



HAL
open science

Polynômes orthogonaux simultanés et systèmes dynamiques infinis

Emmanuel Bourreau

► **To cite this version:**

Emmanuel Bourreau. Polynômes orthogonaux simultanés et systèmes dynamiques infinis. Mathématiques [math]. Université des Sciences et Technologie de Lille - Lille I, 2002. Français. NNT : . tel-00005227

HAL Id: tel-00005227

<https://theses.hal.science/tel-00005227>

Submitted on 9 Mar 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre : 3117

Thèse présentée à
L'UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE
LILLE

pour obtenir
LE TITRE DE DOCTEUR DE L'UNIVERSITÉ
SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES

par
Emmanuel Bourreau

Polynômes orthogonaux simultanés et
systèmes dynamiques infinis

Soutenue le 10 mai 2002, devant la Commission d'Examen :

Présidente : J. VAN ISEGHEM, Université de Lille I
Directeur de thèse : B. BECKERMANN, Université de Lille I
Rapporteurs : A. DRAUX, INSA de Rouen
W. VAN ASSCHE, Université Catholique de Louvain
Examineurs : C. BREZINSKI, Université de Lille I
V. KALIAGUINE, Université de Nizhny Novgorod, Russie
A. APTEKAREV, Institut Keldysh, Moscou, Russie

Remerciements

Alors qu'une étape s'achève, il est temps d'évaluer le chemin parcouru et remercier les différentes personnes qui m'ont accompagné.

Je voudrais tout d'abord remercier Bernhard Beckermann, mon directeur de thèse, pour sa patience à mon égard. J'effectue des recherches sous sa direction depuis mon mémoire de D.E.A. en 1996. Pendant ces six années, sa persévérance, parfois son obstination m'ont permis de ne jamais me décourager malgré mes activités d'enseignant du secondaire qui prennent une grande partie de mon temps.

Même si certaines coquilles typographiques doivent subsister, cette thèse n'aurait pu être améliorée sans la lecture attentive et les remarques de M. Draux et M. Van Assche qui ont accepté d'être mes deux rapporteurs. Qu'ils soient encore remerciés pour la célérité avec laquelle ils ont remis leur rapport, le temps qui leur était imparti était très court.

La présence de Mme Van Iseghem pour présidente et de M. Kaliaguine et M. Aptekarev est pour moi un honneur, en même temps qu'une source d'appréhension : les auteurs des articles sur lesquels j'ai travaillé, parfois avec acharnement, sont aujourd'hui mes juges. J'espère qu'ils seront indulgents.

Je suis heureux que M. Brezinski fasse partie de mon jury. C'est lui qui m'enseigna l'Analyse Numérique en Maîtrise puis pendant le D.E.A. Qu'il sache que je garde précieusement l'article sur Stieljes, paru dans la revue "pour la Science", qu'il m'a gentilement dédié.

J'ai pu aussi compter sur le soutien de mes collègues professeurs. Ils ont réduit ma charge de travail en rédigeant les sujets de D.S. communs ou en élaborant les barèmes.

A leur manière, mes parents, frère et soeurs m'ont encouragé par leurs plaisanteries sur la durée inhabituelle de ma thèse!

Je remercie aussi ma femme Sophie pour sa présence à mes côtés. Heureusement que nous n'avons pas attendu la fin de ma thèse pour avoir de beaux enfants! Pierre et Constance font notre joie à tous les deux. Que Pierre me pardonne de n'avoir pu jouer tous les jours avec lui. A deux ans et demi, il sait déjà que je vais à la "fac" pour voir "M. Betermann".

Table des matières

Table des matières	4
Introduction	6
1 Polynômes matriciellement orthogonaux	11
1.1 Introduction	11
1.2 Biorthogonalité	13
1.2.1 Forme bilinéaire	13
1.2.2 Algèbre linéaire	14
1.2.3 Récurrence à $r + s + 1$ termes	16
1.2.4 Dualité et Christoffel-Darboux	20
1.2.5 Approximants de Padé	25
1.2.6 Cas particuliers des approximants de Padé précédents	27
1.2.6.i Systèmes parfaits	27
1.2.6.ii Cas scalaire	29
1.3 Analyse fonctionnelle	29
1.3.1 Opérateur	29
1.3.2 Résolvante	30
1.3.3 Caractérisation de l'ensemble résolvant	33
1.4 Résultats pour les matrices constantes	34
2 Fractions continues	42
2.1 Formules générales	43
2.2 Vérifications des formules	47
2.2.1 Le cas scalaire	47
2.2.2 Fractions continues généralisées	47
2.2.3 Cas vectoriel (algorithme de Jacobi-Perron)	49
2.2.4 Fractions continues matricielles	51

	5
2.2.5	Fractions continues matricielles par blocs 54
2.3	Accélération de convergence 57
3	Moments modifiés 71
3.1	Introduction 71
3.2	Calcul des coefficients. Algorithme de Chebyshev modifié . . . 72
3.2.1	Introduction 72
3.2.2	Algorithme de Chebyshev modifié scalaire 72
3.2.3	Algorithme de Chebyshev modifié vectoriel 76
3.3	Résultats numériques 79
4	Systèmes dynamiques infinis 86
4.1	Résolutions théorique et pratique du système semi-infini de Toda 87
4.2	Estimations 94
4.3	Applications 95
4.3.1	Mise en oeuvre 96
4.3.2	Exemples 97
4.4	Démonstrations 105
	Conclusion 117
A	Codes Maple 119
A.1	Algorithme de Chebyshev modifié vectoriel 120
A.2	Systèmes dynamiques, algorithme de Chebyshev scalaire . . . 127
	Bibliographie 133

Introduction

A travers l'étude de certains problèmes d'analyse fonctionnelle (équations intégrales, séries de Fourier, problème de Sturm-Liouville et, plus généralement, problèmes aux limites dans les équations aux dérivées partielles) est apparue la notion de système orthogonal de fonctions. Ces problèmes amènent à considérer des espaces hermitiens constitués de fonctions et à déterminer les valeurs propres et les fonctions propres de certains endomorphismes de ces espaces. Dans le cas d'un opérateur hermitien, les sous-espaces propres sont orthogonaux deux à deux. Le problème essentiel consiste alors à chercher des bases hilbertiennes constituées de fonctions propres.

Les polynômes orthogonaux sont un cas simple de système orthogonal.

Historiquement, les polynômes orthogonaux scalaires apparaissent comme un outil pour les fractions continues et sont étudiés par Chebyshev et Stieltjes. Les polynômes orthogonaux apparaissent aussi en interpolation polynomiale, en développement en série de fonctions, en calcul approché ou exact d'intégrales et encore dans l'étude des approximants de Padé scalaires, vectoriels et matriciels et des fractions continues scalaires, vectorielles et matricielles. Ils permettent par exemple d'étudier la convergence de ces différentes quantités. Ils sont aussi un outil important dans la résolution des systèmes dynamiques de Toda-Langmuir et les systèmes associés de Kac-van Moerbeke et les équations de Korteweg-deVries.

Les polynômes orthogonaux formels permettent aussi de justifier et démontrer les algorithmes d'algèbre linéaire comme la méthode du Gradient conjugué, de Lanczos ou encore l'algorithme GMRES.

En lien avec les polynômes orthogonaux, les travaux présentés dans cette thèse suivent deux axes de recherche.

D'un côté, nous nous sommes intéressés aux fractions continues. A partir du cas scalaire, à la fin du vingtième siècle, se sont développées diverses généralisations comme le cas vectoriel donné par Parusnikov [60], les fractions

continues généralisées étudiées par De Bruin, Jacobsen et Levrie [25, 50, 53], les fractions de Thiele et les n -fractions par Levrie et Bultheel [51, 52], les fractions continues matricielles de matrices carrées étudiées par Aptekarev et Nikishin[3] et dernièrement les fractions continues pour les matrices rectangulaires dans l'article de Van Iseghem et Sorokin [72]. Ces développements permettent d'effectuer une approximation rationnelle de vecteurs ou de matrices de fonctions ou de caractériser le spectre d'un opérateur. Dans le cas scalaire, la fraction continue peut être exprimée comme une composition d'homographies. Cette décomposition est généralisée dans certains articles [50].

Cependant, il n'existe pas d'unification de ces différents travaux. Le chapitre 2 traite ainsi du regroupement des différentes expressions scalaires, vectorielles ou matricielles des fractions continues sous l'aspect d'une décomposition d'homographies définies dans le cas matriciel général de matrices rectangulaires $r \times s$. On retrouve dans ce formalisme les divers cas évoqués plus tôt en posant $r = s = 1$ pour obtenir le cas scalaire ou $r = 1, s > 1$ pour le cas vectoriel.

Cette définition d'homographie matricielle permet dans la deuxième partie de ce chapitre de démontrer un résultat souvent recherché en analyse numérique (voir par exemple l'historique présenté par C. Brezinski [22]) et toujours d'actualité (voir [21]) : l'accélération de convergence.

Le deuxième axe de recherche est la résolution du problème spectral inverse de manière numériquement stable : connaissant les moments associés à une mesure, est-il possible de déterminer l'opérateur associé ? C'est-à-dire, si on considère les polynômes orthogonaux relativement à cette mesure, est-il possible de déterminer la matrice infinie regroupant les coefficients de récurrence de ces polynômes orthogonaux et ceci de manière numériquement stable ?

Dans le cas scalaire, la matrice en question est une matrice à trois diagonales dite de Jacobi. Grâce aux divers travaux de Gautschi [32, 33, 34, 35, 36], on sait que le calcul des coefficients n'est généralement pas stable numériquement comme il l'indique dans de nombreux exemples [34, Exemples 4.1, 4.3, 4.8]. Ce calcul n'est pas stable avec l'algorithme de Chebyshev qui utilise les moments ordinaires $\int x^k d\sigma$. Et même si on travaille avec l'algorithme de Chebyshev modifié qui utilise les moments modifiés $\int p_k(x) d\sigma(x)$ où (p_k) est une famille de polynômes relativement à une mesure $d\sigma$, le calcul peut rester numériquement instable. Dans l'article écrit avec B. Beckermann à ce sujet [9], nous avons montré que pour envisager une stabilité numérique, il était nécessaire d'utiliser une mesure $d\sigma$ dont le support était essentiel-

lement semblable à celui de $d\sigma$. Ce résultat a pu être établi grâce aux estimations données par Fischer [30] sur le conditionnement de l'application non-linéaire qui associe les moments aux coefficients de récurrence.

De même que nous avons généralisé les fractions continues dans le cas matriciel, il nous a paru naturel de généraliser l'algorithme de Chebyshev modifié à la détermination de coefficients de récurrence de polynômes vectoriels. On peut noter qu'une telle généralisation existe dans le cas de polynômes matriciels ([69]) et aussi dans le cas vectoriel ([28]). Cependant, dans ce dernier cas, il n'est pas fait mention des polynômes de référence. *A contrario*, nous avons cherché, au travers d'exemples, des polynômes de référence adaptés aux polynômes vectoriels étudiés.

Malheureusement, dans le cas vectoriel, il n'existe pas d'estimations similaires à celles de Fischer dans le cas scalaire. Il nous est donc impossible pour l'instant de démontrer un résultat qui semble visible dans les exemples : les polynômes doivent posséder le même comportement asymptotique .

Dans un dernier temps, afin d'appliquer de manière intéressante, l'algorithme de Chebyshev modifié, nous nous sommes intéressés aux systèmes dynamiques et plus particulièrement au système de Toda semi-infini.

Il s'agit de déterminer l'évolution temporelle d'un système semi-infini de particules alignées sur le semi-axe réel et qui interagissent suivant une loi exponentielle décroissante en fonction de la distance les séparant. Ces particules peuvent posséder une vitesse initiale.

En effectuant un changement de variable, on peut décrire ce système par l'évolution temporelle des coefficients d'une matrice de Jacobi semi-infinie qui est solution d'une équation différentielle. Notre résultat sur l'algorithme de Chebyshev modifié prend alors tout son sens. En effet, la mesure au temps t , $ds(x, t)$ associée à la matrice est la mesure au temps initial $ds(x, 0)$ multipliée par e^{-tx} . Cela signifie que les polynômes orthogonaux possèdent le même comportement asymptotique et que l'algorithme de Chebyshev modifié devrait être numériquement stable.

Puisque l'on connaît l'évolution temporelle de la mesure, il est possible de déterminer les coefficients de récurrence des polynômes orthogonaux en décomposant la fonction de Weyl associée à cette mesure en fraction continue [58]. Cependant, on ne cherche que l'évolution des n premières particules c'est-à-dire des n premières lignes de la matrice semi-infinie.

Pour résoudre ce problème, nous avons donc décidé de tronquer la matrice initiale à N lignes ($N \gg n$) et de travailler avec un système fini. Il existe alors diverses méthodes pour déterminer l'évolution des coefficients.

Mais notre approche toute nouvelle nous permet d'utiliser l'algorithme de Chebyshev modifié, les moments modifiés étant déterminés par la résolution d'un système fini d'équations linéaires. De plus, il nous est surtout possible d'estimer l'erreur commise sur le calcul des coefficients de récurrence en utilisant à nouveau les estimations données par Fischer.

Le travail présenté se décompose donc en quatre chapitres.

Le premier présente l'aspect théorique nécessaire à la compréhension de la thèse. On y définit les polynômes vectoriels matriciellement orthogonaux et on démontre des résultats sur ces polynômes, similaires aux résultats classiques sur les polynômes scalaires (identité de Christoffel-Darboux, récurrence, convergence des approximants de Padé, théorème de Shohat-Favard). Le lecteur familiarisé avec ces objets et ces résultats peut s'intéresser immédiatement à la section 1.4 de ce chapitre, où l'on propose une nouvelle démonstration d'un résultat déjà donné par Duren [29] sur les opérateurs de Toeplitz bande.

Le chapitre 2 présente les transformations de Möbius ainsi que des homographies matricielles définies grâce au chapitre 1. Cette généralisation regroupe les différentes théories existantes comme les cas scalaire, vectoriel ou matriciel par blocs. Cette définition permet dans la deuxième partie de ce chapitre de démontrer un résultat d'accélération de convergence sur les fractions continues matricielles.

L'objet du chapitre 3 est d'étendre l'algorithme de Chebyshev modifié aux polynômes vectoriels. On rappelle l'algorithme dans le cas scalaire avec, en particulier, des résultats sur le choix des polynômes de référence [9]. Ensuite, grâce aux résultats du chapitre 1 au paragraphe 1.2.3, il est possible de décrire l'algorithme de Chebyshev vectoriel qui est alors appliqué à quelques exemples.

Enfin, le chapitre 4 présente une application de l'algorithme de Chebyshev modifié scalaire pour l'étude temporelle d'un système dynamique semi-infini issu d'un maillage de Toda-Langmuir. On étudie de manière théorique puis au travers d'exemples, l'estimation de l'erreur commise sur le calcul des n premières lignes de coefficients de récurrence d'une famille de polynômes $p_k(x, t)$ orthogonaux relativement à la mesure $e^{-tx} ds(x, 0)$. Les polynômes de référence de l'algorithme de Chebyshev modifié scalaire sont les polynômes $p_k(x, 0)$.

En utilisant l'aspect physique du problème, on pourra représenter l'évolution de ces particules. Il est intéressant de voir, grâce à cette représentation, l'ap-

parition de "motifs" c'est-à-dire de regroupement de deux ou trois particules qui se décalent vers la droite. Ces motifs ressemblent aux "solitons" du cas infini (voir par exemple le tutorial sur le système de Toda présenté par G. Teschl[76]).

On constate, aussi, de nouveau, que les estimations proposées sont parfois trop pessimistes et qu'il est possible de dépasser les limites données par la théorie sans trop s'éloigner de la vérité (que l'on peut vérifier par l'orthogonalité des polynômes en jeu).

Polynômes matriciellement orthogonaux

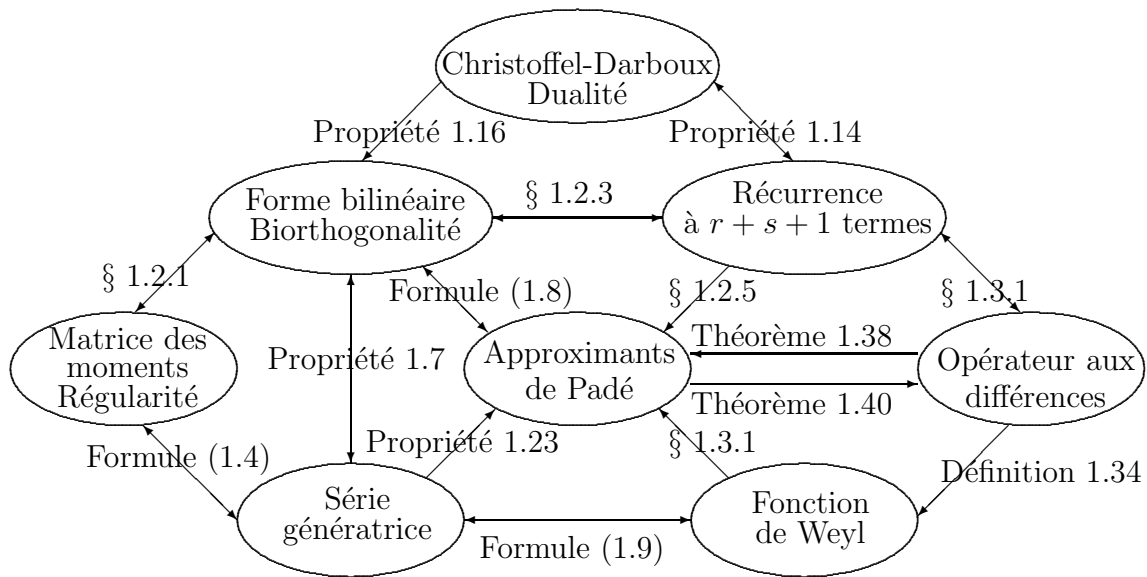
1.1 Introduction

Ce chapitre pourrait posséder trois points de départ :

- La biorthogonalité
- Les approximants de Padé matriciels
- L'opérateur aux différences

Il existe, comme dans le cas scalaire, une interaction entre ces différents aspects comme le présente le schéma ci-après.

Ce diagramme décrit les divers aspects des polynômes matriciellement orthogonaux. Il s'agit en fait de démontrer que ces polynômes vectoriels orthogonaux relativement à une matrice de dimensions $r \times s$ de mesures, possèdent les mêmes propriétés courantes que les polynômes scalaires.



Dans ce chapitre, on a choisi de débiter par l'aspect de la biorthogonalité en définissant une forme bilinéaire pour des polynômes vectoriels avec le paragraphe 1.2.1.

Cette forme bilinéaire permet de définir des moments matriciels (§ 1.2.2) puis une récurrence à $r + s + 1$ termes (§ 1.2.3), les polynômes scalaires possédant, eux, une récurrence à 3 termes. Cette récurrence est équivalente à la donnée de la forme bilinéaire (Théorème 1.13 de Shohat-Favard).

A partir de ces récurrences où tous les coefficients sont uniformément bornés, il est possible de définir un opérateur aux différences sur $\mathcal{L}(\ell^2)$ (§ 1.3.1).

Les récurrences permettent de passer aux fractions continues vectorielles ou matricielles. Le n -ième convergent de ces fractions continues est un approximant de Padé (§ 1.2.5). Dans cette partie du chapitre, nous verrons quelques cas particuliers comme les approximants simultanés d'Hermitte-Padé qui ont servi, par exemple, dans les démonstrations de transcendance ou d'irrationalité.

Ensuite, grâce à la dualité, on peut, comme dans le cas scalaire, démontrer une identité de Christoffel-Darboux. Ainsi que C. Brezinski l'a démontré dans [20], il serait possible de démontrer que l'identité de Christoffel-Darboux est équivalente à la donnée de la récurrence à $r + s + 1$ termes (voir à ce sujet [71, § 8]).

En repartant des moments matriciels, on définit la fonction génératrice $\varphi(z)$ (§ 1.2.1). Cette fonction est liée à la fonction de Weyl, décrite par Bere-

zanskii, que l'on définit à partir de la résolvante de l'opérateur aux différences (§ 1.3.1). On peut aussi revenir en arrière et créer la forme bilinéaire à partir de la fonction génératrice (§ 1.2.2).

Ces diverses notions s'imbriquent étroitement dans la notion centrale des approximants de Padé. On en étudiera la formation à partir de la forme bilinéaire (§ 1.2.5) ou de la récurrence à $r + s + 1$ termes et la convergence vers la fonction génératrice ou la fonction de Weyl en dehors de l'image numérique (§ 1.3.1).

Ces approximants de Padé permettront, en particulier, de caractériser l'ensemble résolvant de l'opérateur aux différences.

En utilisant cette caractérisation, on pourra redémontrer, de manière nouvelle, un résultat de P. Duren [29] sur le spectre des matrices de Toeplitz bande.

Il est à signaler que ce chapitre est très largement inspiré des articles de B. Beckermann [5] et J. Van Iseghem et V. Sorokin [71] en ce qui concerne les définitions et les nombreuses propriétés des polynômes vectoriels orthogonaux et les approximants de Padé matriciels.

1.2 Biorthogonalité

1.2.1 Forme bilinéaire

On se donne une forme bilinéaire $\langle \cdot, \cdot \rangle$ sur $\mathbb{C}^{1 \times r}[z] \times \mathbb{C}^{s \times 1}[z]$ vérifiant pour tous $P \in \mathbb{C}^{1 \times r}[z]$ et $Q \in \mathbb{C}^{s \times 1}[z]$, la "relation de structure"

$$\langle zP(z), Q(z) \rangle = \langle P(z), zQ(z) \rangle \quad (1.1)$$

Il est, par exemple, possible de choisir une intégration par rapport à une matrice de mesures dM sur un contour du plan complexe. Pour tout $F \in \mathbb{C}^{1 \times r}$ et $G \in \mathbb{C}^{s \times 1}$

$$\langle F(x), G(x) \rangle = \int_{\gamma} F(x) dM(x) G(x) \quad (1.2)$$

où dM est une mesure matricielle c'est-à-dire une matrice de mesures. On pourrait choisir $dM(x) = W(x)dx$ avec $W(x)$ une matrice de fonctions poids. On pourrait aussi travailler sur un intervalle de \mathbb{R} au lieu d'un contour du plan complexe.

On définit le degré (diagonal) d'un polynôme vectoriel.

Définition 1.1. Soit $Q \in \mathbb{C}^{s \times 1}[z]$. On définit son degré (diagonal) par $ddeg Q := deg q$ où $q(x) = (1, x, \dots, x^{s-1}) \cdot Q(x^s)$.

De manière similaire, pour $P \in \mathbb{C}^{1 \times r}[z]$, on pose $ddeg P = ddeg P^T$.

On peut remarquer que si $Q \in \mathbb{C}^{s \times 1}[z]$, alors $ddeg(zQ) = ddeg Q + s$ et si $Q \in \mathbb{C}^{1 \times r}[z]$, alors $ddeg(zQ) = ddeg Q + r$.

Exemple 1.2. Une suite (Q_n) d'éléments de $\mathbb{C}^{1 \times 2}[z]$ tels que $ddeg Q_n = n$, $Q_n = (q_{1,n}, q_{2,n})$, possède les bornes suivantes pour le degré de chaque composante

n	0	1	2	3	4	5	6	...
$deg q_{1,n}$	$= 0$	≤ 0	$= 1$	≤ 1	$= 2$	≤ 2	$= 3$...
$deg q_{2,n}$	$= -\infty$	$= 0$	≤ 0	$= 1$	≤ 1	$= 2$	≤ 2	...

C'est-à-dire que le degré diagonal de Q_n impose le degré d'une des composantes et une inégalité pour le degré de l'autre composante.

Grâce à la forme bilinéaire, il est possible de définir les polynômes orthogonaux à droite et à gauche.

Définition 1.3. Un polynôme vectoriel $Q_n^R \in \mathbb{C}^{s \times 1}[z]$ de degré $ddeg Q_n^R \leq n$ sera appelé n -ième polynôme orthogonal à droite si et seulement si $\langle P, Q_n^R \rangle = 0$ pour tout $P \in \mathbb{C}^{1 \times r}[z]$, $ddeg P < n$.

De la même manière, on définirait $Q_n^L \in \mathbb{C}^{1 \times r}[z]$, n -ième polynôme orthogonal à gauche.

Exemple 1.4. Si $r = s = 1$, nous retrouvons le cas scalaire. Si $r = 1, s > 1$, Il s'agit des polynômes orthogonaux multiples ou des polynômes orthogonaux simultanés.[2].

1.2.2 Algèbre linéaire

En regroupant les polynômes vectoriels sous une forme matricielle avec $P \in \mathbb{C}^{j \times r}[z]$ et $Q \in \mathbb{C}^{s \times k}[z]$, on peut définir une forme bilinéaire de $\mathbb{C}^{j \times r}[z] \times \mathbb{C}^{s \times k}[z]$ à valeurs dans $\mathbb{C}^{j \times k}$ par

$$\langle P, Q \rangle = \begin{pmatrix} \langle P_{1..r}, Q_{1..s,1} \rangle & \langle P_{1..r}, Q_{1..s,k} \rangle \\ \vdots & \vdots \\ \langle P_{j..r}, Q_{1..s,1} \rangle & \langle P_{j..r}, Q_{1..s,k} \rangle \end{pmatrix}$$

où $P_{n..r}$ indique la n -ième ligne de P et $Q_{1..s,m}$ la m -ième colonne de Q .

Si l'on note I_r et I_s les matrices identité de taille r et s , on peut alors définir les moments \mathcal{M}_k par $\mathcal{M}_k = \langle I_r, x^k \cdot I_s \rangle$ que l'on regroupe dans une matrice infinie, dite Hankel par blocs :

$$H = \begin{pmatrix} \mathcal{M}_0 & \mathcal{M}_1 & \mathcal{M}_2 & \dots \\ \mathcal{M}_1 & \mathcal{M}_2 & \mathcal{M}_3 & \dots \\ \mathcal{M}_2 & \mathcal{M}_3 & \mathcal{M}_4 & \dots \\ \vdots & \vdots & \vdots & \dots \end{pmatrix}$$

On note $H_{n,m}$ la sous matrice supérieure gauche de H à n lignes et m colonnes et $H_n = H_{n,n}$.

On obtient alors le lemme suivant :

Lemme 1.5. *Supposons que $\det H_n \neq 0$ alors Q_n^L et Q_n^R sont uniques à une constante multiplicative près. De plus, ils possèdent le degré correct, c'est-à-dire que $ddeg Q_n^L = ddeg Q_n^R = n$.*

Démonstration. Les polynômes Q_n^L et Q_n^R existent d'après leurs définitions : il suffit de résoudre un système homogène pour déterminer ces polynômes. Le vecteur a des coefficients du polynôme scalaire $(1, z, \dots, z^{s-1}) \cdot Q_n^R(z^s)$ permet de déterminer complètement le polynôme vectoriel Q_n^R . De plus, ce vecteur vérifie $H_{n,n+1} \cdot a = 0$ d'après la définition du polynôme orthogonal Q_n^R . Comme $\det H_n \neq 0$, le vecteur a est déterminé à une constante multiplicative près, la dernière composante étant non nulle. On obtient bien l'existence de Q_n^R et son unicité. Le raisonnement est le même pour Q_n^L avec $H_{n+1,n}$. \square

Corollaire 1.6. *Pour tout n , $\langle Q_n^L, Q_n^R \rangle \neq 0$ si et seulement si pour tout n , $\det H_n \neq 0$, c'est-à-dire que la matrice H est fortement régulière.*

Démonstration. Démontrons donc la condition suffisante.

On suppose que, pour tout n , $\langle Q_n^L, Q_n^R \rangle \neq 0$, c'est-à-dire que pour tout n , Q_n^L et Q_n^R sont de degré exact n . Si on regroupe les coefficients des polynômes scalaires $(1, z, \dots, z^{s-1}) \cdot Q_n^R(z^s)$ et $Q_n^L(z^r) \cdot (1, z, \dots, z^{r-1})^t$ dans des matrices R et L , celles-ci sont respectivement triangulaires supérieure et inférieure. Les coefficients diagonaux sont tous non nuls car les polynômes sont de degré exact.

On effectue le produit LHR . On obtient, grâce à la définition des polynômes vectoriels orthogonaux, une matrice diagonale $(\langle Q_j^L, Q_k^R \rangle)_{j,k}$, les coefficients diagonaux étant non nuls. Les matrices L et R sont inversibles donc pour tout n , H_n est inversible, ce qui complète la démonstration. \square

On peut définir la base "canonique" de l'espace $\mathbb{C}^{s \times 1}[z]$ par

$$\eta_0^R = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \eta_{s-1}^R = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \eta_s^R = \begin{pmatrix} z \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \eta_{2s}^R = \begin{pmatrix} z^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots$$

C'est-à-dire que si l'on écrit n sous la forme $n = sq + q_0$, $0 \leq q_0 < s$, alors les composantes de η_n^R sont nulles, excepté la composante $q_0 + 1$ qui est égale à z^q .

On pourrait définir, de la même manière, une base canonique $(\eta_0^L, \dots, \eta_r^L, \dots)$ pour les polynômes vectoriels $\mathbb{C}^{1 \times r}[z]$.

On supposera désormais que les polynômes vectoriels orthogonaux vérifient $\langle Q_n^L(z), Q_n^R(z) \rangle = 1$, $n = 0, \dots$ après renormalisation. A cause de l'orthogonalité des Q_n^L et Q_n^R , l'initialisation de ces récurrences est réalisée selon la décomposition de H :

$$LHR = I \quad (1.3)$$

Cette décomposition sera utilisée dans le chapitre 3 sur les moments modifiés comme nous le rappellerons dans la remarque 3.9. Comme indiqué précédemment, la $n + 1$ -ème ligne de L donne les coefficients du polynôme scalaire $Q_n^L(x^r) \cdot (1, \dots, x^{r-1})^T$. La $n + 1$ -ème colonne de R contient les coefficients du polynôme scalaire $(1, \dots, x^{s-1}) \cdot Q_n^R(x^s)$. A partir des moments \mathcal{M}_k , on peut définir la série formelle génératrice

$$\varphi(z) = \sum_{j=0}^{+\infty} \mathcal{M}_j z^{-j-1} \quad (1.4)$$

Cette série est reliée à la forme bilinéaire de la manière suivante.

Propriété 1.7. Soient $P \in \mathbb{C}^{1 \times r}[z]$ et $Q \in \mathbb{C}^{s \times 1}[z]$ deux polynômes vectoriels et $\varphi(z)$ la série génératrice. Alors on obtient l'égalité

$$\langle P, Q \rangle = \text{coeff}(P\varphi Q, z^{-1})$$

Démonstration. Il suffit de vérifier cette égalité pour les éléments de la base. Soient donc η_n^L , $n = jr + j'$ et η_m^R , $m = ks + k'$.

$$\begin{aligned} \langle \eta_n^L(x), \eta_m^R(x) \rangle &= \langle x^j \eta_{j'}^L, x^k \eta_{k'}^R \rangle \text{ d'après la définition des bases} \\ &= \langle \eta_{j'}^L I_r, x^{j+k} I_s \eta_{k'}^R \rangle \text{ d'après la relation (1.1)} \\ &= \eta_{j'}^L \langle I_r, x^{j+k} I_s \rangle \eta_{k'}^R \\ &= \eta_{j'}^L \mathcal{M}_{j+k} \eta_{k'}^R \end{aligned}$$

Si on effectue le produit $\eta_n^L(z)\varphi(z)\eta_m^R(z)$, on obtient $z^{j+k}\eta_{j'}^L\varphi(z)\eta_{k'}^R$. Le coefficient de z^{-1} dans cette série est donc bien $\eta_{j'}^L\mathcal{M}_{j+k}\eta_{k'}^R$. \square

Cette propriété permet le passage de la série génératrice à la forme bilinéaire et fournit ainsi la matrice des moments H .

1.2.3 Récurrence à $r + s + 1$ termes

On définit les numérateurs, polynômes associés aux dénominateurs Q_n^L et Q_n^R .

Définition 1.8.

$$P_n^L(z) = \left\langle \frac{Q_n^L(x) - Q_n^L(z)}{x - z}, I_s \right\rangle$$

$$P_n^R(z) = \left\langle I_r, \frac{Q_n^R(x) - Q_n^R(z)}{x - z} \right\rangle$$

On a les initialisations suivantes

Lemme 1.9. *Les dénominateurs vérifient*

$$\forall k \in \{-s, \dots, -1\}, Q_k^L = 0 \text{ et } \forall k \in \{-r, \dots, -1\}, Q_k^R = 0.$$

tandis que les numérateurs vérifient les égalités suivantes

$$\forall k \in \{0, \dots, r-1\}, P_k^L = 0, \quad \forall k \in \{0, \dots, s-1\}, P_k^R = 0$$

Démonstration. Les dénominateurs Q_k^L , $k \in \{0, \dots, r-1\}$ et Q_k^R , $k \in \{0, \dots, s-1\}$ sont déterminés à partir de la décomposition (1.3). Ce sont des polynômes vectoriels dont les coordonnées ne dépendent pas de z . Ainsi, lorsque l'on applique la définition pour obtenir les numérateurs P_k^L et P_k^R , on obtient bien la valeur 0. Le choix d'annulation des dénominateurs d'indices négatifs est habituel. \square

Propriété 1.10. *Toujours sous l'hypothèse de régularité et de normalisation, ces suites de polynômes vérifient une relation de récurrence à $r+s+1$ termes.*

$$zQ_n^L(z) = \sum_{j=n-s}^{n+r} \langle zQ_n^L(z), Q_j^R(z) \rangle Q_j^L(z) =: \sum_{j=n-s}^{n+r} A_{n,j} Q_j^L(z), \quad n \geq s$$

$$zQ_n^R(z) = \sum_{j=n-r}^{n+s} \langle Q_j^L(z), zQ_n^R(z) \rangle Q_j^R(z) =: \sum_{j=n-r}^{n+s} A'_{j,n} Q_j^R(z), \quad n \geq r$$

Dans le produit, comme indiqué dans la relation de structure (1.1), la variable "x" peut être déplacée d'un côté à l'autre. Donc les quantités $A_{n,j}$ et $A'_{k,m}$ sont en fait les mêmes pour $k = n$ et $j = m$.

Les relations restent valables pour tous $n \geq 0$, puisque l'on a choisi d'annuler les dénominateurs d'indices négatifs.

Démonstration. Comme remarqué précédemment,

$$ddeg zQ_n^L(z) = ddeg Q_n^L(z) + r = n + r.$$

On décompose donc $zQ_n^L(z)$ suivant les polynômes vectoriels Q_k^L de degré inférieur,

$$zQ_n^L(z) = \sum_{j=0}^{n+r} \beta_j Q_j^L(z)$$

On forme alors le produit avec les Q_j^R afin d'obtenir la valeur des β_j . De plus, les β_j sont nuls pour $j < n - r$. En effet, $\langle zQ_n^L(z), Q_j^R(z) \rangle = \langle Q_n^L(z), zQ_j^R(z) \rangle = 0$ pour $j < n - r$ car $zQ_j^R(z)$ est de degré diagonal strictement inférieur à n .

Le raisonnement peut être appliqué de la même manière pour obtenir l'autre récurrence. \square

Comme le degré diagonal est exact, les coefficients extrêmes des décompositions sont non nuls. Cela permet donc de définir la matrice A .

Définition 1.11. On notera A , la matrice contenant les coefficients $(A_{n,j})$. A est une matrice bande avec r surdiagonales et s sous diagonales, les diagonales extrêmes sont constituées d'éléments non nuls.

Le bloc de la matrice A , $\begin{pmatrix} A_{j,k} & \dots & A_{j,m} \\ & & \\ & & \\ A_{n,k} & \dots & A_{n,m} \end{pmatrix}$, sera noté $A \begin{pmatrix} k, \dots, m \\ j, \dots, n \end{pmatrix}$

Cette matrice A représente l'opérateur aux différences.

On regroupe les polynômes Q_n et P_n , ce qui permet de mettre sous forme matricielle les relations de récurrence précédentes :

Theorème 1.12. En posant

$$\mathcal{Q}^R := (Q_0^R, \dots, Q_k^R, \dots) \text{ et } \mathcal{Q}^L := (Q_0^L, \dots, Q_k^L, \dots)^t$$

$$\mathcal{Q}_{n,\lambda}^R := (Q_n^R, \dots, Q_{n+\lambda-1}^R) \in \mathbb{C}^{s \times \lambda}[z] \text{ et } \mathcal{Q}_{n,\lambda}^L := (Q_n^L, \dots, Q_{n+\lambda-1}^L)^t \in \mathbb{C}^{\lambda \times r}[z]$$

on obtient

$$\begin{aligned} z\mathcal{Q}^R(z) &= \mathcal{Q}^R(z)A \\ z\mathcal{Q}^L(z) &= A\mathcal{Q}^L(z) \end{aligned}$$

On définit \mathcal{P}^L , \mathcal{P}^R , $\mathcal{P}_{n,\lambda}^R$ et $\mathcal{P}_{n,\lambda}^L$ de la même manière que pour les Q_n . On obtient alors

$$z\mathcal{P}^R(z) + ((\mathcal{Q}_{0,r}^L(z))^{-1}, 0_{r \times \infty}) = \mathcal{P}^R(z)A, \quad z\mathcal{P}^L(z) + \begin{pmatrix} (\mathcal{Q}_{0,s}^R(z))^{-1} \\ 0_{\infty \times s} \end{pmatrix} = A\mathcal{P}^L(z)$$

De plus, les degrés diagonaux vérifient

$$ddeg P_n^L \leq n - r, \quad n \geq r \text{ et } ddeg P_n^R \leq n - s, \quad n \geq s$$

Si $r = s$, on retrouve les polynômes orthogonaux matriciels de A. Sinap et W. Van Assche [70].

Démonstration. Les bornes sur les degrés diagonaux se déduisent immédiatement de la définition des numérateurs.

Il reste à démontrer les relations de récurrence des numérateurs. Grâce à la symétrie, il suffit de démontrer le théorème pour \mathcal{P}^L . D'après les relations de récurrence du théorème 1.12 et la définition de \mathcal{P}^L , on obtient, de manière formelle

$$\begin{aligned} A\mathcal{P}^L(z) - z\mathcal{P}^L(z) &= (A - zI) \left\langle \frac{\mathcal{Q}^L(x) - \mathcal{Q}^L(z)}{x - z}, I_s \right\rangle \\ &= \left\langle \frac{x\mathcal{Q}^L(x) - z\mathcal{Q}^L(x)}{x - z}, I_s \right\rangle \\ &= \langle \mathcal{Q}^L(x), I_s \rangle \end{aligned}$$

D'après l'orthogonalité, la dernière quantité obtenue est non nulle pour au plus les s premières lignes.

Comme, de plus,

$$I_s = \langle \mathcal{Q}_{0,s}^L, \mathcal{Q}_{0,s}^R \rangle = \langle \mathcal{Q}_{0,s}^L, I_s \rangle \mathcal{Q}_{0,s}^R$$

on obtient bien la récurrence demandée. \square

La donnée d'une forme bilinéaire permet d'obtenir des formules de récurrence à $r + s + 1$ termes pour les polynômes vectoriels orthogonaux et les polynômes associés. Il est donc intéressant de savoir si, comme dans le cas scalaire [58, p.54& sq.], la donnée de récurrence définit l'existence d'une forme bilinéaire unique. Ce théorème, dit de Shohat-Favard, est vérifié dans le cas des polynômes matriciellement orthogonaux. Nous donnons une version bilinéaire de ce théorème tirée de [71, Théorème 2].

Théorème 1.13. *Si $P_n = (P_n^1, \dots, P_n^s)^t$ et $Q_n = (Q_n^1, \dots, Q_n^r)$ sont deux suites de polynômes vectoriels de tailles respectives s et r tels que les s (resp. r) premiers soient de degré diagonal exact, c'est-à-dire $d\deg P_i = i$ et $d\deg Q_j = j$, vérifiant de plus les récurrences suivantes*

$$xP_n = \sum_{k=-r}^s a_{n+k,n} P_{n+k}, \quad P_{-k} = 0 \quad k > 0, \quad P_0, \dots, P_{s-1} \text{ étant fixés}$$

$$xQ_n = \sum_{k=-s}^r a_{n,n+k} Q_{n+k}, \quad Q_{-k} = 0 \quad k > 0, \quad Q_0, \dots, Q_{r-1} \text{ étant fixés}$$

où $a_{n+s,n} \neq 0$ et $a_{n,n+r} \neq 0$ pour tous $n \geq 0$ alors il existe une forme bilinéaire $\langle \cdot, \cdot \rangle$ vérifiant la relation de structure (1.1) telle que

$$\langle Q_m, P_n \rangle = \delta_{m,n}, \quad n, m \geq 0$$

Démonstration. Q_n étant de degré diagonal exact, on peut l'écrire sur la base sous la forme $Q_n = \alpha_n \eta_n^L + \dots$

Calculons $\langle Q_n, P_n \rangle$. En notant $n = kr + k_0$, $k_0 < r$, on obtient

$$\begin{aligned} \langle Q_n, P_n \rangle &= \alpha_n \langle \eta_n^L, P_n \rangle \\ &= \alpha_n \langle x \eta_{n-r}^L, P_n \rangle \text{ d'après la définition de la base} \\ &= \alpha_n \langle \eta_{n-r}^L, x P_n \rangle \text{ d'après la relation de structure (1.1)} \\ &= \alpha_n a_{n-r, n} \langle \eta_{n-r}^L, P_{n-r} \rangle \text{ d'après la formule de récurrence} \\ &= \alpha_n a_{n-r, n} a_{n-2r, n-r} \dots a_{k_0, k_0+r} \langle \eta_{k_0}^L, P_{k_0} \rangle \end{aligned}$$

De plus, on peut exprimer le coefficient α_n en fonction des coefficients de récurrence en comparant les coefficients de tête :

$$\begin{aligned} x Q_n &= \sum_{k=-s}^r a_{n, n+k} Q_{n+k} \\ \alpha_n &= a_{n, n+r} \alpha_{n+r} \\ \alpha_{k_0} &= a_{k_0, k_0+r} \dots a_{n-r, n} \alpha_n \text{ par une récurrence immédiate} \end{aligned}$$

Ce qui nous donne finalement

$$\langle Q_n, P_n \rangle = \langle Q_{k_0}, P_{k_0} \rangle$$

Ainsi, si l'on suppose que $\langle Q_n, P_n \rangle = 1$ pour $n = 0, \dots, r-1$, alors cette relation est vérifiée pour tout n .

Les conditions initiales sont déterminées à partir des $\frac{r(r+1)}{2}$ valeurs $\langle \eta_i^L, \eta_j^R \rangle$, $0 \leq j \leq i \leq r-1$, c'est-à-dire, de manière équivalente, à partir des $\frac{r(r+1)}{2}$ valeurs $\langle Q_i, P_j \rangle$, $0 \leq j \leq i \leq r-1$. Ainsi, savoir que ces valeurs sont égales au symbole de Kronecker revient à définir une unique forme bilinéaire $\langle \cdot, \cdot \rangle$ telle que

$$\forall m, n \geq 0, \langle Q_m, P_n \rangle = \delta_{m, n}$$

□

1.2.4 Dualité et Christoffel-Darboux

Propriété 1.14. *Ces polynômes vérifient la formule de Christoffel-Darboux, démontrée par B. Beckermann dans [5] et par V. N. Sorokin et J. Van Iseghem dans [71]. Pour $n \geq 0$,*

$$\begin{aligned} K_n(x, y) &= \sum_{k=0}^{n-1} Q_k^R(x) Q_k^L(y) \\ &= \frac{Q_{n,s}^R(x) A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} Q_{n-s,s}^L(y) - Q_{n-r,r}^R(x) A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} Q_{n,r}^L(y)}{x - y} \end{aligned}$$

$K_n(x, y)$ est le noyau de Szëgo.

Démonstration. Supposons $n = 0$. Alors la somme est vide donc elle est nulle par convention. La seconde partie de l'égalité comporte $\mathcal{Q}_{-s,s}^L$ et $\mathcal{Q}_{-r,r}$ qui sont nuls d'après l'initialisation. L'égalité est donc vérifiée pour $n = 0$.

Supposons maintenant $n \geq 1$. Notons par E_n la matrice infinie comportant des 1 pour les n premières composantes de la diagonale et 0 ailleurs. Alors

$$\begin{aligned} \sum_{k=0}^{n-1} \mathcal{Q}_k^R(x) \mathcal{Q}_k^L(y) &= \mathcal{Q}^R E_n \mathcal{Q}^L \\ &= \frac{x \mathcal{Q}^R E_n \mathcal{Q}^L - \mathcal{Q}^R E_n y \mathcal{Q}^L}{x - y} \\ &= \frac{\mathcal{Q}^R A E_n \mathcal{Q}^L - \mathcal{Q}^R E_n A \mathcal{Q}^L}{x - y} \\ &= \frac{\mathcal{Q}^R (A E_n - E_n A) \mathcal{Q}^L}{x - y} \end{aligned}$$

On examine plus précisément la forme de $A E_n - E_n A$.

$$\begin{aligned} A E_n - E_n A &= \begin{pmatrix} A \begin{pmatrix} 0, \dots, n-1 \\ 0, \dots, n-1 \end{pmatrix} & 0_{n \times \infty} \\ A \begin{pmatrix} 0, \dots, n-1 \\ n, \dots, \infty \end{pmatrix} & 0_{\infty \times \infty} \end{pmatrix} - \begin{pmatrix} A \begin{pmatrix} 0, \dots, n-1 \\ 0, \dots, n-1 \end{pmatrix} & A \begin{pmatrix} n, \dots, \infty \\ 0, \dots, n-1 \end{pmatrix} \\ 0_{\infty \times n} & 0_{\infty \times \infty} \end{pmatrix} \\ &= \begin{pmatrix} 0_{n \times n} & -A \begin{pmatrix} n, \dots, \infty \\ 0, \dots, n-1 \end{pmatrix} \\ A \begin{pmatrix} 0, \dots, n-1 \\ n, \dots, \infty \end{pmatrix} & 0_{\infty \times \infty} \end{pmatrix} \end{aligned}$$

Supposons ici que $n \geq \max(r, s)$. Comme A est une matrice bande avec r surdiagonales et s sous-diagonales, on obtient finalement la forme suivante :

$$\begin{pmatrix} 0_{(n-r) \times (n-s)} & 0_{(n-r) \times s} & 0_{(n-r) \times r} & 0_{(n-r) \times \infty} \\ 0_{r \times (n-s)} & 0_{r \times s} & -A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} & 0_{r \times \infty} \\ 0_{s \times (n-s)} & A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} & 0_{s \times r} & 0_{s \times \infty} \\ 0_{\infty \times (n-s)} & 0_{\infty \times s} & 0_{\infty \times r} & 0_{\infty \times \infty} \end{pmatrix}$$

En effectuant le calcul matriciel, on obtient bien la formule demandée.

Si $1 \leq n < \max(r, s)$, certaines des matrices nulles présentes dans la forme précédente n'existent plus. Cependant, comme les polynômes \mathcal{Q}_k^L et \mathcal{Q}_k^R d'indices négatifs sont nuls, les produits restent valables. En effet, pour

$$n < s, \quad A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{Q}_{n-s,s}^L(y) = A \begin{pmatrix} 0, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{Q}_{0,n}^L(y)$$

et

$$n < r, \quad \mathcal{Q}_{n-r,r}^R(x) A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} = \mathcal{Q}_{0,n}^R(x) A \begin{pmatrix} n, \dots, n+r-1 \\ 0, \dots, n-1 \end{pmatrix}$$

□

En multipliant par $(x-y)$, et en posant $y = x$, on trouve

Corollaire 1.15. *Les polynômes vérifient l'égalité suivante :*

$$\mathcal{Q}_{n,s}^R(x)A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{Q}_{n-s,s}^L(x) - \mathcal{Q}_{n-r,r}^R(x)A \begin{pmatrix} n-r, \dots, n-1 \\ n, \dots, n+r-1 \end{pmatrix} \mathcal{Q}_{n,r}^L(x) = 0 \quad (1.5)$$

Propriété 1.16. *Le noyau $K_n(x, y)$ est un noyau reproduisant, c'est-à-dire, pour $n \geq 1$,*

$$\langle K_n(x, y), P(y) \rangle = P(x), \quad \forall P \in \mathbb{C}^{s \times 1}[x], \quad ddeg P < n$$

$$\langle Q(y), K_n(y, x) \rangle = Q(x), \quad \forall Q \in \mathbb{C}^{1 \times r}[x], \quad ddeg Q < n$$

Démonstration. Il suffit d'exprimer P ou Q suivant (Q_k^R) ou (Q_k^L) . En utilisant ensuite l'orthogonalité, on obtient les égalités de la propriété. \square

Remarque 1.17. *Dans le cas scalaire, la formule de Christoffel-Darboux se présente sous la forme suivante :*

$$\sum_{k=0}^n P_k(x)P_k(y) = A_{n,n+1} \frac{P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)}{x-y}$$

$$\sum_{k=0}^n P_k^2(x) = A_{n,n+1} [P'_{n+1}(x)P_n(x) - P'_n(x)P_{n+1}(x)]$$

Afin d'étudier simultanément les polynômes matriciellement orthogonaux, on les regroupe

Définition 1.18.

$$B_n^L = \left(\begin{array}{c|c} \mathcal{Q}_{n-s,s}^L & -\mathcal{P}_{n-s,s}^L \\ \mathcal{Q}_{n,r}^L & -\mathcal{P}_{n,r}^L \end{array} \right) \quad \text{et} \quad B_n^R = \left(\begin{array}{c|c} \mathcal{P}_{n-r,r}^R & \mathcal{P}_{n,s}^R \\ \mathcal{Q}_{n-r,r}^R & \mathcal{Q}_{n,s}^R \end{array} \right)$$

On donne une propriété présentant la dualité de ces polynômes matriciellement orthogonaux. Cette propriété sera utilisée dans ce chapitre au cours de la démonstration du théorème 1.44 ainsi que dans la propriété 2.1 du chapitre 2 sur l'accélération de convergence.

Propriété 1.19.

$$\forall n \geq 1, \quad B_n^R \cdot M_n \cdot B_n^L = I_{r+s} \quad (1.6)$$

où

$$M_n = \left(\begin{array}{cc} 0_{r \times s} & -A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} \\ A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} & 0_{s \times r} \end{array} \right)$$

Les coefficients $A_{n,j}$ d'indices négatifs sont artificiels et ne servent qu'à obtenir des formules valables pour $n < \max(r, s)$.

Démonstration. Effectuons le produit matriciel, ce qui donne

$$\forall n \geq 1, B_n^R \cdot M_n \cdot B_n^L := T = \begin{pmatrix} T_{1,1} & T_{1,2} \\ T_{2,1} & T_{2,2} \end{pmatrix}$$

où

$$\begin{aligned} T_{1,1} &= \mathcal{P}_{n,s}^R A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{Q}_{n-s,s}^L - \mathcal{P}_{n-r,r}^R A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} \mathcal{Q}_{n,r}^L \\ T_{1,2} &= -\mathcal{P}_{n,s}^R A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{P}_{n-s,s}^L + \mathcal{P}_{n-r,r}^R A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} \mathcal{P}_{n,r}^L \\ T_{2,1} &= \mathcal{Q}_{n,s}^R A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{Q}_{n-s,s}^L - \mathcal{Q}_{n-r,r}^R A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} \mathcal{Q}_{n,r}^L \\ T_{2,2} &= -\mathcal{Q}_{n,s}^R A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{P}_{n-s,s}^L + \mathcal{Q}_{n-r,r}^R A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} \mathcal{P}_{n,r}^L \end{aligned}$$

Grâce à l'équation (1.5), on obtient $T_{2,1} = 0$.

Examinons $T_{1,1}$ et $T_{2,2}$ qui se traitent tous les deux de la même manière.

On notera $I_s(x) = I_s$ afin d'indiquer la variable sur laquelle agit la forme bilinéaire.

$$\begin{aligned} T_{1,1}(z) &= \left\langle I_r(x), \frac{(\mathcal{Q}_{n,s}^R(x) - \mathcal{Q}_{n,s}^R(z)) A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{Q}_{n-s,s}^L(z)}{x-z} \right\rangle \\ &\quad - \left\langle I_r(x), \frac{(\mathcal{Q}_{n-r,r}^R(x) - \mathcal{Q}_{n-r,r}^R(z)) A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} \mathcal{Q}_{n,r}^L(z)}{x-z} \right\rangle \\ &= \left\langle I_r(x), \frac{\mathcal{Q}_{n,s}^R(x) A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \mathcal{Q}_{n-s,s}^L(z) - \mathcal{Q}_{n-r,r}^R(x) A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} \mathcal{Q}_{n,r}^L(z)}{x-z} \right\rangle \\ & \hspace{20em} (\text{Puisque } T_{2,1} = 0) \\ &= \langle I_r(x), K_n(x, z) \rangle \\ &= I_r(z) \hspace{15em} (\text{d'après la propriété 1.16 du noyau reproduisant}) \end{aligned}$$

On obtient donc $T_{1,1} = I_r$ et $T_{2,2} = I_s$. En appliquant la même technique, à

droite et à gauche, on obtient la forme suivante de $-T_{1,2}(z)$:

$$\begin{aligned}
& \left\langle I_r(u), \frac{\mathcal{Q}_{n,s}^R(u) - \mathcal{Q}_{n,s}^R(z)}{u-z} \right\rangle A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \left\langle \frac{\mathcal{Q}_{n-s,s}^L(v) - \mathcal{Q}_{n-s,s}^L(z)}{v-z}, I_s(v) \right\rangle \\
& - \left\langle I_r(u), \frac{\mathcal{Q}_{n-r,r}^R(u) - \mathcal{Q}_{n-r,r}^R(z)}{u-z} \right\rangle A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} \left\langle \frac{\mathcal{Q}_{n,r}^L(v) - \mathcal{Q}_{n,r}^L(z)}{v-z}, I_s(v) \right\rangle \\
& = \left\langle I_r(u), \left\langle \frac{(\mathcal{Q}_{n,s}^R(u) - \mathcal{Q}_{n,s}^R(z)) A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} (\mathcal{Q}_{n-s,s}^L(v) - \mathcal{Q}_{n-s,s}^L(z))}{(u-z)(v-z)}, I_s(v) \right\rangle \right\rangle \\
& - \left\langle I_r(u), \left\langle \frac{(\mathcal{Q}_{n-r,r}^R(u) - \mathcal{Q}_{n-r,r}^R(z)) A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix} (\mathcal{Q}_{n,r}^L(v) - \mathcal{Q}_{n,r}^L(z))}{(u-z)(v-z)}, I_s(v) \right\rangle \right\rangle
\end{aligned}$$

En développant, et en utilisant l'équation (1.5) et la formule de Christoffel-Darboux, on obtient

$$\begin{aligned}
& = \left\langle I_r(u), \left\langle \frac{K_n(u,v)(u-v) - K_n(z,v)(z-v) - K_n(u,z)(u-z)}{(u-z)(v-z)}, I_s(v) \right\rangle \right\rangle \\
& = \left\langle I_r(u), \left\langle \frac{K_n(u,v)(u-z+z-v) - K_n(z,v)(z-v) - K_n(u,z)(u-z)}{(u-z)(v-z)}, I_s(v) \right\rangle \right\rangle \\
& = \left\langle \left\langle I_r(u), \frac{K_n(u,v) - K_n(u,z)}{v-z} \right\rangle, I_s(v) \right\rangle - \left\langle I_r(u), \left\langle \frac{K_n(u,v) - K_n(z,v)}{u-z}, I_s(v) \right\rangle \right\rangle \\
& = \left\langle \frac{I_r(v) - I_r(z)}{v-z}, I_s(v) \right\rangle - \left\langle I_r(u), \frac{I_s(u) - I_s(z)}{u-z} \right\rangle = 0.
\end{aligned}$$

Pour obtenir la dernière relation, on a utilisé la propriété 1.16 du noyau reproduisant.

Finalement $B_n^R M_n B_n^L = I_{r+s}$. \square

Corollaire 1.20. *On obtient alors*

$$B_n^L B_n^R = M_n^{-1} = \begin{pmatrix} 0_{s \times r} & A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix}^{-1} \\ -A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix}^{-1} & 0_{r \times s} \end{pmatrix}$$

Démonstration. Pour tout n , la matrice M_n est inversible d'après la relation (1.6). \square

Corollaire 1.21. *[Initialisation des récurrences] On a*

$$B_0^R = \begin{pmatrix} -(\mathcal{Q}_{0,r}^L)^{-1} A \begin{pmatrix} 0 \dots r-1 \\ -r \dots -1 \end{pmatrix}^{-1} & 0_{r \times s} \\ 0_{s \times r} & \mathcal{Q}_{0,s}^R \end{pmatrix}$$

et

$$B_0^L = \begin{pmatrix} 0_{s \times r} & A \begin{pmatrix} -s \dots -1 \\ 0 \dots s-1 \end{pmatrix}^{-1} (\mathcal{Q}_{0,s}^R)^{-1} \\ \mathcal{Q}_{0,r}^L & 0_{r \times s} \end{pmatrix}$$

Démonstration. Les quantités $\mathcal{P}_{0,s}^R$, $\mathcal{Q}_{0,s}^R$, $\mathcal{P}_{0,r}^L$ et $\mathcal{Q}_{0,r}^L$ sont connues à partir de l'équation (1.3) et le lemme 1.12. On choisit finalement $\mathcal{P}_{-r,r}^R$ et $\mathcal{P}_{-s,s}^L$, afin que le produit matriciel $B_0^R M_0 B_0^L$ soit égal à I_{r+s} et ainsi la relation (1.6) reste valable même pour $n = 0$. \square

Les coefficients $A_{n,j}$ sont tous d'indices négatifs. Cela signifie que l'on a prolongé artificiellement les diagonales. Dans le cas des diagonales à coefficients constants, les coefficients sont complètement déterminés.

1.2.5 Approximants de Padé

Comme dans le cas scalaire [58, p. 43, §3], on peut considérer le problème suivant d'Hermite-Padé à l'infini [11, 12, §2], [71, §2] :

On se donne une matrice $\mathbb{F}(z)$ de taille $r \times s$ de fonctions

$$\mathbb{F}_{i,j}(z) = \sum_{k=0}^{\infty} \frac{f_{i,j}^{(k)}}{z^{k+1}}, \quad i = 1, \dots, r, \quad j = 1, \dots, s$$

chaque fonction étant une série formelle à coefficients complexes. Soient deux multi-indices $\bar{m} = (m_1, \dots, m_r)$ et $\bar{n} = (n_1, \dots, n_s)$ d'ordre $|\bar{m}| = m_1 + \dots + m_r$ et $|\bar{n}| = n_1 + \dots + n_s$ tels que $|\bar{m}| = |\bar{n}| + \nu$.

On cherche donc des polynômes scalaires Q_1, \dots, Q_s non tous nuls de degré respectivement inférieur à $m_1 - 1, \dots, m_s - 1$ et tels que, pour des polynômes P_1, \dots, P_r , les relations suivantes soient satisfaites :

$$\begin{cases} Q_1 \mathbb{F}_{1,1} + \dots + Q_s \mathbb{F}_{1,s} - P_1 = \mathcal{O}(1/z^{n_1+1}) \\ \vdots \\ Q_1 \mathbb{F}_{r,1} + \dots + Q_s \mathbb{F}_{r,s} - P_r = \mathcal{O}(1/z^{n_r+1}) \end{cases} \quad (1.7)$$

Puisque le problème consiste à résoudre un système linéaire de $|\bar{n}|$ équations à $|\bar{m}| = |\bar{n}| + \nu$ inconnues que sont les coefficients des polynômes $Q_k, k = 1, \dots, s$, il existe toujours une solution non triviale. Les polynômes P_1, \dots, P_r sont la partie polynômiale des séries formelles $Q_1 \mathbb{F}_{k,1} + \dots + Q_s \mathbb{F}_{k,s}, k = 1, \dots, r$. Il est évident, qu'en général, la solution n'est pas unique même à une multiplication près par une constante.

On a étudié l'approximant en l'infini, on pourrait l'étudier en zéro. Dans ce cas, un algorithme de résolution du problème est donné dans [12].

On peut écrire le problème à résoudre sous une forme matricielle en posant $\mathbb{Q} = (Q_1, \dots, Q_s)^t$ et $\mathbb{P} = (P_1, \dots, P_r)^t$ on obtient

$$\mathbb{F}(z)\mathbb{Q}(z) - \mathbb{P}(z) = \mathcal{O}(1/z^{\bar{n}+1})$$

$1/z^{\bar{n}+1}$ étant défini par les relations de (1.7).

Le problème à résoudre est équivalent à la détermination des polynômes $\mathbb{Q} = (Q_1, \dots, Q_s)^t$ vérifiant les relations d'orthogonalité

$$\langle \text{diag}(x^{k_1}, \dots, x^{k_r}), \mathbb{Q}(x) \rangle = 0, \quad 0 \leq k_j \leq n_j - 1, \quad j = 1 \dots r \quad (1.8)$$

Les approximants de Padé sont fortement reliés aux fractions continues. En effet, en définissant l'inverse d'une matrice rectangulaire quelconque comme décrit par J. Van Iseghem et V. Sorokin dans [72], on peut créer une fraction continue à coefficients matriciels à partir de $\mathbb{F}(z)$. Le n -ième convergent, c'est-à-dire la troncature de la fraction continue infinie est un approximant de Padé de $\mathbb{F}(z)$. On retrouve ainsi le cas scalaire [58, Chapter 2, §4]. Les coefficients de la matrice A , c'est-à-dire les coefficients de la récurrence à $r + s + 1$ termes sont ainsi récupérés.

On définit une famille particulière de multi-indices pour répondre à ce problème d'approximation rationnelle d'une série formelle à coefficients matriciels.

Définition 1.22 ([58, 47, 71]). *Soit un multi-indice $\bar{k} = (k_1, \dots, k_r) \in \mathbb{Z}^r$. On dit que \bar{k} est régulier ([47, 71]) ou propre ([58]), si*

$$k_1 \geq k_2 \geq \dots \geq k_r \geq k_1 - 1$$

L'ordre $|\bar{n}|$ d'un multi-indice permet de définir de manière unique un multi-indice régulier : si $|\bar{n}| = rq + r_0$, $r_0 < r$ alors $\bar{n} = (n_i)_{i=1, \dots, r}$ est tel que $n_1 = \dots = n_{r_0} = q + 1$ et $n_{r_0+1} = \dots = n_r = q$.

Désormais, on considère que les indices des polynômes vectoriels orthogonaux sont réguliers. On ne s'intéresse donc qu'aux approximants diagonaux, c'est-à-dire $\nu = 1$.

Dans notre cas, il s'agit de déterminer l'approximant de Padé de la série génératrice formelle

$$\varphi(z) = \sum_{k=0}^{\infty} \langle I_r(x)x^k, I_s(x) \rangle z^{-k-1} = \sum_{k=0}^{\infty} \mathcal{M}_k z^{-k-1} := \left\langle I_r(x), \frac{1}{z-x} I_s(x) \right\rangle$$

sous la forme PQ^{-1} ou $Q^{-1}P$.

Sous l'hypothèse de régularité, ces approximants de Padé diagonaux existent [5, Corollaire 6].

Propriété 1.23. *La fraction rationnelle à valeurs matricielles $\mathcal{P}_{n,s}^R (\mathcal{Q}_{n,s}^R)^{-1} = (\mathcal{Q}_{n,r}^L)^{-1} \mathcal{P}_{n,r}^L$ est un approximant de Padé diagonal autour de l'infini de $\varphi(z)$.*

1.2.6 Cas particuliers des approximants de Padé précédents

1.2.6.i Systèmes parfaits

Dans le cas général précédent, si on pose $r = 1$, on obtient les approximants de Padé de type I, si $s = 1$, on obtient les approximants de Padé de type II ou approximants d'Hermite-Padé simultanés. Ces notions de polynômes orthogonaux multiples sont développées dans des articles de W. Van Assche [79], avec E. Coussement [80], d'A. Aptekarev [2], de V. Kaliaguine [47, 48] et du livre de E. M. Nikishin et V. N. Sorokin [58]. On donne quelques définitions sur les polynômes orthogonaux multiples ainsi que certaines applications de ces polynômes, en particulier en théorie des nombres avec des démonstrations de transcendance ou d'irrationalité. Ce paragraphe donne aussi un cadre théorique à deux exemples qui seront traités dans le chapitre 3, lors de l'étude de l'algorithme de Chebyshev modifié vectoriel.

Définition 1.24. *On appelle fonction de Markov la quantité définie par*

$$f(z) = \int \frac{d\mu(x)}{z - x}$$

Propriété 1.25. *Si le vecteur \vec{f} de fonctions est constitué de fonctions de Markov*

$$\left(\int \frac{d\mu_1(x)}{z - x} dx, \dots, \int \frac{d\mu_r(x)}{z - x} dx \right)$$

alors le dénominateur commun des approximants de Padé simultanés est un polynôme orthogonal de type II vérifiant les relations (1.8)

Contrairement aux polynômes orthogonaux habituels, le polynôme multiple orthogonal n'est pas unique, son degré n'est pas exact. Cependant, il existe deux cas de systèmes de mesures pour lesquels on retrouve certaines propriétés des polynômes orthogonaux habituels.

On considère donc r mesures $d\mu_1, \dots, d\mu_r$ sur l'axe réel. Le support de chaque mesure $d\mu_i$ est un sous-ensemble E_i . On considère que la forme bilinéaire associée peut être exprimée sous la forme d'une intégrale.

Propriété 1.26 (Système d'Angelesco). *Supposons que les supports E_1, \dots, E_r des mesures $d\mu_1, \dots, d\mu_r$ soient des intervalles disjoints. On dit alors que le système de mesures forme un système d'Angelesco.*

Le polynôme orthogonal multiple $P_{\vec{n}}$ relativement à ce système de mesures possède exactement n_i zéros simples dans E_i pour $i = 1, \dots, r$. De ce fait, il est unique à une constante multiplicative près. [2, Théorème 1]

Propriété 1.27 (Système de Nikishin). *Supposons que les supports E_1, \dots, E_r des mesures soient successivement disjoints $E_i \cap E_{i+1} = \emptyset$, $i = 1, \dots, r - 1$. Soit le système de mesures définies comme suit*

$$\left(\int_{E_2} \frac{d\mu_2(x_2)}{x - x_2} \right) d\mu_1(x) \\ \left(\int_{E_2} \int_{E_3} \left(\frac{d\mu_3(x_3)}{x_2 - x_3} \right) \frac{d\mu_2(x_2)}{x - x_2} \right) d\mu_1(x), \dots$$

On dit que le système forme un système de Nikishin.

Si le multi-indice vérifie $n_i \leq n_{i-1} + 1$, $i = 2, \dots, r$ alors le polynôme $P_{\bar{n}}$ relativement au système de Nikishin possède $|\bar{n}|$ zéros dans l'intérieur de E_1 . Et ainsi, de même que précédemment, il est défini de manière unique. [2, Théorème 2]

Remarque 1.28 (Applications). 1. *Les approximants d'Hermite-Padé ont permis à Hermite la démonstration de la transcendance de e [58, p. 129-130].*

2. *Un système particulier de Nikishin construit par V. N. Sorokin permet de retrouver les nombres d'Apery qui permirent de démontrer l'irrationalité de la fonction $\zeta(3)$ [2, Théorème 4].*

3. *Ils seront aussi utilisés pour illustrer le calcul des coefficients de récurrence dans deux exemples associés aux moments modifiés dans la partie 3.3.*

Comme indiqué précédemment, la solution au problème posé n'est pas unique. Pour essayer d'obtenir cette unicité, on doit disposer d'indices particuliers :

Dans le cas des approximants de Padé simultanés

Définition 1.29. *Un multi-indice \bar{n} est normal si tout polynôme Q_n possède le degré exact $n = |\bar{n}|$*

Définition 1.30 ([54, 58]). *Si tous les indices sont normaux alors le système de fonctions \mathbb{F} est dit parfait.*

Si les indices réguliers sont tous normaux alors le système de fonctions \mathbb{F} est dit faiblement parfait.

C'est-à-dire que l'on retrouve, énoncé différemment, le corollaire 1.6 : d'un point de vue algébrique, on dirait que la matrice H des moments est fortement régulière.

1.2.6.ii Cas scalaire

Si $r = s = 1$, on obtient le cas scalaire. La matrice regroupant les coefficients de récurrence porte le nom de matrice de Jacobi. Les valeurs propres de la matrice de Jacobi sont les zéros du polynôme scalaire Q qui est le dénominateur de l'approximant de Padé. Les vecteurs propres de la matrice de Jacobi fournissent les poids de la quadrature de Gauss associée à ce système de polynômes.

Ceci peut se généraliser au cas à $r + s + 1$ diagonales grâce au théorème suivant énoncé et démontré dans [5, Theorem 2]

Theorème 1.31. *det $\mathcal{Q}_{n,s}^R$ est un polynôme de degré n . De plus il existe des vecteurs lignes de polynômes $u_1, \dots, u_\ell \in \mathbb{C}^{1 \times s}[x]$, des nombres complexes $\lambda_1, \dots, \lambda_\ell$ et des entiers ρ_1, \dots, ρ_ℓ tels que*

$$\prod_{j=1}^{\ell} (x - \lambda_j)^{\rho_j} = \frac{1}{const} \cdot \det \mathcal{Q}_{n,s}^R(x)$$

est le polynôme caractéristique de A_n et

$$\frac{\partial^j}{\partial x^j} u_k(x) \cdot \mathcal{Q}_{0,n}^R(x)|_{x=\lambda_k}, \quad j = 0, \dots, \rho_\ell - 1, \quad k = 1, \dots, \ell,$$

sont les vecteurs propres linéairement indépendants et les vecteurs secondaires, à gauche.

On pourrait énoncer le même théorème pour $\mathcal{Q}_{n,r}^L$ avec des vecteurs propres et vecteurs secondaires à droite. En particulier, on obtient

$$\det \mathcal{Q}_{n,s}^R = C \cdot \det \mathcal{Q}_{n,r}^L$$

1.3 Analyse fonctionnelle

1.3.1 Opérateur

On rappelle que l'on note ℓ^2 l'ensemble des suites de nombres complexes de carré sommable :

$$\ell^2 = \{(u_n)_n \in \mathbb{C}^{\mathbb{N}}, \sum_{k=0}^{+\infty} |u_k|^2 < +\infty\}.$$

Cet espace, muni du produit scalaire,

$$\forall (x, y) \in \ell^2 \times \ell^2, (x|y) = \sum_{k=0}^{+\infty} x_k \overline{y_k}.$$

est un espace de Hilbert pour la norme induite. Une base orthonormale associée (e_0, e_1, \dots) est définie par

$$\forall (i, k) \in \mathbb{N}^2, e_{k,i} = \delta_{k,i}$$

où $\delta_{k,i}$ est le symbole de Kronecker.

L'ensemble des applications linéaires continues de ℓ^2 est noté $\mathcal{L}(\ell^2)$. Cet ensemble est aussi un espace de Banach pour la norme induite par la norme de l'espace ℓ^2 .

On se donne une matrice A semi-infinie possédant r sur-diagonales et s sous-diagonales. On associe un opérateur à la matrice A de la manière suivante [1, § 29],[8, §1],[58, p. 75] :

On peut calculer le produit de cette matrice avec des combinaisons linéaires finies d'éléments de la base de ℓ^2 . On peut donc définir un opérateur associé à la matrice A sur l'ensemble des combinaisons linéaires finies. La fermeture de cet opérateur définit un opérateur aux différences \mathcal{A} . Dans le cas scalaire $r = s = 1$, symétrique, réel, l'opérateur est dit opérateur de Jacobi.

Ici, on suppose que les coefficients des $r + s + 1$ diagonales sont uniformément bornés. Alors l'opérateur \mathcal{A} est défini dans $\mathcal{L}(\ell^2)$ et est lui-même borné.

1.3.2 Résolvante

On donne, dans ce paragraphe, des résultats tout-à-fait classiques d'analyse fonctionnelle que l'on peut trouver, par exemple, dans [43, p. 164 et sq.], [65], [66]. Ceci permet de donner un cadre théorique afin d'introduire, finalement, la fonction de Weyl dans le cas matriciel général.

En développant cette fonction en fraction continue, comme indiqué dans le paragraphe précédent sur les approximants de Padé, grâce aux techniques d'inversion de matrice développées par V. Sorokin et J. Van Iseghem dans [72], il est possible de retrouver les coefficients de la matrice associée à l'opérateur. La fonction de Weyl peut être approchée par les approximants de Padé, troncature de la fraction continue matricielle [81]. On retrouve, ainsi, les résultats des cas scalaire et vectoriel [58].

Définition 1.32. Soit $T \in \mathcal{L}(\ell^2)$. On appelle ensemble résolvant $\Omega(T)$, l'ensemble constitué des valeurs λ pour lesquels $(\lambda I - T)$ est inversible dans $\mathcal{L}(\ell^2)$. On note $\sigma(T)$ le complémentaire de $\Omega(T)$ dans \mathbb{C} . Cet ensemble est le spectre de T . $\sigma(T)$ contient les valeurs propres de T .

$\Omega(T)$ est un ouvert du plan complexe et $\sigma(T)$ est fermé.

Dans le cas de la matrice de Jacobi, le spectre de l'opérateur est égal au support de la mesure d'orthogonalité. Ce résultat n'est pas vrai dans le cas vectoriel [79, § 2.4].

Définition 1.33. *L'application $\lambda \mapsto \mathcal{R}(\lambda) = (\lambda I - T)^{-1}$ de \mathbb{C} dans $\mathcal{L}(\ell^2)$ est appelée la résolvante de T . La résolvante est analytique sur $\Omega(T)$.*

On donne ensuite la définition de la fonction de Weyl[14]. Cette fonction permettra de donner une caractérisation de l'ensemble résolvant.

Définition 1.34 (Fonction de Weyl). *Soit \mathcal{A} un opérateur sur ℓ^2 représenté par une matrice A semi-infinie possédant r surdiagonales et s sous-diagonales. On appelle fonction de Weyl [14], la fonction définie sur $\Omega(A)$ par*

$$\Phi(z) = (e_i, (zI - A)^{-1} e_j)_{\substack{j=0, \dots, s-1 \\ i=0, \dots, r-1}}$$

Cette quantité permet de retrouver les données composant la matrice A (voir par exemple [48], [73]).

Dans le cas d'un opérateur symétrique borné, on sait, grâce au théorème spectral que la fonction de Weyl peut être représentée par une fonction de Markov :

$$\Phi(z) = \int_{\mathbb{R}} \frac{d\mu(x)}{z - x}$$

la mesure, appelée mesure spectrale de l'opérateur \mathcal{A} a pour support le spectre de celui-ci.

Définition 1.35. *On appelle résidus les termes*

$$R_k^R = \varphi Q_k^R - P_k^R$$

soit, en regroupant,

$$\mathcal{R}_{n,s}^R = \varphi \mathcal{Q}_{n,s}^R - \mathcal{P}_{n,s}^R \text{ ou } \mathcal{R}_{n-r,r}^R = \varphi \mathcal{Q}_{n-r,r}^R - \mathcal{P}_{n-r,r}^R$$

pour les polynômes à droite, et

$$R_k^L = Q_k^L \varphi - P_k^L$$

soit, en regroupant,

$$\mathcal{R}_{n,r}^L = \mathcal{Q}_{n,r}^L \varphi - \mathcal{P}_{n,r}^L \text{ ou } \mathcal{R}_{n-s,s}^L = \mathcal{Q}_{n-s,s}^L \varphi - \mathcal{P}_{n-s,s}^L$$

pour les polynômes à gauche.

Etant donnés les suites \mathcal{P}^L , \mathcal{P}^R , \mathcal{Q}^L et \mathcal{Q}^R en un z donné, ainsi que la valeur $\varphi(z)$, il est possible de reconstruire l'opérateur résolvant $\mathcal{R}(z)$ grâce au théorème suivant [5, Théorème 4] ou dans la démonstration de [13, Théorème 1].

Théorème 1.36. *Pour $z \in \Omega(A)$, on a $\mathcal{R}(z) = (zI - A)^{-1} = (v_{k,n}(z))_{k=0,1,\dots}^{n=0,1,\dots}$ avec*

$$v_{k,n}(z) = \begin{cases} Q_k^L(z)R_n^R(z) & \text{si } 0 \leq k < n + r, \\ R_k^L(z)Q_n^R(z) & \text{si } 0 \leq n < k + s. \end{cases}$$

La fonction de Weyl $\Phi(z)$ est liée à la fonction génératrice $\varphi(z)$ par la relation suivante :

Propriété 1.37.

$$(\mathcal{Q}_{0,r}^L)^{-1} \Phi(z) (\mathcal{Q}_{0,s}^R)^{-1} = \varphi(z) \quad (1.9)$$

Démonstration. D'après le théorème 1.36, pour tout $0 \leq j \leq r$ et $0 \leq k \leq s$, on a

$$(e_j, (zI - A)^{-1} e_k) = R_j^L(z) \cdot Q_k^R(z)$$

Soit en regroupant les polynômes

$$\Phi(z) = \mathcal{R}_{0,r}^L(z) \cdot \mathcal{Q}_{0,s}^R(z)$$

Or, d'après la définition des polynômes associés, on a

$$\begin{aligned} P_j^L(z) &= \left\langle \frac{Q_j^L(x) - Q_j^L(z)}{x - z}, I_s \right\rangle \\ &= \left\langle \frac{Q_j^L(x)}{x - z}, I_s \right\rangle - Q_j^L(z) \left\langle \frac{1}{x - z} I_r, I_s \right\rangle \\ &= \left\langle \frac{Q_j^L(x)}{x - z}, I_s \right\rangle + Q_j^L(z) \varphi(z) \end{aligned}$$

Pour les résidus, on obtient donc l'égalité formelle suivante

$$R_j^L(z) = \left\langle \frac{Q_j^L(x)}{z - x}, I_s \right\rangle$$

De plus, pour tout polynôme vectoriel $S^R(z) \in \mathbb{C}^{1 \times s}[z]$ de degré diagonal $ddeg S^R \leq j + s - 1$, on a

$$\begin{aligned} 0 &= \left\langle Q_j^L(x), \frac{S^R(z) - S^R(x)}{z - x} \right\rangle \text{ car la fraction est un polynôme en } x \text{ de degré inférieur à } j \\ &= \left\langle Q_j^L(x), \frac{1}{z - x} I_s \right\rangle S^R(z) - \left\langle Q_j^L(x), \frac{S^R(x)}{z - x} \right\rangle \end{aligned}$$

Ce qui nous donne

$$R_j^L(z)S^R(z) = \left\langle Q_j^L(x), \frac{1}{z-x}S^R(x) \right\rangle$$

On choisit $S^R(z) = Q_k^R(z)$, $k \in \{0, \dots, s\}$. En regroupant les polynômes, on obtient alors

$$\mathcal{R}_{0,r}^L(z)\mathcal{Q}_{0,s}^R(z) = \left\langle \mathcal{Q}_{0,r}^L(x)\frac{1}{z-x}, \mathcal{Q}_{0,s}^R(x) \right\rangle$$

Or les polynômes $\mathcal{Q}_{0,r}^L$ et $\mathcal{Q}_{0,s}^R$ ne dépendent pas de x , on peut donc les sortir de la forme bilinéaire. Ce qui prouve l'égalité demandée. \square

D'après la relation (1.9), les approximants de Padé de la fonction de Weyl sont les mêmes que ceux de la série génératrice $\varphi(z)$ à condition de modifier l'initialisation donnée dans le corollaire 1.21. Il suffit de poser $\tilde{\mathcal{Q}}_{0,r}^L = I_r$ et $\tilde{\mathcal{Q}}_{0,s}^R = I_s$ pour obtenir les approximants de Padé de la fonction de Weyl. Les approximants de Padé définis précédemment convergent vers la série génératrice ou la fonction de Weyl localement uniformément comme l'indiquent le théorème démontré par B. Beckermann dans [5, Théorème 10]

Théorème 1.38. *Soit un complexe z_0 et $K(z_0) = \{z \in \hat{\mathbb{C}} : |z - z_0| > \|z_0 I - A\|\}$. La suite d'approximants de Padé $((\mathcal{Q}_{n,r}^L)^{-1}\mathcal{P}_{n,r}^L)_{n \geq 0}$ converge vers la série génératrice $\varphi(z)$ localement uniformément sur $K(z_0)$. En particulier, aucune des fonctions rationnelles matricielles $(\mathcal{Q}_{n,r}^L)^{-1}\mathcal{P}_{n,r}^L$ n'a de pôle dans $K(z_0)$.*

ou le théorème [13, Théorème 3] démontré par B. Beckermann et A. Osipov

Théorème 1.39. *Soit A un opérateur borné. La suite d'approximants de Padé $((\mathcal{Q}_{n,r}^L)^{-1}\mathcal{P}_{n,r}^L)_{n \geq 0}$ converge en capacité vers la fonction de Weyl dans tout sous ensemble compact de la composante non bornée de l'ensemble résolvant de A . De plus, la convergence est uniforme dans tout sous ensemble compact du complémentaire de l'image numérique de A .*

Dans le chapitre 2 sur les transformations de Moebius, on accélère cette convergence dans la section 2.3.

1.3.3 Caractérisation de l'ensemble résolvant

Les polynômes vectoriels orthogonaux permettent de caractériser l'ensemble résolvant grâce à la fonction de Weyl.

On utilisera, pour cela, un théorème [5, Théorème 6] donné par B. Beckermann ou une forme affaiblie du théorème [13, Théorème 1] puisque la matrice

est bornée. Ce théorème sera utilisé dans la démonstration du théorème principal 1.44.

Théorème 1.40. *Soit $z \in \mathbb{C}$. Soient P_k^R, P_k^L, Q_k^R et Q_k^L les polynômes définis par le théorème 1.12 et le lemme 1.9 pour les initialisations.*

Alors $z \in \Omega(A)$ si et seulement si on peut trouver une matrice $\Theta \in \mathbb{C}^{r \times s}$, $\beta \in]1; +\infty[$ et $q \in]0; 1[$ qui dépendent de z tels que

$$\forall 0 \leq k \leq n : \quad |Q_k^L(z)\{\Theta Q_n^R(z) - P_n^R(z)\}| \leq \beta q^{n-k}$$

et

$$\forall 0 \leq n \leq k : \quad |\{Q_k^L(z)\Theta - P_k^L(z)\}Q_n^R(z)| \leq \beta q^{k-n}$$

Remarque 1.41. *La matrice Θ définie précédemment est une valeur de la fonction $\varphi(z)$ (voir la démonstration du théorème 6 dans [5] ou la remarque donnée après le corollaire 1 dans [13]).*

1.4 Résultats pour les matrices constantes

Dans cette partie, en utilisant la fonction génératrice ou la fonction de Weyl définie ci-dessus, on se propose de caractériser l'ensemble résolvant d'un opérateur à $r + s + 1$ diagonales constantes. La démonstration de cette caractérisation utilisera le théorème 1.40 rappelé précédemment. On commencera par déterminer une formule de récurrence pour B_n^R et B_n^L à partir de B_0^R et B_0^L que l'on aura préalablement calculés. A partir de là, en définissant une matrice A^L associée à l'opérateur, on pourra caractériser l'ensemble résolvant de l'opérateur grâce au théorème principal 1.44 : Le nombre complexe z appartient à l'ensemble résolvant si les valeurs propres de la matrice A^L sont scindées en deux groupes : s valeurs propres sont de module strictement inférieur à 1 et les r valeurs propres restantes sont de module strictement supérieur à 1.

Ce théorème n'est en aucun cas une nouveauté. En effet, un énoncé analogue a déjà été donné par P. L. Duren dans [29]. La spécificité réside dans la démonstration. Elle est utile, évidemment pour ce théorème mais aussi pour le théorème 2.8 sur l'accélération de convergence dans le chapitre 2.

Nous supposons dans cette partie que la matrice A est constante par diagonales. On notera $h_j, j = -s, \dots, r$ les composantes des diagonales. On définit alors les matrices de récurrence A^L et A^R de la manière suivante :

Définition 1.42.

$$B_n^L = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ \frac{-h_{-s}}{h_r} & \cdots & \frac{z-h_0}{h_r} & \cdots & \frac{-h_{r-1}}{h_r} \end{pmatrix} \cdot B_{n-1}^L = A^L \cdot B_{n-1}^L$$

et

$$B_n^R = B_{n-1}^R \cdot \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & \frac{-h_r}{h_{-s}} \\ 1 & \ddots & & & \vdots & \vdots \\ 0 & \ddots & \ddots & & \frac{z-h_0}{h_{-s}} & \vdots \\ \vdots & \ddots & & & 0 & \vdots \\ 0 & \cdots & \cdots & 0 & 1 & \frac{-h_{-s+1}}{h_{-s}} \end{pmatrix} = B_{n-1}^R \cdot A^R$$

On cherche à étudier les limites de ces polynômes orthogonaux donc des B_n^L et B_n^R . On recherche donc les valeurs propres et les vecteurs propres des matrices de récurrence.

Dans le cas à diagonales constantes, on notera $\forall n, M_n = M$ où

Propriété 1.43. $A^R \cdot M \cdot A^L = M$

et les valeurs propres de A^L sont les inverses des valeurs propres de A^R

Démonstration. En utilisant la relation 1.6, B_n^R et B_n^L sont inversibles. Dans le cas à diagonales constantes, les relations de récurrence

$$B_n^R = B_{n-1}^R A^R \text{ et } B_n^L = A^L B_{n-1}^L$$

permettent d'obtenir

$$B_{n-1}^R A^R M A^L B_{n-1}^L = I_{r+s} = B_{n-1}^R M B_{n-1}^L$$

L'égalité $A^R M A^L = M$ est donc prouvée.

Le déterminant de M vaut, au signe près, $h_{-s}^s \cdot h_r^r$ ce qui est non nul par définition. Donc M est inversible, de même que A^R et A^L . Alors $M^{-1} \cdot A^R \cdot M = (A^L)^{-1}$ et ainsi les valeurs propres de A^R sont celles de $(A^L)^{-1}$ donc les inverses des valeurs propres de A^L . \square

D'après ce qui précède, il suffit d'étudier la matrice A^L et les résultats s'appliqueront à la matrice A^R .

La matrice A^L est une matrice dite compagnon. Elle est donc semblable à une matrice diagonale par blocs J_i

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda_i & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{pmatrix} \in \mathcal{M}_{m_i}(\mathbb{C})$$

où m_i est la multiplicité de la valeur propre λ_i .

Si les valeurs propres sont distinctes, la matrice est diagonalisable. Les vecteurs propres associés sont $(1, \lambda_i, \dots, \lambda_i^{r+s-1})^t$.

On obtient donc la matrice de passage P :

$$P = \begin{pmatrix} 1 & \cdots & \cdots & 1 \\ \lambda_1 & \cdots & \cdots & \lambda_{r+s} \\ \vdots & & & \vdots \\ \lambda_1^{r+s-1} & \cdots & \cdots & \lambda_{r+s}^{r+s-1} \end{pmatrix}$$

Ceci nous fournit alors l'écriture suivante des matrices B_n^L et B_n^R :

$$B_n^L = P \cdot (D^L)^n \cdot P^{-1} \cdot B_0^L \quad \text{et} \quad B_n^R = B_0^R \cdot (MP) \cdot (D^R)^n \cdot (MP)^{-1}$$

car les matrices de récurrence A^R et A^L sont liées par la formule $A^R = M \cdot (A^L)^{-1} \cdot M^{-1}$ et la matrice diagonale $D^R = (D^L)^{-1}$.

Si certaines valeurs propres possèdent une multiplicité différente de 1 les modifications portent sur la puissance de la matrice, qui possède une sur-diagonale, et sur la matrice de passage. Les vecteurs propres associés restent les mêmes que précédemment mais pour créer une base de l'espace propre associé, on a besoin des vecteurs secondaires. C'est-à-dire que si l'on note $x_i^{(1)}$, le vecteur propre, alors le premier vecteur secondaire (s'il existe) est $x_i^{(2)}$ tel que $(A - \lambda_i \cdot I)x_i^{(2)} = x_i^{(1)}$.

$x_i^{(2)}$ peut être trouvé en dérivant le vecteur propre par rapport à λ_i . Et on recommence jusqu'à la multiplicité de la valeur propre.

On déduit alors l'écriture de la matrice de passage

$$\begin{aligned} P &= (x_1^{(1)}, x_1^{(2)}, \dots, x_1^{(m_1)}, \dots, x_p^{(1)}, \dots, x_p^{(m_p)}) & (1.10) \\ \text{avec } x_i^{(1)} &= (\lambda_i^k)_{k \in \{0, \dots, r+s-1\}}, \\ x_i^{(j)} &= (C_k^{j-1} \lambda^{k-(j-1)})_{k \in \{j-1, \dots, r+s-1\}}. \end{aligned}$$

On remarque que lorsque l'on élève à la puissance n la matrice diagonale par blocs, on obtient une matrice diagonale par blocs où chaque bloc est J_i^n . Pour le calcul explicite de cette matrice, il suffit de dire que $J_i = \lambda_i I_{m_i} + N(m_i)$ avec $N(m_i)$, matrice nilpotente d'ordre m_i . La formule du binôme de Newton permet de déduire la forme de J_i^n puisque $N(m_i)$ commute évidemment avec $\lambda_i I_{m_i}$.

Théorème 1.44 (Calcul de φ). $z \in \Omega(A)$ si et seulement si la matrice A^L défini précédemment possède exactement s valeurs propres de module strictement inférieur à 1 et r valeurs propres de module strictement supérieur à 1.

Auquel cas, on considère la matrice de passage P , définie précédemment, décomposée en blocs et on appelle N l'inverse de P . Notons

$$P = \begin{pmatrix} P_{1,1}(\in \mathbb{C}^{s \times s}) & P_{1,2}(\in \mathbb{C}^{s \times r}) \\ P_{2,1}(\in \mathbb{C}^{r \times s}) & P_{2,2}(\in \mathbb{C}^{r \times r}) \end{pmatrix} \text{ et } N = \begin{pmatrix} N_{1,1}(\in \mathbb{C}^{s \times s}) & N_{1,2}(\in \mathbb{C}^{s \times r}) \\ N_{2,1}(\in \mathbb{C}^{r \times s}) & N_{2,2}(\in \mathbb{C}^{r \times r}) \end{pmatrix}$$

On a alors

$$\varphi = (\mathcal{Q}_{0,r}^L)^{-1} P_{2,1} P_{1,1}^{-1} \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix}^{-1} (\mathcal{Q}_{0,s}^R)^{-1}$$

Il est à noter que l'égalité obtenue pour φ permet d'obtenir une égalité similaire pour Φ grâce à la relation (1.9)

$$\Phi = P_{2,1} P_{1,1}^{-1} \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix}^{-1}$$

Comme on le verra plus loin dans la démonstration avec l'égalité (1.11), tous les calculs pourraient être faits avec la fonction de Weyl $\Phi(z)$ au lieu de la fonction génératrice $\varphi(z)$.

Démonstration. Le but est d'obtenir φ afin d'éliminer les valeurs propres "gênantes". En fait, il faut que le résidu tende vers zéro et que l'on puisse vérifier le Théorème 1.40.

1. Cas des valeurs propres distinctes

On multiplie B_n^R à gauche par $(-I_r \ \varphi)$. En utilisant la définition, on obtient d'une part $(-\mathcal{P}_{n-r,r}^R + \varphi \mathcal{Q}_{n-r,r}^R; -\mathcal{P}_{n,s}^R + \varphi \mathcal{Q}_{n,s}^R)$ et d'autre part, en utilisant les vecteurs propres et valeurs propres,

$$(-I_r \ \varphi) B_0^R M = \left(\varphi \mathcal{Q}_{0,s}^R \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix} ; -(\mathcal{Q}_{0,r}^L)^{-1} \right)$$

On multiplie ensuite ce résultat à droite par $P(D^R)^n$ et on obtient

$$\varphi \mathcal{Q}_{0,s}^R \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix} \begin{pmatrix} \lambda_1^{-n} & \dots & \lambda_{r+s}^{-n} \\ \vdots & & \vdots \\ \lambda_1^{-n+s-1} & \dots & \lambda_{r+s}^{-n+s-1} \end{pmatrix} - (\mathcal{Q}_{0,r}^L)^{-1} \begin{pmatrix} \lambda_1^{-n+s} & \dots & \lambda_{r+s}^{-n+s} \\ \vdots & & \vdots \\ \lambda_1^{-n+r+s-1} & \dots & \lambda_{r+s}^{-n+r+s-1} \end{pmatrix}$$

Supposons que les valeurs propres se séparent en 2 sous-ensembles : a valeurs propres de module inférieur ou égal à 1 et b supérieur strictement à 1. On veut que les a premières valeurs propres disparaissent avec φ car leur module est inférieur à 1. En effet, la puissance qui leur est affectée est négative. Ainsi, les puissances de ces valeurs propres ne tendent pas vers zéro. Les résidus ne tendraient pas non plus vers zéro.

On obtient alors

$$\varphi \mathcal{Q}_{0,s}^R \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ \lambda_1^{s-1} & \dots & \lambda_a^{s-1} \end{pmatrix} = (\mathcal{Q}_{0,r}^L)^{-1} \begin{pmatrix} \lambda_1^s & \dots & \lambda_a^s \\ \vdots & & \vdots \\ \lambda_1^{r+s-1} & \dots & \lambda_a^{r+s-1} \end{pmatrix}$$

On obtient donc r systèmes de a équations à s inconnues. Ces systèmes admettent des solutions si $a \leq s$ et donc puisque $a + b = r + s$, $b \geq r$. Supposons que $a = c + d$ où c indique le nombre de valeurs propres dont le module est strictement inférieur à 1 et d le nombre de valeurs propres de module égal à 1.

On multiplie de la même manière B_n^L à droite par $\begin{pmatrix} \varphi \\ I_s \end{pmatrix}$.

On obtient

$$P(D^L)^n \begin{pmatrix} N_{1,1} \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix}^{-1} (\mathcal{Q}_{0,s}^R)^{-1} + N_{1,2} \mathcal{Q}_{0,r}^L \varphi \\ N_{2,1} \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix}^{-1} (\mathcal{Q}_{0,s}^R)^{-1} + N_{2,2} \mathcal{Q}_{0,r}^L \varphi \end{pmatrix}$$

où

$$P^{-1} := N := \begin{pmatrix} N_{1,1} (\in \mathbb{C}^{c \times s}) & N_{1,2} (\in \mathbb{C}^{c \times r}) \\ N_{2,1} (\in \mathbb{C}^{(d+b) \times s}) & N_{2,2} (\in \mathbb{C}^{(d+b) \times r}) \end{pmatrix}$$

On veut donc éliminer les $d + b$ valeurs propres de module supérieur ou égal à 1. On obtient donc

$$N_{2,2} \mathcal{Q}_{0,r}^L \varphi = -N_{2,1} \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix}^{-1} (\mathcal{Q}_{0,s}^R)^{-1}$$

c'est-à-dire s systèmes de $d + b$ équations à r inconnues. Ces systèmes admettent des solutions si $d + b \leq r$.

On obtient alors $r \geq d + b \geq d + r$ soit $d \leq 0$. Donc $d = 0$. Ce qui signifie que $b = r$ et $a = s$.

En particulier, on obtient

$$\varphi \mathcal{Q}_{0,s}^R \begin{pmatrix} h_{-s} & \dots & h_{-1} \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{-s} \end{pmatrix} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ \lambda_1^{s-1} & \dots & \lambda_s^{s-1} \end{pmatrix} = (\mathcal{Q}_{0,r}^L)^{-1} \begin{pmatrix} \lambda_1^s & \dots & \lambda_s^s \\ \vdots & & \vdots \\ \lambda_1^{r+s-1} & \dots & \lambda_s^{r+s-1} \end{pmatrix}$$

Cette forme de φ peut se définir en utilisant une partition de la matrice de passage P en blocs. On pose

$$P := \begin{pmatrix} P_{1,1}(\in \mathbb{C}^{s \times s}) & P_{1,2}(\in \mathbb{C}^{s \times r}) \\ P_{2,1}(\in \mathbb{C}^{r \times s}) & P_{2,2}(\in \mathbb{C}^{r \times r}) \end{pmatrix} \text{ et } N = \begin{pmatrix} N_{1,1}(\in \mathbb{C}^{s \times s}) & N_{1,2}(\in \mathbb{C}^{s \times r}) \\ N_{2,1}(\in \mathbb{C}^{r \times s}) & N_{2,2}(\in \mathbb{C}^{r \times r}) \end{pmatrix}$$

Cela permet de réécrire $(-I_r \ \varphi) B_0^R M P (D^R)^n$ sous la forme

$$\begin{pmatrix} 0_{r \times s} & (\mathcal{Q}_{0,r}^L)^{-1} (P_{2,1} P_{1,1}^{-1} P_{1,2} - P_{2,2}) \end{pmatrix} \begin{pmatrix} \lambda_1^{-n} & & 0 \\ & \ddots & \\ 0 & & \lambda_{r+s}^{-n} \end{pmatrix}$$

On peut remarquer que $-N_{2,2}$ est l'inverse de $P_{2,1} P_{1,1}^{-1} P_{1,2} - P_{2,2}$.

On veut maintenant appliquer le théorème 1.40.

Effectuons donc le produit $B_m^L \begin{pmatrix} I_r \\ 0_{s \times r} \end{pmatrix} (-I_r \ \varphi) B_n^R$.

D'une part, en utilisant les définitions de B_m^L et B_n^R , on obtient

$$\begin{pmatrix} \mathcal{Q}_{m-s,s}^L (-\mathcal{P}_{n-r,r}^R + \varphi \mathcal{Q}_{n-r,r}^R) & \mathcal{Q}_{m-s,s}^L (-\mathcal{P}_{n,s}^R + \varphi \mathcal{Q}_{n,s}^R) \\ \mathcal{Q}_{m,r}^L (-\mathcal{P}_{n-r,r}^R + \varphi \mathcal{Q}_{n-r,r}^R) & \mathcal{Q}_{m,r}^L (-\mathcal{P}_{n,s}^R + \varphi \mathcal{Q}_{n,s}^R) \end{pmatrix}$$

D'autre part, en utilisant les formules obtenues, on obtient

$$\begin{aligned}
& B_m^L \begin{pmatrix} I_r \\ 0_{s \times r} \end{pmatrix} \begin{pmatrix} -I_r & \varphi \end{pmatrix} B_n^R \\
&= P(D^L)^m P^{-1} \begin{pmatrix} 0_{s \times r} \\ \mathcal{Q}_{0,r}^L \end{pmatrix} \begin{pmatrix} 0_{r \times s} & (\mathcal{Q}_{0,r}^L)^{-1}(P_{2,1}P_{1,1}^{-1}P_{1,2} - P_{2,2}) \end{pmatrix} \begin{pmatrix} \lambda_1^{-n} & & 0 \\ & \ddots & \\ 0 & & \lambda_{r+s}^{-n} \end{pmatrix} (MP)^{-1} \\
&= P(D^L)^m \begin{pmatrix} N_{1,2} \\ N_{2,2} \end{pmatrix} \begin{pmatrix} 0_{r \times s} & P_{2,1}P_{1,1}^{-1}P_{1,2} - P_{2,2} \end{pmatrix} \begin{pmatrix} \lambda_1^{-n} & & 0 \\ & \ddots & \\ 0 & & \lambda_{r+s}^{-n} \end{pmatrix} (MP)^{-1} \\
&= P(D^L)^m \begin{pmatrix} 0_{s \times s} & P_{1,1}^{-1}P_{1,2} \\ 0_{r \times s} & -I_r \end{pmatrix} \begin{pmatrix} \lambda_1^{-n} & & 0 \\ & \ddots & \\ 0 & & \lambda_{r+s}^{-n} \end{pmatrix} (MP)^{-1}
\end{aligned} \tag{1.11}$$

On voit dans ce calcul que les valeurs d'initialisation se simplifient et n'interviennent plus dans la fin de la démonstration comme indiqué à la suite du théorème 1.44. puis un calcul par bloc donne :

$$= P \begin{pmatrix} 0_{s \times s} \begin{pmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_s^m \end{pmatrix} P_{1,1}^{-1}P_{1,2} \begin{pmatrix} \lambda_{s+1}^{-n} & & 0 \\ & \ddots & \\ 0 & & \lambda_{r+s}^{-n} \end{pmatrix} \\ 0_{r \times s} - \begin{pmatrix} \lambda_{s+1}^{m-n} & & 0 \\ & \ddots & \\ 0 & & \lambda_{r+s}^{m-n} \end{pmatrix} \end{pmatrix} (MP)^{-1} \tag{1.12}$$

Le théorème 1.40 est alors vérifié.

2. Cas des valeurs propres de multiplicité supérieure à 1

Ce cas se traite de la même manière que précédemment. Il faut quand même expliciter certaines modifications.

Tout d'abord, comme la matrice diagonale T^L semblable à A^L est constituée de blocs de Jordan, se pose le problème de la forme de T^R . Soit on considère qu'il s'agit tout simplement de l'inverse de T^L , auquel cas,

$$T^R = \begin{pmatrix} J_1^R & & \\ & \ddots & \\ & & J_p^R \end{pmatrix} \text{ et } J_i^R = \begin{pmatrix} \frac{1}{\lambda_i} & -1 & \cdots & \frac{(-1)^{m_i-1}}{\lambda_i^{m_i}} \\ \lambda_i & \lambda_i^2 & \cdots & \vdots \\ & \ddots & \ddots & -1 \\ & & \ddots & \frac{-1}{\lambda_i^2} \\ 0 & & & \frac{1}{\lambda_i} \end{pmatrix}$$

Soit, on considère, de la même façon que pour T^L , que T^R est composée de blocs de Jordan de la même taille mais de valeurs propres inverses. Dans ce cas, on doit calculer la matrice de passage correspondante. Ceci n'était pas nécessaire dans le cas précédent puisque elle était donnée par la relation $A^R = M(A^L)^{-1}M^{-1}$.

On a préféré conserver la première manière de voir afin de bénéficier le plus longtemps possible de l'analogie avec le cas des valeurs propres distinctes.

On obtient facilement la puissance m-ième de J_i^R :

$$(J_i^R)^m = \begin{pmatrix} \frac{1}{\lambda_i^m} & \frac{-C_m^{m-1}}{\lambda_i^{m+1}} & \cdots & \frac{(-1)^{m_i-1}C_{m+m_i-2}^{m-1}}{\lambda_i^{m+m_i-1}} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \frac{-C_m^{m-1}}{\lambda_i^{m+1}} \\ 0 & & & \frac{1}{\lambda_i^m} \end{pmatrix}$$

A partir de là, on peut faire le même raisonnement que précédemment. C'est-à-dire que l'on élimine les valeurs propres gênantes grâce à φ . Cette fois-ci, il faut que la somme des multiplicités des valeurs propres de module inférieur à 1 soit égale à s et on obtient une condition similaire pour r . Soit donc $\{(\lambda_1, m_1), \dots, (\lambda_\nu, m_\nu)\}$ les valeurs propres de module inférieur à 1.

Pour calculer comme en (1.12), il faut regarder si le bloc $P_{1,1}$ de la matrice de passage P , qui est une matrice de Van der Monde généralisée (voir (1.10)), est toujours inversible. En fait, il s'agit d'un problème d'interpolation généralisé d'Hermite. Le déterminant du bloc $P_{1,1}$ est donc non nul.

$$\begin{aligned} & B_m^L \begin{pmatrix} I_r \\ 0_{s \times r} \end{pmatrix} \begin{pmatrix} -I_r & \varphi \end{pmatrix} B_n^R \\ & = P \begin{pmatrix} 0_{s \times s} & \begin{pmatrix} (J_1^L)^m & & 0 \\ & \ddots & \\ 0 & & (J_\nu^L)^m \end{pmatrix} P_{1,1}^{-1} P_{1,2} \begin{pmatrix} (J_{\nu+1}^R)^{-n} & & 0 \\ & \ddots & \\ 0 & & (J_p^R)^{-n} \end{pmatrix} \\ 0_{r \times s} & - \begin{pmatrix} (J_{\nu+1}^L)^{m-n} & & \\ & \ddots & \\ 0 & & (J_p^R)^{m-n} \end{pmatrix} \end{pmatrix} \begin{pmatrix} MP \end{pmatrix}^{-1} \end{aligned}$$

□

et la fin du raisonnement se fait comme précédemment.

Fractions continues

Dans le cas scalaire, sont déjà définies les homographies et les transformations de Möbius permettant une description des fractions continues ainsi que des relations entre les convergents. Dans le cas du nombre d'or $\varphi = \frac{1+\sqrt{5}}{2}$, par exemple, on sait qu'il s'agit d'itérer l'homographie $h(x) = 1 + \frac{1}{x}$ pour obtenir les convergents successifs. Dans ce chapitre, on se propose de définir des transformations de Möbius pour des matrices rectangulaires $r \times s$.

De très nombreux articles considèrent des fractions continues dans les différents cas possibles : scalaire, vectoriel ou matriciel. Cependant, il semble que les diverses définitions soient indépendantes les unes des autres. Il nous a donc paru important, dans un souci d'uniformisation, de lier ces différents articles en créant des fractions continues et des homographies matricielles rectangulaires.

Ces outils, totalement nouveaux, permettent de décrire le processus de création de toute fraction continue comme une succession de compositions d'homographies. Ce développement en fraction continue permet par exemple de retrouver les coefficients d'une matrice représentant un opérateur aux différences. En effet, lorsque l'on connaît la fonction de Weyl définie au chapitre 1 au paragraphe 1.3.2, que l'on peut considérer comme une donnée spectrale, il est possible de retrouver les quantités qui composent la matrice

de départ en utilisant les fractions continues (voir par exemple [7], [73]). Les fractions continues tronquées sont liées aux approximants de Padé. Les transformations de Möbius sont donc un outil permettant d'examiner la convergence des approximants de Padé ou des fractions continues tronquées vers la fonction de Weyl.

De plus, ce processus de création des fractions continues matricielles rectangulaires comme composition d'homographies nous permet d'envisager, dans le futur, l'étude des "value sets" comme dans le cas scalaire.

Ce chapitre comprend deux parties. Dans un premier temps, après avoir défini les nouvelles notions que sont les fractions continues et les homographies matricielles rectangulaires $r \times s$, nous démontrons que nous retrouvons bien les cas particuliers comme le cas scalaire $r = s = 1$, le cas vectoriel avec $r = 1, s > 1$ et d'autres. Notre nouvel outil peut aussi décrire l'algorithme d'inversion de matrice rectangulaire donné par Sorokin et Van Iseghem [72].

Dans un deuxième temps, l'utilisation de ces outils originaux permet d'obtenir un théorème d'accélération de convergence de fractions continues matricielles rectangulaires. Ce théorème est une généralisation d'un théorème d'accélération de convergence existant pour les fractions continues généralisées décrites dans [25].

2.1 Formules générales

On va utiliser les récurrences obtenues dans le chapitre 1 pour définir une transformation de Möbius pour des matrices $W \in \mathbb{C}^{r \times s}$ afin de retrouver le même cadre que dans le cas scalaire par exemple.

Dans ce chapitre, on choisira l'initialisation des polynômes adaptée à chaque exemple.

On commence par expliciter certains résultats du chapitre 1 en faisant apparaître les polynômes vectoriels :

Propriété 2.1. *Les polynômes vectoriels $\mathcal{P}^R, \mathcal{P}^L, \mathcal{Q}^R, \mathcal{Q}^L$ vérifient les relations suivantes*

$$\begin{pmatrix} \mathcal{P}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L - \mathcal{P}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L & -\mathcal{P}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L + \mathcal{P}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L \\ \mathcal{Q}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L - \mathcal{Q}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L & -\mathcal{Q}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L + \mathcal{Q}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L \end{pmatrix} = \begin{pmatrix} I_r & 0_{r \times s} \\ 0_{s \times r} & I_s \end{pmatrix}$$

et

$$\begin{pmatrix} \mathcal{Q}_{n-s,s}^L \mathcal{P}_{n-r,r}^R - \mathcal{P}_{n-s,s}^L \mathcal{Q}_{n-r,r}^R & \mathcal{Q}_{n-s,s}^L \mathcal{P}_{n,s}^R - \mathcal{P}_{n-s,s}^L \mathcal{Q}_{n,s}^R \\ \mathcal{Q}_{n,r}^L \mathcal{P}_{n-r,r}^R - \mathcal{P}_{n,r}^L \mathcal{Q}_{n-r,r}^R & \mathcal{Q}_{n,r}^L \mathcal{P}_{n,s}^R - \mathcal{P}_{n,r}^L \mathcal{Q}_{n,s}^R \end{pmatrix} = \begin{pmatrix} 0_{s \times r} & (C_n^R)^{-1} \\ -(C_n^L)^{-1} & 0_{r \times s} \end{pmatrix}$$

Démonstration. Ces relations proviennent de la formule (1.6)

$$\forall n, B_n^R M_n B_n^L = I_{r+s}$$

Si on développe cette formule par blocs, on obtient les expressions de la propriété. \square

Définition 2.2. *En utilisant les notations du chapitre précédent, soient*

$$S_n^L(W) = \left(\begin{pmatrix} W & C_n^L \end{pmatrix} \begin{pmatrix} \mathcal{Q}_{n-s,s}^L \\ \mathcal{Q}_{n,r}^L \end{pmatrix} \right)^{-1} \begin{pmatrix} W & C_n^L \end{pmatrix} \begin{pmatrix} \mathcal{P}_{n-s,s}^L \\ \mathcal{P}_{n,r}^L \end{pmatrix}$$

$$\text{et } S_n^R(W) = \begin{pmatrix} \mathcal{P}_{n-r,r}^R & \mathcal{P}_{n,s}^R \end{pmatrix} \begin{pmatrix} W \\ C_n^R \end{pmatrix} \left(\begin{pmatrix} \mathcal{Q}_{n-r,r}^R & \mathcal{Q}_{n,s}^R \end{pmatrix} \begin{pmatrix} W \\ C_n^R \end{pmatrix} \right)^{-1}$$

où

$$C_n^L = \varepsilon A \begin{pmatrix} n, \dots, n+r-1 \\ n-r, \dots, n-1 \end{pmatrix}, \quad C_n^R = \varepsilon A \begin{pmatrix} n-s, \dots, n-1 \\ n, \dots, n+s-1 \end{pmatrix} \quad \text{et } \varepsilon = \pm 1$$

Ces formules sont un peu hermétiques mais dans le cas scalaire symétrique $r = s = 1$, on obtient simplement

$$S_n(w) = \frac{wp_{n-1} + a_{n-1}p_n}{wq_{n-1} + a_{n-1}q_n}$$

Remarque 2.3. *La première idée était de poser, pour les polynômes matriciels à gauche, par exemple*

$$S_n^L(W) = \left(\begin{pmatrix} W & I_r \end{pmatrix} \begin{pmatrix} \mathcal{Q}_{n-s,s}^L \\ \mathcal{Q}_{n,r}^L \end{pmatrix} \right)^{-1} \begin{pmatrix} W & I_r \end{pmatrix} \begin{pmatrix} \mathcal{P}_{n-s,s}^L \\ \mathcal{P}_{n,r}^L \end{pmatrix}$$

Et la valeur en $W = 0$ est l'approximant d'ordre n . Mais on voudrait que cette transformation ne dépende pas du type de polynômes : à gauche ou à droite. Pour avoir l'égalité $S_n^L(W) = S_n^R(W)$, il suffit de changer la matrice I_r en une matrice C_n^L et de même pour $S_n^R(W)$ avec la matrice C_n^R .

Propriété 2.4. *On a l'égalité*

$$S_n^L(W) = S_n^R(W)$$

Démonstration. En effet

$$\begin{aligned} & (W\mathcal{P}_{n-s,s}^L + C_n^L\mathcal{P}_{n,r}^L)(\mathcal{Q}_{n-r,r}^R W + \mathcal{Q}_{n,s}^R C_n^R) \\ &= W\mathcal{P}_{n-s,s}^L\mathcal{Q}_{n-r,r}^R W + C_n^L\mathcal{P}_{n,r}^L\mathcal{Q}_{n-r,r}^R W + W\mathcal{P}_{n-s,s}^L\mathcal{Q}_{n,s}^R C_n^R + C_n^L\mathcal{P}_{n,r}^L\mathcal{Q}_{n,s}^R C_n^R \\ &= W\mathcal{Q}_{n-s,s}^L\mathcal{P}_{n-r,r}^R W + C_n^L((C_n^L)^{-1} + \mathcal{Q}_{n,r}^L\mathcal{P}_{n-r,r}^R)W \\ & \quad + W(-(C_n^R)^{-1} + \mathcal{Q}_{n-s,s}^L\mathcal{P}_{n,s}^R)C_n^R + C_n^L\mathcal{Q}_{n,r}^L\mathcal{P}_{n,s}^R C_n^R \quad (\text{d'après la propriété 2.1}) \\ &= W\mathcal{Q}_{n-s,s}^L\mathcal{P}_{n-r,r}^R W + C_n^L\mathcal{Q}_{n,r}^L\mathcal{P}_{n-r,r}^R W + W\mathcal{Q}_{n-s,s}^L\mathcal{P}_{n,s}^R C_n^R + C_n^L\mathcal{Q}_{n,r}^L\mathcal{P}_{n,s}^R C_n^R \\ &= (W\mathcal{Q}_{n-s,s}^L + C_n^L\mathcal{Q}_{n,r}^L)(\mathcal{P}_{n-r,r}^R W + \mathcal{P}_{n,s}^R C_n^R) \end{aligned}$$

Ce qui prouve l'égalité demandée. \square

Définition 2.5. On définit aussi les homographies :

$$s_n^L(W) = \left(\left(W E_{n-1}^L + C_n^L G_{n-1}^L \right) \left(C_{n-1}^L \right)^{-1} \right)^{-1} \left(W D_{n-1}^L + C_n^L F_{n-1}^L \right)$$

et

$$s_n^R(W) = \left(D_{n-1}^R W + F_{n-1}^R C_n^R \right) \left(\left(C_{n-1}^R \right)^{-1} \left(E_{n-1}^R W + G_{n-1}^R C_n^R \right) \right)^{-1}$$

où les matrices notées D, E, F, G sont les blocs des matrices de récurrence A_n^L et A_n^R comme suit :

$$A_{n-1}^L = \begin{matrix} & \overbrace{\hspace{1.5cm}}^s & \overbrace{\hspace{1.5cm}}^r \\ s \{ & \begin{pmatrix} D_{n-1}^L & E_{n-1}^L \\ F_{n-1}^L & G_{n-1}^L \end{pmatrix} & \\ r \{ & \end{matrix} \quad \text{et} \quad A_{n-1}^R = \begin{matrix} & \overbrace{\hspace{1.5cm}}^r & \overbrace{\hspace{1.5cm}}^s \\ r \{ & \begin{pmatrix} D_{n-1}^R & F_{n-1}^R \\ E_{n-1}^R & G_{n-1}^R \end{pmatrix} & \\ s \{ & \end{matrix}$$

c'est-à-dire

$$D_{n-1}^L = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & 0 \end{pmatrix}, \quad E_{n-1}^L = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & & \vdots \\ 1 & 0 & 0 \end{pmatrix},$$

$$\text{et } F_{n-1}^L = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \\ f_1^L & \dots & f_s^L \end{pmatrix}, \quad G_{n-1}^L = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ 0 & \dots & 0 & 1 \\ g_1^L & \dots & \dots & g_r^L \end{pmatrix}$$

et

$$D_{n-1}^R = \begin{pmatrix} 0 & \dots & \dots & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & 0 \end{pmatrix}, \quad E_{n-1}^R = \begin{pmatrix} 0 & \dots & 0 & 1 \\ \vdots & & & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix},$$

$$\text{et } F_{n-1}^R = \begin{pmatrix} 0 & \dots & 0 & f_1^R \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 0 & \dots & 0 & f_r^R \end{pmatrix}, \quad G_{n-1}^R = \begin{pmatrix} 0 & \dots & 0 & g_1^R \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & 0 & \vdots \\ 0 & & 1 & g_s^R \end{pmatrix}$$

Bien évidemment, f_j^L, f_j^R, g_j^L et g_j^R dépendent de n de la manière suivante :

$$\left\{ \begin{array}{l} \forall j \in \{1, \dots, s\}, f_j^L = \frac{-A_{n-1, n-s-2+j}}{A_{n-1, n+r-1}}, \\ g_1^L = \frac{z - A_{n-1, n-1}}{A_{n-1, n+r-1}} \text{ et } \forall j \in \{2, \dots, r\}, g_j^L = \frac{-A_{n-1, n-2+j}}{A_{n-1, n+r-1}} \\ \forall j \in \{1, \dots, r\}, f_j^R = \frac{-A_{n-2-r+j, n-1}}{A_{n-1+s, n-1}}, \\ g_1^R = \frac{z - A_{n-1, n-1}}{A_{n-1+s, n-1}} \text{ et } \forall j \in \{2, \dots, s\}, g_j^R = \frac{-A_{n-2+j, n-1}}{A_{n-1+s, n-1}} \end{array} \right.$$

Encore une fois, appliquons cette définition dans le cas scalaire symétrique $r = s = 1$, on retrouve l'homographie habituelle $s_n(w) = \frac{-a_{n-2}^2}{w + (z - b_{n-1})}$

Propriété 2.6. $S_n^L(W) = S_{n-1}^L(s_n(W))$ et $S_n^R(W) = S_{n-1}^R(s_n(W))$.

Démonstration. On rappelle que

$$\begin{aligned} s_n^L(W) &= \left(\left(W E_{n-1}^L + C_n^L G_{n-1}^L \right) \left(C_{n-1}^L \right)^{-1} \right)^{-1} \left(W D_{n-1}^L + C_n^L F_{n-1}^L \right) \\ &= C_{n-1}^L \left(W E_{n-1}^L + C_n^L G_{n-1}^L \right)^{-1} \left(W D_{n-1}^L + C_n^L F_{n-1}^L \right) \\ &= S_{n-1}^L(s_n(W)) \\ &= \left(\left(s_n(W) \ C_{n-1}^L \right) \left(\begin{array}{c} \mathcal{Q}_{n-1-s, s}^L \\ \mathcal{Q}_{n-1, r}^L \end{array} \right) \right)^{-1} \left(s_n(W) \ C_{n-1}^L \right) \left(\begin{array}{c} \mathcal{P}_{n-1-s, s}^L \\ \mathcal{P}_{n-1, r}^L \end{array} \right) \\ &= \left(C_{n-1}^L \left(\left(W E_{n-1}^L + C_n^L G_{n-1}^L \right)^{-1} \left(W D_{n-1}^L + C_n^L F_{n-1}^L \right) \mathcal{Q}_{n-1-s, s}^L + \mathcal{Q}_{n-1, r}^L \right) \right)^{-1} \\ &\quad \left(C_{n-1}^L \left(\left(W E_{n-1}^L + C_n^L G_{n-1}^L \right)^{-1} \left(W D_{n-1}^L + C_n^L F_{n-1}^L \right) \mathcal{P}_{n-1-s, s}^L + \mathcal{P}_{n-1, r}^L \right) \right)^{-1} \\ &= \left(\left(W D_{n-1}^L + C_n^L F_{n-1}^L \right) \mathcal{Q}_{n-1-s, s}^L + \left(W E_{n-1}^L + C_n^L G_{n-1}^L \right) \mathcal{Q}_{n-1, r}^L \right)^{-1} \\ &\quad \left(\left(W D_{n-1}^L + C_n^L F_{n-1}^L \right) \mathcal{P}_{n-1-s, s}^L + \left(W E_{n-1}^L + C_n^L G_{n-1}^L \right) \mathcal{P}_{n-1, r}^L \right)^{-1} \\ &= \left(W (D_{n-1}^L \mathcal{Q}_{n-1-s, s}^L + E_{n-1}^L \mathcal{Q}_{n-1, r}^L) + C_n^L (F_{n-1}^L \mathcal{Q}_{n-1-s, s}^L + G_{n-1}^L \mathcal{Q}_{n-1, r}^L) \right)^{-1} \\ &\quad \left(W (D_{n-1}^L \mathcal{P}_{n-1-s, s}^L + E_{n-1}^L \mathcal{P}_{n-1, r}^L) + C_n^L (F_{n-1}^L \mathcal{P}_{n-1-s, s}^L + G_{n-1}^L \mathcal{P}_{n-1, r}^L) \right) \\ &= \left(W \mathcal{Q}_{n-s, s}^L + C_n^L \mathcal{Q}_{n, r}^L \right)^{-1} \left(W \mathcal{P}_{n-s, s}^L + C_n^L \mathcal{P}_{n, r}^L \right) = S_n^L(W) \end{aligned}$$

□

2.2 Vérifications des formules

Theorème 2.7. *Les définitions données précédemment sont compatibles avec les autres cas : cas scalaire, cas vectoriel (algorithme de Jacobi-Perron [58, p. 143 & sq.]), fractions continues généralisées [25],[50], fractions continues matricielles [71], [73], fractions matricielles par bloc [3].*

Démonstration. On décrit la matrice de Jacobi associée à ces différents cas, c'est-à-dire la matrice de l'opérateur associé ou encore les coefficients de récurrence des polynômes et, en utilisant les formules générales, on retrouvera les formules connues des homographies dans les divers cas.

2.2.1 Le cas scalaire

Dans le cas scalaire, $r=s=1$ et donc la matrice associée est tridiagonale. Supposons que l'on travaille avec les polynômes orthonormés. C'est-à-dire que l'on a :

$$A = \begin{pmatrix} b_0 & a_0 & 0 & \cdots \\ a_0 & b_1 & a_1 & \\ 0 & \ddots & \ddots & \ddots \end{pmatrix} \text{ et } M_n = \begin{pmatrix} 0 & -a_{n-1} \\ a_{n-1} & 0 \end{pmatrix}$$

Pour l'initialisation, on a $a_{-1} = 1$. Auquel cas, on obtient $\mathcal{P}_{-1,1}^L = \mathcal{P}_{-1,1}^R = p_{-1} = -1$, $\mathcal{Q}_{-1,1}^L = \mathcal{Q}_{-1,1}^R = q_{-1} = 0$, $\mathcal{P}_{0,1}^L = \mathcal{P}_{0,1}^R = p_0 = 0$, $\mathcal{Q}_{0,1}^L = \mathcal{Q}_{0,1}^R = q_0 = 1$ c'est-à-dire que l'on retrouve les initialisations du cas scalaire. La matrice de récurrence est :

$$A_{n-1}^R := \begin{pmatrix} 0 & \frac{-A_{n-2,n-1}}{A_{n,n-1}} \\ 1 & \frac{z - A_{n-1,n-1}}{A_{n,n-1}} \end{pmatrix} = \begin{pmatrix} 0 & \frac{-a_{n-2}}{a_{n-1}} \\ 1 & \frac{z - b_{n-1}}{a_{n-1}} \end{pmatrix}$$

et les homographies sont

$$S_n(w) = \frac{wp_{n-1} + a_{n-1}p_n}{wq_{n-1} + a_{n-1}q_n} \text{ et } s_n(w) = \frac{-a_{n-2}^2}{w + (z - b_{n-1})}$$

2.2.2 Fractions continues généralisées

Pour les articles [25] et [50], on peut faire de même. En effet, on se donne une suite de $(s+1)$ -uplets $(b_k, a_k^{(1)}, \dots, a_k^{(s)})$, $a_k^{(1)} \neq 0$ alors la s -fraction

associée à la suite, écrite sous la forme (cf. [25, p. 161 & 163])

$$K_{k=1}^{\infty} \begin{pmatrix} a_k^{(1)} \\ \vdots \\ a_k^{(s)} \\ b_k \end{pmatrix}$$

est donnée par la suite des approximants $\{A_k^{(1)}/B_k, \dots, A_k^{(s)}/B_k\}_{k=1}^{\infty}$ (s'ils existent), où les numérateurs et dénominateurs vérifient la récurrence à $s+2$ termes :

$$X_k = b_k X_{k-1} + a_k^{(s)} X_{k-2} + \dots + a_k^{(1)} X_{k-s-1}, \quad k = 1, 2, \dots,$$

avec les initialisations

$$A_{-j}^{(i)} = \delta_{i+j, s+1}, \quad B_{-j} = \delta_{s+1+j, s+1} \quad (j = 0, \dots, s \quad i = 1, \dots, s).$$

C'est-à-dire que l'on a la récurrence matricielle suivante :

$$\begin{pmatrix} X_{n-s} \\ \vdots \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ 0 & \dots & 0 & 1 \\ a_n^{(1)} & \dots & a_n^{(s)} & b_n \end{pmatrix} \begin{pmatrix} X_{n-s-1} \\ \vdots \\ \vdots \\ X_{n-1} \end{pmatrix}$$

Comme on a un seul dénominateur, on a $r=1$. Auquel cas, on a $Q_n^L = B_n$ et $P_n^L = (A_n^{(1)}, \dots, A_n^{(s)})$. Avec ces nouvelles notations, on retrouve la forme de la récurrence générale $B_n^L = A_{n-1}^L \cdot B_{n-1}^L$. Donc, par comparaison, on pose $\forall n$, $A_{n-1, n} = 1$, c'est-à-dire que la sur-diagonale est composée de 1. Puis, on a $b_n = x - A_{n-1, n-1}$ et $a_n^{(s)} = -A_{n-1, n-2}, \dots, a_n^{(1)} = -A_{n-1, n-s-1}$. Donc la matrice de Jacobi est de la forme :

$$A = \begin{pmatrix} x - b_1 & 1 & 0 & 0 & \dots \\ -a_2^{(s)} & x - b_2 & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & \\ -a_{s+1}^{(1)} & \dots & -a_{s+1}^{(s)} & x - b_{s+1} & 1 \\ & \ddots & & & \ddots \end{pmatrix}$$

Il faut aussi examiner les initialisations. On a

$$\begin{pmatrix} \mathcal{P}_{-s, s}^L & \mathcal{Q}_{-s, s}^L \\ \mathcal{P}_{0, 1}^L & \mathcal{Q}_{0, 1}^L \end{pmatrix} = \begin{pmatrix} A_{-s}^{(1)} & \dots & A_{-s}^{(s)} & B_{-s} \\ \vdots & & \vdots & \vdots \\ A_0^{(1)} & \dots & A_0^{(s)} & B_0 \end{pmatrix} = I_{s+1}$$

C'est-à-dire que l'on retrouve les bonnes valeurs pour le dénominateur et les numérateurs au rang 0 et qu'il suffit d'initialiser le reste comme demandé. C'est-à-dire que l'on impose que les termes de la matrice de Jacobi d'indice négatif soient ceux imposés par cette initialisation.

Pour les homographies, au lieu d'écrire composante par composante comme indiqué, il suffit de tout regrouper en une matrice de dimension $1 \times s$ c'est-à-dire que, en utilisant leurs notations, on pose $S_n = (S_n^{(1)}, \dots, S_n^{(s)})$, $s_n = (s_n^{(1)}, \dots, s_n^{(s)})$ et $W = (w_1, \dots, w_s)$. Or, en la coordonnée i , on a d'après l'article

$$\begin{aligned} S_n^{(i)}(W) &= S_{n-1}^{(i)}(s_n^{(1)}(W), \dots, s_n^{(s)}(W)) = S_{n-1}^{(i)}(s_n(W)) \\ &= S_{n-1}^{(i)}\left(\frac{a_n^{(1)}}{b_n + w_s}, \frac{a_n^{(2)} + w_1}{b_n + w_s}, \dots, \frac{a_n^{(s)} + w_{s-1}}{b_n + w_s}\right) \\ &= S_{n-1}^{(i)}((b_n + w_s)^{-1} \left[(a_n^{(1)}, \dots, a_n^{(s)}) + (0, w_1, \dots, w_{s-1}) \right]). \end{aligned}$$

La formule générale donne :

$$S_n^L(W) = ((w_s + 1 \cdot b_n)(1)^{-1})^{-1}((0, w_1, \dots, w_{s-1}) + 1 \cdot (a_n^{(1)}, \dots, a_n^{(s)}))$$

c'est-à-dire la même formule.

Pour S_n^L , la formule générale est la même que celle donnée dans l'article [50, formule 5] en regroupant les s formules de l'article en un seul vecteur à s composantes comme dit précédemment.

2.2.3 Cas vectoriel (algorithme de Jacobi-Perron)

Pour l'algorithme de Jacobi-Perron (cf. [58, p. 143] et plus précisément [60]), $r = 1$ et s est quelconque. D'après la formule (5.1) p. 144, il semble normal de prendre

$$s_n(w) = \frac{1}{(p_{1,n}, \dots, p_{s,n}) + (w_1, \dots, w_s)}.$$

Et en utilisant la définition de l'inverse d'un vecteur, on obtient

$$\begin{aligned} s_n(w) &= \left(\frac{1}{p_{s,n} + w_s}, \frac{p_{1,n} + w_1}{p_{s,n} + w_s}, \dots, \frac{p_{s-1,n} + w_{s-1}}{p_{s,n} + w_s} \right) \\ &= (p_{s,n} + w_n)^{-1} \left((1, p_{1,n}, \dots, p_{s-1,n}) + (0, w_1, \dots, w_{s-1}) \right) \end{aligned}$$

On veut savoir quelle est la forme de la matrice de Jacobi associée pour obtenir la matrice de récurrence et ainsi retrouver les formules générales. Ainsi, dans [60, p. 291], la formule (11) indique que les n -ièmes approximants sont de la forme $(B_{1,n}/B_{0,n}, \dots, B_{s,n}/B_{0,n})$ où les $B_{i,n}$ vérifient une relation

de récurrence de la forme

$$B_{i,n} = \sum_{l=0}^s p_{s-l,n} B_{i,s-l-1}, \quad (i = 0, \dots, s)$$

$$\text{et } B_{i,k} = \delta_{i-k,s+1} \quad (i = 0, \dots, s; k = -1, \dots, -s-1)$$

Donc on obtient une récurrence à $s+2$ termes que l'on peut mettre sous forme matricielle :

$$\begin{pmatrix} X_{n-s} \\ \vdots \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ 0 & \dots & 0 & 1 \\ p_{0,n} & \dots & \dots & p_{s,n} \end{pmatrix} \begin{pmatrix} X_{n-s-1} \\ \vdots \\ \vdots \\ X_{n-1} \end{pmatrix}$$

Comme précédemment, on pose $Q_k^L = A_{0,k-1}$ et $P_k^L = (A_{1,k-1}, \dots, A_{s,k-1})$. On retrouve la formule générale. On peut donc en déduire la forme de la matrice de Jacobi

$$A = \begin{pmatrix} x - p_{s,0} & 1 & 0 & & \dots \\ -p_{s-1,1} & x - p_{s,1} & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ -p_{0,s} & \dots & -p_{s-1,s} & x - p_{s,s} & 1 \\ & \ddots & & \ddots & \ddots \end{pmatrix}$$

On doit changer les initialisations de l'article soit

$$A_{0,k} = \delta_{k,-1}, \quad (k = -1, \dots, -s-1)$$

et

$$A_{i,k} = \delta_{i-k,s+2}, \quad (i = 1, \dots, s, k = -1, \dots, -s-1)$$

et on a alors

$$\begin{pmatrix} \mathcal{P}_{-s,s}^L & \mathcal{Q}_{-s,s}^L \\ \mathcal{P}_{0,1}^L & \mathcal{Q}_{0,1}^L \end{pmatrix} = \begin{pmatrix} A_{1,-s-1} & \dots & A_{s,-s-1} & A_{0,-s-1} \\ \vdots & & \vdots & \\ A_{1,-1} & \dots & A_{s,-1} & A_{0,-1} \end{pmatrix} = I_{s+1}$$

En utilisant la formule générale, on retrouve la forme de s_n que l'on supposait correcte. On peut aussi obtenir la forme de S_n puisque les polynômes associés vérifient aussi une relation de récurrence à $s+2$ termes et ainsi on retrouve les formules données dans le cas général.

2.2.4 Fractions continues matricielles

On se donne une matrice de fonctions de taille $r \times s$ et on cherche des polynômes matriciels qui réalise la meilleure approximation rationnelle. Pour cela, dans l'article [71], l'inverse d'une matrice de dimension $r \times s$ est défini comme l'image d'une application T . Cette application est définie comme suit.

Soit une matrice B de dimension $r \times s$. On forme la matrice $\begin{pmatrix} I_s \\ B \end{pmatrix}$ de taille $(r+s) \times s$. On lui applique l'application \tilde{T} qui permute les lignes en plaçant la dernière ligne à la première. On obtient donc une matrice $\begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$ et on pose $T(B) = B_2 B_1^{-1}$. A partir de là, on peut oublier l'application et simplement écrire l'inverse d'une matrice. Le convergent d'ordre n se note alors

$$\Pi_n = \frac{1}{B_1 + \frac{1}{B_2 + \frac{1}{\dots + \frac{1}{B_n}}}}$$

On suppose que la fraction continue est régulière c'est-à-dire que toutes les matrices B_k sont de la forme

$$B_k(z) = \begin{pmatrix} 0 & & 0 & \gamma_{k,1} \\ & \dots & & \vdots \\ 0 & & 0 & \gamma_{k,r-1} \\ \delta_{k,1} & \dots & \delta_{k,s-1} & \alpha_k z + \beta_k \end{pmatrix}$$

On définit une classe d'équivalence sur l'ensemble des matrices de taille $r \times s$ de la manière suivante : les matrices A et B sont équivalentes si et seulement si on peut trouver une matrice inversible C telle que $A = B \cdot C$.

Le but est de retrouver une relation de récurrence à $r + s + 1$ termes et le lien avec le convergent puisque l'on travaille avec des classes d'équivalence. On travaille ici avec les polynômes à droite P_n^R et Q_n^R . En utilisant l'application T , la forme du convergent et des matrices B_k , on peut supposer que $s_n^R(W) = 1/(B_n + W) := T(B_n + W) := (B_n^2 + W^2)(B_n^1 + W^1)^{-1}$

où

$$\tilde{T} \begin{pmatrix} I_s \\ B_n + W \end{pmatrix} = \begin{pmatrix} B_n^1 + W^1 \\ B_n^2 + W^2 \end{pmatrix} \text{ et } T(B_n + W) = (B_n^2 + W^2)(B_n^1 + W^1)^{-1}$$

Examinons la forme de $B_n^1 + W^1$ et $B_n^2 + W^2$, on a

$$B_n^1 + W^1 = \begin{pmatrix} w_{r,1} + \delta_{n,1} & \cdots & w_{r,s-1} + \delta_{n,s-1} & \alpha_n z + \beta_n + w_{r,s} \\ 1 & & & 0 \\ & \ddots & & \\ & & 1 & 0 \end{pmatrix}$$

$$\text{et } B_n^2 + W^2 = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ w_{1,1} & \cdots & w_{1,s-1} & \gamma_{n,1} + w_{1,s} \\ \vdots & & \vdots & \vdots \\ w_{r-1,1} & \cdots & w_{r-1,s-1} & \gamma_{n,r-1} + w_{r-1,s} \end{pmatrix}$$

On va comparer avec la forme générale de s_n^R .

Pour cela, on a besoin de la matrice de Jacobi et de la matrice de récurrence associée. Le convergent Π_n est obtenu par l'utilisation récursive de T mais peut aussi être calculé par \tilde{T} . Or cette application revient à une multiplication par une matrice N_i à chaque étape de la récursion. C'est-à-dire que l'on a

$$\begin{pmatrix} I_s \\ \Pi_n \end{pmatrix} = \prod_{i=1}^n N_i \begin{pmatrix} I_s \\ 0_{r \times s} \end{pmatrix} = M_n \begin{pmatrix} I_s \\ 0_{r \times s} \end{pmatrix}$$

Donc si on note par y_n^k , $k = 1, \dots, r + s$ les colonnes de M_n , alors $\begin{pmatrix} I_s \\ \Pi_n \end{pmatrix} = (y_n^1, \dots, y_n^s)$.

On peut montrer les égalités suivantes entre classes d'équivalence :

$$\mathcal{Cl} \begin{pmatrix} I_s \\ \Pi_n \end{pmatrix} = \mathcal{Cl}(y_n^1, \dots, y_n^s) = \mathcal{Cl}(y_n^1, \dots, y_{n+s-1}^1)$$

où la suite $(y_n^1)_{n \geq 0}$ de vecteurs de taille $r + s$ vérifie la récurrence suivante :

$$y_{n+s}^1 = y_{n+s-1}^1 \delta_{n+s,1} + \cdots + y_{n+1}^1 \delta_{n+2,s-1} + y_n^1 (\alpha_{n+1} z + \beta_{n+1}) \\ + y_{n-1}^1 \gamma_{n+1,r-1} + \cdots + y_{n-r+1}^1 \gamma_{n+1,1} + y_{n-r}^1$$

On connaît même la relation entre les deux dernières matrices :

$$(y_n^1, y_n^2, \dots, y_n^s) = (y_n^1, y_{n+1}^1, \dots, y_{n+s-1}^1) \begin{pmatrix} 1 & -\delta_{n+1,1} & \cdots & -\delta_{n+1,s-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & -\delta_{n+s-1,1} \\ & & & 1 \end{pmatrix}$$

Si on partage les $(y_k^1)_{k \geq 0}$ sous la forme

$$(y_n^1, y_{n+1}^1, \dots, y_{n+s-1}^1) = \begin{pmatrix} Q_n & \cdots & Q_{n+s-1} \\ P_n & \cdots & P_{n+s-1} \end{pmatrix}$$

où P_k est un vecteur de taille r et Q_k , un vecteur de taille s , on retrouve les polynômes à droite P_n^R et Q_n^R .

De plus, on a la récurrence matricielle suivante :

$$(y_n^1, \dots, y_{n+r+s-1}^1) = (y_{n-1}^1, \dots, y_{n+r+s-2}^1) \begin{pmatrix} 0 & & & 0 & 1 \\ 1 & & & & \gamma_{n+r,1} \\ & \ddots & & & \vdots \\ & & & & \gamma_{n+r,r-1} \\ & & \ddots & & \alpha_{n+r}z + \beta_{n+r} \\ & & & \ddots & \delta_{n+r+1,s-1} \\ & & & & \vdots \\ 0 & & & \ddots & 1 & \delta_{n+r+s-1,1} \end{pmatrix}$$

Puisque on doit décaler les indices de r , cela signifie que la dernière colonne de la matrice de récurrence A_{n-1}^R , est

$$(1, \gamma_{n,1}, \dots, \gamma_{n,r-1}, \alpha_n z + \beta_n, \delta_{n+1,s-1}, \dots, \delta_{n-1+s,1})^t.$$

Dans la forme générale de la matrice de récurrence, le coefficient de z dans le seul terme en comportant est $1/A_{n-1+s,n-1}$. On en déduit que $A_{n-1+s,n-1} = 1/\alpha_n$. La colonne indexée par $n-1$ dans la matrice de Jacobi, c'est-à-dire la n -ième colonne est donc

$$-1/\alpha_n(1, \gamma_{n,1}, \dots, \gamma_{n,r-1}, \beta_n, \delta_{n+1,s-1}, \dots, \delta_{n+s-1,1}, -1)^t.$$

La forme de C_n^R est donc

$$C_n^R = \begin{pmatrix} 1 & -\delta_{n+1,1} & -\delta_{n+1,2} & \cdots & -\delta_{n+1,s-1} \\ & 1 & -\delta_{n+2,1} & & \vdots \\ & & \ddots & & \vdots \\ & & & 1 & -\delta_{n+s-1,1} \\ & & & & 1 \end{pmatrix}$$

Donc, dans s_n^R , la matrice $D_{n-1}^R W + F_{n-1}^R C_n^R = D_{n-1}^R W + F_{n-1}^R$ vaut bien $B_n^2 + W^2$. De même, $(C_{n-1}^R)^{-1} E_{n-1}^R W = E_{n-1}^R W = W^1$. Il reste donc à étudier $(C_{n-1}^R)^{-1} G_{n-1}^R C_n^R$. Si on calcule $G_{n-1}^R C_n^R$ et $C_{n-1}^R B_n^1$, on obtient la même matrice

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & \alpha_n z + \beta_n \\ 1 & -\delta_{n+1,1} & \cdots & -\delta_{n+1,s-2} & 0 \\ & \ddots & \ddots & \vdots & \vdots \\ & & \ddots & -\delta_{n+s-2,1} & \vdots \\ & & & 1 & 0 \end{pmatrix}$$

Finalement, la définition de s_n^R entre bien dans le cadre général. A partir de la suite définie par récurrence, on peut obtenir $S_n^R(W)$ et récupérer le convergent Π_n . Il faut prêter attention au fait que dans l'article, il faut changer l'initialisation du théorème où apparaît la suite $(y_n^1)_n$ ou plutôt la compléter. On peut prendre les mêmes initialisations que d'habitude à savoir $(y_0^1, \dots, y_0^{s-1}) = \begin{pmatrix} I_s \\ 0_{r \times s} \end{pmatrix}$ et comme il y a égalité des classes d'équivalence, on obtient le résultat désiré.

2.2.5 Fractions continues matricielles par blocs

On utilise les notations indiquées dans [3]. On se donne la matrice de Jacobi suivante :

$$A = \begin{pmatrix} V_0 & E_0 & 0 & 0 & \dots \\ E_0^* & V_1 & E_1 & 0 & \dots \\ 0 & E_1^* & V_2 & E_2 & \dots \\ \dots & \dots & \ddots & \ddots & \ddots \end{pmatrix} \quad \text{où } E_j \text{ et } V_j \text{ appartiennent à } \mathbb{C}^N$$

Dans l'article, il est supposé que $V_j^* = V_j$ et $\det E_j \neq 0$. On supposera, en plus, pour l'instant, que E_j est une matrice triangulaire inférieure. On montrera que l'on peut toujours écrire la matrice sous cette forme. La condition sur le déterminant de la matrice permet d'affirmer que tous les termes de la diagonale de E_j sont non nuls. La matrice est donc une matrice-bande avec N sous-diagonales et N sur-diagonales.

On définit aussi les polynômes Q_n^L par

$$Q_{-N,N}^L(z) = 0_N, \quad Q_{0,N}^L(z) = I_N,$$

$$E_k Q_{(k+1)N,N}^L(z) = (zI_N - V_k) Q_{kN,N}^L(z) - E_{k+1}^* Q_{(k-1)N,N}^L(z)$$

Les polynômes P_n^L vérifient la même relation de récurrence mais

$$P_{0,N}^L(z) = 0_N, \quad P_{N,N}^L(z) = E_0^{-1}.$$

Alors les fractions continues $\pi_k(z) = (Q_{kN,N}^L(z))^{-1} P_{kN,N}^L(z)$ sont les convergents de la fraction continue suivante [3, p. 331] :

$$\hat{\sigma}(z) = \frac{1}{zI_N - V_0 - E_0 \frac{1}{zI_N - V_1 - E_1 \frac{1}{zI_N - V_2 - \ddots}} E_0^*}$$

On va montrer que chaque "homographie-bloc" dans la fraction continue correspond en fait à la composition de N homographies simples définies dans le cas général. L'homographie-bloc serait donc définie comme suit :

$$\mathcal{S}_k(W) = -E_k(zI_N - V_{k+1} + W)^{-1}E_k^*$$

Etudions donc $s_{kN-N+1}^L \circ \dots \circ s_{kN}^L$. Posons $\beta = kN$. On a alors :

$s_{\beta}^L(W) = C_{\beta-1}^L(W E_{\beta-1}^L + C_{\beta}^L G_{\beta-1}^L)^{-1}(W D_{\beta-1}^L + C_{\beta}^L F_{\beta-1}^L)$ en reprenant les notations du cas général.

Soit, avec une définition évidente, $s_{\beta}^L = C_{\beta-1}^L(Y_{\beta}^L)^{-1}Z_{\beta}^L$. On compose $s_{\beta-1} \circ s_{\beta}$. D'après la forme de s_{β} et après factorisation par $C_{\beta-1}^L(Y_{\beta}^L)^{-1}$:

$s_{\beta-1} \circ s_{\beta}(W) = C_{\beta-2}^L(Z_{\beta}^L E_{\beta-2}^L + Y_{\beta}^L G_{\beta-2}^L)^{-1}(Z_{\beta}^L D_{\beta-2}^L + Y_{\beta}^L F_{\beta-2}^L)$. Par une récurrence constructive immédiate, on obtient que, pour tout k inférieur à β , $s_k^L \circ \dots \circ s_{\beta}^L$ est de la forme $C_{k-1}^L(Y_k^L)^{-1}Z_k^L$.

Les numérateurs Z_k^L et les dénominateurs Y_k^L sont donc définis par récurrence. On montrerait qu'ils sont de la forme :

$$Y_k^L = W H_k^L + C_{\beta}^L N_k^L \text{ et } Z_k^L = W J_k^L + C_{\beta}^L K_k^L \text{ où}$$

$$\begin{pmatrix} J_k^L & H_k^L \\ K_k^L & N_k^L \end{pmatrix} = \begin{pmatrix} J_{k+1}^L & H_{k+1}^L \\ K_{k+1}^L & N_{k+1}^L \end{pmatrix} \begin{pmatrix} D_{k-1}^L & E_{k-1}^L \\ F_{k-1}^L & G_{k-1}^L \end{pmatrix} = \begin{pmatrix} J_{k+1}^L & H_{k+1}^L \\ K_{k+1}^L & N_{k+1}^L \end{pmatrix} A_{k-1}^L$$

avec A_k^L défini comme au chapitre précédent. L'initialisation de la récurrence est réalisée par

$$\begin{pmatrix} J_{\beta}^L & H_{\beta}^L \\ K_{\beta}^L & N_{\beta}^L \end{pmatrix} = \begin{pmatrix} D_{\beta-1}^L & E_{\beta-1}^L \\ F_{\beta-1}^L & G_{\beta-1}^L \end{pmatrix} = A_{\beta-1}^L$$

d'après la forme de Y_{β}^L et Z_{β}^L . On obtient donc :

$$(s_{\beta-N+1}^L \circ \dots \circ s_{\beta}^L)(W) = C_{\beta-N}^L(W H_{\beta-N+1}^L + C_{\beta}^L N_{\beta-N+1}^L)^{-1}(W J_{\beta-N+1}^L + C_{\beta}^L K_{\beta-N+1}^L)$$

D'après la relation de récurrence, on obtient

$$\begin{pmatrix} J_{\beta-N+1}^L & H_{\beta-N+1}^L \\ K_{\beta-N+1}^L & N_{\beta-N+1}^L \end{pmatrix} = \prod_{i=1}^N A_{\beta-i}^L$$

On examine donc le produit matriciel de N matrices compagnons de type A_i^L . Ces matrices carrées sont de dimension $2N$. Donc le produit de N matrices compagnons donne une matrice dont les N dernières lignes sont pleines et les N premières sont le bloc $(0_N, I_N)$. On obtient donc les valeurs suivantes :

$$J_{\beta-N+1}^L = 0_N \text{ et } H_{\beta-N+1}^L = I_N$$

On obtient les formes suivantes pour le numérateur et le dénominateur de la composition :

$$Y_{\beta-N+1}^L = W + C_{\beta}^L N_{\beta-N+1} \text{ et } Z_{\beta-N+1}^L = C_{\beta}^L K_{\beta-N+1}^L.$$

On ne peut trouver directement la forme de $N_{\beta-N+1}^L$ et $K_{\beta-N+1}^L$. On va donc essayer de décrire les formes de $C_{\beta}^L N_{\beta-N+1}^L$ et $C_{\beta}^L K_{\beta-N+1}^L$.

Reprenons le calcul au départ. On commence par multiplier C_{β}^L par $\begin{pmatrix} F_{\beta-1}^L & G_{\beta-1}^L \end{pmatrix}$. On obtient la matrice suivante :

$$\begin{pmatrix} 0 & \dots & 0 & 0 & A_{\beta-N,\beta} & & & 0 \\ \vdots & & \vdots & \vdots & \vdots & A_{\beta-N+1,\beta+1} & & \\ \vdots & & \vdots & \vdots & \vdots & & \ddots & \\ 0 & \dots & 0 & 0 & A_{\beta-2,\beta} & A_{\beta-2,\beta+1} & \dots & A_{\beta-2,\beta+N-2} \\ -A_{\beta-1,\beta-N-1} & \dots & A_{\beta-1,\beta-2} & z - A_{\beta-1,\beta-1} & 0 & \dots & \dots & 0 \end{pmatrix}$$

que l'on décompose en deux blocs :

$$\begin{pmatrix} \overbrace{2N-1} & \overbrace{1} \\ S_{\beta-1}^L & T_{\beta-1}^L \end{pmatrix}$$

On multiplie cette matrice à droite par $A_{\beta-2}^L$, elle aussi décomposée en blocs :

$$2N-1 \left\{ \begin{pmatrix} \overbrace{1} & \overbrace{2N-1} \\ 0_{(2N-1) \times 1} & I_{2N-1} \end{pmatrix} \right. \\ \left. 1 \left\{ \begin{pmatrix} u_{\beta-2} & V_{\beta-2} \end{pmatrix} \right. \right.$$

Ce qui donne $\begin{pmatrix} u_{\beta-2} T_{\beta-1}^L & S_{\beta-1}^L + T_{\beta-1}^L V_{\beta-2} \end{pmatrix}$ c'est-à-dire

$$\begin{pmatrix} 0 & \dots & \dots & 0 & A_{\beta-N,\beta} & & 0 \\ \vdots & & & \vdots & \vdots & \ddots & \\ 0 & \dots & \dots & 0 & A_{\beta-3,\beta} & \dots & A_{\beta-3,\beta+N-3} \\ -A_{\beta-2,\beta-N-2} & \dots & \dots & -A_{\beta-2,\beta-1} & 0 & \dots & 0 \\ 0 & -A_{\beta-1,\beta-N-1} & \dots & z - A_{\beta-1,\beta-1} & 0 & \dots & 0 \end{pmatrix},$$

d'après la valeur des coefficients de $A_{\beta-2}^L$.

On obtient donc une matrice de la même forme que $\begin{pmatrix} C_{\beta}^L F_{\beta-1}^L & C_{\beta}^L G_{\beta-1}^L \end{pmatrix}$ si ce n'est que le bloc triangulaire en haut à droite a perdu une ligne et une colonne et que l'on a rajouté une ligne de coefficients en bas.

Par un raisonnement par récurrence, on obtient donc un "remplissage" de la matrice par le bas. Au bout de N multiplications, on obtient donc :

$$\begin{pmatrix} -A_{\beta-N, \beta-2N} & \dots & \dots & -A_{\beta-N, \beta-N-1} & z - A_{\beta-N, \beta-N} & -A_{\beta-N, \beta-N+1} & \dots & -A_{\beta-N, \beta-1} \\ 0 & & & & & & & \\ \vdots & \ddots & & & & & & \\ 0 & \dots & 0 & -A_{\beta-1, \beta-N-1} & -A_{\beta-1, \beta-N} & \dots & \dots & z - A_{\beta-1, \beta-1} \end{pmatrix}$$

c'est-à-dire que $\begin{pmatrix} C_{\beta}^L K_{\beta-N+1}^L & C_{\beta}^L N_{\beta-N+1}^L \end{pmatrix} = \begin{pmatrix} -E_{k-2}^* & zI_N - V_{k-1} \end{pmatrix}$.

On obtient donc $Y_{kN-N+1}^L = W + zI_N - V_{k-1}$ et $Z_{kN-N+1}^L = -E_{k-2}^*$. Donc la composition des N homographies simples donne

$$\begin{aligned} s_{kN-N+1}^L \circ \dots \circ s_{kN}^L &= C_{kN-N}^L (zI_N - V_{k-1} + W)^{-1} (-E_{k-2}^*) \\ &= E_{k-2} (zI_N - V_{k-1} + W)^{-1} (-E_{k-2}^*) \\ &= -E_{k-2} (zI_N - V_{k-1} + W)^{-1} E_{k-2}^* \end{aligned}$$

ce qui est la formule donnée pour l'homographie-bloc. On retrouve ainsi la formule donnée par Aptekarev et Nikishin.

Au début de cette sous-section de la démonstration, nous avons indiqué que nous pourrions toujours mettre la matrice sous forme d'une matrice bande. En fait, il suffit de réaliser des décompositions LU successives des E_j sous la forme $L_j R_j$ où L_j est une matrice triangulaire inférieure et R_j une matrice triangulaire supérieure. On obtient alors une matrice bande. Il suffit de changer les polynômes P_k^L et Q_k^L . \square

2.3 Accélération de convergence

La convergence de fractions continues a été étudiée par de nombreux auteurs : De Bruin [23], Jacobsen [24, 25, 44], Jones et Waadeland [45], Levrie [50, 51, 52, 53], Kaliaguine [47], Van Iseghem [81] et Sorokin [72]. L'accélération tient une grande part dans ces différents travaux. Dans la même optique, il s'agit dans cette partie de démontrer le résultat suivant d'accélération de convergence de fractions continues matricielles en utilisant les nouveaux outils créés dans la section 2.1 :

Theorème 2.8. *Soient $A_{n, n+p}$, $n \geq 0$, $p \in \{-s, \dots, r\}$ de telle sorte que*

$$\lim_{n \rightarrow \infty} A_{n, n+p} := h_p$$

existe pour $p = -s, \dots, r$ et $h_{-s}.h_r \neq 0$. On supposera que les valeurs propres de la matrice A^L définie dans 1.42 vérifient $|\lambda_1| < |\lambda_2| < \dots < |\lambda_{r+s}|$.

De plus, on suppose que

$$\limsup_{n \rightarrow \infty} \|S_n^L(0)\| \neq \infty$$

Alors la limite

$$\varphi := \lim_{n \rightarrow \infty} S_n^L(0)$$

existe, et si on pose

$$\tilde{\varphi} := - \begin{pmatrix} h_r & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 \\ h_1 & \dots & \dots & h_r \end{pmatrix} \cdot \begin{pmatrix} \lambda_1^s & \dots & \lambda_s^s \\ \vdots & & \vdots \\ \lambda_1^{s+r-1} & \dots & \lambda_s^{s+r-1} \end{pmatrix} \cdot \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ \lambda_1^{s-1} & \dots & \lambda_s^{s-1} \end{pmatrix}^{-1}$$

nous obtenons l'accélération pour tous $x \in \mathbb{C}^{r \times 1}$ et $y \in \mathbb{C}^{s \times 1}$

$$\lim_{n \rightarrow \infty} \frac{|x^t(S_n^L(\tilde{\varphi}) - \varphi)y|}{|x^t(S_n^L(0) - \varphi)y|} = 0.$$

Remarque 2.9. En particulier, si x et y sont des vecteurs canoniques, on obtient une accélération par composante. Grâce à la propriété 1 du Chapitre II de la thèse de B. Germain-Bonne[39], cela fournit alors une accélération en norme

$$\lim_{n \rightarrow \infty} \frac{\|S_n^L(\tilde{\varphi}) - \varphi\|}{\|S_n^L(0) - \varphi\|} = 0$$

Le théorème précédent est une uniformisation d'un théorème donné par Jacobsen et De Bruin [25, Théorème 5, p. 168] :

Théorème 2.10. Soit une fraction généralisée convergente dans \mathbb{C}^n

$$K_{k=1}^\infty \begin{pmatrix} a_k^{(1)} \\ \vdots \\ a_k^{(s)} \\ b_k \end{pmatrix}$$

où les coefficients vérifient

$$\lim_{k \rightarrow +\infty} a_k^{(i)} = a^{(i)} \in \mathbb{C} \quad (i = 1, \dots, s), \quad \lim_{k \rightarrow +\infty} b_k = b \in \mathbb{C}$$

et tels que les solutions r_1, \dots, r_{s+1} de

$$r^{s+1} = br^s + a^{(s)}r^{s-1} + a^{(s-1)}r^{s-2} + \dots + a^{(2)}r + a^{(1)}$$

satisfassent $|r_1| > |r_2| > \dots > |r_s| > |r_{s+1}|$.

Alors les queues des fractions continues convergent vers $w^{(i)}$ ($i = 1, \dots, s$) et on a l'accélération de convergence suivante

$$\lim_{k \rightarrow +\infty} \frac{\xi_0^{(i)} - S_k^{(i)}(w^{(1)}, \dots, w^{(s)})}{\xi_0^{(i)} - S_k^{(i)}(0, \dots, 0)} = 0 \quad (i = 1, \dots, s)$$

si les $\xi_0^{(i)} \neq \infty$, ($i = 1, \dots, s$) sont les valeurs des fractions continues.

Toutes les formules de la section 2.1 dépendent de z . Dans la suite, on peut supposer que $z = 0$.

Avant de démontrer le théorème, on rappelle le théorème dit de Poincaré-Pincherle-Perron [24, Théorème 1.1] qui est un outil majeur de la démonstration. Ce théorème a déjà été utilisé dans [23], [44], [50], [55] pour démontrer la convergence de fractions continues généralisées et leur accélération de convergence.

Théorème 2.11. *On considère une relation de récurrence linéaire de la forme*

$$X_k = a_k^{(n)} X_{k-1} + a_k^{(n-1)} X_{k-2} + \dots + a_k^{(1)} X_{k-n}, \quad a_k^{(1)} \neq 0; k \geq 1$$

tels que les coefficients vérifient $\lim_{k \rightarrow +\infty} a_k^{(j)} = a^{(j)}$, ($j = 1, \dots, n$).

et les racines r_1, \dots, r_n de l'équation auxiliaire $r^n = a^{(n)} r^{n-1} + \dots + a^{(1)}$ soient simples et ordonnées de la manière suivante :

$$|r_1| < |r_2| < \dots < |r_n|$$

Alors il existe une base $\{D_k^{(j)}\} (j = 1, \dots, n)$ de l'ensemble des solutions de la récurrence à n termes qui vérifient

$$\lim_{k \rightarrow +\infty} \frac{D_{k+1}^{(j)}}{D_k^{(j)}} = r_j, \quad (j = 1, \dots, n)$$

Cette base est dite ordonnée par domination car, ainsi qu'on le démontre plus loin,

$$\lim_{k \rightarrow +\infty} \frac{D_k^{(i)}}{D_k^{(j)}} = 0, \quad i < j$$

On démontre maintenant le théorème principal 2.8

Démonstration. Rappelons [79] que les composantes de Q_n^L et P_n^L forment une base de l'ensemble des solutions de la récurrence (où on a choisi $z = 0$)

$$y_{n+r} = -\frac{A_{n,n-s}}{A_{n,n+r}} y_{n-s} - \dots - \frac{A_{n,n+r-1}}{A_{n,n+r}} y_{n+r-1}, \quad n \geq 0. \quad (2.1)$$

Par hypothèse, les coefficients de cette récurrence admettent des limites

$$\lim_{n \rightarrow \infty} \frac{A_{n,n+p}}{A_{n,n+r}} = \frac{h_p}{h_r}$$

Remarquons aussi que les λ_j coïncident avec les racines de l'équation caractéristique

$$\lambda^r = -\frac{h_{-s}}{h_r} \lambda^{-s} - \dots - \frac{h_{r-1}}{h_r} \lambda^{r-1}.$$

Le théorème de Poincaré-Pincherle-Perron 2.11 permet d'affirmer qu'il existe une base $(D_n^{(i)})_{n \geq 0, i = 1, \dots, r+s}$ de l'ensemble des solutions de (2.1) telle que

$$i = 1, \dots, r+s, \quad \lim_{n \rightarrow \infty} \frac{D_{n+1}^{(i)}}{D_n^{(i)}} = \lambda_i. \quad (2.2)$$

Pour cette base, nous avons le lemme suivant montrant qu'il s'agit effectivement d'une base ordonnée par domination.

Lemme 2.12. *Pour $i < j$: $\lim_{n \rightarrow \infty} \frac{D_n^{(i)}}{D_n^{(j)}} = 0$, la décroissance étant géométrique.*

Démonstration. On sait que $\lim_{k \rightarrow +\infty} D_{k+1}^{(i)}/D_k^{(i)} = \lambda_i$, ($i = 1, \dots, r+s$) donc soit $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que :

$$\forall n \in \mathbb{N}, (n > N \Rightarrow |\lambda_i| - \varepsilon < \left| \frac{D_{n+1}^{(i)}}{D_n^{(i)}} \right| < |\lambda_i| + \varepsilon)$$

Donc $(|\lambda_i| - \varepsilon)|D_n^{(i)}| < |D_{n+1}^{(i)}| < (|\lambda_i| + \varepsilon)|D_n^{(i)}|$.

Par récurrence, on obtient

$$\begin{aligned} (|\lambda_i| - \varepsilon)^{(n-N)} |D_{N+1}^{(i)}| &< |D_{n+1}^{(i)}| < (|\lambda_i| + \varepsilon)^{(n-N)} |D_{N+1}^{(i)}|. \\ \text{et } (|\lambda_j| - \varepsilon)^{(n-N)} |D_{N+1}^{(j)}| &< |D_{n+1}^{(j)}| < (|\lambda_j| + \varepsilon)^{(n-N)} |D_{N+1}^{(j)}|. \end{aligned}$$

Finalement, pour $i < j$, $|\lambda_i| < |\lambda_j|$,

$$\frac{|D_{N+1}^{(i)}|}{|D_{N+1}^{(j)}|} \left(\frac{|\lambda_i| - \varepsilon}{|\lambda_j| + \varepsilon} \right)^{(n-N)} < \frac{|D_{n+1}^{(i)}|}{|D_{n+1}^{(j)}|} < \frac{|D_{N+1}^{(i)}|}{|D_{N+1}^{(j)}|} \left(\frac{|\lambda_i| + \varepsilon}{|\lambda_j| - \varepsilon} \right)^{(n-N)}$$

Supposons

$$\varepsilon < \varepsilon_0 := \frac{1}{2} \min_{k \in \{1, \dots, r+s-1\}} (|\lambda_{k+1}| - |\lambda_k|)$$

Alors, pour $j \geq 2$, $|\lambda_j| - \varepsilon > |\lambda_j| - \frac{1}{2}(|\lambda_j| - |\lambda_{j-1}|) > \frac{1}{2}(|\lambda_j| + |\lambda_{j-1}|) > 0$

$$\begin{aligned} \forall 1 < j \leq r+s, \quad 2\varepsilon &< |\lambda_j| - |\lambda_{j-1}| \\ |\lambda_{j-1}| + \varepsilon &< |\lambda_j| - \varepsilon \\ \frac{|\lambda_{j-1}| + \varepsilon}{|\lambda_j| - \varepsilon} &< 1 \end{aligned}$$

Donc $q = \max_{j=2, \dots, r+s} \frac{|\lambda_{j-1}| + \varepsilon}{|\lambda_j| - \varepsilon} < 1$. De plus,

$$\forall 1 \leq i < j \leq r+s, \quad \frac{|\lambda_i| + \varepsilon}{|\lambda_j| - \varepsilon} \leq \frac{|\lambda_{j-1}| + \varepsilon}{|\lambda_j| - \varepsilon} \leq q$$

On obtient donc $\forall 1 \leq i < j \leq r+s$, $\frac{|D_{n+1}^{(i)}|}{|D_{n+1}^{(j)}|} \leq q^{(n-N)}$.

Et finalement $\forall 1 \leq i < j \leq r+s$, $\lim_{n \rightarrow +\infty} \frac{D_{n+1}^{(i)}}{D_{n+1}^{(j)}} = 0$. □

Comme on dispose de deux bases, il existe une matrice P , inversible, telle que, pour tout $n \geq 0$,

$$\begin{aligned} B_n^L &= \begin{pmatrix} Q_{n-s}^L & -P_{n-s}^L \\ \vdots & \vdots \\ Q_{n+r-1}^L & -P_{n+r-1}^L \end{pmatrix} = \begin{pmatrix} D_{n-s}^{(1)} & \cdots & D_{n-s}^{(r+s)} \\ \vdots & & \vdots \\ D_{n+r-1}^{(1)} & \cdots & D_{n+r-1}^{(r+s)} \end{pmatrix} P \\ &= V^{(n-s)} \cdot \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(r+s)}) \cdot P \end{aligned}$$

où

$$V^{(n-s)} = \begin{pmatrix} 1 & \cdots & 1 \\ \frac{D_{n-s+1}^{(1)}}{D_{n-s}^{(1)}} & \cdots & \frac{D_{n-s+1}^{(r+s)}}{D_{n-s}^{(r+s)}} \\ \vdots & & \vdots \\ \frac{D_{n+r-1}^{(1)}}{D_{n-s}^{(1)}} & \cdots & \frac{D_{n+r-1}^{(r+s)}}{D_{n-s}^{(r+s)}} \end{pmatrix}$$

On partitionne P et $V^{(n-s)}$ de la manière suivante :

$$P = \begin{pmatrix} P_{1,1} (\in \mathbb{C}^{s \times r}) & P_{1,2} (\in \mathbb{C}^{s \times s}) \\ P_{2,1} (\in \mathbb{C}^{r \times r}) & P_{2,2} (\in \mathbb{C}^{r \times s}) \end{pmatrix}$$

$$\text{et } V^{(n-s)} = \begin{pmatrix} V_{1,1}^{(n-s)} (\in \mathbb{C}^{s \times s}) & V_{1,2}^{(n-s)} (\in \mathbb{C}^{s \times r}) \\ V_{2,1}^{(n-s)} (\in \mathbb{C}^{r \times s}) & V_{2,2}^{(n-s)} (\in \mathbb{C}^{r \times r}) \end{pmatrix}$$

On peut remarquer que $\lim_{n \rightarrow \infty} V^{(n-s)} = V^{(\infty)}$ existe d'après l'égalité (2.2). C'est une matrice de Vandermonde construite à partir des $\lambda_1, \dots, \lambda_{r+s}$ qui sont distincts et non nuls. En particulier, $V^{(\infty)}$, $V_{1,1}^{(\infty)}$ et $V_{2,2}^{(\infty)}$ sont inversibles. De ce fait, à partir d'un certain rang n , les matrices $V_{1,1}^{(n-s)}$ et $V_{2,2}^{(n-s)}$ sont inversibles et leurs inverses convergent vers $(V_{1,1}^{(\infty)})^{-1}$ et $(V_{2,2}^{(\infty)})^{-1}$.

Nous avons le résultat suivant

Lemme 2.13. *Si $\limsup_{n \rightarrow \infty} \|S_n^L(0)\| < \infty$ alors $P_{2,1}$ est inversible. De plus,*

$$\lim_{n \rightarrow \infty} S_n^L(0) = \varphi := -(P_{2,1})^{-1} P_{2,2}$$

Démonstration. On a

$$\begin{aligned} &\begin{pmatrix} Q_n^L & -P_n^L \\ \vdots & \vdots \\ Q_{n+r-1}^L & -P_{n+r-1}^L \end{pmatrix} = (V_{2,1}^{(n-s)} \quad V_{2,2}^{(n-s)}) \cdot \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(r+s)}) \cdot P \\ &= V_{2,2}^{(n-s)} \cdot \left((V_{2,2}^{(n-s)})^{-1} V_{2,1}^{(n-s)} \quad I_r \right) \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(r+s)}) \cdot P \\ &= V_{2,2}^{(n-s)} \cdot \text{diag}(D_{n-s}^{(s+1)}, \dots, D_{n-s}^{(r+s)}) \cdot \begin{pmatrix} \Delta_n & I_r \end{pmatrix} \cdot P \end{aligned}$$

avec $\|\Delta_n\| \rightarrow 0$ d'après le lemme 2.12. Supposons, maintenant, par l'absurde, qu'il existe $x \in \mathbb{C}^r \neq 0$ tel que $x^t.P_{2,1} = 0$. Comme P est inversible, nous avons $x^t.P_{2,2} \neq 0$. Nous obtenons alors

$$\begin{aligned} \|S_n^L(0)\| &= \sup_{y \in \mathbb{C}^r} \frac{\|y^t \mathcal{P}_{n,r}^L\|}{\|y^t \mathcal{Q}_{n,r}^L\|} \\ &= \sup_{y \in \mathbb{C}^r} \frac{\|y^t \cdot (\Delta_n P_{1,2} + P_{2,2})\|}{\|y^t \cdot (\Delta_n P_{1,1} + P_{2,1})\|} \\ &\geq \frac{\|x^t P_{2,2}\| - \|x^t \Delta_n P_{1,2}\|}{\|x\| \cdot \|\Delta_n\| \cdot \|P_{1,1}\|} \end{aligned}$$

et la dernière quantité tend vers l'infini lorsque n tend vers l'infini, ce qui est en contradiction avec l'hypothèse.

Ceci signifie donc que $P_{2,1}$ est inversible. De plus

$$\begin{aligned} -(P_{2,1})^{-1}P_{2,2} - S_n^L(0) &= -(P_{2,1})^{-1}P_{2,2} - (\mathcal{Q}_{n,r}^L)^{-1}\mathcal{P}_{n,r}^L \\ &= (\mathcal{Q}_{n,r}^L)^{-1}(-\mathcal{Q}_{n,r}^L(P_{2,1})^{-1}P_{2,2} - \mathcal{P}_{n,r}^L) \\ &= (\mathcal{Q}_{n,r}^L)^{-1} \begin{pmatrix} \mathcal{Q}_{n,r}^L & -\mathcal{P}_{n,r}^L \end{pmatrix} \begin{pmatrix} -(P_{2,1})^{-1}P_{2,2} \\ I_s \end{pmatrix} \\ &= (\Delta_n P_{1,1} + P_{2,1})^{-1} \cdot \begin{pmatrix} \Delta_n & I_r \end{pmatrix} \cdot P \begin{pmatrix} -(P_{2,1})^{-1}P_{2,2} \\ I_s \end{pmatrix} \\ &\quad (\text{d'après l'expression de } \mathcal{Q}_{n,r}^L) \\ &= (\Delta_n P_{1,1} + P_{2,1})^{-1} \cdot \begin{pmatrix} \Delta_n & I_r \end{pmatrix} \cdot \begin{pmatrix} P_{1,2} - P_{1,1}(P_{2,1})^{-1}P_{2,2} \\ 0 \end{pmatrix} \\ &= (\Delta_n P_{1,1} + P_{2,1})^{-1}[\Delta_n(P_{1,2} - P_{1,1}(P_{2,1})^{-1}P_{2,2})] \end{aligned}$$

Et la dernière quantité tend vers 0 lorsque n tend vers l'infini. \square

Si l'on pose $R_n^R = \varphi Q_n^R - P_n^R$ et $R_n^L = Q_n^L \varphi - P_n^L$, on obtient la représentation suivante qui permettra d'obtenir l'asymptotique de l'erreur :

Lemme 2.14. Notons $\hat{P} := P_{1,2} - P_{1,1}(P_{2,1})^{-1}P_{2,2}$, alors

$$\begin{aligned} \begin{pmatrix} Q_n^L & R_n^L \\ \vdots & \vdots \\ Q_{n+r-1}^L & R_{n+r-1}^L \end{pmatrix} &= V^{(n-s)} \cdot \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(r+s)}) \cdot \begin{pmatrix} P_{1,1} & \hat{P} \\ P_{2,1} & 0 \end{pmatrix} \\ \begin{pmatrix} -R_{n-r}^R, \dots, -R_{n+s-1}^R \\ Q_{n-r}^R, \dots, Q_{n+s-1}^R \end{pmatrix} &= \begin{pmatrix} 0 & P_{2,1}^{-1} \\ \hat{P}^{-1} & -\hat{P}^{-1}P_{1,1}P_{2,1}^{-1} \end{pmatrix} \cdot \text{diag}\left(\frac{1}{D_{n-s}^{(1)}}, \dots, \frac{1}{D_{n-s}^{(r+s)}}\right) \\ &\quad W^{(n-s)} \begin{pmatrix} 0 & (C_n^R)^{-1} \\ -(C_n^L)^{-1} & 0 \end{pmatrix} \end{aligned}$$

où $W^{(n-s)}$ est l'inverse de $V^{(n-s)}$ et est partitionnée de la même manière.

Démonstration. La première formule se déduit de la démonstration du lemme précédent.

La deuxième s'obtient en utilisant la formule (1.6) et en multipliant à gauche par $\begin{pmatrix} I_r & -\varphi \\ 0 & I_s \end{pmatrix}$ et à droite par $\begin{pmatrix} I_r & \varphi \\ 0 & I_s \end{pmatrix}$. \square

Ce lemme nous permet d'obtenir les asymptotiques suivantes :

Lemme 2.15. $\forall x \in \mathbb{C}^r, \forall y \in \mathbb{C}^s$, en notant $\tilde{x}^t := x^t \cdot P_{2,1}^{-1}$ et $\tilde{y} = \hat{P}y$ et α et β les indices de la première coordonnée non nulle du vecteur \tilde{x} et de la dernière coordonnée non nulle du vecteur \tilde{y} :

$$\tilde{x} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \tilde{x}_\alpha \neq 0 \\ \vdots \end{pmatrix} \quad \text{et} \quad \tilde{y} = \begin{pmatrix} \vdots \\ \tilde{y}_\beta \neq 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\text{On a } \lim_{n \rightarrow \infty} \frac{|x^t(\varphi - S_n^L(0)) \cdot y|}{|(D_{n-s}^{(s+\alpha)})^{-1} \cdot D_{n-s}^{(\beta)}|} \in \mathbb{R} - \{0\}$$

Démonstration. Par définition et d'après le lemme précédent, on a

$$\begin{aligned} & \varphi - S_n^L(0) \\ &= \mathcal{R}_{n,s}^R (\mathcal{Q}_{n,s}^R)^{-1} \\ &= \mathcal{R}_{n,s}^R (\mathcal{R}_{n-s,s}^L \mathcal{Q}_{n,s}^R)^{-1} \mathcal{R}_{n-s,s}^L \\ &= -P_{2,1}^{-1} \text{diag} \left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(r+s)}} \right) W_{2,1}^{(n-s)} (C_n^R)^{-1} C_n^R \\ & \cdot \left(V_{1,1}^{(n-s)} \cdot [W_{1,1}^{(n-s)} - \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) P_{1,1} P_{2,1}^{-1} \text{diag} \left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(s+r)}} \right) W_{2,1}^{(n-s)}] \right)^{-1} \\ & \cdot V_{1,1}^{(n-s)} \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) \hat{P} \end{aligned}$$

D'après le lemme 2.12, on a

$$\lim_{n \rightarrow \infty} \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) P_{1,1} P_{2,1}^{-1} \text{diag} \left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(s+r)}} \right) = 0$$

De plus, la matrice $W_{2,1}^{(n-s)}$ converge vers $W_{2,1}^{(\infty)}$. Donc à partir d'un certain rang, $W_{2,1}^{(n-s)}$ est bornée. Cela signifie donc que, à partir d'un certain rang, on obtient

$$\begin{aligned} \varphi - S_n^L(0) &= -P_{2,1}^{-1} \text{diag} \left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(r+s)}} \right) W_{2,1}^{(n-s)} (V_{1,1}^{(n-s)} (W_{1,1}^{(n-s)} + o(1)))^{-1} \\ & \quad V_{1,1}^{(n-s)} \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) \hat{P} \end{aligned}$$

On a déjà démontré que les matrices $V_{1,1}^{(n-s)}$ et $V_{2,2}^{(n-s)}$ étaient inversibles à partir d'un certain rang. D'après la définition de $W^{(n-s)}$, on a

$$W_{1,1}^{(n-s)}(V_{1,1}^{(n-s)} - V_{1,2}^{(n-s)}(V_{2,2}^{(n-s)})^{-1}V_{2,1}^{(n-s)}) = I_s$$

Cela signifie donc, qu'à partir d'un certain rang, la matrice $W_{1,1}^{(n-s)}$ est inversible d'inverse $V_{1,1}^{(n-s)} - V_{1,2}^{(n-s)}(V_{2,2}^{(n-s)})^{-1}V_{2,1}^{(n-s)}$. Cela nous fournit alors l'égalité :

$$\begin{aligned} & \varphi - S_n^L(0) \\ &= -P_{2,1}^{-1} \text{diag} \left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(r+s)}} \right) W_{2,1}^{(n-s)} (V_{1,1}^{(n-s)} - V_{1,2}^{(n-s)}(V_{2,2}^{(n-s)})^{-1}V_{2,1}^{(n-s)} + o(1)) \\ & \quad \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) \hat{P} \\ &= -P_{2,1}^{-1} \text{diag} \left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(r+s)}} \right) \\ & \quad (-W_{2,2}^{(n-s)}V_{2,1}^{(n-s)} - (I_r - W_{2,2}^{(n-s)}V_{2,2}^{(n-s)})(V_{2,2}^{(n-s)})^{-1}V_{2,1}^{(n-s)} + o(1)) \\ & \quad \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) \hat{P} \\ &= -P_{2,1}^{-1} \text{diag} \left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(r+s)}} \right) (- (V_{2,2}^{(n-s)})^{-1}V_{2,1}^{(n-s)} + o(1)) \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) \hat{P} \end{aligned}$$

Le produit $-(V_{2,2}^{(n-s)})^{-1}V_{2,1}^{(n-s)}$ possède des composantes toutes non nulles puisqu'il s'agit d'interpolation polynomiale. En effet, les matrices limites sont des matrices de VanderMonde, c'est-à-dire que l'on obtient le produit

$$\begin{pmatrix} \lambda_{r+1}^r & \dots & \lambda_{r+s}^r \\ \vdots & & \vdots \\ \lambda_{r+1}^{r+s-1} & \dots & \lambda_{r+s}^{r+s-1} \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1^r & \dots & \lambda_r^r \\ \vdots & & \vdots \\ \lambda_1^{r+s-1} & \dots & \lambda_r^{r+s-1} \end{pmatrix}$$

soit

$$\text{diag}(\lambda_{r+1}^{-r}, \dots, \lambda_{r+s}^{-r}) \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ \lambda_{r+1}^{s-1} & \dots & \lambda_{r+s}^{s-1} \end{pmatrix}^{-1} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ \lambda_1^{s-1} & \dots & \lambda_r^{s-1} \end{pmatrix} \text{diag}(\lambda_1^r, \dots, \lambda_r^r) \quad (2.3)$$

On recherche donc l'inverse d'une matrice de VanderMonde. Notons cet inverse

$$\begin{pmatrix} d_{11} & \dots & d_{1s} \\ \vdots & & \vdots \\ d_{s1} & \dots & d_{ss} \end{pmatrix}$$

Notons $T_k(X) = d_{k1} + \dots + d_{ks}X^{s-1}$, les polynômes de degré au plus $s - 1$ formés à partir des coefficients de cette matrice. Ils vérifient

$$T_k(\lambda_{r+k}) = 1 \quad T_k(\lambda_{r+n}) = 0, \quad n \in \{1, \dots, s\}, \quad n \neq k$$

Ces polynômes d'interpolation existent donc et sont uniques. Ils ne peuvent être nuls en d'autres racines que celles indiquées ci-dessus. Or lorsque l'on effectue le produit (2.3), on obtient

$$\text{diag}(\lambda_{r+1}^{-r}, \dots, \lambda_{r+s}^{-r}) \begin{pmatrix} T_1(\lambda_1) & \dots & T_1(\lambda_r) \\ \vdots & & \vdots \\ T_s(\lambda_1) & \dots & T_s(\lambda_r) \end{pmatrix} \text{diag}(\lambda_1^r, \dots, \lambda_r^r) \quad (2.4)$$

qui est la limite du produit $-(V_{2,2}^{(n-s)})^{-1}V_{2,1}^{(n-s)}$. Ceci qui prouve que toutes les entrées de ce produit sont non nulles à partir d'un certain rang.

Notons le produit $-(V_{2,2}^{(n-s)})^{-1}V_{2,1}^{(n-s)} + o(1)$ par

$$\begin{pmatrix} \gamma_{1,1} & \dots & \gamma_{1,s} \\ \vdots & & \vdots \\ \gamma_{r,1} & \dots & \gamma_{r,s} \end{pmatrix}$$

On multiplie ensuite par les deux diagonales, on obtient alors

$$\begin{pmatrix} \gamma_{1,1}(D_{n-s}^{(s+1)})^{-1}D_{n-s}^{(1)} & \dots & \gamma_{1,s}(D_{n-s}^{(s+1)})^{-1}D_{n-s}^{(s)} \\ \vdots & & \vdots \\ \gamma_{r,1}(D_{n-s}^{(s+r)})^{-1}D_{n-s}^{(1)} & \dots & \gamma_{r,s}(D_{n-s}^{(s+r)})^{-1}D_{n-s}^{(s)} \end{pmatrix}$$

Notons $(c_1, \dots, c_r)^t = \tilde{x}^t$ et $(d_1, \dots, d_s)^t = \tilde{y}^t$. On multiplie ensuite à gauche et à droite par x^t et y pour obtenir

$$|x^t \mathcal{R}_{n,s}^R (\mathcal{R}_{n-s,s}^L \mathcal{Q}_{n,s}^R)^{-1} \mathcal{R}_{n-s,s}^L y| = \left| \sum_{j=1}^s d_j \left(\sum_{k=1}^r c_k \gamma_{k,j} (D_{n-s}^{(s+k)})^{-1} D_{n-s}^{(j)} \right) \right|$$

Divisons par $|(D_{n-s}^{(s+\alpha)})^{-1}D_{n-s}^{(\beta)}|$. Comme $c_k = 0$, $k < \alpha$ et $d_j = 0$, $j > \beta$, on obtient finalement la quantité

$$\left| \sum_{j=1}^{\beta} d_j \left(\sum_{k=\alpha}^r c_k \gamma_{k,j} \frac{D_{n-s}^{(s+\alpha)}}{D_{n-s}^{(s+k)}} \frac{D_{n-s}^{(j)}}{D_{n-s}^{(\beta)}} \right) \right|$$

Grâce au lemme 2.12 sur la base ordonnée par domination et d'après la limite (2.4), on obtient donc pour limite de la somme précédente

$$c_\alpha d_\beta \lim_{n \rightarrow \infty} \gamma_{\alpha,\beta} = c_\alpha d_\beta \lambda_{r+\alpha}^{-r} T_\alpha(\lambda_\beta) \lambda_\beta^r$$

qui est bien un nombre réel non nul. \square

Lemme 2.16. *Nous avons $\varphi_n = -C_n^L \mathcal{R}_{n,r}^L (\mathcal{R}_{n-s,s}^L)^{-1} = -(\mathcal{R}_{n-r,r}^R)^{-1} \mathcal{R}_{n,s}^R C_n^R$ et $S_n^L(\varphi_n) = \varphi$, (c'est-à-dire que φ_n peut être considéré comme la queue d'une fraction continue).*

Finalemnt, $\lim_{n \rightarrow \infty} \varphi_n = \tilde{\varphi}$ existe.

Démonstration. Développons le produit croisé $\mathcal{R}_{n,s}^R C_n^R \mathcal{R}_{n-s,s}^L$.

$$\begin{aligned} \mathcal{R}_{n,s}^R C_n^R \mathcal{R}_{n-s,s}^L &= (\varphi \mathcal{Q}_{n,s}^R - \mathcal{P}_{n,s}^R) C_n^R (\mathcal{Q}_{n-s,s}^L \varphi - \mathcal{P}_{n-s,s}^L) \\ &= \varphi \mathcal{Q}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L \varphi - \mathcal{P}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L \varphi \\ &\quad - \varphi \mathcal{Q}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L + \mathcal{P}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L \end{aligned}$$

D'après la propriété 2.1, on obtient alors

$$\begin{aligned} \mathcal{R}_{n,s}^R C_n^R \mathcal{R}_{n-s,s}^L &= \varphi \mathcal{Q}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L \varphi - (I_r + \mathcal{P}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L) \varphi \\ &\quad - \varphi (\mathcal{Q}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L - I_s) + \mathcal{P}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L \\ &= \varphi \mathcal{Q}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L \varphi - \mathcal{P}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L \varphi \\ &\quad - \varphi \mathcal{Q}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L + \mathcal{P}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L \\ &= (\varphi \mathcal{Q}_{n-r,r}^R - \mathcal{P}_{n-r,r}^R) C_n^L (\mathcal{Q}_{n,r}^L \varphi - \mathcal{P}_{n,r}^L) \\ &= \mathcal{R}_{n-r,r}^R C_n^L \mathcal{R}_{n,r}^L \end{aligned}$$

De plus, d'après le lemme 2.14, on en déduit que les résidus sont inversibles à partir d'un certain rang, ce qui conclut la démonstration de l'égalité.

Considérons maintenant $S_n^L(\varphi_n)$. D'après les formules de la définition 2.2, on obtient :

$$S_n^L(\varphi_n) = (\varphi_n \mathcal{Q}_{n-s,s}^L + C_n^L \mathcal{Q}_{n,r}^L)^{-1} (\varphi_n \mathcal{P}_{n-s,s}^L + C_n^L \mathcal{P}_{n,r}^L)$$

Utilisons la forme de $\varphi_n = -(\mathcal{R}_{n-r,r}^R)^{-1} \mathcal{R}_{n,s}^R C_n^R$. Cela nous donne

$$\begin{aligned} S_n^L(\varphi_n) &= (-\mathcal{R}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L + \mathcal{R}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L)^{-1} (-\mathcal{R}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L + \mathcal{R}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L) \\ &= (\varphi (-\mathcal{Q}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L + \mathcal{Q}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L) - (-\mathcal{P}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L + \mathcal{P}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L))^{-1} \\ &\quad (\varphi (-\mathcal{Q}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L + \mathcal{Q}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L) - (-\mathcal{P}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L + \mathcal{P}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L)) \\ &= \varphi \quad \text{d'après la propriété 2.1} \end{aligned}$$

En utilisant maintenant la formule de φ_n , on peut en déterminer la limite :

$$\varphi_n = -C_n^L \mathcal{R}_{n,r}^L (\mathcal{R}_{n-s,s}^L)^{-1}$$

D'après les limites sur les diagonales, on trouve

$$\lim_{n \rightarrow \infty} C_n^L = \begin{pmatrix} h_r & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 \\ h_1 & \dots & \dots & h_r \end{pmatrix}$$

De plus

$$\begin{aligned}\mathcal{R}_{n,r}^L &= V_{2,1}^{(n-s)} \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) \hat{P} \\ \mathcal{R}_{n-s,s}^L &= V_{1,1}^{(n-s)} \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) \hat{P}\end{aligned}$$

On forme alors le produit $\mathcal{R}_{n,r}^L (\mathcal{R}_{n-s,s}^L)^{-1} = V_{2,1}^{(n-s)} (V_{1,1}^{(n-s)})^{-1}$ et on passe à la limite. On obtient bien la formule demandée. \square

Remarque 2.17. On peut aussi déterminer la forme réduite de φ_n en utilisant $\varphi_n = -(\mathcal{R}_{n-r,r}^R)^{-1} \mathcal{R}_{n,s}^R C_n^R$.

En effet, on a

$$\begin{aligned}\mathcal{R}_{n-r,r}^R &= P_{2,1}^{-1} \text{diag}\left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(s+r)}}\right) W_{2,2}^{(n-s)} (C_n^L)^{-1} \\ \mathcal{R}_{n,s}^R &= -P_{2,1}^{-1} \text{diag}\left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(s+r)}}\right) W_{2,1}^{(n-s)} (C_n^R)^{-1}\end{aligned}$$

Ce qui nous donne finalement $\varphi_n = C_n^L (W_{2,2}^{(n-s)})^{-1} W_{2,1}^{(n-s)}$.

Il s'agit maintenant de démontrer le théorème 2.8. On va donc étudier le numérateur $\varphi - S_n^L(\tilde{\varphi})$

$$\begin{aligned}& \varphi - S_n^L(\tilde{\varphi}) \\ &= S_n^L(\varphi_n) - S_n^R(\tilde{\varphi}) \\ &= S_n^L((\mathcal{R}_{n-r,r}^R)^{-1} (-\mathcal{R}_{n,s}^R) C_n^R) - S_n^R(\tilde{C}_n^L (-\tilde{\mathcal{R}}_{n,r}^L (\tilde{\mathcal{R}}_{n-s,s}^L)^{-1})) \\ &= (-\mathcal{R}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L + \mathcal{R}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L)^{-1} (-\mathcal{R}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L + \mathcal{R}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L) \\ &\quad - (-\mathcal{P}_{n-r,r}^R \tilde{C}_n^L \tilde{\mathcal{R}}_{n,r}^L + \mathcal{P}_{n,s}^R C_n^R \tilde{\mathcal{R}}_{n-s,s}^L) (-\mathcal{Q}_{n-r,r}^R \tilde{C}_n^L \tilde{\mathcal{R}}_{n,r}^L + \mathcal{Q}_{n,s}^R C_n^R \tilde{\mathcal{R}}_{n-s,s}^L)^{-1} \\ &= (-\mathcal{R}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L + \mathcal{R}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L)^{-1} \\ &\quad \left[(-\mathcal{R}_{n,s}^R C_n^R \mathcal{P}_{n-s,s}^L + \mathcal{R}_{n-r,r}^R C_n^L \mathcal{P}_{n,r}^L) (-\mathcal{Q}_{n-r,r}^R \tilde{C}_n^L \tilde{\mathcal{R}}_{n,r}^L + \mathcal{Q}_{n,s}^R C_n^R \tilde{\mathcal{R}}_{n-s,s}^L) \right. \\ &\quad \left. - (-\mathcal{R}_{n,s}^R C_n^R \mathcal{Q}_{n-s,s}^L + \mathcal{R}_{n-r,r}^R C_n^L \mathcal{Q}_{n,r}^L) (-\mathcal{P}_{n-r,r}^R \tilde{C}_n^L \tilde{\mathcal{R}}_{n,r}^L + \mathcal{P}_{n,s}^R C_n^R \tilde{\mathcal{R}}_{n-s,s}^L) \right] \\ &\quad (-\mathcal{Q}_{n-r,r}^R \tilde{C}_n^L \tilde{\mathcal{R}}_{n,r}^L + \mathcal{Q}_{n,s}^R C_n^R \tilde{\mathcal{R}}_{n-s,s}^L)^{-1}\end{aligned}$$

Examinons le crochet :

$$\begin{aligned}[\quad] &= \mathcal{R}_{n,s}^R C_n^R (\mathcal{P}_{n-s,s}^L \mathcal{Q}_{n-r,r}^R - \mathcal{Q}_{n-s,s}^L \mathcal{P}_{n-r,r}^R) \tilde{C}_n^L \tilde{\mathcal{R}}_{n,r}^L \\ &\quad - \mathcal{R}_{n,s}^R C_n^R (\mathcal{P}_{n-s,s}^L \mathcal{Q}_{n,s}^R - \mathcal{Q}_{n-s,s}^L \mathcal{P}_{n,s}^R) C_n^R \tilde{\mathcal{R}}_{n-s,s}^L \\ &\quad + \mathcal{R}_{n-r,r}^R C_n^L (-\mathcal{P}_{n,r}^L \mathcal{Q}_{n-r,r}^R + \mathcal{Q}_{n,r}^L \mathcal{P}_{n-r,r}^R) \tilde{C}_n^L \tilde{\mathcal{R}}_{n,r}^L \\ &\quad + \mathcal{R}_{n-r,r}^R C_n^L (\mathcal{P}_{n,r}^L \mathcal{Q}_{n,s}^R - \mathcal{Q}_{n,r}^L \mathcal{P}_{n,s}^R) C_n^R \tilde{\mathcal{R}}_{n-s,s}^L\end{aligned}$$

En utilisant les formules de la propriété 2.1, on obtient

- $(-\mathcal{P}_{n-s,s}^L \mathcal{Q}_{n-r,r}^R + \mathcal{Q}_{n-s,s}^L \mathcal{P}_{n-r,r}^R) = 0_{s \times r}$
- $(\mathcal{P}_{n-s,s}^L \mathcal{Q}_{n,s}^R - \mathcal{Q}_{n-s,s}^L \mathcal{P}_{n,s}^R) = -(\mathcal{C}_n^R)^{-1}$
- $(-\mathcal{P}_{n,r}^L \mathcal{Q}_{n-r,r}^R + \mathcal{Q}_{n,r}^L \mathcal{P}_{n-r,r}^R) = -(\mathcal{C}_n^L)^{-1}$
- $(\mathcal{P}_{n,r}^L \mathcal{Q}_{n,s}^R - \mathcal{Q}_{n,r}^L \mathcal{P}_{n,s}^R) = 0_{r \times s}$.

Le crochet est donc égal à

$$\begin{aligned} [\quad] &= \mathcal{R}_{n,s}^R \mathcal{C}_n^R \tilde{\mathcal{R}}_{n-s,s}^L - \mathcal{R}_{n-r,r}^R \tilde{\mathcal{C}}_n^L \tilde{\mathcal{R}}_{n,r}^L \\ &= \mathcal{R}_{n-r,r}^R \left((\mathcal{R}_{n-r,r}^R)^{-1} \mathcal{R}_{n,s}^R \mathcal{C}_n^R - \tilde{\mathcal{C}}_n^L \tilde{\mathcal{R}}_{n,r}^L (\tilde{\mathcal{R}}_{n-s,s}^L)^{-1} \right) \tilde{\mathcal{R}}_{n-s,s}^L \end{aligned}$$

Le numérateur est donc égal à

$$\begin{aligned} \varphi - S_n^L(\tilde{\varphi}) &= \mathcal{R}_{n-r,r}^R \left((\mathcal{R}_{n-r,r}^R)^{-1} \mathcal{R}_{n,s}^R \mathcal{C}_n^R - \tilde{\mathcal{C}}_n^L \tilde{\mathcal{R}}_{n,r}^L (\tilde{\mathcal{R}}_{n-s,s}^L)^{-1} \right) \tilde{\mathcal{R}}_{n-s,s}^L \\ &\quad \left(-\mathcal{Q}_{n-r,r}^R \tilde{\mathcal{C}}_n^L \tilde{\mathcal{R}}_{n,r}^L + \mathcal{Q}_{n,s}^R \mathcal{C}_n^R \tilde{\mathcal{R}}_{n-s,s}^L \right)^{-1} \end{aligned}$$

car la première parenthèse est égale à l'identité.

D'après la forme de $\tilde{\varphi}$ et les calculs effectués dans le premier chapitre, on a $\tilde{\varphi} = -\tilde{\mathcal{C}}_n^L \tilde{\mathcal{R}}_{n,r}^L (\tilde{\mathcal{R}}_{n-s,s}^L)^{-1}$.

On modifie ensuite l'écriture à l'intérieur de la dernière parenthèse afin de faire apparaître l'identité ajouté à un terme qui tend vers zéro lorsque n tend vers l'infini.

$$\begin{aligned} \varphi - S_n^L(\tilde{\varphi}) &= \mathcal{R}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) \left(-\mathcal{Q}_{n-r,r}^R \tilde{\mathcal{C}}_n^L \tilde{\mathcal{R}}_{n,r}^L (\tilde{\mathcal{R}}_{n-s,s}^L)^{-1} + \mathcal{Q}_{n,s}^R \mathcal{C}_n^R \right)^{-1} \\ &= \mathcal{R}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) \left(\mathcal{Q}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) + \mathcal{Q}_{n-r,r}^R \varphi_n + \mathcal{Q}_{n,s}^R \mathcal{C}_n^R \right)^{-1} \\ &= \mathcal{R}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) \left(\mathcal{Q}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) - \mathcal{Q}_{n-r,r}^R \mathcal{C}_n^L \mathcal{R}_{n,s}^L (\mathcal{R}_{n-s,s}^L)^{-1} + \mathcal{Q}_{n,s}^R \mathcal{C}_n^R \right)^{-1} \end{aligned}$$

D'après les formules liant les numérateurs et les dénominateurs, on obtient

$$\mathcal{Q}_{n-r,r}^R \mathcal{C}_n^L \mathcal{R}_{n,r}^L = \mathcal{Q}_{n,s}^R \mathcal{C}_n^R \mathcal{R}_{n-s,s}^L - I_s$$

On injecte donc le résultat dans l'expression du numérateur ce qui élimine $\mathcal{Q}_{n,s}^R \mathcal{C}_n^R$ de la parenthèse et on factorise $(\mathcal{R}_{n-s,s}^L)^{-1}$. Ce qui donne

$$\varphi - S_n^L(\tilde{\varphi}) = \mathcal{R}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) (\mathcal{R}_{n-s,s}^L \mathcal{Q}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) + I_s)^{-1} \mathcal{R}_{n-s,s}^L$$

Multiplions par x à gauche et y à droite. On obtient

$$x^t \mathcal{R}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) (\mathcal{R}_{n-s,s}^L \mathcal{Q}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) + I_s)^{-1} \mathcal{R}_{n-s,s}^L y$$

On prend la valeur absolue et on utilise l'inégalité de Cauchy-Schwartz pour aboutir à

$$\|x^t (\varphi - S_n^L(\tilde{\varphi})) y\| \leq \|x^t \mathcal{R}_{n-r,r}^R\| \cdot \|(\tilde{\varphi} - \varphi_n) (\mathcal{R}_{n-s,s}^L \mathcal{Q}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) + I_s)^{-1}\| \cdot \|\mathcal{R}_{n-s,s}^L y\| \quad (2.5)$$

On étudie $x^t \mathcal{R}_{n-r,r}^R$ et $\mathcal{R}_{n-s,s}^L y$. D'après le lemme 2.14, on a

$$\begin{aligned} x^t \mathcal{R}_{n-r,r}^R &= \tilde{x}^t \text{diag} \left(\frac{1}{D_{n-s}^{(s+1)}}, \dots, \frac{1}{D_{n-s}^{(s+r)}} \right) W_{2,2}^{(n-s)} (C_n^L)^{-1} \\ \text{et } \mathcal{R}_{n-s,s}^L y &= V_{1,1}^{(n-s)} \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) \tilde{y} \end{aligned}$$

Notons $(c_1, \dots, c_r)^t = \tilde{x}^t$, $c_k = 0$, $k < \alpha$ et $(d_1, \dots, d_s)^t = \tilde{y}^t$, $d_k = 0$, $k > \beta$.

Donc

$$\begin{aligned} \|x^t \mathcal{R}_{n-r,r}^R\| &\leq \|(0, \dots, 0, c_\alpha (D_{n-s}^{(s+\alpha)})^{-1}, \dots, c_r (D_{n-s}^{(s+r)})^{-1})\| \cdot \|W_{2,2}^{(n-s)} (C_n^L)^{-1}\| \\ \|\mathcal{R}_{n-s,s}^L y\| &\leq \|V_{1,1}^{(n-s)}\| \cdot \|(d_1 D_{n-s}^{(1)}, \dots, d_\beta D_{n-s}^{(\beta)}, 0, \dots, 0)\| \end{aligned}$$

Or

$$\lim_{n \rightarrow \infty} V_{1,1}^{(n-s)} = V_{1,1}^{(\infty)} \text{ et } \lim_{n \rightarrow \infty} W_{2,2}^{(n-s)} (C_n^L)^{-1} = W_{2,2}^{(\infty)} \tilde{C}_0^L$$

Donc leurs normes sont bornées. Notons $K_n = \|W_{2,2}^{(n-s)} (C_n^L)^{-1}\| \cdot \|V_{1,1}^{(n-s)}\|$ et $K = \lim_{n \rightarrow \infty} K_n$.

On utilise par exemple la norme 1 et on divise par $|(D_{n-s}^{(s+\alpha)})^{-1}|$ et $|D_{n-s}^{(\beta)}|$. On obtient alors

$$\frac{\|x^t \mathcal{R}_{n-r,r}^R\| \cdot \|\mathcal{R}_{n-s,s}^L y\|}{|(D_{n-s}^{(s+\alpha)})^{-1}| \cdot |D_{n-s}^{(\beta)}|} \leq K_n \cdot \sum_{j=\alpha}^r \left| c_j \left(\frac{D_{n-s}^{(s+j)}}{D_{n-s}^{(s+\alpha)}} \right)^{-1} \right| \cdot \sum_{j=1}^{\beta} \left| d_j \left(\frac{D_{n-s}^{(j)}}{D_{n-s}^{(\beta)}} \right) \right|$$

Grâce au lemme 2.12 sur la base ordonnée par domination,

$$\lim_{n \rightarrow \infty} \frac{\|x^t \mathcal{R}_{n-r,r}^R\| \cdot \|\mathcal{R}_{n-s,s}^L y\|}{|(D_{n-s}^{(s+\alpha)})^{-1}| \cdot |D_{n-s}^{(\beta)}|} \leq K \cdot c_\alpha \cdot d_\beta. \quad (2.6)$$

De plus, toujours grâce au lemme 2.14, on a

$$\begin{aligned} &\mathcal{R}_{n-s,s}^L \mathcal{Q}_{n-r,r}^R \\ = & - \left(V_{1,1}^{(n-s)} W_{1,2}^{(n-s)} - V_{1,1}^{(n-s)} \text{diag}(D_{n-s}^{(1)}, \dots, D_{n-s}^{(s)}) P_{1,1} P_{2,1}^{-1} \text{diag}(D_{n-s}^{(s+1)}, \dots, D_{n-s}^{(s+r)})^{-1} W_{2,2}^{(n-s)} \right) (C_n^L)^{-1} \end{aligned}$$

Donc la limite existe et est bornée. Donc

$$\lim_{n \rightarrow \infty} \|(\tilde{\varphi} - \varphi_n)(\mathcal{R}_{n-s,s}^L \mathcal{Q}_{n-r,r}^R (\tilde{\varphi} - \varphi_n) + I_s)^{-1}\| = 0 \quad (2.7)$$

Finalement, en utilisant les formules (2.5), (2.6) et (2.7) et le lemme 2.15, on obtient bien l'accélération de convergence :

$$\lim_{n \rightarrow \infty} \frac{|x^t (S_n^L(\tilde{\varphi}) - \varphi) y|}{|x^t (S_n^L(0) - \varphi) y|} = 0$$

□

Corollaire 2.18. *Considérons un opérateur bande $A = (A_{k,j})$ à $r + s + 1$ diagonales vérifiant*

$$\lim_{n \rightarrow \infty} A_{n,n+p} := h_p$$

$p = -s, \dots, r$ et $h_{-s}h_r \neq 0$. Supposons que les valeurs propres, dépendant de z , de la matrice A^L définie dans 1.42 vérifient $|\lambda_1| < |\lambda_2| < \dots < |\lambda_{r+s}|$.

Alors $z \in \Omega(A)$ si et seulement si $z \in \Omega(\tilde{A})$ et l'une des trois conditions équivalentes suivantes est satisfaite

- (i) $P_{2,1}$ est inversible.
- (ii) $\limsup_{n \rightarrow \infty} \|S_n^L(0)\| < \infty$
- (iii) $\lim_{n \rightarrow \infty} S_n^L(0) := \varphi$ existe dans $\mathbb{C}^{r \times s}$.

Notons que $z \in \Omega(\tilde{A})$ signifie que $|\lambda_s| < 1 < |\lambda_{s+1}|$ d'après le théorème 1.44. Sous les hypothèses du corollaire, suivant le raisonnement présenté ci-dessous, on montre que les éléments de $\Omega(\tilde{A})$ ne vérifiant aucune des trois conditions est une valeur propre de A .

Démonstration. D'après le lemme 2.13, on a les implications suivantes :

(ii) \Rightarrow (i) et (i) \Rightarrow (iii)

La dernière implication (iii) \Rightarrow (ii) est évidente.

Supposons que $z \in \Omega(\tilde{A})$. Cela signifie que $|\lambda_s| < 1 < |\lambda_{s+1}|$.

Supposons de plus (i). Cela implique que φ existe et que toutes les formules du lemme 2.14 sont valables.

Grâce à la base ordonnée par domination et à la décroissance géométrique du lemme 2.12, en effectuant complètement les produits, on peut démontrer que les polynômes Q_k et les résidus R_n vérifient le critère du théorème 1.40 du chapitre 1 et ainsi que $z \in \Omega(A)$.

Réciproquement, si $z \in \Omega(A)$, φ existe. On réalise alors la même démonstration que dans le chapitre 1 pour le théorème 1.44, page 37. En effet, par un raisonnement identique, on élimine les solutions ne tendant pas vers zéro. On trouve alors que $(D_{n-s}^{(1)}, \dots, D_{n-s}^{(r+s)})$ sont séparés en deux groupes de tailles respectives r et s .

On trouve, en particulier, $P_{2,1}\varphi + P_{2,2} = 0_{r \times s}$. En utilisant le même argument que dans la démonstration du lemme 2.13, on démontre que $P_{2,1}$ est inversible, c'est-à-dire la condition (i), et ainsi que les formules du lemme 2.14 sont valables. Comme le critère du théorème 1.40 du chapitre 1 est vérifié, on passe à la limite sur les formules. En utilisant alors (2.2), on démontre que $|\lambda_s| < 1 < |\lambda_{s+1}|$ et ainsi que $z \in \Omega(\tilde{A})$. \square

Moments modifiés

3.1 Introduction

Les polynômes orthogonaux apparaissent dans l'étude des fractions continues, les approximants de Padé vectoriels ou matriciels ; ils sont aussi un outil important dans la résolution du maillage de Toda-Langmuir et les systèmes dynamiques associés.

Au lieu de déterminer, de manière explicite, les coefficients des polynômes (bi)orthogonaux définis par une certaine matrice de mesures, on calcule habituellement les coefficients de la relation de récurrence vérifiée par ces polynômes. Ceci permet, en particulier, de diminuer la taille occupée en mémoire. L'évaluation du polynôme en une valeur x_0 peut aussi se réaliser de manière stable grâce à l'algorithme de Clenshaw [63] qui se rapproche de l'algorithme d'Horner. Dans le cas scalaire, une méthode relativement adaptée à la détermination des coefficients de récurrence est l'algorithme de Chebyshev modifié [68], qui a été récemment étendu au cas vectoriel [28] ainsi que dans [18]. La difficulté principale qui réside dans l'application de cet algorithme est un choix convenable des polynômes de référence nécessaires pour les moments modifiés sous-jacents. Nous avons démontré quelques résultats sur le cas scalaire dans [9].

Dans la suite, on présente une version de l'algorithme de Chebyshev modifié pour des polynômes vectoriels, après avoir rappelé l'algorithme scalaire. Ensuite on applique cet algorithme à quelques exemples. Les résultats numériques obtenus sont analysés ainsi que les choix possibles de polynômes de référence.

Cet algorithme sera utilisé de nouveau dans le chapitre suivant car il permet de déterminer les coefficients d'une matrice de Jacobi $\mathcal{J}_N(t)$ qui évolue en fonction du temps, à partir des moments modifiés solutions d'une équation différentielle $\dot{\Phi}_N(t) = -\mathcal{J}_N(0)\Phi_N(t)$.

3.2 Calcul des coefficients. Algorithme de Chebyshev modifié

3.2.1 Introduction

Généralement, pour calculer les coefficients des relations de récurrence, on utilise l'algorithme de Chebyshev modifié au lieu de l'algorithme de Chebyshev. Cet algorithme utilise ce que l'on appelle les *moments modifiés* $\int p_k(x)w(x)dx$ au lieu des moments ordinaires $\int x^k w(x)dx$. Dans le cas scalaire, il a été prouvé dans [9] que le choix des polynômes de référence $(\Pi_k)_k$ est important. En particulier, si ce sont des polynômes orthogonaux, les supports des mesures sous-jacentes doivent être essentiellement les mêmes.

L'algorithme dans le cas scalaire peut être par exemple trouvé dans [32], [68]. Une étude du conditionnement des applications non linéaires qui associent les moments modifiés aux coefficients de récurrence peut être trouvée dans [32], [36]. Fischer [30] puis Beckermann et Bourreau [9] ont amélioré cette étude. Nous rappelons donc tout d'abord cet algorithme et nous le généralisons au cas vectoriel. Nous parlerons en particulier du choix des polynômes de référence. Pour une version par bloc de l'algorithme, on pourra se référer à [69]. Un algorithme dans le cas vectoriel a été décrit dans [28] sans toutefois d'explication sur le choix des polynômes de référence.

3.2.2 Algorithme de Chebyshev modifié scalaire

Soit $d\sigma(t)$ une mesure non négative sur la droite réelle de support fini ou infini.

On suppose que tous les moments

$$\int_{\mathbb{R}} t^k d\sigma(t)$$

existent et sont finis. On peut définir, alors, à partir de ces moments, un système de polynômes π_k orthogonaux relativement à $d\sigma(t)$. Ces polynômes sont uniques si on suppose en plus qu'ils sont unitaires. Alors

$$\begin{aligned} &\pi_k \text{ est de degré } k, \text{ unitaire} \\ &\int_{\mathbb{R}} \pi_m(t)\pi_n(t)d\sigma(t) \begin{cases} = 0 & \text{si } m > n \\ > 0 & \text{si } m = n \end{cases} \end{aligned}$$

Propriété 3.1. *Ces polynômes vérifient une relation de récurrence à trois termes comme suit :*

$$\begin{aligned} \pi_{k+1}(t) &= (t - \alpha_k)\pi_k(t) - \beta_k\pi_{k-1}(t) \\ \pi_{-1}(t) &= 0 \quad \pi_0(t) = 1. \end{aligned} \tag{3.1}$$

α_k et β_k sont des nombres réels et $\beta_k > 0$.

Les coefficients de récurrence se calculent généralement à partir des moments ordinaires $\int x^k d\sigma(x)$ au travers de l'algorithme de Chebyshev [34, 2.4, p. 295]. Or la matrice de Hankel formée des moments ordinaires est très mal conditionnée puisque le conditionnement augmente exponentiellement en fonction de n [6, corollaire 5.15]. L'algorithme est donc numériquement instable. On doit changer de base de polynômes pour obtenir des moments modifiés et essayer de stabiliser l'algorithme.

Soit donc une famille de polynômes orthogonaux $(p_k)_k$, relativement à une mesure s , définis par la relation de récurrence suivante

$$p_{k+1}(t) = (t - a_k)p_k(t) - b_k p_{k-1}(t) ,$$

les coefficients $(a_k)_k$ et $(b_k)_k$ étant supposés connus.

Définition 3.2. *Soient les moments modifiés*

$$m_k = \int_{\mathbb{R}} p_k(t)d\sigma(t) \quad k = 0, 1, \dots$$

Ces moments permettent de calculer les coefficients $(\alpha_k)_k$ et $(\beta_k)_k$ au travers de l'algorithme de Chebyshev modifié.

Lemme 3.3. *Les coefficients de récurrence vérifient les égalités suivantes :*

$$\begin{aligned} \alpha_j &= \frac{\int_{\mathbb{R}} t\pi_j^2(t)d\sigma(t)}{\int_{\mathbb{R}} \pi_j^2(t)d\sigma(t)} \\ \beta_j &= \frac{\int_{\mathbb{R}} \pi_j^2(t)d\sigma(t)}{\int_{\mathbb{R}} \pi_{j-1}^2(t)d\sigma(t)} \end{aligned}$$

Démonstration. Ces formules s'obtiennent aisément en utilisant la formule de récurrence (3.1) ainsi que l'orthogonalité par rapport à $d\sigma$.

Notons $(\cdot, \cdot)_\sigma$ le produit scalaire défini par

$$(f, g)_\sigma = \int_{\mathbb{R}} f(t)g(t)d\sigma(t)$$

Comme π_{j+1} est orthogonal à π_j et π_{j-1} , on obtient les égalités suivantes :

$$\begin{aligned} 0 &= (\pi_{j+1}, \pi_j)_\sigma \\ &= (t\pi_j, \pi_j)_\sigma - \alpha_j(\pi_j, \pi_j)_\sigma - \beta_j(\pi_{j-1}, \pi_j)_\sigma \end{aligned}$$

Ce qui donne la première égalité.

$$\begin{aligned} 0 &= (\pi_{j+1}, \pi_{j-1})_\sigma \\ &= (\pi_j, t\pi_{j-1})_\sigma - \beta_j(\pi_{j-1}, \pi_{j-1})_\sigma \\ &= (\pi_j, \pi_j)_\sigma - \beta_j(\pi_{j-1}, \pi_{j-1})_\sigma \end{aligned}$$

Ce qui conduit à la deuxième égalité. □

Définition 3.4. Soient les moments mixtes

$$\sigma_{j,k} = \int_{\mathbb{R}} \pi_j(t)p_k(t)d\sigma(t), \quad j \leq k$$

qui seront utilisés par l'algorithme des moments modifiés.

Remarque 3.5. On peut noter que $\sigma_{j,k} = 0$ si $j > k$ grâce à l'orthogonalité relativement à $d\sigma$.

On se propose de calculer les coefficients de récurrence (α_k) et (β_k) , $k = 0, \dots, n-1$.

On a donc besoin de $2n$ moments modifiés.

La propriété suivante présente l'algorithme de Chebyshev modifié qui permet le calcul des coefficients de récurrence.

Propriété 3.6. Les relations suivantes servent à initialiser l'algorithme :

$$\begin{aligned} \sigma_{-1,k} &= 0, & k &= 1, 2, \dots, 2n-2 \\ \sigma_{0,k} &= m_k, & k &= 0, 1, \dots, 2n-1 \\ \alpha_0 &= a_0 + \frac{m_1}{m_0} \\ \beta_0 &= m_0 & (\text{par convention}) \end{aligned}$$

Theorème 3.8. 1. Si $\text{supp}(s) \not\subset \text{supp}(\sigma)$ alors le conditionnement de K_n , augmente au minimum avec un taux géométrique

$$\liminf_{n \rightarrow +\infty} \text{cond}_{D_{\text{nor}}}(K_n)^{1/n} \geq \max_{x \in \text{supp}(s)} e^{2g_{\text{supp}(\sigma)}(x)} > 1$$

où g est la fonction de Green relativement au support.

2. Si la mesure s appartient à un ensemble de mesures dénommé **Reg** introduit dans [74], si son support est un intervalle et si $\text{supp}(\sigma) \not\subset \text{supp}(s)$ alors le conditionnement de K_n augmente au minimum avec un taux géométrique

$$\liminf_{n \rightarrow +\infty} \text{cond}_{D_{\text{nor}}}(K_n)^{1/n} \geq \max_{x \in \text{supp}(\sigma)} e^{2g_{\text{supp}(s)}(x)} > 1$$

3. Si $S = \text{supp}(s) = \text{supp}(\sigma)$ alors, avec $M := \max_{x \in S} |x|$,

$$\frac{\sqrt{(\sigma_0)}/2}{\sqrt{\sigma_0} + \sqrt{n} \cdot M} \leq \text{cond}_{D_{\text{nor}}}(K_n) \leq \sqrt{8}M \left[\int K_n^\sigma(t) K_{n-1}^\sigma(t) ds(t) \int K_{2n-1}^s(x) d\sigma(x) \right]^{1/2},$$

K_n^\bullet représentant le noyau de Szëgo de la mesure indiquée. Si, de plus, s et σ appartiennent à **Reg** et si S est, par exemple, une union finie d'intervalles alors $\lim_{n \rightarrow +\infty} \text{cond}_{D_{\text{nor}}}(K_n)^{1/n} = 1$.

L'ensemble **Reg** contient en particulier les mesures de la forme $d\sigma(x) = w(x)dx$, de support un intervalle où w est positive.

3.2.3 Algorithme de Chebyshev modifié vectoriel

Dans cette partie, on définit, sur le même modèle que l'algorithme scalaire, un algorithme permettant de déterminer les coefficients de récurrence de polynômes vectoriels orthogonaux relativement à une matrice de mesures.

Comme dans le cas scalaire, la stabilité numérique semble dépendre, de manière importante, du choix des polynômes dits de référence qui servent au calcul des moments modifiés. Cette conjecture provient simplement de la discussion des exemples qui sont proposés dans la suite de cette partie. Dans le cas scalaire, nous avons démontré que les supports des mesures d'orthogonalité des deux familles de polynômes devaient coïncider. Dans le cas vectoriel, malheureusement, cette supposition n'est, pour l'instant, pas prouvée. Il semble d'ailleurs improbable d'employer la même voie de recherche que celle utilisée par Fischer dans le cas scalaire.

On considère une famille de polynômes orthogonaux $(\Pi_k^R)_{k \in \mathbb{N}}$, avec $d \deg \Pi_k^R = k$. On suppose que les coefficients $N_{j,l}$ des relations de récurrence

$$x \Pi_l^R = \sum_{j=l-r'}^{l+s'} N_{j,l} \Pi_j^R$$

sont connus.

A cause des contraintes portant sur les degrés, on peut montrer que s' doit être choisi égal à s .

En utilisant le produit scalaire défini par (1.2) relativement à une matrice W de fonctions poids, on définit aussi les moments mixtes $\sigma_{k,l} = \langle Q_k^L, \Pi_l^R \rangle_W$. D'après les propriétés d'orthogonalité, on sait que $\sigma_{k,l} = 0$ si k est plus grand que l . De plus, Π_l^R peut être exprimé grâce aux (Q_k^R) par $\Pi_l^R = \sum_{k=0}^l Q_k^R \sigma_{k,l}$. En fait, les $(\sigma_{k,l})$ sont les coefficients de transmission qui permettent d'exprimer les Π_l^R dans la base des $(Q_k^R)_k$. On obtient l'Algorithme 1 pour calculer les n premiers coefficients de récurrence.

Algorithme 1 Algorithme de Chebyshev modifié

Initialisation

$$T := \left\lfloor \frac{n}{r} \right\rfloor s + n + r + s$$

$$\forall k \in \{-s, \dots, -1\}, \forall l \in \mathbb{N}, l < T \quad \sigma_{k,l} = 0$$

$\forall k \in \{0, \dots, r-1\}, \forall k \leq l \leq T$, calcul des $\sigma_{k,l}$ en utilisant une intégration exacte ou une valeur approchée par quadrature.

Calcul des premiers coefficients de récurrence.

Récurrence

for $k = 1$ to n **do**

for $l = k + r - 1$ to $T - \left\lfloor \frac{k}{r} \right\rfloor s$ **do**

$$A_{k-1, k+r-1} \sigma_{k+r-1, l} = - \sum_{j=k-1-s}^{k+r-2} A_{k-1, j} \sigma_{j, l} + \sum_{j=l-r'}^{l+s} N_{j, l} \sigma_{k-1, j}$$

end for

for $l = -s$ to $r - 1$ **do**

$$A_{k, k+l} = \sigma_{k+l, k+l}^{-1} \left(\sum_{j=k+l-r'}^{k+l+s} N_{j, k+l} \sigma_{k, j} - \sum_{p=k-s}^{k+l-1} A_{k, p} \sigma_{p, k+l} \right)$$

end for

end for

Démonstration. D'après le produit scalaire défini précédemment, on obtient :

$$\langle x Q_k^L, \Pi_l^R \rangle_W = \langle Q_k^L, x \Pi_l^R \rangle_W.$$

En utilisant les deux relations de récurrence, on obtient

$$\sum_{j=k-s}^{k+r} A_{k, j} \sigma_{j, l} = \sum_{j=l-r'}^{l+s} N_{j, l} \sigma_{k, j} \quad (3.5)$$

En faisant varier l de $k-s$ à $k+r-1$ dans l'égalité précédente, on obtient un système triangulaire qui permet de calculer les $(A_{k, k-s}, \dots, A_{k, k+r-1})$. Le système est triangulaire car les $\sigma_{k, l}$ sont nuls si k est plus grand que l . On a ainsi explicité la deuxième boucle de la récurrence.

Si on réécrit (3.5), en conservant uniquement $A_{k,k+r}\sigma_{k+r,l}$ dans le membre de gauche, on obtient les relations données dans la première boucle. Et ainsi, on peut déterminer la nouvelle ligne de moments mixtes. \square

Cette démonstration prouve ainsi les relations données dans le cas scalaire avec $r = s = 1$.

Remarque 3.9. Dans le cas scalaire, [32], l'algorithme de Chebyshev modifié utilise comme données initiales le vecteur des moments modifiés

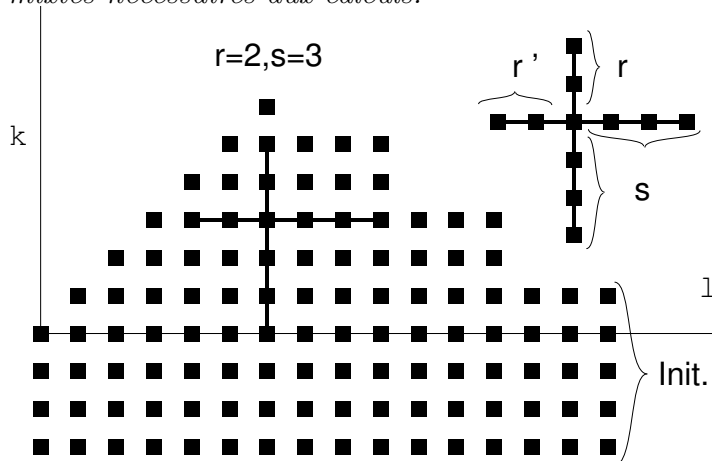
$$m = (m_0, \dots, m_{2n-1})^t, \quad m_k = \int \pi_k(x)w(x)dx$$

En fait, ce sont des moments mixtes $\sigma_{0,k}$ avec $q_0(x) = 1$.

Si $r > 1$, on doit être conscient du problème de l'initialisation. On ne peut pas choisir $(Q_0^L, \dots, Q_{r-1}^L)^t = I_r$ pour calculer les r lignes de données initiales. Mais, en première intention, on suppose que ce sont les valeurs correctes de $(Q_0^L, \dots, Q_{r-1}^L)^t$. On calcule alors seulement $W_0 := (\sigma_{k,l})_{(k,l) \in \{0, \dots, r-1\}}$. Ensuite, comme indiqué dans le chapitre 1, dans l'égalité (1.3), grâce à une décomposition LU de W_0 , on récupère les valeurs correctes de $(Q_0^L, \dots, Q_{r-1}^L)^t$ et on est alors capable de poursuivre les calculs. Si W est symétrique, définie positive, on peut aussi utiliser une décomposition de Cholesky pour récupérer $Q_n^R = (Q_n^L)^T$.

Remarque 3.10. - On possède un degré de liberté dans la normalisation de $A_{k,k+r}$. Par exemple, on peut poser $A_{k,k+r} = c$ ou le définir de telle manière que $\sigma_{k+r,k+r} = 1$. Dans les exemples suivants, $c = 1$ or $c = \frac{1}{2}$. Cette normalisation existe aussi dans le cas scalaire.

- La relation (3.5) est une "règle de la croix". On présente, ci-après un exemple où $r = 2, s = 3$. On veut calculer les relations de récurrence jusqu'à la cinquième récurrence et nous avons dessiné tous les moments mixtes nécessaires aux calculs.



La partie verticale de la croix désigne le membre de gauche de (3.5) et le membre de droite est décrit par la partie horizontale.

- On a besoin de $T := \left\lceil \frac{n}{r} \right\rceil s + n + r + s$ moments initiaux dans les $r + s$ premières lignes afin de calculer les coefficients de la n -ème récurrence. Lors de chaque boucle de la récurrence, on calcule $T - \left\lceil \frac{k}{r} \right\rceil s - k$ moments mixte. Chaque moment mixte nécessite $2r + 2s + 1$ multiplications, $2r + 2s - 1$ additions et une division. Nous avons aussi $r + s$ coefficients. Ils nécessitent $(r + s + 1)(r + s)/2$ multiplications et additions et $r + s$ divisions.

3.3 Résultats numériques

Dans cette section, nous décrivons certaines expériences numériques obtenues par l'application de l'algorithme 3.2.3 à des exemples classiques. Les calculs correspondant ont été réalisés avec Maple V Release 5 sur un Power Macintosh G3/315Mhz. Le code Maple de l'exemple 3.11 est donné, *in extenso*, à l'annexe A.1.

Dans nos deux premiers exemples, on considère des polynômes unitaires Q_n^L d'orthogonalité simultanée relativement à deux poids, c'est-à-dire que $r = 1$ et $s = 2$. Comme indiqué précédemment, ces polynômes vérifient une relation de récurrence à quatre termes

$$Q_{n+1}^L(z) = (z - A_{n,n})Q_n^L(z) - A_{n,n-1}Q_{n-1}^L(z) - A_{n,n-2}Q_{n-2}^L(z) \quad (3.6)$$

où nous voulons déterminer les inconnues $A_{n,n}$, $A_{n,n-1}$ et $A_{n,n-2}$.

Exemple 3.11. Dans [49, section 3&5], les auteurs considèrent le vecteur de poids ($r = 1, s = 2$)

$$W(x) = (h(x)\chi_{[a,0]}(x), h(x)\chi_{[0,1]}(x))$$

où $h(x) = (x - a)^\alpha(1 - x)^\beta|x|^\gamma$ ($\alpha, \beta, \gamma > -1$) et χ_E est la fonction caractéristique de l'intervalle E . Ici, on étudie le cas spécial $a = -1$, $\alpha = -0.5$, $\beta = -0.5$, $\gamma = 0.5$.

Dans cet exemple, on sait que les coefficients des relations de récurrence possèdent des limites périodiques de période 2 :

$$\begin{aligned} -A_{2k,2k} &\rightarrow -2\frac{\sqrt{3}}{9}, & -A_{2k+1,2k+1} &\rightarrow 2\frac{\sqrt{3}}{9}, \\ -A_{2k,2k-1} &\rightarrow -\frac{4}{27}, & -A_{2k+1,2k} &\rightarrow -\frac{4}{27}, \end{aligned}$$

$$-A_{2k,2k-2} \rightarrow -\frac{8\sqrt{3}}{243}, \quad -A_{2k+1,2k-1} \rightarrow \frac{8\sqrt{3}}{243},$$

voir [49, Theorem 2, p.212].

Notre choix pour les polynômes de référence $(\Pi_k^R)_k$ est guidé par les trois critères suivants : tout d'abord, les moments modifiés devraient être disponibles explicitement ou facilement calculables. Ensuite, les polynômes de référence doivent satisfaire une relation de récurrence dont les coefficients soient connus et ainsi ils sont eux-mêmes des polynômes orthogonaux vectoriels. Enfin, d'après le cas scalaire, nous savons déjà [9, Thm 11] que les supports des deux mesures d'orthogonalité doivent coïncider.

Prenant en compte ces critères, nous avons choisi comme polynômes de référence $(\Pi_k^R)_k$ des polynômes de Chebyshev vectoriels convenablement translatés

$$\left(\begin{array}{c} T_0(2x+1) \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ T_0(2x-1) \end{array} \right), \left(\begin{array}{c} T_1(2x+1) \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ T_1(2x-1) \end{array} \right), \dots$$

Les mesures d'orthogonalité sont respectivement $(1 - (2x + 1)^2)^{-1/2} dx$ sur $[-1; 0]$ et $(1 - (2x - 1)^2)^{-1/2} dx$ sur $[0; 1]$. Evidemment, on possède les récurrences ($s = s' = 2$, $r' = 2$) :

$$x\Pi_{2n}^R(x) = \frac{1}{4}\Pi_{2n+2}^R(x) - \frac{1}{2}\Pi_{2n}^R(x) + \frac{1}{4}\Pi_{2n-2}^R(x)$$

$$x\Pi_{2n+1}^R(x) = \frac{1}{4}\Pi_{2n+3}^R(x) + \frac{1}{2}\Pi_{2n+1}^R(x) + \frac{1}{4}\Pi_{2n-1}^R(x)$$

Les résultats numériques sont présentés sur la figure 1. Les calculs des moments initiaux de même que ceux des nouveaux coefficients de récurrence ont été effectués en précision finie avec 20 chiffres décimaux. Pour calculer les moments modifiés initiaux, la fonction a été développée en série de Chebyshev. Ceci a donné les valeurs d'une fonction hypergéométrique. Ces quantités ont été évaluées avec une précision suffisamment grande au moyen de l' ϵ -algorithme appliqué aux sommes partielles.

Les suites des 200 premiers coefficients de récurrence semblent approcher les limites prévues par la théorie.

Pour cet exemple, il existe une formule de Rodrigues explicite pour les polynômes Q_n [49, p.210],

$$Q_{k,k}(z) = M_k^{-1}(z^2 - 1)^{1/2} z^{-1/2} \frac{d^k}{dz^k} ((z^2 - 1)^{k-1/2} z^{k+1/2})$$

$$Q_{k,k-1}(z) = N_k^{-1}(z^2 - 1)^{1/2} z^{-1/2} \frac{d^{k-1}}{dz^{k-1}} ((z^2 - 1)^{k-3/2} z^{k-1/2} (z - t_k)).$$

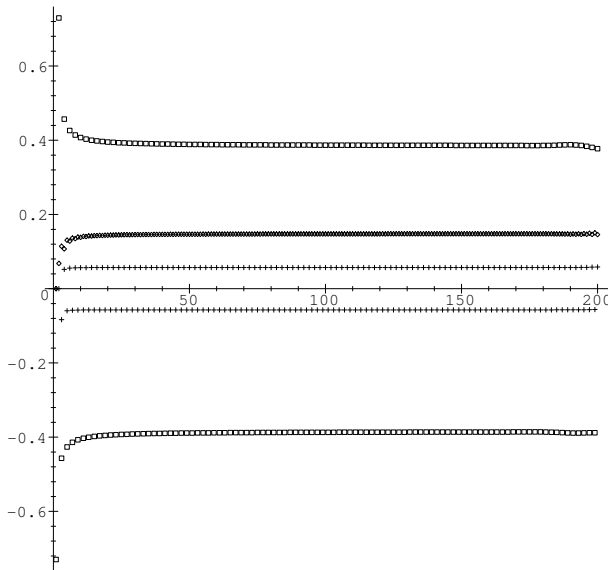


FIG. 3.1 – Exemple 3.11 : 200 coefficients, 20 digits, valeur de $-A_{k,k}$ (carré), $-A_{k,k-1}$ (cercle) et $-A_{k,k-2}$ (croix).

où

$$M_k = (3k + 2)(3k - 1 + 2) \dots (2k + 1 + s),$$

$$N_k = (3k - 2 + 2)(3k - 3 + 2) \dots (2k + 2)$$

et

$$t_k = \frac{-B(k/2 + 3/4, k + 1/2)}{B(k/2 + 1/4, k + 1/2)},$$

$B(p, q)$ étant la fonction Beta. Ces deux formules peuvent être utilisées pour calculer les valeurs exactes des coefficients de récurrence en arithmétique rationnelle. Nous remarquons alors que l'erreur entre les valeurs calculées par l'algorithme et les coefficients exacts est plus petite que 10^{-10} pour $n \leq 100$, et que 10^{-5} pour $n \leq 200$.

Exemple 3.12. *Pineiro [62] (voir aussi Nikishin et Sorokin [58] et Van Assche[79]) ont étudié ce que l'on appelle les polynômes de Jacobi simultanés avec les mesures*

$$d\mu_j(x) = x^{\alpha_j}(1-x)^{\alpha_0} dx, \quad x \in (0, 1),$$

où $\alpha_i > -1$ pour $i = 0, \dots, s$ et $\alpha_i - \alpha_j \notin \mathbb{Z}$ pour $1 \leq i \neq j \leq s$. On considère le cas particulier $s=2$, $\alpha_0 = 0.5$, $\alpha_1 = 0$ et $\alpha_2 = 0.5$. On sait, voir par exemple [79], que, alors, les coefficients de récurrence correspondant de

(3.6) satisfait les limites suivantes :

$$\begin{aligned}\lim_{n \rightarrow +\infty} A_{n,n} &= 3 \left(\frac{4}{27} \right) \\ \lim_{n \rightarrow +\infty} A_{n,n-1} &= 3 \left(\frac{4}{27} \right)^2 \\ \lim_{n \rightarrow +\infty} A_{n,n-2} &= \left(\frac{4}{27} \right)^3\end{aligned}$$

Dans nos expériences numériques, nous avons choisi comme polynômes de référence Π_n^R la suite orthogonale sur $[0, 1]$ suivant la mesure $(1 - (2x - 1)^2)^{-1/2} dx$

$$\begin{pmatrix} T_0(2x-1) \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ T_0(2x-1) \end{pmatrix}, \begin{pmatrix} T_1(2x-1) \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ T_1(2x-1) \end{pmatrix}, \dots$$

En utilisant des calculs élémentaires, on peut déterminer, pour cet exemple, les valeurs exactes des moments modifiés.

$$\int_0^1 T_j(2x-1)(1-x)^{1/2} dx = \frac{1}{4} \int_0^{\pi/2} (\cos(2j+3)\theta + \cos(2j-3)\theta - \cos(2j+1)\theta - \cos(2j-1)\theta) d\theta$$

$$\int_0^1 T_j(2x-1)x^{1/2}(1-x)^{1/2} dx = \frac{1}{4} \int_0^{\pi/2} (\cos(2j\theta) - \frac{1}{2} \cos(2j+4)\theta - \frac{1}{2} \cos(2j-4)\theta) d\theta$$

pour j suffisamment grand. Ce qui prouve que les moments modifiés sont nuls à partir d'un certain rang. Les premiers sont calculables de la même manière, en posant $x = \cos^2(\theta)$. En utilisant alors l'arithmétique rationnelle dans Maple, nous avons, à notre disposition les coefficients de récurrence exacts. Afin de mesurer la sensibilité de l'application non linéaire K_n par rapport aux perturbations des données initiales, c'est-à-dire des moments modifiés, nous avons arrondi les moments modifiés avec une précision de 20 chiffres décimaux. Les calculs suivants ont été réalisés avec une précision de 40 chiffres décimaux suffisante pour négliger les erreurs supplémentaires dues aux calculs en virgule flottante. Il est clair que les résultats du second calcul affichés sur la droite de la figure 2 n'ont aucune signification pour $n \geq 30$. On perd un chiffre significatif à chaque nouvelle étape comme pour les approximants de Padé.

On remarque qu'il est très important de connaître les moments avec la meilleure précision possible. En effet, pour obtenir la même convergence que sur la partie gauche de la figure, on doit calculer les quantités avec une précision décimale de 150 chiffres(!).

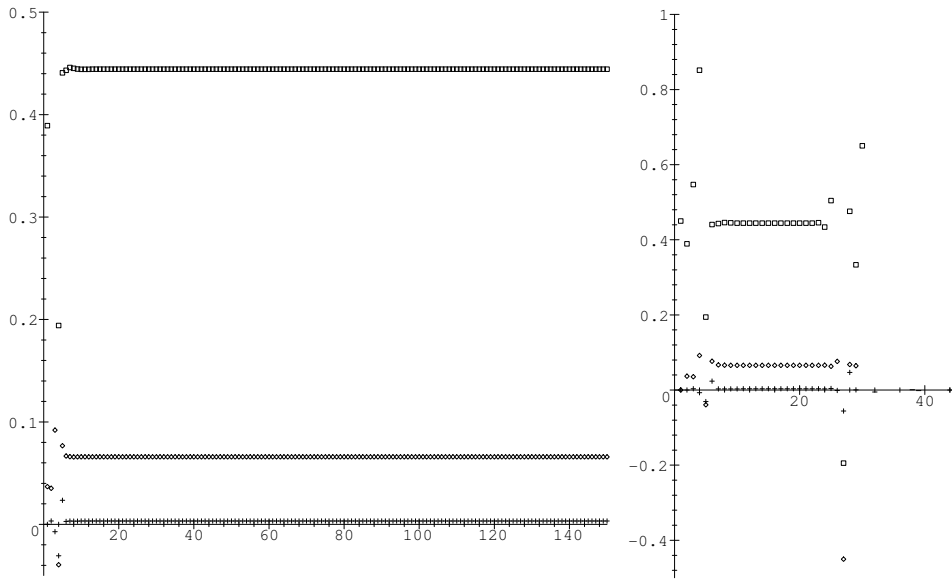


FIG. 3.2 – Exemple 3.12 : Valeur de $A_{k,k}$ (carré), $A_{k,k-1}$ (cercle) and $A_{k,k-2}$ (croix). A gauche, nous avons affiché 150 coefficients obtenus grâce à un calcul en arithmétique exacte. Sur la droite, on trouve les mêmes coefficients après avoir arrondi les données initiales.

Exemple 3.13. Dans ce dernier exemple, nous choisissons $r=s=2$, et le poids est

$$W_C(x) = C \begin{pmatrix} (1-x^2)^{\lambda_1-1/2} & 0 \\ 0 & (1-x^2)^{\lambda_2-1/2} \end{pmatrix} C^T, \quad x \in (-1, 1)$$

où $C \in \mathbb{C}^{2 \times 2}$ est inversible.

On remarque que, dans cet exemple, les polynômes Q_{2n}^L et Q_{2n+1}^L peuvent être regroupés pour construire le n -ième polynôme orthogonal matriciel. Comme vu précédemment, la matrice infinie correspondante $A = (A_{i,j})$ des coefficients de récurrence est pentadiagonale. Si C est la matrice Identité I , seules les entrées sur les diagonales extrêmes sont non nulles. On peut obtenir ces quantités grâce aux coefficients de récurrence de polynômes ultrasphériques adaptés, voir, par exemple, Szegő [75, (4.7.17) p.81]. Ainsi, pour d'autres valeurs de C , on peut explicitement calculer les coefficients de récurrence, de même que les polynômes Q_n^L correspondants; en particulier (dans le cas où le coefficient de récurrence $A_{n,n+2} = 1/2$)

$$\lim_{n \rightarrow +\infty} A_{n,n+1} = \lim_{n \rightarrow +\infty} A_{n,n-1} = 0, \quad \lim_{n \rightarrow +\infty} A_{n,n-2} = \frac{1}{2},$$

et $A_{n,n} = 0$ pour tout $n \geq 0$. Autrement dit, l'opérateur correspondant à A est une perturbation compacte de celui obtenu en choisissant $\lambda_1 = \lambda_2 = 0$ et $C = I$. Les polynômes orthogonaux vectoriels à droite relativement à ce dernier poids sont donnés par

$$\begin{pmatrix} T_0(x) \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ T_0(x) \end{pmatrix}, \begin{pmatrix} T_1(x) \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ T_1(x) \end{pmatrix}, \dots$$

que nous choisirons comme la suite des polynômes de référence Π_k^R . Ils satisfont la relation de récurrence

$$x\Pi_n^R(x) = \frac{1}{2}\Pi_{n+1}^R(x) + \frac{1}{2}\Pi_{n-1}^R(x).$$

Dans nos expériences numériques, nous avons choisi

$$\lambda_1 = 1/2, \lambda_2 = 2, C = \begin{pmatrix} 1 & \gamma \\ \gamma & 1 \end{pmatrix}$$

Comme dans l'exemple précédent, on peut obtenir, explicitement, les moments modifiés. On remarque que $r > 1$, et ainsi nous avons besoin d'appliquer la remarque 3.9 sur l'initialisation.

Les moments modifiés de même que les coefficients de récurrence ont été calculés avec une précision de 10 chiffres décimaux. La norme du vecteur erreur E_n (égal à la différence entre les coefficients de récurrence exacts et calculés) est indiqué dans le tableau suivant pour différentes valeurs de n et γ

n	$\gamma = 1.01$	$\gamma = 1.5$	$\gamma = 2$
10	1.10^{-5}	6.10^{-9}	2.10^{-9}
20	1.10^{-4}	3.10^{-8}	1.10^{-8}
50	2.10^{-4}	2.10^{-7}	2.10^{-8}
100	2.10^{-3}	6.10^{-7}	6.10^{-8}
200	2.10^{-3}	2.10^{-6}	2.10^{-7}
400	1.10^{-3}	1.10^{-5}	1.10^{-6}

On remarque que l'erreur reste petite même pour n proche de 400 et $\gamma = 1.5$ et $\gamma = 2$. Le résultat relativement moins convaincant pour $\gamma = 1.01$ peut être expliqué par le fait que W est proche d'un poids qui n'est plus défini positif sur $[-1, 1]$.

L'algorithme donne de très bons résultats pour l'exemple 3.11 et l'exemple 3.13.

Les faibles résultats de l'exemple 3.12 résultent probablement de comportements asymptotiques différents des polynômes orthogonaux Q_n^R et Π_n^R . On

aurait peut-être dû choisir d'autres polynômes de référence, par exemple les polynômes qui vérifient la récurrence limite mais se pose alors le problème de la détermination des moments modifiés initiaux.

Dans l'exemple 3.13, les propriétés asymptotiques des polynômes orthogonaux et des polynômes de référence sont les mêmes. Cela semble être un aspect théorique important afin de justifier la stabilité numérique assez remarquable de l'algorithme pour cette classe d'exemples.

Il serait intéressant de confirmer nos expériences par une étude théorique du conditionnement numérique de l'application non-linéaire K_n qui associe au vecteur des moments modifiés le vecteur des coefficients de récurrence. Cette étude non triviale dans le cas scalaire (voir [9, 32, 30]) semble encore moins aisée dans le cas vectoriel.

Un choix adapté des polynômes de référence dans le cas vectoriel pour un poids donné reste donc un sujet de recherches à venir.

Systemes dynamiques infinis

On se propose dans cette partie d'étudier l'évolution temporelle d'un système dynamique issu d'un maillage de Toda-Langmuir. Ce maillage ainsi que sa version modifiée dite système de Kac-van Moerbeke peuvent être reliés aux équations de Korteweg-de Vries. Les solutions d'un système de Toda permettent donc d'obtenir les solutions d'autres systèmes [41]. Le maillage de Toda-Langmuir étudié par Moser [57], par Kac et Van Moerbeke [46] ou par Gekhtman [37, 38] est un cas particulier de systèmes plus généraux traités par Berezanskii[14] qui utilise la méthode du problème spectral inverse pour intégrer les équations aux différences, par Bogoyavlenskii [15, 16, 17] ou encore Osipov[59].

Dans le système de Toda-Langmuir, on considère un nombre infini, $k \in \mathbb{Z}$, de particules de même masse situées sur l'axe réel. Physiquement, il s'agit de la description d'un cristal (voir G. Teschl [76]). La particule k se trouve à l'abscisse $x_k = x_k(t)$ qui dépend du temps. On suppose $x_{n-1} < x_n$. Ceci signifie que lors de l'évolution au cours du temps, les particules restent dans l'ordre initial. Ces particules exercent sur leurs voisines immédiates une force exponentielle décroissante en fonction de la distance. La particule x_k est donc soumise à la force $F_k = e^{-(x_k - x_{k-1})} - e^{-(x_{k+1} - x_k)}$. Pour simplifier, on supposera que les particules sont de masse $m = 1$. Les lois de la dynamique permettent

d'écrire les équations suivantes

$$\ddot{x}_k = e^{-(x_k - x_{k-1})} - e^{-(x_{k+1} - x_k)}, \quad k \in \mathbb{Z} \quad (4.1)$$

On suppose que l'on connaît les données initiales, c'est-à-dire la position $x_k(0) = \kappa_k$ et la vitesse $\dot{x}_k = \eta_k$. Ces quantités sont supposées uniformément bornées.

Nous restreindrons notre étude au système de Toda dit semi-infini, c'est-à-dire que l'on suppose que les particules -1 et 0 sont tellement éloignées que leurs interactions deviennent négligeables. Mathématiquement, cela signifie que les équations différentielles (4.1) restent valables pour $k \geq 1$ et on obtient

$$\ddot{x}_0 = -e^{-(x_1 - x_0)}$$

Les particules d'indice négatif, $k \leq -1$, pourraient être traitées de la même manière.

Comme nous le démontrerons ci-après, une solution (globale) de ce système existe et est unique. Elle peut être écrite grâce au théorème spectral mais cette résolution n'est pas constructive et ne se prête pas à un calcul numérique. On propose ici une autre approche, pour laquelle on donnera une analyse d'erreur ainsi que des essais numériques. Cette nouvelle approche est axée sur la résolution effective de systèmes de dimension finie comportant N particules : si on suppose, de plus, que $x_N - x_{N-1} \gg x_{N-1} - x_{N-2}$, alors on obtient un système de N équations différentielles ordinaires non linéaires, c'est-à-dire (4.1) pour $k \in \{1, \dots, N-2\}$ auxquelles s'ajoutent

$$\ddot{x}_0 = -e^{-(x_1 - x_0)} \quad \text{et} \quad \ddot{x}_{N-1} = e^{-(x_{N-1} - x_{N-2})}$$

Un élément essentiel dans notre approche, que l'on veut numériquement stable, est l'utilisation de l'algorithme des moments modifiés, décrit dans le chapitre précédent.

4.1 Résolutions théorique et pratique du système semi-infini de Toda

Lors de la résolution du système de Toda, une étape classique (voir par exemple [58, Chapter 2, §10]) est d'introduire la substitution

$$\begin{aligned} \alpha_k(t) &= \dot{x}_k(t), & k &= 0, \dots \\ \beta_k^2(t) &= e^{-(x_k(t) - x_{k-1}(t))}, & k &= 1, \dots \end{aligned}$$

Le lecteur vérifiera aisément la propriété

Propriété 4.1. *Le système de Toda semi-infini peut être écrit sous la forme suivante*

$$\begin{aligned}\dot{\alpha}_0(t) &= -\beta_1^2(t), \\ \dot{\alpha}_1(t) &= \beta_1^2(t) - \beta_2^2(t), \\ \dot{\alpha}_k(t) &= \beta_k^2(t) - \beta_{k+1}^2(t), \quad k \geq 2 \\ \dot{\beta}_1(t) &= \frac{1}{2}\beta_1(t)(\alpha_0(t) - \alpha_1(t)) \\ \dot{\beta}_k(t) &= \frac{1}{2}\beta_k(t)(\alpha_{k-1}(t) - \alpha_k(t)), \quad k \geq 2\end{aligned}$$

Dans le cas fini à N particules, on obtient pour la dernière quantité $\dot{\alpha}_{N-1}(t) = \beta_{N-1}^2(t)$, l'égalité pour $\dot{\beta}_{N-1}(t)$ restant inchangée.

Définition 4.2. *On pose :*

$$\mathcal{J}(t) = \begin{pmatrix} \alpha_0(t) & \beta_1(t) & & & & & \\ \beta_1(t) & \alpha_1(t) & \beta_2(t) & & & & \\ & \beta_2(t) & \alpha_2(t) & \beta_3(t) & & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \beta_{n-1}(t) & \alpha_{n-1}(t) & \beta_n(t) \\ & & & & & \cdot & \cdot & \cdot \end{pmatrix}$$

et

$$M(t) = \frac{1}{2} \begin{pmatrix} 0 & -\beta_1 & & & & & \\ \beta_1 & 0 & -\beta_2 & & & & \\ & \beta_2 & 0 & -\beta_3 & & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \beta_{n-2} & 0 & -\beta_{n-1} \\ & & & & & \cdot & \cdot & \cdot \end{pmatrix}$$

Les équations de la propriété 4.1 peuvent alors être représentées sous la forme suivante :

$$\dot{\mathcal{J}}(t) = [M(t), \mathcal{J}(t)] := M(t)\mathcal{J}(t) - \mathcal{J}(t)M(t)$$

où $[X, Y]$ représente le crochet de Lie.

Les matrices \mathcal{J} et M forment alors une paire de Lax, $\mathcal{J}(t)$ étant un opérateur symétrique sur ℓ^2 et $M(t)$ un opérateur antisymétrique.

Grâce à des techniques classiques (développement en série), on montre que $\dot{\mathcal{J}}(t) = [M(t), \mathcal{J}(t)]$ admet une solution unique dans l'ensemble des opérateurs bornés sur ℓ^2 dans un voisinage d'un réel t_0 , si on suppose que

$\mathcal{J}(t_0)$ est un opérateur borné (ou, ce qui est équivalent, si les coefficients $\alpha_k(t_0)$, $\beta_k(t_0)$ sont uniformément bornés).

Puisque l'on dispose de la paire de Lax, on peut montrer (voir par exemple [61, 46]) que le spectre de $\mathcal{J}(t)$ ne dépend pas de t . En particulier la norme $\|\mathcal{J}(t)\|$ de l'opérateur ne dépend pas non plus de t .

Nous avons supposé, initialement, que $\mathcal{J}(0)$ est borné. On peut donc en déduire que l'on dispose d'une solution unique pour tout réel t , dite solution globale.

En utilisant les mêmes arguments, ces propriétés valables dans le cas semi-infini le restent pour les systèmes infini ou fini (On pourra consulter [58] pour une étude complète du cas de dimension finie).

Une démarche pour obtenir, au moins théoriquement, la solution du système de Toda semi-infini peut être représentée comme suit :

$$\begin{array}{ccc}
 \mathcal{J}(0) & \xrightarrow{\dot{\mathcal{J}}(t)=(M\mathcal{J}-\mathcal{J}M)(t)} & \mathcal{J}(t) \\
 \downarrow & & \uparrow \\
 ds(x, 0) & \longrightarrow & ds(x, t) = e^{-xt} ds(x, 0) \\
 \mu(z, 0) & \longrightarrow & \mu(z, t) = \frac{\int \frac{ds(x,t)}{z-x}}{\int ds(x,t)} = (e_0, (zI - \mathcal{J}(t))^{-1} e_0)
 \end{array}$$

C'est-à-dire qu'au lieu de résoudre l'équation différentielle ordinaire $\dot{\mathcal{J}}(t) = (M\mathcal{J} - \mathcal{J}M)(t)$, on propose, en premier lieu, de résoudre le problème spectral direct suivant :

Il s'agit de trouver la mesure d'orthogonalité $ds(x, 0)$ associée à la matrice de Jacobi réelle $\mathcal{J}(0)$, que l'on considère comme la matrice des coefficients de récurrence de polynômes orthogonaux. On supposera, sans perte de généralité que $ds(x, 0)$ est une mesure de probabilité, c'est-à-dire $\int ds(x, 0) = 1$.

Cette détermination de la mesure est réalisée théoriquement, grâce au théorème spectral appliqué à $\mathcal{J}(0)$. Pour certaines classes assez restreintes de $\mathcal{J}(0)$, comme la perturbation de la classe trace des opérateurs à coefficients constants, ceci peut être réalisé grâce à des outils comme les déterminants de Túran [40] mais reste un problème pour des classes plus larges de $\mathcal{J}(0)$.

En second lieu, comme indiqué par le diagramme, l'évolution temporelle de la mesure pour le système de Toda est connue [46].

Finalement, il s'agit de résoudre un problème spectral inverse : D'après Nikishin-Sorokin, on peut déterminer la série de Taylor de la fonction de Weyl à l'infini c'est-à-dire calculer les moments ordinaires de $ds(x, t)$ et déterminer

la \mathcal{J} -fraction continue

$$\mu(z, t) = \frac{1}{z - \alpha_0(t) - \frac{\beta_1^2(t)}{z - \alpha_1(t) - \frac{\beta_2^2(t)}{\dots - \frac{\beta_{N-2}^2(t)}{z - \dots}}}}$$

grâce à l'algorithme de Chebyshev qui nécessite, avant tout, le calcul des moments ordinaires associés $\int x^k ds(x, t)$. On peut noter que certains préfèrent travailler avec les moments de la mesure de probabilité $d\tilde{s}(x, t) = \frac{ds(x, t)}{\int ds(x, t)}$ ce qui ne change en rien les coefficients de $\mathcal{J}(t)$ comme indiqué par le lemme 3.3 dans le chapitre précédent.

Dans le cas de la résolution d'un système fini, *a contrario*, on peut proposer des procédés très efficaces et numériquement stables. En suivant le principe de la définition 4.2, on note \mathcal{J}_N et M_N les matrices qui regroupent les inconnues $\alpha_0, \dots, \alpha_{N-1}$ et $\beta_1, \dots, \beta_{N-1}$ d'un système de Toda à N particules.

D'après la propriété 4.1, on obtient de nouveau une paire de Lax, $\dot{\mathcal{J}}_N(t) = [M_N(t), \mathcal{J}_N(t)]$.

$$\begin{array}{ccc} \mathcal{J}_N(0) & & \mathcal{J}_N(t) \\ \downarrow & & \uparrow \\ \mu_N(z, 0) & \longrightarrow & \mu_N(z, t) = (e_0, (zI_N - \mathcal{J}_N(t))^{-1}e_0) \end{array}$$

Rappelons [58, Chapter 2, §10] que $\mu_N(z, t)$ est une fonction rationnelle, plus précisément

$$\mu_N(z, t) = \sum_{j=0}^{N-1} \frac{m_j(t)}{z - \lambda_j} \quad (4.2)$$

où $\lambda_0, \dots, \lambda_{N-1}$ sont les valeurs propres distinctes de $\mathcal{J}_N(t)$, qui, d'après le principe de Lax, ne dépendent pas du temps et $\sqrt{m_j(t)}$ est la première composante du vecteur propre de $\mathcal{J}_N(t)$ associé à λ_j .

Tout d'abord, d'après Nikishin-Sorokin, grâce au principe de Lax, il est possible d'obtenir explicitement l'évolution temporelle de ces coefficients

$$m_k(t) = \frac{m_k(0)e^{-\lambda_k t}}{\sum_{j=0}^{N-1} m_j(0)e^{-\lambda_j t}}, \quad k \in \{0, \dots, N-1\}$$

à partir des conditions initiales.

De la même manière que dans le cas semi-infini, on notera

$$ds_N(x, t) = \sum_{j=0}^{N-1} e^{-t\lambda_j} m_j(0) d\delta_{\lambda_j}(x)$$

la mesure discrète, non normalisée, associée à $\mathcal{J}_N(t)$.

Deuxièmement, le passage des coefficients α_k, β_k aux données de quadrature λ_k, m_k et réciproquement, est une application non linéaire bien conditionnée d'après [9, Formule 7& Theorème 1].

En effet, le problème spectral direct, c'est-à-dire trouver $ds_N(x, 0)$ à partir de $\mathcal{J}_N(0)$, est une activité classique en Algèbre Linéaire : déterminer les valeurs propres et vecteurs propres d'une matrice symétrique. Dans notre cas, où la matrice est de plus tridiagonale, l'algorithme itératif QR [64, §5.4, p.200] est tout-à-fait adapté.

Enfin, pour résoudre le problème spectral inverse, on dispose de trois possibilités. En utilisant l'approche classique exposée par Nikishin et Sorokin [58, Chapter 2, §10], on pourrait calculer les moments $\int x^k ds_N(x, t)$ et ensuite utiliser l'algorithme de Chebyshev pour développer $\mu_N(x, t)$ en \mathcal{J} -fraction continue (finie).

Ici, on propose deux autres possibilités :

La première utilise l'observation faite dans [9, Lemma 2] que la matrice $P(t)$ des vecteurs propres normalisés, qui permet d'écrire évidemment l'égalité

$$\mathcal{J}_N(t) = P(t) \begin{pmatrix} \lambda_0 & & 0 \\ & \ddots & \\ 0 & & \lambda_{N-1} \end{pmatrix} P^t(t)$$

peut être obtenue en calculant la décomposition QR de la matrice

$$\begin{pmatrix} \sqrt{m_0(t)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{m_{N-1}(t)} \end{pmatrix} \begin{pmatrix} 1 & \lambda_0 & \dots & \lambda_0^{N-1} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 1 & \lambda_{N-1} & \dots & \lambda_{N-1}^{N-1} \end{pmatrix}$$

avec $Q = P(t)^t$ orthogonale et R triangulaire supérieure. Cette première méthode peut être schématisé sous la forme suivante.

$$\begin{array}{ccc} \mathcal{J}_N(0) & & \mathcal{J}_N(t) \\ \downarrow & & \uparrow \\ P_N(0), \lambda_j, m_j(0) & \xrightarrow{\text{dec. } QR} & P_N(t) \end{array} \quad (4.3)$$

La deuxième idée, que l'on préfère, utilise les moments modifiés définis par $\varphi_{N,k}(t) = \int p_k(x, 0) ds_N(x, t)$, $k = 0, \dots, N-1$ pour le système de dimension finie et $\varphi_k(t) = \int p_k(x, 0) ds(x, t)$, $k = 0, 1, \dots$ pour le système semi-infini. Les polynômes de référence (notés dans le chapitre précédent $\Pi_k(x)$) sont les polynômes $p_0(x, 0), p_1(x, 0), \dots$ orthonormés par rapport à la mesure $ds_N(x, 0)$

ou $ds(x, 0)$ (les polynômes $p_k(x, t)$ étant orthonormés par rapport à $ds(x, t)$). Cette double utilisation ne devrait pas poser problème car ultérieurement, la matrice $\mathcal{J}_N(0)$ sera la sous-matrice principale d'ordre N de la \mathcal{J}_0 . $ds_N(x, 0)$ correspondra alors à la N -ième quadrature de Gauss associé à $ds(x, 0)$. Ainsi les N premiers polynômes orthogonaux coïncident.

L'avantage de ces moments modifiés est que leur évolution temporelle est simple car elles vérifient une equation différentielle ordinaire linéaire.

Propriété 4.3. *Le vecteur des moments modifiés du système fini de Toda,*

$$\Phi_N(t) = (\varphi_{N,0}(t), \dots, \varphi_{N,N-1}(t))^t$$

est solution du problème de Cauchy

$$\dot{\Phi}_N(t) = -\mathcal{J}_N(0)\Phi_N(t), \quad \Phi_N(0) = e_0 \quad (4.4)$$

Dans le cas du système semi-infini de Toda, le vecteur $\Phi(t) = (\varphi_0(t), \varphi_1(t), \dots) \in \ell^2$ est solution du problème

$$\dot{\Phi}(t) = -\mathcal{J}(0)\Phi(t), \quad \Phi(0) = e_0 \quad (4.5)$$

Les relations de la propriété 4.3 semblent originales. Elles seront démontrées dans le paragraphe 4.4. Il est évidemment possible d'écrire un système d'équations différentielles ordinaires pour les moments classiques de $ds_N(x, t)$ ou $ds(x, t)$ mais la matrice de coefficients que l'on trouve est l'opérateur, non symétrique, "shift" qui se prête beaucoup moins bien à la fois à l'analyse et à la résolution numérique.

En effet, comme $ds(x, t)$ est une modification simple de $ds(x, 0)$ avec, en particulier, le même support, on peut affirmer, d'après [9, Théorème 11], que les passages du schéma suivant

$$\begin{array}{ccc} \mathcal{J}_N(0) & & \mathcal{J}_N(t) \\ & \searrow & \nearrow \\ \dot{\Phi}_N(t) = -\mathcal{J}_N(0)\Phi_N(t) & & \Phi_N(t) \end{array}$$

sont tous bien conditionnés. Les détails algorithmiques de cette approche sont spécifiés dans le chapitre précédent et plus précisément à la remarque 3.7 mais insistons, ici, sur la simplicité de l'approche envisagée avec un système d'équations différentielles ordinaires linéaires à coefficients constants et l'utilisation de l'algorithme de Chebyshev modifié.

Comment peut-on alors tirer profit de ces indications sur la résolution effective d'un système fini-dimensionnel dans la résolution du système semi-infini ?

L'idée de base est que si $\mathcal{J}_N(0)$ est la sous matrice principale d'ordre N de $\mathcal{J}(0)$ alors les n premiers coefficients du système fini à N particules (c'est-à-dire ceux de la sous matrice principale d'ordre n de $\mathcal{J}_N(t)$) diffèrent peu des coefficients correspondant du système semi-infini (c'est-à-dire les coefficients de la sous matrice principale d'ordre n de $\mathcal{J}(t)$) uniformément pour $t \in [0; T]$, au moins si $N (= N(T) \gg n)$ est suffisamment grand. Cette idée de troncature à l'indice N se justifie par la comparaison des systèmes d'équations différentielles linéaires pour les moments modifiés : on montrera implicitement dans la démonstration de la propriété 4.8, avec la majoration (4.12), que, uniformément pour $t \in [0; T]$, le moment modifié $\varphi_N(t)$ tend vers zéro pour $N \rightarrow \infty$ d'une manière très rapide. Le remplacement de l'inconnue $\varphi_N(t)$ par zéro dans le système (4.5) nous donne, pour les premiers moments modifiés, exactement le système (4.4) correspondant à N particules.

Pour être mathématiquement plus précis, nous introduisons les projections suivantes

Définition 4.4. *On définit les projections*

$$E_n : \quad \ell^2 \quad \rightarrow \quad \mathbb{C}^n$$

$$\begin{pmatrix} y_0 \\ y_1 \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \end{pmatrix}$$

On va donc transformer le problème infindimensionnel en un problème de dimension finie sous la forme suivante

$$\begin{array}{ccccc} \mathcal{J}(0) & \longrightarrow & \mathcal{J}(t) & \longrightarrow & E_n \mathcal{J}(t) E_n^* \\ \text{troncature} \downarrow & & & & \simeq \\ \mathcal{J}_N(0) = E_N \mathcal{J}(0) E_N^* & \longrightarrow & \mathcal{J}_N(t) & \longrightarrow & E_n E_N^* \mathcal{J}_N(t) E_N E_n^* \end{array} \quad (4.6)$$

Le but des considérations suivantes est d'estimer l'erreur commise

$$\|E_n E_N^* \mathcal{J}_N(t) E_N E_n^* - E_n \mathcal{J}(t) E_n^*\|. \quad (4.7)$$

c'est-à-dire l'erreur de troncature à l'indice N pour les n premiers coefficients. En particulier nous cherchons à déterminer N , pour n , T , $\mathcal{J}(0)$ donnés, pour lequel l'erreur relative commise reste inférieure à une précision donnée, uniformément pour $t \in [0; T]$.

Comme indiqué précédemment, les moments modifiés obtenus par (4.4) peuvent être interprétés comme la solution d'une petite perturbation du système semi-infini (4.5) et donc comme une perturbation des N premiers moments modifiés du système de Toda semi-infini. On retrouve ainsi le thème

développé dans l'article [9] lorsque l'on cherche le conditionnement de l'application non linéaire qui associe les moments modifiés aux coefficients de récurrence $(\alpha_0, \dots, \alpha_{n-1}, \beta_0, \dots, \beta_{n-1})$.

Ici, on propose une estimation de la quantité

$$\left\| \left(\frac{\partial(\alpha_j, \beta_j)}{\partial \varphi_k} \right)_{\substack{j=0, \dots, n-1 \\ k=0, \dots, 2n-1}} \Delta \Phi_{N,n} \right\|, \quad \Delta \Phi_{N,n}(t) = \begin{pmatrix} \varphi_{N,0}(t) - \varphi_0(t) \\ \vdots \\ \varphi_{N,2n-1}(t) - \varphi_{2n-1}(t) \end{pmatrix}$$

qui en première approximation coïncide avec l'erreur (4.7).

Ensuite, on présentera quelques exemples numériques. La dernière partie du chapitre sera consacrée à la démonstration de l'estimation. Ici, on fera appel à quelques égalités proposées par Fischer [30] ainsi que des techniques développées dans [9].

4.2 Estimations

Theorème 4.5. *Soit le système dynamique infini défini précédemment :*

$$\dot{\mathcal{J}}(t) = M(t)\mathcal{J}(t) - \mathcal{J}(t)M(t)$$

Soient n, N des entiers, $T > 0$ un réel avec $N > \frac{5}{4}T\|\mathcal{J}(0)\| + 4$ et $N \geq 2n$. On suppose que $\mathcal{J}(0)$ est bornée. Alors, en première approximation, l'erreur relative de troncature dans la résolution d'un système semi-infini de Toda peut être estimé uniformément pour $t \in [0; T]$ par :

$$\begin{aligned} & \frac{\|E_n E_N^* \mathcal{J}_N(t) E_N E_n^* - E_n \mathcal{J}(t) E_n^*\|_2}{\|(\alpha_0, \dots, \alpha_{n-1}, \beta_0, \dots, \beta_{n-1})\|_2} \\ & \lesssim e^{2T\|\mathcal{J}(0)\|} \frac{3200}{9} T \|\mathcal{J}(0)\| \left(\frac{5eT\|\mathcal{J}(0)\|}{8(2N - 2n - 4)} \right)^{2N-2n-4} \sqrt{\int \left(\sum_{k=0}^{n-1} p_k^2(x, 0) \right)^2 ds(x, 0)} \end{aligned} \quad (4.8)$$

L'approximation au premier ordre est justifiée car pour tout $t \in [0; T]$, et pour tout $k \leq 2n - 1$, $\Delta \varphi_k(t)$ étant défini par $\varphi_{N,k}(t) - \varphi_k(t)$,

$$|\Delta \varphi_k(t)| \leq \frac{1600}{9} T \left(\frac{5eT\|\mathcal{J}(0)\|}{8(2N - 5 - k)} \right)^{2N-5-k} \quad (4.9)$$

L'intégrale contenant le noyau de Szëgo $\sum_{k=0}^{n-1} p_k(x, 0)^2$ présent dans l'estimation précédente possède une croissance modérée dans de nombreux cas. En effet, si les polynômes de départ sont les polynômes de Chebyshev, on peut majorer l'intégrale par $(n+1)(n+2)$ [9, Exemple 15 & sq].

Grâce à l'estimation (4.8), on peut déduire qu'à T et n fixés, pour obtenir la précision souhaitée, il suffit de choisir un entier N suffisamment grand.

4.3 Applications

On va étudier l'évolution temporelle de la matrice $\mathcal{J}(t)$ dont on connaît seulement la donnée initiale $\mathcal{J}(0)$. On applique l'algorithme des moments modifiés pour déterminer les coefficients des n premières récurrences à un temps T donné.

La matrice $\mathcal{J}(0)$ contient les coefficients de récurrence des polynômes de référence $p_k(x, 0)$. On choisit un entier de troncature $N \geq 2n$ et $N \geq \frac{5}{4}T\|\mathcal{J}(0)\| + 3$.

Comme les polynômes sont connus, cela permet d'obtenir les données de quadrature $m_0(0), \dots, m_{N-1}(0)$ et $\lambda_0, \dots, \lambda_{N-1}$ en effectuant la décomposition QR de la matrice $\mathcal{J}_N(0)$. Ces données sont utiles pour connaître le vecteur des moments modifiés perturbés $\Phi_N(t)$, $t \leq T$.

Les polynômes de référence $p_k(x, 0)$ sont orthonormés. On doit donc effectuer une petite modification de l'algorithme des moments modifiés pour prendre en compte les coefficients de récurrence. Cette modification est une application de l'algorithme du cas vectoriel avec $r = s = 1$.

Propriété 4.6. *Les relations suivantes servent à initialiser l'algorithme (les modifications apportées par rapport à la propriété 3.6 donnée dans la partie sur les moments modifiés sont en gras) :*

$$\begin{aligned}\sigma_{-1,k} &= 0, & k &= 1, 2, \dots, 2n-2 \\ \sigma_{0,k} &= m_k, & k &= 0, 1, \dots, 2n-1 \\ \alpha_0 &= a_0 + \mathbf{b}_1 \frac{m_1}{m_0} \\ \beta_0 &= m_0 \quad (\text{par convention})\end{aligned}$$

Le calcul effectif des coefficients se fait de la manière suivante :

Pour j variant de 1 à $n-1$

Pour k variant de j à $2n-j-1$

$$\begin{aligned}\sigma_{j,k} &= \mathbf{b}_{k+1}\sigma_{j-1,k+1} - (\alpha_{j-1} - a_k)\sigma_{j-1,k} - \beta_{j-1}\sigma_{j-2,k} + b_k\sigma_{j-1,k-1} \\ \alpha_j &= a_j + \mathbf{b}_{j+1} \frac{\sigma_{j,j+1}}{\sigma_{j,j}} - \mathbf{b}_j \frac{\sigma_{j-1,j}}{\sigma_{j-1,j-1}} \\ \beta_j &= \mathbf{b}_j \frac{\sigma_{j,j}}{\sigma_{j-1,j-1}}\end{aligned}$$

L'utilisation de l'algorithme produit les coefficients de récurrence de polynômes unitaires. Pour trouver les coefficients des polynômes orthonormés correspondants, on calculera la racine carrée des $(\beta_k)_k$.

4.3.1 Mise en oeuvre

Comme indiqué précédemment, on détermine les valeurs propres et vecteurs propres de la matrice $\mathcal{J}_N(0)$ grâce à l'algorithme QR . Ensuite, en utilisant les formules (4.11) et (4.14), on obtient

$$\Phi_N(t) = (e^{-t\lambda_0}m_0(0)p_k(\lambda_0, 0)p_0(\lambda_0, 0) + \dots \\ \dots + e^{-t\lambda_{N-1}}m_{N-1}(0)p_k(\lambda_{N-1}, 0)p_0(\lambda_{N-1}, 0))_{k=0, \dots, N-1}$$

Ceci fournit les moments modifiés à tout temps $t \in [0; T]$. Le programme conserve en mémoire les quantités précédentes exceptée l'exponentielle. On peut donc étudier l'évolution en temps comme dans l'algorithme de Runge-Kutta. Grâce à l'algorithme de Chebyshev modifié, on peut donc obtenir les coefficients de récurrence des polynômes $p_k(x, t)$.

Les calculs sont effectués grâce à Maple V Release 5 sur un Macintosh G3 à 400Mhz. Pour effectuer la décomposition QR , on utilise la fonction implémentée dans Maple. Dans les exemples qui suivent, on indique le nombre n de coefficients de récurrence recherchés, le nombre de troncature N , la valeur limite de temps T et le nombre d'étapes de l'évolution dynamique. Lorsque cela est possible, on donne la mesure $ds(x, 0)$ associée à la matrice initiale et donc la mesure $ds(x, t) = e^{-tx}ds(x, 0)$. Cela est réalisé grâce à la formule de Stieltjes-Perron [58, Chapter 2, §9.2]. La connaissance de la mesure permet de vérifier l'orthogonalité par rapport à $ds(x, t)$ des polynômes $p_k(x, t)$, définis à partir des coefficients de récurrence calculés par l'algorithme de Chebyshev modifié.

Puisqu'il s'agit, initialement, d'un problème physique, on essaiera, dans la mesure du possible, de présenter les déplacements des particules.

Dans les exemples qui suivent, la valeur de T est très grande. Avec les quantités choisies, le membre de droite de l'estimation (4.8) est de l'ordre de 10^{10} dans le cas des polynômes de Chebyshev, par exemple. Si T vaut 2 l'estimation passe alors à 10^{-10} . Comme les orthogonalités sont cependant respectées, on peut en conclure que l'estimation ne donne qu'une indication de l'erreur commise et qu'elle est parfois trop pessimiste.

Le code Maple de l'exemple 3 est fourni à l'annexe A.2.

4.3.2 Exemples

1. On considère

n	16
N	32
T	12
boucle	60

$$\mathcal{J}(0) = \begin{pmatrix} 0 & \frac{\sqrt{2}}{2} & 0 & \dots & \\ \frac{\sqrt{2}}{2} & 0 & \frac{1}{2} & 0 & \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ddots \\ & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

Les polynômes de référence initiaux sont les polynômes de Chebyshev de première espèce. La mesure associée est $ds(x, t) = e^{-tx} \frac{1}{\sqrt{1-x^2}}$.

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Configuration initiale.

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 1

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 2

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 3

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 4

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 5

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 6

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 11

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 20

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 30

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 50

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 60.

Explicitons les étapes ci-dessus.

D'après les substitutions définies au début de 4.1 et la forme de la matrice initiale, la diagonale contient les vitesses initiales des particules. Elle est donc nulle pour toutes les particules. Pour déterminer la position relative de chaque particule par rapport à la particule 0, il suffit de déterminer le logarithme de chacun des termes β_k de la surdiagonale et d'effectuer des soustractions successives. La substitution indique aussi que les β_k doivent être compris strictement entre 0 et 1.

Dans tous les exemples, comme la particule 0 n'est soumise à aucune force à gauche, elle part vers la gauche. Les autres se déplaceront aussi, finalement, vers la gauche.

Étudions maintenant notre exemple.

La particule 0 est plus proche de la particule 1 que la particule 2. De ce fait, la particule 1 se déplace vers la droite tandis que la particule 0 se déplace vers la gauche comme indiqué précédemment (Étape 1 à 11).

La particule 1, devenant très proche de la particule 2, va repartir vers la gauche car elle ne subit plus de force de la part de la particule 0. La particule 2 subit une force plus grande de la part de la particule 1 que de la particule 3 car elle est plus proche. La particule 2 part donc vers la droite (Étape 20). Elle transmet sa vitesse à la particule 3 avant de repartir vers la gauche. Et on peut réitérer ces explications avec les particules 3 (Étape 50), 4 (Étape 60), 5...

La vitesse se transmet ainsi de particule en particule, on obtient une sorte de "soliton" (voir [76]).

3. (Code Maple à l'annexe A.2) On considère cette fois-ci des particules équidistantes, la première étant munie d'une vitesse initiale dirigée vers la seconde particule.

n	16
N	32
T	12
boucle	60

$$\mathcal{J}(0) = \begin{pmatrix} 1 & \frac{1}{2} & 0 & \dots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ddots \\ & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

La mesure associée est $ds(x, t) = e^{-tx} \left(\frac{2\sqrt{1-x^2}}{\pi-4x+5} + \frac{3}{4}\delta_{5/4}(x) \right)$.

◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Configuration initiale.

◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 3

◆◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 6

◆◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 9

◆◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 12

◆ ◆◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 15

◆ ◆◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 20

◆ ◆◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 30

◆ ◆◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 50

◆ ◆◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆ ◆

Etape 60

L'évolution est bien conforme à l'intuition physique. La première particule se déplace vers la droite et transmet le mouvement à la seconde particule. Puis, en raison de la force exercée, la première particule repart vers la gauche. Le reste de l'évolution ressemble à l'exemple précédent.

6. On considère cette fois-ci le cas où $N = n$, les particules vont donc se comporter comme si elles étaient en nombre fini, c'est-à-dire que les particules 0 à 7 vont partir vers la gauche, les particules 8 à 15 vont partir vers la droite de manière symétrique à cause du choix de la configuration initiale.

n	16
N	16
T	8
boucle	40

$$\mathcal{J}(0) = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \dots & \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ddots \\ & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊

Configuration initiale.

◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊

Etape 10

◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊

Etape 20

◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊

Etape 30

◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊ ◊

Etape 40

7. Considérons de nouveau l'exemple 1 mais en tronquant à une valeur de N moindre. Les valeurs particulières de T et du nombre de boucles dans le tableau sont dues à une erreur à l'itération 53 qui provoque l'arrêt de l'algorithme : la troncature est trop faible pour poursuivre à un temps T plus important. On va comparer les deux évolutions.

n	16
N	20
T	10.4
boucle	52

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Configuration initiale.

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 30

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 35

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 40

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 45

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 50

◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇ ◇

Etape 52

Jusqu'à l'étape 35, aucune modification, par rapport au premier exemple de référence, n'est décelée. A partir de l'étape 40, c'est-à-dire au temps $t = 8$, la particule de droite commence à se diriger vers la droite, les étapes suivantes ne sont donc plus conformes à la réalité d'un système semi-infini. Il est quand même remarquable que les calculs restent valables jusqu'à une valeur de t relativement importante, malgré une valeur N de troncature assez faible. Si on augmentait la valeur N de troncature, le temps t où les calculs restent valables augmente lui aussi. On peut aussi choisir $N > 2n$, ce qui permet de repousser encore la limite temporelle.

4.4 Démonstrations

Avant de présenter une démonstration du théorème 4.5, rappelons que le spectre $\sigma(\mathcal{J}(t)) = \sigma(\mathcal{J}(0))$ coïncide avec le support de la mesure d'orthogonalité $ds(x, t)$ associée. Montrons que l'on peut simplement considérer le cas $\|\mathcal{J}(0)\| = 1$ (en conséquence, $s(x, t)$ admet un support inclus dans l'intervalle $[-1; 1]$.) :

Effectivement, en utilisant la substitution $\tilde{T} = T \cdot \|\mathcal{J}(0)\|$, $\tilde{t} = t \cdot \|\mathcal{J}(0)\|$, $\tilde{\alpha}_j(\tilde{t}) = \alpha_j(t)/\|\mathcal{J}(0)\|$ et $\tilde{\beta}_j(\tilde{t}) = \beta_j(t)/\|\mathcal{J}(0)\|$ et évidemment $\tilde{\mathcal{J}}(\tilde{t}) = \mathcal{J}(t)/\|\mathcal{J}(0)\| = \mathcal{J}(t)/\|\mathcal{J}(0)\|$ et $\tilde{M}(\tilde{t}) = M(t)/\|\mathcal{J}(0)\|$, nous obtenons $t \in [0; T]$ si et seulement si $\tilde{t} \in [0; \tilde{T}]$ et

$$\frac{\partial \tilde{\mathcal{J}}}{\partial \tilde{t}}(\tilde{t}) = \frac{1}{\|\mathcal{J}(0)\|^2} \frac{\partial \mathcal{J}}{\partial t}(t) = [\tilde{M}(\tilde{t}), \tilde{\mathcal{J}}(\tilde{t})]$$

Dans la suite de cette section, pour éviter les lourdeurs de notation, on omettra les tildes.

Le reste du chapitre est composé de 5 parties. On commence par présenter la démonstration de la propriété 4.3. Ensuite on donne en lemme 4.7 une formule intégrale pour la différence $\Delta\Phi_N(t)$ des deux vecteurs de moments modifiés en fonction de $\varphi_N(t)$. Le comportement de $\varphi_N(t)$ est l'objet de la proposition 4.8 et de ses lemmes annexes 4.9, 4.10 et 4.11. Ceci nous permet dans la quatrième partie d'établir la deuxième estimation (4.9) du théorème. Finalement, nous montrons la première estimation (4.8).

Démonstration. [**Propriété 4.3**] Démontrons les formules du cas semi-infini. $\mathcal{J}(0)$ représente les coefficients de récurrence des polynômes orthogonaux $p_k(x, 0)$. Or $\varphi_k(t) = \int_{-1}^1 p_k(x, 0)e^{-tx} ds(x, 0)$. Dérivons par rapport au temps :

$$\begin{aligned} \frac{d\varphi_k}{dt} &= \int_{-1}^1 (-xp_k(x, 0))e^{-tx} d\sigma(x, 0) \\ &= \int_{-1}^1 ((-\beta_k(0)p_{k-1}(x, 0) - \alpha_k(0)p_n(x, 0) - \beta_{k+1}(0)p_{k+1}(x, 0))e^{-tx} d\sigma(x, 0) \\ &= -\beta_k(0)\varphi_{k-1}(t) - \alpha_k(0)\varphi_k(t) - \beta_{k+1}(0)\varphi_{k+1}(t) \end{aligned}$$

De plus, au temps 0, d'après l'orthogonalité des polynômes $p_k(x, 0)$, on a

$$\begin{aligned} \varphi_0(0) &= \int_{-1}^1 p_0(x, 0) ds(x, 0) = 1 \\ \varphi_k(0) &= \int_{-1}^1 p_k(x, 0) ds(x, 0) = 0, \quad k \leq 1 \end{aligned}$$

ce qui prouve la formule (4.5).

Cette démonstration peut être reproduite pour le cas fini. Pour $k < N-1$, les égalités précédentes restent vraies. Pour $k = N-1$, $\forall t \in \mathbb{R}$, $\varphi_{N,N}(t) = 0$.

En effet, $p_N(x, t)$ est le polynôme caractéristique de $\mathcal{J}_N(t)$. Ce polynôme s'annule donc en toutes les abscisses d'intégration qui sont les valeurs propres de $\mathcal{J}_N(t)$, indépendantes du temps. La formule de Gauss-Christoffel est exacte pour $p_N(x, t)$ donc l'intégrale précédente s'annule bien en tout temps t . Ceci prouve ainsi la formule (4.4). \square

Passons maintenant à un lemme qui décrit l'erreur commise lorsque l'on résout le système de Cauchy de dimension finie en lieu et place du cas semi-infini :

Lemme 4.7. *On considère le vecteur des moments modifiés du système fini de Toda*

$$\Phi_N(t) = (\varphi_{N,0}(t), \dots, \varphi_{N,N-1}(t))^t$$

et les N premières composantes du vecteur $\Phi(t)$ des moments modifiés du cas semi-infini

$$(\varphi_0(t), \dots, \varphi_{N-1}(t))^t$$

L'erreur commise est la suivante

$$\Delta\Phi_{N,N/2}(t) := \begin{pmatrix} \varphi_{N,0}(t) \\ \vdots \\ \varphi_{N,N-1}(t) \end{pmatrix} - \begin{pmatrix} \varphi_0(t) \\ \vdots \\ \varphi_{N-1}(t) \end{pmatrix} = \int_0^t e^{(t-u)(-\mathcal{J}_N(0))} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \beta_N(0)\varphi_N(u) \end{pmatrix} du \quad (4.10)$$

Démonstration. On a démontré à la propriété 4.3 que les vecteurs des moments modifiés vérifiaient des systèmes d'équations différentielles ordinaires linéaires. Dans le cas semi-infini, on a

$$\frac{\partial\Phi(t)}{\partial t} = -\mathcal{J}(0)\Phi(t)$$

On tronque l'équation au rang N en multipliant chaque membre de l'équation à gauche par

$$\begin{pmatrix} I_N & 0_{N \times \infty} \end{pmatrix}$$

On obtient donc une nouvelle équation

$$\frac{\partial}{\partial t} \begin{pmatrix} \varphi_0(t) \\ \vdots \\ \varphi_{N-1}(t) \end{pmatrix} = -\mathcal{J}_N(0) \begin{pmatrix} \varphi_0(t) \\ \vdots \\ \varphi_{N-1}(t) \end{pmatrix} - \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \beta_N(0)\varphi_N(t) \end{pmatrix}$$

où β_N est le coefficient de la surdiagonale de $\mathcal{J}(0)$.

Résolvons cette équation en temps sans l'apport de $(0, \dots, 0, \beta_N(0)\varphi_N(t))$

c'est-à-dire l'équation sans second membre (4.4) vérifiée par le vecteur des moments modifiés du système fini de Toda. La donnée initiale est $(1, 0, \dots)^t$. On trouve donc,

$$\Phi_N(t) = \begin{pmatrix} \varphi_{N,0}(t) \\ \vdots \\ \varphi_{N,N-1}(t) \end{pmatrix} = e^{-\mathcal{J}_N(0)t} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (4.11)$$

D'après la théorie sur les équations différentielles [27, Chapitre 7, §2.4], l'erreur commise est donc

$$\begin{pmatrix} \varphi_{N,0}(t) \\ \vdots \\ \varphi_{N,N-1}(t) \end{pmatrix} - \begin{pmatrix} \varphi_0(t) \\ \vdots \\ \varphi_{N-1}(t) \end{pmatrix} = \int_0^t e^{(t-u)(-\mathcal{J}_N(0))} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \beta_N(0)\varphi_N(u) \end{pmatrix} du$$

□

On désire maintenant étudier le comportement de $\varphi_N(t)$ qui est présent dans la formule d'erreur précédente (4.10). On donne donc la propriété suivante qui sera démontrée grâce aux différents lemmes 4.9, 4.10 et 4.11 que l'on présente.

Propriété 4.8.

$$|\varphi_N(t)| \leq \frac{40R^{-N+2} \exp(t(R + 1/R)/2)}{3} \quad (4.12)$$

On commence par prouver quelques lemmes qui permettront d'alléger la démonstration de cette propriété.

Lemme 4.9. *Soit R un réel strictement supérieur à 1. Soit \mathcal{C} l'ellipse de demi-axes $1/2(R + 1/R)$ et $1/2(R - 1/R)$ de foyers $(-1; 0)$ et $(1; 0)$ et E le segment de l'axe réel $[-1; 1]$. La distance entre l'ellipse \mathcal{C} et le segment E est égale à $1/2(R + 1/R) - 1$.*

Démonstration. On cherche à minimiser la distance entre l'ellipse et le segment. On minimise donc la quantité $|z - x|$ où z est l'affixe d'un point de l'ellipse et x l'abscisse d'un point du segment.

$$z(\theta) = 1/2(R + 1/R) \cos(\theta) + i/2(R - 1/R) \sin(\theta) = J(Re^{i\theta}), \quad \theta \in [0; 2\pi].$$

où $J(y) = 1/2(y + 1/y)$ est la fonction de Joukowski. On peut étudier cette minimisation uniquement dans le premier quadrant grâce aux symétries.

Lorsque l'abscisse du point de l'ellipse est comprise entre 0 et 1, la distance de l'ellipse au segment est égale à l'ordonnée du point de l'ellipse. Donc, pour cette partie de l'ellipse, le minimum est atteint au point d'abscisse 1. L'argument θ de l'affixe z vaut $\arccos(2/(R + 1/R))$. Donc l'ordonnée de ce point, c'est-à-dire le minimum sur cette partie, est

$$\frac{R - 1/R}{2} \cdot \sin(\arccos(\frac{2}{R + 1/R})) = \frac{R - 1/R}{2} \cdot \sqrt{1 - \frac{4}{(R + 1/R)^2}} = \frac{(R - 1/R)^2}{2(R + 1/R)}$$

Si l'abscisse du point de l'ellipse est plus grande que 1, on étudie le minimum sur z de $|z - 1|$. Si on calcule explicitement ce module en fonction de θ , on trouve

$$|z-1| = \sqrt{1/4(R + 1/R)^2 \cos^2(\theta) + 1 - (R + 1/R) \cos(\theta) + 1/4(R - 1/R)^2 \sin^2(\theta)}.$$

On dérive cette fonction, le numérateur vaut

$$-1/4(R + 1/R)^2 \sin(2\theta) + 1/4(R - 1/R)^2 \sin(2\theta) + (R + 1/R) \sin(\theta)$$

soit

$$-\sin(2\theta) + 1/2(R + 1/R)$$

Or $1/2(R + 1/R) > 1$ donc le numérateur est positif. Donc la fonction est croissante sur $[0; \arccos(2/(R + 1/R))]$. Or cette fonction est positive donc le minimum est atteint en 0. Donc le minimum de $|z-1|$ est $1/2(R+1/R)-1$. \square

Lemme 4.10. Pour n entier non nul et $R \geq 2$,

$$R^{2n} - 1 \geq \frac{3}{4}R^{2n}$$

l'égalité n'étant vraie que pour $R = 2$ et $n = 1$.

Démonstration. On minore la quantité $(R^{2n} - 1)/R^{2n}$

$$\begin{aligned} \frac{R^{2n} - 1}{R^{2n}} &= 1 - \frac{1}{R^{2n}} \\ &\geq 1 - \frac{1}{2^{2n}} \\ &\geq 1 - \frac{1}{4} \\ &\geq \frac{3}{4} \end{aligned}$$

\square

Lemme 4.11. Soit $R \geq 2$, n un entier non nul et $n/t \geq 2$. Soit $g_n(R) = \frac{\exp(tR)}{R^n}$. Cette fonction atteint un minimum sur $[2; +\infty[$ en $\frac{n}{t}$ de valeur $\left(\frac{et}{n}\right)^n$.

Démonstration. Dérivons la fonction g_n . On obtient

$$\begin{aligned} g'_n(R) &= \frac{t \exp(tR)}{R^n} - \frac{n \exp(tR)}{R^{n+1}} \\ &= \frac{(Rt - n) \exp(tR)}{R^{n+1}} \end{aligned}$$

Pour $R = n/t$, $g'_n(R) = 0$. Sur $[0; n/t]$, $g'_n < 0$. Sur $[n/t; +\infty[$, $g'_n > 0$. Donc g_n atteint un minimum en $R = n/t$, de valeur $\left(\frac{et}{n}\right)^n$.

Cette quantité est équivalente à $\frac{t^n \sqrt{2\pi n}}{n!}$. □

On peut maintenant démontrer la propriété 4.8

Démonstration. Soit $f(x) = \exp(-xt)$ pour $x \in [-1; 1]$. Soit q_n le polynôme interpolant la fonction f aux zéros x_1, \dots, x_n du polynôme de Chebyshev : $q_n(x) = c_n \prod_j (x - x_j)$. Alors d'après la formule de Newton :

$$\varepsilon := \max_{x \in [-1; 1]} |f(x) - q_n(x)| = \max_{x \in [-1; 1]} \left| \frac{1}{2i\pi} \int_{\mathcal{C}} \frac{c_n(x - x_1) \dots (x - x_n)}{c_n(z - x_1) \dots (z - x_n)} \frac{f(z)}{(z - x)} dz \right|$$

où \mathcal{C} est un contour qui entoure x_1, \dots, x_n, x une seule fois.

Or $\max_{x \in [-1; 1]} |q_n(x)| = 1$ car q_n est un polynôme de Chebyshev. Donc

$$\varepsilon \leq \frac{1}{2\pi} \max_{x \in [-1; 1]} \int_{\mathcal{C}} \left| \frac{f(z)}{(z - x)q_n(z)} \right| dz.$$

On choisit comme contour \mathcal{C} , l'ellipse $z(\theta) = 1/2(R + 1/R) \cos(\theta) + i/2(R - 1/R) \sin(\theta) = J(Re^{i\theta})$, $\theta \in [0; 2\pi]$. $J(y) = 1/2(y + 1/y)$ est la fonction de Joukowski.

Or $q_n(J(y)) = J(y^n)$. Donc $|q_n(J(Re^{it}))| \geq 1/2(R^n - 1/R^n)$, cette valeur étant atteinte en $t = \pi$.

$\forall z, |\exp(-tz)| \leq \exp(t(R + 1/R)/2)$.

On utilise le lemme 4.9 ce qui donne $|z - x| \geq 1/2(R + 1/R) - 1$.

Il nous reste la longueur du contour. On peut la majorer par le périmètre du cercle de rayon le demi grand axe de l'ellipse $1/2(R + 1/R)$. Rassemblons les différentes inégalités :

$$\varepsilon \leq \frac{(R + 1/R)/2 \cdot \exp(t(R + 1/R)/2)R}{1/2(R^n - 1/R^n)(1/2(R + 1/R) - 1)}$$

C'est-à-dire

$$\max_{x \in [-1;1]} |f(x) - q_n(x)| \leq \frac{2R^{n+1}(R^2 + 1) \cdot \exp(t(R + 1/R)/2)}{(R^{2n} - 1)(R - 1)^2} \quad (4.13)$$

On obtient donc pour $\varphi_N(t)$.

$$\begin{aligned} \varphi_N(t) &= \int_{-1}^1 p_N(x, 0) e^{-tx} ds(x, 0) \\ &= \int_{-1}^1 (e^{-tx} - q_{N-1}(x)) p_N(x, 0) ds(x, 0) \\ |\varphi_N(t)|^2 &\leq \max_{x \in [-1;1]} |e^{-tx} - q_{N-1}(x)|^2 \underbrace{\int_{-1}^1 (p_N(x, 0))^2 ds(x, 0)}_{=1} \underbrace{\int_{-1}^1 ds(x, 0)}_{=1} \\ &\leq \frac{4R^{2N}(R^2 + 1)^2 \exp(t(R + 1/R))}{(R^{2N-2} - 1)^2 (R - 1)^4} \end{aligned}$$

Choisissons $R \geq 2$. Comme la fonction $\frac{R^2 + 1}{(R - 1)^2}$ est une fonction positive strictement décroissante sur $[2; +\infty[$, ceci nous donne

$$\frac{R^2 + 1}{(R - 1)^2} \leq 5.$$

Alors

$$|\varphi_N(t)| \leq \frac{10R^N \exp(t(R + 1/R)/2)}{(R^{2N-2} - 1)}$$

D'où, en utilisant le lemme 4.10 précédent

$$|\varphi_N(t)| \leq \frac{40R^{-N+2} \exp(t(R + 1/R)/2)}{3}$$

□

Remarque 4.12. Comme indiqué précédemment dans l'introduction, on a obtenu au cours de la démonstration avec la majoration (4.12), pour $|\varphi_N(t)|$, une décroissance géométrique à T fixé en fonction de N . Ceci permet de justifier la troncature effectuée lors de la résolution de l'équation différentielle.

Démontrons maintenant le théorème.

On utilise l'égalité (4.10) donnant l'erreur commise lorsque l'on résout une équation différentielle avec un second membre si l'on omet ce second membre. Cette erreur influe sur chaque moment modifié calculé jusqu'au rang n . On peut alors calculer les différentielles relativement à ces moments modifiés de chacun des coefficients de récurrence.

Enfin, on calcule la norme 2 de ces quantités afin d'obtenir l'estimation du théorème. La proposition 4.8 sera utilisée à la fin de la démonstration. Précisons quelques notations et propriétés supplémentaires qui seront utilisées :

Définition 4.13. On notera $K_n(x, y, t)$ le noyau de Szëgo [75, (3.1.9)] associé au temps t . C'est-à-dire

$$K_n(x, y, t) = \sum_{k=0}^n p_k(\bar{x}, t) p_k(y, t)$$

Le noyau de Szëgo peut être utilisé pour donner une représentation intégrale de la somme partielle d'indice n du développement en série de Fourier d'une fonction f . Ceci permet d'obtenir la propriété suivante appelée propriété du noyau reproduisant [75, (3.1.12)] :

Propriété 4.14. $\int K_n(x, y, 0) p(x) d\sigma(x, 0) = p(y)$ où p est un polynôme de degré inférieur ou égal à n .

Démonstration. [Estimation (4.8)]

On cherche à déterminer les coefficients de récurrence (α_k, β_k) composant la matrice $\mathcal{J}(t)$ pour $k \leq n$. Pour cela on effectue une troncature du système au rang $N \gg n$. Les coefficients de récurrence sont calculés grâce à l'algorithme de Chebyshev modifié.

Les moments modifiés sont perturbés par la troncature. Comme indiqué précédemment, on cherche donc à déterminer

$$\frac{\partial \alpha_j(t)}{\partial \varphi_k} \Delta \varphi_k \text{ et } \frac{\partial \beta_j(t)}{\partial \varphi_k} \Delta \varphi_k, \quad j = 0 \dots n-1, \quad k = 0 \dots 2n-1$$

Les dérivées partielles de α_j et β_j ont été données par H. J. Fischer dans [30].

$$\begin{aligned} \frac{\partial \alpha_j(t)}{\partial \varphi_k} &= [\beta_{j+1}(t) c_k(p_j(x, t) p_{j+1}(x, t)) - \beta_j(t) c_k(p_j(x, t) p_{j-1}(x, t))] \\ \frac{\partial \beta_j(t)}{\partial \varphi_k} &= \beta_j(t) c_k(p_j^2(x, t) - p_{j-1}^2(x, t)) \end{aligned}$$

où, pour p polynôme, $c_k(p) = \int p_k(x, 0) p(x) d\sigma(x, 0)$. C'est-à-dire que $c_k(p)$ est le coefficient de $p_k(x, 0)$ dans la décomposition de p sur cette base de polynômes.

L'erreur commise sur le calcul de φ_k lors de la résolution de l'équation différentielle est

$$\Delta \varphi_k(t) := \varphi_{N,k}(t) - \varphi_k(t) = \int_0^t \beta_N(0) \varphi_N(u) \left[\sum_{j=0}^{N-1} e^{(t-u)(-\lambda_j)} m_j(0) p_k(\lambda_j, 0) p_{N-1}(\lambda_j, 0) \right] du$$

En effet, la formule (4.10) nous donne l'erreur commise

$$\Delta\Phi_N(t) := \begin{pmatrix} \varphi_{N,0}(t) \\ \vdots \\ \varphi_{N,N-1}(t) \end{pmatrix} - \begin{pmatrix} \varphi_0(t) \\ \vdots \\ \varphi_{N-1}(t) \end{pmatrix} = \int_0^t e^{(t-u)(-\mathcal{J}_N(0))} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \beta_N(0)\varphi_N(u) \end{pmatrix} du$$

Examinons l'erreur. Comme la matrice symétrique $\mathcal{J}_N(t)$ possède trois diagonales, elle représente les coefficients des relations de récurrence de polynômes $p_k(x, t)$ orthogonaux par rapport à la mesure $ds(x, t)$. Cette matrice symétrique réelle est diagonalisable :

$$\mathcal{J}_N(0) = P \begin{pmatrix} \lambda_0 & & 0 \\ & \ddots & \\ 0 & & \lambda_{N-1} \end{pmatrix} P^{-1}$$

où P est constituée des vecteurs propres orthonormés de la matrice $\mathcal{J}_N(0)$ et P^{-1} est en fait la transposée de P . On obtient l'égalité suivante :

$$P = \begin{pmatrix} \sqrt{m_0(0)}p_0(\lambda_0, 0) & \dots & \sqrt{m_{N-1}(0)}p_0(\lambda_{N-1}, 0) \\ \vdots & & \vdots \\ \sqrt{m_0(0)}p_{N-1}(\lambda_0, 0) & \dots & \sqrt{m_{N-1}(0)}p_{N-1}(\lambda_{N-1}, 0) \end{pmatrix}$$

On rappelle que les $m_k(t)$ et les λ_k ont déjà été définis (voir la formule (4.2) et sq.). Grâce à cette formule, on vérifie aisément que P est une matrice orthogonale. Grâce aux relations de récurrence, on montre que chacun des vecteurs est bien un vecteur propre.

Quand on passe à l'exponentielle, on obtient :

$$e^{(t-u)(-\mathcal{J}_N(0))} = P \begin{pmatrix} e^{(t-u)(-\lambda_0)} & & 0 \\ & \ddots & \\ 0 & & e^{(t-u)(-\lambda_{N-1})} \end{pmatrix} P^{-1} \quad (4.14)$$

Et en multipliant à droite par $(0, \dots, 0, 1)^t$:

$$\begin{aligned} & (e^{(t-u)(-\lambda_0)}m_0(0)p_k(\lambda_0, 0)p_{N-1}(\lambda_0, 0) + \dots \\ & \dots + e^{(t-u)(-\lambda_{N-1})}m_{N-1}(0)p_k(\lambda_{N-1}, 0)p_{N-1}(\lambda_{N-1}, 0))_{k=0, \dots, N-1} \end{aligned}$$

L'erreur commise sur le calcul de φ_k est donc bien

$$\Delta\varphi_k(t) := \varphi_{N,k}(t) - \varphi_k(t) = \int_0^t \beta_N(0)\varphi_N(u) \left[\sum_{j=0}^{N-1} e^{(t-u)(-\lambda_j)}m_j(0)p_k(\lambda_j, 0)p_{N-1}(\lambda_j, 0) \right] du \quad (4.15)$$

Grâce aux formules de Fischer et au calcul explicite de $\Delta\varphi_k(t)$, on obtient les égalités suivantes :

$$\begin{aligned} \frac{\partial\alpha_j(t)}{\partial\varphi_k}\Delta\varphi_k &= [\beta_{j+1}(t)c_k(p_j(x,t)p_{j+1}(x,t)) - \beta_j(t)c_k(p_j(x,t)p_{j-1}(x,t))] \\ &\quad \int_0^t \beta_N(0)\varphi_N(u) \left[\sum_{q=0}^{N-1} e^{(t-u)(-\lambda_q)} m_q(0)p_k(\lambda_q,0)p_{N-1}(\lambda_q,0) \right] du \\ \frac{\partial\beta_j(t)}{\partial\varphi_k}\Delta\varphi_k &= \beta_j(t)c_k(p_j^2(x,t) - p_{j-1}^2(x,t)) \\ &\quad \int_0^t \beta_N(0)\varphi_N(u) \left[\sum_{q=0}^{N-1} e^{(t-u)(-\lambda_q)} m_q(0)p_k(\lambda_q,0)p_{N-1}(\lambda_q,0) \right] du \end{aligned}$$

On développe le premier crochet dans l'expression de $\frac{\partial\alpha_j(t)}{\partial\varphi_k}\Delta\varphi_k$ et on s'occupe des quantités associées à β_{j+1} en remplaçant les c_k par la formule intégrale et on somme sur k . Les autres calculs, de même que l'estimation des autres dérivées partielles se feront de la même manière. On obtient alors

$$\begin{aligned} &\beta_{j+1}(t) \sum_{k=0}^{2j+1} \left(\int_{-1}^1 p_k(x,0)p_j(x,t)p_{j+1}(x,t)ds(x,0) \right) \\ &\quad \int_0^t \beta_N(0)\varphi_N(u) \left[\sum_{q=0}^{N-1} e^{(t-u)(-\lambda_q)} m_q(0)p_k(\lambda_q,0)p_{N-1}(\lambda_q,0) \right] du \\ &= \beta_{j+1}(t) \sum_{k=0}^{2j+1} \int_{-1}^1 ds(x,0) \int_0^t du \sum_{q=0}^{N-1} [e^{(t-u)(-\lambda_q)} m_q(0)\beta_N(0)\varphi_N(u)p_j(x,t)p_{j+1}(x,t) \\ &\quad p_k(x,0)p_k(\lambda_q,0)p_{N-1}(\lambda_q,0)] \\ &= \beta_{j+1}(t) \int_0^t du \sum_{q=0}^{N-1} e^{(t-u)(-\lambda_q)} m_q(0)\beta_N(0)\varphi_N(u) \\ &\quad \int_{-1}^1 ds(x,0)p_j(x,t)p_{j+1}(x,t)p_{N-1}(\lambda_q,0) \sum_{k=0}^{2j+1} p_k(x,0)p_k(\lambda_q,0) \\ &= \beta_{j+1}(t) \int_0^t du \sum_{q=0}^{N-1} e^{(t-u)(-\lambda_q)} m_q(0)\beta_N(0)\varphi_N(u)p_{N-1}(\lambda_q,0) \\ &\quad \int_{-1}^1 ds(x,0)K_{2j+1}(x,\lambda_q,0)p_j(x,t)p_{j+1}(x,t) \end{aligned}$$

Or $p_j p_{j+1}$ et, $p_j p_{j-1}$ dans l'autre partie de l'égalité, sont de degré inférieur ou égal à $2j + 1$. Donc, d'après la propriété 4.14 du "noyau reproduisant", on obtient

$$\beta_{j+1}(t) \int_0^t du \beta_N(0)\varphi_N(u) \sum_{q=0}^{N-1} e^{(t-u)(-\lambda_q)} m_q(0)p_{N-1}(\lambda_q,0)p_j(\lambda_q,t)p_{j+1}(\lambda_q,t)$$

Par orthogonalité, en utilisant le polynôme q_n défini lors de la démonstration de l'erreur d'interpolation (4.13), la dernière estimation est égale à

$$\beta_{j+1}(t) \int_0^t du \beta_N(0) \varphi_N(u) \sum_{q=0}^{N-1} (e^{(t-u)(-\lambda_q)} - q_{N-2j-3}(\lambda_q)) m_q(0) p_{N-1}(\lambda_q, 0) p_j(\lambda_q, t) p_{j+1}(\lambda_q, t)$$

On passe en valeur absolue, les β_k sont positifs et on utilise la majoration (4.13) pour obtenir

$$\beta_{j+1}(t) \int_0^t du |\beta_N(0) \varphi_N(u)| \frac{2R^{N-2j-2}(R^2+1) \exp((t-u)(R+1/R)/2)}{(R^{2N-4j-6}-1)(R-1)^2} \sum_{q=0}^{N-1} m_q(0) |p_{N-1}(\lambda_q, 0)| |p_j(\lambda_q, t) p_{j+1}(\lambda_q, t)|$$

On utilise ensuite le lemme 4.10, l'estimation (4.12) et certaines majorations utilisées dans la démonstration de cette estimation. Ce qui conduit à

$$\begin{aligned} & \beta_{j+1}(t) \|\mathcal{J}\| \int_0^t du \frac{40R^{-N+2} \exp(u(R+1/R)/2)}{3} \frac{40R^{-N+2j+4} \exp((t-u)(R+1/R)/2)}{3} \\ & \sum_{q=0}^{N-1} m_q(0) |p_{N-1}(\lambda_q, 0)| |p_j(\lambda_q, t) p_{j+1}(\lambda_q, t)| \\ = & \beta_{j+1}(t) \frac{1600}{9} R^{-2N+2j+6} \int_0^t \exp(t(R+1/R)/2) du \sum_{q=0}^{N-1} m_q(0) |p_{N-1}(\lambda_q, 0)| |p_j(\lambda_q, t) p_{j+1}(\lambda_q, t)| \end{aligned}$$

La fonction $\frac{R^2+1}{2R^2}$ est positive et strictement décroissante sur $[2; +\infty[$. Donc,

pour tout $R \in [2; +\infty[$, $\frac{R^2+1}{2R} \leq \frac{5}{8}R$. L'estimation devient donc

$$\beta_{j+1}(t) \frac{1600}{9} R^{-2N+2j+6} t \exp\left(\frac{5}{8}tR\right) \sum_{q=0}^{N-1} m_q(0) |p_{N-1}(\lambda_q, 0)| |p_j(\lambda_q, t) p_{j+1}(\lambda_q, t)|$$

On majore la dernière somme en utilisant l'inégalité de Cauchy-Schwarz.

$$\begin{aligned} & \sum_{q=0}^{N-1} m_q(0) |p_{N-1}(\lambda_q, 0)| |p_j(\lambda_q, t) p_{j+1}(\lambda_q, t)| \\ & \leq \sqrt{\sum_{q=0}^{N-1} m_q(0) (p_{N-1}(\lambda_q, 0))^2} \sqrt{\sum_{q=0}^{N-1} m_q(0) (p_j(\lambda_q, t) p_{j+1}(\lambda_q, t))^2} \end{aligned}$$

Choisissons $N \geq 2n$. Comme $j+j+1 \leq 2n-1$, la quadrature de Gauss est exacte. Donc la première racine carrée vaut 1 car on obtient la norme de p_{N-1} . La seconde vaut

$$\sqrt{\int p_j^2(x, t) p_{j+1}^2(x, t) ds(x, 0)}$$

De plus, $j \leq n - 1$, donc $R^{2N-2j-6} \geq R^{2N-2n-4}$. R reste un paramètre à déterminer. On minimise donc la majoration grâce au lemme 4.11, on obtient

$$\beta_{j+1}(t) \frac{1600}{9} t \left(\frac{5et}{8(2N-2n-4)} \right)^{2N-2n-4} \sqrt{\int p_j^2(x, t) p_{j+1}^2(x, t) ds(x, 0)}$$

avec $\frac{2N-2n-4}{\frac{5}{8}t} \geq 2$, cette inégalité étant vérifiée dès que $N \geq \frac{5}{4}T + 4$.

On obtient les majorations suivantes, la deuxième se faisant de la même manière :

$$\left| \sum_{k=0}^{2j+1} \frac{\partial \alpha_j(t)}{\partial \varphi_k} \Delta \varphi_k \right| \leq \frac{1600}{9} t \left(\frac{5et}{8(2N-2n-4)} \right)^{2N-2n-4} \left(\beta_{j+1}(t) \sqrt{\int p_j^2(x, t) p_{j+1}^2(x, t) ds(x, 0)} + \beta_j(t) \sqrt{\int p_j^2(x, t) p_{j-1}^2(x, t) ds(x, 0)} \right)$$

$$\left| \sum_{k=0}^{2j+1} \frac{\partial \beta_j(t)}{\partial \varphi_k} \Delta \varphi_k \right| \leq \frac{1600}{9} t \left(\frac{5et}{8(2N-2n-4)} \right)^{2N-2n-4} \beta_j(t) \sqrt{\int (p_j^2(x, t) - p_{j-1}^2(x, t))^2 ds(x, 0)}$$

On obtient donc la majoration suivante de la norme euclidienne.

$$\begin{aligned} & \sum_{j=0}^{n-1} \left(\left(\sum_{k=0}^{2j+1} \frac{\partial \alpha_j(t)}{\partial \varphi_k} \Delta \varphi_k \right)^2 + \left(\sum_{k=0}^{2j+1} \frac{\partial \beta_j(t)}{\partial \varphi_k} \Delta \varphi_k \right)^2 \right) \\ & \leq \left(\frac{1600}{9} T \left(\frac{5eT}{8(2N-2n-4)} \right)^{2N-2n-4} \right)^2 \int \sum_{j=0}^{n-1} \beta_j^2 (p_j^2(x, t) + p_{j-1}^2(x, t))^2 d\sigma(x, 0) \end{aligned}$$

On majore par parties

$$\beta_j^2 (p_j^2(x, t) + p_{j-1}^2(x, t)) \leq \|xI - \mathcal{J}\|_2^2 K_{j-1}(x, t) \leq 2^2 K_{j-1}(x, t)$$

et $p_j^2(x, t) + p_{j-1}^2(x, t) = K_j(x, t) - K_{j-2}(x, t)$.

On obtient donc une somme "télescopique" qui nous donne finalement :

$$\begin{aligned} & \sum_{j=0}^{n-1} \left(\left(\sum_{k=0}^{2j+1} \frac{\partial \alpha_j(t)}{\partial \varphi_k} \Delta \varphi_k \right)^2 + \left(\sum_{k=0}^{2j+1} \frac{\partial \beta_j(t)}{\partial \varphi_k} \Delta \varphi_k \right)^2 \right) \\ & \leq \left(\frac{3200}{9} T \left(\frac{5eT}{8(2N-2n-4)} \right)^{2N-2n-4} \right)^2 \int K_{n-1}(x, x, t) K_{n-2}(x, x, t) ds(x, 0) \end{aligned}$$

On veut estimer maintenant l'intégrale qui contient les noyaux de Szegö. La nouvelle mesure est égale à $ds(x, t) = e^{-xt} ds(x, 0)$. La fonction exponentielle est strictement positive donc on peut trouver un minorant $\rho_{\min} = e^{-t} > 0$ et

un majorant $\rho_{\max} = e^t$ de cette fonction sur le support commun des mesures. Ainsi, pour tout polynôme P ,

$$\sqrt{\rho_{\min}} \|P\|_{L_2(ds(*,0))} \leq \|P\|_{L_2(ds(*,t))} \leq \sqrt{\rho_{\max}} \|P\|_{L_2(ds(*,0))}.$$

Et comme [75, Theorem 3.1.3]

$$K_n(x, x, t) = \max_{d^\circ P \leq n} \left| \frac{P(x)}{\|P\|_{L_2(ds(*,t))}} \right|,$$

on obtient

$$\frac{K_n(x, x, 0)}{\rho_{\max}} \leq K_n(x, x, t) \leq \frac{K_n(x, x, 0)}{\rho_{\min}}.$$

Dans notre cas, $\rho_{\min}\rho_{\max} = e^{-T}e^T = 1$, ce qui donne finalement

$$K_n(x, x, 0)\rho_{\min} \leq K_n(x, x, t) \leq K_n(x, x, 0)\rho_{\max}.$$

On obtient donc un majorant de $\|E_n E_N^* \mathcal{J}_N(t) E_N E_n^* - E_n \mathcal{J}(t) E_n^*\|_2$.

De plus $\|(\alpha_0, \dots, \alpha_{n-1}, \beta_0, \dots, \beta_{n-1})\|$ peut être minoré par $\beta_0(t) = \sqrt{\int e^{-tx} ds(x, 0)}$.

Or

$$\forall (t, x) \in [0; T] \times [-1; 1], \quad e^{-tx} \geq e^{-T}.$$

Ce qui donne finalement l'estimation (4.8). \square

Montrons maintenant que l'approximation au premier ordre est bien justifiée.

Démonstration. [Formule (4.9)] D'après la formule (4.15) définissant les $\Delta\varphi_k$, on obtient $\forall k \in \{0, \dots, N-1\}$,

$$\Delta\varphi_k(t) = \int_0^t \beta_N(0)\varphi_N(u) \left[\sum_{j=0}^{N-1} e^{(t-u)(-\lambda_j)} m_j(0) p_k(\lambda_j, 0) p_{N-1}(\lambda_j, 0) \right] du$$

Appliquons le même raisonnement que dans la démonstration précédente en utilisant, à nouveau le polynôme d'interpolation q_{N-2-k}

$$\Delta\varphi_k(t) = \int_0^t \beta_N(0)\varphi_N(u) \left[\sum_{j=0}^{N-1} (e^{(t-u)(-\lambda_j)} - q_{N-2-k}) m_j(0) p_k(\lambda_j, 0) p_{N-1}(\lambda_j, 0) \right] du$$

En utilisant les mêmes arguments que précédemment, on obtient la majoration

$$\frac{1600}{9} \left(\frac{5eT}{8(2N-5-k)} \right)^{2N-5-k} T$$

D'après la formule de Stirling,

$$\left(\frac{5eT}{8(2N-5-k)} \right)^{2N-5-k} T \sim_N \frac{\sqrt{2\pi(2N-5-k)}}{(2N-9-k)!} \left(\frac{5T}{8} \right)^{2N-8-k} T$$

Ceci prouve que lorsque N tend vers l'infini, à T et n fixé, $\Delta\varphi_k$ tend vers zéro. \square

Conclusion

Voici le moment de conclure, tâche tout aussi difficile que d'introduire le sujet.

Dans le chapitre 4, nous avons donné un bel exemple d'utilisation de l'algorithme de Chebyshev modifié scalaire ainsi qu'une estimation de l'erreur commise. Dans le chapitre 2, nous avons donné une généralisation pour les matrices rectangulaires des homographies ainsi qu'un théorème d'accélération de convergence. Dans le chapitre 3, nous avons généralisé l'algorithme de Chebyshev modifié au cas vectoriel. Ces deux derniers chapitres utilisent les outils définis au premier chapitre.

Cependant, ou heureusement(!), il reste certaines parties à traiter.

Nous avons, dans le premier chapitre, démontré le théorème 1.44 sur les opérateurs de Toeplitz à diagonales constantes. On pourrait essayer de démontrer un théorème similaire pour les opérateurs à diagonales périodiques de période p . Même si une utilisation astucieuse du théorème d'Hamilton-Cayley permet d'associer l'opérateur à $r + s + 1$ diagonales périodiques à un opérateur à $p(r + s) + 1$ diagonales constantes, il n'est pas possible, en appliquant le théorème 1.44 à ce nouvel opérateur, de déterminer une caractérisation de l'ensemble résolvant de l'opérateur initial.

Dans le deuxième chapitre, nous avons généralisé l'algorithme de Chebyshev modifié scalaire au cas vectoriel. Dans le cas scalaire, nous avons démontré [9] que les polynômes de référence ainsi que les polynômes étudiés doivent posséder des mesures d'orthogonalité dont les supports doivent être essentiellement les mêmes. Pour l'instant, ce théorème n'a pu être généralisé au cas vectoriel.

Il resterait aussi à appliquer l'algorithme de Chebyshev vectoriel à un système dynamique pouvant être mis sous la forme d'un opérateur à plus de trois diagonales.

Le théorème d'accélération de convergence du chapitre 2 n'a pas, pour l'instant, été vérifié sur un exemple.

Dans le chapitre sur les systèmes dynamiques, afin d'obtenir des résultats plus précis, on pourrait modifier la démonstration de l'estimation en considérant le système infini comme un système fini soumis à une force extérieure qui évolue au cours du temps. Pour pouvoir résoudre ce système, on approche cette force extérieure par sa valeur au temps initial et on résout le système fini à second membre constant. C'est-à-dire

$$\dot{\mathcal{J}}(t) = [M(t), \mathcal{J}(t)]$$

peut être décrit par

$$\dot{\mathcal{J}}_N(t) = [M_N(t), \mathcal{J}_N(t)] + f(x_{N+1}(t), x_{N+2}(t), \dots) \approx [M_N(t), \mathcal{J}_N(t)] + f(x_{N+1}(0), x_{N+2}(0), \dots)$$

Comme on peut le remarquer, le travail restant est vaste et promet encore d'intéressantes recherches.

A

Codes Maple

On trouvera dans ce chapitre le code Maple de l'exemple 3.11, page 79, sur l'algorithme des moments modifiés vectoriel ainsi que l'exemple 3, page 100 du chapitre 4 des systèmes dynamiques. Les codes sont très commentés. Cela permet de suivre les calculs pas à pas. Dans les deux cas, on crée des fichiers de résultats. Dans le deuxième, de plus, pour créer des animations Quicktime avec Gnuplot, on crée un fichier "batch".

On peut noter que dans l'exemple 3.11, à cause des puissances choisies, on peut optimiser le calcul des moments modifiés initiaux. Le programme n'est donc pas général pour cette initialisation.

Dans le code de l'exemple 3 du chapitre 4, nous avons laissé en commentaire certaines parties du programme utilisées pour d'autres exemples de ce même chapitre.

A.1 Algorithme de Chebyshev modifié vectoriel

```
coefficient:=proc(n)
local A,a,B,alp,bet,gam;
options remember;
with(linalg);
with(orthopoly);

# A est la matrice des moments modifiés
# B celle des coefficients de recurrence de la n-ieme recurrence (fin du
# calcul)
# C est la ligne des nouveaux moments calculés. C'est-à-dire une mémoire
# tampon
# a est la borne inferieure de l'intervalle entre -1 et 0)
# Exemple de Kalyaguin et Ronveaux
# Revise 21/01/2000

Digits:=20;
a:=-1;
#alpha
alp:=-0.5;
#beta
bet:=-0.5;
#gamma
gam:=0.5;
A:=array(sparse,1..3,1..3*n+2+1);
B:=array(sparse,1..4);
C:=array(sparse,1..3*n+2+1);
T:=array(sparse,1..2,1..3*n+2+1);

# debut de calcul des moments
for i to 3*n+2+1 do

# Calcul des Tchebychev adaptés
# en première ligne, ceux sur [0,1]
# en deuxième ligne, ceux sur [a,0]

#initialisation
```

```

T[1,1]:=1;
T[1,2]:=x;

# recurrence
if (i>2 and i<(3*n+3)/2+1) then
  T[1,i]:=sort(collect(expand((2*x+1)*T[1,i-1])-T[1,i-2],x));
fi;

# calcul de la ligne des sigma indice zero
# On prend les Tchebychev vectoriels (u_0,0);(0,u_0);(u_1,0)...
# s'=2 et r'=2
# Donc si i est pair on doit multiplier par T_(i/2) et prendre la première
# densité mu_1
# Si i est impair on doit multiplier par T_((i-1)/2) et utiliser la deuxième
# densité mu_2
# attention, les indices sont décalés de 1
# on calcule les Tchebychev en meme temps pour aller plus vite
# Attention, comme on a pris les chebychev adaptés à chaque
# il faut modifier les récurrences de calcul des moments en fonction de la
# parité de l dans sigma_k,l

if ((i-1) mod 2)=0 then
# calcul des moments en utilisant la formule avec la somme et l'algorithme
# epsilon pour accélérer la convergence
# nombre d'itérations dans l'algorithme, doit etre impair !
pre:=51;
E:=array(sparse,1..pre,1..3);

# calcul du premier moment
if i=1 then
#initialisation
for j from 0 to pre-1 do
  if j=0 then
    E[1,2]:=(-1)^((i-1)/2)*(-1)^j*binomial(2*j+2,j+1)
    *binomial(2*j,j)/2^(4*j+2);
  else
    E[j+1,2]:=E[j,2]+(-1)^((i-1)/2)*(-
    1)^j*binomial(2*j+2,j+1)*binomial(2*j,j)/2^(4*j+2);
  fi;

```

```

od;
#recurrence
for k to pre-1 do
  for j to pre-k do
    E[j,3]:=E[j+1,1]+(E[j+1,2]-E[j,2])^(-1);
  od;
#print(eval(E));
for j to pre do
  if j<pre+1-k then
    E[j,1]:=E[j,2];
    E[j,2]:=E[j,3];
  else
    E[j,1]:=0;
    E[j,2]:=0;
  fi;
  E[j,3]:=0;
od;
od;
# calcul des autres moments i>1
else
for j from (i-1)/2-1 to (i-1)/2-1+(pre-1) do
  if j=(i-1)/2-1 then
    E[1,2]:=(-1)^((i-1)/2)*(-1)^j*binomial(2*j+2,j+1-(i-1)/2)
    *binomial(2*j,j)/2^(4*j+2);
  else
    E[j+1-(i-1)/2+1,2]:=E[j-(i-1)/2+1,2]+(-1)^((i-1)/2)*(-1)^j
    *binomial(2*j+2,j+1-(i-1)/2)*binomial(2*j,j)/2^(4*j+2);
  fi;
od;
#recurrence
for k to pre-1 do
  for j to pre-k do
    E[j,3]:=E[j+1,1]+(E[j+1,2]-E[j,2])^(-1);
  od;
#print(eval(E));
  if k=pre-1 then
    erreur:=abs(E[1,1]-E[2,1]);
  fi;
for j to pre do

```

```

    if j<pre+1-k then
        E[j,1]:=E[j,2];
        E[j,2]:=E[j,3];
    else
        E[j,1]:=0;
        E[j,2]:=0;
    fi;
    E[j,3]:=0;
od;
od;
fi;
# Fin de l'initialisation des moments par l'epsilon-algorithme
# On remplit la matrice A des moments mixtes
# Si i est pair, on utilise le résultat de l'epsilon algorithm
A[3,i]:=evalf(E[1,2]*Pi);
else
# Si i est impair, comme c'est la deuxième quantité calculée
# On utilise le moment précédent (particularité des puissances choisies)
    A[3,i]:=A[3,i-1]*(-1)^((i-2)/2)
fi;
# Affichage des moments à l'écran pour l'évolution
    print(3*n+2+1-i,A[3,i]);
od;

# Matrice des coefficients : 4 coefficients à chaque n
# rangés sous la forme [A_k,k-2;A_k,k-1;A_k,k;A_k,k+1]
# La 4 ième position est donc un 1 puisque les polynomes sont unitaires
# Pour les A_0,j, les deux premières positions sont nulles
# Seule la troisième équivalente à A_0,0 est à calculer et vaut 1/2*
# sigma_0,2-.5*sigma_0,0 divisé par sigma_0,0
B[4]:=1;
B[3]:=(1/2*A[3,3]-1/2*A[3,1])/A[3,1];
B[2]:=0;
B[1]:=0;
fd=fopen('resultat',WRITE);
fprintf(fd,'%a %a %a %a\n',evalf(B[1]),evalf(B[2]),evalf(B[3]),evalf(B[4]));

# Calcul de la ligne des moments d'ordre 1
# le coefficient dominant est égal à 1

```

```

# Calcul de sigma_1,1
C[2]:=(1/B[4])*(-B[3]*A[3,2]-1/2*a*A[3,2]-1/2*a*A[3,4]);
#Calcul de sigma_1,1
for l from 3 to 3*n+1 do
  # ici si l est impair c'est-à-dire que le l dans le sigma_k,l est pair
  # et réciproquement
  if (l mod 2)=1 then
    C[l]:=(1/B[4])*(-B[3]*A[3,1]+1/4*A[3,1-2]-1/2*A[3,1]+1/4*A[3,1+2]);
  else
    C[l]:=(1/B[4])*(-B[3]*A[3,1]-1/4*a*A[3,1-2]-1/2*a*A[3,1]-1/4*a*A[3,1+2]);
  fi;
od;

# Calcul des coefficients d'ordre 1
# le premier est nul
B[1]:=0;
B[2]:=1/2*C[3]/A[3,1];
B[3]:=(-1/2*a*C[2]-1/2*a*C[4]-A[3,2]*B[2])/C[2];
B[4]:=1;

# Décalage des lignes des moments pour n'avoir que 3 lignes à chaque fois
# donc on a, dans A, en première ligne, les moments d'ordre -1,
# en deuxième ligne les moments d'ordre 0 et en troisième les moments
# d'ordre 1

for i to 3*n+2+1 do
  A[1,i]:=A[2,i];
  A[2,i]:=A[3,i];
  A[3,i]:=C[i];
od;
fprintf(fd, '%a %a %a %a\n', evalf(B[1]), evalf(B[2]), evalf(B[3]), evalf(B[4]));

# ici k>=2
# il s'agit donc du cas général de calcul des sigma_k,l
# le coefficient dominant est égal à 1
# pour k=2, on a donc C[3], puis les autres
# le premier calcul est celui de sigma_2,2
# on perd deux calculs de moment en fin de matrice à chaque étape
# à cause de s'=2

```

```

for k from 2 to n do
  if (k mod 10)=0 then
    print(k);
  fi;
# nettoyage de C et de som
  C:=array(sparse,1..3*n+2+1);
  som:=0;
# calcul du moment sigma_{k+1,k+1} mais pour un k fictif =1
# calcul de la somme
  for m from k-2 to k do
    som:=som-B[m-k+2+1]*A[m+3-k,k+1];
  od;
# ici le premier calcul est celui de sigma_{2,2} pour k=2
# donc on inverse les conventions, si k est pair, c'est pareil que dans
# la réalité
  if (k mod 2)=0 then
    som:=som-1/2*A[3,k+1]+1/4*A[3,k+3];
  else
    som:=som-1/2*a*A[3,k+1]-1/4*a*A[3,k+3];
  fi;
  C[k+1]:=(1/B[4])*som;
# calcul des sigma_{k+1,l} pour l>=k+2
  for l from k+2 to 3*n+2+1-2*k do
#calcul de la somme
    som:=0;
    for m from k-2 to k do
      som:=som-B[m-k+2+1]*A[m+3-k,l];
    od;
# pour la toute première boucle et le premier calcul,
# l=4 mais c'est sigma_{2,3} que l'on calcule donc ici
# la convention est si l est pair alors on est inversé par rapport à
# la réalité d'où
    if (l mod 2)=1 then
      som:=som+1/4*A[3,l-2]-1/2*A[3,l]+1/4*A[3,l+2];
    else
      som:=som-1/4*a*A[3,l-2]-1/2*a*A[3,l]-1/4*a*A[3,l+2];
    fi;
    C[l]:=(1/B[4])*som;
  od;
od;

```

```

# calcul des coefficients d'ordre k
# le premier calcul se fait pour les moments d'ordre 2 c'est-à-dire
# effectivement la valeur de k d'où la réalité donc
if (k mod 2)=0 then
  if k=2 then
    B[1]:=1/2*C[k+1]/A[2,k-1];
  else
    B[1]:=1/4*C[k+1]/A[2,k-1];
  fi;
  if k=2 then
    B[2]:=(-1/2*a*C[k+2]-B[1]*A[2,k])/A[3,k];
  else
    B[2]:=(-1/4*a*C[k+2]-B[1]*A[2,k])/A[3,k];
  fi;
  B[3]:=(1/4*C[k+3]-1/2*C[k+1]-B[1]*A[2,k+1]-B[2]*A[3,k+1])/C[k+1];
  B[4]:=1;
else
  if k=3 then
    B[1]:=-1/2*a*C[k+1]/A[2,k-1];
  else
    B[1]:=-1/4*a*C[k+1]/A[2,k-1];
  fi;
  B[2]:=(1/4*C[k+2]-B[1]*A[2,k])/A[3,k];
  B[3]:=(-1/4*a*C[k+3]-1/2*a*C[k+1]-B[1]*A[2,k+1]-B[2]*A[3,k+1])/C[k+1];
  B[4]:=1;
fi;

# On décale
for i to 3*n+2+1 do
  A[1,i]:=A[2,i];
  A[2,i]:=A[3,i];
  A[3,i]:=C[i];
od;

# La ligne des coefficients est placée dans le fichier
fprintf(fd,'%a %a %a %a\n',evalf(B[1]),evalf(B[2]),evalf(B[3]),evalf(B[4]));
od;
fclose(fd);
end;

```

A.2 Systèmes dynamiques, algorithme de Chebyshev scalaire

```

pass8:=proc(n,N,T1,T2,bcl);
with(linalg);
with(plots);
# A et B vecteurs des coefficients de récurrence
# C moments modifiés au départ
# D vecteur temporaire
# J matrice Jacobi
# n indique le nombre de coefficients voulus
# La fréquence de coupure  $N$  doit donc être au minimum de  $2n$ 
# A représente les  $\alpha_0, \dots, \alpha_{n-1}$ 
A:=array(sparse,1..n);
# B représente les  $\beta_0, \dots, \beta_{n-1}$ 
B:=array(sparse,1..n);
# C est la matrice récursive des moments mixtes
# au départ, la ligne 1 de C contient les  $\sigma_{-1}$ 
# La ligne 2 de C contient les  $\sigma_0$ , c'est-à-dire les moments modifiés
#  $m_0, \dots, m_{2*n-1}=0$ 
# coef indique en colonne les différentes évolutions au cours du temps
# et en ligne les différentes particules (elles sont toutes nécessaires
# car on aura des sommes télescopiques pour obtenir la position de la
# dernière particule)
# coef(j,k) indique l'évolution au temps k de la particule j
coef:=array(sparse,1..n,1..bcl);
init:=array(sparse,1..n);
l:=array(sparse,1..n,1..2);
# J matrice de départ
# vecs matrice des vecteurs propres
# K matrice où les lignes de vecs auront été multipliées par la première
# ligne. Cela permettra de calculer plus vite en tout temps t.
# Il s'agit d'un prétraitement. On peut aussi multiplier les colonnes
# par les  $\exp(-\lambda_j t)$ , on pourra remultiplier par
#  $\exp(-\lambda_j (t_2-t_1))$ 
# preparation du fichier batch pour gnuplot
J:=array(symmetric,sparse,1..N+1,1..N+1);
H:=array(sparse,1..N+1,1..N+1);
K:=array(sparse,1..N+2,1..N+2);

```



```

T:=(T2-T1);
Digits:=40;
arret:=0;
# Initialisation de J
# une vitesse positive pour la première particule
J[1,1]:=1;
# Initialisation des positions
for j from 1 to N do
# surdiagonale
  J[j,j+1]:=1/2;
od;
print(eval(J));
# Calcul des valeurs propres et vecteurs propres relatifs à J
vecs:='vecs';
lambda:=evalf(Eigenvals(J,vecs));
# Préparation de la matrice K au temps 0
# K[i,j] indique  $m_j * p_i(\lambda_j, 0) * p_0(\lambda_j, 0)$ 
# H[i,j] indique  $m_j(0) * p_i(\lambda_j, 0) * p_0(\lambda_j, 0)$ 
# première ligne p_0, deuxième ligne p_1... dénoté par i
# première colonne m_0 et lambda_0, deuxième colonne m_1 et lambda_1
# dénotés par j
for i to N+1 do
  for j to N+1 do
    H[i,j]:=vecs[1,j]*vecs[i,j];
  od;
od;
# Configuration de départ
for j from 2 to n do
  init[j]:=init[j-1]-2*ln(J[j-1,j]);
od;
#print(eval(vecs));print(lambda);
# démarrage de la boucle de calcul en temps indiqué par bcl
# h est le pas de temps
h:=evalf(T/bcl);
Digits:=80;
for i to bcl do
print(i);
# C indique les 3 lignes consécutives des moments mixtes
C:='C';

```

```

C:=array(sparse,1..3,1..2*n+1);
# On commence par multiplier chaque colonne j de K par exp(-lambda_j*h)
# Boucle sur les lignes
# Si N>2*n
if arret=0 then
if N>=2*n then
  for j to 2*n do
# Boucle sur les colonnes successives
  for k to N+1 do
    K[j,k]:=H[j,k]*exp(-lambda[k]*(i*h+T1));
# Création des moments modifiés
    if k=1 then
      C[2,j]:=K[j,k]
    else
      C[2,j]:=C[2,j]+K[j,k]
    fi;
  od;
od;
else
# N<2*n, donc il manque des moments modifiés mais  $\phi_N=0$  donc on va
# s'en sortir. Pour l'instant, on récupère  $m_0, \dots, m_{N-1}$ .
  for j to N+1 do
# Boucle sur les colonnes successives
  for k to N+1 do
    K[j,k]:=H[j,k]*exp(-lambda[k]*(i*h+T1));
# Création des moments modifiés
    if k=1 then
      C[2,j]:=K[j,k]
    else
      C[2,j]:=C[2,j]+K[j,k]
    fi;
  od;
od;
  C[2,N+1]:=0;
fi;
# Ainsi C contient les moments modifiés initiaux sur les deux premières
# lignes, la troisième est complétée au fur et à mesure
#print(evalm(C));

```

```

# Algorithme des Moments modifiés

# Calcul des coefficients de la première ligne
nomcoeff:=cat('coeff',i,'.dat');
fc:=fopen(nomcoeff,WRITE);
fprintf(fc,'%a %a %a %a\n',n,T1,T2,bc1);
#Calcul de alpha_0
A[1]:=evalf(J[1,2]*C[2,2]/C[2,1]+J[1,1]);
# Calcul de beta_0
B[1]:=evalf(C[2,1]);
#B[1]:=sqrt(Pi);
#B[1]:=1/sqrt(2);
# itérations
# ligne de la récurrence
  for j from 2 to n do
# Calcul des moments mixtes
# Nous sommes déjà à la deuxième ligne donc il nous en reste 2n-2
  for k from j to 2*n-j+1 do
# Les récurrences de référence influent sur les calculs des moments modifiés
# On a donc laissé en commentaire certains calculs effectués pour d'autres
# exemples
#   if k=2 then
#     C[3,k]:=evalf(1/2*C[2,k+1]-(A[j-1]-J[k,k])*C[2,k]-B[j-1]*C[1,k]
#       +sqrt(1/2)*C[2,k-1]);
#   else
#     C[3,k]:=evalf(1/2*(C[2,k+1]+C[2,k-1])-
#       (A[j-1]-J[k,k])*C[2,k]-B[j-1]*C[1,k]);
#   fi;
  if k<N then
    C[3,k]:=evalf(J[k,k+1]*C[2,k+1]-(A[j-1]-J[k,k])*C[2,k]-B[j-1]*C[1,k]...
      ...+J[k-1,k]*C[2,k-1]);
  fi;
  od;
#   if j=2 then
#     A[j]:=evalf(J[j,j]+1/2*C[3,j+1]/C[3,j]-sqrt(1/2)*C[2,j]/C[2,j-1]);
#     B[j]:=evalf(sqrt(1/2)*C[3,j]/C[2,j-1]);
#   else
#     A[j]:=evalf(J[j,j]+1/2*(C[3,j+1]/C[3,j]-C[2,j]/C[2,j-1]));
#     B[j]:=1/2*C[3,j]/C[2,j-1];

```

```

# fi;(
  A[j]:=evalf(J[j,j]+J[j,j+1]*C[3,j+1]/C[3,j]-J[j-1,j]*C[2,j]/C[2,j-1]);
  B[j]:=evalf(J[j-1,j]*C[3,j]/C[2,j-1]);
  if B[j]<0 then
    print(j);
    print(B[j]);
    arret:=1;
    fi;
  for k from 1 to 2*n do
    if k>j-1 then
      C[1,k]:=C[2,k];
      C[2,k]:=C[3,k];
    fi;
  od;
od;
for j to n do
  if B[j]<0 then
    arret:=1
  else
    B[j]:=sqrt(B[j]);
  fprintf(fc,'%a %a \n',evalf(A[j]),evalf(B[j]));
  fi;
od;
fclose(fc);
if arret=0 then
# La première particule est déterminée par la vitesse au cours du temps
  if i=1 then
    coef[1,i]:=A[1]*h;
    minx:=coef[1,1];
  else
    coef[1,i]:=coef[1,i-1]+A[1]*h;
    minx:=min(minx,coef[1,i]);
  fi;
# Les suivantes par recurrence
  for k from 2 to n do
    coef[k,i]:=-2*ln(B[k])+coef[k-1,i];
  od;
if i>1 then
  maxx:=max(maxx,coef[n,i]);

```

```
else
  maxx:=coef[n,i];
fi;
  k:='k';
fi;
fi;
od;
#print("ouverture de batch");
fb:=fopen('batch',WRITE);
#print("terminal");
fprintf(fb,'%s "%s" %s \n',"set terminal moov landscape color dashes"...
...,Geneva,"12 600 200 5");
fprintf(fb,'%s \n',"set noborder");
fprintf(fb,'%s \n',"set noxtics");
fprintf(fb,'%s \n',"set noytics");
#print("début des boucles");
for i from 0 to bcl do
  nom:=cat('resultat',i,'.dat');
  fprintf(fb,'%s%f:%f%s %s' %s \n',"plot [",minx-1,maxx+1,"]",nom,notitle);
  fd:=fopen(nom,WRITE);
  for k to n do
    if i<>0 then
      #print(k);
      fprintf(fd,'%a %a \n',evalf(coef[k,i]),0);
    else
      fprintf(fd,'%a %a \n',evalf(init[k]),0);
    fi;
  od;
  fclose(fd);
od;
fclose(fb);
end;
```

Bibliographie

- [1] N. I. Akhiezer, I. M. Glazman, *Theory of linear operators in a Hilbert space* volumes I, II, Pitman, Boston, 1981.
- [2] A. I. Aptekarev, Multiple orthogonal polynomials, *J. Comput. Appl. Math.* 99 (1998), 423-447.
- [3] A. I. Aptekarev, E. M. Nikishin, The scattering problem for a discrete Sturm-Liouville operator, *Math. Sbornik* Vol 121(163)(1983), No. 3 (in Russian), English translation *Math. USSR Sbornik* Vol 49(1984), No. 2.
- [4] G. A. Baker, P. Graves-Morris, *Padé Approximants*, deuxième édition, Cambridge University Press, Encyclopaedia of mathematics and its applications, 1996.
- [5] B. Beckermann, Nonsymmetric difference operators and polynomials being orthogonal with respect to rectangular matrix valued measures, *publication ANO 335*, Lille 1995.
- [6] B. Beckermann, On the numerical condition of polynomial bases : Estimates for the condition number of Vandermonde, Krylov and Hankel matrices, Thèse d'Habilitation, Universität Hannover, 1995.
- [7] B. Beckermann, On the classification of the spectrum of second order difference operators, *publication ANO 379*, Lille 1997.
- [8] B. Beckermann, Complex Jacobi matrices, *J. Comput. Appl. Math.* 127 (2001), 17-65.
- [9] B. Beckermann, E. Bourreau, How to choose modified moments?, *J. Comput. Appl. Math.* 98 (1998) 81-98.
- [10] B. Beckerman, V. Kaliaguine, The diagonal of the Padé table and the approximation of the Weyl function of second order difference operators, *Constr. Approx.* 13 (1997) 481-510.

Les publications ANO sont accessibles par téléchargement sur ano.univ-lille1.fr

- [11] B. Beckermann, G. Labahn, A uniform approach for Hermite-Padé and simultaneous Padé approximations and their matrix-type generalisations, *Numer. Algo.* 3 (1992) 45-54.
- [12] B. Beckermann, G. Labahn, A uniform approach for the fast computation of matrix-type Padé approximants, *J. Matrix Anal. Appl.* Vol. 15, No. 3, (1994) 804-823.
- [13] B. Beckermann, A. Osipov, Some spectral properties of infinite band matrices, *publication ANO 434*, Lille 2002.
- [14] Yu. M. Berezanskii, Integration of nonlinear difference equation by the inverse spectral problem method, *Soviet. Mathem. Doklady*, 31 (1985), No. 2, p. 264-267.
- [15] O. I. Bogoyavlenskii, Some constructions of integrable dynamical systems, *Engl. Transl. in Math. USSR Izvestiya*, Vol. 31(1988), No. 1, 47-75.
- [16] O. I. Bogoyavlenskii, Integrable dynamical systems associated with the KdV equation, *Engl. Transl. in Math. USSR Izvestiya*, Vol. 31(1988), No. 3, 435-454.
- [17] O. I. Bogoyavlenskii, The Lax representation with a spectral parameter for certain dynamical systems, *Engl. Transl. in Math. USSR Izvestiya*, Vol. 32(1989), No. 2, 245-268.
- [18] E. Bourreau, Modified moments and matrix orthogonal polynomials, *Acta Applicandae Mathematicae*, 61(2000), 53-64.
- [19] C. Brezinski, *Biorthogonality and its applications to numerical analysis.*, Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker.
- [20] C. Brezinski, A direct proof of the Christoffel-Darboux identity and its equivalence to the recurrence relationship, *J. Comput. Appl. Math.* 32 (1990), 17-25.
- [21] C. Brezinski, difference and differential equations and convergence acceleration algorithms, *publication ANO 419*, Lille 2001.
- [22] C. Brezinski, Convergence acceleration during the 20th century, *publication ANO 421*, Lille 2001.
- [23] M. G. de Bruin, Convergence of generalized C-fractions, *J. Approximation Theory* 24(1978), 177-207.

- [24] M. G. de Bruin, L. Jacobsen, The dominance concept of linear recurrence relations with applications to continued fractions, *Nieuw Archief v. Wiskunde*, No 4 Vol 3(1985), 253-266.
- [25] M. G. de Bruin, L. Jacobsen, Modification of generalized continued fractions. I, *Lecture Notes in Mathematics*, vol 1237 (J. Gilewicz, M. Pindor, W. Siemaszko, Eds.), Springer-Verlag, Berlin, 1987, 161-176.
- [26] K. M. Case, M. Kac, A discrete version of the inverse scattering problem, *J. Math. Phys.*, 14(1973) 594-603.
- [27] J.-P. Demailly, *Analyse numérique et équations différentielles*, Collection Grenoble Sciences, Presses universitaires de Grenoble, 1991.
- [28] Z. Da Rocha, Shohat-Favard and Chebyshev's methods in d -orthogonality, *Numer. Algo.* 20 (1999) 139-164.
- [29] P. L. Duren, On the spectrum of a Toeplitz operator, *Pac. J. Math.*, 14(1964), 21-29.
- [30] H.-J. Fischer, On the condition of orthogonal polynomials via modified moments, *Z. Anal. Anwendungen* 15 (1) (1996) 1-18.
- [31] H. Flashka, The Toda lattice II, *Progr. Theoret. Phys.*, 51(1974) 703-716.
- [32] W. Gautschi, Construction of Gauss-Christoffel quadrature formulas, *Math. Comput.*, 22 (1968) 251-270.
- [33] W. Gautschi, On the construction of Gaussian quadrature rules from modified moments, *Math. Comput.*, 24 (1970) 245-260.
- [34] W. Gautschi, On generating orthogonal polynomials, *Siam J. Sci. Statist. Comput.*, 3(1982) 289-317.
- [35] W. Gautschi, Orthogonal polynomials - Constructive theory and applications, *J. Comput. Appl. Math.* 12&13 (1985) 61-76.
- [36] W. Gautschi, On the sensitivity of orthogonal polynomials to perturbations in the moments, *Numer. Math.* 48 (1986) 369-382.
- [37] M. I. Gekhtman, Integration of non-abelian Toda chains, *Engl. Transl. Func. Anal. Appl.* 24(1990) 3, 231-233.
- [38] M. I. Gekhtman, Solution of infinite Toda chains, *Engl. Transl. Func. Anal. Appl.* 25(1991) 3, 230-232.
- [39] B. Germain-Bonne, Estimation de la limite de suite et formalisation de procédés d'accélération de convergence, Thèse de doctorat, 1978, Université de Lille I.

- [40] J. S. Geronimo, W. Van Assche, Approximating the weight function for orthogonal polynomials on several intervals, *J. of App. Th.* 65 (1991) 341-371.
- [41] F. Gesztesy, H. Holden, B. Simon, Z. Zhao, On the Toda and Kac-van Moerbeke systems, *Transaction of the American Mathematical Society*, v. 339(1993), 2, 849-868.
- [42] P. Hartman, A. Wintner, On the spectra of Toeplitz matrices, *Am. J. Math.* 72(1950), 359-366.
- [43] F. Hirsch, G. Lacombe, *Eléments d'Analyse Fonctionnelle*, Dunod, Sciences Sup, 1999.
- [44] L. Jacobsen, Uniform convergence of limit periodic generalized continued fractions, *Analysis* 8(1988), 95-119.
- [45] L. Jacobsen, W. B. Jones, H. Waadeland, Convergence acceleration for continued fractions $K(a_n/1)$, where $a_n \rightarrow \infty$, *Rational approximation and its applications in mathematics and physics*, Lecture Notes in Mathematics, vol. 1237, 1987, 177-187.
- [46] M. Kac, P. Van Moerbeke, On an explicitly soluble system of nonlinear differential equations related to certain Toda lattices, *Adv. in Math.*, 16, 160-169(1975).
- [47] V. Kaliaguine, The operator moment problem, vector continued fractions and an explicit form of the Favard theorem for vector orthogonal polynomials, *J. Comput. Appl. Math.* 65 (1995) 181-193.
- [48] V. Kaliaguine, Hermite-Padé Approximants and spectral analysis of nonsymmetric operators, *Russian Acad. Sci. Sb. Math.* Vol 82(1995), No. I.
- [49] V. Kaliaguine, A. Ronveaux, On a system of "classical" polynomials of simultaneous orthogonality, *J. Comput. Appl. Math.* 67 (1996) 207-217.
- [50] P. Levrie, L. Jacobsen, Convergence acceleration for generalized continued fractions, *Transactions of the American Mathematical Society*, Volume 305, Number 1, Jan. 1988.
- [51] P. Levrie, Convergence acceleration for n -fractions, *Applied Numerical Mathematics* 7 (1991) 481-492.
- [52] P. Levrie, A. Bultheel, A note on Thiele n -fractions, *Numer. Algo.* 4 (1993) 225-239.
- [53] P. Levrie, Some identities for G -continued fractions and generalized continued fractions, *J. Comput. Appl. Math.* 51 (1994) 85-97.

- [54] K. Mahler, Perfect Systems, *Compos. Math.*, 19(1968), 95-166.
- [55] A. Máté, P. Nevai, A generalization of Poincaré's theorem for recurrence equations, *J. Approximation Theory* 63(1990), 92-97.
- [56] J. M. Monier, *Nouveau cours de Mathématiques, Algèbre 2* (Collection J'intègre, Dunod).
- [57] J. Moser, Three integrable Hamiltonian systems connected with isospectral deformations, *Adv. in Math.*, 16, 197-220(1975).
- [58] E. M. Nikishin, V. N. Sorokin, *Rational Approximations and Orthogonality*, Translations of Mathematical Monographs 92, Amer. Math. Soc., Providence, RI, 1991.
- [59] A. S. Osipov, Discrete Analog of the Korteweg-de Vries(KdV) equation : Integration by the method of the inverse problem, *Mathematical Notes*, Vol. 56, Nos. 5-6, 1994, 1312-1314.
- [60] V. I. Parusnikov, The Jacobi-Perron algorithm and simultaneous approximation of functions, *Mat. Sb.* 114(156) (1981), no. 2, 322-333; English transl. in *Math. USSR Sb.* 42 (1982), 287-296.
- [61] A. M. Perelomov, *Integrable systems of classical mechanics and Lie algebras*, Vol. I, Birkhäuser.
- [62] L. R. Pineiro, On simultaneous approximations for a collection of Markov functions, *Vestnik Mosk. Univ., Ser. I* (1987), no. 2, 67-70 (in Russian); *Moscow Univ. Math. Bull.* 42 (2) (1987), 52-55.
- [63] W. H. Press, B. P. Flannery, S. Teukolsky, W. T. Vetterling, *Numerical Recipes, The art of Scientific Computing*, Cambridge University Press, Cambridge, 1986.
- [64] A. Quarteroni, R. Sacco, F. Saleri, *Numerical Mathematics*, Texts in Applied Mathematics, Vol. 37, Springer-Verlag.
- [65] W. Rudin, *Analyse réelle et complexe*, Masson.
- [66] W. Rudin, *Analyse fonctionnelle*, Ediscience.
- [67] W. Rudin, *Principes d'analyse mathématique*, Masson.
- [68] R.A. Sack, A.F. Donovan, An algorithm for Gaussian quadrature given modified moments, *Numer. Math.* 18 (1971/72) 465-478.
- [69] A. Sinap, Gaussian quadrature for matrix valued functions on the real line, *J. Comput. Appl. Math.* 65 (1995) 369-385.
- [70] A. Sinap, W. Van Assche, Orthogonal matrix polynomials and applications, *J. Comput. Appl. Math.* 66 (1996) 27-52.

- [71] V.N. Sorokin, J. Van Iseghem, Algebraic Aspects of Matrix Orthogonality for Vector Polynomials, *J. of App. Th.* 90 (1997) 97-116.
- [72] V.N. Sorokin, J. Van Iseghem, Matrix Continued Fractions, *J. of App. Th.* 96 (1999) 237-257.
- [73] V.N. Sorokin, J. Van Iseghem, Matrix Hermite-Padé problem and dynamical systems, Publication ANO 397, Laboratoire d'Analyse Numérique, Lille I, Juin 1999.
- [74] H. Stahl, V. Totik, *General Orthogonal polynomials*, Cambridge University Press, New York, 1992.
- [75] G. Szegö, *Orthogonal Polynomials*, AMS, Providence, 1975
- [76] G. Teschl, The Toda Lattice, <http://www.mat.univie.ac.at/~gerald/ftp/book-jac/toda.html> .
- [77] W. Van Assche, *Asymptotics for orthogonal polynomials*, Lecture Notes in Mathematics, Vol. 1265, Springer-Verlag, Berlin.
- [78] W. Van Assche, Christoffel functions and Turan determinants on several intervals, *J. Comput. Appl. Math.* 48 (1993) 207-223.
- [79] W. Van Assche, Nonsymmetric linear difference equations for multiple orthogonal polynomials, Proceedings of SIDEIII, Sabaudia, Italy 1998.
- [80] W. Van Assche, E. Coussement, Some classical multiple polynomials, <http://www.wis.kuleuven.ac.be/wis/applied/walter.html>, fichier : clasmult.ps.
- [81] J. Van Iseghem, Matrix continued fraction for the resolvent function of the band operator, *Acta Applicandae Mathematicae*, 61(2000), 351-365.