Multipoint Schur's algorithm, rational orthogonal functions, asymptotic properties and Schur rational approximation

Vincent Lunot, Laurent Baratchart, Stanislas Kupin, Martine Olivi

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Multipoint Schur's algorithm, rational orthogonal functions, asymptotic properties and Schur rational approximation

Vincent Lunot — Laurent Baratchart — Stanislas Kupin — Martine Olivi

N° 6620

Août 2008

Thème NUM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6620--FR+ENG

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE
INRIA
centre de recherche SOPHIA ANTIPOLIS - MÉDITERRANÉE

Multipoint Schur's algorithm, rational orthogonal functions, asymptotic properties and Schur rational approximation

Vincent Lunot , Laurent Baratchart , Stanislas Kupin , Martine Olivi

Thème NUM — Systèmes numériques
Équipes-Projets Apics

Rapport de recherche n° 6620 — Août 2008 — 83 pages

Abstract: In [20] the connections between the Schur algorithm, the Wall's continued fractions and the orthogonal polynomials are revisited and used to establish some nice convergence properties of the sequence of Schur functions associated with a Schur function. In this report, we generalize some of Krushchev's results to the case of a multipoint Schur algorithm, that is a Schur algorithm where all the interpolation points are not taken in 0 but anywhere in the open unit disk. To this end, orthogonal rational functions and a recent generalization of Geronimus theorem are used. Then, we consider the problem of approximating a Schur function by a rational function which is also Schur. This problem of approximation is very important for the synthesis and identification of passive systems. We prove that all strictly Schur rational function of degree n can be written as the $2n$-th convergent of the Schur algorithm if the interpolation points are correctly chosen. This leads to a parametrization using the multipoint Schur algorithm. Some examples are computed by an L^2 norm optimization process and the results are validated by comparison with the unconstrained L^2 rational approximation.

Key-words: multipoint Schur's algorithm, rational orthogonal functions, Schur functions, Wall continued fraction, asymptotic properties, rational approximation

Il s'agit de la seconde partie de la thèse de Vincent Lunot, soutenue le 05/05/08

Centre de recherche INRIA Sophia Antipolis – Méditerranée
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Algorithme de Schur multipoint, fonctions rationnelles orthogonales, propriétés asymptotiques et approximation rationnelle Schur

Résumé : Dans [20] les relations entre l'algorithme de Schur, les fractions continues de Wall et les polynômes orthogonaux sont revisitées et utilisées pour établir certaines propriétés de convergence de la suite de Schur d'une fonction Schur. Dans ce rapport, certains résultats de Krushchev sont généralisés à l'algorithme de Schur multipoints, c'est-à-dire lorsque les points d'interpolation ne sont plus pris en 0 mais en n'importe quel point du disque unité ouvert. Pour cela, on fait appel aux fonctions rationnelles orthogonales et à une récente généralisation du théorème de Géronimus. On considère ensuite le problème de l'approximation rationnelle Schur d'une fonction Schur. Ce problème revêt une importance particulière dans le domaine de la synthèse et de l'identification de sytèmes passifs. On prouve que toute fonction rationnelle Schur de degré n peut être obtenue comme le $2n$-ième convergent d'un algorithme de Schur dont les points d'interpolation sont convenablement choisis. Cela nous permet de construire un paramétrage des fonctions rationnelles strictement Schur fondé sur l'algorithme de Schur multipoints. Des exemples numériques sont traités par une procédure d'optimisation de la norme L^2 et les résultats validés par comparaison avec l'approximation rationnelle L^2 non-contrainte.

Mots-clés : algorithme de Schur multipoint, fonctions rationnelles orthogonales, fonctions de Schur, fraction continue de Wall, propriétés asymptotiques, approximation rationnelle

1 Introduction

In this work, we are interested in approximating a Schur function f by a rational function which is also Schur. A Schur function is an analytic function whose modulus is bounded by 1 in the unit disk. This problem of approximation is very important for the synthesis and identification of passive systems. The main idea is to use a generalized multipoint Schur algorithm, that is a Schur algorithm where all the reference points are not taken in 0 but are taken at points $(\alpha_j)_{j\geq 1}$ anywhere in the unit disk. Such an algorithm leads to a sequence of Schur rational functions that we are studying all along this paper.
In the first section, we introduce the generalized Schur algorithm, and rewrite it as a continued fraction. We then give some basic properties of the convergents of this continued fraction. In particular, the convergents of even order are Schur rational functions which interpolate f at the points (α_j).
In the next section, we introduce the orthogonal rational functions on the unit circle and give all the basic results needed on this topic. Our main reference is the book [6].
The third section makes a connection between the Schur algorithm and the orthogonal rational functions. This is a generalization of the Geronimus theorem ([13], [21]) which states that the Schur parameters are equal to the Geronimus parameters of the orthogonal polynomials of the measure associated to f by the Herglotz transform.
The first three sections are in fact all the necessary background to study asymptotic properties of the convergents of even order. These properties are given in the fourth section, and are mainly a generalization of the work of Khrushchev ([20]) who studied the L^2-convergence in the case of the classical algorithm. The difficulty here comes from the fact that we let the points go the circle.
In addition, we obtained a "Szegő condition" and a result of convergence for the Schur functions which seems to be asymptotically very close to a BMO convergence.
Finally, in the fifth section, we give some practical ways to approximate a Schur function by a rational function of a given order. We prove that all strictly Schur rational function of degree n can be written as the $2n$-th convergent of the Schur algorithm if the interpolation points are correctly chosen. This leads to a parametrization using the Schur algorithm. We give some details about it, and also explain how to compute effectively the L^2-norm. Some examples are computed using an optimization process, and the results are validated by a comparison with the unconstrained L^2 rational approximation.

2 Notations and first definitions

This section presents some basic notations and definitions that will be used throughout our study.

We denote by $\mathbb{D}$ the unit disc $\mathbb{D} = \{z \in \mathbb{C}, |z| < 1\}$ and by $\mathbb{T}$ the unit circle $\mathbb{T} = \{z \in \mathbb{C}, |z| = 1\}$.
$H(\mathbb{D})$ and $C(\mathbb{D})$ represent respectively the set of analytic functions and the set of continuous functions over $\mathbb{D}$. We denote by $A(\mathbb{D})$ the disk algebra, i.e. the set of analytic functions in $\mathbb{D}$, continuous on $\overline{\mathbb{D}}$.
For a function f, we define the infinity norm $\|\cdot\|_\infty$ by $\|f\|_\infty = \sup_{z\in\mathbb{D}} |f(z)|$.

Definition 2.1 *An analytic function f on $\mathbb{D}$ such that $\|f\|_\infty \leq 1$ is called a Schur function. The set $\mathcal{S}$ of all Schur functions is called the Schur class $\mathcal{S}$.*
If f is an analytic function in $\mathbb{D}$ with $\|f\|_\infty < 1$, we will say that f is strictly Schur.

Let $\{z_n\}$ be a subset of $\mathbb{D} \setminus \{0\}$ and s be a nonnegative integer. A function of the form

$$B(z) = z^s \prod_n \frac{|z_n|}{z_n} \frac{z_n - z}{1 - \bar{z}_n z}$$

is called a Blaschke product. Furthermore, if the set $\{z_n\}$ is finite, it is called a finite Blaschke product.
It is well known (e.g. [12] or [25]) that if $\sum_n (1 - |z_n|) < \infty$, then B is in $H^\infty(\mathbb{D})$, the zeros of B are the points z_n (and 0 if $s > 0$) and $|B| = 1$ almost everywhere on $\mathbb{T}$. Therefore, $\sum_n (1 - |z_n|) < \infty$ is a sufficient condition for the existence of a non-zero function in $H^\infty(\mathbb{D})$ with given zeros $\{z_n\}$. In fact, this is also a necessary condition (e.g. [12] or [25]): the zeros z_n of a non-zero function in $H^\infty(\mathbb{D})$ satisfy $\sum_n (1 - |z_n|) < \infty$.
We will sometimes use the following corollary: if a function in $H^\infty(\mathbb{D})$ has an infinity of zeros at the points z_n and if $\sum_n (1 - |z_n|) = \infty$, then it is the zero function.

For a sequence $\{\alpha_k\}_{k=0}^\infty \subset \mathbb{D}$ with $\alpha_0 = 0$, we define the elementary Blaschke factors

$$\zeta_k = \frac{z - \alpha_k}{1 - \bar{\alpha}_k z} \quad , k \geq 0 \tag{1}$$

and the partial Blaschke products

$$\begin{cases} \mathcal{B}_0(z) = 1 \\ \mathcal{B}_k(z) = \mathcal{B}_{k-1}(z)\zeta_k(z) = \prod_{i=1}^k \frac{z-\alpha_i}{1-\bar{\alpha}_i z} \text{ for } k \geq 1. \end{cases} \tag{2}$$

The functions $\{\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_n\}$ span the space

$$\mathcal{L}_n = \left\{ \frac{p_n}{\pi_n} \ : \ \pi_n(z) = \prod_{k=1}^n (1 - \bar{\alpha}_k z), \quad p_n \in \mathcal{P}_n \right\} \tag{3}$$

where $\mathcal{P}_n$ is the space of algebraic polynomials of degree at most n.
In particular, if all the α_k are equal to 0, the space $\mathcal{L}_n$ coincides with the space $\mathcal{P}_n$.
Note that a function of $\mathcal{L}_n$ is analytic in $\mathbb{D}$.

For any function f, we introduce the parahermitian conjugate f_* defined by

$$f_*(z) = \overline{f(1/\bar{z})}. \tag{4}$$

Two useful and immediate equalities are $\zeta_{n*} = \zeta_n{}^{-1}$ and $\mathcal{B}_{k*} = \mathcal{B}_k{}^{-1}$.
We set for any function $f \in \mathcal{L}_n$:

$$f^* = \mathcal{B}_n f_*. \tag{5}$$

It is immediate to check that f^* is also in $\mathcal{L}_n$.
We denote by $\mathcal{B}_{n,i}$ the product $\prod_{k=i}^{k=n} \zeta_k$. If

$$f = a_n\mathcal{B}_n + a_{n-1}\mathcal{B}_{n-1} + \cdots + a_1\mathcal{B}_1 + a_0$$

then

$$f^* = \bar{a}_0\mathcal{B}_{n,1} + \bar{a}_1\mathcal{B}_{n,2} + \cdots + \bar{a}_{n-2}\mathcal{B}_{n,n-1} + \bar{a}_{n-1}\mathcal{B}_{n,n} + \bar{a}_n.$$

Finally, we remark that the leading coefficient a_n is given by

$$a_n = \overline{f^*(\alpha_n)}$$

and that

$$a_0 = f(\alpha_1).$$

We denote by m the normalized Lebesgue measure on $\mathbb{T}$: $m(\mathbb{T}) = 1$.
Now that all the main notations have been presented, we are able to begin with the study of the Schur algorithm.

3 The Schur algorithm

Starting from a Schur function f, the classical Schur algorithm ([26]) gives a sequence of Schur functions $(f_k)_{k\in\mathbb{N}}$ and a sequence of complex numbers $(\gamma_k)_{k\in\mathbb{N}}$ as follows:

$$\begin{cases} f_0 = f, \\ \gamma_k = f_k(0), \\ f_{k+1}(z) = \dfrac{1}{z}\dfrac{f_k(z) - \gamma_k}{1 - \overline{\gamma_k} f_k(z)}, \end{cases} \quad \text{for } k \geq 0.$$

Note that for every $k \in \mathbb{N}$, $\omega \mapsto \frac{\omega-\gamma_k}{1-\overline{\gamma_k}\omega}$ is a Moebius transform which maps $\mathbb{D}$ onto $\mathbb{D}$, so by the Schwarz lemma ([12]) f_k is a Schur function for every $k \in \mathbb{N}$. An interesting property ([4]) of the Schur algorithm is that it realizes a one-to-one correspondence between the Schur class $\mathcal{S}$ and the sequence of complex numbers $(\gamma_k)_{k\in\mathbb{N}}$ having the properties: $|\gamma_k| \leq 1$ for $k \geq 0$, and if for a certain k_0, $|\gamma_{k_0}| = 1$, $f_{k_0}(z) = \gamma_{k_0}$ is a constant function and then $\gamma_k = 0$ for $k > k_0$.
Note that the Schur algorithm extends to operator-valued functions ([24], [8]).

3.1 Multipoint Schur algorithm

In the classical algorithm, the Schur parameters γ_n are obtained by evaluating the functions f_n at 0. This process can be extended to evaluation points arbitrary in $\mathbb{D}$ (e.g. [18], [21]). We next describe such an algorithm.

Let $\{\alpha_k\}_{k=1}^{\infty}$ be a sequence of points in $\mathbb{D}$ and $\{c_k\}_{k=0}^{\infty}$ be a sequence of points in $\mathbb{T}$ with $c_0 = 1$. Then, the generalized Schur algorithm is :

For $k \geq 0$, f_k and γ_k are defined by

$$\begin{cases} f_0 = f \\ \gamma_k = \bar{c}_k f_k(\alpha_{k+1}) \\ f_{k+1} = \dfrac{1}{\zeta_{k+1}} \dfrac{\bar{c}_k f_k - \gamma_k}{1 - \bar{\gamma}_k \bar{c}_k f_k} \end{cases} \quad \text{for } k \geq 0,$$

where ζ_k is the Moebius transform defined by (1).
If $|\gamma_k| = 1$, the algorithm stops.

The parameters (α_k) are the interpolations points. They are those parameters equal to 0 in the classical Schur algorithm, which are presently taken anywhere in the disk. The parameters (c_k) have modulus equal to 1, and are rotations applied to the f_k at each step of the algorithm. Note that the (c_k) can also be seen as normalization parameters of the Moebius transforms since

$$\begin{aligned} f_{k+1} &= \frac{1}{\zeta_{k+1}} \frac{\bar{c}_k f_k - \gamma_k}{1 - \bar{\gamma}_k \bar{c}_k f_k} \\ &= \frac{\bar{c}_k}{\zeta_{k+1}} \frac{f_k - c_k \gamma_k}{1 - \bar{\gamma}_k \bar{c}_k f_k} \\ &= \frac{1}{c_k \zeta_{k+1}} \frac{f_k - f_k(\alpha_{k+1})}{1 - \overline{f_k(\alpha_{k+1})} f_k}. \end{aligned}$$

As in the classical case, the sequence $(f_n)_{n\in\mathbb{N}}$ is a sequence of Schur functions, therefore the $(\gamma_n)_{n\in\mathbb{N}}$ lie in $\overline{\mathbb{D}}$.

Definition 3.1 *The sequence $(\gamma_n)_{n\in\mathbb{N}}$ is called the sequence of Schur parameters of the Schur function f associated to the sequence (α_k).*

The Schur parameters depend only on the values of f and its derivatives $f^{(j)}$ at the points $(\alpha_k)_k$. More precisely,

Proposition 3.2 *For $k \in \mathbb{N}$, γ_k depends only on the values $f^{(i)}(\alpha_{j+1})$, $0 \leq j \leq k$, $0 \leq i < m_{j+1}$, where m_{j+1} is the multiplicity of α_{j+1} at the k-th step, i.e. m_{j+1} is the cardinality of the set $\{l,\, 0 \leq l \leq k,\, \alpha_{l+1} = \alpha_{j+1}\}$.*

Proof Noticing that $f_j(\alpha_j) = f'_{j-1}(\alpha_j)\bar{c}_{j-1}\bar{z}_j \frac{1-|\alpha_j|^2}{1-|f_{j-1}(\alpha_j)|^2}$, the proof is immediate by induction.

■

The Schur algorithm can be reversed in order to express f_{k-1} as a function of f_k. We obtain

$$f_{k-1} = c_{k-1}\frac{\zeta_k f_k + \gamma_{k-1}}{1+\bar{\gamma}_{k-1}\zeta_k f_k} = c_{k-1}\gamma_{k-1} + \frac{(1-|\gamma_{k-1}|^2)c_{k-1}\zeta_k}{\bar{\gamma}_{k-1}\zeta_k + \frac{1}{f_k}}. \quad (6)$$

We denote by τ_k the map

$$\begin{array}{rccl} \tau_k: & \overline{\mathbb{D}} & \longrightarrow & \mathcal{S} \\ & \omega & \longmapsto & \tau_k(\omega) = \begin{cases} c_k\gamma_k + \frac{(1-|\gamma_k|^2)c_k\zeta_{k+1}}{\bar{\gamma}_k\zeta_{k+1}+\frac{1}{\omega}} & \text{if } \omega \neq 0, \\ c_k\gamma_k & \text{if } \omega = 0. \end{cases} \end{array}$$

Note that we should write $\tau_k(\omega)(z)$ because $\tau_k(\omega)$ is a Schur function of z through ζ_{k+1}. Much of the recursive complexity of the Schur algorithm lies in the fact that we shall substitute to ω a function of z to make $\tau_k(\omega(z))(z)$ a function of z only. In particular, we have $f_k = \tau_k(f_{k+1})$. Therefore, f is equal to

$$f = \tau_0 \circ \tau_1 \circ \cdots \circ \tau_n(f_{n+1}). \quad (7)$$

Proposition 3.3 *The Schur algorithm stops if and only if f is a finite Blaschke product.*

Proof For p a polynomial, denote by $\tilde{p}$ the polynomial $z^n\overline{p\left(\frac{1}{\bar{z}}\right)}$ where n is the degree of p.
Suppose that f_n is a Blaschke product of degree n. Then f_n can be expressed as $\frac{p}{\tilde{p}}$ where p has its roots in $\mathbb{D}$, so

$$f_{n+1} = \frac{1-\overline{\alpha_{n+1}}z}{z-\alpha_{n+1}}\frac{\bar{c}_n p - \gamma_n\tilde{p}}{\tilde{p}-\bar{\gamma}_n\bar{c}_n p}.$$

Let $P_0 = \bar{c}_n p - \gamma_n\tilde{p}$. Then $\tilde{P}_0 = c_n(\tilde{p} - \bar{c}_n\bar{\gamma}_n p)$, so $\frac{\bar{c}_n p - \gamma_n\tilde{p}}{\tilde{p}-\bar{\gamma}_n\bar{c}_n p}$ is of the form $\frac{P}{\tilde{P}}$ for some polynomial P. Note that, since $\bar{c}_n f_n(\alpha_{n+1}) = \gamma_n$, P vanishes at α_{n+1}. Therefore f_{n+1} is a Blaschke product of degree $n-1$. Thus, if f is a Blaschke product of degree n, f_n is a Blaschke product of degree 0, i.e. a constant of modulus 1, and the algorithm stops.
Conversely, if $f_k = \frac{p}{\tilde{p}}$ is a Blaschke product of degree $n-k$, then

$$f_{k-1} = c_{k-1}\frac{(z-\alpha_k)p + \gamma_{k-1}\tilde{p}(1-\overline{\alpha_k}z)}{\tilde{p}(1-\overline{\alpha_k}z) + \bar{\gamma}_{k-1}(z-\alpha_k)p},$$

so f_{k-1} is a Blaschke product of degree at most $n-k+1$. In fact, using the first part of the proof, we get that f_{k-1} is exactly of degree $n-k+1$ (otherwise f_k is not of degree $n-k$). Therefore, if f_n is a constant of modulus 1, f is a Blaschke product of degree n.

■

3.2 Continued fractions

In this section, we give a very short introduction to continued fractions. Many good references, such as [29], can be found on this topic.
A continued fraction is an infinite expression of the form

$$b_0 + \cfrac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{a_3}{b_3 + \frac{a_4}{\cdots}}}}.$$

also denoted for economy of space by

$$b_0 + \frac{a_1}{b_1} \; + \; \frac{a_2}{b_2} \; + \; \frac{a_3}{b_3} \; + \ldots$$

Let $t_0(\omega) = b_0 + \omega$ and

$$t_k(\omega) = \frac{a_k}{b_k + \omega} \quad \text{for } k \geq 1.$$

We call the n-th convergent, and we denote by P_n/Q_n, the fraction

$$\frac{P_n}{Q_n} = t_0 \circ t_1 \circ \cdots \circ t_n(0) = b_0 + \cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \ddots + \frac{a_n}{b_n}}}.$$

Proposition 3.4 *The quantities P_n and Q_n are given by the recurrence relations*

$$\begin{cases} P_{-1} = 1, Q_{-1} = 0, \\ P_0 = b_0, Q_0 = 1, \\ P_{k+1} = b_{k+1}P_k + a_{k+1}P_{k-1} \\ Q_{k+1} = b_{k+1}Q_k + a_{k+1}Q_{k-1} \end{cases}$$

for all non-negative k.
More generally,

$$t_0 \circ t_1 \circ \cdots \circ t_n(\omega) = \frac{P_{n-1}\omega + P_n}{Q_{n-1}\omega + Q_n}.$$

Proof By induction. We have

$$t_0(\omega) = b_0 + \omega = \frac{P_{-1}\omega + P_0}{Q_{-1}\omega + Q_0}.$$

Suppose the statement true for k. Then

$$\begin{aligned} t_0 \circ t_1 \circ \cdots \circ t_{k+1}(\omega) &= \frac{P_{k-1}\frac{a_{k+1}}{b_{k+1}+\omega} + P_k}{Q_{k-1}\frac{a_{k+1}}{b_{k+1}+\omega} + Q_k} \\ &= \frac{P_k\omega + b_{k+1}P_k + a_{k+1}P_{k-1}}{Q_k\omega + b_{k+1}Q_k + a_{k+1}Q_{k-1}} \\ &= \frac{P_k\omega + P_{k+1}}{Q_k\omega + Q_{k+1}}. \end{aligned}$$

This gives the announced result.

■

3.3 Wall rational functions

In this section, we follow the same scheme as in ([20]).
Let $(d_k)_{k\in\mathbb{N}}$ be a sequence of points on the unit circle $\mathbb{T}$, with $d_0 = 1$. We now define the c_k of the Schur algorithm by $c_k = d_k^2$. Let (α_k) be a sequence of points in the unit disk $\mathbb{D}$. Recall from (7) that $f = \tau_0 \circ \tau_1 \circ \cdots \circ \tau_n(f_{n+1})$ with

$$\tau_k(\omega) = c_k\gamma_k + \frac{(1-|\gamma_k|^2)c_k\zeta_{k+1}}{\bar{\gamma}_k\zeta_{k+1} + \frac{1}{\omega}}.$$

A rational Schur function R_n of degree at most n can be obtained by interrupting the Schur algorithm at step n, that is, by replacing f_{n+1} by 0:

$$\begin{aligned} R_n &= \tau_0 \circ \tau_1 \circ \cdots \circ \tau_{n-1} \circ \tau_n(0) \\ &= \tau_0 \circ \tau_1 \circ \cdots \circ \tau_{n-1}(c_n\gamma_n). \end{aligned} \tag{8}$$

The rational functions R_n play a key role in what follows. Indeed, we will see later how to approximate f using the sequence (R_n). Therefore, we will now pay a particular attention to the properties of these rational functions. The first one is an interpolation property:

Theorem 3.5 *The rational function R_n interpolates f at the points α_k, $1 \le k \le n+1$, and has the same $n+1$ first Schur parameters as f.*

Proof Remark that $\tau_k(\omega)(\alpha_{k+1})$ is independent of ω. Indeed, $\tau_k(\omega)(\alpha_{k+1}) = c_k\gamma_k$. Let k be an integer such that $0 \le k \le n$. Then:

$$\begin{aligned} f(\alpha_{k+1}) &= \tau_0 \circ \cdots \circ \tau_k(\tau_{k+1} \circ \cdots \circ \tau_n \circ f_{n+1})(\alpha_{k+1}) \\ &= \tau_0 \circ \cdots \circ \tau_k(\tau_{k+1} \circ \cdots \circ \tau_n(0))(\alpha_{k+1}) \\ &= R_n(\alpha_{k+1}). \end{aligned}$$

Thus, R_n interpolates f at the point α_{k+1}.
We next prove by induction that f and R_n have the same $n+1$ first Schur parameters. Using what precedes, we get that f and R_n have the same first Schur parameter γ_0. Now, suppose that the k first Schur parameters of f and R_n are equal. Then, if we denote by $R_n^{[1]}, \ldots R_n^{[n]}$ the Schur functions of R_n obtained through the Schur algorithm, $R_n^{[k]}$ is equal to $\tau_{k-1}^{-1} \circ \cdots \circ \tau_0^{-1}(R_n)$. Thus,

$$\begin{aligned} R_n^{[k]}(\alpha_{k+1}) &= \tau_{k-1}^{-1} \circ \cdots \circ \tau_0^{-1}(R_n)(\alpha_{k+1}) \\ &= \tau_{k-1}^{-1} \circ \cdots \circ \tau_0^{-1} \circ \tau_0 \circ \tau_1 \circ \cdots \circ \tau_{n-1}(c_n\gamma_n)(\alpha_{k+1}) \\ &= \tau_k \circ \cdots \circ \tau_{n-1}(c_n\gamma_n)(\alpha_{k+1}) = c_k\gamma_k \end{aligned}$$

since $\tau_k(\omega)(\alpha_{k+1}) = c_k\gamma_k$. Therefore, the $k+1$-st Schur parameter of R_n is equal to the $k+1$-st Schur parameter of f.

■

The previous theorem leads to the existence of a function with given Schur parameters:

Corollary 3.6 *Let $\check{\gamma}_i$, $0 \leq i \leq n-1$, be n points in the unit disk $\mathbb{D}$ and c_i, $0 \leq i \leq n-1$, be n points on the unit circle $\mathbb{T}$. Then, there is a Schur function whose n first Schur parameters are the $\check{\gamma}_i$, $0 \leq i \leq n-1$.*

Proof Using the previous theorem, the function

$$\check{R}_n = \check{\tau}_0 \circ \cdots \circ \check{\tau}_{n-1}(c_n\check{\gamma}_n)$$

where

$$\check{\tau}_k(\omega) = c_k\check{\gamma}_k + \frac{(1-|\check{\gamma}_k|^2)c_k\zeta_{k+1}}{\overline{\check{\gamma}_k}\zeta_{k+1} + \frac{1}{\omega}}$$

satisfies the announced condition.

■

We are now going to study the sequence of rational Schur function R_n using continued fractions. We note $\frac{P_n}{Q_n}$ the sequence of convergents associated to the continued fraction

$$c_0\gamma_0 + \frac{(1-|\gamma_0|^2)c_0\zeta_1}{\bar{\gamma}_0\zeta_1} \; + \; \frac{1}{c_1\gamma_1} \; + \; \frac{(1-|\gamma_1|^2)c_1\zeta_2}{\bar{\gamma}_1\zeta_2} \; + \dots \tag{9}$$

so that the R_n are the convergents of even index: $R_n = \frac{P_{2n}}{Q_{2n}}$.
By proposition 3.4, for $n \geq 1$:

$$\begin{aligned} P_{2n} &= c_n\gamma_n P_{2n-1} + P_{2n-2} \\ Q_{2n} &= c_n\gamma_n Q_{2n-1} + Q_{2n-2} \\ P_{2n-1} &= \bar{\gamma}_{n-1}\zeta_n P_{2n-2} + (1-|\gamma_{n-1}|^2)c_{n-1}\zeta_n P_{2n-3} \\ Q_{2n-1} &= \bar{\gamma}_{n-1}\zeta_n Q_{2n-2} + (1-|\gamma_{n-1}|^2)c_{n-1}\zeta_n Q_{2n-3} \end{aligned} \tag{10}$$

with

$$P_{-1} = 1, \quad P_0 = c_0\gamma_0 = \gamma_0, \quad Q_{-1} = 0, \quad Q_0 = 1.$$

Our purpose is now to give explicit formulas in order to compute R_n, that is formulas for P_{2n} and Q_{2n}. The following lemma expresses the relations between the rational functions of even and odd order. We shall make the convention that $Q_{2n}^* = \mathcal{B}_n Q_{2n*}$ and $Q_{2n+1}^* = \mathcal{B}_{n+1}Q_{2n+1*}$ and similarly for P_{2n} and P_{2n+1}. It will actually follow from the lemma that this convention agrees with definition (5), in that we will have $P_{2n+1}, Q_{2n+1} \in \mathcal{L}_{n+1}$ and $P_{2n}, Q_{2n} \in \mathcal{L}_n$ by (10).

Lemma 3.7 *For $n \geq 0$, we have*

$$P_{2n+1} = C_n \zeta_{n+1} Q^*_{2n}, \quad Q_{2n+1} = C_n \zeta_{n+1} P^*_{2n}$$

where $C_n = \prod_{k=0}^{k=n} c_k \in \mathbb{T}$.

Proof For $n = 0$ we have

$$P_1 = \bar{\gamma}_0 \zeta_1 c_0 \gamma_0 + (1 - |\gamma_0|^2) c_0 \zeta_1 = c_0 \zeta_1 Q^*_0$$

and

$$Q_1 = \bar{\gamma}_0 \zeta_1 = c_0 \zeta_1 P^*_0.$$

Assuming the hypothesis is true for all indices smaller than n, we obtain that

$$\begin{aligned}
C_n \zeta_{n+1} Q^*_{2n} &= C_n \zeta_{n+1} (c_n \gamma_n Q_{2n-1} + Q_{2n-2})^* \\
&= C_n \zeta_{n+1} (\bar{c}_n \bar{\gamma}_n Q^*_{2n-1} + \zeta_n Q^*_{2n-2}) \\
&= C_{n-1} \zeta_{n+1} (\bar{\gamma}_n Q^*_{2n-1} + c_n \zeta_n Q^*_{2n-2}) \\
&= C_{n-1} \zeta_{n+1} (\bar{\gamma}_n \bar{C}_{n-1} P_{2n-2} + c_n \bar{C}_{n-1} P_{2n-1}) \\
&= \zeta_{n+1} (\bar{\gamma}_n P_{2n-2} + c_n P_{2n-1}) \\
&= \zeta_{n+1} (\bar{\gamma}_n P_{2n} - c_n |\gamma_n|^2 P_{2n-1} + c_n P_{2n-1}) \\
&= P_{2n+1}.
\end{aligned}$$

This yields the first relation of the lemma. The proof of the other relation is similar.

■

From (10), we have for $n \geq 1$:

$$\begin{aligned}
P_{2n+1} &= \bar{\gamma}_n \zeta_{n+1} P_{2n} + (1 - |\gamma_n|^2) c_n \zeta_{n+1} P_{2n-1} \\
&= \bar{\gamma}_n \zeta_{n+1} (c_n \gamma_n P_{2n-1} + P_{2n-2}) + (1 - |\gamma_n|^2) c_n \zeta_{n+1} P_{2n-1} \\
&= \bar{\gamma}_n \zeta_{n+1} P_{2n-2} + c_n \zeta_{n+1} P_{2n-1}
\end{aligned}$$

and similarly $Q_{2n+1} = \bar{\gamma}_n \zeta_{n+1} Q_{2n-2} + c_n \zeta_{n+1} Q_{2n-1}$ so that

$$\begin{aligned}
\begin{bmatrix} P_{2n+1} & Q_{2n+1} \\ P_{2n} & Q_{2n} \end{bmatrix} &= \begin{bmatrix} c_n \zeta_{n+1} & \bar{\gamma}_n \zeta_{n+1} \\ \gamma_n c_n & 1 \end{bmatrix} \begin{bmatrix} P_{2n-1} & Q_{2n-1} \\ P_{2n-2} & Q_{2n-2} \end{bmatrix} \\
&= \begin{bmatrix} \zeta_{n+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \bar{\gamma}_n \\ \gamma_n & 1 \end{bmatrix} \begin{bmatrix} c_n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P_{2n-1} & Q_{2n-1} \\ P_{2n-2} & Q_{2n-2} \end{bmatrix}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\begin{bmatrix} c_{n+1} & 0 \\ 0 & \zeta_{n+1} \end{bmatrix} \begin{bmatrix} P_{2n+1} & Q_{2n+1} \\ P_{2n} & Q_{2n} \end{bmatrix} \\
&= \frac{\zeta_{n+1}}{\zeta_n} \begin{bmatrix} c_{n+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \bar{\gamma}_n \\ \gamma_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} c_n & 0 \\ 0 & \zeta_n \end{bmatrix} \begin{bmatrix} P_{2n-1} & Q_{2n-1} \\ P_{2n-2} & Q_{2n-2} \end{bmatrix}.
\end{aligned}$$

Thus, using the previous lemma,

$$\begin{bmatrix} C_{n+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Q^*_{2n} & P^*_{2n} \\ P_{2n} & Q_{2n} \end{bmatrix} = \begin{bmatrix} c_{n+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \bar{\gamma}_n \\ \gamma_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} C_n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Q^*_{2n-2} & P^*_{2n-2} \\ P_{2n-2} & Q_{2n-2} \end{bmatrix}. \tag{11}$$

Iterating, we get

$$\begin{bmatrix} C_{n+1}Q^*_{2n} & C_{n+1}P^*_{2n} \\ P_{2n} & Q_{2n} \end{bmatrix} = \left(\prod_{k=n}^{k=1} \begin{bmatrix} c_{k+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \bar{\gamma}_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} \zeta_k & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} c_1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \bar{\gamma}_0 \\ \gamma_0 & 1 \end{bmatrix}. \tag{12}$$

Let $\Sigma_n = \prod_{k=0}^n d_k$. Note that, by definition of c_k, we have $\Sigma_n^2 = C_n$. We choose as representative of R_n the rational function $R_n = \frac{A_n}{B_n}$ with $A_n = \bar{\Sigma}_n P_{2n}$ and $B_n = \bar{\Sigma}_n Q_{2n}$.

Definition 3.8 *A_n and B_n are called the n-th Wall rational functions associated to the Schur function f and the sequences (α_k) and (d_k).*

As pointed out before, R_n plays a key role in the theory. This role will now be emphasized through the Wall rational functions A_n and B_n.
From what precedes, we have :

Proposition 3.9 *The Wall rational functions A_n and B_n are given by the formula*

$$\Sigma_n \begin{bmatrix} c_{n+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B^*_n & A^*_n \\ A_n & B_n \end{bmatrix} = \left(\prod_{k=n}^{k=1} \begin{bmatrix} c_{k+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \bar{\gamma}_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} \zeta_k & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} c_1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \bar{\gamma}_0 \\ \gamma_0 & 1 \end{bmatrix}$$

with

$$\Sigma_n = \prod_{k=0}^{k=n} d_k.$$

Corollary 3.10 *A_n and B_n have the following properties :*

1. $B_n(z)B_n^*(z) - A_n(z)A_n^*(z) = \mathcal{B}_n(z)\omega_n$,
2. $|B_n(\xi)|^2 - |A_n(\xi)|^2 = \omega_n$ *for* $\xi \in \mathbb{T}$,
3. $f(\alpha_i) = \frac{A_n}{B_n}(\alpha_i) = \frac{B_n^*}{A_n^*}(\alpha_i)$ *for all* $1 \leq i \leq n+1$

with

$$\omega_n = \prod_{k=0}^{k=n} (1 - |\gamma_k|^2).$$

Proof By taking the determinant, we obtain from (12) that

$$\begin{aligned} B_n(z)B_n^*(z) - A_n(z)A_n^*(z) &= Q_{2n}(z)Q_{2n}^*(z) - P_{2n}(z)P_{2n}^*(z) \\ &= \mathcal{B}_n(z)\prod_{k=0}^{n}(1 - |\gamma_k|^2). \end{aligned}$$

The conclusion is then immediate.

■

Important properties of the Wall rational functions are:

Proposition 3.11 *For all $n \geq 0$:*

1. *B_n is an analytic function which does not vanish on $\overline{\mathbb{D}}$,*
2. *$\frac{A_n^*}{B_n}$ is a Schur function.*

Proof The proof will be given for P_{2n} and Q_{2n}. Since $P_0 = \gamma_0$ and $Q_0 = 1$, P_0 and Q_0 are two analytic functions and Q_0 does not vanish on $\overline{\mathbb{D}}$. Let us assume that these hypothesis are true for n. Then both functions $\frac{P_{2n}}{Q_{2n}}$ and $\frac{P_{2n}^*}{Q_{2n}}$ are analytic on $\overline{\mathbb{D}}$. From corollary 3.10, and by the maximum principle, these two functions are Schur. Furthermore, from (11), it is immediate that P_{2n+2} and Q_{2n+2} are both analytic in the disk and that

$$\begin{aligned} |Q_{2n+2}(z)| &= |\zeta_{n+1}(z)C_{n+1}\gamma_{n+1}P_{2n}^*(z) + Q_{2n}(z)| \\ &\geq |Q_{2n}(z)|\left(1 - |\gamma_{n+1}|\left|\frac{A_n^*}{B_n}\right|\right) > 0. \end{aligned}$$

■

The Wall rational functions A_n and B_n are related to f by the following formula:

Theorem 3.12 *The Wall rational functions A_n and B_n are rational functions $\in \mathcal{L}_n$ such that*

$$f(z) = \frac{A_n(z) + \zeta_{n+1}(z)B_n^*(z)f_{n+1}(z)}{B_n(z) + \zeta_{n+1}(z)A_n^*(z)f_{n+1}(z)}.$$

Proof Proposition 3.4 applied to the continuous fraction (9) gives us in view of (7)

$$f(z) = \frac{P_{2n}\frac{1}{f_{n+1}} + P_{2n+1}}{Q_{2n}\frac{1}{f_{n+1}} + Q_{2n+1}} = \frac{P_{2n} + P_{2n+1}f_{n+1}}{Q_{2n} + Q_{2n+1}f_{n+1}}.$$

But using lemma 3.7, we get

$$\begin{aligned} f(z) &= \frac{P_{2n} + C_n \zeta_{n+1} Q_{2n}^* f_{n+1}}{Q_{2n} + C_n \zeta_{n+1} P_{2n}^* f_{n+1}} \\ &= \frac{\overline{C_n}^{1/2} P_{2n} + C_n^{1/2} \zeta_{n+1} Q_{2n}^* f_{n+1}}{\overline{C_n}^{1/2} Q_{2n} + C_n^{1/2} \zeta_{n+1} P_{2n}^* f_{n+1}} \\ &= \frac{A_n + \zeta_{n+1} B_n^* f_{n+1}}{B_n + \zeta_{n+1} A_n^* f_{n+1}}. \end{aligned}$$

■

4 Orthogonal rational functions on the unit circle

Orthogonal rational functions have been widely studied ([9], [23], [6]). We recall here the main aspects of this theory. Its remarkable feature is to make connection with the Schur algorithm as we shall see in the next section.

4.1 Reproducing kernel Hilbert spaces

Good references on reproducing kernel Hilbert spaces are [27], [11] and [2]. We recall here, mostly without proof, the properties that will be useful in what follows. We will write RKHS for "Reproducing Kernel Hilbert Space".
A RKHS is a complex-valued function Hilbert space in which pointwise evaluation is a continuous linear function, that is:

Definition 4.1 *Let X be an arbitrary set and H be an Hilbert space of complex valued functions on X. H is a RKHS if and only if the linear map $f \mapsto f(x)$ from H to $\mathbb{C}$ is continuous for each $x \in X$.*

From the Riesz-Fréchet theorem ([25]), for $\omega \in X$ there exists a unique function $k(.,\omega)$ in H such that

$$f(\omega) = \langle f, k(.,\omega) \rangle \quad \forall f \in H.$$

Definition 4.2 *The function $(z,\omega) \mapsto k(z,\omega)$ from $X \times X$ to $\mathbb{C}$ such that*

$$f(\omega) = \langle f, k(.,\omega) \rangle \quad \forall f \in H \tag{13}$$

is called the reproducing kernel of H. The reproducing kernel is clearly unique.

The reproducing kernel is a Hermitian function, that is

$$\forall z \in X, \forall \omega \in X, k(z,\omega) = \overline{k(\omega, z)}.$$

Since in a Hilbert space of finite dimension pointwise evaluation is always continuous, we have

Proposition 4.3 *A Hilbert space of functions of finite dimension is a RKHS.*

The result we mainly use throughout is:

Proposition 4.4 *If H is a RKHS, and if (e_n) is an orthonormal basis, then the reproducing kernel k of H is equal to*

$$k(z,w) = \sum_n e_n(z)\overline{e_n(w)}. \tag{14}$$

Proof First, note that if $\dim(H) = \infty$, $\sum_n e_n(z)\overline{e_n(w)}$ converges in H. Indeed, we have $\sum_n |e_n(w)|^2 = \sum_n \langle e_n(.), k(.,\omega)\rangle = \|k(.,\omega)\|_2 < +\infty$ because $k(.,\omega) \in H$.
We next prove the equality (14). Let f in H. Expressing f in the basis (e_n), we obtain that $f = \sum_n a_n e_n$ for some $a_n \in \mathbb{C}$. Thus,

$$\begin{aligned} \langle f, \sum_n e_n(.)\overline{e_n(w)}\rangle &= \langle \sum_n a_n e_n(.), \sum_n e_n(.)\overline{e_n(w)}\rangle \\ &= \sum_n \langle a_n e_n(.), \overline{e_n(w)} e_n(.)\rangle \\ &= \sum_n a_n e_n(\omega) \\ &= f(\omega). \end{aligned}$$

As the reproducing kernel is unique, we get

$$k(z,w) = \sum_n e_n(z)\overline{e_n(w)}.$$

■

4.2 Christoffel-Darboux formulas in $\mathcal{L}_n$

Let μ be a real probability measure on the unit circle $\mathbb{T}$ with infinite support and $L^2(\mu)$ the familiar Hilbert space with inner product

$$\langle f, g\rangle_\mu = \int_{\mathbb{T}} f(\xi)\overline{g(\xi)} d\mu(\xi).$$

The space $\mathcal{L}_n$ endowed with the inner product $< .,. >_\mu$ is a Hilbert space of finite dimension, so it is a RKHS. Therefore, there exists a reproducing kernel $k_n(z,w)$ such that for every point $w \in \mathbb{D}$, $k_n(z,w) \in \mathcal{L}_n$ as a function of z and

$$\forall f \in \mathcal{L}_n, \forall w \in \mathbb{D}, f(w) = \langle f(.), k_n(.,w)\rangle_\mu . \tag{15}$$

Let us denote by $\{\phi_0, \phi_1, \ldots, \phi_n\}$ an orthonormal basis for $\mathcal{L}_n$ such that $\phi_0 = 1$ and $\phi_k \in \mathcal{L}_k \setminus \mathcal{L}_{k-1}$. Such a basis is easily obtained by the Gram-Schmidt orthonormalization process applied to $\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_n$. We can write

$$\phi_n = a_{n,n}\mathcal{B}_n + a_{n,n-1}\mathcal{B}_{n-1} + \ldots + a_{n,1}\mathcal{B}_1 + a_{n,0}\mathcal{B}_0, \quad a_{n,n} = \kappa_n. \tag{16}$$

Note that $\kappa_n = \overline{\phi_n^*(\alpha_n)}$.
For $0 \le k \le n$, $\mathcal{B}_n\phi_{k*}$ is in $\mathcal{L}_n$. Moreover, $\{\mathcal{B}_n\phi_{0*}, \mathcal{B}_n\phi_{1*}, \ldots, \mathcal{B}_n\phi_{n*}\}$ is also an orthonormal basis, since

$$\langle \mathcal{B}_n\phi_{k*}, \mathcal{B}_n\phi_{l*}\rangle_\mu = \int_{\mathbb{T}} |\mathcal{B}_n(\xi)|^2 \overline{\phi_k(\xi)}\phi_l(\xi)d\mu(\xi) = \delta_{k,l}.$$

Using this new basis to compute the reproducing kernel, we get by (14) that

$$k_n(z,w) = \mathcal{B}_n(z)\overline{\mathcal{B}_n(w)} \sum_{k=0}^{n} \phi_{k*}(z)\overline{\phi_{k*}(w)}. \tag{17}$$

Letting $w \to \alpha_n$, since $\mathcal{B}_n(\alpha_n) = 0$ and no term is singular except if $k = n$, every term in the sum vanishes except for $k = n$, and computing the limit we have

$$\begin{aligned} k_n(z,\alpha_n) &= \mathcal{B}_n(z)\phi_{n*}(z) \lim_{w\to\alpha_n} \overline{\mathcal{B}_n(w)\phi_{n*}(w)} \\ &= \phi_n^*(z)\overline{\phi_n^*(\alpha_n)} \\ &= \kappa_n\phi_n^*(z). \end{aligned} \tag{18}$$

In particular, $k_n(\alpha_n, \alpha_n) = |\kappa_n|^2$. From (17) we may write

$$\frac{k_n(z,w)}{\mathcal{B}_n(z)\overline{\mathcal{B}_n(w)}} - \frac{k_{n-1}(z,w)}{\mathcal{B}_{n-1}(z)\overline{\mathcal{B}_{n-1}(w)}} = \phi_{n*}(z)\overline{\phi_{n*}(w)}, \quad n \ge 1.$$

Multiplying by $\mathcal{B}_n(z)\overline{\mathcal{B}_n(w)}$ gives the following important relation:

$$k_n(z,w) - \zeta_n(z)\overline{\zeta_n(w)}k_{n-1}(z,w) = \phi_n^*(z)\overline{\phi_n^*(w)}. \tag{19}$$

Using (14) with the orthonormal basis $(\phi_0, \ldots, \phi_n)$, we also have that

$$k_n(z,w) = k_{n-1}(z,w) + \phi_n(z)\overline{\phi_n(w)}, \quad n \ge 1. \tag{20}$$

We may use this relation to replace either $k_n(z,w)$ or $k_{n-1}(z,w)$ in relation (19) and then compute the other one. We get this way the following Christoffel-Darboux relations ([6], Theorem 3.1.3):

Proposition 4.5 *For z and w in $\mathbb{C}$ such that z and w do not coincide on $\mathbb{T}$, and for $n \ge 1$, we have*

$$k_{n-1}(z,w) = \frac{\phi_n^*(z)\overline{\phi_n^*(w)} - \phi_n(z)\overline{\phi_n(w)}}{1 - \zeta_n(z)\overline{\zeta_n(w)}} \tag{21}$$

$$k_n(z,w) = \frac{\phi_n^*(z)\overline{\phi_n^*(w)} - \zeta_n(z)\overline{\zeta_n(w)}\phi_n(z)\overline{\phi_n(w)}}{1 - \zeta_n(z)\overline{\zeta_n(w)}}. \tag{22}$$

A direct application of the Christoffel-Darboux relations is ([6], Corollary 3.1.4):

Proposition 4.6 *For all $n \geq 1$, for all $z \in \mathbb{D}$: $\phi_n^*(z) \neq 0$ and* $\left|\frac{\phi_n(z)}{\phi_n^*(z)}\right| < 1$.

Proof From (21), we get for $w = z$ that

$$(1 - |\zeta_n(z)|^2) k_{n-1}(z, z) = |\phi_n^*(z)|^2 - |\phi_n(z)|^2.$$

But

$$k_{n-1}(z, z) = \sum_{k=0}^{n-1} |\phi_k(z)|^2 = 1 + \sum_{k=1}^{n-1} |\phi_k(z)|^2 > 0.$$

Since $k_{n-1}(z, z) > 0$ and $|\zeta_n(z)| < 1$ for $z \in \mathbb{D}$, we deduce that

$$|\phi_n^*(z)| > |\phi_n(z)|$$

and the conclusion is immediate.

■

Using the above proposition, we get $\phi_n^*(\alpha_{n-1}) \neq 0$ for every $n \geq 0$. Therefore, since ϕ_n is uniquely determined up to a multiplicative constant of modulus 1, we can fix ϕ_n uniquely by assuming $\phi_n^*(\alpha_{n-1}) > 0$. In what follows, we denote by ϕ_n the orthogonal rational functions normalized by

$$\phi_n^*(\alpha_{n-1}) > 0. \tag{23}$$

Note that this is not the same normalization as in [6], where it is supposed that $\kappa_n = \phi_n^*(\alpha_n) > 0$.
The Christoffel-Darboux formulas imply a recurrence relation for the ϕ_n, which is the object of the next section.

4.3 Orthogonal rational functions of the first kind

Evaluate (21) at $w = \alpha_{n-1}$ and take into account the equality $k_{n-1}(z, \alpha_{n-1}) = \kappa_{n-1}\phi_{n-1}^*(z)$ (see (18)). This gives the relation

$$\kappa_{n-1}\phi_{n-1}^*(z) = \frac{\phi_n^*(z)\overline{\phi_n^*(\alpha_{n-1})} - \phi_n(z)\overline{\phi_n(\alpha_{n-1})}}{1 - \zeta_n(z)\overline{\zeta_n(\alpha_{n-1})}}, \quad n \geq 1. \tag{24}$$

Then take the superstar conjugate

$$\overline{\kappa_{n-1}}\phi_{n-1}(z) = \frac{\phi_n(z)\phi_n^*(\alpha_{n-1}) - \phi_n^*(z)\phi_n(\alpha_{n-1})}{\zeta_n(z) - \zeta_n(\alpha_{n-1})}$$

and put these equations together into a linear system to obtain

$$\begin{bmatrix} \overline{\phi_n^*(\alpha_{n-1})} & -\phi_n(\alpha_{n-1}) \\ -\overline{\phi_n(\alpha_{n-1})} & \phi_n^*(\alpha_{n-1}) \end{bmatrix} \begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix}$$
$$= \begin{bmatrix} \overline{\kappa_{n-1}} & 0 \\ 0 & \kappa_{n-1} \end{bmatrix} \begin{bmatrix} \zeta_n(z)-\zeta_n(\alpha_{n-1}) & 0 \\ 0 & 1-\overline{\zeta_n(\alpha_{n-1})}\zeta_n(z) \end{bmatrix} \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix}$$

so that we have the recurrence relations

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = T_n(z) \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} \quad \forall n \geq 1,$$

where T_n is equal to

$$T_n = \frac{|\kappa_{n-1}|}{|\phi_n^*(\alpha_{n-1})|^2 - |\phi_n(\alpha_{n-1})|^2} \begin{bmatrix} \overline{\phi_n^*(\alpha_{n-1})} & \phi_n(\alpha_{n-1}) \\ \overline{\phi_n(\alpha_{n-1})} & \phi_n^*(\alpha_{n-1}) \end{bmatrix}$$
$$\begin{bmatrix} \overline{\kappa_{n-1}}/|\kappa_{n-1}| & 0 \\ 0 & \kappa_{n-1}/|\kappa_{n-1}| \end{bmatrix} \begin{bmatrix} \zeta_n-\zeta_n(\alpha_{n-1}) & 0 \\ 0 & 1-\overline{\zeta_n(\alpha_{n-1})}\zeta_n \end{bmatrix}.$$

Now, it is easily checked that

$$\begin{aligned} \zeta_n(z)-\zeta_n(\alpha_{n-1}) &= \frac{(1-|\alpha_n|^2)(z-\alpha_{n-1})}{(1-\bar{\alpha}_n\alpha_{n-1})(1-\bar{\alpha}_n z)}, \\ 1-\overline{\zeta_n(\alpha_{n-1})}\zeta_n(z) &= \frac{(1-|\alpha_n|^2)(1-\bar{\alpha}_{n-1} z)}{(1-\alpha_n\bar{\alpha}_{n-1})(1-\bar{\alpha}_n z)}, \end{aligned}$$

so that

$$\begin{aligned} &\begin{bmatrix} \zeta_n(z)-\zeta_n(\alpha_{n-1}) & 0 \\ 0 & 1-\overline{\zeta_n(\alpha_{n-1})}\zeta_n(z) \end{bmatrix} \\ &= \frac{(1-|\alpha_n|^2)(1-\bar{\alpha}_{n-1} z)}{(1-\alpha_n\bar{\alpha}_{n-1})(1-\bar{\alpha}_n z)} \begin{bmatrix} \eta_n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{25}$$

where

$$\eta_n = \frac{1-\alpha_n\bar{\alpha}_{n-1}}{1-\bar{\alpha}_n\alpha_{n-1}} \in \mathbb{T}. \tag{26}$$

Furthermore,

$$\begin{aligned} &\begin{bmatrix} \overline{\phi_n^*(\alpha_{n-1})} & \phi_n(\alpha_{n-1}) \\ \overline{\phi_n(\alpha_{n-1})} & \phi_n^*(\alpha_{n-1}) \end{bmatrix} \begin{bmatrix} \overline{\kappa_{n-1}}/|\kappa_{n-1}| & 0 \\ 0 & \kappa_{n-1}/|\kappa_{n-1}| \end{bmatrix} \begin{bmatrix} \eta_n & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \overline{\phi_n^*(\alpha_{n-1})}\eta_n\overline{\kappa_{n-1}}/|\kappa_{n-1}| & 0 \\ 0 & \phi_n^*(\alpha_{n-1})\kappa_{n-1}/|\kappa_{n-1}| \end{bmatrix} \begin{bmatrix} 1 & -\overline{\tilde{\gamma}_n} \\ -\tilde{\gamma}_n & 1 \end{bmatrix} \end{aligned} \tag{27}$$

where

$$\tilde{\gamma}_n = -\eta_n \frac{\overline{\phi_n(\alpha_{n-1})}}{\phi_n^*(\alpha_{n-1})} \frac{\overline{\kappa_{n-1}}}{\kappa_{n-1}}, \quad n \geq 1. \tag{28}$$

Note that, by proposition 4.6, $\tilde{\gamma}_n$ is well defined in $\mathbb{D}$.

Definition 4.7 *We call $\tilde{\gamma}_n \in \mathbb{D}$ the n-th Szegő (or Geronimus) parameter of the measure μ associated to the sequence (α_k).*

Evaluating (24) at $z = \alpha_{n-1}$ and taking the square root, we get after a short computation

$$|\kappa_{n-1}| = |1 - \bar{\alpha}_n \alpha_{n-1}| \frac{\sqrt{|\phi_n^*(\alpha_{n-1})|^2 - |\phi_n(\alpha_{n-1})|^2}}{\sqrt{1 - |\alpha_{n-1}|^2}\sqrt{1 - |\alpha_n|^2}},$$

so that, from (28),

$$\frac{|\kappa_{n-1}|}{|\phi_n^*(\alpha_{n-1})|^2 - |\phi_n(\alpha_{n-1})|^2} = \frac{|1 - \bar{\alpha}_n \alpha_{n-1}|}{\sqrt{1 - |\alpha_{n-1}|^2}\sqrt{1 - |\alpha_n|^2}|\phi_n^*(\alpha_{n-1})|\sqrt{1 - |\tilde{\gamma}_n|^2}}. \tag{29}$$

Combining (25), (27) and (29), we finally have that

$$T_n(z) = \sqrt{\frac{1 - |\alpha_n|^2}{1 - |\alpha_{n-1}|^2}} \frac{1}{\sqrt{1 - |\tilde{\gamma}_n|^2}} \frac{1 - \bar{\alpha}_{n-1} z}{1 - \bar{\alpha}_n z} \begin{bmatrix} \lambda_n & 0 \\ 0 & \bar{\lambda}_n \end{bmatrix} \begin{bmatrix} 1 & -\bar{\tilde{\gamma}}_n \\ -\tilde{\gamma}_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \tag{30}$$

where

$$\lambda_n = \frac{|1 - \bar{\alpha}_n \alpha_{n-1}|}{1 - \alpha_n \bar{\alpha}_{n-1}} \frac{\overline{\phi_n^*(\alpha_{n-1})}}{|\phi_n^*(\alpha_{n-1})|} \eta_n \frac{\overline{\kappa_{n-1}}}{|\kappa_{n-1}|} = \frac{1 - \alpha_n \bar{\alpha}_{n-1}}{|1 - \bar{\alpha}_n \alpha_{n-1}|} \frac{\overline{\kappa_{n-1}}}{|\kappa_{n-1}|} \in \mathbb{T}. \tag{31}$$

We have obtained the following result ([6], Theorem 4.1.1, but with another normalization of the orthogonal rational functions):

Proposition 4.8 *The orthogonal rational functions are given by the formula*

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = T_n(z) \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} \qquad \forall n \geq 1,$$

with $T_n(z)$ defined as in (30).

A first application of this formula is to the location of the roots of the orthogonal rational functions. Note that by proposition 4.6, since the set of roots of ϕ_n is the image of the set of roots of ϕ_n^* by the map $z \mapsto 1/\bar{z}$, we already know that the roots are in the closed unit disk $\overline{\mathbb{D}}$.

Corollary 4.9 *The orthogonal rational functions ϕ_n have all their roots in $\mathbb{D}$.*

Proof By induction, we show that ϕ_n^* has no roots in $\overline{\mathbb{D}}$. This is clearly true for $n = 0$. If it is true for n, then the function $\frac{\phi_n}{\phi_n^*}$ is analytic in $\overline{\mathbb{D}}$ and by proposition 4.6, $\left|\frac{\phi_n}{\phi_n^*}\right| \leq 1$ in $\overline{\mathbb{D}}$. Using the previous recurrence formula on ϕ_{n+1}^*, we obtain that

$$\phi_{n+1}^* = \sqrt{\frac{1 - |\alpha_{n+1}|^2}{1 - |\alpha_n|^2}} \frac{1}{\sqrt{1 - |\tilde{\gamma}_{n+1}|^2}} \bar{\lambda}_{n+1} \frac{1 - \bar{\alpha}_n z}{1 - \bar{\alpha}_{n+1} z} \phi_n^* \left(1 - \tilde{\gamma}_{n+1} \zeta_n \frac{\phi_n}{\phi_n^*}\right).$$

Using the induction hypothesis, and since $|\tilde{\gamma}_{n+1}\zeta_n \frac{\phi_n}{\phi_n^*}| \leq |\tilde{\gamma}_{n+1}| < 1$ for all $z \in \overline{\mathbb{D}}$, the latter expression does not have any root in $\overline{\mathbb{D}}$.

■

The recurrence relation can be inverted in order to express ϕ_{n-1}, ϕ_{n-1}^* as functions of ϕ_n, ϕ_n^*.

Corollary 4.10 *The orthogonal rational functions are given by the reverse recurrence formula*

$$\begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix} = T_n^{-1}(z) \begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} \forall n \geq 1,$$

with $T_n^{-1}(z)$ equal to

$$T_n^{-1}(z) = \sqrt{\frac{1-|\alpha_{n-1}|^2}{1-|\alpha_n|^2}} \frac{1}{\sqrt{1-|\tilde{\gamma}_n|^2}} \frac{1-\bar{\alpha}_n z}{1-\bar{\alpha}_{n-1} z} \begin{bmatrix} \frac{1}{\zeta_{n-1}(z)} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \overline{\tilde{\gamma}_n} \\ \tilde{\gamma}_n & 1 \end{bmatrix} \begin{bmatrix} \overline{\lambda_n} & 0 \\ 0 & \lambda_n \end{bmatrix}.$$

Proof Immediate since λ_n is in $\mathbb{T}$ hence

$$\begin{aligned} T_n(z)^{-1} &= \sqrt{\frac{1-|\alpha_{n-1}|^2}{1-|\alpha_n|^2}} \sqrt{1-|\tilde{\gamma}_n|^2} \frac{1-\bar{\alpha}_n z}{1-\bar{\alpha}_{n-1} z} \begin{bmatrix} \frac{1}{\zeta_{n-1}(z)} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\overline{\tilde{\gamma}_n} \\ -\tilde{\gamma}_n & 1 \end{bmatrix}^{-1} \begin{bmatrix} \overline{\lambda_n} & 0 \\ 0 & \lambda_n \end{bmatrix} \\ &= \sqrt{\frac{1-|\alpha_{n-1}|^2}{1-|\alpha_n|^2}} \frac{1}{\sqrt{1-|\tilde{\gamma}_n|^2}} \frac{1-\bar{\alpha}_n z}{1-\bar{\alpha}_{n-1} z} \begin{bmatrix} \frac{1}{\zeta_{n-1}(z)} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \overline{\tilde{\gamma}_n} \\ \tilde{\gamma}_n & 1 \end{bmatrix} \begin{bmatrix} \overline{\lambda_n} & 0 \\ 0 & \lambda_n \end{bmatrix}. \end{aligned}$$

■

For $\omega \in \mathbb{D}$, we denote by $P(., \omega)$ the Poisson kernel

$$P(z, \omega) = \frac{1-|\omega|^2}{|z-\omega|^2}, \quad z \in \mathbb{T}.$$

Note that whenever u is harmonic in $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$, we have

$$u(\omega) = \int_{\mathbb{T}} u(z) P(z, \omega) dm(z).$$

This we call the Poisson identity for harmonic functions.
We now get the orthonormality of $\phi_0, \ldots, \phi_n$ with respect to another measure than μ ([6], Theorem 6.1.9).

Corollary 4.11 *The rational functions $\phi_0, \ldots, \phi_n$ are orthonormal in $L^2\left(\frac{P(., \alpha_n)}{|\phi_n|^2} dm\right)$.*

Proof Let $N = \int_{\mathbb{T}} \frac{P(.,\alpha_n)}{|\phi_n|^2} dm$. Then $\frac{P(.,\alpha_n)}{N|\phi_n|^2} dm$ is a probability measure. For $n \geq 0$ and $k < n$, we have

$$\begin{aligned} \int_{\mathbb{T}} \sqrt{N}\phi_n \overline{\sqrt{N}\phi_k} \frac{P(.,\alpha_n)}{N|\phi_n|^2} dm &= \int_{\mathbb{T}} \frac{\phi_{k*}}{\phi_{n*}} P(.,\alpha_n) dm \\ &= \int_{\mathbb{T}} \frac{\phi_k^*}{\phi_n^*} \zeta_{k+1} \ldots \zeta_n P(.,\alpha_n) dm \\ &= 0 \end{aligned}$$

because we can apply the Poisson identity since ${\phi_n}^*$ has no zero in $\overline{\mathbb{D}}$. We also have

$$\int_{\mathbb{T}} |\sqrt{N}\phi_n|^2 \frac{P(.,\alpha_n)}{N|\phi_n|^2} dm = \int_{\mathbb{T}} P(.,\alpha_n) dm = 1.$$

Therefore, $\sqrt{N}\phi_n$ is orthonormal to $\sqrt{N}\phi_0, \ldots, \sqrt{N}\phi_{n-1}$, that is to $\mathcal{L}_{n-1}$, with respect to the measure $\frac{P(.,\alpha_n)}{N|\phi_n|^2} dm$. But the reverse recurrence formula (corollary 4.10) together with (28) shows that the first $n-1$ orthogonal rational functions normalized by (23) are uniquely determined by the n-th orthogonal rational function and the (α_k). Therefore, the $\sqrt{N}\phi_k$, $0 \leq k \leq n$, are the orthonormal rational functions for the measure $\frac{P(.,\alpha_n)}{N|\phi_n|^2} dm$. In particular,

$$\int_{\mathbb{T}} |\sqrt{N}\phi_0|^2 \frac{P(.,\alpha_n)}{N|\phi_n|^2} dm = \int_{\mathbb{T}} \frac{P(.,\alpha_n)}{|\phi_n|^2} dm = 1.$$

Thus, $N = 1$, and the conclusion is immediate.

■

Iterating the recurrence formula, we obtain an expression of ϕ_n.

Corollary 4.12 *For $n \geq 1$, ϕ_n and ϕ_n^* are given by the relation:*

$$\begin{bmatrix} \phi_n \\ \phi_n^* \end{bmatrix} = \frac{\sqrt{1-|\alpha_n|^2}}{1-\bar{\alpha}_n z} \frac{1}{\Pi_n} \left(\prod_{k=n}^{k=1} \begin{bmatrix} \lambda_k & 0 \\ 0 & \overline{\lambda_k} \end{bmatrix} \begin{bmatrix} 1 & -\overline{\tilde{\gamma}_k} \\ -\tilde{\gamma}_k & 1 \end{bmatrix} \begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with

$$\Pi_n = \prod_{k=n}^{k=1} \sqrt{1-|\tilde{\gamma}_k|^2}.$$

Proof Immediate from proposition 4.8 since $\alpha_0 = 0$ and $\phi_0 = \phi_0^* = 1$.

■

4.4 Orthogonal rational functions of the second kind

As in [6], chapter 4, we now define the sequence (ψ_n) of orthogonal rational functions of the second kind. We shall see later that this sequence satisfies the same recurrence relations than ϕ_n, but with $\tilde{\gamma}_n$ replaced by $-\tilde{\gamma}_n$.

Definition 4.13 *Given μ, (α_k) and (ϕ_n) as before, we call orthogonal rational functions of the second kind the sequence ψ_n such that*

$$\begin{cases} \psi_0 = 1 \\ \psi_n(z) = \int_{\mathbb{T}} \frac{t+z}{t-z} \left(\phi_n(t) - \phi_n(z)\right) d\mu(t) \end{cases} .$$

We will see later that the ψ_n are indeed rational functions. The following proposition ([6], Lemma 4.2.2 and 4.2.3) is very useful for computations.

Proposition 4.14 *For $n \geq 1$, the functions (ψ_n) satisfy the formulas:*

$$\psi_n(z)g(z) = \int_{\mathbb{T}} \frac{t+z}{t-z} \left(\phi_n(t)g(t) - \phi_n(z)g(z)\right) d\mu(t)$$

for all g such that $g_* \in \mathcal{L}_{n-1}$, and moreover we have

$$-\psi_n^*(z)h(z) = \int_{\mathbb{T}} \frac{t+z}{t-z} \left(\phi_n^*(t)h(t) - \phi_n^*(z)h(z)\right) d\mu(t)$$

for all h such that $h_* \in \zeta_n \mathcal{L}_{n-1}$.

Proof We first prove the first equality. If g is constant, the result is immediate. We therefore suppose $n \geq 2$. Let $z \in \mathbb{D}$.
If $z = \alpha_k$ for some k, $1 \leq k \leq n-1$, $g(z) = \infty$. By definition, we have

$$\psi_n(\alpha_k) = \int_{\mathbb{T}} \frac{t+\alpha_k}{t-\alpha_k} \left(\phi_n(t) - \phi_n(\alpha_k)\right) d\mu(t).$$

But, since $\frac{t+\alpha_k}{t-\alpha_k} \in \mathcal{L}_{n-1}$ and $n \geq 2$, we get by orthogonality

$$\psi_n(\alpha_k) = -\phi_n(\alpha_k) \int_{\mathbb{T}} \frac{t+\alpha_k}{t-\alpha_k} d\mu(t)$$

which is the announced result when $g(z) = \infty$.
Suppose $z \neq \alpha_k$ for all k, $1 \leq k \leq n-1$. By density, it is enough to prove the result if $g(z)$ is analytic at z with $g(z) \neq 0$. In order to conclude, using the defintion of ψ_n, we just have to check that

$$\int \frac{t+z}{t-z} \phi_n(t) \frac{g(t)}{g(z)} d\mu(t) = \int \frac{t+z}{t-z} \phi_n(t) d\mu(t) \text{ whenever } g_* \in \mathcal{L}_{n-1}.$$

But $\frac{g(t)}{g(z)} - 1$ vanishes for $t = z$, therefore

$$\frac{g(t)}{g(z)} - 1 = (t - z)\frac{p}{\prod_{k=1}^{k=n-1}(t - \alpha_k)}$$

where p is a polynomial in t of degree at most $n - 2$. Thus,

$$\begin{aligned}
\int \frac{t+z}{t-z}\phi_n(t)\left(\frac{g(t)}{g(z)} - 1\right)d\mu(t) &= \int \frac{t+z}{t-z}(t-z)\frac{p(t)}{\prod_{k=1}^{k=n-1}(t-\alpha_k)}\phi_n(t)d\mu(t) \\
&= \int \frac{(t+z)p(t)}{\prod_{k=1}^{k=n-1}(t-\alpha_k)}\phi_n(t)d\mu(t) \\
&= \int \overline{\left(\frac{t^{n-1}\overline{(t+z)p(t)}}{\prod_{k=1}^{k=n-1}(1-\overline{\alpha_k}t)}\right)}\phi_n(t)d\mu(t) \\
&= 0
\end{aligned}$$

because, since $\bar{t} = \frac{1}{t}$ on $\mathbb{T}$ and $\deg p \leq n - 2$, we have on $\mathbb{T}$:

$$\frac{t^{n-1}\overline{(t+z)p(t)}}{\prod_{k=1}^{k=n-1}(1-\overline{\alpha_k}t)} = \frac{t^{n-1}\overline{(1/\bar{t}+z)p(1/\bar{t})}}{\prod_{k=1}^{k=n-1}(1-\overline{\alpha_k}t)} \in \mathcal{L}_{n-1}$$

Therefore, the first equality is proved.
Since $\mathcal{B}_n h$ is in $\mathcal{L}_{n-1}$, we get from the latter

$$\psi_n(z)\mathcal{B}_{n*}(z)h_*(z) = \int \frac{t+z}{t-z}\left(\phi_n(t)\mathcal{B}_{n*}(t)h_*(t) - \phi_n(z)\mathcal{B}_{n*}(z)h_*(z)\right)d\mu(t).$$

We conclude by taking the lower-$*$ conjugate in z of this expression.

■

We deduce from the following proposition that ψ_n is indeed a rational function (see [6], Theorem 4.2.4).

Proposition 4.15 *The sequences* (ϕ_n) *and* (ψ_n) *satisfy the recurrence relations:*

$$\begin{bmatrix} \phi_n & \psi_n \\ \phi_n^* & -\psi_n^* \end{bmatrix} = \frac{\sqrt{1-|\alpha_n|^2}}{1-\bar{\alpha}_n z}\frac{1}{\Pi_n}\left(\prod_{k=n}^{k=1}\begin{bmatrix} \lambda_k & 0 \\ 0 & \bar{\lambda}_k \end{bmatrix}\begin{bmatrix} 1 & -\overline{\tilde{\gamma}_k} \\ -\tilde{\gamma}_k & 1 \end{bmatrix}\begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

with

$$\Pi_n = \prod_{k=n}^{k=1}\sqrt{1-|\tilde{\gamma}_k|^2}.$$

In particular, ψ_n *is in* $\mathcal{L}_n$.

Proof From Corollary 4.12, we now that this relation holds for (ϕ_n, ϕ_n^*), so we just have to prove it for (ψ_n, ψ_n^*). We first check that this result for $n = 1$. As $\psi_0 = 1$, we want to prove that

$$\psi_1 = \beta_1 \frac{z + \overline{\tilde{\gamma}_1}}{1 - \overline{\alpha_1} z}$$

with

$$\beta_1 = \sqrt{\frac{1 - |\alpha_1|^2}{1 - |\tilde{\gamma}_1|^2}} \lambda_1.$$

We have

$$\begin{aligned}
\psi_1(z) &= \int_{\mathbb{T}} \frac{t+z}{t-z} (\phi_1(t) - \phi_1(z)) d\mu(t) \\
&= \beta_1 \int_{\mathbb{T}} \frac{t+z}{t-z} \left(\frac{t - \overline{\tilde{\gamma}_1}}{1 - \overline{\alpha_1} t} - \frac{z - \overline{\tilde{\gamma}_1}}{1 - \overline{\alpha_1} z} \right) d\mu(t) \\
&= \beta_1 \int_{\mathbb{T}} \frac{t+z}{t-z} \left(\frac{(t-z)(1 - \overline{\alpha_1 \tilde{\gamma}_1})}{(1 - \overline{\alpha_1} t)(1 - \overline{\alpha_1} z)} \right) d\mu(t) \\
&= \beta_1 \int_{\mathbb{T}} \frac{(t+z)(1 - \overline{\alpha_1 \tilde{\gamma}_1})}{(1 - \overline{\alpha_1} t)(1 - \overline{\alpha_1} z)} d\mu(t) \\
&= \beta_1 \frac{1 - \overline{\alpha_1 \tilde{\gamma}_1}}{1 - \overline{\alpha_1} z} \int_{\mathbb{T}} \frac{t+z}{1 - \overline{\alpha_1} t} d\mu(t).
\end{aligned}$$

As ϕ_1 is orthogonal to 1, we also have

$$\int_{\mathbb{T}} \frac{t}{1 - \overline{\alpha_1} t} d\mu(t) = \overline{\tilde{\gamma}_1} \int_{\mathbb{T}} \frac{1}{1 - \overline{\alpha_1} t} d\mu(t). \tag{32}$$

Therefore

$$\psi_1(z) = \beta_1 \frac{1 - \overline{\alpha_1 \tilde{\gamma}_1}}{1 - \overline{\alpha_1} z} (\overline{\tilde{\gamma}_1} + z) \int_{\mathbb{T}} \frac{1}{1 - \overline{\alpha_1} t} d\mu(t).$$

But, by (32),

$$\begin{aligned}
\int_{\mathbb{T}} \frac{1}{1 - \overline{\alpha_1} t} d\mu(t) &= 1 + \overline{\alpha_1} \int_{\mathbb{T}} \frac{t}{1 - \overline{\alpha_1} t} d\mu(t) \\
&= 1 + \overline{\alpha_1} \overline{\tilde{\gamma}_1} \int_{\mathbb{T}} \frac{1}{1 - \overline{\alpha_1} t} d\mu(t)
\end{aligned}$$

thus

$$\int_{\mathbb{T}} \frac{1}{1 - \overline{\alpha_1} t} d\mu(t) = \frac{1}{1 - \overline{\alpha_1} \overline{\tilde{\gamma}_1}}.$$

Therefore,

$$\begin{aligned}
\psi_1(z) &= \beta_1 \frac{1 - \overline{\alpha_1 \tilde{\gamma}_1}}{1 - \overline{\alpha_1} z} (\overline{\tilde{\gamma}_1} + z) \frac{1}{1 - \overline{\alpha_1} \overline{\tilde{\gamma}_1}} \\
&= \beta_1 \frac{z + \overline{\tilde{\gamma}_1}}{1 - \overline{\alpha_1} z}.
\end{aligned}$$

which is the result we want.
We now proceed by induction.
Assume $n > 1$. Proposition 4.14 gives us with n replaced by $n-1$ and $g = 1$ together with $h = \zeta_{n-1*}$,

$$\begin{bmatrix} \psi_{n-1}(z) \\ -\psi^*_{n-1}(z) \end{bmatrix} = \int \frac{t+z}{t-z} \left(\begin{bmatrix} \phi_{n-1}(t) \\ \frac{\zeta_{n-1}(z)}{\zeta_{n-1}(t)} \phi^*_{n-1}(t) \end{bmatrix} - \begin{bmatrix} \phi_{n-1}(z) \\ \phi^*_{n-1}(z) \end{bmatrix} \right) d\mu(t).$$

Multiplying by $T_n(z)$ whose definition was given in (30), we obtain

$$\begin{aligned} & T_n(z) \begin{bmatrix} \psi_{n-1}(z) \\ -\psi^*_{n-1}(z) \end{bmatrix} \\ = & \int \frac{t+z}{t-z} \left(T_n(z) \begin{bmatrix} \phi_{n-1}(t) \\ \frac{\zeta_{n-1}(z)}{\zeta_{n-1}(t)} \phi^*_{n-1}(t) \end{bmatrix} - \begin{bmatrix} \phi_n(z) \\ \phi^*_n(z) \end{bmatrix} \right) d\mu(t) \\ = & \int \frac{t+z}{t-z} \left(\frac{(1-\overline{\alpha_n}t)(1-\overline{\alpha_{n-1}}z)}{(1-\overline{\alpha_n}z)(1-\overline{\alpha_{n-1}}t)} T_n(t) \begin{bmatrix} \frac{\zeta_{n-1}(z)}{\zeta_{n-1}(t)} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{n-1}(t) \\ \frac{\zeta_{n-1}(z)}{\zeta_{n-1}(t)} \phi^*_{n-1}(t) \end{bmatrix} - \begin{bmatrix} \phi_n(z) \\ \phi^*_n(z) \end{bmatrix} \right) d\mu(t) \\ = & \int \frac{t+z}{t-z} \left(\frac{(1-\overline{\alpha_n}t)(z-\alpha_{n-1})}{(1-\overline{\alpha_n}z)(t-\alpha_{n-1})} \begin{bmatrix} \phi_n(t) \\ \phi^*_n(t) \end{bmatrix} - \begin{bmatrix} \phi_n(z) \\ \phi^*_n(z) \end{bmatrix} \right) d\mu(t). \end{aligned}$$

But, by proposition 4.14 applied with $g(z) = (1 - \bar{\alpha}_n z)/(z - \alpha_{n-1})$, the first row in the right handside of the last term is equal to ψ_n. So it only remains to prove that the second row is equal to $-\psi^*_n$. To this effect, observe that

$$\begin{aligned} & \int \frac{t+z}{t-z} \left(\frac{z-\alpha_{n-1}}{t-\alpha_{n-1}} - \frac{z-\alpha_n}{t-\alpha_n} \right) \frac{1-\overline{\alpha_n}t}{1-\overline{\alpha_n}z} \phi^*_n(t) d\mu(t) \\ = & \int \frac{t+z}{t-z} \left(\frac{(t-z)(\alpha_n-\alpha_{n-1})}{(t-\alpha_{n-1})(t-\alpha_n)} \right) \frac{1-\overline{\alpha_n}t}{1-\overline{\alpha_n}z} \phi^*_n(t) d\mu(t) \\ = & \int \frac{(1-\overline{\alpha_n}t)(t+z)(\alpha_n-\alpha_{n-1})}{(t-\alpha_n)(t-\alpha_{n-1})(1-\overline{\alpha_n}z)} \phi^*_n(t) d\mu(t) \\ = & \int \mathcal{B}_{n-1}(t) \frac{(t+z)(\alpha_n-\alpha_{n-1})}{(t-\alpha_{n-1})(1-\overline{\alpha_n}z)} \overline{\phi_n(t)} d\mu(t) \\ = & \ 0 \end{aligned}$$

because

$$\mathcal{B}_{n-1}(t) \frac{(t+z)(\alpha_n-\alpha_{n-1})}{(t-\alpha_{n-1})(1-\overline{\alpha_n}z)} \in \mathcal{L}_{n-1}$$

as a function of t for fixed $z \in \overline{\mathbb{D}}$. Therefore,

$$\begin{aligned} & \int \frac{t+z}{t-z} \left(\frac{(1-\overline{\alpha_n}t)(z-\alpha_{n-1})}{(1-\overline{\alpha_n}z)(t-\alpha_{n-1})} \phi^*_n(t) - \phi^*_n(z) \right) d\mu(t) \\ = & \int \frac{t+z}{t-z} \left(\frac{(1-\overline{\alpha_n}t)(z-\alpha_n)}{(1-\overline{\alpha_n}z)(t-\alpha_n)} \phi^*_n(t) - \phi^*_n(z) \right) d\mu(t) \\ = & -\psi^*_n(z) \end{aligned}$$

by proposition 4.14 with $h(z) = (1 - \bar{\alpha}_n z)/(z - \alpha_n)$. This achieves the induction step.

■

We now show that the sequence (ψ_n) satisfies the same recurrence relations than (ϕ_n), but with $\tilde{\gamma}_n$ replaced by $-\tilde{\gamma}_n$:

Corollary 4.16 *The sequence ψ_n satisfies the recurrence relations:*

$$\begin{bmatrix} \psi_n \\ \psi_n^* \end{bmatrix} = \frac{\sqrt{1-|\alpha_n|^2}}{1-\bar{\alpha}_n z}\frac{1}{\Pi_n}\left(\prod_{k=n}^{k=1}\begin{bmatrix} \lambda_k & 0 \\ 0 & \bar{\lambda}_k \end{bmatrix}\begin{bmatrix} 1 & \overline{\tilde{\gamma}_k} \\ \tilde{\gamma}_k & 1 \end{bmatrix}\begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Proof Note that, by Proposition 4.15,

$$\begin{bmatrix} \psi_n \\ -\psi_n^* \end{bmatrix} = \frac{\sqrt{1-|\alpha_n|^2}}{1-\bar{\alpha}_n z}\frac{1}{\Pi_n}\left(\prod_{k=n}^{k=1}\begin{bmatrix} \lambda_k & 0 \\ 0 & \bar{\lambda}_k \end{bmatrix}\begin{bmatrix} 1 & -\overline{\tilde{\gamma}_k} \\ -\tilde{\gamma}_k & 1 \end{bmatrix}\begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Therefore, since $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^2 = Id$,

$$\begin{aligned}\begin{bmatrix} \psi_n \\ \psi_n^* \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \psi_n \\ -\psi_n^* \end{bmatrix} \\ &= \frac{\sqrt{1-|\alpha_n|^2}}{1-\bar{\alpha}_n z}\frac{1}{\Pi_n} \\ &\quad \left(\prod_{k=n}^{k=1}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \lambda_k & 0 \\ 0 & \bar{\lambda}_k \end{bmatrix}\begin{bmatrix} 1 & -\overline{\tilde{\gamma}_k} \\ -\tilde{\gamma}_k & 1 \end{bmatrix}\begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right)\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{\sqrt{1-|\alpha_n|^2}}{1-\bar{\alpha}_n z}\frac{1}{\Pi_n}\left(\prod_{k=n}^{k=1}\begin{bmatrix} \lambda_k & 0 \\ 0 & \bar{\lambda}_k \end{bmatrix}\begin{bmatrix} 1 & \overline{\tilde{\gamma}_k} \\ \tilde{\gamma}_k & 1 \end{bmatrix}\begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} 1 \\ 1 \end{bmatrix}.\end{aligned}$$

■

Proposition 4.17 *For all z in $\overline{\mathbb{D}}$, it holds that*

$$\phi_n(z)\psi_n^*(z) + \phi_n^*(z)\psi_n(z) = 2\frac{1-|\alpha_n|^2}{(1-\overline{\alpha_n}z)(z-\alpha_n)}z\mathcal{B}_n(z).$$

Proof Taking determinants in the relation of proposition 4.15, we get

$$\begin{aligned}\phi_n(z)\psi_n^*(z) + \phi_n^*(z)\psi_n(z) &= 2\frac{1-|\alpha_n|^2}{(1-\overline{\alpha_n}z)^2}\prod_{k=1}^{k=n}|\lambda_k|^2\zeta_{k-1}(z) \\ &= 2\frac{1-|\alpha_n|^2}{(1-\overline{\alpha_n}z)^2}z\mathcal{B}_n(z)\frac{1-\overline{\alpha_n}z}{z-\alpha_n} \\ &= 2\frac{1-|\alpha_n|^2}{(1-\overline{\alpha_n}z)(z-\alpha_n)}z\mathcal{B}_n(z).\end{aligned}$$

■

In particular, we have:

Corollary 4.18 *For $z \in \mathbb{T}$, one has*

$$\phi_n(z)\psi_n^*(z) + \phi_n^*(z)\psi_n(z) = 2\mathcal{B}_n(z)P(z,\alpha_n) \tag{33}$$

where $P(z,\alpha_n) = \frac{1-|\alpha_n|^2}{|z-\alpha_n|^2}$ *is the Poisson kernel at* α_n.

5 Link between orthogonal rational functions and Wall rational functions

If we glance at Propositions 3.9 and 4.15, we see that the recurrence formulas for the Wall rational functions A_n, B_n and for the orthogonal rational functions ϕ_n, ψ_n look quite similar. In this section, we will use this similarity to prove a generalized Geronimus theorem (see [13] for the original version). We first need to associate to a Schur function f a measure μ: we use for this the Herglotz transform. Next, we prove a Geronimus theorem which states the relation between the Szegő parameters of μ and the Schur parameters of f ([21]).

5.1 The Herglotz transform

We denote by F the Herglotz transform of μ:

$$F(z) = \int_{\mathbb{T}} \frac{\xi + z}{\xi - z} d\mu(\xi). \tag{34}$$

We have ([7], Theorem 3.4):

Proposition 5.1 *The Herglotz transform is related to the orthogonal rational functions ϕ_n, ψ_n associated with μ by a relation of the form*

$$F(z) = \frac{\psi_n^*(z)}{\phi_n^*(z)} + \frac{z\mathcal{B}_n(z)u(z)}{\phi_n^*(z)}$$

where u is an analytic function in $\mathbb{D}$.

Proof Proposition 4.14 gives us with $h(z) = 1/\mathcal{B}_n(z)$

$$\begin{aligned}\frac{F(z)\phi_n^*(z) - \psi_n^*(z)}{\mathcal{B}_n(z)} &= \int \frac{t+z}{t-z}\frac{\phi_n^*(z)}{\mathcal{B}_n(z)}d\mu(t) + \int \frac{t+z}{t-z}\left(\frac{\phi_n^*(t)}{\mathcal{B}_n(t)} - \frac{\phi_n^*(z)}{\mathcal{B}_n(z)}\right)d\mu(t)\\ &= \int \frac{t+z}{t-z}\frac{\phi_n^*(t)}{\mathcal{B}_n(t)}d\mu(t).\end{aligned}$$

This is a Cauchy integral, so it is a holomorphic function of z in $\mathbb{D}$. Evaluating this function at 0, we get

$$\int \frac{\phi_n^*(t)}{\mathcal{B}_n(t)} d\mu(t) = \int \overline{\phi_n(t)} d\mu(t) = 0$$

by orthogonality of ϕ_n and 1. The conclusion is then immediate.

∎

The Riesz-Herglotz theorem [25] states that the Herglotz transform is a one-to-one mapping between the set of probability measures on $\mathbb{T}$ and the set of analytic functions F in $\mathbb{D}$ satisfying

$$F(0) = 1, \quad ReF(z) > 0, \quad z \in \mathbb{D}.$$

$\frac{F-1}{F+1}$ is a Schur function that vanishes at zero, so the Schwarz lemma implies that

$$f(z) = \frac{1}{z} \frac{F(z) - 1}{F(z) + 1}$$

is also a Schur function. Therefore, we obtain a one-to-one correspondence between probability measures μ on $\mathbb{T}$ and Schur functions f via the relation

$$\int_{\mathbb{T}} \frac{\xi + z}{\xi - z} d\mu(\xi) = \frac{1 + zf(z)}{1 - zf(z)}. \tag{35}$$

For fixed $z \in \mathbb{D}$, we denote by Ω_z the map

$$\Omega_z : \omega \mapsto \frac{1}{z} \frac{\omega - 1}{\omega + 1}.$$

Note that $f(z) = \Omega_z(F(z))$.

Definition 5.2 *The function f associated to μ through (35) will be called the Schur function of μ.*

Applying Fatou's theorem on nontangential limits of harmonic functions ([12]) to the real part of (35), we obtain an expression for the Lebesgue derivative μ' of the measure μ in terms of f:

$$\mu'(\xi) = \frac{1 - |f(\xi)|^2}{|1 - \xi f(\xi)|^2} \text{ a.e. on } \mathbb{T}. \tag{36}$$

Since $1 - zf(z)$ is a non-zero function of H^∞, it cannot vanish on a set of positive measure. Therefore, $\mu' > 0$ a.e. on $\mathbb{T}$ if and only if $|f| < 1$ a.e. on $\mathbb{T}$.
The Schur parameters of the function f associated with μ can be computed from the orthogonal rationals functions of μ:

Proposition 5.3 *$f(z)$ and $\Omega_z\left(\frac{\psi_n^*(z)}{\phi_n^*(z)}\right)$ have the same first n Schur parameters.*

Proof From Proposition 5.1, we get

$$F^{(i)}(z) = \left(\frac{\psi_n^*(z)}{\phi_n^*(z)}\right)^{(i)} + \left(\frac{z\mathcal{B}_n(z)u(z)}{\phi_n^*(z)}\right)^{(i)}, \quad i \geq 0. \tag{37}$$

Let j be an integer such that $0 \leq j \leq n-1$. We denote by m_{j+1} the multiplicity of α_{j+1} at the n-th step (see Proposition 3.2). Then, if $0 \leq i < m_{j+1}$, since $\mathcal{B}_n(z) = h(z)\prod_{k=1}^{m_{j+1}}(z-\alpha_{j+1})$ with $h \in \mathcal{L}_n$, we have $B_n^{(i)}(\alpha_{j+1}) = 0$. Therefore, using (37), we obtain

$$F^{(i)}(\alpha_{j+1}) = \left(\frac{\psi_n^*}{\phi_n^*}\right)^{(i)}(\alpha_{j+1}).$$

Since $f(z) = \Omega_z(F(z))$, we conclude using Proposition 3.2.

■

5.2 A Geronimus theorem

Geronimus was first to express the relation between the classical Schur algorithm applied to the Schur function of a measure μ and the orthogonal polynomials of μ. In [21], the connection between the Geronimus parameters of the orthogonal rational functions and the Schur parameters of a multipoint Schur algorithm is detailed. However, the normalisation of the orthogonal rational functions in this reference is different from ours, so the link is made with a multipoint Schur algorithm without the rotations c_k. We chose to keep our generalized multipoint algorithm and we give below another proof of the Geronimus theorem.

Theorem 5.4 *Fix $(\alpha_k)_{k\geq 1} \in \mathbb{D}$ and $f \in \mathcal{S}$.*
We associate with f the measure μ given by (35). We denote by $(\tilde{\gamma}_k)_{k\geq 1}$ the Geronimus parameters of μ (see (28)), and by λ_k the elements of $\mathbb{T}$ defined by (31).
If the parameters $(c_k)_{k\geq 1}$ of the multipoint Schur algorithm are defined by

$$c_k = \lambda_k^2, \quad c_0 = 1,$$

then the Geronimus parameters $(\tilde{\gamma}_k)_{k\geq 1}$ and the Schur parameters $(\gamma_k)_{k\in\mathbb{N}}$ of f are related by

$$\tilde{\gamma}_{k+1} = \gamma_k \text{ for all } k \geq 0.$$

Proof We first study the connection between the recurrence formulas. From proposition 4.15, we have

$$\begin{aligned}
&\begin{bmatrix} \phi_{n+1}(z) & \psi_{n+1}(z) \\ \phi^*_{n+1}(z) & -\psi^*_{n+1}(z) \end{bmatrix} \\
&= \Delta_{n+1} \left(\prod_{k=n+1}^{k=1} \begin{bmatrix} \lambda_k^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\overline{\tilde{\gamma}_k} \\ -\tilde{\gamma}_k & 1 \end{bmatrix} \begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\
&= \Delta_{n+1} \left(\prod_{k=n+1}^{k=1} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_k^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \overline{\tilde{\gamma}_k} \\ \tilde{\gamma}_k & 1 \end{bmatrix} \begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}
\end{aligned}$$

with

$$\Delta_{n+1} = \frac{\sqrt{1-|\alpha_{n+1}|^2}}{1-\bar{\alpha}_{n+1} z} \frac{\prod_{k=1}^{n+1} \bar{\lambda}_k}{\prod_{k=1}^{n+1} \sqrt{1-|\tilde{\gamma}_k|^2}}.$$

Therefore, if the parameters c_k are taken such that $c_k = \lambda_k^2$ for all $k \geq 1$ and if $\frac{U_n}{V_n}$ stands for the n-th convergent of a Schur function with parameters $\gamma_k := \tilde{\gamma}_{k+1}$ for all $k \geq 0$ (such a function exists because of Corollary 3.6), we get from Proposition 3.9 the following expression of ϕ_n, ψ_n with respect to U_n, V_n,

$$\begin{aligned}
&\begin{bmatrix} \phi_{n+1}(z) & \psi_{n+1}(z) \\ \phi^*_{n+1}(z) & -\psi^*_{n+1}(z) \end{bmatrix} \\
&= \Sigma_n \Delta_{n+1} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} c_{n+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_n^* & U_n^* \\ U_n & V_n \end{bmatrix} \begin{bmatrix} \zeta_0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\
&= \Sigma_n \Delta_{n+1} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} c_{n+1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -zV_n^* + U_n^* & -zV_n^* - U_n^* \\ -zU_n + V_n & -zU_n - V_n \end{bmatrix}
\end{aligned}$$

with $\Sigma_n = \prod_{k=1}^n \lambda_k$.
Since

$$\Sigma_n \prod_{k=1}^{n+1} \bar{\lambda}_k = \left(\prod_{k=1}^{n} \lambda_k \right) \prod_{k=1}^{n+1} \bar{\lambda}_k = \left(\prod_{k=1}^{n} |\lambda_k| \right) \bar{\lambda}_{n+1} = \bar{\lambda}_{n+1}$$

and $c_{n+1} = \lambda_{n+1}^2$, we obtain

$$\begin{aligned}
&\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{n+1}(z) & \psi_{n+1}(z) \\ \phi^*_{n+1}(z) & -\psi^*_{n+1}(z) \end{bmatrix} \\
&= \frac{\sqrt{1-|\alpha_{n+1}|^2}}{1-\bar{\alpha}_{n+1} z} \frac{1}{\prod_{k=1}^{n+1} \sqrt{1-|\tilde{\gamma}_k|^2}} \begin{bmatrix} \lambda_{n+1} & 0 \\ 0 & \bar{\lambda}_{n+1} \end{bmatrix} \begin{bmatrix} -zV_n^* + U_n^* & -zV_n^* - U_n^* \\ -zU_n + V_n & -zU_n - V_n \end{bmatrix}.
\end{aligned} \tag{38}$$

In particular, we have

$$\frac{\psi^*_{n+1}}{\phi^*_{n+1}} = \frac{1 + z\frac{U_n}{V_n}}{1 - z\frac{U_n}{V_n}} \tag{39}$$

so

$$\frac{U_n(z)}{V_n(z)} = \Omega_z \left(\frac{\psi^*_{n+1}(z)}{\phi^*_{n+1}(z)} \right).$$

Then, from proposition 5.3, $\frac{U_n}{V_n}$ has the same first $n+1$ Schur parameters as the Schur function f of the measure μ. This gives the expected result.

■

Note that a consequence of the theorem is that the elements U_n and V_n of the proof are equal to the Wall rational functions A_n and B_n of f. In particular, equations (38) and (39) gives us

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{n+1}(z) & \psi_{n+1}(z) \\ \phi^*_{n+1}(z) & -\psi^*_{n+1}(z) \end{bmatrix}$$
$$= \frac{\sqrt{1-|\alpha_{n+1}|^2}}{1-\bar{\alpha}_{n+1}z} \frac{1}{\prod_{k=1}^{n+1} \sqrt{1-|\tilde{\gamma}_k|^2}} \begin{bmatrix} \lambda_{n+1} & 0 \\ 0 & \bar{\lambda}_{n+1} \end{bmatrix} \begin{bmatrix} -zB^*_n + A^*_n & -zB^*_n - A^*_n \\ -zA_n + B_n & -zA_n - B_n \end{bmatrix} \tag{40}$$

and

$$\frac{\psi^*_{n+1}}{\phi^*_{n+1}} = \frac{1+z\frac{A_n}{B_n}}{1-z\frac{A_n}{B_n}}. \tag{41}$$

5.3 Consequences of the Geronimus theorem

The following corollary to Theorem 5.4 gives the expression of the measure associated to the Wall rational functions by the Herglotz transform. This is a generalization to the multipoint case of [20], Corollary 5.2.

Corollary 5.5 $\frac{A_n}{B_n}$ *is the Schur function of the measure* $\frac{P(.,\alpha_{n+1})}{|\phi_{n+1}|^2}dm$.

Proof Indeed, by (33), we have on $\mathbb{T}$:

$$\begin{aligned} Re\left(\frac{\psi^*_{n+1}}{\phi^*_{n+1}}\right) &= \frac{\overline{\mathcal{B}_{n+1}}\left(\psi^*_{n+1}\phi_{n+1} + \phi^*_{n+1}\psi_{n+1}\right)}{2\left|\phi_{n+1}\right|^2} \\ &= \frac{P(.,\alpha_{n+1})}{|\phi_{n+1}|^2}. \end{aligned}$$

Thus, $\frac{\psi^*_{n+1}}{\phi^*_{n+1}}$ and $\int \frac{t+z}{t-z}\frac{P(t,\alpha_{n+1})}{|\phi_{n+1}(t)|^2}dm(t)$ are two analytic functions in $\mathbb{D}$ with the same real part, therefore they are related by

$$\frac{\psi^*_{n+1}}{\phi^*_{n+1}} = \int \frac{t+z}{t-z}\frac{P(t,\alpha_{n+1})}{|\phi_{n+1}(t)|^2}dm(t) + ic$$

where c is a real constant. So by (41),

$$\frac{1+z\frac{A_n}{B_n}}{1-z\frac{A_n}{B_n}} = \int \frac{t+z}{t-z}\frac{P(t,\alpha_{n+1})}{|\phi_{n+1}(t)|^2}dm(t) + ic.$$

Evaluating the above expression at 0 gives us

$$1 = \int \frac{P(., \alpha_{n+1})}{|\phi_{n+1}|^2} dm + ic.$$

Since the integral is real, $c = 0$.

■

In view of Corollary 4.16, the Geronimus theorem also leads to another definition of the orthogonal rational functions of the second kind:

Corollary 5.6 *Up to a normalization, the orthogonal rational functions of the second kind associated to f (or F) are the orthogonal rational functions of the first kind associated to $-f$ (or $\frac{1}{F}$).*

The following theorem gives a useful relation between the Lebesgue derivative μ' of the measure μ, the Schur functions f_n and the orthogonal rational functions ϕ_n. This is a generalization to the multipoint case of [20], Theorem 2.

Theorem 5.7 *Let (ϕ_n) be the orthogonal rational functions of a probability measure μ associated to a sequence (α_n), and (f_n) the Schur functions associated to μ with the choice $c_n = \lambda_n^2$. Then*

$$\mu' = \frac{1 - |f_n|^2}{|1 - \overline{c_n} \zeta_n \frac{\phi_n}{\phi_n^*} f_n|^2} \frac{P(., \alpha_n)}{|\phi_n|^2} \quad a.e. \ on \ \mathbb{T}.$$

Proof From Theorem 3.12, we have:

$$\begin{aligned} 1 - |f|^2 &= 1 - \left| \frac{A_n + \zeta_{n+1} B_n^* f_{n+1}}{B_n + \zeta_{n+1} A_n^* f_{n+1}} \right|^2 \\ &= \frac{|B_n + \zeta_{n+1} A_n^* f_{n+1}|^2 - |A_n + \zeta_{n+1} B_n^* f_{n+1}|^2}{|B_n + \zeta_{n+1} A_n^* f_{n+1}|^2}. \end{aligned} \tag{42}$$

Note that on $\mathbb{T}$, $A_n^* \overline{B_n} = \overline{A_n} B_n^*$ so that

$$\zeta_{n+1} A_n^* f_{n+1} \overline{B_n} + B_n \overline{\zeta_{n+1} A_n^* f_{n+1}} - \overline{A_n} \zeta_{n+1} B_n^* f_{n+1} - A_n \overline{\zeta_{n+1} B_n^* f_{n+1}} = 0.$$

Therefore, on expanding (42), we find that

$$1 - |f|^2 = \frac{(|B_n|^2 - |A_n|^2)(1 - |f_{n+1}|^2)}{|B_n + \zeta_{n+1} A_n^* f_{n+1}|^2}.$$

Furthermore, by Corollary 3.10, we obtain

$$1 - |f|^2 = \frac{\omega_n (1 - |f_{n+1}|^2)}{|B_n + \zeta_{n+1} A_n^* f_{n+1}|^2} \tag{43}$$

where

$$\omega_n = \prod_{k=0}^{k=n} (1 - |\gamma_k|^2).$$

Using again Theorem 3.12, we get

$$\begin{aligned} |1 - zf|^2 &= \left|1 - \frac{zA_n + \zeta_{n+1} z B_n^* f_{n+1}}{B_n + \zeta_{n+1} A_n^* f_{n+1}}\right|^2 \\ &= \left|\frac{B_n - zA_n + \zeta_{n+1} f_{n+1}(A_n^* - zB_n^*)}{B_n + \zeta_{n+1} A_n^* f_{n+1}}\right|^2. \end{aligned}$$

In another connection, we deduce from (40) and Theorem 5.4 that

$$\begin{cases} zB_n^* - A_n^* &= \frac{1-\bar{\alpha}_{n+1} z}{\sqrt{1-|\alpha_{n+1}|^2}} \sqrt{\omega_n} \overline{\lambda_{n+1}} \phi_{n+1} \\ B_n - zA_n &= \frac{1-\bar{\alpha}_{n+1} z}{\sqrt{1-|\alpha_{n+1}|^2}} \sqrt{\omega_n} \lambda_{n+1} \phi_{n+1}^* \end{cases}$$

and therefore

$$|1 - zf|^2 = \left|\frac{1 - \bar{\alpha}_{n+1} z}{\sqrt{1 - |\alpha_{n+1}|^2}} \sqrt{\omega_n}\right|^2 \left|\frac{\lambda_{n+1}\phi_{n+1}^* - \zeta_{n+1} f_{n+1} \overline{\lambda_{n+1}} \phi_{n+1}}{B_n + \zeta_{n+1} A_n^* f_{n+1}}\right|^2 \tag{44}$$

$$= \omega_n \frac{|1 - \bar{\alpha}_{n+1} z|^2}{1 - |\alpha_{n+1}|^2} \left|\frac{\lambda_{n+1}\phi_{n+1}^* - \zeta_{n+1} f_{n+1} \overline{\lambda_{n+1}} \phi_{n+1}}{B_n + \zeta_{n+1} A_n^* f_{n+1}}\right|^2. \tag{45}$$

From what precedes, we deduce that

$$\frac{1 - |f|^2}{|1 - zf|^2} = \frac{1 - |f_{n+1}|^2}{|\phi_{n+1}^* - \overline{c_{n+1}} \zeta_{n+1} f_{n+1} \phi_{n+1}|^2} \frac{1 - |\alpha_{n+1}|^2}{|1 - \bar{\alpha}_{n+1} z|^2}.$$

Since $\mu'(\xi) = \frac{1-|f(\xi)|^2}{|1-\xi f(\xi)|^2}$ a.e. on $\mathbb{T}$ by (36) and $|\phi_{n+1}^*| = |\phi_{n+1}|$ on $\mathbb{T}$, we obtain

$$\mu' = \frac{1 - |f_{n+1}|^2}{|\phi_{n+1}|^2 |1 - \overline{c_{n+1}} \zeta_{n+1} \frac{\phi_{n+1}}{\phi_{n+1}^*} f_{n+1}|^2} \frac{1 - |\alpha_{n+1}|^2}{|\xi - \alpha_{n+1}|^2} \text{ a.e. on } \mathbb{T}.$$

■

6 Some asymptotic properties

In [20], various kinds of convergence for the rational functions $\frac{A_n}{B_n}$ are studied in the case of the classical Schur algorithm. There, it is in particular shown that ([20], Theorem 1):

If $\alpha_k = 0$ for every $k \geq 0$, then $|f| < 1$ a.e. on $\mathbb{T}$ if and only if $\lim_n \int_{\mathbb{T}} |f_n|^2 dm = 0$.

In this section, we study asymptotic properties of the Schur functions f_n and of the Wall rational functions A_n/B_n. Except for an "asymptotic-BMO-type" convergence of the Schur functions f_n, these are mainly generalizations of the results of Khrushchev where errors are integrated against the Poisson kernel of α_n rather than the Lebesgue measure. The difficulty here comes from the fact that we let the points go to the circle.
In order to prove the convergence respect to the Poincaré metric, we first need to solve a Szegő-type problem.

6.1 A Szegő-type problem

6.1.1 Generalities

We denote by μ' the Lebesgue derivative of the positive measure μ.

Definition 6.1 *A measure μ is called a Szegő measure if* $\log(\mu') \in L^1(\mathbb{T})$.

Let μ be a Szegő measure and let S be the Szegő function of μ:

$$S(z) = \exp\left(\frac{1}{2}\int_{\mathbb{T}} \frac{t+z}{t-z}\log(\mu')dm(t)\right).$$

The Szegő function is outer ([12]) and satisfies $|S|^2 = \mu'$ almost everywhere on $\mathbb{T}$.
Szegő proved ([28]) the following relation between the orthonormal polynomials ϕ_n of an absolutely continuous Szegő measure and the Szegő function S :

$$\lim_n \phi_n^*(z)S(z) = 1 \text{ locally uniformly in } \mathbb{D}.$$

This was later extended to non-absolutely continuous Szegő measures (see for example [22]).
A generalization of this theorem is given in [6] (Theorem 9.6.9) for orthogonal rational functions :
If μ is Szegő and if the points (α_n) are compactly included in $\mathbb{D}$, then we have locally uniformly in $\mathbb{D}$

$$\lim_n \left|\frac{S(z)\phi_n^*(z)(1-\bar{\alpha}_n z)}{\sqrt{1-|\alpha_n^2|}}\right| = 1.$$

Szegő also proved that the convergence of the orthonormal polynomials is uniform on the unit circle if the Lebesgue derivative of the measure is everywhere strictly positive on $\mathbb{T}$ and Lipschitz-Dini continuous, i.e. satisfies

$$|\mu'(\theta+\delta) - \mu'(\theta)| < L|\log(\delta)|^{-1-\lambda}$$

where L and λ are fixed positive numbers. Our study is akin to this: indeed, we will prove that if μ is absolutely continuous and Szegő, and if $\sum_{k=0}^{\infty}(1-|\alpha_k|) = \infty$, then the orthogonal rational functions ϕ_n satisfy

$$\lim_n |\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2) = 1$$

as soon as μ' is strictly positive and Dini-continuous. We do not assume here that the α_n are compactly included in $\mathbb{D}$.
A direct consequence of this result is that, under the above hypotheses and if $\lim_n |\alpha_n| = 1$, then $|\phi_n^*(\alpha_n)|$ diverges at the same rate as $(1-|\alpha_n|^2)^{-1}$.
The main tools we will use are reproducing kernels (see section 4.1) and some facts from rational approximation.

6.1.2 An approximation problem

We recall that π_n is defined in (3). We denote by $\mathcal{P}_n\left(\frac{d\mu}{|\pi_n|^2}\right)$ the subspace of $L^2\left(\frac{d\mu}{|\pi_n|^2}\right)$ of polynomials of degree at most n and by $H^2\left(\frac{d\mu}{|\pi_n|^2}\right)$ the closure of the polynomials in $L^2\left(\frac{d\mu}{|\pi_n|^2}\right)$.
The idea here is to express $|\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2)$ in terms of reproducing kernels of the spaces $\mathcal{P}_n\left(\frac{d\mu}{|\pi_n|^2}\right)$ and $H^2\left(\frac{d\mu}{|\pi_n|^2}\right)$. In what follows, we will sometime use the notation $d\mu_n$ for $\frac{d\mu}{|\pi_n|^2}$.

Proposition 6.2 *Let μ be an absolutely continuous Szegő measure. Then, the reproducing kernel E_n of $H^2\left(\frac{d\mu}{|\pi_n|^2}\right)$ is equal to*

$$E_n(\xi,\omega) = \frac{1}{1-\xi\bar{\omega}}\frac{\pi_n(\xi)\overline{\pi_n(\omega)}}{S(\xi)\overline{S(\omega)}}.$$

Proof First of all, it is clear that $E_n(.,\omega)$ is in $H^2(d\mu_n)$ for a fixed ω in $\mathbb{D}$ because on the one hand, $\frac{\pi_n(\xi)}{1-\xi\bar{\omega}}$ can be uniformly approximated by polynomials in $\overline{\mathbb{D}}$, and in the other hand, the fact that S is outer implies by the Beurling theorem ([12]) that there is a sequence (p_k) of polynomials such that $\lim_k \|1-p_kS\|_{L^2(dm)} = 0$. Then,

$$\begin{aligned}\int_{\mathbb{T}}\left(\frac{1}{S}-p_k\right)\frac{d\mu}{|\pi_n|^2} &= \int_{\mathbb{T}}\left(\frac{1}{S}-p_k\right)\frac{|S|^2dm}{|\pi_n|^2} = \int_{\mathbb{T}}(1-p_kS)\,\bar{S}\frac{dm}{|\pi_n|^2}\\ &\leq \frac{\|S\|_{L^2(dm)}\|1-p_kS\|_{L^2(dm)}}{\inf_{\mathbb{T}}|\pi_n|^2}\end{aligned}$$

by the Schwartz inequality. Therefore, we get $\lim_k \|p_k - 1/S\|_{L^2(d\mu_n)} = 0$.
Next, let q be a polynomial. We have

$$\begin{aligned}\int_{\mathbb{T}} q(t)\overline{\left(\frac{1}{1-t\bar{\omega}}\frac{\pi_n(t)\overline{\pi_n(\omega)}}{S(t)\overline{S(\omega)}}\right)}\frac{d\mu(t)}{|\pi_n(t)|^2} &= \int_{\mathbb{T}} q(t)\frac{1}{1-\bar{t}\omega}\frac{\overline{\pi_n(t)}\pi_n(\omega)}{\overline{S(t)}S(\omega)}\frac{|S(t)|^2dm(t)}{|\pi_n(t)|^2}\\ &= \int_{\mathbb{T}}\frac{q(t)}{t-\omega}t\frac{\pi_n(\omega)}{\pi_n(t)}\frac{S(t)}{S(\omega)}dm(t)\\ &= \frac{\pi_n(\omega)}{S(\omega)}\int_{\mathbb{T}}\frac{q(t)S(t)}{(t-\omega)\pi_n(t)}tdm(t).\end{aligned}$$

As $\frac{qS}{\pi_n}$ is in H^2, we obtain by the Cauchy theorem that

$$\int_{\mathbb{T}} \frac{q(t)S(t)}{(t-\omega)\pi_n(t)} t dm(t) = \frac{q(\omega)S(\omega)}{\pi_n(\omega)}.$$

Thus, we get

$$\int_{\mathbb{T}} q(t)\overline{\left(\frac{1}{1-t\bar{\omega}}\frac{\pi_n(t)\overline{\pi_n(\omega)}}{S(t)\overline{S(\omega)}}\right)}\frac{d\mu(t)}{|\pi_n(t)|^2} = q(\omega) \quad \text{for every } q \text{ polynomial.}$$

By density, this is true for every f in $H^2(d\mu_n)$. As the reproducing kernel is unique, the conclusion is immediate.

■

Proposition 6.3 *Let R_n be the reproducing kernel of $\mathcal{P}_n\left(\frac{d\mu}{|\pi_n|^2}\right)$. Then*

$$|\pi_n\phi_n^*| = \frac{|R_n(.,\alpha_n)|}{\|R_n(.,\alpha_n)\|_{L^2(d\mu_n)}}.$$

Proof Let p_{n-1} be a polynomial of degree at most $n-1$. As ϕ_n is orthogonal to $\mathcal{L}_{n-1}$, we have

$$\int_{\mathbb{T}} \overline{\phi_n}\frac{p_{n-1}}{\pi_{n-1}} d\mu = 0.$$

But

$$\begin{aligned}
\int_{\mathbb{T}} \overline{\phi_n}\frac{p_{n-1}}{\pi_{n-1}} d\mu &= \int_{\mathbb{T}} \overline{\phi_n(t)}\frac{p_{n-1}(t)(1-\bar{\alpha}_n t)}{\pi_n(t)} d\mu(t) \\
&= \int_{\mathbb{T}} \frac{\phi_n^*(t)}{\mathcal{B}_n(t)}\frac{p_{n-1}(t)(1-\bar{\alpha}_n t)}{\pi_n(t)} d\mu(t) \\
&= \int_{\mathbb{T}} \phi_n^*(t)\frac{\pi_n(t)}{t^n\overline{\pi_n(t)}}\frac{p_{n-1}(t)(1-\bar{\alpha}_n t)}{\pi_n(t)} d\mu(t) \\
&= \int_{\mathbb{T}} \pi_n(t)\phi_n^*(t)\bar{t}^{n-1}p_{n-1}(t)(\bar{t}-\bar{\alpha}_n)\frac{d\mu(t)}{|\pi_n(t)|^2} \\
&= \int_{\mathbb{T}} \pi_n(t)\phi_n^*(t)\overline{\left(t^{n-1}\overline{p_{n-1}\left(\frac{1}{\bar{t}}\right)}(t-\alpha_n)\right)}\frac{d\mu(t)}{|\pi_n(t)|^2}.
\end{aligned}$$

Therefore, since $t^{n-1}\overline{p_{n-1}(1/\bar{t})}$ ranges over $\mathcal{P}_{n-1}(z)$ as p_{n-1} ranges over the same set, $\pi_n\phi_n^*$ is μ_n-orthogonal to every polynomial of degree at most n which vanishes at α_n. This is also true for $R_n(.,\alpha_n)$. Thus, $\pi_n\phi_n^*$ and $R_n(.,\alpha_n)$ are proportional. We conclude using the following equality

$$\|\pi_n\phi_n^*\|^2_{L^2(d\mu_n)} = \int_{\mathbb{T}} |\pi_n\phi_n^*|^2\frac{d\mu}{|\pi_n|^2} = 1 = \left\|\frac{R_n(.,\alpha_n)}{\|R_n(.,\alpha_n)\|_{L^2(d\mu_n)}}\right\|^2_{L^2(d\mu_n)}.$$

■

We now derive an expression of $|\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2)$ in terms of the reproducing kernels R_n and E_n.

Corollary 6.4 *For every $n \geq 1$,*

$$|\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2) = \frac{R_n(\alpha_n,\alpha_n)}{E_n(\alpha_n,\alpha_n)} \leq 1. \tag{46}$$

Proof By definition of the reproducing kernel, we have

$$\|R_n(.,\alpha_n)\|^2_{L^2(d\mu_n)} = \int_{\mathbb{T}} R_n(t,\alpha_n)\overline{R_n(t,\alpha_n)}\frac{d\mu(t)}{|\pi_n(t)|^2} = R_n(\alpha_n,\alpha_n).$$

Therefore, by proposition 6.3,

$$|\pi_n(\alpha_n)\phi_n^*(\alpha_n)|^2 = \frac{|R_n(\alpha_n,\alpha_n)|^2}{\|R_n(.,\alpha_n)\|^2_{L^2(d\mu_n)}} = R_n(\alpha_n,\alpha_n)$$

and we get the first equality in (46) using the fact that, from proposition 6.2

$$E_n(\alpha_n,\alpha_n) = \frac{1}{1-|\alpha_n|^2}\frac{|\pi_n(\alpha_n)|^2}{|S(\alpha_n)|^2}. \tag{47}$$

Furthermore, as $R_n(.,\omega)$ is the orthogonal projection of $E_n(.,\omega)$ on $\mathcal{P}_n\left(\frac{d\mu}{|\pi_n|^2}\right)$ since $\mathcal{P}_n(d\mu_n) \subset H^2(d\mu_n)$, we have

$$\|R_n(.,\omega)\|_{L^2(d\mu_n)} \leq \|E_n(.,\omega)\|_{L^2(d\mu_n)} \text{ for all } \omega \in \mathbb{D}.$$

Therefore,

$$\|R_n(.,\alpha_n)\|^2_{L^2(d\mu_n)} \leq \|E_n(.,\alpha_n)\|^2_{L^2(d\mu_n)}.$$

As $R_n(\alpha_n,\alpha_n) = \|R_n(.,\alpha_n)\|^2_{L^2(d\mu_n)}$ and $E_n(\alpha_n,\alpha_n) = \|E_n(.,\alpha_n)\|^2_{L^2(d\mu_n)}$, we get

$$\frac{R_n(\alpha_n,\alpha_n)}{E_n(\alpha_n,\alpha_n)} \leq 1.$$

■

We now state our problem in an approximation-theoretic manner.
Because $R_n(.,\alpha_n)$ is the orthogonal projection of $E_n(.,\alpha_n)$ on $\mathcal{P}_n(d\mu_n)$, $R_n(.,\alpha_n)$ is the polynomial of degree at most n which minimizes

$$\min_{r_n\in\mathcal{P}_n} \|E_n(.,\alpha_n) - r_n\|_{L^2(d\mu_n)}.$$

But

$$\begin{aligned}\|E_n(.,\alpha_n)-r_n\|^2_{L^2(d\mu_n)} &= \int_{\mathbb{T}}\left|\frac{1}{1-\overline{\alpha_n}t}\frac{\pi_n(t)\overline{\pi_n(\alpha_n)}}{S(t)\overline{S(\alpha_n)}}-r_n(t)\right|^2\frac{|S(t)|^2}{|\pi_n(t)|^2}dm(t)\\ &= \int_{\mathbb{T}}\left|\frac{1}{1-\overline{\alpha_n}t}\frac{\overline{\pi_n(\alpha_n)}}{\overline{S(\alpha_n)}}-\frac{r_n(t)S(t)}{\pi_n(t)}\right|^2 dm(t).\end{aligned}$$

Thus, finding the polynomial P_n which minimizes

$$\min_{p_n\in\mathcal{P}_n}\left\|\frac{1}{1-\overline{\alpha_n}t}-\frac{p_n(t)S(t)}{\pi_n(t)}\right\|_{L^2(dm)} \tag{48}$$

gives us $R_n(.,\alpha_n)$ by the relation

$$R_n(.,\alpha_n)=\frac{\overline{\pi_n(\alpha_n)}}{\overline{S(\alpha_n)}}P_n.$$

Then, in view of (46) and (47), the quantity $|\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2)$ in which we are interested can be expressed as

$$|\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2)=\left|\frac{P_n(\alpha_n)S(\alpha_n)}{\pi_{n-1}(\alpha_n)}\right|. \tag{49}$$

Now, for every polynomial p_n

$$\begin{aligned}\left\|\frac{1}{1-\bar{\alpha}_nt}-\frac{p_n(t)S(t)}{\pi_n(t)}\right\|^2_{L^2(dm)} &= \left\|\left(1-\frac{p_n(t)S(t)}{\pi_{n-1}(t)}\right)\frac{1}{1-\bar{\alpha}_nt}\right\|^2_{L^2(dm)}\\ &= \left\|\left(1-\frac{p_n(t)S(t)}{\pi_{n-1}(t)}\right)\frac{1}{t-\alpha_n}\right\|^2_{L^2(dm)}\\ &= \left\|\left(1-\frac{p_n(\alpha_n)S(\alpha_n)}{\pi_{n-1}(\alpha_n)}\right)\frac{1}{t-\alpha_n}+\left(\frac{p_n(\alpha_n)S(\alpha_n)}{\pi_{n-1}(\alpha_n)}-\frac{p_n(t)S(t)}{\pi_{n-1}(t)}\right)\frac{1}{t-\alpha_n}\right\|^2_{L^2(dm)}.\end{aligned}$$

Using the orthogonality between analytic and antianalytic functions and the Cauchy theorem, we get

$$\begin{aligned}&\left\|\frac{1}{1-\bar{\alpha}_nt}-\frac{p_n(t)S(t)}{\pi_n(t)}\right\|^2_{L^2(dm)}\\ &=\left|1-\frac{p_n(\alpha_n)S(\alpha_n)}{\pi_{n-1}(\alpha_n)}\right|^2\frac{1}{1-|\alpha_n|^2}+\left\|\left(\frac{p_n(\alpha_n)S(\alpha_n)}{\pi_{n-1}(\alpha_n)}-\frac{p_n(t)S(t)}{\pi_{n-1}(t)}\right)\frac{1}{t-\alpha_n}\right\|^2_{L^2(dm)}.\end{aligned} \tag{50}$$

Therefore, if a sequence of polynomials (p_n) exists such that

$$\left\|\frac{1}{1-\bar{\alpha}_nt}-\frac{p_n(t)S(t)}{\pi_n(t)}\right\|^2_{L^2(dm)}=o\left(\frac{1}{1-|\alpha_n|^2}\right), \tag{51}$$

then by the definition of P_n (see (48)) we also have

$$\left\| \frac{1}{1-\bar{\alpha}_n t} - \frac{P_n(t)S(t)}{\pi_n(t)} \right\|^2_{L^2(dm)} = o\left(\frac{1}{1-|\alpha_n|^2} \right),$$

and using (50), we obtain

$$\lim_n \frac{P_n(\alpha_n)S(\alpha_n)}{\pi_{n-1}(\alpha_n)} = 1.$$

Then, (49) gives

$$\lim_n |\phi_n^*(\alpha_n)|^2 |S(\alpha_n)|^2 (1-|\alpha_n|^2) = 1.$$

Now, suppose that μ' is strictly positive and Dini continuous on $\mathbb{T}$. Then, $\frac{1}{S}$ is an analytic function, continuous on $\mathbb{T}$. If $\sum_{k=0}^{n}(1-|\alpha_k|) = \infty$, then $\cup_{k=0}^{\infty}\mathcal{L}_k$ is dense in the disk algebra $A(\mathbb{D})$ ([1]). Therefore, a sequence of polynomials p_n of degree n exists such that

$$\lim_n \left\| \frac{1}{S} - \frac{p_n}{\pi_n} \right\|_\infty = 0.$$

Thus,

$$\lim_n \left\| 1 - \frac{p_n S}{\pi_n} \right\|_\infty \leq \|S\|_\infty \lim_n \left\| \frac{1}{S} - \frac{p_n}{\pi_n} \right\|_\infty = 0.$$

Since by the Cauchy theorem

$$\left\| \frac{1}{1-\bar{\alpha}_n t} - \frac{p_{n-1}(t)S(t)}{\pi_n(t)} \right\|^2_{L^2(dm)} \leq \left\| 1 - \frac{p_{n-1}S}{\pi_{n-1}} \right\|_\infty \frac{1}{1-|\alpha_n|^2},$$

the sequence (p_{n-1}) satisfies (51). We therefore obtained the following theorem :

Theorem 6.5 *If μ is an absolutely continuous measure such that μ' is strictly positive and Dini continuous on $\mathbb{T}$, and if $\sum_{k=0}^{n}(1-|\alpha_k|) = \infty$, then*

$$\lim_n |\phi_n^*(\alpha_n)|^2 |S(\alpha_n)|^2 (1-|\alpha_n|^2) = 1.$$

Note that in our argument, we uniformly approximate the inverse of S. This leads to quite strong hypotheses. In fact, we only need to find a sequence of polynomials which satisfies the problem defined by (51). This problem is stated in term of L^2 norm, and without inverse of S. Therefore, the hypotheses could be probably weakened using another argument.

6.2 Convergence of the Schur functions f_n

We first give a L^2-convergence property with respect to a varying weight which is the Poisson kernel taken at the points α_j. This leads to the construction of a sequence of interpolation points for which we obtain an asymptotic-BMO-type convergence.

6.2.1 L^2 convergence with respect to a varying weight

We first show a weak-(*) convergence of the measures $\frac{P(.,\alpha_n)}{|\phi_n|^2}dm$:

Lemma 6.6 *If* $\sum_{k=1}^{k=\infty}(1-|\alpha_k|)=\infty$ *then*

$$(*)-\lim_n \frac{P(.,\alpha_n)}{|\phi_n|^2}dm = d\mu.$$

Proof Corollary 4.11 states that $\phi_0,\dots,\phi_n$ are orthonormal in $L^2\left(\frac{P(.,\alpha_n)}{|\phi_n|^2}dm\right)$. Therefore, $\phi_0,\dots,\phi_n$ are orthonormal in $L^2(d\mu)$ and in $L^2\left(\frac{P(.,\alpha_n)}{|\phi_n|^2}dm\right)$. Thus,

$$\int_{\mathbb{T}} \phi_i\overline{\phi_j}\frac{P(.,\alpha_n)}{|\phi_n|^2}dm = \int_{\mathbb{T}} \phi_i\overline{\phi_j}d\mu$$

for all $0\le i,j\le n$. In particular, for all $0\le i\le n$, we have

$$\int_{\mathbb{T}} \phi_i\frac{P(.,\alpha_n)}{|\phi_n|^2}dm = \int_{\mathbb{T}} \phi_i d\mu.$$

As $(\phi_k)_{0\le k\le n}$ is a basis of $\mathcal{L}_n$, for all $g\in\mathcal{L}_n$, we get

$$\int_{\mathbb{T}} g\frac{P(.,\alpha_n)}{|\phi_n|^2}dm = \int_{\mathbb{T}} g d\mu \tag{52}$$

and upon conjugating,

$$\int_{\mathbb{T}} \bar{g}\frac{P(.,\alpha_n)}{|\phi_n|^2}dm = \int_{\mathbb{T}} \bar{g} d\mu. \tag{53}$$

But, as $\sum_{k=1}^{k=\infty}(1-|\alpha_k|)=\infty$, $\cup_{k=0}^{k=\infty}\mathcal{L}_k\bigcup\cup_{k=0}^{k=\infty}\overline{\mathcal{L}_k}$ is dense in $C(\mathbb{T})$, the space of continuous functions in $\mathbb{T}$ ([1]). Therefore,

$$(*)-\lim_n \frac{P(.,\alpha_n)}{|\phi_n|^2}dm = d\mu.$$

■

Note that if the points are compactly included in $\mathbb{D}$ and if I is an open arc on $\mathbb{T}$ such that μ has no mass at the end-points, then we have

$$\lim_n \int_I g\frac{P(.,\alpha_n)}{|\phi_n|^2}dm \le \int_I g d\mu \text{ for every } g\in C(\mathbb{T}). \tag{54}$$

Indeed, let $\epsilon>0$ and let h_I be a continuous positive function such that $h_I(t)=1$ for every t in I and $\int_{\mathbb{T}} h_I d\mu \le \mu(I)+\epsilon$. Then, since all the functions are positive, we have

$$\int_I \frac{P(.,\alpha_n)}{|\phi_n|^2}dm \le \int_{\mathbb{T}} h_I\frac{P(.,\alpha_n)}{|\phi_n|^2}dm.$$

We conclude using the previous lemma since

$$\lim_n \int_{\mathbb{T}} h_I \frac{P(.,\alpha_n)}{|\phi_n|^2} dm = \int_{\mathbb{T}} h_I d\mu \le \mu(I) + \epsilon.$$

Note also that if $\sum_{k=1}^{k=\infty}(1-|\alpha_k|) = \infty$, since $P(z,\alpha_n) = z/(z-\alpha_n) + \bar{\alpha_n}z/(1-\bar{\alpha_n}z)$, $P(.,\alpha_n)$ is in $\mathcal{L}_n + \bar{\mathcal{L}}_n$, then we get using (52) and (53)

$$\int_{\mathbb{T}} P(.,\alpha_n) \frac{P(.,\alpha_n)}{|\phi_n|^2} dm = \int_{\mathbb{T}} P(.,\alpha_n) d\mu. \tag{55}$$

If the interpolation points do not tend "too quickly" toward the circle, we have the following L^2-convergence :

Theorem 6.7 *Let μ be an absolutely continuous measure. If $\sum_{k=1}^{k=\infty}(1-|\alpha_k|) = \infty$ and $\lim_k |\alpha_k| = 1$, and if at every point of accumulation of the (α_k) f is continuous and $|f| < 1$, then*

$$\lim_k \int |f_k|^2 P(.,\alpha_k) dm = 0.$$

Proof Suppose that the limit does not converge to 0. Then, there is $\epsilon > 0$, an infinite set $K \subset \mathbb{N}$ and a sub-sequence of (α_k) which converges to $\alpha \in \mathbb{T}$ such that

$$\forall n \in K, \qquad \int |f_n|^2 P(.,\alpha_n) dm \ge \epsilon.$$

By theorem 5.7, using the elementary equality

$$|1 - \overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n|^2 = 1 + |f_n|^2 - 2Re(\overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n)$$

we get

$$|\phi_n|^2 \mu'(1 + |f_n|^2 - 2Re(\overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n)) = (1 - |f_n|^2) P(.,\alpha_n)$$

and therefore

$$|f_n|^2 = \frac{P(.,\alpha_n) - |\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'} + \frac{2|\phi_n|^2\mu' Re(\overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n)}{P(.,\alpha_n) + |\phi_n|^2\mu'}.$$

Thus, we obtain

$$|f_n|^2 = \frac{P(.,\alpha_n) - |\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'} - \frac{P(.,\alpha_n) - |\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'} Re\left(\overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right) + Re\left(\overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right).$$

Since $\zeta_n(\alpha_n) = 0$, we get by harmonicity

$$\int Re\left(\overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right) P(.,\alpha_n) dm = 0.$$

Consequently,

$$\int |f_n|^2 P(.,\alpha_n)dm = \int \frac{P(.,\alpha_n) - |\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'} \left(1 - Re\left(\overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right)\right) P(.,\alpha_n)dm.$$

But since ζ_n, f_n and $\frac{\phi_n}{\phi_n^*}$ are Schur functions (see proposition 4.6),

$$\left|1 - Re\left(\overline{c_n}\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right)\right| \leq 2$$

and we get

$$\int |f_n|^2 P(.,\alpha_n)dm \leq 2\int \left|1 - \frac{2|\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'}\right| P(.,\alpha_n)dm. \tag{56}$$

Let

$$g_n = \frac{2|\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'}.$$

Using the inequality

$$\frac{4x^2}{(1+x)^2} \leq x \text{ for all } x \geq 0$$

we deduce

$$\begin{aligned}
\int_{\mathbb{T}} g_n^2 P(.,\alpha_n)dm &= \int_{\mathbb{T}} \frac{4(|\phi_n|^2\mu' P(.,\alpha_n)^{-1})^2}{(1 + |\phi_n|^2\mu' P(.,\alpha_n)^{-1})^2} P(.,\alpha_n)dm \\
&\leq \int_{\mathbb{T}} |\phi_n|^2\mu' P(.,\alpha_n)^{-1} P(.,\alpha_n)dm \\
&= \int_{\mathbb{T}} |\phi_n|^2\mu' dm \leq 1
\end{aligned}$$

because of the orthonormality of ϕ_n. By the Schwarz inequality, it follows that

$$\int_{\mathbb{T}} g_n P(.,\alpha_n)dm \leq \left(\int_{\mathbb{T}} g_n^2 P(.,\alpha_n)dm\right)^{1/2} \leq 1. \tag{57}$$

Furthermore, we get again by the Schwarz inequality:

$$\begin{aligned}
\int_{\mathbb{T}} \sqrt{\mu'} P(.,\alpha_n)dm &= \int_{\mathbb{T}} \frac{\sqrt{2}|\phi_n|\sqrt{\mu'}\sqrt{P(.,\alpha_n)}}{\sqrt{P(.,\alpha_n) + |\phi_n|^2\mu'}} \frac{\sqrt{P(.,\alpha_n) + |\phi_n|^2\mu'}\sqrt{P(.,\alpha_n)}}{\sqrt{2}|\phi_n|} dm \\
&\leq \left(\int_{\mathbb{T}} g_n P(.,\alpha_n)dm\right)^{1/2} \left(\frac{1}{2}\int_{\mathbb{T}} \left(\frac{P(.,\alpha_n)}{|\phi_n|^2} + \mu'\right) P(.,\alpha_n)dm\right)^{1/2}.
\end{aligned}$$

Using (55) and the absolutely continuity of the measure, we get

$$\int_{\mathbb{T}} \sqrt{\mu'} P(.,\alpha_n)dm \leq \left(\int_{\mathbb{T}} g_n P(.,\alpha_n)dm\right)^{1/2} \left(\int_{\mathbb{T}} \mu' P(.,\alpha_n)dm\right)^{1/2}. \tag{58}$$

Since by hypothesis, (α_n) converges to $\alpha \in \mathbb{T}$ and μ' is continuous at α, passing to the inferior limit in (58), we get

$$\sqrt{\mu'(\alpha)} \leq \sqrt{\mu'(\alpha)} \liminf_n \left(\int_{\mathbb{T}} g_n P(., \alpha_n) dm \right)^{1/2}.$$

Therefore, we obtain

$$\liminf_n \int_{\mathbb{T}} g_n P(., \alpha_n) dm \geq 1.$$

Combining this last inequality with (57), we obtain

$$\lim_n \int_{\mathbb{T}} g_n P(., \alpha_n) dm = \lim_n \int_{\mathbb{T}} g_n^2 P(., \alpha_n) dm = 1.$$

It follows that

$$\lim_n \int_{\mathbb{T}} (1-g_n)^2 P(., \alpha_n) dm = \int_{\mathbb{T}} P(., \alpha_n) dm - 2 \lim_n \int_{\mathbb{T}} g_n P(., \alpha_n) dm + \lim_n \int_{\mathbb{T}} g_n^2 P(., \alpha_n) dm = 0.$$

Thus, using the Schwarz inequality and (56), we conclude that

$$\lim_n \int_{\mathbb{T}} |f_n|^2 P(., \alpha_n) dm = 0.$$

■

A similar type of convergence is obtained when the (α_n) are compactly in included in $\mathbb{D}$.

Theorem 6.8 *If the (α_k) are compactly included in $\mathbb{D}$ and if $|f| < 1$ a.e. on $\mathbb{T}$, then*

$$\lim_n \int |f_n|^2 P(., \alpha_n) dm = 0.$$

Proof We denote by $\alpha \in \mathbb{D}$ an accumulation point of (α_k). Using the same argument as above, equation (57) still holds. Now, for any open arc I on $\mathbb{T}$ with no mass at the end-points, we get by the Schwarz inequality:

$$\begin{aligned} \frac{1}{m(I)} \int_I \sqrt{\mu'} dm &= \frac{1}{m(I)} \int_I \frac{\sqrt{2}|\phi_n|\sqrt{\mu'}}{\sqrt{P(., \alpha_n) + |\phi_n|^2 \mu'}} \frac{\sqrt{P(., \alpha_n) + |\phi_n|^2 \mu'}}{\sqrt{2}|\phi_n|} dm \quad (59) \\ &\leq \left(\frac{1}{m(I)} \int_I g_n dm \right)^{1/2} \left(\frac{1}{2|I|} \int_I \left(\frac{P(., \alpha_n)}{|\phi_n|^2} + \mu' \right) dm \right)^{1/2} (60) \end{aligned}$$

As $g_n = \frac{2|\phi_n|^2\mu' P(e^{i\theta},\alpha_n)^{-1}}{1+|\phi_n|^2\mu' P(e^{i\theta},\alpha_n)^{-1}}$, we have $0 \leq g_n \leq 2$ a.e on $\mathbb{T}$. Let g be a weak-$(*)$ limit of the bounded sequence $(g_n)_n$ in $L^\infty(\mathbb{T})$. Passing to the limit in (60), and using (54), we obtain

$$\frac{1}{|I|}\int_I \sqrt{\mu'}dm \leq \left(\frac{1}{|I|}\int_I gdm\right)^{1/2}\left(\frac{1}{2}\frac{\mu(I)}{|I|} + \frac{1}{2|I|}\int_I \mu' dm\right)^{1/2}.$$

Thus, by Lebesgue's theorem on differentiation and by Helly's theorem ([10]),

$$\sqrt{\mu'} \leq \sqrt{g}\left(\frac{1}{2}\mu' + \frac{1}{2}\mu'\right)^{1/2} \leq \sqrt{g}\sqrt{\mu'} \text{ a.e. on } \mathbb{T}.$$

Since $\mu' > 0$ a.e. on $\mathbb{T}$, $g \geq 1$ a.e. on $\mathbb{T}$. Combining this last inequality with (57), and using the fact that $\lim_n P(.,\alpha_n) = P(.,\alpha)$ uniformly on $\mathbb{T}$, we obtain

$$\lim_n \int_{\mathbb{T}} g_n P(.,\alpha_n)dm = \lim_n \int_{\mathbb{T}} g_n^2 P(.,\alpha_n)dm = 1.$$

It follows that

$$\lim_n \int_{\mathbb{T}} (1-g_n)^2 P(.,\alpha_n)dm = \int_{\mathbb{T}} P(.,\alpha_n)dm - 2\lim_n \int_{\mathbb{T}} g_n P(.,\alpha_n)dm + \lim_n \int_{\mathbb{T}} g_n^2 P(.,\alpha_n)dm = 0.$$

Thus, using the Schwarz inequality and (56), we conclude that

$$\lim_n \int_{\mathbb{T}} |f_n|^2 P(.,\alpha_n)dm = 0.$$

■

Combining the proofs of the two previous theorems, we obtain:

Corollary 6.9 *Let μ be an absolutely continuous measure. If $\sum_{k=1}^{k=\infty}(1-|\alpha_k|) = \infty$, if $|f| < 1$ a.e. on $\mathbb{T}$ and if at every point of accumulation of the (α_k) in $\mathbb{T}$, f is continuous and $|f| < 1$, then*

$$\lim_k \int |f_k|^2 P(.,\alpha_k)dm = 0.$$

In particular, we obtain a result stated in ([20]) for the classical Schur algorithm:

Corollary 6.10 *If $1 \leq p < \infty$, $|f| < 1$ a.e. on $\mathbb{T}$ and $\alpha_k = 0$ for every $k \geq 1$ then*

$$\lim_n \int_{\mathbb{T}} |f_n|^p \, dm = 0.$$

Proof As $\|f_n\|_\infty \leq 1$ for all n, the sequence f_n is in L^p for all $1 \leq p \leq \infty$. But $\|f_n\|_2$ converges to 0, so for every sequence, we can extract a subsequence such that $\lim_k f_k(t) = 0$ a.e. on $\mathbb{T}$. We conclude using Lebesgue's dominated convergence.

■

6.2.2 An asymptotic-BMO-type convergence

In the following, we will construct a sequence of interpolation points for which the sequence f_n tends in L^1 mean to its average on smaller and smaller intervals.

Theorem 6.11 *Let $(\epsilon_k)_{k\in\mathbb{N}}$ be a sequence of real numbers such that*

$$\begin{cases} 0 < \epsilon_k \le \frac{1}{\pi}, \\ \sum_{k=0}^{k=\infty} \epsilon_k = \infty, \\ \lim_{k\to\infty} \epsilon_k = 0, \end{cases}$$

and f be a continuous Schur function such that $|f| < 1$ on $\mathbb{T}$.
Then the points $(\alpha_k)_k$ can be chosen such that

$$\lim_n \sup_{\alpha\in D_n} \int_{\mathbb{T}} |f_n(t) - f_n(\alpha)|\, P(t,\alpha) dm(t) = 0.$$

where D_n denotes the closed disk of radius $1-\epsilon_n\pi$:

$$D_n = \{z \in \mathbb{C}, |z| \le 1 - \epsilon_n \pi\}.$$

Proof Recall that

$$|f_{n+1}(e^{i\theta})| = \left| \frac{f_n(e^{i\theta}) - f_n(\alpha_{n+1})}{1 - \overline{f_n(\alpha_{n+1})} f_n(e^{i\theta})} \right|.$$

We denote by $\mathcal{I}_n$ the application from $\mathbb{D}$ to $[0,1]$ such that

$$\mathcal{I}_n(\alpha) = \int_{\mathbb{T}} \left| \frac{f_n(t) - f_n(\alpha)}{1 - \overline{f_n(\alpha)} f_n(t)} \right|^2 P(t,\alpha) dm(t).$$

At each step of the Schur algorithm, we may choose $\alpha_{n+1} \in D_n$ which maximizes $\mathcal{I}_n$. Then we have :

$$\begin{aligned} \int_{\mathbb{T}} |f_{n+1}(t)|^2 P(t,\alpha_{n+1}) dm(t) &= \int_{\mathbb{T}} \left| \frac{f_n(t) - f_n(\alpha_{n+1})}{1 - \overline{f_n(\alpha_{n+1})} f_n(t)} \right|^2 P(t,\alpha_{n+1}) dm(t) \\ &= \sup_{\alpha\in D_n} \int_T \left| \frac{f_n(t) - f_n(\alpha)}{1 - \overline{f_n(\alpha)} f_n(t)} \right|^2 P(t,\alpha) dm(t). \end{aligned}$$

As f_n is Schur, $|1 - \overline{f_n(\alpha)} f_n(t)| \le 2$. Therefore,

$$2 \int_{\mathbb{T}} |f_{n+1}(t)|^2 P(t,\alpha_{n+1}) dm(t) \ge \sup_{\alpha\in D_n} \int_{\mathbb{T}} |f_n(t) - f_n(\alpha)|^2 P(t,\alpha) dm(t).$$

Using the Schwarz inequality, we get

$$2 \int_{\mathbb{T}} |f_{n+1}(t)|^2 P(t,\alpha_{n+1}) dm(t) \ge \left(\sup_{\alpha\in D_n} \int_{\mathbb{T}} |f_n(t) - f_n(\alpha)|\, P(t,\alpha) dm(t) \right)^2.$$

Thus, corollary 6.9 gives

$$\lim_n \sup_{\alpha \in D_n} \int_{\mathbb{T}} |f_n(t) - f_n(\alpha)|\, P(t,\alpha) dm(t) = 0.$$

■

Corollary 6.12 *Under the same hypothesis as the previous theorem, the points* $(\alpha_k)_k$ *can be chosen such that*

$$\lim_{n\to\infty} \sup_{m(I)\geq \epsilon_n} \frac{1}{m(I)} \int_I |f_n - (f_n)_I|\, dm = 0$$

where $(f_n)_I$ *is defined by*

$$(f_n)_I = \frac{1}{m(I)} \int_I f_n dm.$$

Proof Let I be an arc of $\mathbb{T}$ such that $m(I) \geq \epsilon_n$.
Suppose first that $m(I) \leq \frac{1}{\pi}$ and define by α_I the point of D_n such that $\alpha_I = (1 - m(I)\pi)e^{i\theta_I}$ where $e^{i\theta_I}$ is the center of I. We have

$$\begin{aligned} P(e^{i\theta}, \alpha_I) &= \frac{1 - |\alpha_I|^2}{1 - 2|\alpha_I|\cos(\theta - \theta_I) + |\alpha_I|^2} \\ &= \frac{1 + |\alpha_I|}{1 - |\alpha_I| + 2|\alpha_I| \frac{1-\cos(\theta-\theta_I)}{1-|\alpha_I|}}. \end{aligned}$$

Suppose that $e^{i\theta} \in I$, that is $|\theta - \theta_I| \leq m(I)\pi$. Using the inequality $1 - \cos(x) \leq \frac{x^2}{2}$, we get

$$\begin{aligned} P(e^{i\theta}, \alpha_I) &\geq \frac{1 + |\alpha_I|}{1 - |\alpha_I| + |\alpha_I| \frac{(\theta-\theta_I)^2}{1-|\alpha_I|}} \\ &\geq \frac{1 + |\alpha_I|}{1 - |\alpha_I| + \frac{|\alpha_I|\pi^2 m(I)^2}{1-|\alpha_I|}} \\ &\geq \frac{2 - \pi m(I)}{\pi m(I) + \frac{(1-\pi m(I))\pi^2 m(I)^2}{\pi m(I)}} \\ &\geq \frac{1}{m(I)\pi}. \end{aligned}$$

Therefore, if χ stands for the characteristic function of I and if $\epsilon_n \leq m(I) \leq \frac{1}{\pi}$, then $\frac{\chi(t)}{m(I)} \leq \pi P(t, \alpha_I)$.

Furthermore, if $m(I) > \frac{1}{\pi}$, we have $\pi P(t,0) = \pi \geq \frac{1}{m(I)}$. Thus, for all arc I of $\mathbb{T}$ such that $m(I) \geq \epsilon_n$, a point α_I in D_n exists such that

$$\frac{\chi(t)}{m(I)} \leq \pi P(t, \alpha_I).$$

Now, remark that $|(f_n)_I - f_n(\alpha_I)| \leq 1/m(I) \int_I |f_n - f_n(\alpha_I)| dm$. Indeed,

$$\begin{aligned} |(f_n)_I - f_n(\alpha_I)| &= \left| \frac{1}{m(I)} \int_I f_n dm - f_n(\alpha_I) \right| = \left| \frac{1}{m(I)} \int_I (f_n - f_n(\alpha_I)) dm \right| \\ &\leq \frac{1}{m(I)} \int_I |f_n - f_n(\alpha_I)| dm. \end{aligned}$$

We conclude using the above theorem and the following inequalities:

$$\begin{aligned} \sup_{m(I)\geq\epsilon_n} \frac{1}{m(I)} \int_I |f_n - (f_n)_I|\, dm &\leq \sup_{m(I)\geq\epsilon_n} \frac{1}{m(I)} \int_I (|f_n - f_n(\alpha_I)| + |f_n(\alpha_I) - (f_n)_I|)\, dm \\ &\leq 2 \sup_{m(I)\geq\epsilon_n} \frac{1}{m(I)} \int_I |f_n - f_n(\alpha_I)|\, dm \\ &= 2 \sup_{m(I)\geq\epsilon_n} \int \frac{\chi}{m(I)} |f_n - f_n(\alpha_I)|\, dm \\ &\leq 2\pi \sup_{m(I)\geq\epsilon_n} \int |f_n - f_n(\alpha_I)|\, P(., \alpha_I) dm \\ &\leq 2\pi \sup_{\alpha\in D_n} \int |f_n - f_n(\alpha)|\, P(., \alpha) dm \end{aligned}$$

■

If no constraint is made on the length of the intervals (i.e. $\epsilon_n = 0$ for each n), then the convergence in the previous corollary is called a BMO convergence. Details about BMO can be found in [12], Chapter 6.
Here, an unsolved question appears: which hypotheses are needed on f in order to obtain a BMO convergence? The difficulty to answer such a question is that the hypotheses made on f have to propagate to every f_n throughout the Schur algorithm.
Note also that we do not obtain a similar result of convergence for the Wall rational functions A_n/B_n. Here, the problem is due to the mean $(f_n)_I$.

6.3 Convergence of the Wall rational functions A_n/B_n

We will now give different kinds of convergence for the Wall rational functions. The first one is convergence on compact subset which is deduced merely from an elementary property satisfied by the zeros of a non-zero function in H^∞. The other three (convergence in the pseudo-hyperbolic distance, the Poincaré metric, and in $L^2(\mathbb{T})$) are implied by the convergence of the Schur functions f_n in $L^2(\mathbb{T})$.

6.3.1 Convergence on compact subsets

Convergence of A_n/B_n on compact subsets of $\mathbb{D}$ is easily obtained, using the fact that the zeros of a non-zero function in H^∞ satisfy the relation $\sum_{k=1}^{\infty}(1-|\alpha_k|) < \infty$ ([25]).

Theorem 6.13 *If $\sum_{k=1}^{k=\infty}(1-|\alpha_k|) = \infty$, $\frac{A_n}{B_n}$ converges to f uniformly on compact subsets of $\mathbb{D}$.*

Proof As $\left\|\frac{A_n}{B_n}\right\|_\infty \leq 1$ for all $n \in \mathbb{N}$, $\left\{\frac{A_n}{B_n}\right\}$ is a normal family. Therefore, a subsequence that converges uniformly on compact subsets can be extracted. We denote by $\check{f}$ the limit of such a subsequence. As $\frac{A_n}{B_n}(\alpha_k) = f(\alpha_k)$ for all $n \geq k-1$, $f(\alpha_k) = \check{f}(\alpha_k)$ for all k. Thus, the function $f - \check{f}$ belongs to H^∞ and the points α_k are its zeros. As $\sum_{k=1}^{k=\infty}(1-|\alpha_k|) = \infty$, we conclude that $f = \check{f}$. Thus, f is the only limit point, and A_n/B_n converges to f, locally uniformly in $\mathbb{D}$

■

6.3.2 Convergence with respect to the pseudohyperbolic distance

The pseudohyperbolic distance ρ on $\mathbb{D}$ is defined by ([12])

$$\rho(z, w) = \left|\frac{z-w}{1-\bar{w}z}\right|.$$

Convergence with respect to the pseudohyperbolic distance is essentially a consequence of the following well-known property.

Property 6.14 *The pseudohyperbolic distance is invariant under Moebius transformations.*

Proof Let $\mathcal{M}$ be the Moebius transform defined by

$$\mathcal{M}(z) = \beta\frac{z-\alpha}{1-\bar{\alpha}z} \text{ with } \alpha \in \mathbb{D} \text{ and } \beta \in \mathbb{T}.$$

We have

$$\begin{aligned}\mathcal{M}(z) - \mathcal{M}(\omega) &= \beta\left(\frac{z-\alpha}{1-\bar{\alpha}z} - \frac{\omega-\alpha}{1-\bar{\alpha}\omega}\right) \\ &= \beta\frac{(1-|\alpha|^2)(z-\omega)}{(1-\bar{\alpha}z)(1-\bar{\alpha}\omega)}\end{aligned}$$

and

$$\begin{aligned}1 - \overline{\mathcal{M}(z)}\mathcal{M}(\omega) &= 1 - \overline{\left(\beta\frac{z-\alpha}{1-\bar{\alpha}z}\right)}\beta\frac{\omega-\alpha}{1-\bar{\alpha}\omega} \\ &= \frac{(1-|\alpha|^2)(1-\bar{z}\omega)}{(1-\alpha\bar{z})(1-\bar{\alpha}\omega)}.\end{aligned}$$

Therefore,

$$\left|\frac{\mathcal{M}(z)-\mathcal{M}(\omega)}{1-\overline{\mathcal{M}(z)}\mathcal{M}(\omega)}\right| = \left|\frac{z-\omega}{1-\bar{z}\omega}\right|.$$

■

The proof of convergence is now immediate ([20], Corollary 2.4 for $\alpha_k = 0$):

Theorem 6.15 *If $|f| < 1$ on $\mathbb{T}$, f continuous, and $\sum_{k=1}^{k=\infty}(1-|\alpha_k|) = \infty$ then*

$$\lim_n \int_{\mathbb{T}} \rho\left(f, \frac{A_n}{B_n}\right)^2 P(., \alpha_{n+1}) dm = 0$$

Proof As the pseudohyperbolic distance is invariant under Moebius transformations, we have in view of (7) and (8),

$$\rho\left(f, \frac{A_n}{B_n}\right) = \rho\left(\tau_0 \circ \cdots \circ \tau_n(f_{n+1}), \tau_0 \circ \cdots \circ \tau_n(0)\right) = \rho(f_{n+1}, 0) = |f_{n+1}|.$$

We conclude using Corollary 6.9.

■

6.3.3 Convergence with respect to the Poincaré metric

In the disk, the Poincaré metric is defined by

$$\mathfrak{P}(z,\omega) = \log\left(\frac{1+\rho(z,\omega)}{1-\rho(z,\omega)}\right) \text{ for } z,\omega \in \mathbb{D}.$$

The following theorem is given in the classical case (i.e. $\alpha_k = 0$) in [20], Theorem 2.6.

Theorem 6.16 *If μ is an absolutely continuous measure such that μ' is positive and Dini continuous on $\mathbb{T}$ and if $\sum_{k=0}^{n}(1-|\alpha_k|) = \infty$, then*

$$\lim_n \int_{\mathbb{T}} \mathfrak{P}\left(f, \frac{A_n}{B_n}\right) P(., \alpha_{n+1}) dm = 0.$$

In particular, this holds if $|f| < 1$ and f is Dini-continuous on $\mathbb{T}$.

Proof Using again the invariance of the pseudohyperbolic distance under Moebius transformations, we get $\rho\left(f, \frac{A_n}{B_n}\right) = |f_{n+1}|$. This gives

$$\mathfrak{P}\left(f, \frac{A_n}{B_n}\right) = \log\left(\frac{1+|f_{n+1}|}{1-|f_{n+1}|}\right). \tag{61}$$

Using Theorem 5.7 and the definition of the Szegő function S, since $|\phi_n| = |\phi_n^*|$ on $\mathbb{T}$, we get

$$|\phi_n^*|^2|S|^2\frac{|1-\bar{\alpha}_n\xi|^2}{1-|\alpha_n|^2} = \frac{1-|f_n|^2}{|1-\overline{c_n}\zeta_n\frac{\phi_n}{\phi_n^*}f_n|^2} \text{ a.e. on } \mathbb{T}. \tag{62}$$

Furthermore, if g is a Schur function, $1-g$ is a function in H^∞ such that $Re(1-g) \geq 0$, and therefore $1-g$ is an outer function (see [12], Corollary 4.8). Thus,

$$\int_{\mathbb{T}} \log|1-g|^2 P(.,\alpha_n)dm = \log(|1-g(\alpha_n)|^2).$$

Consequently, since $\zeta_n(\alpha_n) = 0$,we obtain on putting $g = \overline{c_n}\zeta_n\frac{\phi_n}{\phi_n^*}f_n$ that

$$\int_{\mathbb{T}} \log|1-\overline{c_n}\zeta_n\frac{\phi_n}{\phi_n^*}f_n|^2 P(.,\alpha_n)dm = \log(|1-\overline{c_n}\zeta_n(\alpha_n)\frac{\phi_n(\alpha_n)}{\phi_n^*(\alpha_n)}f_n(\alpha_n)|^2) = \log(1) = 0.$$

Using the previous equation and (62), we get

$$\int_{\mathbb{T}} \log\left(|\phi_n^*|^2|S|^2\frac{|1-\bar{\alpha}_n\xi|^2}{1-|\alpha_n|^2}\right) P(\xi,\alpha_n)dm(\xi) = \int_{\mathbb{T}} \log(1-|f_n|^2)P(\xi,\alpha_n)dm(\xi).$$

As ϕ_n^*, S^2 and $1-\bar{\alpha}_n\xi$ are outer functions, we obtain

$$\log(|\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2)) = \int_{\mathbb{T}} \log(1-|f_n|^2)P(.,\alpha_n)dm,$$

and Theorem 6.5 gives us

$$\lim_n \int_{\mathbb{T}} \log(1-|f_n|^2)P(.,\alpha_n)dm = 0. \tag{63}$$

Using the inequality $\log(1+x) \leq x$ for $x > -1$, we get

$$0 \leq |f_n|^2 \leq -\log(1-|f_n|^2) \tag{64}$$

and

$$0 \leq \log(1+|f_n|) \leq |f_n|. \tag{65}$$

Therefore, by (64) and (63),

$$\lim_n \int_{\mathbb{T}} |f_n|^2 P(.,\alpha_n)dm = 0$$

and, by the previous equation and (65),

$$\lim_n \int_{\mathbb{T}} \log(1+|f_n|)P(.,\alpha_n)dm = 0$$

because, by the Schwarz inequality,

$$0 \leq \int_{\mathbb{T}} \log(1+|f_n|)P(.,\alpha_n)dm \leq \int_{\mathbb{T}} |f_n|P(.,\alpha_n)dm \leq \left(\int_{\mathbb{T}} |f_n|^2 P(.,\alpha_n)dm\right)^{1/2}.$$

Since $\log(1-|f_n|^2) = \log(1-|f_n|) + \log(1+|f_n|)$, we also have

$$\lim_n \int_{\mathbb{T}} \log(1-|f_n|)P(.,\alpha_n)dm = 0.$$

We obtain the expected result by (61).

■

6.3.4 Convergence in $L^2(\mathbb{T})$

Using the relation between f_{n+1} and $\frac{A_n}{B_n}$ and the L^2 convergence of the Schur functions f_n, we shall directly obtain the L^2 convergence of the Wall rational functions $\frac{A_n}{B_n}$ as follows.

Lemma 6.17 *For $t \in \mathbb{T}$, we have*

$$|f_{n+1}(t)| \left|1 - \frac{A_n}{B_n}(t)\overline{f(t)}\right| = \left|f(t) - \frac{A_n}{B_n}(t)\right|.$$

Proof Proposition 3.12 gives

$$f(z) = \frac{A_n(z) + \zeta_{n+1}(z)B_n^*(z)f_{n+1}(z)}{B_n(z) + \zeta_{n+1}(z)A_n^*(z)f_{n+1}(z)}.$$

Therefore,

$$f(z) - \frac{A_n(z)}{B_n(z)} = \zeta_{n+1}(z)f_{n+1}(z)\frac{B_n^*(z) - A_n^*(z)f(z)}{B_n(z)}.$$

Thus, for $t \in \mathbb{T}$,

$$\begin{aligned}
\left|f(t) - \frac{A_n(t)}{B_n(t)}\right| &= |\zeta_{n+1}(t)f_{n+1}(t)| \left|\frac{\mathcal{B}_n(t)(\overline{B_n(t)} - \overline{A_n(t)}f(t))}{B_n(t)}\right| \\
&= |f_{n+1}(t)| \left|\frac{\overline{B_n(t)} - \overline{A_n(t)}f(t)}{B_n(t)}\right| \\
&= |f_{n+1}(t)| \left|1 - \frac{\overline{A_n(t)}}{\overline{B_n(t)}}f(t)\right|.
\end{aligned}$$

■

Proposition 6.18 *The convergence in L^p, $1 \leq p < \infty$, of f_n to zero with respect to the varying weight $P(.,\alpha_n)$ implies the convergence in L^p of $\frac{A_n}{B_n}$ to f with respect to $P(.,\alpha_{n+1})$.*

Proof As f and $\frac{A_n}{B_n}$ are two Schur functions, using the previous lemma, we get

$$\left| f(t) - \frac{A_n}{B_n}(t) \right| \leq 2|f_{n+1}(t)| \text{ for } t \in \mathbb{T}.$$

The conclusion is then immediate by dominated convergence.

■

The two following corollaries are direct applications of the previous results.

Corollary 6.19 *If $\sum_{k=1}^{k=\infty}(1-|\alpha_k|) = \infty$, and if $|f| < 1$ and f is continuous on $\mathbb{T}$, then*

$$\lim_n \int_{\mathbb{T}} \left| f - \frac{A_{n-1}}{B_{n-1}} \right|^2 P(.,\alpha_n) dm = 0.$$

In particular, we obtain a result given in [20] for the classical Schur algorithm:

Corollary 6.20 *If $1 \leq p < +\infty$, $|f| < 1$ a.e. on $\mathbb{T}$, and $\alpha_k = 0$ for every $k \geq 1$, then*

$$\lim_n \int_{\mathbb{T}} \left| f - \frac{A_n}{B_n} \right|^p dm = 0.$$

7 Approximation by a Schur rational function of given degree

The goal of this section is to give practical means of approximating a function by a Schur rational function. We first show that the Schur algorithm leads to a parametrization of all *strictly* Schur rational functions of given degree. We next explain how to compute efficiently the L^2 norm of a rational function analytic in the unit disk. We then have all the necessary information to implement an optimization process. Examples are given, and compared with L^2 unconstrained approximation.

7.1 Parametrization of strictly Schur rational functions

Below, we parametrize the strictly Schur rational functions of order n by their convergents of order n (see section 3.3). Let $(c_k)_{k\geq 0}$ be a sequence on $\mathbb{T}$ with $c_0 = 1$. We denote by $\mathcal{S}_n$ the set of all strictly Schur rational functions of degree at most n and we define the application Γ by

$$\begin{array}{rccc} \Gamma : & \mathbb{D}^{2n+1} & \longrightarrow & \mathcal{S}_n \\ & (\alpha_1, \ldots, \alpha_n, \gamma_0, \ldots, \gamma_n) & \longmapsto & R_n \end{array}$$

where

$$R_n = \tau_0 \circ \tau_1 \circ \cdots \circ \tau_{n-1} \circ \tau_n(0)$$

with

$$\tau_k(\omega) = c_k\gamma_k + \frac{(1-|\gamma_k|^2)c_k\zeta_{k+1}}{\bar{\gamma}_k\zeta_{k+1} + \frac{1}{\omega}}.$$

The next theorem shows that Γ is surjective.
For h a polynomial of degree n, we denote by $\tilde{h}$ the polynomial of degree n defined by $\tilde{h}(z) = z^n\overline{h(\frac{1}{\bar{z}})}$.

Theorem 7.1 *Every* strictly *Schur irreducible rational function* $\frac{p}{q}$ *of degree* n *can be written as a convergent of order* n.
Furthermore, the only possible interpolation points $\alpha_1, \ldots, \alpha_n$ *(counted with multiplicity) are the points in the set*

$$\mathcal{R} = \{z \in \mathbb{D}, (p\tilde{p} - q\tilde{q})(z) = 0\}.$$

Proof We will show that choosing the interpolation points in $\mathcal{R}$ leads to a constant Schur function f_n. We then conclude applying the reverse Schur algorithm.

1. We first prove that $p\tilde{p} - q\tilde{q}$ has n roots in the unit disk $\mathbb{D}$. Suppose that $p\tilde{p} - q\tilde{q}$ is a polynomial of degree $m < 2n$. Then if we put $p = \sum_{k=0}^n a_k z^k$ and $q = \sum_{k=0}^n b_k z^k$, we have

$$a_{n-k}\bar{a}_k - b_{n-k}\bar{b}_k = 0 \text{ for all } 0 \le k < 2n - m$$

   and therefore, 0 is a root of $\tilde{p}p - \tilde{q}q$ with multiplicity $2n-m$. Suppose now that some root ξ is on the unit circle $\mathbb{T}$. As $\frac{p}{q}$ is Schur and irreducible, $q(\xi) \neq 0$. Then $\frac{p\tilde{p}}{q\tilde{q}}(\xi) = \left|\frac{p}{q}(\xi)\right|^2 = 1$, and therefore, $\frac{p}{q}$ is not strictly Schur, a contradiction. Furthermore, if $\xi \neq 0$ is a root of $p\tilde{p} - q\tilde{q}$, $\frac{1}{\bar{\xi}}$ is also a root of $p\tilde{p} - q\tilde{q}$. Therefore, there are exactly n points (counted with multiplicity) in $\mathcal{R}$.

2. We now show that the degree of f_i decreases at each step of the Schur algorithm if and only if the α_j are taken in $\mathcal{R}$.

   Recall that

$$f_1 = \bar{c}_0 \frac{p - c_0\gamma_0 q}{q - \overline{c_0\gamma_0}p}\frac{1 - \overline{\alpha_1}z}{z - \alpha_1}.$$

   First, note that $p - c_0\gamma_0 q$ and $q - \overline{c_0\gamma_0}p$ are relatively prime. Indeed, if α is a common root, we have $p(\alpha) = c_0\gamma_0 q(\alpha)$ and $q(\alpha) - |\gamma_0|^2 q(\alpha) = 0$. Therefore, $q(\alpha) = 0$ and $p(\alpha) = 0$. This contradicts the irreducibility of $\frac{p}{q}$.
   Note also that, if $\deg(p - c_0\gamma_0 q) \le n-1$ and $\deg(q - \overline{c_0\gamma_0}p) \le n-1$, then $\deg p \le n-1$ and $\deg q \le n-1$. Indeed, we get $a_n - c_0\gamma_0 b_n = 0$ and

$b_n - \overline{c_0\gamma_0}a_n = 0$, and therefore $a_n(1-|c_0\gamma_0|^2) = 0$ and $b_n(1-|c_0\gamma_0|^2) = 0$. Since $|c_0\gamma_0| < 1$, we obtain $a_n = b_n = 0$. This contradicts the hypothesis $\deg p/q = n$.
Thus, the degree of f_1 is equal to $n-1$ if and only if

- $z-\alpha_1$ divides $p - c_0\gamma_0 q$, and
- $1-\overline{\alpha_1}z$ divides $q - \overline{c_0\gamma_0}p$ if $\alpha_1 \neq 0$, or else the degree of $q-\overline{c_0\gamma_0}p$ is $\leq n-1$.

Note that, in this case, $d^o f - d^o f_1 = 1$.

Suppose $\alpha_1 \in \mathcal{R}$. Then $(p\tilde{p} - q\tilde{q})(\alpha_1) = 0$. As $\frac{p}{q}$ is irreducible and analytic in $\mathbb{D}$, $q(\alpha_1) \neq 0$. Thus

$$\frac{(q\tilde{q} - p\tilde{p})(\alpha_1)}{q(\alpha_1)} = \tilde{q}(\alpha_1) - c_0\gamma_0\tilde{p}(\alpha_1) = 0. \tag{66}$$

If $\alpha_1 \neq 0$, then

$$\overline{\alpha_1}^n q\left(\frac{1}{\overline{\alpha_1}}\right) - \overline{c_0\gamma_0}\cdot\overline{\alpha_1}^n p\left(\frac{1}{\overline{\alpha_1}}\right) = 0.$$

We deduce that $1-\overline{\alpha_1}z$ divides $q-\overline{c_0\gamma_0}p$. If $\alpha_1 = 0$, by (66), the degree of $q-\overline{c_0\gamma_0}p$ is strictly less than n. Furthermore, by definition of γ_0, $z-\alpha_1$ divides $p - c_0\gamma_0 q$. Thus, $\deg f_1 = n-1$.

Conversely, if $\alpha_1 \neq 0$ with $p(\alpha_1) - c_0\gamma_0 q(\alpha_1) = 0$ and $q(\frac{1}{\overline{\alpha_1}}) - \overline{c_0\gamma_0}p(\frac{1}{\overline{\alpha_1}}) = 0$, then $\tilde{q}(\alpha_1) - c_0\gamma_0\tilde{p}(\alpha_1) = 0$, from which it follows that $\alpha_1 \in \mathcal{R}$. If $\alpha_1 = 0$ and $p(0) = c_0\gamma_0 q(0)$ with $\deg(q - \bar{c}_0\bar{\gamma}_0 p) < n$, then $\tilde{q}(0) - c_0\gamma_0\tilde{p}(0) = 0$ and again $\alpha_1 \in \mathcal{R}$.

3. We finally prove that if $f_1 = \frac{p_1}{q_1}$, then the roots of $p_1\tilde{p_1} - q_1\tilde{q_1}$ that lie in the unit disk are the points of $\mathcal{R}\setminus\{\alpha_1\}$ (counting multiplicity). Since

$$\begin{pmatrix} p_1 & \bar{c}_0\tilde{q_1} \\ q_1 & \bar{c}_0\tilde{p_1} \end{pmatrix} = \begin{pmatrix} z-\alpha_1 & 0 \\ 0 & 1-\overline{\alpha_1}z \end{pmatrix}^{-1} \begin{pmatrix} \bar{c}_0 & -\gamma_0 \\ -\overline{c_0\gamma_0} & 1 \end{pmatrix} \begin{pmatrix} p & \tilde{q} \\ q & \tilde{p} \end{pmatrix},$$

taking determinants, we get

$$p_1\tilde{p_1} - q_1\tilde{q_1} = (1-|\gamma_0|^2)\frac{p\tilde{p} - q\tilde{q}}{(z-\alpha_1)(1-\overline{\alpha_1}z)}.$$

Therefore, the set of the roots of $p_1\tilde{p_1} - q_1\tilde{q_1}$ in $\mathbb{D}$ is $\mathcal{R}\setminus\{\alpha_1\}$.

Iterating this process n times, we get $f_n(z) = \gamma_n$. Conclusion is then immediate. ∎

We endow the space of rational functions of degree n with the differential structure which is naturally inherited from the coefficients of the numerators and denominators. Then it becomes a smooth submanifold of every Hardy space H^p, $1 < p < \infty$, of the disk of dimension $2n+1$ over $\mathbb{C}$ ([3]).

Theorem 7.2 *If $a = (\alpha_1, \dots, \alpha_n, \gamma_0, \dots, \gamma_n)$ is such that the points $\alpha_1, \dots, \alpha_n$ are all distinct and $d^\circ\Gamma(a) = n$, then the derivative $d\Gamma(a)$ at $a \in \mathbb{D}^{2n+1}$ is an isomorphism.*

Proof We give a proof by induction. The result is immediate if $n = 0$. We denote by Γ_i:

$$\Gamma_i(\alpha_{i+1}, \dots, \alpha_n, \gamma_i, \dots, \gamma_n) = \tau_i \circ \dots \circ \tau_n(0).$$

We therefore have

$$\begin{aligned}\Gamma(\alpha_1, \dots, \alpha_n, \gamma_0, \dots, \gamma_n) &= \tau_0 \circ \Gamma_1(\alpha_2, \dots, \alpha_n, \gamma_1, \dots, \gamma_n) \\ &= \frac{\zeta_1\Gamma_1(\alpha_2, \dots, \alpha_n, \gamma_1, \dots, \gamma_n) + \gamma_0}{1 + \overline{\gamma_0}\zeta_1\Gamma_1(\alpha_2, \dots, \alpha_n, \gamma_1, \dots, \gamma_n)}.\end{aligned}$$

Note that, in the following, we will just write Γ_1 for $\Gamma_1(\alpha_2, \dots, \alpha_n, \gamma_1, \dots, \gamma_n)$. On differentiating if the space of rational functions of degree n is viewed as a submanifold of H^p, $1 < p < \infty$, we have

$$\begin{aligned}\frac{\partial\Gamma}{\partial\gamma_0} &= \frac{1}{1 + \overline{\gamma_0}\zeta_1(z)\Gamma_1(z)} \\ \frac{\partial\Gamma}{\partial\overline{\gamma_0}} &= -\frac{\zeta_1(z)\Gamma_1(z)(\zeta_1(z)\Gamma_1(z) + \gamma_0)}{(1 + \overline{\gamma_0}\zeta_1(z)\Gamma_1(z))^2} \\ \frac{\partial\Gamma}{\partial\alpha_1} &= -\frac{\Gamma_1(z)}{(1 + \overline{\gamma_0}\zeta_1(z)\Gamma_1(z))^2}\frac{1 - |\gamma_0|^2}{1 - \overline{\alpha_1}z} \\ \frac{\partial\Gamma}{\partial\overline{\alpha_1}} &= \frac{\zeta_1(z)\Gamma_1(z)}{(1 + \overline{\gamma_0}\zeta_1(z)\Gamma_1(z))^2}\frac{(1 - |\gamma_0|^2)z}{1 - \overline{\alpha_1}z}\end{aligned}$$

and for $k \geq 1$,

$$\begin{aligned}\frac{\partial\Gamma}{\partial\gamma_k} &= \frac{\zeta_1(z)(1 - |\gamma_0|^2)}{(1 + \overline{\gamma_0}\zeta_1(z)\Gamma_1(z))^2}\frac{\partial\Gamma_1}{\partial\gamma_k} \\ \frac{\partial\Gamma}{\partial\overline{\gamma_k}} &= \frac{\zeta_1(z)(1 - |\gamma_0|^2)}{(1 + \overline{\gamma_0}\zeta_1(z)\Gamma_1(z))^2}\frac{\partial\Gamma_1}{\partial\overline{\gamma_k}} \\ \frac{\partial\Gamma}{\partial\alpha_{k+1}} &= \frac{\zeta_1(z)(1 - |\gamma_0|^2)}{(1 + \overline{\gamma_0}\zeta_1(z)\Gamma_1(z))^2}\frac{\partial\Gamma_1}{\partial\alpha_{k+1}} \\ \frac{\partial\Gamma}{\partial\overline{\alpha_{k+1}}} &= \frac{\zeta_1(z)(1 - |\gamma_0|^2)}{(1 + \overline{\gamma_0}\zeta_1(z)\Gamma_1(z))^2}\frac{\partial\Gamma_1}{\partial\overline{\alpha_{k+1}}}\end{aligned}$$

Suppose that the hypothesis is true for $n-1$, that is if $\alpha_2, \dots, \alpha_n$ are all distinct and $d^\circ\Gamma_1(\hat{a}) = n - 1$ then $d\Gamma_1(\hat{a})$ is an isomorphism, with $\hat{a} = (\alpha_2, \dots, \alpha_n, \gamma_1, \dots, \gamma_n)$. Suppose there exists a linear combination such that:

$$\sum_{l=0}^{n-1}\left(\frac{\partial\Gamma}{\partial\gamma_l}d\gamma_l + \frac{\partial\Gamma}{\partial\overline{\gamma_l}}d\overline{\gamma_l} + \frac{\partial\Gamma}{\partial\alpha_{l+1}}d\alpha_{l+1} + \frac{\partial\Gamma}{\partial\overline{\alpha_{l+1}}}d\overline{\alpha_{l+1}}\right) + \frac{\partial\Gamma}{\partial\gamma_n}d\gamma_n + \frac{\partial\Gamma}{\partial\overline{\gamma_n}}d\overline{\gamma_n} = 0$$

Then we have for every z, on multiplying by $(1+\overline{\gamma_0}\zeta_1(z)\Gamma_1(z))^2$,

$$\begin{aligned}0 = \zeta_1(z)(1-|\gamma_0|^2)\sum_{l=1}^{n-1}\left(\frac{\partial\Gamma_1}{\partial\gamma_l}d\gamma_l + \frac{\partial\Gamma_1}{\partial\overline{\gamma_l}}d\overline{\gamma_l} + \frac{\partial\Gamma_1}{\partial\alpha_{l+1}}d\alpha_{l+1} + \frac{\partial\Gamma_1}{\partial\overline{\alpha_{l+1}}}d\overline{\alpha_{l+1}}\right)\\ + \zeta_1(z)(1-|\gamma_0|^2)\left(\frac{\partial\Gamma_1}{\partial\gamma_n}d\gamma_n + \frac{\partial\Gamma_1}{\partial\overline{\gamma_n}}d\overline{\gamma_n}\right)\\ + \zeta_1(z)\Gamma_1(z)\left(\overline{\gamma_0}d\gamma_0 - (\zeta_1(z)\Gamma_1(z)+\gamma_0)d\overline{\gamma_0} + \frac{(1-|\gamma_0|^2)z}{1-\overline{\alpha_1}z}d\overline{\alpha_1}\right)\\ + d\gamma_0 - \frac{(1-|\gamma_0|^2)\Gamma_1(z)}{1-\overline{\alpha_1}z}d\alpha_1.\end{aligned} \tag{67}$$

Evaluating at α_1, we get

$$d\gamma_0 = \frac{\Gamma_1(\alpha_1)(1-|\gamma_0|^2)}{1-|\alpha_1|^2}d\alpha_1 \tag{68}$$

Therefore, the last row in (67) can be expressed as :

$$(1-|\gamma_0|^2)\left(\frac{\Gamma_1(\alpha_1)}{1-|\alpha_1|^2} - \frac{\Gamma_1(z)}{1-\overline{\alpha_1}z}\right)d\alpha_1$$

This can be written as

$$(|\gamma_0|^2-1)\zeta_1(z)\left(g_1(z) + \overline{\alpha_1}\frac{\Gamma_1(\alpha_1)}{1-|\alpha_1|^2}\right)d\alpha_1$$

with

$$g_1(z) = \frac{\Gamma_1(z)-\Gamma_1(\alpha_1)}{z-\alpha_1}.$$

A cancellation by ζ_1 in (67) gives us:

$$\begin{aligned}0 = (1-|\gamma_0|^2)\sum_{l=1}^{n-1}\left(\frac{\partial\Gamma_1}{\partial\gamma_l}d\gamma_l + \frac{\partial\Gamma_1}{\partial\overline{\gamma_l}}d\overline{\gamma_l} + \frac{\partial\Gamma_1}{\partial\alpha_{l+1}}d\alpha_{l+1} + \frac{\partial\Gamma_1}{\partial\overline{\alpha_{l+1}}}d\overline{\alpha_{l+1}}\right)\\ + (1-|\gamma_0|^2)\left(\frac{\partial\Gamma_1}{\partial\gamma_n}d\gamma_n + \frac{\partial\Gamma_1}{\partial\overline{\gamma_n}}d\overline{\gamma_n}\right)\\ + \Gamma_1(z)\left(\overline{\gamma_0}d\gamma_0 - (\zeta_1(z)\Gamma_1(z)+\gamma_0)d\overline{\gamma_0} + \frac{(1-|\gamma_0|^2)z}{1-\overline{\alpha_1}z}d\overline{\alpha_1}\right)\\ + (|\gamma_0|^2-1)\left(g_1(z) + \overline{\alpha_1}\frac{\Gamma_1(\alpha_1)}{1-|\alpha_1|^2}\right)d\alpha_1.\end{aligned} \tag{69}$$

Γ_1 is a rational irreducible function $\frac{p_1}{q_1}$ of degree $n-1$ by Theorem 7.1. Thus $\frac{\partial\Gamma_1}{\partial\Box} \in \frac{\mathcal{P}_{2n-2}}{q_1^2}$ where $\Box$ denotes any of the variable α_j, γ_j, $\bar{\alpha}_j$ or $\bar{\gamma}_j$. In fact, in the previous expression, all terms are in $\frac{\mathcal{P}_{2n-2}}{q_1^2}$, except perhaps

$$-\zeta_1(z)\Gamma_1^2(z)d\overline{\gamma_0}$$

and

$$(1-|\gamma_0|^2)\frac{z\Gamma_1(z)}{1-\overline{\alpha_1}z}d\overline{\alpha_1}.$$

Using (68) and (69), we get

$$\left(-\zeta_1(z)\Gamma_1^2(z)\frac{\overline{\Gamma_1(\alpha_1)}}{1-|\alpha_1|^2}+\frac{z\Gamma_1(z)}{1-\overline{\alpha_1}z}\right)d\overline{\alpha_1}\in\frac{\mathcal{P}_{2n-2}}{q_1(z)^2}. \tag{70}$$

Note that

$$-\zeta_1(z)\Gamma_1^2(z)\frac{\overline{\Gamma_1(\alpha_1)}}{1-|\alpha_1|^2}+\frac{z\Gamma_1(z)}{1-\overline{\alpha_1}z}=p_1(z)\frac{(1-|\alpha_1|^2)zq_1(z)-\overline{\Gamma_1(\alpha_1)}(z-\alpha_1)p_1(z)}{(1-|\alpha_1|^2)(1-\bar{\alpha}_1z)q_1(z)^2}. \tag{71}$$

Suppose that $d\overline{\alpha_1}\neq 0$.
Then, if $\alpha_1\neq 0$, combining (70) and (71), we get

$$p_1(1/\overline{\alpha_1})\left(q_1(1/\overline{\alpha_1})-\overline{\Gamma_1(\alpha_1)}p_1(1/\overline{\alpha_1})\right)=0.$$

If $p_1(1/\overline{\alpha_1})=0$, then

$$\frac{p(z)}{q(z)}=\frac{(z-\alpha_1)p_1(z)+c_0\gamma_0q_1(z)(1-\overline{\alpha_1}z)}{q_1(z)(1-\overline{\alpha_1}z)+\overline{c_0\gamma_0}(z-\alpha_1)p_1(z)}$$

has the same degree than $\frac{p_1}{q_1}$ (because $1-\overline{\alpha_1}z$ is a common factor).
If $q_1(1/\overline{\alpha_1})-\overline{\Gamma_1(\alpha_1)}p_1(1/\overline{\alpha_1})=0$, then $(p_1\tilde{p_1}-q_1\tilde{q_1})(\alpha_1)=0$ and α_1 is a multiple root.
Furthermore, if $\alpha_1=0$, we have $zp_1(z)\left(q_1(z)-\overline{\Gamma_1(0)}p_1(z)\right)\in\mathcal{P}_{2n-2}$ if and only if $\deg(zp_1(z))\leq n-1$ or $\deg(q_1(z)-\overline{\Gamma_1(0)}p_1(z))\leq n-2$, which is equivalent to $\deg(zp_1(z))\leq n-1$ or $(p\tilde{p}-q\tilde{q})(0)=0$.
From what precedes, we deduce that if $\deg p/q=n$ and α_1 is not a multiple root, then the derivative $d\Gamma(a)$ is injective (and therefore surjective counting dimensions).

■

7.2 Computation of the L^2 norm

In order to be able to optimize with respect to the L^2 norm, we will now see how to numerically compute efficiently the Hermitian product $\langle f,g\rangle=\int_{\mathbb{T}}f(t)\overline{g(t)}dm(t)$ for f,g rational functions analytic inside the unit disk. Two kind of methods are presented : the first one uses elementary operations on polynomials, and the other one uses matrix operations.

7.2.1 Two methods using elementary operations on polynomials

The two methods proposed brings the computation of the Hermitian product of two rational functions back to the computation of the Hermitian product of two polynomials. Therefore, they essentially use the elementary property :

Property 7.3 *If* $p = \sum_{k=0}^{k=m} p_k z^k$ *and* $q = \sum_{k=0}^{k=n} q_k z^k$ *are two polynomials then*

$$\langle p, q \rangle = \sum_{k=0}^{\min(m,n)} p_k \overline{q_k}.$$

The first method is very basic and gives an approximation of the Hermitian product. However, it is quite efficient for Schur rational functions of small degree. It simply consists in approximating f and g by their Taylor polynomials of order N, the Hermitian product is then obtained using the previous property. If N is sufficiently big, the result is very good (for the examples presented in the next section, two hundred Taylor coefficients were taken). The Taylor coefficients are easily obtained using the "long" division with respect to increasing powers.
The second method has the advantage of avoiding any truncation. However, it requires to efficiently compute an extended gcd. For a neater notation, the following computation is done for $\frac{a}{b}$ and $\frac{r}{q}$ rational functions *analytic outside the unit disk*, i.e. the roots of b and q are in the unit disk. This is equivalent to the corresponding problem in the disk upon changing z into $1/z$. Here, for a polynomial q, we denote by $\tilde{q}$ the polynomial $\tilde{q} = z^{d^o q}\overline{q\left(\frac{1}{\bar{z}}\right)}$. As $gcd(b, \tilde{q}) = 1$, there exist u and v such that $ub + v\tilde{q} = 1$. Then, if $r = r_1 q + r_0$ with $d^o r_0 < d^o r$,

$$\begin{aligned}
\left\langle \frac{a}{b}, \frac{q}{r} \right\rangle &= \left\langle \frac{a(ub + v\tilde{q})}{b}, \frac{r}{q} \right\rangle \\
&= \left\langle au, \frac{r}{q} \right\rangle + \left\langle \frac{av\tilde{q}}{b}, \frac{r}{q} \right\rangle \\
&= \langle au, r_1 \rangle + \left\langle \frac{av\tilde{q}}{b}, \frac{r}{q} \right\rangle
\end{aligned}$$

where we have taken into account the orthogonality of $H^2(\mathbb{D})$ and $H^2(\mathbb{C} \setminus \overline{\mathbb{D}})$. As $\tilde{q} = z^{d^o q}\overline{q\left(\frac{1}{\bar{z}}\right)}$, we have

$$\left\langle \frac{av\tilde{q}}{b}, \frac{r}{q} \right\rangle = \left\langle \frac{avz^{d^o q}}{b}, r \right\rangle.$$

The euclidean division of $avz^{d^o q}$ by b gives

$$avz^{d^o q} = k_1 b + \rho.$$

Therefore,

$$\left\langle \frac{a}{b}, \frac{q}{r} \right\rangle = \langle au, r_1 \rangle + \langle k_1, r \rangle.$$

Note that the Hermitian product of two rational functions $f = \frac{a_0}{b_0}$ and $g = \frac{r_0}{q_0}$ analytic inside the unit disk is

$$\begin{aligned} \langle f, g \rangle &= \left\langle \frac{a_0}{b_0}, \frac{r_0}{q_0} \right\rangle \\ &= \left\langle \frac{z^{d^o q_0}}{z^{d^o r_0}} \frac{\tilde{r_0}}{\tilde{q_0}}, \frac{z^{d^o b_0}}{z^{d^o a_0}} \frac{\tilde{a_0}}{\tilde{b_0}} \right\rangle \end{aligned}$$

and is therefore obtained as a Hermitian product of two rational function analytic outside the disk.

7.2.2 A method using matrix representations

We now present a method which adopts the matrix point of view. The computation is carried out using a realization of f and g, i.e. by expressing these functions with matrices. More details about realizations and system theory can be found in [19].

Definition 7.4 *A rational function is proper (resp. strictly proper) if the numerator's degree is less or equal (resp. strictly less) than the denominator's degree.*
A matrix is proper rational (resp. strictly proper rational) if its entries are rational proper (resp. strictly proper) functions.

In fact, we will study here how to compute the L^2 norm of proper rational matrices. For this, we first want to express strictly proper rational matrices using 3 complex matrices A, B, C.
Let $H(s)$ be a strictly proper rational matrix $m \times p$ and let $d(s) = s^r + d_1 s^{r-1} + ... + d_r$ be the least common denominator of the entries of $H(s)$. Then $H(s) = \frac{N(s)}{d(s)}$, where $N(s)$ is a matrix $m \times p$ with polynomial entries. As H is strictly proper, there exist complex matrices $m \times p$ $N_1, N_2, ..., N_r$ such that $N(s) = N_1 s^{r-1} + N_2 s^{r-2} + ... + N_r$.
We denote by I_p the $p \times p$ identity matrix.
We define the matrices $A : pr \times pr$, $B : pr \times p$, $C : m \times pr$ by :

$$\left\{ \begin{array}{l} A = \begin{bmatrix} -d_1 I_p & -d_2 I_p & \cdots & -d_r I_p \\ I_p & 0 & \cdots & 0 \\ & \ddots & \ddots & \\ (0) & & I_p & 0 \end{bmatrix}, \\ B = \begin{bmatrix} I_p \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \\ C = \begin{bmatrix} N_1 & N_2 & \cdots & N_r \end{bmatrix}. \end{array} \right.$$

Then:

$$
\begin{aligned}
&(sI-A)\begin{bmatrix} s^{r-1}I_p & s^{r-2}I_p & \cdots & I_p\end{bmatrix}^t \\
&= \begin{bmatrix} (s+d_1)I_p & d_2I_p & \cdots & d_rI_p \\ -I_p & sI_p & & (0) \\ & \ddots & \ddots & \\ (0) & & -I_p & sI_p \end{bmatrix}\begin{bmatrix} s^{r-1}I_p \\ s^{r-2}I_p \\ \vdots \\ I_p \end{bmatrix} \\
&= \begin{bmatrix} (s^r+d_1s^{r-1}+d_2s^{r-2}+...+d_r)I_p \\ (-s^{r-1}+s^{r-1})I_p \\ \vdots \\ -sI+sI \end{bmatrix} \\
&= d(s)\begin{bmatrix} I_p \\ 0 \\ \vdots \\ 0 \end{bmatrix}.
\end{aligned}
$$

We deduce that $(sI-A)^{-1} = \frac{1}{d(s)}\begin{bmatrix} s^{r-1}I_p & * & \cdots & * \\ s^{r-2}I_p & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ I_p & * & \cdots & * \end{bmatrix}$. Therefore,

$$
\begin{aligned}
&C(sI-A)^{-1}B \\
&= \frac{1}{d(s)}\begin{bmatrix} N_1 & N_2 & \cdots & N_r \end{bmatrix}\begin{bmatrix} s^{r-1}I_p & s^{r-2}I_p & \cdots & I_p\end{bmatrix}^t \\
&= \frac{N_1s^{r-1}+N_2s^{r-2}+...+N_r}{d(s)} \\
&= \frac{N(s)}{d(s)} \\
&= H(s).
\end{aligned}
$$

Definition 7.5 *Let $H(s)$ be a proper rational matrix. We call realization of H any 4-tuple (A,B,C,D) of complex matrices such that $H(s) = C(sI-A)^{-1}B + D$.*

From what precedes, a realization of a strictly proper rational matrix always exists. Let now H be proper rational and let $D = \lim_{s\to\infty} H$. Then $H-D$ is strictly proper, so there exists (A,B,C) such that $H-D = C(sI-A)^{-1}B$. Therefore, $H = C(sI-A)^{-1}B+D$. Thus, we have obtained a realization for a proper rational matrix. Note that a proper rational matrix does not have a unique realization.
A realization is called a *minimal realization of* H if the size of A is minimal among all the possible realizations of H.

We now briefly explain how to compute the L^2 norm using a minimal realization. We now suppose that (A, B, C, D) is a minimal realization of a proper rational matrix H whose entries are analytic outside the unit disk and up to the unit circle. It is well-known that the eigenvalues of A are the poles of H ([19], [14]). By analyticity of H, the eigenvalues of A are therefore inside the unit disk. We have

$$(sI - A)^{-1} = s^{-1}\left(I - \frac{A}{s}\right)^{-1} = s^{-1}\sum_{j=0}^{\infty}\left(\frac{A}{s}\right)^{j} = \sum_{j=0}^{\infty} A^j s^{-(j+1)}.$$

Therefore, $H(s) = D + \sum_{j=0}^{\infty} CA^jBs^{-(j+1)}$. Let H_1 and H_2 be two strictly proper rational matrices whose entries are analytical outside the unit disk. From what precedes, we have

$$\begin{cases} H_1(s) = D_1 + \sum_{j=0}^{\infty} C_1A_1^jB_1s^{-(j+1)}, \text{ and} \\ H_2(s) = D_2 + \sum_{j=0}^{\infty} C_2A_2^jB_2s^{-(j+1)}. \end{cases}$$

Thus

$$\begin{aligned} \langle H_1, H_2\rangle &= Tr\left(D_1D_2^* + \sum_{j=0}^{\infty} C_1A_1^jB_1B_2^*(A_2^*)^jC_2^*\right) \\ &= Tr\left(D_1D_2^* + C_1\left(\sum_{j=0}^{\infty} A_1^jB_1B_2^*(A_2^*)^j\right)C_2^*\right). \end{aligned}$$

We denote by P the matrix $P = \sum_{j=0}^{\infty} A_1^jB_1B_2^*(A_2^*)^j$, which is well-defined since A_1 and A_2 have all their eigenvalues in $\mathbb{D}$. It is immediate that P is a solution of the Stein (or Lyapounov) equation: $A_1PA_2^* + B_1B_2^* = P$. Since all the eigenvalues of A_1 and A_2 are in $\mathbb{D}$, no eigenvalue of A_1 is the reciprocal of an eigenvalue of A_2. Therefore, the Stein problem has a unique solution. Since $\langle H_1, H_2\rangle = Tr(D_1D_2^* + C_1PC_2^*)$, solving the Stein problem gives the value of $\langle H_1, H_2\rangle$.
More details about the matrix P and the Stein problem can be found in [5].

7.3 Examples

In order to approximate a function f, we have implemented an optimization process using the parametrization presented in section 7.1. The criterion which is minimized is the relative L^2 error

$$e(\alpha_1, \ldots, \alpha_n, \gamma_0, \ldots, \gamma_n) = \frac{\|f - \Gamma(\alpha_1, \ldots, \alpha_n, \gamma_0, \ldots, \gamma_n)\|_2}{\|f\|_2}.$$

In practice, the points of the unit disk $\alpha_1, \ldots, \alpha_n, \gamma_0, \ldots, \gamma_n$ are parametrized by the application

$$\Lambda : \begin{array}{l} \mathbb{R}^2 \longrightarrow \mathbb{D} \\ (x, y) \mapsto \frac{x}{\sqrt{x^2+y^2+1}} + i\frac{y}{\sqrt{x^2+y^2+1}} \end{array}.$$

This allows to do an unconstrained optimization : to compute a Schur rational function of degree n, we would like to optimize

$$\inf_{(x_{\alpha_1}, y_{\alpha_1}, \ldots, x_{\gamma_n}, y_{\gamma_n}) \in \mathbb{R}^{4n+2}} \|f - \Gamma(\Lambda(x_{\alpha_1}, y_{\alpha_1}), \ldots, \Lambda(x_{\gamma_n}, y_{\gamma_n}))\|_2.$$

This problem depends of $4n+2$ real parameters. Note that, as the parametrization Γ is not defined for parameters of modulus 1, the infimum is not necessarily attained. In the following examples, the initialization of the optimization is done using the asymptotic-BMO-type criterion (see section 6.2.2), that is by computing a sequence of points (α_n) such that α_{n+1} minimizes

$$\mathcal{I}_n(\alpha) = \int_{\mathbb{T}} \left| \frac{f_n(t) - f_n(\alpha)}{1 - \overline{f_n(\alpha)} f_n(t)} \right|^2 P(t, \alpha) dm(t).$$

No refined attempts at solving this optimization problem were made: we simply used a grid search.
The results obtained by this "Schur optimization" are compared with the L^2 unconstrained approximation given by the hyperion software[1] ([15]). In particular, we check that the error of our result s lies between the L^2 error of the result h given by hyperion and the "normalized L^2 error" (i.e. the error of the arl2 function of the hyperion software scaled into the unit disk in order to obtain a Schur function), that is we check that $e(h) \leq e(s) \leq e\left(\frac{h}{\|h\|_\infty}\right)$.

In the following figures, when a function g is plotted, the left graph represents the image by g of the unit circle, and the right graph is the modulus of this image, i.e. we plot:

On the left: $t \mapsto g(e^{it})$ and on the right: $t \mapsto |g(e^{it})|$ for $-\pi \leq t \leq \pi$.

7.3.1 Approximation of Schur functions

Example 1
We are now interested in approximating a polynomial $p30$ of degree 30 plotted in Fig. 1. Note that $p30$ is Schur and $\|p30\|_2 = 0.7852$.
The results given by our optimization process and by hyperion for degrees 7 to 9 are presented in Tab. 1. None of the best L^2-unconstrained approximations is Schur.
Fig. 2 is a good example of what happens when one approximates a Schur function whose modulus is near 1 on an interval of the unit circle: the L^2 unconstrained approximation oscillates (in modulus) around one. Here, where the approximation computed by hyperion exceeds 1 (in modulus), the Schur approximation "hits" one. On this example, the initialization points are not very good (see fig. 3, 5 and 7).

[1]The hyperion software essential feature is to find a rational approximation of McMillan degree n of a stable transfer function given by incomplete frequency measures. Its development has been

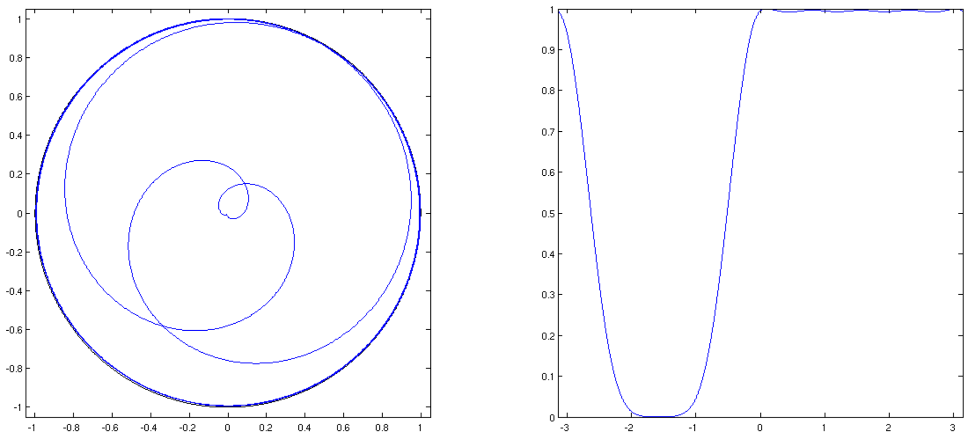


Figure 1: Function $p30$, polynomial of degree 30, Schur.

		Degree 7	Degree 8	Degree 9
L^2	$\|.\|_\infty$	1.0235	1.0056	1.0014
(hyperion)	error	6.72 e-2	1.16 e-2	1.32 e-3
Schur	error	6.89 e-2	1.19 e-2	1.51 e-3
L^2 normalized	error	7.09 e-2	1.29 e-2	1.99 e-3

Table 1: Approximation of the Schur function $p30$: comparison between our Schur process and hyperion

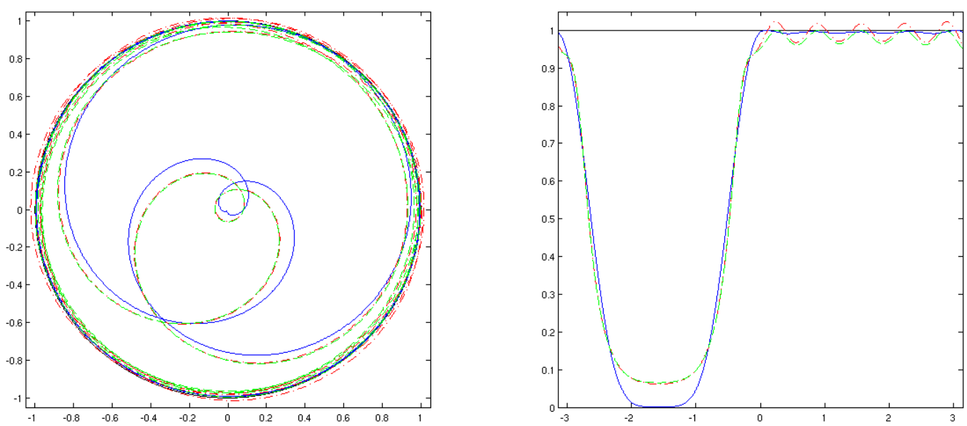


Figure 2: Function $p30$ (blue), Schur approximation (green) and L^2 approximation (red) of degree 7.

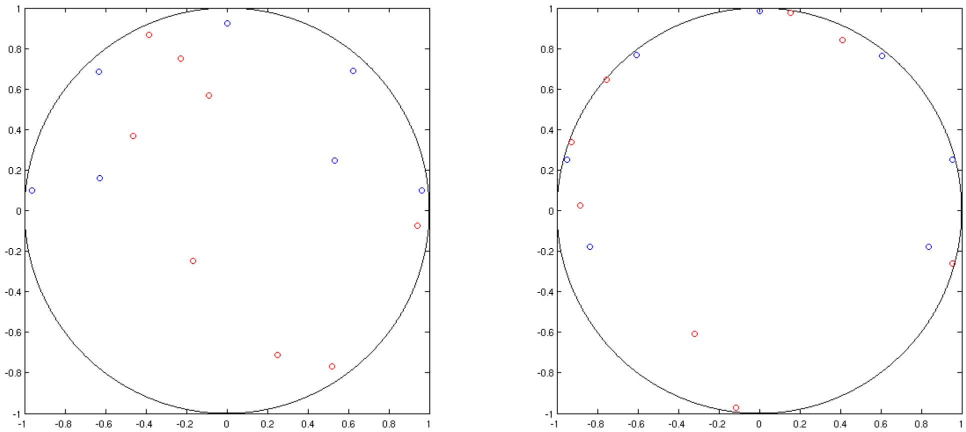


Figure 3: Initialization points (left) and optimized points (right) of the Schur function of degree 7 : parameters α (blue) and γ (red).

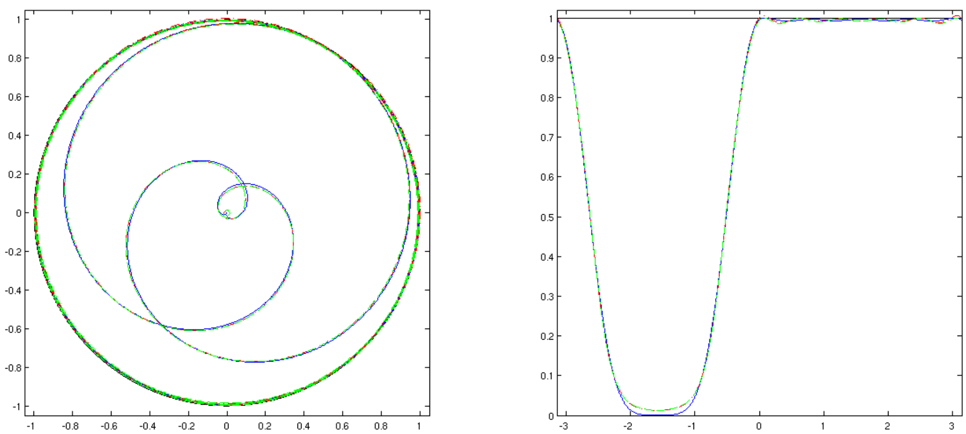


Figure 4: Function $p30$ (blue), Schur approximation (green) and L^2 approximation (red) of degree 8.

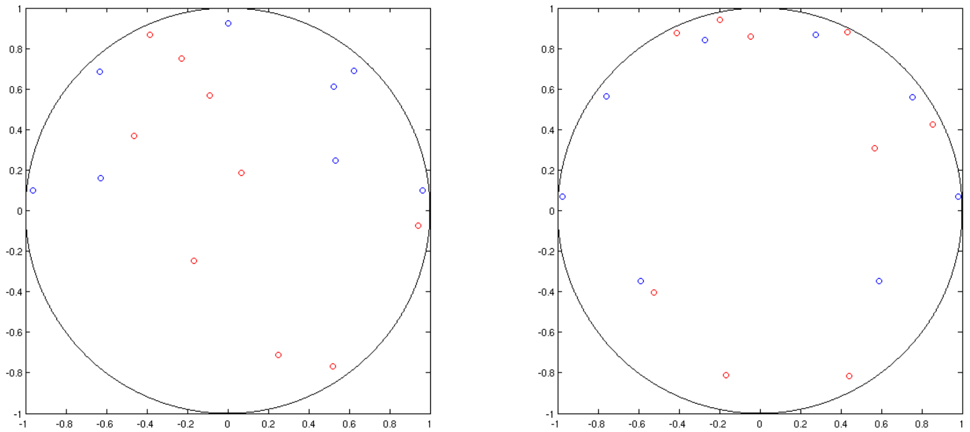


Figure 5: Initialization points (left) and optimized points (right) of the Schur function of degree 8 : parameters α (blue) and γ (red).

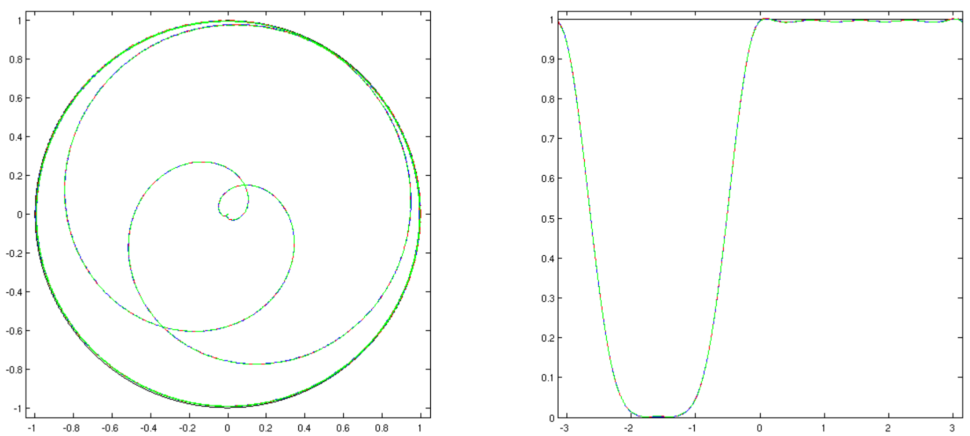


Figure 6: Function $p30$ (blue), Schur approximation (green) and L^2 approximation (red) of degree 9.

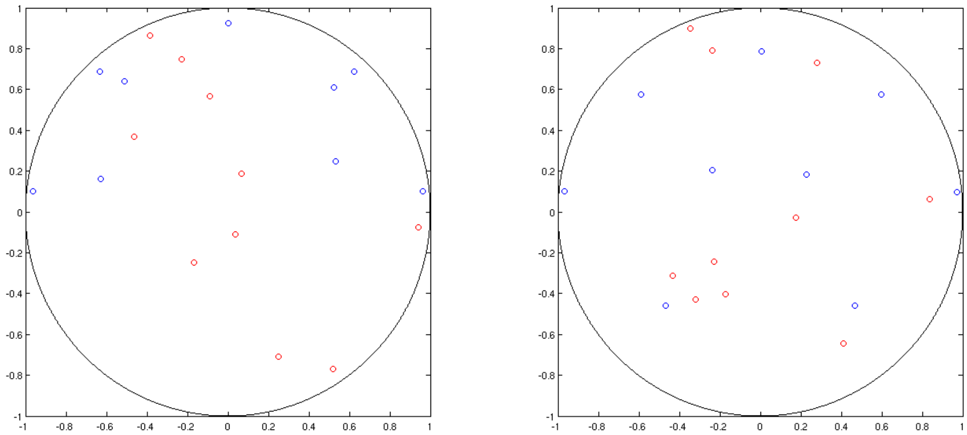


Figure 7: Initialization points (left) and optimized points (right) of the Schur function of degree 9 : parameters α (blue) and γ (red).

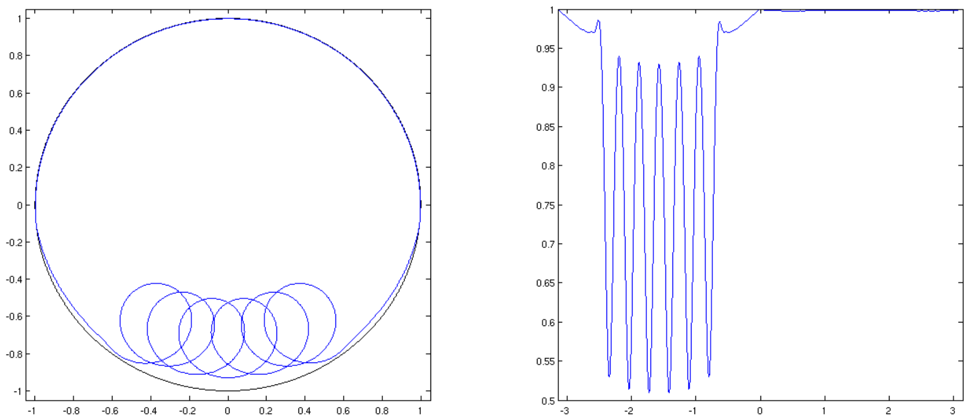


Figure 8: Function f, polynomial of degree 60.

Example 2

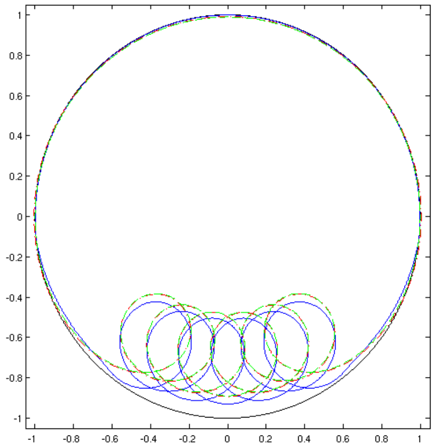
We are now interested in approximating a polynomial $p60$ of degree 60 plotted in fig. 8. Note that $p60$ is Schur and $\|p60\|_2 = 0.9304$.

The approximations of degree 7 to 9 obtained using our Schur process and hyperion are compared in Tab. 2. Note that none of the best L^2-unconstrained approximations is Schur.

abandoned in 2001. The Endymion software, which is still under development, will offer most of the functionalities of hyperion. Note that the author of the hyperion software chose to write “hyperion” in lowercase letters.

		Degree 7	Degree 8	Degree 9
L^2	$\|\|.\|\|_\infty$	1.0053	1.0037	1.0014
(hyperion)	error	2.97 e-2	1.69 e-2	4.5 e-3
Schur	error	3.01 e-2	1.70 e-2	4.7 e-3
L^2 normalized	error	3.02 e-2	1.73 e-2	4.8 e-3

Table 2: Approximation of the Schur function $p60$: comparison between our Schur process and hyperion



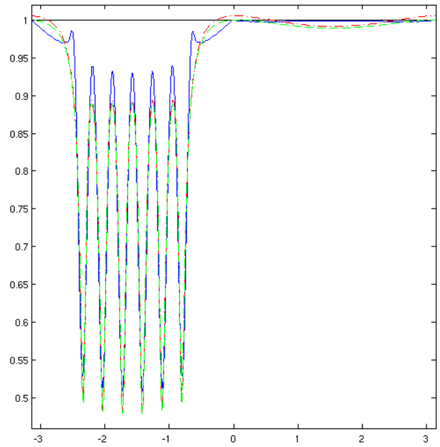


Figure 9: Function $p60$ (blue), Schur approximation (green) and L^2 approximation (red) of degree 7.

For the initialization, we first computed points $\alpha_1, \ldots, \alpha_{10}$ using the asymptotic-BMO-type criterion and chose among them. The initial interpolation points at degree 7 are the points $\alpha_2, \ldots, \alpha_8$, at degree 8 they are $\alpha_1, \ldots, \alpha_8$, and at degree 9 they are $\alpha_2, \ldots, \alpha_{10}$. The initializations for the degrees 7 and 8 are quite good (see fig. 10 and fig. 12).

7.3.2 Approximation of analytic but not Schur functions

In the two following examples, we are interested in approximating analytic, but not Schur, functions. In practice, standard applications arise from the fact that the function is known to be Schur, but some measurement errors occurred and lead to a function with values greater than 1 in modulus at some places.

Example 3

An example is taken of a rational function $r5$ of degree 5 such that $\|r5\|_\infty = 1.01$ and $\|r5\|_2 = 0.6225$. Note that $r5$ is not Schur but is analytic in the unit disk.

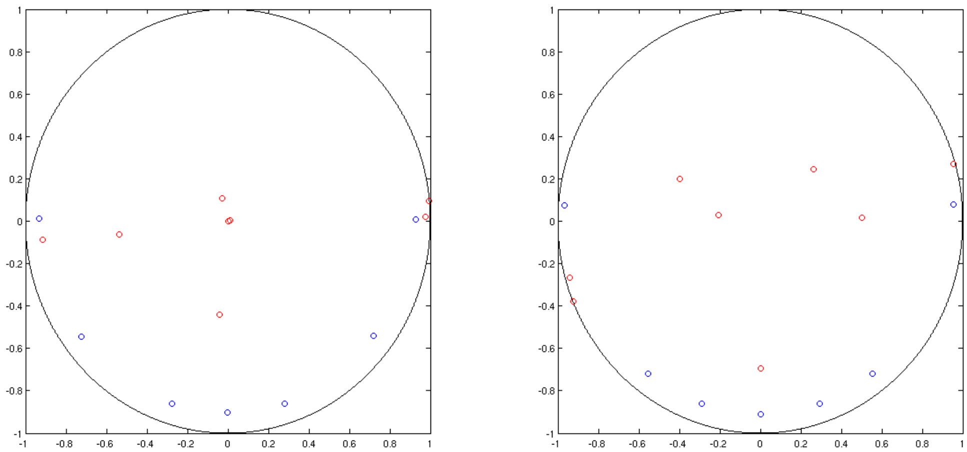


Figure 10: Initialization points (left) and optimized points (right) of the Schur function of degree 7 : parameters α (blue) and γ (red).

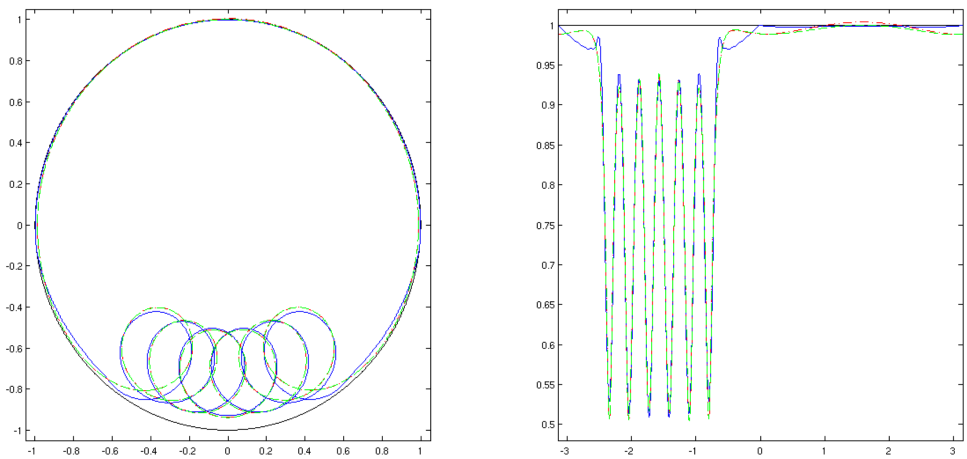


Figure 11: Function $p60$ (blue), Schur approximation (green) and L^2 approximation (red) of degree 8.

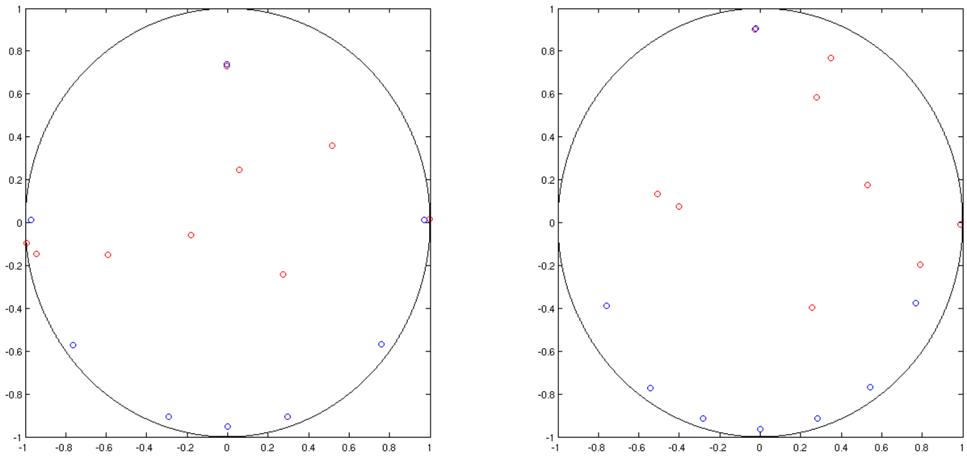


Figure 12: Initialization points (left) and optimized points (right) of the Schur function of degree 8 : parameters α (blue) and γ (red).

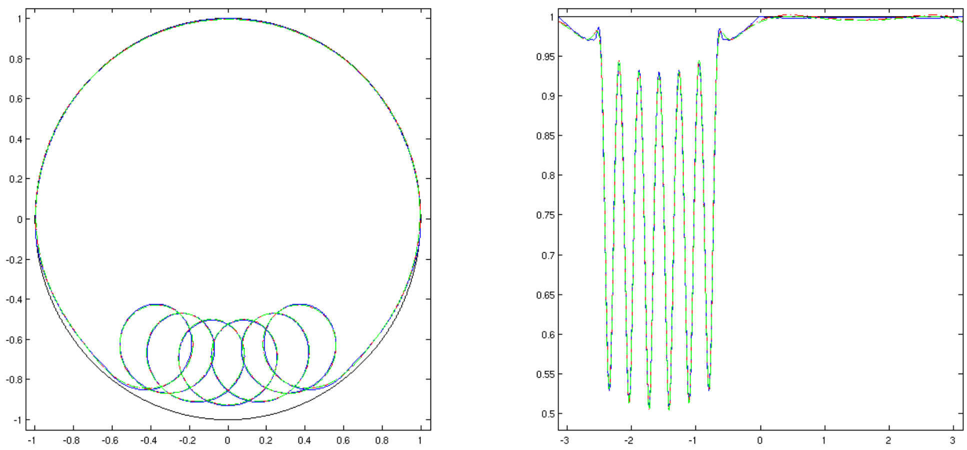


Figure 13: Function $p60$ (blue), Schur approximation (green) and L^2 approximation (red) of degree 9.

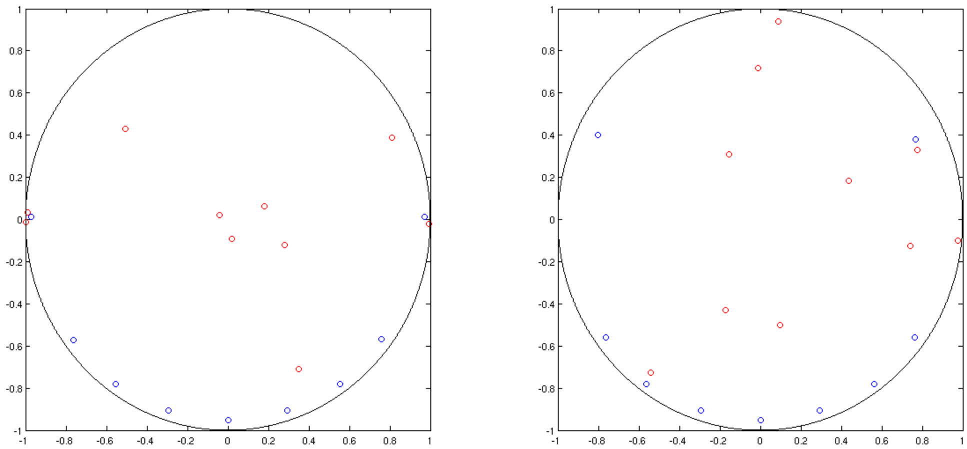


Figure 14: Initialization points (left) and optimized points (right) of the Schur function of degree 9 : parameters α (blue) and γ (red).

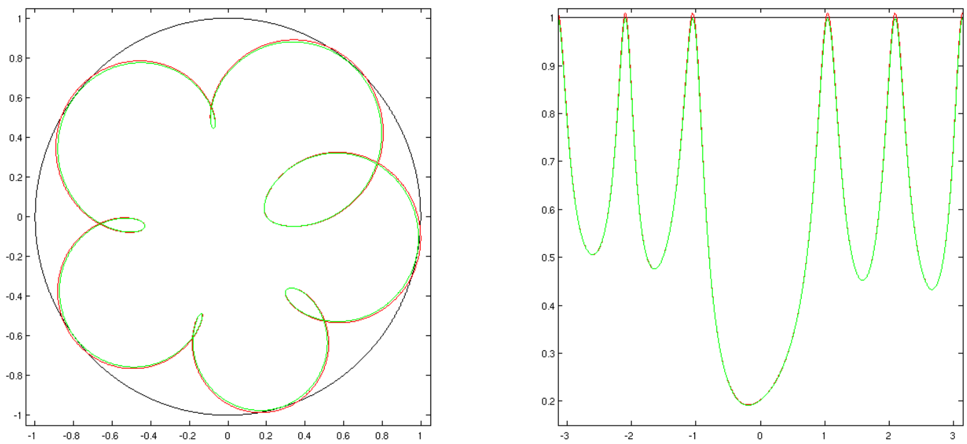


Figure 15: Function $r5$ (red) and Schur approximation (green) of degree 5.

As the asymptotic-BMO-type criterion can be applied only to Schur functions, the initialization was done upon applying it to the Schur function $r5/\|r5\|_\infty$.
Using our optimization process, we obtain an approximation of degree 5 with an error of $7.89e-3$. Scaling $r5$ into the unit disk (i.e. considering the function $\frac{r5}{\|r5\|_\infty}$) gives an error of $9.90e-3$.
Consider the initial and optimized parameters (see fig. 16). In this example, the interpolation points α given by the asymptotic-BMO-type criterion are surprisingly good.
Example 4

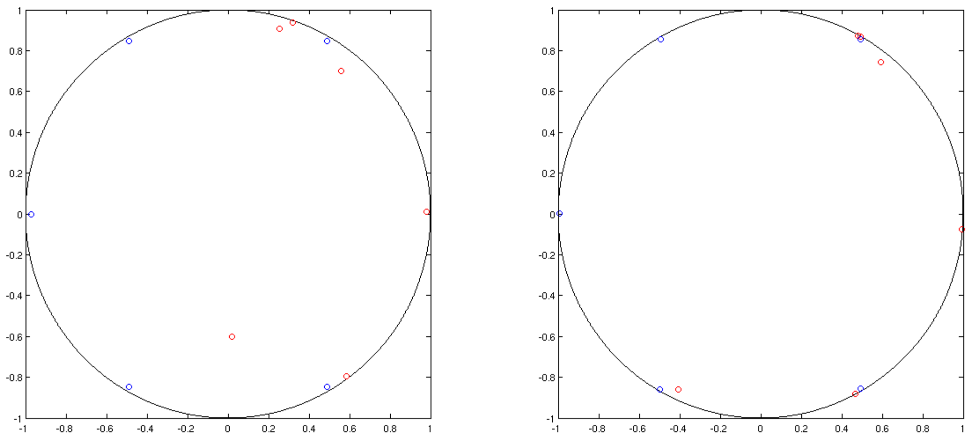


Figure 16: Initialization points (left) and optimized points (right) of the Schur function of degree 5 : parameters α (blue) and γ (red).

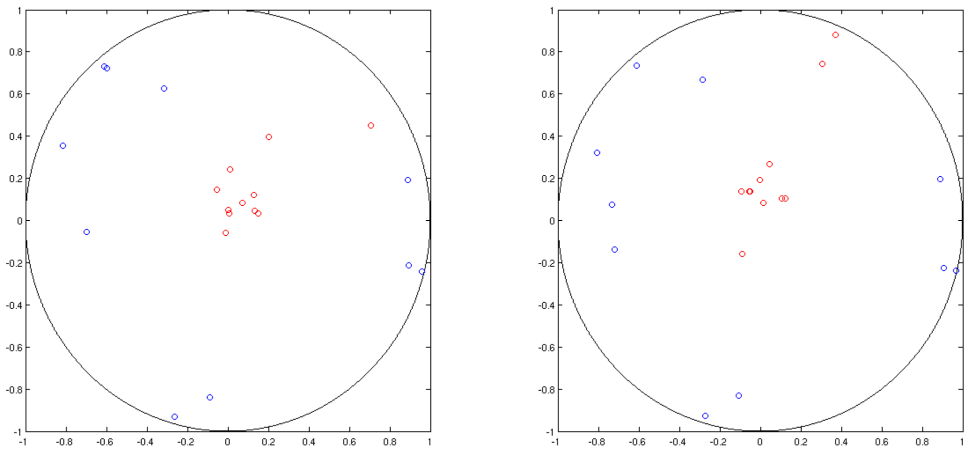


Figure 17: Initialization points (left) and optimized points (right) of the Schur function of degree 10 : parameters α (blue) and γ (red).

We want here to approximate a rational function $r10$ of degree 10, analytic in the unit disk, and such that $\|r10\|_\infty = 1.02$ and $\|r10\|_2 = 0.6772$. The asymptotic-BMO-type criterion applied to $r10/\|r10\|_\infty$ gives a sequence of points with one of multiplicity 3. As such an initialization could numerically leads to some problems, we chose to apply the asymptotic-BMO-type criterion to the strictly Schur function $\frac{r10}{1.05}$. The result is quite good : indeed, only one of the interpolation points α seems to have moved (see fig. 17).

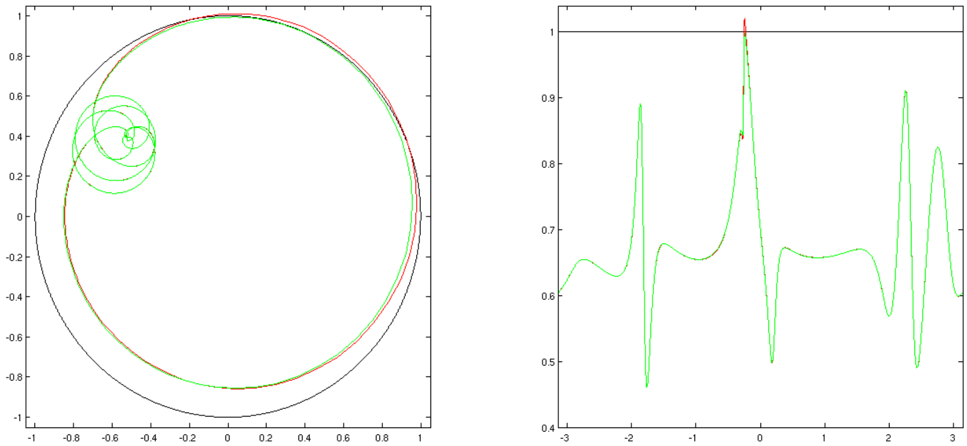


Figure 18: Function $r10$ (red) and Schur approximation (green) of degree 10.

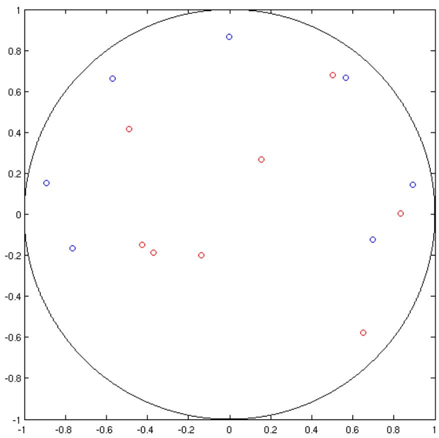


Figure 19: Another initialization for the approximation of degree 7 of $p30$: parameters α (blue) and γ (red).

The error of approximation is $2.58e-3$ (see fig. 18). Scaling $r10$ into the unit disk gives an error of $1.96e-2$.

On the last three examples, at least one initialization for a given degree seems to be quite good. However, all the initial interpolation points of the first example are bad. We chose to compute again an initialization but this time to the scaled strictly Schur function $0.97 \times p30$. This leads to the points plotted in fig. 19 for the degree 7. The interpolation points are "in the same directions" than the optimized points of the fig. 3.

8 Conclusion

In the previous section, we used a parametrization with Schur parameters of modulus strictly less than 1 only. Using this method, only *strictly* Schur rational functions could be represented. Finding a way to parametrize *all* Schur rational functions of given degree would be a great improvement. This is our attempt in this section. We will present an interpolation on the circle, and also another algorithm with Schur parameters strictly less than 1, but which has the advantage to have a limit when the parameters tend toward the circle. How to merge the two types of parametrization into a single one is an open problem as for now.

8.1 J-inner matrices and the Schur algorithm

This section is an introduction to the J-inner matrices and some of their properties.

Definition 8.1 *Let* $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. *A* 2×2 *matrix-valued funtion* θ *is called* J*-inner if it is meromorphic in* $\mathbb{D}$ *and*

- $\theta(z)J\theta(z)^* \leq J$ *at every point* z *of analyticity of* θ *in* $\mathbb{D}$*, and*
- $\theta(z)J\theta(z)^* = J$ *at almost every point* z *of* $\mathbb{T}$*.*

Many properties of J-inner matrices can be found in ([11]). A basic one is the following:

Proposition 8.2 *If* $\theta = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix}$ *is* 2×2 J*-inner and analytic in* $\mathbb{D}$ *and* g *is a Schur function, then* $(\theta_{21}g + \theta_{22})$ *is invertible in* $\mathbb{D}$*. Furthermore, if* $T_\theta(g)$ *is defined by*

$$T_\theta(g) = (\theta_{11}g + \theta_{12})(\theta_{21}g + \theta_{22})^{-1}$$

then $f = T_\theta(g)$ *is a Schur function.*

The result carries to higher sizes of θ but we will not need it.

Proof The proof can be found in different references, e.g. [11] for the matricial case. However, for a better understanding, we choose to give it again.
We first prove that $\theta_{21}g + \theta_{22}$ is invertible at any point of $\mathbb{D}$. As θ is J-inner, we have $\theta J\theta^* \leq J$ that is

$$\begin{pmatrix} |\theta_{11}|^2 - |\theta_{12}|^2 & \theta_{11}\overline{\theta_{21}} - \theta_{12}\overline{\theta_{22}} \\ \theta_{21}\overline{\theta_{11}} - \theta_{22}\overline{\theta_{12}} & |\theta_{21}|^2 - |\theta_{22}|^2 \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{in } \mathbb{D}.$$

This leads to $|\theta_{21}|^2 - |\theta_{22}|^2 \leq -1$, which is equivalent to $|\theta_{22}|^2 \geq 1 + |\theta_{21}|^2$. Therefore, θ_{22} is invertible at any point of $\mathbb{D}$. We thus have

$$1 - \left|\frac{\theta_{21}}{\theta_{22}}\right|^2 \geq \frac{1}{|\theta_{22}|^2} > 0$$

that is $\left|\frac{\theta_{21}}{\theta_{22}}\right|^2 < 1$ at any point of $\mathbb{D}$. We then deduce that $\theta_{21}g + \theta_{22} = \theta_{22}(\theta_{22}^{-1}\theta_{21}g + 1)$ is invertible at any point of $\mathbb{D}$.
We now prove that f is Schur. We have:

$$\begin{pmatrix} f \\ 1 \end{pmatrix} = \begin{pmatrix} \theta_{11}g + \theta_{12} \\ \theta_{21}g + \theta_{22} \end{pmatrix} (\theta_{21}g + \theta_{22})^{-1} = \theta \begin{pmatrix} g \\ 1 \end{pmatrix} (\theta_{21}g + \theta_{22})^{-1}$$

and

$$\begin{pmatrix} f \\ 1 \end{pmatrix}^* J \begin{pmatrix} f \\ 1 \end{pmatrix} = |f|^2 - 1.$$

Therefore,

$$\begin{aligned} |f|^2 - 1 &= (\theta_{21}g + \theta_{22})^{-*} \begin{pmatrix} \overline{g} & 1 \end{pmatrix} \theta^* J \theta \begin{pmatrix} g \\ 1 \end{pmatrix} (\theta_{21}g + \theta_{22})^{-1} \\ &\leq (\theta_{21}g + \theta_{22})^{-*}(|g|^2 - 1)(\theta_{21}g + \theta_{22})^{-1} \\ &\leq 0 \end{aligned}$$

and f is Schur.

■

Note that the multipoint Schur algorithm we used is such that

$$f = \frac{\zeta_1 f_1 + \gamma_0}{1 + \bar{\gamma}_0 \zeta_1 f_1}$$

that is $f = T_{\theta_1}(f_1)$ with

$$\theta_1(z) = \frac{1}{\sqrt{1 - |\gamma_0^2|}} \begin{pmatrix} \zeta_1(z) & \gamma_0 \\ \bar{\gamma}_0 \zeta_1(z) & 1 \end{pmatrix}. \tag{72}$$

It is easy to check that θ_1 is J-inner. Indeed,

$$\begin{aligned}
J-\theta_1(z)J\theta_1^*(z) &= J-\frac{1}{\sqrt{1-|\gamma_0^2|}}\begin{pmatrix}\zeta_1(z) & \gamma_0\\ \bar\gamma_0\zeta_1(z) & 1\end{pmatrix}J\frac{1}{\sqrt{1-|\gamma_0^2|}}\begin{pmatrix}\overline{\zeta_1(z)} & \gamma_0\overline{\zeta_1(z)}\\ \bar\gamma_0 & 1\end{pmatrix}\\
&= J-\frac{1}{1-|\gamma_0^2|}\begin{pmatrix}\zeta_1(z) & \gamma_0\\ \bar\gamma_0\zeta_1(z) & 1\end{pmatrix}\begin{pmatrix}\overline{\zeta_1(z)} & \gamma_0\overline{\zeta_1(z)}\\ -\bar\gamma_0 & -1\end{pmatrix}\\
&= \frac{1}{1-|\gamma_0^2|}\begin{pmatrix}1-|\zeta_1(z)|^2 & -\gamma_0(|\zeta_1(z)|^2-1)\\ -\bar\gamma_0(|\zeta_1(z)|^2-1) & |\gamma_0|^2(1-|\zeta_1(z)|^2)\end{pmatrix}\\
&= \frac{1-|\zeta_1(z)|^2}{1-|\gamma_0^2|}\begin{pmatrix}1 & \gamma_0\\ \bar\gamma_0 & |\gamma_0|^2\end{pmatrix}\\
&= \frac{1-|\zeta_1(z)|^2}{1-|\gamma_0^2|}\begin{pmatrix}1\\ \bar\gamma_0\end{pmatrix}(1\quad \gamma_0)\\
&\geq 0 \text{ for } z\in\mathbb{D} \text{ and } =0 \text{ for } z\in\mathbb{T}.
\end{aligned}$$

The Schur algorithm is based on the following result:

Let f be a Schur function. f satisfies the interpolation property $f(\alpha_1)=\gamma_0$ if and only if $f=T_{\theta_1}(f_1)$ for some Schur function f_1.

This result holds if we replace θ_1 by any J-inner function of the form $\theta_1 H$ where H is a constant matrix satisfying $H^*JH=J$ (such a matrix H is called J-unitary). This is a very particular case of the Nevanlinna-Pick interpolation problem studied for example in [11].

In section 8.3, another choice of J-inner matrix will be proposed.

8.2 Interpolation on the circle

The Schur algorithm studied in the previous section falls short of considering points on the unit circle. We now study an algorithm which manages such an interpolation. The following proposition shows a relation between the value of a Schur function at points of the unit circle, and the value of its angular derivative. The proof can be found in [5].

Proposition 8.3 *Let α_T and γ_T in $\mathbb{T}$. We denote by $f'(\alpha_T)$ the limit $\lim_{z\to\alpha_T} f'(z)$ where z converges to α_T nontangentially. If f is a Schur function such that $f(\alpha_T)=\gamma_T$, then $f'(\alpha_T)=\rho\bar\alpha_T\gamma_T$ where ρ is a positive real constant.*

We now define a J-inner matrix which leads to an interpolation scheme on the circle.

Proposition 8.4 *Let α_T and γ_T be points of the unit circle, ρ be a positive real constant, and x_T be the vector such that $x_T^t=(1\quad \bar\gamma_T)$. Then, the matrix θ_2 defined by*

$$\theta_2(z)=I_2+\frac{1}{2\rho}\frac{z+\alpha_T}{z-\alpha_T}x_Tx_T^*J$$

is J-inner.

Proof We have

$$\begin{aligned} J-\theta_2(z)J\theta_2(z)^* &= J-\left(I_2+\frac{1}{2\rho}\frac{z+\alpha_T}{z-\alpha_T}x_Tx_T^*J\right)J\left(I_2+\frac{1}{2\rho}\overline{\left(\frac{z+\alpha_T}{z-\alpha_T}\right)}Jx_Tx_T^*\right) \\ &= -\frac{1}{2\rho}\frac{z+\alpha_T}{z-\alpha_T}x_Tx_T^*-\frac{1}{2\rho}\overline{\left(\frac{z+\alpha_T}{z-\alpha_T}\right)}x_Tx_T^* \\ &\quad -\frac{1}{(2\rho)^2}\left|\frac{z+\alpha_T}{z-\alpha_T}\right|^2(1-|\gamma_T|^2)x_Tx_T^*. \end{aligned}$$

As $|\gamma_T|=1$, we get

$$J-\theta_2(z)J\theta_2(z)^*=-\frac{1}{2\rho}\left[\frac{z+\alpha_T}{z-\alpha_T}+\overline{\left(\frac{z+\alpha_T}{z-\alpha_T}\right)}\right]x_Tx_T^*=-\frac{1}{\rho}Re\left(\frac{z+\alpha_T}{z-\alpha_T}\right)x_Tx_T^*.$$

But $Re\left(\frac{z+\alpha_T}{z-\alpha_T}\right)=Re\left(\frac{|z|^2+\alpha_T\bar{z}-\bar{\alpha}_Tz-|\alpha_T|^2}{|z-\alpha_T|^2}\right)=\frac{|z|^2-|\alpha_T|^2}{|z-\alpha_T|^2}\leq 0$ for all $z\in\mathbb{D}$, and consequently, $J-\theta_2(z)J\theta_2(z)^*\geq 0$.

■

Proposition 8.5 *If g is a Schur function such that $g(\alpha_T)\neq\gamma_T$ then $f=T_{\theta_2}(g)$ is a Schur function such that $f(\alpha_T)=\gamma_T$ and $f'(\alpha_T)=\rho\overline{\alpha_T}\gamma_T$.*

Proof We have

$$\theta_2(z)=\begin{pmatrix} 1+\frac{1}{2\rho}\frac{z+\alpha_T}{z-\alpha_T} & -\frac{\gamma_T}{2\rho}\frac{z+\alpha_T}{z-\alpha_T} \\ \frac{\bar{\gamma}_T}{2\rho}\frac{z+\alpha_T}{z-\alpha_T} & 1-\frac{1}{2\rho}\frac{z+\alpha_T}{z-\alpha_T} \end{pmatrix}$$

so that

$$f(z)=\frac{(2\rho(z-\alpha_T)+(z+\alpha_T))g(z)-\gamma_T(z+\alpha_T)}{\bar{\gamma}_T(z+\alpha_T)g(z)+2\rho(z-\alpha_T)-(z+\alpha_T)}.$$

Therefore

$$f(\alpha_T)=\frac{2\alpha_T(g(\alpha_T)-\gamma_T)}{2\alpha_T(\bar{\gamma}_Tg(\alpha_T)-1)}=\gamma_T$$

because $g(\alpha_T)\neq\gamma_T$.
A direct computation gives

$$\begin{aligned} f'(\alpha_T) &= \frac{((2\rho+1)g(\alpha_T)+2\alpha_Tg'(\alpha_T)-\gamma_T)}{2\alpha_T(\bar{\gamma}_Tg(\alpha_T)-1)} \\ &\quad -\frac{(\bar{\gamma}_Tg(\alpha_T)+2\alpha_T\bar{\gamma}_Tg'(\alpha_T)+2\rho-1)(2\alpha_T(g(\alpha_T)-\gamma_T)}{(2\alpha_T(\bar{\gamma}_Tg(\alpha_T)-1))^2} \\ &= \frac{2\rho(g(\alpha_T)-\gamma_T)}{2\alpha_T(\bar{\gamma}_Tg(\alpha_T)-1)} \\ &= \rho\bar{\alpha}_T\gamma_T. \end{aligned}$$

■

Note that if $f = p/q$, an interpolation point in the circle is always a root of $p\tilde{p} - q\tilde{q}$. We will now show that if we apply the algorithm associated to θ_2 to a Schur rational function p/q of degree n such that $p/q(\alpha_T) = \gamma_T$ and $(p/q)'(\alpha_T) = \rho\overline{\alpha_T}\gamma_T$, then $g = T_{\theta_2^{-1}}\left(\frac{p}{q}\right)$ is a Schur rational function of degree $n-1$. Indeed,

$$\begin{aligned} g &= \frac{2\rho(z-\alpha_T)\frac{p}{q} - (z+\alpha_T)\left(\frac{p}{q} - \gamma_T\right)}{2\rho(z-\alpha_T) + (z+\alpha_T)\left(1 - \bar{\gamma}_T\frac{p}{q}\right)} \\ &= \frac{2\rho p - (z+\alpha_T)\frac{p-\gamma_T q}{z-\alpha_T}}{2\rho q - (z+\alpha_T)\bar{\gamma}_T\frac{p-\gamma_T q}{z-\alpha_T}}. \end{aligned}$$

But evaluating the numerator and denominator of g at α_T gives

$$2\rho p(\alpha_T) - 2\alpha_T q(\alpha_T) f'(\alpha_T) = 2\rho\gamma_T q(\alpha_T) - 2\alpha_T q(\alpha_T)\rho\bar{\alpha}_T\gamma_T = 0$$

and

$$2\rho q(\alpha_T) - 2\alpha_T\bar{\gamma}_T q(\alpha_T) f'(\alpha_T) = 2\rho q(\alpha_T) - 2\alpha_T\bar{\gamma}_T q(\alpha_T)\rho\bar{\alpha}_T\gamma_T = 0.$$

Therefore, the degree of g is at most $n-1$. Applying the linear transform T_{θ_2} to g increases the degree of at most one. Thus, the degree of g is exactly $n-1$.

8.3 A better algorithm ?

We are now going to study another parametrization whose advantage is to have a limit when points tend towards the circle. The link with the previous Schur algorithm is given.

8.3.1 Another algorithm

Proposition 8.6 *Let α and γ be points of the unit disk $\mathbb{D}$, and x be the vector $(1 \quad \bar{\gamma})^t$. Then, the matrix θ_3 defined by*

$$\theta_3(z) = I_2 + \frac{\zeta_\alpha(z) - 1}{1 - |\gamma|^2} x x^* J \tag{73}$$

is J-inner.

Proof We have

$$\begin{aligned}
J - \theta_3(z)J\theta_3^*(z) &= J - \left(I_2 + \frac{\zeta_\alpha(z)-1}{1-|\gamma|^2}xx^*J\right) J \left(I_2 + Jxx^*\frac{\overline{\zeta_\alpha(z)}-1}{1-|\gamma|^2}\right) \\
&= -\frac{\zeta_\alpha(z)-1}{1-|\gamma^2|}xx^* - xx^*\frac{\overline{\zeta_\alpha(z)}-1}{1-|\gamma|^2} - \frac{|\zeta_\alpha(z)-1|^2}{1-|\gamma^2|}xx^* \\
&= -\left(|\zeta_\alpha(z)-1|^2 + \zeta_\alpha(z) - 1 + \overline{\zeta_\alpha(z)} - 1\right)\frac{xx^*}{1-|\gamma|^2} \\
&= -((\zeta_\alpha(z)-1)(\overline{\zeta_\alpha(z)}-1) + \zeta_\alpha(z) - 1 + \overline{\zeta_\alpha(z)} - 1)\frac{xx^*}{1-|\gamma|^2} \\
&= \frac{1-|\zeta_\alpha(z)|^2}{1-|\gamma|^2}xx^* \\
&\geq 0 \text{ for all } z \in \mathbb{D}.
\end{aligned}$$

■

Proposition 8.7 *Let g be a Schur function. Then $f = T_{\theta_3}(g)$ is a Schur function such that $f(\alpha) = \gamma$.*

Proof We have

$$x^*J\theta_3(\alpha) = x^*J - \frac{1}{1-|\gamma|^2}x^*Jxx^*J$$

and $x^*Jx = 1 - |\gamma|^2$, therefore $x^*J\theta_3(\alpha) = 0$. Thus,

$$x^*J\begin{pmatrix} f(\alpha) \\ 1 \end{pmatrix} = x^*J\theta_3(\alpha)\begin{pmatrix} g(\alpha) \\ 1 \end{pmatrix}((\theta_3)_{21}(\alpha)g(\alpha) + (\theta_3)_{22}(\alpha))^{-1} = 0$$

and we get $f(\alpha) = \gamma$.

■

8.3.2 Relation between the two algorithms

We now show that the J-inner matrix of the "new" algorithm is in fact the J-inner matrix of the previous algorithm multiplied by a constant matrix H.
The proof of the following lemma is immediate.

Lemma 8.8 *Let γ in $\mathbb{D}$ and*

$$H(\gamma) = \frac{1}{\sqrt{1-|\gamma|^2}}\begin{pmatrix} 1 & \gamma \\ \bar{\gamma} & 1 \end{pmatrix}.$$

The matrix $H(\gamma)$ has the following properties:

- *$H(\gamma)$ is J-unitary, i.e. $H(\gamma)JH(\gamma)^* = J$,*
- *$H(\gamma)^{-1} = H(-\gamma)$.*

We now give another expression of the J-inner matrix associated to the "new" algorithm ([16]).

Proposition 8.9 *The matrix θ_3 defined by (73) is of the form*

$$\theta_3(z) = H(\gamma) \begin{pmatrix} \zeta_\alpha & 0 \\ 0 & 1 \end{pmatrix} H(\gamma)^{-1}.$$

Proof We have

$$\begin{aligned} H(-\gamma)\theta_3(z)H(\gamma) &= H(-\gamma)\left(I_2 + \frac{\zeta_\alpha(z)-1}{1-|\gamma|^2}xx^*J\right)H(\gamma) \\ &= I_2 + \frac{\zeta_\alpha(z)-1}{1-|\gamma|^2}H(-\gamma)xx^*JH(\gamma) \\ &= I_2 + \frac{\zeta_\alpha(z)-1}{1-|\gamma|^2}\sqrt{1-|\gamma|^2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\sqrt{1-|\gamma|^2}\begin{pmatrix} 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \zeta_\alpha & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

■

Note that the matrix θ_1 defined by (72) is of the form

$$\theta_1 = H(\gamma)\begin{pmatrix} \zeta_\alpha & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore, the link between the matrix θ_3 and θ_1 is given by

$$\theta_3 = \theta_1 H(-\gamma).$$

8.3.3 Toward a parametrization of all Schur rational functions

We now show that when the point α tends to a point α_T of the unit circle, θ_3 tends to θ_2 ([17]). We have

$$\begin{aligned} \frac{\zeta_\alpha(z)-1}{1-|f(\alpha)|^2} &= \frac{-\frac{|\alpha|}{\alpha}\frac{z-\alpha}{1-\bar{\alpha}z} - 1}{1-|f(\alpha)|^2} \\ &= \frac{\frac{-|\alpha|(z-\alpha)-(\alpha-|\alpha|^2 z)}{\alpha-|\alpha|^2 z}}{1-f(\alpha)\overline{f(\alpha)}} \\ &= \frac{(|\alpha|-1)\frac{\alpha+|\alpha|z}{\alpha(1-\bar{\alpha}z)}}{1-f(\alpha)\overline{f(\alpha)}}. \end{aligned}$$

Using a Taylor expansion, we get

$$f(\alpha) = f(\alpha_T) + (\alpha - \alpha_T)f'(\alpha_T) + o(|\alpha - \alpha_T|).$$

Therefore,

$$\begin{aligned} 1 - f(\alpha)\overline{f(\alpha)} &= -2Re\left[(\alpha - \alpha_T)\overline{f(\alpha_T)}f'(\alpha_T)\right] + o(|\alpha - \alpha_T|) \\ &= -2Re\left[(\alpha - \alpha_T)\bar{\gamma}_T\rho\bar{\alpha}_T\gamma_T\right] + o(|\alpha - \alpha_T|) \\ &= -2Re\left[\rho(\alpha\overline{\alpha_T} - 1)\right] + o(|\alpha - \alpha_T|) \end{aligned}$$

and we get

$$\frac{\zeta_\alpha(z) - 1}{1 - |f(\alpha)|^2} = \frac{(|\alpha| - 1)\frac{z + |\alpha|\alpha}{(|\alpha|^2 z - \alpha)}}{2Re\left[\rho(\alpha\overline{\alpha_T} - 1)\right] + o(|\alpha - \alpha_T|)}.$$

It remains to check that $\frac{|\alpha|-1}{2Re(\alpha\overline{\alpha_T}-1)}$ tends toward $\frac{1}{2}$. Let η be a complex number such that $\alpha = \alpha_T + \eta$. Then

$$|\alpha|^2 = |\alpha_T|^2 + 2Re(\eta\bar{\alpha}_T) + |\eta|^2 = 1 + 2Re(\eta\bar{\alpha}_T) + |\eta|^2$$

and we deduce that

$$|\alpha| = 1 + Re(\eta\bar{\alpha}_T) + o(\eta).$$

Thus $|\alpha| - 1 = Re(\eta\bar{\alpha}_T) + o(\eta)$. As $2Re\,(\alpha\bar{\alpha_T} - 1) = 2Re(\eta\bar{\alpha}_T)$, the conclusion is immediate.

As stated before, only *strictly* Schur rational functions can be represented using the parametrization of the previous section. From what precedes, we see that the algorithm associated to θ_3 could be combined with interpolation on the unit circle, and therefore, parameters could be taken in the closed unit disk $\overline{\mathbb{D}}$. This could be a great improvment. However, new questions arise: could this algorithm be related to orthogonal rational functions ? And in practice, when do you choose to take interpolation points on the circle and how could one compute the parameter ρ?

References

[1] N. I. Achieser. *Theory of approximation.* Dover Publications Inc., New York, 1992. Translated from the Russian and with a preface by Charles J. Hyman, Reprint of the 1956 English translation.

[2] Daniel Alpay. *The Schur algorithm, reproducing kernel spaces and system theory*, volume 5 of *SMF/AMS Texts and Monographs.* American Mathematical Society, Providence, RI, 2001. Translated from the 1998 French original by Stephen S. Wilson.

[3] Daniel Alpay, Laurent Baratchart, and Andrea Gombani. On the differential structure of matrix-valued rational inner functions. In *Nonselfadjoint operators and related topics (Beer Sheva, 1992)*, volume 73 of *Oper. Theory Adv. Appl.*, pages 30–66. Birkhäuser, Basel, 1994.

[4] M. Bakonyi and T. Constantinescu. *Schur's algorithm and several applications*, volume 261 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow, 1992.

[5] Joseph A. Ball, Israel Gohberg, and Leiba Rodman. *Interpolation of rational matrix functions*, volume 45 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1990.

[6] Adhemar Bultheel, Pablo González-Vera, Erik Hendriksen, and Olav Njåstad. *Orthogonal rational functions*, volume 5 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 1999.

[7] Adhemar Bultheel, Pablo González-Vera, Erik Hendriksen, and Olav Njåstad. Orthogonal rational functions on the unit circle: from the scalar to the matrix case. In *Orthogonal polynomials and special functions*, volume 1883 of *Lecture Notes in Math.*, pages 187–228. Springer, Berlin, 2006.

[8] Zoia Ceauşescu and Ciprian Foiaş. On intertwining dilations. V. *Acta Sci. Math. (Szeged)*, 40(1-2):9–32, 1978.

[9] M. M. Djrbashian. Orthogonal systems of rational functions on the unit circle with given set of poles. *Soviet Mathematics Doklady*, 3:1794–1798, 1962.

[10] Peter L. Duren. *Theory of H^p spaces*. Pure and Applied Mathematics, Vol. 38. Academic Press, New York, 1970.

[11] Harry Dym. *J contractive matrix functions, reproducing kernel Hilbert spaces and interpolation*, volume 71 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1989.

[12] John B. Garnett. *Bounded analytic functions*, volume 236 of *Graduate Texts in Mathematics*. Springer, New York, first edition, 2007.

[13] J. Geronimus. On polynomials orthogonal on the circle, on trigonometric moment-problem and on allied Carathéodory and Schur functions. *Rec. Math. [Mat. Sbornik] N. S.*, 15(57):99–130, 1944.

[14] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Invariant subspaces of matrices with applications*, volume 51 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. Reprint of the 1986 original.

[15] J. Grimm. Rational approximation of transfer functions in the hyperion software. *Rapport de Recherche INRIA No. 4002, Sept. 2000*, 2000.

[16] Bernard Hanzon, Martine Olivi, and Ralf L. M. Peeters. Balanced realizations of discrete-time stable all-pass systems and the tangential Schur algorithm. *Linear Algebra Appl.*, 418(2-3):793–820, 2006.

[17] Bernard Hanzon, Martine Olivi, and Ralf L. M. Peeters. Boundary interpolation and parametrization of lossless functions. *Rapport de Recherche INRIA*, 2008.

[18] William B. Jones. Schur's algorithm extended and Schur continued fractions. In *Nonlinear numerical methods and rational approximation (Wilrijk, 1987)*, volume 43 of *Math. Appl.*, pages 281–298. Reidel, Dordrecht, 1988.

[19] Thomas Kailath. *Linear systems.* Prentice-Hall Inc., Englewood Cliffs, N.J., 1980. Prentice-Hall Information and System Sciences Series.

[20] Sergei Khrushchev. Schur's algorithm, orthogonal polynomials, and convergence of Wall's continued fractions in $L^2(\mathbb{T})$. *J. Approx. Theory*, 108(2):161–248, 2001.

[21] Heinz Langer and Andreas Lasarow. Solution of a multiple Nevanlinna-Pick problem via orthogonal rational functions. *J. Math. Anal. Appl.*, 293(2):605–632, 2004.

[22] E. M. Nikishin and V. N. Sorokin. *Rational approximations and orthogonality*, volume 92 of *Translations of Mathematical Monographs.* American Mathematical Society, Providence, RI, 1991. Translated from the Russian by Ralph P. Boas.

[23] K. Pan. On the convergence of rational functions orthogonal on the unit circle. *J. Comput. Appl. Math.*, 76(1-2):315–324, 1996.

[24] V. P. Potapov. The multiplicative structure of J-contractive matrix functions. *Trudy Moskov. Mat. Obšč.*, 4:125–236, 1955.

[25] Walter Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

[26] I. Schur. Uber potenzreihen, die im innern des einheitskreises beschrankt sind, i. j. reine angew. math., 147:205–232, 1917. *English translation in: I. Schur methods in operator theory and signal processing (Operator theory: advances and applications OT 18 (1986), Birkh auser Verlag)*, 1917.

[27] Laurent Schwartz. Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants). *J. Analyse Math.*, 13:115–256, 1964.

[28] Gábor Szegő. *Orthogonal polynomials.* American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

[29] H. S. Wall. *Analytic Theory of Continued Fractions.* D. Van Nostrand Company, Inc., New York, N. Y., 1948.

Centre de recherche INRIA Sophia Antipolis – Méditerranée
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399