

Generalizing the KYP Lemma to the Union of Intervals

Goele Pipeleers[‡], Tetsuya Iwasaki[†], and Shinji Hara[‡]

Abstract—A recent generalization of the Kalman-Yakubovich-Popov (KYP) lemma establishes the equivalence between a semi-infinite inequality on a segment of a circle or straight line in the complex plane and a linear matrix inequality. In this paper we further generalize the KYP lemma to particular curves in the complex plane, which include the union of segments of a circle or line as a special case.

I. INTRODUCTION

The Kalman-Yakubovich-Popov (KYP) lemma [1]–[3] is a key result in modern system and control theory. The classical version of the lemma states the equivalence between a semi-infinite inequality on an entire circle or straight line in the complex plane and a finite-dimensional linear matrix inequality (LMI). Recently, the KYP lemma was extended to inequalities on a segment of a circle or line [4], [5]. In this paper we further generalize the KYP lemma to particular curves in the complex plane, characterized by a polynomial equality and inequality that satisfy certain conditions. As a first application of our result we show that the considered set of curves includes the union of segments on a circle or line as a special case. Although an LMI equivalent to a semi-infinite inequality on a union of segments can also be obtained by application of the results of [4], [5] to each of the intervals separately, the LMI obtained by our novel result is generally smaller in both dimension and number of variables.

The paper is organized as follows: Section II describes the curves in the complex plane that are considered in our generalization of the KYP lemma. The generalization itself is presented in Section III, and compared to existing results in Section IV. Section V concludes the paper.

Notation. The set of positive integers is denoted by \mathbb{I} , and its subset up to $n \in \mathbb{I}$ is $\mathbb{I}_n = \{1, \dots, n\}$. The symbols $\mathbb{R}^{n \times m}$ and $\mathbb{C}^{n \times m}$ denote the sets of $n \times m$ real, respectively complex, matrices, and \mathbb{R}_+ and \mathbb{R}_{++} corresponds to the sets of nonnegative, respectively positive, real numbers. The sets of $n \times n$ real symmetric and complex Hermitian matrices are indicated by \mathbb{S}^n and \mathbb{H}^n , respectively. For a matrix X , \bar{X} denotes its complex conjugate, X^T its transpose, and X^* its complex conjugate transpose: $X^* = \bar{X}^T$. $0_{n,m}$ is the $n \times m$ zero matrix, and I_n the $n \times n$ identity matrix, and the

subscripts are omitted when the dimensions are clear from the context. The matrix Kronecker product is indicated by \otimes , and the imaginary unit is denoted by $i = \sqrt{-1}$.

II. CURVES IN THE COMPLEX PLANE

A. General Form

In this paper we generalize the KYP lemma to curves in the complex plane of the following form:

$$\Lambda(\Phi, \Psi) = \{\lambda \in \mathbb{C} : l_\ell(\lambda)^* \Phi l_\ell(\lambda) = 0, \quad l_\ell(\lambda)^* \Psi l_\ell(\lambda) \geq 0\}, \quad (1)$$

where $\ell \in \mathbb{I}$ is a given integer and $\Phi, \Psi \in \mathbb{H}^{\ell+1}$ are given Hermitian matrices. The mapping $l_\ell(\lambda) : \mathbb{C} \rightarrow \mathbb{C}^{\ell+1}$ is defined as

$$l_\ell(\lambda) = [\lambda^\ell \quad \lambda^{\ell-1} \quad \dots \quad \lambda \quad 1]^T, \quad (2)$$

while $l_0(\lambda) = 1$. In case $\Lambda(\Phi, \Psi)$ is unbounded it is extended with ∞ , and we agree $l_\ell(\infty) = [1 \ 0 \ \dots \ 0]^T$. We will assume the Hermitian matrices Φ, Ψ to satisfy the following two assumptions:

Assumption 1: The matrices $\Phi, \Psi \in \mathbb{H}^{\ell+1}$ admit a decomposition of the following form:

$$\Phi = T^* \Phi_o T, \quad \Psi = T^* \Psi_o T, \quad (3a)$$

where

$$\Phi_o = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \Psi_o = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}, \quad \begin{array}{l} 0 \leq \alpha \leq \gamma, \text{ or} \\ \alpha < 0 < \gamma, \end{array} \quad (3b)$$

for some matrix $T \in \mathbb{C}^{2 \times (\ell+1)}$ of full row rank, and real scalars $\alpha, \beta, \gamma \in \mathbb{R}$. In addition, for each $s \in \Lambda(\Phi_o, \Psi_o)$, the ℓ th degree polynomial equality in λ given by

$$[1 \quad -s] T l_\ell(\lambda) = 0, \quad \text{for } s \neq \infty, \quad (4a)$$

$$[0 \quad 1] T l_\ell(\lambda) = 0, \quad \text{for } s = \infty, \quad (4b)$$

has ℓ distinct roots. These roots are grouped in the set $\mathbb{L}_T(s)$.

Assumption 2: When $\ell \geq 2$, there exists a Hermitian matrix $R \in \mathbb{H}^\ell$ such that

$$l_{\ell-1}(\lambda)^* R l_{\ell-1}(\lambda) > 0, \quad \forall \lambda \in \Lambda(\Phi, \Psi), \quad (5a)$$

$$l_{\ell-1}(\lambda_i)^* R l_{\ell-1}(\lambda_j) = 0, \quad \forall \lambda_i, \lambda_j \in \mathbb{L}_T(s), \quad i \neq j, \quad \forall s \in \Lambda(\Phi_o, \Psi_o). \quad (5b)$$

Decomposition (3) defines a mapping between the curve $\Lambda(\Phi, \Psi)$ and a segment on the imaginary axis, $\Lambda(\Phi_o, \Psi_o)$. Each $s \in \Lambda(\Phi_o, \Psi_o)$ is mapped onto the ℓ roots of (4),

[‡]Goele Pipeleers is a Postdoctoral Fellow of the Research Foundation Flanders affiliated with the Dept. of Mechanical Engineering, Katholieke Universiteit Leuven, Celestijnenlaan 300B, B-3001 Heverlee, Belgium, goele.pipeleers@mech.kuleuven.be.

[†]Tetsuya Iwasaki is with the Dept. of Mechanical and Aerospace Engineering, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095-1597, USA, tiwasaki@ucla.edu.

[‡]Shinji Hara is with the Dept. of Information Physics and Computing, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan, shinji.hara@ipc.i.u-tokyo.ac.jp.

grouped in $\mathbb{L}_T(s) \subset \Lambda(\Phi, \Psi)$, while all $\lambda \in \mathbb{L}_T(s)$ are mapped onto the same s :

$$s = \frac{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} T l_\ell(\lambda)}{T l_\ell(\lambda)} = \frac{t_1(\lambda)}{t_2(\lambda)}, \quad \text{if } t_2(\lambda) \neq 0, \quad (6a)$$

$$= \infty, \quad \text{otherwise.} \quad (6b)$$

Hence, for all $\lambda \in \mathbb{L}_T(s)$ with $s \neq \infty$,

$$T l_\ell(\lambda) = t_2(\lambda) \begin{bmatrix} s \\ 1 \end{bmatrix}. \quad (6c)$$

The curve $\Lambda(\Phi, \Psi)$ is thus characterized as the set of complex numbers λ that belong to $\mathbb{L}_T(s)$ for some $s \in \Lambda(\Phi_o, \Psi_o)$.

For $\ell = 1$, Assumption 2 is irrelevant, and the curves considered in the generalized KYP lemma of [4], [5] are retrieved: $\Lambda(\Phi, \Psi)$ corresponds to a nonempty and nonsingleton segment of a circle or straight line in \mathbb{C} . If in addition $\Psi = 0$ is considered, the curve corresponds to the entire circle or line, retrieving the original KYP lemma [1]–[3].

B. Union of Segments of Circle or Straight Line

As a particular case, the union of ℓ nonintersecting, nonempty and nonsingleton segments on a circle or straight line in \mathbb{C} admits a description of the form (1) with Hermitian matrices $\Phi, \Psi \in \mathbb{H}^{\ell+1}$ that satisfy Assumptions 1 and 2. For the sake of brevity, we focus on the real axis. The extension to other lines and circles in the complex plane essentially relies on a bijective mapping to the real axis [4], and is briefly discussed at the end of this section. To clarify our elaborations, we define the matrices Φ_r, Ψ_r as follows:

$$\Phi_r = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \quad \Psi_r = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad (7)$$

such that $\Lambda(\Phi_r, \Psi_r) = \mathbb{R}_+ \cup \{\infty\}$. In addition, for $\ell \in \mathbb{I}$, $J_\ell \in \mathbb{R}^{2\ell \times (\ell+1)}$ is defined as

$$J_\ell = \begin{bmatrix} I_\ell & 0_{\ell,1} \\ 0_{\ell,1} & I_\ell \end{bmatrix}. \quad (8)$$

Lemma 1: Let 2ℓ scalars $\alpha_i, \beta_i \in \mathbb{R}$, $i \in \mathbb{I}_\ell$, be given that satisfy

$$\alpha_1 < \beta_1 < \alpha_2 < \dots < \beta_{\ell-1} < \alpha_\ell < \beta_\ell. \quad (9)$$

Let the vectors $a, b \in \mathbb{R}^{\ell+1}$ be defined by

$$\prod_{i=1}^{\ell} (\lambda - \alpha_i) = a^\top l_\ell(\lambda), \quad \prod_{i=1}^{\ell} (\lambda - \beta_i) = b^\top l_\ell(\lambda), \quad (10)$$

and set $\tilde{T} = [-a \ b]^\top$. Then, the matrices $\Phi, \Psi \in \mathbb{H}^{\ell+1}$ given by

$$\Phi = \tilde{T}^* \Phi_r \tilde{T}, \quad \Psi = \tilde{T}^* \Psi_r \tilde{T}, \quad (11)$$

satisfy

$$\Lambda(\Phi, \Psi) = \bigcup_{i=1}^{\ell} [\alpha_i, \beta_i]. \quad (12)$$

In addition, there exists $R \in \mathbb{S}^\ell$ such that

$$\Phi = J_\ell^* (\Phi_r \otimes R) J_\ell, \quad (13a)$$

and

$$l_{\ell-1}(\lambda)^* R l_{\ell-1}(\lambda) > 0, \quad \forall \lambda \in \mathbb{R}. \quad (13b)$$

Proof: Decomposition (11) defines a mapping between $\tau \in \Lambda(\Phi_r, \Psi_r) = \mathbb{R}_+ \cup \{\infty\}$ and $\lambda \in \Lambda(\Phi, \Psi)$, where every $\tau \in \mathbb{R}_+$ is mapped onto the ℓ roots of

$$\begin{bmatrix} 1 & -\tau \end{bmatrix} \tilde{T} l_\ell(\lambda) = -a^\top l_\ell(\lambda) - \tau b^\top l_\ell(\lambda) = 0,$$

while for $\tau = \infty$, this equation reads as $b^\top l_\ell(\lambda) = 0$. Hence, $\Lambda(\Phi, \Psi)$ corresponds to the 180° root locus of

$$\frac{b^\top l_\ell(\lambda)}{a^\top l_\ell(\lambda)} = \frac{\prod_{i=1}^{\ell} (\lambda - \beta_i)}{\prod_{i=1}^{\ell} (\lambda - \alpha_i)},$$

and indeed corresponds to the union of the intervals $[\alpha_i, \beta_i]$ for $i \in \mathbb{I}_\ell$. Consequently, (12) holds.

The real symmetric matrix R that satisfies (13a) is unique. Since both Φ_r and Ψ_r are pure imaginary, the only nontrivial constraints relate to the off-diagonal entries of the imaginary part of Φ . The corresponding $0.5\ell(\ell+1)$ linear constraints on the $0.5\ell(\ell+1)$ entries of R can readily be verified to be linearly independent.

From (13a) and definition (8), we obtain

$$l_\ell(\lambda)^* \Phi l_\ell(\lambda) = l_1(\lambda)^* \Phi_r l_1(\lambda) \cdot l_{\ell-1}(\lambda)^* R l_{\ell-1}(\lambda).$$

Solving this equation for the second term of the right-hand side yields

$$l_{\ell-1}(\lambda)^* R l_{\ell-1}(\lambda) = \sum_{i=1}^{\ell} (\beta_i - \alpha_i) \prod_{j \neq i} (\lambda - \beta_j)(\lambda - \alpha_j), \quad (14)$$

which can be shown to satisfy (13b) by induction on ℓ . ■

Remark 1: Note that $\alpha_1 = -\infty$ can be considered by replacing the term $(\lambda - \alpha_1)$ in (10) by 1. Similarly, for $\beta_\ell = \infty$, $(\lambda - \beta_\ell)$ should be replaced -1 .

Corollary 1: The matrices Φ, Ψ given by (11) satisfy Assumptions 1 and 2.

Proof: Using the following common congruence transformation of Φ_r, Ψ_r :

$$\Phi_r = T_r^* \Phi_o T_r, \quad \Psi_r = T_r^* \Psi_o T_r,$$

with

$$T_r = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ i & i \end{bmatrix}, \quad \Psi_o = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

we can write Φ, Ψ given by (11) in the form (3) with Ψ_o given above and $T = T_r \tilde{T}$.

Next, we show that the matrix $R \in \mathbb{S}^\ell$ obtained from decomposition (13a) satisfies (5). Condition (5a) readily follows from (13b). To show (5b) we consider arbitrary $\lambda_i, \lambda_j \in \mathbb{L}_T(s)$, $i, j \in \mathbb{I}_\ell$, $i \neq j$, for some arbitrary $s \in \Lambda(\Phi_o, \Psi_o)$. From the root-locus argument in the proof of Lemma 1 we note that $\lambda_i \in [\alpha_i, \beta_i]$ and $\lambda_j \in [\alpha_j, \beta_j]$, and hence $\lambda_i \neq \lambda_j$. Using the decomposition $\Phi = T^* \Phi_o T$ and result (6c), we obtain

$$l_\ell(\lambda_i)^* \Phi l_\ell(\lambda_j) = t_2(\lambda_i)^* t_2(\lambda_j) \cdot l_1(s)^* \Phi_o l_1(s) = 0,$$

for $s \neq \infty$, while for $s = \infty$

$$l_\ell(\lambda_i)^* \Phi l_\ell(\lambda_j) = t_1(\lambda_i)^* t_1(\lambda_j) \cdot l_1(s)^* \Phi_o l_1(s) = 0.$$

Elaborating the left-hand side using (13a) yields

$$l_\ell(\lambda_i)^* \Phi l_\ell(\lambda_j) = l_1(\lambda_i)^* \Phi_r l_1(\lambda_j) \cdot l_{\ell-1}(\lambda_i)^* R l_{\ell-1}(\lambda_j), \\ = i(\lambda_i - \lambda_j) \cdot l_{\ell-1}(\lambda_i)^* R l_{\ell-1}(\lambda_j).$$

As this must equal 0 while $\lambda_i \neq \lambda_j$, (5b) holds. \blacksquare

To transfer the results above to the union of ℓ nonempty, nonsingleton, nonintersecting segments $\Lambda(\phi, \psi_i)$ on an arbitrary circle or line $\Lambda(\phi, 0)$, we compute a bijective mapping between $\Lambda(\phi, 0)$ and the real axis $\Lambda(\Phi_r, 0)$. To this end, we pick one, say the ℓ th, segment $\Lambda(\phi, \psi_\ell)$ and compute the matrix $T \in \mathbb{C}^{2 \times 2}$ such that

$$\phi = T^* \Phi_r T, \quad \psi_\ell = T^* \Psi_r T.$$

This matrix defines a bijective mapping $\tau(\lambda)$ between $\lambda \in \Lambda(\phi, 0)$ and $\tau \in \Lambda(\Phi_r, 0)$, similar to (6), [4]. $\Lambda(\phi, \psi_\ell)$ is mapped onto \mathbb{R}_+ , while the other $\ell - 1$ segments $\Lambda(\phi, \psi_i)$, $i \in \mathbb{I}_{\ell-1}$ are mapped onto nonintersecting intervals of \mathbb{R}_- . Let $\tilde{\Phi}, \tilde{\Psi} \in \mathbb{H}^{\ell+1}$ and $\tilde{R} \in \mathbb{H}^\ell$ be the result of Lemma 1 applied to the image of $\bigcup_{i=1}^\ell \Lambda(\phi, \psi_i)$ under the mapping $\tau(\lambda)$. Let in addition the matrices $\mathbf{T}_k \in \mathbb{R}^{(k+1) \times (k+1)}$ be defined as follows:

$$\mathbf{T}_k = \begin{bmatrix} T_1 \star T_1 \star \dots \star T_1 \star T_1 \\ T_1 \star T_1 \star \dots \star T_1 \star T_2 \\ \vdots \\ T_1 \star T_2 \star \dots \star T_2 \star T_2 \\ T_2 \star T_2 \star \dots \star T_2 \star T_2 \end{bmatrix},$$

where T_1 and T_2 respectively denote the first and second row of T and \star is the vector convolution. Then, it is readily verified that $\tilde{\Phi}, \tilde{\Psi} \in \mathbb{H}^{\ell+1}$ and $R \in \mathbb{H}^\ell$ defined by

$$\tilde{\Phi} = \mathbf{T}_\ell^* \tilde{\Phi} \mathbf{T}_\ell, \quad \tilde{\Psi} = \mathbf{T}_\ell^* \tilde{\Psi} \mathbf{T}_\ell, \quad R = \mathbf{T}_{\ell-1}^* \tilde{R} \mathbf{T}_{\ell-1},$$

satisfy $\Lambda(\tilde{\Phi}, \tilde{\Psi}) = \bigcup_{i=1}^\ell \Lambda(\phi, \psi_i)$, $\tilde{\Phi} = J_\ell^*(\phi \otimes R) J_\ell$, and $l_{\ell-1}(\lambda)^* R l_{\ell-1}(\lambda) > 0$, for all $\lambda \in \Lambda(\phi, 0)$.

III. GENERALIZED KYP LEMMA

A. Preliminary Results

For given $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{n \times m}$ and $\ell \in \mathbb{I}$, the matrices $F_\ell \in \mathbb{C}^{(\ell+1)n \times (n+m\ell)}$ and $G_\ell \in \mathbb{C}^{\ell(n+m) \times (n+m\ell)}$ are defined as follows:

$$F_\ell = \begin{bmatrix} A^\ell & A^{\ell-1}B & A^{\ell-2}B & \dots & B \\ A^{\ell-1} & A^{\ell-2}B & \dots & B & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A & B & 0 & \dots & 0 \\ I & 0 & \dots & \dots & 0 \end{bmatrix}, \quad (15a)$$

$$G_\ell = \begin{bmatrix} A^{\ell-1} & A^{\ell-2}B & A^{\ell-3}B & \dots & B & 0 \\ 0 & 0 & 0 & \dots & 0 & I \\ A^{\ell-2} & A^{\ell-3}B & \dots & B & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ A & B & 0 & \dots & \dots & 0 \\ 0 & 0 & I & 0 & \dots & 0 \\ I & 0 & \dots & \dots & \dots & 0 \\ 0 & I & 0 & \dots & \dots & 0 \end{bmatrix}. \quad (15b)$$

In addition, we will use the mappings $H_0(\lambda) = (\lambda I - A)^{-1}B$, and $H_\ell(\lambda) : \mathbb{C} \rightarrow \mathbb{C}^{(n+m\ell) \times m}$, $\ell \in \mathbb{I}$, given by

$$H_\ell(\lambda) = \begin{bmatrix} (\lambda I - A)^{-1}B \\ I \\ \lambda I \\ \vdots \\ \lambda^{\ell-1}I \end{bmatrix}, \quad (16)$$

and $H_0(\infty) = B$ and we set $H_\ell(\infty) = [0 \dots 0 I_m]^T$ for $\ell \in \mathbb{I}$. For a given $T \in \mathbb{C}^{2 \times (\ell+1)}$, the mapping $\hat{H}_\ell(s) : \mathbb{C} \rightarrow \mathbb{C}^{(n+m\ell) \times m\ell}$ is defined as

$$\hat{H}_\ell(s) = [H_\ell(\lambda_1) \quad \dots \quad H_\ell(\lambda_\ell)], \quad (17)$$

with $\{\lambda_1, \dots, \lambda_\ell\} = \mathbb{L}_T(s)$.

For the matrices and mappings defined above, the following relations are readily verified:

$$F_\ell H_\ell(\lambda) = l_\ell(\lambda) \otimes H_0(\lambda), \quad (18a)$$

$$G_\ell H_\ell(\lambda) = l_{\ell-1}(\lambda) \otimes H_1(\lambda). \quad (18b)$$

In addition, we will need the following two lemma's to prove our main result.

Lemma 2: Let $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{n \times m}$, $T \in \mathbb{C}^{2 \times (\ell+1)}$, $x \in \mathbb{C}^{n+m\ell}$, and $s \in \mathbb{C}$ be given. Let F_ℓ and $H_\ell(\lambda)$ be defined by (15a) and (16), respectively, and let λ_i with $i \in \mathbb{I}_\ell$ be the roots of (4): $\mathbb{L}_T(s) = \{\lambda_1, \dots, \lambda_\ell\}$. Suppose that T has full row rank, that the roots λ_i are all distinct, and that $\det(\lambda_i I - A) \neq 0$ for all $i \in \mathbb{I}_\ell$. Then, there exists $w \in \mathbb{C}^n$ satisfying

$$\begin{bmatrix} s \\ 1 \end{bmatrix} \otimes w = (T \otimes I_n) F_\ell x, \quad (19)$$

if and only if there exists $u \in \mathbb{C}^{m\ell}$ such that

$$x = \hat{H}_\ell(s) u. \quad (20)$$

The proof of this lemma is presented in Appendix B.

Lemma 3: Let $\Phi_o, \Psi_o \in \mathbb{H}^2$ of the form (3b) be given, as well as $X, Y \in \mathbb{C}^{n \times m}$. Then

$$[X \quad Y] (\Phi_o \otimes I) \begin{bmatrix} X^* \\ Y^* \end{bmatrix} = 0, \quad (21a)$$

$$[X \quad Y] (\Psi_o \otimes I) \begin{bmatrix} X^* \\ Y^* \end{bmatrix} \succeq 0, \quad (21b)$$

hold if and only if X and Y can be factored as

$$X = W \text{diag}(s_1, \dots, s_m) V^*, \quad Y = W V^*, \quad (22)$$

with some $W \in \mathbb{C}^{n \times m}$, unitary $V \in \mathbb{C}^{m \times m}$, and $s_i \in \Lambda(\Phi_o, \Psi_o)$ for all $i \in \mathbb{I}_m$.

This lemma follows from [6, lemma 5] and [7, lemma 5].

B. The Main Theorem

Theorem 1: Let Hermitian matrices $\Phi, \Psi \in \mathbb{H}^{\ell+1}$ and $\Theta \in \mathbb{H}^{m+n}$, and matrices $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{n \times m}$ be given. Suppose Φ and Ψ satisfy Assumptions 1 and 2, and A has no eigenvalues on $\Lambda(\Phi, \Psi)$. Let $R \in \mathbb{H}^\ell$ be a matrix satisfying (5). Then, the following statements are equivalent.

(i) The inequality

$$\begin{bmatrix} (\lambda I - A)^{-1}B \\ I \end{bmatrix}^* \Theta \begin{bmatrix} (\lambda I - A)^{-1}B \\ I \end{bmatrix} \prec 0, \quad (23)$$

holds for all $\lambda \in \Lambda(\Phi, \Psi)$.

(ii) There exist $P, Q \in \mathbb{H}^n$ that satisfy $Q \succ 0$ and

$$F_\ell^*(\Phi \otimes P + \Psi \otimes Q)F_\ell + G_\ell^*(R \otimes \Theta)G_\ell \prec 0. \quad (24)$$

Proof: To prove that (ii) implies (i) we multiply (24) by $H_\ell(\lambda)^*$ on the left and $H_\ell(\lambda)$ on the right. On account of relations (18) and the mixed-product property of the matrix Kronecker product, this yields

$$\begin{aligned} & l_\ell(\lambda)^* \Phi l_\ell(\lambda) \cdot H_0(\lambda)^* P H_0(\lambda) \\ & + l_\ell(\lambda)^* \Psi l_\ell(\lambda) \cdot H_0(\lambda)^* Q H_0(\lambda) \\ & + l_{\ell-1}(\lambda)^* R l_{\ell-1}(\lambda) \cdot H_1(\lambda)^* \Theta H_1(\lambda) \prec 0. \end{aligned}$$

By definition (1), the first terms equals zero for all $\lambda \in \Lambda(\Phi, \Psi)$, and the second term is positive semi-definite. In addition, $l_{\ell-1}(\lambda)^* R l_{\ell-1}(\lambda) > 0$ for all $\lambda \in \Lambda(\Phi, \Psi)$ on account of Assumption 2, and hence, (23) holds.

We prove that (i) implies (ii) by contradiction, using a theorem of alternatives [8, Theorem 1] [9, Theorem 1.3]: from the infeasibility certificate of (24) we construct a $\lambda \in \Lambda(\Phi, \Psi)$ that violates (23). Suppose (24) is infeasible. Then there exists a nonzero positive semidefinite matrix $Z \in \mathbb{H}^{n+m\ell}$ such that

$$\text{tr}(G_\ell^*(R \otimes \Theta)G_\ell Z) \geq 0, \quad (25a)$$

$$\tilde{F}_\ell(\bar{\Phi} \otimes Z)\tilde{F}_\ell^* = 0, \quad (25b)$$

$$\tilde{F}_\ell(\bar{\Psi} \otimes Z)\tilde{F}_\ell^* \succeq 0. \quad (25c)$$

The matrix $\tilde{F}_\ell \in \mathbb{H}^{n \times (\ell+1)(n+m\ell)}$ is defined as

$$\tilde{F}_\ell = [F_1^\ell \quad \cdots \quad F_{\ell+1}^\ell],$$

where $F_i^\ell \in \mathbb{H}^{n \times (n+m\ell)}$ with $i \in \mathbb{I}_{\ell+1}$ denote the block rows of \tilde{F}_ℓ :

$$F_i^\ell = \begin{bmatrix} F_1^\ell \\ \vdots \\ F_{\ell+1}^\ell \end{bmatrix}.$$

Let $Z = \Upsilon \Upsilon^*$ with $\Upsilon \in \mathbb{C}^{(n+m\ell) \times r}$ be a full-rank factorization of Z , where r is the rank. Equality (25b) and inequality (25c) can then be written as

$$\tilde{E}_\ell(\bar{\Phi} \otimes I_r)\tilde{E}_\ell^* = 0, \quad \tilde{E}_\ell(\bar{\Psi} \otimes I_r)\tilde{E}_\ell^* \succeq 0, \quad (26)$$

where

$$\tilde{E}_\ell = [F_1^\ell \Upsilon \quad \cdots \quad F_{\ell+1}^\ell \Upsilon].$$

Making use of the decomposition (3), the equality and inequality in (26) are of the form (21) with

$$\begin{bmatrix} X & Y \end{bmatrix} = \tilde{E}_\ell(T^\top \otimes I), \quad \rightarrow \quad \begin{bmatrix} X \\ Y \end{bmatrix} = (T \otimes I)F_\ell \Upsilon.$$

By Lemma 3, matrices X, Y admit a decomposition of the form (22), which yields

$$(T \otimes I)F_\ell \Upsilon = \begin{bmatrix} W \text{diag}(s_k) \\ W \end{bmatrix} V^*,$$

for some $W \in \mathbb{C}^{n \times r}$, unitary $V \in \mathbb{C}^{r \times r}$, and $s_k \in \Lambda(\Phi_o, \Psi_o)$ with $k \in \mathbb{I}_r$. Let w_k and v_k respectively denote the k th column of W and V , then this equality corresponds to

$$\begin{bmatrix} s_k \\ 1 \end{bmatrix} \otimes w_k = (T \otimes I)F_\ell \Upsilon v_k, \quad \forall k \in \mathbb{I}_r.$$

Then by Lemma 2, there exist $u_k \in \mathbb{C}^{m\ell}$ with $k \in \mathbb{I}_r$ such that

$$\Upsilon v_k = \hat{H}_\ell(s_k) u_k.$$

Note that u_k must be nonzero since Υ has full column rank and v_k is nonzero. Because V is unitary, we have

$$Z = \Upsilon V V^* \Upsilon^* = \sum_{k=1}^r (\hat{H}_\ell(s_k) u_k) (\hat{H}_\ell(s_k) u_k)^*.$$

Substituting this decomposition into (25a) yields

$$\sum_{k=1}^r u_k^* M_k u_k \geq 0, \quad (27a)$$

where

$$M_k = (G_\ell \hat{H}_\ell(s_k))^* (R \otimes \Theta) (G_\ell \hat{H}_\ell(s_k)). \quad (27b)$$

Elaborating the (i, j) th $m \times m$ subblock of M_k , denoted by $M_{i,j}^k$, using relation (18b) yields

$$\begin{aligned} M_{i,j}^k &= (G_\ell H_\ell(\lambda_i^k))^* (R \otimes \Theta) (G_\ell H_\ell(\lambda_j^k)) \\ &= l_{\ell-1}(\lambda_i^k)^* R l_{\ell-1}(\lambda_j^k) \cdot H_1(\lambda_i^k)^* \Theta H_1(\lambda_j^k), \end{aligned}$$

where λ_i^k with $i \in \mathbb{I}_\ell$ denote the ℓ elements of $\mathbb{L}(s_k)$. Since $s_k \in \Lambda(\Phi_o, \Psi_o)$, $\lambda_i^k \in \Lambda(\Phi, \Psi)$. On account of Assumption 2, we have $M_{i,j}^k = 0$ for $i \neq j$, while $M_{i,i}^k$ is a positive-scalar multiple of $H_1(\lambda_i^k)^* \Theta H_1(\lambda_i^k)$. As a result, (27) implies that at least for one (k, i) , $H_1(\lambda_i^k)^* \Theta H_1(\lambda_i^k) \prec 0$ cannot hold, and we reach a contradiction with statement (i). \blacksquare

IV. COMPARISON TO EXISTING RESULTS

A. Generalized KYP Lemma of [4], [5]

The generalization of the KYP lemma presented in [4], [5] provides an LMI equivalent for an inequality of the form (23) on a segment of a circle or straight line in \mathbb{C} , i.e. a curve of the form $\Lambda(\Phi, \Psi)$ for $\Phi, \Psi \in \mathbb{H}_2$ that satisfy Assumption 1. In case the inequality must hold on a union of ℓ segments, an LMI reformulation can be obtained by applying the result of [4], [5] for each of the intervals separately. This LMI involves 2ℓ matrix variables in \mathbb{H}_n , and comprises ℓ inequalities in \mathbb{H}_n and ℓ inequalities in \mathbb{H}_{n+m} . Theorem 1 provides an alternative LMI condition, involving only 2 matrix variables in \mathbb{H}_n and comprising one inequality in \mathbb{H}_n , and one in $\mathbb{H}_{n+m\ell}$. As this LMI is generally smaller in both dimension and number of variables, it is computationally advantageous. To illustrate this advantage, we apply both approaches to a finite impulse response (FIR) filter design example adopted from [4]. Herein, the coefficients $h_i, i = 0, \dots, n$ of the filter

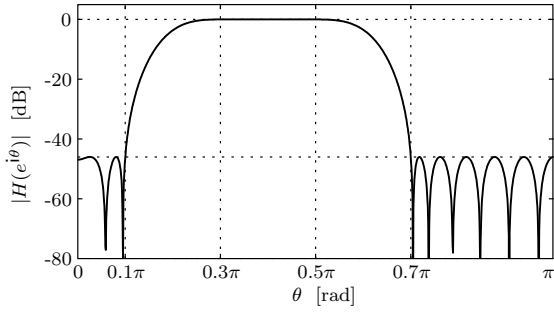


Fig. 1. Solution of the optimal filter design example.

$H(z) = \sum_{i=0}^n h_i z^{-i}$ are computed by solving the following problem:

$$\text{minimize } \varepsilon_p \quad (28a)$$

$$\text{subject to } |H(e^{j\theta})| \leq \varepsilon_s, \quad \forall \theta \in [-\theta_{sl}, \theta_{sl}] \cup [\theta_{sh}, 2\pi - \theta_{sh}] \quad (28b)$$

$$|H(e^{j\theta}) - e^{-id\theta}| \leq \varepsilon_p, \quad \forall \theta \in [\theta_{pl}, \theta_{ph}] \quad (28c)$$

where we use $\varepsilon_s = 5 \cdot 10^{-3}$, $\theta_{sl} = 0.1\pi$, $\theta_{sh} = 0.7\pi$, $\theta_{pl} = 0.3\pi$, $\theta_{ph} = 0.5\pi$, and delay $d = 15$. A filter of length $n = 30$ is designed. The LMI equivalents to the semi-infinite constraints are parsed with YALMIP [10] and solved with SDPT3 [11]. The optimal $\varepsilon_p = 2.5 \cdot 10^{-4}$ and Figure 1 shows the amplitude response of the optimal filter. SDPT3 requires 23 CPU seconds to solve the semidefinite program that results from applying the results of [4], [5] to each of the two intervals considered in (28b), while the LMI obtained by Theorem 1 is solved in 15 CPU seconds. In both cases, all LMI variables could be selected real symmetric instead of complex Hermitian without introducing conservatism [12].

B. Sum-of-Squares Result

As described in [13] for the result of [4], [5], we can apply Theorem 1 to obtain a sum-of-squares condition for the positivity of a univariate polynomial on a union of intervals. The ring of polynomials in $\lambda \in \mathbb{R}$ is denoted by $\mathbb{R}[\lambda]$ and Σ is its subset of sum-of-squares polynomials:

$$\Sigma = \left\{ \sum_i p_i^2 : p_i \in \mathbb{R}[\lambda] \right\}.$$

With $A \in \mathbb{R}^{mn \times mn}$, $B \in \mathbb{R}^{mn \times m}$ given as follows:

$$[A \quad B] = \begin{bmatrix} 0 & I_m & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & I_m & 0 \\ 0 & \cdots & \cdots & 0 & I_m \end{bmatrix}, \quad (29)$$

we obtain

$$\begin{bmatrix} (\lambda I - A)^{-1} B \\ I \end{bmatrix} = \frac{1}{\lambda^n} \hat{l}_n(\lambda) \otimes I_m,$$

with

$$\hat{l}_n(\lambda) = [1 \quad \lambda \quad \cdots \quad \lambda^n]^\top.$$

When applied to a scalar polynomial ($m = 1$), this yields the following result:

Corollary 2: Let the polynomial $\theta \in \mathbb{R}[\lambda]$ of degree $2n$ be given, as well as 2ℓ scalars $\alpha_i, \beta_i \in \mathbb{R}$, $i \in \mathbb{I}_\ell$ that satisfy (9). Set

$$r(\lambda) = l_{\ell-1}(\lambda)^* R l_{\ell-1}(\lambda),$$

as in (14). Then, the following statements are equivalent:

- (i) $\theta(\lambda) \geq 0$ for all $\lambda \in \bigcup_{i=1}^\ell [\alpha_i, \beta_i]$.
- (ii) There exist $z \in \Sigma$ and $q \in \Sigma$ of degrees $2(n + \ell - 1)$ and $2(n - 1)$, respectively, such that

$$r(\lambda)\theta(\lambda) + q(\lambda)\psi(\lambda) = z(\lambda), \quad (30)$$

where

$$\psi(\lambda) = \prod_{i=1}^\ell (\lambda - \alpha_i)(\lambda - \beta_i).$$

Proof: This result directly follows from Theorem 1 with A, B given in (29), Φ, Ψ obtained from Lemma 1 and $\Theta \in \mathbb{S}^{n+1}$ an arbitrary symmetric matrix that satisfies

$$\theta(\lambda) = \hat{l}_n(\lambda)^* \Theta \hat{l}_n(\lambda). \quad (31)$$

Let P, Q be a solution of (24) and set Z equal to the left-hand side of (24), then (30) holds for $q(\lambda) = \hat{l}_{n-1}(\lambda)^* Q \hat{l}_{n-1}(\lambda)$ and $z(\lambda) = \hat{l}_{n+\ell-1}(\lambda)^* Z \hat{l}_{n+\ell-1}(\lambda)$. ■

Remark 2: To gain some insight in the role of the matrix P , we use (13a) and definitions (15) to elaborate its contribution in (24):

$$F_\ell^* (\Phi \otimes P) F_\ell = G_\ell^* (R \otimes (F_1^* (\Phi_r \otimes P) F_1)) G_\ell.$$

As noted in [13], with A, B given by (29), $\Theta(P) = \Theta + F_1^* (\Phi_r \otimes P) F_1$ with $P \in \mathbb{H}^n$ corresponds to an explicit parametrization of all matrices that satisfy (31).

Condition (30) differs from current sum-of-squares certificates for positivity of a univariate polynomial on a union of intervals. Let $V \subset \mathbb{R}[\lambda]$ be given as:

$$V = \{(\beta_\ell - \lambda)(\lambda - \alpha_1), (\lambda - \beta_1)(\lambda - \alpha_2), \dots, (\lambda - \beta_{\ell-1})(\lambda - \alpha_\ell)\},$$

and the corresponding preordering T_V :

$$T_V = \left\{ \sum_{e \in \{0,1\}^\ell} s_e(\lambda) \prod_{i=1}^\ell v_i(\lambda)^{e_i} : s_e(\lambda) \in \Sigma \right\},$$

where $v_i(\lambda)$ are the elements of V . Let $T_{V,n}$ be the subset of T_V obtained by limiting the degree of the terms in the summation to $2n$. Then, the following result is proven in [14]:

Lemma 4: Let V and $T_{V,n}$ be as defined above. Then, a polynomial $\theta(\lambda) \in \mathbb{R}[\lambda]$ of degree $2n$ is nonnegative on $\bigcup_{i=1}^\ell [\alpha_i, \beta_i]$ if and only if $\theta(\lambda) \in T_{V,n}$,

V. CONCLUSION

This paper generalizes the KYP lemma to curves characterized by a particular polynomial equality and inequality in \mathbb{C} . As a first application of our result we show that the considered set of curves includes the union of segments on a circle or line as a special case. When dealing with these special cases, the novel LMI condition is computationally advantageous compared to multiple applications of the result

of [4], [5]. Currently, we are investigating additional applications of our result as well as its potential in cases where assumption 2 does not hold.

APPENDIX B

Proof of Lemma 2: For the sake of brevity the proof is restricted to the case $s \neq \infty$. Condition (19) is reformulated as

$$Z \begin{bmatrix} x \\ w \end{bmatrix} = 0,$$

where

$$Z = \left[(T \otimes I_n) F_\ell \quad - \begin{bmatrix} s \\ 1 \end{bmatrix} \otimes I_n \right].$$

To prove our result, we show that the nullspace of $Z \in \mathbb{C}^{2n \times (2n+m\ell)}$ is spanned by the columns of $N \in \mathbb{C}^{(2n+m\ell) \times m\ell}$, given by:

$$N = \begin{bmatrix} H_\ell(\lambda_1) & \cdots & H_\ell(\lambda_\ell) \\ t_2(\lambda_1)H_0(\lambda_1) & \cdots & t_2(\lambda_\ell)H_0(\lambda_\ell) \end{bmatrix}.$$

This way, $\begin{bmatrix} x \\ w \end{bmatrix}$ satisfies (19) if and only if it can be expressed as Nu for some $u \in \mathbb{C}^{m\ell}$. That is, x is given by (20) and

$$w = [t_2(\lambda_1)H_0(\lambda_1) \quad \cdots \quad t_2(\lambda_\ell)H_0(\lambda_\ell)] u.$$

The columns of N indeed lie in the nullspace of Z , i.e. $ZN = 0$, as for each root λ_i

$$\begin{aligned} Z \begin{bmatrix} H_\ell(\lambda_i) \\ t_2(\lambda_i)H_0(\lambda_i) \end{bmatrix} \\ = (T \otimes I_n) F_\ell H_\ell(\lambda_i) - t_2(\lambda_i) \begin{bmatrix} s \\ 1 \end{bmatrix} \otimes H_0(\lambda_i) = 0, \end{aligned}$$

since

$$\begin{aligned} (T \otimes I_n) F_\ell H_\ell(\lambda_i) &= (T \otimes I_n) (t_\ell(\lambda_i) \otimes H_0(\lambda_i)), \\ &= (T t_\ell(\lambda_i)) \otimes H_0(\lambda_i), \\ &= t_2(\lambda_i) \begin{bmatrix} s \\ 1 \end{bmatrix} \otimes H_0(\lambda_i). \end{aligned}$$

The first equality follows from (18a), while the second one results from the mixed-product property of the matrix Kronecker product. The third equality follows from (6c).

The fact that the columns of N constitute a basis for the nullspace of Z follows from N having full column rank and Z being of full row rank. The $m\ell \times m\ell$ central part of N equals $L \otimes I_m$ with

$$L = [\hat{l}_{\ell-1}(\lambda_1) \quad \cdots \quad \hat{l}_{\ell-1}(\lambda_\ell)].$$

As the ℓ roots λ_i are distinct, the Vandermonde matrix L is nonsingular and hence, N has full column rank. The fact that Z has full row rank is shown by contradiction. Suppose the rows of Z are not linearly independent. Then there exists $v \in \mathbb{C}^{1 \times 2n}$, $v \neq 0$ for which $vZ = 0$. To satisfy the last n entries of this equality, v must be of the form

$$v = [v_1 \quad -sv_1],$$

for some $v_1 \in \mathbb{C}^{1 \times n}$, $v_1 \neq 0$. The first n entries of $vZ = 0$ then read as $v_1 M = 0$ with

$$M = [1 \quad -s] (T \otimes I_n) \begin{bmatrix} A^\ell \\ \vdots \\ A \\ I \end{bmatrix}.$$

Let α_i with $i \in \mathbb{I}_\ell$ denote the eigenvalues of A . Then the eigenvalues μ_i of M are given by

$$\mu_i = [1 \quad -s] T l_\ell(\alpha_i).$$

As we assumed that $\alpha_i \notin \mathbb{L}_T(s)$ for all $i \in \mathbb{I}_\ell$, the corresponding $\mu_i \neq 0$. Hence, $v_1 M = 0$ cannot hold for some $v_1 \neq 0$, and a contradiction is reached. ■

REFERENCES

- [1] R. E. Kalman, "Lyapunov functions for the problem of Lur'e in automatic control," *Proc. of the National Academy of Sciences of the United States of America*, vol. 49, no. 2, pp. 201–205, 1963.
- [2] V. A. Yakubovich, "Solution of certain matrix inequalities encountered in non-linear regulation theory," *Doklady Akademii Nauk SSSR*, vol. 143, pp. 1304–1307, 1962, english translation in *Soviet mathematics - Doklady*, vol. 3, 1962, pp. 620–623.
- [3] V. M. Popov, "Absolute stability of nonlinear systems of automatic control," *Automation and Remote Control*, vol. 22, pp. 857–875, 1962, russian original in August 1961.
- [4] T. Iwasaki and S. Hara, "Generalized KYP lemma: unified frequency domain inequalities with design applications," *IEEE Transactions on Automatic Control*, vol. 50, no. 1, pp. 41–59, 2005.
- [5] C. W. Scherer, "LMI relaxations in robust control," *European Journal of Control*, vol. 12, no. 1, pp. 3–29, 2006.
- [6] A. Rantzer, "On the Kalman-Yakubovich-Popov lemma," *Systems and Control Letters*, vol. 28, no. 1, pp. 7–10, 1996.
- [7] T. Iwasaki, G. Meinsma, and M. Fu, "Generalized S-procedure and finite frequency KYP lemma," *Mathematical Problems in Engineering*, vol. 6, pp. 305–320, 2000.
- [8] V. Balakrishnan and L. Vandenberghe, "Semidefinite programming duality and linear time-invariant systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 1, pp. 30–41, 2003.
- [9] A. Berman and A. Ben-Israel, "More on linear inequalities with applications to matrix theory," *Journal of Mathematical Analysis and Applications*, vol. 33, pp. 482–496, 1971.
- [10] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. [Online]. Available: <http://users.isy.liu.se/johan/yalmip>
- [11] R. Tutuncu, K. Toh, and M. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming Ser. B*, vol. 95, pp. 189–217, 2003.
- [12] G. Pipeleers and L. Vandenberghe, "Generalized KYP lemma with real data," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2940–2944, 2011.
- [13] S. Hara and T. Iwasaki, "Sum-of-squares decomposition via generalized KYP lemma," *IEEE Transactions on Automatic Control*, vol. 54, no. 5, pp. 1025–1029, 2009.
- [14] L. Faybusovich, "On nesterov's approach to semi-infinite programming," *Acta Applicandae Mathematicae*, vol. 74, pp. 195–215, 2002.
- [15] M. Marshall, "Positive polynomials and sums of squares," ser. *Mathematical Surveys and Monographs*. American Mathematical Society, 2008, vol. 146.