2

Construction and Analysis

Bo Wahlberg[1] and Tomás Oliveira e Silva[2]

[1] KTH, Stockholm, Sweden
[2] University of Aveiro, Aveiro, Portugal

2.1 Objectives

In this chapter, we introduce some basics methods to construct and analyze rational orthonormal functions. Two different techniques are described in detail:

- The Gram-Schmidt procedure applied to transfer functions.
- A matrix factorization technique based on state-space models.

2.2 Rational Orthogonal Functions

The theory of orthogonal functions is indeed classical and one of the most developed fields of mathematics. See *e.g.* [40, 61, 96, 139, 201, 204, 261, 294] for introductions to and overviews of this area.

2.2.1 Notation

We will use the notation introduced in Chapter 1 and will start by repeating the most important definitions: The Hardy space of square integrable functions on the unit circle, $\mathbb{T}$, and analytic outside the unit circle including infinity, $\mathbb{E}$, is denoted by $\mathcal{H}_2(\mathbb{E})$. The corresponding inner product of $X(z), Y(z) \in \mathcal{H}_2(\mathbb{E})$ is defined by

$$\langle X, Y \rangle := \frac{1}{2\pi i} \oint_{\mathbb{T}} X(z) Y^*(1/z^*) \frac{dz}{z},$$

where * denotes complex conjugate. Notice that if $|z| = 1$ then $1/z^* = z$. The squared norm of $X(z)$ equals $||X||^2 = \langle X, X \rangle$. Two transfer functions $F_1(z)$ and $F_2(z)$ are orthonormal if $||F_1|| = ||F_2|| = 1$ and $\langle F_1, F_2 \rangle = 0$. For matrix valued functions we will use the cross-Gramian or matrix 'outer product'

$$[\![X, Y]\!] := \frac{1}{2\pi i} \oint_{\mathbb{T}} X(z) Y^*(1/z^*) \frac{dz}{z},$$

where * denotes complex conjugate transpose for matrices. For vector valued functions $X(z), Y(z) \in \mathcal{H}_2^{n\times 1}(\mathbb{E})$, the vector space inner product equals $\langle X, Y\rangle = [\![X^T, Y^T]\!]$. An import class of functions are functions with real valued impulse responses, *i.e.* $Y(z) = \sum_k y_k z^{-k}$, where $\{y_k\}$ are real numbers. Then $Y^*(1/z^*) = Y^T(/z)$, and the cross-Gramian is a real valued matrix satisfying $[[\![X, Y]\!]]^T = [\![Y, X]\!]$.

2.2.2 The Gram-Schmidt Procedure

Consider two linearly independent functions $\bar{F}_1(z), \bar{F}_2(z) \in \mathcal{H}_2(\mathbb{E})$[1], which form a basis for their span. An orthonormal basis for this subspace can easily be obtained using the Gram-Schmidt procedure,

$$\begin{aligned} K_1(z) &= \bar{F}_1(z), \\ F_1(z) &= \frac{K_1(z)}{||K_1||}, \\ K_2(z) &= \bar{F}_2(z) - \langle \bar{F}_2, F_1\rangle F_1(z), \\ F_2(z) &= \frac{K_2(z)}{||K_2||}. \end{aligned}$$

Then $||F_1|| = ||F_2|| = 1$ and

$$\langle F_2, F_1\rangle = \frac{1}{||K_2||}\left(\langle \bar{F}_2, F_1\rangle - \langle \bar{F}_2, F_1\rangle ||F_1||^2\right) = 0,$$

because $||F_1|| = 1$. Hence, $F_1(z)$ and $F_2(z)$ are orthonormal, and they span the same space as $\bar{F}_1(z)$ and $\bar{F}_2(z)$. Notice that $K_2(z)$ is the projection of $\bar{F}_2(z)$ onto the orthogonal complement to the space spanned by $F_1(z)$.

In general, $F_i(z)$ can recursively be determined by normalizing the projection of $\bar{F}_i(z)$ onto the orthogonal complement of $\mathrm{Sp}\{F_1(z) \ldots F_{i-1}(z)\}$. Hence, it is in principle easy to construct orthonormal functions from a given set of functions. Also notice that the orthonormal functions are by no means unique. For example, changing the ordering of the functions will affect the corresponding basis, but the span of the basis functions will remain the same. The following example illustrates this idea for rational functions with real poles.

Example 2.1. Consider the functions $\bar{F}_1(z) = 1/(z-a)$ and $\bar{F}_2(z) = 1/(z-b)$, $-1 < a, b < 1$, $a \neq b$. These functions are linearly independent but are not orthonormal in the space $\mathcal{H}_2(\mathbb{E})$. Using the Gram-Schmidt procedure, we find

[1] Overbar is used here to mark functions that are not orthonormal, and should not be confused by complex conjugate (denoted by * in this chapter).

$$F_1(z) = \frac{1}{z-a} / (1/\sqrt{1-a^2}) = \frac{\sqrt{1-a^2}}{z-a},$$
$$K_2(z) = \frac{1}{z-b} - \frac{\sqrt{1-a^2}}{1-ab} \frac{\sqrt{1-a^2}}{z-a} = \frac{b-a}{1-ab} \frac{1}{z-b} \frac{1-az}{z-a},$$
$$F_2(z) = \frac{\sqrt{1-b^2}}{z-b} \frac{1-az}{z-a}.$$

The Cauchy integral formula can be used to calculate the scalar products, *e.g.*

$$\langle \bar{F}_2, F_1 \rangle = \frac{1}{2\pi i} \oint_{\mathbb{T}} \frac{F_1(1/z)}{z-b} \frac{dz}{z} = F_1(1/b)/b,$$

because the only pole of the integrand inside the unit circle is $z = b$. Note that $z = 0$ is not a pole because $F_1(\infty) = 0$. The functions $F_1(z)$ and $F_2(z)$ are now by construction orthonormal.

The function $F_2(z)$ has a special structure. It can be factored as

$$F_2(z) = \frac{\sqrt{1-b^2}}{z-b} G_b(z),$$

with

$$G_b(z) = \frac{1-az}{z-a}.$$

The transfer function $G_b(z)$ is called a first-order all-pass transfer function, and satisfies $G_b(z)G_b(1/z) = 1$.

These results can be extended by including a third function:

$$\bar{F}_3(z) = \frac{1}{z-c}, \ -1 < c < 1, \ c \neq a, b.$$

The Gram-Schmidt procedure then gives

$$K_3(z) = \bar{F}_3(z) - \langle \bar{F}_3, F_2 \rangle F_2(z) - \langle \bar{F}_3, F_1 \rangle F_1(z),$$
$$F_3(z) = \frac{K_3(z)}{||K_3||},$$

where $\langle \bar{F}_3, F_1 \rangle = F_1(1/c)/c$ and $\langle \bar{F}_3, F_2 \rangle = F_2(1/c)/c$. Further calculations give

$$F_3(z) = \frac{\sqrt{1-c^2}}{z-c} \frac{1-az}{z-a} \frac{1-bz}{z-b},$$

which, by the Gram-Schmidt construction, is orthogonal to $F_1(z)$ and $F_2(z)$. Projecting $\bar{F}_3$ onto the space orthogonal to the first two basis functions corresponds to multiplication by an all-pass function $G_b(z)$ with the same poles as the previous two basis functions,

$$G_b(z) = \frac{(1-az)(1-bz)}{(z-a)(z-b)}, \quad G_b(z)G_b(1/z) = 1.$$

The normalization is the same as for the first basis function, since $||G_b|| = 1$.

This example can be extended to a set of n first-order transfer functions with distinct real poles.

2.2.3 The Takenaka-Malmquist Functions

In the previous section, we showed how to construct orthonormal basis functions using the Gram-Schmidt procedure. For the general case with possibly multiple complex poles, this will result in the so-called Takenaka-Malmquist functions

$$F_k(z) = \frac{\sqrt{1-|\xi_k|^2}}{z-\xi_k} \prod_{i=1}^{k-1} \left[\frac{1-\xi_i^* z}{z-\xi_i} \right], \; k=1,2,\dots, \quad \xi_i \in \mathbb{C}, \; |\xi_i|<1. \quad (2.1)$$

This structure can be anticipated from the Cauchy integral formula,

$$\langle F_k, F_1 \rangle = F_k(1/\xi_1^*)\sqrt{1-|\xi_1|^2} = 0, \quad \forall k > 1,$$

which forces $F_k(z)$ to have a zero at $1/\xi_1^*$. Using this argument for $\langle F_k, F_j \rangle$, $j = 2, \dots k-1$, leads to the zero structure $F_k(1/\xi_j^*) = 0$, $j < k$.
The projection property of all-pass function

$$G_b(z) = \prod_{i=1}^{k-1} \left[\frac{1-\xi_i^* z}{z-\xi_i} \right],$$

i.e.

$$\langle F_j, F G_b \rangle = 0, \; j \leq k-1, \quad \forall\, F(z) \in \mathcal{H}_2\big(\mathbb{E}\big),$$

is known in a somewhat more abstract setting as the Beurling-Lax theorem, [254, 319].
The Takenaka-Malmquist functions will in general have complex impulse responses. The all-pass function $G_b(z)$ will then satisfy the condition

$$G_b(z) G_b^*(1/z^*) = 1.$$

2.2.4 Error Bounds

Let the poles of of the basis functions $\{F_k(z)\}_{k=1,\dots n}$ be $\{\xi_i \in \mathbb{C}, \, |\xi_i| < 1, \, i = 1 \dots n\}$, and define the corresponding all-pass transfer function

$$G_b(z) = \prod_{i=1}^{n} \left[\frac{1-\xi_i^* z}{z-\xi_i} \right].$$

Decompose $G_0(z) \in H_2(\mathbb{T})$ as

$$G_0(z) = G_n(z) + E_n(z) G_b(z),$$

where $G_n(z) \in \mathrm{Sp}\{F_k(z)\}_{k=1,\dots n}$, and $E_n(z) \in H_2(\mathbb{T})$. Then $E_n(z)G_b(z)$ is orthogonal to $G_n(z)$ because of the projection property of $G_b(z)$ and hence $G_n(z)$ is the optimal approximation of $G_0(z)$ in the space spanned by the given basis functions. Now make a (partial fraction expansion) decomposition of

$$G_0(z)[G_b(z)]^{-1} = D_n(z) + E_n(z),$$

where the poles of $D_n(z)$ equals the zeros of $G_b(z)$, which are outside the unit circle, and $E_n(z)$ has the same poles as $G_0(z)$, which are inside the unit circle. Hence, we have decomposed $G_0(z)[G_b(z)]^{-1}$ into its stable part and its completely unstable part. From the Cauchy integral formula we then have

$$E_n(z) = \frac{1}{2\pi i} \oint_{\mathbb{T}} \frac{G_0(\zeta)}{\zeta - z} G_b^*(1/\zeta^*) d\zeta, \quad |z| \geq 1.$$

Here we have used $[G_b(z)]^{-1} = G_b^*(1/z^*)$. Hence

$$G_0(z) - G_n(z) = \left[\frac{1}{2\pi i} \oint_{\mathbb{T}} \frac{G_0(\zeta)}{\zeta - z} G_b^*(1/\zeta^*) d\zeta \right] G_b(z).$$

If

$$G_0(z) = \sum_{j=1}^{n_0} \frac{b_j}{z - p_j} \;\Rightarrow\; E_n(z) = \sum_{j=1}^{n_0} \frac{b_j\, G_b^*(1/p_j^*)}{z - p_j}$$

and thus

$$G_0(z) - G_n(z) = \left[\sum_{j=1}^{n_0} \frac{b_j\, G_b^*(1/p_j^*)}{z - p_j} \right] G_b(z). \tag{2.2}$$

Example 2.2. Let us illustrate the calculations for the a first-order system and the FIR basis

$$G_0(z) = \frac{1}{z - p}, \quad F_k(z) = z^{-k},\; G_b(z) = z^{-n}.$$

Here

$$G_0(z) = z^{-1} + \ldots + p^{n-1} z^{-n} + p^n z^{-n} \frac{1}{z - p},$$

$$D_n(z) = \frac{z^{n-1} + \ldots + p^{n-1}}{z^n}, \quad E_n(z) = \frac{p^n}{z - p},$$

and thus

$$G_0(z) - G_n(z) = \frac{1}{z - p} p^n z^{-n} = G_0(z) G_b(1/p) G_b(z).$$

2.2.5 Completeness

The above expressions for the approximation error have previously been derived in *e.g.* [224]. They are in particular useful to bound the approximation error on the unit circle, *i.e.*

$$|G_0(e^{i\omega}) - G_n(e^{i\omega})| = \left| \sum_{j=1}^{n_0} \frac{b_j}{e^{i\omega} - p_j} G_b^*(1/p_j^*) \right| \leq \sum_{j=1}^{n_0} \left| \frac{b_j}{e^{i\omega} - p_j} \right| \left| G_b^*(1/p_j^*) \right|, \tag{2.3}$$

because $|G_b(e^{\omega})| = 1$. The rate of convergence to zero of the right-hand side of the error expression (2.3) as n tends to infinity is closely coupled to the size of $|G_b^*(1/\zeta^*)|$, $|\zeta| < 1$. In [16] the following upper bound is derived

$$|G_b^*(1/\zeta^*)| \leq \exp\left[-\frac{1}{2}(1 - |\zeta|) \sum_{j=1}^{n} (1 - |\xi_j|) \right], \quad |\zeta| < 1.$$

If

$$\sum_{j=1}^{\infty} (1 - |\xi_j|) = \infty \tag{2.4}$$

this implies that $|G_b^*(1/\zeta^*)| \to 0$ as $n \to \infty$ and the approximation error will tend to zero. It turns out that the so-called Szász condition (2.4) implies that the Takenaka-Malmquist functions are complete in $H_2(\mathbb{T})$ (and also in $H_p(\mathbb{T})$, $1 \leq p < \infty$, and in the disc algebra of functions analytic outside and continuous on the unit circle). This means that the approximation error measured by the corresponding norm can be made arbitrarily small by choosing a large enough value for n. If the Szász condition (2.4) is not satisfied, the Blascke product $G_b(z)$ will converge to an $H_2(\mathbb{T})$ function, which will be orthogonal to $F_k(z)$ for all k. Hence, the Takenaka-Malmquist functions are then not complete, and thus the condition (2.4) is both necessary and sufficient. The Szász condition has a long history and was already known by Takenaka [296]. The original paper, [292], by Szász contains a rather complete analysis of the closure of certain sets of rational functions in various spaces. Recent work in this area includes [16].

2.2.6 Real Impulse Response

The Z-transform $F(z)$ of a real sequence satisfies $F(z) = F^*(z^*)$. Hence, complex poles (and zeros) will appear in complex conjugate pairs. As previously mentioned, the Takenaka-Malmquist functions will in general not have real impulse responses. If $\xi_{j+1} = \xi_j^*$ it is possible to form linear combinations of $F_j(z)$ and $F_{j+1}(z)$ to obtain two orthonormal functions with real impulse responses, which span the same space.

Deeper insights into the construction of orthonormal functions can be gained by using state-space models and matrix algebra. This is the topic of the next section.

2.3 State-space Theory

2.3.1 Introduction

The basic idea of this chapter is to construct and analyze state-space models

$$x(t+1) = Ax(t) + Bu(t),$$

where $u(t)$ is the input signal, and $x(t) \in \mathbb{R}^n$ is the state vector, for which the transfer functions from the input to the states,

$$V(z) = [F_1(z) \dots F_n(z)]^T = (zI - A)^{-1}B,$$

are orthonormal. This means that

$$[\![V, V]\!] = \frac{1}{2\pi i} \oint_{\mathbb{T}} V(z)V^T(1/z)\frac{dz}{z} = I \text{ (the identity matrix)}.$$

Here we will specify the eigenvalues of A, which are the poles of $\{F_k(z)\}$. We will only study state-space models with real A, B-matrices, *i.e.* with real valued impulse response.

To understand the basic ideas of using state-space theory to find rational orthonormal basis functions, consider the following example.

Example 2.3. Returning to Example 2.1, consider the following state-space model

$$\begin{aligned} \bar{x}_1(t+1) &= a\bar{x}_1(t) + u(t) \\ \bar{x}_2(t+1) &= b\bar{x}_2(t) + u(t). \end{aligned}$$

The transfer functions $\bar{F}_1(z) = 1/(z-a)$ and $\bar{F}_2(z) = 1/(z-b)$ are then the transfer functions from the input $u(t)$ to the states $\bar{x}_1(t)$ and $\bar{x}_2(t)$, respectively. Define the state vector $\bar{x}(t) = [\bar{x}_1(t)\ \bar{x}_2(t)]^T$ and the corresponding state-space matrices

$$\bar{A} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The input-to-state vector transfer function

$$\bar{V}(z) = (zI - \bar{A})^{-1}\bar{B}$$

here equals

$$\bar{V}(z) = \begin{bmatrix} \bar{F}_1(z) \\ \bar{F}_2(z) \end{bmatrix} = \begin{bmatrix} \frac{1}{z-a} \\ \frac{1}{z-b} \end{bmatrix}.$$

2.3.2 State Covariance Matrix

Assume that the input signal $u(t)$ to the system is a zero mean white noise process with variance 1, *i.e.*

$$\mathsf{E}\{u(t)u(t+k)\} = \begin{cases} 1, \; k = 0 \\ 0, \; k \neq 0 \end{cases}.$$

The state covariance matrix

$$\bar{P} = \mathsf{E}\{\bar{x}(t)\bar{x}^T(t)\}, \tag{2.5}$$

of a stable state-space model

$$\bar{x}(t+1) = \bar{A}\bar{x}(t) + \bar{B}u(t),$$

satisfies the Lyapunov equation

$$\bar{P} = \bar{A}\bar{P}\bar{A}^T + \bar{B}\bar{B}^T. \tag{2.6}$$

The state covariance matrix $\bar{P}$ also equals the so-called *controllability* Gramian of the state-space model. The reason why we are interested in the state covariance matrix is that $\bar{P}$ corresponds to the matrix 'outer product' of the state vector transfer function $\bar{V}(z)$. That is:

$$\begin{aligned} \bar{P} &= \frac{1}{2\pi i} \oint_{\mathbb{T}} \bar{V}(z)\bar{V}^T(1/z)\frac{dz}{z} \\ &= [\![\bar{V}, \bar{V}]\!] = \begin{bmatrix} \langle \bar{F}_1, \bar{F}_1 \rangle & \langle \bar{F}_1, \bar{F}_2 \rangle \\ \langle \bar{F}_2, \bar{F}_1 \rangle & \langle \bar{F}_2, \bar{F}_2 \rangle \end{bmatrix} \end{aligned}$$

This follows from

$$\bar{P} = \frac{1}{2\pi i} \oint_{\mathbb{T}} (zI - \bar{A})^{-1}\bar{B}\bar{B}^T(z^{-1}I - \bar{A}^T)^{-1}\frac{dz}{z} = \sum_{k=0}^{\infty} \bar{A}^k\bar{B}\bar{B}^T(\bar{A}^T)^k,$$

where the last equality follows using the power series expansion of $(zI - \bar{A})^{-1}$, satisfies Equation (2.6).

Example 2.4. For Example 2.3 we have

$$\bar{P} = \begin{bmatrix} \dfrac{1}{1-a^2} & \dfrac{1}{1-ab} \\ \dfrac{1}{1-ab} & \dfrac{1}{1-b^2} \end{bmatrix}. \tag{2.7}$$

2.3.3 Input Balanced State-space Realization

The basic idea is now to find a new state-space realization for which the state covariance matrix equals the identity matrix, $P = I$. The corresponding input

to state transfer functions will then be orthonormal. Furthermore, they will span the same space as the original functions, as only linear transformations are considered. A state-space realization for which $\bar{P} = I$ is called *input balanced*, [198].
Let $x(t) = T\bar{x}(t)$, where T is a square non-singular transformation matrix, and denote the transformed state covariance matrix by

$$P = \mathsf{E}\{x(t)x^T(t)\} = T\bar{P}T^T.$$

The orthonormalization problem now corresponds to finding a transformation matrix T such that $T\bar{P}T^T = I$, where $\bar{P}$ is a given symmetric positive definite matrix. This is a standard problem in linear algebra. The solution is to take T equal to a square (matrix) root of $\bar{P}^{-1}$, *i.e.* $\bar{P}^{-1} = T^T T$. This follows from

$$T\bar{P}T^T = T[\bar{P}^{-1}]^{-1}T^T = T[T^T T]^{-1}T^T = I.$$

Notice that we have to assume that $\bar{P}$ is nonsingular. This is the same as assuming that the state-space model is controllable, as $\bar{P}$ also equals the controllability Gramian. This corresponds to the assumption that original functions are linearly independent.

The square root of a matrix is by no means unique. Let Q be a orthogonal matrix, $Q^T Q = I$. Then QT is also a square root of $\bar{P}^{-1}$, and thus a permissible transformation. We refer to *e.g.* [103,155] or any other standard book in matrix computations for a detailed discussion of this topic.

An interesting option is to apply the Gram-Schmidt QR-algorithm to a permissible transformation, $T = QR$, where Q is orthogonal and R is upper triangular, and then use the transformation $x(t) = R\bar{x}$. By working on permuted rows and columns of $\bar{P}$, it is also possible to find a lower triangular matrix L (using the Gram-Schmidt QR-algorithm), such that $x(t) = L\bar{x}$ implies that $P = I$. All this is more or less the same as applying the Gram-Schmidt procedure directly to the transfer functions.

Example 2.5. For Example 2.3 we can take

$$L = \begin{bmatrix} \sqrt{1-a^2} & 0 \\ -\dfrac{\sqrt{1-b^2}(1-a^2)}{a-b} & \dfrac{\sqrt{1-b^2}(1-ab)}{a-b} \end{bmatrix},$$

which is obtained by identifying the implicit transformation used in the Gram-Schmidt procedure in Example 2.1. It is easy to verify that $L\bar{P}L^T = I$, where $\bar{P}$ is given by (2.7).

2.3.4 Construction Method

Summarizing the observations in the above example leads to the following method for construction of orthonormal functions:

1. Consider the set spanned by the n linearly independent rational transfer functions
$$\{\bar{F}_1(z) \dots \bar{F}_n(z)\}.$$
2. Determine a controllable state-space model
$$\bar{x}(t+1) = \bar{A}\bar{x}(t) + \bar{B}u(t)$$
with input-to-state transfer functions
$$\bar{V}(z) = [\bar{F}_1(z) \dots \bar{F}_n(z)]^T = (zI - \bar{A})^{-1}\bar{B}.$$
The eigenvalues of $\bar{A}$ equal the poles of $\{\bar{F}_1(z) \dots \bar{F}_n(z)\}$.
3. Determine the state covariance matrix (controllability Gramian) $\bar{P}$ by solving the Lyapunov equation
$$\bar{P} = \bar{A}\bar{P}\bar{A}^T + \bar{B}\bar{B}^T.$$
This is a linear equation in the elements of $\bar{P}$, for which there exists efficient numerical algorithms.
4. Find a matrix square root T of the inverse of $\bar{P}$, *i.e.* $T\bar{P}T^T = I$. The solution is not unique. There exist several numerically robust algorithms for solving this problem.
5. Make a transformation of the states, $x(t) = T\bar{x}(t)$. The transformed state-space model is
$$x(t+1) = Ax(t) + Bu(t), \quad A = T\bar{A}T^{-1},\ B = T\bar{B}.$$
The new input-to-state transfer function equals
$$V(z) = (zI - A)^{-1}B = T\bar{V}(z).$$
The components of $V(z) = [F_1(z) \dots F_n(z)]^T$ then form an orthonormal basis for $\text{Sp}\{\bar{F}_1(z) \dots \bar{F}_n(z)\}$.

2.4 All-pass Transfer Function

As first noted by Mullis and Roberts [257], all-pass transfer functions provide a powerful framework for factorization of state covariance matrices. In this section, we will summarize this theory.

Consider a single input single output asymptotically stable all-pass transfer function $H(z)$ of order m with real valued impulse response,

$$H(z)H(1/z) = 1, \tag{2.8}$$

completely determined by its poles $\{\xi_i \in \mathbb{C},\ |\xi_i| < 1\}$,

$$H(z) = \prod_{i=1}^{m} \left[\frac{1 - \xi_i^* z}{z - \xi_i} \right]. \tag{2.9}$$

Such a transfer function is often called an inner function, and the representation (2.9) is called a Blaschke product. We will use the notation $H(z)$ for an arbitrary all-pass transfer function, and the notation $G_b(z)$, as in the previous section, for all-pass filters directly connected to certain orthogonal basis functions.

Let (A, B, C, D) be a minimal input balanced state-space realization of $H(z)$. The corresponding state covariance matrix satisfies $P = APA^T + BB^T$ for unit variance white noise input. Also recall that we can ensure that $P = I$, by a proper choice of state vector.

2.4.1 Orthogonal State-space Models

The impulse response of the all-pass system can be written in terms of the state-space matrices,

$$h_k = \begin{cases} D, & k = 0 \\ CA^{k-1}B, & k \geq 1 \end{cases}.$$

The all-pass property (2.8) implies

$$\sum_{k=1}^{\infty} h_k h_{k+j} = \begin{cases} 1, & j = 0 \\ 0, & j \geq 1 \end{cases},$$

which can be re-expressed as

$$\begin{aligned} DD^T + CPC^T &= 1, \\ CA^{j-1}(BD^T + APC^T) &= 0, \quad j \geq 1. \end{aligned}$$

Here P is, as before, the controllability Gramian. The last equation implies that $BD^T + APC^T$ is in the null-space of the observability matrix of $H(z)$. Because the state-space realization of $H(z)$ is assumed to be minimal, and hence observable, this implies that $BD^T + APC^T = 0$. Using $P = I$, we then obtain the three conditions

$$\begin{aligned} AA^T + BB^T &= I, \\ DD^T + CC^T &= 1, \\ BD^T + AC^T &= 0, \end{aligned}$$

or more compactly

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^T = I, \tag{2.10}$$

i.e. the corresponding matrix is orthogonal. A state-space realization

$$\begin{bmatrix} x(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix},$$

of $H(z)$ which satisfies (2.10) is called orthogonal.

The above characterization of SISO all-pass filters also holds for multi-input multi-output (MIMO) all-pass filters. A MIMO all-pass system with transfer function $H(z) = D + C(zI - A)^{-1}B$ is orthogonal if and only if it has an orthogonal realization. Note that $H(z)$ is all-pass if and only if $H(z)H^T(1/z) = I$.

2.4.2 Connections of All-pass Filters

Given a set of poles, we can construct an orthogonal state-space realization of the corresponding all-pass transfer function $H(z)$. Different orthogonal realizations and factorizations of

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{2.11}$$

will lead to different orthonormal basis functions. Notice that the matrix (2.11) is by no means unique, as a unitary state-space transformation does not change the orthonormality property.

The key observation of Roberts and Mullis, [257], is that orthogonality and the all-pass property are preserved through several different connections of all-pass filters (some or all of these filters may have more than one input and/or more than one output). We will be most interested in cascade connections, for which the following key result holds.

Theorem 2.1. *Consider two orthogonal all-pass filters $H_1(z)$ and $H_2(z)$ with state vectors $x_1(t)$ and $x_2(t)$, respectively. Then the cascade (serial) connection $H_2(z)\,H_1(z)$ is also all-pass and orthogonal with the state vector:*

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

The result can of course be generalized to an arbitrary cascade of all-pass filter as shown in the figure below.

The proof of Theorem 2.1 is constructive. Define the orthogonal state-space realizations of $H_1(z)$ and $H_2(z)$ as

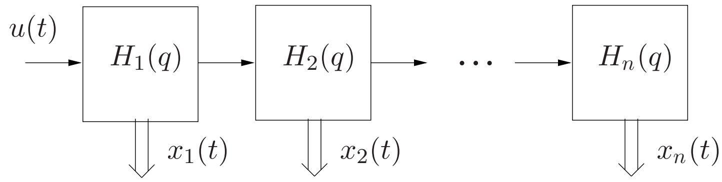

Fig. 2.1. Cascade all-pass filter network. Notice that the transfer functions from input-to-states are orthonormal, and hence the factorization provides a parameterization of different classes of orthonormal basis functions.

$$\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}, \quad \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix},$$

with input signals $u_1(t)$ and $u_2(t)$ and output signals $y_1(t)$ and $y_2(t)$, respectively. By combining the states, we obtain

$$\begin{aligned} \begin{bmatrix} x_1(t+1) \\ x_2(t+1) \\ y_2(t) \end{bmatrix} &= \begin{bmatrix} I & 0 & 0 \\ 0 & A_2 & B_2 \\ 0 & C_2 & D_2 \end{bmatrix} \begin{bmatrix} x_1(t+1) \\ x_2(t) \\ y_1(t) \end{bmatrix} \\ &= \begin{bmatrix} I & 0 & 0 \\ 0 & A_2 & B_2 \\ 0 & C_2 & D_2 \end{bmatrix} \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & I & 0 \\ C_1 & 0 & D_1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ u_1(t) \end{bmatrix} \end{aligned}$$

Hence, the total matrix for the cascaded system

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & A_2 & B_2 \\ 0 & C_2 & D_2 \end{bmatrix} \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & I & 0 \\ C_1 & 0 & D_1 \end{bmatrix} = \begin{bmatrix} A_1 & 0 & B_1 \\ B_2C_1 & A_2 & B_2D_1 \\ D_2C_1 & C_2 & D_2D_1 \end{bmatrix}$$

is orthonormal, as it is the product of two orthonormal matrices. This proves the result.

2.4.3 Feedback Connections of All-pass Filters

Feedback connections in an orthogonal all-pass filter also give rise to orthogonal all-pass filters. To show this, consider Example 10.4.3 in the book by Roberts and Mullis, [257], where feedback is introduced in the orthogonal all-pass filter described by the state variable equations

$$\begin{bmatrix} x(t+1) \\ y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} x(t) \\ u_1(t) \\ u_2(t) \end{bmatrix} \tag{2.12}$$

via the connection $u_2(t) = y_2(t)$. This connection implies that D_{22} is a square matrix. In the above formulas delay free loops will be present unless $D_{22} = 0$. Assuming that $I - D_{22}$ is invertible, elementary algebra yield

$$\begin{bmatrix} x(t+1) \\ y_1(t) \end{bmatrix} = \left\{ \begin{bmatrix} A & B_1 \\ C_1 & D_{11} \end{bmatrix} + \begin{bmatrix} B_2 \\ D_{12} \end{bmatrix} (I - D_{22})^{-1} \begin{bmatrix} C_2 & D_{21} \end{bmatrix} \right\} \begin{bmatrix} x(t) \\ u_1(t) \end{bmatrix}.$$

The matrix inside curly braces is the DC gain of another orthogonal all-pass filter, with state-space realization

$$\left[\begin{array}{c|cc} D_{22} & C_2 & D_{21} \\ \hline B_2 & A & B_1 \\ D_{12} & C_1 & D_{11} \end{array}\right].$$

But the DC gain of orthogonal all-pass filters must be an orthogonal matrix because

$$H(z)H^T(1/z) = I.$$

Thus, feedback connections of the type described above give rise to orthogonal all-pass filters. The same result holds if the feedback connection is made via a general orthogonal all-pass filter (with the appropriate number of input and output signals). Writing the state-space realizations of the two filters together, and using an obvious notation for the (feedback) orthogonal filter, yields

$$\begin{bmatrix} x(t+1) \\ y_1(t) \\ y_2(t) \\ x_f(t+1) \\ y_f(t) \end{bmatrix} = \begin{bmatrix} A_1 & B_1 & B_2 & 0 & 0 \\ C_1 & D_{11} & D_{12} & 0 & 0 \\ C_2 & D_{21} & D_{22} & 0 & 0 \\ 0 & 0 & 0 & A_f & B_f \\ 0 & 0 & 0 & C_f & D_f \end{bmatrix} \begin{bmatrix} x(t) \\ u_1(t) \\ u_2(t) \\ x_f(t) \\ u_f(t) \end{bmatrix}$$

with the connections $u_f(t) = y_2(t)$ and $u_2(t) = y_f(t)$. These connections are of the type used initially, with the D_{22} of the first example replaced by the direct sum of D_{22} and D_f of the second. Thus we have proved the following result, which will be used in Subsection 2.6:

Theorem 2.2. *If $I - D_{22}$, where D_{22} is defined in Equation (2.12), is invertible, then the feedback connection of an orthogonal all-pass filter made via another orthogonal all-pass filter is also orthogonal and all-pass.*

2.4.4 Orthogonal Filters Based on Feedback of Resonators

Define

$$\begin{bmatrix} x(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} A_r & B_r \\ C_r & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ e(t) \end{bmatrix}.$$

Now apply the feedback $e(t) = u(t) - y(t)$ to obtain the closed loop system

$$\begin{bmatrix} x(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} A_r - B_r C_r & B_r \\ C_r & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}.$$

This will be an orthogonal realization if

$$\begin{bmatrix} A_r - B_r C_r & B_r \\ C_r & 0 \end{bmatrix}^T \begin{bmatrix} A_r - B_r C_r & B_r \\ C_r & 0 \end{bmatrix} =$$

$$\begin{bmatrix} A_r^T A_r - C_r^T B_r^T A_r - A_r^T B_r C_r + C_r^T B_r^T B_r C_r + C_r^T C_r & A_r^T B_r - C_r^T B_r^T B_r \\ B_r^T A_r - B_r^T B_r C_r & B_r^T B_r \end{bmatrix}$$

$$= \begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix},$$

with the simple solution

$$B_r^T B_r = 1, \quad C_r = B_r^T A_r, \quad A_r^T A_r = I.$$

Let $G_r(z)$ be the transfer function of the open loop system, $G_r(z) = C_r(zI - A_r)^{-1} B_r$, and $H(z)$ be the transfer function of the corresponding closed loop all-pass system. Then, using some manipulation, we can write the open loop transfer function as:

$$G_r(z) = \frac{H(z)}{1 - H(z)}.$$

Hence, the eigenvalues of A_r (the poles of $G_r(z)$) satisfy $H(\lambda_r) = 1$ and thus $|\lambda_r| = 1$. This can also be seen from the orthonormality condition $A_r^T A_r = I$. Such a system, with all poles on the unit circle, is called a *resonator*. Here we have assumed a delay-free closed loop, *i.e.* no direct term in $G_r(z)$. This will introduce a pure delay in $H(z)$, *i.e.* a pole in $z = 0$, which is a restriction which may have to be removed before construction of the corresponding input-to-state orthogonal transfer functions $V(z)$. This approach has been extensively studied in [240, 241].

2.4.5 Basic Orthogonal All-pass Filters

The simplest all-pass building blocks are:

- First-order all-pass filters with transfer functions

$$H_i(z) = \frac{z - a_i}{1 - a_i z}, \qquad -1 < a_i < 1,$$

  with orthogonal state-space realizations

$$\begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} = \begin{bmatrix} a_i & \sqrt{1 - a_i^2} \\ \sqrt{1 - a_i^2} & -a_i \end{bmatrix}. \tag{2.13}$$

- Second-order all-pass filters with transfer functions

$$H_j(z) = \frac{-c_j z^2 + b_j(c_j - 1)z + 1}{z^2 + b_j(c_j - 1)z - c_j}, \quad -1 < b_j < 1,\ -1 < c_j < 1,$$

  with orthonormal state-space realizations

$$\begin{bmatrix} A_j & B_j \\ C_j & D_j \end{bmatrix} = \left[\begin{array}{cc|c} b_j & c_j\sqrt{1-b_j^2} & \sqrt{1-b_j^2}\sqrt{1-c_j^2} \\ \sqrt{1-b_j^2} & -b_jc_j & -b_j\sqrt{1-c_j^2} \\ \hline 0 & \sqrt{1-c_j^2} & -c_j \end{array}\right]. \tag{2.14}$$

It should be noted that the realization is not unique.

- First-order all-pass section with a pole at the origin and two input and output signals

$$\begin{bmatrix} A_j & B_j \\ C_j & D_j \end{bmatrix} = \left[\begin{array}{c|cc} 0 & 0 & 1 \\ \hline \sqrt{1-\gamma_j^2} & -\gamma_j & 0 \\ \gamma_j & \sqrt{1-\gamma_j^2} & 0 \end{array}\right], \tag{2.15}$$

where $-1 < \gamma_j < 1$. This realization will be used in Subsection 2.6.

2.5 Cascade Structures

We now have a powerful tool to construct simple networks of cascaded all-pass filters, for which the individual state transfer functions form orthogonal basis functions. Let us now consider some special cases.

Finite Impulse Response (FIR) Networks: Consider the tapped delay line network in Figure 2.2. Here $H(z) = z^{-m}$, and $H_i(z) = z^{-1}$. The corresponding orthogonal state-space model is obtained by taking

$$x(t) = [u(t-1)\, u(t-2) \ldots u(t-n)]^T.$$

Consequently

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{aligned}$$

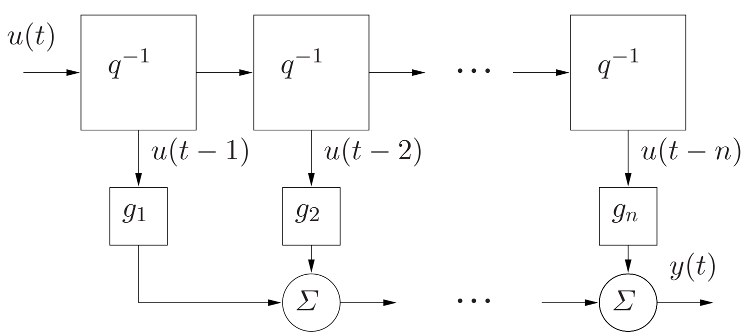

Fig. 2.2. FIR network.

where

$$A = \begin{pmatrix} 0\,0\,\ldots\,0\,0 \\ 1\,0\,\ldots\,0\,0 \\ 0\,1\,\ldots\,0\,0 \\ \vdots \quad \ddots \quad \vdots \\ 0\,0\,\ldots\,1\,0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$C = \begin{pmatrix} 0\,0\,\ldots\,0\,1 \end{pmatrix}, \quad D = 0. \tag{2.16}$$

Laguerre Networks: By taking n identical first-order all-pass filters

$$H_i(z) = \frac{1 - az}{z - a}, \quad -1 < a < 1,$$

using the orthogonal state-space realizations (2.13), the input-to-state transfer functions equal the so-called Laguerre basis functions

$$F_k(z) = \frac{\sqrt{1 - a^2}}{z - a} \left[\frac{1 - az}{z - a} \right]^{k-1}. \tag{2.17}$$

The corresponding Laguerre model has a pole at a with multiplicity n.

Two-parameter Kautz Networks: By using identical second-order all-pass filters,

$$H_i(z) = \frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c}, \quad -1 < b < 1, \; -1 < c < 1. \tag{2.18}$$

with orthogonal state-space realizations (2.14), the so-called two-parameter Kautz basis functions

$$F_{2k-1}(z) = \frac{\sqrt{1 - c^2}(z - b)}{z^2 + b(c-1)z - c} \left[\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right]^{k-1}$$

$$F_{2k}(z) = \frac{\sqrt{1 - c^2}\sqrt{1 - b^2}}{z^2 + b(c-1)z - c} \left[\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right]^{k-1}.$$

$$-1 < b < 1, \quad -1 < c < 1, \quad k = 1, 2 \ldots,$$

are obtained. The corresponding two-parameter Kautz model will have the pole structure $\{\xi_1, \xi_2, \xi_1, \xi_2, \ldots \xi_1, \xi_2\}$, where ξ_1 and ξ_2 can be either real or complex conjugated.

Generalized Orthonormal Basis Functions: The class of generalized orthonormal basis functions (GOBF) is obtained by cascading identical n_b^{th} order all-pass filters,

$$H_i(z) = G_b(z).$$

The corresponding state-space forms of $G_b(z)$ is just restricted to be orthogonal. The corresponding GOBF basis functions can then be written in vector form

$$V_k(z) = (zI - A)^{-1} B [G_b(z)]^{k-1}, \tag{2.19}$$

and thus

$$F_{j+(k-1)nb}(z) = e_j^T (zI - A)^{-1} B [G_b(z)]^{k-1}, \quad j = 1, \ldots nb.$$

Observe that further structure can be imposed by factorizing $G_b(z)$ into first and second order all-pass filters. The vector function

$$V_1(z) = (zI - A)^{-1} B$$

will play an important role in the GOBF analysis to follow, and actually determines the specific structure.
The corresponding GOBF model, illustrated in Figure 2.3 will have the pole structure $\{\xi_1, \ldots \xi_{n_b}, \xi_1 \ldots \xi_{n_b}, \ldots \ldots \xi_1, \ldots \xi_{n_b}\}$.

Takenaka-Malmquist Functions: The general case corresponds to cascading different all-pass blocks of order one or two. These basis functions are sometimes called the Takenaka-Malmquist functions, *cf.* (2.1). The corresponding basis functions for the case with possible complex poles are

$$F_k(z) = \frac{\sqrt{1 - |\xi|_k^2}}{z - \xi_k} \prod_{i=1}^{k-1} \left[\frac{1 - \xi_i^* z}{z - \xi_i} \right]$$

Poles with non-zero imaginary part can be combined in complex conjugated pairs into the corresponding real Kautz form discussed above.

Observe that the ordering of the all-pass filters will influence the basis functions. This freedom can be used to incorporate other consideration, such as compactness. See [26] for a detailed discussion on optimal basis selection.

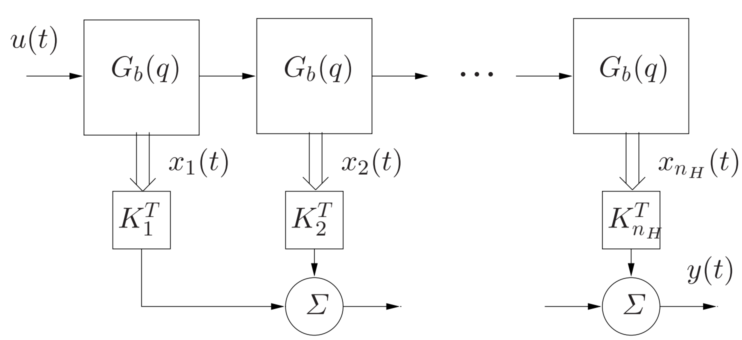

Fig. 2.3. GOBF model. Notice that the gains K_j are $1 \times n_b$ row-vectors.

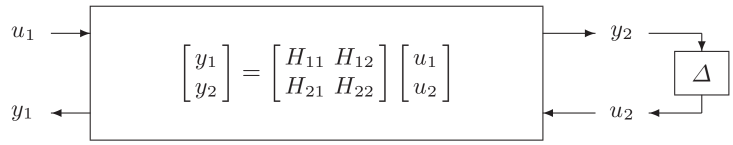

Fig. 2.4. A general feedback connection.

2.6 Ladder Structures

In addition to the cascaded connection of balanced all-pass systems described in the previous subsection, feedback connections of balanced all-pass systems are also important. Let us begin with the feedback connection presented in Figure 2.4. It is quite simple to verify that the transfer function from $u_1(t)$ to $y_1(t)$ is

$$H_{11}(z) + H_{12}(z)\Delta(z)(I - H_{22}(z)\Delta(z))^{-1}H_{21}(z). \tag{2.20}$$

Note that this is the linear fractional transform widely used in robust control theory to describe uncertainties in a system [340]. According to Theorem 2.2, if both

$$H(z) = \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix}$$

and $\Delta(z)$ have orthogonal realizations, then so has the system of Figure 2.4. Consider, in particular, the case $H(z) = H(z,\gamma)$, where

$$H(z,\gamma) = \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} = \begin{bmatrix} \sqrt{1-\gamma^2} & \gamma \\ -\gamma & \sqrt{1-\gamma^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}, \qquad |\gamma| < 1.$$

Note that one possible orthogonal realization of this system is given by (2.15). The block diagram of this all-pass transfer function is depicted in Figure 2.5.

2.6.1 Schur Algorithm

For this $H(z,\gamma)$ the linear fractional transform (2.20) takes the form

$$H_n(z) = -\gamma + \frac{(1-\gamma^2)\Delta\, z^{-1}}{1-\gamma\Delta\, z^{-1}} = \frac{z^{-1}\Delta - \gamma}{1-\gamma\, z^{-1}\Delta}, \tag{2.21}$$

which is a bilinear transformation of $z^{-1}\Delta(z)$. Suppose that $H_n(z)$ in (2.21) is a rational all-pass transfer function of degree n. Using the previous formula we may write, with H_{n-1} instead of Δ,

$$H_n(z) = \frac{H_{n-1}(z) - \gamma_n\, z}{z - \gamma_n\, H_{n-1}(z)} \qquad \text{with} \qquad H_{n-1}(z) = z\,\frac{H_n(z)+\gamma_n}{1+\gamma_n H_n(z)}.$$

Represent $H_n(z)$ in the form

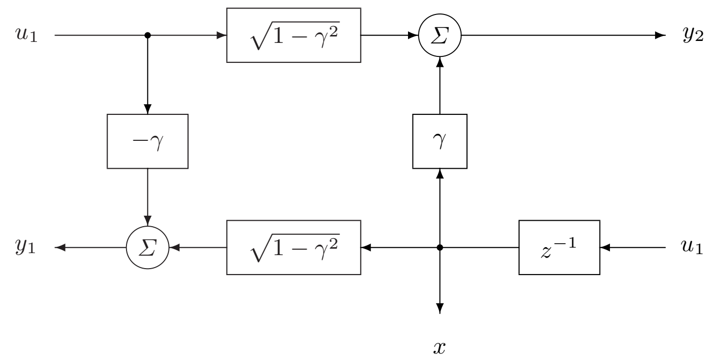

Fig. 2.5. The block diagram of the 2-input 2-output all-pass transfer function $H(z,\gamma)$.

$$H_n(z) = \epsilon D_n(z)/N_n(z),$$

where $N_n(z)$ is the (monic) denominator of $H_n(z)$, $D_n(z) = z^n N_n(1/z)$ is the reciprocal polynomial of $N_n(z)$, and where $|\epsilon| = 1$. It is then possible to verify that $H_{n-1}(z)$ becomes a rational all-pass function of degree less than n if and only if

$$\gamma_n = -\lim_{z\to\infty} H_n(z) = -1/\lim_{z\to 0} H_n(z).$$

For all other values of γ_n, the degree of $H_{n-1}(z)$ remains equal to n. In fact, if $|\gamma_n| \neq 1$, then the degree of $H_{n-1}(z)$ will be exactly $n-1$. This degree reduction step constitutes the first iteration of the famous Schur algorithm [140,272] applied to $H_n(z)$ and is also the first iteration of the well-known Schur-Cohn stability test applied to $N_n(z)$. If the same degree reduction formulas are applied to $H_{n-1}(z)$, $H_{n-2}(z)$, and so on, eventually a constant $H_k(z)$ will be reached. If $H_n(z)$ is asymptotically stable, which, according to Section 3.3, implies that $\big|H_n(z)\big| < 1$ for $|z| > 1$, then $|\gamma_n| < 1$. Because for these values of γ_n the bilinear transform (2.21) maps the exterior of the unit circle into the exterior of the unit circle, *cf.* the Laguerre case of Subsection 3.2, it follows than when $H_{n-1}(z)$ is not a constant, then $\big|H_{n-1}(z)\big| < 1$ for $|z| > 1$; if $H_{n-1}(z)$ is a constant, then $\big|H_{n-1}(z)\big| = 1$ for all z, and no more degree reduction steps can be performed.

Thus, when $H_n(z)$ is asymptotically stable (and all-pass), Schur's iterations can be applied exactly n times, giving rise to asymptotically stable all-pass transfer functions $H_{n-1}(z), \ldots, H_0(z)$, the last of which is constant. Moreover, it also follows that $|\gamma_k| < 1$ for $k = 1, \ldots, n$.
This constitutes the Schur-Cohn stability test. It can be shown that if one of the γ_k has a modulus not strictly less than 1 then $H_n(z)$ is not asymptotically stable.

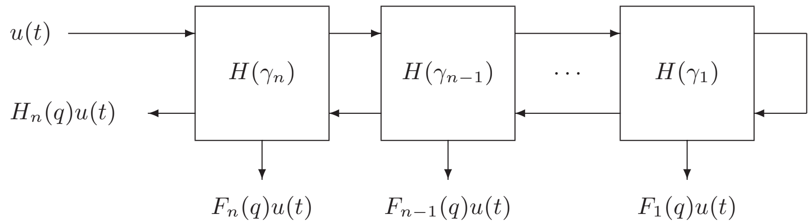

Fig. 2.6. Orthogonal realization of an all-pass scalar transfer function. The $H(\gamma)$ blocks are presented in Figure 2.5. The signals $F_k(q)u(t)$ are the state variables.

2.6.2 Recursion

The degree reduction steps give rise to orthogonal feedback-based realizations for the $H_k(z)$ all-pass transfer functions, as shown in Figure 2.6 for $H_n(z)$. These orthogonal filter structures are precisely the classical Gray-Markel normalized ladder filters [107]. Similar, but more sophisticated orthogonal structures can be found in [76].

The recursion can be applied to calculate the polynomials

$$\begin{bmatrix} N_k(z) \\ D_k(z) \end{bmatrix} = \begin{bmatrix} 1 & -\gamma_k \\ -\gamma_k & 1 \end{bmatrix} \begin{bmatrix} N_{k-1}(z) \\ z\, D_{k-1}(z) \end{bmatrix},$$

with $N_0(k) = D_0(k) = 1$. As mentioned above, the functions

$$F_k(z) = \left[\prod_{l=k}^{n_b} \sqrt{1-\gamma_l^2} \right] \frac{N_{k-1}(z)}{D_{n_b}(z)}$$

are orthogonal and

$$G_b(z) = \frac{N_{n_b}(z)}{D_{n_b}(z)}.$$

Recall that $N_k(z)$ and $D_k(z)$ are reciprocal polynomials, *i.e.* $N_k(z) = z^{-k} D_k(1/z)$. It is worthwhile to note that the all-pass filter used in Laguerre networks corresponds to the case $n = 1$ and $\gamma_1 = a$ and that the all-pass filter used in two-parameter Kautz networks corresponds to the case $n = 2$ and $\gamma_1 = b$ and $\gamma_2 = c$,

$$G_b(z) = \frac{1-\gamma_1 z}{z-\gamma_1},$$

$$G_b(z) = \frac{1-\gamma_1(1-\gamma_2)z-\gamma_2 z^2}{z^2-\gamma_1(1-\gamma_2)z-\gamma_2}.$$

2.6.3 Gray-Markel Normalized Ladder Form

Another way to look at the classical Gray-Markel normalized ladder realization of an all-pass function corresponds to using Givens rotations as elementary factors of the orthogonal system matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

which should have a Hessenberg structure (zeros below the first sub-diagonal). The book [254] discusses how to use this structure as a IIR model structure for system identification and adaptive filtering. It seems that the shift structure obtained by cascading all-pass filters allows a simpler asymptotic theory. Also notice that all ladder orthonormal basis functions are of the same order, while the orders of the basis functions obtained by cascading all-pass filters increases with higher index.

Finally, we would like to mentioned lattice/ladder structures for colored input signals. The corresponding basis function and networks are introduced and used in Chapter 11.

2.7 Reproducing Kernels

The theory of orthogonal functions are closely related to reproducing kernel Hilbert spaces. This section, partly based on [208], gives an overview of how to use reproducing kernels to analyze problems related to orthogonal rational functions.

2.7.1 Derivation

Let X_n denote the space spanned by the functions $\{\bar{F}_1(z)\dots\bar{F}_n(z)\}$, and assume $G(z)\in Y$ where $X_n\subset Y$. Consider the problem of finding the best approximation of $G(z)$ in X_n by solving the optimization problem

$$\min_{\bar{G}\in X_n} ||G-\bar{G}||^2.$$

Because $\bar{G}(z)\in X_n$, we can use the parameterization

$$\bar{G}(z)=\bar{\theta}^T\bar{V}(z), \quad \bar{V}(z)=\left[\bar{F}_1(z)\dots\bar{F}_n(z)\right]^T, \; \bar{\theta}\in\mathbb{R}^n,$$

to obtain

$$||G-\bar{G}||^2=\langle G,G\rangle+\llbracket G,\bar{V}\rrbracket\,\bar{\theta}+\bar{\theta}^T\llbracket\bar{V},G\rrbracket+\bar{\theta}^T\llbracket\bar{V},\bar{V}\rrbracket\,\bar{\theta}. \tag{2.22}$$

This is a quadratic function of $\bar{\theta}$ and it is easy to show that

$$\bar{\theta}_n=\llbracket\bar{V},\bar{V}\rrbracket^{-1}\llbracket G,\bar{V}\rrbracket^T$$

minimizes the cost function (2.22). The corresponding optimal approximation of G in X_n is thus given by

$$G_n(\mu) = [\![G, \bar{V}]\!] \, [\![\bar{V}, \bar{V}]\!]^{-1} \, \bar{V}(\mu)$$

Here we have used the notation subindex n to denote that this is the optimal approximation of $G(z)$ in the n-dimensional subspace X_n. We have also used the argument $z = \mu$ for reasons that soon will become clear.

By defining the so-called reproducing kernel for X_n,

$$K(z, \mu, X_n) = \bar{V}^*(\mu) \, [\![\bar{V}, \bar{V}]\!]^{-1} \, \bar{V}(z),$$

the optimal approximation can be written as

$$G_n(\mu) = \langle G, K(., \mu, X_n) \rangle.$$

Notice that $\langle G, K(., \mu, X_n) \rangle$ is just the orthogonal projection of G on X_n. This means that

$$F(\mu) = \langle F, K(., \mu, X_n) \rangle, \quad \forall F \in X_n,$$

and in particular

$$G_n(\mu) = \langle G_n, K(., \mu, X_n) \rangle = \langle G, K(., \mu, X_n) \rangle.$$

The reproducing kernel for a space is unique and independent of basis. A change of basis $V(z) = T\bar{V}(z)$, for a non-singular square matrix T, gives

$$\begin{aligned} K(z, \mu, X_n) &= \bar{V}^*(\mu) \, [\![\bar{V}, \bar{V}]\!]^{-1} \, \bar{V}(z) \\ &= V^*(\mu)[T^T]^{-1} \left(T^{-1} \, [\![V, V]\!] \, [T^T]^{-1}\right)^{-1} T^{-1} V(z) \\ &= V^*(\mu) \, [\![V, V]\!]^{-1} \, V(z). \end{aligned}$$

2.7.2 Orthonormal Basis

For an orthonormal basis $V(z) = [F_1(z) \dots F_n(z)]^T$ we have $[\![V, V]\!] = I$, and the expression for the reproducing kernel simplifies to

$$K(z, \mu, X_n) = \sum_{k=1}^{n} F_k^*(\mu) F_k(z).$$

The function

$$K(z, z, X_n) = \sum_{k=1}^{n} |F_k(z)|^2 \tag{2.23}$$

plays an important role in calculating certain Toeplitz forms and parameter covariance matrices. First, notice that the normalization gives

$$\frac{1}{2\pi}\int_{\pi}^{\pi} K(e^{i\omega}, e^{i\omega}, X_n)d\omega = n.$$

The matrix

$$\bar{P} = [\![\bar{V}, \bar{V}]\!]$$

can be viewed as a state covariance matrix, see (2.5). We then have the relation

$$\bar{V}^*(z)\bar{P}^{-1}\bar{V}(z) = K(z, z, X_n),$$

where the right-hand side can be calculated using (2.23). For the Takenaka-Malmquist functions (of which FIR, Laguerre and Kautz are special cases), there is an explicit expression of $K(z, z, X_n)$:

$$F_k(z) = \frac{\sqrt{1-|\xi|_k^2}}{z-\xi_k}\prod_{i=1}^{k-1}\left[\frac{1-\xi_i^* z}{z-\xi_i}\right], \quad |\xi_i| < 1, \quad \Rightarrow$$

$$K(z, z, X_n) = \sum_{i=1}^{n}\frac{1-|\xi_i|^2}{|z-\xi_i|^2}. \tag{2.24}$$

2.7.3 Weighted Spaces and Gramians

Weighting can be incorporated in this framework by studying the space X_n^H spanned by

$$\{F_1(z)H(z)\ldots F_n(z)H(z)\}.$$

By defining the Gramian

$$P_H = [\![VH, VH]\!] = \frac{1}{2\pi i}\oint_{\mathbb{T}} V(z)V^T(1/z)\Phi(z)\frac{dz}{z}, \quad \Phi(z) = H(z)H(1/z),$$

the reproducing kernel for X_n^H can be written

$$K_n(z, \mu, X_n^H) = H^*(\mu)V^*(\mu)P_H^{-1}V(z)H(z). \tag{2.25}$$

The right-hand side of this expression has a nice interpretation in terms of Toeplitz covariance matrices. Let $V(z) = [1\ldots z^{-(n-1)}]^T$. The Gramian P_H is then the covariance matrix corresponding to the spectral density $\Phi(z)$, and from (2.25) we have

$$\frac{V^*(z)P_H^{-1}V(z)}{K(z, z, X_n^H)} = \frac{1}{\Phi(z)}, \tag{2.26}$$

which is closely related to (3.20). As mentioned earlier, it is easy to calculate $K(z, z, X_n^H)$ for certain spaces. For example, if $V(z) = [1\ldots z^{-(n-1)}]^T$ and $\Phi(z)$ corresponds to an autoregressive process $H(z) = z^n/A(z)$, where $A(z) = \prod_{i=1}^{n}(z-\xi_i)$, it follows from (2.24) that

$$K(z, z, X_n^H) = \sum_{i=1}^{n}\frac{1-|\xi_i|^2}{|z-\xi_i|^2}. \tag{2.27}$$

This result can be used to give explicit variance expression for estimated models, see [208, 332] and Chapter 5.

Another interesting expression is

$$\begin{aligned}||HK(.,\mu,X_n)||^2 &= \langle HK(.,\mu,X_n), HK(.,\mu,X_n)\rangle \\ &= \frac{1}{2\pi i}\oint_{\mathbb{T}} H(z)\bar{V}^*(\mu)\bar{P}^{-1}\bar{V}(z)\bar{V}^T(1/z)\bar{P}^{-1}\bar{V}(\mu)H(1/z)\frac{dz}{z} \\ &= \bar{V}^*(\mu)\bar{P}^{-1}\bar{P}_H\bar{P}^{-1}\bar{V}(\mu), \end{aligned} \tag{2.28}$$

where

$$\bar{P} = [\![\bar{V},\bar{V}]\!] = \frac{1}{2\pi i}\oint_{\mathbb{T}} \bar{V}(z)\bar{V}^T(1/z)\frac{dz}{z},$$

$$\bar{P}_H = [\![\bar{V}H,\bar{V}H]\!] = \frac{1}{2\pi i}\oint_{\mathbb{T}} \bar{V}(z)\bar{V}^T(1/z)\Phi(z)\frac{dz}{z}.$$

Now

$$\langle HK(.,\mu,X_n), HK_n(.,\mu,X_n)\rangle = \langle \Phi K(.,\mu,X_n), K_n(.,\mu,X_n)\rangle.$$

2.7.4 Convergence

If $\Phi(z)$ would have been an $\mathcal{H}_2(\mathbb{E})$-function and for a complete basis $\{F_k(z)\}_{k=1\ldots\infty}$, we would have that

$$\langle \Phi K_n(.,\mu,X), K_n(.,\mu,X)\rangle \approx \Phi(\mu)K_n(\mu,\mu,X),$$

for large n. Now, $\Phi(z) = H(z)H(1/z)$ has poles outside the unit circle and is not in $\mathcal{H}_2(\mathbb{E})$. However, the result still holds! Define the positive function

$$P(z,\mu,X_n) = \frac{|K(z,\mu,X_n)|^2}{K(\mu,\mu,X_n)},$$

which will behave as a real-valued Dirac function for large n. Also $\Phi(e^{i\omega})$ will be a real valued positive function of ω, and

$$\langle HK(.,\mu,X_n), HK(.,\mu,X_n)\rangle = K(\mu,\mu,X_n)\langle \Phi, P(.,\mu,X_n)\rangle.$$

Now a rather involved analysis, given in [208], proves that

$$\langle \Phi, P(.,\mu,X_n)\rangle \to \Phi(\mu)$$

as the dimension n tends to infinity. This implies

$$\frac{\bar{V}^*(z)\bar{P}^{-1}\bar{P}_H\bar{P}^{-1}\bar{V}(z)}{K(z,z,X_n)} \approx \Phi(z), \quad \text{for large } n.$$

This expression is very useful in calculating the asymptotic variance of model estimates, see Chapter 5.

http://www.springer.com/978-1-85233-956-2

Modelling and Identification with Rational Orthogonal Basis Functions
Heuberger, P.S.C.; van den Hof, P.M.J.; Wahlberg, B. (Eds.)
2005, XXVI, 397 p., Hardcover
ISBN: 978-1-85233-956-2