Calcul rapide d'un coefficient d'une série algébrique en caractéristique positive

Alin Bostan
INRIA
alin.bostan@inria.fr

Gilles Christol
IMJ
gilles.christol@imj-prg.fr

Philippe Dumas
INRIA
philippe.dumas@inria.fr

Une série algébrique est une série formelle $y = f(x) \in \mathbb{Q}[[x]]$ solution d'une équation algébrique $E(x, y) = 0$, où $E(x, y)$ est un polynôme de $\mathbb{Z}[x, y]$. Par exemple, la série énumérative des arbres 2–3

$$f(x) = 1 + 2\,x + 10\,x^2 + 66\,x^3 + 498\,x^4 + 4066\,x^5 + \cdots$$

est algébrique et satisfait à l'équation algébrique $y = 1 + xy^2 + xy^3$ (il y a soixante-six arbres 2–3 à trois sommets internes, comme le montre le coefficient $[x^3]f$ de x^3 dans f).

Étant donnés un nombre entier N et un nombre premier p, nous nous intéressons dans ce travail au calcul efficace de $[x^N]f \bmod p$, le reste modulo p du coefficient de x^N dans f. En d'autres termes, il s'agit de calculer rapidement le coefficient $[x^N]\bar{f}$ d'une série $\bar{f} \in \mathbb{F}_p[[x]]$, racine d'un polynôme à deux variables $\bar{E}[x, y]$ sur le corps fini $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Par exemple, nous souhaitons calculer le milliardième coefficient modulo 9001 de la série des arbres 2–3.

Une idée simple est de substituer le développement à l'ordre N de la série dans l'équation et d'identifier à zéro. Cela fait résoudre un système non linéaire dans lequel chaque terme est déterminé par les précédents. Le calcul du N^{e} terme de $\bar{f}$ par cette méthode demande un nombre d'opérations arithmétiques dans $\mathbb{F}_p$ de l'ordre de N^d, où d est le degré de l'équation $\bar{E} = 0$. L'itération formelle de Newton permet d'accélérer ce calcul, en atteignant une complexité $O(N \log N)$, où d n'apparaît que dans la constante du $O(\,)$.

Mais il y a bien plus efficace pour calculer directement le N^{e} terme de $\bar{f} \in \mathbb{F}_p[[x]]$. On utilise pour ce faire l'automaticité des séries algébriques sur un corps fini [1, Cor. 4.5] : l'entier N se décompose en base p et chacun de ses $\log_p N$ chiffres permet d'effectuer une transition dans un automate, à partir de son état initial, ce qui fournit en sortie la valeur du coefficient désiré. Ce calcul demande de l'ordre de $\log_p N$ opérations arithmétiques dans $\mathbb{F}_p$.

Encore faut-il disposer d'un automate, et son précalcul s'avère coûteux. La construction d'un automate à partir de l'équation algébrique [3] passe par

l'écriture d'une autre équation algébrique, d'une forme particulière (de Mahler), c'est-à-dire utilisant l'opérateur de Frobenius $g(x) \mapsto g(x^p) = g(x)^p$. Le coût du calcul de cette équation de Mahler à partir d'une équation algébrique $\bar{E}(x, y) = 0$ est en général de l'ordre de $p^{d(d-1)/2}$.

Une meilleure approche repose sur la considération d'une fraction rationnelle dont la série algébrique est la diagonale [5], [2, Cor. 13, p. 6-09]. Par exemple la série des arbres 2–3 est la diagonale de la fraction rationnelle

$$\frac{3\,xy^4 + 9\,xy^3 + 9\,xy^2 + 5\,xy + 2\,x - y - 1}{xy^3 + 4\,xy^2 + 5\,xy + 2\,x - 1} =$$
$$\underline{1} + y + (\underline{2\,xy}) + (4\,x^2y) + (8\,x^3y + \underline{10\,x^2y^2} - 4\,xy^3)$$
$$+ (16\,x^4y + 40\,x^3y^2 - 2\,xy^4) + (32\,x^5y + 120\,x^4y^2 + \underline{66\,x^3y^3} - 22\,x^2y^4) + \cdots$$

Le précalcul s'effectue dans un espace de polynômes à deux variables dont les degrés partiels sont bornés par ceux de la fraction rationnelle [2, Th. 32]. Nous montrons ainsi que l'utilisation des diagonales comme structure de données permet d'aboutir à une complexité essentiellement quadratique en p pour le précalcul, ce qui constitue déjà une énorme amélioration.

Nous accélérons encore ce calcul : nous montrons, et ceci est notre principale contribution, que l'évaluation du N^{e} coefficient d'une série algébrique de $\mathbb{F}_p[[x]]$ peut se faire en $O(\log_p N + p \log^2 p)$ opérations dans $\mathbb{F}_p$, où la constante du $O(\,)$ ne dépend que du degré d. Ceci est à comparer au calcul de tous les coefficients jusqu'au rang N, qui ne peut pas demander moins de N opérations, ou au calcul direct du N^{e} coefficient dont la meilleure complexité annoncée (mais non explicitée) est $O(p \times \log_p N)$ [4, p. 121].

Bibliographie

[1] Jean-Paul Allouche et Jeffrey Shallit. The ring of k-regular sequences. *Theoret. Comput. Sci.*, 98(2):163–197, 1992.

[2] Gilles Christol. Éléments analytiques uniformes et multiformes. In *Séminaire Delange-Pisot-Poitou (15e année : 1973/74), Théorie des nombres, Fasc. 1, Exp. No. 6*, page 18. Secrétariat Mathématique, Paris, 1975.

[3] Gilles Christol, Teturo Kamae, Michel Mendès-France et Gérard Rauzy. Suites algébriques, automates et substitutions. *Bull. Soc. Math. France*, 108(4):401–419, 1980.

[4] David V. Chudnovsky et Gregory V. Chudnovsky. Computer algebra in the service of mathematical physics and number theory. In *Computers in mathematics (Stanford, CA, 1986)*, volume 125 of *Lecture Notes in Pure and Appl. Math.*, pages 109–232. Dekker, New York, 1990.

[5] Harry Furstenberg. Algebraic functions over finite fields. *J. Algebra*, 7:271–277, 1967.