# Weighted $H^2$ approximation of transfer functions

Juliette Leblond[*], Martine Olivi[*]

**Abstract:** the aim of this work is to generalize to the weighted case some results and algorithms concerning $L^2$ approximation by analytic and rational functions which are useful to perform the identification of unknown transfer functions of a class of stable (linear causal time–invariant) systems from incomplete frequency data.

**Keywords:** frequency domain identification, weighted Hardy spaces, extremal problems, rational approximation, orthogonal polynomials.

## 1 Introduction

The approximation issues approached in this paper are mainly motivated by a frequency domain identification problem related to a class of linear time-invariant causal single-input/single-output systems in discrete time that possess the stability property of having finite weighted input–output $l^2 \to l^\infty$ (or equivalently $l^2 \to l^\infty$) gain, for a rather general class of weights.

Assume that we are given some of the (possibly noisy) pointwise values of the transfer function of such a system, measured at frequencies belonging to a subset of the unit circle $\mathbb{T}$ that corresponds to its bandwidth. Such measurements may be obtained using harmonic identification procedures. Some rough information concerning the behavior of the system outside the bandwidth may be available. In order to identify the unknown system, we want to find a rational stable function of bounded Mac–Millan degree accounting well enough for these data.

Although the experimental data are discrete values, the stability constraint on the model we are looking for cannot be guaranteed by a discrete least–square criterion as the degree increases, while this is not even a convenient framework to approach convergence and robustness issues.

As in the worst case identification algorithms [GK, HKN, Par], we perform a preliminary interpolation step which consists in getting a robust non–causal interpolant accounting for the given experimental data. Our identification problem can then be approached by two consecutive stages consisting in solving:

- a bounded extremal problem which furnishes the transfer function of an infinite dimensional stable causal model for the system (analytic approximation step),
- a rational approximation problem that provides a transfer function of bounded Mac–Millan degree (model order reduction step).

Both steps are handled here by minimizing a weighted integral quadratic criterion on a Hardy space consisting of transfer functions that possess the above described stability property, or among its rational functions of bounded degree.

[*]INRIA, BP 93, 06902 Sophia–Antipolis Cedex, FRANCE, {leblond}{olivi}@sophia.inria.fr, phone: 33 4 92 38 78 76, fax: 33 4 92 38 78 58

The case of white noise inputs corresponds to the unweighted issues where the cost is the classical $L^2$ criterion for the Lebesgue measure on $\mathbb{T}$. In this situation, the two approximation problems have already been studied and resolution algorithms provided in [ABL, BL] and [BCO, BOW].
Our aim here is to generalize these results to more general inputs or criterion, namely when the cost is induced by the norm

$$\|f\|_\mu = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(e^{i\theta})|^2 d\mu(\theta),$$

for a finite positive measure $\mu$ on $\mathbb{T}$, the quantity $d\mu/d\lambda$ being the spectral density of the noisy input [R] in the classical stochastic framework; this type of weighted $L^2$ approximation problems comes up when minimizing the variance of the output error between the searched model and the "true system". In the literature, such a weighting is also present in criteria induced by either quadratic, uniform, or operator norm, or else in model reduction via frequency balanced realizations for various applications [E, LA]. It arises in different forms of the standard problem of robust control [BGR, F]. For example, when one pursues an identification procedure with the purpose of designing a controller, then the weight represents the control performance specifications [E, Ge]. Moreover, such a criterion is commonly used to weight some frequencies more than the others in order to get through the frequency dependence of the model reduction error or to represent the confidence one has in the available measurements for either identification, filtering or control issues (it is interesting to give some importance to the bandwidth on which we initially got the data).
In section 2, we state our approximation problems and characterize the considered weights. Then, in sections 3 and 4, the weighted analytic and rational approximation problems are studied and resolution algorithms are given.

## 2 Statement of the problems

Since impulse responses are real–valued signals, the associated transfer functions possess the *conjugate–symmetry property* $f(\bar{z}) = \overline{f(z)}$. For this reason we consider here *real* Banach spaces of conjugate–symmetric functions.
Let $\mu$ be any positive finite measure on the unit circle $\mathbb{T}$ satisfying $\mu(\Gamma) = \mu(\bar{\Gamma})$ for any $\Gamma \subset \mathbb{T}$ and let $L^2(\mu)$ be the *real* Hilbert space of functions on $\mathbb{T}$ that are square–summable w.r.t. $\mu$ and satisfy the conjugate–symmetry property (such functions possess real Fourier coefficients); $L^2(\mu)$ is endowed with the inner product defined by:

$$< f\,,\, g >_\mu = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})\, g(e^{-i\theta})\, d\mu(\theta)\,, \tag{1}$$

and with the associated norm $\|\ \|_\mu$.
Define the *real* weighted Hardy spaces $H^2(\mu)$ and $\bar{H}_0^2(\mu)$ as the $L^2(\mu)$ closures of the families $\{z^k,\ k \geq 0\}$ and $\{1/z^k,\ k > 0\}$, respectively. We may identify these functions with their traces on $\mathbb{T}$. Proofs or details about the considerations of this section can be found in [CS2, Ga, H]. The two spaces $H^2(\mu)$ and $\bar{H}_0^2(\mu)$ are isometric under the map defined on $L^2(\mu)$ by:

$$f(z) \quad \mapsto \quad \frac{f(1/z)}{z} = \check{f}(z)\,. \tag{2}$$

Whenever $\mu = \lambda$, the Lebesgue measure, we write for the sake of simplicity: $L^2(\lambda) = L^2(\mathbb{T})$, together with $<\,,\,>_\lambda = <\,,\,>$ and $\|\ \|_\lambda = \|\ \|$, for its inner product and associated norm. Then, $H^2(\lambda) = H^2$ is the classical *real* Hardy space of the unit disk $\mathbb{D}$; it coincides with the closed subspace of $L^2(\mathbb{T})$ of

functions whose Fourier coefficients of negative index are zero. The orthogonal complement of $H^2$ in $L^2(\mathbb{T})$ w.r.t. $\lambda$ is the space $\bar{H}_0^2(\lambda) = \bar{H}_0^2$ of stable transfer functions consisting in $L^2(\mathbb{T})$ functions which possess Fourier coefficients of non–negative index equal to zero (or, equivalently, in functions analytic outside the closed unit disk, vanishing at infinity, and bounded in $L^2$–norm on circles of radius $r > 1$). Let $L^\infty(\mathbb{T})$ be the *real* Banach space of essentially bounded conjugate–symmetric functions and $H^\infty = H^2 \cap L^\infty(\mathbb{T})$ [Ga, II.4].
Concerning weighted Hardy spaces, it first follows from a result due to Szegö [H, ch.4] that $\mu$ is absolutely continuous w.r.t. $\lambda$ and such that $H^2(\mu) \neq L^2(\mu)$ if and only if

$$d\mu(\theta) = |\nu(e^{i\theta})|^2 \, d\theta \tag{3}$$

for some outer $H^2$ function $\nu$; see [CS1, CS2], where Hardy spaces are studied for such Szegö weights. We in fact settle here to the stronger case where:

$$L^2(\mathbb{T}) = L^2(\mu)\,, \tag{4}$$

which is equivalent to $\bar{H}_0^2 = \bar{H}_0^2(\mu)$ or $H^2 = H^2(\mu)$ and provides the simplest assumption for our weighted approximation problems to be well-posed. It turns out [LO, thm.1] that (4) is satisfied if and only if (3) holds for some $\nu \in H^\infty$ invertible in $H^\infty$.

*(H) In the remainder of this paper, we make the standing assumption that $\mu$ satisfies* (3) *for a function $\nu$ belonging to $H^\infty$ and invertible in $H^\infty$.*

In this case,

$$< f \, , \, g >_\mu = \frac{1}{2i\pi} \int_{\mathbb{T}} \check{f}(z) \, g(z) \, |\nu(z)|^2 \, dz \,, \tag{5}$$

and, for any symmetric subset $\Gamma$ of $\mathbb{T}$ (i.e. such that $\bar{\Gamma} = \Gamma$), if $\chi_\Gamma$ denotes its characteristic function, then $L^2(\Gamma)$ can be endowed with the $L^2(\mu)$ norm: $\|f\|_{\Gamma,\mu} = \|\chi_\Gamma \, f\|_\mu$.

We now precisely state in this framework[1] the two approximation problems linked with the two stages of our identification scheme. Assume that the bandwidth on which measurements have been performed is a symmetric subset $K$ of $\mathbb{T}$ (if $K$ is not already symmetric, consider $K \cup \bar{K}$ and conjugate data on $\bar{K}$). Recall that a preliminary interpolation procedure might have been performed, which is supposed to furnish two functions $\varphi \in L^2(K)$ and $h \in L^2(J)$ reflecting both available measurements on $K$ and further information outside $K$ (if nothing is known there, take $h = 0$). Trigonometric polynomials (Jackson, de La Vallée Poussin) provide robust interpolants for pointwise data [MPG, T]. Using the (stable–unstable) transformation (2), the first step of our procedure can be approached in $H^2$ where it amounts to solve:

*$(P_1)$ Given $\varphi \in L^2(K)$, $h \in L^2(J)$, and $M > 0$, find a function $g_0 \in H^2$ which minimizes $\|\varphi - g\|_{K,\mu}$ among the functions $g \in H^2$ which satisfy the constraint $\|h - g\|_{J,\mu} \leq M$.*

Now, the second step of our identification procedure applies to $\check{g}_0$ but we set it up as an approximation issue for an arbitrary $\bar{H}_0^2$ function:

[1] Hardy spaces of the unit disk are appropriate to describe a discrete time system behavior but continuous time systems can also be handled that way using a Möbius transform.

$(P_2)$ *Given $f \in \bar{H}_0^2$ and an integer $n > 0$, find a rational function $r_0$ which minimizes*

$$\|f - r\|_{\mu}, \tag{6}$$

*where $r$ ranges over the rational functions in $\bar{H}_0^2$ of Mac–Millan degree at most $n$.*

# 3 Weighted analytic approximation

In this section, we explain how to get a solution to problem $(P_1)$. For $\mu = \lambda$, it has been solved when $h = 0$ [ABL] and when $\varphi = 0$ [KN]. Since then, it has been approached in the general $H^p$ setting, $1 \leq p < \infty$ [BL] and in $H^\infty$ [BLP]. Existence and a characterization of a solution to problem $(P_1)$ for measures $\mu$ satisfying $(H)$ can be deduced from these results. Let $h \in L^2(J)$, $M > 0$, and define

$$C_M^h(\mu) = \{g_{|_K}\ ,\ g \in H^2\ ,\ \|h - g\|_{J,\mu} \leq M\}\ ,$$

where $g_{|_K}$ denotes the restriction of $g$ to $K$. Denote by $P_{H^2}$ the $\lambda$–orthogonal projection from $L^2(\mathbb{T})$ onto $H^2$ and let $T$ be the Toeplitz operator with symbol $\chi_J$:

$$T(g) = P_{H^2}(\chi_J\, g)\ ,\ \forall g \in H^2\, .$$

**Theorem 1** *Let $K$ be a symmetric subset of $\mathbb{T}$ such that both $K$ and its complementary subset $J$ are of positive $\mu$ measure, where $\mu$ satisfies $(H)$. Then, there exists a unique solution $g_0 \in H^2$ to problem $(P_1)$. Moreover, $\|h - g_0\|_{J,\mu} = M$ whenever $\varphi \notin C_M^h(\mu)$ and $g_0$ is given in this case by the implicit equation:*

$$g_0 = \nu^{-1}\,(1 + l\,T)^{-1}\,P_{H^2}\ (\nu(\chi_K\,\varphi + (l+1)\chi_J\,h))\ , \tag{7}$$

*where $l \in (-1, +\infty)$ is the unique number such that $\|h - g_0\|_{J,\mu} = M$.*

Theorem 1 follows from [BL, thm.2,4] upon multiplication by the $H^\infty$ functions $\nu$ or $\nu^{-1}$. Without a norm constraint on $g_0$ outside $K$, problem $(P_1)$ becomes ill–posed unless $\varphi$ is already the trace on $K$ of an $H^2$ function; in this case, $(P_1)$ can be interpreted when $M \to \infty$ as a recovery issue of the $H^2$ function $\varphi$ from its values on $K$, [BL, prop.3], [BLP, prop.1].
In order to compute $g_0$, we have to get through the implicit character in $M$ of equation (7). To this end, if $\varphi \notin H^2_{|K}$, it can be shown as in [BL, prop.4] that $M$ is a smoothly decreasing function of the Lagrange parameter $l$ from $(-1, \infty)$ onto $(0, \infty)$. Hence, $M$ being given, $g_0$ can be numerically computed using a dichotomy procedure on $l$. Furthermore, as $l \to -1$, the error $e_\mu = \|\varphi - g_0\|_{K,\mu}$ goes to zero while $M \to \infty$. Another characterization of $g_0$ by a Carleman formula [A, Pat] can also be obtained from [BL, cor.1].

# 4 Weighted rational approximation

## 4.1 A criterion depending on the denominators of the approximants.

We first establish a normality result which generalizes [BOW, prop.2.1]. Observe that a rational function $p/q$ belongs to $\bar{H}_0^2$ if and only if $p/q$ is *stable* ($q$ has its roots inside the unit disk $\mathbb{D}$) and *strictly proper* (vanishes at infinity). The Mac–Millan degree of such a rational function is thus the degree of $q$.

**Proposition 1** *If $f \in \bar{H}_0^2$ is not a rational function of degree less than $n$, then the argument of any local minimum of* (6) *is an irreducible fraction whose degree is equal to $n$.*

*Proof:* assume that $p_0/q_0$ is a local minimum of (6) for which $p_0$ and $q_0$ are coprime polynomials and $\deg q_0 < n$. For $a$ small enough and $b$ such that $|b| < 1$, we have that

$$\left\| f - \left(\frac{p_0}{q_0} + \frac{a}{z-b}\right)\right\|_\mu^2 \geq \left\| f - \frac{p_0}{q_0}\right\|_\mu^2 ,$$

or equivalently,

$$a^2 \left\| \frac{1}{z-b} \right\|_\mu^2 - 2\,a \left\langle f - \frac{p_0}{q_0}, \frac{1}{z-b} \right\rangle_\mu \geq 0\,.$$

This holds if and only if $< f - \frac{p_0}{q_0}, \frac{1}{z-b} >_\mu = 0$, and since the family $\{1/(z-b)\,,\ |b|<1\}$ is dense in $H^2(\mu)$, we must have $f = p_0/q_0$. This contradicts the assumption on $f$. □

*We assume in the following that $f$ is not rational of degree less than $n$.* Since by proposition 1, a solution $p/q$ to problem $(P_2)$ over the set of $\bar{H}_0^2$ rational functions of Mac–Millan has exact degree $n$, *we shall therefore assume $q$ to be monic.*

The next step is to eliminate the numerator $p$. Any local minimum $p/q$ of (6) must be the orthogonal projection of $f$ onto $V_q$, the $n$–dimensional linear space of strictly proper rational functions whose denominator is $q$, with respect to $\mu$. Thus $p$ can be computed by solving a linear system and becomes a function of $q$ denoted by $L_n^\mu(q,f)$. Finally, problem $(P_2)$ can be solved by minimizing the function $\psi_n^\mu(.,f)$ defined on the set $\Delta_n$ of real polynomials of degree $n$ whose roots belong to $\mathbb{D}$ by:

$$\psi_n^\mu(q,f) \;=\; \|f - \frac{L_n^\mu(q,f)}{q}\|_\mu^2\,. \tag{8}$$

Here $L_n^\mu(q,f)/q$ is the orthogonal projection of $f$ onto $V_q = \{z^i/q,\ i = 0,\ldots,n-1\}$ in $L^2(\mu)$. When it is clear from the context, the dependence on $f$ will be omitted in $\psi_n^\mu$ and $L_n^\mu$. In the case of the Lebesgue measure, $L_n^\lambda(q)$ can be easily computed as the remainder of some division in $H^2$ [BOW]. Although the general situation is more complicated, we propose below an integral representation formula for $L_n^\mu(q)$.

Define the reciprocal polynomial $\widetilde{P}$ of a real polynomial $P$ of formal degree $k$ by $\widetilde{P}(z) = z^k\,P(1/z)$. Note that $P$ has exact degree $k$ if and only if $\widetilde{P}(0) \neq 0$ and that $P$ and $\widetilde{P}$ always have the same roots on $\mathbb{T}$.
Let $\{\Phi_j^q\}_{j\geq 0}$ denote the system of orthonormal polynomials on $\mathbb{T}$ for the measure $d\mu/|q|^2$ (see [S, XI] and also [BCS]). The orthogonal polynomial $\Phi_j^q$ has precisely degree $j$ and its roots lie in $\mathbb{D}$ [S, thm.11.4.1].

**Proposition 2** *The polynomial $L_n^\mu(q)$ is given by*

$$L_n^\mu(q)(z) = \frac{1}{2i\pi}\int_{\mathbb{T}} \frac{\check{f}}{q}(\xi)\frac{\widetilde{\Phi}_n^q(\xi)\widetilde{\Phi}_n^q(z) - \Phi_n^q(\xi)\Phi_n^q(z)}{1-\xi z}\,|\nu(\xi)|^2\,d\xi\,. \tag{9}$$

*Proof:* by choosing $\{\Phi_j^q/q\},\ j = 0,\cdots,n-1$, as a basis of $V_q$, we get that

$$L_n^\mu(q)(z) = \sum_{j=0}^{n-1} < f, \frac{\Phi_j^q}{q} >_\mu\ \Phi_j^q(z) \;=\; \frac{1}{2i\pi}\int_{\mathbb{T}} \frac{\check{f}}{q}(\xi)\sum_{j=0}^{n-1}\Phi_j^q(\xi)\Phi_j^q(z)\,|\nu(\xi)|^2\,d\xi\,,$$

in view of (5). Using the Christoffel–Darboux formula [S, XI] for the Szegö kernel:

$$\sum_{j=0}^{n-1} \Phi_j^q(\xi)\,\Phi_j^q(z) = \frac{\widetilde{\Phi}_n^q(\xi)\,\widetilde{\Phi}_n^q(z) - \Phi_n^q(\xi)\,\Phi_n^q(z)}{1-\xi\,z}\,, \tag{10}$$

we obtain (9). □

*We shall restrict ourselves to the case where $\nu = 1/\widetilde{w}$, for a monic polynomial $w$ of degree $d$ whose roots lie inside the unit disk $\mathbb{D}$ and which does not vanish at zero.* This class of weights is rich enough to give some freedom in the choice of the desired shape while the technical complexity of the computations remains limited. As we shall see, for a weight of this form, the numerator (9) is given by a formula involving the $d$ orthogonal polynomials $\Phi_n^q, \ldots, \Phi_{n+d-1}^q$, which can be easily computed from $\Phi_{n+d}^q = q\,w$ using a descending recurrence. This is of particular interest when $d$ is not too large.

For such a weight, we have

$$|\nu(z)|^2 = \frac{z^d}{w(z)\,\widetilde{w}(z)}\,,\ z \in \mathbb{T}\,, \tag{11}$$

and the integral representation (9) can be rewritten as:

$$L_n^\mu(q)(z) = \frac{1}{2i\pi}\int_{\mathbb{T}} \frac{\check{f}_w}{qw}(\xi)\frac{\widetilde{\Phi}_n^q(\xi)\widetilde{\Phi}_n^q(z) - \Phi_n^q(\xi)\Phi_n^q(z)}{1-\xi z}\,d\xi\,, \tag{12}$$

where

$$f_w = \frac{f}{w}\,, \tag{13}$$

so that $\check{f}_w(z) = \check{f}(z)\,z^d/\widetilde{w}(z)$ is analytic in $\mathbb{D}$. It is easily seen from (10) that, for $d \geq 1$,

$$L_n^\mu(q,f)(z) + \sum_{j=n}^{n+d-1} < f, \Phi_j^q/q >_\mu \ \Phi_j^q(z) = \frac{1}{2i\pi}\int_{\mathbb{T}} \frac{\check{f}_w}{qw}(\xi)\frac{\widetilde{q}\widetilde{w}(\xi)\widetilde{q}\widetilde{w}(z) - qw(\xi)qw(z)}{1-\xi z}\,d\xi\,,$$

where the right hand–side is precisely the numerator $L_{n+d}^\lambda(qw, f_w)$ of degree $n+d-1$ associated to the denominator $qw$ in the unweighted approximation of $f_w$ at degree $n+d$, so that $\widetilde{L}_{n+d}^\lambda(qw, f_w)$ can be interpreted as the remainder in the division of $\check{f}_w\widetilde{q}\widetilde{w}$ by $qw$ as in [BCO]. We then relate the weighted scheme to the unweighted one by:

$$\left\{\begin{array}{lcl} L_n^\mu(q,f) & = & L_{n+d}^\lambda(qw, f_w) - \sum_{j=n}^{n+d-1} < f_w, \Phi_j^q/qw > \ \Phi_j^q\,, \\ \psi_n^\mu(q,f) & = & \psi_{n+d}^\lambda(qw, f_w) + \sum_{j=n}^{n+d-1} < f_w, \Phi_j^q/qw >^2\,. \end{array}\right.$$

Following [BCO], we prove in the next sections that the function $\psi_n^\mu$ defined by (8) does extend smoothly to an open neighborhood of $\Delta_n$; this will enable us to describe an algorithm to find local minima of $\psi_n^\mu(q,f)$ using a gradient algorithm and proceeding inductively on the degree.

## 4.2 Extension of the criterion.

Whenever $\nu$ satisfies (11), the system $\{\Phi_j^q\}$ can be computed by [S]:

$$\Phi_j^q(z) = z^{j-n-d}q(z)w(z), \quad j \geq n+d\,, \tag{14}$$

together with the induction formulas:

$$\left\{ \begin{array}{rcl} \widetilde{\Phi}^q_j(0)\,\widetilde{\Phi}^q_j(z) & = & \widetilde{\Phi}^q_{j+1}(0)\,\widetilde{\Phi}^q_{j+1}(z) \;-\; \Phi^q_{j+1}(0)\Phi^q_{j+1}(z)\,,\;\; 0 \le j < n+d\,, \\ \\ \widetilde{\Phi}^q_j(0)^2 & = & \widetilde{\Phi}^q_{j+1}(0)^2 \;-\; \Phi^q_{j+1}(0)^2 = \sum_{k=0}^{j} \Phi^q_k(0)^2 \end{array} \right. \tag{15}$$

**Proposition 3** *There exists a neighborhood $\mathcal{V}$ of $\overline{\Delta_n}$ such that, for $q \in \mathcal{V}$ and $j \ge n$, formulas* (14) *and* (15) *define polynomials $\Phi^q_j$ of degree $j$ and the map $q \mapsto \Phi^q_j$ is smooth. Moreover, if $q = uq'$ where $u$ is monic of degree $m$ and has all its roots of modulus* 1 *while $q'$ belongs to $\Delta_{n-m}$, then $\Phi^q_n = u\,\Phi^{q'}_{n-m}$.*

*Proof:* first, for $j \ge n+d$, formula (14) defines polynomials $\Phi^q_j$ of degree $j$ that are clearly smooth functions of $q$ on any open neighborhood of $\Delta_n$. Now, given any $q$, formula (15) allows to smoothly deduce $\Phi^q_j$ from $\Phi^q_{j+1}$ as long as

$$\widetilde{\Phi}^q_{j+1}(0)^2 - \Phi^q_{j+1}(0)^2 \neq 0. \tag{16}$$

Let us prove by induction that (16) is satisfied for $n \le j \le n+d-1$ and $q$ in some neighborhood of $\overline{\Delta_n}$. Assume first that $q$ belongs to $\overline{\Delta_n}$, so that $\Phi^q_{j+1}$, being the limit of orthonormal polynomials whose roots lie inside $\mathbb{D}$, has all its roots in $\bar{\mathbb{D}}$. In this case, $\Phi^q_{j+1}(0)/\widetilde{\Phi}^q_{j+1}(0)$ being the product of the roots of $\Phi^q_{j+1}$ has modulus at most 1, and thus, unless each root of $\Phi^q_{j+1}$ belong to $\mathbb{T}$, (16) is true. Now, it is easily proved from the recurrence formulas that, when defined, the polynomials $\Phi^q_{j+1}$ and $\Phi^q_j$ have same roots on $\mathbb{T}$. From (14), these roots are precisely the roots of $q$ on $\mathbb{T}$. So, $\Phi^q_{j+1}$ has at most $n$ roots on the circle and (16) is true for $j \ge n$. By continuity it is still valid in a neighborhood of $\overline{\Delta_n}$.
To get the second assertion, observe that formula (14) implies that $\Phi^q_{n+d} = u\Phi^{q'}_{n+d-m}$. Moreover, using that $u(0)^2 = \widetilde{u}(0)^2 = 1$ and $u(z) = u(0)\widetilde{u}(z)$, it can be proved by induction from (15) that $\Phi^q_j = u\Phi^{q'}_{j-m}$ for $j \ge n \ge m$. □

**Proposition 4** *Whenever $\check{f}$ and $\nu$ are analytic in a disk $D_r = \{\, z, |z| < r \,\}$ for some $r > 1$, the map $\psi^\mu_n$ smoothly extends to a neighborhood $\mathcal{V}$ of $\overline{\Delta_n}$.*

*Proof:* from (13), the function $\check{f}_w$ is analytic in the disk $D_r$. In the integral representation (12) the unit circle $\mathbb{T}$ can be deformed into any contour $\Gamma$ contained in $D_r$ that encompasses the roots of $q$. Choosing the neighborhood $\mathcal{V}$ of $\Delta_n$ in proposition 3 in order for $q$ to have all its roots in $D_r$ allows this integral on $\Gamma$ to remain defined for $q \in \mathcal{V}$ and yields a smooth extension of $L^\mu_n$.
Furthermore, if $q \in \Delta_n$, properties of the orthogonal projection show that

$$\psi^\mu_n(q) \;=\; \|f - \frac{L^\mu_n(q)}{q}\|^2_\mu \;=\; \|f\|^2_\mu - < f, \frac{L^\mu_n(q)}{q} >_\mu, \tag{17}$$

so that it is sufficient to smoothly extend the map $q \mapsto < f, z^j/q >_\mu$ for every $j$, and this is done by putting

$$< f, z^j/q >_\mu \;=\; \frac{1}{2i\pi} \int_\Gamma \check{f}_w(\xi) \frac{\xi^j}{q(\xi)w(\xi)} d\xi.$$

□

Let us denote by $\nabla_n(q)$ the gradient vector of $\psi^\mu_n$ at the point $q$. The following lemma can be proved as in [BOW], using propositions 3 and 4.

**Lemma 1** *Let $q \in \partial\Delta_n$ and suppose that $q = uq'$ where $u$ is monic of degree $m$ and has all its roots of modulus 1 while $q'$ belongs to $\Delta_{n-m}$. Then $L_n^\mu(q) = uL_{n-m}^\mu(q')$ and $\psi_n^\mu(q) = \psi_{n-m}^\mu(q')$.*
*Moreover, if $q$ belongs to some smooth part of $\partial\Delta_n$, $q' \in \Delta_{n-1}$ is a critical point of $\psi_{n-1}^\mu$, then $\nabla_n(q)$ is orthogonal to $\partial\Delta_n$ and points outwards.*

### 4.3 An algorithm to find a local minimum.

We shall assume in this section that, for $k = 1 \dots n$, $\nabla_k$ does not vanish on $\partial\Delta_k$ and that all the critical points of $\psi_k^\mu$ on $\Delta_k$ are non degenerate[2]. Whenever these assumptions are satisfied, $\psi_k^\mu$ has a finite number of critical points in $\Delta_k$ and an algorithm can be described following the same scheme than in [BCO].

The function $\psi_n^\mu$ is smooth and its local minima belong to $\Delta_n$ by proposition 1, which is open and bounded in $\mathbb{R}^n$ (a polynomial $q(z) = z^n + q_{n-1}z^{n-1} + \dots + q_0$ of $\Delta_n$ is represented by its coefficients $(q_{n-1}, q_{n-2}, \dots, q_0)$). Therefore, local minima are critical points of $\psi_n^\mu$ and can be found by a gradient algorithm. We integrate the vector field $-\nabla_n$ from an initial point. If we meet the boundary of $\Delta_n$ then, by lemma 1, we are led to solve a problem of lower order. Conversely, still by lemma 1, a local minimum of $\psi_k^\mu$, $k < n$, provides a suitable initial point to integrate $-\nabla_{k+1}$. The procedure can thus continue through different orders (strictly positive, since $\psi_0^\mu = \|f\|_\mu^2 \geq \psi_n^\mu$ on $\Delta_n$) whereas the value of the criterion (which is $\psi_k^\mu$ while integrating $-\nabla_k$) decreases. Thus, a local minimum cannot be met twice; since local minima are finite in number, the procedure converges.

**Remark:** in the unweighted case, the approximation problem in $\bar{H}^2$ reduces to the approximation problem in $\bar{H}_0^2$. This is an obvious consequence of the $L^2(\mathbb{T})$–orthogonality between the space of constant valued functions and $\bar{H}_0^2$. This is no more true for the weighted approximation problem in $\bar{H}^2$; however, it may be solved as $(P_2)$, using orthogonal polynomials from degree $n+1$ instead of degree $n$.
In the particular case of weights of *degree one*, Möbius transforms play an important role. Indeed, let $\nu(z) = 1/(1 - w_0\, z)$, $w_0 \in (-1, 1)$. By an easy computation, we obtain:

$$< f, g >_\nu = \frac{1}{1 - w_0^2} < f \circ \phi_{w_0}, g \circ \phi_{w_0} >,$$

where $\phi_{w_0}$ is the Möbius transform of the unit disk defined by:

$$\phi_{w_0}(z) = \frac{z + w_0}{1 + w_0 z}.$$

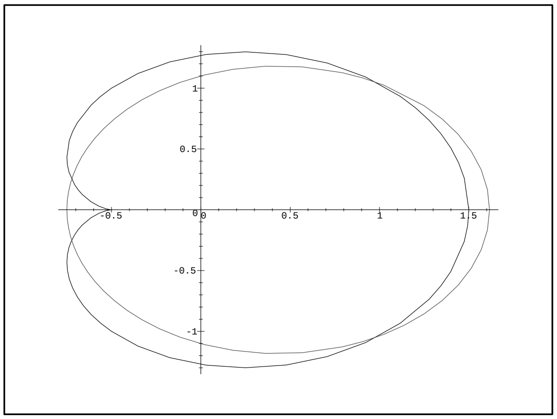
The map $f \mapsto f \circ \phi_{w_0}/\sqrt{1 - w_0}$ is an isometry from $H^2(\mu)$ onto $H^2$ (and also from $\bar{H}^2(\mu)$ onto $\bar{H}^2$) which preserves the Mac-Millan degree of a rational function. It thus allows to handle a version of $(P_2)$ stated in $\bar{H}^2$ by solving analogous unweighted problems.

As an illustration, we numerically solve problem $(P_2)$ with $n = 1$ for the function

$$f(z) = \frac{z + 0.5}{z^2},$$

[2] For the Lebesgue measure, these two properties hold in an open dense subset of the space of $\bar{H}_0^2$ functions that are analytic outside a disk $D_r$ for $r < 1$ [B]. Although this has not been established yet in the weighted case, it seems reasonable that this "genericity" result still holds.

and different weights of degree one, given by (11) for some $w(z) = z - w_0$, $|w_0| < 1$. Figures 1 and 2, show the Nyquist diagrams of $f$ together with its the best rational approximants of degree one for $w_0 = 0$ (unweighted approximation), $w_0 = 0.5$, and $w_0 = 0.9$.



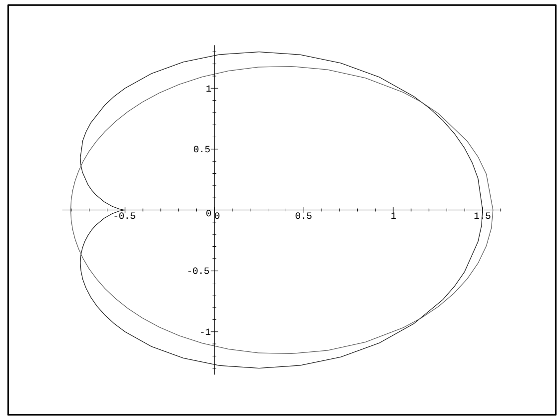
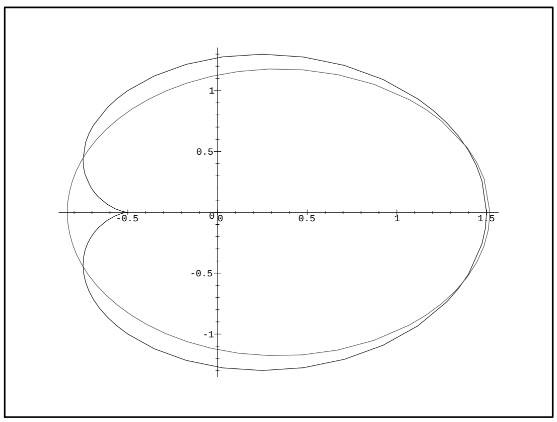




Figure 1: $w_0$=0; a=0.366;

$w_0$=0.5; a=0.3194



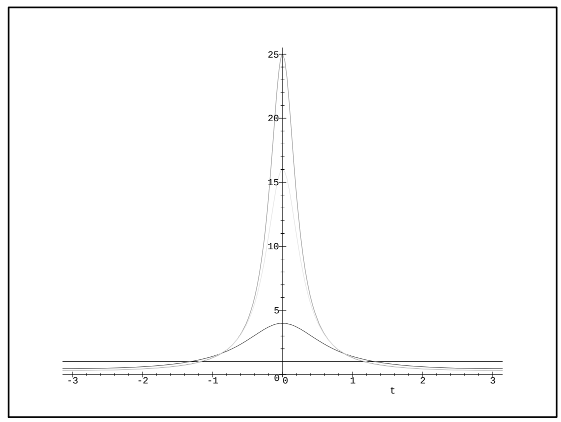




Figure 2: $w_0$=0.9; a=0.289;

$\left|\frac{1}{e^{i\theta}-w_0}\right|^2$ for $w_0 = 0,\ 0.5,\ 0.75$, and $0.8$

Such weights act more or less on low frequencies. This suggests to use them in order to approach the rational approximation problem when one is mainly concerned with the quality of the approximation around 0 or $\pi$ (where the same phenomenon would occur for $w_0 < 0$). The use of a higher degree $w$ would allow to build more refined shapes for the weight and to insist on arbitrary symmetric frequencies by choosing complex conjugate roots.

# 5 Conclusion

For the family of measures satisfying hypothsesis $(H)$, the solution of the weighted bounded extremal problem $(P_1)$ can be deduced from the solution of the unweighted one by an explicit change of

variable. Note that $(P_1)$ could be approached for more general measures on $\mathbb{T}$, namely the ones induced by Szegö weights for which the Adamjan–Arov–Krein theory has already been extended [CS1, CS2]. Concerning the weighted rational approximation issue $(P_2)$, orthogonal polynomials on $\mathbb{T}$ for $d\mu/|q|^2$ are used to express the best numerator in the criterion (8) and to establish its smoothness property if (11) holds. The natural idea to appeal to a basis of orthogonal polynomials on $\mathbb{T}$ has been used in [BCS]. In this note, an unweighted rational approximation problem is studied, which can be expressed as $(P_2)$ for $\mu = \lambda$ with the additional and difficult constraint that the degree of the numerator should be less or equal to some fixed $m < n - 1$. This constraint prevents from smoothly extending the criterion. This relies on the fact that, if $q$ has more than $m$ roots on $\mathbb{T}$, then $\Phi_m^q$ does not extend smoothly to $\overline{\Delta}_n$. Generalizations of our present work and of results in [BCS] to analogous $(m, n)$ weighted rational approximation problems remain under study.
It would also be interesting to answer further theoretical questions such as the consistency problem: if $f$ is already rational of degree $n$, is it the *single* critical point of the problem? Once again, the answer does not come straightfully as in the unweighted scalar case (for which consistency holds) and depends on the measure $\mu$. It appears that consistency may fail when $\mu$ is given by (3) and (11) for a polynomial $w$ of degree larger than 2. This is a relevant question when studying identification schemes, which is classically handled in a stochastic framework, see [L]. For arbitrary $f$ in $\bar{H}_0^2$, the criterion $\psi_n^\mu$ generally has several local minima and despite our algorithm will converge to one of these, we cannot get sure to find them all. This is an additional motivation for introducing a weight in the rational approximation problem $(P_2)$ since it allows to consider the following uniqueness issue: given $f$ in $\bar{H}_0^2$, is it possible to find a measure $\mu$ which ensures uniqueness of the critical points of $\psi_n^\mu$?

## References


[A] L. Aizenberg. *Carleman's formulas in complex analysis.* Kluwer Academic Publishers, 1993.

[ABL] D. Alpay, L. Baratchart, and J. Leblond. Some extremal problems linked with identification from partial frequency data. In J.L. Lions R.F. Curtain, A. Bensoussan, editor, *10th conference on analysis and optimization of systems, Sophia–Antipolis 1992*, volume 185 of *Lect. Notes in Control and Information Sc.*, pages 563–573. Springer-Verlag, 1993.

[BGR] J. Ball, I. Gohberg, and L. Rodman. *Interpolation of rational matrix functions.* Birkhäuser Verlag, Basel, 1990.

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

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

[BCS] L. Baratchart, M. Cardelli, and E.B. Saff. An algorithm for a certain type of rational approximation in $H_2^-$. Unpublished.

[BL] L. Baratchart and J. Leblond. Hardy approximation to $L^p$ functions on subsets of the circle with $1 \leq p < \infty$. *Constructive Approximation*, 14:41–56, 1998.

[BLP] L. Baratchart, J. Leblond, and J.R. Partington. Hardy approximation to $L^\infty$ functions on subsets of the circle. *Constructive Approximation*, 12:423–436, 1996.

[BOW] L. Baratchart, M. Olivi, and F. Wielonsky. On a rational approximation problem in the real hardy space $H_2$. *Theoretical Computer Science*, 94:175–197, 1992.

[CS1] M. Cotlar and C. Sadosky. Abstract, weighted, and multidimensional AAK theorems, and the singular numbers of Sarason commutants. *Integral equations operator theory*, 17:169–201, 1993.

[CS2] M. Cotlar and C. Sadosky. Hankel forms and operators in Hardy spaces with two Szegö weights. In *Proceedings of IWOTA'96*, Operator Theory, adv. and appl. Birkäuser, 1997.

[E] D. Enns. Model reduction for control systems design. Ph.D. thesis, department of Aeronautics and Astronautics, Stanford University, 1984.

[F] B. Francis. *A course in $H^\infty$ control theory.* Lectures notes in control and information sciences. Springer–Verlag, 1987.

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

[Ge] M. Gevers. *Towards a joint design of identification and control?*, pages 111–151. H.L. Trentelman and J.C. Willems editors, Birkhauser, 1993.

[GK] G. Gu and P.P. Khargonekar. Linear and nonlinear algorithms for identification in $H_\infty$ with error bounds. *IEEE Trans. on Aut. Control*, 37:953–963, 1992.

[HKN] A.J. Helmicki, C.A. Jacobson, and C.N. Nett. Control–oriented system identification: a worst-case/deterministic approach in $H_\infty$. *IEEE Trans. on Aut. Cont.*, 36:1163–1176, 1991.

[H] K. Hoffman. *Banach spaces of analytic functions.* Dover, 1988.

[KN] M.G. Krein and P.Y. Nudel'man. Approximation of $L^2(\omega_1, \omega_2)$ functions by minimum– energy transfer functions of linear systems. *Problemy Peredachi Informatsii*, 11(2):37–60, 1975. English translation.

[LA] G.A. Latham and B.D.O. Anderson. Frequency-weighted optimal hankel norm approximation of stable transfer functions. *Systems and Control Letters*, 5:229–236, 1985.

[LO] J. Leblond and M. Olivi. Weighted $H^2$ approximation of transfer functions, 1996. INRIA research report no. 2863.

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

[MPG] P.M. Mäkilä, J.R. Partington, and T.K. Gustafsson. Worst-case control-relevant identification. *Automatica*, 31:1799–1819, 1995.

[Par] J.R. Partington. *Interpolation, identification and sampling.* Oxford University Press, 1997. To appear.

[Pat] D.J. Patil. Representation of $H^p$ functions. *Bull. A.M.S.*, 78(4), 1972.

[R] Y. Rozanov. *Stationary Random Processes.* Holden Day, San Francisco, 1967.

[S] G. Szegö. *Orthogonal Polynomials.* Colloquium Publications. AMS, 1939.

[T] N. Torkhani. Contribution à l'identification fréquentielle robuste des systèmes dynamiques linéaires. Thèse de doctorat, Ecole Nationale des Ponts et Chaussées, décembre 1995.