ADVANCES IN APPLIED MATHEMATICS 3, 407–416 (1982)

Une approche combinatoire pour l'itération de Newton – Raphson

H. DÉCOSTE, G. LABELLE,* ET P. LEROUX*

Département de Mathématiques, Université du Québec à Montréal, Montréal, Québec H3C 3P8, Canada

Starting with an approximation α having a contact of order n with the species A of R-enriched rooted trees (in the sense of Joyal (*Advances in Math*. **42** (1981), 1–82) and Labelle (*Advances in Math*. **42** (1981), 217–247)), a new approximation α^+, having a contact of order $2n+2$ with A, is deduced by a purely combinatorial argumentation. This provides a combinatorial setting for the classical Newton–Raphson iterative scheme. A generalization involving contacts of higher orders is also developed.

0. INTRODUCTION

La méthode itérative (dite de la "tangente") de Newton–Raphson est bien connue. Elle permet d'obtenir des approximations successives α_n, $n = 0, 1, 2, \ldots$, d'une racine a d'une équation de la forme $f(t) = 0$ en partant d'une approximation initiale (bien choisie) α_0 et en définissant les suivantes par $\alpha_{n+1} = \alpha_n^+$, où

$$\alpha^+ = \alpha - [f'(\alpha)]^{-1} f(\alpha) \tag{0.1}$$

est le point d'intersection, avec l'axe des t, de la tangente au graphe de f au point $(\alpha, f(\alpha))$. L'efficacité de la méthode provient du fait que sous des hypothèses très faibles on a convergence "quadratique" de la suite α_n vers la racine a, c'est-à-dire,

$$a - \alpha_{n+1} = O\big((a - \alpha_n)^2\big) \to 0 \qquad \text{si } n \to \infty. \tag{0.2}$$

Par exemple, si f est de classe C^2 au voisinage de a et si $f'(a) \neq 0$ alors un

*Avec l'appui financier du Ministère de l'Education du Québec, subvention d'équipe EQ 1608.

0196-8858/82/040407-10$05.00/0

développement limité de Taylor dans (0.1) au point a montre que

$$a - \alpha^{+} = -\frac{f''(a)}{2f'(a)} \cdot (a - \alpha)^2 + \cdots . \tag{0.3}$$

Ce phénomène de convergence rapide est couramment décrit en langage numérique en disant qu'à toutes fins pratiques, le nombre de décimales exactes pour la racine a est *doublé* à chaque pas d'itération.

La version classique de l'itération de Newton–Raphson que nous venons de rappeler fait partie intégrante du domaine de l'Analyse et ne se prête pas, comme telle, à une interprétation combinatoire directe. Néanmoins, c'est en la traduisant d'abord dans le contexte des séries formelles (section 1) que nous pourrons ensuite, à l'aide du langage et des méthodes de la théorie des espèces de structures au sens de Joyal [2], la faire "basculer" dans le domaine de la Combinatoire proprement dite (section 2). Techniquement parlant, l'itération de Newton–Raphson apparaîtra combinatoirement comme la solution naturelle au problème de la formation d'une suite d'espèces de structures convergeant quadratiquement (dans un sens que l'on précisera) vers l'espèce des arborescences enrichies au sens de Labelle [3]. Finalement, nous introduisons dans la section 3 un concept nouveau qui est celui de *nervure d'ordre k* d'une arborescence enrichie. Ce concept permet de généraliser notre traitement de l'itération de Newton–Raphson au cas de la convergence d'ordre supérieur (i.e., cubique, biquadratique,..., etc.).

1. Le Cas des Séries Formelles

On peut considérer que le problème de résoudre une équation se ramène essentiellement à celui du calcul d'un inverse local pour une fonction. Il est donc naturel, dans le contexte des séries formelles cette fois, d'interpréter l'itération de Newton–Raphson comme une méthode permettant l'inversion rapide de ces séries sous l'opération de substitution. Par exemple, soit $\mathbb{Q}$ le corps des rationnels (ou, plus généralement, un corps commutatif de caractéristique zéro) et soit

$$c = c(x) = c_1 x + c_2 x^2 + \cdots \in \mathbb{Q}[[x]]$$

une série formelle donnée (avec $c_1 \neq 0$). En prenant $f(t) = c(t) - x$ où $t = t(x)$, on voit immédiatement que la solution a de l'équation $f(t) = 0$ est la série formelle

$$a = a(x) = c^{\langle -1 \rangle}(x) \in \mathbb{Q}[[x]]$$

qui est l'inverse de c sous la substitution. Utilisant l'expression (0.1) on peut alors former une suite d'approximations $\alpha_i = \alpha_i(x)$, $i \geq 0$, de la solution $a = a(x)$ en posant $\alpha_0 = 0$ et $\alpha_{i+1} = \alpha_i^+$, $i \geq 0$, avec

$$\alpha^+ = \alpha - \frac{c(\alpha) - x}{c'(\alpha)}. \tag{1.1}$$

Brent et Kung [1] ont étudié l'efficacité de cette méthode d'inversion sous l'angle de la complexité des calculs qu'elle implique. La formule (0.3) étant encore valable on a, en particulier, que le phénomène de la convergence quadratique se traduit comme suit: *si* $\alpha(x)$ *possède un contact d'ordre n avec* $a(x)$ (i.e., les termes de ces séries coincident jusqu'à l'ordre n) *alors* $\alpha^+(x)$ *possède un contact d'ordre* $2n + 1$ *avec* $a(x)$.

C'est donc dire que le nombre de *termes* exacts est plus que *doublé* à chaque pas d'itération.

Même cette version plus algébrique de l'itération de Newton–Raphson ne se relève pas, comme telle, au plan combinatoire. La principale difficulté est due au fait que (sauf dans des cas triviaux) la positivité des coefficients de $c(x)$ n'entraîne pas la positivité des coefficients de $a(x)$.

Une façon de faire disparaître cette difficulté est d'écrire, sans perte de généralité, la série à inverser sous la forme

$$c(x) = x/r(x), \tag{1.2}$$

où $r(x) \in \mathbb{Q}[[x]]$ satisfait $r(0) \neq 0$. Le problème de l'inversion de la série $c(x)$ revient alors à résoudre l'équation

$$f(t) = 0, \qquad \text{où } f(t) = xr(t) - t. \tag{1.3}$$

Avec ce nouveau choix pour f, la formule (0.1) s'écrit

$$\alpha^+ = \alpha + \frac{xr(\alpha) - \alpha}{1 - xr'(\alpha)} \tag{1.4}$$

et on vérifie que la positivité des coefficients de la série $r(x)$ entraîne la positivité des coefficients de la série $a(x)$, inverse de $c(x)$ (par exemple, en invoquant la formule d'inversion de Lagrange, voir [3]). Cette dernière version de l'itération de Newton–Raphson est en fait plus efficace que la précédente puisque (0.3) montre qu'*un contact d'ordre n entre* α *et a est, cette fois, transformé en un contact d'ordre* $2n + 2$ *entre* α^+ *et a*. De plus, une approche combinatoire pour (1.4) est maintenant possible. C'est ce que nous allons voir dans la section qui vient.

2. L'APPROCHE COMBINATOIRE

Pour nous situer dans un cadre adéquat, nous allons faire appel à certains concepts fondamentaux de la récente théorie combinatoire des espèces de structures au sens de Joyal. Cette théorie permet de représenter de façon purement combinatoire (lire ensembliste) les séries génératrices exponentielles ainsi que les opérations usuelles sur celles-ci, principalement l'addition, la multiplication, la substitution et la dérivation. Le lecteur non familier avec le langage et les méthodes de cette théorie pourra consulter [2] qui est l'article de fond sur le sujet.

Plus spécifiquement, nous nous proposons de *déduire combinatoirement* la version (1.4) de l'itération de Newton–Raphson à partir d'un *problème d'approximation* entre espèces de structures.

Par convention, le degré d'approximation d'une espèce donnée M par une espèce N est mesuré par l'*ordre de contact* que les espèces M et N ont entre elles: un contact d'ordre n, noté

$$M \stackrel{n}{=} N, \tag{2.1}$$

signifie que les espèces "tronquées" obtenues de M et N en oubliant les structures portées par les cardinalités $> n$ sont isomorphes. On remarquera, en particulier, que les séries génératrices $M(x)$ et $N(x)$ associées aux espèces M et N ont alors un contact d'ordre n au sens de la section précédente. Dans le cas où les deux espèces M et N sont isomorphes on écrira simplement $M = N$ pour ne pas alourdir les notations.

En faisant appel aux opérations générales (addition, multiplication, substitution, etc) entre espèces de structures, le problème de l'inversion des séries formelles (version (1.3)) se relève au plan combinatoire comme suit: Résoudre pour l'espèce inconnue T l'équation combinatoire

$$\begin{aligned} f(T) = XR(T) - T = 0, \\ \text{c'est-à-dire,} \qquad T = XR(T), \end{aligned} \tag{2.2}$$

où R est une espèce *fixée* quelconque et X est l'espèce des singletons. La solution de cette équation est l'espèce A des arborescences R-enrichies au sens de Labelle [3]: une arborescence (= arbre avec racine) étiquettée par les éléments d'un ensemble fini U est dite *R-enrichie* si la fibre de chaque élément (pour l'orientation canonique des arêtes en direction de la racine) est munie d'une R-structure.

Pour le voir, il suffit de remarquer que l'isomorphisme

$$A = XR(A), \tag{2.3}$$

qui caractérise la solution A s'interprète en disant que la donnée d'une

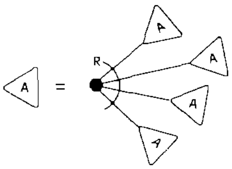

FIGURE 1

A-structure est canoniquement équivalente à la donnée d'un singleton suivi d'une R-assemblée de A structures (voir figure 1). Itérant cette analyse jusqu'à épuisement de l'ensemble fini sous-jacent on arrive à la conclusion qu'une A-structure est bien une arborescence R-enrichie (voir figure 2 dans laquelle les R-structures placées sur les diverses fibres sont symbolisées par des arcs de cercle). Dorénavant, toutes nos arborescences seront R-enrichies sans qu'il soit nécessaire de le spécifier explicitement. *Fixons* un entier $n \geq 0$ et posons-nous le problème d'approximation combinatoire suivant:

A partir d'une espèce α ayant un contact d'ordre n avec l'espèce A(i.e., $\alpha \overset{n}{=} A$), *construire combinatoirement et de façon canonique une meilleure approximation α^+ de A au sens où, cette fois,*

$$\alpha^+ \overset{2n+2}{=\!=} A. \tag{2.4}$$

Etant donné la nature du problème, nous pouvons supposer que α est l'espèce des arborescences *légères* au sens de la définition suivante.

DÉFINITION 1. Une arborescence (R-enrichie) sur un ensemble fini U est dite

(1) *légère* si $|U| \leq n$,

(2) *lourde* si $n + 1 \leq |U| \leq 2n + 2$,

où $|U|$ représente la cardinalité de l'ensemble U.

Désignons par β l'espèce des arborescences lourdes. Etant donné qu'une arborescence quelconque sur un ensemble U de cardinalité $|U| \leq 2n + 2$ est

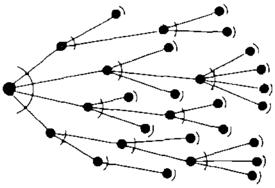

FIGURE 2

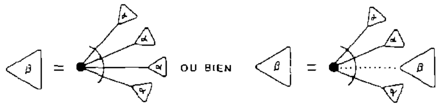

FIGURE 3

soit légère, soit lourde, on a évidemment

$$\alpha + \beta \overset{2n+2}{=\!=} A. \tag{2.5}$$

Ainsi notre problème d'approximation se ramène à la construction de l'espèce β des arborescences lourdes à partir de l'espèce α des arborescences légères. Analysons donc la structure de l'arborescence lourde générique (sur un ensemble de cardinalité $\leq 2n + 2$).

Un argument simple de cardinalité montre qu'à la racine d'une arborescence lourde est attachée une R-assemblée d'arborescences toutes légères sauf *au plus une seule* qui est lourde (voir figure 3). En effet, si au moins deux des arborescences attachées étaient lourdes, la cardinalité de l'ensemble sous-jacent serait $> 2n + 2$ (à cause de la racine), contrairement à l'hypothèse.

Une itération de cette dichotomie jusqu'à épuisement des arborescences lourdes rencontrées en cours de route permet de décomposer *canoniquement* l'arborescence lourde de la façon décrite par la figure 4. Le chemin pointillé contenu dans cette figure permet, en quelque sorte, de "coder" notre décomposition canonique de l'arborescence lourde en arborescences légères. Par simple découpage, on peut résumer la figure 4 à l'aide de la figure 5.

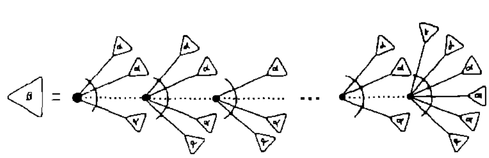

FIGURE 4

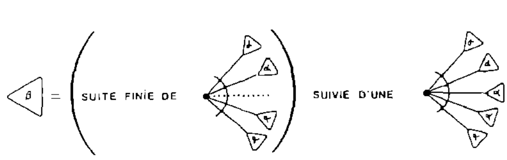

FIGURE 5

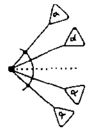

FIGURE 6

En termes plus imagés: sur un ensemble U tel que $n + 1 \leq |U| \leq 2n + 2$, une arborescence lourde équivaut canoniquement à *une suite finie d'arborescences à branches légères ayant chacune une "branche morte" (pointillée)* suivie *d'une arborescence lourde à branches légères.*

Les opérations générales entre espèces permettent de réinterpréter algébriquement les constituants de la figure 5 comme suit:

● L'espèce des *arborescences à branches légères ayant une branche morte* symbolisée génériquement par la figure 6 est isomorphe à l'espèce décrite par la formule

$$XR'(\alpha). \tag{2.6}$$

En effet, cette figure représente un point suivi d'une R'-assemblée de α-structures. On le voit facilement à cause de la présence de la branche morte (segment pointillé) et de la définition de l'espèce dérivée R'.

● L'espèce des *arborescences lourdes à branches légères* symbolisée génériquement par la figure 7 est isomorphe à l'espèce décrite par la formule

$$XR(\alpha) - \alpha. \tag{2.7}$$

La soustraction de l'espèce α dans (2.7) provient du fait qu'une arborescence du genre de celle de la figure 7, bien que de la forme $XR(\alpha)$, *n'est pas* légère étant donné qu'elle est portée par au moins $n + 1$ points (en raison de l'analyse qui nous a conduits successivement aux figures 3, 4, et 5).

Utilisant (2.6) et (2.7) nous pouvons donc écrire

$$\begin{aligned} \beta &\overset{2n+2}{=\!=} (\text{suite finie de } XR'(\alpha)) \cdot (XR(\alpha) - \alpha) \\ &=\!= L(XR'(\alpha)) \cdot (XR(\alpha) - \alpha), \end{aligned} \tag{2.8}$$

où $L(X) = 1/(1 - X)$ désigne l'espèce des ordres linéaires [2].

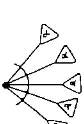

FIGURE 7

Compte tenu de (2.5) nous avons donc obtenu combinatoirement (et sans faire appel à aucune tangente, courbe ou série formelle) la version suivante de l'itération de Newton–Raphson (comparer avec (1.4)).

PROPOSITION 2. Soit L l'espèce des ordres linéaires et soit α l'espèce des arborescences R-enrichies légères (i.e., portées par des cardinalités $\leq n$). Alors l'espèce α^+ définie par

$$\alpha^+ = \alpha + L(XR'(\alpha)) \cdot (XR(\alpha) - \alpha) \tag{2.9}$$

possède un contact d'ordre $2n + 2$ avec l'espèce A des arborescences R-enrichies.

3. CONVERGENCE D'ORDRE SUPÉRIEUR

L'espèce des arborescences lourdes (au sens de la définition 1) nous a permis de reformuler la notion de convergence quadratique et d'aboutir à la version combinatoire (proposition 2) de l'itération de Newton–Raphson. Afin d'obtenir une convergence d'ordre supérieur nous allons d'abord élargir la notion d'arborescence lourde.

DÉFINITION 3. Soient $n \geq 0$ et $k \geq 1$ des entiers fixés. Une arborescence (R-enrichie) sur un ensemble fini U est dite *k-lourde* si

$$n + 1 \leq |U| \leq (k + 1)(n + 1). \tag{3.1}$$

Désignons par β_k l'espèce des arborescences k-lourdes.

Le problème d'approximation d'ordre $(k + 1)$ que nous posons maintenant est celui de décrire une espèce α^+ ayant un contact d'ordre $(k + 1)(n + 1)$ avec l'espèce A des arborescences (R-enrichies), à partir de l'espèce α des arborescences légères.

Etant donné que

$$\alpha + \beta_k \overset{r}{=} A \quad \text{où } r = (k + 1)(n + 1), \tag{3.2}$$

il suffit, comme plus haut, d'exprimer l'espèce β_k à l'aide de l'espèce α. A la racine d'une arborescence k-lourde (générique sur au plus $(k + 1)(n + 1)$ points) sont attachées des arborescences toutes légères sauf, au plus, k qui sont k-lourdes. C'est-à-dire qu'il existe un et un seul i, satisfaisant $0 \leq i \leq k$, pour lequel on ait la situation décrite par la figure 8. Itérant cette analyse jusqu'à épuisement des arborescences k-lourdes rencontrées on obtient génériquement la situation décrite par la figure 9. Cette fois-ci, au lieu d'une simple colonne vertébrale, c'est plutôt une *nervure d'ordre k* (i.e., la

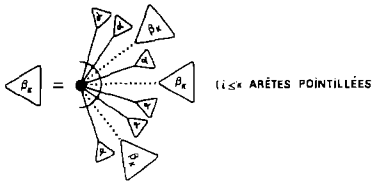

FIGURE 8. (ici $i = 3$).

sous-arborescence pointillée de la figure 9) qui permet de "coder" notre décomposition canonique de l'arborescence k-lourde en arborescences légères. Pour $i \geq 1$ fixé, (et sur un ensemble quelconque) le membre de droite de la figure 8 décrit une structure générique d'espèce

$$XR^{(i)}(\alpha)\beta_k^i/i!. \tag{3.3}$$

Cette formule provient de la définition générale des dérivées successives d'une espèce et du fait que les i segments pointillés de la figure 8 ne sont pas ordonnés (d'où la division par $i!$).

Compte tenu de (2.7) et (3.3) la figure 8 (sommée pour $i = 0, 1, \ldots, k$) montre que l'on a

$$\beta_k \overset{k}{=} \gamma, \quad r = (k + 1)(n + 1), \tag{3.4}$$

où γ est l'espèce définie implicitement par l'équation combinatoire (polynomiale en γ):

$$\gamma = (XR(\alpha) - \alpha) + \sum_{i=1}^{k} \frac{XR^{(i)}(\alpha)}{i!}\gamma^i. \tag{3.5}$$

Nous avons donc obtenu la version combinatoire suivante pour l'itération de Newton–Raphson d'ordre supérieur.

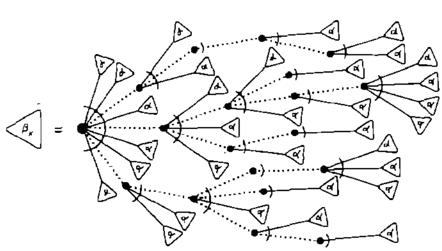

FIGURE 9. (ici $k \geq 7$).

PROPOSITION 4. *Soit α l'espèce des arborescences R-enrichies légères (i.e. portées par des cardinalités $\leq n$) et soit γ l'espèce définie implicitement par l'équation combinatoire polynomiale* (3.5). *Alors l'espèce α^+ définie par $\alpha^+ = \alpha + \gamma$ possède un contact d'ordre $(k+1)(n+1)$ avec l'espèce A des arborescences R-enrichies.*

On remarquera que dans le cas spécial où $k = 1$, on retrouve bien la version (2.9) de l'itération de Newton–Raphson établie dans la section précédente. En effet, l'équation polynomiale (3.5) est alors linéaire en γ et peut se résoudre explicitement.

En guise de retour final à l'Analyse classique (en supposant que $f(t) = xR(t) - t$ où $R(t)$ est une fonction de classe C^{k+1} ou une série formelle), le lecteur vérifiera que la proposition 4 que nous venons de démontrer conduit naturellement à la situation suivante: Si α est une approximation d'une racine a de l'équation $f(t) = 0$ alors on obtient une nouvelle approximation α^+ en annulant le polynôme de Taylor de degré k

$$f(\alpha) + \frac{f'(\alpha)}{1!}(t-\alpha) + \cdots + \frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k \tag{3.6}$$

de la fonction $f(t)$ (développée autour du point $t = \alpha$). De plus, la convergence, lorsqu'elle a lieu, est d'ordre $k+1$ au sens où, pour $\alpha_{n+1} = \alpha_n^+$, $n = 0, 1, 2 \ldots$, on a

$$a - \alpha_{n+1} = 0\left((a - \alpha_n)^{k+1}\right). \tag{3.7}$$

On retrouve évidemment le cas classique de la "tangente" en prenant $k = 1$.

RÉFÉRENCES

1. R. P. BRENT AND H. T. KUNG $0((n \log n)^{3/2})$ algorithms for composition and reversion of power series, *in* "Analytic Computational Complexity," Proceedings of the Symposium on Analytical Computational Complexity, Carnegie-Mellon Univ., Pittsburgh, Penn., 1975, pp. 217–225, Academic Press, New York. 1976.
2. A. JOYAL, Une théorie combinatoire des séries formelles, *Advances in Math.* **42**, No. 1 (1981), 1–82.
3. G. LABELLE, Une nouvelle démonstration combinatoire des formules d'inversion de Lagrange, *Advances in Math.* **42**, No. 3 (1981), 217–247.