RICE UNIVERSITY

Model Reduction of Second Order Linear Dynamical Systems

by

Cong Teng

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Dan Sorensen, Chairman
Noah Harding Professor of Computational and Applied Mathematics, Rice University

Athanassios Antoulas
Professor of Electrical and Computer Engineering, Rice University

Matthias Heinkenschloss
Associate Professor of Computational and Applied Mathematics, Rice University

Nathaniel Dean
Chair and Professor of Mathematics, Texas Southern University

HOUSTON, TEXAS

JUNE 2004
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
Abstract

Model Reduction on Second Order Linear Dynamical Systems

by

Cong Teng

In this thesis, we propose several algorithms for model reduction of second order dynamical systems. These various projection methods are based on singular value decomposition, Krylov projection, and balanced truncation. In many cases models are given in second order form, and the goal is to produce a reduced order system which is in second order form, gives an accurate approximation of the original system, and maintains some important properties such as stability and passivity. Model reduction on first order linear time invariant dynamical systems has been extensively studied, algorithms and theory are well-developed. People usually study and deal with second order system by transforming to first order form which doubles the dimension. This can be inefficient and generally does not respect the second order form. The reduced model is not realizable as a second order system. So far only a very few
algorithms have been proposed for second order model reduction. Most of these are not practical for large scale settings, and no error bounds have been provided. In this thesis, a global error bound is given for some of the algorithms based on SVD and balanced truncation, the error bound is bounded by a constant times the summation of the neglected singular (or Hankel singular) values, that means those second order model reduction algorithms provide accurate approximations to the original systems. The structures of controllability and observability Gramians $P$ and $Q$ are discussed. All algorithms developed in this thesis have been implemented and shown to be numerically efficient, and applicable to large scale settings. All algorithms are implemented in Matlab, some of them are implemented in Fortran and C separately for which we use LAPACK. In this thesis, we apply our algorithms to three real models. The performance of our algorithms is compared with some of the previously existing algorithms. It turns out that most of our algorithms are very competitive with existing methods.
Acknowledgements

First I want to thank my advisor Dr. Dan Sorensen for his good guidance in research and study, for his patience, for choosing a good topic for my thesis. I really learned a lot from him. I am feeling lucky to have an advisor like him. I would also like to thank Dr. Nathaniel Dean whom I worked with for my first two years in my Ph.D. program at Rice, it was a lot of fun. It is really enjoyable to take Rice undergraduate courses, I took many classes in a wide range from four departments, it was a good experience for me, it let me gain some ability and feel more confident to face the future challenges. During my stay at Rice for five years, I really feel thankful for a good environment in research, study and life. The student life at Rice is wonderful, I really enjoy attending different events, all kinds of sports, Shepherd School music concerts, dance classes, and many many others.

I would like to thank CAAM Department for providing a good environment and many many services so that I can grow up in career. I would like to thank all Duncan Hall people for being so kind and nice to me, it is a great pleasure to be a member of them. In the end, I will give my deep thanks to my parents, I am really indebt to them, they always support me in any cases, there is a special spot for them in my heart forever.
# Contents

Abstract ........................................... ii

Acknowledgements ................................ iv

List of Figures ................................... viii

List of Tables ................................... xi

1 Introduction ................................... 1

1.1 Problem Overview ................................ 1

1.2 Methods, Approaches and Contributions .......... 4

1.3 Implementations and Applications .................... 6

1.4 Thesis Outline .................................... 7

2 Model Reduction of First Order Linear Dynamical Systems 9

2.1 Some Basic Theory ................................ 10

2.2 First Order Model Reduction ...................... 19
2.2.1 Gramian Based Method: Balanced Truncation and
Hankel-norm Approximation .................................. 20

2.3 Moment Matching: Krylov Method ....................... 23

3 Second Order Linear Dynamical Systems .................. 32

3.1 Introduction ............................................. 32
3.2 Second Order Linear Dynamical Systems ................. 36

4 Second Order Model Reduction Based on Balanced Truncation and
Singular Value Decomposition ................................ 47

4.1 Algorithms of Second Order Model Reduction Based on Balanced Truncation ................................................. 48

4.2 Algorithms of Second Order Model Reduction Based on Singular Value Decomposition ........................................... 62

4.3 Computing the Blocks of $\mathcal{P}$ and $\mathcal{Q}$ .................. 76

4.4 Some Discussions of the Proposed Algorithms .............. 83

5 Error Bounds .............................................. 87

5.1 Error Systems .......................................... 88

5.2 Error Bounds Associated with $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$ .............. 93

5.3 Error Bounds Associated with $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ .............. 100

6 Algorithms for Second Order Model Reduction Based on Krylov
Projection ..................................................... 110
6.1 Some Algorithms for Second Order Model Reduction Based on Krylov Projection ........................................... 110

7 Conclusions and Future Work ...................................... 121

7.1 Summary and Conclusions ........................................ 121

7.2 Future work ....................................................... 122

A Applications and Implementations ................................ 124

Bibliography ................................................................ 129
List of Figures

4.1 Frequency response on building model with size $n = 24$ by balanced truncation on $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ from the paper of Meyer and Srinivasan [34] 54

4.2 Frequency response on building model with size $n = 24$ by balanced truncation on $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$ ................................................. 56

4.3 Frequency response on building model with size $n = 24$ by balanced truncation on $(\mathcal{P}_{11}, \mathcal{P}_{22})$ and $(\mathcal{Q}_{11}, \mathcal{Q}_{22})$ from the paper by VanDooren etc. [53] ................................................................. 57

4.4 Frequency response on building model with size $n = 21$ by balanced truncation on $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ from the paper of Meyer and Srinivasan [34] 58

4.5 Frequency response on building model with size $n = 21$ by balanced truncation on $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$ ................................................. 59

4.6 Frequency response on building model with size $n = 21$ by balanced truncation on $(\mathcal{P}_{11}, \mathcal{P}_{22})$ and $(\mathcal{Q}_{11}, \mathcal{Q}_{22})$ from the paper of VanDooren etc. [53] ................................................................. 60

4.7 Frequency response on building model with size $n = 24$ by SVD on $\mathcal{P}_{11}$ . 66
4.8 Frequency response on building model with size \( n = 24 \) by SVD on \( P_{22} \). 67

4.9 Frequency response on building model with size \( n = 24 \) by SVD on \( Q_{11} \). 68

4.10 Frequency response on building model with size \( n = 24 \) by SVD on \( P_{22} \). 69

4.11 Frequency response on building model with size \( n = 21 \) by SVD on \( P_{11} \). 70

4.12 Frequency response on building model with size \( n = 21 \) by SVD on \( P_{22} \). 71

4.13 Frequency response on building model with size \( n = 21 \) by SVD on \( Q_{11} \). 72

4.14 Frequency response on building model with size \( n = 21 \) by SVD on \( Q_{22} \). 73

4.15 Frequency response on Aluminum model with size \( n = 1734 \) by SVD on 
\( P_{11} \), top figure is of order \( r = 240 \), bottom figure is of order \( r = 140 \). 74

4.16 Frequency response on Aluminum model with size \( n = 1734 \) by SVD on 
\( P_{22} \), with different ranges on frequencies. 75

4.17 Frequency response on Aluminum model with size \( n = 1734 \) by SVD on 
\( Q_{11} \), with different ranges on frequencies. 77

4.18 Frequency response on Aluminum model with size \( n = 1734 \) by SVD on 
\( Q_{22} \), with different ranges on frequencies, both are of order \( r = 140 \). 78

6.1 Frequency response on building model with size \( n = 24 \) by Krylov method on \( K^{-1}M \). 115

6.2 Frequency response on building model with size \( n = 24 \) by Krylov method on \( M^{-1}K \). 116

6.3 Frequency response on building model with size \( n = 21 \) by applying Arnoldi on \( K^{-1}M \). 117
6.4 Frequency response on building model with size $n = 21$ by applying Arnoldi on $M^{-1}K$ ........................................ 118

6.5 Frequency response on Aluminum model with size $n = 1734$, Krylov method, Applying Arnoldi to $K^{-1}M$ ................................. 119

6.6 Frequency response on Aluminum model with size $n = 1734$, Krylov method, Applying Arnoldi to $M^{-1}K$ ................................. 120

A.1 Frequency response on building model with size $n = 24$ by SVD on $\mathcal{P}_{11}$ and $\mathcal{P}_{22}$ ............................................. 125

A.2 Frequency response on building model with size $n = 24$ by SVD on $\mathcal{Q}_{11}$ and $\mathcal{Q}_{22}$ ............................................. 126

A.3 Frequency response on building model with size $n = 21$ by SVD on $\mathcal{P}_{11}$ and $\mathcal{P}_{22}$ ............................................. 127

A.4 Frequency response on building model with size $n = 21$ by SVD on $\mathcal{Q}_{11}$ and $\mathcal{Q}_{22}$ ............................................. 128
List of Tables

4.1 $\mathcal{H}_2$ norm of building model with size $n = 24$ by balancing on $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ from the paper of Meyer and Srinivasan [34] .................................. 55

4.2 $\mathcal{H}_2$ norm of building model with size $n = 24$ by balancing on $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$ ........................................................................................................ 56

4.3 $\mathcal{H}_2$ norm of building model with size $n = 24$ by balancing on $(\mathcal{P}_{11}, \mathcal{P}_{22})$ and $(\mathcal{Q}_{11}, \mathcal{Q}_{22})$ from the paper by VanDooren etc. [53] .................. 57

4.4 $\mathcal{H}_2$ norm of building model with size $n = 21$ by balancing on $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ from the paper of Meyer and Srinivasan [34] .............................. 58

4.5 $\mathcal{H}_2$ norm of building model with size $n = 21$ by balancing on $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$ ........................................................................................................ 59

4.6 $\mathcal{H}_2$ norm of building model with size $n = 21$ by balancing on $(\mathcal{P}_{11}, \mathcal{P}_{22})$ and $(\mathcal{Q}_{11}, \mathcal{Q}_{22})$ from the paper of VanDooren etc. [53] .................. 60

4.7 $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $\mathcal{P}_{11}$ ................ 66

4.8 $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $\mathcal{P}_{22}$ ................ 67

4.9 $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $\mathcal{Q}_{11}$ .................. 68
4.10 $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $Q_{22}$  

4.11 $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $P_{11}$  

4.12 $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $P_{22}$  

4.13 $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $Q_{11}$  

4.14 $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $Q_{22}$  

6.1 $\mathcal{H}_2$ norm of building model with size $n = 24$ by applying Arnoldi on $K^{-1}M$  

6.2 $\mathcal{H}_2$ norm of building model with size $n = 24$ by applying Arnoldi on $M^{-1}K$  

6.3 $\mathcal{H}_2$ norm of building model with size $n = 21$ by applying Arnoldi on $K^{-1}M$  

6.4 $\mathcal{H}_2$ norm of building model with size $n = 21$ by applying Arnoldi on $M^{-1}K$  

A.1 $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $P_{11}$ and $P_{22}$  

A.2 $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $Q_{11}$ and $Q_{22}$  

A.3 $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $P_{11}$ and $P_{22}$  

A.4 $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $Q_{11}$ and $Q_{22}$
Chapter 1

Introduction

1.1 Problem Overview

In this thesis we consider the following second order linear dynamical systems

\[ \Sigma_2 : \left\{ \begin{array}{l}
    M\dddot{q}(t) + G\dddot{q}(t) + K\dddot{q}(t) = B_0 u(t) \\
    y(t) = E_0\ddot{q}(t) + F_0\dddot{q}(t)
\end{array} \right. \quad (1.1.1) \]

where \( M, G, K \in \mathbb{R}^{n \times n}, B_0 \in \mathbb{R}^{n \times m}, E_0, F_0 \in \mathbb{R}^{p \times n}, u : [0, \infty) \mapsto \mathbb{R}^m \) is the input, \( y : [0, \infty) \mapsto \mathbb{R}^p \) is the output, \( q : [0, \infty) \mapsto \mathbb{R}^n \) are the internal states, \( n \) is the state-space dimension.

Many applications, such as circuit simulation and power systems lead to very large models, specific examples are the NASA space network antenna [21], Euler-Lagrangian equations [7]. Thus we are interested in cases where \( n \) is very large, while \( m, p \ll n \). Due to limitations on time and computer storage, it is often difficult or impossible to directly simulate or control these large scale systems.
Usually people study and deal with second order systems by transforming to first order systems,

\[
\Sigma_1 : \begin{cases}
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t)
\end{cases} \iff \Sigma_1 := \begin{pmatrix} A & B \\ C & D \end{pmatrix}
\] (1.1.2)

where \( A \in \mathbb{R}^{N \times N}, B \in \mathbb{R}^{N \times m}, C \in \mathbb{R}^{p \times N}, D \in \mathbb{R}^{p \times m} \), and \( u \) is the input, \( y \) is the output, \( x \) is the internal state space, \( N \) is the state-space dimension.

However this is inefficient since the corresponding first order system has double dimension \( N, N = 2n \), and moreover the reduced system is generally no longer in second order form. This is not recommended from physical point of view since the system will lose the original physical meanings in some aspects.

It is desirable to reduce the second order system to smaller order which preserves some important properties of the original one. It is also desirable for reduced system

1. to preserve the second order structure;

2. to satisfy an error bound assuming that \( \frac{\|y - y_{\text{red}}\|}{\|u\|} < \text{tol} \), i.e. outputs corresponding to the same inputs are close;

3. to preserve important system properties such as stability, passivity, etc..

This thesis addresses these issues.

It is convenient to study this problem in the frequency domain. Taking the Laplace transformation in (1.1.2) and (1.1.1), we can get the transfer functions for first order
and second order systems respectively,

\[ G(s) = C(sI - A)^{-1}B + D \]  \hspace{1cm} (1.1.3)

\[ H(s) = (E_0 + F_0s)(Ms^2 + Gs + K)^{-1}B_0 \]  \hspace{1cm} (1.1.4)

and the relations of inputs and outputs in Laplace domain are then,

\[ Y(s) = G(s)U(s) \]

and

\[ Y(s) = H(s)U(s) \]

First order model reduction has been extensively studied. Two effective ways to obtain first order model reduction are **moment matching** and **balanced truncation**. A number of methods on moment matching based on Krylov projection have been proposed, the algorithms are efficient because of the properties of Krylov projection methods [3], [17]. Balanced truncation [35] is a well-known and well-developed method, it has a good error bound which is 2 times the summation of the neglected Hankel singular values. The algorithms are not computationally efficient when the system is in large scale setting since the cost is \( O(n^3) \), and many approximation algorithms based on balanced truncation method have been proposed in order to solve large scale problems.

The idea of balancing is to try to diagonalize **controllability and observability** gramians \( P \) and \( Q \) simultaneously, i.e. to find balanced transformation matrices \( W \)
and $V$, where $W, V \in \mathbb{R}^{N \times N}$ and $W^T V = I_N$, such that

$$\begin{align*}
W^T \mathcal{P} W = V^T Q V = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_N).
\end{align*}$$

The idea of balanced truncation is to let the reduced system keep the $k$ largest Hankel singular values $\sigma_1, \sigma_2, \cdots, \sigma_k$ of the original system. Therefore in the balanced truncation method, we need to compute the projection matrices $W_1, V_1 \in \mathbb{R}^{N \times k}$, where $W_1$ and $V_1$ consist of the first $k$ columns of $W$ and $V$ respectively, then perform projection to truncate the original system.

So far very few algorithms for second order model reduction have been proposed, most of which are not practical for large scale systems, no error bounds are provided for the error systems, and for the algorithms based on Krylov projection methods, no rational interpolation (moment matching) has been performed or proved. See Chapter 3 and Chapter 6.

1.2 Methods, Approaches and Contributions

Since the algorithms and theory on first order model reduction are well developed, it is natural to derive the algorithms on second order model reduction by using the existing techniques on an equivalent first order system.

The existing algorithms on second order model reduction are derived by two approaches: gramian based and Krylov projection methods [34] [52] [53] [18].

For a given second order system, suppose $\mathcal{P}$ and $\mathcal{Q}$ are the controllability and
observability gramians for the corresponding first order system, partition $\mathcal{P}$ and $\mathcal{Q}$ to four blocks of equal size: $\mathcal{P} = \begin{pmatrix} P_{11} & P_{12} \\ P_{T12} & P_{22} \end{pmatrix}$, $\mathcal{Q} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{T12} & Q_{22} \end{pmatrix}$. Balanced truncation method is a gramian based method that tries to block diagonalize the corresponding first order gramians $\mathcal{P}$ and $\mathcal{Q}$, or suppose they are block diagonal, i.e. $P_{12} = 0$ and $Q_{12} = 0$.

In this thesis, we present several algorithms for second order model reduction. Three general approaches are considered: singular value decomposition, balanced truncation, and Krylov projection.

The SVD based algorithms are derived by a new approach, namely by the formula of $\mathcal{H}_2$ norm of second order systems, and combining with the optimization problems associated with second order linear systems. The main idea of these algorithms is to take SVD on diagonal blocks of the corresponding first order controllability and observability gramians, then form the projection basis using the singular vectors corresponding to the $r$ largest singular values.

One algorithm is proposed based on balanced truncation method. Suppose $\mathcal{P}$ and $\mathcal{Q}$ are the controllability and observability gramians of the corresponding first order system, if we partition $\mathcal{P}$ and $\mathcal{Q}$ as four blocks with equal size, let $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$ are the $(2,2)$-block of $\mathcal{P}$ and $\mathcal{Q}$ respectively. The main idea of our algorithm is to balance $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$, and get the projection matrices.

For some of the algorithms based on SVD and balanced truncation methods, we give a global error analysis to show that the $\mathcal{H}_2$ norm of the error system is bounded.
by a scalar times the summation of the neglected singular (or Hankel singular) values. This provides a theoretical basis that these model reduction algorithms produce reduced systems that give accurate approximations to the original system.

Several algorithms are proposed based on Krylov projection. These algorithms work fast and efficiently because of the properties of Krylov projection methods.

In all algorithms, the reduced systems maintain second order form. The structures of $P$ and $Q$ are also discussed, where $P$ and $Q$ are the controllability and observability gramians for the corresponding first order system. Under the assumption of $F_0 = 0$, we observe that in any case (i.e. any coordinate system), $Q$ cannot be block diagonal, i.e. the off diagonal blocks are not zeros. Therefore when applying the well known first order balanced truncation technique to the corresponding first order system, i.e. to diagonalize both $P$ and $Q$, some important information of the original system can be lost in reduced system.

1.3 Implementations and Applications

All algorithms except one, the passivity preserving, are numerically efficient and suitable for large scale settings. All algorithms are implemented in Matlab, some of them are implemented in Fortran and C separately for which we use LAPACK and ARPACK. In Appendix A of this thesis, we apply our algorithms to three real models, the performance is compared with some of the previously existing algorithms in sigma plots, it turns out that most of our algorithms are competitive with those
existing ones.

1.4 Thesis Outline

- Chapter 1 is an overview of this thesis.

- Chapter 2 gives a brief description on first order model reduction, including basic theory and two most effective ways for model reduction, namely balanced truncation and moment matching.

- Chapter 3 studies second order linear system. We introduce some important quantities of second order linear system, such as stability, gramians, and $\mathcal{H}_2$ norm. We make an observation on the structures of $P$ and $Q$. A concise expression of $\mathcal{H}_2$ norm for second order linear systems is derived.

- Chapter 4 proposes some algorithms based on balanced truncation and singular value decomposition. We also derive some properties related to equivalence and $\mathcal{H}_2$ norms of the reduced and original systems.

- Chapter 5 provides global error bounds for some of the algorithms based on balanced truncation and SVD. The error bounds associated with $P_{11}$ and $Q_{11}$ (i.e. the $(1,1)$-block of $P$ and $Q$ respectively, where $P$ and $Q$ are the controllability and observability gramians of the corresponding first order system) are bounded by a scalar times the summation of the neglected singular (or Hankel singular) values, which is a desired result.
• Chapter 6 presents some algorithms based on Krylov projection method which are fast and efficient because of the properties of Krylov projection method.

• Chapter 7 gives conclusions and describes future work.

• Appendix A and the corresponding chapters shows the plots of frequency response of the algorithms that we presented in this thesis to three real models. The $\mathcal{H}_2$ norms of the reduced and error systems are listed as tables.
Chapter 2

Model Reduction of First Order

Linear Dynamical Systems

First order model reduction is well-developed, but only a few algorithms and little theory exist for second order systems. Because of the close relations between first and second order systems, it is natural to derive algorithms and theory for second order model reduction based on the algorithms of reducing the corresponding first order systems. Therefore the basic theory and algorithms for first order model reduction are presented.

In this chapter, first we introduce the basic theory and establish the essential definitions and notation. Then, we describe two most effective ways in first order model reduction: balanced truncation and moment matching based on Krylov projection.

This chapter is based on the lecture notes of ELEC 501 in Fall 2001 [3].
2.1 Some Basic Theory

In this thesis, we consider the following first order continuous linear dynamical systems:

\[
\Sigma \begin{cases} 
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t) 
\end{cases} \tag{2.1.1}
\]

or for simplicity:

\[
\Sigma := \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{2.1.2}
\]

where \( A \in \mathbb{R}^{N \times n} \), \( B \in \mathbb{R}^{N \times m} \), \( C \in \mathbb{R}^{p \times N} \), \( D \in \mathbb{R}^{p \times m} \) are given matrices, \( m \) and \( p \) are the number of inputs and outputs respectively, \( u(t) \in \mathbb{R}^m \) is the given vector of inputs, \( y(t) \in \mathbb{R}^p \) is the unknown vector of outputs, \( x(t) \in \mathbb{R}^N \) is the unknown vector of internal variables, \( N \) is the dimension of the system.

With some modifications, the definitions and results in continuous linear systems can apply to discrete linear systems,

\[
\begin{cases} 
x(t + 1) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t) 
\end{cases} \tag{2.1.3}
\]

where \( t \in \mathbb{R} \). See [3] [7] for more details.

For a vector-valued or single variable function \( f(t) \), we denote by

\[
F(s) = \int_0^\infty f(t)e^{-st} \, dt \tag{2.1.4}
\]

the Laplace transformation of \( f \). Here \( s \) is in general complex. When it is purely
imaginary, \( s = j \omega \ (\omega \geq 0) \) is physically meaningful and corresponds to the frequency \( \omega \).

Throughout this discussion, we assume \( x(0) = 0 \). The Laplace transformation of \( \dot{x}(t) \) is then
\[
\int_0^\infty \dot{x}(t)e^{-st} \, dt = sX(s).
\]

Applying Laplace transformation to linear system (2.1.1), we get the following frequency-domain expression:
\[
\begin{aligned}
    sX(s) &= AX(s) + BU(s) \\
    Y(s) &= CX(s) + DU(s)
\end{aligned}
\tag{2.1.5}
\]

Eliminating \( X(s) \) in (2.1.5) results in
\[
Y(s) = (C(sI - A)^{-1}B + D)U(s)
\]

which is the frequency-domain input-output relation,
\[
Y(s) = G(s)U(s) \tag{2.1.6}
\]

where
\[
G(s) = C(sI - A)^{-1}B + D \tag{2.1.7}
\]

is called transfer function, one of the most important quantities in linear system theory. Note that \( G(s) : \mathbb{C} \mapsto \mathbb{C}^{p \times m} \) is a matrix-valued function. The dual system is defined as the transpose of a transfer matrix \( G(s) \) ([56]):
\[
G(s) \rightarrow G^T(s) = B^T(sI - A^T)^{-1}C^T + D^T
\]
or equivalently

\[
\begin{bmatrix}
A & B \\
C & D
\end{bmatrix} \rightarrow
\begin{bmatrix}
A^T & C^T \\
B^T & D^T
\end{bmatrix}
\]

The conjugate system is defined as follows:

\[ G \rightarrow \bar{G}(s) = G^T(-s) = B^T(-sI - A^T)^{-1}C^T + D^T \]

or equivalently

\[
\begin{bmatrix}
A & B \\
C & D
\end{bmatrix} \rightarrow
\begin{bmatrix}
-A^T & -C^T \\
B^T & D^T
\end{bmatrix}
\]

It is easy to see that system \( \Sigma \) is also a linear operator mapping inputs \( u(.) \in \mathcal{L}_2[-\infty, +\infty] \) to outputs \( y(.) \in \mathcal{L}_2[-\infty, +\infty] \). The following is the definition for dual operator.

**Definition 2.1.1** Suppose \( U \) and \( Y \) are Hilbert spaces equipped with inner products \(<.,>_U \) and \(<.,>_Y \) respectively. Let \( L \) be a linear operator from \( U \) to \( Y \), \( L : U \mapsto Y \). The dual of \( L \) is defined as the linear operator from \( Y \) to \( U \), \( L^* : Y \mapsto U \), such that

\[ <Lu, y>_Y = <u, L^*y>_U \]

for all \( y \in Y \) and all \( u \in U \).

It is not hard to verify (see [56]) that the dual operator for system (2.1.1) is the linear operator for its dual system.

For linear system (2.1.1), especially in large scale settings, the goal is usually to obtain input-output mapping \( u(t) \mapsto y(t) \) (or \( U(s) \mapsto Y(s) \)), rather than computing all the components of internal state variable \( x(t) \) (or \( X(s) \)), since what we mostly
concerned with are the outputs for different given inputs. Obviously, transfer function plays a key role in these cases.

For a given system \( \begin{pmatrix} A & B \\ C & D \end{pmatrix} \), if we transform the state (coordinate transformation) as following,

\[ \tilde{x} = Tx, \quad \text{det} \, T \neq 0 \]

(2.1.1) in the new state \( \tilde{x} \) becomes

\[
\begin{align*}
\dot{\tilde{x}} &= \underbrace{TAT^{-1}}_{\hat{A}} \tilde{x} + \underbrace{TB}_{\hat{B}} u \\
y &= \underbrace{CT^{-1}}_{\hat{C}} \tilde{x} + Du
\end{align*}
\]

where \( D \) remains unchanged. The corresponding triples \( \begin{pmatrix} A & B \\ C & D \end{pmatrix} \) and \( \begin{pmatrix} \hat{A} & \hat{B} \\ \hat{C} & D \end{pmatrix} \) are called **equivalent**.

Under equivalence,

\[
G_\Sigma(s) = D + C(sI - A)^{-1} = D + CT^{-1}T(sI - A)^{-1}T^{-1}TB
\]

\[
= D + CT^{-1}(sI - TAT^{-1})^{-1}TB = \hat{D} + \hat{C}(sI - \hat{A})^{-1}\hat{B}
\]

\[
= G_{\hat{\Sigma}}(s).
\]

So equivalent triples have the same transfer function. From here, we can see that for a system with given inputs and outputs, there are infinitely many internal state descriptions.

By solving the differential equation in (2.1.1), it is not hard to get the solution
for $x(t)$ ([7]),

$$x(t) = e^{(t-t_0)A}x(t_0) + \int_{t_0}^{t} e^{(t-\tau)A}Bu(\tau) \, d\tau$$

(2.1.8)

The output in time domain is then:

$$y(t) = Ce^{(t-t_0)A}x(t_0) + C \int_{t_0}^{t} e^{(t-\tau)A}Bu(\tau) \, d\tau + Du(t)$$

(2.1.9)

Compare it with the input-output relation in frequency domain which is much simpler:

$$Y(s) = G(s)U(s) = (C(sI - A)^{-1}B + D)U(s)$$

(2.1.10)

**Stability** of the system is another important concept, see [7].

The origin is *stable* if for any $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that if $||x(0)|| < \delta$ then $||x(t)|| < \varepsilon$ for all $t > 0$. Otherwise the origin is *unstable*.

The origin is *asymptotically stable* if it is stable, and if there exists $\delta' > 0$ such that if $||x(0)|| < \delta'$ then $\lim_{t \to \infty} ||x(t)|| = 0$.

**Theorem 2.1.2** [7] *First order linear system* (2.1.1) *is asymptotically stable* $\iff$ *all eigenvvalues of* $A$ *have negative real parts.*

In practice, we do not design systems to explode, all practical linear systems are designed to be stable.

For input $u \in \mathcal{L}_2(-\infty, 0]$, assume the initial condition is zero, i.e. $\lim_{t \to -\infty} x(t) = 0$, from (2.1.8), the state at time $0$ is then

$$x(0) = \int_{-\infty}^{0} e^{-At}Bu(t) \, dt = C_0 u(t)$$

(2.1.11)
Assume a zero input for system (2.1.1) when \( t > 0 \), i.e. \( u(t) = 0 \) \( (t > 0) \), then from (2.1.9), for all \( t \geq 0 \) the output \( y(t) \in \mathcal{L}_2[0, +\infty] \) of system (2.1.1) is:

\[
y(t) = Ce^{At}x(0) = O_b x(0)
\]  

(2.1.12)

\( C_0 \) is called reachability operator, \( C_0 : \mathcal{L}_2[-\infty, 0] \mapsto \mathbb{R}^n \), which maps the past inputs \( u(.) \) to the present state; and \( O_b \) is called observability operator, \( O_b : \mathbb{R}^n \mapsto \mathcal{L}_2[0, +\infty] \), which maps the present state to future outputs \( y(.) \).

Given two arbitrary vectors \( x(t_0) \) and \( x_f \) in state space, if there exists the input vector \( u(t) \) that will take the initial state \( x(t_0) \) to the final state \( x(t) = x_f \), then the system will be called reachable. If for all possible known input and output vectors \( u(t) \) and \( y(t) \), we can determine the value of the initial state \( x(t_0) \), then the system is called observable.

In the definition of reachability, if we let \( x_f = 0 \), then the controllability can be defined. In continuous-time linear systems, reachability and controllability are identical.

**Theorem 2.1.3** A system is controllable if and only if

\[
\text{rank}[B, AB, \ldots, A^{n-1}B] = n.
\]

A system is observable if and only if

\[
\text{rank}[C, CA, \ldots, CA^{n-1}]^T = n.
\]

In linear systems (2.1.1), there are two important so-called Lyapunov equations:

\[
AP + PA^T + BB^T = 0
\]  

(2.1.13)
\[ A^T Q + QA + C^T C = 0 \]  

(2.1.14)

under assumption that the system is stable, it is well-known that the reachability and observability gramians \( \mathcal{P} \) and \( \mathcal{Q} \) are the unique solutions to Lyapunov equations (2.1.11) and (2.1.12), both \( \mathcal{P} \) and \( \mathcal{Q} \) are symmetric positive definite, and \( \mathcal{P} \) and \( \mathcal{Q} \) can be expressed as:

\[
\mathcal{P} := \int_{0}^{\infty} e^{A\tau} BB^T e^{A^T \tau} d\tau \\
\mathcal{Q} := \int_{0}^{\infty} e^{A^T \tau} C^T C e^{A\tau} d\tau
\]

It is also easy to show that (see [7]) \( \mathcal{P} = C_0^* C_0 \) and \( \mathcal{Q} = O_0 O_0^* \).

We have the following results:

**Proposition 2.1.4** The system is reachable if and only if \( \mathcal{P} \) is positive definite. The system is observable if and only if \( \mathcal{Q} \) is positive definite.

The unit impulse \( \delta : [0, \infty) \mapsto \mathbb{R} \) is defined as

\[ \delta(t) = 0 \text{ for } t \neq 0, \text{ and } \int_{-\infty}^{\infty} \delta(t) \, dt = 1 \]

By using the sifting property of the impulse, if \( a < \tau < b \) and \( f(t) \) is continuous at \( t = \tau \), then

\[
\int_{a}^{b} f(t) \delta(t - \tau) \, dt = f(\tau). \tag{2.1.15}
\]

By (2.1.15), equation (2.1.9) becomes

\[
y(t) = \int_{0}^{t} (Ce^{(t-\tau)A} B + D\delta(t - \tau)) u(t) \, d\tau
\]
\begin{equation}
= \int_0^t h(t - \tau)u(\tau) d\tau
\end{equation}

This has the form of a convolution integral, and \( h(t) \) is called the \textit{impulse response matrix}, also called \textit{weighting matrix}, which has the form

\begin{equation}
h(t) = Ce^{tA}B + D\delta(t)
\end{equation}

The \textbf{Hankel operator} \( \mathcal{H} \) is defined as:

\begin{equation}
\mathcal{H} : \mathcal{L}^m(\mathbb{R}_-) \to \mathcal{L}^p(\mathbb{R}_+), u_- \mapsto y_+(t) := \mathcal{H}(u_-) = \int_{-\infty}^{0} h(t - \tau)u_- (\tau) d\tau, \; t \geq 0.
\end{equation}

Now we can define another two important quantities of the system.

\textbf{Definition 2.1.5} \textit{The Hankel singular values of the stable system} \( \Sigma \), denoted by \( \sigma_1, \sigma_2, \ldots, \sigma_r \), are the singular values of \( \mathcal{H} \) defined by (2.1.17). \textit{The Hankel norm of} \( \Sigma \) \textit{is the largest singular value of} \( \mathcal{H} \),

\[ ||\Sigma||_{\mathcal{H}} = \sigma_{\max}(\mathcal{H}) = \sigma_1 \]

\textit{The} \( \mathcal{H}_2 \textit{ norm} \) \textit{is defined as:}

\[ ||\Sigma||_{\mathcal{H}_2} := ||h(t)||_{\mathcal{L}^2(\mathbb{R}^+)} \]

[3] showed that the \( \mathcal{H}_2 \) norm of the system exists only if \( D = 0 \) and \( \Sigma \) is stable, in this case it can be expressed as:

\[ ||\Sigma||_{\mathcal{H}_2}^2 = \int_0^\infty \text{trace}\{h^T(t)h(t)\} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{\hat{h}^T(-j\omega)\hat{h}(j\omega)\} d\omega \]
where $\hat{h}(s)$ is the Laplace transformation of $h(t)$, and the $\mathcal{H}_2$ norm can also be expressed as

$$||\Sigma||_{\mathcal{H}_2}^2 = \text{trace}\{B^TQB\} = \text{trace}\{C^TPC^T\}$$  \hspace{1cm} (2.1.19)

The $\mathcal{H}_\infty$-norm of system $\Sigma$ is

$$||\Sigma||_{\mathcal{H}_\infty} = \sup_{\omega \in (-\infty, \infty)} \sigma_{\max}(G(j\omega))$$  \hspace{1cm} (2.1.20)

Recall the definition of equivalent systems. Under equivalence ($\tilde{x} = Tx$), the gramians are transformed as follows:

$$\tilde{P} = TP^T, \tilde{Q} = T^{-T}QT^{-1} \Rightarrow \tilde{P}\tilde{Q} = T(PQ)T^{-1}$$  \hspace{1cm} (2.1.21)

Therefore the product of the two gramians of equivalent systems are similar, and so have the same eigenvalues.

**Lemma 2.1.6** The **Hankel singular values** of a reachable, observable and stable linear system $\Sigma$ are the square roots of the eigenvalues of the products of gramians $P$ and $Q$:

$$\sigma_i = \sqrt{\lambda_i(PQ)}, \quad i = 1, 2, \cdots, n$$  \hspace{1cm} (2.1.22)

Therefore $\sigma_1$ is the **Hankel norm** of $\Sigma$,

$$||\Sigma||_{\mathcal{H}} = \sqrt{\lambda_{\max}(PQ)} = \sigma_1$$  \hspace{1cm} (2.1.23)

From (2.19), it is easy to see the Hankel singular values are transformation invariant.

We close this section by presenting a lemma.
Lemma 2.1.7 Let $\mathcal{P}$ and $\mathcal{Q}$ denote the infinite gramians of a stable linear system.

1. The minimal energy required to steer the state of the system from 0 to $x_r$ is:

$$x_r^* \mathcal{P}^{-1} x_r$$

2. The maximal energy produced by observing the output of the system whose initial state is $x_0$ is:

$$x_0^* \mathcal{Q} x_0$$

3. The states which are difficult, i.e. require large amounts of energy to reach are in the span of those eigenvectors of $\mathcal{P}$ which correspond to small eigenvalues. Furthermore, the states which are difficult to observe, i.e. produce small observation energy, are in the span of those eigenvectors of $\mathcal{Q}$ which correspond to small eigenvalues.

Lemma 2.1.6 and Lemma 2.1.7 are standard results in the theory of linear systems. For proof and more details, see [3] and standard textbooks.

### 2.2 First Order Model Reduction

The basic idea of model reduction is to produce a lower dimensional system of the same type, that will approximate the outputs of the original system, i.e. $\frac{\|y - y_r\|}{\|u\|} < tol$.

The basic approximation methods for model reduction on linear systems are gramian based and Krylov based. There are two ways for gramian based method:
balanced truncation and Hankel-norm approximation. Krylov based method is to try to do the moment matching in order for reduced system to approximate the original system.

The final stage of model reduction is to produce \((A_r, B_r, C_r)\), this is usually done by projection method.

\begin{center}
\textbf{First order model reduction:}
\end{center}

\begin{enumerate}
\item \textbf{Step 1.} Form the transformation matrices \(W_k, V_k \in \mathbb{K}^{n \times k}\) such that \(W_k^T V_k = I_k\).
\item \textbf{Step 2.} Perform the projection to \((A, B, C)\) and get \((A_k, B_k, C_k)\):
\[ A_k = W_k^T A V_k, \quad B_k = W_k^T B, \quad C_k = CV_k. \]
\end{enumerate}

If \(W_k = V_k\), then it is called orthogonal projection.

### 2.2.1 Gramian Based Method: Balanced Truncation and Hankel-norm Approximation

The widely used approach in first order model reduction \textbf{balanced truncation} which was introduced by Moore [35] and now appears in many standard textbooks.

From Lemma 2.5, we know that if \(P = Q = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)\), then a state which is difficult to reach is also difficult to observe. It is desirable to eliminate such states from the reduced system. The basic idea of balanced truncation is to find a realization for system (2.1.1) so that in the new coordinate system both \(P\) and \(Q\)
are diagonal and equal, then we can truncate the system by keeping $r$ largest Hankel singular values.

**Definition 2.2.1** The stable system $\Sigma$ is balanced if and only if $\mathcal{P} = \mathcal{Q}$. $\Sigma$ is principle-axis balanced if and only if

$$\mathcal{P} = \mathcal{Q} = \Sigma := \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)$$

In this thesis, we should use the term balanced to mean principle-axis balanced simply balanced.

The problem is to find a balancing transformation matrix, then truncate. There are several ways to do this. Next we list a numerically stable implementation for balanced truncation method.

**Algorithm 2.2.1 Balanced Truncation** [35]

*input* $A, B, C, k$; *output* $A_k, B_k, C_k$.

1. *Compute* $\mathcal{P}, \mathcal{Q}$.

2. *Take* Cholesky factorizations of $\mathcal{P}$ and $\mathcal{Q}$,

$$\mathcal{P} = UU^T, \quad \mathcal{Q} = LL^T$$

where $U$ is upper triangular and $L$ is lower triangular.

3. *Compute* SVD of $U^TL$,

$$U^TL = U_2S_2V_2^T$$
4. Compute the balanced transformations as follows:

\[ W^T = S_2^{-\frac{1}{2}} V_2^T L^T \]

\[ V = U U_2 S_2^{-\frac{1}{2}} \]

Form the projection matrices,

\[ W_k = W[:, 1 : k], \quad V_k = V[:, 1 : k]. \]

i.e. \( W_k \) and \( V_k \) consist of the first \( k \) columns of \( W \) an \( V \), respectively.

5. Perform the projection:

\[ A_k = W_k^T A V_k, \quad B_k = W_k^T B, \quad C_k = C V_k \]

Balanced truncation method has a good error bound, see following theorem. The \( H_2 \) norm of the error system is bounded by 2 times the sum of the neglected Hankel singular values. So in general, the reduced system has a good accuracy in approximating the original system.

**Theorem 2.2.2** [35] Suppose \( \sigma_i \), \( i = 1, 2 \cdots , n \) be the Hankel singular values of the system, then the \( H_\infty \)-norm of the error system obtained from balanced truncation method is bounded by twice the summation of the neglected Hankel singular values.

\[ ||\Sigma - \Sigma_{\text{red}}||_{H_\infty} \leq 2(\sigma_{k+1} + \cdots + \sigma_n) \quad (2.2.1) \]

Even though Algorithm 2.1 is numerically stable, it is not efficient since we need to compute SVD of an \( N \times N \) matrix. A number of approximation algorithms have been proposed based on balanced truncation to solve different large scale real problems.
For a given stable system $\Sigma$, the \textbf{Hankel-norm approximation} is to seek stable approximants $\Sigma_*$ satisfying
\[
\sigma_{k+1}(\Sigma) \leq \|\Sigma - \Sigma_*\|_H < \sigma_k(\Sigma) \tag{2.2.2}
\]

\section{2.3 Moment Matching: Krylov Method}

The main idea of moment matching is to match the first $r$ coefficients of some power series expansions of the transfer function $H(s)$ for an appropriate $r$.

For simplicity, in this section we consider SISO systems, i.e. $m = p = 1$ and $\Sigma = \begin{pmatrix} A & b \\ c \end{pmatrix}$. During the discussion, we shall indicate when the methods apply to or can be extended to MIMO systems. A $j$-th Krylov subspace corresponding to some matrix $G$ and vector $g$ is denoted by $\mathcal{K}_j(G, g)$:
\[
\mathcal{K}_j(G, g) = \text{span}\{g, Gg, G^2g, \cdots, G^{j-1}g\}.
\]

By using Neumann expansion \cite{10}
\[
(I - \eta G)^{-1} = \sum_{j=0}^{\infty} (\eta G)^j \tag{2.3.1}
\]
we can expand the transfer function $G(s)$ as following:
\[
G(s) = C(sI - A)^{-1}B = s^{-1}C(I - s^{-1}A)^{-1}B = s^{-1}C \left( \sum_{j=0}^{\infty} (s^{-1}A)^j \right) B = \sum_{j=0}^{\infty} (CA^j B)s^{-(j+1)}
\]
\[ = \mu^0 + \mu^1 s^{-1} + \mu^2 s^{-2} + \mu^3 s^{-3} + \cdots \] (2.3.2)

where the coefficients are known as Markov parameters. It is easy to see that \( \mu_j = CA^{j-1}B \). A reduced order model of size \( k \) whose Markov parameters \( \hat{\mu}_j = \mu_j \) for \( j = 1, 2, \ldots, 2k \), is known as a partial realization [25].

There are three matrices associated with system \( \Sigma \). The observability matrix \( \mathcal{O}_t \) is defined as:

\[
\mathcal{O}_t = \mathcal{O}_t(c, A) = \begin{bmatrix}
  c \\
  cA \\
  cA^2 \\
  \vdots \\
  cA^{t-1}
\end{bmatrix},
\]

the controllability (reachability) matrix \( \mathcal{R}_t \) is:

\[
\mathcal{R}_t = \mathcal{R}_t(A, b) = \begin{bmatrix}
  b & Ab & A^2b & \cdots & A^{t-1}b
\end{bmatrix},
\]

the \( t \times t \) Hankel matrix and its shift are:

\[
\mathcal{H}_t := \begin{pmatrix}
  \mu^1 & \mu^2 & \cdots & \mu^t \\
  \mu^2 & \mu^3 & \cdots & \mu^{t+1} \\
  \vdots & \ddots & \ddots & \vdots \\
  \mu^t & \mu^{t+1} & \cdots & \mu^{2t-1}
\end{pmatrix} \in \mathbb{R}^{t \times t}, \quad \mathcal{\overrightarrow{H}}_t := \begin{pmatrix}
  \mu^2 & \mu^3 & \cdots & \mu^{t+1} \\
  \mu^3 & \mu^4 & \cdots & \mu^{t+2} \\
  \vdots & \ddots & \ddots & \vdots \\
  \mu^{t+1} & \mu^{t+2} & \cdots & \mu^{2t}
\end{pmatrix} \in \mathbb{R}^{t \times t}.
\]

It is easy to see that

\[ \mathcal{H}_t = \mathcal{O}_t \mathcal{R}_t \quad \mathcal{\overrightarrow{H}}_t = \mathcal{O}_t A \mathcal{R}_t. \]
Applying $k$-step **two-sided Lanczos** algorithm to matrix $A \in \mathbb{R}^{n \times n}$ produces three matrices $W, V \in \mathbb{R}^{n \times k}$, $H \in \mathbb{R}^{k \times k}$ and two vectors $f, g \in \mathbb{R}^n$ such that

$$AV = VH + fe_k^*, \quad A^*W = WH^* + ge_k^*$$  \hspace{1cm} (2.3.3)

where $H$ is tridiagonal, $H = W^*AV$, $W^*V = I_k$, $W^*f = 0$, $V^*g = 0$, $W = \text{span} \ \text{col} \ K_k(A, b)$, $V = \text{span} \ \text{col} \ K_k(A^*, c^*)$, $W^*W = I_k$, $V^*V = I_k$, $e_k$ denotes the $k$-th unit vector in $\mathbb{R}^k$. We then use $W$ and $V$ as projection matrices to reduce the system and get

$$\Sigma = \begin{pmatrix} \hat{A} & \hat{b} \\ \hat{c} \\ \end{pmatrix}$$

where

$$\hat{A} = W^*AV \quad \hat{b} = W^*b \quad \hat{c} = cV$$

**Theorem 2.3.1** The reduced system defined above matches the original system for the first $2k$ Markov parameters:

$$\hat{\mu}^j = \mu^j, \quad j = 1, 2, 3, \ldots, 2k.$$ 

Moreover, $\hat{A}$ is tridiagonal, and $\hat{b}, \hat{c}^*$ are multiples of unit vector $e_1$.

**Two-sided Lanczos algorithm**

1. $\beta_1 = \sqrt{|b^*c^*|}$, $\gamma_1 = \text{sgn}(b^*c^*)\beta_1$, $v_1 = b/\beta_1$, $w_1 = c^*/\gamma$

2. for $j = 1, \ldots, k$
3. \( \alpha_j = w_j^* A v_j \)

4. \( r_j = A v_j - \alpha_j v_j - \gamma_j v_{j-1}, q_j = A^* w_j - \alpha_j w_j - \beta_j w_{j-1} \)

5. \( \beta_{j+1} = \sqrt{|r_j^* q_j|}, \gamma_{j+1} = \text{sgn}(r_j^* q_j) \beta_{j+1} \)

6. \( v_{j+1} = r_j / \beta_{j+1}, w_{j+1} = q_j / \gamma_{j+1} \).

7. end for

8. end

From the above two-sided Lanczos algorithm, we have

\[
V_k = \begin{pmatrix} v_1 & v_2 & \cdots & v_k \end{pmatrix}, W_k = \begin{pmatrix} w_1 & w_2 & \cdots & w_k \end{pmatrix}
\]

such that

\[
AV_k = V_k H_k + \beta_{k+1} v_{k+1} e_k^* , \quad A^* W_k = W_k H_k^* + \gamma_{k+1} w_{k+1} e_k^*
\]

\[
H_k = \begin{pmatrix} \alpha_1 & \gamma_2 \\
\beta_2 & \alpha_2 & \ddots \\
& \ddots & \ddots & \gamma_k \\
& & \beta_k & \alpha_k
\end{pmatrix}
\]

\[
r_k \in \mathcal{R}_{k+1}(A, b), \quad q_k^* \in \mathcal{Q}_{k+1}(c, A)
\]

The \( k \)-step Arnoldi factorization of a matrix \( A \in \mathbb{R}^{n \times n} \) produces two matrices \( V \in \mathbb{R}^{n \times k}, H \in \mathbb{R}^{k \times k} \), a vector \( f \in \mathbb{R}^n \) such that:

\[
AV = VH + fe_k^*
\]
where $H$ is in upper Hessenberg form, $H = V^* AV$, $V = \text{span} \text{col} K_k(A, b)$, $V^* V = I_k$, $V^* f = 0$, $e_k$ is the $k$-th unit vector in $\mathbb{R}^k$. We then use $V$ as the orthogonal projection matrix to get the reduced system,

$$\hat{\Sigma} = \begin{pmatrix} \hat{A} & \hat{b} \\ \hat{c} \end{pmatrix}$$

where

$$\hat{A} = V^* AV \quad \hat{b} = V^* b \quad \hat{c} = cV$$

**Theorem 2.3.2** The reduced system $\hat{\Sigma}$ defined above matches the original system for the first $k$ Markov parameters,

$$\hat{\mu}_j = \mu_j, \quad j = 1, 2, \ldots, k.$$  

Moreover, $\hat{A}$ is upper Hessenberg, $\hat{b}$ is a multiple of the unit vector $e_1$.

The Arnoldi algorithm

1. $v_1 = \frac{b}{\|b\|}, \ w = Av_1, \ \alpha_1 = v_1^* w, \ f_1 = w - v_1 \alpha_1, \ V_1 = [v_1], \ H_1 = [\alpha_1]$

2. for $j = 1, 2, \ldots, k - 1$

3. $\beta_j = \|f_j\|, \ v_{j+1} = \frac{f_j}{\beta_j}$

4. $V_{j+1} = \begin{bmatrix} V_j & v_{j+1} \end{bmatrix}, \ \hat{H}_j = \begin{bmatrix} H_j \\ \beta_j e_j^* \end{bmatrix}$

5. $w = Av_{j+1}, \ h = V_{j+1}^* w, \ f_{j+1} = w - V_{j+1} h$
6. \[ H_{j+1} = \begin{bmatrix} \hat{H}_j & h \end{bmatrix} \]

7. end for

8. end

**Remarks on Lanczos and Arnoldi procedures:**

1. The number of operations is \( O(kn^2) \), only matrix-vector multiplications, no matrix-matrix multiplications, no matrix factorizations. For gramian based or SVD based method, the cost is \( O(n^3) \). So Krylov projection based method is relatively efficient and favorable for large scale settings.

2. There is no need to compute the reduced order model by matrix-matrix product: \( \hat{A} = W^*AV \), \( \hat{A} \) is obtained directly from Lanczos or Arnoldi algorithms, which is \( H \) in (2.3.3) and (2.3.4). This eliminates ill-conditioning.

3. Lanczos breaks down if det \( \mathcal{H}_t = 0 \). The remedy is Look-ahead Lanczos [19] [36].

4. The reduced system \( \hat{\Sigma} \) may not be stable even if the original system is stable. The remedy is to use implicitly restarted Lanczos or Arnoldi to eliminate unwanted eigenvalues [50].

The expansion of \( G(s) \) in (2.3.2) is in the form of partial realization which emphasizes behavior at \( t = 0 \), such a model is for the extremely rapidly decaying dynamics of the system according to [26], but it may not accurately reproduce the behavior at
some later time. The remedy is to consider the power series expansion at \( s = 0 \) or use a shift \( \sigma \) which are known as Padé and shifted Padé expansions, respectively:

\[
G(s) = c(sI - A)^{-1}b = -c(I - sA^{-1})^{-1}A^{-1}b \\
= -c \left( \sum_{j=0}^{\infty} (sA^{-1})^j \right) A^{-1}b \\
= \sum_{j=0}^{\infty} (-cA^{-(j+1)}b) s^j \\
= \mu_0^0 + \mu_1^1 s + \mu_2^2 s^2 + \mu_3^3 s^3 + \cdots \tag{2.3.5}
\]

which is Padé expansion and \( \mu_0^j = -cA^{-(j+1)}b \), and

\[
G(s) = c(sI - A)^{-1}b = c(sI - \sigma I + \sigma I - A)^{-1}b \\
= -c \left( I - (s - \sigma)(A - \sigma I)^{-1} \right)^{-1} (A - \sigma I)^{-1}b \\
= -c \left( \sum_{j=0}^{\infty} ((s - \sigma)(A - \sigma I)^{-1})^j \right) (A - \sigma I)^{-1}b \\
= \sum_{j=0}^{\infty} (-c(A - \sigma I)^{-(j+1)}b) s^j \\
= \mu_0^\sigma + \mu_1^\sigma (s - \sigma) + \mu_2^\sigma (s - \sigma)^2 + \mu_3^\sigma (s - \sigma)^3 + \cdots \tag{2.3.6}
\]

which is shifted Padé expansion and \( \mu_0^\sigma = -c(A - \sigma I)^{-(j+1)}b \). The shift \( \sigma \) is called an interpolation point.

(2.3.2) and (2.3.5) can be looked as special cases of shifted Padé (2.3.6) with shift \( \sigma = \infty \) and \( \sigma = 0 \), respectively. So partial realization associated with (2.2.2) produces a reduced order model which tends to approximate the original model for high frequencies, the error may be significant for other frequencies, especially for \( \sigma = 0 \), the steady-state error.
The algorithms for getting projection matrices from Padé or shifted Padé expansions are very similar to that of partial realization, in the above Lanczos and Arnoldi algorithms, we simply use \((\sigma I - A)^{-1}\) to replace \(A\), \((\sigma I - A)^{-1}b\) to replace \(b\). Note that we need to use LU factorization for \((\sigma I - A)^{-1}\) once in Lanczos or Arnoldi algorithms.

The Krylov projection method we outlined above can only be used to match the moments for a single point. In order to approximate the original system in a wider frequency range, we may consider to use multiple interpolation points \(\sigma_k, k = 1, 2, \cdots, K\), which is known as multipoint Padé or rational interpolant:

\[
G(s) = \sum_{j_k=0}^{\infty} (-c(A - \sigma_k I)^{-j_k}b) s^{j_k}
= \mu_{\sigma_k}^0 + \mu_{\sigma_k}^1 (s - \sigma_k) + \mu_{\sigma_k}^2 (s - \sigma_k)^2 + \mu_{\sigma_k}^3 (s - \sigma_k)^3 + \cdots \tag{2.3.7}
\]

where \(\mu_{\sigma_k}^{j_k} = -c(A - \sigma_k I)^{-j_k}(A - \sigma_k I)^{-1} b, (k = 1, 2, \cdots, K)\). This may be implemented through the rational Krylov method by Ruhe [41].

The goal is still to find the projection matrices \(W, V \in \mathbb{C}^{n \times k}\) as before.

**Lemma 2.3.3** [20] *Given two matrices \(G \in \mathbb{C}^{n \times n}\), \(H \in \mathbb{C}^{n \times m}\), let \(V \in \mathbb{C}^{n \times k}\) be a full rank matrix such that \(K_j(G, H) \subseteq \mathcal{V} = \text{span} \text{ col} (V)\), and let \(W\) be any matrix in \(\mathbb{R}^{n \times k}\) orthogonal to \(V\), \(W^TV = I_k\). Then the projected matrices \(\hat{G} = W^TGV\) and \(\hat{H} = W^TH\) satisfy*

\[
G^iH = V\hat{G}^i\hat{H}, \quad i = 0, 1, 2, \cdots, j - 1.
\]

Based on the above lemma, the following theorem holds for SISO systems [26] and for MIMO systems [20].
Theorem 2.3.4 [26] If

\[ \bigcup_{k=1}^{K} \mathcal{V} \left( (A - \sigma_k I)^{-1}, (A - \sigma_k I)^{-1} B \right) \subseteq \mathcal{V} = \text{span} \, \text{col}(V) \]  

(2.3.8)

and

\[ \bigcup_{k=1}^{K} \mathcal{V} \left( (A - \sigma_k I)^{-T}, (A - \sigma_k I)^{-T} C^T \right) \subseteq \mathcal{W} = \text{span} \, \text{col}(W) \]  

(2.3.9)

where \( W^T V = I \) and \( \sigma_k \) are chosen so that \( (A - \sigma_k I) \) are invertible for any \( k \), \( 1 \leq k \leq K \), then the moments of \( \Sigma \) and \( \hat{\Sigma} \) satisfy

\[ \mu_{\sigma_k}^{j_k} = \hat{\mu}_{\sigma_k}^{j_k}, \text{ for } j_k = 1, 2, \cdots, b_k + c_k \text{ and } k = 1, 2, \cdots, K \]  

(2.3.10)

provided the matrices \( (\sigma_k I - \hat{A}_r) \) are invertible.

For implementations of the above theorem to find the projection matrices \( W \) and \( V \) in an efficient and numerically stable way, see [26] for three algorithms: dual rational Arnoldi algorithm, rational Lanczos method, and rational power Krylov method.
Chapter 3

Second Order Linear Dynamical Systems

3.1 Introduction

Models of linear dynamical systems are often given in second order form

\[ M\ddot{q}(t) + G\dot{q}(t) + Kq(t) = B_0u(t) \]
\[ y(t) = E_0q(t) + F_0\dot{q}(t) \]  \hspace{1cm} (3.1.1)

where \( M, G, K \in \mathbb{R}^{n \times n}, B_0 \in \mathbb{R}^{n \times m}, E_0, F_0 \in \mathbb{R}^{p \times n} \) are given matrices, \( u(t) \in \mathbb{R}^m \) is the given vector of inputs, \( y(t) \in \mathbb{R}^p \) is the unknown vector of outputs, \( q(t) \in \mathbb{R}^n \) is the unknown vector of internal variables, \( n \) is the dimension of the system, and \( M \) is assumed to be invertible.

Models of mechanical systems in particular are usually of second order form (3.1.1). For such a system, \( M = M^T, G \) and \( K = K^T \) are respectively the mass,
damping and stiffness matrices, \( u(t) \in \mathbb{R}^m \) is the input, \( B_0(t) u(t) = f(t) \in \mathbb{R}^n \) is the vector of external forces, and \( q(t) \in \mathbb{R}^n \) is the vector of internal variables (see [40] [54] [11] [57] for more information on such models).

In civil engineering or aeronautics, the size \( n \) of the system (obtained using for instance finite elements techniques [40] [54]) is often so large that many analysis and design problems can not be solved within a reasonable computing time. It is then advisable to construct a reduced order system that keeps the "mechanical" structure of the system.

If we let \( x = \begin{bmatrix} q \\ \dot{q} \end{bmatrix} \), (3.1.1) is then transformed to a corresponding \( 2n \)-dimensional first-order system

\[
\Sigma := \begin{pmatrix} A & B \\ C & D \end{pmatrix}
\]
i.e.

\[
\dot{x}(t) = Ax(t) + Bu(t) \\
y(t) = Cx(t) + Du(t)
\]

where

\[
A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}G + K_M & G_M \end{bmatrix} = \begin{bmatrix} 0 & I \\ K_M & G_M \end{bmatrix} \\
B = \begin{bmatrix} 0 \\ M^{-1}B_0 \\ B_M \end{bmatrix} = \begin{bmatrix} 0 \\ B_M \end{bmatrix}
\]

(3.1.2)
\[ C = \begin{bmatrix} E_0 & F_0 \end{bmatrix} \tag{3.1.4} \]
\[ D = 0 \tag{3.1.5} \]

In this chapter, we assume second order system (3.1.1) has dimension \( n \), the transformed first order system has dimension \( N \), therefore \( N = 2n \).

Since the theory and methods for first order model reduction are well-developed, people often study second order system of (3.1.1) by transforming to first order form, and apply the techniques of first order model reduction. However this is not efficient since the transformed first order system has double dimension. More importantly the reduced-order corresponding first order system generally loses the second order structure and the corresponding physical interpretation.

Presently very few algorithms exist for second order model reduction (see [34] [52] [53] [18]). Existing methods are typically based on two approaches: balanced truncation and Krylov projection. These methods appear to perform well in experiments, but no error bounds have been provided for any of them.

The objective of this thesis is to develop algorithms for direct second order model reduction which satisfy the following:

- the reduced order system preserves the second order structure;
- the reduced order system has sufficient accuracy in approximating the original system, \( \frac{\|y-y_{\text{real}}\|}{\|u\|} < tol \);
- the algorithms are computationally efficient and numerically stable;
the reduced system preserves some important properties such as stability, passivity etc..

Recall that first order model reduction is mainly done by projection method:

**First order model reduction**

**Step 1.** Form the projection matrices \( W, V \in \mathbb{R}^{N \times k} \) such that \( W^T V = I_k \).

**Step 2.** Perform the projection to \((A, B, C, D)\) and get \((\hat{A}, \hat{B}, \hat{C}, \hat{D})\):

\[
\hat{A} = W^T AV, \quad \hat{B} = W^T B, \quad \hat{C} = CV, \quad \hat{D} = D.
\]

For a given second order system \((3.1.1)\), consider the transformed first order system, it is easy to see a sufficient condition for reduced order system to preserve second order structure is the projection matrices are in the following form:

\[
W = \begin{pmatrix} W_q \\ W_q \end{pmatrix}, \quad V = \begin{pmatrix} V_q \\ V_q \end{pmatrix}
\]

where \( W, V \in \mathbb{R}^{N \times k}, W_q, V_q \in \mathbb{R}^{n \times r}, k = 2r, W_q^T V_q = I_r \), and therefore \( W^T V = I_k \).

So in reduced system,

\[
\hat{A} = W^T AV = \begin{pmatrix} W_q^T \\ W_q^T \end{pmatrix} \begin{pmatrix} 0 & I \\ -M^{-1}K & -M^{-1}G \end{pmatrix} \begin{pmatrix} V_q \\ V_q \end{pmatrix} = \begin{pmatrix} 0 & I_r \\ -W_q^T M^{-1}KV_q & -W_q^T M^{-1}GV_q \end{pmatrix}
\]

\[
\hat{B} = WB = \begin{pmatrix} W_q^T \\ W_q^T \end{pmatrix} \begin{pmatrix} 0 \\ M^{-1}B_0 \end{pmatrix} = \begin{pmatrix} 0 \\ W_q^T M^{-1}B_0 \end{pmatrix}
\]
\[
\hat{C} = CV = \begin{bmatrix} E_0 & F_0 \\ \end{bmatrix} \begin{bmatrix} V_q \\ V_q \end{bmatrix} = \begin{bmatrix} E_0 V_q & F_0 V_q \end{bmatrix}.
\]

In this thesis, we mainly focus on developing algorithms through the following projection methods:

**Second order model reduction**

**Step 1.** Form the projection matrices \( W_q, V_q \in \mathbb{R}^{n \times r} \) such that \( W_q^T V_q = I_r \).

**Step 2.** Perform the projection to \( (M, G, K, B_0, E_0, F_0) \) and get \( (\hat{M}, \hat{G}, \hat{K}, \hat{B}_0, \hat{E}_0, \hat{F}_0) \):

\[
\hat{M} = I_r, \quad \hat{G} = W_q^T M^{-1} G V_q, \quad \hat{K} = W_q^T M^{-1} K V_q,
\]

\[
\hat{B}_0 = W_q^T M^{-1} B_0, \quad \hat{E}_0 = E_0 V_q, \quad \hat{F}_0 = F_0 V_q.
\]

### 3.2 Second Order Linear Dynamical Systems

Transfer function, stability, gramians, and system \( \mathcal{H}_2 \) norm etc. are important quantities in first order linear systems. In this section, we study these concepts in second order linear systems which are in some special forms, we also derive some properties for these quantities.

By assuming zero initial conditions, \( q(0) = 0 \) and \( \dot{q}(0) = 0 \), if we take Laplace
transformation for the second order system (3.1.1), then

\[
\int_0^\infty \dot{q}(t)e^{-st} \, dt = q(t)e^{-st}|_0^\infty - \int_0^\infty q(t) \, de^{-st} = 0 + s \int_0^\infty q(t)e^{-st} \, dt = sQ(s),
\]

\[
\int_0^\infty \ddot{q}(t)e^{-st} \, dt = \dot{q}(t)e^{-st}|_0^\infty - \int_0^\infty \dot{q}(t) \, de^{-st} = 0 + s \int_0^\infty \dot{q}(t)e^{-st} \, dt = s^2Q(s).
\]

So in frequency domain, (3.1.1) becomes to be the following,

\[
\begin{align*}
s^2M Q(s) + sGQ(s) + KQ(s) &= B_0U(s) \\
Y(s) &= E_0Q(s) + sF_0Q(s)
\end{align*}
\]

(3.2.1)

Eliminating \(Q(s)\) in (3.2.1), we get the input-output relation in Laplace domain,

\[
Y(s) = (E_0 + sF_0) (s^2M + sG + K)^{-1} B_0U(s).
\]

(3.2.2)

The **transfer function** for second order system (3.1.1) is then,

\[
H(s) = (E_0 + sF_0) (s^2M + sG + K)^{-1} B_0.
\]

(3.2.3)

**Stability** is another important concept for linear systems. All poles of the system must lie in the open left half plane and these are characterized as solutions to an eigenvalue problem. For second order system, stability can be characterized by the quadratic eigenvalue problem.
The **quadratic eigenvalue problem** [51] is to find scalars $\lambda$ and nonzero vectors $x$ and $y$ such that

$$(\lambda^2 M + \lambda G + K)x = 0, \quad y^T(\lambda^2 M + \lambda G + K) = 0$$

where $M, G, K \in \mathbb{R}^{n \times n}$, and $x, y$ are the right and left eigenvectors corresponding to the eigenvalue $\lambda$, i.e. the roots (zeros) of $\det(Ms^2 + Gs + K)$.

For simplicity, we denote

$$P(s) = Ms^2 + Gs + K \quad (3.2.4)$$

**Theorem 3.2.1** ([51]) *Given a second order system (3.1.1), suppose the transformed first order system is*

$$
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix}
$$

*Then $\lambda$ is an eigenvalue of $A$ if and only if $\lambda$ is a root of $\det(P(s))$.*

Actually it is not hard to verify that $(\lambda, x)$ is an eigenpair of quadratic matrix polynomial $P(s)$ if and only if $(\lambda, x, \lambda x)$ is an eigenpair of $A$.

From Theorem 2.1.2, we know that a first order system is stable if and only if all eigenvalues of $A$ have negative real parts, so from Theorem 3.2.1, the stability of second order systems can be characterized by quadratic eigenvalue problems which is stated in the following corollary.

**Corollary 3.2.2** *Second order system (3.1.1) is asymptotically stable $\iff$ any eigenvalue $\lambda$ of quadratic matrix polynomial $Ms^2 + Gs + K$ has negative real part, i.e. $\text{Re}(\lambda) < 0$.***
Next theorem gives a sufficient condition for second order system to be stable.

**Theorem 3.2.3** [51] *If* $M, G, K$ *are all symmetric positive definite, then any eigenvalue* $\lambda$ *of quadratic matrix polynomial* $Ms^2 + Gs + K$ *has negative real part, i.e. Re$(\lambda) < 0$.*

As mentioned earlier, most models in mechanic engineering are in second order form of (3.1.1), and satisfy $M, K$ being symmetric positive definite. If $G$ is also symmetric positive definite, from Theorem 3.2.3, we then know these systems must be stable.

In the theory of first order linear systems, gramians play important roles. There are several ideas concerning a proper notation of gramians in second order linear systems.

For a given second order system (3.1.1), suppose $P, Q \in \mathbb{R}^{2n \times 2n}$ are controllability and observability gramians for the corresponding first order system. Partition $P$ and $Q$ each into four equal blocks, with each block of size $n \times n$.

\[ P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} \quad \text{(3.2.5)} \]

\[ Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{pmatrix} \quad \text{(3.2.6)} \]

Under the assumption that $u(\tau) = \delta(\tau)$ and $x(0) = 0$, controllability gramian $P$ for first order system is:

\[ P = \int_0^\infty x(\tau)x(\tau)^T d\tau \]
Note that for a given second order system $\Sigma_2$, if we let $x = \begin{bmatrix} q \\ \dot{q} \end{bmatrix}$, then a second order system can be transformed into a first order system. Similarly we can also define the controllability gramian $P_q$ for second order system as:

$$P_q = \int_0^\infty q(\tau)q(\tau)^T d\tau$$

It is easy to see $P_q = P_{11}$, i.e. the second order controllability gramian $P_q$ is actually the first $n$ by $n$ principle submatrix of the corresponding first order controllability gramian $P$. This is one idea in defining the second order controllability gramian. In Chapter 5, we will use this idea to provide the error bound for some gramian based algorithms.

According to Glover [22], the controllability gramian $P$ of first order system (2.1.1) arises from the optimization problem

$$\min_{u(t)\in L_2[-\infty,0]} \int_{-\infty}^{0} u^2(t) dt$$

subject to

$$\dot{x} = Ax + Bu \quad x(0) = x_0$$

for which the optimal value is $x_0^T P^{-1} x_0$. Symmetrically we can get the observability gramian $Q$ from the dual optimization problem

$$\min_{y(t)\in L_2[0,\infty]} \int_{0}^{\infty} y^2(t) dt$$

subject to

$$\dot{x} = A^T x + C^T u \quad y(t) = B^T x(t) \quad x(0) = x_0 \quad u(t) = 0 \; (t > 0)$$
for which the optimal value is $x_0^TQ^{-1}x_0$

Analogous to this, Meyer and Srinivasan in [34], VanDooren etc. in [53] propose that the optimization problem associated with the second order system is,

$$\min_{\dot{q}(0) = q_0 \in \mathbb{R}^n} \min_{u(t) \in L_T^2(-\infty, 0]} \int_{-\infty}^{0} u^2(t) \, dt$$

subject to

$$M\ddot{q}(t) + G\dot{q}(t) + Kq(t) = B_0u(t), \quad q(0) = q_0$$

and the optimal value is $q_0^TP_{11}^{-1}q_0$. The solution of the dual problem has optimal value $q_0^TQ_{11}^{-1}q_0$. According to Meyer and Srinivasan [34], in this case, the system has no requirement in velocity at $t = 0$, and we call $P_{11}$ and $Q_{11}$ the free velocity gramians of second order systems; similarly the zero velocity second order gramians can be defined by the following optimization problem, in which the system has zero velocity at $t = 0$:

$$\min_{u(t) \in L_T^2(-\infty, 0]} \int_{-\infty}^{0} u^2(t) \, dt$$

subject to

$$M\ddot{q}(t) + G\dot{q}(t) + Kq(t) = B_0u(t), \quad q(0) = q_0, \quad \dot{q}(0) = 0$$

and the optimum is $q_0^T(P_{11} - P_{12}P_{22}^{-1}P_{12}^T)^{-1}q_0$, the optimum for the dual problem is $q_0^T(Q_{11} - Q_{12}Q_{22}^{-1}Q_{12}^T)^{-1}q_0$, and so the zero velocity gramians are defined as

$$P_{ZV} = P_{11} - P_{12}P_{22}^{-1}P_{12}^T, \quad Q_{ZV} = Q_{11} - Q_{12}Q_{22}^{-1}Q_{12}^T$$
compare to the free velocity gramians:

\[ P_{FV} = P_{11}, \quad Q_{FV} = Q_{11} \]

VanDooren etc. in [53] notice that the optimization problem

\[
\min_{\dot{q}_0} \min_{u(t)} \int_{-\infty}^{0} u^2(t) \, dt \tag{3.2.11}
\]

subject to

\[ M\ddot{q}(t) + G\dot{q}(t) + Kq(t) = B_0 u(t), \quad y(t) = E_0 q(t) + F_0 \dot{q}(t), \quad \dot{q}(0) = \dot{q}_0, \quad u(t) = 0 \quad (t > 0) \]

has the optimal value \( \dot{q}_0^T P_{22}^{-1} \dot{q}_0 \), and the dual problem has optimum \( q_0^T Q_{22}^{-1} \dot{q}_0 \). So the pair \( (P_{11}, Q_{11}) \) represents the information of second order systems corresponding to positions \( q(t) \), and the pair \( (P_{22}, Q_{22}) \) represents the information corresponding to velocities \( \dot{q}(t) \).

For a given second order system, the following theorem gives some observations on structures of gramians \( P \) and \( Q \) in the corresponding first order system.

**Theorem 3.2.4** Given second order system (3.1.1), suppose \( P \) and \( Q \) are the controllability and observability gramians for the transformed first order system. Then the following statements hold:

1. There is a coordinate system for which \( P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} \), where \( P_{12} \) is real, tridiagonal, and skew symmetric matrix, and hence all diagonal elements are 0.

2. If \( F_0 = 0 \), then in any coordinate system, there is no state space transformation which will give a block diagonal \( Q \).
**Proof:** Suppose the Schur decomposition of $P_{12}$ is $P_{12} = U P_{12s} U^T$ where $U$ is orthogonal matrix, $P_{12s}$ is in real Schur form, i.e. quasi upper triangular. Let the transformation (contragradient transformation) matrix be

$$W = \begin{pmatrix} U \\ U \end{pmatrix},$$

and the projection matrix be

$$W_r = \begin{pmatrix} U_r \\ U_r \end{pmatrix},$$

where $U_r$ consists of the first $r$ columns of $U$. Then

$$W^T PW^{-T} = \begin{pmatrix} U^T \\ U^T \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} \begin{pmatrix} U \\ U \end{pmatrix}$$

$$= \begin{pmatrix} U^T P_{11} U & U^T P_{12} U \\ U^T P_{12}^T U & U^T P_{22} U \end{pmatrix} = \begin{pmatrix} U^T P_{11} U & P_{12s} \\ P_{12s}^T & U^T P_{22} U \end{pmatrix}$$

For simplicity, we may suppose $P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}$, where $P_{12}$ is in real Schur form, i.e. $P_{12}$ is quasi upper triangular (since there is a coordinate system on which $P$ is in this form). From Lyapunov equation

$$AP + PA^T + BB^T = 0,$$

we have

$$\begin{pmatrix} 0 & I \\ K_M & G_M \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} + \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix} \begin{pmatrix} 0 & K_M^T \\ I & G_M^T \end{pmatrix} + \begin{bmatrix} 0 \\ B_M \end{bmatrix} \begin{bmatrix} 0 & B_M^T \end{bmatrix} = 0$$
or
\[
\begin{pmatrix}
P_{12}^T & P_{22} \\
K_M P_{11} + G_M P_{12}^T & K_M P_{12} + G_M P_{22}
\end{pmatrix} + 
\begin{pmatrix}
P_{12} & P_{11} K_M^T + P_{12} G_M^T \\
P_{22} & P_{12} K_M^T + P_{22} G_M^T
\end{pmatrix} + 
\begin{pmatrix}
0 & 0 \\
0 & B_M B_M^T
\end{pmatrix} = 0.
\]
Equating each block,
\[
P_{12}^T + P_{12} = 0 \quad (3.2.12)
\]
\[
K_M P_{11} + G_M P_{12}^T + P_{22} = 0 \quad (3.2.13)
\]
\[
K_M P_{12} + G_M P_{22} + P_{12} K_M^T + P_{22} G_M^T + B_M B_M^T = 0 \quad (3.2.14)
\]
Recall that \( P_{12} \) is in real Schur form, from (3.2.12) one can get that \( P_{12} \) is tridiagonal with diagonal elements all 0, and \( P_{12} \) is skew symmetric matrix.

From the observability Lyapunov equation
\[
A^T Q + QA + C^T C = 0, \quad (3.2.15)
\]
one can get
\[
\begin{pmatrix}
0 & K_M^T \\
I & G_M^T
\end{pmatrix} 
\begin{pmatrix}
Q_{11} & Q_{12} \\
Q_{12}^T & Q_{22}
\end{pmatrix} + 
\begin{pmatrix}
Q_{11} & Q_{12} \\
Q_{12}^T & Q_{22}
\end{pmatrix} 
\begin{pmatrix}
0 & I \\
K_M & G_M
\end{pmatrix} + 
\begin{bmatrix}
C_M^T \\
C_M 
\end{bmatrix} = 0
\quad (3.2.16)
\]
Equating each block of (3.2.16) gives
\[
K_M^T Q_{12} + Q_{12} K_M + C_M^T C_M = 0 \quad (3.2.17)
\]
\[
K_M^T Q_{22} + Q_{11} + Q_{12} G_M = 0 \quad (3.2.18)
\]
\[
Q_{12} + G_M^T Q_{22} + Q_{12}^T + Q_{22} G_M = 0 \quad (3.2.19)
\]
From (3.2.17), we know $Q_{12}$ is not zero, otherwise $C_M = 0$, and then from (3.2.15) we get the solution for this Lyapunov equation is $Q = \int_0^\infty e^{AT} C^T C e^{Ar} \, dt = 0$. This contradicts with $Q$ being symmetric positive definite. So $Q$ is not block diagonal.

$\mathcal{H}_2$ norm is an important quantity for bounding linear systems.

**Lemma 3.2.5** [3] $\mathcal{H}_2$ norm of a first order system $\Sigma = \begin{pmatrix} A & B \\ C \\ D \end{pmatrix}$ can be expressed as

$$||\Sigma||_{\mathcal{H}_2}^2 = tr\{CPCT\} = tr\{B^TQB\} \quad (3.2.20)$$

The $\mathcal{H}_2$ norm for second order system (3.1.1) with $F_0 = 0$ is stated in the following proposition.

**Proposition 3.2.6** Given a second order system (3.1.1) with $F_0 = 0$, suppose $\Sigma = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is the transformed first order system, where

$$A = \begin{bmatrix} 0 & I \\ K_M & G_M \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_M \end{bmatrix}, \quad C = \begin{bmatrix} C_M & 0 \end{bmatrix}, \quad D = 0$$

Then the following holds:

$$||\Sigma||_{\mathcal{H}_2}^2 = tr\{B_M^T Q_{22} B_M\} = tr\{C_M P_{11} C_M^T\} \quad (3.2.21)$$

**Proof:** From Lemma 3.2.5,

$$||\Sigma||_{\mathcal{H}_2}^2 = \text{trace}\left[ \begin{bmatrix} 0 & B_M^T \\ Q_{12} & Q_{11} \end{bmatrix} \begin{bmatrix} 0 \\ B_M \end{bmatrix} \begin{bmatrix} 0 \\ Q_{12} \end{bmatrix} \begin{bmatrix} 0 \\ B_M \end{bmatrix} \right]$$
\[
= \text{trace}\{ \begin{bmatrix} B_M^T Q_{12}^T & B_M^T Q_{22} \\ B_M \\ 0 \end{bmatrix} \} = \text{trace}\{ B_M^T Q_{22} B_M \}
\]

and

\[
\|\Sigma\|_{\mathcal{H}_2}^2 = \text{trace}\{ \begin{bmatrix} C_M & 0 \\ P_{11} & P_{12} \\ P_{12}^T & P_{22} \\ C_M^T \end{bmatrix} \} = \text{trace}\{ C_M P_{11} \}
\]

From the discussions in this section, we can see some important quantities in second order linear systems have some special forms. Stability is closely related to quadratic eigenvalue problems; under the assumption of \( F_0 = 0 \), system \( \mathcal{H}_2 \) norm can be characterized by the parameters \( M^{-1} B_0 \) and \( Q_{22} \), or the parameters \( E_0 \) and \( P_{11} \); the gramians for the corresponding first order system have some special properties. From Theorem 3.2.4, we know that under the assumption of \( F_0 = 0 \), no state space transformation gives a block diagonal \( Q \) in its corresponding first order system. So the first order balanced truncation method is not accurate when applied to second order model reduction, while some existing methods on second order model reduction use the idea of first order balanced truncation, i.e. try to diagonalize \( P \) and \( Q \) in the same time. From Theorem 3.2.4, we can also see that the technique in deriving the error bound for first order balanced truncation method cannot be extended to second order model reduction.
Chapter 4

Second Order Model Reduction

Based on Balanced Truncation and

Singular Value Decomposition

In this chapter, first we present algorithms on second order model reduction based upon two approaches: balanced truncation and singular value decomposition. Then we discuss the equivalence of transformed systems to original systems, and the difference of the $\mathcal{H}_2$ norms of the original and reduced systems.
4.1 Algorithms of Second Order Model Reduction

Based on Balanced Truncation

Recall the notation, for a given second order system,

\[
\begin{align*}
M\ddot{q}(t) + G\dot{q}(t) + Kq(t) &= B_0u(t) \\
y(t) &= E_0q(t) + F_0\dot{q}(t)
\end{align*}
\]

or

\[
\begin{align*}
M\ddot{q}(t) + G\dot{q}(t) + Kq(t) &= B_0u(t) \\
y(t) &= E_0q(t)
\end{align*}
\]

by assuming \(F_0 = 0\), suppose \(P\) and \(Q\) are the controllability and observability gramians for the corresponding first order system, partition \(P\) and \(Q\) into four equal size blocks:

\[
P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{pmatrix}, \quad Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{pmatrix}
\]

From (3.2.9), (3.2.10) and their dual problems, we know the pair \((P_{11}, Q_{11})\) represents the information on second order systems corresponding to positions \(q(t)\), and the pair \((P_{22}, Q_{22})\) represents the information corresponding to the velocities \(\dot{q}(t)\).

Papers [34] and [53] use this idea to provide variants of balanced truncation and give two algorithms. The main idea of the algorithm in [34] is to balance \(P_{11}\) and \(Q_{11}\) to get the balanced transformation matrices \(W_1\) and \(V_1\) where \(W_1^TP_{11}W_1 = V_1^TQ_{11}V_1 = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)\) and \(W_1^TV_1 = I_n\). Then use \(W = \begin{pmatrix} W_1r \\ W_{1r} \end{pmatrix}\) and
\[ V = \begin{pmatrix} V_{1r} \\ V_{1r} \end{pmatrix} \] as projection matrices for the corresponding first order system, where \( W_{1r} \) and \( V_{1r} \) consist of the first \( r \) columns of \( W_1 \) and \( V_1 \) respectively. Then for the corresponding first order system, the reduced system with order \( k = 2r \) is:

\[
\begin{aligned}
\hat{A} &= W^TAV = \begin{pmatrix} W_{1r}^T \\ W_{1r}^T \end{pmatrix} \begin{pmatrix} 0 & I \\ -M^{-1}K & -M^{-1}G \end{pmatrix} \begin{pmatrix} V_{1r} \\ V_{1r} \end{pmatrix} \\
&= \begin{pmatrix} 0 & I \\ -W_{1r}^{-1}M^{-1}KV_{1r} & -W_{1r}^{-1}M^{-1}GV_{1r} \end{pmatrix} \\
\hat{B} &= WB = \begin{pmatrix} W_{1r}^T \\ W_{1r}^T \end{pmatrix} \begin{bmatrix} 0 \\ M^{-1}B_0 \end{bmatrix} = \begin{bmatrix} 0 \\ W_{1r}^TM^{-1}B_0 \end{bmatrix} \\
\hat{C} &= CV = \begin{bmatrix} E_0 & F_0 \end{bmatrix} \begin{pmatrix} V_{1r} \\ V_{1r} \end{pmatrix} = \begin{bmatrix} E_0V_{1r} & F_0V_{1r} \end{bmatrix}
\end{aligned}
\]

Therefore we get the reduction rules for second order system,

\[
\begin{aligned}
\hat{M} &= I_r \\
\hat{K} &= W_{1r}^TM^{-1}KV_{1r} \\
\hat{G} &= W_{1r}^TM^{-1}GV_{1r} \\
\hat{B}_0 &= W_{1r}^TM^{-1}B_0 \\
\hat{E}_0 &= E_0V_{1r} \\
\hat{F}_0 &= F_0V_{1r}
\end{aligned}
\]

(4.1.4)

Similar idea is used in the algorithm of paper [53], which is to balance \( P_{11} \) and \( Q_{11} \) to get the balancing transformation matrices \( W_1 \) and \( V_1 \), balance \( P_{22} \) and \( Q_{22} \) to
get the balanced transformation matrices \( W_2 \) and \( V_2 \). Then use \( W = \begin{pmatrix} W_{1r} \\ W_{2r} \end{pmatrix} \)

and \( V = \begin{pmatrix} V_{1r} \\ V_{2r} \end{pmatrix} \) as balanced truncation matrices for the corresponding first

order system, where \( W_{1r}, W_{2r}, V_{1r}, V_{2r} \) consist of the first \( r \) columns of \( W_1, W_2, V_1 \)

\( V_2 \) respectively, and so \( W^T V = I_k \). Then the reduced system with order \( k = 2r \) is:

\[
\begin{align*}
\hat{A} &= W^T AV = \begin{pmatrix} W_{1r}^T \\ W_{2r}^T \end{pmatrix} \begin{pmatrix} 0 & I \\ -M^{-1}K & -M^{-1}G \end{pmatrix} \begin{pmatrix} V_{1r} \\ V_{2r} \end{pmatrix} \\
&= \begin{pmatrix} 0 & W_{1r}^T V_{2r} \\ -W_{2r}^T M^{-1}K V_{1r} & -W_{2r}^T M^{-1}G V_{2r} \end{pmatrix} \\
\hat{B} &= WB = \begin{pmatrix} W_{1r}^T \\ W_{2r}^T \end{pmatrix} \begin{bmatrix} 0 \\ M^{-1}B_0 \end{bmatrix} = \begin{bmatrix} 0 \\ W_{2r}^T M^{-1}B_0 \end{bmatrix} \\
\hat{C} &= CV = \begin{bmatrix} E_0 & F_0 \end{bmatrix} \begin{pmatrix} V_{1r} \\ V_{2r} \end{pmatrix} = \begin{bmatrix} E_0 V_{1r} & F_0 V_{2r} \end{bmatrix}
\end{align*}
\](4.1.5)

The reduction rules for second order system are then,

\[
\begin{align*}
\hat{M} &= I_r \\
\hat{K} &= W_{2r}^T M^{-1}K V_{1r} \\
\hat{G} &= W_{2r}^T M^{-1}G V_{2r} \\
\hat{B}_0 &= W_{2r}^T M^{-1}B_0 \\
\hat{E}_0 &= E_0 V_{1r} \\
\hat{F}_0 &= F_0 V_{2r}
\end{align*}
\](4.1.6)
Algorithm 4.1.1 [34] (Balanced truncation method based on \( \mathcal{P}_{11} \) and \( \mathcal{Q}_{11} \))

1. Compute \( \mathcal{P}_{11} \) and \( \mathcal{Q}_{11} \);

2. Compute the balanced truncation matrices (projection matrices) \( W_1, V_1 \in \mathbb{R}^{n \times r} \) such that \( W_1^T V_1 = I_r \) and \( W_1^T \mathcal{P}_{11} W_1 = V_1^T \mathcal{Q}_{11} V_1 = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r) \), where \( \sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2 \) are the \( r \) largest eigenvalues of \( \mathcal{P}_{11} \mathcal{Q}_{11} \).

3. Perform the projection to \((M, G, K, B_0, E_0, F_0)\) and get \((\hat{M}, \hat{G}, \hat{K}, \hat{B}_0, \hat{E}_0, \hat{F}_0)\) by using the formulas in (4.1.4).

4. End.

Algorithm 4.1.2 [53] (Balanced truncation method based on \( \begin{pmatrix} \mathcal{P}_{11} \\ \mathcal{P}_{22} \end{pmatrix} \) and \( \begin{pmatrix} \mathcal{Q}_{11} \\ \mathcal{Q}_{22} \end{pmatrix} \))

1. Compute \( \mathcal{P}_{11}, \mathcal{P}_{22}, \mathcal{Q}_{11}, \mathcal{Q}_{22} \);

2. Compute the balanced truncation matrices for \( \mathcal{P}_{11} \) and \( \mathcal{Q}_{11} \) to get \( W_1, V_1 \in \mathbb{R}^{n \times r} \), balanced transformation matrices for \( \mathcal{P}_{22} \) and \( \mathcal{Q}_{22} \) to get \( W_2, V_2 \in \mathbb{R}^{n \times r} \).

3. Perform the projection as in (4.1.6) to get the reduced system \((\hat{M}, \hat{G}, \hat{K}, \hat{B}_0, \hat{E}_0, \hat{F}_0)\).

4. End.

Very naturally, we may think about applying the balanced truncation method to \( P_{22} \) and \( Q_{22} \), and get the balanced projection matrices \( W_1, V_1 \) such that \( W_1^T V_1 = I_r \). Then use (4.1.3) and (4.1.4) to reduce the system.
Algorithm 4.1.3 (Balanced truncation method based on \( P_{22} \) and \( Q_{22} \))

1. Compute \( P_{22} \) and \( Q_{22} \);

2. Compute the balanced projection matrices \( W_{i_r}, V_{i_r} \in \mathbb{R}^{n \times r} \) such that \( W_{i_r}^T V_{i_r} = I_r \).

3. Use formulas (4.1.4) to perform projection and get reduced system

\[
(\hat{M}, \hat{G}, \hat{K}, \hat{B}_0, \hat{E}_0, \hat{F}_0).
\]

4. End.

Algorithms 4.1.1-4.1.3 all need to compute \( M^{-1} \). The computation of inverse of a matrix is not recommended from the numerical point of view. The reasons are two: the computational error may be larger than necessary; the computation is not efficient. For PDE problem, mass matrix is well conditioned, mass-lumping is applied (some spectral methods), so no need to compute \( M^{-1} \). In [34], Meyer and Srinivasan give another way to avoid computing \( M^{-1} \), they give the following implementations for the above projections in step 3 of Algorithm 4.1.1, which is also suitable for steps 3 in Algorithms 4.1.2 and 4.1.3. In (4.1.1), take singular value decomposition of \( M \) and get \( M = Z_1 \Lambda Z_2^T \). So \( Z_1 \) and \( Z_2 \) are orthogonal matrices, i.e. \( Z_1^T Z_1 = I_n \) and \( Z_2^T Z_2 = I_n \). Define \( \Phi_2 = Z_2 \Lambda^{-\frac{1}{2}} \) and \( \Phi_1 = \Lambda^{-\frac{1}{2}} Z_1^T \). Multiply the differential equation by \( \Phi_1 \) on the left, and coordinate transform by \( \Phi_2 \)

\[
\begin{align*}
\Phi_1 M \Phi_2 \ddot{q}(t) + \Phi_1 G \Phi_2 \dot{q}(t) + \Phi_1 K \Phi_2 q(t) &= \Phi_1 B_0 u(t) \\
y(t) &= E_0 \Phi_2 q(t) + F_0 \Phi_2 \dot{q}(t)
\end{align*}
\]

(4.1.7)
Note that $Z_1$ and $Z_2$ are orthogonal matrices,

$$\Phi_1 M \Phi_2 = (\Lambda^{-\frac{1}{2}} Z_1^T)(Z_1 \Lambda Z_2^T)(Z_2 \Lambda^{-\frac{1}{2}}) = \Lambda^{-\frac{1}{2}} \Lambda \Lambda^{-\frac{1}{2}} = I$$

In the realization of (4.1.7),

$$\{M, G, K, B_0, E_0, F_0\} = \{I_n, \Phi_1 G \Phi_2, \Phi_1 K \Phi_2, \Phi_1 B_0, E_0 \Phi_2, F_0 \Phi_2\}. \quad (4.1.8)$$

$M$ is identity matrix, so there is no need to compute the inverse of $M$.

**Algorithm 4.1.4** [34] (Implementation of steps 3 in Algorithms 4.1.1—4.1.3)

1. Perform SVD to $M$ and get $M = Z_1 \Lambda Z_2^T$. Define $\Phi_2 = Z_2 \Lambda^{-\frac{1}{2}}$ and $\Phi_1 = \Lambda^{-\frac{1}{2}} Z_1^T$;

2. In steps 3 of Algorithms 4.1.1—4.1.3, perform projections to the realization of second order system:

$$\{M, G, K, B_0, E_0, F_0\} = \{I_n, \Phi_1 G \Phi_2, \Phi_1 K \Phi_2, \Phi_1 B_0, E_0 \Phi_2, F_0 \Phi_2\}.$$

3. End.

It is easy to see that Algorithm 4.1.2 is equivalent to applying balanced truncation method to $P = \begin{pmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{pmatrix}$ and $Q = \begin{pmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} \\ \mathcal{Q}_{21} & \mathcal{Q}_{22} \end{pmatrix}$ in the corresponding first order system, in which it assumes $P$ and $Q$ are both block diagonal.

Here we apply these three balancing algorithms to two real second order problems: building models with sizes $n = 24$ and $n = 21$ respectively. The plots are frequency responses, about the relative error on $\mathcal{H}_2$-norm, we list as tables.
In the plots, dotted lines represent the frequency responses of the reduced systems, and the solid lines represent the frequency responses of the original systems. We can see for the algorithms of applying balanced truncation methods to $P_{11}$ and $Q_{11}$ or to $P_{22}$ and $Q_{22}$, i.e. Algorithms 4.1.1 and 4.1.3, the plots fit quite good, and the relative error on $H_2$-norm are relatively small. But the method of applying balanced truncation to $(P_{11}, P_{22})$ and $(Q_{11}, Q_{22})$, i.e. algorithm 4.1.2 does not have good performance in experiments, see related figures and tables below. One reason is for the corresponding first order system, the reduced systems from Algorithms 4.1.1 and 4.1.3 preserve second order structure, but the one from Algorithm 4.1.2 is no longer in second order form, and so the second order error system is relatively big in $H_2$-norm.

Figure 4.1: Frequency response on building model with size $n = 24$ by balanced truncation on $P_{11}$ and $Q_{11}$ from the paper of Meyer and Srinivasan [34]
Table 4.1: $\mathcal{H}_2$ norm of building model with size $n = 24$ by balancing on $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ from the paper of Meyer and Srinivasan [34]

<table>
<thead>
<tr>
<th>$n = 24$</th>
<th>$|\Sigma_{red}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma_{err}|_{\mathcal{H}_2}^2$</th>
<th>$\frac{|\Sigma_{err}|_{\mathcal{H}<em>2}^2}{|\Sigma|</em>{\mathcal{H}_2}^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 4$</td>
<td>2.2979e-05</td>
<td>0.1126</td>
<td></td>
</tr>
<tr>
<td>$r = 8$</td>
<td>2.0702e-05</td>
<td>0.0166</td>
<td></td>
</tr>
<tr>
<td>$r = 12$</td>
<td>2.0571e-05</td>
<td>0.0018</td>
<td></td>
</tr>
<tr>
<td>$r = 16$</td>
<td>2.0524e-05</td>
<td>5.9177e-05</td>
<td></td>
</tr>
<tr>
<td>$r = 20$</td>
<td>2.0522e-05</td>
<td>2.2513e-07</td>
<td></td>
</tr>
<tr>
<td>$r = 24$</td>
<td>2.0521e-05</td>
<td>1.3093e-13</td>
<td></td>
</tr>
</tbody>
</table>

Four algorithms are presented in a paper by VanDooren etc. [52] based on variants of this balanced truncation method. The main ideas are as following.

Recall that balanced truncation method of first order model reduction is to find transformation matrices $W, V \in \mathbb{R}^{N \times N}$ such that the transformed gramians $\bar{\mathcal{P}}$ and $\bar{\mathcal{Q}}$ are diagonal and equal:

$$\bar{\mathcal{P}} = \bar{\mathcal{Q}} = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_N)$$

where $\bar{\mathcal{P}} = W^T \mathcal{P} W$, $\bar{\mathcal{Q}} = V^T \mathcal{Q} V$, and $W^T V = I$. So $W^T \mathcal{P} \mathcal{Q} V = \Lambda = \text{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_N^2)$, where $\sigma_i$ (1 ≤ $i$ ≤ N) are the Hankel singular values of the system. Then

$$\begin{cases}
\mathcal{P} \mathcal{Q} V = V \Lambda \\
W^T \mathcal{P} Q = \Lambda W^T
\end{cases} \quad (4.1.9)$$
Figure 4.2: Frequency response on building model with size $n = 24$ by balanced truncation on $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{\text{red}}|_{H_2}^2$</th>
<th>$|\Sigma_{\text{err}}|_{H_2}^2$</th>
<th>$|\Sigma|_{H_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2.5889e-05</td>
<td>0.2037</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.0648e-05</td>
<td>0.0164</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2.0558e-05</td>
<td>0.0017</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>2.0524e-05</td>
<td>5.8799e-05</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2.0521e-05</td>
<td>2.2370e-07</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>2.0521e-05</td>
<td>3.5794e-13</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: $H_2$ norm of building model with size $n = 24$ by balancing on $\mathcal{P}_{22}$ and $\mathcal{Q}_{22}$
Figure 4.3: Frequency response on building model with size $n = 24$ by balanced truncation on $(\mathcal{P}_{11}, \mathcal{P}_{22})$ and $(\mathcal{Q}_{11}, \mathcal{Q}_{22})$ from the paper by VanDooren etc. [53]

<table>
<thead>
<tr>
<th>$n = 24$</th>
<th>$| \Sigma_{\text{red}} |_{\mathcal{H}_2}^2$</th>
<th>$\frac{| \Sigma_{\text{err}} |_{\mathcal{H}<em>2}^2}{| \Sigma |</em>{\mathcal{H}_2}^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 4$</td>
<td>3.3552e-05</td>
<td>2.0368</td>
</tr>
<tr>
<td>$r = 8$</td>
<td>3.9957e-05</td>
<td>2.0742</td>
</tr>
<tr>
<td>$r = 12$</td>
<td>4.3813e-05</td>
<td>2.1614</td>
</tr>
<tr>
<td>$r = 16$</td>
<td>4.4411e-05</td>
<td>2.1680</td>
</tr>
<tr>
<td>$r = 20$</td>
<td>4.4450e-05</td>
<td>2.1681</td>
</tr>
<tr>
<td>$r = 24$</td>
<td>4.4451e-05</td>
<td>2.1680</td>
</tr>
</tbody>
</table>

Table 4.3: $\mathcal{H}_2$ norm of building model with size $n = 24$ by balancing on $(\mathcal{P}_{11}, \mathcal{P}_{22})$ and $(\mathcal{Q}_{11}, \mathcal{Q}_{22})$ from the paper by VanDooren etc. [53]
Figure 4.4: Frequency response on building model with size $n = 21$ by balanced truncation on $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ from the paper of Meyer and Srinivasan [34]

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{\text{red}}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma_{\text{err}}|_{\mathcal{H}<em>2}^2 / |\Sigma|</em>{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.0275</td>
<td>2.3262e-04</td>
</tr>
<tr>
<td>6</td>
<td>0.0274</td>
<td>9.0939e-05</td>
</tr>
<tr>
<td>9</td>
<td>0.0274</td>
<td>1.1980e-04</td>
</tr>
<tr>
<td>12</td>
<td>0.0273</td>
<td>1.4357e-04</td>
</tr>
<tr>
<td>15</td>
<td>0.0274</td>
<td>8.5728e-06</td>
</tr>
<tr>
<td>18</td>
<td>0.0274</td>
<td>5.5039e-07</td>
</tr>
<tr>
<td>21</td>
<td>0.0274</td>
<td>1.6632e-09</td>
</tr>
</tbody>
</table>

Table 4.4: $\mathcal{H}_2$ norm of building model with size $n = 21$ by balancing on $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$ from the paper of Meyer and Srinivasan [34]
Figure 4.5: Frequency response on building model with size $n = 21$ by balanced truncation on $P_{22}$ and $Q_{22}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{red}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma_{err}|_{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.0277</td>
<td>2.4571e-04</td>
</tr>
<tr>
<td>6</td>
<td>0.0274</td>
<td>8.0149e-06</td>
</tr>
<tr>
<td>9</td>
<td>0.0274</td>
<td>3.2930e-07</td>
</tr>
<tr>
<td>12</td>
<td>0.0274</td>
<td>7.1254e-08</td>
</tr>
<tr>
<td>15</td>
<td>0.0274</td>
<td>2.0001e-08</td>
</tr>
<tr>
<td>18</td>
<td>0.0274</td>
<td>1.7831e-09</td>
</tr>
<tr>
<td>21</td>
<td>0.0274</td>
<td>1.6714e-09</td>
</tr>
</tbody>
</table>

Table 4.5: $\mathcal{H}_2$ norm of building model with size $n = 21$ by balancing on $P_{22}$ and $Q_{22}$
Figure 4.6: Frequency response on building model with size $n = 21$ by balanced truncation on $(P_{11}, P_{22})$ and $(Q_{11}, Q_{22})$ from the paper of VanDooren etc. [53]

<table>
<thead>
<tr>
<th>$n = 21$</th>
<th>$| \Sigma_{red} |_{\mathcal{H}_2}^2$</th>
<th>$| \Sigma_{err} |_{\mathcal{H}_2}^2$</th>
<th>$| \Sigma |_{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 3$</td>
<td>0.0044</td>
<td>1.1537</td>
<td></td>
</tr>
<tr>
<td>$r = 6$</td>
<td>0.0044</td>
<td>1.1559</td>
<td></td>
</tr>
<tr>
<td>$r = 9$</td>
<td>0.0045</td>
<td>1.1584</td>
<td></td>
</tr>
<tr>
<td>$r = 12$</td>
<td>0.0045</td>
<td>1.1593</td>
<td></td>
</tr>
<tr>
<td>$r = 15$</td>
<td>0.0046</td>
<td>1.1596</td>
<td></td>
</tr>
<tr>
<td>$r = 18$</td>
<td>0.0046</td>
<td>1.1611</td>
<td></td>
</tr>
<tr>
<td>$r = 21$</td>
<td>0.0049</td>
<td>1.1726</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6: $\mathcal{H}_2$ norm of building model with size $n = 21$ by balancing on $(P_{11}, P_{22})$ and $(Q_{11}, Q_{22})$ from the paper of VanDooren etc. [53]
Equating the first $k$ columns of $\mathcal{P}\mathcal{Q}V = V\Lambda$, we have

$$\mathcal{P}\mathcal{Q}V_k = V \begin{bmatrix} \Lambda_+ \\ 0 \end{bmatrix},$$

where $V_k$ consists of the first $k$ columns of $V$, and $\Lambda_+ = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_k)$. Then it is easy to see that

$$\mathcal{P}\mathcal{Q}V_k = V_k\Lambda_+.$$

Similarly, by equating the first $k$ rows of $W^T\mathcal{P}\mathcal{Q} = \Lambda W^T$, we have

$$W_k^T\mathcal{P}\mathcal{Q} = \begin{bmatrix} \Lambda_+ \\ 0 \end{bmatrix} W^T = \Lambda_+ W_k^T$$

Clearly $W_k$ and $V_k$ are projection matrices in balanced truncation method. For simplicity, we denote them by $W$ and $V$ respectively. Since the second order system always has even order in its corresponding first order system, we may suppose $k = 2r$. Recall that $N = 2n$. Partition $W, V \in \mathbb{R}^{2n \times 2r}$ as following, with each block of size $n \times r$:

$$W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}, \quad V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$$

Then we may perform projection on $(A, B, C)$ and get,

$$\hat{A} = W^TAV, \quad \hat{B} = W^TB, \quad \hat{C} = CV$$

In order to let the reduced system keep second order structure, one of the sufficient conditions is the projection matrices are block diagonal, i.e.

$$W_{12} = W_{21} = V_{12} = V_{21} = 0.$$

(4.1.10)
Based on this idea and combining with balanced truncation, [52] proposes three algorithms.

These three algorithms plus Algorithm 4.1.2 are supposing $\mathcal{P}$ and $\mathcal{Q}$ are block diagonal, or try to block diagonal $\mathcal{P}$ and $\mathcal{Q}$. From Theorem 3.2.4, we know under the assumption of $F_0 = 0$, there is no state space transformation gives block diagonal $\mathcal{Q}$, so the reduced systems produced by these four algorithms may lose some information contained in the original system.

## 4.2 Algorithms of Second Order Model Reduction

### Based on Singular Value Decomposition

In this section we propose some algorithms based on SVD method.

For a given second order system (4.1.2), from Proposition 3.2.6, we know

$$||\Sigma||_{\mathcal{H}_2}^2 = tr\{B_M^TQ_{22}B_M\} = tr\{C_M\mathcal{P}_{11}C_M^T\} \quad (4.2.1)$$

So $\mathcal{P}_{11}$ and $\mathcal{Q}_{22}$ are important quantities in weighting the $\mathcal{H}_2$ norm of second order systems, from its dual system, symmetrically $\mathcal{Q}_{11}$ and $\mathcal{P}_{22}$ are crucial in weighting the system $\mathcal{H}_2$ norms. Also from (3.2.9), (3.2.10) and their dual problems, we know that the pair $(\mathcal{P}_{11}, \mathcal{P}_{22})$ represents the information corresponding to the positions $q(t)$, and the pair $(\mathcal{Q}_{11}, \mathcal{Q}_{22})$ represents the information corresponding to the velocities $\dot{q}(t)$. we would like the reduced system to retain as much information as possible from $\mathcal{P}_{11}$, $\mathcal{P}_{22}$, $\mathcal{Q}_{11}$ or $\mathcal{Q}_{22}$. Based on this idea and combining with SVD method, we introduce
the following algorithms.

For instance, we take as much information as possible from \( P_{11} \) in reduced system. The approach is to perform SVD to \( P_{11} \), get \( P_{11} = W_1 S V_1^T \) where \( W_1 = V_1 \) are \( \mathbb{R}^{n \times n} \) orthogonal matrices and \( W_1^T V_1 = I_n \) since \( P \) is symmetric positive definite and so \( P_{11} \) is symmetric positive definite. Let the transformation matrices for the corresponding first order system be

\[
W = \begin{pmatrix} W_1 \\ W_1 \end{pmatrix}, \quad V = W
\]

and the projection matrices be

\[
W_k = \begin{pmatrix} W_{1r} \\ W_{1r} \end{pmatrix}, \quad V_k = W_k
\]

where \( W_{1r} \) consists of the first \( r \) columns of \( W_1 \). It is easy to see that \( W^T V = I_n \) and \( W_k^T V_k = I_k \). For simplicity, we denote \( W_k \) and \( V_k \) by \( W \) and \( V \), respectively. Then perform projection as in (4.1.3) and (4.1.4). In Chapter 5, we will derive an error bound for this approximation. We shall show \( \mathcal{H}_2 \) norm of the error system is bounded by a scalar times the largest neglected singular value of \( P_{11} \).

Similarly we can apply the same procedures to \( P_{22} \), \( Q_{11} \), and \( Q_{22} \), this results in three other algorithms.

Algorithm 4.2.1 (Second order model reduction — SVD on \( P_{11} \))
(or \( P_{22} \), \( Q_{11} \), \( Q_{22} \) respectively)

1. Compute \( P_{11} \) (or \( P_{22} \), \( Q_{11} \), \( Q_{22} \) respectively), we denote it as \( P \) for simplicity;
2. Take SVD on $P$ and get $P = W_1 S_1 W_1^T$. Form the orthogonal projection matrices $W_{1r} = V_{1r}$ which consists of the first $r$ columns of $W_1$;

3. Perform the projection as in (4.1.4) to get the reduced system $(\hat{M}, \hat{G}, \hat{K}, \hat{B}_0, \hat{E}_0, \hat{F}_0)$.

4. End.

From (3.2.9) and (3.2.10), we know $P_{11}$ and $P_{22}$ represent the information on second order systems corresponding to positions $q(t)$ and velocities $\dot{q}(t)$, so while taking the information from $P_{11}$, we would also like to combine with the information of $P_{22}$. The approach is to take SVD on $P_{11}$ and $P_{22}$, get $P_{11} = W_1 S_1 W_1^T$, $P_{22} = W_2 S_2 W_2^T$. Let the orthogonal projection matrices for the corresponding first order system be

$$\begin{pmatrix} W_{1r} \\ W_{2r} \end{pmatrix} V = W_k$$

where $W_{1r}, W_{2r}$ consist of the first $r$ columns of $W_1$ and $W_2$ respectively. Then perform projection as in (4.1.5) and (4.1.6). Similarly we can apply the same procedures to $Q_{11}$ and $Q_{22}$. This results in the following algorithm:

Algorithm 4.2.2 (Second order model reduction — SVD on $P_{11}$ and $P_{22}$)
(or, $Q_{11}$ and $Q_{22}$ respectively)

1. Compute $P_{11}$ and $P_{22}$ (or, $Q_{11}$ and $Q_{22}$ respectively). For simplicity, we denote $R_1 = P_{11}$ (or $Q_{11}$ respectively), $R_2 = P_{22}$ (or $Q_{22}$ respectively).
2. Perform SVD to \( R_1 \) and \( R_2 \), get \( R_1 = W_1 S_1 W_1^T \) and \( R_2 = W_2 S_2 W_2^T \). Form the projection matrices \( W_1_r \) and \( W_2_r \), which consist of the first \( r \) columns of \( W_1 \) and \( W_2 \), respectively. Let \( V_1_r = W_1_r \), \( V_2_r = W_2_r \).

3. Perform the projection as in (4.1.6) to get the reduced system \((\hat{M}, \hat{G}, \hat{K}, \hat{B}_0, \hat{E}_0, \hat{F}_0)\).

4. End.

Algorithms 4.2.1 has very good performance when applied to real problems. The reduced system produced by Algorithm 4.2.2 does not keep second order structure in its corresponding first order system, see (4.1.5) in which the \((1,2)\)-block is not identity matrix. When we perform full order transformation instead of truncation, the transformed second order systems are not equivalent to the original system (4.1.1), the performance of Algorithm 4.2.2 is not ideal in experiment. One reason is for the corresponding first order system, the reduced systems from Algorithms 4.2.1 preserve second order structure, but the one from Algorithm 4.2.2 is no longer in second order form, and so the second order error system is relatively big in \( \mathcal{H}_2 \)-norm.

The plots here are for the algorithms of applying SVD to \( P_{11}, P_{22}, Q_{11} \) and \( Q_{22} \) on two building models of size 24 and 21, respectively, and an aluminum plate model which has size 1734. The plots are frequency responses, about the relative error on \( \mathcal{H}_2 \)-norm, we list as tables. In the plots, dotted lines represent the frequency responses of the reduced systems, and the solid lines represent the frequency responses of the original systems. From the plots we can see the algorithms work good, and the relative errors are acceptable.
Figure 4.7: Frequency response on building model with size $n = 24$ by SVD on $\mathcal{P}_{11}$

<table>
<thead>
<tr>
<th>$n$ = 24</th>
<th>$|\Sigma_{\text{red}}|^2_{\mathcal{H}_2}$</th>
<th>$|\Sigma_{\text{err}}|^2_{\mathcal{H}<em>2}/|\Sigma|^2</em>{\mathcal{H}_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 4$</td>
<td>2.5509e-05</td>
<td>0.2087</td>
</tr>
<tr>
<td>$r = 8$</td>
<td>2.5487e-05</td>
<td>0.1898</td>
</tr>
<tr>
<td>$r = 12$</td>
<td>2.0539e-05</td>
<td>0.0017</td>
</tr>
<tr>
<td>$r = 16$</td>
<td>2.0522e-05</td>
<td>7.9056e-05</td>
</tr>
<tr>
<td>$r = 20$</td>
<td>2.0521e-05</td>
<td>2.7265e-07</td>
</tr>
<tr>
<td>$r = 24$</td>
<td>2.0521e-05</td>
<td>4.0615e-13</td>
</tr>
</tbody>
</table>

Table 4.7: $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $\mathcal{P}_{11}$
Figure 4.8: Frequency response on building model with size $n = 24$ by SVD on $\mathcal{P}_{22}$

<table>
<thead>
<tr>
<th>$n = 24$</th>
<th>$|\Sigma_{\text{red}}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma_{\text{err}}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma|_{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 4$</td>
<td>2.8402e-05</td>
<td>0.3231</td>
<td></td>
</tr>
<tr>
<td>$r = 8$</td>
<td>2.5415e-05</td>
<td>0.1574</td>
<td></td>
</tr>
<tr>
<td>$r = 12$</td>
<td>2.3333e-05</td>
<td>0.0895</td>
<td></td>
</tr>
<tr>
<td>$r = 16$</td>
<td>2.3685e-05</td>
<td>0.1168</td>
<td></td>
</tr>
<tr>
<td>$r = 20$</td>
<td>2.0484e-05</td>
<td>2.3613e-04</td>
<td></td>
</tr>
<tr>
<td>$r = 24$</td>
<td>2.0521e-05</td>
<td>6.9673e-13</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.8: $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $\mathcal{P}_{22}$
Figure 4.9: Frequency response on building model with size $n = 24$ by SVD on $Q_{11}$

<table>
<thead>
<tr>
<th>$n = 24$</th>
<th>$| \Sigma_{red} |^2_{\mathcal{H}_2}$</th>
<th>$| \Sigma_{err} |^2_{\mathcal{H}_2}$</th>
<th>$| \Sigma |^2_{\mathcal{H}_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 4$</td>
<td>1.3623e-05</td>
<td>0.5649</td>
<td></td>
</tr>
<tr>
<td>$r = 8$</td>
<td>1.2925e-05</td>
<td>0.5421</td>
<td></td>
</tr>
<tr>
<td>$r = 12$</td>
<td>2.4278e-05</td>
<td>0.1331</td>
<td></td>
</tr>
<tr>
<td>$r = 16$</td>
<td>2.3824e-05</td>
<td>0.0865</td>
<td></td>
</tr>
<tr>
<td>$r = 20$</td>
<td>2.0520e-05</td>
<td>1.4253e-06</td>
<td></td>
</tr>
<tr>
<td>$r = 24$</td>
<td>2.0521e-05</td>
<td>4.0450e-14</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.9: $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $Q_{11}$
Figure 4.10: Frequency response on building model with size $n = 24$ by SVD on $Q_{22}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{\text{red}}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma_{\text{err}}|_{\mathcal{H}<em>2}^2/|\Sigma|</em>{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$2.8490e-05$</td>
<td>0.3205</td>
</tr>
<tr>
<td>8</td>
<td>$2.5399e-05$</td>
<td>0.1574</td>
</tr>
<tr>
<td>12</td>
<td>$2.3490e-05$</td>
<td>0.0964</td>
</tr>
<tr>
<td>16</td>
<td>$2.3625e-05$</td>
<td>0.1130</td>
</tr>
<tr>
<td>20</td>
<td>$2.0630e-05$</td>
<td>$3.3623e-04$</td>
</tr>
<tr>
<td>24</td>
<td>$2.0521e-05$</td>
<td>$3.8650e-13$</td>
</tr>
</tbody>
</table>

Table 4.10: $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $Q_{22}$
Figure 4.11: Frequency response on building model with size $n = 21$ by SVD on $\mathcal{P}_{11}$

<table>
<thead>
<tr>
<th>$n = 21$</th>
<th>$| \Sigma_{\text{red}} |_{\mathcal{H}_2}^2$</th>
<th>$| \Sigma_{\text{err}} |_{\mathcal{H}<em>2}^2 / | \Sigma |</em>{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 3$</td>
<td>0.0260</td>
<td>0.0021</td>
</tr>
<tr>
<td>$r = 6$</td>
<td>0.0270</td>
<td>0.0097</td>
</tr>
<tr>
<td>$r = 9$</td>
<td>0.0270</td>
<td>0.0052</td>
</tr>
<tr>
<td>$r = 12$</td>
<td>0.0271</td>
<td>0.0052</td>
</tr>
<tr>
<td>$r = 15$</td>
<td>0.0276</td>
<td>7.0015e-05</td>
</tr>
<tr>
<td>$r = 18$</td>
<td>0.0274</td>
<td>1.0061e-07</td>
</tr>
<tr>
<td>$r = 21$</td>
<td>0.0274</td>
<td>1.9009e-08</td>
</tr>
</tbody>
</table>

Table 4.11: $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $\mathcal{P}_{11}$
Figure 4.12: Frequency response on building model with size $n = 21$ by SVD on $P_{22}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{\text{red}}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma_{\text{err}}|_{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.0281</td>
<td>0.0017</td>
</tr>
<tr>
<td>6</td>
<td>0.0266</td>
<td>0.0011</td>
</tr>
<tr>
<td>9</td>
<td>0.0271</td>
<td>0.0074</td>
</tr>
<tr>
<td>12</td>
<td>0.0272</td>
<td>0.0048</td>
</tr>
<tr>
<td>15</td>
<td>0.0276</td>
<td>8.8357e-05</td>
</tr>
<tr>
<td>18</td>
<td>0.0274</td>
<td>8.4319e-07</td>
</tr>
<tr>
<td>21</td>
<td>0.0274</td>
<td>5.9075e-08</td>
</tr>
</tbody>
</table>

Table 4.12: $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $P_{22}$
Figure 4.13: Frequency response on building model with size $n = 21$ by SVD on $Q_{11}$

<table>
<thead>
<tr>
<th>$n = 21$</th>
<th>$|\Sigma_{red}|_2^2$</th>
<th>$|\Sigma_{err}|_2^2$</th>
<th>$|\Sigma|_2^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 3$</td>
<td>$3.7114e-16$</td>
<td>$1.0000$</td>
<td></td>
</tr>
<tr>
<td>$r = 6$</td>
<td>$3.2814e-16$</td>
<td>$1.0000$</td>
<td></td>
</tr>
<tr>
<td>$r = 9$</td>
<td>$0.0106$</td>
<td>$0.2866$</td>
<td></td>
</tr>
<tr>
<td>$r = 12$</td>
<td>$0.0242$</td>
<td>$0.0603$</td>
<td></td>
</tr>
<tr>
<td>$r = 15$</td>
<td>$0.0278$</td>
<td>$0.0021$</td>
<td></td>
</tr>
<tr>
<td>$r = 18$</td>
<td>$0.0279$</td>
<td>$7.4569e-04$</td>
<td></td>
</tr>
<tr>
<td>$r = 21$</td>
<td>$0.0274$</td>
<td>$6.5396e-07$</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.13: $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $Q_{11}$
Figure 4.14: Frequency response on building model with size $n = 21$ by SVD on $Q_{22}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{red}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma_{err}|_{\mathcal{H}<em>2}^2/|\Sigma|</em>{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$1.0321e-15$</td>
<td>1.0000</td>
</tr>
<tr>
<td>6</td>
<td>$5.9236e-13$</td>
<td>1.0000</td>
</tr>
<tr>
<td>9</td>
<td>0.0287</td>
<td>0.0041</td>
</tr>
<tr>
<td>12</td>
<td>0.0276</td>
<td>$6.2968e-04$</td>
</tr>
<tr>
<td>15</td>
<td>0.0274</td>
<td>$1.2348e-06$</td>
</tr>
<tr>
<td>18</td>
<td>0.0274</td>
<td>$1.3800e-06$</td>
</tr>
<tr>
<td>21</td>
<td>0.0274</td>
<td>$9.5546e-07$</td>
</tr>
</tbody>
</table>

Table 4.14: $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $Q_{22}$
Figure 4.15: Frequency response on Aluminum model with size $n = 1734$ by SVD on $P_{11}$, top figure is of order $r = 240$, bottom figure is of order $r = 140$
Figure 4.16: Frequency response on Aluminum model with size $n = 1734$ by SVD on $P_{22}$, with different ranges on frequencies.
4.3 Computing the Blocks of $\mathcal{P}$ and $\mathcal{Q}$.

Most algorithms in Sections 1 and 2 need to compute the diagonal blocks of controllability and observability gramians: $\mathcal{P}_{11}$, $\mathcal{P}_{22}$, $\mathcal{Q}_{11}$ and $\mathcal{Q}_{22}$. There exist many methods to compute $\mathcal{P}$ and $\mathcal{Q}$, but not the blocks. One may take the $(1,1)$-block, the first $n \times n$ principle submatrix of $\mathcal{P}$ to form $\mathcal{P}_{11}$ after getting $\mathcal{P}$. But obviously this is not efficient, and so not practical for large scale matrices.

Sorensen in [46] proposes an efficient algorithm to compute controllability gramian $\mathcal{P}$, which can be used to compute the blocks of $\mathcal{P}$ efficiently.

**Theorem 4.3.1** [46] *Given a first order system*

$$
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix}
$$

*and some step size $h \in \mathbb{R}$, assume zero initial condition $x(0) = 0$ and the unit impulse input $u(t) = \delta(t)$. Let

$$x_1 = (I - \frac{h}{2}A)^{-1}B
$$

and

$$x_{j+1} = (I - \frac{h}{2}A)^{-1}(I + \frac{h}{2}A)x_j \text{ for } j = 1, 2, 3 \cdots
$$

Then the controllability gramian $\mathcal{P}$ is,

$$
\mathcal{P} = \int_0^\infty x(\tau)x(\tau)^T d\tau = h \sum_{j=0}^\infty x_{j+1}x_{j+1}^T.
$$

(4.3.1)

Single input single out (SISO) systems are simple but important in real problems.

For SISO case, in order to compute $\mathcal{P}_{11}$, the first $n \times n$ submatrix of $\mathcal{P}$, we can take the first half of $x_j$ ($j \geq 1$) in each step, then take the product and summation as in (4.3.1). The following is an algorithm to compute $\mathcal{P}_{11}$ for SISO systems. The
Figure 4.17: Frequency response on Aluminum model with size $n = 1734$ by SVD on $\Omega_{11}$, with different ranges on frequencies.
Figure 4.18: Frequency response on Aluminum model with size $n = 1734$ by SVD on $Q_{22}$, with different ranges on frequencies, both are of order $r = 140$. 
computation can be easily generalized to multiple input multiple output (MIMO) systems.

**Algorithm 4.3.1** [46] (Computation of $P_{11}$ by using Theorem 4.3.1 for SISO systems)

*input: $A \in \mathbb{R}^{N \times N}$, $B \in \mathbb{R}^N$, step size $h$, $N = 2n$; output: $P_{11}$.*

1. $x = (I - \frac{h}{2}A)^{-1}B$;

2. $y = x[1:n,1]$;

3. $P_{11} = yy^T$;

4. $j = 1$;

5. while ($\|y\| > tol \|P_{11}\|$ and $j < maxiter$),

6. $x = (I - \frac{h}{2}A)^{-1}((I + \frac{h}{2}A)x)$;

7. $y = x[1:n,1]$ (i.e. take the first $n$ components);

8. $P_{11} = P_{11} + yy^T$;

9. $j = j+1$;

10. end while;

11. end;
For MIMO systems, instead of letting \( y = x[1:n,1] \) in Algorithm 4.3.1, which is the first half of \( x \) in step 2 and step 7, we take the first \( n \) rows of \( x \), i.e. \( y = [1:n,:) \). To compute \( P_{22} \), we then take the second half of \( x \) in step 2 and step 7.

In Algorithm 4.3.1, it needs to do matrix factorization to get \((I - \frac{h}{2}A)^{-1}\) which takes \( O(N^3) \) time, but it only needs to do this once in the whole algorithm, all other steps take time \( O(N^2) \). Step 6 uses BLAS (see [2] [8]) level 2 computations, i.e. matrix-vector products, which has \( \text{iter} \times N^2 \) operations. Steps 3 and 8 use matrix-matrix product, BLAS level 3 computations, but each is the product of an \( n \times 1 \) matrix and a \( 1 \times n \) matrix, so step 3 and step 8 take time \( n^2 \) and \( \text{iter} \times n^2 \) respectively. Step 1 and step 6 are involved in solving linear equations, each of which costs \( O(N^2) \) assuming we know \((I - \frac{h}{2}A)^{-1}\). Note that \( N = 2n \), \( O(N^2) = O(n^2) \). So the whole algorithm takes \( O(n^2) \) plus a one-time \( O(n^3) \) operation, which makes the algorithm favorable for large scale matrices.

In experiments, Algorithm 4.3.1 works quite good in accuracy, but for some of the real problems, convergence is not fast. For instance, for a second order Aluminum Plate model, each iteration in computing \( P \) takes about 16 minutes, the convergence rate becomes an important factor in getting an accurate \( P \), or the diagonal blocks. In order to speed convergence, we may use multiple shifts in computing \( P \) or \( Q \).

The basic idea of Theorem 4.3.1 is to use quadrature rule to get \( P \), and it is the same as the result that we get by using Smith’s method. Consider Lyapunov equation

\[
AP + PA^T + BB^T = 0.
\]
We can get the solution $\mathcal{P}$ by using the well known Smith’s method:

$$\mathcal{P} = 2\mu \sum_{j=0}^{\infty} A_{\mu}^j B_{\mu} B_{\mu}^T (A_{\mu}^T)^j$$

where $\mu \in \mathbb{R}$ is a shift we use, $A_{\mu} = (A - \mu I)^{-1}(A + \mu I)$ and $B_{\mu} = (A - \mu I)^{-1}B$. In the $k$-th iteration,

$$\mathcal{P}_k^S = 2\mu \sum_{j=0}^{k} A_{\mu}^j B_{\mu} B_{\mu}^T (A_{\mu}^T)^j.$$

According to Penzl [37], computational experience indicates that Smith’s method (also called ADI method with single shift) converges very slowly, a moderate increase of the number of shifts would speed up the convergence. However it is also indicated that a further increase of the number of shifts would hardly improve the acceleration. So it is recommended to use Smith-$l$ iterations, i.e. to use $l$ shifts cyclicly. In the $k$-th index of iteration, we get

$$\mathcal{P}_k^{Sl} = \sum_{j=0}^{k-1} A_d^j T (A_d^j)^T$$

where

$$A_d = \prod_{i=1}^{l} (A - \mu_i^* I)(A + \mu_i I)^{-1}$$

and

$$T = -2\rho_i \sum_{i=1}^{l} A_{\mu_i} B B^T A_{\mu_i}^*$$

where $\rho_i = \text{Real}(\mu_i)$ and $A_{\mu_i} = \prod_{j=1}^{i} (A - \mu_j^* I)(A + \mu_j I)^{-1}$

By quadrature, we can get the same result for $\mathcal{P}$. Assume $u(\tau) = \delta(\tau)$ and $x(0) = 0$. By definition,

$$\mathcal{P} = \int_{0}^{\infty} x(\tau)x(\tau)^T d\tau.$$
We then partition $t$ into small intervals, the simplest way is to equally divide $t$ with each interval of length $h$, then we get $x(0), x(h), x(2h), \cdots, x(jh), x((j + 1)h), \cdots$.

By the definition of integral,

$$
P = \int_0^\infty x(\tau) x(\tau)^T d\tau = h \sum_{j=0}^\infty x_{j+1} x_{j+1}^T
$$

By letting $\mu = \frac{2}{h}$, we then obtain the following,

$$
P = h \sum_{j=0}^\infty x_{j+1} x_{j+1}^T = 2\mu \sum_{j=0}^\infty A_\mu^j B_\mu B_\mu^T (A_\mu^T)^j
$$

which is the same as the result as get from Smith’s method. From the approach of quadrature rule, we can also try to speed up the convergence of $P$ by using multiple step size $h_1, h_2, \cdots$. Here $h_i$ represents the step size, it must be very small since by the definition of integral, the step size should be approaching zero. Also we know $\mu = \frac{2}{h}$, so the shifts $\mu_i$ that we choose in the methods of Smith or Smith-$l$ iteration, should be relatively big.

A method for computing $Q$, is given in Theorem 4.3.1 and its dual problem, which we state as a corollary:

**Corollary 4.3.2** With the same conditions as in Theorem 4.3.1, let

$$x_1 = (I - \frac{h}{2} A^T)^{-1} C^T$$

and

$$x_{j+1} = (I - \frac{h}{2} A^T)^{-1}(I + \frac{h}{2} A^T)x_j \text{ for } j = 1, 2, 3 \cdots$$
Then the observability gramian $Q$ is,

\[ Q = \int_{0}^{\infty} x(\tau)x(\tau)^T d\tau = h \sum_{j=0}^{\infty} x_{j+1}x_{j+1}^T. \]

**Proof:** From Lyapunov equations

\begin{align}
AP + PA^T + BB^T &= 0 \quad (4.3.2) \\
A^TQ + QA + C^TC &= 0 \quad (4.3.3)
\end{align}

It is easy to see that $Q$ is essentially the controllability gramian for the dual problem,

\[
\begin{cases}
\dot{x} &= A^Tx + C^Tu \\
y &= B^Tx + D^Tu 
\end{cases} \quad (4.3.4)
\]

Then from Theorem 4.3.1, the results of this Corollary hold.

\[ \blacksquare \]

### 4.4 Some Discussions of the Proposed Algorithms

In the following proposition, we list the properties of reduced systems produced by Algorithms 4.1.1—4.1.3 and 4.2.1—4.2.3, i.e. all the algorithms we discussed in this thesis so far, including two existing algorithms.

**Proposition 4.4.1** For a given second order system (4.1.1), the following statements hold for Algorithms 4.1.1—4.1.3 and 4.2.1—4.2.3:

1. The full order transformed second order systems produced by Algorithms 4.1.1, 4.1.3, 4.2.1 and 4.2.3, i.e. by the reduction formulas of (4.1.4), are equivalent to
the original systems, those produced by Algorithms 4.1.2 and 4.2.2, i.e. by the reduction formulas (4.1.6), are not equivalent to the original system in general.

2. All algorithms in the reduced systems preserve second order structure.

3. For the corresponding first order system, the reduced systems produced by Algorithms 4.1.1, 4.1.3, 4.2.1 and 4.2.3, i.e. by the formulas of (4.1.4), preserve second order form, the reduced system produced by Algorithms 4.1.2 and 4.2.2, i.e. by formulas in (4.1.6) are not in second order form.

Proof: Recall that for the corresponding first order system, Algorithms 4.1.1, 4.1.3, 4.2.1 and 4.2.3 have transformation matrices

\[ W = \begin{pmatrix} W_1 \\ W_1 \end{pmatrix}, \quad V = \begin{pmatrix} V_1 \\ V_1 \end{pmatrix} \]

where \( W_1, V_1 \in \mathbb{R}^{n \times n} \) and \( W_1^T V_1 = I_n \), therefore \( W, V \in \mathbb{R}^{N \times N} \) and \( W^T V = I_N \).

Algorithms 4.1.2 and 4.2.2 have transformation matrices

\[ W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}, \quad V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \]

where \( W_1, V_1, W_2, V_2 \in \mathbb{R}^{n \times n} \), \( W_1^T V_1 = I_n \) and \( W_2^T V_2 = I_n \), therefore \( W, V \in \mathbb{R}^{N \times N} \) and \( W^T V = I_N \). In general, \( W_1 \neq W_2, V_1 \neq V_2 \).

1. From second order system

\[
\begin{align*}
M \ddot{q}(t) + G\dot{q}(t) + Kq(t) &= B_0u(t) \\
y(t) &= E_0q(t) + F_0\dot{q}(t)
\end{align*}
\]  
(4.4.1)
Multiply the differential equation in (4.4.1) by \( M^{-1} \) on the left,

\[
\begin{align*}
\ddot{q}(t) + M^{-1}G\dot{q}(t) + M^{-1}Kq(t) &= M^{-1}B_0u(t) \\
y(t) &= E_0q(t) + F_0\dot{q}(t)
\end{align*}
\] (4.4.2)

Multiply the differential equation in (4.4.2) by \( W_1^T \) on the left, by noticing that \( V_1^{-1}V_1 = I_n \), we get

\[
\begin{align*}
W_1^TV_1V_1^{-1}\ddot{q}(t) + W_1^TM^{-1}GV_1V_1^{-1}\dot{q}(t) + W_1^TM^{-1}KV_1V_1^{-1}q(t) &= W_1^TM^{-1}B_0u(t) \\
y(t) &= E_0V_1V_1^{-1}q(t) + F_0V_1V_1^{-1}\dot{q}(t)
\end{align*}
\] (4.4.3)

Coordinate transform by \( V_1 \), and denote \( z(t) = V_1^{-1}q(t) \). Then in the new coordinate system, (4.4.3) becomes

\[
\begin{align*}
\ddot{z}(t) + W_1^TM^{-1}GV_1\dot{z}(t) + W_1^TM^{-1}KV_1z(t) &= W_1^TM^{-1}B_0u(t) \\
y(t) &= E_0V_1z(t) + F_0V_1\dot{z}(t)
\end{align*}
\] (4.4.4)

The realization for second order system (4.4.4) is

\[\{M, G, K, B_0, E_0, F_0\} = \{I_n, W_1^TM^{-1}GV_1, W_1^TM^{-1}KV_1, E_0V_1, F_0V_1\}\]

This is the same as the full order transformed system produced by Algorithms 4.1.1, 4.1.3, 4.2.1 and 4.2.3, i.e. the same reduction rules as in equations (4.1.4).

Algorithms 4.1.2 and 4.2.2 use reduction rules of (4.1.6) in which

\[
\hat{K} = W_2^TM^{-1}KV_1, \\
\hat{G} = W_2^TM^{-1}GV_2,
\]

and in general \( V_{1r} \neq V_{2r} \). So these two algorithms generally cannot make the full order transformed system equivalent to the original system.
2. All these algorithms use reduction formulas of (4.1.4) or (4.1.6), in which we can easily see the reduced systems are in second order structure.

3. Algorithms 4.1.1, 4.1.3, 4.2.1 and 4.2.3 use the reduction rules given by (4.1.3) and (4.1.4), from (4.1.3), it is easy to see the corresponding first order reduced system is in second order structure. Algorithms 4.1.2 and 4.2.2 are developed by equations (4.1.5) and (4.1.6). In (4.1.6), the (1, 2)-block is not identity matrix, so the reduced system is not in second order form for the corresponding first order system. (4.1.5).

In this chapter, we developed several algorithms based on balancing transformation and SVD. We also discussed two existing algorithms. All these are gramian based. Algorithms 4.1.1, 4.1.3 and 4.2.1 have good performances when applying to real problems, Algorithms 4.1.2 and 4.2.2 do not perform well in experiments. One reason is for Algorithms 4.1.1, 4.1.3 and 4.2.1, the reduced systems keep second order structure in the corresponding first order system, but Algorithms 4.1.2 and 4.2.2 do not have this property. From this we can conclude that when we work on second order model reduction, it is important to let reduced system in the corresponding first order system keep second order structure. So when we develop algorithms and theory, it is essential to study the corresponding first order system first, and try to let the reduced system preserve second order structure.
Chapter 5

Error Bounds

In this chapter, we will give a global error bound for a second order system which has $\mathcal{P}_{11}$ (or $\mathcal{Q}_{11}$) diagonal in the transformed systems. There are several ways to achieve this, for example, by the Algorithms 4.2.1 and 4.2.2, performing SVD to $\mathcal{P}_{11}$ (or $\mathcal{Q}_{11}$), or by Algorithms 4.1.1 and 4.1.2, applying balanced truncation method to $\mathcal{P}_{11}$ and $\mathcal{Q}_{11}$. See Chapters 3 and 4. This global error bound, which is the $\mathcal{H}_2$ norm of the error system, is bounded by a scalar times the summation of the neglected singular (or Hankel singular) values of $\mathcal{P}_{11}$ (or $\mathcal{Q}_{11}$).

We also provide a global error bound for a second order error system which has $\mathcal{P}_{22}$ (or $\mathcal{Q}_{22}$) diagonal in transformed systems. However this error bound is not really satisfactory.

For a given second order system (4.1.1), assume $F_0 = 0$, a special case of (4.1.1). From Theorem 3.2.4, we know there is no state space transformation which gives block
diagonal \( Q \). So the technique in deriving the error bound for balanced truncation method in first order systems cannot be applied to second order systems.

5.1 Error Systems

In this section, we introduce the second order error systems, and close by giving expressions of \( \mathcal{H}_2 \) norm of the error systems. First we shall recall some previous notation.

For a given second order system,

\[
\Sigma : \begin{cases}
M \ddot{q}(t) + G \dot{q}(t) + K q(t) = B_0 u(t) \\
y(t) = E_0 q(t)
\end{cases}
\]  
(5.1.1)

its corresponding first order system is,

\[
\Sigma_1 : \begin{bmatrix}
0 & I \\
K_M & G_M \\
C_M & 0
\end{bmatrix}
\begin{bmatrix}
0 \\
B_M \\
D
\end{bmatrix}
\]

where

\[K_M = M^{-1}K, \ G_M = M^{-1}G, \ B_M = M^{-1}B_0, \ C_M = E_0, \ D = 0\]

Suppose second order system \( \Sigma \) in (5.1.1) has dimension \( n \), then its corresponding first order system \( \Sigma_1 \) has dimension \( N = 2n \). Partition first order gramians \( \mathcal{P} \) and \( \mathcal{Q} \) into four equal blocks accordingly,

\[
\mathcal{P} = \begin{pmatrix}
\mathcal{P}_{11} & \mathcal{P}_{12} \\
\mathcal{P}_{12}^T & \mathcal{P}_{22}
\end{pmatrix}, \quad \mathcal{Q} = \begin{pmatrix}
\mathcal{Q}_{11} & \mathcal{Q}_{12} \\
\mathcal{Q}_{12}^T & \mathcal{Q}_{22}
\end{pmatrix}
\]
From Proposition 3.2.6, the $\mathcal{H}_2$ norm of the second order system can be expressed as:

$$||\Sigma||_{\mathcal{H}_2}^2 = \text{trace}\left\{ B_M^T \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} B_M \right\} = \text{trace}\{B_M^T Q_{22} B_M\} \quad (5.1.2)$$

or

$$||\Sigma||_{\mathcal{H}_2}^2 = \text{trace}\left\{ C_M^T \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix} C_M \right\} = \text{trace}\{C_M P_{11} C_M^T\} \quad (5.1.3)$$

In order to let the corresponding first order reduced system preserve second order structure, a sufficient condition is the projection matrices for the corresponding first order system are in the form of:

$$W_k = \begin{pmatrix} W_r \\ \end{pmatrix} \quad V_k = \begin{pmatrix} V_r \\ \end{pmatrix}$$

where $W_r, V_r \in \mathbb{R}^{n \times r}$, $W_r^T V_r = I_r$, $k = 2r$ and so $W_k^T V_k = I_k = I_{2r}$. Then the reduced system is:

$$A_1 = W_k^T A V_k = \begin{bmatrix} 0 & I_r \\ W_r^T K_M V_r & W_r^T G_M V_r \end{bmatrix} = \begin{bmatrix} 0 & I_r \\ K_{M1} & G_{M1} \end{bmatrix}$$

$$B_1 = W_k^T B = \begin{bmatrix} 0 \\ W_r^T B_M \end{bmatrix} = \begin{bmatrix} 0 \\ B_{M1} \end{bmatrix}$$

$$C_1 = C V_k = \begin{bmatrix} C_M V_r & 0 \end{bmatrix} = \begin{bmatrix} C_{M1} & 0 \end{bmatrix}$$

And the actual reduction rules are:

$$M_1 = I_r$$
\[ K_{M1} = W_r^T K_M V_r \]
\[ G_{M1} = W_r^T G_M V_r \]
\[ B_{M1} = W_r^T B_M \]
\[ C_{M1} = C_M V_r \]

In the reduced system, the controllability and observability gramians are,

\[
P_1 = W_k^T \mathcal{P} W_k = \begin{bmatrix} W_r^T \mathcal{P}_{11} W_r & W_r^T \mathcal{P}_{12} W_r \\ W_r^T \mathcal{P}_{21} W_r & W_r^T \mathcal{P}_{22} W_r \end{bmatrix} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21}^T & \mathcal{P}_{22} \end{bmatrix} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21}^T & \mathcal{P}_{22} \end{bmatrix} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21}^T & \mathcal{P}_{22} \end{bmatrix}
\]

\[
Q_1 = V_k^T \mathcal{Q} V_k = \begin{bmatrix} V_r^T \mathcal{Q}_{11} V_r & V_r^T \mathcal{Q}_{12} V_r \\ V_r^T \mathcal{Q}_{21} V_r & V_r^T \mathcal{Q}_{22} V_r \end{bmatrix} = \begin{bmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} \\ \mathcal{Q}_{21}^T & \mathcal{Q}_{22} \end{bmatrix} = \begin{bmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} \\ \mathcal{Q}_{21}^T & \mathcal{Q}_{22} \end{bmatrix}
\]

The error system is:

\[
\begin{bmatrix} A & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} B \\ B_1 \end{bmatrix} = \begin{bmatrix} 0 & I \\ K_M & G_M \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 \\ B_M \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

Note that its transfer function is

\[
H_{err}(s) = C(sI - A)^{-1}B - C_1(sI - A_1)^{-1}B_1
\]

The gramians for the error system are,

\[
\mathcal{P}_{err} = \begin{pmatrix} Q & \gamma \\ \gamma^T & Q_1 \end{pmatrix}, \quad \mathcal{Q}_{err} = \begin{pmatrix} Q & -Z \\ -Z^T & Q_1 \end{pmatrix}
\]
and the Lyapunov equations are:

\[
\begin{pmatrix}
A & 0 \\
0 & A_1
\end{pmatrix}
\begin{pmatrix}
P & Y \\
Y^T & P_1
\end{pmatrix}
+ \begin{pmatrix}
P & Y \\
Y^T & P_1
\end{pmatrix}
\begin{pmatrix}
A^T & 0 \\
0 & A_1^T
\end{pmatrix}
+ \begin{bmatrix}
B \\
B_1
\end{bmatrix}
\begin{bmatrix}
B^T & B_1^T
\end{bmatrix} = 0
(5.1.4)
\]

\[
\begin{pmatrix}
A^T & 0 \\
0 & A_1^T
\end{pmatrix}
\begin{pmatrix}
Q & -Z \\
-Z^T & Q_1
\end{pmatrix}
+ \begin{pmatrix}
Q & -Z \\
-Z^T & Q_1
\end{pmatrix}
\begin{pmatrix}
A & 0 \\
0 & A_1
\end{pmatrix}
+ \begin{bmatrix}
C^T \\
-C_1^T
\end{bmatrix}
\begin{bmatrix}
C \\
-C_1
\end{bmatrix} = 0
(5.1.5)
\]

Equating the blocks of (5.1.4) and (5.1.5), we have

\[
AP + PA^T + BB^T = 0
(5.1.6)
\]

\[
A_1 P_1 + P_1 A_1^T + B_1 B_1^T = 0
(5.1.7)
\]

\[
AY + YA_1^T + BB_1^T = 0
(5.1.8)
\]

and

\[
A^T Q + QA + C^T C = 0
(5.1.9)
\]

\[
A_1^T Q_1 + Q_1 A_1 + C_1^T C_1 = 0
(5.1.10)
\]

\[
A^T Z + ZA_1 + C^T C_1 = 0
(5.1.11)
\]

Partition \( Y \) and \( Z \) into four equal blocks

\[
Y = \begin{pmatrix}
Y_{11} & Y_{12} \\
Y_{21} & Y_{22}
\end{pmatrix}
\quad Z = \begin{pmatrix}
Z_{11} & Z_{12} \\
Z_{21} & Z_{22}
\end{pmatrix}
\]

Notice that

\[
\begin{bmatrix}
B^T & B_1^T
\end{bmatrix}
\begin{bmatrix}
Q & -Z \\
-Z^T & Q_1
\end{bmatrix}
\begin{bmatrix}
B \\
B_1
\end{bmatrix} = 0
\]
\[
\begin{bmatrix}
0 & B_M^T & 0 & B_{M1}^T
\end{bmatrix}
\begin{bmatrix}
Q & -Z \\
-Z^T & Q_1
\end{bmatrix}
\begin{bmatrix}
0 \\
B_M \\
0 \\
B_{M1}
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & B_M^T \\
B_M
\end{bmatrix}
Q
\begin{bmatrix}
0 \\
B_M^T \\
0 \\
B_{M1}
\end{bmatrix}
Z
\begin{bmatrix}
0 \\
B_{M1}
\end{bmatrix}
+ 
\begin{bmatrix}
0 & B_{M1}^T \\
B_M
\end{bmatrix}
Q_1
\begin{bmatrix}
0 \\
B_{M1}
\end{bmatrix}
\]

\[
= B_M^T Q_{22} B_M - 2 B_M^T Z_{22} B_{M1} + B_{M1}^T Q_{22} B_{M1}
\]

and

\[
\begin{bmatrix}
C & -C_1 \\
\gamma^T & \mathcal{P}_1
\end{bmatrix}
\begin{bmatrix}
P & \gamma \\
\gamma^T & \mathcal{P}_1
\end{bmatrix}
\begin{bmatrix}
C_M^T \\
C_{M1}^T
\end{bmatrix}
\]

\[
= \begin{bmatrix}
C_M & 0 & -C_{M1} & 0
\end{bmatrix}
\begin{bmatrix}
P & \gamma \\
\gamma^T & \mathcal{P}_1
\end{bmatrix}
\begin{bmatrix}
C_M^T \\
C_{M1}^T \\
0
\end{bmatrix}
\]

\[
= \begin{bmatrix}
C_M & 0
\end{bmatrix}
\mathcal{P}
\begin{bmatrix}
C_M^T \\
0
\end{bmatrix}
- 2 \begin{bmatrix}
C_M & 0
\end{bmatrix}
\gamma
\begin{bmatrix}
C_{M1}^T \\
0
\end{bmatrix}
\]

\[
+ \begin{bmatrix}
C_{M1} & 0
\end{bmatrix}
\mathcal{P}_1
\begin{bmatrix}
C_{M1}^T \\
0
\end{bmatrix}
\]

\[
= C_M \mathcal{P}_{11} C_M^T - 2 C_M \gamma_{11} C_{M1}^T + C_{M1} \mathcal{P}_{11} C_{M1}^T
\]

which result in the following lemma.

**Lemma 5.1.1** $\mathcal{H}_2$ norm of the error system of second order system (5.1.1) can be
expressed as

$$\|\Sigma_{err}\|_{H_2} = \text{trace}\{B_M^T Q_{22} B_M - 2 B_M^T Z_{22} B_{M1} + B_{M1}^T Q_{221} B_{M1}\}$$

or

$$\|\Sigma_{err}\|_{H_2} = \text{trace}\{C_M P_{11} C_M^T - 2 C_M Y_{11} C_{M1}^T + C_{M1} P_{11} C_{M1}^T\}.$$  

5.2 Error Bounds Associated with $P_{22}$ and $Q_{22}$

In this section, we give an error bound for a second order linear system which has $Q_{22}$ diagonal. There are many ways to make $Q_{22}$ diagonal, for example, SVD on $Q_{22}$, or applying balanced truncation method to $P_{22}$ and $Q_{22}$. Symmetrically by dual system, we can give an error bound associated with $P_{22}$.

**Theorem 5.2.1** Suppose $Q_{22} = S = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}$, where $S_1 = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r)$, $S_2 = \text{diag}(\sigma_{r+1}, \sigma_{r+2}, \cdots, \sigma_n)$. Then for SISO system,

$$\|\Sigma_{err}\|_{H_2} \leq c_1 \sigma_{r+1} + c_2 \|J\|_F$$

where $J = Z_{12} + Z_{21} - Q_{12}^T \cdot (Q_{12}^T)^r$, $Q_{12}$ and $(Q_{12}^T)^r$ consist of the first $r$ columns of $Q_{12}$ and $Q_{12}^T$ respectively.

**Proof:** From Lemma 5.1.1,

$$\|\Sigma_{err}\|_{H_2} = \text{trace}\{B_M^T Q_{22} B_M - 2 B_M^T Z_{22} B_{M1} + B_{M1}^T Q_{221} B_{M1}\}$$

$$= \text{trace}\{2B_{M1}^T S_1 B_{M1} - 2 B_M^T Z_{22} B_{M1} + B_{M1}^T S_2 B_{M2}\} \quad (5.2.1)$$
where $Z$ satisfies

$$A^T Z + Z A_1 + C^T C_1 = 0$$

So

$$
\begin{pmatrix}
0 & K_M^T \\
I & G_M^T
\end{pmatrix}
\begin{pmatrix}
Z_{11} & Z_{12} \\
Z_{21} & Z_{22}
\end{pmatrix}
+ \begin{pmatrix}
Z_{11} & Z_{12} \\
Z_{21} & Z_{22}
\end{pmatrix}
\begin{pmatrix}
0 & I \\
K_M & G_M
\end{pmatrix}
+ \begin{pmatrix}
C_M^T \\
0
\end{pmatrix}
\begin{pmatrix}
C_M \\
0
\end{pmatrix} = 0
$$

(5.2.2)

Equating each block of (5.2.2), we get

$$K_M^T Z_{21} + Z_{12} K_M + C_M^T C_M = 0$$

(5.2.3)

$$K_M^T Z_{22} + Z_{11} + Z_{12} G_M = 0$$

(5.2.4)

$$Z_{11} + G_M^T Z_{21} + Z_{22} K_M = 0$$

(5.2.5)

$$Z_{12} + G_M^T Z_{22} + Z_{21} + Z_{22} G_M = 0$$

(5.2.6)

Partition the matrices in each side of $A^T Q + QA + C^T C = 0$,

$$
\begin{pmatrix}
0 & K_M^T \\
I & G_M^T
\end{pmatrix}
\begin{pmatrix}
Q_{11} & Q_{12} \\
Q_{12}^T & S
\end{pmatrix}
+ \begin{pmatrix}
Q_{11} & Q_{12} \\
Q_{12}^T & S
\end{pmatrix}
\begin{pmatrix}
0 & I \\
K_M & G_M
\end{pmatrix}
+ \begin{pmatrix}
C_M^T \\
0
\end{pmatrix}
\begin{pmatrix}
C_M \\
0
\end{pmatrix} = 0.
$$

(5.2.7)

Equating each block of (5.2.7), we get

$$K_M^T Q_{12} + Q_{12} K_M + C_M^T C_M = 0$$

(5.2.8)

$$K_M^T S + Q_{11} + Q_{12} G_M = 0$$

(5.2.9)

$$Q_{12} + G_M^T S + Q_{12}^T + S G_M = 0$$

(5.2.10)
The first $r$ columns of $G_M$ is $G_M^r = \begin{bmatrix} (G_M^r)_1 \\ (G_M^r)_2 \end{bmatrix}$ where $(G_M^r)_1$ is the first $r$ rows of $G_M^r$. So $(G_M^r)_1$ is the first $r \times r$ principle submatrix of $G_M$, and then $(G_M^r)_1 = G_{M1}$.

Partition matrix $G_M$ accordingly, and denote

$$G_M = \begin{bmatrix} G_{M1} & G_{M3} \\ G_{M2} & G_{M4} \end{bmatrix}$$

So $G_M^r = \begin{bmatrix} G_{M1} \\ G_{M2} \end{bmatrix}$. Equating the first $r$ columns of (5.2.10), we get the following,

$$Q_{12}^r + (Q_{12}^T)^r + G_M^T \begin{bmatrix} S_1 \\ 0 \end{bmatrix} + \begin{bmatrix} S_1 G_{M1} \\ S_2 G_{M2} \end{bmatrix} = 0$$

or

$$Q_{12}^r + (Q_{12}^T)^r + G_M^T \begin{bmatrix} S_1 \\ 0 \end{bmatrix} + \begin{bmatrix} S_1 \\ 0 \end{bmatrix} G_{M1} = - \begin{bmatrix} 0 \\ S_2 G_{M2} \end{bmatrix} \quad (5.2.11)$$

(5.2.6) – (5.2.11) gives

$$G_M^T H + HG_{M1} = \begin{bmatrix} 0 \\ S_2 G_{M2} \end{bmatrix} + J \quad (5.2.12)$$

where $J = Q_{12}^r + (Q_{12}^T)^r - Z_{12} - Z_{21}$ and $H = Z_{22} - \begin{bmatrix} S_1 \\ 0 \end{bmatrix}$. Let $W_1 = \begin{bmatrix} 0 \\ S_2 G_{M2} \end{bmatrix} + J$.

It is well known that

$$H = - \int_0^\infty e^{G_M^T \tau} W_1 e^{G_{M1}^T \tau} d\tau$$
Hence

\[ Z_{22} = \begin{pmatrix} S_1 \\ 0 \end{pmatrix} - M, \quad \text{where } M := \int_0^\infty e^{G_{m \tau}} W_1 e^{G_{m1 \tau}} d\tau \]  \hspace{1cm} (5.2.13)

So

\[ B_M^T Z_{22} B_{M1} = B_{M1}^T S_1 B_{M1} - B_M^T M B_{M1} \]  \hspace{1cm} (5.2.14)

Plug (5.2.14) into (5.2.1),

\[ ||\Sigma_{err}||_{\mathcal{H}_2} = \text{trace} \left\{ 2B_M^T MB_{M1} + B_{M2}^T S_2 B_{M2} \right\} \]  \hspace{1cm} (5.2.15)

Note that,

\[ \text{trace} \left\{ B_M^T MB_{M1} \right\} = \text{trace} \left\{ \int_0^\infty B_M^T e^{G_{m \tau}} W_1 e^{G_{m1 \tau}} B_{M1} d\tau \right\} \]

\[ = \text{trace} \left\{ \int_0^\infty e^{G_{m1 \tau}} B_{M1} B_M^T e^{G_{m \tau}} d\tau W_1 \right\} \]

\[ = \text{trace} \left\{ X^T W_1 \right\} \]  \hspace{1cm} (5.2.16)

where \( X = \int_0^\infty e^{G_{m \tau}} B_M B_{M2}^T e^{G_{m1 \tau}} d\tau \), and so \( X \) solves

\[ G_M X + X G_{M1}^T + B_M B_{M1}^T = 0 \]  \hspace{1cm} (5.2.17)

Recall \( W_1 = \begin{bmatrix} 0 \\ S_2 G_{M2} \end{bmatrix} + J \), and \( J = Z_{21} + Z_{12} - (Q_{12}^T)^r \). Therefore

\[ X^T W_1 = X^T \begin{bmatrix} 0 \\ S_2 G_{M2} \end{bmatrix} + J. \]

Plugging into (5.2.16), we obtain

\[ \text{trace} \left\{ B_M^T MB_{M1} \right\} = \text{trace} \left\{ X^T W_1 \right\} = \text{trace} \left\{ X^T \begin{bmatrix} 0 \\ S_2 G_{M2} \end{bmatrix} \right\} + \text{trace} \left\{ X^T J \right\} \]  \hspace{1cm} (5.2.18)
Notice that $Q_{22} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r, \sigma_{r+1}, \sigma_{r+2}, \cdots, \sigma_n)$. Plugging (5.2.18) into (5.2.15), gives

\[
||\Sigma_{err}||_{\mathcal{H}_2} = \text{trace} \left\{ 2B_M^TMB_{M1} + B_{M2}^T S_2 B_{M2} \right\}
= 2 \text{trace} \{ B_M^TMB_{M1} \} + \text{trace} \{ B_{M2}^T S_2 B_{M2} \}
= 2 \text{trace} \{ X^T \begin{bmatrix} 0 \\ S_2 G_{M2} \end{bmatrix} \} + 2 \text{trace} \{ X^T J \} + \text{trace} \{ B_{M2}^T S_2 B_{M2} \}
\leq 2 \| X \|_{\mathcal{F}} \| S_2 G_{M2} \|_{\mathcal{F}} + 2 \| X \|_{\mathcal{F}} \| J \|_{\mathcal{F}} + \| B_{M2}^T \|_{\mathcal{F}} \| S_2 B_{M2} \|_{\mathcal{F}}
\leq 2 \| X \|_{\mathcal{F}} \| G_{M2} \|_{\mathcal{F}} \sigma_{r+1} + \| B_{M2} \|_{\mathcal{F}}^2 \sigma_{r+1} + 2 \| X \|_{\mathcal{F}} \| J \|_{\mathcal{F}}
= c_1 \sigma_{r+1} + c_2 \| J \|_{\mathcal{F}},
\]

where $c_1 = 2 \| X \|_{\mathcal{F}} \| G_{M2} \|_{\mathcal{F}} + \| B_{M2} \|_{\mathcal{F}}^2$ and $c_2 = 2 \| X \|_{\mathcal{F}}$. 

From theorem 5.2.1, the first term is bounded by a constant times $\sigma_{r+1}$ where $\sigma_{r+1}$ is the largest neglected singular values or Hankel singular values of $Q_{22}$. The problem is the second term $c_2 \| J \|_{\mathcal{F}}$, if we could bound this term by a scalar times $\sigma_{r+1}$, then Theorem 5.2.1 gives a desired error bound.

Symmetrically from dual system, we can get the error bound associated with $P_{22}$.

**Corollary 5.2.2** Suppose $P_{22} = T = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$, where $T_1 = \text{diag}(\delta_1, \delta_2, \cdots, \delta_r)$, $T_2 = \text{diag}(\delta_{r+1}, \delta_{r+2}, \cdots, \delta_n)$. Then for SISO system,

\[
||\Sigma_{err}||_{\mathcal{H}_2} \leq c_1 \sigma_{r+1} + c_2 \| J \|_{\mathcal{F}}
\]
where \( J = \gamma_{12} + \gamma_{21} - P_{12}^r - (P_{12}^T)^r \), \( P_{12}^r \) and \( (P_{12}^T)^r \) consist of the first \( r \) columns of \( P_{12} \) and \( P_{12}^T \) respectively.

From Lemma 5.1.1, we know

\[
\|\Sigma_{\text{err}}\|_{\mathcal{H}_2} = \text{trace}\{B_M^T Q_{22} B_M - 2 B_M^T Z_{22} B_{M1} + B_{M1}^T Q_{22,1} B_{M1}\} = \text{trace}\{C_M P_{11} C_M^T - 2 C_M \gamma_{11} C_{M1}^T + C_{M1} P_{11,1} C_{M1}^T\}.
\]

From the above equations we can see in bounding the \( \mathcal{H}_2 \) norm of the error system, \( P_{11} \) and \( Q_{22} \) play the similar roles. We may try the same approach for \( P_{11} \).

Unfortunately, this approach will not apply to the reachability gramian. A similar derivation working with the Lyapunov equation for the reachability Gramian would result in the following argument.

Suppose

\[
AY + YA_1^T + BB_1^T = 0
\]

Partition the matrices,

\[
\begin{pmatrix}
0 & I \\
K_M & G_M
\end{pmatrix}
\begin{pmatrix}
Y_{11} & Y_{12} \\
Y_{21} & Y_{22}
\end{pmatrix} +
\begin{pmatrix}
Y_{11} & Y_{12} \\
Y_{21} & Y_{22}
\end{pmatrix}
\begin{pmatrix}
0 & K_{M1}^T \\
I & G_{M1}^T
\end{pmatrix} +
\begin{bmatrix}
0 \\
B_M
\end{bmatrix}
\begin{bmatrix}
0 \\
B_{M1}^T
\end{bmatrix} = 0
\]

Then

\[
Y_{21} + Y_{12} = 0 \quad (5.2.19)
\]

\[
K_M Y_{11} + G_M Y_{21} + Y_{22} = 0 \quad (5.2.20)
\]

\[
Y_{22} + Y_{11} K_{M1}^T + Y_{12} G_{M1}^T = 0 \quad (5.2.21)
\]
\[ K_M Y_{12} + G_M Y_{22} + Y_{21} K_{M1}^T + Y_{22} G_{M1}^T + B_M B_{M1}^T = 0 \quad (5.2.22) \]

From \( AP + P A^T + B B^T = 0 \), we have
\[
\begin{pmatrix}
0 & I \\
K_M & G_M
\end{pmatrix}
\begin{pmatrix}
P_{11} & P_{12} \\
P_{12}^T & P_{22}
\end{pmatrix}
+ 
\begin{pmatrix}
P_{11} & P_{12} \\
P_{12}^T & P_{22}
\end{pmatrix}
\begin{pmatrix}
0 & K_M^T \\
I & G_M^T
\end{pmatrix}
+ 
\begin{pmatrix}
0 & B_B^T \\
B_M & 0
\end{pmatrix}
= 0
\]

or
\[
\begin{pmatrix}
P_{12}^T & P_{22} \\
K_M P_{11} + G_M P_{12}^T & K_M P_{12} + G_M P_{22}
\end{pmatrix}
+ 
\begin{pmatrix}
P_{12} & P_{11} K_M^T + P_{12} G_M^T \\
P_{22} & P_{12}^T K_M^T + P_{22} G_M^T
\end{pmatrix}
+ 
\begin{pmatrix}
0 & 0 \\
0 & B_M B_B^T
\end{pmatrix}
= 0
\]

By equating each block, we obtain
\[ P_{12}^T + P_{12} = 0 \quad (5.2.23) \]
\[ K_M P_{11} + G_M P_{12}^T + P_{22} = 0 \quad (5.2.24) \]
\[ K_M P_{12} + G_M P_{22} + P_{12}^T K_M^T + P_{22} G_M^T + B_M B_B^T = 0 \quad (5.2.25) \]

Suppose \( P_{22} = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix} \). Equating the first \( r \) columns of (5.2.25) gives
\[ K_M P_{12}^r + G_M \begin{pmatrix} T_1 \\ 0 \end{pmatrix} + P_{12}^T (K_M)^r + \begin{pmatrix} T_1(G_M^T)^r \\ T_2(G_M^T)^r \end{pmatrix} + B_M B_B^T = 0 \]

It is easy to see \((G_M^T)^r = G_M^T \), and we denote \((G_M^T)^r = \begin{pmatrix} G_{M1}^T \\ G_{M3}^T \end{pmatrix} \). So
\[ K_M P_{12}^r + G_M \begin{pmatrix} T_1 \\ 0 \end{pmatrix} + P_{12}^T (K_M)^r + \begin{pmatrix} T_1 \\ 0 \end{pmatrix} G_M^T + B_M B_B^T = \begin{pmatrix} 0 \\ T_2 G_{M3}^T \end{pmatrix} \quad (5.2.26) \]
(5.2.22) – (5.2.26) gives

\[
G_M J + JG^T_{M1} = \begin{bmatrix} 0 \\ T_2G^T_{M3} \end{bmatrix} - U
\]

(5.2.27)

where \( U = K_M(Y_{12} - P^T_{12}) + (Y_{21} - P^T_{12})G^T_{M1} \), and \( J = Y_{22} - \begin{bmatrix} T_1 \\ 0 \end{bmatrix} \). Let \( W_2 = \begin{bmatrix} 0 \\ T_2G^T_{M3} \end{bmatrix} + U \). It is well known that

\[
J = - \int_0^\infty e^{G_M \tau} W_2 e^{G^T_{M1} \tau} d\tau
\]

Hence

\[
Y_{22} = \begin{bmatrix} T_1 \\ 0 \end{bmatrix} - N \quad \text{where} \quad N = \int_0^\infty e^{G_M \tau} W_2 e^{G^T_{M1} \tau} d\tau
\]

(5.2.28)

Note that in (5.2.19) to (5.2.25), the only equations related to Lyapunov equations are (5.2.22) and (5.2.25), and the \(||\Sigma_{err}||_{H_2}|| is only related to \( \mathcal{P}_{11} \) and has nothing to do with \( \mathcal{P}_{22} \). Thus, this approach has failed to provide a bound.

### 5.3 Error Bounds Associated with \( \mathcal{P}_{11} \) and \( \mathcal{Q}_{11} \)

In this section, we will give a global error bound for a second order error system in which \( \mathcal{P}_{11} \) (or \( \mathcal{Q}_{11} \)) is diagonal.
For a given second order system

\[
\begin{align*}
M\ddot{q}(t) + G\dot{q}(t) + Kq(t) &= B_0 u(t) \\
y(t) &= E_0 q(t)
\end{align*}
\]  

we can transform it to first order system \( \Sigma_1 = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \) by letting \( x = \begin{pmatrix} q \\ \dot{q} \end{pmatrix} \):

\[
\begin{align*}
\dot{x} &= Ax + Bu \\
y &= Cx
\end{align*}
\]  

where

\[
A = \begin{pmatrix} 0 & I \\ -K_M & -G_M \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ B_M \end{pmatrix}, \quad C = \begin{pmatrix} E_0 & 0 \end{pmatrix}
\]

and \( K_M = M^{-1}K, \ G_M = M^{-1}G, \ B_M = M^{-1}B_0 \). We know the transfer functions for first and second order systems are respectively,

\[
G(s) = C(sI - A)^{-1}B
\]

and

\[
H(s) = E_0 (s^2 M + sG + K)^{-1} B_0.
\]

So

\[
Y(s) = G(s)U(s) = H(s)U(s).
\]

For impulse response \((u(t) = \delta(t) \text{ and } U(s) = 1)\),

\[
Y(s) = H(s)U(s) = H(s) = E_0 (s^2 M + sG + K)^{-1} B_0.
\]
From the second equation of (5.3.1), in Frequency domain we have

\[ Y(s) = E_0 Q(s). \]

Where \( Q(s) \) is the Laplace transformation of \( q(t) \). So

\[ Q(s) = (Ms^2 +Gs + K)^{-1}B_0. \]

Without loss of generality, we may suppose

\[ M = I_n, \quad G = M^{-1}G, \quad K = M^{-1}K, \]

to be consistent with our second order model reduction algorithms based on balanced truncation or SVD. See Chapters 3 and 4. We denote the reduced system by \( \hat{\Sigma} \), and the realization by \( (\hat{M}, \hat{G}, \hat{K}, \hat{B}_0, \hat{E}_0) \). Let

\[ F(s) = (Ms^2 +Gs + K)^{-1}B_0. \]

Therefore

\[ \hat{F}(s) = (\hat{M}s^2 +\hat{G}s + \hat{K})^{-1}\hat{B}_0. \]

By Parseval’s theorem [56] [30],

\[
\mathcal{P}_{11} = \int_{0}^{\infty} q(t) q^*(t) \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(i\omega) F^*(i\omega) \, d\omega \\
\mathcal{P}_{22} = \int_{0}^{\infty} \dot{q}(t) \dot{q}^*(t) \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 F(i\omega) F^*(i\omega) \, d\omega.
\]

For first order system (5.3.2), by definition, the controllability gramian is

\[ \mathcal{P} = \int_{-\infty}^{\infty} x(t) x^*(t) \, dt \]
where \( x(t) \) is the internal state. So for second order system (5.3.1), the controllability gramian is

\[
P = \int_{-\infty}^{\infty} q(t) q^*(t) \, dt
\]

where \( q(t) \) is the internal state for second order system. Therefore \( P = \mathcal{P}_{11} \).

In this section, we suppose \( \Sigma \) has dimension \( n \), and the reduced system \( \hat{\Sigma} \) has dimension \( r \).

**Theorem 5.3.1** For a SISO second order system (5.3.1), suppose \( \mathcal{P}_{11} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \) is diagonal, where \( S_1 = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r) \), \( S_2 = \text{diag}(\sigma_{r+1}, \sigma_{r+2}, \ldots, \sigma_n) \), and \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \). Then

\[
\| \Sigma_{err} \|^2_{H_2} = \| \Sigma - \hat{\Sigma} \|^2_{H_2} \leq c_0(\sigma_{r+1} + \sigma_{r+2} + \cdots + \sigma_n)
\]  

(5.3.3)

where \( c_0 \) is a constant.

**Proof:** Let \( A(s) = Ms^2 + Gs + K \), and partition

\[
A(s) = \begin{bmatrix} A_{11}(s) & A_{12}(s) \\ A_{21}(s) & A_{22}(s) \end{bmatrix}, \quad E_0 = \begin{bmatrix} E_{01} & E_{02} \end{bmatrix}, \quad B_0 = \begin{bmatrix} B_{01} \\ B_{02} \end{bmatrix}
\]

where \( A_{11}(s) \) is the first \( r \times r \) principle submatrix of \( A(s) \), \( E_{01} \) consists of the first \( r \) columns of \( E_0 \), and \( B_{01} \) consists of the first \( r \) rows of \( B_0 \). So

\[
A(s)F(s) = B_0 \tag{5.3.4}
\]

Partition accordingly, \( F(s) = \begin{bmatrix} F_1(s) \\ F_2(s) \end{bmatrix} \). Then \( S = \mathcal{P}_{11} = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(i\omega)F^*(i\omega) \, d\omega \),
and

\[ S_1 = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_1(i\omega)F_1^\ast(i\omega) \, d\omega \]
\[ S_2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_2(i\omega)F_2^\ast(i\omega) \, d\omega \]
\[ 0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_2(i\omega)F_1^\ast(i\omega) \, d\omega \]

Moreover,

\[ \text{trace}\{S_1\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \|F_1(i\omega)\|^2_F \, d\omega, \quad \text{trace}\{S_2\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \|F_2(i\omega)\|^2_F \, d\omega \]

In reduced system $\hat{\Sigma}$, define

\[ \hat{A}(s) := \hat{M}s^2 + \hat{G}s + \hat{K} = A_{11}(s) \]

and define $\hat{F}$ via

\[ \hat{A}(s)\hat{F}(s) = B_0. \tag{5.3.5} \]

Observe from previous definition (5.3.4),

\[
\begin{pmatrix}
A_{11}(s) & A_{12}(s) \\
A_{21}(s) & A_{22}(s)
\end{pmatrix}
\begin{pmatrix}
F_1(s) \\
F_2(s)
\end{pmatrix}
= \begin{pmatrix}
A_{11}(s)F_1(s) + A_{12}(s)F_2(s) \\
A_{21}(s)F_1(s) + A_{22}(s)F_2(s)
\end{pmatrix}
= B_0 = \begin{pmatrix}
B_{01} \\
B_{02}
\end{pmatrix}
\]

Equating (1,1)–block,

\[ A_{11}(s)F_1(s) + A_{12}(s)F_2(s) = B_{01}. \]

Therefore combining with (5.3.5), we have

\[ F_1(s) = A_{11}^{-1}(s)[B_{01} - A_{12}F_2(s)] = \hat{F}(s) - A_{11}^{-1}(s)A_{12}F_2(s). \]
Let \( W(s) := A_{11}^{-1}(s)A_{12}(s) \). Thus
\[
\hat{F}(s) = F_1(s) + W(s)F_2(s).
\]

Let \( y = E_0 q \) and \( \dot{y} = \hat{E}_0 \dot{q} \) be the outputs for full and reduced systems respectively for the same input \( u \). Then in time domain,
\[
y(t) - \dot{y}(t) = E_0 q(t) - \hat{E}_0 \dot{q}(t),
\]
and in frequency domain,
\[
Y(s) - \dot{Y}(s) = E_0 (Ms^2 + Gs + K)^{-1} B_0 U(s) - \hat{E}_0 (\hat{M}s^2 + \hat{G}s + \hat{K})^{-1} \hat{B}_0 U(s) = E_0 F(s)U(s) - \hat{E}_0 \hat{F}(s)U(s).
\]

Thus
\[
H_e(s) = E_0 F(s) - \hat{E}_0 \hat{F}(s)
\]
is the transfer function of the error system \( \Sigma_{err} = \Sigma - \hat{\Sigma} \)

Using the formula for \( H_2 \) norm in frequency domain,
\[
\|\Sigma_{err}\|^2_{H_2} = \text{trace}\left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} H_e(i\omega) H_e^*(i\omega) \, d\omega \right\}
\]
\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{ (E_0 F(i\omega) - \hat{E}_0 \hat{F}(i\omega)) (E_0 F(i\omega) - \hat{E}_0 \hat{F}(i\omega))^* \} \, d\omega
\]
\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{ E_0 F(i\omega) (E_0 F(i\omega))^* \} \, d\omega
\]
\[
- 2\frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{ E_0 F(i\omega) (\hat{E}_0(i\omega) \hat{F}(i\omega))^* \} \, d\omega
\]
\[
+ \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{ \hat{E}_0 \hat{F}(i\omega) (\hat{E}_0(i\omega) \hat{F}(i\omega))^* \} \, d\omega.
\]

Notice that
\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{ E_0 F(i\omega)(E_0 F(i\omega))^* \} \, d\omega
\]
\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_{01} F_1(i\omega)(E_{01} F_1(i\omega))^*\} \, d\omega + \\
\frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_{02} F_2(i\omega)(E_{02} F_2(i\omega))^*\} \, d\omega
\]

and

\[
= \text{trace}\{E_{01} S_1 E_{01}^*\} + \text{trace}\{E_{02} S_2 E_{02}^*\},
\]

and

\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_0 F(i\omega)(\hat{E}_0 \hat{F}(i\omega))^*\} \, d\omega
\]

\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_0 F(i\omega)(\hat{E}_0 (F_1(i\omega) + W(i\omega)F_2(i\omega)))^*\} \, d\omega
\]

\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_0 F(i\omega)(E_0 F_1(i\omega))^*\} \, d\omega + \\
\frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_0 F(i\omega)F_2^*(i\omega)W^*(i\omega)E_{01}^*\} \, d\omega
\]

\[
= \text{trace}\{E_{01} S_1 E_{01}^*\} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_0 F(i\omega)F_2^*(i\omega)W^*(i\omega)E_{01}^*\} \, d\omega,
\]

and

\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{\hat{E}_0 \hat{F}(i\omega) (\hat{E}_0 \hat{F}(i\omega))^*\} \, d\omega
\]

\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_0 (F_1(i\omega) + W(i\omega)F_2(i\omega)) (F_1(i\omega) + W(i\omega)F_2(i\omega))^*E_{01}^*\} \, d\omega
\]

\[
= \text{trace}\{E_{01} S_1 E_{01}^*\} + \frac{1}{2\pi} \int_{-\infty}^{\infty} 2 \text{trace}\{E_0 F_1(i\omega)F_2^*(i\omega)W^*(i\omega)E_{01}^*\} \, d\omega
\]

\[
+ \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{(E_0 W(i\omega)F_2(i\omega))(E_0 W(i\omega)F_2(i\omega))^*\} \, d\omega.
\]

It follows that

\[
\|\Sigma_{err}\|_{F,2}^2 = \text{trace}\{E_{02} S_2 E_{02}^*\} - \frac{2}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{EF(i\omega)F_2^*(i\omega)W^*(i\omega)E_{01}^*\} \, d\omega
\]

\[
+ \frac{1}{2\pi} \int_{-\infty}^{\infty} 2 \text{trace}\{E_0 F_1(i\omega)F_2^*(i\omega)W^*(i\omega)E_{01}^*\} \, d\omega
\]

\[
+ \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{(E_0 W(i\omega)F_2(i\omega))(E_0 W(i\omega)F_2(i\omega))^*\} \, d\omega
\]
\[ \text{trace}\{E_{02}S_2E^{*}_{02}\} - \]
\[ \frac{2}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{\left(E_{01}F_1(i\omega) + E_{02}F_2(i\omega)\right)F^{*}_2(i\omega)W^{*}(i\omega)E^{*}_{01}\} \, d\omega \]
\[ + \frac{2}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{E_{01}F_1(i\omega)F^{*}_2(i\omega)W^{*}(i\omega)E^{*}_{01}\} \, d\omega \]
\[ + \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{\left(E_{01}W(i\omega)F_2(i\omega)\right)\left(E_{01}W(i\omega)F_2(i\omega)\right)^{*}\} \, d\omega \]
\[ = \text{trace}\{E_{02}S_2E^{*}_{02}\} - \frac{2}{2\pi} \int_{-\infty}^{\infty} 2 \text{trace}\{E_{02}F_2(i\omega)F^{*}_2(i\omega)W^{*}(i\omega)E^{*}_{01}\} \, d\omega \]
\[ + \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{\left(E_{01}W(i\omega)F_2(i\omega)\right)\left(E_{01}W(i\omega)F_2(i\omega)\right)^{*}\} \, d\omega \]
\[ = \text{trace}\{E_{02}S_2E^{*}_{02}\} + \]
\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{\left(E_{01}W(i\omega) - 2E_{02}\right)F_2(i\omega)\left(E_{01}W(i\omega)F_2(i\omega)\right)^{*}\} \, d\omega \]

Finally we can bound
\[
\left| \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{\left(E_{01}W(i\omega) - 2E_{02}\right)F_2(i\omega)\left(E_{01}W(i\omega)F_2(i\omega)\right)^{*}\} \right|
\]
\[ = \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{\left(E_{01}W(i\omega)\right)^{*}\left(E_{01}W(i\omega) - 2E_{02}\right) F_2(i\omega)F^{*}_2(i\omega)\} \right|
\]
\[ \leq \sup_{\omega}\|\left(E_{01}W(i\omega)\right)^{*}\left(E_{01}W(i\omega) - 2E_{02}\right)\|_2 \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}\{F_2(i\omega)F^{*}_2(i\omega)\} \, d\omega \]
\[ \leq \sup_{\omega}\|\left(E_{01}W(i\omega)\right)^{*}\left(E_{01}W(i\omega) - 2E_{02}\right)\|_2 \text{trace}\{S_2\}. \]

Plug into previous equations to obtain
\[
\|\Sigma_{\text{err}}\|_{\mathcal{H}_2}^2 \leq c_0 \text{trace}\{S_2\} = c_0(\sigma_{r+1} + \sigma_{r+2} + \cdots + \sigma_n)
\]

with \(c_0 = \sup_{\omega}\|\left(E_{01}W(i\omega)\right)^{*}\left(E_{01}W(i\omega) - 2E_{02}\right)\|_2 + \|E_0\|_2^2.\)

We know transfer functions for first and second order systems are
\[
G(s) = C(sI - A)^{-1}B
\]
and

\[ H(s) = E(s^2M + sG + K)^{-1}B_0 \]

So

\[ Y(s) = G(s)U(s) = H(s)U(s). \]

For dual system of first order system in (5.3.2),

\[ Y_d(s) = G^*(s)U(s) = B^*(sI - A^*)^{-1}C^*U(s). \]

So for dual system of second order system in (5.3.1),

\[ Y_d(s) = H^*(s)U(s) = B_0^*(s^2M^* + sG^* + K^*)^{-1}E_0^*. \]

We know the second order controllability gramian is defined to be the gramian of the state \( q(t) \) under impulse response:

\[ P = \int_0^\infty q(t)q^*(t) \, dt \]

which is the leading \( n \times n \) principal submatrix of the first order gramian \( P \) for the corresponding first order system, i.e. \( P = P_{11} \). By symmetry, we can see the controllability gramian \( P_d \) for the second order dual system is the leading \( n \times n \) principal submatrix of the corresponding gramian \( P_d \) which is \( Q \). So \( P_d \) is actually \( Q_{11} \), the leading \( n \times n \) principal submatrix of first order gramian \( Q \). So when applying SVD on \( Q_{11} \), and doing the same reduction procedures as to \( P = P_{11} \), we get the similar global error bound.
Theorem 5.3.2 For a SISO second order system (5.3.1), suppose $Q_{11} = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$ is diagonal, where $T_1 = \text{diag}(\delta_1, \delta_2, \cdots, \delta_r)$, $T_2 = \text{diag}(\delta_{r+1}, \delta_{r+2}, \cdots, \delta_n)$, and $\delta_1 \leq \delta_2 \leq \cdots \leq \delta_n$. Then

$$\|\Sigma_{\text{err}}\|_{\mathcal{H}_2}^2 = \|\Sigma - \hat{\Sigma}\|_{\mathcal{H}_2}^2 \leq c_1(\delta_{r+1} + \delta_{r+2} + \cdots + \delta_n) \quad (5.3.6)$$

where $c_1 = \sup_{\omega} \|W(i\omega)B_0_1(B_{01}^*W(i\omega)^{-2B_{02}})\|_2 + \|B_{02}\|_2^2$, $W(s) = A_{11}(s)^{-1}A_{12}(s)$, $A(s) = Ms^2 + Gs + K$, and $A(s) = \begin{bmatrix} A_{11}(s) & A_{12}(s) \\ A_{21}(s) & A_{22}(s) \end{bmatrix}$ with $A_{11}(s) \in \mathbb{C}^{r \times r}$. 

Similarly, by definitions of $\mathcal{P}$, $\mathcal{P}_{22}$, and using Parseval's theorem, we obtain

$$\mathcal{P}_{22} = \int_0^\infty \dot{x}(t)\dot{x}^*(t) \, dt = \int_{-\infty}^{\infty} \omega^2 F(i\omega)F^*(i\omega) \, d\omega.$$ 

One may think to use the same procedures as above to give the error bound for algorithms associated with $\mathcal{P}_{22}$. But unfortunately, the same approach cannot be applied to derive the error bound associated with $\mathcal{P}_{22}$ (and $Q_{22}$ symmetrically). But to provide a good error bound for algorithms associated with $\mathcal{P}_{22}$ and $Q_{22}$ is what we want, that will be for future work.
Chapter 6

Algorithms for Second Order Model Reduction Based on Krylov Projection

In this chapter, we will develop some algorithms based on Krylov projection. In Section 1, two algorithms are proposed based on Arnoldi factorization.

6.1 Some Algorithms for Second Order Model Reduction Based on Krylov Projection

The cost for Krylov projection is $\mathcal{O}(kn^2)$ where $k$ is the index of the Krylov subspace used to produce the projection matrices. And the gramian based method costs $\mathcal{O}(n^3)$. 
This makes algorithms based on Krylov projection, more efficient and favorable than balanced truncation or SVD methods in large scale problems. In experiments, Krylov based method works very fast, but the performance or accuracy is not as good as gramian based method.

There are two algorithms of Krylov based second order model reduction from [18]. In this section, we propose two algorithms.

Consider second order system

\[
\begin{align*}
M\ddot{q}(t) + G\dot{q}(t) + Kq(t) &= B_0 u(t) \\
y(t) &= E_0 q(t) + F_0 \dot{q}(t)
\end{align*}
\] (6.1.1)

Multiply the differential equation in (6.1.1) by $K^{-1}$ on the left, (6.1.1) then becomes

\[
\begin{align*}
K^{-1}M\ddot{q}(t) + K^{-1}G\dot{q}(t) + q(t) &= K^{-1}B_0 u(t) \\
y(t) &= E_0 q(t) + F_0 \dot{q}(t)
\end{align*}
\] (6.1.2)

Apply $r$-step Arnoldi factorization to $K^{-1}M$,

\[
K^{-1}MV = VH + fe_r^T.
\] (6.1.3)

By the properties of Arnoldi factorization, see [47] [48], we know $V^TV = I_r$, $H$ is upper Hessenberg with all subdiagonal elements non-negative, and $V^Tf = 0$. Therefore,

\[
V^TK^{-1}MV = V^TVH + V^Tfe_r^T = H
\] (6.1.4)

Suppose $V$ consists of the first $r$ columns of the orthogonal matrix $U \in \mathbb{R}^{n\times n}$. If the differential equation of (6.1.2) is multiplied by $U^T$ on the left, (6.1.2) becomes

\[
U^TK^{-1}MUU^T\ddot{q} + U^TK^{-1}GUU^T\dot{q} + U^TTUU^Tq = U^TK^{-1}B_0 u
\] (6.1.5)
\[ y = E_0 U U^T q + F_0 U U^T \dot{q} \]  

(6.1.6)

If we put \( z(t) = U^T q(t) \), then (6.1.5) and (6.1.6) become

\[ U^T K^{-1} M U \dot{z}(t) + U^T K^{-1} G U \dot{z}(t) + z(t) = U^T K^{-1} B_0 u(t) \]  

(6.1.7)

\[ y(t) = E_0 U z(t) + F_0 U z(t) \]  

(6.1.8)

Let the projection matrices be \( V \), which consists of the first \( r \) columns of \( U \). The reduction rules are then:

\[
\begin{align*}
\dot{\tilde{M}} &= V^T K^{-1} M V = H \\
\dot{\tilde{G}} &= V^T K^{-1} G V \\
\dot{\tilde{K}} &= I_r \\
\dot{\tilde{B}}_0 &= V^T K^{-1} B_0 \\
\dot{\tilde{E}}_0 &= E_0 V \\
\dot{\tilde{F}}_0 &= F_0 V.
\end{align*}
\]  

(6.1.9)

This results in the following Algorithm and Proposition.

**Algorithm 6.1.1** (Second order model reduction — Krylov method, applying Arnoldi factorization to \( K^{-1} M \))

1. compute \( r \)-step Arnoldi factorization to \( K^{-1} M \)

\[ K^{-1} M V = V H + f e_r^T; \]

2. let \( V \) be the projection matrix;

3. perform the projection as in (6.1.9) to get the reduced system \((\tilde{M}, \tilde{G}, \tilde{K}, \tilde{B}_0, \tilde{E}_0, \tilde{F}_0)\).
Proposition 6.1.1 The full order transformed second order system produced by Algorithm 6.1.1 is equivalent to the original system.

When applying to real systems, Algorithm 6.1.1 is very fast and efficient. The experimental results indicate that the accuracy in approximating the original systems is good. However it is not as good as gramian based algorithms that we presented in previous chapters. We could also think about applying the same approach to $G^{-1}M$, which has similar performance in experiments. However this does not make much sense from a physical point view. Algorithm 6.1.1 has no corresponding reduction rules for the corresponding first order model. It is natural to think about applying Arnoldi factorization to $M^{-1}K$. This approach preserves structure, but does not give good approximations in some of the real problems..

Algorithm 6.1.2 (Second order model reduction — Krylov method, applying Arnoldi factorization to $M^{-1}K$)

1. compute r-step Arnoldi factorization to $M^{-1}K$

\[ M^{-1}KV = VH + f e_r^T; \]

2. let $V$ be the projection matrix;

3. perform the projection:

\[ \hat{M} = I_r \]

\[ \hat{G} = V^T M^{-1} GV \]
\[
\hat{K} = H \\
\hat{B}_0 = V^T M^{-1} B_0 \\
\hat{E}_0 = E_0 V \\
\hat{F}_0 = F_0 V
\]

**Proposition 6.1.2** In Algorithm 6.1.2, the full order transformed system is equivalent to the original system. The orthogonal projection matrix for the corresponding first order system is:

\[
W = \begin{pmatrix} V \\ V \end{pmatrix}
\]

where \( V \) is from the \( r \)-step Arnoldi factorization of \( M^{-1} K \), i.e.,

\[
M^{-1} KV = VH + fe_r^T.
\]

**Proof:** The equivalence can be proved very similar to that in Proposition 6.1.1. The second part can be easily seen from equations (4.1.3) and (4.1.4).

Both algorithms work fast for large scale problems. Algorithm 6.1.1 in reduced system is not in second order form for corresponding first order one. Algorithm 6.1.2 has all the desired properties, the reduced system preserves second order structure for both second order system and the corresponding first order system. But in experiments, Algorithm 6.1.1 has a better approximation than Algorithm 6.1.2. See the following plots and tables by applying the algorithms to three real second order models.
Figure 6.1: Frequency response on building model with size $n = 24$ by Krylov method on $K^{-1}M$

<table>
<thead>
<tr>
<th>$n = 24$</th>
<th>$|\Sigma_{red}|^2_{\mathcal{H}_2}$</th>
<th>$|\Sigma_{err}|^2_{\mathcal{H}<em>2} / |\Sigma|^2</em>{\mathcal{H}_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 4$</td>
<td>$8.7359e-06$</td>
<td>$0.5059$</td>
</tr>
<tr>
<td>$r = 8$</td>
<td>$1.9895e-05$</td>
<td>$0.5522$</td>
</tr>
<tr>
<td>$r = 12$</td>
<td>$1.9930e-05$</td>
<td>$0.0181$</td>
</tr>
<tr>
<td>$r = 16$</td>
<td>$2.0339e-05$</td>
<td>$0.0027$</td>
</tr>
<tr>
<td>$r = 20$</td>
<td>$2.0453e-05$</td>
<td>$7.8839e-04$</td>
</tr>
<tr>
<td>$r = 24$</td>
<td>$2.0521e-05$</td>
<td>$1.1557e-13$</td>
</tr>
</tbody>
</table>

Table 6.1: $\mathcal{H}_2$ norm of building model with size $n = 24$ by applying Arnoldi on $K^{-1}M$
Figure 6.2: Frequency response on building model with size $n = 24$ by Krylov method on $M^{-1}K$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{\text{red}}|_{\mathcal{H}_2}^2$</th>
<th>$\frac{|\Sigma_{\text{err}}|_{\mathcal{H}<em>2}^2}{|\Sigma|</em>{\mathcal{H}_2}^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2.1782e-07</td>
<td>0.9985</td>
</tr>
<tr>
<td>8</td>
<td>3.0160e-07</td>
<td>0.9968</td>
</tr>
<tr>
<td>12</td>
<td>3.1163e-06</td>
<td>0.7845</td>
</tr>
<tr>
<td>16</td>
<td>3.9117e-06</td>
<td>0.5952</td>
</tr>
<tr>
<td>20</td>
<td>4.0057e-06</td>
<td>0.5720</td>
</tr>
<tr>
<td>24</td>
<td>2.0521e-05</td>
<td>2.5294e-13</td>
</tr>
</tbody>
</table>

Table 6.2: $\mathcal{H}_2$ norm of building model with size $n = 24$ by applying Arnoldi on $M^{-1}K$
Figure 6.3: Frequency response on building model with size $n = 21$ by applying Arnoldi on $K^{-1}M$

$$
\begin{array}{|c|c|c|}
\hline
n = 21 & \|\Sigma_{\text{red}}\|_{\mathcal{H}_2}^2 & \|\Sigma_{\text{err}}\|_{\mathcal{H}_2}^2 \\
\hline
r = 3 & 345.5242 & 1.2616e+04 \\
\hline
r = 6 & 6.8975e+05 & 2.5180e+07 \\
\hline
r = 9 & 0.1274 & 2.3860 \\
\hline
r = 12 & 6.5582e+06 & 2.3941e+08 \\
\hline
r = 15 & 2.5424 & 92.2554 \\
\hline
r = 18 & 0.0274 & 9.5265e-05 \\
\hline
r = 21 & 0.0274 & 1.2489e-07 \\
\hline
\end{array}
$$

Table 6.3: $\mathcal{H}_2$ norm of building model with size $n = 21$ by applying Arnoldi on $K^{-1}M$
Figure 6.4: Frequency response on building model with size $n = 21$ by applying Arnoldi on $M^{-1}K$

<table>
<thead>
<tr>
<th>$n = 21$</th>
<th>$|\Sigma_{red}|_{H_2}^2$</th>
<th>$|\Sigma_{err}|_{H_2}^2$</th>
<th>$|\Sigma|_{H_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 3$</td>
<td>$8.0713e-05$</td>
<td>$1.0031$</td>
<td></td>
</tr>
<tr>
<td>$r = 6$</td>
<td>$3.2724e-05$</td>
<td>$1.0014$</td>
<td></td>
</tr>
<tr>
<td>$r = 9$</td>
<td>$0.1115$</td>
<td>$4.9806$</td>
<td></td>
</tr>
<tr>
<td>$r = 12$</td>
<td>$5.3147e+07$</td>
<td>$1.9402e+09$</td>
<td></td>
</tr>
<tr>
<td>$r = 15$</td>
<td>$2.7287e+14$</td>
<td>$9.9613e+15$</td>
<td></td>
</tr>
<tr>
<td>$r = 18$</td>
<td>$4.6801e+11$</td>
<td>$1.7085e+13$</td>
<td></td>
</tr>
<tr>
<td>$r = 21$</td>
<td>$0.0274$</td>
<td>$1.7047e-07$</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.4: $H_2$ norm of building model with size $n = 21$ by applying Arnoldi on $M^{-1}K$
Figure 6.5: Frequency response on Aluminum model with size $n = 1734$, Krylov method, Applying Arnoldi to $K^{-1}M$
Figure 6.6: Frequency response on Aluminum model with size $n = 1734$, Krylov method, Applying Arnoldi to $M^{-1}K$
Chapter 7

Conclusions and Future Work

7.1 Summery and Conclusions

In this thesis, we focus on model reduction of second order systems. A number of algorithms are proposed based on gramian and Krylov projection. Gramian based algorithms are developed by approaches of balanced truncation and singular value decomposition, some of the algorithms have good error bounds. Krylov projection based algorithms are developed by means of Arnoldi factorization and spectral zeros, two of the algorithms are passivity preserving. Most of the algorithms proposed in this thesis have good performance in experiments.

First order model reduction is well developed, very little exist for second order
model reduction. We know second order system

\[
\begin{align*}
M\ddot{q}(t) + G\dot{q}(t) + K &= B_0u(t) \\
y(t) &= E_0q(t) + F_0\dot{q}(t)
\end{align*}
\tag{7.1.1}
\]

can be transformed to first order system

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t)
\end{align*}
\tag{7.1.2}
\]

by letting \(x = \begin{bmatrix} q \\ \dot{q} \end{bmatrix}\), where

\[
A = \begin{pmatrix} 0 & I \\ -M^{-1}K & -M^{-1}G \end{pmatrix}, \quad B = \begin{bmatrix} 0 \\ M^{-1}B_0 \end{bmatrix}, \quad C = \begin{bmatrix} E_0 & F_0 \end{bmatrix}
\]

When reducing second order systems, we apply projections to

\((I, M^{-1}G, M^{-1}K, M^{-1}B_0, E_0, F_0)\).

Even though by doing this way, symmetry of the data can get lost, but the performance in experiments are much better than applying projections to

\((M, G, K, B_0, E_0, F_0)\)

which keeps symmetry of the data.

### 7.2 Future work

Since a very few algorithms exist for second order model reduction, there are a lot to do for applying projections directly to second order systems. In Chapter 3, we proved
that in any case (any coordinate systems) the transformed first order controllability gramian $Q$ cannot be diagonal, so I would pay more attention on Krylov projection based second order model reduction.

So far, we proposed a few algorithms based on Krylov projections for second order model reduction. Many Krylov based algorithms for first order model reduction might be able to extend to second order ones by solving quadratic eigenvalue problems. There are be a lot of work to do in this direction.
Appendix A

Applications and Implementations

In this appendix, we apply our algorithms of performing SVD on the combinations of diagonal blocks of gramians in the corresponding first order system to two real second order systems: building models with sizes $n = 24$ and $n = 21$ respectively.
Figure A.1: Frequency response on building model with size $n = 24$ by SVD on $P_{11}$ and $P_{22}$.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{\text{red}}|_{\mathcal{H}_2}^2$</th>
<th>$|\Sigma_{\text{err}}|_{\mathcal{H}<em>2}^2/|\Sigma|</em>{\mathcal{H}_2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$5.5683e-04$</td>
<td>28.1274</td>
</tr>
<tr>
<td>8</td>
<td>0.0133</td>
<td>6.4923e+02</td>
</tr>
<tr>
<td>12</td>
<td>$5.6774e+21$</td>
<td>$2.7666e+26$</td>
</tr>
<tr>
<td>16</td>
<td>$8.2169e+04$</td>
<td>$4.0041e+09$</td>
</tr>
<tr>
<td>20</td>
<td>$2.4361e+27$</td>
<td>$1.1871e+32$</td>
</tr>
<tr>
<td>24</td>
<td>10.0482</td>
<td>$4.8965e+05$</td>
</tr>
</tbody>
</table>

Table A.1: $\mathcal{H}_2$ norm of building model with size $n = 24$ by SVD on $P_{11}$ and $P_{22}$.
Figure A.2: Frequency response on building model with size $n = 24$ by SVD on $Q_{11}$ and $Q_{22}$

<table>
<thead>
<tr>
<th>$n = 24$</th>
<th>$|\Sigma_{\text{red}}|_{\mathcal{H}^2}^2$</th>
<th>$\frac{|\Sigma_{\text{err}}|<em>{\mathcal{H}^2}^2}{|\Sigma|</em>{\mathcal{H}^2}^2}$</th>
<th>$\frac{|\Sigma_{\text{err}}|<em>{\mathcal{H}^2}^2}{|\Sigma</em>{\text{red}}|_{\mathcal{H}^2}^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 4$</td>
<td>0.83410200989824</td>
<td>4.064589211862721 $e + 04$</td>
<td>1.00001266470931</td>
</tr>
<tr>
<td>$r = 8$</td>
<td>0.02949278561274</td>
<td>1.434780869034447 $e + 03$</td>
<td>0.99833843457894</td>
</tr>
<tr>
<td>$r = 12$</td>
<td>0.00429163190511</td>
<td>2.121469297810374 $e + 02$</td>
<td>1.01443048866325</td>
</tr>
<tr>
<td>$r = 16$</td>
<td>0.02232504569758</td>
<td>1.08896292973860 $e + 03$</td>
<td>1.00100918981542</td>
</tr>
<tr>
<td>$r = 20$</td>
<td>5.069564607094859 $e + 32$</td>
<td>2.470373695836959 $e + 37$</td>
<td>1</td>
</tr>
<tr>
<td>$r = 24$</td>
<td>1.215557469363124 $e + 37$</td>
<td>5.923351275354629 $e + 41$</td>
<td>1</td>
</tr>
</tbody>
</table>

Table A.2: $\mathcal{H}^2$ norm of building model with size $n = 24$ by SVD on $Q_{11}$ and $Q_{22}$
Figure A.3: Frequency response on building model with size $n = 21$ by SVD on $\mathcal{P}_{11}$ and $\mathcal{P}_{22}$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$|\Sigma_{red}|_{\mathcal{H}_2}^2$</th>
<th>$\frac{|\Sigma_{err}|_{\mathcal{H}<em>2}^2}{|\Sigma|</em>{\mathcal{H}_2}^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.1051</td>
<td>4.6260</td>
</tr>
<tr>
<td>6</td>
<td>0.1323</td>
<td>3.5769</td>
</tr>
<tr>
<td>9</td>
<td>0.0195</td>
<td>0.7853</td>
</tr>
<tr>
<td>12</td>
<td>0.1677</td>
<td>6.7449</td>
</tr>
<tr>
<td>15</td>
<td>1.1633</td>
<td>42.3894</td>
</tr>
<tr>
<td>18</td>
<td>$7.2064e+04$</td>
<td>$2.6308e+06$</td>
</tr>
<tr>
<td>21</td>
<td>$3.2426e+11$</td>
<td>$1.1837e+13$</td>
</tr>
</tbody>
</table>

Table A.3: $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $\mathcal{P}_{11}$ and $\mathcal{P}_{22}$
Figure A.4: Frequency response on building model with size $n = 21$ by SVD on $Q_{11}$ and $Q_{22}$

<table>
<thead>
<tr>
<th>$n = 21$</th>
<th>$| \Sigma_{\text{red}} |_{\mathcal{H}^2}^2$</th>
<th>$| \Sigma_{\text{err}} |_{\mathcal{H}^2}^2$</th>
<th>$| \Sigma_{\text{err}} |<em>{\mathcal{H}^2}^2 / | \Sigma</em>{\text{red}} |_{\mathcal{H}^2}^2$</th>
<th>$| \Sigma_{\text{err}} |_{\mathcal{H}^2}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 3$</td>
<td>1.2680e-12</td>
<td>1.0000</td>
<td>2.1604e+10</td>
<td>1</td>
</tr>
<tr>
<td>$r = 6$</td>
<td>1.6832e-10</td>
<td>1.0000</td>
<td>1.6275e+08</td>
<td>1</td>
</tr>
<tr>
<td>$r = 9$</td>
<td>5.3448e+05</td>
<td>1.9511e+07</td>
<td>1.0000</td>
<td>1</td>
</tr>
<tr>
<td>$r = 12$</td>
<td>2.2399e+07</td>
<td>8.1770e+08</td>
<td>1.0000</td>
<td>1</td>
</tr>
<tr>
<td>$r = 15$</td>
<td>8.1315e+15</td>
<td>2.9685e+17</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$r = 18$</td>
<td>3.8042e+23</td>
<td>1.3888e+25</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$r = 21$</td>
<td>5.2531e+33</td>
<td>1.9177e+35</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table A.4: $\mathcal{H}_2$ norm of building model with size $n = 21$ by SVD on $Q_{11}$ and $Q_{22}$
Bibliography


