N° d'ordre : H 342

Université des Sciences et Technologies de Lille 1

Habilitation à diriger des recherches

spécialité **Mathématiques**

APPROXIMATION RATIONNELLE DANS L'ESPACE DE HARDY DU DISQUE, INTERPOLATION RATIONNELLE DE LA FONCTION EXPONENTIELLE ET APPROXIMATION DE HERMITE–PADÉ

Franck WIELONSKY [1]

Soutenue en Octobre 2002 devant le jury

Jean–Bernard BAILLON	Professeur à l'université de Paris 1	
Laurent BARATCHART	Dir. de Recherche à l'INRIA Sophia Antipolis	
Claude BREZINSKI	Professeur à l'université de Lille	
Michel CROUZEIX	Professeur à l'université de Rennes	
Wieslaw PLESNIAK	Professeur à l'université de Cracovie	rapporteur
Vilmos TOTIK	Professeur à l'université de Szeged	rapporteur
Alain YGER	Professeur à l'université de Bordeaux	rapporteur

1. Laboratoire de Mathématiques appliquées, FRE CNRS 2222, UST Lille 1, Math, F-59655 Villeneuve d'Ascq CEDEX, France, e-mail: wielonsk@ano.univ-lille1.fr

Table des matières

1 Introduction **5**

2 Approximation rationnelle dans l'espace de Hardy H^2 du disque et unicité d'un meilleur approximant **11**

- 2.1 Motivations provenant de la théorie des systèmes . . . 12
  - 2.1.1 Systèmes discrets et identification stochastique . . . 13
  - 2.1.2 systèmes continus et identification harmonique . . . 14
- 2.2 Premières propriétés . . . 15
- 2.3 Le problème de l'unicité d'un meilleur approximant . . . 21
  - 2.3.1 Approximation de type fixé dans des disques décroissants vers l'origine . . . 23
  - 2.3.2 Un théorème du type Montessus de Ballore dans l'espace de Hardy $H_{2,\mathbb{R}}$ . . . 24
  - 2.3.3 Approximation diagonale de la fonction exponentielle . . . 25
- 2.4 Le problème de l'unicité pour la classe des fonctions de Markov . . . 26
- 2.5 Estimées asymptotiques pour les approximants rationnels L^2 de fonctions de Markov 31

3 Interpolation rationnelle de la fonction exponentielle et zéros des polynômes exponentiels réels **35**

- 3.1 Points d'interpolation réels . . . 36
- 3.2 Points d'interpolation complexes . . . 39

4 Asymptotique des approximants de Hermite-Padé et applications **43**

- 4.1 Asymptotique des approximants de Hermite-Padé . . . 43
- 4.2 Une application à la théorie des nombres . . . 45
- 4.3 Distribution des zéros et asymptotique normalisée . . . 47

Nous présentons dans ce rapport les résultats publiés dans les articles suivants. Les articles [1*, 2*, 3*, 5*, 8*, 10*, 11*] font l'objet du Chapitre 2. Les articles [4*, 12*, 13*] et [6*, 7*, 9*, 14*] sont décrits dans les chapitres 3 et 4 respectivement.

[1*] L. Baratchart, M. Olivi, F. Wielonsky, *Asymptotic properties in rational* l^2*-approximation*, Proc. of the 9th Conf. on Analysis and Optimization of Systems, Antibes, lect. notes in control and information sciences 144, J.L. Lions and A. Bensoussan Eds, (1990), 477-486.

[2*] L. Baratchart, M. Olivi, F. Wielonsky, *On a rational approximation problem in the real Hardy space* H_2, Theoretical Computer Science 94 (1992), 175-197.

[3*] L. Baratchart, F. Wielonsky, *Rational approximation problem in the real Hardy space* H_2 *and Stieltjes integrals: a uniqueness theorem*, Constr. Approx. 9 (1993), 1-21.

[4*] L. Baratchart, E.B. Saff, F. Wielonsky, *Rational interpolation of the exponential function*, Canad. J. of Math. 47 (6), (1995), 1121-1147.

[5*] L. Baratchart, E.B. Saff, F. Wielonsky, *A criterion for uniqueness of a critical point in* H_2 *rational approximation*, J. d'Analyse Math. 70 (1996) , 225-266.

[6*] F. Wielonsky, *Asymptotics of diagonal Hermite-Padé approximants to* e^z, J. Approx. Theory 90 (1997), 283-298.

[7*] F. Wielonsky, *Hermite-Padé approximants to exponential functions and an inequality of Mahler*, J. Number Th. 74 (1999), 230-249.

[8*] L. Baratchart, H. Stahl, F. Wielonsky, *Non-uniqueness of rational best approximants*, J. Comp. and Appl. Math. 105 (1999), 141-154.

[9*] F. Wielonsky, *Some properties of Hermite-Padé approximants to* e^z, Continued Fractions: From Analytic Number Theory to Constructive Approximation , Contemporary Mathematics 236 (1999), 369-379.

[10*] L. Baratchart, H. Stahl, F. Wielonsky, *Asymptotic uniqueness of best rational approximants of given degree to Markov functions in* L^2 *of the circle*, Constr. Approx. 17 (2001), 103-138.

[11*] L. Baratchart, H. Stahl, F. Wielonsky, *Asymptotic error estimates for* L^2 *rational best approximants to Markov functions*, J. Approx. Theory 108 (2001), 53-96.

[12*] F. Wielonsky, *A Rolle's theorem for real exponential polynomials in the complex domain*, J. Math. Pures Appl. 80 (2001), 389-408.

[13*] F. Wielonsky, *On rational approximation to the exponential function with complex conjugate interpolation points*, J. Approx. Theory 111 (2001), 344-368.

[14*] A.B. Kuijlaars, W. Van Assche, and F. Wielonsky, *Quadratic Hermite-Padé approximation to the exponential function: a Riemann-Hilbert approach*, manuscrit, 2002.

Chapitre 1

Introduction

Ce manuscrit est constitué de trois parties, traitant chacune de différents thèmes en approximation de fonctions. L'approximation des fonctions peut se voir comme la branche de l'analyse mathématique qui consiste à approcher au sens de diverses normes des fonctions individuelles ou appartenant à certaines classes données, par des fonctions ne dépendant souvent que d'un nombre fini de paramètres. On peut par exemple approcher des éléments pris dans la classe des fonctions continues, lisses ou analytiques, ou des fonctions individuelles comme la fonction exponentielle ou possédant des singularités comme la fonction valeur absolue ou racine carrée en l'origine. Les approximants peuvent être choisis dans la famille des polynômes, des sommes d'exponentielles, des fractions rationnelles, des fonctions méromorphes, des splines, ..., au sens de diverses normes (par exemple L^p, $1 \leq p \leq \infty$) ou par des propriétés d'interpolation. Les problèmes que l'on se pose alors sont usuellement des questions d'existence, de caractérisation, d'unicité des meilleurs approximants, de normalité (à savoir si l'on utilise la totalité des paramètres dont on dispose pour obtenir les meilleurs approximants). On peut également chercher à estimer la vitesse de convergence des approximants, les propriétés d'interpolation de ces derniers ou de manière symétrique, les propriétés d'approximation des interpolants.

Le lien entre approximants et polynômes orthogonaux est classique, par exemple lorsqu'on considère la norme L^2 ou des interpolants rationnels, en particulier les fractions continues. Ces deux thèmes ont été intensément étudiés grâce aux outils fournis par la théorie du potentiel logarithmique qui s'applique directement à l'étude des polynômes dans le plan complexe. Rappelons ici simplement l'observation élémentaire que pour tout polynôme $P_n(z) = \prod_{j=1}^n (z - z_j)$, la quantité

$$\frac{1}{n} \log \frac{1}{|P_n(z)|} = \frac{1}{n} \sum_j \log \frac{1}{|z - z_j|} = \int \log \frac{1}{|z - t|} d\delta(t)$$

est le potentiel logarithmique associé à la mesure de probabilité discrète δ, somme des mesures de Dirac de masse $1/n$ en chaque zéro de P_n. De plus, comme il est bien connu, un polynôme orthogonal minimise une norme L^2, c'est à dire que lorsque son degré tend vers l'infini et sous certaines hypothèses de régularité sur la mesure, le potentiel associé est solution d'un problème extrémal correspondant à la minimisation d'une énergie. Ainsi, il existe une analogie entre le comportement des zéros d'un polynôme orthogonal pour une mesure ν et celui à l'équilibre de charges électriques situées sur le support de ν. De même, les zéros et les pôles d'une fonction rationnelle extrémale, comme par exemple la solution du problème de minimisation de Zolotarev, peuvent s'interpréter comme des charges positives et négatives disposées respectivement sur les deux plateaux d'un condensateur.

Une technique différente, faisant un lien avec l'étude des problèmes inverses, s'est également développée dans les années 90. Elle consiste à caractériser les polynômes orthogonaux comme solutions d'un problème de Riemann-Hilbert, puis à effectuer une série de transformations qui réduisent ce problème à un problème élémentaire que l'on sait résoudre explicitement. Cette technique a permis de raffiner très précisément différentes estimées sur certains polynômes orthogonaux P_n, par exemple sur l'axe réel, en obtenant des asymptotiques de Plancherel-Rotach, dans des voisinages du plus petit et du plus grand zéros de P_n.

Parmi les problèmes classiques qui font intervenir des approximants qui sont des fonctions analytiques dans des espaces de Hardy, on peut citer le problème d'extension de Nehari dont le problème de Carathéodory-Fejer et le problème d'interpolation de Nevanlinna-Pick sont des cas particuliers. Il existe une généralisation de cette théorie pour des fonctions méromorphes avec un nombre prescrit de pôles. Elle a été établie par Adamjan, Arov et Krein (théorie AAK) dans les années 70. Pour les meilleurs approximants méromorphes, il existe une caractérisation, via la théorie des opérateurs, en particulier les opérateurs de Hankel dans H_2. Les meilleurs approximants rationnels, plus contraints que les approximants méromorphes, n'admettent pas de caractérisation analogue. On décrit dans ce rapport des résultats obtenus pour trois problèmes, tous liés à l'approximation rationnelle :

– approximation rationnelle en norme L^2 dans le disque unité,

– interpolation rationnelle de la fonction exponentielle,

– approximation de Hermite–Padé de la fonction exponentielle.

Le problème d'approximation abordé dans le Chapitre 2 consiste à minimiser, pour une fonction g dans l'espace de Hardy $H_2(U)$ du disque unité U, l'intégrale

$$\frac{1}{2i\pi}\int_0^{2\pi}|g(e^{it})-r(e^{it})|^2dt,$$

où r décrit l'ensemble des fonctions rationnelles d'un degré donné, analytiques dans l'adhérence $\overline{U}$ de U. Ce problème se comporte, vis à vis de nombreux aspects, de manière bien différente de problèmes d'approximation "linéaires", par exemple de l'approximation de Padé dont la solution, unique, fournie par la résolution d'un système d'équations linéaires, est d'ordinaire plus facile à calculer. De plus, le meilleur approximant rationnel L^2 n'est pas forcément unique et, en général, des minima locaux peuvent apparaître. Cependant, malgré cette complexité accrue, les approximants L^2 se révèlent particulièrement bien adaptés lorsque la fonction g est déterminée naturellement sur le cercle unité plutôt que par les valeurs de ses coefficients de Taylor en l'origine. Du point de vue de la théorie du potentiel, on peut d'ailleurs noter que le cercle est un ensemble de capacité positive, contrairement au point à l'origine qui est un ensemble polaire (de capacité nulle).
Le critère de minimisation étant un critère intégral, on peut le dériver en fonction des variables paramétrant la fonction r, par exemple ses coefficients. On peut prouver que ce critère est lisse et donc appliquer les outils différentiels pour en faire son étude. Ainsi, on peut parler des points critiques du critère, i.e. les fonctions rationnelles en lesquelles toutes les dérivées partielles de ce critère s'annulent.
La motivation originelle pour l'étude de ce problème d'approximation au sein du projet MIAOU à l'INRIA Sophia–Antipolis provient de la théorie des systèmes linéaires. D'abord, ces systèmes peuvent être décrits, dans le domaine fréquentiel (i.e. lorsqu'on considère la transformation de

Laplace de l'équation régissant les entrées-sorties du système) comme un opérateur de multiplication par la fonction dite de transfert. Cette fonction appartient à un certain espace de Hardy du demi–plan droit, suivant les propriétés de stabilité du système. Ensuite, les fonctions rationnelles de degré n sont précisément les fonctions de transfert des systèmes de dimension finie, i.e. ceux régis par une équation différentielle linéaire d'ordre n. Notre problème s'interprète donc comme l'approximation d'un système S_∞ de dimension infinie par un système S_n de dimension finie de telle sorte que l'énergie du système $S_\infty - S_n$ (au sens des fonctions de transfert) soit minimisée. Un autre domaine d'applications pour lequel l'approximation rationnelle L^2 se révèle être un outil très intéressant est celui des problèmes inverses, tels ceux qui interviennent en théorie du contrôle non destructif. On peut trouver dans [41] l'étude d'une approche, présentée également dans [10], consistant à reformuler en terme du problème d'approximation L^2, celui de la détection d'une fissure à l'intérieur d'une structure plane, à l'aide de données sur le bord.

D'un point de vue numérique, un algorithme de recherche de minima locaux par intégration du gradient a été implémenté par le projet MIAOU dans le logiciel *Hyperion*, voir [28]. Pour un degré donné, le comportement de cet algorithme dépend grandement de la configuration des minima locaux. En particulier, plus le nombre de ces minima est réduit et plus les chances de trouver le minimum absolu en partant d'un certain nombre de conditions initiales est grand. Bien sûr, la configuration optimale est obtenue lorsqu'il existe un unique minimum local, à savoir le meilleur approximant lui-même. La question se pose donc de savoir pour quelles classes de fonctions notre problème d'approximation, à degré fixé, admet un unique minimum local. Pour étudier cette question, abordée dans le paragraphe 2.3, on utilise une approche, voir [6], issue de la topologie différentielle, à savoir un théorème de l'indice qui peut se voir comme un analogue du théorème de Poincaré–Hopf. Ce théorème fournit une contrainte, indépendante de la fonction g approchée, sur l'ensemble des points critiques à un degré donné n. Il permet de traduire le problème global de l'unicité d'un minimum local en une question locale, à savoir : chaque point critique est–il un minimum local ? La réponse demande une analyse précise de la matrice hessienne en ces points critiques. Cette analyse permet d'obtenir une condition de positivité pour cette matrice : le numérateur p et le dénominateur q du point critique considéré doivent rester suffisamment premiers entre eux relativement à l'erreur $\|g - p/q\|_\infty$. Finalement, cette condition se trouve être vérifiée si l'on peut construire un interpolant rationnel de g de degré $n-1$ qui soit suffisamment *mauvais partout sur le cercle unité* relativement à l'erreur du point critique en degré n. Cela demande une connaissance très précise des erreurs commises par tous ces approximants. Lorsqu'on s'intéresse à l'unicité asymptotique, i.e. lorsque le degré n devient grand, les estimées en racine n-ième, celles qui sont le plus couramment fournies par la théorie du potentiel, sont insuffisantes. Des asymptotiques fortes, dites de Szegő, sont donc nécessaires. Dans le cas de fonctions de Markov et de la fonction exponentielle, de telles asymptotiques peuvent être obtenues. Elles permettent de déduire pour ces fonctions l'unicité asymptotique d'un minimum local pour notre problème d'approximation rationnel dans l'espace de Hardy du disque $H_2(U)$. Mentionnons qu'une approche différente, basée sur la théorie des opérateurs de Hankel, fournit un critère d'unicité, voir [13, Theorem 7], lorsque la fonction approchée g est une fonction rationnelle de degré n. Ce critère ne nécessite pas la connaissance d'estimées de l'erreur partout sur le cercle. Il s'appuie sur la décroissance géométrique des erreurs en norme L^2 d'une suite de points critiques de degrés 0 à $n-1$. De cette décroissance, on obtient l'unicité d'un point critique et donc d'un meilleur approximant au degré $n-1$.

Le Chapitre 3 concerne l'interpolation rationnelle de la fonction exponentielle. On décrit comment le résultat de convergence uniforme des approximants de Padé de la fonction exponentielle (théorème de Padé) peut se généraliser aux interpolants rationnels, ou approximants de Padé multi-points.

En général, l'occurrence de pôles pour les interpolants qui ne correspondent pas à des singularités de la fonction interpolée ("spurious poles" en anglais) détruisent la convergence uniforme de ces interpolants. La convergence en mesure ou en capacité (plus faible que la convergence uniforme) est la notion adaptée pour prendre en compte le phénomène des "spurious poles", voir à ce sujet le théorème de Nuttall–Pommerenke.
Ici, la généralisation obtenue repose fortement sur les propriétés de la fonction exponentielle. Dans le cas de points d'interpolation choisis sur un compact de l'axe réel, la preuve consiste à dériver suffisamment de fois la relation d'interpolation en utilisant le théorème de Rolle. Cela est décrit dans le paragraphe 3.1. Des résultats de géométrie des polynômes, assez proches du théorème de Laguerre sur les zéros des dérivées polaires, permettent alors de déduire la convergence de *tous* les pôles des interpolants vers l'infini, et par suite la convergence uniforme des interpolants elle-même. Le cas des points d'interpolation dans un compact du plan complexe, symétriques par conjugaison est abordé dans le paragraphe 3.2. Il nécessite un théorème de Rolle complexe qui s'applique aux polynômes exponentiels réels. Un tel résultat peut s'obtenir à partir de la connaissance de la répartition asymptotique des zéros des polynômes exponentiels donnée par le théorème de Polya-Dickson, et en adaptant le théorème de Jensen sur les zéros de la dérivée d'un polynôme réel. Cela permet une généralisation partielle du théorème de Padé aux points d'interpolation complexes.

Le Chapitre 4 s'intéresse à la généralisation du théorème de Padé dans une autre direction, celle des approximants de Hermite–Padé. Comme il est bien connu, les approximants rationnels simultanés (ou approximants de Hermite–Padé de type II, selon une terminologie introduite par Mahler) ont été considérés pour la première fois par Hermite pour prouver la transcendance de e. Les liens entre l'approximation de Padé et l'irrationalité ou la transcendance des nombres sont illustrés par les deux assertions suivantes, empruntées à [67].

Soit x un nombre réel. Supposons qu'il existe des entiers p_n et q_n, $n \in \mathbb{N}$, tels que
1) $q_n x - p_n \neq 0, \quad n \in \mathbb{N}$,
2) $\lim_{n\to\infty}(q_n x - p_n) = 0$.
Alors x est irrationnel.

Soit x un nombre réel. Supposons que pour tout entier $m \in \mathbb{N}$ et pour tous entiers $a_0, a_1, \ldots, a_m \in \mathbb{Z}$, il existe des entiers $p_{0,n}, p_{1,n}, \ldots, p_{m,n}$, $n \in \mathbb{N}$, tels que
1) $\sum_{k=0}^{m} a_k p_{k,n} \neq 0, \quad n \in \mathbb{N}$,
2) $\lim_{n\to\infty}(p_{0,n} x^k - p_{k,n}) = 0, \quad k = 1,2,\ldots,m$.
Alors x est transcendant.

Les approximants de type I, ceux que l'on étudie ici, ont été également introduits par Hermite. Le lien entre les approximants de chacun des deux types est classique et donné par les relations de Mahler. On s'intéresse à un ensemble de polynômes $A_0, \ldots, A_m$, de degrés strictement inférieurs à n tels que

$$\sum_{p=0}^{m} A_p(z) e^{pz} = O(z^{mn+n-1}).$$

On prouve pour ces approximants l'analogue du théorème de Padé, à savoir la convergence de chacun des coefficients A_p vers des exponentielles, lorsque $n \to \infty$, voir le paragraphe 4.1. La démonstration consiste à appliquer la méthode du col aux expressions intégrales des A_p,

$$A_p(z) = \frac{1}{2i\pi} \int_{C_0} \frac{e^{vz}dv}{\prod_{l=0}^{m}(v+p-l)^n}, \quad 0 \leq p \leq m,$$

où C_0 désigne un cercle centré en l'origine et de rayon inférieur à 1.
On applique ces résultats de convergence pour améliorer une inégalité de Mahler concernant une minoration de la distance à $\mathbb{N}$ des exponentielles de nombres entiers. On prouve, pour tout entier g positif suffisamment grand, l'inégalité

$$|e^g - a| > g^{-20g}, \quad a \in \mathbb{N}.$$

Cela s'obtient en choisissant un ensemble de polynômes $A_0, \ldots, A_m$, dont les degrés ne sont pas constants (cas "diagonal") mais fonctions de l'indice.
Enfin, on étudie la distribution asymptotique des zéros des polynômes $A_0, \ldots, A_m$. Les expérimentations numériques décrivent la répartition bien particulière de ces zéros dans le plan complexe. H. Stahl a montré comment la méthode du col permet encore d'expliquer, au moins dans le cas quadratique, cette répartition asymptotique : cela passe par l'étude d'une fonction harmonique définie sur la surface de Riemann paramétrant les points critiques de la fonction

$$zv - \sum_{l=0}^{m} \log(v-l),$$

à comparer avec l'intégrande dans les expressions intégrales précédentes des A_p, $p = 0, \ldots, m$. En conjonction avec la méthode déjà évoquée précédemment, reposant sur l'étude d'un problème de Riemann-Hilbert, on a pu utiliser cette analyse pour obtenir des asymptotiques très précises des approximants de Hermite–Padé ainsi que de leurs zéros.

Chapitre 2

Approximation rationnelle dans l'espace de Hardy H^2 du disque et unicité d'un meilleur approximant

Notations

$\mathbb{T}$, U, V	:	cercle unité, disque unité ouvert, complémentaire dans $\overline{\mathbb{C}}$ du disque unité fermé.
T_r, U_r, V_r	:	cercle de rayon r, disque ouvert de rayon r, complémentaire dans $\overline{\mathbb{C}}$ du disque fermé de rayon r (centrés à l'origine).
$\mathcal{P}_n$	:	espace des polynômes réels de degré au plus n ; paramétrés par leurs coefficients, $\mathcal{P}_n$ est équipé de la topologie euclidienne de $\mathbb{R}^{n+1}$.
$\mathcal{M}_n$	:	polynômes réels unitaires de degré n.
$\widetilde{\mathcal{M}}_n$	:	polynômes réels de degré au plus n de coefficient constant 1.
$\mathcal{M}_n^r$	:	polynômes réels unitaires de degré n ayant toutes leurs racines dans U_r.
$\widetilde{\mathcal{M}}_n^r$	:	polynômes réels de degré au plus n de coefficient constant 1 ayant toutes leurs racines dans V_r.
Δ_n	:	polynômes réels unitaires de degré n ayant toutes leurs racines dans $\overline{U}$; adhérence de $\mathcal{M}_n^1$ par rapport à la topologie euclidienne de $\mathcal{P}_n$.
$\widetilde{\Delta}_n$	:	polynômes réels de degré au plus n de coefficient constant 1 ayant toutes leurs racines dans $\overline{V}$; adhérence de $\widetilde{\mathcal{M}}_n^1$ par rapport à la topologie euclidienne de $\mathcal{P}_n$.
$\Vert\cdot\Vert_\infty$, $\Vert\cdot\Vert_2$	:	normes dans $L_\infty(\mathbb{T})$, et $L_2(\mathbb{T})$ respectivement.
$< \cdot,\cdot >$	:	produit scalaire dans $L_2(\mathbb{T})$.
$L_{2,\mathbb{R}}(\mathbb{T})$	:	sous-espace réel de $L_2(\mathbb{T})$ des fonctions à coefficients de Fourier réels.
$C_{\mathbb{R}}(\mathbb{T})$	:	sous-espace réel de $L_{2,\mathbb{R}}(\mathbb{T})$ des fonctions continues.
$H_{2,\mathbb{R}}(U)$	:	espace de Hardy réel du disque unité des fonctions dans $L_{2,\mathbb{R}}(\mathbb{T})$ dont les coefficients de Fourier d'indices négatifs sont nuls.
$H^0_{2,\mathbb{R}}(V)$	:	espace de Hardy réel du complémentaire du disque unité dont les fonctions s'annulent à l'infini; complément orthogonal de $H_{2,\mathbb{R}}(U)$ dans $L_{2,\mathbb{R}}(\mathbb{T})$.
$\mathbb{P}_+$, $\mathbb{P}_-$	:	projections orthogonales $L_{2,\mathbb{R}}(\mathbb{T}) \to H_{2,\mathbb{R}}(U)$ et $L_{2,\mathbb{R}}(\mathbb{T}) \to H^0_{2,\mathbb{R}}(V)$, respectivement.
$H_{\infty,\mathbb{R}}(U)$	:	sous-espace réel de $H_{2,\mathbb{R}}(U)$ des fonctions essentiellement bornées.
$H^n_{\infty,\mathbb{R}}(U)$	:	ensemble des fonctions méromorphes h/q avec $h \in H_{\infty,\mathbb{R}}(U)$ et $q \in \mathcal{M}_n^1$.
$\mathcal{R}^0_{m,n}(V)$	:	sous-ensemble de $H^0_{2,\mathbb{R}}(V)$ des fonctions rationnelles $p/z^{m-n+1}q$ avec $p \in \mathcal{P}_m$ et $q \in \mathcal{M}_n^1$.
$\widetilde{\mathcal{R}}_{m,n}(U)$	:	sous-ensemble de $H_{2,\mathbb{R}}(U)$ des fonctions rationnelles $\widetilde{p}/\widetilde{q}$ avec $\widetilde{p} \in \mathcal{P}_m$ et $\widetilde{q} \in \widetilde{\mathcal{M}}_n^1$.

On considère les deux problèmes d'approximation suivants :

Pb(V,m,n): *Soit $f \in H^0_{2,\mathbb{R}}(V)$ et deux entiers positifs m, n avec $m \geq n-1$, minimiser*

$$\left\| f - \frac{p}{z^{m-n+1}q} \right\|_2$$

lorsque $p/z^{m-n+1}q$ parcourt l'ensemble $\mathcal{R}^0_{m,n}(V)$.

Pb(U,m,n): *Soit $g \in H_{2,\mathbb{R}}(U)$ et deux entiers positifs m, n avec $m \geq n-1$, minimiser*

$$\left\| g - \frac{\widetilde{p}}{\widetilde{q}} \right\|_2$$

lorsque $\widetilde{p}/\widetilde{q}$ parcourt l'ensemble $\widetilde{\mathcal{R}}_{m,n}(U)$.

Ces deux problèmes d'approximation sont en fait équivalents. Soit

$$h^\sigma(z) := \frac{1}{z} h(\frac{1}{z}),$$

l'application $h \to h^\sigma$ est une isométrie involutive de $L_{2,\mathbb{R}}(\mathbb{T})$ échangeant $H_{2,\mathbb{R}}(U)$ (resp. $\widetilde{\mathcal{R}}_{m,n}(U)$) et $H^0_{2,\mathbb{R}}(V)$ (resp. $\mathcal{R}^0_{m,n}(V)$), de sorte que **Pb**(U,m,n) est le tranformé de **Pb**(V,m,n) par cette application.
Le problème **Pb**(U,m,n) est classique en théorie de l'approximation (voir e.g. [11, 21, 23, 24, 35, 72], [3*]) et se trouve d'autre part lié à différentes questions issues des sciences appliquées, telles que la théorie des systèmes dynamiques, et plus précisément à la théorie du contrôle [22], ainsi qu'au traitement du signal [19, 36, 51].
Comme le lecteur peut le remarquer, l'énoncé de ces deux problèmes est donné dans des espaces de fonctions à coefficients réels. Cela tient aux domaines d'applications qui viennent d'être évoqués, et pour lesquels seul le cas réel prend un sens. Toutes les questions abordées pourraient l'être dans des espaces de fonctions à coefficients complexes. Cependant, ce cadre plus général peut amener quelques différences par rapport au cas réel. Notamment, le meilleur approximant rationnel complexe d'une fonction réelle n'est pas forcément réel. Une autre remarque est que l'on considère ici des fonctions rationnelles p/q de type (m,n) (i.e. $p \in \mathcal{P}_m$ et $q \in \mathcal{M}^1_n$) avec $m \geq n-1$, c'est à dire le cas super-diagonal, qui est une généralisation relativement directe du cas diagonal $m = n-1$. La généralisation au cas sous-diagonal $m < n-1$ est moins aisée.
Le problème d'approximation étudié ici admet une généralisation matricielle difficile mais particulièrement importante pour les domaines d'applications mentionnés précédemment, voir [12].

2.1 Motivations provenant de la théorie des systèmes

Pour donner à notre problème d'approximation un cadre plus concret et pour en motiver son étude, nous commencerons par un bref rappel des notions élémentaires intervenant en théorie des systèmes. On peut consulter [6], en particulier le Chapitre 1, pour plus de détails. La théorie des systèmes dynamiques, ou automatique, peut se définir de manière très générale comme l'étude d'opérateurs entre fonctions du temps.

2.1.1 Systèmes discrets et identification stochastique

Un système dynamique discret est une application $u \to y$ où l'entrée

$$u = (\dots, u_{k-1}, u_k, u_{k+1}, \dots)$$

est une fonction à valeurs réelles du temps discret k, qui génère une sortie

$$y = (\dots, y_{k-1}, y_k, y_{k+1}, \dots)$$

du même type, où y_k ne dépend que de u_j pour $j \leq k$. Le système est dit stationnaire si une translation dans le temps du signal d'entrée produit une translation identique pour le signal de sortie. On s'intéresse aux systèmes linéaires

$$y_k = \sum_{j=0}^{\infty} f_j u_{k-j},$$

où la sortie à un instant k est une combinaison linéaire des entrées passées avec des coefficients constants $f_j \in \mathbb{R}$. Considérons les fonctions génératrices

$$u(z) = \sum_{k \in \mathbb{Z}} u_k z^{-k}, \quad y(z) = \sum_{k \in \mathbb{Z}} y_k z^{-k},$$

et définissons la fonction de transfert du système

$$f(z) = \sum_{k=0}^{\infty} f_k z^{-k}.$$

Alors, le comportement entrée-sortie du système est décrit par l'équation

$$y(z) = f(z) u(z),$$

et de plus
(i) le système est un opérateur borné $l^2 \to l^2$ si et seulement si $f \in \overline{H}_{\mathbb{R}}^{\infty}$, on dit que le système est (l^2, l^2)-stable;
(ii) le système est un opérateur borné $l^2 \to l^\infty$ si et seulement si $f \in \overline{H}_{\mathbb{R}}^{2}$, on dit que le système est (l^2, l^∞)-stable.
Un système linéaire est dit de dimension finie n s'il peut être décrit en terme d'une variable d'état $x_k \in \mathbb{R}^n$,

$$x_{k+1} = A x_k + B u_k, \quad y_k = C x_k + D u_k,$$

où A est une matrice réelle $n \times n$, B (resp. C) un vecteur colonne (resp. ligne) à n entrées réelles, et D un nombre réel, n étant le plus petit entier tel que ces équations soient vérifiées. Un système linéaire stationnaire est de dimension n si et seulement si sa fonction de transfert est rationnelle de degré n et analytique à l'infini (cf. [31], [52]). La fonction de transfert s'écrit alors

$$f(z) = D + C(zI_n - A)^{-1} B.$$

Soit un processus stochastique décrit par une série temporelle de nombres réels

$$(\dots, y_{k-1}, y_k, y_{k+1}, \dots).$$

Si l'on suppose que ce processus est régulier, i.e. purement aléatoire en un certain sens, cf. [57], on peut le représenter sous la forme

$$y_k = \sum_{j=0}^{\infty} f_j u_{k-j}, \tag{2.1.1}$$

où le processus u est un bruit blanc (i.e. u_i est d'espérance nulle, et $\mathrm{cov}(u_i,u_j) = 1$ si $i = j$ et 0 sinon), et $(f_j)_{j\geq 0}$ une suite de nombres réels de carré sommable,

$$\sum_{j=0}^{\infty} f_j^2 < \infty.$$

Les modèles ne possédant qu'un nombre fini de paramètres, en l'occurrence les modèles ARMA (auto-régressifs et moyenne mobile), sont intéressants en pratique car ce sont principalement ceux pour lesquels les calculs peuvent être menés. Ils sont décrits par une équation du type

$$y_k + q_{n-1}y_{k-1} + \cdots + q_0 y_{k-n} = p_n u_k + \cdots + p_0 u_{k-n}. \tag{2.1.2}$$

Soit

$$p(z) = p_n z^n + \cdots + p_0, \quad q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0,$$

et

$$p/q(z) = \sum_{i=0}^{\infty} h_i z^{-i},$$

le développement de la fraction rationnelle propre p/q, alors le système (2.1.2) se réécrit sous la forme

$$y_k = \sum_{j=0}^{\infty} h_j u_{k-j}. \tag{2.1.3}$$

On peut alors approcher le modèle (2.1.1) par un modèle du type (2.1.2), en choisissant de minimiser la variance de l'erreur entre la sortie du système (2.1.1) et celle du modèle prédictif de type (2.1.2), c'est à dire en résolvant le problème $\mathbf{Pb}(V,n,n)$ pour la fonction $f(z)/z$.

2.1.2 systèmes continus et identification harmonique

Un système dynamique continu est une application continue $\sigma : u(t) \to y(t)$, où u et y sont des fonctions définies sur $[0,\infty)$, à valeurs dans $\mathbb{R}$, et où l'on choisit, par exemple, comme entrées admissibles les fonctions $u \in L^1$ à support compact. On suppose que le système σ est linéaire et stationnaire, i.e.

$$\Delta_r(\sigma(u)) = \sigma(\Delta_r(u)), \quad \forall r \geq 0,$$

où Δ_r est l'opérateur de translation

$$\Delta_r(g(t)) = g(t-r),$$

g une fonction sur $\mathbb{R}$. Alors, on peut vérifier que l'application σ est un opérateur de convolution

$$y(t) = \int_0^{\infty} h(t-\tau)u(\tau)d\tau. \tag{2.1.4}$$

La fonction $h : [0,\infty) \to \mathbb{R}$ est la réponse impulsionnelle du système, et correspond formellement à la sortie engendrée par une impulsion de Dirac. Si u et h sont à croissance exponentielle, il en est de même de y et les transformées de Laplace $U(s)$, $Y(s)$ et $H(s)$ sont définies sur un demi-plan droit $\{\mathrm{Re} z > a\}$. Dans le domaine fréquentiel, l'équation de convolution (2.1.4) se réécrit

$$Y(s) = H(s)U(s),$$

où H est la fonction de transfert du système. Ainsi, les fonctions de transfert rationnelles de degré n correspondent à des opérateurs différentiels linéaires d'ordre n. Soit $\mathcal{H}^p_{\mathbb{R}}$, $1 \leq p \leq \infty$, les espaces de Hardy réels du demi-plan droit, le comportement entrée-sortie du système, ou plus précisément sa stabilité, peut être caractérisé en terme de la fonction de transfert de la manière suivante :
$\sigma : L^2[0,\infty) \to L^2[0,\infty)$ si et seulement si $H \in \mathcal{H}^\infty_{\mathbb{R}}$,
$\sigma : L^2[0,\infty) \to L^\infty[0,\infty)$ si et seulement si $H \in \mathcal{H}^2_{\mathbb{R}}$,
$\sigma : L^\infty[0,\infty) \to L^\infty[0,\infty)$ si et seulement si $H \in \mathcal{W}_{\mathbb{R}}$, l'algèbre de Wiener du demi-plan droit, soit l'ensemble des transformées de Laplace des fonctions sommables $[0,\infty) \to \mathbb{R}$.
Dans le cadre de notre étude, il est bien entendu que ce sont les systèmes $\sigma : L^2[0,\infty) \to L^\infty[0,\infty)$ correspondant à des fonctions de transfert $H \in \mathcal{H}^2_{\mathbb{R}}$ qui nous intéressent.
L'identification harmonique consiste à appliquer une entrée périodique $u(t) = e^{i\omega t}$ au système σ et à observer la sortie à l'équilibre,

$$y(t) = \lambda e^{i\Phi} e^{i\omega t},$$

où l'on a :

$$\lambda = |H(i\omega)|, \quad \Phi = \arg H(i\omega).$$

On peut ainsi déterminer certaines valeurs de H sur l'axe imaginaire et reconstruire, par exemple en résolvant un problème extrémal borné, un modèle de dimension infinie dont la fonction de transfert soit une fonction de $\mathcal{H}^2_{\mathbb{R}}$. Il reste à approcher cette fonction par une fraction rationnelle analytique dans $\{\mathrm{Re}(s) > 0\}$, au sens de la norme de l'espace $\mathcal{H}^2_{\mathbb{R}}$. Comme l'application

$$H(s) \to f(z) = 2\sqrt{\pi}\frac{(H \circ \varphi)(z)}{z-1},$$

où

$$\varphi(z) = \frac{z+1}{z-1},$$

est une isométrie de $\mathcal{H}^2_{\mathbb{R}}$ sur $H^0_{2,\mathbb{R}}(V)$ qui préserve les fonctions rationnelles et leur degré, on est ramené à notre problème d'approximation $\mathbf{Pb}(V,n,n)$.

2.2 Premières propriétés

En premier lieu, le problème $\mathbf{Pb}(V,m,n)$, où de manière équivalente le problème $\mathbf{Pb}(U,m,n)$, admet toujours une solution, c'est à dire qu'un *meilleur approximant existe* (voir [72, 23, 5]). Cependant, celui-ci n'est pas forcément unique. Comme l'on peut s'y attendre, des exemples de *non-unicité* sont basés sur l'existence de symétrie pour la fonction f à approcher. Considérons, par exemple, une fonction f paire qui ne soit pas rationnelle, et soit $r(z)$ un meilleur approximant de f. Alors $r(-z)$ est également un meilleur approximant. Supposons que $r(z)$ s'écrive sous forme irréductible $p/q(z)$, et que le degré du dénominateur q soit impair, ce qui sera le cas si le degré d'approximation est impair, comme nous le verrons dans la suite (propriété de normalité). Alors l'équation

$$p(z)q(-z) = p(-z)q(z) \tag{2.2.1}$$

implique $q(z) = -q(-z)$, puisque q est unitaire de degré impair. Donc, par (2.2.1), il en est de même de p, ce qui induit un facteur commun entre p et q, et une contradiction avec leur coprimarité.
Le problème $\mathbf{Pb}(V,m,n)$ peut se décomposer en deux parties, une partie linéaire correspondant au numérateur de la fraction rationnelle, et une partie non linéaire correspondant à son dénominateur. En effet, comme l'espace $H^0_{2,\mathbb{R}}(V)$ est muni d'une norme issue d'un produit scalaire, on obtient, pour un dénominateur q fixé, que la fraction rationnelle la plus proche de f est la *projection orthogonale* de f sur l'espace vectoriel

$$V_q := \mathcal{P}_m/z^{m-n+1}q \subset H^0_{2,\mathbb{R}}(V). \tag{2.2.2}$$

Le véritable paramètre à ajuster est donc le dénominateur q, la partie non linéaire dans notre famille d'approximants. Mentionnons que notre problème, comme tout problème non linéaire, admet des solutions qui sont les meilleurs approximants de f, mais également tout un ensemble de points critiques (que nous allons définir plus précisement) parmi lesquels des minima et des maxima locaux.
Une propriété essentielle, concernant le dénominateur q, est celle de *normalité*. Elle stipule que, pour un ordre d'approximation (m,n) fixé et pour une fonction f qui n'est pas une fraction rationnelle dans $\mathcal{R}^0_{m-1,n-1}(V)$, tout minimum local $p/z^{m-n+1}q$ pour le problème $\mathbf{Pb}(V,m,n)$ est tel que $\deg q = n$ avec p et q premiers entre eux (voir [2*, Prop. 2.1] ou [34, Prop. 2.2]). Autrement dit, pour construire le dénominateur d'une fraction rationnelle qui soit un minimum local pour $\mathbf{Pb}(V,m,n)$, il faut nécessairement utiliser tous les paramètres dont on dispose. Mentionnons, en passant, que cette propriété de normalité n'est pas toujours vérifiée lorsqu'on s'intéresse à d'autres normes, le meilleur approximant rationnel d'un produit de Blaschke étant, par exemple dans le cas de la norme du Sup, la fonction nulle.
A présent, nous allons tirer parti de la différentiabilité de la norme L^2 pour étudier la régularité de notre critère, et en particulier définir de manière plus précise la notion de *point critique* pour $\mathbf{Pb}(V,m,n)$. Rappelons que $\mathcal{P}_m$ et $\mathcal{M}^1_n$ sont munis de la topologie euclidienne de $\mathbb{R}^{m+1}$ et $\mathbb{R}^n$ respectivement. Alors, l'application $\mathcal{P}_m \times \mathcal{M}^1_n \to \mathbb{R}$ donnée par

$$(p,q) \to \left\| f - \frac{p}{z^{m-n+1}q} \right\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \left| f - \frac{p}{z^{m-n+1}q} \right|^2 d\theta \tag{2.2.3}$$

est lisse, comme on peut le voir en dérivant sous l'intégrale. Un couple (p_c,q_c) sera dit critique si la dérivée de l'application précédente s'y annule. Si pour $q \in \mathcal{M}^1_n$, on note $L^f_m(q) \in \mathcal{P}_m$, le numérateur de la projection orthogonale de f sur le sous-espace $\mathcal{P}_m/z^{m-n+1}q \subset H^0_{2,\mathbb{R}}(V)$, on peut considérer l'application $\mathcal{M}^1_n \to \mathbb{R}$ donnée par

$$\psi^f_{m,n}(q) := \left\| f - L^f_m(q)/z^{m-n+1}q \right\|_2^2 = \|f\|_2^2 - \|L^f_m(q)/z^{m-n+1}q\|_2^2, \tag{2.2.4}$$

dont on peut vérifier qu'elle est lisse comme fonction de q. De plus, une paire critique (p_c,q_c) est nécessairement de la forme $(L^f_m(q_c),q_c)$, où q_c est un point critique de $\psi^f_{m,n}$. Ainsi, une fraction rationnelle de $\mathcal{R}^0_{m,n}(V)$ sera dite critique si elle est de la forme $L^f_m(q_c)/q_c$ avec q_c un point critique de $\psi^f_{m,n}$. On peut noter que, les points critiques étant définis localement, leur étude est plus aisée que celle des meilleurs approximants.
La propriété suivante concerne l'approximation d'une fraction rationnelle à son propre ordre. On dira que $\mathbf{Pb}(V,m,n)$ est un *problème consistant* si pour toute fraction rationnelle $f(z) = p/z^{m-n+1}q \in \mathcal{R}^0_{m,n}(V)$, avec p et q premiers entre eux, l'application $\psi^f_{m,n}$ admet un unique point critique, à savoir la fonction f elle-même. On peut montrer que $\mathbf{Pb}(V,m,n)$, $n \geq 1$, est consistant

si $m \geq n-2$ et non consistant sinon (cf. [6] pour le cas $m = n-1$ et [34, Theorem 4.2] pour le cas général).
Le fait pour une fraction rationnelle d'être un point critique vis à vis du problème $\mathbf{Pb}(V,m,n)$ peut s'énoncer en terme de division, ou de manière équivalente en terme de *propriétés d'interpolation*. Pour cela, il nous faut d'abord introduire la notation

$$\widetilde{p}(z) := z^k p(1/z),$$

qui définit le polynôme réciproque de $p \in \mathcal{P}_k$. Remarquons que cette définition a un sens, uniquement si l'on précise l'espace $\mathcal{P}_k$ ambiant, celui dont le polynôme p est vu comme un élément. Par exemple, dans la suite, il est sous-entendu que $q \in \mathcal{P}_n$ et $L_m^f(q) \in \mathcal{P}_m$. On peut à présent caractériser la projection $L_m^f(q)$ pour un polynôme $q \in \mathcal{M}_n^1$ fixé par la proposition suivante, qui est essentiellement une reformulation de [72, Chapter 9, Theorem 2] (cf. aussi [11]) dans un cadre plus général.

Proposition 2.2.1 [5*, Lemma 2.1] *Soit $f \in H_{2,\mathbb{R}}^0(V)$ et $q \in \mathcal{M}_n^1$. Ecrivons la division de $f^\sigma \widetilde{q}$ par qz^{m-n+1} :*

$$f^\sigma \widetilde{q} = v_q q z^{m-n+1} + r, \tag{2.2.5}$$

où $r \in \mathcal{P}_m$ et $v_q \in H_{2,\mathbb{R}}(U)$. Alors $L_m^f(q) = \widetilde{r}$.

Autrement dit, pour un dénominateur q fixé, le meilleur approximant de f est l'interpolant de cette fonction aux réfléchis à travers $\mathbb{T}$ des racines de q, ainsi qu'à l'infini avec la multiplicité $m-n+1$. Il existe également une caractérisation analogue pour les points critiques de $\psi_{m,n}^f$:

Proposition 2.2.2 [5*, Proposition 2.3] *Soit $q \in \mathcal{M}_n^1$ et considérons à nouveau la division (2.2.5). Alors q est un point critique de $\psi_{m,n}^f$ si et seulement si q divise $v_q L_m^f(q)$ dans $H_{2,\mathbb{R}}(U)$. En particulier, si q est un point critique irréductible, alors q divise v_q dans $H_{2,\mathbb{R}}(U)$.*

A nouveau, on peut réinterpréter la Proposition 2.2.2 en terme d'interpolation. En effet, en remplaçant r par $\widetilde{L_m^f}(q)$ et en changeant z en $1/z$ dans (2.2.5), on obtient

$$f - \frac{L_m^f(q)}{z^{m-n+1}q} = \frac{v_q^\sigma \widetilde{q}}{z^{m-n+1}q}.$$

Ainsi, q est un point critique si et seulement si pour chaque zéro ξ de q de multiplicité n_1 qui soit un pôle d'ordre $n_2 \leq n_1$ de $L_m^f(q)/q$, la fraction $L_m^f(q)/qz^{m-n+1}$ interpole f en $1/\xi$ à un ordre $n_1 + n_2$, ainsi qu'à l'infini avec la multiplicité $m-n+1$. En particulier,
si q est un point critique irréductible, $L_m^f(q)/q$ est l'approximant de Padé multipoint de $z^{m-n+1}f$ avec interpolation d'ordre 2 aux réciproques des racines de q et d'ordre $m-n+1$ à l'infini.
On peut également remarquer qu'un point critique réductible correspond à une fraction rationnelle qui interpole $z^{m-n+1}f$ avec un défaut de normalité, c'est à dire que la forme irréductible p_1/p_2 de la fraction $L_m^f(q)/q$ interpole $z^{m-n+1}f$ dans $\overline{\mathbb{C}} \setminus \overline{U}$ en plus de $\deg p_1 + \deg p_2 + 1$ points, en comptant les multiplicités.
Les points critiques vérifiant des propriétés d'interpolation par rapport à la fonction f, on peut s'attendre à ce qu'ils convergent vers cette fonction, lorsqu'on considère une suite de points critiques tel que le degré m tende vers l'infini. Avant d'énoncer ce résultat, nous allons introduire l'hypothèse selon laquelle f non seulement appartient à $H_{2,\mathbb{R}}^0(V)$ mais s'étend également de manière analytique à un voisinage de $\overline{V}$:

(H) il existe un $\eta > 0$ tel que $f(z)$ est analytique pour $|z| > 1-\eta$, $f(\overline{z}) = \overline{f(z)}$ et $f(\infty) = 0$.

Proposition 2.2.3 [5*, Proposition 2.4] *Si l'hypothèse* **(H)** *est satisfaite, les applications*

$$L_m^f : \mathcal{M}_n^1 \to \mathcal{P}_m, \qquad \psi_{m,n}^f : \mathcal{M}_n^1 \to \mathbb{R}$$

s'étendent de manière lisse à un voisinage de Δ_n dans $\mathcal{M}_n$.

En particulier, l'équation de division (2.2.5),

$$f^\sigma \widetilde{q} = v_q q z^{m-n+1} + \widetilde{L}_m^f(q), \tag{2.2.6}$$

reste vérifiée pour tout polynôme q de Δ_n. On peut à présent énoncer notre résultat de convergence sur le fermé $\overline{V}$.

Proposition 2.2.4 [5*, Proposition 2.8] *Supposons l'hypothèse* **(H)** *vérifiée avec f analytique dans un domaine Ω simplement connexe contenant $\overline{V}$. Soit $\check{\Omega} = \{1/z; z \in \Omega\}$. Pour $q \in \Delta_n$, écrivons*

$$f^\sigma \widetilde{q} = v_q q z^{m-n+1} + \widetilde{L}_m^f(q)$$

la division de $f^\sigma \widetilde{q}$ par qz^{m-n+1}. Alors, l'ensemble

$$\{v_q; \quad q \in \Delta_n,\, n \in \mathbb{N},\, m \in \mathbb{N},\, m \geq n-1\}$$

est une famille normale de fonctions dans $\check{\Omega}$. Si (m_k,n_k) est une suite de couples d'entiers positifs tel que $m_k \geq n_k - 1$ et $\lim_{k\to\infty} m_k = \infty$, et si, pour chaque k, $q_{n_k} \in \Delta_{n_k}$ est un point critique de ψ_{m_k,n_k}^f, alors

$$\lim_{k\to\infty} \frac{L_{m_k}^f(q_{n_k})}{q_{n_k}} = f \tag{2.2.7}$$

uniformément sur $\overline{V}$.

Une conséquence de cette proposition est la suivante. Il existe un degré à partir duquel les points critiques ne sont plus des maxima locaux pour notre critère. Cela s'obtient en analysant localement la variation au second ordre de ce critère et par un calcul de la matrice hessienne en un point critique, voir [1*].

D'un point de vue numérique, le calcul des approximants H_2 dans le cas diagonal $m = n - 1$ a été implémenté dans le logiciel *Hyperion* au sein du projet MIAOU. De nombreuses autres fonctionnalités sont incluses dans *Hyperion*, par exemple, le calcul d'approximants pondérés ou bien matriciels, ainsi qu'un programme de complétion analytique, voir [28, 9]. Ce logiciel permet en particulier de résoudre un problème d'identification de filtres hyperfréquences par approximation dans le plan complexe.

On suppose comme précédemment que l'hypothèse **(H)** de prolongement est vérifiée. L'algorithme scalaire [8], [2*] que nous allons décrire est un algorithme de gradient basé tout d'abord sur une description des domaines compacts Δ_n. Rappelons qu'on utilise les coefficients des polynômes pour paramétrer dans $\mathbb{R}^n$ les éléments de Δ_n. Il est clair que Δ_1 est le segment $[-1,1]$. Le domaine Δ_2 est le triangle de sommet $(-2,1)$, $(2,1)$ et $(0,-1)$, voir Figure 1, et le domaine Δ_3 est représenté sur la Figure 2.

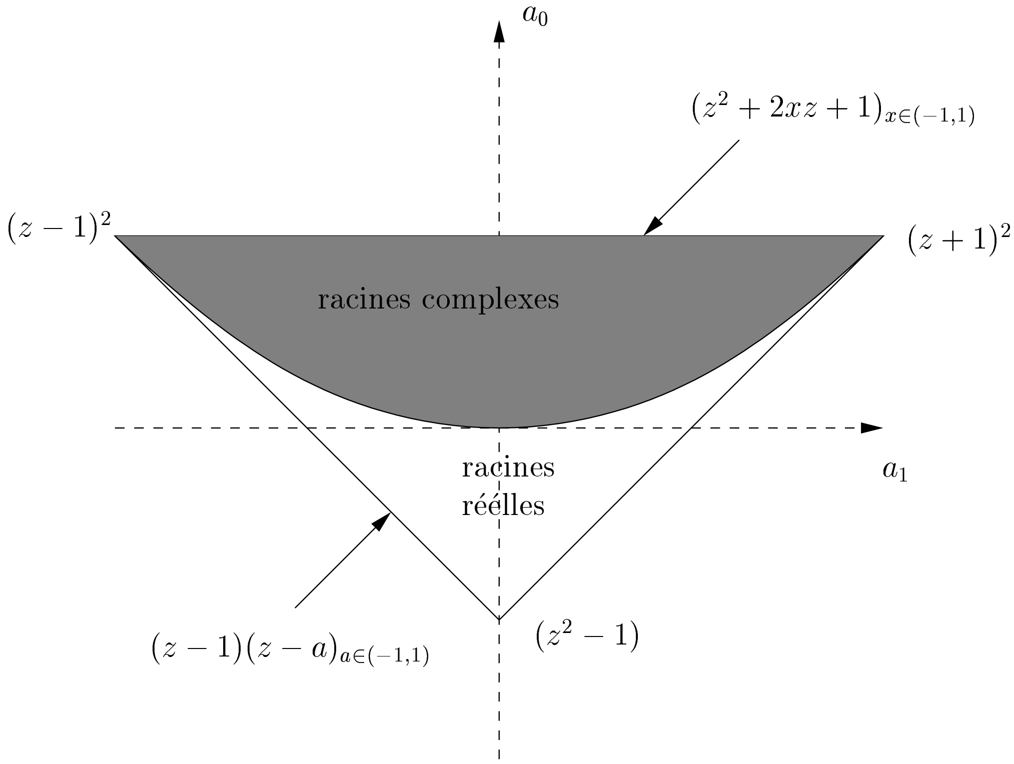

Figure 1

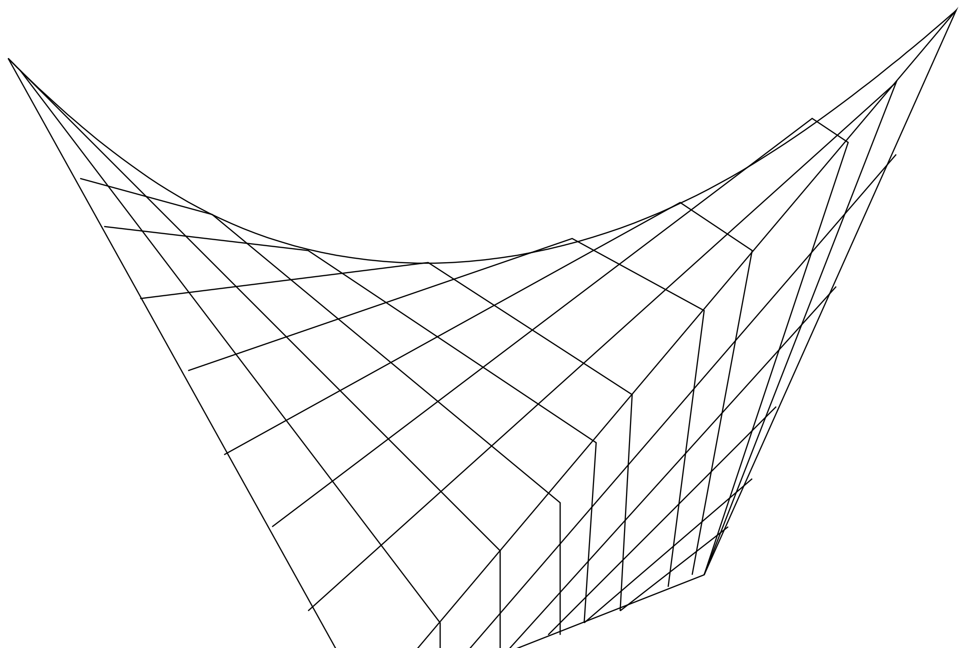

Figure 2

Le domaine Δ_n est homéomorphe à la boule de dimension n et le bord $\partial\Delta_n$ de Δ_n consiste en les polynômes de Δ_n ayant au moins une racine de module 1. Parmi ces polynômes, on peut distinguer ceux ayant 1 ou -1 comme racine, et ceux n'ayant pas de racines réelles de module 1, mais qui sont divisibles par un polynôme de la forme $z^2+2xz+1$ avec $x \in (-1,1)$. Ainsi, le bord $\partial\Delta_n$ est l'union de $(z-1)\Delta_{n-1}$, $(z+1)\Delta_{n-1}$ et $(z^2+2xz+1)\Delta_{n-2}$, $x \in (-1,1)$, donc est constitué de deux copies de Δ_{n-1} et d'une infinité de copies de Δ_{n-2}. D'autre part, $\partial\Delta_n$ possède des coins et n'est donc pas lisse. La partie lisse de $\partial\Delta_n$ se caractérise comme étant constituée des polynômes de Δ_n ayant exactement un facteur irréductible sur $\mathbb{R}$ avec des racines de module 1. Enonçons ensuite un lemme qui découle facilement de l'équation de division (2.2.6), et un corollaire.

Lemme 2.2.5 [8, Lemma1], [2*, Proposition 3.3] *Soit $q \in \partial\Delta_n$, $q = q_u q_i$ tel que q_u est un polynôme unitaire de degré $k \leq n$ ayant toutes ses racines de module 1 et $q_i \in \mathcal{M}^1_{n-k}$. Alors*

$$\psi^f_{n-1,n}(q) = \psi^f_{n-k-1,n-k}(q_i).$$

Notons $\nabla_n(q)$ le gradient de $\psi^f_{n-1,n}$ au point $q \in \Delta_n$. Le Lemme 2.2.5 et la propriété de normalité entrainent le corollaire suivant.

Corollaire 2.2.6 [8, Corollary of Lemma 1], [2*, Proposition 3.5] *Soit q comme dans le Lemme 2.2.5, appartenant à la partie lisse de $\partial\Delta_n$, et tel que q_i est un point critique de Δ_{n-k} ($k=1$ ou 2). Alors, s'il n'est pas nul, le gradient $\nabla_n(q)$ est orthogonal à $\partial\Delta_n$ et pointe vers l'extérieur du domaine.*

Supposons que la fonction f ne soit pas rationnelle de degré inférieur à n. Nous ferons également une hypothèse supplémentaire qui est vérifiée génériquement [11], c'est à dire sur un ouvert dense de l'espace des fonctions f analytiques dans une couronne $\{1-\epsilon < |z| < 1+\epsilon\}$ et continues sur son adhérence. Pour $1 \leq k \leq n$, cette hypothèse est que tous les points critiques de $\psi_{k-1,k}$ dans Δ_k sont non dégénérés, i.e. que la dérivée seconde de $\psi_{k-1,k}$ en ces points est une forme quadratique non dégénérée. Cela entraine en particulier que les points critiques dans Δ_k sont en nombre fini. Cela entraine également que ∇_k ne s'annule pas sur $\partial\Delta_k$, car on peut montrer que les points critiques sur le bord du domaine Δ_k sont toujours dégénérés.

L'algorithme de recherche d'un minimum local procède de la manière suivante :

(0) On choisit un point initial $q_0 \in \mathcal{M}^1_n$.

(1) On intègre le champs de vecteur $-\nabla_n$ avec la condition initiale $(q_0,\psi_{n-1,n}(q_0))$.
  - On atteint un minimum local $\rightsquigarrow$ FIN.
  - On atteint le bord de $\Delta_n \rightsquigarrow$ (2).

(2) On est sur le point q_b de $\partial\Delta_n$:

$$q_b = q_u q_i,\ q_u(\alpha) = 0 \Rightarrow |\alpha| = 1,\ q_i \in \mathcal{M}^1_k$$

On intègre le champs de vecteur $-\nabla_k$ avec la condition initiale $(q_i,\psi_{k-1,k}(q_i))$.
  - On atteint un minimum d'ordre $k < n \rightsquigarrow$ (3).
  - On atteint le bord de $\Delta_k \rightsquigarrow$ (2) en remplaçant n par k.

(3) On est sur un minimum $q_m \in \mathcal{M}^1_k$ de ψ_k:
On intègre le champs de vecteur $-\nabla_{k+1}$ avec la condition initiale $((z+1)q_m,\psi_{k-1,k}(q_m))$.
  - On atteint un minimum d'ordre $k+1$
    - Si $k+1 < n \rightsquigarrow$ (3) en remplaçant k par $k+1$.
    - If $k+1 = n \rightsquigarrow$ FIN.
  - On atteint le bord de $\Delta_{k+1} \rightsquigarrow$ (2) en remplaçant n par $k+1$.

Faisons quelques remarques à propos de cet algorithme :
- Les points cols étant des points critiques instables, l'intégration du gradient, ou plutôt de son opposé, ne s'arrête que sur des minima.
- Tout au long de l'algorithme, la valeur du critère, donnée par ψ_k sur Δ_k pour $k \leq n$, ne cesse de décroitre. Ainsi, on ne rencontre jamais deux fois le même minimum. Avec l'hypothèse faite sur la non dégénérescence des points critiques, il ne peut y avoir qu'un nombre fini de minima. Cela assure la terminaison de l'algorithme.

2.3 Le problème de l'unicité d'un meilleur approximant

La description de l'algorithme précédent montre tout l'intérêt, au moins d'un point de vue numérique, d'une situation où la fonction f que l'on souhaite approcher n'engendre pour le problème associé $\mathbf{Pb}(V,m,n)$ qu'un unique point critique, à savoir le meilleur approximant lui-même. On désire donc montrer qu'une telle situation se produit quelquefois. Cependant, mettre en évidence des fonctions f exhibant cette propriété demande une analyse locale détaillée des points critiques ainsi qu'un argument de topologie différentielle permettant de déduire de cette analyse le résultat global d'unicité. Cette section ainsi que la suivante sont consacrées à cette étude.
Nous commencerons par énoncer le résultat donnant un invariant sur les points critiques pour une fonction et un degré fixés et qui permet, pour le problème de l'unicité, de déduire une conclusion globale à partir d'assertions locales. Nous utiliserons la notion d'indice de Morse d'un point critique non dégénéré, à savoir le nombre de valeurs propres négatives de la matrice hessienne en ce point. Cet indice est invariant par changement de coordonnées.

Théorème 2.3.1 [Théorème de l'indice [11, 7]] *On suppose que l'hypothèse* **(H)** *est satisfaite et que $\psi^f_{m,n}$ n'a que des points critiques non dégénérés dans Δ_n, aucun ne se trouvant sur le bord $\partial\Delta_n$. Soit $\mathcal{C}$ l'ensemble de ces points critiques et $\varepsilon(q)$ l'indice de Morse de $q \in \mathcal{C}$. Alors*

$$\sum_{q\in\mathcal{C}} (-1)^{\varepsilon(q)} = 1.$$

Ce théorème peut s'interpréter comme un analogue du théorème de Poincaré-Hopf, sachant que Δ_n est homéomorphe à la boule de dimension n. Pour une fonction spécifique f donnée, on prouve l'unicité éventuelle d'un point critique par la méthode suivante. On montre que chaque point critique q est un minimum local non dégénéré, donc possède un indice de Morse $\varepsilon(q)$ qui est nul. Le Théorème 2.3.1 permet alors de conclure. Une condition assurant qu'un point critique est un minimum local non dégénéré est donnée dans le Théorème 2.3.2. Il énonce que $L^f_m(q)$ et q doivent être "suffisamment premiers entre eux", relativement à l'erreur linéarisée correspondant à ce point critique.

Théorème 2.3.2 [5*, Theorem 2.10] *Soit $f \in H^0_{2,\mathbb{R}}(V)$ et $q \in \mathcal{M}^1_n$ un point critique irréductible de $\psi^f_{m,n}$ tel qu'il existe une relation de coprimarité*

$$bL^f_m(q) + cq = 1, \quad b,c \in H_{\infty,\mathbb{R}}(U), \tag{2.3.1}$$

avec

$$\|(fq - L^f_m(q)/z^{m-n+1})b\|_\infty < 1/2. \tag{2.3.2}$$

Alors, q est un minimum local non dégénéré de $\psi^f_{m,n}$.

La preuve consiste à se ramener au cas diagonal $m = n-1$, puis à utiliser l'expression de la variation au second ordre de $\psi^f_{n-1,n}$ dans un voisinage du point critique q, donnée, dans un système de coordonnées bien choisi, par la forme quadratique

$$H(p,p) = \left\|\frac{p}{q}\right\|^2 - < f - \frac{L^f_{n-1}(q)}{q}, 2\frac{\widetilde{q}^2 bp^2}{q^3} > , \qquad p \in \mathcal{P}_{n-1},$$

où b est le coefficient de $L^f_{n-1}(q)$ dans l'équation de Bezout (2.3.1), avec $m = n-1$. L'inégalité

$$\left| < f - \frac{L^f_{n-1}(q)}{q}, 2\frac{\widetilde{q}^2 bp^2}{q^3} > \right| \leq 2 \left\|(fq - L^f_{n-1}(q))b\right\|_\infty \left\|\frac{p}{q}\right\|_2^2,$$

montre que (2.3.2) entraine la positivité de H.
Il nous reste à décrire une méthode pratique pour trouver, lorsque cela est possible, des coefficients b et c satisfaisant (2.3.1) et (2.3.2). Cette méthode repose sur la construction d'une fraction rationnelle qui interpole f en suffisamment de points mais qui néanmoins approche cette fonction partout sur $\mathbb{T}$ moins bien que ne le fait le point critique. On note $Ord_\infty(h)$ l'ordre à l'infini d'une fonction méromorphe h, i.e. l'entier ν tel que

$$h(z) = O(z^\nu) \qquad as \quad |z| \to \infty.$$

On note aussi $Z_V(h)$ le nombre fini de zéros de h dans V, en comptant leur multiplicité.
Lemme 2.3.3 [5*, Lemma 2.11] *On suppose que l'hypothèse* **(H)** *est satisfaite. Soit $q \in \mathcal{M}^1_n$ premier avec $L^f_m(q)$. Si B/A est une fonction rationnelle irréductible à coefficients réels tel que*

$$[Z_V - Ord_\infty](z^{m-n+1}Af - B) \geq n, \tag{2.3.3}$$

et si de plus

$$|f - L^f_m(q)/z^{m-n+1}q| < |f - B/z^{m-n+1}A| \qquad sur \quad \mathbb{T}, \tag{2.3.4}$$

alors le polynôme $AL^f_m(q) - Bq$ n'a pas de zéros dans $\overline{U}$ et on peut poser dans (2.3.1),

$$b = \frac{A}{AL^f_m(q) - Bq}, \qquad c = \frac{-B}{AL^f_m(q) - Bq}.$$

Si (2.3.4) *est remplacé par l'inégalité plus forte*

$$3|f - L^f_m(q)/z^{m-n+1}q| < |f - B/z^{m-n+1}A| \qquad sur \quad \mathbb{T}, \tag{2.3.5}$$

alors (2.3.2) *est également satisfaite.*
Pour la preuve, on utilise (2.3.4) et le théorème de Rouché pour montrer que l'indice $W(AL^f_m(q) - Bq)$ de la courbe $(AL^f_m(q) - Bq)(\mathbb{T})$ autour de l'origine est égal à $W(z^{m-n+1}Af - B) + n$. L'inégalité (2.3.3) entraine que $AL^f_m(q) - Bq$ n'a pas de zéros dans $\overline{U}$. De l'inégalité (2.3.5) et de l'inégalité du triangle, on déduit facilement que (2.3.2) est vérifiée avec b choisi comme dans le théorème.
De ce qui précède, on déduit :
Théorème 2.3.4 [5*, Theorem 2.12] (critère d'unicité pour **Pb**(V,m,n)) *On suppose que l'hypothèse* **(H)** *est satisfaite, et que tout point critique q de $\psi^f_{m,n}$ vérifie*

(i) *q est irréductible,*
(ii) *il existe une fonction rationnelle B/A irréductible, à coefficients réels tel que*

$$3|f - L_m^f(q)/z^{m-n+1}q| < |f - B/z^{m-n+1}A| \qquad sur \quad \mathbb{T}, \tag{2.3.6}$$

et $B/z^{m-n+1}A$ interpole f dans V de sorte que

$$[Z_V - Ord_\infty](z^{m-n+1}Af - B) \geq n. \tag{2.3.7}$$

Alors, $\psi_{m,n}^f$ admet un unique point critique $q^* \in \mathcal{M}_n^1$, et $L_m^f(q^*)/q^*$ est l'unique minimum de $\mathbf{Pb}(V,m,n)$.

Bien entendu, il existe un résultat équivalent pour le problème $\mathbf{Pb}(U,m,n)$, voir [5*, Theorem 2.13]. On peut remarquer que le Théorème 2.3.4 assure en quelque sorte l'unicité d'un point critique lorsque la décroissance des erreurs en interpolation rationnelle est suffisamment régulière avec le degré, puisque on demande de gagner un facteur plus grand que 3 partout sur $\mathbb{T}$, lorsqu'on passe d'un interpolant de type $(m-1,n-1)$ à un interpolant de type (m,n). Typiquement, dans l'utilisation du Théorème 2.3.4, on choisira pour fonction rationnelle B/A, l'approximation de Padé en l'infini de f de type $(m-1,n-1)$, pour lequel on a égalité dans (2.3.7), et pour lequel des estimées de l'erreur sur $\mathbb{T}$ peuvent quelquefois être connues.

Nous décrivons à présent trois applications du Théorème 2.3.4, qui donnent des résultats d'unicité dans des situations concrètes. En fait, il est plus naturel d'énoncer ces résultats par rapport au problème $\mathbf{Pb}(U,m,n)$ d'approximation dans le disque unité et c'est ce que nous ferons.

2.3.1 Approximation de type fixé dans des disques décroissants vers l'origine

On considère le problème d'approximation suivant, et son cas limite lorsque le rayon r tend vers zéro.

$\mathbf{Pb}(U_r,m,n)$: *soit g analytique dans U_{r_0} telle que $g(\overline{z}) = \overline{g(z)}$, m, n, deux entiers positifs tels que $m \geq n-1$, et $0 < r < r_0$, minimiser*

$$\frac{1}{2\pi}\int_0^{2\pi} \left| g - \frac{\widetilde{p}}{\widetilde{q}} \right|^2 (re^{i\theta})d\theta$$

lorsque $\widetilde{p}$ décrit $\mathcal{P}_m$ et $\widetilde{q}$ décrit $\widetilde{\mathcal{M}}_n^r$.

Posons $f_r(z) := f(rz)$, pour toute fonction f et tout $r > 0$. Alors $\mathbf{Pb}(U_r,m,n)$ pour une fonction g est équivalent à $\mathbf{Pb}(U,m,n)$ pour la fonction g_r. Ainsi, la terminologie introduite pour $\mathbf{Pb}(U,m,n)$, en particulier la notion de point critique, se traduit, grâce à cette équivalence, directement pour $\mathbf{Pb}(U_r,m,n)$. La théorie développée précédemment permet d'obtenir le résultat suivant.

Théorème 2.3.5 [5*, Theorem 3.1] *Soit $g(z) = \sum_{l=0}^{\infty} g_l z^l$, avec $g_l \in \mathbb{R}$, et définissons les déterminants*

$$\mathcal{G}_{l,k} := \begin{vmatrix} g_l & g_{l-1} & \cdots & g_{l-k+1} \\ \vdots & \vdots & \vdots & \vdots \\ g_{l+k-1} & g_{l+k-2} & \cdots & g_l \end{vmatrix}$$

en convenant que $g_s = 0$ pour $s < 0$. Supposons que

$$\mathcal{G}_{k+m-n,k} \neq 0, \qquad 1 \leq k \leq n. \tag{2.3.8}$$

Alors $\mathbf{Pb}(U_r,m,n)$ *admet une unique paire critique d'où une unique solution, pour* r *suffisamment petit. De plus, si on dénote par* $(\widetilde{p}^*(z,r),\widetilde{q}^*(z,r))$ *cette paire et si* P_m^0/Q_n^0, $Q_n^0(0) = 1$, *désigne l'approximant de Padé de type* (m,n) *de* g, *on a*

$$\widetilde{p}^*(z,r) \to P_m^0(z) \quad \text{dans} \quad \mathcal{P}_m, \qquad \widetilde{q}^*(z,r) \to Q_n^0(z) \quad \text{dans} \quad \widetilde{\mathcal{M}}_n \quad \text{lorsque} \quad r \to 0. \tag{2.3.9}$$

La preuve consiste d'abord à montrer que tout point critique irréductible $\widetilde{p}/\widetilde{q}$ satisfait les propriétés de convergence (2.3.9). Cela s'obtient en utilisant les relations d'interpolations vérifiées par $\widetilde{p}/\widetilde{q}$, à savoir

$$g\widetilde{q} - \widetilde{p} = O(z^{m-n+1}q_{1/r^2}^2), \tag{2.3.10}$$

et le fait que les $m+n+1$ points d'interpolation concernés se trouvent dans le disque U_r qui décroît vers l'origine avec r. On utilise également la non nullité de $\mathcal{G}_{m,n}$. On montre ensuite que tous les points critiques sont irréductibles pour r assez petit. Sinon, on pourrait construire une suite de points critiques irréductibles de type $(l+m-n,l)$ pour un certain $l < n$ fixé, qui interpole g en plus de $2l+m-n+1$ points de $\overline{U}_r$. Comme ces points critiques convergent vers l'approximant de Padé P_{l+m-n}^0/Q_l^0 de g de type $(l+m-n,l)$, et que celui-ci interpole g en l'origine avec la multiplicité $2l+m-n+1$ exactement en vertu de (2.3.8), on obtient une contradiction. Enfin, l'unicité d'un point critique pour r assez petit se prouve en appliquant l'analogue du Théorème 2.3.4 pour $\mathbf{Pb}(U,m,n)$ et la fonction g_r. On choisit pour fraction rationnelle B/A l'approximant de Padé de g de type $(m-1,n-1)$, qui fournit une erreur de l'ordre $O(z^{m+n-1})$ lorsque le module de z tend vers 0.

2.3.2 Un théorème du type Montessus de Ballore dans l'espace de Hardy $H_{2,\mathbb{R}}$

Le résultat décrit ici est un analogue pour l'espace de Hardy du théorème classique de Montessus de Ballore (cf. [4]) concernant les approximants de Padé des fonctions méromorphes et leur comportement asymptotique lorsque le degré du numérateur tend vers l'infini.

Théorème 2.3.6 [5*, Theorem 4.1] *Soit* n *un entier positif et* $R > 1$. *Soit*

$$Q_n(z) = \prod_{j=1}^{n}(z - x_j), \qquad 1 < |x_j| < \rho < R$$

un polynôme réel, et $g = G/Q_n$ *où* G *est analytique dans* U_R, *et satisfait* $G(\overline{z}) = \overline{G(z)}$. *De plus, supposons que* $G(x_j) \neq 0$ *pour tout* j. *Alors, l'application* $\widetilde{\psi}_{m,n}^g : \widetilde{\Delta}_n \to \mathbb{R}$ *définie par* $\widetilde{\psi}_{m,n}^g(\widetilde{q}) = \psi_{m,n}^{g^\sigma}(q)$ *admet un unique point critique* $\widetilde{q}_n^* \in \widetilde{\Delta}_n$ *pour* m *assez grand; en particulier, le problème* $\mathbf{Pb}(U,m,n)$ *admet une unique solution* $\widetilde{p}_m^*/\widetilde{q}_n^*$. *En outre, toujours pour* m *assez grand,* $\widetilde{p}_m^*/\widetilde{q}_n^*$ *admet exactement* n *pôles dans* $\mathbb{C}$ *et*

$$\widetilde{q}_n^* \to Q_n/Q_n(0) \qquad \textit{dans } \mathcal{P}_n \qquad \textit{lorsque} \quad m \to \infty. \tag{2.3.11}$$

De plus,

$$\widetilde{p}_m^*/\widetilde{q}_n^* \to g \qquad \textit{lorsque} \quad m \to \infty, \tag{2.3.12}$$

localement uniformément dans $U_R' := U_R \setminus \cup_{j=1}^k \{x_j\}$. *Plus précisément, si* $K \subset U_R'$ *est un ensemble compact,*

$$\limsup_{m\to\infty} \|g - \widetilde{p}_m^*/\widetilde{q}_n^*\|_K^{1/m} \leq \max\{|z|,z \in K\}/R, \tag{2.3.13}$$

où $\|\cdot\|_K$ *désigne la norme du sup sur* K.

On montre d'abord que les relations (2.3.11), (2.3.12) et (2.3.13) sont vérifiées pour toute suite de points critiques irréductibles. Pour cela, on utilise le théorème de Montessus de Ballore étendu aux approximants de Padé multipoint (voir [53, Theorem 2]). Ensuite, on prouve par contradiction que tout point critique est irréductible, pour m assez grand. Finalement, le résultat d'unicité s'obtient en utilisant l'analogue du Théorème 2.3.4 pour $\mathbf{Pb}(U,m,n)$ et en choisissant pour fraction rationnelle B/A l'approximant de Padé $P^0_{m,n-1}/Q^0_{m,n-1}$ de g de type $(m,n-1)$. La majoration de l'erreur $|g - \widetilde{p}^*_m/\widetilde{q}^*_n|$ sur $\mathbb{T}$ est fournie par (2.3.13) avec $K = \overline{U}$. Pour l'erreur relative à l'approximant de Padé, on montre par contradiction la minoration suivante

$$|g - P^0_{m,n-1}/Q^0_{m,n-1}| \geq (\rho + \epsilon)^{-m} \quad \text{sur } \mathbb{T},$$

avec $\epsilon > 0$ et $\rho + \epsilon < R$ (voir [5*, Lemma 4.2]).

2.3.3 Approximation diagonale de la fonction exponentielle

Cette sous-section présente un résultat d'unicité asymptotique pour la fonction exponentielle.

Théorème 2.3.7 [5*, Theorem 5.1] *Soit $g = e^z$, il existe un entier $n_0 \geq 1$ tel que pour $n \geq n_0$, le problème $\mathbf{Pb}(U,n-1,n)$ admet une unique paire critique $(\widetilde{p}^*_{n-1},\widetilde{q}^*_n)$, en particulier un unique minimum local et global. Si l'on pose $R^*_n = \widetilde{p}^*_{n-1}/\widetilde{q}^*_n$, on a*

$$\lim_{n\to\infty} R^*_n(z) = e^z$$

localement uniformément dans $\mathbb{C}$. Plus précisément, soit $K \subset \mathbb{C}$ un ensemble compact, il existe deux constantes $\widetilde{C}_1 = \widetilde{C}_1(K)$ et $\widetilde{C}_2 = \widetilde{C}_2(K)$ telles que pour n grand et $z \in K$,

$$\widetilde{C}_1 |q^*_n(z)|^2 \leq \delta_n^{-1} |e^z - R^*_n(z)| \leq \widetilde{C}_2 |q^*_n(z)|^2,$$

où

$$\delta_n := \frac{n!(n-1)!}{(2n)!(2n-1)!}.$$

De plus, lorsque $n \to \infty$, on a

$$\widetilde{p}^*_{n-1}(z) \to e^{z/2} \quad \text{et} \quad \widetilde{q}^*_n(z) \to e^{-z/2}$$

localement uniformément dans $\mathbb{C}$. Enfin, pour toute constante $\kappa > 3$ et n assez grand, les zéros de $\widetilde{q}^*_n$ et $\widetilde{p}^*_{n-1}$ sont à moins de κ/n des zéros de Q^0_n et P^0_{n-1} respectivement, où P^0_{n-1}/Q^0_n désigne l'approximant de Padé de type $(n-1,n)$ de e^z.

Le fait que les points critiques soient irréductibles est une conséquence d'un résultat de normalité pour la fonction e^z dans le plan complexe, voir [49]. La convergence séparée des points critiques s'obtient par une technique de projection, [44], et en utilisant des résultats du même type sur les interpolants rationnels réels de e^z, voir [4*]. La partie difficile de la démonstration concerne la majoration de l'erreur pour les points critiques. On l'obtient principalement par des arguments de famille normale, le théorème de de La Vallée-Poussin pour les fonctions rationnelles et une comparaison avec les interpolants réels. La preuve du Théorème 2.3.7 ne permet pas d'estimer l'entier n_0. Cependant, les essais numériques laissent supposer que $n_0 = 1$.

2.4 Le problème de l'unicité pour la classe des fonctions de Markov

Cette section est consacrée à l'approximation rationnelle des fonctions de Markov, i.e. des fonctions du type

$$f_\mu(z) = \int \frac{d\mu(t)}{t-z}, \tag{2.4.1}$$

où μ est une mesure positive à support compact dans $\mathbb{R}$. Selon une terminologie assez répandue, on parle de fonctions de Markov, de Stieltjes ou bien de Hamburger, suivant que le support de μ est un compact de $\mathbb{R}$, une demi-droite réelle ou l'ensemble $\mathbb{R}$ tout entier. Les fonctions de ce type sont bien connues en analyse classique, puisqu'elles établissent un lien entre le problème des moments, la théorie des polynômes orthogonaux, la théorie des fractions continues et les approximants de Padé. Ici, pour que la fonction f_μ soit dans $H^0_{2,\mathbb{R}}(V)$, on suppose que

$$\operatorname{supp}(\mu) \subset (-1,1), \tag{2.4.2}$$

et on se limite au cas de l'approximation rationnelle diagonale, i.e. aux fonctions rationnelles de type $(n-1,n)$. Les différents résultats obtenus ayant trait à l'unicité d'un meilleur approximant pour une fonction de Markov, sont décrits par les quatre théorèmes suivants.

Théorème 2.4.1 [3*, Theorem 3] *Soit $[t_1,t_2] \subset (-1,1)$, le plus petit intervalle contenant $\operatorname{supp}(\mu)$. Si t_1 et t_2 satisfont*

$$(t_2 - t_1) < \sqrt{2}(1 - \max(t_1^2,t_2^2)),$$

et si $d\mu$ possède au moins n points dans son support, alors il existe un unique point critique pour le problème $\mathbf{Pb}(V,n-1,n)$, $n \geq 1$, associé à la fonction f_μ définie par (2.4.1).

Ce théorème n'est certainement pas optimal. Cependant, il n'est pas difficile de vérifier que certaines hypothèses sur $\operatorname{supp}(\mu)$ doivent être faites pour assurer l'unicité d'un meilleur approximant, voir la remarque dans l'introduction de [3*]. On peut également se poser la question de l'unicité asymptotique d'un meilleur approximant, à savoir s'il existe un entier $n_0 \in \mathbb{N}$ tel qu'il y ait un unique meilleur approximant de f_μ pour $\mathbf{Pb}(V,n-1,n)$ dès que $n \geq n_0$. La réponse est affirmative si la mesure $d\mu$ est suffisamment régulière, plus précisément dans la classe de Szegő.

Définition 2.4.2 *Une mesure positive μ sur $\mathbb{R}$ satisfait la condition de Szegő si*
(i) $\operatorname{supp}(\mu)$ *est un intervalle $[a,b] \subset \mathbb{R}$.*
(ii) Si $d\mu(t) = \mu'(t)dt + d\mu_s(t)$ est la décomposition de Radon-Nykodym , où μ_s est une mesure singulière et $\mu' \in L^1([a,b])$, alors

$$\int_a^b \frac{\log \mu'(t)}{\sqrt{(t-a)(b-t)}} dt > -\infty. \tag{2.4.3}$$

Le résultat est le suivant.

Théorème 2.4.3 [10*, Theorem 1.3] *Si μ est une mesure positive sur $[a,b] \subset (-1,1)$ satisfaisant la condition de Szegő, alors, pour n assez grand, il existe un unique point critique pour le problème $\mathbf{Pb}(V,n-1,n)$ associé à la fonction de Markov f_μ définie par (2.4.1). En particulier, le meilleur approximant est unique.*

L'hypothèse du Théorème 2.4.3 est-elle vraiment nécessaire ? Le résultat suivant montre que la mesure μ doit satisfaire au moins une hypothèse en plus de (2.4.2), pour que l'unicité asymptotique soit vérifiée.

Théorème 2.4.4 [8*, Theorem 3] *Il existe une mesure positive μ,* $\mathrm{supp}(\mu) \subset (-1,1)$*, telle que pour chaque degré impair n, le problème* $\mathbf{Pb}(V,n-1,n)$ *pour f_μ admet deux meilleurs approximants distincts.*

Remarquons d'abord qu'une fonction de Markov non nulle n'est jamais paire, ce qui exclut les exemples immédiats de non unicité décrits au début de la Section 1.2. Comme on va le voir dans la suite, on prouve le Théorème 2.4.4 en construisant une mesure discrète, i.e. supportée par un ensemble dénombrable de points. De ce théorème, on déduit par un argument simple de perturbation un résultat de non unicité dans le cas de mesures lisses, mais pour un nombre fini de degrés d'approximation seulement.

Théorème 2.4.5 [8*, Theorem 4] *Pour tout entier $n_0 \in \mathbb{N}$ impair, il existe une mesure positive μ satisfaisant la condition de Szegő, telle que pour chaque degré impair $1 \leq n \leq n_0$, le problème* $\mathbf{Pb}(V,n-1,n)$ *admet pour f_μ deux meilleurs approximants distincts.*

Une propriété de base qui gouverne le comportement des points critiques pour le problème $\mathbf{Pb}(V,n-1,n)$ associé à une fonction de Markov définie par (2.4.1) est la suivante. Les points critiques vérifient des relations d'orthogonalité à poids variables par rapport à la mesure μ. Pour un point irréductible q de degré n, ces relations s'écrivent

$$\int \frac{q}{\widetilde{q}^2}(t)P_{n-1}(t)d\mu(t) = 0, \quad \forall P_{n-1} \in \mathcal{P}_{n-1},$$

où la quantité $\widetilde{q}^2$ est considérée comme un poids variable avec le degré de q. Des relations d'orthogonalité pour les points critiques, on tire les propriétés suivantes.

Proposition 2.4.6 [3*, Proposition 5] *Soit $q \in \Delta_n$ et f une fonction de Markov avec $f \notin \mathcal{R}^0_{n-1,n}(V)$. Si q est un point critique de $\psi^f_{n-1,n}$, alors q est dans $\mathcal{M}^1_n$ et*
(i) le point critique q est irréductible.
(ii) les racines de q sont réelles.
(iii) les racines de q sont distinctes.
(iv) les racines de q sont incluses dans le plus petit segment contenant le support de la mesure $d\mu$.

D'autre part, localement, on peut calculer la variation au second ordre du critère dans un voisinage d'un point critique.

Proposition 2.4.7 [3*, Proposition 7] *Soit $q(z) = \prod_{i=1}^n (z-\xi_i)$, un point critique irréductible. La variation au second ordre du critère dans un voisinage de q est donnée par la forme quadratique*

$$\frac{1}{2}\delta_q = \sum_{i,j} \frac{X_i X_j}{1-\xi_i\xi_j} - 2\sum_i \left[\frac{q'(\xi_i)}{\tilde{q}(\xi_i)}\right]^2 Q_i X_i^2, \tag{2.4.4}$$

où les X_i, $i = 1,\dots,n$, forment un système de coordonnées convenablement choisi, et

$$Q_i := \frac{\displaystyle\int (q^2(t)/\tilde{q}^2(t))(d\mu/(1-t\xi_i))}{\displaystyle\int {q_i}^2(t)/\tilde{q}^2(t)d\mu}. \tag{2.4.5}$$

Le Théorème 2.4.1 s'obtient, d'une part par le théorème de l'indice (Théorème 2.3.1), et d'autre part, en montrant que la forme quadratique (2.4.4) est positive dès que

$$\sum_i \frac{2Q_i}{1-\xi_i^2} < 1.$$

Cette dernière assertion se déduit du théorème de Pick ([25, Theorem 2.2]) qui donne un critère d'existence d'une solution à un problème d'interpolation dans le disque U par une fonction analytique de U dans $\bar{U}$ en terme de positivité d'une certaine forme quadratique. Finalement, on utilise la majoration suivante du terme de gauche

$$\sum_i \frac{2Q_i}{1-\xi_i^2} < 2\frac{\lambda^2}{(1-\lambda^2)^2}, \tag{2.4.6}$$

où $\lambda > 0$ est tel que le support de μ soit inclus dans le segment $[-\lambda,\lambda]$. Cette inégalité s'obtient en interprétant, grâce aux relations d'orthogonalité, l'expression (2.4.5) de Q_i principalement comme la norme dans $L^2(d\mu)$ de la projection de la fonction $zq/\tilde{q}(z)$ sur l'espace de dimension n engendré par les $q_i/\tilde{q}$ où $q_i(z) = q(z)/(z-\xi_i)$, $i = 1,\ldots,n$.

L'obtention du Théorème 2.4.3 est basée sur un raffinement du critère énoncé dans le Théorème 2.3.4.

Théorème 2.4.8 [10*, Theorem 5.3] *Soit f satisfaisant l'hypothèse* **(H)**. *Soit $q \in \mathcal{M}_n^1$, un point critique irréductible de $\psi_{n-1,n}^f$, tel que $f - L_{n-1}^f(q)/q$ ne s'annule pas sur $\mathbb{T}$. Alors q est un minimum local non dégénéré de $\psi_{n-1,n}^f$ s'il existe une fonction $\beta_{n-1} \in H_{\infty,\mathbb{R}}^{n-1}(U) \cap C_{\mathbb{R}}(\mathbb{T})$ satisfaisant*

$$2 < \left| 1 - \frac{f-\beta_{n-1}}{f - L_{n-1}^f(q)/q}(z) \right|, \qquad \forall z \in \mathbb{T}, \tag{2.4.7}$$

et telle que l'indice de la courbe $(f-\beta_{n-1})(\mathbb{T})$ autour de l'origine soit inférieur où égal à $1-2n$.

Remarque 1 : si la fonction β_{n-1} est une fraction rationnelle de type $(n-2,n-1)$, la condition sur l'indice de la courbe $(f-\beta_{n-1})(\mathbb{T})$ se réinterprète comme une condition d'interpolation de la fonction f par la fonction β_{n-1} en $2n-2$ points de $\mathbb{C} \setminus \overline{U}$.

Remarque 2 : si f est une fonction de Markov, on peut montrer que l'existence d'une fonction β_{n-1} vérifiant les hypothèses du Théorème 2.4.8 est *équivalente* au fait que le point critique q est un minimum local non dégénéré.

La preuve du Théorème 2.4.8 s'obtient grâce à une expression de la variation au second ordre du critère au voisinage d'un point critique q faisant intervenir l'opérateur de Hankel

$$\Gamma_{\Phi_q} : H_{2,\mathbb{R}}^0(V) \to H_{2,\mathbb{R}}^0(V)$$

défini par

$$\Gamma_{\Phi_q}(f) = \mathbb{P}_-(\Phi_q f^\sigma), \qquad f \in H_{2,\mathbb{R}}^0(V),$$

associé au symbole

$$\Phi_q = L_{n-1}^f(q)/w_q q\tilde{q}.$$

La fonction $w_q \in H_{2,\mathbb{R}}(U)$ est déterminée par la relation de division

$$f^\sigma \widetilde{q} = w_q q^2 + \widetilde{L}_{n-1}^f(q),$$

qui est une réécriture de (2.2.6) en tenant compte du fait que le point critique irréductible q divise v_q dans $H_{2,\mathbb{R}}(U)$, voir Proposition 2.2.2. La fonction Φ_q ayant des coefficients réels, l'opérateur Γ_{Φ_q} est auto-adjoint. En outre, w_q ne s'annule pas dans $\overline{U}$ et la fonction Φ_q possède n pôles dans $\overline{U}$, à savoir les racines de q, donc l'opérateur Γ_{Φ_q} est de rang n [46, Theorem 3.11]. Son image est l'espace V_q défini par (2.2.2), avec ici $m = n-1$. Cela induit un isomorphisme

$$\Gamma_{\Phi_q} : V_q \to V_q$$

et permet de définir le pseudo-inverse que l'on notera $\Gamma^\sharp_{\Phi_q}$.

Théorème 2.4.9 [10*, Proposition 5.2] *Soit f satisfaisant l'hypothèse* **(H)**. *Soit $q \in \mathcal{M}^1_n$, un point critique irréductible de $\psi^f_{n-1,n}$, tel que $f - L^f_{n-1}(q)/q$ ne s'annule pas sur $\mathbb{T}$. La variation au second ordre de $\psi^f_{n-1,n}$ en q est donnée par la forme quadratique*

$$H(p,p) = 2 < (I - 2\Gamma^\sharp_{\Phi_q})\frac{\nu}{q},\frac{\nu}{q} > ,$$

où le polynôme $\nu \in \mathcal{P}_{n-1}$ s'obtient à partir de la perturbation $p \in \mathcal{P}_{n-1}$ par une certaine transformation linéaire dépendant uniquement du point q.

Pour que le point critique q soit un minimum local non dégénéré, il suffit, au vu du théorème précédent, que les valeurs propres de l'opérateur auto-adjoint $\Gamma^\sharp_{\Phi_q}$ soient de module strictement inférieur à 1/2. De manière équivalente, les modules des valeurs propres de l'isomorphisme Γ_{Φ_q} : $V_q \to V_q$, c'est à dire les valeurs singulières non nulles de l'opérateur de Hankel Γ_{Φ_q}, doivent être de module strictement supérieur à 2. La théorie d'Adamjan-Arov-Krein [2, 3] permet d'interpréter ces valeurs singulières comme les distances de la fonction f aux meilleures approximations méromorphes en norme L^∞. Le Théorème 2.4.8 se prouve alors en appliquant le principe de de La Vallée-Poussin aux approximants méromorphes.

Il reste à construire, pour n assez grand, une fonction β_{n-1} vérifiant les hypothèses du Théorème 2.4.8. On considère une fonction rationnelle de type $(n-2,n-1)$ qui interpole f_μ en $n-1$ points $y_{n,1}, \dots ,y_{n,n-1}$ de I^{-1}, $I = [a,b]$, avec multiplicité 2, les points d'interpolation devant être choisis de sorte que l'inégalité (2.4.7) soit vérifiée. Pour cela, on utilise l'estimée asymptotique donnée dans le théorème suivant. Pour l'énoncer, nous introduisons l'application conforme $\Psi : \overline{\mathbb{C}} \setminus I \to U$, normalisée par $\psi(\infty) = 0$, $\psi(b) = 1$ et la fonction de Szegő D_μ pour la mesure μ (voir [10*, p. 125] pour une définition précise).

Théorème 2.4.10 [64, Chapter 16] *Soit une mesure μ sur un intervalle réel $[a,b]$ satisfaisant la condition de Szegő. Pour chaque $n \in \mathbb{N}$, soit $S_n = \{z_{n,1}, \dots ,z_{n,2n}\}$ un ensemble de $2n$ points dans $\overline{\mathbb{C}} \setminus [a,b]$ invariant par conjugaison, tel que $\cup S_n$ est sans point d'accumulation sur $[a,b]$. Soit $R_{n-1,n}$ l'interpolant rationnel de type $(n-1,n)$ de f_μ en les points de S_n. Alors*

$$f_\mu(z) - R_{n-1,n}(z) = (1 + o(1))2\pi\Psi(z)D^2_\mu(z)\prod_{j=1}^{2n}\frac{\Psi(z) - \Psi(z_{n,j})}{1 - \Psi(z_{n,j})\Psi(z)}, \qquad n \to \infty, \tag{2.4.8}$$

localement uniformément dans $\overline{\mathbb{C}} \setminus [a,b]$.

On déduit de ce théorème l'égalité

$$\frac{f_\mu(z) - \beta_{n-1}(z)}{f_\mu(z) - L^f_{n-1}(q)/q(z)} = (1 + o(1))\left[\prod_{j=1}^{n-1}\frac{\Psi(z) - \Psi(y_{n,j})}{1 - \Psi(y_{n,j})\Psi(z)}\Big/\prod_{k=1}^{n}\frac{\Psi(z) - \Psi(\xi^{-1}_{n,k})}{1 - \Psi(\xi^{-1}_{n,k})\Psi(z)}\right]^2, \tag{2.4.9}$$

localement uniformément dans $\overline{\mathbb{C}} \setminus I$, les $\xi_{n,k}$ désignant les racines de q. On observe alors que le logarithme du module de l'expression à droite de (2.4.9) peut s'interpréter comme un potentiel de Green dans $\overline{\mathbb{C}} \setminus I$, associé à une mesure signée discrète dont le support est contenu dans I^{-1} et dont la masse vaut 2. On montre que toute fonction de la forme

$$h_\nu(z) := \exp\left\{\int_{I^{-1}} \log \frac{1-\Psi(t)\Psi(z)}{\Psi(z)-\Psi(t)} d\nu(t)\right\}, \tag{2.4.10}$$

où ν est une mesure signée sur I^{-1} avec $\nu(I^{-1}) = 2$, est limite localement uniforme dans $\overline{\mathbb{C}} \setminus \{I \cup I^{-1}\}$ du membre de droite de (2.4.9) pour un choix approprié des points $y_{n,j}$ relativement aux points $\xi_{n,k}$. Cela correspond à une discrétisation du potentiel $\log|h_\nu(z)|$. Finalement, on montre qu'il existe une fonction du type (2.4.10) vérifiant

$$|1 - h_\nu(z)| > 2, \qquad \forall z \in \mathbb{T}.$$

Elle est donnée explicitement comme la composée de l'application conforme qui envoie $\overline{\mathbb{C}} \setminus \{I \cup I^{-1}\}$ sur une couronne dont le cercle intérieur est $\mathbb{T}$ avec une transformation de Möbius convenablement choisi. L'expression de cette fonction sous la forme (2.4.10) s'obtient par le théorème de représentation de Riesz des fonctions superharmoniques.

En ce qui concerne le Théorème 2.4.4 de non unicité, il se prouve d'abord en considérant l'exemple élémentaire d'une fonction de Markov impaire qui est une fonction rationnelle de degré 2,

$$f_0(z) = \int \frac{d\mu_0(t)}{t-z}, \quad \mu_0 = (\delta_{z_0} + \delta_{-z_0})/2, \quad z_0 \in (-1{,}1),$$

et qui admet deux meilleurs approximants en degré 1. On note $\mathcal{M}_m$ l'ensemble des fonctions de Markov dont la mesure associée est supportée en m points de l'intervalle $(-1{,}1)$, et $\mathcal{M}_m^{sym} \subset \mathcal{M}_m$, le sous-ensemble de celles dont la mesure associée est paire. On montre alors qu'il est possible de "remonter" l'exemple précédent en une fonction de Markov de $\mathcal{M}_{2m+2}^{sym}$ approchée au degré $2m+1$ par deux meilleures approximations.

Lemme 2.4.11 [8*, Lemma 7] *Soit $f_\mu \in \mathcal{M}_{2m}^{sym}$, alors il existe un $\alpha_0 > 0$ tel que pour tout $0 < \alpha \leq \alpha_0$, la fonction*

$$f_1(z) = \int \frac{d(\mu + \alpha\mu_0)(t)}{t-z}$$

admet deux meilleurs approximants en degré $2m+1$, et de plus

$$\operatorname{dist}(f_1, \mathcal{M}_{2m+1}^{sym}) > \operatorname{dist}(f_1, \mathcal{M}_{2m+1}). \tag{2.4.11}$$

Finalement, on prouve que l'inégalité (2.4.11) reste vérifiée localement autour de f_1 pour toute fonction de Markov impaire, qui admet par conséquent également deux meilleurs approximants (de la forme $R(z)$ et $-R(-z)$ en degré $2m+1$). Cela permet de construire de manière récursive une mesure discrète

$$\mu := \sum_{j=1}^{\infty} \alpha_j \mu_{0j}, \quad \mu_{0j} = \frac{1}{2}(\delta_{-z_{0j}} + \delta_{z_{0j}}), \quad z_{0j} \in (-1{,}1),$$

avec des coefficients $\alpha_j > 0$ qui décroissent de manière convenable pour que la fonction associée f_μ admette en chaque degré impair deux meilleurs approximants distincts.

2.5 Estimées asymptotiques pour les approximants rationnels L^2 de fonctions de Markov

Dans le prolongement de la preuve du Théorème 2.4.3, il est possible de donner par des techniques similaires, issues de la théorie du potentiel, des estimées asymptotiques du taux de convergence des approximants L^2 étudiés. Les résultats peuvent s'énoncer en trois théorèmes, chacun correspondant à une hypothèse différente de régularité sur la mesure μ. Le premier résultat ne fait pas d'hypothèse sur μ autre que (2.4.2). On note S le support de μ.

Théorème 2.5.1 [11*, Theorem 2] *Soit f une fonction de Markov dont la mesure μ satisfait (2.4.2), et soit $I(\mu) = [a,b]$ le plus petit intervalle contenant S. Alors, pour toute suite de points critiques $(R_n)_n$ correspondant au $\mathbf{Pb}(V,n-1,n)$, on a*

$$\limsup_{n\to\infty} |f(z) - R_n(z)|^{1/2n}$$

$$\leq \begin{cases} \exp\Big[-p_{S,\mathbb{T}}(z)\Big] & \text{pour} \quad z \in \overline{\mathbb{D}} \setminus I(\mu) \\ \max\Big(\Big|\frac{1-az}{a-z}\Big|,\Big|\frac{1-bz}{b-z}\Big|\Big) \exp\Big[\frac{-1}{\operatorname{cap}(S,\mathbb{T})}\Big] & \text{pour} \quad z \in \overline{\mathbb{C}} \setminus \overline{\mathbb{D}} \end{cases} \tag{2.5.1}$$

localement uniformément dans $\overline{\mathbb{C}} \setminus I(\mu)$, où $p_{S,\mathbb{T}}$ désigne le potentiel du condensateur $(S,\mathbb{T})$.

Pour la définition d'un condensateur et du potentiel associé, on peut consulter [61, Appendice VIII] ou [54, Chapitre VIII].

On dira que la mesure μ est régulière, $\mu \in \mathbf{Reg}$, voir [61, Chapitre 3] si les polynômes orthogonaux pour μ ont un comportement asymptotique en racine n-ième régulier, i.e.

$$\lim_{n\to\infty} \frac{1}{n} \log |p_n(z)| = g_{\overline{\mathbb{C}}\setminus S}(z,\infty),$$

localement uniformément sur $\overline{\mathbb{C}} \setminus I(\mu)$. Ici, la fonction $g_{\overline{\mathbb{C}}\setminus S}(z,\infty)$ désigne la fonction de Green du domaine $\overline{\mathbb{C}} \setminus S$, avec pôle à l'infini.

Théorème 2.5.2 [11*, Theorems 3,4] *Si $\mu \in \mathbf{Reg}$, alors*

$$\lim_{n\to\infty} \|f - R_n\|_2^{1/2n} = \exp\Big[\frac{-1}{\operatorname{cap}(S,\mathbb{T})}\Big].$$

De plus,

$$\lim_{n\to\infty} |f(z) - R_n(z)|^{1/2n} \;=\; \exp\Big[-p_{S,S^{-1}}(z)\Big], \tag{2.5.2}$$

localement uniformément dans $\overline{\mathbb{C}} \setminus \Big(I(\mu) \cup I(\mu)^{-1}\Big)$.

Remarque La symétrie du condensateur (S,S^{-1}) par rapport à $\mathbb{T}$ entraine que

$$p_{S,\mathbb{T}}(z) = p_{S,S^{-1}}(z), \quad z \in \overline{\mathbb{D}} \setminus I(\mu),$$

de sorte que dans $\overline{\mathbb{D}} \setminus I(\mu)$, la majoration (2.5.1) coincide avec la limite (2.5.2).

Il est possible également de déterminer la distribution asymptotique des zéros, des pôles et des points d'interpolation des approximants.

Théorème 2.5.3 [11*, Theorem 6] *Soient $P_n = P(R_n)$, $Z_n = Z(R_n)$, et $I_n = Z(f - R_n)$ les ensembles des pôles de R_n, des zéros de R_n, et des points d'interpolation de f par R_n respectivement.*

Si $\mu \in$ **Reg***, alors les mesures de comptage associées convergent faiblement, lorsque n tend vers l'infini, vers les limites suivantes :*

$$\frac{1}{n}\nu_{P_n} \xrightarrow{*} \omega_{S,\mathbb{T}} = \omega_{S,S^{-1}},$$

$$\frac{1}{n}\nu_{Z_n} \xrightarrow{*} \omega_{S,\mathbb{T}} = \omega_{S,S^{-1}},$$

$$\frac{1}{2n}\nu_{I_n} \xrightarrow{*} \omega_{S^{-1},S},$$

où $\omega_{S,S^{-1}}$ et $\omega_{S^{-1},S}$ désignent les mesures d'équilibre du condensateur (S,S^{-1}) sur S et S^{-1} respectivement.

Les preuves des trois théorèmes précédents reposent, entre autres, sur une estimée en racine n-ième de l'erreur pour l'interpolation rationnelle d'une fonction de Markov, [27, 37, 61]. C'est une version affaiblie, mais suffisante ici, de l'estimée donnée dans le Théorème 2.4.10, déjà vue dans la section précédente. Elle fait intervenir le potentiel de Green du domaine $\overline{\mathbb{C}} \setminus S$ associé à une mesure α :

$$g(\overline{\mathbb{C}} \setminus S,\alpha,z) = \int g_{\overline{\mathbb{C}}\setminus S}(z,x)d\alpha(x).$$

Théorème 2.5.4 [61, Theorem 6.1.6] *Soit $\mathcal{A}$ un schéma d'interpolation, c'est à dire une matrice triangulaire de points $(a_{i,j})_{i=1,2,\dots,j=1,\dots,i}$ dans $\overline{\mathbb{C}}$, symétrique par rapport à l'axe réel (i.e. chaque ensemble $(a_{i,j})$ pour un indice i fixé est invariant par conjugaison), et tel que les points d'accumulation de $\mathcal{A}$ n'intersectent pas $I(\mu)$. On suppose que $\mathcal{A}$ admet une distribution asymptotique donnée par une mesure de probabilité α. Soit $R_{n-1,n}(f,\mathcal{A},z)$, l'interpolant de type $(n-1,n)$ de la fonction de Markov f relativement au schéma $\mathcal{A}$. Alors, on a*

$$\limsup_{n\to\infty} |f(z) - R_{n-1,n}(f,\mathcal{A},z)|^{1/2n} \leq \exp[-g(\overline{\mathbb{C}} \setminus S,\alpha,z)] \tag{2.5.3}$$

localement uniformément dans $\overline{\mathbb{C}} \setminus I(\mu)$. De plus, si la mesure μ est régulière, on a égalité dans (2.5.3) et on peut remplacer la limite supérieure par une limite simple.

Si la mesure μ vérifie la condition de Szegő, alors on peut obtenir des asymptotiques fortes (ou asymptotiques de Szegő). Pour énoncer le résultat, nous avons besoin de la moyenne géométrique de μ,

$$D(\mu) := \exp\left[\int \log\dot{\mu}(x)\, d\omega_{[a,b],\mathbb{T}}(x)\right],$$

$$d\mu(x) = \dot{\mu}(x)\, d\omega_{[a,b]}(x) + d\mu_s(x), \quad x \in [a,b] := I(\mu),$$

où $\omega_{[a,b]}$ et $\omega_{[a,b],\mathbb{T}}$ désignent la mesure d'équilibre de $[a,b]$ et la mesure d'équilibre sur $[a,b]$ du condensateur $([a,b],\mathbb{T})$ respectivement. Nous avons également besoin d'une fonction de Szegő $D(\mu,z)$ associée à la mesure μ et au condensateur $([a,b],[a,b]^{-1})$. Cette fonction, analytique dans $\overline{\mathbb{C}} \setminus ([a,b] \cup [a,b]^{-1})$, est de module 1 sur $\mathbb{T}$, et sa propriété principale est de reconstruire $\dot{\mu}$ sur $[a,b]$. Plus précisément, elle admet des limites non-tangentielles presque partout sur $[a,b] \cup [a,b]^{-1}$, et

$$|D(\mu; x \pm i\,0)|^2 = \begin{cases} \dot{\mu}(x)/D(\mu) & \text{pour presque tout} \quad x \in [a,b], \\ D(\mu)/\dot{\mu}(1/x) & \text{pour presque tout} \quad x \in [a,b]^{-1}. \end{cases}$$

L'existence d'une telle fonction s'obtient en considérant la solution d'un problème de Dirichlet généralisé dans le domaine $\overline{\mathbb{C}} \setminus ([a,b] \cup [a,b]^{-1})$, généralisé dans le sens où la fonction sur le bord n'est pas continue mais seulement intégrable. Une preuve de l'existence d'une fonction de Szegő dans une couronne est également donnée dans [56, Theorem 9].

Théorème 2.5.5 [11*, Theorems 7,8] *Soit μ une mesure de Szegő de support* $\mathrm{supp}(\mu) = [a,b] \subseteq (-1,1)$, *et soit*

$$\rho \; = \; \rho_{[a,b]} := \; \exp\Big[\frac{-1}{\mathrm{cap}([a,b],\mathbb{T})}\Big].$$

Alors, on a

$$\lim_{n\to\infty} \|f - R_n\|_2 \; \rho_{[a,b]}^{-2n} = \sqrt{\frac{8K}{\pi(1-ab)}}\, D(\mu),$$

où K désigne l'intégrale elliptique complète de première espèce de module $k = (b-a)/(1-ab)$. Soit la représentation conforme $\varphi : \overline{\mathbb{C}} \setminus ([a,b] \cup [a,b]^{-1}) \to \{\rho < |z| < 1/\rho\}$ avec $\varphi(b) = \rho = \rho_{[a,b]}$. Alors, lorsque n tend vers l'infini,

$$f(z) - R_n(z) \; = \; \big(1 \; + \; o(1)\,\big)\, \rho_{[a,b]}^{2n}\, \varphi(z)^{-2n} \;\; D(\mu)\, D(\mu; z)^2 \frac{-2}{\sqrt{(z-a)(z-b)}}, \tag{2.5.4}$$

localement uniformément dans $\overline{\mathbb{C}} \setminus ([a,b] \cup [a,b]^{-1})$.

Remarque Le potentiel du condensateur $([a,b],[a,b]^{-1})$ s'exprime à l'aide de la représentation conforme φ. On a

$$p_{[a,b],[a,b]^{-1}}(z) = \log|\varphi(z)| - \log\rho, \quad z \in \overline{\mathbb{C}}.$$

En considérant la racine $2n$-ième du module de part et d'autre de (2.5.4), on retrouve bien (2.5.2) lorsque le support de μ est l'intervalle $[a,b]$.

La preuve du Théorème 2.5.5 s'appuie essentiellement sur l'estimée donnée dans le Théorème 2.4.10, et sur les propriétés de la fonction de Szegő $D(\mu,z)$ introduite précédemment.

Chapitre 3

Interpolation rationnelle de la fonction exponentielle et zéros des polynômes exponentiels réels

Au cours du premier chapitre, nous avons étudié les approximants L^2 de fonctions dans l'espace de Hardy du disque. Ils ont la propriété d'être des interpolants rationnels (ou approximants de Padé multipoint) de la fonction qu'ils approchent, voir Proposition 2.2.2 et les commentaires qui suivent. Remarquons, en passant, que les approximants de Padé sont des cas particuliers d'interpolants rationnels, puisque dans ce cas, tous les points d'interpolation sont situés en l'origine. On s'intéresse ici aux propriétés analytiques des interpolants rationnels, i.e. principalement à leurs propriétés de convergence. Il n'existe pas de résultat général dans ce domaine, i.e. qui s'applique indépendamment de la fonction interpolée. La principale difficulté provient de l'existence de pôles pour les interpolants rationnels qui ne reflètent pas des singularités de la fonction ("spurious poles", en anglais). Ces pôles détruisent localement la convergence des interpolants. On peut trouver par exemple dans [70], un exemple d'une fonction entière dont l'ensemble des pôles de ses approximants de Padé est dense dans le plan complexe. On doit donc se restreindre à des résultats plus faibles que la convergence uniforme, telle que la convergence en mesure, ou en capacité, considérée par exemple dans le théorème de Nuttall-Pommerenke. Principalement, ce théorème affirme que les "spurious poles" ne détruisent la convergence des interpolants que sur des ensembles de petite capacité, voir [45, 50, 26, 32, 71]. En ce qui concerne la convergence uniforme, la conjecture de Baker-Gammel-Wills a été proposée dans les années 1970 :

Soit f une fonction méromorphe dans $|z| < 1$, analytique en l'origine. Il existe une suite infinie $\mathcal{I}$ d'entiers positifs tels que

$$\lim_{\substack{n\to\infty \\ n\in\mathcal{I}}} [n/n](z) = f(z),$$

localement uniformément dans $|z| < 1$, en dehors des pôles de f, où $[n/n]$ désigne l'approximant de Padé de f de type (n,n).

Cette conjecture a été infirmée récemment par D.S. Lubinsky puis par V.I. Buslaev, voir [18]. Il n'existe essentiellement que deux classes de fonctions pour lesquelles les approximants de Padé, ou plus généralement les interpolants rationnels satisfont de bonnes propriétés de convergence. La première classe est celle des fonctions de Markov, déjà étudiées dans les sections 1.4 et 1.5.

Le résultat initial, énoncé en termes de fractions continues de Chebyshev, remonte à Markov lui-même [43]. La convergence a été ensuite généralisée aux interpolants rationnels, voir [27, 61], et des asymptotiques exactes ont finalement été données dans [64, 59], c'est le Théorème 2.4.10. La deuxième classe consiste en les séries fréquences de Polya, c'est à dire les fonctions du type

$$Cz^{\lambda}e^{\gamma z}\prod_{i=1}^{\infty}(1+\alpha_i z)/\prod_{i=1}^{\infty}(1-\beta_i z),$$

où $C \geq 0$, $\lambda \in \mathbb{N}$, $\gamma \geq 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, $\sum(\alpha_i+\beta_i) < \infty$, l'élément de base de cette classe étant la fonction exponentielle. Les approximants de Padé de l'exponentielle ont été originellement considérés par Hermite, en relation avec sa preuve de la transcendance de e. Puis, Padé, reprenant la suite des travaux de Hermite, a prouvé la convergence de ces approximants vers la fonction exponentielle. On peut trouver une preuve de la convergence séparée du numérateur et du dénominateur dans [47] et une généralisation de ce résultat aux fréquences de Polya dans [1].
Nous décrivons ici des résultats qui étendent le théorème de Padé aux interpolants rationnels.

3.1 Points d'interpolation réels

On considère d'abord le cas de points d'interpolation situés sur un intervalle compact de l'axe réel.
Théorème 3.1.1 [4*, Theorem 2.1] *Soit* $B^{(m+n)} := \{x_k^{(m+n)}\}_{k=0}^{m+n}$, $m = m_\nu$, $n = n_\nu$ *une suite triangulaire de points d'interpolation réels (pas forcément distincts) contenus dans l'intervalle* $[-\rho,\rho]$ *telle que*

$$\lim_{\nu\to\infty} m_\nu + n_\nu = \infty, \tag{3.1.1}$$

et soit $R_{m,n} = P_{m,n}/Q_{m,n}$ *la fonction rationelle de type (m,n) qui interpole* e^z *en* $B^{(m+n)}$. *Alors*

$$\lim_{\nu\to\infty} R_{m_\nu,n_\nu}(z) = e^z$$

localement uniformément dans $\mathbb{C}$. *De plus, si (m,n) satisfait*

$$\lim_{\nu\to\infty}\frac{m_\nu}{n_\nu} = \lambda \quad (0 \leq \lambda \leq +\infty), \tag{3.1.2}$$

on a, lorsque $\nu \to \infty$,

$$P_{m_\nu,n_\nu}(z) \to e^{\lambda z/(1+\lambda)} \quad et \quad Q_{m_\nu,n_\nu}(z) \to e^{-z/(1+\lambda)} \tag{3.1.3}$$

localement uniformément dans $\mathbb{C}$, *où* Q_{m_ν,n_ν} *est normalisé par la condition* $Q_{m_\nu,n_\nu}(0) = 1$.
Lorsque tous les points d'interpolation sont égaux à zéro, le Théorème 3.1.1 n'est autre que le théorème classique de Padé (voir [47, p. 434]).
Théorème 3.1.2 [4*, Theorem 2.2] *Soit* $B^{(m+n)}$, $m = m_\nu$, $n = n_\nu$ *et* $R_{m,n}$ *comme dans le Théorème* 2.1, *où l'on suppose que l'on a* (3.1.1) *et* (3.1.2). *Pour* $K \subset \mathbb{C}$ *un ensemble compact, on définit*

$$c_0 = \min_{z\in K}|e^z|, \quad c_1 = \max_{z\in K}|e^z|,$$

et on pose

$$C_0 = e^{-\rho}c_0^{2/(1+\lambda)}, \quad C_1 = e^{\rho}c_1^{2/(1+\lambda)}.$$

Alors, pour tout nombre réel positif $\alpha < 1$, *il existe un entier positif* L *tel que l'interpolant rationnel* $R_{m,n}(z)$ *de* e^z *vérifie pour tout* $z \in K$,

$$\alpha C_0 \prod_{k=0}^{m+n} |z - x_k^{(m+n)}| \leq \frac{(m+n)!(m+n+1)!}{m!n!} |e^z - R_{m,n}(z)| \leq \frac{C_1}{\alpha} \prod_{k=0}^{m+n} |z - x_k^{(m+n)}|, \qquad (3.1.4)$$

dès que $m + n \geq L$.
On peut remarquer que dans le cas où le compact K ne contient qu'un point $z = \sigma + it$, et si l'on choisit $\rho = 0$, i.e. $x_k^{(m+n)} = 0$ pour tout k, m et n, on déduit de (3.1.4), lorsque α tend vers 1, que

$$|e^z - R_{m,n}(z)| = \frac{m!n!}{(m+n)!(m+n+1)!} |z|^{m+n+1} e^{2\sigma/(1+\lambda)} (1 + o(1)), \qquad (3.1.5)$$

qui redonne bien l'asymptotique correcte pour les approximants de Padé, voir [17, p. 138]. En remplaçant z par $z - \xi$, pour un certain $\xi \in \mathbb{R}$, et en multipliant par e^ξ, on obtient une estimée de l'erreur pour l'approximant de Padé de e^z en ξ. Elle montre que toute estimée uniforme par rapport à des points d'interpolation dans $[-\rho,\rho]$ doit contenir un facteur exponentiel en ρ, comme il apparait effectivement dans (3.1.4). La preuve du Théorème 3.1.1 repose principalement sur le lemme

Lemme 3.1.3 [4*, Lemma 2.4] *Soit* $s_n(z)$ *une suite de polynômes unitaires avec* $\deg s_n = n$ *et supposons que les racines de* s_n *soient toutes dans un disque* $|z| \leq \rho$. *Soit* $q_{m,n}(z)$ *le polynôme unitaire de degré* n *tel que*

$$(I + D)^{m+1} q_{m,n} = s_n, \quad m \in \mathbb{N},$$

où D *désigne la différentiation. Alors les assertions suivantes sont vérifiées :*
(i) *toutes les racines de* $q_{m,n}$ *tendent vers l'infini quand* $m + n$ *tend vers l'infini,*
(ii) *pour* $m + n$ *assez grand, (de sorte que* $q_{m,n}(0) \neq 0$, *voir* (i)*), la famille* $\{q_{m,n}/q_{m,n}(0)\}$ *est normale dans le plan complexe,*
(iii) *si* $m = m_\nu$, $n = n_\nu$ *satisfont* (3.1.1) *et* (3.1.2), *alors*

$$\lim_{\nu \to \infty} q_{m,n}(z)/q_{m,n}(0) = e^{-z/(1+\lambda)} \qquad (3.1.6)$$

localement uniformément dans le plan complexe.
On fait alors l'observation suivante. Soit

$$q_{m+n+1}(z) = \prod_{k=0}^{m+n} (z - x_k^{(m+n)}),$$

et soit l'équation d'interpolation

$$\widetilde{Q}_{m,n}(z) e^z - \widetilde{P}_{m,n}(z) = O(q_{m+n+1}(z)), \qquad (3.1.7)$$

où l'on suppose le polynôme $\widetilde{Q}_{m,n}$ unitaire. Alors, en dérivant $(m+1)$ fois cette identité, on déduit par le théorème de Rolle qu'il existe un polynôme unitaire π_n de degré n avec toutes ses racines dans $[-\rho,\rho]$ tel que

$$(I + D)^{m+1} \widetilde{Q}_{m,n} = \pi_n. \qquad (3.1.8)$$

C'est l'équation considérée dans le Lemme 3.1.3. La preuve du Lemme 3.1.3 repose elle-même sur deux résultats en géométrie des polynômes.

Théorème (Walsh, [42], Theorem 18.1, p.81): *Soit*

$$f(z)=\sum_{j=0}^{n}a_jz^j,\quad g(z)=\sum_{j=0}^{n}b_jz^j=b_n\prod_{j=1}^{n}(z-\beta_j),$$

et

$$h(z)=\sum_{j=0}^{n}(n-j)!b_{n-j}f^{(j)}(z).$$

Si tous les zéros de $f(z)$ sont dans un disque A, alors tous les zéros de $h(z)$ sont dans un ensemble constitué de n régions circulaires obtenues en translatant A par les vecteurs β_j.

Théorème (Szegő, [42], Theorem 16.1, p.65): *Soit*

$$f(z)=\sum_{k=0}^{n}\binom{n}{k}A_kz^k,\quad g(z)=\sum_{k=0}^{n}\binom{n}{k}B_kz^k$$

et

$$h(z)=\sum_{k=0}^{n}\binom{n}{k}A_kB_kz^k.$$

Si tous les zéros de $f(z)$ sont dans un disque A, alors tout zéro γ de $h(z)$ est de la forme $\gamma=-\alpha\beta$ où α est un point dans A et β un zéro de $g(z)$.

L'observation principale consiste à remarquer que le théorème de Walsh s'applique avec $f(z)=s_n(z)$, $g(z)=\frac{1}{n!}q^0_{m,n}(z)$ et $h(z)=q_{m,n}(z)$, où $q^0_{m,n}$ désigne le dénominateur unitaire de l'approximant de Padé de type (m,n). Comme les zéros de $q^0_{m,n}$ tendent vers l'infini avec $m+n$, et que les zéros de s_n restent dans l'intervalle $[-\rho,\rho]$, cela entraine que les zéros de $q_{m,n}$ tendent également vers l'infini avec $m+n$. Le théorème de Szegő est utilisé pour prouver que la famille $\{q_{m,n}/q_{m,n}(0)\}$ est normale dans $\mathbb{C}$.

L'estimée (3.1.4) dans le Théorème 3.1.2 s'obtient par une technique initialement utilisée dans [14].On décrit seulement le cas d'approximants diagonaux de type (l,l). Pour un entier positif n fixé, on considère le schéma d'interpolation C_n dont la $(2n+2k)$-ième ligne est obtenue en rajoutant à l'ensemble $B^{(2n)}=\{x_i^{(2n)}\}_{i=0}^{2n}$ le point zéro avec multiplicité $2k$ pour $k\geq 0$. En appliquant le Théorème 3.1.1 avec le schéma C_n, on obtient

$$e^z-R_n^B(z)=\sum_{k=0}^{\infty}\left[R^{C_n}_{n+k+1}(z)-R^{C_n}_{n+k}(z)\right], \tag{3.1.9}$$

et

$$D_k(z):=R^{C_n}_{n+k+1}(z)-R^{C_n}_{n+k}(z)=\frac{\beta_{k+1}z^{2k}\prod_{i=0}^{2n}(z-x_i^{(2n)})}{\widetilde{Q}^{C_n}_{n+k}(z)\widetilde{Q}^{C_n}_{n+k+1}(z)}, \tag{3.1.10}$$

où β_{k+1} est le coefficient dominant de $\widetilde{P}^{C_n}_{n+k+1}\widetilde{Q}^{C_n}_{n+k}-\widetilde{P}^{C_n}_{n+k}\widetilde{Q}^{C_n}_{n+k+1}$. On utilise alors des inégalités prouvées par ailleurs pour obtenir la majoration, pour tout $0<\delta<1$,

$$|D_k(z)|\leq 2\frac{e^\rho c_1}{\delta}\frac{(n+k+1)!(n+k)!}{(2n+2k+2)!(2n+2k)!}|z|^{2k}\prod_{i=0}^{2n}|z-x_i^{(2n)}|,$$

où n est assez grand et k quelconque. Le fait que le quotient $|D_{k+1}(z)/D_k(z)|$ soit majoré par un terme d'ordre $(n+k)^{-2}$ termine la preuve de la majoration dans (3.1.4). La minoration s'obtient de manière similaire.

3.2 Points d'interpolation complexes

Il semble naturel de vouloir étendre les résultats précédents au cas d'un ensemble de points d'interpolation dans le plan complexe, symétrique par rapport à l'axe réel. La principale difficulté de cette généralisation, si l'on veut utiliser la même approche que celle décrite après l'énoncé du Lemme 3.1.3, est l'obtention d'un théorème de Rolle complexe. Nous donnons une version d'un tel théorème, valable pour les polynômes exponentiels réels, i.e. pour les expressions du type

$$g(z) = \sum_{j=1}^{n} p_j(z) e^{\alpha_j z}, \tag{3.2.1}$$

avec des fréquences réelles

$$0 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_n,$$

et des coefficients polynômiaux réels p_j. Soit

$$m_j = \deg p_j, \qquad j = 1,\ldots,n,$$

on définit le degré de g,

$$\deg g := n - 1 + \sum_{j=1}^{n} m_j,$$

et son diamètre, le réel positif $\alpha_n - \alpha_1$. On note aussi $\mathcal{E}_{\rho,p}$ l'ellipse

$$\mathcal{E}_{\rho,p} := \{x + iy : \frac{x^2}{p} + y^2 = \rho^2\},$$

et $\mathcal{F}_{\rho,p}$, l'intérieur de cette ellipse. Pour M et N, deux points du plan complexe, on pose

$$Q(M,N) := \frac{d^2(N,M_0) - d^2(M,M_0)}{d^2(N,M) d^2(N,\overline{M})}, \tag{3.2.2}$$

où M_0 désigne la projection de M sur l'axe réel $\mathcal{O}x$, $\overline{M}$, le point conjugué de M, et $d(.\ ,.)$, la distance euclidienne usuelle entre deux points du plan complexe.

Théorème 3.2.1 (Un théorème de Rolle complexe) [12*, Theorem 2.1] *Soit g un polynôme exponentiel réel de diamètre inférieur ou égal à un nombre réel positif fixé α et soit ρ et ρ' deux nombres réels positifs tels que $0 < \rho < \rho' < 2\pi/\alpha$. Finalement, soit $p \geq 1$ un nombre entier. On suppose que g a $L \geq \deg g$ zéros dans le domaine $\mathcal{F}_{\rho,p}$. Alors la dérivée de g a au moins $L-1$ zéros dans le domaine $\mathcal{F}_{\rho',p+1}$ dès que*

$$\deg g \geq \frac{1}{\rho'^2}\left(1 - \frac{\alpha\rho'}{2}\cot\frac{\alpha\rho'}{2}\right) \max_{\substack{M\in\mathcal{F}_{\rho,p}\\ N\in\mathcal{E}_{\rho',p+1}}} Q(M,N)^{-1}. \tag{3.2.3}$$

On peut donner une borne supérieure explicite pour le maximum dans le membre de droite de (3.2.3) en fonction de ρ, ρ' et p. Notons aussi qu'il existe une version du théorème précédent pour des domaines $\mathcal{E}_{\rho,p}$ de hauteur ρ arbitraire. Dans ce cas, le résultat est moins précis car la minoration sur le degré à partir duquel le théorème de Rolle s'applique n'est plus explicite.

La preuve du Théorème 3.2.1 repose d'abord sur le fait que les zéros de grand module d'un polynôme exponentiel g se distribuent de manière régulière le long de l'axe imaginaire, plus précisement,

leur partie imaginaire croît de manière arithmétique (théorème de Polya-Dickson). Cela a pour conséquence que dans la factorisation de Hadamard de g, il n'est pas nécessaire d'inclure des facteurs exponentiels pour assurer la convergence du produit infini. La dérivée logarithmique g'/g se calcule donc principalement comme celle d'un polynôme et cela permet d'adapter la preuve classique de Jensen, voir [42, Theorem 7.1, p. 26], au cas des polynômes exponentiels réels. On considère un contour $\mathcal{E}_{\rho',p+1}$, $\rho' > \rho$, qui contient tous les cercles de Jensen associés aux zéros de g dans $\mathcal{E}_{\rho,p}$. On montre que si le nombre de zéros z_j de g dans $\mathcal{E}_{\rho,p}$ est assez grand, alors la somme des parties imaginaires des termes $1/(z - z_j) + 1/(z - \overline{z}_j)$ correspondants à ces zéros devient prépondérante par rapport à celle relative aux zéros hors de $\mathcal{E}_{\rho,p}$. On en déduit que l'image de l'ellipse $\mathcal{E}_{\rho',p+1}$ par la fonction g'/g ne peut faire qu'une fois le tour de l'origine et par le principe de l'argument, la différence entre le nombre de zéros de g et de g' dans $\mathcal{E}_{\rho',p+1}$ ne peut pas excéder un.
Dans le problème de l'interpolation de la fonction exponentielle par une fonction rationnelle de type (m,n), il faut utiliser le théorème de Rolle $m+1$ fois pour faire disparaitre le numérateur, voir (3.1.8). Il nous reste donc à donner une version du théorème de Rolle complexe pour les dérivées d'ordre supérieur d'un polynôme exponentiel réel g. On considère plutôt la dérivée du quotient de $g(z)$ par $\exp(\alpha z)$, où α est la plus petite fréquence de g, que l'on note $g^{\{1\}}$. On note de même $g^{\{r\}}$ les itérés de cette transformation. Contrairement à la dérivée, le degré de $g^{\{1\}}$ est toujours inférieur de 1 au degré de g.

Théorème 3.2.2 [12*, Theorem 3.2] *Soient ϵ, α, et ρ trois nombres réels positifs tels que $\rho < 2\pi/\alpha$. Soit $(g_\nu)_{\nu\in\mathbb{N}}$, une suite de polynômes exponentiels réels de diamètre inférieur ou égal à α telle que*

$$\lim_{\nu\to\infty} \deg g_\nu = \infty.$$

Soit r_ν une suite d'entiers tels que $1 \leq r_\nu \leq \deg g_\nu - 1$, satisfaisant

$$4\left(1 - \frac{\alpha\rho}{2}\cot\frac{\alpha\rho}{2}\right) < \inf_{\nu\in\mathbb{N}} \frac{\deg g_\nu + 1 - r_\nu}{r_\nu + 1}. \tag{3.2.4}$$

Pour chaque ν, on suppose que g_ν possède $L_\nu \geq \deg g_\nu$ zéros dans le domaine $\mathcal{F}_{\rho,1}$. Alors, il existe un entier positif C qui ne dépend que de ϵ, α, et ρ tel que le polynôme exponentiel $g_\nu^{\{r_\nu\}}$ possède au moins $L_\nu - r_\nu$ zéros à l'intérieur de $\mathcal{E}_{\rho+\epsilon,r_\nu+1}$, dès que le degré de g_ν est plus grand que C.
De plus, si

$$\lim_{\nu\to\infty} r_\nu = \infty \quad \text{et} \quad \lim_{\nu\to\infty} \frac{\deg g_\nu}{r_\nu} = \mu \quad (1 \leq \mu \leq \infty), \tag{3.2.5}$$

l'hypothèse (3.2.4) *peut être remplacée par*

$$4\left(1 - \frac{\alpha\rho}{2}\cot\frac{\alpha\rho}{2}\right) < \mu - 1. \tag{3.2.6}$$

La preuve consiste à appliquer le Théorème 3.2.1 plusieurs fois. La difficulté vient de ce que l'on doit considérer une suite d'ellipses concentriques dont les distances mutuelles tendent vers zéro, alors que la longueur de leur demi-axe parallèle à Ox tend vers l'infini. On vérifie que l'inégalité (3.2.4) (ou (3.2.6)) entraine une inégalité de type (3.2.3) à chaque utilisation du Théorème 3.2.1.
Le Théorème 3.2.2 permet d'appliquer pour le problème de l'interpolation rationnelle complexe de e^z la même approche que celle utilisée pour le problème de l'interpolation réelle. Il est à noter ici un point essentiel pour la preuve, à savoir que, d'après le Théorème 3.2.2, les zéros de $g_\nu^{(r_\nu)}$, $g_\nu(z) = Q_{m_\nu,n_\nu}(z)e^z - P_{m_\nu,n_\nu}(z)$, $r_\nu = m_\nu + 1$, restent dans une ellipse de demi-axe $\sqrt{m_\nu}$, lorsque m_ν tend vers l'infini avec ν. Cela permet d'utiliser le théorème de Walsh pour montrer que les

pôles des interpolants complexes tendent encore vers l'infini. En effet, rappelons que les pôles des approximants de Padé tendent vers l'infini comme $m_\nu + n_\nu \gg \sqrt{m_\nu}$. On obtient ainsi le résultat suivant :

Théorème 3.2.3 [13*, Theorem 2.1] *Soit* $B^{(m+n)} := \{z_k^{(m+n)}\}_{k=0}^{m+n}$, $m = m_\nu$, $n = n_\nu$ *un tableau triangulaire de points d'interpolation complexes, pas nécessairement distincts, symétriques i.e.* $\overline{B^{(m+n)}} = B^{(m+n)}$, *inclus dans le domaine* $\mathcal{F}_{\rho,1}$, $0 < \rho < 2\pi$, *et tel que*

$$\lim_{\nu\to\infty} m_\nu + n_\nu = \infty. \tag{3.2.7}$$

Soit $R_{m,n} = P_{m,n}/Q_{m,n}$ *un approximant de Padé multipoint (i.e. un interpolant rationnel linéarisé de type* (m,n) *de* e^z *en les points de* $B^{(m+n)}$ *tel que*

$$Q_{m,n}(z)e^z - P_{m,n}(z) = O(\omega_{m+n+1}(z)), \tag{3.2.8}$$

avec

$$\omega_{m+n+1}(z) = \prod_{k=0}^{m+n} (z - z_k^{(m+n)}).$$

Supposons que la suite (m_ν,n_ν) *satisfait*

$$4\left(1 - \frac{\rho}{2}\cot\frac{\rho}{2}\right) < \inf_{\nu\geq\nu_0} \max\left(\frac{n_\nu+1}{m_\nu+2},\frac{m_\nu+1}{n_\nu+2}\right), \tag{3.2.9}$$

où ν_0 *est un entier positif. Alors, les quatre assertions suivantes sont vérifiées :*
(i) *Pour* ν *grand, les polynômes* $P_{m,n}$ *et* $Q_{m,n}$ *tels que* (3.2.8) *sont de degré* m *et* n *respectivement.*
(ii) *Tous les zéros et les pôles de* $R_{m,n}$ *tendent vers l'infini, lorsque* ν *devient grand. En particulier, pour* ν *grand, aucun pôle de* $R_{m,n}$ *ne se trouve dans le domaine* $\mathcal{F}_{\rho,1}$ *contenant les points d'interpolation. Donc, divisant* (3.2.8) *par* $Q_{m,n}$, *on obtient que* $R_{m,n}$ *est un interpolant rationnel de* e^z *en les points de* $B^{(m+n)}$ *satisfaisant*

$$e^z - R_{m,n}(z) = O(\omega_{m+n+1}(z)).$$

(iii) *Lorsque* $\nu \to \infty$, *on a, localement uniformément dans* $\mathbb{C}$,

$$R_{m_\nu,n_\nu}(z) \to e^z. \tag{3.2.10}$$

(iv) *De plus, si* (m,n) *satisfait*

$$\lim_{\nu\to\infty} \frac{m_\nu}{n_\nu} = \lambda \quad (0 \leq \lambda \leq +\infty), \tag{3.2.11}$$

alors le numérateur et le dénominateur convergent séparément, c'est à dire que l'on a, lorsque $\nu \to \infty$,

$$P_{m_\nu,n_\nu}(z) \to e^{\lambda z/(1+\lambda)} \quad \text{et} \quad Q_{m_\nu,n_\nu}(z) \to e^{-z/(1+\lambda)} \tag{3.2.12}$$

localement uniformément dans $\mathbb{C}$, *où* Q_{m_ν,n_ν} *est normalisé par* $Q_{m_\nu,n_\nu}(0) = 1$.
Si (3.2.11) *est vérifiée, la condition* (3.2.9) *se réduit à*

$$4\left(1 - \frac{\rho}{2}\cot\frac{\rho}{2}\right) < \max(\lambda,1/\lambda). \tag{3.2.13}$$

A titre d'exemple, lorsqu'on considère des interpolants diagonaux, c'est à dire de type (n,n), on a $\lambda = 1$ et la condition (3.2.13) est vérifiée si $\rho < 1{,}689$... On prouve également des estimées de l'erreur similaires à celles obtenues dans le Théorème 3.1.2.
Pour clore ce chapitre, nous ferons une dernière remarque concernant le nombre de zéros maximal que peut avoir un polynôme exponentiel dans un compact du plan complexe, par exemple un disque. Si $N(g,z_0,r)$ désigne le nombre de zéros du polynôme exponentiel

$$g(z) = \sum_{j=1}^{n} q_j(z) e^{\omega_j z}, \quad \omega_j \in \mathbb{C},$$

(où les coefficients q_j sont des polynômes complexes) dans le disque fermé de rayon r, centré en z_0. Alors, on a, voir Tijdeman [63], Waldschmidt [69], Voorhoeve [68],

$$N(g,z_0,r) \leq 4\Omega r/\pi + 2 \deg g, \tag{3.2.14}$$

où

$$\Omega = \max\{|\omega_j|, j = 1, \dots ,n\}.$$

Pour un polynôme exponentiel réel de diamètre α, une majoration est plus facile à calculer. On a

$$N(g,z_0,r) \leq \alpha r/\pi + \deg g. \tag{3.2.15}$$

On peut se demander comment se comporte cette majoration asymptotiquement, i.e. lorsque z_0 tend vers l'infini ou lorsque le degré de g tend vers l'infini. Le théorème de Polya-Dickson, qui donne la distribution asymptotique des zéros de grand module, montre qu'en dehors d'un compact autour de l'origine, le membre de droite de (3.2.15) se simplifie en $\alpha r/\pi$. Par contre, si z_0 est fixé, tandis que $\deg g$ tend vers l'infini, une application du Théorème 3.2.2 (lorsqu'elle est possible, i.e. lorsque le degré de l'un des coefficients polynômiaux de g est plus grand que la somme des degrés des autres coefficients) laisse supposer que le membre de droite de (3.2.15) se simplifie en $\deg g$. Cela peut suggérer la question suivante.

Question : *Soit un disque fermé, centré à l'origine, de rayon r, et un diamètre α. Existe-t-il un entier C ne dépendant que de r et α tel que, pour tout polynôme exponentiel réel g de diamètre inférieur ou égal à α, on ait*

$$N(g,0,r) \leq \deg g,$$

dès que le degré de g est plus grand que C ?

Chapitre 4

Asymptotique des approximants de Hermite-Padé et applications

4.1 Asymptotique des approximants de Hermite-Padé

Dans le Chapitre 3, nous avons vu comment le théorème de Padé sur les approximants de la fonction exponentielle pouvait se généraliser au cas de certains interpolants rationnels. Il existe un autre type d'approximants, les approximants de Hermite-Padé pour lesquels on peut également obtenir une version de ce théorème. Nous présentons dans ce chapitre les différents résultats concernant cette étude. Rappelons d'abord la définition des approximants de Hermite-Padé, dits latins, ou de type I, selon une terminologie due à Mahler. Soit $(f_0, \dots, f_m)$ un vecteur de $m+1$ fonctions analytiques à l'origine. Soit le multi-indice $n = (n_0, \dots, n_m) \in \mathbb{N}^{m+1}$, l'approximant de Hermite–Padé de degré n, de type I, est le vecteur non nul de polynômes

$$(A_0, \dots, A_m) \in \prod_{j=0}^{m} \mathcal{P}_{n_j - 1}$$

tel que

$$\sum_{i=0}^{m} A_i(z) f_i(z) = O(z^{|n|-1}) \qquad \text{lorsque} \quad z \to 0, \tag{4.1.1}$$

où $|n| = \sum_{i=0}^{m} n_i$. Mentionnons seulement qu'il existe un autre type d'approximants de Hermite-Padé, dit germain ou de type II, consistant en un vecteur d'approximants simultanés. Ce sont ces derniers qui ont été utilisés par Hermite pour prouver la transcendance du nombre e. Nous ne considérons ici que les approximants de type I, avec le choix des fonctions $(1, e^z, \dots, e^{mz})$. Une propriété importante de ce cas est qu'il constitue un système parfait, voir [40]. Cela signifie que pour tout multi-indice $n = (n_0, \dots, n_m)$, une solution $A_0, \dots, A_m$ de (4.1.1) satisfait

$$\deg A_j = n_j - 1, \quad j = 0, \dots, m.$$

On en déduit que la solution de (4.1.1) est unique, à un facteur constant près. Le choix d'un vecteur d'exponentielles $(1, e^z, \dots, e^{mz})$ est un des rares cas où il existe des expressions intégrales pour les solutions de (4.1.1). En effet, on montre facilement que les formules

$$A_p(z) = \frac{1}{2i\pi} \int_{C_0} \frac{e^{\zeta z} d\zeta}{\prod_{l=0}^{m} (\zeta + p - l)^{n_l}}, \qquad 0 \le p \le m, \tag{4.1.2}$$

où C_0 est un cercle centré à l'origine et de rayon inférieur à 1, définissent des polynômes de degrés $n_p - 1$ satisfaisant (4.1.1).
Nous donnons à présent un résultat de convergence des coefficients $A_0, \ldots, A_m$ qui étend le théorème de Padé aux approximants de Hermite-Padé.
Théorème 4.1.1 [6*, Theorem 2.1] *Soit $A_0, \ldots, A_m$ les approximants de Hermite–Padé de la fonction exponentielle donnés par (4.1.2), avec $n_l = n$, $0 \le l \le m$, de degré inférieur à n. Soit $\{-p - \eta_p, 0 < \eta_p < 1\}_{p=0}^{m-1}$ l'ensemble des m points critiques, c'est à dire les m racines de la dérivée du polynôme de Pochhammer*

$$(z)_m = z(z+1)\cdots(z+m-1).$$

Alors, il existe des constantes non nulles (que l'on peut calculer explicitement) $\mu_{p,n}$, $0 \le p < m/2$, telles que, lorsque $n \to \infty$,

$$A_p(0) \sim (-1)^{mn}\mu_{p,n},\ A_{m-p}(0) \sim (-1)^{n-1}\mu_{p,n},\ 0 \le p < m/2.$$

Par conséquent, pour n grand, on peut définir $\widetilde{A}_p$, le polynôme obtenu en divisant A_p par son coefficient constant. Alors,

$$\widetilde{A}_p(z) \to e^{\eta_p z},\ \widetilde{A}_{m-p}(z) \to e^{-\eta_p z},\ 0 \le p < m/2,$$

localement uniformément dans $\mathbb{C}$. Si m est pair, on définit $A_{m/2}^{(1)}$ (resp. $A_{m/2}^{(2)}$) la sous-suite des polynômes $A_{m/2}$ correspondant à des indices n pairs (resp. impairs). Alors, $A_{m/2}^{(1)}$ est un polynôme impair, $A_{m/2}^{(2)}$ un polynôme pair, et on peut prouver comme pour les autres coefficients, la convergence de ces sous-suites vers certaines exponentielles localement uniformément dans $\mathbb{C}$.
La preuve consiste à utiliser la méthode du col pour estimer les expressions intégrales (4.1.2) des polynômes A_p. En utilisant de la même manière l'expression intégrale du reste R,

$$R(z) = \frac{1}{2i\pi}\int_{C_\infty} \frac{e^{\zeta z} d\zeta}{\prod_{l=0}^{m}(\zeta - l)^{n_l}}, \tag{4.1.3}$$

où C_∞ est un cercle centré en l'origine de rayon plus grand que m, on peut aussi montrer que dans le cas diagonal,

$$R(z) \sim \frac{z^{mn+n-1}e^{mz/2}}{(mn+n-1)!},$$

localement uniformément dans $\mathbb{C}$, lorsque $n \to \infty$. Cette estimée permet de résoudre le problème de minimisation suivant.

Problème : *soit une valeur de n donnée, minimiser uniformément dans le disque $\{z \in \mathbb{C},\ |z| \le \rho\}$,*

$$r_n(z) = \sum_{p=0}^{m} a_{p,n}(z)e^{pz},$$

où les coefficients $a_{p,n}$ sont des polynômes de degrés inférieurs ou égaux à n, et $a_{m,n}$ est unitaire.

Pour résoudre ce problème, on considère les translatés des approximants de Hermite–Padé suivants :

$$a_{p,n}^*(z) = n!(m!)^{n+1}A_p\left(z - \frac{m\rho^2}{2(mn+m+n)}\right), \quad 0 \le p \le m,$$

$$r_n^*(z) = n!(m!)^{n+1} R\left(z - \frac{m\rho^2}{2(mn+m+n)}\right),$$

où le facteur $n!(m!)^{n+1}$ provient de la normalisation par le coefficient dominant de A_m, de sorte que $a_{m,n}^*$ soit unitaire.
Soit $\|\cdot\|_\rho$, la norme du Sup sur $\{z \in \mathbb{C},\ |z| \leq \rho\}$, on a
Théorème 4.1.2 [6*, Theorem 3.2]
(i) $\|\sum_{p=0}^m a_{p,n}^*(z)e^{pz}\|_\rho \sim \frac{n!(m!)^{n+1}}{(mn+m+n)!}\rho^{mn+m+n}$.
(ii) *Soit* $\rho < \pi/m$ *et*

$$w_n^* = \min_{\substack{\deg a_{p,n} \leq n,\ 0 \leq p \leq m \\ a_{m,n} = z^n + \cdots}} \|\sum_{p=0}^m a_{p,n}(z)e^{pz}\|_\rho.$$

Alors

$$w_n^* \sim \frac{n!(m!)^{n+1}}{(mn+m+n)!}\rho^{mn+m+n}.$$

La preuve reprend celles utilisées dans [16] et [15], où le cas des approximants de Padé, respectivement le cas des approximants de Padé quadratiques, avaient été considérés. On suppose qu'il existe des polynômes $\tilde{a}_{p,n}$, $0 \leq p \leq m$, avec $\deg \tilde{a}_{p,n} \leq n$ et $\tilde{a}_{m,n}$ unitaire tels que pour $|z| = \rho$,

$$|\sum_{p=0}^m \tilde{a}_{p,n}(z)e^{pz}| < |\sum_{p=0}^m a_{p,n}^*(z)e^{pz}|.$$

Alors, par le théorème de Rouché, la différence

$$\sum_{p=0}^m (\tilde{a}_{p,n} - a_{p,n}^*)e^{pz} \tag{4.1.4}$$

possède au moins $mn+m+n$ zéros dans $\{|z| \leq \rho\}$ pour n grand. Mais cela contredit le fait qu'une expression non nulle $\sum_{p=0}^m a_{p,n}(z)e^{pz}$ avec $\sum_{p=0}^m \deg a_{p,n} = h$ peut avoir au plus $h+m+m\rho/\pi$ zéros dans le disque $\{|z| \leq \rho\}$, voir [49, Problem 206.2]. En effet, la somme des degrés des polynômes $\tilde{a}_{p,n} - a_{p,n}^*$ est au plus $mn+n-1$. Donc, l'expression (4.1.4) peut avoir au plus $mn+m+n-1$ zéros dans $\{|z| \leq \rho\}$ puisque l'on suppose $m\rho/\pi < 1$. Ceci achève la preuve du Théorème 4.1.2.
Il serait très intéressant de s'affranchir de la condition $\rho < \pi/m$ limitant le rayon du disque sur lequel on fait la minimisation.

4.2 Une application à la théorie des nombres

Comme il est bien connu, les approximants de Hermite-Padé sont d'usage fréquent en théorie analytique des nombres. Nous décrivons à présent une application de ces approximants, en vue d'améliorer la minoration due à Mahler, voir [39],

$$|e^g - a| > g^{-33g}, \qquad a \in \mathbb{N},$$

où g est un entier positif suffisamment grand. La méthode de preuve, utilisée par Mahler, est la suivante. On considère les approximants de Hermite-Padé d'un vecteur d'exponentielles,

$$R(z) := \sum_{p=0}^m A_p(z)e^{pz}, \tag{4.2.1}$$

tels que le degré de chaque coefficient A_p est inférieur à n. Les coefficients des polynômes A_p sont des nombres rationnels. En multipliant par un commun multiple des dénominateurs de ces nombres, on obtient une approximation

$$\widetilde{R}(z) = \sum_{p=0}^{m} \widetilde{A}_p(z) e^{pz}, \tag{4.2.2}$$

où tous les polynômes $\widetilde{A}_p$ sont à coefficients entiers. On évalue (4.2.2) pour $z = g$, et l'on fait dépendre les paramètres m et n de cet entier g,

$$n = [g/\beta], \qquad m = [\alpha \log n], \tag{4.2.3}$$

où $[x]$ désigne la partie entière de x. Ici, α et β sont des réels positifs à ajuster de manière optimale. On écrit $e^g = a + \delta$, où a est l'entier le plus proche de e^g, $|\delta| < 1/2$. A l'aide des relations (4.2.3), on vérifie aisément que pour $1 \leq p \leq m$,

$$(a+\delta)^p = a^p + p\delta a^{p-1}(1 + o(1)) \qquad \text{lorsque} \quad g \to +\infty,$$

où la convergence du terme en $o(1)$ vers zéro est uniforme par rapport à p. Alors l'équation (4.2.2) devient

$$\widetilde{R}(g) = \sum_{p=0}^{m} \widetilde{A}_p(g) a^p + \sum_{p=1}^{m} \widetilde{A}_p(g) p\delta a^{p-1}(1 + o(1)). \tag{4.2.4}$$

Si α et β sont choisis de manière convenable, $\widetilde{R}(g)$ tend vers zéro, lorsque g devient grand. On suppose que $|\widetilde{R}(g)| \leq 1/2$. Comme la première somme à droite de (4.2.4) est un entier, cela entraine que

$$|\widetilde{R}(g)| \leq |\sum_{p=1}^{m} \widetilde{A}_p(g) p\delta a^{p-1}(1 + o(1))| \leq (1 + o(1))|\delta| \sum_{p=1}^{m} |pa^{p-1}\widetilde{A}_p(g)|. \tag{4.2.5}$$

Pour obtenir la minoration désirée sur $|\delta|$, on a besoin d'une minoration et d'une majoration de $|\widetilde{R}(g)|$ (la majoration pour assurer la condition $|\widetilde{R}(g)| \leq 1/2$), ainsi que d'une majoration des $|\widetilde{A}_p(g)|$, $1 \leq p \leq m$. Pour cela, on utilise à nouveau les expressions intégrales (4.1.2) et (4.1.3) des coefficients et du reste, auxquelles on applique la méthode du col, mais, cette fois, pour optimiser la minoration sur $|\delta|$, on utilise des approximants non diagonaux, c'est à dire que les degrés des coefficients A_p varient avec l'indice p. Pour fixer les idées, on a choisi les degrés de la manière suivante :

$$\deg A_p = n - 1, \quad 0 \leq p < q \quad \text{ou} \quad m - q < p \leq m,$$
$$\deg A_p = (s+1)n - 1, \quad q \leq p \leq m - q,$$

avec $s = 50$, $q = [\mu m]$, $\mu = 0{,}07$. Il reste encore à estimer un multiple commun des dénominateurs des coefficients des polynômes A_p, $0 \leq p \leq m$. Cela se fait à l'aide d'une technique développée par Hata [29, 30], pour estimer un commun multiple aux dénominateurs des coefficients intervenant dans la décomposition en fractions élémentaires de certaines fonctions rationnelles. Finalement, on obtient le résultat :

Théorème 4.2.1 [7*, Theorem 1.1] *Soit g un entier positif suffisamment grand. Alors, on a*

$$|e^g - a| > g^{-19{,}183g}, \qquad a \in \mathbb{N}.$$

On a également

$$|\log g - a| > g^{-19{,}183 \log\log g}, \qquad a \in \mathbb{N}.$$

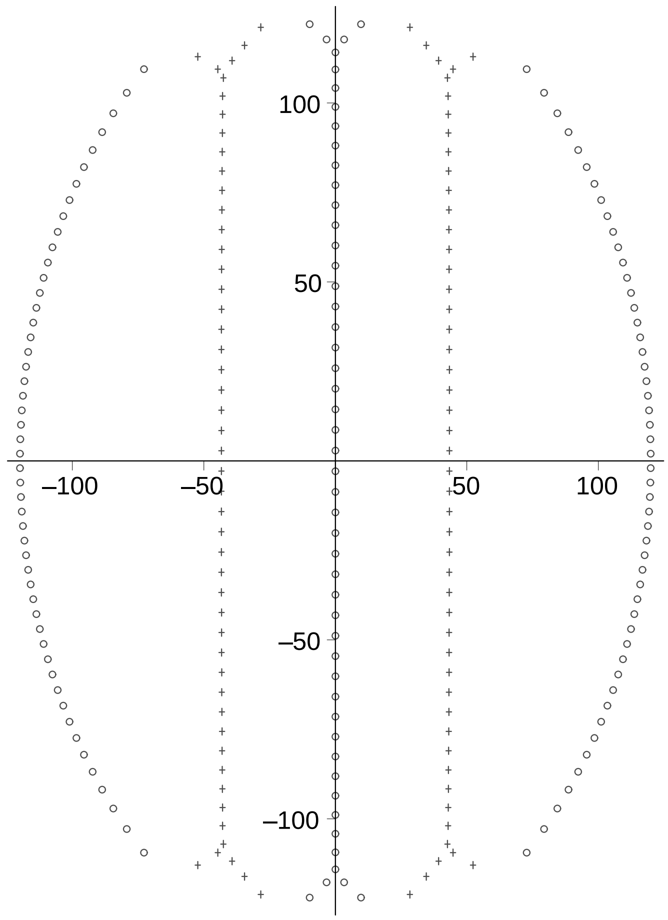

Figure 3

4.3 Distribution des zéros et asymptotique normalisée

Le premier article concernant la distribution des zéros d'approximants de la fonction exponentielle est sans doute l'article de Szegő [62], à propos de la somme partielle $s_n(z) = \sum_{k=0}^{n} z^k/k!$ de la série de Taylor de e^z. Dans cet article, l'auteur prouve que la somme normalisée $S_n(nz)$ a tous ses zéros dans $|z| \leq 1$, $n \geq 1$, et que $\hat{z}$ est un point d'accumulation des zéros de $\{s_n(nz)\}_{n=1}^{\infty}$ si et seulement si

$$|\hat{z}e^{1-\hat{z}}| = 1 \text{ et } |\hat{z}| \leq 1.$$

Cette relation détermine la *courbe de Szegő*. Saff et Varga ont généralisé ce résultat aux approximants de Padé de l'exponentielle, et mis en évidence différentes courbes ("eye-shaped curves"), constituées des points d'accumulation des zéros et pôles des approximants normalisés. On peut alors se demander comment se comportent les zéros des coefficients polynômiaux des approximants de Hermite-Padé d'un vecteur d'exponentielles $(1,e^z,\ldots,e^{mz})$. Nous donnons quelques résultats numériques obtenus dans [9*]. La Figure 3 représente les zéros des polynômes A_0, A_1, A_2, A_3, A_4, de degrés 50, tels que

$$A_0 + A_1 e^z + A_2 e^{2z} + A_3 e^{3z} + A_4 e^{4z} = O(z^{254}). \qquad (4.3.1)$$

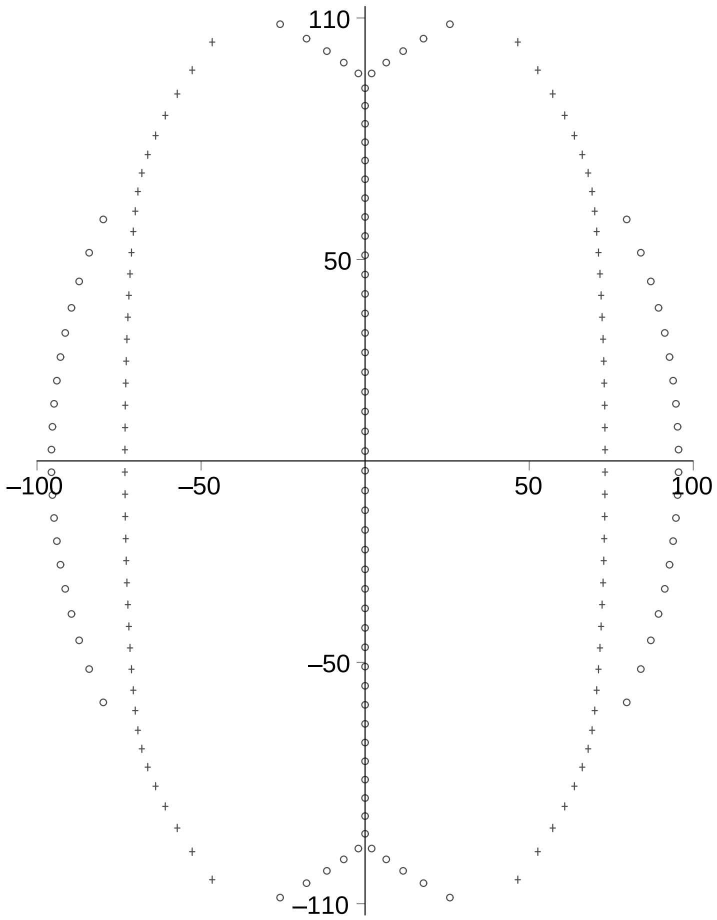

Figure 4

Les 50 zéros des polynômes A_0 à A_4 apparaissent en 5 groupes de gauche à droite de la figure. Les zéros de A_0, A_2, A_4 sont représentés par des cercles “o”. Ceux de A_1 et A_3 sont représentés par des croix “+”.
Les Figures 4 et 5 représentent encore les zéros de 5 polynômes tels que l'expression (4.3.1) s'annule en l'origine avec un ordre maximal, mais, à présent, on considère des approximations non diagonales. En effet, on choisit

$$\deg A_0 = \deg A_4 = 20,\ \deg A_1 = \deg A_3 = 40,\ \deg A_2 = 60,$$

et

$$\deg A_0 = 12,\ \deg A_1 = 24,\ \deg A_2 = 36,\ \deg A_3 = 48,\ \deg A_4 = 60,$$

respectivement. D'un point de vue théorique, les distributions des zéros peuvent être explicitées, au moins dans le cas quadratique. L'approche qui est décrite maintenant est entièrement tirée de l'article de H. Stahl [60]. Cette approche est reprise dans [14*], en conjonction avec la méthode de Riemann-Hilbert. On considère un changement d'échelle et les polynômes associés P_n, Q_n, et R_n définis par

$$P_n(z) = p_n(3nz), \quad Q_n(z) = q_n(3nz), \quad R_n(z) = r_n(3nz),$$

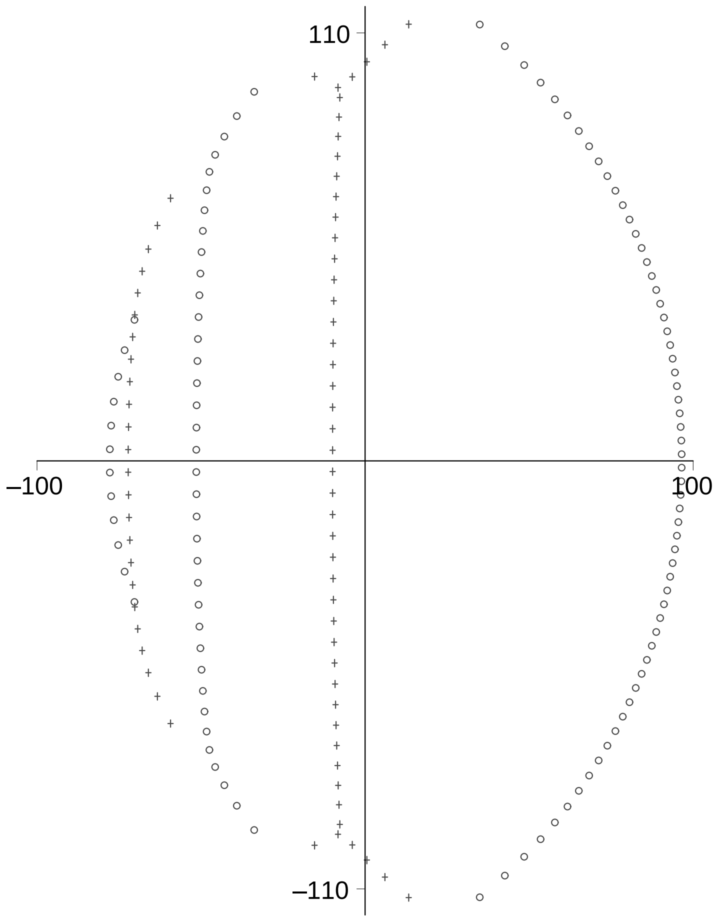

Figure 5

où

$$p_n(z) + q_n(z)e^z + r_n(z)e^{2z} = O(z^{3n+2}) \quad \text{lorsque} \quad z \to 0. \tag{4.3.2}$$

Ainsi, l'on a

$$P_n(z)(e^{-3z})^n + Q_n(z) + R_n(z)(e^{3z})^n = O(z^{3n+2}) \quad \text{lorsque} \quad z \to 0,$$

avec les expressions intégrales

$$R_n(z) = \frac{n!\,2^{n+1}\,e^{-3nz}}{2\pi i\;3^n\,n^n} \oint_{C_1} \frac{e^{3nz\,v}\;dv}{v^{n+1}(v^2-1)^{n+1}}, \tag{4.3.3}$$

$$Q_n(z) = \frac{n!\,2^{n+1}}{2\pi i\;3^n\,n^n} \oint_{C_0} \frac{e^{3nz\,v}\;dv}{v^{n+1}(v^2-1)^{n+1}}, \tag{4.3.4}$$

$$P_n(z) = \frac{n!\,2^{n+1}\,e^{3nz}}{2\pi i\;3^n\,n^n} \oint_{C_{-1}} \frac{e^{3nz\,v}\;dv}{v^{n+1}(v^2-1)^{n+1}}, \tag{4.3.5}$$

$$E_n(z) = \frac{n!\,2^{n+1}\,e^{3nz}}{2\pi i\;3^n\,n^n} \oint_{C_\infty} \frac{e^{3nz\,v}\;dv}{v^{n+1}(v^2-1)^{n+1}}, \tag{4.3.6}$$

où C_1, C_0, C_{-1} et C_∞ sont des courbes autour des points 1, 0, -1 et ∞ respectivement. Si l'on exprime la fonction sous l'intégrale comme une exponentielle, la fonction

$$u_z(v) = 3zv - \log(v-1)v(v+1)$$

apparait naturellement. Les points critiques de cette fonction sont les solutions de

$$3z = \frac{1}{v-1} + \frac{1}{v} + \frac{1}{v+1},$$

ou

$$z = z(v) = \frac{v^2 - 1/3}{v(v^2-1)}.$$

Si l'on reporte cette valeur de z dans $u_z(v)$, on définit une nouvelle fonction

$$u(v) = \frac{3v^2-1}{v^2-1} + \log\frac{1}{v(v^2-1)}.$$

On est ainsi naturellement amené à l'étude de la surface de Riemann $\mathcal{R}$ et de la fonction algébrique bijective $\psi : \mathcal{R} \to \overline{\mathbb{C}}$ tel que le diagramme suivant commute

$$\begin{tikzcd}
\zeta \in \mathcal{R} \arrow[r, "\psi"] \arrow[d, "\pi"'] & v \in \overline{\mathbb{C}} \arrow[dl, "z"] \\
z(v) \in \overline{\mathbb{C}} &
\end{tikzcd}$$

π désignant la projection canonique. La surface $\mathcal{R}$ est de genre 0 et admet 3 feuillets connectés par les points de branchement que sont les 4 racines z_1, z_2, z_3, z_4 de l'équation

$$z'(v) = -\frac{v^4 + 1/3}{v^2(v^2-1)^2} = 0. \tag{4.3.7}$$

On vérifie facilement que

$$\psi \circ \pi^{-1}(\infty) = \{-1,0,1\}.$$

On introduit les fonctions

$$h^*(\zeta) = u \circ \psi(\zeta), \quad h(\zeta) = \mathrm{Re}(u \circ \psi)(\zeta).$$

Soit $\zeta_0 = \psi^{-1}(\infty)$. La fonction h est harmonique sur $\mathcal{R} \setminus (\{\zeta_0\} \cup \pi^{-1}(\infty))$ et sous-harmonique en ζ_0.

Lemme 4.3.1 [60, Lemma 3.3] *Il existe 4 courbes de Jordan $\Gamma_{\infty j}$, $j = 1, \dots, 4$, chacune issue d'un point de branchement z_j et allant à l'infini,*

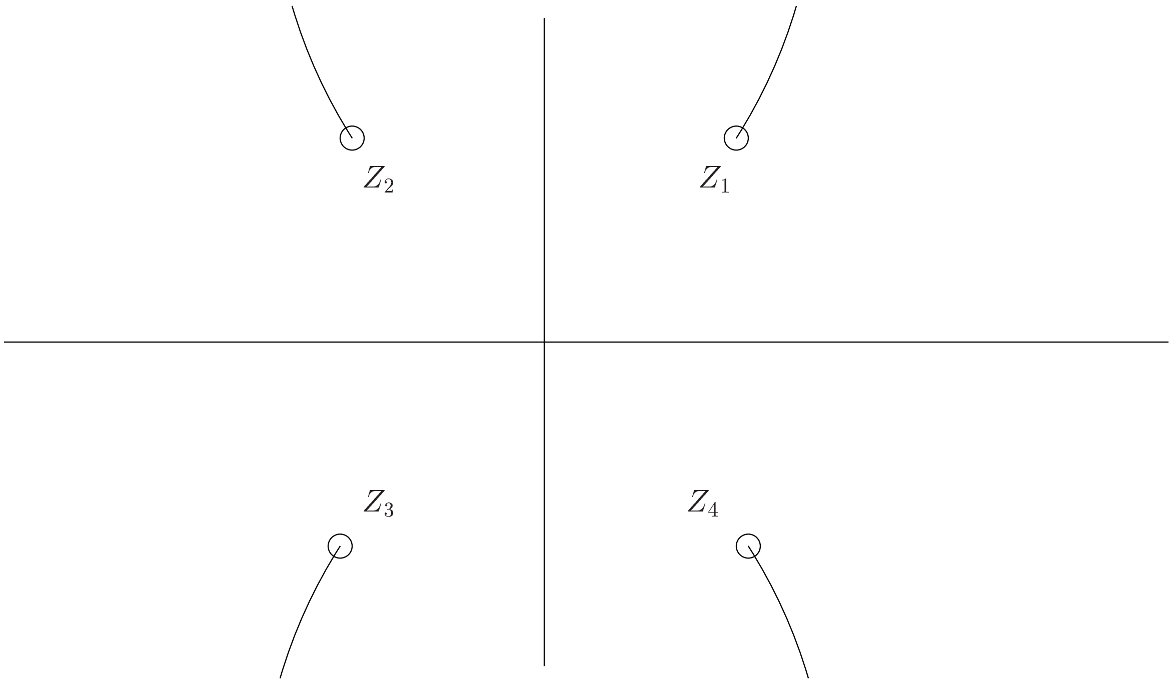

Figure 6

déterminées de manière unique par les 2 conditions suivantes :
(i) *Soit $\hat{D}_1 \subset \overline{\mathbb{C}}$, le domaine $\overline{\mathbb{C}} \setminus (\Gamma_{\infty 1} \cup \Gamma_{\infty 4})$ et $\hat{\pi}_1^{-1}$ la branche de π^{-1} dans $\hat{D}_1$ telle que*

$$\psi \circ \hat{\pi}_1^{-1}(0) = \sqrt{1/3}.$$

La fonction $\hat{h}_1 := h \circ \hat{\pi}_1^{-1}$ admet un prolongement continu à tout $\mathbb{C}$.
(ii) *Soit $\hat{D}_{-1} \subset \overline{\mathbb{C}}$, le domaine $\overline{\mathbb{C}} \setminus (\Gamma_{\infty 2} \cup \Gamma_{\infty 3})$ et $\hat{\pi}_{-1}^{-1}$ la branche de π^{-1} dans $\hat{D}_{-1}$ telle que*

$$\psi \circ \hat{\pi}_{-1}^{-1}(0) = -\sqrt{1/3}.$$

La fonction $\hat{h}_{-1} := h \circ \hat{\pi}_{-1}^{-1}$ admet un prolongement continu à tout $\mathbb{C}$.
Soit $\hat{D}_0 \subset \overline{\mathbb{C}}$, le domaine $\overline{\mathbb{C}} \setminus (\Gamma_{\infty 1} \cup \Gamma_{\infty 2} \cup \Gamma_{\infty 3} \cup \Gamma_{\infty 4})$ et $\hat{\pi}_0^{-1}$ la branche de π^{-1} dans $\hat{D}_0$ telle que

$$\psi \circ \hat{\pi}_0^{-1}(0) = \infty.$$

Alors si les 2 conditions (i) *et* (ii) *sont satisfaites, la fonction $\hat{h}_0 := h \circ \hat{\pi}_0^{-1}$ admet un prolongement continu à tout $\mathbb{C} \setminus \{0\}$ avec une singularité logarithmique en 0.*

La Figure 6 décrit les 4 courbes $\Gamma_{\infty j}$, $j = 1, \dots, 4$.

Lemme 4.3.2 [60, Lemma 3.5] *L'ensemble*

$$\{z \in \overline{\mathbb{C}},\, \hat{h}_0(z) = \hat{h}_1(z)\}$$

est constitué de 3 arcs analytiques Γ_{1k}, $k = 1{,}2{,}3$, connectant les 2 points de branchement z_1 et z_4.

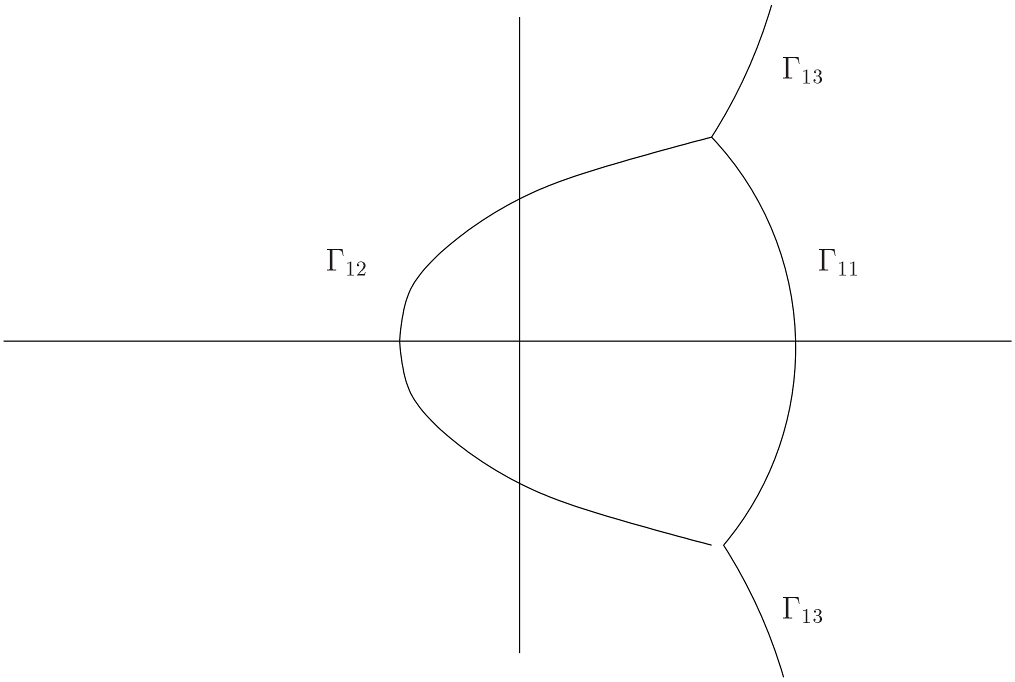

Figure 7

On a $\Gamma_{13} = \Gamma_{\infty 1} \cup \Gamma_{\infty 4}$. De même, l'ensemble

$$\{z \in \overline{\mathbb{C}},\, \hat{h}_0(z) = \hat{h}_{-1}(z)\}$$

est constitué de 3 arcs analytiques $\Gamma_{-1,k}$, symétriques des arcs Γ_{1k} par rapport à l'axe imaginaire.

La Figure 7 décrit les 3 arcs $\Gamma_{1,k}$, $k = 1{,}2{,}3$. Soit la fonction

$$h_{\max}(z) = \max\{h(\zeta),\, \zeta \in \mathcal{R},\, \pi(\zeta) = z\}.$$

La fonction $h_{\max}$, en tant que maximum de fonctions sous-harmoniques, est elle-même sous-harmonique. Elle est harmonique en tout point de $\mathbb{C}$, sauf en les points où les différentes branches de $h \circ \pi^{-1}$ coincident. Le lemme suivant précise les propriétés de la fonction $h_{\max}$.

Lemme 4.3.3 [60, Lemma 2.6] *Il existe un ensemble Γ d'arcs de Jordan analytiques tels que la fonction $h_{\max}$ soit harmonique exactement dans $\mathbb{C} \setminus \Gamma$. De la formule de Poisson-Jordan en théorie du potentiel, et de la sous-harmonicité de $h_{\max}$, on déduit l'existence d'une mesure positive ν sur Γ telle que pour tout $R > 0$, on a*

$$h_{\max}(z) = h_R(z) + \int_{|t|<R} \log\left|\frac{R(z-t)}{R^2 - \bar{t}z}\right| d\nu(t), \quad \text{pour} \quad |z| \leq R,$$

où h_R est une fonction harmonique dans $\{|z| < R\}$. L'ensemble Γ

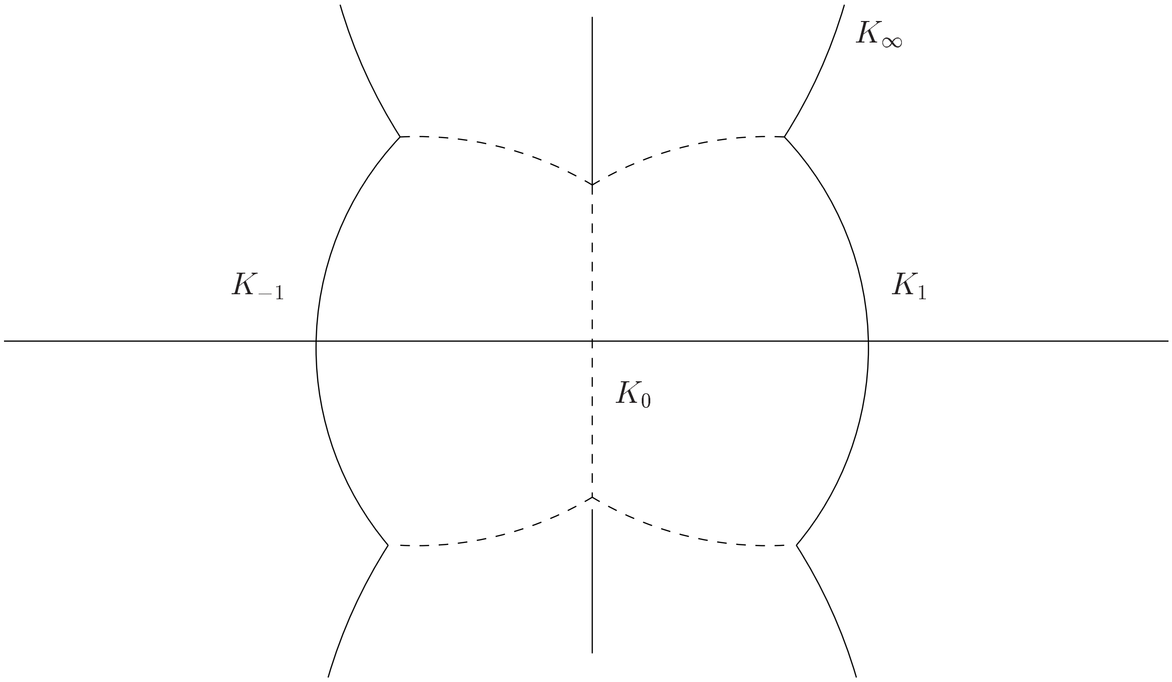

Figure 8

est symétrique par rapport aux axes réel et imaginaire, et se décompose en

$$\Gamma = K_{-1} \cup K_0 \cup K_1 \cup K_\infty,$$

avec

$$K_{-1} = \Gamma_{-11}, \quad K_1 = \Gamma_{11}, \quad K_\infty = \Gamma_{13} \cup \Gamma_{-13} = \Gamma_{\infty 1} \cup \Gamma_{\infty 2} \cup \Gamma_{\infty 3} \cup \Gamma_{\infty 4}$$

et

$$K_0 = (\Gamma_{12} \cap \{\mathrm{Re} z > 0\}) \cup (\Gamma_{-12} \cap \{\mathrm{Re} z < 0\}) \cup [-iy_1, iy_1], \quad y_1 > 0.$$

L'ensemble Γ est décrit par la Figure 8. A partir des arcs définis dans le lemme précédent, on introduit les nouvelles fonctions suivantes. Soit $D_j := \overline{\mathbb{C}} \backslash K_j$, $j = -1,0,1,\infty$. On définit les différentes branches ψ_j, h_j et h_j^*, $j = -1,0,1,\infty$, des fonctions multivaluées $\psi \circ \pi^{-1}$, $h \circ \pi^{-1}$ et $h^* \circ \pi^{-1}$ sur chacun des domaines D_j, $j = -1,0,1,\infty$, par les conditions

$$\psi_j(\infty) = j, \quad j = -1,0,1, \quad \psi_\infty(0) = \infty,$$

$$h_j^* = u \circ \psi_j = h^* \circ \pi_j^{-1}, \quad h_j = \mathrm{Re}(u \circ \psi_j) = h \circ \pi_j^{-1}, \quad j = -1,0,1,\infty.$$

Alors, on a le résultat suivant

Théorème 4.3.4 [60, Theorem 2.8] *Les 3 restrictions $\nu_j = \nu_{|K_j}$, $j = -1,0,1$, de la mesure ν introduite dans le Lemme 4.3.3 sont des mesures de probabilité. La mesure positive $\nu_\infty = \nu_{|K_\infty}$ est de masse infinie. On a les représentations*

$$h_j(z) = 3j\mathrm{Re}(z) + \log\binom{2}{j+1} + \int \log|z - t| d\nu_j(t), \quad j = -1,0,1,$$

et dans chaque disque fixé de rayon $R > 0$, il existe également une représentation intégrale de h_∞ faisant intervenir la mesure ν_∞.

Il est alors possible de construire pour chaque point z dans $\mathbb{C}$, 4 courbes fermées qui peuvent être utilisées pour estimer les intégrales (4.3.3), (4.3.4), (4.3.5) et (4.3.6) par la méthode du col.

Lemme 4.3.5 [60, Lemma 4.6] *Soit*

$$G(v) = G(v,z) = e^{3zv}/v(v^2 - 1),$$

et soit $j \in \{-1,0,1,\infty\}$, $z \in D_j$, $z \neq 0$ *si* $j = \infty$. *Il existe une courbe de Jordan* C_j, *analytique par morceaux, telle que*

(i) La courbe C_j *entoure le point* j *et ce point* j *seulement. Cette courbe peut être utilisée comme contour d'intégration dans l'expression intégrale correspondante.*

(ii) Le point critique $v_j = \psi_j(z)$ *appartient à* C_j *et*

$$|G(v_j)| > |G(v)|, \quad \forall v \in C_j \setminus \{v_j\}.$$

Grâce au lemme précédent, la méthode du col s'applique avec les chemins critiques C_j pour donner des asymptotiques fortes des approximants de Hermite-Padé de la fonction exponentielle.

Théorème 4.3.6 [60, Theorem 2.9] *On a les égalités suivantes :*

$$P_n(z) = f_{-1}(z)e^{n(h^*_{-1}(z)+3z)}\left(1 + O(1/n)\right), \quad n \to \infty, \quad z \in D_{-1},$$

$$Q_n(z) = f_0(z)e^{nh^*_0(z)}\left(1 + O(1/n)\right), \quad n \to \infty, \quad z \in D_0,$$

$$R_n(z) = f_1(z)e^{n(h^*_1(z)-3z)}\left(1 + O(1/n)\right), \quad n \to \infty, \quad z \in D_0,$$

$$E_n(z) = f_\infty(z)e^{nh^*_\infty(z)}\left(1 + O(1/n)\right), \quad n \to \infty, \quad z \in D_\infty, \quad z \neq 0,$$

où

$$f_j(z) = 2/\sqrt{3\psi_j(z)^4 + 1}, \quad j = -1,0,\infty.$$

Les estimations en $O(1/n)$ *sont vérifiées localement uniformément en* z.

Du théorème précédent et des expressions de $h_j = \mathrm{Re}(h^*_j)$ données dans le Théorème 4.3.4, on déduit les asymptotiques en racine n-ième suivantes.

Théorème 4.3.7 [60, Theorem 2.2] *Localement uniformément, on a*

$$\lim_{n\to\infty} \frac{1}{n}\log|P_n(z)| = \int \log|z-t|d\nu_{-1}(t), \quad z \in \mathbb{C} \setminus \mathrm{supp}(\nu_{-1}),$$

$$\lim_{n\to\infty} \frac{1}{n}\log|Q_n(z)| = \log 2 + \int \log|z-t|d\nu_0(t), \quad z \in \mathbb{C} \setminus \mathrm{supp}(\nu_0),$$

$$\lim_{n\to\infty} \frac{1}{n}\log|R_n(z)| = \int \log|z-t|d\nu_1(t), \quad z \in \mathbb{C} \setminus \mathrm{supp}(\nu_1),$$

$$\lim_{n\to\infty} \frac{1}{n}\log|E_n(z)| = h(\nu_\infty,z), \quad z \in \mathbb{C} \setminus \mathrm{supp}(\nu_\infty),$$

où $h(\nu_\infty,z)$ *admet aussi une représentation intégrale explicite faisant intervenir la mesure* ν_∞.

Des résultats précédents découlent également la distribution asymptotique recherchée des zéros de P_n, Q_n, R_n et de l'erreur E_n.

Théorème 4.3.8 [60, Theorem 2.1] *Soient* ν_{P_n}, ν_{Q_n}, ν_{R_n}, *les mesures discrètes normalisées associées aux zéros de* P_n, Q_n *et* R_n. *Alors on a les convergences faibles suivantes, lorsque* $n \to \infty$,

$$\nu_{P_n} \to \nu_{-1}, \quad \nu_{Q_n} \to \nu_0, \quad \nu_{R_n} \to \nu_1.$$

Soit ν_{E_n}, la mesure discrète (de support infini) associée aux zéros de l'erreur E_n. Alors lorsque $n \to \infty$,

$$\frac{1}{n}\nu_{E_n,R} \to \nu_{\infty,R} + 3\delta_0,$$

où $\nu_{E_n,R}$ et $\nu_{\infty,R}$ désignent les restrictions de ν_{E_n} et ν_∞ au disque de rayon R.

Pour conclure ce chapitre, nous allons décrire brièvement comment les asymptotiques précédentes peuvent s'obtenir et même se préciser grâce à une méthode générale développée dans les années 90, principalement par P. Deift, A.S. Fokas, A.R. Its, A.V. Kitaev, T. Kriecherbauer, K. T-R. Mclaughlin, S. Venakides et X. Zhou. Cette méthode s'applique à de nombreuses familles de polynômes exhibant des propriétés d'orthogonalité. Elle consiste à caractériser ces polynômes comme les solutions d'un problème de Riemann-Hilbert bidimensionnel, et de là, d'en tirer des informations sur leur asymptotique, grâce à une succession de transformations qui permettent de décomposer le problème et finalement d'utiliser une méthode du col.

Soit $h : \mathbb{R} \to \mathbb{R}$, une fonction intégrable. On définit sa transformée de Cauchy

$$\mathcal{C}[h](z) := \frac{1}{2i\pi}\int_{-\infty}^{+\infty} \frac{h(t)}{t-z}dt, \quad z \in \mathbb{C} \setminus \mathbb{R}.$$

Soit $w : \mathbb{R} \to [0,\infty)$ tel que chaque moment $\int_{\mathbb{R}} x^j w(x)dx$, $j \in \mathbb{N}$, soit fini et chaque $w(s)s^j$, $j \in \mathbb{N}$, satisfait une condition de Lipschitz uniforme sur $\mathbb{R}$,

$$|w(s)s^j - w(x)x^j| \leq H_j|s-x|^{\alpha_j}, \quad 0 < \alpha_j.$$

On note

$$p_n(x) = p_n(w,x) = \gamma_n x^n + \dots, \quad \gamma_n > 0,$$

le n-ième polynôme orthogonal associé à $w(x)$.

Théorème 4.3.9 (voir [20, 38]) *Le problème de Riemann-Hilbert suivant :*

(I) $\quad Y : \mathbb{C} \setminus \mathbb{R} \to \mathbb{C}^{2\times 2}$ est analytique,

(II) $$Y_+(x) = Y_-(x)\begin{pmatrix} 1 & w(x) \\ 0 & 1 \end{pmatrix}, \quad x \in \mathbb{R},$$

(III) $$Y(z)\begin{pmatrix} z^{-n} & 0 \\ 0 & z^n \end{pmatrix} = I + O\left(\frac{1}{|z|}\right), \quad |z| \to \infty,$$

où I désigne la matrice unité 2×2, admet l'unique solution

$$Y(z) = \begin{pmatrix} p_n(z)/\gamma_n & \mathcal{C}[p_n w](z)/\gamma_n \\ -2i\pi\gamma_{n-1}p_{n-1}(z) & -2i\pi\gamma_{n-1}\mathcal{C}[p_{n-1}w](z) \end{pmatrix}. \tag{4.3.8}$$

On donne quelques éléments de la preuve. Soit Y une solution du problème de Riemann-Hilbert. Alors (II) entraine que

$$\det Y_+(x) = \det Y_-(x), \quad x \in \mathbb{R},$$

et chaque limite est continue, par le théorème de Privalov sur les intégrales singulières. Donc, $\det Y$ est analytique sur $\mathbb{C} \setminus \mathbb{R}$ et continue sur $\mathbb{R}$. C'est une fonction entière telle que $\det Y(z) \to 1$,

$|z| \to \infty$. Par le théorème de Liouville, $\det Y(z) = 1$ sur $\mathbb{C}$. Par conséquent, $Y(z)^{-1}$ est analytique sur $\mathbb{C} \setminus \mathbb{R}$. Soit Z une autre solution. On considère la matrice $R(z) = Z(z)Y(z)^{-1}$ qui vérifie

$$R_+(x) = R_-(x), \quad x \in \mathbb{R}.$$

A nouveau, R est une fonction entière et (III) entraine que

$$R(z) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + O((1/|z|)\,, \quad |z| \to \infty.$$

Donc, $R = I$, ce qui prouve l'unicité d'une solution. D'autre part, la matrice $Y(z)$ dans (4.3.8) vérifie évidemment (I). La propriété (II) s'obtient par le théorème de Plemelj. La propriété (III) est moins évidente en ce qui concerne la 2-ième colonne de Y. C'est là où intervient l'orthogonalité des polynômes p_n. Nous ne rentrerons pas dans le détail des estimations.
Une fois le Théorème 4.3.9 établi, on peut appliquer à la solution Y une succession de transformations

$$Y \to U \to T \to S \to R$$

qui aboutit finalement à un problème de Riemann-Hilbert simple résoluble explicitement. La première transformation $Y \to U$ est un changement d'échelle qui permet de se ramener ultérieurement à un problème de Riemann-Hilbert sur l'intervalle $[-1{,}1]$. La transformation $U \to T$ consiste à introduire une "fonction phase" g analytique sur $\mathbb{C} \setminus \mathbb{R}$, vérifiant certaines conditions sur $\mathbb{R}$, en particulier $g_+ - g_-$ y est purement imaginaire. Cette fonction s'interprète comme le potentiel correspondant à une mesure de probabilité, solution d'un problème extrémal en théorie du potentiel logarithmique avec champ externe. L'effet de la fonction g est de ramener le saut à des termes oscillatoires sur $[-1{,}1]$ et à des termes à décroissance exponentielle sur $\mathbb{R} \setminus [-1{,}1]$, lorsque $n \to \infty$. L'étape $T \to S$ consiste à déformer le contour $\mathbb{R}$ en un contour $\sum_S = \cup_{i=1}^5 \sum_i$,

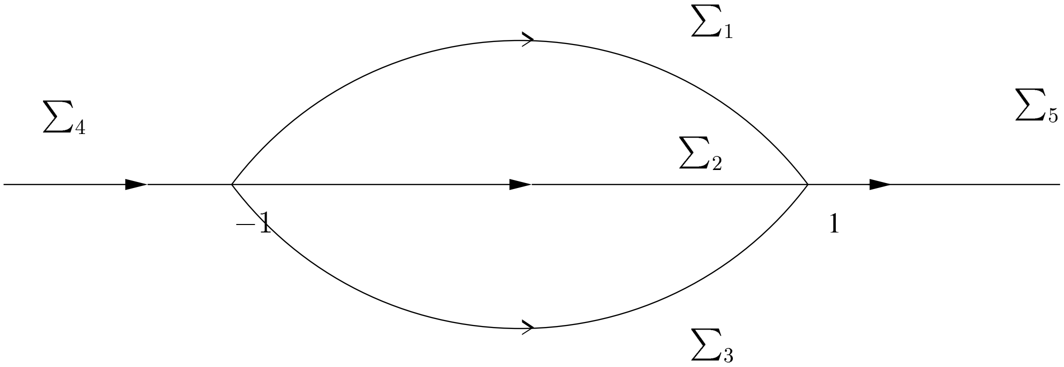

Figure 9

voir la Figure 9, de telle sorte que le saut considéré dans le problème de Riemann-Hilbert pour S converge vers la matrice identité sur $\sum_S \setminus \sum_2$, les termes oscillatoires sur $[-1{,}1]$ devenant à décroissance exponentielle sur $\sum_1$ et $\sum_3$. Ainsi, on est ramené, lorsque $n \to \infty$, à un problème de Riemann-Hilbert sur l'intervalle $[-1{,}1]$ uniquement, qui se résout de manière explicite. La transformation $S \to R$ consiste à se ramener à ce problème élémentaire, en étudiant comment le passage à la limite influe sur la solution du problème initial. Cela demande en particulier d'examiner localement la solution au voisinage des singularités -1 et 1. Ceci se fait à l'aide de fonctions de Airy.
Les approximants de Padé étudiés dans ce chapitre vérifient des propriétés d'orthogonalité. Par exemple, dans le cas quadratique, les polynômes définis par (4.3.2) vérifient les relations

$$\int_{\mathcal{C}} \frac{p_n(z)e^{-z} + r_n(z)e^{z}}{z^{3n+2}} z^k dz = 0, \quad k = 0{,}1, \ldots ,2n,$$

où $\mathcal{C}$ est un contour autour de l'origine, orienté positivement. Cela permet d'appliquer la méthode précédente, modifiée convenablement, pour obtenir des asymptotiques très précises, notamment sur le contour Γ défini dans le Lemme 4.3.3, lieu d'accumulation des zéros des différents approximants, et en particulier au voisinage des 4 points de branchement z_j, solutions de (4.3.7). Cela fait l'objet de l'étude menée dans [14*].

Bibliographie

[1] R. J. Arms and A. Edrei, *The Padé tables and continued fractions generated by totally positive sequences.* In "Mathematical Essays dedicated to A.J. Macintyre", 1–21, Ohio Univ. Press, Athens, Ohio, 1970.

[2] V.M. Adamjan, D.Z. Arov and M.G. Krein, *Infinite Hankel matrices and generalized Caratheodory-Fejer and Riesz problems*, Functional Anal. Appl. 2: 1-18, 1968

[3] V.M. Adamjan, D.Z. Arov and M.G. Krein, *Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem*, Math. USSR-Sb. 15: 31-73, 1971.

[4] G.A. Baker, Jr., Essentials of Padé Approximants, Academic Press, New York, 1975

[5] L. Baratchart, *Existence and generic properties of L^2 approximants for linear systems*, I.M.A. Jour. of Math. Control & Information, Vol. 3:89-101, 1986.

[6] L. Baratchart, *Sur l'approximation rationnelle l^2 pour les systèmes dynamiques linéaires*, Thèse de doctorat d'état, Université de Nice, 1987.

[7] L. Baratchart, *Rational and meromorphic approximation in L^p of the circle: System-theoretic motivations, critical points and error rates*, Computational Methods in Function Theory (CMFT'97), N. Papamichael, S. Ruscheweyh, E. Saff (éditeurs), World Scientific Publishing Co. Pte. Ltd., p. 1-34, 1998.

[8] L. Baratchart, M. Cardelli, and M. Olivi, *Identification and rational L^2 approximation : a gradient algorithm*, Automatica 27, No.2, 413–418, 1991.

[9] L. Baratchart, J. Grimm, J. Leblond, M. Olivi, F. Seyfert, F. Wielonsky, *Identification d'un filtre hyperfréquences par approximation dans le domaine complexe*, rapport technique INRIA no. 219, 1998.

[10] L. Baratchart, J. Leblond, F. Mandréa, E.B. Saff, *How can the meromorphic approximation help to solve some 2D inverse problems for the Laplacian?*, Inverse Problems 15, 79–90, 1999.

[11] L. Baratchart and M. Olivi, *Index of critical points in l^2-approximation*, Systems & Control Letters 10:167-174, 1988.

[12] L. Baratchart and M. Olivi, *Critical points and error rank in best H^2 matrix rational approximation of fixed McMillan degree*, Const. Approx. 14, 273-300, 1998.

[13] L. Baratchart and F. Seyfert, *An L^p analog to the AAK theory for $p \geq 2$*, to appear in J. Funct. Analysis, 2002.

[14] P.B. Borwein, *Rational interpolation to e^x, II*, SIAM. J. Math. Anal. Vol.16, No. 3, pp. 656-662, 1985.

[15] P.B. Borwein, *Quadratic Hermite-Padé approximation to the exponential function*, Const. Approx. 2, 291-302, 1986.

[16] D. Braess, *On the conjecture of Meinardus on rational approximation of e^x, II*, J. Approx. Theory 40, pp. 375-379, 1984.

[17] D. Braess, Nonlinear Approximation Theory. Springer Series in Computational Mathematics, vol. 7, Springer Verlag, Berlin, 1986.

[18] V.I. Buslaev, *Simple counterexample to the Baker-Gammel-Wills conjecture*, East J. On Approx. 7, pp. 515-517, 2001.

[19] P.E. Caines, Linear Stochastic Systems, Probability and Mathematical Statistics, Wiley, 1988.

[20] P. Deift, T. Kriecherbauer, K. T-R. Mclaughlin, S. Venakides and X. Zhou, *Strong asymptotics of orthogonal polynomials with respect to exponential weights*, Comm. Pure Appl. Math. 52, pp. 1491-1552, 1999.

[21] J. Della Dora, *Contribution à l'approximation de fonctions de la variable complexe au sens de Hermite–Padé et de Hardy*, Thèse d'état, Univ. Scient. et Medicale de Grenoble, 1980.

[22] J.C.Doyle, B. A. Francis and A. R. Tannenbaum, Feedback Control Theory, Macmillan Publishing Company, 1992.

[23] M. Duc–Jacquet, *Approximation des fonctionelles linéaires sur les espaces Hilbertiens à noyaux reproduisants*, Thèse d'état, Univ. Scient. et Medicale de Grenoble, 1973.

[24] V.D. Erohin, *On the best approximation of analytic functions by rational functions*, Doklady Akad.Nauk SSSR 128, pp. 29-32, (in Russian), 1959.

[25] J. B. Garnett, Bounded Analytic Functions, Academic Press, 1981.

[26] A.A. Gonchar, *On the convergence of Padé approximants.* Math. U.S.S.R. Sbornik 21, 155–166, 1973.

[27] A.A. Gonchar and G. López Lagomasino, *On Markov's theorem for multipoint Padé approximants.* Math. USSR Sb., 34, 449-59, 1978.

[28] J. Grimm, *Rational approximation of transfer functions in the Hyperion software*, Rapport de recherche 4002, INRIA, 2000.

[29] M. Hata, *Legendre type polynomials and irrationality measures*, J. Reine Angew. Math. 407, 99-125, 1990.

[30] M. Hata, *A lower bound for rational approximations to π*, J. Number Theory 43, 51-67, 1993.

[31] R.E. Kalman, P.L. Falb, and M.A. Arbib, Topics in mathematical system theory, Mc Graw Hill, 1969.

[32] J. Karlsson, *Rational interpolation and best rational approximation.* J. Math. Anal. Appl. 53, 38–52, 1976.

[33] S. Lang, Complex Analysis, Addison-Wesley, Reading, MA., 1977.

[34] J. Leblond, E.B. Saff, F. Wielonsky, *Weighted H^2 rational approximation and consistency*, Numer. Math. 90, pp. 521-553, 2002.

[35] A. L. Levin, *The distribution of poles of rational functions of best approximation and related questions*, Math. USSR Sbornik 9, n^0 2, pp.267–274, 1969.

[36] L. Ljung, System identification : Theory for the user, Prentice–Hall, 1987.

[37] G. López Lagomasino, *On the asymptotics of the ratio of orthonormal polynomials and the convergence of multipoint Padé approximants.* Math. USSR Sbornik 56, 216-29, 1987.

[38] D.S. Lubinsky, *Asymptotics of orthogonal polynomials: some old, some new, some identities*, Acta Appl. Math. 61, 207–256, 2000.

[39] K. Mahler, *Applications of some formulas by Hermite to the approximation of exponentials and logarithms*, Math. Ann. 168, 200-227, 1967.

[40] K. Mahler, *Perfect systems*, Comp. Math. 19, 95–166, 1968.

[41] F. Mandréa, *Comportement asymptotique des pôles d'approximants rationnels et méromorphes : application aux problèmes inverses du Laplacien 2D*, Thèse de doctorat, Université de Nice, 2001.

[42] M. Marden : "Geometry of Polynomials" Mathematical Surveys No. 3, American Mathematical Society, Providence, RI, 1966.

[43] A. Markov, *Deux démonstrations de la convergence de certaines fractions continues.* Act. Math. 19, 93–104, 1895.

[44] D.J. Newman, Approximation with Rational Functions, Regional Conf. Series 41. AMS Providence, Rhode Island, 1979.

[45] J. Nuttall, *The convergence of Padé approximants of meromorphic functions.* J. Math. Anal. Appl. 31, 147–153, 1970.

[46] J.R. Partington, An introduction to Hankel operators, Cambridge University Press, Cambridge, 1988.

[47] O. Perron, Die Lehre von den Kettenbrüchen, 3rd ed., vol. 2, B. G. Teubner, Stuttgart, 1957.

[48] P.P. Petrushev and V.A. Popov, Rational Approximation of Real Functions, Encyclopedia of Mathematics and its Applications, Cambridge University Press 28, Cambridge, 1987.

[49] G. Pólya and G. Szegö, Problems and Theorems in Analysis, I, Springer Verlag, Berlin, 1972.

[50] C. Pommerenke, *Padé approximants and convergence in capacity.* J. Math. Anal. Appl. 41, 775–780, 1973.

[51] P.A. Regalia, Adaptative IIR Filtering in Signal Processing and Control, Dekker, 1995.

[52] H.H. Rosenbrock, State space and multivariable theory, Wiley, New York, 1970.

[53] E.B. Saff, *An extension of Montessus de Ballore's theorem on the convergence of interpolating rational functions.* J. Approx. Theory, 6:63-67, 1972.

[54] E.B. Saff and V. Totik, Logarithmic potentials with external fields, *Grundlehren der math. wiss.* 316, Springer-Verlag, 1997.

[55] E. B. Saff, R. S. Varga, *On the zeros and poles of Padé approximants to* e^z. II., Padé and rational approximations: theory and applications (E.B. Saff, R.S. Varga, eds.), pp. 195-213. New York, Academic Press, 1977.

[56] D. Sarason, The H^p spaces of an annulus. Mem. Amer. Math. Soc., No.56, 1965.

[57] A.N. Shiryayev, Probability, Graduate Texts in Math., Springer Verlag, New York, 1984.

[58] T. Söderström, *On the uniqueness of maximum likelihood identification*, Automatica, pp. 193-197, 1975.

[59] H. Stahl, *Strong asymptotics for orthogonal polynomials with varying weights*, Acta Sci. Math. (Szeged) 66, 147–192, 2000.

[60] H. Stahl, *Quadratic Hermite-Padé polynomials associated with the exponential function*, manuscript, 2001.

[61] H. Stahl and V. Totik, General orthogonal polynomials, *Encyclopedia of Mathematics*, 43, Cambridge university press, New York, 1992.

[62] G. Szegő, *Über eine eigenschaft der exponentialreihe*, Sitzungsber. Berl. Math. Ges. 23, 50–64, 1924.

[63] R. Tijdeman, *On the number of zeros of general exponential polynomials*, Indag. Math. 33, 1–7, 1971.

[64] V. Totik, Weighted approximation with varying weight, Lecture notes in Math., 1569, Springer Verlag, Heidelberg, 1994.

[65] L.N. Trefethen, *The asymptotic accuracy of rational approximations to e^z on a disk*, J. Approx. Theory 40, pp. 380-383, 1984.

[66] M. Tsuji, Potential Theory in Modern Function Theory, Maruzen, Tokyo, 1959.

[67] W. Van Assche, *Multiple orthogonal polynomials, irrationality and transcendence*, Continued Fractions: From Analytic Number Theory to Constructive Approximation, Contemporary Mathematics 236, 325–342, 1999.

[68] M. Voorhoeve, *On the oscillation of exponential polynomials*, Math. Z. 151 (3), 277–294, 1976.

[69] M. Waldschmidt, *Indépendance algébrique des valeurs de la fonction exponentielle*, *Bull. Soc. Math. France* 99, 285–304, 1971.

[70] H. Wallin, *The convergence of Padé approximants and the size of the power series coefficients.* Applicable Anal. 4, no. 3, 235–251, 1974.

[71] H. Wallin, *Potential theory and approximation of analytic functions by rational interpolation*, Proc. Colloq., Univ. Joensuu, Joensuu, Lecture Notes in Math., 747, Springer, Berlin, 434–450, 1979.

[72] J.L. Walsh, Interpolation and Approximation by Rational Functions in the Complex Domain, AMS colloq. Pub. XX, 1969.

Résumé

Les travaux présentés portent sur différents aspects de l'approximation rationnelle dans le plan complexe :
– approximation rationnelle en norme L^2 dans le disque unité,
– interpolation rationnelle de la fonction exponentielle,
– approximation de Hermite–Padé de la fonction exponentielle.

Le problème d'approximation abordé dans la première partie consiste à minimiser, pour une fonction g dans l'espace de Hardy $H_2(U)$ du disque unité U, l'intégrale

$$\frac{1}{2\pi}\int_0^{2\pi}|g(e^{it})-r(e^{it})|^2dt,$$

où r décrit l'ensemble des fonctions rationnelles d'un degré donné, analytiques dans l'adhérence $\overline{U}$ de U. Nous motivons l'étude de ce problème non linéaire, en relation avec l'identification de systèmes dynamiques en automatique. Nous en donnons les propriétés de base, notamment de différentiabilité, puis nous décrivons un algorithme de calcul des minima locaux. Nous nous intéressons ensuite particulièrement au problème de l'unicité d'un point critique pour un degré d'approximation fixé. Nous décrivons deux critères, en terme de décroissance des erreurs d'approximation, assurant qu'un point critique est un minimum local. En conjonction avec un théorème de l'indice (analogue du théorème de Poincaré–Hopf sur une variété lisse) donnant une contrainte globale sur l'ensemble des points critiques, on en déduit un critère d'unicité d'un point critique, et en particulier d'un meilleur approximant. Nous appliquons ce critère à différentes situations, notamment pour l'approximation au sens de Montessus de Ballore, l'approximation diagonale de la fonction exponentielle ainsi que des fonctions de Markov (ou transformées de Cauchy) de la forme

$$f_\mu(z)=\int\frac{d\mu(t)}{t-z}.$$

En ce qui concerne ces fonctions, la régularité de la mesure μ est un facteur déterminant dans le type de résultat obtenu. Cela se reflète également dans les estimées asymptotiques des erreurs d'approximation, obtenues avec l'aide de différents outils de la théorie du potentiel.

Dans la deuxième partie de notre travail, nous nous intéressons à une généralisation du résultat classique (théorème de Padé) de convergence des approximants de Padé de la fonction exponentielle. Dans un premier temps, en s'appuyant sur des résultats de géométrie des polynômes, tels que les théorèmes de Walsh et Szegő, nous étendons ce théorème au cas des interpolants rationnels en des points réels de la fonction exponentielle. Puis, en nous basant sur un théorème de Rolle complexe pour les polynômes exponentiels réels, nous donnons une généralisation partielle pour les interpolants rationnels en des points complexes.

Dans la troisième partie, nous étendons le théorème de Padé au cas des approximants de Hermite-Padé d'un vecteur d'exponentielles :

$$R(z)=\sum_{p=0}^{m}A_p(z)e^{pz}=O(z^{mn+n-1}),\quad \deg A_p<n,\quad 0\le p\le m.$$

Pour une normalisation convenable, nous déterminons la limite de chaque coefficient A_p et du reste R, lorsque n tend vers l'infini. En conjonction avec des résultats de Hata sur la nature arithmétique des coefficients A_p, cette étude nous permet d'améliorer une inégalité due à Mahler concernant la distance aux entiers de puissances du nombre e. Nous montrons que

$$|e^g-a|\ge g^{-19{,}183g},\quad a{,}g\in\mathbb{N},\quad g\ge g_0.$$

Finalement, nous décrivons la distribution asymptotique des zéros de ces approximants, obtenus en interprétant ceux–ci comme les solutions d'un problème de Riemann–Hilbert tridimensionnel. Cette approche nouvelle s'inscrit dans les travaux récents de Deift, Kuijlaars, Van Assche et de leurs coauteurs.