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, H^2 norm, H^∞ norm, Tight Bound, 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, e.g., robust control, one is more interested in other norms like the supremum or H^∞ norm. Thus, linking these two norms can lead to valuable insights for these applications. This problem has been first addressed in the engineering context in [1] where one derived results such as

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

where 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 $\operatorname{Re}(a_i) > 0$ are 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 been noted that the bound has its limitations as it is no longer necessarily tight in the real rational case, i.e., if one restricts the coefficients of the linear combination (2), i.e., the b_i's, to be such that $M(s)$ is a transfer function of a system with real-valued impulse response. Complex coefficients then correspond to complex poles and, like the poles, come in complex conjugate pairs.

This paper provides a tight bound $\rho(\jmath\omega_0)\|M\|_2^2$, for the value of $|M(\jmath\omega_0)|^2$ for any fixed frequency ω_0, as well as a tight bound $\|\rho\|_\infty \|M\|_2^2$ for $\|M\|_\infty^2$, in the real rational case. In the continuous-time case it 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 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 $\kappa(s) = k(s,s)$.

The link to reproducing kernels is of interest because these objects have been studied extensively in the mathematical literature, see e.g., [2] for an overview. Specifically for the space of all strictly proper rational functions with common denominator q, which we refer to as a *rational module* and denote it by X^q, the reproducing kernel (RK) takes a particularly simple form since X^q is a coinvariant subspace of H^2, see e.g. [3], [4]. With this background, 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

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

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

with similar results for discrete-time. The ideas derived with this machinery generalize seamlessly to the case of $\mathbb{C}^n$ or $\mathbb{C}^{n\times m}$ instead of $\mathbb{C}$-valued functions. In the context of norm bounds a special vector-valued case has been studied in [5].

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. After some remarks on the general vector-valued case in Section VII, we conclude in Section VIII.

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

Let Ω denote an abstract set such as, e.g., the unit circle or the imaginary axis in the complex plane. 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.

In the following we will embed the linear space X over $\mathbb{R}$ in the smallest linear space $\mathbf{X}$ over $\mathbb{C}$ containing it. Assume that $\mathrm{X} \cap \jmath\mathrm{X} = \{0\}$ and let

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

denotes the complexification of X. Any $f \in \mathbf{X}$ then has a unique representation as $f = f_1 + \jmath f_\jmath$ with $f_1, f_\jmath \in \mathrm{X}$.

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. In order to represent this linear functional by an element in $\mathbf{X}$ we introduce a complex valued inner product $\langle\cdot,\cdot\rangle$ on $\mathbf{X}$ via

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

where we have extended the real valued $(\cdot,\cdot)$ to $\mathbf{X}$ 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 2-norm $\|\cdot\|_2$ induced by $\langle\cdot,\cdot\rangle$ and the supremum norm $\|\cdot\|_\infty$ defined by

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

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 $w \in \Omega$. In other words k is independent of the particular choice of ONB [2]. Note that $k(w, z) = k(z, w)^*$.

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

Theorem 1 *Let $\kappa(w) = k(w, w)$. 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 $\mathbf{X}$.

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. □

The inequality (11) fails to be tight on $\mathrm{X} \subseteq \mathbf{X}$ since k_w being an element in $\mathbf{X}$ does not suffice for $k_w \in \mathrm{X}$. Actually it is easy to check that $k_w = k_{w,1} + \jmath k_{w,\jmath}$ with $k_{w,\jmath} \ne 0$ in general. We expand k_w into $k_{w,1} + \jmath k_{w,\jmath}$ and note that

$$\begin{aligned} |f(w)|^2 &= |\langle f, k_w\rangle|^2 \\ &= |(f, k_w) + \jmath(\jmath f, -k_w)|^2 \\ &= (f, k_{w,1})^2 + (f, k_{w,\jmath})^2, \end{aligned} \tag{13}$$

where the last equality holds if and only if $f \in \mathrm{X}$.

We maximize (13) over the unit ball in X to obtain a new tight bound on X given by $\rho(w)$ as defined in Theorem 2 which is the abstract version of our main result presented as Theorem 9 below.

Theorem 2 *Let $k_w = k_{w,1} + \jmath k_{w,\jmath}$ with $k_{w,1}, k_{w,\jmath} \in \mathrm{X}$ for all $w \in \Omega$. Moreover define*

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

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

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

The maximum eigenvalue of G is given by $\rho(w)$ which follows by a simple calculation. So it remains to check

$$\lambda_{\max}(G) = \sup_{f\in X}\{(f, k_{w,1})^2 + (f, k_{w,\jmath})^2 \,|\, \|f\|_2^2 = 1\} =: \sigma.$$

Supremizing over X and supremizing over X_w yields to the same value σ where X_w denotes the 2-dimensional subspace generated by $k_{w,1}, k_{w,\jmath} \in \mathrm{X}$. 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

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

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 3 The proof of Theorem 2 together with

$$\lambda_1(G) + \lambda_2(G) = \|k_{w,1}\|_2^2 + \|k_{w,\jmath}\|_2^2 = \langle k_w, k_w\rangle, \tag{16}$$

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

$$\kappa(w)/2 \le \rho(w) \le \kappa(w). \tag{17}$$

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 $\mathrm{R}L^2_\bullet$ of $L^2_\bullet$ denoted by $\mathrm{R}L^2_d$ for discrete-time and $\mathrm{R}L^2_c$ for continuous-time. Since every single-input single-output linear time invariant (LTI) system admits an input-output decomposition, its controllable part is represented by its transfer function which is a real rational function [6]. In the following we establish the fact that the second summand in (14) is given by the absolute value of $k(w,w^{-1})/2$ for discrete-time and $k(w,-w)/2$ for continuous-time: see (21) and (24) below.

A. Discrete Time

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

$$\|f\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(\mathrm{e}^{\jmath\omega})|^2 \,\mathrm{d}\omega < \infty. \tag{18}$$

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

$$\mathrm{R}L^2_d = \{f \in \mathbb{R}(z) \mid f \text{ has no pole in } \mathbb{D}\} \tag{19a}$$
$${}^{\mathrm{c}}\mathrm{R}L^2_d = \{f \in \mathbb{C}(z) \mid f \text{ has no pole in } \mathbb{D}\}, \tag{19b}$$

The following fact is elementary; so we skip its proof.

Lemma 4 *Let $f \in {}^{\mathrm{c}}\mathrm{R}L^2_d$ and*

$$f_1(z) = \frac{f(z) + f^*(z^{-1})}{2},\; f_\jmath(z) = \frac{f(z) - f^*(z^{-1})}{2\jmath}, \tag{20}$$

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

Theorem 5 *Let $k : \mathbb{D}\times\mathbb{D} \to \mathbb{C}$ be the kernel which reproduces the complexification of a finitely generated subspace $\mathrm{X} \subseteq \mathrm{R}L^2_d$. Then*

$$k(w,w^{-1})/2 = \|k_{w,1}\|_2^2 - \|k_{w,\jmath}\|_2^2 - \jmath\, 2(k_{w,1}, k_{w,\jmath}) \tag{21}$$

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 4, we have $b_i^* \in \mathrm{R}L^2_d$. This implies, again by Lemma 4, 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 we have 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 (21). □

B. Continuous Time

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

$$\|f\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |f(\jmath\omega)|^2 \,\mathrm{d}\omega < \infty. \tag{22}$$

We define the real rational subspace $\mathrm{R}L^2_c$ and its complexificaton ${}^{\mathrm{c}}\mathrm{R}L^2_c \subseteq L^2_c$:

$$\mathrm{R}L^2_c = \{f \in \mathbb{R}(z) \mid f \text{ s.p., no pole in } \jmath\mathbb{R}\} \tag{23a}$$
$${}^{\mathrm{c}}\mathrm{R}L^2_c = \{f \in \mathbb{C}(z) \mid f \text{ s.p., no pole in } \jmath\mathbb{R}\}, \tag{23b}$$

with s.p. meaning strictly proper.

Theorem 6 is the continuous-time version of Theorem 5. The proof is completely analogous and therefore skipped.

Theorem 6 *Let $k : \jmath\mathbb{R}\times\jmath\mathbb{R} \to \mathbb{C}$ be the kernel which reproduces the complexification of a finitely generated subspace $\mathrm{X} \subseteq \mathrm{R}L^2_c$. Then*

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

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 $k_w = k_{-w}$.

So we have established the fact that ρ, defined in (14), is given by $\rho(z) = k(z,z)/2 + |k(z,z^{-1})|/2$ discrete-time and $\rho(s) = k(s,s)/2 + |k(s,-s)|/2$ for continous-time.

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

In this section we specialize the subspace $\mathrm{X} \subseteq \mathrm{R}L^2_\bullet$ to be a real rational module. This will allow us to compute the reproducing kernel of its complexification and thus turn the previously derived abstract formulas into concrete closed form expressions. 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 $\operatorname{Re} a < 0$ (resp. $|a| < 1$). We define the real rational Hardy spaces as subspaces of $\mathrm{R}L_c^2$ and $\mathrm{R}L_d^2$ respectively

$$\mathrm{R}H_c^2 = \{f \mid f = p/q \text{ strictly proper}, q \text{ is } c\text{-stable}\}, \quad (25a)$$

$$\mathrm{R}H_d^2 = \{f \mid f = p/q \text{ strictly proper}, q \text{ is } d\text{-stable}\}. \quad (25b)$$

For $q \in \mathbb{R}[x]$ define its polynomial module $\mathrm{X}_q = \{p \in \mathbb{R}[x], \deg(p) < \deg(q)\}$ and its rational module

$$\mathrm{X}^q = \left\{\frac{p}{q} \in \mathbb{R}(x) : p \in \mathrm{X}_q\right\}, \quad (26)$$

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 $\mathrm{X}^q \subseteq \mathrm{R}H_c^2$ and $\mathrm{X}^q \subseteq \mathrm{R}H_d^2$ 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. [3], [4]) states that $\mathbf{X}^q$ is *coinvariant*, i.e.,

$$^{\mathrm{c}}\mathrm{R}H_c^2 \ominus \mathbf{X}^q = \frac{q^\star}{q}\,^{\mathrm{c}}\mathrm{R}H_c^2, \quad ^{\mathrm{c}}\mathrm{R}H_d^2 \ominus \mathbf{X}^q = \frac{q^\star}{q}\,^{\mathrm{c}}\mathrm{R}H_d^2, \quad (27)$$

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

$$q^\star(s) = q(-s), \quad \text{and} \quad q^\star(z) = z^n q(z^{-1}), \quad (28)$$

if q is c-stable and d-stable, respectively, and $n = \deg(q)$. The importance of (27) is the corollary

$$\mathbf{X}^{pq} \ominus \mathbf{X}^q = \frac{q^\star}{q}\mathbf{X}^p, \quad (29)$$

whenever p, q are two c-stable or d-stable polynomials (cf. Corollary 5 in [7]). It is (29) that enables us to derive an explicit form for the reproducing kernel of $\mathbf{X}^q$ in discrete-time in Theorem 7 and continuous-time in Theorem 8. Due to the recursive structure of their computation these kernels are called Christoffel-Darboux kernels [8].

Theorem 7 *Let $q = \prod_{i=1}^n (z - a_i) \in \mathbb{R}[z]$, d-stable, and $k : \mathbb{D} \times \mathbb{D} \to \mathbb{C}$ be defined via*

$$k(z,w) = \frac{1}{q(z)q^\star(w)} \cdot \frac{q^\star(z)q(w) - q(z)q^\star(w)}{1 - z\bar{w}}, \quad (30)$$

for $z \neq w$ and

$$k(z,z) = z\left(\frac{q'(z)}{q(z)} - \frac{q^{\star\prime}(z)}{q^\star(z)}\right) = \sum_{i=1}^n \frac{1 - |a_i|^2}{|z - a_i|^2}, \quad (31)$$

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

Proof: Let $q = q_3$ with $q_3 = q_1 \cdot q_2$ and $q_1, q_2 \in \mathbb{R}[z]$. Moreover let

$$m_i(z,w) = \frac{q_i^\star(z)}{q_i(z)}\frac{q_i^\star(w)}{q_i(w)},$$

for $i = 1, 2, 3$. Let k_i be the reproducing kernel of $\mathbf{X}^{q_i}$. Then

$$k_3(z,w) = \frac{m_1(z,w) - 1}{1 - z\bar{w}} + \frac{m_3(z,w) - m_1(z,w)}{1 - z\bar{w}}$$
$$= \frac{m_3(z,w) - 1}{1 - z\bar{w}} = \frac{\frac{q^\star(z)q(w)}{q(z)q^\star(w)} - \frac{q(z)q^\star(w)}{q(z)q^\star(w)}}{1 - z\bar{w}}$$

by induction hypotheses. Here we used the facts that $k_3(z,w) = k_1(z,w) + m_1(z,w)k_2(z,w)$, due to (29). More precisely: due to the *orthogonal* decomposition (29) k_3 is the sum of k_1 and the RK, say $\tilde{k}_2$, of $n_1\mathrm{X}^{q_2}$ with $n_1 = q_1^\star/q_1$. Since n_1 is all-pass, i.e., $n_1^* n_1 = 1$, it follows by (9) that $\tilde{k}_2(z,w) = n_1(z)k_2(z,w)n_1^*(w)$ which equals $m_1 k_2$.

It is easy to check that (30) holds for $q = (z - a_1)$. Thus, by induction, we have proven (30). The diagonal readily follows from the fact that $k(z,w)$ is continuous; apply l'Hospital's rule to calculate $k(w,w)$ via $\lim_{z \to w} k(z,w)$. □

Theorem 8 *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*

$$k(s,w) = \frac{1}{q(s)q^\star(s)} \cdot \frac{q^\star(s)q(w) - q(s)q^\star(w)}{w - s}, \quad (32)$$

for $s \neq w$ and

$$k(s,s) = \frac{q'(s)}{q(s)} - \frac{q^{\star\prime}(s)}{q^\star(s)} = -\sum_{i=1}^n \frac{2\,\mathrm{Re}\,a_i}{|s - a_i|^2}, \quad (33)$$

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

Proof: We obtain (32) by the same reasoning as we obtained (30) in the proof of Theorem 7. Since $k(s,w)$ is continuous we can calculate $k(w,w)$ via $\lim_{s \to w} k(s,w)$ and l'Hospital's rule, i.e., $k(w,w)$ equals

$$-\frac{q^{\star\prime}(w)}{q^\star(w)} + \frac{q'(w)}{q(w)} = \sum \frac{-1}{-w - \bar{a}_i} + \frac{1}{w - a_i}$$
$$= \sum \frac{w - a_i - w - \bar{a}_i}{(-w - \bar{a}_i)(w - a_i)},$$

which, replacing w in $k(w,w)$ by s, concludes the proof. □

We summarize the results in Theorem 9, which is Theorem 2 specialized to rational modules, using Theorem 6, 8 for continuous-time and Theorem 5, 7 for discrete-time.[1]

Theorem 9 *If $\mathbf{X} = \mathrm{X} + \jmath\mathrm{X}$ and $\mathrm{X} = \mathrm{X}^q$ (equipped with L_c^2 norm) for some c-stable $q \in \mathbb{R}[s]$, $q = \prod(s - a_i)$, then Theorem 1 and Theorem 2 hold with*

$$\kappa(s) = \frac{q'(s)}{q(s)} - \frac{q^{\star\prime}(s)}{q^\star(s)} = -\sum_{i=1}^n \frac{2\,\mathrm{Re}\,a_i}{|s - a_i|^2}, \quad (34a)$$

$$\rho(s) = \frac{1}{2}\left|\frac{1 - (q(s)^{-1}q(-s))^2}{2s}\right| + \frac{\kappa(s)}{2} \quad (34b)$$

Similarly, in the discrete-time case, Theorem 1 and Theorem 2 hold with

$$\kappa(z) = \frac{zq'(z)}{q(z)} - \frac{zq^{\star\prime}(z)}{q^\star(z)} = \sum_{i=1}^n \frac{1 - |a_i|^2}{|z - a_i|^2}, \quad (35a)$$

$$\rho(z) = \frac{1}{2}\left|\frac{(z^n q(z)^{-1}q(z^{-1}))^2 - 1}{1 - z^2}\right| + \frac{\kappa(z)}{2} \quad (35b)$$

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

[1] Our results in (34a, 35a) correspond to the results of De Bruyne at al (2.4, 3.4) as found in [1] since $\beta_{ij} = \langle 1/(s - a_i), 1/(s - a_j)\rangle$ implies

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

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 $k(z,w)$.

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 (34), (35) for continuous- and discrete-time respectively.

Lemma 10 *Let $q \in \mathbb{R}[z]$ be d-stable with $\deg(q) > 0$ and k denote the reproducing kernel of $\mathbf{X}^q$. Then $\{k(z,\cdot) \,|\, z \in \mathbb{D}\}$ separates points in $\mathbb{D}$. That is for all $w_i \in \mathbb{D}$ with $w_1 \neq w_2$ there exists $z \in \mathbb{D}$ 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 10 and property 3) of Theorem 5 and Theorem 6, respectively, we obtain a necessary and sufficient condition for $\|\kappa\|_\infty = \|\rho\|_\infty$ in the form of Corollary 11 and an easy sufficient condition given by Corollary 12.[2]

Corollary 11 *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 12 *Let $q = \prod(x - a_i) \in \mathbb{R}[x]$ be d-stable ($x = z$) or c-stable ($x = s$). Then the condition $a_i = \mathrm{Re}\, a_i$ 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 $\mathrm{Re}\, z$ on $\{z \in \mathbb{D} \,|\, \mathrm{Re}(z) > 0\}$ since each of its summands

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

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 that $\|\kappa\|_\infty = \|\rho\|_\infty$. □

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 $\|\kappa\|_\infty = \|\rho\|_\infty$. 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].

[2] To see that in general $\|\kappa\|_\infty > \|\rho\|_\infty$ take, e.g., a c-stable $q \in \mathbb{R}[s]$ with $q = (s + 1/2)(s + \mathrm{e}^{\jmath\omega_0})(s + \mathrm{e}^{-\jmath\omega_0})$. 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 11.

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 $\|\kappa\|_\infty \approx \kappa(\mathrm{e}^{\jmath\,0.72}) \approx 3.43382$ and $\|\rho\|_\infty = \rho(\mathrm{e}^{\jmath\,0}) = 2.8$; 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 $\|\kappa\|_\infty = \|\rho\|_\infty = \kappa(\mathrm{e}^{\jmath\,0}) \approx 7.905$; see Fig. 3.

Example 3 For $q(s) = (s + a)(s + a^*)$ with $a = 1/2 - 2\jmath$ 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 $B_2 = \{f_\theta \,|\, \theta \in \mathbb{D}\}$; see Fig. 4.

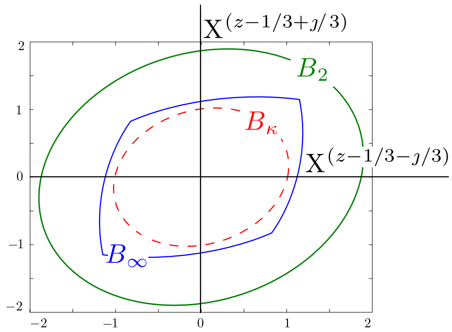

Fig. 1. Norm balls B_2, B_∞ and B_κ corresponding to Example 1. Due to $\|\rho\|_\infty < \|\kappa\|_\infty$ we observe that $\partial B_\infty \cap B_\kappa = \varnothing$.

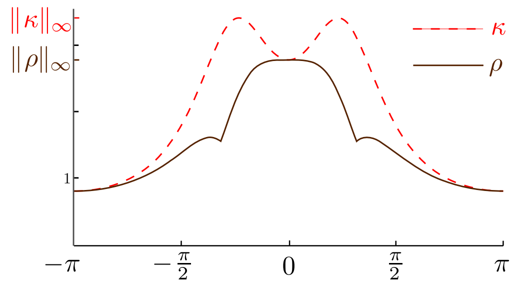

Fig. 2. Norm bounds $\kappa(\mathrm{e}^{\jmath\omega})$ and $\rho(\mathrm{e}^{\jmath\omega})$ corresponding to Example 1. As predicted by Corollary 11 we observe that $\kappa(0) < \|\kappa\|_\infty$.

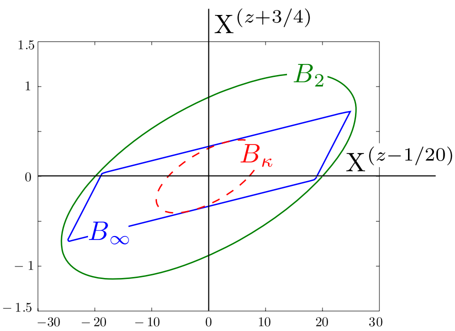

Fig. 3. Norm balls B_2, B_∞ and B_κ corresponding to Example 2. Due to $\|\rho\|_\infty = \|\kappa\|_\infty$ we observe that $\partial B_\infty \cap B_\kappa \neq \varnothing$.

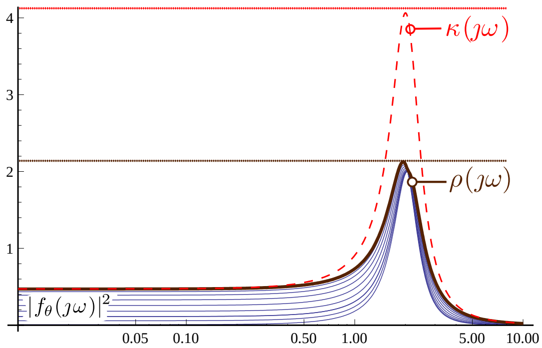

Fig. 4. Norm bounds $\kappa(\jmath\omega)$, $\rho(\jmath\omega)$ and $|f_\theta(\jmath\omega)|^2$ of all elements in $f_\theta \in B_2$ corresponding to Example 3. Note that $\rho(\jmath\omega)$ is the envelope of the $|f_\theta(\jmath\omega)|^2$'s and, as predicted by Remark 3, $\rho(\jmath\omega) \in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

VII. Remarks on the Vector Valued Case

In this section we provide a natural extension of the complex bound for the vector-valued case which shows the flexibility of the integral kernel approach. To keep the paper reasonably short we choose not to discuss issues of loosing tightness of the complex bound when dealing with linear spaces over the reals.

Let $(\mathbf{X}, \langle\cdot,\cdot\rangle)$ denote a finite dimensional complex inner product space of $\mathcal{A}$-valued functions defined on some set Ω which are bounded in the sense that

$$\|f\|_\infty = \sup_{w\in\Omega} \|f(w)\|_{\mathcal{A}} = \sup_{\omega\in\Omega} \sup_{\alpha\in\mathcal{A}_1} [\![f(w),\alpha]\!] < \infty,$$

for all $f \in \mathbf{X}$. Here $(\mathcal{A}, [\![\cdot,\cdot]\!])$ denotes a finitely generated complex inner product space, $\|\alpha\|^2_{\mathcal{A}} = [\![\alpha,\alpha]\!]$ the induced norm, and $\mathcal{A}_1$ the unit ball. For $\alpha \in \mathcal{A}$ let α^* denote the functional defined by $\alpha^*(\beta) = [\![\beta,\alpha]\!]$ for all $\beta \in \mathcal{A}$. In the scalar case $\mathcal{A} = \mathbb{C}$ and α^* is the complex conjugate of α.

Mimicking the scalar case, given an ONB $\{b_i\}_{i=1}^n$ of $\mathbf{X}$ let $K(z,w) = \sum_{i=1}^n b_i(z)b_i^*(w)$ and $K_{w,\alpha} = K(z,w)\alpha$. Then K is uniquely determined by its properties: 1) $K_{w,\alpha} \in \mathbf{X}$ and 2) $\langle f, K_{w,\alpha}\rangle = \alpha^* f(w)$ for all $w \in \Omega, \alpha \in \mathcal{A}$, $f \in \mathbf{X}$; called the reproducing kernel of $(\mathbf{X}, \langle\cdot,\cdot\rangle)$. In particular $K : \Omega\times\Omega \to L(\mathcal{A})$ is independent of the chosen ONB and takes its values in the space $L(\mathcal{A})$ of linear maps from $\mathcal{A}$ to $\mathcal{A}$.

Theorem 1 is a special case of the statement found in Theorem 13 for the vector-valued case.

Theorem 13 *Let $R(w) = K(w,w)^{1/2}$, i.e., $R(w) \in L(\mathcal{A})$ symmetric (Hermitian), such that $R(w)R(w) = K(w,w)$ and let $\|\cdot\|_{L(\mathcal{A})}$ denote the operator norm on $L(\mathcal{A})$ induced by the norm on $\mathcal{A}$. For all $w \in \Omega, f \in \mathbf{X}$ there holds $\|f(w)\|^2_{\mathcal{A}} \le \|R(w)\|^2_{L(\mathcal{A})}\|f\|^2_2$ and this bound is tight. In particular*

$$\|\cdot\|^2_\infty \le \|R\|^2_\infty \cdot \|\cdot\|^2_2, \tag{36}$$

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

Proof: For any $\alpha \in \mathcal{A}$, the Cauchy-Bunyakovsky-Schwarz inequality in $\mathbf{X}$ yields

$$|[\![\alpha, f(w)]\!]|^2 \le \|K_{w,\alpha}\|^2_2 \cdot \|f\|^2_2. \tag{37}$$

Utilizing $\langle K_{w,\alpha}, K_{w,\alpha}\rangle = [\![\alpha, K(w,w)\alpha]\!]$, which equals $\|R(w)\alpha\|^2_{\mathcal{A}}$, and supremizing over $\alpha \in \mathcal{A}_1$ yields

$$\|f(w)\|^2_{\mathcal{A}} \le \|R(w)\|^2_{L(\mathcal{A})} \cdot \|f\|^2_2, \tag{38}$$

The rest follows by the same reasoning as in the scalar case discussed in Theorem 1. □

We conclude this section with a simple example where the computation of $R(w) = K(w,w)^{1/2}$ can be reduced to the scalar case.

Example 4 Let $(\mathbb{C}^{n\times m}, [\![\cdot,\cdot]\!])$ be defined by $[\![A,B]\!] = \mathrm{tr}(B^{\mathrm{H}}A)$ for all $A, B \in \mathbb{C}^{n\times m}$, $q \in \mathbb{R}[z]$ d-stable and

$$\mathbf{X}^q \otimes \mathbb{C}^{n\times m} := \{P/q \,|\, P \in \mathbb{C}^{n\times m}[z], \deg(P_{ij}) < \deg(q)\},$$

which is a subspace[3] of $\left(L^2(\mathbb{D}, \mathbb{C}^{n\times m}), \langle\cdot,\cdot\rangle\right)$, with inner product $\langle f,g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} [\![f(\mathrm{e}^{\jmath\omega}), g(\mathrm{e}^{\jmath\omega})]\!]\,\mathrm{d}\omega$. Let k denote the kernel reproducing the scalar space $\mathbf{X}^q$. Then $\langle f, k_w\cdot\alpha\rangle$ equals $\alpha^* f(w)$ because of[4]

$$\int [\![f(\mathrm{e}^{\jmath\omega}), k_w(\mathrm{e}^{\jmath\omega})\cdot\alpha]\!]\,\mathrm{d}\omega = \int [\![f(\mathrm{e}^{\jmath\omega}),\alpha]\!]\cdot k_w^*(\mathrm{e}^{\jmath\omega})\,\mathrm{d}\omega.$$

In other words $K(z,w) = k(z,w)\cdot I$ reproduces $\mathbf{X}^q\otimes\mathbb{C}^{n\times m}$ where I denotes the identity in $L(\mathbb{C}^{n\times m})$. Using (38) this brings us to the conclusion that $\|\cdot\|^2_\infty \le \|\kappa\|_\infty \|\cdot\|^2_2$, with $\kappa(w) = k(w,w)$ gives a tight bound on $\mathbf{X}^q \otimes \mathbb{C}^{n\times m}$.

VIII. Conclusions

We have solved the problem of bounding the absolute value of a continuous-time or discrete-time system transfer function (and thus also its H^∞ norm) from above using its H^2 norm and the reproducing kernel of the function class considered. We have shown how closed form expressions can be obtained by restricting the class of transfer functions to those forming real rational modules. Furthermore we have provided the basis for future work on this topic regarding vector and matrix-valued transfer functions.

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

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

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

[5] 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.

[6] J. C. Willems and J. W. Polderman, *Introduction to Mathematical Systems Theory – A Behavioral Approach*. Springer-Verlag New York, Inc., 1998.

[7] 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*, 2008.

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

[3] Note that, by partial fraction expansion, all elements $f \in \mathbf{X}^q \otimes \mathbb{C}^{n\times m}$ are of the form $f = \sum \frac{A_i}{z-a_i}$ $A_i \in \mathbb{C}^{n\times m}$ and $q = \prod(z-a_i)$ assuming no multiple zeros which is the case studied by S. Hara; cf. Theorem 2 [5].

[4] We use reproducing property of k and the fact that $[\![f,\alpha]\!] \in \mathbf{X}^q$ for all $f \in \mathbf{X}^q \otimes \mathbb{C}^{n\times m}$ and all $\alpha \in \mathbb{C}^{n\times m}$.