Itération de Newton-Schröder-Householder et nouveau nouvel (?) algorithme pour la décomposition de Jordan-Chevalley-Dunford effective

Jean-François Burnol, 10 avril 2017

Table des matières

Ce texte prend la suite de

http://jf.burnol.free.fr/agreg170408NewDun.pdf

Je commence par un rappel, et au final j'introduis ce qui est possiblement (mais j'en doute) une nouvelle nouvelle idée.

1 Approche naïve à la décomposition de Dunford

On se place sur un corps $\mathbf{K}$ (de caractéristique nulle ou suffisamment grande pour certaines des formules qui suivent) et on considère un endomorphisme A d'un espace vectoriel V de dimension finie. Sur $V \otimes_{\mathbf{K}} \overline{\mathbf{K}}$, il y a une décomposition en espaces caractéristiques V_λ. En définissant D comme l'endomorphisme agissant par λ sur V_λ on obtient $A = D + N$, avec $DN = ND$, N nilpotent, D diagonalisable sur $\overline{\mathbf{K}}$. C'est la décomposition de Jordan-Chevalley-Dunford. Elle est en fait définie sur K au sens où $D \in K[A]$ (si la caractéristique est nulle).

Notre objectif est d'obtenir effectivement un tel polynôme. Pour cela soit P le polynôme caractéristique unitaire de A, et $Q = P/(P, P')$ sa version réduite (cette formule demande la caractéristiquen nulle) , et soit n tel que P divise Q^n.

J'explique dans

http://jf.burnol.free.fr/agreg170405Dunford.pdf

que sur la clôture algébrique le problème est résolu par la fraction rationnelle

$$F_N = \frac{\sum_j \frac{\lambda_j}{(X-\lambda_j)^N}}{\sum_j \frac{1}{(X-\lambda_j)^N}}$$

car il est clair que $F_N - \lambda_j$ a en λ_j un zéro de multiplicité au moins N, donc une fois mise sous la forme

$$F_N = \frac{S_N}{R_N} = \frac{\sum_j \lambda_j \prod_{i\neq j}(X-\lambda_i)^N}{\sum_j \prod_{i\neq j}(X-\lambda_i)^N}$$

le dénominateur R_N est un polynôme premier avec Q, et si T_N est son inverse multiplicatif modulo Q^N alors $D_N = T_N R_N$ donne une formule qui exhibe le D de la décomposition de Chevalley-Dunford comme un polynôme en A.

De plus il est à coefficients de K par un argument de polynômes symétriques élémentaires (la caractéristique zéro intervenant ici pour être sûr que Q est dans $\mathbf{K}[X]$ si P l'est ; mais si l'on sait que $Q \in \mathbf{K}[X]$, alors le résultat est vrai sans condition sur la caractéristique), ou plus directement à cause des formules

$$\begin{aligned} F_n &= X + (n-1)\frac{U_{n-1}}{U_n}Q \\ U_{n+1} &= -nQ'U_n + QU_n' \\ U_1 &= Q' \end{aligned}$$

qui cependant ne fonctionnent bien qu'en caractéristique $p \geqslant N$, car

$$U_n = (-1)^{n-1}(n-1)!R_n$$

et donc si $N > p$, les formules de récurrence aboutissent à $U_N = 0$ et on est bloqué. J'ai expliqué cela dans

http://jf.burnol.free.fr/agreg170406Dunford_NewAlgo.pdf

Il reste à la dernière étape à inverser U_N modulo Q^N (ou même seulement modulo Q^{N-1}), par une identité de Bézout, qu'on peut toujours trouver car U_N est premier avec Q.

Comme je l'ai expliqué par la suite dans

http://jf.burnol.free.fr/agreg170408NewDun.pdf

on obtient d'autres fractions rationnelles tout aussi valable pour la décomposition de Dunford avec n'importe quelle valeur initiale U_1, à partir du moment où U_1 est premier avec Q. Le choix $U_1 = 1$ donne les plus petits polynômes U_n.

2 Itérations de Newton, Halley et Schröder

Jusqu'à présent nous avons suivi un chemin vraiment élémentaire et pas subtil. Par contre Chevalley lui a une approche bien plus ingénieuse : *trouver la partie semi-simple D, c'est trouver une racine simple de l'équation $Q(x) = 0$ qui diffère de l'indéterminée X par un multiple de Q (donc par un nilpotent modulo (Q^N).)*

Il utilise l'idée de la méthode de Newton $x - Q(x)/Q'(x)$, $x_0 = X$, mais en travaillant avec des polynômes. Pour cela il choisit un polynôme H qui inverse Q' modulo Q et itère les substitutions $X \leftarrow X - H(X)Q(X)$.

C'est assez dantesque sur les polynômes exacts, mais on travaille bien sûr modulo Q^N et on peut même travailler modulo Q^2 puis Q^4, Q^8, etc... dans les étapes intermédiaires.

Dès que $2^n \geqslant N$, on a notre solution D modulo Q^N.

Il se trouve, comme je l'explique dans

http://jf.burnol.free.fr/agreg170408NewDun.pdf

que l'approche moins subtile de la première section est *aussi liée à l'idée de* Chevalley !

En effet, si l'on choisit $U_1 = 1$, on obtient

$$F_2 = X - \frac{Q}{Q'}$$

et

$$F_3 = X - 2\frac{QQ'}{2(Q')^2 - QQ''}$$

or $x - f/f'$ est la formule de Newton et $x - 2ff'/(2(f')^2 - ff'')$ celle de Halley (1694). Si l'on part suffisamment proche d'un zéro *simple* d'une fonction suffisamment dérivable, les itérations de Newton convergent quadratiquement et celles de Halley cubiquement.

Il apparaît que notre $F_N = X + (N-1)\frac{U_{N-1}Q}{U_N}$ est du même type. Car les récurrences des U_n expriment le fait que

$$U_n = Q^n \left(\frac{d}{dX}\right)^{n-1} \frac{U_1}{Q}$$

donc on a

$$F_N = X + (N-1)\frac{U_{N-1}}{U_N}Q = X + (N-1)\frac{(U_1/Q)^{(N-2)}}{(U_1/Q)^{(N-1)}}$$

or l'on doit semble-t-il à
E. Schröder *Über unendliche viele Algorithmen zur Auflösung der Gleichungen*, Math. Ann., 2 :317-365. (1870)
le fait que la formule

$$x + (p-1)\frac{(1/f)^{(p-2)}}{(1/f)^{(p-1)}}$$

donne un schéma itératif d'ordre $p \geqslant 2$ au voisinage d'un zéro simple d'une fonction suffisamment dérivable, en particulier analytique.

Voir le chapitre 4, section *Iterations of higher order* du livre :

A. S. Householder, *The Numerical Treatment of a Single Nonlinear Equation*, McGraw-Hill, New York, (1970)

Householder souligne dans ce contexte le fait que l'on a parfois avantage à utiliser la variante

$$x + (p-1)\frac{(g/f)^{(p-2)}}{(g/f)^{(p-1)}}$$

avec g quelconque ne s'annulant pas sur l'intervalle où l'on travaille, proche d'une racine simple de f. Si f a un zéro multiple, il est naturel d'y prendre $g = f'$, et alors on retrouve le contexte exact de la première section avec Q y jouant le rôle de f.

On aurait pu y utiliser les multiplicités des valeurs propres λ, et cela se traduirait par $g = P'$, $f = P$ (et $x = X\ldots$) dans la formule de Schröder-Householder. Mais il faut toujours y prendre le p tel que P divise Q^p.

Ainsi j'avais évoqué $U_1 = 1$, $U_1 = Q'$, ou ses puissances modulo P, mais il y a aussi $U_1 = QP'/P$ qui est utilisable comme point de départ des récurrences.

3 Première formule de Schröder

Quel que soit le point de départ, les fractions rationnelles obtenues F_n sont uniques modulo Q^N et il est naturel de les écrire sous forme d'une série de puissances de Q avec

des puissances de Q' au dénominateur. Voici le résultat :

$$F_1 = X$$
$$F_2 = X - \frac{Q}{Q'} \pmod{Q^2}$$
$$F_3 = X - \frac{Q}{Q'} - \frac{Q''Q^2}{2(Q')^3} \pmod{Q^3}$$
$$F_4 = X - \frac{Q}{Q'} - \frac{Q''Q^2}{2(Q')^3} - \frac{(3(Q'')^2 - Q'Q''')Q^3}{6(Q')^5} \pmod{Q^4}$$
$$F_5 = X - \frac{Q}{Q'} - \frac{Q''Q^2}{2(Q')^3} - \frac{(3(Q'')^2 - Q'Q''')Q^3}{6(Q')^5} - \frac{(15(Q'')^3 - 10Q'''Q''Q' + Q^{(4)}(Q')^2)Q^4}{24(Q')^7} \pmod{Q^5}$$

On voit ces développements, dans le contexte des schémas de Newton d'ordres supérieurs dans ce papier de Gourdon-Sebah (2001) :

P. Sebah, X. Gourdon *Newton's method and high order iterations*, October 3, 2001
http://numbers.computation.free.fr/Constants/Algorithms/newton.html

C'est assez laborieux de les obtenir à la main si le point de départ est la formule de Schröder-Householder évoquée précédemment, c'est-à-dire notre formule avec les U_{n-1}/U_n : car il faut développer ce quotient en puissances de Q/Q'...

Une méthode plus rapide pour les calculs à la main est expliquée dans l'article de Sebah-Gourdon, mais on n'y devine peut-être pas le principe général qui est pourtant, *a posteriori*, trivial !

En effet, et cela figure dans le livre de Householder précédemment cité (et qui est en référence de l'article de Sebah-Gourdon) voici une autre contribution de Schröder.

Considérons la fonction analytique f dans un voisinage de r, avec un zéro simple en $z = r$. Soit ψ la fonction inverse avec $\psi(0) = r$. Si z est suffisamment proche de r, et $w = f(z)$, alors $z = \psi(w)$. Si w est suffisamment proche de 0, le développement de Taylor suivant est licite :

$$r = \psi(0) = \psi(w - w) = \psi(w) - w\psi'(w) + \sum_{k=2}^{\infty} (-1)^k \frac{w^k}{k!} \left(\frac{d}{dw}\right)^k \psi(w)$$

Et maintenant en exprimant à nouveau en la variable $z = \psi(w)$, cela donne

$$r = z - f(z)\frac{1}{f'(z)} + \sum_{k=2}^{\infty} (-1)^k \frac{f(z)^k}{k!} \left(\frac{1}{f'(z)}\frac{d}{dz}\right)^k z$$

$$r = z - \frac{f}{f'} - \sum_{k=2}^{\infty} \left(\left(-\frac{1}{f'} \frac{d}{dz} \right)^{k-1} \frac{1}{f'} \right) \frac{f^k}{k!}$$

L'idée est enfantine et peut-être nous l'avons tous eue à un moment ou un autre et nous l'avons oubliée ! La série pour un z donné suffisamment proche de r converge, mais banalement comme peut le faire toute série de Taylor pour une fonction analytique, rien de transcendant là-dedans, mais ce qui est fascinant bien sûr c'est sa propriété auto-améliorante : si l'on tronque à n'importe quel ordre $p \geqslant 2$ (au sens de conserver jusqu'au terme en f^{p-1}) on obtient un schéma itératif d'ordre p.

En effet examinons :

$$F_p(z) = z - \frac{f}{f'} - \sum_{k=2}^{p-1} \left(\left(-\frac{1}{f'} \frac{d}{dz} \right)^{k-1} \frac{1}{f'} \right) \frac{f^k}{k!} = z - \gamma_1 f - \sum_{k=2}^{p-1} \gamma_k \frac{f^k}{k!}$$

Certainement $F_p(r) = r$, considérons $F_p'(z)$:

$$\begin{aligned} F_p'(z) =& 1 - \gamma_1(z) f'(z) \\ & - \gamma_1'(z) f(z) - \gamma_2(z) f(z) f'(z) \\ & - \gamma_2'(z) \frac{f^2(z)}{2!} - \gamma_3(z) \frac{f^2(z)}{2!} f'(z) \\ & - \gamma_3'(z) \frac{f^3(z)}{3!} - \gamma_4(z) \frac{f^3(z)}{3!} f'(z) \\ & \cdots \\ & - \gamma_{p-1}'(z) \frac{f^{p-1}(z)}{(p-1)!} \end{aligned}$$

Or par construction

$$\begin{aligned} 0 &= 1 - \gamma_1(z) f'(z) \\ 0 &= \gamma_1'(z) + \gamma_2(z) f'(z) \\ 0 &= \gamma_2'(z) + \gamma_3(z) f'(z) \\ 0 &= \gamma_3'(z) + \gamma_4(z) f'(z) \\ & \cdots \end{aligned}$$

et ainsi

$$F_p'(z) = -\gamma_{p-1}'(z) \frac{f^{p-1}(z)}{(p-1)!}$$

$$F_p(r+h) = r - \frac{1}{(p-1)!} \int_0^h \gamma_{p-1}'(r+t) \underbrace{f^{p-1}(r+t)}_{=\mathcal{O}(t^{p-1})} dt = r + \mathcal{O}(h^p)$$

Ce qui établit que $z_{n+1} = F_p(z_n)$ est une itération d'ordre au moins p.

En fait toute cette vérification semble un peu superflue car dans la variable $w = f(z)$ c'est trivial et a été construit pour. Mais cela autorise une variante. Supposons que les fonctions $\gamma_1, \gamma_2, \ldots, \gamma_{p-1}$ sont choisies au voisinage de r avec les contraintes

$$\begin{aligned} 1 - \gamma_1(z)f'(z) &= \mathcal{O}(f^{p-1}) \\ \gamma_1'(z) + \gamma_2(z)f'(z) &= \mathcal{O}(f^{p-2}) \\ \gamma_2'(z) + \gamma_3(z)f'(z) &= \mathcal{O}(f^{p-3}) \\ &\cdots \\ \gamma_{p-2}'(z) + \gamma_{p-1}(z)f'(z) &= \mathcal{O}(f) \end{aligned}$$

alors le raisonnement ci-dessus fonctionne.

À propos je ne comprends pas bien un passage du livre de Householder qui semble dire qu'il « suffit que les égalités entre les gamma et leurs dérivées marchent au point $z = r$ », ce qui semble devoir dire que partout on a $\mathcal{O}(f)$ comme dans la dernière condition. J'ai relu plusieurs fois ce passage, et je ne comprends pas ce qu'il veut dire car moi j'ai besoin apparemment des conditions ci-dessus et pas seulement de $\mathcal{O}(f)$.

Bref, une fois que $F_p(r+h) = r + \mathcal{O}(h^p)$ est établi, il en résulte que $f(F_p(z))) = f(r + \mathcal{O}((z-r)^p)) = \mathcal{O}((z-r)^p) = \mathcal{O}(f(z)^p)$. Ce qui en Algèbre va devenir la résolution au problème de congruences équivalent à la construction du D de Chevalley-Dunford.

4 Retour aux polynômes et à Dunford

Les formules de récurrence pour les γ_n sont, dans leur version stricte, $\gamma_1 = \frac{1}{f'}$ et $\gamma_{n+1} = -\frac{1}{f'}\gamma_n'$. Posons

$$\gamma_n = \frac{\alpha_n}{(f')^{2n-1}} \qquad \alpha_1 = 1$$

Alors

$$\frac{\alpha_{n+1}}{(f')^{2n+1}} = -\frac{1}{f'}\left(\alpha_n' \frac{1}{(f')^{2n-1}} - (2n-1)\frac{\alpha_n f''}{(f')^{2n}}\right) = \frac{-f'\alpha_n' + (2n-1)f''\alpha_n}{(f')^{2n+1}}$$

$$\boxed{\alpha_{n+1} = (2n-1)f''\alpha_n - f'\alpha_n'}$$

Nous obtenons :

Théorème. *Soit Q un polynôme à coefficients dans $\mathbf{K}$, à racines simples dans $\overline{\mathbf{K}}$.*

Soit $\alpha_1 = 1$, $\alpha_2 = Q''$, $\alpha_3 = -Q'Q''' + 3(Q'')^2$, et plus généralement

$$n \geqslant 1 \implies \alpha_{n+1} = (2n-1)Q''\alpha_n - Q'\alpha_n'$$

ou de manière équivalente avec $\gamma_n = \alpha_n/(f')^{2n-1}$:

$$\gamma_{n+1} = \frac{\alpha_{n+1}}{(Q')^{2n+1}} = -\frac{1}{Q'}\frac{d}{dX}\left(\frac{\alpha_n}{(Q')^{2n-1}}\right) = -\frac{d\gamma_n}{dQ} \qquad \gamma_1 = \frac{1}{Q'}$$

Alors la fraction rationnelle de Newton-Schröder, *qui est dans $\mathbf{K}(X)$:*

$$F_n = X - \sum_{k=1}^{n-1} \frac{\alpha_k}{(Q')^{2k-1}}\frac{Q^k}{k!} = X - \sum_{k=1}^{n-1} \gamma_k \frac{Q^k}{k!}$$

vérifie $F_n(Q) \equiv 0 \bmod Q^n$, au sens d'avoir en chaque racine de Q dans $\overline{\mathbf{K}}$ un zéro de multiplicité au moins n.

En effet cela traduit des relations polynomiales universelles que nous avons vérifiées sur $\mathbf{C}$ déjà. On peut aussi suivre le schéma de preuve précedent avec le télescopage dans la dérivée F_n.

Corollaire 1. *Soit H vérifiant une identité de Bézout $HQ' + TQ^{n-1} = 1$. On peut le prendre de degré $< (n-1)\deg Q$. Puis définissons* [1]

$$\begin{aligned}
\Gamma_1 &\equiv H && (\bmod\ Q^{n-1})\\
\Gamma_2 &\equiv -H\Gamma_1' \equiv -HH' && (\bmod\ Q^{n-2})\\
\Gamma_3 &\equiv -H\Gamma_2' && (\bmod\ Q^{n-3})\\
&\cdots && \cdots\\
\Gamma_{n-1} &\equiv -H\Gamma_{n-2}' && (\bmod\ Q)
\end{aligned}$$

Alors le polynôme

$$D_n \equiv X - \Gamma_1 Q - \Gamma_2\frac{Q^2}{2} - \Gamma_3\frac{Q^3}{6} - \cdots - \Gamma_{n-1}\frac{Q^{n-1}}{(n-1)!} \pmod{Q^n}$$

est la partie semi-simple de X dans $\mathbf{K}[X]/(Q^n)$.

Dans la pratique on n'est pas obligé de faire les réductions modulaire, mais il semble recommandable de la faire au moins modulo Q^{n-1} tout du long. Cette méthode nécessite de calculer les puissances de Q de toute façon.

1. Remarque : si $T = Q^k Z$, alors $HT' = HQ^k Z' + kHQ'Q^{k-1}Z$ est dans (Q^{k-1}).

Je n'ai pas cherché à comparer l'algorithme de Chevalley, celui de

http://jf.burnol.free.fr/agreg170406Dunford_NewAlgo.pdf

(plutôt à faire avec $U_1 = 1$ d'ailleurs) et celui-ci.

Il n'est pas évident que Chevalley ait vraiment l'avantage : sa formule de substitution semble signifier des calculs beaucoup plus compliqués. Il y a moins d'étapes mais elles coûtent de plus en plus cher, et je rappelle que $1 + 2 + 4 + \cdots + 2^k = 2^{k+1} - 1$.

Addendum (12 avril) : *on trouvera au bas de la page 14 l'explication que l'algorithme de* Chevalley *est semble-t-il en* $q^3n^2\log n$ *(*$q = \deg Q$*), s'il est bien implémenté, tandis que celui de cette fiche précédente, et celui de la section suivante sont en* q^2n^2 *(avec des multiplications, divisions, Euclide quadratiques en les degrés). Par contre celui du Corollaire précédent semble être cubique en* n *et serait donc le pire de tous.*

Ici on ne fait à chaque étape que des multiplications de polynômes, des dérivations et des divisions euclidiennes. L'algorithme de

http://jf.burnol.free.fr/agreg170406Dunford_NewAlgo.pdf

semble pas mal, comme ça à vue de nez, car on a des polynômes qui sont petits à chaque étape et ne nécessite aucune réduction (et faut y prendre $U_1 = 1$ ..., bon j'ai déjà dit). Il n'y a qu'une inversion à faire à la fin.

Celui du théorème ci-dessus semble moins bon. Mais bon, pas testé dans la pratique. Et puis surtout il y a une bien meilleure chose à faire. Nous y venons maintenant.

5 Une nouvelle (?) idée, et un nouveau nouvel (?) algorithme pour Dunford

Je reviens à l'idée « enfantine »

$$r = \psi(0) = \psi(w - w) = \psi(w) - w\psi'(w) + \dots$$

Je suis un peu gêné par la précision qu'il faut ensuite dans les γ_1, γ_2, etc..., comme nous l'avons vu précédemment, et sauf erreur de ma part. Cela a donné le résultat un peu alambiqué de la section précédente.

Donc essayons de mettre en place une récurrence plus astucieuse. Au départ $\psi(w) - r$ a un zéro simple en $w = 0$. La fonction $G(w) = \psi(w) - w\psi'(w)$ est $r + c_2w^2 + \dots$. Clairement il faut passer à $K(w) = G(w) - \frac{1}{2}wG'(w) = r + \mathcal{O}(w^3)$. Puis on passe à $K(w) - \frac{1}{3}wK'(w)$.

Cela veut donc dire qu'on pose $\psi_1(w) = \psi(w)$, puis $\psi_{n+1}(w) = \psi_n(w) - \frac{1}{n}w\psi_n'(w)$, $n \geqslant 1$.

Théorème. *On a* $\forall n \geqslant 1 \quad \psi_n(w) = r + \mathcal{O}(w^n)$.

Preuve. Déjà faite. □

Traduisons cela dans la variable z. Mais attention, je ne prends pas comme définition bien sûr que $f(F_n(z)) = \psi_n(f(z))$ qui n'aurait pas du tout de sens (je rappelle que $\psi_n(0) =$

$\psi(0) = r$), ni même $\psi_n(f(z)) - r$. Non, je prends $F_n(z) = \psi_n(f(z))$. Ainsi $F_{n+1}(z) = F_n(z) - \frac{1}{n} f(z) \frac{1}{f'(z)} \frac{d}{dz} F_n(z)$.

En utilisant une notation opératorielle, (avec 1 pour l'opérateur identité) :

$$\begin{aligned} F_1 &= z \\ F_2 &= z - \frac{f}{f'} = (1 - f\frac{d}{df})F_1 \\ F_3 &= (1 - \frac{1}{2} f\frac{d}{df})F_2 = (1 - \frac{1}{2} f\frac{d}{df})(1 - f\frac{d}{df}) \cdot z \\ &\cdots \\ F_n &= \frac{1}{(n-1)!}(n-1-N)(n-2-N)\cdots(1-N)\cdot z \end{aligned}$$

avec N l'opérateur $f\frac{d}{df} = f/f'd/dx$.

C'est le « number operator » au sens où $N(f^n) = nf^n$ (mais je ne peux plus utiliser N pour un entier maintenant.) L'avantage de la formulation opératorielle est que l'on peut calculer le produit dans n'importe quel ordre. Ainsi :

$$F_n = (-1)^{n-1} \frac{1}{(n-1)!}(N-1)(N-2)\cdots(N-n+1)z = (-1)^{n-1}\binom{N-1}{n-1} \cdot z$$

Notons que l'opérateur définissant F_n annule f, f^2, …, f^{n-1}.

Théorème.

$$F_n(r+h) = r + \mathcal{O}(h^n)$$

et par conséquent $z_{k+1} = F_n(z_k)$ définit une itération vers r d'ordre au moins n.

Preuve. Par définition $F_n(r+h) = \psi_n(f(r+h)) = \psi_n(w)$ et par construction $\psi_n(w) = r + \mathcal{O}(w^n) = r + \mathcal{O}(h^n)$. □

Le grand avantage c'est que si l'on modifie F_2 par quelque chose d'ordre deux, alors F_3 ne bouge qu'à l'ordre trois etc... C'est la même stabilité qu'on avait dans la discussion des U_n de ma fiche précédente, et qui avait été perdue avec les γ_j de la section précédente.

Revenons à la variable w et à la fonction ψ_n. Elle est obtenue à partir d'une certaine fonction analytique $\psi(w)$ par l'action d'un opérateur linéaire qui transforme w^k en $1/(n-1)!(n-1-k)(n-2-k)\cdots(1-k)w^k$, c'est donc tout simplement l'opérateur

$(-1)^{n-1}/(n-1)!w^n(d/dw)^{n-1}w^{-1}$. Notons que cet opérateur laisse invariante la constante r, et annule w, w^2, …, w^{n-1}.

Si l'on revient à la variable $z = \psi(w)$, $w = f(z)$, notre fonction $F_n(z)$ est donc tout bêtement

$$F_n = \frac{(-1)^{n-1}}{(n-1)!} f^n (d/df)^{n-1} \left(\frac{z}{f}\right)$$

Théorème. *Soit f une fonction analytique (ou ne différant d'une telle fonction que par une fonction plate en r) ayant un zéro simple en $z = r$.*

Soit N l'opérateur $f\frac{d}{df} = \frac{f}{f'}\frac{d}{dz}$.

Soit $n \geqslant 2$. On pose

$$F_n = (1 - \frac{1}{n-1}N)\cdots(1 - \frac{1}{2}N)(1-N)\cdot z = \frac{(-1)^{n-1}}{(n-1)!} f^n (\frac{1}{f'}\frac{d}{dz})^{n-1} \left(\frac{z}{f}\right)$$

et aussi on conviendra que $F_1 = z$. On a $F_2 = z - f/f'$.

Pour tout $n \geqslant 2$, l'itération $z_{k+1} = F_n(z_k)$ est un schéma de Newton-Halley *d'ordre au moins n, autrement dit $F_n(z) - r$ possède en $z = r$ un zéro d'ordre au moins n. De plus si dans la construction itérative, on modifie F_k en lui ajoutant une fonction analytique ayant aussi un zéro d'ordre au moins k en $z = r$ (c'est-à-dire une fonction du type $f^k g$) alors F_{k+1} aura encore un zéro d'ordre au moins $k+1$ en $z = r$ et ainsi de suite.*

Lorsque $f = Q$ est une fraction rationnelle dans $\mathbf{K}(X)$ possèdant des racines simples dans une clôture algébrique,

$$F_n = \frac{(-1)^{n-1}}{(n-1)!} Q^n (\frac{1}{Q'}\frac{d}{dX})^{n-1} \left(\frac{X}{Q}\right)$$

est une fraction rationnelle qui est définie aux zéros λ_j de Q, et est telle que $Q(F_n)$ possède en λ_j un zéro de multiplicité au moins n, autrement dit F_n donne modulo Q^n la partie semi-simple de X.

On obtient de manière commode une solution polynomiale en choisissant tout d'abord un polynôme H représentant un inverse multiplicatif de Q' modulo Q (et pas nécessairement modulo Q^n ; on prendra en général l'unique H vérifiant $\deg H < \deg Q$). Puis on définit le polynôme D_n par la formule :

$$D_n = (1 - \frac{1}{n-1}HQ\frac{d}{dX})\cdots(1 - \frac{1}{2}HQ\frac{d}{dX})(1 - HQ\frac{d}{dX})\cdot X$$

qui peut s'évaluer itérativement, avec éventuellement des réductions modulo Q^k pour chaque D_k. Alors D_n donne la partie semi-simple de l'endomorphisme de multiplication par X modulo Q^n.

Dans l'ordre itératif, avec $D_1 = X$,

$$D_{n+1} = (1 - \frac{1}{n} HQ \frac{d}{dX}) D_n$$

on peut modifier D_n par un élément de (Q^n), ceci ne modifiera D_{n+1} que modulo (Q^{n+1}).

Si l'on prend pour H l'unique polynôme de degré $< \deg Q$ inverse multiplicatif de Q', il sera en général de degré $\deg Q - 1$, donc on voit qu'en général $\deg D_{n+1} = 2\deg Q - 2 + \deg D_n$ (on suppose $\deg Q > 1$, car si $Q = a + bX$, on aura $H = b^{-1}$, $D_2 = X - b^{-1}(a + bX) = -b^{-1}a$, et tous les $D_n = D_2$), donc $\deg D_n = (2\deg Q - 2)n - 2\deg Q + 3$, si l'on procède sans réduction modulo Q^n.

Preuve. Je vais juste expliquer l'histoire du H, tout le reste ayant déjà été fait, essentiellement. Tout d'abord nous avons nos fractions rationnelles.

$$F_n = (1 - \frac{1}{n-1} \frac{Q}{Q'} \frac{d}{dX}) \cdots (1 - \frac{1}{2} \frac{Q}{Q'} \frac{d}{dX})(1 - \frac{Q}{Q'} \frac{d}{dX}) \cdot X$$

provenant directement de notre schéma de Newton d'ordre supérieur. Montrons par récurrence : *pour tout $n \geqslant 1$, $F_n - D_n$ est de la forme $Q^n Z$ avec une fraction rationnelle Z définie en chacune des racines λ_j de Q dans la clôture algébrique $\overline{\mathbf{K}}$.*

C'est vrai pour $n = 1$, puisque $F_1 = D_1 = X$. Supposons-le vrai pour n. Alors

$$\begin{aligned} F_{n+1} &= (1 - \frac{1}{n} \frac{Q}{Q'} \frac{d}{dX})(D_n + Q^n Z) \\ &= (1 - \frac{1}{n} \frac{Q}{Q'} \frac{d}{dX}) D_n + Q^n Z - \frac{1}{n} \frac{Q}{Q'} (n Q^{n-1} Q' Z + Q^n Z') \\ &= D_{n+1} - \frac{1}{n} (\frac{1}{Q'} - H) Q D_n' - \frac{1}{n} Q^{n+1} \frac{Z'}{Q'} \end{aligned}$$

Or $\frac{1}{Q'} - H = \frac{1 - HQ'}{Q'}$ et par définition de H, $1 - HQ' = QT$ pour un certain T. Donc il ne reste plus qu'à examiner une contribution du type $Q^2 D_n'$. Or, nous savons que pour tout j, $F_n - \lambda_j$ s'annule à l'ordre n au moins en $X = \lambda_j$. Par conséquent et compte tenu de l'hypothèse de récurrence c'est aussi le cas de $D_n - \lambda_j$. Donc, la dérivée D_n' s'annule au moins à l'ordre $n - 1$ (je suppose $n \geqslant 2$), et lorsque l'on multiplie par Q^2, cela donne un zéro d'ordre au moins $n + 1$. Et si $n = 1$, alors $D_1 = X$, et $Q^2 D_1' = Q^2$ a bien un zéro d'ordre 2 en chaque λ_j. Il est donc établi que $Q^2 D_n'$, qui est un polynôme, est divisible par Q^{n+1} dans $\overline{\mathbf{K}}[X]$, et donc aussi dans $\mathbf{K}[X]$. Ce qui achève la preuve de la validité de $F_{n+1} \equiv D_{n+1} \bmod Q^{n+1}$ dans $\mathbf{K}(X)$.

Le théorème est démontré. □

Voici à nouveau une table des premiers F_n.

$$F_n = (1 - \frac{1}{n-1}\frac{Q}{Q'}\frac{d}{dX})\cdots(1 - \frac{1}{2}\frac{Q}{Q'}\frac{d}{dX})(1 - \frac{Q}{Q'}\frac{d}{dX})\cdot X$$

$$F_1 = X$$
$$F_2 = X - \frac{Q}{Q'}$$
$$F_3 = X - \frac{Q}{Q'} - \frac{Q''Q^2}{2(Q')^3}$$
$$F_4 = X - \frac{Q}{Q'} - \frac{Q''Q^2}{2(Q')^3} - \frac{(3(Q'')^2 - Q'''Q')Q^3}{6(Q')^5}$$
$$F_5 = F_4 - \frac{15(Q'')^3 - 10Q'Q''Q''' + Q^{(4)}(Q')^2}{(Q')^7}\cdot\frac{Q^4}{24}$$
$$F_6 = F_5 - \frac{105(Q'')^4 - 105Q'(Q'')^2Q''' + 10(Q')^2(Q''')^2 + 15(Q')^2Q''Q^{(4)} - (Q')^3Q^{(5)}}{(Q')^9}\cdot\frac{Q^5}{120}$$

Terminons sur la démonstration que le développement en puissances de Q est bien celui donné par ce que j'ai appelé la première formule de Schröder :
Théorème. *Les F_n ainsi définies vérifient :*

$$F_{n+1} = F_n - \left(\left(-\frac{1}{Q'}\frac{d}{dX}\right)^{n-1}\frac{1}{Q'}\right)\frac{Q^n}{n!}$$

Preuve. On a $F_{n+1} = F_n - \frac{1}{n}\frac{Q}{Q'}\frac{d}{dX}F_n$ et de plus nous savons que F_n s'exprime sous la forme

$$F_n = \frac{(-1)^{n-1}}{(n-1)!}Q^n(\frac{1}{Q'}\frac{d}{dX})^{n-1}\left(\frac{X}{Q}\right)$$

donc

$$F_{n+1} = F_n - \frac{(-1)^{n-1}}{n!}\frac{Q}{Q'}\frac{d}{dX}Q^n(\frac{1}{Q'}\frac{d}{dX})^{n-1}\left(\frac{1}{Q}\cdot X\right)$$

Il faut maintenant observer que F_n est un certain polynôme en l'opérateur $Q\frac{d}{dQ}$ agissant sur X, et que cette action commute avec celle de l'opérateur $Q\frac{d}{dQ}$, donc, en permutant,

$$F_{n+1} = F_n - \frac{(-1)^{n-1}}{n!}Q^n(\frac{1}{Q'}\frac{d}{dX})^{n-1}\left(\frac{1}{Q}\frac{Q}{Q'}\frac{d}{dX}X\right) = F_n - \frac{Q^n}{n!}(-\frac{1}{Q'}\frac{d}{dX})^{n-1}\frac{1}{Q'}$$

ce qu'il fallait démontrer. □

Ainsi, l'approche de cette section redonne les mêmes fractions rationnelles que le développement obtenu à partir de l'idée $\psi(0) = \psi(w - w)$, mais la formule itérative est plus stable car elle permet de remplacer $\frac{1}{Q'}$ par une approximation modulo Q seulement.

Ajout le 11 avril 2017

On peut extraire des raisonnements précédents la conclusion suivante. La formule :

$$D_n = (1 - \frac{1}{n-1} H_{n-1} Q \frac{d}{dX}) \cdots (1 - \frac{1}{2} H_2 Q \frac{d}{dX})(1 - H_1 Q \frac{d}{dX}) \cdot X$$

donne modulo Q^n la partie semi-simple de X sous la seule condition que $\forall k \quad H_k Q' \equiv 1 \bmod Q$. Encore plus généralement :

Théorème. *Soit $Q \in \mathbf{K}[X]$ un polynôme à racines simples dans $\overline{\mathbf{K}}$. Soit D_1 un polynôme quelconque et soit D_n des polynômes vérifiant une récurrence ;*

$$D_{n+1} = (1 - \frac{1}{n} H_n Q \frac{d}{dX})(D_n + Q^n Z_n)$$

avec des $Z_n \in \mathbf{K}[X]$ arbitraires et les $H_n \in \mathbf{K}[X]$ tels que $H_n Q' \equiv 1 \bmod Q$. Alors, pour tout $n \geqslant 1$, $D_n \in \mathbf{K}[X]$ donne modulo Q^n la partie semi-simple de D_1.

Preuve. Soient $\lambda_1, \ldots, \lambda_k$ les racines de Q dans $\overline{\mathbf{K}}$. Il s'agit de prouver les congruences $D_n \equiv D_1(\lambda_j) \bmod (X - \lambda_j)^n$. C'est vrai pour $n = 1$. Supposons le vrai pour n et montrons le pour $n+1$. Fixons $1 \leqslant j \leqslant k$. Alors $D_n - D_1(\lambda_j) \in (X - \lambda_j)^n \overline{\mathbf{K}}[X]$. Et $D_n + Q^n Z_n$ est de la forme $D_1(\lambda_j) + (X - \lambda_j)^n T_n$. Nous voyons donc que

$$\begin{aligned} D_{n+1} &= D_1(\lambda_j) + (X - \lambda_j)^n T_n - \frac{1}{n} H_n Q \cdot \Big((X - \lambda_j)^n T_n \Big)' \\ &= D_1(\lambda_j) + (X - \lambda_j)^n \left(1 - H_n \frac{Q}{(X - \lambda_j)} \right) T_n - \frac{1}{n} H_n Q \cdot (X - \lambda_j)^n T_n' \end{aligned}$$

Or $1 - H_n Q'(\lambda_j)$ (et $H_n Q$) s'annulent en λ_j, et comme $Q'(\lambda_j) - \frac{Q}{X - \lambda_j}$ également il en résulte que $D_{n+1} \in D_1(\lambda_j) + (X - \lambda_j)^{n+1} \overline{\mathbf{K}}[X]$, et c'est ce qu'il fallait établir. □

On peut traduire la même idée dans le contexte des itérations de Newton d'ordres supérieurs. Il suffit de s'assurer que F_1 est de la forme $z + f g_1$, $F_2 = F_1 - \frac{f}{f'} F_1' + f^2 g_2$, $F_3 = F_2 - \frac{1}{2} \frac{f}{f'} F_2' + f^3 g_3$ etc… alors les itérations $z_{k+1} = F_n(z_k)$ sont d'ordre au moins n. Mais en fait, je crois avoir plus ou moins déjà dit cela…

Au final, j'ai la sourde impression d'avoir déjà fait tout ceci il y a trente ans comme exercice lorsque j'apprenais l'algèbre commutative !

J'ai un vague raisonnement douteux que si les multiplications de polynômes sont faites naïvement (donc avec un coût quadratique en le degré), alors le coût de mon algorithme comme ci-dessus est quadratique en le N final de Q^N, car chaque itération a un coût linéaire ; rappelons qu'on peut faire la réduction modulo Q^N tout à la fin, car comme je l'ai déjà indiqué le degré final est linéaire en N et donc cette dernière réduction aura aussi un coût quadratique en N (et en $\deg Q$, voir paragraphe suivant). Je ne tiens pas compte ici de la taille des coefficients qui sont supposés être traités comme des flottants ou des entiers modulo quelque chose. Par contre la méthode de Chevalley semble pire malgré son nombre d'étapes en $\log N$. Suffit de regarder la dernière itération et toutes les multiplications de polynômes qu'il faudra alors y effectuer pour s'en convaincre…

… ou pas : ce n'est peut-être pas si catastrophique pour Chevalley. Soit $q = \deg Q$. La dernière étape nous demande d'évaluer un polynôme de degré (a priori) $2q - 1$ en un polynôme de degré au pire $qN - 1$ (si l'on a fait les réductions modulaires). Chaque multiplication suivi d'une réduction modulo Q^N va coûter $\mathcal{O}(q^2 N^2)$. Bien sûr si on néglige de réduire modulo Q^N c'est plus cher. Par un schéma de Horner pour évaluer $HQ(T)$, on peut donc considérer qu'il y a $\mathcal{O}(q)$ opérations coûtant chacune $\mathcal{O}(q^2 N^2)$, donc au total $\mathcal{O}(q^3 N^2)$ (qui englobe le calcul initial de H et de HQ). Et il faut rajouter un facteur $\log N$. Par comparaison pour l'algorithme de cette dernière section, je vois un coût de $\mathcal{O}(qN^2)$ pour les récurrences mais la réduction modulo Q^N semble devoir coûter $\mathcal{O}(q^2 N^2)$ (remarque : il ne faut **pas** faire de réductions intermédiaires modulo les Q^k, car les degrés n'augmentent que linéairement de toute façon). Donc, sauf erreur d'analyse, la comparaison montre une différence par rapport à $\deg Q$ mais pas tant en ce qui concerne N. Si l'on s'y prend mal (ne pas réduire modulo Q^N pour Chevalley, ou ne pas y utiliser un schéma de Horner pour HQ) la comparaison peut en être modifiée. Avantage tout de même aux récurrences de cette fiche, ou de l'autre avec le U_{N-1}/U_N, par rapport à Chevalley.

Addendum (12 avril) : on peut implémenter la formule $D_n = (1 - \frac{1}{n-1} HQ \frac{d}{dX}) \ldots (1 - HQ \frac{d}{dX}) X$ de manière à calculer les « chiffres » en base Q : $D_n = X - HQ + \gamma_2 Q^2 - \cdots + (-1)^{n-1} \gamma_{n-1} Q^{n-1}$ d'une manière efficace. On a $\deg H < \deg Q$ et la récurrence est $D_{n+1} = D_n - (HQD_n')/n$. Supposons qu'on a D_n avec $\deg D_n < n \deg Q$. On sait que D_n' est nul modulo Q^{n-1}. Cela veut dire que la division euclidienne de D_n' (de degré au plus $n \deg Q - 2$) par Q^{n-1} donne un résultat exact de degré au plus $\deg Q - 2$. On n'a donc besoin que de conserver les $\deg Q - 1$ coefficients dominants de D_n' (plus précisément les monômes de degrés allant de $(n-1)q$ à $nq - 2$ pour D_n', ceux parmi eux de plus hauts degrés pouvant être nuls ; on a noté $q = \deg Q$) et de Q^{n-1} (disons K_n) et de calculer le quotient monôme par monôme (en laissant tomber à chaque fois le terme dans K_n de plus bas degré). Cela donnera à un coût $\mathcal{O}(q^2)$ un polynôme A_n de degré au plus $\deg Q - 2$ tel que $D_n' - A_n Q^{n-1}$ est de degré $< (n-1) \deg Q$, donc nul, sans avoir à le vérifier ! On fait alors le produit HA_n, puis la réduction modulo Q, ce qui coûte à nouveau $\mathcal{O}(q^2)$ et donne le γ_n. À l'étape suivante il nous faudra les $q - 1$ coefficients dominants de Q^n, leur obtention coûte $\mathcal{O}(q^2)$, et on aura à calculer les coefficients de X^{nq+1} à X^{nq+q-1} dans le produit $\gamma_n Q^n$ or on connaît γ_n et les $q - 1$ coefficients dominants de Q^n, donc ici aussi $\mathcal{O}(q^2)$. Ainsi, en se focalisant sur la seule connaissance des $\deg Q - 1$ coefficients dominants des D_k et des Q^{k-1}, mis à jour de manière itérative, on obtient les chiffres de D_n en base Q pour seulement $\mathcal{O}(q^2 n)$. Bon, je sens que je vais devoir faire une nouvelle fiche avec un algorithme fonctionnant mod Q uniquement…