Using H^2 norm to bound H^∞ norm from above on Real Rational Modules

Tzvetan Ivanov*, Brian D.O. Anderson⋆, P.-A. Absil*, Michel Gevers*

Abstract— **Various optimal control strategies exist in the literature. Prominent approaches are Robust Control and Linear Quadratic Regulators, the first one being related to the H^∞ norm of a system, the second one to the H^2 norm. In 1994, F. De Bruyne et al [1] showed that assuming knowledge of the poles of a transfer function one can derive upper bounds on the H^∞ norm as a constant multiple of its H^2 norm. We strengthen these results by providing tight upper bounds also for the case where the transfer functions are restricted to those having a real valued impulse response. Moreover the results are extended by studying spaces consisting of transfer functions with a common denominator polynomial. These spaces, called rational modules, have the feature that their analytic properties, captured in the integral kernel reproducing them, are accessible by means of purely algebraic techniques.**

Keywords: Robust Control, LQR, Confidence Region, Weighted H^2 norm, Supremum norm, H^∞ norm, Tight Bound, Real Rational Module, Christoffel-Darboux, Reproducing Kernel

I. Introduction

It is well known that norms induced by inner products, such as the H^2 norm, are important because they lend themselves to computations and geometric interpretations. However in many applications one is more interested in other norms like the supremum or H^∞ norm. For example in system identification for control, ellipsoids are the most natural shape of confidence regions from a statistician's perspective. This is because in this case well known results on asymptotic normality theory and Fisher information can be used to quantify the resulting error probability [2]. The shape and size of the ellipsoid represents the mismatch between the true system and the estimated one with respect to the H^2 norm weighted by the Signal-to-Noise ratio [3]. Yet from the robust control perspective, confidence regions having the shape of H^∞ balls are more interesting. A natural question in this context is how to find the largest ellipsoid which is contained in a given supremum norm ball. Another application is model order reduction. By choosing to minimize an H^2 norm approximation criterion one may use steepest descent and gradient methods. Moreover the resulting solution can be analyzed using techniques from differential geometry [4]. Having said this it is however also the case that it is physically more meaningful to minimize an H^∞ norm approximation criterion. Thus one is interested to quantify the deviation of the H^2 optimal approximation from an H^∞ optimal approximation.

The problem of bounding the H^∞ norm from above by a constant multiple of the H^2 norm has been first addressed in the engineering context in [1]. In that paper one derived results such as

$$|M(s)|^2 \leq \kappa(s) \cdot \|M\|_2^2, \quad \kappa(s) = \sum_{i=1}^{n} \frac{2 \cdot \operatorname{Re} a_i}{|s + a_i|^2}. \tag{1}$$

In the formula above M is the strictly proper transfer function of a stable continuous-time system ($s = \jmath\omega$, $\jmath^2 = -1$) of the form

$$M(s) = \frac{b_1}{s + a_1} + \cdots + \frac{b_n}{s + a_n}, \tag{2}$$

where the b_i's are arbitrary complex numbers and the a_i's with $\operatorname{Re}(a_i) > 0$ fix distinct pole locations in the left half plane. Analogous results have been derived in the discrete-time setting with $M(z)$ and $z = e^{\jmath\omega}$. Moreover $\|\cdot\|_\infty^2 \leq \|\kappa\|_\infty \|\cdot\|_2^2$, with κ defined in (1), has been recognized as the tight bound, i.e., the best upper bound which holds for all functions satisfying (2). However in [1] it has also been noted that the bound has its limitations because in the real rational case this bound may not be tight. By real rational we mean rational with real-valued impulse response. This corresponds to restricting the coefficients of the linear combination (2), i.e., the b_i's, to be such that the inverse Laplace transform of $M(s)$ is real.[1]

This paper shows that it is possible to close the gap, i.e., tighten the bound such that it can be achieved by a real rational transfer function, removing the restriction of distinct poles while still giving closed form expressions. Specifically the tight bound for the value of $|M(\jmath\omega_0)|^2$ as a multiple of the H^2 norm, say $\rho(\jmath\omega_0) \cdot \|M\|_2^2$, can be obtained for any fixed frequency ω_0. By supremizing over all frequencies we obtain $\|M\|_\infty^2 \leq \|\rho\|_\infty \|M\|_2^2$ which is tight in the sense that equality is achieved by some strictly proper $M(s)$ with a real valued impulse response and poles prescribed by (2). For continuous-time systems the bound is given by

$$\rho(s) = \frac{k(s, s)}{2} + \frac{|k(s, -s)|}{2}, \tag{3}$$

where $k(s, w)$ is the integral kernel reproducing the space of functions defined by (2). We provide an analogous result

†This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.

*Tzvetan Ivanov, P.-A. Absil and Michel Gevers are with the Center for Systems Engineering and Applied Mechanics (CESAME) Universite Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. (http://www.inma.ucl.ac.be/˜{ivanov,gevers,absil}

⋆Brian D.O. Anderson is with National ICT Australia and the Australian National University Research School of Information Sciences & Engineering, Canberra, ACT 2601, Australia (http://users.rsise.anu.edu.au/˜briandoa)

[1] A complex coefficient b_i then necessarily corresponds to a complex pole a_i and those come in conjugate pairs in the sense that $(b_j, a_j) = (b_i^*, a_i^*)$ for some index $j = 1, \ldots, n$.

ISBN 978-963-311-369-1

for discrete-time with s replaced by z and $-s$ replaced by z^{-1}. From this point of view the older bound κ for complex rational functions is given by $\kappa(s) = k(s, s)$.

Reproducing kernels (RK) have been extensively studied in the mathematical literature, see e.g., [5] for an overview. Linking our results to RK theory is of interest because it shows that our results are applicable to any 2-norm and weighted H^2 norms in particular.[2] If, however, the function space carries additional structure, such as symmetry relations with respect to the domain where the functions are defined, more can be said. Specifically, for the space of all strictly proper rational functions with common denominator q, which we refer to as a rational module and denote by X^q, the reproducing kernel takes a particularly simple form since X^q is a coinvariant subspace of H^2, see e.g. [6], [7]. With this background, we shall show that the bounds κ and ρ defined in (1) and (3), respectively, can be expressed in terms of the coefficients of the constant denominator term q given by, e.g., $q(s) = \prod_{i=1}^{n}(s + a_i)$ for the space defined by (2), via[3]

$$\kappa(s) = \frac{q(s)'}{q(s)} - \frac{q(-s)'}{q(-s)}, \tag{4a}$$

$$\rho(s) = \left|\frac{q(-s)^2}{s \cdot q(s)^2} - \frac{1}{s}\right| + \frac{\kappa(s)}{2} \tag{4b}$$

with similar results for discrete-time. Due to space limitations we just state and do not demonstrate that the ideas derived with this machinery can be seamlessly generalized to the case of $\mathbb{C}^n$ or $\mathbb{C}^{n\times m}$ instead of $\mathbb{C}$-valued functions. A special case of this has been studied in [8].

The paper is structured as follows. In Section II we study the bound for general linear subspaces over the reals whose elements are complex valued functions. In Section III we turn to real rational complex valued functions whose domain is the imaginary axis for continuous-time systems and the unit circle for discrete-time systems. In Section IV we specialize to real rational modules. We give conditions for the complex and real bound to coincide in Section V, and several examples illustrating this in Section VI. We conclude in Section VII.

II. Real Linear Subspaces of $\mathbb{C}$-valued functions

At this point of the paper we do not fix the domain Ω of the functions to be studied nor do we fix a particular 2-norm $\|\cdot\|_2$. This is indeed important since both quantities depend on the application. For example in control theory discrete time systems are naturally modeled as functions on the unit circle whereas continuous time systems live on the imaginary axis. In both cases we do not need to restrict to standard H^2 norms but may also choose to include weights.

Let Ω denote an abstract set. Consider a finitely generated linear space X over the reals consisting of bounded complex valued functions $f : \Omega \to \mathbb{C}$ equipped with an inner product $(\cdot, \cdot)$ which is $\mathbb{R}$-linear in both arguments. We emphasize that due to this assumption X is only closed under linear combinations with real coefficients, e.g., multiplying an element in X by the imaginary unit $\jmath$ may result in a function which is not in X. In fact from now on we make the stronger assumption that $\mathrm{X} \cap \jmath\mathrm{X} = \{0\}$, i.e., the only element in X which when multiplied by $\jmath$ remains in X is the zero vector. We will embed the $\mathbb{R}$-linear space X in the smallest linear space $\mathbf{X}$ over $\mathbb{C}$ containing it. For this let

$$\mathbf{X} = {}^{\mathrm{c}}\mathrm{X}, \text{ where } {}^{\mathrm{c}}\mathrm{X} = \mathrm{X} + \jmath\mathrm{X} \tag{5}$$

denotes the complexification of X. Any $f \in \mathbf{X}$ then has a unique representation as $f = f_1 + \jmath f_\jmath$ with $f_1, f_\jmath \in \mathrm{X}$. It is important not to confuse $f_1(z)$ with $\mathrm{Re}(f(z))$.

The evaluation of f at $w \in \Omega$, i.e., the map ev_w given by $\mathbf{X} \to \mathbb{C}, f \mapsto \mathrm{ev}_w(f) = f(w)$, is then a linear functional which makes it easy to study, as opposed to the evaluation restricted to X. The introduction of a complex valued inner product $\langle\cdot,\cdot\rangle$ on $\mathbf{X}$ allows us to represent linear functionals by vectors. In order for the complex inner product to extend the initial real inner product we define

$$\langle f, g\rangle = (f, g) + \jmath(\jmath f, -g), \tag{6}$$

where we also extended the real inner product $(\cdot,\cdot)$ by

$$(f, g) = (f_1, g_1) + (f_\jmath, g_\jmath), \tag{7}$$

for all $f, g \in \mathbf{X}$.

On $\mathbf{X}$ there exist now two natural norms: the 2-norm $\|\cdot\|_2$ induced by $\langle\cdot,\cdot\rangle$ and the sup norm $\|\cdot\|_\infty$ defined by

$$\|f\|_\infty = \sup\{|f(w)| \mid w \in \Omega\}. \tag{8}$$

In order to link these two norms on $\mathbf{X}$ we need the notion of a reproducing kernel k for $\mathbf{X}$. For this let $\{b_i\}_{i=1}^n$ denote an orthonormal basis (ONB) of $\mathbf{X}$ w.r.t. the complex inner product $\langle\cdot,\cdot\rangle$ and define $k : \Omega \times \Omega \to \mathbb{C}$ via

$$k(z, w) = \sum_{i=1}^{n} b_i(z) b_i^*(w). \tag{9}$$

Let $k_w(z) = k(z, w)$, and think of $k_w \in \mathbf{X}$ as a function of z, then by the Riesz-Representation theorem for Hilbert spaces k is uniquely determined by its properties 1) $k_w \in \mathbf{X}$ and 2) $f(w) = \langle f, k_w\rangle$ which hold for all $f \in \mathbf{X}$ and $w \in \Omega$. Note that k satisfies $k(w, z) = k(z, w)^*$.

Definition 1 *Let $\{b_i\}_{i=1}^n$ denote an ONB of $(\mathbf{X}, \langle\cdot,\cdot\rangle)$. The two variable function $k(z, w)$ given by (9) does not depend on the particular choice of ONB and is called the reproducing kernel of $\mathbf{X}$ w.r.t. $\langle\cdot,\cdot\rangle$. Moreover we denote the one-variable function $k(\cdot, w)$ by k_w.*

The statement of Theorem 2 is the abstract version of concrete inequalities such as (1) found in [1].

Theorem 2 *Let k denote the reproducing kernel of $(\mathbf{X}, \langle\cdot,\cdot\rangle)$ and $\kappa(w) = k(w, w)$. For all $w \in \Omega$, $f \in \mathbf{X}$, there holds $|f(w)|^2 \le \kappa(w)\|f\|_2^2$ and this bound is tight. In particular*

$$\|\cdot\|_\infty^2 \le \|\kappa\|_\infty \cdot \|\cdot\|_2^2, \tag{10}$$

is a tight bound on $\mathbf{X}$.

[2] For the general theory there is even no need to impose a restriction on the domain where the functions are defined or to restrict them to have a particular form.

[3] Here $q(s)'$ denotes the usual derivative of q with respect to s, e.g., $q(s)' = ns^{n-1}$ for $q(s) = s^n$.

Proof: The Cauchy-Bunyakovsky-Schwarz inequality

$$|f(w)|^2 \le \|k_w\|_2^2 \, \|f\|_2^2, \tag{11}$$

is tight since it becomes an equality for $f = k_w \in \mathbf{X}$. Utilizing $\langle k_w, k_w\rangle = k(w,w)$ we obtain

$$\|f\|_\infty^2 \le \|\kappa\|_\infty \, \|f\|_2^2, \tag{12}$$

which is obviously tight on $\mathbf{X}$ because (11) was tight. □

We now address the question how to obtain a tight bound on the subspace $\mathrm{X} \subseteq \mathbf{X}$. We first split k_w into $k_w = k_{w,1} + \jmath k_{w,\jmath}$ with $k_{w,1}, k_{w,\jmath} \in \mathrm{X}$ and note that for all $f \in \mathbf{X}$

$$\operatorname{Re} f(w) = (f, k_{w,1}) \quad \text{and} \quad \operatorname{Im} f(w) = (f, k_{w,\jmath}), \tag{13}$$

i.e., they are Riesz-representation of the real part and imaginary part of the evaluation at w. In general $k_{w,\jmath} \neq 0$ and the preceding proof does not show that (11) is achieved on X but merely on $\mathbf{X}$. In fact we will see that (11) is not tight on X in general.

The following result in Theorem 3 demonstrates that it is possible to close the gap between the optimal bound, call it $\rho(w)$, and $k(w,w)$. Despite its abstract nature Theorem 3 illuminates the main idea and leads to practically important statements such as Theorem 10 in Section IV.

Theorem 3 *Let k denote the reproducing kernel of $\mathbf{X}$ and $k_w = k_{w,1} + \jmath k_{w,\jmath}$ with $k_{w,1}, k_{w,\jmath} \in \mathrm{X}$ for all $w \in \Omega$. Define $\rho : \Omega \to \mathbf{R}$ via*

$$\rho(w) = \tfrac{k(w,w) + \left| \|k_{w,1}\|_2^2 - \|k_{w,\jmath}\|_2^2 - \jmath\, 2 (k_{w,1}, k_{w,\jmath}) \right|}{2}. \tag{14}$$

Then for all $w \in \Omega, f \in \mathrm{X}$ there holds $|f(w)|^2 \le \rho(w)\|f\|_2^2$ and this bound is tight. In particular

$$\|\cdot\|_\infty^2 \le \|\rho\|_\infty \cdot \|\cdot\|_2^2, \tag{15}$$

is a tight bound on X.

Proof: We need to check that $\rho(w) = \sigma$ with

$$\sigma = \sup\{(f, k_{w,1})^2 + (f, k_{w,\jmath})^2 \mid \|f\|_2^2 = 1, f \in \mathrm{X}\}. \tag{16}$$

Let $G \in \mathbb{R}^{2\times 2}$ be the Gramian defined via

$$G = \begin{bmatrix} (k_{w,1}, k_{w,1}) & (k_{w,1}, k_{w,\jmath}) \\ (k_{w,\jmath}, k_{w,1}) & (k_{w,\jmath}, k_{w,\jmath}) \end{bmatrix},$$

and simple calculation shows that the maximum eigenvalue of G is given by $\rho(w)$. If $k_{w,1}$ and $k_{w,\jmath}$ are $\mathbb{R}$-linearly dependent it follows that G is singular and $\rho(w) = k(w,w) = \operatorname{trace}(G)$ and this bound becomes an equality for $f = k_{w,1}$ or $f = k_{w,\jmath}$. If $k_{w,1}$ and $k_{w,\jmath}$ are $\mathbb{R}$-linearly independent then G is non-singular since

$$\det(G) = \|k_{w,1}\|_2^2 \|k_{w,\jmath}\|_2^2 - (k_{w,1}, k_{w,\jmath})^2 > 0,$$

by the Cauchy-Bunyakovsky-Schwarz inequality. Supremizing over X according to (16) and supremizing over X_w, where X_w denotes the 2-dimensional subspace generated by $k_{w,1}, k_{w,\jmath} \in \mathrm{X}$, yields the same value σ. Let

$$x^{\mathrm{T}} = [(f, k_{w,1}), (f, k_{w,\jmath})] \in \mathbb{R}^{1\times 2} \quad \text{with} \quad f \in \mathrm{X}_w,$$

denote the coordinates of f in the $\{k_{w,1}, k_{w,\jmath}\}$ basis. Then

$$\begin{aligned} \sigma &= \sup\{x^{\mathrm{T}} x \mid x \in \mathbb{R}^2, x^{\mathrm{T}} G^{-1} x = 1\} \\ &= \sup\{y^{\mathrm{T}} G y \mid y \in \mathbb{R}^2, y^{\mathrm{T}} y = 1\} = \lambda_{\max}(G). \end{aligned}$$

The second part of the theorem, i.e., (15), follows by supremizing $|f(w)|^2 \le \rho(w)\|f\|_2^2$ over $w \in \Omega$. □

Remark 4 The proof of Theorem 3 together with

$$\lambda_1(G) + \lambda_2(G) = \|k_{w,1}\|_2^2 + \|k_{w,\jmath}\|_2^2 = \langle k_w, k_w \rangle, \tag{17}$$

and $\lambda_{\max} > (\lambda_1 + \lambda_2)/2$ reveals that

$$\kappa(w)/2 \le \rho(w) \le \kappa(w). \tag{18}$$

In other words the bound in the real case is at most two times smaller than the bound for the complexification.

III. Real Rational Transfer Functions of Linear Time Invariant Systems

In this section we first introduce the real rational subspace $\mathrm{R}L_{\mathrm{d}}^2$ of L_{d}^2 for discrete-time, and $\mathrm{R}L_{\mathrm{c}}^2$ of L_{c}^2 for continuous-time. These spaces have natural symmetry relations which induce symmetry relations on the reproducing kernels. To be precise, we establish that the correction term given by

$$\left| \|k_{w,1}\|_2^2 - \|k_{w,\jmath}\|_2^2 - \jmath (k_{w,1}, k_{w,\jmath}) \right|, \tag{19}$$

in (14) in Theorem 3 is given by

$$|k(w, w^{-1})| \quad \text{and} \quad |k(w, -w)|, \tag{20}$$

for discrete-time and continuous-time, respectively.

A. Discrete Time

Let $L_{\mathrm{d}}^2 = L^2(\mathbb{T}, \mathbb{C})$ be the space of all complex valued functions on the unit circle $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ with

$$\|f\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\mathrm{e}^{\jmath\omega})|^2 \, \mathrm{d}\omega < \infty. \tag{21}$$

The starting point for an algebraic theory is the real rational subspace and its complexification $\mathrm{R}L_{\mathrm{d}}^2 \subseteq {}^{\mathrm{c}}\mathrm{R}L_{\mathrm{d}}^2 \subseteq L_{\mathrm{d}}^2$

$$\mathrm{R}L_{\mathrm{d}}^2 = \{f \in \mathbb{R}(z) \mid f \text{ has no pole in } \mathbb{T}\} \tag{22a}$$

$${}^{\mathrm{c}}\mathrm{R}L_{\mathrm{d}}^2 = \{f \in \mathbb{C}(z) \mid f \text{ has no pole in } \mathbb{T}\}, \tag{22b}$$

In the following elementary result f^* denotes the complex conjugate, i.e., $f^*(z) = \overline{f(\bar{z})}$. The proof is immediate.

Lemma 5 *Let $f \in {}^{\mathrm{c}}\mathrm{R}L_{\mathrm{d}}^2$ and*

$$f_1(z) = \frac{f(z) + f^*(z^{-1})}{2}, \; f_\jmath(z) = \frac{f(z) - f^*(z^{-1})}{2\jmath}, \tag{23}$$

then $f = f_1 + \jmath f_\jmath$ with $f_1, f_\jmath \in \mathrm{R}L_{\mathrm{d}}^2$. In particular the following three statements are equivalent: 1) $f \in \mathrm{R}L_{\mathrm{d}}^2$, 2) $f^* \in \mathrm{R}L_{\mathrm{d}}^2$ and 3) $f^*(z^{-1}) = f(z)$.

Theorem 6 *Let $k : \mathbb{T} \times \mathbb{T} \to \mathbb{C}$ be the reproducing kernel of the complexification of a finitely generated subspace $\mathrm{X} \subseteq \mathrm{R}L_{\mathrm{d}}^2$. Then*

$$k(w, w^{-1})/2 = \|k_{w,1}\|_2^2 - \|k_{w,\jmath}\|_2^2 - \jmath\, 2 (k_{w,1}, k_{w,\jmath}) \tag{24}$$

and $k(z,w)$ possesses the properties:

1) $k(z,w) = k(w^{-1}, z^{-1})$,
2) $k(w,w) = k(w^{-1}, w^{-1})$,
3) $|k(w,w^{-1})| \le k(w,w)$,

with equality in 3) if and only if $k_w = k_{w^{-1}}$.

Proof: Let $\{b_i\}_{i=1}^n$ denote a basis of X. Then, due to the equivalence of 2) and 3) in Lemma 5, we have $b_i^* \in \mathrm{R}L_\mathrm{d}^2$. This implies, again by Lemma 5, that

$$\begin{aligned} k(w^{-1}, z^{-1}) &= \sum b_i(w^{-1})\, b_i^*(z^{-1}) \\ &= \sum b_i^*(z^{-1})\, b_i^{**}(w^{-1}) \\ &= \sum b_i(z)\, b_i^*(w) = k(z,w), \end{aligned}$$

which proves 1) and 2).

From the Cauchy-Bunyakovsky-Schwarz inequality, it follows that $|k(w,w^{-1})|^2$ is bounded from above by

$$\begin{aligned} |\langle k_{w^{-1}}, k_w\rangle|^2 &\le \langle k_{w^{-1}}, k_{w^{-1}}\rangle \langle k_w, k_w\rangle \\ &= k(w^{-1}, w^{-1}) k(w,w) = k(w,w)^2, \end{aligned}$$

with equality if and only if $k_w = k_{w^{-1}}$. Thus we have checked 3).

Let $u = k_{w,1}$ and $v = k_{w,\jmath}$; then $k_{w^{-1}} = u - \jmath v$ since

$$\begin{aligned} 2\,(k_{w^{-1},1})(z) &= k(z, w^{-1}) + k^*(z^{-1}, w^{-1}) \\ &= k^*(w^{-1}, z) + k(w^{-1}, z^{-1}) \\ &= k^*(z^{-1}, w) + k(z,w) = 2\,u(z), \end{aligned}$$

and similarly $k_{w^{-1},\jmath} = -v$. From this it follows that

$$\begin{aligned} k(w,w^{-1}) &= \langle u - \jmath v, u + \jmath v\rangle \\ &= (u,u) - (v,v) - \jmath(v,u) - \jmath(u,v) \\ &= \|u\|_2^2 - \|v\|_2^2 - \jmath\, 2(u,v). \end{aligned}$$

which proves (24). □

B. Continuous Time

Let $L_\mathrm{c}^2 = L^2(\jmath\mathbb{R}, \mathbb{C})$ be the space of all complex valued functions on the imaginary axis with

$$\|f\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |f(\jmath\omega)|^2\, \mathrm{d}\omega < \infty. \tag{25}$$

We define the real rational subspace $\mathrm{R}L_\mathrm{c}^2$ and its complexification ${}^\mathrm{c}\mathrm{R}L_\mathrm{c}^2 \subseteq L_\mathrm{c}^2$:

$$\mathrm{R}L_\mathrm{c}^2 = \{f \in \mathbb{R}(z) \mid f \text{ s.p., no pole in } \jmath\mathbb{R}\} \tag{26a}$$
$${}^\mathrm{c}\mathrm{R}L_\mathrm{c}^2 = \{f \in \mathbb{C}(z) \mid f \text{ s.p., no pole in } \jmath\mathbb{R}\}, \tag{26b}$$

with s.p. meaning strictly proper.

Theorem 7 is the continuous-time version of Theorem 6. The proof is completely analogous and therefore skipped.

Theorem 7 *Let* $k : \jmath\mathbb{R}\times\jmath\mathbb{R} \to \mathbb{C}$ *be the kernel which reproduces the complexification of a finitely generated subspace* $\mathrm{X} \subseteq \mathrm{R}L_\mathrm{c}^2$*. Then*

$$k(w,-w)/2 = \|k_{w,1}\|_2^2 - \|k_{w,\jmath}\|_2^2 - \jmath\, 2(k_{w,1}, k_{w,\jmath}) \tag{27}$$

and $k(s,w)$ *possesses the properties:*

1) $k(s,w) = k(-w,-s)$,
2) $k(w,w) = k(-w,-w)$,
3) $|k(w,-w)| \le k(w,w)$,

with equality in 3) if and only if $k_w = k_{-w}$.

IV. The Christoffel-Darboux Kernel of a Real Rational Module

So far we have seen that the bound we are seeking, i.e., $\rho(w)$ as defined in Theorem 3 is given by

$$\frac{k(w,w) + |k(w,w^{-1})|}{2} \quad \text{and} \quad \frac{k(w,w) + |k(w,-w)|}{2},$$

in discrete-time and continuous-time, respectively. In this section we want to redeem the promise to provide closed form formulas for $k(z,w)$ and the related quantities $\kappa(w)$ and $\rho(w)$. We therefore restrict our attention to the case where the subspace $\mathrm{X} \subseteq \mathrm{R}L_\mathrm{c}^2$ (or $\mathrm{X} \subseteq \mathrm{R}L_\mathrm{d}^2$), which was general up to now, turns out to be a real rational module.

In the following we treat the continuous and discrete-time case in parallel in order to emphasize that they possess the same structural properties. We call a polynomial $q \in \mathbb{R}[s]$ *c*-stable (resp. $q \in \mathbb{R}[z]$ *d*-stable) if $q(a) = 0$ implies $\mathrm{Re}\, a < 0$ (resp. $|a| < 1$). We define the real rational Hardy spaces as subspaces of $\mathrm{R}L_\mathrm{c}^2$ and $\mathrm{R}L_\mathrm{d}^2$, respectively

$$\mathrm{R}H_\mathrm{c}^2 = \{f \mid f = p/q \text{ strictly proper}, q \text{ is } c\text{-stable}\}, \tag{28a}$$
$$\mathrm{R}H_\mathrm{d}^2 = \{f \mid f = p/q \text{ strictly proper}, q \text{ is } d\text{-stable}\}. \tag{28b}$$

We now make the assumption that we have a priori knowledge about the poles $\{a_i\}$ of the system. In other words we concentrate on systems with a fixed denominator $q = \prod(x - a_i)$. Since we are interested in the real rational case we restrict to $q \in \mathbb{R}[x]$. The associated polynomial module is given by $\mathrm{X}_q = \{p \in \mathbb{R}[x], \deg(p) < \deg(q)\}$ and the associated rational module is

$$\mathrm{X}^q = \left\{\frac{p}{q} \in \mathbb{R}(x) : p \in \mathrm{X}_q\right\}. \tag{29}$$

All spaces come together with the corresponding complexifications $\mathbf{X}_q = \mathrm{X}_q + \jmath\mathrm{X}_q$, $\mathbf{X}^q = \mathrm{X}^q + \jmath\mathrm{X}^q$. Then $\mathrm{X}^q \subseteq \mathrm{R}H_\mathrm{c}^2$ and $\mathrm{X}^q \subseteq \mathrm{R}H_\mathrm{d}^2$ if $q \in \mathbb{R}[s]$ is *c*-stable and $q \in \mathbb{R}[z]$ is *d*-stable, respectively. Then Beurling's theorem on invariant subspaces (cf. [6], [7]) states that $\mathbf{X}^q$ is *coinvariant*, i.e.,

$${}^\mathrm{c}\mathrm{R}H_\mathrm{c}^2 \ominus \mathbf{X}^q = \frac{q^\star}{q}{}^\mathrm{c}\mathrm{R}H_\mathrm{c}^2, \quad {}^\mathrm{c}\mathrm{R}H_\mathrm{d}^2 \ominus \mathbf{X}^q = \frac{q^\star}{q}{}^\mathrm{c}\mathrm{R}H_\mathrm{d}^2, \tag{30}$$

respectively, where the para-adjoint $q^\star$ is given by

$$q^\star(s) = q(-s), \quad \text{and} \quad q^\star(z) = z^n q(z^{-1}), \tag{31}$$

if q is *c*-stable and *d*-stable, respectively, and $n = \deg(q)$. The importance of (30) is the corollary

$$\mathbf{X}^{pq} \ominus \mathbf{X}^q = \frac{q^\star}{q}\mathbf{X}^p, \tag{32}$$

whenever p, q are two *c*-stable or *d*-stable polynomials (cf. Corollary 5 in [3]). It is (32) that enables us to derive an explicit form for the reproducing kernel of $\mathbf{X}^q$ in discrete-time in Theorem 8 and continuous-time in Theorem 9. Due to the recursive structure of their computation these kernels are special cases of so called Christoffel-Darboux kernels [9].

Theorem 8 *Let $q = \prod_{i=1}^n (z - a_i) \in \mathbb{R}[z]$, d-stable, and $k : \mathbb{T} \times \mathbb{T} \to \mathbb{C}$ be defined via*

$$k(z,w) = \frac{q^\star(w)q(z) - q(w)q^\star(z)}{1 - \bar{w}z} \cdot \frac{-1}{q^\star(w)q(z)}, \quad (33)$$

for $z \neq w$ and

$$k(z,z) = z\left(\frac{q'(z)}{q(z)} - \frac{q^{\star\prime}(z)}{q^\star(z)}\right) = \sum_{i=1}^n \frac{1 - |a_i|^2}{|z - a_i|^2}, \quad (34)$$

for $z = w$. Then k is the reproducing kernel of $\mathbf{X}^q$.

Proof: An inductive proof of (33) can be found in [10, Lemma 2.1, (13)] and (34) readily follows from the fact that $k(z,w)$ is continuous and application of l'Hospital's rule to calculate $k(w,w)$ via $\lim_{z \to w} k(z,w)$. □

Theorem 9 *Let $q = \prod_{i=1}^n (s - a_i) \in \mathbb{R}[s]$, c-stable, and $k : \jmath\mathbb{R} \times \jmath\mathbb{R} \to \mathbb{C}$ be defined via*

$$k(s,w) = \frac{q^\star(w)q(s) - q(w)q^\star(s)}{w - s} \cdot \frac{-1}{q^\star(w)q(s)}, \quad (35)$$

for $s \neq w$ and

$$k(s,s) = \frac{q'(s)}{q(s)} - \frac{q^{\star\prime}(s)}{q^\star(s)} = -\sum_{i=1}^n \frac{2\operatorname{Re} a_i}{|s - a_i|^2}, \quad (36)$$

for $s = w$. Then k is the reproducing kernel of $\mathbf{X}^q$.

Proof: An inductive proof of (35) can be found in [11, Lemma 5] and the rest follows as in the proof of Theorem 8. □

We summarize the results in Theorem 10, which is Theorem 3 specialized to rational modules, using Theorem 7, 9 for continuous-time and Theorem 6, 8 for discrete-time.[4]

Theorem 10 *If $\mathbf{X} = \mathrm{X} + \jmath\mathrm{X}$ and $\mathrm{X} = \mathrm{X}^q$ (equipped with L^2_c norm) for some c-stable $q \in \mathbb{R}[s]$, $q = \prod(s - a_i)$, then Theorem 2 and Theorem 3 hold with*

$$\kappa(s) = \frac{q'(s)}{q(s)} - \frac{q^{\star\prime}(s)}{q^\star(s)} = -\sum_{i=1}^n \frac{2\operatorname{Re} a_i}{|s - a_i|^2}, \quad (37a)$$

$$\rho(s) = \frac{1}{2}\left|\frac{1 - (q(s)^{-1}q(-s))^2}{2s}\right| + \frac{\kappa(s)}{2} \quad (37b)$$

Similarly, in the discrete-time case, Theorem 2 and Theorem 3 hold with

$$\kappa(z) = \frac{zq'(z)}{q(z)} - \frac{zq^{\star\prime}(z)}{q^\star(z)} = \sum_{i=1}^n \frac{1 - |a_i|^2}{|z - a_i|^2}, \quad (38a)$$

$$\rho(z) = \frac{1}{2}\left|\frac{(z^n q(z)^{-1}q(z^{-1}))^2 - 1}{1 - z^2}\right| + \frac{\kappa(z)}{2} \quad (38b)$$

[4]Our results in (37a, 38a) correspond to the results of De Bruyne et al (2.4, 3.4) as found in [1] since $\beta_{ij} = \langle 1/(s - a_j), 1/(s - a_i)\rangle$ implies

$$k(s,w) = [\tfrac{1}{-w-a_1}, \cdots, \tfrac{1}{-w-a_n}] \begin{bmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{s-a_1} \\ \vdots \\ \frac{1}{s-a_n} \end{bmatrix}.$$

To see this, note that for $q = \prod(s - a_i)$ $(a_i \neq a_j)$ an ONB of X^q is given by $\{\sum_{ij} \frac{g_{ij}}{s - a_j}, i = 1, \ldots, n\}$ with $g = \beta^{-1/2}$ and thus k can be computed from (9) and analogously for $k(z,w)$.

if $\mathbf{X} = \mathrm{X} + \jmath\mathrm{X}$, and $\mathrm{X} = \mathrm{X}^q$ (equipped with L^2_d norm) for some d-stable $q = \prod(z - a_i) \in \mathbb{R}[z]$.

V. On Conditions for bounds on a real rational module and its complexification to coincide

In this section we examine when the tight bound $\|\rho\|_\infty$ on a real rational module X^q coincides with the tight bound $\|\kappa\|_\infty$ of its complexification $\mathbf{X}^q$. Recall that κ and ρ are defined in (37), (38) for continuous- and discrete-time respectively.

Lemma 11 *Let $q \in \mathbb{R}[z]$ be d-stable with $\deg(q) > 0$ and k denote the reproducing kernel of $\mathbf{X}^q$. Then $\{k(z, \cdot) \mid z \in \mathbb{T}\}$ separates points in $\mathbb{T}$. That is for all $w_i \in \mathbb{T}$ with $w_1 \neq w_2$ there exists $z \in \mathbb{T}$ with $k(z, w_1) \neq k(z, w_2)$. The same holds mutatis mutandis for $q \in \mathbb{R}[s]$ being c-stable.*

Proof: Since $\deg(q) > 0$ there exists some $a \in \mathbb{C}$ such that $q(a) = 0$ and thus, by partial fraction expansion, $1/(z - a) \in \mathbf{X}^q$. Since $1/(w_1 - a) = 1/(w_2 - a)$ implies $w_1 = w_2$ we can conclude that $k_{w_1} \neq k_{w_2}$ since they take different values on $1/(z - a) \in \mathbf{X}^q$. □

Using Lemma 11 and property 3) of Theorem 6 and Theorem 7, respectively, we obtain a necessary and sufficient condition for $\|\kappa\|_\infty = \|\rho\|_\infty$ in the form of Corollary 12 and an easy sufficient condition given by Corollary 13.[5]

Corollary 12 *Let $q \in \mathbb{R}[z]$ be d-stable; then $\|\kappa\|_\infty = \|\rho\|_\infty$ if and only if $\|\kappa\|_\infty \in \{\kappa(-1), \kappa(1)\}$. Similarly for $q \in \mathbb{R}[s]$ being c-stable, $\|\kappa\|_\infty = \|\rho\|_\infty$ iff $\|\kappa\|_\infty = \kappa(0)$.*

Corollary 13 *Let $q = \prod(x - a_i) \in \mathbb{R}[x]$ be d-stable $(x = z)$ or c-stable $(x = s)$. Then the condition $a_i = \operatorname{Re} a_i$ for all i is sufficient (but not necessary) for $\|\kappa\|_\infty = \|\rho\|_\infty$.*

Proof: Let $q \in \mathbb{R}[s]$ be c-stable. Since all a_i are assumed to be real, $|s - a_i|^2$ is minimized by $s = 0$. This is sufficient for $\|\kappa\|_\infty = \|\rho\|_\infty$ since $\|\kappa\|_\infty = \kappa(0)$. To check the claim for discrete-time, let $q \in \mathbb{R}[z]$ be d-stable. Then κ is a convex function of $\operatorname{Re} z$ on $\{z \in \mathbb{T} \mid \operatorname{Re}(z) > 0\}$ since each of its summands

$$\frac{1 - |a_i|^2}{|z - a_i|^2} = \frac{1 - a_i^2}{(\operatorname{Re} z - a_i)^2},$$

is convex in that sense. So κ attains its maximum on $\{-1, 1\}$, i.e, $\|\kappa\|_\infty \in \{\kappa(-1), \kappa(1)\}$ which implies the claim. □

VI. Numerical Examples

Let B_2, B_κ and B_∞ denote the unit balls of $\|\cdot\|_2$, $\|\cdot\|_\kappa$ and $\|\cdot\|_\infty$ in the space X^q where the κ norm is the scaled H^2 norm given by $\|f\|_\kappa^2 = \|\kappa\|_\infty \|f\|_2^2$. Note that $B_\kappa \subseteq B_\infty$ is equivalent to the statement that $\|\kappa\|_\infty \|f\|_2^2 = 1$ implies $\|f\|_\infty^2 \leq 1$ and is thus equivalent to (10). Let ∂B_∞ denote the border of B_∞ then $B_\kappa \cap \partial B_\infty \neq \varnothing$ is equivalent to

[5]To see that in general $\|\kappa\|_\infty > \|\rho\|_\infty$ take, e.g., a c-stable $q \in \mathbb{R}[s]$ with $q = (s + 1/2)(s + e^{\jmath\omega_0})(s + e^{-\jmath\omega_0})$. By calculation

$$\kappa(\jmath\omega) = \begin{cases} 2 + 4\cos\omega_0 & \text{for } \omega = 0, \\ 1 + \frac{2}{\cos\omega_0} & \text{for } \omega = 1. \end{cases}$$

which implies $\kappa(0) < \kappa(1) \leq \|\kappa\|_\infty$ for $\omega_0 \in (\pi/2 - \varepsilon, \pi/2)$ and $\varepsilon > 0$ sufficiently small. In this case $\kappa(0) \neq \|\kappa\|_\infty > \|\rho\|_\infty$ by Corollary 12.

$\|\kappa\|_\infty = \|\rho\|_\infty$. We provide three examples: Example 1 and Example 2 are in discrete-time; Example 3 is continuous-time and similar to the one in Section 5 of [1].

Example 1 For $q = q_1 q_2$ with $q_1 = (z - 1/3 - \jmath/3)$ and $q_2 = (z - 1/3 + \jmath/3)$ we have $\|\kappa\|_\infty \approx \kappa(\mathrm{e}^{\jmath \cdot 0.72}) \approx 3.43382$ and $\|\rho\|_\infty = \rho(\mathrm{e}^{\jmath \cdot 0}) = 2.8$; see Fig. 1 and 2.

Example 2 For $q = q_1 q_2$ with $q_1 = (z - 1/20)$ and $q_2 = (z + 3/4)$ the numerical result is $\|\kappa\|_\infty = \|\rho\|_\infty = \kappa(\mathrm{e}^{\jmath \cdot 0}) \approx 7.905$; see Fig. 3.

Example 3 For $q(s) = (s + a)(s + a^*)$ with $a = 1/2 - 1/2\jmath$ we have $\|\kappa\|_\infty \approx \kappa(\jmath 2.0) \approx 4.062$ and $\|\rho\|_\infty \approx \rho(\jmath 1.94) \approx 2.13$. Let $g_\theta = \frac{\theta}{s+a} + \frac{\theta^*}{s+a^*}$, and normalize $f_\theta = g_\theta / \|g_\theta\|_2$ then the unit ball B_2 in X^q is given by $B_2 = \{f_\theta \,|\, \theta \in \mathbb{T}\}$; see Fig. 4.

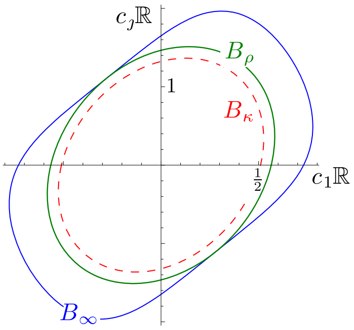

Fig. 1. Norm balls B_ρ, B_∞ and B_κ corresponding to Example 1 with $q_1^{-1} = c_1 + \jmath c_\jmath$ and $c_1, c_\jmath \in \mathrm{X}_q$. Due to $\|\rho\|_\infty < \|\kappa\|_\infty$ we observe that $\partial B_\infty \cap B_\kappa = \varnothing$.

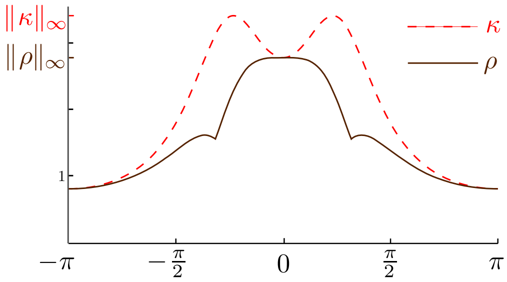

Fig. 2. Norm bounds $\kappa(\mathrm{e}^{\jmath\omega})$ and $\rho(\mathrm{e}^{\jmath\omega})$ corresponding to Example 1. As predicted by Corollary 12 we observe that $\|\rho\|_\infty = \kappa(0) < \|\kappa\|_\infty$.

VII. Conclusions

We have addressed the problem of bounding the absolute value of a complex valued function by its 2-norm assuming the function lives in a finite dimensional linear space. We emphasized rational functions describing stable continuous-time or discrete-time systems with real valued impulse response. The absolute value of such transfer functions (and thus also their H^∞ norm) can be bounded from above assuming knowledge of their poles and their H^2 norm. We have provided the abstract tools to treat the weighted H^2 norm case as well. In the unweighted H^2 case closed form expressions can be obtained for real rational modules. We have provided the basis for future work on this topic regarding vector and matrix-valued transfer functions. Due to the close relations between rational modules and realization theory, it is natural to further exploit state space methods in general and classical stability analysis results on Bézout, Lyapunov, and Stein equations in particular.

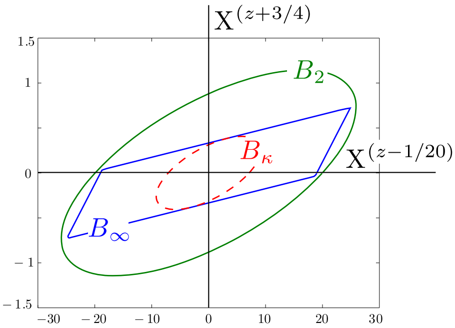

Fig. 3. Norm balls B_2, B_∞ and B_κ corresponding to Example 2. Due to $\|\rho\|_\infty = \|\kappa\|_\infty$ we observe that $\partial B_\infty \cap B_\kappa \neq \varnothing$.

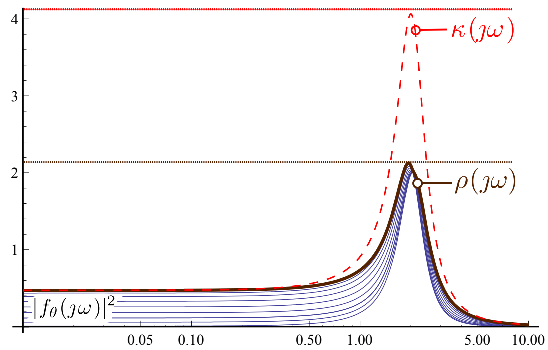

Fig. 4. Norm bounds $\kappa(\jmath\omega)$, $\rho(\jmath\omega)$ and $|f_\theta(\jmath\omega)|^2$ of all elements in $f_\theta \in B_2$ corresponding to Example 3. Note that $\rho(\jmath\omega)$ is the envelope of the $|f_\theta(\jmath\omega)|^2$'s and, as predicted by Remark 4, $\rho(\jmath\omega) \in [\kappa(\jmath\omega)/2, \kappa(\jmath\omega)]$.

References

[1] F. De Bruyne, B.D.O. Anderson, and M. Gevers, "Relating H_2 and H_∞ bounds for finite dimensional systems," *Systems & Control Letters*, vol. 24, pp. 173–181, January 1994.

[2] S. Douma and P. Van den Hof, "Relations between uncertainty structures in identification for robust control," *Automatica*, no. 3, pp. 439–457, 2005.

[3] T. Ivanov, P.-A. Absil, B. Anderson, and M. Gevers, "Application of real rational modules in system identification," in *47th IEEE Conference on Decision and Control*, Dezember 2008, pp. 111–116.

[4] P. Van Dooren, K. A. Gallivan, and P.-A. Absil, "H_2-optimal model reduction of MIMO systems," *Appl. Math. Lett.*, vol. 21, no. 12, pp. 1267–1273, 2008.

[5] B. Okutmustur, "Reproducing Kernel Hilbert Spaces," Ph.D. dissertation, Bilkent University, http://www.thesis.bilkent.edu.tr/0002953.pdf, August 2005.

[6] P. A. Fuhrmann, *A Polynomial Approach to Linear Algebra*. Springer-Verlag New York, Inc., 1996.

[7] A. Beurling, "On two problems concerning linear transformations in Hilbert space," in *Acta Math.*, vol. 81, 1949, pp. 239–255.

[8] S. Hara, B. D. Anderson, and H. Fujioka, "Relating H^2 and H^∞ norm bounds for sampled-data systems," in *IEEE Transactions on Automatic Control*, vol. 42, 1997, pp. 858–863.

[9] B. Simon, "The Christoffel–Darboux kernel," 2008, arXiv:0806.1528v1.

[10] H. Akçay, "On the uniform approximation of discrete–time systems by generalized fourier series," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1461–1467, 2001.

[11] H. Akçay and B. Ninness, "Orthonormal basis functions for modelling continuous-time systems," *Elsevier Signal Processing*, vol. 77, pp. 261 – 274, 1999.