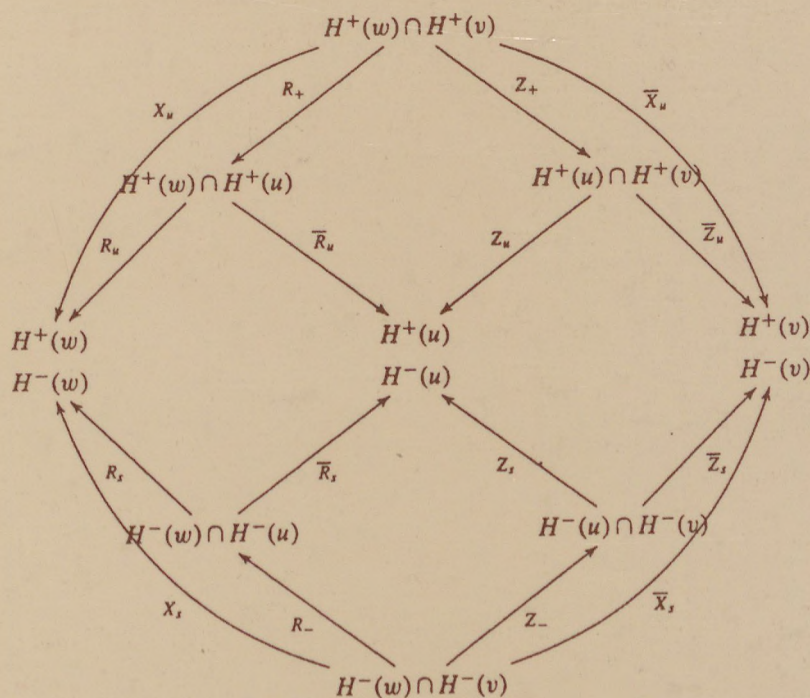


Representability of Stochastic Systems

by

Gy. Michaletzky, J. Bokor, P. Várlaki



Akadémiai Kiadó, Budapest

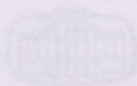
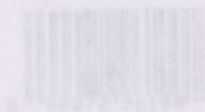
Representability of Stochastic Systems

To

Márta, Ildikó, Kati

Copyright © 1985 by P. Varadi

MTAK



Magyar Tudományos Akadémia, Budapest

182833

130810

Representability of Stochastic Systems

by

Gy. Michaletzky, J. Bokor, P. Várlaki

MTAK



Akadémiai Kiadó, Budapest

018051

Representability of Stochastic Systems

by

Gy. Michaleltzky, J. Bokor, P. Várlaki

MATYÁN
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

ISBN 963 05 7546 9

© Gy. Michaleltzky, J. Bokor, P. Várlaki 1998

Published by Akadémiai Kiadó
H-1519 Budapest, P.O.Box 245.

All rights reserved. No part of this book may be reproduced in a retrieval system,
or transmitted in any form or by any means without the prior permission of the publisher

Printed in Hungary

M. TUD. AKADÉMIAI KÖNYVTÁRA
Könyvtár 6829/19 98 SZ.

Contents

List of Figures	7
Frequently Used Notation	9
1. Introduction	13
1.1. Historical background	13
1.2. The method and structure of the book	19
2. Spectral representation and Wold decomposition of stationary processes	21
2.1. Spectral representation of stationary time series	21
2.2. Spectral density and rank of processes	23
2.3. Regularity of stationary processes. Wold decomposition	28
3. Purely nondeterministic stationary processes with rational spectral density	33
3.1. Parametrization of spectral factors (ARMA, MFD, state-space representations)	34
3.2. Weak state-space realizations	41
3.3. Strong state-space realizations	49
4. System invariants and canonical forms in SSR, ARMA and MFD representations	63
4.1. Observability invariants	64
4.2. Constructibility invariants and canonical state-space representations	65
4.3. Stochastic state-space representations and observability invariants	69
4.4. Reachability invariants and canonical state-space realizations	71
4.5. Controllability invariants and canonical state-space realizations	73
4.6. Stochastic MFD and ARMA representations, transfer relations between the various forms	74
4.7. Correspondences among stochastic state-space, MFD and ARMA representations	86
5. "Geometric" theory of state-space realizations	91
5.1. Splitting subspaces	91

5.2. Minimal (Markovian) splitting subspaces	94
5.3. State-space realizations	97
5.4. Algebraic characterization of various subspaces	100
5.5. Generalized inner product representation	112
5.6. Acausal state-space representations	114
5.7. Zeros of spectral factors	116
6. The structure of all-pass functions	123
6.1. Factorization of inner functions	123
6.2. Characterization and all-pass factorization of all-pass functions	129
6.3. All-pass factorization of all-pass functions	137
7. Generalized Wiener–Hopf factorization of all-pass functions	151
7.1. Generalized Wiener–Hopf factorization of an all-pass function without minimality conditions	152
7.2. Various subspaces connected to an all-pass function	158
7.3. Minimal all-pass functions	163
7.4. The general case – generalized Wiener–Hopf factorization. Transfer function approach	164
7.5. Reconstruction of the output process from input noises. Geometric approach	173
7.6. State-space description of the factorization	188
8. Canonical reconstruction of stationary processes	211
Author Index	217
Subject Index	219
References	223

List of Figures

2.1. The generation of the process $y(t)$ and its spectral density applying the filter with impulse response $\phi(U)$	25
3.1. The flow-diagram of a forward state-space realization	36
3.2. The flow diagram of a backward state-space realization	36
3.3. Forward ARMA and dual forward ARMA representations of the process y	39
3.4. Factorization corresponding to the <i>reachability</i> , <i>observability</i> , <i>controllability</i> and <i>constructibility</i> operators	55
3.5. Factorization corresponding to the <i>reachability</i> , <i>observability</i> , <i>controllability</i> and <i>constructibility</i> operators (backward system)	59
3.6. The filter relating the forward and backward innovations	63
4.1. Transfer relations for a forward spectral factor	90
4.2. Transfer relations for a backward spectral factor	91
4.3. The transfer relations between the various representations	92
6.1. Connections between the various subspaces	143
8.1. Stable, minimum phase (forward innovation) realization	213
8.2. Stable, maximum phase realization	213
8.3. Unstable, maximum phase (backward innovation) realization	214
8.4. Unstable, minimum phase realization	214
8.5. Past and future subspaces generated by the process y and the (forward, backward) innovations	215
8.6. Connections between the various "extremal" generating white noises	217
8.7. Reachability, observability, controllability and constructibility operators in case of "extremal" realizations	218

Frequently Used Notation

$\mathbb{Z}, \mathbb{N}$

the set of integers and non-negative integers, respectively

$y(t), t \in \mathbb{Z}$

vector valued stationary stochastic process with zero mean

$H = H(y) =$

Hilbert space generated by the components of the process y , as a closed subspace of L_2

$= \text{span}\{y(s), s \in \mathbb{Z}\}$

closed subspace of $H(y)$ generated by the past of the process y

$H_t^- = H_t^-(y) =$

$= \text{span}\{y(s), s < t\}$

closed subspace of $H(y)$ generated by the future of the process y

$H_t^+ = H_t^+(y) =$

$= \text{span}\{y(s), s \geq t\}$

forward shift operator

$Uy(t) \triangleq y(t+1)$

backward shift operator

$U^*y(t) \triangleq y(t-1)$

$g(s), s \in \mathbb{Z}$

impulse response sequence of a discrete filter

$g(U^*) = \sum_{s \in \mathbb{Z}} g(s)(U^*)^s$

impulse response operator of a discrete filter

$g(e^{i\lambda}) = \sum_{s \in \mathbb{Z}} g(s)e^{is\lambda}$

frequency response function of a filter

$R(t) \triangleq E[y(s+t)y^T(s)], t \in \mathbb{Z}$

covariance (matrix) sequence of $y(t)$

$f(\lambda) = \sum_{t \in \mathbb{Z}} R(t)e^{i\lambda t}$

spectral density matrix

$f(\lambda) = \frac{1}{2\pi} B(e^{-i\lambda})B^*(e^{-i\lambda})$

spectral factorization

$B(z) = \sum_{k=0}^{\infty} g(k)z^k$

analytic (causal) spectral factor

$B^*(z) = \sum_{k=-\infty}^0 g(k)z^k$

coanalytic (anticausal) spectral factor

$\text{Pr}_{H_t^-} y(t)$

projection of a vector $y(t)$ to the space H_t^-

$\text{Pr}_{H_t^-} H_t^+ \subset H_t^-$

projection of the space H_t^+ to H_t^-

$\mathcal{H}_{n,m}, n, m \in \mathbb{N}$

block-Hankel matrix with n block-rows and m block-columns

$R(k) = CA^{k-1}G, k \geq 1$

matrices in the covariance factorization

A, B, C, D

matrices of a forward state-space representation

$\bar{A} = A^T, \bar{B}, \bar{C} = G^T, \bar{D}$

matrices of a backward state-space representation

$\mathcal{O}(A, C) = [C^T, A^T C^T, \dots]^T$

observability matrix of the (C, A) pair

$\mathcal{C}(A, B) = [B, AB, \dots]$

reachability matrix of the (A, B) pair

$\mathcal{O} = \mathcal{O}(A, C)$

observability matrix of a (forward) state-space representation

$\mathcal{R} = C(A, B)$	reachability matrix of a (forward) state-space representation
$C = C(A^{-1}, A^{-1}B)$	controllability matrix of a (forward) state-space representation
$C_0 = \mathcal{O}(A^{-1}, A^{-1}C)$	constructibility matrix of a (forward) state-space representation
$M(k) = CA^{k-1}B, k \geq 1$	Markov parameters of a (forward) state-space representation
$\{\alpha_{ijk}, v_i\}$	observability invariants
$\{\beta_{ijk}, \mu_i\}$	constructibility invariants
$\{\gamma_{ijk}, \lambda_i\}$	reachability invariants
$\{\delta_{ijk}, \kappa_i\}$	controllability invariants
$A(U^*), B(U^*)$	forward AR and MA operators (polynomial matrices of the shift operator) of a FARMA form (left FARMA)
$\bar{A}(U^*), \bar{B}(U^*)$	backward AR and MA operators (polynomial matrices of the shift operator) of a BARMA form (left BARMA)
$P(U), Q(U)$	denominator and numerator operators of a FMFD form (left FMFD)
$\bar{P}(U), \bar{Q}(U)$	denominator and numerator operators of a BMFD form (left BMFD)
$P_r(U), Q_r(U)$	denominator and numerator operators of a dFMFD form (right FMFD)
$\bar{P}_r(U), \bar{Q}_r(U)$	denominator and numerator operators of a dBMFD form (right BMFD)
$A_r(U^*), B_r(U^*)$	forward AR and MA operators (polynomial matrices of the shift operator) of a dFARMA form (right FARMA)
$\bar{A}_r(U^*), \bar{B}_r(U^*)$	backward AR and MA operators (polynomial matrices of the shift operator) of a dBARMA form (right BARMA)
$H_t^-(y) \ominus H_{t-1}^-(y)$	forward innovation subspace
$H_{t+1}^+(y) \ominus H_t^+(y)$	backward innovation subspace
$w(t)$	forward generating noise process
$\bar{w}(t)$	backward generating noise process
$w_-(t)$	forward innovation process with covariance P_-
$w_+(t)$	forward generating noise process with covariance P^+
$\bar{w}_-(t)$	backward generating noise process with covariance matrix $\bar{P}_-$

$\bar{w}_+(t)$	backward innovation process with covariance matrix $\bar{P}_+$
$x(t), \bar{x}(t)$	state vectors of a forward and backward state-space representation, respectively
$x_-(0), x_+(0)$	state vectors of a forward minimal and forward maximal state-space representation
$\bar{x}_-(0), \bar{x}_+(0)$	backward maximal and backward minimal state-space representation
$\xi(t)$	partial state vector of a dual (right) ARMA or MFD form
$P \triangleq E[x(t)x^T(t)], \forall t$	forward state covariance matrix
$\bar{P} \triangleq E[\bar{x}(t)\bar{x}^T(t)], \forall t$	backward state covariance matrix
$P_- \triangleq E[x_-(t)x_-^T(t)], \forall t$	minimal forward state covariance matrix
$P_+ \triangleq E[x_+(t)x_+^T(t)], \forall t$	maximal forward state covariance matrix
$\bar{P}_- \triangleq E[\bar{x}_-(t)\bar{x}_-^T(t)], \forall t$	maximal backward state covariance matrix
$\bar{P}_+ \triangleq E[\bar{x}_+(t)\bar{x}_+^T(t)], \forall t$	minimal backward state covariance matrix
$\Phi(z)$	phase function
$W_-(z), W_+(z)$	stable minimum phase and stable maximum phase spectral factor
$\bar{W}_-(z), \bar{W}_+(z)$	antistable minimum phase and antistable maximum phase spectral factor
$(S, \bar{S})$	(invariant) subspaces generating Markovian representations
$S \perp \bar{S} X$	the subspaces S and $\bar{S}$ are conditionally orthogonal with respect to the subspace X
X	splitting subspace, state-space of a Markovian realization
$X \cap (H^+)^{\perp}$	the unobservable subspace of X
$X \cap (H^-)^{\perp}$	the unconstructible subspace of X
$N^+ = \bar{S} \cap (H^-)^{\perp}$	unpredictable part of $\bar{S}$
$N^- = S \cap (H^+)^{\perp}$	part of S which cannot be (backward) predicted from H^+
$D + C(z^{-1}I - A)^{-1}B$	rational forward spectral factor
$\bar{D} + \bar{C}(zI - \bar{A})^{-1}\bar{B}$	rational backward spectral factor
$S \ominus H^-$	the subspace corresponding to the left zeros with modulus larger than 1 of a forward realization

$\bar{S} \ominus H^+$

the subspace corresponding to the left zeros with modulus less than 1 of a forward realization

 $\Pr_{\bar{S}^\perp} (H^+)^\perp$

the subspace corresponding to the right zeros with modulus larger than one of a forward realization

Abbreviations

SSR	State-Space Representation
FSSR	Forward SSR
BSSR	Backward SSR
dFSSR	dual-Forward SSR
dBSSR	dual-Backward SSR
ARMA	AutoRegressive Moving Average representation
FARMA	Forward ARMA (left FARMA) representation
BARMA	Backward ARMA (left BARMA) representation
dFARMA	dual-Forward ARMA (right FARMA) representation
dBARMA	dual-Backward ARMA (right BARMA) representation
MFD	Matrix Fraction Description form
FMFD	Forward MFD (left FMFD) form
BMFD	Backward MFD (left BMFD) form
dFMFD	dual-Forward MFD (right FMFD) form
dBMFD	dual-Backward MFD (right BMFD) form

1. Introduction

1.1. Historical background

Research in the possible representations of stochastic systems and processes developed during the past more than 70 years in a relatively spontaneous way, based on different mathematical and conceptual approaches, almost in an “evolutionary” manner.

Considering the development of spontaneous feature, we may state that, besides statistical time-series analysis in economics as well as in other fields, the probability theoretical approach invented by *A. Kolmogorov* and *N. Wiener* which evolved parallel at the start, after World War II became bifurcated into two descriptive approaches. These two directions are the general probability theory of stochastic processes, and the theory of stochastic dynamical systems with noisily measured output signals and possible stochastic and deterministic input excitations.

The theory of stochastic linear systems, introduced primarily by *N. Wiener* and playing an ever growing role in signals, systems and control, had been, naturally, closely related to the results in the field of deterministic systems described by transfer functions (where major development can be traced during the 50s), which goes back to analysis of dynamical systems using complex function analysis elaborated by *Bode*, *Nyquist* et al. in course of the 30s, mostly using the feedback loop transfer function characteristics of the given control system. (In Appendix B and C of *Wiener's* famous book “Extrapolation, interpolation and smoothing of stationary time series”, *N. Levinson* gave heuristically a preliminary concept of stochastic systems, as well as a possible approach to system identification.)

The deterministic and stochastic linear dynamical system concepts developed this way jointly, were revolutionized by the new *environment and system* concept of *Kalman*, and the filtering theory of *Kalman–Bucy*, later by the Kalmanian realization theory, leading to the concept of state-space and system representation in state equation forms.

As worded by *J. C. Willems* (Willems, 1991):

The development of the field of System Theory as a scientific discipline owes more to the vision and to the research work of Kalman than to that of any other individual.

We agree with *Willems* stating that the two most important concepts of Dynamical Systems are *controllability* and *observability*, i.e.

[...] the important notions of controllability and observability. Controllability refers to the possibility of matching any feasible past trajectory to any feasible future trajectory, while observability refers to the possibility of deducing the latent variable trajectory from the manifest variable trajectory.

These two concepts provided a theoretical basis for the investigation of the various forms of the possible representations of stochastic systems, as well.

Specifically, the discussion of representations in the statistical time-series theory is done through a description of typically discrete ARMA (AutoRegressive Moving Average) type representations, while in the theory of dynamical systems in control, representations of continuous matrix fraction description type are dominated.

The ARMA representation approach was elaborated mainly by British-American statistical time-series schools, and was summarized in the famous and very influential book "Time series analysis, forecasting and control" by Box and Jenkins, see (Box and Jenkins, 1970).

The second one, i.e. the process conceptual was "cultivated" in first line by Soviet-Russian mathematicians and the most important results was summarized and further developed by Rozanov in his excellent book "Stationary stochastic processes" (Rozanov, 1990).

In the 60–70s era of control theory called "modern control", however, the state space description got increasingly into the focus of analysis, as sometimes opposed to the use of discrete transfer function representation. This modelling and representation approach has been having from the "beginnings" a very international (almost universal) character.

In the "classical-modern" era of control and system theory we experienced the reunification of the use of state-space and rational (matrix) function representations in a synergic way. This can be observed today in the "post modern" era, where most recent problems (e.g. in robust filtering and control) can be formulated or solved using either state-space or matrix rational function (Hardy-space) formalism, as well.

It is common in all the three theories that the measurable output processes are – explicitly or implicitly – derived from a white noise process, see (Wold, 1938).

The development of the realization and "representability" theory is particularly accelerated by the unique fast development of the theory of signals and systems and the adaptation of the theory during the past three decades, inspired greatly by the new problems in the field of system and control science. Significant results are originated also in the domain of the algebraic and topologic structure of linear multivariable systems, successfully applied then in solving the problems of control theory and practice, system identification or signal processing and lately in change and fault detection.

This really extremely fast theoretical development resulted in the significant converging of the previously developed threefold approach (linear systems, stochastic

processes, and the theory of time-series), as a result of which the statistical theory of linear systems emerged in parallel, the realization theory of multivariable stochastic systems was tried to be enframed first by two monographs, by *Hannan* and *Deistler* (Hannan and Deistler, 1988) and *Caines* (Caines, 1988).

The relationships between the internal type state space models describing the signals and systems and the realization and structural features of the input/output (i.e. external) type Matrix Fraction and ARMA models, as well as the equivalent system descriptions are studied very intensively. These results are related mainly to the realization theory of ARMA models playing an important role in applications, and the deduction of conversion (transfer) relations between the equivalent realizations, in any forms discussed above.

It is known that finite-dimensional linear systems have several equivalent representations. For the sake of brevity, we refer to all these with the abbreviations common in the international literature:

- State Space Representation = SSR
- Matrix Fraction Description = MFD
- AutoRegressive Moving Average = ARMA

In the solution of problems in control theory, system identification, signal processing etc., there exists basically an adequate or favoured representation that fits mostly into the specifics of the task. For this reason, the results showing the conditions and system-features under which one representation can be transformed into another one, which is more advantageous and even necessary to achieve a satisfactory solution of a specific problem, are of importance.

The realization theory of stochastic processes and systems were dealt with *Akaike* (Akaike, 1974a; Akaike, 1976; Akaike, 1977), *Faurre* (Faurre, 1976), *Lindquist* and *Picci* (Lindquist and Picci, 1977; Lindquist and Picci, 1979; Lindquist and Picci, 1985b; Lindquist and Picci, 1990), *Anderson* and *Kailath* (Anderson and Kailath, 1979), *Lindquist* and *Pavon* (Lindquist and Pavon, 1984), *Michaletzky* and *Tusnády* (Michaletzky and Tusnády, 1991), and the summary of the early results can be found in the 4th Chapter in *Caines* (Caines, 1988). The results of system and realization-theory were summarized by *Willems* (Willems, 1986a; Willems, 1986b; Willems, 1987), and later (Willems, 1991), in a rather modern and abstract way.

At the earliest the statistical literature dealt with ARMA representations, mostly from estimation-theory point of view. Their applications in control-theory can be found in *Åström* (Åström, 1970) first, and the same in system-identification in *Eykhoff* (Eykhoff, 1974), *Box* and *Jenkins* (Box and Jenkins, 1970), *Ljung* and *Söderström* (Ljung and Söderström, 1983) and *Ljung* (Ljung, 1987).

Literature in the 70s dealt primarily with the single-variable case (scalar process). The structure of the ARMA models of multivariable (vector) processes, their parameter estimation were studied by *Hannan* (Hannan, 1976), *Dunsmuir and Hannan* (Dunsmuir and Hannan, 1976), *Hannan and Kavalieris* (Hannan and Kavalieris, 1984), *Clark* (Clark, 1976), *Ljung and Caines* (Ljung and Caines, 1979), *Deistler and Hannan* (Deistler and Hannan, 1981), *Gevers and Wertz* (Gevers and Wertz, 1984; Gevers and Wertz, 1987).

Naturally, the literature enumerated is not but a small fragment of the literature on the field. The special emphasis on them is due to their outstanding contribution, monographic character or the direct relation with the subject of the book.

By the early 80s the topologic and algebraic structures of finite dimensional irreducible state space and matrix fraction representations, furthermore, the correspondence between their canonical (controllability, observability) forms became widely known. For this reason, it is rather surprising that the connections between the (deterministic and stochastic) ARMA and the SSR, furthermore, the MFD representations was studied first in the 80s only, see *Bokor and Keviczky* (Bokor and Keviczky, 1982), *Wolovich and Elliott* (Wolovich and Elliott, 1983), *Gevers* (Gevers, 1986), *Janssen* (Janssen, 1987; Janssen, 1988c; Janssen, 1988b), *Deistler and Gevers* (Deistler and Gevers, 1989), *van den Hof* (van den Hof, 1989).

Janssen (Janssen, 1988a), following the principle elaborated by *Forney* (Forney, 1975) for MFDs, derived the ARMA representations directly from a minimal polynomial base of the right-hand side null-space of the system's extended (enlarged) transfer matrix.

ARMA models derived in such a way had several unfavourable features summarized by *Gevers* (Gevers, 1986). That is to say, in many system identification and control tasks the use of *unique* (canonical) ARMA representations is favourable, which feature does not hold for these models.

The introduction of canonical state space representations defined by the constructibility invariants, *Bokor and Keviczky* (Bokor and Keviczky, 1987), was an essential further innovation.

By their help, it was proved that in case the system has no pole in the origin, a canonical ARMA model can always be given, which is uniquely determined by the constructibility invariants.

The extension of these results towards to the ARMA description of general stochastic processes and systems needed some further consideration, due to the fact that stochastic processes can be described with forward and backward state-space representations, as well. It can be demonstrated that the canonical form of the backward ARMA representation of a given process is uniquely determined by the backward SSR observability invariants, while in forward ARMA case, by the forward SSR constructibility invariants, see *Bokor and Keviczky* (Bokor and Keviczky, 1987), *Bokor, Keviczky and Bányász* (Bokor et al., 1989b; Bokor

et al., 1989a). Naturally, the issue of the connections between various equivalent representations rises repeatedly, see also *van den Hof* (van den Hof, 1989).

The first theoretical analysis regarding the "geometric" construction of the forward and backward state-space representations is due primarily to *Lindquist* and *Picci* (Lindquist and Picci, 1977; Lindquist and Picci, 1979; Lindquist and Picci, 1985a; Lindquist and Picci, 1985b; Lindquist and Picci, 1990), in the discrete time case *Lindquist* and *Pavon* (Lindquist and Pavon, 1984), as well as was developed later by *Michaletsky*, see (Lindquist et al., 1995; Lindquist and Michaletsky, 1997).

On the basis of these extended, as well as rather involved methods and approaches, on the present stage, one may attempt to describe the complete "meta-system" of the possible representations of stochastic systems in the *narrow* and *broader* sense.

A representational metasystem in the *narrow sense* is meant by the most complete possible description of the possible (external) input/output (matrix fraction and autoregressive moving average) and (internal) state-space representations uniquely definable.

The description of a representational metasystem in the *broader sense* is meant by the unique definition of all the possible representations with vector white noise pairs (or, in an even broader sense, with vector white noise pair-groups), where each element (i.e. a basic representation of the representational metasystem) taken in a narrow sense can be uniquely defined from the given vector white noise pairs or the vector white noise pair-groups. At the same time, the output processes can be reconstructed (as to their statistical features) or, using our terminology, in fact re-presented from the vector white noise pairs.

The concepts introduced heuristically in the above can be organically considered the Representability Theory of the Stationary Stochastic Systems, due to their close relation to the mathematical system theory principles. The above concepts are as follows:

- the representability, i.e. the description of the representations of all possible input/output and state space of linear stationary stochastic systems; and their
- re-presentability, i.e. the re-representation or reconstruction of the entire statistics of the output process, knowing only the vector white noise processes related to the given representations.

Now let us characterize shortly and draftly the essential features of this possible representability theory of stationary stochastic systems.

We start from the natural approach, namely while observing the output of the stationary stochastic systems, we intend to explain the change in the specific and actual realization of the stochastic output signals, processes, based on some generating mechanism. Then we may apply the concept that the primary statistical

characteristics of the stationary output processes (in the broader sense) is given by the output covariance-function. Applying the filtering theory – i.e. generating every process from white noise – the general task in this case is to find the generating mechanism which results in the given or the wanted covariance statistics and output realization starting from white noise, based on some natural and general criteria.

In the case of multivariable systems, obviously, further problems emerge. The realization of multivariable systems, in the case of stochastic processes of rational spectrum, demands some inner description of the system. Then, we intend to find the most concise parametrization reflecting the basic system features belonging to the state space of the finite dimensional systems' equivalence classes. These will lead to the parametrization using system invariants which were introduced based on the four *Kalman* criteria, i.e. controllability, observability, reachability and constructibility. In harmony with the principle of state-space descriptions, the projections of the future of the output process to its past or to the past of an unmeasurable white noise input or projecting the past to the future space provide a method of determining the various system invariants.

It is obvious that the state-space (internal) and input/output (external) descriptions will be given with the same invariants. Thus these descriptions are only the different outer forms of appearance of the essential system features.

Besides the representability theory of state space (internal) and the corresponding input/output (external) models, the further representability problem deriving from the white noise concept is how it is possible to generate the equivalent system description of the system invariants, while the white noise is not fixed. From mathematical point of view, this means the revealing of stochastic relations between certain white noise "representations", types (e.g. forward, backward vector white noises), leading to the analysis of the all-pass transfer functions. In other words, besides the analysis of representability and the revealing of state space and input/output representations, the relations between those white noise processes which can "drive" them will be analysed.

Obviously, these two tasks closely "correlate", since the white noise processes from the given output can be identified only by the help of the given state space and input/output representations. On the other hand, the relationships between the white noise processes, i.e. the phase and the all-pass functions describing this connection (both from structural and parametrical point of view) contain implicitly, in principle, the complete representation system, together with the ability of the re-presentability of the output processes from these white noise sources.

The book is aimed to at least partially give solution and explanations to these major and not clear-cut tasks, pointing out to the fact that we intend to represent system features, naturally unrecognizable "in themselves" because of the epistemological limit, but only through abstract notions. After all, the methodology of the book is determined by the preciseness and correctness of these abstract notions, and

the as much as possible definite provision of the relations between them – regarded from the above double aspect of representability.

It can be seen that under such sophisticated modelling circumstances only these abstract concepts and paradigmatic frame can give the inevitable base of applying the empirical system identification procedures and stochastic control methods in the most different practical fields of modern science and technology.

1.2. The method and structure of the book

Concretely, on the basis of the above preliminaries and aims we would like to give an introduction to the realization theory of stationary time series with rational spectral density in discrete time case with the main emphasis on the connection between the input white noises and the output signal. We also analyze the problem of the re-presentability of the original time series from the connections of generating white noises. Our analysis forges together the *Faurre*-type approach leading to weak realizations, the methods based on spectral factors advocated for a long time by *Kučera* and geometric methods initiated by *Wonham* in deterministic setting and developed in the works of *Lindquist* and *Picci* for the stochastic case. We also compare the state space equation, ARMA and MFD descriptions mainly using the ideas of *Bokor* and *Keviczky*. Although some parts of this book are well-known, we wanted to give a comprehensive introduction of this important area.

In Chapter 2 we study the spectral representation of weakly stationary stochastic processes using the spectral decomposition of a unitary operator on *Hilbert*-spaces and the *Wold*-decomposition. We also discuss the *Kronecker*-theorem assuring the rationality of the spectral density.

In Chapter 3 we define the various representations of stochastic systems. These include the state-space representations, ARMA and MFD representations with their dual forms. Weak and strong state-space realizations of stationary processes are analysed, as well. We show how to construct the forward (backward) innovation realization as solutions of a *Riesz*-representation problem. The connections between various minimal state-space realizations are also discussed.

Chapter 4 compares the various (canonical) parametrization possibilities of a rational (analytic or coanalytic, causal or anticausal) spectral factor using system invariants. It also shows how to transform the forward (backward) state-space representations into forward (backward) ARMA and MFD forms. The problem of controllability, observability, constructibility and reachability, and their relationships to the internal and external model types are also discussed in detail.

The splitting subspaces defining the state-spaces of minimal Markovian realizations are discussed in Chapter 5. We show that they can be used for the analysis of acausal spectral factors, when there is no specification on the pole structure of the corresponding transfer function. An immediate consequence of this fact is that

different spectral factors may share the same state-space. The zero structure of the spectral factors is also mentioned proving that it depends only on the state-space.

Chapter 6 deals with the characterization and factorizations of inner functions. We give a short “geometric” proof for the theorem proved originally by *Fuhrmann* and *Hoffmann* (Fuhrmann and Hoffmann, 1998) characterizing the inner factors in terms of solutions of a finite dimensional matrix equation (in some cases this equation can be transformed into a *Riccati*-equation) and enabling us to generalize it to a statement concerning the ordering of these inner factors. These results – using the geometric argument – are also extended to the all-pass case. We show that the minimal all-pass factors of an all-pass function are characterized by the solutions of two simultaneous algebraic Riccati equations. Nonminimal realizations of these all-pass factors in terms of projections are constructed, as well.

The generalized *Wiener–Hopf* factorizations of a given all-pass function are considered in Chapter 7. This problem is connected with the reconstructibility question of the output process given a forward and a backward white noise in its state-space realizations. Especially, we show that if the all-pass function is minimal then the only factorization is determined by the forward and backward innovations of an output process.

Chapter 8 shortly summarizes the possible *canonical* re-presentations of stationary stochastic processes. These canonical cases are obtained via using the extremal possibilities of the generating (input) white noises.

2. Spectral representation and Wold decomposition of stationary processes

In this chapter first we shortly outline the spectral representation of weakly stationary processes in discrete time using the spectral representation of the unitary operator determined by the “time shift”. Then we introduce the notion of spectral density and define the notion of full rank stationary processes. The connection between the dimension of the innovation subspace defined in the *Wold* decomposition and the rank of the process is analysed.

Let $y(t)$, $t \in \mathbb{Z}$ be a p -dimensional vector-valued stationary time series with zero mean. Denote

$$R(t) = E[y(s+t)y(s)^T], \quad t \in \mathbb{Z}, \quad (2.1)$$

the covariance sequence of y . Although we shall investigate only second order properties of $y(t)$, $t \in \mathbb{Z}$, in order to facilitate the language we assume that the finite dimensional distributions of this time series are Gaussian ones.

It is well known that the sequence $R(t)$, $t \in \mathbb{Z}$, is positive semidefinite, i.e.

$$\sum z_j^* R(t_i - t_j) z_i \geq 0,$$

for all $t_1, \dots, t_n \in \mathbb{Z}$, $z_1, \dots, z_n \in \mathbb{C}^p$.

2.1. Spectral representation of stationary time series

An important notion is the spectral representation of y , which means that $y(t)$ can be written as

$$y(t) = \int_{-\pi}^{\pi} e^{it\lambda} \Phi(d\lambda), \quad (2.2)$$

where Φ is a stochastic spectral measure. There are several nice constructions leading to the definition of Φ . In our opinion the most elegant although not the most elementary one is based on the *Stone* representation theorem (see *Rozanov* (Rozanov, 1990)). This fits well to the geometric constructions we shall use later.

For the vector process y , denote $y_k(t)$, $k = 1, \dots, p$ ($t \in \mathbb{Z}$) the coordinate random variables. Define $H(y)$ as the closed subspace of $L_2(P)$ (i.e. a subspace

of the scalar valued random variables which are square integrable with respect to the probability measure denoted by P generated by $\{y_k(t), k = 1, \dots, p, t \in \mathbb{Z}\}^1$. Similarly

$$\begin{aligned} H_t^-(y) &= \langle y_k(s), k = 1, \dots, p, s < t \rangle \\ H_t^+(y) &= \langle y_k(s), k = 1, \dots, p, s \geq t \rangle, \end{aligned} \quad (2.3)$$

where $\langle \cdot \rangle$ denotes the generated closed subspace. Denote shortly

$$H^+ = H_0^+(y), \quad H^- = H_0^-(y), \quad H_0 = H(y). \quad (2.4)$$

Observe that the definitions of the $-$ and $+$ subspaces are not symmetric. This is a speciality of the discrete time case.

We also remark here that we shall apply these notations for arbitrary stationary time series not only for the process y .

There is a unitary operator U defined on the Hilbert space H_0 for which

$$U y_k(t) = y_k(t+1), \quad k = 1, \dots, p, t \in \mathbb{Z}. \quad (2.5)$$

The previous equation defines an isometric operator on finite dimensional subspaces of H_0 which can be extended isometrically to the closure of these subspaces. This extension is unitary. We shall call U as the shift operator, its inverse is the so-called backward shift operator. Since U is unitary it has a spectral representation

$$U = \int_{-\pi}^{\pi} e^{i\lambda} \Pi(d\lambda), \quad (2.6)$$

where $\Pi(d\lambda)$ is a projector-measure, i.e. for any Borel-measurable subset $\Delta \subset [-\pi, \pi]$, the operator $\Pi(\Delta)$ is a projection depending additively on Δ , and

$$\Pi(\Delta)\Pi(\Delta') = 0, \quad \Pi(\Delta \cup \Delta') = \Pi(\Delta) + \Pi(\Delta'),$$

if $\Delta \cap \Delta' = \emptyset$. (I.e. for disjoint subsets of $[-\pi, \pi]$ the ranges of the corresponding projections are orthogonal.) We obtain that

$$U^t = \int_{-\pi}^{\pi} e^{it\lambda} \Pi(d\lambda). \quad (2.7)$$

Consequently

$$y_k(t) = U^t y_k(0) = \int_{-\pi}^{\pi} e^{it\lambda} \Pi(d\lambda) y_k(0). \quad (2.8)$$

¹ We might as well define the Hilbert space $H(y)$ corresponding to the k -dimensional stationary process y as a subspace of $L_2^k(P)$, this latter is the space of $\mathbb{R}^k$ -valued random variables with finite covariance matrix.

Applying the projections $\Pi(\Delta)$ to the random variables $y_k(0)$, $k = 1, \dots, p$, we obtain a so called stochastic spectral measure:

$$\Phi_k(\Delta) = \Pi(\Delta)y_k(0),$$

or shortly the column vector

$$\Phi(\Delta) = (\Phi_k(\Delta), k = 1, \dots, p),$$

so

$$y(t) = \int_{-\pi}^{\pi} e^{it\lambda} \Phi(d\lambda). \quad (2.9)$$

Introducing the matrix-valued set function

$$F(\Delta) = E \left(\left[\Pi(\Delta)y(0) \right] \left[\Pi(\Delta)y(0) \right]^T \right),$$

we can conclude that

$$R(t) = E[y(s+t)y(s)^T] = \int_{-\pi}^{\pi} e^{it\lambda} F(d\lambda). \quad (2.10)$$

Obviously the matrix $F(\Delta)$ is positive semidefinite for every $\Delta \subset [-\pi, \pi]$. If all components of $F(\Delta)$ are absolutely continuous then

$$f(\lambda) = \frac{F(d\lambda)}{d\lambda} \quad (2.11)$$

is the so called spectral density. It is a matrix-valued function. For any $\lambda \in [-\pi, \pi]$, $f(\lambda)$ can be defined to be positive semidefinite.

2.2. Spectral density and rank of processes

In order to make the picture more complete we list several definitions and theorems without proofs in connection with the spectral density, cf. (Rozanov, 1990).

DEFINITION 2.1. *The p -dimensional process y is of rank q , if it has a spectral density $f(\lambda)$, $-\pi \leq \lambda \leq \pi$, such that*

$$\text{rank } f(\lambda) = q, \quad (2.12)$$

for almost every λ .

DEFINITION 2.2. The p -dimensional process y is of full-rank, if it is of rank $q = p$.

THEOREM 2.3. If the p -dimensional stationary process y is of rank q , then it has the representation

$$y(t) = \sum_{s=-\infty}^{+\infty} g(t-s)w(s) = \sum_{s=-\infty}^{+\infty} g(s)w(t-s), \quad (2.13)$$

where $w(s)$, $s \in \mathbb{Z}$ is a q -dimensional stationary process with zero mean and with components in $H_0 = H(y)$. Moreover

$$\text{cov}(w(t), w(s)) = I_q \delta_{t,s}. \quad (2.14)$$

Also

$$f(\lambda) = \frac{1}{2\pi} \varphi(\lambda) \varphi^*(\lambda), \quad (2.15)$$

where

$$\varphi(\lambda) = \sum_{s=-\infty}^{+\infty} g(s) e^{-is\lambda} \quad (2.16)$$

and

$$\varphi^*(\lambda) = \overline{\varphi(\bar{\lambda}^{-1})}^T.$$

In this case

$$R(t) = \int_{-\pi}^{\pi} e^{it\lambda} f(\lambda) d\lambda. \quad (2.17)$$

Observe that $\varphi(-\lambda)$ is the discrete Fourier transform of the sequence $g(s)$, $s \in \mathbb{Z}$.

DEFINITION 2.4. The q -dimensional stationary process w with zero mean is called white noise process, if it satisfies equation 2.14.

An important subclass of the stationary processes is the set of stationary processes with rational spectral density, i.e. when the component functions of the matrix-valued spectral density $f(\lambda)$, $-\pi \leq \lambda \leq \pi$ are rational functions of $e^{-i\lambda}$. The spectral analysis of these processes can be based on the following lemma.

LEMMA 2.5. Every (scalar-valued) nonnegative function $f(\lambda)$, which is rational in $e^{-i\lambda}$, can be written in the form

$$f(\lambda) = \frac{|Q(e^{-i\lambda})|^2}{|P(e^{-i\lambda})|^2}, \quad (2.18)$$

where – with the substitution $z = e^{-i\lambda}$ –

$$P(z) = \sum_{k=0}^n p_k z^k, \quad Q(z) = \sum_{k=0}^m q_k z^k$$

are polynomials of z without zeros inside the complex unit circle.

Moreover, if $f(\lambda) = f(-\lambda)$, then the coefficients of P , Q can be chosen real numbers.

REMARK 2.6. Equation 2.13 can be interpreted (see e.g. the control engineering literature) such that the process $y(t)$ is generated as the output of a linear filter driven by the white noise process $w(t)$, see Figure 2.1. The filter “impulse response” function is given by

$$\phi(U) = \sum_{s=-\infty}^{+\infty} g(s) U^{-s} \quad (2.19)$$

and the filter “transfer function” is given as

$$\phi(z) = \sum_{s=-\infty}^{+\infty} g(s) z^s. \quad (2.20)$$

The function $\phi(e^{-i\lambda})$ (i.e. replacing $z = e^{-i\lambda}$) is called usually the filter frequency function. The relation between the spectral densities of driving and that of output processes (called Wiener–Hintchin formulae) can be illustrated as shown on Figure 2.1.

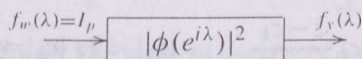
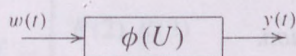


Figure 2.1. The generation of the process $y(t)$ and its spectral density applying the filter with impulse response $\phi(U)$

In the multidimensional case, when f is a $p \times p$ matrix-valued function which is rational in $e^{-i\lambda}$ and positive semidefinite for a.e. λ the representation 2.18 has the form

$$f(\lambda) = \frac{1}{2\pi} B(e^{-i\lambda}) B^*(e^{-i\lambda}), \quad (2.21)$$

where the elements of B are rational in $z = e^{-i\lambda}$ and analytic inside the unit circle. An important observation is that the equation 2.21 does not define the function B

in a unique way, because for any rational function $U(z)$ which is analytic inside the unit circle and for which

$$U(e^{-i\lambda})U^*(e^{-i\lambda}) = I \quad a.e. \quad (2.22)$$

holds we get the factorization

$$f(\lambda) = \frac{1}{2\pi} B(e^{-i\lambda}) U(e^{-i\lambda}) [B(e^{-i\lambda}) U(e^{-i\lambda})]^*.$$

(Let us remark that in the system engineering literature it is more common to use the substitution $z = e^{i\lambda}$ instead of $z = e^{-i\lambda}$ applied here. Accordingly, in that literature the terminology *analytic function* – without explicitly specifying the domain – refers to functions being analytic out of the unit circle. The convention applied here is in coherence with the usual terminology used in complex function theory. Thus, using the phrase *analytic function* – without explicitly defining the domain – we refer to functions which are analytic inside the unit circle. Furthermore, *coanalytic* refers to functions which are analytic out of the unit circle. Because of this unfortunate circumstance we usually define explicitly the domain where the function is analytic. Moreover, sometimes we use the expression *causal function* as a synonym to analytic functions, and *anticausal function* as a synonym to coanalytic functions.)

DEFINITION 2.7. *The $q \times q$ rational matrix function $U(z)$, $z \in \mathbb{C}$, which is analytic inside the unit circle and satisfying the equation 2.22, i.e.*

$$U(e^{-i\lambda})U^*(e^{-i\lambda}) = I \quad a.e.,$$

is called rational inner function.

In the one-dimensional case the general form of a rational inner function is

$$\prod_{k=1}^l c_k \frac{z - z_k}{1 - \bar{z}_k z}, \quad |z_k| < 1, |c_k| = 1,$$

the so-called finite *Blaschke-product*. Observe, that this function has zeros at z_k , $k = 1, \dots, l$, i.e. inside the unit circle and poles at $\bar{z}_k^{-1}$, i.e. outside of the unit circle. So the multiplication with this function may eliminate some zeros of modulus larger than one and produces its “mirror images” as zeros inside the unit circle. Later on in Chapter 6 we give a characterization of the matrix valued rational inner functions.

Returning to the factorization 2.21, since $B(z)$ is analytic inside the unit circle the function

$$g(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{it\lambda} B(e^{-i\lambda}) d\lambda \quad (2.23)$$

is zero if $t < 0$, consequently

$$B(e^{-i\lambda}) = \sum_{t=0}^{\infty} g(t)e^{-i\lambda t}. \quad (2.24)$$

This gives that in the representation 2.13

$$y(t) = \sum_{s=-\infty}^t g(t-s)w(s) = \sum_0^{\infty} g(s)w(t-s), \quad (2.25)$$

where w is an appropriate white noise.

The following theorem characterizes the rationality of the spectral density in terms of the covariance sequence $R(t)$, $t \in \mathbb{Z}$. In order to formulate the theorem we need the notations below. For arbitrary $n, m \in \mathbb{N}$ denote

$$\mathcal{H}_{n,m} = \begin{bmatrix} R(1) & R(2) & R(3) & \cdots & R(m) \\ R(2) & R(3) & & & \\ R(3) & & & & \\ \vdots & & & & \\ & & & & R(n+m-1) \end{bmatrix}$$

A matrix having the structure above is called *block-Hankel* matrix.

Also denote by $\mathcal{H}$ the infinite dimensional block-Hankel matrix:

$$\mathcal{H} = \begin{bmatrix} R(1) & R(2) & R(3) & \cdots \\ R(2) & R(3) & & \\ R(3) & & & \\ \vdots & & & \end{bmatrix}. \quad (2.26)$$

THEOREM 2.8. (Kronecker) *The spectral density of a stationary time series $y(t)$, $t \in \mathbb{Z}$, is rational if and only if*

$$\sup_{n,m} [\text{rank } \mathcal{H}_{n,m}] < \infty. \quad (2.27)$$

Let us introduce the quantity

$$\text{rank } \mathcal{H} = \sup_{n,m} [\text{rank } \mathcal{H}_{n,m}].$$

Note, that the matrix $\mathcal{H}$ is infinite dimensional, thus - without any further elaboration - we might consider $\text{rank } \mathcal{H}$ only as a notation.

Thus the Kronecker's theorem can be stated as the finiteness of $\text{rank } \mathcal{H}$ is equivalent to the rationality of the spectral density.

2.3. Regularity of stationary processes. Wold decomposition

We are going to show that $\text{rank } \mathcal{H}$ can also be described as the dimension of the range of an appropriate linear operator. To this aim let us analyse the structure of the subspaces $H_t^-(y)$, $H_t^+(y)$, $t \in \mathbb{Z}$. Obviously

$$U H_t^-(y) = H_{t+1}^-(y),$$

$$U^* H_t^+(y) = H_{t-1}^+(y)$$

and

$$H_{t+1}^-(y) \supset H_t^-(y), \quad H_{t-1}^+(y) \supset H_t^+(y).$$

Especially $H^-(y)$ is invariant under the backward shift U^* , $H^+(y)$ is invariant under the (forward) shift operator U .

DEFINITION 2.9. *The process y is deterministic (or singular) if*

$$H_t^-(y) = H_{t+1}^-(y).$$

REMARK 2.10. *If the equality holds for a $t \in \mathbb{Z}$ then it is valid for all $t \in \mathbb{Z}$.*

DEFINITION 2.11. *The process y is nondeterministic (or regular) if*

$$H_t^-(y) \neq H_{t+1}^-(y).$$

for some $t \in \mathbb{Z}$ (and consequently for all $t \in \mathbb{Z}$).

DEFINITION 2.12. *The process y is purely nondeterministic (or completely regular) if*

$$H_{-\infty}(y) = \bigcap_t H_t^-(y) = \{0\}.$$

These properties can be characterized in terms of the spectral measure. The following theorem was proved by Kolmogorov. (See Caines (Caines, 1988).)

THEOREM 2.13. *Assume that y is a full rank process. Then it is purely nondeterministic if and only if its spectral density f satisfies*

$$\int_{-\pi}^{\pi} \log [\det f(\lambda)] d\lambda > -\infty. \quad (2.28)$$

Let us observe that $H_{-\infty}(y)$ is U and U^* invariant, i.e. the restriction of U to this subspace is a unitary operator. It can be proved, that it is the maximal subspace of H^- with this property.

In general, the following decomposition holds

$$H^- = H_{-\infty}(y) \oplus \sum_{n \geq 0} U^{*n} D, \quad (2.29)$$

where

$$D = H^- \ominus U^* H^-.$$

is the *innovation* space. This is the so-called *Wold-Kolmogorov* decomposition.

DEFINITION 2.14. *The restriction of the operator U^* to the subspace H^- is called restricted shift operator and denoted by S .*

I.e. S is considered as an operator from H^- to H^- . It is obviously an isometric operator. Its adjoint is

$$S^* = \text{Pr}_{H^-} U,$$

where Pr_{H^-} is the projection to H^- . Observe that $D = \ker S^*$. If y is purely nondeterministic then

$$H^- = \sum_{n \geq 0} U^{*n} D$$

and

$$H_0 = \sum_{n=-\infty}^{+\infty} U^{*n} D.$$

Obviously, $\dim D < p$.

Now let us choose an orthonormal basis in D . Denote its elements by $w_j(-1)$, $j = 1, \dots, q$ ($q = \dim D$) and define

$$w_j(t) = U^{t+1} w_j(-1), \quad t \in \mathbb{Z}.$$

These vectors form an orthonormal basis in H_0 , and the process $w(t)$, $t \in \mathbb{Z}$, is a white noise process. The vectors $y(t)$ can be written as linear combinations of these vectors. Due to stationarity we obtain the following representation (Wold decomposition):

$$y(t) = \sum_{s=-\infty}^t g(t-s)w(s), \quad (2.30)$$

(cf. equations 2.13 and 2.25). This means also that $y(t)$ is generated by a causal filter.

Observe, that any process with the representation above is purely non-deterministic, because

$$H_t^-(y) \subseteq H_t^-(w) \quad (2.31)$$

and $\bigcap_t H_t^-(w)$ is orthogonal to the vectors $w_j(s)$, $j = 1, \dots, q$, $s \in \mathbb{Z}$, consequently

$$\bigcap_t H_t^-(w) = \{0\}.$$

Let us mention that in the present case, when we have constructed $w(t)$ using the subspace D , we have that $H_t^-(w) = H_t^-(y)$, but to assure that y is purely nondeterministic only the inclusion $H_t^-(y) \subset H_t^-(w)$ was needed.

THEOREM 2.15. *If the purely nondeterministic process y is of rank q , then $q = \dim D$, where $D = H^- \ominus U^* H^-$.*

Proof. Let $q' = \dim D$. Theorem 2.3 shows that if y is of rank q then it can be written as

$$y(t) = \sum_{s=-\infty}^{+\infty} \tilde{g}(t-s) \tilde{w}(s)$$

(see equation 2.13), where $\tilde{w}$ is a q -dimensional white noise process the components of which are in $H(y)$. Consequently the vectors $\tilde{w}_j(s)$, $j = 1, \dots, q$, $s \in \mathbb{Z}$ generate $H(y)$, moreover they form an orthonormal basis in it.

Also the process $w(t)$, $t \in \mathbb{Z}$, constructed earlier via choosing first an orthonormal basis in D , then shifting it, determines another orthonormal basis in $H(y)$.

$$y(t) = \sum_{s=-\infty}^t g(t-s) w(s).$$

(See equation 2.30.)

It follows that the spectral density of y can be written as

$$f(\lambda) = \frac{1}{2\pi} B(e^{-i\lambda}) B^*(e^{-i\lambda}),$$

where

$$B(z) = \sum_{k=0}^{\infty} g(k) z^k$$

is a $p \times q'$ matrix function, thus the rank of f cannot be larger, then the size of B , i.e.

$$q \leq q'. \quad (2.32)$$

On the other hand the elements of the process w can be expressed as linear combinations of the values of $\tilde{w}$

$$w(t) = \sum_{s=-\infty}^{+\infty} c(t-s)\tilde{w}(s).$$

But the spectral density of w is $I_{q'}/2\pi$, so

$$I_{q'} = C(e^{-i\lambda})C^*(e^{-i\lambda}), \quad (2.33)$$

where

$$C(z) = \sum_{k=-\infty}^{+\infty} c(k)z^k.$$

The matrix function C is of size $q' \times q$ so comparing the ranks of both sides we get that

$$q' \leq q,$$

which together with the inequality 2.32 leads to the equality $q = q'$. ■

Observe that the function C should be a square-matrix function satisfying the equation 2.33 which is identical to equation 2.22 in the definition of inner functions except that in the present case it is not necessarily analytic inside the unit circle.

DEFINITION 2.16. A $q' \times q$ matrix function C satisfying

$$C(e^{-i\lambda})C^*(e^{-i\lambda}) = I_{q'}$$

is called all-pass function.

Observe that if $q' = q = 1$ then the transformation defined by the coefficients of an all-pass function as

$$\xi \mapsto \sum_{t=-\infty}^{\infty} c(-t)U^t\xi,$$

where

$$C(z) = \sum_{t=-\infty}^{\infty} c(t)z^t$$

maps an orthonormal basis into another one, so in geometric terms it is an orthogonal transformation.

Notice that with some care this can be generalized to the case when q, q' are arbitrary. We do not want to go into details but let us point out that this generalization is not straightforward due to the fact that, when C is a matrix valued function then the underlying Hilbert space on which the corresponding operator is defined should contain q -dimensional vectors as its elements.

3. Purely nondeterministic stationary processes with rational spectral density

In this chapter the basic representation possibilities of the various spectral factors are considered, namely state-space representation, ARMA type, MFD type representations with their dual forms. Let us point out that similarly to the general state-space representations no assumptions concerning the zero structure are prescribed in the ARMA and MFD case, as well. After this weak- and strong state-space representations are analysed in detail, starting from the factorization of the Hankel matrix defined by the covariance sequence of the stationary process and presenting the matrix inequalities introduced by Faurre. But continuing the analysis of these inequalities we present a solution of the strong state-space realization problem, constructing the state vector process using the Riesz representation theorem of linear functionals. Forward and backward state-space representations, and a transformation formula between them are considered. We close this chapter with the so-called inner product representation enabling us to construct forward and backward state vector processes at the same time. The basic observation behind this construction is that the factorization of Hankel matrix is nothing else then expressing the elements of this matrix, which are bone fide covariances, i.e. scalar products in a Hilbert space as scalar products of finite dimensional vectors.

In the sequel we shall analyse the structure of purely nondeterministic stationary processes with rational spectral density having full rank. As we have seen they can be expressed as linear combinations of the elements of a white noise process w . The coefficients in these linear combinations determine the so-called transfer function,

$$B(z) = \sum_{k=-\infty}^{\infty} g(k)z^k \quad (3.1)$$

(cf. equation 2.16) or in our case a (left) spectral factor of the spectral density f (cf. equation 2.15). If

$$H_0^-(y) \subset H_0^-(w)$$

then the corresponding coefficients are zero for the negative indices, thus we get an *analytic* (inside the unit circle) spectral factor:

$$B_a(z) = \sum_{k=0}^{\infty} g(k)z^k. \quad (3.2)$$

If

$$H_0^+(y) \subset H_0^+(w),$$

then

$$y(t) = \sum_{s=t}^{\infty} g(t-s)w(s) = \sum_{-\infty}^0 g(s)w(t-s) \quad (3.3)$$

thus

$$B_{ca}(z) = \sum_{k=-\infty}^0 g(k)z^k \quad (3.4)$$

is analytic outside of the unit circle, it is a so-called *coanalytic* spectral factor. The terminology expresses the fact that B^* (where $B^*(z) = \overline{B(\bar{z}^{-1})}^T$) is analytic inside the unit circle.

Instead of analytic (coanalytic) spectral factors sometimes we use the terminology *causal* (*anticausal*) spectral factor. These notations are related to the following situation. In case of analytic transfer function the value of the process y at t is generated from the *past* (with respect to t) of the white noise w and this is called as the causal situation. If the transfer function is coanalytic, then the corresponding value $y(t)$ is generated from the *future* of $w(t)$ leading to the anticausal case.

In the general case when the coefficients in the linear combination in equation 3.1 produce a double infinite series, then they define an *acausal* spectral factor.

3.1. Parametrization of spectral factors (ARMA, MFD, state-space representations)

We are going to introduce several types of possible parametrizations of the spectral factors. Let us mention that usually they do not provide a minimal set of parameters, i.e. they give an overparametrization of the set of stationary time series with rational spectral density. (For various types of minimal parametrizations see Ober (Ober, 1996) and Hanzon and Ober (Ober and Hanzon, 1997).)

The spectral factors can be written in state-space equation forms (SSR), in autoregressive-moving average (ARMA) form, in matrix fractional form (MFD) and in dual ARMA and dual MFD form, as well.

These notions can be defined directly as forms of the spectral factor, but also using the stationary process y and a white noise process w . We start with this second possibility.

3.1.1. STATE-SPACE REPRESENTATIONS (SSR)

If there exists a stationary process $x(t)$, $t \in \mathbb{Z}$ such that the equations

$$\begin{aligned} x(t+1) &= Ax(t) + Bw(t) \\ y(t) &= Cx(t) + Dw(t) \end{aligned} \quad (3.5)$$

hold, where A is a stable matrix, i.e. its spectral radius $\varrho(A) < 1$, and $w(t+1)$, $w(t+2)$, ... are uncorrelated with $x(t)$, $x(t-1)$, ... then the system 3.5 defines a (forward) state-space representation of y . In this case the process x is the so-called state process. From the first equation – state-equation – we get that

$$x(t) = A^k x(t-k) + \sum_{j=1}^k A^{j-1} B w(t-j). \quad (3.6)$$

Taking the limit $k \rightarrow \infty$, using the stability of A we get that $A^k x(t-k) \rightarrow 0$ in L_2 sense, so

$$x(t) = \sum_{s=-\infty}^{t-1} A^{t-s-1} B w(s). \quad (3.7)$$

The second equation – observation-equation – gives that

$$y(t) = Dw(t) + \sum_{s=-\infty}^{t-1} CA^{t-s-1} B w(s). \quad (3.8)$$

The corresponding spectral factor is

$$D + \sum_{j=1}^{\infty} CA^{j-1} B z^j = D + C(z^{-1}I - A)^{-1}B. \quad (3.9)$$

The stability of A ensures that the poles of this spectral factor are outside of the unit circle, i.e. we have got an analytic spectral factor.

Figure 3.1 represents the flow-diagram of a forward state-space realization.

In a symmetric way, if $\bar{w}(t)$, $t \in \mathbb{Z}$ is a white noise process, $\bar{x}(t)$, $t \in \mathbb{Z}$, is a stationary process such that the equations

$$\begin{aligned} \bar{x}(t) &= \bar{A}\bar{x}(t+1) + \bar{B}\bar{w}(t) \\ y(t) &= \bar{C}\bar{x}(t+1) + \bar{D}\bar{w}(t) \end{aligned} \quad (3.10)$$

hold, ($\bar{A}$ is again a stable matrix, but now the variables $\bar{w}(t)$, $\bar{w}(t-1)$, ... and $\bar{x}(t+1)$, $\bar{x}(t+2)$, ... are uncorrelated) defining a so-called backward state-space representation, then they lead to the infinite series representation

$$y(t) = \bar{D}\bar{w}(t) + \sum_{s=t+1}^{\infty} \bar{C}\bar{A}^{s-t-1}\bar{B}\bar{w}(s) \quad (3.11)$$

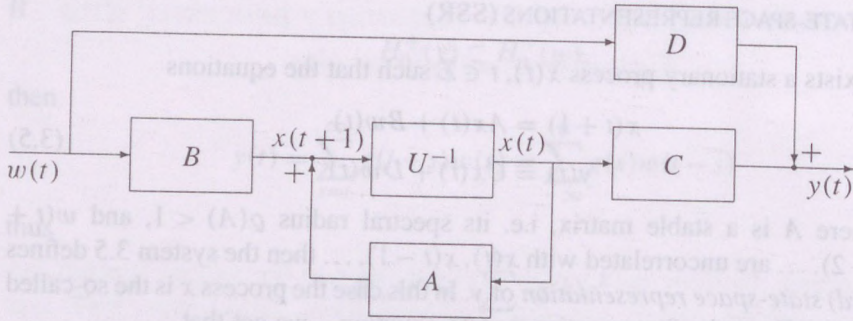


Figure 3.1. The flow-diagram of a forward state-space realization

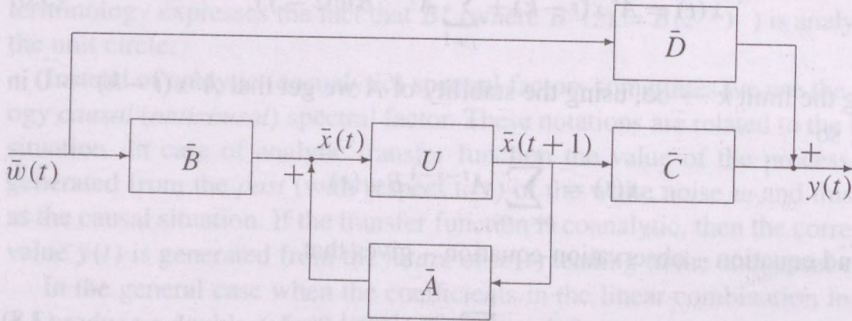


Figure 3.2. The flow diagram of a backward state-space realization

determining to the spectral factor

$$\bar{D} + \sum_{j=1}^{\infty} \bar{C} \bar{A}^{j-1} \bar{B} z^{-j} = \bar{D} + \bar{C} (zI - \bar{A})^{-1} \bar{B}. \quad (3.12)$$

This is a coanalytic spectral factor.

Figure 3.2 represents the flow-diagram of a backward state-space realization.

As we shall see later in various situations it is natural to consider a common generalization of forward and backward state-space realizations. So consider the system of equations

$$\begin{aligned} x_1(t+1) &= A_1 x_1(t) + B_1 w(t) \\ \bar{x}_2(t) &= \bar{A}_2 \bar{x}_2(t+1) + \bar{B}_2 w(t) \\ y(t) &= C_1 x_1(t) + \bar{C}_2 \bar{x}_2(t) + D w(t) \end{aligned} \quad (3.13)$$

where $A_1, \bar{A}_2$ are stable matrices. From this we get the representation

$$y(t) = D w(t) + \sum_{s=-\infty}^{t-1} C_1 A_1^{t-s-1} B_1 w(s) + \sum_{s=t+1}^{\infty} \bar{C}_2 \bar{A}_2^{s-t-1} \bar{B}_2 w(s) \quad (3.14)$$

defining an acausal spectral factor

$$D + C_1 (z^{-1}I - A_1)^{-1} B_1 + \overline{C}_2 (zI - \overline{A}_2)^{-1} \overline{B}_2. \quad (3.15)$$

Let us remark that equations 3.5, 3.10 and 3.13 determine state-space representations or in other words realizations of the process $y(t)$, $t \in \mathbb{Z}$. At the same time when we use the terminology state-space realization without explicitly specifying whether it is forward, backward or general representation then we refer to forward state-space representations.

3.1.2. ARMA REPRESENTATIONS

The autoregressive-moving average (ARMA) representation can be described as follows:

$$\begin{aligned} y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = \\ b_0 w(t) + b_1 w(t-1) + \dots + b_m w(t-m), \end{aligned} \quad (3.16)$$

where $a_1, \dots, a_n, b_0, \dots, b_m$ are matrices. Forward ARMA representations relate the *most recent past* elements of $y(t)$ and $w(t)$ at a given t .

Introducing the matrix polynomials

$$\begin{aligned} I + a_1 z + \dots + a_n z^n &= A(z) \\ b_0 + b_1 z + \dots + b_m z^m &= B(z) \end{aligned}$$

we may write shortly

$$A(U^{-1})y(t) = B(U^{-1})w(t), \quad (3.17)$$

where U^{-1} is the backward shift operator. The corresponding spectral factor can be written as

$$A(e^{-i\lambda})^{-1} B(e^{-i\lambda}). \quad (3.18)$$

Obviously, for arbitrary $a_1, \dots, a_n, b_0, \dots, b_m$ matrices the ARMA equation 3.16 may not have a stationary solution.

DEFINITION 3.1. *The representation 3.16 is called forward ARMA (FARMA) representation, if the spectral factor $A(z)^{-1}B(z)$ is analytic (inside the unit circle), in other words it is causal.*

DEFINITION 3.2. *The representation 3.16 is called backward ARMA (BARMA) representation, if the spectral factor $A(z)^{-1}B(z)$ is coanalytic (i.e. analytic outside of the unit circle), in other words it is anticausal.*

3.1.3. MFD REPRESENTATIONS

Similarly to the representation 3.16 we may consider the equations:

$$\begin{aligned} y(t) + a_1 y(t+1) + \dots + a_n y(t+n) = \\ b_0 w(t) + b_1 w(t+1) + \dots + b_m w(t+m), \end{aligned} \quad (3.19)$$

where $a_1, \dots, a_n, b_0, \dots, b_m$ are matrices. This is the so-called matrix fractional description (MFD). Thus the MFD representations relate the *most recent future* elements of $y(t)$ and $w(t)$ at a given value t .

Introducing the matrix polynomials

$$\begin{aligned} I + a_1 z + \dots + a_n z^n &= P(z) \\ b_0 + b_1 z + \dots + b_m z^m &= Q(z) \end{aligned}$$

we may write shortly

$$P(U)y(t) = Q(U)w(t), \quad (3.20)$$

where U is the forward shift operator. The corresponding spectral factor can be written as

$$P(e^{i\lambda})^{-1} Q(e^{i\lambda}). \quad (3.21)$$

Obviously, for arbitrary $a_1, \dots, a_n, b_0, \dots, b_m$ matrices the MFD equation 3.19 may not have a stationary solution.

DEFINITION 3.3. *The representation 3.19 is called forward MFD (FMFD) representation, if the spectral factor $P(z^{-1})^{-1} Q(z^{-1})$ is analytic (inside the unit circle), in other words it is causal.*

DEFINITION 3.4. *The representation 3.19 is called backward MFD (BMFD) representation, if the spectral factor $P(z^{-1})^{-1} Q(z^{-1})$ is coanalytic (i.e. analytic outside of the unit circle), in other words it is anticausal.*

3.1.4. DUAL ARMA AND MFD REPRESENTATIONS

To define the dual representations we should consider another weakly stationary process denoted by $\xi(t)$, $t \in \mathbb{Z}$, $E\xi(t) = 0$. Both the output process and the input process are described using this auxiliary process.

3.1.4.1. *Dual ARMA representations*

The dual ARMA representation is defined by the following equations:

$$\begin{aligned} y(t) &= a_0 \xi(t) + a_1 \xi(t-1) + \dots + a_n \xi(t-n) \\ w(t) &= b_0 \xi(t) + b_1 \xi(t-1) + \dots + b_m \xi(t-m), \end{aligned} \quad (3.22)$$

where $a_0, a_1, \dots, a_n, b_0, \dots, b_m$ are matrices, $\xi(t), t \in \mathbb{Z}$ is a q -dimensional weakly stationary process, called the partial state, $E\xi(t) = 0$.

Introducing the matrix polynomials

$$\begin{aligned} a_0 + a_1 z + \dots + a_n z^n &= A_d(z) \\ b_0 + b_1 z + \dots + b_m z^m &= B_d(z) \end{aligned} \quad (3.23)$$

we may write shortly

$$\begin{aligned} y(t) &= A_d(U^{-1})\xi(t) \\ w(t) &= B_d(U^{-1})\xi(t), \end{aligned} \quad (3.24)$$

where U^{-1} is the backward shift operator. The corresponding spectral factor can be written as

$$A_d(e^{-i\lambda})B_d(e^{-i\lambda})^{-1}. \quad (3.25)$$

DEFINITION 3.5. *The representation 3.22 is called dual forward ARMA (dFARMA) representation, if the spectral factor $A_d(z)B_d(z)^{-1}$ is analytic (inside the unit circle), in other words it is causal.*

The Figure 3.3 compares the forward ARMA and dual forward ARMA representations.

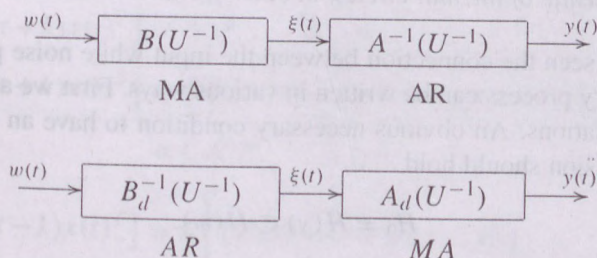


Figure 3.3. Forward ARMA and dual forward ARMA representations of the process y . The process ξ is the partial state process

DEFINITION 3.6. *The representation 3.22 is called dual backward ARMA (dBARMA) representation, if the spectral factor $A_d(z)B_d(z)^{-1}$ is coanalytic (i.e. analytic outside of the unit circle), in other words it is anticausal.*

3.1.4.2. Dual MFD representations

Similarly to the representation 3.22 we may consider the equations:

$$\begin{aligned} y(t) &= a_0 \xi(t) + a_1 \xi(t+1) + \dots + a_n \xi(t+n) \\ w(t) &= b_0 \xi(t) + b_1 \xi(t+1) + \dots + b_m \xi(t+m), \end{aligned} \quad (3.26)$$

where $a_0, a_1, \dots, a_n, b_0, \dots, b_m$ are matrices, $\xi(t), t \in \mathbb{Z}$ is a q -dimensional weakly stationary process, $E\xi(t) = 0$. This is the so-called dual MFD form.

Introducing the matrix polynomials

$$\begin{aligned} a_0 + a_1 z + \dots + a_n z^n &= P_d(z) \\ b_0 + b_1 z + \dots + b_m z^m &= Q_d(z) \end{aligned} \quad (3.27)$$

we may write shortly

$$\begin{aligned} y(t) &= P_d(U)\xi(t) \\ w(t) &= Q_d(U)\xi(t), \end{aligned} \quad (3.28)$$

where U is the forward shift operator. The corresponding spectral factor can be written as

$$P_d(e^{i\lambda}) Q_d(e^{i\lambda})^{-1}. \quad (3.29)$$

DEFINITION 3.7. *The representation 3.26 is called dual forward MFD (dFMFD) representation, if the spectral factor $P_d(z^{-1})Q_d(z^{-1})^{-1}$ is analytic (inside the unit circle), in other words it is causal.*

DEFINITION 3.8. *The representation 3.26 is called dual backward MFD (dBMFD) representation, if the spectral factor $P_d(z^{-1})Q_d(z^{-1})^{-1}$ is coanalytic (i.e. analytic outside of the unit circle), in other words it is anticausal.*

As we have seen the connection between the input white noise process and the output stationary process can be written in various ways. First we analyse the state space representations. An obvious necessary condition to have an SSR is that the following inclusion should hold

$$H_0 = H(y) \subset H(w).$$

In the next section we prove that if this condition is fulfilled then it is always possible to construct a state-space representation. If in addition the connection

$$H_t^-(y) \subset H_t^-(w)$$

holds then we get an analytic or causal spectral factor. Similarly, if

$$H_t^+(y) \subset H_t^+(w)$$

holds then we obtain coanalytic or anticausal spectral factor.

In general, there are several state-space descriptions of a stationary process. The state-space can always be enlarged because in the observation equation defining the matrix C in an appropriate way we can nullify the unnecessary parts of the state vector. This raises the question of minimality of state-space representations.

DEFINITION 3.9. *A state-space representation is called minimal, if the dimension of the state vector is minimal among all state-space descriptions, generating the same output process.*

As we shall see in the next sections the minimality effects on the one hand the possible choice of the input white noise process and on the other hand the actual state-space description of the corresponding spectral factor.

3.2. Weak state-space realizations

Let us first consider the second-order connections between the various elements of a causal state-space representation.

$$x(t+1) = Ax(t) + Bw(t)$$

$$y(t) = Cx(t) + Dw(t).$$

Denote

$$P = E[x(t)x(t)^T]. \quad (3.30)$$

Then

$$E[x(t+k)x(t)^T] = A^k P, \quad k \geq 0$$

$$\begin{aligned} E[y(t+k)x(t)^T] &= E[(Cx(t+k) + Dw(t+k))x(t)^T] = \\ &= CA^k P, \quad k \geq 0 \end{aligned}$$

$$E[y(t-k)x(t)^T] = E\left[(Cx(t-k) + Dw(t-k))\right]$$

$$\cdot \left(A^k x(t-k) + \sum_{j=1}^k A^{j-1} Bw(t-j) \right)^T \Bigg]$$

$$= (CPA^T + DB^T)(A^T)^{k-1}, \quad k \geq 1$$

$$E[y(t)y(t-k)^T] = CA^{k-1}(APC^T + BD^T) = R(k), \quad k \geq 1$$

$$E[y(t)y(t)^T] = CPC^T + DD^T$$

$$E[x(t+1)x(t+1)^T] = P = APA^T + BB^T.$$

Let us introduce the notation

$$G = APC^T + BD^T.$$

We have obtained that

$$\begin{aligned} R(k) &= CA^{k-1}G, \quad k \geq 1 \\ R(0) &= CPC^T + DD^T \\ P &= APA^T + BB^T. \end{aligned} \quad (3.31)$$

Let us analyse first the so-called *Ljapunov-equation* (in discrete time case)

$$P = APA^T + BB^T, \quad (3.32)$$

together with its stochastic counterpart

$$x(t+1) = Ax(t) + Bw(t). \quad (3.33)$$

LEMMA 3.10. *If $\varrho(A) < 1$ then the stochastic equation has a stationary solution for which $w(t), w(t+1), \dots$ are uncorrelated to $x(t), x(t-1), \dots$*

Proof. Obviously the series

$$x(t) = \sum_{s=-\infty}^t A^{t-s-1} Bw(s)$$

is convergent in L_2 -sense and gives a solution of the equation. ■

REMARK 3.11. *This solution is unique since from the equation 3.33 it follows that*

$$x(t) = A^k x(t-k) + \sum_{j=1}^k A^{j-1} Bw(t-j) \quad (3.34)$$

and $\text{Var}(A^k x(t-k)) = A^k P (A^T)^k \rightarrow 0$.

LEMMA 3.12. *If $BB^T > 0$ then for any matrices P, A satisfying the Ljapunov-equation 3.32 we have that $P > 0, \varrho(A) < 1$.*

Proof. Since $P \geq BB^T$ we obtain that $P > 0$. Consider now a left eigen-vector of A :

$$\lambda z^T = z^T A.$$

Then

$$z^T P z = z^T A P A^T z + z^T B B^T z$$

so

$$(1 - |\lambda|^2) z^T P z = z^T B B^T z > 0,$$

proving that $|\lambda| < 1$. ■

LEMMA 3.13. *If there exists a purely nondeterministic solution $x(t)$, $t \in \mathbb{Z}$ of the equation 3.33, for which $P > 0$ then $\varrho(A) < 1$.*

Proof. Let λ be a left eigenvalue of A . Since for any non-zero vector z we have that $z^T P z > 0$ we get that $|\lambda| \leq 1$. On the other hand

$$\Pr_{H_s^-(x)} x(t) = A^{t-s} x(s)$$

which tends to zero in view of $\bigcap_s H_s^-(x) = \{0\}$. This implies that

$$A^j P (A^T)^j \rightarrow 0, \quad \text{if } j \rightarrow \infty.$$

P is positive definite, so $\varrho(A) < 1$. ■

Let us denote by

$$C_N(A, B) = [B, AB, \dots, A^{N-1}B]. \quad (3.35)$$

LEMMA 3.14. *If for some N the matrix $C_N(A, B)$ is of full rank, then $P > 0$.*

Proof. Iterating the Ljapunov-equation 3.32 we get that

$$\begin{aligned} P &= A^{N+1} P (A^T)^{N+1} + B B^T + A B B^T A^T + \dots + A^N B B^T (A^T)^N \\ &= A^{N+1} P (A^T)^{N+1} + C_N(A, B) C_N(A, B)^T. \end{aligned}$$

Since the last term is nonsingular we get that $P > 0$. ■

Assume that A is an $n \times n$ -matrix. The Cayley–Hamilton theorem gives that

$$\text{rank } C_N(A, B) = \text{rank } C_n(A, B), \quad \text{if } N \geq n.$$

DEFINITION 3.15. *The pair (A, B) is controllable² if*

$$\text{rank } C_n(A, B) = n. \quad (3.37)$$

² The terminology *controllable* is connected to deterministic systems. Namely, let us consider the system

$$\begin{aligned} x(t+1) &= A x(t) + B u(t) \\ x(0) &= x_0 \end{aligned} \quad (3.36)$$

Let us remark – without proof – the following theorem which describes the connection between the number of positive and negative eigenvalues of P and the number of eigenvalues of A inside or outside of the unit circle.

Assume that the Ljapunov-equation 3.32 holds, where $P = P^T$. For the symmetric matrix P let us denote by $\pi(P)$, $\nu(P)$, $\delta(P)$ the number of positive, negative and zero eigenvalues of P . For the matrix A $\pi(A)$, $\nu(A)$, $\delta(A)$ denote the number eigenvalues of A outside of, inside of and on the unit circle. The main part of the following theorem was proved by Glover (Glover, 1984) in continuous time setting which can be transformed to our discrete time case via the Cayley-transform.

THEOREM 3.16. (Cf. Glover (Glover, 1984)) *Consider the Ljapunov-equation*

$$P = APA^T + BB^T, \quad P = P^T. \quad (3.38)$$

1. *There exists a unique solution of 3.38 if and only if $\lambda_i(A)\overline{\lambda_j(A)} \neq 1$ for all pairs of eigenvalues $\lambda_i(A)$, $\lambda_j(A)$ of the matrix A .*
2. *If $\delta(P) = 0$ (i.e. P is nonsingular) then $\pi(A) \leq \nu(P)$, $\nu(A) \leq \pi(P)$.*
3. *If $\delta(A) = 0$ then $\pi(P) \leq \nu(A)$, $\nu(P) \leq \pi(A)$.*
4. *If the pair (A, B) is controllable, then $\pi(A) = \nu(P)$, $\nu(A) = \pi(P)$, $\delta(A) = \delta(P) = 0$.*
5. *$\pi(P) + \nu(P) \geq \text{rank } C_n(A, B)$.*
6. *If $\pi(A) = n$ then $\pi(P) = 0$.*
7. *If $\nu(A) = n$ then $\nu(P) = 0$.*

As we have seen the existence of a causal state-space realization implies that $R(t)$ can be factorized as

$$R(t) = CA^{t-1}G, \quad t \geq 1. \quad (3.39)$$

(cf. equation 3.31)

where $x_0 \in \mathbb{R}^n$ is the initial state, $u(t)$, $t \in \mathbb{Z}$ is a deterministic input sequence.

The system 3.36 is called *reachable*, if for any state vector $x \in \mathbb{R}^n$ there exists a finite sequence of input vectors, such that using the initial condition $x_0 = 0$ we can reach the prescribed state x in finite time.

The system 3.36 is called *controllable*, if for any state vector $x \in \mathbb{R}^n$ there exists a finite sequence of input vectors, such that using the prescribed initial condition $x_0 = x$ we can control back the system to the origin in finite time.

The condition 3.37 is equivalent to the reachability of the system. But because of historical reasons, in that case the pair (A, B) is called controllable.

THEOREM 3.17. *The covariance sequence $R(t)$ of a stationary stochastic process $y(t)$, $t \in \mathbb{Z}$ can be written in the form of equation 3.39 if and only if $\text{rank}(\mathcal{H}) < \infty$.*

Proof. If there exist matrices A , C , G of size $n \times n$, $p \times n$, $n \times p$ solving 3.39 then

$$\begin{aligned} \mathcal{H}_{k,m} &= \begin{bmatrix} R(1) & R(2) & \dots & R(m) \\ R(2) & & & \\ \vdots & & & \\ R(k) & & & R(m+k-1) \end{bmatrix} \\ &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} [G \ AG \ \dots \ A^{m-1}G]. \end{aligned} \quad (3.40)$$

Since $C_m(A, G)$ has n rows we obtain that

$$\text{rank } \mathcal{H}_{k,m} \leq n, \quad \text{for all } k, m.$$

So

$$\text{rank } \mathcal{H} \leq n.$$

Assume now that $\text{rank } \mathcal{H} = n < \infty$. First we show that there are matrices $M_1, M_2, \dots$ of size $p \times n$ and $V_1, V_2, \dots$ of size $n \times p$ such that

$$\mathcal{H}_{k,m} = \mathcal{O}_k C_m, \quad \text{for all } k, m \geq 1,$$

where

$$C_m = [V_1, V_2, \dots, V_m]$$

$$\mathcal{O}_k^T = [M_1^T, M_2^T, \dots, M_k^T].$$

To this aim let us choose a number l such that

$$n = \text{rank } \mathcal{H}_{ll} = \text{rank } \mathcal{H}_{k,m} \quad \text{for all } k, m \geq l.$$

Let us factorize the matrix $\mathcal{H}_{ll}$ as

$$\mathcal{H}_{ll} = \mathcal{O}_l C_l,$$

where C_l has exactly n rows. This can be done for example choosing first a maximal linearly independent set of row vectors of $\mathcal{H}_{ll}$ and they form the rows of C_l . Then every row of $\mathcal{H}_{ll}$ can be written as linear combinations of the rows C_l . The coefficients in these linear combinations form the corresponding rows of $\mathcal{O}_l$.

This construction can easily be extended because if $k \geq l$ then all row vectors of $\mathcal{H}_{k,l}$ are linear combinations of the rows of C_l determining in this way the rows of

$\mathcal{O}_k$. Obviously the columns of $\mathcal{O}_k$ form a basis in the linear subspace generated by the columns of $\mathcal{H}_{k,m}$. Writing each column as a linear combination of these vectors forming the basis we get the corresponding columns of C_l .

Of course another possibility would be first to choose a linearly independent set of column-vectors of $\mathcal{H}_{ll}$ they form the matrix $\mathcal{O}_l$ and the elements in the columns of C_l determine the appropriate linear combinations of these column-vectors to produce an arbitrary column of $\mathcal{H}_{ll}$.

Forming the $p \times n$, $n \times p$ block matrices in C_m , $\mathcal{O}_l$ we get the sequences $M_1, M_2, \dots, V_1, V_2, \dots$

It remains to show that the equations

$$M_k = CA^{k-1}, \quad V_k = A^{k-1}G, \quad k \geq 1$$

can be solved. Define

$$C = M_1, \quad G = V_1.$$

The column vectors of $\mathcal{O}_l$ form a basis in the space generated by the column vectors of $H_{l,l+1}$. Especially, every column vector of

$$\begin{bmatrix} R(2) & \dots & R(l+1) \\ \vdots & & \\ R(l+1) & \dots & R(2l+1) \end{bmatrix} = \begin{bmatrix} M_2 \\ \vdots \\ M_{l+1} \end{bmatrix} [V_1 \dots V_l]$$

is in this subspace. Consequently

$$\text{range } \mathcal{O}_l \supset \text{range} \begin{bmatrix} M_2 \\ \vdots \\ M_{l+1} \end{bmatrix}.$$

Thus there exists a unique $n \times n$ matrix A for which

$$\mathcal{O}_l A = \begin{bmatrix} M_2 \\ \vdots \\ M_{l+1} \end{bmatrix},$$

in other words

$$M_{j+1} = M_j A, \quad j = 1, \dots, l.$$

So

$$M_j = CA^{j-1}, \quad j = 1, \dots, l+1.$$

This argument can be repeated for any $k \geq l$. The uniqueness of the solution guarantees that we get always the same matrix A . I.e.

$$M_j = CA^{j-1}, \quad j \geq 1,$$

implying that

$$R(k) = CA^{k-1}G, \quad k \geq 1.$$

Similar argument works on the matrices C_j , $j \geq l$. From the uniqueness we can conclude that

$$V_j = A^{j-1}G, \quad j \geq 1.$$

Observe, that the actual form of the matrices C , A , G depends on the choice of the factorization

$$\mathcal{H}_l = \mathcal{O}_l C_l.$$

Under the minimality of the number of columns in $\mathcal{O}_l$ this is unique except a nonsingular transformation $\mathcal{O}_l T^{-1}$, $T C_l$. This gives that C , A , G are unique except the transformation CT^{-1} , TAT^{-1} , TG .

In the previous construction

$$\text{rank } C_j = \text{rank } C_j(A, G) = n, \quad \text{if } j \geq l.$$

so the pair (A, G) is controllable. Also introducing the notation

$$\mathcal{O}_j(A, C) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{j-1} \end{bmatrix}$$

we obtain that

$$\text{rank } \mathcal{O}_j = \text{rank } \mathcal{O}_j(A, C) = n, \quad \text{if } j \geq l. \quad (3.41)$$

In this case we say that the pair (C, A) is observable³.

Since the size of the matrix A gives the dimension of the state-space it follows that this dimension is always at least $n = \text{rank } \mathcal{H}$.

³ The expression *observability* is again connected to deterministic systems. Namely, let us consider the system

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \\ x(0) &= x_0 \end{aligned} \quad (3.42)$$

where $x_0 \in \mathbb{R}^n$ is the initial condition.

The system 3.42 is called *observable* if for any unknown initial vector $x_0 \in \mathbb{R}^n$ observing the output for a long enough, but finite time the initial condition can be determined, assuming that the input sequence $u(t)$ is known.

The system 3.42 is called *constructible* if for any unknown initial vector $x_0 \in \mathbb{R}^n$ observing the output for a long enough, but finite time the value of the state variable at the end of the observation period can be determined, assuming that the input sequence $u(t)$ is known.

Condition 3.41 is equivalent to the constructibility of the system. But because of historical reasons, in this case the pair (C, A) is called observable.

Let us fix the matrices C , A , G . In order to get the other matrices in the state-space equations it remains to solve the equations

$$\begin{aligned} P &= APA^T + BB^T \\ G &= APC^T + BD^T \\ R(0) &= CPC^T + DD^T. \end{aligned} \quad (3.43)$$

for B , D and P .

Equation 3.43 can be written as follows

$$\begin{bmatrix} P - APA^T & G - APC^T \\ G^T - CPA^T & R(0) - CPC^T \end{bmatrix} = \begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B^T & D^T \end{bmatrix}.$$

The right hand side is positive semidefinite. Moreover every positive semidefinite matrix can be factorized in the form above. So it is enough to solve the following matrix inequalities

$$M(P) \geq 0, \quad P \geq 0, \quad (3.44)$$

where

$$M(P) = \begin{bmatrix} P - APA^T & G - APC^T \\ G^T - CPA^T & R(0) - CPC^T \end{bmatrix}. \quad (3.45)$$

Observe that in the inequalities 3.44 the only unknown quantity is P . After solving them finally we have to factorize $M(P)$. There are infinitely many factorizations of the form

$$M(P) = \begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B^T & D^T \end{bmatrix}.$$

But putting an additional requirement the number of rows in $[B^T, D^T]$ be minimal one – this defines the dimension of the input white noise – the remaining nonuniqueness is only a nonsingular transformation.

We do not carry out a complete analysis of this system of inequalities. But let us quote a theorem which was proved by *Faurre et al.* (cf. (Faurre et al., 1979)). Denote

$$\mathcal{P} = \{P \geq 0 | M(P) \geq 0\}. \quad (3.46)$$

THEOREM 3.18. (See *Faurre, Clerget and Germain* (Faurre et al., 1979)) $\mathcal{P}$ is a convex set. There exist $P_-, P_+ \in \mathcal{P}$ such that for any $P \in \mathcal{P}$ we have that

$$P_- \leq P \leq P_+. \quad (3.47)$$

(See also equation 5.15)

3.3. Strong state-space realizations

Instead of proving *Faurre's* theorem via direct analysis of the matrix inequality 3.44, we try to find immediately a state vector process $x(t)$, $t \in \mathbb{Z}$. As a side result of this construction in Chapter 5 we get a proof of the inequality 3.47.

THEOREM 3.19. *Let y be a purely nondeterministic full-rank stationary process for which $\text{rank } \mathcal{H} = n < \infty$. Let us assume that via the factorization of the Hankel-matrices $\mathcal{H}_{l,j}$ described in the proof of Theorem 3.17 we have determined the matrices C, A, G . Then*

1. $\varrho(A) < 1$,

2. *there exists a stationary process $x(t)$, $t \in \mathbb{Z}$, and a white noise $w(t)$, $t \in \mathbb{Z}$, providing a state space realization of $y(t)$, $t \in \mathbb{Z}$, i.e.*

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

3. *the coordinate variables of $x(0)$ are in H^- .*

Proof. First observe that $R(0) > 0$, because if there existed a nonzero vector z for which $R(0)z = 0$ then we would have that $z^T y(t) = 0$ for all $t \in \mathbb{Z}$, consequently $R(t)z = 0$ for all $t \in \mathbb{Z}$ implying that the rank of the spectral density should be less than p for all λ . This would contradict to the assumption that y is of full-rank.

At the same time

$$\Pr_{H_t^-(y)} y(0) \rightarrow 0, \quad \text{if } t \rightarrow -\infty,$$

since y is a purely nondeterministic process. Consequently

$$\Pr_{\{y(t)\}} y(0) \rightarrow 0, \quad \text{if } t \rightarrow -\infty,$$

i.e.

$$R(t)R(0)^{-1} \rightarrow 0.$$

In other words

$$CA'^{-1}G \rightarrow 0, \quad \text{if } t \rightarrow \infty.$$

This obviously implies that

$$\mathcal{O}_n(A, C)A^j C_n(A, G) \rightarrow 0, \quad \text{if } j \rightarrow \infty.$$

Using that (A, G) is controllable, (C, A) is observable, so

$$\text{rank } C_n(A, G) = \text{rank } \mathcal{O}_n(A, C) = n,$$

we obtain that

$$A^j \rightarrow 0, \quad \text{if } j \rightarrow \infty.$$

This means that $\varrho(A) < 1$ proving (1).

As concerns (2), and (3) we have to find a state vector $x(0)$ with components in H^- such that

$$E[x(0)y(-t)^T] = A^{t-1}G, \quad t \geq 1,$$

or writing componentwise

$$E[x_j(0)y_k(-t)] = (A^{t-1}G)_{jk}, \quad t \geq 1.$$

Thus fixing j the scalar products of $x_j(0)$ with $y_k(-t)$, $t \geq 1$ are already given. But $y_k(-t)$, $t \geq 1$, $k = 1, \dots, p$ form a generator system in H^- . Moreover, if for a finite linear combination

$$\sum \mu_t^T y(-t) = 0,$$

then

$$E\left[y(s) \left(\sum \mu_t^T y(-t)\right)^T\right] = 0, \quad s \geq 0,$$

i.e.

$$CA^s \left(\sum A^{t-1}G\mu_t\right) = 0.$$

Observability of (C, A) implies that $\sum A^{t-1}G\mu_t = 0$. Thus the functional

$$y_k(t) \rightarrow (A^{t-1}G)_{jk}, \quad t \geq 1$$

is linear, furthermore and $\varrho(A) < 1$ implies that it is continuous. So it can be extended to $H^- = H_0^-(y)$. But any continuous linear functional can be represented as a scalar product. So there exists a vector in H^- – denote it by $x_j(0)$ – for which

$$E[x_j(0)y_k(-t)] = (A^{t-1}G)_{jk}, \quad t \geq 1.$$

Applying the shift operator U we get a stationary process

$$x(t) = U^t x(0).$$

It remains to check that $x(t)$, $t \in \mathbb{Z}$ determines a state space representation for y . But

$$\text{cov}[x(t+1) - Ax(t), y(s)] = A^{t-s}G - AA^{t-s-1}G = 0, \quad \text{if } s \leq t-1.$$

Thus

$$\Pr_{H_t^-(y)} x(t+1) = Ax(t).$$

Similarly

$$\text{cov}[y(t) - Cx(t), y(s)] = R(t-s) - CA^{t-s-1}G = 0, \quad \text{if } s \leq t-1.$$

Thus

$$\Pr_{H_t^-(y)} y(t) = Cx(t).$$

Observe that the coordinates of $x(t+1) - Ax(t)$, $y(t) - Cx(t)$ are in $H_{t+1}^-(y) \ominus H_t^-(y)$ (moreover they generate it). Choosing an arbitrary orthonormal basis in $H_0^-(y) \ominus H_{-1}^-(y)$ and applying the shift operator we get a white noise sequence $w(t)$.

Since every elements in the subspaces $H_{t+1}^-(y) \ominus H_t^-(y)$ can be written as linear combination of the components of $w(t)$, we get that there exist matrices B , D for which the equations

$$x(t+1) = Ax(t) + Bw(t)$$

$$y(t) = Cx(t) + Dw(t)$$

hold.

REMARK 3.20. The covariance matrix P of the state vector $x(t)$ obviously satisfies

$$P > 0, \quad M(P) \geq 0.$$

Observe that $\text{rank } M(P) = \dim[H_0^-(y) \ominus H_{-1}^-(y)] = p$ using that y is of full rank.

REMARK 3.21. Since the realization constructed above has the special property that the components of $x(0)$ are in H^- we shall denote it by $x_-(t)$ and its covariance matrix by P_- . Similarly the corresponding white noise will be denoted by $w_-(t)$ and the matrices by B_- , D_- . Thus

$$\begin{aligned} x_-(t+1) &= Ax_-(t) + B_-w_-(t) \\ y(t) &= Cx_-(t) + D_-w_-(t). \end{aligned} \tag{3.48}$$

The components of $w_-(t)$ form a basis in the so-called innovation subspace $H_{t+1}^-(y) \ominus H_t^-(y)$, so this realization is called (forward) innovation realization, $w_-(t)$ is the (forward) innovation process.

REMARK 3.22. The dimension of x_- is $n = \text{rank } \mathcal{H}$, so 3.48 is a minimal realization. Especially, $\text{rank } \mathcal{H}$ gives the dimension of the state space of any minimal realization.

We would like to prove that the matrix P_- is the smallest one among the covariance matrices of state vectors of minimal realization.

To this aim let us consider an arbitrary (forward) minimal realization of y .

$$\begin{aligned} x(t+1) &= Ax(t) + Bw(t) \\ y(t) &= Cx(t) + Dw(t). \end{aligned} \quad (3.49)$$

As we have seen an arbitrary realization provides a factorization of the Hankel-matrices. But this factorization is unique up to a nonsingular transformation. So applying a transformation in the state space choosing another coordinate system we can achieve that the matrices C , A , G be the same ones we have used in the construction of the innovation realization.

Since $\varrho(A) < 1$ we get that

$$H_t^-(y) \subset H_t^-(w).$$

On the other hand

$$x(0) = \sum_{s=1}^{\infty} A^{s-1} Bw(-s),$$

so the components of $x(0)$ are in $H_0^-(w)$. Also

$$y(t) - CA^t x(0) = Dw(t) + \sum_{s=0}^{t-1} CA^{t-s} Bw(s), \quad t \geq 0.$$

Consequently

$$\Pr_{H_0^-(w)} y(t) = CA^t x(0), \quad t \geq 0.$$

Denote the subspace spanned by the coordinates of $x(0)$ by X . The observability of the pair (A, C) implies that

$$\Pr_{H_0^-(w)} H_0^+(y) = X. \quad (3.50)$$

Especially, if X_- denotes the subspace spanned by the coordinates of $x_-(0)$, then

$$\Pr_{H_0^-(w_-)} H_0^+(y) = X_-. \quad (3.51)$$

Observe that

$$H_0^-(w_-) = H_0^-(y). \quad (3.52)$$

So $\text{rank } \mathcal{H} = \dim(\Pr_{H_0^-(y)} H_0^+(y))$.

THEOREM 3.23. *If the process $x(t)$, $t \in \mathbb{Z}$ determines a minimal state space realization then*

$$P = E[x(t)x^T(t)] \geq P_-.$$

Proof.

$$\Pr_{H_0^-(w_-)} \Pr_{H_0^-(w)} H_0^+(y) = \Pr_{H_0^-(w_-)} X,$$

and also it is equal to

$$\Pr_{H_0^-(w_-)} H_0^+(y) = X_-.$$

So

$$\Pr_{H_0^-(w_-)} x(0) = x_-(0).$$

This implies that

$$P \geq P_-.$$

If the stationary process y has the representation defined in equation 3.49 then the elements of this process can be written as follows

$$y(t) = Dw(t) + \sum_{k=1}^{\infty} CA^{k-1}Bw(t-k) \quad (3.53)$$

Since the process w is a white noise, the coefficients in this infinite series representation are the covariances between $y(t)$ and $w(t-k)$. Considering these elements for $k \geq 1$ we obtain that

$$\text{cov}(y(t), w(t-k)) = CA^{k-1}B.$$

Observe that this factorization is similar to the factorization 3.39 of the covariance function of y . Arguing similarly as in Theorem 3.17 we can form another Hankel-matrix $\mathcal{M}$ containing these covariances denoted by $M(k) = \text{cov}(y(t), w(t-k))$ – called as Markov parameters – and construct a factorization similar to 3.40. In fact,

$$\begin{aligned} \mathcal{M}_{k,m} &= \begin{bmatrix} M(1) & M(2) & \dots & M(m) \\ M(2) & & & \\ \vdots & & & \\ M(k) & & & M(m+k-1) \end{bmatrix} \\ &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} [B \ AB \ \dots \ A^{m-1}B]. \end{aligned} \quad (3.54)$$

The first factor is $\mathcal{O}_k(C, A)$, while the second factor is $C_m(A, B)$.

An immediate geometric interpretation can be given to these matrices. Concerning the observability matrix $\mathcal{O}(C, A)$ the elements of it can be obtained via projecting $y(t)$ to the state-space X , for $t \geq 0$.

$$\Pr_X y(t) = CA^t x(0) \quad t \geq 0.$$

We can look at this connection in another way:

$$y(t) = CA^t x(0) + \sum_{j=1}^t CA^{j-1} Bw(t-j) + Dw(t).$$

Loosely speaking – eliminating the terms containing the noise elements $w(0), w(1), \dots, w(t)$ from the right hand side the remaining part of the output process is obtained from the state $x(0)$ using the elements of the *observability* matrix. In the present case this “elimination” can be performed via the orthogonal projection onto X . At the same time we can consider the behaviour of the process y under the condition that the value of the future noises $w(0), w(1), \dots, w(t)$ are given, for example they are zero.

Consider now the second factor $C_m(A, B)$. The formula

$$x(0) = \sum_{j=1}^{\infty} A^{j-1} Bw(-j)$$

shows that the matrices $A^{j-1}B$ $j \geq 1$ determine the operator giving $x(0)$ from the past values of the noise: $w(-1), w(-2), \dots$. Formulating in another way, the expression

$$x(0) = A^k x(t-k) + \sum_{j=1}^k A^{j-1} Bw(-j)$$

shows that under the condition $x(t-k) = 0$ – considering as an initial condition – the elements of $C_k(A, B)$ determine the behaviour of $x(0)$. The condition $\text{rank } C_k(A, B) = n$ for large enough k is equivalent to the *reachability* of the system 3.49.

If the state transition matrix A is regular then the equations in 3.49 can be written as follows:

$$\begin{aligned} x(t) &= A^{-1}x(t+1) - A^{-1}Bw(t) \\ y(t) &= CA^{-1}x(t+1) - CA^{-1}Bw(t) + Dw(t). \end{aligned} \quad (3.55)$$

Iterating these expressions starting from the value $t = 0$

$$\begin{aligned} x(0) &= A^{-t}x(t) - \sum_{j=0}^{t-1} A^{-j-1}Bw(j) \\ y(-t) &= CA^{-t}x(0) - \sum_{j=1}^t CA^{-t+j-1}Bw(-j) + Dw(-t) \end{aligned} \quad (3.56)$$

for $t \geq 1$. Focusing on the second equation we can observe that “eliminating” again the noise terms from the right hand side the remaining part of the output process y for the time points from the “past” can be obtained from the state vector $x(0)$ using the *constructibility* matrix $[CA^{-1}, CA^{-2}, \dots, CA^{-t}]$. Let us point out that in the present case this “elimination” has no longer the interpretation

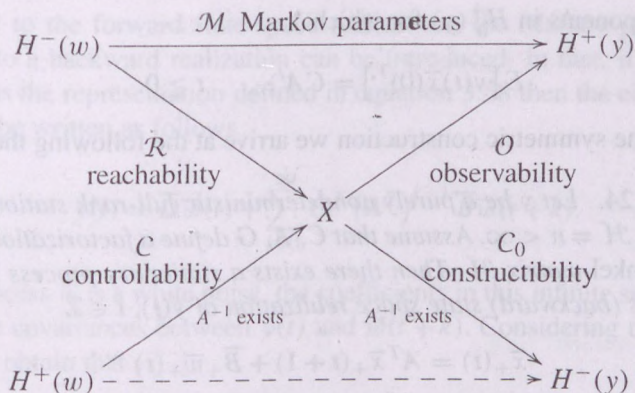


Figure 3.4. Factorization corresponding to the *reachability*, *observability*, *controllability* and *constructibility* operators. The line from $H^+(w)$ to $H^-(y)$ is dashed because the connection between the future of the input noise and the past of the output process which would be valid for deterministic systems is only informal for the present stochastic case

as a projection, but it can still be interpreted as “conditioning” on the values of the noise $w(-1), w(-2), \dots$. This “conditioning” argument is borrowed from the deterministic system theory, where the value of the input is not determined by some stochastic law but it can be arbitrarily controlled. Although there is – in some sense – similar notion in the probability theory: conditional probability, conditional distribution; but we should warn the reader that in the present case this should be handled with care, because $x(0)$ and $w(-1), w(-2), \dots, w(-t)$ are not stochastically independent, thus the “conditioning” on these values of the noise process changes the stochastic behaviour of $x(0)$. This problem did not arise in the previous cases when we have factorized the Markov-parameters, because in those examples the conditions were (stochastically) independent from the remaining elements.

Applying a similar argument for the first equation in 3.56, that is “conditioning” on $x(t) = 0$ for a value $t \geq 1$ we obtain that the value of the state vector $x(0)$ is obtained from the “future” values of the input noise w using the matrix $[A^{-1}B, A^{-2}B, \dots, A^{-t}B]$. (Here, again, the condition is (stochastically) not independent from the noise values appearing in the formula.)

Observe, that the controllability of the system is connected to the fact that the rank of this matrix is n – the dimension of the state-space – for large enough t .

Figure 3.4 summarizes these observations.

In our previous construction after the factorization of the *Hankel*-matrix we have used $C(A, G)$ to produce the state vector process $x_-(t), t \in \mathbb{Z}$. Symmetrically, based on the matrix $[C, CA, \dots]^T$, we can state the problem of finding a vector

$\bar{x}(0)$ with components in $H_0^+(y)$ for which

$$E[y(t)\bar{x}(0)^T] = CA^t, \quad t \geq 0.$$

Carrying out the symmetric construction we arrive at the following theorem.

THEOREM 3.24. *Let y be a purely nondeterministic full-rank stationary process for which $\text{rank } \mathcal{H} = n < \infty$. Assume that C, A, G define a factorization of minimal size of the Hankel-matrix $\mathcal{H}$. Then there exists a stationary process $\bar{x}_+(t)$, $t \in \mathbb{Z}$ which provides (backward) state space realization of $y(t)$, $t \in \mathbb{Z}$*

$$\begin{aligned} \bar{x}_+(t) &= A^T \bar{x}_+(t+1) + \bar{B}_+ \bar{w}_+(t) \\ y(t) &= G^T \bar{x}_+(t+1) + \bar{D}_+ \bar{w}_+(t) \end{aligned} \quad (3.57)$$

such that the coordinate vectors of $\bar{x}_+(0)$ are in H^+ .

Denote $\bar{P}_+ = E[\bar{x}_+(t)\bar{x}_+(t)^T]$.

REMARK 3.25. $\bar{w}_+(t)$, $t \in \mathbb{Z}$ is the backward innovation process.

The components of $\bar{w}_+(t)$ form an orthonormal basis in

$$H_t^+(y) \ominus H_{t+1}^+(y).$$

The subspace $H_0^+(y) \ominus H_1^+(y)$ corresponding to the case $t = 0$ is the so-called backward innovation subspace. Similarly to the equations 3.50, 3.51 and Theorem 3.23 we can formulate the following theorem for the backward case.

THEOREM 3.26. *If the process $\bar{x}(t)$, $t \in \mathbb{Z}$, determines a minimal backward state-space realization*

$$\begin{aligned} \bar{x}(t) &= A^T \bar{x}(t+1) + \bar{B} \bar{w}(t) \\ y(t) &= G^T \bar{x}(t+1) + \bar{D} \bar{w}(t) \end{aligned} \quad (3.58)$$

then

$$H_t^+(\bar{w}) \supset H_t^+(y) = H_t^+(\bar{w}_+). \quad (3.59)$$

Denote $\bar{X}$ the subspace spanned by the components of $\bar{x}(0)$. We have that

$$\Pr_{H_0^+(\bar{w})} H_0^-(y) = \bar{X}. \quad (3.60)$$

Also

$$\bar{P} \geq \bar{P}_+, \quad (3.61)$$

where

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

Similarly to the forward state-space realizations the Markov parameters corresponding to a backward realization can be introduced. In fact, if the stationary process y has the representation defined in equation 3.58 then the elements of this process can be written as follows

$$y(t) = \overline{D}\overline{w}(t) + \sum_{k=1}^{\infty} G^T (A^T)^{k-1} \overline{B}\overline{w}(t+k). \quad (3.62)$$

Since the process $\overline{w}$ is a white noise, the coefficients in this infinite series representation are the covariances between $y(t)$ and $\overline{w}(t+k)$. Considering these elements for $k \geq 1$ we obtain that

$$\text{cov}(y(t), \overline{w}(t+k)) = G^T (A^T)^{k-1} \overline{B}.$$

Now we can form the Hankel-matrix $\mathcal{M}^b$ containing these covariances denoted by $M^b(k) = \text{cov}(y(t), \overline{w}(t+k))$ – called as backward Markov parameters.

$$\begin{aligned} \mathcal{M}_{k,m}^b &= \begin{bmatrix} M^b(1) & M^b(2) & \dots & M^b(m) \\ M^b(2) & & & \\ \vdots & & & \\ M^b(k) & & & M^b(m+k-1) \end{bmatrix} \\ &= \begin{bmatrix} G^T \\ G^T A^T \\ \vdots \\ G^T (A^T)^{k-1} \end{bmatrix} \begin{bmatrix} \overline{B} & A^T \overline{B} & \dots & (A^T)^{m-1} \overline{B} \end{bmatrix}. \end{aligned} \quad (3.63)$$

The first factor is $\mathcal{O}_k^b(G^T, A^T)$, while the second factor is $C_m^b(A^T, \overline{B})$.

The elements of $\mathcal{O}^b(G^T, A^T)$ can be obtained via projecting $y(-t)$ to the state-space X , for $t \geq 1$

$$\text{Pr}_X y(-t) = G^T (A^T)^{t-1} x(0), \quad t \geq 1.$$

Formulating in the other way:

$$y(-t) = G^T (A^T)^{t-1} x(0) + \sum_{j=1}^{t-1} G^T (A^T)^{j-1} \overline{B}\overline{w}(-t+j) + \overline{D}\overline{w}(-t).$$

The elimination of the noise terms from the right hand side can be performed via the orthogonal projection onto X , or considering the conditional behaviour of the past of y with respect to the fixed past values of the noise $\overline{w}$. The remaining part of the output process is determined by the (backward) observability matrix $\mathcal{O}^b(G^T, A^T)$.

Concerning the second factor $C_m^b(A^T, \bar{B})$, the formula

$$\bar{x}(0) = \sum_{j=0}^{\infty} (A^T)^j \bar{B} \bar{w}(j)$$

shows that the matrices $(A^T)^j \bar{B}$ $j \geq 0$ determine the operator giving $\bar{x}(0)$ from the future values of the noise: $\bar{w}(0), \bar{w}(1), \dots$. The condition $\text{rank } C_k^b(A^T, \bar{B}) = n$ for large enough k is equivalent to the *reachability* of the system 3.58.

If the state transition matrix A^T is regular then the equations in 3.58 can be written in a “forward” form:

$$\begin{aligned} \bar{x}(t+1) &= (A^T)^{-1} \bar{x}(t) - (A^T)^{-1} \bar{B} \bar{w}(t) \\ y(t) &= G^T (A^T)^{-1} \bar{x}(t) - G^T (A^T)^{-1} \bar{B} \bar{w}(t) + \bar{D} \bar{w}(t). \end{aligned} \quad (3.64)$$

Iterating these expressions

$$\begin{aligned} \bar{x}(0) &= (A^T)^{-1} \bar{x}(-1) - \sum_{j=1}^{\infty} (A^T)^{-j} \bar{B} \bar{w}(-j) \\ y(t) &= G^T (A^T)^{-t-1} \bar{x}(0) - \sum_{j=0}^{\infty} G^T (A^T)^{-t+j-1} \bar{B} \bar{w}(j) + \bar{D} \bar{w}(t) \end{aligned} \quad (3.65)$$

for $t \geq 0$. Thus “eliminating” the noise terms from the right hand side of the second equation the remaining part of the output process y for the “future” time points can be obtained from the state vector $\bar{x}(0)$ using the *constructibility* matrix $[G^T (A^T)^{-1}, G^T (A^T)^{-2}, \dots, G^T (A^T)^{-t}]$. This “elimination” can be interpreted as “conditioning” on the values of the noise $\bar{w}(0), \bar{w}(1), \dots$.

From the first equation in 3.65, we obtain that “conditioning” on $\bar{x}(-t) = 0$ for a value $t \geq 0$ the value of the state vector $\bar{x}(0)$ is determined by the “past” values of the input (backward) noise $\bar{w}$ using the matrix $[(A^T)^{-1} \bar{B}, (A^T)^{-2} \bar{B}, \dots, (A^T)^{-t} \bar{B}]$.

Figure 3.5 summarizes these observations.

Let us consider briefly the question of uniqueness of the processes x_- and $\bar{x}_+$. The first step in the construction above is the factorization of the Hankel-matrix. We have already pointed out that – under the condition of minimality – the only nonuniqueness is a nonsingular transformation T , i.e. if C, A, G determine a factorization of the covariance sequence $R(k), k \in \mathbb{Z}$ (where A is of minimal size) then all possible factorizations (of minimal size) are written in the form $CT, T^{-1}AT, T^{-1}G$, where T is an arbitrary nonsingular matrix. The coordinates of $x_-(0)$ are determined by continuous linear functionals defined on the closed subspace H^- . Since the matrices C, A define these functionals on the generator system $y_j(t), t \leq -1$ in a unique way the corresponding extensions and consequently the state vector $x_-(0)$ are already uniquely determined. Similarly, choosing the matrices A, G the state vector (for a backward realization) $\bar{x}_+(0)$ is uniquely determined. Consequently, the forward (backward) innovation realizations are determined uniquely up to a

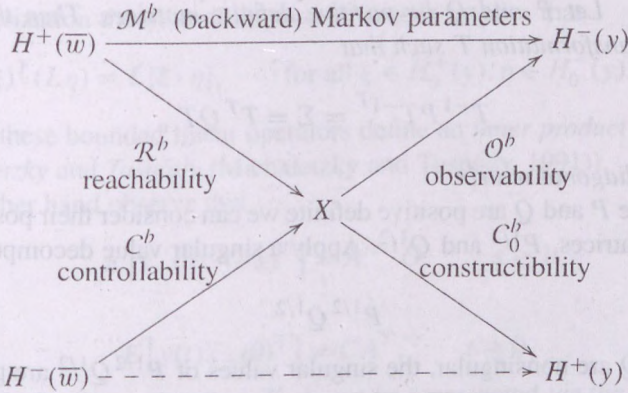


Figure 3.5. Factorization corresponding to the *reachability*, *observability*, *controllability* and *constructibility* operators (backward system). The line from $H^-(\bar{w})$ to $H^+(y)$ is dashed because the connection between the past of the input (backward) noise and the future of the output process which would be valid for deterministic systems is only informal for the present stochastic case

nonsingular transformation T . Observe, that if instead of the matrices C, A, G we use $CT, T^{-1}AT, T^{-1}G$ then the state vector $x_-(0)$ is mapped into $T^{-1}x_-(0)$, while $\bar{x}_+(0)$ into $T^T\bar{x}_+(0)$. The transformation on the corresponding covariance matrices obviously given by the equations

$$P_- \mapsto T^{-1}P_-T^{-1T}, \quad (3.66)$$

$$\bar{P}_+ \mapsto T^T\bar{P}_+T. \quad (3.67)$$

In the next chapter we consider several possibilities of defining the transformation T or more precisely choosing the matrices C, A, G in a “canonical” way. The underlining idea in those constructions is somehow specifying the structure of these system matrices. But equation 3.66 gives the possibility of reducing the state covariance matrices. These considerations lead to the *stochastic balanced* realization introduced by *Desai and Pal* (Desai and Pal, 1982; Desai and Pal, 1984).

DEFINITION 3.27. *The forward and the corresponding backward realizations determined by the state vector processes $x_-(t)$, $t \in \mathbb{Z}$ and $\bar{x}_+(t)$, $t \in \mathbb{Z}$ (based on the same factorization of the Hankel matrix) are stochastically balanced if their covariance matrices are diagonal and equal, i.e.*

$$\Sigma = P_- = \bar{P}_+,$$

where Σ is a diagonal matrix.

LEMMA 3.28. *Let P and Q be positive definite matrices. Then there exists a nonsingular transformation T such that*

$$T^{-1}PT^{-1T} = \Sigma = T^TQT, \quad (3.68)$$

where Σ is a diagonal matrix.

Proof. Since P and Q are positive definite we can consider their positive definite square-root matrices, $P^{1/2}$ and $Q^{1/2}$. Apply a singular value decomposition to the matrix

$$P^{1/2}Q^{1/2}.$$

Since P and Q are nonsingular, the singular values of $P^{1/2}Q^{1/2}$ are positive.

$$P^{1/2}Q^{1/2} = U\Sigma V,$$

where $UU^T = VV^T = I$. Set $T = \Sigma^{1/2}VQ^{-1/2}$.

Obviously, $T = \Sigma^{-1/2}U^TP^{1/2}$. Consequently,

$$T^{-1T}PT^{-1} = \Sigma^{1/2}U^TP^{-1/2}PP^{-1/2}U\Sigma^{1/2} = \Sigma, \quad (3.69)$$

$$TQT^T = \Sigma^{1/2}VQ^{-1/2}QQ^{-1/2}V^T\Sigma^{1/2} = \Sigma. \quad (3.70)$$

Applying this argument to the positive definite matrices P_- and $\bar{P}_+$ we obtain that in an appropriate coordinate system the covariance matrices P_- and $\bar{P}_+$ are equal and diagonal.

Let us point out that this common diagonal matrix is defined via the singular values of the product $P_-^{1/2}\bar{P}_+^{1/2}$ or equivalently the square-roots of the eigenvalues of the product $P_- \bar{P}_+$.

Let us return to the factorization of the *Hankel*-matrix for a short time. Every element in the *Hankel*-matrix is a covariance, in other words it is a scalar product of two vectors in $H(y)$, one of which is in $H_0^+(y)$, the other one is in $H_0^-(y)$.

The factorization produces this number as a scalar product of two vectors in $\mathbb{R}^n$ – the corresponding row vector in $\mathcal{O}(A, C)$ and the corresponding column vector in $\mathcal{C}(A, G)$.

We define two continuous linear operators

$$K : H_0^+(y) \rightarrow \mathbb{R}^n, \quad L : H_0^-(y) \rightarrow \mathbb{R}^n$$

via extending the mappings given on generator systems

$$\begin{aligned} K : y_j(t) &\rightarrow j\text{-th row of } CA^t, & t \geq 0 \\ L : y_j(-s) &\rightarrow j\text{-th column of } A^{s-1}G, & s \geq 1. \end{aligned}$$

The factorization theorem of the *Hankel*-matrix implies that

$$(K\xi)^T(L\eta) = E[\xi \cdot \eta], \quad \text{for all } \xi \in H_0^+(y), \eta \in H_0^-(y). \quad (3.71)$$

We say that these bounded linear operators define an *inner product representation* (see *Michaletsky and Tusnády* (Michaletsky and Tusnády, 1991)).

On the other hand observe that

$$E[x_-(0)y(-s)^T] = A^{s-1}G, \quad s \geq 1$$

and

$$E[y(t)\bar{x}_+(0)^T] = CA^T, \quad t \geq 0.$$

In other words the linear operators K, L can be represented via the vectors $x_-(0), \bar{x}_+(0)$. We say that the components of $x_-(0)$ and $\bar{x}_+(0)$ form a *representative vector system* for the inner product representation K, L .

Especially

$$E[\xi \cdot \eta] = E[\xi \bar{x}_+(0)^T] E[x_-(0)\eta], \quad \text{for all } \xi \in H_0^+(y), \eta \in H_0^-(y). \quad (3.72)$$

(This is nothing else then the factorization of the *Hankel*-matrix.)

But the components of $\bar{x}_+(0)$ and $x_-(0)$ are in $H_0^+(y)$ and $H_0^-(y)$, respectively. Applying the identity 3.72 to these vectors we get that

$$E[\bar{x}_+(0)x_-(0)^T] = E[\bar{x}_+(0)\bar{x}_+(0)^T] E[x_-(0)x_-(0)^T] = \bar{P}_+P_-. \quad (3.73)$$

Immediate consequence of this equation that using stochastically balanced realizations the covariance matrix between $x_-(0)$ and $\bar{x}_+(0)$ becomes diagonal. Normalizing the covariance matrices of $x_-(0)$ and $\bar{x}_+(0)$ we obtain that

$$\text{cov}(P_-^{-1/2}x_-(0), \bar{P}_+^{-1/2}\bar{x}_+(0)) = \Sigma \quad (3.74)$$

Since the coordinates of $x_-(0)$ span the subspace $X_- = \text{Pr}_{H^-} H^+$, the coordinates of $\bar{x}_+(0)$ span the subspace $X_+ = \text{Pr}_{H^+} H^-$ and in a stochastically balanced realization the component random variables of $x_-(0)$ and $\bar{x}_+(0)$ are biorthogonal to each other, the diagonal elements of the matrix Σ are the so-called (non-zero) *canonical correlations* between H^- and H^+ , or – in other words – they are the cosines of the angles between the subspaces H^- and H^+ .

We may apply the inner product representation identity to the elements $\bar{w}_+(t)$, $t \geq 0$ and $w_-(s)$, $s < 0$. The components of these vectors are in H^+ and H^- , respectively. Thus we obtain the following identity

$$E[\bar{w}_+(t)w_-(-s)^T] = E[\bar{w}_+(t)\bar{x}_+(0)^T] E[x_-(0)w_-(-s)^T] \\ t \geq 0, \quad s \geq 1.$$

But

$$\begin{aligned} E[x_-(0)w_-(-s)^T] &= A^{s-1}B, & s \geq 1 \\ E[\bar{x}_+(0)\bar{w}_+(t)^T] &= (A^T)^t \bar{B}, & t \geq 0. \end{aligned}$$

Consequently

$$E[\bar{w}_+(t)w_-(-s)^T] = \bar{B}^T A^{t+s-1}B. \quad (3.75)$$

This implies that

$$\Pr_H \bar{w}_+(t) = \sum_{s=1}^{\infty} \bar{B}^T A^{t+s-1}B w_-(-s).$$

Another way of looking at this identity is the following. Since the components of $w_-(s)$, $s \in \mathbb{Z}$, form an orthonormal basis in H_0 , $\bar{w}_+(t)$ can be written as a linear transform of the forward innovation process:

$$\bar{w}_+(t) = \sum_{s=-\infty}^{+\infty} g(t-s)w_-(s) = \sum_{s=-\infty}^{+\infty} g(s)w_-(t-s).$$

Denote the corresponding – all-pass – transfer function by

$$\Phi(z) = \sum_{k=-\infty}^{+\infty} g(k)z^k. \quad (3.76)$$

The previous identity gives the values of $g(k) = \bar{B}^T A^{k-1}B$, $k \geq 1$.

In other words the so-called strictly stable part of Φ (i.e. the part containing the terms with positive power in z) is

$$\sum_{k=1}^{\infty} \bar{B}^T A^{k-1}B z^k = \bar{B}^T (z^{-1}I - A)^{-1}B. \quad (3.77)$$

DEFINITION 3.29. The function $\Phi(z)$ defined in equation 3.76 is the so-called phase function of the stationary process y .

Formulating in another way the phase function $\Phi(z)$ is an *all-pass extension* of the function $\bar{B}^T (z^{-1}I - A)^{-1}B$. This phase function maps the forward innovation process to the backward one as shown on the Figure 3.6.

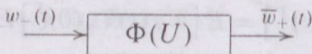


Figure 3.6. The filter relating the forward and backward innovations

4. System invariants and canonical forms in SSR, ARMA and MFD representations

For any fixed input noise $w(t)$, $t \in \mathbb{Z}$ the matrices A , B , C , D in the state-space description of the stationary process y overparametrize the corresponding transfer function. In this chapter we investigate the so-called canonical forms of state-space models, together with ARMA and MFD models and their dual forms, as well. These canonical forms are based on system invariants. Together with the parametrization we point out the special structure of the corresponding state vector process.

First we start with some general definitions about invariant functions.

DEFINITION 4.1. Let us consider a set $\mathcal{R}$ and an equivalence relation $\sim$ on $\mathcal{R}$. Let S be another set. We call the function $f : \mathcal{R} \rightarrow S$ the invariant of the equivalence relation $\sim$ if

$$x_1 \sim x_2 \implies f(x_1) = f(x_2)$$

holds. We call the function $f : \mathcal{R} \rightarrow S$ the complete invariant of the equivalence relation if

$$x_1 \sim x_2 \iff f(x_1) = f(x_2)$$

holds.

DEFINITION 4.2. Let $S_1, \dots, S_k$ be sets. We call the functions $f_1, \dots, f_k$; $f_i : \mathcal{R} \rightarrow S_i$ the complete invariant system of the equivalence relation $\sim$ if the function $f = (f_1, \dots, f_k) : \mathcal{R} \rightarrow S_1 \times \dots \times S_k$ is a complete invariant of the equivalence relation $\sim$.

DEFINITION 4.3. Invariants $f_1, \dots, f_k$ are independent if the invariant $f = (f_1, \dots, f_k)$ is surjective.

According to the general terminology in the literature we term invariant indices if the image of invariants is integer and invariant parameters if this image is real number.

A nonsingular transformation in the state-space has no effect on the transfer function. In order to find a minimal parametrization first analyse the behaviour of the observability matrix under the nonsingular state-space transformations. Let us consider the set of observable pairs. Denote it by

$$\mathcal{R} = \{(A, C) : A \in \mathbb{R}^{n \times n}; C \in \mathbb{R}^{p \times n}; (A, C) \text{ observable}\}$$

and let us define the relation $\sim$ as follows.

$(A, C) \sim (\tilde{A}, \tilde{C})$ if and only if there exist a $T \in \mathbb{R}^{n \times n}$, $\det T \neq 0$, such that $\tilde{A}T = TA$ and $\tilde{C} = CT$. We are looking for invariant functions defined on the set $\mathcal{R}$ in the sense of Definition 4.1.

4.1. Observability invariants

If the pair (A, C) is observable, then the observability matrix $\mathcal{O}_n(A, C)$ is of full rank. Let us denote the rows of C by $c_1^T, \dots, c_p^T$, i.e.

$$C = \begin{bmatrix} c_1^T \\ \vdots \\ c_p^T \end{bmatrix}.$$

Consider the rows of the observability matrix in the following order:

$$\begin{array}{cccc} c_1^T & \dots & c_p^T \\ c_1^T A & \dots & c_p^T A \\ c_1^T A^2 & \dots & c_p^T A^2 \\ \vdots & & \end{array}$$

and select those vectors from left to the right and row by row, which are linearly independent from the set of previously selected vectors. Hence $\text{rank } \mathcal{O}_n(A, C) = n$, in such a way we select a basis in $\mathbb{R}^n$.

DEFINITION 4.4. (See Kailath (Kailath, 1980)) *The above mentioned (basis) selection rule is called Scheme II selection plan.*

Denote v_i , $i = 1, \dots, p$, the smallest nonnegative integer for which the vector $c_i^T A^{v_i}$ already depends on the earlier chosen linearly independent vectors in the Scheme II selection plan. The numbers v_i , $i = 1, \dots, p$, are called left-hand side Kronecker indices (Popov (Popov, 1972)) or observability indices (Kailath (Kailath, 1980)). Obviously

$$v_1 + \dots + v_p = n$$

holds. Consequently for appropriate coefficients α_{ijk} we have that

$$c_i^T A^{v_i} = \sum_{j=1}^p \sum_{k=0}^{v_j} \alpha_{ijk} c_j^T A^k, \quad (4.1)$$

where

$$v_{ij} = \begin{cases} \min(v_i, v_j) - 1, & \text{if } i \leq j \\ \min(v_i, v_j - 1), & \text{if } i > j. \end{cases}$$

Let us denote by

$$\nu(A, C) = \{v_i; i = 1, \dots, p\}$$

the set of indices and by

$$\alpha(A, C) = \{\alpha_{ijk} : i = 1, \dots, p; j = 1, \dots, p; k = 0, \dots, v_{ij}\}$$

the set of coefficients included in dependency relation 4.1.

One can realize that ν, α remain unchanged if we consider the pair $(\tilde{A}, \tilde{C})$ instead of the pair (A, C) , i.e.

$$(A, C) \sim (\tilde{A}, \tilde{C}) \implies \{\nu(A, C), \alpha(A, C)\} = \{\nu(\tilde{A}, \tilde{C}), \alpha(\tilde{A}, \tilde{C})\},$$

thus Scheme II selection plan is the invariant of the equivalence relation $\sim$ on the set $\mathcal{R}$. It can be proved (cf. *Popov* (Popov, 1972), Proposition 5–6, Theorem 1), that the functions ν, α determine a complete and independent set of invariants.

In other words, the numbers $\nu(A, C), \alpha(A, C)$ determine the equivalence class in $\mathcal{R}$ which (A, C) belongs to.

Since the selected vectors form a basis in $\mathbb{R}^n$ the linear transformation mapping the elements of the original basis into this new one is nonsingular so we can apply this as a basis transformation. In Section 4.3 we return to the observability invariants and derive the corresponding canonical form together with a state-space representation. But first let us consider another type of invariants.

4.2. Constructibility invariants and canonical state-space representations

Assume that the $n \times n$ matrix A is nonsingular and the $p \times n$ matrix C is of rank p . We can define the so-called constructibility matrix C_0 as

$$C_0 = \begin{bmatrix} CA^{-1} \\ CA^{-2} \\ \vdots \\ CA^{-n} \end{bmatrix}$$

and assume that $\text{rank } C_0 = n$, in other words the pair (A, C) is constructible. Then one can search for a new, unique basis in the state-space among the rows of the

constructibility matrix C_0 in the following order:

$$\begin{array}{c} c_1^T A^{-1} \dots c_p^T A^{-1} \\ c_1^T A^{-2} \dots c_p^T A^{-2} \\ c_1^T A^{-3} \dots c_p^T A^{-3} \\ \vdots \end{array}$$

and select the first linearly independent vectors from the left to the right, and row by row exactly the same way as described in the Scheme II selection plan in Kailath (Kailath, 1980).

DEFINITION 4.5. For $1 \leq i \leq p$ the i -th constructibility index μ_i is defined as the smallest positive integer such that $c_i^T A^{-\mu_i-1}$ is a linear combination of its antecedents.

It can be easily verified that the constructibility indices satisfy the relation

$$\mu_1 + \dots + \mu_p = n.$$

An immediate consequence of Definition 4.5 is the following proposition.

PROPOSITION 4.6. There exist unique numbers β_{ijk} such that the following dependency relations hold:

$$c_i^T A^{-\mu_i-1} = \sum_{j=1}^p \sum_{k=0}^{\mu_{ij}} \beta_{ijk} c_j^T A^{-k-1}, \quad (4.2)$$

where

$$\mu_{ij} = \begin{cases} \min(\mu_i, \mu_j) - 1, & \text{if } i \leq j \\ \min(\mu_i, \mu_j - 1), & \text{if } i > j. \end{cases}$$

PROPOSITION 4.7. The numbers $\{\beta_{ijk}\}$, $\{\mu_i\}$ remain unchanged under change of basis in the state space.

Proof. Let T be a non-singular matrix of the new basis vectors. Consider the matrices $\tilde{C} = CT$, $\tilde{A} = T^{-1}AT$, so $\tilde{A}^{-1} = T^{-1}A^{-1}T$. The new constructibility matrix has the block entries $\tilde{C}\tilde{A}^{-k} = CA^{-k}T$. Thus the vectors in the dependency relations 4.2 are all multiplied by T from the right. Since T is non-singular, the numbers β_{ijk} , μ_i remain unchanged. ■

REMARK 4.8. In view of Definition 4.2 the statement of Proposition 4.7 is that the numbers $\{\beta_{ijk}\}$, $\{\mu_i\}$ are invariants for the equivalence class $\tilde{C} = CT$, $\tilde{A} = T^{-1}AT$ with nonsingular A and T .

PROPOSITION 4.9. *The numbers $\{\beta_{ijk}\}$, $\{\mu_i\}$ constitute a complete system of independent invariants for the equivalence class of observable pairs (C, A^{-1}) .*

Proof. We show that any fixed pair of matrices (A, C) , where C is of full row rank can be transformed using a nonsingular basis transformation into matrices of special structure the elements of which are functions of the set of invariants $\{\beta_{ijk}\}$, $\{\mu_i\}$.

Since we have assumed that C is of full rank we can suppose that $\mu_i \geq 1$, $i = 1, \dots, p$.

In order to examine the structure of the state-space representation in the new basis form the following matrix of the previously selected linearly independent vectors:

$$T = \begin{bmatrix} c_1^T A^{-1} \\ \vdots \\ c_1^T A^{-\mu_1} \\ \vdots \\ c_p^T A^{-1} \\ \vdots \\ c_p^T A^{-\mu_p} \end{bmatrix}.$$

Calculate $\tilde{C} = CT^{-1}$, $\tilde{A} = TAT^{-1}$. It can easily be verified that if the elements of $\tilde{C}$ are denoted in the following way

$$\tilde{C} = \left[\begin{array}{cccc|cccc} \tilde{c}_{110} & \tilde{c}_{111} & \dots & \tilde{c}_{11\mu_1-1} & \dots & \tilde{c}_{1p0} & \tilde{c}_{1p1} & \dots & \tilde{c}_{1p\mu_p-1} \\ \vdots & \vdots & & \vdots & \dots & \vdots & \vdots & & \vdots \\ \tilde{c}_{p10} & \tilde{c}_{p11} & \dots & \tilde{c}_{p1\mu_p-1} & \dots & \tilde{c}_{pp0} & \tilde{c}_{pp1} & \dots & \tilde{c}_{pp\mu_p-1} \end{array} \right] \quad (4.3)$$

and observing that the vector $c_j^T A^{-k} T^{-1}$ coincides with $\tilde{c}_j^T$, if $k = 0$, and it has only one nonvanishing element, if $k = 1, \dots, \mu_j$, with value 1 at the position $\sum_{i=1}^{j-1} \mu_i + k$ the matrix $\tilde{A}$ is of the form

$$\tilde{A} = \left[\begin{array}{cccc|cccc} \tilde{c}_{110} & \tilde{c}_{111} & \dots & \tilde{c}_{11\mu_1-1} & \dots & \tilde{c}_{1p0} & \tilde{c}_{1p1} & \dots & \tilde{c}_{1p\mu_p-1} \\ 1 & 0 & 0 & 0 & \dots & & & & \\ \vdots & \ddots & & \vdots & \dots & & & & \\ 0 & 0 & 1 & 0 & \dots & & 0 & & \\ \hline & & & & \ddots & & & & \\ \tilde{c}_{p10} & \tilde{c}_{p11} & \dots & \tilde{c}_{p1\mu_p-1} & \dots & \tilde{c}_{pp0} & \tilde{c}_{pp1} & \dots & \tilde{c}_{pp\mu_p-1} \\ & & & & \dots & 1 & 0 & 0 & 0 \\ & & & & & \vdots & \ddots & & \vdots \\ & & & & & 0 & 0 & 1 & 0 \end{array} \right] \quad (4.4)$$

Thus $\tilde{A}$ is in a "top-companion" form and the non-trivial entries of $\tilde{A}$ are the entries of $\tilde{C} = CT^{-1}$. To show that the invariants $\{\mu_i\}$ and $\{\beta_{ijk}\}$ determine the elements of these matrices multiply equation 4.2 by AT^{-1} from the right and collect each term containing $\beta_{ij0}\tilde{c}_i^T$ to the right hand side of the equation. We obtain the following formula:

$$-\tilde{A}_p S + I_{\mu_1 \dots \mu_p} = \tilde{A}_0 \tilde{C} \quad (4.5)$$

where the elements of $\tilde{A}_p$ and $\tilde{A}_0$ consist of the β_{ijk} coefficients:

$$\tilde{A}_p = \begin{bmatrix} \beta_{110} & \beta_{111} & \dots & \beta_{11\mu_1-1} & \dots & \beta_{1p0} & \beta_{1p1} & \dots & \beta_{1p\mu_p-1} \\ \vdots & \vdots & & \vdots & \dots & \vdots & \vdots & & \vdots \\ \beta_{p10} & \beta_{p11} & \dots & \beta_{p1\mu_1-1} & \dots & \beta_{pp0} & \beta_{pp1} & \dots & \beta_{pp\mu_p-1} \end{bmatrix}$$

$$\tilde{A}_0 = \begin{bmatrix} \beta_{110} & \dots & \beta_{1p0} \\ \vdots & & \vdots \\ \beta_{p10} & \dots & \beta_{pp0} \end{bmatrix}$$

(by definition $\beta_{ijk} = 0$ if $\mu_{ij} < k \leq \mu_j - 1$);

$$I_{\mu_1 \dots \mu_p} = \text{blockdiag} \{ [0 \dots 0 \ 1], \dots, [0 \dots 0 \ 1] \}$$

(the size of the i -th block is μ_i) and

$$S = \text{blockdiag} \{ S_1, \dots, S_p \},$$

where S_i is the $(\mu_i \times \mu_i)$ -dimensional *Toeplitz*-matrix

$$S_i = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix}.$$

The columns determined by the nonzero elements in $I_{\mu_1 \dots \mu_p}$ are linearly independent, consequently the rank of the matrix on the left hand side of 4.5 is p , thus $\tilde{A}_0$ must be invertible, so

$$\tilde{C} = \tilde{A}_0^{-1} (I_{\mu_1 \dots \mu_p} - \tilde{A}_p S). \quad (4.6)$$

This proves that $\tilde{A}$ and $\tilde{C}$ are determined by the constructibility invariants. ■

DEFINITION 4.10. The state-space representation given in top-companion form 4.3 and 4.4 is called *constructibility form*.

4.3. Stochastic state-space representations and observability invariants

Let us return to the observability invariants. To find the corresponding canonical form we investigate the state-space representations of the process $y(t)$, $t \in \mathbb{Z}$. The stationary process $y(t)$ can be written in the form

$$y(t) = \sum_{s=-\infty}^{t-1} g(t-s)w(s) = \sum_{s=0}^{\infty} g(s)w(t-s),$$

where $w(t)$ is a (forward) white noise process, or equivalently

$$\begin{aligned} x(t+1) &= Ax(t) + Bw(t), \\ y(t) &= Cx(t) + Dw(t). \end{aligned}$$

Observe that the state-space $X = \Pr_{H_0^-(w)} H_0^+(y)$ so it is generated by the projections of the coordinates of $y(t)$, $t \geq 0$ onto $H_0^-(w)$. The elements of the *Hankel*-matrix $\mathcal{O}(A, C)C(A, B)$ are the covariances between $y(t)$, $t \geq 0$ and $w(s)$, $s \leq -1$, the latter being equal to the covariances between $\Pr_{H_0^-(w)} y(t)$, $t \geq 0$, and $w(s)$, $s \leq -1$. Thus the linear dependency relations between the vectors $\Pr_{H_0^-(w)} y_j(t)$ coincide with that of the corresponding rows of the observability matrix. Immediate consequence of this observation is the following lemma.

LEMMA 4.11. *Assume that the stationary process $y(t)$ has a finite dimensional realization of minimum order $n = \text{rank}(\mathcal{H})$. Let us apply the Scheme II selection plan to the row vectors of the observability matrix $\mathcal{O}(A, C)$, defining the observability invariants $\{v_i\}$, $\{\alpha_{ijk}\}$. Then the vectors*

$$\Pr_{H_0^-(w)} y_i(j), \quad j = 0, \dots, v_i - 1; \quad i = 1, \dots, p \quad (4.7)$$

form a basis in X .

Therefore the basis selection strategy for state-space X can be related to the basis selection (in $\mathbb{R}^n$) from the rows of observability matrix $\mathcal{O}$.

REMARK 4.12. *Let us point out that if $v_i = 0$ for some i then the corresponding coordinate of y is not used in 4.7.*

PROPOSITION 4.13. *Let the basis vectors selected by Scheme II selection plan from the rows of the observability matrix $\mathcal{O}$ be arranged in the following chain:*

$$c_1^T, c_1^T A, \dots, c_1^T A^{v_1-1}, \dots, c_p^T, \dots, c_p^T A^{v_p-1}, \quad (4.8)$$

where $v_1, \dots, v_p$ are the so-called observability indices, (if $v_i = 0$ for some i then the corresponding row c_i^T of C is not used in 4.7) and define the process $x(t)$ as

follows,

$$x(t) = \left[\Pr_{H_t^-(w)} y_1(t), \Pr_{H_t^-(w)} y_1(t+1), \dots, \Pr_{H_t^-(w)} y_1(t+\nu_1-1), \dots, \Pr_{H_t^-(w)} y_p(t+\nu_p-1) \right]^T. \quad (4.9)$$

In this basis the process $y(t)$ has the following forward state-space representation:

$$\begin{aligned} x(t+1) &= A_o x(t) + B_o w(t), \\ y(t) &= C_o x(t) + D_o w(t), \end{aligned} \quad (4.10)$$

where the structure of the matrices A_o, C_o is

$$A_o = \left[\begin{array}{cccc|ccc|cccc} 0 & 1 & \dots & 0 & \dots & & & & & & \\ \vdots & \ddots & \ddots & \vdots & \dots & & & & & & 0 \\ 0 & \dots & 0 & 1 & \dots & & & & & & \\ \hline \alpha_{110} & \alpha_{111} & \dots & \alpha_{11\nu_1-1} & \dots & \alpha_{1p0} & \alpha_{1p1} & \dots & \alpha_{1p\nu_p-1} & & \\ \hline & & & & \ddots & & & & & & \\ \hline & & & & \dots & 0 & 1 & \dots & 0 & & \\ & & & & & \vdots & \ddots & \ddots & \vdots & & \\ & & 0 & & & \dots & \ddots & \ddots & \dots & & \\ & & & & & \dots & 0 & \dots & 0 & 1 & \\ \hline \alpha_{p10} & \alpha_{p11} & \dots & \alpha_{p1\nu_p-1} & \dots & \alpha_{pp0} & \alpha_{pp1} & \dots & \alpha_{pp\nu_p-1} & & \end{array} \right], \quad (4.11)$$

and

$$C_o = \left[\begin{array}{cccc|ccc|cccc} 1 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 & & \\ \vdots & \vdots & & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & & \\ 0 & 0 & \dots & 0 & \dots & 1 & 0 & \dots & 0 & & \end{array} \right], \quad (4.12)$$

if $\nu_i \geq 1$ for all $i = 1, \dots, p$ whereas if $\nu_i = 0$ for some i then the corresponding row of C_o is $[\alpha_{i10}, 0, \dots, 0, \alpha_{i20}, 0, \dots, 0, \dots, \alpha_{i, i-1, 0}, 0, \dots, 0]$. $\{\alpha_{ijk}\}$ are the observability invariants, and – by definition – $\alpha_{ijk} = 0$ if $\nu_{ij} < k \leq \nu_j - 1$, $j = 1, \dots, p$, $i = 1, \dots, p$.

Proof. For vectors, selected from rows of the observability matrix using the selection plan Scheme II, the dependency relations 4.1 holds. The dependency relations among the vectors $\Pr_{H_t^-(w)} y_i(t+k)$, $i = 1, \dots, p$, $k = 1, \dots, n-1$ and among the rows of the observability matrix having the form $c_i^T A^k$ coincide. Applying the identity

$$\Pr_{H_t^-(w)} \Pr_{H_{t+1}^-(w)} y_i(t+1+k) = \Pr_{H_t^-(w)} y_i(t+1+k)$$

for $k = 0, \dots, \nu_i - 2$, $i = 1, \dots, p$ and equation 4.1 for $k = \nu_i - 1$, we obtain that A_o has the form above.

The structure of the matrix C_o if $\nu_i \geq 1$ for any $i = 1, \dots, p$ is obtained by the direct use of the equation $y(t) = \Pr_{H_t^-(w)} y(t) + D w(t)$. Since all components

of $\Pr_{H_1^-(w)} y(t)$ are elements of the state vector $x(t)$, the rows of C_0 are the corresponding unit vectors. If $v_i = 0$ for some i then the corresponding row of C should be expressed as linear combination of the preceding rows, the coefficients are defined in equation 4.1.

REMARK 4.14. *The state-space representation given in Proposition 4.13 is called observability canonical form.*

Similar construction leads to the backward canonical forms.

4.4. Reachability invariants and canonical state-space realizations

Let us assume that the pair (A, B) is controllable, or in other words $\text{rank}[B, \dots, A^{n-1}B] = n$ and apply the Scheme II selection plan to the vectors

$$b_1, \dots, b_q, Ab_1, \dots, Ab_q, \dots$$

Denote $\lambda_i, i = 1, \dots, q$ the smallest nonnegative integer for which the vector $A^{\lambda_i} b_i$ already depends on the set of earlier chosen (linearly independent) vectors. The numbers $\lambda_1, \dots, \lambda_q, i = 1, \dots, q$ are the reachability indices. Obviously

$$\lambda_1 + \dots + \lambda_q = n,$$

and $\text{rank } B$ is equal to the number of nonzero reachability indices. Denote

$$\begin{aligned} I_B &= \{i | \lambda_i > 0\}, \\ q_B &= \text{rank } B. \end{aligned} \quad (4.13)$$

According to the definition

$$A^{\lambda_i} b_i = \sum_{j=1}^q \sum_{k=1}^{\lambda_{ij}} \gamma_{ijk} A^{k-1} b_j \quad (4.14)$$

for appropriate coefficients γ_{ijk} , where

$$\lambda_{ij} = \begin{cases} \min(\lambda_i, \lambda_j), & \text{if } i \leq j, \\ \min(\lambda_i + 1, \lambda_j), & \text{if } i > j. \end{cases} \quad (4.15)$$

Note that in the first summation instead of $j = 1, \dots, q$ we can take $j \in I_B$.

Set $\gamma_{ijk} = 0$, if $k > \lambda_{ij}$. Introduce the matrix

$$T = [b_1, \dots, A^{\lambda_1-1} b_1, \dots, b_q, \dots, A^{\lambda_q-1} b_q], \quad (4.16)$$

where b_i is not included, if $\lambda_i = 0$. Then

$$B_r = T^{-1}B = \text{blockdiag} \{ [1, 0, \dots, 0]^T, \dots, [1, 0, \dots, 0]^T \}, \quad (4.17)$$

where the size of the i -th block is λ_i , if $\lambda_i \geq 1$ for all $i = 1, \dots, q$, whereas if $\lambda_i = 0$ for some i then the corresponding column of B_r is

$$[\gamma_{i11}, 0, \dots, 0, \gamma_{i21}, 0, \dots, 0, \dots, \gamma_{i, i-1, 1}, 0, \dots, 0, 0, \dots, 0]^T,$$

i.e. this column is a linear combination of the previous columns in B_r ;

$$A_r = T^{-1}AT = [A_{ij}^{(r)}]_{i,j \in I_B}, \quad (4.18)$$

where

$$A_{ii}^{(r)} = \begin{bmatrix} 0 & \dots & 0 & \gamma_{ii1} \\ 1 & 0 & \dots & \gamma_{ii2} \\ & \ddots & & \vdots \\ 0 & & 1 & \gamma_{ii\lambda_i} \end{bmatrix}, \quad (4.19)$$

$$A_{ij}^{(r)} = \begin{bmatrix} 0 & \dots & 0 & \gamma_{ji1} \\ 0 & \dots & 0 & \gamma_{ji2} \\ \vdots & & \vdots & \\ 0 & \dots & 0 & \gamma_{ji\lambda_i} \end{bmatrix}. \quad (4.20)$$

(Equations 4.17 and 4.18–4.20 are immediate consequences of the identities $B = TB_r$ and $AT = TA_r$ which are obvious from the definition of the transformation T and 4.14.)

DEFINITION 4.15. *The state-space representation determined by the matrices A_r, B_r is called reachability canonical form.*

REMARK 4.16. *Since, for $l \geq 1$*

$$A^{l-1}B = \text{cov}(x(0), w(-l)) = \text{cov}(x(0), \text{Pr}_X w(-l)),$$

we obtain that the vectors

$$\text{Pr}_X w_1(-1), \dots, \text{Pr}_X w_1(-\lambda_1), \dots, \text{Pr}_X w_q(-1), \dots, \text{Pr}_X w_q(-\lambda_q) \quad (4.21)$$

are linearly independent and generate the subspace X . The transformation T defined in 4.16 transforms the matrix of selected columns of $C_n(A, B) = [B, AB, \dots, A^{n-1}B]$ into the unit matrix, consequently the coordinates of $T^{-1}x(0)$ – defining the state vector in the reachability canonical form – constitute a basis in X biorthogonal to the basis defined in 4.21.

4.5. Controllability invariants and canonical state-space realizations

Similarly to the constructible invariants and constructible canonical forms for a controllable pair (A, B) we can define the controllable invariants and controllable canonical forms assuming that the $n \times n$ matrix A is nonsingular and the $n \times q$ matrix B is of rank q . In this case the matrix $[A^{-1}B, \dots, A^{-n}B] = A^{-n}[A^{n-1}B, \dots, B]$ is of full rank so the Scheme II selection plan can be applied considering the column vectors in the following order

$$A^{-1}b_1, \dots, A^{-1}b_q, A^{-2}b_1, \dots, A^{-n}b_q,$$

and selecting those which are linearly independent from the previous ones. Denote κ_i , $1 \leq i \leq q$ the smallest nonnegative integer for which $A^{-\kappa_i-1}b_i$ already linearly depends on its antecedents. The numbers $\kappa_1, \dots, \kappa_q$ are the controllability indices.

It can easily be verified that the controllability indices satisfy the relation

$$\kappa_1 + \dots + \kappa_q = n.$$

According to the definition

$$A^{-\kappa_i-1} = \sum_{i=1}^q \sum_{k=0}^{\kappa_{ij}} \delta_{ijk} A^{-k-1} b_j \quad (4.22)$$

for appropriate coefficients δ_{ijk} , where

$$\kappa_{ij} = \begin{cases} \min(\kappa_i, \kappa_j) - 1, & \text{if } i \leq j \\ \min(\kappa_i, \kappa_j - 1), & \text{if } i > j. \end{cases}$$

Set $\delta_{ijk} = 0$, if $k > \kappa_{ij}$. Introduce the matrix

$$T = [A^{-1}b_1, \dots, A^{-\kappa_1}b_1, \dots, A^{-\kappa_q}b_q].$$

The proof of the following proposition is similar to that of Proposition 4.9, therefore it is omitted.

PROPOSITION 4.17. *Introduce the notations*

$$\begin{aligned} A_c &= T^{-1}AT, \\ B_c &= T^{-1}B. \end{aligned} \quad (4.23)$$

Then

$$B_c = \left(I_{\kappa_1 \dots \kappa_q}^{(m)} - S^{(m)} \tilde{D}_q \right) \tilde{D}_0^{-1}, \quad (4.24)$$

where

$$\tilde{D}_q^T = \left[\begin{array}{ccc|ccc} \delta_{110} & \delta_{111} & \dots & \delta_{11(\kappa_1-1)} & \delta_{1q0} & \delta_{1q1} & \dots & \delta_{1q(\kappa_q-1)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \delta_{q10} & \delta_{q11} & \dots & \delta_{q1(\kappa_1-1)} & \delta_{qq0} & \delta_{qq1} & \dots & \delta_{qq(\kappa_q-1)} \end{array} \right],$$

$$\tilde{D}_0 = \begin{bmatrix} \delta_{110} & \dots & \delta_{q10} \\ \vdots & & \vdots \\ \delta_{1q0} & \dots & \delta_{qq0} \end{bmatrix},$$

$$I_{\kappa_1 \dots \kappa_q}^{(m)} = \text{blockdiag} \{ [0 \dots 0 \ 1]^T, \dots, [0 \dots 0 \ 1]^T \},$$

$$S^{(m)} = \text{blockdiag} \left\{ \begin{bmatrix} 0 & 1 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 1 \end{bmatrix}, \dots, \begin{bmatrix} 0 & 1 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 1 \end{bmatrix} \right\},$$

$$A_c = B_c \text{blockdiag} \{ [1, 0, \dots, 0], \dots, [1, 0, \dots, 0] \} + S^{(m)},$$

where the size of the i -th block in the blockdiagonal matrices is κ_i .

DEFINITION 4.18. The state-space representation determined by the matrices A_c and B_c is called *controllability canonical form*.

4.6. Stochastic MFD and ARMA representations, transfer relations between the various forms

There are some alternative approaches to construct irreducible MFDs. *Forney's* approach (see (Forney, 1975)) is based on selecting a minimal polynomial basis for the right null space of the extended transfer function $[I_P | K(z)]$, *Wolovich* (Wolovich, 1974) and *Kailath* (Kailath, 1980) showed the derivation of MFDs from SSRs. The canonical forms of MFDs obtained from SSRs are discussed e.g. *Guidorzi* (Guidorzi, 1981), pseudocanonical forms are derived by *Gevers* and *Wertz* (Gevers and Wertz, 1984).

In the approach discussed here the left irreducible MFDs are obtained from stochastic canonical SSRs via transfer relations. This reduces the problem of obtaining MFDs for second order processes to an algebraic problem.

The use of monic ARMA forms is dominant e.g. in system identification, prediction and certain control applications and received a renewed interest, see *Guidorzi* and *Wertz* (Gevers and Wertz, 1984), *Deistler* and *Gevers* (Deistler and Gevers, 1989). Here we consider mainly the construction of forward and backward monic ARMA forms.

In earlier papers the deterministic ARMA forms were derived from left MFDs, see e.g. *Wolovich* and *Elliott* (Wolovich and Elliott, 1983), *Bokor* and *Keviczky* (Bokor and Keviczky, 1987). *Gevers* (Gevers, 1986), however, pointed out that this procedure generally leads to nonmonic ARMA forms. *Janssen* (Janssen, 1988a) applied *Forney's* approach (Forney, 1975) to obtain ARMA forms from the transfer

function and concluded that the resulting ARMA forms are generally not monic. Bokor and Keviczky (Bokor and Keviczky, 1987) considered the realization of deterministic ARMA forms from constructibility canonical SSRs. They showed, that in the case when the matrix A in the SSR is invertible, monic ARMA forms can always be obtained via transfer relations. This concept can be extended to obtain forward monic ARMA forms from stochastic forward SSRs.

In addition it will be shown here that the backward monic ARMA forms are naturally related to backward stochastic SSRs via transfer relations.

The transfer relations between forward and backward realizations having the same state process are given in Chapter 5.3.

THEOREM 4.19. (FSSR–FMFD transfer relation) *Let the n -dimensional, observable state-space representation*

$$\begin{aligned} x(t+1) &= A_o x(t) + B_o w(t), \\ y(t) &= C_o x(t) + D_o w(t). \end{aligned} \quad (4.25)$$

be given in observable canonical form. Then the associated canonical MFD is given by the transfer relations

$$y(t) = P^{-1}(U) Q(U) w(t) + D_o w(t), \quad (4.26)$$

where

$$P(z) = \text{diag}\{z^{v_1}, \dots, z^{v_p}\} - A_P S_P(z), \quad (4.27)$$

$$Q(z) = [S_Q(z) - A_P S(z)] B_o, \quad (4.28)$$

the integers $v_1, \dots, v_p$ are the observability (left Kronecker) indices, the elements of A_P are the invariant parameters α_{ijk} :

$$A_P = \left[\begin{array}{ccc|ccc} \alpha_{110} & \alpha_{111} & \dots & \alpha_{11(v_1-1)} & \dots & \alpha_{1p0} & \alpha_{1p1} & \dots & \alpha_{1p(v_p-1)} \\ \vdots & & & \vdots & \dots & \vdots & & & \vdots \\ \alpha_{p10} & \alpha_{p11} & \dots & \alpha_{p1(v_1-1)} & \dots & \alpha_{pp0} & \alpha_{pp1} & \dots & \alpha_{pp(v_p-1)} \end{array} \right] \quad (4.29)$$

and

$$S_P(z) = \text{blockdiag} \left\{ (1, z, \dots, z^{v_1-1})^T, \dots, (1, z, \dots, z^{v_p-1})^T \right\}, \quad (4.30)$$

$$S_Q(z) = \text{blockdiag} \left\{ (z^{v_1-1}, \dots, z, 1)^T, \dots, (z^{v_p-1}, \dots, z, 1)^T \right\}, \quad (4.31)$$

$$S(z) = \text{blockdiag} \{ S_1(z), \dots, S_p(z) \}, \quad (4.32)$$

where $S_k(z)$ is a $v_k \times v_k$ -dimensional Toeplitz-matrix

$$S_k(z) = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & & \\ z & \ddots & \ddots & \\ \vdots & \ddots & 1 & 0 & 0 \\ z^{v_k-2} & \dots & z & 1 & 0 \end{bmatrix}. \quad (4.33)$$

Proof. This is constructive, the key point is to exploit that the SSR is given in an observability canonical form. Using equation 4.25 recursively we can write that

$$\begin{aligned} y(t) &= C_o x(t) + D_o w(t) \\ y(t+1) &= C_o A_o x(t) + C_o B_o w(t) + D_o w(t+1) \\ &\vdots \\ y(t+n-1) &= C_o A_o^{n-1} x(t) + C_o A_o^{n-2} B_o w(t) + \dots + D_o w(t+n-1) \end{aligned}$$

Collecting together from these equations only the rows corresponding to the Scheme II selection plan, i.e. those containing the vectors $y_i(j)$, $j = t, \dots, t + v_i - 1$ we can write in view of 4.9 that

$$S_P(U)y(t) = x(t) + S(U)B_o w(t) + S_P(U)D_o w(t). \quad (4.34)$$

Furthermore, expressing the state vector from 4.34 and substituting into the equation

$$\begin{aligned} U^{v_i} y_i(t) &= c_{oi}^T A_o^{v_i-1} x(t) + \\ &+ (c_{oi}^T A_o^{v_i-1} + U c_{oi}^T A_o^{v_i-2} + \dots + U^{v_i-1} c_{oi}^T) B_o w(t) + \\ &+ U^{v_i} D_o w(t), \quad i = 1, \dots, p \end{aligned}$$

we obtain – observing that the rows of A_P are equal to $c_{oi}^T A_o^{v_i-1}$ – that

$$P(U)y(t) = [Q(U) + P(U)D_o]w(t), \quad (4.35)$$

which is equivalent to 4.26 because $\deg[\det P(z)] = n$.

The polynomial matrices S_P , S_Q , S are defined by Equations 4.30–4.33. Consequently their structures are defined uniquely through observability indices $v_1, \dots, v_p$, hence FMFD is uniquely determined by the observability invariants of A_o , C_o and by the matrices B_o , D_o . ■

REMARK 4.20. In Theorem 4.19 the MFD is uniquely determined by observability invariants of the forward state space representation and by the (forward) white

noise process. Therefore this MFD is called observability canonical forward MFD (FMFD). The relationship between observability canonical state-space representations and FMFD-s occurring in Theorem 4.19 is a bijection. In addition, it can be shown that if the SSR in question is a minimal representation then the corresponding FMFD is irreducible, i.e. the polynomials $P(z)$, $Q(z)$ are left coprime.

In the following we demonstrate that the BSSR derived in the discussion preceding Theorem 5.18 is naturally related to a backward ARMA representation. This is formally stated in the following theorem.

THEOREM 4.21. (BSSR-BARMA transfer relation) *Let the n -dimensional (observable) backward SSR*

$$\begin{aligned}\bar{x}(t) &= \bar{A}_o \bar{x}(t+1) + \bar{B}_o \bar{w}(t), \\ y(t) &= \bar{C}_o \bar{x}(t+1) + \bar{D}_o \bar{w}(t),\end{aligned}\quad (4.36)$$

be given in observable canonical form, with $\bar{A}_o$ invertible. Then the associated canonical backward ARMA form is given by the transfer relations

$$y(t) = \bar{A}^{-1}(U^{-1})\bar{B}(U^{-1})\bar{w}(t) + \bar{D}_o \bar{w}(t), \quad (4.37)$$

where

$$\bar{A}(z^{-1}) = I_p - \bar{C}_o \bar{A}_o^{-1} \bar{S}_A(z^{-1}), \quad (4.38)$$

$$\bar{B}(z^{-1}) = -\bar{C}_o \bar{A}_o^{-1} \bar{S}_B(z^{-1}) \bar{B}_o. \quad (4.39)$$

and

$$\bar{S}_A(z^{-1}) = \text{blockdiag} \{ (z^{-1}, \dots, z^{-v_1})^T, \dots, (z^{-1}, \dots, z^{-v_p})^T \}, \quad (4.40)$$

$$\bar{S}_B(z^{-1}) = \text{blockdiag} \{ \bar{S}_{B1}(z^{-1}), \dots, \bar{S}_{Bp}(z^{-1}) \}, \quad (4.41)$$

where $\bar{S}_{Bk}(z^{-1})$ is a $v_k \times v_k$ -dimensional Toeplitz-matrix

$$\bar{S}_{Bk}(z^{-1}) = \begin{bmatrix} 1 & 0 & \dots & 0 \\ z^{-1} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ z^{-v_k+1} & \dots & z^{-1} & 1 \end{bmatrix}. \quad (4.42)$$

(Recall that the integers $v_1, \dots, v_p$ denote the observability indices of the pair $(\bar{C}_o, \bar{A}_o)$.)

Proof. From equation 4.36 we can write the following recursion

$$y(t-k) = \overline{C} \overline{A}_o^k \overline{x}(t+1) + \overline{C}_o \sum_{i=0}^{k-1} \overline{A}_o^{k-i-1} \overline{B}_o \overline{w}(t-i) + \overline{D}_o \overline{w}(t-k) \quad (4.43)$$

for $k = 1, \dots, n$. Collecting together from these equations only the rows corresponding to the Scheme II selection plan, i.e. those containing the vectors $y_i(t-j)$, $j = 1, \dots, v_i$ we can write that

$$\overline{S}_A(U^{-1})y(t) = \overline{A}_o \overline{x}(t+1) + \overline{S}_B(U^{-1})\overline{B}_o \overline{w}(t) + \overline{S}_A(U^{-1})\overline{D}_o \overline{w}(t). \quad (4.44)$$

Furthermore, expressing the state vector from 4.44 and substituting into the observational equation of 4.36 we obtain that

$$\overline{A}(U^{-1})y(t) = [\overline{B}(U^{-1}) + \overline{A}(U^{-1})\overline{D}_o] \overline{w}(t),$$

concluding the proof. ■

REMARK 4.22. *The 4.37 BARMA representation is uniquely determined by the observability invariants of the BSSR and by the backward white noise process $\overline{w}(t)$, $t \in \mathbb{Z}$ (through the matrices $\overline{B}_o$, $\overline{D}_o$). Therefore this BARMA representation is called observability canonical BARMA representation. The correspondence between observability canonical BSSR-s and observability canonical BARMA forms is a one-to-one relationship.*

Forward ARMA representations can be derived from forward SSR-s utilizing the constructibility invariants introduced in this chapter.

Now, for the sake of completeness, we state the appropriate transfer relations between the observability and constructibility canonical forms of the state-space equations and the corresponding ARMA, MFD forms, but we neglect the detailed verification, because it can be accomplished in the same way as in Theorems 4.19 and 4.21.

THEOREM 4.23. (FSSR-FARMA transfer relation) *Let the n -dimensional (minimal) forward SSR*

$$\begin{aligned} x(t+1) &= A_c x(t) + B_c w(t), \\ y(t) &= C_c x(t) + D_c w(t), \end{aligned} \quad (4.45)$$

be given in constructible canonical form, and assume that A_c is invertible and $\text{rank } C_c = p$. Then the associated canonical forward ARMA form is given by the transfer relations

$$y(t) = A^{-1}(U^{-1})B(U^{-1})w(t) + D_c w(t), \quad (4.46)$$

where

$$A(z^{-1}) = I_p - C_c S_A(z^{-1}),$$

$$B(z^{-1}) = -C_c S_B(z^{-1}) B_c,$$

and

$$S_A(z^{-1}) = \text{blockdiag} \{ [z^{-1}, \dots, z^{-\mu_1}]^T, \dots, [z^{-1}, \dots, z^{-\mu_p}]^T \}, \quad (4.47)$$

$$S_B(z^{-1}) = \text{blockdiag} \{ S_{B1}(z^{-1}), \dots, S_{Bp}(z^{-1}) \}, \quad (4.48)$$

where $S_{Bk}(z^{-1})$ is a $\mu_k \times \mu_k$ -dimensional Toeplitz-matrix

$$S_{Bk}(z^{-1}) = \begin{bmatrix} z^{-1} & 0 & \dots & 0 \\ z^{-2} & z^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ z^{-\mu_k} & \dots & z^{-2} & z^{-1} \end{bmatrix}. \quad (4.49)$$

The integers $1, \dots, \mu_p$ are the constructibility indices.

REMARK 4.24. The FARMA representation above is uniquely determined by the constructibility invariants of the FSSR and by the (forward) white noise process $w(t)$, $t \in \mathbb{Z}$ (through the matrices B_c , D_c). Therefore this FARMA representation is called constructible canonical FARMA representation.

The correspondence between constructibility forward state-space representations and constructibility canonical FARMA forms is a one-to-one relationship.

THEOREM 4.25. (BSSR-BMFD transfer relation) *Let the n -dimensional (minimal) backward SSR*

$$\bar{x}(t) = \bar{A}_c \bar{x}(t+1) + \bar{B}_c \bar{w}(t), \quad (4.50)$$

$$y(t) = \bar{C}_c \bar{x}(t+1) + \bar{D}_c \bar{w}(t),$$

be given in constructible canonical form, and assume that $\bar{A}_c$ is invertible and $\text{rank } \bar{C}_c = p$. Then the associated canonical backward MFD form is given by the transfer relations

$$y(t) = \bar{P}^{-1}(U) \bar{Q}(U) \bar{w}(t) + \bar{D}_c \bar{w}(t), \quad (4.51)$$

where

$$\bar{P}(z) = \text{diag} \{ z^{\mu_1}, \dots, z^{\mu_p} \} - \bar{A}_p \bar{S}_p(z),$$

$$\bar{Q}(z) = (\bar{A}_p \bar{S}(z) - \bar{S}_Q(z)) \bar{B}_c,$$

the integers $1, \dots, \mu_p$ are the constructibility indices, the elements of $\bar{A}_p$ are the invariant parameters $\{\beta_{ijk}\}$ of $(\bar{C}_c, \bar{A}_c^{-1})$ and

$$\bar{S}_p(z) = \text{blockdiag} \left\{ [1, z, \dots, z^{\mu_1-1}]^T, \dots, [1, z, \dots, z^{\mu_p-1}]^T \right\}, \quad (4.52)$$

$$\bar{S}_Q(z) = \text{blockdiag} \left\{ [z^{\mu_1}, \dots, z]^T, \dots, [z^{\mu_p}, \dots, z]^T \right\}, \quad (4.53)$$

$$\bar{S}(z) = \text{blockdiag} \left\{ \bar{S}_1(z), \dots, \bar{S}_p(z) \right\}, \quad (4.54)$$

where $\bar{S}_k(z)$ is a $\mu_k \times \mu_k$ -dimensional Toeplitz-matrix

$$\bar{S}_k(z) = \begin{bmatrix} 0 & 0 & \dots & 0 \\ z & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ z^{\mu_k-1} & \dots & z & 0 \end{bmatrix}. \quad (4.55)$$

REMARK 4.26. The BMFD representation above is uniquely determined by the constructibility invariants of the backward state-space representation 4.50 and by the (backward) white noise process $\bar{w}(t)$, $t \in \mathbb{Z}$ (through the matrices $\bar{B}_c$, $\bar{D}_c$). Therefore this BMFD representation is called constructible canonical BMFD representation.

It follows from the preceding statements that instead of deriving FARMA representations from FMFD-s, see e.g. Wolovich and Elliott (Wolovich and Elliott, 1983), Gevers (Gevers, 1986), or from transfer function, Janssen (Janssen, 1988a), it is better to obtain ARMA models from constructible canonical forms by interpreting the so-called constructibility invariants. Furthermore, the backward ARMA forms can be derived from backward state-space representations by the use of observability invariants.

REMARK 4.27. The analysis constructing canonical state-space realizations from the observability and constructibility invariants was carried out under the blanket assumption that the input noise was fixed. But since the observability and constructibility invariants defined using the matrices C , A or $\bar{C}$, $\bar{A}$ in the forward or backward cases, respectively, which are invariant for all minimal stochastic realizations of the process $y(t)$, $t \in \mathbb{Z}$, (see Chapter 5) these invariants are determined by the output process y and do not depend on the input noise w .

THEOREM 4.28. (FSSR–dual MFD transfer relations) Let the n -dimensional, controllable state-space realization

$$\begin{aligned} x(t+1) &= A_r x(t) + B_r w(t), \\ y(t) &= C_r x(t) + D_r w(t), \end{aligned} \quad (4.56)$$

be in reachable canonical form. Then the associated dual forward MFD form is given by the transfer relations

$$y(t) = P_r(U)Q_r(U)^{-1}w(t) + D_rw(t), \quad (4.57)$$

where

$$\begin{aligned} P_r(z) &= C_r S_{P,r}(z), \\ S_{P,r}(z) &= \text{blockdiag} \left\{ [z^{\lambda_1-1}, \dots, z, 1]^T, \dots, [z^{\lambda_q-1}, \dots, z, 1]^T \right\}, \\ Q_r(z) &= \text{diag}\{z^{\lambda_1}, \dots, z^{\lambda_q}\} - A_{Q,r} S_{Q,r}(z), \\ S_{Q,r}(z) &= \text{blockdiag} \left\{ [1, z, \dots, z^{\lambda_1-1}]^T, \dots, [1, z, \dots, z^{\lambda_q-1}]^T \right\}, \\ A_{Q,r} &= - \begin{bmatrix} \gamma_{111} & \gamma_{112} & \dots & \gamma_{11\lambda_1} & \dots & \gamma_{1q1} & \gamma_{1q2} & \dots & \gamma_{1q\lambda_q} \\ \vdots & & & & & \vdots & & & \\ \gamma_{q11} & \gamma_{q12} & \dots & \gamma_{q1\lambda_1} & \dots & \gamma_{qq1} & \gamma_{qq2} & \dots & \gamma_{qq\lambda_q} \end{bmatrix}. \end{aligned}$$

the integers $\lambda_1, \dots, \lambda_q$ are the reachability (right Kronecker) indices, the elements of $A_{Q,r}$ are the invariant parameters γ_{ijk} (for $\lambda_i = 0$ the corresponding blocks in $S_{P,r}, A_{Q,r}, S_{Q,r}$ are missing).

Proof. (See Kučera (Kučera, 1991), pp. 51–53 and Theorem 2.17.) Denote e_{ij} , $j = 1, \dots, \lambda_i$, $i \in I_B$ (where I_B were defined in 4.13) the unit vectors in $\mathbb{R}^n$ (in this order). Consider the upper triangular matrix G of size $q_B \times q_B$ defined as follows:

$$G = \begin{bmatrix} 1 & & & & \\ 0 & \ddots & -\gamma_{i,j,\lambda_i+1} & & \\ \vdots & \ddots & & 1 & \\ 0 & \dots & 0 & & 1 \end{bmatrix}_{j,i \in I_B}.$$

(i.e. j is the row index, i is the column index). Set

$$I_r = [e_{i1}]_{i \in I_B},$$

(an $n \times q_B$ dimensional matrix), and denote B_r^0 the matrix of size $q_B \times q_B$ the i -th column of which is the corresponding unit vector, if $\lambda_i \geq 1$, whereas it is $[\gamma_{i11}, \dots, \gamma_{i,i-1,1}, 0, \dots, 0]^T$ if $\lambda_i = 0$. Obviously $B_r = I_r B_r^0$. We determine a new basis in $\mathbb{R}^n$ consisting of the vectors t_{ij} , $j = 1, \dots, \lambda_i$, $i \in I_B$ in the following recursive way:

$$t_{i1} = e_{i1} - \sum_{j < i, j \in I_B} \gamma_{i,j,\lambda_i+1} e_{j1}, \quad (4.58)$$

or shortly

$$T_r = [t_{i1}]_{i \in I_B} = I_r G.$$

Denote by $\gamma_{i,k}$ the q_B dimensional column vector with the coordinates γ_{ijk} , $j \in I_B$, and define $\tilde{\gamma}_{i,k}$ as follows:

$$\tilde{\gamma}_{i,k} = G^{-1} \gamma_{i,k}.$$

The coordinates of $\tilde{\gamma}_{i,k}$ are denoted by $\tilde{\gamma}_{ilk}$, $j \in I_B$. Obviously

$$\sum_{j \in I_B} \gamma_{ijk} e_{j1} = \sum_{l \in I_B} \tilde{\gamma}_{ilk} t_{l1}. \quad (4.59)$$

Set

$$t_{ik} = A_r t_{i,k-1} - \sum_{l \in I_B} \tilde{\gamma}_{i,l,(\lambda_i-k+2)} t_{l1} \text{ if } k \geq 2, \quad 2 \leq k \leq \lambda_i. \quad (4.60)$$

In other words

$$t_{ik} = A_r^{k-1} t_{i,1} - \sum_{l \in I_B} \sum_{j=1}^{k-1} \tilde{\gamma}_{i,l,(\lambda_i-k+j+1)} A^{j-1} t_{l1}. \quad (4.61)$$

Observe that in the present case equation 4.14 can be written as

$$A_r^{\lambda_i} \left(e_i - \sum_{j < i, j \in I_B} \gamma_{i,j,(\lambda_i+1)} e_{j,1} \right) = \sum_{j \in I_B} \sum_{k=1}^{\lambda_i} \gamma_{ijk} A^{k-1} e_{j1}, \quad (4.62)$$

(recall that $\gamma_{ijk} = 0$, if $k > \lambda_{ij}$). Using equations 4.58 and 4.59 we obtain that

$$A_r^{\lambda_i} t_{i,1} = \sum_{l \in I_B} \sum_{k=1}^{\lambda_i} \tilde{\gamma}_{ilk} A^{k-1} t_{l1}. \quad (4.63)$$

Equations 4.60 and 4.63 imply that in the basis defined by the vectors $t_{i,j}$, $j = 1, \dots, \lambda_i$, $i \in I_B$ the matrix A_r has the following form – denoted by $A^{(m)}$ –:

$$A^{(m)} = \left[A_{li}^{(m)} \right]_{l,i \in I_B}, \quad (4.64)$$

where

$$A_{ii}^{(m)} = \begin{bmatrix} \tilde{\gamma}_{ii\lambda_i} & \dots & \tilde{\gamma}_{ii1} \\ 1 & 0 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & \ddots & \ddots & \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}, \quad (4.65)$$

$$A_{li}^{(m)} = \begin{bmatrix} \tilde{\gamma}_{il\lambda_i} & \dots & \tilde{\gamma}_{il1} \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix}, \quad l \neq i,$$

and B_r is transformed into

$$B^{(m)} = I_r G^{-1} B_r^0. \quad (4.66)$$

Partitioning the coordinates of the transformed state vector – denoted by $x^{(m)}$ – accordingly to the blocks of $A^{(m)}$ and introducing the double indexed notation

$$x^{(m)}(t) = \left(x_{ij}^{(m)}(t) \right)_{j=1, \dots, \lambda_i, i \in I_B},$$

(but still considered as a column vector) we obtain from the form of the matrix $A^{(m)}$ that

$$x_{i,j+1}^{(m)}(t+1) = x_{ij}^{(m)}(t), \quad \text{if } 1 \leq j \leq \lambda_i - 1.$$

Introduce the q -dimensional process $\xi(t) = (\xi_i(t))_{i=1, \dots, q}$, $t \in \mathbb{Z}$, for which

$$\xi_i(t) = \begin{cases} x_{i\lambda_i}^{(m)}(t), & \text{if } i \in I_B, \\ w_i(t), & \text{if } i \notin I_B. \end{cases}$$

We obtain that

$$U^j \xi_i(t) = x_{i\lambda_i}^{(m)}(t+j) = x_{i,\lambda_i-j}^{(m)}(t), \quad \text{if } 0 \leq j \leq \lambda_i - 1 \quad \text{and } i \in I_B.$$

Also

$$\begin{aligned} x_{i1}^{(m)}(t+1) &= \sum_{j \in I_B} \sum_{k=1}^{\lambda_j} \tilde{\gamma}_{j,i,(\lambda_j+1-k)} x_{jk}(t) \\ &= U^{\lambda_i} \xi_i(t) - \sum_{j \in I_B} \sum_{k=1}^{\lambda_j} \tilde{\gamma}_{j,i,(\lambda_j+1-k)} U^{\lambda_j-k} \xi_j(t) \end{aligned}$$

is equal to the i -th coordinate of $G^{-1} B_r^0 w(t)$, if $i \in I_B$. Consequently,

$$w_i(t) = \begin{cases} U^{\lambda_i} \xi_i(t) - \sum_{j < i, j \in I_B} \gamma_{j,i,\lambda_j+1} U^{\lambda_j} \xi_j(t) \\ \quad - \sum_{j \in I_B} \sum_{k=1}^{\lambda_j} \gamma_{j,i,(\lambda_j+1-k)} U^{(\lambda_j-k)} \xi_j(t), & \text{if } i \in I_B \\ \xi_i(t), & \text{if } i \notin I_B. \end{cases}$$

Thus

$$y(t) = D_r w(t) + C_r S_r(U) \xi(t)$$

and

$$w(t) = Q_r(U) \xi(t)$$

proving the representation 4.57. Elementary application of the identity

$$\det(F - G R^{-1} H) \det R = \det(R - H F^{-1} G) \det F$$

where F, G, R, H are matrices shows that $\det(Q_r(z)) = \det(zI - A_r)$ proving the theorem. ■

REMARK 4.29. *The functions*

$$P_r(z) \text{diag}(z^{-\lambda_1}, \dots, z^{-\lambda_q}), \quad Q_r(z) \text{diag}(z^{-\lambda_1}, \dots, z^{-\lambda_q})$$

are obviously matrix polynomials in z^{-1} . Denote them by $A_r(z^{-1})$ and $B_r(z^{-1})$, respectively. Then

$$y(t) = (A_r(U^{-1})B_r^{-1}(U^{-1}) + D_r)w(t). \quad (4.67)$$

If $Q_r(z)$ has no zeros at the origin, i.e. A is a regular matrix, then the multiplication by the diagonal matrix $\text{diag}\{z^{-\lambda_1}, \dots, z^{-\lambda_q}\}$ do not eliminate either of the original zeros, i.e. the representation 4.67 is a dual forward ARMA representation.

Note that since $U^{\lambda_i-1}\xi_i(0) = x_{i1}(0) \in X$ we have that $x_{i,\lambda_i}(0) = \xi_i(0) \in X \cap U^{-\lambda_i-1}H^-(w)$. It can be proved by induction that the vectors

$$U^j \xi_i(0) \quad j = 0, 1, \dots, \lambda_i - 1 - l, \quad i \in I_B \text{ for which } \lambda_i > l$$

form a basis in the subspace $\mathcal{K}_l = X \cap U^{-l}X$. This provides the possibility of defining the set of reachability indices $-\lambda_1, \dots, -\lambda_q$ in a geometric way. (See Chapter 7.) Furthermore, since obviously $\dim \mathcal{K}_l = \dim[U^l X \cap X]$, it can easily be proved that if X is also a state-space of a backward realization (cf. Chapter 5) then the reachability indices of the backward and forward realizations coincide (modulo permutation), in other words the indices are determined by the state-space X itself.

Similar construction gives the proof of the following theorem.

THEOREM 4.30. (BSSR-dual BARMA transfer relations) *Let the n -dimensional, controllable backward state-space realization*

$$\begin{aligned} \bar{x}(t) &= \bar{A}_r \bar{x}(t+1) + \bar{B}_r \bar{w}(t), \\ y(t) &= \bar{C}_r \bar{x}(t+1) + \bar{D}_r \bar{w}(t), \end{aligned} \quad (4.68)$$

be in reachable canonical form. Then the associated dual backward ARMA form is given by the transfer relations

$$y(t) = [\bar{A}_r(U^{-1})\bar{B}_r(U^{-1})^{-1} + \bar{D}_r]\bar{w}(t), \quad (4.69)$$

where

$$\begin{aligned} \bar{A}_r(z^{-1}) &= \bar{C}_r \bar{S}_{A,r}(z^{-1}), \\ \bar{S}_{A,r}(z^{-1}) &= \text{blockdiag} \left\{ [z^{-(\lambda_1-1)}, \dots, 1]^T, \dots, [z^{-(\lambda_q-1)}, \dots, 1]^T \right\}, \\ \bar{B}_r(z^{-1}) &= \text{diag}\{z^{-\lambda_1}, \dots, z^{-\lambda_q}\} - \bar{A}_{B,r} \bar{S}_{B,r}(z^{-1}), \\ \bar{S}_{B,r}(z^{-1}) &= \text{blockdiag} \left\{ [1, \dots, z^{-(\lambda_1-1)}]^T, \dots, [1, \dots, z^{-(\lambda_q-1)}]^T \right\}, \end{aligned}$$

$$\bar{A}_{B,r} = - \left[\begin{array}{ccc|ccc} \gamma_{111} & \gamma_{112} & \dots & \gamma_{11\lambda_1} & \dots & \gamma_{1q1} & \gamma_{1q2} & \dots & \gamma_{1q\lambda_q} \\ \vdots & & & & & \vdots & & & \\ \gamma_{q11} & \gamma_{q12} & \dots & \gamma_{q1\lambda_1} & \dots & \gamma_{qq1} & \gamma_{qq2} & \dots & \gamma_{qq\lambda_q} \end{array} \right],$$

the integers $\lambda_1, \dots, \lambda_q$ are the reachability (right Kronecker) indices of the pair $(\bar{A}_r, \bar{B}_r)$, the elements of $\bar{A}_{B,r}$ are the invariant parameters γ_{ijk} (for $\lambda_i = 0$ the corresponding blocks in $\bar{S}_{A,r}, \bar{A}_{B,r}, \bar{S}_{B,r}$ are missing).

REMARK 4.31. Similarly to Remark 4.29, if $\bar{A}_r$ is nonsingular, then the dual BARMA representation can be transformed into a dual BMFD representation. To this define

$$\begin{aligned} \bar{P}_r(z) &= \bar{A}_r(z^{-1}) \text{diag}(z^{\lambda_1}, \dots, z^{\lambda_q}), \\ \bar{Q}_r(z) &= \bar{B}_r(z^{-1}) \text{diag}(z^{\lambda_1}, \dots, z^{\lambda_q}). \end{aligned} \quad (4.70)$$

The following theorem show the connection between forward state-space realizations in controllable canonical form and dual forward ARMA representations.

THEOREM 4.32. (FSSR–dual FARMA transfer relations) Let the n -dimensional, controllable state-space realization

$$\begin{aligned} x(t+1) &= A_c x(t) + B_c w(t), \\ y(t) &= C_c x(t) + D_c w(t), \end{aligned} \quad (4.71)$$

be in controllable canonical form assuming that A_c is nonsingular and $\text{rank } B_c = q$. Then the associated dual forward ARMA form is given by the transfer relations

$$y(t) = [A_c(U^{-1})B_c(U^{-1})^{-1} + D_c]w(t), \quad (4.72)$$

where

$$\begin{aligned} A_c(z^{-1}) &= C_c S_{A,c}(z^{-1}), \\ S_{A,c}(z^{-1}) &= \text{blockdiag} \left\{ [z^{-(\kappa_1-1)}, \dots, 1]^T, \dots, [z^{-(\kappa_q-1)}, \dots, 1]^T \right\}, \\ B_c(z^{-1}) &= \text{diag} \{ z^{-\kappa_1}, \dots, z^{-\kappa_q} \} - A_{B,c} S_{B,c}(z^{-1}), \\ S_{B,c}(z^{-1}) &= \text{blockdiag} \left\{ [1, \dots, z^{-(\kappa_1-1)}]^T, \dots, [1, \dots, z^{-(\kappa_q-1)}]^T \right\}, \\ A_{B,c} &= - \left[\begin{array}{ccc|ccc} \delta_{110} & \delta_{111} & \dots & \delta_{11(\kappa_1-1)} & \dots & \delta_{1q0} & \delta_{1q1} & \dots & \delta_{1q(\kappa_q-1)} \\ \vdots & & & & & \vdots & & & \\ \delta_{q10} & \delta_{q11} & \dots & \delta_{q1(\kappa_1-1)} & \dots & \delta_{qq0} & \delta_{qq1} & \dots & \delta_{qq(\kappa_q-1)} \end{array} \right], \end{aligned}$$

the integers $\kappa_1, \dots, \kappa_q$ are the controllability (right Kronecker) indices of the pair (A_c, B_c) , the elements of $A_{B,c}$ are the invariant parameters δ_{ijk} .

Proof. The proof of this theorem is similar to that of Theorem 4.28 therefore just the key steps are mentioned. First we transform equation 4.71 into the following form:

$$\begin{aligned}x(t) &= A_c^{-1}x(t+1) - A_c^{-1}B_cw(t), \\y(t) &= C_cA_c^{-1}x(t+1) + (D_c - C_cA_c^{-1}B_c)w(t).\end{aligned}\tag{4.73}$$

Now observe that the construction of the controllable canonical form gives that the column vectors of the matrix $A_c^{-1}B_c$ are unit vectors. Note that the assumption $\text{rank } B_c = q$ implies that now the restriction of the indices to the set I_B is not needed. Transforming these unit vectors using the controllability invariants $\delta_{ji\kappa_j}$ as in 4.58 we obtain the vectors $t_{l1}, l = 1, \dots, q$. From this set of linearly independent vectors the basis $t_{lj}, j = 1, \dots, \kappa_l, l = 1, \dots, q$ is constructed similarly to equation 4.61 but using A_c^{-1} instead of A_r in the recursive definition. As a consequence in this basis the form of the matrices A_c^{-1} and $A_c^{-1}B_c$ becomes similar to that of listed in 4.64–4.66.

Applying the corresponding basis transformation we obtain a transformed state vector denoted by $x^{(m)}$. Partitioning its coordinates and using double indices according to the indices used in the definition of the basis vectors t_{lj} the auxiliary process ξ is defined as follows:

$$\xi_i(t) = x_{i\kappa_i}^{(m)}(t+1).$$

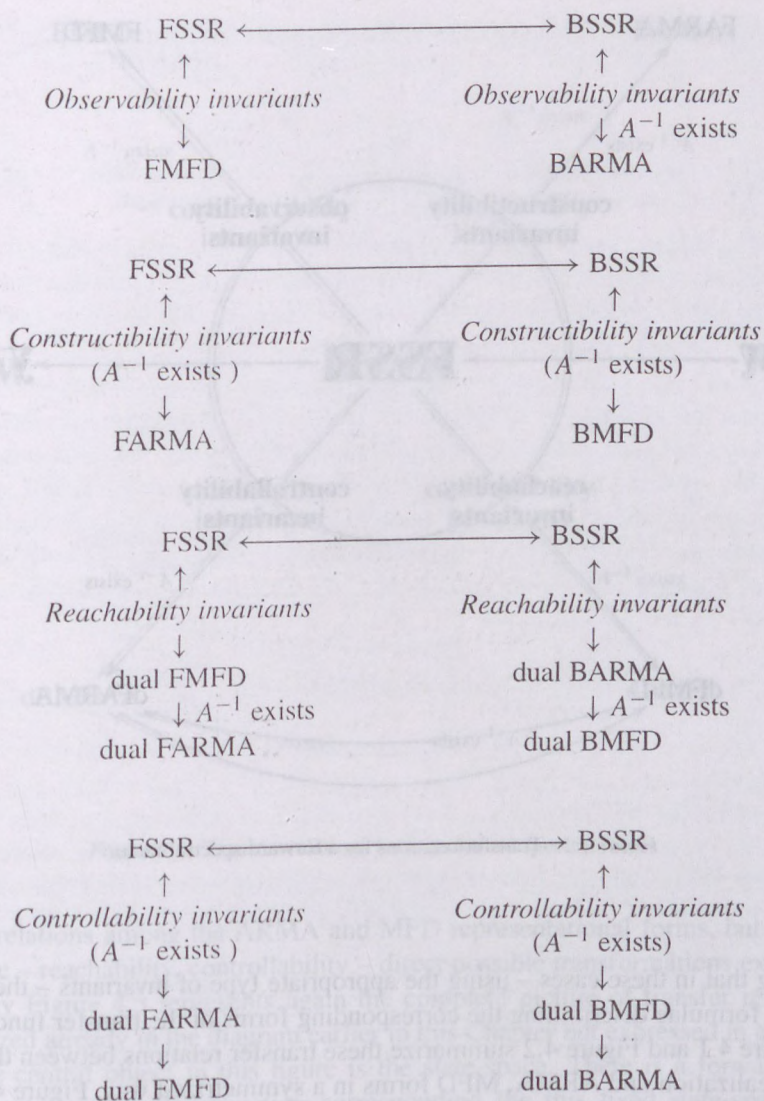
Straightforward, but tedious bookkeeping steps show that the processes $x(t)$ (consequently $y(t)$ as well) and $w(t)$ can be expressed using the process $\xi(t)$ giving the representation 4.72. ■

REMARK 4.33. Similarly to Remark 4.31, this dual forward ARMA representation can be transformed into a dual forward MFD form multiplying the factors $A_c(z^{-1})$ and $B_c(z^{-1})$ with the matrix $\text{diag}(z^{\kappa_1}, \dots, z^{\kappa_q})$.

To make the picture more complete we mention that starting from a backward state-space realization given in controllable canonical form a dual backward MFD representation can be constructed which can be transformed applying a diagonal factor into a dual backward ARMA form.

4.7. Correspondences among stochastic state-space, MFD and ARMA representations

The correspondences among stochastic state-space representations and MFD, ARMA forms can be illustrated by the following diagrams.



Let us point out that in the previous diagram the vertical arrows indicate a transfer relation between two representations of the same transfer function. At the same time the horizontal arrows connect forward and backward state-space realizations, i.e. the do not have the same spectral factor, transfer function, but these realizations have the same state space.

Let us emphasize that starting from the spectral density of the output stationary process all possible spectral factors together with its various forms (state-space, ARMA, MFD) can be constructed, although usually this leads to very complicated calculations. The arrows above show the transfer relations discussed in this Chapter,

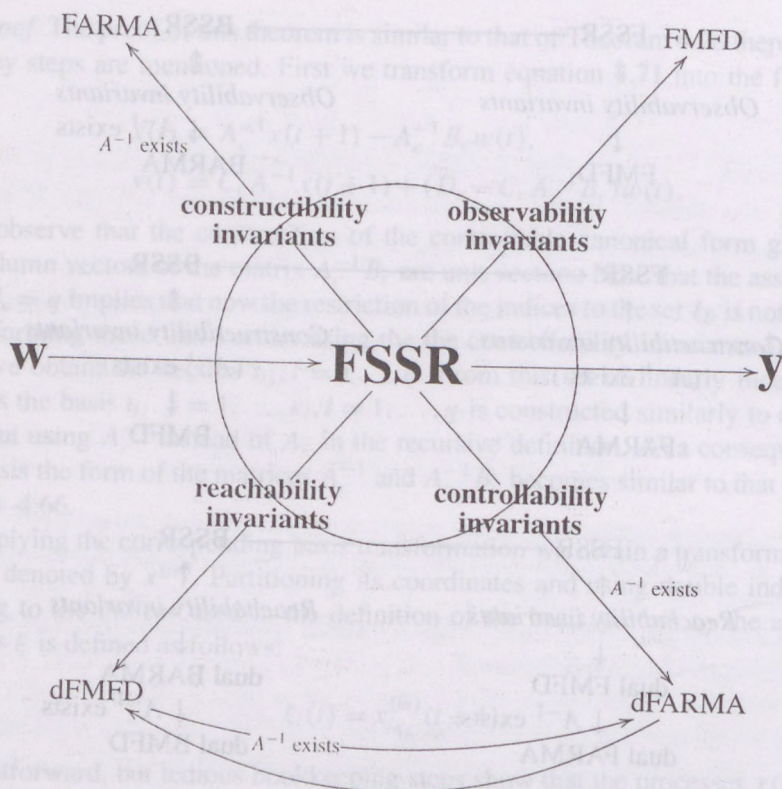


Figure 4.1. Transfer relations for a forward spectral factor

showing that in these cases – using the appropriate type of invariants – there exist explicit formulae determining the corresponding forms of the transfer function.

Figure 4.1 and Figure 4.2 summarize these transfer relations between the state-space realization and ARMA, MFD forms in a symmetrized way. Figure 4.1 for a causal – analytic – spectral factor, while Figure 4.2 for an anticausal – coanalytic – spectral factor.

Each of these figures correspond to the various forms of the same spectral factor. The horizontal arrows in these figures leading from the fixed generating input (forward or backward) white noise to the output via the state-space express this fact. From these figures it can be seen what kind of "computable" transfer relations exist based on various system invariants corresponding to the four Kalman criteria.

The interesting feature of this figures is that the "network" of the transfer relations does not form a simple symmetric structure, because on the side of the output generating principles – constructibility, observability – there are no direct

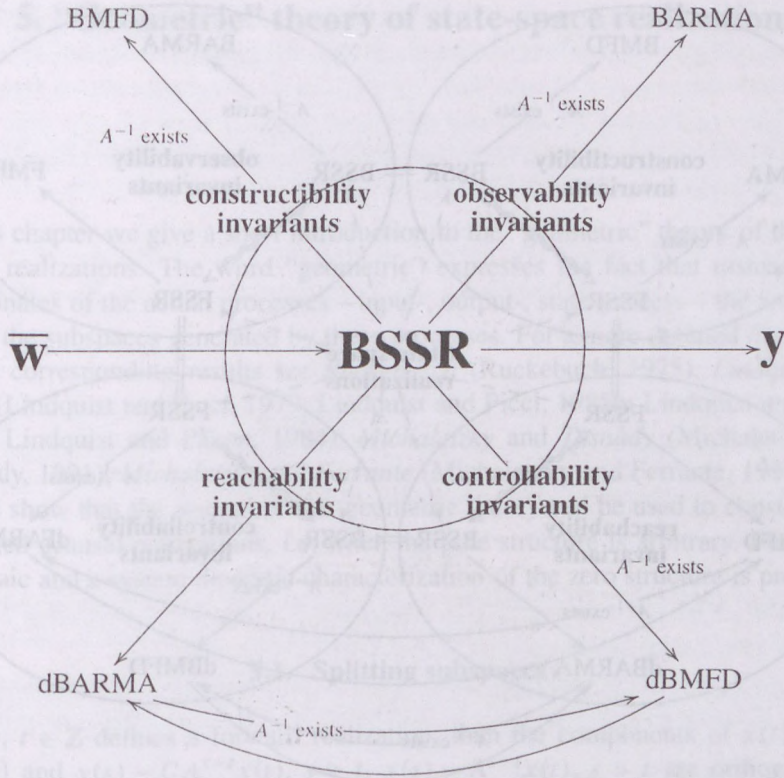


Figure 4.2. Transfer relations for a backward spectral factor

transfer relations among the ARMA and MFD representational forms, but on the input side – reachability, controllability – direct possible transformations exist.

Finally Figure 4.3 represents again the complete picture of transfer relations summarized already in the diagram earlier in this Chapter but expressed in another way. The central object in this figure is the state-space. There is a forward and a backward state-space realization corresponding to this fixed state-space. In the figure broken arrows are pointing from this "common" state-space to the corresponding forward and backward state-space realizations represented by the vertices of an octagon. (Because of dimensional limitations – a picture should be two-dimensional – the four "FSSR" vertices represent the same forward spectral factor, similarly for the four "BSSR" vertices.) The various system invariants are represented by the big ellipses. Each ellipse and the octagon has two vertices in common – corresponding to the forward and the backward state-space realization.

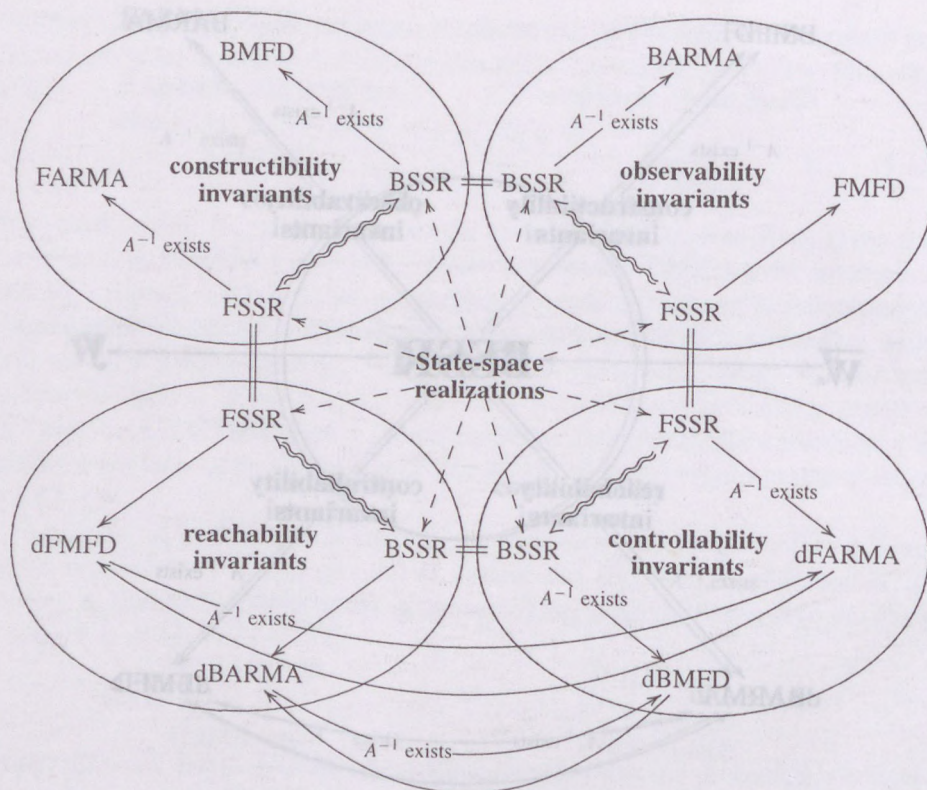


Figure 4.3. The transfer relations between the various representations. The big ellipses represent the various sets of invariants. Solid lines (curves) indicate that they provide another form of the "same" transfer function. The curly lines indicate the connections between realizations with the same state-space.

5. "Geometric" theory of state-space realizations

In this chapter we give a short introduction to the "geometric" theory of the state-space realizations. The word "geometric" expresses the fact that instead of the coordinates of the actual processes – input-, output-, state-process – the analysis is based the subspaces generated by these processes. For a more detailed description of the corresponding results see *Ruckebusch* (Ruckebusch, 1975), *Lindquist* and *Picci* (Lindquist and Picci, 1979; Lindquist and Picci, 1985b; Lindquist and Picci, 1990; Lindquist and Pavon, 1984), *Michaletsky* and *Tusnády* (Michaletsky and Tusnády, 1991), *Michaletsky* and *Ferrante* (Michaletsky and Ferrante, 1995).

We show that the arsenal of this geometric theory can be used to construct the so-called acausal realizations, i.e. when the pole structure is arbitrary. Finally, an algebraic and a system theoretic characterization of the zero structure is provided.

5.1. Splitting subspaces

If $x(t)$, $t \in \mathbb{Z}$ defines a forward realization, then the components of $x(t)$ are in $H_t^-(w)$ and $y(s) - CA^{s-t}x(t)$, $s \geq t$, $x(s) - A^{s-t}x(t)$, $s > t$ are orthogonal to $H_t^-(w)$, so the process $x(t)$ is *Markovian* (in L_2 -sense) for the pair of processes $(x(t), y(t))$.

Denote by S the closed subspace generated by the components of $y(t)$, $t \leq -1$, and $x(t)$, $t \leq 0$. Similarly $\bar{S}$ denotes the closed subspace generated by the components of $y(t)$, $t \geq 0$, $x(t)$, $t \geq 0$. (Observe that the set of indices in the definition of S and $\bar{S}$ for the process y are disjoint but for the process x the value $t = 0$ is included in both sets.) The Markovian property can be expressed that S and $\bar{S}$ are conditionally orthogonal given X , i.e.

$$E[(\xi - \Pr_X \xi)(\eta - \Pr_X \eta)] = 0, \quad \forall \xi \in \bar{S}, \eta \in S. \quad (5.1)$$

Obviously

$$\begin{aligned} H^- &\subset S, \quad H^+ \subset \bar{S} \\ U^{-1}S &\subset S, \quad U\bar{S} \subset \bar{S} \end{aligned} \quad (5.2)$$

and

$$X = S \cap \bar{S}. \quad (5.3)$$

Denote $H = S \vee \bar{S}$. The conditional orthogonality property 5.1 can be expressed as

$$H = \bar{S}^\perp \oplus (S \cap \bar{S}) \oplus S^\perp. \quad (5.4)$$

We can repeat the same argument for the backward realizations. This leads also to conditionally orthogonal subspaces $S, \bar{S}$ with the properties above. If we are able to show that any pair of subspaces $(S, \bar{S})$ with the properties above determine a state-space realization forward and backward together then we also prove that the state-spaces of forward and backward realizations coincide. (See Lindquist and Picci (Lindquist and Picci, 1979; Lindquist and Pavon, 1984).)

THEOREM 5.1. (Lindquist and Picci (Lindquist and Picci, 1990)) *Let $H \supset H_0$ be a Hilbert-space, U be a shift operator on H extending the shift on H_0 . Let $S, \bar{S}$ be two subspaces of H such that*

$$\begin{aligned} H^- \subset S, & \quad H^+ \subset \bar{S} \\ U^{-1}S \subset S, & \quad U\bar{S} \subset \bar{S} \end{aligned} \quad (5.5)$$

and $S, \bar{S}$ are conditionally orthogonal with respect to

$$X = S \cap \bar{S}. \quad (5.6)$$

Then there exists a forward (and also a backward) state-space realization of y for which X is spanned by the components of $x(0), (\bar{x}(0))$.

Proof. Choose an orthonormal basis in $US \ominus S$ this should form the components of $w(0)$. Choose an arbitrary basis in X . This defines the coordinates of $x(0)$. Set

$$x(t) = U^t x(0), \quad w(t) = U^t w(0).$$

Let us check that the equations in the state-space realization are fulfilled. Projecting $x(1)$ onto S we get that the image is in X so it can be expressed as linear combinations of the components of $x(0)$. The projection error is in $US \ominus S$, so it can be expressed in terms of $w(0)$. Denote the corresponding coefficients by A and B , i.e.

$$Ax(0) = \text{Pr}_S x(1),$$

$$Bw(0) = \text{Pr}_{US \ominus S} x(1).$$

Consequently – applying the shift U^t –

$$x(t+1) = Ax(t) + Bw(t).$$

Similarly, projecting the components of $y(0)$ onto S we get the equation

$$y(t) = Cx(t) + Dw(t).$$

The same construction works for backward realizations. ■

DEFINITION 5.2. *The subspace X in the previous construction – i.e. the subspace satisfying equations 5.1 and 5.5–5.6 – is called Markovian splitting subspace (for the process $y(t)$, $t \in \mathbb{Z}$).*

REMARK 5.3. *The S and $\bar{S}$ spaces can be unnecessarily large. Especially, the same-state space X can be produced by various $(S, \bar{S})$ pairs. But of course*

$$S \supset H^- \vee X^-$$

where X^- is generated by the subspaces $U^{-t}X$, $t \geq 0$.

Similarly

$$\bar{S} \supset H^+ \vee X^+,$$

where X^+ is generated by the subspaces U^tX , $t \geq 0$.

Changing S to $H^- \vee X^-$, $\bar{S}$ to $H^+ \vee X^+$ we obtain the same state-space. Because of these we shall always assume that

$$S = H^- \vee X^-, \quad \bar{S} = H^+ \vee X^+. \quad (5.7)$$

At the same time even the space X can be unnecessarily large. This raises again the question of minimality – in geometric terms.

DEFINITION 5.4. *The subspace $X \cap (H^+)^{\perp}$ is the unobservable subspace of X . The space X is called observable if $X \cap (H^+)^{\perp} = \{0\}$.*

DEFINITION 5.5. *The subspace $X \cap (H^-)^{\perp}$ is the unconstructible subspace of X . The space X is called constructible if $X \cap (H^-)^{\perp} = \{0\}$.*

DEFINITION 5.6. *A Markovian splitting subspace X is called minimal if no proper subspace of X is Markovian splitting subspace.*

THEOREM 5.7. (See Lindquist and Picci (Lindquist and Picci, 1990)). *The Markovian splitting subspace X is minimal if and only if it is observable and constructible.*

REMARK 5.8. *Let us observe that the notion of minimality used here is the minimality in the subspace inclusion sense. Earlier, for stationary processes with rational spectral density we have defined the minimality in terms of the dimension of the state-spaces (see: Definition 3.15). But it is easy to see that now these two notions coincide. To this aim first observe that if X is a minimal Markovian splitting subspace then*

$$\text{Pr}_S H^+ = X$$

in view of the observability of X , so

$$\text{Pr}_{H^-} X = \text{Pr}_{H^-} \text{Pr}_S H^+ = \text{Pr}_{H^-} H^+ = X_{-}. \quad (5.8)$$

But because no nonzero elements in X are orthogonal to H^- – constructibility – we obtain that

$$\dim X = \dim X_-.$$

Thus – in the rational case – the minimality of X in subspace inclusion sense implies that it is of minimal dimension.

REMARK 5.9. In the previous construction the subspace $S \ominus U^{-1}S$ was spanned by the components of $w(-1)$. This implies that

$$S \supset H_0^-(w).$$

It may happen even under the assumption $S = H^- \vee X^-$ that these two subspaces do not coincide. An obvious necessary and sufficient condition to have that $S = H_0^-(w)$ is that $\bigcap_t U^{-t}S = \{0\}$. Similarly, for the backward realization the coordinates of $\bar{w}(0)$ span the subspace $\bar{S} \ominus U\bar{S}$, consequently

$$\bar{S} \supset H_0^+(w).$$

To have equality here $\bigcap_t U^t\bar{S} = \{0\}$ should hold.

DEFINITION 5.10. A Markovian splitting subspace X is called proper if $\bigcap_t U^{-t}S = 0$ and $\bigcap_t U^t\bar{S} = 0$.

5.2. Minimal (Markovian) splitting subspaces

As we have seen a pair of conditionally orthogonal subspaces $(S, \bar{S})$ determine a state-space X . But in case of minimality $\text{Pr}_S H^+ = X$ should hold, so this naturally raises the question whether it would be possible to define a *minimal* state-space using only S .

To this aim first we need a lemma.

LEMMA 5.11. (See Lindquist and Picci (Lindquist and Pavon, 1984).) Assume that X is a Markovian splitting subspace determined by the pair $(S, \bar{S})$. Then X is observable if and only if $\bar{S} = H^+ \vee S^\perp$. Similarly, X is constructible if and only if $S = H^- \vee \bar{S}^\perp$.

Proof. The conditional orthogonality of $S, \bar{S}$ given $X = S \cap \bar{S}$ implies the orthogonal decomposition

$$H = S \vee \bar{S} = \bar{S}^\perp \oplus X \oplus S^\perp.$$

The observability of X means that

$$X \cap H^{+\perp} = \{0\}.$$

Taking the orthocomplement we get that

$$H = X^\perp \vee H^+ = (\bar{S}^\perp \oplus S^\perp) \vee H^+.$$

But $H^+ \subset \bar{S}$ so H^+ is orthogonal to $\bar{S}^\perp$. Thus the observability is equivalent to the equality

$$H = \bar{S}^\perp \oplus (H^+ \vee S^\perp)$$

or in other words to

$$\bar{S} = H^+ \vee S^\perp.$$

The proof for the constructibility is similar. ■

DEFINITION 5.12. Denote

$$N^+ = H^+ \cap (H^-)^\perp, \quad N^- = H^- \cap (H^+)^\perp. \quad (5.9)$$

LEMMA 5.13. Let X be a Markovian splitting subspace. Then, if X is observable then $X \perp N^-$, and if X is constructible then $X \perp N^+$.

Proof. Observability implies that $\bar{S}^\perp = S \cap H^{+\perp} \supset N^-$. Thus $X \subset \bar{S}$ is orthogonal to N^- . The rest follows by symmetry. ■

The previous lemma implies that if we start with a single S space and would like to obtain a minimal Markovian splitting subspace, then we have to assume that $S \perp N^+$.

THEOREM 5.14. Assume that we are given a Hilbert-space $H \supset H_0$ with a shift operator U which is an extension of the shift defined on H_0 . Let $S \supset H_-$ be a subspace of H for which

$$U^{-1}S \subset S, \quad S \perp N^+. \quad (5.10)$$

Assume also that

$$\bigcap_t U^{-t}S = \{0\}.$$

Then the equation

$$X = \text{Pr}_S H^+, \quad (5.11)$$

defines a minimal Markovian splitting subspace.

Proof. Let us try to mimic the proof of Theorem 5.1. First choose an orthonormal basis in $US \ominus S$. They give the components of $w(0)$. Choose an arbitrary basis in X . This defines the coordinates of $x(0)$. Put $x(t) = U^t x(0)$, $w(t) = U^t w(0)$. Projecting $x(1)$ – the components of which are in US – onto S we have to prove that the image is contained in X . But

$$UX = \text{Pr}_{US} UH^+,$$

thus

$$\text{Pr}_S UX = \text{Pr}_S \text{Pr}_{US} UH^+ = \text{Pr}_S UH^+ \subset \text{Pr}_S H^+ = X,$$

using that $UH^+ \subset H^+$. Consequently, $\text{Pr}_S x(1)$ can be expressed as linear combinations of the components of $x(0)$. The projection error is in $US \ominus S$, i.e. linear transform of $w(0)$. Applying the U^t shift we get the equations

$$x(t+1) = Ax(t) + Bw(t).$$

Similarly, projecting $y(0)$ onto S we can make use of the fact that the components of $y(0)$ are also in US , so we arrive at the equation

$$y(t) = Cx(t) + Dw(t).$$

The observability of X is immediate from the construction since $X = \text{Pr}_S H^+$ implies that $X \cap (H^+)^{\perp} = \{0\}$.

To prove the minimality of X it remains to check the constructibility. Although this is true in general (cf. *Linquist and Picci* (Lindquist and Picci, 1990)) we prove it only under the assumption of rationality of the spectral density when it is enough to check the dimension of X . The subspace

$$X_+ = \text{Pr}_{H^+} H^- = H^+ \ominus N^+ \quad (5.12)$$

is a minimal Markovian splitting subspace so $\dim X_+ = \text{rank } H = n$. But – using that $S \perp N^+$ – we have that

$$X = \text{Pr}_S H^+ = \text{Pr}_S X_+.$$

Consequently,

$$\dim X \leq \dim X_+.$$

Since X is a Markovian splitting subspace $\dim X \geq n$ implying

$$\dim X = n,$$

i.e. X is minimal. ■

REMARK 5.15. *The proof of minimality in the general case can be based on a theorem due to Lindquist and Picci, see (Lindquist and Picci, 1990). Namely, let X be a proper Markovian splitting subspace. Then X is minimal if and only if X is observable and $X \perp N^+$ (or X is constructible and $X \perp N^-$).*

Obviously the symmetric version of the previous theorem is also true. If $\bar{S} \supset H^+$, $U\bar{S} \subset \bar{S}$, $\bigcap_i U^i \bar{S} = \{0\}$ and $\bar{S} \perp N^-$ then the equation

$$X = \text{Pr}_{\bar{S}} H^- \quad (5.13)$$

defines a minimal Markovian splitting subspace.

DEFINITION 5.16. *A Markovian splitting subspace X is called internal if $X \subset H_0$, in which case $H = H_0$.*

REMARK 5.17. *Let us note that if X is an internal splitting subspace, i.e. $X \subset H_0$ then instead of 5.7 identities*

$$\begin{aligned} S &= H^- \vee X = H^- \vee (X \cap H^+) \\ \bar{S} &= H^+ \vee X = H^+ \vee (X \cap H^-) \end{aligned} \quad (5.14)$$

hold. (See Theorem 3.1 in Lindquist and Picci (Lindquist and Picci, 1990).)

5.3. State-space realizations

Let us harvest several conclusions of this geometric construction for the matrices playing role in state-space realizations. We shall assume that $\text{rank } H < \infty$. We have seen that each Markovian splitting subspace defines a forward and also a backward state-space realization of y . Let us assume that $x(t)$ is the state vector of a minimal forward realization. How to construct explicitly the matrices in a corresponding backward realization? Of course *any* basis in X would serve as appropriate choice for the backward process. For example $x(0)$ itself does. But computing

$$\text{Pr}_X x(-1) = P A^T P^{-1} x(0)$$

and

$$\begin{aligned} \text{Pr}_X y(-1) &= \text{cov}(y(-1), x(0)) P^{-1} x(0) = \\ &= \text{cov}(C x(-1) + D w(-1), x(0)) P^{-1} x(0) = \\ &= (C P A^T + D B^T) P^{-1} x(0) = G^T P^{-1} x(0), \end{aligned}$$

we can see that

$$\bar{x}(t) = P^{-1}x(t)$$

is a more suitable choice. In this case

$$\bar{P} = E[\bar{x}(t)\bar{x}(t)^T] = P^{-1}.$$

Since that $\bar{P}_+ \leq \bar{P}$ for any minimal backward realizations (see equation 3.61), we obtain that

$$\bar{P}_+^{-1} \geq P.$$

Denote $P_+ = \bar{P}_+^{-1}$ we get *Faurre's* theorem (cf. Theorem 3.18)

$$P_+ \geq P \geq P_- \quad (5.15)$$

(cf. equation 3.47).

Consequently in the backward state-space realization

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

$$\bar{y}(t) = \bar{C}\bar{x}(t+1) + \bar{D}\bar{w}(t)$$

the matrices $\bar{A} = A^T$, $\bar{C} = G^T$ are already defined. In order to determine the two other matrices we have to solve the following factorization equation:

$$\begin{bmatrix} P^{-1} - A^T P^{-1} A & C^T - A^T P^{-1} G \\ C - G^T P^{-1} A & R(0) - G^T P^{-1} G \end{bmatrix} = \begin{bmatrix} \bar{B} \\ \bar{D} \end{bmatrix} \begin{bmatrix} \bar{B}^T & \bar{D}^T \end{bmatrix}.$$

Direct calculation shows that

$$\bar{B} = -(I + A^T)P^{-1}(I + A)^{-1}B,$$

$$\bar{D} = D - (C + G^T P^{-1})(I + A)^{-1}B$$

determine a factorization with the same number of columns as in the matrix B , i.e. the dimension of the backward noise is the same as that of $w(t)$ (see *Michaletsky* (Michaletsky, 1992)).

As a special case, we can consider a backward realization corresponding to the state space X_- . Denote by $\bar{w}_-(t)$, $t \in \mathbb{Z}$ the corresponding input white noise and $\bar{x}_-$ the state process:

$$\bar{x}_-(t) = \bar{A}\bar{x}_-(t+1) + \bar{B}_-\bar{w}_-(t),$$

$$\bar{y}(t) = \bar{C}\bar{x}_-(t+1) + \bar{D}_-\bar{w}_-(t)$$

Similarly, there is also a forward realization with state-space $X_+ = \text{Pr}_{H^+} H^-$. Denote by w_+ the corresponding input white noise and x_+ the state process:

$$x_+(t+1) = Ax_+(t) + B_+w_+(t),$$

$$y(t) = Cx_+(t) + D_+w_+(t).$$

Also, $\overline{W}_-(z)$ and $W_+(z)$ denote the corresponding backward and forward spectral factors, respectively.

THEOREM 5.18. *Let X be a proper minimal Markovian splitting subspace corresponding to a full rank stationary process with rational spectral density. Choose an arbitrary basis in X and denote by P the corresponding covariance matrix. Then X is internal (see: Definition 5.16) if and only if*

$$\text{rank } M(P) = p. \quad (5.16)$$

REMARK 5.19. *This means that the dimension $q = \text{rank } M(P)$ of the corresponding white noise is p , i.e. it defines a square spectral factor.*

Proof. Define a white noise $w(t)$, $t \in \mathbb{Z}$ for which

$$H_t^-(w) = U^t S.$$

Since $H^- = H_0^-(w_-) \subset S$, $w_-(t)$ can be expressed as

$$w_-(t) = \sum_{s=-\infty}^t g(t-s)w(s) = \sum_0^{\infty} g(s)w(t-s).$$

Defining $C(z) = \sum_{k=0}^{\infty} g(k)z^k$ and using that both $w_{-j}(t)$, $j = 1, \dots, p$, $t \in \mathbb{Z}$ and $w_j(t)$, $j = 1, \dots, q$, $t \in \mathbb{Z}$ form an orthonormal set we obtain that

$$I_p = C(e^{-i\lambda})C^*(e^{-i\lambda}), \quad \text{for a.e. } \lambda.$$

Since C is an (analytic) $p \times q$ matrix function this implies that $q \geq p$. If X is internal then we may converse the previous argument, $w(t)$ can be expressed using $w_-(s)$, $s \in \mathbb{Z}$, leading to the opposite inequality $q \leq p$. Thus in the internal case $q = p$. Conversely, if $q = p$ then C is a $p \times p$ rational inner function. Since it is a square matrix function for almost every λ we also have that

$$I_p = C(e^{-i\lambda})C^*(e^{-i\lambda}).$$

Performing the multiplication and comparing the coefficients of $e^{-i\lambda}$ on both sides we obtain that

$$\sum_{u=\max(-t,0)}^{\infty} g^*(u+t)g(u) = I_p \delta_{t,0}.$$

Consequently

$$\begin{aligned}
 \sum_{s=t}^{\infty} g^*(s-t)w_-(s) &= \sum_{s=t}^{\infty} g^*(s-t) \sum_{r=-\infty}^s g(s-r)w(r) = \\
 &= \sum_{r=-\infty}^{+\infty} \left[\sum_{s=\max(t,r)}^{\infty} g^*(s-t)g(s-r) \right] w(r) = \\
 &= \sum_{r=-\infty}^{+\infty} \left[\sum_{u=\max(t-r,0)}^{\infty} g^*(u+r-t)g(u) \right] w(r) = w(t).
 \end{aligned}$$

So $w(t)$ is a linear transform of $w_-(s)$, $s \in \mathbb{Z}$, i.e. its components are in H_0 . But X is a proper splitting subspace, thus $x(0)$ can be expressed in terms of $w(s)$, $s < 0$, consequently $X \subset H_0$, i.e. X is an internal splitting subspace. ■

5.4. Algebraic characterization of various subspaces

Let us assume that the full-rank stationary process $y(t)$, $t \in \mathbb{Z}$ with rational spectral density has the following state-space realization

$$\begin{aligned}
 x(t+1) &= Ax(t) + Bw(t), \\
 y(t) &= Cx(t) + Dw(t),
 \end{aligned} \tag{5.17}$$

where the eigenvalues of the matrix A are of modulus less than one, i.e. $\rho(A) < 1$. X denotes the state-space. (For the time being we do not assume the minimality of the realization defined in the equation 5.17.) It is generated by the two conditionally orthogonal subspaces $S, \bar{S}$, where now $S = H^-(w)$. Obviously $S^\perp = H^+(w)$. Consider the subspace

$$H^+(w) \ominus [H^+(w) \cap H^+(y)] = \text{Pr}_{H^+(w)} [H^+(y)^\perp]. \tag{5.18}$$

This is invariant under the operator $\text{Pr}_{H^+(w)} U^{-1}$.

Note that if the system of equations 5.17 defines a constructible realization then

$$\begin{aligned}
 H^+(w) \cap H^+(y) &= S^\perp \cap H^+ = (H^- \vee \bar{S}^\perp)^\perp \cap H^+ = \\
 &= H^{-\perp} \cap H^+ = N^+
 \end{aligned}$$

using Lemma 5.13 and the inclusion $H^+ \subset \bar{S}$. In other words, in this case

$$H^+(w) \ominus [H^+(w) \cap H^+(y)] = S^\perp \ominus N^+.$$

LEMMA 5.20. *There exist matrices F , K_0 such that $\rho(F) < 1$ holds, for which the coordinates of*

$$\sum_{k=0}^{\infty} F^k K_0 w(k)$$

generate the subspace $\text{Pr}_{H^+(w)}[H^+(y)^\perp]$.

Proof. Denote $l = \dim[H^+(w) \ominus (H^+(w) \cap H^+)]$. Taking an arbitrary generator system consisting of $l' \geq l$ vectors in the subspace $H^+(w) \ominus [H^+(w) \cap H^+]$ and expressing its elements as linear combinations of $w(t)$, $t \geq 0$ we can define a sequence of matrices $K_0, K_1, \dots$ such that the coordinates of

$$\sum_{k=0}^{\infty} K_t w(t)$$

generate $H^+(w) \ominus [H^+(w) \cap H^+]$. Since this subspace is invariant under $\text{Pr}_{H^+(w)} U^{-1}$ for every $\alpha \in \mathbb{R}^l$ there exists a $\beta \in \mathbb{R}^l$ such that

$$\alpha^T \sum_{k=1}^{\infty} K_t w(t-1) = \beta^T \sum_{k=0}^{\infty} K_t w(t),$$

i.e.

$$\alpha^T K_{t+1} = \beta^T K_t \text{ for every } t \geq 0,$$

in other words

$$\text{range} \begin{bmatrix} K_1^T \\ K_2^T \\ \vdots \end{bmatrix} \subset \text{range} \begin{bmatrix} K_0^T \\ K_1^T \\ \vdots \end{bmatrix}.$$

This implies the existence of a matrix F such that

$$K_t = F K_{t-1} \text{ for every } t \geq 1,$$

i.e.

$$K_t = F^t K_0 \text{ for every } t \geq 0.$$

Let us point out that since we have considered only a generator system the rows of the matrix $[K_0, K_1, \dots]$ are not necessarily linearly independent, thus the matrix F should not be uniquely defined. On the other hand, we can assume that

$$\text{range } F = \text{range} [K_1, K_2, \dots]. \quad (5.19)$$

Introduce the notation

$$\mathcal{K} = \sum_{k=0}^{\infty} K_t K_t^T. \quad (5.20)$$

Obviously the *Ljapunov*-equation

$$\mathcal{K} - F\mathcal{K}F^T = K_0K_0^T \quad (5.21)$$

holds true. We show that $\rho(F) < 1$. In fact, if for a nonzero vector $x \in \mathbb{R}^l$

$$x^T F = \lambda x^T \quad (5.22)$$

then from 5.20 it follows that

$$(1 - |\lambda|^2)x^T \mathcal{K}x = x^T K_0K_0^T x.$$

Consequently, if $|\lambda| \geq 1$ then $x^T K_0 = 0$ holds, thus

$$x^T K_j = x^T F^j K_0 = \lambda^j x^T K_0 = 0$$

for every $j \geq 0$. Using condition 5.19 we obtain that

$$x^T F = 0$$

contradicting to 5.22 and the assumption $|\lambda| \geq 1$. ■

Let us emphasize that the proof was based on the observation that $H^+(w) \ominus [H^+(w) \cap H^+]$ is invariant under the operator $\text{Pr}_{H^+(w)} U^{-1}$. Especially, it does not depend on the process y itself, only on the subspace generated by its “future”: $y(t), t \geq 0$.

The matrices F and K_0 constructed in Lemma 5.20 are not unique, they depend on the previously chosen generator system in $\text{Pr}_{H^+(w)}[H^{+\perp}]$. This connection is the topic of the following proposition.

PROPOSITION 5.21. *Consider a basis $\xi_1, \dots, \xi_l$ in the subspace $\text{Pr}_{H^+(w)}[H^{+\perp}]$ and denote F^b, K_0^b the corresponding matrices constructed according to the proof of Lemma 5.20. Let F, K_0 be arbitrary matrices for which the coordinates of $\sum_{k=0}^{\infty} F^k K_0 w(k)$ generate $\text{Pr}_{H^+(w)}[H^{+\perp}]$. Then there exists a matrix L of full column rank such that equations*

$$\begin{aligned} FL &= LF^b, \\ K_0 &= LK_0^b \end{aligned} \quad (5.23)$$

hold.

Proof. Let us express the coordinates of $\sum_{k=0}^{\infty} F^k K_0 w(k)$ as linear combinations of the coordinates of $\sum_{k=0}^{\infty} (F^b)^k K_0^b w(k)$. Arranging the corresponding coefficients into a matrix we obtain that

$$\sum_{k=0}^{\infty} F^k K_0 w(k) = L \sum_{k=0}^{\infty} (F^b)^k K_0^b w(k). \quad (5.24)$$

Using that the first system forms a generator system while the second one forms a basis it follows that the matrix L is of full column rank. Since $w(t)$, $t \in \mathbb{Z}$ is a white noise from 5.24 for $t = 0$ we obtain that

$$K_0 = LK_0^b. \quad (5.25)$$

Substituting 5.25 into the equation 5.24 for $t = 1$ we get that

$$FLK_0^b = LF^bK_0^b \quad (5.26)$$

Consider now the equation for $t = 2$. Substituting 5.25 and 5.26 it follows that

$$L(F^b)^2K_0^b = F^2LK_0^b = FLF^bK_0^b.$$

Continuing similarly finally we arrive at the equations

$$LF^b(F^b)^kK_0^b = FL(F^b)^kK_0^b.$$

Since the row vectors of $[K_0^b, F^bK_0^b, \dots]$ are linearly independent the equation

$$FL = LF^b$$

follows concluding the proof. ■

Let us point out that since L is of full column rank the equation 5.23 implies that the spectrum of F_b is contained in the spectrum of F . (See *Gantmacher* (Gantmacher, 1988).)

Furthermore, since the coordinates of $\sum_{t=0}^{\infty} K_t w(t)$ defined in Lemma 5.20 form a generator system in

$$H^+(w) \ominus (H^+(w) \cap H^+(y)),$$

the coordinates of $\sum_{t=1}^{\infty} K_t w(t-1)$ generate the subspace

$$\Pr_{H^+(w)} U^{-1} (H^+(w) \ominus (H^+(w) \cap H^+(y))).$$

Consequently, the dimension of $\text{range}[K_1, K_2, \dots]$ coincides with the dimension of this subspace, thus this is equal to the rank of the matrix F defined in Lemma 5.20. In other words

$$\text{rank } F = \text{rank } F_b,$$

which together with the previous observation gives that the nonzero part of the spectrum of F and F_b coincide.

We show that if we omit the rank condition 5.19 the matrix F can be chosen in a uniform way, i.e. not depending on the minimal splitting subspace X . For this we need the following lemma.

LEMMA 5.22. *Let $\xi \in (H^{-\perp} \ominus N^+)$ be an arbitrary vector. Then*

$$\Pr_{S^\perp} [\Pr_{H^{-\perp} \ominus N^+} U^{-1} \xi] = \Pr_{S^\perp \ominus N^+} U^{-1} [\Pr_{S^\perp} \xi]. \quad (5.27)$$

Proof. If $\xi \in H^{-\perp} \ominus N^+$ then $U^{-1} \xi \perp N^+$, thus

$$\Pr_{H^{-\perp} \ominus N^+} U^{-1} \xi = \Pr_{H^\perp} U^{-1} \xi.$$

The inclusion $S^\perp \subset H^{-\perp}$ implies that the left hand side is $\Pr_{S^\perp} U^{-1} \xi$. On the other hand $\Pr_{S^\perp} \xi \perp N^+$ thus

$$\Pr_{S^\perp \ominus N^+} U^{-1} \Pr_{S^\perp} \xi = \Pr_{S^\perp} U^{-1} \Pr_{S^\perp} \xi.$$

At the same time $(\xi - \Pr_{S^\perp} \xi) \perp S^\perp$, consequently the orthogonality relation $U^{-1}(\xi - \Pr_{S^\perp} \xi) \perp S^\perp$ holds, proving equation 5.27. ■

The following lemma gives another characterization of the subspace $\Pr_{H^+(w)}[H^{+\perp}]$.

LEMMA 5.23. *If for the realization 5.17 the identity $H(w) = H(y)$, holds then*

$$H^+(w) \ominus [H^+(w) \cap H^+] = \Pr_{H^+(w)} [\Pr_{H^+} S] \quad (5.28)$$

Proof. Obviously

$$\begin{aligned} H^+(w) \ominus \Pr_{H^+(w)} [\Pr_{H^+} S] &= H^+(w) \cap [\Pr_{H^+} S]^\perp = \\ &= H^+(w) \cap [H^{+\perp} \oplus (H^+ \cap H^+(w))], \end{aligned} \quad (5.29)$$

But

$$H^+(w) \cap H^{+\perp} \subset H^{-\perp} \cap H^{+\perp} = [H^- \vee H^+]^\perp = \{0\},$$

using the assumption that the realization 5.17 is internal. Consequently 5.28 is equal to

$$H^+ \cap H^+(w),$$

concluding the proof. ■

Note that if equation 5.17 defines a minimal realization then the assumption $H(w) = H(y)$ means that it is an internal one.

COROLLARY 5.24. *If equation 5.17 defines a minimal and internal realization then*

$$\Pr_{H^+} S \subset \Pr_{H^+(w)} [H^{+\perp}] \oplus X.$$

Proof. Projecting $\text{Pr}_{H^+} S$ onto $H^+(w)$ and S we obtain $\text{Pr}_{H^+(w)}[H^{+\perp}]$ according to Lemma 5.23 and $X = \text{Pr}_S H^+$. ■

REMARK 5.25. Note that the projection of the n -dimensional subspace $X_+ = \text{Pr}_{H^+} S$ (where $n = \dim X$) onto $H^+(w)$ is not necessarily n -dimensional because it may contain vectors orthogonal to $H^+(w)$, i.e. contained in S as well. Consequently

$$\dim(H^+(w) \ominus [H^+(w) \cap H^+]) = n - \dim(H^+ \cap S) = n_u^{(z)},$$

this latter is equal to the number of strictly unstable zeros (i.e. with modulus larger than 1) counting multiplicities, too. (Cf. Chapter 5.7)

REMARK 5.26. Observe that Corollary 5.24 implies that if the realization 5.17 is minimal and internal then there exists a matrix T_0 such that the subspace $\text{Pr}_{H^+} S$ is spanned by the coordinates of

$$x(0) + T_0 \sum_{k=0}^{\infty} F^k K_0 w(k). \quad (5.30)$$

THEOREM 5.27.

1. If τ , K_0 and F ($\rho(F) < 1$) are arbitrary matrices satisfying the equation

$$\begin{bmatrix} \tau P & K_0 \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} = F \begin{bmatrix} \tau P & 0 \end{bmatrix}, \quad (5.31)$$

then the coordinates of the vector $\sum_{k=0}^{\infty} F^k K_0 w(k)$ are in $H^+(w) \ominus [H^+(w) \cap H^+]$.

2. Conversely, assume that the state-space realization 5.17 is observable. Denote P the covariance matrix of $x(0)$ in the state space representation 5.17. Consider the matrices F , K_0 defined in Lemma 5.20 using a previously fixed generator system in $\text{Pr}_{H^+(w)}[H^{+\perp}]$. Then there exists a matrix τ such that equation 5.31 holds.

The matrices F , K_0 defined in Lemma 5.20 are maximal solutions of 5.31. (Maximality is meant in the sense that the rank of the controllability matrix $C(F, K_0)$ is maximal.)

Furthermore, if the realization is minimal and internal then equation

$$\tau P = -\mathcal{K} T_0^T \quad (5.32)$$

holds true, where $\mathcal{K}$ and T_0 are defined in 5.20 and 5.30, respectively.

Proof. (1) Written in componentwise equation 5.31 gives that

$$\tau PC^T + K_0 D^T = 0 \quad (5.33)$$

and

$$\tau PA^T - K_0 B^T = F\tau P. \quad (5.34)$$

Since according to the assumption $\rho(F) < 1$ we can consider the vector

$$\sum_{k=0}^{\infty} F^k K_0 w(k) + \tau x(0).$$

Taking the covariances with $y(0)$ we obtain

$$K_0 D^T + \tau PC^T,$$

which is zero in view of equation 5.33. Computing the covariance with

$$y(1) = Dw(1) + CBw(0) + CAx(0)$$

we get

$$FK_0 D^T + K_0 B^T C^T + \tau PA^T C^T.$$

But equations 5.33 and 5.34 imply that

$$FK_0 D^T + K_0 B^T C^T + \tau PA^T C^T = FK_0 D^T + F\tau PC^T = 0.$$

Continuing similarly we obtain that

$$\left(\sum_{k=0}^{\infty} F^k K_0 w(k) + \tau x(0) \right) \perp y(t)$$

for every $t \geq 0$. Projecting onto $H^+(w)$ we obtain that the coordinates of

$$\sum_{k=0}^{\infty} F^k K_0 w(k) \text{ are in } \Pr_{H^+(w)} H^{+\perp}$$

proving (1).

(2) Conversely, if η belongs to the subspace $\Pr_{H^+(w)}[H^{+\perp}]$ then it can be written as

$$\eta = \Pr_{H^+(w)} \xi, \quad (5.35)$$

where $\xi \perp H^+$. On the other hand $\Pr_{S \cap (H^+)^{\perp}} \xi$ is orthogonal to $H^+(w) \vee H^+$ thus subtracting it from ξ if necessary we can assume that

$$\Pr_S \xi \in \Pr_S H^+ \subset X, \quad (5.36)$$

where X is the splitting subspace corresponding to the realization 5.17. Now, if $\xi_1, \xi_2 \perp H^+$ satisfying both 5.35 and 5.36 then

$$(\xi_1 - \xi_2) \perp H^+$$

and

$$(\xi_1 - \xi_2) \in S = H^-(w)$$

i.e. 5.36 implies that

$$(\xi_1 - \xi_2) \in X.$$

Now, from the observability of X it follows that $\xi_1 - \xi_2 = 0$. Thus ξ is uniquely defined by η and also

$$\xi = \eta + \text{Pr}_S \xi. \quad (5.37)$$

Repeating this construction for the coordinates of $\sum_{k=0}^{\infty} F^k K_0 w(k)$ as η , where F, K_0 are defined in Lemma 5.20 we obtain that there exists a matrix τ such that

$$\sum_{k=0}^{\infty} F^k K_0 w(k) + \tau x(0) \perp y(t) \text{ for every } t \geq 0. \quad (5.38)$$

Taking the covariances we get that

$$F^t K_0 D^T + \sum_{k=0}^{t-1} F^k K_0 B^T (A^T)^{t-k-1} + \tau P (A^T)^t C^T = 0. \quad (5.39)$$

For $t = 0$ we obtain that

$$K_0 D^T + \tau P C^T = 0. \quad (5.40)$$

Considering the difference of the equation 5.39 for $t+1$ and $F \times$ the equation for t we get that

$$(K_0 B^T + \tau P A^T - F \tau P) (A^T)^t C^T = 0, \quad t \geq 0.$$

The observability of (C, A) implies that

$$K_0 B^T + \tau P A^T - F \tau P = 0. \quad (5.41)$$

Thus equation 5.31 is fulfilled. This argument together with part (1) gives the maximality of the solution defined by the matrices F, K_0 provided by Lemma 5.20.

It remains to prove that

$$\tau P + \mathcal{K} T_0^T = 0 \quad (5.42)$$

holds true.

To this aim observe that the coordinates of

$$x(0) + T_0 \sum_{k=0}^{\infty} F^k K_0 w(k)$$

span the subspace $\text{Pr}_{H^+(y)} S$, especially there are in the subspace H^+ , and the coordinates of

$$\tau x(0) + \sum_{k=0}^{\infty} F^k K_0 w(k)$$

belong to $H^{+\perp}$. Taking the covariances we obtain that equation 5.42 holds true utilizing that $\mathcal{K} = \sum_{k=0}^{\infty} (F^T)^k K_0^T K_0 F^k$. ■

REMARK 5.28. *Introducing the notation*

$$\xi(t) = \sum_{k=0}^{\infty} F^k K_0 w(k+t) \quad (5.43)$$

for $t \in \mathbb{Z}$ we obtain that

$$\xi(t) = F\xi(t+1) + K_0 w(t). \quad (5.44)$$

The coordinates of $\xi(0)$ are in $H^+(w)$, in particular

$$E(\xi(0)x(0)^T) = 0.$$

As we have seen in the proof of the previous theorem the second equation expresses the fact that

$$\xi(0) + \tau x(0) \perp y(0) = Cx(0) + Dw(0)$$

while the first equation claims that

$$[\xi(0) + \tau x(0)] - F[\xi(1) + \tau x(1)] \perp x(1) = Ax(0) + Bw(0).$$

REMARK 5.29. Let us point out that the definition of the matrices F, K_0 depends on the generator system used in Lemma 5.20. On the other hand the coordinates of $x(0)$ form a basis in X . These two systems can be connected together. In fact, choosing a basis $\eta_1, \dots, \eta_n \in \text{Pr}_{H^+} S$ for which $\text{Pr}_S \eta_i$ coincide with the coordinates of $x(0)$ the vectors $\text{Pr}_{S^\perp} \eta_i$ form a generator system (possibly containing zero or in general linearly dependent vectors) in $H^+(w) \ominus [H^+(w) \cap H^+]$. Constructing the corresponding matrices F, K_0 and $\mathcal{K}$ according to Lemma 5.20 we obtain that in this case the coordinates of

$$x(0) + \sum_{k=0}^{\infty} F^k K_0 w(k)$$

span the subspace $\text{Pr}_{H^+} S$ (i.e. now $T_0 = I$). Moreover, F and K_0 define a maximal solution of 5.31 and

$$\tau P + \mathcal{K} = 0.$$

Thus

$$\begin{bmatrix} -\mathcal{K} & K_0 \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} = F \begin{bmatrix} -\mathcal{K} & 0 \end{bmatrix} \quad (5.49)$$

holds true.

(In view of Theorem 5.47 this equation determines the left stable zeros of the function $D^T + B^T(zI - A^T)^{-1}C^T$ or in other words the right unstable zeros of the spectral factor $D + C(z^{-1}I - A)^{-1}B$.)

Note that the proof of Theorem 5.27 shows that the coordinates of $\tau x(0)$ span the subspace

$$\text{Pr}_S[\bar{S} \ominus H^+].$$

Straightforward calculation gives that this latter equals to

$$X \cap (S \cap H^+)^\perp.$$

Similarly, it can be proved that the coordinates of

$$\sum_{k=0}^{\infty} F^k K_0 w(k) + \tau x(0)$$

span the subspace

$$\bar{S} \ominus H^+.$$

Consider now the subspace $S \ominus H^-$.

LEMMA 5.30. *There exist matrices R, L_0 , $\rho(R) < 1$ such that the coordinates of*

$$\sum_{k=0}^{\infty} R^{k-1} L_0 w(-k) \text{ generate the subspace } S \ominus H^-. \quad (5.45)$$

Furthermore for any solution of 5.45 there exists a matrix μ such that the equation

$$\begin{bmatrix} R\mu & L_0 \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} = \begin{bmatrix} \mu & 0 \end{bmatrix} \quad (5.46)$$

is fulfilled.

Proof. Consider an arbitrary generator system in $S \ominus H^-$. Since this subspace is invariant under the operator $\text{Pr}_S U$ the argument used in Lemma 5.20 can be repeated leading to the construction of the matrices R, L_0 satisfying the assumption $\rho(R) < 1$.

The property

$$\sum_{k=1}^{\infty} R^{k-1} L_0 w(-k) \perp y(-t) \text{ for every } t \geq 1$$

can be expressed as

$$R^{t-1} L_0 D^T + \sum_{k=t+1}^{\infty} R^{k-1} L_0 B^T (A^T)^{k-t-1} C^T = 0. \quad (5.47)$$

Introduce the notation

$$\mu = \sum_{j=0}^{\infty} R^j L_0 B^T (A^T)^j.$$

Then equation $L_0 B^T + R \mu A^T = \mu$ holds true. From 5.47 for $t = 1$ we obtain that

$$L_0 D^T + R \mu C^T = 0$$

concluding the proof. ■

Note that the size of the matrix R equal the number of vectors in the previously chosen generator system, consequently the matrix μ is not necessarily a square matrix.

(Invoking again Theorem 5.47 the equation 5.46 determines the right stable zeros of $D + C(z^{-1} - A)^{-1}B$.)

We shall show in Remark 5.32 that the matrix R can be defined in a “uniform” way, i.e. not depending on the minimal splitting subspace X . To this we need the following lemma.

Consider now the splitting subspace $X_+ = \text{Pr}_{H^+} H^-$. The corresponding “ S ” space is $S_+ = H^- \vee X_+$ while the “ $\bar{S}$ ” space is $\bar{S}_+ = H^+$.

LEMMA 5.31. *Let $\xi \in S_+ \ominus H^-$ be an arbitrary vector. Then*

$$\text{Pr}_S [\text{Pr}_{S_+ \ominus H^-} U \xi] = \text{Pr}_{S \ominus H^-} U [\text{Pr}_S \xi]. \quad (5.48)$$

Proof. Observe that if $\xi \perp H^-$ then $U \xi \perp H^-$, thus

$$\text{Pr}_{S_+ \ominus H^-} U \xi = \text{Pr}_{S_+} U \xi.$$

The inclusion $S \subset S_+$ implies that the left hand side of 5.48 is $\text{Pr}_S U \xi$. On the other hand $\text{Pr}_S \xi \in H^-$ thus

$$\text{Pr}_{S \ominus H^-} U \text{Pr}_S \xi = \text{Pr}_S U \text{Pr}_S \xi.$$

At the same time $\xi - \text{Pr}_S \xi \perp S$ consequently $U(\xi - \text{Pr}_S \xi) \perp S$ proving 5.48. ■

REMARK 5.32. Choosing an arbitrary generator system $\xi_1, \dots, \xi_m$ in $S_+ \ominus H^-$ there exists a matrix R such that

$$\text{Pr}_{S_+ \ominus H^-} U\xi = R\xi, \quad (5.49)$$

where $\xi = [\xi_1 \dots \xi_m]^T$. The previous lemma implies that defining the vectors $\eta_i = \text{Pr}_S \xi_i$ – where $\eta_1, \dots, \eta_m$ form a generator system in $S \ominus H^-$ – we have that

$$\text{Pr}_S U\eta = R\eta,$$

where $\eta = [\eta_1 \dots \eta_m]^T$, i.e. the matrix provided by Lemma 5.30 can be defined via equation 5.49.

Similarly to Lemma 5.22 the following lemma gives the possibility of using the matrix R defined via the splitting subspace X_+ for the subspaces $\bar{S}^\perp \ominus N^-$.

LEMMA 5.33. Let $\xi \in [S_+ \ominus H^-]$ be an arbitrary vector. Then

$$\text{Pr}_{\bar{S}^\perp} \text{Pr}_{S_+ \ominus H^-} U\xi = \text{Pr}_{\bar{S}^\perp \ominus N^-} U \text{Pr}_{\bar{S}^\perp} \xi. \quad (5.50)$$

Proof. Since $\bar{S} \supset H^+$ it follows that

$$\bar{S}^\perp \subset H^{+\perp} = S_+ \ominus X_+ \subset S_+. \quad (5.51)$$

First we prove that

$$\text{Pr}_{\bar{S}^\perp} (S_+ \ominus H^-) = \bar{S}^\perp \ominus N^- \quad (5.52)$$

The inclusion $N^- \subset H^-$ implies that

$$\text{Pr}_{\bar{S}^\perp} (S_+ \ominus H^-) \perp N^-,$$

i.e. it is contained in $\bar{S}^\perp \ominus N^-$. On the other hand if

$$\lambda \in \bar{S}^\perp \cap N^{-\perp} \text{ and } \lambda \perp (S_+ \ominus H^-)$$

then $\lambda \in (\bar{S}^\perp \cap N^{-\perp} \cap [H^- \oplus S_+^\perp])$. There exist $\zeta \in H^-$, $\eta \in S_+^\perp$ such that $\lambda = \zeta + \eta$.

Equation 5.51 implies that $\lambda \perp \eta$ thus $\eta = \lambda - \zeta \in (\bar{S}^\perp \vee H^-) \perp S_+^\perp$, i.e. $\eta = 0$.

Consequently $\lambda = \zeta \in (H^- \cap \bar{S}^\perp) = N^-$, proving that $\lambda = 0$.

Now using that $\xi \perp H^-$ implies that $U\xi \perp H^-$ we get that

$$\text{Pr}_{\bar{S}^\perp} \text{Pr}_{S_+ \ominus H^-} U\xi = \text{Pr}_{\bar{S}^\perp} \text{Pr}_{S_+} U\xi = \text{Pr}_{\bar{S}^\perp} U\xi.$$

On the other hand

$$U(\xi - \text{Pr}_{\bar{S}^\perp} \xi) \perp \bar{S}^\perp \text{ and } \text{Pr}_{\bar{S}^\perp \ominus N^-} U\xi = \text{Pr}_{\bar{S}^\perp} U\xi$$

using that $U\xi \perp N^-$ in view of 5.52, thereby concluding the proof. ■

REMARK 5.34. Similarly to the Remark 5.32 an arbitrary generator system $\xi_1, \dots, \xi_m$ in $S_+ \ominus H^-$ can be "carried over" to the subspace $\bar{S}^\perp \ominus N^-$ as follows. Set $\eta_i = \text{Pr}_{\bar{S}^\perp \ominus N^-} \xi_i$, $i = 1, \dots, m$, and $\eta = [\eta_1 \dots \eta_m]^T$. Then

$$\text{Pr}_{\bar{S}^\perp \ominus N^-} U\eta = R\eta. \quad (5.53)$$

5.5. Generalized inner product representation

The inner product representation introduced at the end of Chapter 3 (see equation 3.71) can be extended to general state-spaces, too.

THEOREM 5.35. (cf. Michaletzky and Tusnády (Michaletzky and Tusnády, 1991)) Let us assume that X_1, X_2 are proper, minimal Markovian splitting subspaces defined by the pairs $(S_1, \bar{S}_1), (S_2, \bar{S}_2)$ of a full-rank stationary purely nondeterministic process y with rational spectral density. Suppose that $S_2 \supset S_1$. Assume that the components of $x_2(0)$ and $x_1(0)$ form bases in X_1, X_2 , respectively, such that $\text{Pr}_{S_1} x_2(0) = x_1(0)$. Define $\bar{x}_2(0) = P_2^{-1} x_2(0)$.

Then the operators

$$\begin{aligned} K : \bar{S}_2 &\rightarrow \mathbb{R}^n & K\xi &= E[\bar{x}_2(0)\xi] \\ L : S_1 &\rightarrow \mathbb{R}^n & L\eta &= E[x_1(0)\eta] \end{aligned}$$

determine an inner product representation of $\bar{S}_2$ and S_1 , i.e.

$$E[\xi\eta] = (K\xi)^T (L\eta), \quad \forall \xi \in \bar{S}_2, \eta \in S_1.$$

Proof. First let us remark that the condition $\text{Pr}_{S_1} x_2(0) = x_1(0)$ can always be fulfilled, namely first choosing a basis in X_2 then projecting it to S_1 . Since $X_2 = \text{Pr}_{S_2} H^+$, so $X_1 = \text{Pr}_{S_1} H^+ = \text{Pr}_{S_1} X_2$, these projected vectors form a basis in X_1 .

Observe that the elements of $\bar{S}_2 \ominus X_2$ are orthogonal to S_1 because the conditional orthogonality of $S_2, \bar{S}_2$ given X_2 implies that they are orthogonal to S_2 and $S_2 \supset S_1$. Since $X_1 = \text{Pr}_{S_1} X_2$, the subspace $S_2 \ominus X_1$ is orthogonal to X_2 , and the previous observation implies that it is orthogonal to $\bar{S}_2$. Thus, if $\xi \in \bar{S}_2, \eta \in S_1$ then

$$\begin{aligned} E[\xi\eta] &= E[\text{Pr}_{X_2} \xi \cdot \text{Pr}_{X_1} \eta] = \\ &= \text{cov}(\xi, x_2(0)^T) P_2^{-1} \text{cov}(x_2(0), x_1(0)) P_1^{-1} \text{cov}(x_1(0), \eta). \end{aligned}$$

The assumption

$$x_1(0) = \text{Pr}_{S_1} x_2(0)$$

implies that

$$\text{cov}(x_2(0), x_1(0)) = P_1.$$

Using that $\bar{x}_2(0) = P_2^{-1}x_2(0)$ we obtain that

$$E[\xi\eta] = E[\xi x_2(0)^T] E[x_1(0)\eta] \quad (5.54)$$

concluding the proof. ■

REMARK 5.36. *Instead of $S_2 \supset S_1$ we may assume that $\bar{S}_1 \supset \bar{S}_2$ and $\text{Pr}_{\bar{S}_2} \bar{x}_1(0) = \bar{x}_2(0)$. Then again the covariances with $\bar{x}_2(0)$ and $x_1(0)$, respectively, define an inner product representation of $\bar{S}_2$ and S_1 , where now $x_1(0) = \bar{P}_1^{-1}\bar{x}_1(0)$.*

COROLLARY 5.37. *Under the assumptions of the previous theorem (or remark) we have that*

$$E[\bar{x}_2(0)x_1(0)^T] = \bar{P}_2 P_1. \quad (5.55)$$

Proof. The inner product representation can be applied to the components of $\bar{x}_2(0)$ and $x_1(0)$ because they are in $\bar{S}_2$ and S_1 , respectively. ■

COROLLARY 5.38. Cf. *Green* (Green, 1988) for continuous-time systems and *Michaletsky* (Michaletsky, 1992) for discrete-time systems. *Under the assumptions of the previous theorem (or remark) the realizations defined by the processes x_1 and $\bar{x}_2$ can be stochastically balanced, i.e. choosing an appropriate (minimal rank) factorization of the covariance sequence $R(k)$, $k \geq 1$ in the form $R(k) = CA^{k-1}G$ the corresponding covariance matrices of $x_1(0)$ and $\bar{x}_2(0)$ are equal and diagonal.*

Proof. Since the transformation $C \mapsto CT$, $A \mapsto T^{-1}AT$, $G \mapsto T^{-1}G$ maps the covariance matrix P_1 into $T^{-1}P_1T^{-1T}$ and $\bar{P}_2$ into $T^T\bar{P}_2T$ the Lemma 3.28 can be applied. Thus choosing the nonsingular matrix T in an appropriate manner we can achieve that P_1 and $\bar{P}_2$ be equal and diagonal. Furthermore, an immediate consequence of the previous corollary that the diagonal elements in this stochastically balanced realization are the square roots of the eigenvalues of the product $P_1\bar{P}_2$. ■

In Chapter 3 we have defined the stochastically balanced realization of the forward and backward innovation realizations (see Definition 3.27). In this case the covariance matrices of $x_-(0)$ and $\bar{x}_+(0)$ have this special diagonal form. Thus the previous corollary generalizes this notion.

Remark that the transformation T "balancing" a pair of forward and backward realization depends on these realizations, in general. In other words, balancing e.g. the (forward and backward) innovation realizations does not mean that all possible pairs of forward and backward realizations are immediately balanced at the same time.

Obviously, the equation $\Pr_{S_1} x_2(0) = x_1(0)$ used in the previous theorem implies that $P_2 \leq P_1$. It can be proved that this inequality between the covariance matrices can be formulated in a coordinate-free way.

DEFINITION 5.39. (See Definition 6.1 in *Lindquist and Picci* (Lindquist and Picci, 1990).) *Given two minimal Markovian splitting subspaces, X_1, X_2 , let $X_1 < X_2$ denote the ordering*

$$\|\Pr_{X_1} \xi\| \leq \|\Pr_{X_2} \xi\| \quad \text{for all } \xi \in H^+. \quad (5.56)$$

THEOREM 5.40. *Assume that the coordinates of $x_+(0)$ form a basis in $X_+ = \Pr_{H^+} H^-$ and define $x(0) = \Pr_X x_+(0)$ for every minimal Markovian splitting subspace X and denote $P = E(x(0)x(0)^T)$ the corresponding covariance matrix. Then*

$$X_1 < X_2 \text{ if and only if } P_1 < P_2.$$

In Chapter 7 we need the following theorem.

THEOREM 5.41. (See Theorem 6.8 in *Lindquist and Picci* (Lindquist and Picci, 1990).) *Let the minimal Markovian splitting subspaces X_1 and X_2 be generated by the pair of subspaces $(S_1, \bar{S}_1), (S_2, \bar{S}_2)$. Then*

1. *If X_1 is internal, then*

$$X_1 < X_2 \text{ if and only if } S_1 \subset S_2.$$

2. *If X_2 is internal, then*

$$X_1 < X_2 \text{ if and only if } \bar{S}_2 \subset \bar{S}_1.$$

Note that under the assumptions of the previous theorem $S_1 \subset S_2$ implies that $\Pr_{S_1} X_2 = X_1$ and $\bar{S}_2 \subset \bar{S}_1$ implies that $\Pr_{\bar{S}_2} X_1 = X_2$.

5.6. Acausal state-space representations

Until now we have analysed how to construct forward and backward realizations from a Markovian splitting subspace X leading to analytic (causal) and coanalytic (anticausal) spectral factors. This is very natural if we start with the factorization of the *Hankel*-matrix since this gives a stable state-transition matrix A , i.e. $\rho(A) < 1$. But looking at the geometric setup, we have seen that the state-space itself does not define the type of the realization, whether it is forward or backward one. This is determined by another assumption, namely, we obtained a forward realization

assuming that $S = H^-(w)$ which contains X and $H^-(y)$, as well, leading to an analytic spectral factor; the assumption $\bar{S} = H^+(\bar{w}) \supset H^+(\bar{y})$ leads to a backward realization or coanalytic spectral factor.

Let us approach the same problem from an other point of view. Assume that we have a state-space realization of y

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

$$y(t) = Cx(t) + Dw(t),$$

where $w(t)$, $t \in \mathbb{Z}$ is a white noise process; $x(t)$, $t \in \mathbb{Z}$ is a stationary process, but we do not assume anything about the spectrum of A (except that it has no eigenvalues of modulus 1), in other words about the orthogonality properties between the state process and the input noise.

Straightforward calculation gives that $x(0)$ defines a Markovian splitting subspace. Obviously, in general neither of the inclusions $X \subset H^-(w)$, $X \subset H^+(w)$ hold.

Again direct calculation gives that

$$X \cap H^-(w) \supset \Pr_{H^-(w)} H^+ \quad (5.57)$$

and

$$X \cap H^+(w) \supset \Pr_{H^+(w)} H^-. \quad (5.58)$$

An obvious geometric reformulation of these observations leads to the following theorem (see *Michaletzky and Ferrante* (Michaletzky and Ferrante, 1995)).

THEOREM 5.42. Suppose that we are given a Hilbert-space $H \supset H_0$, with a shift operator U extending the shift on H_0 . Assume that $\mathcal{A}$, $\mathcal{B}$ are closed subspaces of H such that $H = \mathcal{A} \oplus \mathcal{B}$.

Define

$$X_s = \Pr_{\mathcal{A}} H^+, \quad X_u = \Pr_{\mathcal{B}} H^-, \quad X = X_s \oplus X_u.$$

Then

1. X is a splitting subspace, i.e. H^- and H^+ are conditionally orthogonal given X .
2. if $U^{-1}\mathcal{A} \subset \mathcal{A}$, $UB \subset \mathcal{B}$, then X is a Markovian splitting subspace.
3. if $\mathcal{A} \ominus U^{-1}\mathcal{A}$ is finite dimensional and $\bigcap_t U^{-t}\mathcal{A} = \{0\}$, $\bigcap_t U^t\mathcal{B} = \{0\}$, then there exists an acausal realization of y , with an input noise $w(t)$, $t \in \mathbb{Z}$ for which

$$\mathcal{A} = H^-(w), \quad \mathcal{B} = H^+(w). \quad (5.59)$$

This realization defines a spectral factor with no specification on the pole structure.

Proof. The proof of this Theorem is sheer computation, so we omit it. ■

REMARK 5.43. *It is easy to see that defining*

$$S = \mathcal{A} \oplus X_u, \quad \bar{S} = \mathcal{B} \oplus X_s \quad (5.60)$$

we get the usual $(S, \bar{S})$ pair of subspaces connected with the Markovian splitting subspace X .

REMARK 5.44. *Since in the analysis of acausal spectral factors we use the same Markovian splitting subspaces as earlier, the same notion of minimality can easily be applied. Let us mention an immediate corollary of this.*

COROLLARY 5.45. *If the Markovian splitting subspace X determined by the pair of orthogonal subspaces $\mathcal{A}, \mathcal{B}$ is minimal, then*

$$N^+ \perp \mathcal{A}, \quad N^- \perp \mathcal{B}. \quad (5.61)$$

5.7. Zeros of spectral factors

Let us close this chapter with some remarks on the zeros of a spectral factor (cf. Ball, Gohberg and Rodman (Ball et al., 1988), Michaletzky and Ferrante (Michaletzky and Ferrante, 1995)). Assume that $W_a(z)$ is an acausal spectral factor of the form

$$W_a(z) = D + C_1(z^{-1}I - A_1)^{-1}B_1 + \bar{C}_2(zI - \bar{A}_2)^{-1}\bar{B}_2, \quad (5.62)$$

where $A_1, \bar{A}_2$ are stable matrices and this form of W_a defines a minimal realization of W_a with the state-space equations:

$$\begin{aligned} x_1(t+1) &= A_1 x_1(t) + B_2 w(t), \\ x_2(t) &= \bar{A}_2 x_2(t+1) + \bar{B}_2 w(t), \\ y(t) &= C_1 x_1(t) + \bar{C}_2 x_2(t+1) + D w(t). \end{aligned} \quad (5.63)$$

DEFINITION 5.46. *The row vector-valued complex function $\phi(z)$ is called left zero function of W_a at z_0 if it is analytic in a neighbourhood of z_0 and*

$$\lim_{z \rightarrow z_0} \phi(z) W_a(z) = 0. \quad (5.64)$$

The order of the zero function is the largest positive integer k , for which

$$\lim_{z \rightarrow z_0} [(z - z_0)^{-k+1} \Phi(z) W_a(z)] = 0. \quad (5.65)$$

Obviously, it is enough to analyse the zero functions of order k in the form

$$\sum_{i=0}^{n+k-1} a_i^T (z - z_0)^i \quad (5.66)$$

in the sense that a function ϕ is a left zero function if and only if the polynomial defined by the formula above using $a_i^T = (1/i)\phi(z_0)^{(i)}$ as coefficients is a left zero function. The vectors a_i^T , $i = 1, \dots, k$ form a so-called *left zero chain* of $W(z)$ at z_0 .

Denote shortly by M the matrix with rows a_i^T , $i = 1, \dots, k$. Also, let Λ be the *Jordan-block* of order k determined by z_0 :

$$\Lambda = \begin{bmatrix} z_0 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \vdots \\ & \ddots & \ddots & 0 \\ & & 1 & z_0 \end{bmatrix}. \quad (5.67)$$

THEOREM 5.47. (See *Michaletsky and Ferrante* (Michaletsky and Ferrante, 1995).) *The matrix M determines a left zero chain of W at z_0 if and only if there exist two matrices Π_1 , Π_2 satisfying the matrix equation*

$$[\Lambda \Pi_1 \ \Pi_2 \ M] \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & \bar{A}_2 & \bar{B}_2 \\ C_1 & \bar{C}_2 & D \end{bmatrix} = [\Pi_1 \ \Lambda \Pi_2 \ 0]. \quad (5.68)$$

Proof. By symmetry we may assume that $I - z_0 A_1$ is a regular matrix. Also assume that z_0 is different from ∞ . The proof goes by induction. Let ϕ be a left zero function at z_0 . In view of 5.64

$$a_0^T D + a_0^T C_1 z_0 (I - z_0 A_1)^{-1} B_1 + \lim_{z \rightarrow z_0} \phi(z) \bar{C}_2 (z_0 I - \bar{A}_2)^{-1} \bar{B}_2 = 0. \quad (5.69)$$

Denote

$$\begin{aligned} \xi_0^T &= a_0^T C_1 (I - z_0 A_1)^{-1}, \\ \eta_0(z)^T &= \phi(z) \bar{C}_2 (zI - \bar{A}_2)^{-1}. \end{aligned}$$

We have

$$a_0^T C_1 + z_0 \xi_0^T A_1 = \xi_0^T \quad (5.70)$$

$$\phi(z) \bar{C}_2 + \eta_0(z)^T \bar{A}_2 = z \eta_0(z)^T \quad (5.71)$$

Equation 5.69 gives that $\lim_{z \rightarrow z_0} \eta_0(z)^T \bar{B}_2$ exists. Multiplying 5.71 by $\bar{B}_2$ from the right we get that $\lim_{z \rightarrow z_0} \eta_0(z)^T \bar{A}_2 \bar{B}_2$ exists. Multiplying consecutively by $\bar{A}_2 \bar{B}_2$, $\bar{A}_2^2 \bar{B}_2$, ... we obtain that

$$\lim_{z \rightarrow z_0} \eta_0(z)^T C(\bar{A}_2, \bar{B}_2)$$

exists. Controllability of $(\bar{A}_2, \bar{B}_2)$ implies that $\lim_{z \rightarrow z_0} \eta_0(z)^T = \eta_0^T$ exists and satisfies the system of equations

$$\begin{aligned} a_0^T \bar{C}_2 + \eta_0^T \bar{A}_2 &= z_0 \eta_0^T \\ a_0^T D + z_0 \xi_0^T B_1 + \eta_0^T \bar{B}_2 &= 0. \end{aligned} \quad (5.72)$$

Then $\Pi_1 = [\xi_0^T]$, $\Pi_2 = [\eta_0^T]$, $M = [a_0^T]$ gives the solution of 5.68 at z_0 .

Consider the case when $k \geq 2$. Write ϕ in the form

$$\phi(z) = a_0^T + (z - z_0)\phi_1(z).$$

Equations 5.70 and 5.72 give that

$$\begin{aligned} \phi(z)W(z) &= (z - z_0)[\phi_1(z)D + (\phi_1(z)zC_1 + \xi_0^T)(I - zA_1)^{-1}B_1 \\ &\quad + (\phi_1(z)\bar{C}_2 - \eta_0^T)(zI - \bar{A}_2)^{-1}\bar{B}_2]. \end{aligned}$$

Denoting

$$\begin{aligned} \xi_1^T &= (a_1^T C_1 + \xi_0^T A_1)(I - z_0 A_1)^{-1}, \\ \eta_1(z)^T &= (\phi_1(z)\bar{C}_2 - \eta_0^T)(zI - \bar{A}_2)^{-1}, \end{aligned}$$

we have similarly as before

$$\begin{aligned} a_1^T C_1 + (\xi_0^T + z_0 \xi_1^T)A_1 &= \xi_1^T \\ \phi_1(z)\bar{C}_2 + \eta_1(z)^T \bar{A}_2 &= z\eta_1(z)^T + \eta_0^T. \end{aligned} \quad (5.73)$$

$$a_1^T D + (\xi_0^T + z_0 \xi_1^T)B_1 + \lim_{z \rightarrow z_0} \eta_1(z)^T \bar{B}_2 = 0,$$

implying that $\lim_{z \rightarrow z_0} \eta_1(z)^T \bar{B}_2$ exists. Multiplying the second equation of 5.73 by $\bar{B}_2$, $\bar{A}_2 \bar{B}_2$, ... we obtain that $\lim_{z \rightarrow z_0} \eta_1(z)^T C(\bar{A}_2, \bar{B}_2)$, consequently

$$\lim_{z \rightarrow z_0} \eta_1(z)^T = \eta_1^T$$

exists. Obviously

$$\begin{aligned} a_1^T \bar{C}_2 + \eta_1^T \bar{A}_2 &= z_0 \eta_1^T + \eta_0^T \\ a_1^T D + (\xi_0^T + z_0 \xi_1^T)B_1 + \eta_1^T \bar{B}_2 &= 0. \end{aligned}$$

Using the matrices

$$\Pi_1 = \begin{bmatrix} \xi_0^T \\ \xi_1^T \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} \eta_0^T \\ \eta_1^T \end{bmatrix}, \quad M = \begin{bmatrix} a_0^T \\ a_1^T \end{bmatrix}$$

we obtain the solution of 5.68 for $k = 2$ at z_0 . Similar argument can be applied for $k \geq 3$. Let us point out that up to now we have used only the controllability of $(\bar{A}_2, \bar{B}_2)$.

Conversely, let us assume that equation 5.68 holds, and construct a left zero function of order at least k at z_0 . To this aim it is enough to determine $a_i^T, i = 0, \dots, n+k-1$. Let the rows of the matrix M define the first k elements of this sequence. We define two other sequences $\xi_i, i = 0, \dots, n+k-1, \eta_i, i = 0, \dots, n+k-1$ as follows. Again the first k elements are given by the row vectors of the matrices Π_1, Π_2 , respectively. The remaining elements of these sequences can be defined successively as solutions of the equations

$$\begin{aligned} a_i^T C_1 - \xi_i^T (I - z_0 A_1) &= -\xi_{i-1}^T A_1 \\ a_i^T \bar{C}_2 - \eta_i^T (z_0 I - \bar{A}_2) &= \eta_{i-1}^T \end{aligned} \quad (5.74)$$

$i = k, \dots, (n+k-1)$. (Equation 5.68 implies that 5.74 are valid for $i = 1, \dots, k-1$.) The observability of the pair $(C_1, A_1), (\bar{C}_2, \bar{A}_2)$ implies that 5.74 do have solutions.

Defining $\phi(z)$ by 5.66 we obtain that

$$\begin{aligned} \phi(z)W(z) &= \sum_{i=0}^{n+k-1} a_i^T (z - z_0)^i D \\ &\quad + \sum_{i=0}^{n+k-1} a_i^T (z - z_0)^i C_1 z (I - z A_1)^{-1} B_1 \\ &\quad + \sum_{i=0}^{n+k-1} a_i^T (z - z_0)^i \bar{C}_2 (z I - \bar{A}_2)^{-1} \bar{B}_2. \end{aligned}$$

Using the equations

$$a_0^T C_1 = \xi_0^T (I - z_0 A_1), \quad a_0^T \bar{C}_2 = \eta_0^T (z_0 I - \bar{A}_2)$$

we obtain that

$$a_0^T C_1 z (I - z A_1)^{-1} B_1 - z_0 \xi_0^T B_1 = (z - z_0) \xi_0^T (I - z A_1)^{-1} B_1$$

and

$$a_0^T \bar{C}_2 (z I - \bar{A}_2)^{-1} \bar{B}_2 - \eta_0^T \bar{B}_2 = -(z - z_0) \eta_0^T (z I - \bar{A}_2)^{-1} \bar{B}_2.$$

Consequently the first terms containing a_0^T can be eliminated as follows:

$$\begin{aligned}
 a_0^T D + a_0^T C_1 z (I - z A_1)^{-1} B_1 + a_0^T \bar{C}_2 (z I - \bar{A}_2)^{-1} \bar{B}_2 \\
 = a_0^T D + z_0 \xi_0^T B_1 + \eta_0^T \bar{B}_2 \\
 + (z - z_0) \xi_0^T (I - z A_1)^{-1} B_1 - (z - z_0) \eta_0^T (z I - \bar{A}_2)^{-1} \bar{B}_2 \\
 = (z - z_0) [\xi_0^T (I - z A_1)^{-1} - \eta_0^T (z I - \bar{A}_2)^{-1} \bar{B}_2].
 \end{aligned}$$

From 5.68 and 5.74 it follows that

$$(z a_1^T C_1 + \xi_0^T) (I - z A_1)^{-1} B_1 - (z_0 \xi_1 + \xi_0) B_1 = (z - z_0) \xi_1^T (I - z A_1)^{-1} B_1$$

and

$$(a_1^T \bar{C}_2 - \eta_0^T) (z I - \bar{A}_2)^{-1} \bar{B}_2 - \eta_1^T \bar{B}_2 = -(z - z_0) \eta_1^T (z I - \bar{A}_2)^{-1} \bar{B}_2.$$

Thus the terms containing a_1^T can be written as

$$\begin{aligned}
 a_1^T D + a_1^T C_1 z (I - z A_1)^{-1} B_1 + a_1^T \bar{C}_2 (z I - \bar{A}_2)^{-1} \bar{B}_2 \\
 + \xi_0^T (I - z A_1)^{-1} B_1 - \eta_0^T (z I - \bar{A}_2)^{-1} \bar{B}_2 = \\
 = a_1^T D + (\xi_1^T z_0 + \xi_0^T) B_1 + \eta_1^T \bar{B}_2 \\
 + (z - z_0) \xi_1^T (I - z A_1)^{-1} B_1 - (z - z_0) \eta_1^T (z I - \bar{A}_2)^{-1} \bar{B}_2 \\
 = (z - z_0) [\xi_1^T (I - z A_1)^{-1} - \eta_1^T (z I - \bar{A}_2)^{-1} \bar{B}_2].
 \end{aligned}$$

Continuing in this way after k steps we arrive at the equation

$$\begin{aligned}
 \phi(z) W(z) = (z - z_0)^k [\xi_{k-1}^T (I - z A_1)^{-1} B_1 - \eta_{k-1}^T (z I - \bar{A}_2)^{-1} \bar{B}_2] \\
 + \sum_{i=k}^{n+k-1} (z - z_0)^i a_i^T W(z).
 \end{aligned}$$

Note that in the present case $I - z_0 A_1$ is nonsingular, so it is enough to transform the term containing $(z I - \bar{A}_2)^{-1}$ further. Using successively the identity

$$(a_i^T \bar{C}_2 - \eta_{i-1}^T) (z I - \bar{A}_2)^{-1} \bar{B}_2 - \eta_i^T \bar{B}_2 = -(z - z_0) \eta_i^T (z I - \bar{A}_2)^{-1} \bar{B}_2$$

we obtain finally that

$$\begin{aligned}
 \phi(z) W(z) = (z - z_0)^k \xi_{k-1}^T (I - z A_1)^{-1} B_1 \\
 - (z - z_0)^{n+k} \eta_n^T (z I - \bar{A}_2)^{-1} \bar{B}_2
 \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=k}^{n+k-1} (z - z_0)^i a_i^T C_1 (I - zA_1)^{-1} B_1 \times \\
& \quad \times \sum_{i=k}^{n+k-1} (z - z_0)^i (a_i^T D + \eta_i^T \overline{B}_2)
\end{aligned}$$

proving 5.65. ■

REMARK 5.48. Introducing the notation $\Pi = [-\Pi_1 \ \Pi_2]$ we get that the previous equation is equivalent to

$$\Lambda \Pi x(t+1) = \Pi x(t) + My(t). \quad (5.75)$$

REMARK 5.49. In the case, when $z_0 = \infty$ similar argument can be used. Denote Λ_0 the Jordan block of order k determined by the eigenvalue 0. Then instead of 5.68 we get that

$$\begin{bmatrix} \Pi_1 & \Lambda_0 \Pi_2 & M \end{bmatrix} \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & \overline{A}_2 & \overline{B}_2 \\ C_1 & \overline{C}_2 & D \end{bmatrix} = \begin{bmatrix} \Lambda_0 \Pi_1 & \Pi_2 & 0 \end{bmatrix}, \quad (5.76)$$

which is equivalent to

$$\Pi x(t+1) = \Lambda_0 \Pi x(t) + My(t). \quad (5.77)$$

REMARK 5.50. An immediate consequence of this identity is that the zeros are determined by the state-space X . The components of $x(0)$ define only a coordinate system in X not affecting the value of Λ or Λ_0 in equations 5.75 or 5.77. Stated in another form, spectral factors with the same minimal Markovian splitting subspace share the same zeros.

REMARK 5.51. Also in the light of the previous identity it is not surprising that the stable zeros (of modulus not greater than one) are determined by $X \cap H^+$, the unstable zeros (of modulus not less than one) are determined by $X \cap H^-$.

Especially, the zeros of modulus exactly 1 are given by $X \cap H^- \cap H^+$, see Hannan and Poskitt (Hannan and Poskitt, 1988), Green (Green, 1988), Michaletzky (Michaletzky, 1992).

6. The structure of all-pass functions

In this chapter we first analyse the structure of inner functions giving a geometric proof of the characterizing algebraic equation. This geometric approach enables us to give a short concise proof of the factorization theorem. Next we consider the all-pass case. The basic idea is again that the characterizing algebraic equation can be transformed into a geometric one, namely it is equivalent to the existence of an orthogonal transformation in a finite dimensional subspace determined by state vector and the output process. A little bit more involved analysis leads to the factorization theorem of all-pass functions. These factorizations are determined by solutions of two simultaneous *Riccati*-equations. As a side result a direct geometric proof of the fact that in case of the phase function the minimal all-pass factorizations define minimal spectral factors and vice versa is obtained.

Let us remark that because in this chapter the output process $y(t)$ is usually a white noise process, instead of denoting it by y we shall use the notation v .

6.1. Factorization of inner functions

In this section we address the following problems: What is a general form an inner function? How to characterize the inner factors of an inner function?

THEOREM 6.1. *Let $\Phi(z) = D + C(z^{-1}I - A)^{-1}B$ be a $p \times p$ rational function. Assume that $\rho(A) < 1$, so Φ is analytic (inside the unit circle). Let P be the solution of the Ljapunov-equation*

$$P = APA^T + BB^T. \quad (6.1)$$

Then if the equation

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} = \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \quad (6.2)$$

holds then Φ is an inner function. Conversely, if Φ is an inner function and (C, A) is observable then the equation 6.2 holds.

Proof. Consider a p -dimensional white noise process $w(t)$, $t \in \mathbb{Z}$ and via the equations

$$x(t+1) = Ax(t) + Bw(t)$$

$$v(t) = Cx(t) + Dw(t)$$

define the stationary processes $x(t)$, $t \in \mathbb{Z}$ and $v(t)$, $t \in \mathbb{Z}$. Take the covariances of both sides

$$\begin{aligned} E \begin{bmatrix} x(t+1) \\ v(t) \end{bmatrix} \begin{bmatrix} x(t+1)^T & v(t)^T \end{bmatrix} &= \\ &= \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix}. \end{aligned} \quad (6.3)$$

So the equation above is equivalent to

$$\begin{aligned} E[x(t+1)v(t)^T] &= 0 \\ E[v(t)v(t)^T] &= I_p. \end{aligned} \quad (6.4)$$

First assume that these are fulfilled. Then $x(t+1)$ is orthogonal to $v(t)$. Using the state space equation and the stability of A we get that $x(t+2)$, $x(t+3)$, $\dots$ are also orthogonal to $v(t)$, i.e.

$$E[x(t+k)v(t)^T] = 0, \quad k \geq 1.$$

The observation equation now implies that

$$E[v(t+k)v(t)^T] = 0, \quad k \geq 1,$$

i.e. $v(t)$, $t \in \mathbb{Z}$ is a white noise, thus Φ is an inner function.

Now, let us assume the observability of the pair (C, A) and Φ be an inner function, i.e. $v(t)$, $t \in \mathbb{Z}$ is a white noise. Then obviously

$$E[v(t)v(t)^T] = I_p.$$

On the other hand

$$R(t) = E[v(s+t)v(s)^T] = 0, \quad \text{if } t \geq 1.$$

But

$$R(t) = CA^{t-1}G.$$

So the observability of (C, A) implies that

$$G = 0.$$

This concludes the proof because

$$G = E[x(t+1)v(t)^T].$$

REMARK 6.2. Since $H(w) = H(v)$ we obtained that under that assumptions of the previous theorem the inclusion $X \subset H^+(v)$ holds. ■

REMARK 6.3. The state space X of an observable realization of Φ can be written as

$$X = E^{H^-(w)} H^+(v). \quad (6.5)$$

If Φ is an inner function then $X \subset H^+(v)$ and (6.5) imply that $X = H^+(v) \cap H^-(w)$ thus $H^+(v)$ and $H^-(w)$ are conditionally orthogonal subspaces with respect to their intersection X and especially $H^-(w) = H^-(v) \oplus X$.

This implies that the inverse transfer function Φ^{-1} which "maps" $v(t)$ to $w(t)$ has a realization with the same state space X , because

$$X = E^{H^+(v)} H^-(w). \quad (6.6)$$

Now we analyse the problem of factorization of an inner function Φ as $\Phi = \Phi_1 \Phi_2$ where Φ_1, Φ_2 are also inner functions. Φ_1 is called *left inner factor* of Φ .

THEOREM 6.4. Assume that $\Phi(z) = D + C(z^{-1}I - A)^{-1}B$ a $p \times p$ rational inner function. $\rho(A) < 1$. Moreover (C, A) is observable and (A, B) is controllable. Then Φ_1 is a left inner factor of Φ if and only if there are solutions P_1, B_1, D_1 of the equation

$$\begin{bmatrix} A & B_1 \\ C & D_1 \end{bmatrix} \begin{bmatrix} P_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B_1^T & D_1^T \end{bmatrix} = \begin{bmatrix} P_1 & 0 \\ 0 & I \end{bmatrix} \quad (6.7)$$

and

$$\Phi_1(z) = D_1 + C(z^{-1}I - A)^{-1}B_1. \quad (6.8)$$

Proof. First consider a solution of the equation above. We know already that this defines an inner function Φ_1 . We have to prove that it is a left factor of Φ . Take again a state-space system corresponding to the matrices (A, B_1, C, D_1) but assume that the process $v(t)$, $t \in \mathbb{Z}$ is matched to the output process constructed in the previous theorem (this can be done because both are p -dimensional white noises). So

$$\begin{aligned} x_1(t+1) &= Ax_1(t) + B_1 w_1(t), \\ v(t) &= Cx_1(t) + D_1 w_1(t). \end{aligned} \quad (6.9)$$

Denote X_1 the state space of the process $x_0(t)$, $t \in \mathbb{Z}$. Then from Remark 6.3 it follows that $X, X_1 \subset H^+(v)$ and

$$\begin{aligned} H^-(w) &= X \oplus H^-(v) \\ H^-(w_0) &= X_1 \oplus H^-(v). \end{aligned}$$

We show that $X_1 \subset X$ holds. Define a linear transformation

$$T: X \rightarrow X_1$$

for which the components of

$$CA^k x(0)$$

are mapped to the corresponding components of

$$CA^k x_1(0).$$

This has a linear extension because if

$$\sum_{t \geq 0} \lambda(t) CA^T x(0) = 0$$

then

$$\sum_{t \geq 0} \lambda(t) CA^T = 0$$

so

$$\sum_{t \geq 0} \lambda(t) CA^T x_1(0) = 0.$$

Here we have used that the controllability of (A, B) implies that $P > 0$. Observe that

$$CA^k x(0) = \text{Pr}_X v(k)$$

and

$$CA^k x_1(0) = \text{Pr}_{X_0} v(k).$$

The linear extension gives that if $\xi \in H^+(v)$ then

$$T(\text{Pr}_X \xi) = \text{Pr}_{X_1} \xi.$$

Especially if $\xi \in X \subset H^+(v)$ then

$$T\xi = \text{Pr}_{X_1} \xi,$$

i.e. T is a projection operator. (In particular $Tx^i(0) = x_1^i(0)$.) Consequently, for any $\xi \in H^+$

$$\text{Pr}_{X_1} \xi = \text{Pr}_{X_1} (\text{Pr}_X \xi)$$

thus if $\xi \perp X$ then $\xi \perp X_1$, in other words

$$X_1 \subset X,$$

consequently

$$H^-(w) \supset H^-(w_1)$$

which means that the transfer function Φ_2 "mapping" $w(t)$ to $w_1(t)$ is inner. Obviously

$$\Phi = \Phi_1 \Phi_2.$$

For the proof of the converse statement see *Fuhrmann* (Fuhrmann, 1981), *Fuhrmann and Hoffmann* (Fuhrmann and Hoffmann, 1998). It is based on the fact which can be found in the standard literature, that every left inner factor is determined by an invariant subspace of X . ■

The factors Φ_1, Φ_2 can be expressed explicitly using the matrix P_1 appearing in the solution of 6.7.

THEOREM 6.5. (cf. *Bart, Gohberg and Kaashoek* (Bart et al., 1984).) Let $\Phi(z) = D + C(z^{-1}I - A)^{-1}B$ be a $p \times p$ rational inner function, $\rho(A) < 1$. Furthermore, assume that (C, A) is observable, (A, B) is controllable. Let P_1 be an arbitrary solution of 6.7, P be the unique positive definite solution. Then

1. $\Pi = P_1 P^{-1}$ is a projection, and

2.

$$\Phi_1(z) = D_1 + C(z^{-1}I - A)^{-1}B_1$$

$$\Phi_2(z) = D_2 + C_2(z^{-1}I - A)^{-1}B$$

define the corresponding factorization $\Phi = \Phi_1 \Phi_2$, where B_1, D_1, C_2, D_2 are solutions of the equation

$$\begin{bmatrix} B_1 \\ D_1 \end{bmatrix} \begin{bmatrix} C_2 & D_2 \end{bmatrix} = \begin{bmatrix} \Pi A(I - \Pi) & \Pi B \\ C(I - \Pi) & D \end{bmatrix}.$$

REMARK 6.6. If the normalizing condition $D = I$ is fulfilled then we may assume that $D_1 = D_2 = I$ so we obtain that

$$B_1 = \Pi B, \quad C_2 = C(I - \Pi)$$

leading to the factorization identities of Bart, Gohberg and Kaashoek (Bart et al., 1984).

Proof. Consider the state-space systems corresponding to Φ and Φ_1 . As we have seen in the proof of the previous theorem we can achieve that

$$H^-(w) \supset H^-(w_1) \supset H^-(v).$$

$$X_1 = \text{Pr}_{H^-(w_1)} H^+(v) = H^-(w_1) \ominus H^-(v)$$

is the state space of Φ_1 . Denote

$$X_2 = \text{Pr}_{H^-(w)} H^+(w_1) = H^-(w) \ominus H^-(w_1)$$

the state space of Φ_2 . Then

$$X = X_1 \oplus X_2$$

and X_2 is an $\text{Pr}_{H^-(w)} U$ -invariant subspace of X . Thus we can construct a nonsingular transformation $T = [T_1 \ T_2]^T$ for which

$$\begin{bmatrix} \hat{x}_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} x(t),$$

where the coordinates of $\hat{x}_1(0)$ and $x_2(0)$ form a basis in X_1 and X_2 , respectively, so the processes $\hat{x}_1(t)$ and $x_2(t)$, $t \in \mathbb{Z}$ are state-processes of Φ_1 and Φ_2 , respectively. (Recall that the process $x_1(t)$ defined in 6.9 is also a state process of a realization of Φ_1 .) Denote

$$U = [U_1 \ U_2] = T^{-1}.$$

We obtain that

$$\begin{aligned} \begin{bmatrix} \hat{x}_1(t+1) \\ x_2(t+1) \end{bmatrix} &= T A T^{-1} \begin{bmatrix} \hat{x}_1(t) \\ x_2(t) \end{bmatrix} + T B w(t) \\ v(t) &= C T^{-1} \begin{bmatrix} \hat{x}_1(t) \\ x_2(t) \end{bmatrix} + D w(t). \end{aligned} \quad (6.10)$$

Since X_2 is invariant the identity $T_2 A U_1 = 0$ holds. The matrix

$$\Pi = U_1 T_1$$

obviously satisfies the identity $\Pi^2 = \Pi$ thus Π is a projection and $I - \Pi = U_2 T_2$. The equation $E[\hat{x}_1(t)x_2(t)^T] = 0$ implies that

$$\Pi P (I - \Pi)^T = 0. \quad (6.11)$$

Observe that $\hat{x}_1(t)$ and $x_2(t)$, $t \in \mathbb{Z}$ determine (minimal) realizations of Φ_1 and Φ_2 . Consequently

$$\begin{aligned} x_2(t+1) &= T_2 A U_2 x_2(t) + T_2 B w(t) \\ w_1(t) &= C_z x_2(t) + D_z w(t) \end{aligned} \quad (6.12)$$

and

$$\begin{aligned} \hat{x}_1(t+1) &= T_1 A U_1 \hat{x}_1(t) + B_x w_1(t) \\ v(t) &= C U_1 \hat{x}_1(t) + D_x w_1(t) \end{aligned} \quad (6.13)$$

with appropriate matrices C_z , D_z , B_x , D_x .

In other words

$$\Phi_1(z) = D_x + C U_1 (z^{-1} I - T_1 A U_1)^{-1} B_x \quad (6.14)$$

and

$$\Phi_2(z) = D_z + C_z (z^{-1}I - T_2AU_2)^{-1} T_2B. \quad (6.15)$$

The identity $T_2AU_1 = 0$ implies that $U_1T_1AU_1 = AU_1$ and $T_2AU_2T_2 = T_2A$, consequently

$$\begin{aligned} \Phi_1(z) &= D_x + C (z^{-1}I - A)^{-1} U_1B_x \\ \Phi_2(z) &= D_z + C_zT_2 (z^{-1}I - A)^{-1} B. \end{aligned} \quad (6.16)$$

The observability of C, A implies that $x_1(t) = U_1\hat{x}_1(t) = \Pi x(t)$ implying that

$$P_1 = \Pi P \Pi^T,$$

which in turn together with (6.11) gives that $\Pi_1 = P_1 P^{-1}$.

Compare now (6.12) and (6.13) to the realization (6.10). We obtain that

$$B_x C_z x_2(t) + B_x D_z w(t) = T_1 A U_2 x_2(t) + T_1 B w(t)$$

and

$$D_x C_z x_2(t) + D_x D_z w(t) = C U_2 x_2(t) + D w(t).$$

The uncorrelatedness of $x_2(t), w(t)$ implies that the equation

$$\begin{bmatrix} B_x \\ D_x \end{bmatrix} \begin{bmatrix} C_z & D_z \end{bmatrix} = \begin{bmatrix} T_1 A U_2 & T_1 B \\ C U_2 & D \end{bmatrix}$$

holds true.

Denote $D_1 = D_x, B_1 = U_1 B_x, D_2 = D_z, C_2 = C_z T_2$. We obtain that

$$\begin{bmatrix} B_1 \\ D_1 \end{bmatrix} \begin{bmatrix} C_2 & D_2 \end{bmatrix} = \begin{bmatrix} \Pi A(I - \Pi) & \Pi B \\ C(I - \Pi) & D \end{bmatrix}$$

holds. ■

6.2. Characterization and all-pass factorization of all-pass functions

Now let us consider the general – all-pass – case. First we would like to prove an extension of Theorems 6.4 and 6.5. After that we characterize the left all-pass factors in terms of solutions of two *Riccati*-equations.

Finally, we show that the minimal factorizations of the so-called phase function corresponding to a stationary process y determine the minimal acausal spectral factors of v .

THEOREM 6.7. Assume that the $p \times q$ transfer function Φ has a realization

$$\Phi = D + C(z^{-1}I - A)^{-1}B$$

where the matrix A has no eigenvalues of modulus 1. If there exists a solution of the equation

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} = \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \quad (6.17)$$

where $P = P^T$, then Φ is an all-pass function.

Proof. By definition

$$\Phi^*(z) = \sum_{k=-\infty}^{+\infty} \varphi_{-k}^T z^k = D^T + B^T(zI - A^T)^{-1}C^T.$$

Denote

$$\Lambda(z) = (z^{-1}I - A)^{-1}B$$

$$\Delta(z) = (zI - A^T)^{-1}C^T$$

$$T(z) = C(z^{-1}I - A)^{-1}.$$

Using that

$$P - APA^T = (I - zA)P + AP(zI - A^T)$$

we obtain

$$\begin{aligned} \Lambda(z)\Phi^*(z) &= \\ &= (z^{-1}I - A)^{-1}B \left[B^T(zI - A^T)^{-1}C^T + D^T \right] = \\ &= (z^{-1}I - A)^{-1} [BB^T + APA^T - P] (zI - A^T)^{-1}C^T + \\ &\quad + zP(zI - A^T)^{-1}C^T + (zI - A^T)^{-1}(APC^T + BD^T) \end{aligned}$$

and

$$\begin{aligned} \Phi(z)\Phi^*(z) &= \\ &= (D + C\Lambda(z))\Phi^*(z) = \\ &= T(z)(BB^T + APA^T - P)\Delta(z) + T(z)(APC^T + BD^T) + \\ &\quad + DD^T + DB^T(zI - A^T)^{-1}C^T + CPz(zI - A^T)^{-1}C^T = \\ &= T(z)(BB^T + APA^T - P)\Delta(z) + T(z)(APC^T + BD^T) + \\ &\quad + (CPA^T + DB^T)\Delta(z) + (CPC^T + DD^T). \end{aligned}$$

So, under the assumption 6.17,

$$\Phi(z)\Phi^*(z) = I_p.$$

Consequently $p \leq q$ and Φ is an all-pass transfer function. ■

REMARK 6.8. *As a side result we obtained that under the assumptions of the previous theorem*

$$\Lambda(z)\Phi^*(z) = zP\Delta(z). \quad (6.18)$$

If $p = q$ then $\Phi^{-1} = \Phi^*(z)$ thus $w(t) = \sum_{s=-\infty}^{\infty} \phi_{-k}^T y(t-s)$ i.e. in terms of the realizations of Φ and Φ^*

$$x(t+1) = Ax(t) + Bw(t)$$

$$v(t) = Cx(t) + Dw(t),$$

$$\bar{x}(t) = A^T \bar{x}(t+1) + C^T y(t)$$

$$w(t) = B^T \bar{x}(t+1) + D^T y(t).$$

Thus (6.18) implies that

$$P\bar{x}(t) = x(t).$$

This observation is the corner-stone of the converse theorem.

THEOREM 6.9. *Assume that*

$$\begin{aligned} x(t+1) &= Ax(t) + Bw(t) \\ v(t) &= Cx(t) + Dw(t) \end{aligned} \quad (6.19)$$

defines a minimal realization of a $p \times p$ all-pass transfer function, where the matrix A has no eigenvalues of modulus 1.

Then there exists a nonsingular matrix P , $P = P^T$, such that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} = \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix}. \quad (6.20)$$

Proof. We may assume that

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

where A_1 and A_2^{-1} are stable matrices (i.e. with eigenvalues of modulus less than 1). Denote n_1 and n_2 the number of stable and unstable eigenvalues of A , respectively.

$n = n_1 + n_2$ is the dimension of $x(t)$. The minimality implies that (A_1, B_1) , (A_2, B_2) are controllable and (C_1, A_1) , (C_2, A_2) are observable, where

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \ C_2]$$

are the corresponding partitions. We may write

$$x_1(t+1) = A_1 x_1(t) + B_1 w(t)$$

$$x_2(t) = A_2^{-1} x_2(t+1) - A_2^{-1} B_2 w(t)$$

$$v(t) = C_1 x_1(t) + C_2 A_2^{-1} x_2(t) + (D - C_2 A_2^{-1} B_2) w(t).$$

The realization above determines an acausal spectral factor of v . According to Theorem 5.35 its state space can be split into two parts

$$X_s = \Pr_{H^-(w)} H^+(v), \quad X_u = \Pr_{H^+(w)} H^-(v) \quad (6.21)$$

defining the stable and unstable parts. Set

$$X = X_s \oplus X_u. \quad (6.22)$$

Similarly, for the inverse transfer function, which "maps" $y(t)$, $t \in \mathbb{Z}$ to $w(t)$, $t \in \mathbb{Z}$, the usual construction can be applied, since in our case $v(t)$ is also a white noise, thus its state space is generated by

$$\bar{X}_s = \Pr_{H^-(v)} H^+(w), \quad \bar{X}_u = \Pr_{H^+(v)} H^-(w), \quad (6.23)$$

and

$$\bar{X} = \bar{X}_s \oplus \bar{X}_u. \quad (6.24)$$

First we prove that

$$X = \bar{X}. \quad (6.25)$$

This is immediate from the identities

$$H^-(w) = X_s \oplus (H^-(w) \cap H^-(v))$$

$$H^+(w) = X_u \oplus (H^+(w) \cap H^+(v))$$

$$H^-(v) = \bar{X}_s \oplus (H^-(w) \cap H^-(v))$$

$$H^+(v) = \bar{X}_u \oplus (H^+(w) \cap H^+(v))$$

giving that

$$X = H(v) \ominus (N^- \oplus N^+) = \bar{X},$$

where

$$N^- = H^-(w) \cap H^-(v), \quad N^+ = H^+(w) \cap H^+(v).$$

Observe that

$$\begin{aligned} n_1 &= \dim X_s = \dim \bar{X}_u, \\ n_2 &= \dim X_u = \dim \bar{X}_s. \end{aligned} \quad (6.26)$$

Denote by V_0 and W_0 the subspace spanned by the coordinates of $v(0)$ and $w(0)$, respectively. Then the subspaces W_0, X_s, UX_u are pairwise orthogonal. Similarly, using that $v(t), t \in \mathbb{Z}$ is a white noise, we obtain that $V_0, \bar{X}_s, U\bar{X}_u$ are pairwise orthogonal, as well. Also,

$$\dim[W_0 \oplus X_s \oplus UX_u] = p + n_1 + n_2 = \dim[V_0 \oplus U\bar{X}_u \oplus \bar{X}_s].$$

Observe, that

$$V_0, UX_s, X_u \subset W_0 \oplus X_s \oplus UX_u$$

and also

$$W_0, \bar{X}_u, U\bar{X}_s \subset V_0 \oplus U\bar{X}_u \oplus \bar{X}_s.$$

This implies that

$$W_0 \oplus X_s \oplus UX_u = V_0 \oplus U\bar{X}_u \oplus \bar{X}_s = (W_0 \vee X = V_0 \vee U\bar{X}).$$

So we can define an orthogonal transformation on the space $W_0 \vee X$ which maps W_0, X_s, UX_u onto $V_0, U\bar{X}_u, \bar{X}_s$ respectively. In order to construct it explicitly first consider an orthogonal transformation on X which maps X_s onto $\bar{X}_u, X_u$ onto $\bar{X}_s$, respectively. Assume that the coordinates of $\tilde{z}_1(0), \tilde{z}_2(0), z_1(0), z_2(0)$ generate $\bar{X}_u, \bar{X}_s, X_s, X_u$, respectively, and moreover they have unit covariance matrices. In view of 6.25 there exists an orthonormal matrix $\mathcal{P}$ such that

$$\begin{bmatrix} \tilde{z}_1(0) \\ \tilde{z}_2(0) \end{bmatrix} = \mathcal{P} \begin{bmatrix} z_1(0) \\ z_2(0) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} z_1(0) \\ z_2(0) \end{bmatrix}.$$

On the other hand

$$z_1(0) = q_1 x_1(0), \quad z_2(0) = q_2 x_2(0)$$

for appropriate nonsingular matrices q_1, q_2 . Using $\mathcal{P}$ the corresponding orthogonal transformation on $W_0 \vee X$ can be defined as

$$\begin{aligned} \tilde{z}_1(1) &= p_{11} z_1(1) + p_{12} z_2(1) = \\ &= p_{11} q_1 A_1 q_1^{-1} z_1(0) + p_{12} z_2(1) + p_{11} q_1 B_1 w(0), \\ \tilde{z}_2(0) &= p_{21} z_1(0) + p_{22} z_2(0) = \\ &= p_{21} z_1(0) + p_{22} q_2 A_2^{-1} q_2^{-1} z_2(1) - p_{22} q_2 A_2^{-1} B_2 w(0), \\ y(0) &= C_1 q_1^{-1} z_1(0) + C_2 A_2^{-1} q_2^{-1} z_2(1) + \left(D - C_2 A_2^{-1} B_2 \right) w(0). \end{aligned}$$

Consequently

$$\begin{bmatrix} p_1 q_1 A_1 q_1^{-1} & p_{12} & p_1 q_1 B_1 \\ p_{21} & p_2 q_2 A_2^{-1} q_2^{-1} & -p_2 q_2 A_2^{-1} B_2 \\ C_1 q_1^{-1} & C_2 A_2^{-1} q_2^{-1} & D - C_2 A_2^{-1} B_2 \end{bmatrix} \quad (6.27)$$

is an orthonormal matrix. We show that p_2 is nonsingular. This is equivalent to saying that

$$X_u^\perp \cap \bar{X}_s = 0,$$

or in other words

$$X_s \cap \bar{X}_s = 0.$$

But

$$H^-(w) = X_s \oplus (H^-(v) \cap H^-(w))$$

$$H^-(v) = \bar{X}_s \oplus (H^-(v) \cap H^-(w))$$

proving the previous identity. Symmetrically, the relationship

$$X_u \cap \bar{X}_u = 0$$

implies that the matrix p_1 is nonsingular.

The orthogonality of 6.27 implies that its *Redheffer*-transform is J-unitary (see *Dym* (Dym, 1988)). Denoting this transform by $\mathcal{R}$:

$$\mathcal{R} = [\mathcal{R}_1 \ \mathcal{R}_2 \ \mathcal{R}_3],$$

where

$$\begin{aligned} \mathcal{R}_1 &= \begin{bmatrix} p_1 q_1 A_1 q_1^{-1} - p_{12} \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} p_{21} & p_{21} \\ - \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} p_{21} & p_{21} \\ C_1 q_1^{-1} - \left(C_2 A_2^{-1} q_2^{-1} \right) \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} p_{21} & p_{21} \end{bmatrix}, \\ \mathcal{R}_2 &= \begin{bmatrix} p_{12} \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} \\ \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} \\ C_2 A_2^{-1} q_2^{-1} \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} \end{bmatrix}, \\ \mathcal{R}_3 &= \begin{bmatrix} p_1 q_1 B_1 - p_{12} \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} \left(-p_2 q_2 A_2^{-1} B_2 \right) \\ \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} p_2 q_2 A_2^{-1} B_2 \\ D - C_2 A_2^{-1} B_2 + C_2 A_2^{-1} q_2^{-1} \left(p_2 q_2 A_2^{-1} q_2^{-1} \right)^{-1} p_2 q_2 A_2^{-1} B_2 \end{bmatrix}. \end{aligned}$$

We can reduce the above formula to

$$\begin{aligned} \mathcal{R} &= \\ &= \begin{bmatrix} p_1 q_1 A_1 q_1^{-1} - p_{12} q_2 A_2 q_2^{-1} p_2^{-1} p_{21} & p_{12} q_2 A_2 q_2^{-1} p_2^{-1} & p_1 q_1 B_1 + p_{12} q_2 B_2 \\ -q_2 A_2 q_2^{-1} p_2^{-1} p_{21} & q_2 A_2 q_2^{-1} p_2^{-1} & q_2 B_2 \\ C_1 q_1^{-1} - C_2 q_2^{-1} p_2^{-1} p_{21} & C_2 q_2^{-1} p_2^{-1} & D \end{bmatrix} \end{aligned}$$

Let us observe that

$$\mathcal{R} \begin{bmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & I \end{bmatrix} \mathcal{R}^T = \begin{bmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & I \end{bmatrix}$$

and

$$\mathcal{R} = \begin{bmatrix} p_1 q_1 & p_2 q_2 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ C_1 & C_2 & D \end{bmatrix} \begin{bmatrix} q_1^{-1} & 0 & 0 \\ -q_2^{-1} p_2^{-1} p_{21} & q_2^{-1} p_2^{-1} & 0 \\ 0 & 0 & I \end{bmatrix}.$$

This implies that

$$\begin{aligned} &\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ C_1 & C_2 & D \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_1^T & 0 & C_1^T \\ 0 & A_2^T & C_2^T \\ B_1^T & B_2^T & D^T \end{bmatrix} = \\ &= \begin{bmatrix} q_1^{-1} q_1^{T-1} & q_1^{-1} p_1^{-1} p_{12} q_2^{T-1} & 0 \\ q_2^{-1} p_{12}^T p_1^{T-1} q_1^{T-1} & -q_2^{-1} q_2^{T-1} & 0 \\ 0 & 0 & I \end{bmatrix}, \end{aligned}$$

using that $p_{21} p_{21}^T + p_2 p_2^T = I$ and $p_1 p_1^T + p_{12} p_{12}^T = I$, where P denotes

$$P = \begin{bmatrix} q_1^{-1} q_1^{T-1} & -q_1^{-1} p_{21}^T p_2^{-1} q_2^{T-1} \\ -q_2^{-1} p_2^{-1} p_{12} q_1^{T-1} & -q_2^{-1} q_2^{T-1} \end{bmatrix}$$

But the orthogonality of $\mathcal{P}$ also implies that

$$p_1 p_{21}^T + p_{12} p_2^T = 0$$

i.e.

$$p_1^{-1} p_{12} = -p_{21}^T p_2^{T-1}.$$

Thus we obtain the equation

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} = \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix}$$

proving the theorem. ■

REMARK 6.10. *Introducing another coordinate system defined as*

$$\begin{bmatrix} \bar{x}_1(0) \\ \bar{x}_2(0) \end{bmatrix} = \begin{bmatrix} q_1^T p_1^T & 0 \\ 0 & -q_2^T p_2^T \end{bmatrix} \begin{bmatrix} \tilde{z}_1(0) \\ \tilde{z}_2(0) \end{bmatrix}$$

we get that

$$\begin{aligned} P\bar{x}(0) &= \begin{bmatrix} q_1^{-1} q_1^{T-1} & -q_1^{-1} p_{21}^T p_2^{T-1} q_2^{T-1} \\ q_2^{-1} p_{12}^T p_1^{T-1} q_1^{T-1} & -q_2^{-1} q_2^{T-1} \end{bmatrix} \\ &\times \begin{bmatrix} q_1^T p_1^T & 0 \\ 0 & -q_2^T p_2^T \end{bmatrix} \begin{bmatrix} p_1 & p_{12} \\ p_{21} & p_2 \end{bmatrix} \begin{bmatrix} q_1 x_1(0) \\ q_2 x_2(0) \end{bmatrix} = x(0). \end{aligned}$$

Consequently, applying both sides of 6.20 to $[\bar{x}(1) \ v(0)]^T$ and solving the equation to $[\bar{x}(0) \ w(0)]^T$ we obtain that

$$\bar{x}(0) = A^T \bar{x}(1) + C^T v(0)$$

$$w(0) = B^T \bar{x}(1) + D^T v(0),$$

giving a backward realization of the transfer function Φ^* .

Since Φ^* is all-pass, as well, Theorem 6.9 implies that

$$\begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} Q & 0 \\ 0 & I \end{bmatrix} \quad (6.28)$$

holds and moreover

$$PQ = I. \quad (6.29)$$

6.3. All-pass factorization of all-pass functions

Let us consider now the question of factorization of all-pass functions in the form

$$\Phi = \Phi_1 \Phi_2,$$

where Φ_1 and Φ_2 are also all-pass functions. Since $\Phi_1^* \Phi$ is all-pass, the problem formulated in this way is too general. Instead of this we may consider the so-called minimal factorizations (see Bart, Gohberg and Kaashoek (Bart et al., 1984)).

DEFINITION 6.11. Assume that Φ , Φ_1 , Φ_2 are rational transfer functions for which $\Phi = \Phi_1 \Phi_2$. This factorization is called minimal, if

$$n = n_1 + n_2,$$

where n , n_1 , n_2 are the dimensions of the state spaces of minimal realizations of Φ , Φ_1 and Φ_2 , respectively.

THEOREM 6.12. Let Φ , Φ_1 , Φ_2 be $p \times p$ rational all-pass transfer functions for which

$$\Phi = \Phi_1 \Phi_2. \quad (6.30)$$

Consider a p -dimensional white noise sequence $w(t)$, $t \in \mathbb{Z}$ and apply the transformation determined by Φ and Φ_2 to the sequence w , defining in this way the stationary sequences $v(t)$, $t \in \mathbb{Z}$ and $u(t)$, $t \in \mathbb{Z}$, respectively. Denote by X , Z , R the state spaces of the minimal realizations of Φ , Φ_1 , Φ_2 in the Hilbert-space $H(w)$.

Then

$$1. X \subset Z \vee R$$

2. the factorization 6.30 is minimal, if and only if

$$H^+(u) \supset H^+(v) \cap H^+(w)$$

and

$$H^+(u) = [H^+(v) \cap H^+(u)] \vee [H^+(w) \cap H^+(u)].$$

Proof. Since Φ and Φ_2 are $p \times p$ all-pass functions we have that

$$H = H(v) = H(w) = H(u).$$

At the same time

$$H \ominus X = (H^-(v) \cap H^-(w)) \oplus (H^+(v) \cap H^+(w)),$$

$$H \ominus Z = (H^-(v) \cap H^-(u)) \oplus (H^+(v) \cap H^+(u)),$$

$$H \ominus R = (H^-(u) \cap H^-(w)) \oplus (H^+(u) \cap H^+(w)).$$

The inclusions

$$[H^-(v) \cap H^-(u)] \cap [H^-(u) \cap H^-(w)] \subset H^-(v) \cap H^-(w)$$

$$[H^+(v) \cap H^+(u)] \cap [H^+(u) \cap H^+(w)] \subset H^+(v) \cap H^+(w)$$

imply that

$$X \subset Z \vee R$$

proving (1) and the inequality $n \leq n_1 + n_2$.

It follows that the factorization is minimal, if and only if

$$X = Z \vee R, \quad (6.31)$$

and

$$Z \cap R = 0. \quad (6.32)$$

The previous argument gives that

$$X = Z \vee R,$$

is equivalent to

$$H^-(u) \supset H^-(v) \cap H^-(w) \quad (6.33)$$

and

$$H^+(u) \supset H^+(v) \cap H^+(w). \quad (6.34)$$

On the other hand

$$Z \cap R = 0$$

means that

$$Z^\perp \vee R^\perp = H$$

i.e.

$$H^+(u) = [H^+(v) \cap H^+(u)] \vee [H^+(w) \cap H^+(u)], \quad (6.35)$$

and

$$H^-(u) = [H^-(v) \cap H^-(u)] \vee [H^-(w) \cap H^-(u)]. \quad (6.36)$$

We prove that 6.34 and 6.35 imply 6.33 and 6.36. Obviously, from 6.35 it follows that

$$H^-(u) = (H^-(v) \vee H^-(u)) \cap (H^-(w) \vee H^-(u)) \supset H^-(v) \cap H^-(w)$$

so (6.30) is fulfilled. Also, taking the intersection of both sides with $H^-(v)$ we obtain that

$$H^-(u) \cap H^-(v) = H^-(v) \cap (H^-(w) \vee H^-(u)).$$

But 6.34 implies that

$$H^-(u) \subset H^-(v) \vee H^-(w).$$

If $a \in H^-(u)$ then in view of the rationality of Φ_1, Φ_2 there exists $b \in H^-, c \in H^-(w)$ such that $a = b + c$. But in this case

$$b = a - c \in H^-(v) \cap (H^-(u) \vee H^-(w)).$$

Consequently, $b \in H^-(u) \cap H^-(v)$. Similarly, $c \in H^-(w) \cap H^-(v)$, proving (6.33). This concludes the proof of (2). ■

REMARK 6.13. *Symmetric argument gives that the minimality of the factorization is equivalent to*

$$H^-(u) \supset H^-(v) \cap H^-(w)$$

$$H^-(u) = [H^-(v) \cap H^-(u)] \vee [H^-(w) \cap H^-(u)].$$

Using the notation

$$N^+ = H^+(v) \cap H^+(w), \quad N^- = H^-(v) \cap H^-(w)$$

we see that an arbitrary minimal factorization is characterized by the pair of subspaces

$$Z_+ = (H^+(v) \cap H^+(u)) \ominus N^+$$

$$R_+ = (H^+(w) \cap H^+(u)) \ominus N^+.$$

Obviously

$$Z_+ \subset H^+(v) \ominus N^+ = \bar{X}_u$$

$$R_+ \subset H^+(w) \ominus N^+ = X_u.$$

But we may write that

$$\bar{X}_u = [H^-(v) \vee H^-(w)] \ominus H^-(v),$$

$$Z_+ = [H^-(v) \vee H^-(w)] \ominus [H^-(v) \vee H^-(u)],$$

$$X_u = [H^-(v) \vee H^-(w)] \ominus H^-(w),$$

$$R_+ = [H^-(v) \vee H^-(w)] \ominus [H^-(w) \vee H^-(u)].$$

Thus these subspaces are invariant under the operator

$$\text{Pr}_{H^-(v) \vee H^-(w)} U|_{H^-(v) \vee H^-(w)}.$$

Since $Z_+ \subset \bar{X}_u$ is an invariant subspace this gives the possibility of characterizing it via a left inner factor of an inner function, or as a solution of a matrix equation.

In order to be able to formulate this statement precisely let us first analyse the factorizations of $\Phi = \Phi_1 \Phi_2$ when Φ_1 is an inner function. Obviously, this is equivalent to the fact that the unstable part of the state space Z is trivial, i.e.

$$Z_u = \text{Pr}_{H^+(u)} H^-(v) = 0.$$

Thus

$$H^+(u) \subset H^+(v).$$

Let us observe that in this case the minimality condition is reduced to

$$H^+(u) \supset N^+.$$

Let us denote by Φ_s the left inner factor of determined by the selection

$$H^+(u) = N^+.$$

In this case the corresponding state space is

$$Z = \Pr_{H^-(u)} H^+(v) = \bar{X}_u.$$

In view of Theorem 6.4 the invariant subspaces of $\bar{X}_u$ can be characterized in terms of the left inner functions of Φ_s .

Denote the right all-pass factor of Φ in the factorization determined by Φ_s as Φ_r . Since $H^+(u) \subset H^+(w)$ this is a coinner function, i.e. the stable part of its state space is trivial,

$$R = R_u = \Pr_{H^+(w)} H^-(u).$$

Taking the adjoint transfer function, Φ_r^* , this is inner, it has the same state space

$$\bar{R} = R.$$

Let us observe that

$$\Pr_{H^+(w)} H^-(u) = H^+(w) \ominus [H^+(w) \cap H^+(u)] = H^+(w) \ominus N^+ = X_u.$$

Furthermore,

$$X_u = \Pr_{H^-(v) \vee H^-(w)} H^-(w).$$

Invoking again Theorem 6.4 we have the possibility of characterizing the invariant subspaces of X_u in terms of left inner factors of Φ_r^* . Figure 6.1 shows the connections between the various subspaces.

THEOREM 6.14. *The left all-pass factors defining minimal factorizations of Φ are determined via solutions of the equation*

$$\begin{bmatrix} A_1 & b_1 \\ C_1 & d_1 \end{bmatrix} \begin{bmatrix} p_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_1^T & C_1^T \\ b_1^T & d_1^T \end{bmatrix} = \begin{bmatrix} p_1 & 0 \\ 0 & I \end{bmatrix} \quad (6.37)$$

and

$$\begin{bmatrix} A_2^{T-1} & c_2 \\ B_2^T A_2^{T-1} & d_2 \end{bmatrix} \begin{bmatrix} p_2 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_2^{-1} & A_2^{-1} B_2 \\ c_2^T & d_2^T \end{bmatrix} = \begin{bmatrix} p_2 & 0 \\ 0 & I \end{bmatrix} \quad (6.38)$$

in terms of b_1, d_1, c_2, d_2 , and p_1, p_2 ($p_1 = p_1^T, p_2 = p_2^T$) $p_1 \geq 0, p_2 \geq 0$.

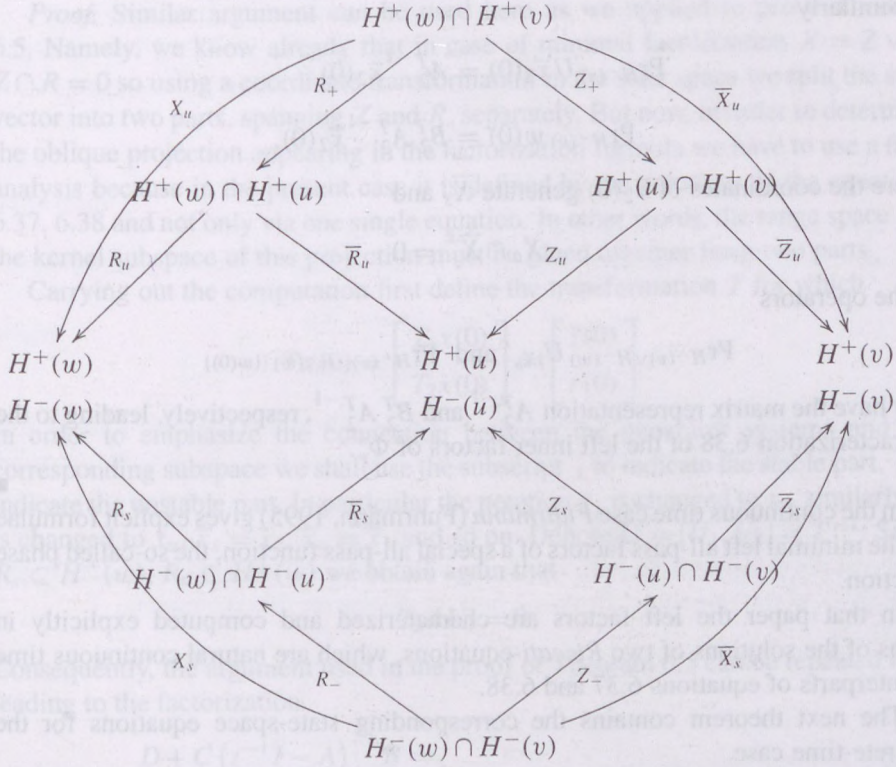


Figure 6.1. Connections between the various subspaces. The subspaces written on the various arrows and bended arrows are denoting the orthogonal difference of the subspaces positioned at the two ends of these arrows

Proof. In view of the argument before this theorem and Theorem 6.4 it is enough to compute the matrices in a minimal realization of Φ_s and Φ_r^* . But

$$\text{Pr}_{H^-(w)} U x_1(0) = A_1 x_1(0)$$

$$\text{Pr}_{H^-(w)} v(0) = C_1 x_1(0),$$

where the coordinates of $x_1(0)$ generate X_s and

$$\bar{X}_u \cap X_s^\perp = 0$$

so for the matrix form of the operators

$$\text{Pr}_{H^-(v) \vee H^-(w)} U|_{\bar{X}_u} \text{ and } \text{Pr}_{H^-(v) \vee H^-(w)} \{v(0)\}$$

we may use the same matrices A_1 and C_1 , respectively. Theorem 6.4 gives that the left inner factors of Φ_s are determined by the solutions of 6.37.

Similarly

$$\Pr_{H^-(v)} U \bar{x}_2(0) = A_2^{T^{-1}} \bar{x}_2(0)$$

$$\Pr_{H^-(v)} w(0) = B_2^T A_2^{T^{-1}} \bar{x}_2(0),$$

where the coordinates of $\bar{x}_2(0)$ generate $\bar{X}_s$ and

$$X_u \cap \bar{X}_s^\perp = 0$$

so the operators

$$\Pr_{H^-(v) \vee H^-(w)} U|_{X_u} \text{ and } \Pr_{H^-(v) \vee H^-(w)} \{w(0)\}$$

may have the matrix representation $A_2^{T^{-1}}$ and $B_2^T A_2^{T^{-1}}$, respectively, leading to the characterization 6.38 of the left inner factors of Φ_r^* . ■

In the continuous time case *Fuhrmann* (Fuhrmann, 1995) gives explicit formulae for the minimal left all-pass factors of a special all-pass function, the so-called phase function.

In that paper the left factors are characterized and computed explicitly in terms of the solutions of two *Riccati*-equations, which are natural continuous time counterparts of equations 6.37 and 6.38.

The next theorem contains the corresponding state-space equations for the discrete-time case.

THEOREM 6.15. Assume that $\Phi(z) = D + C(z^{-1}I - A)^{-1}B$ a $p \times p$ rational all-pass function, where A has no eigenvalues of modulus 1. Furthermore, assume that (C, A) is observable, (A, B) is controllable. Let p_1, p_2 be arbitrary solutions of 6.37 and 6.38 and P is a nonsingular solution of 6.20. $Q = P^{-1}$. Define the projection Π determining its range and kernel subspace as follows:

$$\text{range}(\Pi) = \text{range} \begin{bmatrix} p_1 & 0 \\ 0 & P_2 - P_2 p_2 P_2 \end{bmatrix}$$

$$\ker(\Pi) = \ker \left(\begin{bmatrix} p_1 & 0 \\ 0 & P_2 - P_2 p_2 P_2 \end{bmatrix} Q \right).$$

Then the corresponding minimal factorization $\Phi = \Phi_1 \Phi_2$ is defined as

$$\Phi_1(z) = d_1 + C(z^{-1}I - A)^{-1}b_1$$

$$\Phi_2(z) = d_2 + c_2(z^{-1}I - A)^{-1}B,$$

where

$$\begin{bmatrix} b_1 \\ d_1 \end{bmatrix} \begin{bmatrix} c_2 & d_2 \end{bmatrix} = \begin{bmatrix} \Pi A (I - \Pi) & \Pi B \\ C (I - \Pi) & D \end{bmatrix}.$$

Proof. Similar argument can be used here as we applied to prove Theorem 6.5. Namely, we know already that in case of minimal factorization $X = Z \vee R$, $Z \cap R = 0$ so using a coordinate transformation in the state space we split the state vector into two parts, spanning Z and R , separately. But now, in order to determine the oblique projection appearing in the factorization formula we have to use a finer analysis because in the present case it is defined by the solutions of the equations 6.37, 6.38 and not only via one single equation. In other words, the range space and the kernel subspace of this projection must be glued together form two parts.

Carrying out the computation first define the transformation T for which

$$Tx(0) = \begin{bmatrix} T_1 x(0) \\ T_2 x(0) \end{bmatrix} = \begin{bmatrix} \bar{z}(0) \\ r(0) \end{bmatrix}.$$

In order to emphasize the connection between the generator systems and the corresponding subspace we shall use the subscript s to indicate the stable part, u to indicate the unstable part. In particular the notation x_1 is changed to x_s , similarly x_2 is changed to x_u , $\bar{x}_s = \bar{x}_2$, $\bar{x}_u = \bar{x}_1$ and so on. Denote $U = [U_1, U_2] = T^{-1}$. Since $R_s \subset H^-(w)$, $R_u \subset H^+(w)$ we obtain again that

$$T_2 A U_1 = 0.$$

Consequently, the argument used in the proof of Theorem 6.5 can be repeated here leading to the factorization

$$\begin{aligned} D + C(z^{-1}I - A)^{-1}B &= \\ &= \left(d_1 + C(z^{-1}I - A)^{-1}b_1\right) \left(d_2 + c_2(z^{-1}I - A)^{-1}B\right), \end{aligned}$$

where

$$\begin{bmatrix} b_1 \\ d_1 \end{bmatrix} [c_2 \ d_2] = \begin{bmatrix} \Pi A(I - \Pi) & \Pi B \\ C(I - \Pi) & D \end{bmatrix}.$$

The projection is determined via the formula

$$\Pi = U_1 T_1.$$

Especially

$$\text{range}(\Pi) = \ker(T_2)$$

$$\ker(\Pi) = \ker(T_1).$$

To finish the proof of the theorem we have to determine these subspaces. To this end let us look at the finer structure of the transformation T . Let

$$T = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{13} & T_{14} \\ T_{21} & T_{22} \\ T_{23} & T_{24} \end{bmatrix}.$$

according to the decomposition

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{13} & T_{14} \\ T_{21} & T_{22} \\ T_{23} & T_{24} \end{bmatrix} \begin{bmatrix} x_s(0) \\ x_u(0) \end{bmatrix} = \begin{bmatrix} \bar{z}_u(0) \\ \bar{z}_s(0) \\ r_s(0) \\ r_u(0) \end{bmatrix}$$

$R_u \subset X_u$ and $R_s \subset X_s$ imply that $T_{22} = 0$, $T_{23} = 0$.

Let us consider first the transformations T_{21} , T_{11} , T_{12} determined by the equations

$$T_{21}x_s(0) = r_s(0)$$

$$T_{11}x_s(0) + T_{12}x_u(0) = \bar{z}_u(0).$$

We know that $\bar{X}_u$ is the state space of the maximal left inner factor Φ_s of Φ . (Recall that $\Phi = \Phi_s \Phi_r$.) On the other hand $\bar{X}_u \cap H^+(u)$ is a $\text{Pr}_{H^-(v) \vee H^-(w)} U$ invariant subspace of $\bar{X}_u$. So Theorem 6.4 connects this subspace to a solution of equation 6.37. Furthermore, Theorem 6.5 shows that this solution determines a factorization of Φ_s which is intimately connected to a projection. Consequently to find the kernel subspaces of the operators above we have to determine this projection.

Since

$$\bar{x}_u(0) = Q_1 x_s(0) + Q_{12} x_u(0), \quad (6.39)$$

where

$$Q = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{21} & Q_2 \end{bmatrix} = P^{-1} = \begin{bmatrix} P_1 & P_{12} \\ P_{21} & P_2 \end{bmatrix}$$

choosing $Q_1^{-1} \bar{x}_u(t)$, $t \in \mathbb{Z}$ as the state vector process of a realization of Φ_r , the equations

$$\bar{X}_u = \text{Pr}_{H^-(v) \vee H^-(w)} H^+$$

$$X_s = \text{Pr}_{H^-(w)} H^+$$

assure that the matrix form of the operators $\text{Pr}_{\bar{X}_u} U|_{\bar{X}_u}$ and $\text{Pr}_{\bar{X}_u} |_{\{v(0)\}}$ is A_1 , C_1 .

Furthermore

$$\bar{X}_u \cap H^+(u) = E^{H^-(v) \vee H^-(w)} (H^+(v) \cap H^+(u))$$

and in view of the minimality of the factorization

$$R_s = E^{H^-(w)} (H^+(v) \cap H^+(u))$$

so we obtain that

$$X_s = E^{H^-(w)} \bar{X}_u$$

$$R_s = E^{H^-(w)} (\bar{X}_u \cap H^+(u)).$$

Now consider the solution p_1 of equation 6.37. Applying the argument used in Theorem 6.5 for the factorization of inner functions it is possible to introduce a state-space transformation in $\bar{X}_u$ in coherence with the decomposition

$$\bar{X}_u = \bar{Z}_u \oplus (\bar{X}_u \cap H^+(u))$$

defined by

$$Q_1^{-1} \bar{x}_u(0) \mapsto \begin{bmatrix} t_1 Q_1^{-1} \bar{x}_u(0) \\ t_2 Q_1^{-1} \bar{x}_u(0) \end{bmatrix},$$

where

$$t_1 Q_1^{-1} \bar{x}_u(0) = \bar{z}_u(0).$$

The coordinates of $t_2 Q_1^{-1} \bar{x}_u(0)$ form a basis in $\bar{X}_u \cap H^+(u)$. Moreover projecting $t_2 Q_1^{-1} \bar{x}_u(0)$ onto $H^-(w)$ we obtain a basis in R_s . I.e. we may assume that

$$r_s(0) = \text{Pr}_{H^-(w)} t_2 Q_1^{-1} \bar{x}_u(0) = t_2 x_s(0).$$

using 6.39 and the fact that $x_u(0)$ is orthogonal to $H^-(w)$. Consequently

$$T_{21} = t_2$$

and

$$\begin{bmatrix} T_{11} & T_{12} \end{bmatrix} = t_1 Q_1^{-1} \begin{bmatrix} Q_1 & Q_{12} \end{bmatrix}.$$

At the same time invoking Theorem 6.5 the kernel subspaces of t_1, t_2 can be determined. Namely, the matrix

$$\Pi_1 = p_1 \left[\text{Var} \left(Q_1^{-1} \bar{x}_u(0) \right) \right]^{-1} = p_1 Q_1$$

is an oblique projection and

$$\ker(t_1) = \ker(\Pi_1) = \ker p_1 Q_1$$

and

$$\ker(t_2) = \text{range}(\Pi_1) = \text{range } p_1 Q_1.$$

Similarly, X_u is the state-space of Φ_r^* , where Φ_r is the maximal right co-inner factor of Φ .

$$X_u = E^{H^-(v) \vee H^-(w)} H^+(w)$$

$$\bar{X}_s = E^{H^-(v)} H^+(w)$$

consequently the equation

$$x_u(0) = P_{21} \bar{x}_u(0) + P_2 \bar{x}_s(0) \quad (6.40)$$

gives that – introducing the state-vector process $P_2^{-1}x_u(0)$ – the matrix form of the operators $\Pr_{H^-(v) \vee H^-(w)} U|_{X_u}$ and $\Pr_{H^-(v) \vee H^-(w)} \{w(0)\}$ is $(A_2^T)^{-1}$, $(A_2^T)^{-1} B_2^T$. Considering the solution p_2 of equation (6.35) the coordinate transformation

$$P_2^{-1}x_u(0) \mapsto \begin{bmatrix} l_1 P_2^{-1}x_u(0) \\ l_2 P_2^{-1}x_u(0) \end{bmatrix}$$

determined by the decomposition

$$X_u = R_u \oplus (X_u \cap H^+(u))$$

is connected to the oblique projection

$$\Pi_2 = p_2 \left[\text{Var} \left(P_2^{-1}x_u(0) \right) \right]^{-1} = p_2 P_2.$$

Especially, let

$$r_u(0) = l_1 P_2^{-1}x_u(0)$$

so

$$T_{24} = l_1 P_2^{-1}.$$

Also

$$X_u \cap H^+(u) = \Pr_{H^-(v) \vee H^-(w)} (H^+(u) \cap H^+(w))$$

$$\bar{Z}_s = \Pr_{H^-(v)} (H^+(u) \cap H^+(w)),$$

consequently

$$\bar{Z}_s = \Pr_{H^-(v)} [X_u \cap H^+(u)].$$

The coordinates of $l_2 P_2^{-1}x_u(0)$ form a basis in $X_u \cap H^+(u)$. Projecting this to $H^-(v)$ we get a basis in $\bar{Z}_s$. So we may assume that

$$\bar{z}_s(0) = \Pr_{H^-(v)} l_2 P_2^{-1}x_u(0) = l_2 \bar{x}_s(0),$$

using 6.40 and the orthogonality of $\bar{x}_u(0)$ and $H^-(v)$. This implies that

$$\begin{bmatrix} T_{13} & T_{14} \end{bmatrix} = l_2 \begin{bmatrix} Q_{21} & Q_{22} \end{bmatrix}.$$

Furthermore

$$\ker(l_1) = \ker(\Pi_2) = \text{range}(I - p_2 P_2)$$

$$\ker(l_2) = \text{range}(\Pi_2) = \ker(I - p_2 P_2)$$

Summing up

$$\text{range}(\Pi) = \ker(T_2) = \left\{ \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \middle| \alpha \in \text{range } p_1 Q_1, \beta \in P_2 \text{range}(I - p_2 P_2) \right\}$$

and

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{13} & T_{14} \end{bmatrix} = \begin{bmatrix} t_1 Q_1^{-1} & 0 \\ 0 & I_2 \end{bmatrix} \begin{bmatrix} Q_1 & Q_{12} \\ Q_{21} & Q_2 \end{bmatrix}$$

thus

$$\begin{aligned} \ker(\Pi) &= \ker(T_1) = Q^{-1} \left\{ \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \middle| \alpha \in Q_1 \ker p_1 Q_1, \beta \in \ker(I - p_2 P_2) \right\} \\ &= Q^{-1} \left\{ \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \middle| \alpha \in \ker p_1, \beta \in \ker(I - p_2 P_2) \right\}. \end{aligned}$$

In other words

$$\begin{aligned} \text{range}(\Pi) &= \text{range} \begin{bmatrix} p_1 & 0 \\ 0 & P_2 - P_2 p_2 P_2 \end{bmatrix} \\ \ker(\Pi) &= \ker \left(\begin{bmatrix} p_1 & 0 \\ 0 & P_2 - P_2 p_2 P_2 \end{bmatrix} Q \right). \end{aligned}$$

REMARK 6.16. *Let us remark that*

$$\begin{aligned} p_1 &= \Pi_1 P_1 = P_1 \Pi_1^T \\ P_2 - P_2 p_2 P_2 &= P_2 (I - \Pi_2) = (I - \Pi_2)^T P_2. \end{aligned}$$

Our construction in a sense generalizes Theorem 8.1 in *Fuhrmann* (Fuhrmann, 1995) because not every all-pass function can be considered as a phase function. *Green and Anderson* (Green and Anderson, 1987a) proved that a necessary and sufficient condition for this is the all-pass function should be a “minimal” one. This notion is defined as follows. Consider the rational all-pass function Φ .

DEFINITION 6.17. *The all-pass function is called minimal all-pass if*

$$\dim X_s = \dim X_u + \dim [H^+(v) \cap H^-(w)].$$

REMARK 6.18. *Green and Anderson's (Green and Anderson, 1987a) original formulation is given in terms of the stable and unstable poles of Φ , and the number of Hankel singular values of which equal 1. But since this Hankel-operator is defined as $\text{Pr}_{H^-(w)} : H^+(v) \rightarrow H^-(w)$ these numbers coincide with $\dim X_s$, $\dim X_u$, $\dim [H^+(v) \cap H^-(w)]$, respectively.*

Let us remark that in case of phase function the minimal factorization is connected to the minimal realizations. Namely, assume that now y is a purely-nondeterministic full-rank stationary process with rational spectral density. $w_-(t)$, $t \in \mathbb{Z}$ is the innovation process, $\bar{w}_+(t)$, $t \in \mathbb{Z}$ is the backward innovation process. The corresponding transfer functions are W_- and $\bar{W}_+$, respectively. Consider the all-pass function

$$\Phi = \bar{W}_+^{-1} W_-$$

the phase function. Let $\Phi = \Phi_1 \Phi_2$ be an arbitrary all-pass factorization. Apply the transformation defined by Φ_2 to the sequence $w_-(t)$, $t \in \mathbb{Z}$ defining in this way the white noise process $w(t)$, $t \in \mathbb{Z}$. Since $W_- \Phi_2^{-1}$ obviously a (possibly acausal) spectral factor of y the subspace

$$X = \Pr_{H^-(w)} H^+(y) \oplus \Pr_{H^+(w)} H^-(y)$$

is a Markovian splitting subspace.

PROPOSITION 6.19. *The factorization $\Phi = \Phi_1 \Phi_2$ is a minimal factorization if and only if X is a minimal splitting subspace.*

Proof. Set $S = H^-(w) \oplus \Pr_{H^+(w)} H^-(y)$, $\bar{S} = H^+(w) \oplus \Pr_{H^-(w)} H^+(y)$. The minimality of X is equivalent to

$$S = H^-(y) \vee \bar{S}^\perp, \quad \bar{S} = H^+(y) \vee S^\perp.$$

Since the inclusions

$$S \supset H^-(y) \vee \bar{S}^\perp, \quad \bar{S} \supset H^+(y) \vee S^\perp$$

are always fulfilled we have to examine the converse inclusions.

Observe that

$$H^-(y)^\perp = H^+(w_-)$$

and

$$H^+(y)^\perp = H^-(\bar{w}_+).$$

Consequently

$$\bar{S}^\perp = H^-(w) \cap H^-(\bar{w}_+), \quad S^\perp = H^+(w) \cap H^+(w_-).$$

On the other hand if

$$[H^-(w) \cap H^-(\bar{w}_+)] \vee H^-(y) \supset H^-(w) \quad (6.41)$$

then it contains $\Pr_{H^+(w)} H^-(y)$, as well. In fact, if $\alpha \in H^-(y)$ then 6.41 implies that $\Pr_{H^-(w)} \alpha$ is an element of the set on the left hand side consequently this set contains $\Pr_{H^+(w)} \alpha = \alpha - \Pr_{H^-(w)} \alpha$ as well. Symmetric argument works for $\bar{S}$, too.

Summing up, until now we have proved that the minimality of X is equivalent to the inclusions

$$[H^-(w) \cap H^-(\overline{w}_+)] \vee H^-(y) \supset H^-(w)$$

and

$$[H^+(w) \cap H^+(w_-)] \vee H^+(y) \supset H^+(w).$$

Now taking intersections with $H^-(w)$ and $H^+(w)$, respectively, we arrive at the equations

$$[H^-(w) \cap H^-(\overline{w}_+)] \vee [H^-(w) \cap H^-(y)] = H^-(w)$$

$$[H^+(w) \cap H^+(w_-)] \vee [H^+(w) \cap H^+(y)] = H^+(w)$$

which are the identities of the minimal factorizations.

7. Generalized Wiener–Hopf factorization of all-pass functions

In the previous Chapter the characterization of all-pass and inner functions was discussed. Using the obtained results, in this Chapter we would like to consider a general factorization problem of all-pass functions connected to the reconstruction problem of an output process from the existing connections between two generating input – forward and backward – white noises.

Specifically, first the generalized *Wiener–Hopf* factorization of all-pass functions will be discussed without minimality conditions, afterward the various subspaces connected to an all-pass function will be analyzed and the minimal all-pass function will be investigated. Finally, after the transfer function approach to the generalized *Wiener–Hopf* factorization, we deal in detail with the reconstruction of the output process from input white noises via a geometric approach and a state-space description of the factorization, as well.

Assume that y is a full rank purely nondeterministic stationary process with rational spectral density. Let us consider two minimal, internal realizations of y :

$$x_1(t+1) = Ax_1(t) + B_1w_1(t)$$

$$y(t) = Cx_1(t) + D_1w_1(t)$$

and

$$\bar{x}_2(t) = A^T\bar{x}_2(t+1) + \bar{B}_2\bar{w}_2(t)$$

$$y(t) = G^T\bar{x}_2(t+1) + \bar{D}_2\bar{w}_2(t)$$

not necessarily sharing the same state space. Denote the corresponding spectral factors by $W_1(z)$, $\bar{W}_2(z)$. These are $p \times p$ matrix functions. Since the realizations are internal the components of $w_1(t)$, $t \in \mathbb{Z}$ and that of $\bar{w}_2(t)$, $t \in \mathbb{Z}$ form orthonormal bases in H_0 . Especially $\bar{w}_2(t)$ can be written as a linear transform of $w_1(s)$, $s \in \mathbb{Z}$:

$$\bar{w}_2(t) = \sum_{s=-\infty}^{+\infty} \varphi(t-s)w_1(s).$$

Obviously

$$\Phi(z) = \sum_{k=-\infty}^{+\infty} \varphi(k)z^k = \bar{W}_2^{-1}(z)W_1(z).$$

Since the transfer function $\overline{W}_2^{-1} W_1$ transforms the white noise $w_1(s)$, $s \in \mathbb{Z}$ into the white noise $\overline{w}_2(t)$, $t \in \mathbb{Z}$, the function $\overline{W}_2^{-1} W_1$ is an all-pass function.

In this chapter we analyse the problem how to reconstruct the process y – or its covariance structure – from the all-pass function Φ .

If $\overline{D}_2$ is nonsingular, then it is easy to give an explicit form of $\overline{W}_2^{-1} W_1$. Namely, using the matrix inversion lemma we obtain that

$$\begin{aligned} \overline{W}_2^{-1} W_1 &= \left[\overline{D}_2 + G^T (zI - A^T)^{-1} \overline{B}_2 \right]^{-1} \left[D_1 + C (z^{-1}I - A)^{-1} B_1 \right] = \\ &= \left[\overline{D}_2^{-1} - \overline{D}_2^{-1} G (zI - A^T + \overline{B}_2 \overline{D}_2^{-1} G^T)^{-1} \overline{B}_2 \overline{D}_2^{-1} \right] \cdot \\ &\quad \cdot \left[D_1 + C (z^{-1}I - A)^{-1} B_1 \right]. \end{aligned}$$

The identity

$$C = G^T \overline{P}_2 A + \overline{D}_2 \overline{B}_2^T$$

implies that

$$\overline{B}_2 \overline{D}_2^{-1} C = \overline{P}_2 (I - zA) + (I - z^{-1} \overline{\Gamma}_2^T) \overline{P}_2 - (I - z^{-1} \overline{\Gamma}_2^T) \overline{P}_2 (I - zA),$$

where $\overline{\Gamma}_2^T = A^T - \overline{B}_2 \overline{D}_2^{-1} G^T$. Consequently

$$\begin{aligned} \overline{W}_2^{-1} W_1 &= \overline{D}_2^{-1} D_1 - \overline{D}_2^{-1} G^T (zI - \overline{\Gamma}_2^T)^{-1} \overline{B}_2 \overline{D}_2^{-1} D_1 \\ &\quad + \overline{D}_2^{-1} C (z^{-1}I - A)^{-1} B_1 - \overline{D}_2^{-1} G^T \overline{P}_2 B_1 \\ &\quad - \overline{D}_2^{-1} G^T \overline{P}_2 A (z^{-1}I - A)^{-1} B_1 \\ &\quad - \overline{D}_2^{-1} G^T (zI - \overline{\Gamma}_2^T)^{-1} \overline{\Gamma}_2^T \overline{P}_2 B_1 \\ &= \overline{D}_2^{-1} (D_1 - G^T \overline{P}_2 B_1) + \overline{B}_2^T (z^{-1}I - A)^{-1} B_1 \\ &\quad - \overline{D}_2^{-1} G^T (zI - \overline{\Gamma}_2^T)^{-1} (\overline{B}_2 \overline{D}_2^{-1} D_1 + \overline{\Gamma}_2^T \overline{P}_2 B_1). \end{aligned} \tag{7.1}$$

Let us remark that

$$\overline{D}_2^{-1} (D_1 - G^T \overline{P}_2 B_1) = E [\overline{w}_2(0) w_1(0)^T]$$

and

$$\overline{B}_2 \overline{D}_2^{-1} D_1 + \overline{\Gamma}_2^T \overline{P}_2 B_1 = E [\overline{x}_2(0) w_1(0)^T].$$

Let us point out that the realization above is – in general – a nonminimal realization of the transfer function $\overline{W}_2^{-1} W_1$.

7.1. Generalized Wiener-Hopf factorization of an all-pass function without minimality conditions

We would like to analyse the reverse problem. Assuming that we are given the all-pass function

$$\Phi = \bar{W}_2^{-1} W_1,$$

how to find – if it is possible – the corresponding spectral factors determining minimal, internal (forward and backward) realizations of the process y , where W_1 is an analytic, $\bar{W}_2$ is a coanalytic function. Or assume that the white noises $w_1(t)$, $t \in \mathbb{Z}$ and $\bar{w}_2(t)$, $t \in \mathbb{Z}$ are given, how to reconstruct the process $y(t)$, $t \in \mathbb{Z}$. Similar question was analysed by *M. Green* and *B.D.O. Anderson* (Green and Anderson, 1987a; Green and Anderson, 1987b).

Let us remark immediately that since

$$\bar{W}_2^{-1} W_1 = (T \bar{W}_2)^{-1} (T W_1)$$

for every nonsingular transformation T , the process y cannot be exactly reconstructed because the processes y and $T y$ lead to the same all-pass function.

As a first attempt we may try to reverse the calculations leading to the form above 7.1 of the function Φ . Namely, let us assume that

$$\Phi(z) = D_0 + \bar{B}_2 (z^{-1} I - A)^{-1} B_1 + C_0 (z I - A_0)^{-1} B_0, \quad (7.2)$$

where A is stable, A_0 is nonsingular and stable and 7.2 gives a *minimal* realization of Φ . We can obtain a *class* of factorizations

$$\Phi = \bar{W}_2^{-1} W_1$$

assuming that

$$W_1(z) = D_1 + C (z^{-1} I - A)^{-1} B_1, \quad (7.3)$$

$$\bar{W}_2(z) = \bar{D}_2 + G^T (z^{-1} I - A)^{-1} \bar{B}_2. \quad (7.4)$$

Note that here we have used the matrices A , B_1 , $\bar{B}_2$ determining the stable part of the function Φ . Comparing this to equation 7.1 we can see that the specialty in this choice lies in the fact that we have tacitly assumed that the term containing $\bar{\Gamma}^T$ does not provide new stable poles. This immediately holds in the case when for example $\bar{W}_2 = \bar{W}_+$, i.e. all the zeros are out of the unit circle. But in general this assumption means that the stable zeros $\bar{W}_2$ providing the stable poles of $\bar{W}_2^{-1}$ are cancelled by the appropriate zeros of W_1 .

Assuming the form defined in equations 7.3 and 7.4 we have to determine the matrices C , G^T , D_1 , $\bar{D}_2$. Since W_1 , $\bar{W}_2$ should be spectral factors of the same process y the equations

$$G^T = C P_1 A^T + D_1 B_1^T, \quad (7.5)$$

$$C = G^T \bar{P}_2 A + \bar{D}_2 \bar{B}_2^T \quad (7.6)$$

should hold, where P_1 and $\bar{P}_2$ are the solutions of the *Ljapunov*-equations

$$P_1 - A P_1 A^T = B_1 B_1^T,$$

$$\bar{P}_2 - A^T \bar{P}_2 A = \bar{B}_2 \bar{B}_2^T.$$

Furthermore, equation 7.1 suggest an additional equation

$$D_1 = \bar{D}_2 D_0 + G^T \bar{P}_2 B_1. \quad (7.7)$$

We shall prove that solving these equations we obtain factorizations of Φ . During the proof we shall use Theorem 6.9 providing a characterization of all-pass functions. Namely, transforming Φ into the form

$$\begin{aligned} \Phi(z) = & \left(D_0 - C_0 A_0^{-1} B_0 \right) + \bar{B}_2^T (z^{-1} I - A)^{-1} B_1 \\ & - C_0 A_0^{-1} (z^{-1} I - A_0)^{-1} A_0^{-1} B_0 \end{aligned}$$

equation 6.20 implies that

$$S \begin{bmatrix} P_1 & P_{12} & 0 \\ P_{21} & P_{22} & 0 \\ 0 & 0 & I \end{bmatrix} S^T = \begin{bmatrix} P_1 & P_{12} & 0 \\ P_{21} & P_{22} & 0 \\ 0 & 0 & I \end{bmatrix}, \quad (7.8)$$

where

$$S = \begin{bmatrix} A & 0 & B_1 \\ 0 & A_0^{-1} & A_0^{-1} B_0 \\ \bar{B}_2^T & -C_0 A_0^{-1} & D_0 - C_0 A_0^{-1} B_0 \end{bmatrix}.$$

Similarly,

$$S^T \begin{bmatrix} \bar{P}_1 & \bar{P}_{12} & 0 \\ \bar{P}_{21} & \bar{P}_{22} & 0 \\ 0 & 0 & I \end{bmatrix} S = \begin{bmatrix} \bar{P}_1 & \bar{P}_{12} & 0 \\ \bar{P}_{21} & \bar{P}_{22} & 0 \\ 0 & 0 & I \end{bmatrix}. \quad (7.9)$$

LEMMA 7.1. Assume that the matrix P_{12} has full rank. Then equations 7.5, 7.6 and 7.7 imply that

$$G^T \bar{P}_{12} - \bar{D}_2 C_0 = 0. \quad (7.10)$$

Proof. Computing the sum $7.5 + 7.6 \times P_1 A^T + 7.7 \times B_1^T$ we obtain

$$G^T = G^T \bar{P}_2 P_+ \bar{D}_2 D_0 B_1^T + \bar{D}_2 \bar{B}_2^T P_1 A^T. \quad (7.11)$$

On the other hand (1-3) of 7.8 gives

$$A P_1 \bar{B}_2 - A P_{12} A_0^{T-1} C_0^T + B_1 D_0^T - B_1 B_0^T A_0^{T-1} C_0^T = 0.$$

At the same time from (1-2) of 7.8 it follows that

$$A P_{12} A_0^{T-1} + B_1 B_0^T A_0^{T-1} = P_{12}.$$

Thus

$$A P_{12} \bar{B}_2 + B_1 D_0^T = P_{12} C_0^T. \quad (7.12)$$

Substituting 7.12 into 7.11 we obtain that

$$G^T (I - \bar{P}_2 P_1) - \bar{D}_2 C_0 P_{21} = 0.$$

Equation $I - \bar{P}_2 P_1 = \bar{P}_{12} P_{21}$ and the assumption on the rank of P_{21} proves 7.10. ■

LEMMA 7.2. Assume that the matrix P_{12} has full rank. Then equations 7.5, 7.6, 7.7 imply that

$$\bar{D}_2 = D_1 D_0^T + C P_1 \bar{B}_2 \quad (7.13)$$

holds true.

Proof. From 7.7 it follows that

$$D_1 D_0^T = \bar{D}_2 D_0 D_0^T + G^T \bar{P}_2 B_1 D_0^T. \quad (7.14)$$

On the other hand invoking (3-2) of 7.8

$$\bar{B}_2^T P_{12} A_0^{T-1} - C_0 A_0^{-1} P_{22} A_0^{T-1} + (D_0 - C_0 A_0^{-1} B_0) B_0^T A_0^{T-1} = 0.$$

Thus

$$\bar{B}_2^T P_{12} A_0^{T-1} + D_0 B_0^T A_0^{T-1} - C_0 P_{22} = 0 \quad (7.15)$$

using (2-2) of 7.8.

At the same time (3-3) of 7.8 gives that

$$\begin{aligned} & \bar{B}_2^T P_1 \bar{B}_2 - \bar{B}_2^T P_{12} A_0^{T-1} C_0^T - C_0 A_0^{-1} P_{21} \bar{B}_2 + C_0 A_0^{-1} P_{22} A_0^{T-1} C_0^T \\ & + (D_0 - C_0 A_0^{-1} B_0) (D_0 - C_0 A_0^{-1} B_0)^T = I \end{aligned}$$

Expressing $D_0 B_0^T A_0^{T-1}$ via 7.15 and using again (2-2) of 7.8 we get that

$$\bar{B}_2^T P_1 \bar{B}_2 - C_0 P_{22} C_0^T + D_0 D_0^T = I.$$

Substituting this into 7.14

$$\begin{aligned} D_1 D_0^T &= \bar{D}_2 - \bar{D}_2 \bar{B}_2^T P_1 \bar{B}_2 + \bar{D}_2 C_0 P_{22} C_0^T + G^T \bar{P}_2 B_1 D_0^T \\ &= \bar{D}_2 - C P_1 \bar{B}_2 + G^T \bar{P}_2 A P_1 \bar{B}_2 + G^T \bar{P}_{12} P_{22} C_0^T + G^T \bar{P}_2 B_1 D_0^T, \end{aligned}$$

where we have used 7.6 and 7.10. But

$$A P_{12} \bar{B}_2 + B_1 D_0^T = P_{12} C_0^T.$$

Consequently,

$$D_1 D_0^T = \bar{D}_2 - C P_1 \bar{B}_2 + G^T [\bar{P}_2 P_{12} + \bar{P}_{12} P_{22}] C_0^T.$$

Since

$$\bar{P}_2 P_{12} + \bar{P}_{12} P_{22} = 0$$

we have proved 7.13. ■

PROPOSITION 7.3. Assume that equations 7.5, 7.6, 7.7 hold and the matrix P_{12} is of full rank. Define W_1 and $\bar{W}_2$ using 7.3 and 7.4. Then

$$\bar{W}_2 \Phi = W_1.$$

Proof. Since

$$\begin{aligned} \bar{W}_2(z) \Phi(z) &= \\ &= \bar{D}_2 D_0 + G^T (zI - A^T)^{-1} \bar{B}_2 D_0 + \bar{D}_2 \bar{B}_2^T (z^{-1}I - A)^{-1} B_1 \\ &\quad + G^T (zI - A^T)^{-1} \bar{B}_2 \bar{B}_2^T (z^{-1}I - A)^{-1} B_1 + \bar{D}_2 C_0 (zI - A_0)^{-1} B_0 \\ &\quad + G^T (zI - A^T)^{-1} \bar{B}_2 C_0 (zI - A_0)^{-1} B_0, \end{aligned}$$

using the identity

$$\bar{B}_2 \bar{B}_2^T = (I - z^{-1} A^T) \bar{P}_2 + \bar{P}_2 (I - zA) - (I - z^{-1} A^T) \bar{P}_2 (I - zA)$$

and 7.6, 7.7 we obtain that

$$\begin{aligned} \bar{W}_2(z) \Phi(z) &= \\ &= D_1 + C (z^{-1}I - A)^{-1} B_1 + G^T (zI - A^T)^{-1} (\bar{B}_2 D_0 + A^T \bar{P}_2 B_1) \\ &\quad + \bar{D}_2 C_0 (zI - A_0)^{-1} B_0 + G^T (zI - A^T)^{-1} \bar{B}_2 C_0 (zI - A_0)^{-1} B_0. \end{aligned}$$

On the other hand (2-1) of 7.9 implies that

$$A^T \bar{P}_{12} A_0^{-1} - \bar{B}_2 C_0 A_0^{-1} = \bar{P}_{12}, \quad (7.16)$$

thus

$$\bar{B}_2 C_0 = \bar{P}_{12} (zI - A_0) - (zI - A^T) \bar{P}_{12}.$$

Consequently,

$$\begin{aligned} \bar{W}_2(z) \Phi(z) &= W_1(z) + (\bar{D}_2 C_0 - G^T \bar{P}_{12}) (zI - A_0)^{-1} B_0 \\ &\quad + G^T (zI - A^T)^{-1} [\bar{B}_2 D_0 + A^T \bar{P}_2 B_1 + \bar{P}_{12} B_0]. \end{aligned}$$

But (3-1) of 7.9 using 7.16 gives that

$$\bar{B}_2 D_0 + A^T \bar{P}_2 B_1 + \bar{P}_{12} B_0 = 0.$$

Applying Lemma 7.2 we obtain the equation

$$\bar{W}_2 \Phi = W_1,$$

concluding the proof. ■

Later we show that the assumption concerning the rank of P_{21} is always fulfilled if Φ is a minimal all-pass function.

REMARK 7.4. *Similarly to Lemma 7.1 it can be proved that if $\bar{P}_{21}$ has full row rank (which always holds when Φ is a minimal all-pass function) then*

$$C P_{12} + D_1 B_0^T = 0. \quad (7.17)$$

Since 7.17 and 7.5, 7.7 are equivalent to 7.5, 7.6, 7.7 to obtain a class of factorizations of Φ first we have to find the solutions of 7.17 and then define G^T and $\bar{D}_2$ using 7.5 and 7.7.

We have seen that the (all-pass) transfer function between a forward and a backward generating noise of the same output process y under some assumption contains the full pole information of the spectral density, but only a partial information about the zero pattern. This pole-information is carried by the stable part of Φ , while the zero-information is conveyed by its antistable part. Because of possible cancellations – W_1 and $\bar{W}_2$ may have common zeros – the information about the zeros is not necessarily complete. In the extremal case, i.e. $\bar{W}_+^{-1}$ and W_- , assuming that there are no zeros on the perimeter of the complex unit circle, there is no cancellation, consequently the covariance structure of y can be fully recovered.

But even in this case, if there are zeros of modulus one, the structure of the process y cannot be uniquely determined.

Let us observe that if $\overline{W}_2 = \overline{W}_1$, i.e. W_1 and $\overline{W}_2$ are forward and backward spectral factors corresponding to the same state space, then Φ is inner, so there is no zero information in this function. (The usual notation in the literature for these inner functions is $K = \overline{W}^{-1}W$.)

If instead of taking a forward and a backward spectral factor we consider two spectral factors of the same type – forward–forward or backward–backward – then the transfer function between the corresponding white noises contains only zero information – there is no pole information in these functions. Especially, the inner function $W_-^{-1}W$ carries the information about the strictly unstable zeros of the spectral factor W , while the information about the strictly stable zeros of W is contained in the inner function $W^{-1}W_+$.

With some care these notions can be defined for noninternal spectral factors, as well. For more information consult Lindquist and Picci (Lindquist and Pavon, 1984; Lindquist and Picci, 1990), or Fuhrmann and Gombani (Fuhrmann and Gombani, 1998).

7.2. Various subspaces connected to an all-pass function

Let us point out that in the construction of the previous section nothing guarantees that the realizations obtained for W_1 and $\overline{W}_2$ are minimal and furthermore they determine minimal realizations of the process y defined as

$$y(t) = W_1(z)w_1(t) = \overline{W}_2(z)\overline{w}_2(t).$$

In the sequel we develop a geometric structure of this factorization problem giving a condition which assures the pole–zero cancellation explained after 7.3 and 7.4 providing a possibility of defining first the zeros of $\overline{W}_2$ and W_1 .

During this discussion we return to the notations applied in Chapter 6, namely assume that

$$\Phi(z) = D + (z^{-1}I - A)^{-1}B$$

is a minimal realization, where

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

where A_1 and A_2^{-1} are stable matrices,

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1, C_2].$$

We restrict ourselves to the internal case, consequently we assume that $H(w_1) = H(\bar{w}_2)$ holds. According to Theorem 5.42 the state space of a minimal realization of the all pass function Φ is given by

$$X = X_u \oplus X_s = \text{Pr}_{H^+(w_1)} H^-(\bar{w}_2) \oplus \text{Pr}_{H^-(w_1)} H^+(\bar{w}_2). \quad (7.18)$$

Also the inverse transfer function is connected to the state space

$$\bar{X} = \bar{X}_u \oplus \bar{X}_s = \text{Pr}_{H^+(\bar{w}_2)} H^-(w_1) \oplus \text{Pr}_{H^-(\bar{w}_2)} H^+(w_1). \quad (7.19)$$

Equation 6.25 gives that $X = \bar{X}$.

Denote n_s, n_u the number of stable and unstable poles of Φ , respectively. Furthermore, let r, s denote the number of *Hankel*-singular values which are equal to 1 of the *Hankel*-operator $\text{Pr}_{H^-(w)} : H^+(\bar{w}_2) \rightarrow H^-(w_1)$ and $\text{Pr}_{H^-(\bar{w}_2)} : H^+(w_1) \rightarrow H^-(\bar{w}_2)$, respectively. These are the *Hankel*-operators defined by Φ and Φ^* . Obviously,

$$\begin{aligned} r &= \dim [H^+(\bar{w}_2) \cap H^-(w_1)], \\ s &= \dim [H^+(w_1) \cap H^-(\bar{w}_2)], \\ n_s &= \dim X_s = \dim \bar{X}_u, \\ n_u &= \dim X_u = \dim \bar{X}_s. \end{aligned} \quad (7.20)$$

Set

$$n = n_s + n_u.$$

Let us denote the state vector process by $x^\Phi(t)$, $t \in \mathbb{Z}$ assuming that its first n_s coordinates span X_s , the remaining coordinates span X_u , i.e.

$$x^\Phi(t) = \begin{bmatrix} x_s^\Phi(t) \\ x_u^\Phi(t) \end{bmatrix}.$$

Let us partition accordingly the matrices in the state-space representation:

$$\begin{aligned} x_s^\Phi(t+1) &= A_1 x_s^\Phi(t) + B_1 w_1(t) \\ x_u^\Phi(t) &= A_2^{-1} x_u^\Phi(t+1) - A_2^{-1} B_2 w_1(t) \\ \bar{w}_2(t) &= C_1 x_s^\Phi(t) + C_2 A_2^{-1} x_u^\Phi(t+1) + (D - C_2 A_2^{-1} B_2) w_1(t). \end{aligned} \quad (7.21)$$

Denote P the solution of the corresponding *Ljapunov*-equation:

$$P - A P A^T = B B^T$$

or in partitioned form

$$\begin{aligned} \begin{bmatrix} P_1 & P_{12} \\ P_{21} & P_2 \end{bmatrix} - \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} P_1 & P_{12} \\ P_{21} & P_2 \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}^T = \\ = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} [B_1^T \ B_2^T]. \end{aligned} \quad (7.22)$$

Similarly, the state-space representation of the inverse transfer function can be written in the following form:

$$\begin{aligned} \bar{x}_u^\Phi(t) &= A_1^T \bar{x}_u^\Phi(t+1) + C_1^T \bar{w}_2(t) \\ \bar{x}_s^\Phi(t+1) &= (A_2^{-1})^T \bar{x}_s^\Phi(t) - (A_2^{-1})^T C_2^T \bar{w}_2(t) \\ w_1(t) &= B_1^T \bar{x}_u^\Phi(t+1) + B_2^T (A_2^{-1})^T \bar{x}_s^\Phi(t) \\ &\quad + (D^T - B_2^T (A_2^{-1})^T C_2^T) \bar{w}_2(t). \end{aligned} \quad (7.23)$$

Denote Q the solution of the corresponding *Ljapunov*-equation:

$$Q - A^T Q A = C^T C$$

or in partitioned form

$$\begin{aligned} \begin{bmatrix} Q_1 & Q_{12} \\ Q_{21} & Q_2 \end{bmatrix} - \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}^T \begin{bmatrix} Q_1 & Q_{12} \\ Q_{21} & Q_2 \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} = \\ = \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix} [C_1 \ C_2]. \end{aligned} \quad (7.24)$$

Consequently, the coordinates of $x_s^\Phi(0)$, $x_u^\Phi(0)$, $\bar{x}_u^\Phi(0)$ and $\bar{x}_s^\Phi(0)$ generate the subspaces X_s , X_u , $\bar{X}_u$ and $\bar{X}_s$, respectively, and the equation

$$PQ = I \quad (7.25)$$

holds true.

PROPOSITION 7.5.

1.

$$X = (X_u \vee \bar{X}_s) \oplus [H^+(\bar{w}_2) \cap H^-(w_1)] \quad (7.26)$$

2.

$$X = (X_s \vee \bar{X}_u) \oplus [H^+(w_1) \cap H^-(\bar{w}_2)]. \quad (7.27)$$

Proof. We are going to prove (1), (2) follows by symmetry.

(1) Since the right hand side of 7.26 is obviously a subspace of X it is enough to prove that no element in X is orthogonal to this subspace.

Assume that $\alpha \in X$ is orthogonal to $X_u, \bar{X}_s$ and $H^+(\bar{w}_2) \cap H^-(w_1)$. Since $X = X_u \oplus X_s$ it follows that $\alpha \in X_s = H^-(w_1) \ominus [H^-(w_1) \cap H^-(\bar{w}_2)]$. α is orthogonal to $\bar{X}_s$, consequently it is also orthogonal to

$$H^-(\bar{w}_2) = \bar{X}_s \oplus [H^-(w_1) \cap H^-(\bar{w}_2)].$$

Thus $\alpha \in H^+(\bar{w}_2)$. Moreover $\alpha \in H^+(\bar{w}_2) \cap H^-(w_1)$. But according to our assumption it is also orthogonal to this subspace, so $\alpha = 0$, concluding the proof. ■

COROLLARY 7.6. $n_s - r = n_u - s$.

Proof. Since $X_u \cap \bar{X}_s = H^+(w_1) \cap H^-(\bar{w}_2)$ Proposition 7.5 implies that

$$n = 2n_u - s + r. \quad (7.28)$$

But $n = n_u + n_s$, so $n_s - r = n_u - s$. ■

Essentially this identity can be found in *Green and Anderson* (Green and Anderson, 1987b) – in continuous time setting. Their proof is based on the factorization indices introduced by *Clancey and Gohberg* (Clancey and Gohberg, 1981). See also *Dym and Gohberg* (Dym and Gohberg, 1983).

If there exists a p -dimensional stationary process $y(t)$, $t \in \mathbb{Z}$ such that $H = H(y) = H(w_1) = H(\bar{w}_2)$, for which the white noise $w_1(t)$, $t \in \mathbb{Z}$ determines an analytic spectral factor and $\bar{w}_2(t)$, $t \in \mathbb{Z}$ gives a coanalytic spectral factor then the inclusions

$$H^- \subset H^-(w_1), \quad H^+ \subset H^+(\bar{w}_2) \quad (7.29)$$

must hold.

PROPOSITION 7.7. If $H = H(y)$ and 7.29 holds then the all-pass function transforming the sequence $w_1(t)$, $t \in \mathbb{Z}$ to $\bar{w}_2(t)$, $t \in \mathbb{Z}$ is a minimal all-pass function.

Proof. 7.29 implies that

$$H^-(\bar{w}_2) \cap H^+(w_1) = 0$$

(since $H^+ \vee H^- = H$), consequently in view of the previous proposition

$$n_s = n_u + r.$$

Thus Φ is minimal. ■

REMARK 7.8. Since

$$x_u^\Phi(0) = P_{21}\bar{x}_u^\Phi(0) + P_{22}\bar{x}_s^\Phi(0),$$

the condition $X_u \cap \bar{X}_s = 0$ implies that the matrix P_{21} is of full row rank. Similarly, the minimality of Φ implies that $\bar{P}_{21}$ is of full row rank. These assumptions were used several times in Lemmata 7.2, 7.1 and Proposition 7.3.

We need the following lemma in the sequel.

LEMMA 7.9. Consider the state-space equation given in 7.21 for $x_s^\Phi(t)$, $t \in \mathbb{Z}$. The corresponding backward equation can be written as follows

$$\begin{aligned} P_1^{-1}x_s^\Phi(t) &= A_1^T P_1^{-1}x_s^\Phi(t+1) \\ &+ \left(C_1^T + P_1^{-1}P_{12}C_2^T \right) \left(\bar{w}_2(t) - C_2A_2^{-1}Q_2^{-1}\bar{x}_s^\Phi(t+1) \right). \end{aligned} \quad (7.30)$$

Proof. Since

$$P_1^{-1}x_s^\Phi(t) = \bar{x}(t) + P_1^{-1}P_{12}\bar{x}_s^\Phi(t),$$

using the first two equation in 7.23 we obtain that

$$\begin{aligned} P_1^{-1}x_s^\Phi(t) &= A_1^T P_1^{-1}x_s^\Phi(t+1) + \left(P_1^{-1}P_{12}A_2^T - A_1^T P_1^{-1}P_{12} \right) x_s^\Phi(t+1) \\ &+ \left(C_1^T + P_1^{-1}P_{12}C_2^T \right) \bar{w}_2(t). \end{aligned} \quad (7.31)$$

On the other hand

$$\begin{aligned} P_1^{-1}P_{12}A_2^T - A_1^T P_1^{-1}P_{12} &= P_1^{-1}P_{12}A_2^T + A_1^T Q_{12}Q_2^{-1} = \\ &= P_1^{-1} \left(P_1A_1^T Q_{12} + P_{12}A_2^T Q_2 \right) Q_2^{-1}. \end{aligned}$$

But the equation

$$PA^TQA + PC^TC = I$$

implies

$$\begin{bmatrix} P_1 & P_{12} \end{bmatrix} \begin{bmatrix} A_1^T & 0 \\ 0 & A_2^T \end{bmatrix} \begin{bmatrix} Q_{12} \\ Q_2 \end{bmatrix} A_2 + \begin{bmatrix} P_1 & P_{12} \end{bmatrix} \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix} C_2 = 0,$$

i.e.

$$P_1A_1^T Q_{12} + P_{12}A_2^T Q_2 + (P_1C_1^T + P_{12}C_2^T)C_2A_2^{-1} = 0.$$

Substituting this into 7.31 we get equation 7.30. ■

Direct calculation gives that the covariance matrix of the variable $\bar{w}_2(t) - C_2 A_2^{-1} Q_2^{-1} \bar{x}_2^\Phi(t+1)$, $t \in \mathbb{Z}$ is $I + C_2 A_2^{-1} Q_2^{-1} (A_2^T)^{-1} C_2^T$.

Observe that

$$\left[I + C_2 A_2^{-1} Q_2^{-1} (A_2^T)^{-1} C_2^T \right]^{-1} = I - C_2 Q_2^{-1} C_2^T.$$

Let us point out that – invoking the second equation in 7.23 –

$$\Pr_{\bar{X}_s} \bar{w}_2(-1) = C_2 A_2^{-1} Q_2^{-1} \bar{x}_s^\Phi(0).$$

It can easily be proved that

$$\bar{w}_2(-1) - C_2 A_2^{-1} Q_2^{-1} \bar{x}_s^\Phi(0) \in (H^-(w_1) \cap H^-(\bar{w}_2))$$

and

$$\left[\bar{w}_2(0) - C_2 A_2^{-1} Q_2^{-1} \bar{x}_2^\Phi(1) \right] \perp (H^-(w_1) \cap H^-(\bar{w}_2)),$$

proving that

$$\bar{w}_2(t) - C_2 A_2^{-1} Q_2^{-1} \bar{x}_s^\Phi(t+1), \quad t \in \mathbb{Z}$$

is an uncorrelated, stationary sequence and moreover the coordinates of

$$\bar{w}_2(t) - C_2 A_2^{-1} Q_2^{-1} \bar{x}_s^\Phi(t+1), \quad t \leq -1 \tag{7.32}$$

span the subspace $H^-(w_1) \cap H^-(\bar{w}_2)$.

Thus, if the process x_s^Φ , $t \in \mathbb{Z}$ is imbedded into an all-pass system, then the noise in its backward equation can be given explicitly – based on a noise in another backward equation.

7.3. Minimal all-pass functions

THEOREM 7.10. Assume that $r = s = 0$. Then there exists a p -dimensional stationary process $y(t)$, $t \in \mathbb{Z}$ such that $H = H(y)$ and for which

$$H^- \subset H^-(w_1), \quad H^+ \subset H^+(\bar{w}_2).$$

Moreover, this process is essentially unique. Furthermore $w_1(t)$, $t \in \mathbb{Z}$ is the forward innovation process of $y(t)$, $t \in \mathbb{Z}$, $\bar{w}_2(t)$, $t \in \mathbb{Z}$ is the backward one. I.e. $w_1(t)$, $t \in \mathbb{Z}$ determines a minimum phase, analytic spectral factor, $\bar{w}_2(t)$, $t \in \mathbb{Z}$, determines the maximum phase, coanalytic spectral factor.

REMARK 7.11. Obviously, if the process $y(t)$, $t \in \mathbb{Z}$ satisfies the theorem then $T y(t)$ is also an appropriate selection, where T is a fixed nonsingular $p \times p$ matrix. The essential uniqueness in the theorem refers to this fact.

Proof. Observe that $r = s = 0$ implies that

$$\begin{aligned} H^-(w_1) \vee H^+(\bar{w}_2) &= H, \\ H^-(w_1) \cap H^+(\bar{w}_2) &= 0. \end{aligned} \quad (7.33)$$

Apply the shift operator U . Since $w_1(t)$, $t \in \mathbb{Z}$ is a p -dimensional white noise we have that

$$p = \dim [UH^-(w_1) \ominus H^-(w_1)].$$

The assumptions in 7.33 imply that

$$p = \dim [UH^-(w_1) \cap H^+(\bar{w}_2)].$$

Denote this intersection by Y_0 and choose an arbitrary basis in Y_0 . Consider the elements of this basis as the coordinate vectors of $y(0)$. Set

$$y(t) = U^t y(0).$$

Obviously,

$$\begin{aligned} H^- &= H^-(y) \subset H^-(w_1), \\ H^+ &= H^+(y) \subset H^+(\bar{w}_2). \end{aligned}$$

Thus y is a purely nondeterministic process. $y(0)$ can be written as linear combinations of $w_1(t)$, $t \leq 0$, so the corresponding spectral factor is analytic. Symmetrically, $y(0)$ is a linear combination of the vectors $\bar{w}_2(t)$, $t \geq 0$, leading to a coanalytic spectral factor. Since $Y_0 \cap H^-(y) = 0$ and $\dim Y_0 = p$

$$\dim [UH^-(y) \ominus H^-(y)] \geq p.$$

But it cannot be larger than p because $w_1(t)$, $t \in \mathbb{Z}$, is a p -dimensional white noise process and $H(y) \subset H(w_1)$. Consequently, y is a full-rank process, and $H(y) = H(w_1)$. Especially,

$$\begin{aligned} H^-(y) &= H^-(w_1), \\ H^+(y) &= H^+(\bar{w}_2). \end{aligned}$$

In other words $w_1(t)$, $t \in \mathbb{Z}$, is the forward innovation process of $y(t)$, and $\bar{w}_2(t)$, $t \in \mathbb{Z}$, is the backward one.

The essential uniqueness is immediate from the observation that the coordinates of $y(0)$ must stay in $UH^-(w_1) \cap H^+(\bar{w}_2)$ which is – as we have seen it – p -dimensional. ■

REMARK 7.12. Let us emphasize that the previous theorem states that if Φ is a minimal all-pass function for which $r = 0$ then a factorization of the form

$$\Phi = \overline{W_2}^{-1} W_1$$

where W_1 is analytic, $\overline{W_2}$ is coanalytic ensures that W_1 must be of minimum phase, $\overline{W_2}$ must be of maximum phase.

7.4. The general case – generalized Wiener-Hopf factorization. Transfer function approach

In this section we analyse what can be preserved from the previous construction in the general case, i.e. when r is arbitrary.

Obviously, the coordinates of $y(0)$ must be in $UH^-(w_1) \cap H^+(\overline{w_2})$. This ensures the analyticity of W_1 and the coanalyticity of $\overline{W_2}$. But now $\dim[UH^-(w_1) \cap H^+(\overline{w_2})] = r + p$, so $y(0)$ is not (essentially) unique. We have to choose a p -dimensional subspace Y of this intersection but in such a way that

$$H(y) = H(w_1) = H(\overline{w_2})$$

hold.

In order to be able to find all the p -dimensional stationary processes fulfilling these requirements let us repeat briefly in our Hilbert space setting the construction of the factorization indices. See Clancey and Gohberg (Clancey and Gohberg, 1981), Dym and Gohberg (Dym and Gohberg, 1983), and also Ball and Helton (Ball and Helton, 1984), Ball and Ran (Ball and Ran, 1987).

Set

$$\mathcal{K}_0 = H^+(\overline{w_2}) \cap UH^-(w_1)$$

and

$$\mathcal{K}_i = H^+(\overline{w_2}) \cap U^{1-i}H^-(w_1), \quad i \geq 1.$$

Obviously

$$\mathcal{K}_{i+1} \subset \mathcal{K}_i, \quad U\mathcal{K}_{i+1} \subset \mathcal{K}_i,$$

or in general

$$U^j \mathcal{K}_l \subset \mathcal{K}_i, \quad \text{if } 0 \leq j \leq l, \quad 0 \leq i \leq l - j.$$

Since $\dim \mathcal{K}_0 = r + p$ there exists a largest integer ρ for which $\mathcal{K}_\rho \neq 0$. Define

$$\mathcal{N}_\rho = \mathcal{K}_\rho,$$

$$\mathcal{N}_i = \mathcal{K}_i \ominus (\mathcal{K}_{i+1} \vee U\mathcal{K}_{i+1}), \quad 0 \leq i < \rho.$$

It can be checked easily by induction that

$$\mathcal{K}_i = \mathcal{N}_i \oplus (\mathcal{N}_{i+1} \vee U\mathcal{N}_{i+1}) \oplus (\mathcal{N}_{i+2} \vee U\mathcal{N}_{i+2} \vee U^2\mathcal{N}_{i+2}) \oplus \dots \\ \oplus (\mathcal{N}_\rho \vee U\mathcal{N}_\rho \vee \dots \vee U^{\rho-i}\mathcal{N}_\rho)$$

or in other words, introducing the notation

$$\mathcal{M}_i = \mathcal{N}_i \oplus U\mathcal{N}_{i+1} \oplus U^2\mathcal{N}_{i+2} \oplus \dots \oplus U^{\rho-i}\mathcal{N}_\rho,$$

we get that

$$\mathcal{K}_i = (\mathcal{K}_{i+1} \vee U\mathcal{M}_{i+1}) \oplus \mathcal{N}_i = \mathcal{K}_{i+1} \vee \mathcal{M}_i.$$

This implies that

$$\dim \mathcal{K}_i = \dim \mathcal{K}_{i+1} + \dim \mathcal{N}_i + \dots + \dim \mathcal{N}_\rho. \quad (7.34)$$

Let us denote

$$Y = \mathcal{N}_\rho \oplus \mathcal{N}_{\rho-1} \oplus \dots \oplus \mathcal{N}_0$$

and

$$\bar{Y} = \mathcal{M}_0 = \mathcal{N}_0 \oplus U\mathcal{N}_1 \oplus \dots \oplus U^\rho \mathcal{N}_\rho.$$

Then 7.34 implies that

$$\dim Y = \dim \bar{Y} = \dim \mathcal{K}_0 - \dim \mathcal{K}_1 = p.$$

Choose an orthonormal basis $\xi_1, \dots, \xi_p$ in Y such that the corresponding vectors span in order $\mathcal{N}_\rho, \mathcal{N}_{\rho-1}, \dots, \mathcal{N}_0$. For each ξ_i denote k_i the index of the space in the sequence $\mathcal{N}_0, \dots, \mathcal{N}_\rho$ to which it belongs. Then the vectors

$$\{U^j \xi_i, 0 \leq j \leq k_i, i = 1, \dots, p\}$$

form a basis in $\mathcal{K}_0$. Denote

$$\eta_i = U^{k_i} \xi_i, \quad i = 1, \dots, p.$$

Observe that $\eta_1, \dots, \eta_p$ generate $\mathcal{M}_0$.

The numbers $k_1, \dots, k_p$ are the so-called (*nonnegative*) *factorization indices*. Clearly

$$\dim \mathcal{K}_1 = r = \sum_{i=1}^p k_i.$$

The construction gives that the vectors

$$U^{-j} \eta_i, \quad j = 0, \dots, k_i, \quad i = 1, \dots, p$$

generate $\mathcal{K}_0$ so

$$\bigvee_{j=0}^{\infty} U^{-j} \bar{Y} = UH^-(w_1),$$

and similarly

$$\bigvee_{j=0}^{\infty} U^j Y = H^+(\bar{w}_2).$$

This implies that expressing $\eta = (\eta_1, \dots, \eta_p)$ as a linear combination of $w_1(t)$, $t \leq 0$ we obtain an analytic, minimum phase, function W_0 and expressing $\xi = (\xi_1, \dots, \xi_p)$ in terms of $\bar{w}_2(t)$, $t \geq 0$, we obtain a coanalytic, maximum phase function $\bar{W}_0$. Furthermore, introducing the so-called winding matrix

$$D = \begin{bmatrix} z^{k_1} & & \\ & \ddots & \\ & & z^{k_p} \end{bmatrix}$$

we get the so-called *generalized Wiener-Hopf factorization* of Φ as

$$\Phi = \bar{W}_0^{-1} D W_0.$$

Every element in $UH^-(w_1) \cap H^+(\bar{w}_2)$ can be written in the form of linear combination of $\{U^j \xi_i, 0 \leq j \leq k_i, i = 1, \dots, p\}$ and that of $\{U^{-j} \eta_i, 0 \leq j \leq k_i, i = 1, \dots, p\}$. Since the coordinates of $y(0)$ must lie in this intersection the representations

$$y(0)_i = \sum_{j=1}^p \sum_{l=0}^{k_j} \lambda_{ijl} U^l \xi_j = \sum_{j=1}^p \sum_{l=0}^{k_j} \lambda_{ijl} U^{l-k_j} \eta_j$$

must hold. Let us introduce the matrix functions $D_+(z)$, $D_-(z)$ with entries defined by

$$[D_+(z)]_{ij} = \sum_{l=0}^{k_j} \lambda_{ijl} z^{k_j-l}$$

$$[D_-(z)]_{ij} = \sum_{l=0}^{k_j} \lambda_{ijl} z^{-l}$$

D_+ is a polynomial in z , while D_- is a polynomial in z^{-1} . Obviously

$$D_+(z) = D_-(z) D(z).$$

The full rank condition imposed on W_1 and $\bar{W}_2$ is clearly equivalent to the assumption that

$$\det D_+(e^{-i\lambda}), \quad \det D_-(e^{-i\lambda})$$

are not identically zero. In this case the functions

$$W_1(z) = D_+(z)W_0(z)$$

$$\overline{W}_2(z) = D_-(z)\overline{W}_0(z)$$

provide a factorization of $\Phi(z)$ as

$$\Phi(z) = \overline{W}_2^{-1}(z)W_1(z).$$

As a summary of the previous consideration with some further refinement we can formulate the following theorem.

THEOREM 7.13. *Assume that Φ is a rational minimal all-pass function. Then it has the generalized Wiener-Hopf factorization*

$$\Phi = \overline{W}_0^{-1} D W_0, \quad (7.35)$$

where

$$D = \begin{bmatrix} z^{k_1} & & \\ & \ddots & \\ & & z^{k_p} \end{bmatrix} \quad (7.36)$$

is a winding matrix, $W_0(z)$ is analytic, minimum phase, $\overline{W}_0(z)$ is coanalytic, maximum phase function.

Also there exists a factorization

$$\Phi(z) = \overline{W}_2^{-1}(z)W_1(z),$$

where W_1 is analytic $\overline{W}_2$ is coanalytic if and only if there exists a matrix polynomial $D_+(z)$ for which the degree of $[D_+(z)]_{ij}$ is no greater than k_j such that $\det D_+(e^{-i\lambda})$ is not identically zero and

$$\begin{aligned} W_1(z) &= D_+(z)W_0(z) \\ \overline{W}_2(z) &= D_-(z)\overline{W}_0(z) \end{aligned} \quad (7.37)$$

where $D_-(z) = D_+(z)D^{-1}(z)$.

Furthermore W_1 is a minimum phase, $\overline{W}_2$ is a maximum phase function if and only if the zeros of the polynomial $P(z) = \det D_+(z)$ are of modulus 1.

Proof. It remains to prove only the last extension. $W_1(z)$ is of minimum phase if it has no zeros inside the unit circle which is equivalent to that $\det W_1(z)$ has no zeros there. But $W_0(z)$ is a minimum phase function so it remains to control $P(z) = \det D_+(z)$. Similar reasoning gives that $\overline{W}_2(z)$ is of maximum phase if $\det D_-(z) = P(z)z^{-r}$ has no zeros outside the unit circle. Consequently, all zeros of $P(z)$ must have modulus 1. ■

REMARK 7.14. *Because of its beauty let us mention a special possibility. Set*

$$y(0)_i = \begin{cases} \xi_i, & \text{if } k_i = 0 \\ \eta_i - \xi_i & \text{if } k_i > 0 \end{cases}$$

i.e. in the second case the difference of the farthest elements in the chain $U^j \xi_i$, $j \in \mathbb{Z}$ inside $UH^-(w_1) \cap H^+(\bar{w}_2)$.

In this case $D_+(z)$ is a diagonal matrix with elements

$$\begin{aligned} 1 & \quad \text{if } k_i = 0, \\ 1 - z^{k_i} & \quad \text{if } k_i > 0. \end{aligned}$$

Thus the zeros of $\det D_+(z)$ are of modulus 1, so for the corresponding process $y(t)$, $t \in \mathbb{Z}$ defined by this special factorization

$$H^-(y) = H^-(w_1)$$

$$H^+(y) = H^+(\bar{w}_2)$$

hold.

The coordinates of $x_s^\Phi(0)$ form a basis in $\text{Pr}_{H^-(w_1)} H^+(\bar{w}_2)$. If $H^-(w_1) \cap H^+(\bar{w}_2) \neq 0$ then – possibly applying a coordinate transformation – we may assume that the first r coordinates of $x_s^\Phi(0)$ span $H^-(w_1) \cap H^+(\bar{w}_2)$ and the remaining $n_s - r$ coordinates are orthogonal to $H^-(w_1) \cap H^+(\bar{w}_2)$. In this case the covariance matrix P_1 of $x_s^\Phi(0)$ is block diagonal:

$$P_1 = \begin{bmatrix} P_1^{11} & 0 \\ 0 & P_1^{22} \end{bmatrix}.$$

Let us split $x_s^\Phi(0)$ into two parts according to this

$$x_s^\Phi(0) = \begin{bmatrix} x_{s1}^\Phi(0) \\ x_{s2}^\Phi(0) \end{bmatrix}.$$

Since in this case the coordinates of $x_{s1}^\Phi(0)$ are also in $\text{Pr}_{H^+(\bar{w}_2)} H^-(w_1)$ they can be expressed using only $\bar{x}_u^\Phi(0)$. In other words in the formula

$$x_s^\Phi(0) = P_{11} \bar{x}_u^\Phi(0) + P_{12} \bar{x}_s^\Phi(0)$$

the first $r \times (n_s - r)$ block of P_{12} is zero and the second $(n_s - r) \times (n_s - r)$ block is invertible. Now using the identity

$$A_1 P_{12} A_2^T + B_1 B_2^T = P_{12}$$

and introducing the block structure in A_1, B_1, P_{12} according to the decomposition of $x_s^\Phi(0)$ above

$$A_1 = \begin{bmatrix} A_1^{11} & A_1^{12} \\ A_1^{21} & A_1^{22} \end{bmatrix}, \quad B_1 = \begin{bmatrix} B_{11} \\ B_{12} \end{bmatrix}, \quad P_{12} = \begin{bmatrix} 0 \\ P_{12}^\times \end{bmatrix}$$

we obtain that

$$A_1^{12} P_{12}^\times A_2^T + B_{11} B_2^T = 0.$$

Thus

$$A_1^{12} = -B_{11} \left[B_2^T (A_2^T)^{-1} (P_{12}^\times)^{-1} \right].$$

Let us remark that in the canonical form of P introduced by Glover (Glover, 1984) $P_{12}^\times = I$.

The state space equation for $x_{s1}^\Phi(t)$, $t \in \mathbb{Z}$ can be written as

$$x_{s1}^\Phi(t+1) = A_1^{11} x_{s1}^\Phi(t) + B_{11} \left[w_1(t) - B_2^T (A_2^T)^{-1} (P_{12}^\times)^{-1} x_{s2}^\Phi(t) \right]$$

implying the representation

$$x_{s1}^\Phi(0) = \sum_{j=1}^{\infty} (A_1^{11})^{j-1} B_{11} \left(w(-j) - B_2^T (A_2^T)^{-1} (P_{12}^\times)^{-1} x_{s2}^\Phi(-j) \right). \quad (7.38)$$

Similarly, using the block diagonal structure of P_1 we obtain that $x_{s1}^\Phi(0)$ can be expressed using the first r coordinates of $\bar{x}_u^\Phi(0)$. I.e. if

$$\bar{x}_u^\Phi(0) = \begin{bmatrix} \bar{x}_{u1}^\Phi(0) \\ \bar{x}_{u2}^\Phi(0) \end{bmatrix}$$

in partitioned form then the coordinates of $\bar{x}_{u1}^\Phi(0)$ span $H^-(w_1) \cap H^+(\bar{w}_2)$. Repeating the argument above we obtain that in

$$\bar{x}_u^\Phi(0) = Q_1 x_s^\Phi(0) + Q_{12} x_u^\Phi(0)$$

the first $r \times (n_s - r)$ block of Q_{12} is zero and the second $(n_s - r) \times (n_s - r)$ block is invertible. Thus introducing a block structure

$$C_1 = [C_{11}, C_{12}] \quad Q_{12} = \begin{bmatrix} 0 \\ Q_{12}^\times \end{bmatrix}$$

from the equation

$$A_1^T Q_{12} A_2 + C_1^T C_2 = Q_{12}$$

we obtain that

$$A_1^{21T} Q_{12}^\times A_2 + C_{11}^T C_2 = 0.$$

Consequently,

$$A_1^{21T} = -C_{11}^T C_2 A_2^{-1} (Q_{12}^\times)^{-1} \quad (7.39)$$

using that $Q_{12}^\times$ is nonsingular. Substituting this into the state space equation for $\bar{x}_{u1}^\Phi$ we obtain that

$$\bar{x}_{u1}^\Phi(t) = A_1^{11T} \bar{x}_{u1}^\Phi(t+1) + C_{11}^T \left[\bar{w}_2(t) - C_2 A_2^{-1} (Q_{12}^\times)^{-1} \bar{x}_{u2}^\Phi(t+1) \right]. \quad (7.40)$$

At the same time from 7.21 we obtain that

$$\begin{aligned} \bar{w}_2(t) - C_2 A_2^{-1} Q_{12}^\times{}^{-1} \bar{x}_{u2}^\Phi(t+1) &= \\ &= C_{11} x_{s1}^\Phi(t) + C_{12} x_{s2}^\Phi(t) + C_2 x_u^\Phi(t) + D w_1(t) \\ &\quad - C_2 A_2^{-1} Q_{12}^\times{}^{-1} \bar{x}_{u2}^\Phi(t+1). \end{aligned}$$

Using the equations

$$\begin{aligned} \bar{x}_{u2}^\Phi(t+1) &= Q_1^{22} x_{s2}^\Phi(t+1) + Q_{12}^\times x_u^\Phi(t+1) \\ x_{s2}^\Phi(t+1) &= A_1^{21} x_{s1}^\Phi(t) + A_1^{22} x_{s2}^\Phi(t) + B_{12} w_1(t) \\ x_u^\Phi(t+1) &= A_2 x_u^\Phi(t) + B_2 w_1(t) \end{aligned}$$

we get that

$$\begin{aligned} \bar{w}_2(t) - C_2 A_2^{-1} Q_{12}^\times{}^{-1} \bar{x}_{u2}^\Phi(t+1) &= \\ &= \left(C_{11} - C_2 A_2^{-1} Q_{12}^\times{}^{-1} Q_1^{22} A_1^{21} \right) x_{s1}^\Phi(t) \\ &\quad + \left(D - C_2 A_2^{-1} Q_{12}^\times{}^{-1} Q_1^{22} B_{12} - C_2 A_2^{-1} B_2 \right) w_1(t) \\ &\quad + \left(C_{12} - C_2 A_2^{-1} Q_{12}^\times{}^{-1} Q_1^{22} A_1^{22} \right) x_{s2}^\Phi(t). \end{aligned} \quad (7.41)$$

From 7.39 it follows that

$$C_{11} - C_2 A_2^{-1} Q_{12}^\times{}^{-1} Q_1^{22} A_1^{21} = \left(I + A_2^{-1} Q_{12}^\times{}^{-1} Q_1^{22} Q_{12}^{\times T-1} A_1^{T-1} \right) C_{11}.$$

On the other hand equation 6.17 implies that

$$A_1^{22} P_{12}^\times A_2^T + B_{12} B_2^T = P_{12}^\times$$

and

$$C_{12} P_{12}^\times A_2^T + C_2 P_2 A_2^T + D B_2^T = 0.$$

Expressing A_1^{22} and utilizing that $Q_{12}^{\times-1} Q_1^{22} = P_2 P_{12}^{\times-1}$ we obtain that

$$\begin{aligned}
 C_{12} - C_2 A_2^{-1} Q_{12}^{\times-1} Q_1^{22} A_1^{22} &= \\
 &= C_{12} + C_2 A_2^{-1} P_2 P_{12}^{\times-1} \left(P_{12}^{\times} A_2^{T-1} P_{12}^{\times-1} - B_2 B_2^T A_2^{T-1} P_{12}^{\times-1} \right) = \\
 &= C_{12} + C_2 \left(P_2 + A_2^{-1} B_2 B_2^T A_2^{T-1} \right) P_{12}^{\times-1} \\
 &\quad - C_2 A_2^{-1} P_2 P_{12}^{\times-1} B_{12} B_2^T A_2^{T-1} P_{12}^{\times-1} = \\
 &= \left(-D + C_2 A_2^{-1} B_2 + C_2 A_2^{-1} Q_{12}^{\times-1} Q_1^{22} B_{12} \right) B_2^T A_2^{T-1} P_{12}^{\times-1}.
 \end{aligned}$$

Substituting these back into equation 7.41 we can conclude that

$$\begin{aligned}
 \bar{w}_2(t) - C_2 A_2^{-1} Q_{12}^{\times-1} \bar{x}_{u2}^{\Phi}(t+1) &= \\
 &= \left(I + C_2 A_2^{-1} Q_{12}^{\times-1} Q_1^{22} Q_{12}^{\times T-1} A_2^{T-1} C_2^T \right) C_{11} x_{s1}^{\Phi}(t) \quad (7.42) \\
 &\quad + \left(D - C_2 A_2^{-1} B_2 - C_2 A_2^{-1} Q_{12}^{\times-1} Q_1^{22} B_{12} \right) \\
 &\quad \times \left[w_1(t) - B_2^T A_2^{T-1} P_{12}^{\times-1} x_{s2}^{\Phi}(t) \right].
 \end{aligned}$$

Denote shortly

$$\begin{aligned}
 \Sigma_2 &= \left[I + C_2 A_2^{-1} Q_{12}^{\times-1} Q_1^{22} Q_{12}^{\times T-1} A_2^{T-1} C_2^T \right] = \\
 &= \text{Var} \left(\bar{w}_2(t) - C_2 A_2^{-1} Q_{12}^{\times-1} \bar{x}_{u2}^{\Phi}(t+1) \right) = \\
 \Sigma_1 &= \left[I + B_2^T A_2^{T-1} P_{12}^{\times-1} P_1^{22} P_{12}^{\times T-1} A_2^{-1} B_2 \right] = \\
 &= \text{Var} \left(w_1(t) - B_2^T A_2^{T-1} P_{12}^{\times-1} x_{s2}^{\Phi}(t) \right) + \\
 D_W &= D - C_2 A_2^{-1} B_2 - C_2 A_2^{-1} Q_{12}^{\times-1} Q_1^{22} B_{12}.
 \end{aligned}$$

Then equation 7.42 implies a product factorization

$$\Phi(z) = \bar{W}_0^{-1} L W_0,$$

where

$$\bar{W}_0 = \left[I - C_2 A_2^{-1} Q_{12}^{\times-1} [0, I] (zI - A^T)^{-1} C^T \right]$$

$$L(z) = \left[D_W + \Sigma_2 C_{11} (z^{-1} I - A_1^{11})^{-1} B_{11} \right]$$

$$W_0 = \left[I - B_2^T A_2^{T-1} P_{12}^{\times-1} [0, I] (z^{-1} I - A)^{-1} B \right].$$

(Cf. Theorem 4.1 in *M. Green and B.D.O. Anderson* (Green and Anderson, 1987b).)

Since

$$A + B B_2^T A_2^{T-1} P_{12}^{\times-1} [0, I] = \begin{bmatrix} A_1^{11} & 0 \\ A_1^{21} & P_{12}^{\times} A_2^{T-1} P_{12}^{\times-1} \end{bmatrix}$$

is a stable matrix, the poles and zeros of the transfer function W_0 are in the unit circle. Similarly, from

$$A^T + C^T C_2 A_2^{-1} Q_{12}^{\times-1} [0, I] = \begin{bmatrix} A_1^{11T} & 0 \\ A_1^{12T} & Q_{12}^{\times} A_2^{-1} Q_{12}^{\times-1} \end{bmatrix}$$

it follows that the poles and zeros of $\bar{W}_0$ are outside of the unit circle.

Finally the orthogonality of $x_{s1}^{\Phi}(t)$ and $w_1(t)$, $x_{s2}^{\Phi}(t)$ and that of $x_{s1}^{\Phi}(t+1)$ and $\bar{w}_2(t)$, $\bar{x}_{u2}^{\Phi}(t+1)$ implies that

$$\begin{bmatrix} A_1^{11} & B_{11} \\ \Sigma_2 C_{11} & D_W \end{bmatrix} \begin{bmatrix} P_1^{11} & 0 \\ 0 & \Sigma_1 \end{bmatrix} \begin{bmatrix} A_1^{11} & B_{11} \\ \Sigma_2 C_{11} & D_W \end{bmatrix}^T = \begin{bmatrix} P_1^{11} & 0 \\ 0 & \Sigma_2 \end{bmatrix}.$$

Thus $\Sigma_2^{-\frac{1}{2}} L(z) \Sigma_1^{\frac{1}{2}}$ is a stable all-pass, i.e. inner function.

The following proposition is the discrete time version of Theorem 4.4 in *Green-Anderson* (Green and Anderson, 1987b).

PROPOSITION 7.15. (Cf. the short discussion after Remark 4.29.) *The number of positive factorization indices equals to the rank of B_{11} . The positive factorization indices coincide with the reachability invariants of (A_1^{11}, B_{11}) .*

Proof. The coordinates of $x_{s1}^{\Phi}(0)$ span the subspace $\mathcal{K}_1$. The positive factorization indices provide a special basis in this spaces. Namely, introducing the notation

$$U^{-1} \eta_i = \eta'_i$$

for those indices i for which $k_i > 0$ then the vectors

$$U^{-j} \eta'_i, \quad j = 0, \dots, k_i - 1 \quad (k_i > 0)$$

form a basis in $\mathcal{K}_1$.

It can be proved by induction that if

$$\alpha^T x_{s1}^{\Phi}(0) \in \mathcal{K}_j = H^+(\bar{w}_2) \cap U^{1-j} H^-(w_1) \quad (j \geq 2)$$

i.e. it is orthogonal to $w_1(-1), w_1(-2), \dots, w_1(-j+1)$ then

$$\alpha^T B_{11} = 0, \alpha^T A_1^{11} B_{11} = 0, \dots, \alpha^T (A_1^{11})^{j-2} B_{11} = 0.$$

Consequently the linear dependency structure in the subspace $\mathcal{K}_1$ coincide with that of the columns of the controllability matrix

$$[B_{11}, A_1^{11} B_{11}, \dots]$$

proving both assertions of the proposition. ■

7.5. Reconstruction of the output process from input noises. Geometric approach

We return to the problem of factorizing the $p \times p$ all-pass function Φ in the form

$$\Phi = \overline{W}_2^{-1} W_1 \quad (7.43)$$

where W_1 is a $p \times p$ analytic function, $\overline{W}_2$ is a $p \times p$ coanalytic one. First let us fix a factorization and define the stationary process $y(t)$, $t \in \mathbb{Z}$, as

$$y(t) = W_1 w_1(t) \quad (7.44)$$

we obtain that

$$y(t) = \overline{W}_2 \overline{w}_2(t), \quad (7.45)$$

especially W_1 and $\overline{W}_2$ define spectral factors of the process y .

We would like to characterize the factorizations for which $W_1, \overline{W}_2$ determine *minimal* (forward and backward) *realizations* of y .

Considering the subspaces $H^-(y)$, $H^+(y)$ we can define the state spaces of these spectral factors as

$$X_{w_1} = \Pr_{H^-(w_1)} H^+(y) \quad (7.46)$$

and

$$X_{\overline{w}_2} = \Pr_{H^+(\overline{w}_2)} H^-(y). \quad (7.47)$$

We restrict ourselves to a special class of solutions of the factorization problem above, namely we assume that $X_{w_1}, X_{\overline{w}_2}$ are *minimal, internal* Markovian splitting subspaces and moreover

$$X_{\overline{w}_2} > X_{w_1}$$

in the sense of ordering defined by Lindquist and Picci (Lindquist and Picci, 1990). (See Definition 5.39.) Theorem 5.41 shows that in the present case the assumption $X_{\overline{w}_2} > X_{w_1}$ is equivalent to the relation $\Pr_{H^-(w_1)} X_{\overline{w}_2} = X_{w_1}$.

The next statement claims that these assumptions determine uniquely the state spaces $X_{w_1}, X_{\overline{w}_2}$.

LEMMA 7.16. *If X_{w_1} , $X_{\bar{w}_2}$ are minimal internal Markovian splitting subspaces for which the relation*

$$\Pr_{H^-(w_1)} X_{\bar{w}_2} = X_{w_1} \quad (7.48)$$

holds then

$$X_{w_1} = \Pr_{H^-(w_1)} H^+(\bar{w}_2) \quad (= X_s),$$

$$X_{\bar{w}_2} = \Pr_{H^+(\bar{w}_2)} H^-(w_1) \quad (= \bar{X}_u).$$

Proof. Since X_{w_1} , $X_{\bar{w}_2}$ determine internal realizations in view of Remark 5.17 we have that

$$H^-(w_1) = X_{w_1} \vee H^-(y),$$

$$H^+(\bar{w}_2) = X_{\bar{w}_2} \vee H^+(y).$$

Moreover, invoking Theorem 5.41 the assumption 7.48 implies that

$$H^-(y) \subset H^-(w_1) \subset H^-(y) \vee X_{\bar{w}_2}.$$

Thus

$$X_{\bar{w}_2} = \Pr_{H^+(\bar{w}_2)} H^-(y) = \Pr_{H^+(\bar{w}_2)} H^-(w_1).$$

Using that

$$\begin{aligned} H^+(\bar{w}_2) &= [H^+(\bar{w}_2) \cap H^+(w_1)] \oplus \Pr_{H^+(\bar{w}_2)} H^-(w_1) = \\ &= [H^+(\bar{w}_2) \cap H^+(w_1)] \oplus X_{\bar{w}_2} \end{aligned}$$

from 7.48 it follows that

$$X_{w_1} = \Pr_{H^-(w_1)} H^+(\bar{w}_2).$$

Note that the previous Lemma shows that the assumption $X_{\bar{w}_2} > X_{w_1}$ implies that the state processes of W_1 and $\bar{W}_2$ are already fixed, moreover the stable part of the function Φ determines processes

$$x_s^\Phi(t+1) = A_1 x_s^\Phi(t) + B_1 w_1(t),$$

$$\bar{x}_u^\Phi(t) = A_1^T \bar{x}_u^\Phi(t+1) + C_1^T \bar{w}_2(t).$$

Thus Lemma 7.16 provides a condition under which the zero-pole cancellation explained after 7.3 and 7.4 occurs.

As a next step we give a geometric construction of the subspaces $H^-(y)$, $H^+(y)$ following the ideas developed in *Ball-Ran* (Ball and Ran, 1987).

Since

$$H^+(y) = [H^+(y) \cap H^+(w_1)] \oplus \Pr_{H^+(y)} H^-(w_1) \quad (7.49)$$

the construction of $H^+(y)$ is equivalent to the construction of the subspaces

$$H^+(y) \cap H^+(w_1) \text{ and } \Pr_{H^+(y)} H^-(w_1).$$

Corollary 5.24 and Remark 5.26 show that the subspace $\Pr_{H^+(y)} H^-(w_1)$ can be written as an angular subspace over X_{w_1} , i.e. there exists a linear operator

$$T^\times : X_{w_1} \mapsto H^+(w_1)$$

such that

$$\Pr_{H^+(y)} H^-(w_1) = \{f + T^\times f : f \in X_{w_1}\} = (I + T^\times)X_{w_1},$$

where I is the identity operator. At the same time Lemma 5.23 implies that

$$H^+(y) \cap H^+(w_1) = H^+(w_1) \ominus T^\times X_{w_1}.$$

Similarly to 7.49

$$H^-(y) = [H^-(y) \cap H^-(\bar{w}_2)] \oplus \Pr_{H^-(y)} H^+(\bar{w}_2).$$

Again Corollary 5.24 and Remark 5.26 imply – but now using them for backward realization – that there exists a linear transformation

$$T : X_{\bar{w}_2} \mapsto H^-(\bar{w}_2)$$

such that

$$\Pr_{H^-(y)} H^+(\bar{w}_2) = \{g + Tg : g \in X_{\bar{w}_2}\} = (I + T)X_{\bar{w}_2}.$$

The following theorem shows that a construction based on the previous observations can be used for the definition of the subspaces $H^-(y)$ and $H^+(y)$. Let us point out that first we determine only the subspaces H^- , H^+ without defining the generating process y . Based on Lemma 7.16 the state spaces of the factors W_1 and $\bar{W}_2$ are uniquely determined:

$$X_{w_1} = \Pr_{H^-(w_1)} H^+(\bar{w}_2), \quad X_{\bar{w}_2} = \Pr_{H^+(\bar{w}_2)} H^-(w_1).$$

Obviously $X_{w_1} = \Pr_{H^-(w_1)} X_{\bar{w}_2}$.

THEOREM 7.17. Assume that we are given two linear operators

$$T^\times : X_{w_1} \mapsto H^+(w_1), \quad T : X_{\bar{w}_2} \mapsto H^-(\bar{w}_2)$$

such that

- (i) the subspace $TX_{\bar{w}_2}$ is invariant under $\text{Pr}_{H^-(\bar{w}_2)}U$,
- (ii) the subspace $T^\times X_{w_1}$ is invariant under $\text{Pr}_{H^+(w_1)}U^{-1}$,
- (iii) $g + Tg \in H^-(w_1)$, if $g \in X_{\bar{w}_2}$,
- (iv) $f + T^\times f \in H^+(\bar{w}_2)$, if $f \in X_{w_1}$,
- (v) $\{g + Tg : g \in X_{\bar{w}_2}\} \vee \{f + T^\times f : f \in X_{w_1}\} \supset X$, where X is defined in 7.18 or 7.19,
- (vi) $\text{Pr}_{H^-(\bar{w}_2)}U^{-1}g + U^{-1}Tg - T[\text{Pr}_{H^+(\bar{w}_2)}U^{-1}g]$ is orthogonal to $TX_{\bar{w}_2}$ for every $g \in X_{\bar{w}_2}$,
- (vii) $\text{Pr}_{H^+(w_1)}Uf + UT^\times f - T^\times[\text{Pr}_{H^-(w_1)}Uf]$ is orthogonal to $T^\times X_{w_1}$ for every $f \in X_{w_1}$.

Define

$$H^+ = [H^+(w_1) \ominus T^\times X_{w_1}] \oplus \{f + T^\times f : f \in X_{w_1}\} \quad (7.50)$$

$$H^- = [H^-(\bar{w}_2) \ominus TX_{\bar{w}_2}] \oplus \{g + Tg : g \in X_{\bar{w}_2}\}. \quad (7.51)$$

Then

- a) H^+ is a U -invariant, H^- is a U^{-1} -invariant subspaces,
- b) $H^- \vee H^+ = H$,
- c) $\text{Pr}_{H^-(w_1)}H^+ = X_{w_1}$, $\text{Pr}_{H^+(\bar{w}_2)}H^- = X_{\bar{w}_2}$,
- d) X_{w_1} , $X_{\bar{w}_2}$ are minimal Markovian splitting subspaces of H^- , H^+ .

Proof. Obviously

$$\text{Pr}_{H^-(w_1)}H^+ = \text{Pr}_{H^-(w_1)}\{f + T^\times f : f \in X_{w_1}\} = \text{Pr}_{H^-(w_1)}X_{w_1} = X_{w_1}.$$

Similarly, we get that

$$\text{Pr}_{H^+(\bar{w}_2)}H^- = X_{\bar{w}_2},$$

proving c).

As concerns a) the assumption (i) implies that $H^-(\bar{w}_2) \ominus TX_{\bar{w}_2}$ is a U^{-1} -invariant subspace. So to prove that H^- is also U^{-1} -invariant we have to show that for any $g \in X_{\bar{w}_2}$ there exists an $h \in X_{\bar{w}_2}$ and $\xi \in H^-(\bar{w}_2) \ominus TX_{\bar{w}_2}$ such that

$$U^{-1}(g + Tg) = (h + Th) + \xi.$$

Since $U^{-1}Tg, Th, \xi \in H^-(\bar{w}_2), h \in H^+(\bar{w}_2)$ we must have that

$$h = \Pr_{H^+(\bar{w}_2)} U^{-1}g.$$

So

$$\xi = \Pr_{H^-(\bar{w}_2)} U^{-1}g + U^{-1}Tg - T[\Pr_{H^+(\bar{w}_2)} U^{-1}g].$$

The right hand side is in $H^-(\bar{w}_2)$ so it is enough to require that it be orthogonal to $TX_{\bar{w}_2}$ which is our assumption (vi).

Similar reasoning gives that from (ii) and (vii) it follows that H^+ is a U -invariant subspace.

To prove b) observe that

$$\begin{aligned} H^- \vee H^+ &= (H^-(\bar{w}_2) \ominus TX_{\bar{w}_2}) \\ &\oplus [\{g + Tg : g \in X_{\bar{w}_2}\} \vee \{f + T^\times f : f \in X_{w_1}\}] \\ &\oplus [H^+(w_1) \ominus T^\times X_{w_1}]. \end{aligned} \quad (7.52)$$

Assumption (v) implies that this contains $X = X_{w_1} \vee X_{\bar{w}_2}$, so the subspaces $TX_{\bar{w}_2}, TX_{w_1}$ are also contained in $H^- \vee H^+$. Comparing this with the first and third term of the decomposition in 7.52 we get that

$$H^- \vee H^+ \supset H^-(\bar{w}_2) \vee H^+(w_1) \vee X = H.$$

Comparing 7.52 to

$$H = [H^-(w_1) \cap H^-(\bar{w}_2)] \oplus X \oplus [H^+(w_1) \cap H^+(\bar{w}_2)]$$

we obtain that

$$H^-(\bar{w}_2) \ominus TX_{\bar{w}_2} \subset H^-(\bar{w}_2) \cap H^-(w_1). \quad (7.53)$$

In fact $H^-(\bar{w}_2) \ominus TX_{\bar{w}_2}$ is orthogonal to X and $H^+(\bar{w}_2) \cap H^+(w_1)$. 7.53 together with (iii) gives that $H^- \subset H^-(w_1)$. Similar reasoning gives that $H^+ \subset H^+(\bar{w}_2)$.

Thus from c) it already follows that $X_{w_1}, X_{\bar{w}_2}$, are Markovian splitting subspaces of H^-, H^+ .

To prove the minimality first we show that

$$X_{\bar{w}_2} \cap \{f + T^\times f : f \in X_{w_1}\}^\perp = 0, \quad (7.54)$$

$$X_{w_1} \cap \{g + Tg : g \in X_{\bar{w}_2}\}^\perp = 0. \quad (7.55)$$

In fact, assume that $\xi \in X_{\bar{w}_2}$ and $\xi \perp f + T^\times f$ for all $f \in X_{w_1}$. Consider the element $f = \Pr_{H^-(w_1)} \xi \in X_{w_1}$. Let us observe that

$$(f + T^\times f) - \xi = T^\times f - \Pr_{H^+(w_1)} \xi \in H^+(w_1) \cap H^+(\bar{w}_2) = H^+(\bar{w}_2) \ominus X_{\bar{w}_2}.$$

Thus it is orthogonal to $X_{\bar{w}_2}$ containing ξ . This implies that $\xi \perp \xi$ i.e. $\xi = 0$ proving 7.54.

The proof of 7.55 is similar.

Since $\{f + T^\times f : f \in X_{w_1}\}^\perp \supset H^+(y)^\perp$, 7.54 ensures that

$$X_{\bar{w}_2} \cap H^+(y)^\perp = 0. \quad (7.56)$$

On the other hand

$$X_{\bar{w}_2} \cap H^-(y) = 0 \quad (7.57)$$

because if $\xi \in X_{\bar{w}_2} \subset H^+(\bar{w}_2)$ and $\xi \perp H^-$ then $\xi \perp X_{\bar{w}_2}$ using that $H^-(\bar{w}_2) \ominus TX_{\bar{w}_2} \subset H^-(\bar{w}_2)$ and $TX_{\bar{w}_2} \subset H^-(\bar{w}_2)$, implying that $\xi = 0$.

Equations 7.56 and 7.57 give that $X_{\bar{w}_2}$ is a minimal splitting subspace.

The minimality of X_{w_1} can be proved similarly. ■

REMARK 7.18. *Immediate consequence of the previous theorem that*

$$H^-(\bar{w}_2) \ominus TX_{\bar{w}_2} = H^-(\bar{w}_2) \cap H^-. \quad (7.58)$$

In fact, since $H^- \subset H^-(\bar{w}_2)$ the inclusion

$$H^-(\bar{w}_2) \ominus TX_{\bar{w}_2} \subset H^-(\bar{w}_2) \cap H^-$$

obviously holds. On the other hand $X_{\bar{w}_2} \subset H^-(\bar{w}_2)$ and $TX_{\bar{w}_2} \perp H^-(\bar{w}_2)$, thus $(I + T)X_{\bar{w}_2} \cap H^-(\bar{w}_2) = 0$ proving the identity 7.58 above.

Similarly,

$$H^+(w_1) \ominus T^\times X_{w_1} = H^+(w_1) \cap H^+. \quad (7.59)$$

REMARK 7.19. *Observe that since $\text{Pr}_{H^-(\bar{w}_2)} X = \bar{X}_s$ conditions (iii), (iv) and (v) in Theorem 7.17 imply that*

$$TX_{\bar{w}_2} \supset \bar{X}_s \quad (7.60)$$

proving again that $H^-(\bar{w}_2) \ominus TX_{\bar{w}_2} \subset H^-(\bar{w}_2) \ominus \bar{X}_s = H^-(\bar{w}_2) \cap H^-(w_1)$ again.

Similar reasoning proves that conditions (iii), (iv) and (v) imply the inclusion

$$T^\times X_{w_1} \supset X_u. \quad (7.61)$$

PROPOSITION 7.20. *Under the assumptions of Theorem 7.17*

(i)

$$TX_{\bar{w}_2} = \bar{X}_s \oplus \{TX_{\bar{w}_2} \cap [H^-(\bar{w}_2) \cap H^-(w_1)]\} \quad (7.62)$$

(ii)

$$T^\times X_{w_1} = X_u \oplus \{T^\times X_{w_1} \cap [H^+(\bar{w}_2) \cap H^+(w_1)]\} \quad (7.63)$$

Proof. (i) Since

$$TX_{\bar{w}_2} \subset H^-(\bar{w}_2) = \bar{X}_s \oplus [H^-(\bar{w}_2) \cap H^-(w_1)],$$

the inclusion 7.60 implies 7.62.

(ii) Equation 7.63 can be proved via similar reasoning. ■

COROLLARY 7.21. *Under the assumptions of Theorem 7.17*

$$(i) H^- \vee X_{w_1} = H^-(w_1) \subset H^- \vee X_{\bar{w}_2}$$

$$(ii) H^+ \vee X_{\bar{w}_2} = H^+(\bar{w}_2) \subset H^+ \vee X_{w_1}.$$

Proof. (i) As we have seen $H^- \supset H^-(w_1) \cap H^-(\bar{w}_2)$. At the same time $[H^-(w_1) \cap H^-(\bar{w}_2)] \oplus X_{w_1} = H^-(w_1)$. Furthermore, using 7.51 we obtain that

$$H^- \vee X_{\bar{w}_2} = H^-(\bar{w}_2) \oplus X_{\bar{w}_2} \supset H^-(w_1).$$

The latter inclusion follows from the observation that

$$[H^-(\bar{w}_2) \oplus X_{\bar{w}_2}]^\perp = H^+(\bar{w}_2) \cap H^+(w_1).$$

Symmetric argument proves (ii). ■

COROLLARY 7.22. *Under the assumptions of Theorem 7.17*

$$(i) (I + T)X_{\bar{w}_2} = \text{Pr}_{H^-} X_{\bar{w}_2}$$

$$(ii) (I + T^\times)X_{w_1} = \text{Pr}_{H^+} X_{w_1}.$$

Proof. (i) The definition of H^- implies that $(I + T)X_{\bar{w}_2} = \text{Pr}_{H^-} H^+(\bar{w}_2)$. But $H^+(\bar{w}_2) \ominus X_{\bar{w}_2}$ is orthogonal to $H^-(w_1)$ which contains H^- thereby proving (i).

Symmetric argument gives the proof of (ii). ■

Let us introduce the notations

$$X_- = \text{Pr}_{H^-} X_{\bar{w}_2}, \quad (7.64)$$

$$X_+ = \text{Pr}_{H^+} X_{w_1}. \quad (7.65)$$

Since under the assumptions of Theorem 7.17 X_{w_1} and $X_{\bar{w}_2}$ are minimal Markovian splitting subspaces these notations are in coherence with 5.8 and 5.12.

Let us observe that

$$X_{\bar{w}_2} = \text{Pr}_{H^+(\bar{w}_2)} X_-,$$

and

$$TX_{\bar{w}_2} = \text{Pr}_{H^-(\bar{w}_2)} X_-.$$

Similarly

$$X_{w_1} = \text{Pr}_{H^-(w_1)} X_+,$$

and

$$T^\times X_{w_1} = \text{Pr}_{H^+(w_1)} X_+.$$

Let us point out that Theorem 7.17 shows how to construct the subspaces using the operators $T^\times, T$. At the same time Corollary 5.24 and Remark 5.26 show that the converse statement is also true. The subspaces H^+, H^- determine the operators $T^\times, T$, respectively.

REMARK 7.23. *The special case $H^- = H^-(w_1)$ (or/and $H^+ = H^+(\bar{w}_2)$) can easily be described in terms of T ($T^\times$) as follows. In case $H^- = H^-(w_1)$ we must have that*

$$X_{w_1} = \text{Pr}_{H^-(w_1)} H^+(\bar{w}_2) = \{g + Tg : g \in X_{\bar{w}_2}\}$$

thus if $\xi \in X_{w_1}$ then defining

$$g = \text{Pr}_{H^+(\bar{w}_2)} \xi$$

we obtain that

$$Tg = \text{Pr}_{H^-(\bar{w}_2)} \xi.$$

This determines T uniquely. Let us observe that in this case

$$TX_{\bar{w}_2} = \text{Pr}_{H^-(\bar{w}_2)} X_{w_1} = \bar{X}_s.$$

We would like to point out that determining the subspaces H^-, H^+ the process y for which

$$H^- = H^-(y), \quad H^+ = H^+(y) \quad (7.66)$$

is not necessarily uniquely determined, except the case when $H^- \cap H^+ = 0$ (apart from nonsingular p -dimensional transformation defined on the coordinates of the process y). This latter statement is immediate from Theorem 7.10.

Denote by T_0 the linear operator mapping $X_{\bar{w}_2}$ to $\bar{X}_s$ defined as follows:

$$T_0 \left(\text{Pr}_{X_{\bar{w}_2}} \xi \right) = \text{Pr}_{H^-(\bar{w}_2)} \xi, \quad (7.67)$$

where $\xi \in X_{w_1} = X_s$. According to the previous remark this corresponds to the special case $H^- = H^-(w_1)$.

Let us observe that

$$\ker T_0 = [H^-(w_1) \cap H^-(\bar{w}_2)] \quad \text{and} \quad \text{range } T_0 = \bar{X}_s.$$

The following proposition describes the structure of T in the general case, i.e. when $H^- \subset H^-(w_1)$.

PROPOSITION 7.24. *There exists a linear operator*

$$t : [H^-(w_1) \cap H^+(\bar{w}_2)] \mapsto [H^-(w_1) \cap H^-(\bar{w}_2)] , \quad (7.68)$$

and a projection

$$R : X_{\bar{w}_2} \mapsto [H^-(w_1) \cap H^+(\bar{w}_2)] \quad (7.69)$$

such that for every $\xi \in X_{\bar{w}_2}$

$$T\xi = T_0\xi + tR\xi , \quad (7.70)$$

where

$$\ker R = \Pr_{H^+(\bar{w}_2)}(X_- \cap X_s) ,$$

$$\text{range } R \subset [H^-(w_1) \cap H^+(\bar{w}_2)] ,$$

but otherwise $\ker R$ and $\text{range } R$ determine an arbitrary direct sum decomposition of $X_{\bar{w}_2}$.

Proof. Since $(I + T)X_{\bar{w}_2} = X_-$, the operator T can be determined in the following way:

$$T(\Pr_{H^+(\bar{w}_2)}\xi) = \Pr_{H^-(\bar{w}_2)}\xi , \quad \xi \in X_- .$$

On the other hand (see equation 7.62)

$$TX_{\bar{w}_2} = \bar{X}_s \oplus \{TX_{\bar{w}_2} \cap [H^-(w_1) \cap H^-(\bar{w}_2)]\} .$$

Now, for every $\xi \in X_- \subset H^-(w_1)$

$$\Pr_{H^-(\bar{w}_2)}\xi \in (H^-(w_1) \cap H^-(\bar{w}_2))$$

if and only if

$$\Pr_{H^+(\bar{w}_2)}\xi \in (H^-(w_1) \cap H^+(\bar{w}_2)) .$$

Since $\Pr_{H^+(\bar{w}_2)}X_- = X_{\bar{w}_2} \supset (H^-(w_1) \cap H^+(\bar{w}_2))$ and no element in X_- is orthogonal to $X_{\bar{w}_2}$, the operator

$$t : [H^-(w_1) \cap H^+(\bar{w}_2)] \mapsto [H^-(w_1) \cap H^-(\bar{w}_2)]$$

defined by the condition

$$\zeta + t\zeta \in X_- , \quad \zeta \in (H^-(w_1) \cap H^+(\bar{w}_2)) \quad (7.71)$$

is well-defined.

In particular,

$$T\zeta = t\zeta , \quad \text{if } \zeta \in (H^-(w_1) \cap H^+(\bar{w}_2)) . \quad (7.72)$$

Observe that in this case

$$T_0 \zeta = 0. \quad (7.73)$$

On the other hand, if $\xi \in X_-$, then

$$\Pr_{H^-(\bar{w}_2)} \xi \in \bar{X}_s \quad \text{if and only if} \quad \xi \in (X_- \cap X_s).$$

In fact, since $\bar{X}_s \perp [H^-(w_1) \cap H^-(\bar{w}_2)]$ the property $\Pr_{H^-(\bar{w}_2)} \xi \in \bar{X}_s$ implies that $\xi \perp [H^-(w_1) \cap H^-(\bar{w}_2)]$ thus

$$\xi \in \{H^-(w_1) \ominus [H^-(w_1) \cap H^-(\bar{w}_2)]\} = X_s.$$

Conversely, if $\xi \in X_s \cap X_- \perp [H^-(w_1) \cap H^-(\bar{w}_2)]$, then

$$\Pr_{H^-(\bar{w}_2)} \xi \perp [H^-(w_1) \cap H^-(\bar{w}_2)],$$

i.e. $\Pr_{H^-(\bar{w}_2)} \xi \in \bar{X}_s$.

Formulating in another way, if

$$\eta \in \Pr_{H^+(\bar{w}_2)} [X_- \cap X_s], \quad \text{then} \quad T\eta = T_0\eta. \quad (7.74)$$

Since the projection from X_- to $X_{\bar{w}_2}$ is bijective and $\Pr_{H^-(\bar{w}_2)} X_- = TX_{\bar{w}_2}$ the decomposition 7.62 implies that

$$X_{\bar{w}_2} = \Pr_{H^+(\bar{w}_2)} [X_- \cap X_s] \vee [H^-(w_1) \cap H^+(\bar{w}_2)].$$

Let us point out that this decomposition is not necessarily a direct sum decomposition.

Consequently, a projection operator $R : X_{\bar{w}_2} \mapsto [H^-(w_1) \cap H^+(\bar{w}_2)]$, for which

$$\ker R = \Pr_{H^+(\bar{w}_2)} [X_- \cap X_s]$$

and

$$\text{range } R \subset [H^-(w_1) \cap H^+(\bar{w}_2)]$$

can be defined. Obviously, this operator R is not necessarily unique.

Then, for $\xi \in X_{\bar{w}_2}$,

$$\begin{aligned} T\xi &= T(I - R)\xi + TR\xi = T_0(I - R)\xi + tR\xi = \\ &= T_0\xi + tR\xi \end{aligned} \quad (7.75)$$

from equations 7.74, 7.72 and 7.73 proving 7.70. ■

Observe that

$$T_0 X_{\bar{w}_2} \perp t [H^-(w_1) \cap H^+(\bar{w}_2)].$$

Similar representation holds for the operator $T^\times$. Namely, denote by

$$T_0^\times : X_{w_1} \mapsto X_u$$

the operator for which

$$T_0^\times \left(\Pr_{X_{w_1}} \xi \right) = \Pr_{H^+(w_1)} \xi,$$

if $\xi \in X_{\bar{w}_2} = \bar{X}_u$.

PROPOSITION 7.25. *There exists a linear operator*

$$t^\times : [H^-(w_1) \cap H^+(\bar{w}_2)] \mapsto [H^+(w_1) \cap H^+(\bar{w}_2)],$$

and a projection

$$R^\times : X_{w_1} \mapsto [H^-(w_1) \cap H^+(\bar{w}_2)] \quad (7.76)$$

such that for every $\xi \in X_{w_1}$

$$T^\times \xi = T_0^\times \xi + t^\times R^\times \xi, \quad (7.77)$$

where

$$\ker R^\times = \Pr_{H^-(w_1)} (X_+ \cap \bar{X}_u),$$

$$\text{range } R^\times \subset [H^-(w_1) \cap H^+(\bar{w}_2)],$$

but otherwise $\ker R^\times$ and $\text{range } R^\times$ determine an arbitrary direct sum decomposition of X_{w_1} .

Propositions 7.24 and 7.25 show that the operators T and $T^\times$ determining the subspaces H^- and H^+ can be defined using the projections R and $R^\times$ and the operators t and $t^\times$. The next theorem expresses the conditions (i), (ii) and (v) of Theorem 7.17 in terms of these operators.

THEOREM 7.26. *Assume that we are given two linear operators*

$$t^\times : [H^-(w_1) \cap H^+(\bar{w}_2)] \mapsto [H^+(w_1) \cap H^+(\bar{w}_2)],$$

$$t : [H^-(w_1) \cap H^+(\bar{w}_2)] \mapsto [H^-(\bar{w}_2) \cap H^-(w_1)],$$

and two projections

$$R^\times : X_{w_1} \mapsto [H^-(w_1) \cap H^+(\bar{w}_2)],$$

$$R : X_{\bar{w}_2} \mapsto [H^-(w_1) \cap H^+(\bar{w}_2)].$$

such that

$$\ker R^\times \cap [H^-(w_1) \cap H^+(\bar{w}_2)] = \ker t^\times,$$

$$\ker R \cap [H^-(w_1) \cap H^+(\bar{w}_2)] = \ker t.$$

Define the operators T as follows $T^\times$.

$$T^\times = T_0^\times + t^\times R^\times \quad T = T_0 + tR.$$

Then conditions (i), (ii) and (v) of Theorem 7.17 are equivalent to the following:

(i) the subspace

$$t(H^-(w_1) \cap H^+(\bar{w}_2)) \text{ is invariant under } \Pr_{H^-(w_1) \cap H^-(\bar{w}_2)} U, \quad (7.78)$$

(ii) the subspace

$$t^\times(H^-(w_1) \cap H^+(\bar{w}_2)) \text{ is invariant under } \Pr_{H^+(w_1) \cap H^+(\bar{w}_2)} U^{-1}, \quad (7.79)$$

(iii)

$$\ker t \vee \ker t^\times = H^-(w_1) \cap H^+(\bar{w}_2), \quad (7.80)$$

Proof. Let us observe that if $\xi \in [H^-(w_1) \cap H^+(\bar{w}_2)]$, then $tR\xi = t\xi$. In fact, $(I - R)\xi \in (\ker R \cap [H^-(w_1) \cap H^+(\bar{w}_2)])$. But this latter coincide with $\ker t$.

(i) Condition (i) of Theorem 7.17 states that $TX_{\bar{w}_2}$ is invariant under $\Pr_{H^-(\bar{w}_2)} U$.

The construction of T gives that $\text{range } T = \text{range } T_0 \oplus \text{range } t$. Furthermore, $\text{range } T_0 = \bar{X}_s = H^-(\bar{w}_2) \ominus [H^-(w_1) \cap H^-(\bar{w}_2)]$.

Consequently

$$H^-(\bar{w}_2) \ominus TX_{\bar{w}_2} = [H^-(w_1) \cap H^-(\bar{w}_2)] \ominus t(H^-(w_1) \cap H^+(\bar{w}_2)) \quad (7.81)$$

and the corresponding invariance properties of the subspaces

$$TX_{\bar{w}_2} \quad \text{and} \quad t(H^-(w_1) \cap H^+(\bar{w}_2))$$

are equivalent to the U^{-1} invariance of the subspaces standing on the left hand side and the right hand side of equation 7.81, respectively. Thus condition (i) in Theorem 7.17 is equivalent to condition (i) in Theorem 7.26.

The equivalence of (ii) and condition (ii) of Theorem 7.17 can be proved similarly.

(iii) First assume that condition (v) of Theorem 7.17 is fulfilled, i.e.

$$(I + T)X_{\bar{w}_2} \vee (I + T^\times)X_{w_1} \supset X.$$

Thus, if $\xi \in X$ then there exist $f \in X_{w_1}$, $g \in X_{\bar{w}_2}$, such that

$$f + T^\times f + g + Tg = \xi. \quad (7.82)$$

Since in this case $f, g, \xi \in X$ the relation $T^\times f + Tg \in X$ should hold. At the same time equations 7.62 and 7.63 show that $TX_{\bar{w}_2}$ (and TX_{w_1}) can be orthogonally decomposed into a subspace which is in X and another one which is orthogonal to X . So $T^\times f + Tg \in X$ implies that

$$T^\times f \in X_u, \quad Tg \in \bar{X}_s. \quad (7.83)$$

Consequently, $Tg = T_0g$, $T^\times f = T_0^\times f$ and

$$f + T^\times f \in X_{\bar{w}_2}, \quad g + Tg \in X_{w_1}. \quad (7.84)$$

Moreover, if $\xi \in (H^-(w_1) \cap H^+(\bar{w}_2)) = X_{w_1} \cap X_{\bar{w}_2}$, then 7.82 and 7.84 imply that

$$f + T^\times f, g + Tg \in X_{w_1} \cap X_{\bar{w}_2}.$$

In particular, since $T^\times f \in X_u$, $f \in X_{w_1} = X_s$, we have that

$$f \perp T^\times f \quad \text{and} \quad f + T^\times f \perp T^\times f,$$

i.e. $T^\times f = 0$.

Similar argument gives that $Tg = 0$. Using now that

$$f + g = \xi, \quad f \in X_{w_1}, \quad g \in X_{\bar{w}_2},$$

we obtain that

$$f, g \in X_{w_1} \cap X_{\bar{w}_2}$$

and from 7.83 and 7.84 it follows that

$$tg = 0.$$

Similarly,

$$t^\times f = 0.$$

In other words

$$\ker t \vee \ker t^\times = H^-(w_1) \cap H^+(\bar{w}_2),$$

proving (iii).

Conversely, assume that (iii) holds. Obviously, if

$$\xi \in (\ker R \cap [H^-(w_1) \cap H^+(\bar{w}_2)]), \quad \text{then } T\xi = 0, \text{ i.e. } (I + T)\xi = \xi.$$

Applying a similar argument to the operator $T^\times$ we obtain that condition (iii) assures that

$$(I + T)X_{\bar{w}_2} \vee (I + T^\times)X_{w_1} \supset [H^-(w_1) \cap H^+(\bar{w}_2)]. \quad (7.85)$$

Furthermore, on the subspace $\ker R$ the operators T and T_0 coincide. On the other hand, $\text{range } R \subset [H^-(w_1) \cap H^+(\bar{w}_2)]$ this latter equal $\ker T_0$. Now, if $\xi \in X_{\bar{w}_2}$ then

$$\begin{aligned}(I + T_0)\xi &= (I + T_0)R\xi + (I + T_0)(I - R)\xi = \\ &= R\xi + (I + T)(I - R)\xi.\end{aligned}$$

Applying a similar argument to the operator $T^\times$ and using equation 7.85 and the identities $(I + T)X_{\bar{w}_2} = X_{w_1}$ and $(I + T^\times)X_{w_1} = X_{\bar{w}_2}$ we obtain that

$$\begin{aligned}(I + T)X_{\bar{w}_2} \vee (I + T^\times)X_{w_1} &\supset [H^-(w_1) \cap H^+(\bar{w}_2)] \vee X_{\bar{w}_2} \vee X_{w_1} \\ &= X,\end{aligned}$$

concluding the proof. ■

Let us point out that straightforward calculation gives that for the operators T and $T^\times$ constructed in Theorem 7.26 conditions (iii) and (iv) of Theorem 7.17 are automatically fulfilled.

REMARK 7.27. Consider shortly conditions (vi) and (vii) of Theorem 7.17. Recall, that – based on the assumptions that $(I + T)X_{\bar{w}_2} \subset H^-(w_1)$, and $TX_{\bar{w}_2}$ is invariant under the operator $\text{Pr}_{H^-(\bar{w}_2)}U$ – condition (vi) was equivalent to the U^{-1} -invariance of H^- . I.e. for every $g \in X_{\bar{w}_2}$

$$U^{-1}(g + Tg) = h + Th + \xi,$$

for appropriate $h \in X_{\bar{w}_2}$ and $\xi \in H^-(\bar{w}_2) \ominus TX_{\bar{w}_2}$. Now, according to 7.70

$$Tg = T_0g + tRg.$$

Apply this to the vector h , as well.

$$U^{-1}(g + T_0g + tRg) = h + T_0h + tRh + \xi.$$

Observe, that $U^{-1}tRg$, tRh and ξ belong to the subspace $H^-(w_1) \cap H^-(\bar{w}_2)$, but $h + T_0h \in X_{w_1} \perp [H^-(w_1) \cap H^-(\bar{w}_2)]$. Let us introduce the notations $\eta = g + T_0g$, $\zeta = h + T_0h$. Then

$$\zeta = \text{Pr}_{X_s} U^{-1} \eta$$

should hold. Moreover, $g = \text{Pr}_{H^+(\bar{w}_2)} \eta$, $h = \text{Pr}_{H^+(\bar{w}_2)} \zeta$. Thus,

$$\xi = U^{-1} \eta - \text{Pr}_{X_s} U^{-1} \eta + U^{-1} tR \text{Pr}_{H^+(\bar{w}_2)} \eta - tR \text{Pr}_{H^-(\bar{w}_2)} \text{Pr}_{X_s} U^{-1} \eta.$$

Since this latter is an element of $H^-(w_1)$ and obviously orthogonal to X_s the condition the ξ is perpendicular to $TX_{\bar{w}_2}$ is equivalent to

$$\xi \perp tRX_{\bar{w}_2},$$

using either 7.62 or Theorem 7.26. This proves that condition (vi) in Theorem 7.17 can be changed to the following one:

(vi') for every $\eta \in X_s$

$$\begin{aligned} & [U^{-1}\eta - \Pr_{X_s} U^{-1}\eta + U^{-1}tR\Pr_{H^+(\bar{w}_2)}\eta - tR\Pr_{H^-(\bar{w}_2)}\Pr_{X_s} U^{-1}\eta] \\ & \perp tRX_{\bar{w}_2}. \end{aligned}$$

Similarly, condition (vii) in Theorem 7.17 is equivalent to the following:

(vii') for every $\eta \in \bar{X}_s$

$$\begin{aligned} & [U\eta - \Pr_{\bar{X}_s}\eta + Ut^\times R^\times \Pr_{H^-(w_1)}\eta - tR\Pr_{H^+(w_1)}\Pr_{\bar{X}_s} U\eta] \\ & \perp t^\times R^\times X_{w_1}. \end{aligned}$$

The following lemma describes the difference between the subspaces $H^-(w_1)$, H^- . (Cf. Lemmata 5.30 and 5.31.)

LEMMA 7.28. Under the assumptions of Theorem 7.17

$$(i) \quad H^- \subset H^-(w_1) \subset H^- \oplus \left\{ \eta - \Pr_{(I+T)X_{\bar{w}_2}} \eta : \eta \in X_{\bar{w}_2} \right\},$$

$$(ii) \quad H^+ \subset H^+(\bar{w}_2) \subset H^+ \oplus \left\{ \xi - \Pr_{(I+T^\times)X_{w_1}} \xi : \xi \in X_{w_1} \right\}.$$

Proof. (i) Since $\Pr_{(I+T)X_{\bar{w}_2}} \eta = \Pr_{H^-} \eta$ if $\eta \in X_{\bar{w}_2}$, thus $H^- \perp \eta - \Pr_{(I+T)X_{\bar{w}_2}} \eta$ and $H^- \oplus \{ \eta - \Pr_{(I+T)X_{\bar{w}_2}} \eta : \eta \in X_{\bar{w}_2} \}$ containing $H^-(w_1)$ according to Theorem 5.41 using the assumption $X_{\bar{w}_2} > X_{w_1}$

Similar argument proves (ii). ■

7.6. State-space description of the factorization

According to Theorem 7.17 the conditions (i)–(vii) ensure the existence of the subspaces H^- , H^+ satisfying conditions a)–d) in that theorem. (Note that they are equivalent to each other.) In this section first we reformulate conditions (i)–(vii) in terms of matrices. This argument is divided into four steps, and finally we provide state-space representation of stationary processes y for which $H^-(y) = H^-$ and $H^+(y) = H^+$ hold true in this way describing the generalized Wiener-Hopf factorizations of the all-pass function $\Phi(z)$.

The conditions (i)–(vii) list some properties of the operators

$$T : X_{\bar{w}_2} \mapsto H^-(\bar{w}_2), \quad T^\times : X_{w_1} \mapsto H^+(w_1).$$

To give a matrix representation of these operators, it is enough to determine their value on the elements of a basis in $X_{\bar{w}_2}$ (and X_{w_1} , respectively), so for example on the coordinate vectors of $\bar{x}_u^\Phi(0)$ and $x_s^\Phi(0)$.

Let us fix the subspaces H^+ and H^- determined by Theorem 7.17

Step 1. Conditions (i)–(ii) of Theorem 7.17.

Condition (ii) of Theorem 7.17 implies that Lemma 5.20 can be applied thus invoking Remark 5.29 we obtain that there exist matrices $S^\times$, H_0 of sizes $n_s \times n_s$ and $n_s \times p$, respectively, for which $\rho(S^\times) < 1$, such that

the coordinates of $x_s^\Phi(0)$ are mapped

$$\text{to the coordinates of } \eta(0) = \sum_{l=0}^{\infty} (S^\times)^l H_0 w(l) \text{ under } T^\times,$$

these latter forming a generator system in $H^+(w_1) \ominus (H^+(w_1) \cap H^+)$.

Similarly, condition (i) of Theorem 7.17 implies applying Lemma 5.20 and Remark 5.29 in “backward” setting that there exist matrices S , G_0 of sizes $n_s \times n_s$ and $n_s \times p$, respectively, for which $\rho(S) < 1$ such that

the coordinates of $\bar{x}_u^\Phi(0)$ are mapped

$$\text{to the coordinates of } \xi(0) = \sum_{l=1}^{\infty} S^{l-1} G_0 \bar{w}_2(-l) \text{ under } T,$$

these latter forming a generator system in $H^-(\bar{w}_2) \ominus (H^-(\bar{w}_2) \cap H^-)$.

Let us introduce the notations

$$\mathcal{G} = \sum_{i=0}^{\infty} S^i G_0 G_0^T (S^T)^i \geq 0,$$

$$\mathcal{H} = \sum_{i=0}^{\infty} S^{\times i} H_0 H_0^T (S^{\times T})^i \geq 0.$$

These are the covariance matrices of $\xi(0) = T \bar{x}_u^\Phi(0)$ and $\eta(0) = T^\times x_s^\Phi(0)$, respectively, satisfying the *Ljapunov*-equations

$$\begin{aligned} \mathcal{G} - S\mathcal{G}S^T &= G_0 G_0^T, \\ \mathcal{H} - S^\times \mathcal{H} S^{\times T} &= H_0 H_0^T. \end{aligned} \tag{7.86}$$

An immediate consequence of Proposition 7.24 that the covariance matrix $\mathcal{G}$ of $\xi(0)$ can be decomposed as follows.

PROPOSITION 7.29. *For the covariance matrix $\mathcal{G}$ the following inequality holds:*

$$\mathcal{G} \geq P_1^{-1} - Q_1. \quad (7.87)$$

Moreover

$$\begin{aligned} \ker(P_1^{-1} - Q_1) \vee \ker \mathcal{G}_0 &= \mathbb{R}^n \\ \text{range}(P_1^{-1} - Q_1) \cap \text{range} \mathcal{G}_0 &= \{0\}, \end{aligned} \quad (7.88)$$

where $\mathcal{G}_0 = \mathcal{G} - [P_1^{-1} - Q_1]$.

Furthermore, define a projection $\Pi : \mathbb{R}^n \mapsto \mathbb{R}^n$ in the following way:

$$\begin{aligned} \ker \Pi &= \ker \mathcal{G}_0 \\ \text{range} \Pi &\subset \ker(P_1^{-1} - Q_1), \end{aligned} \quad (7.89)$$

but otherwise an arbitrary complement of $\ker \Pi$. Then for every $\alpha \in \ker \Pi$

$$T[\alpha^T \bar{x}_u^\Phi(0)] = T_0[\alpha^T \bar{x}_u^\Phi(0)] \in \bar{X}_s \quad (7.90)$$

and for every $\alpha \in \text{range} \Pi$

$$T[\alpha^T \bar{x}_u^\Phi(0)] = t[\alpha^T \bar{x}_u^\Phi(0)] \in H^-(w_1) \cap H^-(\bar{w}_2). \quad (7.91)$$

Proof. Consider first the operator T_0 defined in 7.67. Since

$$P_1^{-1} x_s^\Phi(0) = \bar{x}_u^\Phi(0) + P_1^{-1} P_{12} \bar{x}_s^\Phi(0),$$

we obtain that for every $\alpha \in \mathbb{R}^n$

$$T_0[\alpha^T \bar{x}_u^\Phi(0)] = \alpha^T P_1^{-1} P_{12} \bar{x}_s^\Phi(0).$$

But the covariance matrix of $P_1^{-1} P_{12} \bar{x}_s^\Phi(0)$ is

$$\begin{aligned} -P_1^{-1} P_{12} Q_2 P_{21} P_1^{-1} &= P_1^{-1} P_{12} Q_{21} P_1 P_1^{-1} = \\ &= P_1^{-1} (I - P_1 Q_1) = P_1^{-1} - Q_1, \end{aligned}$$

proving that

$$\mathcal{G} \geq P_1^{-1} - Q_1$$

in view of 7.70, using that $\text{range} T_0 = \bar{X}_s \perp \text{range} t$.

Next observe that, if

$$\alpha^T \bar{x}_u^\Phi(0) \in \text{Pr}_{H^+(\bar{w}_2)}(X_- \cap X_s),$$

then

$$t(\alpha^T \bar{x}_u^\Phi(0)) = 0 \quad \text{i.e.} \quad T(\alpha^T \bar{x}_u^\Phi(0)) = T_0(\alpha^T \bar{x}_u^\Phi(0)),$$

in other words

$$\alpha \in \ker G_0.$$

On the other hand, if

$$\beta^T \bar{x}_u^\Phi(0) \in H^-(w_1) \cap H^+(\bar{w}_2),$$

then

$$T_0(\beta^T \bar{x}_u^\Phi(0)) = 0 \quad \text{i.e.} \quad T(\beta^T \bar{x}_u^\Phi(0)) = t(\beta^T \bar{x}_u^\Phi(0)),$$

or in other words

$$\beta \in \ker(P_1^{-1} - Q_1).$$

Now Proposition 7.24 implies that

$$\ker G_0 \vee \ker(P_1^{-1} - Q_1) = \mathbb{R}^n.$$

Consider now an arbitrary projection Π satisfying 7.89. Then 7.90 is an immediate consequence of the decomposition 7.70.

To prove 7.91 we show that if $\alpha \in \ker(P_1^{-1} - Q_1)$, then

$$T\alpha^T \bar{x}_u^\Phi(0) = t\alpha^T \bar{x}_u^\Phi(0).$$

To this aim it is enough to observe that in this case

$$\alpha^T \bar{x}_u^\Phi(0) \in [H^-(w_1) \cap H^+(\bar{w}_2)] = X_s \cap \bar{X}_u.$$

But

$$\bar{x}_u^\Phi(0) = Q_1 x_s(0) + Q_{12} x_u(0),$$

thus

$$\alpha^T \bar{x}_u^\Phi(0) \in [X_s \cap \bar{X}_u] \quad \text{if and only if} \quad \alpha^T Q_{12} = 0.$$

On the other hand

$$P_1^{-1} - Q_1 = (I - Q_1 P_1) P_1^{-1} = Q_{12} P_{21} P_1^{-1}.$$

Since $X_u \cap \bar{X}_s = \{0\}$ the rows of P_{21} are linearly independent proving that

$$\ker(P_1^{-1} - Q_1) = \ker Q_{21},$$

concluding the proof of 7.91. ■

Similarly, the covariance matrix $\mathcal{H}$ of $\eta(0)$ has the following decomposition.

PROPOSITION 7.30. For the covariance matrix $\mathcal{H}$ the inequality

$$\mathcal{H} \geq Q_1^{-1} - P_1$$

holds true. Moreover,

$$\ker(Q_1^{-1} - P_1) \vee \ker \mathcal{H}_0 = \mathbb{R}^n,$$

where $\mathcal{H}_0 = \mathcal{H} - [Q_1^{-1} - P_1]$.

Furthermore, define the projection $\Pi^\times : \mathbb{R}^n \mapsto \mathbb{R}^n$ in the following way:

$$\ker \Pi^\times = \ker \mathcal{H}_0,$$

and

$$\text{range } \Pi^\times \subset \ker(Q_1^{-1} - P_1),$$

but otherwise an arbitrary complement of $\ker \Pi^\times$. Then, for every $\alpha \in \ker \Pi^\times$

$$T^\times \alpha^T x_s^\Phi(0) = T_0^\times \alpha^T x_s^\Phi(0) \in X_u,$$

and for every $\alpha \in \text{range } \Pi^\times$

$$T^\times \alpha^T x_s^\Phi(0) = t\alpha^T x_s^\Phi(0) \in [H^+(w_1) \cap H^+(\bar{w}_2)].$$

Step 2. Conditions (iii)–(iv) of Theorem 7.17.

LEMMA 7.31.

(i) The assumption $g + Tg \in H^-(w_1)$ for every $g \in X_{\bar{w}_2}$ is equivalent to

$$Q_{12} + \sum_{l=1}^{\infty} S^{l-1} G_0 C_2 (A_2^{-1})^l = 0. \quad (7.92)$$

(ii) the assumption $f + T^\times f \in H^+(\bar{w}_2)$ for every $f \in X_{w_1}$ is equivalent to

$$P_{12} + \sum_{l=0}^{\infty} (S^\times)^l H_0 B_2^T (A_2^T)^{-(l+1)} = 0. \quad (7.93)$$

Proof. (i) The assumption (iii) of Theorem 7.17 can be formulated as the coordinates of

$$\bar{x}_u^\Phi(0) + \sum_{l=1}^{\infty} S^{l-1} G_0 \bar{w}_2(-l)$$

are in $H^-(w_1)$. Using that

$$\begin{aligned}\Pr_{H^+(w_1)} \bar{x}_u^\Phi(0) &= Q_{12} x_u^\Phi(0) \\ \Pr_{H^+(w_1)} \bar{w}_2(-l) &= C_2 \left(A_2^{-1}\right)^l x_u^\Phi(0), \quad l \geq 1\end{aligned}$$

we obtain that

$$\left[Q_{12} + \sum_{l=1}^{\infty} S^{l-1} G_0 C_2 \left(A_2^{-1}\right)^l \right] x_u^\Phi(0) = 0, \quad l \geq 1$$

implying the required identity.

(ii) can be proved similarly. ■

Equations 7.92 and 7.93 are obviously equivalent to

$$Q_{12} A_2 - S Q_{12} + G_0 C_2 = 0 \quad (7.94)$$

and

$$P_{12} A_2^T - S^\times P_{12} + H_0 B_2^T = 0. \quad (7.95)$$

Step 3. Conditions (vi)–(vii) of Theorem 7.17.

LEMMA 7.32.

(i) Under the assumption (i) of Theorem 7.17 the relation

$$\Pr_{H^-(\bar{w}_2)} U^{-1} g + U^{-1} T g - T [\Pr_{H^+(\bar{w}_2)} U^{-1} g] \perp T X_{\bar{w}_2} \quad (7.96)$$

for every $g \in X_{\bar{w}_2}$ is equivalent to the

$$G_0 C_1 + S G - G A_1 = 0. \quad (7.97)$$

(ii) Under the assumption (ii) of Theorem 7.17 the relation

$$\Pr_{H^+(w_1)} U f + U T^\times f - T^\times [\Pr_{H^-(w_1)} U f] \perp T^\times X_{w_1} \quad (7.98)$$

for every $f \in X_{w_1}$ is equivalent to

$$H_0 B_1^T + S^\times \mathcal{H} - \mathcal{H} A_1^T = 0. \quad (7.99)$$

Proof. (i) Expressing the vectors in 7.96 in terms of linear combinations $\bar{w}_2(l)$, $l \in \mathbb{Z}$ we obtain – using the identity $\bar{x}_u^\Phi(0) = \sum_{l=0}^{\infty} (A_1^T)^l C_1^T \bar{w}_2(l)$ – that

$$\begin{aligned}\Pr_{H^-(\bar{w}_2)} \left[\sum_{l=0}^{\infty} A_1^{Tl} C_1^T \bar{w}_2(l-1) \right] + \sum_{l=1}^{\infty} S^{l-1} G_0 \bar{w}_2(-l-1) \\ - A_1^T \sum_{l=1}^{\infty} S^{l-1} G_0 \bar{w}_2(-l) \perp \sum_{l=1}^{\infty} S^{l-1} G_0 \bar{w}_2(-l)\end{aligned}$$

must hold. In other words

$$G_0 C_1 + \sum_{l=2}^{\infty} S^l G_0 G_0^T S^{T(l-1)} - \sum_{l=1}^{\infty} S^{(l-1)} G_0 G_0^T S^{T(l-1)} A_1 = 0.$$

i.e.

$$G_0 C_1 + S\mathcal{G} - \mathcal{G}A_1 = 0.$$

(ii) can be proved similarly. ■

Step 4. We incorporate the analysis of condition (v) of Theorem 7.17 into the proof of the following theorem which gives an equivalent form of Theorem 7.17 in terms of matrices.

THEOREM 7.33. *Assume that the operators $T, T^\times$ are given by*

$$T(\alpha^T \bar{x}_u^\Phi(0)) = \alpha^T \sum_{l=1}^{\infty} G_{l-1} \bar{w}_2(-l), \text{ if } \alpha \in \mathbb{R}^{n_s},$$

$$T^\times(\alpha^T x_s^\Phi(0)) = \alpha^T \sum_{l=0}^{\infty} H_l w(l), \text{ if } \alpha \in \mathbb{R}^{n_s}.$$

Then the assumptions (i)–(vii) are equivalent to the equations:

$$G_l = S^l G_0, \quad H_l = S^{\times l} H_0$$

where $\rho(S) < 1, \rho(S^\times) < 1$,

$$\begin{aligned} H_0 \begin{bmatrix} B_1^T & B_2^T \end{bmatrix} + [-\mathcal{H} \quad P_{12}] \begin{bmatrix} A_1^T & 0 \\ 0 & A_2^T \end{bmatrix} - S^\times [-\mathcal{H} \quad P_{12}] &= 0 \\ G_0 \begin{bmatrix} C_1 & C_2 \end{bmatrix} + [-\mathcal{G} \quad Q_{12}] \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} - S [-\mathcal{G} \quad Q_{12}] &= 0, \end{aligned} \quad (7.100)$$

where

$$\mathcal{G} = \sum_{i=0}^{\infty} G_i G_i^T,$$

$$\mathcal{H} = \sum_{i=0}^{\infty} H_i H_i^T,$$

solving the Ljapunov-equations

$$\begin{aligned} \mathcal{G} - S\mathcal{G}S^T &= G_0 G_0^T \\ \mathcal{H} - S^\times \mathcal{H} S^{\times T} &= H_0 H_0^T \end{aligned} \quad (7.101)$$

and for every $a \in \mathbb{R}^{n_s}$, $b \in \mathbb{R}^{n_s-r}$ there exist $\alpha, \beta \in \mathbb{R}^{n_s}$ such that the identities

$$\begin{aligned} (\mathcal{G} + Q_1 - P_1^{-1})\alpha &= 0, \\ (\mathcal{H} + P_1 - Q_1^{-1})\beta &= 0, \\ Q_{21}Q_1^{-1}\beta &= b, \\ P_1^{-1}\alpha + \beta &= a \end{aligned} \quad (7.102)$$

hold true.

Proof. In view of Lemmata 7.31 and 7.32 it is enough to consider the condition

$$\{g + Tg : g \in X_{\bar{w}_2}\} \vee \{f + T^\times f : f \in X_{w_1}\} \supset X. \quad (7.103)$$

An arbitrary vector in X can be written as

$$\xi + \eta$$

where $\xi \in X_s$, $\eta \in X_u$. Assume that there exist $f \in X_{w_1}$, $g \in X_{\bar{w}_2}$ such that

$$f + T^\times f + g + Tg = \xi + \eta.$$

But

$$f + g + Tg \in H^-(w_1), \quad T^\times f = H^+(w_1)$$

so we have that

$$\eta = T^\times f$$

and

$$\xi = f + g + Tg.$$

Projecting onto $H^-(\bar{w}_2)$ and $H^+(\bar{w}_2)$ we obtain that

$$\text{Pr}_{H^+(\bar{w}_2)}\xi = g + \text{Pr}_{H^+(\bar{w}_2)}f,$$

$$\text{Pr}_{H^-(\bar{w}_2)}\xi = Tg + \text{Pr}_{H^-(\bar{w}_2)}f.$$

Writing ξ as $a^T x_s^\Phi(0)$, η as $b^T x_u^\Phi(0)$ there exist $\alpha, \beta \in \mathbb{R}^{n_s}$ such that defining $f = \beta^T x_s^\Phi(0)$, $g = \alpha^T \bar{x}_u^\Phi(0)$

$$b^T x_u^\Phi(0) = \beta^T \sum_{l=0}^{\infty} H_l w(l), \quad (7.104)$$

$$a^T P_1 \bar{x}_u^\Phi(0) = \alpha^T \bar{x}_u^\Phi(0) + \beta^T P_1 \bar{x}_u^\Phi(0), \quad (7.105)$$

$$a^T P_{12} \bar{x}_s^\Phi(0) = \alpha^T \sum_{l=1}^{\infty} G_{l-1} \bar{w}_2(-l) + \beta^T P_{12} \bar{x}_s^\Phi(0). \quad (7.106)$$

Equation 7.105 can be shortly written as $a = P^{-1}\alpha + \beta$.

Taking the covariances with $x_u^\Phi(0)$ and $\sum_{l=0}^{\infty} H_l w(l)$ of the terms in 7.104 and using 7.99 we conclude that 7.104 is equivalent to the following pair of equations:

$$-b^T P_2 = \beta^T P_{12}$$

$$b^T P_{21} = \beta^T \mathcal{H}.$$

The identities

$$P_{21} Q_1 + P_2 Q_{21} = 0$$

$$P_{12} Q_{21} + P_1 Q_1 = I$$

imply that 7.104 is equivalent to

$$b^T = -\beta^T P_{12} P_2^{-1}$$

$$\beta^T (\mathcal{H} + P_1 - Q_1^{-1}) = 0.$$

Similar computation can be applied to 7.106 leading to the relations

$$a^T P_{12} = -\alpha^T Q_{12} Q_2^{-1} + \beta^T P_{12} \quad (7.107)$$

$$\alpha^T (\mathcal{G} + Q_1 - P_1^{-1}) = 0.$$

But

$$Q_{12} Q_2^{-1} = -P_1^{-1} P_{12}$$

so 7.106 follows already from 7.104.

Summarizing we have obtained that surjectivity of the transformation determined by the matrix

$$\begin{bmatrix} P_1^{-1} & I \\ 0 & Q_{21} Q_1^{-1} \end{bmatrix} \quad (7.108)$$

from $\ker(\mathcal{G} + Q_1 - P_1^{-1}) \oplus \ker(\mathcal{H} + P_1 - Q_1^{-1})$ to $\mathbb{R}^{2n_s-r}$ is equivalent to 7.103, concluding the proof of the Theorem. ■

It can be proved by similar reasoning that condition 7.60 i.e. $TX_{\bar{w}_2} \supset \bar{X}_s$ - assuming that (iii) of Theorem 7.17 holds - $(I+T)X_{\bar{w}_2} \subset H^-(w_1)$ - is equivalent to the following:

— for every $a \in \mathbb{R}^{n_s-r}$ there exists a vector $\alpha \in \mathbb{R}^{n_s}$ such that

$$a = Q_2^{-1} Q_{21} \alpha$$

$$(\mathcal{G} + Q_1 - P_1^{-1}) \alpha = 0.$$

Similarly, condition 7.61 i.e. $T^\times X_{w_1} \supset X_u$ under the assumption that (iv) of Theorem 7.17 – $(I + T^\times)X_{w_1} \subset H^+(\bar{w}_2)$ – holds is equivalent to the following:

- for every $a \in \mathbb{R}^{n_s-r}$ there exists a vector $\alpha \in \mathbb{R}^{n_s}$ such that

$$\begin{aligned} a &= P_2^{-1} P_{21} \alpha \\ (\mathcal{H} + P_1 - Q_1^{-1}) \alpha &= 0. \end{aligned}$$

REMARK 7.34. Observe that 7.80 can be expressed using the projections Π , $\Pi^\times$ as follows. For every $a \in \ker(Q_1^{-1} - P_1)$ there exists an

$$\alpha \in \ker(P_1^{-1} - Q_1) \cap \ker \Pi \quad \text{and} \quad \beta \in \ker(Q_1^{-1} - P_1) \cap \ker \Pi^\times,$$

such that

$$a = P_1^{-1} \alpha + \beta.$$

In fact, $\alpha^T \bar{x}_u^\Phi(0) \in \ker t$ if and only if $\alpha \in \ker(P_1^{-1} - Q_1) \cap \ker \Pi$, and $\beta^T x_s^\Phi(0) \in \ker t^\times$ if and only if $\beta \in \ker(Q_1^{-1} - P_1) \cap \ker \Pi^\times$. Furthermore,

$$\begin{aligned} a^T x_s^\Phi(0) &= \alpha^T \bar{x}_u^\Phi(0) + \beta^T x_s^\Phi(0) = \\ &= \alpha^T P_1^{-1} x_s^\Phi(0) + \beta^T x_s^\Phi(0) \end{aligned}$$

should hold, i.e.

$$a^T = \alpha^T P_1^{-1} + \beta^T.$$

REMARK 7.35. As we have seen in Remark 7.23 the special case $H^- = H^-(w_1)$ can be expressed in terms of $T : X_{\bar{w}_2} \mapsto H^-(\bar{w}_2)$. Since

$$x_s^\Phi(0) = P_1 \bar{x}_u^\Phi(0) + P_{12} \bar{x}_s^\Phi(0)$$

we obtain that in this case

$$T(\alpha^T \bar{x}_u^\Phi(0)) = \alpha^T P_1^{-1} P_{12} \bar{x}_s^\Phi(0)$$

must hold. Immediate calculation gives that

$$\mathcal{G} = -P_1^{-1} P_{12} Q_2 P_{21} P_1^{-1} = P_1^{-1} - Q_1.$$

So in this case

$$\ker(\mathcal{G} + Q_1 - P_1^{-1}) = \mathbb{R}^n.$$

Similarly, $H^+ = H^+(\bar{w}_2)$ implies that

$$\mathcal{H} = Q_1^{-1} - P_1$$

i.e.

$$\ker(\mathcal{H} + P_1 - Q_1^{-1}) = \mathbb{R}^n.$$

Now consider the second equation in 7.100 written as follows:

$$\begin{bmatrix} A_1^T & 0 \\ 0 & A_2^T \end{bmatrix} \begin{bmatrix} \mathcal{G} \\ -Q_{21} \end{bmatrix} = \begin{bmatrix} \mathcal{G} \\ -Q_{21} \end{bmatrix} S^T + \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix} G_0^T.$$

In particular,

$$\begin{bmatrix} A_1^T & 0 \\ 0 & A_2^T \end{bmatrix} \text{range} \begin{bmatrix} \mathcal{G} \\ -Q_{21} \end{bmatrix} \subset \begin{bmatrix} \mathcal{G} \\ -Q_{21} \end{bmatrix} \vee \text{range} \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix}, \quad (7.109)$$

thus $\text{range}[\mathcal{G} - Q_{21}]^T$ is an (A^T, C^T) -invariant subspace. (Cf. Wonham (Wonham, 1977).) But $\mathcal{G} = (P_1^{-1} - Q_1) + G_0$ (see Proposition 7.29). Using that

$$P_1^{-1} - Q_1 = -Q_{12} Q_2^{-1} Q_{21},$$

we obtain

$$\text{range} \begin{bmatrix} P_1^{-1} - Q_1 \\ -Q_{21} \end{bmatrix} = \text{range} \begin{bmatrix} Q_{12} \\ Q_2 \end{bmatrix} Q_2^{-1} Q_{21} = \text{range} \begin{bmatrix} Q_{12} \\ Q_2 \end{bmatrix}$$

due to the row independence of Q_{21} . The second column equation in 7.24 implies that

$$\begin{bmatrix} A_1^T & 0 \\ 0 & A_2^T \end{bmatrix} \begin{bmatrix} Q_{12} \\ Q_2 \end{bmatrix} = \begin{bmatrix} Q_{12} \\ Q_2 \end{bmatrix} A_2^{-1} - \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix} C_2 A_2^{-1}$$

showing that $\text{range}[Q_{12} \ Q_2]^T$ is an (A^T, C^T) -invariant subspace.

Furthermore,

$$\begin{aligned} \text{range} \begin{bmatrix} \mathcal{G} \\ -Q_{21} \end{bmatrix} &= \text{range} \begin{bmatrix} \mathcal{G} \\ -Q_{21} \end{bmatrix} (I - \Pi) \vee \text{range} \begin{bmatrix} \mathcal{G} \\ -Q_{21} \end{bmatrix} \Pi = \\ &= \text{range} \begin{bmatrix} Q_{12} \\ Q_2 \end{bmatrix} \vee \text{range} \begin{bmatrix} G_0 \\ 0 \end{bmatrix} \end{aligned}$$

using that $G_0(I - \Pi) = 0$ and $G\Pi = (I - \Pi + \Pi)^T G\Pi = \Pi^T G\Pi = G_0$.

Thus 7.109 is equivalent to

$$A^T \text{range} \begin{bmatrix} G_0 \\ 0 \end{bmatrix} \subset \text{range} \begin{bmatrix} G_0 \\ 0 \end{bmatrix} \vee \text{range} \begin{bmatrix} Q_{12} \\ Q_2 \end{bmatrix} \vee \text{range } C^T. \quad (7.110)$$

Obviously,

$$A^T \begin{bmatrix} G_0 \\ 0 \end{bmatrix} = \begin{bmatrix} A_1^T G_0 \\ 0 \end{bmatrix}.$$

Since $Q_2^T \alpha + C_2^T \beta = 0$ implies that $\alpha = -Q_2^{-1} C_2^T \beta$ the equation 7.88 can be written as

$$A_1^T \text{range } G_0 \subset \text{range } G_0 \vee \text{range} (C_1^T + P_1^{-1} P_{12} C_2^T). \quad (7.111)$$

using the identity $Q_{12} Q_2^{-1} + P_1^{-1} P_{12} = 0$.

Observe, that $\text{range } G_0 = \text{range } \Pi^T$ in view of 7.89.

Let us point out that 7.111 can be proved in another way, as well. Set

$$\xi_0(0) = T [\Pi^T \bar{x}_u^\Phi(0)] = t [\Pi^T \bar{x}_u^\Phi(0)] \in [H^-(w_1) \cap H^+(\bar{w}_2)].$$

The coordinates of the vector $\xi_0(0)$ should span the invariant subspace $t(H^-(w_1) \cap H^+(\bar{w}_2))$, thus 7.32 implies the existence of appropriate matrices s, g_0 such that

$$\xi_0(0) = \sum_{l=1}^{\infty} s^{l-1} g_0 (\bar{w}_2(-l) - C_2 A_2^{-1} Q_2^{-1} \bar{x}_s^\Phi(-l+1)),$$

i.e.

$$\xi_0(t+1) = s \xi_0(t) + g_0 [\bar{w}_2(-l) - C_2 A_2^{-1} Q_2^{-1} \bar{x}_s^\Phi(-l+1)].$$

Obviously,

$$G_0 = \sum_{l=1}^{\infty} s^{l-1} g_0 (I - C_2 A_2^{-1} Q_2^{-1} (A_2^T)^{-1} C_2^T) g_0^T (s^T)^{l-1},$$

in other words

$$G_0 = s G_0 s^T + g_0 (I - C_2 A_2^{-1} Q_2^{-1} (A_2^T)^{-1} C_2^T) g_0^T, \quad (7.112)$$

so $\text{rank } G_0 = \text{rank } C(s, g_0)$, where $C(s, g_0)$ is the controllability matrix of the pair s, g_0 .

Straightforward calculation gives that condition (vi) of Theorem 7.17 can be expressed as

$$(C_1^T + P_1^{-1} P_{12} C_2^T) g_0^T + G_0 s^T - A_1^T G_0^T = 0, \quad (7.113)$$

implying 7.111

Arguing similarly as in Lemma 5.20 we can assume that

$$\text{range } s = \text{range } [s g_0, s^2 g_0, \dots]$$

and $\rho(s) < 1$, implying that $\text{range } s \subset \text{range } \mathcal{G}_0$, especially $\text{range } \mathcal{G}_0$ is an s -invariant subspace, and s is regular on $\text{range } s$, consequently,

$$\text{range } s + \ker s = \mathbb{R}^n.$$

Especially, choosing a basis in $\text{range } \mathcal{G}_0$ and extending it to a basis in $\mathbb{R}^n$ using vectors from $\ker s$ we can form a matrix X containing these basis vectors as columns.

Then

$$X^{-1} s X = \begin{bmatrix} s_0 & 0 \\ 0 & 0 \end{bmatrix},$$

and

$$\mathcal{G}_0 (X^T)^{-1} = [\tilde{\mathcal{G}}_0 \ 0] \quad g_0^T (X^T)^{-1} = [\tilde{g}_0 \ 0].$$

Now 7.113 implies that

$$\begin{aligned} & (C_1^T + P_1^{-1} P_{12} C_2^T) g_0^T (X^T)^{-1} \\ & + \mathcal{G}_0 (X^T)^{-1} X^T s^T (X^T)^{-1} - A_1^T \mathcal{G}_0 (X^T)^{-1} = 0, \end{aligned}$$

thus

$$(C_1^T + P_1^{-1} P_{12} C_2^T) \tilde{g}_0^T + \tilde{\mathcal{G}}_0 s_0^T - A_1^T \tilde{\mathcal{G}}_0 = 0, \quad (7.114)$$

Equations 7.113 and 7.114 show that the subspace $\text{range } \mathcal{G}_0 = \text{range } \tilde{\mathcal{G}}_0$ is an $(A_1^T, C_1^T + P_1^{-1} P_{12} C_2^T)$ -invariant subspace. Equation 7.89 indicates that this subspace has trivial intersection with $\text{range } (P_1^{-1} - Q_1)$. On the other hand every $(A_1^T, C_1^T + P_1^{-1} P_{12} C_2^T)$ -invariant subspace can be considered as an invariant subspace of the matrix

$$A_1^T - (C_1^T + P_1^{-1} P_{12} C_2^T) F,$$

where F is an appropriate matrix – a “friend” – using the terminology of *Wohnam* (Wonham, 1977). In particular

$$(A_1^T - (C_1^T + P_1^{-1} P_{12} C_2^T) F) \tilde{\mathcal{G}}_0 = \tilde{\mathcal{G}}_0 s_0^T.$$

Since the columns of $\mathcal{G}_0$ are linearly independent, the eigenvalues of s_0^T coincide with those of $A_1^T - (C_1^T + P_1^{-1} P_{12} C_2^T) F$ restricted to the subspace $\text{range } \tilde{\mathcal{G}}_0$.

This observation gives the key for a *converse* construction. Namely, choose an arbitrary matrix F and a subspace $L \subset \mathbb{R}^n$, such that

1.

$$L \cap \text{range} \left(P_1^{-1} - Q_1 \right) = \{0\}, \quad (7.115)$$

2. L is an invariant subspace of $A_1^T - \left(C_1^T + P_1^{-1} P_{12} C_2^T \right) F$ and the corresponding eigenvalues are of modulus less than 1.

Consider a matrix $\widetilde{G}_0$ with columns forming a basis in L . Define the matrix s_0^T of size $\dim L \times \dim L$ as the solution of

$$\left(A_1^T - \left(C_1^T + P_1^{-1} P_{12} C_2^T \right) F \right) \widetilde{G}_0 = \widetilde{G}_0 s_0^T. \quad (7.116)$$

Set $\widetilde{g}_0 = F \widetilde{G}_0$. Furthermore, assume that

3.

$$\text{rank } C(s_0, \widetilde{g}_0) = \dim L = \text{rank } \widetilde{G}_0. \quad (7.117)$$

Form the solution $\widehat{G}_0$ of the equation

$$\begin{aligned} \widehat{G}_0 - \begin{bmatrix} s_0 & 0 \\ 0 & 0 \end{bmatrix} \widehat{G}_0 \begin{bmatrix} s_0^T & 0 \\ 0 & 0 \end{bmatrix} &= \\ &= \begin{bmatrix} \widetilde{g}_0 \\ 0 \end{bmatrix} \left(I - C_2 A_2^{-1} Q_2^{-1} (A_2^T)^{-1} C_2^T \right) [\widetilde{g}_0^T \ 0]. \end{aligned} \quad (7.118)$$

The rank condition 7.117 implies that there exists a regular matrix X such that $[\widehat{G}_0 \ 0]^T = X \widetilde{G}_0$. Set

$$G_0 = [\widehat{G}_0 \ 0] (X^T)^{-1}$$

$$s = X^{-1} \begin{bmatrix} s_0 & 0 \\ 0 & 0 \end{bmatrix} X$$

$$g_0 = X^{-1} \begin{bmatrix} \widetilde{g}_0 \\ 0 \end{bmatrix}.$$

Then

$$\begin{aligned} &\left(C_1^T + P_1^{-1} P_{12} C_2^T \right) g_0^T + G_0 s^T - A_1^T G_0 = \\ &= \left(\left(C_1^T + P_1^{-1} P_{12} C_2^T \right) [\widetilde{g}_0^T \ 0] + [\widetilde{G}_0 s_0^T \ 0] - [A_1^T \widetilde{G}_0 \ 0] \right) \\ &\quad \times (X^T)^{-1} = 0, \end{aligned}$$

invoking 7.117. At the same time

$$\begin{aligned} & \mathcal{G}_0 - s\mathcal{G}_0s^T - g_0 \left(I - C_2A_2^{-1}Q_2^{-1}(A_2^T)^{-1}C_2^T \right) g_0^T = \\ & X^{-1} \left(\begin{bmatrix} \widehat{\mathcal{G}}_0 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} s_0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \widehat{\mathcal{G}}_0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} s_0^T & 0 \\ 0 & 0 \end{bmatrix} \right. \\ & \quad \left. - \begin{bmatrix} \widetilde{g}_0 \\ 0 \end{bmatrix} \left(I - C_2A_2^{-1}Q_2^{-1}(A_2^T)^{-1}C_2^T \right) \begin{bmatrix} \widetilde{g}_0^T & 0 \end{bmatrix} \right) (X^T)^{-1} \\ & = 0, \end{aligned}$$

using 7.118. Consequently,

$$\begin{aligned} \xi_0(0) &= \sum_{l=1}^{\infty} s^{l-1} g_0 \left(\bar{w}_2(-l) - C_2A_2^{-1}Q_2^{-1}\bar{x}_s^{\Phi}(-l+1) \right) \\ &\in [H^-(w_1) \cap H^-(\bar{w}_2)] \end{aligned}$$

is well-defined and equations 7.113 and 7.112 are satisfied. Thus, defining

$$\xi(0) = P_1^{-1}P_{12}\bar{x}_s(0) + \xi_0(0),$$

the corresponding transformation T for which

$$T(\alpha^T \bar{x}_u^{\Phi}(0)) = \alpha^T \xi(0)$$

maps $X_{\bar{w}_2}$ into $H^-(\bar{w}_2)$ and conditions (i), (iii) and (vi) of Theorem 7.17 are fulfilled. Observe, that

$$\text{range } \mathcal{G}_0 = L.$$

Similarly, starting with a matrix $F^{\times}$ and a subspace $L^{\times} \subset \mathbb{R}^n$ for which

$$(i) \quad L^{\times} \cap \text{range} \left(Q_1^{-1} - P_1 \right) = \{0\},$$

$$(ii) \quad L^{\times} \text{ is an } \left[A_1 - \left(B_1 + Q_1^{-1}B_2 \right) F^{\times} \right] \text{-invariant subspace and the corresponding eigenvalues are of modulus less than 1,}$$

$$(iii) \quad \text{rank } C(s_0^{\times}, \widetilde{h}_0) = \dim L^{\times} = \text{rank } \widetilde{\mathcal{H}}_0,$$

where the columns of $\widetilde{\mathcal{H}}_0$ form a basis in $L^{\times}$, the equation

$$\left[A_1 - \left(B_1 + Q_1^{-1}Q_{12}B_2 \right) F^{\times} \right] \widetilde{\mathcal{H}}_0 = \widetilde{\mathcal{H}}_0 (s_0^{\times})^T$$

defines $s_0^\times$ and

$$\widetilde{h}_0^T = F^\times \widetilde{\mathcal{H}}_0,$$

the operator $T^\times$ satisfying conditions (ii), (iv) and (vii) of Theorem 7.17 can be constructed. Moreover, for this construction

$$\text{range } \mathcal{H}_0 = L^\times.$$

It remains to fulfil condition (vii) in Theorem 7.17. To this aim observe that

$$\ker \mathcal{G}_0 = L^\perp \quad \text{and} \quad \ker \mathcal{H}_0 = (L^\times)^\perp.$$

Step 5. Let us return to the state-space description of the generalized Wiener-Hopf factorizations under minimality conditions of the all-pass function Φ . In Theorem 7.33 we have summarized the analysis how to express the conditions originally given in operator form in Theorem 7.17 for the future and past spaces (H^+ and H^-) of the appropriate process y giving the factorization in terms of matrices.

Now let us assume that we have fixed the matrices considered in Theorem 7.33, i.e. we fixed the subspaces H^+ and H^- and characterize the stationary process $y(t)$, $t \in \mathbb{Z}$ for which $H^+(y) = H^+$ and $H^-(y) = H^-$.

To this aim first observe that since $\Pr_{H^-(w_1)} H^+ = X_{w_1} = X_s$ the process y has the form

$$y(t) = C_y x_s^\Phi(t) + D_y w(t). \quad (7.119)$$

This together with the system-equation

$$x_s^\Phi(t+1) = A_1 x_s^\Phi(t) + B_1 w(t) \quad (7.120)$$

defines a minimal realization of y .

In the following lemmata we essentially repeat the construction leading to Theorem 7.17 and 7.33 but now expressing the various building blocks in terms of the matrices in the state-space representation 7.119–7.120. In order to make the proof more transparent the steps are listed in advance.

First we consider the subspace $H^+(y)$. The analysis of the identity $H^+(y) = H^+$ is put together of various building blocks.

- (i) $H^+(y) \subset H^+(\overline{w}_2)$,
- (ii) $H^+(w_1) \cap H^+(y) = H^+(w_1) \cap H^+$.

Observe that Remark 7.37 implies that the subspace

$$H^+(w_1) \ominus [H^+(w_1) \cap H^+]$$

already determines the operator $T^\times$.

- (iii) $H^+(y) \ominus [H^+(y) \cap H^+(w_1)] = H^+ \ominus [H^+ \cap H^+(w_1)]$ using Corollary 7.22
(ii) this latter is equal to the subspace $(I + T^\times)X_{w_1}$.

Obviously (i)–(iii) together imply that $H^+(y) = H^+$.

Next we consider the subspace $H^-(y)$. The analysis of the identity $H^-(y) = H^-$ is carried out as follows:

- (i) $H^-(y) \subset H^-(w_1)$,
(ii) $H^-(w_1) \ominus H^-(y) = H^-(w_1) \ominus H^-$.

LEMMA 7.36. *The assumption $H^+(y) \subset H^+(\bar{w}_2)$ is equivalent to the equation*

$$C_y P_{12} A_2^T + D_y B_2^T = 0. \quad (7.121)$$

Proof. Using the equations 7.119 and 7.120 we can write that

$$\begin{aligned} y(t) &= C_y (P_1 \bar{x}_u^\Phi(t) + P_{12} \bar{x}_s^\Phi(t)) \\ &\quad + D_y (D^T \bar{w}_2(t) + B_1^T \bar{x}_u^\Phi(t+1) + B_2^T \bar{x}_s^\Phi(t+1)) = \\ &= (D_y D^T + C_y P_1 C_1^T + C_y P_{12} C_2^T) \bar{w}_2(t) \\ &\quad + (D_y B_1^T + C_y P_1 A_1^T) \bar{x}_u^\Phi(t+1) \\ &\quad + (D_y B_2^T + C_y P_{12} A_2^T) \bar{x}_s^\Phi(t+1). \end{aligned}$$

The assumption $y(t) \in H^+(\bar{w}_2)$ is fulfilled if and only if the last term vanishes i.e.

$$D_y B_2^T + C_y P_{12} A_2^T = 0. \quad \blacksquare$$

REMARK 7.37. *Introducing the notations*

$$\begin{aligned} D_y D^T + C_y P_1 C_1^T + C_y P_{12} C_2^T &= \bar{D}_y^T \\ D_y B_1^T + C_y P_1 A_1^T &= \bar{C}_y^T \end{aligned} \quad (7.122)$$

we obtain that

$$y(t) = \bar{D}_y^T \bar{w}_2(t) + \bar{C}_y^T \bar{x}_u^\Phi(t+1)$$

and

$$\begin{bmatrix} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ C_1 & C_2 & D \end{bmatrix} \begin{bmatrix} P_1 & P_{12} & 0 \\ P_{21} & P_2 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} C_y^T \\ 0 \\ D_y^T \end{bmatrix} = \begin{bmatrix} \bar{C}_y^T \\ 0 \\ \bar{D}_y^T \end{bmatrix}. \quad (7.123)$$

According to Remark 7.18 we know that

$$T^\times X_{w_1} = H^+(w_1) \ominus [H^+(w_1) \cap H^+]. \quad (7.124)$$

Theorem 7.33 characterizes this subspace and the operator $T^\times$ in terms of the matrices $S^\times, H_0$. On the other hand Theorem 5.27 describes the subspace $H^+(w_1) \ominus [H^+(w_1) \cap H^+(y)]$. Furthermore condition

$$(I + T^\times)X_{w_1} = H^+(y) \ominus [H^+(y) \cap H^+(w_1)] \quad (7.125)$$

should hold as well. Immediate application of Theorem 5.27 gives the following assertion summarizing the analysis of equations 7.124 and 7.125.

COROLLARY 7.38. *Consider the subspace H^+ defined in Theorem 7.17. Assume that a minimal realization of the process $y(t)$ is determined by the equation*

$$y(t) = C_y x_s^\Phi(t) + D_y w(t).$$

Then the identity

$$H^+(y) = H^+$$

is equivalent to the following equations:

- (i) the matrices $\tau = -\mathcal{H} P_1^{-1}$, $h_0 = H_0$, $s^\times = S^\times$ define a maximal – under the condition that $\rho(s^\times) < 1$ – solution of the

$$\begin{bmatrix} \tau P_1 & H_0 \end{bmatrix} \begin{bmatrix} A_1^T & C_y^T \\ B_1^T & D_y^T \end{bmatrix} = s^\times \begin{bmatrix} \tau P_1 & 0 \end{bmatrix}, \quad (7.126)$$

where the maximality is meant in the sense as explained in Theorem 5.27,

(ii)

$$A_2 P_{21} C_y^T + B_2 D_y^T = 0. \quad (7.127)$$

Finally, we need a characterization of $H^-(y)$ in terms of the system matrices. Immediate consequence of 7.119 that

$$H^-(y) \subset H^-(w_1).$$

Thus the equality

$$H^-(w_1) \ominus H^- = H^-(w_1) \ominus H^-(y)$$

is equivalent to the required identity

$$H^-(y) = H^-.$$

Lemma 5.30 characterizes the subspace $H^-(w_1) \ominus H^-(y)$. The following lemma describes the subspace $H^-(w_1) \ominus H^-$ using the matrices defining the operator T .

LEMMA 7.39. *The subspace $H^-(w_1) \ominus H^-$ is spanned by the coordinates of the vector*

$$\sum_{k=1}^{\infty} S^{k-1} (\mathcal{G}B_1 - G_0D - Q_{12}B_2) w_1(-k). \quad (7.128)$$

Proof. Let us point out that Lemma 5.30 shows that the subspace $H^-(w_1) \ominus H^-$ can be generated by the coordinates of the vector $\sum_{k=1}^{\infty} R^{k-1} L_0 W_1(-k)$ where according to Remark 5.32 the matrix R can be chosen to be equal to S used in the definition of the operator T or in other words in the definitions of the subspace $H^-(\bar{w}_2) \ominus [H^-(\bar{w}_2) \cap H^-]$. Thus it remains to determine the corresponding matrix L_0 .

In view of Lemma 7.28 we have to characterize those elements of $H^-(w_1)$ which are orthogonal to $\bar{x}_u^\Phi(0) - \text{Pr}_{(I+T)X_{\bar{w}_2}} \bar{x}_u^\Phi(0)$. Observe that

$$\begin{aligned} & \bar{x}_u^\Phi(0) - \text{Pr}_{(I+T)X_{\bar{w}_2}} \bar{x}_u^\Phi(0) \\ &= \bar{x}_u^\Phi(0) - Q_1(Q_1 + \mathcal{G})^{-1} \left(\bar{x}_u^\Phi(0) + \sum_{j=1}^{\infty} S^{j-1} G_0 \bar{w}_2(-j) \right). \end{aligned}$$

Let us project it to $H^-(w_1)$. Since $\bar{x}_u^\Phi(0) = Q_1 x_s^\Phi(0) + Q_{12} x_u^\Phi(0)$ equations in 7.21 imply that we get the vector

$$\begin{aligned} & [Q_1 - Q_1(Q_1 + \mathcal{G})^{-1} Q_1] x_s^\Phi(0) \\ & - Q_1(Q_1 + \mathcal{G})^{-1} \sum_{j=1}^{\infty} S^{j-1} G_0 (D w_1(-j) + C_1 x_s^\Phi(-j)) \\ & - \sum_{s=0}^{j-1} C_2 (A_2^{-1})^{s+1} B_2 w_1(-j+s) \\ &= Q_1(Q_1 + \mathcal{G})^{-1} \sum_{k=1}^{\infty} S^{k-1} (\mathcal{G}B_1 - G_0D - Q_{12}B_2) w_1(-k) \end{aligned}$$

using the identities

$$\sum_{j=0}^{k-1} S^{j-1} G_0 C_1 A_1^{k-j-1} = S^{k-1} \mathcal{G} - \mathcal{G} A_1^{k-1}$$

and

$$\sum_{s=0}^{\infty} S^s G_0 C_2 (A_2^{-1})^{s+1} = -Q_{12}.$$

(cf. Lemmata 7.31 and 7.32.)

The invertibility of $Q_1(Q_1 + \mathcal{G})^{-1}$ proves the lemma. $\blacksquare$

LEMMA 7.40. The subspace $\text{Pr}_{X_s}(H^-(w_1) \ominus H^-)$ is generated by the coordinates of

$$(\mathcal{G} + Q_1 - P_1^{-1})x_s^\Phi(0).$$

Proof. Since $\text{cov}(w_1(-j), x_s^\Phi(0)) = B_1^T (A_1^T)^{j-1}$ for $j \geq 1$ the statement of the lemma can be formulated in the following way as well:

$$\sum_{j=0}^{\infty} S^j (\mathcal{G}B_1 - G_0D - Q_{12}B_2) B_1^T (A_1^T)^j = (\mathcal{G} + Q_1 - P_1^{-1}) P_1. \quad (7.129)$$

Using the identities

$$B_1 B_1^T = P_1 - A_1 P_1 A_1^T$$

$$DB_1^T + C_1 P_1 A_1^T + C_2 P_{21} A_1^T = 0$$

$$B_2 B_1^T + A_2 P_{21} A_1^T = P_{21}$$

and

$$\mathcal{G}A_1 = S\mathcal{G} + G_0C_1$$

$$SQ_{12} = Q_{12}A_2 + G_0C_2$$

we obtain that

$$(\mathcal{G}B_1 - G_0D - Q_{12}B_2) B_1^T = \mathcal{G}P_1 - S\mathcal{G}P_1 A_1^T + SQ_{12}P_{21} A_1^T - Q_{12}P_{21}$$

giving that

$$\sum_{j=0}^{\infty} S^j (\mathcal{G}B_1 - G_0D - Q_{12}B_2) B_1^T A_1^{Tj} = \mathcal{G}P_1 - Q_{12}P_{21} = (\mathcal{G} + Q_1 - P_1^{-1}) P_1$$

concluding the proof. ■

COROLLARY 7.41. Consider the subspace H^- defined in Theorem 7.17. Assume that a minimal realization of the process $y(t)$ is determined by the

$$y(t) = C_y x_s^\Phi(t) + D_y w(t).$$

Then the identity

$$H^-(y) = H^- \quad (7.130)$$

is equivalent to the following.

The matrices

$$\mu = (G + Q_1 - P_1^{-1})P_1,$$

$$g_0 = (GB_1 - Q_{12}B_2 - G_0D)$$

$$s = S$$

define a maximal solution of the

$$\begin{bmatrix} s\mu & g_0 \end{bmatrix} \begin{bmatrix} A_1^T & C_y^T \\ B_1^T & D_y^T \end{bmatrix} = \begin{bmatrix} \mu & 0 \end{bmatrix} \quad (7.131)$$

under the condition that $\rho(s) < 1$.

Let us point out that although Corollary 7.38 and 7.41 provide necessary and sufficient conditions for the identities $H^+(y) = H^+$ and $H^-(y) = H^-$, respectively, the equations 7.126 and 7.131 are of different structure. The reason of this lies in the fact that in both cases we used a forward realization for $y(t)$, i.e. they are written in terms of the matrices C_y , D_y and A_1 , B_1 . But the following computation shows that 7.131 can be transformed into another equation which is similar to 7.126. In fact,

$$(G + Q_1 - P_1^{-1}) = GP_1 + Q_1P_1 - I = GP_1 - Q_{12}P_{21}.$$

Thus

$$\begin{aligned} & S(G + Q_1 - P_1^{-1})P_1C_y^T + (GB_1 - Q_{12}B_2 - G_0D)D_y^T \\ &= SG P_1C_y^T - SQ_{12}P_{21}C_y^T + GB_1D_y^T - Q_{12}B_2D_y^T - G_0DD_y^T \\ &= (GA_1 - G_0C_1)P_1C_y^T - (Q_{12}A_2 + G_0C_2)P_{21}C_y^T \\ &\quad + GB_1D_y^T - Q_{12}B_2D_y^T - G_0DD_y^T \\ &= G(A_1P_1C_y^T + B_1D_y^T) - G_0(C_1P_1C_y^T + C_2P_{21}C_y^T + DD_y^T) \\ &\quad - Q_{12}(A_2P_{21}C_y^T + B_2D_y^T) \\ &= G\bar{C}_y - G_0\bar{D}_y, \end{aligned}$$

where at the second equation 7.97 and 7.94 were used, while at the fourth 7.123. Consequently, the second equation in 7.131 is equivalent to

$$\begin{bmatrix} -G & G_0 \end{bmatrix} \begin{bmatrix} \bar{C}_y \\ \bar{D}_y \end{bmatrix} = 0,$$

which together with 7.97 can be written as

$$\begin{bmatrix} -G & G_0 \end{bmatrix} \begin{bmatrix} A_1 & \bar{C}_y \\ C_1 & \bar{D}_y \end{bmatrix} = S \begin{bmatrix} -G & G_0 \end{bmatrix},$$

thus characterizing the identity $H^-(y) = H^-$ using the backward minimal realization

$$\bar{x}_u^\Phi(t) = A_1^T \bar{x}_u^\Phi(t+1) + C_1^T \bar{w}_2(t)$$

$$y(t) = \bar{C}_y^T \bar{x}_u^\Phi(t+1) + \bar{D}_y^T \bar{w}_2(t)$$

of the process $y(t)$, $t \in \mathbb{Z}$.

As a summary the following theorem gives an implicit characterization of the identities $H^- = H^-(y)$, $H^+ = H^+(y)$ in terms of equations for the matrices in the observation equation 7.119 C_y, D_y using the previously defined matrices $S, G_0, S^\times, H_0$ determining the operators $T, T^\times$ and in this way the subspaces H^-, H^+ .

THEOREM 7.42. Assume that H^- and H^+ are defined as in Theorem 7.17. Consider the equations

$$\begin{bmatrix} \tau P_1 & h_0 \end{bmatrix} \begin{bmatrix} A_1^T & C_y^T \\ B_1^T & D_y^T \end{bmatrix} = s^\times \begin{bmatrix} \tau P_1 & 0 \end{bmatrix} \quad (7.132)$$

$$\begin{bmatrix} s\mu & g_0 \end{bmatrix} \begin{bmatrix} A_1^T & C_y^T \\ B_1^T & D_y^T \end{bmatrix} = \begin{bmatrix} \mu & 0 \end{bmatrix} \quad (7.133)$$

in $C_y, D_y, \tau, s^\times, h_0, \mu, s, g_0$.

Define

$$y(t) = C_y x_s^\Phi(t) + D_y w(t).$$

Then the identities $H^- = H^-(y)$, $H^+ = H^+(y)$ are equivalent to the following:

- (i) the matrices $s^\times = S^\times$, $h_0 = H_0$ determine a maximal solution of 7.132 under the condition $\rho(s^\times) < 1$,
- (ii) the corresponding matrix τ can be written

$$\tau P_1 = -\mathcal{H},$$

- (iii) the matrices $s = S$, $g_0 = GB_1 - G_0D - Q_{12}B_2$ determine a maximal solution of 7.133 under the condition $\rho(s) < 1$,

(iv) the corresponding matrix μ can be written as

$$\mu = (G + Q_1 - P_1^{-1}) P_1$$

$$(v) A_2 P_{21} C_y^T + B_2 D_y^T = 0.$$

REMARK 7.43. Lemma 7.39 gives that condition $H^- = H^-(w_1)$ is equivalent to the identity

$$GB_1 - G_0 D - Q_{12} B_2 = 0$$

implying that $G = P_1^{-1} - Q_1$ i.e. $\mu = 0$.

8. Canonical reconstruction of stationary processes

In this final concluding-like chapter using the results presented in Chapters 4 and 7 we list some canonical cases of the reconstruction of stationary processes from two given – forward and backward – input white noise processes.

Let $y(t), t \in \mathbb{Z}$ be a zero-mean, full rank stationary process with rational spectral density. As we have seen in Chapter 5 there are usually several forward and backward realizations of this process. Among the forward realizations there are two “extremal” ones, those which are driven by the white noises $w_-(t), t \in \mathbb{Z}$ and $w_+(t), t \in \mathbb{Z}$, respectively. The corresponding state spaces are X_- and X_+ , the first one determines the forward innovation realization or in other words the spectral factor $W_-(z)$ is stable and minimum phase (there are no poles and zeros inside the unit circle), the second one generates the spectral factor $W_+(z)$, which is stable, but maximum phase (the poles are out of the unit circle but the zeros are inside the unit circle). The extremality is meant in the sense of 5.15, i.e. the state covariance matrices satisfy the inequalities

$$P_- \leq P \leq P_+,$$

where P belongs to an arbitrary minimal forward realization (assuming that a uniform choice of bases is applied).

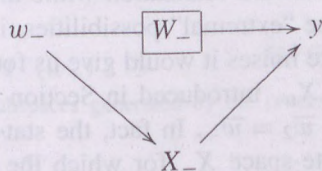


Figure 8.1. Stable, minimum phase (forward innovation) realization

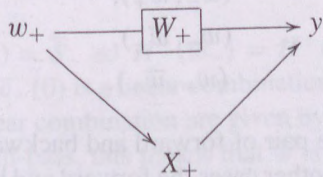


Figure 8.2. Stable, maximum phase (forward) realization

Similarly, the white noises $\bar{w}_+(t), t \in \mathbb{Z}$ and $\bar{w}_-(t), t \in \mathbb{Z}$ determine the “extremal” backward realizations with state spaces X_+ and X_- , respectively. The

corresponding spectral factors are $\bar{W}_+(z)$ and $\bar{W}_-(z)$, respectively. The first one determines the backward innovation realization, i.e. $\bar{W}_+$ is antistable and maximum phase (there are no poles and zeros out of the unit circle), while $\bar{W}_-$ is antistable and minimum phase (the poles are inside the unit circle but the zeros are out of the unit circle).

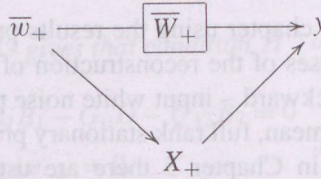


Figure 8.3. Antistable, maximum phase (backward innovation) realization

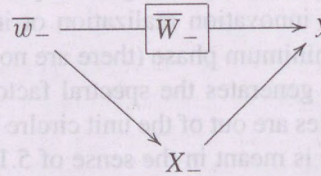


Figure 8.4. Antistable, minimum phase (backward) realization

In Chapter 7 we have analysed the problem of reconstruction of the output process y using the transfer relations between two generating white noises, one of which corresponding to a forward realization while the other one to a backward realizations. Considering the “extremal” possibilities, i.e. pairing the above listed forward and backward white noises it would give us four different different cases, but the assumption $X_{\bar{w}_2} > X_{w_1}$ introduced in Section 7.3 excludes one of these, namely the pair $w_1 = w_+$, $\bar{w}_2 = \bar{w}_-$. In fact, the stat-space associated to w_+ is X_+ , while $\bar{W}_-$ has the state-space X_- for which the opposite inequality holds, $X_- = X_{\bar{w}_2} < X_{w_1} = X_+$. Thus the allowable canonical pairs are

$$\begin{aligned} (w_-, \bar{w}_+), \\ (w_-, \bar{w}_-), \\ (w_+, \bar{w}_+). \end{aligned}$$

So the first possibility is the pair of forward and backward innovation realizations. At the same time in the two other cases the forward and backward realizations share the same state space. Namely, for w_- and $\bar{w}_-$ the state space is X_- , while for w_+ and $\bar{w}_+$ the state space is X_+ .

Figure 8.5 shows the various past and future subspaces arising in the “extremal” realizations of a stationary process y . The past of the output process y and that

of the forward innovation noise w_- coincide. Similarly, for the future subspaces $H^+(y) = H^+(\bar{w}_+)$.

The past subspace of the input noise w^+ , generating the stable, maximum phase spectral factor is larger than the past $H^-(y)$. It contains the state-space X_+ , as well. More precisely, $H^-(\bar{w}_+) = H^-(y) \vee X_+$. Let us recall, that $X_+ = \text{Pr}_{H^+(y)} H^-(y) = H^-(w_+) \cap H^+(\bar{w}_+)$.

Similarly, the future subspace of the input noise $\bar{w}_-$, generating the unstable minimum phase spectral factor contains the future-space $H^+(y)$. In fact, $H^+(\bar{w}_-) = H^+(y) \vee X_-$. Here $X_- = \text{Pr}_{H^-(y)} H^+(y) = H^-(w_-) \cap H^+(\bar{w}_-)$.

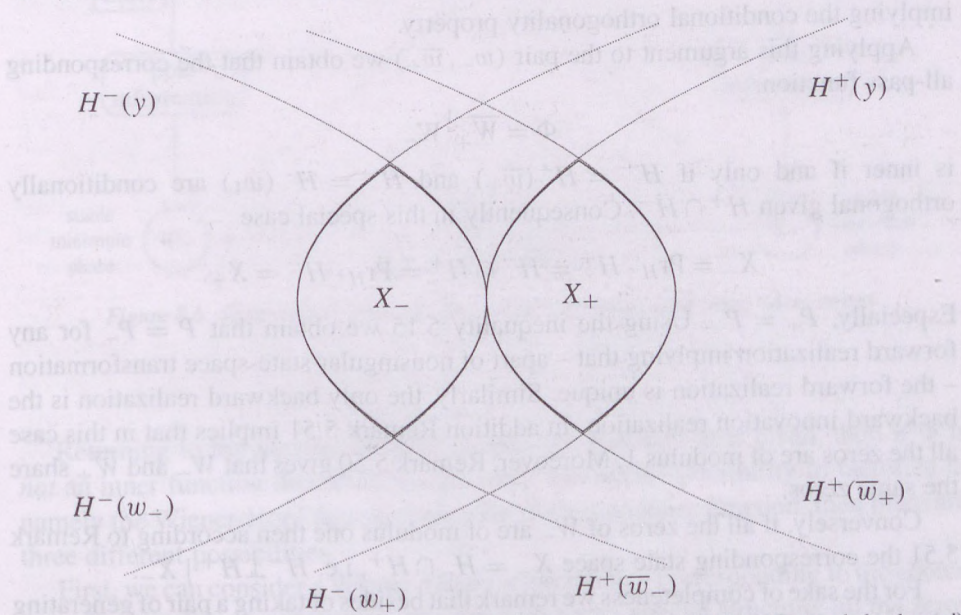


Figure 8.5. Past and future subspaces generated by the process y and the (forward, backward) innovations

Consider first the pair $(w_-, \bar{w}_-)$. In this case the corresponding all-pass function

$$\Phi(z) = \bar{W}_-^{-1} W_-$$

is inner. In fact, $H^+(\bar{w}_-) = \bar{S}_-$ so $H^-(\bar{w}_-) = H^+(\bar{w}_-)^{\perp} = \bar{S}_-^{\perp} = N_- \subset S_- = H^-(w_-)$. Consequently, $\bar{w}_-(0)$ is a linear combination of $w_-(0), w_-(-1), \dots$. But the coefficients in this linear combination are given by the coefficients of $\Phi(z)$, i.e. Φ is analytic. Since it is all-pass, this means that Φ is an inner function.

Similarly, for the pair $w_+, \bar{w}_+$ the all-pass function $\bar{W}_+^{-1} W_+$ is inner, because $H^-(\bar{w}_+) = (H^+)^{\perp} \subset (N^+)^{\perp} = S_+ = H^-(w_+)$.

Observe that the previous argument can be reversed. If Φ , originating from a pair $(w_1, \bar{w}_2)$ is an inner function, then $\bar{w}_2$ is a linear combination of $w_1(0), w_1(-1), \dots$,

consequently $H^-(\bar{w}_2) \subset H^-(w_1)$, or in other words the subspaces $H^+(\bar{w}_2) = \bar{S}_2$ and $H^-(w_1) = S_1$ are perpendicularly intersecting each other, i.e. S_1 and $\bar{S}_2$ are conditionally orthogonal given $S_1 \cap \bar{S}_2$. In fact, let $\xi \in S_1$, then

$$\xi = \text{Pr}_{\bar{S}_2} \xi + \text{Pr}_{\bar{S}_2^\perp} \xi.$$

But $\text{Pr}_{\bar{S}_2^\perp} \xi \in \bar{S}_2^\perp \subset S_1$, so $\text{Pr}_{\bar{S}_2} \xi \in S_1$. Thus

$$\text{Pr}_{\bar{S}_2} S_1 \subset (S_1 \cap \bar{S}_2)$$

implying the conditional orthogonality property.

Applying this argument to the pair $(w_-, \bar{w}_+)$ we obtain that the corresponding all-pass function

$$\Phi = \bar{W}_+^{-1} W_-$$

is inner if and only if $H^+ = H^+(\bar{w}_+)$ and $H^- = H^-(w_-)$ are conditionally orthogonal given $H^+ \cap H^-$. Consequently in this special case

$$X_- = \text{Pr}_{H^-} H^+ = H^- \cap H^+ = \text{Pr}_{H^+} H^- = X_+.$$

Especially, $P_+ = P_-$. Using the inequality 5.15 we obtain that $P = P_-$ for any forward realization implying that – apart of nonsingular state-space transformation – the forward realization is unique. Similarly, the only backward realization is the backward innovation realization. In addition Remark 5.51 implies that in this case all the zeros are of modulus 1. Moreover, Remark 5.50 gives that W_- and $\bar{W}_+$ share the same zeros.

Conversely, if all the zeros of W_- are of modulus one then according to Remark 5.51 the corresponding state space $X_- = H_- \cap H^+$, i.e. $H^- \perp H^+ | X_-$.

For the sake of completeness we remark that besides of taking a pair of generating white noises of opposite types – a forward and a backward noise – it is possible to consider pairs of input noises of the same type. Namely, the transfer function between the the extremal forward noises, thus from w_+ to w_- , i.e.

$$W_-^{-1} W_+$$

is an inner function and similarly the transfer function between the two extremal backward noises, thus from $\bar{w}_-$ to $\bar{w}_+$, i.e.

$$\bar{W}_+^{-1} \bar{W}_-$$

is a conjugate inner function. These functions contain the information about the zero structure. Figure 8.6 demonstrates the possibilities listed above.

Naturally, the directions of the arrows above can be reversed, i.e. we can consider the inverses of the transfer functions above; this transformation for example produces a conjugate inner function from an inner one.

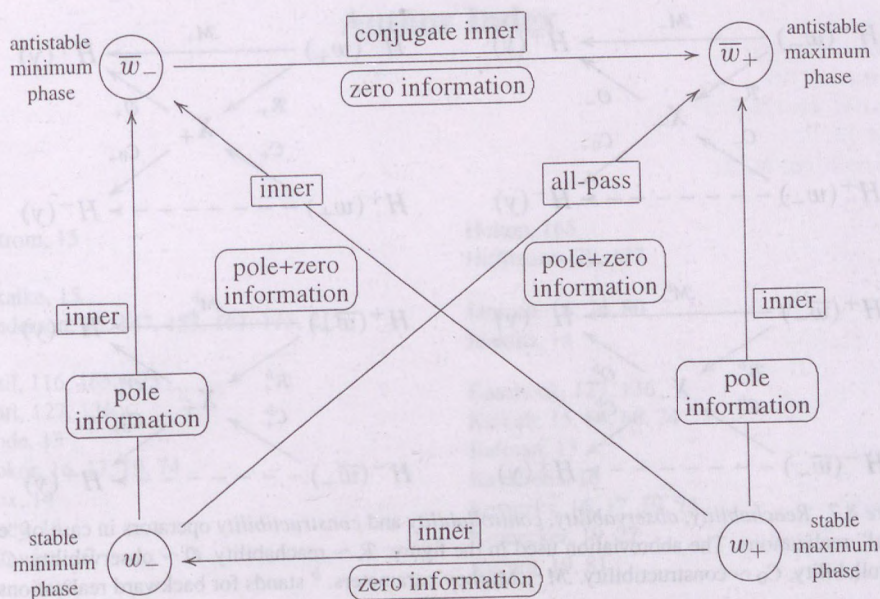


Figure 8.6. Connections between the various "extremal" generating white noises

Returning to the inverse problem, if only the function Φ is given, then if Φ is *not* an inner function then there is only one "canonical" possibility to factorize it, namely the Wiener-Hopf factorization. But, if Φ is an inner function, then there are three different possibilities.

First, we can consider a Wiener-Hopf factorization – corresponding to the choice $(w_-, \bar{w}_+)$ leading to a stationary process with a very special structure: all the zeros of W_- are of modulus 1. (Cf. Corollary 4.2 in *Green and Anderson* (Green and Anderson, 1987b) in the continuous time case.)

As a second case, we can either consider the pair $(W_-, \bar{W}_-)$ or the pair $(W_+, \bar{W}_+)$. Let us point out that $(W_-, \bar{W}_-)$ $(W_+, \bar{W}_+)$ still share the same zeros, but these are out of the unit circle (inside the unit circle, respectively) and not necessarily on the unit circle.

After the factorization $\Phi = \bar{W}_2^{-1} W_1$ we can choose a canonical state-space representation (observability, reachability, constructibility, controllability canonical forms) for the spectral factors W_1 and $\bar{W}_2$ according to the constructions listed in Chapter 4.

Figure 8.7 summarizes the various "extremal" operators corresponding to these canonical parametrizations.

As we have seen to define the constructibility and controllability canonical forms some nonsingularity assumptions are needed.

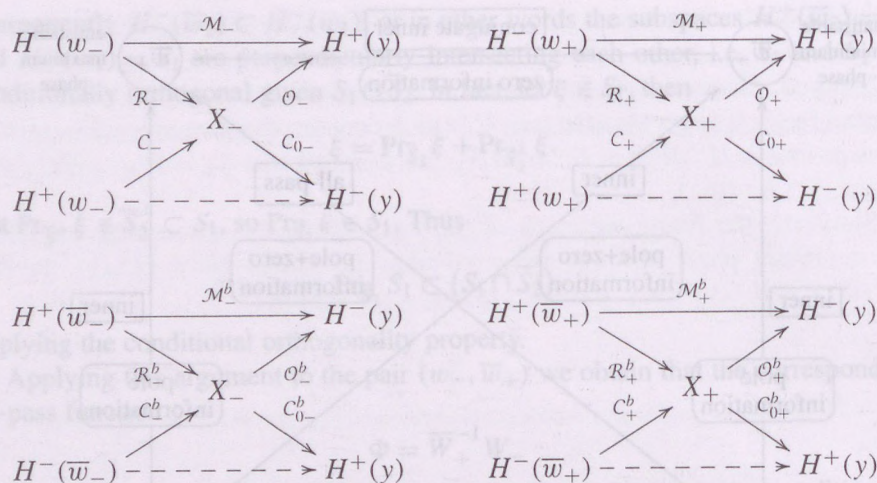


Figure 8.7. Reachability, observability, controllability and constructibility operators in case of “extremal” realizations. The abbreviation used in the figure: $\mathcal{R} \sim$ reachability, $\mathcal{O} \sim$ observability, $\mathcal{C} \sim$ controllability, $\mathcal{C}_0 \sim$ constructibility. $\mathcal{M} \sim$ Markov parameters. b stands for backward realizations

In the statistical identification of the stationary stochastic systems the search of the canonical forms of state-space representations play a significant role. As we have seen canonical parametrization can be based on the notions of reachability, observability, controllability and constructibility. This well known part of the identification procedure and its validity analysis is possible in case of any type of white noise input process. Then the efficiency of the identification of the stochastic process y can be evaluated on the basis of the “whiteness” of the *estimated* source noise time series.

At the same time together with canonical parametrizations we can consider extremal (forward, backward) realizations, spectral factors. This corresponds to the *apriori* specification of the position of poles and zeros with respect to the unit circle. This gives us altogether $4 \times 4 = 16$ different possibilities. But the analysis of these possibilities from the identification point of view may already lead us to the subject of new studies.

Author Index

- Åstrom, 15
- Akaike, 15
- Anderson, 15, 147, 153, 161, 173, 215
- Ball, 116, 165, 175
- Bart, 127, 136
- Bode, 13
- Bokor, 16, 17, 19, 74
- Box, 14
- Bucy, 13
- Bányász, 17
- Caines, 15, 28
- Clancey, 161, 165
- Clark, 16
- Clerget, 48
- Deistler, 15, 16, 74
- Desai, 59
- Dunsmuir, 16
- Dym, 134, 161
- Elliott, 16, 74, 80
- Eykhoff, 15
- Faurre, 15, 19, 48, 98
- Ferrante, 91, 115–117
- Forney, 16, 74
- Fuhrmann, 20, 127, 142, 147, 158
- Gantmacher, 103
- Germain, 48
- Gevers, 16, 74, 80
- Glover, 44, 170
- Gohberg, 116, 127, 136, 161, 165
- Gombani, 158
- Green, 121, 147, 153, 161, 173, 215
- Guidorzi, 74
- Hannan, 15, 16, 121
- Hanzon, 34
- Helton, 165
- Hoffmann, 20, 127
- Janssen, 16, 74, 80
- Jenkins, 14
- Kaashoek, 127, 136
- Kailath, 15, 64, 66, 74
- Kalman, 13
- Kavalieris, 16
- Keviczky, 16, 17, 19, 74
- Kronecker, 19, 27
- Kučera, 19, 81
- Levinson, 13
- Lindquist, 15, 17, 19, 91–94, 96, 97, 114, 158, 174
- Ljung, 15
- Michaletzky, 15, 17, 61, 91, 98, 112, 115–117, 121
- Nyquist, 13
- Ober, 34
- Pal, 59
- Pavon, 15, 17
- Picci, 15, 17, 19, 91–94, 96, 97, 114, 158, 174
- Popov, 64, 65
- Poskitt, 121
- Ran, 165, 175
- Riesz, 19
- Rodman, 116
- Rozanov, 14, 21
- Ruckebusch, 91
- Söderström, 15
- Tusnády, 15, 61, 91, 112
- van den Hof, 16
- Várlaki, 227

Subject Index

- acausality, 114
- all-pass
 - extension, 62
- all-pass factor
 - minimal, left, 142
- all-pass factorization, 129
- all-pass function, 31, 123
- ARMA, 34, 37
 - dual, 39
 - monic, 75
 - nonmonic, 74
- Autoregressive Moving Average, *see* ARMA
- basis
 - orthonormal, 92
- Blaschke-product, 26
- Borel measurable set, 22
- canonical correlation, 61
- canonical form
 - constructibility, 68
 - controllability, 74
 - observability, 71
 - reachability, 72
- canonical forms, 63
- Cayley-Hamilton theorem, 43
- condition
 - for minimality, 140
- conditional orthogonality, 91
- constructibility, 65
 - canonical form, 68
 - subspace, 93
- controllability, 43, 47, 73
 - canonical form, 74
- convex set, 48
- covariance sequence, 21
- dual representations
 - ARMA, 38
 - MFD, 38
- eigenvalue, 131
 - stable, 131
 - unstable, 131
- equivalence class, 65
- extension
 - linear, 126
- factorization
 - all-pass, 136
 - minimal, 136, 139
 - Wiener-Hopf, 151
- Faurre-theorem, 49
- generator system, 60
- geometric theory of SSR's, 91
- Hankel matrix, 55
 - block, 27
 - factorization theorem, 61
- Hankel singular value, 147
- Hilbert space, 22, 92
- inner function
 - factorization, 123
- inner product representation, 61, 112
 - generalized, 112
- innovation
 - process, backward, 56
 - process, forward, 51
 - realization, forward, 51
 - subspace, 51
- invariant
 - complete, 63
 - complete system of $\sim$ s, 67
 - constructibility, 65
 - controllability, 73
 - observability, 64, 69
 - reachability, 71
 - surjective, 63
- Jordan block, 117
- left zero function, 116

linear

- extension, 126
- transformation, 125

linear transformation, 65

Ljapunov equation, 123

Ljapunov-equation, 42

mapping, 126

Markovian process, 91

matrix

- constructibility, 65
- nonsingular, 131
- observability, 64
- orthonormal, 133
- stable, 131
- Toeplitz, 68

Matrix Fractional Description, *see* MFD

MFD, 34, 38

- dual, 40

- Forney's approach, 74

minimality, 47, 132

minimality condition, 140

observability, 47, 50, 64, 69

- canonical form, 71, 76
- subspace, 93

operator

- backward shift, 22
- isometric, 22, 29
- kernel subspace, 144
- linear, 28
- restricted shift, 29
- shift, 22, 28, 50, 92
- unitary, 22, 29

orthogonal

- pairwise, 133

orthonormal basis, 51

phase function, 62, 142

process

- completely regular, 28
- deterministic, 28
- full rank, 28
- full-rank, 148
- nondeterministic, 28
- purely nondeterministic, 28, 43, 148
- rank, 23
- regular, 28
- singular, 28

spectral density, 23

stationary, 148

projection, 127

projection error, 96

projector measure, 22

rational inner function, 26

reachability, 44, 71

- canonical form, 72

realization

- backward, 92
- forward, 91

Redheffer transform, 134

regularity, 28

relation

- equivalence, 63

representative vector system, 61

Riccati equation, 20, 129

scalar product, 50

Scheme II, 64

spectral density, 23

spectral factor

- acausal, 116, 132
- analytic, 33, 115
- coanalytic, 34, 115
- zeros, 116

spectral measure

- stochastic, 23

spectral representation, 21

splitting subspace, 91

- Markovian, 93

splitting subspaces

- minimal, 94

SSR, 34, 35, 97

- acausal, 114

- backward, 97

- canonical, 65

- forward, 97

- minimal, 41, 94

- minimal, backward, 56

- strong, 49

- structure, 67

- weak, 41

State Space Representation, *see* SSRstationary process, *see* process

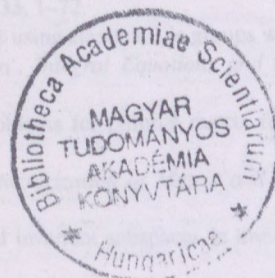
Stone representation theorem, 21

subspace, 51

- constructible, 93

- minimal splitting, 148

- observable, 93
- shift invariant, 128
- time series, 21
- top-companion form, 68
- transfer function, 33
 - inverse, 125
- transfer relation
 - BSSR-BARMA, 77
 - BSSR-BMFD, 79
 - BSSR-dual BARMA, 84
 - FSSR-dual FARMA, 85
 - FSSR-dual MFD, 80
 - FSSR-FARMA, 78
 - FSSR-FMFD, 75
- transfer relations, 74
- transformation
 - linear, 125
 - nonsingular, 47
 - orthogonal, 133
- unit circle, 25
- white noise sequence, 51
- Wiener-Hopf factorization, 20, 151
- Wold decomposition, 21, 28, 29
- Wold-Kolmogorov decomposition, *see* Wold decomposition



References

- Akaike, H. (1974a), 'A new look at the statistical model identification problem'. *IEEE Trans. Automat. Contr.*, **AC-19**, 716–723.
- Akaike, H. (1974b), 'Stochastic theory of minimal realization'. *IEEE Transactions on Automatic Control*, **AC-19**, 667–674.
- Akaike, H. (1975), 'Markovian representation of stochastic processes by canonical variables'. *SIAM Journal of Control*, **13**, 162–173.
- Akaike, H. (1976), 'Canonical correlation analysis of time series and the use of an information criterion'. In: R.K. Mehra and D.G. Lainiotis (eds.), *Systems Identification, Advances and Case Studies*. Academic Press, New York, 27–96.
- Akaike, H. (1977), 'On entropy maximisation principle'. In: P. Krishnaiah (ed.), *Applications of Statistics*. North-Holland, 27–41.
- Anderson, B.D.O. (1967a), 'An algebraic solution to the spectral factorization problem'. *IEEE Transactions on Automatic Control*, **AC-12**, 410–414.
- Anderson, B.D.O. (1967b), 'A system theory criterion for positive real matrices'. *SIAM Journal of Control and Optim.*, **5**, 171–172.
- Anderson, B.D.O. (1969), 'The inverse problem of stationary covariance generation'. *J. Statistical Physics*, **1**, 133–147.
- Anderson, B. D. O. (1973), 'Algebraic properties of minimal degree spectral factors'. *Automatica*, **9**, 491–500.
- Anderson, B.D.O. and T. Kailath (1979), 'Forwards, backwards and dynamically reversible Markovian models of second-order processes'. *IEEE Transactions on Circuits and Systems*, **CAS-26**, 956–965.
- Aoki, M. (1990), *State Space Modeling of Time Series*. Springer-Verlag, 2nd ed.
- Åström, K.J. (1970), *Introduction to stochastic control theory*. Academic Press.
- Åström, K.J. and P. Eykhoff (1971), 'System identification – a survey'. *Automatica*, **7**, 123–162.
- Badawi, F. (1980), 'PhD Thesis'.
- Badawi, F., A. Lindquist and M. Pavon (1979), 'A stochastic realization approach to the smoothing problem'. *IEEE Trans. Automatic Control*, **AC-24**, 878–888.
- Ball, J.A. (1984), 'Invariant Subspace Representations, Unitary Interpolants and Factorization Indices'. Vol. 12 of *Operator Theory*. Birkhäuser, Basel.
- Ball, J.A., I. Gohberg and L. Rodman (1988), 'Realization and interpolation of rational matrix functions'. *Operator Theory, Advances and Applications*, **33**, 1–72.
- Ball, J.A. and W. Helton (1984), 'Beurling-Lax representations using classical Lie groups with many applications II, $(Gl(n, C))$ and Wiener-Hopf factorization'. *Integral Equations and Operator Theory*, **7**, 291–309.
- Ball, J.A. and A.C. Ran (1987), 'Global inverse spectral problems for rational matrix functions'. *Linear Algebra and its Applications*, **86**, 237–282.
- Bart, H., I. Gohberg and M.A. Kaashoek (1984), *Minimal factorization of Matrix and Operator Functions*, Operator Theory. Birkhäuser, Basel.
- Basile, G. and G. Marro (1973), 'Controlled and conditioned invariant subspaces in linear system theory'. *J. Optimization Theory and Appl.*, **3**, 306–316.

- Bittanti, S., A.J. Laub and J. Willems (1991), *The Riccati Equation*. Springer-Verlag.
- Bokor, J. and L. Keviczky (1982), 'Structural properties and structure determination of vector difference equations'. *International Journal of Control*, **36**, 461-475.
- Bokor, J. and L. Keviczky (1987), 'ARMA canonical forms obtained from constructibility invariants'. *International Journal of Control*, **45**, 861-873.
- Bokor, J., L. Keviczky and C. Bányász (1989a), 'Monic ARMA realizations of second order stochastic processes'. In: *21st ISCIE Symp. on Stochastic Systems Theory and its Applications*.
- Bokor, J., L. Keviczky and C. Bányász (1989b), 'Realization of stochastic processes in state-space, MFD and ARMA forms'. In: *Proceedings of the IEEE Conference on Control and Applications*, Jerusalem, WC-5-4.
- Bokor, J., M. Tanyi and G. Terdik (1990), 'Forward and backward Markovian state space models of second order processes'. *Computers Math. Applic.*, **19**, 21-29.
- Box, G. E.P. and G.M. Jenkins (1970), *Time series Analysis, Forecasting and Control*. Holden Day.
- Bucy, R.S., D. Rappaport and L.M. Silverman (1970), 'Correlated noise filtering and invariant directions for the Riccati equation'. *IEEE Trans. Automatic Control*, **AC-15**, 535-540.
- Caines, P.E. (1979), *Linear Stochastic Systems*. Dunod.
- Caines, P.E. (1988), *Linear Stochastic Systems*. Wiley.
- Caines, P.E. and D. Delchamps (1980), 'Splitting subspaces, spectral factorization and the positive real equation, (structural features of the stochastic realization problem'. In: *Proc. 1980 IEEE Conf. Decision and Control*, 358-362.
- Clancey, K. and I. Gohberg (1981), *Factorization of Matrix Functions and Singular Integral Operators*, Vol.3 of *Operator Theory*. Birkhuser Verlag, Basel.
- Clark, J. M.C. (1976), 'The consistent selection of parametrization in system identification'. In: *Proceedings of the Joint Automatic Control Conference*.
- Deistler, M. and M. Gevers (1989), 'Properties of the Parametrization of Monic ARMAX Systems'. *Automatica*, **25**, 87-96.
- Deistler, M. and E.J. Hannan (1981), 'Some properties of the parametrization of ARMA systems with unknown order'. *J. Multivariate Anal.*, **11**, 474-484.
- Desai, U.B. and D. Pal (1982), 'A realization approach to stochastic model reduction and balanced stochastic realization'. In: *Preprint IEEE Conf. on Decision and Control*. 1105-1112.
- Desai, U.B. and D. Pal (1984), 'A transformation approach to stochastic model reduction'. *IEEE Trans. Automat. Control*, **29**, 1097-1100.
- Desoer, A. and J.D. Schulman (1974), 'Zeros and poles of matrix transfer functions and their dynamical interpretations'. *IEEE Trans. Circuits and Systems*, **CAS-21**, 3-8.
- Dunsmuir, W. and J. Hannan (1976), 'Vector linear time series models'. *Adv. Appl. Probab.*, **8**, 339-364.
- Dym, H. (1988), *J-contractive Matrix Functions, Reproducing Kernel Hilbert Spaces and Interpolation*. Wietmann Institute.
- Dym, H. and I. Gohberg (1983), 'Unitary interpolants, factorization indices and infinite Hankel block matrices'. *J. Func. Anal.*, **54**, 229-289.
- Dym, H. and H.P. McKean (1976), *Gaussian Processes, Function Theory and the Inverse Spectral Problem*. Academic Press.
- Eykhooff, P. (1974), *System Identification, (Parameter and State Estimation*. Wiley & Sons, London.
- Faibusovich, E. (1987), 'Matrix Riccati inequality, (Existence of solutions'. *Systems & Control Letters*, **9**, 59-64.
- Faurre, P., M. Clerget and F. Germain (1979), *Opérateurs Rationnels Positifs*. Dunod.
- Faurre, P.L. (1976), 'Stochastic realization algorithm'. In: R. Mehra and D.F. Lainiotis (eds.), *System Identification, Advances and Case Studies*. Academic Press.
- Ferrante, A. (1994a), 'A parametrization of minimal stochastic realizations'.

- Ferrante, A. (1994b), 'A parametrization of the minimal square spectral factors of a nonrational spectral density'. *Journal of Math. Systems, Estim. & Control*.
- Ferrante, A., G. Michaletzky and M. Pavon (1993), 'Parametrization of all minimal square spectral factors'. *Systems & Control Letters*, **21**, 249–254.
- Finesso, L. and G. Picci (1981), 'A characterization of minimal spectral factors'. *IEEE Trans. Autom. Control*, **AC-27**, 122–127.
- Forney, D. (1975), 'Minimal Bases of Rational Vector Spaces with Applications to Multivariable Linear Systems'. *SIAM Journal of Control*, **13**, 493–520.
- Fuhrmann, P. (1981), *Linear Operators and Systems in Hilbert Space*. McGraw-Hill.
- Fuhrmann, P. (1995), 'On the characterization and parametrization of minimal spectral factors'. *J. of Mathematical Systems, Estimation and Control*, **5**, 383–444.
- Fuhrmann, P. and A. Gombani (1998), 'On a Hardy-space approach to the analysis of spectral factors'. *Automatica*.
- Fuhrmann, P. and J. Hoffmann (1998), 'Factorization theory for stable discrete-time inner functions'. *J. of Math. Systems, Estim. and Control*.
- Gantmacher, F.R. (1988), *Theory of matrices*. Hauka. 4th ed. (in Russian).
- Gevers, M. (1986), 'ARMA Models, Their Kronecker Indices and Their McMillan Degree'. *Int. Journal of Control*, **43**, 1745–1761.
- Gevers, M. and V. Wertz (1984), 'Uniquely Identifiable State Space and ARMA Parametrization for Multivariable Linear Systems'. *Automatica*, **20**, 333–347.
- Gevers, M. and V. Wertz (1987), 'Techniques for the selection of identifiable parametrizations for multivariable linear systems'. In: C. Leondes (ed.), *Control and Dynamic Systems Vol. 26*. Academic Press, New York, 35–86.
- Glover, K. (1984), 'All optimal Hankel-norm approximations of linear multivariate systems and their L^∞ -norm bound'. *Internat. J. Control*, **39**, 1115–1193.
- Green, M. (1988), 'Balanced stochastic realizations'. *Linear Algebra and Appl.*, **98**, 211–247.
- Green, M. and B.D.O. Anderson (1986), 'The approximation of power spectra by phase matching'. In: *Proc. of 25-th Conf. on Decision and Control*. Greece.
- Green, M. and B.D.O. Anderson (1987a), 'Factorization of all-pass matrix functions'. *Int. J. of Control*, **45**, 215–235.
- Green, M. and B.D.O. Anderson (1987b), 'State-space formulae for the factorization of all-pass matrix functions'. *Int. J. of Control*, **45**, 1575–1602.
- Guidorzi, P. (1981), 'Invariants and Canonical Forms for Systems, (Structural and Parameter Identification'. *Automatica*, **17**, 117–133.
- Hannan, E.J. (1976), 'The identification and parametrization of ARMAX and state space forms'. *Econometrica*, **44**, 713–723.
- Hannan, E.J. (1976), 'The convergence of some time-series recursions'. *Ann. Stat.*, **4**, 1258–1270.
- Hannan, E.J. and M. Deistler (1988), *The Statistical Theory of Linear Systems*. John Wiley & Sons, New York.
- Hannan, E.J. and L. Kavalieris (1984), 'Multivariate linear time series models'. *Adv. Applied Probability*, **16**, 492–561.
- Hannan, E.J. and D.S. Poskitt (1988), 'Unit canonical correlations between future and past'. *Annals of Statistics*, **16**, 784–790.
- Hautus, M. L.J. and L.M. Silverman (1974), 'System structure and and singular control'. *Linear Algebra and Applications*, **50**, 3–8.
- Hazewinkel, M. and R.E. Kalman (1976), 'On invariants, canonical forms and moduli for linear constant finite dimensional dynamical systems'. In: G. Marchesini and S. Mitter (eds.), *Lecture Notes in Econ. – Mathem. System Theory, Vol. 131*. Springer Verlag, Berlin, 48–80.

- Ho, L. and R.E. Kalman (1965), 'Effective construction of linear state variable models from input/output data'. In: *Proceedings of the 3rd Allerton Conference*. 449–459.
- Isidori, A. and C. Moog (1986), 'On the nonlinear equivalent of the notion transmission zeros'. In: C.I. Byrnes and A. Kurzhansky (eds.), *Modelling and Adaptive Control*. Springer-Verlag.
- Janssen, P. (1988a), 'General Results on the McMillan Degree and Kronecker Indices of ARMA and MFD Models'. *Int. Journal of Control*, **46**.
- Janssen, P. H.M. (1987), 'MFD models and time delays; some consequences for identification'. *Int. J. Control*, **45**, 1179–1196.
- Janssen, P. H.M. (1988b), 'General results on the McMillan degree and the Kronecker indices of ARMA and MFD models'. *Int. J. Control*, **48**, 591–608.
- Janssen, P. H.M. (1988c), 'On Model Parametrization and Model Structure Selection for Identification of MIMO-Systems'.
- Kailath, T. (1980), *Linear Systems*. Prentice-Hall, Englewood Cliffs, NY.
- Kučera, V. (1991), *Analysis and design of discrete linear control systems*. Prentice Hall, New York, London, Toronto, Sydney.
- Lindquist, A. and G. Michaletzky (1997), 'Output-induced subspaces, invariant directions and interpolation in linear discrete time stochastic systems'. *SIAM J. Control and Optimization*, **35/3**, 810–859.
- Lindquist, A., G. Michaletzky and G. Picci (1995), 'Zeros of spectral factors, the geometry of splitting subspaces and the algebraic Riccati inequality'. *SIAM J. Control Optim.*, **33**, 365–401.
- Lindquist, A. and M. Pavon (1984), 'On the structure of state-space models for discrete-time stochastic vector processes'. *IEEE Trans. Automatic Control*, **AC-29**, 418–432.
- Lindquist, A. and G. Picci (1977), 'On the structure of minimal splitting subspaces in Stochastic realization Theory'. In: *Proc. of the 1977 Conference Decision and Control*. 42–48.
- Lindquist, A. and G. Picci (1979), 'On the stochastic realization problem'. *SIAM J. Control Optim.*, **17**, 365–389.
- Lindquist, A. and G. Picci (1985a), 'Forward and backward semimartingale representations for stationary increment processes'. *Stochastics*, **15**, 1–50.
- Lindquist, A. and G. Picci (1985b), 'Realization theory for multivariate stationary Gaussian processes'. *SIAM J. Control Optim.*, **23**, 809–857.
- Lindquist, A. and G. Picci (1990), 'A geometric approach to modelling and estimation of linear stochastic systems'. *Journal of Mathematical Systems, Estimation and Control*, **1**, 241–333.
- Ljung, L. (1987), *System Identification, Theory for the User*. Prentice-Hall Inc., Englewood Cliffs, N.J.
- Ljung, L. and P. Caines (1979), 'Asymptotic normality of prediction error estimation for approximate system models'. *Stochastics*, **31**, 29–46.
- Ljung, L. and T. Kailath (1976), 'Backwards Markovian models for second order stochastic processes'. *IEEE Transactions on Automatic Control*, **IT-22**, 488–491.
- Ljung, L. and T. Söderström (1983), *Theory and practice of recursive identification*. The MIT Press.
- Luenberger, D.G. (1967), 'Canonical forms for linear multivariable systems'. *IEEE Trans. Autom. Control*, **AC-12**, 290–293.
- Mårtensson, K. (1971), 'On the matrix Riccati equation'. *Infom. Sci.*, **3**, 17–49.
- Michaletsky, G. (1992), 'Zeros of (non-square) spectral factors and canonical correlations'. In: *Proc. 11th IFAC World Congress*. 167–172. Tallinn, Estonia, Also in Preprint collection (1990), 221–226.
- Michaletsky, G. and A. Ferrante (1995), 'Splitting subspaces and acausal spectral factors'. *Journal of Math. Systems, Estimation and Control*, **5**, 363–366. Summary, long paper retrieval code, (89459).

- Michaletsky, G. and G. Tusnády (1991), 'Representation of inner products and stochastic realization'. In: L. Gerencsér and P.E. Caines (eds.), *Topics in Stochastic Systems, (Modelling, Estimation and Adaptive Control*. Springer-Verlag.
- Molinari, P. (1977), 'The time-invariant linear-quadratic optimal-control problem'. *Automatica*, **13**, 347–357.
- Ober, R. (1987), 'Balanced realizations, (canonical form, parametrization, model reduction'. *Internat. J. Control*, **46**, 643–670.
- Ober, R. (1996), 'Balanced canonical forms'. In: S. Bittanti and G. Picci (eds.), *Identification, Adaptation, Learning*. Springer Verlag, 120–183.
- Ober, R. and B. Hanzon (1997), 'Overlapping block-balanced canonical forms and parametrizations, (the stable SISO case'. *SIAM J. Control and Optim.*, **35**, 228–242.
- Pavon, M. (1980), 'Stochastic realization and invariant directions of the matrix Riccati equation'. *SIAM J. Control Optim.*, **18**, 155–180.
- Pavon, M. (1982), 'A new algorithm for optimal interpolation of discrete-time stationary processes'. In: A. Bensoussan and J.L. Lions (eds.), *Analysis and Optimization of Systems*. Springer-Verlag, New York, 701–718.
- Pavon, M. (1984), 'New results on the interpolation problem for continuous-time stationary increments processes'. *SIAM J. Control Optim.*, **22**, 133–142.
- Picci, G. (1976), 'Stochastic realization of Gaussian processes'. In: *Proceedings IEEE* 64, 112–122.
- Picci, G. and S. Pinzoni (1991), 'Acausal models of stationary processes'. In: *Proc. of the First European Control Conference*.
- Picci, G. and S. Pinzoni (1994), 'Acausal models and balanced realizations of stationary processes'. *Linear Algebra and Applications*, **205–206**, 997–1043.
- Popov, M. (1972), 'Invariant Description of Linear Time Invariant Controllable Systems'. *SIAM Journal of Control*, **10**, 252–264.
- Potter, E. (1966), 'Matrix quadratic solutions'. *SIAM J. Appl. Math.*, **14**, 496–501.
- Rozanov, N.I. (1967), *Stationary Random Processes*. Holden Day.
- Rozanov, N.I. (1990), *Stationary Random Processes*. Nauka. in Russian, 2nd ed.
- Ruckebuch, G. (1975), 'Représentations Markoviennes de processus Gaussiens Stationnaires'.
- Sand, J.Å. (1995), 'Zeros of discrete-time spectral factors and the internal part of a Markovian splitting subspace'. *SIAM J. Control Optim.*
- Shayman, M. (1983), 'Geometry of the algebraic Riccati equation'. *SIAM J. Control and Optim.*, **21**, 375–394.
- Sz.-Nagy, B. and C. Foias (1970), *Harmonic Analysis of Operators on Hilbert Space*. North-Holland.
- Terdik, G. (1985), 'Transfer functions and conditions for stationarity of bilinear models with Gaussian residuals'. *Proc. Roy. Statist. Soc. Ser. A*, **400**, 315–330.
- Terdik, G. (1991), 'Bilinear state space realization for polynomial systems'. *Computers Math. Applic.*, **22**, 69–83.
- Terdik, G., P. Várlaki and J. Bokor (1992), 'On problem of stochastic bilinear realization'. *Computational systems analysis*. Elsevier Science Publisher, 67–72.
- van den Hof, P. M.J. (1989), 'A criterion based equivalence for equation error models'. *IEEE Trans. Autom. Control*, **AC-34**.
- Verghese, G. and T. Kailath (1979), 'A further note on backwards Markovian models'. *IEEE Transactions on Information Theory*, **IT-25**, 121–124. Correction to the above paper in *ibid IT-25*, p. 50.
- Wiener, N. (1958a), 'The prediction theory of multivariate stochastic processes, Part I, The regularity condition'. *Acta Mathematica*, **98**, 11–150.
- Wiener, N. (1958b), 'The prediction theory of multivariate stochastic processes, Part II, The linear predictor'. *Acta Mathematica*, **99**, 93–137.

- Willems, J.C. (1971), 'Least squares stationary optimal control and the algebraic Riccati equation'. *IEEE Trans. Autom. Control*, **AC-16**, 621-634.
- Willems, J.C. (1986a), 'From time series to linear systems, Part I'. *Automatica*, **22**, 561-580.
- Willems, J.C. (1986b), 'From time series to linear systems, Part II'. *Automatica*, **22**, 675-694.
- Willems, J.C. (1987), 'From time series to linear systems, Part III'. *Automatica*, **23**, 87-115.
- Willems, J.C. (1991), 'Dynamical systems, Controllability and Observability, (A Post-Modern Point of View'. In: C.A. Athanasios (ed.), *Mathematical System Theory. The Influence of R. E. Kalman*. Springer Verlag, Berlin, Heidelberg, New York, 19-40.
- Wold, A. (1938), *A Study in the Analysis of Stationary Time Series*. Uppsala, (Almqvist and Wiksell.
- Wolovich, W.A. (1974), *Linear Multivariable Systems*, Applied Mathematical Sciences, No.11. Springer Verlag, New York.
- Wolovich, W.A. and H. Elliott (1983), 'Discrete models for linear multivariable systems'. *International Journal of Control*, **38**, 337-357.
- Wonham, W.M. (1977), *Linear Multivariable Control, a Geometric Approach*. Springer-Verlag.
- Youla, D.C. (1961), 'On the factorization of rational matrices'. *IEEE Transactions on Information Theory*, **IT-7**, 172-189.

6350. /

The purpose of this book is to investigate modeling and representation approaches of stationary stochastic phenomena. These approaches have their origins in the theory of multivariate stochastic processes, time series analysis and in the algebraic geometric theory of stochastic systems. The stochastic representations most frequently used are the auto-regressive-moving-average (ARMA), matrix-fraction-descriptions (MFD) and the state-space representations. It is shown how to derive these – forward and backward – representations and their dual forms from the analytic and co-analytic spectral factors. These representations are parametrized by system invariants reflecting the four basic Kalmanian principles of controllability, observability, reachability and reconstructibility.

Detailed structure of the state-space realizations is provided using geometric (Hilbert-space) principles including the analysis of the zero structure and balanced realizations.

The structure of generalized Wiener-Hopf factorization is studied in details, first using geometric consideration, then computing the system matrices.

The perspective provided by this can be interesting for those who are doing research in signal processing, stochastic modeling and system identification, or in control system design. Parts of the text can also be useful in courses on stochastic systems, filtering, prediction or on realization theory.

I SEN 963-05-7546-9



9 799630 575460



VS99.8VO

Gy. Michaletzky, J. Bokor, P. Várlaki • Representability of Stochastic Systems