Université Catholique De Louvain
Faculté des sciences appliquées
Département d'ingénierie mathématique

Center for Systems Engineering and Applied Mechanics

Model Reduction of Linear Systems, an Interpolation point of View

Antoine Vandendorpe

Thesis submitted in partial fulfillment
of the requirements for the degree of
Docteur en sciences appliquées

Dissertation committee:
Vincent Blondel (chair)
Adhemar Bultheel
Yves Genin
Kyle Gallivan
Paul Van Dooren (Promotor)
Jan Willems

December 1, 2004

Acknowledgements

First of all, I would like to thank my promotor, Paul Van Dooren. It has been a pleasure, beside an honor, to collaborate with him. His knowledge, his enthusiasm and his friendship have always been a great support for me these last four years. He also gave me the opportunity to work with Yves Genin, and for this, I am also very grateful! I can only thank Yves for our collaboration these two last years. I will always remember our passionate discussions and his fruitful advices. I met Paul and Yves as a student, I leave them as a friend.

I would like to thank Kyle Gallivan for his support. My visit at Florida State University was the starting point of this thesis. After two weeks at FSU, I have already had enough material for two years of fruitful research. Many other researchers contributed to this work. To cite only a few, I would like to express my gratitude to Thanos Antoulas, Hugo Woerdeman and Brian Anderson for helpful discussions. Two other researchers contributed a lot to this thesis: Younes Chahlaoui and Damien Lemonnier. I really enjoyed to work with them in our *model reduction team* at CESAME.

I am indebted to all the members of the jury of this thesis, Profs. Adhemar Bultheel, Kyle Gallivan, Jan Willems and Vincent Blondel, for their critical reading which contributed to improve the quality of this manuscript.

This research is supported by the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. I also gratefully acknowledge support from the Belgian National Fund for Scientific Research.

The scientific athmosphere at the Centre for Systems Engineering and Applied Mechanics was particularly motivating. Many thanks to Dominique, Etienne, Gido, Isabelle, Lydia, Michou and Yannick for their administrative and technical assistance. I especially enjoyed the friendship of Yvan Hachez and Jacques Theys who both shared my office.

I dedicate this thesis to my wife, Alexandra, and our coming child.

Contents

Table of Contents ii

Notation v

Introduction 1

1 Preliminary Results 11
1.1 Linear Algebra 11
1.2 System Theory 19

2 Model Reduction of Linear Systems 27
2.1 Model Reduction by State Space Projection 28
2.2 Modal Approximation 30
2.3 Krylov Approximation Techniques 31
2.4 Gramian-based Techniques 35
2.5 Other Sparse Model Reduction techniques 37

3 Model Reduction by Moment Matching, the SISO case 41
3.1 Some facts about Krylov subspaces 41
3.2 Model Reduction via Rational Interpolation 48
3.3 Model Reduction via Truncation 53
3.4 Sylvester Equations, Krylov Subspaces and Projection 56
3.5 Concluding Remarks 58
3.6 Notes and References 59

4 Model Reduction by Moment Matching, the MIMO case . . . 61
4.1 Introduction 61
4.2 Preliminary results 66
4.3 Auxiliary results 71
4.4 The Multipoint Padé Reduced Order Transfer Function 78

4.5 Extension to Generalized State Space Realizations 83
4.6 Second Formulation of the Tangential Interpolation Problem . . . 85
4.7 Concluding remarks 88
4.8 Notes and References 89

5 On the Embedding of State Space Realizations 91
5.1 Introduction 91
5.2 A necessary rank condition 92
5.3 Necessary and sufficient conditions 93
5.4 The Pencil Approach 99
5.5 Link between the Pencil of Halevi and Tangential Interpolation . 104
5.6 Concluding remarks 105
5.7 Notes and References 105

6 Model Reduction of Second Order Systems 107
6.1 Introduction 107
6.2 Second order structure preserving model reduction 111
6.3 Second order Balanced Truncation 112
6.4 Second Order Structure Preserving Krylov Techniques 114
6.5 Numerical Experiments 121
6.6 Concluding Remarks 127
6.7 Notes and References 128

7 Model Reduction of Interconnected Systems 129
7.1 Introduction 129
7.2 Interconnected Systems Balanced Truncation 131
7.3 Krylov techniques for interconnected systems 135
7.4 Examples of Structured Model Reduction Problems 140
7.5 Concluding Remarks 143
7.6 Notes and References 143

Conclusion 145

References 149

Notation glossary

Basics

$\mathbb{N}$	set of nonnegative integer numbers
$\mathbb{N}_0$	set of strictly positive integer numbers
$\mathbb{Z}$	set of integer numbers
$\mathbb{R}$	field of real numbers
$\mathbb{C}$	field of complex numbers
j	complex unit, $j = \sqrt{-1}$
$\mathcal{R}e(z)$	real part of z
$\mathcal{I}m(z)$	complex part of z
$\bar{z}$	complex conjugate of z
$\lvert z \rvert$	modulus of z
$\Longleftrightarrow$	if and only if
$:=$	equal by definition to
δ_{ij}	Kronecker delta: equals 1 if $i = j$ and zero otherwise
$\dim X$	dimension of X
$< \cdot, \cdot >$	generic inner product
$\gcd(\cdot, \cdot)$	Greatest Common Divisor
$\lVert \cdot \rVert_p$	p-norm ($1 \leq p \leq +\infty$)
$\lVert \cdot \rVert_2$	Euclidean norm (vectors) / spectral norm

Matrix theory

I_k	$k \times k$ identity matrix
X^T	transpose of matrix X
X^*	conjugate transpose of matrix X
$x_{i,j}$	element located at the ith row and the jth column of X
$x_{i,:}$	i^{th} row of the matrix X
$x_{:,j}$	j^{th} column of the matrix X
$Im(X)$	subspace generated by the columns of X
$Ker(X)$	$\{x \in \mathbb{C}^m : Ax = 0\}$
$vec(X)$	vector formed by stacking the columns of X into one vector
$rank(X)$	rank of matrix X
$\det X$	determinant of square matrix X
$trace(X)$	trace of square matrix X
$\lambda_k(X)$	k-th eigenvalue of matrix X
$\lambda_{\min}(X)$	eigenvalue of X of smallest real part
$\lambda_{\max}(X)$	eigenvalue of X of largest real part
$\Lambda(X)$	set of eigenvalues of the matrix X
$\rho(X)$	$\max_i \lvert\lambda_i(X)\rvert$
$J_{w,\delta}$	Jordan matrix of size δ with eigenvalue w
$J_{w,\delta,k}$	Block Jordan matrix of size $k\delta$ with eigenvalue w
$\sigma_{\max}(X)$	maximal singular value of matrix X
$\sigma_{\min}(X)$	minimal singular value of matrix X
$\otimes$	Kronecker product

System Theory

$T_{i,j}(s)$	scalar rational function located at the position (i,j) of $T(s)$
$normal\ rank\ (T(s))$	maximal rank of $T(s)$ over $\mathbb{C}$
$Z_{C,A,B}(s)$	system zero matrix of the state space realization (C,A,B)

Abbreviations and acronyms

SISO	Single Input Single Ouptut
MIMO	Multiple Input Multiple Output
BT	Balanced Truncation
SOBT	Second Order Balanced Truncation
ISBT	Interconnected System Balanced Truncation

Introduction

These days, the modelling of physical processes gives rise to mathematical systems of increasing complexity. Good mathematical models have the following characteristics. On the one hand, they have to reproduce the physical process as precisely as possible. On the other hand, they have to be used efficiently, i.e. the computing time and the storage resources needed to simulate the mathematical model are limited. As a consequence, there must be a tradeoff between accuracy and computational constraints (speed of computation and storage limitation).

At the present time, one is often faced with systems that have an unacceptably high level of complexity. It is then desirable to approximate such systems by systems of lower complexity. This is the Model Reduction Problem.

Examples

Model reduction of large scale dynamical systems has received a lot of attention during the last decade : it is a crucial tool in reducing the computational complexity of e.g. analysis and design of Micro-Electro-Mechanical Systems (MEMS) [67], in simulation of electronic devices [24], in weather prediction [25] and in control of partial differential equations [55].

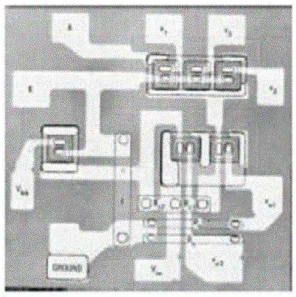
We motivate this with a few examples.

VLSI Circuits

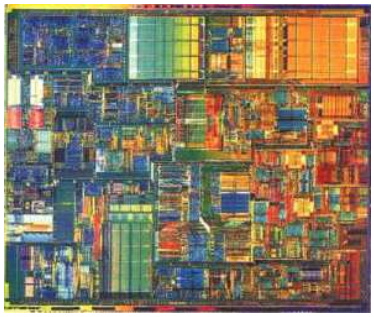
At the present time, the chip complexity of high frequency sub-micron VLSI circuits increases by at least 50 percent each year. In Figure 0.1, one of the first circuits produced in the sixties is compared to the more recent Pentium IV processor. Due to the increase in operating frequency and the decrease of metal width, simulations are required to verify that there is no significant delay or distortion of the circuit signal. Simulation of such complex *interconnected*

systems gives rise to models (obtained by spatial discretization of Maxwell's equations) resulting in a system of $n \approx 10^6$ equations. In Chapter 6, a more

Fig. 0.1. VLSI Circuits



Integrated Circuit, 1960



Pentium IV, 2001

detailed description of the equations appearing in RLC systems is given. Such large scale systems are in practice approximated using interpolation techniques [4, 28, 31, 9].

Weather Forecasting

The following more general problem is solved using ideas similar to the model reduction techniques presented in this thesis. The problem consists roughly of approximating large scale dense matrices by low rank ones in order to reduce the computational complexity.

In order to prevent flooding in the part of the Netherlands below sea level, one wants to predict rapidly storm surges in the North Sea. The equations governing the evolution of the wave surge are in this case the *shallow water* equations (see Figure 0.2 for a justification of *shallow*). In the same figure, the horizontal and vertical axes indicate the number of discretization points, while the color on the right-hand side indicates the depth of the sea at various locations of interest. Part of the discretization grid used in this case is shown in figure 0.3. The problem is to predict the wave surge in the North Sea based on a few localized measurements. This is achieved by means of a Kalman Filter which reconstructs the full state using a few noisy observations.

The Finite Element discretization of the shallow water equations yields roughly 60000 equations, and the resulting computational time is several times the allowed limit of 6 hours. Therefore reduced order modelling techniques are needed.

Fig. 0.2. Depth of the North Sea

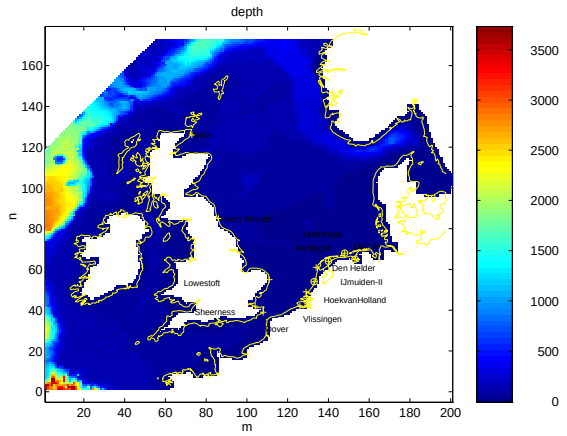
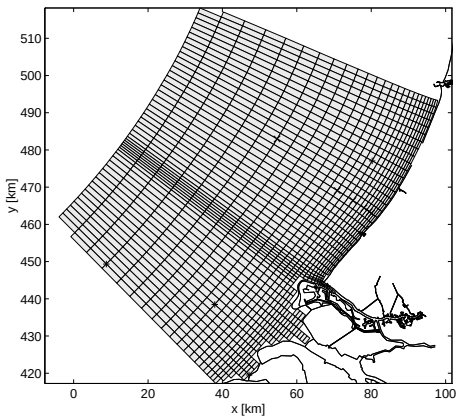


Fig. 0.3. Wave surge prediction problem: discretization grid close to the coast



The discretized state vector $q(t) = \begin{bmatrix} h(t) \; u(t) \; v(t) \end{bmatrix}$ is composed of three components: the water level $h(t)$ and the current velocities $u(t), v(t)$. The generic state space system to be considered is the following:

$$\dot{q}(t) = F(q(t), w(t)),$$
$$y(t) = G(q(t), z(t)),$$

where the output $y(t)$ is the set of measurements, $z(t)$ is the output noise and $w(t)$ is the set of (stochastic) inputs. A linearized version of this provides a

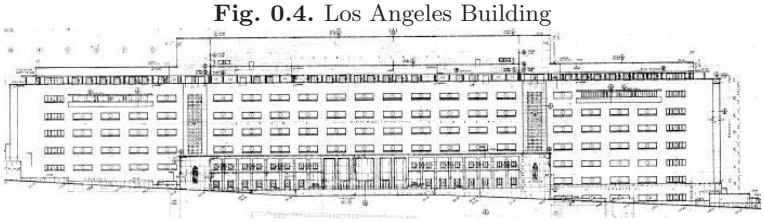
linear time varying system for which the state $q(t)$ is evaluated using Kalman filtering. In order to meet the computational constraints, the covariance matrix is approximated by a low rank approximation which is adapted at each time step. This amount essentially to time varying *model reduction*. This problem has been studied by Verlaan and Heemink at Delft University in the Netherlands; for details see [88]. A recent and more general study of model reduction techniques for linear time varying systems may be found in [18].

Building Active Damping Control

These days, active damping devices facilitate reducing the vibration of the structure of large buildings. This is crucial in order to build large structures in areas of significant seismic activity. Dampers are based on Magneto-Rheological fluids with viscosity that changes in milliseconds, when exposed to a magnetic field. The modelling of the vibrating structure of large buildings can lead to very large systems of equations. In order to control in real time the dampers, reduced order models are required.

For instance, in Chapter 6, the vibrating structure of a building in Los Angeles gives rise to a system of $n \approx 50000$ equations. As we can see in Figure 0.4, this large building is composed of about 10 large floors. This building is

Fig. 0.4. Los Angeles Building



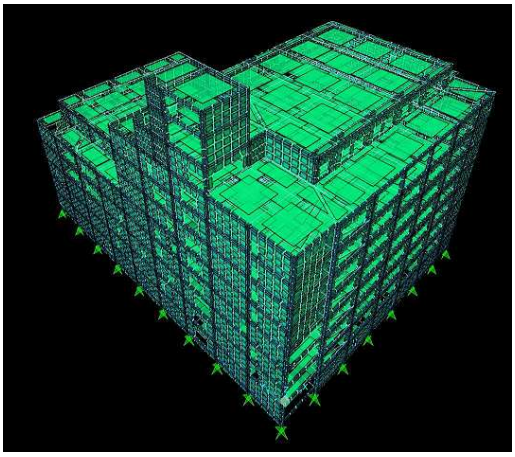
modelled as a set of girders. Each girder is discretized in two points: the top and the bottom of the girder. This corresponds for the above example to a state space of dimension $n \approx 50000$ (see Figure 0.5). The input $u(t)$ is the set of external forces applied to the system, in this case the force produced by an earthquake. The data used in this model in Chapter 6 are taken from the August 17th 1999 earthquake in Turkey.

This is a typical example of second order mechanical system. Indeed, the equation of position vector $q(t)$ of the system is of the form

$$M\ddot{q}(t) + D\dot{q}(t) + Sq(t) = F^{in}u(t),$$

where M, D and S are respectively the mass, damping and stiffness matrices. The output is chosen as an arbitrary position in the building.

Fig. 0.5. Modelling of the Building



As we will see, it is possible to approximate efficiently such complex systems with the Model Reduction techniques studied in this thesis. For more details, see section 6.5.

Formulation of the Problem

The systems considered in this thesis have the following form :

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases}$$

where the dimension n of the state vector $x(t)$ is very large. The objective of the Model Reduction procedure is to approximate such a system by another system

$$\begin{cases} \dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{B}u(t), \\ \hat{y}(t) = \hat{C}\hat{x}(t) + \hat{D}u(t), \end{cases}$$

where the dimension of the new state vector $\hat{x}(t)$ is $k \ll n$.

Such systems are represented by rational matrix functions $T(s) := C(sI_n - A)^{-1}B$ and $\hat{T}(s) := \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ linking the input $u(t)$ to the output $y(t)$ in the Laplace domain.

In this thesis, we choose to model a system with inputs and outputs. This is sometimes too restrictive. First, in some cases, the distinction between inputs and output is not *a priori* clear. Second, it is desirable to be able to treat the different representations of a given system in a unified way. In the *behavioral* setting, the basic variables considered are the *external* or *manifest* variables $w(t)$, which consists of $u(t)$ and $y(t)$, without distinction between them. The *collection of trajectories* describing the evolution of $w(t)$ over time defines a dynamical system (see [70]). This approach is not considered here.

The construction of the reduced order model typically passes via the derivation of one or two projective subspaces of the state space in which the original system is modelled. There are several approaches to find such projective subspaces. Depending on the size of the original system, two classes of model reduction techniques are competing.

For systems of moderate size, for which $O(n^3)$ algorithms can be used, SVD-techniques are preferred, such a the well known Balanced Truncation algorithm or the Optimal Hankel Norm Approximation algorithm. With these two gramian based methods, there exists a global a priori error bound between the original and the reduced order transfer functions.

When only $O(n^2)$ or $O(n)$ algorithms are possible, other techniques have been developed. In this area, interpolation techniques, also called Krylov techniques, are mostly used. These appear to work very well in practice, even if there exists no global error bound. Even worse, stability of the original system can be lost. It should be pointed out that if the state space realization at hand is dense, finite interpolation points are chosen and exact solvers are used, the complexity of Krylov techniques is also $O(n^3)$ but there exist very efficient iterative solvers, and large scale state space realization are mostly sparse, reducing considerably the computational complexity of such techniques. Another set of techniques of low complexity that approximate the gramians also exist. These are the Smith-ADI techniques. We will come back in more details on this in Chapter 2.

At the present time, active research is under way to find model reduction techniques that combine the qualities of both techniques, i.e. low complexity and global error bounds.

Thesis Outline

The objective of this thesis is twofold, corresponding to two different parts of the thesis.

First, there is a need for a better understanding of Krylov techniques. The exact potential of model reduction using interpolation conditions, and more generally state space projection, is studied in Chapters 3, 4 and 5. The second part of this thesis, corresponding to Chapters 6 and 7, concerns the study of structure preserving model reduction techniques.

In Chapter 3, the generality of Krylov techniques for model reduction of SISO systems is studied in details, giving rise to the following surprising (at least to us) result: any transfer function can be obtained by matching a set of interpolation points from any transfer function of larger Mc Millan degree. In other words, for SISO systems, Krylov approximation technique are universal. As a side effect, a link between Sylvester equations and Krylov subspaces is also described. This chapter corresponds to the following published papers :

- K. Gallivan, A. Vandendorpe and P. Van Dooren. Model Reduction via truncation : an interpolation point of view. *Linear Algebra Appl.*, 375:115-134, 2003. Preliminary version presented at IFAC World Congress 2002 [37].
- K. Gallivan, A. Vandendorpe and P. Van Dooren. Sylvester equations and projection-based model reduction. *J. Comp. Appl. Math., Special Issues,* 162:213-229, 2004.

In Chapter 4, a generalization of existing interpolation techniques for MIMO systems is developed. Instead of imposing interpolation conditions of the type

$$T(\lambda_i) = \hat{T}(\lambda_i), \qquad 1 \leq i \leq 2k,$$

more general tangential interpolation conditions can be imposed between the original and the reduced order system:

$$x_i T(\lambda_i) = x_i \hat{T}(\lambda_i), \quad T(\lambda_{i+k}) y_i = \hat{T}(\lambda_{i+k}) y_i, \qquad 1 \leq i \leq k,$$

where the x_i and y_i are respectively the left and right tangential interpolation directions. Such interpolation conditions appear naturally for MIMO systems when projecting via Sylvester equations. The corresponding journal paper is

- K. Gallivan, A. Vandendorpe and P. Van Dooren. Model reduction of MIMO systems via tangential interpolation. *SIAM Journal on Matrix Analysis and Applications*, 2004. Accepted. Preliminary version presented at MTNS 2002 [36].

In Chapter 5, the generality of projection techniques for model reduction of MIMO systems is studied. In contrast with the SISO case, given two $m \times p$ transfer functions, it is not always possible to construct a state space realization

of the transfer function of smaller Mc Millan degree by a projection of a state space realization of the transfer function of larger Mc Millan degree. Necessary and sufficient conditions for the embedding of two transfer functions are given for the case $n = k+1$. As shown by an example and an open problem, this topic requires further investigation. The following papers are related to this chapter :

- A. Vandendorpe and P. Van Dooren. Projection of state space realizations. contribution to the book : " *Unsolved problems in Mathematical Systems and Control Theory*", by V.D. Blondel and A. Megretski), Princeton University Press. 2004.
- Y. Genin and A. Vandendorpe. On the embedding of state space realizations, 2004. In preparation. Preliminary version presented at the SIAM Conference on Applied Linear Algebra [44].

This concludes the first part of this work.

Most large scale systems have a particular structure. They can be modelled as a set of subsystems that interconnect to each other. It then makes sense to develop model reduction techniques that preserve the structure of the original system. This is the objective of the second part of this thesis.

In Chapter 6, second order structure preserving model reduction techniques are developed. A new Second Order Balanced Truncation algorithm is derived from second order gramians. Krylov techniques are also developed in order to preserve the second order structure. These new techniques are then applied in a numerical example. Two publications are related to this chapter :

- Y. Chahlaoui, D. Lemonnier, A. Vandendorpe and P. Van Dooren. Second Order Balanced Truncation. *Linear Algebra Appl.*, Special Issue, 2003. Accepted. Preliminary version presented at MTNS 2002 [23].
- Y. Chahlaoui, K. Gallivan, A. Vandendorpe and P. Van Dooren. Model Reduction of Second Order Systems. In *Model Reduction of Dynamical Systems*, Eds. P. Benner et al., Springer Verlag, 2004. Submitted. Preliminary version presented at MTNS 2004 [22].
- A. Vandendorpe and P. Van Dooren. Krylov Techniques for Model Reduction of Second Order Systems. CESAME, Université catholique de Louvain, Technical Report 07-2004, 2004.

In Chapter 7, the second order structure preserving techniques are generalized to *interconnected systems*. It is shown that several structured model reduction techniques existing in the literature can be seen as particular case of the *Interconnected Systems Balanced Truncation* technique developed in that chapter. The following related paper will soon be submitted for publication :

- A. Vandendorpe and P. Van Dooren. Model Reduction of Interconnected Systems, 2004. In preparation. Preliminary version as conference paper published at MTNS 2004 [84].

For the reader that is not familiar with the Theory of Linear Systems, Chapter 1 introduces the basic concepts, with references to the literature. A survey on Model Reduction techniques is also given in Chapter 2.

For the interested reader, preliminary version of our publications are available online at the following webpage: http://www.auto.ucl.ac.be/∼vandendorpe/

1

Preliminary Results

In this chapter, the basic theory used throughout this thesis is introduced. For the sake of conciseness, proofs are omitted and references to the literature are given for the interested reader.

Section 1.1 reviews basic linear algebra results, with an emphasis on canonical forms of matrices and Section 1.2 introduces the Theory of Linear Systems.

1.1 Linear Algebra

Let us first introduce some notation. Let $\mathbb{R}$ denote the real scalar field and $\mathbb{C}$ the complex scalar field. Unless otherwise stated, all the matrices considered throughout this thesis have coefficients in the field $\mathbb{C}$, or in the polynomial ring $\mathbb{C}[s]$, in which case they are called $s-matrices$ and denoted $A(s), B(s), \ldots$, i.e. each element of the matrix is a polynomial in s. The preceding notation stands also for *rational* matrices. A rational matrix $A(s)$ is a matrix whose elements are rational functions, i.e. can be written as a fraction:

$$A_{i,j}(s) = \frac{\alpha_{i,j}(s)}{\beta_{i,j}(s)},$$

where $\alpha_{i,j}(s)$ and $\beta_{i,j}(s)$ are scalar polynomials. We use $A, B, \ldots$ exclusively for constant matrices. Greek letters $\alpha, \beta, \ldots$ are used for scalars in the field $\mathbb{C}$. For two scalar polynomials $\alpha(s)$ and $\beta(s)$, the symbol $\alpha(s)|\beta(s)$ means that $\alpha(s)$ divides $\beta(s)$, , i.e. there exists a scalar polynomial $\gamma(s)$ such that $\alpha(s)\gamma(s) = \beta(s)$. Given a matrix A, $a_{i,:}$ is the ith row, $a_{:,j}$ is the jth column and $a_{i,j}$ is the element located at the ith row and the jth column of A.

Linear Subspaces

Let $x_1, x_2, \ldots, x_k \in \mathbb{C}^n$. A vector of the form $\alpha_1 x_1 + \ldots + \alpha_k x_k$ is a *linear combination* over $\mathbb{C}$ of $x_1, x_2, \ldots, x_k$. The set of all linear combinations of

$x_1, x_2, \ldots x_k \in \mathbb{C}^n$ is a subspace called the *span* of $x_1, x_2, \ldots, x_k$, denoted by

$$span\{x_1, x_2, \ldots, x_k\} := \{x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k : \alpha_i \in \mathbb{C}\}.$$

Let a_i denote the i-th column of the matrix $A \in \mathbb{C}^{p\times m}$. The span of $a_1, \ldots, a_m$ is called the image of A and is denoted by

$$Im(A) := \{y = Ax : x \in \mathbb{C}^m\}.$$

A set of vectors $x_1, x_2, \ldots, x_k \in \mathbb{C}^n$ are said to be *linearly dependent* over $\mathbb{C}$ if there exist $\alpha_1, \ldots, \alpha_k \in \mathbb{C}$ not all zero such that $\alpha_1 x_1 + \ldots + \alpha_k x_k = 0$; otherwise they are said to be *linearly independent*.

Let S be a subspace of $\mathbb{C}^n$, then a set of vectors $\{x_1, \ldots, x_k\} \in S$ is called a *basis* for S if $x_1, x_2, \ldots, x_k$ are linearly independent and $S = span\{x_1, x_2, \ldots, x_k\}$. The number k is called the *dimension* of S, denoted by $\dim(S)$. The *kernel* of a matrix $A \in \mathbb{C}^{p\times m}$ is defined by

$$Ker(A) := \{x \in \mathbb{C}^m : Ax = 0\}.$$

The rank of a matrix A is defined by

$$rank(A) = \dim(Im(A)).$$

A matrix $A \in \mathbb{C}^{p\times m}$ is said to have *full row rank* if $p \leq m$ and $rank(A) = p$. Analogously, A is said to have *full column rank* if $m \leq p$ and $rank(A) = m$. A full rank square matrix is called a *nonsingular matrix*. Nonsingularity and invertiblility are two different notions. For instance, a polynomial matrix $P(s)$ such that its determinant is a polynomial of degree strictly larger than 0 is nonsingular but does not admit a polynomial inverse.

For any matrix $A \in \mathbb{C}^{m\times p}$, its conjugate transpose $A^* \in \mathbb{C}^{p\times m}$ is such that for any integer i, j such that $1 \leq i \leq p$, $1 \leq j \leq m$,

$$a^*(i, j) = \overline{a(j, i)}.$$

Definition 1.1. *A square matrix $U \in \mathbb{C}^{n\times n}$ whose columns form an orthonormal basis for $\mathbb{C}^n$ is called an unitary matrix (or orthogonal matrix for the real case), and it satisfies $U^*U = I_n = UU^*$.*

Definition 1.2. *A square matrix $A \in \mathbb{C}^{n\times n}$ is hermitian (or symmetric in the real case) if $A^* = A$. A square hermitian matrix $A = A^*$ is said to be positive definite (semi definite), denoted by $A > 0$ (≥ 0), if $x^*Ax > 0$ (≥ 0) for all $x \neq 0$.*

By extension, given two hermitian matrices $A, B \in \mathbb{C}^{n\times n}$, one write $A < B$ (≤ 0) if the hermitian matrix $B - A$ is (semi) positive definite, i.e. $B - A > 0$ (≥ 0).

The following lemma concerns the rank of the product of two matrices.

Lemma 1.3. *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times k}$. Then*

$$rank(A) + rank(B) - n \leq rank(AB) \leq \min\{rank(A), rank(B)\}.$$

Matrix Calculus

The following result, known as the Schur complement, is very useful.

Lemma 1.4. *Consider the matrix*

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

with $A \in \mathbb{C}^{n_1 \times n_1}$ *invertible,* $B \in \mathbb{C}^{n_1 \times n_2}$, $C \in \mathbb{C}^{n_2 \times n_1}$, $D \in \mathbb{C}^{n_2 \times n_2}$ *and with the additional assumption that*

$$\Gamma := D - CA^{-1}B,$$

is invertible. Then, the following matrix algebraic relation is well known:

$$M^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B\Gamma^{-1}CA^{-1} & -A^{-1}B\Gamma^{-1} \\ -\Gamma^{-1}CA^{-1} & \Gamma^{-1} \end{bmatrix}. \tag{1.1}$$

Let $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times q}$, the *Kronecker product* of A and B is defined as

$$A \otimes B := \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix} \in \mathbb{C}^{mp \times nq}.$$

Let $X \in \mathbb{C}^{m \times n}$ and let $vec(X)$ denote the vector formed by stacking the columns of X into one long vector:

$$vec(X) := \begin{bmatrix} x_{11} \; x_{21} \; \dots \; x_{m1} \; x_{12} \; x_{22} \dots x_{1n} \; \dots \; x_{mn} \end{bmatrix}^T.$$

Then, for any matrices $A \in \mathbb{C}^{k \times m}$, $B \in \mathbb{C}^{n \times l}$ and $X \in \mathbb{C}^{m \times n}$, the following relation holds [57]:

$$vec(AXB) = (B^T \otimes A)vec(X).$$

Consequently, if $k = m$ and $l = n$, then

$$vec(AX + XB) = (B^T \otimes I_m + I_n \otimes A)vec(X).$$

Let $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{m \times m}$, and let $\{\lambda_i, i = 1, \dots, n\}$ be the eigenvalues of A and $\{\mu_j, j = 1, \dots, m\}$ be the eigenvalues of B. Then the following properties hold:

1. The eigenvalues of $A \otimes B$ are the mn numbers $\lambda_i \mu_j$, $i = 1, \dots, n$, $j = 1, \dots, m$.
2. The eigenvalues of $B^T \otimes I_m + I_n \otimes A$ are the mn numbers $\lambda_i + \mu_j$, $i = 1, \dots, n$, $j = 1, \dots, m$.

The following lemma is a consequence of the preceding properties.

Lemma 1.5. *Consider the Sylvester equation*

$$AX + XB = C \tag{1.2}$$

where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{m\times m}$ *and* $C \in \mathbb{C}^{n\times m}$ *are given matrices. There exists a unique solution* $X \in \mathbb{C}^{n\times m}$ *if and only if* $\lambda_i(A) + \lambda_j(B) \neq 0$, $\forall i = 1, \ldots, n$ *and* $\forall j = 1, \ldots m$.

In particular, if $B = A^*$, (1.2) *is called the "Lyapunov Equation" and the necessary and sufficient condition for the existence of a unique solution is that* $\lambda_i(A) + \bar{\lambda}_j(A) \neq 0$, $\forall i, j = 1, \ldots, n$.

The proof consists in rewriting (1.2) as a linear equation in $vec(X)$ by using Kronecker products and imposing all the eigenvalues to be nonzero.

Canonical Forms

Depending on the set of transformations allowed, several canonical form of matrices have been developed. As we will see in the sequel, these canonical forms play a central role in Matrix Theory. For this reason, some of them are briefly recalled here.

The Eigenvalues and the Jordan Form

Let $A \in \mathbb{C}^{n\times n}$, then the eigenvalues of A are the n roots of its characteristic polynomial $p(s) := \det(sI - A)$. This set of roots is called the *spectrum* of A and is denoted by $\Lambda(A)$ (not to be confused with singular values defined later). The maximal modulus of the eigenvalues is called the *spectral radius*, denoted by

$$\rho(A) := \max_{i=1}^{n} |\lambda_i|.$$

If $\lambda \in \Lambda(A)$ then any nonzero vector $x \in \mathbb{C}^n$ that satisfies

$$Ax = \lambda x$$

is referred to as a *right eigenvector* of A. Analogously, a nonzero vector y is called a *left eigenvector* of A if

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

It is a well known fact in linear algebra that any complex matrix admits a *Jordan Canonical* representation:

Theorem 1.6. *For any square complex matrix $A \in \mathbb{C}^{n\times n}$, there exists an invertible matrix T such that*

$$A = TJT^{-1}$$

where

$$J = diag\{J_1, J_2, \ldots, J_l\}; \tag{1.3}$$

$$J_i = diag\{J_{i,1}, J_{i,2}, \ldots, J_{i,m_i}\}; \tag{1.4}$$

$$J_{i,j} = \begin{bmatrix} \lambda_i & -1 & & & \\ & \lambda_i & -1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & -1 \\ & & & & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_{i,j}\times n_{i,j}}; \tag{1.5}$$

with $\sum_{i,j} n_{i,j} = n$ and with $\{\lambda_i : i = 1, \ldots, l\}$ as the distinct eigenvalues of A.

Definition 1.7. *When the Jordan form of a matrix is diagonal, we call the matrix nondefective (or diagonalizable).*

It is known that nondefective matrices are dense in $\mathbb{C}^{n\times n}$. Another class of matrices is defined below.

Definition 1.8. *A square matrix $A \in \mathbb{C}^{n\times n}$ is called cyclic if the Jordan canonical form of A has one and only one Jordan block associated with each distinct eigenvalue.*

Another equivalent definition [92] is that a matrix $A \in \mathbb{C}^{n\times n}$ is cyclic if and only of there exists a vector $v \in \mathbb{C}^n$ such that the matrix

$$\begin{bmatrix} v & Av & \ldots & A^{n-1}v \end{bmatrix} \tag{1.6}$$

is nonsingular.

In other words, a matrix A is cyclic if its Jordan form has $m_i = 1, i = 1, \ldots, l$, where the integer m_i is defined as the number of Jordan blocks of eigenvalue λ_i (see (1.4)). It should be pointed out that if a matrix A is cyclic, then almost any nonzero vector v is such that the matrix in (1.6) is nonsingular [94].

Clearly, a square matrix A with all distinct eigenvalues is cyclic and diagonalizable :

$$A \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

In this case, each column of the similarity transformation matrix T that diagonalize A is a right eigenvector of A, and each row of T^{-1} a left eigenvector of A.

Definition 1.9. *A matrix $A \in \mathbb{C}^{n\times n}$ is called Hurwitz if all its eigenvalues have a strictly negative real part.*

Jordan matrices will play an important role in the sequel and we therefore introduce the following compact notation.

Definition 1.10. *The matrix $J_{w,\delta,k} \in \mathbb{C}^{k\delta\times k\delta}$ is defined to be*

$$J_{w,\delta,k} := \begin{bmatrix} wI_k & -I_k & & \\ & \ddots & \ddots & \\ & & \ddots & -I_k \\ & & & wI_k \end{bmatrix}. \tag{1.7}$$

When $k = 1$, $J_{w,\delta,1}$ is simply a Jordan matrix of size $\delta \times \delta$ at eigenvalue w and is written $J_{w,\delta}$.

The Smith Form

Let us first introduce the invariant polynomials and elementary divisors of a polynomial matrix $A(s)$ [41].

Definition 1.11. *Let $A(s)$ be a $p \times m$ polynomial matrix of rank r. We denote by $D_j(s)$ the greatest common divisor of all the minors of order j in $A(s)$ ($j = 1, 2, \ldots, r$). The corresponding quotients will be denoted by $i_1(s), i_2(s), \ldots, i_r(s)$:*

$$i_1(s) := \frac{D_r(s)}{D_{r-1}(s)}, \quad i_2(s) := \frac{D_{r-1}(s)}{D_{r-2}(s)}, \ldots, i_r(s) := D_1(s). \tag{1.8}$$

The polynomials $i_1(s), i_2(s), \ldots, i_r(s)$ defined by (1.8) are called the invariant polynomials of the rectangular matrix $A(s)$.

We decompose the invariant polynomials $i_1(s), i_2(s), \ldots, i_r(s)$ over irreducible factors over $\mathbb{C}$:

$$\begin{aligned} i_1(s) &= [\phi_1(s)]^{c_1}[\phi_2(s)]^{c_2}\ldots[\phi_t(s)]^{c_t}, \\ i_2(s) &= [\phi_1(s)]^{d_1}[\phi_2(s)]^{d_2}\ldots[\phi_t(s)]^{d_t}, \\ &\vdots \\ i_r(s) &= [\phi_1(s)]^{l_1}[\phi_2(s)]^{l_2}\ldots[\phi_t(s)]^{l_t}. \end{aligned} \tag{1.9}$$

Here $\phi_1(s), \phi_2(s), \ldots, \phi_t(s)$ are all the distinct factors irreducible over $\mathbb{C}$ (and with highest coefficient 1) that occur in $i_1(s), i_2(s), \ldots, i_r(s)$. The polynomials $\phi_i(s)$ have all a unique zero that is not equal to the zeros of the other polynomials $\phi_j(s)$, $j \neq i$.

Definition 1.12. *All the powers among* $[\phi_1(s)]^{c_1}, \ldots, [\phi_t(s)]^{l_t}$ *in* (1.9)*, as far as they are distinct from 1, are called the elementary divisors of the matrix* $A(s)$ *in* $\mathbb{C}$*.*

The invariant polynomials of the matrix $A(s)$ appear as diagonal elements of the Smith form of this matrix. Let us introduce the SMith form of a polynomial matrix.

Definition 1.13. *A polynomial matrix* $U(s) \in \mathbb{C}^{n\times n}[s]$ *is called unimodular if it is nonsingular and has a polynomial inverse.*

A necessary and sufficient condition for a matrix $U(s)$ to be unimodular is that $\det U(s)$ is a nonzero constant.

The definition of the canonical Smith form of a polynomial matrix $A(s)$ is the following [60],

Theorem 1.14. *For any matrix* $A(s) \in \mathbb{C}^{m\times n}[s]$*, there exists unimodular matrices* $M(s) \in \mathbb{C}^{m\times m}[s]$ *and* $N \in \mathbb{C}^{n\times n}[s]$ *that diagonalize* $A(s)$ *as follows :*

$$M(s)A(s)N(s) = \left[\begin{array}{ccc|c} e_1(s) & & & \\ & \ddots & & 0_{r,n-r} \\ & & e_r(s) & \\ \hline & 0_{m-r,r} & & 0_{m-r,n-r} \end{array}\right];$$

where

$$e_1(s)|e_2(s)|\ldots|e_r(s).$$

It is not difficult to prove that $\forall iin\mathbb{N}:\ 1 \leq i \leq r,\ e_i(s) = i_{r-i+1}(s)$.

The Kronecker Form

For a pair of matrices $A, B \in \mathbb{C}^{m\times n}$, the set of vectors $x \in \mathbb{C}^n$ (resp. $y \in \mathbb{C}^m$) such that there exists $\lambda \in \mathbb{C}$ with

$$Ax = \lambda Bx \qquad (resp.\ y^T A = y^T B\lambda), \tag{1.10}$$

are called *generalized right (resp. left) eigenvectors* of the pair (A, B). The set of scalars $\lambda \in \mathbb{C}$ such that it is possible to find eigenvectors that satisfy (1.10) is the set of *generalized eigenvalues* of (A, B). In contrast with the classical eigenvalue problem, the set of generalized eigenvalues of (A, B) can be equal to $\mathbb{C}$. Even infinity can be seen as a generalized eigenvalue (take for instance $B = 0$ in (1.10)).

The canonical form associated with the generalized eigenvalue problem is called the Kronecker form and is briefly recalled below [82].

Theorem 1.15. *Let $A, B \in \mathbb{C}^{m\times n}$. There exist invertible matrices $S \in \mathbb{C}^{m\times m}$ and $T \in \mathbb{C}^{n\times n}$ such that the pencil $S(sB - A)T$ admits the following form :*

$$S(sB - A)T = diag\{0_{k_1,k_2}, L_{l_1}, \ldots, L_{l_s}, L^T_{r_1}, \ldots, L^T_{r_l}, I - sN, sI - J\}$$

where

1. *$0_{k_1,k_2}$ is the zero matrix of dimension $k_1 \times k_2$;*
2. *L_k is the $(k + 1) \times k$ bidiagonal pencil*

$$L_k := \begin{bmatrix} s & & & \\ -1 & s & & \\ & \ddots & \ddots & \\ & & -1 & s \\ & & & -1 \end{bmatrix};$$

3. *N is nilpotent, i.e. all its eigenvalues are zero;*
4. *N and J are in Jordan canonical form.*

The elementary divisors of $sI - J$ are called the *finite elementary divisors* of $sB - A$, the elementary divisors of $\mu I - N$ the *infinite elementary divisors* of $sB - A$, and the index sets $\{l_1, \ldots, l_s\}$ and $\{r_1, \ldots, r_t\}$ the nontrivial *left* and *right Kronecker indices* of $sB - A$. A pencil (A, B) is called *regular* if the matrices A and B are square and

$$\det(sB - A) \neq 0$$

for some values of $s \in \mathbb{C}$. The Kronecker form of regular matrix pencils is only composed of blocs of the form $sI - J$ and $I - sN$, with N, J Jordan matrices.

The Smith-Mc Millan Form

This is a natural extension of the Smith form for rational matrices [60]. Let us first define the notion of normal rank for rational matrices.

Definition 1.16. *Let $H(s)$ be a $p \times m$ rational matrix. The normal rank of $H(s)$ is the maximal rank of $H(s)$ for $s \in \mathbb{C}$. $H(s)$ is said to have full column (resp. row) normal rank if its normal rank is equal to m (resp. p).*

The normal rank of $H(s)$ is obtained for any value of s that excepted for a finite number of points, namely the transmission zeros and the poles of $H(s)$ (these are respectively the zeros of the polynomial $\epsilon_i(s)$ and $\psi_i(s)$ appearing below, see also Subsection 1.2).

Lemma 1.17. *Given any proper rational matrix $H(s)$ of normal rank r, there exist unimodular matrices $U(s)$, $V(s)$ such that*

$$U(s)H(s)V(s) = M(s) := \begin{bmatrix} diag\{\epsilon_i(s)/\psi_i(s)\} & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.11}$$

where the scalar polynomials ϵ_i and $\psi_i(s)$ are such that the pairs $\{\epsilon_i(s), \psi_i(s)\}$ are coprime (i.e. their greatest common polynomial divisor is 1), $i = 1, \ldots, r$ and

$$\epsilon_1(s)|\epsilon_2(s)|\ldots|\epsilon_r(s), \quad \psi_r(s)|\psi_{r-1}(s)|\ldots|\psi_1(s).$$

The matrix $M(s)$ is uniquely defined and is called the Smith-Mc Millan form of $H(s)$.

Let $d(s)$ be the monic least common multiple of the denominators of the entries of $H(s)$. The Smith Mc Millan form $M(s)$ of $H(s)$ can be deduced from the Smith form of the polynomial matrix $d(s)H(s)$.

1.2 System Theory

Let us consider the following generalized state-space model describing a finite dimensional continuous linear time invariant system

$$S \begin{cases} E\dot{x}(t) = Ax(t) + Bu(t), \\ \quad y(t) = Cx(t) + Du(t), \end{cases} \tag{1.12}$$

with input $u(t) \in \mathbb{C}^m$, state $x(t) \in \mathbb{C}^n$, output $y(t) \in \mathbb{C}^p$, and with system matrices A, B, C, D, E that belong to respectively $\mathbb{C}^{n\times n}$, $\mathbb{C}^{n\times m}$, $\mathbb{C}^{p\times n}$, $\mathbb{C}^{p\times m}$ and $\mathbb{C}^{n\times n}$. Discrete time systems (where the time variable t belongs to $\mathbb{Z}$ instead of $\mathbb{R}$) are not studied here. The time varying case (with matrices $(C(t), A(t), B(t))$ evolving in time) is not considered either (see for instance [18] and references therein for a recent study of model reduction techniques of such systems). The system (1.12) is related to its *transfer function*,

$$T(s) := C(sE - A)^{-1}B + D,$$

that links the inputs to the outputs in the Laplace domain. The element at the *ith* row and the *jth* column of $T(s)$ is denoted by $T_{i,j}(s)$. Transfer functions of linear time invariant systems (also called LTI systems) are *rational* matrix functions. A transfer function $T(s)$ is called *proper* if $\lim_{s\to\infty} T(s) < \infty$ and *strictly proper* if $\lim_{s\to\infty} T(s) = 0$.

When the matrix E is equal to the identity matrix, the model (1.12) is called a standard state space realization of $T(s)$ (or simply a state space realization). For the sake of simplicity, only standard state space realizations are considered

in this thesis. Nevertheless, it is shown how to extend our results to generalized state space realizations when necessary.

When a quadruple (a triple) of matrices (C, A, B, D) (or simply (C, A, B)) is such that the (strictly) proper transfer function $T(s)$ satisfies

$$T(s) = C(sI_n - A)^{-1}B + D \quad \big(T(s) = C(sI_n - A)^{-1}B\big)\,,$$

one says that the quadruple (or triple) *realizes* $T(s)$ and one writes

$$(C, A, B, D) \sim T(s) \quad (\textit{or simply } (C, A, B) \sim T(s))\,.$$

The dimension of the matrix D does not depend on the dimension of the state vector $x(\cdot)$. Typically, the matrix $\hat{D}$ of the reduced order system $\hat{C}(sI_k - \hat{A})^{-1}\hat{B} + \hat{D}$ is chosen to be equal to D. As a consequence, this matrix is generally assumed to be equal to zero in the model reduction problem. For this reason, we mainly consider standard state space realizations of strictly proper transfer functions, i.e. triples (C, A, B). They correspond thus to a system S of the following form :

$$S \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t). \end{cases} \tag{1.13}$$

The order of a state space realization (C, A, B) is defined as the dimension n of the square matrix $A \in \mathbb{C}^{n,n}$. Clearly, there exist infinitely many state space realizations corresponding to the same transfer function. The minimal order n over all the state space realization of a transfer function $T(s)$ is called the *Mc Millan degree* of $T(s)$.

Two well known matrices associated with a state space realization (C, A, B) of order n of a $p \times m$ transfer function $T(s)$ are the controllability matrix $Contr(A, B) \in \mathbb{C}^{n \times nm}$ and the observability matrix $Obs(C, A) \in \mathbb{C}^{pn \times n}$ defined by

$$Contr(A, B) := \begin{bmatrix} B \ \dots \ A^{n-1}B \end{bmatrix} \quad , \quad Obs(C, A) := \begin{bmatrix} C^T \ \dots \ (A^T)^{n-1}C^T \end{bmatrix}^T .$$

A pair of matrices (A, B) is called *controllable* if the controllability matrix $Contr(A, B)$ has full row rank. Analogously, a pair of matrices (C, A) is called *observable* if the observability matrix $Obs(C, A)$ has full column rank.

A well known result is the following.

Lemma 1.18. *A state space realization (C, A, B) is minimal if and only if (C, A) is observable and (A, B) is controllable.*

It turns out that any strictly proper rational matrix admits a minimal standard state space realization. The following result is well known [13, 94],

Lemma 1.19. *Let (C_1, A_1, B_1) and (C_2, A_2, B_2) be two minimal state space realizations of the transfer function $T(s)$ of Mc Millan degree n. Then, there exists an invertible matrix $S \in \mathbb{C}^{n \times n}$ such that*

$$(C_2, A_2, B_2) = (C_1 S, S^{-1}A_1 S, S^{-1}B_1).$$

Note that Lemma 1.19 is not true for nonminimal state space realizations.

The concept of zero

Let us consider a $p \times m$ strictly proper transfer function $T(s)$. By considering the Smith-Mc Millan form (1.11) of $T(s)$, it appears that the rank of $T(s)$ drops at points in the complex plane corresponding to the zeros of the polynomials $\epsilon_i(s)$. This leads to the concept of *transmission zero*.

Definition 1.20. *The roots of all the polynomials $\epsilon_i(s)$ in the Smith-Mc Millan form of $T(s)$ are called the transmission zeros of $T(s)$. A complex number $\lambda \in \mathbb{C}$ is called a blocking zero of $T(s)$ if $T(\lambda) = 0$.*

It is clear that a blocking zero is a transmission zero. Moreover, for a scalar transfer function, the blocking zeros and the transmission zeros are the same. A general interpretation of transmission zero is given by the following lemma,

Lemma 1.21. *Let $T(s)$ be a $p \times m$ proper transfer function with full column normal rank. Then $\lambda \in \mathbb{C}$ is a transmission zero of $T(s)$ if and only if there exists a nonzero vector $u \in \mathbb{C}^m$ such that $T(\lambda)u = 0$.*

For a simple proof, see for instance [94]. Similarly, we have the following lemma:

Lemma 1.22. *Let $T(s)$ be a $p \times m$ proper transfer function with full row normal rank. Then λ is a transmission zero of $T(s)$ if and only if there exists a nonzero $v \in \mathbb{C}^p$ such that $v^T T(\lambda) = 0$.*

In the case where the transmission zero is not a pole, we can give a useful alternative characterization of the transmission zeros. The following lemma is easy to show from the definition of zeros.

Lemma 1.23. *Suppose $\lambda \in \mathbb{C}$ is not a pole of $T(s)$. Then, λ is a transmission zero if and only if*

$$\textit{rank } T(\lambda) < \textit{normal rank } T(s).$$

Corollary 1.24. *Let $T(s)$ be a square $m \times m$ proper transfer function and $\det T(s) \not\equiv 0$. Suppose $\lambda \in \mathbb{C}$ is not a pole of $T(s)$. Then λ is a transmission zero of $T(s)$ if and only if $\det T(\lambda) = 0$.*

Let us introduce the concept of *invariant zero*.

Definition 1.25. *Let (C, A, B) be a state space realization of the $p \times m$ transfer function $T(s)$. A complex number $\lambda \in \mathbb{C}$ is called an invariant zero of the realization (C, A, B) if it satisfies*

$$\textit{rank} \begin{bmatrix} \lambda I - A & B \\ C & 0 \end{bmatrix} < \textit{normal rank} \begin{bmatrix} sI - A & B \\ C & 0 \end{bmatrix}.$$

The matrix

$$Z_{(C,A,B)}(s) := \begin{bmatrix} sI - A & B \\ C & 0 \end{bmatrix}$$

is called the system zero matrix of (C, A, B) *(see [72] for a more general concept of system zero matrix).*

It is clear that invariant zeros do not change under similarity transformation. The following lemma is obvious.

Lemma 1.26. *Suppose* $Z_{(C,A,B)}(s)$ *has full column rank. Then* $\lambda \in \mathbb{C}$ *is an invariant zero of* (C, A, B) *if and only if there exist* $0 \neq x \in \mathbb{C}^n$ *and* $u \in \mathbb{C}^m$ *such that*

$$\begin{bmatrix} \lambda I_n - A & B \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = 0.$$

Moreover, if $u = 0$*, then* λ *is also an unobservable mode.*

Proof. By definition, λ is an invariant zero if there is a nonzero vector $\begin{bmatrix} x \\ u \end{bmatrix} \neq 0$ such that

$$\begin{bmatrix} \lambda I_n - A & B \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = 0$$

since $Z_{C,A,B}(s)$ has full column rank.

On the other hand, suppose λ is an invariant zero, then there is a vector $\begin{bmatrix} x \\ u \end{bmatrix} \neq 0$ such that

$$\begin{bmatrix} \lambda I - A & B \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = 0.$$

We claim that $x \neq 0$. Otherwise, $Bu = 0$ or $u = 0$ since $Z_{C,A,B}(s)$ has full column rank. Finally, note that if $u = 0$, then

$$\begin{bmatrix} \lambda I_n - A \\ C \end{bmatrix} x = 0$$

and λ is an unobservable mode.

□

Lemma 1.27. $T(s)$ *has full column (row) normal rank if and only if* $Z_{(C,A,B)}(s)$ *has full column (row) normal rank.*

Proof. This follows by noting that

$$\begin{bmatrix} sI - A & B \\ C & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ C(sI - A)^{-1} & I \end{bmatrix} \begin{bmatrix} sI - A & B \\ 0 & T(s) \end{bmatrix},$$

and

$$\textit{normal rank } Z_{(C,A,B)}(s) = n + \textit{normal rank } T(s).$$

□

Theorem 1.28. *Let $T(s)$ be a rational proper transfer function and let the quadruple (C, A, B, D) be a corresponding minimal realization. Then a complex number λ is a transmission zero of $T(s)$ if and only if it is an invariant zero of the minimal realization.*

Norms

Before defining the system norms that are used here, let us first recall the definition of vector and matrix norms.

Vector Norm

Let X be a vector space, a real-valued function $\|\cdot\|$ defined on X is said to be a *norm* on X if it satisfies the following properties :

1. $\|x\| \geq 0$ (positivity)
2. $\|x\| = 0$ if and only if $x = 0$ (positive definiteness);
3. $\|\alpha x\| = |\alpha|\|x\|$, for any scalar α (homogeneity);
4. $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality);

for any $x, y \in X$.

Let $x \in \mathbb{C}^n$. The vector $p-$norm of x are defined as

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, \quad for\ 1 \leq p \leq \infty.$$

Matrix Norm

Let $A \in \mathbb{C}^{m\times n}$, then the matrix norm *induced* by a vector $p-$norm is defined as

$$\|A\|_p := \sup_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p}.$$

Another important matrix norm is the *Frobenius* norm defined as follows :

$$\|A\|_F = \sqrt{\sum_{i,j} |a_{i,j}|^2} = \sqrt{trace(A^*A)} = \sqrt{trace(AA^*)}.$$

The Singular Value Decomposition

A very useful tool in matrix analysis is the *Singular Value Decomposition.*

Theorem 1.29. *For any matrix $A \in \mathbb{C}^{m\times n}$, there exist unitary matrices $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ such that*

$$A = U\Sigma V^*; \tag{1.14}$$

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & O_{r\times(n-r)} \\ 0 & & \sigma_r & \\ \hline & O_{(m-r)\times r} & & O_{(m-r)\times(n-r)} \end{array}\right], \tag{1.15}$$

where the singular values σ_i are real and non increasing scalars :

$$\sigma_1 \geq \cdots \geq \sigma_r > 0. \tag{1.16}$$

For a matrix A, $\sigma_k(A)$ is defined as the $k-th$ largest singular value of A. This allows us to write the dyadic expansion

$$A = \sum_{i=1}^{r} u_i \sigma_i v_i^*,$$

where u_i (resp. v_i) is the $i-th$ column of U (resp; V) and r is the rank of A.

An important application of the singular value decomposition in approximation theory is the following theorem.

Theorem 1.30. *Let $A \in \mathbb{C}^{m\times n}$ be a rank r matrix. The best approximation of A by a matrix $B \in \mathbb{C}^{m\times n}$ of rank $s < r$ satisfies*

$$\min_{rank(B)\leq s} \|A - B\|_2 = \sigma_{s+1}(A).$$

This minimum is e.g. obtained by taking $B = \sum_{i=1}^{s} u_i \sigma_i v_i^*$.

System Norms

Let $T(s)$ be a $p \times m$ proper strictly stable (i.e. all its poles lie in the left half plane) transfer function. The classic H_2 and H_∞ norm of $T(s)$ are defined as

$$\begin{aligned} \|T(s)\|_{H_2} &:= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} trace(T^*(j\omega)T(j\omega))d\omega}, \\ \|T(s)\|_{H_\infty} &:= \sup_{\omega\in\mathbb{R}} \sigma_{max}\left(T(j\omega)\right). \end{aligned} \tag{1.17}$$

Another very important norm that appears in the model reduction framework is the *Hankel* norm. For this we introduce the concept of dual of an operator.

Definition 1.31. *Let L be a linear operator acting from a Hilbert space U to a Hilbert space Y equipped respectively with the inner products $< \ , \ >_U$ and $< \ , \ >_Y$. The dual of L, denoted by L^*, is defined as the linear operator acting from Y to U such that $< Lu, y >_Y = < u, L^*y >_U$ for all $y \in Y$ and all $u \in U$.*

A stable transfer function $T(s) := C(sI - A)^{-1}B$ is also a linear (convolution) operator mapping square integrable inputs $u(.) \in \mathcal{L}_2[-\infty, +\infty]$ to square integrable outputs $y(.) \in \mathcal{L}_2[-\infty, +\infty]$. It is clear from Definition 1.31 that the dual of $T(s)$ is given by $T^*(s) := B^*(sI - A^*)^{-1}C^*$. Indeed, for any input $u(.)$ and output $y(.)$,

$$< T(.)u(.), y(.) >= \int_{-\infty}^{\infty} u^*(j\omega)T^*(j\omega)y(j\omega)d\omega =< u(.), T^*(.)y(.) > .$$

Two matrices are associated with the linear system (1.13). These are the "controllability gramian" P and the "observability gramian" Q. If A is Hurwitz, they are the unique solutions of the following Lyapunov equations :

$$AP + PA^T + BB^T = 0 \quad , \quad A^TQ + QA + C^TC = 0. \tag{1.18}$$

If we apply an input $u(.) \in \mathcal{L}_2[-\infty, 0]$ to the system (1.13) for $t < 0$, the position of the state at time $t = 0$ (by assuming the zero initial condition $x(-\infty) = 0$) is equal to

$$x(0) = \int_{-\infty}^{0} e^{-At}Bu(t)dt := \mathcal{C}_o u(t).$$

By assuming that a zero input is applied to the system for $t > 0$, then for all $t \geq 0$, the output $y(.) \in \mathcal{L}_2[0, +\infty]$ of the system (1.13) is equal to

$$y(t) = Ce^{At}x(0) := \mathcal{O}_b x(0).$$

The so-called controllability operator $\mathcal{C}_o : \mathcal{L}_2[-\infty, 0] \mapsto \mathbb{R}^n$ (mapping past inputs $u(.)$ to the present state) and observability operator $\mathcal{O}_b : \mathbb{R}^n \mapsto \mathcal{L}_2[0, +\infty]$ (mapping the present state to future outputs $y(.)$) also have dual operators, respectively $\mathcal{C}_o^*$ and $\mathcal{O}_b^*$. It can be shown that the controllability and observability gramians are related to those via the identities $P = \mathcal{C}_o^*\mathcal{C}_o$ and $Q = \mathcal{O}_b\mathcal{O}_b^*$ ([94]).

If the state space coordinates of the system are changed to $z := Tx$ for some nonsingular T then $\dot{z} = TAT^{-1}z + TBu$, $y = CT^{-1}z$. Furthermore, the controllability and observability gramians become TPT^* and $T^{-*}QT^{-1}$, respectively. This leads us to the following concept.

Definition 1.32. *Consider two matrices $X, Y \in \mathbb{C}^{n\times n}$ related to a particular state $x \in \mathbb{C}$. The pair (X, Y) is said to perform a contragradient transformation under the state space change of coordinate $\bar{x} := Tx$ if*

$$\bar{X} = TXT^* \qquad \textit{and} \qquad \bar{Y} = T^{-*}YT^{-1}. \tag{1.19}$$

The gramians P and Q depend strongly on the state-space coordinates but their product performs the contragradient transformation (1.19), implying that

$$PQ \to TPQT^{-1} \tag{1.20}$$

in the transformed coordinates. Therefore the eigenvalues of PQ are invariant under state-space transformation, and are hence input-output invariant.

Definition 1.33. *Let A be Hurwitz. The Hankel singular values of $T(s) := C(sI - A)^{-1}B$ are defined as [45]*

$$\sigma_i\left(T(s)\right) := \{\lambda_i(PQ)\}^{1/2}$$

where by convention $\sigma_i\left(T(s)\right) \geq \sigma_{i+1}\left(T(s)\right)$. P and Q are defined in (1.18). *The Hankel norm of the stable strictly proper transfer function $T(s)$ is defined as the largest Hankel singular value of $T(s)$, i.e. $\sigma_1\left(T(s)\right)$.*

It should be pointed out that it is always possible to find a minimal state space realization of $T(s)$ such that the controllability and observability gramians are equal and diagonal. This is the so-called *balanced realization*. The concepts of gramians and Hankel map play a central role in the model reduction framework, as we will see in the following chapters.

Notes and References

A classic reference to Matrix Theory is [56]. For a deeper study of the Kronecker Form, we refer to [42] where a proof of Theorem 1.15 is available. An excellent reference providing many numerical algorithms to solve linear algebra problems is [46].

This chapter is certainly not sufficient for the reader that is not familiar with the theory of linear systems. We refer to [94] for an excellent introduction to the general framework of System Theory. Most of the results presented here are given in more details in Chapters 2 and 3 of that book. An earlier reference to linear systems is [60] where simple derivations of the Smith and Smith-Mc Millan form are available.

2

Model Reduction of Linear Systems

When the system order n of is too large for solving various control problems within a reasonable computing time, it is natural to consider approximating it by a reduced order system

$$\hat{S} \ : \ \begin{cases} \dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{B}u(t), \\ \hat{y}(t) = \hat{C}\hat{x}(t), \end{cases} \tag{2.1}$$

driven with the same input $u(t) \in \mathbb{C}^m$, but having a different output $\hat{y}(t) \in \mathbb{C}^p$ and state $\hat{x}(t) \in \mathbb{C}^k$. The matrix $\hat{A}$ belongs to $\mathbb{C}^{k\times k}$. The order k of the reduced order system is also assumed to be much smaller than the order n of the original system: $k \ll n$. The objective of the reduced order model is to reduce the dimension of the state-space (of dimension n) of the system to a lower dimension k in such a way that the "behavior" of the reduced order model is sufficiently close to that of the full order system. For a same input $u(t)$, we thus want $\hat{y}(t)$ to be close to $y(t)$. One shows that in the frequency domain, this is equivalent to imposing conditions on the frequency responses of both systems [94]: one wants to find a reduced order model such that the transfer functions of both models, i.e.

$$T(s) = C(sI_n - A)^{-1}B \quad , \quad \hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}, \tag{2.2}$$

are such that the error

$$E(s) := T(s) - \hat{T}(s)$$

is minimal for a given criterion.

Most popular model reduction techniques for linear systems fit into the framework of either SVD-based or Krylov subspace-based techniques [4]. Perhaps the most popular model reduction technique for linear systems is the Balanced Truncation method. This SVD-based technique has many advantages: the stability of the original system is preserved and there exists an a priori

global bound on the error between the original and the reduced system. The main drawback is that the technique cannot be applied to large-scale systems of order n, i.e., those systems where $O(n^3)$ computations is an unacceptably large cost. On the other hand, Krylov subspace-based techniques that are based on imposing moment matching conditions between the original and the reduced transfer function, such as rational/tangential interpolation methods, can be applied to large-scale systems but do not provide global error bounds and depend significantly on the choice of certain parameters. Even worse, the stability of the reduced order system is not guaranteed.

The outline of this chapter is as follows. In Section 2.1, the projection framework for model reduction of linear systems is presented. In Section 2.2, the Modal Approximation technique is introduced. Krylov and SVD techniques are presented in Section 2.3 and 2.4 respectively. Other model reduction techniques are briefly sketched in Section 2.5.

2.1 Model Reduction by State Space Projection

There exist many ways of constructing a reduced order system, but most of them construct a projection of the state space into a subspace that contains the essential dynamics of the system. When a transfer function is constructed by projecting a state space realization of another transfer function, one says that the former is *embedded* into the latter.

Definition 2.1. *The state space realization $(\hat{C}, \hat{A}, \hat{B})$ of order k, with $\hat{C} \in \mathbb{C}^{p\times k}$ and $\hat{B} \in \mathbb{C}^{k\times m}$, is said to be embedded into the state space realization (C, A, B) (with $C \in \mathbb{C}^{p\times n}$ and $B \in \mathbb{C}^{n\times m}$) when there exist projecting matrices $Z, V \in \mathbb{C}^{n\times k}$ such that $Z^T V = I_k$ and*

$$\{\hat{A}, \hat{B}, \hat{C}\} = \{Z^T AV, Z^T B, CV\}. \tag{2.3}$$

A transfer function $\hat{T}(s)$ is said to be embedded into another transfer function if any minimal state space realization of $\hat{T}(s)$ is embedded into any minimal state space realization of $T(s)$.

The notion of embedding does not depend on the choice of the minimal state space realizations. More precisely,

Lemma 2.2. *Let $T(s)$ and $\hat{T}(s)$ be two $p \times m$ strictly proper transfer function of respective Mc Millan degree n and k. If there exists one minimal state space realization $(\hat{C}, \hat{A}, \hat{B})$ of $\hat{T}(s)$ that is embedded in one minimal state space realization (C, A, B) of $T(s)$, then any minimal state space realization of $\hat{T}(s)$ is embedded in any minimal state space realization of $T(s)$.*

Proof. Let us assume that $(\hat{C}, \hat{A}, \hat{B})$ is embedded into (C, A, B). From Definition 2.1, there exists $Z, V \in \mathbb{C}^{n\times k}$ such that $Z^T V = I_k$ and (2.3) is satsified. Let (C_2, A_2, B_2) and $(\hat{C}_2, \hat{A}_2, \hat{B}_2)$ be two minimal state space realizations of $T(s)$ and $\hat{T}(s)$ respectively. From Lemma 1.19, there exists $S \in \mathbb{C}^{n\times n}$ and $\hat{S} \in \mathbb{C}^{k\times k}$ such that

$$(\hat{C}_2\hat{S}, \hat{S}^{-1}\hat{A}_2\hat{S}, \hat{S}^{-1}\hat{B}_2) = (\hat{C}, \hat{A}, \hat{B}), \quad (C_2, A_2, B_2) = (CS, S^{-1}AS, S^{-1}B). \tag{2.4}$$

Using (2.3) and (2.4), one finds that

$$\begin{aligned}(\hat{C}_2, \hat{A}_2, \hat{B}_2) &= (\hat{C}\hat{S}^{-1}, \hat{S}\hat{A}\hat{S}^{-1}, \hat{S}\hat{B}) \\ &= (CV\hat{S}^{-1}, \hat{S}Z^T AV\hat{S}^{-1}, \hat{S}Z^T B) \\ &= (C_2S^{-1}V\hat{S}^{-1}, \hat{S}Z^T SA_2S^{-1}V\hat{S}^{-1}, \hat{S}Z^T SB_2).\end{aligned}$$

By defining

$$Z_2 = S^T Z\hat{S}^T, \quad V_2 = S^{-1}V\hat{S}^{-1},$$

clearly $Z_2^T V_2 = I_k$ and

$$\{\hat{A}_2, \hat{B}_2, \hat{C}_2\} = \{Z_2^T A_2 V_2, Z_2^T B_2, C_2 V_2\}.$$

□

If we change the state-space coordinate basis of the system (1.13) by choosing

$$\tilde{x} = Sx,$$

with the matrix $S \in \mathbb{C}^{n\times n}$ invertible, the system (1.13) is equivalent to the system

$$\begin{cases} \dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u \\ y = \tilde{C}\tilde{x}, \end{cases} \tag{2.5}$$

where

$$\{\tilde{A}, \tilde{B}, \tilde{C}\} = \{SAS^{-1}, SB, CS^{-1}\}. \tag{2.6}$$

It can be shown that the reduced-order system $\hat{T}(s)$ is embedded into $T(s)$ if and only if there exists a state-space coordinate basis in which the matrices of the original system $T(s)$ are (see for instance [68] for a proof)

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \tag{2.7}$$

and the matrices of the reduced-order model are taken to be

$$\hat{A} = A_{11}, \quad \hat{B} = B_1, \quad \hat{C} = C_1.$$

For instance, from equations (2.3) and (2.6), we can choose the projecting matrices $Z^T \in \mathbb{C}^{k\times n}$ and $V \in \mathbb{C}^{n\times k}$ to be respectively the first k rows of S and

the first k columns of S^{-1}, where S is the coordinate basis change that put the original system in the form (2.7).

All the model reduction techniques that are introduced in this chapter use (at least implicitly) this *projection* of state space in order to construct the reduced order transfer function. The generality of the projection technique will be discussed in more details later.

2.2 Modal Approximation

This model reduction technique consists of taking the part of the transfer function with the poles that are the closest to the imaginary axis and throwing away the others. This is probably the oldest model reduction technique developed for linear systems. Due to its simplicity, the Modal Approximation method is still used for some applications.

For simplicity, let us consider a stable SISO strictly proper transfer function of Mc Millan degree n,

$$T(s) := \frac{n(s)}{d(s)},$$

where the degrees of the polynomials $d(s)$ and $n(s)$ are respectively n and $n-1$, devoid of common zeros. Let us assume that all the roots of $d(s)$ are distinct: $d(s) := \prod_{i=1}^{n}(s-\lambda_i)$ and that the roots are ordered such that $0 > Re(\lambda_1) \geq Re(\lambda_2) \geq \ldots \geq Re(\lambda_n)$. Using the well known *partial fraction expansion* formula, one can rewrite $T(s)$ as a sum of n transfer functions of Mc Millan degree one :

$$T(s) = \sum_{i=1}^{n} T_i(s) = \sum_{i=1}^{n} \frac{\gamma_i}{s-\lambda_i}, \tag{2.8}$$

where $T_i(s) := \frac{\gamma_i}{s-\lambda_i}$. Intuitively, from the expression (1.17) of the H_2 and H_∞ norms, closer is λ_i to the imaginary axis, higher is its contribution to the overall energy of $T(s)$. The idea of modal approximation consists in taking $\sum_{i=1}^{k} T_i(s)$ as an approximation of Mc Millan degree k of $T(s)$.

In a state space framework, given a state space realization (C, A, B) of a transfer function, it is always possible to perform a state space transformation $(\bar{C}, \bar{A}, \bar{B}) = (CS, S^{-1}AS, S^{-1}B)$ such that $\bar{A}$ is block diagonal :

$$\bar{A} = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_l \end{bmatrix},$$

where the matrices A_i, $1 \leq i \leq l$ have distinct eigenvalues. Rewriting with appropriate dimensions

$$\bar{C} = \left[\, C_1 \ldots C_l \,\right] \quad , \quad \bar{B} = \left[\, B_1^T \ldots B_l^T \,\right]^T ,$$

one obtains the following *generalized* partial fraction expansion [59]

$$T(s) = \sum_{i=1}^{l} C_i (sI - A_i) B_i .$$

Techniques in order to perform such a block diagonalization of the state matrix A have been developed in the literature. The objective is to divide the set of eigenvalues of A, i.e. the poles of $T(s)$ into two sets: one to be discarded and the other to be kept in the reduced order system.

The major drawback of Modal Approximation is the lack of guaranteed bounds for the approximation error and the poor flexibility of this technique. For these reasons amongst others, Modal Approximation has been replaced by other techniques that are presented below. Nevertheless, this procedure can be used in a state space framework as an intermediate step in order to throw away the unstable part of a reduced order system obtained by another technique (see for instance [58] and [49]). This method is also useful when the state matrix A is of very large dimension but sparse, permitting the application of cheap Modal Approximation techniques in order to obtain an intermediate transfer function of acceptable degree, which can be reduced once again using other techniques (see Section 6.5).

When the original transfer function $T(s)$ is not proper, one often copies to polynomial part of $T(s)$ into the reduced order system. This can also be viewed as a *modal approximation* technique, with a pole at infinity.

The modal approximation technique introduced here is by no way the only existing scheme. For instance, one does not take into account the value of the numerator γ_i in (2.8) in the selection process. Enhanced techniques have been proposed in the literature. We refer to [86] and references therein for a deeper study of Modal Approximation.

2.3 Krylov Approximation Techniques

A popular class of model reduction techniques is based on constructing a reduced order transfer function that satisfies a set of interpolation conditions with respect to the transfer function one wants to approximate. These are the Krylov model reduction techniques.

First, some words about the notation. We say that a rational matrix function $R(s)$ is $O(\lambda - s)^k$ in s with $k \in \mathbb{Z}$ if its Taylor expansion about the point λ can be written as follows :

$$R(s) = O(\lambda - s)^k \Longleftrightarrow R(s) = \sum_{i=k}^{+\infty} R_i (\lambda - s)^i , \tag{2.9}$$

where the coefficients R_i are constant matrices. If $R_k \neq 0$, then we say that $R(s) = \Theta(\lambda - s)^k$. As a consequence, if $R(s) = \Theta(\lambda - s)^k$ and k is strictly negative, then λ is a pole of $R(s)$ and if k is strictly positive, then λ is a zero of $R(s)$. Analogously, we say that $R(s)$ is $O(s^{-1})^k$ if the following condition is satisfied :

$$R(s) = O(s^{-1})^k \Longleftrightarrow R(s) = \sum_{i=k}^{+\infty} R_i s^{-i}, \tag{2.10}$$

where the coefficients R_i are constant matrices. It should be stressed that, in general, $R(s)$ being $O(s)^{-k}$ is not equivalent to $R(s)$ being $O(s^{-1})^k$. To convince yourself, replace λ by zero and k by -1 in (2.9) and compare the expression with (2.10) and $k = 1$.

Definition 2.3. *Given a pair (A, B), the Krylov matrix of order $k \in \mathbb{N}_0$, written $K_k(A, B)$, is defined as*

$$K_k(A, B) := \left[B \ AB \ A^2B \ \dots \ A^{k-1}B \right].$$

The corresponding Krylov subspace of order $k \in \mathbb{N}_0$, written $\mathcal{K}_k(A, B)$, is defined as as

$$\mathcal{K}_k(A, B) := Im\left(K_k(A, B)\right).$$

If $k \leq 0$, then we define

$$\mathcal{K}_k(A, B) = \{0\}.$$

One can expand the transfer function $T(s) = C(sI - A)^{-1}B$ at infinity as follows :

$$T(s) = \sum_{k=0}^{+\infty} CA^kBs^{-k-1} := \sum_{k=0}^{\infty} M_k^{(\infty)} s^{-k-1}, \tag{2.11}$$

where the coefficients $M_k^{(\infty)} := CA^kB$ are called the *Laurent* coefficients (also called *Markov* parameters) of $T(s)$. Note that M_0^{∞} corresponds to the coefficient of s^{-1} and not s^0 in (2.11) because $T(s)$ is strictly proper. If one wants to find a reduced transfer function, $\hat{T}(s) := \hat{C}(sI - \hat{A})^{-1}\hat{B} := \sum_{k=0}^{\infty} \hat{M}_k^{(\infty)} s^{-k-1}$, that approximates as well as possible the original transfer function for large frequency $s \to \infty$, it makes sense to choose $\hat{T}(s)$ such that for $0 \leq k \leq K - 1$,

$$\hat{M}_k^{(\infty)} = M_k^{(\infty)}.$$

It is equivalent to say that

$$T(s) - \hat{T}(s) = O(s^{-1})^{K+1}.$$

This problem is known as the *partial realization problem.*

In order to obtain a good approximation in the low frequency domain, one might prefer to construct a transfer function

$$\hat{T}(s) = \hat{C}(sI - \hat{A})^{-1}\hat{B} = \sum_{k=0}^{\infty} \hat{M}_k^{(\lambda)}(\lambda - s)^k,$$

such that

$$\hat{M}_k^{(\lambda)} = M_k^{(\lambda)} \qquad for\ 0 \leq k \leq K, \tag{2.12}$$

with

$$M_k^{(\lambda)} := C(\lambda I - A)^{-k-1}B, \quad \hat{M}_k^{(\lambda)} := \hat{C}(\lambda I - \hat{A})^{-k-1}\hat{B}.$$

Equation (2.12) can be rewritten more compactly as follows

$$T(s) - \hat{T}(s) = O(\lambda - s)^{K+1}.$$

The transfer function $\hat{T}(s)$ is called a *Padé* approximation if it interpolates $T(s)$ at $\lambda = 0$ and a *shifted Padé* approximation if $\lambda \neq 0$.

A natural generalization of the preceding problems is the following : how to construct a transfer function $\hat{T}(s)$ that interpolates $T(s)$ at several points in the complex plane, up to given successive derivatives? This is the *rational interpolation* problem.

The main results concerning this problem are summarized in the following Theorem (given here for standard state space realizations) :

Theorem 2.4. *Let the original system be*

$$T(s) := C(sI - A)^{-1}B, \tag{2.13}$$

and the reduced system be

$$\hat{T}(s) := \hat{C}(sI - \hat{A})^{-1}\hat{B} = CV\left(Z^T(sI - A)V\right)^{-1} Z^T B, \tag{2.14}$$

with the bi-orthogonality constraint

$$Z^T V = I.$$

If

$$\bigcup_{k=1}^{K} \mathcal{K}_{b_k}((\sigma_k I - A)^{-1}, (\sigma_k I - A)^{-1}B) \subseteq Im(V) \tag{2.15}$$

and

$$\bigcup_{k=1}^{K} \mathcal{K}_{c_k}((\sigma_k I - A)^{-T}, (\sigma_k I - A)^{-T}C^T) \subseteq Im(Z) \tag{2.16}$$

where the interpolation points σ_k are chosen such that the matrices $\sigma_k I - A$ are invertible $\forall k \in \{1, \ldots, K\}$ then the moments of the systems (2.13) *and* (2.14) *at the points σ_k satisfy*

$$T(s) - \hat{T}(s) = O(s - \sigma_k)^{b_k + c_k}, \tag{2.17}$$

for $k = 1, 2, \ldots, K$, provided these moments exist, i.e. provided the matrices $\sigma_k \hat{I} - \hat{A}$ are invertible.

For a proof, see for instance [26]. □

Remark 2.5. The preceding Theorem remains valid in the case of generalized state space realizations where the bi-orthogonality constraint $Z^T V = I$ is replaced by $Z^T E V = \hat{E}$, where $(C, E, A, B) \sim T(s)$ and $(\hat{C}, \hat{E}, \hat{A}, \hat{B}) \sim \hat{T}(s)$ [48].

The Krylov subspace-based methods that produce reduced order models based on rational interpolation can be applied to MIMO systems efficiently as long as the number of inputs and outputs, m and p, stay suitably moderate in size. For MIMO systems where m and p are too large, a more general *tangential interpolation* problem has recently been considered (see [39]). Given a $p \times m$ strictly proper transfer function $T(s)$, one shows how to construct a $p \times m$ reduced order transfer function $\hat{T}(s)$ of Mc Millan degree k that satisfies a set of *tangential interpolation* conditions of the form

$$x_i(\lambda_i)T(\lambda_i) = x_i(\lambda_i)\hat{T}(\lambda_i) \quad , \quad T(\lambda_j)y_j(\lambda_j) = T(\lambda_j)y_j(\lambda_j),$$

with chosen left and right polynomial directions vectors $x_i(s) \in \mathbb{C}^{1\times p}[s], y_j \in \mathbb{C}^{m\times 1}[s]$ and interpolation points $\lambda_i, \lambda_j \in \mathbb{C} \cup \{\infty\}$, $1 \leq i \leq k$, $k+1, \leq j \leq 2k$. Such interpolation conditions will correspond to *generalized Krylov* subspaces. As with rational interpolation, higher order tangential interpolation conditions can be imposed at each point to improve the approximation. This topic is developed in Chapter 4.

Krylov techniques have one major advantage: they can be implemented by using efficient numerical techniques (related to the Arnoldi [48] and Lanczos [33] algorithms). All that need to be computed are the Krylov subspaces that must be contained in the image of the projection matrices Z and V. These roughly require only sparse matrix-vector products. As a consequence, Krylov model reduction techniques can be applied to reduce systems of very large order. That is the reason why Krylov techniques are very popular in circuit simulation [30]. The major drawback concerning interpolation techniques is that they are local by nature. For instance, there exists no global error bound between the original and the reduced order system obtained by a Krylov technique. Even worse, stability of the original system can be lost in the reduced order system. Concerning stability, one remedy consists of using a modal approximation technique in order to delete the unstable modes of the reduced order system, as it is done implicitly in [49]. It is also possible to guarantee stability of the reduced order system by choosing appropriate interpolation points [6], but this can affect local approximation errors.

The local approximation property means that good approximations can be achieved in specific regions over a wide dynamic range typically at the cost of larger global error. This requires however, that the interpolation points and their corresponding order of matching must be specified. For some applications,

the user may have such information but for blackbox library software a heuristic automatic selection strategy is needed (see [48]) and the design of such a strategy is still an open question. The other main drawback is the lack of an error bound on the global quality of the approximation, e.g., the H_∞-norm of the difference between original and reduced transfer functions. Recent research has begun to address the evaluation of the H_∞-norm given a reduced order model that may help in selecting points [19]. If the lack of global error bounds is an important problem for theoretical purposes, this may not be the main concern for practical purposes, where only selected parts of the frequency domain are important, advocating for model reduction techniques with local approximation properties at selected points.

2.4 Gramian-based Techniques

We consider here linear time-invariant systems modelled by the system of equations (1.13) which therefore have a transfer function $T(s) = C(sI_n - A)^{-1}B$ that links the inputs to the outputs in the Laplace domain.

The Balanced Truncation method is based on the following physical interpretation of the gramians defined in (1.18). The controllability matrix arises from the following optimization problem. Let

$$J(v(t), a, b) := \int_a^b v(t)^T v(t) dt$$

be the *energy* of the vector function $v(t)$ in the interval $[a, b]$. Then (see [45])

$$\min_{\mathcal{C}_0 u(t)=x_0} J(u(t), -\infty, 0) = x_0^T P^{-1} x_0, \tag{2.18}$$

and, symmetrically, we have the dual property

$$\min_{\mathcal{O}_b^* y(t)=x_0} J(y(t), -\infty, 0) = x_0^T Q^{-1} x_0, \tag{2.19}$$

where $\mathcal{O}_b^*$ is the dual of the operator O_b defined in page 25. This dual operator corresponds to the controllability operator of the dual of the transfer function $T(s)$ (seen as an operator mapping the input space to the output space).

Remark 2.6. Another way to define the observability gramian from the observability operator $\mathcal{O}_b$ has already been described in page 25. We have chosen to give here the *dual* version because it can easily be generailzed for second order gramians and interconnected systems in Chapters 6 and 7.

Two essential algebraic properties of gramians P and Q are as follows. First, it has already been pointed out that the gramians enjoy a contragradient property

when a coordinate transformation $x(t) = S\bar{x}(t)$ is applied to the system. This implies that the eigenvalues of the product $\bar{\mathcal{P}}\bar{\mathcal{Q}} = S^{-1}\mathcal{P}\mathcal{Q}S$ depends only on the transfer function $T(s)$ and not on a particular choice of state-space realization. This implies also that there exists a state-space realization $(C_{bal}, A_{bal}, B_{bal})$ of $T(s)$ such that the corresponding gramians are equal and diagonal $\bar{P} = \bar{Q} = \Sigma$ [94]. Secondly, because these gramians appear in the solutions of the optimization problems (2.18) and (2.19), they say something about the energy that goes through the system, and more specifically, about the distribution of this energy among the state variables.

Smaller is $x_0^T P^{-1} x_0$ is, more "controllable" is the state x_0, because it can be reached with a input of small energy. By duality, smaller is $x_0^T Q^{-1} x_0$, more "observable" is the state x_0. Thus when both gramians are equal and diagonal, the order of magnitude of a diagonal value of the product PQ is a good measure for the influence of the corresponding state variable in the mapping $y(.) = \mathcal{O}_b\mathcal{C}_o u(.)$ which maps past inputs $u(t) \in \mathcal{L}_2[-\infty, 0]$ to future outputs $y(t) \in \mathcal{L}_2[0, +\infty]$ passing via that particular state at time $t = 0$.

Given a transfer function $T(s)$, the popular balanced truncation model reduction method consists of finding a state-space realization $(C_{bal}, A_{bal}, B_{bal})$ of $T(s)$ such that the gramians are equal and diagonal (this is the so-called *balanced realization*) and then the reduced model is constructed by keeping the states corresponding to the largest eigenvalues of the product PQ in it. In other words, the widely used balanced truncation technique chooses Z and V such that $Z^T V = I$, and

$$\begin{cases} PQV = V\Lambda_+, \\ QPZ = Z\Lambda_+, \end{cases} \tag{2.20}$$

where Λ_+ is a square diagonal matrix containing the largest eigenvalues of PQ. Then a state-space realization of the reduced transfer function is given by $(CV, Z^T AV, Z^T B)$. The idea of the balanced truncation technique thus consists of keeping those states that are most controllable and observable according to the gramians satisfying the optimization problems (2.18) and (2.19).

Remark 2.7. The Hankel operator that maps the past input to the future output is defined as follows : $\mathcal{H} := \mathcal{O}_b\mathcal{C}_o$. Since $PQ = \mathcal{C}_o\mathcal{C}_o^*\mathcal{O}_b^*\mathcal{O}_b$ and $QP = \mathcal{O}_b^*\mathcal{O}_b\mathcal{C}_o\mathcal{C}_o^*$, the dominant eigenspaces $\mathcal{V}$ of PQ and $\mathcal{Z}$ of QP are linked with the dominant eigenspaces $\mathcal{X}$ of $\mathcal{H}\mathcal{H}^*$ and $\mathcal{Y}$ of $\mathcal{H}^*\mathcal{H}$: indeed it holds that $\mathcal{X} = \mathcal{O}_b\mathcal{V}$ and $\mathcal{Y} = \mathcal{C}_o^*\mathcal{Z}$. Therefore projecting on the spaces $\mathcal{V}$ and $\mathcal{Z}$ also approximates the Hankel map $\mathcal{H}$ well.

An important advantage of the Balanced Truncation technique is the following error bound.

Theorem 2.8. *Let $T(s)$ be a stable strictly proper transfer function of Mc Millan degree n. Let $\hat{T}(s)$ be the strictly proper reduced order transfer function of Mc Millan degree k of $T(s)$ obtained by balanced truncation. If the $k-th$ Hankel*

singular value $\sigma_k(T(s))$ *is strictly larger than* $\sigma_{k+1}(T(s))$*, then* $\hat{T}(s)$ *is stable and*

$$\|T(s) - \hat{T}(s)\|_{H_\infty} \leq 2 \sum_{i=k+1}^{n}{}' \sigma_i \left(T(s)\right),$$

where $\sum'$ *indicates that Hankel singular value of multiplicity larger than one are only counted once.*

The major drawback is that the computational complexity for computing a reduced order transfer function is $O(n^3)$, where n is the Mc Millan degree of the transfer function to be approximated. Recent work [79] considers this problem and describes an Approximate Balanced Truncation approach for large-scale linear systems.

Many model reduction techniques are related to Balanced Truncation. For instance, the Optimal Hankel norm approximation method provides as a reduced order transfer function of order k the generically unique transfer function $\hat{T}_H(s)$ that minimizes the Hankel norm of the error :

$$\hat{T}_H(s) = \arg\min_{\hat{T}(s)} \|T(s) - \hat{T}(s)\|_H,$$

over all stable strictly proper transfer $\hat{T}(s)$ of Mc Millan degree k. We refer to [94], [45] for a deeper study of the balanced truncation and Optimal Hankel norm approximation techniques. The complexity is also $O(n^3)$.

Other SVD-based model reduction techniques are further discussed in [50].

2.5 Other Sparse Model Reduction techniques

At the present time, active research is pursuing model reduction techniques that benefit from the advantage of Krylov techniques, i.e. low computational complexity, and SVD-techniques, i.e. existence of a global error bound between the original and the reduced order system.

Instead of exactly computing the gramians (1.18) that require $O(n^3)$ floating point operations for a system of order n, why not computing *approximate* solutions by using iterative solvers? The Smith-ADI iteration tries to perform a power iteration in order to approximate the gramians as follows.

First, for every scalar $\mu < 0$, the solution P of the Lyapunov equation $AP + PA^* + BB^* = 0$ is also the solution of the *Stein* equation

$$P - A_\mu P A_\mu^* = B_\mu B_\mu^*,$$

where

$$A_\mu := (A - \mu I)(A + \mu I)^{-1} \quad , \quad B_\mu = \sqrt{-2\mu}(A + \mu I)^{-1} B.$$

The gramian can then be rewritten as

$$P = \sum_{i=0}^{\infty} A_\mu^i B_\mu B_\mu^* (A_\mu^i)^*.$$

If the sum converges quickly,

$$P_k := \sum_{i=0}^{k} A_\mu^i B_\mu B_\mu^* (A_\mu^i)^* \tag{2.21}$$

is a good approximation of P, if k is large enough. Of course, the choice of the shift μ influences the convergence of the iteration. Once the approximate gramians have been computed, one proceeds as for Balanced Truncation by first computing the left and right dominant the eigenspaces of the product $P_k Q_k$ and then projecting.

In order to enhance the convergence of the algorithm, multiple shifts strategies have been developed in [69]. Low-Rank and low-rank square root versions have been further studied in [61] and [52].

If the truncation parameter k in (2.21) is large enough, $P_k \approx P$ and Smith-ADI model reduction techniques become equivalent to the classic Balanced Truncation Technique. For a small value of k, Smith-ADI techniques are similar to Krylov techniques (see for instance [61], Section 10.1, where a connection between moment matching methods and an ADI technique is given for symmetric systems).

Another SVD-technique that approximates Balanced Truncation is based on iteratively computing the cross gramian. The *cross gramian* X for square systems $m = p$, is defined as the solution of the Sylvester equation

$$AX + XA + BC = 0.$$

A restarting algorithm based on solving projected Sylvester equations can be found in [79].

Another set of model reduction techniques for linear systems has not been introduced yet. These are the *structure preserving* model reduction techniques. Large scale systems are often composed of subsystems that interconnect to each other and it can be important to preserve the structure of the interconnections in the reduced order systems. Examples of interconnected systems are second order systems, plants with controller (controller reduction), weighted systems, etc. This topic will be studied in Chapters 6 and 7.

All the model reduction methods introduced so far are explicit *projection* techniques, in the sense of Section 3.3 (except Optimal Hankel Norm Approximation). Another method that does not perform a classic projection is the *singular perturbation* technique (see [68] and references therein). This last model reduction method is not considered here.

Notes and References

Many papers have been written in the area of model reduction of linear systems. A recent and rather complete study of this topic with an emphasis on Krylov and SVD-Krylov techniques can be found in [3, 8]. The excellent work [94] gives a more detailed study of the Balanced Truncation and Optimal Hankel Norm approximation methods (see also [45]). It should be pointed out that the list of model reduction techniques presented here is by no way exhaustive. Rather than giving a detailed study of a limited number of model reduction technique, our choice was to present a broader class of methods and to refer to the literature for a deeper study.

3

Model Reduction by Moment Matching, the SISO case

In this chapter, we focus our attention on scalar transfer¡ functions, i.e. with one input and one output (SISO LTI systems). General results concerning Krylov subspaces are developed in Section 3.1. Using these results, a new proof of Theorem 2.4 is constructed in Section 3.2. Earlier proofs of this theorem have appeared in the literature [34, 26], but our derivation of the theorem yields new insights into the problem. As a consequence, it is shown in Section 3.3 that for any SISO strictly proper transfer function $T(s)$ of Mc Millan degree n and any strictly proper SISO transfer function $\hat{T}(s)$ of Mc Millan degree $k < n$, $\hat{T}(s)$ can always be constructed via truncation of the system $T(s)$. Finally, a link between Sylvester equations, Krylov subspaces and projection matrices is drawn in Section 3.4. This chapter is based mainly on [38] and [40]. Parts of the theory developed in this chapter are more deeply studied in the general MIMO case in Chapter 4. For this reason, some details are omitted and some proofs are shortened.

3.1 Some facts about Krylov subspaces

To begin, let us develop some preliminary results concerning Krylov subspaces. The main objective of this section is to prove that the image of the controllability and observability spaces can be reconstructed using Krylov subspaces and eigenspaces. This is obtained in Lemma 3.11.

It should be pointed out that results similar to Lemma 3.9 also appear in [62], Chapter 6 and [1].

Lemma 3.1. *Consider an arbitrary pair of matrices (A, B) with $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times 1}$. Consider n polynomials of degree at most $n-1$,*

$$\phi_j(x) = \sum_{i=0}^{n-1} \alpha_{i,j} x^i, \qquad 1 \le j \le n,$$

an define the matrix $M \in \mathbb{R}^{n\times n}$ such that

$$M(i,j) = \alpha_{i-1,j}, \qquad 1 \le i,j \le n.$$

If M is invertible (i.e. if the polynomials are independent), then

$$Im\left(\left[\,\phi_1(A)B \ldots \phi_n(A)B\,\right]\right) = \mathcal{K}_n(A,B).$$

Proof. Because the functions $\phi_j(x)$ are polynomial and of finite degree, they are analytic in a neighborhood of the spectrum of A and the functions $\phi_j(A)$ are well defined. The proof of the lemma now follows from the following equation

$$\left[\,\phi_1(A)B \ldots \phi_n(A)B\,\right] = \left[\,B \ldots A^{n-1}B\,\right] M$$

and the fact that M is invertible.

□

Remark 3.2. By the Cayley-Hamilton Theorem and by considering the Jordan canonical form of the matrix $A \in \mathbb{C}^{n\times n}$, it is well known that any function $\phi(\cdot)$ analytic in a neighborhood of the spectrum of A, denoted by $\Lambda(A)$, can be written as a polynomial function of A of degree $n-1$. Hence, the functions ϕ_i can be more general than polynomials.

This leads to the following definition.

Definition 3.3. *Let A be a square matrix of dimension n, let $\phi(\cdot)$ be a function analytic in a neighborhood of the spectrum of A, the polynomial function of minimal degree, $r(\cdot)$ (obtained via Cayley-Hamilton), such that the matrices $r(A)$ and $\phi(A)$ are equal, is called the interpolating polynomial of $\phi(\cdot)$ with respect to the matrix $A \in \mathbb{C}^{n\times n}$.*

Lemma 3.4. *Consider an arbitrary pair of matrices (A,B) with $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times 1}$. Let $\phi(\cdot)$ be any function such that the matrix $\phi(A) \in \mathbb{C}^{n\times n}$ is invertible. Then*

$$\phi(A)\mathcal{K}_n(A,B) = \mathcal{K}_n(A,B).$$

Proof. By Cayley-Hamilton,

$$\phi(A)\mathcal{K}_n(A,B) = r(A)\mathcal{K}_n(A,B) \subseteq \mathcal{K}_n(A,B),$$

where $r(A)$ is the interpolating polynomial of $\phi(A)$. By invertibility of $\phi(A)$,

$$dim(\phi(A)\mathcal{K}_n(A,B)) = dim(\mathcal{K}_n(A,B)).$$

Equality of the 2 subspaces follows.

□

Definition 3.5. *An* ***interpolation set*** *I*

$$I = \{(s_1, m_1), \ldots, (s_r, m_r)\},$$

is defined as a set of pairs (s_i, m_i) where the points $s_i \in \mathbb{C} \cup \infty$ are distinct and the indices $m_i \in \mathbb{N}_0$. The size of the interpolation set I, denoted by $size(I)$ is defined as

$$size(I) := \sum_{i=1}^{r} m_i.$$

*An interpolation set I is called a $T(s)$-**admissible interpolation set** when no interpolation point s_i is a pole of $T(s)$. A **minimal** $T(s)$-**admissible interpolation set** is a $T(s)$-admissible interpolation set of size n, where n is the Mc Millan degree of $T(s)$.*

Definition 3.6. *A couple of $T(s)$-admissible interpolation sets (I_1, I_2), denoted by*

$$I_1 = \{(z_1, \mu_1), \ldots, (z_{r_1}, \mu_{r_1})\} \quad , \quad I_2 = \{(w_1, \nu_1), \ldots, (w_{r_2}, \nu_{r_2})\},$$

*is called a **separation** of I if the set of points of I is the union of those of I_1 and I_2 and if their corresponding indices add up. By that, we mean that for each couple $(s_k, m_k) \in I$ belonging to I_1 and I_2 we have*

$$z_i = w_j = s_k \Rightarrow \mu_i + \nu_j = m_k,$$

and for each couple $(s_k, m_k) \in I$ belonging to only one set I_1 or I_2, we have (e.g. for I_1)

$$z_i = s_k \Rightarrow \mu_i = m_k.$$

As a consequence, we have

$$size(I_1) + size(I_2) = size(I).$$

*A separation (I_1, I_2) is called **symmetric** when $size(I_1) = size(I_2)$.*

The quantities occurring in $Contr(A, B)$ and $Obs(C, A)$

$$\Gamma_{A,B}(\infty, k) := A^{k-1}B \quad \Delta_{C,A}(\infty, k) := CA^{k-1} \tag{3.1}$$

can be seen as "moments" of $(sI - A)^{-1}B$ and $C(sI - A)^{-1}$ about infinity. Similarly, we define the moments about a finite expansion point $\lambda \in \mathbb{C}$

$$\Gamma_{A,B}(\lambda, k) := (\lambda I - A)^{-k}B, \quad \Delta_{C,A}(\lambda, k) := C(\lambda I - A)^{-k}. \tag{3.2}$$

Definition 3.7. *Let I be a $T(s)$-admissible interpolation set. For any state-space realization (C, A, B) of $T(s)$, we define the generalized controllability matrix $\mathcal{C}_{A,B}$ by the following equations*

$$\mathcal{C}_{A,B}(s_i, m_i) := \left[\, \Gamma_{A,B}(s_i, 1)\ \Gamma_{A,B}(s_i, 2)\ \dots\ \Gamma_{A,B}(s_i, m_i) \,\right], \tag{3.3}$$

$$\mathcal{C}_{A,B}(I) := \left[\, \mathcal{C}_{A,B}(s_1, m_1)\ \mathcal{C}_{A,B}(s_2, m_2)\ \dots\ \mathcal{C}_{A,B}(s_r, m_r) \,\right], \tag{3.4}$$

and generalized observability matrix to be

$$\mathcal{O}_{C,A}(s_i, m_i) := \left[\, \Delta_{C,A}(s_i, 1)^T\ \Delta_{C,A}(s_i, 2)^T\ \dots\ \Delta_{C,A}(s_i, m_i)^T \,\right]^T, \tag{3.5}$$

$$\mathcal{O}_{C,A}(I) := \left[\, \mathcal{O}_{C,A}(s_1, m_1)^T\ \mathcal{O}_{C,A}(s_2, m_2)^T\ \dots\ \mathcal{O}_{C,A}(s_r, m_r)^T \,\right]^T. \tag{3.6}$$

Let us introduce a final notation. Let (A, B) be a pair of matrices with $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times 1}$. If $s_i \neq \infty$ is not an eigenvalue of A, then define the matrix $A_i \in \mathbb{C}^{n\times n}$ by

$$A_i = (s_i I - A)^{-1}, \quad B_i = (s_i I - A)^{-1} B.$$

If $s_i = \infty$, then define

$$A_i = A, \quad B_i = B.$$

The following lemma is a straightforward consequence of the partial fraction expansion of a rational matrix. It will prove to be useful in the sequel.

Lemma 3.8. *Consider an arbitrary pair of matrices (A, B) with $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times 1}$. Let i and j be two non negative integers such that $i + j \geq 1$.*

1. *If $s_1 \neq \infty$, $s_2 \neq \infty$ and $s_1 \neq s_2$, then*

$$Im(A_1^i A_2^j B) \subseteq \mathcal{K}_i(A_1, B_1) + \mathcal{K}_j(A_2, B_2). \tag{3.7}$$

2. *If $s_1 \neq \infty$ and $s_2 = \infty$, then*

$$Im(A_1^i A_2^j B) \subseteq \mathcal{K}_i(A_1, B_1) + \mathcal{K}_{j-i+1}(A_2, B_2). \tag{3.8}$$

3. *a) If $s_1 = s_2 \neq \infty$ then,*

$$Im(A_1^i A_2^j B) \subseteq \mathcal{K}_{i+j}(A_1, B_1). \tag{3.9}$$

b) If $s_1 = s_2 = \infty$ then,

$$Im(A_1^i A_2^j B) \subseteq \mathcal{K}_{i+j+1}(A, B). \tag{3.10}$$

Proof. The third part of the Lemma is obvious. Let us prove the two first parts.

First, we suppose that $s_1 \neq s_2$ and that s_1 and s_2 are both different from ∞. We obtain by partial fraction expansion the identity

$$(s_1 I - A)^{-1}(s_2 I - A)^{-1} = (s_1 I - A)^{-1}\frac{1}{s_2 - s_1} + (s_2 I - A)^{-1}\frac{1}{s_1 - s_2}. \tag{3.11}$$

By recursively applying this equation, we find that

$$
\begin{aligned}
&(s_1 I - A)^{-i}(s_2 I - A)^{-j} \\
&= \frac{1}{s_2 - s_1}(s_1 I - A)^{-i}(s_2 I - A)^{-j+1} \\
&\quad + \frac{1}{s_1 - s_2}(s_2 I - A)^{-i+1}(s_2 I - A)^{-j} \qquad (3.12) \\
&= \sum_{k=1}^{i} \alpha_k (s_1 I - A)^{-k} + \sum_{l=1}^{j} \beta_l (s_2 I - A)^{-l}, \qquad (3.13)
\end{aligned}
$$

where the last equation is obtained by recursively applying equation (3.12). The coefficients α_i and β_j are not explicitly given here. The important point is that they depend only on the points s_i and s_j, i.e. they are the same for any matrix A. Moreover, it is clear that the coefficients related to the highest moments of the partial fraction expansion of $(s_1 I - A)^{-i}(s_2 I - A)^{-j}$, i.e. α_i and β_j, are different from zero. Multiply both sides of equation (3.13) by B, and equation (3.7) is satisfied.

Secondly, suppose that $s_2 = \infty$. Then,

$$(s_1 I - A)^{-1} A = -I + s_1 (s_1 I - A)^{-1}. \qquad (3.14)$$

By recursively applying this equation to $(s_1 I - A)^{-i} A^j$ and following the same reasoning as before, we find that

$$
\begin{aligned}
&Im\left((s_1 I - A)^{-i} A^j B\right) \subseteq \mathcal{K}_i(A_1, B_1) \quad if\ i > j, \qquad (3.15) \\
&Im\left((s_1 I - A)^{-i} A^j B\right) \subseteq \mathcal{K}_i(A_1, B_1) + \mathcal{K}_{j-i+1}(A, B) \quad if\ i \leq j. \qquad (3.16)
\end{aligned}
$$

Hence, equation (3.8) is satisfied.

□

Another proof of the following lemma may be found in [1].

Lemma 3.9. *Let $T(s)$ be a strictly proper SISO LTI transfer function of Mc Millan degree n with a state space realization $T(s) = C(sI - A)^{-1}B$. Let*

$$I = \{(s_1, m_1), \ldots, (s_r, m_r)\},$$

be a minimal $T(s)$-admissible interpolation set. Then

1. $Im\,(\mathcal{C}_{A,B}(I)) = Im\,(Contr(A, B))$.
2. $Ker\,(\mathcal{O}_{C,A}(I)) = Ker\,(Obs(C, A))$.

Proof. In the sequel, we drop the subscripts C, A, B. We prove only the first statement, the second one follows by transposition. For simplicity, we suppose that there is no point at infinity. This case can be treated similarly but requires more tedious notation. The proof consists of showing that the condition of Lemma 3.1 is satisfied.

From the set I, define $\forall i:\ 1 \leq i \leq m_i$

$$\tilde{\Gamma}(1,i) := \Gamma(s_1, i),$$

where $\Gamma(\lambda, k)$ is defined in equations (3.2) and (3.1).

Define $\forall i:\ 2 \leq i \leq r$, $\forall k:\ 1 \leq k \leq m_i$,

$$\tilde{\Gamma}(i,k) := \left(\prod_{j=1}^{i-1} A_j^{m_j} \right) \Gamma(s_i, k).$$

Define the matrix

$$\tilde{\mathcal{C}}(I) = \left[\tilde{\Gamma}(1,1), \tilde{\Gamma}(1,2), \ldots, \tilde{\Gamma}(r, m_r) \right].$$

As a consequence of Lemma 3.8, we obtain

$$Im\left(\tilde{\mathcal{C}}(I)\right) = Im\left(\mathcal{C}(I)\right).$$

Now, we use Lemmata 3.1 and 3.4. The matrix

$$N = \prod_{i=1}^{r} (A_i)^{-m_i} \tilde{\mathcal{C}}(I),$$

satisfies the condition of Lemma 3.1 because every column is a polynomial function of A of a different order, with degree smaller than n. Hence, $Im(N) = Im(Contr(A,B))$. By Lemma 3.4, $Im(N) = Im\left(\mathcal{C}(I)\right)$. This concludes the proof.

□

Lemma 3.10. *Consider an arbitrary pair of matrices (A, B) with $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times 1}$. Let $\mathcal{X}$ be a right invariant subspace of A. If*

$$Im(A^i B) \subseteq \mathcal{K}_i(A,B) + \mathcal{X}, \tag{3.17}$$

then, $\forall\ k \in \mathbb{N}$,

$$Im(A^{i+k} B) \subseteq \mathcal{K}_i(A,B) + \mathcal{X}.$$

Proof. Let us prove it for $k = 1$.

$$\begin{aligned} Im(A^{i+1}B) &= A Im(A^i B) \\ &\subseteq A\mathcal{K}_i(A,B) + A\mathcal{X} \\ &\subseteq \mathcal{K}_i(A,B) + Im(A^i B) + \mathcal{X} \\ &= \mathcal{K}_i(A,B) + \mathcal{X}. \end{aligned}$$

An easy induction completes the proof.

□

Lemma 3.11. *Let $T(s)$ be a strictly proper SISO transfer function of Mc Millan degree n. Let $\mathcal{X}$ (resp. $\mathcal{Y}$) be a right (resp. left) invariant subspace of A of dimension K. Let the columns of the matrix $X \in \mathbb{C}^{n\times K}$ (resp. $Y \in \mathbb{C}^{K\times n}$) be a basis of $\mathcal{X}$ (resp. $\mathcal{Y}$). Let I be a $T(s)$-admissible interpolation set of size $n-K$, denoted again by*

$$I = \{(s_1, m_1), \ldots, (s_r, m_r)\}.$$

Let the triple (C, A, B) be a minimal realization of $T(s)$. Then

1. $rank\left(\left[X\ \mathcal{C}_{A,B}(I)\right]\right) = n.$
2. $rank\left(\begin{bmatrix} Y \\ \mathcal{O}_{C,A}(I)\end{bmatrix}\right) = n.$

Proof. Only the first part of the Lemma will be proved; the second one follows by transposition. As a consequence of Lemma 3.9,

$$rank\,(\mathcal{C}_{A,B}(I)) = n - K.$$

Indeed, the $T(s)$-admissible interpolation set I may be seen as a subset of a minimal $T(s)$-admissible interpolation set of $T(s)$. Hence, the columns of $\mathcal{C}_{A,B}(I)$ must be linearly independent. Let us consider the first column of $\mathcal{C}_{A,B}(I)$. The matrices A_i and B_i associated with the point s_i are defined as usual. Suppose that

$$dim\,(\mathcal{X} + \mathcal{K}_{m_1}(A_1, B_1) + \ldots + \mathcal{K}_{m_r}(A_r, B_r)) = q < K + m_1 + \ldots + m_r = n.$$

Then, necessarily, $\exists p:\ 1 \le p \le r$ and $k_p:\ 0 \le k_p \le m_p - 1$ such that

$$Im(A_p^{k_p} B_p) \subseteq \mathcal{X} + \mathcal{K}_{m_1}(A_1, B_1) + \ldots + \mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) + \mathcal{K}_{k_p}(A_p, B_p). \tag{3.18}$$

Some care must be taken when $s_p = \infty$. Firstly, suppose that $s_p = \infty$ and $k_p = 0$. Then, multiply both sides of equation (3.18) by A. From Lemmata 3.8 and 3.10, and equations (3.15) to (3.16), we obtain the following relations :

$$\begin{aligned} Im(AB) &\subseteq A\mathcal{X} + A\mathcal{K}_{m_1}(A_1, B_1) + \ldots + A\mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) \\ &\subseteq \mathcal{X} + \mathcal{K}_{m_1}(A_1, B_1) + \ldots + \mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) + Im(B) \\ &= \mathcal{X} + \mathcal{K}_{m_1}(A_1, B_1) + \ldots + \mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}), \end{aligned}$$

where the last equation comes from equation (3.18) with $s_p = \infty$ and $k_p = 0$. But this implies that

$$dim\,(\mathcal{K}_n(A, B)) \le q < n.$$

This contradicts the fact that the pair (A, B) is controllable.

If $s_p = \infty$ and $k_p > 0$, then

$$\begin{aligned} Im(A^{k_p+1}B) &\subseteq A\mathcal{X} + A\mathcal{K}_{m_1}(A_1, B_1) + \ldots \\ &\quad + A\mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) + A\mathcal{K}_{k_p}(A, B) \\ &\subseteq \mathcal{X} + \mathcal{K}_{m_1}(A_1, B_1) + \ldots + \mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) \\ &\quad + \mathcal{K}_{k_p}(A_p, B_p) + Im(A^{k_p}B) \\ &= \mathcal{X} + \mathcal{K}_{m_1}(A_1, B_1) + \ldots \\ &\quad + \mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) + \mathcal{K}_{k_p}(A_p, B_p), \end{aligned}$$

and again, the transfer function $T(s)$ is not of Mc Millan degree n.

Suppose now that $\forall i: \ 1 \leq i \leq p, s_i \neq \infty$, and multiply again both sides of equation (3.18) by A_p. From Lemmata 3.8 and 3.10, we find that

$$Im(A_p^{k_p+1}B_p) \subseteq \mathcal{X} + \mathcal{K}_{m_1}(A_1, B_1) + \ldots + \mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) + \mathcal{K}_{k_p}(A_p, B_p). \tag{3.19}$$

This implies that

$$dim\,(\mathcal{K}_n(A_p, B_p)) \leq q < n.$$

But Lemma 3.9 implies $dim\,(\mathcal{K}_n(A_p, B_p)) = n$ and this is impossible. Finally, suppose that $\exists\ 1 \leq i \leq p$ such that $s_i = \infty$. From our previous discussion, $i < p$. For simplicity, suppose that $s_1 = \infty$. In such a case, by following the same reasoning as before,

$$\begin{aligned} Im(A_p^{k_p+1}B_p) &\subseteq A_p\mathcal{X} + \mathcal{K}_{m_1}(A, B) + A_p\mathcal{K}_{m_2}(A_2, B_2) + \ldots \\ &\quad + A_p\mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) + \mathcal{K}_{k_p}(A_p, B_p) + Im(A_p^{k_p}B_p) \\ &\subseteq \mathcal{X} + \mathcal{K}_{m_1}(A, B) + \mathcal{K}_{m_2}(A_2, B_2) + \ldots \\ &\quad + \mathcal{K}_{m_{p-1}}(A_{p-1}, B_{p-1}) + \mathcal{K}_{k_p}(A_p, B_p). \end{aligned}$$

This is again a contradiction with the controllability of the pair (A, B).

□

3.2 Model Reduction via Rational Interpolation

In this section, the general problem of constructing a transfer function $\hat{T}(s)$ of Mc Millan degree k that satisfies a set of $2k$ interpolation conditions with respect to a transfer function $T(s)$ of Mc Millan degree n is solved using Krylov subspaces. $T(s)$ and $\hat{T}(s)$ are both assumed to be SISO strictly proper transfer functions. Let us begin with a definition.

Definition 3.12. *Let $T(s)$ and $\hat{T}(s)$ be two strictly proper SISO transfer functions of respective Mc Millan degree n and k. Let I be a $T(s)$-admissible interpolation set of size $2k$, denoted by*

$$I = \{(s_1, m_1), \ldots, (s_r, m_r)\}.$$

We say that $\hat{T}(s)$ interpolates $T(s)$ at I when the following conditions are satisfied :

1. $\forall i: \ 1 \leq i \leq r$ *such that* $s_i \neq \infty$,

$$T(s) - \hat{T}(s) = O(s - s_i)^{m_i}. \tag{3.20}$$

2. *If ∞ is a point of I, say $s_k = \infty$, then*

$$T(s) - \hat{T}(s) = O(s^{-1})^{m_k+1}. \tag{3.21}$$

Remark 3.13. Because I is $T(s)$-admissible, the interpolation points s_i are not poles of $T(s)$. As a consequence, if $\hat{T}(s)$ interpolates $T(s)$ at I, then the interpolation points are not poles of $\hat{T}(s)$, i.e. I is $\hat{T}(s)$-admissible as well.

Let us consider a minimal realization (C, A, B) of $T(s)$ and a minimal realization $(\hat{C}, \hat{A}, \hat{B})$ of a transfer function $\hat{T}(s)$ that interpolates $T(s)$ at an interpolation set I. Writing equation (3.20) is equivalent to imposing the m_i first coefficients of the Taylor expansions of $\hat{T}(s)$ and $T(s)$ about s_i to be equal, i.e.

$$\hat{C}(s_i I_k - \hat{A})^{-k}\hat{B} = C(s_i I_n - A)^{-k}B, \quad \forall k: \ 1 \leq k \leq m_i.$$

Equation (3.21) is equivalent to imposing the m_k first Markov parameters of both transfer functions to be equal, i.e.

$$\hat{C}\hat{A}^i\hat{B} = CA^iB, \quad \forall i: \ 0 \leq i \leq m_k - 1.$$

Lemma 3.14. *Let $T(s) = C(sI_n - A)^{-1}B$ be any strictly proper SISO transfer function and let I be a $T(s)$-admissible interpolation set. Let $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ be any strictly proper SISO transfer function. Then $\hat{T}(s)$ interpolates $T(s)$ at I if and only if either of the following two equivalent conditions hold :*

$$C\mathcal{C}_{A,B}(I) = \hat{C}\mathcal{C}_{\hat{A},\hat{B}}(I) \quad , \quad \mathcal{O}_{C,A}(I)B = \mathcal{O}_{\hat{C},\hat{A}}(I)\hat{B}. \tag{3.22}$$

Proof. It is simply another way to write down the interpolation conditions of Definition 3.12.

□

Lemma 3.15. *Let $T(s) = C(sI_n - A)^{-1}B$ be a strictly proper SISO transfer function. Let I be a $T(s)$-admissible interpolation set and (I_1, I_2) be a symmetric separation of I. If the strictly proper SISO transfer function $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ interpolates $T(s)$ at I, then*

$$\mathcal{O}_{C,A}(I_1)\mathcal{C}_{A,B}(I_2) = \mathcal{O}_{\hat{C},\hat{A}}(I_1)\mathcal{C}_{\hat{A},\hat{B}}(I_2). \tag{3.23}$$

Proof. Define A_i and B_i as usual, and consider one element of the matrix equality (3.23). We have to prove that

$$CA_i^{k_1}A_j^{k_2}B = \hat{C}\hat{A}_i^{k_1}\hat{A}_j^{k_2}\hat{B}. \tag{3.24}$$

From Lemma 3.8, we can rewrite this equation by partial fraction expansion as a linear combination of equation (3.22). This completes the proof of the lemma. □

For an arbitrary interpolation set $I = \{I_l, I_r\}$ of size $2k$, the $k \times k$ matrix appearing in (3.23),

$$\mathcal{L}_{T(s)}(I) := \mathcal{O}_{C,A}(I_l)\mathcal{C}_{A,B}(I_r), \tag{3.25}$$

is the key to solving the interpolation problem. This matrix that does not depend on the choice of the particular state space realization (C, A, B) of $T(s)$ but only on the value of $T(s)$ at the interpolation points of I is called a Loewner matrix. Generically, this square matrix is nonsingular. In such a case, there exists generically only one transfer function of Mc Millan degree k that interpolates $T(s)$ at I, and there is no interpolating transfer function of Mc Millan degree smaller than k. Let us prove this last assertion. From Lemma 3.15, if $\hat{T}(s) := \hat{C}(sI - \hat{A})^{-1}\hat{B}$ interpolates $T(s)$ at I, then

$$\mathcal{L}_{T(s)}(I) = \mathcal{O}_{\hat{C},\hat{A}}(I_l)\mathcal{C}_{\hat{A},\hat{B}}(I_r). \tag{3.26}$$

Because $\mathcal{L}_{T(s)}(I)$ is nonsingular, the rank of $\mathcal{O}_{\hat{C},\hat{A}}(I_l)$ and the rank of $\mathcal{C}_{\hat{A},\hat{B}}(I_r)$ must be at least equal to k, and this is only possible if the Mc Millan degree of $\hat{T}(s)$ is larger than k. Uniqueness will be proved in Theorem 3.17.

Lemma 3.16. *Let $T(s) = C(sI_n - A)^{-1}B$ be a strictly proper SISO transfer function of Mc Millan degree n. Let I be a $T(s)$-admissible interpolation set of size $2k$ and (I_1, I_2) be a symmetric separation of I. Suppose that $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ is a strictly proper SISO transfer function of Mc Millan degree k, which interpolates $T(s)$ at I. Then*

$$\mathcal{O}_{C,A}A\mathcal{C}_{A,B} = \mathcal{O}_{\hat{C},\hat{A}}\hat{A}\mathcal{C}_{\hat{A},\hat{B}}. \tag{3.27}$$

Proof. Define A_i and B_i as usual, and consider again one element of the matrix equality (3.27). We have to prove that

$$CA_i^{k_1}AA_j^{k_2}B = \hat{C}\hat{A}_i^{k_1}\hat{A}\hat{A}_j^{k_2}\hat{B}. \tag{3.28}$$

The idea is that using partial fraction expansion it is possible to rewrite equation (3.28) as a linear combination of equations (3.22) and (3.23).

The point at infinity requires more care. We show it for instance when $A_i = A$. From Definition 3.7, this implies that one of the points of I_1, say $s_{1,1}$ is equal to ∞. Then,

$$CA^{u-1}B = \hat{C}\hat{A}^{u-1}\hat{B}, \quad \forall u: \ 1 \leq u \leq m_{1,1}.$$

If $A_j = A$, then the point ∞ is also a point of I_2, say $s_{2,1} = \infty$. Then,

$$CA^{v-1}B = \hat{C}\hat{A}^{v-1}\hat{B}, \quad \forall v: \ 1 \leq v \leq m_{2,1}.$$

Clearly, the point ∞ must be a point of I, say $s_1 = \infty$. Because (I_1, I_2) is a separation of I, $m_{1,1} + m_{2,1} = m_1$, and

$$CA^{w-1}B = \hat{C}\hat{A}^{w-1}\hat{B}, \quad \forall w: \ 1 \leq w \leq m_1. \tag{3.29}$$

Now, $k_1 + 1 + k_2 \leq m_{1,1} + m_{2,1} - 1 = m_1 - 1$, and equality (3.28) follows from equation (3.29). This concludes the proof for the case $A_i = A_j = A$. Suppose now that $A_j = (s_j I - A)^{-1}$ and $A_i = A$. Then,

$$CA_j^v B = \hat{C}\hat{A}_j^v\hat{B}, \quad \forall v: \ 1 \leq v \leq m_{2,j}.$$

From partial fraction expansion, it follows then that

$$CA^{k_1}AA_j^{k_2}B = -CA^{k_1}A_j^{k_2-1}B + s_j CA^{k_1}A_j^{k_2}B$$

Now, equation (3.28) follows from Lemmas 3.15 and 3.14. This completes the proof for $A_i = A$. The case $A_i \neq A$ is shown by using similar arguments and is omitted here.

□

This leads to the main result of this section.

Theorem 3.17. *Let $T(s) = C(sI_n - A)^{-1}B$ be a strictly proper SISO transfer function of Mc Millan degree n. Let I be a $T(s)$-admissible interpolation set of size $2k$ and let (I_1, I_2) be a symmetric separation of I. Suppose that $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ is a strictly proper SISO transfer function of Mc Millan degree k, which interpolates $T(s)$ at I. Then $\hat{T}(s)$ can be obtained by truncation of $T(s)$ with*

$$Z^T = \mathcal{O}_{\hat{C},\hat{A}}(I_1)^{-1}\mathcal{O}_{C,A}(I_1) \quad , \quad V = \mathcal{C}_{A,B}(I_2)\mathcal{C}_{\hat{A},\hat{B}}(I_2)^{-1}. \tag{3.30}$$

Moreover, $\hat{T}(s)$ is the unique transfer function of minimal Mc Millan degree k that interpolates $T(s)$ at I.

Proof. I_1 and I_2 are a both minimal $\hat{T}(s)$-admissible interpolation sets. From Lemma 3.9, the matrices $\mathcal{O}_{\hat{C},\hat{A}}(I_1)$ and $\mathcal{C}_{\hat{A},\hat{B}}(I_2)$ are invertible. From Lemmas 3.14 to 3.16, it is easy to check that conditions (2.3) of Definition 2.1 are satisfied with Z and V defined in equation (3.30).

Let us prove the uniqueness assertion. Assume that $T_1(s) := C_1(sI_k - A_1)^{-1}B_1$ and $T_2(s) := C_2(sI_k - A_2)^{-1}B_2$ are two transfer functions of degree k

that interpolate $T(s)$ at I. From the preceding discussion, there exist two pairs (Z_1, V_1) and (Z_2, V_2) such that

$$Im(V_1) = Im(V_2) = Im\left(\mathcal{C}_{A,B}(I_r)\right), \quad Im(Z_1) = Im(Z_2) = Im\left(\mathcal{O}_{A,B}(I_r)^T\right), \tag{3.31}$$

with the biorthogonality constraint

$$Z_1^T V_1 = Z_2^T V_1 = I_k, \tag{3.32}$$

that satisfy the following properties :

$$(C_1, A_1, B_1) = (CV_1, Z_1^T AV_1, Z_1^T B), \quad (C_2, A_2, B_2) = (CV_2, Z_2^T AV_2, Z_2^T B). \tag{3.33}$$

From (3.31), there exist $k \times k$ invertible matrices L and R such that

$$Z_1^T = LZ_2^T, \quad V_1 = V_2 R.$$

From (3.32), $L = R^{-1}$ and from (3.33),

$$(C_1, A_1, B_1) = (C_2 R,\ R^{-1} A_2 R,\ R^{-1} B_2).$$

This clearly implies that $T_1(s) = T_2(s)$.

□

Remark 3.18. Theorem 3.17 provides a practical way to construct a transfer function of minimal Mc Millan degree that interpolates a SISO transfer function $T(s) := C(sI_n - A)^{-1}B$ of degree n at an interpolation set I of size $2k$ as follows.

1. First, compute the $n \times k$ Krylov matrices $Z^T = \mathcal{O}_{C,A}(I_l),\ V = \mathcal{C}_{A,B}(I_r)$ where $\{I_l, I_r\}$ is a symmetric separation of I.
2. If $Z^T V$ is invertible, compute the SVD decomposition $Z^T V = \hat{U}\Sigma\hat{V}^T$ and the projecting matrices $\tilde{Z} = Z\hat{U}\Sigma^{-1/2}, \quad \tilde{V} = V\hat{V}\Sigma^{-1/2}$.
3. Compute the reduced order state space realization

$$(C_{MP}, A_{MP}, B_{MP}) := (C\tilde{V}, \tilde{Z}^T A\tilde{V}, \tilde{Z}^T B).$$

   If the transfer function $T_{MP}(s) := C_{MP}(sI_k - A_{MP})^{-1}B_{MP}$, called the *Multipoint Pade* transfer function, has no pole at the interpolation points belonging to I, then $T_{MP}(s)$ is the unique transfer function of Mc Millan degree k that interpolates $T(s)$ at I.

Remark 3.19. The preceding theorem shows that if there exists a transfer function $\hat{T}(s)$ of Mc Millan degree k that interpolates a transfer function $T(s)$ of larger Mc Millan degree at a interpolation set I of size $2k$, then $\hat{T}(s)$ is the unique interpolating transfer function of degree k, there exist no interpolating transfer function of smaller Mc Millan degree and $\hat{T}(s)$ can be constructed by

a projection using Krylov subspaces corresponding to the interpolation conditions. But it does not say what are the necessary and sufficient conditions for the existence of a unique interpolating transfer function of Mc Millan degree k. First, from equation (3.26) and the discussion following it, there exists no interpolating transfer function of Mc Millan degree smaller than the rank of $\mathcal{L}_{T(s)}(I)$. So, a necessary condition for the existence and uniqueness of an interpolating transfer function of Mc Millan degree k is that the Loewner matrix $\mathcal{L}_{T(s)}(I) \in \mathbb{C}^{k\times k}$ is invertible. A necessary and sufficient condition for the existence and uniqueness of an interpolating transfer function of Mc Millan degree k is that the interpolation points are not poles of the Multipoint Pade transfer function $T_{MP}(s)$. This is verified if for every interpolation point $s_i \in I$, the matrix $\mathcal{O}_{C,A}(I_l)(s_iI - A)\mathcal{C}_{A,B}(I_r)$ is nonsingular (see [26, 48] for a proof). A generalization of the preceding discussion related to the system zero pencil is given in the next chapter. For the case where $\mathcal{L}_{T(s)}(I)$ is singular, we refer to [5].

As a consequence, if the SISO strictly proper transfer function $\hat{T}(s)$ of Mc Millan degree k interpolates the SISO strictly proper transfer function $T(s)$ of degree $n > k$ at an interpolation set of size $2k$, then $\hat{T}(s)$ can be obtained from $T(s)$ by a projection using Krylov subspaces. The special case where the interpolation set between $T(s)$ and $\hat{T}(s)$ is of size less than $2k$ (i.e. when the Mc Millan degree of $T(s) - \hat{T}(s)$ is less than or equal to $2k$) is treated in the next section.

3.3 Model Reduction via Truncation

The following lemma is well-known in the literature (see for instance [92] and references therein for a proof).

Lemma 3.20. *Let the pair of matrices $A \in \mathbb{C}^{n\times n}$ and $C \in \mathbb{C}^{1\times n}$ be observable. Let $\mathcal{X}$ be a right invariant subspace of A of dimension K and let the matrix $X \in \mathbb{C}^{n\times K}$ be full rank with $\mathcal{X} = Im(X)$. Define the matrices $\tilde{A} \in \mathbb{C}^{K\times K}$ and $\tilde{C} \in \mathbb{C}^{1\times K}$ by the following equations :*

$$AX = X\tilde{A}, \quad CX = \tilde{C}.$$

Then, the pair $(\tilde{C}, \tilde{A})$ is observable and

$$\Lambda\left(\tilde{A}\right) \subseteq \Lambda\left(A\right). \tag{3.34}$$

Remark 3.21. Since changing the basis X to XS results in a transformed pair $(S^{-1}AS, S^{-1}C)$, it is always possible to choose the basis X of the invariant subspace $\mathcal{X}$ such that the pair $(\tilde{A}, \tilde{C})$ is in observable canonical form (see for instance [94]).

Theorem 3.22. *Choose $T(s) = C(sI_n - A)^{-1}B$, an arbitrary strictly proper SISO transfer function of Mc Millan degree n. Choose $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$, an arbitrary strictly proper SISO transfer function of Mc Millan degree $k < n$. Then $\hat{T}(s)$ can be constructed from $T(s)$ via truncation.*

Proof. By a recursive argument, it is not difficult to see that this theorem is true for every $n > k$ if and only if it is true for $n = k+1$. We therefore prove it for $n = k+1$ only. The proof is constructive : we construct Z and V such that the conditions of Definition 2.1 are satisfied. Define

$$T(s) := \frac{n(s)}{d(s)}, \quad \hat{T}(s) := \frac{\hat{n}(s)}{\hat{d}(s)},$$

where $d(s)$ and $\hat{d}(s)$ are monic polynomials of degree $k+1$ and k, and where $degree\,(n(s)) < k+1$ and $degree\,(\hat{n}(s)) < k$. Because the Mc Millan degree of $T(s)$ is $k+1$ and that of $\hat{T}(s)$ is k, the polynomials $n(s)$ and $d(s)$ are coprime, and $\hat{n}(s)$ and $\hat{d}(s)$ are coprime as well. Define the error transfer function $E(s)$ to be

$$E(s) := T(s) - \hat{T}(s) = \frac{n(s)\hat{d}(s) - \hat{n}(s)d(s)}{d(s)\hat{d}(s)} := \frac{n_E(s)}{d_E(s)},$$

with

$$\begin{aligned} K &= degree\left(\gcd\left(d(s), \hat{d}(s)\right)\right) \\ degree\,(d_E(s)) &= 2k+1-K \\ degree\,(n_E(s)) &< 2k+1-K. \end{aligned}$$

We can write

$$\begin{aligned} \gcd\left(d(s), \hat{d}(s)\right) &= (s-\beta_1)^{\nu_1}\cdots(s-\beta_p)^{\nu_p}, \\ \sum_{i=1}^{p} \nu_i &= K. \end{aligned}$$

Without loss of generality we can also write

$$\begin{aligned} d(s) &= (s-\beta_1)^{n_1}\cdots(s-\beta_q)^{n_q} \\ \hat{d}(s) &= (s-\hat{\beta}_1)^{\hat{n}_1}\cdots(s-\hat{\beta}_q)^{\hat{n}_{\hat{q}}}; \end{aligned}$$

where, $\forall i:\ 1 \le i \le p$,

$$\hat{\beta}_i = \beta_i, \quad \min(n_i, \hat{n}_i) = \nu_i.$$

Clearly, $E(s)$ has $2k - K + 1$ zeros, with at least one zero at ∞. Those zeros are the points where $\hat{T}(s)$ interpolates $T(s)$. More precisely, we can write

$$n_E(s) := \kappa(s-\alpha_1)^{\sigma_1}\cdots(s-\alpha_z)^{\sigma_z},$$
$$\sum_{i=1}^{z}\sigma_i \;=\; 2k+1-K-\sigma_{z+1},$$

where $\sigma_{z+1} \in \mathbb{N}_0$ is the multiplicity of the zero at ∞ of $E(s)$. Indeed, it is not difficult to check from our definitions that

$$\lim_{s\to\infty} T(s)s^{\sigma_{z+1}} = \kappa,$$

where $\kappa \in \mathbb{C}$ is the gain of the transfer function $E(s)$. Moreover, $T(s)$ and $\hat{T}(s)$ have K poles in common. If $K = 0$, then $\hat{T}(s)$ can be constructed from truncation of $T(s)$ via rational interpolation (see Theorem 3.17). We now suppose that $K > 0$.

Clearly, $K \leq k$ and from Lemma 3.20, it is always possible to find a full rank matrix $X_1 \in \mathbb{C}^{(k+1)\times K}$ such that the following relations hold :

$$AX_1 = X_1\tilde{A}, \quad CX_1 = \tilde{C},$$

where

$$\tilde{C} = \begin{bmatrix}1, 0, \ldots, 0\end{bmatrix} \quad , \quad \tilde{A} = \begin{bmatrix} -a_1 & 1\;0\ldots 0 \\ -a_2 & 0\;1\ldots 0 \\ \vdots & \vdots\;\vdots \quad \vdots \\ -a_{K-1} & 0\;0\ldots 1 \\ -a_K & 0\;0\ldots 0 \end{bmatrix}, \tag{3.35}$$

and

$$s^K + a_1 s^{K-1} + \ldots + a_K = (s-\beta_1)^{\nu_1}\cdots(s-\beta_p)^{\nu_p}.$$

This is indeed the observer canonical form associated with the common spectrum of $T(s)$ and $\hat{T}(s)$. Similarly, there exists a matrix $\hat{X}_1 \in \mathbb{C}^{k\times K}$ such that the following relations hold :

$$\hat{A}\hat{X}_1 = \hat{X}_1\tilde{A}, \quad \hat{C}\hat{X}_1 = \tilde{C}. \tag{3.36}$$

Now, we focus our attention on the $2k-K+1$ interpolation conditions. If $\sigma_{z+1} = 1$, we define the interpolation set I to be

$$I := \{(\alpha_1,\sigma_1),\ldots,(\alpha_z,\sigma_z)\}.$$

Otherwise, $\sigma_{z+1} > 1$, and we then define I to be

$$I := \{(\alpha_1,\sigma_1),\ldots,(\alpha_z,\sigma_z),(\infty,\sigma_{z+1}-1)\}.$$

Clearly, I is a $T(s)$-admissible set of size $2k-K$. We separate this set into two $T(s)$-admissible sets. The first one, I_1, is of size k and the second one, I_2, is of size $k-K$. Define

$$Y := \mathcal{O}_{C,A}(I_1), \quad \hat{Y} := \mathcal{O}_{\hat{A},\hat{C}}(I_1), \quad X_2 := \mathcal{C}_{A,B}(I_2), \quad \hat{X}_2 := \mathcal{C}_{\hat{A},\hat{B}}(I_2). \tag{3.37}$$

From Lemma 3.11, the matrices $\hat{Y}$ and

$$\hat{X} := \left[\, \hat{X}_1, \hat{X}_2 \,\right]$$

are invertible. Finally, define

$$X = \left[\, X_1, X_2 \,\right].$$

Now, we check that

$$CX = \hat{C}\hat{X}, \quad YB = \hat{Y}\hat{B}, \quad YX = \hat{Y}\hat{X}, \quad YAX = \hat{Y}\hat{A}\hat{X}. \tag{3.38}$$

To verify the first part of equation (3.38),

$$CX = \left[\, CX_1, CX_2 \,\right] = \left[\, \hat{C}\hat{X}_1, \hat{C}\hat{X}_2 \,\right] = \hat{C}\hat{X},$$

where the last equation follows from the construction of X_1 and $\hat{X}_1$ and Lemma 3.14. The second part of equation (3.38) follows from Lemma 3.14. Finally,

$$YX = \left[\, YX_1, YX_2 \,\right].$$

Let $\phi(A)$ be a polynomial function of A, then

$$C\phi(A)X_1 = CX_1\phi(\tilde{A}) = \hat{C}\hat{X}_1\phi(\tilde{A}) = \hat{C}\phi(\hat{A})\hat{X}_1,$$

where the matrix $\tilde{A} \in \mathbb{C}^{K\times K}$ is defined in (3.35). Hence, $YX_1 = \hat{Y}\hat{X}_1$ and the third part of equation (3.38) follows from Lemma 3.15. For the same reasons, $YAX_1 = \hat{Y}\hat{A}\hat{X}_1$, and the fourth part of equation (3.38) follows from Lemma 3.16. Choose then

$$V := X\hat{X}^{-1}, \quad Z^T := \hat{Y}^{-1}Y, \tag{3.39}$$

which implies that equations (2.3) of Definition 2.1 are satisfied. Hence, $\hat{T}(s)$ can be constructed by truncation of $T(s)$.

□

3.4 Sylvester Equations, Krylov Subspaces and Projection

In the preceding section, it was shown that any SISO strictly proper transfer function can be obtained by a projection of any SISO strictly proper transfer function of larger Mc Millan degree. Moreover, by looking at the proof of Theorem 3.22, the projecting matrices can be chosen such that their image is a sum of Krylov subspaces and eigenspaces (see (3.37)). In this section, it will

be shown that such projecting matrices can always be constructed by solving Sylvester equations of the following form

$$AVR_1 - VR_2 + BY = 0, \quad L_1 Z^T A - L_2 Z^T + XC = 0, \tag{3.40}$$

where (C, A, B) is a minimal state space realization of $T(s)$.

First assume that $V \in \mathbb{C}^{n\times k}$ is a Krylov matrix at a finite point $\lambda \in \mathbb{C}$, then

$$AV = (A - \lambda I + \lambda I)\left[(\lambda I - A)^{-1}B \ldots (\lambda I - A)^{-k}B\right] \tag{3.41}$$

$$= VJ_{0,k} + \lambda V - B\left[1\ 0 \ldots 0\right] = VJ_{\lambda,k} - B\left[1\ 0 \ldots 0\right], \tag{3.42}$$

i.e. equation (3.40) is satisfied with

$$(R_1, R_2) = (I_k, J_{\lambda,k}), \quad Y = \left[1\ 0 \ldots 0\right].$$

Note that the pair (Y, R_2) is observable.

Let us assume that V is a Krylov matrix with an infinite interpolation point. Then,

$$-AVJ_{0,k} + B\left[1\ 0 \ldots 0\right] = \left[B \ldots A^{k-1}B\right] = V, \tag{3.43}$$

i.e. equation (3.40) is satisfied with

$$(R_1, R_2) = (-J_{\lambda,k}, I_k), \quad Y = \left[1\ 0 \ldots 0\right].$$

Note that the pair (Y, R_1) is observable. Finally, if V is an eigenspace of A, then clearly there exists a matrix $R \in \mathbb{C}^{k\times k}$ whose eigenvalues are a subset of the eigenvalues of A such that

$$AV - VR = 0,$$

i.e. equation (3.40) is satisfied with

$$(R_1, R_2) = (I_k, R), \quad Y = \left[0 \ldots 0\right].$$

Note that the pairs (Y, R_1) and (Y, R_2) are unobservable. The preceding discussion leads to the following result.

Theorem 3.23. *Let $T(s) := C(sI_n - A)^{-1}B$ and $\hat{T}(s) := \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ be two SISO strictly proper transfer function of respective Mc Millan degree n and $k < n$. There always exist regular matrix pencils (L_1, L_2) and (R_1, R_2) and vectors X and Y such that*

$$(\hat{C}, \hat{A}, \hat{B}) = (CV, Z^T AV, Z^T B),$$

with projecting matrices $Z^T V = I_k$ satisfying the following Sylvester equations

$$AVR_1 - VR_2 + BY^T = 0, \quad L_1 Z^T A - L_2 Z^T + XC = 0, \tag{3.44}$$

Finally, the generalized eigenvalues of the pairs (R_1, R_2) and (V_1, V_2) are either interpolation points between $T(s)$ and $\hat{T}(s)$ or the inverse of common poles between $T(s)$ and $\hat{T}(s)$, depending on whether they are associated with the controllable and observable part of the pairs (Y, R_i) and (L_i, X) or not.

Proof. First, it is worth mentioning that if (3.40) is valid for Z and V, then for any invertible matrix S, it remains valid for the pair $(SZ^T, VS^{-1}$ by replacing the pairs (L_1, L_2) and (R_1, R_2) with (L_1S^{-1}, L_2S^{-1}) and (SR_1, SR_2). From Theorem 3.22, equations (3.36), (3.37) and (3.39), it is always possible to choose projecting matrices Z and V such that their images are a sum of Krylov subspaces of the pair (A, B) and eigenspaces of A. This can be proved by decomposing each of these subspaces, using equations (3.41), (3.43) and the properties of eigenspaces, and looking at the particular structures of the matrices L_i, R_i and vectors X and Y appearing in these equations.

□

Conversely, if a reduced order transfer function $\hat{T}(s) := \hat{C}(sI_k - \hat{A})^{-1}\hat{B} = CV(sI_k - Z^TAV)^{-1}Z^TB$ is constructed from projecting matrices satisfying Sylvester equations of the form (3.40) (with (C, A, B) a minimal state space realization of $T(s)$), then the generalized eigenvalues of the pairs (L_1, L_2) and (R_1, R_2) yield a subset of interpolation points and common poles between $T(s)$ and $\hat{T}(s)$.

Remark 3.24. If one wants $\hat{T}(s)$ to interpolate $T(s)$ at a point λ that tends to be equal to a pole of $T(s)$ (i.e. an eigenvalue of A), it seems reasonable that in the limit, $\hat{T}(s)$ would also have a pole at λ. It is possible to implement algorithms that perform this automatically, but some care must be taken. Indeed, a naive implementation of Multipoint Padé would be to compute the Krylov subspaces directly, giving rise to obvious numerical problems when trying to compute $(\lambda I - A)v = b$ with $\lambda I - A$ close to singularity. The key to overcome this is again to look at the Sylvester equations. One can for instance normalize the Sylvester equations (3.44) in such a way that the norm of the first column of V, call it v_1, is equal to one:

$$(A - \lambda I)v_1 = by_1.$$

As a consequence, the closer λ is from the spectrum of A, the smaller y_1 is. In the limit, if λ is an eigenvalue of A, y_1 is put to zero and v_1 is an eigenvector of A corresponding to the eigenvalue λ. If $v_1 \in Im(V)$, λ is also an eigenvalue of $\hat{A} = Z^TAV$ and $\hat{T}(s)$ admits a pole at $s = \lambda$.

3.5 Concluding Remarks

Generically, two SISO transfer functions $T(s)$ and $\hat{T}(s)$, of order $n + 1$ and n respectively, do not have common poles. Hence, almost every strictly proper SISO transfer function of Mc Millan degree k can be obtained from any strictly proper SISO transfer function of Mc Millan degree $n > k$ via *Multipoint Padé* interpolation. This implies that a reduced order transfer function constructed

using Krylov subspaces may yield an error of arbitrarily large norm. As a consequence, the interpolation points must be chosen with care when trying to construct a reduced order transfer function via *Multipoint Padé* techniques.

Another important result developed in this chapter is the link between Sylvester equations, Krylov subspaces and interpolation.

There are many open questions. Given a strictly proper SISO transfer function $T(s)$ of Mc Millan degree n, and a strictly proper SISO transfer function $\hat{T}(s)$ of Mc Millan degree $k < n$, we have constructed one set of projecting matrices Z and V such that $\hat{T}(s)$ can be obtained from truncation of $T(s)$. The solution set for the matrices V and Z is much larger. For instance, when there are more than $2k$ interpolation points we can choose any subset of $2k$ zeros to construct a pair of projectors V and Z. Partial answers to these questions are given in Chapter 5.

A more practical question about *Multipoint Padé* approximation is how to find interpolation conditions that ensure to have a global error bound between the original and the reduced order transfer functions? For instance, is it possible to find an easy characterization of the interpolation points between a transfer function and a reduced order system obtained by *balanced truncation* or *optimal Hankel norm approximation* technique? How to choose interpolation points such that the reduced order transfer function is stable, is also not yet completely answered, despite the fact that particular choices of interpolation points guarantee stability [4].

In order to simplify the proofs, another approach has been investigated related to the Chinese Remainder Theorem. Unfortunately, it has not been possible yet to simplify all the results using such an approach.

Generalizations of the preceding results in the MIMO case are developed in the next chapters. More precisely, the link between Sylvester equations, *generalized* Krylov subspaces and the problem of *tangential interpolation* is treated in Chapter 4. The generality of the embedding of state space realization in the MIMO case is studied in more details in Chapter 5.

3.6 Notes and References

This Chapter is based on the recent papers [40, 38] and is the result of a collaborative work with Kyle Gallivan and Paul Van Dooren.

Padé techniques for model reduction in linear system theory probably appear in the sixties (see [16] for a survey). Rational interpolation of SISO transfer functions in a state space context appears in [5]. A study of the partial realization problem is available in [47]. In 1987, de Villemagne and Skelton [26] where the first who proposed to construct reduced order systems by imposing a set of interpolation conditions, but they did not realize the generality of this approach. In his PhD thesis, Grimme [48, 34] studied in more details Krylov

techniques for model reduction in a generalized state space context and made connections with the Lanczos and Arnoldi algorithms. Since then, many authors have studied Krylov techniques for model reduction, see for instance the work of Freund [11], Bai [9] and Jaimoukha [58] to cite only a few.

4

Model Reduction by Moment Matching, the MIMO case

In this chapter, we address the problem of constructing a reduced order transfer function of minimal Mc Millan degree that satisfies a set of tangential interpolation conditions with respect to another transfer function. The resulting reduced order transfer function appears to be generically unique and we present a simple and efficient technique to construct this interpolating reduced order system. This is a generalization of the *Multipoint Padé* technique which is particularly suited to handle multi-input multi-output systems.

4.1 Introduction

Tangential interpolation of given input/output data has already been treated in the literature [7, 13]. Here, we address the case where these data are themselves obtained from tangential information of a given (large-scale) transfer function, which to our knowledge has not been considered.

In this chapter, we consider $p \times m$ *strictly proper* transfer functions $T(s)$, i.e. where $\lim_{s\to\infty} T(s) = 0$. This implies that the point at infinity is a zero of $T(s)$. For this reason, a separate treatment of the point at infinity is required.

We must use the well-established concept of zero of a system (see e.g. [72]) and the following related definition (in this chapter, the term *zero* corresponds to the term *transmission zero* of Definition 1.20).

Definition 4.1. *Suppose that $T(s)$ is a $p \times m$ rational function. A $m \times 1$ polynomial vector $y(s)$ is a right zero direction of order k at the zero λ if $y(\lambda) \neq 0$ and*

$$T(s)y(s) = O(\lambda - s)^k. \tag{4.1}$$

Analogously, a $1 \times p$ polynomial vector $x(s)$ is a left zero direction of $T(s)$ when $x^*(s)$ is a right zero of $T^*(s)$. The order of a zero is defined as the maximum order of the zero directions at this point. A zero λ is called a blocking zero of $T(s)$ if $T(\lambda) = 0$.

Note that for SISO systems, any zero is a blocking zero. Generically, MIMO transfer functions do not have blocking zeros. It then makes sense to develop Krylov techniques for MIMO systems that do not impose the error $T(s)-\hat{T}(s)$ to have blocking zeros but rather chosen left and right zero directions. This is the purpose of this chapter.

Remark 4.2. For MIMO systems, a zero can also be a pole. If λ is not a pole of $T(s)$, only the k first Taylor coefficients of $y(s)$ about λ are important. If λ is a pole of $T(s)$, the situation is more complicated. Indeed, assume that λ is a pole of order p of $T(s)$ and that $y(s)$ has an expansion about λ, then

$$T(s)y(s) = \left(\sum_{i=-p}^{+\infty} T_i(\lambda - s)^i\right)\left(\sum_{j=0}^{\infty} y_j(\lambda - s)^j\right). \tag{4.2}$$

We see that the first $k+p$ terms in the Taylor expansion of $y(s)$ are important to ensure that the product (4.2) has a zero of order k. This case will not be discussed here, but a few remarks will be made to indicate how it complicates the problem.

We now present the concept of tangential interpolation that will be considered in this chapter. Three concepts are defined, namely left, right and two-sided tangential interpolation. Interpolation at the point at infinity is considered as a special case.

Let z be a finite point in the complex plane. Let $T(s)$ and $\hat{T}(s)$ be two $p\times m$ strictly proper transfer functions that do not have a pole at $s=z$.

Left Tangential Interpolation

Let $x(s)$ be a $1\times p$ polynomial vector of degree $\beta-1$ and not equal to zero at $s=z$. We say that $\hat{T}(s)$ interpolates $T(s)$ at $(z, x(s))$ if

$$x(s)\left(T(s)-\hat{T}(s)\right) = O(z-s)^{\beta}. \tag{4.3}$$

Let $x(s)$ be a $1\times p$ polynomial vector in s^{-1}, of degree $\beta-1$ in s^{-1} and not equal to zero at $s=\infty$. We say that $\hat{T}(s)$ interpolates $T(s)$ at $(\infty, x(s))$ if

$$x(s)\left(T(s)-\hat{T}(s)\right) = O(s^{-1})^{\beta+1}. \tag{4.4}$$

Right Tangential Interpolation

Let $y(s)$ be a $m\times 1$ polynomial vector of degree $\delta-1$ and not equal to zero at $s=z$. We say that $\hat{T}(s)$ interpolates $T(s)$ at $(z, y(s))$ if

$$\left(T(s)-\hat{T}(s)\right)y(s) = O(z-s)^{\delta}. \tag{4.5}$$

Let $y(s)$ be a $m\times 1$ polynomial vector in s^{-1}, of degree $\delta-1$ in s^{-1} and not equal to zero at $s=\infty$. We say that $\hat{T}(s)$ interpolates $T(s)$ at $(\infty, y(s))$ if the following condition is satisfied :

$$\left(T(s) - \hat{T}(s)\right) y(s) = O(s^{-1})^{\delta+1}. \tag{4.6}$$

Two Sided Tangential Interpolation

Let $x(s)$ be a $1 \times p$ polynomial vector of degree $\beta - 1$ and not equal to zero at $s = z$. Let $y(s)$ be a $m \times 1$ polynomial vector of degree $\delta - 1$ and not equal to zero at $s = z$. We say that $\hat{T}(s)$ interpolates $T(s)$ at $(z, x(s), y(s))$ if the following condition is satisfied :

$$x(s) \left(T(s) - \hat{T}(s)\right) y(s) = O(z - s)^{\beta+\delta}. \tag{4.7}$$

Let $x(s)$ be a $1 \times p$ polynomial vector in s^{-1}, of degree $\beta - 1$ in s^{-1} and not equal to zero at $s = \infty$. Let $y(s)$ be a $m \times 1$ polynomial vector in s^{-1}, of degree $\delta - 1$ in s^{-1} and not equal to zero at $s^{-1} = 0$. We say that $\hat{T}(s)$ interpolates $T(s)$ at $(\infty, x(s), y(s))$ if the following condition is satisfied :

$$x(s) \left(T(s) - \hat{T}(s)\right) y(s) = O(s^{-1})^{\beta+\delta+1}. \tag{4.8}$$

The objective of this chapter is the following. We are given a transfer function $T(s)$ and a set of tangential interpolation conditions of the type (4.3) to (4.8) at a number of points of the complex plane, and we want to construct the transfer function of minimal Mc Millan degree that satisfies these interpolation conditions. In order to make the problem more precise, we need to introduce the following concepts.

Definition 4.3. *Let $z_1, \ldots, z_{k_{left}}$ be points in the extended complex plane, not necessarily distinct or finite. For each finite z_α, a $1 \times p$ polynomial vector $x_\alpha(s)$ of degree $\beta_\alpha - 1$ and not equal to zero at $s = z_\alpha$ is given :*

$$x_\alpha(s) = \sum_{j=0}^{\beta_\alpha - 1} x_\alpha^{[j]} (z_\alpha - s)^j, \quad x_\alpha^{[0]} \neq 0. \tag{4.9}$$

If $z_\alpha = \infty$, then a $1 \times p$ polynomial vector in s^{-1}, $x_\alpha(s)$ of degree $\beta_\alpha - 1$ in s^{-1} and not equal to zero at $s = \infty$ is given :

$$x_\alpha(s) = \sum_{j=0}^{\beta_\alpha - 1} x_\alpha^{[j]} s^{-j}, \quad x_\alpha^{[0]} \neq 0. \tag{4.10}$$

The left interpolation set I_{left} is defined as follows

$$I_{left} := \left\{ (z_1, x_1(s)), \ldots, \left(z_{k_{left}}, x_{k_{left}}(s)\right) \right\}. \tag{4.11}$$

The size of I_{left}, written $size(I_{left})$, is defined as follows :

$$size(I_{left}) := \sum_{i=1}^{k_{left}} \beta_i. \tag{4.12}$$

Finally, the set of interpolation points of I_{left}, written $p(I_{left})$ is defined as follows :

$$p(I_{left}) = \{z_1, \ldots, z_{k_{left}}\}. \tag{4.13}$$

Analogously, a right tangential interpolation set

$$I_{right} := \left\{(w_1, y_1(s)), \ldots, \left(w_{k_{right}}, y_{k_{right}}(s)\right)\right\}, \tag{4.14}$$

with the points $w_1, \ldots, w_{k_{right}}$ arbitrarily chosen in $\mathbb{C} \cup \infty$ and each $m \times 1$ polynomial vector $y_\alpha(s)$, $1 \leq \alpha \leq k_{right}$ of degree $\delta_\alpha - 1$ in s if w_α is finite (of degree $\delta_\alpha - 1$ in s^{-1} otherwise) defined with the same conventions as above.

Let I_l be a left tangential interpolation set. Let I_r be a right tangential interpolation set. The set

$$I = \{I_l, I_r\} \tag{4.15}$$

is called a tangential interpolation set. The set of interpolation points of I, written $p(I)$, is defined by

$$p(I) := p(I_l) \cup p(I_r). \tag{4.16}$$

Let $T(s)$ be a transfer function, then we say that the tangential interpolation set I is $T(s)$-admissible if $T(s)$ has m inputs and p outputs and no point belonging to $p(I)$ is a pole of $T(s)$, i.e. no interpolation point is a pole of $T(s)$.

Let the tangential interpolation set $I = \{I_l, I_r\}$ be defined as above. If some $z_\alpha \in I_l$ is equal to some $w_\gamma \in I_r$, say $\xi_{\alpha,\gamma} = z_\alpha = w_\gamma$, then define $x_\alpha^{(f)}(s)$ to be the polynomial vector of size $1 \times p$ of degree f obtained by keeping the first f terms in the Taylor expansion of $x_\alpha(s)$ about z_α, and analogously for $y_\gamma^{(g)}(s)$:

$$x_\alpha^{(f)}(s) := \sum_{j=0}^{f-1} x_\alpha^{[j]}(z_\alpha - s)^j \quad , \quad y_\gamma^{(g)}(s) := \sum_{j=0}^{g-1} y_\gamma^{[j]}(w_\gamma - s)^j. \tag{4.17}$$

Use the same notation if z_α or w_γ is equal to ∞ :

$$x_\alpha^{(f)}(s) := \sum_{j=0}^{f-1} x_\alpha^{[j]} s^{-j} \quad , \quad y_\gamma^{(g)}(s) := \sum_{j=0}^{g-1} y_\gamma^{[j]} s^{-j}. \tag{4.18}$$

We are now able to define the tangential interpolation problem.

Definition 4.4. *Let $T(s)$ and $\hat{T}(s)$ be two* ***strictly proper*** *$p \times m$ transfer functions. $\hat{T}(s)$ interpolates $T(s)$ at I if the three following conditions are satisfied :*

1. *$\hat{T}(s)$ interpolates $T(s)$ at any couple $(z_\alpha, x_\alpha(s))$ belonging to I_l,*

2. *$\hat{T}(s)$ interpolates $T(s)$ at any couple $(w_\gamma, y_\gamma(s))$ belonging to I_r,*
3. *Finally, for every $z_\alpha = w_\gamma := \xi_{\alpha,\gamma}$, we impose that for all $f = 1, \ldots, \beta_\alpha$; $g = 1, \ldots, \delta_\gamma$, $\hat{T}(s)$ interpolates $T(s)$ at $(\xi_{\alpha,\gamma}, x_\alpha^{(f)}(s), y_\gamma^{(g)}(s))$.*

Remark 4.5. There is a notable difference of notation between the definition of the tangential interpolation problem in the SISO and in the MIMO case. In Definition 4.4, the derivative order of the the interpolation conditions at an interpolation point is implicitly given by the polynomial degree of the tangential interpolation direction vectors $x_i(s), y_j(s)$ and not explicitly given as for the SISO interpolation sets of Definition 3.5. Adding new tangential interpolation vectors $x_i(s)$ at the same interpolation point λ does not necessarily result in imposing higher order interpolation conditions at this point but more generally interpolation conditions at other directions. This clearly only makes sense in the MIMO setting.

Two other remarks are in order. In this chapter, we consider only the simple case when the interpolation set I is $T(s)$-admissible and $\hat{T}(s)$-admissible. Secondly, the tangential interpolation problem has been studied in a slightly different form in the literature, e.g., in [13], and the reader is directed there for general results about the theory of interpolation of rational matrix functions. At first sight, one could think that our definition of the two sided tangential interpolation problem is not the same as the one treated in [13]. A lemma showing the equivalence between the two formulations is proved in Section 4.6.

The problem solved in this chapter can be stated as follows:

Problem 4.6. We are given a strictly proper $p \times m$ transfer function $T(s)$ of Mc Millan degree n, and a corresponding minimal state space realization (C, A, B), such that

$$T(s) = C(sI_n - A)^{-1}B,$$

with $C \in \mathbb{C}^{p\times n}$, $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times m}$. We are also given a $T(s)$-admissible tangential interpolation set I. We want to construct a $p \times m$ reduced order transfer function $\hat{T}(s)$ of minimal Mc Millan degree k,

$$\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}, \tag{4.19}$$

with $\hat{C} \in \mathbb{C}^{p\times k}$, $\hat{A} \in \mathbb{C}^{k\times k}$, $\hat{B} \in \mathbb{C}^{k\times m}$ such that I is $\hat{T}(s)$-admissible and $\hat{T}(s)$ tangentially interpolates $T(s)$ at I.

The remainder of this chapter is organized as follows. In Section 4.2, the tangential interpolation problem is solved for two simple sets of interpolation conditions. In Section 4.3, the background necessary to solve the general problem 4.6 is introduced. In Section 4.4, the Multipoint Padé approximation is constructed and its main properties are analyzed. An extension of the preceding results for generalized state space realizations is given in Section 4.5. In Section 4.6, another formulation of the Tangential Interpolation Problem is presented. Concluding remarks are given in Section 4.7.

4.2 Preliminary results

In this section, we present the solution of Problem 4.6 for two particular interpolation sets. The general results are given in Sections 4.3 and 4.4.

One set of n distinct right interpolation conditions

The first simpler problem solved in this section is the following :

Problem 4.7. Let $T(s)$ be a $p \times m$ transfer function of Mc Millan degree n. Let $\{\lambda_1, \ldots, \lambda_k\}$ be k (where $k < n$) distinct finite points in the complex plane that are not poles of $T(s)$. Let $\{y_1, \ldots, y_k\}$ be k $m \times 1$ nonzero vectors. We want to construct a $p \times m$ transfer function $\hat{T}(s)$ of Mc Millan degree k such that for all $1 \leq i \leq k$,

$$T(\lambda_i)y_i = \hat{T}(\lambda_i)y_i. \tag{4.20}$$

Let (C, A, B) be a minimal state space realization of the $p \times m$ transfer function $T(s)$. In order to solve the problem, we construct the $n \times k$ matrix $V := \begin{bmatrix} v_1 \ldots v_k \end{bmatrix}$ that satisfies the following Sylvester equation :

$$A \begin{bmatrix} v_1 \ldots v_k \end{bmatrix} - \begin{bmatrix} v_1 \ldots v_k \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} + B \begin{bmatrix} y_1 \ \ldots \ y_k \end{bmatrix} = 0. \tag{4.21}$$

In the SISO case, it follows from Lemma 3.9 that for any set of nonzero scalars $y_1, \ldots, y_k$, V that solves (4.21) must be a full rank matrix if the pair (A, B) is controllable. In the MIMO case, the situation is more complicated. Even if the vectors $y_1, \ldots, y_k$ are linearly independent, V may not be of full rank, even if the pair (A, B) is controllable. This is related to the controllability indices of the pair (A, B) (see for instance [60], Section 6.4.6).

Let us assume that V has full column rank k. Construct $Z \in \mathbb{C}^{n \times k}$ such that

$$Z^T V = I_k.$$

Construct $\hat{C} \in \mathbb{C}^{p \times k}$, $\hat{A} \in \mathbb{C}^{k \times k}$ and $\hat{B} \in \mathbb{C}^{k \times m}$ as follows :

$$\hat{C} := CV \quad , \quad \hat{A} := Z^T AV \quad , \quad \hat{B} := Z^T B.$$

Define the transfer function

$$\hat{T}(s) := \hat{C}(sI_k - \hat{A})^{-1}\hat{B}.$$

Let us prove that if the interpolation points λ_i are not poles of $\hat{T}(s)$, then $\hat{T}(s)$ solves Problem 4.7. First note that for any $1 \leq i \leq k$ the columns of V can be computed as follows :

$$v_i = (\lambda_i I_n - A)^{-1} B y_i.$$

We will also use the following well known result.

Lemma 4.8. *Let $V \in \mathbb{C}^{n\times k}$. If the vector v belongs to the column span of the matrix V. Then, for any matrix $W \in \mathbb{C}^{n\times k}$ such that $W^T V = I_k$,*

$$v = VW^T v.$$

Proof. Because v belongs to the linear span of the columns of V, there exists a vector $\hat{v} \in \mathbb{C}^k$ such that $v = V\hat{v}$. For any W^T satisfying $W^T V = I_k$, we have $\hat{v} = W^T v$. This in turn implies that $v = VW^T v$.

□

Let us consider an arbitrary interpolation point, say λ_1. Defining W by

$$W^T := \left(Z^T(\lambda_1 I_n - A)V\right)^{-1} Z^T(\lambda_1 I_n - A) \tag{4.22}$$

clearly yields $W^T V = I_k$ and applying the preceding lemma, we obtain the following equalities :

$$\begin{aligned} T(\lambda_1)y_1 &= C(\lambda_1 I_n - A)^{-1}By_1 && (4.23)\\ &= CVW^T(\lambda_1 I_n - A)^{-1}By_1 && (4.24)\\ &= CV(\lambda_1 I_k - Z^T AV)^{-1}Z^T By_1 && (4.25)\\ &= \hat{T}(\lambda_1)y_1. && (4.26) \end{aligned}$$

Following the same reasoning with $\lambda_2, \ldots, \lambda k$ (corresponding λ_1 by another interpolation point in (4.22)), this proves that $\hat{T}(s)$ solves Problem 4.7.

Remark 4.9. 1. This reasoning is very similar to the technique used in the SISO case in [26] and [48] . These papers develop techniques to construct a SISO transfer function of Mc Millan degree k that satisfies a set of (scalar) interpolation conditions with respect to an original transfer function.
2. It should be pointed out that the transfer function $\hat{T}(s)$ of Mc Millan degree k that solves Problem 4.7 is not unique. This is due to the fact that there exist infinitely many matrices $Z \in \mathbb{C}^{n\times k}$ such that $Z^T V = I_k$, where V satisfies (4.21) and is generically unique. We will see in the sequel that, by imposing k additional left interpolation conditions, one generically determines a unique reduced order transfer function $\hat{T}(s)$ of Mc Millan degree k.
3. The condition that no interpolation point λ_i is a pole of $\hat{T}(s)$ implicitly appears in (4.22) where it is assumed that the matrix $Z^T(\lambda_1 I_n - A)V = \lambda_1 I_k - \hat{A}$ is invertible.

One unique two-sided interpolation condition

We next consider the case where the interpolation set consists of only one finite interpolation point $\alpha \in \mathbb{C}$, i.e., in terms of the parameters of Problem 4.6,

$$k_{left} = k_{right} = 1 \quad , \quad \beta_1 = \delta_1 = k \quad , \quad z_1 = w_1 = \alpha. \tag{4.27}$$

Moreover, we assume that α is not a pole of $T(s)$. Deleting the subscripts not required due to the simpler conditions to clarify the notation allows the problem to be stated as follows:

Problem 4.10. Given $T(s) = C(sI_n - A)^{-1}B$, $\alpha \in \mathbb{C}$, $x(s) := \sum_{i=0}^{k-1} x^{[i]}(\alpha - s)^i$ and $y(s) := \sum_{i=0}^{k-1} y^{[i]}(\alpha - s)^i$, construct a reduced order transfer function $\hat{T}(s)$ of Mc Millan degree k such that

$$x(s)T(s) = x(s)\hat{T}(s) + O(\alpha - s)^k, \tag{4.28}$$

$$T(s)y(s) = \hat{T}(s)y(s) + O(\alpha - s)^k, \tag{4.29}$$

and for all $f = 1, \ldots, k$; $g = 1, \ldots, k$,

$$x^{(f)}(s)(T(s) - \hat{T}(s))y^{(g)}(s) = O(\alpha - s)^{f+g}. \tag{4.30}$$

In order to solve the problem, we first rewrite equations (4.28) to (4.30) as matrix equations. Note that for any $\alpha \in \mathbb{C}$ that is not a pole of $T(s)$, we can write

$$T(s) = C(sI_n - A)^{-1}B = C\left((s-\alpha)I_n + \alpha I_n - A\right)^{-1} B \tag{4.31}$$

$$= C(\alpha I_n - A)^{-1}\left(I_n - (\alpha - s)(\alpha I_n - A)^{-1}\right)^{-1} B \tag{4.32}$$

$$= \sum_{k=0}^{\infty} C(\alpha I_n - A)^{-k-1}B(\alpha - s)^k. \tag{4.33}$$

Let us consider the left interpolation conditions corresponding to equation (4.28). By imposing the k first coefficients of the Taylor expansion of the product $x(s)(T(s) - \hat{T}(s))$ to be zero, we find the following system of equations :

$$\begin{aligned} & x^{[0]}C(\alpha I - A)^{-1}B \\ & = x^{[0]}\hat{C}(\alpha I - \hat{A})^{-1}\hat{B} \end{aligned} \tag{4.34}$$

$$\begin{aligned} & x^{[1]}C(\alpha I - A)^{-1}B + x^{[0]}C(\alpha I - A)^{-2}B \\ & = x^{[1]}\hat{C}(\alpha I - \hat{A})^{-1}\hat{B} + x^{[0]}\hat{C}(\alpha I - \hat{A})^{-2}\hat{B} \end{aligned} \tag{4.35}$$

$$\vdots$$

$$\begin{aligned} & x^{[k-1]}C(\alpha I - A)^{-1}B + \ldots + x^{[0]}C(\alpha I - A)^{-k}B \\ & = x^{[k-1]}\hat{C}(\alpha I - \hat{A})^{-1}\hat{B} + x^{[0]}\hat{C}(\alpha I - \hat{A})^{-k}\hat{B} \end{aligned} \tag{4.36}$$

Defining the matrix $X \in \mathbb{C}^{k \times kp}$ and the generalized observability matrix $\mathcal{O}_{C,A} \in \mathbb{C}^{kp \times n}$ as follows :

$$X := \begin{bmatrix} x^{[0]} & & \\ \vdots & \ddots & \\ x^{[k-1]} & \ldots & x^{[0]} \end{bmatrix} \quad ; \quad \mathcal{O}_{C,A} := \begin{bmatrix} C(\alpha I - A)^{-1} \\ \vdots \\ C(\alpha I - A)^{-k} \end{bmatrix} \tag{4.37}$$

and defining matrix $\mathcal{O}_{\hat{C},\hat{A}} \in \mathbb{C}^{kp\times k}$ analogously by replacing the matrices C and A by $\hat{C}$ and $\hat{A}$ in (4.37), we are able to state the following lemma.

Lemma 4.11. *A $p \times m$ transfer function $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ satisfies the interpolation conditions (4.28) if and only if*

$$X\mathcal{O}_{\hat{C},\hat{A}}\hat{B} = X\mathcal{O}_{C,A}B. \tag{4.38}$$

Proof. Equation (4.38) is simply a matrix form of the system (4.34)-(4.36). □

We can transpose the preceding reasoning to the right interpolation condition (4.29). Defining

$$Y = \begin{bmatrix} y^{[0]} & \dots & y^{[k-1]} \\ & \ddots & \vdots \\ & & y^{[0]} \end{bmatrix} \quad ; \quad \mathcal{C}_{A,B} = \begin{bmatrix} (\alpha I - A)^{-1}B & \dots & (\alpha I - A)^{-k}B \end{bmatrix} \tag{4.39}$$

and following the same reasoning as before, we obtain the lemma.

Lemma 4.12. *A $p \times m$ transfer function $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ satisfies the interpolation conditions (4.29) if and only if*

$$\hat{C}\mathcal{C}_{\hat{A},\hat{B}}Y = C\mathcal{C}_{A,B}Y. \tag{4.40}$$

At this point, all that we have done is to rewrite the left and right interpolation conditions into matrix equations. Next, we define the generalized Loewner matrix as

$$\mathcal{L}_{T(s)} = X\mathcal{O}_{C,A}\mathcal{C}_{A,B}Y. \tag{4.41}$$

The matrix $\mathcal{L}_{\hat{T}(s)}$ is defined as $\mathcal{L}_{T(s)}$ by replacing the matrices C, A and B by $\hat{C}, \hat{A}$ and $\hat{B}$. By rewriting the two-sided interpolation conditions corresponding to (4.30), we obtain the following lemma.

Lemma 4.13. *A $p \times m$ transfer function $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ satisfies the interpolation conditions (4.30) if and only if*

$$\mathcal{L}_{\hat{T}(s)} = \mathcal{L}_{T(s)}. \tag{4.42}$$

The following result can be proven using partial fraction expansion and Lemmas 4.11 to 4.12.

Proposition 4.14. *Every transfer function $\hat{T}(s)$ that satisfies the equations (4.28),(4.29) and* (4.30) *is such that*

$$X\mathcal{O}_{C,A}A\mathcal{C}_{A,B}Y = X\mathcal{O}_{\hat{C},\hat{A}}\hat{A}\mathcal{C}_{\hat{A},\hat{B}}Y. \tag{4.43}$$

The main result of this section can now be stated as the follows:

Proposition 4.15. *If the matrix $\mathcal{L}_{T(s)}$ is invertible, then every transfer function that satisfies the interpolation conditions (4.28)-(4.30) has a Mc Millan degree greater than or equal to k. Moreover, the transfer function of degree k that satisfies the equations (4.28)-(4.30) is unique if it exists and it can be constructed by the projection matrices V and Z that satisfy :*

$$Im(V) = Im\left(\mathcal{C}_{C,A}Y\right) \tag{4.44}$$
$$Ker(Z^T) = Ker\left(X\mathcal{O}_{A,B}\right) \tag{4.45}$$
$$Z^TV = I_k, \tag{4.46}$$

if α is not a pole of $\hat{A}$.

Sketch of the proof :

Suppose that there exists a transfer function of Mc Millan degree k such that equations (4.28) to (4.30) are satisfied. It follows that

$$X\mathcal{O}_{\hat{C},\hat{A}}\hat{B} = X\mathcal{O}_{C,A}B \tag{4.47}$$
$$\hat{C}\mathcal{C}_{\hat{C},\hat{A}}U = C\mathcal{C}_{C,A}Y \tag{4.48}$$
$$X\mathcal{O}_{\hat{C},\hat{A}}\hat{A}\mathcal{C}_{\hat{A},\hat{B}}Y = X\mathcal{O}_{C,A}A\mathcal{C}_{A,B}Y. \tag{4.49}$$

Because of the invertibility of $\mathcal{L}_{T(s)}$, the matrices $X\mathcal{O}_{\hat{C},\hat{A}} \in \mathbb{C}^{k\times k}$ and $\mathcal{C}_{\hat{A},\hat{B}}Y \in \mathbb{C}^{k\times k}$ are invertible. If we define

$$M = \left(X\mathcal{O}_{\hat{C},\hat{A}}\right)^{-1} \tag{4.50}$$
$$N = \left(\mathcal{C}_{\hat{A},\hat{B}}Y\right)^{-1} \tag{4.51}$$
$$Z^T = MX\mathcal{O}_{C,A} \tag{4.52}$$
$$V = \mathcal{C}_{A,B}YN, \tag{4.53}$$

it is straightforward to show that

$$\hat{A} = Z^TAV, \quad \hat{B} = Z^TB, \quad \hat{C} = CV, \quad Z^TV = I_k. \tag{4.54}$$

Because the image of Z, V is uniquely defined by the interpolation conditions, the uniqueness of an interpolating transfer function of Mc Millan degree k follows. Finally, there cannot exist interpolating transfer functions of Mc Millan degree smaller than k. Indeed, if $\hat{T}(s) := \hat{C}(sI_r - \hat{A})^{-1}\hat{B}$, of Mc Millan degree $r < k$, satisfies the interpolation conditions, the rank of $\mathcal{O}_{\hat{C},\hat{A}}$ and $\mathcal{C}_{\hat{A},\hat{B}}$ are both smaller or equal to r, contradicting the fact that $\mathcal{L}_{T(s)} \in \mathbb{C}^{k\times k}$ is nonsingular.

4.3 Auxiliary results

In this section, we define a generalized Loewner matrix that will allow us to construct explicitly the solution of the interpolation problem (4.6) under some mild conditions. This generalized Loewner matrix is inspired by the discussion in [5]. For the SISO case previous results based on [1], [35], and [40] may be found in [38].

In this section, we are given a strictly proper transfer function $T(s)$ and a $T(s)$-admissible interpolation set $I = \{I_l, I_r\}$ as defined in Section 4.1. The objective of this section is to find a way to characterize the set of strictly proper transfer functions $\hat{T}(s)$ such that I is $\hat{T}(s)$-admissible (the interpolation points are not poles of $\hat{T}(s)$) and $\hat{T}(s)$ tangentially interpolates $T(s)$ at I.

We define first matrices that will be used in the development. Consider the set I_l and associate with the pair $(z_\alpha, x_\alpha(s)) \in I_l$ defined in (4.9) - (4.10) the matrix $X_\alpha \in \mathbb{C}^{\beta_\alpha \times p\beta_\alpha}$

$$X_\alpha := \begin{bmatrix} x_\alpha^{[0]} & & \\ \vdots & \ddots & \\ x_\alpha^{[\beta_\alpha - 1]} & \dots & x_\alpha^{[0]} \end{bmatrix}. \tag{4.55}$$

and define the matrix $X(I_l) \in \mathbb{C}^{size(I_l) \times p\ size(I_l)}$ by

$$X(I_l) := diag\{X_\alpha\}_{\alpha=1}^{k_{left}}. \tag{4.56}$$

Analogously, with the pair $(w_\alpha, y_\alpha(s)) \in I_r$, we associate the matrix

$$Y_\alpha := \begin{bmatrix} y_\alpha^{[0]} & \dots & y_\alpha^{[\delta_\alpha - 1]} \\ & \ddots & \vdots \\ & & y_\alpha^{[0]} \end{bmatrix} \tag{4.57}$$

and define

$$Y(I_r) := diag\{Y_\alpha\}_{\alpha=1}^{k_{right}}. \tag{4.58}$$

To simplify the notation, we do not write explicitly that the matrices X_α and Y_α are related to respectively the left and the right interpolation sets I_r and I_l.

Using Definition 1.10, we easily obtain the following lemma.

Lemma 4.16.

$$J_{w,\delta,m} Y_\alpha = Y_\alpha J_{w,\delta} \quad , \quad J_{w,\beta}^T X_\alpha = X_\alpha J_{w,\beta,p}^T \tag{4.59}$$

Proof. The case $w = 0$ is nothing but the shift invariance property of block Toeplitz matrices. It then also follows for $J_{w,\delta,m} = wI + J_{0,\delta,m}$ since we add the same term on both sides of the equation (4.59).

□

We use the same notation as in Chapter 3. More precisely, by extension to the SISO case, we define the generalized observability and controllability matrices associated with respectively the left interpolation set I_l and the right interpolation set I_r (defined in Definition 4.3) as follows :

$$\mathcal{O}_{C,A}(I_l) := \mathcal{O}_{C,A}(\{(z_1,\beta_1),\ldots,(z_{k_{left}},\beta_{k_{left}})\}), \tag{4.60}$$
$$\mathcal{C}_{A,B}(I_r) := \mathcal{C}_{A,B}(\{(w_1,\delta_1),\ldots,(w_{k_{right}},\delta_{k_{right}})\}), \tag{4.61}$$

where the generalized observability and controllability matrices on the right hand side of (4.60) and (4.61) satisfy Definition 3.7.

We associate with the tangential interpolation set I the *generalized Loewner* matrix $\mathcal{L}_{T(s)}(I) \in \mathbb{C}^{size(I_l)\times size(I_r)}$ defined by

$$\mathcal{L}_{T(s)}(I) := X(I_l)\mathcal{O}_{C,A}(I_l)\mathcal{C}_{A,B}(I_r)Y(I_r), \tag{4.62}$$

where (C,A,B) is a minimal realization of $T(s)$.

It is straightforward to verify then that $\mathcal{L}_{T(s)}(I)$ does not depend on the particular state space realization of $T(s)$ but only of the value of $T(s)$ at the interpolation points and directions belonging to the set I. Next we derive a series of lemmas that are needed for our main result in Theorem 4.23.

Lemma 4.17. *If $z_\alpha \neq w_\gamma$ and both interpolation points are finite,*

$$\begin{aligned}
&\mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)\\
&= \frac{1}{w_\gamma - z_\alpha}\mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\left(\begin{bmatrix} B & 0 & \ldots & 0\end{bmatrix} - \mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)J_{0,\delta_\gamma,m}\right)\\
&\quad + \frac{1}{z_\alpha - w_\gamma}\left(\begin{bmatrix} C\\ 0\\ \vdots\\ 0\end{bmatrix} - J^T_{0,\beta_\alpha,p}\mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\right)\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma).
\end{aligned} \tag{4.63}$$

If $z_\alpha \neq w_\gamma$ and z_α is infinite, then

$$\begin{aligned}
&\mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)\\
&= \begin{bmatrix} C\\ 0\\ \vdots\\ 0\end{bmatrix}\mathcal{C}_{A,B}(z_\alpha,\delta_\alpha) - J^T_{0,\beta_\alpha}\mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)J_{0,\delta_\gamma,m}\\
&\quad -w_\gamma J^T_{0,\beta}\mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma) + J^T_{0,\beta_\alpha}\mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\begin{bmatrix} B & 0 & \ldots & 0\end{bmatrix}
\end{aligned} \tag{4.64}$$

Proof. We first prove (4.63). Recall that if $\alpha \neq \beta \in \mathbb{C}$, then

$$(\alpha I - A)^{-1}(\beta I - A)^{-1} = \frac{1}{\beta - \alpha}(\alpha I - A)^{-1} + \frac{1}{\alpha - \beta}(\beta I - A)^{-1}. \tag{4.65}$$

This permits us to write that

$$\begin{aligned}
&\mathcal{O}_{C,A}(z_\alpha, \beta_\alpha)\mathcal{C}_{A,B}(w_\gamma, \delta_\gamma) \\
&= \begin{bmatrix} C(z_\alpha I - A)^{-1} \\ \vdots \\ C(z_\alpha I - A)^{-\beta_\alpha} \end{bmatrix} \begin{bmatrix} (w_\gamma I - A)^{-1}B \ldots (w_\gamma I - A)^{-\delta_\gamma}B \end{bmatrix} && (4.66) \\
&= \frac{1}{w_\gamma - z_\alpha} \begin{bmatrix} C(z_\alpha I - A)^{-1} \\ \vdots \\ C(z_\alpha I - A)^{-\beta_\alpha} \end{bmatrix} \begin{bmatrix} (B \ldots (w_\gamma I - A)^{-\delta_\gamma + 1}B \end{bmatrix} + && (4.67) \\
&\quad \frac{1}{z_\alpha - w_\gamma} \begin{bmatrix} C \\ \vdots \\ C(z_\alpha I - A)^{-\beta_\alpha + 1} \end{bmatrix} \begin{bmatrix} (w_\gamma I - A)^{-1}B \ldots (w_\gamma I - A)^{-\delta_\gamma}B \end{bmatrix}.
\end{aligned}$$

This last equation is equal to (4.63). This concludes the proof for the finite case.

Next, consider the case $z_\alpha = \infty$. The proof is similar but uses the following equality

$$A(\lambda I - A)^{-1} = -I + \lambda(\lambda I - A)^{-1}. \tag{4.68}$$

This permits us to write that

$$\begin{aligned}
&\mathcal{O}_{C,A}(z_\alpha, \beta_\alpha)\mathcal{C}_{A,B}(w_\gamma, \delta_\gamma) \\
&= \begin{bmatrix} C \\ \vdots \\ CA^{\beta_\alpha - 1} \end{bmatrix} \begin{bmatrix} (w_\gamma I - A)^{-1}B \ldots (w_\gamma I - A)^{-\delta_\gamma}B \end{bmatrix} && (4.69) \\
&= \begin{bmatrix} C \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathcal{C}_{A,B}(w_\gamma, \delta_\gamma) \\
&\quad - J_{0,\beta_\alpha}^T \mathcal{O}_{C,A}(z_\alpha, \beta_\alpha)(A - w_\gamma I + w_\gamma I)\mathcal{C}_{A,B}(w_\gamma, \delta_\gamma) && (4.70) \\
&= \begin{bmatrix} C \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathcal{C}_{A,B}(w_\gamma, \delta_\gamma) - w_\gamma J_{0,\beta_\alpha}^T \mathcal{O}_{C,A}(z_\alpha, \beta_\alpha)\mathcal{C}_{A,B}(w_\gamma, \delta_\gamma) \\
&\quad + J_{0,\beta_\alpha}^T \mathcal{O}_{C,A}(z_\alpha, \beta_\alpha) \left(\begin{bmatrix} B \; 0 \ldots 0 \end{bmatrix} - \mathcal{C}_{A,B}(w_\gamma, \delta_\gamma) J_{0,\delta} \right). && (4.71)
\end{aligned}$$

This last term is equal to the right-hand side of equation (4.64).

□

To prove Theorem 4.23, we need the important result that the matrix $\mathcal{L}_{\hat{T}(s)}(I)$ is invariant for any matrix $\hat{T}(s)$ interpolating $T(s)$ at I (for which I is $\hat{T}(s)$-admissible). However, to show this result, we need the following lemmas.

Lemma 4.18. *Let $T(s) = C(sI - A)^{-1}B$ and $\hat{T}(s) = \hat{C}(sI - \hat{A})^{-1}\hat{B}$ be two $p \times m$ strictly proper transfer functions. Let I_l be a left interpolation set that is $T(s)$- and $\hat{T}(s)$-admissible. Then, $\hat{T}(s)$ interpolates $T(s)$ at I_l if and only if*

$$X(I_l)\mathcal{O}_{\hat{C},\hat{A}}(I_l)\hat{B} = X(I_l)\mathcal{O}_{C,A}(I_l)B. \tag{4.72}$$

Proof. Because of the diagonal structure of X, if we prove (4.72) for one diagonal block of X, say for instance X_α, we prove it for the entire equation (4.72). So we consider the block associated with X_α, and we drop α and I_l from $x_\alpha(s), X_\alpha, \mathcal{O}_{C,A}(I_l), \mathcal{O}_{\hat{C},\hat{A}}(I_l)$ to make the notation simpler. In other words, we consider the case where there is only one vector $x(s)$ of degree $\beta - 1$ associated with one interpolation point z in the left interpolation set I_l. We assume that z is finite (appropriate change must be made for the case $z = \infty$). We have to show that (4.3) is satisfied if and only if

$$X\mathcal{O}_{\hat{C},\hat{A}}\hat{B} = X\mathcal{O}_{C,A}B. \tag{4.73}$$

We can write that

$$T(s) = \sum_{i=0}^{+\infty} C(zI - A)^{-i-1}B(z - s)^i \quad , \quad \hat{T}(s) = \sum_{i=0}^{+\infty} \hat{C}(zI - \hat{A})^{-i-1}\hat{B}(z - s)^i. \tag{4.74}$$

Equation (4.3) says that $x(s)$ is a left zero of $T(s) - \hat{T}(s)$. This means that the first β Taylor coefficients of $x(s)(T(s) - \hat{T}(s))$ at $s = z$ are zero. In other words, for all $1 \leq i \leq \beta$, the following equation must be satisfied :

$$\sum_{k=0}^{i-1} x^{[k]}\hat{C}(zI - \hat{A})^{i-k}\hat{B} = \sum_{k=0}^{i-1} x^{[k]}C(zI - A)^{i-k}B, \tag{4.75}$$

and this equation turns out to be exactly the ith row of equation (4.73).

□

Analogously, for the right interpolation conditions, we have the following lemma :

Lemma 4.19. *Let $T(s) = C(sI - A)^{-1}B$ and $\hat{T}(s) = \hat{C}(sI - \hat{A})^{-1}\hat{B}$ be two $p \times m$ strictly proper transfer functions. Let I_r be a right interpolation set that is $T(s)$- and $\hat{T}(s)$-admissible. Then, $\hat{T}(s)$ interpolates $T(s)$ at I_r if and only if*

$$\hat{C}\mathcal{C}_{\hat{A},\hat{B}}Y = C\mathcal{C}_{A,B}Y. \tag{4.76}$$

The proof is similar to the proof of Lemma 4.18.

Lemma 4.20. *Let $T(s) = C(sI - A)^{-1}B$ and $\hat{T}(s) = \hat{C}(sI - \hat{A})^{-1}\hat{B}$ be two $p \times m$ strictly proper transfer functions. Let $I = \{I_l, I_r\}$ be an interpolation set that is $T(s)$- and $\hat{T}(s)$-admissible. If $\hat{T}(s)$ interpolates $T(s)$ at I then, for every pair of indices α, γ such that $z_\alpha = w_\gamma = \xi$ (where ξ is finite),*

$$X_\alpha \mathcal{O}_{\hat{C},\hat{A}}(z_\alpha, \beta_\alpha)\mathcal{C}_{\hat{A},\hat{B}}(w_\gamma, \delta_\gamma)Y_\gamma = X_\alpha \mathcal{O}_{C,A}(z_\alpha, \beta_\alpha)\mathcal{C}_{A,\hat{B}}(w_\gamma, \delta_\gamma)Y_\gamma; \quad (4.77)$$

and for every pair of indices α, γ such that $z_\alpha = w_\gamma = \xi$ (where $\xi = \infty$),

$$X_\alpha \mathcal{O}_{\hat{C},\hat{A}}(z_\alpha, \beta_\alpha)\hat{A}\mathcal{C}_{\hat{A},\hat{B}}(w_\gamma, \delta_\gamma)Y_\gamma = X_\alpha \mathcal{O}_{C,A}(z_\alpha, \beta_\alpha)A\mathcal{C}_{A,B}(w_\gamma, \delta_\gamma)Y_\gamma. \quad (4.78)$$

Proof. We consider the finite case. To simplify the notation, we drop the subscripts α, γ. Let us choose two integers f, g such that $1 \le f \le \beta$ and $1 \le g \le \delta$. Condition 3 of Definition 4.4 applied to $x(s) = x_\alpha^{(f)}(s)$ and $y(s) = y_\gamma^{(g)}(s)$ says that the $f+g$ first derivatives of $x^{(f)}(s)\left(T(s) - \hat{T}(s)\right)y^{(g)}(s)$ at $s = \xi$ are zero. The condition corresponding to the derivative of highest order is

$$\begin{aligned}
&\frac{1}{(f+g-1)!}\frac{d^{f+g-1}}{ds^{f+g-1}}\left\{x^{(f)}(s)\hat{T}(s)y^{(g)}(s)\right\}\Big|_{s=\xi} = \\
&= \sum_{k=0}^{f-1}\sum_{l=0}^{g-1} x^{[k]}C(\xi I - A)^{k+l-f-g}By^{[l]} && (4.79)\\
&= \sum_{k=0}^{f-1}\sum_{l=0}^{g-1}\left(x^{[k]}C(\xi I - A)^{k-f}\right)\left(\xi I - A)^{l-g}Bu^{[l]}\right) && (4.80)\\
&= \left(X\mathcal{O}_{C,A}\mathcal{C}_{A,B}Y\right)_{f,g}. && (4.81)
\end{aligned}$$

Thus, (4.77) is a consequence of the interpolation conditions. The proof is similar for the infinite interpolation point.

□

Equations (4.73), (4.40), (4.77) and (4.78) are just a matrix version of the interpolation conditions of Definition 4.4. We now proceed to prove that (4.73) and (4.40) imply as well that $X\mathcal{O}_{\hat{C},\hat{A}}\mathcal{C}_{\hat{A},\hat{B}}Y = X\mathcal{O}_{C,A}\mathcal{C}_{A,B}Y$ and $X\mathcal{O}_{\hat{C},\hat{A}}\hat{A}\mathcal{C}_{\hat{A},\hat{B}}Y = X\mathcal{O}_{C,A}A\mathcal{C}_{A,B}Y$, provided the two-sided interpolation condition 3 of Definition 4.4 is added for every pair $z_\alpha = w_\gamma$. This may seem surprising but it is a simple consequence of Lemma 4.20 when $z_{\alpha \ne w_\gamma}$ and follows from the two-sided condition when $z_\alpha = w_\alpha$.

Lemma 4.21. *If the strictly proper transfer function $\hat{T}(s) = \hat{C}(sI - \hat{A})^{-1}\hat{B}$ interpolates $T(s)$ at $I = \{I_l, I_r\}$ (where the interpolation set I is $T(s)$- and $\hat{T}(s)$-admissible), then*

$$X\mathcal{O}_{\hat{C},\hat{A}}\mathcal{C}_{\hat{A},\hat{B}}Y = X\mathcal{O}_{C,A}\mathcal{C}_{A,B}Y. \quad (4.82)$$

Proof. The proof will be done block by block. If $z_\alpha = w_\gamma = \xi_{\alpha,\gamma}$ and $\xi_{\alpha,\gamma}$ is finite, the proof follows from Lemma 4.20. Let us consider the case $\xi_{\alpha,\gamma}$ infinite.

$$\begin{aligned}
&X_\alpha \mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)Y_\gamma \\
&= X_\alpha \mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\left[B \dots A^{\delta_\gamma-1}\right]\begin{bmatrix} y_\gamma^{[0]} & \dots & y_\gamma^{\delta_\gamma-1} \\ & \ddots & \vdots \\ & & y_\gamma^{[0]} \end{bmatrix} && (4.83)\\
&= X_\alpha \mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)B\left[y_\gamma^{[0]} \dots y_\gamma^{\delta_\gamma-1}\right] && (4.84)\\
&\quad -X_\alpha \mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)A\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)Y_\gamma J_{0,\delta} && (4.85)\\
&= X_\alpha \mathcal{O}_{\hat{C},\hat{A}}(z_\alpha,\beta_\alpha)\mathcal{C}_{\hat{A},\hat{B}}(w_\gamma,\delta_\gamma)Y_\gamma. && (4.86)
\end{aligned}$$

Secondly, we suppose that

$$z_\alpha \neq w_\gamma. \tag{4.87}$$

We assume that z_α and w_γ are finite. The idea is to recursively use equation (4.63). We want to show that

$$X_\alpha \mathcal{O}_{\hat{C},\hat{A}}(z_\alpha,\beta_\alpha)\hat{B} = X_\alpha \mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)B \tag{4.88}$$

and

$$\hat{C}\mathcal{C}_{\hat{A},\hat{B}}(w_\gamma,\delta_\gamma)Y_\gamma = C\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)Y_\gamma \tag{4.89}$$

imply

$$X_\alpha \mathcal{O}_{\hat{C},\hat{A}}(z_\alpha,\beta_\alpha)\mathcal{C}_{\hat{A},\hat{B}}(w_\gamma,\delta_\gamma)Y_\gamma = X_\alpha \mathcal{O}_{C,A}(z_\alpha,\beta_\alpha)\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)Y_\gamma. \tag{4.90}$$

We drop again $\alpha, \gamma, (z_\alpha,\beta_\alpha), (w_\gamma,\delta_\gamma)$ to simplify the notation.

$$\begin{aligned}
&X\mathcal{O}_{C,A}\mathcal{C}_{A,B}Y \\
&= \frac{1}{w-z}X\mathcal{O}_{C,A}\left(\left[B\ 0 \dots 0\right] - \mathcal{C}_{A,B}J_{0,\delta,m}\right)Y \\
&\quad + \frac{1}{z-w}X\left(\begin{bmatrix} C \\ 0 \\ \vdots \\ 0 \end{bmatrix} - J_{0,\beta,p}^T\mathcal{O}_{C,A}\right)\mathcal{C}_{A,B}Y && (4.91)\\
&= \frac{1}{w-z}\left[X\mathcal{O}_{C,A}B\ 0 \dots 0\right]Y + \frac{1}{z-w}X\begin{bmatrix} C\mathcal{C}_{A,B}Y \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
&\quad -\frac{1}{w-z}X\mathcal{O}_{C,A}\mathcal{C}_{A,B}YJ_{0,\delta} - \frac{1}{z-w}J_{0,\beta}X\mathcal{O}_{C,A}\mathcal{C}_{A,B}Y. && (4.92)
\end{aligned}$$

From Lemmas 4.18 and 4.19 we deduce,

$$\frac{1}{w-z}\left[X\mathcal{O}_{\hat{C},\hat{A}}\hat{B}\ 0\ldots 0\right]Y=\frac{1}{w-z}\left[X\mathcal{O}_{\hat{C},\hat{A}}\hat{B}\ 0\ldots 0\right]Y, \tag{4.93}$$

$$\frac{1}{z-w}X\begin{bmatrix}C\mathcal{C}_{A,B}Y\\0\\\vdots\\0\end{bmatrix}=\frac{1}{z-w}X\begin{bmatrix}\hat{C}\mathcal{C}_{\hat{A},\hat{B}}Y\\0\\\vdots\\0\end{bmatrix}. \tag{4.94}$$

By using a recursive argument, it can be shown that

$$X\mathcal{O}_{C,A}\mathcal{C}_{A,B}YJ_{0,\delta}=X\mathcal{O}_{\hat{C},\hat{A}}\mathcal{C}_{\hat{A},\hat{B}}YJ_{0,\delta}, \tag{4.95}$$

$$J_{0,\beta}X\mathcal{O}_{C,A}\mathcal{C}_{A,B}Y=J_{0,\beta}X\mathcal{O}_{\hat{C},\hat{A}}\mathcal{C}_{\hat{A},\hat{B}}Y. \tag{4.96}$$

Finally, we have to consider the case with one infinite interpolation point, say for instance $z_\alpha=\infty$ and the other point w_γ finite. This can be treated similarly by using recursively equation (4.64).

□

Lemma 4.22. *If the strictly proper transfer function $\hat{T}(s)=\hat{C}(sI-\hat{A})^{-1}\hat{B}$ interpolates $T(s)$ at $I=\{I_l,I_r\}$ and I is $T(s)$- and $\hat{T}(s)$-admissible, then*

$$X\mathcal{O}_{\hat{C},\hat{A}}\hat{A}\mathcal{C}_{\hat{A},\hat{B}}Y=X\mathcal{O}_{C,A}A\mathcal{C}_{A,B}Y. \tag{4.97}$$

Proof. We recall that

$$A\mathcal{C}_{A,B}Y=\left[A\mathcal{C}_{A,B}(w_1,\delta_1)Y_1\ \ldots\ A\mathcal{C}_{A,B}(w_s,\delta_s)Y_s\right] \tag{4.98}$$

The proof will again be done block by block. Let us prove it for the block of $\mathcal{C}_{\hat{A},\hat{B}}(I_r)Y$ corresponding to w_γ. Two cases must be considered.

Assuming that w_γ is finite yields

$$\begin{aligned}&A\mathcal{C}_{C,A}(w_\gamma,\delta_\gamma)Y_\gamma\\&=(A-w_\gamma I+w_\gamma I)\mathcal{C}_{C,A}(w_\gamma,\delta_\gamma)Y_\gamma &(4.99)\\&=-\left[B\ \ldots\ (w_\gamma I_n-A)^{-\delta_\gamma+1}B\right]Y_\gamma+w_\gamma\mathcal{C}_{C,A}(w_\gamma,\delta_\gamma)Y_\gamma &(4.100)\\&=-B\left[y^{[0]}\ \ldots\ y^{[\delta_\gamma-1]}\right]+\mathcal{C}_{C,A}(w_\gamma,\delta_\gamma)Y_\gamma J_{w_\gamma,\delta_\gamma}. &(4.101)\end{aligned}$$

This allows us to write that

$$\begin{aligned}&X\mathcal{O}_{\hat{C},\hat{A}}\hat{A}\mathcal{C}_{\hat{A},\hat{B}}(w_\gamma,\delta_\gamma)Y_\gamma\\&=X\mathcal{O}_{\hat{C},\hat{A}}\left(-\hat{B}[y^{[0]}\ \ldots\ y^{[\delta_\gamma-1]}]+\mathcal{C}_{\hat{A},\hat{B}}(w_\gamma,\delta_\gamma)Y_\gamma J_{w_\gamma,\delta_\gamma}\right) &(4.102)\\&=X\mathcal{O}_{C,A}\left(-B[y^{[\delta_\gamma-1]}\ \ldots\ y^{[0]}]+\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)Y_\gamma J_{w_\gamma,\delta_\gamma}\right) &(4.103)\\&=X\mathcal{O}_{C,A}A\mathcal{C}_{A,B}(w_\gamma,\delta_\gamma)Y_\gamma, &(4.104)\end{aligned}$$

where the first part of (4.103) is a consequence of Lemma 4.18 and the second part of (4.103) is a consequence of Lemma 4.20.

Secondly, assume that $w_\gamma = \infty$. Two cases must be considered. If z_α is finite, then the proof is done by transposing the preceding results. If $\xi_{\alpha,\gamma} = \infty$, then this follows from Lemma 4.20.

□

Putting together the preceding results, we obtain the following theorem that gives the main result of the section.

Theorem 4.23. *Let (C_1, A_1, B_1) be a minimal state space realization of the strictly proper transfer function $T_1(s)$ and (C_2, A_2, B_2) be a minimal state space realization of the strictly proper transfer function $T_2(s)$. Let the interpolation set $I = \{I_l, I_r\}$ be $T_1(s)$ and $T_2(s)-$admissible (i.e. the interpolation points are neither poles of $T_1(s)$ nor $T_2(s)$). Then, $T_1(s)$ interpolates $T_2(s)$ at I if and only if the following equations are satisfied :*

$$C_1\mathcal{C}_{A_1,B_1}(I_r)Y(I_r) = C_2\mathcal{C}_{A_2,B_2}(I_r)Y(I_r) \tag{4.105}$$

$$X(I_l)\mathcal{O}_{C_1,A_1}(I_l)B_1 = X(I_l)\mathcal{O}_{C_2,A_2}(I_l)B_2 \tag{4.106}$$

$$X(I_l)\mathcal{O}_{C_1,A_1}(I_l)\mathcal{C}_{A_1,B_1}(I_r)Y(I_r) = X(I_l)\mathcal{O}_{C_2,A_2}(I_l)\mathcal{C}_{A_2,B_2}(I_r)Y(I_r) \tag{4.107}$$

$$X(I_l)\mathcal{O}_{C_1,A_1}(I_l)A_1\mathcal{C}_{A_1,B_1}(I_r)Y(I_r) = X(I_l)\mathcal{O}_{C_2,A_2}(I_l)A_2\mathcal{C}_{A_2,B_2}(I_r)Y(I_r) \tag{4.108}$$

Proof. Follows from the preceding results.

□

4.4 The Multipoint Padé Reduced Order Transfer Function

In this section, we give a practical way of constructing a minimal state space realization of the transfer function of minimal Mc Millan degree that interpolates $T(s)$ at the interpolation set I when the corresponding Loewner matrix $\mathcal{L}_{T(s)}(I)$ is invertible. The interpolating transfer function of minimal Mc Millan degree will be called the Multipoint Padé reduced order transfer function $\hat{T}_{MP}(s)$. A minimal state space realization $(\hat{C}_{MP}, \hat{A}_{MP}, \hat{B}_{MP})$ of $\hat{T}_{MP}(s)$ will be obtained by a projection technique. More precisely, the state space realization $(\hat{C}_{MP}, \hat{A}_{MP}, \hat{B}_{MP})$ will be constructed by projecting a minimal state space realization (C, A, B) of $T(s)$ with two projecting matrices $Z, V \in \mathbb{C}^{n\times k}$ as follows :

$$\hat{C}_{MP} = CV \quad , \quad \hat{A}_{MP} = Z^T AV \quad , \quad \hat{B}_{MP} = Z^T B \quad , \quad Z^T V = I_k.$$

It will be shown that the projecting matrices Z, V can be obtained by solving Sylvester equations.

In order to prove these facts, we first introduce two new pairs of matrices. Let us consider the left tangential interpolation set I_l defined in (4.11). For any integer α such that $1 \leq \alpha \leq k_{left}$, define the matrices $(L_\alpha^{(l)}, L_\alpha^{(r)})$ as follows :

1. If the interpolation point z_α is finite, then take

$$L_\alpha^{(l)} := I_{\beta_\alpha} \quad , \quad L_\alpha^{(r)} := J_{z_\alpha,\beta_\alpha}^T . \tag{4.109}$$

2. If the interpolation point z_α is infinite, then define

$$L_\alpha^{(l)} := -J_{0,\beta_\alpha}^T \quad , \quad L_\alpha^{(r)} := I_{\beta_\alpha} . \tag{4.110}$$

Moreover, define the matrix $\mathcal{X}_\alpha$ as follows :

$$\mathcal{X}_\alpha = \begin{bmatrix} x_\alpha^{[0]} \\ \vdots \\ x_\alpha^{[\beta_\alpha - 1]} \end{bmatrix} . \tag{4.111}$$

Finally, define the matrices $L^{(l)}(I_l)$, $L^{(r)}(I_l)$ and $\mathcal{X}(I_l)$ as follows

$$L^{(l)}(I_l) := diag\{L_\alpha^{(l)}\}_{\alpha=1}^{k_{left}} \quad , \quad L^{(r)}(I_l) := diag\{L_\alpha^{(r)}\}_{\alpha=1}^{k_{left}} , \tag{4.112}$$

$$\mathcal{X}(I_l) := \begin{bmatrix} \mathcal{X}_1 \\ \vdots \\ \mathcal{X}_{k_{left}} \end{bmatrix} . \tag{4.113}$$

Let us consider the right tangential interpolation set I_r defined in (4.14). For any integer α such that $1 \leq \alpha \leq k_{right}$, define the matrices $(R_\alpha^{(l)}, R_\alpha^{(r)})$ as follows :

1. If the interpolation point w_α is finite, then take

$$R_\alpha^{(l)} := I_{\delta_\alpha} \quad , \quad R_\alpha^{(r)} := J_{w_\alpha,\delta_\alpha} . \tag{4.114}$$

2. If the interpolation point w_α is infinite, then define

$$R_\alpha^{(l)} := -J_{0,\delta_\alpha} \quad , \quad R_\alpha^{(r)} := I_{\delta_\alpha} . \tag{4.115}$$

Moreover, define

$$\mathcal{Y}_\alpha := \begin{bmatrix} y_\alpha^{[0]} \ldots y_\alpha^{[\delta_\alpha - 1]} \end{bmatrix} . \tag{4.116}$$

Finally, define the matrices $R^{(l)}(I_r)$, $R^{(r)}(I_r)$ and $\mathcal{Y}(I_r)$ as follows

$$R^{(l)}(I_r) := diag\{R_\alpha^{(l)}\}_{\alpha=1}^{k_{right}} \quad , \quad R^{(r)}(I_r) := diag\{R_\alpha^{(r)}\}_{\alpha=1}^{k_{right}} \tag{4.117}$$

$$\mathcal{Y}(I_r) := \begin{bmatrix} \mathcal{Y}_0 \ldots \mathcal{Y}_{k_{right}} \end{bmatrix} . \tag{4.118}$$

As a consequence of these definitions we have

$$L^{(l)}L^{(r)} = L^{(r)}L^{(l)} \quad , \quad R^{(l)}R^{(r)} = R^{(r)}R^{(l)} \tag{4.119}$$

and we can now derive the following lemma that introduces the related Sylvester equations.

Lemma 4.24. *Let (C, A, B) be a state space realization of the transfer function $T(s)$. Let us consider a $T(s)-$admissible interpolation set $I = \{I_l, I_r\}$. Then,*

$$N = \mathcal{C}_{A,B}(I_r)Y(I_r) \Longleftrightarrow ANR^{(l)}(I_r) - NR^{(r)}(I_r) + B\mathcal{Y}(I_r) = 0, \tag{4.120}$$

$$M = X(I_l)\mathcal{O}_{C,A}(I_l) \Longleftrightarrow L^{(l)}(I_l)MA - L^{(r)}M + \mathcal{X}C = 0 \tag{4.121}$$

Proof. Let us prove equation (4.120) for only one interpolation condition $I_r = \{(w, y(s))\}$ at a finite point w.

$$\begin{aligned} & ANR^{(l)}(I_r) - NR^{(r)}(I_r) + B\mathcal{Y}(I_r) = 0 \\ \Longleftrightarrow\ & A\left[\, n_1 \ldots n_k \,\right] - \left[\, n_1 \ldots n_k \,\right] J_{w,k} \\ & \quad + B\left[\, y^{[0]} \ldots y^{[k-1]} \,\right] = 0. \end{aligned} \tag{4.122}$$

Let us solve this linear equation for N column by column from n_1 up to n_k. We find recursively that

$$(wI - A)n_1 = By^{[0]} \tag{4.123}$$

$$(wI - A)n_{i+1} = By^{[i]} + n_i. \tag{4.124}$$

Moreover, the matrix $wI - A$ is invertible because we always assume here that the interpolation set I is $T(s)$-admissible. This proves that $N = \mathcal{C}_{A,B}(I_r)Y(I_r)$ for one finite interpolation condition $I_r = \{(w, y(s))\}$.

Let us prove equation (4.120) for only one interpolation condition $I_r = \{(w, y(s))\}$ at an infinite point $w = \infty$.

$$\begin{aligned} & ANR^{(l)}(I_r) - NR^{(r)}(I_r) + B\mathcal{Y}(I_r) = 0 \\ \Longleftrightarrow\ & A\left[\, n_1 \ldots n_k \,\right] \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \ldots & \ldots & \ldots & 0 \end{bmatrix} \\ & \quad - \left[\, n_1 \ldots n_k \,\right] + B\left[\, y^{[0]} \ldots y^{[k-1]} \,\right] = 0. \end{aligned} \tag{4.125}$$

Again, by solving this equation column by column we find that $N = \mathcal{C}_{A,B}(I_r)Y(I_r)$ for one interpolation condition $I_r = \{(\infty, y(s))\}$. If the interpolation set I_r contains more than one pair, say k_r pairs, because of the block diagonal structure of

$R^{(l)}, R^{(r)}$ and $Y(I_r)$, and the block structure of $\mathcal{Y}(I_r)$, we can split the columns of N into k_r blocks and prove the result for each pair $(w_\gamma, y_\gamma(s)) \in I_r$ in order to prove that

$$\begin{aligned} N &= \left[N_1 \ldots N_{k_r} \right] \\ &= \left[\mathcal{C}_{A,B}\left(w_1, y_1(s)\right) Y\left(w_1, y_1(s)\right) \ldots \mathcal{C}_{A,B}\left(w_{k_r}, y_{k_r}(s)\right) Y\left(w_{k_r}, y_{k_r}(s)\right) \right] \\ &= \mathcal{C}_{A,B}(I_r) Y(I_r). \end{aligned} \tag{4.126}$$

□

The main result of this chapter can now be formalized.

Theorem 4.25. *Consider a transfer function $T(s)$ and a $T(s)$-admissible tangential interpolation set $I := \{I_l, I_r\}$ and assume that the corresponding Loewner matrix $\mathcal{L}_{T(s)}(I) \in \mathbb{C}^{k\times k}$ is invertible. Define then two invertible matrices $M, N \in \mathbb{C}^{k\times k}$ such that*

$$\mathcal{L}_{T(s)} := X(I_l)\mathcal{O}_{C,A}(I_l)\mathcal{C}_{A,B}(I_r)Y(I_r) = MN, \tag{4.127}$$

and define the "Multipoint Padé" reduced order transfer function $\hat{T}_{MP}(s)$ via its state space realization $\{\hat{A}_{MP}, \hat{B}_{MP}, \hat{C}_{MP}\}$ given by the equations :

$$\hat{C}_{MP}N = C\mathcal{C}_{A,B}Y, \tag{4.128}$$

$$M\hat{B}_{MP} = X\mathcal{O}_{C,A}B, \tag{4.129}$$

$$M\hat{A}_{MP}N = X\mathcal{O}_{C,A}A\mathcal{C}_{A,B}Y. \tag{4.130}$$

If the interpolation points are not poles of $\hat{T}_{MP}(s)$, i.e. if the interpolation set I is $\hat{T}_{MP}(s)-$admissible, then $\hat{T}_{MP}(s)$ interpolates $T(s)$ at I. Moreover, $\hat{T}_{MP}(s)$ is the unique transfer function of Mc Millan degree $size(I_l) = size(I_r)$ that interpolates $T(s)$ at I and there exists no such transfer function of lower Mc Millan degree.

Proof. First, note that it is always possible to find a couple of invertible matrices M, N that satisfy (4.127) because of the invertibility of $\mathcal{L}_{T(s)}(I)$. Second, it can be verified that $\hat{T}_{MP}(s)$ is uniquely defined and does not depend on the particular choice of matrices M, N satisfying (4.127).

The proof consists of showing that $M = X(I_l)\mathcal{O}_{\hat{C}_{MP},\hat{A}_{MP}}(I_l)$ and that $N = \mathcal{C}_{\hat{A}_{MP},\hat{B}_{MP}}(I_r)Y(I_r)$. From the preceding results, it is equivalent to show that M and N are solutions of the Sylvester equations of Lemma 4.24. First, from equations (4.127) to (4.130) and Lemma 4.24, we have

$$\begin{aligned} &\hat{A}_{MP}NR^{(l)} - NR^{(r)} + \hat{B}_{MP}\mathcal{Y} \\ &= M^{-1}X\mathcal{O}_{C,A}\left(A\mathcal{C}_{A,B}YR^{(l)} - \mathcal{C}_{A,B}YR^{(r)} + B\mathcal{Y}\right) = 0. \end{aligned} \tag{4.131}$$

This implies also from Lemma 4.24 that $N = \mathcal{C}_{\hat{A}_{MP},\hat{B}_{MP}}(I_r)Y(I_r)$. Analogously, $M = X(I_l)\mathcal{O}_{\hat{C}_{MP},\hat{A}_{MP}}(I_l)$. The proof follows now from Proposition 4.23. Indeed, (4.128) is equivalent to saying that the right tangential interpolation conditions are satisfied, (4.129) corresponds to the left tangential equations and equations (4.127) and (4.130) are equivalent to the two-sided interpolation conditions. Hence, $\hat{T}_{MP}(s)$ interpolates $T(s)$ at I.

We have still to prove that $\hat{T}_{MP}(s)$ is the unique transfer function of Mc Millan degree k that satisfies the interpolation conditions with respect to $T(s)$, and that there exist no transfer function of Mc Millan degree smaller than k that satisfies the interpolation conditions. To do this, first assume that there exists $\hat{T}(s)$ of Mc Millan degree $\hat{k} < k$ that satisfies the interpolation conditions. Let $(\hat{C}, \hat{A}, \hat{B})$ be a minimal state space realization of $\hat{T}(s)$. Clearly,

$$rank\ \mathcal{C}_{\hat{A},\hat{B}}(I_r)Y(I_r) \leq rank\ \mathcal{C}_{\hat{A},\hat{B}}(I_r) = rank\ Contr(\hat{A}, \hat{B}) = \hat{k} < k. \quad (4.132)$$

From the interpolation conditions, we must have that $\mathcal{L}_{T(s)}(I) = \mathcal{L}_{\hat{T}(s)}(I)$. This implies that

$$k = rank\ \mathcal{L}_{T(s)}(I) = rank\ \mathcal{L}_{\hat{T}(s)}(I) \leq \hat{k}. \quad (4.133)$$

This proves that it is not possible to find an interpolating transfer function of Mc Millan degree smaller than k.

If we assume that there exists another interpolating transfer function $\hat{T}(s)$ of Mc Millan degree k, it is not difficult to verify that the procedure given for constructing a minimal state space realization $(\hat{C}, \hat{A}, \hat{B})$ of $\hat{T}(s)$ will produce a state space realization that is similar to $(\hat{C}_{MP}, \hat{A}_{MP}, \hat{B}_{MP})$. This implies that $\hat{T}(s) = \hat{T}_{MP}(s)$ and concludes the proof.

□

By inverting the matrices M and N into the equations (4.127) to (4.130), if we define

$$Z^T = M^{-1}X\mathcal{O}_{C,A} \quad , \quad V = \mathcal{C}_{A,B}YN^{-1}, \quad (4.134)$$

we see that

$$Z^TV = I_k \quad , \quad CV = \hat{C}_{MP} \quad , \quad Z^TB = \hat{B}_{MP} \quad , \quad Z^TAV = \hat{A}_{MP}. \quad (4.135)$$

As in the SISO case, projecting with Sylvester equations is equivalent to solving an interpolation problem and, possibly, performing Modal Approximation. Here is a MIMO version of Theorem 3.23.

Theorem 4.26. *Let $\hat{T}(s) := \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ be a $p \times m$ transfer function of Mc Millan degree k and $T(s) := C(sI_n - A)^{-1}B$ be a $p \times m$ transfer function of Mc Millan degree n. There exist full rank matrices $Z, V \in \mathbb{C}^{n\times k}$ (with Z^TV nonsingular) satisfying Sylvester equations of the form*

$$AVR^{(l)} - VR^{(r)} + B\mathcal{Y} = 0, \quad L^{(l)}Z^TA - L^{(r)}Z^T + \mathcal{X}C = 0, \quad (4.136)$$

with regular pencils $sR^{(l)} - R^{(r)}$ *and* $sL^{(l)} - L^{(r)}$ *if and only if* $\hat{T}(s)$ *solves a tangential interpolation problem with respect to* $T(s)$ *(with possible additional modal approximation properties).*

Sketch of the Proof :
The idea is essentially the same as for Theorem 3.23. Let us consider the image of V. We have to prove that V satisfies (4.136) if and only if its image is a sum of generalized Krylov subspaces of the pair (A, B) and eigensubspaces of A.

The "if" part is a consequence of Lemma 4.24. To prove the "only if part", consider the Kronecker form of the regular pencil $sR^{(l)} - R^{(r)}$ and then decompose the image of V with respect to the diagonal blocks of this Kronecker form. By solving the Sylvester equation (4.136) block by block, one can decompose the image of V as a sum of generalized Krylov subspaces.

Modal approximation appears when the pair $(R^{(l)}, \mathcal{Y})$ is unobservable.

Remark 4.27. The main difference between Theorem 4.26 and Theorem 3.23 is that projecting with Sylvester equations is not universal anymore, i.e. there exist pairs of transfer functions $\left(T(s), \hat{T}(s)\right)$ such that the transfer function of smaller Mc Millan degree cannot be obtained via Sylvester equations from the transfer function of larger Mc Millan degree.

4.5 Extension to Generalized State Space Realizations

Krylov techniques can easily be extended to generalized state space systems, also called descriptor systems. Let us briefly sketch how to handle this case.

Assume that (C, E, A, B) is a minimal generalized state space realization of the MIMO transfer function $T(s)$. Assume also that the pencil $sE - A$ is regular. Similarly to the standard case (see equations (3.1) to (3.5)), define the following *moment* matrices:

$$\Gamma_{E,A,B}(\lambda, k) := \left((\lambda E - A)^{-1}E\right)^{k-1} (\lambda E - A)^{-1}B, \tag{4.137}$$

$$\Gamma_{E,A,B}(\infty, k) := (E^{-1}A)^{k-1}E^{-1}B, \tag{4.138}$$

$$\Delta_{C,E,A}(\lambda, k) := C(\lambda E - A)^{-1}\left(E(\lambda E - A)^{-1}\right)^{k-1}, \tag{4.139}$$

$$\Delta_{C,E,A}(\infty, k) := CE^{-1}(AE^{-1})^{k-1}. \tag{4.140}$$

From these definitions, the generalized controllability and observability matrices are defined as usual.

Definition 4.28. *Let* (C, E, A, B) *be a generalized state space realization of the SISO transfer function* $T(s)$ *(with* $sE - A$ *a regular pencil). Let* I *be a* $T(s)$*-admissible interpolation set (see Definition 3.5). We define the generalized controllability matrix* $\mathcal{C}_{E,A,B}(I)$ *by the following equations*

$$\mathcal{C}_{E,A,B}(s_i, m_i) := \left[\, \Gamma_{E,A,B}(s_i, 1) \ldots \Gamma_{E,A,B}(s_i, m_i) \,\right], \tag{4.141}$$
$$\mathcal{C}_{E,A,B}(I) := \left[\, \mathcal{C}_{E,A,B}(s_1, m_1) \ldots \mathcal{C}_{E,A,B}(s_r, m_r) \,\right], \tag{4.142}$$

and generalized observability matrix $\mathcal{O}_{C,E,A}(I)$ *to be*

$$\mathcal{O}_{C,E,A}(s_i, m_i) := \left[\, \Delta_{C,E,A}(s_i, 1)^T \ldots \Delta_{C,E,A}(s_i, m_i)^T \,\right]^T, \tag{4.143}$$
$$\mathcal{O}_{C,E,A}(I) := \left[\, \mathcal{O}_{C,E,A}(s_1, m_1)^T \ldots \mathcal{O}_{C,E,A}(s_r, m_r)^T \,\right]^T. \tag{4.144}$$

If $T(s)$ *is a MIMO transfer function and* $\{I_{left}, I_{right}\}$ *is a* $T(s)$*-admissible tangential interpolation set, define*

$$\mathcal{O}_{C,E,A}(I_l) := \mathcal{O}_{C,E,A}(\{(z_1, \beta_1), \ldots, (z_{k_{left}}, \beta_{k_{left}})\}), \tag{4.145}$$
$$\mathcal{C}_{E,A,B}(I_r) := \mathcal{C}_{E,A,B}(\{(w_1, \delta_1), \ldots, (w_{k_{right}}, \delta_{k_{right}})\}), \tag{4.146}$$
$$\mathcal{L}_{T(s)}(I) := X(I_l)\mathcal{O}_{C,E,A}(I_l)\mathcal{C}_{E,A,B}(I_r)Y(I_r) \tag{4.147}$$

With the preceding extensions, a generalization of Proposition 4.25 for generalized state space realizations is the following.

Theorem 4.29. *Consider a transfer function* $T(s)$ *with generalized state space realization* (C, E, A, B) *(with a regular pencil* $sE - A$*) and a* $T(s)$*-admissible tangential interpolation set* $I := \{I_l, I_r\}$*. Assume that the corresponding Loewner matrix* $\mathcal{L}_{T(s)}(I) \in \mathbb{C}^{k\times k}$ *is invertible. Construct two projecting matrices* $Z, V \in \mathbb{C}^{n\times k}$ *such that there exists invertible matrices* $M, N \in \mathbb{C}^{k\times k}$ *with*

$$VM = \mathcal{C}_{E,A,B}(I_r)Y(I_r), \tag{4.148}$$
$$NZ^T = X(I_l)\mathcal{O}_{C,E,A}(I_l). \tag{4.149}$$

Define the "Multipoint Padé" reduced order transfer function $\hat{T}_{MP}(s)$ *via its state space realization* $\{\hat{C}_{MP}, \hat{E}_{MP}, \hat{A}_{MP}, \hat{B}_{MP}\}$ *given by the equations :*

$$\hat{C}_{MP} = CV, \tag{4.150}$$
$$\hat{B}_{MP} = Z^T B, \tag{4.151}$$
$$\hat{A}_{MP} = Z^T AV, \tag{4.152}$$
$$\hat{E}_{MP} = Z^T EV. \tag{4.153}$$

If the interpolation points are not poles of $\hat{T}_{MP}(s)$*, i.e. if the interpolation set* I *is* $\hat{T}_{MP}(s)-$*admissible, then* $\hat{T}_{MP}(s)$ *interpolates* $T(s)$ *at* I*. Moreover,* $\hat{T}_{MP}(s)$ *is the unique transfer function of Mc Millan degree* $size(I_l) = size(I_r) = k$ *that interpolates* $T(s)$ *at* I *and there exists no such transfer function of lower Mc Millan degree.*

Remark 4.30. The proof of Theorem 4.29 is not given here because it is similar to the proof of Theorem 4.25, but with more tedious notation. A simple way

to understand the preceding results consists in assuming that the matrix E is invertible. Thus, if (C, E, A, B) realizes $T(s)$, so does the standard state space realization $(C, E^{-1}A, E^{-1}B)$. Then, it remains to observe that the transfer function $T_{MP}(s)$ given by Theorem 4.29 is the same as the transfer function $T_{MP}(s)$ given in 4.25. The main advantage of using generalized state space realizations is that the state space realization of the original transfer function $T(s)$ is often available directly in generalized form. So, it makes sense to use model reduction techniques that can be applied directly to the generalized state space realization. We refer to [40] for other results in this direction.

It should be also pointed out that the bi-orthogonality condition $Z^T V = I_k$ is implicitly replaced in Theorem 4.29 by the equation $\hat{E}_{MP} = Z^T EV$.

4.6 Second Formulation of the Tangential Interpolation Problem

The Tangential Interpolation Problem has been studied in details in [13]. Let us show here that their formulation of the tangential interpolation problem is the same as ours.

Lemma 4.31. *Let $T(s)$ and $\hat{T}(s)$ be two* ***strictly proper*** *$p \times m$ transfer functions. $\hat{T}(s)$ tangentially interpolates $T(s)$ at I with respect to Definition 4.4 if and only if the three following conditions are satisfied :*

for all finite $z_\alpha, 1 \leq \alpha \leq r$, for any $1 \leq i \leq \beta_\alpha$:

$$\frac{d^{i-1}}{ds^{i-1}} \left\{ x_\alpha(s) \left(T(s) - \hat{T}(s) \right) \right\} \Big|_{s=z_\alpha} = 0 \tag{4.154}$$

for all $z_\alpha = \infty, 1 \leq \alpha \leq r$,

$$x_\alpha(s) \left(T(s) - \hat{T}(s) \right) = O(s^{-1})^{\beta_\alpha + 1} \tag{4.155}$$

for all finite w_α, $1 \leq \alpha \leq s$, for any $1 \leq i \leq \delta_\alpha$,

$$\frac{d^{i-1}}{ds^{i-1}} \left\{ \left(T(s) - \hat{T}(s) \right) y_\alpha(s) \right\} \Big|_{s=w_\alpha} = 0 \tag{4.156}$$

for all $w_\alpha = \infty, 1 \leq \alpha \leq s$,

$$\left(T(s) - \hat{T}(s) \right) y_\alpha(s) = O(s^{-1})^{\delta_\alpha + 1} \tag{4.157}$$

for all finite $\xi_{\alpha,\gamma}$, for all $f = 1, \ldots, \beta_\alpha$; $g = 1, \ldots, \delta_\gamma$,

$$\frac{d^{f+g-1}}{ds^{f+g-1}} \left\{ x_\alpha^{(f)}(s) \left(T(s) - \hat{T}(s) \right) y_\gamma^{(g)}(s) \right\} \Big|_{s=\xi_{\alpha,\gamma}} = 0 \tag{4.158}$$

for all infinite $\xi_{\alpha,\gamma}$*, the coefficient* $e^{[f+g]}$ *of* s^{-f-g} *of the product*

$$x_\alpha^{(f)}(s)\left(T(s)-\hat{T}(s)\right)y_\gamma^{(g)}(s) \doteq \sum_{k=1}^{+\infty} e^{[k]}s^{-k} \tag{4.159}$$

is zero, where $f=1,\ldots,\beta_\alpha$*;* $g=1,\ldots,\delta_\gamma$*.*

Proof. It is easy to see that the left tangential interpolation conditions (4.154)-(4.155) and condition 1 of Definition 4.4 are equivalent. For the same reasons, the right tangential interpolation conditions (4.156)-(4.157) and conditions 2 of Definition 4.4 are equivalent. Moreover, it is not difficult to see that the two-sided tangential interpolation condition 3 of Definition 4.4 implies conditions (4.158) and (4.159). The proof will be completed by showing that conditions (4.154) through (4.159) imply the conditions 1, 2 and 3 of Definition 4.4.

Let us first consider the case with a finite left and right interpolation point $z \in \mathbb{C}$. As usual, we assume that this point is admissible for $T(s)$ and $\hat{T}(s)$, i.e. it is neither a pole of $T(s)$, nor a pole of $\hat{T}(s)$. So, we assume that we are given two polynomial vectors $x(s)$ and $y(s)$ of respective degree $\beta-1$ and $\delta-1$ such that

$$x(s)\left(T(s)-\hat{T}(s)\right) = O(s-z)^\beta, \quad x(z)\neq 0 \tag{4.160}$$

$$\left(T(s)-\hat{T}(s)\right)y(s) = O(s-z)^\delta, \quad y(z)\neq 0 \tag{4.161}$$

and for all $1\le f\le\beta$, $1\le g\le\delta$,

$$\frac{d^{f+g-1}}{ds^{f+g-1}}\left\{x^{(f)}(s)\left(T(s)-\hat{T}(s)\right)y^{(g)}(s)\right\}\Big|_{s=z} = 0 \tag{4.162}$$

We want to prove that this implies for all $1\le f\le\beta$, $1\le g\le\delta$,

$$x^{(f)}(s)\left(T(s)-\hat{T}(s)\right)y^{(g)}(s) = O(s-z)^{f+g}. \tag{4.163}$$

By using Lemma 4.20, equation (4.163) is equivalent to the equation

$$X\mathcal{O}_{C,A}\mathcal{C}_{A,B}Y = X\mathcal{O}_{\hat{C},\hat{A}}\mathcal{C}_{\hat{A},\hat{B}}Y. \tag{4.164}$$

The proof will be completed if we show that for all $1\le f\le\beta$, $1\le g\le\delta$, for all integer k such that $1\le k\le f+g-1$, the derivative

$$\frac{d^{f+g-k-1}}{ds^{f+g-k-1}}\left\{x^{(f)}(s)\left(T(s)-\hat{T}(s)\right)y^{(g)}(s)\right\}\Big|_{s=z} = 0. \tag{4.165}$$

Let us first verify (4.165) for $k=1$. First, straightforward calculation gives

$$\frac{d^{f+g-2}}{ds^{f+g-2}}\left\{x^{(f)}(s)T(s)y^{(g)}(s)\right\}\Big|_{s=z}$$
$$= \sum_{k=0}^{f-1}\sum_{l=0}^{g-1} x^{[k]}C(zI-A)^{k+l-f-g+1}By^{[l]} \tag{4.166}$$
$$= \sum_{k=0}^{f-1}\sum_{l=0}^{g-1}\left(x^{[k]}C(zI-A)^{k-f}\right)(zI-A)\left(zI-A)^{l-g}By^{[l]}\right) \tag{4.167}$$
$$= \left(X\mathcal{O}_{C,A}(zI-A)\mathcal{C}_{A,B}Y\right)_{f,g}. \tag{4.168}$$

From Lemmas 4.20 and 4.22,

$$\left(X\mathcal{O}_{C,A}(zI-A)\mathcal{C}_{A,B}Y\right) = \left(X\mathcal{O}_{\hat{C},\hat{A}}(zI-\hat{A})\mathcal{C}_{\hat{A},\hat{B}}Y\right). \tag{4.169}$$

This concludes the proof for the case $k = 1$. Now, we assume that for all $1 \le f \le \beta$ and $1 \le g \le \delta$, and for all $0 \le r \le \min(k, f+g-1)$,

$$\frac{d^{f+g-r-1}}{ds^{f+g-r-1}}\left\{x^{(f)}(s)\left(T(s)-\hat{T}(s)\right)y^{(g)}(s)\right\}\Big|_{s=z} = 0, \tag{4.170}$$

and we want to prove that (4.170) is still true for $r = \min(k+1, f+g-1)$. So, we choose $1 \le f \le \beta$ and $1 \le g \le \delta$ such that $f+g-1 \ge k+1$. We obtain the following equations

$$\frac{d^{f+g-k-2}}{ds^{f+g-k-2}}\left\{x^{(f)}(s)\left(T(s)-\hat{T}(s)\right)y^{(g)}(s)\right\}\Big|_{s=z}$$
$$= \frac{d^{f-1+g-k-1}}{ds^{f-1+g-k-1}}\left\{x^{(f-1)}(s)\left(T(s)-\hat{T}(s)\right)y^{(g)}(s)\right\}\Big|_{s=z} \tag{4.171}$$
$$+\frac{d^{f-1+g-k-1}}{ds^{f-1+g-k-1}}\left\{(z-s)^{f-1}x^{[f-1]}\left(T(s)-\hat{T}(s)\right)y^{(g)}(s)\right\}\Big|_{s=z}. \tag{4.172}$$

By the recursive argument,

$$\frac{d^{f-1+g-k-1}}{ds^{f-1+g-k-1}}\left\{x^{(f-1)}(s)\left(T(s)-\hat{T}(s)\right)y^{(g)}(s)\right\}\Big|_{s=z} = 0. \tag{4.173}$$

Moreover, we know from (4.156) that

$$\left(T(s)-\hat{T}(s)\right)y^{(g)}(s) = O(z-s)^g. \tag{4.174}$$

This implies that

$$\frac{d^{f+g-k-2}}{ds^{f+g-k-2}}\left\{x^{(f)}(s)\left(T(s)-\hat{T}(s)\right)y^{(g)}(s)\right\}\Big|_{s=z} = 0. \tag{4.175}$$

The case at infinity can be treated in a similar way.

□

4.7 Concluding remarks

An interesting property of Multipoint Padé approximation that is worth mentioning is the following. Assume that one has constructed a reduced order transfer function $\hat{T}_1(s)$ that interpolates the original transfer function $T(s)$ at the interpolation set I_1 with the projecting matrices Z_1 and V_1. If one wants to add new interpolation conditions, say I_2, all that we have to do is to compute the generalized Krylov subspaces corresponding to the new interpolation set I_2 and to construct new projecting matrices Z_2, V_2 that contain respectively the column span of Z_1 and V_1 and the new respectively left and right generalized Krylov subspaces.

Another important result that can easily be derived is that we only need the projecting matrices Z, V to contain some subspaces, but they can contain other subspaces as well! For instance, Theorem 4.25 can be generalized as follows :

Theorem 4.32. *Consider a transfer function $T(s) := C(sI - A)^{-1}B$ and a $T(s)-$admissible tangential interpolation set $I := \{I_l, I_r\}$. Let us assume that the projecting matrices Z, V (such that $Z^T V = I_k$) are such that*

$$Colsp\left(\mathcal{C}_{A,B}(I_r)Y(I_r)\right) \subseteq Colsp(V),$$
$$Colsp\left(\mathcal{O}^T_{C,A}(I_l)X^T(I_l)\right) \subseteq Colsp(Z^T).$$

Then, if the interpolation point of I are not poles of $\hat{T}(s) := CV(sI_k - Z^T AV)^{-1}Z^T B$, the transfer function $\hat{T}(s)$ interpolates $T(s)$ at I.

Again, it is possible to rewrite Theorem 4.32 for generalized state space realizations by replacing the condition $Z^T V = I_k$ by $Z^T EV = \hat{E}$ and by replacing the generalized observability matrices $\mathcal{O}_{C,A}$ and $\mathcal{C}_{A,B}$ with respectively $\mathcal{O}_{C,E,A}$ and $\mathcal{C}_{E,A,B}$.

In Theorem 4.25 and 4.29, it is always assumed that there is the same number of left and right tangential interpolation conditions. Indeed, the Loewner matrix $\mathcal{L}_{T(s)}(I)$ is assumed to be square. This ensures to define uniquely the interpolating transfer function of minimal Mc Millan degree. Let us discuss how to handle the nonsymmetric case. Assume that there are k left tangential interpolation conditions and $k - r$ right interpolation conditions. From these conditions, construct $Z \in \mathbb{C}^{n\times k}$ and $V_1 \in \mathbb{C}^{n\times(k-r)}$ such that their images are equal to the corresponding generalized Krylov subspaces. Then, choose $V_2 \in \mathbb{C}^{n\times r}$ such that

$$Z^T \begin{bmatrix} V_1 & V_2 \end{bmatrix} = I_k,$$

and project using (Z, V). From 4.32, the interpolation conditions are satisfied. Because there is a degree of freedom in choosing V_2, the interpolating transfer function of minimal Mc Millan degree is not unique.

Finally, we have shown that the projecting matrices Z, V yielding a state space realization of $\hat{T}_{MP}(s)$ are solutions of Sylvester equations. Moreover, we

have shown that generically, constructing a reduced order transfer function with projecting matrices that are solutions of a Sylvester equation with respect to a state space realization of the original transfer function is equivalent to solving a particular tangential interpolation problem (plus possible Modal Approximation properties). In Chapter 5, another interpretation of the Sylvester equations appearing naturally in the Multipoint Padé approximation framework will be given.

If the transfer function is not strictly proper, its polynomial part is often copied in the reduced order transfer function. Assume that $T(s) := C(sE - A)^{-1}B + D$ has a polynomial part. This implies that E is singular with a kernel V_∞ of a particular dimension k_∞. By imposing $Im(V_\infty) \subseteq Im(V)$ (where V is the right projecting matrix), the reduced order system will have an $\hat{E}$ matrix with a kernel of the same dimension and hence will retain these algebraic equations [35].

4.8 Notes and References

This chapter is based on [39]. A simpler version is available in [36]. This is the result of a collaborative work with Kyle Gallivan and Paul Van Dooren.

As already pointed out, we refer to the book [13] for an extensive study of the general problem of interpolation of rational matrix functions, in a more abstract mathematical setting. Related work focusing on the tangential interpolation problem is available in [14, 7].

5

On the Embedding of State Space Realizations

In this chapter, the generality of the particular model reduction method, known as the projection of state space realization, is investigated. Given two transfer functions, one wants to find the necessary and sufficient conditions for the embedding of a state space realization of the transfer function of smaller Mc Millan degree into a state space realization of the transfer function of larger Mc Millan degree. Two approaches are considered, both in the MIMO case. First, when the difference of Mc Millan degree between the transfer functions is equal to one and there is no common pole, necessary and sufficient conditions are provided. Then, the generic case is considered using a pencil approach related to Krylov subspaces. A conjecture concerning the generality of Krylov techniques is also given. Finally, it is shown that the condition of embedding, related to the eigenstructure of a pencil, is related to the tangential interpolation problem discussed in the preceding chapter.

5.1 Introduction

We consider two $p \times m$ strictly proper transfer functions

$$T(s) = C(sI_{n+k} - A)^{-1}B, \quad \hat{T}(s) = \hat{C}(sI_n - \hat{A})^{-1}\hat{B}, \tag{5.1}$$

of respective Mc Millan degrees $n + k$ and k. We want to characterize the set of projecting matrices $Z, V \in \mathbb{C}^{(n+k)\times k}$ such that

$$\hat{C} = CV, \quad \hat{A} = Z^T AV, \quad \hat{B} = Z^T B, \quad Z^T V = I_n. \tag{5.2}$$

Somewhat surprisingly, if the model reduction problem for linear systems has been widely studied for more than fifty years, the exact potential of the projection technique for model reduction has received prior to this date little attention only. In this chapter, we state this problem in the following terms (the notion of embedding for transfer functions is introduced in Definition 2.1).

Problem 5.1. Given two $p \times m$ strictly proper transfer functions : $T(s)$ of Mc Millan degree $n+k$ and $\hat{T}(s)$ of Mc Millan degree n, what are the necessary and sufficient conditions in order to ensure that $\hat{T}(s)$ is embedded in $T(s)$.

We will need some additional notation. Any transfer function $T(s)$ can be decomposed into its strictly proper part $\lfloor T(s) \rfloor$ and its polynomial part $\lceil T(s) \rceil$:

$$T(s) = \lfloor T(s) \rfloor + \lceil T(s) \rceil.$$

The formal degree of a polynomial matrix $F(s)$ is defined as the highest degree in s of $F(s)$. Let us recall that the normal rank of a $p \times m$ rational function $R(s)$ is the maximal rank of $R(s)$ for $s \in \mathbb{C}$.

In the SISO case, it has been proved in Chapter 3, Theorem 3.22 that for any pair of strictly proper transfer functions, the transfer function of smaller Mc Millan degree is embedded in the transfer function of larger Mc Millan degree. This result has been proved with the help of Krylov techniques related to interpolation theory. As shown in the sequel, it appears that the question is significantly more complex in the MIMO case.

The outline of this chapter is as follows. A necessary rank condition for the embedding is given in section 5.2. Necessary and sufficient conditions for the embedding when $\hat{T}(s)$ and $T(s)$ do not have common poles are derived in section 5.3. Another approach, based on the Kronecker form of some pencils and an open problem are formulated in section 5.4. A connection between Krylov Subspaces and the Pencil Approach of Halevi is given in Subsection 5.5. Concluding remarks are given in section 5.6.

Because we consider in this chapter mainly transfer functions $T(s)$ and $\hat{T}(s)$ that differ only by one in terms of Mc Millan degree, we have chosen for ease of notation to write that the Mc Millan degree of $\hat{T}(s)$ is n and the Mc Millan degree of $T(s)$ is $n+k$, with most of the time $k=1$. This is in contrast with the notation adopted in the other chapters where the Mc Millan degrees of $T(s)$ and $\hat{T}(s)$ are respectively equal to n and k.

5.2 A necessary rank condition

Let us first investigate a straightforward consequence of the embedding of one transfer function into another.

Proposition 5.2. *Let $T(s)$ and $\hat{T}(s)$ be $p \times m$ strictly proper transfer functions of Mc Millan degree $n+k$ and n respectively. If $\hat{T}(s)$ is embedded in $T(s)$, the error $E(s) := T(s) - \hat{T}(s)$ can be factorized as follows :*

$$E(s) = L(s)C(s)R(s), \quad \textit{where} \tag{5.3}$$

1. *$L(s)$ is a $p \times k$ proper transfer function of Mc Millan degree less than or equal to n;*

2. *$C(s)$ is a $k \times k$ transfer function;*
3. *$R(s)$ is a $k \times m$ proper transfer function of Mc Millan degree less than or equal to n.*

Proof. The proof is based on a simple Schur complement argument. Since $\hat{T}(s)$ is embedded in $T(s)$, there exists a minimal state space realization $(\hat{C}, \hat{A}, \hat{B})$ of $\hat{T}(s)$ and a minimal state space realization (C, A, B) of $T(s)$ such that

$$C = \begin{bmatrix} \hat{C} \; C_2 \end{bmatrix} \quad , \quad A = \begin{bmatrix} \hat{A} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad , \quad B = \begin{bmatrix} \hat{B} \\ B_2 \end{bmatrix}. \tag{5.4}$$

Setting $\Delta(s) := (sI_k - A_{22}) - A_{21}(sI_n - \hat{A})^{-1}A_{12}$,, $A := (sI_n - \hat{A})$, $B := -A_{21}$, $C := -A_{12}$ and $D := sI_k - A_{22}$, one deduces from (1.1) by straightforward algebraic manipulations that the difference between $T(s)$ and $\hat{T}(s)$ can be expressed as

$$\begin{aligned} &T(s) - \hat{T}(s) \\ &= \left(\hat{C}(sI_n - \hat{A})^{-1}A_{12} + C_2\right) \Delta(s)^{-1} \left(A_{21}(sI_n - \hat{A})^{-1}\hat{B} + B_2\right). \end{aligned} \tag{5.5}$$

This completes the proof of Proposition 5.2.

□

As a first consequence of this result, it appears that the normal rank of the error $E(s)$ must be smaller than or equal to k. This rank condition on the error transfer function is not sufficient to guarantee the existence of an embedding solution, as shown later.

5.3 Necessary and sufficient conditions

Let us first investigate the state space realization properties of transfer functions related by unimodular transformations. We begin with a simple, but rather surprising result.

Lemma 5.3. *Let $T(s)$ be a $p \times m$ strictly proper transfer function. Let $U(s)$, $V(s)$ be two unimodular polynomial matrices of appropriate dimension. Then,*

$$\left\lfloor U^{-1}(s) \left\lfloor U(s)T(s)V(s) \right\rfloor V^{-1}(s) \right\rfloor = T(s), \tag{5.6}$$

Proof. Let us define

$$T_1(s) := \lceil U(s)T(s)V(s) \rceil, \quad T_2(s) := \lfloor U(s)T(s)V(s) \rfloor.$$

We then use the fact that $T_1(s)$ is a polynomial matrix, that the inverse of a unimodular matrix is a polynomial matrix and that the product of two polynomials is a polynomial and $T(s)$ is strictly proper to obtain

$$\begin{aligned}
T(s) &= U^{-1}(s)U(s)T(s)V(s)V^{-1}(s) \\
&= \lfloor U^{-1}(s)U(s)T(s)V(s)V^{-1}(s) \rfloor \\
&= \left\lfloor U^{-1}(s)\left(T_1(s)+T_2(s)\right)V^{-1}(s) \right\rfloor \\
&= \left\lfloor U^{-1}(s)T_2(s)V^{-1}(s) \right\rfloor \\
&= \left\lfloor U^{-1}(s)\left\lfloor U(s)T(s)V(s) \right\rfloor V^{-1}(s) \right\rfloor .
\end{aligned}$$

□

Lemma 5.4. *Let $U(s) := \sum_{i=0}^{k_u} U_i s^i$, $V(s) := \sum_{j=0}^{k_v} V_j s^j$ be two unimodular polynomial matrices of appropriate dimension. Let (C, A, B) be a minimal state space realization of the strictly proper transfer function $T(s)$. Then the Mc Millan degree of $\lfloor U(s)T(s)V(s) \rfloor$ is equal to the Mc Millan degree of $T(s)$ and*

$$\left(\sum_{i=0}^{k_u} U_i C A^i, A, \sum_{j=0}^{k_v} A^j B V_j \right) \sim \lfloor U(s)T(s)V(s) \rfloor . \tag{5.7}$$

Proof. The proof is straightforward from the equality

$$U(s)C(sI_n - A)^{-1}B = \sum_i U_i C \left(A + (sI_n - A)\right)^i (sI_n - A)^{-1}B$$

which yields directly

$$\left\lfloor U(s)C(sI_n - A)^{-1}B \right\rfloor = \left(\sum_i U_i C A^i \right) (sI_n - A)^{-1}B.$$

The full proof is then achieved by applying the same argument on the right. This result implies that the Mc Millan degree of $\lfloor U(s)T(s)V(s) \rfloor$ is less than or equal to the Mc Millan degree of $T(s)$. From the relation (5.6), one deduces by duality that the Mc Millan degree of $T(s)$ is smaller than or equal to the Mc Millan degree of $\lfloor U(s)T(s)V(s) \rfloor$. Therefore, the transfer functions $T(s)$ and $\lfloor U(s)T(s)V(s) \rfloor$ do have the same Mc Millan degree.

□

In order to find the necessary and sufficient conditions for the embedding problem of state space realizations in the MIMO case, we focus on the case where $T(s) := C(sI_{n+1} - A)^{-1}B$ has Mc Millan degree $n+1$, $\hat{T}(s) := \hat{C}(sI_n - \hat{A})^{-1}\hat{B}$ has Mc Millan degree n and $T(s)$ and $\hat{T}(s)$ do not have common poles. Without this assumption, the problem appears to be considerably more complex. Define

$$d(s) := \det(sI_{n+1} - A), \quad \hat{d}(s) := \det(sI_n - \hat{A}).$$

Conformably with Proposition 5.2, the normal rank of the difference $E(s) := T(s) - \hat{T}(s)$ will be assumed to be equal to one. Because $T(s)$ and $\hat{T}(s)$ are

devoid of common poles, the Mc Millan degree of $E(s)$ is clearly equal to $2n+1$. As a result, there exist two unimodular matrices $U_E(s)$ and $V_E(s)$ such that

$$E(s) = U_E(s) \begin{bmatrix} \frac{n_E(s)}{d(s)\hat{d}(s)} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} V_E(s).$$

As $E(s)$ is strictly proper ($E(s)$ is the difference between two strictly proper transfer functions, implying that it is also a strictly proper transfer function) and since the polynomial matrices $U_E(s)$ and $V_E(s)$ are unimodular, the rational function $\frac{n_E(s)}{\hat{d}(s)d(s)}$ is strictly proper. Indeed, if the scalar transfer function had a polynomial part, then $E(s)$ would also have a polynomial part, contradicting the fact that $E(s)$ is strictly proper. Moreover, $E(s)$ can be rewritten by partial fraction expansion as follows :

$$\frac{n_E(s)}{\hat{d}(s)d(s)} = \frac{n_1(s)}{d(s)} - \frac{n_2(s)}{\hat{d}(s)},$$

with $n_1(s)/d(s)$ and $n_2(s)/\hat{d}(s)$ both strictly proper. As a result, one has the expressions:

$$T(s) = U_E(s) \begin{bmatrix} \frac{n_1(s)}{d(s)} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} V_E(s), \quad \hat{T}(s) = U_E(s) \begin{bmatrix} \frac{n_2(s)}{\hat{d}(s)} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} V_E(s).$$

and, consequently, the following lemma.

Lemma 5.5. *If $T_1(s)$ and $T_2(s)$ are two strictly proper transfer functions with no common poles such that the normal rank of $E(s) := T_1(s) - T_2(s)$ is equal to 1, then the normal rank of $T_1(s)$ and the normal rank of $T_2(s)$ are both equal to one. Moreover, the same unimodular matrices can put $T_1(s), T_2(s)$ and $E(s)$ into their canonical Smith-Mc Millan form.*

Under the stated assumptions, the following result can then be proved.

Theorem 5.6. *Let $T(s)$ and $\hat{T}(s)$ be $p \times m$ strictly proper transfer functions, devoid of common poles and of Mc Millan degree $n+1$ and n respectively. Then, $\hat{T}(s)$ is embedded in $T(s)$ if and only if there exist two polynomial vectors of degree less or equal to n, denoted by $l(s)$ and $r(s)$ such that*

$$E(s) = \frac{l(s)r^T(s)}{d(s)\hat{d}(s)}. \tag{5.8}$$

Proof. In view of section 5.2, these conditions are necessary. Let us prove that they are also sufficient. From Lemma 5.5, there exist two unimodular matrices $U(s)$ and $V(s)$ that put the three rational matrices $T(s), \hat{T}(s)$ and $E(s)$ into their canonical Smith-Mc Millan form so that one can write

$$U^{-1}(s)T(s)V^{-1}(s) = \begin{bmatrix} \frac{n_1(s)}{d(s)} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}, \quad U^{-1}(s)\hat{T}(s)V^{-1}(s) = \begin{bmatrix} \frac{n_2(s)}{\hat{d}(s)} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix},$$

$$U^{-1}(s)E(s)V^{-1}(s) = \begin{bmatrix} \frac{n_E(s)}{\hat{d}(s)d(s)} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}.$$

If $u(s)$ stands for the first column of $U(s)$ and $v(s)^T$ for the the first row of $V(s)$,

$$u(s) := \sum_{i=0}^{k_u} u_i s^i, \quad v^T(s) := \sum_{j=0}^{k_v} v_j^T s^j,$$

then there clearly exist two scalar polynomials of formal degree less than or equal to n, $\phi(s)$ and $\psi(s)$, such that $l(s) = u(s)\phi(s)$ and $r^T(s) = v^T(s)\psi(s)$. If $(\hat{c}, \hat{A}, \hat{b})$ is a minimal state space realization of $n_2(s)/\hat{d}(s)$, it appears that vector a_r and scalar γ must also exist such that

$$\hat{c}(sI_n - \hat{A})^{-1}a_r + \gamma = \frac{\phi(s)}{\hat{d}(s)}. \tag{5.9}$$

Indeed, from the canonical observability form of the pair $(\hat{c}, \hat{A})$, one can verify that it is always possible to construct a SISO proper transfer function such that its poles are a subset of the zeroes of $\hat{d}(s)$ with a state space realization of the form (5.9) by an appropriate choice of a_r and γ. By duality, there also exist a_l and β such that

$$a_l(sI_n - \hat{A})^{-1}\hat{b} + \beta = \frac{\psi(s)}{\hat{d}(s)}.$$

Setting

$$\hat{C} := \sum_{i=0}^{k_u} u_i \hat{c}\hat{A}^i \quad , \quad \hat{B} := \sum_{j=0}^{k_v} \hat{A}^j \hat{b} v_j^T, \tag{5.10}$$

one then deduces from Lemma 5.4 that $(\hat{C}, \hat{A}, \hat{B})$ is a minimal state space realization of $\hat{T}(s)$ and, moreover, that one has

$$\frac{l(s)}{\hat{d}(s)} = \hat{C}(sI_n - \hat{A})^{-1}a_r + u_0\gamma \quad , \quad \frac{r(s)}{\hat{d}(s)} = a_l(sI_n - \hat{A})^{-1}\hat{B} + \beta v_0^T.$$

Defining $s - \alpha$ as the polynomial quotient of the division of $d(s)$ by $\hat{d}(s)$ and $\tilde{T}(s)$ as the transfer function of state space realization

$$\left(\left[\hat{C} \; u_0\gamma \right], \begin{bmatrix} \hat{A} & a_r \\ a_l^T & \alpha \end{bmatrix}, \begin{bmatrix} \hat{B} \\ \beta v_0^T \end{bmatrix} \right),$$

one can define the expression

$$\tilde{T}(s) := \frac{\tilde{N}(s)}{\tilde{d}(s)},$$

where $\tilde{N}(s)$ is a $p \times m$ matrix of formal degree less than or equal to n and where

$$\tilde{d}(s) := \det \begin{bmatrix} sI_n - \hat{A} & -a_r \\ -a_l^T & s - \alpha \end{bmatrix}.$$

Defining

$$N(s) := u(s)n_1(s)v^T(s), \quad \hat{N}(s) := u(s)n_2(s)v^T(s),$$

one obtains

$$T(s) = \frac{N(s)}{d(s)}, \quad \hat{T}(s) = \frac{\hat{N}(s)}{\hat{d}(s)}.$$

Applying the same argument as in the proof of (5.5), one then obtains the polynomial expression

$$\tilde{N}(s)\hat{d}(s) - \hat{N}(s)\tilde{d}(s) = l(s)r^T(s). \tag{5.11}$$

Since $T(s) = N(s)/d(s)$, it follows that (5.11) is satisfied as well for $N(s)$ and $d(s)$ substituted for $\tilde{N}(s)$ and $\tilde{d}(s)$ respectively.

Actually, $T(s)$ and $\tilde{T}(s)$ must be the same transfer function. Indeed, let us assume that $\hat{N}(s), \hat{d}(s), l(s), r(s)$ and α are given and characterize the set of polynomials $d^*(s)$ of degree equal to $n+1$ and of polynomial matrices $N^*(s)$ of formal degree less than or equal to n such that equation (5.11) is satisfied and with the additional condition that $s - \alpha$ is the polynomial quotient of the division of $d^*(s)$ by $\hat{d}(s)$. As $\hat{T}(s)$ is a rank one transfer function, $\hat{N}(s)$ and $\hat{d}(s)$ cannot have common zeros for otherwise the Mc Millan degree of $\hat{T}(s)$ would be less than n. At each of the n zeros s_i of $\hat{d}(s)$, it thus appears from (5.11) that one has the relation $-\hat{N}(s_i)d^*(s_i) = l(s_i)r^T(s_i)$ and this clearly imposes n interpolation constraints on the polynomial $d^*(s)$ (if $\hat{d}(s)$ admits multiple zeros, then successive derivatives of $d^*(s)$ at each multiple zero of $\hat{d}(s)$ can be found recursively). Finally, the condition relative to the division of $d^*(s)$ by $\hat{d}(s)$ is

seen to determine $d^*(s)$ in a unique manner. Once $d^*(s)$ is fixed, one deduces $N^*(s)$ from (5.11):

$$N^*(s) = \frac{\hat{N}(s)d^*(s) + l(s)r^T(s)}{\hat{d}(s)}.$$

This implies the equality $\tilde{T}(s) = T(s)$, which completes the proof of the theorem.

□

We conclude this section by giving an example showing that the rank condition is not sufficient to ensure the embedding.

Let $T(s) := C(sI_{n+1} - A)^{-1}B$ and $\hat{T}(s) := \hat{C}(sI_n - \hat{A})^{-1}\hat{B}$ be defined as follows

$$T(s) := \begin{bmatrix} \frac{1}{(s-1)^3} - \frac{2}{(s-1)^2} + \frac{3}{s-1} \\ \frac{1}{s-1} \end{bmatrix} \quad , \quad \hat{T}(s) := \begin{bmatrix} \frac{1}{s^2} + \frac{3}{s} \\ \frac{1}{s} \end{bmatrix}. \tag{5.12}$$

Clearly $T(s)$ and $\hat{T}(s)$ are two strictly proper transfer functions of Mc Millan degree respectively equal to 3 and 2, having no common pole such that the error,

$$E(s) := T(s) - \hat{T}(s) := \frac{l(s)r^T(s)}{d(s)\hat{d}(s)},$$

has rank one, where

$$d(s) := \det(sI_3 - A) = (s-1)^3, \quad \hat{d}(s) := \det(sI_2 - \hat{A}) = s^2.$$

It appears that $\hat{T}(s)$ is not embeddable into $T(s)$ for there exists no factorization of $l(s)r^T(s)$ as a product of two polynomial matrices of formal degree less than 2. To see this, consider minimal state space realizations (C, A, B) and $(\hat{C}, \hat{A}, \hat{B})$ of $T(s)$ and $\hat{T}(s)$ respectively. The error

$$E(s) := T(s) - \hat{T}(s) := \frac{l(s)r^T(s)}{d(s)\hat{d}(s)} = \begin{bmatrix} \frac{1}{s^2(s-1)^3} \\ \frac{1}{s(s-1)} \end{bmatrix},$$

has rank one but $E(s)d(s)\hat{d}(s)$ cannot be factorized into the product of two polynomial matrices of formal degree less than 2. Indeed, one has :

$$E(s)d(s)\hat{d}(s) = l(s)r^T(s) = \begin{bmatrix} 1 \\ (s-1)^2 s \end{bmatrix}$$

so that the only possible factorization is

$$l(s) = \frac{1}{\alpha}\begin{bmatrix} 1 \\ (s-1)^2 s \end{bmatrix} \quad , \quad r(s) = \alpha,$$

where α is any nonzero complex number. Therefore, $\hat{T}(s)$ cannot be obtained by truncation from $T(s)$.

5.4 The Pencil Approach

The objective of this section is to find *generic* conditions by studying the Kronecker form of particular matrix pencils. This section is based on ideas developed in [85, 53].

From now on, assume that $T(s) = C(sI_{n+k} - A)^{-1}B$ is a $p \times m$ transfer function of Mc Millan degree $n + k$ and that $\hat{T}(s) = \hat{C}(sI_n - \hat{A})^{-1}\hat{B}$ is a $p \times m$ transfer function of Mc Millan degree n, i.e. the state space realizations (C, A, B) and $(\hat{C}, \hat{A}, \hat{B})$ are both minimal. Let us also assume that B and C are full rank matrices. This is equivalent to impose that there is no zero block in the Kronecker form of $T(s)$.

The Approach of Halevi

Let us first briefly review some general results. If $\hat{T}(s)$ is embedded in $T(s)$, conditions (5.2) must be satisfied, which yield a quadratic system of $2n^2 + n(m + p)$ scalar equations in the $2n(n + k)$ entries of the unknown projections matrices $Z, V \in \mathbb{C}^{(n+k)\times n}$. Therefore, the number of degrees of freedom is larger, equal or smaller than the number of problem constraints depending on whether $2k > m + p$, $2k = m + p$ or $2k < m + p$.

Define $B^+ \in \mathbb{C}^{m\times(n+k)}$ to be a left inverse of B and $B_\perp \in \mathbb{C}^{(n+k-m)\times(n+k)}$ to be the left kernel of B. Analogously, define $C^+ \in \mathbb{C}^{(n+k)\times p}$ to be a right inverse of C and $C_\perp \in \mathbb{C}^{(n+k)\times(n+k-p)}$ to be the right kernel of C,

$$\begin{bmatrix} B^+ \\ B_\perp \end{bmatrix} B = \begin{bmatrix} I_m \\ 0_{(n+k-m)\times m} \end{bmatrix}, \quad C \begin{bmatrix} C^+ & C_\perp \end{bmatrix} = \begin{bmatrix} I_p & 0_{p\times(n+k-p)} \end{bmatrix}. \tag{5.13}$$

If we rewrite

$$Z^T = \hat{B}B^+ + XB_\perp, \tag{5.14}$$

$$V = C^+\hat{C} + C_\perp Y, \tag{5.15}$$

with X and Y arbitrary matrices of dimensions $n\times(n+k-m)$ and $(n+k-p)\times n$ respectively, then conditions on $\hat{C}$ and $\hat{B}$ of (5.2) are automatically satisfied. In order to satisfy the two other conditions, we must have

$$Z^T AV = \hat{A} \tag{5.16}$$

$$\Longleftrightarrow (\hat{B}B^+ + XB_\perp)A(C^+\hat{C} + C_\perp Y) = \hat{A} \tag{5.17}$$

$$\Longleftrightarrow \begin{bmatrix} X & I_n \end{bmatrix} \begin{bmatrix} B_\perp AC_\perp & B_\perp AC^+\hat{C} \\ \hat{B}B^+AC_\perp & \hat{B}B^+AC^+\hat{C} - \hat{A} \end{bmatrix} \begin{bmatrix} Y \\ I_n \end{bmatrix} = 0. \tag{5.18}$$

By rewriting the biorthogonality constraint with the same technique, we obtain the following equivalent notation :

$$Z^T V = I_n \tag{5.19}$$

$$\iff \begin{bmatrix} X & I_n \end{bmatrix} \begin{bmatrix} B_\perp C_\perp & B_\perp C^+ \hat{C} \\ \hat{B} B^+ C_\perp & \hat{B} B^+ C^+ \hat{C} - I_n \end{bmatrix} \begin{bmatrix} Y \\ I_n \end{bmatrix} = 0. \tag{5.20}$$

Equations (5.18) and (5.20) can be rewritten in more condensed form as follows :

$$\begin{cases} Z^T A V = \hat{A} \\ Z^T V = I_n \end{cases} \iff \begin{bmatrix} X & I_n \end{bmatrix} (\mathcal{A}_H - \mathcal{B}_H s) \begin{bmatrix} Y \\ I_n \end{bmatrix} = 0, \tag{5.21}$$

by defining

$$\mathcal{A}_H := \begin{bmatrix} B_\perp A C_\perp & B_\perp A C^+ \hat{C} \\ \hat{B} B^+ A C_\perp & \hat{B} B^+ A C^+ \hat{C} - \hat{A} \end{bmatrix}, \quad \mathcal{B}_H := \begin{bmatrix} B_\perp C_\perp & B_\perp C^+ \hat{C} \\ \hat{B} B^+ C_\perp & \hat{B} B^+ C^+ \hat{C} - I_n \end{bmatrix}. \tag{5.22}$$

The following result follows from the preceding discussion [54].

Lemma 5.7. *With the preceding notation, $\hat{T}(s)$ is embedded in $T(s)$ if and only if there exists $Z_1 \in \mathbb{C}^{n\times(n+k-m)}$, $V_1 \in \mathbb{C}^{(n+k-p)\times n}$, and invertible matrices $V_2, Z_2 \in \mathbb{C}^{n\times n}$ such that*

$$Z(\mathcal{A}_H - \mathcal{B}_H s)V := \begin{bmatrix} Z_1 & Z_2 \end{bmatrix} (\mathcal{A}_H - \mathcal{B}_H s) \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = 0. \tag{5.23}$$

By studying the *Generic Kronecker Form* of the pencil $\mathcal{A}_H - \mathcal{B}_H s$, Halevi found *generic* conditions for the embedding between two $p \times m$ transfer functions $T(s)$ and $\hat{T}(s)$ of Mc Millan degree $n + k$ and n depending on m, p, n and k. The idea consists in finding the generic conditions for the existence of a zero block of dimension $n \times n$ in the Kronecker Form of $\mathcal{A}_H - \mathcal{B}_H s$.

Unfortunately there is no interpretation concerning the pencil $\mathcal{A}_H - \mathcal{B}_H s$ in terms of quantities related to the embedding problem. An answer to this last question will be given in Subsection 5.5.

A Tangential Interpolation Approach

Define the following state space realization

$$(C_E, A_E, B_E) := \left\{ \begin{bmatrix} C & -\hat{C} \end{bmatrix}, \begin{bmatrix} A & \\ & \hat{A} \end{bmatrix}, \begin{bmatrix} B \\ \hat{B} \end{bmatrix} \right\}. \tag{5.24}$$

Clearly, (C_E, A_E, B_E) is a state space realization (not necessarily minimal) of $E(s) := T(s) - \hat{T}(s)$. Let us consider its corresponding system zero matrix,

$$s\mathcal{E}_I - \mathcal{A}_I := \begin{bmatrix} sI_{n+k} - A & 0 & B \\ 0 & sI_n - \hat{A} & \hat{B} \\ C & -\hat{C} & 0 \end{bmatrix}. \tag{5.25}$$

From Definition 1.25, the zeros of this pencil are invariant zeros of the state space realization (C_E, A_E, B_E). If $T(s)$ and $\hat{T}(s)$ do not have common poles, the generalized eigenvalues of $s\mathcal{E}_I - \mathcal{A}_I$ are also transmission zeros of $E(s)$ because the state space realization (C_E, A_E, B_E) is then a minimal state space realization of $E(s)$ together with (C, A, B) and $(\hat{C}, \hat{A}, \hat{B})$ minimal by assumption. This would contradict our hypothesis. As we will see, the Kronecker form of this pencil is another key tool to study the problem of embedding.

Let us consider here only the square case $m = p$. Let us further assume that the pencil $s\mathcal{E}_I - \mathcal{A}_I$ is regular, i.e. its rank is equal to $2n + k + m$. Then, there exist invertible matrices $\mathcal{Z}, \mathcal{V}$ that put the pencil $s\mathcal{E}_I - \mathcal{A}_I$ into the following Kronecker Form [42]:

$$\mathcal{Z}(s\mathcal{E}_I - \mathcal{A}_I)\mathcal{V} = \begin{bmatrix} sI_{2n+k-m} - J & 0 \\ 0 & I_{2m} \end{bmatrix} := s\tilde{\mathcal{E}}_I - \tilde{\mathcal{A}}_I, \tag{5.26}$$

with J a block Jordan matrix (generically diagonal) with diagonal elements $J(i,i) := \lambda_i$. If the eigenvalues λ_i are not equal to poles of $T(s)$ or $\hat{T}(s)$ [39], there exists a non zero vector v_i such that

$$T(\lambda_i)v_i = \hat{T}(\lambda_i)v_i, \quad 1 \le i \le 2n + k - m.$$

Because the matrices $\tilde{\mathcal{A}}_I$ and $\tilde{\mathcal{E}}_I$ commute, and $\mathcal{Z}, \mathcal{V}$ are nonsingular,

$$\left\{\mathcal{Z}\mathcal{A}_I\mathcal{V} = \tilde{\mathcal{A}}_I,\ \mathcal{Z}\mathcal{E}_I\mathcal{V} = \tilde{\mathcal{E}}_I\right\} \Longrightarrow \left\{\mathcal{A}_I\mathcal{V}\tilde{\mathcal{E}}_I = \mathcal{E}_I\mathcal{V}\tilde{\mathcal{A}}_I,\ \tilde{\mathcal{E}}_I\mathcal{Z}\mathcal{A}_I = \tilde{\mathcal{A}}_I\mathcal{Z}\mathcal{E}_I\right\}. \tag{5.27}$$

Let us decompose the matrices $\mathcal{Z}$ and $\mathcal{V}$ as follows:

$$\begin{aligned} \mathcal{Z} &:= \begin{array}{c} \\ (2n+k-m) \\ (2m) \end{array} \begin{array}{c} \begin{array}{ccc} (n+k) & (n) & (m) \end{array} \\ \begin{bmatrix} Z_{1f} & Z_{2f} & Z_{3f} \\ Z_{1i} & Z_{2i} & Z_{3i} \end{bmatrix} \end{array} := \begin{bmatrix} \mathcal{Z}_f \\ \mathcal{Z}_i \end{bmatrix}, \\ \mathcal{V} &:= \begin{array}{c} \\ (n+k) \\ (n) \\ (m) \end{array} \begin{array}{c} \begin{array}{cc} (2n+k-m) & (2m) \end{array} \\ \begin{bmatrix} V_{1f} & V_{1i} \\ V_{2f} & V_{2i} \\ V_{3f} & V_{3i} \end{bmatrix} \end{array} := \begin{bmatrix} \mathcal{V}_f & \mathcal{V}_i \end{bmatrix}. \end{aligned} \tag{5.28}$$

From (5.27),

$$\mathcal{A}_I \begin{bmatrix} V_{1f} & 0 \\ V_{2f} & 0 \\ V_{3f} & 0 \end{bmatrix} = \begin{bmatrix} V_{1f} & V_{1i} \\ V_{2f} & V_{2i} \\ 0 & 0 \end{bmatrix} \tilde{\mathcal{A}}_I, \quad \mathcal{A}_I^T \begin{bmatrix} Z_{1f}^T & 0 \\ Z_{2f}^T & 0 \\ Z_{3f}^T & 0 \end{bmatrix} = \begin{bmatrix} Z_{1f}^T & Z_{1i}^T \\ Z_{2f}^T & Z_{2i}^T \\ 0 & 0 \end{bmatrix} \tilde{\mathcal{A}}_I^T. \tag{5.29}$$

One deduces that V_{1i}, V_{2i}, Z_{1i} and Z_{2i} are equal to zero. Let us look at the finite part of the spectrum:

$$\begin{bmatrix} A & 0 & -B \\ 0 & \hat{A} & -\hat{B} \\ -C & \hat{C} & 0 \end{bmatrix} \begin{bmatrix} V_{1f} \\ V_{2f} \\ V_{3f} \end{bmatrix} = \begin{bmatrix} V_{1f} \\ V_{2f} \\ 0 \end{bmatrix} J, \quad \begin{bmatrix} A & 0 & -B \\ 0 & \hat{A} & -\hat{B} \\ -C & \hat{C} & 0 \end{bmatrix}^T \begin{bmatrix} Z_{1f}^T \\ Z_{2f}^T \\ Z_{3f}^T \end{bmatrix} = \begin{bmatrix} Z_{1f}^T \\ Z_{2f}^T \\ 0 \end{bmatrix} J^T.$$

One deduces that

$$CV_{1f} = \hat{C}V_{2f}, \quad Z_{1f}B = -Z_{2f}\hat{B}. \tag{5.30}$$

Moreover, from (5.26)

$$\mathcal{Z}_f(s\mathcal{E}_I - \mathcal{A}_I)\mathcal{V}_f = sI_{2n+k-m} - J.$$

Let us consider a set of n rows of $\mathcal{Z}_f$, $\begin{bmatrix} Z_1 & Z_2 & Z_3 \end{bmatrix}$, and n columns of $\mathcal{V}_f$, $\begin{bmatrix} V_1^T & V_2^T & V_3^T \end{bmatrix}^T$ such that

$$\begin{bmatrix} Z_1 & Z_2 & Z_3 \end{bmatrix} (s\mathcal{E}_I - \mathcal{A}_I) \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = 0. \tag{5.31}$$

If $k \geq m$, it is always possible to find such matrices. For instance, the last n columns of $\mathcal{V}_f$ and the first n rows of $\mathcal{Z}_f$ would satisfy (5.31), but other choices are possible, depending on the structure of J. If J is diagonal, just take rows of $\mathcal{Z}_f$ and columns of $\mathcal{V}_f$ of different indices. If the square matrices V_2 and Z_2 are invertible (generically, this is the case), then $\hat{T}(s)$ is embedded in $T(s)$. Indeed, from (5.30) and (5.31),

$$Z_1AV_1 = -Z_2\hat{A}V_2, \quad Z_1V_1 = -Z_2V_2, \quad Z_1B = -Z_2\hat{B}, \quad CV_1 = \hat{C}V_2.$$

By defining $Z^T := -Z_2^{-1}Z_1$ and $V := V_1V_2^{-1}$, conditions (5.2) are satisfied and $\hat{T}(s)$ is embedded in $T(s)$. To conclude, if $k \geq m$, then $\hat{T}(s)$ is *generically* embedded in $T(s)$. If $k < m$, then $\hat{T}(s)$ is *generically* not embedded in $T(s)$, because given two arbitrary $m \times m$ transfer functions $T(s)$ and $\hat{T}(s)$ of Mc Millan degree $n+k$ and n, the rank of the difference $T(s) - \hat{T}(s)$ is generically equal to m instead of k as it should be if $\hat{T}(s)$ were embedded in $T(s)$, as shown in Proposition 5.2. It can be shown that in the special case $m = k$, there exists only a finite number of projecting matrices and these come from rows and columns of $\mathcal{Z}$ and $\mathcal{V}$.

In general, the following conjecture [85] has been proposed.

Conjecture 5.8. A minimal state space realization of the strictly proper transfer function $\hat{T}(s)$ of Mc Millan degree n is embedded in the strictly proper transfer function $T(s)$ of Mc Millan degree $n+k$ if and only if there exist two regular pencils, $R^{(r)} - sR^{(l)}$ and $L^{(r)} - sL^{(l)}$ such that the matrices $M, \hat{M}, R, \hat{R}, \mathcal{X}$ and $\mathcal{Y}$ of the following equations

$$\begin{bmatrix} A - sI_{n+k} & 0 & B \\ 0 & \hat{A} - sI_n & \hat{B} \\ C & -\hat{C} & 0 \end{bmatrix} \begin{bmatrix} NR^{(l)} \\ \hat{N}R^{(l)} \\ \mathcal{Y} \end{bmatrix} = \begin{bmatrix} N \\ \hat{N} \\ 0 \end{bmatrix} (R^{(r)} - sR^{(l)}), \tag{5.32}$$

$$\begin{bmatrix} L^{(l)}M & -L^{(l)}\hat{M} & \mathcal{X} \end{bmatrix} \begin{bmatrix} A - sI_{n+k} & 0 & B \\ 0 & \hat{A} - sI_n & -\hat{B} \\ C & \hat{C} & 0 \end{bmatrix}$$
$$= (L^{(r)} - sL^{(l)}) \begin{bmatrix} M & -\hat{M} & 0 \end{bmatrix}, \tag{5.33}$$

satisfy the following conditions :

1. $\left[L^{(l)}M \middle| -L^{(l)}\hat{M} \middle| \mathcal{X} \right] (s\mathcal{E}_I - \mathcal{A}_I) \begin{bmatrix} NR^{(l)} \\ \hat{N}R^{(l)} \\ \mathcal{Y} \end{bmatrix} = 0,$
2. $\dim \left(Im(\hat{N}R^{(l)}) \right) = \dim \left(Im(\hat{M}^T {L^{(l)}}^T) \right) = n.$

The conditions given by Conjecture 5.8 are at least sufficient. Indeed, from equations (5.33), (5.32) and the regularity assumption of $R^{(r)} - sR^{(l)}$ and ${L^{(r)}}^T - s{L^{(l)}}^T$, it follows that

$$CNR^{(l)} = \hat{C}\hat{N}R^{(l)} \quad , \quad L^{(l)}MB = L^{(l)}\hat{M}\hat{B}. \tag{5.34}$$

Then, from condition 1,

$$L^{(l)}MNR^{(l)} = L^{(l)}\hat{M}\hat{N}R^{(l)} \quad , \quad L^{(l)}MANR^{(l)} = L^{(l)}\hat{M}\hat{A}\hat{N}R^{(l)}. \tag{5.35}$$

Finally, conditions 1 and 2 imply that the matrices $\hat{N}R^{(l)}$ and $\hat{M}^T {L^{(l)}}^T$ are right invertible. Defining $Z, V \in \mathbb{C}^{n+k \times n}$ by

$$Z^T = (L^{(l)}\hat{M})^{-l} L^{(l)}M, \quad V = NR^{(l)}(\hat{N}R^{(l)})^{-r}, \tag{5.36}$$

we can easily verify equations (5.1) and (5.2).

Equations (5.32) and (5.33) give us the following Sylvester equations :

$$ANR^{(l)} - NR^{(r)} + B\mathcal{Y} = 0 \quad , \quad L^{(l)}MA - L^{(r)}M + \mathcal{X}C = 0. \tag{5.37}$$

These Sylvester equations correspond to generalized left and right eigenspaces of the system zero matrix (5.25) (see Lemma 4.24). The choice of matrices $L^{(r)}, L^{(l)}, R^{(r)}, R^{(l)}, \mathcal{X}$ and $\mathcal{Y}$ correspond to respectively left and right tangential interpolation conditions at the eigenvalues σ_i of $(M_r - sN_r)$ and γ_j of $(M_l - sN_l)$, that are satisfied between $T(s)$ and $\hat{T}(s)$ (see Chapter 4). These eigenspaces correspond to disjoint parts of the spectrum of $M - Ns$ such that the product $L^{(l)}MNR^{(l)} = L^{(l)}\hat{M}\hat{N}N_r$ is invertible.

Conjecture 5.8 states that any projected reduced-order transfer function can be obtained by imposing some interpolation conditions or some modal approximation conditions with respect to the original transfer function. In other words, if a transfer function is embedded into another then it must be possible to construct the projecting matrices from Sylvester equations.

5.5 Link between the Pencil of Halevi and Tangential Interpolation

In Chapters 3 and 4, interpolation techniques for model reduction have been studied in details. At first sight, there seems to be no common points between the problem of embedding of two transfer functions and the problem of interpolation. It turns out that these two problems are closely related. In this section, it is shown that the pencil of Halevi is a subpencil of the system zero matrix of the error. Let us state this formally.

Lemma 5.9. *The zeros of the pencil $\mathcal{A}_H - \mathcal{B}_H s$ (defined in* (5.22)*) are invariant zeros of the state space realization (C_E, A_E, B_E) (defined in* (5.24)*).*

Proof. The pencil $s\mathcal{E}_I - \mathcal{A}_I$ defined in (5.25) is strictly equivalent to the new pencil $\mathcal{A}_1 - \mathcal{B}_1 s$ defined by

$$\begin{bmatrix} \begin{bmatrix} B^+ \\ B_\perp \end{bmatrix} & & \\ & I_n & \\ & & I_p \end{bmatrix} (s\mathcal{E}_I - \mathcal{A}_I) \begin{bmatrix} \begin{bmatrix} C^+ & C_\perp \end{bmatrix} & & \\ & I_n & \\ & & I_m \end{bmatrix}$$

$$= \begin{bmatrix} B^+(A-sI)C^+ & B^+(A-sI)C_\perp & 0 & I_m \\ B_\perp(A-sI)C^+ & B_\perp(A-sI)C_\perp & 0 & 0 \\ 0 & 0 & \hat{A} - sI_n & \hat{B} \\ I_p & 0 & -\hat{C} & 0 \end{bmatrix} \tag{5.38}$$

$$\doteq \mathcal{A}_1 - \mathcal{B}_1 s. \tag{5.39}$$

This pencil, in turn, is strictly equivalent to the pencil $\mathcal{A}_2 - \mathcal{B}_2$ defined by

$$\begin{bmatrix} I_m & & & \\ & I_{n+k-m} & & \\ \hat{B} & & -I_n & \\ & & & I_p \end{bmatrix} (\mathcal{A}_1 - \mathcal{B}_1 s) \begin{bmatrix} I_p & & \hat{C} & \\ & I_{n+k-p} & & \\ & & I_n & \\ & & & I_m \end{bmatrix}$$

$$= \begin{bmatrix} B^+(A-sI_{n+k})C^+ & B^+(A-sI_{n+k})C_\perp & B^+(A-sI_{n+k})C^+\hat{C} & I_m \\ B_\perp(A-sI_{n+k})C^+ & B_\perp(A-sI_{n+k})C_\perp & B_\perp(A-sI_{n+k})C^+\hat{C} & 0 \\ \hat{B}B^+(A-sI_{n+k})C^+ & \hat{B}B^+(A-sI_{n+k})C_\perp & -\hat{A} + sI_n + \hat{B}B^+(A-sI_{n+k})C^+\hat{C} & 0 \\ I_p & 0 & 0 & 0 \end{bmatrix}$$

$$= \mathcal{A}_2 - \mathcal{B}_2 s \tag{5.40}$$

$$\doteq \begin{bmatrix} * & * & I_m \\ * & \mathcal{A}_H - \mathcal{B}_H s & \\ I_p & & \end{bmatrix} \tag{5.41}$$

This concludes the proof.

From (5.41), it is clear that the invariant zeros of C_E, A_E, B_E are the zeros of $\mathcal{A}_H - \mathcal{B}_H s$ plus $p + m$ zeros at infinity. We think that this lemma can help to prove Conjecture 5.8.

5.6 Concluding remarks

Concerning the two pencil approaches, namely the approach of Halevi and the tangential interpolation approach, one cannot exclude the possibility that for any choice of matrices Z and V satisfying (5.23) or (5.31), one of the submatrices Z_2 or V_2 is singular. This is the reason why we say that the results of that section are only generic.

The main results of this chapter are the following. On the one hand, necessary and sufficient conditions for the problem of embedding have been found in the particular case where the difference of Mc Millan degree between transfer functions devoid of common poles is equal to one. In such a case, the embedding of $\hat{T}(s)$ into $T(s)$ turns out to be equivalent to a factorization property of $T(s) - \hat{T}(s)$. On the other hand, more general results are obtained by studying the structure of well chosen system zero pencils, namely the pencil of Halevi and the system zero matrix. An open problem states that if a transfer function is embedded into another, then it is possible to construct the projecting matrices from Sylvester equations. Finally, it is shown that the pencil of Halevi (that appears in the problem of embedding) is a submatrix of the system zero pencil of the error (that appears in the tangential interpolation problem). This links the problem of embedding of this chapter with the tangential interpolation problem of Chapter 4.

Another reasonable conjecture in view of the preceding results is the following.

Conjecture 5.10. If the following condition holds:

$$n - k \geq m,$$

then any regular $m \times m$ transfer function $\hat{T}(s)$ of Mc Millan degree k is embedded in any $m \times m$ regular transfer function $T(s)$.

In other words, for practical model reduction problems of regular systems, the projection technique is conjectured to be universal.

5.7 Notes and References

The existence of projecting matrices Z, V satisfying (5.1,5.2) is therefore related to the above sub-matrix problem. A square matrix $\hat{A}$ is said to be *embedded* in a square matrix A when there exists a change of coordinates S such that $\hat{A} - sI_k$ is a sub-matrix of $S^{-1}(A - sI_n)S$. Necessary and sufficient conditions for the embedding of such monic pencils are given in [81, 65].

As for monic pencils, we say that the pencil $\hat{M} - \hat{N}s$ is embedded in the pencil $M - Ns$ when there exist invertible matrices Le, Ri such that $\hat{M} - \hat{N}s$ is

a sub-matrix of $Le(M - Ns)Ri$. Finding necessary and sufficient conditions for the embedding of such general pencils is still an open problem [64]. Nevertheless, one obtains from [81, 65, 64] *necessary* conditions on $(\hat{C}, \hat{A}, \hat{B})$ and (C, A, B) for $Z_{\hat{C},\hat{A},\hat{B}}(s)$ to be embedded in $Z_{C,A,B}(s)$. These obviously give *necessary* conditions for the existence of projecting matrices Z, V satisfying (5.1,5.2).

This chapter is the result of a collaboration with Yves Genin [43]. We would like to thank Yoram Halevi for helpful discussions around this topic.

6

Model Reduction of Second Order Systems

6.1 Introduction

In this chapter, the problem of constructing a reduced order system while preserving the second order structure of the original system is discussed. After a brief introduction on second order systems, two classes of second order structure preserving model reduction techniques – Krylov subspace-based and SVD-based – are presented. For the Krylov techniques, conditions on the projectors that guarantee the reduced second order system tangentially interpolates the original system at given frequencies are derived and an algorithm is described. For SVD-based techniques, a Second Order Balanced Truncation method is derived from *second order gramians*.

Let us first discuss two areas where second order systems appear.

Mechanical Systems

Second order systems arise naturally in many areas of engineering (see, for example, [71, 73, 90]) with the following form :

$$\begin{cases} M\ddot{q}(t) + D\dot{q}(t) + Sq(t) = F^{in}\ u(t), \\ \qquad\qquad\qquad\qquad y(t) = F^{out}\ q(t). \end{cases} \tag{6.1}$$

We assume that $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$, $q(t) \in \mathbb{R}^n$, $F^{in} \in \mathbb{R}^{n\times m}$, $F^{out} \in \mathbb{R}^{p\times n}$, and $M, D, S \in \mathbb{R}^{n\times n}$ with M invertible. For mechanical systems the matrices M, D and S represent, respectively, the *mass* (or *inertia*), *damping* and *stiffness* matrices, $u(t)$ corresponds to the vector of *external* forces, F^{in} is the input distribution matrix, $y(\cdot)$ is the output measurement vector, F^{out} is the output measurement matrix, and $q(t)$ to the vector of *internal generalized coordinates*.

General RLC Circuits

Another area where (generalized) second order systems naturally appear is in Circuit Theory. Let us consider this in more details, following the description in [28].

The connectivity of a circuit can be captured using a directional graph. The nodes of the graph correspond to the nodes of the circuits, and the edges of the graph correspond to each circuit element. By convention, a direction is assigned to graph edges in order to distinguish between the source and destination nodes. The adjacency matrix, $\mathbf{A}$, of the directional graph describes the connectivity of a circuit. Each row of the matrix corresponds to a graph edge and, therefore, to a circuit element. Each column of the matrix corresponds to a graph or circuit node. The column corresponding to the datum (ground) node of the circuit is omitted in order to remove redundancy. By convention, a row of the connectivity matrix will contain $+1$ in the column corresponding to the source node, -1 in the columns corresponding to the destination node, and 0 everywhere else. It is easy to see that Kirchhoff's laws, which depend only on connectivity can be expressed using the adjacency matrix

$$\begin{aligned} Kirchhoff's\ Current\ Law: &\quad A^T i_b = 0, \\ Kirchhoff's\ Voltage\ Law: &\quad A v_n = v_b, \end{aligned} \tag{6.2}$$

where i_b and v_b are the vectors of branch currents and voltages, respectively, and v_n is the vector of the non-datum node voltages.

We are interested in analysing RLC circuits and for simplicity we assume that the circuit is excited just by current sources. In this case the adjacency matrix and the branch current and voltage vectors can be partitioned accordingly to circuit-element types as follows:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_i \\ \mathbf{A}_g \\ \mathbf{A}_c \\ \mathbf{A}_l \end{bmatrix}, \quad \mathbf{v}_b = \begin{bmatrix} \mathbf{v}_i \\ \mathbf{v}_g \\ \mathbf{v}_c \\ \mathbf{A}_l \end{bmatrix}, \quad \mathbf{i}_b = \begin{bmatrix} \mathbf{i}_i \\ \mathbf{i}_g \\ \mathbf{i}_c \\ \mathbf{i}_l \end{bmatrix}.$$

Here, the subscripts i, g, c and l stand for branches containing current sources, resistors, capacitors, and inductors, respectively.

The set of circuit equations is completed by adding the so-called *branch constitutive relationships* (BCR's), which describe the physical behavior of the circuit elements. In the case of RLC circuits, the BCR's are as follows:

$$\mathbf{i}_i = -\mathbf{i}_t(t), \quad \mathbf{i}_c = \mathcal{G}\mathbf{v}_g, \quad \mathbf{i}_c = \mathcal{C}\frac{d}{dt}\mathbf{v}_c, \quad \mathbf{v}_l = \mathcal{L}\frac{d}{dt}\mathbf{i}_l. \tag{6.3}$$

Here, $\mathbf{i}_t(t)$ is the vector of current-source values, $\mathcal{G}$ and $\mathcal{C}$ are appropriately-sized diagonal matrices whose diagonal entries are the conductance and capacitance

values of each element. It is clear that these values are positive for any physical circuit. The matrix $\mathcal{L}$ is also diagonal in the absence of inductive coupling. Inductive coupling introduces off-diagonal terms in the inductance matrix, but $\mathcal{L}$ remains symmetric and positive definite.

The modified nodal formulation (MNA) of the circuit equations is obtained by combining the Kirchhoff equations (6.2) with the BCRs (6.3), and eliminating as many current unknows as possible. For the case of RLC circuits only inductor currents need to be left as unknowns. The resulting MNA equations are as follows:

$$\mathbf{A}_g^T \mathcal{G} \mathbf{A}_g \mathbf{v}_n + \mathbf{A}_c^T \mathcal{C} \mathbf{A}_c \frac{d}{dt} \mathbf{v}_n + \mathbf{A}_l^T \mathbf{i}_l = \mathbf{A}_i^T \mathbf{i}_t(t), \tag{6.4}$$

$$\mathbf{A}_l \mathbf{v}_n - \mathcal{L} \frac{d}{dt} \mathbf{i}_l = 0. \tag{6.5}$$

In matrix form, we obtain

$$\underbrace{\begin{bmatrix} A_g^T \mathcal{G} A_g & A_l^T \\ A_l & 0 \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} \mathbf{v}_n \\ \mathbf{i}_l \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} A_c^T \mathcal{C} A_c & 0 \\ 0 & -\mathcal{L} \end{bmatrix}}_{C} \underbrace{\frac{d}{dt} \begin{bmatrix} \mathbf{v}_n \\ \mathbf{i}_l \end{bmatrix}}_{\frac{d}{dt}x} = \underbrace{\begin{bmatrix} A_i^T \\ 0 \end{bmatrix}}_{B} \mathbf{i}_t(t).$$

Assume for simplicity that the matrix $\mathbf{A}_l$ is left invertible, and denote $\mathbf{A}_l^{-left}$ its left inverse. By injecting (6.5) into (6.4), one obtains the second order equation

$$\mathbf{A}_g^T \mathcal{G} \mathbf{A}_g \mathbf{A}_l^{-left} \mathcal{L} \frac{d}{dt} \mathbf{i}_l + \mathbf{A}_c^T \mathcal{C} \mathbf{A}_c \mathbf{A}_l^{-left} \mathcal{L} \frac{d^2}{dt^2} \mathbf{i}_l + \mathbf{A}_l^T \mathbf{i}_l = \mathbf{A}_i^T \mathbf{i}_t(t).$$

The input is chosen as the vector of current source values $x(t) := \mathbf{i}_t(t)$ and the output is chosen as the vector of voltages across the excitation sources, $y(t) := v_i(t) = A_i v_n = B^T x(t)$. In the Laplace domain, this gives rise to the following transfer function

$$y(s) = B^T (G + sC)^{-1} B x(s).$$

This is not strictly speaking a second order system. Nevertheless, by choosing $\mathbf{i}_l$ as the output, one retrieves the second order form. The important point here is that the matrices B, G and C describing the system are structured, and it makes to preserve such a structure in the reduced order system.

Description of the Problem

Let us adopt the notation corresponding to a mechanical system. The transfer function associated with the system (6.1) is

$$T(s) := F^{out} P(s)^{-1} F^{in}, \tag{6.6}$$

where

$$P(s) := Ms^2 + Ds + S \tag{6.7}$$

is the *characteristic polynomial matrix*. The zeros of $\det P(s)$ are also known as the *characteristic frequencies* of the system and play an important role in model reduction, e.g., the system is stable if these zeros lie in the left half plane.

Often, the original system is too large to allow the efficient solution of various control or simulation tasks. In order to address this problem, techniques that produce a reduced system of size $k \ll n$ that possesses the essential properties of the full order model have been developed. Such a reduced model can then be used effectively, e.g., in real-time, for controlling or simulating the phenomena described by the original system. If the original system has a second order form, it makes sense to construct a reduced order system that preserves this second order form. We therefore need to build a reduced model,

$$\begin{cases} \hat{M}\ddot{\hat{q}}(t) + \hat{D}\dot{\hat{q}}(t) + \hat{S}\hat{q}(t) = \hat{F}^{in}u(t) \\ \hat{y}(t) = \hat{F}^{out}\hat{q}(t) \end{cases} \tag{6.8}$$

where $\hat{q}(t) \in \mathbb{R}^k$, $\hat{M}$, $\hat{D}$, $\hat{S} \in \mathbb{R}^{k\times k}$, $\hat{F}^{in} \in \mathbb{R}^{k\times m}$, $\hat{F}^{out} \in \mathbb{R}^{p\times k}$, such that its transfer function is "close" to the original transfer function.

Another important property that one often wants to preserve is the symmetry. If the matrices M, D, S are symmetric, and $F^{out} = {F^{in}}^T$, the transfer function $T(s)$ is symmetric. There also exist Model Reduction techniques for second order systems that preserve this symmetry for the reduced order system [23, 66]. This point is not treated here.

Second order systems can be seen as a particular class of linear systems. Indeed, by rewriting the system (6.1) as follows

$$\begin{cases} \dot{x}(t) = \begin{bmatrix} 0 & I_n \\ -S_M & -D_M \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ F_M^{in} \end{bmatrix} u(t) \\ \\ y(t) = \begin{bmatrix} F_M^{out} & 0 \end{bmatrix} x(t) \end{cases} \tag{6.9}$$

where the state $x(t)$ is $\begin{bmatrix} q(t)^T & \dot{q}(t)^T \end{bmatrix}^T$, and where we have chosen a coordinate system in which the mass matrix M is the identity (for simplicity, the mass matrix M is assumed to be invertible, and one can write for example: $S_M = M^{-1}S$, $D_M = M^{-1}D$, $F_M^{in} = M^{-1}F^{in}$, $F_M^{out} = F^{out}$), one recovers the form (1.13) of a linear system. We can thus rewrite the transfer function defined in (6.6) as

$$T(s) = C(sI_{2n} - A)^{-1}B \tag{6.10}$$

by defining

$$A := \begin{bmatrix} 0 & I_n \\ -S_M & -D_M \end{bmatrix} \quad , \quad B := \begin{bmatrix} 0 \\ F_M^{in} \end{bmatrix} \quad , \quad C := \begin{bmatrix} F_M^{out} & 0 \end{bmatrix}. \tag{6.11}$$

Note that if the dimension of the state $q(t)$ of the original second order system (6.1) is equal to n, the order of its corresponding linearized state space realization (6.11) (also called the Mc Millan degree of $T(s)$ if the (C, A, B) is minimal) is equal to $2n$.

A reduced model for the second order system (6.1) could be produced by applying standard linear model reduction techniques to (C, A, B) in (6.11) to yield a small linear system $(\hat{C}, \hat{A}, \hat{B})$. Unfortunately, there is no guarantee that the matrices defining the reduced system $(\hat{C}, \hat{A}, \hat{B})$ have the nonzero structure necessary to preserve the second order form of the original system. Such a guarantee requires the development of second order structure preserving model reduction techniques. Other reasons advocating for second order structure preserving model reduction methods are given in the concluding remarks of this chapter.

This chapter is organized as follows. In Section 6.2, a simple sufficient condition for constructing reduced order systems that preserve the second order structure is developed. Generalizations of Balanced Truncation and Krylov subspace-based methods that enforce this sufficient condition for second order systems are presented in Sections 6.3 and and 6.4 respectively. After some numerical experiments in Section 6.5, concluding remarks are made in Section 6.6.

6.2 Second order structure preserving model reduction

In this section, a simple sufficient condition for obtaining a second order reduced system from a second order system is presented. The following result can be found in a slightly different form in [21].

Lemma 6.1. *Let (C, A, B) be the state space realization defined in* (6.11)*. If one projects such a state space realization with $2n \times 2k$ bloc diagonal matrices*

$$\mathcal{Z} := \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix}, \quad \mathcal{V} := \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}, \quad \mathcal{Z}^T \mathcal{V} = I_{2k},$$

where $Z_1, V_1, Z_2, V_2 \in \mathbb{R}^{n \times k}$, then the reduced transfer function

$$\hat{T}(s) := C\mathcal{V} \left(\mathcal{Z}^T (sI_{2n} - A)\mathcal{V} \right)^{-1} \mathcal{Z}^T B$$

is a second order transfer function, provided the matrix $Z_1^T V_2$ is invertible.

Proof. First, notice that the transfer function does not change under any similarity transformation of the system matrices. Let us consider the similarity transformation $W \in \mathbb{R}^{2k \times 2k}$ such that

$$W := \begin{bmatrix} X & \\ & Y \end{bmatrix},$$

with $X, Y \in \mathbb{R}^{k\times k}$ verifying

$$X^{-1}(Z_1^T V_2)Y = I_k.$$

From the preceding results,

$$\begin{aligned}\hat{T}(s) &:= C\mathcal{V}W\left(W^{-1}\mathcal{Z}^T(sI_{2n}-A)\mathcal{V}W\right)^{-1}W^{-1}\mathcal{Z}^T B\\ &= F_M^{out}V_1X\left(s^2I_k + sY^{-1}Z_2^T D_M V_2 Y + Y^{-1}Z_2^T S_M V_1 X\right)^{-1}Y^{-1}Z_2^T F_M^{in}.\end{aligned}$$

This is clearly a second order transfer function. □

□

6.3 Second order Balanced Truncation

The earliest balanced truncation technique for second order systems known to the authors is described in [66]. Based on this work, an alternative technique was developed in [21]. In this section an overview of the latter method, called SOBT (Second Order Balanced Truncation), is given.

The first step in the development of SOBT, based on a balance and truncate process similar to that discussed in Section 2.4, involves the definition of *two* pairs of $n \times n$ gramians ("second order gramians") that change according to contragradient transformations, and that have some energetic interpretation. The first pair $(\mathcal{P}_{pos}, \mathcal{Q}_{pos})$ corresponds to an energy optimization problem depending only on the *positions* $q(t)$ and not on the *velocities* $\dot{q}(t)$. Reciprocally, the second pair $(\mathcal{P}_{vel}, \mathcal{Q}_{vel})$ correspond to an optimization problem depending only on the velocities $\dot{q}(t)$ and not the on the positions $q(t)$. By analogy to the first order case, the gramians $\mathcal{Q}_{pos}$ and $\mathcal{Q}_{vel}$ are defined from the dual systems. Given the gramians, a balancing step in the method is defined by transforming to a coordinate system in which the second order gramians are equal and diagonal : $\bar{\mathcal{P}}_{pos} = \bar{\mathcal{Q}}_{pos} = \Sigma_{pos}, \quad \bar{\mathcal{P}}_{vel} = \bar{\mathcal{Q}}_{vel} = \Sigma_{vel}$. Their diagonal values enable us to identify the *important* positions and the *important* velocities, i.e. those with (hopefully) large effect on the I/O map. Once identified the reduced second order model follows by truncation of all variables not identified as important.

In order to define a pair of second order gramians measuring the contribution of the position coordinates (independently of the velocities) with respect to the I/O map, consider an optimization problem naturally associated with the second order system (see [66]) of the form

$$\min_{\dot{q}_0\in\mathbb{R}^n}\ \min_{u(t)}\ J(u(t), -\infty, 0), \tag{6.12}$$

subject to

$$\ddot{q}(t) + D_M\dot{q}(t) + S_M q(t) = F_M^{in}u(t),\ q(0) = q_0.$$

One easily sees that the optimum is $q_0^T {\mathcal{P}_{11}}^{-1} q_0$, where $\mathcal{P}_{11}$ is the $n \times n$ left upper block of $\mathcal{P}$, where $\mathcal{P}$ is the controllability gramian satisfying equation (1.18) with (C, A, B) given in (6.11). Starting with (2.18) we must solve

$$\min_{\dot{q}_0 \in \mathbb{R}^n} \; J_{q_0}(\dot{q}_0) \;=\; \begin{bmatrix} q_0^T & \dot{q}_0^T \end{bmatrix} \mathcal{P}^{-1} \begin{bmatrix} q_0 \\ \dot{q}_0 \end{bmatrix}$$

Partitioning $\mathcal{P}^{-1}$ as follows

$$\mathcal{P}^{-1} \;=\; \begin{bmatrix} R_1 & R_2 \\ R_2^T & R_3 \end{bmatrix}$$

and annihilating the gradient of $J_{q_0}(\dot{q}_0)$ gives the relation $\dot{q}_0 = -R_3^{-1} R_2^T q_0$. The value of J_{q_0} at this point is $q_0^T (R_1 - R_2 R_3^{-1} R_2^T) q_0$. This is simply the Schur complement of R_3 which is ${\mathcal{P}_{11}}^{-1}$. Similarly, the solution of the dual problem corresponds to $q_0^T {\mathcal{Q}_{11}}^{-1} q_0$, where $\mathcal{Q}_{11}$ is the $n \times n$ left upper block of $\mathcal{Q}$ (1.18).

Note that the transfer function is seen as a linear operator acting between two Hilbert spaces. The dual of such an operator is defined in Definition 1.31. It follows that the dual of a second order transfer function might not be a second order transfer function. This has no consequences here since only the energy transfer interpretation between the inputs, the outputs, the initial positions and velocities is important. Under the change of coordinates $q(t) = T\bar{q}(t)$, it is easy to verify that this pair of gramians undergo a contragradient transformation :

$$(\bar{\mathcal{P}}_{11}, \bar{\mathcal{Q}}_{11}) = (T^{-1} \mathcal{P}_{11} T^{-T}, T^T \mathcal{Q}_{11} T).$$

This implies that there exists a new coordinate system such that both $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ are equal and diagonal. Their energetic interpretation is seen by considering the underlying optimization problem. In (6.12), the energy necessary to reach the given position q_0 over all past inputs and initial velocities is minimized. Hence, these gramians describe the distribution of the I/O energy among the positions.

A pair of second order gramians that give the contribution of the velocities with respect to the I/O map can be defined analogously. The associated optimization problem is

$$\min_{q_0 \in \mathbb{R}^n} \min_{u(t)} \; J(u(t), -\infty, 0) \tag{6.13}$$

subject to

$$\ddot{q}(t) + D_M \dot{q}(t) + S_M q(t) \;=\; F_M^{in} u(t), \quad \dot{q}(0) \;=\; \dot{q}_0.$$

Following the same reasoning as before for the optimization problem (6.12), one can show that the solution of (6.13) is $\dot{q}_0^T {\mathcal{P}_{22}}^{-1} \dot{q}_0$, where $\mathcal{P}_{22}$ is the $n \times n$ right lower block of $\mathcal{P}$. The solution of the dual problem is $\dot{q}_0^T {\mathcal{Q}_{22}}^{-1} \dot{q}_0$, where $\mathcal{Q}_{22}$

is the $n \times n$ right lower block of $\mathcal{Q}$. As before, under the change of coordinates $q(t) = T\bar{q}(t)$ one can check that this pair of gramians undergo a contragradient transformation and the energetic interpretation is given by considering the underlying optimization problem. In (6.13), the energy necessary to reach the given velocity $\dot{q}_0$ over all past inputs and initial positions is minimized. Hence, these gramians describe the distribution of the I/O energy among the velocities.

Given the interpretation above these second order gramians are good candidates for balancing and truncation. Therefore, we choose :

$$(\mathcal{P}_{pos},\ \mathcal{Q}_{pos}) \ = \ (\mathcal{P}_{11},\ \mathcal{Q}_{11}) \quad \text{and} \quad (\mathcal{P}_{vel},\ \mathcal{Q}_{vel}) \ = \ (\mathcal{P}_{22},\ \mathcal{Q}_{22})\ . \tag{6.14}$$

It is not possible to balance both pairs of second order gramians at the same time with a single change of coordinates of the type $q(t) = T\bar{q}(t)$. A change of coordinates is required for both positions and velocities (unlike the approach in [66]). Therefore, we work in a state-space context, starting with the system (6.11). The SOBT method, therefore, first computes both pairs of second order gramians, $(\mathcal{P}_{pos},\ \mathcal{Q}_{pos})$ and $(\mathcal{P}_{vel},\ \mathcal{Q}_{vel})$. Given the gramians, the contragradient transformations that make $\mathcal{P}_{pos} = \mathcal{Q}_{pos} = \Lambda_{pos}$ and $\mathcal{P}_{vel} = \mathcal{Q}_{vel} = \Lambda_{vel}$, where Λ_{pos} and Λ_{vel} are positive definite diagonal matrices, are computed. Finally, truncate the positions corresponding to the smallest eigenvalues of Λ_{pos} and the velocities corresponding to the smallest eigenvalues of Λ_{vel}.

At present, there exists no a priori global error bound for SOBT and the stability of the reduced system is not guaranteed. Nevertheless, SOBT yields good numerical results, providing reduced transfer functions with approximation error comparable with the traditional Balanced Truncation technique.

6.4 Second Order Structure Preserving Krylov Techniques

The Krylov subspace-based methods discussed in Chapters 3 and 4 do not preserve second order structure when applied to the linear system (6.11). It is possible to modify them to satisfy the constraint presented in Section 6.2 and thereby produce a second order reduced system. Section 6.4 summarizes the earliest Krylov subspace-based method for second order systems [80]. The simple technique constructs, via projection, a second order reduced transfer function that matches the Markov parameters ($\lambda = \infty$) of the original transfer function. The limitation of the technique when applied to an arbitrary complex interpolation point is also discussed. Section 6.4, addresses this limitation using a generalization that allows multipoint rational interpolation. Finally, the problem of second order structure preserving tangential interpolation is solved in 6.4.

A particular case : Matching the Markov parameters

Su and Craig proposed a Krylov subspace-based projection method that preserves second order structure while matching the Markov parameters of the original transfer function [80]. The method is based on the observation that the right Krylov subspaces corresponding to interpolation at $\lambda = \infty$ for the system (6.11) has the form

$$\begin{bmatrix} B & AB & A^2B & \dots \end{bmatrix} = \begin{bmatrix} 0 & F_M^{in} & -D_M F_M^{in} & \dots \\ F_M^{in} & -D_M F_M^{in} & -S_M F_M^{in} + D_M^2 F_M^{in} & \dots \end{bmatrix} \tag{6.15}$$

$$= \begin{bmatrix} 0 & Q_{v,0} & Q_{v,1} & \dots \\ Q_{v,0} & Q_{v,1} & Q_{v,2} & \dots \end{bmatrix}. \tag{6.16}$$

and that if we write

$$\mathcal{K}_k(A, B) = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix},$$

it follows that

$$Im(V_1) \subseteq Im(V_2).$$

So by projecting the state space realization (6.11) with

$$\mathcal{V} := \begin{bmatrix} V_2 & 0 \\ 0 & V_2 \end{bmatrix}, \quad \mathcal{Z} := \begin{bmatrix} Z & 0 \\ 0 & Z \end{bmatrix}$$

such that $Z^T V_2 = I_k$, we obtain an interpolating second order transfer function of the form

$$\hat{T}(s) = F_M^{out} V_2 \left(Z^T (s^2 I_n + s D_M + S_M)^{-1} V_2 \right) Z^T F_M^{in}. \tag{6.17}$$

Hence, a second order system with the same k first Markov parameters as the original second order system can be constructed by projecting with $Z, V \in \mathbb{R}^{n \times k}$ such that $Z^T V = I_k$ and the image of V contains the image of $Q_{v,0}, \dots, Q_{v,k-1}$. Since $\mathcal{K}_k(A, B) \subseteq \mathcal{V}$, it follows from Theorem 2.4 that the first k Markov parameters of $T(s)$ and $\hat{T}(s)$ are equal.

If we apply the construction for any interpolation points $\lambda \in \mathbb{R}$, the corresponding right Krylov space is such that

$$\mathcal{K}_k((\lambda I - A)^{-1}, (\lambda I - A)^{-1} B) = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix},$$

with A and B defined in (6.11) and

$$Im(V_1) \subseteq Im(V_2). \tag{6.18}$$

Let us prove this. First, using the Schur Complement formula (1.1), let us compute the inverse of $\lambda I - A$:

$$\begin{bmatrix} \lambda I & -I \\ S_M & \lambda I + D_M \end{bmatrix}^{-1} = \begin{bmatrix} \lambda^{-1}(I_n - \Gamma^{-1} S_M \lambda^{-1}) & \lambda^{-1}\Gamma^{-1} \\ -\Gamma^{-1} S_M \lambda^{-1} & \Gamma^{-1} \end{bmatrix}, \tag{6.19}$$

with $\Gamma := \lambda I_n + D_M + S_M \lambda^{-1}$.

Let us first compute $z_1 := \Gamma^{-1} B$. It is clear that

$$(\lambda I - A)^{-1} B = \begin{bmatrix} \lambda^{-1} z_1 \\ z_1 \end{bmatrix}.$$

Let us define the vectors $y_k, z_k \in \mathbb{R}^n$ as follows,

$$\begin{bmatrix} y_k \\ z_k \end{bmatrix} := (\lambda I - A)^{-k} B.$$

From (6.19),

$$y_k = \lambda^{-1}(y_{k-1} - z_k).$$

This implies (6.18).

Unfortunately, a similar statement can not be made for the *left* Krylov subspaces $\mathcal{K}_k((\lambda I - A)^{-T}, (\lambda I - A)^{-T} C^T)$. This implies that when the second order Krylov technique is extended to interpolation at arbitrary points in the complex plane by projecting as in (6.17), only k interpolation conditions can be imposed for a reduced second order system of Mc Millan degree $2k$.

Second Order Rational Interpolation

The projection technique of Su and Craig has been generalized independently in [32] and [83] to solve the rational interpolation problem that produces a *second order* transfer function of order k, i.e., of Mc Millan degree $2k$, $\hat{T}(s)$, that interpolates $T(s)$ at $2k$ points in the complex plane. (See also the references in the end of this chapter.) The conditions that determine the projections are given in Theorem 6.2 and the associated algorithm is presented.

By combining the results of Theorem 2.4 and Section 6.2, the following Theorem can be proven.

Theorem 6.2. *Let* $T(s) := F_M^{out}(s^2 I_n + D_M s + S_M)^{-1} F_M^{in} = C(sI_{2n} - A)^{-1} B$, *with*

$$A := \begin{bmatrix} 0 & I_n \\ -S_M & -D_M \end{bmatrix}, \quad B := \begin{bmatrix} 0 \\ F_M^{in} \end{bmatrix}, \quad C := \begin{bmatrix} F_M^{out} & 0 \end{bmatrix},$$

be a second order transfer function of Mc Millan degree $2n$, *i.e.* $S_M, D_M \in \mathbb{R}^{n \times n}$*). Let* $Z, V \in \mathbb{R}^{2n \times k}$ *be defined as*

$$V := \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad , \quad Z := \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix},$$

with V_1, V_2, Z_1 *and* $Z_2 \in \mathbb{R}^{n \times k}$ *such that*

$$Z_1^T V_1 = Z_2^T V_2 = I_k.$$

Let us define the $2n \times 2k$ *projecting matrices*

$$\mathcal{V} := \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} \quad , \quad \mathcal{Z} := \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix}.$$

Define the second order transfer function $\hat{T}(s)$ *of order* k *(and of Mc Millan degree* $2k$*) by*

$$\begin{aligned} \hat{T}(s) &:= C\mathcal{V}\left(\mathcal{Z}^T(sI_{2n} - A)\mathcal{V}\right)^{-1}\mathcal{Z}^T B \\ &:= \hat{C}(sI_{2k} - \hat{A})^{-1}\hat{B}. \end{aligned} \tag{6.20}$$

If

$$\bigcup_{i=1}^{K} \mathcal{K}_{J_{b_i}}((\lambda_i I_{2n} - A)^{-1}, (\lambda_i I_{2n} - A)^{-1}B) \subseteq Im(V) \tag{6.21}$$

and

$$\bigcup_{i=1}^{K} \mathcal{K}_{J_{c_i}}((\lambda_i I_{2n} - A)^{-T}, (\lambda_i I_{2n} - A)^{-T}C^T) \subseteq Im(Z) \tag{6.22}$$

where the interpolation points λ_i *are chosen such that the matrices* $\lambda_i I_{2n} - A$ *are invertible* $\forall i \in \{1, \ldots, K\}$ *then, if the matrix* $Z_1^T V_2$ *is invertible,*

$$T(s) - \hat{T}(s) = O(s - \lambda_i)^{J_{b_i} + J_{c_i}} \tag{6.23}$$

for the finite points λ_i*, provided these moments exist, i.e. provided the matrices* $\lambda_i I_{2k} - \hat{A}$ *are invertible and*

$$T(s) - \hat{T}(s) = O(s^{-1})^{J_{b_i} + J_{c_i}} \tag{6.24}$$

if $\lambda_i = \infty$*.*

Proof. Clearly, $\mathcal{Z}^T\mathcal{V} = I_{2k}$. The second order structure of $\hat{T}(s)$ follows from Lemma 6.1. It is clear that

$$Im(V) \subseteq Im(\mathcal{V}) \quad , \quad Im(Z) \subseteq Im(\mathcal{Z}).$$

The interpolation conditions are then satisfied because of Theorem 2.4. □

The form of the projectors allows the development of an algorithm similar to the Rational Krylov family of algorithms for first order systems [48]. The algorithm, shown below, finds a *second order* transfer function of order k, i.e. of Mc Millan degree $2k$, $\hat{T}(s)$, that interpolates $T(s)$ at $2k$ interpolation points λ_1 up to λ_{2k}, i.e.,

$$T(s) - \hat{T}(s) = O(\lambda_i - s) \qquad for\ 1 \leq i \leq 2k, \tag{6.25}$$

We assume for simplicity that the interpolation points are finite, distinct and not poles of $T(s)$. The algorithm is easily modified to impose higher order conditions at the interpolation points.

Algorithm 6.4.1 *1. Construct Z and V such that*

$$V = \left[\, (\lambda_1 I_{2n} - A)^{-1} B \ldots (\lambda_k I_{2n} - A)^{-1} B \,\right] = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

$$Z^T = \begin{bmatrix} C(\lambda_{k+1} I_{2n} - A)^{-1} \\ \vdots \\ C(\lambda_{2k} I_{2n} - A)^{-1} \end{bmatrix} = \left[\, Z_1^T \ Z_2^T \,\right],$$

where $V_1, V_2 \in \mathbb{R}^{n \times k}$ are the first n rows and the last n rows of V respectively and $Z_1, Z_2 \in \mathbb{R}^{n \times k}$ are the first n rows and the last n rows of Z respectively. Choose the matrices $M_1, M_2, N_1, N_2 \in \mathbb{R}^{k \times k}$ such that $N_1^T Z_1^T V_1 M_1 = N_2^T Z_2^T V_2 M_2 = I_k$.

2. Construct

$$\mathcal{V} := \begin{bmatrix} V_1 M_1 & \\ & V_2 M_2 \end{bmatrix} \quad , \quad \mathcal{Z} := \begin{bmatrix} Z_1 N_1 & \\ & Z_2 N_2 \end{bmatrix}.$$

3. Construct the matrices

$$\hat{C} := C\mathcal{V} \quad , \quad \hat{A} := \mathcal{Z}^T A \mathcal{V} \quad , \quad \hat{B} := \mathcal{Z}^T B.$$

and define the reduced transfer function

$$\hat{T}(s) := \hat{C}(sI_{2k} - \hat{A})^{-1} \hat{B}.$$

From Theorem 6.2, $\hat{T}(s)$ is a second order transfer function of order k that satisfies the interpolation conditions (6.25). The algorithm above has all of the freedom in the method of forming the bases and selecting interpolation points and their associated orders found in the Rational Krylov family of algorithms [48]. As a result, the second order rational interpolation problem can be solved while exploiting the sparsity of the matrices and parallelism of the computing platform in a similar fashion.

Second order Structure Preserving Tangential Interpolation

It is possible to generalize the earlier results for MIMO systems to perform tangential interpolation and preserve second order structure. This is accomplished by replacing Krylov subspaces at each interpolation point, λ_i, with *generalized* Krylov subspaces as done in [39]. The spaces are defined as follows:

Definition 6.3. *Let $M \in \mathbb{R}^{n\times n}$, $X \in \mathbb{R}^{n\times m}$, $y^{[i]} \in \mathbb{R}^m$, $i = 0, \ldots, k-1$ and define $Y \in \mathbb{R}^{km\times k}$ as*

$$Y = \begin{bmatrix} y^{[0]} & \ldots & y^{[k-1]} \\ & \ddots & \vdots \\ & & y^{[0]} \end{bmatrix}.$$

A generalized Krylov subspace of order k, denoted $\mathcal{K}_k(M, X, Y)$, is the image of the matrix $\left[\, X \; MX \; \ldots \; M^{k-1}X \,\right] Y$.

For example, by using Algorithm 6.4.2 below to compute bases for *generalized* Krylov subspaces and forming the appropriate projections, one can construct a second order transfer function $\hat{T}(s)$ of order n that satisfies the following interpolation conditions with respect to the second order transfer function $T(s)$ of order n:

$$x_i \left(T(s) - \hat{T}(s)\right) = O(\lambda_i - s) \quad , \quad \left(T(s) - \hat{T}(s)\right) x_{i+n} = O(\lambda_{i+n} - s), \tag{6.26}$$

where $x_1, \ldots, x_n \in \mathbb{R}^{1\times p}$ and $x_{n+1}, \ldots, x_{2n} \in \mathbb{R}^{m\times 1}$.

Algorithm 6.4.2 *1. Construct Z and V such that*

$$V = \left[\, (\lambda_{n+1} I_{2n} - A)^{-1} B x_{n+1} \; \ldots \; (\lambda_{2n} I_{2n} - A)^{-1} B x_{2n} \,\right] = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

$$Z^T = \begin{bmatrix} x_1 C(\lambda_1 I_{2n} - A)^{-1} \\ \vdots \\ x_n C(\lambda_n I_{2n} - A)^{-1} \end{bmatrix} = \left[\, Z_1^T \; Z_2^T \,\right],$$

where $Z_1, Z_2, V_1, V_2 \in \mathbb{R}^{n\times k}$. Choose the matrices $M_1, M_2, N_1, N_2 \in \mathbb{R}^{n\times n}$ such that $N_1^T Z_1^T V_1 M_1 = N_2^T Z_2^T V_2 M_2 = I_n$.

2. Construct

$$\mathcal{V} := \begin{bmatrix} V_1 M_1 & \\ & V_2 M_2 \end{bmatrix} \quad , \quad \mathcal{Z} := \begin{bmatrix} Z_1 N_1 & \\ & Z_2 N_2 \end{bmatrix}.$$

3. Construct the matrices

$$\hat{C} := C\mathcal{V} \quad , \quad \hat{A} := \mathcal{Z}^T A \mathcal{V} \quad , \quad \hat{B} := \mathcal{Z}^T B.$$

and define the reduced transfer function

$$\hat{T}(s) := \hat{C}(sI_{2n} - \hat{A})^{-1}\hat{B}.$$

It can be shown that $\hat{T}(s)$ is a second order transfer function of order n that satisfies the interpolation conditions (6.26) (see [39]).

It is also possible to impose higher order conditions while preserving the structure of the algorithm and the reduced order system. Consider, for instance, right tangential interpolation conditions of higher order (similar results hold for left tangential interpolation). Let the polynomial vector $x(s) := \sum_{i=0}^{k-1} x^{[i]}(s-\lambda)^i$. To impose the tangential interpolation condition

$$\Big(T(s) - \hat{T}(s)\Big)\, x(s) = O(s-\lambda)^k,$$

we construct $\hat{T}(s)$ as in Algorithm 6.4.2 using the generalized Krylov subspace $\mathcal{K}((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B, X)$ where X is formed from the $x^{[i]}$, $i = 0, \ldots, k-1$, i.e.,

$$Im\left\{ \left[\, (\lambda I - A)^{-1}B \ldots (\lambda I - A)^{-k}B \,\right] \begin{bmatrix} x^{[0]} & \ldots & x^{[k-1]} \\ & \ddots & \vdots \\ & & x^{[0]} \end{bmatrix} \right\} \subseteq Im\left\{ \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \right\}.$$

We refer to [39] for more details on this topic.

Real Transfer functions and Complex interpolation

First, we recall that for any pair of real matrices R_1, R_2, if

$$(R_1 + R_2 j)(S_1 + S_2 j) = I,$$

with S_1, S_2 real matrices, then $S_1 - S_2 j$ is the inverse of the matrix $R_1 - R_2 j$.

The second order transfer function that will be studied in the next section 6.5 is *real*, i.e. the matrices F^{out}, F^{in}, M, D and S are real. This implies that for any complex number λ,

$$T(\overline{\lambda}) = C(\overline{\lambda} I_{2n} - A)^{-1}B = C\overline{(\lambda I - A)^{-1}}B = \overline{T(\lambda)}. \tag{6.27}$$

Because the original transfer function is real, it makes sense to impose the reduced order transfer function $\hat{T}(s)$ to be real as well. In order to guarantee the matrices $\hat{F}^{out}, \hat{F}^{in}, \hat{M}, \hat{D}$ and $\hat{S}$ to be real, it is sufficient to use real projecting matrices $\mathcal{Z}, \mathcal{V}$ to construct them.

It is not difficult to see that the gramians of a real transfer function are real. This implies that if one applies the Second Order Balanced Truncation technique to a real second order transfer function, the projecting matrices will automatically be real as well, resulting in a real reduced order transfer function.

Some care must be taken concerning Krylov techniques. Indeed, if A, B are real matrices and $\lambda \in \mathbb{C}$, the matrix $K_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right)$ is in general complex.

Assume that $T(s)$ and $\hat{T}(s)$ are real transfer function. It follows that the error transfer function $E(s)$ is also real. From (6.27), if $\lambda \in \mathbb{C}$ is an interpolation point between $T(s)$ and $\hat{T}(s)$, then $\overline{\lambda}$ is also a zero of $E(s)$.

From the preceding discussion, it can recursively be proved that if A and B are real matrices of compatible dimension, then, for any complex number λ,

$$K_k\left((\overline{\lambda}I - A)^{-1}, (\overline{\lambda}I - A)^{-1}B\right) = \overline{K_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right)}.$$

To conclude, if one wants to construct a real transfer function $\hat{T}(s)$ that interpolates another real transfer function $T(s) = C(sI - A)^{-1}B$ at a complex number λ, $\hat{T}(s)$ must also interpolate $T(s)$ at $\overline{\lambda}$. This can be done using only real projecting matrices as follows.

1. First compute the Krylov matrix

$$V_1 + V_2 j := K_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right),$$

   with V_1 and V_2 real matrices.
2. Project with a right projecting matrix $\mathcal{V}$ such that

$$Im\left[V_1 \; V_2\right] \subseteq \mathcal{V}.$$

This ensures that $\mathcal{K}_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right)$ and $\mathcal{K}_k\left((\overline{\lambda}I - A)^{-1}, (\overline{\lambda}I - A)^{-1}B\right)$ are both subspaces of $\mathcal{V}$.

6.5 Numerical Experiments

In this section, model reduction techniques are applied to a large scale second order system representing the vibrating structure of a building. The objective is to compare the performance of second order structure preserving model reduction techniques, namely the SOBT technique introduced in Section 6.3 and the Second Order Krylov technique introduced in Section 6.4, with respect to the standard first order techniques, namely the Balanced Truncation and the Multipoint Padé techniques.

The characteristics of the second order system to be reduced are the following. The stiffness and mass matrix S and M are of dimension $n = 26394$. The mass matrix M is diagonal and the stiffness matrix S is symmetric and sparse (S contains approximately 2×10^5 non zero elements). The input vector is the transpose of the output vector:

$$F^{out} = {F^{in}}^T = \left[1 \ldots 1\right].$$

The damping matrix is *proportional*, meaning it is a linear combination of the mass matrix M and the stiffness matrix S:

$$D := \alpha M + \beta S.$$

The second order transfer function of Mc Millan degree $2n = 52788$ to be reduced is

$$T(s) := F^{out}(s^2M + sD + S)^{-1}F^{in} = F^{out}(s^2M + s(\alpha M + \beta S) + S)^{-1}F^{in}.$$

Given the structure of M we normalize by transforming the equation so that the mass matrix is the identity as follows :

$$T(s) = F^{out}M^{-1/2}\left(s^2I + s(\alpha I + \beta M^{-1/2}SM^{-1/2}) + M^{-1/2}SM^{-0.5}\right)^{-1} M^{-1/2}F^{in}.$$

So, we redefine $\bar{S} = M^{-1/2}SM^{-1/2}$ and $\bar{F}^{in} = M^{-1/2}F^{in} = \bar{F}^{out^T}$.

One intermediate system and five reduced order systems will be constructed from $T(s)$.

Three reasons led us to construct an intermediate transfer function. First, concerning the SVD techniques, it is not possible to apply the Balanced Truncation or the Second Order Balanced Truncation methods directly to the transfer function $T(s)$ because its Mc Millan degree $2n$ is too large for applying $O(n^3)$ algorithms. Second, the intermediate transfer function, assumed very close to $T(s)$, will also be used to approximate of the error bound between the different reduced transfer functions and the original transfer function $T(s)$. Finally, the intermediate transfer function will also be used in order to choose interpolation points for the Krylov techniques.

For these reasons, an intermediate second order transfer function of order 200 (i.e. of Mc Millan degree 400), called $T_{200}(s)$, is first constructed from $T(s)$ using Modal Approximation as follows. $T_{200}(s)$ is constructed by projecting $\bar{S}$ onto its eigenspace corresponding to its 200 eigenvalues of smallest magnitude. This corresponds to keeping the 400 eigenvalues of $s^2I + s(\alpha I + \beta\bar{S}) + \bar{S}$ the closest to the imaginary axis. Let $V_{f200} \in \mathbb{R}^{26364\times 200}$ be the projection matrix corresponding to the 200 eigenvalues of $\bar{S}$ the closest to the imaginary axis (with $V_p^TV_p = I_{200}$) (V_{f200} is computed with the Matlab function *eigs*). The intermediate transfer function is

$$\begin{aligned}
&T_{200}(s)\\
&:= \tilde{F}^{out}\left(s^2I_{200} + s\tilde{D} + \tilde{S}\right)^{-1}\tilde{F}^{in} = F^{out}M^{-1/2}V_p\\
&\quad\left(s^2I + s(\alpha I + \beta V_p^TM^{-1/2}SM^{-1/2}V_p) + V_p^TM^{-1/2}SM^{-1/2}V_p\right)^{-1} V_p^TM^{-1/2}F^{in}.
\end{aligned}$$

By checking the difference between $T(s)$ and $T_{200}(s)$ at different points in the complex plane, it has been verified that both transfer functions are very close to each other. The Hankel singular values of $T_{200}(s)$ are shown in Figure 6.1.

Let us construct the reduced order transfer functions.

First, from $T_{200}(s)$, we compute the reduced transfer function of Mc Millan degree 20 obtained by using balanced truncation (with the *sysred* Matlab function of the SLICOT library [15]), called $T_{bt}(s)$. Note that $T_{bt}(s)$ is no longer in

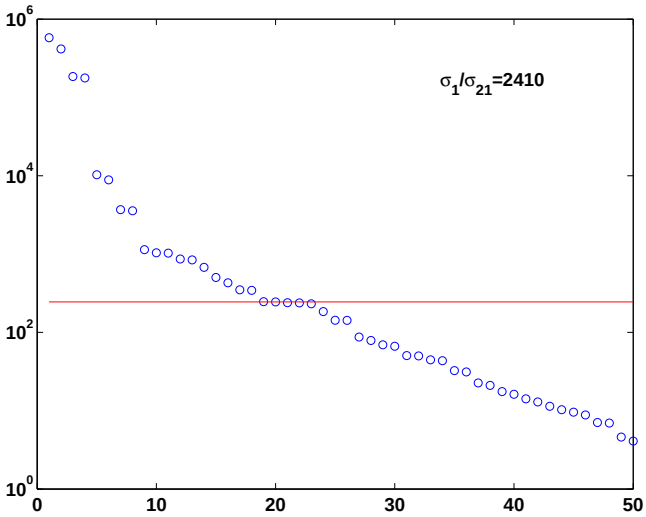


Fig. 6.1. Hankel singular values of $T_{200}(s)$

second order form. Another second order transfer function of order 20 (and Mc Millan degree 40), called $T_{sobt}(s)$, is constructed from $T_{200}(s)$ using the SOBT algorithm [21].

Concerning the Krylov techniques, the reduced order transfer functions are computed directly from the original transfer function $T(s)$. Three reduced order systems are compared. The first one is constructed using the standard first order Krylov procedure. The two other reduced systems (corresponding to different choices of interpolation points) are constructed using a second order Krylov technique.

In order to apply Krylov techniques, a first important step consists in choosing the interpolation points. Indeed, the quality of the reduced order system is very sensitive to the choice of interpolation points.

An interesting fact is that there are 42 interpolation points between $T_{200}(s)$ and $T_{bt}(s)$ that have a positive real part (among the 420 zeros of $T_{200}(s) - T_{bt}(s)$). From a dozen experiments, it has been observed that when using the standard Balanced Truncation technique, the number of interpolation points in the right-half plane between the original and the reduced transfer function is roughly equal to twice the Mc Millan degree of the reduced transfer function. The interpolation points in the right-half plane have the advantage that they are neither close to the poles of the system to be reduced nor to the poles of the Balanced Truncation reduced system because both transfer functions are stable. This implies that both transfer functions do not vary too much there and this is preferable in order to avoid numerical instability.

Because the Mc Millan degree of $T_{bt}(s)$ is equal to 20, it is well known that 40 points are sufficient in order to describe $T_{bt}(s)$. In other words, the only transfer function of Mc Millan degree smaller than 20 that interpolates $T_{bt}(s)$ at 40 points in the complex plane is $T_{bt}(s)$ itself [38]. So, we take the 40 interpolation points between $T_{200}(s)$ and $T_{bt}(s)$ with largest real part as our choice for computing the transfer function of Mc Millan degree 20, denoted $T_{Kryl}(s)$, that interpolates the original transfer function $T(s)$ at these points. The poles and interpolation points are shown in Figure 6.2. Because $T_{200}(s)$ is very close to $T(s)$, $T_{Kryl}(s)$ should be close to $T_{bt}(s)$. Using the second

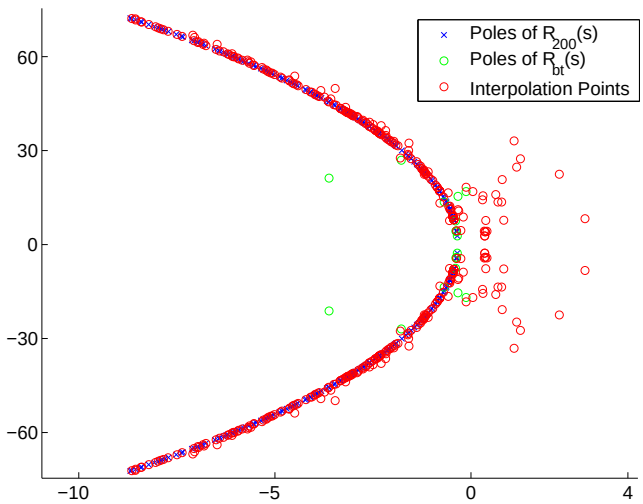


Fig. 6.2. Poles and interpolation points for $T_{200}(s)$ and $T_{bt}(s)$

order Krylov technique, a reduced second order transfer function $T_{sokryl}(s)$ of Mc Millan degree 28 is also constructed. Its Mc Millan degree was first chosen to be 20 but the resulting reduced transfer function was not stable. For this reason, additional interpolation conditions were added until the reduced transfer function was stable, resulting in a Mc Millan degree equal to 28. The transfer function $T_{sokryl}(s)$ interpolates $T(s)$ at the 28 rightmost interpolation points between $T_{200}(s)$ and $T_{bt}(s)$. These 28 interpolation points are 14 pairs of complex conjugated numbers. The Krylov subspaces are computed for only one half of interpolation points, say those with positive complex part and the real projecting matrices are obtained by taking both the real and complex part of the Krylov subspaces, following the scheme explained in Subsection 6.4.

For comparison purposes a set of interpolation points randomly generated (with symmetry with respect to the real axis in order to obtain a real interpolating transfer function) in a rectangle delimited by the extreme zeros in the left half plane of $T_{200}(s) - T_{bt}(s)$ is also used in the second order Krylov method to generate $T_{randsokryl}(s)$. These two sets of interpolation points are shown in Figure 6.3. The transfer functions $T_{200}(s)$, $T_{bt}(s)$, $T_{sobt}(s)$, $T_{randsokryl}(s)$, $T_{kryl}(s)$

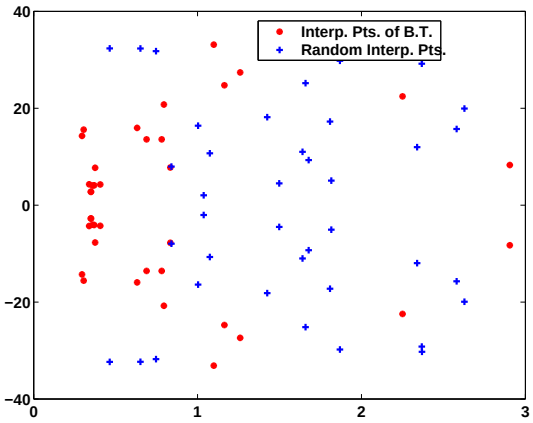


Fig. 6.3. Interpolation points for $T_{bt}(s)$, $T_{sokryl}(s)$ and $T_{randsokryl}(s)$

and $T_{sokryl}(s)$ are plotted in Figure 6.4. Recall, that $T_{200}(s)$ is used here as computationally tractable approximation of $T(s)$. More can be learned by considering the the H_∞-norm errors relative to $\|T_{200}(s)\|_\infty$ shown in Table 6.1.

Table 6.1. Relative errors for reduced order models

Reduced Transfer function	Model Reduction technique	Mc Millan degree	$\frac{\|T_{200}(s)-T_{reduced}(s)\|_\infty}{\|T_{200}(s)\|_\infty}$
$T_{bt}(s)$	Balanced Truncation	20	$4.3\ 10^{-4}$
$T_{sobt}(s)$	Second Order Balanced Truncation	40	$2.6\ 10^{-4}$
$T_{kryl}(s)$	Krylov	20	$8.3\ 10^{-4}$
$T_{sokryl}(s)$	Second Order Krylov	28	$5.8\ 10^{-2}$
$T_{randsokryl}(s)$	Random Second Order Krylov	20	$7\ 10^{-2}$

As a first observation, it looks as if the six transfer functions are close to each other, especially for frequencies smaller than 10 rad/sec (where the bode mag-

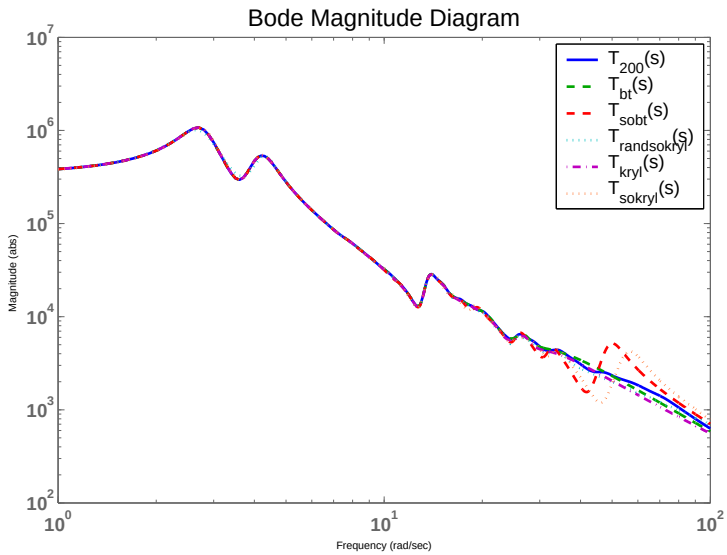


Fig. 6.4. The six transfer functions

nitude diagrams are undistinguishable, see Figure 6.4). This is a good news because they should all approximate well the same transfer function $T(s)$.

One observes from Table 6.1 that the SVD techniques perform better than the Krylov techniques. Two remarks are in order. First, it should be kept in mind that only the Krylov reduced transfer functions are directly computed from the original data of $T(s)$. Second, concerning the Krylov techniques, the quality of the approximation depends strongly on the choice of the interpolation points. Because for SISO systems, any transfer function can be constructed from Krylov subspaces from any transfer function of larger Mc Millan degree, there should exist interpolation conditions that produce reduced order transfer functions with smaller error bound than what can be obtained with balanced techniques, but of course, it is not easy to find such interpolation conditions.

A surprising fact concerning SVD techniques is that the best approximation is obtained with $T_{sobt}(s)$ and not $T_{bt}(s)$. Nevertheless, one should not forget that the Mc Millan degree of $T_{sobt}(s)$ is twice as large as the Mc Millan degree of $T_{bt}(s)$.

In contrast with SVD techniques, the error obtained with the first order transfer function $T_{kryl}(s)$ is 100 times smaller than for the second order transfer functions $T_{sokryl}(s)$ and $T_{randsokryl}(s)$. This tends to indicate that Second Order Krylov techniques perform quite poorly compared to the first order techniques, perhaps indicating that a more sophisticated algorithm for choosing the interpolation points for these methods is needed.

Finally, by choosing random interpolation points, the error remains roughly the same than by taking the *balanced truncation* interpolation points: 0.058 for $T_{sokryl}(s)$ and 0.07 for $T_{randsokryl}(s)$. This is probably due to the fact that the area chosen to generate the interpolation points for $T_{randsokryl}(s)$ contains good information about the original transfer function.

6.6 Concluding Remarks

Concerning the second order Krylov technique, the following observation is worth mentioning. For SISO systems, it has been shown in [17] and [66] that for every first order system (c, A, b) of pair Mc Millan degree such that $cb = 0$, there exists a state space transformation that puts it into a second order form. In other words, every SISO system (with first Markov parameter equal to zero) can be rewritten as a second order system. This implies that in the SISO case, it is possible to impose $4n - 1$ interpolation conditions for a reduced second order system of Mc Millan degree $2n$ by first using the classic Multipoint Padé technique of Theorem 3.1 and then reconstructing a second order form with an appropriate state space coordinate transformation. Currently, no proof is available for the MIMO case. If it is sufficient to impose the first Markov parameter to be zero in order to ensure the second order property, one can wonder if the second order structure preserving schemes presented here are useful. Here are some reasons why we still believe that our approach is interesting.

1. First, it appears that some of our second order model reduction techniques are as good as, or even better than existing standard techniques. For instance, in Table 6.1, the smallest error bound is obtained for the reduced order system obtained using the SOBT algorithm, and not the standard Balanced Truncation algorithm. If the Mc Millan degree of $T_{sobt}(s)$ is twice larger than the Mc Millan degree of $T_{bt}(s)$, its complexity is roughly the same, due to its preserved second order structure.
2. As we will see in the next chapter, the second order structure preserving techniques presented have been extended to more general structured systems, permitting to unify several structure preserving model reduction methods existing in the literature.
3. From a practical point of view, several softwares in mechanical engineering need a system written in second order form. It is then simpler to provide a system directly in second order form.
4. Another important point that has not been discussed yet is that other important characteristics of the original system, such as symmetry and passivity, can be maintained by projecting with block diagonal matrices (by projecting with equal left and right projecting matrices).

As for generalized state space realizations of first order systems, it is also possible to apply Krylov technique to second order systems without requiring the

mass matrix M to be equal to the identity. Concerning the SOBT technique, special care must taken in deriving the second order gramians.

For second order balanced truncation, numerical results are very encouraging, but many important questions remain open. For instance, does there exist an a priori global error bound with SOBT, as for Balanced Truncation? Even simpler, is stability of the reduced system always guaranteed? If the answer to the preceding questions is negative, does there exist a better choice of second order gramians? Also, the development of an approximate version applicable to large scale systems is needed.

As we will see in Chapter 7, there is a natural extension of the second order structure preserving techniques presented in this chapter for the general framework of *interconnected systems*.

6.7 Notes and References

This chapter is a resume of three papers.

The first paper [21] (see also [22]) is the result of a team work with Y. Chahlaoui, D. Lemonnier and Paul Van Dooren. We would like also to than Karl Meerbergen for a preliminary version of that work [23].

The second-order preserving Krylov techniques have been studied in [83].

A more recent work involving Y. Chahlaoui, K. Gallivan and P. Van Dooren is the survey paper [20].

It should be pointed out that several authors have recently considered the problem of using Krylov techniques while preserving the second order form. See for instance the work of Bai [10, 9, 12], Freund [11, 32] and others [74].

This is in contrast with second order balanced truncation techniques that, to our knowledge, have only been considered in [66]. See also [78].

7

Model Reduction of Interconnected Systems

Large scale linear systems have often a particular structure. In this chapter, we consider a particular class of structured systems that can be modelled as a set of input/output subsystems that interconnect to each other, in the sense that outputs of some subsystems are inputs of other subsystems. Sometimes, it is important to preserve this structure into the reduced order system. Instead of reducing the entire system as a black box, it makes sense to reduce each subsystem (or a few of them) by taking into account its interconnection with the other subsystems in order to approximate the entire system. The purpose of this chapter is to present model reduction techniques that preserve the structure of the interconnections. With our approach, several structured model reduction techniques existing in the literature appear as special cases of our methods, permitting to unify the theory.

7.1 Introduction

A preliminary remark is worth mentioning. As for the rest of this thesis, we only consider here systems as input/output systems. This is not always the best choice for representing systems, especially in the context of complex interconnected systems. Indeed, inputs and outputs do not always naturally appear in the description of a system, and it is not always possible to model a structured system as it is suggested in this chapter. A more general definition of interconnected systems appears in the behavioral approach [70]. Nevertheless, very interesting examples of structured systems can be modelled as proposed here.

Well known input/output structured model reduction problems are the problems of weighted model reduction, controller reduction and second order model reduction. For each of these structured problems, specialized model reduction techniques have been developed. It turns out that all these structured problems can be modelled as examples of *interconnected* systems defined below.

Here is the description of the *interconnected system* considered in this chapter. Let us call this large scale linear system $G(s)$. This system is composed of an interconnection of k sub-systems $T_i(s)$. Each subsystem is assumed to be a linear MIMO transfer function. Subsystem $T_j(s)$ has α_j inputs denoted by the vector a_j and β_j outputs denoted by the vector b_j :

$$b_i(s) = T_i(s)a_i(s). \tag{7.1}$$

Define $\alpha := \sum_{i=1}^{k} \alpha_i$ and $\beta := \sum_{j=1}^{k} \beta_j$. The inputs of each subsystem are either outputs of other subsystems or external input that do not depend on the other subsystems.

One can rewrite a transfer function from its subsystems via the use of an "interconnection matrix"

$$a_i(s) = u_i(s) + \sum_{j=1}^{k} K_{i,j} b_j(s). \tag{7.2}$$

Sometimes it is preferable to define the external output $u_i(s)$ as a linear combination of a global external output $u(s)$. This is written as $u_i(s) = H_i u(s)$, where $H_i \in \mathbb{R}^{\alpha_i \times m}$. The output of $G(s)$, denoted by $y(s)$ is a linear function of the outputs of the subsystems:

$$y(s) := \sum_{i=1}^{k} F_i b_i(s),$$

with $F_i \in \mathbb{R}^{p \times \beta_i}$. Define

$$a(s) := \begin{bmatrix} a_1(s)^T \ \dots \ a_k(s)^T \end{bmatrix}^T, \qquad b(s) := \begin{bmatrix} b_1(s)^T \ \dots \ b_k(s)^T \end{bmatrix}^T,$$

$$T(s) := \begin{bmatrix} T_1(s) & & \\ & \ddots & \\ & & T_k(s) \end{bmatrix}, H := \begin{bmatrix} H_1 \\ \vdots \\ H_k \end{bmatrix}, F = \begin{bmatrix} F_1 \ \dots \ F_k \end{bmatrix}.$$

and finally the connectivity matrix K as follows

$$K := \begin{bmatrix} K_{1,1} & \dots & K_{1,k} \\ \vdots & \ddots & \vdots \\ K_{k,1} & \dots & K_{k,k} \end{bmatrix}. \tag{7.3}$$

The Mc Millan degree of $T_i(s)$ is n_i and (A_i, B_i, C_i, D_i) is a minimal state space realization of $T_i(s)$. From these definitions, $T(s) = C(sI - A)^{-1}B + D$ with

$$C := \begin{bmatrix} C_1 & & \\ & \ddots & \\ & & C_k \end{bmatrix}, A := \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix},$$
$$B := \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{bmatrix}, D := \begin{bmatrix} D_1 & & \\ & \ddots & \\ & & D_k \end{bmatrix}.$$

The preceding equations can be rewritten as follows :

$$a(s) = Hu(s) + Kb(s), \quad b(s) = T(s)a(s), \quad y(s) = Fb(s). \tag{7.4}$$

From this,

$$y(s) = F(I - T(s)K)^{-1}T(s)Hu(s). \tag{7.5}$$

In others words, $G(s) = F(I - T(s)K)^{-1}T(s)H$, with $F \in \mathbb{R}^{p\times\beta}$. Hence, a state space realization of $G(s)$ is given by (A_G, B_G, C_G, D_G) defined by (see for instance [94], pg 66)

$$\begin{aligned} C_G &:= F(I - DK)^{-1}C \ , A_G := A + BK(I - DK)^{-1}C, \\ B_G &:= B(I - KD)^{-1}H \ , D_G := FD(I - KD)^{-1}H. \end{aligned} \tag{7.6}$$

If all the transfer functions are strictly proper, i.e. $D = 0$, the state space realization (7.6) of $G(s)$ is simpler :

$$C_G = FC, \quad A_G = A + BKC, \quad B_G = BH, \quad D_G = 0.$$

Let us finally remark that if all systems are in parallel, i.e. $K = 0$, then $G(s) = FT(s)H$.

Figure 7.1 gives an example of an interconnected system $G(s)$ composed of three subsystems. The problem of *interconnected systems model reduction* proposed here consists in reducing the subsystems $T_i(s)$ in order to approximate the global mapping from $u(s)$ to $y(s)$ and not the internal mappings from $a_i(s)$ to $b_i(s)$.

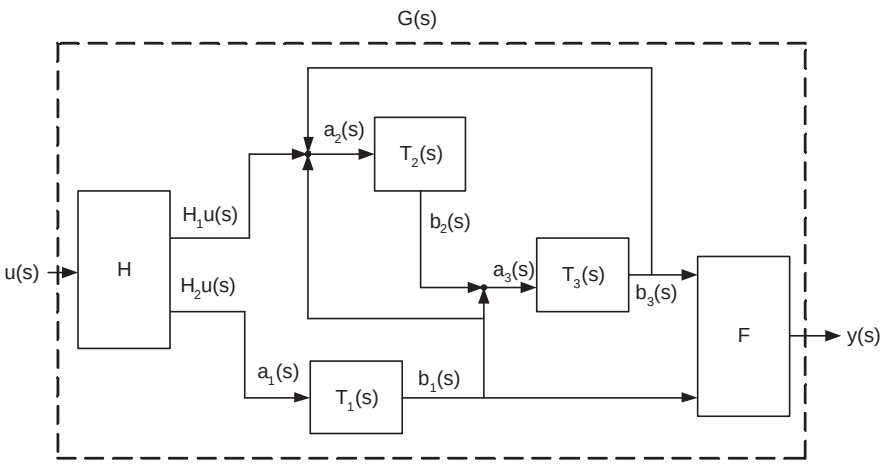
This chapter is organized as follows. After some preliminary results, a Balanced Truncation framework for interconnected systems is derived in Section 7.2. Krylov model reduction techniques for interconnected systems are presented in Section 7.3. In Section 7.4, several connections with existing model reduction techniques for structured systems are given. Concluding remarks are made in Section 7.5.

7.2 Interconnected Systems Balanced Truncation

This section is an extension of Section 6.3. For an introduction on Balanced Truncation, see Section 2.4 (see also [66, 21]).

Let us start with some basic lemmas.

Fig. 7.1. Example of interconnected system



Lemma 7.1. *Let $x_i \in \mathbb{R}^{n_i}$ and $M_{i,j} \in \mathbb{R}^{n_i \times n_j}$ for $1 \leq i \leq k$. Define*

$$x := \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \quad , \quad M := \begin{bmatrix} M_{1,1} & \dots & M_{1,k} \\ \vdots & \ddots & \vdots \\ M_{k,1} & \dots & M_{k,k} \end{bmatrix}.$$

Assume that the matrix M is positive definite. Let us consider the product

$$J(x, M) := x^T M^{-1} x.$$

Then, for any fixed $x_i \in \mathbb{R}^{n_i \times n_i}$,

$$\min_{x_j, j \neq i} J(x, M) = x_i^T M_{i,i}^{-1} x_i.$$

Proof. Without loss of generality, let us assume that $i = 1$. For ease of notation, define $y := \begin{bmatrix} x_2^T & \dots & x_k^T \end{bmatrix}^T$ and $\begin{bmatrix} N_{1,1} & N_{1,2} \\ N_{1,2}^T & N_{2,2} \end{bmatrix} := M^{-1}$ with $N_{1,1} \in \mathbb{R}^{n_1 \times n_1}$. We obtain the following expression

$$J(x, M) = x_1^T N_{1,1} x_1 + 2 x_1^T N_{1,2} y + y^T N_{2,2} y. \tag{7.7}$$

Because M is positive definite, N is also positive definite. This implies that $N_{1,1}$ and $N_{2,2}$ are positive definite. $J(x, M)$ is a quadratic form and the Hessian of $J(x, M)$ with respect to y is equal to $N_{2,2}$. This implies that the minimum is obtained by annihilating the gradient :

$$\frac{\partial J(x,M)}{\partial y} = 2N_{1,2}^T x_1 + 2N_{2,2} y.$$

The minimum is obtained for $y^* = -N_{2,2}^{-1} N_{1,2}^T x_1$, which yields the optimal value

$$\min_y J(x,M) = x_1^T N_{1,1} x_1 - x_1^T N_{1,2} N_{2,2}^{-1} N_{1,2}^T x_1 = x_1^T M_{1,1}^{-1} x_1,$$

where the last equality was obtained by using the Schur complement.

□

Another optimization result is often used in the context of weighted model reduction. Instead of finding the minimum of J with respect to the other variables, one might be interested in finding the value of J by putting the other states equal to zero. This gives rise to the following result :

Lemma 7.2. *Let $x_i \in \mathbb{R}^{n_i}$ and $M_{i,j} \in \mathbb{R}^{n_i \times n_j}$ for $1 \le i \le 2$. Define*

$$x := \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad , \quad M := \begin{bmatrix} M_{1,1} & M_{1,2} \\ M_{2,1} & M_{2,2} \end{bmatrix}.$$

Assume that the matrix M is positive definite. Let us consider the product

$$J(x,M) := x^T M^{-1} x.$$

Then, for any fixed $x_i \in \mathbb{R}^{n_i \times n_i}$,

$$J(x,M)_{x_j=0, j\neq i} = x_i^T \left(M_{i,i} - M_{i,j} M_{j,j}^{-1} M_{j,i} \right)^{-1} x_i.$$

Proof. The proof consists in rewriting the inverse of M by using the Schur Complement Formula (1.1).

□

The generalization to k different states is obvious.

Let us consider the controllability and observability gramians of $G(s)$:

$$A_G P_G + P_G A_G^T + B_G B_G^T = 0, \quad A_G^T Q_G + Q_G A_G + C_G^T C_G = 0. \tag{7.8}$$

Let us decompose

$$P_G = \begin{bmatrix} P_{1,1} & \dots & P_{1,k} \\ \vdots & \ddots & \vdots \\ P_{k,1} & \dots & P_{k,k} \end{bmatrix} \quad , \quad Q_G = \begin{bmatrix} Q_{1,1} & \dots & Q_{1,k} \\ \vdots & \ddots & \vdots \\ Q_{k,1} & \dots & Q_{k,k} \end{bmatrix},$$

where $P_{i,j} \in \mathbb{R}^{n_i \times n_j}$. If we perform a state space transformation Φ_i to the state $\bar{x}_i(t) = \Phi_i x_i(t)$ of each interconnected transfer function $T_i(s)$, we actually perform a state space transformation

$$\Phi := \begin{bmatrix} \Phi_1 & & \\ & \ddots & \\ & & \Phi_k \end{bmatrix}$$

to the realization $(\bar{A}, \bar{B}, \bar{C}, \bar{D}) = (\Phi A \Phi^{-1}, \Phi B, C\Phi^{-1}, D)$ of $T(s)$. This, in turn implies that $(\bar{A}_G, \bar{B}_G, \bar{C}_G, \bar{D}_G) = (\Phi A_G \Phi^{-1}, \Phi B_G, C_G \Phi^{-1}, D_G)$. From this,

$$(\bar{P}_G, \bar{Q}_G) = (\Phi P_G \Phi^T, \Phi^{-T} Q_G \Phi^{-1}),$$

i.e. they also perform a contragradient transformation. This implies that $(\bar{P}_{i,i}, \bar{Q}_{i,i}) = (\Phi_i P_{i,i} \Phi_i^T, \Phi_i^{-T} Q_{i,i} \Phi_i^{-1})$, which is a contra-gradient transformation that only depends on the state space transformation on x_i, i.e. on the state space associated to $T_i(s)$.

Let us recall (see (2.18)) that the minimal past energy necessary to reach $x_i(0) = x_i$ for each $1 \leq i \leq k$ with the pair (A_G, B_G) is given by the expression

$$\begin{bmatrix} x_1^T & \dots & x_k^T \end{bmatrix} P_G^{-1} \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}.$$

The following result is a consequence of Lemma 7.1.

Lemma 7.3. *With the preceding notation, the minimal past input energy*

$$J := \int_{-\infty}^{0} u(t)^T u(t) dt$$

to apply to the interconnected transfer function $G(s)$ in order that for subsystem i at time $t = 0$, $x_i(0) = \bar{x}_i$ over all initial input condition $x_j(0), j \neq i$, is

$$\bar{x}_i^T P_{i,i}^{-1} \bar{x}_i.$$

Moreover, the minimal input energy necessary in order that for subsystem i at time $t = 0$, $x_i(0) = \bar{x}_i$ and that for all the other subsystems, $x_j(0) = 0, j \neq i,$, is

$$\bar{x}_i^T (P_G^{-1})_{i,i} \bar{x}_i,$$

where $(P_G^{-1})_{i,i}$ is the i, i block of the inverse of P_G, and this block is equal to the inverse of the Schur Complement of $P_{i,i}$.

Finally,

$$0 < P_{i,i}^{-1} \leq (P_G^{-1})_{i,i}. \tag{7.9}$$

Proof. The two first results are direct consequences of Lemma 7.1. The inequality (7.9) follows from the Schur complement but a more intuitive alternative proof is the following. For any nonzero vector $\bar{x}_i$, the energy necessary for subsystem i at time $t = 0$ to reach $x_i(0) = \bar{x}_i$ over all initial input condition

$x_j(0), j \neq i$, cannot be larger than by imposing $x_j(0) = 0, j \neq i$. This implies that for any nonzero vector x,

$$x^T \left((P_G^{-1})_{i,i} - P_{i,i}^{-1} \right) x \geq 0.$$

□

Similar energy interpretations hold for the diagonal blocks of the observability matrix Q_G and of its inverse.

From Lemma 7.3, it makes sense to truncate the part of the state x_i of each subsystem $T_i(s)$ corresponding to the smallest eigenvalues of the product $P_{i,i}Q_{i,i}$ (other choices of gramians are possible, depending on the chosen energy optimization problem). We can thus perform a bloc diagonal transformation in order to make the gramians $P_{i,i}$ and $Q_{i,i}$ both equal and diagonal :$P_{i,i} = Q_{i,i} = \Sigma_i$. Then, we can truncate each subsystem $T_i(s)$ by deleting the states corresponding to the smallest eigenvalues of Σ_i.

To resume, our *Interconnected Systems Balanced Truncation* (in short ISBT) algorithm, proceeds as follows.

Algorithm 7.2.1 *Let $(C_G, A_G, B_G, D_G) \sim G(s)$, where $G(s)$ is an interconnection of k subsystems $(C_i, A_i, B_i, D_i) \sim T_i(s)$ of order n_i. In order to construct a reduced order system $\hat{G}(s)$ while preserving the interconnections, perform as follows.*

1. *Compute the gramians P_G and Q_G satisfying* (7.8).
2. *For each subsystem $T_i(s)$, perform the contragradient transformation Φ_i in order to make the gramians $P_{i,i}$ and $Q_{i,i}$ equal and diagonal.*
3. *For each subsystem (C_i, A_i, B_i, D_i), keep only the state corresponding to the largest eigenvalues of $P_{i,i} = Q_{i,i} = \Sigma_i$, giving the reduced subsystems $\hat{T}_i(s)$.*
4. *Define*

$$\hat{G}(s) = F(I - \hat{T}(s)K)^{-1}T(s)H, \quad \textit{with} \quad \hat{T}(s) := diag\{\hat{T}_i(s)\}.$$

Remark 7.4. A variant of Algorithm 7.2.1 consists in performing a *balance and truncate* procedure for each subsystem $T_i(s)$ with respect to the Schur complements of $P_{i,i}$ and $Q_{i,i}$ instead of $P_{i,i}$ and $Q_{i,i}$. From Lemma 7.3, this corresponds to sorting the state-space of each system C_i, A_i, B_i with respect to the optimization problem $\min_u \parallel u(t) \parallel^2$ such that $x_i(0) = x_0$ and $x_j = 0$ for $j \neq i$. Mixed strategies are also possible.

7.3 Krylov techniques for interconnected systems

Krylov techniques have already been considered in the literature for particular cases of structured systems. See for instance [77] in the controller reduction framework, or [80] in the second-order model reduction framework. This last

case has been revisited recently in [32] and [83] (cf. Section 6.4). But, to our knowledge, it is the first time they are studied in the general framework of *Interconnected Systems*.

The problem is the following. If one projects the state-space realizations (C_i, A_i, B_i) of the interconnected transfer functions $T_i(s)$ with projecting matrices Z_i, V_i containing Krylov subspaces, giving rise to reduced-order transfer functions $\hat{T}_i(s)$ that satisfy interpolation conditions with respect to $T_i(s)$, what are the resulting relations between $\hat{G}(s)$ and $G(s)$?

If one imposes the same interpolation conditions for every pair of subsystems $T_i(s)$ and $\hat{T}_i(s)$, then the same interpolation conditions hold between the block diagonal transfer functions $T(s)$ and $\hat{T}(s)$ as well. Let us investigate what this implies for $G(s)$ and $\hat{G}(s)$. Let us assume that

$$(\hat{C}, \hat{A}, \hat{B}) = (CV, Z^T AV, Z^T B)$$

such that $Z^T V = I$ and

$$\mathcal{K}_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1} B\right) \subseteq Im(V).$$

From Theorem 2.4, $\hat{T}(s) := \hat{C}(sI - \hat{A})^{-1}\hat{B}$ interpolates $T(s)$ at $s = \lambda$ up to the k first derivatives. Concerning $G(s)$, the matrices F, K, D and H are unchanged. One obtains

$$\hat{G}(s) = C_G V(sI - Z^T A_G V)^{-1} Z^T B_G + D_G.$$

At first sight, there seem to be no reason why V should contain the subspace

$$\mathcal{K}_k\left((\lambda I - A_G)^{-1}, (\lambda I - A_G)^{-1} B_G\right),$$

except for the case of interpolation at infinity. Indeed, it can easily be proven recursively that

$$\mathcal{K}_k\left(A + BK(I - DK)^{-1}C, B(I - KD)^{-1}H\right) \subseteq \mathcal{K}_k\left(A, B\right).$$

It turns out that the preceding result holds for arbitrary points in the complex plane, as shown in the following lemma.

Lemma 7.5. *Let $\lambda \in \mathbb{C}$ be a point that is neither a pole of A nor a pole of A_G (defined in* (7.6)*). Then*

- $$\begin{aligned} &\mathcal{K}_k\left((\lambda I - A_G)^{-1}, (\lambda I - A_G)^{-1} B_G\right) \\ &\subseteq \mathcal{K}_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1} B\right), \end{aligned} \tag{7.10}$$

- $$\begin{aligned} &\mathcal{K}_k\left((\lambda I - A_G)^{-T}, (\lambda I - A_G)^{-T} C_G^T\right) \\ &\subseteq \mathcal{K}_k\left((\lambda I - A)^{-T}, (\lambda I - A)^{-T} C^T\right). \end{aligned} \tag{7.11}$$

Proof. Only (7.10) will be proven. An analog proof can be given for (7.11). First, let us prove that the image of $(\lambda I - A_G)^{-1}B_G$ is included in the image of $(\lambda I - A)^{-1}B$. In order to simplify the notation, let us define the matrix $X := K(I - DK)^{-1}C$. From the preceding definitions,

$$\begin{aligned}
&(\lambda I - A_G)^{-1}B_G \\
&= (\lambda I - A - BX)^{-1} B(I - KD)^{-1}H \\
&= \left(I - (\lambda I - A)^{-1}BX\right)^{-1} (\lambda I - A)^{-1}B(I - KD)^{-1}H \\
&= \sum_{i=0}^{\infty} \left((\lambda I - A)^{-1}BX\right)^{i} (\lambda I - A)^{-1}B(I - KD)^{-1}H \\
&= (\lambda I - A)^{-1}B \left(I - X(\lambda I - A)^{-1}B\right)^{-1} (I - KD)^{-1}H.
\end{aligned}$$

This clearly implies that the image of $(\lambda I - A_G)^{-1}B_G$ is included in the image of $(\lambda I - A)^{-1}B$. Let us assume that

$$\mathcal{K}_{k-1}\left((\lambda I - A_G)^{-1}, (\lambda I - A_G)^{-1}B_G\right) \subseteq \mathcal{K}_{k-1}\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right),$$

and prove that this implies that

$$\mathcal{K}_{k}\left((\lambda I - A_G)^{-1}, (\lambda I - A_G)^{-1}B_G\right) \subseteq \mathcal{K}_{k}\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right). \quad (7.12)$$

Because $(\lambda I - A_G)^{-k+1}B_G$ belongs to $\mathcal{K}_{k-1}\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right)$, there exists a matrix Y such that

$$(\lambda I - A_G)^{-k+1}B_G = K_{k-1}\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right) Y.$$

One obtains then that

$$\begin{aligned}
&(\lambda I - A_G)^{-k}B_G \\
&= (\lambda I - A_G)^{-1}(\lambda I - A_G)^{-k+1}B_G \\
&= \sum_{i=0}^{\infty} \left((\lambda I - A)^{-1}BX\right)^{i} (\lambda I - A)^{-1}K_{k-1}\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right) Y.
\end{aligned}$$

Note that

$$\begin{aligned}
&(\lambda I - A)^{-1}K_{k-1}\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right) \\
&= K_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right) \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & \ddots & 0 \\ & & & 1 \end{bmatrix}.
\end{aligned}$$

Moreover, for any natural number $i > 0$, it is clear that

$$\left((\lambda I - A)^{-1}BX\right)^i \in \mathcal{K}_1\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right).$$

This proves that (7.12) is satisfied.

□

Thanks to the preceding lemma, there are at least two ways to project the subsystems $T_i(s)$ in order to satisfy a set of interpolation conditions using Krylov subspaces as follows.

Lemma 7.6. *Let $\lambda \in \mathbb{C}$ be neither a pole of $T(s)$ nor a pole of $G(s)$. Define*

$$V \in \mathbb{C}^{n\times r} := \left[\, V_1^T \ \dots \ V_k^T \,\right]^T,$$

such that $V_i \in \mathbb{C}^{n_i\times r}$. Assume that either

$$\mathcal{K}_k\left((\lambda I - A_G)^{-1}, (\lambda I - A_G)^{-1}B_G\right) \subseteq Im(V). \tag{7.13}$$

or

$$\mathcal{K}_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right) \subseteq Im(V). \tag{7.14}$$

Construct left projecting matrices $Z_i \in \mathbb{C}^{n_i\times r}$ such that $Z_i^T V_i = I_r$. Project each subsystem as follows :

$$(\hat{C}_i, \hat{A}_i, \hat{B}_i) := (C_i V_i, Z_i^T A_i V_i, Z_i^T B_i).$$

Then, $\hat{G}(s)$ interpolates $G(s)$ at λ up to the first k derivatives.

Proof. As a consequence of Lemma 7.5, first note that (7.14) implies (7.13). Let us assume that (7.13) is satisfied. The preceding operation corresponds to projecting C_G, A_G, B_G with

$$\mathcal{Z} := \begin{bmatrix} Z_1 & & \\ & \ddots & \\ & & Z_k \end{bmatrix} \quad , \quad \mathcal{V} := \begin{bmatrix} V_1 & & \\ & \ddots & \\ & & V_k \end{bmatrix}.$$

This implies that $\mathcal{Z}^T\mathcal{V} = I$ and $Im(V) \subseteq Im(\mathcal{V})$. This concludes the proof.

□

In some contexts, such as controller reduction or in the presence of weighting functions, one does not construct a reduced order transfer function $\hat{G}(s)$ by projecting the state spaces of all the subsystems (C_i, A_i, B_i) but one only project some of one subsystem. Let us consider this last possibility.

Corollary 7.7. *Define*

$$V \in \mathbb{C}^{n\times r} := \left[\, V_1^T \ \dots \ V_k^T \,\right]^T,$$

such that $V_i \in \mathbb{C}^{n_i \times r}$. *Assume that either*

$$\mathcal{K}_k\left((\lambda I - A)^{-1}, (\lambda I - A)^{-1}B\right) \subseteq Im(V),$$

or

$$\mathcal{K}_k\left((\lambda I - A_G)^{-1}, (\lambda I - A_G)^{-1}B_G\right) \subseteq Im(V).$$

Construct a reduced order transfer function $\hat{G}(s)$ *by only projecting one subsystem, say* (C_i, A_i, B_i), *as follows. Let* $Z_i \in \mathbb{C}^{n_i \times r}$ *such that* $Z_i^T V_i = I_r$. *Project subsystem* (C_i, A_i, B_i) *as follows :*

$$(\hat{C}_i, \hat{A}_i, \hat{B}_i) := (C_i V_i, Z_i^T A_i V_i, Z_i^T B_i), \tag{7.15}$$

and keep all the other subsystems unchanged. Then, $\hat{G}(s)$ *interpolates* $G(s)$ *at* λ *up to the first* k *derivatives.*

Proof. Again, note that (7.14) implies (7.13). Let us assume that (7.13) is satisfied. The operation (7.15) corresponds to projecting C_G, A_G, B_G with

$$\mathcal{Z} := \begin{bmatrix} I_{\sum_{j=1}^{n_{i-1}} n_j} & & \\ & Z_i & \\ & & I_{\sum_{j=i+1}^{k} n_j} \end{bmatrix},$$

$$\mathcal{V} := \begin{bmatrix} I_{\sum_{j=1}^{n_{i-1}} n_j} & & \\ & V_i & \\ & & I_{\sum_{j=i+1}^{k} n_j} \end{bmatrix}.$$

This implies that $\mathcal{Z}^T\mathcal{V} = I$ and $Im(V) \subseteq Im(\mathcal{V})$. This concludes the proof. □

Remark 7.8. It is also possible to project the subsystems $T_i(s)$ in such a way that the reduced interconnected transfer function $\hat{G}(s)$ satisfies a set of tangential interpolation conditions with respect to the original interconnected transfer function $G(s)$, but special care must be taken.

Indeed, Lemma 7.5 is generically not true anymore for generalized Krylov subspaces corresponding to tangential interpolation conditions. In other words, the column space of the matrix

$$\mathcal{K}_k\left((\lambda I - A_G)^{-1}B_G, (\lambda I - A_G)^{-1}, Y\right) :=$$

$$\left[\, (\lambda I - A_G)^{-1}B_G \ \dots \ (\lambda I - A_G)^{-k}B_G \,\right] \begin{bmatrix} y_0 & \cdots & y_{k-1} \\ & \ddots & \vdots \\ & & y_0 \end{bmatrix}$$

is in general not contained in the column space of the matrix

$$\mathcal{K}_k\left((\lambda I - A)^{-1}B, (\lambda I - A)^{-1}, Y\right) := \\ \left[(\lambda I - A)^{-1}B \ldots (\lambda I - A)^{-k}B\right] \begin{bmatrix} y_0 & \cdots & y_{k-1} \\ & \ddots & \vdots \\ & & y_0 \end{bmatrix}.$$

In such a case, interchanging matrices C_G, A_G, B_G by C, A, B, as done in Lemma 7.6 and Corollary 7.7 is not always permitted. Nevertheless, Lemma 7.6 and Corollary 7.7 can be extended to the tangential interpolation framework by projecting the state space realizations (C_i, A_i, B_i) with generalized Krylov subspaces of the form $\mathcal{K}_k\left((\lambda I - A_G)^{-1}B_G, (\lambda I - A_G)^{-1}, Y\right)$ and not of the form $\mathcal{K}_k\left((\lambda I - A)^{-1}B, (\lambda I - A)^{-1}, Y\right)$.

7.4 Examples of Structured Model Reduction Problems

As we will see in this section, many structured systems can be modelled as *interconnected systems*. Three well known structured systems are presented, namely *weighted* systems, *second-order* systems and *controlled* systems. For each of these specific cases one recovers well-known formulas. It turns out that several existing model reduction techniques for structured systems are particular cases of our ISBT algorithm.

The preceding list is by no means exhaustive. For instance, because linear fractional transforms correspond to making a constant feedback to a part of the state, this can also be described by an interconnected system. Periodic systems are also a typical example of interconnected system that is not considered below.

Weighted Model Reduction

As a first example, let us consider the following *weighted* transfer function :

$$y(s) = W_{out}(s)T(s)W_{in}(s)u(s) := G(s)u(s).$$

Let (C_o, A_o, B_o, D_o), (C, A, B, D) and (C_i, A_i, B_i, D_i) be the state space realizations of respectively $W_{out}(s)$, $T(s)$ and $W_{in}(s)$, of respective order n_o, n and n_i. A state space realization (C_G, A_G, B_G, D_G) of $G(s)$ is given by

$$\left[\begin{array}{c|c} A_G & B_G \\ \hline C_G & D_G \end{array}\right] := \left[\begin{array}{ccc|c} A_o & B_oC & B_oDC_i & B_oDD_i \\ 0 & A & BC_i & BD_i \\ 0 & 0 & A_i & B_i \\ \hline C_0 & D_oC & D_oDC_i & D_oDD_i \end{array}\right]. \tag{7.16}$$

The transfer function $G(s)$ corresponds to the *interconnected* system $\mathcal{S}$ with

$$\mathcal{S} \; : \; \begin{cases} b_1(s) = W_o(s)a_1(s), \quad b_2(s) = T(s)a_2(s), \quad b_3(s) = W_i(s)a_3(s) \\ \quad y(s) = b_1(s), \quad a_1(s) = b_2(s), \quad a_2(s) = b_3(s), \quad a_3 = u(s) \end{cases},$$

and

$$H = \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix} \quad , \quad K = \begin{bmatrix} 0 & I & 0 \\ 0 & 0 & I \\ 0 & 0 & 0 \end{bmatrix} \quad , \quad F = \begin{bmatrix} I & 0 & 0 \end{bmatrix}.$$

A frequency weighted balanced reduction method was first introduced by Enns [27, 94]. Its strategy is the following. Note that Enns assumes that $D = 0$ (otherwise D can be added to $\hat{T}(s)$).

Algorithm 7.4.1 *1. Compute the gramians P_G and Q_G satisfying* (7.8) *with C_G, A_G, B_G, D_G defined in* (7.16).

2. Perform a state space transformation on C, A, B in order to obtain $P = Q = \Sigma$ diagonal, where P and Q are the diagonal blocs of P_G and Q_G corresponding to the $T(s)$:

$$P = \begin{bmatrix} 0_{n,n_o} & I_n & 0_{n,n_i} \end{bmatrix} P_G \begin{bmatrix} 0_{n_o,n} \\ I_n \\ 0_{n_i,n} \end{bmatrix}, \quad Q = \begin{bmatrix} 0_{n,n_o} & I_n & 0_{n,n_i} \end{bmatrix} Q_G \begin{bmatrix} 0_{n_o,n} \\ I_n \\ 0_{n_i,n} \end{bmatrix}.$$

3. Truncate (C, A, B) by keeping only the part of the state space corresponding to the largest eigenvalues of Σ.

It is clear the algorithm of Enns is exactly the same as the ISBT Algorithm applied to weighted systems. As for the ISBT Algorithm 7.2.1, there is generally no known a priori error bound for the approximation error and the reduced order model is not guaranteed to be stable either.

There exists other weighted model reduction techniques. See for instance [89] where an elegant error bound is derived.

A generalization of weighted systems are *cascaded systems*. If we assume that the interconnected systems are such that the output of $T_i(s)$ is the input of $T_{i+1}(s)$, we obtain a structure similar than for the weighted case. For instance, the matrix K has the form

$$K = \begin{bmatrix} 0 & & & \\ I_{\beta_1} & \ddots & & \\ & \ddots & \ddots & \\ & & I_{\beta_{k-1}} & 0 \end{bmatrix}.$$

Second-Order systems

Second-Order systems 6.1 (the mass matrix M is assumed equal to the identity matrix) can be seen as an interconnection of two subsystems as follows.

Define $T_1(s)$ and $T_2(s)$ corresponding to the following system :

$$\begin{cases} \dot{x}_1(t) = -Dx_1(t) - Sy_2(t) + F_{in}u(t) \\ y_1(t) = x_1(t) \end{cases}, \quad \begin{cases} \dot{x}_2(t) = 0x_2(t) + y_1(t) \\ y_2(t) = x_2(t) \end{cases} \tag{7.17}$$

From this, $y_1(s) := T_1(s)a_1(s) = (sI_n + D)^{-1}a_1(s)$ with $a_1(s) := u_1(t) - Sy_2(s)$ (with the convention $u_1(t) = F_{in}u(t)$) and $y_2(s) = F_{out}s^{-1}a_2(s) := T_2(s)a_2$ with $a_2(s) = y_1(s)$. Matrices F, H, K are given by

$$F := \begin{bmatrix} 0 & F_{out} \end{bmatrix}, \quad H := \begin{bmatrix} F_{in} \\ 0 \end{bmatrix}, \quad K := \begin{bmatrix} 0 & -S \\ I & 0 \end{bmatrix}.$$

From the preceding definitions, one obtains

$$C = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \quad A = \begin{bmatrix} -D & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},$$
$$C_G = \begin{bmatrix} 0 & F_{out} \end{bmatrix}, A_G = \begin{bmatrix} -D & -K \\ I & 0 \end{bmatrix}, B_G = \begin{bmatrix} F_{in} \\ 0 \end{bmatrix}.$$

The matrices C_G, A_G, B_G are clearly a state space realization of $F_{out}(s^2I_n + Ds + S)^{-1}F_{in}$. At this point, it is not difficult to verify that the Second-Order Balanced Truncation technique proposed in [21] is exactly the same as our Interconnected Balanced Truncation technique applied to $T_1(s)$ and $T_2(s)$ (see Section 6.3). In general, systems of order k can be rewritten as an interconnection of k subsystems by generalizing the preceding ideas.

Controller Order Reduction

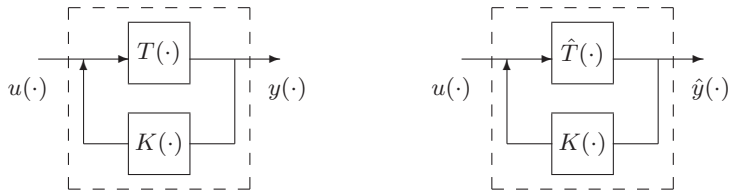
The Controller Reduction problem introduced by Anderson and Liu [2] is the following. Most high-order linear plants $T(s)$ are controlled with a high order linear system $K(s)$. In order to model such *structured* systems by satisfying the computational constraints, it is sometimes needed to approximate either the plant, or the controller, or both systems by reduced order systems, denoted respectively by $\hat{T}(s)$ and $\hat{K}(s)$.

The objective of Controller Order Reduction is to find $\hat{T}(s)$ and/or $\hat{K}(s)$ that minimize the *structured* error

$$\|G(s) - \hat{G}(s)\|, \qquad with$$
$$G(s) := (I - T(s)K(s))^{-1}T(s), \quad \hat{G}(s) := (I - \hat{T}(s)\hat{K}(s))^{-1}\hat{T}(s).$$

Balanced Truncation model reduction techniques have also been developed for this problem. Again, most of these techniques are very similar to the ISBT Algorithm. See for instance [87] for recent results. Depending on the choice of the pair of gramians, it is possible to develop balancing strategies that ensure the stability of the reduced system, under certain assumptions [63].

Fig. 7.2. Controller Order Reduction



7.5 Concluding Remarks

In this chapter, several model reduction techniques have been developed in order to perform model reduction of interconnected systems, either by SVD or by Krylov techniques. For instance, the ISBT Algorithm is a very generic tool for performing structured preserving balanced truncation.

As we have seen, our approach permits to unify several existing structured model reduction techniques in the same framework.

It should be pointed out that our approach, as any other input/output model reduction technique, cannot be apply to every interconnected systems but only a subset of it.

- First, only input/output systems (and subsystems) are studied. As pointed out in the beginning of this chapter, this is not always the best choice of representing a system since interconnections do not always correspond to a causal relation where outputs of one system are inputs to another system.
- Second, the techniques explained here can only be applied to systems composed of a small number of large scale linear systems that are connected to each other. If the interconnected system to be reduced is composed of a large number of small subsystems, then other model reduction techniques should be used. Assume for instance that a system $G(s)$ is composed of 1000 subsystems $T_i(s)$ of small Mc Millan degree, say for instance 10 so that the Mc Millan degree of $G(s)$ is 10000. Assume that you want to approximate $G(s)$ by a transfer function of degree 100. Then you cannot keep 1000 subsystems, i.e. you need another strategy. This example can seem artificial but this is exactly what happens in Circuit Simulation where one is facing systems of Mc Millan degree larger than 10^5 that are actually composed of a very large set of small subsystems.

7.6 Notes and References

A preliminary version of this chapter is available in [84]. This is the result of a collaborative work with Paul Van Dooren.

Interconnected systems, also called aggregated systems, have been studied in the eighties [29] in the model reduction framework, but received recently only little attention [76]. This is in contrast with controller and weighted SVD-based model reduction techniques that have been widely studied in the literature [75, 2, 27]. Controller reduction Krylov techniques have also been considered recently in [51].

Conclusion

The two major contributions of this thesis are the development of new interpolation techniques for model reduction, and structure preserving schemes for reducing *interconnected* systems.

The common point between these two contributions is that they both enlarge the set of available model reduction procedures in a unified framework, permitting to develop more flexible model reduction techniques.

Interpolation and Model Reduction

In the SISO case, the following points have been proved. Let (C, A, B) be a state space realization of a SISO transfer function $T(s)$.

- Given two transfer functions, it is always possible to construct a state space realization of the transfer function of smallest Mc Milan degree by projecting a state space realization of the transfer function of largest Mc Millan degree using Krylov subspaces and/or eigenspaces.
- The image of the matrix V is a sum of Krylov subspaces of the pair A, B and of eigensubspaces of the matrix A if and only if it is a solution of a Sylvester equation of the form

$$AVR_1 - VR_2 + BY = 0. \tag{7.18}$$

In other words, projecting with Sylvester equations is a universal model reduction framework for SISO systems.

The generalization of the preceding results for the MIMO case is the following. Let (C, A, B) be a state space realization of a $p \times m$ transfer function $T(s)$.

- V is a sum of generalized Krylov subspaces of the pair (A, B) and/or eigensubspaces of A if and only if it is a solution of a Sylvester equation of the form (7.18).

- Projecting (C, A, B) with generalized Krylov subspaces results in solving a tangential interpolation problem.
- Concerning the problem of embedding in the MIMO case, necessary and sufficient conditions have been found when the transfer functions differ only by one degree and do not have common poles. Moreover, generic solutions to the problem corresponding to some matrix pencils have been sketched. Finally, a connection between the pencil of Halevi and the system zero pencil is made.

Model Reduction of Structured Systems

Some large scale systems $G(s)$ are structured in the sense that they can be modelled as a set of subsystems $T_i(s)$ that interconnect to each other. SVD as well as Krylov techniques have been developed in order to preserve the structure of the original system.

- Concerning SVD-techniques, new pairs of *interconnected gramians* have been defined for each subsystem $T_i(s)$. These pairs of gramians are related to an energy-based optimization problem with respect to the overall system $G(s)$. From these gramians, the ISBT algorithm is derived. This algorithm turns out to be a generalization of existing controller and weighted balanced truncation techniques.
- Both interpolation and tangential interpolation techniques that preserve the structure of the original system have been developed.
- These techniques have been compared in a real-life application: the construction of reduced order models of the vibrating structure of a building. It turns out that the SOBT Algorithm performs very well in that example. Concerning Krylov technique, second order structure preserving techniques performed poorly compared to SVD techniques. This can be due to a bad selection of interpolation points.

Open Questions and Future Work

Many questions have appeared during the elaboration of this thesis. Amongst them, the following are of particular interest.

- Concerning Krylov techniques, a central question is "What interpolation conditions should we impose in order to obtain *good* reduced order systems?". This remains an important open problem for practical applications. A more precise question is "What are the interpolation conditions that correspond to Balanced Truncation?". In several numerical examples, one observes that interpolation points are the mirror images with respect to the imaginary axis of the poles of the original transfer function nearest to the imaginary axis. As a consequence, the error tends to behave as an

all-pass transfer function in the imaginary axis. This phenomenon has also been observed by other researchers (see recent work of T. Antoulas and D. Sorensen) and requires further investigation.

- As we have seen, the problem of embedding in the MIMO case is more complex and certainly not completely understood. A generalization of this problem would be to consider the embedding of generalized state space realizations of not necessarily proper transfer functions. More generally, there exist several open problems concerning the embedding of matrix polynomials that are under active investigation (see for instance [93] for recent results).
- The open problem 5.8 concerning the generality of Krylov techniques in the MIMO case [85] is still under investigation. It is now clear that the existence of projecting matrices is related to the eigenstructure of the system zero pencil of $T(s) - \hat{T}(s)$. Krylov subspaces appear then as *generic* subspaces for model reduction. There is a hope for unifying all the existing model reduction methods in the setting of Krylov subspaces.
- Concerning the ISBT technique of Chapter 7, does there exist another choice of gramians that provides an a priori global error bound? A more general question concerns the problem of approximating systems that are composed of a large number of small subsystems.

The preceding list is by no way exhaustive. Here are some more general considerations concerning the limitation of the results developed in this thesis.

A first limitation of the model reduction techniques studied here is that only linear time invariant systems have been considered. It turns out that many large scale systems have a nonlinear behavior that cannot be neglected. At present time, the study of Model Reduction techniques for nonlinear systems is not as well developed as the linear case. This is certainly a topic of central importance for practical applications.

Another limitation is that we only consider input/output systems represented by a state space realization. As already pointed out, this is not the only possible choice. We refer to [70] and references therein for another approach, namely the *behavioral* approach. Several results concerning the problem of model order reduction have been developed recently. For instance, algorithms have been derived that pass directly from the differential equations describing the behavior of a finite-dimensional linear system to a balanced state space representation have been derived in [91]. The study of model reduction of interconnected systems in this framework is an interesting way for future research.

All the model reduction methods developed for MIMO linear systems assume that the number of inputs m and outputs p is small compared to the dimension of the state space n. In many cases, the whole state needs to be approximated, and input or outputs do not always have a sense. Let us consider for instance the building problem discussed in Chapter 6. The state represents

a spatial discretization of the vibrating structure of a building. What is important in the active damping problem is to be able to attenuate the overall vibration of the structure and not some specific places, or an average position. To our knowledge, the problem of reducing large scale systems such that all the state is an input and/or an output has not received enough attention in the literature.

Clearly, numerous challenging problems need still to be solved in this area.

References

1. B. D. O. Anderson and A. C. Antoulas. Rational Interpolation and State-Variable Realizations. *Linear Algebra and its Applications*, 137:479–509, 1990.
2. B.D.O. Anderson and Y. Liu. Controller reduction: Concepts and approaches. *IEEE Trans. Autom. Control*, 34(8):802–812, 1989.
3. A. C. Antoulas. Lectures on optimal approximation of linear systems. Lecture Notes, October 1996.
4. A. C. Antoulas. *Lectures on the Approximation of Large-Scale Dynamical Systems*. SIAM, 2004.
5. A. C. Antoulas and B. D. O. Anderson. On the scalar rational interpolation problem. *IMA J. Math. Control Inform.*, 3(2,3):61–88, 1986.
6. A. C. Antoulas and B. D. O. Anderson. On the problem of stable rational interpolation. *Linear Algebra Appl.*, 122/123/124:301–329, 1989.
7. A. C. Antoulas, J. A. Ball, J. Kang, and J. C. Willems. On the Solution of the Minimal Rational Interpolation Problem. *Linear Algebra and its Applications*, 137:511–573, 1990.
8. A.C. Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: An overview. *Int. J. Appl. Math. Comput. Sci.*, 11(5):1093–1121, 2001.
9. Z. Bai. Krylov subspace technique for reduced-order modeling of large-scale dynamical systems. *Applied Numerical Mathematics*, 43:9–44, 2002.
10. Z. Bai. Structure preserving model reduction of very large second-order dynamical systems and related development. Seminar given at the BIRS workshop on Model Reduction Problems and Matrix Methods, 2004.
11. Z. Bai, P. M. Dewilde, and R. W. Freund. Reduced-order modeling. Numerical Analysis Manuscript 02–4–13, Bell Laboratories, 2002.
12. Z. Bai and Y. Su. Dimension reduction of second order dynamical systems via a second-order arnoldi method. Technical Report CSE-2004-1, University of California, Davis, 2004.
13. J. A. Ball, I. Gohberg, and L. Rodman. *Interpolation of rational matrix functions*. Birkhäuser Verlag, Basel, 1990.
14. J. A. Ball and J. Kang. Matrix polynomial solutions of tangential Lagrange-Sylvester interpolation conditions of low Mc Millan degree. *Linear Algebra and its Applications*, 137/138:699–746, 1990.

15. P. Benner, V. Mehrmann, V. Sima, S. Van Huffel, and A. Varga. SLICOT – a subroutine library in systems and control theory. In *Datta, Biswa Nath (ed.), Applied and computational control, signals, and circuits. Vol. 1. Basel: Birkhäuser. 499-539* . 1999.
16. A. Bultheel and M. Van Barel. Padé techniques for model reduction in linear system theory : a survey. *J. Comput. Appl. Math.*, 14:401–438, 1986.
17. A. Bunse-Gerstner, B. Salimbahrami, R. Grotmaack, and B. Lohmann. Existence and Computation of Second Order Reduced Systems using Krylov Subspace Methods. In *MTNS 2004 (16th Symp. on the Mathematical Theory of Networks and Systems)*, 2004.
18. Y. Chahlaoui. *Low-rank approximation and model reduction.* PhD thesis, Université catholique de Louvain, 2003.
19. Y. Chahlaoui, K. Gallivan, and P. Van Dooren. The H_∞ norm calculation for large sparse systems. In *Proceedings of 16th Symp. on the Mathematical Theory of Networks and Systems*, Katholieke Universiteit Leuven, Belgium, 2004.
20. Y. Chahlaoui, K. Gallivan, A. Vandendorpe, and P. Van Dooren. Model reduction of second order systems. In *Model Reduction of Dynamical Systems.* Eds. P. Benner et al., Springer Verlag, 2004. submitted.
21. Y. Chahlaoui, D. Lemonnier, A. Vandendorpe, and P. Van Dooren. Second-Order Balanced Truncation. *Linear Algebra Appl., Special Issue*, 2004. To appear.
22. Y. Chahlaoui, D. Lemonnier, A. Vandendorpe, and P. Van Dooren. Second Order Structure Preserving Balanced Truncation. In *MTNS 2004 (16th Symp. on the Mathematical Theory of Networks and Systems)*, 2004.
23. Y. Chahlaoui, K. Meerbergen, D. Lemonnier, A. Vandendorpe, and P. Van Dooren. Model reduction of second order system. In *MTNS 2002 (15th Symp. on the Mathematical Theory of Networks and Systems)*, 2002.
24. E. Chiprout and M. S. Nakhla. *Asymptotic Waveform Evaluation and Moment Matching for Interconnect Analysis.* The Kluwer International Series in Engineering and Computer Science., 1994.
25. S. E. Cohn. An introduction to estimation theory. *Journal of the Meteorological Society of Japan*, 75:257–288, 1997.
26. C. de Villemagne and R. E. Skelton. Model reductions using a projection formulation. *Internat. J. Control*, 46(6):2141–2169, 1987.
27. D. Enns. *Model Reduction for Control System Design.* PhD thesis, Department of Aeronautics and Astronautics, Stanford University, Stanford, 1984.
28. P. Feldmann and R.W. Freund. Reduced-order modeling of large linear passive multi-terminal circuits using matrix-padé approximation, 1997. http://www.cs.bell-labs.com/cm/cs/doc/97/4-03.ps.gz.
29. A. Feliachi and C. Bhurtun. Model reduction of large-scale interconnected systems. *Int. J. Syst. Sci.*, 18:2249–2259, 1987.
30. R. W. Freund. Krylov-subspace methods for reduced-order modeling in circuit simulation. *J. Comput. Appl. Math.*, 123(1-2):395–421, 2000.
31. R. W. Freund. Model reduction methods based on krylov subspaces. Technical report, Bell Laboratories, Lucent Technologies, 2003.
32. R.W. Freund. Padé-Type Reduced Order Modeling of Higher Order Systems. Seminar given at Stanford University, USA, 2003.
33. K. Gallivan, E. Grimme, and P. Van Dooren. Asymptotic waveform evaluation via a Lanczos method. *Appl. Math. Lett.*, 7(5):75–80, 1994.

34. K. Gallivan, E. Grimme, and P. Van Dooren. On some recent developments in projection-based model reduction. In *ENUMATH 97. Proceedings of the 2nd European conference on numerical mathematics and advanced applications held in Heidelberg, Germany*, pages 98–113. Bock, Hans Georg (ed.) et al., 1998.
35. K. Gallivan, E. Grimme, and P. Van Dooren. Model reduction of large-scale systems, rational Krylov versus balancing techniques. In *Error control and adaptivity in scientific computing*, pages 177–190. Kluwer Acad. Publ., Dordrecht, 1999.
36. K. Gallivan, A. Vandendorpe, and P. Van Dooren. Model reduction via tangential interpolation. In *MTNS 2002 (15th Symp. on the Mathematical Theory of Networks and Systems)*, page 6, 2002. Session WA4, paper WA4-5.
37. K. Gallivan, A. Vandendorpe, and P. Van Dooren. On the generality of Multipoint Padé approximations. In *15th IFAC World Congress on Automatic Control*, page 6, July 2002. Session T-Th-13, paper 2443.
38. K. Gallivan, A. Vandendorpe, and P. Van Dooren. Model Reduction via truncation : an interpolation point of view. *Linear Algebra Appl.*, 375:115–134, 2003.
39. K. Gallivan, A. Vandendorpe, and P. Van Dooren. Model reduction of MIMO systems via tangential interpolation. *SIAM Journal on Matrix Analysis and Applications*, 2004. To appear, available at http://www.auto.ucl.ac.be/vandendorpe/.
40. K. Gallivan, A. Vandendorpe, and P. Van Dooren. Sylvester equations and projection-based model reduction. *J. Comp. Appl. Math., Special Issue*, 162:213–229, 2004.
41. F.R. Gantmacher. *Theory of Matrices*, volume 1. Chelsea Publishing Company, Chelsea, New York, 1959.
42. F.R. Gantmacher. *Theory of Matrices*, volume 2. Chelsea Publishing Company, Chelsea, New York, 1959.
43. Y. Genin and A. Vandendorpe. On the Projection Technique for Model Order Reduction. In Preparation.
44. Y. Genin and A. Vandendorpe. Projection and embedding of rational transfer functions. In *SIAM Conference on Applied Linear Algebra*, The College of William and Mary, Williamsburg, July 2003. Abstract.
45. K. D. Glover. All optimal Hankel-norm approximation of linear multivariable systems and their L^∞-error bounds. *Int. J. Control*, 39(6):1115–1193, 1984.
46. G. Golub and C. F. Van Loan. *Matrix computations.3rd ed.* Baltimore, The Johns Hopkins Univ. Press. xxvii, 694 p. , 1996.
47. W. B. Gragg and A. Lindquist. On the partial realization problem. *Linear Algebra Appl.*, 50:277–319, 1983.
48. E. Grimme. *Krylov projection methods for model reduction.* PhD thesis, Department of Electrical Engineering, University of Illinois at Urbana-Champaign, 1997.
49. E.J. Grimme, D.C. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numer. Algorithms*, 12(1-2):1–31, 1996.
50. S. Gugercin and A.C. Antoulas. A survey of model reduction by balanced truncation and some new results. to appear in International Journal of Control, 2004.
51. S. Gugercin, C.A. Beattie, A.C. Antoulas, and E. Gildin. Controller Reduction by Krylov Projection Methods. In *Proceedings of 16th Symp. on the Mathematical Theory of Networks and Systems*, Katholieke Universiteit Leuven, Belgium, July 2004.

52. S. Gugercin, D.C. Sorensen, and A.C. Antoulas. A modified low-rank Smith method for large-scale Lyapunov equations. *Numer. Algorithms*, 32(1):27–55, 2003.
53. Y. Halevi. On model order reduction via projection. In *15th IFAC World Congress on Automatic Control*, page 6, July 2002.
54. Y. Halevi. On the class of reduced order models obtainable by projection. *Automatica*, 2004. submitted.
55. P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry.* Cambridge Monographs on Mechanics and Applied Mathematics., 1996.
56. R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press. XIII, 561 p., 1985.
57. R. A. Horn and C. R. Johnson. *Topics in matrix analysis.* Cambridge University Press. viii, 607 p. , 1991.
58. I. Jaimoukha and E. Kasenally. Implicitly restarted Krylov subspace methods for stable partial realizations. *SIAM Journal on Matrix Analysis and Applications*, 18(3):633–652, 1997.
59. B. Kågström and P. Van Dooren. A generalized state-space approach for the additive decomposition of a transfer matrix. *Journal of Numerical linear algebra with applications*, 1(2):165–181, 1992.
60. T. Kailath. *Linear Systems.* Information and System Sciences. Prentice-Hall, 1980.
61. J.-R. Li. *Model Reduction of Large Linear Systems via Low Rank System Gramians.* PhD thesis, Department of Mathematics, Massachusetts Institute of Technology, 2000.
62. J.-R. Li and J. White. Low rank solution of Lyapunov equations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):260–280, January 2003.
63. C.-A. Lin and T.-Y. Chiu. Model reduction via frequency weighted balanced realization. *Control-Theory and Advanced Technology*, 8(2):341–351, 1992.
64. J. J. Loiseau, S. Mondié, I. Zaballa, and P. Zagalak. Assigning the Kronecker invariants of a matrix pencil by row or column completions. *Linear Algebra Appl.*, 278(1-3):327–336, 1998.
65. E. Marques de Sá. Imbedding conditions for λ-matrices. *Linear Algebra Appl.*, 24:33–50, 1979.
66. D. G. Meyer and S. Srinivasan. Balancing and model reduction for second-order form linear systems. *IEEE Trans. Automat. Control*, 41(11):1632–1644, 1996.
67. T. Mukherjee, G. Fedder, D. Ramaswamy, and J. White. Emerging simulation approaches for micromachined devices. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 19(12):1572 – 1589, December 2000.
68. G. Obinata and B.D.O. Anderson. *Model reduction for control system design.* Communications and Control Engineering Series. London: Springer., 2001.
69. T. Penzl. A cyclic low rank Smith method for large sparse Lyapunov equations. *SIAM J. Sci. Comput.*, 21(4):1401–1418 (electronic), 1999/00.
70. J.W. Polderman and J.C. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach.* Springer Verlag, 1998.
71. A. Preumont. *Vibration control of active structures.* Kluwer Academic Publishers, 1997.

72. H. H. Rosenbrock. *State-space and multivariable theory.* John Wiley & Sons, Inc. [Wiley Interscience Division], New York, 1970.
73. M. F. Rubinstein. *Structural systems-statics, dynamics and stability.* Prentice-Hall, Inc, 1970.
74. B. Salimbahrami and B. Lohmann. Structure Preserving Order Reduction of Large Scale Second Order Systems. In *10th IFAC/IFORS/IMACS/IFIP Symposium on Large Scale Systems, Theory and Applications*, number 14, pages 260–267, July 2004.
75. G. Schelfhout. *Model reduction for control design.* PhD thesis, Electrotechnical Engineering ESAT, K.U. Leuven, 1996.
76. B. Shapiro. Frequency Weighted Sub-System Model Truncation to Minimize Systems Level Model Reduction Errors. submitted to IEEE, 2001.
77. R.E. Skelton and B.D.O. Anderson. q-Markov covariance equivalent realizations. *Int. J. Control*, 44:1477–1490, 1986.
78. C. Soize. Reduced models in the medium frequency range for general dissipative structural dynamics systems. *Eur. J. Mechan. & Solids*, 17:657–685, 1998.
79. D.C. Sorensen and A. C. Antoulas. The Sylvester equation and approximate balanced reduction. *Linear Algebra and its Applications*, 351-352:671–700, 2002.
80. T.-J. Su and R.R. Craig Jr. Model Reduction and Control of Flexible Structures Using Krylov Vectors. *J. Guidance*, 14(2):260–267, 1991.
81. R. C. Thompson. Interlacing Inequalities for Invariant Factors. *Linear Algebra Appl.*, 24:1–31, 1979.
82. P. Van Dooren. The generalized eigenstructure problem in linear system theory. *IEEE Trans. Autom. Control*, 26:111–129, 1981.
83. A. Vandendorpe and P. Van Dooren. Krylov techniques for model reduction of second-order systems. Technical Report 07, CESAME, Université catholique de Louvain, 2004.
84. A. Vandendorpe and P. Van Dooren. Model reduction of interconnected systems. In *MTNS 2004 (16th Symp. on the Mathematical Theory of Networks and Systems)*, 2004.
85. A. Vandendorpe and P. Van Dooren. Projection of state space realizations. In *Unsolved Problems in Mathematical Systems and Control Theory*, volume 2, pages 58–64. Princeton University Press, 2004.
86. A. Varga. Enhanced modal approach for model reduction. *Math. Model. Syst.*, 1(2):91–105, 1995.
87. A. Varga and B.D.O. Anderson. Accuracy-enhancing methods for balancing-related frequency-weighted model and controller reduction. *Automatica*, 39(5):919–927, 2003.
88. M. Verlaan. *Efficient Kalman Filtering Algorithms for Hydrodynamic models.* PhD thesis, Faculteit Technische Wiskunde en Informatica, Technische Universiteit Delft, 1998.
89. G. Wang, V. Sreeram, and W.Q. Liu. A new frequency-weighted balanced truncation method and an error bound. *IEEE Trans. Autom. Control*, 44(9):1734–1737, 1999.
90. W. Weaver Jr and P. R. Johnston. *Structural dynamics by finite elements.* Prentice-Hall, Inc, 1987.

91. J.C. Willems and P. Rapisarda. Balanced state representation with polynomial algebra. In *Directions in Mathematical Systems Theory and Optimization*, volume 286, pages 345–357. Springer Lecture Notes in Control and Information Sciences, 2002.
92. W. M. Wonham. *Linear Multivariable Control: A Geometric Approach*, volume 10 of *Applications of Mathematics*. Springer-Verlag, 3rd edition, 1985.
93. I. Zaballa. Feedback invariants of matrix quadruple completions. *Linear Algebra Appl.*, 292(1-3):73–97, 1999.
94. K. Zhou, J. C. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall, Inc, Upper Saddle River, NJ, 1996.