Weighted H^2 approximation of transfer functions

Juliette Leblond and Martine Olivi
INRIA
BP 93, 06902 Sophia–Antipolis Cedex, France, e-mail :
leblond,olivi@sophia.inria.fr

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

Keywords
Hardy spaces, weighted rational approximation, frequency domain identification.

1 INTRODUCTION

Let $L^2(\mathbb{T})$ be the *real* Hilbert space of square–summable functions on the unit circle $\mathbb{T}$ satisfying the conjugate–symmetry property $f(e^{-i\theta}) = \overline{f(e^{i\theta})}$, or equivalently whose Fourier coefficients are real. The space $L^2(\mathbb{T})$ is endowed with its classical inner product $<\ ,\ >$ and associated norm $\|\ \|$. The Hardy space H^2 of the unit disk $\mathbb{D}$ is the closed subspace of $L^2(\mathbb{T})$ which contains functions whose Fourier coefficients of negative index are zero. These functions admit an analytic extension in $\mathbb{D}$ and H^2 is isometric to the space of analytic functions in $\mathbb{D}$ which satisfy the growth condition:

$$\sup_{r<1} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(r\, e^{i\theta})|^2 d\theta < \infty\,, \tag{1}$$

which provides a norm on this space (see Garnett (1981), II.3). Furthermore, $\bar{H}_0^2$ denotes the orthogonal complement of H^2 in $L^2(\mathbb{T})$. It consists in $L^2(\mathbb{T})$ functions which possesses Fourier coefficients of non–negative index equal to zero or, equivalently, in functions analytic outside the closed unit disk and vanishing at infinity satisfying the growth condition (1) for $r > 1$. The two spaces H^2 and $\bar{H}_0^2$ are isometrical under the map:

$$\begin{array}{rcl} L^2(\mathbb{T}) & \to & L^2(\mathbb{T}) \\ f(z) & \mapsto & \dfrac{f(1/z)}{z} = \check{f}(z)\,. \end{array} \tag{2}$$

Functions belonging to $\bar{H}_0^2$ are transfer functions of linear strictly causal stationary discrete time systems of finite variance for a white noise input or, equivalently, which possess an $l^1(\mathbb{N})$–input / $l^2(\mathbb{N})$–output stability property. For an unknown $\bar{H}_0^2$ transfer function, assume that we are only given some of its (possibly noisy) pointwise values at some frequencies belonging to a symmetric subset K of $\mathbb{T}$ (which corresponds to the bandwidth of the associated system). Such measurements may be obtained using harmonic identification procedures. In order to identify the unknown system, we want to find a rational $\bar{H}_0^2$ function of bounded Mac–Millan degree accounting well enough for these data, in a sense to be made precise below.

Let μ be some positive finite measure on the unit circle $\mathbb{T}$, absolutely continuous with respect to the Lebesgue measure λ (for which we write $d\lambda(\theta) = d\theta$) and satisfying $d\mu(-\theta) = -d\mu(\theta)$. In this work, we want to identify the unknown transfer function by solving approximation problems with respect to this measure μ. This will take place in the *real* Hilbert space $L^2(\mu)$ of conjugate–symmetric functions that are square–summable on $\mathbb{T}$ with respect to the measure μ endowed with the inner product defined by:

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

and with the associated norm $\|\ \|_\mu$. This will be done for every measure μ such that $L^2(\mathbb{T}) = L^2(\mu)$, space equality which we assume to hold during the sequel of this introduction and for which we will provide a necessary and sufficient condition in section 2. This condition is indeed required for problems (P_1) and (P_2) below to make sense.

In the classical stochastic framework, this type of approximation problems come up when minimizing the variance of the output error between the searched model and the "true system", the quantity $d\mu/d\lambda$ being the spectral density of the noisy input (when this input is a white noise, then $\mu = \lambda$). Such a weighting μ may also be present in the criterion when one pursues an identification procedure with the purpose of designing a controller, in which case μ represents the control performance criterion. Moreover, these weighted L^2 criteria may arise in control problems, when computing stable optimal controllers under some parametrization, for example. Such a μ may be simply used to weight some frequencies more than the others and to represent the confidence one has in the available measurements for either identification, filtering or control issues.

For any symmetric subset Γ of $\mathbb{T}$, $L^2(\Gamma)$ stands for the *real* Lebesgue space of square summable conjugate–symmetric functions on Γ. If χ_Γ denotes the characteristic function of Γ, $\|f\|_{\Gamma,\mu}$ is defined to be equal to $\|\chi_\Gamma f\|_\mu$; it induces a norm on the space of square summable functions on Γ with respect to μ.

A preliminary step is to get an interpolant $\phi \in L^2(K)$ for the given experimental data and a function $\kappa \in L^2(J)$, $J = \mathbb{T} \setminus K$, which reflects the behaviour of the system to be identified outside the bandwidth (if nothing is known, one can take $\kappa = 0$). Now, the identification procedure we are concerned with splits in two approximation issues. The first one furnishes an $\bar{H}_0^2$ approximant to the prescribed data and can be formulated as a bounded (dual) extremal problem:

(P_1) *Given $\phi \in L^2(K)$, $\kappa \in L^2(J)$, and $M > 0$, find a function $f_0 \in \bar{H}_0^2$ which minimizes $\|\phi - f\|_{K,\mu}$ among the functions $f \in \bar{H}_0^2$ which satisfy the constraint $\|\kappa - f\|_{J,\mu} \leq M$.*

The second issue is the following rational approximation problem which may be associated to a model reduction step, in the overall identification procedure.

(P_2) *Given $f \in \bar{H}_0^2$ and an integer $n > 0$, find a rational function $r \in \bar{H}_0^2$ of Mac–Millan degree less than n which minimizes $\|f - p/q\|_\mu$, where p/q ranges over the rational functions in $\bar{H}_0^2$ of Mac–Millan degree less than n.*

Problems (P_1) and (P_2) are stated here in the Hardy space $\bar{H}_0^2$ where they may be associated to identification issues for discrete time systems. However, analogous identification questions for continuous time systems, which naturally take place in Hardy spaces of the right half–plane, can also be formulated this way up to a Möbius transformation.

We give now the main statements of our results, the proofs of which are detailed in Leblond and Olivi (1995). We first state a necessary and sufficient condition for $L^2(\mathbb{T}) = L^2(\mu)$ to hold in section 2. We then solve the analytic approximation problem (P_1) and provide an explicit characterization of its solution in section 3. Finally, section 4 is devoted to the rational approximation problem (P_2) for the solution of which we lay the foundation stone of an algorithm.

2 WEIGHTED H^2 SPACES

Let $H^2(\mu)$ and $\bar{H}_0^2(\mu)$ be the weighted *real* Hardy spaces respectively defined to be the $L^2(\mu)$ closure of polynomials and the $L^2(\mu)$ closure of $\{1/z^k,\ k > 1\}$. Let $L^\infty(\mathbb{T})$ be the *real* Banach space of essentially bounded conjugate–symmetric functions. The Hardy space H^∞ is defined to be the $L^\infty(\mathbb{T})$ closure of polynomials; it verifies $H^\infty = H^2 \cap L^\infty(\mathbb{T})$ (Garnett (1981), II.4). We then have:

Theorem 1 *Let μ be a finite positive measure absolutely continuous with respect to the Lebesgue measure and such that $d\mu(-\theta) = -d\mu(\theta)$. Then, $L^2(\mathbb{T}) = L^2(\mu)$ if and only if*

$$d\mu = |\nu|^2 \, d\lambda \tag{3}$$

for a function ν which belongs to H^∞ and is invertible in H^∞. In this case, we also have:

$$H^2 = H^2(\mu) \quad and \quad \bar{H}_0^2 = \bar{H}_0^2(\mu)\,.$$

The proof of theorem 1 relies in particular on Szegö's and Beurling's theorems, see e.g. Garnett (1981), Hoffman (1988).

We assume in the sequel that μ verifies (3). Consequently, the above approximation problems (P_1) and (P_2) deal with and lead to $\bar{H}_0^2$ transfer functions and then respects the usual stability property of the systems; this is required for identification purposes though not for mathematical reasons. Moreover, for any symmetric subset Γ of $\mathbb{T}$, the space of square–summable functions on Γ with respect to μ (that is of finite $\| \; \|_{\Gamma,\mu}$ norm) coincides with $L^2(\Gamma)$.

For sake of simplicity but without loss of generality, we suppose from now on that the measure μ is such that:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} d\mu(\theta) = 1\,,$$

or equivalently that $\|\nu\| = 1$.

3 WEIGHTED H^2 APPROXIMATION

In this section, we get a solution to problem (P_1) by solving the following analogous problem (P_1') in H^2. Indeed, thanks to the isometry (2), problems (P_1) and (P_1') are equivalent.

(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 $\|h - g\|_{J,\mu} \leq M$.*

In the unweighted case where $\mu = \lambda$, problem (P_1') has been solved in Alpay et al. (1993) and, in a somewhat dual form, in Krein and Nudel'man (1975). Moreover, it has been approached in the general H^p setting, $1 \leq p \leq \infty$, in Baratchart et al. (1994).

Assume further that K is a subset of $\mathbb{T}$ such that both K and its complementary J are of positive Lebesgue measure. Theorem 1 allows us to deduce the following solution to problem (P_1') from results obtained in the unweighted case (see Alpay et al. (1993)). Let

$$C_M^h(\mu) = \{g_{|K}\,,\; g \in H^2 \text{ s.t. } \|h - g\|_{J,\mu} \leq M\}\,.$$

Theorem 2 *If μ is given by* (3) *for some $\nu \in H^\infty$, invertible in H^∞, then, there exists a unique solution $g_0 \in H^2$ to problem (P_1'). Moreover, if $\varphi \notin C_M^h(\mu)$, $\|h - g_0\|_{J,\mu} = M$.*

Note that when $M \to \infty$, problem (P_1') becomes ill–posed. This follows from the fact that $H^2_{|K}$, the space of traces on K of H^2 functions, is dense in $L^2(K)$. Moreover, let (g_n) be a sequence of H^2 such that $\|\varphi - g_n\|_{K,\mu}$ tends to 0; then, if φ is not the trace on K of an H^2 function, then $\lim_{n\to\infty} \|g_n\|_{J,\mu} = \infty$.

Now, the solution g_0 of problem (P_1') can be explicitly characterized. Let $T \;:\; H^2 \to H^2$ be the Toeplitz operator with symbol χ_J:

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

where P_{H^2} denotes the orthogonal projection from $L^2(\mathbb{T})$ onto H^2. The operator T is bounded, self-adjoint, positive, with norm 1 and spectrum equal to $[0,1]$. We have that:

$$g_0 = \nu^{-1}(1+lT)^{-1} P_{H^2}\left(\nu(\chi_K \varphi + (l+1)\chi_J h)\right) , \tag{4}$$

for some $l \in (-1,+\infty)$ such that $\|g_0 - h\|_{J,\mu} = M$.

This is obtained from the analogous expression of the solution to problem (P_1') for $\mu = \lambda$ which can be recovered from (4) by setting $\nu = 1$ (Alpay et al. (1993), Baratchart and Leblond (1993)). Hence, for given functions φ, h and a measure μ, the solution (4) of problem (P_1') is equal to the product of ν^{-1} by the solution of the analogous unweighted problem for the functions $\varphi\,\nu$ and $h\,\nu$.

It can also be shown that M is decreasing with respect to l so that g_0 can be numerically computed using a dichotomy procedure. Furthermore, as $l \to -1$, then the error $\|\varphi - g_0\|_{K,\mu}$ goes to zero while $M \to \infty$ if $\varphi \notin H^2_{|K}$. However, if $\varphi \in H^2_{|K}$, then it follows from a result of Patil (1972) that, as $l \to -1$, $g_0 \to \varphi$ in H^2.

Finally, thanks to (2), this provides us with the solution to problem (P_1). Indeed, if we take $\varphi = \check{\phi}$, $h = \check{\kappa}$, and then build the associated solution g_0 to (P_1'), then $f_0 = \check{g}_0$ is the solution to (P_1) we are looking for.

4 WEIGHTED RATIONAL APPROXIMATION

About the rational approximation problem (P_2), as in the case of the Lebesgue measure, a result of normality has been proved: unless f is rational of degree less than n, a rational local best approximant to f at order n has degree exactly n.

We assume that f is not rational of degree less than n. In this case, (P_2) becomes:

(P_2) given $f \in \bar{H}^2_0$ and an integer $n > 0$, find a rational function $r \in \bar{H}^2_0$ of Mac-Millan degree n which minimizes $\|f - p/q\|_\mu$, where p/q ranges over the rational functions in $\bar{H}^2_0$ of Mac-Millan degree n.

Recall that a rational function p/q lies in $\bar{H}^2_0$ if and only if $d^\circ p < d^\circ q$ and q is a Schur polynomial (i.e. has all its roots in the unit disk). Moreover, we shall assume that q is monic. The first step is to eliminate the numerator p. Now, the set of approximants has a manifold structure and by differentiating the above criterion with respect to the coefficients of p, we get that every critical point (p,q) satisfies:

$$< f - \frac{p}{q}, \frac{z^i}{q} >_\mu = 0 \ , \ i = 0, \cdots, n-1 .$$

Hence, if V_q denotes the n-dimensional linear subspace of $\bar{H}^2_0$ generated by $\{z^i/q\}$, $i = 0, \cdots, n-1$, the rational approximant p/q is the orthogonal projection of f onto V_q with respect to μ. In this way, p becomes a function of q denoted by $L_\mu(q,f)$ or simply by $L_\mu(q)$ when the dependence on f is clear from the context.

Let now $\{\Phi_j\}$, $j \geq 0$, be the system of orthonormal polynomials on $\mathbb{T}$ for the measure $d\mu/|q|^2$ (see Szegö (1939), XI). By choosing $\{\Phi_j/q\}$, $j = 0, \cdots, n-1$, as a basis of V_q, we get that:

$$L_\mu(q) = \sum_{j=0}^{n-1} < f, \frac{\Phi_j}{q} >_\mu \Phi_j \,. \tag{5}$$

Define the reciprocal polynomial $\tilde{P}$ of a polynomial P of degree k by

$$\tilde{P}(z) = z^k P(1/z) \,.$$

Although $L_\mu(q)$ is a polynomial of degree possibly less than $n-1$, its reciprocal polynomial will still be considered to be:

$$\widetilde{L_\mu(q)}(z) = z^{n-1} L_\mu(q)(1/z) \,.$$

Let g_μ be the orthogonal projection on H^2 of $g\,|\nu|^2$ which can also be expressed as:

$$g_\mu(z) = \frac{1}{2i\pi} \int_{\mathbb{T}} g(\xi) \frac{d\mu(\xi)}{\xi - z} \,, \tag{6}$$

Using the Christoffel–Darboux formula (Szegö (1939), XI):

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

together with the fact that $\Phi_n = q$, we deduce from (7) the following result.

Proposition 1 *The reciprocal polynomial $\widetilde{L_\mu(q)}$ is given by $\widetilde{L_\mu(q)} = \tilde{q}\, g_\mu - q\, v_\mu(q)$, for some function $v_\mu(q)$ which belongs to H^2. Equivalently, $\widetilde{L_\mu(q)}$ is the remainder of the division in H^2 of $\tilde{q}\, g_\mu$ by q.*

An immediate consequence of proposition 1 is that:

$$L_\mu(q, f) = L_\lambda(q, f_\mu) \,, \tag{7}$$

where $f_\mu = \breve{g}_\mu$ and $L_\lambda(q, f_\mu)$ is the numerator associated with q when solving (P_2) for the Lebesgue measure λ and the function f_μ (see Baratchart et al. (1992)).

We are thus led to minimize the function $\Psi_\mu(., f)$ defined on the set Δ_n of monic polynomials q of degree n whose roots belong to $\mathbb{D}$ by:

$$\begin{array}{rccl} \Psi_\mu(., f): & \Delta_n & \longrightarrow & \mathbb{R} \\ & q & \longmapsto & \|f - \frac{L_\mu(q)}{q}\|_\mu^2 \,. \end{array} \tag{8}$$

It follows from (7) that

$$\Psi_\mu(q, f) = \Psi_\lambda(q, f_\mu) + \|f\|_\mu^2 - \|f_\mu\|^2 \,, \tag{9}$$

where $\Psi_\lambda(q, f_\mu)$ is the criterion to be minimized in the unweighted case for the function f_μ. Hence, minimizing the weighted criterion for some function f is just the same than minimizing the usual L^2 one for the associated function f_μ. We can thus apply differential tools as in the case of the Lebesgue measure λ for which problem (P_2) has been solved in Baratchart et al. (1991) and (1992).
The procedure goes as follows. From (8), the function $\Psi_\mu(., f)$ is smooth so that its local minima can be found by a gradient algorithm. Moreover, if g and ν are analytic in a disk $D_r = \{ z, |z| < r \}$ for some $r > 1$, then $\Psi_\mu(., f)$ extends to a neighbourhood of Δ_n and possesses recursive properties: when we meet the boundary of Δ_n we are led to solve a problem of lower order and the solution of a problem of order $k < n$ provides a boundary initial point to search a minimum at order $k+1$. Thus, this procedure continues through different orders until we find a local minimum at order n.

Theoretical questions could also be considered, 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 case and depends on the measure μ. It would be interesting to link our further results with the ones obtained in a stochastic framework, see Ljung (1987).

To conclude, let us mention that we already got good numerical results in the unweighted case for the identification at order 8 of hyperfrequencies filters from experimental data provided by the french CNES.

As a consequence of the results of sections 3 and 4, the algorithms which solve the weighted approximation problems (P_1) and (P_2) are the same than in the unweighted case, up to changes of variables that link the function f to be approximated to the weight μ. These latter are to be implemented.

REFERENCES

Alpay, D., Baratchart, L. and Leblond, J. (1993) 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, L.N.C.I.S.*, Springer-Verlag, **185**, 563–573.

Baratchart, L., Cardelli, M. and Olivi, M. (1991) Identification and rational L^2 approximation : a gradient algorithm. *Automatica*, **27(2)**:413–418.

Baratchart, L. and Leblond, J. (1993) Characterization of solutions to a class of bounded extremal problems in L^2. Unpublished.

Baratchart, L., Leblond, J. and Partington, J.R. (1994) Hardy approximation to L^p functions on subsets of the circle. INRIA research report **2377**.

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

Cardelli, M. and Saff, E.B. (1992) An algorithm for a certain type of rational approximation in H_2^-. Unpublished.

Garnett, J.B. (1981) *Bounded analytic functions.* Academic Press.

Hoffman, K. (1988) *Banach spaces of analytic functions.* Dover.

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

Leblond, J. and Olivi, M. (1995) Hardy approximation in weighted L^2 spaces of an arc. In preparation.

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

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

Szegö, G. (1939) Orthogonal polynomials. *Col. Pub. A.M.S*, **23**.