶ Soit $N > 0$ un entier. Soit $\boldsymbol{\alpha} \in \mathbb{N}^{d+1}$ tel que $\sum_i \alpha_i < N$. Il existe un polynôme $R_{\boldsymbol{\alpha}}$ en les variables $x, y_1, \ldots, y_d$, à coefficients entiers, tel que

$$Q^N \partial_{y_1}^{\alpha_1} \ldots \partial_{y_d}^{\alpha_d} \partial_x^{\alpha_0} F = R_{\boldsymbol{\alpha}}.$$

De plus

$$\deg_{y_i}(R_{\boldsymbol{\alpha}}) \leqslant N(M_i - 1)$$
$$\|R_{\boldsymbol{\alpha}}\|_1 \leqslant 2^s s! \|Q\|_1^{N-1} \|P\|_1 \prod_{i=0}^{d} (M_i - 1)^{\alpha_i}$$ ◀

Démonstration. Par le Lemme 8.22, nous savons que $Q^{s+1} \partial_{\boldsymbol{\alpha}} F = P_{\boldsymbol{\alpha}}$, avec $s = \|\boldsymbol{\alpha}\|_1$. On a donc $R_{\boldsymbol{\alpha}} = Q^{N-s-1} P_{\boldsymbol{\alpha}}$.
Ainsi $\deg_{y_i} R_{\boldsymbol{\alpha}} \leqslant (N - s - 1) \deg_{y_i}(Q) + \deg_{y_i}(P_{\boldsymbol{\alpha}}) \leqslant N(M_i - 1)$.
Et $\|R_{\boldsymbol{\alpha}}\|_1 \leqslant 2^s s! \|Q\|_1^{N-1} \|P\|_1 \prod_{i=0}^{d} (M_i - 1)^{\alpha_i}$.
■

Lemme 8.24 (Comparaison des dimensions).
▶ Soit $N > 1$ un entier quelconque.

- On note $\mathcal{D}$ l'ensemble des opérateurs $\partial_{y_1}^{\alpha_1} \ldots \partial_{y_d}^{\alpha_d} \partial_x^{\alpha_0}$, tels que $\sum_i \alpha_i < N$. Alors $|\mathcal{D}| = \binom{N+d}{d+1}$.
- On note $\mathcal{M}$ l'ensemble des monômes de la forme $y_1^{i_1} \ldots y_d^{i_d}$, avec $i_1 \leqslant N(M_1 - 1)$, …, $i_d \leqslant N(M_d - 1)$. Alors $|\mathcal{M}| = \prod_{i=1}^{d} \big((N(M_i - 1) + 1\big) < N^d \prod_{i=1}^{d} M_i$.
- Pour $N = (d+1)! \prod_{i=1}^{d} M_i - d$, on a l'inégalité stricte $|\mathcal{D}| > |\mathcal{M}|$.

◀

Démonstration. Les deux premiers points sont classiques. Il reste juste à vérifier que pour $N = (d+1)! \prod_{i=1}^{d} M_i - d$, on ait bien $\binom{N+d}{d+1} > N^d \prod_{i=1}^{d} M_i$:

$$\begin{aligned}
\frac{\binom{N+d}{d+1}}{N^d \prod_{i=1}^{d} M_i} &= \frac{(N+d)(N+d-1)\cdots(N+1)N}{(d+1)! N^d \prod_{i=1}^{d} M_i} \\
&= \frac{N+d}{(d+1)! \prod_{i=1}^{d} M_i} \prod_{i=0}^{d-1} (1 + \frac{i}{N}) \\
&\geqslant \frac{N+d}{(d+1)! \prod_{i=1}^{d} M_i} = 1\,.
\end{aligned}$$ ■

Exemple 8.25 (Sanity check).
▶ Dans l'exemple de l'automate $\mathcal{A}_{\text{ex}}$, nous avons $d = 2$ et $M_1 = M_2 = 3$. Ainsi

$$|\mathcal{D}| = \binom{N+2}{3} = N(N+1)(N+2)/6 \quad \text{et} \quad |\mathcal{M}| = (2N+1)^2\,.$$

$N = 22$ le rate de peu : $|\mathcal{D}| = 2024$ tandis que $|\mathcal{M}| = 2025$.

On peut calculer à la main que $N = 23$ est le plus petit N tel que $|\mathcal{D}| > |\mathcal{M}|$. On a alors $|\mathcal{D}| = 2300 > 2209 = |\mathcal{M}|$).

La valeur que nous annonçons dans la proposition précédente est $N = 6 \times 9 - 2 = 52$, qui est bien une borne supérieure de 23. ◀

Proposition 8.26 (Argument de dimension de Lipshitz).
▶ Soit $N = (d+1)!|\mathcal{A}|^d \prod_{i=1}^{d} \|\mathcal{A}\|_{i,\infty} - d$. La famille $\{\partial_{y_1}^{\alpha_1} \dots \partial_{y_d}^{\alpha_d} \partial_x^{\alpha_0} F \ : \ \sum_i \alpha_i < N\}$ est liée sur $\mathbb{Z}[x]$. Plus précisément, il existe une relation de dépendance non triviale de la forme

$$\sum_{\|\boldsymbol{\alpha}\|_1 < N} v_{\boldsymbol{\alpha}}(x)\ \partial_{y_1}^{\alpha_1} \dots \partial_{y_d}^{\alpha_d} \partial_x^{\alpha_0} F = 0,$$

où chaque $v_{\boldsymbol{\alpha}}(x)$ est un polynôme dans $\mathbb{Z}[x]$, de degré en x inférieur à $(d|\mathcal{A}|\|\mathcal{A}\|_\infty)^{O(d^2)}$, et $\log \|v_{\boldsymbol{\alpha}}\|_1 \leqslant (d|\mathcal{A}|\|\mathcal{A}\|_\infty)^{O(d^2)}$. ◀

Démonstration. Nous rappelons que $M_0 = |Q_{\mathcal{A}}| + 1 \leqslant |\mathcal{A}|$, et $M_i = |\mathcal{A}|\|\mathcal{A}\|_{i,\infty}$. Ainsi nous bornons les M_i par $|\mathcal{A}|\|\mathcal{A}\|_\infty$. Avec ces notations,

$$N \leqslant ((d+1)|\mathcal{A}|\|\mathcal{A}\|_\infty)^d \text{ et } |\mathcal{M}| \leqslant (N|\mathcal{A}|\|\mathcal{A}\|_\infty)^d \leqslant (d+1)^{d^2}(|\mathcal{A}|\|\mathcal{A}\|_\infty)^{d(d+1)}\,.$$

Par les Lemmes 8.23 et 8.24, pour tout $\partial_{\boldsymbol{\alpha}} \in \mathcal{D}$, $Q^N \partial_{\boldsymbol{\alpha}} F$ peut s'écrire comme une combinaison linéaire sur $\mathbb{Z}[x]$ d'éléments de $\mathcal{P}$. On peut donc écrire leur décomposition dans une matrice $\mathcal{R}$, qui est la matrice des vecteurs R_α, pour $\|\alpha\|_1 < N$, dans la base $\mathcal{M}$. Ainsi, $\mathcal{R}$ est une matrice de taille $|\mathcal{M}| \times |\mathcal{D}|$ à coefficients dans $\mathbb{Z}[x]$, qui de plus, par le lemme 8.24, a plus de colonnes que de lignes. Les colonnes de cette matrice sont donc liées.

Nous pouvons alors appliquer les formules de Cramer (cf Proposition 8.9) : il existe une solution non nulle $\boldsymbol{v}$ à l'équation $\mathcal{R}\boldsymbol{v} = \mathbf{0}$, telle que chaque coordonnée de $\boldsymbol{v}$ est un mineur de $\mathcal{R}$ ou l'opposé d'un mineur de $\mathcal{R}$ d'ordre $|\mathcal{M}|$. Comme $\mathcal{R}$ est une matrice à coefficients dans $\mathbb{Z}[x]$, les coordonnées de $\boldsymbol{v}$ sont aussi dans $\mathbb{Z}[x]$.

Ainsi, nous pouvons avons trouvé une relation de dépendance de la forme :

$$Q^N \sum_{\|\boldsymbol{\alpha}\|_1 < N} v_{\boldsymbol{\alpha}}(x)\ \partial_{y_1}^{\alpha_1} \dots \partial_{y_d}^{\alpha_d} \partial_x^{\alpha_0} F = 0.$$

Comme Q^N est un polynôme non nul, et la somme $\sum_{\|\boldsymbol{\alpha}\|_1 < N} v_{\boldsymbol{\alpha}}(x)\ \partial_{y_1}^{\alpha_1} \dots \partial_{y_d}^{\alpha_d} \partial_x^{\alpha_0} F$ est un élément du module $\mathcal{M}'$, par la Proposition 8.16, nous en déduisons que

$$\sum_{\|\boldsymbol{\alpha}\|_1 < N} v_{\boldsymbol{\alpha}}(x)\ \partial_{y_1}^{\alpha_1} \dots \partial_{y_d}^{\alpha_d} \partial_x^{\alpha_0} F = 0.$$

De plus pour tout α,

$$\deg_x(v_{\boldsymbol{\alpha}}) \leqslant |\mathcal{M}|N(M_0 - 1) \leqslant (d+1)^{d(d+1)}|\mathcal{A}|^{d(d+2)+1}\|\mathcal{A}\|_\infty^{d(d+2)}\,.$$

Nous pouvons alors appliquer les bornes du Lemme 8.7 aux mineurs de la matrice (en les appliquant à la transposée pour faire intervenir les normes des colonnes) : chaque mineur est borné par le produit des sommes des normes 1 de chacune de ses colonnes. Or chaque colonne du mineur est extrait d'une colonne de $\mathcal{R}$, qui est simplement la décomposition d'un $R_{\boldsymbol{\beta}}$ dans la base canonique de $\mathcal{M}$. Ainsi la somme des normes 1 dans $\mathbb{Z}[x]$ des composantes d'une colonne $\boldsymbol{\beta} \in \mathcal{D}$ de $\mathcal{R}$ est égale à $\|R_{\boldsymbol{\beta}}\|_1$ dans $\mathbb{Z}[x, y_1, \dots, y_d]$.

Pour tout β tel que $\|\beta\|_1 < N$, par les bornes du Lemme 8.23, nous avons

$$\|R_{\boldsymbol{\beta}}\|_1 \leqslant 2^{N-1}(N-1)!\|Q\|_1^{N-1}\|P\|_1(|\mathcal{A}|\|\mathcal{A}\|_\infty)^N\,.$$

Ainsi en utilisant la borne du Lemme 8.7 sur les déterminants :

$$\|v_{\boldsymbol{\alpha}}\|_1 \leqslant \left(2^{N-1}(N-1)!\|Q\|_1^{N-1}\|P\|_1(|\mathcal{A}|\|\mathcal{A}\|_\infty)^N\right)^{|\mathcal{M}|}. \tag{8.3}$$

En passant au log, et en utilisant les bornes sur N et $\mathcal{M}$, et le fait que $\|P\|_1, \|Q\|_1 < 2^{|\mathcal{A}|}$, nous obtenons finalement

$$\begin{aligned}\log(\|v_{\boldsymbol{\alpha}}\|_1) &\leqslant |\mathcal{M}|\big(N-1+(N-1)\log(N)+N|\mathcal{A}|+N\log(|\mathcal{A}|\|\mathcal{A}\|_\infty)\big)\\ &\leqslant (d|\mathcal{A}|\|\mathcal{A}\|_\infty)^{O(d^2)}\,(1+|\mathcal{A}|+d\log(d+1)+(d+1)\log(|\mathcal{A}|\|\mathcal{A}\|_\infty))\\ &= (d|\mathcal{A}|\|\mathcal{A}\|_\infty)^{O(d^2)}\end{aligned}$$

■

Exemple 8.27 .

Cela explique pourquoi j'ai simplifié F dans l'introduction.

▶ Pour l'exemple de l'automate $\mathcal{A}_{\text{ex}}$, je trouve en fait une relation de dépendance pour $N = 8$. L'équation différentielle satisfaite par F est trop grande pour être écrite ici. Elle vérifie $\deg_x(v_{\boldsymbol{\alpha}}) \leqslant 45$ et $\|v_{\boldsymbol{\alpha}}\|_1 \leqslant 348\,086\,586\,267\,256\,320$ pour tout $\|\alpha\|_1 < 8$.

Si on applique les bornes annoncées dans la preuve, avec $N = 8$, $\|Q\|_1 = 18$, $|\mathcal{M}| = 289$, $M_0 = M_1 = M_2 = 3$ (en prenant les degrés exacts de Q), nous obtenons $\deg_x(v_{\boldsymbol{\alpha}}) \leqslant 4624$ et par l'équation (8.3), $\|v_{\boldsymbol{\alpha}}\|_1 \leqslant (101\,108\,579\,083\,223\,040)^{289}$. Même s'il s'agit d'un exemple, qui n'est en aucun cas représentatif, il est probable que nos bornes surestiment beaucoup la taille des coefficients. ◀

8.3.3 Équation différentielle satisfaite par un langage de Parikh faiblement non ambigu

Nous pouvons désormais démontrer la proposition principale de cette section :

Proposition 8.3 (Taille de l'équation différentielle satisfaite par un langage de Parikh faiblement non ambigu).
▶ Soit $\mathcal{A}$ un automate de Parikh faiblement non ambigu. Alors la série $L(x)$ du langage de $\mathcal{A}$ satisfait une équation différentielle linéaire de la forme :

$$q_s(x)\partial_x^s L(x) + \cdots + q_0(x)L(x) = 0\,,$$

avec $s \leqslant (d|\mathcal{A}|\,\|\mathcal{A}\|_\infty)^{O(d)}$, et pour tout $i \in [0, s]$,

$$\begin{aligned}\deg(q_i) &\leqslant (d|\mathcal{A}|\,\|\mathcal{A}\|_\infty)^{O(d^2)}\,,\\ \log\|q_i\|_1 &\leqslant (d|\mathcal{A}|\,\|\mathcal{A}\|_\infty)^{O(d^2)}\,.\end{aligned}$$

◀

Démonstration. D'après la Proposition 8.26, il existe une relation de dépendance non triviale de la forme

$$\sum_{\|\boldsymbol{\alpha}\|_1<N} v_{\boldsymbol{\alpha}}(x)\;\partial_{y_1}^{\alpha_1}\ldots\partial_{y_d}^{\alpha_d}\partial_x^{\alpha_0}F = 0,$$

où $N \leqslant (d|A|\,\|A\|_\infty)^{O(d)}$, chaque $v_{\boldsymbol{\alpha}}(x)$ est un polynôme dans $\mathbb{Z}[x]$, de degré en x inférieur à $(d|\mathcal{A}|\|\mathcal{A}\|_\infty)^{O(d^2)}$, et $\log \|v_{\boldsymbol{\alpha}}\|_1 \leqslant (d|\mathcal{A}|\|\mathcal{A}\|_\infty)^{O(d^2)}$. Nous réécrivons cette équation différentielle sous la forme suivante :

$$\sum_{\|\boldsymbol{\beta}\|_1 < N} p_{\boldsymbol{\beta}}(x, \partial_x)\ \partial_{y_1}^{\beta_1} \dots \partial_{y_d}^{\beta_d} F = 0, \tag{8.4}$$

où pour $\boldsymbol{\beta} \in \mathbb{N}^d$, $p_{\boldsymbol{\beta}}(x, \partial_x) = \sum_{\alpha_0=0}^{N-\|\boldsymbol{\beta}\|_1} v_{\alpha_0,\boldsymbol{\beta}}(x)\partial_x^{\alpha_0}$.

On rappelle que F a été choisie telle que $[y_1^{-1}\cdots y_d^{-1}]F(x,\boldsymbol{y}) = L(x)$. Soit $\mathcal{B}$ l'ensemble des vecteurs qui interviennent dans la relation de dépendance, c'est-à-dire l'ensemble des $\boldsymbol{\beta} \in \mathbb{N}^d$ tels que $\|\boldsymbol{\beta}\|_1 < N$ et $p_{\boldsymbol{\beta}}(x, \partial_x) \neq 0$. Soit $\boldsymbol{\beta^{\min}}$ le plus petit vecteur de $\mathcal{B}$, pour l'ordre lexicographique.

Nous montrons alors que $p_{\boldsymbol{\beta^{\min}}}(x, \partial x)(L) = 0$. Nous rappelons les notations : $\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}} = y_1^{-\beta_1^{\min}} \cdots y_d^{-\beta_d^{\min}}$, et $\boldsymbol{y^{-1}} = y_1^{-1}\cdots y_d^{-1}$.

Pour cela, nous admettons un instant l'égalité suivante : pour tout $\boldsymbol{\beta} \in \mathbb{N}^d$,

$$[\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}]\partial_{y_1}^{\beta_1}\cdots\partial_{y_d}^{\beta_d}F = \begin{cases} 0 & \text{si } \boldsymbol{\beta} \neq \boldsymbol{\beta^{\min}} \\ (-1)^{\beta_1+\dots\beta_d}\beta_1!\dots\beta_d![\boldsymbol{y^{-1}}]F & \text{si } \boldsymbol{\beta} = \boldsymbol{\beta^{\min}} \end{cases} \tag{8.5}$$

On remarque alors que comme les $p_{\boldsymbol{\beta}}(\boldsymbol{x}, \partial x)$ ne dépendent pas des variables $\boldsymbol{y}$, en extrayant dans l'Équation (8.4) les coefficients en $\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}$, on obtient l'égalité :

$$(-1)^{\beta_1+\dots\beta_d}\beta_1!\dots\beta_d!\cdot p_{\boldsymbol{\beta^{\min}}}(x,\partial)[\boldsymbol{y^{-1}}]F = 0.$$

Comme $L(x) = [\boldsymbol{y^{-1}}]F$, nous avons établi que $p_{\boldsymbol{\beta^{\min}}}(x, \partial x)(L) = 0$, ce qui conclut la preuve – les bornes sur les degrés, ordre et coefficients venant directement des bornes sur l'équation de F.

Il reste donc uniquement à prouver l'Équation (8.5). Rappelons que $\boldsymbol{\beta^{\min}}$ est le plus petit vecteur, pour l'ordre lexicographique, qui apparaît dans l'Équation (8.4). Rappelons que F appartient à $\mathbb{Q}[[x, y_1, y_1^{-1}, \dots, y_d^{-1}, y_d]]$. En regroupant les termes en x, on peut écrire F sous la forme :

$$F = \sum_{\boldsymbol{\alpha}\in\mathbb{Z}^d} H_{\boldsymbol{\alpha}}(x) y_1^{\alpha_1}\cdots y_d^{\alpha_d}.$$

Ainsi :

$$[\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}]\partial_{y_1}^{\beta_1}\cdots\partial_{y_n}^{\beta_n}F = [\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}]\sum_{\boldsymbol{\alpha}\in\mathbb{Z}^d} H_{\boldsymbol{\alpha}}(\boldsymbol{x})\partial_{y_1}^{\beta_1}\cdots\partial_{y_n}^{\beta_n} y_1^{\alpha_1}\cdots y_n^{\alpha_n}.$$

On peut alors remarquer que dès qu'un α_i est positif, $\partial_{y_1}^{\beta_1}\cdots\partial_{y_d}^{\beta_d} y_1^{\alpha_1}\cdots y_d^{\alpha_d}$ ne peut contribuer au coefficient de $\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}$. On peut donc limiter la somme précédente aux exposants $\boldsymbol{\alpha}$ qui n'ont que des composantes strictement négatives (donc inférieures ou égales à -1). De plus, si $\boldsymbol{\beta} \in \mathcal{B}$ est différent de $\boldsymbol{\beta^{\min}}$, alors il existe une coordonnée i telle que $\beta_i > \beta_i^{\min}$, par définition de $\boldsymbol{\beta^{\min}}$. Alors l'exposant en y_i de $\partial_{y_1}^{\beta_1}\cdots\partial_{y_d}^{\beta_d} y_1^{\alpha_1}\cdots y_d^{\alpha_d}$ est $\alpha_i - \beta_i < -1 - \beta_i^{\min}$: donc ce terme ne contribue pas non plus au coefficient de $\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}$. De même, par minimalité de $\boldsymbol{\beta^{\min}}$, $\boldsymbol{\alpha}$ doit être égal à $-\mathbf{1}$ pour pouvoir contribuer au terme $\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}$. Finalement :

$$\begin{aligned}
[\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}]\partial_{y_1}^{\beta_1}\cdots\partial_{y_n}^{\beta_n}F &= [\boldsymbol{y}^{-\boldsymbol{\beta^{\min}}-\mathbf{1}}]H_{-\mathbf{1}}(\boldsymbol{x})\partial_{y_1}^{\beta_1^{\min}}\cdots\partial_{y_n}^{\beta_n^{\min}}\boldsymbol{y}^{-1} \\
&= H_{-\mathbf{1}}(\boldsymbol{x})(-1)^{\beta_1^{\min}+\dots+\beta_n^{\min}}\beta_1^{\min}!\cdots\beta_n^{\min}! \\
&= (-1)^{\beta_1^{\min}+\dots+\beta_n^{\min}}\beta_1^{\min}!\cdots\beta_n^{\min}!\ [\boldsymbol{y^{-1}}]F\,,
\end{aligned}$$

qui est bien l'égalité annoncée à l'Équation (8.5). ■

Remarque 8.28 (Comparaison avec la borne de notre article ICALP).
▶ Dans notre article publié à ICALP, nous avions une meilleure borne sur le degré : $\deg(q_i) \leqslant (d|\mathcal{A}|\,\|\mathcal{A}\|_\infty)^{O(d)}$. Cette différence vient du fait que nous avions appliqué l'algorithme de Lipshitz dans $\mathbb{Q}$, et non dans $\mathbb{Q}[x]$. Avec cette méthode, la matrice associée à F pour le calcul de l'équation différentielle est plus grosse, mais les degrés des polynômes sont bornés en x, comme pour l'ordre, par N et non par la taille de la matrice. J'ai décidé d'utiliser des matrices dans $\mathbb{Q}[x]$ dans ce chapitre, car cela permet d'essayer la méthode sur des exemples simples, pas trop gros (la taille de la matrice est un facteur très limitant en pratique). Les tailles des coefficients des polynômes ont dans les deux versions la même borne, qui est au moins exponentiellement plus grande que celle du degré, si bien que cette différence n'a pas de conséquence dans la suite du chapitre . ◀

8.4 Le problème de l'inclusion des automates de Parikh faiblement non ambigus

Nous pouvons enfin attaquer le problème principal de ce chapitre, à savoir le problème de l'inclusion. Nous rappelons le contexte : soient deux automates de Parikh $\mathcal{A}$ et $\mathcal{B}$ faiblement non ambigus, donnés en entrée sous la forme précisée au début de chapitre. Nous notons $d_\mathcal{A}$ et $d_\mathcal{B}$ leurs degrés. Nous cherchons à décider si $\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B})$. Nous rappelons la méthode inspirée de celle de [SI85] pour résoudre le problème :

a. Nous bornons les ordres, degrés, et coefficients des équations différentielles de la série génératrice $A(x) = \sum_n a_n x^n$ de $\mathcal{L}(\mathcal{A})$, et de la série génératrice $C(x) = \sum_n c_n x^n$ de $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$ en sous-section 8.4.1.

b. En sous-section 8.4.1, nous bornons l'ordre, le degré, et les coefficients de l'équation différentielle satisfaite par $D(x) := A(x) - C(x)$, la série de comptage des mots qui sont dans $\mathcal{L}(\mathcal{A})$ mais pas dans $\mathcal{L}(\mathcal{B})$.

c. Nous bornons l'ordre et les racines du polynôme de tête de la récurrence linéaire satisfaite par la suite $d_n := a_n - c_n$, en sous-section 8.4.2. Nous en déduisons la borne sur la taille du plus petit témoin de non inclusion annoncée dans le Théorème 8.4.

d. Nous développons un algorithme élémentaire (qui n'utilise pas d'algorithmes de calcul formel) d'énumération pour résoudre le problème d'inclusion, à l'aide de la borne sur la taille du mot témoin. Cette partie est présentée en sous-section 8.4.3.

8.4.1 Bornes sur les équations différentielles

Dans un premier temps, nous calculons les bornes sur l'équation différentielle satisfaite par la série génératrice $C(x)$ du langage $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$:

Proposition 8.29 (Taille de l'intersection).
▶ La série $C(x)$ du langage $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$ satisfait une équation différentielle linéaire de la forme :

$$q_s(x)\partial_x^s C(x) + \cdots + q_0(x)C(x) = 0\,,$$

avec $s \leqslant ((d_{\mathcal{A}} + d_{\mathcal{B}})|\mathcal{A}||\mathcal{B}| \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty))^{O(d_{\mathcal{A}}+d_{\mathcal{B}})}$, et pour tout $i \in [0, s]$,

$$\deg(q_i) \leqslant ((d_{\mathcal{A}} + d_{\mathcal{B}})|\mathcal{A}||\mathcal{B}| \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty))^{O((d_{\mathcal{A}}+d_{\mathcal{B}})^2)},$$
$$\log \|q_i\|_1 \leqslant ((d_{\mathcal{A}} + d_{\mathcal{B}})|\mathcal{A}||\mathcal{B}| \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty))^{O((d_{\mathcal{A}}+d_{\mathcal{B}})^2)}.$$

◀

Démonstration. Il suffit de spécialiser les bornes de la Proposition 8.3 pour l'automate produit $\mathcal{C} = \mathcal{A} \times \mathcal{B}$ qui calcule l'intersection : on a alors $|\mathcal{C}| \leqslant |\mathcal{A}| \times |\mathcal{B}|$, $d_{\mathcal{C}} = d_{\mathcal{A}} + d_{\mathcal{B}}$ et $\|\mathcal{C}\|_\infty = \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)$. ■

Pour borner l'ordre, le degré, et les coefficients de l'équation différentielle satisfaite par $D(x) := A(x) - C(x)$, nous avons besoin du lemme suivant :

Proposition 8.30 (Somme/soustraction de séries holonomes [Kau14]).
▶ Soient deux séries holonomes $A(x)$ et $C(x)$ satisfaisant deux équations différentielles linéaires non triviales de la forme :

$$p_r^A(x)\partial^r A(x) + \cdots + p_0^A(x)A(x) = 0 \quad \text{et} \quad p_r^C(x)\partial^r C(x) + \cdots + p_0^C(x)C(x) = 0.$$

Soit $D_{\max} = \max_{i\in[r]}(\deg(p_i^A), \deg(p_i^C))$ et $S_\infty = \max_{i\in[r]}(\|p_i^A)\|_\infty, \|p_i^C\|_\infty)$. Alors, $D(x) = A(x) - C(x)$ vérifie une équation différentielle linéaire non triviale de la forme

$$q_{2r}(x)\partial^{2r} D(x) + \cdots + q_0(x)D(x) = 0,$$

où pour tout $i \in [2r]$, $\log(\|q_i\|_\infty) \leqslant O(r^2(1 + \log(r) + \log(D_{\max}) + \log(S_\infty))$
et $\deg(q_i) \leqslant 2(r+1)D_{\max}$.

◀

Démonstration. Nous appliquons le résultat de [Kau14, Theorem 2], qui nous permet d'affirmer que $D(x)$ satisfait une équation différentielle non triviale de la forme :

$$q_{2r}\partial^{2r} D(x) + \cdots + q_0 D(x) = 0$$

où pour chaque $i \in [2r]$, $\deg(q_i) \leqslant 2(r+1)D_{\max}$. De plus, pour tout $i \in [2r]$,

$$\begin{aligned}\mathrm{ht}(\|q_i\|_\infty) \leqslant\ & \mathrm{ht}(2r) + \mathrm{ht}((2r+1)!) + (2r+1)\mathrm{ht}(D_{\max}) \\ & + (2r+2)((2r)(\mathrm{ht}(1) + \mathrm{ht}(D_{\max})) + \mathrm{ht}(S_\infty))\end{aligned}$$

avec $\mathrm{ht}(x) = \log(1 + |x|)$. Comme pour tout $x \geqslant 1$, $\log(1 + x) \leqslant 1 + \log(x)$, nous pouvons en déduire que :

$$\log(\|q_i\|_\infty)) \leqslant 8(r+1)^2 + (2r+2)\log(r) + (4r^2+6r+1)\log(D_{\max}) + (2r+2)\log(S_\infty).$$

■

En appliquant ce lemme à $A(x)$ et $C(x)$, nous déduisons directement la proposition suivante :

Proposition 8.31 (Bornes sur l'équation satisfaite par $D(x)$).
▶ La série $D(x) := A(x) - C(x)$ satisfait une équation différentielle non triviale de la forme

$$q_s(x)\partial_x^s D(x) + \cdots + q_0(x)D(x) = 0\,,$$

avec $s \leqslant ((d_{\mathcal{A}} + d_{\mathcal{B}})|\mathcal{A}||\mathcal{B}|\ \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty))^{O(d_{\mathcal{A}}+d_{\mathcal{B}})}$, et pour tout $i \in [0, s]$,

$$\deg(q_i) \leqslant ((d_{\mathcal{A}} + d_{\mathcal{B}})|\mathcal{A}||\mathcal{B}|\ \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty))^{O((d_{\mathcal{A}}+d_{\mathcal{B}})^2)}\,,$$
$$\log \|q_i\|_\infty \leqslant ((d_{\mathcal{A}} + d_{\mathcal{B}})|\mathcal{A}||\mathcal{B}|\ \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty))^{O((d_{\mathcal{A}}+d_{\mathcal{B}})^2)}\,.$$

◀

Démonstration. On applique la Proposition 8.30, associée aux Propositions 8.3 et 8.29. Les grands O en exposants absorbent beaucoup les détails plus fins des expressions de la Proposition 8.30, et on obtient alors les mêmes types de bornes pour $D(x)$ que pour $C(x)$. ■

8.4.2 Bornes sur la récurrence et témoin de non inclusion

En une variable, la suite des coefficients d'une série holonome $D(x)$ satisfait une équation de récurrence linéaire à coefficients polynomiaux. Nous voulons dans cette section établir des bornes sur le degré et la taille des coefficients du polynôme de tête de la récurrence satisfaite par la suite $d_n = a_n - c_n$, qui compte le nombre de mots qui sont dans $\mathcal{L}(A)$ mais pas dans $\mathcal{L}(B)$. Pour cela, nous avons besoin de transformer l'équation différentielle satisfaite par $D(x)$ en récurrence sur d_n. Nous commençons par une proposition technique :

Proposition 8.32 (Conversion d'une équation différentielle en récurrence linéaire).
▶ Soit $H(x) = \sum u_n x^n$ une série satisfaisant l'équation différentielle suivante :

$$q_r(x)\partial_r H(x) + \ldots + q_0(x)H(x) = 0\,.$$

Alors la suite (u_n) satisfait une récurrence de la forme :

$$\sum_{k=-s}^{S} t_k(n)u_{n+k} = 0, \text{ pour } n \geqslant s, \text{ avec } t_S \neq 0,$$

et de plus $0 \leqslant s \leqslant \max_i(\deg(q_i)), \qquad 0 \leqslant S \leqslant r,$

$$\|t_S\|_\infty \leqslant \max_i(\|q_i\|_\infty) \cdot r^{r+1}, \qquad \deg(t_S) \leqslant r\,.$$

◀

Démonstration. Sans perte de généralité, nous pouvons supposer que pour un certain indice $k_0 \in [0, r]$, $q_{k0}(0) \neq 0$. Sinon, nous pouvons nous ramener à ce cas en divisant l'équation différentielle par le plus grand monôme x^m qui divise tous les polynômes q_k. Remarquons que cette opération de division par x^m ne change pas la norme infinie des coefficients polynomiaux, et diminue leur degré.

Soit $D = \max_i \deg(q_i)$ le plus grand degré des coefficients polynomiaux de l'équation différentielle. Nous pouvons réécrire cette équation satisfaite par H sous la forme suivante :

$$\sum_{k=0}^{r}\sum_{k'=0}^{D} a_{k,k'}x^{k'}\partial_x^k H(x) = 0.$$

Cette équation se traduit alors en la relation de récurrence suivante, valable pour $n \geqslant D$, sur les coefficients de $H(x)$

$$\sum_{k=0}^{r}\sum_{k'=0}^{D} a_{k,k'}(n-k'+1)(n-k'+2)\dots(n-k'+k)u_{n-k'+k} = 0.$$

En posant $j = k - k'$, nous pouvons réécrire cette égalité sous la forme :

$$\sum_{j=-D}^{r} t_j(n)u_{n+j} = 0 \text{ avec } t_j(n) = \sum_{k=\max(0,j)}^{\min(r,D+j)} a_{k,k-j}\prod_{\ell=1}^{k}(n+j-\ell+1)$$

Cependant, certains polynômes t_j peuvent être nuls ; nous devons préciser le terme de tête de la récurrence.

Posons $S := \max\{k-k' \mid k \in [0,r], k' \in [0,D] \text{ et } a_{k,k'} \neq 0\}$. L'hypothèse $q_{k_0}(0) \neq 0$ implique que $a_{k_0,0} \neq 0$, si bien que $S \geqslant k_0 \geqslant 0$.

Montrons que $t_S(n)$ est bien le polynôme de tête dans la récurrence. Par maximalité de S, $t_j(n) = 0$ pour $S < j \leqslant r$. De plus par définition, $t_S(n) = \sum_{k=S}^{r} r_k(n)$ où $r_k(n) = a_{k,k-S}\prod_{\ell=1}^{k}(n+S-\ell+1)$ pour tout $k \in [S,r]$. Par définition de S, au moins un des r_k est non nul. Comme pour tout $k \in [S,r]$, $\deg(r_k) = k$ dès que $a_{k,k-S} \neq 0$, nous savons que $t_S(n) \neq 0$ et $t_S(n)$ est de degré au plus r.

Il ne reste plus qu'à borner la taille des coefficients de t_S. Nous remarquons qu'en développant le produit $\prod_{\ell=1}^{k}(n+S-\ell+1) = \sum_{m=0}^{k} b_m n^m$, nous avons, pour tout $m \in [k]$, $b_m = \sum_{I\subseteq[k],|I|=k-m}\prod_{\ell\in I}(S-\ell+1)$, si bien que $|b_m| \leqslant \binom{k}{m} r^{k-m} \leqslant k^m r^{k-m} \leqslant r^r$. Ainsi :

$$\begin{array}{rcl} \|t_S\|_\infty & \leqslant & \sum_{k=S}^{r}|a_{k,k-S}| \cdot \|\prod_{\ell=1}^{k}(n+S-\ell+1)\|_\infty \\ & \leqslant & \sum_{k=S}^{r}|a_{k,k-S}| \cdot r^r \\ & \leqslant & \max_i(\|q_i\|_\infty) \cdot r^{r+1}. \end{array}$$

■

Pour pouvoir localiser les racines du polynôme de tête, nous avons besoin du lemme suivant :

Lemme 8.33 (Localisation des racines).
▶ Soit $P \neq 0$ un polynôme de $\mathbb{Z}[x]$. Alors $P(n) \neq 0$ pour tout $n \geqslant \|P\|_\infty + 1$. ◀

Démonstration. Soit $n \geqslant \|P\|_\infty + 1$. On écrit $P(x) = \sum_{k=0}^{d} a_k x^k$. Supposons par l'absurde que $P(n) = 0$. Alors :

$$n^d \leqslant \left|a_d n^d\right| = \left|\sum_{k=0}^{d-1} a_k n^k\right| \leqslant \|P\|_\infty \frac{n^d-1}{n-1} \leqslant n^d - 1$$

ce qui mène à une contradiction. ■

Nous pouvons désormais borner la récurrence satisfaite par d_n, et en déduire un majorant de la taille du plus petit mot témoin de non inclusion :

Théorème 8.4 (Taille du plus petit témoin de non inclusion).
▶ Si $\mathcal{L}(A)$ n'est pas inclus dans $\mathcal{L}(B)$, alors il existe un mot w qui soit dans $\mathcal{L}(\mathcal{A})$ mais pas dans $\mathcal{L}(\mathcal{B})$ tel que :

$$|w| \leqslant 2^{(dM)^{O(d^2)}},$$

avec $d = d_{\mathcal{A}} + d_{\mathcal{B}}$, et $M = |\mathcal{A}||\mathcal{B}|\max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)$. ◀

Démonstration. Pour $n \in \mathbb{N}$, on note a_n (resp c_n) le nombre de mots de taille n dans $\mathcal{L}(\mathcal{A})$ (resp. $\mathcal{L}(\mathcal{B})$). On pose $d_n = a_n - c_n$ le nombre de mots de taille n qui sont dans $\mathcal{L}(\mathcal{A})$ mais pas dans $\mathcal{L}(B)$. Alors par les Proposition 8.31 et Proposition 8.32, la suite (d_n) satisfait une récurrence linéaire de la forme

$$\sum_{k=-s}^{S} t_k(n) d_{n+k} = 0, \quad \text{pour } n \geqslant s, \text{ avec } t_S \neq 0$$

que l'on réécrit sous la forme :

$$t_S(n-S)d_n = -\sum_{k=1}^{S+s} t_{S-k}(n-S)d_{n-k} \quad \text{pour } n \geqslant S+s,$$

avec $S + s \leqslant (dM)^{O(d)}$, $\deg(t_S) \leqslant (dM)^{O(d)}$ et $\log(\|t_S\|_\infty) \leqslant (dM)^{O(d^2)}$.

Soit $n_0 \in \mathbb{N}$. Si $t_S(n_0 - S)$ n'est pas nul, et si $u_{n_0-1} = \cdots = u_{n_0-S-s} = 0$ alors $u_{n_0} = 0$. Pour être sûr que $n_0 - S$ n'est pas une racine de t_S, il suffit de choisir $n_0 - S \geqslant \|t_S\|_\infty + 1$ par le Lemme 8.33). En posant $W = s + S + \|t_S\|_\infty + 1$, nous avons donc l'équivalence :

$$D(x) = 0 \quad \text{si et seulement si} \quad \forall n \leqslant W,\ d_n = 0\,.$$

Autrement dit, du point de vue des langages, $L(\mathcal{A}) \not\subseteq L(\mathcal{B})$ si et seulement si $L(\mathcal{A}) \setminus L(\mathcal{B})$ contient un mot de longueur au plus $W \leqslant 2^{(dM)^{O(d^2)}}$. ■

8.4.3 Borne de complexité du problème de l'inclusion

Dans cette section, nous développons un algorithme de décision du problème de l'inclusion. Cet algorithme est élémentaire dans le sens où il ne demande aucun outil de calcul formel. Le but est de prouver le corollaire suivante :

Corollaire 8.5 (Borne de complexité du problème de l'inclusion).
▶ Soient deux automates de Parikh faiblement non ambigus $\mathcal{A}$ and $\mathcal{B}$ de dimensions $d_\mathcal{A}$ et $d_\mathcal{B}$. On peut décider le problème de l'inclusion $L(\mathcal{A}) \subseteq L(\mathcal{B})$ en temps $2^{2^{O(d^2 \log(dM))}}$ où $d = d_\mathcal{A} + d_\mathcal{B}$ et $M = |\mathcal{A}|\,|\mathcal{B}|\,\max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)$. ◀

D'après le Théorème 8.4, si $\mathcal{L}(\mathcal{A}) \not\subseteq \mathcal{L}(\mathcal{B})$ alors on peut trouver un mot w de taille $|w| \leqslant W = 2^{(dM)^{O(d^2)}}$, avec $d = d_\mathcal{A} + d_\mathcal{B}$ et $M = |\mathcal{A}||\mathcal{B}| \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)$, qui soit dans $\mathcal{L}(\mathcal{A})$ mais pas dans $\mathcal{L}(\mathcal{B})$. L'algorithme de résolution du problème de l'inclusion est simple : nous allons compter sur les automates le nombre de mots de taille n dans $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$ et dans $\mathcal{L}(\mathcal{A})$, jusqu'à la taille $n = W$. Autrement dit nous calculons sur les automates les W premiers termes des séries génératrices des langages $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$ et $\mathcal{L}(\mathcal{A})$. Si les deux valeurs calculées correspondent jusqu'à la taille W, alors il y a bien inclusion $\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{B})$, sinon on a trouvé la taille du plus petit mot dans $\mathcal{L}(\mathcal{B}) \setminus \mathcal{L}(\mathcal{A})$.

Une approche naïve consiste à énumérer tous les calculs possibles de longueur n des automates, et de vérifier si le calcul est acceptant ; cependant cette énumération a un surcoût exponentiel en n, et comme n va aller jusqu'à W, qui est déjà doublement exponentiel, on aimerait éviter une exponentielle supplémentaire.

La proposition suivante explique comment compter simplement les calculs de longueur n de l'automate :

Proposition 8.34 (Dénombrement élémentaire des calculs de l'automate).
▶ Soit $\mathcal{A} = (\Sigma, Q, q_I, F, C, \Delta)$ un automate de Parikh faiblement non ambigu de dimension $d \geqslant 1$, donné sous la forme précisée au début du chapitre. Soit $W \in \mathbb{N}$ tel que $W \geqslant d|\mathcal{A}|\|\mathcal{A}\|_\infty$. Alors on peut compter le nombre total de calculs de $\mathcal{A}$ de longueur inférieure à W qui joignent q_I à un état final, groupés par valeur du vecteur qui les étiquette, en temps $W^{O(d)}$. ◀

Démonstration. Dans un premier temps, nous remarquons que tout calcul de longueur n dans $\mathcal{A}$ est étiqueté par un vecteur $\boldsymbol{v}$ de norme infinie inférieure à $n\|\mathcal{A}\|_\infty$.

Pour tout $q \in Q$, tout entier $n \leqslant W$, tout vecteur $\boldsymbol{v} \in \mathbb{N}^d$, nous notons $a_q(n, \boldsymbol{v})$ le nombre de calculs de $\mathcal{A}$ de longueur n, partant de q, finissant dans un état final, et étiquetés par le vecteur $\boldsymbol{v}$. On a facilement $a_q(n, \boldsymbol{v}) \leqslant |\Delta|^n$.

Pour $\boldsymbol{v}, \boldsymbol{u}$ deux vecteurs de $\mathbb{N}^d$, nous notons $\boldsymbol{u} \preceq \boldsymbol{v}$ si pour tout $i \in [d]$, $u_i \leqslant v_i$.

On vérifie simplement que les $a_q(n, \boldsymbol{v})$ vérifient la récurrence suivante, pour tout $q \in Q$, $n \geqslant 1$, et $\|\boldsymbol{v}\|_\infty \leqslant n\|\mathcal{A}\|_\infty$:

Il s'agit simplement du système linéaire satisfait par les séries génératrices, exprimé du point de vue des coefficients.

$$a_q(n, \boldsymbol{v}) = \sum_{\substack{(q,(a,\boldsymbol{u}),q') \in \Delta \\ \text{avec } \boldsymbol{u} \preceq \boldsymbol{v}}} a_{q'}(n-1, \boldsymbol{v} - \boldsymbol{u}). \tag{8.6}$$

avec comme conditions initiales $a_q(0, \mathbf{0}) = 1$ si $q \in F$, 0 sinon.

L'énumération des vecteurs $\boldsymbol{v}$ de norme infinie inférieure à $W\|\mathcal{A}\|_\infty$, dans l'ordre lexicographique, se fait en $(W\|\mathcal{A}\|_\infty + 1)^d \leqslant W^{O(d)}$ itérations. Pour chaque vecteur $\boldsymbol{v}$, et pour tout n allant de 1 à W, on utilise l'équation 8.6 pour calculer $a_q(n, \boldsymbol{v})$ (l'ordre lexicographique assure que toutes les valeurs de la somme de l'équation 8.6 ont bien été calculées avant) : on effectue une itération sur les transitions sortantes de q, qui sont au plus $|\Delta|$; dans chacune de ces transitions, on calcule la valeur du vecteur $\boldsymbol{v} - \boldsymbol{u}$ en temps $O\big(d \log(n\|\mathcal{A}\|_\infty)\big) \leqslant O(W^2)$, puis on ajoute la valeur de $a_{q'}(n, \boldsymbol{u})$ à $a_q(n, \boldsymbol{v})$, en temps $O\big(n \log(|\Delta|)\big) \leqslant O(W^2)$.

On peut donc calculer tous les $a_q(n, \boldsymbol{v})$, pour tout $n \leqslant W$, tout $q \in Q$ et $\|\boldsymbol{v}\|_\infty \leqslant W\|\mathcal{A}\|_\infty$ en $W^{O(d)}$ opérations.

■

Proposition 8.35 (Énumération des vecteurs du semilinéaire).
▶ Soit $C = \cup_{i=1}^{p} C_i$ un semilinéaire de dimension $d \geqslant 1$, représenté sous une forme non ambiguë, avec $C_i = \boldsymbol{c_i} + P_i^*$. Soit $r \in \mathbb{N}$ tel que $r \geqslant dp$. Alors on peut énumérer tous les vecteurs de C de norme infinie plus petite que r en temps $r^{O(\sum_i |P_i|)}$. ◀

Démonstration. Soit $1 \leqslant i \leqslant p$ tel que $r \geqslant \|\boldsymbol{c_i}\|_\infty$. On note $P_i = \{\boldsymbol{p_1}, \ldots, \boldsymbol{p_{|P_i|}}\}$ (on omet les indices i pour plus de lisibilité). Quitte à retirer certains de vecteurs P_i, on peut supposer que $r \geqslant \max_{\boldsymbol{p} \in P_i}(\|\boldsymbol{p}\|_\infty)$.

Pour tout vecteur $\boldsymbol{u}$ dans $[0, r]^{|P_i|}$, on note $b_i(\boldsymbol{u})$ le vecteur $\boldsymbol{c_i} + \sum_{k=1}^{|P_i|} u_k \boldsymbol{p_k}$. On énumère ainsi tous les vecteurs de C_i par une simple boucle à $(r+1)^{|P_i|}$ itérations. Si $\boldsymbol{u}$ est tel que $u_k \geqslant 1$ alors $b_i(\boldsymbol{u}) = b_i(\boldsymbol{u} - \boldsymbol{e_k}) + \boldsymbol{p_k}$ se calcule en temps $O(d \log(r))$.

On peut donc énumérer tous les vecteurs de C_i en temps $O((r+1)^{|P_i|} d \log(r)) \leqslant r^{O(|P_i|)}$, et donc tous les vecteurs de C en $r^{O(\sum_i |P_i|)}$ opérations. ■

Nous pouvons désormais démontrer le Corollaire 8.5 :

Preuve du Corollaire 8.5. On pose $W = 2^{(dM)^{O(d^2)}}$, avec $d = d_{\mathcal{A}} + d_{\mathcal{B}}$ et $M = |\mathcal{A}||\mathcal{B}| \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)$. On rappelle que $\mathcal{C}$ désigne l'automate produit qui calcule $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$.

Nous calculons dans un premier temps les vecteurs du semilinéaire de $\mathcal{A}$ et de $\mathcal{C}$ de norme infinie inférieure à $W \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)$. D'après la Proposition 8.35, cela se fait en $(W \max(\|\mathcal{A}\|_\infty, \|\mathcal{B}\|_\infty)^{O(|\mathcal{A}|+|\mathcal{B}|)}$ opérations.

Nous comptons ensuite tous les calculs de longueur inférieure à W joignant l'état initial à un état final, groupés par le vecteur qui les étiquette, de l'automate $\mathcal{A}$ et de l'automate produit $\mathcal{C}$. Cela se fait d'après la Proposition 8.34 en $W^{O(d)}$ opérations. On modifie légèrement la procédure de la Proposition 8.34, sans changer sa complexité, pour additionner lors de l'énumération des vecteurs, tous les $a_{q_I}(n, \boldsymbol{v})$ calculés tels que $\boldsymbol{v} \in C$ afin d'avoir à la fin de la procédure, pour chaque $n \leqslant W$, le nombre de calculs acceptants de longueur n.

On peut ainsi calculer en $2^{(dM)^{O(d^2)}}$ opérations le nombre de mots de longueur inférieure à W dans $\mathcal{L}(\mathcal{A})$ et dans $\mathcal{L}(\mathcal{A}) \cap \mathcal{L}(\mathcal{B})$, et ainsi décider le problème de l'inclusion en $2^{(dM)^{O(d^2)}}$ opérations . ■

8.5 Conclusion, et ouverture

Dans ce chapitre, nous avons utilisé le caractère holonome de la série génératrice des automates de Parikh faiblement non ambigus pour en déduire un algorithme élémentaire de résolution du problème de l'inclusion, avec une première borne sur sa complexité. Comme nous l'avons expliqué, nous sommes allés au chemin le plus direct pour obtenir ces bornes, qui s'avèrent sur des exemples simples surestimer un peu trop la taille des coefficients.

Je pense qu'il est possible, en suivant la même démarche que nous avons utilisée, d'obtenir des majorations plus petites, à l'aide de bornes plus précises lors des calculs, et notamment pour l'étude de l'algorithme de Lipshitz. Il arrive aussi en calcul formel que les coefficients des calculs intermédiaires d'un algorithme grossissent artificiellement pour s'effondrer à la toute fin ; c'est le cas par exemple pour le calcul du pgcd par l'algorithme d'Euclide [VZGG13]. Il faudrait s'assurer que nous ne sommes pas confrontés au même problème dans notre analyse de l'algorithme de Lipshitz.

Une autre possibilité, pour continuer le travail dans une autre direction, serait d'étudier des algorithmes plus efficaces de calcul de diagonales de fractions rationnelles, comme ceux à base de *télescopage créatif* ([Chy14, BCLS18]) : peut-être permettront-ils d'obtenir des bornes plus fines sur les équations différentielles. C'est un travail que j'aimerais bien continuer, mais qui demandera un investissement mathématique important. Une autre idée serait d'essayer de prendre en compte le caractère $\mathbb{N}$-rationnel des séries étudiées dans les algorithmes de calcul de diagonales.

Dans ce chapitre, nous avons utilisé les algorithmes sur les séries holonomes pour borner uniquement la taille du plus petit témoin de non inclusion : l'algorithme final que nous fournissons pour résoudre l'inclusion est élémentaire et repose uniquement sur cette borne. Il serait intéressant de résoudre le problème de l'inclusion en calculant en pratique les équations différentielles avec des outils de calcul formel, et d'étudier sa complexité.

Pour ce qui concerne la complexité du problème de l'inclusion, il nous manque des familles de langages pour tester nos bornes, qui vérifieraient que la taille du plus petit témoin de non inclusion est "grande" : si possible exponentielle ou doublement exponentielle en la taille des automates. Avant d'étudier la complexité théorique du problème de l'inclusion des automates de Parikh faiblement non ambigus, il est

plus intéressant d'étudier celle de l'universalité : étant donné un automate de Parikh faiblement non ambigu $\mathcal{A}$, est-ce que $\mathcal{L}(\mathcal{A}) = \Sigma^*$? En effet, dans [Cle20], l'auteur explique comment le problème de l'inclusion $L \subseteq M$ est souvent équivalent au problème de l'universalité de $(M \cap L) \cup (\Sigma^* \setminus L)$, où L peut être supposé sans perte de généralité déterministe. Son approche présente ainsi le problème de l'universalité comme central dans l'étude du problème de l'inclusion pour de nombreuses classes d'automates, dont les automates de Parikh faiblement non ambigus : comme les automates de Parikh déterministes sont clos par complémentaire, et les automates de Parikh faiblement non ambigus sont clos par union disjointe et intersection, le langage $(M \cap L) \cup (\Sigma^* \setminus L)$ est bien reconnu par un automate de Parikh faiblement non ambigu.

Enfin, il est naturel d'étendre la démarche de ce chapitre avec les automates de Parikh à pile. Le problème de l'inclusion étant indécidable pour les langages algébriques non ambigus, la méthode ne se généralise pas, mais elle peut s'appliquer au problème de l'universalité $\mathcal{L}(\mathcal{A}) = \Sigma^*$. Deux directions sont possibles pour aborder le problème :

— reprendre l'algorithme de Lipshitz depuis le début, et l'analyser dans le cas où la série holonome dont on calcule une diagonale est algébrique ;

— sinon réutiliser l'étude du cas rationnel, en utilisant le fait que toute série algébrique est une diagonale de fraction rationnelle [DL87].

La seconde me semble dans un premier temps plus envisageable que la première.