UNIVERSITÉ DE NICE SOPHIA ANTIPOLIS

Mémoire

présenté pour obtenir le diplôme d'

Habilitation à Diriger des Recherches

en Sciences Fondamentales et Appliquées

Spécialité: Mathématiques

par

Martine OLIVI

Parametrization of rational lossless matrices with applications to linear systems theory

soutenue le 25 Octobre 2010 devant le jury composé de

Laurent BARATCHART	INRIA Sophia-Antipolis	Directeur
Bernhard BECKERMANN	Université de Lille	Rapporteur
Hugue GARNIER	Université de Nancy (CRAN)	Examinateur
Tryphon GEORGIOU	University of Minnesota	Rapporteur
Jan MACIEJOWSKI	University of Cambridge	Examinateur
Paul VAN DOOREN	Université Catholique de Louvain	Rapporteur

Contents

Introduction

The works presented in this habilitation thesis were initially motivated by the rational approximation issue in system theory.
Rational matrix-valued functions arise with the transfer function description of a finite dimensional, linear, time-invariant system. Rational approximation seems to me the main issue in the fundamental modelization problem: given some data about a system, find the best model in the popular class of LTI systems. This problem can be understood in many different ways and resulted in many approaches depending on the nature of the data (time series, frequency data), on the purpose of the model (prediction, control), etc. . The current approaches mainly fall in two sub-classes, the projection methods (subspace method, balanced truncation) easy to implement but not always efficient, and optimization methods (Hankel, H_2) which result in difficult non-linear problems. Model reduction and identification are still very active fields in system theory as numerous recent publications show [3, 24, 58].

Our contribution was dedicated to rational approximation in the Hardy space H_2. This space is interesting both for its underlying stochastic interpretation and the Hilbertian framework it provides. It corresponds for the transfer function to a $L^2 \mapsto L^\infty$ stability requirement (a finite energy input produces a bounded energy output). Given a transfer function in H_2, the rational approximation problem is thus to minimize the distance in L^2-norm to the set of rational stable transfer functions of order at most n. This problem features some serious difficulties : it is a non-linear optimization problem, the set of approximants has a complicated structure and the L^2-error may possess a great number of local minima in which a descent algorithm can get stuck. To cope with these problems, we have developed an original approach [BCO91, FO98] based on

- a compactification of the optimization set which makes use of the Douglas-Shapiro-Shields factorization and the projection theorem in Hilbert space. Lossless matrix valued functions play in the matrix case the role of the denominator in the scalar case. The "denominator" is thus optimized while the "numerator" is computed by projection: the optimization runs over the set of lossless functions which thus enter the picture and bring their rich structure. Surprisingly enough, this method is seldom used in approximation; it is known, in the context of least-square approximation, under the name "separable least-squares".

- the use of an atlas of charts to describe the optimization set: the space of lossless functions of fixed McMillan degree. This type of representation is recommended if one wants to use differential calculus tools for solving an optimization prob-

lem. The first atlas of charts for lossless matrix valued functions were obtained in [1] from a tangential Schur algorithm. An important part of this document is devoted to the description of such atlases.

Rational approximation was our first motivation to study lossless functions and their parametrizations. However, these functions play an important role in system theory anyway because of the Douglas-Shapiro-Shields factorization. They also are of independent interest, being the transfer functions of conservative systems: the scattering matrix of a frequency filter, the polyphase matrix of an orthogonal filter bank are lossless. Overall, they lie at the heart of Schur analysis, a rich theory studying interpolation problems for functions satisfying a metric constraint, namely Schur functions (see e.g. [18, 4, 20]). This gives rise to multiple interactions with system theory [35, 40] and has widely inspired this work.

Concerning the parametrization of lossless (matrix-valued) functions, we aimed at connecting two types of representations

- atlases of charts derived from interpolation problems. The parameters are thus interpolation values. This is a very efficient way to take into account the Schur metric constraint and subsequently the stability requirement.

- state-space representations and balanced canonical forms. State-space representations involve basically linear algebra techniques and their efficiency for numerical computations has been demonstrated.

In [31], such a connection was established between the Hessenberg canonical form for discrete-time scalar systems and the Schur algorithm. The Hessenberg canonical form is balanced and such that, in addition, the associated controlability matrix is upper-triangular. The state-space realization can be written as a product of Givens rotations whose parameters are precisely the sequence Schur parameters. In chapter 2, the results of [31] are generalized to the matrix case following the approach developed in [HOP06, PHO07, HOP09]. Overlapping balanced canonical forms are associated with a suitable tangential Schur algorithm. Realization matrices are obtained as a product of unitary matrices that can be parametrized by a sequence of vectorial Schur parameters. Moreover, for a canonical choice of interpolation values and interpolation directions in the Schur algorithm, the realization matrix possesses a sub-diagonal pivot structure. It generalizes the Hessenberg form of the SISO (single input-single output) case and should be a useful tool for model reduction.

Chapter 3 is dedicated to H_2 rational approximation and its practical implementation in the software RARL2. Another atlas of charts is used here, which has been constructed from more general interpolation data [MO07]. In this atlas, a chart is indexed by a lossless transfer function whose parameters are all zero, which explains the denomination "lossless mutual encoding" for this parametrization. It exhibits several nice features, in particular computational simplicity and the fact that it better suits the representation of real or conjugate symmetric lossless functions. This is an important point, since real-world systems and their transfer functions are usually real. The software RARL2 [MO04] is mainly used for two dedicated applications, for which it is combined with other tools into specific softwares: PRESTO-HF for the identification of microwave filters from frequency data [60] and FINDSOURCE3D for the localiza-

tion of dipolar sources in electroencephalography [8]. Potential applications abound, among which wavelet approximation and aircraft flutter modal parameter identification are already under investigation.

More recently, motivated by microwave filter design applications, we worked on a completion problem for matrix valued lossless systems. Frequency filters are usually described by scattering matrices which are not only lossless but also symmetric, due to the reciprocity law attached to wave propagation. Since the design problem bears on a submatrix (the transmission), it would be interesting to characterize the submatrices of a lossless symmmetric matrix. This problem can be stated as follows: under which conditions can a symmetric Schur matrix be extended into a lossless one? We recast the celebrated Darlington synthesis problem [15] with an additional symmetry requirement which refines the completion problem. In Chapter 4 we summarize the contribution of [BEGO07] which significantly improves previous results [2] and provides both a minimal degree extension of double dimension and a minimal dimension extension of the same degree. These results brings a new understanding of the structure of 3×3 reciprocal matrices and the derivation of efficient algorithm for the synthesis of 3-ports devices is being studied.

Chapter 1

Lossless functions in system theory

Lossless systems and their transfer functions are involved in many areas of system theory, from electrical network theory and related synthesis problems [2] to modern digital signal processing [63]. A lossless system is a device which conserves energy, as electrical networks with no resistive elements. The usefulness of lossless transfer functions is also and mainly due to the Douglas-Shapiro-Shields factorization. They indeed appear as generalized denominators of rational matrix-valued functions.

In an engineering context, linear systems have been studied since the 1930s. An input/output and frequency domain viewpoint prevailed and SISO systems were modeled by their transfer functions. In the sixties, the move to state space models open the way to new mathematical fields. The input/state/output systems had much more modeling power, were far richer mathematically and took into consideration initial conditions, something that transfer functions fail to do. Mathematical system theory was rapidly growing. The algebraic setting and the concept of module turned out to be essential to deal with multivariable or MIMO (multi-input, multi-output) systems [37, 36]. The works of Fuhrmann, that we mainly follow, aimed at synthesizing the algebraic approach of Kalman, the state-space approach as well as the polynomial approach of Rosenbrock [56]. Functional analysis methods and geometry, relying on a stability assumption, happened to be relevant in many purposes and nowadays classical [23]. From a mathematical viewpoint, lossless functions arise with this geometric framework, namely associated with the invariant subspaces of the Hardy space H_2. The reader interested by the history of mathematical systems theory should have a look at Willems's essay [64].

In this chapter, we go over the now classical representations of finite-dimensional LTI systems from transfer functions to state-space descriptions [23], and then glide to interpolation theory for rational systems as developed in [4]. The state and the pole-zero structure of the transfer function are the key concepts and lossless functions the heroes of this chapter. We focus on discrete-time systems that is, from a complex analysis viewpoint, to the framework of the unit disk which is in some sense simpler. Our studies mostly bear on discrete-time lossless matrices, except in Chapter 4. The discrete-time and continuous-time settings are connected through a bilinear (or Möbius) transformation of the variable (see e.g. [27] and section 3.4).

1.1 Transfer functions and their factorizations

From an input/output viewpoint, a discrete-time system is modeled by a map :

$$\sigma : (u_k)_{k\in\mathbb{Z}} \mapsto (y_k)_{k\in\mathbb{Z}}, \quad u_k \in \mathbb{F}^m, \quad y_k \in \mathbb{F}^p, \quad \mathbb{F} = \mathbb{R} \text{ or } \mathbb{C}.$$

Linearity together with time-invariance of the system induces a discrete convolution operator. Putting more structure on the input and output signals spaces, allows to rephrase systems properties (causality, stability) in mathematical terms. The two approaches described in this section follow two parallel lines, splitting the time axis in two parts, modeling the state and translating shift-invariance into commutation properties.

1.1.1 The algebraic framework

Assuming the system initially at rest, algebra enters the picture representing input and output by truncated Laurent series and the system as a map

$$\begin{array}{rccc} \sigma : & \mathbb{F}((1/z))^m & \longrightarrow & \mathbb{F}((1/z))^p \\ & u(z) = \sum u_k z^{-k} & \mapsto & y(z) = \sum y_k z^{-k} \end{array} .$$

The set of truncated Laurent series $\mathbb{F}((1/z))$ is the quotient field of the set of formal series $\mathbb{F}[[1/z]]$. Linearity and time-invariance of the system imply that σ is a linear transformation of vector spaces. The input/output behavior of the system is then completely described by the multiplication by a matrix-valued function called the *transfer function* of the system

$$y(z) = T(z)u(z).$$

Causality of the system implies that $T(z)$ is proper, $T(z) \in \mathbb{F}[[1/z]]^{p\times m}$.

With the concept of state, the module structure appears as the basic structure in this context. The state is the memory of the system: at $t = k$, the state x_k is the information that together with u_k determines uniquely the output $y(t)$ for all $t > k$. To address this concept, the time axis is divided in two parts, writing $\mathbb{F}((1/z))$ as a direct sum

$$\mathbb{F}((1/z) = \mathbb{F}[z] \oplus 1/z\mathbb{F}[1/z],$$

in which $\mathbb{F}[z]$ denotes the ring of polynomials over $\mathbb{F}$. This decomposition is obtained from the exact sequence of $\mathbb{F}[z]$-modules:

$$0 \to \mathbb{F}[z] \xrightarrow{injection} \mathbb{F}((1/z)) \xrightarrow{projection} \underbrace{\mathbb{F}((1/z))/\mathbb{F}[z]}_{1/z\mathbb{F}[[1/z]]} \to 0$$

The past-input to future-output map $\bar{\sigma}$ is then defined by the following commutative diagram

$$\begin{array}{ccc} \mathbb{F}((1/z))^m & \xrightarrow{\sigma} & \mathbb{F}((1/z))^p \\ \pi^+ \downarrow & & \downarrow \pi^- \\ \mathbb{F}[z]^m & \xrightarrow{\bar{\sigma}} & 1/z\mathbb{F}[[1/z]]^p \end{array}$$

The kernel of $\bar{\sigma}$, $\ker \bar{\sigma}$, is a sub-module of the free module $\mathbb{F}[z]^m$ over the PID (principal ideal domain) $\mathbb{F}[z]$.

$$\ker \bar{\sigma} = \{u(z) \in \mathbb{F}[z]^m; T(z)u(z) \in \mathbb{F}[z]^p\}\,. \tag{1.1}$$

The structure of such sub-modules is well-known and described by the following theorem ([34, II, Ch. 3], [23]).

Theorem 1.1.1 (Adapted basis) *Every sub-module $\mathcal{E}$ of a finitely generated free module $\mathcal{M}$ over a PID $\mathcal{R}$, is free. There exists a basis $(e_1, e_2, \ldots, e_m)$ of $\mathcal{M}$ and a set of elements $d_1, d_2, ...d_s$, $s \leq m$ of $\mathcal{R}$ such that d_i/d_{i+1} and $(d_1e_1, d_2e_2, \ldots, d_se_s)$ is a basis of $\mathcal{E}$.*

In our case, any basis of $\mathbb{F}[z]^m$ consists in the columns of a *unimodular matrix*, a matrix which is invertible in $\mathbb{F}(z)^{m\times m}$ (its determinant is a constant). Theorem 1.1.1 asserts that $\ker \bar{\sigma}$ is spanned by the columns of a $m \times m$ polynomial matrix $D(z)$, $\ker \bar{\sigma} = D(z)\mathbb{F}[z]^m$ and $D(z)$ is of the form

$$D(z) = U(z)\mathrm{diag}(d_1(z), d_2(z), \ldots, d_s(z), 0, \ldots, 0)V(z) \tag{1.2}$$

- $U(z)$ is a unimodular matrix,
- $d_1(z), d_2(z), \ldots, d_s(z)$, are uniquely defined polynomials, with unit leading coefficients, satisfying the divisibility conditions d_i/d_{i+1} for $i = 1, \ldots, s-1$, called the *invariant factors*,
- $V(z)$ is a unimodular matrix that can be chosen arbitrarily.

Formula (1.2) is known as the Smith form of a polynomial matrix. The product $d_1...d_k$ of the invariant factors is the $g.c.d.$ of the $k \times k$ minors of $D(z)$ [34, II, Chap. 3, Sec. 9]. The Smith form also exists for a rectangular matrix.

Clearly, the state-space concept can be modeled by the quotient module

$$\mathcal{X} = \mathbb{F}[z]^m / \ker \bar{\sigma},$$

that is the set of equivalence classes for the relation induced by $\ker \bar{\sigma}$ in $\mathbb{F}[z]^m$: $f \sim g$ if $\bar{\sigma}f = \bar{\sigma}g$. We get the following commutative diagram, where π is a projection and ι an injection. The state-space can be viewed either as the quotient module $\mathcal{X}$ or as the

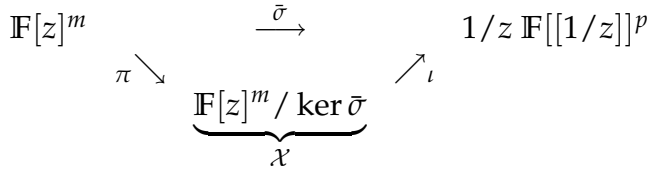

Figure 1.1: The state space.

range of the injection ι and thus as a sub-module of $1/z\,\mathbb{F}[[1/z]]^p$ or else as a vector space over $\mathbb{F}$. This triple nature is at the heart of realization theory (see section 1.2).

A LTI system is called finite order if $\mathcal{X}$ is a finite dimensional vector space over $\mathbb{F}$. Finite-dimensional LTI systems are characterized by rational transfer functions which is a consequence of

Theorem 1.1.2 *The following assertions are equivalent:*

- $\ker \bar{\sigma}$ *is a full sub-module (m generators)*
- $D(z)$ *is non-singular*
- $\mathcal{X}$ *is a torsion module. It is a direct sum of cyclic sub-modules:* $\mathcal{X} = \bigoplus \frac{\mathbb{F}[z]}{d_i \mathbb{F}[z]}$, *the* d_i*'s being the invariant factors of* $D(z)$.

In this case, $\mathcal{X}$ *is a finite dimensional vector space over* $\mathbb{F}$. *The order of the system is defined as* $\dim_{\mathbb{F}} \mathcal{X} = \sum_i \deg d_i = \deg(\det D(z))$.

From now on, we restrict ourselves to the study of finite dimensional LTI systems. A number of important representation results follow from the structure of $\ker \bar{\sigma}$, as the matrix fraction description [16] or the Smith-McMillan form [50].

The columns of $D(z)$ belong to $\ker \bar{\sigma}$ and thus $T(z)D(z)$ must be a polynomial matrix $N(z)$; $D(z)$ being invertible, we get the matrix fraction description (MFD)

$$T(z) = N(z)D(z)^{-1} \tag{1.3}$$

This factorization is called a *right* MFD, it has been obtained working with the columns of the matrix, which is quite natural in this input/output context. However, working with the rows yields a dual representation, the *left MFD*, $T(z) = \tilde{D}(z)^{-1}\tilde{N}(z)$.

The Smith-McMillan form is a simple extension of the Smith form (1.2) in the rectangular case. It is obtained from $T(z) = P(z)/q(z)$, where $q(z)$ is a common denominator of the rational entries of $T(z)$ and $P(z)$ a polynomial matrix, writing $P(z)$ in Smith form.

Theorem 1.1.3 (Smith-McMillan form) *Any rational function matrix* $T(z)$ *can be written in the form*

$$T(z) = U(z)\text{diag}\,\{n_1(z)/d_1(z), ..., n_r(z)/d_r(z), 0, ..., 0\}V(z) \tag{1.4}$$

with

- $U(z)$ *and* $V(z)$ *unimodular matrices*
- n_i *and* d_i *are coprime polynomials, with unit leading coefficients, satisfying the divisibility properties*

$$n_1|n_2|...|n_r \quad \text{and} \quad d_r|d_{r-1}|...|d_1.$$

The matrix $\Lambda(z) = \{n_1(z)/d_1(z), ..., n_r(z)/d_r(z), 0, ...\}$ *is called the Smith-McMillan form of* $T(z)$ *and is uniquely defined: it is* canonic.

The non-negative integer r is the *normal rank* of $T(z)$: there exists at least one minor of order r which does not vanish identically, and all minors of order greater than r vanish identically.

From a naive point of view, a pole of a rational matrix function is a pole of any of its entry, but this definition neglects the problem of assigning multiplicities. The Smith McMillan form is the simplest way to define poles and zeros for a rational matrix function:

- the *pole polynomial* is defined as $d(z) = d_1(z)d_2(z)\dots d_r(z)$ and its roots are the poles of the transfer matrix $T(z)$.
- the *zero polynomial* is defined as $n(z) = n_1(z)n_2(z)\dots n_r(z)$ and its roots are the zeros or *transmission zeros* of the transfer matrix $T(z)$.

A zero is a point where the local rank of $N(z)$ in (1.3) drops below the normal rank [16]. It is important to notice that $n(z)$ and $d(z)$ are not necessarily relatively prime. A rational matrix can have a pole and a zero at the same location $z = \omega$ which do not cancel. The McMillan degree of a rational matrix is defined to be $\sup\{\deg n, \deg d\}$. In the case of a proper (analytic at ∞) transfer function, the McMillan degree is the total number of poles (with multiplicity). It is given by $\deg[d(z)]$ and coincides with the order of the system.

The location and multiplicity of poles and zeros are essential to fully understand the characteristics of a multivariable system. If we write, for some $\omega \in \mathbb{C}$,

$$\frac{n_i}{d_i}(z) = (z-\omega)^{\sigma_i(\omega)} l_i(z), \quad \sigma_i(\omega) \in \mathbb{Z}, \tag{1.5}$$

where $l_i(z)$ has neither poles nor zeros at ω, we get from the divisibility properties

$$\sigma_1(\omega) \leq \sigma_2(\omega) \leq \cdots \leq \sigma_r(\omega).$$

These numbers are called the *structural indices* or *partial multiplicities* of $T(z)$ at ω, and ω is a pole if and only if some index is negative and a zero if and only if some index is positive.

Note that, the relevant algebraic object in the study of the pole structure of $T(z)$ is the torsion module $\mathcal{X}$, that is the state-space of the system. For a module theoretic approach to pole-zero theory we refer the reader to [65, 66]. In a broad sense, system zeros have been widely studied in the literature and many different definitions can be found, corresponding to intuitive physical interpretations: system zeros, invariant zeros, blocking zeros, decoupling zeros (see [49] and the recent book [62] for an overview on this topic).

1.1.2 The geometric framework

In this section, a system is modeled by an operator in Hilbert spaces. This framework is suitable for an extension to infinite dimensional systems [23]. However, it brings some insights even in the finite dimensional case. Adding structure on signal and system spaces, we get a sharpened description. There is a great similarity in the development of the geometric and the algebraic approach.

We now assume the input and output signals to have finite energy, that is to belong to the Hilbert spaces of square summable series l^2. Using Parseval theorem, the input and the output can be represented by vectors with elements in $L^2(\mathbb{T})$, the space of functions defined on the unit circle $\mathbb{T}$ and square integrable. We get similar expressions for the input and the output signals, but in which the variable z now belongs to the unit circle. Note that the assumption is in fact about the stability of the system which, in this

case, produces a finite energy output for a finite energy input. This notion of stability is called $L^2 - L^2$ stability and it implies that the transfer function T is a bounded operator

$$\|T\|_\infty = \sup_\theta ||T(e^{i\theta})||, \tag{1.6}$$

where $||\,.\,||$ denotes the operator norm $\mathbb{C}^m \to \mathbb{C}^p$.

The relevant spaces of functions are thus the Hardy spaces of the disk, H_2 and H_∞. This plead for the discrete-time setting, since the Hardy spaces of the disk are simpler than their half-plane analogs. They are spaces of analytic functions in $\mathbb{D}$ satisfying some metric constraint, but they can be viewed alternatively as subspaces of the space $L^2(\mathbb{T})$. The Hardy space H_2 (resp. H_∞) consists of functions in $L^2(\mathbb{T})$ (resp $L^\infty(\mathbb{T})$) whose Fourier coefficients (a_n) satisfy $a_n = 0$ for $n < 0$. Symmetrically, the conjugate Hardy space $\bar{H}_2$ (resp. $\bar{H}_\infty$) consists of functions in $L^2(\mathbb{T})$ (resp $L^\infty(\mathbb{T})$) whose Fourier coefficients satisfy $a_n = 0$ for $n > 0$. The transfer function of a $L^2 - L^2$ stable and causal systems thus belong to $\bar{H}_\infty^{p\times m}$. The Hardy space H_2 is an Hilbert space and as such, it provides a very interesting setting for approximation problems, as we shall see in Chapter 3.

In this context, dividing the time-axis in two parts gives rise, in terms of signals, to the orthogonal decomposition

$$L^2(\mathbb{T}) = H_2 \oplus H_2^\perp,$$

where $H_2^\perp$ (the subspace of $\bar{H}_2$ of functions vanishing at ∞) is the orthogonal complement of H_2 in $L^2(\mathbb{T})$. The past-input to future-output map H, or *Hankel operator*, can then be defined by the following commutative diagram

$$\begin{array}{ccc} L^2(\mathbb{T})^m & \xrightarrow{\sigma} & L^2(\mathbb{T})^p \\ \pi^+ \downarrow & & \downarrow \pi^- \\ H_2^m & \xrightarrow{H} & (H_2^p)^\perp \end{array}$$

and the kernel of H is now a subspace of H_2^m,

$$\ker H = \left\{ u(z) \in H_2^m ; T(z)u(z) \in H_2^p \right\}. \tag{1.7}$$

This subspace possesses the fundamental property to be shift-invariant, that is an invariant subspace of the shift operator $S : u(z) \to zu(z)$. The Beurling-Lax theorem [23, Th. 12.22] characterizes shift-invariant subspaces of H_2^m by means of inner matrix-valued functions. A matrix $Q(z)$ in $H_\infty^{m\times m}$ is called *inner* if it takes unitary values a.e. on the unit circle.

Theorem 1.1.4 (Beurling-Lax) *Each closed shift-invariant subspace $\mathcal{M}$ of full range of H_2^m is of the form*

$$\mathcal{M} = QH_2^m,$$

for some $m \times m$ inner matrix $Q(z)$. Moreover $Q(z)$ is unique up to right multiplication by some constant unitary matrix.

The theorem is stated here in the simple (finite dimension) vectorial case and for full ranges subspaces ($\{u(z),\ u \in \mathcal{M},\ z \in \mathbb{T}\}$ spans $\mathbb{C}^m$). It mainly relies on the Wold decomposition [23, Th. 8.2.]

$$\mathcal{M} = \oplus_n S^n \mathcal{L}, \quad \mathcal{L} = \mathcal{M} \ominus S\mathcal{M}.$$

A more general version can be found in [32]. Note that this result can be adapted to the real Hardy space, the subspace of H_2 consisting of functions with real Fourier coefficients, which is relevant for real systems [BO91].

As a consequence of Beurling-Lax theorem, $\ker H$ is of the form QH_2^m for some inner matrix $Q(z)$, unique, up to a right unitary constant multiplier. The state-space is isomorphic to the orthogonal complement $H(Q)$ of QH_2^m in H_2^m.

$$H_2^m = QH_2^m \oplus H(Q).$$

It is finite dimensional over $\mathbb{F}$ if and only if $Q(z)$ is rational.

Recall that an inner function is invertible a.e. and its inverse, given by

$$Q(z)^{-1} = Q\left(1/\bar{z}\right)^*. \tag{1.8}$$

This formula is obtained by analytic continuation from the identity $Q(z)Q(z)^* = I$ for $|z| = 1$. Throughout this document, we shall use the isometric transformation

$$G(z) \mapsto G^\sharp(z) = G\left(1/\bar{z}\right)^*. \tag{1.9}$$

so that $Q(z)^{-1} = Q^\sharp(z)$. The matrix $Q^\sharp(z)$ is analytic in the complement of the closed unit disk and takes unitary values on the circle. Such matrix functions are called *co-inner* or *stable allpass* or *lossless* since they are the transfer functions of lossless systems. We shall use this terminology for both discrete time and continuous time transfer functions: for us, a lossless function is analytic in the stability domain and takes unitary values on its boundary.

Another immediate consequence is the Douglas-Shapiro-Shields factorization [17] of the transfer function $T(z)$, which is obtained by observing that the columns of $Q(z)$ clearly belongs to $\ker H$, so that $T(z)Q(z) = C(z)$ is in $H_2^{p\times m}$.

Theorem 1.1.5 (Douglas-Shapiro-Shields) *Any $p \times m$ rational matrix function $T(z)$ analytic in the complement of the unit disk, can be represented as*

$$T(z) = C(z)Q^\sharp(z), \tag{1.10}$$

where $Q(z)$ is a $m \times m$ inner function, $C(z) \in H_2^{p\times m}$ and $Q^\sharp(z)$ as same degree as $T(z)$. The lossless matrix-valued function $Q^\sharp(z)$ is called the right lossless factor *of $T(z)$.*

This factorization must be compared with the matrix fraction description (1.3). We have less freedom in the choice of the inner matrix function $Q(z)$, which is unique up to a left *unitary* factor, than in the choice of a polynomial denominator $D(z)$, which is unique up to a left *unimodular* factor. The Douglas-Shapiro-Shields factorization of rational matrix achieves the closest analogy with irreducible fractions in the scalar case.

1.1.3 Inner matrix-valued functions

The algebra of inner matrix-valued functions is relatively simple and close to polynomial algebra except for commutativity.

The determinant of a rational inner function $Q(z)$ is inner that is a finite Blaschke product :

$$\det Q(z) = \frac{q(z)}{\tilde{q}(z)}, \quad \tilde{q}(z) = z^n \overline{q(1/\bar{z})}$$

where $q(z)$ is a Schur or stable polynomial (roots in $\mathbb{D}$) of degree n and $\tilde{q}(z)$ its reciprocal polynomial. The zeros of $Q(z)$ belong to the open unit disk and its poles to the complement of the closed unit disk, so that no pole-zero cancellation can occur when multiplying rational inner matrices. For any two inner matrices Q_1 and Q_2, we thus have the following relation on the McMillan degree :

$$\deg(Q_1 Q_2) = \deg Q_1 + \deg Q_2.$$

The following factorization will be used in several occasions. It results from the work of Potapov on the multiplicative structure of J-contractive matrix-valued functions ([55]). It is obtained by induction on the zeros (see [21]). Let w be a zero of the inner matrix function $Q(z)$ of McMillan degree n. Let u be some unit vector in the kernel of $Q(w)$. Then one can extract from Q a left inner factor of the form

$$B_{w,u}(z) = I + (b_w(z) - 1)\, uu^*, \quad b_w(z) = \frac{z - w}{1 - \bar{w}z}, \tag{1.11}$$

and $Q = B_{w,u} Q_1$ for some Q_1, still inner and of degree $n - 1$. Note that any inner factor of McMillan degree 1 is of the form (1.11), up to a right unitary matrix [18]. We call these inner factors *elementary inner factors* or *Potapov factors*.

Proposition 1.1.1 (Potapov factorization) *Any inner matrix-valued function $Q(z)$ can be written as the product of elementary inner factors of the form (1.11)*

$$Q(z) = B_{w_n,u_n}(z) B_{w_{n-1},u_{n-1}} \dots B_{w_1,u_1}(z) Q_0 \tag{1.12}$$

where $w_1, w_2, \dots, w_n$ belong to the open unit disk and $u_1, u_2, \dots, u_n$ are unit vectors.

Using the transformation (1.9), analog results can be stated for rational lossless matrices.

1.2 Realization theory and balanced canonical forms

The state-space description of a system has an important modeling potential and is nowadays the most used in the engineer community. A finite order linear time invariant system in discrete time is then described by a pair of dynamical equations

$$\begin{cases} x_{k+1} &= A\, x_k + B\, u_k, \\ y_k &= C\, x_k + D\, u_k, \end{cases} \tag{1.13}$$

with $k \in \mathbb{Z}$, $x_k \in \mathbb{F}^n$ for some non negative integer n (the state space dimension), $u_k \in \mathbb{F}^m$, the input space, and $y_k \in \mathbb{F}^p$, the output space. This state space representation of a system is explicit and is referred to as an internal description.

The transfer matrix associated with the linear system (1.13) is easily computed to be

$$T(z) = D + \sum_{k=0}^{\infty} CA^k B\, z^{-(k+1)} = D + C(zI - A)^{-1}B \tag{1.14}$$

Conversely, any rational matrix function which is proper (analytic at infinity) can be written in this form [4]. The quadruple (A, B, C, D) is called a *realization* of $T(z)$. If the system is assumed to be strictly causal, the transfer function is strictly proper $T(\infty) = 0$ and $D = 0$. Note that any change of basis in the state-space, associated with a change-of-basis matrix T, provides a similar realization $(T^{-1}AT, T^{-1}B, CT, D)$. A realization is far from being unique and the choice of a canonical realization is a central issue in system theory.

Transfer functions and state-space representations are complementary descriptions of finite-dimensional LTI systems. The functional framework is rather used for theoretical developments, while the state-space description is very important for numerical computations, since it involves basic linear algebra techniques. We hope this work is a convincing demonstration of the interest of combining the two approaches.

1.2.1 Realization theory

A state space representation can be obtained from the input/output description of the previous section in many ways. We briefly recall the unifying point of view developed in [22, 23] and, in the stable case, connect it to the reproducing kernel Hilbert spaces approach developed in [19] for example.

Past inputs					State	Future outputs			
z^{n+1}	z^n	$\dots$	z	1		z^{-1}	z^{-2}	z^{-3}	$\dots$
	u_0	$\dots$	u_{n-1}	u_n	x_n	y_{n+1}	y_{n+2}	y_{n+3}	$\dots$
u_0	u_1	$\dots$	u_n	0	Ax_n	y_{n+2}	y_{n+3}	$\dots$	$\dots$

Figure 1.2: Time invariance

Time-invariance, which is illustrated in Figure 1.2, can be expressed in terms of commutation properties with the shift operator [23]:

$$\begin{array}{ccccc} \mathbb{F}[z]^m & \xrightarrow{\pi} & \mathcal{X} & \xrightarrow{\bar{\sigma}} & 1/z\,\mathbb{F}[[1/z]]^p \\ \Big\downarrow{\scriptstyle S} & & \Big\downarrow{\scriptstyle S_D} & & \Big\downarrow{\scriptstyle S^*} \\ \mathbb{F}[z]^m & \xrightarrow{\pi} & \mathcal{X} & \xrightarrow{\bar{\sigma}} & 1/z\,\mathbb{F}[[1/z]]^p \end{array}$$

where S is the shift operator in H_2^m :

$$S(u_+(z)) = zu_+(z),$$

S^* is the *backward shift operator* in $(H_2^p)^\perp$,

$$S^*(y_-(z)) = zy_-(z) - zy_-(z)|_\infty$$

and S_D the *restricted shift operator*

$$S_D(\pi(u_+(z)) = \pi(zu_+(z)).$$

The restricted shift operator is thus a linear map on a finite dimensional vector space. A basis of $\mathcal{X}$ being chosen, S_D can be represented by a $n \times n$ matrix A, the restriction of the projection π to $\mathbb{F}^m$ by a $n \times m$ matrix B, and the map $x \in cX \to z\bar{\sigma}(x)(z)|_\infty \in \mathbb{F}^p$ by a $p \times n$ matrix C. The input/output behavior characterized by the transfer matrix $T(z)$ can be alternatively described by the linear system (1.13) in which $D = 0$. The D-matrix of a realization does not affect the dynamical behavior of the system.

The realization obtained in this way is said to be a *minimal realization*. The dimension of the state space in any other realization is greater than the dimension of $\mathcal{X}$ (the *order* of the system or *McMillan degree*). The minimality of the realization relies on some important properties of the matrices A, B and C [35]:

- the set of vectors of the form $\sum A^i B u_i, u_i \in \mathbb{F}^p$ spans the state space, so that

$$\sum \text{Im } A^i B = \mathbb{F}^p, \tag{1.15}$$

  we would say that the pair (A, B) is controllable or reachable.

- the map $\bar{\sigma}$ is injective, so that

$$\bigcap_{j=0}^{n-1} \ker CA^j = 0, \tag{1.16}$$

  we would say that the pair (C, A) is observable.

Attached to any realization, the observability matrix O and the controllability matrix K are defined by

$$O = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}, \quad K = [B, AB, \ldots, A^{n-1}B] \tag{1.17}$$

The pair (A, B) is reachable if and only if the controllability matrix K has full row rank n, while the pair (C, A) is observable if and only if the observability O has full column rank n. Minimality holds if and only if both controllability and observability hold.

Two minimal realizations (A, B, C, D) and $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ associated with a given transfer function are always *similar*: there exists a unique T invertible such that

$$(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}) = (TAT^{-1}, TB, CT^{-1}, D). \tag{1.18}$$

The map :

$$\Phi_T : (A, B, C, D) \mapsto (TAT^{-1}, TB, CT^{-1}, D), \tag{1.19}$$

is called a *state isomorphism* and a *state isometry* if in addition T is unitary.

If the eigenvalues of A all belong to the open unit disk, then the matrix A is called (discrete-time) asymptotically stable, and (A, B, C, D) an asymptotically stable realization. In this case, the controllability Gramian W_c and the observability Gramian W_o are well defined as the convergent series

$$W_c = \sum_{k=0}^{\infty} A^k BB^* (A^*)^k, \quad W_o = \sum_{k=0}^{\infty} (A^*)^k C^* C A^k. \tag{1.20}$$

The Gramians are characterized as the unique (and positive semi-definite) solutions of the respective Lyapunov-Stein equations

$$W_c - AW_cA^* = BB^*, \tag{1.21}$$

$$W_o - A^*W_oA = C^*C. \tag{1.22}$$

Moreover, under asymptotic stability of A it holds that W_c is positive definite if and only if the pair (A, B) is reachable, and W_o is positive definite if and only if the pair (C, A) is observable.

In [19], the connections between a reproducing kernel Hilbert spaces (RKHS) approach to interpolation, and methods based on realization theory as in [4], are clarified. The central role played by finite dimensional $\mathcal{R}_\alpha$ invariant subspaces in realization theory is emphasized. The generalized backward shift operator $\mathcal{R}_\alpha$ is defined on matrix valued functions $F(z)$ by

$$\mathcal{R}_\alpha = \begin{cases} \frac{F(z)-F(\alpha)}{z-\alpha} & \text{if} \quad z \neq \alpha \\ F'(\alpha) & \text{if} \quad z = \alpha \end{cases}$$

We now stress some connections with the RKHS approach in the geometric setting of section 1.1.2, in which the matrix A is asymptotically stable. As previously mentioned, the state space can be viewed either as a the quotient subspace $H(Q)$ of H_2^m or as a subspace of $(H_2^p)^\perp$, the range space $\mathcal{M}$ of $\bar{H}$. The space $H(Q)$ is a finite dimensional RKHS $\mathcal{R}_0$ invariant, while $\mathcal{M}$ is a finite dimensional RKHS $\mathcal{R}_\infty$ invariant.

The range space $\mathcal{M}$ of $\bar{H}$ is spanned by the columns of some $p \times n$ matrix-valued function $F(z)$. Since $\mathcal{M}$ is backward shift invariant, we have

$$S^*(F(z)) = zF(z) - zF(z)|_\infty = F(z)A.$$

But $zF(z)|_\infty = C$ so that $F(z) = C(zI - A)^{-1}$ and $\mathcal{M}$ is spanned by the columns of $C(zI - A)^{-1}$ (see [19, Th. 3.1.]). It must be noticed that the n columns of a $p \times n$ matrix valued function are linearly independent if and only if the pair (C, A) is observable, that is a null kernel pair in the terminology of [4].

Rather than $H(Q)$, we now consider $H(Q)^\sharp = \{v(z);\ \ v^\sharp(z) \in H(Q)\}$ which also satisfies

$$(H_2^{1\times m})^\perp = (H_2^{1\times m})^\perp Q^\sharp \oplus H(Q)^\sharp.$$

As previously, the n-dimensional space $H(Q)^\sharp$ is spanned by the rows of some $n \times m$ matrix $F(z)$ and by shift invariance, we now get $F(z) = (zI - A)^{-1}B$. The space $H(Q)$ is spanned by the column of $B^*(I - zA^*)^{-1}$ which is the general form for a finite dimensional shift invariant subspace of H_2^m (see [19]).

It follows from the Douglas-Shapiro-Shields factorization that the columns of $T(z)$ belong to $H(Q)^\sharp$, as well as the columns of $Q^\sharp(z) - Q^\sharp(\infty)$. We thus get the following realizations

$$T(z) = C(zI - A)^{-1}B, \quad Q^\sharp(z) = \check{D} + \check{C}(zI - A)^{-1}B.$$

Thus $T(z)$ and $Q^\sharp(z)$ share the same pair (A, B).

1.2.2 Null-pole triples and pole-zero structure

The notions of reachable and observable pairs is also essential in studying the pole-zero structure of a rational matrix. We introduce the concept of pole triple and null triple [4, Chap. 3.3] the right generalization of poles and zeros to the matrix case. They serve to address the fundamental interpolation problem: find a rational matrix with a given null pole structure in some domain.

In this section, we only consider square matrices. Given a square rational matrix-valued function $T(z)$ which is regular ($\det T(z)$ does not vanish identically), a triple (C, A, B) is a *pole triple* of $T(z)$ at some pole z_0 of $T(z)$ if and only if

- the pair (A, B) is reachable (full range pair) (1.15)
- the pair (C, A) is observable (null kernel pair) (1.16)
- $T(z) - C(zI - A)^{-1}B$ is analytic at z_0.

More generally, (C, A, B) is called a pole triple of $T(z)$ relative to a compact set K, if the spectrum $\sigma(A)$ of A lies within K and $T(z) - C(zI - A)^{-1}B$ admits an analytic continuation to the whole K. The pair (C, A) is then called a *right pole pair* and the pair (A, B) a *left pole pair* relative to K. The existence of a pole triple can be shown using the local Smith-McMillan form [4, Th. 3.3.1] which clarifies the relation with the pole structure of $T(z)$. Note that any triple $(CT, T^{-1}AT, T^{-1}B)$ similar to (C, B, A) is also a pole triple of $T(z)$ relative to K. The triple (C, A, B) of a minimal realization (1.14) of $T(z)$ is clearly a *global pole triple* of $T(z)$, i.e. a pole triple relative to the whole complex plane.

The definition of null triple is analogous. A triple $(\tilde{C}, \tilde{A}, \tilde{B})$ is a *null triple* of $T(z)$ at some zero z_0 of $T(z)$ iff it is a *pole triple* of $T(z)^{-1}$. If the matrix D in (1.14) is invertible, then a minimal realization of $T(z)^{-1}$ can be computed as

$$T(z)^{-1} = D^{-1} - D^{-1}C(zI - (A - BD^{-1}C))^{-1}BD^{-1}, \tag{1.23}$$

and clearly $(-D^{-1}C, A - BD^{-1}C, BD^{-1})$ is a *global null triple* of $T(z)$. If $(\tilde{C}, \tilde{A}, \tilde{B})$ is a *null triple* of $T(z)$ relative to some compact set K, the pair $(\tilde{A}, \tilde{B})$ is then called a *left null pair* and the pair $(\tilde{C}, \tilde{A})$ a *right null pair* relative to K.

Poles and zeros at infinity can be handled using a Möbius transformation [4, Chap. 3.5].

A scalar function is uniquely determined up to a constant complex number by its poles and zeros. In the matrix case, we may ask the following question: does a right pole pair (C, A) and a left null pair $(\tilde{A}, \tilde{B})$ determine, up to a constant factor, a rational matrix $T(z)$ (analytic and invertible at infinity)? The answer is no. An extra condition is required which is that the Sylvester equation

$$SA - \tilde{A}S = \tilde{B}C \tag{1.24}$$

must have an invertible solution, which is called the *null-pole coupling matrix* [4, Th. 4.3.1]. In this case the unique matrix $T(z)$ such that $T(\infty) = I$, (C, A) is a right pole pair and $(\tilde{A}, \tilde{B})$ is a left null pair is given by

$$T(z) = I + C(zI - A)^{-1}S^{-1}\tilde{B}, \text{ while } T(z)^{-1} = I - CS^{-1}(zI - \tilde{A})^{-1}\tilde{B}.$$

A null-pole triple contains left-side information about $T(z)$ in the sense that it determines $T(z)$ uniquely, up to a right invertible constant matrix factor.

Given a subset $K \subset \mathbb{C}$ and a rational matrix function $T(z)$, analytic and invertible at infinity, we refer to a set $(C, A; \tilde{A}, \tilde{B}; S)$ as a *null-pole triple* for over K, if

- (C, A) is a right pole pair for $T(z)$ with respect to K
- $(\tilde{A}, \tilde{B})$ is a left null pair for $T(z)$ with respect to K
- S is the associated null pole coupling matrix, i.e. the invertible matrix satisfying (1.24).

1.2.3 Balanced realizations and canonical forms

A transfer function possesses many minimal realizations.

A *canonical form* on a set endowed with an equivalence relation consists in the choice of a unique element within every class. In our case, it is the choice of a unique representative among all the similar realizations associated with a transfer function in some class. Many canonical forms are known, as the Jordan form, the companion form, observer and controller forms, Popov and echelon form [35]. In the SISO case, a same canonical form can be used for the whole class of stable systems, but not in the MIMO case. In engineering, systems are usually modeled by a realization with some structure, whose entries are related to the physical parameters. The concept of canonical form serves to clarify the connection between such structured realizations and a mathematical model and to address the identifiability issue. Depending on the underlying application, one or the other canonical form would better suit.

A minimal and asymptotically stable realization (A, B, C, D) of a transfer function is called *balanced* if W_o and W_c, its observability and controllability Gramians (1.20), are both diagonal and equal. Any minimal and asymptotically stable realization is similar to a balanced realization. The concept of balanced realizations was first introduced in [51] in the continuous time case. In [53] the same was done for the discrete time case. Balanced realizations are now a well-established tool for model reduction which often exhibit good numerical properties. Two distinct balanced realizations associated with

the same function are related by a state isometry. With balancedness we have made some progress towards a canonical form.

A system is called *input-normal* if $W_c = I$ and it is called *output-normal* if $W_o = I$. Balanced realizations are directly related to input-normal and output-normal realizations, respectively, by diagonal state isomorphism. The property of input-normality (resp. output-normality) is preserved under state isometry.

For a lossless function, balanced realizations present a particular interest. To a realization of the form (1.14), we associate the block-partitioned matrix

$$R = \begin{bmatrix} D & C \\ B & A \end{bmatrix} \tag{1.25}$$

which we call the *realization matrix*[1]. The following proposition characterizes the balanced realizations of rational lossless functions in discrete time.

Proposition 1.2.1 *[HOP06] (i) For any minimal balanced realization of a $m \times m$ rational lossless function the observability and controllability Gramians are both equal to the identity matrix and* the associated realization matrix (1.25) is unitary.
(ii) Conversely, if the realization matrix associated with a realization (A, B, C, D) of order n of some $m \times m$ rational function G is unitary, then G is lossless of McMillan degree $\leq n$. The realization is minimal if and only if A is asymptotically stable and then it is balanced.

The first point is classical (see e.g. [25]). The second point asserts that unitary realizations matrices correspond to possibly non-minimal realizations of lossless functions and give an interpretation of limit points in a balanced canonical form.

1.3 Interpolation and parametrization

Analytic interpolation originates in the work of I. Schur [59] and his famous algorithm. The Schur algorithm is a nice recursive test for checking the boundedness of an analytic function $f(z)$ in the disk: define a sequence of functions by $f_0 = f$ and

$$f_{k+1}(z) = \frac{f_k(z) - \gamma_k}{z(1 - \gamma_k^* f_k(z))}, \quad \gamma_k = f_k(0). \tag{1.26}$$

This algorithm establishes a one-to-one correspondence between the Schur class

$$\mathcal{S} = \{f(z);\ |f(z)| \leq 1 \text{ for } |z| \leq 1\}$$

and the sequences of complex numbers (γ_k) which satisfy:

- $|\gamma_k| \leq 1$ for all $k \geq 0$
- when $|\gamma_n| = 1$ for a certain n, then for all $k > n$, $\gamma_k = 0$.

[1]This positioning of the block matrices is rather unusual but it is very convenient for exhibiting a triangular structure in the sub-matrix $\begin{bmatrix} B & A \end{bmatrix}$.

This last situation appears exactly when the function $f(z)$ is a finite Blaschke product of degree n, that is a scalar inner function. In this case, both a pole and a zero cancellation arises in the Schur recursion so that f_{k+1} has one degree less than that of f_k. The set of inner functions of McMillan degree n can thus be parametrized by the sequence of Schur numbers $(\gamma_0, \gamma_1, \ldots, \gamma_n)$.

An important part of this work is concerned with the parametrization of inner (or lossless) matrix-valued functions derived from analytic interpolation theory. In the matrix case, the same two ingredients are involved :

- under some conditions, all the solutions to the interpolation problem can be represented by means of a linear fractional transformation (LFT) which generalizes (1.26).
- for rational inner functions, the LFT induces a decrease of degree which corresponds to the "size" of the interpolation condition.

In this section we present the background concerning interpolation problems and their use for parametrization issues. A matrix function is Schur if it belongs to the unit ball of the Banach space $(L^\infty(\mathbb{T}))^{p\times m}$ endowed with the norm

$$\|F\|_\infty = \sup_\theta ||F(e^{i\theta})||,$$

where $||\ .\ ||$ denotes the operator norm $\mathbb{C}^m \to \mathbb{C}^p$. For a pair of matrices P and Q, the notation $P \leq Q$ (resp $P < Q$) means $Q - P$ positive semi-definite (resp. positive definite). With this convention, a $p \times m$ rational function $S(z)$ is Schur iff it is analytic and contractive in the open unit disk $S(z)^* S(z) \leq I$ for $z \in \mathbb{D}$, that is to say $\|S\|_\infty \leq 1$.

Interpolation theory deals with functions analytic in the disk while system theory deals with functions analytic outside the unit disk. To relate these two situations, we use the isometric transformation (1.9).

Linear fractional transformations play a crucial role in this story. We begin with some results on these maps and we specify which linear fractional transformations preserve the metric constraint.

1.3.1 J-inner functions and linear fractional transformations.

We now introduce the concept of J-inner function, where J is generally a signature matrix. A matrix-valued function$\Theta(z)$ is a J-inner function if at every point of analyticity z of $\Theta(z)$ it satisfies

$$\Theta(z)^* J \Theta(z) \leq J, \quad |z| < 1, \tag{1.27}$$
$$\Theta(z)^* J \Theta(z) = J, \quad |z| = 1, \tag{1.28}$$
$$\Theta(z)^* J \Theta(z) \geq J, \quad |z| > 1. \tag{1.29}$$

A J-lossless function is a function which is $(-J)$-inner. Note that J-inner and J-lossless functions in general may have poles everywhere in the complex plane. Condition

(1.29) which is both satisfied by J-inner and J-lossless functions implies that at any points where $\Theta(z)$ is analytic and invertible

$$\Theta(z)^{-1} = J\Theta^{\sharp}(z)J. \tag{1.30}$$

For a further background on lossless systems and their realizations, see e.g. [25].

When $J = I_p \oplus -I_q$, an interesting physical link can be stressed between lossless and J-lossless functions. Assume that $G(z)$ is the lossless transfer function of a passive system with $(p+q)$-inputs and $(p+q)$-outputs described by a quadripole (see Figure 1.3).

Figure 1.3: Quadripole

We thus have

$$\begin{bmatrix} a_2 \\ b_1 \end{bmatrix} = G \begin{bmatrix} a_1 \\ b_2 \end{bmatrix}, \text{ or equivalently } \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = \Theta \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}.$$

where $\Theta(z)$ is a J-lossless function easily computed from $G(z)$ (Ginzburg transform). In system theory, $G(z)$ is usually referred as the scattering matrix of the system, while $\Theta(z)$ is the chain matrix. In terms of chain matrices, the cascade of two such devices is just a product, which makes the interest of the latest compared to scattering matrices [40].

From now on, we assume that $J = I_m \oplus -I_m$. Our interest in J-inner functions relies on the following result.

Proposition 1.3.1 *Let $\Theta(z)$ be J-inner of size $2m \times 2m$ and of McMillan degree k*

$$\Theta(z) = \begin{bmatrix} \Theta_{11}(z) & \Theta_{12}(z) \\ \Theta_{21}(z) & \Theta_{22}(z) \end{bmatrix}.$$

Then, the linear fractional transformation

$$T_\Theta(F) = (\Theta_{11}F + \Theta_{12})(\Theta_{21}F + \Theta_{22})^{-1}, \tag{1.31}$$

is defined for every rational $m \times m$ Schur function $F(z)$ and preserves the metric constraint:

$$\|F\|_\infty \leq 1 \Longrightarrow \|T_\Theta(F)\|_\infty \leq 1.$$

If $Q(z)$ is $m \times m$ inner of McMillan degree n, then $T_\Theta(Q)$ is also inner of McMillan degree $\leq n + k$.

In the literature, this result is usually stated for J-inner functions analytic in the disk [4], but this condition is not necessary [20].

The linear fractional transformation T_H associated with a constant J-unitary matrix H

$$H^*JH = J \tag{1.32}$$

is a bijection from the set of inner functions [55] to itself which preserves the McMillan degree [HOP06]. It will be called a *generalized Möbius transform*.

Every *constant* J-unitary matrix can be represented in a unique way (see [18, Th. 1.2]) as follows:

$$M = H(E) \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}, \tag{1.33}$$

where P and Q are $m \times m$ unitary matrices and $H(E)$ denotes the Halmos extension of a strictly contractive $m \times m$ matrix E (i.e., such that $I - E^*E > 0$)

$$\begin{aligned} H(E) &= \begin{bmatrix} I & E \\ E^* & I \end{bmatrix} \begin{bmatrix} (I-EE^*)^{-1/2} & 0 \\ 0 & (I-E^*E)^{-1/2} \end{bmatrix} && (1.34) \\ &= \begin{bmatrix} (I-EE^*)^{-1/2} & 0 \\ 0 & (I-E^*E)^{-1/2} \end{bmatrix} \begin{bmatrix} I & E \\ E^* & I \end{bmatrix} && (1.35) \end{aligned}$$

1.3.2 J-lossless and lossless embedding

The J-unitary property (1.28) implies some relations between poles and zeros of a rational function, which can be stated as follows: the set $(C, A; \tilde{A}, \tilde{B}; S)$ is a null pole triple for $\Theta(z)$ over $\mathbb{D}$ if and only if $(-J\tilde{B}^*, \tilde{A}^*; A^*, C^*; S^*)$ is a null pole triple for $\Theta(z)$ over $\mathbb{C} \setminus \mathbb{D} \bigcup \{\infty\}$ [4, Lemma 7.4.1]. A J-lossless function is thus completely determined, up to a constant J-unitary factor, from a null pole triple over $\mathbb{D}$ [4, Th. 7.4.2]. If in addition the J-lossless is assumed to be analytic outside the unit disk, a null pole triple over $\mathbb{D}$ consists just in a right pole pair, which happens to be global, and we have the following result that can also be found in [43].

Proposition 1.3.2 (J-lossless embedding) *Given a reachable pair (A, B) with A asymptotically stable, there exits a unique Hermitian matrix P that satisfies the Stein equation*

$$P - APA^* = BJB^*.$$

If $P > 0$, then the matrix-valued functions $\Theta(z) = \check{D} + \check{C}(zI - A)^{-1}B$, with

$$\begin{aligned} \check{C} &= -JB^*(I - \nu A^*)^{-1}P^{-1}(A - \nu I) \\ \check{D} &= I - JB^*(I - \nu A^*)^{-1}P^{-1}B \end{aligned}$$

is J-lossless for every ν such that $|\nu| = 1$.
The J-lossless function $\Theta(z)$ satisfies $\Theta(\nu) = I$ and it is the only J-lossless function with global left pole pair (A, B) that satisfies this property.
All other J-lossless functions with the same left pole pair is given by $H\Theta(z)$ where H is a constant J-unitary function: $H^*JH = J$.

The matrix $\Theta(z)$ in Proposition 1.3.2 can be written in the form

$$\Theta(z) = I - (z-\nu)JB^*(I-\nu A^*)^{-1}P^{-1}(zI-A)^{-1}B. \tag{1.36}$$

In the case $J = I$, the lossless embedding provides a one-to-one correspondence between the set of reachable pair (A, B), A asymptotically stable, up to similarity, and the set of lossless functions up to a left unitary matrix, defined by

$$(A, B) \mapsto I - (z-\nu)B^*(I-\nu A^*)^{-1}P^{-1}(zI-A)^{-1}B.$$

This correspondence is a diffeomorphism (see [1, Cor.2.1]). If in addition the matrix $\begin{bmatrix} B & A \end{bmatrix}$ is assumed to have orthonormal rows ($AA^* + BB^* = I$), the lossless embedding consists in completing it into a unitary realization matrix

$$\begin{bmatrix} B & A \end{bmatrix} \mapsto \begin{bmatrix} D & C \\ B & A \end{bmatrix}, \quad \begin{cases} C &= -B^*(I-\nu A^*)^{-1}P^{-1}(A-\nu I), \\ D &= I - B^*(I-\nu A^*)^{-1}P^{-1}B. \end{cases}$$

There are many ways to perform such a completion. In Chapter 3 we propose an interesting method based on a Cholesky factorization.

1.3.3 The Nudelman interpolation problem.

The most general form of a one-sided (left) interpolation condition for a matrix-valued Schur function $F(z)$ is in term of a contour integral [4]

$$\frac{1}{2i\pi}\int_{\mathbb{T}}(zI - W^*)^{-1}U^*F(z)\,dz = V^*, \tag{1.37}$$

where (U, W) is an observable pair and W is asymptotically stable (U is $m \times k$ and W is $k \times k$). Note that if W is a diagonal matrix

$$\begin{aligned} W &= \operatorname{diag}(\overline{w}_1, \overline{w}_2, \ldots, \overline{w}_k),\ w_i \neq w_j, \\ U &= \begin{bmatrix} u_1 & u_2 & \ldots & u_k \end{bmatrix}, \\ V &= \begin{bmatrix} v_1 & v_2 & \ldots & v_k \end{bmatrix}, \end{aligned}$$

this problem reduces to a Nevanlinna-Pick problem :

$$u_i^*F(w_i) = v_i^*, i = 1, \ldots, k. \tag{1.38}$$

Theorem 1.3.1 *Let (U, W) be an observable pair.*

(i) the left interpolation problem (1.37) admits a solution $F(z)$ if and only if the solution P of the symmetric Stein equation

$$P - W^*PW = U^*U - V^*V \tag{1.39}$$

is positive semi-definite.

(ii) If $P > 0$, the set of all solutions $F(z)$ of the Nudelman interpolation problem (1.37) is given by

$$\{F = T_\Theta[G] : G \text{ Schur } \}$$

where T_Θ is the LFT (1.31) associated with the J-inner function

$$\Theta(z) = I + (z - \nu) \begin{bmatrix} U \\ V \end{bmatrix} (I - Wz)^{-1} P^{-1} (\nu I - W^*)^{-1} \begin{bmatrix} U \\ V \end{bmatrix}^* J, \tag{1.40}$$

uniquely determined up to an arbitrary unit complex number ν and an arbitrary constant right J-unitary factor.

This result can be found in many works, following different approaches. In [4, Ch.18], only strictly Schur solutions are considered and it is shown that a strictly Schur solution of (1.37) does exit if and only if $P > 0$. In this book, only concerned with rational matrix-valued functions, an approach based on realization theory is chosen. A parametrization by means of an LFT is searched a priori and it is shown that the interpolation problem (1.37) is to find a J-inner function $\Theta(z)$ with a prescribed global left null pair $(W^*, [U^*\ V^*]J)$. Equivalently, $(W^*, [U^*\ V^*]J)$ must be a global left pole pair of the J-lossless function $\Theta(z)^{-1}$. By the Lossless Embedding (Proposition 1.3.2), $\Theta(z)^{-1}$ is uniquely determined up to a left unitary factor H by the formula (1.36), which yields (1.40) using the relation $\Theta(z)^{-1} = J\Theta^\sharp(z)J$.

In a RKHS approach (see e.g. [18] or [19]), the k-dimensional subspace $\mathcal{M}$ of H_2^k built from the interpolation data

$$\mathcal{M} = \text{span} \left\{ \begin{bmatrix} U \\ V \end{bmatrix} (I - W z)^{-1} \right\},$$

plays a central role. Endowed with the metric induced by $P > 0$ (J-inner product), the space $\mathcal{M}$ is a RKHS with a reproducing kernel of the form

$$K_\mathcal{M}(z, w) = \frac{J - \Theta(z) J \Theta(w)^*}{1 - \bar{w} z}.$$

If $F(z)$ is a solution of (1.37), then the map of multiplication by $[I - F]$ is an isometry form $\mathcal{M}$ into $H(F)$. This isometry forces an LFT between $F(z)$ and $\Theta(z)$.

1.3.4 The tangential Schur algorithm vs Potapov factorization

The tangential Schur algorithm recursively handles interpolation constraints of the form

$$Q(w)^* u = v, \tag{1.41}$$

for an inner matrix-valued function $Q(z)$, and a triple (w, u, v), $w \in \mathbb{D}$, u and v m-vectors, $\|u\| = 1$ and $\|v\| < 1$.

In this case, the solution of (1.39) is the strictly positive matrix

$$P = \frac{1 - \|v\|^2}{1 - |w|^2}.$$

The J-inner function (1.40) is given by

$$\Theta(z) = I + (z-\nu)\begin{bmatrix} u \\ v \end{bmatrix}(1-z\bar{w})^{-1}P^{-1}(\nu-w)^{-1}\begin{bmatrix} u \\ v \end{bmatrix}^* J. \tag{1.42}$$

In this section, we propose a specific treatment for this particular case which brings some insights and will be useful in Chapter 2.

For any constant J-unitary matrix H, $Q(z)$ satisfies the interpolation condition (1.41) if and only if $T_H(Q)(z)$ satisfies

$$T_H(Q)(w)^* x = y, \quad \text{with} \begin{bmatrix} x \\ y \end{bmatrix} = H \begin{bmatrix} u \\ v \end{bmatrix}.$$

Choosing $H = H(uv^*)$, the Halmos extension (1.34) of the contractive matrix uv^*, it is easily checked that

$$H(uv^*)J\begin{bmatrix} u \\ v \end{bmatrix} = \sqrt{1-\|v\|^2}\begin{bmatrix} u \\ 0 \end{bmatrix},$$

so that the interpolation condition (1.41) becomes $\hat{Q}(w)^* u = 0$. Finally, $Q(z)$ is a solution of (1.41) if and only if w is a zero of $\hat{Q}(z)$ associated to the left kernel vector u^*.

This condition can be handled by the Potapov method. An elementary inner factor of the form (1.11) can be extracted from $\hat{Q}(z)$, so that $\hat{Q} = B_{w,u}Q_1$, where $Q_1(z)$ is inner. This operation can be written in an LFT form, $\hat{Q} = T_{S_{w,u}}(Q_1)$, where $S_{w,u}(z)$ is the block diagonal matrix $S_{w,u} = B_{w,u} \oplus I$. We thus get the following linear fractional representation for the solutions of (1.41)

$$Q = T_{\Theta_{w,u,v,H}}(Q_1), \quad \Theta_{w,u,v,H}(z) = H(uv^*)S_{w,u}(z)H, \tag{1.43}$$

for some inner function Q_1. The matrix function $\Theta_{w,u,v,H}$ includes an arbitrary constant J-unitary factor on the right which produces a generalized Möbius transform T_H and provides equivalent representations of the solutions (see section 1.3.1).

Formula (1.42) with $\nu = 1$ can be written in the following form [18, Th. 1.4.]

$$\Theta_{w,u,v}(z) = I + \left(\frac{b_w(z)}{b_w(1)} - 1\right)\frac{\begin{bmatrix} u \\ v \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}^* J}{1-\|v\|^2}, \tag{1.44}$$

which possesses nice multiplicative properties [18, Th. 1.3.]. This form was used in [1] and [FO98]. It corresponds to $\Theta_{w,u,v,H}(z)$ in (1.43) with $H = H(uv^*)^{-1}$ and satisfies $\Theta_{w,u,v,H}(1) = I$, so that the associated LFT preserves the value at 1 of the inner function.

In [HOP06], the freedom in the choice of the matrix H has been used to associate with the Schur algorithm a nice recursive construction of balanced realizations. The J-inner matrix is

$$\hat{\Theta}_{w,u,v}(z) = H(uv^*)S_{w,u}(z)H(\bar{w}uv^*)$$

These results will be presented in Chapter 2.

The tangential Schur algorithm with interpolation points anywhere in the disk has been scarcely used in system theory literature. In [39], a similar Schur algorithm is presented which gives rise to a circuit theoretical interpretation. It is described in the case of continuous-time transfer functions and it corresponds in our discrete-time setting to the choice of the J-inner functions $\hat{\Theta}_{0,u,v}(z)$.

1.3.5 Parametrizations from the tangential Schur algorithm.

If the Schur algorithm (1.26) was devised to characterize and parametrize scalar Schur functions, interpolation theory is scarcely used for parametrization issues in the matrix case. The first attempt in this direction was to our knowledge [1], in which the tangential Schur algorithm is used to construct an atlas of charts for the set of matrix-valued lossless functions of fixed McMillan degree. More recently, interpolation theory has been used for a more constrained problem: the parametrization of Schur functions with a degree constraint and satisfying some interpolation conditions (see e.g. [13] and the bibliography therein).

We now provide an overview of [1], which emphasizes on the advantages of an interpolation theory approach to parametrization issues:

- atlases of charts are obtained, which are the nice parametrizations in view of optimization
- the Schur constraint is easily handled.

A (smooth) manifold $\mathcal{M}$ is a mathematical space in which every point has a neighborhood which is homeomorphic to the Euclidean space $\mathbb{R}^d$, where d is the *dimension* of the manifold. The structure of a manifold is encoded by a collection of charts that form an atlas[2], that is a collection of coordinate maps

$$\phi_i : \mathcal{V}_i \to \mathbb{R}^d, \quad \mathcal{V}_i \subset \mathcal{M} \quad \text{open} \tag{1.45}$$

such that the $\mathcal{V}_i$'s cover the manifold and the *transition maps* or *change of coordinates* $\phi_i \circ \phi_j^{-1}$ are smooth [61]. Differential calculus can be extended to a manifold structure.

The set of $m \times m$ inner matrix functions of McMillan degree n has a manifold structure and the tangential Schur algorithm provides an atlas of charts for this manifold [1]. Given an inner function $Q(z)$ of McMillan degree n, a sequence of interpolation points $\mathbf{w} = (w_n, w_{n-1}, \ldots, w_1)$ and a sequence of interpolation direction vectors $\mathbf{u} = (u_n, u_{n-1}, \ldots, u_1)$, the tangential Schur algorithm (section 1.3.4) yields a sequence of functions defined by $Q_n = Q$ and for $k = n, n-1, \ldots$ let

$$v_k = Q_k(w_k)^* u_k.$$

- if $\|v_k\| < 1$, let

$$Q_{k-1} = T^{-1}_{\Theta_{w_k,u_k,v_k}} Q_k, \tag{1.46}$$

  in which $\Theta_{w_k,u_k,v_k}(z)$ is defined by (1.44).

[2] in analogy with an atlas consisting of charts of the surface of the Earth

- if v_k has norm 1, then stop.

The algorithm stops if v_k fails to have norm strictly less than 1. As in the scalar case, the degree of $Q_{k-1}(z)$ is one less than that of $Q_k(z)$, so that the algorithm stops after at most n steps. In the matrix case, an inner function may fail to be strictly contractive in some direction. If the algorithm meets such a situation then it stops. This is an important difference with the Schur algorithm for scalar inner functions (1.26) which always performs n steps. The manifold structure of the set of matrix inner functions of fixed McMillan degree is not trivial and several coordinate maps are necessary to describe the whole set.

A chart is attached to a sequence of interpolation points $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ and a sequence of interpolation directions $\mathbf{u} = (u_1, u_2, \ldots, u_n)$. An inner function $Q(z)$ belongs to the domain $\mathcal{V}_{\mathbf{u},\mathbf{v}}$ of this chart if and only if the tangential Schur algorithm (1.46) stops after n steps

$$Q = Q_n \xrightarrow{w_n,u_n} \cdots Q_k \xrightarrow{w_k,u_k} \cdots Q_1 \xrightarrow{w_1,u_1} \cdots Q_0,$$

where Q_0 is a constant unitary matrix, and thus provides a complete sequence of interpolation vectors v_k, $k = n, \ldots, 1$ such that $\|v_k\| < 1$, the *Schur parameter vectors*. The coordinate map is

$$\phi_{\mathbf{w},\mathbf{u}} : \; Q \in \mathcal{V}_{\mathbf{w},\mathbf{u}} \to (Q_0, v_1, v_2, \ldots v_n).$$

The dimension of the manifold is $2nm + m^2$.

This atlas is very rich and flexible since it possesses an infinite number of charts. Given an inner function $Q(z)$, it is possible to zoom on it by choosing what we call an *adapted chart*, a chart centered at $Q(z)$, in which all the Schur vectors are 0. It is easily obtained from a Potapov factorization (1.11)

$$Q(z) = B_{w_n,u_n}(z) B_{w_{n-1},u_{n-1}} \ldots B_{w_1,u_1}(z) Q_0$$

which can be viewed as a particular case of the tangential Schur algorithm in which all the interpolation vectors are zero, $Q_k(w_k)^* u_k = 0$, for $k = n, n-1, \ldots, 1$ (see section 1.3.4). In practice, a realization in Schur form yields the interpolation points $w_1, w_2, \ldots, w_n$ and the interpolation directions $u_1, u_2, \ldots, u_n$ [MOHP02, MO07]. It can be very conveniently used for optimization purposes, moving from one chart to another when the conditioning becomes to bad. On the other way, a chart contains "almost" all the inner functions so that changes of charts should not happen very often.

In many applications, the Douglas-Shapiro-Shield factorization of a proper transfer function brings into play quotient spaces of inner functions up to a right (or left) constant unitary matrix. Atlases of charts for these quotient spaces are easily deduced from the following formula: for any $Q(z)$ matrix-valued inner function, Λ and Π constant unitary matrices

$$T_{\Theta_{w,\Lambda u,\Pi v}}(\Lambda Q \Pi^*) = \Lambda T_{\Theta_{w,u,v}}(Q) \Pi^*. \tag{1.47}$$

Indeed, if $Q = \phi^{-1}_{\mathbf{w},\mathbf{u}}(v_n, v_{n-1}, \ldots, v_1, Q_0)$, then for any unitary matrix Π we have that $Q\Pi^* = \phi_{\mathbf{w},\mathbf{u}}(\Pi^* Q_0, \Pi v_1, \Pi v_2, \ldots, \Pi v_n)$, so that we can choose a representative of $Q(z)$

by fixing the final unitary matrix Q_0 in the Schur algorithm. A chart of the right quotient is thus attached to a triple $(\mathbf{w}, \mathbf{u}, Q_0)$ and the associated homeomorphism is

$$\phi_{\mathbf{w},\mathbf{u},Q_0}(Q) = (v_1, v_2, \ldots, v_n).$$

Formula (1.47) also holds true for $\hat{\Theta}_{w,u,v}(z)$.

Chapter 2

Schur vectors and balanced canonical forms for lossless systems

In this chapter, the outcome of a longstanding collaboration with B. Hanzon and R. Peeters is reported. This collaboration started during the MTNS 1995 where our respective results on the parametrization of lossless functions were presented. My work was on the use of the Schur atlas (section 1.3.5) for matrix rational approximation [FO98], while their contribution established a connection between the Hessenberg canonical form and the Schur algorithm in the scalar case [31]. We agreed on the interest of looking for such a connection in the matrix case: a representation which combines the properties of balanced canonical forms (section 1.2.3) and that of the Schur algorithm (section 1.3.4) should bring some new insights and prove useful tool in many applications.

The Hessenberg canonical form for a lossless scalar function of McMillan degree n is a balanced realization (A, b, c, d) such that the sub-matrix $\begin{bmatrix} b & A \end{bmatrix}$ is positive upper triangular. Such a realization is uniquely defined and the triangular structure in $\begin{bmatrix} b & A \end{bmatrix}$ induces a triangular structure in the reachability matrix $K = [b, Ab, A^2b, \ldots, A^{n-1}b]$.

The realization matrix $R = \begin{bmatrix} d & c \\ b & A \end{bmatrix}$ is then unitary (Proposition 1.2.1) and a factorization is obtained by means of Givens rotations in a recursive way: let $\gamma_n = d$, $|\gamma_n| < 1$, and put $\kappa_n = \sqrt{1 - |\gamma_n|^2}$, then

$$\begin{bmatrix} \bar{\gamma}_n & \kappa_n & 0 \\ \kappa_n & -\gamma_n & 0 \\ 0 & 0 & I_{n-1} \end{bmatrix} \underbrace{\left[\begin{array}{c|c} d & c \\ \hline * & \\ 0 & A \end{array}\right]}_{R} = \begin{bmatrix} 1 & 0 \\ 0 & R_{n-1} \end{bmatrix}, \tag{2.1}$$

where R_{n-1} is in Hessenberg form of order $n-1$.
Repeating this process, we get a sequence of realization matrices $(R_k)_{k=n,\ldots,0}$, R_k of order k still in Hessenberg form and a sequence of parameters $(\gamma_k)_{k=n,\ldots,0}$, with $|\gamma_k| <$

1 for $k > 0$ and $|\gamma_0|$=1, which parametrize R:

$$R = \left[\begin{array}{c|cccccc} \gamma_n & \kappa_n\gamma_{n-1} & \kappa_n\kappa_{n-1}\gamma_{n-2} & \cdots & & \kappa_n\kappa_{n-1}\cdots\kappa_1\gamma_0 \\ \hline \kappa_n & -\bar{\gamma}_n\gamma_{n-1} & -\bar{\gamma}_n\kappa_{n-1}\gamma_{n-2} & \cdots & & -\bar{\gamma}_n\kappa_{n-1}\cdots\kappa_1\gamma_0 \\ 0 & \kappa_{n-1} & -\bar{\gamma}_{n-1}\gamma_{n-2} & \cdots & & -\bar{\gamma}_{n-1}\kappa_{n-2}\cdots\kappa_1\gamma_0 \\ 0 & 0 & \kappa_{n-2} & & & -\bar{\gamma}_{n-2}\kappa_{n-3}\cdots\kappa_1\gamma_0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \ddots & \kappa_2 & -\bar{\gamma}_2\gamma_1 & -\bar{\gamma}_2\kappa_1\gamma_0 \\ 0 & 0 & \cdots & 0 & \kappa_1 & -\bar{\gamma}_1\gamma_0 \end{array}\right] \tag{2.2}$$

The interesting point established in [31] is that the sequence of lossless functions

$$g_k(z) = \gamma_k + c_k(zI_k - A_k)^{-1}b_k, \quad k = n, \ldots, 0$$

associated with the sequence of realization matrices $(R_k)_{k=n,\ldots,0}$ is obtained by a Schur algorithm from $g(z) = d + c(zI - A)^{-1}b$. This algoritm is an analog of (1.26) but for functions contractive outside the unit disk, so that the interpolation condition is at ∞:

$$\left\{\begin{array}{rcl} g_k(\infty) & = & \gamma_k \\ g_{k-1}(z) & = & \frac{(g_k(z)-\gamma_k)z}{1-\bar{\gamma}_k g_k(z)}. \end{array}\right. \tag{2.3}$$

The parameters $\gamma_n, \gamma_{n-1}, \ldots, \gamma_1, \gamma_0$ in the Hessenberg form are thus Schur parameters.

The generalization of this result to the matrix case is far from being straightforward and a complete description of structured realizations in terms of a Schur algorithm is obtained in three steps

- a recursive construction of balanced realization is obtained from an adapted tangential Schur algorithm [HOP06].
- specifying canonical interpolation points and directions in the Schur algorithm, a sub-diagonal pivot structure is obtained for the realization matrix. It generalizes the triangular structure of the Hessenberg from [HOP09].
- a condition on the sequence of interpolation directions is derived, which ensures that a pivot structure in the realization matrix induces a similar structure in the reachability matrix [PHO07].

2.1 Balanced canonical forms from the tangential Schur algorithm.

This section presents the results of [HOP06]. A connection is established between a tangential Schur algorithm and a recursive construction of balanced realizations which generalizes (2.1). The interpolation points can be chosen anywhere in the complement of the disk (in (2.1) interpolation points are at ∞). We get more flexibility in the choice

of a chart which is interesting from an optimization point of view (see Chapter 3), but the triangular structure is lost.

To deal with balanced canonical forms we must move from the framework of inner functions to that of lossless functions using the transformation (1.9): $Q \mapsto Q^\sharp$. Since we have

$$\tilde{Q} = T_\Theta(Q) \Leftrightarrow \tilde{Q}^\sharp = \mathcal{T}_\Theta(Q^\sharp),$$

where $\mathcal{T}_\Theta$ is the linear fractional transformation

$$\mathcal{T}_\Theta(G) = (\Theta_{22}G + \Theta_{21})(\Theta_{12}G + \Theta_{11})^{-1}, \tag{2.4}$$

the results of section 1.3 can be immediately translated in terms of lossless functions.

If $G(z)$ is $(m \times m)$ lossless of McMillan degree n and $\Theta(z)$ $(2m \times 2m)$ J-inner of McMillan degree k, then $\mathcal{T}_\Theta(G)$ exists and the matrix function $\tilde{G} = \mathcal{T}_\Theta(G)$ is also lossless and of McMillan degree $\leq n + k$ (Proposition 1.3.1). The left interpolation condition (1.41) changes into a right interpolation condition for $\tilde{G}(z)$

$$\tilde{G}\,(1/\bar{w})\,u = v. \tag{2.5}$$

The solutions can be represented by the following LFT (compare with (1.43))

$$\tilde{G} = \mathcal{T}_{\Theta_{w,u,v,H}}(G), \quad \Theta_{w,u,v,H}(z) = H(uv^*)S_{w,u}(z)H \tag{2.6}$$

In [PHO01], a unified framework is presented in which linear fractional transformations on transfer functions are represented by corresponding linear fractional transformations on state-space realization matrices. However, these formulas are rather complicated and involve matrix inversions. The question is thus: can we attach to (2.6) a simpler computation of balanced realizations which generalizes (2.1)?

In the particular case where the Schur vector v in (2.5) is the null vector, the LFT is just a matrix product (see section 1.3.4)

$$\tilde{G}(z) = G(z)B^\sharp_{w,u}(z),$$

and the cascade realization $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ of $\tilde{G}(z)$ can be computed from a realization (A, B, C, D) of $G(z)$ by [25, 43]

$$\begin{bmatrix} \tilde{D} & \tilde{C} \\ \tilde{B} & \tilde{A} \end{bmatrix} = \begin{bmatrix} D & 0 & C \\ 0 & 1 & 0 \\ B & 0 & A \end{bmatrix} \begin{bmatrix} I_p - (1+\bar{w})uu^* & \sqrt{1-|w|^2}\,u & 0 \\ \sqrt{1-|w|^2}\,u^* & w & 0 \\ 0 & 0 & I_{n-1} \end{bmatrix} \tag{2.7}$$

A cascade decomposition can be easily obtained from a realization in Schur form that is a realization for which the matrix A is upper triangular. This representation could be used for parametrization purposes. Then, the parameters would be the zeros and corresponding directions, and it is well known that zeros does not behave nicely as optimization parameters.

Comparing (2.1) and (2.7), a generalized recursion for balanced realization matrices can be conjectured in the form:

$$\begin{bmatrix} \widetilde{D} & \widetilde{C} \\ \widetilde{B} & \widetilde{A} \end{bmatrix} = \begin{bmatrix} \mathcal{V} & 0 \\ 0 & I_{n-1} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & D & C \\ 0 & B & A \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & I_{n-1} \end{bmatrix}, \tag{2.8}$$

where $\mathcal{U}$ and $\mathcal{V}$ are $(m+1)\times(m+1)$ unitary matrices. This state-space recursion actually defines a mapping on lossless systems,

$$\mathcal{F}_{\mathcal{U},\mathcal{V}}:\ G(z)=D+C(zI-A)^{-1}B \longrightarrow \tilde{G}(z)=\tilde{D}+\tilde{C}(zI-\tilde{A})^{-1}\tilde{B},$$

which coincides with a linear fractional transformation .

Proposition 2.1.1 *[HOP06, Th. 6.1] Let $\mathcal{U}$ and $\mathcal{V}$ be partitioned as*

$$\mathcal{U}\ =\ \begin{bmatrix} \alpha_u & M_u \\ K_u & \beta_u^* \end{bmatrix},\quad \mathcal{V}\ =\ \begin{bmatrix} \alpha_v & M_v \\ K_v & \beta_v^* \end{bmatrix}, \tag{2.9}$$

where K_u and K_v are scalar, α_u, α_v, β_u and β_v are m-vectors and M_u and M_v are $m\times m$. Assume that $K_v - K_u z$ is invertible.
Then, the mapping $\mathcal{F}_{\mathcal{U},\mathcal{V}}$ coincides with the linear fractional transformation $\widetilde{G}=\mathcal{T}_{\Phi_{\mathcal{U},\mathcal{V}}}(G)$, associated with the $2m\times 2m$ J-inner function

$$\Phi_{\mathcal{U},\mathcal{V}}(z)=\begin{bmatrix} M_u & 0 \\ 0 & M_v \end{bmatrix}+\begin{bmatrix} \alpha_u \\ \alpha_v \end{bmatrix}(K_v-K_u z)^{-1}\begin{bmatrix} \beta_u \\ \beta_v \end{bmatrix}^* J\begin{bmatrix} z\,I_m & 0 \\ 0 & I_m \end{bmatrix}. \tag{2.10}$$

Moreover $\Phi_{\mathcal{U},\mathcal{V}}(z)$ has McMillan degree 1, except if $|K_u|=|K_v|=1$.

In fact, this result is still valid and can be proved similarly [MO07, Prop. 1] when the mapping $\mathcal{F}_{\mathcal{U},\mathcal{V}}$ is associated with a more general recursion

$$\begin{bmatrix} \widetilde{D} & \widetilde{C} \\ \widetilde{B} & \widetilde{A} \end{bmatrix}=\begin{bmatrix} \mathcal{V} & 0 \\ 0 & I_{n-k} \end{bmatrix}\begin{bmatrix} I_k & 0 & 0 \\ 0 & D & C \\ 0 & B & A \end{bmatrix}\begin{bmatrix} U^* & 0 \\ 0 & I_{n-k} \end{bmatrix}. \tag{2.11}$$

This version can be used in the setting of Nudelman interpolation.

We then have to determine H in (2.6) and $\mathcal{U}$, $\mathcal{V}$ in Proposition (2.1.1) so that $\Theta_{w,u,v,H}$ coincides with $\Phi_{\mathcal{U},\mathcal{V}}$ given by (2.10).

First observe that there are many possibilities to do this. Since a left pole pair over $\mathbb{D}$ uniquely determines a stable J-inner function (Proposition 1.3.2) up to a right J-unitary constant factor, the left pole pairs (u,v) and $(\alpha_u K_v^{-1},\alpha_v K_v^{-1})$ associated with the pole $w=K_u K_v^{-1}$ must be similar. In particular K_v must be invertible and the orthonormality of the columns $\begin{bmatrix}\alpha_u & K_u\end{bmatrix}^T$ and $\begin{bmatrix}\alpha_v & K_v\end{bmatrix}^T$ completely fix them. Any unitary completion of these columns provides an admissible pair $(\mathcal{U},\mathcal{V})$, the matrix H being determined by evaluating $\Theta_{w,u,v,H}$ and $\Phi_{\mathcal{U},\mathcal{V}}$ at some point of the circle.

In [HOP06, Th. 6.4] a method specific to the tangential case is used. The left constant J-unitary factor H in (2.6) is searched a priori in the general form (1.33), $H(E)(P\oplus Q)$, where E is a contractive matrix and P and Q unitary. It is shown that E must be equal to $\bar{w}uv^*$ and choosing $P=Q=I_m$ the J-inner function $\Theta_{w,u,v,H}$ attains the form

$$\hat{\Theta}_{w,u,v}=H(uv^*)S_{w,u}(z)H(\bar{w}uv^*). \tag{2.12}$$

The corresponding matrices $\mathcal{U}$ and $\mathcal{V}$ are given, in terms of the interpolation data w, u, v, by

$$\mathcal{U} = \begin{bmatrix} \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}u & I_m - (1 + \frac{w\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}})uu^* \\ \frac{\overline{w}\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}} & \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}u^* \end{bmatrix}, \tag{2.13}$$

$$\mathcal{V} = \begin{bmatrix} \frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}v & I_m - (1 - \frac{\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}})\frac{vv^*}{\|v\|^2} \\ \frac{\sqrt{1-\|v\|^2}}{\sqrt{1-|w|^2\|v\|^2}} & -\frac{\sqrt{1-|w|^2}}{\sqrt{1-|w|^2\|v\|^2}}v^* \end{bmatrix}. \tag{2.14}$$

Note that if v is the null vector, then (2.8) is precisely the cascade decomposition (2.7).

A family of overlapping balanced canonical forms can thus be attached to a tangential Schur algorithm. The algorithm is that of section 1.3.4 for the lossless case and in which the LFT's J-inner symbols are chosen in the form (2.12): let $G_n = G$ and for $k = n, n-1, \ldots, 1$, let

$$v_k = G_k(1/\bar{w}_k)u_k.$$

- if $\|v_k\| < 1$, let
$$G_{k-1} = T^{-1}_{\hat{\Theta}_{w_k,u_k,v_k}}(G_k),$$
- if v_k has norm 1, then stop.

An atlas of charts for the manifold $\mathbb{L}^m_n$ of lossless $m \times m$ functions of McMillan degree n can be constructed as in section (1.3.5). A canonical form $\mathcal{C}_{\mathbf{w},\mathbf{u}}$ is attached to each chart of this atlas, i.e. to a sequence of interpolation points $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ and a sequence of interpolation directions $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, as described below

$$\mathcal{C}_{\mathbf{w},\mathbf{u}} : (v_1, v_2, \ldots, v_n, D_0) \mapsto R,$$

where the unitary realization matrix R, computed using (2.8), is a *product of unitary matrices* of size $(m+n) \times (m+n)$:

$$R = \Gamma_n\Gamma_{n-1}\cdots\Gamma_1\Gamma_0\Delta_1^T\Delta_2^T\cdots\Delta_n^T, \tag{2.15}$$

where for $k = 1, \ldots, n$:

$$\Gamma_k = \begin{bmatrix} I_{n-k} & 0 & 0 \\ 0 & V_k & 0 \\ 0 & 0 & I_{k-1} \end{bmatrix}, \tag{2.16}$$

$$\Delta_k = \begin{bmatrix} I_{n-k} & 0 & 0 \\ 0 & U_k & 0 \\ 0 & 0 & I_{k-1} \end{bmatrix}. \tag{2.17}$$

The unitary matrix blocks U_k and V_k are given by (2.13) and (2.14) respectively, in which the triple (w, u, v) must be replaced by (w_k, u_k, v_k), and furthermore

$$\Gamma_0 = \begin{bmatrix} I_n & 0 \\ 0 & D_0 \end{bmatrix}.$$

This computation of a realization matrix only involves products of unitary matrices and thus behaves nicely from a numerical point of view. These canonical forms have been used to represent lossless matrices for rational approximation purposes (see Chapter 3).

Given a balanced realization of a lossless function, it is not easy to decide whether it is or not in canonical form with respect to a chart. However, a realization (D, C, B, A) in Schur form, that is with a triangular dynamic matrix, happens to be canonical in a particular chart. This chart is easily determined, the interpolation points being the eigenvalues of A and the interpolation directions deduced from B. It is an adapted chart, in the sense that the parameters of (D, C, B, A) are null interpolation vectors. From a Schur algorithm point of view, it corresponds to the Potapov factorization (1.12)[MOHP02, MO07]. The Schur form provides a nice way to find an *adapted chart* for a given lossless functions which is clearly allowed by the freedom we have in choosing the interpolation points anywhere. However, in general, the realizations we obtain in this way have no structure. To get some structure, we must restrict the atlas. This will be the object of the next section.

2.2 Canonical forms with a pivot structure

In this section we report the results obtained in [PHO07] and [HOP09]. Choosing the interpolation point at ∞ ($w = 0$) and the interpolation directions as standard basis vectors yields a particular pivot structure for the sub-matrix $[B\ A]$. This structure, we call sub-diagonal, presents a lot of interests in itself and has been studied in [HOP09]. Contrary to the scalar case, the associated controllability matrix K may not have a particular pivot structure. Such a pivot structure is guaranteed if the matrix A has a staircase form [PHO07] which is obtained for some particular sequences of interpolation directions.

2.2.1 Subdiagonal canonical forms from the Schur algorithm

We consider the case $w_k = 0,\ k = 1, \ldots, n$. Hence, each balanced canonical form is determined by the choice of direction vectors. Each such balanced canonical form is given by (2.15) in which the blocks U_k and V_k attain a simpler form

$$U_k = \begin{bmatrix} u_k & I_m - u_k u_k^T \\ 0 & u_k^T \end{bmatrix}, \quad V_k = \begin{bmatrix} v_k & I_m - (1 - \sqrt{1 - \|v_k\|^2}\, \frac{v_k v_k^T}{\|v_k\|^2} \\ \sqrt{1 - \|v_k\|^2} & -v_k^T \end{bmatrix}.$$

It is important to note and not too difficult to see that the unitary matrix product

$$\Gamma = \Gamma_n \Gamma_{n-1} \cdots \Gamma_1 \Gamma_0 \tag{2.18}$$

in fact forms a positive m-upper Hessenberg matrix. An $(m + n) \times (m + n)$ matrix is called *positive m-upper Hessenberg* if the m-th subdiagonal only has positive entries and

the last $n-1$ subdiagonals are all zero. It also follows almost directly that if the direction vectors $u_1, \ldots, u_n$ are taken to be standard basis vectors, then the matrix product

$$\Delta^T = \Delta_1^T \Delta_2^T \cdots \Delta_n^T \tag{2.19}$$

yields a permutation matrix. Hence, in that case, the balanced realization matrix R is obtained as a column permutation of an unitary positive m-upper Hessenberg matrix. This structure is more precisely described using the concepts of pivot vectors and full pivot structure.

Following the original papers, we present the results for real systems. However, they can easily be transposed to the complex case. We denote by e_k the k-th standard basis vector in $\mathbb{R}^n$, whose entries are all zero except for the k-th entry which is 1.

Consider a vector $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$. The vector x is called a *pivot vector* with pivot at position k, $1 \leq k \leq n$, or pivot-k vector for short, if $x_k > 0$ and if the x_j's with $j > k$ are all zero. An $n \times r$ matrix M, $r \geq n$ is said to have a *full pivot structure* $J = \{j_1, j_2, \ldots, j_n\}$ if for each $k \in \{1, 2, \ldots, n\}$ it holds that column j_k of M is a pivot-k vector.

For example, the following matrix has a full pivot structure $J = \{7, 1, 5, 3, 6\}$.

$$M = \begin{bmatrix} * & * & * & * & * & * & + & * \\ + & * & * & * & * & * & 0 & * \\ 0 & * & * & * & + & * & 0 & * \\ 0 & * & + & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & + & 0 & * \end{bmatrix}$$

where $*$ denotes an arbitrary number and $+$ denotes a positive number.

Consider the partitioned matrix $[B|A]$ in $\mathbb{R}^{n \times (m+n)}$. We say this has a *subdiagonal pivot structure* if

(i) $[B|A]$ has a full pivot structure

(ii) the prescribed pivot columns of A have the property that a column with pivot at position k has column number $p_k < k$ in A (hence column number $j_k = m + p_k < m + k$ in $[B|A]$), for each $k \in \{1, 2, \ldots, n\}$.

The matrix

$$[B|A] = \left[\begin{array}{cc|cccc} + & * & * & * & * & * \\ 0 & * & + & * & * & * \\ 0 & + & 0 & * & * & * \\ 0 & 0 & 0 & * & + & * \end{array}\right]$$

has a subdiagonal pivot structure, while this one

$$[B|A] = \left[\begin{array}{cc|cccc} + & * & * & * & * & * \\ 0 & * & * & * & + & * \\ 0 & + & * & * & 0 & * \\ 0 & 0 & * & * & 0 & + \end{array}\right]$$

does not have a subdiagonal pivot structure.

Following the action of each permutation matrix Δ_k on each pivot in Γ, we get the following result.

Theorem 2.2.1 *[HOP09, Th. 4] Let R be given by (2.15) and let the direction vectors be standard basis vectors:*

$$u_{n-k+1} = e_{i(k)}, \quad i(k) \in \{1, \dots, m\}, \quad k = 1, \dots, n.$$

Then, the sub-matrix $[B|A]$ has a subdiagonal pivot structure completely determined by the sequence $i(1), i(2), \dots, i(n)$ as follows: for $k = 1, 2, \dots, n$,

(i) if for all $j < k$, $i(j) \neq i(k)$, then $j_k = i(k)$ and pivot-k is in B;

(ii) otherwise let $l = \sup\{j |\ j < k,\ i(j) = i(k)\}$, then $j_k = l + m$ and pivot-k is in one of the first $k - 1$ columns of A.

Conversely, to any specified subdiagonal pivot structure for the matrix $[B\ A]$ corresponds a sequence of direction vectors, more precisely

Theorem 2.2.2 *[HOP09, Th. 5] Let $J = (j_1, j_2, \dots, j_n)$ be a given subdiagonal pivot structure for $n \times (m + n)$ matrices. For each $k = 1, 2, \dots, n$, choose the direction vector $u_{n+1-k} = e_{i(k)}$ in the Schur algorithm by induction: $i(1) := j_1$, and for $k = 2, \dots, n$*

(i) if $j_k \leq m$, then $i(k) := j_k$,

(ii) if $m < j_k \leq m + n$ define $p_k := j_k - m < k$ (subdiagonal structure); then $i(k) := i(p_k)$.

For any choice of the Schur parameter vectors $v_1, v_2, \dots, v_n$ (all of norm < 1) and for any choice of the unitary matrix D_0, consider the $(m + n) \times (m + n)$ unitary realization matrix R given by (2.15). Then, the sub-matrix $[B|A]$ possesses the subdiagonal pivot structure J.

We have got a one-to-one correspondance between the set of subdiagonal pivot structure for $n \times (m + n)$ matrices and the set of sequences of n standard m-vectors. If we restrict the infinite atlas of section 2.1 to the charts associated with sequences of null interpolation points and standard interpolation direction, it is not difficult to see that we get a finite sub-atlas. In this atlas, a local canonical form is attached to a subdiagonal pivot structure. This atlas possesses m^n charts (see 2.2.1).

2.2.2 Subdiagonal canonical forms under state isometry

The interest of subdiagonal pivot structure goes beyond the parametrization of lossless functions in connection with the Schur algorithm. In many examples, a normalization of the system by a state isomorphism (1.19) leaves the freedom of applying an orthogonal transformation. The system can be normalized in the sense that its finite controllability (resp. observability) matrix has orthonormal rows, or to be input (resp. output) normal, balanced, LQG-balanced, etc. (see [HOP09],[30]). In all these cases it makes sense to look for canonical forms under *state isometry*, a state isomorphism in

which the change-of-basis matrix is orthogonal. The subdiagonal pivot structures correspond to local canonical forms under state isometry, and these local canonical forms are covering all cases. These results hold true both for discrete-time and continuous systems.

Theorem 2.2.3 *[HOP09, Th. 3] Suppose that $[B|A]$ and $[QB|QAQ^*]$ have the same subdiagonal pivot structure $J = \{j_1, j_2, ..., j_n\}$ and that Q is orthogonal. Then $Q = I$.*

To show that every controllable pair (A, B) can be mapped to a subdiagonal pivot structure by an appropriate state isometry, a recursive algorithm "which cannot get stuck, if controllability holds" has been presented in [HOP09]. It was largely inspired by the discrete-time lossless case.
First step: choose a non-zero column j_1 in B (existence is ensured by controllability). Find a unitary matrix Q_1 which maps it into a pivot-1 vector and perform the corresponding state isometry. We get a matrix $[B_1|A_1]$ with a pivot-1 vector in column j_1.
Recursion step: suppose that a state-isometry Q_i has been determined, so that

$$[B_i|A_i] = [Q_i B_{i-1}, Q_i A_{i-1} Q_i^*]$$

has a pivot-k in column j_k with the subdiagonal property $j_k < m + k$, for $k = 1, 2, \ldots, i$. Then, consider the sub-matrix formed by the $n - i$ last rows of $[B_i|A_i]$. Choose a non-zero column j_{i+1} (existence is again ensured by controllability). Find a unitary matrix $\tilde{Q}_{i+1}$ which maps it into a pivot-1 vector, then $Q_{i+1} = I_i \oplus \tilde{Q}_{i+1}$ maps column j_{i+1} of $[B_i|A_i]$ into a pivot-$(i+1)$ vector. Then $[Q_{i+1}B_i|Q_{i+1}A_iQ_{i+1}^*]$ possesses a pivot-k in column j_k for $k = 1, 2, \ldots, i+1$. Because of the subdiagonal structure, right multiplication by Q_{i+1}^* cannot destroy the pivot structure.
Applying the recursion step for each $k = 1, 2, \ldots, n$, we can construct a state isometry that brings $[B|A]$ into a subdiagonal pivot structure.

This algorithm provides an effective method to find a chart for a given lossless system in the atlas of section 2.2.1. This also gives a controllability test. An interesting practical question is: how to chose the non-zero column at each step? Take the one with the largest norm? For numerical reasons, small pivots (the positive kth entry of a pivot-k) must be avoided, since they correspond to a form which is poorly controllable. However, we have no selection strategy which controls the size of the pivots. Note that this algorithm is rather insensitive to small perturbations.

We have got a set of overlapping local canonical forms for input normal pairs which works both in discrete-time and in continuous-time. In discrete-time, the tangential Schur algortihm with interpolation condition suitably chosen provides parameters for these forms. In continuous-time, subdiagonal pivot structures can be parametrized following the ideas developed in [30] for the more constrained *staircase* pivot structure, which ensures that also the controllability matrix presents a pivot structure.

2.2.3 Staircase canonical forms

We come back to the setting of discrete-time lossless functions (section 2.2.1). We are now looking for balanced canonical forms which ensure a full pivot structure in the

reachability matrix. It is not difficult to see that if the realization matrix R itself happens to be positive m-upper Hessenberg, then the first n columns of the corresponding controllability matrix $K = [B|AB|A^2B|\ldots]$ also form a positive upper triangular matrix. In the SISO case, you get in this way a *global* canonical form. In the MIMO case, you only get a *local* canonical form, associated with the full pivot structure $J = \{1, 2, \ldots, n\}$ for $[B|A]$. The object of this section is to complete it into a family of overlapping local canonical forms, each of these forms being associated with a pivot structure for the controllability matrix K.

The question is thus: which subdiagonal pivot structures for the $n \times (m+n)$ matrix $[B|A]$ imply that the associated controllability matrix contains a column with a pivot at position k for each $k = 1, 2, \ldots, n$? The answer is that the subdiagonal pivot structure must be of the following particular form: if the columns of B contain p_B pivots then the remaining $p_A = n - p_B$ pivots have to be located in the first p_A columns of A with increasing pivot positions. We shall say that A has a *staircase structure*. A pivot structure of this form will be called an *admissible* pivot structure. Clearly, an *admissible* pivot structure J for $[B, A]$ is totally determined by the induced pivot structure for B.

For several purposes, the induced pivot structures for A and B are more conveniently described in terms of the column-oriented description $Q = \{q_1, \ldots, q_{m+n}\}$ for $[B, A]$. Note that each column ℓ is a pivot-q_ℓ vector, where 'a pivot-0 vector' is synonymous to 'not a pivot vector'. For the matrix A it holds that the associated column-oriented pivot structure $S = \{s_1, \ldots, s_n\}$ satisfies $s_k = q_{m+k}$ for all $k = 1, \ldots, n$.

Consider the full pivot structure $J = \{3, 1, 5, 6, 4, 7\}$ for the $6 \times (4+6)$ partitioned matrix

$$[B, A] = \left[\begin{array}{cccc|cccccc} * & * & + & * & * & * & * & * & * & * \\ + & * & 0 & * & * & * & * & * & * & * \\ 0 & * & 0 & * & + & * & * & * & * & * \\ 0 & * & 0 & * & 0 & + & * & * & * & * \\ 0 & * & 0 & + & 0 & 0 & * & * & * & * \\ 0 & * & 0 & 0 & 0 & 0 & + & * & * & * \end{array}\right]. \tag{2.20}$$

The column-oriented description is given by $Q = \{2, 0, 1, 5, 3, 4, 6, 0, 0, 0\}$ and $S = \{3, 4, 6, 0, 0, 0\}$ for A.

If v is a pivot-k vector, then the staircase structure of A implies that $w = Av$ is a pivot-s_k vector. For this reason, the map $S : k \mapsto s(k) = s_k$, $k = 1, \ldots, n$ is called the *successor* function (for convenience we also define $S(0) = 0$). If v is a pivot-k vector, then the sequence of pivot positions in $v, Av, A^2v, A^3v, \ldots$ is given by $k, S(k), S^2(k), S^3(k), \ldots$ In this way, an admissible pivot structure J for $[B, A]$ generates a uniquely specified full pivot structure $\widetilde{J}$ for the controllability matrix K. To visualize this, it is helpful to introduce an $m \times n$ array $Y = (y_{i,j})$, defined as follows: entry $y_{i,j}$ denotes the pivot position of vector i in the j-th block $A^{j-1}B$ of K. This array will be called an *admissible Young diagram*.

The reachability matrix $K = [B, AB, A^2B, \ldots]$ associated with (2.20) is of the form:

$$K = \left[\begin{array}{cccc|cccc|cccc|cc} * & * & + & * & * & * & * & * & * & * & * & * & \ldots & \ldots \\ + & * & 0 & * & * & * & * & * & * & * & * & * & \ldots & \ldots \\ 0 & * & 0 & * & * & * & + & * & * & * & * & * & \ldots & \ldots \\ 0 & * & 0 & * & + & * & 0 & * & * & * & * & * & \ldots & \ldots \\ 0 & * & 0 & + & 0 & * & 0 & * & * & * & * & * & \ldots & \ldots \\ 0 & * & 0 & 0 & 0 & * & 0 & * & * & * & + & * & \ldots & \ldots \end{array}\right]$$

The corresponding 4×6 array Y is:

$$Y = \begin{array}{|c|c|c|c|c|c|} \hline 2 & 4 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 3 & 6 & 0 & 0 & 0 \\ \hline 5 & 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

The first column of Y specifies the pivot structure of B. The other entries of Y satisfy the rule $y_{i,j+1} = S(y_{i,j})$. The admissible Young diagram displays the pivot structure of $[B, A]$ as well as that of K.

Theorem 2.2.4 *[PHO07, Th. 4.1] Let J be an admissible full pivot structure for the block-partitioned matrix $[B, A]$, with a column-oriented description $Q = \{q_1, \ldots, q_{m+n}\}$ and the successor function S given by $S = \{s_1, \ldots, s_n\} = \{q_{m+1}, \ldots, q_{m+n}\}$ and $S(0) = 0$. Then J induces a full pivot structure $\widetilde{J}$ for the (finite) controllability matrix*

$$K = [B, AB, \ldots, A^{n-1}B]$$

which is specified in terms of the $m \times n$ array Y associated with Q as follows:
(i) $y_{i,1} = q_i$ for $i = 1, \ldots, m$;
(ii) $y_{i,j+1} = S(y_{i,j})$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n-1$.

Conversely, for every non-admissible full pivot structure J there exists an $n \times (m+n)$ matrix $[B, A]$ having the full pivot structure J, for which $K = [B, AB, A^2B, \ldots]$ does not have a full pivot structure [PHO07, Th.3.4].

This study is strongly connected to the concepts of nice selection and dynamical indices.
The set $\mathcal{D}(m, n)$ is defined as the set of all multi-indices

$$\mathcal{D}(m, n) = \{(d_1, d_2, \ldots, d_m) \in \mathbb{N}^m \text{ such that } d_1 + d_2 + \ldots + d_m = n\}.$$

A selection of n columns from an $n \times nm$ controllability matrix $K = [B, AB, \ldots, A^{n-1}B]$ is called a *nice selection* if there exists a multi-index $d \in \mathcal{D}(m, n)$ for which the selected set of columns is given by

$$\{A^{j-1}Be_i \,|\, j \in \{1, 2, \ldots, d_i\} \text{ for } i = 1, 2, \ldots, m\}.$$

The full pivot structure $\tilde{J}$ for K, obtained from an admissible pivot structure for $[B|A]$, constitutes a nice selection of columns.

The concept of Young diagram, a left-aligned $m \times n$ binary array corresponding to a nice selection with an associated vector of dynamical indices $d = (d_1, \ldots, d_m)$ is used in [26] for example. We define a *numbered Young diagram* as a Young diagram in which the unit entries are replaced by the numbers $1, 2, \ldots, n$ in some arbitrary order, so that they all occur exactly once. The *admissible Young diagrams* associated with a full pivot structure $\tilde{J}$ for K obtained from an admissible pivot structure for $[B|A]$ is a numbered Young diagram. The number of nonzero entries d_i of the i-th row of the admissible Young diagram are the dynamical indices corresponding to the nice selection $\tilde{J}$.

Among all the numbered Young diagrams that can be obtained from a given vector of dynamical indices, we may characterized the *admissible Young diagrams*. This characterization is in terms of the *right-aligned* version $Y_r = (y^r_{i,j})$ of the numbered Young diagram Y obtained by shifting the nonzero entries of each row $n - d_i$ positions to the right, $y^r_{i,j} := y_{i,j+d_i-n}$.

Proposition 2.2.1 *[PHO07, Prop. 4.5.] An $m \times n$ numbered Young diagram corresponding to a nice selection with a vector of dynamical indices $d = (d_1, \ldots, d_m)$ is admissible if and only if there exists an $m \times m$ permutation matrix Π for the associated right-aligned array Y_r such that the nm-vector $vec(\Pi Y_r) = \left((\Pi Y_r e_1)^T, (\Pi Y_r e_2)^T, \ldots, (\Pi Y_r e_n)^T\right)^T \in \mathbb{R}^{nm}$ obtained by stacking the n columns of the array ΠY_r, has the property that if the zero entries are deleted then the n-vector $(1, 2, 3, \ldots, n)^T$ is obtained.*

For the previous example, we get

$$Y = \begin{array}{|c|c|c|c|c|c|}\hline 2&4&0&0&0&0\\\hline 0&0&0&0&0&0\\\hline 1&3&6&0&0&0\\\hline 5&0&0&0&0&0\\\hline\end{array} \quad Y_r = \begin{array}{|c|c|c|c|c|c|}\hline 0&0&0&0&2&4\\\hline 0&0&0&0&0&0\\\hline 0&0&0&1&3&6\\\hline 0&0&0&0&0&5\\\hline\end{array} \quad \Pi = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}$$

The permutation Π makes that each column of the right-aligned matrix has the property that the nonzero entries in the column form an increasing sequence.

Conversely, starting from a given vector of dynamical indices d and an arbitrary choice of Π permuting the nonzero rows of the associated Young diagram, completely determines an unique admissible Young diagram.

Restricting the atlas of subdiagonal forms of section 2.2.1 to the chart corresponding to admissible pivot structures, we get a sub-atlas. Indeed, this set of charts is covering $\mathbb{L}^m_n$ in the sense that for each input-normal $[B, A]$ with A asymptotically stable, there exists an admissible numbered Young diagram Y and an orthogonal matrix Q such that $[QB, QAQ^T]$ has the admissible pivot structure associated with Y. Each staircase canonical form can thus be associated with a recursive Schur algorithm with interpolation points at 0 and interpolation direction selected among standard basis vectors following the rule : for each $k = 1, 2, \ldots, n$, $u_{n+k-1} = e_{i(k)}$, where $(i(k), j(k))$ denotes the unique pair of indices such that $y_{i(k),j(k)} = k$ in the admissible Young diagram Y.

This atlas is not minimal, in the sense that no further local canonical form can be left out without losing the property of covering the family. To obtain a minimal sub-atlas,

one has to choose one of the local canonical forms for each $d \in \mathcal{D}(m,n)$. One possible choice is the unique permutation for which the permuted dynamical indices form a non-increasing sequence, while the order of the rows which have the same dynamical index is kept the same. With hindsight, one can say that this particular choice to obtain a minimal atlas for continuous-time lossless systems was used in [30].

The number of charts in the various atlases are

subdiagonal forms	straicase forms	minimal atlas
m^n	$\sum_{\ell=1}^{\min\{m,n\}} \ell! \begin{pmatrix} m \\ \ell \end{pmatrix} \begin{pmatrix} n-1 \\ \ell-1 \end{pmatrix}$	$\begin{pmatrix} m+n-1 \\ m-1 \end{pmatrix}$

To demonstrate the structure of the charts that constitute the various atlases, a simple example ($m = 2, n = 3$) is presented in Figure 2.1. The charts marked with a * form a minimal atlas. A Young diagram representation is also used for subdiagonal but non-staircase forms. In this case, the numbering in the Young diagram specifies the pivot structure of $[B\ A]$ according to Theorem 2.2.4 but gives no information on the controllability matrix.

2.3 Perspectives

In this chapter, we have described (structured) canonical forms which can be parametrized using Schur vectors. This approach combines the practical advantages of state-space representations with the efficiency and simplicity of Schur analysis. It has a wide range of applications, most of which have not yet been explored.

In [31] the structure of the canonical form was used to show that the set of lossless functions of order $\leq n$ with the topology induced by H_2 has the structure of an $2n+1$ hypersphere S^{2n+1}. The subspace of lossless systems of order $\leq k < n$ appears as an embedded hypersphere. Looking at the Hessenberg form (2.2) parametrized by the sequence of Schur parameters, the situation can be analyzed as follows: if the norm of some Schur parameter, say γ_j, goes to 1, the corresponding pivot κ_j goes to 0 and the dynamic matrix A then presents a diagonal block-structure with a unitary lower block. The realization is no more minimal, but represents a $n-j$ order (still lossless) system. The lossless systems of order $< n$ appear as boundary points in $\mathbb{L}_n^1$ and the Hessenberg structure can be used for model reduction purposes.

In the MIMO case, because of the non-trivial manifold structure of $\mathbb{L}_n^m$, a pivot in a local subdiagonal form can be zero for two completely different reasons: either the local chart is not adapted or the underlying system fails to be controllable which is the boundary case. To say it differently, the boundary of a chart (local subdiagonal chart) is mostly composed by systems of order n, for which another chart should be used, and some of order $< n$, which are boundary points of the manifold. It would be very helpful, both for identification and model reduction purposes, to better understand the topological structure of the set of $m \times m$ lossless systems of order $\leq n$. Subdiagonal forms could bring some insights in this study. They could also provide simple methods for model order reduction for lossless systems as well as other classes of systems,

Chart	$[B\|A]$	Pivot structure	Young Diagram	Canonical form
(e_1, e_1, e_1)	$\left[\begin{array}{cc\|ccc} + & * & * & * & * \\ 0 & * & + & * & * \\ 0 & * & 0 & + & * \end{array}\right]$	$\{1,3,4\}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 3 \\ \hline \end{array}$	staircase*
(e_1, e_1, e_2)	$\left[\begin{array}{cc\|ccc} * & + & * & * & * \\ + & 0 & * & * & * \\ 0 & 0 & * & + & * \end{array}\right]$	$\{2,1,4\}$	$\begin{array}{\|c\|c\|c\|}\hline 2 & 3 & \\ \hline 1 & & \\ \hline \end{array}$	subdiagonal
(e_1, e_2, e_1)	$\left[\begin{array}{cc\|ccc} + & * & * & * & * \\ 0 & + & * & * & * \\ 0 & 0 & + & * & * \end{array}\right]$	$\{1,2,3\}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 3 & \\ \hline 2 & & \\ \hline \end{array}$	staircase
(e_2, e_1, e_1)	$\left[\begin{array}{cc\|ccc} + & * & * & * & * \\ 0 & * & + & * & * \\ 0 & + & 0 & * & * \end{array}\right]$	$\{1,3,2\}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & \\ \hline 3 & & \\ \hline \end{array}$	staircase*
(e_2, e_2, e_1)	$\left[\begin{array}{cc\|ccc} + & * & * & * & * \\ 0 & + & * & * & * \\ 0 & 0 & * & + & * \end{array}\right]$	$\{1,2,4\}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & & \\ \hline 2 & 3 & \\ \hline \end{array}$	subdiagonal
(e_1, e_2, e_2)	$\left[\begin{array}{cc\|ccc} * & + & * & * & * \\ * & 0 & + & * & * \\ + & 0 & 0 & * & * \end{array}\right]$	$\{2,3,1\}$	$\begin{array}{\|c\|c\|c\|}\hline 3 & & \\ \hline 1 & 2 & \\ \hline \end{array}$	staircase*
(e_2, e_1, e_2)	$\left[\begin{array}{cc\|ccc} * & + & * & * & * \\ + & 0 & * & * & * \\ 0 & 0 & + & * & * \end{array}\right]$	$\{2,1,3\}$	$\begin{array}{\|c\|c\|c\|}\hline 2 & & \\ \hline 1 & 3 & \\ \hline \end{array}$	staircase
(e_2, e_2, e_2)	$\left[\begin{array}{cc\|ccc} * & + & * & * & * \\ * & 0 & + & * & * \\ * & 0 & 0 & + & * \end{array}\right]$	$\{2,3,4\}$	$\begin{array}{\|c\|c\|c\|}\hline & & \\ \hline 1 & 2 & 3 \\ \hline \end{array}$	staircase*

Figure 2.1: Subdiagonal canonical forms for $m = 2$, $n = 3$.

based on this observation: if $[B|A]$ is in subdiagonal pivot form and we truncate the last $n-i$ rows and columns, we again have a controllable pair in subdiagonal pivot form.

One of the main practical reasons for studying atlases of charts was that of iterative optimization algorithms using differential tools in the MIMO case. Even if our preference bears on infinite and flexible atlases (see Chapter 3), a finite atlas of structured canonical forms could present some interest. The issue of selecting a better chart when switching becomes necessary and of monitoring the conditioning of a chart are fundamental in this connection. For the atlas of subdiagonal canonical forms, the algorithm of section 2.2.2 provides a chart for a given system. However, we don't know what could be a good choice of the non-zero vector (the pivot column) at each step of the algorithm. This essential question may have different answers from a model reduction or an optimization perspective. A chart selection algorithm is missing for the atlas of staircase forms and will be highly desirable.

Subdiagonal pivot structures could be even more interesting in the continuous-time setting. Real world systems are often described by means of realization with a particular structure in which the physical parameters are displayed (see section 3.4). In some applications, subdiagonal forms could offer an easy connection between optimization parameters and physical parameters. In the SISO case, the single subdiagonal canonical form is the well-known Schwarz-Ober form in which the dynamic matrix A is tridiagonal. It can be parametrized by the n positive pivots [52]. It is possible to interpret these parameters as interpolation values in a recursive Schur algorithm, but the interpolations values must be taken at ∞ which belongs to the analyticity boundary in continuous-time [OHP08]. Multivariable lossless systems can analogously be parametrized using interpolation data on the imaginary line [4, Chap. 21]. However, the overlapping subdiagonal canonical forms are not easily recovered from these parametrizations. It must be noticed that vanishing moments, diagonal Markov parameters, can be interpreted in term of boundary interpolation conditions. Boundary interpolation sounds a promising tool and its use in the representation of both continuous-time and discrete-time systems is under investigation.

Chapter 3

Rational H_2 approximation and lossless mutual encoding

Rational approximation was the first topic I studied some twenty years ago. It is at the origin of my interest for lossless functions and their parametrizations.

Since finite order LTI systems and their rational transfer functions are the most used models in system theory, rational approximation is at the heart of modeling problems. The challenge is to select a model that is close enough to a physical system and yet simple enough to be studied analytically. Modeling is thus at the origin of a fertile interaction between rational approximation methods and system theory. A great number of methods have been proposed in the literature, that divide into two main groups: projection methods and optimization methods (see [12] for a nice oriented historical survey).

Among these methods, rational approximation in the Hardy space $\bar{H}_2$ presents a number of interesting features. Assuming a transfer function belongs to $\bar{H}_2$ corresponds to a stability condition for the underlying physical system: a bounded energy input produces a bounded output. The Hardy space $\bar{H}_2$ possesses a very rich structure which combines analyticity properties and an Hilbert space framework. However, $\bar{H}_2$ rational approximation is a difficult non-linear problem due to the complexity of the set approximants (rational matrices of fixed order) and the existence of many local minima.

A specific approach has been developed at INRIA to cope with this problem, which is based on the following points

- the optimization range is reduced to the set of lossless functions,

- an atlas of chart is used to parametrize the set of lossless functions.

The exposition here is definitely application oriented, so that the emphasis will be put on the effective implementation of the method. It is based on state-space formulas and a suitable atlas of chart which was presented in [MO07].

3.1 The rational approximation problem

The L^2-norm of a matrix-valued function $F(z)$ whose entries belong to $L^2(\mathbb{T})$ is given by

$$\|F\|_2^2 = \frac{1}{2\pi}\mathrm{Tr}\int_0^{2\pi} F(e^{it})F(e^{it})^* dt.$$

The rational approximation problem we consider is, given a matrix-valued function $F(z) \in \bar{H}_2^{p\times m}$, to minimize the L^2 distance to the set of rational functions in $\bar{H}_2^{p\times m}$ of McMillan degree less than or equal to n.

First remark that any solution H must satisfy $H(\infty) = F(\infty)$. Thus, we may restrict our study to the case of strictly proper transfer functions, that is to the space $(H_2^{p\times m})^\perp$.

A number of qualitative results are available which assert that the problem is well-posed and pave the way to convergent algorithms. It was proved in [5] that *the global minimum of the L^2 criterion does exist*, as well as the *normality property*: if $F(z)$ is not of McMillan degree strictly less than n, then the best approximant $H(z)$ at order less than or equal to n has effective order n, so that the problem can now be stated as:

Rational approximation problem. Given $F(z) \in (H_2^{p\times m})^\perp$, find $\hat{H}(z)$ such that

$$\hat{H} = \mathrm{argmin}_{H\in\Xi_n}\|F - H\|_2^2 \tag{3.1}$$

where

$$\Xi_n = \{H \in (H_2^{p\times m})^\perp, \deg H = n\}$$

is the set of rational strictly proper stable transfer functions of exact degree n.

The following *consistency* result should also be mentioned: if $F(z)$ has McMillan degree n, then the only local minima (and even the only critical point) of the L^2 criterion is $F(z)$ itself [BO98].

The present approach was first proposed in the SISO case [BCO91] and then in the MIMO case [FO98]. The first step is the elimination of the linear variable by means of the Douglas-Shapiro-Shields factorization (Theorem 1.1.5). Any rational matrix function $H(z) \in (H_2^{p\times m})^\perp$ can be written in the form

$$H(z) = C(z)G(z),$$

where $G(z)$ is lossless of McMillan degree n and $C(z) \in H_2^{p\times m}$. Any approximant $H(z)$ of $F(z)$ thus minimizes the distance from $F(z)$ to the vector space

$$\mathrm{Vec}(G) = \{H \in (H_2^{p\times m})^\perp;\quad H = C\,G,\quad C \in H_2^{p\times m}\}.$$

The Projection theorem in a Hilbert space asserts that $H(z)$ is completely determined by $G(z)$ as the orthogonal projection of $F(z)$ onto $\mathrm{Vec}(G)$. Equivalently, $C(z)$ is the orthogonal projection of $F(z)G^\sharp(z)$ onto $H_2^{p\times m}$, $C(z) = \pi_+(FG^\sharp)$. In least-square optimization and using a state-space formulation, this elimination step is classical and

known under the name of Separable Least Square. It presents some important advantages: the dimension of the parameters space is reduced and mostly, lossless functions enter the picture. The rational approximation problem is then to minimize the criterion

$$\psi_n : \ G \mapsto \|F - \pi_+(FG^\sharp)G\|^2 \tag{3.2}$$

over the set $\mathbb{L}_n^m \setminus \mathbb{U}_m$ the right quotient of the set of $m \times m$ lossless functions of McMillan degree n by unitary constant matrices.

The second step is to find a clever parametrization of the set $\mathbb{L}_n^m$ and its right quotient $\mathbb{L}_n^m \setminus \mathbb{U}_m$. In order to use differential tools (as a gradient algorithm) to solve this optimization problem, an atlas of charts happens to be the desirable representation. Over the years, different atlases of charts have been experimented and implemented. The first one was the atlas derived form the tangential Schur algorithm constructed in [1]. A polynomial representation of lossless matrices were used as described in [FO98] and implemented in the software Hyperion [29]. However, a state-space representation is often preferred for computations and the balanced canonical forms of section 2.1 were used in the software RARL2 [MO04]. However, the natural framework for this atlas is that of complex valued functions. In particular, an adapted chart, that is a chart centered at a given lossless function requires complex interpolation points, its (possibly complex) poles (see section 2.1).

However, physical systems are often real-valued and their transfer functions $T(z)$ are *real*, that is, they satisfy the relation $\overline{T(z)} = T(\bar{z})$. It must be mentioned that the best complex approximation of a real function may fail to be real. For example the function $f(z) = 1/z^3 - 1/z$ admits three minima: a real one and two complex, which achieve the best relative error. In the next section, we propose an atlas which better suits the representation of real lossless functions. It is based on the more general Nudelman interpolation problem (see section 1.3.3).

3.2 Nudelman interpolation and lossless mutual encoding

We first built an atlas of chart for the set $\mathbb{L}_n^p$. To this end, we consider the interpolation problem (see section 1.3.3) now stated for lossless functions: given an observable pair (U, W), where W is $n \times n$ stable, find all the lossless function $G(z) \in \mathbb{L}_n^m$ satisfying the interpolation condition

$$\frac{1}{2i\pi} \int_{\mathbb{T}} (zI - W^*)^{-1} U^* G^\sharp(z) \, dz = V^*. \tag{3.3}$$

Note that two equivalent triples (W, U, V) and $(TWT^{-1}, UT^{-1}, VT^{-1})$ yield the same interpolation condition (3.3). Thus, we may assume the observable pair (U, W), W is stable, to be output normal :

$$U^*U + W^*W = I. \tag{3.4}$$

By Theorem 1.3.1, if the solution P to the Stein equation (1.39) is positive definite, then the set of all lossless solutions $G(z)$ of McMillan degree n of (3.3) is given by

$$\{G(z) \in \mathbb{L}_n^m;\ G = \mathcal{T}_{\Theta_{W,U,V}H}(G_0),\ G_0 \text{ constant unitary matrix}\}$$

where $\Theta_{W,U,V}$ is the J-inner matrix (1.40) with $\nu = 1$ and H is a right constant J-unitary multiplier, that can be freely chosen.

A chart of $\mathbb{L}_n^m$ is thus attached to an output normal pair (U, W), and a lossless function $G(z)$ belongs to the domain $\mathcal{V}_{U,W}$ of this chart if and only if the solution P to the Stein equation (1.39) is positive definite. The coordinate map is

$$\phi_{U,W} :\ G \in \mathcal{V}_{U,W} \to (G_0, V).$$

The matrix H may depend on the interpolation triple (W, U, V). In this case we shall require $H(W, U, V)$ to be differentiable with respect to V and to satisfy the relation

$$\begin{bmatrix} \Lambda & 0 \\ 0 & \Pi \end{bmatrix} H(W, U, V) \begin{bmatrix} \Lambda^* & 0 \\ 0 & \Pi^* \end{bmatrix} = H(W, \Lambda U, \Pi V) \tag{3.5}$$

for any unitary matrices Λ and Π. This equality ensures that the quotient $\mathbb{L}_n^m \setminus \mathbb{U}_m$ can be perform within each chart by fixing G_0.
The matrix H and unitary matrices $\mathcal{U}$ and $\mathcal{V}$ can be constructed following the approach of section 2.1 so that a balanced canonical form is computed by the multiplicative formula (2.11) in which $d = n$. However, in the present case the balanced canonical form can be obtained directly and nicely interpreted. The interpolation value V in (3.3) as a very simple state-space formulation. An analog formula has also been derived in continuous-time in [28].

Proposition 3.2.1 *[MO07, Prop. 4] Let (U, W) be an output normal pair, W stable, and let $G(z) = D + C(zI_n - A)^{-1}B$ be a balanced realization of $G(z) \in \mathbb{L}_n^m$. Let Q be the unique solution to the Stein equation*

$$Q - AQW = BU \tag{3.6}$$

Then, the interpolation value V in (3.3) and the associated solution P to the Stein equation (1.39) are given by

$$V = DU + CQW \tag{3.7}$$

$$P = Q^*Q. \tag{3.8}$$

Formula (3.7) is easily established and can be rewritten with (3.6) in a matrix form

$$\begin{bmatrix} D & C \\ B & A \end{bmatrix} \begin{bmatrix} U \\ QW \end{bmatrix} = \begin{bmatrix} V \\ Q \end{bmatrix}. \tag{3.9}$$

The realization of $G(z)$ being balanced by assumption, we get

$$U^*U + W^*QQ^*W = V^*V + Q^*Q,$$

so that $Q^*Q = P$, the unique solution of (1.39).

Under the assumption $P > 0$ and thus $K = V^*V + P$ positive definite, we would like to reverse the process and specify a unique way to compute a balanced realization of $G(z)$ from the interpolation matrix V and a constant unitary matrix (the analog of G_0 in the LFT representation).

We first fix the square root Q of P to be the (uniquely defined) Hermitian square root, $Q = P^{1/2}$, and we use it to normalize the interpolation data

$$(W, U, V) \mapsto (\tilde{W}, \tilde{U}, \tilde{V}) = (QWQ^{-1}, UQ^{-1}, VQ^{-1}) \tag{3.10}$$

Then, equation (3.9) yields

$$\begin{bmatrix} \tilde{V} \\ I \end{bmatrix}^* \begin{bmatrix} D & C \\ B & A \end{bmatrix} \begin{bmatrix} \tilde{U} \\ \tilde{W} \end{bmatrix} = \tilde{V}^*\tilde{V} + I.$$

Using the Hermitian square root $K^{1/2}$ of $K = \tilde{V}^*\tilde{V} + I$, we get

$$\begin{bmatrix} \tilde{V}K^{-1/2} \\ K^{-1/2} \end{bmatrix}^* \begin{bmatrix} D & C \\ B & A \end{bmatrix} \begin{bmatrix} \tilde{U}K^{-1/2} \\ \tilde{W}K^{-1/2} \end{bmatrix} = I.$$

For any unitary completions $\mathcal{U}$, $\mathcal{V}$ of these orthonormal columns,

$$\mathcal{U} = \begin{bmatrix} \tilde{U}K^{-1/2} & * \\ \tilde{W}K^{-1/2} & * \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} \tilde{V}K^{-1/2} & * \\ K^{-1/2} & * \end{bmatrix}$$

we have that

$$\mathcal{V}^* \begin{bmatrix} D & C \\ B & A \end{bmatrix} \mathcal{U} = \begin{bmatrix} I & 0 \\ 0 & D_0 \end{bmatrix} \tag{3.11}$$

for some unitary matrix D_0.

Now, the unitary completions $\mathcal{U}$ and $\mathcal{V}$ have to be specified. In [MO07] the matrix $\mathcal{V}$ is chosen accordingly to Proposition 1.3.2 in which $J = I$ and $\nu = -1$:

$$\mathcal{V} = \begin{bmatrix} \tilde{V}(I + \tilde{V}^*\tilde{V})^{-1/2} & (I + \tilde{V}\tilde{V}^*)^{-1/2} \\ (I + \tilde{V}^*\tilde{V})^{-1/2} & -\tilde{V}^*(I + \tilde{V}\tilde{V}^*)^{-1/2} \end{bmatrix}. \tag{3.12}$$

For the matrix $\mathcal{U}$, another method is used which makes use of a Cholesky factorization, and has been implemented in the software RARL2.
First remark that when $D_0 = I$ and $V = 0$ we get

$$\begin{bmatrix} D & C \\ B & A \end{bmatrix} = \begin{bmatrix} * & * \\ U^* & W^* \end{bmatrix},$$

whence the idea to attach a chart to a unitary balanced realization

$$\Omega = (W^*, U^*, Y^*, X^*) \tag{3.13}$$

rather than just an output normal pair (W, U).

The matrix $\mathcal{U}$ is then computed from Ω as follows:

- perform a state isomorphism with matrix $Q = P^{1/2}$:

$$(W, Y, U, X) \longrightarrow (\tilde{W}, \tilde{Y}, \tilde{U}, X) = (QWQ^{-1}, QY, UQ^{-1}, X).$$

- compute a Cholesky factorization of the matrix

$$\begin{bmatrix} K & L \\ L^* & N \end{bmatrix} = \begin{bmatrix} \tilde{U} & X \\ \tilde{W} & \tilde{Y} \end{bmatrix}^* \begin{bmatrix} \tilde{U} & X \\ \tilde{W} & \tilde{Y} \end{bmatrix}. \tag{3.14}$$

  using the well-known formula [18, Sec. 0.2], we get

$$\begin{aligned} \begin{bmatrix} K & L \\ L^* & N \end{bmatrix} &= \begin{bmatrix} I & 0 \\ L^*K^{-1} & I \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & Z^{-1} \end{bmatrix} \begin{bmatrix} I & K^{-1}L \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} K^{1/2} & K^{-1/2}L \\ 0 & Z^{-1/2} \end{bmatrix}^* \begin{bmatrix} K^{1/2} & K^{-1/2}L \\ 0 & Z^{-1/2} \end{bmatrix} \end{aligned}$$

  where $Z^{-1} = N - L^*K^{-1}L$ can be computed by inverting the matrix (3.14) [18, Formula (0.8)].

The matrix

$$\mathcal{U} = \begin{bmatrix} \tilde{U} & X \\ \tilde{W} & \tilde{Y} \end{bmatrix} \begin{bmatrix} K^{-1/2} & -K^{-1}LZ^{1/2} \\ 0 & Z^{1/2} \end{bmatrix} \tag{3.15}$$

is then a unitary completion of the orthonormal columns of $\begin{bmatrix} \tilde{U}K^{-1/2} & \tilde{W}K^{-1/2} \end{bmatrix}^T$.

The matrices K, L and Z are given by

$$L = \tilde{U}^*X + \tilde{W}^*\tilde{Y} \tag{3.16}$$
$$K = \tilde{U}^*\tilde{U} + \tilde{W}^*\tilde{W} = I + \tilde{V}^*\tilde{V} \tag{3.17}$$
$$Z = X^*X + \tilde{Y}^*\tilde{Y}. \tag{3.18}$$

and it is easily checked that K and Z are positive definite.

Proposition 3.2.2 *[MO07, Prop. 5] A lossless function $G(z)$, given by a balanced realization (A, B, C, D), is in the domain $\mathcal{D}_\Omega$ of the chart associated with $\Omega = (W^*, U^*, Y^*, X^*)$ if and only if the solution Q to the Stein equation (3.6) is positive definite. A realization $\tilde{R}$ is in canonical form with respect to this chart if and only if the solution Q to (3.6) is $P^{1/2}$, P being a solution of (1.39).*

The coordinate map is

$$\phi_\Omega : \ G(z) \in \mathcal{D}_\Omega \to (D_0, V),$$

in which V and D_0 are computed as follows :

- first compute V and Q by (3.7) and (3.6)
- then compute $P = Q^*Q$ and $\mathcal{U}$ and $\mathcal{V}$ as in (3.15) and (3.12)

- finally compute $T = QP^{-1/2}$ and $(\tilde{A}, \tilde{B}, \tilde{C}, D) = (T^{-1}AT, T^{-1}B, CT, D)$ a balanced realization for $G(z)$ in canonical form. Then D_0 is given by

$$\mathcal{V}^* \begin{bmatrix} \tilde{D} & \tilde{C} \\ \tilde{B} & A \end{bmatrix} \mathcal{U} = \begin{bmatrix} I & 0 \\ 0 & D_0 \end{bmatrix}. \tag{3.19}$$

The balanced realization $\Omega = (W^*, U^*, Y^*, X^*)$ defines a lossless function

$$G_\Omega(z) = X^* + U^*(zI_n - W^*)^{-1}Y^* \in \mathbb{L}_n^m. \tag{3.20}$$

Since two similar balanced realizations provide the same chart up to a rotation in the parameters space, the chart is in fact attached to an element of $\mathbb{L}_n^m$ which explains the denomination lossless mutual encoding.

The solution Q of (3.6) has the following integral representation ([4, Th. A.2.2.])

$$Q = \frac{1}{2i\pi} \int_{\mathbb{T}} (z\,I - A)^{-1} B U (I - z\,W)^{-1} dz.$$

This formula can be interpreted as a matrix-valued inner product

$$Q = \left\langle (zI - A)^{-1}B, (zI - W^*)^{-1}U^* \right\rangle$$

defined on the left Hilbert module of $n \times m$ square integrable matrix-valued functions on the unit circle [10]. Since the pairs (W^*, U^*) and (A, B) are input normal, their distance in the norm associated with the matrix-valued inner product is

$$\| |(zI - A)^{-1}B - (zI - W^*)^{-1}U^* | \| = 2I - Q - Q^*.$$

Hence, the quality of a chart can be estimated by computing $\mathrm{m}(P) \in]0, 1]$:

$$\mathrm{m}(P) = \inf\{|\lambda|, \lambda \text{ eigenvalue of } P)\}.$$

The closer $\mathrm{m}(P)$ is to 1, the better is the chart, the best choice being $P = I$. This choice provides us with an adapted chart.

Proposition 3.2.3 (Adapted chart) *[MO07] The chart associated with a balanced realization $\Omega = (A, B, C, D)$ of $G(z)$ is an adapted chart for $G(z)$, i.e. the interpolation matrix V in (3.7) is the zero matrix and $D_0 = I$.*

If $\phi_\Omega(G) = (D_0, V)$ and Σ is a constant unitary matrix, then $\phi_\Omega(\Sigma G) = (\Sigma D_0, \Sigma V)$. An atlas for the (left) quotient space $\mathbb{L}_n^m \setminus \mathbb{U}_m$ is thus obtained by fixing D_0 within each chart.

3.3 RARL2 : a rational approximation software

The Matlab based software RARL2[1] solves the minimization problem (3.1) for a function $F(z)$ given in one of the following forms

[1] The software RARL2 is described in [MOHP02] and available on the web page `http:www-sop.inria.fr/apics/RARL2/rarl2-eng.html`.

(1) a realization $F(z) = \mathcal{D} + \mathcal{C}(zI_N - \mathcal{A})^{-1}\mathcal{B}$

(2) Fourier coefficients $F(z) = \sum_{j=0}^{N} F_j z^{-j}$.

It is mostly a model reduction tool. Pointwise values on the circle

$$F(e^{i\theta_j}), j = 1, \ldots, N$$

can also be handled, but in this case the L^2-norm is replaced by a least-square criterion. Then, the result is not guaranteed and very dependent on the data. When the data is band-limited a first completion step is highly desirable (see section 3.4). In this connection, other representations of functions, using alternative basis of $L^2(\mathbb{T})$ as orthogonal rational functions [12], could better suit the completion issue.

The implementation is based on a state space representation (A, B, C, D) of LTI stable systems approximants. The concentrated criterion (3.2) is thus defined on the set of lossless functions or equivalently input normal pairs (A, B). It is computed, as well as the gradient, using state-space formulas.

Let $H(z) = \mathcal{D} + \Gamma(zI - A)^{-1}B$ with (A, B) input normal. The error $F - H$ has realization

$$\tilde{A} = \begin{bmatrix} \mathcal{A} & 0 \\ 0 & A \end{bmatrix}, \tilde{B} = \begin{bmatrix} \mathcal{B} \\ B \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} \mathcal{C} & -\Gamma \end{bmatrix}, \tilde{D} = 0,$$

and L^2 norm

$$\|F - H\|_2^2 = \mathrm{Tr}\left(\tilde{C} W_c \tilde{C}^*\right),$$

where W_c is the reachability gramian of the error. Fixing A and B, the L^2 error is minimal for $\Gamma = \mathcal{C}W_{12}$, where W_{12} is the submatrix of W_c solution to

$$\mathcal{A}W_{12}A^* + \mathcal{B}B^* = W_{12},$$

and the concentrated criterion is

$$J(A, B) = \|F\|_2^2 - \mathrm{Tr}\left(\mathcal{C}W_{12}W_{12}^*\mathcal{C}^*\right). \tag{3.21}$$

Let W_{21} be the solution of the Lyapunov equation

$$A^*W_{21}\mathcal{A} - \Gamma^*\mathcal{C} = W_{21}, \quad \Gamma = \mathcal{C}W_{12}$$

The gradient of the concentrated criterion can then be computed as

$$\Delta J(A, B).(\partial A, \partial B) = 2\,\mathrm{Re}\left(\mathrm{Tr}\left(W_{21}\mathcal{A}\, W_{12}\, \partial A^*\right)\right) + 2\,\mathrm{Re}\left(\mathrm{Tr}\left(W_{21}\mathcal{B}\, \partial B^*\right)\right) \tag{3.22}$$

The software divides into two libraries

- `arl2lib` contains all the computations concerning the L^2 criterion.
  The functions `arl2SS` and `arl2COEFF` compute the criterion and the gradient when the function is given by a realization and Fourier coefficients respectively.

- `boplib` is concerned with the parametrization of lossless functions (balanced output pairs) by means of the lossless mutual encoding method described in section 3.2. It also provides a minimization process which could handle **any criterion** defined over the manifold $\mathbb{L}_n^m \setminus \mathbb{U}_m$.

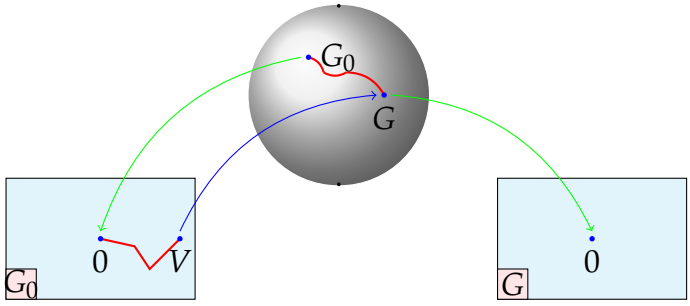

Figure 3.1: Optimization over a manifold

The minimization process makes use of the Matlab solver `fmincon`. It starts at some initial point G_0 which is encoded in its adapted chart $\Omega = (A, B, C, D)$ (see Proposition 3.2.3). Then `fmincon` performs the optimization of the criterion submitted to the nonlinear constraint $P > 0$, where P is the solution to

$$P - APA^* = BB^* - V^*V$$

and V the parameter of the current point in the chart. This constraint ensures we remain within the domain of the chart. When a constraint violation occurs, a new adapted chart is computed for the current point and the optimization pursues within this new chart (see Figure 3.1), until a minimum is reached. A scheme of the whole minimization process is provided in Figure (3.2).

The convergence of the algorithm has been proved under mild assumptions in the SISO case [BCO91] but never in the MIMO case. The main obstruction to the convergence is if the boundary of the manifold is reached, that is to say if the constraint violation (P singular) corresponds to the non-minimality of the canonical realization, that is to a drop of degree for the lossless functions. This would results in changing chart indefinitely.

Since the criterion may possess many local minima, the choice of an initial point in the optimization process is essential. The approximant provided by another model reduction method, mainly the balanced truncation method of [41], is often used as a starting point. An interesting recursive method on the order of the approximant can also be used (see [BCO91] in the scalar case and [FO98, PHO02] in the matrix case). This strategy has been implemented (function `RARL2`) on the bases of the following observation.

If $\hat{G}(z)$ is factor as $\hat{G} = B_{w,u}\, G$, where $G(z)$ is lossless of degree k, and $B_{w,u}(z)$ is the Potapov factor (1.11), we get an interesting formula for the criterion (3.2)

$$\psi_{k+1}(\hat{G}) = \psi_k(G) - (1 - |w|^2)\|R(w)^* u\|^2,$$

where $R(w)$ is the smooth function of w

$$R(w) = \frac{1}{2i\pi} \int_0^{2\pi} \frac{\pi^-(FG^\sharp)(z)}{1 - \bar{w}z}\, dz.$$

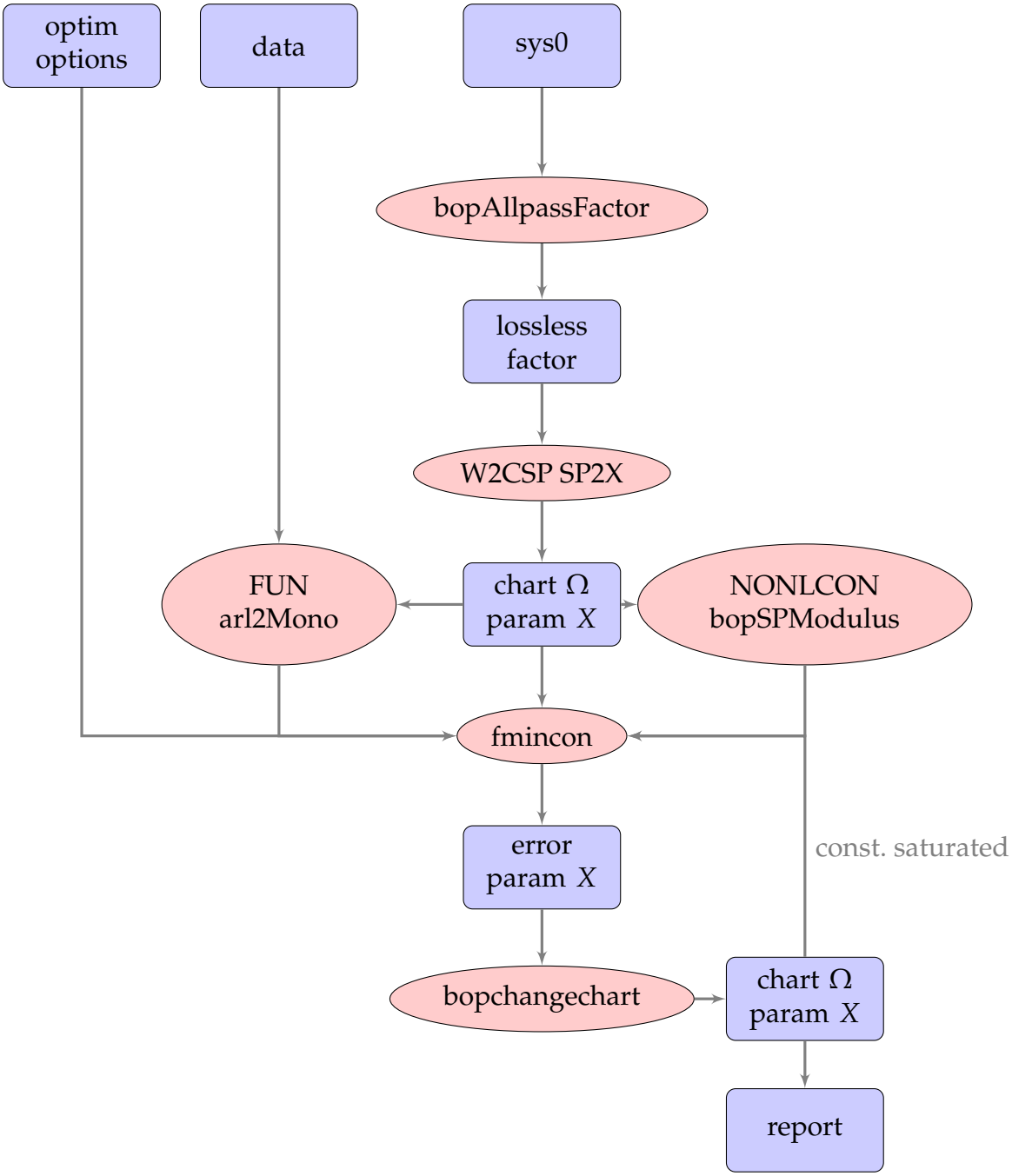

Figure 3.2: Minimization process in RARL2

If w converges to the unit circle, $B_{w,u}(z)$ converges to a unitary matrix X and the function $\hat{G}(z)$ of degree $k+1$, to $XG(z)$ which has degree k. We thus reach the boundary of the manifold $\mathbb{L}_n^m$. Conversely, if $G(z)$ is a minimum of the criterion of order k, then we may choose any $\hat{G} = B_{w,u}\, G$, with w close enough to the unit circle so that $\psi_{k+1}(\hat{G}) < \psi_k(G)$, as starting point for a minimization at order $k+1$. Starting at degree 0 and choosing several starting points at each iteration allows to compute a great number of local minima and to improve the chances to reach the global one.

3.4 Identification of microwave filters.

RARL2 is a generalist software which efficiently performs model reduction. It is intended to be associated with other tools in order to solve more specific inverse problems. We discuss in this section a dedicated application which is now mainly developed by F. Seyfert. It illustrates the possibilities of the software.

The problem is to recover the transfer function of a filter from frequency data. These data are estimate values of the transfer function at pure imaginary points, obtained from the steady-state outputs of the filter to harmonic inputs. Frequency data are an interesting starting point in the identification process. The LTI assumption is already integrated in some sense and the quality of the measurements can be estimated [54].

A two stage approach has been carried out to deal with this problem: in a first step, a stable transfer function of high degree approximating the data is searched. Then, a model reduction step is performed (see Figure 3.3).

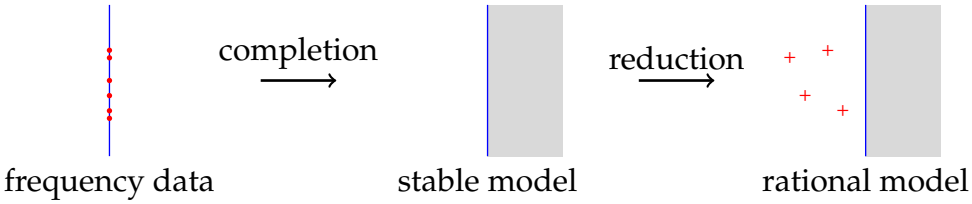

Figure 3.3: Identification vs model reduction

To carry over this problem into our discrete-time setting, we use an isometry from the Hardy space $\mathcal{H}_2$ of the right half-plane to the Hardy space $H_2^\perp$. Note that a function $\tilde{F}(s) \in \mathcal{H}_2^{m\times m}$ is square integrable on the imaginary axis and thus strictly proper (it vanishes at infinity). The function $F(z)$ of $(H_2^{m\times m})^\perp$ which is associated with $\tilde{F}(s)$ is

$$F(z) = \frac{\sqrt{2}}{z-1}\tilde{F}\left(\frac{z+1}{z-1}\right). \tag{3.23}$$

It satisfies $\|F\|_{L^2(\mathbb{T})} = \|\tilde{F}\|_{L^2(\mathbb{I})}$ and if $\tilde{F}(s)$ is rational, $\deg F = \deg \tilde{F}$. The formulas which allow to derive a realization $(A, B, C, 0)$ of $F(z)$ from a realization $(\tilde{A}, \tilde{B}, \tilde{C}, 0)$ of

$\tilde{F}(s)$ and reciprocally are completely symmetric:

$$\left\{\begin{array}{rcl} C &=& \tilde{C} \\ A &=& -(I-\tilde{A})^{-1}(I+\tilde{A}) \\ B &=& (I-\tilde{A})^{-1}\tilde{B} \end{array}\right. \qquad \left\{\begin{array}{rcl} \tilde{C} &=& C \\ \tilde{A} &=& -(I-A)^{-1}(I+A) \\ \tilde{B} &=& (I-A)^{-1}B \end{array}\right. \tag{3.24}$$

The data on the imaginary axis can be sent on the unit circle via the Möbius transformation $iw_k \mapsto \frac{iw_k+1}{iw_k-1}$ and we get an equivalent problem in the framework of the disk (see Figure 3.4). However, if the original system fails to be strictly proper, then the direct feedthrough (value at infinity of the transfer function) must be evaluated.

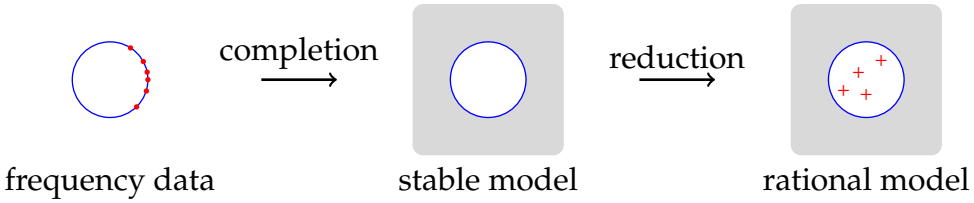

Figure 3.4: Two steps identification in discrete-time

Since the data are usually band-limited, this problem is not well-posed and a metric constraint must be imposed outside the bandwidth. This completion problem is a difficult task, for which a lot of methods have been proposed, whose efficiency mostly depend on the quality of the data and require some engineering ability. The methods divide in two classes, interpolation methods among which boundary Nevanlinna-Pick interpolation algorithms sounds interesting [14], and approximation methods more inclined to take into account the quality of the data.

For the identification of microwave filters, we chose the approximation approach and solved a bounded extremal problem. The whole process is described in [BGL+98]. The model reduction step is then performed by the software RARL2 as explained in section 3.3. The microwave filters that we consider are used in telecommunication satellites for channel multiplexing. A longstanding cooperation with the space agency CNES resulted in a dedicated software PRESTO-HF [60] that wraps RARL2 into a package which is now fully integrated in the design and tuning process. In Figure 3.5, a 8th order model of a MIMO 2×2 microwave filter is shown, obtained from 800 pointwise data.

3.5 More applications

The rational approximation method and the software described in this chapter have been used in several different situations which demonstrates the adaptability of this approach.

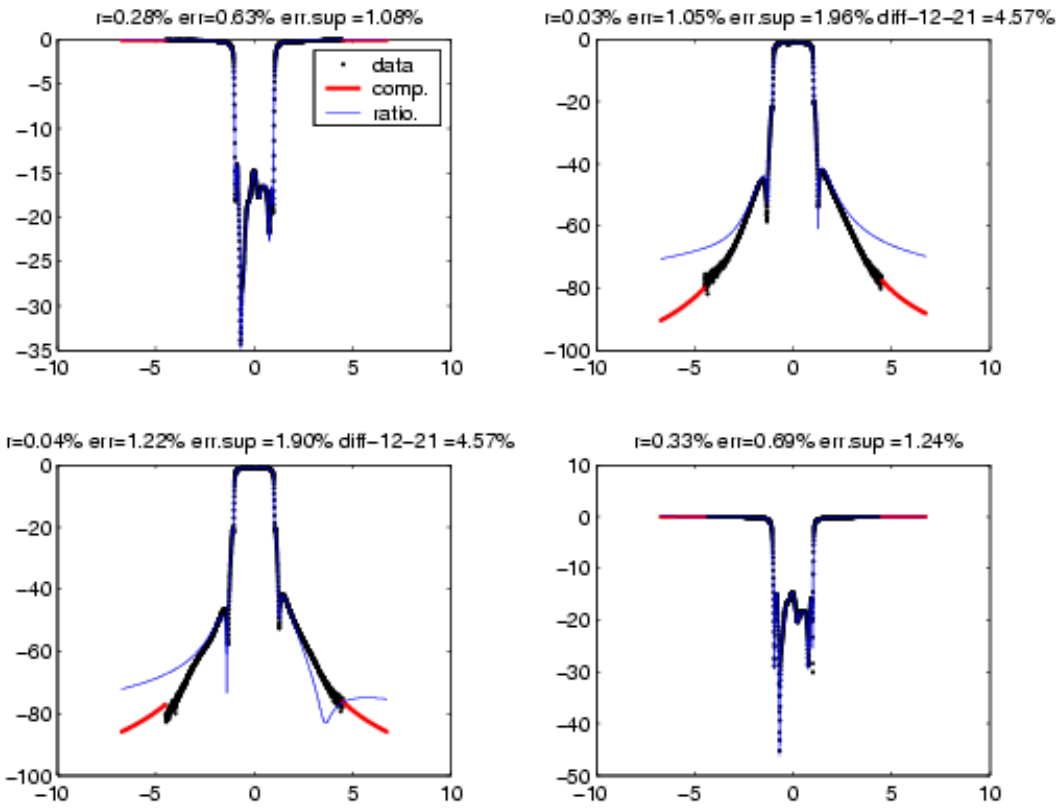

Figure 3.5: Bode diagram of the data and the approximant at order 8.

3.5.1 Localization of dipolar sources in electroencephalography.

This research is conducted by J. Leblond and L. Baratchart from the APICS-INRIA team. From measurements by electrodes of the electric potential $u(z)$ on the scalp, the problem is to recover a distribution of m pointwise dipolar current sources $C_k \in \mathbb{R}^3$ with moments $p_k \in \mathbb{R}^3$ located in the brain (modeling the presence of epileptic foci). The head is modeled as a set of three spherical nested regions (brain, skull, scalp) and in each region constant conductivities are assumed. A macroscopic model and quasi-static approximation of Maxwell-equations are used to describe the spacial behavior of $u(z)$. In the brain, the electric potential $u(z)$ assumes the form

$$u(z) = h(z) + \sum_{k=1}^{m} \frac{< p_k,\ x - C_k >}{4\pi \|x - C_k\|^3},$$

where $h(z)$ is harmonic, while the other component (anti-harmonic part $u_a(z)$) bears the information on the sources.

The inverse problem can be approach in two steps: get first $u_a(z)$ from the data and then recover the localizations and the moments of the sources from $u_a(z)$. Rational approximation is involved in this second step [6] and performs on planar sections (see Figure 3.6).

Note that $u_a(z)$ is not rational and only has singularities. However, the restrictions

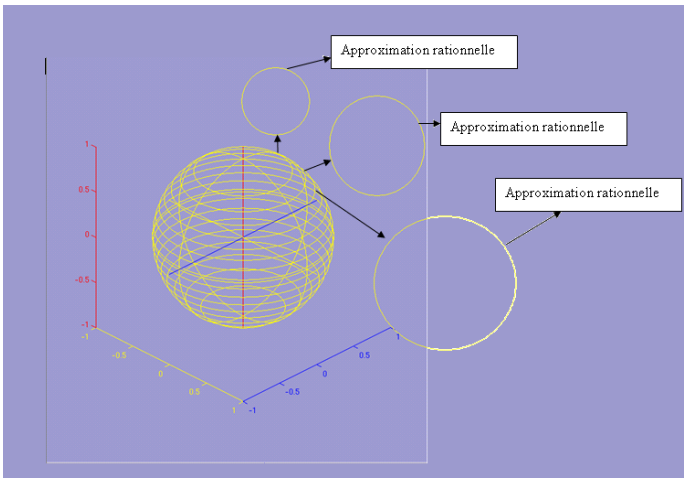

Figure 3.6: Rational approximation on planar sections

$f_p(z)$ of $u_a(z)^2$ to planar sections Γ_p of the brain possesses triple poles and singularities at the same location. These singularities are strongly and explicitly linked with the sources, and the poles of the best L^2 rational approximations to f_p on Γ_p accumulate to these singularities [7] (see Figure 3.7).

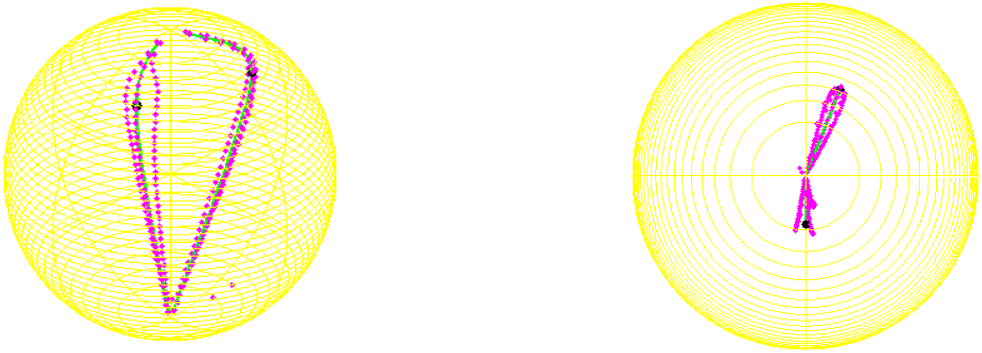

Figure 3.7: Localization of sources: the singularities (green) are aligned and the sources (black) correspond to the maximum modulus. The rational approximations (red) accumulate to these singularities.

One of the advantages of this method is that it does not require the a priori knowledge of the exact number of sources. If the order of the approximation is larger than the number of sources, the extra poles accumulate to the boundary of the domain Γ_p. For this application, the software RARL2 has been modified to impose triple poles for the approximant. This was an easy trick thanks to the flexibility of the parametrization of section 3.2.

3.5.2 Multi-objective control.

In [57], revisited in a chain-scattering perspective in [DMO05], it is shown that if the pair (C, A) of the Youla parameter $Y(z) = D + C(zI - A)^{-1}B$ is fixed, then the search over the parameters (B, D) can be reduced to an efficiently solvable LMI problem. Limiting the search of the parameter $Y(z)$ to the FIR form

$$Y(z) = Y_0 + Y_1\frac{1}{z} + \ldots + Y_p\frac{1}{z^p}$$

provides solutions to the multi-objective control problem. However, this is also the main limitation of the approach as high order expansions might be necessary, due notably to the fact that the poles structure is fixed through the pair (C, A). Such a drawback can be avoided if the search is performed over all the parameters $Y(z)$ of fixed McMillan degree. This can be done using the atlas of section 3.2 to parametrize the corresponding pairs (C, A). The modularity of RARL2 allows to use this parametrization (library `boplib`) to optimize a different criterion. This algorithm for multi-objective control has been implemented by M. Bordier and J.P. Marmorat.

3.5.3 Wavelets approximation

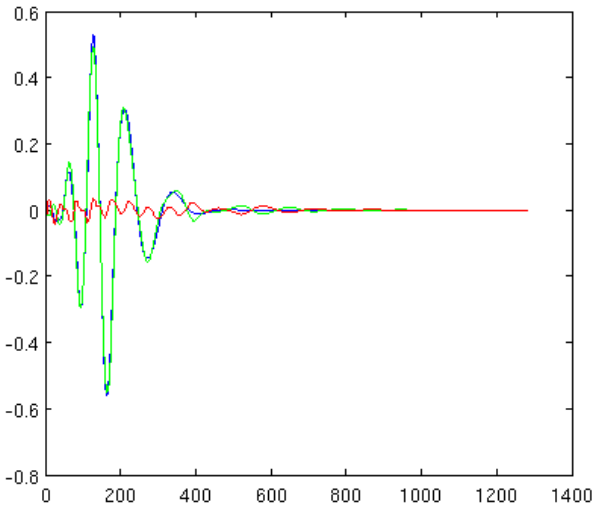

Figure 3.8: Daubechies wavelet db7 (blue) and its degree 8 rational approximation (green).

This application of our rational approximation methods to orthogonal wavelets was proposed by R. Peeters. The problem is to implement wavelets in analog circuits in view of medical signal processing applications. A dedicated method has been developed [38] based on an L^2-approximation of the wavelet by the impulse response of a

stable, causal, low order filter. However, this method fails to find an accurate and sufficiently small order approximation in some difficult cases (Daubechies wavelets db7 and db3).

The idea was to use the software RARL2 to perform a model reduction on an accurate high order (100-200) approximation. However, an admissibility condition for wavelets is that the integral of a wavelet equals zero, which means that it has a vanishing moment of order 0. The low order approximation is still required to have an integral zero, otherwise undesired bias will show up when the wavelet is used in an application. We thus had to adapt a version of the RARL2 software to address this constraint. Since we are dealing with a quadratic optimization problem under a linear constraint, this can be solved analytically. We could thus reformulate the problem of L^2-approximation subject to this constraint into an optimization problem over the class of lossless systems. This could be handled by the software with only minor changes and we were able to perform an accurate approximation of order 8 for db7 (Figure 3.8).

3.6 Perspectives.

The examples presented in section 3.5 show the wide range of applicability of the method and suggest a number of improvements as well as new lines of research. In the applications to multi-objective control and wavelet approximation, the criterion has been modified to address a different linear constraint.

In the inverse problem of source recovery, the function to be approximated is not rational but non integer rational (it involves fractional powers). A interesting topic for future research would be the approximation by non integer rational functions. Fractional (or non integer) systems and fractional orthonormal basis are active domain of research which could be connected in some way to our works. However, I have no knowledge in this domain, and this is presently a very prospective subject.

At contrary, enforcing passivity in rational approximation tools is already under investigation. Passive devices play an important role in a lot of application areas: telecommunication, chemical process control, economy, biomedical processes. Network simulation software packages (as ADS or SPICE) require passive models for their components. However, enforcing passivity while identifying a model from (band limited) frequency data is still an open and challenging problem. The current approaches are mostly "ad hoc" methods. The possibility to handle this problem using the approximation techniques developed in the APICS team seems promising. Up to now, if the data comes from a passive system, the passivity of the model is not guaranteed by our approach.

The passivity constraint brings new difficulties and rises a number of interesting theoretical and practical questions, both in the interpolation and in the approximation step. A passive system has, in case of a scattering representation, a transfer function which is Schur (contractive in the stability domain, see section 1.3). The research could be carried out in many directions. The Schur approximation problem should be studied from a theoretical viewpoint: existence of an approximant, behavior of the criterion

at the boundary of the domain, etc. in order to design a specific algorithm. It must be noticed that approximation from band-limited data becomes a well-posed problem with the Schur constraint.

The parametrization issue for Schur functions should also be addressed. In V. Lunot PHD thesis [48], a parametrization of all strictly Schur rational functions of degree n is constructed from a multipoint Schur algorithm (the recursive algorithm (1.26) in which the interpolation points can be chosen anywhere in the open unit disk), the parameters being both the interpolation values and interpolation points. Examples are computed by an L^2 norm optimization process and the results are validated by comparison with the unconstrained L^2 rational approximation. Choosing the interpolation points with respect to the approximated Schur function so as to yield the best convergence possible remains a major open issue. An alternative approach relies on the observation that a Schur function can be viewed as a sub-block of a lossless function of higher dimension (see Chapter 4). A parametrization of Schur functions could be deduced from that of lossless functions. This approach must be connected to some works on degree-constrained analytic interpolation (see e.g [44, 45]). The use of boundary interpolation could be proved necessary in this approach to handle functions which are not strictly Schur.

Chapter 4

Symmetric lossless matrix-valued extensions of rational contractions

In this chapter, the application we have in mind is the synthesis of resonant filters, mainly microwave or SAW filters used in telecommunication devices. These filters are usually represented by a scattering matrix which is a lossless matrix-valued function. The lossless property is the expression of the conservation law. The reciprocity law attached to the wave propagation phenomenon makes this matrix symmetric. We are thus interested in an useful description of symmetric lossless matrices in view of filters design applications.

Since the synthesis often concerns the transmission, a submatrix of the scattering matrix, we aim at characterizing the sub-matrices of a scattering matrix. Any sub-matrix of a lossless matrix is contractive in the analyticity domain, that is a Schur function (see Section 1.3). Conversely, on which conditions a symmetric Schur matrix $S(s)$ can be extended into a symmetric lossless one? Without the symmetry requirement, such an extension is known under the name of Darlington synthesis. This problem has been widely studied in the sixties for a circuit synthesis purpose. Darlington proved the existence of such an extension.

The symmetric Darlington synthesis problem has been little studied and the existing solutions require an important increase of the size of the extension [2]. In [BEGO07], we proved that a Schur matrix $S(s)$ do possess a symmetric lossless extension if and only if its zeros have even multiplicity. If this condition is not satisfied, given a $m \times m$ symmetric Schur matrix $S(s)$ of McMillan degree n, two dual extension problems can be formulated:

- either we fix the size of the extension to $2m \times 2m$ and we look for a minimal degree extension,
- or we fix the degree n of the extension and we look for a minimum size extension.

The solutions obtained in [BEGO07, BEGO10] for each of these problems improve significantly the results in the literature.

In this chapter, we stick to the continuous-time case, which is relevant in filter design applications, so that our lossless functions are analytic in the right half-plane. The

relevant spaces of functions are the Hardy spaces of the right half-plane $\mathcal{H}_2$ and $\mathcal{H}_\infty$. A Schur matrix function belongs to the unit ball of $\mathcal{H}_\infty^{m\times m}$. We set

$$W^*(s) = W(-\bar{s})^*,$$

the pseudo-Hermitian conjugate of $W(s)$.

4.1 The scalar case

The case where $S(s)$ is a scalar function is easily handled and gives a first illustration of these problems. If $S(s) = \frac{p(s)}{q(s)}$ is a Schur rational function, then any 2×2 rational lossless extension can be written in the form [9]

$$\hat{S} = \frac{1}{q}\begin{bmatrix}\xi & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}p^* & -r^*\\ r & p\end{bmatrix} \tag{4.1}$$

where $\xi \in \mathbb{T}$ and $r(s)$ is a polynomial solution to the spectral factorization problem:

$$rr^* = qq^* - pp^*. \tag{4.2}$$

The Schur property ensures that (4.2) has a solution. Any solution consists in sharing the conjugate zeros $(\alpha, -\bar{\alpha})$ of $qq^* - pp^*$ between the two factors (the zeros of $qq^* - pp^*$ on the imaginary axis have even multiplicity). Then, a symmetric extension do exist if and only if the zeros of $qq^* - pp^*$ have even multiplicities. In this case, it will be essentially unique.

The scattering matrix of a 2-port is of this form. A nice description of these matrices is in term of the polynomial numerators $p(s)$ and $r(s)$. For the matrix (4.1) to be symmetric, the polynomial $r(s)$ must satisfy $r^* = -\bar{\xi}r$. The scattering matrix of a two-port satisfies in addition $S(\infty) = I$, so that $\xi = (-1)^n$ and it can thus be parametrized in the form

$$\frac{1}{q}\begin{bmatrix}(-1)^n p^* & r\\ r & p\end{bmatrix} \tag{4.3}$$

where $p(s)$ ranges over the monic polynomials of degree n and $r(s)$ over the $n-1$ degree polynomials satisfying the condition $r^* = (-1)^{n-1}r$. This form is particularly interesting for synthesis purposes [11, 46, 47], since the modulus of the transmission $S_{12}(s) = \frac{r(s)}{q(s)}$ can only be expressed in terms of $p(s)$ and $r(s)$

$$|S_{12}|^2 = \frac{1}{1 + \frac{|p|^2}{|r|^2}}.$$

From this application viewpoint, completion is closely related to a parametrization issue: given a part (a block, a column, an element, two numerators) of a lossless matrix, is it possible to rebuilt the whole matrix? In an unique way?

If the zeros of $qq^* - pp^*$ fail to have even multiplicities, an extension either of higher degree or of higher dimension must be searched.

We may choose for $r(s)$ the spectral factor whose roots are in the closed left half-plane, unique up to a unit complex number. The extension (4.1) can thus be made symmetric multiplying it by an appropriate lossless factor (to simplify the writing we set $\xi = 1$):

$$\tilde{S} = \begin{bmatrix} -\frac{p^* r^*}{qr} & \frac{r^*}{q} \\ \frac{r^*}{q} & \frac{p}{q} \end{bmatrix} = \begin{bmatrix} -\frac{p^*}{q} & \frac{r^*}{q} \\ \frac{r}{q} & \frac{p}{q} \end{bmatrix} \begin{bmatrix} \frac{r^*}{r} & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.4}$$

This extension has degree $2n - n_0$ where n_0 is the number of roots of $r(s)$ on the imaginary axis, counted with multiplicity.

But we can be cleverer and write $qq^* - pp^* = r_1^2 r_2 r_2^*$, where $r_1(s)$ is auto-reciprocal $r_1^* = r_1$, and $r_2(s)$ stable with only simple roots. Then, the following extension is symmetric:

$$\Sigma = \begin{bmatrix} -\frac{p^*}{q}\frac{r_2^*}{r_2} & \frac{r_1 r_2^*}{q} \\ \frac{r_1 r_2^*}{q} & \frac{p}{q} \end{bmatrix} = \begin{bmatrix} -\frac{p^*}{q} & \frac{r_1 r_2^*}{q} \\ \frac{r_1 r_2}{q} & \frac{p}{q} \end{bmatrix} \begin{bmatrix} \frac{r_2^*}{r_2} & 0 \\ 0 & 1 \end{bmatrix}.$$

and in fact it has minimal possible degree.

4.2 From Darlington synthesis to symmetric extensions

Since the notion of a zero is more complicated in the matrix case, the solution of the extension problem will be more involved. The basic extension problem is, given a $m \times m$ Schur matrix $S(s)$ of McMillan degree n, to find a $2m \times 2m$ lossless extensions of the same degree

$$\hat{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S \end{bmatrix}.$$

The Darlington synthesis can still be obtained in two steps: first compute $S_{21}(s)$ such that $S_{21} S_{21}^* = I - SS^*$ is a spectral factorization, which makes $\begin{bmatrix} S_{21} & S \end{bmatrix}$ isometric, and next, compute a lossless extension of it.

Concerning the spectral factorization in the matrix case, we have the following result [67]:

Theorem 4.2.1 (Youla) *Let $\Phi(s)$ be a rational $m \times m$ paraconjugate Hermitian matrix, i.e.*

$$\Phi^*(s) = \Phi(s),$$

of normal rank k which is non-negative on the imaginary axis. Then, there exists a $m \times k$ matrix $\tilde{W}(s)$ such that
(1) $\Phi(s) = \tilde{W}(s)\tilde{W}^*(s)$

(2) *$\tilde{W}(s)$ and its left inverse $\tilde{W}(s)^{-1}$ are both analytic in Re $s > 0$.*

(3) *$\tilde{W}(s)$ is unique up to a constant, unitary $k \times k$ matrix multiplier on the right.*

(4) *Any factorization of the form $\Phi(s) = W(s)W^*(s)$ in which $W(s)$ is $m \times k$, rational and analytic in Re $s > 0$, is given by $W(s) = \tilde{W}(s)Q(s)$ where $Q(s)$ is inner.*
$\tilde{W}(s)$ is called the minimum phase *spectral factor and $W(s)$ a spectral factor.*

Since $S(s)$ is contractive, the matrix $\Phi(s) = I - S(s)S^*(s)$ has all the required properties. We assume that the normal rank of $\Phi(s)$ is m, which amounts to say that the $S(s)$ is strictly contractive at some point of the imaginary axis. The spectral factors are square $m \times m$ matrices whose determinants does not vanish identically. They are thus invertible as rational matrices. We will be mainly interested with those spectral factors for which the extension $\begin{bmatrix} S_{21} & S \end{bmatrix}$ has same degree n than $S(s)$ and which are called *minimal spectral factors*.

In [42], a state-space construction of all the lossless extensions preserving the McMillan degree is presented. It involves the solutions to a Riccati equation which is closely connected to the spectral factorization through the Bounded Real Lemma. This lemma [2] asserts that a function given by a minimal realization, $S(s) = D + C(sI - A)^{-1}B$, is a Schur function strictly contractive at ∞ if and only if there exist P, $\hat{B}$, and D_{21} such that

$$\begin{aligned} AP + PA^* + BB^* + \hat{B}\hat{B}^* &= 0 \\ PC^* + BD^* + \hat{B}D_{21}^* &= 0 \\ DD^* + D_{21}D_{21}^* &= I \end{aligned}$$

and P is positive definite. Then

$$S_{21}(s) = D_{21} + C(sI - A)^{-1}\hat{B}$$

is a left spectral factor of $I - SS^*$.

Let $D_{21} = (I - DD^*)^{1/2}$, then $\hat{B} = -(PC^* + BD^*)D_{21}^{-1}$ and P is solution to the Riccati equation

$$\mathcal{R}(P) = P\gamma P + \alpha P + P\alpha^* + \beta = 0 \tag{4.5}$$

where

$$\left\{ \begin{array}{rcl} \alpha & = & A + BD^*(I - DD^*)^{-1}C, \\ \beta & = & B(I - D^*D)^{-1}B^*, \\ \gamma & = & C^*(I - DD^*)^{-1}C. \end{array} \right.$$

Theorem 4.2.2 *[42] The following statements are equivalent*

(i) $S(s)$ is a Schur function which is strictly contractive at infinity.

(ii) the Riccati equation (4.5) has an Hermitian solution

(iii) All pure imaginary eigenvalues of $\mathcal{A}$ have even multiplicity

In this case, there is a one to one correspondence between the Hermitian solutions of (4.5) and the lossless extensions of $S(s)$ of degree n, whose value at infinity is prescribed by

$$D_{21} = (I - DD^*)^{1/2}, \quad D_{12} = (I - D^*D)^{1/2}, \quad D_{11} = -D^*$$

The extension is given by

$$\hat{S}_P(s) = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D \end{bmatrix} + \begin{bmatrix} \hat{C} \\ C \end{bmatrix} (sI - A)^{-1} \begin{bmatrix} \hat{B} & B \end{bmatrix}.$$

Associated with the Riccati equation (4.5), the matrix

$$\mathcal{A} = \begin{bmatrix} -\alpha^* & -\gamma \\ \beta & \alpha \end{bmatrix}$$

is in fact a (possibly non-minimal) dynamic matrix of $(I - SS^*)^{-1}$. It is Hamiltonian and its eigenvalues are thus symmetric with respect to the imaginary axis, counting multiplicities. It satisfies the important similarity relation

$$\begin{bmatrix} I & 0 \\ -P & I \end{bmatrix} \mathcal{A} \begin{bmatrix} I & 0 \\ P & I \end{bmatrix} = \begin{bmatrix} -(\alpha + P\gamma)^* & -\gamma \\ 0 & \alpha + P\gamma \end{bmatrix}.$$

The spectrum of $\mathcal{A}$ splits as $\sigma(\mathcal{A}) = \sigma(\alpha + P\gamma) \cup \sigma(-(\alpha + P\gamma)^*)$. It can be verified that S_{21}^{-1} has dynamic matrix $\alpha + P\gamma$ while S_{12}^{-1} has dynamic matrix $-P(\alpha + P\gamma)^* P^{-1}$. The block diagonal form shows us that the matrix P induces a distribution of the zeros of $I - SS^*$ (eigenvalues of $\mathcal{A}$) between S_{12} and S_{21}, in such a way that S_{12} and S_{21} have conjugate zeros with respect to the imaginary axis. This also follows from the next equality which generalizes in some sense (4.1) to the matrix case [BEGO10]

$$\det \hat{S} = -\det S_{12} \det(S_{21}^*)^{-1}. \tag{4.6}$$

However, the state-space construction of Theorem 4.2.2 requires the condition $S(\infty)$ strictly contractive and symmetric realizations to address the symmetric extension problem. However, the scattering matrix of a m-port usually satisfies $S(\infty) = I$, while a real minimal realization of a symmetric function may fail to be symmetric, which complicates the study of the extension problem for matrices with real coefficients.

We give a more general frequency domain formulation which can be transposed to the discrete-time setting (functions analytic in the unit disk). Since $S(s)$ is strictly contractive at some point of the imaginary axis, for almost every $\nu \in \mathbb{T}$, the matrix $\nu I - S(s)$ is invertible in the closed right half-plane, so that $(\nu I - S(s))^{-1}$ is stable. Then, every inner extension of $\begin{bmatrix} S_{21} & S \end{bmatrix}$ is obtained as follows.

Theorem 4.2.3 *[BEGO10] Let $S(s)$ be a $m \times m$ Schur function such that $\nu I - S(s)$ is invertible in $\mathcal{H}_\infty^{m\times m}$. Let $S_{21}(s)$ be a spectral factor of $I - SS^*$. Every lossless extension $\Sigma(s)$ of $[S_{21}\ S]$ can be written as:*

$$\hat{S} = \begin{bmatrix} M - MS_{21}^*(I - \nu S^*)^{-1} S_{21} & MS_{21}^*(I - \nu S^*)^{-1}(\nu I - S) \\ S_{21} & S \end{bmatrix} \tag{4.7}$$

where $M(s)$ is a right Douglas-Shapiro-Shields inner factor of $\beta = (\nu I - S)^{-1} S_{21}$, i.e.

$$\beta = \alpha^* M, \tag{4.8}$$

with $\alpha(s)$ (as $\beta(s)$) in $\mathcal{H}_2^{m\times m}$, i.e. stable (see Theorem 1.1.5). The extension $\Sigma(s)$ has same degree as $[S_{21}\ S]$ if and only if $M(s)$ has minimal degree, i.e. $M(s)$ and $\alpha(s)$ are left coprime.

Formula (4.7) offers a frequency domain expression that seems to be lacking. It is closely related to the *mixed matrix* representation of a m-port. In such a representation, the input/output couples are of two types, voltages and currents on one hand, incoming and outgoing waves on the other hand. It naturally arises in the description of a surface acoustic wave filter in which two distinct types of energy are involved. However, using a well-known transformation, any scattering matrix can be put into this form. It also calls to mind the lossless embedding formula (see Proposition 1.3.2).

We also have the following description of all the lossless extensions of a Schur function in terms of minimal degree extensions:

Proposition 4.2.1 *[BEGO07, Prop.2] All rational lossless extensions of a Schur function S, contractive at ∞, can be written on the form*

$$\begin{bmatrix} L & 0 \\ 0 & I \end{bmatrix} \hat{S} \begin{bmatrix} R & 0 \\ 0 & I \end{bmatrix} \tag{4.9}$$

where L, R are lossless, and $\hat{S}(s)$ is a lossless extension of $S(s)$ at the same McMillan degree.

4.3 Minimal degree vs minimal size symmetric extensions

An analog of the extension (4.4) in the scalar case can be obtained in the matrix case as follows:

Proposition 4.3.1 *[BEGO07, Prop.4] Let*

$$\tilde{S} = \begin{bmatrix} \tilde{S}_{11} & \tilde{S}_{12} \\ \tilde{S}_{21} & S \end{bmatrix} \tag{4.10}$$

be a minimal degree extension (Theorem 4.2.3) associated with the minimum phase spectral factor $\tilde{S}_{21}(s)$ of $I - SS^*$. Let n_0 denote the number of zeros on the imaginary axis of $\tilde{S}_{21}(s)$. The extension

$$\tilde{\Sigma} = \begin{bmatrix} \tilde{S}_{11} & \tilde{S}_{12} \\ \tilde{S}_{21} & S \end{bmatrix} \begin{bmatrix} \tilde{Q} & 0 \\ 0 & I \end{bmatrix}, \quad \tilde{Q} = \tilde{S}_{21}^{-1} \tilde{S}_{12}^T, \tag{4.11}$$

is symmetric, lossless and has degree $2n - n_0$.

It is easily seen that $\tilde{Q}(s)$ is unitary and analytic in the right half-plane, since $\tilde{S}_{21}(s)$ is minimum phase. The assertion on the degree follows from (4.6). This extension is the starting point for the construction of minimal degree and minimal size extensions. It possesses a very particular property among all the lossless extensions of $S(s)$ that will help us to understand the algorithm of the next section.

Let $\hat{S}(s)$ be any minimal degree extension of $S(s)$, then

$$\hat{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S \end{bmatrix} = \begin{bmatrix} L^* & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \tilde{S}_{11} & \tilde{S}_{12} \\ \tilde{S}_{21} & S \end{bmatrix} \begin{bmatrix} R & 0 \\ 0 & I \end{bmatrix} \tag{4.12}$$

in which $R(s)$ and $L(s)$ are lossless. Indeed, $R = \tilde{S}_{21}^{-1} S_{21}$ is analytic in the right half-plane as $\tilde{S}_{21}^{-1}$, while $L^* = S_{12}\tilde{S}_{12}^{-1}$ is analytic in the left half-plane as $\tilde{S}_{12}^{-1}$. Moreover, since $\hat{S}(s)$ and $\tilde{S}(s)$ have the same determinant, we must have

$$\det L(s) = \det R(s).$$

Let $\Sigma(s)$ be the symmetric unitary extension associated with $\hat{S}(s)$ by

$$\Sigma = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & I \end{bmatrix}, \quad Q = S_{21}^{-1} S_{12}^T. \tag{4.13}$$

It is easily verified that we have $R^* \tilde{Q} (L^*)^T = Q$, and thus

$$\det Q = \det \tilde{Q} (\det R^*)^2.$$

Since only zeros of $\det \tilde{Q}(s)$ with even multiplicity can be canceled, if κ denote the number of distinct zeros of $\det \tilde{Q}(s)$ with odd multiplicity, then $\det Q(s)$ has McMillan degree greater than or equal to κ. Using Proposition 4.2.1, the following result is easily deduced.

Lemma 4.3.1 *Let $\tilde{S}(s)$ be the lossless extension (4.11). Let κ denote the number of distinct zeros of* $\det \tilde{Q}(s)$ *with odd multiplicity. Then, any symmetric lossless extension of $S(s)$ has degree greater than or equal to $n + \kappa$.*

Note that the extension (4.13) is not always lossless. In [BEGO07, Prop. 3] , it is proved that the matrix $Q(s)$ is lossless if and only if $P^{-T} - P$ is positive definite, P being the solution to the Riccati equation (4.5) associated with $\hat{S}(s)$. As a consequence, we recover the fact that $\tilde{Q}(s)$ is lossless, since the solution $\tilde{P}$ of the Riccati equation (4.5) associated with $\tilde{S}(s)$ is minimal, $P - \tilde{P} > 0$, for any other solution.

As in the scalar case, the extension (4.13), $\tilde{\Sigma}(s)$, is the worse that one can do, concerning the degree of the extension. In order to reduce the degree while keeping the extension symmetric, it is possible to factor out symmetrically Potapov factors (1.12) associated with double zeros. In the right half-plane setting that we consider in this chapter, the Potapov factors are of the form

$$B_{w,u}(s) = I + (b_w(s) - 1)\, uu^*, \quad b_w(s) = \frac{s - w}{s + \overline{w}}. \tag{4.14}$$

Lemma 4.3.2 (Symmetric Potapov factorization) *Let $T(s)$ be a $m \times m$ symmetric inner function of McMillan degree N. The following assertions are equivalent:*

1. *$T(s)$ has a zero ω of* multiplicity *strictly greater than* 1
2. *there exists a unit vector which satisfies the conditions*

$$\begin{aligned} T(\omega)u &= 0 \\ u^T T'(\omega)u &= 0 \end{aligned}$$

3. *the matrix $T(s)$ can be factored as*

$$T(s) = B_{\omega,u}(s)^T\, R(s)\, B_{\omega,u}(s)$$

for some rational inner matrix $R(s)$ of degree $N-2$

The assertion 1. $\Rightarrow$ 2. makes use of the Takagi factorization [33, Cor. 4.4.4] which is a special singular value decomposition for symmetric matrices: a symmetric matrix Λ can be written in the form

$$\Lambda = U^T \Delta U, \tag{4.15}$$

in which Δ is a positive diagonal matrix and U is unitary. Two different situations may occur which are illustrated on the following simple examples:

- some partial multiplicity (see section 1.1.1) is even:

$$\begin{bmatrix} b_\omega(s)^2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} b_\omega(s) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_\omega(s) & 0 \\ 0 & 1 \end{bmatrix}$$

- all the partial multiplicities are 1:

$$\begin{bmatrix} 0 & b_\omega(s) \\ b_\omega(s) & 0 \end{bmatrix} = \begin{bmatrix} b_\omega(s) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b_\omega(s) & 0 \\ 0 & 1 \end{bmatrix} \tag{4.16}$$

$$\begin{bmatrix} b_\omega(s) & 0 \\ 0 & b_\omega(s) \end{bmatrix} = B_{w,u}(s)^T B_{w,u}(s), \quad u = \begin{bmatrix} \frac{1}{2} \\ -\frac{i}{2} \end{bmatrix} \tag{4.17}$$

$$\begin{bmatrix} b_\omega(s) & 0 \\ 0 & -b_\omega(s) \end{bmatrix} = B_{w,u}(s)^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} B_{w,u}(s), \quad u = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \tag{4.18}$$

If some partial multiplicity is greater than 1, say $\sigma_i(\omega)$, and $V(s)$ is the left unimodular matrix in the Smith-McMillan factorization of $T(s)$, then we may choose for u the ith column vector of $V(\omega)$. Otherwise, the Takagi factorization of $T'(\omega)$ provides the vector u.

The assertion 2. $\Rightarrow$ 3. can be viewed as a particular case of the Schur algorithm for symmetric lossless (or inner) functions described in [OHP05].

Theorem 4.3.1 (Minimal degree symmetric extension) *Let $S(s)$ be a symmetric Schur function, strictly contractive at infinity and let $\widetilde{\Sigma}(s)$ be its minimal extension (4.11) with $\tilde{S}_{21}(s)$ minimum phase ; define $Q := \tilde{S}_{21}^{-1}\tilde{S}_{12}^T$, and let κ be the number of distinct zeros of* $\det Q(s)$ *with odd algebraic multiplicity. Then $S(s)$ has a symmetric inner completion of degree $n+\kappa$. This extension of $S(s)$ has minimal degree among all the symmetric extensions of $S(s)$.*

This result is proved by a recursive application of the symmetric Potapov factorization (Lemma 4.3.2). Since $Q(s)$ has $n - n_0 = \kappa + 2l$ zeros, we can perform l iterations. The zeros of $Q(s)$ lying within the m first columns of $\tilde{\Sigma}(s)$, the Potapov factors are all of the form $B_{w,u}(s) \oplus I$. We finally get

$$\tilde{\Sigma} = \begin{bmatrix} B^T & 0 \\ 0 & I \end{bmatrix} \check{\Sigma} \begin{bmatrix} B & 0 \\ 0 & I \end{bmatrix},$$

where $B(s)$, the product of the Potapov factors, has degree l and $\check{\Sigma}(s)$ is a symmetric lossless extension of $S(s)$ of degree $n+\kappa$.

In [2] it was shown that it is possible to construct a symmetric extension of exact degree n by increasing the size of the extension to $2m+n$. This result can be significantly improve. Indeed, we may extend the inner matrix (4.11) into a $(2m+1)\times(2m+1)$ matrix of degree $n+2(n-n_0)$

$$\hat{\Sigma} := \begin{bmatrix} \tilde{S}_{11}Q & 0 & \tilde{S}_{12} \\ 0 & \det Q & 0 \\ \tilde{S}_{12}^T & 0 & S \end{bmatrix}.$$

The matrix $Q \oplus \det Q$ has exactly $n-n_0$ double zeros. But then, in view of Lemma 4.3.2, we can obtain a reduction of degree by $2(n-n_0)$.

Theorem 4.3.2 (Minimal dimension symmetric extension) *Let $S(s)$ be a strictly contractive symmetric $m\times m$ Schur function of McMillan degree n. Then $S(s)$ has a symmetric inner extension of dimension $(2m+1)\times(2m+1)$ and McMillan degree n.*

4.4 Perspectives

It should be stressed that the above results rely on the fact that we work over the complex field. The situation for real coefficients functions is more complicated. Example (4.17 with w real) shows that the symmetric Potapov factorization may require complex Potapov factors even if the starting function is real. In the real case, an extra increase of the extension degree (or size) is necessary which depends on some indices attached to the real zeros of $S(s)$ [BEGO10].

From the applications viewpoint, the derivation of efficient algorithms for the synthesis of microwave multiplexers calls for a nice polynomial description of a $m\times m$ scattering matrix which generalizes (4.3). Theorem 4.3.2 provides an interesting method to built a 3×3 lossless matrix from a Schur scalar function p/q, where $p(s)$ and $q(s)$ are polynomials. It can be proved that any 3×3 lossless extension can be obtained in this way. This approach provides a polynomial description in which some divisibility properties are involved. This result answers a long standing question in the filter community: under some assumptions, the scattering 3×3 matrix is completely determined from one of its diagonal entries (a reflection). However, since the symmetric Potapov factorization can be performed in many ways, several extensions can be built from the same pair (p,q). Their numbers and their inter-connections depends on the number of double zeros of the stable polynomial $r(s)$ in the spectral factorization (4.2). These issues are currently under investigation.

From a practical synthesis point of view the extension process that starts with the numerator polynomials p_1 and p_2 of the non-diagonal entries (transmissions) of a row, is more relevant. However, this approach remains technically problematic: some issues concerning the stability of the derived polynomial $q(s)$ are still unsolved for the moment. Such a description is available for some particular form of the polynomial

model in connection with some special circuit topologies currently used for the implementation of the diplexers. This form corresponds to a scattering matrix $S(s)$ for which the matrix $M(s)$ in (4.7) is diagonal. The polynomial model is then closely related to the internal structure of the filter and provides an efficient description in view of frequency design and identification purposes. Promising algorithms are being developed and implemented along this line.

Publications

[BCO91] L. Baratchart, M. Cardelli, and M. Olivi. Identification and rational L^2 approximation: a gradient algorithm. *Automatica*, 27(2):413–418, 1991.

[BEGO07] L. Baratchart, P. Enqvist, A. Gombani, and M. Olivi. Minimal symmetric Darlington synthesis. *MCSS*, 4:283–311, 2007.

[BEGO10] L. Baratchart, P. Enqvist, A. Gombani, and M. Olivi. Minimal symmetric Darlington synthesis: the real case. In *Proceedings of the MTNS, Budapest, Hungary*, 2010. To appear.

[BGL+98] L. Baratchart, J. Grimm, J. Leblond, M. Olivi, F. Seyfert, and F. Wielonsky. Identification d'un filtre hyperfréquence par approximation dans le domaine complexe. Technical Report 219, INRIA, 1998.

[BO91] L. Baratchart and M. Olivi. Inner-unstable factorization of stable rational transfer functions. In G.B. Di Masi, A. Gombani, and A. B. Kurzhansky, editors, *Modeling, estimation and control of systems with uncertainty*. Birkhäuser, Boston, 1991.

[BO98] L. Baratchart and M. Olivi. Critical points and error rank in best H^2 matrix rational approximation. *Constructive Approximation*, 14:273–300, 1998.

[DMO05] R. Drai, J.P. Marmorat, and M. Olivi. A chain-scattering approach to LMI multi-objective control. In *16th IFAC World Congress*, Prague, July 2005.

[FO98] P. Fulcheri and M. Olivi. Matrix rational H^2-approximation: a gradient algorithm based on Schur analysis. *SIAM Journal on Control and Optimization*, 36(6):2103–2127, 1998.

[HOP06] B. Hanzon, M. Olivi, and R.L.M. Peeters. Balanced realizations of discrete-time stable all-pass systems and the tangential Schur algorithm. *Linear Algebra and its Applications*, 2006.

[HOP09] B. Hanzon, M. Olivi, and R.L.M. Peeters. Subdiagonal pivot structures and associated canonical forms under state isometries. In *Proceedings of SYSID, St Malo, France*, 2009.

[MO04] J. P. Marmorat and M. Olivi. RARL2: a Matlab based software for H^2 rational approximation, 2004.

[MO07] J.-P. Marmorat and M. Olivi. Nudelman interpolation, parametrizations of lossless functions and balanced realizations. *Automatica*, 43:1329–1338, 2007.

[MOHP02] J.-P. Marmorat, M. Olivi, B. Hanzon, and R.L.M. Peeters. Matrix rational H^2 approximation: a state-space approach using Schur parameters. In *Proceedings of the CDC, Las-Vegas, USA*, 2002.

[OHP05] M. Olivi, B. Hanzon, and R. Peeters. A Schur algorithm for symmetric inner functions. In *Proceedings of the CDC, Seville, Spain*, 2005.

[OHP08] M. Olivi, B. Hanzon, and R.L.M. Peeters. Lossless scalar functions: boundary interpolation, Schur algorithm and Ober's canonical form. In *Proceedings of the CDC, Cancun, Mexico*, 2008.

[PHO01] R. Peeters, B. Hanzon, and M. Olivi. Linear fractional transformations and balanced realizations of discrete time stable allpass systems. In *1st IFAC Symposium on Systems Structure and Control, Prague*, 2001.

[PHO02] R.L.M. Peeters, B. Hanzon, and M. Olivi. On a recursive state-space method for discrete-time H^2-approximation. In *Prooceedings of the MTNS*, Notre-dame, 2002.

[PHO07] R. Peeters, B. Hanzon, and M. Olivi. Canonical lossless state-space systems: staircase forms and the Schur algorithm. *Linear Algebra and its Applications*, 425:404–433, 2007.

Most of these papers are available on the web page:
`http://www-sop.inria.fr/members/Martine.Olivi/publis.html`

Bibliography

[1] D. Alpay, L. Baratchart, and A. Gombani. On the differential structure of matrix-valued rational inner functions. *Operator Theory: Advances and Applications*, 73:30–66, 1994.

[2] B.D.O Anderson and V. Vongpanitlerd. *Network analysis and synthesis- A modern system theory approach*. Prentice Hall, 1973.

[3] A. Antoulas. *Approximation of large-scale dynamical systems*. SIAM, 2005.

[4] J.A. Ball, I. Gohberg, and L. Rodman. *Interpolation of rational matrix functions*, volume 45 of *Operator Theory: Advances and Applications*. Birkhäuser, 1990.

[5] L. Baratchart. Existence and generic properties for L^2 approximants of linear systems. *I.M.A. Journal of Math. Control and Identification*, 3:89–101, 1986.

[6] L. Baratchart, J. Leblond, and J.-P. Marmorat. Sources identification in 3D balls using meromorphic approximation in 2D disks. *Electronic Transactions on Numerical Analysis (ETNA)*, 25:41–53, 2006.

[7] L. Baratchart and M. Yattselev. Convergent interpolation to Cauchy integrals over analytic arcs with Jacobi type weights. Submitted for publication.

[8] R. Bassila, J.P. Marmorat, J. Leblond, and M. Clerc. FINDSOURCE3D: a Matlab software prototype dedicated to the resolution of inverse problems of source detection, 2009.

[9] V. Belevitch. Elementary applications of scattering formalism to network design. *IRE Trans. Circuit Theory*, 1956.

[10] A. Ben-Artzi and I. Gohberg. Orthogonal polynomials over Hilbert modules. In *Operator Theory: Advances and Applications*, pages 96–126. Birkhäuser Verlag, 1994.

[11] S. Bila, R.J. Cameron, P. Lenoir, V. Lunot, and F. Seyfert. Chebyshev synthesis for multi-band microwave filters. In *2006 IEEE MTT-S International Microwave Symposium Digest*, pages 1221–1224, 2006.

[12] A. Bultheel, P. Gonzales-Vera, E. Hendriksen, and O. Njstad. *Orthogonal Rational Functions*. Cambridge University Press, 1999.

[13] C.I. Byrnes, T.T. Georgiou, and A. Lindquist. A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. *IEEE Trans. Autom.Control*, 46(6):822–839, 2001.

[14] C. Coelho, L. Silveira, and J. Phillips. Passive constrained rational approximation algorithm using nevanlinna-pick interpolation. In *Proceedings of the conference on Design, automation and test in Europe*, page 923, 2002.

[15] S. Darlington. Synthesis of reactance 4-poles which produce a prescribed insertion loss characteristics. *J. Math. Physics*, 18:257–355, 1939.

[16] C.A. Desoer and J.D. Schulman. Zeros and poles of matrix transfer functions and their dynamical interpretation. *IEEE Trans. Circuits Systems*, CAS-21:3–8, 1974.

[17] R. Douglas, H. Shapiro, and A. Shields. Cyclic vectors and invariant subspaces for the backward shift operator. *Annales de l'Institut Fourier (Grenoble)*, 20:37–76, 1970.

[18] H. Dym. *J-contractive matrix functions, reproducing kernel spaces and interpolation*, volume 71 of *CBMS lecture notes*. American Mathematical Society, Rhode Island, 1989.

[19] H. Dym. Shifts, realizations and interpolation, redux. *Operator Theory : Advances and Applications*, 73:182–243, 1994.

[20] H. Dym and D. Arov. *J-contractive matrix valued functions and related topics*. Cambridge University Press, 2008.

[21] H. Dym and N. Young. A Schur-Cohn theorem for matrix polynomials. *Proceedings of the Edinburgh Mathematical Society*, 33:337–366, 1990.

[22] P. Fuhrmann. Algebraic system theory: an analyst point of view. *Journal of Franklin institute*, 1976.

[23] P. Fuhrmann. *Linear systems and operators in Hilbert spaces*. Mc Graw Hill, 1981.

[24] H. Garnier and L. Wang. *Identification of continuous time models from sampled data*. Springer, 2008.

[25] Y. Genin, P. Van Dooren, T. Kailath, J.-M. Delosme, and M. Morf. On Σ-lossless transfer functions and related questions. *Linear Algebra and its Applications*, 50:251–275, 1983.

[26] M. Gevers. ARMA models, their Kronecker indices and their McMillan degree. *Int. J. Control*, 43(6):1745–1761, 1986.

[27] K. Glover. All optimal Hankel norm approximations of linear multivariable systems and their L^∞-error bounds. *Int. J. Contr.*, 39:1115–1193, 1984.

[28] A. Gombani and Gy. Michaletzky. On the nevanlinna-pick interpolation problem: analysis of the mcmillan degree of the solutions. *Linear Algebra and Applications*, 425:486–517, 2007.

[29] J. Grimm. Rational approximation of transfer functions in the Hyperion software. Technical Report 4002, INRIA, 2000.

[30] B. Hanzon and R.J. Ober. Overlapping block-balanced canonical forms for various classes of linear systems. *Linear Algebra and its Applications*, 281:171–225, 1998.

[31] B. Hanzon and R.L.M. Peeters. Balanced parametrizations of stable SISO all-pass systems in discrete-time. *MCSS*, 13:240–276, 2000.

[32] H. Helson. *Lectures on invariant subspaces*. Academic press, 1964.

[33] R. A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[34] N. Jacobson. *Lecture in abstract algebra*, volume II. Linear Algebra. Springer, New York, 1952.

[35] T. Kailath. *Linear Systems*. Prentice-Hall, 1980.

[36] R. Kalman, P. Falb, and A. Arbib. *Topics in Mathematical System Theory*. McGraw-Hill, New-York, 1969.

[37] R.E. Kalman. Mathematical description of linear dynamical systems. *SIAM journal on control*, 1(2):152–192, 1963.

[38] J. Karel, S. Haddad, R. Westra, W. Serdijn, and R. Peeters. Implementing wavelets in continuous-time analog circuits with dynamic range optimization. *IEEE Trans. on circuits and systems*, 2009.

[39] H. Kimura. Directional interpolation approach to H^∞-optimization and robust stabilization. *IEEE transactions on automatic control*, AC-32(12):1085–1093, 1987.

[40] H. Kimura. *Chain-Scattering Approach to H^∞-Control*. Birkhäuser, 1997.

[41] S.-Y. Kung and D.W. Lin. Optimal hankel-norm model reductions: Multivariable systems. *IEEE trans. on Automatic control*, AC-26(4):832–852, 1981.

[42] P. Lancaster and L. Rodman. *Algebraic Riccati Equations*. Oxford University Press, 1995.

[43] H. Lev-Ari and T. Kailath. State-space approach to factorization of lossless transfer functions and structured matrices. *Linear Algebra and its Applications*, 162–164:273–295, 1992.

[44] A. Linquist, C. Byrnes, and T. Georgiou. A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. *IEEE Trans. Automatic Control*, AC-46:822–839, 2001.

[45] A. Linquist and T. Georgiou. Kullback-Leibler approximation of spectral density functions. *IEEE Transactions on Information Theory*, pages 2910–2917, 2003.

[46] V. Lunot, S. Bila, and F. Seyfert. Optimal synthesis for multi-band microwave filters. In *2007 IEEE MTT-S International Microwave Symposium Digest*, pages 115–118, 2007.

[47] V. Lunot, F. Seyfert, S. Bila, and A. Nasser. Certified computation of optimal multi-band filtering functions. *IEEE Transactions on Microwave Theory and Techniques*, 56(1):105–112, 2008.

[48] Vincent Lunot. *Techniques d'approximation rationnelle en synthèse fréquentielle : problème de Zolotarov et algorithme de Schur*. PhD Thesis, Université de Provence, 2008.

[49] A. Macfarlane and N. Karcanias. Poles and zeros of linear multivariable systems : a survey of the algebraic, geometric and complex-variable theory. *Int. J. Control*, 24(1):33–74, 1976.

[50] B. Macmillan. Introduction to formal realizability theory. *Bell. Syst. Tech. J.*, 31, 1952.

[51] B.C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, AC-26:17–32, 1981.

[52] R. Ober. Asymptotically stable allpass transfer functions: canonical form, parametrization and realization. In *Proc. of the IFAC World Congress, Munich*, 1987.

[53] L. Pernebo and L.M. Silverman. Model reduction via balanced state space representations. *IEEE Trans. Autom. Control*, pages 382–387, 1982.

[54] R. Pintelon and J. Schoukens. *System Identification*. IEEE Press, 2001.

[55] V.P. Potapov. The multiplicative structure of *J*-contractive matrix functions. *Amer. Math. Soc. Transl.*, 15(2):131–243, 1960.

[56] H. Rosenbrock. *State-space and multivariable theory*. London: Thomas Nelson and Sons, 1970.

[57] C. Scherer. An efficient solution to multi-objective control problems with LMI objectives. *System and Control Letters*, pages 75–84, 2000.

[58] W. Schilders, H. van der Vorst, and J. Rommes. *Model Order Reduction*. Springer, 2008.

[59] I. Schur. On power series which are bounded in the interior of the unit circle. *I. Schur. methods in operator theory and signal processing, Operator theory : advances and applications*, 1986. Translation from J. Reine Angew. Math. 147, 205-232 (1917).

[60] F. Seyfert. PRESTO-HF: a Matlab based toolbox dedicated to the identification problem of low pass coupling parameters of band pass microwave filters, 2004.

[61] M. Spivak. *Calculus on manifolds: a modern approach to classical theorems of advanced calculus*. Harper Collins, 1965.

[62] J. Tokarzewski. *Finite zeros in discrete time control systems*. Springer-Verlag, 2006.

[63] P. Vaidyanathan and Z. Doganata. The role of lossless systems in modern digital signal porcessing. *IEEE Transaction on Education*, 32(3), 1989.

[64] J.C. Willems. In control, almost from the beginning until the day after tomorrow. *European Journal of Control*, 13:71–81, 2007.

[65] B. Wyman, K. Sain, G. Conte, and A.-M. Perdon. On the zeros and poles of a transfer function. *Linear Algebra and its Applications*, 124:123–144, 1989.

[66] B. Wyman, K. Sain, G. Conte, and A.-M. Perdon. Poles and zeros of matrices of rational functions. *Linear Algebra and its Applications*, 157:113–139, 1991.

[67] D.C. Youla. On the factorization of rational matrices. *IRE Transactions on Information Theory*, 7:172–189, 1961.