A Generalized Orthonormal Basis for Linear Dynamical Systems

Peter S. C. Heuberger, Paul M. J. Van den Hof, *Member, IEEE*, and Okko H. Bosgra

***Abstract*—In many areas of signal, system, and control theory, orthogonal functions play an important role in issues of analysis and design. In this paper, it is shown that there exist orthogonal functions that, in a natural way, are generated by stable linear dynamical systems and that compose an orthonormal basis for the signal space ℓ_2^n. To this end, use is made of balanced realizations of inner transfer functions. The orthogonal functions can be considered as generalizations of, e.g., the pulse functions, Laguerre functions, and Kautz functions, and give rise to an alternative series expansion of rational transfer functions. It is shown how we can exploit these generalized basis functions to increase the speed of convergence in a series expansion, i.e., to obtain a good approximation by retaining only a finite number of expansion coefficients. Consequences for identification of expansion coefficients are analyzed, and a bound is formulated on the error that is made when approximating a system by a finite number of expansion coefficients.**

I. Introduction

CONSIDER a linear time-invariant stable discrete-time system G, represented by its proper transfer function $G(z)$ in the Hilbert space $\mathcal{H}_2$, i.e., $G(z)$ is analytic outside the unit circle, $|z| \geq 1$. A general and common representation of $G(z)$ is in terms of its Laurent expansion around $z = \infty$, as

$$G(z) = \sum_{k=0}^{\infty} G_k z^{-k} \tag{1}$$

with $\{G_k\}_{k=0,1,\cdots}$ the sequence of Markov parameters.

In constructing this series expansion we have employed a set of orthogonal functions: $\{z^0, z^{-1}, z^{-2}, \cdots\}$, where orthogonality is considered in terms of the inner product in $\mathcal{H}_2$. In a generalized form we can write (1) as

$$G(z) = \sum_{k=0}^{\infty} L_k f_k(z) \tag{2}$$

with $\{f_k(z)\}_{k=0,1,2,\cdots}$ a sequence of orthogonal functions.

Manuscript received November 17, 1992; revised December 3, 1993. recommended by Past Associate Editor, B. Pasik-Duncan. This work was supported in part by Shell Research B.V., The Hague, and the Center for Industrial Control Science, The University of Newcastle, Newcastle NSW, Australia.

P. S. C. Heuberger was with the Mechanical Engineering Systems and Control Group, Delft University of Technology, 2 628 CD Delft, The Netherlands and is now with the Dutch National Institute of Public Health and Environmental Protection (RIVM), P.O. Box 1, 3720 BA Bilthoven, The Netherlands.

P. M. J. Van den Hof and O. H. Bosgra are with the Mechanical Engineering Systems and Control Group, Delft University of Technology, Mekelweg 2, 2 628 CD Delft, The Netherlands.

IEEE Log Number 9408269.

There are a number of research areas that deal with the question of either approximating a given system G with a finite number of coefficients in a series expansion as in (2), or (approximately) identifying an unknown system in terms of a finite number of expansion coefficients through

$$\hat{G}(z) = \sum_{k=0}^{N} \hat{L}_k f_k(z). \tag{3}$$

The problem that will be analyzed in this paper is the following.

Can we construct a sequence of orthogonal basis functions $\{f_{k,\bar{G}}(z)\}_{k=0,\cdots\infty}$ with $\bar{G} \in \mathcal{H}_2$, such that

a) to some extent, the basis can be adapted to a linear stable system $\bar{G}$ to be described, implying that $\bar{G}$ can be accurately described by only a small number of coefficients in the expansion, and
b) the basis allows the construction of an error bound for the approximation of a linear stable system G by a finite length expansion in the basis $f_{k,\bar{G}}$, i.e., an upper bound on $\|G(z) - \sum_{k=0}^{N} L_k f_{k,\bar{G}}(z)\|$ in some prechosen norm, whenever G and $\bar{G}$ do not match exactly.

The use of orthogonal functions with the aim of adapting the system and signal representation to the specific properties of the systems and signals at hand has a long history. The classical work of Lee and Wiener during the 1930's on network synthesis in terms of Laguerre functions [24], [46] is summarized in [25]. Laguerre functions have been used in the 1950's and 1960's to represent transient signals [45], [7]. During the past decades, the use of orthogonal functions has been studied in problems of filter synthesis [22], [30] and for system identification [23], [32], [31], [6] and approximation [35], [36]. In these approaches to system identification, the input and output signals are transformed to a (Laguerre) transformed domain and standard identification techniques are applied to the signals in this domain. Data reduction has been the main motivation in these studies. Identification of continuous-time models with the aid of orthogonal functions is considered in e.g., [38] and [29]. In recent years, a renewed interest in Laguerre functions has emerged. The approximation of (infinite dimensional) systems in terms of Laguerre functions has been considered in [27], [28], [12], [13], and [15]. In the identification of coefficients in finite length series expansions, Laguerre function representations have been considered from a statistical analysis point of view in [43], [42], and [16].

The use of Laguerre-function-based identification in adaptive control and controller tuning is studied in [47] and [9]. A second-order extension to the basic Laguerre functions using the so called Kautz functions [21] is subject of discussion in [41] and [44].

In this paper we will expand and generalize the orthogonal functions as basis functions for dynamical system representations. Specifically we will generalize the Laguerre functions and Kautz functions to a situation where a higher degree of flexibility is present in the choice of basis functions, and where consequently a smaller error bound as meant in part b) of the problem can be obtained. Laguerre functions are specifically appropriate for accurate modeling of systems with dominant first-order dynamics, whereas Kautz functions are directed toward systems with dominant second-order resonant dynamics. The generalized basis functions, introduced in this paper, will be suited also for systems with a wide range of dominant dynamics, i.e., dominant high frequency and low frequency behavior.

We will restrict attention to the transfer function space $\mathcal{H}_2$ being equipped with the usual inner product. This choice, rather than the $\mathcal{H}_\infty$-space where orthogonality is abandoned, is motivated by the fact that our main intended application of these results is in the area of approximate system identification. As the main stream of approaches in system identification is directed toward prediction error methods and the use of least-squares types of identification criteria, [26], the choice of a two-norm is quite straightforward and natural in this respect.

Note that the two problems a) and b) should be treated as a joint problem. One of the (trivial) solutions to problem a) only is the use of a Gram–Schmidt orthogonalization procedure on the impulse response of the system G itself [1]. In that case the system can be described by a series expansion of only one single term. In this situation, however, no results are available for part b) of the problem.

In an identification context, the use of the orthogonal functions as in (1) leads to the so-called *finite impulse response* (FIR)-model [26]

$$y(t) = \sum_{k=0}^{N} G_k(\theta)u(t-k) + \varepsilon(t) \tag{4}$$

where $\varepsilon(t)$ is the one-step-ahead prediction error, and $\{y(t), u(t)\}$ are samples of the output, input of the dynamical system to be identified. The identification of the unknown coefficients $\{G_k(\theta)\}_{k=0,\cdots,N}$ through least squares minimization of $\varepsilon(t)$ over the time interval is an identification method that has some favorable properties. First, it is a linear regression scheme, which leads to a simple analytical solution; second, it is of the type of output-error-method, which has the advantage that the input/output system $G(z)$ can be estimated consistently whenever the unknown noise disturbance on the output data is uncorrelated with the input signal [26].

It is well known, however, that for moderately damped systems, and/or in situations of high sampling rates, it may take a large value of N, the number coefficients to be estimated, to capture the essential dynamics of the system G into its model. If we would be able to improve the basis functions in such a way that an accurate description of the model to be estimated can be achieved by a small number of coefficients in a series expansion, then this is beneficial from both aspects of bias and variance of the model estimate.

For the series expansion in (1) with $f_k = z^{-k}$, it is straightforward to show that a system G will have a finite length series expansion if and only if all system poles are at $z = 0$. Moreover, in the scalar case the length of the expansion, i.e., the index of the last nonzero coefficient, equals the total number of poles at $z = 0$.

As a generalized situation, we can consider Laguerre polynomials [37] that are known to generate a sequence of orthogonal functions [14]

$$f_k(z) = \sqrt{1-a^2}\, z \frac{(1-az)^k}{(z-a)^{k+1}}, \qquad |a| < 1. \tag{5}$$

Similar to above, a system G will have a finite length series expansion if and only if all system poles are at $z = a$, with the length of the expansion being equal to the total number of poles at $z = a$.

In dealing with the problem of finding similar results for any general stable dynamical system $G(z)$, we have considered the question of whether a linear system in a natural way gives rise to a set of orthogonal functions. The answer to this question appears to be affirmative. It will be shown that every stable system gives rise to a complete set of orthonormal functions based on input (or output) balanced realizations, or equivalently based on a singular value decomposition of a corresponding Hankel matrix. These generalized orthogonal basis functions will be shown to provide solutions to problems a) and b).

In Section III we will first briefly state the main result of this paper. Next in Section IV it will be shown how inner functions generate two sets of orthonormal functions that are complete in the signal space ℓ_2. This is the basic ingredient of the main result. Next an interpretation of these results is given in terms of balanced state-space representations. After showing the relations of the new basis functions with existing ones, we will focus on the dynamics that implicitly are involved in the inner functions generating the basis. It will be shown that if the dynamics of a stable system match the dynamics of the inner function that generates the basis, then the representation of this system in terms of this basis becomes extremely simple. Consequences for a related identification and approximation problem are discussed in Section VIII.

Due to space limitations, a complete statistical analysis of the related system identification problems that result from these basis functions can not be given in this paper. A statistical analysis along similar lines as [43] and [44] is presented elsewhere [39].

The proofs of all results are collected in an appendix.

II. PRELIMINARIES

We will use the following notation.

$(\cdot)^T$ Transpose of a matrix.

$(\cdot)^*$	Complex conjugate transpose of a matrix.
$\mathbf{C}^{p\times m}$	Set of complex-valued matrices of dimension $p \times m$.
$\mathbf{R}^{p\times m}$	Real-valued matrix with dimension $p \times m$.
$\mathbf{Z}_+$	Set of nonnegative integers.
$\ell_2[0,\infty)$	Space of squared summable sequences on the time interval $\mathbf{Z}_+$.
$\ell_2^{m\times n}[0,\infty)$	Space of matrix sequences $\{F_k \in \mathbf{C}^{m\times n}\}_{k=0,1,2,\cdots}$ such that $\sum_{k=0}^{\infty} tr(F_k^* F_k)$ is finite.
$\mathcal{H}_2^{p\times m}$	Set of real $p \times m$ matrix functions, analytic for $\|z\| \geq 1$, that are squared integrable on the unit circle.
$\mathcal{RH}_2^{p\times m}$	Set of real rational $p \times m$ matrix functions, analytic for $\|z\| \geq 1$, that are squared integrable on the unit circle.
$\|\cdot\|_2$	Induced 2-norm or spectral norm of a constant matrix, i.e., its maximum singular value.
$\|\cdot\|_\infty$	H_∞-norm.
$Vec(\cdot)$	Vector-operation on a matrix, stacking its columns on top of each other.
$\otimes$	Kronecker matrix product.
$\mathcal{H}(G)$	(Block) Hankel matrix related to transfer function $G = \sum_{k=0}^{\infty} G_k z^{-k}$, defined by $\mathcal{H}_{ij}(G) = G_{i+j-1}$ being the (i,j)-block element.
e_i	ith Euclidian basis vector in $\mathbf{R}^n$.
I_n	$n \times n$ Identity matrix.

In this paper we will consider discrete-time signals and systems. A linear time-invariant finite-dimensional system will be represented by its rational transfer function $G \in \mathcal{RH}_2^{p\times m}$, with m the number of inputs in u, and p the number of outputs in y. State-space realizations will be considered of the form

$$x(k+1) = Ax(k) + Bu(k) \tag{6}$$

$$y(k) = Cx(k) + Du(k) \tag{7}$$

with $A \in \mathbf{C}^{n\times n}$, $B \in \mathbf{C}^{n\times m}$, $C \in \mathbf{C}^{p\times n}$, and $D \in \mathbf{C}^{p\times m}$. (A,B,C,D) is an n-dimensional realization of G if $G(z) = C(zI-A)^{-1}B + D$. A realization is stable if all eigenvalues of A lie strictly within the unit circle. If a realization is stable, the controllability gramian P and observability gramian Q are defined as the solutions to the Lyapunov equations $APA^* + BB^* = P$ and $A^*QA + C^*C = Q$, respectively. A stable realization is called (internally) balanced if $P = Q = \Sigma$, with $\Sigma = diag(\sigma_1, \cdots, \sigma_n)$, $\sigma_1 \geq \cdots \geq \sigma_n$, a diagonal matrix with the positive Hankel singular values as diagonal elements. A stable realization is called input balanced if $P = I$, $Q = \Sigma^2$, and output balanced if $P = \Sigma^2$, $Q = I$.

A system $G \in \mathcal{RH}_2^{p\times m}$ is called inner if it satisfies $G^T(z^{-1})G(z) = I$. As G is analytic outside and on the unit circle, it has a Laurent series expansion $\sum_{k=0}^{\infty} G_k z^{-k}$.

III. The Main Result

We will start the technical part of this paper by giving the basic result first and then consecutively give the analysis that provides the ingredients for making the result plausible.

Theorem 3.1: Let G be an $m \times m$ inner transfer function with McMillan degree $n > 0$, having a Laurent expansion $G(z) = \sum_{k=0}^{\infty} G_k z^{-k}$ and satisfying $\|G_0\|_2 < 1$, and let (A,B,C,D) be a balanced realization of $G(z)$. Denote

$$V_k(z) = z(zI-A)^{-1}BG^k(z). \tag{8}$$

Then the set of functions $\{e_i^T V_k(z)\}_{i=1,\cdots n;k=0,\cdots\infty}$ constitutes an orthonormal basis of the function space $\mathcal{H}_2^{1\times m}$.

A direct consequence of this theorem is the following corollary. □

Corollary 3.2: Let G be an inner function with McMillan degree n as in Theorem 3.1, with a corresponding sequence of basis functions $V_k(z)$. Then for every proper stable transfer function $H \in \mathcal{H}_2^{p\times m}$ there exist unique $D_s \in \mathbf{R}^{p\times m}$, and $L = \{L_k\}_{k=0,1,\cdots} \in \ell_2^{p\times n}[0,\infty)$, such that

$$H(z) = D_s + z^{-1}\sum_{k=0}^{\infty} L_k V_k(z). \tag{9}$$

We refer to D_s, L_k as the orthogonal expansion coefficients of $H(z)$. □

Note that due to the fact that $V_k(z)$ is an $n \times m$-matrix of transfer functions, the dimension of each L_k is $p \times n$.

IV. Orthonormal Functions Generated by Inner Transfer Functions

In this section we will show that a square and inner transfer function gives rise to an infinite set of orthonormal functions. This derivation is based on the fact that a singular value decomposition of the Hankel matrix associated to a linear system induces a set of left (right) singular vectors that are orthogonal. Considering the left (right) singular vectors as discrete time functions, they are known to be orthogonal in ℓ_2-sense, thus generating a number of orthogonal functions being equal to the McMillan degree of the corresponding system. We will embed an inner function with McMillan degree n into a sequence of inner functions with McMillan degree kn, for which the left (right) singular vectors of the Hankel matrix span a space with dimension kn. If we let $k \to \infty$ the set of left (right) singular vectors will yield an infinite number of orthonormal functions, which can be shown to be complete in ℓ_2.

First we have to recapitulate some properties of inner transfer functions.

Proposition 4.1: Let $G(z)$ be an inner transfer function with a Laurent expansion $G(z) = \sum_{k=0}^{\infty} G_k z^{-k}$. Then

$$\sum_{k=0}^{\infty} G_{k+i}^T G_k = I \qquad \text{for } i = 0; \tag{10}$$

$$= 0 \qquad \text{for } i > 0. \tag{11}$$

□

The Hankel matrix of an inner transfer function has some specific properties, reflected in the following two results.

Proposition 4.2: Let $G(z)$ be an inner function with McMillan degree $n > 0$. Then a singular value decomposition (svd) of $\mathcal{H}(G)$ satisfies

$$\mathcal{H}(G) = U_0 V_0^*$$

with $U_0, V_0 \in \boldsymbol{C}^{\infty \times n}$ unitary,[1] and the pair (U_0, V_0) is unique modulo postmultiplication with a unitary matrix $T \in \boldsymbol{C}^{n \times n}$. □

The proposition states that an inner transfer function has all Hankel singular values equal to one. For continuous-time systems, this is proven in [11]. The discrete-time version follows straightforwardly by applying a bilinear transformation.

Proposition 4.3: Let $G(z)$ be a square inner function, having a Laurent expansion $G(z) = \sum_{k=0}^{\infty} G_k z^{-k}$. Denote the block Toeplitz matrices

$$T_v = \begin{bmatrix} G_0 & G_1 & G_2 & \cdots & \cdots \\ 0 & G_0 & G_1 & G_2 & \cdots \\ 0 & 0 & G_0 & G_1 & \cdots \\ \vdots & \vdots & 0 & G_0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{12}$$

$$T_u = \begin{bmatrix} G_0 & 0 & 0 & \cdots & \cdots \\ G_1 & G_0 & 0 & \cdots & \cdots \\ G_2 & G_1 & G_0 & \ddots & \cdots \\ \vdots & \vdots & G_1 & G_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}. \tag{13}$$

Then

i) $T_v T_v^T = T_u^T T_u = I$;

ii) $T_v V_0 = T_u^T U_0 = 0$, for any unitary matrices U_0, V_0 satisfying $U_0 V_0^* = \mathcal{H}(G)$. □

Lemma 4.4: Let $G(z)$ be a square inner function with McMillan degree n. Then for all $k \in \boldsymbol{Z}_+$, $G^k(z)$ is an inner function with McMillan degree kn. □

The result in the lemma is quite straightforward if one realizes that an inner function has all poles within the unit circle, and all zeros outside the circle.

Considering Proposition 4.2, it follows that the rows of V_0^* and the columns of U_0, are n mutually orthonormal vectors of infinite dimension. Additionally Lemma 4.4 shows that we can construct an inner transfer function with increasing McMillan degree, by repeatedly multiplying the transfer function with itself, and thus implicitly creating an increasing number of orthogonal vectors. The following result shows how we can increase this number of vectors, by embedding the svd of $\mathcal{H}(G)$ into a sequence of svd's of $\mathcal{H}(G^k)$.

Theorem 4.5: Let $G(z)$ be a square inner function with McMillan degree $n > 0$. Then

a) There exist unitary matrices $U_i, V_i \in \boldsymbol{C}^{\infty \times n}$, $i = 0, 1, \cdots$, such that for every $0 \neq k \in \boldsymbol{Z}_+$, the matrices

$$\Gamma_k^o = [\, U_{k-1} \ \cdots \ U_1 \ U_0] \quad \text{and} \tag{14}$$

$$\Gamma_k^c = \begin{bmatrix} V_0^* \\ V_1^* \\ \vdots \\ V_{k-1}^* \end{bmatrix} \tag{15}$$

constitute a singular value decomposition of $\mathcal{H}(G^k)$, through

$$\mathcal{H}(G^k) = \Gamma_k^o \Gamma_k^c. \tag{16}$$

b) The matrix sequence $\{U_i, V_i\}_{i=0,1,\cdots}$ is unique up to postmultiplication of each U_i and V_i with one and the same unitary matrix.

c) Let $G(z)$ have a Laurent expansion $G(z) = \sum_{i=0}^{\infty} G_i z^{-i}$, and consider the block Toeplitz matrices T_u, T_v as in (12), (13) then the matrix sequence $\{U_i, V_i\}_{i=0,1,\cdots}$ satisfies

$$V_k^* = V_{k-1}^* T_v \tag{17}$$

$$U_k = T_u U_{k-1} \qquad \text{for } k = 1, 2, \cdots. \tag{18}$$

□

The theorem shows the construction of orthogonal matrices Γ_k^o, Γ_k^c that have a nesting structure. The suggested svd of $\mathcal{H}(G^k)$ incorporates svd's of $\mathcal{H}(G^i)$ for all $i < k$. In this way orthogonal matrices Γ_k^o and Γ_k^c are constructed with an increasing rank. Note that the restriction on the structure of the consecutive svd's is so strong that, according to b), given a singular value decomposition $\mathcal{H}(G) = U_0 V_0^*$, the matrix sequence $\{U_i, V_i, \ i = 1, 2, \cdots\}$ is uniquely determined. Note also that there is a clear duality between the controllability part Γ_k^c and the observability part Γ_k^o. To keep the exposition and the notation as simple as possible we will further restrict attention to the controllability part of the problem. Dual results exist for the observability part.

Proposition 4.6: Let $G(z)$ be an $m \times m$ inner function with McMillan degree $n > 0$, and consider any sequence of unitary matrices $\{V_i\}_{i=0,1,\cdots}$ satisfying (17) in Theorem 4.5. Denote for $k \in \boldsymbol{Z}_+$

$$V_k(z) = \sum_{i=0}^{\infty} M_k(i) z^{-i}, \text{ with } M_k(i) \in \boldsymbol{C}^{n \times m} \text{ defined by}$$

$$V_k^* =: [M_k(0) \ M_k(1) \ M_k(2) \ \cdots]. \tag{19}$$

Then

$$V_k(z) = V_0(z) G^k(z).$$

□

The proposition actually is a z-transform-equivalent of the result in Theorem 4.5. It shows the construction of the controllability matrix Γ_k^c.

In the next stage we show that this controllability matrix generates a sequence of orthogonal functions that is complete in ℓ_2^n.

Theorem 4.7: Let $G(z)$ be an $m \times m$ inner function with McMillan degree $n > 0$, such that $\|G_0\|_2 < 1$; consider a sequence of unitary matrices $\{V_i\}_{i=0,1,\cdots}$ as meant in Theorem 4.5. For each $k \in \boldsymbol{Z}_+$ consider the function $\phi_k : \boldsymbol{Z}_+ \to \boldsymbol{C}^n$, defined by

$$[\phi_k(0) \ \phi_k(1) \ \phi_k(2) \ \cdots] = V_k^*.$$

Then the set of functions $\Psi(G) := \{\phi_k\}_{k=0}^{\infty}$ constitutes an orthonormal basis of the signal space $\ell_2^n[0, \infty)$. □

[1] With slight abuse of notation we will use this notation to indicate an operator $\boldsymbol{C}^n \to \ell_2[0, \infty)$.

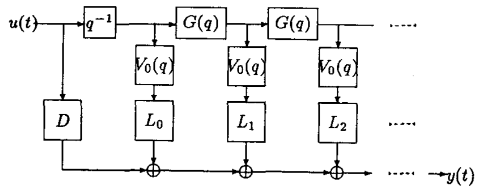


Fig. 1. Series expansion of a transfer function in terms of an orthonormal basis $\Psi(G)$.

Remark 4.8: This basis has been derived from the singular value decomposition of the Hankel matrix $\mathcal{H}(G)$. As stated in Proposition 4.2 this svd is unique up to postmultiplication of U_0, V_0 with a unitary matrix. Consequently—within this context—both V_k^*, $V_k(z)$, and the corresponding basis functions $\{\phi_k\}$ are unique up to unitary premultiplication.

For use later on we will formalize the class of inner functions that have the property as mentioned in the previous theorem.

Definition 4.9: We define the class of functions: $\mathcal{G}_1 := \{$all square inner functions G with McMillan degree > 0 such that $\|G_0\|_2 < 1\}$. □

As a result of the fact that the proposed orthonormal functions constitute a basis of ℓ_2^n, each square inner function generates an orthonormal basis that provides a unique transformation of ℓ_2^q-signals to an orthogonal domain. Similarly, when given such an orthonormal basis, each stable rational function can be expanded in a series expansion of basis functions $V_k(z)$ as defined in Proposition 4.6.

Corollary 4.10: Let $G \in \mathcal{G}_1$, and let $\Psi(G)$ be as defined in Theorem 4.7. Then

a) For every signal $x \in \ell_2^q[0,\infty)$ there exists a unique transform $X = \{X_k\}_{k=0,1,\cdots} \in \ell_2^{q\times n}[0,\infty)$ such that

$$x(t) = \sum_{k=0}^{\infty} X_k \phi_k(t). \tag{20}$$

We refer to $X_k \in \boldsymbol{C}^{q\times n}$ as the orthogonal expansion coefficients of x.

b) For every proper stable transfer function $H(z) \in \mathcal{H}_2^{p\times m}$ there exist unique $D_s \in \mathbf{R}^{p\times m}$, and $L = \{L_k\}_{k=0,1,\cdots} \in \ell_2^{p\times n}[0,\infty)$, such that

$$H(z) = D_s + z^{-1}\sum_{k=0}^{\infty} L_k V_k(z). \tag{21}$$

We refer to D_s, L_k as the orthogonal expansion coefficients of $H(z)$. □

We will refer to the sequence $\{V_k(z)\}_{k=0,1,\cdots}$, as defined in Proposition 4.6, as the sequence of generating transfer functions for the orthonormal basis $\Psi(G)$.

The series expansion as reflected in (21) is schematically depicted in the diagram in Fig. 1, where q reflects the time shift, $qu(t) = u(t+1)$.

To find appropriate ways to calculate the orthogonal functions, as well as to determine the transformations as meant in the corollary, we will now first analyze the results presented so far in terms of state-space realizations.

V. Balanced State-Space Representations

To represent the orthogonal controllability matrix in a state-space form, we will use a balanced state-space realization of G. We first present the following, rather straightforward, lemma.

Lemma 5.1: Let G be a square inner transfer function with minimal realization (A,B,C,D). Then the realization is (internally) balanced if and only if it is both input balanced and output balanced. □

Next we examine how the property that a transfer function is inner, is reflected in a state-space realization of the function.

Proposition 5.2: Let G be a transfer function with realization (A,B,C,D), such that (A,B) is a controllable pair, and the realization is output balanced, i.e., $A^*A + C^*C = I$. Then $G^T(z^{-1})G(z) = I$ if and only if

i) $D^*C + B^*A = 0$, and

ii) $D^*D + B^*B = I$. □

Note that for this proposition there also exists a dual form, concerning the transfer function G^T with realization (A^*, C^*, B^*, D^*), that can be applied if G is square inner.

The characterization of the inner property in the above proposition is made for output balanced realizations. Since, according to Lemma 5.1 output balancedness is implied by balancedness, it also refers to balanced realizations.

The class of functions $\mathcal{G}_1$ can simply be characterized in terms of a balanced realization.

Proposition 5.3: Let G be an $m\times m$ inner function with minimal balanced realization (A,B,C,D). Then $G \in \mathcal{G}_1$ if and only if rank $B = m$, or equivalently rank $C = m$. □

The following proposition shows that we can use a balanced realization of G to construct a balanced realization for any power of G.

Proposition 5.4: Let G be a square inner transfer function with minimal balanced realization (A,B,C,D) having state dimension $n > 0$. Then for any $k > 1$ the realization (A_k, B_k, C_k, D_k) with

$$A_k = \begin{bmatrix} A & 0 & \cdots & \cdot & 0 \\ BC & A & 0 & \cdot & 0 \\ BDC & BC & \cdot & \cdot & 0 \\ \vdots & \vdots & \cdot & \ddots & 0 \\ BD^{k-2}C & BD^{k-1}C & \cdots & BC & A \end{bmatrix} \tag{22}$$

$$B_k = \begin{bmatrix} B \\ BD \\ BD^2 \\ \vdots \\ BD^{k-1} \end{bmatrix} \tag{23}$$

$$C_k = [D^{k-1}C \quad D^{k-2}C \quad \cdots \quad DC \quad C] \tag{24}$$

$$D_k = D^k \tag{25}$$

is a minimal balanced realization of G^k with state dimension $n \cdot k$. □

Examining the realization in the above proposition, reveals a similar structure of observability and controllability matrices, as has been discussed in the previous section; e.g., taking the situation $k = 2$, it shows that the controllability matrix of (A_2, B_2) contains the controllability matrix of (A, B) as its first block row.

Proposition 5.5: Let $G(z)$ be an $m \times m$ inner transfer function with McMillan degree $n > 0$, whose Hankel matrix has an svd $\mathcal{H}(G) = U_0 V_0^*$; let (A, B, C, D) be a minimal balanced realization of G such that $V_0^* = [B\ AB\ A^2B\ \cdots]$. Then the unique sequence of orthogonal matrices $\{\Gamma_k^c\}_{k=1,2,\cdots}$ as considered in Theorem 4.5 is determined by

$$\Gamma_k^c = [B_k\ A_kB_k\ A_k^2B_k\ \cdots] \tag{26}$$

with A_k, B_k as defined in (22), (23). □

The above result shows how a minimal balanced realization of G actually generates the sequence of orthogonal matrices Γ_k^c, the rows of which are the basis functions in our orthonormal basis of ℓ_2^n.

We will show that there exist recursive formulae for constructing the orthogonal functions.

Proposition 5.6: Let G be an inner function, $G \in \mathcal{G}_1$, and consider the assumptions and notation as in Theorem 4.5 and Proposition 5.5. Denote

$$X = BC \quad \text{and} \tag{27}$$

P any matrix satisfying

$$PB = BD. \tag{28}$$

Then the elements of Γ_k^c are determined by the following recursive equations

$$M_0(0) = B \tag{29}$$

$$M_k(i+1) = AM_k(i) + \sum_{j=1}^{k} P^{j-1} X M_{k-j}(i),\ i \geq 0; \tag{30}$$

$$M_k(0) = PM_{k-1}(0) \tag{31}$$

with Γ_k^c as in (15) with

$$V_k^* = [M_k(0)\ M_k(1)\ M_k(2)\ \cdots]$$

as in (19). □

The recursive equations show how we can simply construct the set of orthogonal functions. Note that the matrix P in (28) is nonunique. The result (29)-(31) however is unique. A straightforward choice for P satisfying (28) is

$$P = BD(B^*B)^{-1}B^*. \tag{32}$$

Note that, as a result of Proposition 5.3, the matrix B^*B is invertible whenever $G \in \mathcal{G}_1$.

The orthogonal functions $\Psi(G)$ generated by an inner function G can be represented in terms of their generating functions $V_k(z)$, as defined in Proposition 4.6. These generating transfer functions can also be realized in terms of a minimal balanced realization of G. This is reflected in the following theorem.

Theorem 5.7: Let G be an inner function, $G \in \mathcal{G}_1$, with a minimal balanced realization (A, B, C, D). Let this inner function generate an orthonormal basis with corresponding generating functions $V_k(z)$, as defined in Proposition 4.6.

1) Let F be a matrix determined by

$$F = X - PA \tag{33}$$

with X defined in (27) and P any matrix satisfying (28). Then, for $k \in \boldsymbol{Z}_+$,

a) $V_k(z) = \left[(zI - A)^{-1}F(I - zA^*)\right]^k z(zI - A)^{-1}B$;
b) $V_k(z)$ is unique, i.e., it is not dependent on the specific choice of P in (28).

2) If there exists a matrix R such that $B = RC^*$, then $F = R$ satisfies the conditions of Part 1 of this theorem. □

Now we come to the construction of a series expansion of any stable proper rational transfer function, in terms of the new orthonormal basis.

Theorem 5.8: Let G be an inner function, $G \in \mathcal{G}_1$, with a minimal balanced realization (A, B, C, D). Let this inner function generate an orthonormal basis with corresponding generating functions $V_k(z)$, as defined in Proposition 4.6. Let $H \in \mathcal{H}_2^{p \times m}$ be any proper and stable transfer function with a minimal realization (A_s, B_s, C_s, D_s). Then

$$H(z) = D_s + z^{-1} \sum_{k=0}^{\infty} L_k V_k(z) \tag{34}$$

with $L_k \in \boldsymbol{C}^{p \times n}$ determined by

$$L_k = C_s Q_k \tag{35}$$

$$Q_0 = A_s Q_0 A^* + B_s B^* \tag{36}$$

$$Q_{i+1} = A_s Q_{i+1} A^* + A_s Q_i F^* - Q_i A F^* \tag{37}$$

with F as defined in (33). □

In Section VII we will show that specific choices of $G(z)$ in relation with $H(z)$, i.e., specific relations between the inner function G producing the orthonormal basis and a transfer function H that should be described in this basis, will lead to very simple representations.

VI. A Generalization of Classical Basis Functions

In this section we show three examples of well-known sets of orthogonal functions that are frequently used in the description of linear time-invariant dynamical systems and that occur as special cases in the framework that is discussed in this paper.

Pulse Functions

Consider the inner function $G(z) = z^{-1}$, $G \in \mathcal{G}_1$. The Hankel matrix of G satisfies

$$\mathcal{H}(G) = \begin{bmatrix} 1 & 0 & \cdots & \cdot & 0 \\ 0 & 0 & 0 & \cdot & 0 \\ 0 & 0 & 0 & \cdot & 0 \\ \vdots & \vdots & \cdot & \ddots & \cdot \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix} [1\quad 0\quad 0\quad \cdots] = U_0 V_0^*. \tag{38}$$

As a result $V_0(z) = 1$, and with Proposition 4.6 the generating transfer functions $V_k(z)$ satisfy $V_k(z) = G^k(z) = z^{-k}$, $k = 0, 1, \cdots$. The corresponding set of basis functions $\Psi(G)$

is determined by $\phi_k(t) = \delta(t-k)$ with $\delta(\tau)$ the Kronecker delta function.

The inner function G can be realized by the minimal balanced realization $(A, B, C, D) = (0, 1, 1, 0)$. The equation $PB = BD$ is satisfied by $P = 0$, and the corresponding result for F is $F = BC = 1$. Applying Theorem 5.8 shows the classical result that $L_k = C_s A_s^k B_s$.

Laguerre Functions

Consider the inner function $G(z) = \dfrac{1-az}{z-a}$, with some real-valued a, $|a| < 1$, and denote $\eta = 1 - a^2$. A minimal balanced realization of G is given by $(A, B, C, D) = (a, \sqrt{\eta}, \sqrt{\eta}, -a)$. Equation $PB = BD$ is satisfied by $P = -a$, leading to $F = BC - PA = \eta + a^2 = 1$. Taking account of the fact that for one-dimensional scalar G, $M_k(i) = \phi_k(i)$, it follows from Proposition 5.6 that

$$\phi_0(0) = \sqrt{\eta} \tag{39}$$

$$\phi_k(i+1) = a\phi_k(i) + \eta \sum_{j=1}^{k} (-a)^{j-1} \phi_{k-j}(i) \tag{40}$$

$$\phi_k(0) = -a\phi_{k-1}(0). \tag{41}$$

These equations exactly match the equations that generate the normalized discrete-time Laguerre polynomials with discount factor a, [14], [32].

The corresponding generating transfer functions $V_k(z)$ can be analyzed with the result of either Proposition 4.6 or Theorem 5.7

$$V_k(z) = \sqrt{\eta} z \frac{(1-az)^k}{(z-a)^{k+1}}. \tag{42}$$

This exactly fits with the formulation of the generating transfer functions of discrete-time Laguerre polynomials in, e.g., [23].

Kautz Functions

Consider the inner function $G(z) = \dfrac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c}$ with some real-valued b, c satisfying $|c|, |b| < 1$.

A balanced realization of $G(z)$ can be found to be equal to

$$A = \begin{bmatrix} b & \sqrt{(1-b^2)} \\ c\sqrt{(1-b^2)} & -bc \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ \sqrt{(1-c^2)} \end{bmatrix}$$
$$C = [\gamma_2 \quad \gamma_1] \quad D = -c$$

with $\gamma_1 = -b\sqrt{(1-c^2)}$ and $\gamma_2 = \sqrt{(1-c^2)(1-b^2)}$.

With the expression for $V_0(z)$ from Theorem 5.7-a) it follows that

$$z^{-1} V_0(z) = \frac{\sqrt{(1-c^2)}}{z^2 + b(c-1)z - c} \begin{bmatrix} \sqrt{(1-b^2)} \\ z - b \end{bmatrix} \tag{43}$$

which exactly equals $\begin{bmatrix} \psi_2(z, b, c) \\ \psi_1(z, b, c) \end{bmatrix}$ representing the orthogonal Kautz functions, as represented in [41], [44]. Postmultiplication with $G^k(z)$ is equivalent to the situation in the case of Kautz functions.

VII. Orthonormal Functions Originating from General Dynamical Systems

We have shown that any square inner transfer function $G \in \mathcal{G}_1$ generates an orthonormal basis for the signal space ℓ_2^n. One of the reasons for developing this generalized bases was to find out whether we can yield a more suitable representation of a general dynamical system, when the basis within which we describe the system is more or less adapted to the system dynamics. In view of the results presented so far, this aspect relates to the question whether we can construct an inner transfer function generating a basis that incorporates dynamics of a general system to be represented within this basis.

There are several ways of connecting general transfer functions to inner functions, as e.g., inner/outer factorization [10], [5], normalized coprime factorization [8], [40], [33], [4], or inner-unstable factorization [2]. Even if the corresponding inner functions are not square, they can always be embedded in a square inner function [11]. In this paper, however, we will explore a different connection, where a general stable dynamical system with input balanced realization $(A, B, \tilde{C}, \tilde{D})$ will induce a square inner function through retaining the matrices (A, B) and constructing (C, D) such that (A, B, C, D) is inner. This implies that the poles of the stable dynamical system are retained in the corresponding inner function. The following result shows the existence and construction of such an inner function.

Proposition 7.1: Let (A, B) be the system matrix and input matrix of an input balanced realization of a transfer function $H \in \mathcal{RH}_2^{p \times m}$ with McMillan degree $n > 0$, and with rank $B = m$. Then

a) There exist matrices C, D such that (A, B, C, D) is a minimal balanced realization of a square inner function $G \in \mathcal{G}_1$.
b) A realization (A, B, C, D) has the property mentioned in a) if and only if

$$C = UB^*(I_n + A^*)^{-1}(I_n + A) \tag{44}$$

$$D = U[B^*(I_n + A^*)^{-1}B - I_m] \tag{45}$$

with $U \in \mathbb{R}^{m \times m}$ any unitary matrix.

c) For a realization satisfying (44), (45) a valid choice of matrix F satisfying (33) is given by

$$F = [I_n + B(U - I_m)(B^*B)^{-1}B^*](I_n + A)(I_n + A^*)^{-1}. \tag{46}$$

□

In the proposition all inner functions are characterized that can be constructed in the way as described above, by retaining the matrices (A, B) of any given stable system. Note that the extension C, D is not unique. The nonuniqueness is reflected by a possible unitary premultiplication of the inner function. Note also that when choosing $U = I_m$, expression (46) reduces to $F = (I_n + A)(I_n + A^*)^{-1}$.

We will now present a result that is very appealing. It shows that when we want to describe the dynamical system H in terms of the basis that it has generated, as presented in Proposition 7.1, then the series expansion in the new orthogonal basis becomes extremely simple.

Theorem 7.2: Let $H \in \mathcal{RH}_2^{p\times m}$ have an input balanced realization (A_s, B_s, C_s, D_s), having all controllability indexes > 0. Let $G \in \mathcal{G}_1$ be a square inner function with minimal balanced realization (A, B, C, D) such that $A = A_s$ and $B = B_s$, generating an orthonormal basis with generating transfer functions $V_k(z)$. Then

$$H(z) = D_s + z^{-1} \sum_{k=0}^{\infty} L_k V_k(z) \tag{47}$$

with

$$L_0 = C_s \quad \text{and} \tag{48}$$
$$L_k = 0 \quad \text{for } k > 0. \tag{49}$$

Proof: The proof follows by applying Theorem 5.8. With $A = A_s$, $B = B_s$ (36) becomes $Q_0 = AQ_0A^* + BB^*$. Since (A, B) is input balanced, the solution to this equation is $Q_0 = I$, leading to $L_0 = C_s$. Substituting $Q_0 = I$ in (37) and using the stability of A shows that $Q_i = 0$ for $i > 0$. □

The theorem shows that when we use a general stable and proper dynamical system to generate an orthonormal basis as described above, then the system itself has a very simple representation in terms of this basis. It is represented in a series expansion with only two nonzero expansion coefficients, being equal to the system matrices C_s and D_s.

In the next section we will discuss the results of this paper regarding their relevance to problems of system identification and system approximation.

It has to be stressed that, so far, we have only used the generalized orthonormal basis to study the series expansion of a given stable transfer function. Similar to the case of the pulse functions and Laguerre functions, the presented generalized functions induce a transformation of ℓ_2-signals to a transform domain, compare e.g., with the z-domain when pulse functions are used. In this transform domain dynamical system equations can be derived, leading to transform pairs of time-domain and orthogonal-domain system representations. In the case of a Laguerre basis, these kinds of transformations actually have been used frequently also in an identification context, by first transforming the measured input/output signals to the Laguerre domain, and consecutively identifying a system in this domain; see e.g., [22], [23], [32], [31].

For the generalized basis, results along these lines have been presented in [18], [19]. An analysis of the system transformations between time domain and generalized transform domain is treated in [19] and [39].

VIII. System Approximation and Identification

We will now discuss the way in which the introduced orthogonal basis functions provide a solution to problem b) as mentioned in the introduction, i.e., the quantification of an error bound for finite length expansion approximants.

We will present results showing that the speed of convergence in an orthogonal series expansion can be quantified and that an increase of speed is obtained as the dynamics of system and basis approach each other. To formulate these results we need an alternative formulation of Theorem 5.8 in terms of Kronecker products.

Proposition 8.1: Let $H \in \mathcal{RH}_2^{p\times m}$ be a transfer function with an input balanced realization (A_s, B_s, C_s, D_s), and let (A, B) be an input balanced pair that generates an $m \times m$ inner transfer function $G \in \mathcal{G}_1$, leading to an orthonormal basis $\Psi(G)$.

Then the orthogonal expansion coefficients L_k satisfying $H(z) = D_s + z^{-1}\sum_{k=0}^{\infty} L_k V_k(z)$ are determined by

$$Vec(L_k) = ZX^kY \tag{50}$$

with

$$Z = (I \otimes C_s)M^{-1} \tag{51}$$
$$Y = Vec(B_sB^*) \tag{52}$$
$$X = NM^{-1} \tag{53}$$
$$M = I \otimes I - A \otimes A_s \tag{54}$$
$$N = F \otimes A_s - FA^* \otimes I. \tag{55}$$

Note that due to (50) we can consider $Vec(L_k)$ as a sequence of Markov parameters of a dynamical system with a state-space realization given by $(X, Y, Z, 0)$. By examining the eigenvalues of this realization, we create the possibility of drawing some conclusions on the speed of convergence of the series expansion. The following result is taken from [19].

Proposition 8.2: Consider the situation of Proposition 8.1 with $H(z)$ and $G(z)$ having McMillan degree n_s, n, respectively, and $m = 1$. Let μ_i, $i = 1, \cdots, n_s$ denote the eigenvalues of A_s, and ρ_j, $j = 1, \cdots, n$ denote the eigenvalues of A. The dynamical system $Z(zI - X)^{-1}Y$ has a realization $(X_o, Y_o, Z_o, 0)$ that satisfies

a) X_o has dimension n_s;

b) X_o has eigenvalues λ_i, $i = 1, \cdots, n_s$ that satisfy

$$|\lambda_i| = \prod_{j=1}^{n} \left| \frac{\mu_i - \rho_j}{1 - \mu_i \rho_j} \right|. \tag{56}$$

Since the proof of this proposition is somewhat outside the scope of this paper, the reader is referred to [19].

The above proposition shows that we can draw conclusions on the convergence rate of the sequence of expansion coefficients $\{L_k\}_{k=0,\cdots}$, when given the eigenvalues of the original system $H(z)$ and the eigenvalues of the inner function $G(z)$ that generates the basis. Note that when the sets of eigenvalues $\{\mu_i\}$, $\{\rho_j\}$ coincide, then $\lambda_i = 0$, for all i, and consequently the sequence $\{L_k\}$ will have a finite number of elements unequal to zero. The above result also enables the determination of an upper bound on the error that is made, when we approximate a given system $H(z)$ through a finite number of its expansion coefficients.

Theorem 8.3: Consider the situation of Proposition 8.2, and denote

$$\hat{H}^N(z) = D_s + z^{-1} \sum_{k=0}^{N-1} L_k V_k(z)$$

and $\lambda := \max_i |\lambda_i|$. Then there exists a finite $c \in \mathbf{R}$ such that for any $\eta \in \mathbf{R}$, $\eta > \lambda$

$$\|H(z) - \hat{H}^N(z)\|_\infty \le c\frac{\eta^{N+1}}{1-\eta}. \tag{57}$$

□

Since λ is a measure for the "closeness" of system dynamics and basis dynamics, the above theorem shows that the error that is made when neglecting the tail of a series expansion, becomes smaller as λ becomes smaller. As a result, when restricting to a fixed number of expansion coefficients, the approximation error gets smaller the more accurate the basis dynamics is "adapted" to the system.

In the final part of this paper we will briefly comment on how these results could be employed in an approximate identification framework. As mentioned in the introduction, identification of a finite impulse model (FIR) (4), has some important advantages; however, it fails to be successful when the number of coefficients to be estimated becomes large. This may happen in situations of high sampling rates, moderately damped systems, as well as systems that have dominant dynamics in both the high-frequent and low-frequent region (e.g., multitime-scale systems). An alternative way to attain the advantages of this identification method, is to exploit the model structure

$$y(t) = D(\theta) + \sum_{k=0}^{N-1} L_k(\theta)V_k(q)u(t) + \varepsilon(t) \tag{58}$$

where $\varepsilon(t)$ is the one-step-ahead prediction error, $D(\theta), L_k(\theta)$ the parameterized expansion coefficients, and with $V_k(z)$ representing an appropriately chosen basis.

Note that this model structure can simply be written as

$$y(t) = D(\theta) + \sum_{k=0}^{N-1} L_k(\theta)\tilde{u}_k(t) + \varepsilon(t) \tag{59}$$

where $\tilde{u}_k(t)$ can simply be calculated by applying $u(t)$ to the -known- filters $V_k(q)$, compare Fig. 1.

Identifying θ through least squares optimization of $\varepsilon(t)$over the time interval, is a similar problem as in the case of a FIR-model. With appropriately chosen basis functions, however, the convergence rate of the series expansion can become extremely fast; with only a few coefficients to be estimated a very accurate approximate model can be obtained. This is of course interesting and appealing from both aspects of bias (accurate approximation is possible) and variance (few parameters to be estimated from data). An analysis of bias and variance errors in these identification schemes is presented in [39].

Additionally, when comparing these "orthogonal FIR" model structures with nonlinearly parameterized model structure as e.g., a Box Jenkins or ARMAX model ([26]), we avoid problems of possible occurrence of local (nonglobal) minima in the quadratic identification criterion. Moreover the freedom in the choice of basis functions allows the fruitful use of "a priori information" concerning the system dynamics.

Very often an identification experimenter has a -rough-knowledge about the dynamics of the system under consideration, e.g., from previous experiments or from physical insight into the process dynamics. It would be favorable to exploit this knowledge in an identification procedure. The method suggested above, shows that this a priori knowledge can be exploited in terms of the basis functions that are chosen. When we have -rough- knowledge about the poles of the system, we can construct basis functions that are based on this set of poles. The more accurate the poles are, i.e., the more accurate our a priori information is, the better we can adapt the basis functions to the system dynamics. As a result, see Theorem 8.3, the estimated model can become more accurate when restricting to a prespecified number of coefficients to be estimated.

Effectively the identification problem now reflects the identification of the mismatch between the system under consideration and the knowledge that already was available, represented in the basis functions. This actually is very appealing, as the priori information simplifies the identification procedure. Note that in the way described above, the *a priori* information does not have to be exact, i.e., it is not of the type of fixing *a priori* a constraint on the model parameters, as e.g., the steady-state gain. The information can be uncertain. The only result is that the more accurate it is, the more simple the system representation will be.

This discussion also motivates the use of an iterative scheme, where the identification of parameters θ is performed iteratively, using the model that is estimated in step $i-1$ for constructing the basis functions for step i. An example of such an iterative scheme has been shown in [19].

One remark that has to be made in this respect, is a remark on the model order of a system represented by a finite number of expansion coefficients. The McMillan degree of this system, as in the case of an FIR-representation, will generally be large. This results from the following observation.

Proposition 8.4: Consider the transfer function

$$\hat{H}^N(z) = \hat{D} + \sum_{k=0}^{N-1} \hat{L}_k V_k(z)$$

with $V_k(z)$ the generating transfer functions of an orthonormal basis $\Psi(G)$, where the inner function $G \in \mathcal{G}_1$ has a minimal balanced realization (A, B, C, D) with dimension n. Then $\hat{H}^N(z)$ has a state-space realization $(A_{N-1}, B_{N-1}, K, \hat{D})$, with A_{N-1},B_{N-1} defined in (22), (23), and $K = [\, \hat{L}_0 \; \hat{L}_1 \; \hat{L}_2 \; \cdots \; \hat{L}_{N-1} \,]$. □

The proof of this proposition follows by inspection.

With $\hat{L}_i$ being the result of an unconstrained optimization in an identification procedure, the state-space dimension of the model will generically be equal to Nn. Consequently, if one wants to represent the model again in a traditional state-space form of low dimension, a model reduction procedure will have to be used to arrive at a reduced dimension. This also motivates a further analysis of the realization problem in terms of orthogonal expansion coefficients $\{L_k\}$.

IX. Conclusions

We have developed a theory on orthogonal functions as basis functions for general linear time-invariant stable systems. The basic ingredient is that every square inner transfer function in a very natural way induces two sets of orthogonal functions that form a basis of the signal space ℓ_2. The ordinary pulse functions and the classical Laguerre and Kautz polynomials are special cases in this theory of inner functions.

With this concept we have explored the connection between a general dynamical system and an inner function, by letting the inner function be determined through a specified set of poles. An important property of the resulting orthonormal functions is that they—to some extent—incorporate the dynamic behavior of the underlying system. We have developed a theory on these system based orthogonal functions, both on an input–output level and in terms of balanced state-space realizations. Furthermore we have shown how the alternative basis can be fruitfully used in problems of system approximation and identification, leading to simplified identification schemes, in which *a priori* knowledge about the process dynamics can be utilized by incorporating the information into the basis.

APPENDIX

Lemma A1: Let $G(z)$, $F(z)$, and $R(z)$ be stable transfer functions wit Laurent expansions $G(z) = \sum_{k=0}^{\infty} G_k z^{-k}$, $F(z) = \sum_{k=0}^{\infty} F_k z^{-k}$, and $R(z) = \sum_{k=0}^{\infty} R_k z^{-k}$. Then $R(z) = F(z)G(z)$ if and only if

$$[R_0\ R_1\ R_2\ \cdots] = [F_0\ F_1\ F_2\ \cdots]T_v \tag{A.1}$$

with T_v as defined in (12). □

Proof: The equality $R(z) = (\sum_{k=0}^{\infty} F_k z^{-k})\ (\sum_{k=0}^{\infty} G_k z^{-k})$ is equivalent to $\sum_{k=0}^{\infty} R_k z^{-k} = \sum_{k=0}^{\infty} [\sum_{i=0}^{k} F_i G_{k-i}]z^{-k}$, and to $R_k = \sum_{i=0}^{k} F_i G_{k-i}$, which exactly matches (A.1). □

Lemma A2 [11]: Given matrices $X \in \boldsymbol{C}^{n\times m}$, $Y \in \boldsymbol{C}^{n\times r}$, $r \geq m$, with $XX^* = YY^*$; then there exists a $W \in \boldsymbol{C}^{m\times r}$ such that $Y = XW$ and $WW^* = I$. □

Lemma A3: Let $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Then $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ and $\sum_{i=0}^{t}\binom{k+i}{k} = \binom{k+t+1}{k+1}$. □

Proof: By simple calculation. □

Lemma A4: Let $G(z)$ be a square inner function such that $\|G_0\| < 1$, with $\|\cdot\|$ any induced matrix norm. Let G^k have a Laurent expansion $G^k(z) = \sum_{i=0}^{\infty} G_i^{(k)} z^{-i}$. Then $\|G_i^{(k)}\| \leq \binom{k+i-1}{k-1}\|G_0\|^{k-i}$ for $0 \leq i \leq k-1$. □

Proof: The proof will be given by induction. For $k = 1$, validity is trivial. Suppose that the statement holds true for $k \leq n$. Now we consider two cases.

i) Consider $G_i^{(n+1)}$, where $i < n$

$$G_i^{(n+1)} = G_0 G_i^{(n)} + G_1 G_{i-1}^{(n)} + \cdots + G_i G_0^{(n)};$$

$$\begin{aligned}
\|G_i^{(n+1)}\| &\leq \\
&\leq \|G_0\|\|G_i^{(n)}\| + \|G_{i-1}^{(n)}\| + \cdots + \|G_1^{(n)}\| + \|G_0^{(n)}\| \\
&\leq \|G_0\|\binom{n+i-1}{n-1}\|G_0\|^{n-i} + \\
&\quad + \binom{n+i-2}{n-1}\|G_0\|^{n-i+1} + \\
&\quad + \cdots + \binom{n}{n-1}\|G_0\|^{n-1} + \|G_0\|^{n}
\end{aligned} \tag{A.2}$$

$$\begin{aligned}
&\leq \|G_0\|^{n+1-i}\sum_{j=0}^{i}\binom{n-1+j}{n-1} \\
&= \|G_0\|^{n+1-i}\binom{n+i}{n} \qquad \text{by Lemma A3.}
\end{aligned} \tag{A.3}$$

ii) Consider the case $i = n$

$$G_n^{(n+1)} = G_0 G_n^{(n)} + G_1 G_{n-1}^{(n)} + \cdots + G_n G_0^{(n)}; \tag{A.4}$$

$$\begin{aligned}
\|G_n^{(n+1)}\| &\leq \\
&\leq \|G_0\| + \|G_{n-1}^{(n)}\| + \cdots + \|G_1^{(n)}\| + \|G_0^{(n)}\| \\
&\leq \|G_0\|\left[1 + \binom{2n-1}{n-1} + \binom{2n-3}{n-1} + \cdots + \right. \\
&\qquad \left. + \binom{n}{n-1} + 1\right] \\
&= \|G_0\|\left[1 + \binom{2n-1}{n}\right] \\
&\leq \|G_0\|\left[\binom{2n-1}{n-1} + \binom{2n-1}{n}\right] = \\
&= \|G_0\|\binom{2n}{n}.
\end{aligned}$$

We have shown that $\|G_i^{(n+1)}\| \leq \|G_0\|^{n+1-i}\binom{n+i}{n}$ for $i \leq n$, which proves the result. □

Lemma A5: Let $G(z)$ be an $m \times m$ inner function such that $\|G_0\| < 1$, with $\|\cdot\|$ any induced matrix norm. Let G^k have a Laurent expansion $G^k(z) = \sum_{i=0}^{\infty} G_i^{(k)} z^{-i}$, and Hankel matrix $\Pi_k := \mathcal{H}(G^k)$. Then for all i

$$\lim_{k\to\infty} \max_j \|(\Pi_k^*\Pi_k)_{ij} - \delta_{ij}\| = 0.$$

Proof: Consider $R_k(i) = \sum_{t=0}^{i}\|G_t^{(k)}\|$. With Lemma A4 it follows that $R_k(i) \leq$

$$\begin{aligned}
\sum_{t=0}^{i}\binom{k+t-1}{k-1}\|G_0\|^{k-t} &\leq \left[\sum_{t=0}^{i}\binom{k+t-1}{k-1}\right]\|G_0\|^{k-i} \\
&= \binom{k+i}{k}\|G_0\|^{k-i} \leq (k+1)^i\|G_0\|^{k-i}.
\end{aligned}$$

Since $\|G_0\| < 1$, this implies that $R_k(i) \to 0$ for $k \to \infty$.

Now consider the (i,j)-block element of $\Pi_k^*\Pi_k$ with $j \geq i$. $(\Pi_k^*\Pi_k)_{ij} = \sum_{s=0}^{\infty}(G_{i+s}^{(k)})^*\ G_{j+s}^{(k)} = \delta_{ij}I_m - \sum_{s=0}^{i-1}(G_s^{(k)})^*\ G_{s+j-i}^{(k)}$. Consequently $\|(\Pi_k^*\ \Pi_k)_{ij} - \delta_{ij}\ I_m\| = \|\sum_{s=0}^{i-1}(G_s^{(k)})^*\ G_{s+j-i}^{(k)}\| \leq \sum_{s=0}^{i-1}\|G_s^{(k)}\| = R_k\ (i-1) \to 0$ for $k \to \infty$.

For $j < i$ it holds that $R_k(j-1) < R_k(i-1)$, which implies that for all j, $\|(\Pi_k^*\Pi_k)_{ij} - \delta_{ij}\| \leq R_k(i-1)$. □

Proof of Proposition 4.1: Denote $L_i = \sum_{k=0}^{\infty} G_{k+i}^T G_k$, for $i \in \boldsymbol{Z}$, with $G_j := 0, j < 0$. Then $G^T(z^{-1})\,G(z) = \sum_{\ell=0}^{\infty} G_\ell^T\ z^\ell \sum_{k=0}^{\infty} G_k z^{-k}$. This expression equals $\sum_{j=-\infty}^{\infty}(\sum_{k=0}^{\infty} G_{k+j}^T G_k)\ z^j = \sum_{j=-\infty}^{\infty} L_j z^j$. Since G is inner, $G^T\ (z^{-1})\ G(z) = I$, and evaluation of the former expression for $j \geq 0$ proves the result. □

Proof of Proposition 4.3: Part i) follows directly from Proposition 4.1. For Part ii), consider $T_v(\mathcal{H}(G))^*$. Applying Proposition 4.1, shows that $T_v(\mathcal{H}(G))^* = 0$, which implies that $T_v V_0 U_0^* = 0$ and $T_v V_0 U_0^* U_0 = 0$, leading to $T_v V_0 = 0$. The proof for T_u follows analogously, employing the fact that $G^T(z)$ is inner too. □

Proof of Lemma 4.4: If G is inner, then for any $k > 1$, $(G^k)^T(z^{-1})\, G^k(z) = (G^{k-1})^T(z^{-1})\, G^T(z^{-1})\, G(z)\, G^{k-1}(z) = (G^{k-1})^T(z^{-1})\, G^{k-1}(z)$, and by induction it follows that G^k is inner. A proof for the McMillan degree of G^k is contained in the proof of Proposition 5.4. □

Proof of Theorem 4.5:

Part A: A constructive proof will be given in three steps.

i) The choices for U_j and V_j^* as in (17), (18) lead to matrices Γ_k^o and Γ_k^c in (14), (15), that are unitary;

ii) The constructed matrix $\Gamma_k^o \Gamma_k^c$ has a block Hankel structure;

iii) $\Gamma_k^o \Gamma_k^c = \mathcal{H}(G^k)$.

Proofs:

i) Note that the (i,j)-block-element of $\Gamma_k^c(\Gamma_k^c)^*$ equals $V_0^* T_v^{i-1}(T_v^*)^{j-1} V_0$. With Proposition 4.3 it follows that this equals I for $i = j$, and zero elsewhere.

ii) This proof will be given by complete induction. For $k = 1$ the statement is true by definition. Assume that it holds for $k-1$, i.e., $\Pi_{k-1} := \Gamma_{k-1}^o \Gamma_{k-1}^c$ is a Hankel matrix. We have to show that Π_k is a Hankel matrix too, with $\Pi_k = [U_{k-1} \;\; \Gamma_{k-1}^o] \begin{bmatrix} \Gamma_{k-1}^c \\ V_{k-1}^* \end{bmatrix}$.

The Markov parameters of the system $G^{k-1}(z)$ will be denoted by $H_0, H_1, H_2, \cdots$.

With U_j and V_j^* chosen as in (17), (18), it follows that

$$\Pi_k = [T_u \Gamma_{k-1}^o \;\; U_0] \begin{bmatrix} \Gamma_{k-1}^c \\ V_{k-1}^* \end{bmatrix}$$

showing that $\Pi_k = T_u \Pi_{k-1} + U_0 V_0^* T_v^{k-1}$.

The matrix Π_k has a block Hankel structure if and only if $S\Pi_k = \Pi_k S^*$, with

$$S = \begin{bmatrix} 0 & I & 0 & 0 & \cdots \\ 0 & 0 & I & 0 & \cdots \\ \vdots & \vdots & 0 & \ddots & \vdots \end{bmatrix}.$$

Evaluation of $S\Pi_k$ shows that $S\Pi_k = ST_u\Pi_{k-1} + SU_0V_0^*T_v^{k-1} =$

$$= \begin{bmatrix} G_1 & \\ G_2 & T_u \\ \vdots & \end{bmatrix} \Pi_{k-1} + \begin{bmatrix} G_2 & G_3 & \cdots \\ G_3 & G_4 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} T_v^{k-1} \tag{A.5}$$

$$\begin{aligned} &= \begin{bmatrix} G_1 & \\ G_2 & T_u \\ \vdots & \end{bmatrix} \begin{bmatrix} H_1 \;\; H_2 \;\; H_3 \;\; \cdots \\ \Pi_{k-1} S^* \end{bmatrix} + \\ &\quad + \begin{bmatrix} G_2 & G_3 & \cdots \\ G_3 & G_4 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} T_v^{k-1} \\ &= T_u \Pi_{k-1} S^* + \\ &\quad + \begin{bmatrix} G_1 & G_2 & \cdots \\ G_2 & G_3 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} H_1 \;\; H_2 \;\; H_3 \;\; \cdots \\ T_v^{k-1} \end{bmatrix}. \end{aligned} \tag{A.6}$$

From Lemma A1 we can deduct that the ith block row of T_v^{k-1} corresponds to the Markov parameters of the transfer function $z^{-i+1}G^{k-1}(z)$. So

$$T_v^{k-1} = \begin{bmatrix} H_0 & H_1 & H_2 & \cdots \\ 0 & H_0 & H_1 & \cdots \\ \vdots & 0 & H_0 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}.$$

As a result (A.6) can be written as

$$S\Pi_k = T_u \Pi_{k-1} S^* + U_0 V_0^* T_v^{k-1} S^* = \Pi_k S^*$$

which proves that Π_k is a block Hankel matrix.

iii) The proof follows by induction, similarly as in step ii). Consider the first block row of $\Pi_k = \Gamma_k^o \Gamma_k^c$. This equals

$$G_0[H_1 \;\; H_2 \;\; \cdots] + [G_1 \;\; G_2 \;\; \cdots] T_v^{k-1}.$$

Lemma A1 shows that this is equivalent to

$$G_0[H_1 \;\; H_2 \;\; \cdots] + [W_0 \;\; W_1 \;\; \cdots]$$

where W_i is such that

$$\begin{aligned} \sum_{k=0}^{\infty} W_k z^{-k} &= \left[\sum_{j=0}^{\infty} G_{j+1} z^{-j}\right] G^{k-1}(z) = \\ &= \left[\sum_{j=0}^{\infty} G_{j+1} z^{-j}\right] \left[\sum_{i=0}^{\infty} H_i z^{-i}\right]. \end{aligned}$$

Hence the first block row of Π_k corresponds to the Markov parameters of $\left[\sum_{j=0}^{\infty} G_j z^{-j}\right] \left[\sum_{i=0}^{\infty} H_i z^{-i}\right] = G^k(z)$.

Part B: Since $\Pi_k = [U_{k-1} \;\; \Gamma_{k-1}^o] \begin{bmatrix} \Gamma_{k-1}^c \\ V_{k-1}^* \end{bmatrix}$ is an svd of Π_k, it follows that $U_0^* \Pi_k = V_{k-1}^*$, and $\Pi_k V_0 = U_{k-1}$ which shows the uniqueness of U_{k-1} and V_{k-1} for a given Γ_{k-1}^o and Γ_{k-1}^c. Since U_0, V_0 are unique up to unitary postmultiplication, this holds for the whole sequence of matrices $\{U_i, V_i\}_{i=0,1,\cdots}$.

Part C: The proof is given by construction in part a). □

Proof of Proposition 4.6: With $V_k^* = V_{k-1}^* T_v$ the result follows immediately from Lemma A1. □

Proof of Theorem 4.7: The result of the theorem follows if the set of basis functions is complete in ℓ_2, i.e., if for any $x \in \ell_2[0, \infty)$, the following implication holds

$$(< \phi_k, x > = 0 \text{ for all } k) \Rightarrow x = 0$$

with $< \cdot, \cdot >$ the inner product in ℓ_2.

If $< \phi_k, x > = 0$ for all k, then $(\Gamma_k^c)^* \Gamma_k^c y = 0$ for all k, with $y := [x(0) \; x(1) \; \cdots]^*$. Consider the ith row of this equation: $[(\Gamma_k^c)^* \Gamma_k^c]_{i*} y = 0$ for all k, with $[(\Gamma_k^c)^* \Gamma_k^c]_{i*}$ the ith row of the corresponding matrix, then

$$\|([(\Gamma_k^c)^* \Gamma_k^c]_{i*} - e_i^*) y\| \le \|[(\Gamma_k^c)^* \Gamma_k^c]_{i*} - e_i^*\| \cdot \|y\|.$$

Since, $(\Gamma_k^c)^*\Gamma_k^c = (\mathcal{H}(G^k))^*\mathcal{H}(G^k)$, it follows from Lemma A5 that $\lim_{k\to\infty} \|[(\Gamma_k^c)^*\Gamma_k^c]_{i*} - e_i^*\| = 0$, which implies that $y_i = 0$. □.

Proof of Corollary 4.10: Part a) follows directly from the completeness of the basis. For part b), consider the ith row of $H(z) - D$, with $D = \lim_{z\to\infty} H(z)$, and $H(z) - D$ written as $\sum_{k=1}^{\infty} h^T(k)z^{-k}$, with $h(k) \in \mathbb{R}^m$.

Consider the scalar time series $\{w(t)\}_{t=0,1,\cdots}$ defined by

$$[w(0)\ w(1)\ w(2)\ ] = [h^T(1)\ h^T(2)\ \cdots].$$

Applying part a) delivers $w(t) = \sum_{k=0}^{\infty} W_k^T\phi_k(t)$, with $W_k \in \mathbb{R}^n$, $\phi_k(t) \in \mathbb{R}^n$. As a result $h^T(j+1) = \sum_{k=0}^{\infty} W_k^T[\phi_k(mj+1)\ \cdots\ \phi_k(m(j+1))\ ]$.

In the notation of Proposition 4.6 this leads to $h^T(j+1) = \sum_{k=0}^{\infty} W_k^T M_k(j)$. Consequently $\sum_{i=1}^{\infty} h^T(i)z^{-i} = \sum_{i=1}^{\infty}\sum_{k=0}^{\infty} W_k^T M_k(i-1)z^{-i} = z^{-1}\sum_{k=0}^{\infty} W_k^T V_k(z)$. Since this applies to each row of $H(z) - D$, this proves the result.

Proof of Lemma 5.1: From Proposition 4.2 it follows that for the realization of an inner function, the controllability and observability grammians have to satisfy $PQ = I$, while stability requires that $P, Q \geq 0$. In a balanced realization $P = Q$ and diagonal, which implies $P = Q = I$. □

Proof of Proposition 5.2:

$$\begin{aligned} G^T(z^{-1})G(z) &= [B^*(z^{-1}I - A^*)^{-1}C^* + D^*] \\ &\quad \cdot [C(zI - A)^{-1}B + D] \\ &= B^*(z^{-1}I - A^*)^{-1}C^*C(zI - A)^{-1}B + \\ &\quad + D^*C(zI - A)^{-1}B \\ &\quad + B^*(z^{-1}I - A^*)^{-1}C^*D + D^*D. \quad \text{(A.7)} \end{aligned}$$

Using $A^*A + C^*C = I$, we can rewrite the first term of the right-hand side by employing $I - A^*A = A^*(zI - A) + (z^{-1}I - A^*)A + (z^{-1}I - A^*)(zI - A)$.

Substitution of this in (A.7) shows that

$$\begin{aligned} G^T(z^{-1})G(z) &= (D^*C + B^*A)(zI - A)^{-1}B \\ &\quad + B^*(z^{-1}I - A^*)^{-1}(C^*D + A^*B) \\ &\quad + B^*B + D^*D. \end{aligned}$$

Since (A, B) is a controllable pair, it follows that $G^T(z^{-1})G(z) = I$ if and only if $B^*B + D^*D = I$ and $D^*C + B^*A = 0$. □

Proof of Proposition 5.3: Using Proposition 5.2, and its dual version, it follows that $DD^* + CC^* = D^*D + B^*B = I$. Now $\|D\|_2 < 1$ is equivalent to the smallest singular value of B^*B being greater than zero which is equivalent to rank $B = m$. The result for rank C follows analogously. □

Proof of Proposition 5.4: We use complete induction on k to prove this proposition. Note that we can write

$$A_k = \begin{bmatrix} A_{k-1} & 0 \\ BC_{k-1} & A \end{bmatrix} \quad B_k = \begin{bmatrix} B_{k-1} \\ BD^{k-1} \end{bmatrix}$$
$$C_k = [DC_{k-1}\quad C]$$

with $(A_1, B_1, C_1, D_1) = (A, B, C, D)$. Validity of the statement for (A_1, B_1, C_1, D_1) is straightforward. Assuming validity for $k - 1$, we have to show that the statement holds for k. First we show that (A_k, B_k, C_k, D_k) is indeed a realization of $G^k(z)$

$$\begin{aligned} C_k(zI - A_k)^{-1}B_k + D_k &= [DC_{k-1}\quad C] \cdot \\ &\begin{bmatrix} (zI - A_{k-1})^{-1} & 0 \\ (zI - A)^{-1}BC_{k-1}(zI - A_{k-1})^{-1} & (zI - A)^{-1} \end{bmatrix} \\ &\cdot \begin{bmatrix} B_{k-1} \\ BD^{k-1} \end{bmatrix} + D^k \\ &= [DC_{k-1} + C(zI - A)^{-1}BC_{k-1}](zI - A_{k-1})^{-1}B_{k-1} + \\ &\quad + C(zI - A)^{-1}BD^{k-1} + D^k \\ &= [D + C(zI - A)^{-1}B]C_{k-1}(zI - A_{k-1})^{-1}B_{k-1} + \\ &\quad + [C(zI - A)^{-1}B + D]D^{k-1} \\ &= [D + C(zI - A)^{-1}B][C_{k-1}(zI - A_{k-1})^{-1}B_{k-1} + D_{k-1}] \\ &= G(z)G^{k-1}(z) = G^k(z). \end{aligned}$$

Balancedness of the realization (A_k, B_k, C_k) can be shown by evaluating: $A_kA_k^* + B_kB_k^*$. For brevity of notation, we will write $(A_{k-1}, B_{k-1}, C_{k-1}, D_{k-1}) = (A, B, C, D)$

$$\begin{aligned} A_kA_k^* + B_kB_k^* &= \\ &\begin{bmatrix} AA^* + BB^* & AC^*B^* + BD^*B^* \\ BCA^* + BDB^* & BCC^*B + AA^* + BDD^*B^* \end{bmatrix} \\ &= \begin{bmatrix} AA^* + BB^* & (AC^* + BD^*)B^* \\ B(CA^* + DB^*) & B(CC^* + DD^*)B^* + AA^* \end{bmatrix}. \end{aligned}$$

Employing Proposition 5.2 together with $AA^* + BB^* = I$ shows that the above expression equals the identity matrix. In a similar way, using the dual forms, it can be shown that $A_k^*A_k + C_k^*C_k = I$, which proves that the realization is balanced and minimal. □

Proof of Proposition 5.5: Since $V_0^* = [B\ AB\ A^2B\ \cdots]$ and $\mathcal{H}(G) = U_0V_0^*$ is an svd, it follows that $U_0^* = [C^*\ A^*C^*\ \cdots]$. Similarly it holds for any k that $\mathcal{H}(G^k) = \Gamma_k^o\Gamma_k^c$ is an svd, with $(\Gamma_k^o)^* = [C_k^*\ A_k^*C_k^*\ \cdots]$. Since Γ_k^o and Γ_k^c satisfy the recursion property of Theorem 4.5–a), the given solution has to be the unique one. □

Proof of Proposition 5.6: With $X = BC$ and P any matrix satisfying $PB = BD$, the matrices A_k, B_k as in Proposition 5.4 will take the form

$$A_k = \begin{bmatrix} A & 0 & \cdots & \cdot & 0 \\ X & A & 0 & \cdot & 0 \\ PX & X & \cdot & \cdot & 0 \\ \vdots & \vdots & \cdot & \ddots & 0 \\ P^{k-2}X & P^{k-1}X & \cdots & X & A \end{bmatrix} \text{ and}$$

$$B_k = \begin{bmatrix} B \\ PB \\ P^2B \\ \vdots \\ P^{k-1}B \end{bmatrix}.$$

We can write

$$A_k^j B_k = \begin{bmatrix} M_0(j) \\ M_1(j) \\ \vdots \\ M_{k-1}(j) \end{bmatrix} = A_k(A_k^{j-1}B_k) = = A_k \begin{bmatrix} M_0(j-1) \\ M_1(j-1) \\ \vdots \\ M_{k-1}(j-1) \end{bmatrix}.$$

With the above representation of A_k this leads to the recursive relation (30). Relations (29), (31) follow directly from B_k. □

Proof of Theorem 5.7: 1–a). Using Proposition 4.6 we have to show that

$$z(zI-A)^{-1}B[D+C(zI-A)^{-1}B]^k = = [(zI-A)^{-1}F(I-zA^*)]^k z(zI-A)^{-1}B.$$

Note that it is sufficient to show that this holds for $k=1$, since successive application of the equality for $k=1$ shows the result for any k. The equality for $k=1$ is equivalent to

$$\begin{aligned}&\{B[D+C(zI-A)^{-1}B] = F(I-zA^*)(zI-A)^{-1}B\}\\ &\Leftrightarrow \{BD + BC(zI-A)^{-1}B = BC(zI-A)^{-1}B+\\ &-PA(zI-A)^{-1}B - zFA^*(zI-A)^{-1}B\},\\ &\Leftrightarrow \{BD = -PA(zI-A)^{-1}B - zFA^*(zI-A)^{-1}B\}.\end{aligned}$$

With $PB = BD$ it suffices to show that

$$P = -PA(zI-A)^{-1} - zFA^*(zI-A)^{-1}.$$

This is equivalent to $\{P(zI-A) = -PA - zFA^*\} \Leftrightarrow \{P = -FA^*\} \Leftrightarrow \{P = -BCA^* + PAA^*\} \Leftrightarrow \{P = BDB^* + P - PBB^*\} \Leftrightarrow \{PBB^* = BDB^*\}$ which is known to be true since $PB = BD$.

1–b) This follows directly from Proposition 4.6.

2) Take $P = -RA^*$. Then $PB = -RA^*B$, which with Proposition 5.2–i) equals RC^*D. With $B = RC^*$ it follows that $PB = BD$ and thus this choice of P satisfies (28).

Now it has to be shown that for this P, $BC - PA = R$. This follows from $BC - PA = BC + RA^*A = BC + R(I - C^*C) = BC + R - BC = R$. □

Proof of Theorem 5.8: Denote the infinite-dimensional matrices $B_\infty := B_k$, $k \to \infty$, and $A_\infty := A_k$, $k \to \infty$. Then we can rewrite (34) as

$$\begin{aligned}&[C_sB_s\ C_sA_sB_s\ C_sA_s^2B_s\ \cdots] =\\ &= [L_0\ L_1\ \cdots][B_\infty\ A_\infty B_\infty\ A_\infty^2 B_\infty\ \cdots].\end{aligned} \tag{A.8}$$

Because of the orthonormality of $[B_\infty\ A_\infty B_\infty\ \cdots]$ postmultiplication of (A.8) with $[B_\infty\ A_\infty B_\infty\ \cdots]^*$ provides

$$[C_sB_s\ C_sA_sB_s\ \cdots][B_\infty\ A_\infty B_\infty\ \cdots]^* = [L_0\ L_1\ L_2\ \cdots]$$

leading to $[L_0\ L_1\ L_2\ \cdots] = \sum_{k=0}^\infty C_sA_s^kB_sB_\infty^*(A_\infty^*)^k$.

We define

$$Q := [Q_0\ Q_1\ Q_2\ \cdots] = \sum_{k=0}^{\infty} A_s^k B_s B_\infty^* (A_\infty^*)^k \tag{A.9}$$

and as a result, $L_k = C_sQ_k$, which equals (35).

Based on (A.9) we can write $A_sQA_\infty^* = Q - B_sB_\infty^*$. With $X = BC$ and P any matrix satisfying $PB = BD$, this leads to

$$A_s[Q_0\ Q_1\ Q_2\ \cdots]\begin{bmatrix} A^* & X^* & X^*P^* & \cdots & \cdots \\ 0 & A^* & X^* & X^*P^* & \cdots \\ 0 & 0 & A^* & X^* & \cdots \\ \vdots & \vdots & 0 & A^* & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = [Q_0\ Q_1\ Q_2\ \cdots] - B_s[B^*\ B^*P^*\ B^*(P^*)^2\ \cdots].$$

The first element of this equation shows $A_sQ_0A^* = Q_0 - B_sB^*$, which equals (36). The ith element leads to

$$A_s[Q_iA^* + \sum_{j=1}^{i} Q_{i-j}X^*(P^*)^{j-1}] = Q_i - B_sB^*(P^*)^i. \tag{A.10}$$

Postmultiplication with P^* gives

$$A_s[Q_iA^*P^* + \sum_{j=1}^{i} Q_{i-j}X^*(P^*)^j] = Q_iP^* - B_sB^*(P^*)^{i+1}. \tag{A.11}$$

Writing (A.10) for $i \to i+1$, shows that

$$A_s[Q_{i+1}A^* + \sum_{j=1}^{i+1} Q_{i-j+1}X^*(P^*)^{j-1}] = Q_{i+1} - B_sB^*(P^*)^{i+1} \tag{A.12}$$

and subtracting (A.11) from (A.12) delivers

$$A_sQ_{i+1}A^* - A_sQ_i[A^*P^* - X^*] = Q_{i+1} - Q_iP^*. \tag{A.13}$$

Note that $P = P(AA^* + BB^*) = PAA^* + BDB^* = PAA^* - BCA^* = -FA^*$, and since $F = X - PA$, (A.13) leads to (37). □

Proof of Proposition 7.1: Part a) A similar result for continuous-time systems is proven in [11]. The discrete-time version follows by applying a bilinear transformation, as is shown in [19].

Part b) The proof is based on an equivalence relation as employed in [11], based on the bilinear transformation. If (A_d, B_d, C_d, D_d) is a balanced realization of a square discrete-time system G_d, then (A_c, B_c, C_c, D_c) is a (continuous-time) balanced realization of a continuous-time system G_c, where

$$\begin{aligned}
A_c &= [A_d - I][A_d + I]^{-1} & A_d &= [I + A_c][I - A_c]^{-1}\\
B_c &= \sqrt{2}[A_d + I]^{-1}B_d & B_d &= \sqrt{2}[I - A_c]^{-1}B_c\\
C_c &= \sqrt{2}C_d[A_d + I]^{-1} & C_d &= \sqrt{2}C_c[I - A_c]^{-1}\\
D_c &= D_d - C_d[A_d + I]^{-1}B_d & D_d &= D_c + C_c[I - A_c]^{-1}B_c.
\end{aligned}$$

Furthermore G_d is inner if and only if G_c is inner. From [11] it follows that G_c is inner if and only if the following conditions are satisfied:

1) $A_c + A_c^* + B_cB_c^* = A_c + A_c^* + C_c^*C_c = 0$
2) $D_cD_c^* = D_c^*D_c = I$
3) $D_c^*C_c + B_c^* = D_cB_c^* + C = 0$.

Given A_d, B_d we can construct A_c, B_c with the equations above, additionally choosing $C_c := B_c^*$, $D_c := -I$ it follows that (A_c, B_c, C_c, D_c) is a balanced realization of an inner function.

Now transforming the continuous-time realization back to the discrete-time domain, with the expressions for C_d, D_d

as given before and employing the relation $(I - A_c)^{-1} = \frac{1}{2}(I + A_d)$, shows that

$$C_d = B_d^*(A_d^* + I)^{-1}(A_d + I) \tag{A.14}$$

$$D_d = B_d^*(A_d^* + I)^{-1}B_d - I \tag{A.15}$$

which completes the balanced realization (A_d, B_d, C_d, D_d).

Using the fact that B_d and C_d have full rank, Proposition 5.2 now implies that premultiplication of C_d and D_d with any unitary matrix U characterizes the intended class of balanced realizations.

Part c) Consider a realization (A, B, C, D) satisfying (44), (45) with $U = I$. For $\bar{F} = [I + A][I + A^*]^{-1}$ it follows immediate that $\bar{F}C^* = B$, and the result follows with Theorem 5.7-2).

Denote $F(\lambda, U) = [\lambda I + B(U - \lambda I)(B^*B)^{-1}B^*]\bar{F}$, with U a unitary matrix and $\lambda \in \boldsymbol{C}$. Then by substitution it can be verified that $F(\lambda, U)C^*U^* = B$. This means that for any $\tilde{C} = UC$ we have constructed a λ-family $F(\lambda, U)$ that satisfies $F(\lambda, U)\tilde{C}^* = B$. Again with Theorem 5.7–2) and choosing $\lambda = 1$ this proves the result. □

Proof of Proposition 8.1: This result follows directly from rewriting the Lyapunov equations in Theorem 5.8 in terms of Kronecker products: see [3], [19]. □

Proof of Theorem 8.3: Consider a single scalar entry of the rational matrix function $W(z) = Z_o(zI - X_o)^{-1}Y_o$, written as $w(z) = \sum_{k=1}^{\infty} w_k z^{-k}$. Then $w(z)$ is convergent for $|z| > \lambda$. Consequently, according to basic theory of power series, see e.g., [20], as employed also in [17], there exists an $\alpha \in \mathbf{R}$ such that for each $\eta > \lambda$, $|w_k \eta^{-k}| \leq \alpha$, leading to $|w_k| \leq \alpha\eta^k$. Since this holds for any entry of $W(z)$, and $W(z) = \sum_{k=1}^{\infty} Vec(L_{k-1})z^{-k}$, there exist scalars α_{ij} such that $|L_k(i,j)| \leq \alpha_{ij}\eta^{k+1}$, with $L_k(i,j)$ being the (i,j)-entry in L_k. Denoting $E(z) := H(z) - \hat{H}^N(z) = \sum_{k=N}^{\infty} L_k V_k(z)$, it follows that $\|E(z)\|_\infty \leq \|\sum_{k=N}^{\infty} L_k V_k(z)\|_\infty \leq \|V_0(z)\|_\infty \sum_{k=N}^{\infty} \|L_k\|_\infty$.

Using the above upper bound for $|L_k(i,j)|$ together with the well-known relation between the ∞-norm and the Frobenius norm, $\|\cdot\|_\infty \leq \|\cdot\|_F$, it follows that

$$\|L_k\|_\infty \leq \eta^{k+1}\sqrt{\sum_{i,j} \alpha_{ij}^2}. \tag{A.16}$$

Substituting this latter upper bound in the derived upper bound for $\|E(z)\|_\infty$, the result of the theorem follows with $c = \|V_0(z)\|_\infty \sqrt{\sum_{i,j} \alpha_{ij}^2}$. □

REFERENCES

[1] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space,* vol. 1. Boston, MA: Pitman Adv. Publ. Program, 1981.

[2] L. Baratchart and M. Olivi, "Inner-unstable factorization of stable rational transfer functions," in *Modeling, Estimation and Control of Systems with Uncertainty,* G.B. DiMasi, A. Gombani and A.B. Kurzhanski, Eds. Boston, MA: Birkhäuser Verlag, 1991, pp. 22–39.

[3] R. Bellman, *Introduction to Matrix Computations.* New York: McGraw-Hill, 1970.

[4] P. M. M. Bongers and P. S. C. Heuberger, "Discrete normalized coprime factorization," in *Proc. 9th Int. Conf. Analysis and Optimization of Systems,* Antibes, France, June 12–15, 1990, pp. 307–313.

[5] C.-C. Chu, "On discrete inner-outer and spectral factorizations," in *Proc. Amer. Contr. Conf.,* Atlanta, GA, 1988, pp. 1699-1700.

[6] P. R. Clement, "Laguerre functions in signal analysis and parameter identification," *J. Franklin Inst.,* vol. 313, no. 2, pp. 85–95, 1982.

[7] G. J. Clowes, "Choice of the time scaling factor for linear system approximations using orthonormal Laguerre functions," *IEEE Trans. Automat. Contr.,* vol. AC-10, no. 5, pp. 487–489, 1965.

[8] C. A. Desoer, R-W Liu, J. Murray and R. Seaks, "Feedback system design: the fractional representation approach to analysis and synthesis," *IEEE Trans. Automat. Contr.,* vol. AC-25, pp. 399–412, 1980.

[9] G. A. Dumont, Y. Fu, and A.-L. Elshafei, "Orthonormal functions in identification and adaptive control," in *Intelligent Tuning and Adaptive Control, Selected Papers from the IFAC Symposium* Oxford, England: Pergamon Press, 1991, pp. 193–198.

[10] B. A. Francis, *A Course in H_∞ Control Theory* (Lecture Notes in Control Information Sciences) vol. 88. Berlin: Springer-Verlag, 1987.

[11] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L_∞-error bounds," *Int. J. Contr.,* vol. 39, 1115–1193, 1984.

[12] K. Glover, J. Lam and J. R. Partington, "Rational approximation of a class of infinite-dimensional systems. I. Singular values of Hankel operators," *Math. Contr. Signals Syst.,* vol. 3, no. 4, pp. 325–344, 1990.

[13] ——, "Rational approximation of a class of infinite-dimensional systems. II. Optimal convergence rates of L_∞ approximants," *Math. Contr. Signals Syst.,* vol. 4, no. 3, pp. 233–246, 1991.

[14] M. J. Gottlieb, "Concerning some polynomials orthogonal on finite or enumerable set of points," *Amer. J. Math.,* vol. 60, pp. 453–458, 1938.

[15] G. Gu, P. P. Khargonekar, and E. B. Lee, "Approximation of infinite-dimensional systems," *IEEE Trans. Automat. Contr.,* vol. 34, no. 6, pp. 610–618, 1989.

[16] S. Gunnarsson and B. Wahlberg, "Some asymptotic results in recursive identification using Laguerre models," *Int. J. Adaptive Contr. Signal Processing,* vol. 5, no. 5, pp. 313–333, 1991.

[17] A. J. Helmicki, C. A. Jacobson, and C. N. Nett, "Control oriented system identification: a worst-case/deterministic approach in H_∞," *IEEE Trans. Automat. Contr.,* vol. 36, no. 10, pp. 1163–1176, 1991.

[18] P. S. C. Heuberger and O. H. Bosgra, "Approximate system identification using system based orthonormal functions," in *Proc. 29th IEEE Conf. Decis. Contr.,* Honolulu, HI, 1990, pp. 1086–1092.

[19] P. S. C. Heuberger, "On approximate system identification with system based orthonormal functions," Ph.D. dissertation, Delft University of Technology, The Netherlands, 1991.

[20] E. Hille, *Analytic Function Theory,* vol. 1. Boston, MA: Ginn, 1959.

[21] W. H. Kautz, "Transient synthesis in the time domain," *IRE Trans. Circuit Theory,* vol. CT-1, pp. 29–39, 1954.

[22] R. E. King and P. N. Paraskevopoulos, "Digital Laguerre filters," *Int. J. Circuit Theory and Appl.,* vol. 5, pp. 81–91, 1977.

[23] ——, "Parametric identification of discrete time SISO systems," *Int. J. Contr.,* vol. 30, pp. 1023–1029, 1979.

[24] Y. W. Lee, "Synthesis of electrical networks by means of the Fourier transforms of Laguerre functions," *J. Math. Physics,* vol. 11, pp. 83–113, 1933.

[25] ——, *Statistical Theory of Communication.* New York: Wiley, 1960.

[26] L. Ljung, *System Identification—Theory for the User.* Englewood Cliffs, NJ: Prentice-Hall, 1987.

[27] P. M. Mäkilä, "Approximation of stable systems by Laguerre filters," *Automatica,* vol. 26, pp. 333–345, 1990.

[28] ——, "Laguerre series approximation of infinite dimensional systems," *Automatica,* vol. 26, no. 6, pp. 985–995, 1990.

[29] ——, "Laguerre methods and H_∞ identification of continuous-time systems," *Int. J. Contr.,* vol. 53, no. 3, pp. 689–707, 1991.

[30] M. Masnadi-Shirazi and N. Ahmed, "Optimum Laguerre networks for a class of discrete-time systems," *IEEE Trans. Signal Processing,* vol. 39, no. 9, pp. 2104–2108, 1991.

[31] Y. Nurges, "Laguerre models in problems of approximation and identification of discrete systems," *Automat. Remote Contr.,* vol. 48, 346–352, 1987.

[32] Y. Nurges and Y. Yaaksoo, "Laguerre state equations for a multivariable discrete system," *Automat. Remote Contr.,* vol. 42, 1601–1603, 1981.

[33] R. Ober and D. McFarlane, "Balanced canonical forms for minimal systems: A normalized coprime factor approach," *Linear Algebra Applicat.,* vol. 122/123/124, pp. 23–64, 1989.

[34] P. N. Paraskevopoulos, "System analysis and synthesis via orthogonal polynomial series and Fourier series," *Math. Comput. Simulation,* vol. 27, 453–469, 1985.

[35] M. Schetzen, "Power-series equivalence of some functional series with applications," *IEEE Trans. Circuit Theory,* vol. CT-17, no. 3, pp. 305–313, 1970.

[36] ——, "Asymptotic optimum Laguerre series," *IEEE Trans. Circuit Theory,* vol. CT-18, no. 5, pp. 493–500, 1971.

[37] G. Szegö, *Orthogonal Polynomials*, 4th ed. Providence, RI: American Mathematical Soc. 1975.
[38] H. Unbehauen and G. P. Rao, "Continuous-time approaches to system identification," *Prepr. 8th IFAC/IFORS Symposium Identification and System Param. Estim.*, Beijing, China, 1988, pp. 60–68.
[39] P. M. J. Van den Hof, P. S. C. Heuberger, and J. Bokor, "Identification with generalized orthonormal basis functions—Statistical analysis and error bounds," *Preprints 10th IFAC Symposium on System Identification*, Copenhagen, vol. 3, 1994, pp. 207–212, to appear in *Automatica*, vol. 31, no. 12
[40] M. Vidyasagar, *Control Systems Synthesis: A Factorization Approach.* Cambridge, MA: MIT Press, 1985.
[41] B. Wahlberg, "On the use of orthogonalized exponentials in system identification," Dept. Electr. Eng., Linköping University, Sweden, Rep. LiTH-ISY-1099, 1990.
[42] B. Wahlberg and E. J. Hannan, "Parametric signal modeling using Laguerre filters," *Annals Appl. Prob.*, vol. 3, pp. 467–496, 1993.
[43] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans. Automat. Contr.*, vol. 36, 551–562, 1991.
[44] ——, "System identification using Kautz models," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1276–1282, 1994.
[45] E. E. Ward, "The calculation of transients in dynamical systems,"*Proc. Cambridge Philos. Soc.,* vol. 50, pp. 49–59, 1954.
[46] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series.* Cambridge, MA: MIT Press, 1949.
[47] C. Zervos, P. R. Bélanger and G.A. Dumont, "On PID controller tuning using orthonormal series identification," *Automatica,* vol. 24, no. 2, pp. 165–175, 1988.

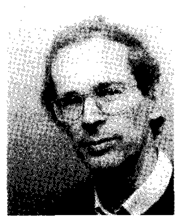
Peter S. C. Heuberger was born in Maastricht, The Netherlands, in 1957. He obtained the M.Sc. degree in Mathematics from the Groningen State University in 1983 and the Ph.D. degree in 1991 from Delft University of Technology.

From 1983 until 1990 he was a Research Assistant in the Mechanical Engineering Systems and Control Group at Delft University of Technology. Since 1991, he has been a staff member of the Dutch National Institute of Public Health and Environmental Protection in Bilthoven, The Netherlands, dealing with modeling and calibration of environmental systems. His main research interests are in issues of system identification and model reduction.

Paul M. J. Van den Hof (S'85–M'88) was born in Maastricht, The Netherlands, in 1957. He obtained the M.Sc. and Ph.D. degrees both from the Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands, in 1982 and 1989, respectively.

From 1986 to 1990 he was an Assistant Professor in the Mechanical Engineering Systems and Control Group at Delft University of Technology, The Netherlands. Since 1991 he has been working in this same group as an Associate Professor. In 1992 he has held a short term visiting position at the Centre for Industrial Control Science, The University of Newcastle, NSW, Australia. His research interests are in issues of system identification, parameterization, and the interplay between identification and robust control design, with applications in mechanical servo systems and industrial process control systems.

Dr. Van den Hof is an Associate Editor of *Automatica.* For his M.Sc thesis he received the 1983 Control Systems Award *(Regeltechniekprijs)* of the *Royal Dutch Institute of Engineers (KIvI).*



Okko H. Bosgra was born in Groningen, The Netherlands, in 1944. He obtained the M.Sc. degree in mechanical engineering from Delft University of Technology in 1968.

From 1981–1985 he held a Professorship in the Department of Physics at the Agricultural University of Wageningen, The Netherlands. Since 1986 he is a Full Professor in control engineering, heading the Mechanical Engineering Systems and Control Group of Delft University of Technology, The Netherlands. His current research interests are inthe theory of robust identification and control design and their application to industrial problems in the process control field and in the field of mechanical servo motion control.

Prof. Bosgra is a member of the EUCA (European Union Control Association) Governing Board, a founding member of the Dutch Systems and Control Theory Network, and an Editor-At-Large of the new *European Journal of Control.*