

ACCELERATION DE LA CONVERGENCE EN ANALYSE NUMERIQUE.

- I. Théorie des récurrences.**
- II. Leçons sur les fractions continues.**
- III. Leçons sur les approximants de Padé.**
- IV. Les algorithmes d'accélération de la
convergence des suites.**
- V. Applications diverses.**

**Cours proposé par André Hautot aux étudiants de dernière année d'études de physique à
l'Université de Liège.**

Plan du cours.

Ce cours s'articule en cinq chapitres :

1. Etude des récurrences ou équations aux différences finies. Un grand nombre d'algorithmes numériques sont récurrents. Leurs propriétés se déduisent de la théorie des récurrences. Cette théorie fait l'objet du premier chapitre.
2. Etude des fractions continues, simples et généralisées. Celles-ci constituent une méthode classique quoique trop peu connue, d'accélération de la convergence.
3. Etude des méthodes classiques de triangles. Les algorithmes de triangles sont en fait des méthodes sommatoires dont certaines, assez anciennes, remontent à Euler, Newton et MacLaurin. D'autres (Borel, Cesaro), plus récentes, datent du début de ce siècle. Ces méthodes sont basées sur un ensemble de schémas **linéaires** de construction des suites $\{t_i\}$ à partir des suites $\{s_i\}$ à accélérer. Ils sont relativement peu efficaces. Pourtant, il est particulièrement intéressant d'analyser les raisons de cette relative inefficacité car cela permet de mettre en oeuvre les améliorations possibles.
4. Etude des algorithmes récents E, ε , ρ , Δ^2 , θ , Levin, GBW, Steffensen. Ceux-ci sont **non linéaires**. Ils tiennent d'ailleurs leur efficacité de cette non-linéarité. Une attention particulière est accordée à l'algorithme ε particulièrement utile aux physiciens.
5. Etude de quelques exemples qui montrent l'intérêt des méthodes non linéaires en analyse numérique (calculs de racines, de valeurs propres, d'intégrales, de fonctions, etc...).

Afin de ne pas alourdir l'exposé, nous admettrons certains théorèmes, sans démonstration. Les nombreux exemples exposés permettent d'ailleurs d'en comprendre très rapidement l'énoncé avec un maximum d'efficacité.

Références.

Voici quelques livres de références incontournables qui parlent des sujets classiques abordés dans ce cours. Les sujets plus pointus ne sont couverts que par la littérature spécialisée.

W. B. Jones and W. J. Thron : Continued fractions. Encyclopedia of Math. and its Applications. Vol. 11. Addison-Wesley (1980).

G. A. Baker Jr and P. Graves-Morris : Pade approximants. Ibidem, Vol. 13&14 (1981).

C. Brezinski : Accélération de la convergence numérique. Lecture Notes in Math. Vol 584. Springer-Verlag, Berlin (1977).

J. Wimp : Sequence transformations and their applications. Math in Science & Engineering. Vol. 514. Academic Press, N-Y (1984).

J. Wimp : Computations with recurrence relations. Pitman, Boston (1984).

Introduction.

Un algorithme peut être considéré comme un ensemble ordonné d'opérations arithmétiques et logiques qui transforment un ensemble de données en un ensemble de résultats.

Données → Algorithme → Résultats

Un grand nombre d'algorithmes sont itératifs. Cela signifie qu'on calcule les résultats par approximations successives, engendrant à chaque itération un terme s_i , de la suite, scalaire ou vectorielle, $\{s_0, s_1, s_2, \dots\}$, qui converge vers le résultat cherché s_∞ , limite de la suite. Ce schéma, simple en théorie, soulève en fait deux questions essentielles :

- Le résultat final est-il fiable ?
- Est-il atteint après un temps de calcul raisonnable ?

1. Fiabilité des résultats.

Les données étant encodées avec n chiffres significatifs (donc à la précision 10^{-n}), le problème se pose tout naturellement de savoir si on est en droit d'espérer que les résultats seront également précis avec n chiffres. En fait, il n'en est rien pour deux raisons :

a) Certains **problèmes** sont mathématiquement mal conditionnés. Considérons l'exemple très simple du système linéaire de deux équations à deux inconnues :

$$\begin{cases} x + 3y = 14 \\ 3.5x + 10.501y = 49.004 \end{cases}$$

dont la solution exacte vaut : $x = 2$ et $y = 4$. Si on perturbe légèrement le second membre en y remplaçant le nombre 49.004 par 49.001, on constate que la solution s'en trouve modifiée radicalement. Elle vaut, cette fois : $x = 11$ et $y = 1$! On sait qu'un ordinateur encode généralement les données d'un problème en commettant une erreur de troncature correspondant au premier chiffre négligé. Un problème mal conditionné présente une grande sensibilité à toute erreur d'encodage des données en sorte que les résultats sont entachés d'une erreur amplifiée parfois considérablement. Dans l'exemple traité, la cause de l'erreur réside dans la valeur proche de zéro du déterminant de la matrice du système.

Il en résulte que la matrice inverse possède de grands coefficients qui amplifient la moindre incertitude sur la valeur des termes du second membre. En effet, on a :

$$A\vec{x} = \vec{b} \Rightarrow \vec{x} = A^{-1}\vec{b}$$

$$A\vec{x}^* = \vec{b} + \vec{\delta b} \Rightarrow \vec{x}^* = A^{-1}\vec{b} + A^{-1}\vec{\delta b} \quad d'o\grave{u}$$

$$\vec{\delta x} = \vec{x}^* - \vec{x} = A^{-1}\vec{\delta b}$$

Cette erreur cesse d'être négligeable dès que $\text{dtm}(A)$ est proche de zéro. Dans l'exemple choisi, on a :

$$\vec{\delta b} = \begin{pmatrix} 0 \\ -0.003 \end{pmatrix} \quad d'o\grave{u}: \quad \vec{\delta x} = \begin{pmatrix} 10501 & -3000 \\ -3500 & 1000 \end{pmatrix} \begin{pmatrix} 0 \\ -0.003 \end{pmatrix} = \begin{pmatrix} 9 \\ -3 \end{pmatrix}$$

Un autre exemple classique est celui de la matrice $n \times n$ de Hilbert $H_{ij} = (1/(i+j))$ dont les valeurs propres, et par conséquent aussi le déterminant, sont tous proches de zéro dès que n est grand. Aucun algorithme travaillant avec un nombre limité de chiffres significatifs ne peut éviter l'écueil que représente le mauvais conditionnement d'un problème. Sauf si l'on peut recourir à un logiciel, tel que Mathematica, qui travaille en arithmétique exacte, la seule parade consiste à augmenter, parfois de façon considérable, le nombre des chiffres significatifs encodés en recourant aux précisions doubles (16 chiffres) ou quadruples (33 chiffres).

Les systèmes chaotiques, en physique, sont, eux aussi, des systèmes mal conditionnés : ils présentent la sensibilité aux conditions initiales qui rend toute prédiction de l'évolution du système impossible au-delà d'un temps parfois très court.

b) Certains **algorithmes** sont instables.

On se gardera de confondre cette situation avec celle décrite au point a) ci-dessus. Cette fois, le problème mathématique posé est bien conditionné mais c'est l'algorithme mis en oeuvre pour le résoudre qui ne convient pas car il est instable.

On rappelle que le calcul par ordinateur engendre inévitablement deux sources d'erreurs :

- les erreurs de troncature qui résultent de l'arithmétique finie du processeur qui l'amène à remplacer la fraction $1/3$ par 0.33333333 , par exemple.
- les erreurs d'arrondi qui proviennent de la soustraction de quantités voisines avec la perte de chiffres significatifs que cela implique.

Que deviennent ces erreurs au cours d'un calcul un peu long, comportant des milliers d'opérations arithmétiques élémentaires ? En d'autres termes, comment se propagent les erreurs de troncature et d'arrondi ? On sait que le processeur arrondit par excès ou par défaut chaque résultat intermédiaire en sorte que, naïvement, on pourrait penser qu'au terme d'un calcul comportant un très grand nombre, N , d'opérations élémentaires, la compensation est à peu près exacte. Statistiquement, chacun sait qu'il n'en est rien et que ce type d'erreur croît lentement comme \sqrt{N} . Il est plutôt rare cependant que cette erreur altère le résultat de façon grossière. L'erreur qu'introduit un algorithme instable est beaucoup plus grave que cela. Elle peut croître selon une loi $f(N)$, exponentielle, par exemple, beaucoup plus rapidement croissante que \sqrt{N} . Nous verrons des exemples où une erreur de troncature initiale explose littéralement en cours de calcul au point que plus aucun chiffre calculé n'est correct dans le résultat final.

Lorsqu'un algorithme présente une telle instabilité, il est hors de question de l'utiliser et il faut en chercher un autre qui effectue le même travail de manière plus stable. Si cet algorithme n'est pas connu, la seule ressource est, à nouveau, de travailler avec une réserve accrue de chiffres significatifs.

En résumé, on distinguera :

- l'instabilité intrinsèque d'un problème liée à son mauvais conditionnement mathématique qu'aucun algorithme ne peut lever;
- l'instabilité extrinsèque, liée au choix d'un mauvais algorithme pour traiter un problème qui, par ailleurs, est mathématiquement bien conditionné.

Remarques pratiques : En l'absence d'erreurs, entraînant l'arrêt du programme (division par zéro, overflow, underflow, ...) un ordinateur poursuit son calcul jusqu'à son terme et affiche un résultat sans se préoccuper d'aberrations éventuelles liées au mauvais conditionnement du problème ou à l'instabilité de l'algorithme utilisé. Si on ne dispose d'aucune procédure de vérification destinée à s'assurer du caractère plausible du résultat, il est prudent de procéder à quelques essais numériques qui auront au moins le mérite de mettre en évidence d'éventuels problèmes de précision:

- En modifiant légèrement les conditions initiales (données) tout en maintenant la précision de l'ordinateur constante, on pourra détecter si le problème présente la sensibilité aux conditions initiales. Ce sera le cas toutes les fois que la solution perturbée s'écartera progressivement et parfois même violemment de la solution non perturbée.
 - En maintenant les conditions initiales mais en augmentant le nombre de chiffres significatifs utilisés au cours des calculs, on s'assurera du degré d'instabilité induit par l'algorithme lui-même en comparant le nombre de chiffres altérés dans les résultats.
- On agira en conséquence.

2. Problèmes liés à la vitesse de convergence de l'algorithme.

Il n'est pas toujours facile d'estimer le temps qui sera nécessaire à un ordinateur pour effectuer un programme donné. Cela dépend en effet de la rapidité avec laquelle le processus itératif mis en oeuvre converge vers sa limite. Lorsque ce temps devient prohibitif, il est déraisonnable de poursuivre les calculs sur base d'une vitesse de convergence aussi faible. Une stratégie plus efficace consiste à rechercher par voie théorique dans quelle mesure il serait possible de remplacer la suite $\{s_0, s_1, s_2, \dots\}$ lentement convergente, par une autre suite $\{t_0, t_1, t_2, \dots\}$ qui convergerait beaucoup plus vite vers la même limite. De telles transformations $\{s_i\} \rightarrow \{t_i\}$ existent effectivement, valables pour une classe étendue d'algorithmes lentement convergents. Certaines sont très anciennes et remontent à Euler et Newton. D'autres, nettement plus efficaces il est vrai, sont beaucoup plus récentes. C'est un des buts de ce cours de faire le point sur les techniques existantes.

CHAPITRE I

THEORIE DES RECURRENCES.

On appelle récurrence d'ordre j toute relation du type suivant:

$$f(C_k, C_{k-1}, C_{k-2}, \dots, C_{k-j}, k) = h_k$$

qui relie $j+1$ valeurs consécutives de la suite des C_k pour toutes valeurs entières de k . Cette récurrence est linéaire si la relation f est linéaire sur les C_k ; elle est non linéaire dans le cas contraire. Dans le cas où elle est linéaire, elle est homogène si le terme indépendant des C_k est absent et elle est à coefficients constants si f ne dépend pas explicitement de k . On reconnaît une terminologie déjà utilisée dans la théorie des équations différentielles. On peut accroître l'analogie en introduisant la notation de différence finie Δ . Pour être précis, nous introduisons les opérateurs E et Δ suivants:

$$EC_k = C_{k+1} \quad d'o\grave{u} \text{ on tire : } \quad E^n C_k = C_{k+n}$$

$$\Delta C_k = C_{k+1} - C_k \quad d'o\grave{u} \quad \Delta^2 C_k = C_{k+2} - 2C_{k+1} + C_k \quad etc ...$$

En fait, on a de toute évidence: $\Delta = E - 1$.

Aux ordres plus élevés, on trouve sans peine:

$$\Delta^2 = (E - 1)^2 = E^2 - 2E + 1$$

$$\Delta^3 = (E - 1)^3 = E^3 - 3E^2 + 3E - 1 \quad etc...$$

Ces relations s'inversent sans peine sous la forme:

$$E = \Delta + 1$$

$$E^2 = \Delta^2 + 2\Delta + 1$$

$$E^3 = \Delta^3 + 3\Delta^2 + 3\Delta + 1 \quad etc...$$

Elles permettent d'effectuer les changements de notations suivants.

Soit par exemple la récurrence d'ordre trois :

$$C_{k+3} + kC_{k+2} - C_{k+1} + k^2C_k = 0$$

On peut la réécrire sous la forme équivalente :

$$(E^3 + kE^2 - E + k^2)C_k = 0 \quad (\text{I-1})$$

ou encore, sous cette autre forme, dite aux différences finies, d'ordre trois :

$$\Delta^3 C_k + (k+3)\Delta^2 C_k + (2k+2)\Delta C_k + (k^2 + k)C_k = 0 \quad (\text{I-2})$$

Les écritures (I-1) et (I-2) sont équivalentes. L'écriture (I-2) paraît avoir l'avantage de simuler un parallélisme plus étroit avec la théorie des équations différentielles, l'opérateur Δ prenant la place de l'opérateur de dérivation D . Nous verrons au §I-1-1 qu'il est pourtant préférable de s'en tenir à la forme (I-1).

L'intérêt majeur des récurrences provient du fait qu'elles sont facilement programmées sur ordinateur. Alors qu'il est plutôt rare de considérer des équations différentielles d'ordre supérieur à deux, nous verrons que les récurrences d'ordre plus élevé sont d'un usage courant. Dans cette étude, nous passerons en revue les récurrences linéaires à coefficients constants et variables. Par contre nous passerons sous silence les récurrences non linéaires qui sont nettement plus difficiles à traiter et au sujet desquelles il n'existe que très peu de résultats connus utilisables.

I-1. Récurrence linéaire d'ordre n.

Elle s'écrit sous la forme générale suivante:

$$A_k^{(n)} C_{k+1} + A_k^{(n-1)} C_k + \dots + A_k^{(0)} C_{k-n+1} = h_k \quad (\text{I-3})$$

D'ordre n , elle contient $n+1$ termes plus, éventuellement, un second membre, h_k , indépendant des C_k . Ses coefficients peuvent, en général, dépendre de k . On peut montrer que les résultats établis à propos des équations différentielles se transposent sans difficulté. En particulier on peut montrer que :

- cette équation possède n solutions linéairement indépendantes $f_i(k) = C_k^{(i)}$ ($i=1,2,\dots,n$). On contrôle l'indépendance linéaire de ces n solutions en vérifiant que le déterminant Casoratien des $f_i(k) = C_k^{(i)}$ est différent de zéro :

$$D[f_1 \dots f_n] = \begin{vmatrix} f_1(k) & f_2(k) & \dots & f_n(k) \\ f_1(k+1) & f_2(k+1) & \dots & f_n(k+1) \\ \vdots & \vdots & & \vdots \\ f_1(k+n-1) & f_2(k+n-1) & \dots & f_n(k+n-1) \end{vmatrix} \neq 0$$

L'analogie avec la théorie du Wronskien, valable à propos des équations différentielles, est mieux mise en évidence si on recourt à la forme équivalente suivante :

$$D[f_1 \dots f_n] = \begin{vmatrix} f_1 & f_2 & \dots & f_n \\ \Delta f_1 & \Delta f_2 & \dots & \Delta f_n \\ \vdots & \vdots & & \vdots \\ \Delta^{n-1} f_1 & \Delta^{n-1} f_2 & \dots & \Delta^{n-1} f_n \end{vmatrix}$$

- la solution générale de l'équation homogène est la combinaison linéaire des n solutions de base.
- la solution générale de l'équation avec second membre s'écrit comme la superposition de la solution générale de l'équation homogène plus une solution particulière de l'équation avec second membre : $C_k = \lambda_1 C_k^{(1)} + \dots + \lambda_n C_k^{(n)} + sol. part. \quad \varphi_k$.
- on détermine la solution de façon unique en imposant un jeu de n conditions initiales (ou aux limites), par exemple en fixant les valeurs de C_0, C_1, \dots, C_{n-1} ce qui permet de déterminer la valeur des coefficients λ_i .

Le cas où les coefficients de la récurrence sont constants est particulièrement important car il autorise une solution exacte de la récurrence.

I-1.1. Récurrence linéaire homogène à coefficients constants.

La récurrence linéaire homogène à coefficients constants s'écrit:

$$A^{(n)} C_{k+1} + A^{(n-1)} C_k + \dots + A^{(1)} C_{k-n+1} = 0 \quad (I-4)$$

On peut montrer que l'équation (I-4) possède des solutions exactes qui s'écrivent sous la forme d'exponentielles-polynômes. Commençons par rechercher les solutions qui s'écrivent sous la forme d'exponentielles pures de la forme $C_k = z^k$ où z est une constante à déterminer. Introduisant cette forme dans l'équation (I-4), on trouve de suite l'équation polynomiale caractéristique dont les racines livrent les valeurs permises pour z :

$$A^{(n)} z^n + A^{(n-1)} z^{n-1} + \dots + A^{(1)} z + A^{(0)} = 0$$

Ses racines, réelles ou complexes, se notent z_1, \dots, z_n .

Lorsque ces racines sont toutes distinctes, on a manifestement trouvé les solutions cherchées sous la forme :

$$C_k^{(i)} = z_i^k \quad (i = 1, \dots, n)$$

Elles sont certainement linéairement indépendantes car leur Casoratien se réduit à un déterminant de Vandermonde, non nul :

$$D = \begin{vmatrix} z_1^k & \cdots & z_n^k \\ \vdots & & \vdots \\ z_1^{k+n-1} & \cdots & z_n^{k+n-1} \end{vmatrix} = z_1^k \cdots z_n^k \begin{vmatrix} 1 & \cdots & 1 \\ z_1 & \cdots & z_n \\ \vdots & & \vdots \\ z_1^{n-1} & \cdots & z_n^{n-1} \end{vmatrix} = z_1^k \cdots z_n^k \prod_{i < j} (z_j - z_i) \neq 0$$

La solution générale s'écrit alors sous la forme :

$$C_k = \lambda_1 z_1^k + \cdots + \lambda_n z_n^k$$

où les coefficients λ_i dépendent des conditions initiales.

Lorsque deux ou plusieurs racines sont confondues, on montre qu'une racine de multiplicité μ donne naissance à μ solutions indépendantes qui s'écrivent sous la forme :

$$z^k k^w \quad \text{avec } w = 0, 1, \dots, \mu - 1$$

Exemple 1 : Pour fixer les idées, résolvons la récurrence suivante :

$$C_{k+1} - 5C_k + 8C_{k-1} - 4C_{k-2} = 0$$

L'équation caractéristique s'écrit : $z^3 - 5z^2 + 8z - 4 = 0$.

Ses racines valent :

$$z_1 = 2 \text{ (double)} \rightarrow \begin{cases} C_k^{(1)} = k 2^k \\ C_k^{(2)} = 2^k \end{cases}$$

$$z_2 = 1 \text{ (simple)} \rightarrow C_k^{(3)} = 1^k = 1$$

On vérifie facilement que les 3 solutions trouvées satisfont la récurrence. La solution générale s'écrit donc sous la forme :

$$C_k = \lambda + \mu 2^k + \nu k 2^k$$

Les coefficients λ , μ et ν dépendent des conditions initiales. Par exemple, si on impose : $C_0 = 1$, $C_1 = 3$ et $C_2 = -1$, on écrit le système :

$$\begin{cases} C_0 = \lambda + \mu = 1 \\ C_1 = \lambda + 2\mu + 2\nu = 3 \\ C_2 = \lambda + 4\mu + 8\nu = -1 \end{cases}$$

qui se résout en : $\lambda = -9$, $\mu = 10$ et $\nu = -4$.

Exemple 2 : cet autre exemple montre ce qui se passe lorsque les racines sont complexes.
Soit à résoudre :

$$C_{k+1} - 2C_k + 2C_{k-1} = 0$$

L'équation caractéristique s'écrit :

$$z^2 - 2z + 2 = 0, \text{ dont les racines, simples, valent: } z_{1,2} = 1 \pm i \text{ (simples).}$$

La solution générale de la récurrence peut s'écrire sous la forme:

$$C_k = \lambda(1+i)^k + \mu(1-i)^k$$

On peut lui substituer une forme équivalente qui élude la notation complexe. En effet, on a:

$$1 \pm i = \sqrt{2}e^{\pm i\pi/4} \quad \text{d'où:}$$

$$C_k = \lambda\sqrt{2}^k e^{ik\pi/4} + \mu\sqrt{2}^k e^{-ik\pi/4}$$

soit encore, en groupant les termes autrement :

$$C_k = \lambda'\sqrt{2}^k \cos \frac{k\pi}{4} + \mu'\sqrt{2}^k \sin \frac{k\pi}{4}$$

Si les conditions initiales imposées par le problème sont, par exemple : $C_{-1}=0$ et $C_0=1$, on trouve facilement :

$$C_{-1} = (\lambda' / \sqrt{2}) \cos \frac{\pi}{4} - (\mu' / \sqrt{2}) \sin \frac{\pi}{4} = 0 \quad \text{d'où :} \quad \lambda' = \mu'$$

$$C_0 = \lambda' = 1$$

$$\text{d'où : } C_k = \sqrt{2}^k (\cos \frac{k\pi}{4} + \sin \frac{k\pi}{4})$$

soit la suite : $C_{-1} = 0$; $C_0 = 1$; $C_1 = 2$; $C_2 = 2$; $C_3 = 0$; $C_4 = -4$; ...

qui résout le problème posé. Cette forme exacte permet le calcul de C_{100} , par exemple, sans avoir à itérer la récurrence.

I-1.2. Récurrence linéaire à coefficients constants avec second membre.

Lorsqu'on adjoint un second membre à la récurrence (I-4), elle cesse d'être homogène. Il convient donc d'ajouter à la solution générale de l'équation sans second membre, une solution particulière de l'équation complète. Cela peut se faire par la méthode de variation des constantes, habituelle en théorie des équations différentielles. Nous n'insistons pas sur ce point. Toutefois si le second membre est une exponentielle polynôme du type $Z^k P_m(k)$, une

solution particulière φ_k possède la forme : $\mathbb{Z}^k \text{Pol}_m(k)k^\alpha$, où α est la multiplicité de \mathbb{Z} comme racine de l'équation caractéristique de la récurrence homogène.

Exemple : Nous avons appliqué cette méthode à la récurrence suivante, présentant trois seconds membres distincts. Nous avons trouvé successivement :

$$\begin{aligned} k^2 + 1 &\Rightarrow \varphi_k = k(\alpha k^2 + \beta k + \gamma) \\ C_{k+1} - 5C_k + 8C_{k-1} - 4C_{k-2} &= (-1)^k (k^2 + 1) \Rightarrow \varphi_k = (-1)^k (\alpha k^2 + \beta k + \gamma) \\ 2^k (k^2 + 1) &\Rightarrow \varphi_k = 2^k k^2 (\alpha k^2 + \beta k + \gamma) \end{aligned}$$

Dans chacun des trois cas envisagés, il reste à substituer la forme trouvée pour φ_k dans la récurrence inhomogène afin de trouver la valeur des coefficients α, β et γ qui conviennent.

I-1.3. Récurrence linéaire à coefficients variables.

Lorsque les coefficients de la récurrence sont variables avec k , les choses se compliquent singulièrement. Il n'est, en général, plus possible de trouver une solution exacte à la récurrence, même homogène :

$$A_k^{(n)} C_{k+1} + A_k^{(n-1)} C_k + \dots + A_k^{(0)} C_{k-n+1} = 0 \quad (\text{I-5})$$

Seul le cas de la récurrence d'ordre 1 est soluble exactement. En effet, on a :

$$A_k^{(1)} C_{k+1} + A_k^{(0)} C_k = h_k \quad (\text{I-6})$$

L'équation homogène se réécrit :

$$C_{k+1} = -[A_k^{(0)} / A_k^{(1)}] C_k$$

que l'on peut sommer exactement sous la forme :

$$C_{k+1} = (-1)^{k+1} C_0 \prod_{i=0}^k [A_i^{(0)} / A_i^{(1)}]$$

Reste à trouver une solution particulière de l'équation complète. On vérifie sans peine que la forme suivante convient :

$$\varphi_{k+1} = (-1)^k \prod_{i=0}^k [A_i^{(0)} / A_i^{(1)}] \sum_{\ell=0}^k \frac{(-1)^\ell h_\ell}{A_i^{(1)} \prod_0^\ell [A_i^{(0)} / A_i^{(1)}]}$$

La solution générale de l'équation (I-6) s'obtient en combinant les deux expressions trouvées:

$$C_{k+1} = (-1)^{k+1} \prod_{i=0}^k [A_i^{(0)} / A_i^{(1)}] \left\{ C_0 - \sum_{\ell=0}^k \frac{(-1)^\ell h_\ell}{A_i^{(1)} \prod_0^\ell [A_i^{(0)} / A_i^{(1)}]} \right\}$$

Elle contient évidemment une constante arbitraire C_0 qui est la valeur initiale de C_k en $k = 0$.

Remarque : En posant $k = -1$ dans l'expression de la solution générale, on retrouve bien la valeur initiale C_0 à condition de poser conventionnellement : $\sum_{\ell=0}^{-1} = 0$ et $\prod_{i=0}^{-1} = 1$.

Exemple : soit à résoudre la récurrence :

$$C_{k+1} + (k+1)C_k = k$$

On trouve la solution générale :

$$C_{k+1} = (-1)^{k+1} (k+1)! \left[C_0 - \sum_{\ell=0}^k \frac{(-1)^\ell \ell}{(\ell+1)!} \right]$$

Dans les cas où aucune solution exacte n'est connue, on est amené à rechercher la solution par voie numérique ce qui soulève un certain nombre de problèmes que nous allons étudier.

I-2. Comportement asymptotique des solutions d'une récurrence.

Considérons la récurrence homogène à coefficients variables (I-5). Elle possède n solutions linéairement indépendantes que nous noterons $C_k^{(1)}, C_k^{(2)}, \dots, C_k^{(n)}$, prises dans un ordre que nous préciserons ultérieurement. Il n'est généralement pas possible de trouver une forme exacte pour les $C_k^{(i)}$. Toutefois, on verra qu'on peut, dans un grand nombre de cas, déterminer leur comportement asymptotique $as^{(i)}(k)$ défini de telle manière que : $\lim C_k^{(i)} / as^{(i)}(k) = \text{constante}$, pour k tendant vers l'infini et pour tout $i = 1, \dots, n$.

La récurrence (I-5) étant linéaire, toute combinaison linéaire de ces n solutions est encore solution de (I-5). Au total, la récurrence possède une n -uple infinité de solutions, toutes de la forme :

$$\lambda_1 C_k^{(1)} + \dots + \lambda_n C_k^{(n)}.$$

Le point fondamental est le suivant : toutes ces solutions ne possèdent pas nécessairement le même comportement asymptotique. En fait, il y a, au plus, n asymptotes différentes. Nous appellerons système fondamental de solutions $C_k^{(1)}, C_k^{(2)}, \dots, C_k^{(n)}$ un ensemble de n solutions linéairement indépendantes tel que le produit de leurs asymptotes sera de croissance minimum

lorsque k tend vers l'infini. La numérotation des asymptotes se fait conventionnellement dans l'ordre des dominances décroissantes. Symboliquement, on note:

$$as^{(1)}(k) \geq as^{(2)}(k) \geq \dots \geq as^{(n)}(k)$$

Lorsque deux asymptotes appartenant au système fondamental ne sont pas contrastées en module, elles le sont obligatoirement en phase. Le signe = note symboliquement ce cas.

La solution $C_k^{(1)}$, d'asymptote $as^{(1)}(k)$, est dite dominante;
la solution $C_k^{(n)}$, d'asymptote $as^{(n)}(k)$, est dite dominée;
les autres solutions $C_k^{(2)}, \dots, C_k^{(n-1)}$ sont dites intermédiaires, avec $C_k^{(i)}$ qui ne domine jamais $C_k^{(j)}$ lorsque $i > j$.

On définit $(n-1)$ facteurs de contraste entre solutions contiguës $C_k^{(i)}$ et $C_k^{(i+1)}$:

$$\rho_k^{(i)} = C_k^{(i+1)} / C_k^{(i)} \quad \text{pour } i = 1, 2, \dots, n-1.$$

Lorsque $k \rightarrow \infty$, les $\rho_k^{(i)}$ ne tendent vers zéro que si les solutions $C_k^{(i)}$ et $C_k^{(i+1)}$ sont contrastées en module.

Pour clarifier ces notions, nous reprenons l'exemple de la récurrence, déjà envisagée:

$$C_{k+1} - 5C_k + 8C_{k-1} - 4C_{k-2} = 0$$

Sa solution générale s'écrit sous la forme :

$$C_k = \lambda + \mu 2^k + \nu k 2^k \tag{I-7}$$

Les trois solutions contrastées sont manifestement dans l'ordre des dominances décroissantes :

$$C_k^{(1)} = k 2^k$$

$$C_k^{(2)} = 2^k$$

$$C_k^{(3)} = 1$$

Les contrastes valent respectivement $1/k$ et 2^{-k} . Dans ce cas très simple, les asymptotes coïncident avec les solutions exactes mais cela cesse d'être le cas lorsque la récurrence n'est plus à coefficients constants.

La solution générale (I-7) possède un comportement dominant dès que $\nu \neq 0$ et un comportement sous-dominant dès que $\nu = 0$. Pour y voir plus clair, déterminons les coefficients λ , μ et ν en fonction des conditions initiales.

On écrit par exemple :

$$\begin{cases} C_0 = \lambda + \mu \\ C_{-1} = \lambda + \frac{1}{2}\mu - \frac{1}{2}v \\ C_{-2} = \lambda + \frac{1}{4}\mu - \frac{1}{2}v \end{cases} \Rightarrow \begin{cases} \lambda = C_0 - 4C_{-1} + 4C_{-2} \\ \mu = 4C_{-1} - 4C_{-2} \\ v = 2C_0 - 6C_{-1} + 4C_{-2} \end{cases}$$

- Toutes les solutions dominées sont caractérisées par $v=0$ et $\mu=0$. Elles exigent que l'on ait: $C_0=C_{-1}=C_{-2}$. Il y a donc une simple infinité de solutions dominées, toutes proportionnelles entre elles.

- Les solutions intermédiaires correspondent à $v=0$ et $\mu \neq 0$ d'où on tire :

$2C_0-6C_{-1}+4C_{-2}=0$ et $C_{-1} \neq C_{-2}$. Il y a donc une double infinité de solutions intermédiaires.

- Les autres solutions, caractérisées par $v \neq 0$, sont dominantes. Il y en a une triple infinité.

Il est clair que si l'on choisit les conditions initiales au hasard, on a toutes les chances d'engendrer une solution dominante. Autrement dit, les solutions dominées sont beaucoup plus « rares » que les autres. Il faut savoir que dans un grand nombre d'applications physiques, ce sont précisément les solutions dominées qui sont importantes du fait de leur comportement asymptotique privilégié.

Un problème similaire se présente, en physique, lorsqu'on résout l'équation différentielle de Schrödinger: $\psi''+(E-V)\psi=0$. Etant du second ordre, elle possède deux solutions linéairement indépendantes ψ_1 et ψ_2 qui possèdent des comportements asymptotiques différents aux grandes valeurs de x . Précisément, la solution physiquement acceptable est celle qui décroît le plus vite à l'infini : c'est la solution dominée (cfr §V-4 pour un exemple).

On formule la généralisation de ces résultats de la manière suivante. Revenons, pour ce faire, à la récurrence (I-5) de solution générale : $C_k = \lambda_1 C_k^{(1)} + \lambda_2 C_k^{(2)} + \dots + \lambda_n C_k^{(n)}$.

Cette solution possède un comportement asymptotique dominant $as^1(k)$ si $\lambda_1 \neq 0$. Si $\lambda_1=0$ et $\lambda_2 \neq 0$, elle possède le comportement asymptotique sous-dominant $as^{(2)}(k)$. Si $\lambda_1=\lambda_2=0$ et $\lambda_3 \neq 0$, etc... . Il existe donc une n-uple infinité de solutions dominantes possédant l'asymptote $as^{(1)}(k)$, une (n-1)-uple infinité de solutions d'asymptote $as^{(2)}(k)$, ...enfin une simple infinité de solutions dominées d'asymptote $as^{(n)}(k)$, d'ailleurs toutes identiques entre elles à une constante multiplicative près.

Il est possible d'exprimer ces considérations autrement encore : pour engendrer de façon récurrente une solution particulière de (I-5), il faut partir de n conditions initiales; disons qu'il faut se donner $C_0, C_{-1}, \dots, C_{1-n}$.

- Si on veut engendrer une solution dominée $C_k^{(n)}$, il n'est pas question de choisir ces valeurs arbitrairement. Seul C_0 peut être quelconque mais une fois qu'il est fixé, C_{-1}, \dots, C_{1-n} sont imposés. Nous verrons plus loin, un procédé numérique permettant de les déterminer.

- Si l'on veut construire une solution du type $C_k^{(n-1)}$ qui est dominée par toutes les autres sauf par $C_k^{(n)}$ qu'elle domine, il est possible de choisir C_0 et C_{-1} arbitrairement tandis que C_{-2}, \dots, C_{1-n} ne peuvent pas prendre de valeurs arbitraires. Parmi toutes les valeurs que C_{-1} peut prendre, on exclura cependant celle qui engendre une solution dominée $C_k^{(n)}$.

- Et ainsi de suite pour toutes les solutions intermédiaires.

- Enfin, si l'on veut construire une solution dominante $C_k^{(1)}$, on peut partir de conditions initiales arbitraires sauf qu'il faut prendre soin que n'existe pas entre elles une des relations qui caractérise les solutions dominées.

Les deux exemples qui suivent illustrent la notion de contraste dans deux cas où les solutions exactes de la récurrence ne sont pas connues. Elles sont contrastées en amplitude dans le premier exemple et en phase dans le deuxième exemple.

Exemple 1 : soit la récurrence d'ordre 3 :

$$(k^2 - 2k + 4)C_{k+1} - 3k^2 C_k + (3k^2 + 2k)C_{k-1} - (k^2 + 1)C_{k-2} = 0$$

Elle possède un système fondamental de solutions asymptotiques à (pour $k \rightarrow \infty$):

$$C_k^{(1)} \approx k^{11/8} \exp(4\sqrt{k}) = as^{(1)}(k) \Rightarrow \text{solution dominante}$$

$$C_k^{(2)} \approx k^{3/4} = as^{(2)}(k) \Rightarrow \text{solution intermédiaire}$$

$$C_k^{(3)} \approx k^{11/8} \exp(-4\sqrt{k}) = as^{(3)}(k) \Rightarrow \text{solution dominée}$$

On verra plus loin comment obtenir ces comportements asymptotiques. Les solutions sont contrastées en module. Les facteurs de contraste valent respectivement:

$$\rho_k^{(1)} = k^{-5/8} \exp(-4\sqrt{k})$$

$$\rho_k^{(2)} = k^{5/8} \exp(-4\sqrt{k})$$

Exemple 2 : soit la récurrence d'ordre 2 :

$$(k + 2)C_{k+1} - (2k + 1)C_k + kC_{k-1} = 0$$

Elle possède un système fondamental de solutions asymptotiques à (pour $k \rightarrow \infty$) :

$$C_k^{(1)} \approx k^{-3/4} \exp(2i\sqrt{k}) \Rightarrow |as^{(1)}(k)| = k^{-3/4}$$

$$C_k^{(2)} \approx k^{-3/4} \exp(-2i\sqrt{k}) \Rightarrow |as^{(2)}(k)| = k^{-3/4}$$

Ces deux solutions ne sont pas contrastées en module et leur numérotation peut, de ce fait, être inversée. Aucune combinaison linéaire ne peut évidemment faire apparaître de contraste. Toutefois il existe un contraste de phase qui vaut $\exp(-4i\sqrt{k})$. La notion de contraste est essentielle dans la discussion de l'instabilité numérique des calculs récursifs.

I-3. Instabilité numérique des calculs récurrents.

L'informatisation d'un algorithme récursif peut réserver de fâcheuses surprises. Le lecteur curieux peut procéder à l'expérience numérique suivante, facile à mettre en oeuvre. On programme, en arithmétique flottante, dans le mode progressif, la récurrence que voici :

$$C_{k+1} - 5C_k + 8C_{k-1} - 4C_{k-2} = 0 \quad (k = 0, 1, \dots, 100)$$

en partant de deux ensembles de conditions initiales apparemment similaires :

- 1er cas: $C_2=C_1=C_0=1$ (ou $1/2$)
- 2ème cas: $C_2=C_1=C_0=1/3$ (ou $1/7$)

Dans les deux cas, on devrait trouver que la suite C_k reste constante à sa valeur initiale. En pratique, on constate que ce n'est effectivement vrai que dans le premier cas où l'encodage binaire des conditions initiales est exact. Dans le deuxième cas, une erreur se propage et s'amplifie de façon dramatique dans le calcul.

Le programme Mathematica suivant fait quelque chose de similaire en travaillant sur des conditions initiales exactes ou approchées.

```
co[k_]:=co[k]=5 co[k-1]-8 co[k-2]+4 co[k-3]
{co[0]=1/7;co[1]=1/7;co[2]=1/7};
Table[co[i],{i,0,9}]
```

```
{1/7,1/7,1/7,1/7,1/7,1/7,1/7,1/7,1/7,1/7}
```

```
Clear[co]
co[k_]:=co[k]=5 co[k-1]-8 co[k-2]+4 co[k-3]
{co[0]=N[1/7,6];co[1]=N[1/7,6];co[2]=N[1/7,6]};
Table[N[co[i],6],{i,10}]
```

```
{0.142857,0.142857,0.14286,0.1429,0.1429,0.143,0.14,0.1,0.×10-1,0.×10-1}
```

On observe un phénomène de pollution progressive des résultats.

Pour élucider ce phénomène, nous étudions l'exemple similaire de la tabulation de la fonction de Bessel $J_n(x)$. Le problème se pose de la façon suivante : on désire publier une table de la fonction $J_n(x)$ pour diverses valeurs de la variable x (dans ce qui suit, on se limitera à la valeur $x=1$) et pour $n = 0, 1, 2, \dots, 10$. On sait qu'entre les J_n d'indices n successifs, il existe une relation récurrente d'ordre deux que l'on compte mettre à profit pour calculer $J_2, J_3, J_4, \dots, J_{10}$ une fois connues les valeurs de J_0 et J_1 . Celles-ci doivent faire l'objet d'un calcul préalable par une méthode quelconque.

Cette récurrence s'écrit :

$$J_{n+1} - \frac{2n}{x} J_n + J_{n-1} = 0 \quad (\text{I-8})$$

Concrètement, on injecte les valeurs J_0 et J_1 , calculées initialement avec m chiffres significatifs, dans la récurrence (I-8) et on espère que les valeurs suivantes J_2, J_3, \dots sortiront avec m chiffres exacts. Par exemple pour $x=1$, on a avec huit chiffres corrects :

$J_0(1) = 0.76519768\dots$ et $J_1(1) = 0.44005058\dots$

La récurrence (I-8) appliquée progressivement à ces conditions initiales fournit les valeurs C_n qui sont consignées dans la Table I-1, que l'on comparera aux valeurs exactes $J_n(1)$.

n	$J_n(1)$	C_n
0	7.6519768E-01	7.6519768E-01
1	4.4005058E-01	4.4005058E-01
2	1.1490348E-01	1.1490348E-01
3	1.9563354E-02	1.9563340E-02
4	2.4766389E-03	2.476560E-03
5	2.4975773E-04	2.49140E-04
6	2.0938338E-05	1.4840E-05
7	1.5023258E-06	-7.106E-05
8	9.4223441E-08	-1.00E-03

Table I-1.

On constate une forte dégradation des résultats à mesure que n augmente. C'est un exemple de pollution catastrophique des résultats qui est typique des calculs récursifs mal maîtrisés. Cette pollution n'a rien à voir avec la classique propagation des erreurs d'arrondis, inhérente à tout calcul numérique, mais qui ne présente jamais un caractère aussi dramatique.

Pour étudier de plus près les causes profondes de ce type d'instabilité numérique, nous partons de la récurrence (I-8) qui a servi de base à notre exemple :

$$C_{n+1} - \frac{2n}{x} C_n + C_{n-1} = 0 \quad (\text{I-9})$$

Etant d'ordre deux, cette récurrence possède deux solutions linéairement indépendantes, à savoir les fonctions de Bessel de première et deuxième espèce $J_n(x)$ et $Y_n(x)$. Sa solution générale s'écrit donc :

$$C_n = \lambda J_n(x) + \mu Y_n(x)$$

Les coefficients λ et μ se déterminent d'après les conditions initiales imposées. Du fait que l'on a choisi $C_0=J_0$ et $C_1=J_1$, on s'attend à ce que $C_n=J_n$ pour tout n . Ce serait effectivement le

cas si le calcul était mené rigoureusement en arithmétique infinie. Toutefois la réalité est très différente puisque l'ordinateur travaille toujours en précision limitée (8 chiffres dans notre exemple). Il en résulte que les conditions initiales réellement encodées sont un peu différentes de celles que l'on avait en vue. En réalité, elles s'écrivent :

$$\begin{aligned} C_0 &= J_0(1 - \varepsilon) = \lambda J_0 + \mu Y_0 \\ C_1 &= J_1(1 - \eta) = \lambda J_1 + \mu Y_1 \end{aligned} \quad \text{et}$$

où ε et η représentent la précision des conditions initiales, soit ici $\varepsilon, \eta \approx 10^{-8}$ (des valeurs plus exactes sont $\varepsilon \approx 8.6 \cdot 10^{-9}$ et $\eta \approx 1.3 \cdot 10^{-8}$). En théorie λ devrait valoir 1 et μ devrait valoir 0. En réalité leurs valeurs encodées sont :

$$\lambda = 1 + \frac{\eta J_1 Y_0 - \varepsilon J_0 Y_1}{J_0 Y_1 - J_1 Y_0} \quad \text{et} \quad \mu = (\eta - \varepsilon) \frac{J_0 J_1}{J_1 Y_0 - J_0 Y_1}$$

Il en résulte que la solution, que ces conditions initiales légèrement perturbées engendrent, diffère de $C_n = J_n$ et qu'elle vaut, en réalité :

$$C_n = \left[1 + \frac{\eta J_1 Y_0 - \varepsilon J_0 Y_1}{J_0 Y_1 - J_1 Y_0} \right] J_n + (\eta - \varepsilon) \frac{J_0 J_1}{J_1 Y_0 - J_0 Y_1} Y_n$$

A première vue, on pourrait se montrer satisfait puisque cette expression ne semble différer de $C_n = J_n$ que par des termes de l'ordre de ε et de η , soit la précision des conditions initiales. Si on y regarde de plus près, on constate que cela est tout à fait faux car la solution Y_n domine très fortement J_n en sorte que lorsque n augmente, la solution dominante non désirée Y_n pollue progressivement et dramatiquement la solution dominée cherchée J_n . Ceci est dû à la différence de comportements asymptotiques de Y_n et de J_n . On verra plus loin comment déterminer les asymptotes d'une récurrence telle que (I-9); il suffit pour l'instant de savoir qu'elles s'écrivent respectivement :

$$\begin{aligned} Y_n &\approx as^{(1)}(n) = \left(\frac{x}{2}\right)^n \Gamma(n) \quad (\text{asymptote dominante}) \\ J_n &\approx as^{(2)}(n) = \left(\frac{x}{2}\right)^n / \Gamma(n+1) \quad (\text{asymptote dominée}) \end{aligned}$$

Le contraste entre J_n et Y_n est saisissant; il vaut:

$$\rho_n \approx \frac{(x/2)^{2n}}{\Gamma(n)\Gamma(n+1)}$$

C'est lui qui fixe le degré de pollution de J_n par Y_n : l'erreur absolue commise est de l'ordre de PY_n , où P est la précision de l'ordinateur (soit 10^{-8} dans l'exemple traité ci-dessous).

L'erreur relative s'en déduit : erreur relative $\approx P |Y_n/J_n| = P/\rho_n$

Pour $x = 1$, on trouve : erreur relative $\approx n!(n-1)!10^{-8}/0.5^{2n}$ (I-10)

soit en pratique les ordres de grandeur suivants :

n	erreur relative
1	$4. 10^{-8}$
2	$3.2 10^{-7}$
3	$7.6 10^{-6}$
4	$3.6 10^{-4}$
5	$3. 10^{-2}$
6	3.5 !!

Table I-2

Comparant les valeurs de J_n et celles de C_n telles qu'elles figurent dans la table I-1, on remarque que la formule (I-10) prédit correctement l'ordre de grandeur de la catastrophe lorsque n est suffisamment grand. On y voit un premier intérêt de pouvoir déterminer les asymptotes d'une récurrence. On en découvrira un deuxième lors de l'étude de l'algorithme de Miller. L'exemple précédent révèle une propriété générale des récurrences linéaires, à savoir :

le calcul récursif progressif n'est stable que pour les solutions dominantes $C_k^{(1)}$ de la récurrence étudiée; il est instable pour toutes les solutions non dominantes $C_k^{(i)}$ ($i>1$) et l'erreur relative commise vaut à chaque pas de l'ordre de $P |C_k^{(1)}/C_k^{(i)}|$.

Autrement dit, les solutions non dominantes ne peuvent être calculées de façon stable par la procédure récursive progressive : elles sont en effet automatiquement polluées par la solution dominante, l'ordre de grandeur de la pollution étant donné par le facteur de contraste entre la solution cherchée et la solution dominante. Si ce contraste n'est pas trop sévère, on peut quand même envisager le calcul progressif avec une erreur modérée; s'il est sévère, on doit recourir à un algorithme spécial dû à Miller, que nous allons étudier.

I-4. Algorithme de Miller simple.

L'algorithme de Miller calcule de façon stable $n+1$ valeurs C_0, C_1, \dots, C_n de la solution dominée d'une récurrence. L'idée est des plus simples: elle consiste à essayer la récurrence à rebours, donc dans le sens régressif, en sorte que la solution dominante devienne dominée et inversement. Deux questions se posent alors:

- A quelle valeur N de n faut-il démarrer le calcul régressif ?
- Quelles conditions initiales faut-il adopter pour C_N et C_{N+1} ?

Nous verrons plus loin comment répondre à la première question. En attendant contentons-nous de prendre N suffisamment grand, quitte à augmenter sa valeur par la suite pour mesurer la modification que cela entraîne sur les valeurs cherchées de C_0, C_1, \dots, C_n .

Miller répond à la deuxième question en choisissant des conditions initiales arbitraires, par exemple $C_N=1$ et $C_{N+1}=0$! Cela implique qu'au début, l'algorithme calcule une combinaison linéaire arbitraire des solutions dominantes et dominées. Mais, très vite, la solution dominée va polluer toutes les autres au point qu'elle sera seule présente à l'arrivée pour peu qu'on ait choisi N suffisamment grand. Il suffit alors de renormaliser la suite trouvée en ajustant C_0 à la condition initiale imposée. Si le fait d'augmenter N modifie la suite $C_n, C_{n-1}, \dots, C_1, C_0$ à la précision demandée, e^{-p} , c'est que N n'a pas été choisi suffisamment grand. En pratique, il faut choisir N tel que le contraste ρ entre les solutions dominantes et dominées satisfasse la relation suivante qui fixe N :

$$\rho_{N+1} / \rho_n = e^{-p}$$

Justifions ce qui vient d'être dit.

La solution générale de la récurrence (I-9) étudiée s'écrit sous la forme :

$$C_n = \lambda J_n + \mu Y_n$$

Considérons la solution particulière correspondant aux conditions particulières imposées par l'algorithme de Miller :

$$C_N = 1 = \lambda J_N + \mu Y_N$$

$$C_{N+1} = 0 = \lambda J_{N+1} + \mu Y_{N+1}$$

On en déduit les valeurs de λ et de μ :

$$\lambda = -\frac{Y_{N+1}}{Y_N J_{N+1} - Y_{N+1} J_N}$$

$$\mu = \frac{J_{N+1}}{Y_N J_{N+1} - Y_{N+1} J_N}$$

d'où on déduit la solution engendrée dans ces conditions :

$$C_n = \frac{-J_{N+1} Y_n + Y_{N+1} J_n}{-Y_N J_{N+1} + Y_{N+1} J_N}$$

expression que l'on modifie comme suit :

$$\begin{aligned}
C_n &= \frac{Y_{N+1}}{J_N Y_{N+1} - Y_N J_{N+1}} \left(J_n - Y_n \frac{J_{N+1}}{Y_{N+1}} \right) \\
&= \frac{Y_{N+1}}{J_N Y_{N+1} - Y_N J_{N+1}} J_n \left(1 - \frac{Y_n}{J_n} \frac{J_{N+1}}{Y_{N+1}} \right) \\
&= \lambda J_n \left(1 - \frac{\rho_{N+1}}{\rho_n} \right)
\end{aligned}$$

Si le contraste, ρ , est suffisant et que N est suffisamment grand par rapport à n , on voit que la méthode génère, aux basses valeurs de n , la suite J_n demandée à une constante multiplicative, λ , près que l'on détermine en se servant de la condition initiale $C_0 = J_0$ qui doit évidemment être donnée. On peut recommencer le même calcul en partant de conditions initiales quelconques pour C_N et C_{N+1} : les conclusions restent valables. Par cette méthode, J_n et, a fortiori J_0, J_1, \dots, J_{n-1} , seront corrects à la précision P , à condition de choisir N tel que :

$$\frac{\rho_{N+1}}{\rho_n} \approx P$$

Revenons à l'exemple de la récurrence (I-9) de Bessel. Le calcul de J_n pour $n = 0, 1, \dots, 10$ avec huit chiffres exacts, exige le recours à l'algorithme de Miller. Le facteur de contraste a été évalué à :

$$\rho_n = \frac{1}{2^{2n} (n-1)! n!}$$

On détermine la valeur de N à partir de laquelle l'algorithme de Miller doit démarrer en résolvant l'équation transcendante suivante :

$$\frac{\rho_{N+1}}{\rho_n} = \frac{2^{2n} (n-1)! n!}{2^{2N+2} N! (N+1)!} = 10^{-8} \quad \text{avec} \quad n = 10$$

On trouve $N = 12$. Il importe de bien voir que l'algorithme engendre la suite complète, normalisée à $C_0 = J_0$, soit :

$$C_{13} = 0, C_{12} = *, C_{11}, \mathbf{C_{10}}, \mathbf{C_9}, \dots, \mathbf{C_2}, \mathbf{C_1}, \mathbf{C_0} = \mathbf{J_0}.$$

et que cette valeur de N est telle que C_{10} est tout juste correct avec 8 chiffres; les termes C_n précédents ($n > 10$) sont d'autant moins précis que n est grand tandis que les termes C_n suivants ($n < 10$) sont de précision croissante à mesure que n diminue.

Cela est visible dans la table I-3 qui liste les valeurs obtenues par cette procédure. Les chiffres en gras sont exacts.

n	$C_n J_0(1)/C_0$
13	0.000000000000000E+00
12	4.991679516745441E-13
11	1.198003084018906E-11
10	2.630615105324847E-10
9	5.249250179809506E-09
8	9.422344172603863E-08
7	1.502325817436808E-06
6	2.093833800238928E-05
5	2.497577302112345E-04
4	2.476638964109956E-03
3	1.956335398266841E-02
2	1.149034849319005E-01
1	4.400505857449337E-01
0	7.651976865579670E-01

Table I-3

Remarque : Cet algorithme présente une caractéristique tout à fait remarquable que l'on peut d'ailleurs observer sur l'exemple du calcul de J_n . Alors que le calcul progressif de J_n aurait nécessité la connaissance de deux conditions initiales, par exemple J_0 et J_1 , le calcul régressif basé sur l'algorithme de Miller n'en nécessite plus qu'une seule : J_0 dans l'exemple choisi. En particulier, la valeur de $C_1 = J_1$ surgit tout naturellement de l'algorithme de Miller. En résumé, quand on calcule la solution dominée d'une récurrence par l'algorithme de Miller, une seule condition initiale suffit, ce qui est bien naturel puisque la solution dominée est unique à un facteur multiplicatif près. Théoriquement, on pourrait même imaginer se passer de toute condition initiale du fait que les fonctions J_n obéissent, comme d'ailleurs la plupart des fonctions transcendentes, à des formules sommatoires dont un exemple est donné par la relation :

$$J_0(z) + 2J_2(z) + 2J_4(z) + \dots = 1/2 \quad \text{quel que soit } z.$$

La mise en pratique de cette remarque exige toutefois de prendre en compte la vitesse de convergence de cette série. A cet égard, la formule sommatoire suivante devrait être plus performante, vu le comportement asymptotique rapidement décroissant de J_n :

$$J_0^2(z) + 2J_1^2(z) + 2J_2^2(z) + \dots = 1$$

I-4.1 Application au calcul d'intégrales compliquées.

Soit à calculer les intégrales suivantes (n entier positif) :

$$I_n = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-t} t^{n-1/2}}{(t+1/2)^n} dt \quad (\text{I-11})$$

Analytiquement, seule I_0 se calcule aisément. On trouve classiquement :

$$I_0 = \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-t} \frac{dt}{\sqrt{t}} = 1$$

On peut montrer que les I_n obéissent à une récurrence d'ordre 2 :

$$nI_{n+1} - 2nI_n + (n-1/2)I_{n-1} = 0 \quad (\text{I-12})$$

La démonstration en est la suivante. On considère, simultanément, les familles d'intégrales I_n et J_n définies par les relations:

$$I_n = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-t} t^{n-1/2}}{(t+1/2)^n} dt \quad \text{et} \quad J_n = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-t} t^{n+1/2}}{(t+1/2)^n} dt$$

On commence par intégrer I_n par parties. On pose, pour cela :

$$u = \frac{e^{-t}}{(t+\frac{1}{2})^n} \Rightarrow du = -\frac{e^{-t}}{(t+\frac{1}{2})^n} - n \frac{e^{-t}}{(t+\frac{1}{2})^{n+1}}$$

$$dv = t^{n-1/2} dt \Rightarrow v = \frac{1}{n+1/2} t^{n+1/2}$$

On trouve :

$$I_n = uv \Big|_0^{\infty} + \frac{1}{\sqrt{\pi}} \frac{1}{n+1/2} \left[\int_0^{\infty} \frac{e^{-t} t^{n+1/2}}{(t+1/2)^n} dt + n \int_0^{\infty} \frac{e^{-t} t^{n+1/2}}{(t+1/2)^{n+1}} dt \right]$$

(=0)

d'où on tire :

$$(n+\frac{1}{2})I_n = J_n + nI_{n+1} \quad (\text{I-13})$$

Par ailleurs, on a aussi, en travaillant J_n :

$$\begin{aligned}
J_n &= \int_0^\infty \frac{e^{-t} t^{n+1/2}}{(t+1/2)^n} dt = \int_0^\infty \frac{e^{-t} t^{n-1/2} (t+1/2-1/2)}{(t+1/2)^n} dt \\
&= \int_0^\infty \frac{e^{-t} t^{n-1/2}}{(t+1/2)^{n-1}} dt - \frac{1}{2} \int_0^\infty \frac{e^{-t} t^{n-1/2}}{(t+1/2)^n} dt
\end{aligned}$$

qui se réduit à :

$$J_n = J_{n-1} - \frac{1}{2} I_n \quad (\text{I-14})$$

Eliminant J_n entre (I-13) et (I-14), on trouve la relation cherchée :

$$nI_{n+1} - 2nI_n + (n-1/2)I_{n-1} = 0$$

Accessoirement, on trouve, simultanément, la récurrence satisfaite par les J_n :

$$nJ_{n+1} - (2n+1)J_n + (n+1/2)J_{n-1} = 0$$

La récurrence satisfaite par la suite des I_n possède deux solutions linéairement indépendantes. On verra plus loin comment déterminer leurs asymptotes qui valent respectivement :

$$\begin{aligned}
as^{(1)}(n) &\approx e^{\sqrt{2n}} & (\text{solution dominante}) \\
as^{(2)}(n) &\approx e^{-\sqrt{2n}} & (\text{solution dominée})
\end{aligned}$$

Le contraste entre ces deux asymptotes vaut : $\rho_n = e^{-2\sqrt{2n}}$.

Il est donc assez fortement polluant ($\sim 7.6 \cdot 10^3$ pour $n = 10$), soit une perte de 4 chiffres significatifs si on calcule par erreur la solution dominée de façon progressive. Toute la question est donc de savoir si la suite I_n est solution dominante ou dominée de la récurrence (I-12). Le recours à l'algorithme de Miller s'avérera indispensable dans la deuxième éventualité. Précisément, on détermine le comportement asymptotique des intégrales I_n en posant dans leur définition (I-11) : $t = nu^2$ d'où : $dt = 2nu du$.

On trouve :

$$I_n = \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-nu^2} n^{n-1/2} u^{2n-1} 2nu du}{(nu^2 + 1/2)^n}$$

Aux grandes valeurs de n , le dénominateur se comporte comme :

$$n^n u^{2n} \left(1 + \frac{1}{2nu^2}\right)^n \approx n^n u^{2n} \exp(1/2u^2)$$

L'intégrale I_n est dès lors asymptotique à :

$$I_n \approx 2\sqrt{n/\pi} \int_0^\infty \exp(-nu^2 - 1/2u^2) du \approx e^{-\sqrt{2n}}$$

où on a utilisé ce résultat connu : $\int_0^\infty \exp(-ax^2 - b/x^2) dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} \exp(-2\sqrt{ab})$

On observe que l'intégrale I_n se comporte effectivement comme une solution dominée de la récurrence (I-12), ce qui entraîne le recours obligé à l'algorithme de Miller. Supposons, pour fixer les idées, que l'on veuille calculer la suite I_0, I_1, \dots, I_{10} avec 8 chiffres. Il faut démarrer le calcul régressif à partir de $I_N = 1$ et $I_{N+1} = 0$ (par exemple), avec N solution de l'équation transcendante suivante :

$$\frac{e^{-2\sqrt{2N+2}}}{e^{-2\sqrt{20}}} \approx 10^{-8} \quad d'où \quad N = 92.$$

En fin de calcul régressif, on renormalise toute la suite I_0, I_1, \dots, I_{10} obtenue en la divisant par le facteur, I_0 , qu'il faut pour que la condition initiale $I_0 = 1$ soit satisfaite.

Remarques:

- 1) Bien que la récurrence soit d'ordre 2, on a eu à considérer qu'une seule condition initiale, à savoir $I_0 = 1$; ceci était dû au fait que la suite cherchée était dominée. Si on l'avait trouvée dominante, deux conditions initiales, soit par exemple I_0 et I_1 , auraient été nécessaires et on n'aurait pas eu besoin de recourir à l'algorithme de Miller.
- 2) On peut tenter de bénéficier de la connaissance du comportement asymptotique de la solution dominée pour accélérer légèrement l'algorithme de Miller : il suffit de partir des conditions initiales suivantes, plus optimales pour des raisons évidentes :

$$C_N = as^{(2)}(N) \quad et \quad C_{N+1} = as^{(2)}(N+1)$$

Le gain en précision n'excède cependant guère un chiffre décimal en sorte que cette remarque est ignorée la plupart du temps.

1-5. Algorithme de Miller généralisé.

Le cas des récurrences d'ordre 2 est simple en ce qu'une solution dominante se calcule par le mode progressif tandis qu'une solution dominée se calcule par le mode régressif. Le cas des récurrences d'ordre supérieur à 2 est nettement plus complexe du fait de l'existence de solutions intermédiaires qui ne peuvent se calculer de façon stable dans aucun des deux modes. Il faut, pour les déterminer, recourir à un algorithme de portée plus générale que nous nous bornons à présenter sans démonstration. On procède comme suit : soit les n solutions de la récurrence rangées dans l'ordre des dominances décroissantes :

$$C_k^{(1)} \geq C_k^{(2)} \geq \dots \geq C_k^{(n)}.$$

Devant déterminer de façon stable la solution $C_k^{(i)}$, on réécrit la récurrence sous la forme d'un système linéaire infini obtenu en posant successivement dans la récurrence : $k = i - 1, i, i + 1, \dots$. Hors de ce système infini, on extrait le système tronqué $K \times K$ (K suffisamment grand) qui néglige les C_k d'indice $k > K$. En notation matricielle, cela donne :

$$\begin{pmatrix} A_{i-1}^{(n-i+1)} & \dots & A_{i-1}^{(n)} & 0 & 0 & 0 & 0 \\ \vdots & \ddots & & \ddots & 0 & 0 & 0 \\ A_n^{(0)} & \dots & \ddots & \dots & A_n^{(n)} & 0 & 0 \\ 0 & \ddots & & \ddots & & \ddots & 0 \\ 0 & 0 & A_{K-1}^{(0)} & \dots & \ddots & \dots & A_{K-1}^{(n)} \\ 0 & 0 & 0 & \ddots & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & A_{K+i-2}^{(0)} & \dots & A_{K+i-2}^{(n-i+1)} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ C_K \end{pmatrix} = - \begin{pmatrix} A_{i-1}^{(n-i)} C_0 + \dots + A_{i-1}^{(0)} C_{i-n} \\ A_i^{(n-i-1)} C_0 + \dots + A_i^{(0)} C_{i+1-n} \\ \vdots \\ A_{n-1}^{(0)} C_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Les coefficients $C_0, C_{-1}, \dots, C_{i-n}$ qui figurent au second membre sont les conditions initiales qui, vu l'état de dominance de la solution que l'on a en vue, peuvent être choisies arbitrairement afin que le problème n'ait qu'une solution. Une fois ces valeurs décidées, il suffit de résoudre le système pour obtenir $C_1^{(i)}, \dots, C_K^{(i)}$ de façon généralement stable. Si on augmente la valeur de K , les $C_k^{(i)}$ d'indice k petit tendent vers des valeurs fixes qui sont les valeurs cherchées. Naturellement, ce procédé ne donne correctement $C_k^{(i)}$ que pour k suffisamment petit devant K . Il est clair, en effet, que $C_K^{(i)}$ ne peut en aucun cas être correct puisque dans les dernières équations du système, on a ignoré des termes. En pratique, l'erreur relative commise sur $C_k^{(i)}$ vaut dans tous les cas où il existe un contraste entre les solutions $C_k^{(i)}$ et $C_k^{(i-1)}$:

$$\text{précision sur } C_k^{(i)} \approx \frac{\rho_K^{(i-1)}}{\rho_k^{(i-1)}} = \frac{\left| \frac{C_K^{(i)}}{C_K^{(i-1)}} \right| \left| \frac{C_k^{(i-1)}}{C_k^{(i)}} \right|}{\left| \frac{C_K^{(i-1)}}{C_K^{(i)}} \right| \left| \frac{C_k^{(i)}}{C_k^{(i-1)}} \right|} \quad (i = 2, 3, \dots, n).$$

Cette relation est utile en ce qu'elle permet d'estimer la dimension K du sous-système à résoudre pour obtenir $C_1^{(i)}, \dots, C_K^{(i)}$ avec une précision donnée. L'erreur relative sur $C_k^{(1)}$ est évidemment nulle puisque l'algorithme de Miller généralisé est équivalent, dans ce cas, au calcul progressif qui est stable pour la solution dominante.

Exemple 1: soit la récurrence d'ordre trois déjà évoquée précédemment :

$$(k^2 - 2k + 4)C_{k+1} - 3k^2 C_k + (3k^2 + 2k)C_{k-1} - (k^2 + 1)C_{k-2} = 0$$

qui, rappelons-le, possède trois solutions linéairement indépendantes respectivement asymptotiques à :

$$C_k^{(1)} \approx k^{11/8} \exp(4\sqrt{k}) = as^{(1)}(k) \Rightarrow \text{solution dominante}$$

$$C_k^{(2)} \approx k^{3/4} = as^{(2)}(k) \Rightarrow \text{solution intermédiaire}$$

$$C_k^{(3)} \approx k^{11/8} \exp(-4\sqrt{k}) = as^{(3)}(k) \Rightarrow \text{solution dominée}$$

Toute solution intermédiaire est du type $\mu C_k^{(2)} + \nu C_k^{(3)}$ et exige deux conditions initiales pour fixer μ et ν . Pour obtenir cette solution avec dix chiffres exacts dans le domaine $k = 1, 2, \dots, 50$, il y a lieu de recourir à l'algorithme de Miller généralisé avec K fixé à la valeur, qui est solution de l'équation transcendante suivante :

$$10^{-10} \approx \rho_k^{(1)} / \rho_{50}^{(1)} = \frac{K^{-5/8} \exp(-4\sqrt{K})}{50^{-5/8} \exp(-4\sqrt{50})}$$

On trouve $K = 160$. Il faut donc résoudre un sous-système de dimension 160 pour obtenir $C_1^{(2)}, \dots, C_{50}^{(2)}$ avec la précision de 10^{-10} exigée. En fait les $C_k^{(2)}$ tels que $k < 50$ sont corrects avec une précision meilleure que 10^{-10} tandis que les $C_k^{(2)}$ (avec $k > 50$) sont corrects avec une précision moindre.

Exemple 2 : Reconsidérons la récurrence à coefficient constant, étudiée au §I-3 :

$$C_{k+1} - 9C_k + 26C_{k-1} - 24C_{k-2} = 0$$

Soit à déterminer numériquement les solutions suivantes :

- La solution dominante telle que : $C_{-2} = C_{-1} = C_0 = 1$.

Son expression exacte est : $C_k^{(1)} = 12 \cdot 2^k - 27 \cdot 3^k + 16 \cdot 4^k$ d'où: $C_{10}^{(1)} = 15195181$.

- La solution intermédiaire telle que : $C_{-1} = C_0 = 1$.

Son expression exacte est : $C_k^{(2)} = 4 \cdot 2^k - 3 \cdot 3^k$ d'où: $C_{10}^{(2)} = -173051$.

- La solution dominée telle que $C_0 = 1$.

Son expression exacte est : $C_k^{(3)} = 2^k$ d'où: $C_{10}^{(3)} = 1024$.

Démarrant l'algorithme de Miller généralisé à $K = 40$, on trouve respectivement :

$$C_{10}^{(1)} = 15195181 \text{ (valeur exacte).}$$

$$C_{10}^{(2)} = -173029.26... \text{ (p = 9).}$$

$$C_{10}^{(3)} = 1023.993... \text{ (p = 11.9).}$$

On vérifie sans peine que les précisions p obtenues sont compatibles avec le modèle théorique. En effet, les facteurs de contraste valent respectivement $(3/4)^k$ et $(2/3)^k$ et on a bien :

$$e^{-p} \approx (3/4)^{K-k} \text{ d'où } p_{\text{int}} \approx (K-k) \ln(4/3) = 8.6, \text{ pour la solution intermédiaire.}$$

$$e^{-p} \approx (2/3)^{K-k} \text{ d'où } p_{\text{dominée}} \approx (K-k) \ln(3/2) = 12.2, \text{ pour la solution dominée.}$$

L'algorithme de Miller simple ne fonctionnerait que pour la solution dominée avec le même indice de performance.

On voit sur ces exemples que la notion de contraste entre les solutions d'une récurrence est très importante car elle permet d'évaluer le degré de précision des algorithmes de Miller simple et généralisé. Encore faut-il pouvoir déterminer les comportements asymptotiques des solutions d'une récurrence. Ce point fait l'objet de l'étude des paragraphes suivants.

Avant cela, il reste un mot à dire des récurrences non homogènes. Le principe de la discussion dans ce cas un peu plus complexe est analogue: il faut connaître le comportement asymptotique de la solution particulière de l'équation avec second membre et le comparer à ceux des n solutions de l'équation homogène. L'algorithme de Miller doit alors être adapté en fonction des dominances observées. Nous n'en disons pas davantage ici du fait que nous ne rencontrerons pas ce cas dans la suite du cours. Le lecteur intéressé se rapportera à l'ouvrage fondamental de Wimp cité en référence.

Une alternative existe cependant : soit la récurrence inhomogène :

$$A_k^{(n)} C_{k+1} + \dots + A_k^{(0)} C_{k-n+1} = h_k$$

en abrégé : $L(C_k) = h(k)$ soit encore : $L(C_k) / h(k) = 1$

On a aussi : $L(C_{k+1}) / h(k+1) = 1$

On peut alors ramener le problème au calcul des solutions d'une équation homogène d'ordre supérieur d'une unité à celui de l'équation de départ :

$$L(C_{k+1}) / h(k+1) - L(C_k) / h(k) = 0.$$

I-6. Détermination des asymptotes d'une classe étendue de récurrences linéaires homogènes.

Il est effectivement possible de déterminer les asymptotes d'une classe assez vaste de récurrences du type (I-5) telles que leurs coefficients peuvent s'écrire sous la forme canonique générale:

$$A_k^{(j)} = \sum_{p=0}^{\infty} a_{j,p} k^{-p/r} \quad (\text{I-15})$$

où r est un entier fixé et où les $a_{j,0}$ ne sont pas tous simultanément nuls. Une telle récurrence est dite de Poincaré de type r . Le cas particulier $r=1$ est important car il regroupe les récurrences à coefficients rationnels sur la variable k . On a l'exemple suivant:

$$(k^2 + 1)C_{k+1} + (3k - 1)C_k + k^2 C_{k-1} = 0$$

d'une récurrence qui peut effectivement se réécrire sous la forme canonique de type un :

$$(1 + \frac{1}{k^2})C_{k+1} + (\frac{3}{k} - \frac{1}{k^2})C_k + C_{k-1} = 0 \quad (\text{I-16})$$

Une récurrence de Poincaré de type r est dite régulière si on a en plus que $a_{0,0} \neq 0$ et que $a_{n,0} \neq 0$. Elle est dite irrégulière dans le cas contraire. La récurrence (I-16) est régulière.

I-6.1. Asymptotes des récurrences régulières.

On appelle équation caractéristique de la récurrence régulière (I-5) l'équation algébrique :

$$a_{n,0}z^n + a_{n-1,0}z^{n-1} + \dots + a_{0,0} = 0$$

Vu que $a_{0,0} \neq 0$ et que $a_{n,0} \neq 0$, cette équation est polynomiale de degré n. Elle possède donc n racines, distinctes ou non, réelles ou complexes, peu importe. Dans le cas particulier où la récurrence est à coefficients constants, on retrouve l'équation caractéristique déjà étudiée et les solutions asymptotiques sont exactes dans ce cas. Plusieurs études, dues à Nörlund, Birkhoff, Culmer et Turrittin, ont montré que les asymptotes des n solutions linéairement indépendantes de la récurrence sont toutes du type :

$$as(k) = z^k k^w \exp \left[\alpha k^{\frac{m-1}{m}} + \beta k^{\frac{m-2}{m}} + \dots + \eta k^{\frac{1}{m}} \right] (\ln k)^g \quad (\text{I-17})$$

où m est un entier inférieur ou égal au produit μ , μ étant la multiplicité de z comme solution de l'équation caractéristique, et où les constantes w, α , β , ..., η , g doivent être déterminées par le calcul. En théorie, on procède à cette détermination comme suit : on introduit l'asymptote (I-17) dans la récurrence et on développe tous les termes en puissances de $k^{-1/m}$. Ensuite, on identifie à zéro les coefficients des puissances de k des rangs les plus élevés, ce qui livre en général les relations qui permettent de déterminer les constantes w, α , β , ..., η , g. En pratique, on a effectué ces calculs lourds, une fois pour toutes, dans le plus grand nombre de cas possibles. Les tables I-4 en annexe livrent le résultat complet des calculs pour les récurrences régulières au sens de Poincaré du type un (racine z simple, double, triple ou quadruple) et des types deux, trois et quatre (racine z simple). Le mode d'emploi de ces tables est le suivant : soit la récurrence régulière (I-5) telle que ses coefficients revêtent tous la forme (I-15). On forme l'équation caractéristique possédant n racines :

$$a_{n,0}z^n + a_{n-1,0}z^{n-1} + \dots + a_{0,0} = 0$$

Chaque racine z est obtenue avec sa multiplicité μ . On peut alors modifier la récurrence en posant :

$$C_k = z^k \gamma_k.$$

On obtient une nouvelle récurrence en γ_k dont l'équation caractéristique possède obligatoirement une racine de multiplicité μ en $z = 1$. Les tables I-4 ne sont valables que

lorsque cette opération préalable a été effectuée. En fait, elle doit être recommencée pour chaque racine z . Il suffit alors d'écrire directement l'asymptote en se conformant scrupuleusement aux branchements prévus dans les tableaux I-4.

Tous les résultats sont exprimés en termes des symboles $\sigma(p,q)$ définis par :

$$\sigma(p, q) = \sum_{j=0}^n j^p a_{j,q}$$

(avec $0^0 = 1$ lorsque j et p sont simultanément nuls).

Exemple 1 : Considérons la récurrence régulière au sens de Poincaré, de type un :

$$(1 - 2k^{-1} + 4k^{-2})C_{k+1} - 3C_k + (3 + 2k^{-1})C_{k-1} - (1 + k^{-2})C_{k-2} = 0$$

$j = 3$	$j = 2$	$j = 1$	$j = 0$
$a_{3,0} = 1$	$a_{2,0} = -3$	$a_{1,0} = 3$	$a_{0,0} = -1$
$a_{3,1} = -2$	$autres = 0$	$a_{1,1} = 2$	$a_{0,1} = 0$
$a_{3,2} = 4$		$autres = 0$	$a_{0,2} = -1$
$autres = 0$			$autres = 0$

L'équation caractéristique s'écrit : $z^3 - 3z^2 + 3z - 1 = 0$
dont la racine $z = 1$ est triple. On calcule sans peine :

$$\begin{aligned} \sigma(0,0) &= \sigma(1,0) = \sigma(2,0) = 0 \\ \sigma(3,0) &= 6 & \sigma(0,1) &= 0 & \sigma(2,1) &= -16 \\ \sigma(4,0) &= 36 & \sigma(1,1) &= -4 & \sigma(0,2) &= 3 \end{aligned}$$

Consultant les tables I-4, on trouve les asymptotes correspondant à ce cas sous la rubrique: type un, racine triple, avec $Q_2 = -5/2$. Les asymptotes valent :

$$\begin{aligned} C_k^{(1)} &\approx k^{11/8} \exp(4\sqrt{k}) \\ C_k^{(2)} &\approx k^{3/4} \\ C_k^3 &\approx k^{11/8} \exp(-4\sqrt{k}) \end{aligned}$$

Exemple 2 : Pour bien fixer les idées, traitons encore cet exemple où une racine de l'équation caractéristique n'est pas égale à 1. Soit la récurrence régulière de type un :

$$C_{k+1} - (3 + \frac{2}{k})C_k + 2C_{k-1} = 0$$

$$j=2 \quad j=1 \quad j=0$$

$$a_{2,0} = 1 \quad a_{1,0} = -3 \quad a_{0,0} = 2$$

$$a_{2,1} = 0 \quad a_{1,1} = -2 \quad a_{0,1} = 0$$

$$autres = 0 \quad autres = 0 \quad autres = 0$$

L'équation caractéristique s'écrit : $z^2 - 3z + 2 = 0$. Ses racines sont $z_1 = 1$ et $z_2 = 2$. L'asymptote qui correspond à $z_1 = 1$ se détermine immédiatement en consultant les tables I-4 à la rubrique type un, racine simple.

On trouve sans peine :

$$C_k \approx k^w \text{ avec } w = -\sigma(0,1)/\sigma(1,0) = -(-2)/(-1) = -2$$

La détermination de l'autre asymptote, qui correspond à $z_2 = 2$, nécessite de poser préalablement $C_k = 2^k \gamma_k$. On cherche alors la récurrence satisfaite par γ_k .

On trouve sans difficulté :

$$2\gamma_{k+1} - (3 + \frac{2}{k})\gamma_k + \gamma_{k-1} = 0$$

$$j=2 \quad j=1 \quad j=0$$

$$a_{2,0} = 2 \quad a_{1,0} = -3 \quad a_{0,0} = 1$$

$$a_{2,1} = 0 \quad a_{1,1} = -2 \quad a_{0,1} = 0$$

d'où: $\sigma(0,1) = -2$ et $\sigma(1,0) = 1$.

Consultant les tables I-4, on trouve cette fois: $\gamma_k \approx k^2$ d'où: $C_k \approx 2^k k^2$.

La récurrence initiale possède donc deux solutions contrastées asymptotiques à :

$$C_k^{(1)} \approx 2^k k^2 \quad (\text{solution dominante}) \quad \text{et}$$

$$C_k^{(2)} \approx k^{-2} \quad (\text{solution dominée}).$$

I-6.2. Asymptotes des récurrences irrégulières.

Dans les cas très nombreux où la récurrence n'est pas régulière au sens de Poincaré, soit que $a_{n,0}$ est nul ou que $a_{0,0}$ l'est (ou les deux), on procède comme suit. On commence par former l'équation caractéristique :

$$A_k^{(n)} z^n + A_k^{(n-1)} z^{n-1} + \dots + A_k^{(0)} = 0$$

en ne retenant que le terme principal en k dans chacun des coefficients $A_k^{(j)}$, c'est-à-dire le premier terme $a_{j,p}k^{-p/r}$ non nul. Il est possible de montrer que l'équation caractéristique possède n racines de la forme approchée: $z \approx \rho k^\tau$ (k étant supposé très grand), de multiplicité v .

Pour découvrir les v asymptotes associées à chacune des racines revêtant cette forme, on pose: $C_k = \rho^k [\Gamma(k)]^\tau \gamma_k$, que l'on introduit dans la récurrence de départ. On obtient ainsi une nouvelle récurrence portant sur γ_k .

L'équation caractéristique de cette nouvelle récurrence possède nécessairement une racine de multiplicité v en $z = 1$. Les asymptotes qui lui sont associées se calculent par la méthode décrite plus haut comme si la récurrence satisfaite par γ_k était régulière au sens de Poincaré. Son type dépend des valeurs trouvées pour τ .

Exemple 1 : Soit la récurrence :

$$qC_{k+1} + 4(k+1)^2 C_k + qC_{k-1} = 0$$

où q est une constante quelconque; cette récurrence intervient dans la théorie des fonctions de Mathieu. Elle n'est pas régulière. On forme son équation caractéristique :

$$qz^2 + 4k^2 z + q = 0$$

dont on cherche le comportement asymptotique des racines sous la forme : $z = \rho k^\tau$. On a donc:

$$q\rho^2 k^{2\tau} + 4\rho k^{\tau+2} + qk^0 = 0$$

- $2\tau = 0$ est impossible.
- $2\tau = \tau + 2$, d'où $\tau = 2$ est possible à condition de prendre : $\rho = -4/q$.
- $\tau + 2 = 0$, d'où $\tau = -2$ est possible à condition de prendre : $\rho = -q/4$.

Les racines cherchées sont donc asymptotiques à : $z_1 \approx -(4/q)k^2$ et $z_2 \approx -(q/4)k^{-2}$

Considérant z_1 , on a: $\rho = -4/q$ et $\tau = 2$. On pose donc :

$$C_k = (-4/q)^k [\Gamma(k)]^2 \gamma_k$$

On trouve pour γ_k , la récurrence :

$$\gamma_{k+1} - (1 + k^{-1})^2 \gamma_k + \frac{q^2}{16} k^{-2} (k-1)^{-2} \gamma_{k-1} = 0$$

$$j=2 \quad j=1 \quad j=0$$

$$a_{2,0} = 1 \quad a_{1,0} = -1 \quad a_{0,0} = 0$$

$$a_{2,1} = 0 \quad a_{1,1} = -2 \quad a_{0,1} = 0$$

$$a_{2,2} = 0 \quad a_{1,2} = -1 \quad a_{0,2} = 0$$

On détermine l'asymptote correspondante comme indiqué précédemment (cfr tables I-4, type un, racine simple), à savoir : $\gamma_k \approx k^2$, soit au total :

$$C_k^{(1)} \approx (-4/q)^k k!^2$$

Considérant l'autre racine z_2 , on trouverait pour l'autre solution le comportement asymptotique suivant :

$$C_k^{(2)} \approx (-q/4)^k k!^{-2} k^{-2}$$

On note que la numérotation des solutions tient compte de ce que $C_k^{(1)}$ domine effectivement $C_k^{(2)}$.

Exemple 2 : Considérons la récurrence de Bessel déjà envisagée précédemment :

$$C_{k+1} - \frac{2k}{x} C_k + C_{k-1} = 0$$

L'équation caractéristique s'écrit :

$$z^2 - \frac{2k}{x} z + 1 = 0$$

On pourrait trouver le comportement asymptotique de ses racines en procédant comme dans l'exemple précédent. On peut aussi résoudre l'équation du second degré et trouver :

$$z_{1,2} = \frac{k}{x} \pm \sqrt{\frac{k^2}{x^2} - 1}$$

Aux grandes valeurs de k , on trouve respectivement :

$$z_1 \approx 2k/x \text{ soit : } \rho = 2/x \text{ et } \tau = 1$$

$$z_2 \approx x/(2k) \text{ soit : } \rho = x/2 \text{ et } \tau = -1.$$

Traitant séparément chaque racine, on trouve les asymptotes :

$$C_k^{(1)} \approx (2/x)^k \Gamma(k) \quad \text{et} \quad C_k^{(2)} \approx (x/2)^k / \Gamma(k+1)$$

Nous livrons, en annexe, le contenu des tables I-4 qui permettent de déterminer le comportement asymptotique des solutions d'une récurrence linéaire de Poincaré.

RECURRENCE DE POINCARÉ DU TYPE 1

Racine $z = 1$ *simple* : $\sigma(0, 0) = 0$ et $\sigma(1, 0) \neq 0$
 $\rightarrow C_k \sim k^w$

Racine $z = 1$ *double* : $\sigma(0, 0) = \sigma(1, 0) = 0$ et $\sigma(2, 0) \neq 0$

1. $\sigma(0, 1) \neq 0$
 $\rightarrow C_k \sim k^w \exp [\alpha k^{1/2}]$

2. $\sigma(0, 1) = 0$
 $\rightarrow C_k \sim k^w$

Remarque : si $w_1 = w_2$
 $\rightarrow C_k^{(1)} \sim C_k^{(2)} \ln k$

Racine $z = 1$ *triple* : $\sigma(0, 0) = \sigma(1, 0) = \sigma(2, 0) = 0$ et $\sigma(3, 0) \neq 0$

1. $\sigma(0, 1) \neq 0$
 $\rightarrow C_k \sim k^w \exp [\alpha k^{2/3} + \beta k^{1/3}]$

2. $\sigma(0, 1) = 0$

2.1. $\sigma(1, 1) \neq 0$
 $\rightarrow C_k \sim k^w \exp [\alpha k^{1/2}]$

2.2. $\sigma(1, 1) = 0$

2.2.1. $\sigma(0, 2) \neq 0$
 $\rightarrow C_k \sim k^w \exp [\beta k^{1/3}]$

2.2.2. $\sigma(0, 2) = 0$
 $\rightarrow C_k \sim k^w$

Remarque : si $w_1 = w_2$
 $\rightarrow C_k^{(1)} \sim C_k^{(2)} \ln k$

si $w_1 = w_2 = w_3$
 $\rightarrow C_k^{(1)} \sim C_k^{(2)} \ln k \sim C_k^{(3)} (\ln k)^2$

Racine $z = 1$ *quadruple* : $\sigma(0, 0) = \sigma(1, 0) = \sigma(2, 0) = \sigma(3, 0) = 0$ et $\sigma(4, 0) \neq 0$

1. $\sigma(0, 1) \neq 0$
 $\rightarrow C_k \sim k^w \exp [\alpha k^{3/4} + \beta k^{2/4} + \gamma k^{1/4}]$

$$w = -\sigma(0, 1)/\sigma(1, 0)$$

$$\alpha^2 = -8\sigma(0, 1)/\sigma(2, 0)$$

$$w = 1/4 - 1/24 \alpha^2 \sigma(3, 0)/\sigma(2, 0) - \sigma(1, 1)/\sigma(2, 0)$$

$$1/2 w(w-1)\sigma(2, 0) + w\sigma(1, 1) + \sigma(0, 2) = 0$$

$$\alpha^3 = -81/4 \sigma(0, 1)/\sigma(3, 0)$$

$$\beta = -(1/\alpha) [9\sigma(1, 1) + 1/9 \alpha^3 \sigma(4, 0)]/\sigma(3, 0)$$

$$w = 1/3 - Q_1/\sigma(3, 0)$$

Soit :

$$\alpha = 0$$

$$w = -\sigma(0, 2)/\sigma(1, 1)$$

ou :

$$\alpha^2 = -24\sigma(1, 1)/\sigma(3, 0)$$

$$w = 3/4 + Q_2/\sigma(1, 1)$$

$$\beta^3 = -162\sigma(0, 2)/\sigma(3, 0)$$

$$w = 2/3 - \sigma(2, 1)/\sigma(3, 0)$$

$$1/6 w(w-1)(w-2)\sigma(3, 0) + 1/2 w(w-1)\sigma(2, 1) + w\sigma(1, 2) + \sigma(0, 3) = 0$$

$$\alpha^4 = -(2048/27)\sigma(0, 1)/\sigma(4, 0)$$

$$\beta = -(1/\alpha^2)Q_3/\sigma(4, 0)$$

$$\gamma = -(1/\alpha^3)Q_4/\sigma(4, 0)$$

$$w = -(1/\alpha^3)Q_5/\sigma(4, 0)$$

2. $\sigma(0, 1) = 0$

2.1. $\sigma(1, 1) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\alpha k^{2/3} + \beta k^{1/3}]$$

Soit :	ou :
$\alpha = 0$	$\alpha^3 = -81\sigma(1, 1)/\sigma(4, 0)$
$\beta = 0$	$\beta = Q_6/\sigma(1, 1)$
$w = -\sigma(0, 2)/\sigma(1, 1)$	$w = 1/3 Q_{10}/\sigma(1, 1)$

2.2. $\sigma(1, 1) = 0$

Former l'équation

$$1/384 y^2 \sigma(4, 0) + 1/8 y \sigma(2, 1)$$

$$+ \sigma(0, 2) = 0 \rightarrow y = y_1, y_2$$

Pour chaque racine y_i , discuter :

$$\rightarrow C_k \sim k^w \exp [\beta k^{1/2}]$$

$$\beta_i^2 = y_i$$

$$w = -Q_7/[1/24 \beta^2 \sigma(4, 0) + \sigma(2, 1)]$$

$$1/2 (w^2 - w) \sigma(2, 1) + w \sigma(1, 2) + \sigma(0, 3) = 0$$

Remarque : si $w_1 = w_2$

2.2.2. $y_1 = y_2$

$$[1/24 y \sigma(4, 0) + \sigma(2, 1) = 0]$$

$$2.2.2.1. \sigma(0, 2) \neq 0 \text{ et } Q_7 \neq 0 \rightarrow C_k \sim k^w \exp [\beta k^{2/4} + \gamma k^{1/4}]$$

$$2.2.2.2. \sigma(0, 2) \neq 0 \text{ et } Q_7 = 0 \rightarrow C_k \sim k^w \exp [\beta k^{1/2}]$$

Remarque : si $w_1 = w_2$

2.2.2.3. $\sigma(0, 2) = 0$

2.2.2.3.1. $\sigma(1, 2) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\beta k^{1/3}]$$

Soit :	ou :
$\beta = 0$	$\beta^3 = -648\sigma(1, 2)/\sigma(4, 0)$
$w = -\sigma(0, 3)/\sigma(1, 2)$	$w = -Q_8 \left[\frac{1}{162} \beta^3 \sigma(4, 0) \right. \\ \left. + \sigma(1, 2) \right]$

2.2.2.3.2. $\sigma(1, 2) = 0$ et

$\sigma(0, 3) \neq 0$

$$\gamma^4 = -6144 \sigma(0, 3)/\sigma(4, 0)$$

$$w = 9/8 - \sigma(3, 1)/\sigma(4, 0)$$

2.2.2.3.3. $\sigma(1, 2) = 0$ et
 $\sigma(0, 3) = 0$

$$\rightarrow C_k \sim k^w$$

$$\begin{aligned} & (1/24) w(w-1)(w-2)(w-3)\sigma(4, 0) \\ & + (1/6) w(w-1)(w-2)\sigma(3, 1) \\ & + (1/2) w(w-1)\sigma(2, 2) + w\sigma(1, 3) \\ & + \sigma(0, 4) = 0 \end{aligned}$$

Remarques : si $w_1 = w_2 \rightarrow C_k^{(1)} \sim C_k^{(2)} \ln k$

si $w_1 = w_2 = w_3 \rightarrow C_k^{(1)} \sim C_k^{(2)} \ln k \sim C_k^{(3)} (\ln k)^2$

si $w_1 = w_2 = w_3 = w_4 \rightarrow C_k^{(1)} \sim C_k^{(2)} \ln k \sim C_k^{(3)} (\ln k)^2 \sim C_k^{(4)} (\ln k)^3$

RECURRENCE DE POINCARÉ DU TYPE 2

Racine $z = 1$ simple : $\sigma(0, 0) = 0$ et $\sigma(1, 0) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\alpha k^{1/2}]$$

$$\begin{aligned} \alpha &= -2\sigma(0, 1)/\sigma(1, 0) \\ w &= -S_1/\sigma(1, 0) \end{aligned}$$

Racine $z = 1$ double : $\sigma(0, 0) = \sigma(1, 0) = 0$ et $\sigma(2, 0) \neq 0$

1. $\sigma(0, 1) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\alpha k^{3/4} + \beta k^{2/4} + \gamma k^{1/4}]$$

$$\begin{aligned} \alpha^2 &= -32/9 \sigma(0, 1)/\sigma(2, 0) \\ \beta &= -[3/16 \alpha^2 \sigma(3, 0) + 2\sigma(1, 1)]/\sigma(2, 0) \\ \gamma &= -(1/\alpha) D_1/\sigma(2, 0) \\ w &= -(1/\alpha) D_2/\sigma(2, 0) \end{aligned}$$

2. $\sigma(0, 1) = 0$

Former l'équation : $1/8 \beta^2 \sigma(2, 0) + 1/2 \beta \sigma(1, 1) + \sigma(0, 2) = 0 \rightarrow$

2.1. $\beta_1 \neq \beta_2$

$$[\sigma^2(1, 1) \neq 2\sigma(2, 0)\sigma(0, 2)] \rightarrow C_k \sim k^w \exp [\beta k^{1/2}]$$

2.2. $\beta_1 = \beta_2$

$$[\sigma^2(1, 1) = 2\sigma(2, 0)\sigma(0, 2)]$$

2.2.1. $H_1 \neq 0$

$$\rightarrow C_k \sim k^w \exp [\beta k^{2/4} + \gamma k^{1/4}]$$

$$\begin{aligned} \beta &= \beta_1, \beta_2 \\ \beta &= \beta_i \\ w &= -H_1/[1/2 \beta \sigma(2, 0) + \sigma(1, 1)] \end{aligned}$$

2.2.2. $H_1 = 0$

$$\rightarrow C_k \sim k^w \exp [\gamma k^{1/2}]$$

Remarque : si $w_1 = w_2$

$$\rightarrow C_k^{(1)} \sim C_k^{(2)} \ln k$$

$$\begin{aligned} \beta &= -2\sigma(1, 1)/\sigma(2, 0) \\ \gamma^2 &= -32 H_1/\sigma(2, 0) \\ w &= -D_3/\sigma(2, 0) \\ \gamma &= -2\sigma(1, 1)/\sigma(2, 0) \\ 1/2 w^2 \sigma(2, 0) + w D_4 + H_2 &= 0 \end{aligned}$$

Racine $z = 1$ triple : $\sigma(0, 0) = \sigma(1, 0) = \sigma(2, 0) = 0$ et $\sigma(3, 0) \neq 0$

$$1. \sigma(0, 1) \neq 0 \quad \rightarrow C_k \sim k^w \exp [\alpha k^{5/6} + \beta k^{4/6} + \gamma k^{3/6} + \delta k^{2/6} + \varepsilon k^{1/6}]$$

$$\begin{aligned} \alpha^3 &= -1296/125 \sigma(0, 1)/\sigma(3, 0) \\ \beta &= -(1/\alpha) [25/288 \alpha^3 \sigma(4, 0) \\ &\quad + 18/5 \sigma(1, 1)]/\sigma(3, 0) \end{aligned}$$

$$\begin{aligned} \gamma &= -(1/\alpha^2) D_5/\sigma(3, 0) \\ \delta &= -(1/\alpha^2) D_6/\sigma(3, 0) \\ \varepsilon &= -(1/\alpha^2) D_7/\sigma(3, 0) \\ w &= -(1/\alpha^2) D_8/\sigma(3, 0) \end{aligned}$$

2. $\sigma(0, 1) = 0$

2.1. $\sigma(1, 1) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\alpha k^{3/4} + \beta k^{2/4} + \gamma k^{1/4}]$$

$$\text{Soit :} \quad \alpha = 0 \quad \text{ou :} \quad \alpha^2 = -32/3 \sigma(1, 1)/\sigma(3, 0)$$

$$\beta = -2 \sigma(0, 2)/\sigma(1, 1) \quad \beta = D_9/\sigma(1, 1)$$

$$\gamma = 0 \quad \gamma = D_{10}/\sigma(1, 1)$$

$$w = -[H_1] \quad w = \frac{1}{2} [H_1 + D_{11}]$$

$$+ 1/8 \beta \sigma(2, 0)]/\sigma(1, 1) \quad \sigma(1, 1)$$

2.2. $\sigma(1, 1) = 0$

2.2.1. $\sigma(0, 2) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\beta k^{4/6} + \gamma k^{3/6} + \delta k^{2/6} + \varepsilon k^{1/6}]$$

$$\beta^3 = -81/4 \sigma(0, 2)/\sigma(3, 0)$$

$$\gamma = -2 \sigma(2, 1)/\sigma(3, 0)$$

$$\delta = -(1/\beta) D_{12}/\sigma(3, 0)$$

$$\varepsilon = -(1/\beta^2) D_{13}/\sigma(3, 0)$$

$$w = -(1/\beta^2) D_{14}/\sigma(3, 0)$$

2.2.2. $\sigma(0, 2) = 0$

Former l'équation :

$$1/48 \gamma^3 \sigma(3, 0) + 1/8 \gamma^2 \sigma(2, 1) + 1/2 \gamma \sigma(1, 2) + \sigma(0, 3) = 0 \rightarrow \gamma_i (i = 1, 2, 3)$$

Pour chaque racine γ_i , discuter les cas suivants :

2.2.2.1. γ_i racine simple

$$[1/8 \gamma_i^2 \sigma(3, 0) + 1/2 \gamma_i \sigma(2, 1) + \sigma(1, 2) \neq 0] \rightarrow C_k \sim k^w \exp [\gamma_i k^{1/2}]$$

$$\gamma = \gamma_i$$

$$w = -H_2/[1/8 \gamma^2 \sigma(3, 0) + 1/2 \gamma \sigma(2, 1) + \sigma(1, 2)]$$

2.2.2.2. γ_t racine double
 $[1/8\gamma_t^2\sigma(3, 0) + 1/2\gamma_t\sigma(2, 1)$
 $+ \sigma(1, 2) = 0$ et
 $1/2\gamma_t\sigma(3, 0) + \sigma(2, 1) \neq 0]$
2.2.2.2.1. $H_2 \neq 0$

$$\rightarrow C_k \sim k^w \exp [\gamma k^{2/4} + \delta k^{1/4}]$$

$$\begin{aligned}\gamma &= \gamma_t \\ \delta^2 &= -32 H_2/[1/2\gamma\sigma(3, 0) + \sigma(2, 1)] \\ w &= -1/24 - [H_3 + 1/96\delta^2\sigma(3, 0)]/ \\ &\quad [1/2\gamma\sigma(3, 0) + \sigma(2, 1)]\end{aligned}$$

2.2.2.2.2. $H_2 = 0$

$$\rightarrow C_k \sim k^w \exp [\gamma k^{1/2}]$$

$$\begin{aligned}\gamma &= \gamma_t \\ (w^2/2 - w/12) [1/2\gamma\sigma(3, 0) + \sigma(2, 1)] \\ &\quad + wH_3 + H_4 = 0\end{aligned}$$

Remarque : si $w_1 = w_2$

2.2.2.3. γ_t racine triple
 $[1/8\gamma_t^2\sigma(3, 0) + 1/2\gamma_t\sigma(2, 1)$
 $+ \sigma(1, 2) = 0$ et
 $1/2\gamma_t\sigma(3, 0) + \sigma(2, 1) = 0]$
2.2.2.3.1. $H_2 \neq 0$

$$\rightarrow C_k \sim k^w \exp [\gamma k^{3/6} + \delta k^{2/6} + \varepsilon k^{1/6}]$$

$$\begin{aligned}\gamma &= -2\sigma(2, 1)/\sigma(3, 0) \\ \delta^3 &= -162 H_2/\sigma(3, 0) \\ \varepsilon &= -(1/\delta) 36 H_3/\sigma(3, 0) \\ w &= -(1/\delta^2) [3\varepsilon H_3 + D_{15}]/\sigma(3, 0)\end{aligned}$$

2.2.2.3.2. $H_2 = 0$

2.2.2.3.2.1. $H_3 \neq 0$

$$\rightarrow C_k \sim k^w \exp [\gamma k^{2/4} + \delta k^{1/4}]$$

$$\begin{array}{l} \text{Soit :} \\ \delta = 0 \\ w = -H_4/H_3 \end{array} \quad \left| \begin{array}{l} \text{ou :} \\ \delta^2 = -96 H_3/\sigma(3, 0) \\ w = H_4/(2 H_3) + 9/8 \\ -3 \frac{[1/8\gamma^2\sigma(4, 0) + 1/2\gamma\sigma(3, 1) + \sigma(2, 2)]}{2\sigma(3, 0)} \end{array} \right.$$

2.2.2.3.2.2. $H_3 = 0$
 $H_4 \neq 0$

$$\rightarrow C_k \sim k^w \exp [\gamma k^{3/6} + \varepsilon k^{1/6}]$$

$$\begin{aligned}\gamma &= -2\sigma(2, 1)/\sigma(3, 0) \\ \varepsilon^3 &= -1296 H_4/\sigma(3, 0) \\ w &= -[1/8\gamma^2\sigma(4, 0) - 5/6\sigma(3, 0) + 1/2\gamma\sigma(3, 1) + \sigma(2, 2)]/ \\ &\quad \sigma(3, 0)\end{aligned}$$

2.2.2.3.2.3. $H_3 = 0$
 $H_4 = 0$

$$\rightarrow C_k \sim k^w \exp [\gamma k^{1/2}]$$

$$\begin{aligned} \gamma &= -2 \sigma(2, 1) / \sigma(3, 0) \\ 1/6w (w-1) (w-2) \sigma(3, 0) &+ 1/2 (w^2 - w) \\ [1/8 \gamma^2 \sigma(4, 0) &+ 1/2 \gamma \sigma(3, 1) + \sigma(2, 2)] + w D_{16} + \\ D_{17} &= 0 \end{aligned}$$

Remarque : si $w_1 = w_2 \neq w_3$
 si $w_1 = w_2 = w_3$

$$\begin{aligned} \rightarrow C_k^{(1)} &\sim C_k^{(2)} \ln k \\ \rightarrow C_k^{(1)} &\sim C_k^{(2)} \ln k \sim C_k^{(3)} (\ln k)^2 \end{aligned}$$

RECURRENCE DE POINCARÉ DU TYPE 3

Racine $z = 1$ simple : $\sigma(0, 0) = 0$ et $\sigma(1, 0) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\alpha k^{2/3} + \beta k^{1/3}]$$

$$\begin{aligned} \alpha &= -3/2 \sigma(0, 1) / \sigma(1, 0) \\ \beta &= -S_2 / \sigma(1, 0) \\ w &= -S_3 / \sigma(1, 0) \end{aligned}$$

Racine $z = 1$ double : $\sigma(0, 0) = \sigma(1, 0) = 0$ et $\sigma(2, 0) \neq 0$

$$1. \sigma(0, 1) \neq 0 \quad \rightarrow C_k \sim k^w \exp [\alpha k^{5/6} + \beta k^{4/6} + \gamma k^{3/6} + \delta k^{2/6} + \varepsilon k^{1/6}]$$

$$\begin{aligned} \alpha^2 &= -72/25 \sigma(0, 1) / \sigma(2, 0) \\ \beta &= -[25/144 \alpha^2 \sigma(3, 0) + 3/2 \sigma(1, 1)] / \sigma(2, 0) \\ \gamma &= -(1/\alpha) T_1 / \sigma(2, 0) \\ \delta &= -(1/\alpha) T_2 / \sigma(2, 0) \\ \varepsilon &= -(1/\alpha) T_3 / \sigma(2, 0) \\ w &= -(1/\alpha) T_4 / \sigma(2, 0) \end{aligned}$$

2. $\sigma(0, 1) = 0$

Former l'équation :

$$2/9 \beta^2 \sigma(2, 0) + 2/3 \beta \sigma(1, 1) + \sigma(0, 2) = 0$$

$$\rightarrow \beta = \beta_1, \beta_2$$

2.1. $\beta_1 \neq \beta_2$
 $[\sigma^2(1, 1) \neq 2 \sigma(2, 0) \sigma(0, 2)]$

$$\rightarrow C_k \sim k^w \exp [\beta k^{2/3} + \delta k^{1/3}]$$

$$\begin{aligned} \beta &= \beta_t \\ \delta &= -H_5 / [2 \beta \sigma(2, 0) + 3 \sigma(1, 1)] \\ w &= -T_5 / [2 \beta \sigma(2, 0) + 3 \sigma(1, 1)] \end{aligned}$$

$$2.2. \beta_1 = \beta_2$$

$$[\sigma^2(1, 1) = 2 \sigma(2, 0) \sigma(0, 2)]$$

$$2.2.1. H_5 \neq 0$$

$$\rightarrow C_k \sim k^w \exp [\beta k^{4/6} + \gamma k^{3/6} + \delta k^{2/6} + \varepsilon k^{1/6}]$$

$$\beta = -3/2 \sigma(1, 1)/\sigma(2, 0)$$

$$\gamma^2 = -8/9 H_5/\sigma(2, 0)$$

$$\delta = -T_6/\sigma(2, 0)$$

$$\varepsilon = -(1/\gamma) [4 T_5 + \beta \gamma^2 \sigma(3, 0) + 3/2 \gamma^2 \sigma(2, 1)]/\sigma(2, 0)$$

$$w = -(1/\gamma) T_7/\sigma(2, 0)$$

$$2.2.2. H_5 = 0$$

Former l'équation :

$$1/6 \delta^2 \sigma(2, 0) + \delta [2/9 \beta^2 \sigma(3, 0) + 2/3 \beta \sigma(2, 1) + \sigma(1, 2)] - 1/3 \beta \sigma(2, 0) + 2/81 \beta^4 \sigma(4, 0) + 4/27 \beta^3 \sigma(3, 1) + 2/3 \beta^2 \sigma(2, 2) + 2 \beta \sigma(1, 3) + 3 \sigma(0, 4) = 0$$

$$\rightarrow \delta = \delta_1, \delta_2$$

$$2.2.2.1. \delta_1 \neq \delta_2$$

$$\rightarrow C_k \sim k^w \exp [\beta k^{2/3} + \delta k^{1/3}]$$

$$\beta = -3/2 \sigma(1, 1)/\sigma(2, 0)$$

$$\delta = \delta_f$$

$$w = -9 H_6/[2 \beta^2 \sigma(3, 0) + 3 \delta \sigma(2, 0) + 6 \beta \sigma(2, 1) + 9 \sigma(1, 2)]$$

$$2.2.2.2. \delta_1 = \delta_2$$

$$2.2.2.2.1. H_6 \neq 0$$

$$\rightarrow C_k \sim k^w \exp [\beta k^{4/6} + \delta k^{2/6} + \varepsilon k^{1/6}]$$

$$\beta = -3/2 \sigma(1, 1)/\sigma(2, 0)$$

$$\delta = -[2/3 \beta^2 \sigma(3, 0) + 2 \beta \sigma(2, 1) + 3 \sigma(1, 2)]/\sigma(2, 0)$$

$$\varepsilon^2 = -72 H_6/\sigma(2, 0)$$

$$w = 5/12 - T_8/\sigma(2, 0)$$

$$\beta = -3/2 \sigma(1, 1)/\sigma(2, 0)$$

$$\delta = -[2/3 \beta^2 \sigma(3, 0) + 2 \beta \sigma(2, 1) + 3 \sigma(1, 2)]/\sigma(2, 0)$$

$$1/2 (w^2 - w) \sigma(2, 0) + w T_8 + T_9 = 0$$

$$2.2.2.2.2. H_6 = 0$$

$$\rightarrow C_k \sim k^w \exp [\beta k^{2/3} + \delta k^{1/3}]$$

$$\text{Remarque : si } w_1 = w_2$$

$$\rightarrow C_k^{(1)} \sim C_k^{(2)} \ln k$$

RECURRENCE DE POINCARÉ DU TYPE 4

Racine $z = 1$ *simple* : $\sigma(0, 0) = 0$ et $\sigma(1, 0) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\alpha k^{3/4} + \beta k^{2/4} + \gamma k^{1/4}]$$

$$\begin{aligned}\alpha &= -4/3 \sigma(0, 1)/\sigma(1, 0) \\ \beta &= -S_4/\sigma(1, 0) \\ \gamma &= -S_5/\sigma(1, 0) \\ w &= -S_6/\sigma(1, 0)\end{aligned}$$

RECURRENCE DE POINCARÉ DU TYPE 5

Racine $z = 1$ *simple* : $\sigma(0, 0) = 0$ et $\sigma(1, 0) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\alpha k^{4/5} + \beta k^{3/5} + \gamma k^{2/5} + \delta k^{1/5}]$$

$$\begin{aligned}\alpha &= -5/4 \sigma(0, 1)/\sigma(1, 0) \\ \beta &= -S_7/\sigma(1, 0) \\ \gamma &= -S_8/\sigma(1, 0) \\ \delta &= -S_9/\sigma(1, 0) \\ w &= -S_{10}/\sigma(1, 0)\end{aligned}$$

RECURRENCE DE POINCARÉ DU TYPE 6

Racine $z = 1$ *simple* : $\sigma(0, 0) = 0$ et $\sigma(1, 0) \neq 0$

$$\rightarrow C_k \sim k^w \exp [\alpha k^{5/6} + \beta k^{4/6} + \gamma k^{3/6} + \delta k^{2/6} + \varepsilon k^{1/6}]$$

$$\begin{aligned}\alpha &= -6/5 \sigma(0, 1)/\sigma(1, 0) \\ \beta &= -S_{11}/\sigma(1, 0) \\ \gamma &= -S_{12}/\sigma(1, 0) \\ \delta &= -S_{13}/\sigma(1, 0) \\ \varepsilon &= -S_{14}/\sigma(1, 0) \\ w &= -S_{15}/\sigma(1, 0)\end{aligned}$$

$$\begin{aligned}
Q_1 &= 2/405 \alpha^3 \sigma(5, 0) + \sigma(2, 1) + 1/18 \alpha \beta \sigma(4, 0) \\
Q_2 &= 1/768 \alpha^4 \sigma(4, 0) + 1/16 \alpha^2 \sigma(2, 1) + 1/2 \sigma(0, 2) \\
Q_3 &= 9/160 \alpha^4 \sigma(5, 0) + 64/3 \sigma(1, 1) \\
Q_4 &= 2 \alpha^2 \beta^2 \sigma(4, 0) + 3/8 \alpha^4 \beta \sigma(5, 0) + 9/640 \alpha^6 \sigma(6, 0) + 16 \alpha^2 \sigma(2, 1) + 256/9 \beta \sigma(1, 1) \\
Q_5 &= 1/2 \alpha^2 \gamma \beta \sigma(4, 0) + 2/9 \alpha \beta^3 \sigma(4, 0) + 1/8 \alpha^3 \beta^2 \sigma(5, 0) + 3/64 \alpha^4 \gamma \sigma(5, 0) \\
&\quad + 9/640 \alpha^5 \beta \sigma(6, 0) + 27/71680 \alpha^7 \sigma(7, 0) - 3/8 \alpha^3 \sigma(4, 0) + 16/3 \alpha \beta \sigma(2, 1) \\
&\quad - 32/9 \gamma \sigma(1, 1) + \alpha^3 \sigma(3, 1) \\
Q_6 &= 4/3645 \alpha^5 \sigma(5, 0) + 2/9 \alpha^2 \sigma(2, 1) \\
Q_7 &= -1/32 \beta^2 \sigma(4, 0) + 1/1920 \beta^4 \sigma(5, 0) - 1/4 \sigma(2, 1) + 1/24 \beta^2 \sigma(3, 1) + \sigma(1, 2) \\
Q_8 &= -1/162 \beta^3 \sigma(4, 0) + 1/162 \beta^3 \sigma(3, 1) + \sigma(0, 3) \\
Q_9 &= 3/256 \beta \sigma(4, 0) - 7/384 \beta^2 \sigma(4, 0) + 1/768 \beta \gamma \sigma(4, 0) + 1/1536 \beta^4 \sigma(5, 0) \\
&\quad - 3/32 \sigma(2, 1) + 1/32 \beta^2 \sigma(3, 1) + 1/4 \sigma(1, 2) \\
Q_{10} &= -2/81 \alpha^3 \sigma(4, 0) + 1/81 \alpha^2 \beta^2 \sigma(4, 0) + 2/729 \alpha^4 \beta \sigma(5, 0) + 4/32805 \alpha^6 \sigma(6, 0) \\
&\quad + 2/9 \alpha \beta \sigma(2, 1) + 4/81 \alpha^3 \sigma(3, 1) + \sigma(0, 2) \\
Q_{11} &= 1/8 \beta^2 \sigma(3, 1) - 1/16 \beta^2 \sigma(4, 0) + 1/384 \beta^4 \sigma(5, 0) + \sigma(1, 2) \\
Q_{12} &= 5/128 \beta^2 \sigma(4, 0) - 1/384 \beta^4 \sigma(5, 0) + 1/46080 \beta^6 \sigma(6, 0) + 1/384 \beta^4 \sigma(4, 1) \\
&\quad - 1/16 \beta^2 \sigma(3, 1) + 1/8 \beta^2 \sigma(2, 2) + \sigma(0, 3) \\
S_1 &= 1/8 \alpha^2 \sigma(2, 0) + 1/2 \alpha \sigma(1, 1) + \sigma(0, 2) \\
S_2 &= 2/3 \alpha^2 \sigma(2, 0) + 2 \alpha \sigma(1, 1) + 3 \sigma(0, 2) \\
S_3 &= 2/9 \alpha \beta \sigma(2, 0) + 4/81 \alpha^3 \sigma(3, 0) + 2/9 \alpha^2 \sigma(2, 1) + 1/3 \beta \sigma(1, 1) + 2/3 \alpha \sigma(1, 2) \\
&\quad + \sigma(0, 3) \\
S_4 &= 9/16 \alpha^2 \sigma(2, 0) + 3/2 \alpha \sigma(1, 1) + 2 \sigma(0, 2) \\
S_5 &= 3/2 \alpha \beta \sigma(2, 0) + 9/32 \alpha^3 \sigma(3, 0) + 9/8 \alpha^2 \sigma(2, 1) + 2 \beta \sigma(1, 1) + 3 \alpha \sigma(1, 2) \\
&\quad + 4 \sigma(0, 3) \\
S_6 &= 1/8 \beta^3 \sigma(2, 0) + 3/16 \alpha \gamma \sigma(2, 0) + 27/2048 \alpha^4 \sigma(4, 0) + 9/64 \alpha^2 \beta \sigma(3, 0) \\
&\quad + 3/8 \alpha \beta \sigma(2, 1) + 1/4 \gamma \sigma(1, 1) + 9/128 \alpha^3 \sigma(3, 1) + 9/32 \alpha^2 \sigma(2, 2) + 1/2 \beta \sigma(1, 2) \\
&\quad + 3/4 \alpha \sigma(1, 3) + \sigma(0, 4) \\
S_7 &= 8/15 \alpha^2 \sigma(2, 0) + 4/3 \alpha \sigma(1, 1) + 5/3 \sigma(0, 2) \\
S_8 &= 6/5 \alpha \beta \sigma(2, 0) + 16/75 \alpha^3 \sigma(3, 0) + 3/2 \beta \sigma(1, 1) + 4/5 \alpha^2 \sigma(2, 1) + 2 \alpha \sigma(1, 2) \\
&\quad + 5/2 \sigma(0, 3) \\
S_9 &= 9/10 \beta^2 \sigma(2, 0) + 8/5 \alpha \gamma \sigma(2, 0) + 24/25 \alpha^2 \beta \sigma(3, 0) + 32/375 \alpha^4 \sigma(4, 0) \\
&\quad + 2 \gamma \sigma(1, 1) + 12/5 \alpha \beta \sigma(2, 1) + 32/75 \alpha^3 \sigma(3, 1) + 3 \beta \sigma(1, 2) + 8/5 \alpha^2 \sigma(2, 2) \\
&\quad + 4 \alpha \sigma(1, 3) + 5 \sigma(0, 4) \\
S_{10} &= 4/25 \alpha \delta \sigma(2, 0) + 6/25 \beta \gamma \sigma(2, 0) + 18/125 \alpha \beta^2 \sigma(3, 0) + 16/125 \alpha^2 \gamma \sigma(3, 0) \\
&\quad + 32/625 \alpha^3 \beta \sigma(4, 0) + 128/46875 \alpha^5 \sigma(5, 0) + 1/5 \delta \sigma(1, 1) + 9/50 \beta^2 \sigma(2, 1) \\
&\quad + 8/25 \alpha \gamma \sigma(2, 1) + 24/125 \alpha^2 \beta \sigma(3, 1) + 32/1875 \alpha^4 \sigma(4, 1) + 2/5 \gamma \sigma(1, 2) \\
&\quad + 12/25 \alpha \beta \sigma(2, 2) + 32/375 \alpha^3 \sigma(3, 2) + 3/5 \beta \sigma(1, 3) + 8/25 \alpha^2 \sigma(2, 3) \\
&\quad + 4/5 \alpha \sigma(1, 4) + \sigma(0, 5) \\
S_{11} &= 25/48 \alpha^2 \sigma(2, 0) + 5/4 \alpha \sigma(1, 1) + 3/2 \sigma(0, 2) \\
S_{12} &= 10/9 \alpha \beta \sigma(2, 0) + 125/648 \alpha^3 \sigma(3, 0) + 25/36 \alpha^2 \sigma(2, 1) + 4/3 \beta \sigma(1, 1) \\
&\quad + 5/3 \alpha \sigma(1, 2) + 2 \sigma(0, 3) \\
S_{13} &= 2/3 \beta^2 \sigma(2, 0) + 5/4 \alpha \gamma \sigma(2, 0) + 25/36 \alpha^2 \beta \sigma(3, 0) + 625/10368 \alpha^4 \sigma(4, 0) \\
&\quad + 5/3 \alpha \beta \sigma(2, 1) + 3/2 \gamma \sigma(1, 1) + 125/432 \alpha^3 \sigma(3, 1) + 25/24 \alpha^2 \sigma(2, 2) \\
&\quad + 2 \beta \sigma(1, 2) + 5/2 \alpha \sigma(1, 3) + 3 \sigma(0, 4)
\end{aligned}$$

$$\begin{aligned}
S_{14} &= 5/3 \alpha \delta \sigma(2, 0) + 2 \beta \gamma \sigma(2, 0) + 25/24 \alpha^2 \gamma \sigma(3, 0) + 10/9 \alpha \beta^2 \sigma(3, 0) \\
&+ 125/324 \alpha^3 \beta \sigma(4, 0) + 3125/155520 \alpha^5 \sigma(5, 0) + 4/3 \beta^2 \sigma(2, 1) + 2 \delta \sigma(1, 1) \\
&+ 5/2 \alpha \gamma \sigma(2, 1) + 25/18 \alpha^2 \beta \sigma(3, 1) + 625/5184 \alpha^4 \sigma(4, 1) + 10/3 \alpha \beta \sigma(2, 2) \\
&+ 3 \gamma \sigma(1, 2) + 125/216 \alpha^3 \sigma(3, 2) + 25/12 \alpha^2 \sigma(2, 3) + 4 \beta \sigma(1, 3) + 5 \alpha \sigma(1, 4) \\
&- 6 \sigma(0, 5) \\
S_{15} &= 1/8 \gamma^2 \sigma(2, 0) + 5/36 \alpha \varepsilon \sigma(2, 0) + 2/9 \beta \delta \sigma(2, 0) - 25/216 \alpha^2 \delta \sigma(3, 0) \\
&+ 4/81 \beta^3 \sigma(3, 0) + 5/18 \alpha \beta \gamma \sigma(3, 0) + 25/324 \alpha^2 \beta^2 \sigma(4, 0) + 125/2592 \alpha^3 \gamma \sigma(4, 0) \\
&+ 625/46656 \alpha^4 \beta \sigma(5, 0) + 3125/6718464 \alpha^6 \sigma(6, 0) + 5/18 \alpha \delta \sigma(2, 1) + 1/3 \beta \gamma \sigma(2, 1) \\
&+ 1/6 \varepsilon \sigma(1, 1) + 25/144 \alpha^2 \gamma \sigma(3, 1) + 5/27 \alpha \beta^2 \sigma(3, 1) + 125/1944 \alpha^3 \beta \sigma(4, 1) \\
&+ 3125/933120 \alpha^5 \sigma(5, 1) + 2/9 \beta^3 \sigma(2, 2) + 1/3 \delta \sigma(1, 2) - 5/12 \alpha \gamma \sigma(2, 2) \\
&+ 25/108 \alpha^2 \beta \sigma(3, 2) + 625/31104 \alpha^4 \sigma(4, 2) + 5/9 \alpha \beta \sigma(2, 3) + 1/2 \gamma \sigma(1, 3) \\
&+ 125/1296 \alpha^3 \sigma(3, 3) + 25/72 \alpha^2 \sigma(2, 4) + 2/3 \beta \sigma(1, 4) + 5/6 \alpha \sigma(1, 5) + \sigma(0, 6) \\
D_1 &= 9/128 \alpha^4 \sigma(4, 0) + 3/2 \alpha^2 \sigma(2, 1) - 3/4 \alpha^2 \beta \sigma(3, 0) + 2/3 \beta^2 \sigma(2, 0) - 8/3 \beta \sigma(1, 1) \\
&- 16/3 \sigma(0, 2) \\
D_2 &= 27/10240 \alpha^5 \sigma(5, 0) + 3/32 \alpha^3 \sigma(3, 1) + 3/64 \alpha^3 \beta \sigma(4, 0) - 3/32 \alpha^2 \gamma \sigma(3, 0) \\
&- 1/8 \alpha \sigma(2, 0) + \alpha \sigma(1, 2) + 1/2 \alpha \beta \sigma(2, 1) + 1/8 \alpha \beta^2 \sigma(3, 0) + 1/6 \beta \gamma \sigma(2, 0) \\
&+ 1/3 \gamma \sigma(1, 1) \\
D_3 &= 1/8 \beta^2 \sigma(3, 0) + 1/2 \beta \sigma(2, 1) - 3/8 \sigma(2, 0) + \sigma(1, 2) \\
D_4 &= 1/8 \gamma^2 \sigma(3, 0) - 1/2 \sigma(2, 0) + 1/2 \gamma \sigma(2, 1) + \sigma(1, 2) \\
D_5 &= 16/15 \alpha \beta^2 \sigma(3, 0) + 10/27 \alpha^3 \beta \sigma(4, 0) - 25/1296 \alpha^5 \sigma(5, 0) + 2 \alpha^2 \sigma(2, 1) \\
&+ 96/25 \beta \sigma(1, 1) \\
D_6 &= 32/75 \beta^3 \sigma(3, 0) + 12/5 \alpha \beta \gamma \sigma(3, 0) + 2/3 \alpha^2 \beta^2 \sigma(4, 0) + 5/12 \alpha^3 \gamma \sigma(4, 0) \\
&+ 25/216 \alpha^4 \beta \sigma(5, 0) + 125/31104 \alpha^6 \sigma(6, 0) + 24/5 \alpha \beta \sigma(2, 1) + 108/25 \gamma \sigma(1, 1) \\
&+ 5/6 \alpha^3 \sigma(3, 1) + 216/25 \sigma(0, 2) \\
D_7 &= 9/5 \alpha \gamma^2 \sigma(3, 0) + 16/5 \alpha \beta \delta \sigma(3, 0) + 48/25 \beta^2 \gamma \sigma(3, 0) + 5/9 \alpha^3 \delta \sigma(4, 0) \\
&+ 2/3 \alpha^2 \beta \gamma \sigma(4, 0) + 32/45 \alpha \beta^3 \sigma(4, 0) + 4/3 \alpha^2 \beta \gamma \sigma(4, 0) + 25/144 \alpha^4 \gamma \sigma(5, 0) \\
&+ 10/27 \alpha^3 \beta^2 \sigma(5, 0) + 25/648 \alpha^5 \beta \sigma(6, 0) + 625/653184 \alpha^7 \sigma(7, 0) + 96/25 \beta^3 \sigma(2, 1) \\
&+ 144/25 \delta \sigma(1, 1) + 36/5 \alpha \gamma \sigma(2, 1) + 4 \alpha^2 \beta \sigma(3, 1) + 25/72 \alpha^4 \sigma(4, 1) + 72/5 \alpha \sigma(1, 2) \\
D_8 &= 16/75 \beta^3 \delta \sigma(3, 0) + 4/15 \alpha \beta \varepsilon \sigma(3, 0) + 6/25 \beta \gamma^2 \sigma(3, 0) + 2/5 \alpha \gamma \delta \sigma(3, 0) \\
&+ 1/8 \alpha^2 \gamma^2 \sigma(4, 0) + 16/675 \beta^4 \sigma(4, 0) + 5/108 \alpha^3 \varepsilon \sigma(4, 0) + 2/9 \alpha^2 \beta \delta \sigma(4, 0) \\
&+ 4/15 \alpha^3 \beta^2 \sigma(4, 0) - 25/1296 \alpha^4 \delta \sigma(5, 0) + 4/81 \alpha^2 \beta^3 \sigma(5, 0) + 5/54 \alpha^3 \beta \gamma \sigma(5, 0) \\
&+ 5/648 \alpha^4 \beta^2 \sigma(6, 0) - 1/6 \alpha^2 \sigma(3, 0) + 25/5184 \alpha^5 \gamma \sigma(6, 0) + 5/972 \alpha^4 \beta^2 \sigma(6, 0) \\
&+ 125/139968 \alpha^6 \beta \sigma(7, 0) + 3125/188116992 \alpha^8 \sigma(8, 0) + 4/5 \alpha \delta \sigma(2, 1) \\
&+ 24/25 \beta \gamma \sigma(2, 1) + 12/25 \varepsilon \sigma(1, 1) + 1/2 \alpha^2 \gamma \sigma(3, 1) + 8/15 \alpha \beta^2 \sigma(3, 1) \\
&+ 5/27 \alpha \beta \sigma(4, 1) + 25/2592 \alpha^3 \sigma(5, 1) + \alpha^2 \sigma(2, 2) + 48/25 \beta \sigma(1, 2) \\
D_9 &= 27/2048 \alpha^4 \sigma(4, 0) + 9/32 \alpha^2 \sigma(2, 1) + \sigma(0, 2) \\
D_{10} &= 3/16 \alpha \beta^2 \sigma(3, 0) + 9/128 \alpha^3 \beta \sigma(4, 0) - 81/20480 \alpha^5 \sigma(5, 0) + 3/4 \alpha \beta \sigma(2, 1) \\
&+ 9/64 \alpha^3 \sigma(3, 1) + 3/2 \alpha \sigma(1, 2) \\
D_{11} &= 3/32 \alpha \beta \gamma \sigma(3, 0) - 9/128 \alpha^2 \sigma(3, 0) + 9/256 \alpha^2 \beta^2 \sigma(4, 0) + 9/512 \alpha^3 \gamma \sigma(4, 0) \\
&+ 27/4096 \alpha^4 \beta \sigma(5, 0) - 81/327680 \alpha^6 \sigma(6, 0) + 3/16 \alpha \gamma \sigma(2, 1) + 27/2048 \alpha^4 \sigma(4, 1) \\
&+ 9/64 \alpha^2 \beta \sigma(3, 1) + 9/32 \alpha^2 \sigma(2, 2) + 1/8 \beta \sigma(2, 0) \\
D_{12} &= 27/24 \gamma^2 \sigma(3, 0) + 1/9 \beta^3 \sigma(4, 0) + 9/2 \gamma \sigma(2, 1) + 9 \sigma(1, 2) \\
D_{13} &= 3 \beta \gamma \delta \sigma(3, 0) + 9/16 \gamma^3 \sigma(3, 0) + 2/3 \beta^3 \gamma \sigma(4, 0) + 27/8 \gamma^2 \sigma(2, 1) + 6 \beta \delta \sigma(2, 1) \\
&+ 4/3 \beta^3 \sigma(3, 1) + 27/2 \gamma \sigma(1, 2) + 27 \sigma(0, 3) \\
D_{14} &= 1/6 \beta \delta^2 \sigma(3, 0) - 1/3 \beta^2 \sigma(3, 0) + 1/4 \beta \gamma \varepsilon \sigma(3, 0) + 3/16 \gamma^2 \delta \sigma(3, 0) \\
&+ 1/8 \beta^2 \gamma^2 \sigma(4, 0) + 2/27 \beta^3 \delta \sigma(4, 0) + 2/405 \beta^5 \sigma(5, 0) + 1/2 \beta \varepsilon \sigma(2, 1) \\
&+ 3/4 \gamma \delta \sigma(2, 1) + 1/2 \beta^2 \gamma \sigma(3, 1) + \beta^2 \sigma(2, 2) + 3/2 \delta \sigma(1, 2)
\end{aligned}$$

$$\begin{aligned}
D_{15} &= 1/12 \delta \varepsilon^2 \sigma(3, 0) - 2/3 \delta^2 \sigma(3, 0) + 1/8 \gamma^2 \delta^2 \sigma(4, 0) + 1/2 \gamma \delta^2 \sigma(3, 1) + \delta^2 \sigma(2, 2) \\
D_{16} &= 1/384 \gamma^4 \sigma(5, 0) - 1/16 \gamma^2 \sigma(4, 0) - 1/8 \gamma \sigma(3, 1) + 1/48 \gamma^3 \sigma(4, 1) + 1/8 \gamma^2 \sigma(3, 2) \\
&\quad + 1/2 \gamma \sigma(2, 3) + \sigma(1, 4) \\
D_{17} &= \gamma^6/16080 \sigma(6, 0) + 1/3840 \gamma^5 \sigma(5, 1) - \gamma^4/384 \sigma(5, 0) + 5/128 \gamma^3 \sigma(4, 0) \\
&\quad + 1/16 \gamma \sigma(3, 1) - 1/64 \gamma^3 \sigma(4, 1) - 1/16 \gamma^2 \sigma(3, 2) + 1/384 \gamma^4 \sigma(4, 2) - 1/8 \gamma \sigma(2, 3) \\
&\quad + 1/48 \gamma^3 \sigma(3, 3) + 1/2 \gamma \sigma(1, 5) + \sigma(0, 6) + 1/8 \gamma^2 \sigma(2, 4) \\
H_1 &= \sigma(0, 3) + 1/2 \beta \sigma(1, 2) + 1/8 \beta^2 \sigma(2, 1) - 1/8 \beta \sigma(2, 0) + 1/48 \beta^3 \sigma(3, 0) \\
H_2 &= 1/384 \gamma^4 \sigma(4, 0) + 1/8 \gamma^2 \sigma(2, 2) + 1/2 \gamma \sigma(1, 3) + \sigma(0, 4) - 1/16 \gamma^2 \sigma(3, 0) \\
&\quad - 1/8 \gamma \sigma(2, 1) + 1/48 \gamma^3 \sigma(3, 1) \\
H_3 &= \gamma^3/48 \sigma(4, 0) + \gamma^2/8 \sigma(3, 1) + \gamma/2 \sigma(2, 2) + \sigma(1, 3) - \gamma/3 \sigma(3, 0) - 5/12 \sigma(2, 1) \\
H_4 &= 1/3840 \gamma^5 \sigma(5, 0) - 1/64 \gamma^3 \sigma(4, 0) + 1/16 \gamma \sigma(3, 0) - 1/16 \gamma^2 \sigma(3, 1) \\
&\quad + 1/384 \gamma^4 \sigma(4, 1) - 1/8 \gamma \sigma(2, 2) + 1/48 \gamma^3 \sigma(3, 2) + 1/8 \gamma^2 \sigma(2, 3) + 1/2 \gamma \sigma(1, 4) \\
&\quad + \sigma(0, 5) \\
H_5 &= 4/9 \beta^3 \sigma(3, 0) + 9 \sigma(0, 3) + 2 \beta^2 \sigma(2, 1) + 6 \beta \sigma(1, 2) \\
H_6 &= -1/9 \delta \sigma(2, 0) - 2/27 \beta^2 \sigma(3, 0) + 1/27 \beta \delta^2 \sigma(3, 0) - 4/243 \beta^3 \delta \sigma(4, 0) \\
&\quad + 4/3645 \beta^5 \sigma(5, 0) + 1/18 \delta^2 \sigma(2, 1) - 1/9 \beta \sigma(2, 1) + 2/27 \beta^2 \delta \sigma(3, 1) \\
&\quad + 2/243 \beta^4 \sigma(4, 1) + 2/9 \beta \delta \sigma(2, 2) + 4/81 \beta^3 \sigma(3, 2) + 2/9 \beta^2 \sigma(2, 3) + 1/3 \delta \sigma(1, 3) \\
&\quad - 2/3 \beta \sigma(1, 4) + \sigma(0, 5) \\
T_1 &= 8/15 \beta^2 \sigma(2, 0) + 5/9 \alpha^2 \beta \sigma(3, 0) + 125/2592 \alpha^4 \sigma(4, 0) + 5/6 \alpha^2 \sigma(2, 1) \\
&\quad + 8/5 \beta \sigma(1, 1) + 12/5 \sigma(0, 2) \\
T_2 &= 6/5 \beta \gamma \sigma(2, 0) + 5/8 \alpha^2 \gamma \sigma(3, 0) + 2/3 \alpha \beta^2 \sigma(3, 0) + 25/108 \alpha^3 \beta \sigma(4, 0) \\
&\quad + 125/10368 \alpha^5 \sigma(5, 0) + 2 \alpha \beta \sigma(2, 1) + 9/5 \gamma \sigma(1, 1) + 25/72 \alpha^3 \sigma(3, 1) - 3 \alpha \sigma(1, 2) \\
T_3 &= 9/10 \gamma^2 \sigma(2, 0) + 8/5 \beta \delta \sigma(2, 0) - 5/6 \alpha^3 \delta \sigma(3, 0) + 16/45 \beta^3 \sigma(3, 0) - 2 \alpha \beta \gamma \sigma(3, 0) \\
&\quad + 5/9 \alpha^2 \beta^2 \sigma(4, 0) + 25/72 \alpha^3 \gamma \sigma(4, 0) - 125/1296 \alpha^4 \beta \sigma(5, 0) + 625/186624 \alpha^6 \sigma(6, 0) \\
&\quad + 8/5 \beta^2 \sigma(2, 1) + 12/5 \delta \sigma(1, 1) + 3 \alpha \gamma \sigma(2, 1) + 5/3 \alpha^2 \beta \sigma(3, 1) + 125/864 \alpha^4 \sigma(4, 1) \\
&\quad + 5/2 \alpha^2 \sigma(2, 2) + 24/5 \beta \sigma(1, 2) + 36/5 \sigma(0, 3) \\
T_4 &= 2/15 \beta \varepsilon \sigma(2, 0) + 1/5 \gamma \delta \sigma(2, 0) - 1/12 \alpha \sigma(2, 0) + 1/8 \alpha \gamma^2 \sigma(3, 0) - 2/9 \alpha^3 \delta \sigma(3, 0) \\
&\quad + 5/72 \alpha^2 \varepsilon \sigma(3, 0) + 2/15 \beta^2 \gamma \sigma(3, 0) + 25/648 \alpha^3 \delta \sigma(4, 0) + 5/108 \alpha^2 \beta \gamma \sigma(4, 0) \\
&\quad + 4/81 \alpha \beta^3 \sigma(4, 0) + 5/54 \alpha^2 \beta \gamma \sigma(4, 0) + 125/10368 \alpha^4 \gamma \sigma(5, 0) + 25/972 \alpha^3 \beta^2 \sigma(5, 0) \\
&\quad + 125/46656 \alpha^5 \beta \sigma(6, 0) + 3125/47029248 \alpha^7 \sigma(7, 0) + 1/3 \alpha \delta \sigma(2, 1) + 2/5 \beta \gamma \sigma(2, 1) \\
&\quad + 1/5 \varepsilon \sigma(1, 1) + 5/24 \alpha^2 \gamma \sigma(3, 1) + 2/9 \alpha \beta^2 \sigma(3, 1) + 25/324 \alpha^3 \beta \sigma(4, 1) \\
&\quad + 125/31104 \alpha^5 \sigma(5, 1) + 2/3 \alpha \beta \sigma(2, 2) + 3/5 \gamma \sigma(1, 2) + 25/216 \alpha^3 \sigma(3, 2) + \alpha \sigma(1, 3) \\
T_5 &= 1/6 \delta^2 \sigma(2, 0) - 1/3 \beta \sigma(2, 0) - 2/9 \beta^2 \delta \sigma(3, 0) + 2/81 \beta^4 \sigma(4, 0) + 2/3 \beta \delta \sigma(2, 1) \\
&\quad + 4/27 \beta^3 \sigma(3, 1) + 2/3 \beta^2 \sigma(2, 2) + \delta \sigma(1, 2) + 2 \beta \sigma(1, 3) - 3 \sigma(0, 4) \\
T_6 &= 2/3 \beta^2 \sigma(3, 0) + 2 \beta \sigma(2, 1) + 3 \sigma(1, 2) \\
T_7 &= 1/9 \delta \varepsilon \sigma(2, 0) - 1/4 \gamma \sigma(2, 0) + 2/27 \beta^2 \varepsilon \sigma(3, 0) + 2/9 \beta \gamma \delta \sigma(3, 0) + 1/24 \gamma^3 \sigma(3, 0) \\
&\quad + 4/81 \beta^2 \gamma \sigma(4, 0) + 2/9 \beta \varepsilon \sigma(2, 1) + 1/3 \gamma \delta \sigma(2, 1) + 2/9 \beta^2 \gamma \sigma(3, 1) + 2/3 \beta \gamma \sigma(2, 2) \\
&\quad + 1/3 \varepsilon \sigma(1, 2) + \gamma \sigma(1, 3) \\
T_8 &= 4/81 \beta^3 \sigma(4, 0) + 2/9 \beta \delta \sigma(3, 0) + 1/3 \delta \sigma(2, 1) + 2/9 \beta^2 \sigma(3, 1) + 2/3 \beta \sigma(2, 2) \\
&\quad + \sigma(1, 3) \\
T_9 &= -1/9 \beta \delta \sigma(3, 0) - 2/81 \beta^3 \sigma(4, 0) + 1/162 \delta^3 \sigma(3, 0) + 1/81 \beta^2 \delta^2 \sigma(4, 0) \\
&\quad + 2/729 \beta^4 \delta \sigma(5, 0) + 4/32805 \beta^6 \sigma(6, 0) - 2/27 \beta^2 \sigma(3, 1) - 1/9 \delta \sigma(2, 1) \\
&\quad + 1/27 \beta \delta^2 \sigma(3, 1) + 4/243 \beta^3 \delta \sigma(4, 1) + 4/3645 \beta^5 \sigma(5, 1) + 2/243 \beta^4 \sigma(4, 2) \\
&\quad + 2/27 \beta^2 \delta \sigma(3, 2) + 1/18 \delta^2 \sigma(2, 2) - 1/9 \beta \sigma(2, 2) + 2/9 \beta \delta \sigma(2, 3) + 4/81 \beta^3 \sigma(3, 3) \\
&\quad + 2/9 \beta^2 \sigma(2, 4) + 1/3 \delta \sigma(1, 4) + 2/3 \beta \sigma(1, 5) + \sigma(0, 6)
\end{aligned}$$

CHAPITRE II

LECONS SUR LES FRACTIONS CONTINUES.

II-1. Fractions continues simples.

Historiquement, les fractions continues (fc) ont vu le jour au XVI^e siècle dans le but de créer un algorithme performant de calcul de la racine carrée d'un nombre positif A. Soit à calculer \sqrt{A} et soit a le plus grand entier dont le carré n'excède pas A. On a successivement :

$$\begin{aligned}A &= a^2 + r \\(\sqrt{A} + a)(\sqrt{A} - a) &= r \\ \sqrt{A} &= a + \frac{r}{a + \sqrt{A}}\end{aligned}$$

Cette expression peut se compliquer de proche en proche :

$$\sqrt{A} = a + \frac{r}{a + a + \frac{r}{a + \sqrt{A}}} = a + \frac{r}{2a + \frac{r}{a + \sqrt{A}}} = a + \frac{r}{2a + \frac{r}{2a + \frac{r}{\sqrt{A}}}}$$

En poursuivant indéfiniment le processus, on trouverait au moins formellement :

$$\sqrt{A} = a + \frac{r}{2a + \frac{r}{2a + \frac{r}{2a + \dots}}}$$

Une telle expression est une fraction continue que l'on préfère noter pour des raisons typographiques évidentes :

$$\sqrt{A} = a + r \rangle 2a + r \rangle 2a + r \rangle 2a + \dots$$

Ce développement est purement formel et il importe, pour le légitimer, de s'assurer que le passage à la limite est licite. Avant de procéder à l'étude théorique qui s'impose, rien n'empêche de regarder fonctionner l'algorithme suivant. Pour ce faire, on calcule les approximatifs successifs de la fc, obtenus en tronquant son expression après chaque terme $2a$. On obtient de la sorte la suite que voici :

$$a, a + \frac{r}{2a}, a + \frac{r}{2a + \frac{r}{2a}}, \dots \text{ etc } \quad \text{et on espère qu'elle converge rapidement vers } \sqrt{A}.$$

Exemple numérique : soit à calculer $\sqrt{26} = 5.09901952\dots$. On a: $A = 26, a = 5, r = 1$.
On trouve la suite: **5, 5.1, 5.09900990..., 5.09901961...** qui semble indiquer que l'algorithme est viable.

II-1.1. Théorie des fractions continues simples: calcul des approximatifs.

Considérons la récurrence linéaire d'ordre deux :

$$A_k^{(2)}C_{k+1} + A_k^{(1)}C_k + A_k^{(0)}C_{k-1} = 0$$

Divisons la par $A_k^{(0)}$ et rebaptisons ses coefficients a_k et $-b_k$ pour plus de simplicité :

$$-b_k C_{k+1} + a_k C_k + C_{k-1} = 0 \quad (\text{II-1})$$

Notons φ_k et ψ_k deux solutions linéairement indépendantes de cette récurrence. La solution générale revêt donc la forme :

$$C_k = \lambda \varphi_k + \mu \psi_k$$

Formons le quotient C_{k-1}/C_k en divisant la récurrence par C_k . On trouve successivement :

$$\begin{aligned} -C_{k-1}/C_k &= -\frac{\lambda \varphi_{k-1} + \mu \psi_{k-1}}{\lambda \varphi_k + \mu \psi_k} = a_k + \frac{b_k}{-C_k/C_{k+1}} = a_k + \frac{b_k}{a_{k+1} + \frac{b_{k+1}}{-C_{k+1}/C_{k+2}}} = \dots \\ &= a_k + \frac{b_k}{a_{k+1} + \frac{b_{k+1}}{a_{k+2} + \frac{b_{k+2}}{\ddots + \frac{b_{k+\ell}}{-C_{k+\ell}/C_{k+\ell+1}}}}} \end{aligned}$$

En règle générale, le passage à la limite, pour ℓ augmentant indéfiniment, qui consiste à négliger le dernier terme est illicite. Cela résulte clairement du fait que le premier membre devrait dépendre de λ et de μ alors que le deuxième membre n'en dépend visiblement pas. En fait, on verra lors de l'étude du théorème de Pincherle que ce passage à la limite n'est précisément autorisé que pour la seule valeur du rapport λ/μ qui rend la solution $C_k = \lambda\varphi_k + \mu\psi_k$ dominée.

Avant de présenter le théorème de Pincherle, voyons d'abord la procédure utilisée pour le calcul de la suite des approximants d'une fc. Le calcul direct d'un approximant donné est immédiat si on procède au calcul en commençant par la queue de la fraction continue et qu'on remonte vers la tête en suivant la notation polonaise inversée. Cette procédure directe n'est cependant pas utilisée si on veut connaître la suite des approximants : dans ce cas, en effet, on serait obligé de recommencer le calcul complet de chaque approximant d'ordre k sans jamais pouvoir profiter des calculs déjà effectués aux ordres inférieurs. Il est bien préférable de disposer d'un algorithme récursif de calcul des approximants. Cet algorithme existe effectivement. Il repose sur le théorème suivant :

Théorème : les approximants successifs fc_k de la fc : $fc = a_0 + b_0 \rangle a_1 + b_1 \rangle a_2 + \dots$ s'écrivent comme quotients d'un numérateur N_k par un dénominateur D_k qui obéissent tous deux à la récurrence (II-1) associée à la fc. Ces suites sont initialisées comme suit :

$$N_{-1} = D_0 = 0 \text{ et } N_0 = D_{-1} = 1.$$

Les premiers éléments de ces suites sont les suivants :

$$N_{-1} = 0, \quad N_0 = 1, \quad N_1 = a_0 / b_0, \quad N_2 = \frac{a_0 a_1 + b_0}{b_0 b_1}, \quad \dots$$

$$D_{-1} = 1, \quad D_0 = 0, \quad D_1 = 1 / b_0, \quad D_2 = \frac{a_1}{b_0 b_1}, \quad \dots$$

On a donc :

$$fc_1 = \frac{N_1}{D_1} = a_0, \quad fc_2 = \frac{N_2}{D_2} = \frac{a_0 a_1 + b_0}{a_1}, \quad fc_3 = \frac{N_3}{D_3}, \quad \dots, \quad fc_k = \frac{N_k}{D_k}, \quad \dots$$

Preuve : elle se fait par induction. Le théorème est certainement vrai pour fc_1 et fc_2 . Supposons qu'il soit vrai pour fc_n et montrons qu'il reste vrai pour fc_{n+1} .

Or fc_{n+1} se déduit de fc_n en y remplaçant a_{n-1} par $a_{n-1} + b_{n-1}/a_n$. En fait, a_{n-1} n'apparaît dans le calcul de fc_n que lorsqu'on écrit :

$$fc_n = N_n / D_n = \frac{a_{n-1} N_{n-1} + N_{n-2}}{b_{n-1}} / \frac{a_{n-1} D_{n-1} + D_{n-2}}{b_{n-1}}$$

Par conséquent, on trouve pour fc_{n+1} :

$$fc_{n+1} = [(a_{n-1} + b_{n-1}/a_n)N_{n-1} + N_{n-2}] / [(a_{n-1} + b_{n-1}/a_n)D_{n-1} + D_{n-2}]$$

$$= (a_n N_n + N_{n-1}) / (a_n D_n + D_{n-1}) = N_{n+1} / D_{n+1}$$

II-1.2. Théorème de Pincherle.

Le théorème fondamental qui régit la convergence des fractions continues est dû à Pincherle. Il s'énonce comme suit :

Théorème : soit la fraction continue : $fc = a_0 + b_1/a_1 + b_2/a_2 + \dots$

et la récurrence associée : $-b_k C_{k+1} + a_k C_k + C_{k-1} = 0$,

(i) la fc converge ssi la récurrence associée possède deux solutions contrastées à l'infini, soit $C_k^{(1)}$, dominante, et $C_k^{(2)}$, dominée. On note ρ_k le facteur de contraste :

$$\rho_k = |C_k^{(2)} / C_k^{(1)}|.$$

(ii) la valeur exacte de la fc qui converge est donnée par l'expression suivante :

$$fc = -C_{-1}^{(2)} / C_0^{(2)} \quad \text{pour autant que } C_0^{(2)} \text{ diffère de zéro.}$$

(iii) la vitesse de convergence avec laquelle fc_k tend vers fc est donnée par :

$$\left| \frac{fc_k - fc}{fc} \right| = e^{-p} = \left| C_k^{(2)} / C_k^{(1)} \right| \frac{1}{C_{-1}^{(2)}} \quad \text{avec:} \quad \begin{cases} C_0^{(1)} = 0 \\ C_{-1}^{(1)} = 1 \end{cases}$$

Son ordre de grandeur vaut donc le facteur de contraste et on peut écrire approximativement $e^{-p} \sim \rho_k$. Ce théorème est capital dans l'étude des fc. Il permet une excellente estimation de leur vitesse de convergence par simple détermination du comportement asymptotique des solutions de la récurrence associée.

Démonstration :

a) Si la fc converge, montrons que la récurrence associée possède deux solutions contrastées.

Si fc_k converge, c'est que la limite suivante existe : $\lim_{k \rightarrow \infty} N_k / D_k = \alpha$.

Posons : $X_k = N_k - \alpha D_k$ qui est solution de la récurrence associée (II-1) et formons son quotient avec une autre solution Y_k indépendante de X_k , par exemple : $Y_k = \beta N_k + \gamma D_k$ (avec: $\gamma + \alpha\beta \neq 0$).

On a :

$$\lim_{k \rightarrow \infty} (X_k / Y_k) = \lim_{k \rightarrow \infty} \frac{N_k - \alpha D_k}{\beta N_k + \gamma D_k} = \lim_{k \rightarrow \infty} \frac{(N_k / D_k) - \alpha}{\beta (N_k / D_k) + \gamma} = 0$$

ce qui montre qu'il existe une solution dominée X_k telle que $X_0 \neq 0$, vu que $X_0 = N_0 = 1$.

b) Si la récurrence associée possède une solution dominée X_k , telle que $X_0 \neq 0$, montrons que la fc converge. Dans ce cas, on peut écrire, en effet :

$$X_k = X_0 N_k + X_{-1} D_k$$

Calculons ensuite : $\lim_{k \rightarrow \infty} (X_k / D_k) = X_0 \lim_{k \rightarrow \infty} (N_k / D_k) + X_{-1} = 0$ puisque X_k est dominée. On en déduit que $\lim_{k \rightarrow \infty} (N_k / D_k)$ existe ce qui assure la convergence de la fraction continue vers la valeur $-X_{-1} / X_0$. La vitesse de convergence s'évalue sans peine en formant l'expression :

$$e^{-p} = \left| \frac{fc_k - fc}{fc} \right| = \left| -\frac{N_k}{D_k} \frac{C_0^{(2)}}{C_{-1}^{(2)}} - 1 \right| = \left| \frac{-C_{-1}^{(2)} D_k - C_0^{(2)} N_k}{C_k^{(1)} C_{-1}^{(2)}} \right| = \left| \frac{C_k^{(2)}}{C_k^{(1)} C_{-1}^{(2)}} \right|$$

où $C_k^{(1)}$ est la solution dominante initialisée comme D_k .

Concrètement, l'application de ce théorème exige que l'on soit en mesure de déterminer les comportements asymptotiques des solutions de la récurrence associée. Précisément, nous avons vu, au chapitre I, comment y parvenir lorsque les coefficients de celle-ci sont rationnels. Le paragraphe suivant étudie donc la convergence des fc à coefficients rationnels.

II-1.3. Convergence des fc à coefficients rationnels.

Jusqu'à présent, nous avons noté les fc sous la forme : $fc = a_0 + b_0 \rangle a_1 + b_1 \rangle a_2 + \dots$

et la récurrence associée sous la forme : $-b_k C_{k+1} + a_k C_k + C_{k-1} = 0$.

On peut simplifier les notations en posant :

$$C_k = X_k / \prod_{j=1}^k a_j \quad \text{et} \quad A_k = \frac{b_k}{a_k a_{k+1}}$$

La fc peut alors se réécrire, à un facteur a_0 inessentiel près, sous la forme canonique :

$$fc = 1 + A_0 \rangle 1 + A_1 \rangle 1 + \dots$$

La récurrence associée prend la forme canonique suivante :

$$-A_k X_{k+1} + X_k + X_{k-1} = 0 \tag{II-2}$$

A_k est une fraction rationnelle en la variable k , en sorte qu'il est certainement possible d'écrire son développement asymptotique, valable pour k grand, sous la forme :

$$A_k = k^\nu (\lambda_0 + \lambda_1/k + \lambda_2/k^2 + \dots)$$

où ν est un entier positif, négatif ou nul.

La récurrence associée est alors de Poincaré, type I, et on peut déterminer ses asymptotes par la méthode présentée au chapitre I. L'équation caractéristique s'obtient en ne retenant que le terme dominant dans A_k :

$$-\lambda_0 k^\nu z^2 + z + 1 = 0$$

Le comportement asymptotique de ses racines z_1 et z_2 dépend du signe de ν :

$$\text{si } \nu > 0 : z_{1,2} \approx \pm k^{-\nu/2} / \sqrt{\lambda_0}$$

$$\text{si } \nu = 0 : z_{1,2} \approx \frac{1 \pm \sqrt{1 + 4\lambda_0}}{2\lambda_0}$$

$$\text{si } \nu < 0 : z_1 \approx -1 \quad \text{et} \quad z_2 \approx \frac{k^{-\nu}}{\lambda_0}$$

Il correspond à ces trois cas des vitesses de convergence complètement différentes.

Le tableau (II-1) ci-contre livre le résultat des calculs asymptotiques dans chaque cas ainsi que les conclusions qu'on en tire au sujet des vitesses de convergence des fc correspondantes.

Détaillons à titre d'exemple le cas $\nu=0$:

si $\nu=0$, les racines de l'équation caractéristique sont données par :

$$z_{1,2} = (1 \pm \sqrt{1 + 4\lambda_0}) / (2\lambda_0)$$

On les ramène en $z=1$ en posant dans la récurrence (II-1) : $X_k = z^k \gamma^k$.

On obtient la nouvelle récurrence pour γ_k :

$$\begin{array}{l} -z^2(\lambda_0 + \lambda_1/k + \lambda_2/k^2 + \dots)\gamma_{k+1} + z\gamma_k + \gamma_{k-1} = 0 \\ \begin{array}{ccc} j=2 & j=1 & j=0 \\ a_{2,0} = -\lambda_0 z^2 & a_{1,0} = z & a_{0,0} = 1 \\ a_{2,1} = -\lambda_1 z^2 & a_{1,1} = 0 & a_{0,1} = 0 \\ a_{2,2} = -\lambda_2 z^2 & a_{1,2} = 0 & a_{0,2} = 0 \\ \vdots & nuls & nuls \end{array} \end{array}$$

d'où on tire :

$$\begin{aligned}
\sigma(0,0) &= 0 & \sigma(0,1) &= -\lambda_1 z^2 & \sigma(0,2) &= -\lambda_2 z^2 \\
\sigma(1,0) &= -2\lambda_0 z^2 + z & \sigma(1,1) &= -2\lambda_1 z^2 & \sigma(1,2) &= -2\lambda_2 z^2 \\
\sigma(2,0) &= -4\lambda_0 z^2 + z & \sigma(2,1) &= -4\lambda_1 z^2 & \sigma(2,2) &= -4\lambda_2 z^2 \\
\sigma(3,0) &= -8\lambda_0 z^2 + z & \sigma(3,1) &= -8\lambda_1 z^2 & \sigma(3,2) &= -8\lambda_2 z^2
\end{aligned}$$

La discussion qui suit est basée sur l'étude des comportements asymptotiques telle que détaillée au chapitre I :

Valeur de v	Solution dominante, X_k et dominée, Y_k	$\rho_k \approx Y_k / X_k$	Vitesse de convergence
$v = 0$ $1+4a_0 \neq 0$	$\left\{ \begin{array}{l} X_k \\ Y_k \end{array} \right\} \approx \left(\frac{1 \pm \sigma}{2a_0} \right)^k k^{a_1(-1 \mp 1/\sigma)/(2a_0)} \left\{ \begin{array}{l} P(1/k) \\ Q(1/k) \end{array} \right.$ $\sigma = \sqrt{1+4a_0}$	$\left(\frac{1-\sigma}{1+\sigma} \right)^k k^{a_1/(a_0\sigma)}$	Rapide si $1+4a_0 > 0$ Divergence sinon
$v = 0$ $1+4a_0 = 0$ $a_1 \neq 0$	$\left\{ \begin{array}{l} X_k \\ Y_k \end{array} \right\} \approx (-2)^k k^{2a_1+1/4} \exp[\pm 4\sqrt{a_1}k] \left\{ \begin{array}{l} P(1/\sqrt{k}) \\ Q(1/\sqrt{k}) \end{array} \right.$	$\exp[-8\sqrt{a_1}k]$	Bonne si $a_1 > 0$ Divergence sinon
$v = 0$ $1+4a_0 = 0$ $a_1 = 0$ $1+16a_2 \neq 0$	$\left\{ \begin{array}{l} X_k \\ Y_k \end{array} \right\} \approx (-2)^k k^{(1 \pm \tau)/2} \left\{ \begin{array}{l} P(1/k) \\ Q(1/k) \end{array} \right.$ $\tau = \sqrt{1+16a_2}$	$k^{-\tau}$	Lente si $1+16a_2 > 0$ Divergence sinon
$v = 0$ $1+4a_0 = 0$ $a_1 = 0$ $1+16a_2 = 0$	$X_k \approx (-2)^k \sqrt{k} [\ln k \ P(1/k) + Q(1/k)]$ $Y_k \approx (-2)^k \sqrt{k} \ R(1/k)$	$\frac{1}{\ln k}$	Ultra lente
$v = -2, -3, \dots$	$X_k \approx k^{v-a_1/a_0} k!^{-v} a_0^{-k} P(1/k)$ $Y_k \approx (-1)^k Q(1/k)$	$(-a_0)^k k!^v k^{a_1/a_0-v}$	Ultra rapide si a_0 modéré
$v = -1$	$X_k \approx k^{a_0-1-a_1/a_0} k! a_0^{-k} P(1/k)$ $Y_k \approx (-1)^k k^{-a_0} Q(1/k)$	$k^{1-2a_0+a_1/a_0} (-a_0)^k / k!$	Rapide si a_0 modéré
$v = 1$	$\left\{ \begin{array}{l} X_k \\ Y_k \end{array} \right\} \approx \frac{k^{1/4-a_1/(2a_0)}}{(\pm\sqrt{a_0})^k \sqrt{k}!} \exp[\pm\sqrt{k/a_0}] \left\{ \begin{array}{l} P(1/\sqrt{k}) \\ Q(1/\sqrt{k}) \end{array} \right.$	$(-1)^k \exp[-2\sqrt{k/a_0}]$	Bonne si $a_0 > 0$ Divergence sinon
$v = 2$	$\left\{ \begin{array}{l} X_k \\ Y_k \end{array} \right\} \approx \frac{k^{1/2-a_1/(2a_0) \pm 1/(2\sqrt{a_0})}}{(\pm\sqrt{a_0})^k k!} \left\{ \begin{array}{l} P(1/\sqrt{k}) \\ Q(1/\sqrt{k}) \end{array} \right.$	$(-1)^k k^{-1/\sqrt{a_0}}$	Lente si $a_0 > 0$ Divergence sinon
$v = 3, 4, \dots$	$\left\{ \begin{array}{l} X_k \\ Y_k \end{array} \right\} \approx \frac{k^{v/4-a_1/(2a_0)}}{(\pm\sqrt{a_0})^k \sqrt{k}!^v} \left\{ \begin{array}{l} P(1/\sqrt{k}) \\ Q(1/\sqrt{k}) \end{array} \right.$	$(-1)^k$	Divergence oscillatoire

Table II-1 : comportements asymptotiques des solutions de la récurrence associée à une fraction continue à coefficients rationnels et vitesse de convergence de cette fraction continue. Le

nombre de chiffres népériens corrects est donné par la relation $p \approx -\ln|\rho_k|$. P, Q et R sont des polynômes.

1er cas : $1 + 4\lambda_0 \neq 0$: les racines sont simples, d'où $\gamma_k \sim k^w$ avec $w = -\sigma(0,1)/\sigma(1,0)$.

Les deux asymptotes s'écrivent donc :

$$\begin{aligned} X_k^{(1)} &= z_1^k k^{w_1} & (dominante) \\ X_k^{(2)} &= z_2^k k^{w_2} & (dominée) \end{aligned}$$

et leur contraste vaut : $\rho_k \approx |z_2/z_1|^k k^{w_2-w_1}$

Les valeurs des coefficients w_1 et w_2 se déduisent de l'examen des tables (I-4).

La fc ne converge que si $1 + 4\lambda_0 > 0$, soit dans le cas des racines réelles et distinctes.

2e cas : $1 + 4\lambda_0 = 0$: la racine $z = 1/(2\lambda_0)$ est double dans ce cas. Il faut alors discuter selon que $\sigma(0,1)$ est nul ou non, ce qui revient à faire de même sur λ_1 .

a) $\lambda_1 \neq 0$: $\gamma_k \approx k^w \exp(\alpha\sqrt{k})$ où $\alpha^2 = -8\sigma(0,1)/\sigma(2,0) = 16\lambda_1$

d'où : $X_k \approx (-2)^k k^w \exp(\pm 4\sqrt{\lambda_1 k})$

et le contraste vaut : $\rho_k \approx \exp(-8\sqrt{\lambda_1 k})$ (convergence si $\lambda_1 > 0$ uniquement)

b) $\lambda_1 = 0$ et $1 + 16\lambda_2 \neq 0$: $\gamma_k \sim k^w$

avec : $w(w-1)\sigma(2,0) + 2w\sigma(1,1) + 2\sigma(0,2) = 0$

d'où $w^2 - w - 4\lambda_2 = 0$ soit $w_{1,2} = (1 \pm \sqrt{1 + 16\lambda_2})/2$.

On trouve, cette fois, : $X_k^{(1,2)} \approx (-2)^k k^{w_{1,2}}$

et le contraste vaut : $\rho_k \approx k^{w_2-w_1} = k^{-\sqrt{1+16\lambda_2}}$ (convergence si $1 + 16\lambda_2 > 0$)

c) $\lambda_1 = 0$ et $1 + 16\lambda_2 = 0$: $\gamma_k^{(2)} \sim k^{1/2}$ (dominée) et $\gamma_k^{(1)} \sim k^{1/2} \ln k$ (dominante) et le contraste vaut : $\rho_k \approx 1/\ln(k)$.

En résumé, l'examen du tableau (II-1), révèle l'existence de fractions continues présentant des vitesses de convergence complètement différentes allant du très rapide, en $p \approx k \ln k$, au très lent, en $p \approx \ln(\ln k)$, en passant par tous les stades intermédiaires, $p \approx k$, $p \approx \sqrt{k}$ et enfin $p \approx \ln k$, sans oublier tous les cas de divergence.

Exemple numérique : Discutons, à titre d'exemple, le cas de la fc : $fc = I + A_0 \rangle I + A_1 \rangle I + \dots$
où les coefficients A_k sont donnés par :

$$A_k = -\frac{1}{4} \frac{k^2 + \beta}{k^2 + 1} \quad (\beta \text{ réel quelconque})$$

On développe sans peine A_k sous la forme asymptotique suivante :

$$A_k \approx -\frac{1}{4} + \frac{0}{k} - \frac{1}{4} \frac{\beta - 1}{k^2} + \dots$$

On se trouve donc dans le cas : $\nu = 0, 1 + 4\lambda_0 = 0, \lambda_1 = 0, \lambda_2 = -\frac{1}{4}(\beta - 1)$

La fc converge à la vitesse $\begin{cases} p \approx \sqrt{5-4\beta} \ln k & \text{si } \beta < 5/4 \\ p \approx \ln(\ln k) & \text{si } \beta = 5/4 \end{cases}$

Elle diverge si $\beta > 5/4$.

Supposons $\beta = -11$: la fc s'obtient avec 10 chiffres exacts ($p = 23$) après $e^{23/7} \approx 27$ approximants. Si $\beta = -1$, il en faut déjà nettement plus car : $e^{23/3} \approx 2136$. Si $\beta = 1$, il en faut $9.7 \cdot 10^9$! Enfin, le cas $\beta = 1.25$ est particulièrement dramatique puisque $\exp(9.7 \cdot 10^9)$ approximants sont nécessaires (1 suivi de 4 milliards de zéros !!). On conçoit, à la vue de tels exemples, que les méthodes d'accélération de la convergence soient de toute première utilité.

II-2. Intérêt des fractions continues simples.

Dans l'immense majorité des applications pratiques, les fc que l'on rencontre sont à coefficients rationnels et réels d'où l'intérêt de la table (II-1) qui renseigne immédiatement sur leur vitesse de convergence. On se trouve heureusement très souvent dans le domaine où la convergence est rapide. Il en résulte un grand nombre d'algorithmes efficaces pour le calcul des fonctions élémentaires et transcendentes dont certains sont effectivement utilisés par les ordinateurs modernes. Passons-en quelques-uns en revue.

II-2.1. Racine carrée.

Revenons à la fc définissant la racine carrée de $A = a^2 + r$ (où a^2 est le plus grand carré parfait contenu dans A) :

$$\sqrt{A} = a + r \rangle 2a + r \rangle 2a + \dots$$

La récurrence associée s'écrit : $-rC_{k+1} + 2aC_k + C_{k-1} = 0$

ou encore :
$$-\frac{r}{4a^2} X_{k+1} + X_k + X_{k-1} = 0$$

en posant $C_k = X_k/(2a)^k$.

On est dans le cas : $v = 0$, $1 + 4\lambda_0 = 1 + r/a^2 \neq 0$, d'où la vitesse de convergence :

$$p \approx k \ln \left| \frac{a + \sqrt{a^2 + r}}{a - \sqrt{a^2 + r}} \right| = k \ln \frac{2a^2 + r + 2a\sqrt{a^2 + r}}{r}$$

Le tableau (II-2), qui suit, explicite quelques exemples simples:

A = 2	a = 1	r = 1	p ≈ 1.76 k
A = 3	a = 1	r = 2	p ≈ 1.31 k
A = 5	a = 2	r = 1	p ≈ 2.76 k
A = 6	a = 2	r = 2	p ≈ 2.29 k
A = 7	a = 2	r = 3	p ≈ 1.97 k
A = 8	a = 2	r = 4	p ≈ 1.76 k
A = 10	a = 3	r = 1	p ≈ 3.58 k
A = 11	a = 3	r = 2	p ≈ 2.99 k

Table II-2

Vu que : $r \leq 2a$, on note que p est d'autant meilleur que a (donc aussi A) est élevé et que r est petit. Cette remarque permet d'améliorer l'algorithme grâce à l'astuce suivante. Soit à calculer $\sqrt{2} = 1.414213562\dots$.

L'algorithme simple fournit le résultat suivant (3ème approximant de la fc avec a=1 et r=1):

$$fc_3 = 1 + \frac{1}{2 + 1/2} = 1.4$$

Il y a intérêt à calculer en lieu et place la quantité équivalente : $10^{-1}\sqrt{200}$. On a, cette fois :

a = 14 et r = 4 d'où :

$$fc_3 = 14 + \frac{4}{28 + 4/28} = 14.14213198\dots$$

d'où on tire $\sqrt{2}$ avec 7 chiffres corrects.

On note que les prédictions à caractère asymptotique sont déjà très bonnes, même pour des valeurs très basses de k. On trouve, en effet, dans l'exemple ci-dessus p = k 5.28. Si k = 3, cela donne 15.86 chiffres népériens soit 7 chiffres décimaux.

Pour information, il existe d'autres algorithmes, de type itératif, qui conviennent pour le calcul de \sqrt{A} . Par exemple :

$$x_{n+1} = \frac{x_n^2 + A}{2x_n} \quad \text{initialisé en } x_0 \text{ donné.}$$

La formule suivante, plus compliquée, est encore plus performante :

$$x_{n+1} = \frac{x_n(x_n^2 + 3A)}{3x_n^2 + A} \quad (x_0 \text{ donné})$$

II-2.2. Tangente trigonométrique.

Gauss a étudié la fc suivante et a montré qu'elle tendait vers la tangente trigonométrique :

$$tg(z) = z \{ 1 - z^2 \} 3 - z^2 \} 5 - \dots$$

Transformons-la légèrement afin de retrouver les notations standards :

$$-1 - z \{ 1 - z^2 \} 3 - z^2 \} 5 - \dots = a_0 + \frac{b_0}{a_1 + \frac{b_1}{a_2 + \dots}}$$

où on doit poser : $a_k = 2k - 1$ et $b_k = -z^2$. On trouve, dans ces conditions,

$$A_k = -\frac{z^2}{4k^2 - 1} \approx -\frac{z^2}{4} \frac{1}{k^2} \left(1 + \frac{1}{4k^2} + \dots \right)$$

On se trouve dans le cas : $v = -2$, $\lambda_0 = -z^2/4$, $\lambda_1 = 0$ d'où :

$$e^{-p} \approx \left(\frac{z^2}{4} \right)^k \frac{k^2}{k!^2}$$

On constate deux choses :

- pour z modéré (< 2), la convergence est très rapide, de type factorielle. Par exemple $tg(\pi/4)$ est correct avec 11 chiffres à l'ordre $k = 6$.
- pour $z > 2$, la convergence se détériore aux petites valeurs de k . Il en résulte que l'utilisation optimale de cet algorithme requiert la réduction préalable de l'argument au premier quadrant.

II-2.3. Fonction hypergéométrique de GAUSS.

Rappelons la définition de la fonction hypergéométrique de GAUSS :

$${}_2F_1(a, b; c; z) = 1 + \frac{ab}{c} \frac{z}{1!} + \frac{a(a+1)b(b+1)}{c(c+1)} \frac{z^2}{2!} + \dots = \sum_{n=0}^{\infty} \frac{\Gamma(n+a)\Gamma(n+b)\Gamma(c)}{\Gamma(a)\Gamma(b)\Gamma(n+c)} \frac{z^n}{n!}$$

Un grand nombre de fonctions élémentaires et transcendentes sont des cas particuliers de cette fonction générale. On a, par exemple :

$$\begin{aligned} \frac{1}{z} \ln(1+z) &= F(1, 1; 2; -z) \\ \frac{1}{z} \operatorname{arctg} z &= F\left(\frac{1}{2}, 1; \frac{3}{2}; -z^2\right) \\ \frac{1}{z} \operatorname{arcsin} z &= F\left(\frac{1}{2}, \frac{1}{2}; \frac{3}{2}; z^2\right) \\ \frac{\sin(2a-1)z}{(2a-1)\sin z} &= F\left(a, 1-a; \frac{3}{2}; \sin^2 z\right) \\ \frac{\cos(2a-1)z}{\cos z} &= F\left(a, 1-a; \frac{1}{2}; \sin^2 z\right) \end{aligned}$$

Gauss a montré que la fonction hypergéométrique satisfait l'équation différentielle :

$$z(1-z)w'' + [c - (a+b+1)z]w' - abw = 0$$

La même fonction obéit à la relation récurrente suivante :

$$[c-1+(a+b+1-2c)z]F(a, b; c; z) = (c-1)(1-z)F(a, b; c-1; z) - \frac{1}{c}(c-a)(c-b)zF(a, b; c+1; z)$$

En d'autres termes, la suite : $u_n = F(a, b; c+n; z)$ obéit à la récurrence d'ordre deux :

$$(c+n-a)(c+n-b)zu_{n+1} + (c+n)[c+n-1+(a+b+1-2c-2n)z]u_n + (z-1)(c+n-1)(c+n)u_{n-1} = 0 \quad (\text{II-3})$$

Plus simplement, la suite $v_n = F(a, b; c+n; z)/\Gamma(n+c)$ obéit à :

$$z(c+n-a)(c+n-b)v_{n+1} + [c+n-1+(a+b+1-2c-2n)z]v_n - (1-z)v_{n-1} = 0$$

Montrons que v_n est solution dominée de cette récurrence à condition que l'on ait $|1-z^{-1}| > 1$, ce qui entraîne que u_n est à son tour solution dominée de (II-3). Celle-ci est régulière de type I. Son équation caractéristique s'écrit : $t^2z + (1-2z)t + (z-1) = 0$. Ses racines valent : $t_1 = 1$ et $t_2 = (z-1)/z$. La recherche des asymptotes se fait à partir du tableau classique :

$j = 2$ $a_{2,0} = t^2 z$ $a_{2,1} = (2c - a - b)t^2 z$ $a_{2,2} = (c - a)(c - b)t^2 z$	$j = 1$ $a_{1,0} = (1 - 2z)t$ $a_{1,1} = [2c - 1 + (a + b + 1 - 4c)z]t$ $a_{1,2} = c[c - 1 + (a + b + 1 - 2c)z]t$	$j = 0$ $a_{0,0} = z - 1$ $a_{0,1} = (2c - 1)(z - 1)$ $a_{0,2} = c(c - 1)(z - 1)$
--	--	--

On calcule sans peine :

$$\sigma(1,0) = 2t^2 z + (1 - 2z)t$$

$$\sigma(0,1) = (2c - a - b)t^2 z + [2c - 1 + (a + b + 1 - 4c)z]t + (2c - 1)(z - 1)$$

Pour chaque racine (simple), on peut écrire : $u_n \sim n^w$ avec $w = -\sigma(0,1)/\sigma(1,0)$.

- Racine $t_1 = 1$: on trouve : $u_n^{(1)} \approx 1$

- Racine $t_2 = (z-1)/z$: on trouve : $u_n^{(2)} \approx \left(\frac{z-1}{z}\right)^n n^{(a+b-1)(z-1)/z}$

Il est clair que la fonction hypergéométrique considérée, à savoir :

$$u_n = F(a, b; c + n; z) = 1 + \frac{ab}{c + n} \frac{z}{1!} + \dots$$

est asymptotique à l'unité aux grandes valeurs de n , donc $u_n \sim u_n^{(1)}$. Ce comportement n'est dominé que si : $|(z-1)/z| = |1 - z^{-1}| > 1$. Ce n'est donc qu'à cette condition que le théorème de Pincherle autorise le développement en fc du quotient de deux fonctions hypergéométriques de troisième indice consécutif. Sous cette condition, on a le développement suivant :

$$-(c-1) \frac{F(a, b; c-1; z)}{F(a, b; c; z)} = a_0 + \frac{b_0}{a_1 + \frac{b_1}{a_2 + \dots}}$$

avec les coefficients suivants :

$$\begin{cases} a_n = [(c + n - 1) + (a + b + 1 - 2c - 2n)z]/(z - 1) \\ b_n = -\frac{z}{z - 1} (n + c - a)(n + c - b) \end{cases}$$

Le facteur de contraste définit la vitesse de convergence, il vaut :

$$\rho_k \approx \frac{1}{\left|1 - z^{-1}\right|^n \left|n^{(a+b-1)(z-1)/z}\right|} = e^{-p}$$

Application numérique : illustrons la théorie qui précède par le calcul de la cotangente. On part de l'identité :

$$(2a-1) \operatorname{tg}(z) \cot g[(2a-1)z] = \frac{F(a, 1-a; \frac{1}{2}; \sin^2 z)}{F(a, 1-a; \frac{3}{2}; \sin^2 z)}$$

Ce quotient vaut par la théorie qui précède :

$$-2fc = -2[a_0 \rangle a_1 + b_1 \rangle a_2 + \dots] \quad \text{où :}$$

$$\begin{cases} a_n = (n + \frac{1}{2})(\operatorname{tg}^2 z - 1) \\ b_n = (n + \frac{3}{2} - a)(n + \frac{1}{2} + a)\operatorname{tg}^2 z \end{cases}$$

Posant : $(2a-1)z = u$, on trouve :

$$\cot g(u) = -\frac{2z}{u} \cot g(z) [a_0 \rangle a_1 + b_1 \rangle a_2 + \dots] \quad \text{avec :}$$

$$\begin{cases} a_n = (n + \frac{1}{2})(\operatorname{tg}^2 z - 1) \\ b_n = \operatorname{tg}^2 z \left[(n+1)^2 - \frac{u^2}{4z^2} \right] \end{cases}$$

La valeur de z doit être choisie en sorte que la convergence soit rapide. Or celle-ci vaut :

$$e^{-p} \approx |\rho_k| \approx |\operatorname{tg}^{2k} z|$$

$$\text{soit : } p \approx k \ln |\cot g^2 z|$$

Par exemple, si on choisit $z = \pi/6$, on trouve : $p \sim k \, 1.0986$.

Si on choisit $z = \pi/12$, on trouve : $p \sim k \, 2.6339$.

La théorie qui précède souffre de l'inconvénient majeur de n'être applicable qu'au quotient de deux fonctions hypergéométriques d'indice consécutif (ici le troisième, c, mais des formules similaires existent pour les autres indices a et b). On souhaiterait disposer d'une procédure algorithmique plus générale, capable de s'appliquer directement au développement en série de la fonction hypergéométrique et d'ailleurs plus généralement à n'importe quel développement en série de MacLaurin. Une telle procédure existe effectivement et porte le nom d'algorithme qd.

II-3. Algorithme qd.

Soit la fonction $f(z)$, définie par son développement en série, valable au voisinage de l'origine:

$$f(z) = \sum_0^{\infty} \gamma_k z^k \quad (\text{II-4})$$

On a que la fc associée suivante, si elle converge, converge vers la même limite $f(z)$:

$$\sum_0^{\infty} \gamma_k z^k = \frac{\gamma_0}{1 - \frac{q_1^{(0)} z}{1 - \frac{e_1^{(0)} z}{1 - \frac{q_2^{(0)} z}{1 - \frac{e_2^{(0)} z}{1 - \ddots}}}}} \quad (\text{II-5})$$

Le calcul des quantités q et e s'effectue récursivement grâce aux règles de l'algorithme qd. On part des conditions initiales:

$$\begin{cases} e_0^{(n)} = 0 & (n = 0, 1, \dots) \\ q_1^{(n)} = \gamma_{n+1} / \gamma_n & (n = 0, 1, \dots) \end{cases}$$

Puis on calcule, de proche en proche, :

$$\begin{cases} e_k^{(n)} = e_{k-1}^{(n+1)} + q_k^{(n+1)} - q_k^{(n)} \\ q_{k+1}^{(n)} = e_k^{(n+1)} q_k^{(n+1)} / e_k^{(n)} \end{cases} \quad (k = 1, 2, \dots; n = 0, 1, \dots)$$

Pour la commodité des calculs, on peut disposer les symboles q et e dans un tableau triangulaire de ce type :

$e_0^{(0)} = 0$	$q_1^{(0)} = \gamma_1/\gamma_0$	$e_1^{(0)}$	$q_2^{(0)}$	$e_2^{(0)}$	$q_3^{(0)}$	$e_3^{(0)}$
$e_0^{(1)} = 0$	$q_1^{(1)} = \gamma_2/\gamma_1$	$e_1^{(1)}$	$q_2^{(1)}$	$e_2^{(1)}$	$q_3^{(1)}$	
$e_0^{(2)} = 0$	$q_1^{(2)} = \gamma_3/\gamma_2$	$e_1^{(2)}$	$q_2^{(2)}$	$e_2^{(2)}$		
$e_0^{(3)} = 0$	$q_1^{(3)} = \gamma_4/\gamma_3$	$e_1^{(3)}$	$q_2^{(3)}$			
$e_0^{(4)} = 0$	$q_1^{(4)} = \gamma_5/\gamma_4$	$e_1^{(4)}$				
$e_0^{(5)} = 0$	$q_1^{(5)} = \gamma_6/\gamma_5$			*	*	
$e_0^{(6)} = 0$			*	*		

Table II-3

On voit que l'algorithme relie entre elles les quantités qui figurent aux sommets d'un losange. Seul le coin supérieur gauche du tableau peut être construit de cette façon. Voyons de plus près à quoi ressemblent les premiers approximants de la fc associée :

- 1er approximant :
$$\frac{\gamma_0}{1 - q_1^{(0)}z} = \frac{\gamma_0}{1 - \gamma_1 z / \gamma_0} \approx \gamma_0 + \gamma_1 z + O(z^2)$$

C'est une fraction rationnelle, notée [0,1] pour rappeler qu'elle consiste dans le quotient de deux polynômes en z de degré 0 et 1 respectivement. Elle se comporte, au voisinage de l'origine, comme la série de MacLaurin de départ. On caresse l'espoir que la fc converge, loin de l'origine, plus vite que cette série.

- 2ème approximant :

$$\begin{aligned} \frac{\gamma_0}{1 - \frac{q_1^{(0)}z}{1 - e_1^{(0)}z}} &= \frac{\gamma_0(1 - e_1^{(0)}z)}{1 - (e_1^{(0)} + q_1^{(0)})z} \\ &\approx \gamma_0(1 - e_1^{(0)}z) [1 + (e_1^{(0)} + q_1^{(0)})z + (e_1^{(0)} + q_1^{(0)})^2 z^2 + O(z^3)] \\ &\approx \gamma_0 + \gamma_0(e_1^{(0)} + q_1^{(0)} - e_1^{(0)})z + \gamma_0(e_1^{(0)} + q_1^{(0)})^2 z^2 - \gamma_0 e_1^{(0)}(e_1^{(0)} + q_1^{(0)})z^2 + O(z^3) \\ &\approx \gamma_0 + \gamma_1 z + \gamma_2 z^2 + O(z^3) \end{aligned}$$

C'est une fraction rationnelle, notée [1,1] qui se comporte à nouveau comme la série de départ, au voisinage de l'origine, mais cette fois jusqu'au terme en z^2 inclus.

Nous admettons, sans démonstration, que ces conclusions restent vraies à tous les ordres d'approximation : l'algorithme construit une fc dont les approximants successifs sont des fractions rationnelles notées : [0,0], [0,1], [1,1], [1,2], [2,2], ... qui adoptent, près de l'origine, le même développement en série de puissance de z que la fonction f(z) donnée. Sous certaines conditions, que nous avons appris à reconnaître grâce au théorème de Pincherle, cette fc converge vers f(z) plus vite que ne le faisait son développement de MacLaurin.

La fc associée écrite sous la forme (II-5) est inconvenue à programmer numériquement du fait que son écriture fait intervenir deux suites distinctes, $e_k^{(0)}$ et $q_k^{(0)}$. Il est plus agréable de contracter cette fc sous la forme :

$$\sum_0^\infty \gamma_k z^k = \frac{\gamma_0}{1 - \alpha_0 z - \frac{\beta_1 z^2}{1 - \alpha_1 z - \frac{\beta_2 z^2}{1 - \alpha_2 z - \ddots}}}$$

dont la suite des approximants engendre des fractions rationnelles du type [n-1, n].

Les symboles α_k et β_k sont définis par les relations :

$$\alpha_k = e_k^{(0)} + q_{k+1}^{(0)} \quad \text{et} \quad \beta_k = e_k^{(0)} q_k^{(0)} \quad (\text{II-6})$$

II-3.1. Solutions exactes de l'algorithme qd.

La fonction $f(z)$ étant donnée par son développement de MacLaurin (II-4), on construit généralement la fc associée en développant l'algorithme qd par voie numérique. Dans un certain nombre de cas importants, l'algorithme est soluble exactement par voie analytique. Ce résultat est précieux dans la mesure où il permet d'éviter l'instabilité numérique qui caractérise l'algorithme. Passons en revue quelques cas de solubilité exacte. Une classe, très générale, est fournie par l'expression suivante des coefficients γ_n :

$$\gamma_n = \prod_{j=0}^n \frac{A - q^{\alpha+j}}{B - q^{\gamma+j}} \quad (\text{II-7})$$

Les cas particuliers suivants sont importants :

Exemple 1 : $\gamma_n = \Gamma(n + a)$

Poser dans la formule (II-7) $A = 1$, $B = 1 - q$, $\alpha = a - 1$ puis faire $\gamma \rightarrow -\infty$, $q \rightarrow 1$. On trouve :

$$q_k^{(n)} = n + k + a - 1 \quad \text{et} \quad e_k^{(n)} = k$$

d'où :

$$\alpha_k = 2k + a \quad \text{et} \quad \beta_k = k(k + a - 1)$$

Certes, on pourrait se demander quel intérêt il peut y avoir à considérer une fonction définie par la série :

$$f(z) = \sum_0^{\infty} \Gamma(n + a) z^n$$

dont le rayon de convergence est nul. Pourtant, il n'est pas rare, en physique, qu'un traitement perturbatif fasse apparaître un développement de ce type. Tout le problème consiste alors à extraire l'information qui est contenue dans la série divergente. Un exemple simple est le suivant : en intégrant illicitement, terme à terme, la fonction suivante, on trouve :

$$I = \int_0^{\infty} \frac{e^{-t} dt}{t+1} = 0.596347... = \int_0^{\infty} \sum_0^{\infty} (-t)^k e^{-t} dt = \sum_0^{\infty} (-1)^k k!$$

Cette série est du type $\sum_0^{\infty} z^n \Gamma(n + a)$ avec $z = -1$ et $a = 1$. Le point important est que, bien

que la permutation illicite des passages à la limite ait eu pour conséquence de remplacer l'intégrale I qui existe par une série divergente qui n'existe pas, l'information que constitue la valeur de I reste présente dans les sommes partielles de la série divergente.

Précisément, l'algorithme qd est capable de décrypter cette information en retrouvant la valeur de I . L'algorithme qd est soluble exactement dans ce cas et il donne :

$$I = \frac{\gamma_0}{1 - \alpha_0 z - \frac{\beta_1 z^2}{1 - \alpha_1 z - \dots}}$$

avec : $\alpha_k = 2k + 1$ et $\beta_k = k^2$ soit :

$$I = \frac{1}{2 - \frac{1}{4 - \frac{4}{6 - \frac{9}{8 - \dots}}}}$$

Par exemple, l'approximant représenté vaut 0.593301... . La vitesse de convergence se détermine le plus aisément en posant que la récurrence associée à I^{-1} s'écrit :

$$-b_k C_{k+1} + a_k C_k + C_{k-1} = 0 \quad \text{avec} \quad \begin{cases} a_k = 2k + 2 \\ b_k = -(k + 1)^2 \end{cases}$$

d'où :

$$A_k = \frac{b_k}{a_k a_{k+1}} = \frac{-(k + 1)}{4(k + 2)} \approx -\frac{1}{4} + \frac{1}{4k} + O(1/k^2)$$

Se reportant à la table (II-1), on voit que l'on se trouve dans le cas:
 $v = 0, \lambda_0 = -1/4, \lambda_1 = 1/4 \neq 0$.

Le facteur de contraste définit la vitesse de convergence qui vaut :

$$e^{-p} \approx \rho_k \approx \exp(-4\sqrt{k})$$

Exemple 2 : $\gamma_n = 1 / \Gamma(n + a)$.

On trouve :

$$q_k^{(n)} = \frac{n + a + k - 2}{(n + a + 2k - 3)(n + a + 2k - 2)}$$

$$e_k^{(n)} = \frac{-k}{(n + a + 2k - 2)(n + a + 2k - 1)}$$

d'où on tire α_k et β_k grâce aux relations (II-6).

La fonction exponentielle correspond à ce cas avec $a = 1$ puisque l'on a : $e^z = \sum_0^{\infty} z^n / n!$.

On en déduit l'approximation suivante pour cette fonction.

Remarquons d'abord que si a vaut 1, on a :

$$q_k^{(n)} = \frac{n+k-1}{(n+2k-2)(n+2k-1)} \quad e_k^{(n)} = \frac{-k}{(n+2k-1)(n+2k)}$$

On trouve en conséquence :

$$e^z = \frac{\gamma_0}{1 - \alpha_0 z - \frac{\beta_1 z^2}{1 - \alpha_1 z - \frac{\beta_2 z^2}{1 - \alpha_2 z - \ddots}}}$$

avec :

$$\alpha_k = -1/(4k^2 - 1) \quad \beta_k = -\frac{1}{4}/(2k-1)^2 \quad \text{sauf}^* \quad \beta_1 = -1/2$$

L'exception marquée d'une * provient curieusement du fait que :

$$\lim_{k \rightarrow 1} \lim_{n \rightarrow 0} q_k^{(n)} e_k^{(n)} = -1/4 \quad \text{tandis que} : \quad \lim_{n \rightarrow 0} \lim_{k \rightarrow 1} q_k^{(n)} e_k^{(n)} = -1/2$$

Par exemple, si $z = 1$, on trouve :

$$e^1 = 2.71828... = \frac{1}{0 + \frac{1/2}{4/3 + \frac{1/36}{16/15 + \ddots}}}$$

où l'approximant représenté vaut **2.71875**...

La vitesse de convergence est élevée ainsi qu'on peut le voir en formant la récurrence associée:

$$-\frac{1}{4} z^2 / (2k+1)^2 C_{k+1} + [1 + z/(4k^2 - 1)] C_k + C_{k-1} = 0$$

$$A_k = b_k / (a_k a_{k+1}) = \frac{z^2}{4(2k+1)^2 \left(1 + \frac{z}{4k^2-1}\right) \left(1 + \frac{z}{4k^2+8k+3}\right)} \approx \frac{z^2}{16} k^{-2}$$

Le facteur de contraste définit la vitesse de convergence (cfr table (II-1) , v = -2) :

$$e^{-p} \approx (z^2 / 16)^k k^w / k!^2$$

Peu importe la valeur de w, en fait w = 1, la convergence est de type factorielle.

Exemple 3 : $\gamma_n = \Gamma(n+a) / \Gamma(n+c)$

Poser dans la formule (II-7) : A = 1, B = 1, $\alpha = a - 1$, $\gamma = c - 1$ puis faire $q \rightarrow 1$. On trouve :

$$q_k^{(n)} = \frac{(a+n+k-1)(c+n+k-2)}{(c+n+2k-2)(c+n+2k-3)}$$

$$e_k^{(n)} = \frac{k(k+c-a-1)}{(c+n+2k-2)(c+n+2k-1)}$$

d'où α_k et β_k grâce aux relations (II-6). La fonction hypergéométrique ${}_2F_1(a, 1; c; z)$ correspond à ce cas puisqu'on a, par définition:

$${}_2F_1(a, 1; c; z) = \frac{\Gamma(c)}{\Gamma(a)} \sum_0^\infty \frac{\Gamma(n+a)}{\Gamma(n+c)} z^n$$

On a, dans ce cas : $\gamma_0 \Gamma(c) / \Gamma(a) = 1$. On trouve :

$${}_2F_1(a, 1; c; z) = \frac{1}{1 - \alpha_0 z - \frac{\beta_1 z^2}{1 - \alpha_1 z - \frac{\beta_2 z^2}{1 - \alpha_2 z - \ddots}}}$$

avec :

$$\alpha_k = \frac{k(k+c-a-1)}{(2k+c-2)(2k+c-1)} + \frac{(k+a)(k+c-1)}{(2k+c)(2k+c-1)}$$

$$\beta_k = \frac{k(k+c-a-1)(k+a-1)(k+c-2)}{(2k+c-1)(2k+c-2)^2(2k+c-3)}$$

Traitons, par exemple, le cas de la fonction :

$$z^{-1} \arctg(z) = F(1/2, 1; 3/2; -z^2)$$

Il correspond au choix suivant des paramètres : $a = 1/2$, $c = 3/2$ et $z \rightarrow -z^2$.

Dans ce cas particulier, il se fait qu'il est plus simple de considérer la forme non contractée (II-5). On a, en effet, que :

$$q_k^{(0)} = \frac{(2k-1)^2}{(4k-1)(4k-3)} \quad \text{et} \quad e_k^{(0)} = \frac{(2k)^2}{(4k-1)(4k+1)}$$

d'où on tire le développement suivant :

$$z^{-1} \operatorname{arctg}(z) = \frac{1}{1 + \frac{z^2/(1 \cdot 3)}{1 + \frac{4z^2/(3 \cdot 5)}{1 + \frac{9z^2/(5 \cdot 7)}{1 + \ddots}}}} \quad (\text{II-8})$$

8)

Si $z = 1$, l'approximant représenté vaut 0.7843137..., tandis que la fonction vaut exactement $\pi/4 = 0.785398...$

La récurrence associée s'écrit : $-A_k X_{k+1} + X_k + X_{k-1} = 0$ avec :

$$A_k = \frac{(k+1)^2 z^2}{(2k+1)(2k+3)} \approx \frac{z^2}{4} + O(1/k)$$

Se rapportant à la table (II-1), on constate qu'on est dans le cas $v = 0$, $1 + 4\lambda_0 = 1 + z^2 > 0$, d'où une convergence rapide de l'ordre de :

$$e^{-p} \approx \rho_k \approx \left(\frac{1 - \sqrt{1 + z^2}}{1 + \sqrt{1 + z^2}} \right)^k \quad (\text{II-9})$$

Si $z = 1$, on a, en particulier : $e^{-p} \approx \left(\frac{1 - \sqrt{2}}{1 + \sqrt{2}} \right)^k \approx (1/5.828)^k$

soit près de 8 chiffres décimaux dès que $k = 10$. En comparaison, la série de puissances définissant l'arctg, à savoir, $z - z^3/3 + z^5/5 - \dots$ nécessiterait 500.000 termes pour n'obtenir que 6 chiffres exacts.

La formule (II-9) montre que la vitesse de convergence diminue à mesure que z augmente. Toutefois, si $z > 1$ et réel, on peut profiter de l'identité suivante :

$$\operatorname{arctg} z + \operatorname{arctg} z^{-1} = \pi/2 \quad \text{pour} \quad z \in (0,1)$$

L'extension de l'approximation (II-8), à tout le plan complexe, est autorisée. On montre, en particulier, que les points $z=\pm i$ apparaissent comme singuliers, ce qu'ils sont effectivement, puisque la fonction $\text{arctg } n'y$ est pas définie.

Exemple 4 : $\gamma_n = \rho^{an^2+bn}$

On trouve : $q_k^{(n)} = \rho^{2an+4ak-3a+b}$ et $e_k^{(n)} = (\rho^{2ak} - 1)\rho^{2an+2ak-a+b}$

d'où α_k et β_k grâce à (II-6).

II-4. Fractions continues généralisées.

L'étude des fractions continues simples est, nous l'avons vu, étroitement associée à celle des récurrences linéaires, homogènes, d'ordre deux. Il est possible de généraliser la notion de fc, associées à des récurrences linéaires, homogènes, d'ordre, n , supérieur à deux. On parle alors de fraction continue généralisée (fcg). Il y a plusieurs manières d'envisager cette généralisation. Celle que nous présentons ici est due, principalement, à Rutishauser et Magnus.

II-4.1. Forme déterminante des approximants d'une fc simple.

On montre facilement que la suite des déterminants que voici :

$$\Delta_k = \begin{vmatrix} a_0 & -b_0 & & & & \\ 1 & a_1 & -b_1 & & & \\ & 1 & a_2 & -b_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & a_{k-2} & -b_{k-2} \\ & & & & 1 & a_{k-1} \end{vmatrix}_{k \times k}$$

obéit à la récurrence: $\Delta_k = a_{k-1}\Delta_{k-1} - b_{k-2}\Delta_{k-2}$

Si on pose : $\Delta_k = N_k b_0 b_1 \dots b_{k-1}$, on trouve que la suite N_k satisfait la récurrence (II-1) :

$$b_k N_{k+1} = a_k N_k + N_{k-1}$$

Ses valeurs initiales, $N_1=a_0/b_0$ et $N_2=(a_0 a_1 + b_0)/(b_0 b_1)$, autorisent son identification à la suite des numérateurs, N_k , déjà rencontrée dans le calcul des approximants fc_k .

Par ailleurs les continuants Λ_k s'identifient de la même façon, à un facteur près, aux dénominateurs D_k . En effet, on a que :

$$\Lambda_k = \begin{vmatrix} a_1 & -b_1 & & & \\ 1 & a_2 & -b_2 & & \\ & 1 & a_3 & -b_3 & \\ & & \ddots & \ddots & \ddots \\ & & & 1 & a_{k-2} & -b_{k-2} \\ & & & & 1 & a_{k-1} \end{vmatrix}_{(k-1) \times (k-1)}$$

obéit à la même récurrence que Δ_k et de plus on a que :

$$\Lambda_k = D_k b_0 b_1 \dots b_{k-1} .$$

Il en résulte que la fc simple, associée à la récurrence (II-1), peut s'écrire sous la forme du quotient de deux déterminants tridiagonaux :

$$fc_k = \frac{\begin{vmatrix} a_0 & -b_0 & & & \\ 1 & a_1 & -b_1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & a_{k-2} & -b_{k-2} \\ & & & 1 & a_{k-1} \end{vmatrix}_k}{\begin{vmatrix} a_1 & -b_1 & & & \\ 1 & a_2 & -b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & a_{k-2} & -b_{k-2} \\ & & & 1 & a_{k-1} \end{vmatrix}_{k-1}} \quad (\text{II-10})$$

C'est cette forme déterminante qui sert le plus commodément de base à la généralisation des fc.

II-4.2. Forme déterminante des approximants d'une fcg.

A toute récurrence linéaire homogène d'ordre n, écrite sous la forme canonique :

$$A_k^{(n)} C_{k+1} + A_k^{(n-1)} C_k + \dots + A_k^{(0)} C_{k-n+1} = 0 \quad (\text{II-11})$$

on associe n-1 suites d'approximants généralisés, d'ordre s, notées $fcg_k^{(s)}$ ($s=0, 1, \dots, n-2$) :

$$fcg_k^{(s)} = \frac{\begin{vmatrix} A_s^{(n-s-1)} & \dots & A_s^{(n)} \\ \vdots & & \ddots \\ A_{n-1}^{(0)} & & A_{k-2}^{(n)} \\ & \ddots & \vdots \\ & A_{k+s-1}^{(0)} & \dots & A_{k+s-1}^{(n-s-1)} \end{vmatrix}_k}{\begin{vmatrix} A_{s+1}^{(n-s-1)} & \dots & A_{s+1}^{(n)} \\ \vdots & & \ddots \\ A_n^{(0)} & & A_{k-2}^{(n)} \\ & \ddots & \vdots \\ & A_{k+s-1}^{(0)} & \dots & A_{k+s-1}^{(n-s-1)} \end{vmatrix}_{k-1}} \quad (II-12)$$

Sous réserve que ces suites convergent, lorsque k tend vers l'infini, on définit de la sorte n-1 composantes de fcg associées à la récurrence (II-11). On vérifie sans peine que lorsque $n=2$ (d'où $s=0$), les relations (II-12) se réduisent à la seule équation (II-10) et on retrouve la définition classique des approximants d'une fc simple.

II-4.3. Calcul récursif des approximants d'une fcg.

Les numérateurs et dénominateurs des approximants (II-12) peuvent être calculés récursivement à condition de définir s+2 suites auxiliaires $N_k^{(j)}$ ($j=1, 2, \dots, s+1; n$), qui satisfont toutes la récurrence (II-11) avec les conditions initiales suivantes :

$$N_k^{(j)} = \delta_{k, l-j} \quad (k = s+1-n, \dots, s)$$

On peut en effet démontrer l'identité suivante :

$$\begin{vmatrix} A_s^{(n-s-1)} & \dots & A_s^{(n)} \\ \vdots & & \ddots \\ A_{n-1}^{(0)} & & A_{k-2}^{(n)} \\ & \ddots & \vdots \\ & A_{k+s-1}^{(0)} & \dots & A_{k+s-1}^{(n-s-1)} \end{vmatrix} = (-1)^{k(s+1)} A_s^{(n)} A_{s+1}^{(n)} \dots A_{k+s-1}^{(n)} \begin{vmatrix} N_k^{(1)} & \dots & N_k^{(s+1)} \\ \vdots & & \vdots \\ N_{k+s}^{(1)} & \dots & N_{k+s}^{(s+1)} \end{vmatrix}$$

qui exprime avantageusement le premier déterminant, de dimension variable avec k, sous la forme d'un deuxième déterminant de dimension **fixe** (s+1) quel que soit k.

Elle a pour conséquence que l'approximant (II-12) peut revêtir la forme condensée :

$$fcg_k^{(s)} = \frac{\begin{vmatrix} N_k^{(1)} & \dots & N_k^{(s+1)} \\ \vdots & & \vdots \\ N_{k+s}^{(1)} & \dots & N_{k+s}^{(s+1)} \end{vmatrix}}{\begin{vmatrix} N_k^{(1)} & \dots & N_k^{(s)} & N_k^{(n)} \\ \vdots & & \vdots & \vdots \\ N_{k+s}^{(1)} & \dots & N_{k+s}^{(s)} & N_{k+s}^{(n)} \end{vmatrix}} \quad (s=0, 1, \dots, n-2) \quad (\text{II-13})$$

II-4.4. Théorème de Pincherle généralisé.

Il règle la convergence des (n-1) composantes (II-13) de la fcg associée à la récurrence (II-11).
Il s'énonce comme suit :

- (i) La composante d'indice s de la fcg associée à la récurrence (II-11) converge ssi il existe s+1 solutions indépendantes de la récurrence, notées $f_k^{(i)}$ ($i=1, \dots, s+1$), qui dominent les n-s-1 autres, notées $f_k^{(i)}$ ($i=s+2, \dots, n$).
- (ii) Lorsqu'une composante converge, sa valeur est exactement égale à :

$$fcg^{(s)} = (-1)^{n-s-1} \frac{\begin{vmatrix} f_{-1}^{(s+2)} & \dots & f_{s+1-n}^{(s+2)} \\ \vdots & & \vdots \\ f_{-1}^{(n)} & \dots & f_{s+1-n}^{(n)} \end{vmatrix}}{\begin{vmatrix} f_0^{(s+2)} & \dots & f_{s+2-n}^{(s+2)} \\ \vdots & & \vdots \\ f_0^{(n)} & \dots & f_{s+2-n}^{(n)} \end{vmatrix}}$$

Cette formule est, comme il se doit, invariante vis-à-vis de toute combinaison linéaire des suite f_k choisies.

- (iii) La vitesse de convergence de la suite des approximants de la composante d'indice s est de l'ordre du (s+1)^{ième} facteur de contraste de la récurrence, soit :

$$e^{-p} \approx \left| \frac{C_k^{(s+2)}}{C_k^{(s+1)}} \right|$$

La démonstration de ce théorème ne diffère pas essentiellement de celle du théorème de Pincherle présenté section II-1.2, qui ne concernait que les fc simples (n=2, s=0).

On note que certaines composantes peuvent ne pas converger et que la condition pour qu'elles convergent toutes est que les asymptotes de la récurrence soient toutes contrastées à l'infini.

II-4.5. Application des fcg à la méthode de Bernoulli.

Les fcg se rencontrent toutes les fois que le problème traité fait intervenir une récurrence linéaire homogène d'ordre n supérieur à deux. Une illustration est fournie par l'amélioration suivante de la méthode de Bernoulli de détermination des racines d'un polynôme. Considérons l'exemple d'une récurrence linéaire, d'ordre $n=4$, dont les coefficients, constants, sont précisément les coefficients du polynôme dont on cherche les racines S, T, U, V , supposées de modules distincts. Son équation caractéristique possède donc précisément ces quatre racines S, T, U et V que l'on peut ranger par ordre de modules croissants :

$$|S| < |T| < |U| < |V|.$$

Il est connu que la racine, V , de plus grand module se détermine sans peine par l'algorithme de Bernoulli, qui consiste à calculer numériquement la limite du quotient de deux solutions dominantes quelconques de la récurrence associée. Toutefois, la méthode de Bernoulli ne convient pas pour les autres racines. C'est ici qu'interviennent les fcg associées à la même récurrence. On se convainc en effet facilement, par application du théorème de Pincherle généralisé, que les trois composantes $fcg^{(s)}$ valent respectivement :

$$s=0: fcg^{(0)} = -1/(S T U)$$

$$s=1: fcg^{(1)} = 1/(S T)$$

$$s=2: fcg^{(2)} = -1/S$$

Concrètement, chacune de ces fcg se calcule en estimant numériquement la limite des suites d'approximants (II-13) ce qui permet la détermination des racines par voie entièrement récursive.

Si deux ou plusieurs racines sont comodulaires, la méthode reste valable à condition de considérer une généralisation supplémentaire des fcg basées sur une extension de la formule (II-13).

Concrètement, à chaque ordre s , on associe non pas une mais $n-s-1$ composantes de fcg, notées cette fois :

$$fcg_k^{(i,s)} = \frac{\begin{vmatrix} N_k^{(1)} & \dots & N_k^{(s)} & N_k^{(i)} \\ \vdots & & \vdots & \vdots \\ N_{k+s}^{(1)} & \dots & N_{k+s}^{(s)} & N_{k+s}^{(i)} \end{vmatrix}}{\begin{vmatrix} N_k^{(1)} & \dots & N_k^{(s)} & N_k^{(n)} \\ \vdots & & \vdots & \vdots \\ N_{k+s}^{(1)} & \dots & N_{k+s}^{(s)} & N_{k+s}^{(n)} \end{vmatrix}} \quad (i = s+1, \dots, n-1)$$

Cette généralisation ne pose pas de problème particulier si ce n'est qu'il faut étendre le schéma d'initialisation des suites $N_k^{(i)}$:

$$N_k^{(j)} = \delta_{k,l-j} \quad \text{aux cas: } j=1, 2, \dots, n.$$

On montre, sans trop de difficultés, que dans le cas de l'exemple traité, les diverses composantes de la fcg valent respectivement :

$$s=0: \text{fcg}^{(1,0)} = -1/(S T U); \text{fcg}^{(2,0)} = 1/(T U) + 1/(S U) + 1/(S T); \text{fcg}^{(3,0)} = -(1/U + 1/T + 1/S)$$

$$s=1: \text{fcg}^{(2,1)} = 1/(S T); \text{fcg}^{(3,1)} = -(1/T + 1/S)$$

$$s=2: \text{fcg}^{(3,2)} = -1/S$$

Ces formules offrent l'avantage de rester valables lorsque deux ou plusieurs racines sont comodulaires: une racine de plus grand module disparaît chaque fois qu'on augmente s d'une unité et les autres restent, dont certaines sont éventuellement de même module que celle qui a disparu.

Au chapitre V, réservé aux applications plus élaborées, nous illustrerons la théorie des fcg par l'exemple spectaculaire du calcul des valeurs propres de l'équation de Schrödinger pour un potentiel anharmonique.

CHAPITRE III

LECONS SUR LES APPROXIMANTS DE PADE.

III-1. Un peu d'histoire.

Bien que la thèse de Padé remonte à 1892, il aura fallu attendre les années 1950 pour que les approximants rationnels qu'elle étudiait soient remis à la mode grâce à une série d'applications physiques. Le problème posé par Padé est le suivant : soit une fonction $f(z)$ donnée par son développement en série de puissances valable au voisinage de l'origine :

$$f(z) = c_0 + c_1z + c_2z^2 + \dots \quad (\text{III-1})$$

Il s'agit de trouver un ensemble de fractions rationnelles dépendantes de z qui approximent valablement $f(z)$ au voisinage de l'origine. Les approximants de Padé offrent **une** solution à ce problème en des termes que nous allons expliciter. Cette solution ne prétend pas être optimale et elle ne l'est effectivement pas : elle présente simplement l'avantage de se laisser calculer assez facilement tout en accélérant la convergence de la série initiale, loin de l'origine, dans un certain nombre de cas que nous apprendrons à reconnaître. Nous évoquerons au passage d'autres approximations rationnelles plus précises mais, en contrepartie, généralement plus difficiles à mettre en oeuvre. Signalons que la génération des systèmes IBM 360 utilisaient les approximants de Padé pour l'évaluation de certaines fonctions transcendentes, la fonction exponentielle notamment.

Un cours complet sur les approximants de Padé nous entraînerait fort loin. Nous nous limiterons ici à l'essentiel en nous efforçant de replacer cette théorie dans le contexte général des méthodes modernes d'accélération de la convergence. Cela nous permettra de remettre les approximants de Padé à une place relativement plus modeste que celle qu'ils occupent, un peu indûment aujourd'hui, au détriment d'autres méthodes plus puissantes qui seront étudiées au chapitre IV. L'ouvrage de référence le plus complet est dû à Baker et Graves-Morris (2ème Tome de l'Encyclopedia of Mathematics parue chez Addison-Wesley).

III-2. Calcul analytique des approximants de Padé.

On note traditionnellement $[L/M]$ l'approximant suivant :

$$[L/M] = \frac{a_0 + a_1 z + \dots + a_L z^L}{b_0 + b_1 z + \dots + b_M z^M} \quad (\text{III-2})$$

Les coefficients a_i et b_j sont tels que $a_L \neq 0$ et $b_M \neq 0$ et tels que le développement en série de MacLaurin de $[L/M]$ coïncide avec celui de $f(z)$ aussi loin que possible, c'est-à-dire généralement jusqu'au terme en z^{L+M} inclus. « Généralement » signifie que, dans certains cas, une dégénérescence peut se produire, due à des valeurs particulières des coefficients c_i de f , qui interdit de trouver la coïncidence en question jusqu'à l'ordre $L+M$ inclus. L'étude de ces cas pathologiques ne nous intéresse pas ici car ils sont étrangers au problème général de l'accélération de la convergence. Nous admettrons donc l'existence des approximants $[L/M]$ sous la forme (III-2) tels que l'on a :

$$[L/M] - f(z) = O(z^{L+M+1}) \quad (\text{III-3})$$

Pour obtenir les coefficients a_i et b_i de l'approximant $[L/M]$, on procède comme suit : on réécrit la relation (III-3) sous la forme :

$$(a_0 + a_1 z + \dots + a_L z^L) - (c_0 + c_1 z + \dots + c_{L+M} z^{L+M})(b_0 + b_1 z + \dots + b_M z^M) = O(z^{L+M+1})$$

Cette relation devant être valable pour tout z , on obtient le système linéaire suivant, en annulant les coefficients de z^{L+1}, \dots, z^{L+M} :

$$\begin{pmatrix} c_{L-M+1} & c_{L-M+2} & \dots & c_L \\ c_{L-M+2} & c_{L-M+3} & \dots & c_{L+1} \\ \vdots & \vdots & & \vdots \\ c_L & c_{L+1} & \dots & c_{L+M-1} \end{pmatrix} \begin{pmatrix} b_M \\ b_{M-1} \\ \vdots \\ b_1 \end{pmatrix} = - \begin{pmatrix} c_{L+1} \\ c_{L+2} \\ \vdots \\ c_{L+M} \end{pmatrix} \quad (\text{III-4})$$

où b_0 peut être posé égal à 1 sans que cela altère la généralité du développement puisque $f(z)$ doit être définie à l'origine ($b_0 \neq 0$). Les a_i se déduisent alors des relations obtenues en annulant pareillement les coefficients de z^0, z^1, \dots, z^L , soit:

$$\begin{aligned} a_0 &= c_0 \\ a_1 &= c_1 + b_1 c_0 \\ a_2 &= c_2 + b_1 c_1 + b_2 c_0 \\ &\vdots \\ a_L &= c_L + \sum_{i=1}^{\min(L, M)} b_i c_{L-i} \end{aligned} \quad (\text{III-5})$$

On peut proposer pour l'écriture de l'approximant $[L/M]$ une forme plus synthétique que voici:

$$[L, M] = \frac{P_{(z)}^{L, M}}{Q_{(z)}^{L, M}}$$

où les polynômes P et Q se déduisent de la règle de Cramer relative au système (III-4):

$$Q^{L, M}(z) = \begin{vmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_{L+1} \\ c_{L-M+2} & c_{L-M+3} & \cdots & c_{L+2} \\ \vdots & \vdots & & \vdots \\ c_{L-1} & c_L & \cdots & c_{L+M+1} \\ c_L & c_{L+1} & \cdots & c_{L+M} \\ z^M & z^{M-1} & \cdots & z^0 \end{vmatrix} \quad (\text{III-6})$$

$$P^{L, M}(z) = \begin{vmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_{L+1} \\ c_{L-M+2} & c_{L-M+3} & \cdots & c_{L+2} \\ \vdots & \vdots & & \vdots \\ c_{L-1} & c_L & \cdots & c_{L+M+1} \\ c_L & c_{L+1} & \cdots & c_{L+M} \\ \sum_{i=0}^{L-M} c_i z^{M+i} & \sum_{i=0}^{L-M+1} c_i z^{M+i-1} & \cdots & \sum_{i=0}^L c_i z^i \end{vmatrix} \quad (\text{III-7})$$

La table (III-1) présente, dans l'ordre, les premiers approximants de Padé.

M\L	0	1	2	3	4
0	[0/0]	[1/0]	[2/0]	[3/0]	[4/0]
1	[0/1]	[1/1]	[2/1]	[3/1]	...
2	[0/2]	[1/2]	[2/2]	...	
3	[0/3]	[1/3]	...		
4	[0/4]	...			

Table III-1

Supposons que la fonction f(z) soit connue par son développement de MacLaurin (III-1) jusqu'au terme $c_4 z^4$ inclus. Il est possible de dresser la table triangulaire ci-dessus où, en première ligne, figurent des sommes partielles successives de (III-1).

Exemple : La fonction exponentielle. Résolvant les systèmes (III-6) et (III-7), on obtient la table III-2.

$M \backslash L$	0	1	2	3	4
0	1	$1+z$	$1+z+\frac{z^2}{2}$	$1+z+\frac{z^2}{2}+\frac{z^3}{6}$	$1+z+\frac{z^2}{2}+\frac{z^3}{6}+\frac{z^4}{24}$
1	$\frac{1}{1-z}$	$\frac{2+z}{2-z}$	$\frac{6+4z+z^2}{6-2z}$	$\frac{24+18z+6z^2+z^3}{24-6z}$	
2	$\frac{2-2z+z^2}{6}$	$\frac{6+2z}{6-4z+z^2}$	$\frac{12+6z+z^2}{12-6z+z^2}$		
3	$\frac{6-6z+3z^2-z^3}{24}$	$\frac{24+6z}{24-18z+6z^2-z^3}$			
4	$\frac{4-24z+12z^2-4z^3+z^4}{24}$				

Table III-2

Cette table permet d'estimer l'information maximum que les approximants de Padé sont capables d'extraire à chaque stade $[L/M]$.

Si, pour fixer les idées, on choisit $z = 1$, on obtient le tableau suivant des précisions p :

M\L	0	1	2	3	4
0	0.46 →	1.33 ↓	2.52	3.96	5.61
1	- ∞	2.26 →	4.45 ↓	6.53	
2	1.33	3.96	6.52 →	↓	
3	2.26	5.71		→	↓
4	3.96				

Table III-3

On peut noter sur ce tableau des précisions, une tendance qui est générale et qui serait encore plus apparente si on prolongeait la table dans l'angle inférieur droit : ce sont les approximants diagonaux $[L/L]$ et paradiagonaux $[L\pm 1/L]$ qui accélèrent le mieux la convergence de la série de départ. Ce fait n'est pas particulier à la fonction exponentielle et il se généralise largement en pratique.

III-3. Formes analytiques exactes des approximants de Padé.

La fonction exponentielle est la seule à se laisser traiter analytiquement assez facilement : dans ce cas, le calcul des déterminants (III-6) et (III-7) est facilité par la forme simple des coefficients $c_i = 1/i!$. On trouve:

$$e^z \approx [L, M] = \frac{{}_1F_1(-L; -L-M; z)}{{}_1F_1(-M; -L-M; -z)}$$

Pour rappel : ${}_1F_1(a; b; z) = 1 + \frac{a}{b} \frac{z}{1!} + \frac{a(a+1)}{b(b+1)} \frac{z^2}{2!} + \dots$

est la fonction hypergéométrique dégénérée.

On peut également montrer, mais le calcul des déterminants (III-6) et (III-7) est loin d'être aisé dans ce cas, qu'une forme exacte existe, également, pour les approximants de Padé des fonctions f dont les coefficients c_n du développement de MacLaurin autorise la résolution exacte de l'algorithme qd, c'est-à-dire rappelons-le:

$$\begin{aligned} c_n &= \Gamma(a+n) \quad \text{ou} \\ c_n &= 1/\Gamma(a+n) \quad \text{ou} \\ c_n &= \Gamma(a+n)/\Gamma(b+n) \\ c_n &= \rho^{an^2+bn} \end{aligned}$$

ou encore, plus généralement:

$$c_n = \prod_{j=0}^n \left[(A - q^{\alpha+j}) / (B - q^{\gamma+j}) \right]$$

En particulier, les trois premières formes correspondent aux fonctions hypergéométriques suivantes:

$$\begin{aligned} f(z) &= {}_2F_0(a, l; -; z) \quad \text{ou} \\ f(z) &= {}_1F_1(l; a; z) \quad \text{ou} \\ f(z) &= {}_2F_1(a, l; b; z) \end{aligned}$$

On trouvera dans l'ouvrage de Luke (The special functions and their approximations, vol II) le détail des numérateurs et des dénominateurs des approximants de Padé de ces fonctions particulières.

Un autre exemple est le suivant, tiré de l'encyclopédie de Baker: on considère la fonction

$$f(z) = \sqrt{\frac{1+z/2}{1+2z}} \approx 1 - \frac{3}{4}z + \frac{39}{32}z^2 + \dots$$

On calcule sans peine l'approximant suivant:

$$[1/1] = \frac{a+bz}{1+cz} \approx a + z(b-ac) + z^2(ac^2-bc) + O(z^3)$$

d'où on tire: $a = 1$, $b = 7/8$ et $c = 13/8$, soit:

$$[1/1] = \frac{1+7/8z}{1+13/8z}$$

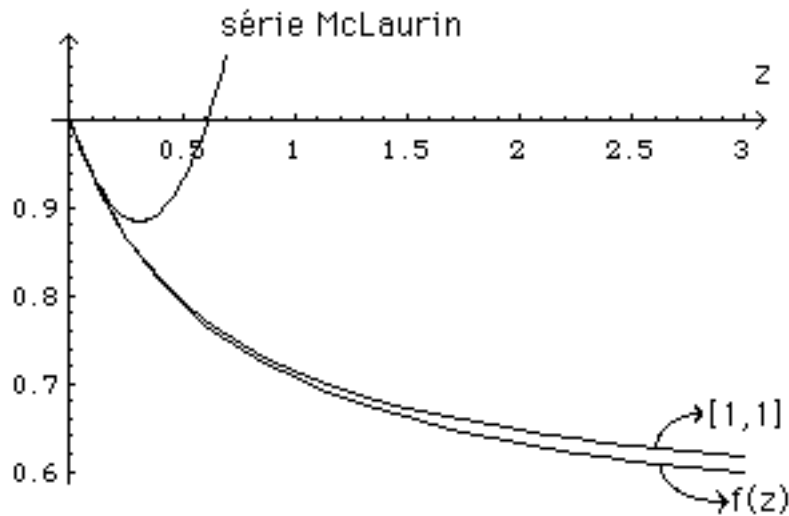


Fig. III-1

La figure III-1 compare le graphe exact de $f(z)$ à celui de la série de MacLaurin limitée au terme en z^2 et à celui de l'approximant $[1/1]$. Il y a évidemment peu de différences près de l'origine. Par contre, loin de l'origine, l'approximant de Padé affirme sa supériorité: malgré les faibles valeurs retenues pour L et M , le comportement asymptotique est déjà presque bon:

$\lim_{z \rightarrow \infty} f(z) = 1/2$, tandis que:

$\lim_{z \rightarrow \infty} [L,M] = 7/13$.

III-4. Approximants de Padé et fractions continues.

Il existe une connexion évidente entre f_c et approximants de Padé proches de la diagonale principale dans la table de Padé. Nous avons vu, au chapitre précédent, une première méthode d'accélération de la convergence d'une série de puissances basées sur l'algorithme qd :

$$f(z) = \sum_0^{\infty} \gamma_k z^k = \frac{\gamma_0}{1 - \frac{q_1^{(0)} z}{1 - \frac{e_1^{(0)} z}{1 - \frac{q_2^{(0)} z}{1 - \ddots}}}}$$

où les quantités $e_k^{(0)}$ et $q_k^{(0)}$ se déduisent de l'algorithme qd. Cet algorithme possède la propriété que son développement en série de puissances de z coïncide avec celui de $f(z)$ aussi loin que possible. Il est facile de vérifier que les approximants de la f_c sont des fractions rationnelles en la variable z . Il devient dès lors clair que l'algorithme qd construit la suite des approximants de Padé suivants :

[0/0], [0/1], [1/1], [1/2], ...

Ces éléments étant paradiagonaux, c'est-à-dire proches de l'optimum, on conclut que l'algorithme qd extrait à lui seul l'essentiel de l'information contenue dans la table de Padé. Celle-ci, sous sa forme complète, n'apporte donc en général qu'un faible complément d'information. Par contre, il peut se faire que les algorithmes de construction de la table de Padé soient plus stables que l'algorithme qd, auquel cas il aurait lieu de les préférer.

III-5. Calcul récursif des approximants de Padé.

Lorsqu'on doit calculer une table de Padé un peu étendue, il est tout à fait inconfortable de recourir aux formes déterminantes (III-6) et (III-7). L'extrême longueur des calculs se double d'une instabilité numérique notoire. Une méthode plus sûre consiste à revenir au système (III-4) et à le résoudre par la méthode d'élimination de Gauss avec pivotement complet, ceci afin de minimiser les erreurs d'arrondi. A quelques raffinements près, c'est une des méthodes réellement utilisées en pratique. Signalons une règle empirique selon laquelle on perd en moyenne M chiffres décimaux lorsqu'on calcule $[L/M]$ en sorte qu'il peut y avoir lieu de recourir à une arithmétique de précision étendue.

Une autre méthode de calcul des approximants de Padé est purement récursive. Elle est basée sur la règle de la croix à laquelle obéissent les approximants voisins : (III-8)

$$\frac{1}{[L+1/M] - [L/M]} + \frac{1}{[L-1/M] - [L/M]} = \frac{1}{[L/M+1] - [L/M]} + \frac{1}{[L/M-1] - [L/M]}$$

assortie des conditions initiales :
$$\begin{cases} [L/-1] = \infty & [-1/M] = 0 \\ [L/0] = c_0 + c_1x + \dots + c_Lx^L \end{cases}$$

Vu la disposition relative des approximants impliqués dans cette relation, à savoir :

$$\begin{array}{ccccc} [L/M-1] & & & & N \\ [L-1/M] & [L/M] & [L+1/M] & \equiv & 0 \quad C \quad E \\ [L/M+1] & & & & S \end{array}$$

La formule (III-8) peut se réécrire de façon plus suggestive en utilisant une notation géographique pour les approximants voisins dans la table. On écrit dès lors :

$$\frac{1}{E-C} + \frac{1}{0-C} = \frac{1}{S-C} + \frac{1}{N-C}$$

Cette relation permet de calculer de proche en proche les lignes successives de la table de Padé comme indiqué dans la table III-4 qui suit :

M\L	-1	0	1	2
-1		∞	∞	∞
0	0	c_0	$\begin{array}{c} \uparrow \\ \leftarrow c_0 + c_1z \rightarrow \\ \downarrow \end{array}$	$c_0 + c_1z + c_2z^2$
1	0		*	
2	0			

Table III-4

Cette règle récursive n'est valable qu'à la condition que tous les approximants qui y figurent existent. Le problème se complique dans certains cas marginaux de dégénérescence, que nous n'abordons pas ici.

Il est intéressant d'utiliser la récurrence (III-8) pour calculer la première ligne de la table de Padé. On pose donc $M = 0$ et on trouve :

$$\frac{1}{[L+1/0]-[L/0]} + \frac{1}{[L-1/0]-[L/0]} = \frac{1}{[L/1]-[L/0]} + 0$$

d'où on peut extraire l'information $[L/1]$ demandée :

$$[L/1] = \frac{[L+1/0] [L-1/0] - [L/0]^2}{[L+1/0] + [L-1/0] - 2[L/0]}$$

qui montre que la première ligne de la table de Padé n'est autre que le résultat de l'application du procédé Δ^2 d'Aitken à la suite $[L/0]$. Rappelons qu'étant donné une suite S_n , le procédé Δ^2 fabrique une nouvelle suite Σ_n telle que :

$$\Sigma_n = \frac{S_{n+1}S_{n-1} - S_n^2}{S_{n+1} + S_{n-1} - 2S_n}$$

qui, sous certaines conditions que nous étudierons au chapitre suivant, converge plus vite que la suite S_n .

En ce qui concerne la convergence des éléments de la table de Padé, un théorème dû à de Montessus affirme que : $\lim_{L \rightarrow \infty} [L/M] = f(z)$, sauf aux pôles de f , pourvu qu'ils soient exactement en nombre M (multiplicité comptée) dans le disque $|z| \leq R$ (convergence par ligne). En dehors de ce théorème assez restrictif, rien n'est connu qui assure la convergence des approximants $[L/M]$ vers f , que ce soit en ligne, en colonne, en diagonale ou encore via n'importe quel chemin dans la table de Padé. Cela est dû au fait qu'il est assez facile de construire des fonctions $f(z)$ telles que les approximants ne convergent vers rien du tout. Le principe de ces contre-exemples est heureusement assez tordu : il consiste à choisir des c_n de formes différentes selon la valeur de n (parité, modulo 3, ...). Baker et Gammel ont tenté une conjecture jamais infirmée à ce jour, à savoir qu'il existe en tous cas une sous-suite d'approximants qui converge vers f . Ajoutons quand même que, pour les fonctions $f(z)$ « honnêtes », la propriété de convergence s'étend à la suite tout entière. Quant à savoir s'il y a réellement accélération, c'est une toute autre histoire qui ne sera complètement élucidée qu'au chapitre suivant. Quelques essais numériques permettent de se faire une première idée à ce sujet.

III-6. Performances et limitations des approximants de Padé.

Si l'on se place au seul point de vue de l'accélération de la convergence, il y a lieu de se demander si les approximants de Padé sont efficaces. Commençons par rappeler ce fait d'expérience selon lequel les meilleurs approximants de Padé sont proches de la diagonale principale de la table de Padé, soit, dans le détail, $[0/0]$, $[0/1]$, $[1/1]$, $[1/2]$, Vu que ceux-ci coïncident avec les approximants de la fraction continue associée, le théorème de Pincherle

permet d'estimer la vitesse de convergence du procédé. Par exemple, nous avons étudié, au § II-3.1, le cas de la fonction exponentielle pour laquelle nous avons trouvé un développement en fraction continue dont la vitesse de convergence, très élevée, est donnée par la relation :

$$e^{-p} \approx (z^2 / 16)^k k / k!^2$$

Elle est très supérieure à la vitesse de convergence de la série de puissances définissant la fonction exponentielle qui est de l'ordre de $|z|^{k+1}/(k+1)!$. Si $z = 1$ et $k = 5$, la fc fait gagner 7 chiffres décimaux !

L'exemple de la fonction exponentielle compte parmi ceux qui produisent une accélération notable de la convergence. Il faut reconnaître que le résultat n'est pas toujours aussi spectaculaire. Considérons cet autre exemple fourni par la fonction dilogarithme :

$$\frac{\text{dil}(z)}{z} = 1 + \frac{z}{4} + \frac{z^2}{9} + \dots + \frac{z^k}{(k+1)^2} + \dots$$

Si on calcule son approximant de Padé [3/3], on trouve:

$$[3/3] = \frac{1}{144} \frac{7908921z^3 - 265113200z^2 + 970942500z - 825375600}{316784z^3 - 3248100z^2 + 8175600z - 5731775}$$

Cet approximant vaut, en $z = 1$:

$$[3/3]_{z=1} = \frac{111637379}{70198704} = 1.590305...$$

qui n'approche la valeur exacte $\pi^2/6 = 1.64493406...$ qu'à 3%, soit guère mieux que la série de départ. Autant dire que, dans cet exemple, la table de Padé n'est d'aucune utilité. Nous verrons au chapitre IV comment reconnaître les cas favorables.

Revenons à ces cas favorables pour constater qu'ils ne sont pas nécessairement optimaux pour la cause. Afin d'expliciter ce point, il convient de préciser le sens qu'on peut prêter à l'approximant optimal. Un exemple clarifiera au mieux la situation. Supposons que l'on désire construire un algorithme de calcul de la fonction $[\text{th}(\mu x)]/x$ avec $x \in (-1,1)$ et $\mu = 1/2\ln(3)$. Une telle approximation est effectivement utilisée par certains calculateurs. La série de MacLaurin correspondante s'écrit :

$$\frac{\text{th}\mu x}{x} = \mu - \frac{\mu^3}{3} z^2 + \frac{2\mu^5}{15} z^4 - \frac{17\mu^7}{315} z^6 + \dots$$

A partir de ce développement, on construit l'approximant [2/4] suivant :

$$\text{Pade}[2/4] = \frac{0.5493061443 + 0.0157853448z^2}{1 + 0.1293159601z^2 + 0.0008670987z^4}$$

Il est intéressant de tester le degré d'approximation obtenu dans l'intervalle $x \in (-1,1)$. On trouve, pour l'erreur relative, le graphe de la figure III-2:

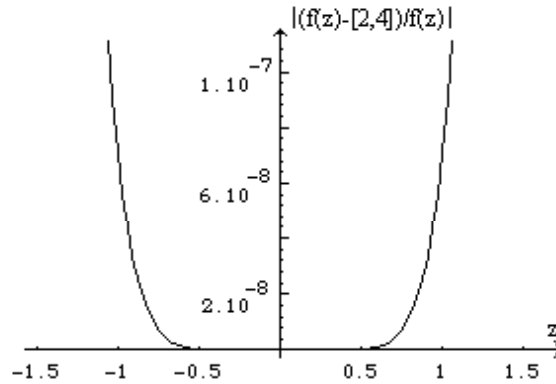


Fig III-2

En d'autres termes, si un tel algorithme est implanté sur ordinateur, on sera sûr de 7 chiffres significatifs dans l'intervalle $(-1,1)$. En dehors de cet intervalle, il y a lieu de prévoir une procédure de réduction qui ramène l'argument dans cet intervalle. Nous n'entrons pas dans ces détails qui nous éloigneraient de notre propos. Il est par contre intéressant de voir s'il est possible de trouver un approximant rationnel du type $(2/4)$ qui posséderait un meilleur graphe d'erreur relative. L'existence de ces fractions rationnelles optimales est effectivement garantie par le théorème du minimax de Tchebychev :

« Quels que soient les entiers m et n , il existe une fraction rationnelle du type (m/n) telle que $\max \left| \frac{(m,n) - f(x)}{f(x)} \right|$ est minimum sur l'intervalle (a,b) donné ».

Revenant à notre exemple, on peut montrer que l'approximant minimax est donné par :

$$\minimax(2/4) = \frac{0.54930614401 + 0.01574011995x^2}{1 + 0.12923360954x^2 + 0.00085891904x^4}$$

dont le graphe d'erreur relative, similaire à celui de la figure III-2, n'excède jamais $0.59 \cdot 10^{-9}$, soit un gain de deux chiffres significatifs. L'approximant minimax sera donc préféré pour l'implantation sur ordinateur. Son inconvénient est que, si la précision est jugée insuffisante, il faut recommencer tous les calculs afin d'augmenter les degrés du numérateur et du dénominateur car le calcul des approximants minimax, contrairement à celui des approximants de Padé, n'est pas récursif.

Les approximants de Padé ordinaires peuvent être généralisés de plusieurs manières.

III-7. Approximants de Padé généralisés.

Nous avons vu qu'un approximant de Padé ordinaire est le résultat de l'extrapolation de la fonction $f = \sum c_i x^i$ au moyen d'une fraction rationnelle qui coïncide, à l'origine, avec $f(x)$ ainsi que ses dérivées jusqu'à un certain ordre n fixé. On pourrait imaginer d'autres types d'extrapolation : nous allons les passer rapidement en revue.

III-7.1. Approximants de Padé multipoints.

Encore appelé problème d'interpolation de Cauchy-Jacobi, il s'agit, cette fois, de construire une fraction rationnelle qui approxime une fonction f donnée en lui étant égale, au point d'abscisse z_0, \dots, z_n . Les abscisses z_i peuvent être toutes distinctes ou bien, certaines d'entre elles peuvent être confondues, auquel cas on parle de confluence. Dans ce cas, il faut comprendre que les fonctions et leurs dérivées doivent être égales. Le cas limite est celui où tous les points z_i sont confondus : on retrouve les approximants de Padé ordinaires avec l'identification de la fonction et de toutes ses dérivées, en $z=z_i$, jusqu'à l'ordre n inclus.

La construction de la table de Padé multipoints résulte d'un jeu d'écritures assez fastidieuses utilisant des différences réciproques peu agréables à manipuler. Heureusement, toute la table de Padé ne doit pas nécessairement être construite et, comme pour les approximants de Padé ordinaires, les éléments diagonaux et paradiagonaux de la table sont bien suffisants. On trouve alors qu'une représentation sous forme de fc existe à nouveau que nous présentons dans deux cas classiques :

1) Les n points sont distincts : algorithme ρ et fc de Thiele.

Le problème se pose de la façon suivante : soit une fonction f qui prend les valeurs S_0, \dots, S_n au point z_0, \dots, z_n ($f(z_i) = S_i$, pour $i = 0, 1, \dots, n$). On cherche la fc qui prend les mêmes valeurs aux mêmes points. On formule implicitement l'espoir que cette fc interpolera, voire extrapolera, fidèlement la fonction, en tous points.

Cette fc existe effectivement et porte le nom de fc de Thiele. Initialement, elle fut introduite en terme de différences réciproques qui sont des quantités malaisées à manipuler. Nous préférons une approche récursive qui porte le nom d'algorithme ρ . On a :

$$f(z) \approx \rho_0^{(0)} + \frac{z - z_0}{\rho_1^{(0)} + \frac{z - z_1}{\rho_2^{(0)} - \rho_0^{(0)} + \frac{z - z_2}{\rho_3^{(0)} - \rho_1^{(0)} + \dots}}}$$

Les quantités $\rho_k^{(n)}$ se calculent récursivement, de la manière suivante :

$$\begin{cases} \rho_{-1}^{(n)} = 0 & \rho_0^{(n)} = f(z_n) \\ \rho_{k+1}^{(n)} = \rho_{k-1}^{(n+1)} + \frac{z_{k+n+1} - z_n}{\rho_k^{(n+1)} - \rho_k^{(n)}} \end{cases}$$

Il est facile de constater, sur quelques approximants simples, que $f(z_i)$ est bien la valeur prise par la fc aux points $z = z_i$. Nous passons la démonstration sous silence.

2) Algorithme qd généralisé.

Nous avons vu, au chapitre précédent, comment l'algorithme qd permet de construire la fc associée à un développement en série de MacLaurin lorsque le but poursuivi est la coïncidence des développements à l'origine jusqu'au degré n inclus. Une généralisation de cet algorithme qd existe, due à Wuytack et Graves-Morris, qui permet de construire la fc d'interpolation dans le cas général où les x_i ne sont pas tous confondus. Elle nécessite qu'on détermine préalablement les coefficients c_i du polynôme d'interpolation de f , écrit sous la forme :

$$P_n(z) = c_0 + c_1(z - z_0) + c_2(z - z_0)(z - z_1) + \dots + c_n \prod_0^{n-1} (z - z_i)$$

Alors, on a que :

$$f(z) = \frac{c_0}{1 - \frac{q_1^{(0)}(z - z_0)}{1 - \frac{e_1^{(0)}(z - z_1)}{1 - \frac{q_2^{(0)}(z - z_2)}{1 - \frac{e_2^{(0)}(z - z_3)}{1 - \dots}}}}}$$

où les quantités e et q se calculent récursivement par l'algorithme qd généralisé, dont les règles s'écrivent comme suit. On part des conditions initiales ($j = 0, 1, \dots, n-1$):

$$\begin{cases} Z_1^{(j)} = z_{j+1} - z_j \\ e_0^{(j+1)} = 0 \\ q_1^{(j)} = \frac{1}{Z_1^{(j)} + c_j / c_{j+1}} \end{cases} \quad e_1^{(j)} = -q_1^{(j)} - q_1^{(j+1)}(q_1^{(j)} Z_1^{(j)} - 1)$$

On a ensuite les règles de l'algorithme qd généralisé proprement dit ($i=2, 3, \dots$):

$$\begin{aligned} Z_i^{(j)} &= z_{j+2i-1} - z_{j+2i-2} \\ q_i^{(j)} &= \frac{1}{Z_i^{(j)} - \frac{e_{i-1}^{(j)} q_{i-1}^{(j+1)} + e_{i-2}^{(j+1)} Z_i^{(j)} e_{i-1}^{(j+1)} - 1}{e_{i-1}^{(j)} + q_{i-1}^{(j)} \frac{q_{i-1}^{(j+1)}}{e_{i-1}^{(j+1)}}}} \\ e_i^{(j)} &= -q_i^{(j)} + \frac{(Z_i^{(j)} q_i^{(j)} - 1)(e_{i-1}^{(j+1)} + q_i^{(j+1)})}{(Z_i^{(j)} e_{i-1}^{(j+1)} - 1)} \end{aligned}$$

Cet algorithme est particulièrement intéressant car il contient tous les cas particuliers de confluence $z_j \rightarrow z_i$. En particulier, si tous les z_i sont confondus, on peut montrer qu'on retrouve l'algorithme qd classique.

III-7.2. Approximants de Padé-Hermite.

Etant donné la fonction $f(z) = c_0 + c_1z + c_2z^2 + \dots$, le principe de la table de Padé est de déterminer deux polynômes P_m et Q_n tels que $f \approx P_m/Q_n$ près de l'origine c'est-à-dire tels qu'on peut écrire:

$$P_m - fQ_n = O(z^{m+n+1}).$$

Une façon de généraliser cette approche est de rechercher trois polynômes Q_q , R_r et S_s tels que

$$Q_q f^2 + 2R_r f + S_s = O(z^{q+r+s+2})$$

$$\text{d'où } f \approx \frac{-R \pm \sqrt{R^2 - QS}}{Q} = \frac{S}{-R \mp \sqrt{R^2 - QS}}$$

où le choix de la branche s'effectue en fonction du contexte. Cet approximant est, en principe, meilleur que l'approximant de Padé ordinaire.

Cette approche est particulièrement intéressante lorsqu'on veut approximer des fonctions qui présentent des singularités telles que des points de branchements, que les approximants de Padé ordinaires sont évidemment impuissants à simuler puisqu'ils sont rationnels.

Exemple :

Soit à approximer la fonction: $\text{arctg}(z) = z - z^3/3 + z^5/5 - z^7/7 + \dots$

Calculons l'approximant de Padé-Hermite [2,2,2]. On a successivement:

$$Q_2 f^2 + 2R_2 f + S_2 = O(z^8) \quad \text{avec} \quad f \approx z - \frac{z^3}{3} + \frac{z^5}{5} - \frac{z^7}{7} + O(z^8)$$

$$\text{et} \quad f^2 \approx z^2 - \frac{2}{3}z^4 + \frac{23}{45}z^6 + O(z^8)$$

$$\text{Si on pose: } \begin{cases} Q_2 = 1 + q_1 z + q_2 z^2 \\ R_2 = r_0 + r_1 z + r_2 z^2 \\ S_2 = s_0 + s_1 z + s_2 z^2 \end{cases}$$

on trouve de suite, en posant $z=0$, que $s_0 = 0$. Il reste dès lors:

$$(1 + q_1 z + q_2 z^2) \left(z - \frac{2}{3} z^3 + \frac{23}{45} z^5 \right) + (2r_0 + 2r_1 z + 2r_2 z^2) \left(1 - \frac{z^2}{3} + \frac{z^4}{5} - \frac{z^6}{7} \right) + s_1 + s_2 z = O(z^7)$$

On trouve, après identification :

$$s_1 + 2r_0 = 0$$

$$1 + 2r_1 + s_2 = 0$$

$$q_1 - \frac{2}{3}r_0 + 2r_2 = 0$$

$$q_2 - \frac{2}{3} - \frac{2}{3}r_1 = 0$$

$$-\frac{2}{3}q_1 + \frac{2}{5}r_0 - \frac{2}{3}r_2 = 0$$

$$\frac{23}{45} - \frac{2}{3}q_2 + \frac{2}{5}r_1 = 0$$

$$\frac{23}{45}q_1 - \frac{2}{7}r_0 + \frac{2}{5}r_2 = 0$$

$$\text{d'où : } Q_2 = 1 + \frac{5}{3}z^2 \quad R_2 = \frac{3}{2}z \quad S_2 = -4z^2$$

$$\text{Finalement, on obtient: } [2,2,2] = \frac{8z}{3 + \sqrt{25 + 80z^2/3}}$$

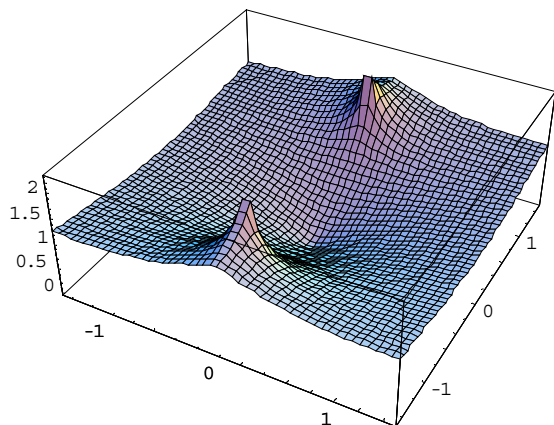
Le signe + , devant le radical, a été retenu afin d'être en accord avec le comportement asymptotique de f. On a, en effet, :

$$\lim_{z \rightarrow \infty} \arctg(z) = \frac{\pi}{2} = 1.57079...$$

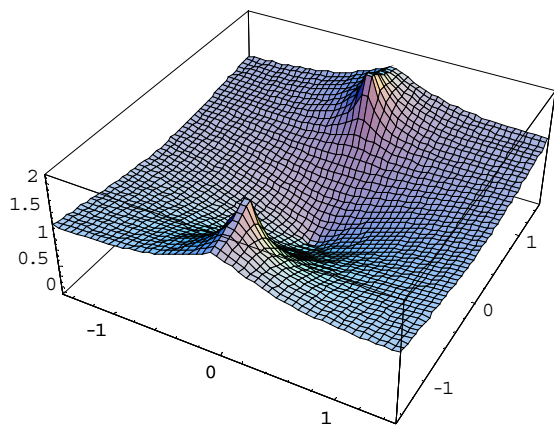
$$\lim_{z \rightarrow \infty} [2,2,2] = \frac{8}{\sqrt{80/3}} = 1.54919...$$

De plus, on constate que l'approximant présente une singularité essentielle en $z = \pm i\sqrt{15/16}$ déjà fort proche de la singularité de $\arctg(z)$, située en $z = \pm i$. Le programme Mathematica qui suit compare les graphes 3D du module de la fonction \arctg et de son approximant de Padé-Hermite, dans le plan complexe.

**Plot3D[Abs[ArcTan[x+y I]],{x,-1.5,1.5},{y,-1.5,1.5},
PlotPoints->50]**



**Plot3D[Abs[(8x+8y I)/(3+Sqrt[25+80(x+y I)^2/3])],
{x,-1.5,1.5},{y,-1.5,1.5},
PlotPoints->50]**



D'autres types de généralisations des approximants de Padé existent, basées sur des équations fonctionnelles d'ordre supérieur à deux mais les calculs se compliquent assez rapidement. Quelques cas réellement étudiés dans la littérature exigent de trouver des polynômes $P_i(z)$ de degré n_i tels que :

- (i) $P_1(z)f(z) + P_2(z)f'(z) = O(z^{n_1+n_2+1})$ ou
- (ii) $P_1(z) + P_2(z)f(z) + P_3(z)f'(z) = O(z^{n_1+n_2+n_3+2})$ ou
- (iii) $P_1(z)f + P_2(z)f'(z) + P_3(z)f''(z) = O(z^{n_1+n_2+n_3+2})$ ou
- (iv) $P_1(z) + P_2(z)f(z) + P_3(z)f^2(z) = O(z^{n_1+n_2+n_3+2})$

III-7.3. Approximation des fonctions transcendentes par des approximants de Padé-Gauss.

Il est possible de raffiner la méthode de Padé et de la rendre très précise dans le domaine intermédiaire $1 < z < 10$ où ni le développement de Taylor d'une fonction, $f(z)$, au voisinage de l'origine ni son développement asymptotique ne donnent des résultats précis. L'idée est la suivante. Commençons par rappeler que l'algorithme qd permet le calcul de l'approximant de Padé ordinaire $[n-1, n]$ de la fonction suivante, sous la forme d'une fc :

$$f(z) = \sum_0^{\infty} c_k z^k = \frac{c_0}{1 - \alpha_0 z - \frac{\beta_1 z^2}{1 - \alpha_1 z - \frac{\beta_2 z^2}{1 - \alpha_2 z - \ddots}}} \quad (\text{III-9})$$

Nous avons vu que les symboles α_k et β_k sont définis par les relations :

$$\alpha_k = e_k^{(0)} + q_{k+1}^{(0)} \quad \text{et} \quad \beta_k = e_k^{(0)} q_k^{(0)}$$

Il revient au même de considérer que cette méthode réalise l'approximation de $f(z)$ par un développement en fractions simples du type :

$$f(z) = \sum_0^{\infty} c_k z^k = \sum_1^s \frac{\lambda_i}{1 - \omega_i z}$$

où les ω et les λ peuvent être calculés récursivement à partir de la récurrence associée à la fc. Le point essentiel de la nouvelle méthode, est que ces valeurs de ω et λ ne changent pas si on cherche un développement plus général du type :

$$f(z) = \sum_0^{\infty} c_k z^k = \sum_1^s \lambda_i F(\omega_i z)$$

où $F(z) = \sum_0^{\infty} \sigma_k z^k$, à condition de poser :

$$c_k = \sigma_k \sum_1^s \lambda_i \omega_i^k = \sigma_k \gamma_k$$

Concrètement, la méthode s'organise comme suit. On désire une approximation de

$$f(z) = \sum_0^{\infty} c_k z^k \quad \text{sous la forme,}$$

$$f(z) = \sum_0^{\infty} c_k z^k = \sum_1^s \lambda_i F(\omega_i z), \quad \text{où la fonction } F \text{ est définie par le développement :}$$

$$F(z) = \sum_{k=0}^{\infty} \frac{c_k}{\gamma_k} z^k$$

On applique l'algorithme q-d à la suite γ_k et on en déduit les coefficients α_k et β_k de la récurrence satisfaite par les polynômes P_s , de degré $s-1$, et Q_s , de degré s , respectivement numérateur et dénominateur de l'approximant de Padé $[s-1, s]$ qui coïncide avec l'approximant de même rang de la fc (III-9). Les ω_i sont précisément les racines du dénominateur, Q_s , et les λ_i sont appelés symboles de Christoffel, dont on peut montrer qu'ils peuvent s'écrire sous l'une ou l'autre forme équivalente :

$$\lambda_i = \frac{P_s(\omega_i)}{Q'_s(\omega_i)} = - \frac{\prod_{k=1}^s \beta_k}{Q_{s+1}(\omega_i) Q'_s(\omega_i)}$$

Cette procédure peut être menée de façon entièrement récursive avec les risques inhérents d'instabilité. On les évite en se restreignant aux cas où l'algorithme q-d est soluble exactement. Trois cas particuliers sont faciles à mettre en œuvre dont deux se révèlent particulièrement utiles.

$$\underline{1^{\text{er}} \text{ cas}} : \gamma_k = \frac{1}{\Gamma(k+a)}$$

On trouve sans peine :

$$Q_s(z) = 1 F 1(-s; -2s - a + 2; -1/z) z^s$$

$$Q_s(\omega_i) = 0 \quad \text{et} \quad \lambda_i = \frac{(-1)^{s+1} \Gamma(a) \Gamma(s+1) \Gamma(s+a-1)}{\Gamma(2s+a-1) \Gamma(2s+a)} \frac{1}{Q_{s+1}(\omega_i) Q'_s(\omega_i)}$$

Ce cas n'est cité que pour mémoire car il paraît peu intéressant, en particulier parce que les racines ω_i sont complexes.

$$\underline{2^{\text{ème}} \text{ cas}} : \gamma_k = \Gamma(k+a)$$

On trouve sans peine :

$$Q_s(z) = 2 F 0(-s, -s - a + 1; -; -1/z) z^s$$

$$Q_s(\omega_i) = 0 \quad \text{et} \quad \lambda_i = - \frac{s! \Gamma(s+a)}{\Gamma(a)} \frac{1}{Q_{s+1}(\omega_i) Q'_s(\omega_i)}$$

Le programme Mathematica suivant montre la précision de la méthode sur le calcul de la fonction hypergéométrique ${}_1F_1(2/3; 1/2; z)$, pour $z=2$ par exemple, sous la forme d'un développement en termes de ${}_0F_1(-; 1/2; z) = \cos(\sqrt{-4z})$. L'approximant, $s=10$, fournit 18 chiffres corrects.

```
Q[a_,s_,u_]:=Simplify[HypergeometricPFQ[{-s,-s-a+1},{},-1/u] u^s]
roots[a_,s_]:=u/.NSolve[Q[a,s,u]==0,u,40]
lambdas[a_,s_]:=-(Factorial[s] Gamma[a+s])/Gamma[a]/(Q[a,s+1,u]
D[Q[a,s,u],u])/u->roots[a,s]
(*Exemple numérique*)
N[HypergeometricPFQ[{2/3},{1/2},2],30]
```

10.5165167727834659194822352679

roots[2/3,10]

{0.0843006124882692328901371045080685435745,0.604642222623998892101182544443885
35670,1.6174338033433653927052210206104175337,3.14854184617440224813984232949350
5016,5.24080339745256866067714889474406704,7.96198070473096477151372514131091964
,11.42055901762907631152911205305373936,15.8012888723279027285754741161493202,21.
46241977187136072230263057993368938,29.32469641802475770623219288241905460}

lambdas[2/3,10]

{0.43294723139238119758704692837149304586,0.36419694437698996229902981114047346
3,0.1580768107360282754369146667792428,0.038982043769953093467283729896297,0.0053
87545582386889430129766287405,0.0003952708450491097565715197719700,0.00001395300
845890863012191527059960,1.99461701041990800445989988771×10⁻⁷,8.26623734179859197603396720744×10⁻¹⁰,4.2778722224201888913313301159174×10⁻¹³}

F[z_]:=Cos[Sqrt[-4 z]]

lambdas[2/3,10].F[2 roots[2/3,10]]

10.516516772783465902443573250844

$$\text{3}^{\text{ème}} \text{ cas : } \gamma_k = \frac{\Gamma(k+a)}{\Gamma(k+c)}$$

On trouve sans peine :

$$Q_s(z) = {}_2F_1(-s, -s-a+1; -2s-c+2; 1/z) z^s$$

$$Q_s(\omega_i) = 0 \quad \text{et} \quad \lambda_i = -\frac{s! \Gamma(c) \Gamma(a+s) \Gamma(c+s-1) \Gamma(c+s-a)}{\Gamma(c+2s) \Gamma(c+2s-1) \Gamma(a) \Gamma(c-a)} \frac{1}{Q_{s+1}(\omega_i) Q'_s(\omega_i)}$$

Le programme Mathematica suivant calcule la fonction hypergéométrique ${}_2F_2(1/2, 1; 2/3, 3; z)$, pour $z=2$ par exemple, sous la forme d'un développement en termes des fonctions de base ${}_1F_1(1; 3; z) = 2(\exp(z)-1-z)/z^2$. L'approximant, $s=10$, fournit 24 chiffres corrects.

```

Q[a_,c_,s_,u_]:=Simplify[HypergeometricPFQ[{-s,-s-a+1},{-2 s-c+2},1/u] u^s]
roots[a_,c_,s_]:=u/.NSolve[Q[a,c,s,u]==0,u,40]
lambdas[a_,c_,s_]:=-(Factorial[s] Gamma[c] Gamma[a+s] Gamma[c+s-1] Gamma[c+s-
a])/((Gamma[c+2 s] Gamma[c+2 s-1] Gamma[a] Gamma[c-a]))/(Q[a,c,s+1,u]
D[Q[a,c,s,u],u])/u->roots[a,c,s]
(*Exemple numérique*)
N[HypergeometricPFQ[{1/2,1},{2/3,3},2],30]

```

1.84589396420159737775290310457

roots[1/2,2/3,10]

{0.006373136507621069986316896678464846,0.056388671513678957819234987935658316,0.15135791232402277374184891311939969,0.28166947448348357988053922564421219,0.43413511511168413146427240411489956,0.59332438754627275723847259678133842,0.7431260888240776655130547400545663,0.8683780052416165051022763043813587,0.9563961159895162250327249846413712,0.9981368067437406199355446609344451}

lambdas[1/2,2/3,10]

{0.0439682627354369798472881980739401,0.044733130005263847186041470027398,0.04634628541494372417126643499118,0.0490005974377662854094285228322,0.05306784366836420319037726613,0.0592684652621342454625615704,0.069132870550792226375249626,0.08657173587515943447633541,0.1261340807252776754860716,0.421776728324861378395380}

```

F[z_]:=2 (Exp[z]-1-z)/z^2
lambdas[1/2,2/3,10].F[2 roots[1/2,2/3,10]]

```

1.84589396420159737775290...

CHAPITRE IV

LES ALGORITHMES

D'ACCELERATION DE LA

CONVERGENCE DES SUITES.

IV-1. Convergence des suites.

Jusqu'à présent, nous n'avons considéré que l'accélération de la convergence de suites un peu particulières, du type :

$$f(z) = \sum_{n=0}^{\infty} c_n z^n \quad (\text{IV-1})$$

en considérant qu'une telle série peut être étudiée comme suite de ses sommes partielles. Nous envisageons, à présent, la convergence des suites quelconques.

Considérons la suite $\{S_n\} = S_0, S_1, S_2, \dots$ qui converge vers la limite S . Les questions essentielles qui se posent sont les suivantes :

- Comment estimer quantitativement la vitesse de convergence de cette suite ?
- S'il s'avère que cette vitesse est faible, est-il possible de l'augmenter en remplaçant la suite $\{S_n\}$ par une autre, $\{\sigma_n\}$ par exemple, qui tendrait, plus rapidement, vers la même limite?

Une réponse satisfaisante à la deuxième question est loin d'être évidente. Nous allons présenter un grand nombre d'algorithmes qui poursuivent ce but. Leur nombre indique qu'aucun algorithme universel n'existe qui serait valable pour tous les cas : cela signifie qu'il y a lieu d'étudier chaque cas séparément afin de dégager les règles qui permettent d'opérer le choix de l'algorithme optimal. Cette situation était déjà illustrée par les approximations de Padé qui accélèrent la série exponentielle mais échouaient complètement en face de la série dilogarithme.

Dans le but d'illustrer au maximum les algorithmes très variés qui vont être présentés, nous avons sélectionné quelques séries "test" destinées à subir ces divers algorithmes. Nous les définissons une fois pour toutes et nous y référerons systématiquement en cours d'exposé. Nous avons choisi ces suites afin d'illustrer le plus grand nombre de cas susceptibles de se présenter en pratique. Elles sont extraites du livre de Wimp (Sequence transformations and their applications.- Academic Press).

Suite S_n Limite S (asymptote principale)

$$(LN2)_n = \sum_0^n \frac{(-1)^k}{k+1}$$

$$\ln 2 = 0.693147180559945... (1/n)$$

$$(RC)_n = \sum_0^n \frac{(-1)^k}{\sqrt{k+1}}$$

$$0.604898643421630... (1/\sqrt{n})$$

$$(PIC)_n = \sum_0^n \frac{1}{(k+1)^2}$$

$$\pi^2/6 = 1.644934066846559... (1/n)$$

$$(FAC)_n = \sum_0^n (-1)^k k!$$

$$\text{divergente, prolongée par } \int_0^\infty \frac{e^{-t}}{t+1} dt = 0.5963473611...$$

$$(FOU)_n = \sum_0^n \frac{\sin(k)}{k}$$

$$\frac{\pi-1}{2} = 1.070796327...$$

$$(IT1)_n \text{ générée par } S_{n+1} = \frac{20}{S_n^2 + 2S_n + 10} \quad (S_0 = 1)$$

$$1.368808107...$$

$$(IT2)_n \text{ générée par } S_{n+1} = \frac{20 - 2S_n^2 - S_n^3}{10} \quad (S_0 = 1)$$

$$\text{divergente}$$

$$(LUB)_n = \sum_0^n \frac{(-1)^{\left\lceil \frac{k}{2} \right\rceil}}{k+1}$$

$$1.131971754$$

Ces suites ont été choisies suffisamment lentes, au départ, pour exiger un procédé d'accélération. Pour quantifier la vitesse de convergence d'une suite, il suffit d'étudier son comportement asymptotique. On peut écrire, dans tous les cas :

$$S_n = S + as(n)$$

en sorte que l'erreur relative que l'on commet sur S en se limitant à S_n vaut :

$$e^{-p} \approx \left| \frac{S - S_n}{S} \right| = \left| \frac{as(n)}{S} \right|$$

La connaissance du comportement asymptotique de la suite permet donc d'en estimer la convergence. En général, $as(n)$ n'est connue qu'à une constante multiplicative près, en sorte que l'on se borne à écrire, $p = -\ln |as(n)|$, où on a négligé le logarithme de la constante S qui est généralement petit. p indique le nombre de chiffres népériens corrects dans S_n , soit 2.3

fois le nombre de chiffres décimaux. Pour fixer les idées, nous appellerons p la précision. En fin de chapitre IV, au terme de l'étude des divers procédés d'accélération, nous leur soumettrons l'ensemble des suites "test" et nous dresserons le tableau comparatif, IV-4, des précisions avant et après transformation.

Il n'existe malheureusement pas de règle absolument générale pour obtenir l'asymptote $as(n)$ par voie analytique. Lorsque la suite est composée des approximants d'une fraction continue, nous avons vu au chapitre II comment le théorème de Pincherle permet d'obtenir une expression pour $as(n)$: il suffit d'estimer le contraste entre les solutions dominantes et dominées de la récurrence associée. Pour ce qui est des séries, le calcul de $as(n)$ est déjà moins simple. On trouvera dans l'ouvrage de De Bruijn (Asymptotic methods in Analysis - North Holland) les bases de la théorie sous-jacente. Une procédure simple consiste à essayer la formule sommatoire de MacLaurin :

$$\sum_{k=1}^n f_k \approx \int_0^n f(k) dk - \frac{1}{2}[f(0) + f(n)] + \frac{1}{12}[f'(n) - f'(0)] - \frac{1}{720}[f'''(n) - f'''(0)] + \dots$$

Exemple : Soit la suite : $S_n = (PIC)_n = \sum_0^n \frac{1}{(k+1)^2} = \sum_1^{n+1} \frac{1}{k^2}$

Il est clair que $f(k) = 1/k^2$, que ses dérivées tendent asymptotiquement vers 0 et que sa primitive fournit le terme dominant. On est donc certain du comportement asymptotique suivant :

$$(PIC)_n \approx S + \lambda/n \text{ où } \lambda \text{ est une constante inconnue et } S \text{ vaut } \pi^2/6.$$

On voit que cette suite converge lentement puisque 10^6 termes garantissent à peine 6 chiffres exacts. Il y a lieu de remarquer sur cet exemple que la formule de Mac Laurin permet de raffiner le développement asymptotique. On peut, en effet, écrire :

$$(PIC)_n \approx S + \lambda_1/n + \lambda_2/n^2 + \lambda_3/n^3 + \dots \quad \text{où les constantes } \lambda_i \text{ sont inconnues.}$$

On voit ainsi apparaître le concept d'échelle asymptotique :

$$S_n \approx S + \lambda_1 as^{(1)}(n) + \lambda_2 as^{(2)}(n) + \lambda_3 as^{(3)}(n) + \dots \quad \text{où } as^{(1)}(n) \text{ est l'asymptote dominante et } as^{(i)}(n), (i>1), \text{ sont les asymptotes dominées.}$$

$$\text{On a, de plus, que : } \lim_{n \rightarrow \infty} [as^{(j)}(n)/as^{(i)}(n)] = 0 \text{ si } j > i.$$

On verra plus loin (algorithme E) que la connaissance de l'échelle asymptotique complète permet, en quelque sorte, d'optimiser l'accélération de la convergence de la suite. Malheureusement, elle n'est pas toujours totalement accessible et il faut souvent se contenter de $as^{(1)}(n)$. On notera que l'échelle asymptotique ne fixe pas, de façon univoque, la valeur numérique des coefficients λ_i du fait que toute combinaison linéaire d'asymptotes distinctes, d'ordre i et plus, peut être considérée elle-même comme une asymptote d'ordre i . Cela n'hypothèque heureusement en rien les possibilités d'accélération de la convergence de la suite S_n .

IV-2. Les procédés de sommation linéaire.

Initialement conçus pour attribuer une limite aux séries divergentes (cfr Hardy, Divergent series, Oxford Press) par une sorte de prolongement analytique, les procédés de sommation peuvent être appliqués aux séries lentement convergentes dans le but de les accélérer. La classe la plus importante des procédés de sommation linéaire est basée sur l'application d'une matrice triangulaire $A = (a_{ij})$ qui transforme la suite originale, S_n , en une suite modifiée, T_n , définie comme suit :

$$\begin{pmatrix} T_0 \\ T_1 \\ T_2 \\ \vdots \end{pmatrix} = A \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ \vdots \end{pmatrix}$$

avec
$$A = \begin{pmatrix} a_{00} & & & \\ a_{10} & a_{11} & & 0 \\ a_{20} & a_{21} & a_{22} & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Appelons A ce procédé de sommation, du nom de la matrice qui le symbolise. On dit qu'il est régulier si, la suite S_n convergeant vers S, la suite transformée, T_n , converge vers la même limite. Un théorème dû à Toeplitz permet d'affirmer la régularité du procédé A sous trois hypothèses :

$$\begin{aligned} (i) \quad & \sum_{k=0}^{\infty} |a_{nk}| < M \quad \forall n \\ (ii) \quad & \lim_{n \rightarrow \infty} a_{nk} = 0 \quad \forall k \\ (iii) \quad & \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} a_{nk} = 1 \end{aligned}$$

Ces conditions sont suffisantes sans être nécessaires car il existe des procédés qui y dérogent et qui sont cependant réguliers pour certaines familles de suites. En fait, le théorème de Toeplitz s'applique également lorsque la matrice A n'est pas triangulaire, mais en pratique, c'est toujours le cas triangulaire qui est utilisé.

Les principaux procédés réguliers étudiés sont :

1) Le procédé (H,1) de Hölder. La matrice A est donnée par :

$$a_{ij} = \begin{cases} 1/(i+1) & \text{pour } j=1, \dots, i. \\ =0 & \text{pour } j>i \end{cases}$$

Un procédé plus évolué, noté (H,K), utilise la matrice A^K (K entier).

Concrètement, on écrit :

$$(H,1) \equiv \begin{pmatrix} 1/1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

d'où la suite transformée, T_n , des moyennes arithmétiques :

$$\begin{aligned} T_0 &= S_0 \\ T_1 &= \frac{S_0 + S_1}{2} \\ T_2 &= \frac{S_0 + S_1 + S_2}{3} \\ &\vdots \end{aligned}$$

Ensuite, on définit la matrice $(H,2)$ par la relation :

$$(H,2) = (H,1) * (H,1)$$

d'où la nouvelle suite U_n :

$$\begin{aligned} U_0 &= T_0 \\ U_1 &= \frac{T_0 + T_1}{2} \\ U_2 &= \frac{T_0 + T_1 + T_2}{3} \end{aligned}$$

Exemples : soit les deux suites divergentes oscillantes :

a) $S_n = \{1, 0, 1, 0, 1, \dots\}$, d'expression générale : $S_n = \sum_0^n (-1)^k = \frac{1 + (-1)^n}{2}$

On trouve, facilement : $T_n = \{1, 1/2, 2/3, 2/4, 3/5, 3/6, \dots\} \rightarrow 1/2$

b) $S_n = \{1, -1, 2, -2, 3, \dots\}$, d'expression générale : $S_n = \sum_0^n (-1)^k (k+1) = \frac{1 + (-1)^n (2n+3)}{4}$

On trouve successivement : $T_n = \{1, 0, 2/3, 0, 3/5, 0, 4/7, 0, \dots\} \equiv \frac{1}{4} \frac{n+2}{n+1} [1 + (-1)^n]$

$$U_n = \{1, 1/2, 5/9, 5/12, 34/75, 34/90, \dots\} \rightarrow 1/4$$

On peut légitimement s'interroger sur le sens à donner à la limite d'une série divergente. Les limites des suites prises en exemple sont en fait données par le prolongement analytique des fonctions $\sum_0^{\infty} z^k$ et $\sum_0^{\infty} (k+1)z^k$ en $z=-1$. On pourrait penser que les fonctions $(1-z)^{-1}$ et $\sum_0^{\infty} z^k$ contiennent la même quantité d'information. Cependant $(1-z)^{-1}$ est définie partout, sauf en $z=1$, tandis que $\sum_0^{\infty} z^k$ n'est définie, dans le plan complexe, que dans le cercle $|z| < 1$ et l'idée du procédé de sommation est précisément de retrouver l'information perdue.

Le deuxième exemple concerne, similairement, la fonction $(1-z)^{-2}$ et la série équivalente $\sum_0^{\infty} (k+1)z^k$.

2) Le procédé de Cesaro d'ordre K (C,K).

Sa matrice d'ordre K, notée symboliquement (C,K), vaut: $(C,K)=H(1)*L^{K-1}$, où la matrice L est triangulaire inférieure gauche, avec tous éléments égaux à l'unité.

3) Le procédé d'Euler.

Les éléments de la matrice A, du procédé d'Euler, sont donnés par :

$$a_{ij} = \frac{q^{i-j}}{(q+1)^i} C_i^j \quad \text{pour } j=0,1,\dots,i$$

$$= 0 \quad \text{pour } j > i$$

$$\left(\sum_{j=0}^i a_{ij} = 1 \right)$$

où le paramètre q peut être choisi arbitrairement (souvent on prend q=1).

Il est facile de voir que ces trois procédés sont réguliers. Ils sont même totaux, c'est-à-dire que l'on a en plus que $\sum_{k=0}^n a_{nk} = 1 \quad \forall n$, ce qui garantit automatiquement la troisième hypothèse du théorème de Toeplitz.

On peut présenter de deux autres manières la théorie des procédés de sommation.

IV-2.1. Moyennes de Nörlund.

Une première manière consiste à définir la suite transformée sous la forme d'une moyenne pondérée des S_n :

$$t_n = \frac{p_n S_0 + p_{n-1} S_1 + \cdots + p_0 S_n}{p_0 + \cdots + p_n}$$

Le procédé est régulier si $\frac{p_i}{\sum p_k} \rightarrow 0 \quad \forall i$.

Le procédé (C,K) de Césaro rentre dans cette catégorie avec le choix pondéral suivant :

$$p_m = \frac{\Gamma(m+K)}{\Gamma(m+1)\Gamma(K)}$$

De même, le procédé d'Euler correspond au choix pondéral : $p_m = q^m C_n^m$.

Par contre les procédés d'Hölder, (H,K), y échappent sauf évidemment (H,1) \equiv (C,1). La théorie des moyennes d'Hausdorff est cependant plus générale en sorte que nous préférons nous attarder sur celle-ci.

IV-2.2. Moyennes d'Hausdorff.

La suite transformée s'écrit sous la forme :

$$t_k^{(0)} = \sum_{m=0}^k (-1)^{k-m} C_k^m (\Delta^{k-m} \mu_m) S_m$$

où μ_m est une suite auxiliaire (telle que $\mu_0 = 1$), choisie en sorte que le procédé qui en résulte soit régulier. On peut adopter un point de vue encore plus général en appliquant le procédé à la suite S considérée comme démarrant à un rang n quelconque, ce qui donne la forme canonique du procédé d'Hausdorff :

$$t_k^{(n)} = \sum_{m=0}^k (-1)^{k-m} C_k^m (\Delta^{k-m} \mu_m) S_{n+m} \quad (\text{IV-2})$$

Tous calculs faits, on dispose les résultats dans un tableau à double entrée de ce type :

$$\begin{array}{cccc}
t_0^{(0)} = S_0 & t_1^{(0)} & t_2^{(0)} & \dots \\
t_0^{(1)} = S_1 & t_1^{(1)} & t_2^{(1)} & \dots \\
t_0^{(2)} = S_2 & t_1^{(2)} & t_2^{(2)} & \dots \\
\vdots & \vdots & \vdots & \ddots
\end{array}$$

Dans cette table, la première colonne reproduit la suite de départ car $\mu_0=1$. La deuxième colonne livre le résultat de la première étape de l'algorithme et ainsi de suite en suivant les colonnes successives qui accélèrent éventuellement la convergence de la suite de départ. On trouvera dans l'ouvrage de Hardy, déjà cité, les conditions que doit remplir la suite μ_n afin que le procédé de sommation soit régulier.

Reprenons les trois exemples déjà étudiés :

- procédé (H, K) de Hölder : $\mu_m = \frac{I}{(m+1)^K}$
- procédé (C,K) de Cesaro : $\mu_m = \frac{I}{C_{m+K}^K}$
- procédé d'Euler : $\mu_m = \frac{I}{(q+1)^m}$

L'application de ces procédés exige le calcul des différences successives de la suite μ_m . Ce calcul est trop complexe dans le cas Hölder pour être envisagé par voie analytique. Par contre, il est faisable dans les deux autres cas. On trouve successivement :

- cas Césaro (C,K) : $\Delta^p \mu_n = \frac{K! n!}{(n+K+p)!} \frac{(K+p-1)!}{(K-1)!} (-1)^p$

$$\text{d'où : } t_k^{(n)} = \frac{\sum_{m=0}^k C_{K+k-m-1}^{K-1} S_{n+m}}{C_{K+k}^K} \quad (\text{IV-3})$$

- cas Euler : $\Delta^p \mu_n = (-1)^p \frac{q^p}{(q+1)^{p+n}}$

$$\text{d'où : } t_k^{(n)} = \sum_{m=0}^k C_k^m \frac{q^{k-m}}{(q+1)^k} S_{n+m} \quad (\text{IV-4})$$

Les formules (III-3) et (III-4), faciles à programmer, sont peu utilisées en pratique à cause de leurs performances modestes. Leur étude théorique n'est cependant pas dénuée d'intérêt dans la mesure où elle permet de comprendre pourquoi les méthodes linéaires accélèrent rarement les suites de façon spectaculaire. Le diagnostic d'impuissance étant posé, il devient alors possible de construire d'autres procédés, généralement non linéaires, qui évitent tous les inconvénients.

IV-3. Noyau d'un procédé de sommation.

Le formalisme d'Hausdorff est particulièrement adapté à la détermination du noyau des procédés de sommation linéaire. On appelle noyau d'un procédé de sommation l'ensemble des suites que ce procédé transforme en la suite constante $\{S\}$ à une étape, k , quelconque mais fixée de son développement. On a donc $t_k^{(n)} = S$ pour tout n et pour k fixé.

L'algorithme est donc exact pour les suites qui font partie de son noyau. On conçoit que, plus le noyau d'un procédé est riche, plus ce procédé risque d'être intéressant en ce qu'il accélérera une classe plus étendue de suites.

La détermination du noyau d'Hausdorff nécessite le lemme suivant :

Lemme : Quelle que soit la suite μ_n , on a l'identité :
$$\sum_{m=0}^k (-1)^{k-m} C_k^m (\Delta^{k-m} \mu_{n+m}) = \mu_n$$

Preuve : considérons l'opérateur $(E - \Delta)^k$ qui se réduit par définition à l'identité.

Il s'écrit encore :
$$(E - \Delta)^k = \sum_{m=0}^k (-1)^{k-m} C_k^m E^m \Delta^{k-m} \equiv I$$

Il suffit d'appliquer cette identité à la suite μ_n pour trouver le résultat annoncé.

Les suites S_n qui font partie du noyau du procédé d'Hausdorff sont telles qu'à une certaine étape k , on a :
$$t_k^{(n)} = S \quad \forall n$$

On a donc, se reportant à (IV-2) et utilisant le résultat du lemme :

$$\begin{aligned} \sum_{m=0}^k (-1)^{k-m} C_k^m (\Delta^{k-m} \mu_m) S_{n+m} &= S \\ &= S \sum_{m=0}^k (-1)^{k-m} C_k^m (\Delta^{k-m} \mu_m) \end{aligned}$$

$$\text{soit encore : } \sum_{m=0}^k (-1)^{k-m} C_k^m (\Delta^{k-m} \mu_m) (S_{n+m} - S) = 0$$

Ceci n'est rien d'autre qu'une récurrence à coefficients constants, satisfaite par la suite $S_n - S$. Sa solution générale s'écrit donc, à chaque étape k , dans le cas probable de racines distinctes de l'équation caractéristique :

$$S_n - S = \lambda_1 z_1^n + \dots + \lambda_k z_k^n$$

Le noyau des procédés d'Hausdorff est donc constitué de combinaisons **quelconques** d'exponentielles (éventuellement flanquées d'un polynôme si une ou plusieurs racines sont multiples) dont les arguments z_i sont **imposés** par le procédé. Les z_i changent d'ailleurs de valeur à chaque étape k .

On conçoit que l'imposition des valeurs des z_i est une restriction sévère qui rend le noyau tellement ténu que le procédé y perd une bonne partie de son utilité. On comparera ultérieurement avec l'algorithme ε qui lève cette restriction.

Exemple 1 : considérons le procédé (C, 1). On a :

$$t_k^{(n)} = \frac{\sum_{m=0}^k C_{k-m}^0 S_{n+m}}{C_{k+1}^1} = \frac{\sum_{m=0}^k S_{n+m}}{k+1}$$

Si $t_k^{(n)} = S$, pour tout n , k fixé, on voit que $S_{n+m} - S$ obéit à une récurrence à coefficients constants d'équation caractéristique :

$$z^k + z^{k-1} + \dots + z + 1 = 0 = \frac{z^{k+1} - 1}{z - 1}$$

d'où il résulte que les suites, qui appartiennent au noyau de (C, 1), sont de la forme :

$$S_n = S + \lambda_1 z_1^n + \dots + \lambda_k z_k^n \quad (\text{IV-5})$$

où les z_i sont les racines $(k+1)^{\text{èmes}}$ de l'unité (1 étant exclu) et où les λ_i sont quelconques (réels ou complexes).

Exemple 2 : considérons le procédé d'Euler.

Procédant de même, on trouve l'équation caractéristique :

$$\sum_{m=0}^k C_k^m q^{k-m} z^m = 0 \quad \text{qui se réduit à: } (1 + z/q)^k = 0, \text{ dont la racine est } -q \text{ de multiplicité } k.$$

Les suites faisant partie du procédé d'Euler, dans sa $k^{\text{ème}}$ étape, sont donc de la forme :

$$S_n = S + (-q)^n P_{k-1}(n) \quad (\text{IV-6})$$

où $P_{k-1}(n)$ désigne un polynôme quelconque de degré $k - 1$ de la variable n .

En résumé, le procédé d'Euler somme exactement les suites S_n exponentielles polynômes de type (IV-6) tandis que le procédé de Cesaro somme exactement les combinaisons linéaires d'exponentielles de type (IV-5) d'arguments imposés.

IV-4. Autres procédés de sommation linéaire.

Tous les procédés linéaires ne sont pas inintéressants même si leur noyau est relativement plus pauvre que celui de leurs homologues non linéaires. Assez curieusement, les plus intéressants sont en fait, non réguliers, ce qui ne les empêche pas de préserver la limite d'un certain nombre de suites. On ne les utilise donc que pour ces suites-là. Ils s'écrivent sous la forme d'une combinaison linéaire, à coefficients variables avec n , des termes de la suite S_{n+m} .

Un exemple utile est fourni par le procédé de Salzer :

$$t_k^{(n)} = \sum_{m=0}^k \frac{(-1)^{m+k}}{k!} (n+m+\gamma)^k C_k^m S_{n+m}$$

où γ est un paramètre dont la valeur n'a guère d'importance, on pose généralement $\gamma = 1$.

Il tire son intérêt du fait que ses coefficients a_{km} dépendent de n ; cela a pour conséquence que l'équation récurrente qui définit le noyau n'est plus à coefficients constants. Les suites qui font partie du noyau ne se réduisent plus dès lors à la seule classe des exponentielles polynômes comme dans les cas Cesaro et Euler. Il est d'ailleurs facile de voir que la $k^{\text{ème}}$ étape de cet algorithme possède comme noyau l'ensemble des suites :

$$S_n = S + \sum_{i=1}^k \lambda_i / (n+\gamma)^i$$

en sorte que l'application répétée de l'algorithme doit permettre d'accélérer les suites se comportant comme :

$$S_n \approx S + \frac{\lambda_1}{n+\gamma} + \frac{\lambda_2}{(n+\gamma)^2} + \dots$$

Tel est le cas, par exemple, de la suite-test notée $(PIC)_n$.

Nous n'insistons pas davantage ici car ce type de procédé apparaîtra plus loin comme un cas particulier de l'algorithme E.

IV-5.1. Algorithmes deltoïdes linéaires.

Les algorithmes de triangle du type : $t_k^{(n)} = \sum_{m=0}^k a_{km} S_{n+m}$ peuvent, dans certains cas, être

avantageusement reformulés en termes d'algorithmes deltoïdes qui présentent l'avantage d'être récursifs. L'un d'eux est bien connu : il s'agit du procédé d'extrapolation de Richardson :

$$\begin{cases} t_{k+1}^{(n)} = \frac{x_n t_k^{(n+1)} - x_{n+k+1} t_k^{(n)}}{x_n - x_{n+k+1}} \\ t_0^{(n)} = S_n \end{cases} \quad (n, k = 0, 1, \dots)$$

où la suite auxiliaire x_n tend vers 0.

Il est connu que $t_k^{(n)}$ est la valeur en $x = 0$ du polynôme d'interpolation de degré k qui passe par les $k+1$ couples : $(x_n, S_n), (x_{n+1}, S_{n+1}), \dots, (x_{n+k}, S_{n+k})$.

Il existe une relation entre la matrice (a_{km}) et la suite x_n :

$$a_{km} = \prod_{i=0, i \neq m}^k \frac{x_{n+i}}{x_{n+i} - x_{n+m}} \quad \text{avec} \quad \sum_{m=0}^k a_{km} = 1$$

Exemple : Le choix $x_i = 1/(i+1)$ mène aux éléments de matrice :

$$a_{km} = \frac{(-1)^{m+k}}{k!} (n+m+1)^k C_k^m$$

qui correspondent au procédé de Salzer avec le choix standard $\gamma = 1$. Celui-ci admet donc la formulation récursive suivante :

$$\begin{cases} t_{k+1}^{(n)} = \frac{(n+k+2)t_k^{(n+1)} - (n+1)t_k^{(n)}}{k+1} \\ t_0^{(n)} = S_n \end{cases} \quad (n, k \geq 0)$$

Rappelons que le procédé de Salzer n'est pas régulier, ce qui ne l'empêche pas d'être très efficace sur un ensemble restreint de suites.

La forme générale des algorithmes deltoïdes est la suivante :

$$\begin{cases} t_{k+1}^{(n)} = \alpha_k t_k^{(n+1)} + \beta_k t_k^{(n)} \\ t_0^{(n)} = S_n \end{cases} \quad n, k \geq 0 \quad (\text{IV-7})$$

La condition $\alpha_k + \beta_k = 1$ est suffisante pour assurer la régularité de l'algorithme. Le procédé de Salzer n'y obéit pas. Les procédés de sommation réguliers sont caractérisés par une matrice triangulaire d'éléments a_{ij} tels que :

$$a_{k+1,m} = \alpha_k a_{k,m-1} + (1 - \alpha_k) a_{k,m}$$

Cette relation s'obtient de suite en posant $t_k^{(n)} = \sum_{m=0}^k a_{k,m} S_{n+m}$ dans l'algorithme (IV-7).

Elle s'inverse en :

$$\alpha_k = \frac{a_{k+1,m} - a_{k,m}}{a_{k,m-1} - a_{k,m}} \quad (\text{IV-8})$$

α_k doit être indépendant de m afin d'assurer la régularité de l'algorithme deltoïde.

Exemple : La méthode de sommation d'Euler rentre dans ce cas et l'application de la formule (IV-8) fournit $\alpha_k = \frac{1}{q+1}$, vu que $a_{k,m} = \frac{q^{k-m}}{(q+1)^k} C_k^m$.

La méthode d'Euler admet donc l'algorithme récursif suivant :

$$\begin{cases} t_{k+1}^{(n)} = \frac{1}{q+1} t_k^{(n+1)} + \frac{q}{q+1} t_k^{(n)} \\ t_0^{(n)} = S_n \end{cases}$$

IV-5.2. Algorithme linéaire de losange.

Un certain nombre de procédés de sommation linéaire n'admettent pas de forme récursive deltoïde. Par contre, ils acceptent un algorithme récursif de losange, c'est-à-dire, à 4 termes :

$$\begin{cases} s_{k+1}^{(n)} = \alpha_k s_k^{(n)} + \beta_k s_k^{(n+1)} + \gamma_k s_{k-1}^{(n)} \\ s_{-1}^{(n)} = 0 \quad s_0^{(n)} = S_n \end{cases}$$

La régularité de l'algorithme est assurée par la condition $\alpha_k + \beta_k + \gamma_k = 1$ et il est possible d'exprimer les éléments a_{ij} de la matrice de sommation en fonction des α_k , β_k et γ_k : il suffit d'introduire la forme $t_k^{(n)} = \sum_{m=0}^k a_{km} S_{n+m}$ dans l'algorithme de losange et d'identifier les termes.

On trouve :
$$\begin{cases} a_{k+1,m} - \alpha_k a_{k,m} - \beta_k a_{k,m-1} - \gamma_k a_{k-1,m} = 0 \\ \alpha_k + \beta_k + \gamma_k = 1 \end{cases}$$

IV-6. L'algorithme ϵ .

Nous abordons, dans ce paragraphe et les suivants, les méthodes d'accélération non linéaires. Elles sont d'invention relativement récentes puisque la doyenne, l'algorithme ϵ , fut publiée en 1956. Ce qui les rend indispensables n'est pas tant leur puissance accélératrice, car un procédé de sommation linéaire bien choisi peut parfois se révéler très performant, que la plus grande richesse de leur noyau qui les rend aptes à être efficaces dans un plus grand nombre de cas.

IV-6.1. Le procédé Δ^2 d'Aitken.

L'algorithme ε est en fait une généralisation d'un procédé d'accélération connu, depuis 1926, sous le nom de procédé Δ^2 d'Aitken. Ou si on préfère formuler les choses autrement, le procédé Δ^2 d'Aitken apparaît comme la première étape de l'algorithme ε .

Etant donnée une suite S_n , le procédé Δ^2 d'Aitken la transforme en une suite $\varepsilon_2^{(n)}$ qui vaut:

$$\varepsilon_2^{(n)} = \frac{S_n S_{n+2} - S_{n+1}^2}{S_{n+2} - 2S_{n+1} + S_n} \quad (\text{IV-9})$$

Une forme déterminante équivalente est très utile pour l'étude théorique du procédé :

$$\varepsilon_2^{(n)} = \frac{\begin{vmatrix} S_n & S_{n+1} \\ \Delta S_n & \Delta S_{n+1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ \Delta S_n & \Delta S_{n+1} \end{vmatrix}}$$

Par exemple, celle-ci permet immédiatement de se rendre compte que le procédé Δ^2 d'Aitken conserve la limite de la suite de départ. Il suffit, pour s'en convaincre, d'y introduire la suite écrite sous la forme: $S_n = S + as(n)$. On trouve, en effet :

$$\varepsilon_2^{(n)} = S + \frac{\begin{vmatrix} as(n) & as(n+1) \\ \Delta as(n) & \Delta as(n+1) \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ \Delta as(n) & \Delta as(n+1) \end{vmatrix}} = S + as'(n)$$

Par contre, rien dans cette formule ne permet d'affirmer que la suite $\varepsilon_2^{(n)}$ converge plus vite que S_n . Il faudrait, pour cela, que l'on ait : $\lim_{n \rightarrow \infty} \frac{as'(n)}{as(n)} = 0$, ce qui n'est certainement pas vrai en toute généralité. Le cas $as(n) = 1/n$ fournit un contre-exemple facile à vérifier :

$$\text{si } S_n \approx S + 1/n + \dots \quad \text{alors } \varepsilon_2^{(n)} \approx S + 1/(2n) + \dots$$

On constate que l'asymptote est simplement amortie d'un facteur 2, sans avoir disparu pour autant ; cette amélioration est bien trop minime pour être intéressante en pratique.

La question se pose alors de savoir quel type de suite la transformation $\varepsilon_2^{(n)}$ accélère valablement. C'est la détermination du noyau de la transformation qui permet de répondre à cette question.

Recherchons pour quelle suite S_n on a exactement : $\varepsilon_2^{(n)} = S$, à l'ordre n .

Il faut et il suffit que :
$$\begin{vmatrix} S_n & S_{n+1} \\ \Delta S_n & \Delta S_{n+1} \end{vmatrix} = \begin{vmatrix} S & S \\ \Delta S_n & \Delta S_{n+1} \end{vmatrix} \quad \forall n > N$$

ou encore :
$$\begin{vmatrix} S_n - S & S_{n+1} - S \\ S_{n+1} - S_n & S_{n+2} - S_{n+1} \end{vmatrix} = 0$$

ou enfin :
$$\begin{vmatrix} S_n - S & S_{n+1} - S \\ S_{n+1} - S & S_{n+2} - S \end{vmatrix} = 0$$

Il faut donc et il suffit qu'il existe deux constantes a_0 et a_1 telles que l'on ait :

$$a_0(S_n - S) + a_1(S_{n+1} - S) = 0 \quad \forall n > N \quad (\text{IV-10})$$

qui est une récurrence linéaire homogène d'ordre un. Sa solution est évidente; elle s'écrit :

$$S_n = S + \alpha \lambda^n \quad (\lambda \neq 1)$$

Tel est le noyau du procédé $\Delta^2 \equiv \varepsilon_2^{(n)}$. Il contient toutes les suites exponentiellement convergentes et divergentes quelles que soient les valeurs de α mais aussi de λ . Le progrès est évident par rapport aux méthodes linéaires qui ne pouvaient accélérer que les suites du type $S_n = S + \alpha \lambda^n$ à la condition que le paramètre λ ait une valeur précise et imposée.

Le procédé Δ^2 se présente, en fait, comme une méthode d'extrapolation par une exponentielle: étant donnés trois termes consécutifs d'une suite S_n , S_{n+1} et S_{n+2} , il calcule le meilleur ajustage exponentiel du type:

$$\begin{cases} S_n = S + \alpha \lambda^n \\ S_{n+1} = S + \alpha \lambda^{n+1} \\ S_{n+2} = S + \alpha \lambda^{n+2} \end{cases}$$

On retrouve l'écriture (IV-9) du procédé Δ^2 en résolvant ce système par rapport à S .

Vu la forme du noyau du procédé, on ne s'étonne pas qu'il soit inefficace dans le cas des suites du type $S_n \approx S + \lambda/n + \dots$.

Remarque: si on applique le procédé d'Aitken à la suite des sommes partielles d'une série de Mac Laurin, $S_n = c_0 + c_1 z + \dots + c_n z^n$, on trouve :

$$\begin{aligned}
\mathcal{E}_2^{(n)} &= c_0 + c_1 z + \dots + c_n z^n + \frac{\begin{vmatrix} 0 & c_{n+1} z^{n+1} \\ c_{n+1} z^{n+1} & c_{n+2} z^{n+2} \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ c_{n+1} z^{n+1} & c_{n+2} z^{n+2} \end{vmatrix}} \\
&= c_0 + c_1 z + \dots + c_n z^n + \frac{c_{n+1}^2 z^{n+1}}{c_{n+1} - c_{n+2} z} = [n+1/1]
\end{aligned}$$

Le procédé Δ^2 fournit donc, dans ce cas, la première ligne de la table de Padé.

IV-6.2. La transformation de Shanks.

Shanks a recherché quelle transformation de suite est susceptible de généraliser la relation (IV-10) sous la forme :

$$\sum_{i=0}^k a_i (S_{n+i} - S) = 0 \quad \forall n > N \quad (\text{IV-11})$$

Il a évidemment trouvé qu'il fallait qu'on ait :

$$\begin{vmatrix} S_n - S & \dots & S_{n+k} - S \\ S_{n+1} - S & \dots & S_{n+k+1} - S \\ \vdots & & \vdots \\ S_{n+k} - S & \dots & S_{n+2k} - S \end{vmatrix} = 0$$

En d'autres termes, il a trouvé une transformation, notée $\mathcal{E}_{2k}^{(n)}$, qui transforme la suite S_n , obéissant à (IV-11), en la suite constante $\{S\}$:

$$\mathcal{E}_{2k}^{(n)} = \frac{\begin{vmatrix} S_n & S_{n+1} & \dots & S_{n+k} \\ \Delta S_n & \Delta S_{n+1} & \dots & \Delta S_{n+k} \\ \vdots & \vdots & & \vdots \\ \Delta S_{n+k-1} & \Delta S_{n+k} & \dots & \Delta S_{n+2k-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \dots & 1 \\ \Delta S_n & \Delta S_{n+1} & \dots & \Delta S_{n+k} \\ \vdots & \vdots & & \vdots \\ \Delta S_{n+k-1} & \Delta S_{n+k} & \dots & \Delta S_{n+2k-1} \end{vmatrix}} \quad (\text{IV-12})$$

Lorsque $k = 1$, on retrouve évidemment le procédé Δ^2 . Lorsque $k > 1$, on découvre de nouveaux procédés d'accélération plus puissants que lui. Voyons à quels types de suites ils s'appliquent préférentiellement en déterminant leur noyau.

La relation (IV-11) montre clairement que S_n -S obéit à une récurrence linéaire à coefficients constants dont la solution générale revêt nécessairement la forme d'exponentielles-polynômes. Le noyau de la transformation de Shanks est donc constitué par l'ensemble des suites du type:

$$S_n = S + \sum_{i=1}^N p_i(n) e^{z_i n} \quad (IV-13)$$

où les $p_i(n)$ sont des polynômes en n et où les z_i sont des nombres quelconques, réels ou complexes. La grosseur relative de ce noyau explique son importance pratique.

Cependant on notera que les suites, fréquemment rencontrées, du type :

$$S_n \approx S + \lambda_1 / n + \lambda_2 / n^2 + \dots$$

n'en font pas partie, en sorte qu'elles ne bénéficient pas de l'application du procédé de Shanks.

IV-6.3. Transformation de Shanks et approximants de Padé.

Si on applique la transformation de Shanks aux séries de puissances $S_n = \sum_{i=0}^n c_i z^i$, en tenant compte de ce que: $\Delta S_n = S_{n+1} - S_n = c_{n+1} z^{n+1}$, on trouve que la relation (IV-12) se réduit à la forme déterminante de l'approximant de Padé $[n+k/k]$. On a donc l'identité importante:

$$\mathcal{E}_{2k}^{(n)} = [n + k / k] \quad (IV-14)$$

Cette identité a une conséquence importante: la table de Padé n'est utile que dans les cas où la suite $S_n = \sum_{i=0}^n c_i z^i$ possède un comportement asymptotique du type (IV-13).

Ceci explique pourquoi la table de Padé est impuissante à accélérer la série dilogarithme

$$S_n = \sum_{i=0}^n \frac{z^i}{(i+1)^2} \text{ qui se comporte asymptotiquement comme } z^{n+1}/(n+1).$$

C'est l'explication des limitations évoquées au § III-6. On notera que le calcul des quantités $\mathcal{E}_{2k}^{(n)}$ nécessite la connaissance de $2k+1$ termes de la suite, soit : $S_n, S_{n+1}, \dots, S_{n+2k}$.

La transformation de Shanks réalise, en fait, l'extrapolation d'une suite donnée S_n par une somme d'exponentielles-polynômes du type (IV-13) et ajuste automatiquement non seulement les coefficients mais aussi les exposants. Tel est le bénéfice qu'on retire de la non-linéarité de l'algorithme.

IV-6.4. L'algorithme ϵ .

Dû à Wynn, cet algorithme procède au calcul récursif des quantités $\epsilon_{2k}^{(n)}$ de la transformation de Shanks définies par la relation (IV-12). Il est rendu indispensable par le fait que les calculs de déterminants sont généralement proscrits, en analyse numérique, à cause des trop fréquentes erreurs de troncature qu'ils provoquent. Rappelons, en effet, que le calcul d'un déterminant $n \times n$ implique en gros $n!/2$ additions et $n!/2$ soustractions de quantités voisines.

Les règles de l'algorithme ϵ sont les suivantes :

$$\begin{cases} \epsilon_{-1}^{(n)} = 0 & \epsilon_0^{(n)} = S_n & (n = 0, 1, \dots) \\ \epsilon_{k+1}^{(n)} = \epsilon_{k-1}^{(n+1)} + \frac{1}{\epsilon_k^{(n+1)} - \epsilon_k^{(n)}} & (k, n = 0, 1, \dots) \end{cases} \quad (\text{IV-15})$$

Les quantités d'indice inférieur impair ne sont que des intermédiaires de calcul. Seules les quantités d'indice pair, $\epsilon_{2k}^{(n)}$, sont intéressantes au point de vue de l'accélération de la convergence des suites. Nous ne reproduisons pas la démonstration de la relation (IV-15) qui exige des manipulations fastidieuses de déterminants.

IV-6.5. Convergence de l'algorithme ϵ .

Il serait faux de croire que l'algorithme ϵ ne s'applique avec succès qu'aux suites du type combinaison exponentielles-polynômes (IV-13). Un premier indice, à l'appui de cette thèse, est fourni par la fonction exponentielle pour laquelle on a :

$$S_n = \sum_{i=0}^n x^i / i! \approx e^x + O\left(\frac{x^{n+1}}{(n+1)!}\right)$$

On a également remarqué qu'en pratique, les séries alternées étaient bien accélérées par l'algorithme ϵ : elles correspondent à des suites qui oscillent autour de leur limite. Par contre, les suites qui tendent monotonément vers leur limite se laissent plus difficilement accélérer par l'algorithme ϵ sauf évidemment si elles sont du type (IV-13).

Cette remarque est illustrée par les deux exemples suivants :

$$1) \quad S_n \approx S + \sum_{i=0}^{\infty} \frac{a_i}{(n+b)^i} \quad (a_i \neq 0)$$

L'asymptote principale de cette suite est du type : $as(n) = \frac{1}{n+b}$

$$\text{Pour } k \text{ fixé, on trouve : } \epsilon_{2k}^{(n)} \approx S + \frac{a_1}{(k+1)(n+k+b)}$$

$$\text{d'asymptote principale : } as'(n) = \frac{1}{(k+1)(n+k+b)}$$

La suite s'en trouve fort peu accélérée puisque : $\lim_{n \rightarrow \infty} \frac{as'(n)}{as(n)} = \frac{1}{k+1} \neq 0$

2) Par contre, dans le cas de la suite oscillante :

$$S_n \approx S + (-1)^n \sum_{i=1}^{\infty} \frac{a_i}{(n+b)^i} \quad (a_i \neq 0)$$

d'asymptote principale : $as(n) = \frac{(-1)^n}{n+b}$

on trouve : $\varepsilon_{2k}^{(n)} \approx S + \frac{(-1)^n k!^2 a_1}{2^{2k} (n+b)^{2k+1}}$

avec cette fois : $as'(n) = \frac{(-1)^n k!^2}{2^{2k} (n+b)^{2k+1}}$

et l'accélération est, de toute évidence, spectaculaire.

Cette situation est générale et explique, en particulier, pourquoi les séries de Fourier sont accélérées par l'algorithme ε .

L'étude de la stabilité de l'algorithme ε appelle une remarque similaire : à cause de la différence des quantités $\varepsilon_k^{(n+1)}$ et $\varepsilon_k^{(n)}$ qui intervient dans son écriture (IV-15), il est, en général, instable pour les suites monotones et stable pour les suites oscillantes. Il n'existe toutefois pas de théorème valable en toute généralité dans ce domaine.

IV-6.6 Généralisations de l'algorithme ε .

Deux généralisations de l'algorithme ε existent qui font usage d'une suite auxiliaire x_n .

- Première généralisation : ses règles s'écrivent :

$$\begin{cases} \varepsilon_{-1}^{(n)} = 0 & \varepsilon_0^{(n)} = S_n & (n = 0, 1, \dots) \\ \varepsilon_{k+1}^{(n)} = \varepsilon_{k-1}^{(n+1)} + \frac{x_{n+1} - x_n}{\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}} & (n, k = 0, 1, \dots) \end{cases}$$

Appliquons-les à l'exemple numérique suivant emprunté à Brezinski : soit la suite

$$S_n = 1 + 3e^{-1.4x_n} \quad \text{avec} \quad x_n = 1.1^{n-1} \quad \text{qui converge très lentement vers 1.}$$

On a, en effet, $S_0 = 1.73979\dots$; $S_1 = 1.64314\dots$; ... ; $S_4 = 1.38631\dots$; ...

On trouve $\varepsilon_4^{(0)} = 1.73799\dots$ par l'algorithme ε simple.

$\varepsilon_4^{(0)} = 1.00272\dots$ par sa première généralisation.

- Deuxième généralisation : ses règles s'écrivent :

$$\begin{cases} \varepsilon_{-l}^{(n)} = 0 & \varepsilon_0^{(n)} = S_n & (n = 0, 1, \dots) \\ \varepsilon_{k+l}^{(n)} = \varepsilon_{k-l}^{(n+1)} + \frac{x_{n+k+l} - x_{n+k}}{\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}} & (n, k = 0, 1, \dots) \end{cases}$$

Cette variante est également efficace pour l'exemple précédent car on trouve $\varepsilon_4^{(0)} = 1.00224\dots$

Autre exemple numérique : soit la suite $S_n = n \sin \frac{1}{n}$ qui tend vers 1.

On a, par exemple $S_{37} = 0.99987\dots$

On trouve successivement $\varepsilon_{36}^{(0)} = 0.9999938\dots$ par l'algorithme ε simple.

$\varepsilon_{36}^{(0)} = 1.0000000027\dots$ par sa première généralisation et

$\varepsilon_{36}^{(0)} = 0.99999999976\dots$ par sa deuxième généralisation
à la condition de bien choisir la suite x_n , soit $x_n = \ln(n+1)$.

Ce choix judicieux de la suite auxiliaire x_n correspond évidemment au comportement asymptotique de la suite S_n mais la relation est pratiquement impossible à établir avec certitude dans la mesure où les noyaux des généralisations de l'algorithme ε ne sont pas connus explicitement. En pratique, on se contente de mettre quelques suites caractéristiques au banc d'essai, quitte à changer de méthode si cela ne donne rien.

IV-7. L'algorithme θ .

C'est un des algorithmes les plus puissants de la littérature.

Jusqu'à présent, les algorithmes non linéaires étudiés étaient tous de la forme dite de losange :

$$\begin{cases} \theta_{-l}^{(n)} = 0 & \theta_0^{(n)} = S_n & (n = 0, 1, \dots) \\ \theta_{k+l}^{(n)} = \theta_{k-l}^{(n+1)} + D_k^{(n)} & (n, k = 0, 1, \dots) \end{cases} \quad (\text{IV-16})$$

où la forme de $D_k^{(n)}$ variait selon que l'on considérait l'algorithme ρ , ε ou ses généralisations.

C'est par un raisonnement purement intuitif que Brezinski a découvert les règles de l'algorithme θ . Partant du schéma (IV-16) ci-dessus, il a tenté la comparaison des vitesses de convergence des suites $\theta_{2k+2}^{(n)}$ et $\theta_{2k}^{(n)}$ censées converger vers la même limite, la première plus vite que la deuxième. On traduit cette exigence en posant que (Δ porte sur l'indice n) :

$$\lim_{n \rightarrow \infty} \frac{\Delta \theta_{2k+2}^{(n)}}{\Delta \theta_{2k}^{(n)}} = 0 \quad (\text{IV-17})$$

Cela provient en effet de ce que, lorsqu'une suite T_n converge plus vite que S_n vers la même limite, on a très généralement que :

$$\lim_{n \rightarrow \infty} \frac{\Delta T_n}{\Delta S_n} = 0$$

Or les règles (IV-16) donnent :

$$\theta_{2k+2}^{(n)} = \theta_{2k}^{(n+1)} + D_{2k+1}^{(n)}$$

et la condition (IV-17) devient :

$$\lim_{n \rightarrow \infty} \frac{\Delta D_{2k+1}^{(n)}}{\theta_{2k}^{(n+1)}} = -1 \quad (\text{IV-18})$$

L'astuce consiste maintenant à modifier les règles (IV-16) en introduisant un paramètre w_k , variable avec k que l'on ajuste de façon convenable :

$$\begin{cases} \theta_{-1}^{(n)} = 0 & \theta_0^{(n)} = S_n \\ \theta_{2k+1}^{(n)} = \theta_{2k-1}^{(n+1)} + D_{2k}^{(n)} \\ \theta_{2k+2}^{(n)} = \theta_{2k}^{(n+1)} + w_k D_{2k+1}^{(n)} \end{cases}$$

La condition (IV-18) se transcrit de façon évidente :

$$\lim_{n \rightarrow \infty} \frac{\Delta \theta_{2k}^{(n+1)}}{\Delta D_{2k+1}^{(n)}} = -w_k$$

En pratique, cette limite n'est pas connue, d'où l'idée de poser simplement

$$w_k = -\frac{\Delta \theta_{2k}^{(n+1)}}{\Delta D_{2k+1}^{(n)}}$$

en reprenant pour $D_k^{(n)}$, l'expression en vigueur dans l'algorithme ε , soit :

$$D_k^{(n)} = \frac{1}{\theta_k^{(n+1)} - \theta_k^{(n)}}$$

Tenant compte de toutes ces indications, on trouve les règles définitives de l'algorithme θ , où l'opérateur Δ agit sur la variable n :

$$\theta_{-1}^{(n)} = 0 \quad \theta_0^{(n)} = S_n \quad (n = 0, 1, \dots)$$

$$\theta_{2k+1}^{(n)} = \theta_{2k-1}^{(n+1)} + \frac{1}{\Delta \theta_{2k}^{(n)}}$$

$$\theta_{2k+2}^{(n)} = \frac{\Delta [\theta_{2k}^{(n+1)} \Delta \theta_{2k+1}^{(n)}]}{\Delta^2 \theta_{2k+1}^{(n)}}$$

Les θ d'indice inférieur impair sont des intermédiaires de calcul et seuls les θ d'indice pair sont retenus.

Il importe de noter que la démarche suivie est purement intuitive et que la seule justification véritable de cet algorithme réside, a posteriori, dans son extrême efficacité. Facile à programmer, cet algorithme est très difficile à étudier du point de vue théorique. On note que :

- Ce n'est plus un algorithme de losange. Par exemple, le calcul de $\theta_2^{(n)}$, première étape de l'algorithme, nécessite la connaissance de quatre termes consécutifs : S_n , S_{n+1} , S_{n+2} et S_{n+3} alors que les algorithmes ε ou ρ n'avaient besoin que de trois termes successifs.
- Sauf si $k = 1$, il n'existe pas de forme déterminante pour $\theta_{2k}^{(n)}$ comme c'était le cas pour $\varepsilon_{2k}^{(n)}$, (c'était la transformation de Shanks) ou pour $\rho_{2k}^{(n)}$ (nous ne l'avons pas donnée mais elle existe). C'est cette carence qui rend difficile l'étude théorique de l'algorithme θ sauf dans sa première étape.
- Pour la même raison, il semble impossible de déterminer le noyau de l'algorithme θ , sauf dans sa première étape. On a en effet que :

$$\theta_2^{(n)} = \frac{\begin{vmatrix} S_n & S_{n+1} & S_{n+2} \\ \Delta S_n & \Delta S_{n+1} & \Delta S_{n+2} \\ 2\Delta S_n & \Delta S_{n+1} & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ \Delta S_n & \Delta S_{n+1} & \Delta S_{n+2} \\ 2\Delta S_n & \Delta S_{n+1} & 0 \end{vmatrix}} = \frac{\Delta \left[S_{n+1} \Delta \frac{1}{\Delta S_n} \right]}{\Delta^2 \left[\frac{1}{\Delta S_n} \right]}$$

$$\theta_2^{(n)} = S \quad \forall n \quad \text{ssi} :$$

$$\Delta \left[(S_{n+1} - S) \Delta \frac{1}{\Delta S_n} \right] = 0 \quad d'où \quad (S_{n+1} - S) \Delta \frac{1}{\Delta S_n} = v$$

où v est une constante arbitraire.

$$(S_{n+1} - S) \left[\frac{1}{S_{n+2} - S_{n+1}} - \frac{1}{S_{n+1} - S_n} \right] = v$$

que l'on réécrit, un peu différemment, sous la forme :

$$\frac{S_{n+1} - S}{S_{n+2} - S_{n+1}} - \left[\frac{S_{n+1} - S}{S_{n+1} - S_n} - I \right] = \nu + I$$

Si on pose $\tau_n = \frac{S_n - S}{S_{n+1} - S_n}$

l'équation précédente devient : $\tau_{n+1} - \tau_n = \nu + I$, récurrence inhomogène du premier ordre à coefficients constants dont la solution générale s'écrit :

$$\tau_n = (\nu + I)n + \mu$$

On trouve donc :

$$I + \frac{I}{\tau_n} = \frac{S_{n+1} - S}{S_n - S} = \frac{(\nu + I)n + (\mu + I)}{(\nu + I)n + \mu} \quad (IV-19)$$

Cette récurrence d'ordre 1 qui porte sur $S_n - S$, se résout à nouveau sans difficulté. En résumé, on trouve que les suites S_n qui constituent le noyau de la première étape de l'algorithme θ s'écrivent sous l'une des formes :

$$S_n = S + \lambda \frac{\Gamma(n+a)}{\Gamma(n+b)} \quad (si \ \nu \neq -I)$$

$$S_n = S + \lambda \rho^n \quad (si \ \nu = -I)$$

On comparera ce noyau à celui de la première étape de l'algorithme ε qui s'écrivait simplement $S_n = S + \lambda \rho^n$ et on mesure le progrès accompli par l'algorithme θ . En particulier, on comprend pourquoi les suites du type $S_n \approx S + \frac{\lambda}{n+b} + \dots$ sont bien accélérées par l'algorithme θ puisqu'elles appartiennent au noyau avec le choix $b = a + 1$.

IV-8. L'algorithme E.

Dû à Havie et Brezinski, l'algorithme E permet de comprendre comment fonctionnent les méthodes non-linéaires d'accélération de la convergence.

Supposons que l'on connaisse l'échelle asymptotique complète d'une suite donnée S_n :

$$S_n = S + a_1 g_1(n) + a_2 g_2(n) + a_3 g_3(n) + \dots$$

avec : $\lim_{n \rightarrow \infty} \frac{g_{i+1}(n)}{g_i(n)} = 0 \quad \forall i$

L'idée vient tout naturellement à l'esprit d'extrapoler la suite S_n au moyen de son échelle asymptotique propre. Il y a lieu, pour cela, de résoudre le système :

$$S_{n+i} = S + a_1 g_1(n+i) + \dots + a_k g_k(n+i) \quad i = 0, 1, \dots, k$$

C'est un système de $k + 1$ équations pour les $k + 1$ inconnues : a_1, \dots, a_k et S . Sauf le cas exceptionnel où S_n serait exactement égale à : $S + a_1 g_1(n) + \dots + a_k g_k(n)$, la solution obtenue en résolvant ce système dépend de n et de k . Notons $E_k^{(n)}$, la valeur obtenue pour S de cette façon. La règle de Cramer permet de trouver une forme déterminante pour $E_k^{(n)}$:

$$E_k^{(n)} = \frac{\begin{vmatrix} S_n & \dots & S_{n+k} \\ g_1(n) & \dots & g_1(n+k) \\ \vdots & & \vdots \\ g_k(n) & \dots & g_k(n+k) \end{vmatrix}}{\begin{vmatrix} 1 & \dots & 1 \\ g_1(n) & \dots & g_1(n+k) \\ \vdots & & \vdots \\ g_k(n) & \dots & g_k(n+k) \end{vmatrix}} = \frac{dtm(S_n \ g_1(n) \dots g_k(n))}{dtm(1 \ g_1(n) \dots g_k(n))} \quad (IV-20)$$

On voit, grâce à cette forme déterminante, que, si les fonctions $g_i(n)$ font réellement partie de l'échelle asymptotique propre de S_n , la suite $E_k^{(n)}$ convergera vers S beaucoup plus vite que S_n . En effet, vu les propriétés des déterminants, on trouve que le comportement asymptotique de la suite $E_k^{(n)}$ est de l'ordre de grandeur de la première asymptote négligée,

$$\text{soit : } E_k^{(n)} \approx S + O(g_{k+1}(n))$$

Ceci représente l'optimum imaginable, quelle que soit la procédure accélératrice employée. On voit donc que, lorsqu'on doit accélérer une suite S_n , l'idéal est d'en connaître l'échelle asymptotique propre. La connaissance des fonctions $g_i(n)$ suffit : il n'est pas nécessaire de connaître les coefficients a_i qui pondèrent les asymptotes. Cela est heureux car ceux-ci sont d'accès beaucoup plus difficile en pratique. Si certains coefficients a_i étaient connus, il est clair qu'on pourrait regrouper plusieurs fonctions $g_i(n)$ en une seule asymptote et l'accélération s'en trouverait accrue mais ce cas est très rare.

Lorsque l'échelle asymptotique est inaccessible ou simplement inconnue, il convient d'envisager une des stratégies suivantes :

- soit essayer une échelle asymptotique qu'on se donne, à priori, en espérant qu'elle ne sera pas trop éloignée de l'échelle véritable. Il y a lieu de noter que, quand le choix effectué ne donne pas satisfaction, on peut toujours essayer une des douze généralisations mises au point par Brezinski qui sont à l'algorithme E ce que θ est à ε . Ces généralisations ont pour but de tenter de rectifier

le choix médiocre effectué pour les $g_i(n)$. Nous entrons ici dans un domaine trop spécialisé pour qu'il figure en détail dans ce cours.

- soit recourir à une forme implicite de l'algorithme. Par exemple, l'analogie entre l'équation (IV-20) et la transformation de Shanks indique qu'il suffit de poser : $g_i(n) = \Delta S_{n+i-1}$ dans (IV-20) pour retrouver l'algorithme ε qui apparaît, de ce fait, comme un cas particulier de l'algorithme E, où on considère que la suite S_n se comporte asymptotiquement comme :

$$S_n \approx S + a_1 \Delta S_n + a_2 \Delta S_{n+1} + \dots + a_k \Delta S_{n+k-1}$$

La forme est dite implicite parce que S_n figure dans l'écriture des asymptotes $g_i(n)$.

On retrouve bien la récurrence à coefficients constants qui définit le noyau de l'algorithme ε et dont la solution générale est une somme d'exponentielles-polynômes.

La généralité de l'algorithme est telle qu'il englobe beaucoup de procédés connus d'accélération. Un autre exemple est la transformation de Richardson qui correspond au choix $g_i(n) = x_n^i$. Nous verrons plus loin d'autres exemples encore.

En résumé, l'algorithme E propose un choix : soit recourir à une échelle asymptotique explicite, qu'elle soit exacte (cas optimal) ou approchée, ou choisir une échelle asymptotique implicite où les S_i interviennent dans l'écriture des $g_i(n)$. Ce dernier cas est pratiqué toutes les fois qu'on n'a aucune idée du comportement asymptotique véritable de la suite S_n . On espère ainsi que le choix implicite se rapproche de la réalité. La grande variété des algorithmes existants n'a d'autre justification que d'offrir le plus vaste éventail de méthodes plus ou moins aptes à accélérer des suites aux comportements asymptotiques les plus divers. Cela signifie également qu'il n'existe pas d'algorithme d'accélération universelle puisque toute procédure doit être adaptée au comportement asymptotique de la suite considérée.

IV-9.1. L'algorithme E récursif.

Il n'est jamais recommandé de calculer de grands déterminants du type (IV-20). Heureusement, il existe un algorithme récursif qui calcule les quantités $E_k^{(n)}$. Il se note :

$$\left\{ \begin{array}{l} E_0^{(n)} = S_n \quad g_{0,i}^{(n)} = g_i(n) \quad \left(\begin{array}{l} n = 0, 1, \dots \\ i = 1, 2, \dots \end{array} \right) \\ E_k^{(n)} = \frac{E_{k-1}^{(n)} g_{k-1,k}^{(n+1)} - E_{k-1}^{(n+1)} g_{k-1,k}^{(n)}}{g_{k-1,k}^{(n+1)} - g_{k-1,k}^{(n)}} \\ g_{k,i}^{(n)} = \frac{g_{k-1,i}^{(n)} g_{k-1,k}^{(n+1)} - g_{k-1,i}^{(n+1)} g_{k-1,k}^{(n)}}{g_{k-1,k}^{(n+1)} - g_{k-1,k}^{(n)}} \quad (k \geq 1, n \geq 0, i \geq k+1) \end{array} \right. \quad (IV-21)$$

IV-9.2. Illustration d'une échelle asymptotique exacte.

Considérons l'exemple de la suite : $S_n = (\text{PIC})_n$:

$$S_n = \sum_{i=0}^n \frac{1}{(i+1)^2} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{(n+1)^2}$$

Cette suite tend vers $\pi^2/6 = 1.644934\dots$

Nous avons vu au paragraphe IV-1. que son comportement asymptotiques est donné par :

$$S_n \approx S + \lambda_1/n + \lambda_2/n^2 + \dots \quad \text{d'où} \quad g_i(n) = n^{-i}.$$

L'algorithme optimal prescrit donc de calculer les expressions suivantes :

$$E_k^{(n)} = \frac{\text{dtm}\left(S_n, \frac{1}{n+\gamma}, \frac{1}{(n+\gamma)^2}, \dots, \frac{1}{(n+\gamma)^k}\right)}{\text{dtm}\left(1, \frac{1}{n+\gamma}, \frac{1}{(n+\gamma)^2}, \dots, \frac{1}{(n+\gamma)^k}\right)}$$

Le dénominateur se calcule sans peine par la règle de Van der Monde. Il reste à développer le numérateur selon les éléments de la première colonne et à utiliser la même règle pour obtenir le résultat du quotient final :

$$E_k^{(n)} = \sum_{m=0}^k \frac{(-1)^{m+k}}{k!} (n+m+\gamma)^k C_k^m S_{n+m}$$

On retrouve le procédé de Salzer qui apparaît comme optimal pour les suites qui se comportent comme :

$$S_n \approx S + a_1/n + a_2/n^2 + \dots$$

On peut calculer les $E_k^{(n)}$ et les disposer dans une table de ce type :

$E_0^{(0)} = S_0$	$E_1^{(0)}$	$E_2^{(0)}$	$E_3^{(0)}$	$E_4^{(0)}$...
$E_0^{(1)} = S_1$	$E_1^{(1)}$	$E_2^{(1)}$	$E_3^{(1)}$
$E_0^{(2)} = S_2$	$E_1^{(2)}$	$E_2^{(2)}$
$E_0^{(3)} = S_3$	$E_1^{(3)}$
$E_0^{(4)} = S_4$
...

Table IV-1.

Dans le cas de l'exemple numérique étudié, on trouve les valeurs, $E_k^{(n)}$, suivantes ainsi que le tableau des précisions :

1	3/2	13/8	355/216	1895/1152
5/4	19/12	59/36	1421/864	
49/36	29/18	473/288		
205/144	1169/720			
5269/3600				

Table IV-2.

0.9	2.4	4.4	7.1	10.9
1.4	3.3	5.6	8.8	
1.7	3.9	6.5		
2.0	4.3			
2.2				

Table IV-3.

L'accélération est optimale pour le terme $E_4^{(0)}$ qui s'avère correct avec cinq chiffres décimaux. On voit donc que l'information que constituent ces cinq chiffres est déjà présente dans les cinq premiers termes de la suite S_0, S_1, \dots, S_4 .

IV-8.3. Application de l'algorithme E au calcul des intégrales définies par la méthode des trapèzes.

La méthode des trapèzes permet de calculer des intégrales définies grâce à la formule suivante :

$$I = \int_0^1 f(x) dx = \lim_{n \rightarrow \infty} S_n \quad \text{où} \quad S_n = \frac{1}{n} \sum_{k=0}^n f\left(\frac{k}{n}\right)$$

Σ'' signifie que l'on compte les termes $k = 0$ et $k = n$ pour moitié seulement. On montre, en analyse numérique, que la suite S_n se comporte asymptotiquement comme :

$$S_n \approx S + \frac{C_1}{n^2} + \frac{C_2}{n^4} + \frac{C_3}{n^6} + \dots$$

$$C_1 = \frac{1}{12} (f'(1) - f'(0))$$

avec :

$$C_2 = -\frac{1}{720} (f'''(1) - f'''(0)) \quad \text{etc...}$$

En application de l'algorithme E, on trouve la formule sommatoire suivante :

$$E_k^{(n)} = \frac{dtm(S_n \ n^{-2} n^{-4} \dots n^{-2k})}{dtm(1 \ n^{-2} n^{-4} \dots n^{-2k})}$$

Les déterminants se calculent à nouveau par la règle de Van der Monde et on trouve :

$$E_k^{(n)} = \frac{\sum_{j=0}^k (-1)^j S_{n+j} C_k^j (n+j)^{2k+1} \frac{(2n+j-1)!}{(2n+k+j)!}}{\sum_{j=0}^k (-1)^j C_k^j (n+j)^{2k+1} \frac{(2n+j-1)!}{(2n+k+j)!}} = 2 \sum_{j=0}^k (-1)^{k+j} \frac{(2n+j-1)! (n+j)^{2k+1}}{j! (k-j)! (2n+k+j)!} S_{n+j}$$

Exemple : Soit à calculer l'intégrale : $\int_0^1 \frac{dx}{x+1} = \ln 2 = 0.6931471805599453094\dots$

Douze termes dans la suite fournissent treize chiffres décimaux corrects :

$$E_{11}^{(0)} = \mathbf{0.693147180559888\dots}$$

IV-8.4. Les transformations de Levin.

On pourrait reproduire le schéma précédent sur un grand nombre d'échelles asymptotiques afin de déduire, chaque fois, de (IV-20), la formule sommatoire qui s'applique. Il suffit que les déterminants soient calculables exactement pour que le calcul récursif devienne superflu. Par exemple, la méthode d'interpolation polynomiale de Richardson se retrouve par ce procédé, en posant, dans (IV-20), $g_i(n) = x_n^i$ et en calculant à nouveau les déterminants par la formule de Van der Monde.

Un autre exemple, d'une importance extrême, est constitué par les transformations de Levin qui, avec l'algorithme θ , s'avèrent être actuellement les meilleurs algorithmes automatiques d'accélération de la convergence. Pour obtenir les deux transformations de Levin nommées u et t , on commence par faire le choix d'une forme générale pour les $g_i(n)$: $g_i(n) = x_n^{i-1} \xi_n$ où la suite x_n est à éléments distinct et ξ_n est à éléments quelconques.

Ce choix permet de calculer les déterminants qui définissent $E_k(n)$ grâce à la règle des mineurs et celle de Van der Monde :

$$E_k^{(n)} = \frac{\sum_{m=0}^k \frac{S_{n+m}}{\xi_{n+m}} \pi_n^{k,m}}{\sum_{m=0}^k \frac{1}{\xi_{n+m}} \pi_n^{k,m}} \quad (\text{IV-22})$$

où les
$$\pi_n^{k,m} = \sum_{\substack{r=0 \\ r \neq m}}^k \frac{I}{x_{n+r} - x_{n+m}}$$

Reste à trouver un bon choix pour les suites auxiliaires x_n et ξ_n .

Levin en a trouvé deux qui constituent les transformations, t et u , qui portent son nom.

a) La transformation t de Levin.

Elle se déduit de (IV-22) en posant : $\xi_n = \Delta S_n$ et $x_n = 1/(n+1)$

On trouve la suite transformée t (pour rappel : $t_0^{(n)} = S_n$) :

$$t_k^{(n)} = \frac{\sum_{m=0}^k \frac{S_{n+m}}{\Delta S_{n+m}} (-1)^m C_k^m (n+m+1)^{k-1}}{\sum_{m=0}^k \frac{I}{\Delta S_{n+m}} (-1)^m C_k^m (n+m+1)^{k-1}}$$

b) La transformation u de Levin.

Elle se déduit de (IV-22) en posant : $\xi_n = (n+1) \Delta S_n$ et $x_n = 1/(n+1)$

On trouve la suite transformée u (pour rappel : $u_0^{(n)} = S_n$) :

$$u_k^{(n)} = \frac{\sum_{m=0}^k \frac{S_{n+m}}{\Delta S_{n+m}} (-1)^m C_k^m (n+m+1)^{k-2}}{\sum_{m=0}^k \frac{I}{\Delta S_{n+m}} (-1)^m C_k^m (n+m+1)^{k-2}}$$

La transformation t agit mieux sur les suites oscillantes alors que u est efficace sur les suites monotones. Il ne faut guère chercher de justification théorique rigoureuse au choix des suites auxiliaires qui ont engendré ces algorithmes dont on n'a même pas pu démontrer qu'ils étaient réguliers. L'expérience numérique révèle cependant qu'ils sont efficaces et c'est ce qui leur confère toute leur importance.

Le noyau de la transformation, u , par exemple, se détermine assez aisément en posant $u_k^{(n)} = S \quad \forall n$ d'où

$$\sum_{m=0}^k \frac{S_{n+m} - S}{\Delta S_{n+m}} (-1)^m C_k^m (n+m+1)^{k-2} = 0$$

Posons encore :
$$h_n = \frac{S_n - S}{\Delta S_n} (n+1)^{k-2} \quad (\text{IV-23})$$

Il vient :
$$\sum_{m=0}^k h_{n+m} (-1)^m C_k^m = 0$$

réurrence d'ordre k , à coefficients constants, d'équation caractéristique $(z-1)^k = 0$.

Il en résulte que : $h_n = a_0 + a_1 n + \dots + a_{k-1} n^{k-1}$ et on a la possibilité de remonter à S_n en sommant la récurrence (IV-23) qui est d'ordre 1. Par exemple, le noyau de la deuxième étape ($k = 2$) apparaît identique à celui de θ_2 . On a, en effet, successivement : $h_n = a_0 + a_1 n$, d'où :

$$\frac{S_n - S}{S_{n+1} - S_n} = a_0 + a_1 n \quad \Rightarrow \quad 1 + \frac{1}{a_0 + a_1 n} = \frac{S_{n+1} - S}{S_n - S} = \frac{a_1 n + (a_0 + 1)}{a_1 n + a_0} = \frac{n + \lambda}{n + \mu} \quad \text{si } a_1 \neq 0$$

$$= \rho \quad \text{si } a_1 = 0$$

si $a_1 = 0$, on trouve un morceau de noyau : $S_n = S + A\rho^n$

Si $a_1 \neq 0$, on trouve l'autre morceau : $S_n = S + A \frac{\Gamma(n + \lambda)}{\Gamma(n + \mu)}$

IV-9. Algorithmes pour suites itératives.

Il existe une catégorie de suites qui apparaissent fréquemment en pratique, ce sont les suites itératives. Considérons l'exemple suivant : soit à trouver une approximation de la racine proche de 1.4 de l'équation cubique : $x^3 + 2x^2 + 10x - 20 = 0$. Sa vraie valeur est : $x = 1.368808107...$ Plusieurs schémas itératifs sont envisageables. Nous en avons retenu deux :

$$S_{n+1} = \frac{20}{S_n^2 + 2S_n + 10} \quad S_0 = 1 \quad (\text{par exemple})$$

$$T_{n+1} = \frac{20 - 2T_n^2 - T_n^3}{10} \quad T_0 = 1 \quad (\text{idem})$$

La suite S_n converge lentement : $S_n = \{1, 1.358, 1.295, 1.401, 1.354, \dots\}$.

La suite T_n diverge.

Peut-on espérer trouver une méthode accélératrice pour S_n et pourquoi pas une méthode sommatoire pour T_n ? Deux algorithmes existent à ces fins.

IV-9.1. La transformation GBW.

Due à Germain-Bronne et Wimp, elle est basée sur le recours à un algorithme de type deltoïde :

$$\begin{cases} \sigma_{k+1}^{(n)} = \frac{x_n \sigma_k^{(n+1)} - x_{n+k+1} \sigma_k^{(n)}}{x_n - x_{n+k+1}} & (n, k \geq 0) \\ \sigma_0^{(n)} = S_n \end{cases} \quad (\text{IV-24})$$

où la suite x_n est à particulariser convenablement.

Rappelons que ce schéma n'est rien d'autre que la forme récursive de l'algorithme de triangle :

$$\sigma_k^{(n)} = \sum_{m=0}^k \mu_{km} S_{n+m}$$

$$\text{avec } \mu_{km} = \prod_{i=0, i \neq m}^k \frac{x_{n+i}}{x_{n+i} - x_{n+m}} \quad (\text{et } \sum \mu_{km} = 1)$$

Précisément, le choix qui convient pour x_n est : $x_n = \Delta S_n$. Il permet de construire un algorithme dont le noyau est très différent de ceux étudiés jusqu'à présent. La sommation par (IV-24), avec $x_n = \Delta S_n$ est exacte si S_n est de la forme

$$S_n = S + \lambda_1 \Delta S_n + \lambda_2 (\Delta S_n)^2 + \dots + \lambda_k (\Delta S_n)^k \quad (\text{IV-25})$$

En d'autres termes, si S_n est solution de cette récurrence, on peut montrer que $\sigma_k^{(n)} = S \quad \forall n$. Malheureusement, on ne peut pas expliciter les solutions de cette récurrence qui est hautement non-linéaire. On note qu'un algorithme équivalent s'obtient en appliquant l'algorithme E avec le choix suivant des $g_i(n)$: $g_i(n) = (\Delta S_n)^i$

L'intérêt de la définition (IV-25) du noyau de la transformation repose sur l'argumentation suivante : supposons que nous soyons intéressés par l'accélération d'une suite itérative définie par :

$$S_{n+1} = \phi(S_n).$$

Rappelons que cette récurrence tend vers une racine de $z = \phi(z)$ aux conditions que :

- que $|\phi'(z)| < 1$, au voisinage de S
- que S_0 soit suffisamment proche de S

Si on développe ϕ en série de Taylor, au voisinage de S , on obtient :

$$S_{n+1} = S + c_1(S_n - S) + c_2(S_n - S)^2 + \dots$$

$$\text{posant : } R_n = S_n - S, \text{ on trouve : } \Delta R_n = (c_1 - 1)R_n + c_2 R_n^2 + \dots$$

$$\text{Inversant ce développement, on trouve : } R_n = c_1' \Delta R_n + c_2' (\Delta R_n)^2 + c_3' (\Delta R_n)^3 + \dots$$

$$\text{c'est-à-dire : } S_n = S + c_1' \Delta S_n + c_2' (\Delta S_n)^2 + \dots$$

On retrouve la définition (IV-25) du noyau de la transformation GBW. On s'attend donc à ce que celui-ci fonctionne sur les suites itératives. Les applications numériques le confirment au-delà de tout espoir et on a, en plus, que l'algorithme fonctionne également comme méthode sommatoire lorsque la série itérative ne diverge pas trop violemment.

IV-9.2. Le procédé d'Overholt.

La littérature mentionne également cette autre méthode apparentée à GBW mais qui lui semble inférieure en pratique. Elle est basée sur un algorithme deltoïde du type (IV-24) avec un autre choix de la suite x_n . Plus précisément, ses règles s'énoncent :

$$\begin{cases} V_{k+1}^{(n)} = \frac{(\Delta S_{n+k-1})^{k+1} V_k^{(n+1)} - (\Delta S_{n+k})^{k+1} V_k^{(n)}}{(\Delta S_{n+k-1})^{k+1} - (\Delta S_{n+k})^{k+1}} \\ V_0^{(n)} = S_n \end{cases}$$

Les résultats théoriques manquent sur cet algorithme.

IV-9.3. Méthode de Steffensen.

Pour clore ces paragraphes consacrés aux suites itératives, signalons une approche très différente et encore plus puissante.

Devant calculer une racine de l'équation : $x = \phi(x)$, on peut soit :

- recourir au schéma itératif classique $x_{n+1} = \phi(x_n)$, quitte à utiliser conjointement l'algorithme GBW pour en accélérer la convergence.
- recourir à des algorithmes du type Newton ou Laguerre qui requièrent le calcul d'une ou plusieurs dérivées de $\phi(x)$.
- recourir à des méthodes plus élaborées qui consistent à combiner le schéma itératif $x_{n+1} = \phi(x_n)$ avec un algorithme d'accélération classique. Par exemple, si on le combine, avec la première étape de l'algorithme ε , on a besoin de trois termes consécutifs de la suite définie par:

$x_{n+1} = \phi(x_n)$, c'est-à-dire :

$$S_n, S_{n+1} = \phi(S_n), \text{ et } S_{n+2} = \phi(S_{n+1}) = \phi(\phi(S_n))$$

On peut alors imaginer remplacer l'itération $x_{n+1} = \phi(x_n)$ par cette autre, nettement plus performante, dite de Steffensen :

$$S_{n+1} = \frac{S_n \phi(\phi(S_n)) - \phi^2(S_n)}{\phi(\phi(S_n)) - 2\phi(S_n) + S_n} \quad (\text{IV-26})$$

Exemple numérique : Considérons l'itération suivante :

$$\begin{cases} S_{n+1} = S_n^2 - 1 \\ S_0 = 1.5 \end{cases}$$

destinée à fournir la racine de l'équation $x^2 - x - 1 = 0$, qui vaut $(1 + \sqrt{5})/2 = 1.618033989...$

En fait, cette suite S_n diverge. Par contre, l'algorithme de Steffensen converge et il le fait même plus vite que ne le ferait l'algorithme GBW en resommant la suite divergente.

Naturellement, le croisement que nous avons fait entre l'itération classique $x_{n+1} = \phi(x_n)$ et le procédé Δ^2 peut être généralisé à d'autres algorithmes plus puissants, par exemple, à l'algorithme ε , qui généralise le procédé Δ^2 .

IV-11. Performances comparées des divers algorithmes étudiés.

Nous avons testé les divers algorithmes étudiés sur les suites prédéfinies au paragraphe IV-1.

Dans chaque cas, nous avons retenu treize termes dans la suite à accélérer $\{S_0, S_1, \dots, S_{12}\}$. L'optimum atteint par chaque algorithme est indiqué dans le tableau IV-4 ci-dessous, par le nombre p de chiffres népériens exacts obtenus.

Suites	S_{12}	ε	θ	C/1	C/2	Eul	Salz	ρ_{n+1}	L/t	L/u	gbw	Ov
$\sum_0^n (-1)^k / (k+1)$	2.9	22.4	29.4	6.1	4.3	15.3	D	D	32.4	31.9	9.7	10.2
$\sum_0^n (-1)^k / \sqrt{k+1}$	1.4	22.5	27.9	5.4	3.2	15.8	D	D	32.6	32.7	16.3	11.4
$\sum_0^n 1/(k+1)^2$	3.1	4.4	25.5	D	D	D	28	29.9	4.8	26.6	4.3	4.5
$\sum_0^n (-1)^k k!$	D	7.1	15.5	D*	D*	D*	D	D*	16.7	15.9	1.8	D
$\sum_0^n \sin(k)/k$	3	7.5	3.8	5.1	4.3	4.5	D	D	3.5	3.5	D	4
$S_{n+1} = \frac{20}{S_n^2 + 2S_n + 10}$	11.0	48.6	34.3	D	D	15.5	D	D	22.6	21.8	74	73.2
$S_{n+1} = \frac{20 - 2S_n^2 - S_n^3}{10}$	DO	2.8	3.0	2.7	3.6	2.9	D	0.9	2.7	2.7	16.0	4.1
$\sum_0^n (-1)^{[k/2]} / (k+1)$	6.0	11.8	D	D	D	8.6	D	DO	D	D	D	D

Table IV-4. Légende : D(O) : divergence (oscillante)

D* : divergence atténuée.

On peut faire les commentaires suivants :

- comme prévu, les procédés linéaires, Césaro C/1 et C/2, ou Euler, ne sont jamais indispensables quoique ce dernier sauve l'honneur avec la suite $(LUB)_n$.

- θ et Levin/u sont en concurrence étroite avec un léger avantage à ce dernier qui le fait préférer. De même, on peut généralement écarter Levin/t au profit de Levin/u.
- GBW et Overholt ont des performances fort semblables sur les suites itératives, avec un avantage au premier que l'on retiendra donc.
- Les séries de Fourier ne se laissent accélérer, encore que modérément, que par l'algorithme ε . On étudiera, au chapitre V, une alternative basée sur l'algorithme E.

CHAPITRE V

APPLICATIONS DIVERSES.

Ce chapitre illustre les techniques étudiées sur quelques exemples plus élaborés. Les contrôles numériques et la programmation des algorithmes sont écrits en langage Mathematica.

V-1. Résolution des systèmes algébriques linéaires via l'algorithme ϵ .

Soit à résoudre le système algébrique linéaire:

$$H \vec{x} = \vec{b}$$

où H représente une matrice $m \times m$ donnée et \vec{b} un vecteur également donné. La solution s'écrit:

$$\vec{x} = H^{-1} \vec{b}$$

On évite l'inversion matricielle en recourant au procédé itératif suivant:

$$\vec{x}_{i+1} = H \vec{x}_i + \vec{x}_i - \vec{b} \quad (\vec{x}_0 \text{ quelconque non nul}) \quad (\text{V-1})$$

Si on pose: $\vec{\xi}_i = \vec{x}_i - \vec{x}$

on trouve: $\vec{\xi}_{i+1} = H \vec{\xi}_i + \vec{\xi}_i$

ou encore: $H \vec{\xi}_i = \Delta \vec{\xi}_i$

On sait, par ailleurs, que toute matrice annule son polynôme minimum, $Pol(\lambda) = \sum_0^v C_k \lambda^k$.

D'où: $\sum_0^v C_k H^k = 0$

soit encore: $\sum_0^v C_k H^k \vec{\xi}_i = 0$ ou enfin: $\sum_0^v C_k \Delta^k \vec{\xi}_i = 0$

Chaque composante de $\vec{\xi}_i$ obéit donc à une récurrence à coefficients constants, d'ordre v . Il en résulte que chaque composante de la suite $\vec{x}_i = \vec{x} + \vec{\xi}_i$ appartient au noyau de l'algorithme ε et que sa limite, la composante de même rang de \vec{x} , est atteinte exactement à l'étape ε_{2v} . Lorsque les valeurs propres de la matrice H sont distinctes, on a $v=n$ et il faut poursuivre le calcul jusqu'à l'ordre $2n$.

Exemple: soit à résoudre le système:

$$\begin{pmatrix} 1 & 0 & 2 & -1 \\ 3 & 2 & -1 & 1 \\ 0 & 2 & -1 & 1 \\ -2 & 1 & -3 & 2 \end{pmatrix} \vec{x} = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 3 \end{pmatrix}, \text{ dont la solution vaut exactement: } \vec{x} = \begin{pmatrix} -1 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

On part de $\vec{x}_0 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, par exemple. On itère la relation (V-1) et on applique l'algorithme

ε aux quatre composantes de la suite vectorielle qui en résulte. La session Mathematica qui suit vérifie, que l'on trouve effectivement la solution cherchée, soit:

$$\varepsilon_8^{(n)} = \vec{x} = \begin{pmatrix} -1 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

**H={{1,0,2,-1},{3,2,-1,1},{0,2,-1,1},{-2,1,-3,2}};b={0,0,3,3};
x[k_]:=x[k]=H.x[k-1]+x[k-1]-b;x[0]={1,2,3,4};Table[x[n],{n,0,10}]**

**{{ {1,2,3,4},{4,10,5,0},{18,37,17,-16},{86,132,55,-101},{383,498,160,511},
{1597,1972,482,-2284},{6442,7941,1657,-9523},{25721,31969,6356,-38486},
{102640,128228,25449,-154002},{410180,513153,102451,-615408},
{1640670,2052140,410895,-2460787} }}**

**eps[k_,n_]:=eps[k,n]=eps[k-2,n+1]+1/(eps[k-1,n+1]-eps[k-1,n])
eps[-1,n_]:=0;
eps[0,n_]:=x[n];
Table[eps[8,n],{n,4}]**

{{{-1,1,2,3},{-1,1,2,3},{-1,1,2,3},{-1,1,2,3} }}

Certes, l'exemple numérique choisi est trop simple pour mériter la mise en oeuvre d'un algorithme aussi évolué. En pratique, cette méthode est utile dans les cas où la matrice H est creuse et de grande dimension. Le fait que presque tous ses éléments soient nuls, dans ce cas, se prête, en effet, bien au calcul récursif des éléments de la suite $\vec{\xi}_i$.

V-2. Détermination des racines d'un polynôme.

L'algorithme ε permet également de trouver les racines d'un polynôme. Un exemple fera mieux comprendre le principe de la méthode.

Soit à calculer les racines $\{-1, 2, 3, -4\}$ du polynôme suivant:

$$P(z) = z^4 - 15z^2 + 10z + 24$$

On suppose, dans un premier temps, qu'elles sont simples et de modules distincts. Si certaines racines sont multiples, il est toujours possible de remplacer P par le pgcd de P et de sa dérivée P' . Si certaines racines restent comodulaires (cas particulier des racines complexes conjuguées), on peut envisager de translater l'origine dans le plan complexe. En résumé, on suppose, dans un premier temps, que les racines peuvent être rangées dans l'ordre des modules décroissants:

$$|z_1| > |z_2| > |z_3| > |z_4|$$

On considère, à présent, la récurrence homogène dont les coefficients, constants, coïncident avec ceux du polynôme de départ:

$$C_{k+1} - 15C_{k-1} + 10C_{k-2} + 24C_{k-3} = 0$$

On l'initialise de façon quelconque, par exemple: $C_0=1, C_{-1}=C_{-2}=C_{-3}=0$.

Calculant la suite C_k de proche en proche, on engendre, sauf conditions initiales exceptionnelles, une solution dominante, donc stable, du type:

$$C_k = \lambda_1 z_1^k + \lambda_2 z_2^k + \lambda_3 z_3^k + \lambda_4 z_4^k \quad (V-2)$$

Si on forme les quotients successifs:

$$\frac{C_{k+1}}{C_k} = \frac{\lambda_1 z_1^{k+1} + \lambda_2 z_2^{k+1} + \lambda_3 z_3^{k+1} + \lambda_4 z_4^{k+1}}{\lambda_1 z_1^k + \lambda_2 z_2^k + \lambda_3 z_3^k + \lambda_4 z_4^k}$$

ils tendent vers z_1 lorsque $k \rightarrow \infty$ du fait que $|z_1|$ est maximum en module. C'est la méthode de Bernoulli, valable uniquement pour cette racine. On trouve les autres racines en appliquant l'algorithme ε à la suite C_k .

On forme, dans ce but, la table (V-1).

On est certain de trouver $\varepsilon_8^{(n)} = 0, \forall n$, car la suite C_k fait partie du noyau de ε_8 et vu la relation (V-2), elle est sommée exactement à la valeur limite 0. En principe, la connaissance de C_0, \dots, C_7 (par exemple) ne permet de construire que l'angle supérieur gauche de la table (V-1), à cause de la structure de la règle récurrente qui construit la table ε , vue au §III-5 dans une notation légèrement différente:

$$\frac{1}{C-E} - \frac{1}{E-G} = \frac{1}{H-E} - \frac{1}{E-B}$$

où les symboles littéraux réfèrent à la position relative des éléments dans la table ϵ :

A B C
D E F
G H I

$\epsilon_0^{(0)} = C_0$	$\epsilon_2^{(0)}$	$\epsilon_4^{(0)}$	$\epsilon_6^{(0)}$	$\epsilon_8^{(0)} = 0$
$\epsilon_0^{(1)} = C_1$	$\epsilon_2^{(1)}$	$\epsilon_4^{(1)}$	$\epsilon_6^{(1)}$	$\epsilon_8^{(1)} = 0$
$\epsilon_0^{(2)} = C_2$	$\epsilon_2^{(2)}$	$\epsilon_4^{(2)}$	*	$\epsilon_8^{(2)} = 0$
$\epsilon_0^{(3)} = C_3$	$\epsilon_2^{(3)}$	$\epsilon_4^{(3)}$	*	$\epsilon_8^{(3)} = 0$
$\epsilon_0^{(4)} = C_4$	$\epsilon_2^{(4)}$	*	*	$\epsilon_8^{(4)} = 0$
$\epsilon_0^{(5)} = C_5$	$\epsilon_2^{(5)}$	*	*	$\epsilon_8^{(5)} = 0$
$\epsilon_0^{(6)} = C_6$	*	*	*	$\epsilon_8^{(6)} = 0$
$\epsilon_0^{(7)} = C_7$	*	*	*	$\epsilon_8^{(7)} = 0$

Table V-1

Toutefois, le fait de savoir que $\epsilon_8^{(n)} = 0, \forall n$, permet précisément de compléter la table dans son angle inférieur droit, simplement en utilisant la même récurrence, prise dans l'ordre inverse.

On sait déjà que:

$$\lim_{n \rightarrow \infty} \frac{\epsilon_0^{(n+1)}}{\epsilon_0^{(n)}} = z_1$$

On peut montrer que l'on a de la même manière:

$$\lim_{n \rightarrow \infty} \frac{\epsilon_2^{(n+1)}}{\epsilon_2^{(n)}} = z_2 \quad \lim_{n \rightarrow \infty} \frac{\epsilon_4^{(n+1)}}{\epsilon_4^{(n)}} = z_3 \quad \lim_{n \rightarrow \infty} \frac{\epsilon_6^{(n+1)}}{\epsilon_6^{(n)}} = z_4$$

Bien que nous passions la démonstration sous silence, indiquons, pour fixer les idées ce qui se passe si on considère la suite:

$$S_n = \epsilon_0^{(n)} = u^n + v^n + w^n$$

On trouve successivement:

$$\mathcal{E}_2^{(n)} = \frac{u^n v^n (u-v)^2 + u^n w^n (u-w)^2 + v^n w^n (v-w)^2}{u^n (u-I)^2 + v^n (v-I)^2 + w^n (w-I)^2}$$

$$\mathcal{E}_4^{(n)} = \frac{u^n v^n w^n (u-v)^2 (u-w)^2 (v-w)^2}{u^n v^n (u-v)^2 (u-I)^2 (v-I)^2 + u^n w^n (u-w)^2 (u-I)^2 (w-I)^2 + v^n w^n (v-w)^2 (v-I)^2 (w-I)^2}$$

$$\mathcal{E}_6^{(n)} = 0$$

Si $|u| > |v| > |w|$, on voit facilement que:

$$\mathcal{E}_0^{(n+1)} / \mathcal{E}_0^{(n)} = u \left[I + O\left((v/u)^n \right) \right]$$

$$\mathcal{E}_2^{(n+1)} / \mathcal{E}_2^{(n)} = v \left[I + O\left(\text{Worst}\left((w/v)^n, (v/u)^n \right) \right) \right]$$

$$\mathcal{E}_4^{(n+1)} / \mathcal{E}_4^{(n)} = w \left[I + O\left((w/v)^n \right) \right]$$

La vitesse de la convergence dépend donc des contrastes existant entre les modules des racines. Lorsqu'il y a quatre racines, on trouve similairement que:

$$\mathcal{E}_0^{(n+1)} / \mathcal{E}_0^{(n)} \rightarrow z_1 \quad \text{vitesse} = |z_2 / z_1|^n$$

$$\mathcal{E}_2^{(n+1)} / \mathcal{E}_2^{(n)} \rightarrow z_2 \quad \text{vitesse} = \text{Worst}\left(|z_2 / z_1|^n, |z_3 / z_2|^n \right)$$

$$\mathcal{E}_4^{(n+1)} / \mathcal{E}_4^{(n)} \rightarrow z_3 \quad \text{vitesse} = \text{Worst}\left(|z_3 / z_2|^n, |z_4 / z_3|^n \right)$$

$$\mathcal{E}_6^{(n+1)} / \mathcal{E}_6^{(n)} \rightarrow z_4 \quad \text{vitesse} = |z_4 / z_3|^n$$

La règle est générale. On peut encore accélérer le processus en réappliquant l'algorithme ε aux suites $\mathcal{E}_{2k}^{(n+1)} / \mathcal{E}_{2k}^{(n)}$ ainsi générées. C'est évident, vu leur comportement asymptotique de type exponentiel. Par exemple, revenant au polynôme: $P(z) = z^4 - 15z^2 + 10z + 24$, on trouve que les racines sont données par les limites suivantes:

$$\mathcal{E}_0^{(n+1)} / \mathcal{E}_0^{(n)} \rightarrow -4 \quad \text{vitesse} = (3/4)^n$$

$$\mathcal{E}_2^{(n+1)} / \mathcal{E}_2^{(n)} \rightarrow 3 \quad \text{vitesse} = (3/4)^n$$

$$\mathcal{E}_4^{(n+1)} / \mathcal{E}_4^{(n)} \rightarrow 2 \quad \text{vitesse} = (2/3)^n$$

$$\mathcal{E}_6^{(n+1)} / \mathcal{E}_6^{(n)} \rightarrow -1 \quad \text{vitesse} = (1/2)^n$$

La session Mathematica suivante le montre.

```
co[n_]:=co[n]=15 co[n-2]-10 co[n-3]-24 co[n-4]
co[0]=1;co[-1]=co[-2]=co[-3]=0;
eps[-1,n_]:=0;
eps[0,n_]:=co[n];
eps[k_,n_]:=eps[k,n]=eps[k-2,n+1]-1/(eps[k-1,n+1]-eps[k-1,n])
Print[Table[N[eps[2 m,10 k+1]/eps[2 m,10 k],10],{m,0,3},{k,9}]]
```

$\epsilon_0^{(10k+1)}/\epsilon_0^{(10k)}$	$\epsilon_2^{(10k+1)}/\epsilon_2^{(10k)}$	$\epsilon_4^{(10k+1)}/\epsilon_4^{(10k)}$	$\epsilon_6^{(10k+1)}/\epsilon_6^{(10k)}$
-3.328037074	3.096979706	1.998288264	-1.001464483
-3.958114702	3.005162691	1.99998231	-1.00000143
-3.997627621	3.00028656	1.999999705	-1.000000001
-3.99986636	3.000016068	1.999999995	-1.
-3.999992474	3.000000904	2.	-1.

Table V-2

V-2.1. Cas de racines comodulaires.

En fait, les rapports $\epsilon_{2k}^{(n+1)}/\epsilon_{2k}^{(n)}$ peuvent être calculés de façon tout à fait équivalente comme racines de l'équation déterminante du premier degré en u:

$$\begin{vmatrix} 1 & c_n \\ u & c_{n+1} \end{vmatrix} = 0$$

où les $c_n = \epsilon_{2k}^{(n)}$

Dans le cas où deux racines sont comodulaires, il apparaît que ces quotients ne convergent pas. Lorsque cela se produit, il y a lieu de remplacer le calcul de ces rapports par celui des racines de l'équation du deuxième degré en u:

$$\begin{vmatrix} 1 & c_n & c_{n-1} \\ u & c_{n+1} & c_n \\ u^2 & c_{n+2} & c_{n+1} \end{vmatrix} = 0$$

ce qui fournit, automatiquement, les deux racines comodulaires.

Lorsqu'il apparaît que ce processus ne converge pas davantage, c'est que trois racines sont sans doute comodulaires et il faut passer au calcul des racines de l'équation du troisième degré en u:

$$\begin{vmatrix} 1 & c_n & c_{n-1} & c_{n-2} \\ u & c_{n+1} & c_n & c_{n-1} \\ u^2 & c_{n+2} & c_{n+1} & c_n \\ u^3 & c_{n+3} & c_{n+2} & c_{n+1} \end{vmatrix} = 0 \quad \text{etc...}$$

V-2.2. Calcul des valeurs propres d'une matrice par l'algorithme ε.

La même méthode permet le calcul des valeurs propres d'une matrice. On part de l'identité bien connue suivante:

$$\text{tr}(A^k) = \sum_{i=1}^n \lambda_i^k$$

On construit la suite des itérés: $S_k = n, \text{tr}(A), \text{tr}(A^2), \text{tr}(A^3), \dots$ et on lui applique l'algorithme ε. Comme elle appartient, de toute évidence, à son noyau, on finit par trouver, au plus tard à l'ordre $2n$, une colonne de zéros. Les valeurs propres, λ_i , se déduisent en travaillant sur les colonnes de la table ε, comme expliqué au paragraphe précédent. Cette stratégie a été améliorée par Brezinski, dans le but d'éviter les produits de matrices. Nous passons cette variante sous silence.

V-3. Accélération des séries de Fourier.

Il ressort des résultats des expériences numériques consignées dans la table (IV-4) que les séries de Fourier constituent une classe importante de suites difficiles à accélérer. Seul l'algorithme ε y parvient avec un succès d'ailleurs mitigé. Dans le cas fréquent où les coefficients de la série de Fourier sont rationnels, on peut écrire son échelle asymptotique sous la forme suivante (γ fixé quelconque, souvent on prend $\gamma=1$):

$$g_{2i}(n) = \frac{\cos nx}{(n+\gamma)^i} \quad \text{et} \quad g_{2i-1}(n) = \frac{\sin nx}{(n+\gamma)^i} \quad (i=1,2,\dots)$$

Il revient au même de les réécrire sous cette autre forme, équivalente, qui simplifie les calculs ultérieurs de déterminants:

$$g_{2i}(n) = \frac{e^{jnx}}{(n+\gamma)^i} \quad \text{et} \quad g_{2i-1}(n) = \frac{e^{-jnx}}{(n+\gamma)^i}$$

On applique l'algorithme E et on trouve ($q = e^{jx}$):

$$E_k^{(n)} = \frac{\text{dtm}(S_n, g_1(n), g_2(n), \dots, g_k(n))}{\text{dtm}(1, g_1(n), g_2(n), \dots, g_k(n))}$$

qui se réécrit, du fait du dédoublement des asymptotes:

$$E_{2k}^{(n)} = \frac{\text{dtm}((n+\gamma)^k S_n, q^n (n+\gamma)^{k-1}, q^{-n} (n+\gamma)^{k-1}, q^n (n+\gamma)^{k-2}, q^{-n} (n+\gamma)^{k-2}, \dots, q^n, q^{-n})}{\text{dtm}((n+\gamma)^k, q^n (n+\gamma)^{k-1}, q^{-n} (n+\gamma)^{k-1}, q^n (n+\gamma)^{k-2}, q^{-n} (n+\gamma)^{k-2}, \dots, q^n, q^{-n})}$$

Calculons, d'abord, le numérateur; le dénominateur suivra en posant tous les S_n égaux à un. Remplaçons, dans un premier temps, $(n+\gamma)^k S_n$ par u^0 , et plus généralement remplaçons $(n+\gamma+\ell)^k S_{n+\ell}$ par u^ℓ . Le numérateur se réduit à un polynôme en u facile à évaluer. Il suffira, ensuite, de faire le remplacement inverse $u^\ell \Rightarrow (n+\gamma+\ell)^k S_{n+\ell}$ pour

obtenir la valeur du déterminant cherché. La dépendance en u du polynôme représenté par le numérateur est immédiate du fait qu'il s'annule, ainsi que toutes ses dérivées jusqu'à l'ordre $(k-1)$ inclus, en $u=q$ et en $u=q^{-1}$. Le numérateur est donc proportionnel à :

$$\text{numérateur} = [u^2 + 1 - u(q + q^{-1})]^k = [u^2 + 1 - 2u \cos x]^k$$

Il n'est pas nécessaire de connaître le coefficient de proportionnalité du fait qu'il est également présent au dénominateur.

Il est connu que $(u^2 + 1 - 2u \cos x)^k$ est en fait la génératrice des polynômes de Gegenbauer:

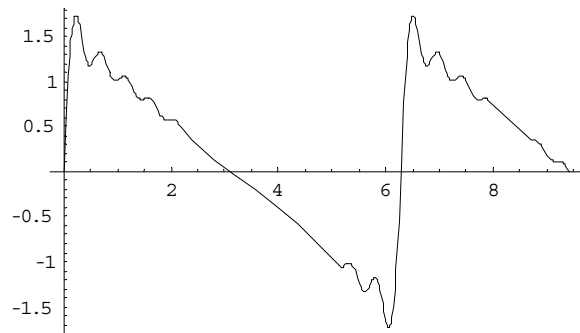
$$(u^2 + 1 - 2u \cos x)^k = \sum_{\ell=0}^{2k} C_{2k}^{-k}(\cos x) u^\ell = \sum_{\ell=0}^{2k} (-1)^\ell C_{2k}^\ell {}_2F_1(-\ell, \ell - 2k; \frac{1}{2} - k; \frac{1 - \cos x}{2}) u^\ell$$

d'où on déduit la formule sommatoire qui convient aux séries de Fourier:

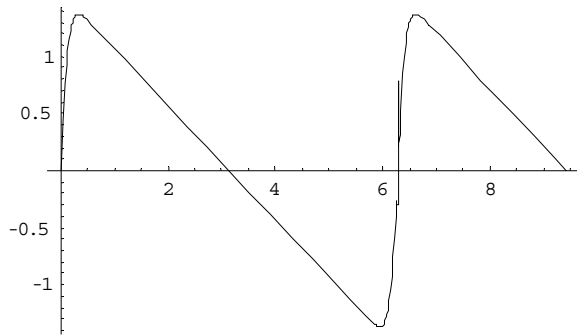
$$E_{2k}^{(n)} = \frac{\sum_{m=0}^{2k} (-1)^m C_{2k}^m {}_2F_1(-m, m - 2k; \frac{1}{2} - k; \frac{1 - \cos x}{2}) (n + \gamma + m)^k S_{n+m}}{\sum_{m=0}^{2k} (-1)^m C_{2k}^m {}_2F_1(-m, m - 2k; \frac{1}{2} - k; \frac{1 - \cos x}{2}) (n + \gamma + m)^k}$$

La session Mathematica, qui suit, compare les graphes de la série de Fourier d'une fonction en dents de scie, tronquée après 13 termes, à la meilleure suite, $E_{12}^{(0)}$, que l'algorithme E est capable d'en extraire.

`Plot[N[Sum[Sin[kx]/k,{k,1,13}]],{x,0,3 Pi}]`



`Plot[Evaluate[N[Sum[(-1)^m Binomial[12,m] HypergeometricPFQ[{-m,m-12},{-11/2},(1-Cos[x])/2] (m+1)^6 Sum[Sin[kx]/k,{k,1,m+1}],{m,0,12}]/Sum[(-1)^m Binomial[12,m] HypergeometricPFQ[{-m,m-12},{-11/2},(1Cos[x])/2] (m+1)^6,{m,0,12}]],{x,0,3 Pi}]]`



On note l'amélioration, très sensible, de la convergence. Seuls subsistent, fort atténués, aux bornes de l'intervalle de périodicité de la fonction représentée, les problèmes inhérents au phénomène de Gibbs.

V-4. Calcul des valeurs propres de l'équation de Schrödinger.

Les fcg, étudiées au §II-4, permettent un calcul numérique élégant des valeurs propres de l'équation différentielle de Schrödinger. On considère, à titre d'exemple, le cas d'un électron voyageant dans un potentiel anharmonique à une dimension:

$$V(x) = x^4$$

L'équation à résoudre est celle de Schrödinger :

$$\psi'' + (E - x^4)\psi = 0 \quad (\text{V-3})$$

où la fonction ψ doit être de carré intégrable entre $-\infty$ et $+\infty$. Cette équation est d'ordre 2; elle possède donc deux solutions indépendantes qui, en $+\infty$ et en $-\infty$, se comportent comme

$$\psi_{div} \approx |x|^{-1} \exp(|x|^3 / 3)$$

$$\psi_{conv} \approx |x|^{-1} \exp(-|x|^3 / 3)$$

Si E est quelconque, la solution qui converge en $x = +\infty$ diverge généralement en $x = -\infty$ et inversement. Au total, aucune solution n'est de carré intégrable. Toutefois, pour certaines valeurs de E appelées valeurs propres, une même fonction ψ converge à la fois en $x = +\infty$ et en $x = -\infty$. Ce sont ces valeurs de E qui nous intéressent ainsi que les fonctions propres correspondantes.

Nous cherchons ces fonctions propres sous l'une ou l'autre des formes suivantes.

V-4.1. Développement en série de Taylor.

$$\psi = \exp(-\omega x^2) \sum_0^{\infty} C_k x^{2k} \quad (V-4)$$

Cette fonction ne décrit que les états pairs. Une fonction similaire où x^{2k} serait remplacée par x^{2k+1} décrirait les états impairs. Nous nous limitons à la détermination des états pairs. Le rôle joué par le paramètre muet ω sera précisé ultérieurement.

Introduisant le développement (V-4) dans l'équation (V-3) et identifiant à zéro le coefficient de x^{2k} , pour tout k , on trouve que les C_k obéissent à la récurrence d'ordre trois :

$$(2k+1)(2k+2)C_{k+1} + (E - 2\omega - 8\omega k)C_k + 4\omega^2 C_{k-1} - C_{k-2} = 0 \quad (V-5)$$

($k = 0, 1, 2, \dots$).

Les conditions initiales sont, à un facteur de normalisation près : $C_{-2} = 0$, $C_{-1} = 0$ et $C_0 = 1$. La récurrence (V-5) étant d'ordre trois, elle possède trois solutions linéairement indépendantes notées : $C_k^{(1)}$, $C_k^{(2)}$ et $C_k^{(3)}$.

On déterminera plus loin leur asymptote et on verra qu'elles sont partiellement contrastées, soit: $|C_k^{(1)}| > |C_k^{(2)}| \approx |C_k^{(3)}|$.

La solution générale s'écrit sous la forme :

$$C_k = \lambda C_k^{(1)} + \mu C_k^{(2)} + \nu C_k^{(3)}$$

où λ , μ et ν sont des coefficients qui dépendent de E . Quand on impose les conditions initiales $C_{-2} = 0$, $C_{-1} = 0$ et $C_0 = 1$, on engendre généralement une solution de type dominant car $\lambda(E) \neq 0$. La fonction ψ correspondante est de type divergent car ce n'est pas en introduisant des C_k dominants dans le développement (V-4) que l'on va construire une fonction ψ de type convergent. Toutefois, pour certaines valeurs de E , les valeurs propres cherchées, $\lambda(E) = 0$ et les C_k deviennent sous-dominants. Il nous faut donc construire la solution sous-dominante, c'est-à-dire intermédiaire, compatible avec les conditions initiales imposées. C'est précisément ce que l'algorithme de Miller généralisé permet de faire.

En notation matricielle, il s'écrit :

$$\begin{pmatrix} E-2\omega & 2 & 0 & 0 & 0 & 0 \\ 4\omega^2 & E-10\omega & 12 & 0 & 0 & 0 \\ -1 & 4\omega^2 & E-18\omega & 30 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 2K(2K-1) \\ 0 & 0 & 0 & -1 & 4\omega^2 & E-2\omega-8\omega K \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ \vdots \\ \vdots \\ C_K \end{pmatrix} = - \begin{pmatrix} 4\omega^2 C_{-1} - C_{-2} \\ -C_{-1} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

Or, le second membre est nul vu les conditions initiales. Le système est donc, en réalité, homogène et l'annulation de son déterminant fournit les valeurs permises pour E .

Théoriquement, il y aurait lieu d'annuler un déterminant infini qui, dans la littérature, porte le nom de déterminant de Hill, du nom de celui qui l'a utilisé pour la première fois, dans un contexte, il est vrai, fort différent. Toutefois, en pratique, on se contente de le tronquer à une valeur de K suffisamment grande, cette valeur étant à déterminer en fonction de la précision requise. Le calcul des asymptotes des solutions $C_k^{(i)}$ est indispensable pour atteindre cet objectif. En effet, nous avons vu que la précision fournie par l'algorithme de Miller généralisé dépendait du facteur de contraste $|C_k^{(2)}/C_k^{(1)}|$.

Passons à la détermination des asymptotes de la récurrence (V-5). Elle est irrégulière. Son équation caractéristique s'écrit :

$$4k^2 z^3 - 8\omega k z^2 + 4\omega^2 z - 1 = 0$$

Recherchons ses trois racines sous la forme : $z = \rho k^\tau$. On trouve :

$$4\rho^3 k^{3\tau+2} - 8\omega\rho^2 k^{2\tau+1} + 4\omega^2 \rho k^\tau - 1 = 0$$

- $3\tau + 2 = 2\tau + 1$ d'où $\tau = -1$ est impossible
- $3\tau + 2 = \tau$ d'où $\tau = -1$ est impossible
- $3\tau + 2 = 0$ d'où $\tau = -2/3$ est possible à condition de poser : $4\rho^3 = 1$ d'où :

$$\rho = \sqrt[3]{1/4} \exp(2i\pi\ell/3) \quad (\ell = 0, 1, 2) \quad \text{et} \quad \tau = -2/3.$$

Il convient donc de poser :

$$C_k^{(\ell)} = \rho^k [\Gamma(k)]^\tau \gamma_k^{(\ell)} = (1/4)^{k/3} \exp(2i\pi k\ell/3) [\Gamma(k)]^{-2/3} \gamma_k^{(\ell)}$$

que l'on introduit dans la récurrence de base (V-5) afin d'obtenir la récurrence satisfaite par $\gamma_k^{(\ell)}$.

On trouve après réductions :

$$\begin{aligned} & (2k+1)(2k+2)(1/4)^{(k+1)/3} \exp[2i\pi\ell(k+1)/3] \gamma_{k+1} + (E-2\omega-8\omega k)(1/4)^{k/3} k^{2/3} \exp(2i\pi\ell k/3) \gamma_k \\ & + 4\omega^2 (1/4)^{(k-1)/3} \exp[2i\pi\ell(k-1)/3] k^{2/3} (k-1)^{2/3} \gamma_{k-1} \\ & - (1/4)^{(k-2)/3} \exp[2i\pi\ell(k-2)/3] k^{2/3} (k-1)^{2/3} (k-2)^{2/3} \gamma_{k-2} = 0 \end{aligned}$$

Pour chacune des trois racines ($\ell = 0, 1, 2$), on a une récurrence de Poincaré type trois vu les développements en puissances de $k^{-1/3}$.

Les tables (I-4) d'asymptotes nous apprennent que l'on a dans ce cas :

$$\gamma_k \approx k^w \exp[\alpha k^{2/3} + \beta k^{1/3}]$$

Tous calculs faits, on trouve les asymptotes suivantes:

$$C_k^{(\ell)} \approx \frac{\left[\exp(2i\pi\ell/3)(1/9)^{1/3} \right]^k}{\Gamma(2k/3)} k^{-4/3} \cdot \exp \left[\omega \exp(4i\pi\ell/3)(2k)^{2/3} - \frac{2}{3} \omega^2 \exp(2i\pi\ell/3)(2k)^{1/3} \right]$$

On voit facilement que $C_k^{(\ell=0)}$ est dominante tandis que $C_k^{(\ell=1)} \approx C_k^{(\ell=2)}$ sont dominées.

Le facteur de contraste vaut :

$$\rho_k \approx \left| C_k^{(\ell=1)} / C_k^{(\ell=0)} \right| \approx \exp \left[-\frac{3}{2} \omega (2k)^{2/3} + \omega^2 (2k)^{1/3} \right]$$

Ce contraste est maximum si on choisit la valeur du paramètre muet ω en sorte que : $\partial \rho / \partial \omega = 0$ soit :

$$\omega_{opt} \approx \frac{3}{4} (2k)^{1/3}$$

Il est alors facile de voir que la précision obtenue dans le calcul des valeurs propres comme racines du déterminant $K \times K$ est de l'ordre de $\rho_K \sim e^{-P}$.

Concrètement, on évite le calcul des déterminants en remarquant qu'ils valent exactement le numérateur des approximants de la fcg d'ordre zéro, associée à la récurrence (V-5) du problème. Les valeurs propres cherchées annulent donc aussi cette fcg. Or la composante d'ordre $s=0$ de la fcg se calcule très facilement par voie récursive. En application de l'identité (II-13), les racines cherchées sont, tout simplement, identiques à la limite des racines de la suite auxiliaire $N_k^{(1)} = C_k$.

V-4.2. Développement en série de Weber-Hermite.

La variante suivante fait appel à des fcg d'ordre $s=1$. Elle consiste à développer la fonction d'onde en série des fonctions d'onde, D_{2k} , de l'oscillateur harmonique, dites de Weber-Hermite (états pairs seulement):

$$\Psi = \sum_{k=0}^{\infty} C_k D_{2k}(\omega x) / k!$$

Sans entrer dans les détails, on trouve que les C_k obéissent, cette fois, à une récurrence d'ordre quatre qui s'écrit in extenso ($k=1, 2, \dots$) :

$$\begin{aligned} & (16k^2 - 4)C_{k+1} + [32k^2 - (24 + \omega^6)k + (4 + \omega^6/2)]C_k \\ & + [24k^2 + (\omega^6 - 36)k + (15 - 3\omega^6/4 - E\omega^4)]C_{k-1} \\ & + [8k^2 - (18 + \omega^6/4)k + (10 + \omega^6/4)]C_{k-2} + (k-1)(k-2)C_{k-3} = 0 \end{aligned} \quad (V-7)$$

Cette récurrence possède quatre solutions linéairement indépendantes, respectivement asymptotiques à ($\ell = 0, 1, 2, 3$):

$$C_k^{(\ell)} \approx (-1/2)^k k^{-3/8} \exp \left[\frac{2}{3} \omega^{3/2} e^{i\pi(2\ell+1)/4} (2k)^{3/4} - \frac{\omega^{9/2}}{24} e^{-i\pi(2\ell+1)/4} (2k)^{1/4} \right]$$

On constate que les solutions numérotées $C_k^{(\ell=0)}$ et $C_k^{(\ell=3)}$ dominent $C_k^{(\ell=2)}$ et $C_k^{(\ell=1)}$.

Si on construit la fonction d'onde sur base de coefficients C_k dominants, on n'obtiendra sûrement pas le bon comportement asymptotique pour Ψ (cfr § V-4.1). En fait, il faut utiliser des C_k dominés qui respectent les conditions initiales évidentes: $C_{-2} = 0$, $C_{-1} = 0$ et $C_0 = 1$. L'algorithme de Miller généralisé permet de les calculer à partir d'un système linéaire analogue à celui du § V-4.1.

Le déterminant de ce système coïncide, cette fois, avec le numérateur de la fcg d'ordre $s=1$, associée à la récurrence (V-7). La convergence de cette fcg est assurée du fait qu'il existe un contraste d'amplitude entre les solutions de la récurrence notées $C_k^{(\ell=0)}$ et $C_k^{(\ell=1)}$. Ce contraste vaut:

$$\rho_k = \exp \left[-\frac{2}{3} \sqrt{2} \omega^{3/2} (2k)^{3/4} + \frac{\sqrt{2}}{24} \omega^{9/2} (2k)^{1/4} \right]$$

En résumé, les valeurs propres, E , se calculent comme racines des déterminants successifs d'ordre K de la matrice de Miller, ou ce qui revient au même comme racine des approximants de la fcg d'ordre un associée à la récurrence (V-7), la vitesse de la convergence étant donnée par la relation: $\rho_K \sim e^{-p}$. On voit facilement qu'elle est maximum lorsqu'on choisit le paramètre muet, ω , à la valeur optimale:

$$\omega_{opt} = (48p / \sqrt{6})^{1/6}.$$

En pratique, on évite, le calcul des déterminants en recourant, à nouveau, au calcul récursif des approximants d'ordre un de la fcg associée, selon la procédure vue au §II-4. Vu que $s=1$, on doit cependant calculer deux suites auxiliaires, $N_k^{(1)}$ et $N_k^{(2)}$, et les combiner conformément à l'identité (II-13).