SOME EXTREMAL PROBLEMS LINKED WITH IDENTIFICATION FROM PARTIAL FREQUENCY DATA

Daniel Alpay,
Department of Mathematics, Ben–Gurion University of the Negev, POB 653, 84105 Beer–Sheva, Israël,

Laurent Baratchart, Juliette Leblond,
INRIA, 2004 route des Lucioles, Sophia–Antipolis, 06565 Valbonne, France.

Abstract

We aim at tackling a robust identification problem for a linear dynamical control system with incomplete frequency data. Its mathematical formulation is a generalization of dual extremal problems in H^p to a subarc of the unit circle $\mathbb{T}$. Specifically, if I is such a subarc and J its complement in $\mathbb{T}$, we study existence and uniqueness of a H^p function of bounded $L^p(J)$ norm which is a best approximation to a given function in $L^p(I)$. Finally, we consider the problem of computing the solution in the case $p = 2$.

1 Introduction

The author's motivation originates with some questions in identification of linear dynamical control systems. Given a stable linear system with unknown transfer function $\mathcal{F}$, harmonic identification procedures allow one to obtain, at least in principle, the values of $\mathcal{F}$ at given frequencies. In practice, the frequencies ω_k's that are accessible in this way lie in a finite interval Ω of the positive imaginary axis corresponding to the bandwidth of the system and one looks for some stable model accounting satisfactorily for these experiments. The type of stability which is sought depends of course on the applications one has in mind. In this paper, we shall be interested in models whose transfer functions belong to some Hardy space $\mathcal{H}^p$ of the half plane for some p such that $1 \leq p \leq \infty$. Note that the behaviour of the model outside Ω is not regarded as very crucial in general, except being stable. Reasons include that frequencies not belonging to Ω often lie beyond the restitution power of the system or outside the validity domain of the linear approximation. Hence, to infer $\mathcal{F}$ from the values $\mathcal{F}(w_k)$, $k = 1, \cdots, N$ is an approximation problem in Ω, still with an eye on what is going on outside Ω, which can be tackled by various methods, a classical account of which can be found in [17].

In this classical approach, the class of models is usually confined within some approximating family reflecting the preconception one has of the system. Optimization is then performed with respect to the parameters of this family. A popular class of models is for example the rational one corresponding to finite dimensional systems for which an extensive amount of control and observation devices has been derived (see e.g. [13], [1], [25], and [26]). In other instances, it may be necessary to include also exponential functions in order to account for delays. For the approximating family to be efficient, it ought to be more or less dense in the class of systems one is interested in, and this is to the

effect that the values $\mathcal{F}(w_k)$ could always be matched by using sufficiently many parameters. Models are thus further discriminated by the trade–off between their complexity and their behaviour in between the frequencies w_k's. At first, it is tempting to minimize the complexity of the model but this may lead to serious difficulties. Well–known examples are polynomial Lagrange interpolation or rational Padé approximation whose behaviour can be very wild and sometimes fail to converge in many reasonable senses if additional experiments are performed, even in the ideal case where the data are noise free and the transfer function $\mathcal{F}$ truly belongs to the closure of the model class (see e.g. [20], [18]). This not to mention the problem of keeping poles and zeroes in appropriate regions of the complex plane. To remedy this, one usually drops the requirement that the model should match the experimental values exactly since these, after all, are subject to measurement errors. Keeping the complexity at a desired level, one rather tries to fit the experimental points as well as possible while meeting some constraints on the global behaviour of the model. One issue, however, remains unclear namely to what extend this global pattern depends on the original approximating family.
To avoid dealing with all these constraints at the same time, two–steps algorithms have been recently advocated in [9] and [11] where a first identification is performed in L^∞ and further approximated by a H^∞ one using Nehari extension. This allows one to pass from discrete data to a continuous description of the behaviour of the system without having to worry about the belonging of the model to the given class at this early stage. Pursuing this point of view further, one may even argue that the identification in the first step above need not be performed outside the bandwidth Ω where no data is available anyway. The second step amounts then to the following question : suppose the desired behaviour on all of Ω is explicitly specified ; what is the best possible approximation satisfying the stability constraint ?
Let us be more precise and assume we deal with a single–input single–output system whose desired behaviour on Ω is given by some function f belonging to $L^p(\Omega)$. Assume further that we want the model to be $\mathcal{H}^p$ stable. If we ask directly what is a function $g \in \mathcal{H}^p$ which is closest to f in $L^p(\Omega)$, the question has usually no answer when $p < \infty$, unless f is already the trace on Ω of some $\mathcal{H}^p$ function. But if we recall that we do not want the model to spread off too much outside Ω, and consequently if we bound the L^p norm of g on the complement of Ω, the problem becomes well–posed. Suppose we solved it and let g be a solution. Then g is a causal and stable model for the system. Furthermore, if f and g are not close enough in $L^p(\Omega)$, we know that the model cannot at the same time meet the expected behaviour and still remain as small as we wanted outside Ω. This might be an indication that further experiments have to be performed around certain frequencies to determine wether this discrepancy is due to a loose description of the desired behaviour or to an overoptimistic estimation of the system's stability.
A dual problem has been studied by Krein and Nudel'man in [15] for $p = 2$: given a function $f \in L^2(\Omega)$ and $\varepsilon > 0$, they establish existence and uniqueness of a function in $\mathcal{H}^2$ which is closer than ε to f and which minimizes the $L^2(i\,\mathbb{R}_+ \setminus \Omega)$ (or the $L^2(i\,\mathbb{R}_+))$ norm. Although the above mentioned authors rely more on spectral theory than we do, the solution they derive for this dual problem is quite reminiscent of the approach in section 3 of the present paper. The main difference in their formulation is that it takes place in time domain and thus leads to integral rather than spectral equations involving some implicit parameter μ which is an analogue to the Lagrange parameter λ that we

introduce in section 3.
Of course, the above formulation is somewhat naive in practice. Indeed, one would certainly plug in further constraints like bounds on the derivative of g to prevent the model from oscillating too much as in [3], and modify the error criterion itself so as to weight frequencies or explicitly include the feedback law as in [8]. From the point of view of robust identification, a framework of which is proposed in [10] and [19], robustness with respect to bounded perturbations should be analyzed. Nevertheless, we shall stick to the simple formulation above except that we shall carry it over to the unit disk where it would correspond to parallel considerations for discrete time systems. By direct inspection or using conformal mapping (see e.g. [12, chap.8]), the results we prove have their counterparts in $\mathcal{H}^p$ but we shall not attempt to state them here.

The issue under investigation may be considered as a generalization to a closed subarc of the unit circle $\mathbb{T}$ of certain classical extremal problems in Hardy spaces H^p of the unit disk $\mathbb{D}$ which have already proved to be useful in control, especially the Nehari extension (see e.g. [6]). More precisely, if I is a subarc of $\mathbb{T}$, and $J = \mathbb{T} \setminus I$, the question we adress can be stated as follows :

For $p \geq 1$, let $f \in L^p(I)$ be given together with some positive M. Find some function $g \in H^p$ whose norm in $L^p(J)$ does not exceed M and which is as close as possible to f in the $L^p(I)$ metric under this constraint.

When $I = \mathbb{T}$, the constraint becomes void, and we recognize a classical dual extremal problem (see e.g. [5, chap.8], [7, chap.IV], [14, chap.VII]). When I is a strict subset of $\mathbb{T}$ and $p < \infty$, it is usually necessary to set $M < \infty$ for the problem to be well–posed. But it is not so when $p = \infty$, and the problem may still be nontrivially adressed in this case even if $M = +\infty$.

Finally, observe that if $f \in L^p(I)$ is *real*, meaning that the arc I is symmetric and $\overline{f(z)} = f(\bar{z})$ (this is the case for transfer functions), the H^p approximant g can be construed so as to be real since $\frac{1}{2}\left(g(z) + \overline{g(\bar{z})}\right)$ does the job as well. When g is unique, which will turn out to be the case when $p < \infty$, this actually proves that it is real.
The proofs in section 2 are omitted. Details will appear in [2].

2 Bounded extremal problems

Let us begin with the following basic result. Recall $C(I)$ is the space of continuous functions on I, $L^p(I)$ is the Lebesgue space of functions whose modulus to the power p is summable on I, H^p is the familiar Hardy space of the unit disk for $1 \leq p \leq \infty$, and $\mathcal{A}$ is the disk algebra.

Theorem 1 *Let $I \subset \mathbb{T}$ be a closed arc of circle of length $l(I) \in (0, 2\pi)$.*
(i) Let f be in $C(I)$. Then, for every $\varepsilon > 0$, there exists $\tilde{h}$ in $\mathcal{A}$ such that

$$\|f - \tilde{h}\|_{L^\infty(I)} \leq \varepsilon \,.$$

(ii) Let $1 \leq p < \infty$, let f be in $L^p(I)$. Then, for every $\varepsilon > 0$, there exists $\tilde{h}$ in H^p such that

$$\|f - \tilde{h}\|_{L^p(I)} \leq \varepsilon\,.$$

In general, however, ε cannot be taken to be 0.

Theorem 1 is an easy consequence of the Runge theorem [23, thm.13.9], a classical result in analytic function theory. It may also be deduced from deeper results in harmonic analysis like Levinson's theorem [24, III.2,thm.II] or, for $p \geq 2$, the Szegö theorem [7, IV,thm.3.1].
Statement (ii) asserts that traces of H^p functions are dense in $L^p(I)$ for $1 \leq p < \infty$. It is no longer so when $p = \infty$:

Proposition 1 *The trace of H^∞ is not dense in $L^\infty(I)$.*

In fact, it can be shown that Blaschke products whose zeroes have an accumulation point lying in the interior of I have an inverse (belonging to $L^\infty(I)$) which is at distance bigger than 1 to H^∞.
The next result shows that the approximants provided by theorem 1 are bound to have wild behaviour outside I. In the sequel, I will always denote a proper closed subarc of $\mathbb{T}$ as in theorem 1 while J will stand for the closure of $\mathbb{T} \setminus I$.

Proposition 2 *Let $1 \leq p \leq \infty$. Let f be in $L^p(I)$ and let $(g_n)_{n>0}$ be a sequence of H^p functions converging to f in $L^p(I)$. If f is not the trace of an H^p function, then $\lim_{n\to\infty} \|g_n\|_{L^p(J)} = \infty$.*

A consequence of proposition 2 is that the $L^p(J)$ norm of the approximant in theorem 1 goes to infinity as ε goes to 0, unless f is already the trace of some function in H^p (resp. $\mathcal{A}$) in which case the whole question is very trivial. It is therefore natural, and serves system–theoretic purposes as explained in the introduction, to bound the $L^p(J)$ norm of the approximant so as to end up with a well–posed problem :

Theorem 2 *Let f be in $L^p(I)$, for $1 \leq p \leq \infty$. For every M satisfying $0 \leq M < \infty$, there exists $g_0 \in H^p$, such that $\|g_0\|_{L^p(J)} \leq M$ and*

$$(1) \qquad \|f - g_0\|_{L^p(I)} = \min_{\substack{g \in H^p \\ \|g\|_{L^p(J)} \leq M}} \|f - g\|_{L^p(I)}\,.$$

When $1 \leq p < \infty$, such a g_0 is unique and satisfies $\|g_0\|_{L^p(J)} = M$ unless $f = g_0$ a.e. on I. When $p = \infty$, neither of these assertions need be true.

Remark 1 For the classical extremal problem with $p = \infty$, it is well known that uniqueness holds whenever f is continuous on $\mathbb{T}$ [7]. Moreover, g is known to belong to $\mathcal{A}$ if f is Dini–continuous. We know no analogue of these results in our case.

3 Characterization in the L^2 case

In this section, we recast the case $p = 2$ in an operator–theoretic framework. This will lead us, on two occasions, to endow $L^2(\mathbb{T})$ and H^2 with the *real* Hilbert space structure induced by the bilinear map $\mathrm{Re}\,(.,.)_{L^2(\mathbb{T})}$. We shall make this distinction explicit when needed.

For $0 < M < \infty$, let

$$\Gamma_M = \{g \in H^2\,,\ \|g\|_{L^2(J)} = M\}\,.$$

We first claim that Γ_M is a smooth submanifold of H^2. Indeed, define

$$\begin{array}{rlcl} \Phi: & H^2 & \longrightarrow & \mathbb{R} \\ & g & \longmapsto & \|g\|^2_{L^2(J)} \end{array}.$$

The map Φ is smooth and for any $g \in H^2$, $u \in H^2$,

$$D_g\,\Phi(u) = 2\,\mathrm{Re}\,(g,u)_{L^2(J)}\,.$$

It is plain that $D_g\,\Phi$ is surjective whenever $g \neq 0$. Being a closed subspace of the *real* Hilbert space H^2, $\mathrm{Ker}\,D_g\,\Phi$ splits in H^2 so that Φ is submersive on $H^2 \setminus \{0\}$ and in particular on $\Gamma_M = \Phi^{-1}(M^2)$ thereby establishing the claim (see e.g. [16, II,2,prop.2]). Moreover, for any $g \in \Gamma_M$ the tangent space to Γ_M at g is $T_M(g) = \mathrm{Ker}\,D_g\,\Phi$.

Let

$$\mathcal{C}_M = \{g|_I\,,\ g \in \Gamma_M\} = \{g|_I\,,\ g \in H^2\,,\ \|g\|_{L^2(J)} = M\}\,,$$

where $g|_I$ denotes restriction to I. It is easily shown that $\mathcal{C}_M$ is a closed subset of $L^2(I)$.

From now on, we assume that $f \in L^2(I) \setminus \mathcal{C}_M$ *.*

We denote by π the orthogonal projection onto the closed convex subset $\mathcal{C}_M$ of $L^2(I)$ and we have that $g_0|_I = \pi\,f$ is the unique solution of (1). Define $\xi\ :\ H^2 \to \mathbb{R}$ by

$$\xi(g) = \|f - g\|^2_{L^2(I)}\,.$$

ξ induces a smooth function on Γ_M and $g \in \Gamma_M$ is a critical point of ξ if and only if

(2) $\quad \mathrm{Re}\,(f - g, u)_{L^2(I)} = 0\,,\ \forall u \in T_M(g)\,.$

Proposition 3 *When $p = 2$, the best approximation g_0 to f in the sense of* (1) *is the unique critical point of ξ on Γ_M satisfying*

(3) $\quad \mathrm{Re}\,(f - g_0, g_0)_{L^2(I)} \geq 0\,.$

Proof : It is a well–known property of the projection (see e.g. [4, 2,I,prop.2]) that $g_0|_I = \pi\, f$ is characterized by

(4) $\quad \text{Re}\,(f-g_0, w-g_0)_{L^2(I)} \leq 0\ ,\ \forall w \in H^2$ such that $w|_I \in \mathcal{C}_M$.

Plugging $w = 0$ in (4) gives (3). On another hand, we know from theorem 2 that the constraint is saturated, i.e. $g_0 \in \Gamma_M$. Since it is the minimum of ξ on Γ_M, g_0 is a critical point.
Conversely, let g be a critical point of ξ on Γ_M satisfying (3). Take any w such that $w|_I \in \mathcal{C}_M$ and define $v \in H^2$ by

$$v = w - \frac{\text{Re}\,(w,g)_{L^2(J)}}{M^2}\, g\,.$$

We have $\text{Re}\,(v,g)_{L^2(J)} = 0$ so that $v \in T_M(g)$. Hence $\text{Re}\,(f-g,v)_{L^2(I)} = 0$, which implies

$$\text{Re}\,(f-g, w-g)_{L^2(I)} = \left[\frac{\text{Re}\,(w,g)_{L^2(J)}}{M^2} - 1\right] \text{Re}\,(f-g,g)_{L^2(I)}\,.$$

Now, $\text{Re}\,(w,g)_{L^2(J)} \leq M^2$ since $w|_I \in \mathcal{C}_M$. Together with (3) it implies (4) with $g_0 = g$. ■

Our next goal is to use proposition 3 to produce a more explicit formula for g_0. Set

$$\tilde{f} = \begin{cases} f & \text{on } I \\ 0 & \text{on } J \end{cases}\ .$$

Let P_{H^2} be the orthogonal projection from $L^2(\mathbb{T})$ onto H^2 and let further $\bar{H}_0^2$ denote the orthogonal complement to H^2 in $L^2(\mathbb{T})$. The latter is nothing but the subspace of the Hardy space of $\mathbb{C} \setminus \bar{\mathbb{D}}$ consisting of functions vanishing at infinity.
Since members $u \in H^2$ of $T_M(g_0)$ are characterized by the property :

$$\text{Re}\,(g_0,u)_{L^2(J)} = 0\,,$$

it follows from proposition 3 and (2) that g_0 satisfies

$$\text{Re}\,(\tilde{f}-g_0,u)_{L^2(\mathbb{T})} = 0\ ,\ \forall u \in T_M(g_0)\,,\ \text{and}\ \text{Re}\,(\tilde{f}-g_0,g_0)_{L^2(\mathbb{T})} \geq -M^2\,,$$

or, since u and g_0 belong to H^2,

(5) $\quad \text{Re}\,(P_{H^2}\tilde{f}-g_0,u)_{L^2(\mathbb{T})} = 0\ ,\ \forall u \in T_M(g_0)\,,$ and $\text{Re}\,(P_{H^2}\tilde{f}-g_0,g_0)_{L^2(\mathbb{T})} \geq -M^2\,.$

Introducing χ_J to be the characteristic function of the arc J, define $\phi\,:\ H^2 \to H^2$ to be the Toeplitz operator with symbol χ_J :

$$\phi(g) = P_{H^2}\,(\chi_J\, g)\,.$$

The obvious virtue of ϕ is that for all $g, h \in H^2$,

$$(\phi(g),h)_{L^2(\mathbb{T})} = (g,h)_{L^2(J)}\,.$$

In particular,

(6) $\quad \text{Re}\,(\phi(g_0),u)_{L^2(\mathbb{T})} = 0\,,\ \forall u \in T_M(g_0)\,.$

Comparing (6) and (5) shows that $\phi(g_0)$ and $P_{H^2}\,\tilde{f} - g_0$ are two vectors of the *real* Hilbert space H^2 both orthogonal to $T_M(g_0)$ whose real codimension is 1. Observe $\phi(g_0) \not\equiv 0$ because ϕ is injective since no non–zero function of $\bar{H}^2_0$ can be 0 on I. Hence, there exists $\lambda \in \mathbb{R}$ such that $\lambda\,\phi(g_0) = P_{H^2}\,\tilde{f} - g_0$, or equivalently such that

(7) $\quad g_0 + \lambda\,\phi(g_0) = P_{H^2}\,\tilde{f}\,.$

Moreover, from (7) :

$$(P_{H^2}\,\tilde{f} - g_0, g_0)_{L^2(\mathbb{T})} = \lambda\,(\phi(g_0), g_0)_{L^2(\mathbb{T})} = \lambda\,(g_0, g_0)_{L^2(J)} = \lambda\,M^2\,.$$

Therefore, in view of (5), $\lambda \geq -1$. Now, the function $\chi_J \in L^\infty(\mathbb{T})$ is real valued, piecewise constant, $\chi_J \geq 0$, [ess inf χ_J, ess supχ_J] $= [0,1]$, and $\|\chi_J\|_{L^\infty(\mathbb{T})} = 1$. Therefore, ϕ is a bounded self–adjoint positive operator with norm 1 and spectrum $\sigma(\phi) = [0,1]$ (see e.g. [21, chap.3]).
Thus, $I + \lambda\,\phi$ has a bounded inverse for $\lambda \in (-1, +\infty)$. Hence,

(8) $\quad g_0 = (1 + \lambda\,\phi)^{-1}\,P_{H^2}\,\tilde{f}\,,$

where $\lambda \in (-1, +\infty)$ is a Lagrange multiplier adjusted so that $\|g_0\|_{L^2(J)} = M$. The difficulty with this formula is of course its implicit character in λ, and it seems to be a hard point to express λ as a function of M. On another hand, it is easy to obtain a differential equation for λ. To this effect, we shall now start to consider M as a variable ranging over $[0, +\infty)$, so that g_0 and λ become functions of M. This also entails that *f is not the trace on I of any H^2 function so as to keep everything well–defined. This hypothesis will remain in force untill the end of the paper.*

First, we claim that

(9) $\quad (f, g_0)_{L^2(I)} = (1 + \lambda)\,M^2 + \|g_0\|^2_{L^2(I)}\,.$

To establish this, simply take the scalar product of (7) with g_0 on $\mathbb{T}$.

On another hand, since ϕ is a self–adjoint operator, it follows from (8) that

$$\begin{aligned} M^2 &= \|g_0\|^2_{L^2(J)} = (g_0, g_0)_{L^2(J)} = (\phi\,g_0, g_0)_{L^2(\mathbb{T})} \\ &= (\phi\,(1 + \lambda\,\phi)^{-2}\,P_{H^2}\,\tilde{f}, P_{H^2}\,\tilde{f})_{L^2(\mathbb{T})}\,. \end{aligned}$$

From the above equation observe that for $\lambda \in (-1, +\infty)$, the function $M^2(\lambda)$ is smooth . Differentiating with respect to λ leads to :

$$\frac{dM^2}{d\lambda} = -2\,(\phi^2\,(1 + \lambda\,\phi)^{-3}\,P_{H^2}\,\tilde{f}, P_{H^2}\,\tilde{f})_{L^2(\mathbb{T})} < 0\,,$$

where the inequality is a consequence of the spectral theorem (see e.g. [22]) and the injectivity of ϕ. Monotoneity implies that λ is in turn a smooth function of M^2 whose derivative is $\left(\frac{dM^2}{d\lambda}\right)^{-1}$. Therefore we get the

Proposition 4 *As a function of* M^2, λ *satisfies the differential equation :*

$$\frac{d\lambda}{dM^2} = \frac{-1}{2\,(\phi^2\,(1+\lambda\,\phi)^{-3}\,P_{H^2}\,\tilde{f}, P_{H^2}\,\tilde{f})_{L^2(\mathbb{T})}}\,,$$

with initial condition

$$\lambda(\|P_{H^2}\,\tilde{f}\|^2_{L^2(J)}) = 0\,.$$

This equation allows one in principle to compute λ from M, but its practical value is unclear at this point. A more amenable way to look at the problem is perhaps to choose λ as parameter and to estimate M, together with the $L^2(I)$ approximation error :

$$e(\lambda) = \|f - g_0\|^2_{L^2(I)}$$

as functions of λ. It is plain that e is increasing with λ, because e obviously increases when M decreases and we have seen that M decreases when λ increases. Moreover, as we have shown in theorem 1, e goes to 0 as M goes to ∞, and so, from (9), as λ goes to -1. These considerations suggest a tentative algorithm as follows :
(i) Take some $\lambda > -1$ but not too far from -1 and compute g_0 given by (8).
(ii) Compute M. If M is too big, increase λ, if not, decrease it. Go to (i).

An important issue is to estimate the behaviour of $e(\lambda)$ and $M(\lambda)$ as $\lambda \to -1$. While the authors do not know yet of asymptotic formulae, there is a simple result showing that $M(\lambda)$ cannot behave arbitrarily, an analogue of which can be found in [15]. On one hand, we have :

$$e(\lambda) = \|P_{\bar{H}^2_0}\,\tilde{f}\|^2_{L^2(\mathbb{T})} + \lambda^2\,(\phi^2\,(1+\lambda\,\phi)^{-2}\,P_{H^2}\,\tilde{f}, P_{H^2}\,\tilde{f})_{L^2(\mathbb{T})} - M^2\,,$$

and

$$\frac{de}{d\lambda} = 2\,(\lambda+1)\,(\phi^2\,(1+\lambda\,\phi)^{-3}\,P_{H^2}\,\tilde{f}, P_{H^2}\,\tilde{f})_{L^2(\mathbb{T})}\,,$$

so that

$$(10)\quad \frac{de}{d\lambda} = -(\lambda+1)\,\frac{dM^2}{d\lambda}\,.$$

On the other hand,

$$(11)\quad \lambda \to \infty \Rightarrow \begin{cases} g \to 0 & \text{by } (8) \\ e(\lambda) \to \|f\|^2_{L^2(I)} & \\ M(\lambda) \to 0 & \\ \lambda\,M^2 \to 0 & \text{by } (9)\,. \end{cases}$$

Moreover we also have from (9) that

$$(12)\quad \lim_{\lambda\to-1}\,(\lambda+1)\,M^2 = 0\,.$$

Now, integrating (10) by parts between any $\lambda_0, \lambda \in (-1, +\infty)$, and letting $\lambda_0 \to -1$, we obtain

$$(13)\quad e(\lambda) = -(\lambda+1)\,M^2(\lambda) + \int_{-1}^{\lambda} M^2(\tau)\,d\tau\,.$$

Letting $\lambda \to \infty$ in (13) and using (11) leads to the following result :

Proposition 5 *The function $M(\lambda)$ belongs to $L^2(-1,+\infty)$ and*

$$(14) \quad \int_{-1}^{+\infty} M^2(\lambda)\, d\lambda = \|f\|^2_{L^2(I)}\,.$$

In particular, the above proposition shows that M cannot increase too fast as λ approaches -1.

4 Conclusion

We studied a class of approximation problems that was reminiscent of classical (dual) extremal problems in Hardy spaces. These questions are attractive from the point of view of system identification because they predict in some sense the stability of underspecified linear models. They also offer some interesting theoretical features. However, a number of open questions still remain.

From the function theoretic viewpoint, the characterization of the $L^\infty(I)$ closure of H^∞ as well as the pending remark 1 are perhaps the major ones. On the computational side, the estimation of the convergence rate in the L^2 case is still to be established in order to implement an algorithm based on the procedure described here. In the L^∞ case, almost everything remains to be done. This bears particular significance if one observes that two main classes of transfer functions are of particular interest in automatic control, namely H^2 transfer functions and H^∞ ones. The subclass $\mathcal{A}$ should further be singled out since it more or less corresponds to compact Hankel operators which are the ones that lend themselves to rational approximation.

A constructive procedure in the H^p case, for $p < \infty$, also remains to be found and could probably be given along the same lines than what we did here for $p = 2$. This may asymptotically help study the H^∞ case.

From the point of view of system theory, it is certainly sound to study such problems in more restrictive functions spaces, for instance in Hardy–Sobolev spaces or in weighted L^p spaces. An account of the Hardy–Sobolev case of exponent 2 may be found in [3].

Acknowledgments. Laurent Baratchart and Juliette Leblond acknowledge helpful discussions with Jean–Pierre Kahane and Martin Zerner.

References

[1] Outils et modèles mathématiques pour l'automatique, l'analyse des systèmes et le traitement du signal, 1981. I.D. Landau ed.

[2] D. Alpay, L. Baratchart, and J. Leblond. Hardy approximation in L^p spaces of an arc. In preparation.

[3] L. Baratchart and J. Leblond. Identification harmonique et trace des classes de hardy sur un arc de cercle. In *Actes du Colloque en l'honneur du 60^e anniversaire du professeur Jean Céa, Sophia–Antipolis*. CEPADUES, avril 1992. A paraitre.

[4] B. Beauzamy. *Introduction to Banach spaces and their geometry.* Mathematics studies. North–Holland, 1985.

[5] P.L. Duren. *Theory of H^p functions.* Academic Press, 1970.

[6] B. Francis. *A course in H^∞ control theory.* Lectures notes in control and information sciences. Springer–Verlag, 1987.

[7] J.B. Garnett. *Bounded analytic functions.* Academic Press, 1981.

[8] M. Gevers. Connecting identification and robust control : a new challenge. In *9th IFAC symposium on identification, Budapest, vol. 1*, pages 1–10, 1991.

[9] G. Gu, P.P. Khargonekar, and Y. Li. Robust convergence of two–stage nonlinear algorithms for identification in $\mathcal{H}_\infty$. *Systems and Control Letters*, 18:253–263, 1992.

[10] A.J. Helmicki, C.A. Jacobson, and C.N. Nett. Control oriented system identification : a worst–case / deterministic approach in H_∞. In *Proceedings of the A.C.C. 1990*, pages 386–391.

[11] A.J. Helmicki, C.A. Jacobson, and C.N. Nett. Control oriented system identification : a worst–case / deterministic approach in H_∞. *IEEE Trans. Automat. Control*, 36(10):1163–1176, 1991.

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

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

[14] P. Koosis. *Introduction to H_p spaces.* Cambridge University Press, 1980.

[15] 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.

[16] S. Lang. *Differential manifolds.* Addison–Wesley, 1972.

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

[18] D.S. Lubinsky. Spurious poles in diagonal rational approximants. In A.A. Gonchar and E.B. Saff, editors, *Proceedings of the first conference US–USSR conference on approximation theory, Tampa, Florida, 1990*, pages 191–213. Springer–Verlag, 1992. To appear.

[19] J.R. Partington. Robust identification in H_∞. Technical report, University of Leeds, department of pure mathematics, 1990.

[20] T.J. Rivlin. *Approximation of functions.* Dover, 1981.

[21] M. Rosenblum and J. Rovnyak. *Hardy classes and operator theory.* Oxford University Press, 1985.

[22] W. Rudin. *Functional analysis*. Series in higher mathematics. Mc Graw Hill, 1973.

[23] W. Rudin. *Real and complex analysis*. Mc Graw Hill, 1982.

[24] L. Schwartz. *Etude des sommes d'exponentielles*. Hermann, 1959.

[25] E.D. Sontag. *Mathematical control theory*. Texts in applied mathematics. Springer–Verlag, 1990.

[26] W.M. Wonham. *Linear multivariable control : a geometric approach*. Applications of mathematics. Springer–Verlag, 1979.