Model reduction methods

Citation for published version (APA):

Document status and date:
Published: 01/01/1993

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Model Reduction Methods
in theory and applied on a hybrid-drive line

S.P. van den Bosch

Eindhoven, 24 September 1993

WFW Rapport 93.133
Under the direction of Dr. Ir. F. E. Veldpaus
Prof. Dr. Ir. J.J. Kok

Eindhoven University of Technology, Faculty of Mechanical Engineering
Systems and Control Group
Abstract

The problem of approximating a high-order state-space system by a model of reduced order without incurring too much error is considered. For balanced truncation and optimal Hankel-norm approximation a multivariable transfer function $G(s)$ of order $n$ is approximated by a function $\hat{G}(s)$ of order $k < n$.

A balanced truncated realisation is derived, where the states are ordered by the influence of injected signals towards the output. This is calculated by means of controllability and observability grammians. Any "weak" subsystem, which contributes little to the impulse response matrix, is eliminated. The Hankel-norm of $(G(s) - \hat{G}(s))$ is minimised to get an optimal Hankel-norm approximation. This is an optimisation of the balanced truncation technique. The solution contains the sum of a causal and anticausal function that minimise the upper error-bound to the $(k+1)$-th Hankel singular value via an all-pass system.

For both balanced truncation and Hankel-norm approximation the reduction is made for the moduli of the transfer functions. Both algorithms are straightforward to implement. It is seen that the necessary stable system maintains stable. No poles are allowed on the imaginary axis, while these will be lost. For both methods error-bounds are given which are independent of frequencies. Hence, the frequency band one is interested in will not necessarily be the one who is the best estimated.

Also a static reduction procedure, called Guyan reduction is considered. The lower eigenfrequencies are approached by only considering the "external" coordinates. The approximation can be expanded by a dynamical part of the other coordinates, as is done by Graig-Bampton. There is also no theoretic error-bound given and the method can only be applied on very large scale systems ($n > 30$).

The model reduction techniques are employed on a hybrid-drive line. This system is nonlinear while the reduction methods are developed for linear systems. The system is reduced in one working point. It is seen that the Guyan reduction does not satisfy in this specific case.

The system is also divided into two linear components, which are reconnected after reduction with the central CVT. It is seen that there is no adequately approximation of the full system after reduction via balanced truncation. This is probably the result of the lacking frequency weighted error-criterium.
Contents

ABSTRACT ................................................................... 2

1 INTRODUCTION ............................................................ 5

2 MODEL REDUCTION BY MOORE ................................................ 6
   Linear independence of functions 6
   Linear dynamic systems 7
   Controllability 8
   Observability 10
   State-space transformations 12
   Balanced realisations 12
   Balanced truncation 13

3 OPTIMAL HANKEL-NORM APPROXIMATIONS ............................ 15
   The Hankel singular Values 15
   The Hankel operator 15
   The induced Hankel-norm 16
   The optimal anticausal function 17
   Solutions to the optimal Hankel-norm approximations 19
   One-step-at-the-time Hankel-norm approximations 20
   The error-bound for balanced truncation 20
   The $L^\infty$-error for optimal Hankel-norm approximations 22

4 LARGE-SCALE LINEAR SYSTEMS REDUCTION .................................... 23
   Guyan reduction 23
   Graig-Bampton 24

5 A FLYWHEEL-HYBRID DRIVE LINE ............................................. 25
   Dynamical equations 25
   Constant CVT-ratio 26
   The nonlinear CVT ratio 27
   The drive submodel 28
   The load submodel 29
   Nonlinear reduced model 29

6 CONCLUSIONS AND RECOMMENDATIONS ........................................ 31
   Conclusions 31
   Recommendations for further research 32
REFERENCES ................................................................. 33

NOMENCLATURE .............................................................. 34

APPENDIX A ................................................................. 35
   The Forbinius-norm 35
   The Lebesgue space 35
   Hardy spaces 35
   Theorem of Parseval 36

APPENDIX B ................................................................. 37
   All-pass transferfunctions 37

APPENDIX C ................................................................. 38
   Proof of theorem 1.4.2 38
   Proof $U$ is an unitary matrix 39
   Derivation Lyapunov equations for an error system 40

APPENDIX D ................................................................. 41
   A reduction algorithm for Hankel-norm approaches 41
1 Introduction

In the practice of an engineer are often high-order state-space system derived. Not seldom it is required to replace such a system by a reduced order model without incurring too much error. As a result of model reduction computationally less demanding and possible numerically more reliable control systems could be obtained. It could involve fewer components or computing resources and save in implementation while the performance is preserved.

An example where model reduction is required is the hybrid-drive line developed by Van der Graaf [1987]. It can be used in a passenger car, where there is next to the engine an additional flywheel and a Continue Variable Transmission. The time-varying CVT-ratio makes the system nonlinear. This makes the reduction procedure more complex while all current reduction methods are developed for linear systems. It is possible to reduce the system in a working point of the CVT-ratio, whereafter the obtained model is made nonlinear again. It is also possible to divide the system into several components, which become reduced separately if they are linear. Subsequently the components could be reconnected.

First model reduction by balanced truncation and Hankel-norm approximations will be considered. By these techniques it is possible to calculate the achievable error between the frequency responses of the full-order model and any reduced order model. Also a lower error can be computed. The description of both reduction methods is aimed at a comprehensible survey for control engineers.

Moore made "the first natural step in model reduction" to determine his, so called, second order modes or Hankel singular values. Therefor he reviewed the system from a signal injection viewpoint rather then by model parameters \((A,B,C,D)\). The transfer will be characterised by controllability and observability. A balanced realisation will be derived, where the states are ordered by the influence of injected signals toward the output. After that any "weak" subsystem, which contributes little to the impulse response matrix, will be eliminated.

The balanced truncation technique can be optimised by optimal Hankel-norm approximation. Therefor the largest Hankel singular value of the difference between the transfer function and its approximation is minimised. Whereas this is the upper bound. The error-system is made all-pass, while the moduli of such a system are constant. By approximating the system by the sum of a causal and anticausal function the achievable error is minimised. A anticausal function is not stable and hence it cannot be used for reduction. However the lower bound can nearly be reached.

We will also consider a static reduction procedure, called Guyan reduction. The lower eigenfrequencies are approached by only considering the "external" coordinates, which are the ones we are interested in. The approximation can be expanded by a dynamical part of the other coordinates, as is done by Graig-Bampton.

We released the model reduction techniques on the hybrid-drive line and came to some conclusions and recommendations. These are given in the second part of this report.
Model reduction by Moore

The main goal of model reduction techniques is to find a lower order model of the original system which has effectively the same impulse response matrix. This is a contradiction. Therefore a measure of the transfer is searched. Moore characterises the transfer in terms of responses to injected signals rather than by model parameters \((A,B,C,D)\).

In the controllability grammian will be found a measurement for the transfer from input signals to states. A measurement of the transfer from states to output is given by the observability grammian. Through join these both grammians is made distinct which input signals have a great influence on the output. In this manner a system balanced realisation is developed. It can be used to trace (a combination of) states that have none or less influence on the impulse response matrix.

2.1 Linear independence of functions

First we will define linear independency. Let \(f(t): \mathbb{R} \rightarrow \mathbb{R}^n\), with \(i = 1,2,...,n\), be some functions which are elements of the column \(f\). Let \(\lambda \in \mathbb{R}^n\) be a column with \(n\) constant numbers as elements. Now is said that the functions \(f_1,f_2,...,f_n\) are linear independent on the interval \([t_0,t_1]\) in case \(\lambda^T f(t) = 0\) for all \(t \in [t_0,t_1]\) gives \(\lambda = 0\).

To appoint if the functions in the column \(f\) are linear independent we use the so called grammian or Gramm-matrix:

\[
G(t_0,t_1) = \int_{t_0}^{t_1} f(t) f^T(t) \, dt \quad (2.1.1)
\]

The grammian \(G(t_0,t_1)\) is positif, semi definite. By analogy the grammian is defined for a matrix \(F: \mathbb{R} \rightarrow \mathbb{R}^n \times \mathbb{R}^m\)

\[
G(t_0,t_1) = \int_{t_0}^{t_1} F(t) F^T(t) \, dt \quad (2.1.2)
\]

We will prove that the next theorem contains a necessary and sufficient condition for linear independency of functions.

**Theorem:**
the functions \(f_i(t): \mathbb{R} \rightarrow \mathbb{R}^n\), with \(i = 1,2,...,n\), are linear independent on \([t_0,t_1]\) if and only if \(\det G(t_0,t_1) \neq 0\), thus as the grammian is non-singular.

**Proof:**
If the functions \(f_i(t)\) are linear dependent on \([t_0,t_1]\) than there exist an constant vector \(\lambda \neq 0\) in such a manner that \(\lambda^T f(t) = 0\) for all \(t \in [t_0,t_1]\). This implies that

---

* In this paper we will only use real rational functions. The derivations made can also be made for complex rational functions. In practice all functions will be real.
\[ \lambda^T G(t_0, t_1) = \int_{t_0}^{t_1} \lambda^T f(t) f'(t) \, dt = 0 \]

thus \( \text{det} [G(t_0, t_1)] = 0 \). A necessary condition for the linear independence of the functions \( f(t) \) is that the grammian \( G(t_0, t_1) \) must be non-singular. It is also sufficient. For \( \text{det} [G(t_0, t_1)] = 0 \) there exist at least one constant vector \( \lambda \neq 0 \) such that \( \lambda^T G(t_0, t_1) \lambda = 0 \). Then

\[ 0 = 0 \cdot \lambda = \lambda^T G(t_0, t_1) \lambda = \int_{t_0}^{t_1} \lambda^T f(t) f'(t) \lambda \, dt = \int_{t_0}^{t_1} \| \lambda^T f(t) \|^2 \, dt \]

which implies that \( \lambda^T f(t) = 0 \) for all \( t \in [t_0, t_1] \). Hence, the functions \( f(t) \) are linear independent.

Analogues we compose the next theorem for function matrices without proof.

**Theorem:**
the columns of the matrix \( F : \mathbb{R} \rightarrow \mathbb{R}^r \times \mathbb{R}^m \) are linear independent on \( t \in [t_0, t_1] \) if and only if \( \text{det} [G(t_0, t_1)] = 0 \).

### 2.2 Linear dynamic systems

Look at the standard linear dynamical system \((A(t), B(t), C(t), D(t))\), i.e. a system in state space approach described by

\[ \begin{align*}
    x(t) &= A(t)x(t) + B(t)u(t) \quad ; \quad t \geq t_0 \quad ; \quad x(t_0) = x_0 \\
    y(t) &= C(t)x(t) + D(t)u(t)
\end{align*} \tag{2.2.1} 
\]

with \( x(t) \in \mathbb{R}^n, \; u(t) \in \mathbb{R}^m, \; y(t) \in \mathbb{R}^p \) and \( A, B, C, D \) matrices to match. The transfer function of the system is:

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

The common solution is given by

\[ \begin{align*}
    x(t) &= \Phi(t)\Phi^{-1}(t_0) x_0 + \int_{t_0}^{t} \Phi(t)\Phi^{-1}(\tau) B(\tau) u(\tau) \, d\tau \quad , \quad t \geq t_0 \\
    y(t) &= C(t)\Phi(t_0, t_0) x_0 + C(t) \int_{t_0}^{t} \Phi(t)\Phi^{-1}(\tau) B(\tau) u(\tau) \, d\tau + D(t)u(t) \quad , \quad t \geq t_0
\end{align*} \tag{2.2.2} \]

where \( \Phi(t) = A(t)\Phi(t) \) for \( t \geq 0; \; \Phi(0) = I \).

While the impulse response is central to the development of the model reduction techniques this specific response is highlighted. Take \( u(t) = e_i \delta(t - t_0) \) where \( \delta(t) \) is the impulse function and \( e_i \) is a column with \( (e_i)_j = \delta_{ij} \). Then by (2.2.2)

\[ x(t) = \Phi(t)\Phi^{-1}(t_0) x_0 + \Phi(t)\Phi^{-1}(t_0) B(t_0) e_i \tag{2.2.3} \]

We define the response to an impulse at input \( i \) on moment \( t_0 \):
\[ h_i(t,\tau) = \Phi(t) \Phi^{-1}(\tau) B(\tau) e_i \]
\[ H(t,\tau) = [h_1(t,\tau), h_2(t,\tau), \ldots, h_m(t,\tau)] = \Phi(t) \Phi^{-1}(\tau) B(\tau) \]

Substituting this definition in (2.2.3) gives
\[ x(t) = \Phi(t) \Phi^{-1}(t_0) x_0 + \int_{t_0}^{t} H(t,\tau) u(\tau) \, d\tau \]

To get some insights on the matrix \( H \) take \( x_0 = 0 \). This implies that
\[ x(t) = \int_{t_0}^{t} H(t,\tau) u(\tau) \, d\tau \] (2.2.4)

and thus matrix \( H(t,\tau) \) fully determines how input \( u \) works on the state \( x \). This immediately leads to the question if there exist an input \( u : [t_0, t] \to \mathbb{R}^m \) such that the state \( x_0 = 0 \) changes into \( x(t) = x_\tau \) on moment \( t \) for every \( x_\tau \).

### 2.3 Controllability

For system (2.2.1) controllability is defined.

**Definition:**

the state \( x_0 \) of system (2.2.1) is called controllable on the interval \([t_0, t]\) with \( t > t_0 \) if there exist at least one input \( u : [t_0, t] \to \mathbb{R}^m \) which the system sends from state \( x_0 \) on moment \( t_0 \) to state \( 0 \) on moment \( t \).

**Definition:**

the system (2.2.1) is completely controllable on the interval \([t_0, t]\) with \( t > t_0 \) if every state \( x_0 \) of that system is controllable on the interval.

Notice that the definition of controllability also can be defined as follows:

**Definition:**

the system (2.2.1) is completely controllable on the interval \([t_0, t]\) with \( t > t_0 \) if it from state \( 0 \) on moment \( t_0 \) can be send to each state \( x_\tau \in \mathbb{R}^n \) on moment \( t \).

Without proof we remark that these two definitions are equivalent. We will use the latter definition. It is seen by (2.2.4) that the system (2.2.1) is completely controllable on the interval \([t_0, t]\) if and only if there for each \( x_\tau \in \mathbb{R}^n \) exist an input \( u : [t_0, t] \to \mathbb{R}^m \) such that
\[ \int_{t_0}^{t} H(t,\tau) u(\tau) \, d\tau = x_\tau \] (2.3.1)

We expect that this can only be true if the columns of the matrix \( H(t,\tau) \) on the interval \([t_0, t]\) are linear independent, what can be interpreted by the following theorem.

**Theorem:**

the system (2.2.1) is completely controllable on the interval \([t_0, t]\) if and only if the controllability grammian \( P(t_0, t) \)
is non-singular.

**Proof:**
To proof sufficiency, let \( P(t_0,t) \) be regular. We choose as input
\[
u(t) = H^T(t,\tau)P^{-1}(t_0,\tau)x,\]

It is seen that this input satisfies equation (2.3.1). Thus the system is completely controllable on \([t_0,t]\) if \( P(t_0,t) \) is non-singular, what implies that the rows of \( H(t,\tau)\Phi(t)\Phi^{-1}(\tau)B(t) \) must be linear independent on \([t_0,t]\). It is also proven to be a necessary condition by Jamshidi [1986].

We will proof that the above mentioned requirement on the controllability grammian is equivalent to the next theorem for a constant system (Van Campen [1990]).

**Theorem:**
the system (2.2.1) is completely controllable if the rank of the controllability matrix \( M \),
\[
M = [B, AB, A^2B, ..., A^{n-1}B],
\]
is equal to \( n \).

**Proof:**
Suppose that the rows of \( \Phi^{-1}(t)B \) are linear dependent for system (2.2.1). Than there is at least one constant vector \( \lambda \neq 0 \) such that \( \lambda^T e^{-At}B=0 \) for all \( t \). Thus for \( t=0 \) this implies \( \lambda^TB=0 \). Besides
\[
\frac{d}{dt}k(\lambda^T e^{-At}B) |_{t=0} = (-1)^T \lambda^T A^kB = 0
\]
\[
\lambda^T [B, AB, A^2B, ..., A^{n-1}B] = \lambda^TM = 0
\]
with \( M = [B, AB, A^2B, ..., A^{n-1}B] \), that is the controllability matrix. Here the rank of \( M \) is smaller then \( n \) while \( A \) is in the \( n \)-dimensional space. It is seen that when the rows of \( \Phi^{-1}(t)B \) want to be linear independent, the rank of \( M \) must be equal to \( n \).

This implies that when the rows of \( H(t,\tau) \) are linear independent the state \( x \), on moment \( t \), can be send inside the whole state space. We say that the couple \( \{A,B\} \) is completely controllable.

It is easily verified that the controllability grammian \( P \) satisfies linear matrix equations (Lyapunov equations) by considering the corresponding matrix differential equations.
\[
\frac{\partial P(t_0,t)}{\partial t} = A(t)P(t_0,t) + P(t_0,t)A^T(t) + B(t)B^T(t) ; \quad P(t_0,t_0) = 0
\]
if \( A \) and \( B \) are constants. For a stable system (2.2.1) holds that \( \lim_{t \to \infty} P(t_0,t)=0 \). Let \( P_s=P(t_0,t) \) be a stationary value for \( t \to \infty \) what gives
\[
AP_s + P_sA^T + BB^T = 0
\]

To give some physical insights on the controllability grammian consider the following minimum energy problem with regard to the system
\[ x(t) = Ax(t) + Bu(t) ; \quad x(t) = \int_{0}^{t} e^{-At}Bu(t) \, dt \]

We desire to send the system into a certain state \( x_0 \) with an optimal input \( u(t) \), that is with minimum energy, thus

\[ \text{Min} \ J(u) = \int_{0}^{T} u^T(t)u(t) \, dt \]

Hereby is seen that \( x_0 = \int_{0}^{\infty} e^{At}B u(-t) \, dt \) and \( P(0,\infty) = \int_{0}^{\infty} e^{At}BB^T e^{A^T} \, dt \). By multiplying this last equation on both sides with \( P^{-1}x_0 \) is shown:

\[ x_0 = \int_{0}^{\infty} e^{AT}BB^T e^{A^T} P^{-1} x_0 \, dt \]

what implies that

\[ u(t) = B^T e^{-A^T} P^{-1} x_0 \]

and thus

\[ J(u_{opt}) = x_0^T P^{-1} x_0 \]

Hence, if \( P \) is "small" there will be some states that can only be reached if a large input energy is used.

The matrix \( P \) gives a measurement for the controllability of states.

### 2.4 Observability

For system (2.2.1) is observability considered.

**Definition:**

the state \( x_{opt} \neq 0 \) of system (2.2.1) is called observable on the interval \([t_0,t]\) with \( t > t_0 \) if the knowledge of input \( u[t_0,t] \to \mathbb{R}^m \) and output \( y[t_0,t] \to \mathbb{R}^n \) is enough to determine \( x(t) \) on moment \( t \).

**Definition:**

the system (2.2.1) is completely observable on the interval \([t_0,t]\) with \( t > t_0 \) if every state \( x_{opt} \neq 0 \) of that system is observable on the interval.

The common solution of system (2.2.1) is given by (2.2.2) while the free response is given by

\[ y_c(t) = C(t) \Phi(t) \Phi^{-1}(t_0) x_0 , \quad t \geq t_0 \]

and the forced response by

\[ y_f(t) = C(t) \int_{t_0}^{t} \Phi(t) \Phi^{-1}(\tau) B(\tau) u(\tau) \, d\tau + D(t) u(t) , \quad t \geq t_0 \]

For a given input the forced response in directly to calculate, therefor we only consider the free response.

Hence we look at a linear, continue, time invariant system without input signals.

* See appendix (A.2) for the definition of the Lebesgue-space.
\[
\dot{x}(t) = A(t)x(t) \quad t \geq t_0, \quad x(t_0) = x_0
\]
\[
y(t) = C(t)x(t)
\]

(2.4.1)

Analogues to controllability can be said to observability that:

**Theorem:**

the system (2.2.1) is completely observable on the interval \([t_0, t]\) if and only if the observability gramian \(Q(t_0, t)\)

\[
Q(t_0, t) = \int_{t_0}^{t} \Phi^{-T}(t_0) \Phi^{-T}(\tau) C^T(\tau) C(\tau) \Phi(\tau) \Phi^{-1}(t_0) \, d\tau
\]

is non-singular.

**Proof:**

For (2.4.1) is derived that \(y(t) = C(t)\Phi(t)\Phi^{-1}(t_0)x_0\). By multiplying both sides of this equation by \(\Phi^{-1}(t_0)\Phi^{-1}(t)\) and integrating over \([t_0, t]\) applies

\[
\int_{t_0}^{t} \Phi^{-T}(t_0) \Phi^{-T}(\tau) C^T(\tau) y(\tau) \, d\tau = Q(t_0, t)x_0
\]

where \(Q(t_0, t)\) is the observability gramian. If this gramian is regular applies

\[
x_0 = Q^{-1}(t_0, t) \int_{t_0}^{t} \Phi^{-T}(t_0) \Phi^{-T}(\tau) C^T(\tau) y(\tau) \, d\tau
\]

Hence, the couple \(\{A, C\}\) is completely observable on \([t_0, t]\) if the columns of \(C(t)\Phi(t)\Phi^{-1}(t_0)\) are linear independent on \([t_0, t]\).

In analogy by controllability there can be proven that the above requirement is equivalent to the following theorem (Van Campen [1990]). The proof is not threaded.

**Theorem:**

the system (2.2.1) is completely observable if the rank of the observability matrix \(N\)

\[
N = [C, CA, CA^2, \ldots, CA^{n-1}]^T
\]

is equal to \(n\).

By considering the corresponding matrix differential equations it is seen that for a stable system (2.2.1) the observability gramian satisfies the following Lyapunov equation

\[
A^TQ_x + Q_xA + C^TC = 0
\]

when \(A, C\) are constants and \(Q_x = Q(t_0, t)\) for \(t \to \infty\).

If the observability gramian is nearly singular then is seen that some initial states are poor to observe. For "small" \(Q\) the system is poor to observe. Because when the system is released from \(x(0) = x_0\) with \(u(t) = 0, \quad t \geq t_0\) then

\[
\int_{t_0}^{t} y^T(t) y(t) \, dt = x_0^T Q x_0
\]

If \(Q\) is nearly singular then some initial conditions will have little effect on the output.
2.5 State-space transformations

It is desirable to reduce the original system by eliminating a "weak" subsystem that contributes a little to the impulse response matrix. Therefore is needed to reorder the coordinates by state-space transformations. We restrict the continuation by only considering time-invariant systems.

\[ \dot{x}(t) = Ax(t) + Bu(t) , \quad t \geq t_0 ; \quad x(t_0) = x_0 \]  

\[ y(t) = Cx(t) + Du(t) \]

This system is called a realisation of the system with state \( x \). It is possible to reorder the state-space coordinates by changing \( x \) into \( x^* = Tx \), with \( T \) a square and non-singular matrix. Then applies

\[ \dot{x}^* = A^*x^* + B^*u \quad ; \quad A^* = TAT^{-1} \quad ; \quad B^* = TB \]

\[ y = C^*x^* + D^*u \quad ; \quad C^* = CT^{-1} \quad ; \quad D^* = D \]

The latter realisation of the system with state \( x^* \) has the same input-output characteristics as model (2.5.1). The transfer function and eigenvalues of the system matrix are invariant under state-space transformations. It is easily shown how other values are under state-space transformation.

\[ \Phi^*(t_0, t) = T\Phi(t_0, t)T^{-1} \]

\[ P^*(t_0, t) = TP(t_0, t)T^T \]

\[ Q^*(t_0, t) = T^{-T}Q(t_0, t)T^{-1} \]

and \( P^*Q^* = TPQT^T \). Hence, also the eigenvalues of \( PQ \) are invariant under state-space transformations and are therefore input-output invariants.

2.6 Balanced realisations

Notice that the controllability grammian represents a transformation from input to states, while the observability gives a transformation from states to outputs. It is desirable to couple these two grammians. While the eigenvalues of \( PQ \) are invariant under state-space transformations it shows that these eigenvalues are a mapping from past inputs to future outputs. Therefore we try to make \( PQ \) diagonal by there eigenvalues. Consider the eigenvalue problem of \( PQ \).

\[ (PQ)\cdot v = \lambda v \]

where \( \lambda \) is an eigenvalue and \( v \) the corresponding eigenvector. For the derivation of a balanced realisation we only consider minimal systems, i.e. systems which are completely controllable and observable. Since the observability grammian is symmetric and semi positive-definite \( Q = Q^T \) and \( Q \geq 0 \). Let \( Q \) have the Cholesky factorisation \( Q = R^T R \).

\[ PR^TRv = \lambda v \]

When multiplying this equation by \( R \) and substitute \( w = Rv \) is seen

\[ (RPR^T)\cdot w = \lambda w \]

and notice that \( RPR^T \) will be a semi positive-definite and symmetric matrix and thus the eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \) will be non-negative. Let \( \sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_n^2 \) be the eigenvalues of \( RPR^T \) and \( W = [w_1, w_2, \ldots, w_n] \) an
orthogonal matrix with the corresponding eigenvectors. Now the diagonal form of $RPR^T$ is given by

$$RPR^T = W\Sigma^2 W^T ; \quad \Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

Since the transformation of coordinates for the controllability grammian was given by (2.5.2), let

$$P = T^{-1}\Sigma T^{-T}$$

thus

$$(RT^{-1})\Sigma (RT^{-1})^T = W\Sigma^2 W^T$$

$$RT^{-1} = W\Sigma^{1/2}$$

The balancing transformation is represented by

$$T = \Sigma^{1/2} W^T R$$

(2.6.1)

and is not unique, because it is determined by the Cholesky factorisation. Since

$$Q = R^T R = T^T \Sigma^{1/2} W^T W \Sigma^{1/2} T = T^T \Sigma T = Q$$

the controllability and observability grammians are equal and diagonal

$$P_{bal} = Q_{bal} = \Sigma$$

(2.6.2)

The elements on the diagonal of both grammians determine a measurement for the transformation from inputs to future outputs for the system in transformed coordinates. Therefore these grammians can successfully be used to reduce the original system.

For the latter procedure it is necessary that the system (2.5.1) is stable and minimal, while $P$ and $Q$ must be non-singular. Hence, the system need to be fully controllable and observable. Glover proposes in his paper a similar method which is able to balance also non-minimal systems. Therefore the system is partitioned in four components: controllable and observable, controllable and not observable, observable and not controllable, and not controllable, not observable and not controllable.

### 2.7 Balanced truncation

For the reduction of a model is needed to partitioned the balanced realisation

$$A_{bal} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B_{bal} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C_{bal} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \quad D_{bal} = D$$

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$$

with $A_{11} \in \mathbb{R}^{2n \times 2n}$, $B_i \in \mathbb{R}^{2n \times m}$, $C_i \in \mathbb{R}^{m \times 2n}$, $D \in \mathbb{R}^{m \times m}$ and $k < n$. There ought to originate a reduced system with almost the same input-output characteristics as the original system:

$$\dot{x}_r = A_{11} x_r + B_1 u$$

$$y = C_1 x_r + D u$$

(2.7.1)

The main idea underlying the model reduction work is to eliminate any "weak" subsystem which contributes little to the impulse response matrix. This implicitly defines the meaning of a "dominant"
subsystem: it is one whose impulse response matrix is close to that of the full model. It is not feasible to use this definition directly, we introduce instead the concept of "internal" dominance. This while every dominant subsystem is at least internal dominant. To determine such a system the Forbinius-norm is used Definition 1.7.1:

the system \( (A_n,B,C) \) is an internally dominant subsystem if the original full model \( (A,B,C) \) can be balanced and organised such that

\[
\mathbf{1} \Sigma^2_1 \mathbf{1}_r > \mathbf{1} \Sigma^2_2 \mathbf{1}_r
\]

This means that signal injections in an internal dominance subsystem involve much stronger signal outputs than for any weak subsystem. The result is that there exist an internally dominant subsystem of order \( k \) if and only if

\[
\left( \sum_{i=1}^{k} \sigma_i^2 \right)^{1/2} > \left( \sum_{i=k+1}^{n} \sigma_i^2 \right)^{1/2}
\]

and the full system can be balanced resulting in \( P_{\text{BAL}}=Q_{\text{BAL}}=\Sigma \).

If there are some states inferior to control and to observe, the matrix \( PQ \) is nearly singular. This implies that the balanced realisation is mathematical hard to make. This is a contradiction while it should be simple to reduce the system. To use the Schur-decomposition Safonov and Chiang [1988] develop a method to obtain a truncated model without a balanced realisation.

To review the model reduction method is needed to determine the error in the transfer function, \( G(s) - \hat{G}(s) \). Where the original transfer function \( G(s) \) is approximated by \( \hat{G}(s) \). Glover determines an upper bound for model reduction problems by a balanced realisation. The \( L^\infty \)-error is given by

\[
\| G(j\omega) - \hat{G}(j\omega) \|_\infty = \sigma_{\text{max}}(G(j\omega) - \hat{G}(j\omega)) \leq 2 \sum_{i=1}^{n} \sigma_i \| G(j\omega) \| \quad \forall \omega
\]

The derivation is given in the next chapter. It will be seen that the given method for model reduction by Moore is not optimal. Notice that for balanced truncation the moduli are approximated, while the phase-change is neglected.

---

* See appendix (A.1) for the definition of the Forbinius-norm.
** See appendix (A.2) for norms in the Lebesgue-space.
3 Optimal Hankel-norm approximations

On the basis of the balanced state-space realisations, derived in section (2), an optimal Hankel-norm approximation can be determined. Therefore the Hankel-norm difference between the original transfer function and its approximation is minimised. To split up the original system into a causal and anticausal system the optimal reduction is achieved. In the subsequent analysis the anticausal system can not be applied. Nevertheless the lower error bound can be well estimated.

3.1 The Hankel singular Values

Consider a standard linear, time-invariant dynamical system \((A, B, C, D)\) where the transfer function of the system is \(G(s) = D + C(sI - A)^{-1}B\). For such a system the controllability and observability grammians are defined by (2.3.2) and (2.4.2). While the eigenvalues of \(PQ, (\lambda_j(PQ))\), are invariant under state-space transformations the following definition is used.

**Definition:**
For a stable system the Hankel singular values of \(G(s)\) are defined as

\[
\sigma_j(G(s)) = |(\lambda_j(P))|^{1/2}
\]

where by convention \(\sigma_1(G(s)) \geq \sigma_\infty(G(s))\).

3.2 The Hankel operator

Consider a linear, time-invariant dynamical system \((A, B, C)\) with input \(u \in L^2(-\infty, \infty)\) and output \(y \in L^2(-\infty, \infty)\), where \(y = Cx\).

**Definition:**
The Hankel operator for the considered system is defined as: \(\Gamma_G : L^2[0, \infty) \to L^2[0, \infty)\) with

\[
(\Gamma_G v)(t) = C\Phi(t) \int_0^t \Phi(t) Bv(t) \, dt
\]

and denote

\[
(\Gamma_G^* v)(t) = B^T \Phi^*(t) \int_0^t \Phi^*(t) C^T v(t) \, dt
\]

The Hankel operator is a mapping from \(L^2[0, \infty) \to L^2[0, \infty)\), i.e. to \(v \in L^2[0, \infty)\) is added by \(\Gamma_G\) a function \(w \in L^2[0, \infty)\) where \(w = (\Gamma_G v)\). For the interpretation of the operator we consider an input \(v\), such that \(v(-t) = u(t)\) for \(t \leq t_0\) and \(v(t) = 0\) for all \(t \geq t_0\), where we from now on consider \(t_0 = 0\).

* See appendix (A.2) for the definition of the Lebesgue-space.
Figure 3.1 \( v(-t) = u(t) \)

This modified input signal gives

\[
(\Gamma_G v)(t) = C\Phi(t) \int_0^t \Phi^{-1}(\tau) Bu(\tau) \, d\tau
\]

such that \( (\Gamma_G v)(t) \) for all \( t \geq 0 \) is equal to the output \( y(t) \) which is the result of the imposed input before \( t=0 \). With other words: \( (\Gamma_G v)(t) \) gives the "future output" for \( t \geq 0 \) as outcome of the "past input" \( v(-t) = u(t) \) for \( t < 0 \).

We define the singular values of the Hankel operator as the eigenvalues \( \sigma_i \) with \( v \) as matching eigenvectors of \( \Gamma_T^T \Gamma_G \), i.e. \( \Gamma_T^T \Gamma_G v = \sigma_i^2 v \). Assume that \( w(t) = \Gamma_G v = C\Phi(t)x_0 \) with

\[
x_0 = \int_0^\infty \Phi(t) Bv(t) \, dt
\]

(3.2.2)

We note that \( x_0 \) is the state on \( t=0 \) as result of the input for \( t \leq 0 \), i.e. the past input. Now

\[
(\Gamma_G^T \Gamma_G v)(t) = B^T \Phi^T(t) \int_0^t \Phi(t) C^T C\Phi(t) \int_0^\xi \Phi(\xi) Bv(\xi) \, d\xi \, dt
\]

\[
= B^T \Phi^T(t) Q x_0
\]

\[
= \sigma_i^2 v(t)
\]

thus

\[
v(t) = \sigma_i^2 B^T \Phi^T(t) Q x_0
\]

Substitution in (3.2.2) yields \( PQ x_0 = \sigma_i^2 x_0 \). Hence, \( x_0 \) is an eigenvector of \( PQ \) with matching eigenvalue \( \sigma_i^2 \).

It is seen that the singular values of the Hankel operator are equal to the Hankel singular values, \( \sigma_i^2 (\Gamma_G) = \lambda_i (PQ) \).

### 3.3 The Induced Hankel-norm

The induced norm of \( \Gamma_G : L^2[0,\infty) \to L^2[0,\infty) \), often called the Hankel-norm, is defined.

**Definition:**

For a stable strictly proper system the Hankel-norm of the transfer function \( G(s) \) is given by the largest Hankel singular value of \( PQ \).

\[
\|G(s)\|_H = \sigma_{\max} (\Gamma_G) = \sqrt{\lambda_{\max} (PQ)}
\]
It is seen for the past input $v \in L^2[0,\infty)$ and future output $w \in L^2[0,\infty)$ that
\[ \| G(s) \| _H^2 = \left( \sigma_{\max}(\Gamma_o) \right)^2 = \max_v \frac{\langle \Gamma_o v, \Gamma_o v \rangle}{\langle v, v \rangle} \]
and thus
\[ \| G(s) \| _H = \sup_{v \in L^2} \left( \frac{\| y \| _{L^2[0,\infty)}}{\| v \| _{L^2[0,\infty)}} \right) \]  
(3.3.1)
what implies that the Hankel-norm is the, so called, $L_2$-gain from past input to future output.

3.4 The optimal anticausal function

To find an optimal approximation the Hankel-norm of the error, $G(s) - \hat{G}(s)$, is minimised, where the transfer function $G(s)$ is approximated by $\hat{G}(s)$. To achieve this we try to make the error system all-pass. The moduli of such an system are constant for the whole frequency range, while only phase changes are taking place. This implies that the magnitude for the error system will be bounded. By means of an anticausal function this error-bound can be minimised. We will first look at the Hankel-norm of the transfer function.

Theorem 3.4.1:
Given a rational stable transfer function $G(s)$, then for any anticausal function $F \in H_{-\infty}$
\[ \| G(j\omega) - F(j\omega) \| _{L^2} \geq \| G(s) \| _H \]
Proof:
By (3.3.1) and Parseval applies
\[ \| G(s) \| _H = \sup_v \frac{\| w(t) \| _{L^2}}{\| v(t) \| _{L^2}} = \sup_v \frac{\| W(j\omega) \| _{L^2}}{\| V(j\omega) \| _{L^2}} \]
Where $W(j\omega)$ and $V(j\omega)$ are the Fourier transformed to $w(t)$ and respectively $v(t)$. Since $G(j\omega)$ is bounded, $y(j\omega) \in L^2$ and thus $y(t) \in L^2$ and $w(t) \in L^2$, $W(j\omega) \in L^2$. This yields for the inner product of $w(t)$

---

* See appendix (C).
** For Hardy-spaces see appendix (A.3).
*** See appendix (B.4) for the theorem of Parseval.
which gives for the Hankel-norm of the transfer function $G(s)$ that

$$
\|G(s)\|_H = \sup_{\omega} \frac{\|W(j\omega)\|_\infty}{\|V(j\omega)\|_\infty} \leq \|G(j\omega) - F(j\omega)\|_{L_2}.
$$

Theorem (3.4.1) can be interpreted as an lower bound of the $L^\infty$-error if the transfer function is approximated by an anticausal function. This function must be determined to reach the lower error-bound.

**Theorem 3.4.2:**

Let $G(s)$ be a rational stable transfer function, then it can be approximated by a function $\hat{H}(s)$ such that

$$
\frac{(G(s) - \hat{H}(s))}{\|G(s)\|_H} = \text{all-pass}
$$

The proof is given in appendix (C) for a system $(A,B,C,D)$ which is reduced to a system $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ of order $k$. The reduced system matrices are thus chosen to make the error system all-pass. For the determination of the reduced system matrices is referred to Glover [1984]. Note that the used balanced realisation is not on principle equal to the one proposed for balanced truncation. A proper choice of $D_e$, the $D$-matrix of the error system, makes $(G(j\omega) - \hat{G}(j\omega))$ all-pass. It is seen that the difference between $G(s)$ and the approximation have constant moduli and thus are bounded.

**Theorem 3.4.3:**

For the reduced model $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ applies:

(a) $\hat{A}$ has no poles on the imaginary axis if $A$ does not neither.

(b) $\hat{A}$ has as many poles in the right respectively left half-plane as $(\Sigma_r \Lambda)$ and $(\Sigma_l \Lambda)$. If the full system is fully controllable and observable then all the states of the reduced system are in the left half-plane. All neither controllable nor observable modes are in the right half-plane.

(c) If either $\Sigma_r \Lambda > 0$, $\Sigma_l \Lambda > 0$ or $\Sigma_r \Lambda < 0$, $\Sigma_l \Lambda < 0$ then the approximation is a minimal realisation, i.e. completely controllable and observable.

The proof is given by Glover [1984].

---

* For simplicity we assume that the multiplicity of all Hankel singular values is equal to one. In practice this will not always be the case. However the derivations made are not fundamental different. For the derivation with multiple Hankel singular values is referred to Glover [1984].
Theorem 3.4.4:
Let \( G(s) \) be a rational stable transfer function, then there exist an anticausal approximation \( F \in H_- \) such that

\[
\| G(j\omega) - F(j\omega) \|_{L_\infty} = \sigma_1(G(s)) = \| G(s) \|_H
\]

Proof:
By theorem (3.4.1) is \( \| G(s) \|_H \leq \| G(j\omega) - F(j\omega) \|_{L_\infty} \) for all \( F \in H_- \). We need to exploit the construction in appendix C to find a minimizing \( F(s) \). Let \( G(s) \) be a stable and rational transfer function with balanced realisation \((A,B,C,D)\) and Hankel singular values \( \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0 \). Also let

\[
\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \sigma_1 \end{bmatrix} \quad \text{met} \quad \Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)
\]

Because \( (\Sigma_1^2 - \sigma_1^2 I) < 0 \) applies that the reduced model has as many poles in the left and right half-plane as \((-\Sigma_1, \Sigma_1^2 I))\) by theorem (3.4.3.b). Thus while all poles of \( G(s) + F(s) = \hat{G} + \hat{C}(sI - \hat{A})^{-1} \hat{B} \) are in the right half-plane is \( F(s) \in H_- \) and \( G(s) = 0 \). By theorem (3.4.2) is now this theorem proven.

This result implies that if we want to approximate the causal transfer function \( G(j\omega) \) with an anticausal function \( F(j\omega) \), the smallest \( L^\infty(-\infty, \infty) \)-error that can be reached is equal to the Hankel-norm of \( G(s) \).

3.5 Solutions to the optimal Hankel-norm approximations

For a given rational stable transfer function \( G(s) \) of order \( n \), there is to determine \( \hat{G}(s) \) of order \( k < n \) such that \( \| G(s) - \hat{G}(s) \|_H \) is minimal. The Hankel-norm gives the \( L^2 \)-gain from past-input to future-output.

Theorem:
Let \( G(s) \) have the Hankel singular values \( \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > \sigma_{k+1} > \sigma_{k+2} > \ldots > \sigma_n > 0 \). Then \( \hat{G}(s) \) is an optimal Hankel-norm approximation if and only if there exist a \( F(s) \in H_- \) such that the error system \( E(s) \triangleq G(s) - \hat{G}(s) - F(s) \) is all-pass, in which case

\[
\| G(s) - \hat{G}(s) \|_H = \sigma_{k+1}
\]

Proof:
For all \( F \in H_- \) is given by theorem (3.4.4):

\[
\| G(j\omega) - \hat{G}(j\omega) - F(j\omega) \|_{L_\infty} = \| G(s) - \hat{G}(s) \|_H
\]

While the Hankel-norm is given by the largest singular value this yields

\[
\| G(s) - \hat{G}(s) \|_H = \sigma_1(G(s)) - \sigma_{k+1}(G(s)) = \sigma_{k+1}(G(s))
\]

by lemma 7.1 of Glover [1984].

Notice that here the characterization of all optimal Hankel-norm approximations is given. If we reduce \( G(j\omega) \) by the sum of a causal and anticausal function \( \hat{G}(s) + F(s) \) then the lower error bound \( (\sigma_{k+1}(G(s))) \) is reached. The choice of the matrix \( D \) for the approximation does not affect the Hankel-norm. However, \( \| G(j\omega) - \hat{G}(j\omega) \|_{L_\infty} \), depends on \( D \).
A specific solution is now given. Let \((A, B, C)\) be a balanced realisation of \(G(s)\) with 
\[ \Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_{n-k+1}). \]
We define \((\hat{A}, \hat{B}, \hat{C}, \hat{D})\) by (C.1.3). Thus 
\[ \hat{G}(s) + F(s) = \hat{D} + \hat{C}(sI - \hat{A})^{-1}\hat{B} \]
with optimal Hankel-norm approximation \(\hat{G}(s) \in H^\infty_+\) and the anticausal function \(F(s) \in H^-_+\). The order of \(F(s)\) is equal to the order of \((G(s) - \hat{G}(s))\) minus one (the multiplication of \(\sigma_1 (G(s) - \hat{G}(s))\)). It is seen that \(\hat{A}\) has \(n+k-1\) poles in the right half-plane and \(k\) poles in the left half-plane.

### 3.6 One-step-at-the-time Hankel-norm approximations

To come to an upper bound of the error in a reduced system we consider the problem of one-step-at-the-time Hankel-norm reduction. By each step the order of the system is decreased by one.

**Theorem 3.6.1:**

Let \(G(s) \in \mathbb{R}^{nxn}\) be a stable rational transfer function. Also let \(\hat{G}(s)\) be an optimal Hankel-norm approximation of order \(n-1\) with the anticausal function \(F(s) = 0\), then 
\[ \frac{(G(s) - \hat{G}(s))}{\sigma_n(G(s))} \text{ is all-pass} \]

**Proof:**

Theorem (3.4.3.b) gives that all poles of \(\hat{A}\) are in the left half-plane. Then \(F(s)\) exist and can be chosen constant or zero. Thus by theorem (3.4.2) is given that \((G(s) - \hat{G}(s))/\sigma_n(G(s))\) is all-pass.

**Theorem:**

Let \(G(s) \in \mathbb{R}^{nxn}\) be a stable rational transfer function with Hankel singular values \(\sigma_1, \sigma_2, \ldots, \sigma_n\), then \(G(s)\) can be reflected by series of all-pass systems \(E_i(s)\); 
\[ G(s) = D_0 + \sigma_1 E_1(s) + \sigma_2 E_2(s) + \ldots + \sigma_n E_n(s) \]

**Proof:**

The procedure used in the proof of theorem (3.6.1) can be employed and repeated until \(\hat{G}(s)\) is of order zero. Hence, if \(\hat{G}_k(s)\) is the optimal Hankel-norm approximation of \(\hat{G}_{k-1}(s)\) then 
\[ (G_k(s) - \hat{G}_k(s))/\sigma_k(G(s)) \text{ is all-pass.} \]

The Hankel singular values of \(\hat{G}_k(s)\) are \(\sigma_1, \sigma_2, \ldots, \sigma_k\). To choose 
\(D_k = \hat{G}_0(s)\), which will be a constant, the theorem is proven.

### 3.7 The error-bound for balanced truncation

As an immediate consequence of section (3.6) we can obtain theoretically error bounds for the balanced truncation reduction method.
Theorem 3.7.1:
Let \((A_{tr}, B_{tr}, C_{tr}, D)\) be a balanced truncated realisation of a stable rational transfer function \(G(s)\) by Moore (2.7.1) and define
\[
G_{b}(s) = C_{b}(sI - A_{b})^{-1}B_{b},
\]
Also let \((\hat{A}, \hat{B}, \hat{C}, \hat{D})\) be an optimal Hankel-norm approximation by (C.1.3) with \(\Sigma_{1} = \Sigma_{2}\) and
\[
G_{b}(s) = \hat{D} + \hat{C}(sI - \hat{A})^{-1}\hat{B}
\]
Then \((G_{b}(s) - G_{b}(s))/\sigma_{n}(G(s))\) is all-pass when
\[
\frac{G_{b}(s) - G_{b}(s)}{\sigma_{n}(G(s))} = D + C(sI - A)^{-1}B
\]
where
\[
\hat{A} = \begin{bmatrix} A_{b} & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} B_{b} \\ \hat{B} \end{bmatrix}, \quad \hat{C} = \begin{bmatrix} C_{b} & -\hat{C} \end{bmatrix}, \quad \hat{D} = -\hat{D}
\]

Proof:
Similarly as in the second part of appendix (C) is proven that there exist a \(\hat{D}\) such that \((G_{b}(s) - G_{b}(s))/\sigma_{n}(G(s))\) is all-pass.

Theorem 3.7.2:
For the in theorem (3.7.1) proposed models is true that:
\[
\| G(j\omega) - G_{b}(j\omega) \|_{L_{\infty}} \leq 2\sigma_{n}(G(s))
\]
and if \(\sigma_{1}, \sigma_{2}, \ldots, \sigma_{n} \leq \sigma_{n}\) then
\[
\| G(s) - G_{b}(s) \|_{L_{\infty}} \leq 2\sigma_{n}(G(s))
\]

Proof:
For the proof of this theorem we consider
\[
\sigma^{-1}(G(s) - G_{b}(s)) = \sigma^{-1}(G(s) - G_{b}(s)) = \sigma^{-1}(G_{b}(s) - G_{b}(s))
\]
The first term on the right-hand side is all-pass by theorem (3.4.2). The second term is all-pass by theorem (3.7.1). Hence the first part of the theorem is proven. Similarly by the fact that all-pass matrices have unity Hankel-norms gives that the second part is proven.

Theorem:
Let \(G(s) \in \mathbb{R}^{nn}\) be a stable rational transfer function with Hankel singular values \(\sigma_{1}, \sigma_{2}, \ldots, \sigma_{n}\), which is reduced to \(\hat{G}(s) \in \mathbb{R}^{nn}\) by balanced truncation, then
\[
\| G(s) - \hat{G}(s) \|_{L_{\infty}} \leq 2(\sigma_{k+1} + \sigma_{k+2} + \ldots + \sigma_{n})
\]
\[
\| G(j\omega) - \hat{G}_{b}(j\omega) \|_{L_{\infty}} \leq 2(\sigma_{k+1} + \sigma_{k+2} + \ldots + \sigma_{n})
\]

Proof:
Similarly with section (3.6) a balanced model, \(\hat{G}(s)\), can be obtained by truncation of the balanced realisation \(\hat{G}_{1}(s)\). Thus by theorem (3.7.2) is seen that \(\| G(s) - \hat{G}_{b}(s) \|_{H_{\infty}, L_{\infty}} = 2\sigma_{i,1}\). This implies that the theorem is proven.
The obtained upper error-bound for model reduction by balanced truncation, gives a strong theoretically support for the employ of this method.

3.8 The $\mathcal{L}_2$-error for optimal Hankel-norm approximations

The lower error-bound for Hankel-norm reduction can be achieved by using the anticausal function $F(s)$. However this can not be done while the approximation would not be stable and minimal. Therefor is seen that the upper error-bound for all Hankel-norm approximations is in theory the same as for balanced truncation.

$$\| G(j\omega) - \hat{G}(j\omega) \|_2 \leq 2 (\sigma_{k+1} + \sigma_{k+2} + \ldots + \sigma_N)$$

(3.8.1)

Nevertheless it is seen by Glover [1984] that there exist an appropriate matrix $D_0$ such that the upper error-bound can be lowered.

$$\| G(j\omega) - \hat{G}(j\omega) - D_0 \|_2 \leq \sigma_{k+1}(G(s)) + \delta$$

with

$$\delta = \| F(j\omega) - D_0 \|_2 \leq \sum_{1 \leq i \leq T} \sigma_i(F(-s))$$

If $\sigma_i(G(s))$ decreases fast, $\delta$ shall be relatively small compared to $\sigma_{k+1}(G(s))$. Notice that this choice of $D_0$ does not effect the Hankel-norm itself.
4 Large-scale linear systems reduction

A large scale system can be approximated by Guyan reduction. We aim to find a lower order model with an equivalent dynamical behaviour. Therefore the coordinates are partitioned in internal and external coordinates, which we use for the static reduction method. It can be expanded with the dynamical part of the internal coordinates. This is done by Graig-Bampton.

4.1 Guyan reduction

Often one is only interested in a small number of lower eigenfrequencies with eigenforms to match. Therefore we divide the complete system into two smaller parts. Consider a system which can be described by \( n \) differential equations in Lagrange-form

\[
M(q(t)) + K(q(t)) = f(t)
\]

where \( q \in \mathbb{C}^n \) with \( M, K \) matching matrices. We can divide the coordinates in external \((q_i)\) and internal \((q)\) degrees of freedom. The external coordinates are connected to other (sub)systems. We assume that the internal coordinates can be written as a linear combination of the external coordinates. Thus

\[
q = \begin{bmatrix} q_i \\ q_o \end{bmatrix} = \begin{bmatrix} T_e \\ 0 \end{bmatrix} q_e = T q_e
\]

Where \( T \) is a transformation matrix which describes the conversion of \( n \) degrees of freedom into \( n_e \) coordinates \((n_e < n)\). Conform we partitioned the differential equations

\[
\begin{bmatrix} M_{ii} & M_{io} \\ M_{oi} & M_{oo} \end{bmatrix} \begin{bmatrix} \dot{q}_i \\ \dot{q}_o \end{bmatrix} + \begin{bmatrix} K_{ii} & K_{io} \\ K_{oi} & K_{oo} \end{bmatrix} \begin{bmatrix} q_i \\ q_o \end{bmatrix} = \begin{bmatrix} f_i \\ f_o \end{bmatrix}
\]

(4.1.1)

The Lagrange equation in external coordinates is described as follows:

\[
M_e q_e + K_e q_e = f_e
\]

with \( M_e = T^T M T, \ K_e = T^T K T \) and \( f_e = T^T f \). The reduction depends on the choose of the transformation matrix \( T \).

**Theorem:**

When all stiffness terms \( K_i q_i \) and \( K_o q_o \) are much larger than the sum of all mass terms \( M_i q_i + M_o q_o \) and the forces \( f_e \), then the transformation matrix \( T \) is given by:

\[
T = -K_i^{-1} K_o
\]

(4.1.2)

**Proof:**

When the above mentioned conditions are justified, the internal part of equation (4.1.1) can be approximate by

\[
K_i q_i + K_o q_o = 0
\]

This implies that when \( K_i \) is regular that \( q_i = -K_i^{-1} K_o q_o = T q_o \).

With this result the reduced matrices and forces are derived:
For the choice of the external degrees of freedom there must be taken some rules into account. 
• To approximate the \( p \) lowest eigenfrequencies adequately, at least three times \( p \) external coordinates must be chosen. This implies that the reduction can only be used effectively for very large scale systems, (degrees of freedom > ±30).
• The external coordinates must be chosen such that they can describe the eigenforms of the eigenfrequencies to determine.

Notice that in contradiction to reduction by balanced truncation and Hankel-norm, the approximation can be made for each frequency band desired. Also note that for this kind of model reduction there are no error bounds available.

### 4.2 Graig-Bampton

The Graig-Bampton method can be seen as an extension on the Guyan reduction method. The complete movement \( q \) is divided into a static movement \( q_s \) and a part \( q_d \) which involves the movement resulted by the internal inertia forces \( (q=q_s+q_d) \). Assumed is that the static reduction is completely determined by the external coordinates as by Guyan.

For the contribution of the "dynamical" part \( q_d \) is used the so called "constraint modes" or "fixed-interface normal modes". These are the columns of the undamped, free vibrating components where all external coordinates are suppressed. That is

\[
M_d q_i + K_d q_i = 0 
\]  
(4.2.1)

We store the eigenvalues \( \omega_k, k=1,2,..,n-p \) of (4.2.1) in a diagonal matrix \( \Omega \) and the matching eigenforms \( u_k, k=1,2,..,n-p \) in \( U \), which can be normalised as \( U^T M U = I \).

\[
\Omega = \text{diag} [\omega_1, \omega_2, .., \omega_{n-p}] \ ; \ U = [u_1, u_2, .., u_{n-p}] 
\]

When \( \xi_k, k=1,..,n-p \) are the internal natural coordinates then the transformation of all coordinates is given by

\[
q = q_s + q_d = \begin{bmatrix} I \\ T_{\text{stat}} \end{bmatrix} q_s + \begin{bmatrix} O \\ U \end{bmatrix} \xi = T q_{\text{red}} 
\]  
(4.2.2)

where \( q_{\text{red}}=[q_s, \xi]^T \) and

\[
T = \begin{bmatrix} I & O \\ T_{\text{stat}} & U \end{bmatrix} 
\]

It is not possible to use all constraint modes. We will only use the \( k \) modes which match with the eigenfrequencies one is interested in. Now the matrices which describes the reduced system are derived:

\[
M_{\text{red}} = \begin{bmatrix} M_d & (M_d - K_u K^{-1}_u M_u) U_k \\ U_k^T (M_d - K_u K^{-1}_u M_u) & I \end{bmatrix} ; \quad K_{\text{red}} = \begin{bmatrix} K_d & 0 \\ 0 & \Omega_k^2 \end{bmatrix} ; \quad f_{\text{red}} = \begin{bmatrix} f_o \\ U_k^T f_i \end{bmatrix} 
\]  
(4.2.3)

where \( M_{\text{red}}, K_{\text{red}} \in \mathbb{C}^{(n-k)(p+k)} \) and \( f_{\text{red}} \in \mathbb{C}^{(n-k)x1} \). The reduced matrices \( M_d, K_u, f_o \) are by Guyan derived.
5 A flywheel-hybrid drive line

A possible system where reduction of the extensive model is required is the flywheel-hybrid drive line. It can be used in a passenger car, where there is next to the engine an additional flywheel to recuperate the energy otherwise spoiled by breaking. A drive-line of this type is developed by Van der Graaf [1987]. A model of this hybrid drive line is complex and contains the main components, being the i.c. engine, the flywheel and the vehicle, and the connecting elements such as clutches, shafts, gear trains and CVT. Reduction of the extensive model towards a more simple model is required, without loss of accuracy in the lower frequency range.

5.1 Dynamical equations

Van der Ven [1988] analyses the hybrid drive line for several operating modes of the vehicle. A torsional model for the "low-speed" hybrid mode is shown in figure (5.1).

![Figure 5.1 Model of "low-speed" hybrid mode.](image)

The equations of motions for the linear model can be derived by using Newton's second law:

\[ J \ddot{\phi} + R \dot{\phi} + K \phi = T_i \]

Where \( \phi \) expresses the rotational angle, \( J \) the mass matrix, \( R \) the damping matrix and \( K \) the stiffness matrix. \( T_i \) represents the external forces. While the system is not attached to the fixed world, the stiffness matrix is singular. If the damping is considered as a second order effect, the torsional motions can be decoupled of a rigid-body mode (Meirovtich [1967]) to make the stiffness matrix non-singular. By eliminating the rigid-body mode the system's degrees-of-freedom is reduced by one.

Obviously is the model nonlinear by the changing CVT-ratio. It is possible to approximate the flywheel system by several methods. We can consider the drive line as two separate linear submodels connected by a nonlinear CVT ratio. Reconnection of reduced submodels obtains the complete model. It is also possible to reduce the complete system in several working points of the CVT ratio. To determine the
dependence of the parameters of the reduced system towards the CVT ratio a nonlinear model is obtained. First the complete system, with constant CVT ratio, will be reduced.

5.2 Constant CVT-ratio

The CVT ratio, $i$, is set to one. We reduce the entire linear system of order $n=18$ to a smaller system of order $k$. The distinct reduction algorithms for this system are considered. Therefore we need to define the in and outputs of the system. The mean engine torque and the inertia torque generated by the CVT ratio are chosen as input signals. The torque generated in the drive shaft is chosen as output signal. For balanced truncation and Hankel-norm approximation is also demanded that the system is stable. While the full system includes the rigid body mode, two eigenvalues are equal to zero when damping is left out of consideration. It is sufficient to eliminate the rigid body mode.

By Moore is said that there exist an adequately approximation if the Froebinius-norm of a system is much larger the of the remaining system. These calculated norms and Hankel singular values for all $k$'s are shown in table (5.1). Also the matching $L^2$-error is given. A suitable approximation can be made for $k=2$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\sigma_k$</th>
<th>$|\Sigma^2_1|_F$</th>
<th>$|\Sigma^2_2|_F$</th>
<th>$2^{\chi} (\sigma_k^2+...+\sigma_1^2)$</th>
<th>$\sigma_{k+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.8594.10^3</td>
<td>1.8594.10^3</td>
<td>1.8577.10^3</td>
<td>3.7466.10^3</td>
<td>1.8577.10^3</td>
</tr>
<tr>
<td>2</td>
<td>1.8587.10^3</td>
<td>2.2108.10^3</td>
<td>5.2283.10^3</td>
<td>2.9112.10^3</td>
<td>4.5870.10^3</td>
</tr>
<tr>
<td>3</td>
<td>4.5870.10^3</td>
<td>2.2108.10^3</td>
<td>4.7061.10^3</td>
<td>1.9938.10^3</td>
<td>4.5844.10^3</td>
</tr>
<tr>
<td>4</td>
<td>4.5844.10^3</td>
<td>2.2108.10^3</td>
<td>2.541.10^3</td>
<td>1.769.10^3</td>
<td>2.2317.10^3</td>
</tr>
<tr>
<td>5</td>
<td>2.2317.10^3</td>
<td>2.2108.10^3</td>
<td>2.2319.10^3</td>
<td>6.3060.10^3</td>
<td>2.2308.10^3</td>
</tr>
<tr>
<td>6</td>
<td>2.2308.10^3</td>
<td>2.2108.10^3</td>
<td>4.7443.10^3</td>
<td>1.8444.10^3</td>
<td>3.9893.10^3</td>
</tr>
<tr>
<td>7</td>
<td>3.9893.10^3</td>
<td>2.2108.10^3</td>
<td>3.9897.10^3</td>
<td>1.0465.10^3</td>
<td>3.9889.10^3</td>
</tr>
<tr>
<td>8</td>
<td>3.9889.10^3</td>
<td>2.2108.10^3</td>
<td>6.7809.10^3</td>
<td>2.4878.10^3</td>
<td>5.7538.10^3</td>
</tr>
<tr>
<td>9</td>
<td>5.7538.10^3</td>
<td>2.2108.10^3</td>
<td>5.6487.10^3</td>
<td>1.3370.10^3</td>
<td>5.6486.10^3</td>
</tr>
<tr>
<td>10</td>
<td>5.6486.10^3</td>
<td>2.2108.10^3</td>
<td>5.3108.10^3</td>
<td>2.0726.10^3</td>
<td>4.5742.10^3</td>
</tr>
<tr>
<td>11</td>
<td>4.5742.10^3</td>
<td>2.2108.10^3</td>
<td>4.3325.10^3</td>
<td>1.1578.10^3</td>
<td>4.3309.10^3</td>
</tr>
<tr>
<td>12</td>
<td>4.3309.10^3</td>
<td>2.2108.10^3</td>
<td>8.3879.10^3</td>
<td>2.9160.10^3</td>
<td>7.1266.10^4</td>
</tr>
<tr>
<td>13</td>
<td>7.1266.10^4</td>
<td>2.2108.10^3</td>
<td>6.9778.10^4</td>
<td>1.4907.10^4</td>
<td>6.9778.10^4</td>
</tr>
<tr>
<td>14</td>
<td>6.9778.10^4</td>
<td>2.2108.10^3</td>
<td>2.9286.10^3</td>
<td>9.5140.10^5</td>
<td>2.8295.10^6</td>
</tr>
<tr>
<td>15</td>
<td>2.8295.10^6</td>
<td>2.2108.10^3</td>
<td>2.9741.10^5</td>
<td>4.7482.10^6</td>
<td>2.3741.10^5</td>
</tr>
<tr>
<td>16</td>
<td>2.3741.10^5</td>
<td>2.2108.10^3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 5.1

The Hankel singular values are plotted in figure (5.2) and it is seen that $\sigma_2=\sigma_2$ and $\sigma$ decreases quite quickly. Hence, we make a second order Hankel-norm approximation. The $L^2$-error is $2(\sigma_{k+1}+...+\sigma_1^2)\leq3.75.10^3$ by equation (3.8.1). It is possible to reduce the error by approximating the anticausal part of the transfer function by calculating the $D$-matrix.

For reduction by Guyan the order of the approximation by trial and error must be determined. As well as the number of external coordinates. We experience that the best approximation for this tool is made by choosing two external coordinates. That is the angle of the flywheel, for its great mass, and the angle of the drive shaft, while it is measured.

The different approximations retrieved by the three above algorithms are shown in a bode plot (5.3). The approximations by balanced truncation and Hankel-norm are almost equal and closely corresponds to the output torque of the full system. However the reduced models diverges from the original model for
higher frequencies. Both algorithms are very effective and yields in a reduced order model and a corresponding $L^\infty$-error.

![Bode plot](image)

Figure 5.3 Bode plot for hybrid-drive line of a) original, b) balanced truncation, c) hankel-norm and d) Guyan reduction.

The approximation obtained by Guyan is insufficient. There seems to be a rotation in the magnitude plot. Besides, the approximation is of 4th-order. Whereas a theoretic error can not be determined for this estimation no direct comparison is justified.

5.3 The nonlinear CVT ratio

It is hard to implement the time-varying CVT-ratio after reduction by balanced truncation or Hankel-norm. A possibility is to reduce the system for several constant CVT ratio and determine the dependence of the different system parameters. For each $i$ the reduction yields in a simple two mass system to represent the hybrid-drive line when the rigid-body mode is included. Notice that the integration of the rigid body mode is not necessary for the required output signals. An other possibility is to consider the drive line as a combination of two separate submodels and the CVT. The submodels represent respectively the drive, i.e. engine and CVT, and the load, i.e. the vehicle. To approximate both subsystems and reconnecting it with the CVT, a nonlinear reduced model can be obtained.
5.4 The drive submodel

The drive submodel represents the engine and CVT and is joined by a spring to the primary CVT pulley (figure 5.4). For reduction is assumed that this spring is attached to the fixed-world. It is an 8th-order model. As input signals the mean engine torque and the torque generated by the angle of the primary CVT pulley are chosen. The torque working on the CVT is chosen as output signal. The Hankel singular values as well as the Forbinius-norms for this submodel are shown in table (5.2).

![Figure 5.4 Drive submodel](attachment:image)

**Table 5.2**

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\sigma_k$</th>
<th>$|\Sigma_1|_F$</th>
<th>$|\Sigma_2|_F$</th>
<th>$2\chi$ ($\sigma_{k+1}+\ldots+\sigma_1$)</th>
<th>$\sigma_{k+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.5227 \times 10^6$</td>
<td>$1.5227 \times 10^6$</td>
<td>$1.5227 \times 10^6$</td>
<td>$3.3515 \times 10^5$</td>
<td>$1.5227 \times 10^6$</td>
</tr>
<tr>
<td>2</td>
<td>$1.5227 \times 10^6$</td>
<td>$1.8109 \times 10^6$</td>
<td>$9.0960 \times 10^5$</td>
<td>$3.0606 \times 10^4$</td>
<td>$7.6488 \times 10^3$</td>
</tr>
<tr>
<td>3</td>
<td>$7.6488 \times 10^3$</td>
<td>$1.8109 \times 10^6$</td>
<td>$7.6488 \times 10^5$</td>
<td>$1.5309 \times 10^4$</td>
<td>$7.6488 \times 10^3$</td>
</tr>
<tr>
<td>4</td>
<td>$7.6488 \times 10^3$</td>
<td>$1.8109 \times 10^6$</td>
<td>$2.6540 \times 10^5$</td>
<td>$1.1290 \times 10^4$</td>
<td>$2.2878 \times 10^3$</td>
</tr>
<tr>
<td>5</td>
<td>$2.2878 \times 10^3$</td>
<td>$1.8109 \times 10^6$</td>
<td>$2.1710 \times 10^5$</td>
<td>$6.7145 \times 10^3$</td>
<td>$2.1648 \times 10^3$</td>
</tr>
<tr>
<td>6</td>
<td>$2.1648 \times 10^3$</td>
<td>$1.8109 \times 10^6$</td>
<td>$7.0919 \times 10^4$</td>
<td>$2.3850 \times 10^3$</td>
<td>$6.0300 \times 10^1$</td>
</tr>
<tr>
<td>7</td>
<td>$6.0300 \times 10^1$</td>
<td>$1.8109 \times 10^6$</td>
<td>$5.8947 \times 10^1$</td>
<td>$1.1789 \times 10^2$</td>
<td>$5.8947 \times 10^1$</td>
</tr>
<tr>
<td>8</td>
<td>$5.8947 \times 10^1$</td>
<td>$1.8109 \times 10^6$</td>
<td>$0$</td>
<td>$0$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

It is seen that only approximations of 4th-order and higher satisfy the requirement for an internally balanced system as described in definition (2.7.1). Even so approximations according to balanced truncation are made for $k=2, 4$ and $6$. These are shown together with the full sub-system in figure (5.5).

![Figure 5.5 Bode plot of drive submodel for a) original, b) $k=6$, c) $k=4$, d) $k=2$ by balanced truncation](attachment:image)

For $k=2$ the lowest eigenfrequency is approached while for $k=4$ also the supreme eigenfrequency is approached. Hence the error of both approximations is equivalent except for frequencies above the supreme resonance frequency. Especially in this case where specific interest is paid to the lower frequencies, it is expected that the model may be reduced to a second order approximation. It is also seen that the Hankel singular values are not founded on the eigenfrequencies and accompanying eigenforms. Notice that the reduction, in section (2), was based on the controllability and observability of states.
However it is not plain why the approach of eigenfrequencies in the figure is not in accordance with the resonance frequency with the largest moduli.

5.5 The load submodel

The load submodel represents the vehicle including the secondary CVT pulley and is joined by the metal V-belt to the primary CVT pulley. For reduction is assumed that this spring, i.e. V-belt, is attached to the fixed-world. It is an 8th-order model. As input signals are chosen the torque generated by the angle of the primary CVT pulley. The torque working on the CVT and the torque generated in the drive shaft are chosen as output signal.

By definition (2.7.1) is at least an approximation of 4th-order required to get an adequately reduction by balanced truncation. Bode plots for \( k=2, 4 \) and 6 are showed in figure (5.7). The plots indeed illustrates that at least an 4th-order approximation is needed.

![Figure 5.6 Load submodel](image)

Notice that in this especially case the lower frequency range is unsatisfactory approximated. Often one in especially interested in this lower frequency range, while for small parameter variation the higher frequency range diverges strongly.

5.6 Nonlinear reduced model

It is possible to reconnect the reduced models with the central CVT. The connection will be actualised by well-defined internal torques. A schema of the full reconnection is given in figure (5.8).
A flywheel hybrid drive line

The mean engine torque together with the variable CVT ratio are defined as input signals. The required output of the system is the torque generated in the drive shaft.

Figure 5.8 Schematic submodel reconnection with nonlinear CVT.

A bode plot is made for the reconnected model with constant CVT ratio. The transfer from the torque generated by the CVT ratio to drive shaft is shown in figure (5.9). It is seen that the reduction does not yield an acceptable approximation of the full model for the lower frequencies. While for frequencies from 10 to $10^3$ the approximation is acceptable. We note that for increasing $i$ the difference for lower frequencies will enlarge.

Figure 5.9 Bode plot for full hybrid-drive line of a) original and b) reconnected submodels of both, the drive and load submodel, 4-th order by balanced truncation.
6 Conclusions and recommendations

6.1 Conclusions

It was seen in section (2) that the controllability grammian $P$ gives a measurement for the controllability of states. Whereas the observability grammian $Q$ gives a measurement for the effect of initial states toward the output. The eigenvalues of the product of both grammians are invariant under state-space transformations and thus give a mapping from input signals to output. By these eigenvalues balanced realisation of a full model can be achieved. Therefore is needed that the system is stable and minimal, i.e. completely controllable and observable. When the transfer function $G(s)$ of order $n$ is reduced to a function $\hat{G}(s)$ of order $k$ the $L^\infty$-error is given by

$$\| G(j\omega) - \hat{G}(j\omega) \|_{L^\infty} \leq \sigma_{\max} \left( \frac{G(j\omega)}{\hat{G}(j\omega)} \right) \leq 2 \sum_{k=1}^{n} \sigma_k \left( G(j\omega) \right) \quad \forall \omega$$

where $\sigma_k$, $i=1,2,...,n$, are the eigenvalues of $PQ$.

The balanced reduction method can be optimized by the Hankel-norm, which represents the $L^2$-gain from "past" input to "future" output. Therefore is attempted to minimalise the Hankel-norm of the difference between the transfer functions of the full model and its approximation. The error system is made all-pass, where the moduli are bounded. To reduce by the sum of a causal and anticausal function the Hankel-norm is minimised and the error made equal to the $(k+1)$-the Hankel singular value of the original transfer function. Notice that the anticausal transfer function in unstable and thus can't be used in the approximation. The $L^\infty$-error is equal to balanced truncation if there is no approximation made for this function. The lower bound can nearly be achieved by an appropriate choice of the matrix $D$ of the approximation.

By Guyan's reduction method the lower eigenfrequencies are approximated. This is to use only external coordinates, which are connected to other (sub)systems. Assumed is that the each of the stiffness terms are much larger than the sum of all mass terms and all forces. A dynamical extension by the internal coordinates to this static reduction can be made by Graig-Bampton.

For both balanced truncation and Hankel-norm approximation is seen that the reduction is made for the moduli of the transfer functions. No phase-matching is made. It is also seen that the necessary stable system maintains stable. No poles are allowed on the imaginary axis, while these will be lost. For both reduction methods some error-bounds are given.

We establish that both algorithms are straightforward to implement and to carry out, while there are no variables to appoint. This brings about some disadvantages. It is seen that the reduction is over all frequencies. Very often one is interested in a specific frequency band. Therefore is desired a variable to fix the reduction towards this frequency band. Note that a small influence can be made by scaling a system with multiple in- and outputs. Also notice that it can not be said out of a bode plot which eigenfrequency will be approximated.
CONCLUSIONS AND RECOMMENDATIONS

By the reduction made for the hybrid-drive line is seen that the Guyan reduction does not satisfy. There is also no theoretic error-bound given.

For an appropriate reconnection of several reduced submodels was always assumed that an adequately approximation of the full system would be obtained. However, when we use this procedure for the hybrid-drive line, this assumption seems not to be justified. For the hybrid-drive line this was not true. The moduli in the lower frequencies are not decently estimated. This is probably the consequence of an improper approximation of the load submodel for the lower frequencies. It could be said out of the bode plot that the load submodel is dominant compared to the drive components.

6.2 Recommendations for further research

It is seen by the conclusions that there are quite a few questions kept unanswered by this report. It even seems that model reduction techniques, as balanced truncation and optimal Hankel-norm approximations, are good-for-nothing. However, this mainly is the result of a not frequency depending balanced realisation. Besides by these reduction algorithms theoretically error bounds are obtained and they are easy to implement.

It is suggested to try to expand the reduction algorithms with some design-variables. By these it should be made possible to influence the frequencies where the transfer-error for is minimised. We assume that it should be made possible to influence the reordering of states and the Hankel singular values in such a manner that this aim is reached. An other possibility is probably to add a function to the full system which filters a specific frequency band.

In relation to the latter suggestion is seen that the states which are relatively well to control and to observe do not match with the eigenfrequencies with the largest magnitude. This seems a contradiction while the Hankel singular values are a measurement for the transfer from input signals to states and output. We recommend to look at a single input and single output system to research the determination of controllable and observable states.

The physical interpretation of an anticausal function is not made comprehensible by this report. It is not obvious why a transfer function can be approximated by such a function. Therefore an investigation about the physical representation of an anticausal function is demanded.

For the hybrid-drive line it is recommendable to reduce the order of the system by approximating the distinct submodels. The reconnection with the "nonlinear" CVT is simple to make. The introduction of the CVT-ratio in the approximation of the full model is hard to make.

This latter recommendation should only be followed up if it is justified that the reconnection of reduced submodels gives an adequately approximation of the full system. For the calculated approximations in this report this was not true. This is probably the consequence of an improper approximation of the load submodel for the lower frequencies. When the reduction algorithms are adjusted with a variable to approximate the lower frequency range, then a correct reconnection could be made. Notice that the reconnection here was made after reduction by balanced truncation.

It is recommendable to examine other reduction methods which give strong error-bounds. It is desirable that these model reduction methods to examine, consists of a straightforward algorithm. It should be easy to implement. An other example of such a method is phase-matching (Tieland [1992]).
References


GLOVER, K., All optimal Hankel-norm approximations of linear multivariable systems and their $L^\infty$-error bounds., Int. J. Control, 1984, Vol. 39, nr. 6, pp 1115.


Nomenclature

$\mathbb{R}^n$, $\mathbb{C}^n$ n-dimensional real and complex euclidean spaces
$\mathbb{C}^{\text{nom}}$ space of $n \times m$ complex matrices
$L^2$, $L^\infty$ Lebesgue-spaces
$\| G(s) \|_p$ Lebesgue-norm
$<\ldots, \cdot >_2$ inner product in $L^2$
$H^\alpha$, $H^\beta$ Hardy-spaces
$\| \Sigma \|_F$ Frobenius-norm of diagonal block matrix $\Sigma \in \mathbb{C}^{n \times n}$
$(A,B,C,D)$ state-space realisation
$A^T$, $A^*$, $A'$ transpose, complex conjugate transpose, inverse of $A \in \mathbb{C}^{n \times n}$
$\lambda_i(A)$ $i$-th eigenvalue of $A \in \mathbb{C}^{n \times n}$
$\lambda_{\text{max}}(A)$ eigenvalue of $A \in \mathbb{C}^{n \times n}$ with maximum modulus
$\text{diag}(A_1, A_2, \ldots, A_m)$ block diagonal matrix with $A_i \in \mathbb{C}^{n_i \times n_i}$
$\det(A)$ determinant of $A \in \mathbb{R}^{n \times n}$
$I_n$ $n \times n$ identity matrix
$T$ transition matrix to the state-space realisation
$G(s)$ transfer function
$\hat{G}(s)$ reduced-order approximation to $G(s)$
$\sigma_H(G(s))$ $H$th Hankel singular value of $G(s)$
$\| G(s) \|_H$ Hankel-norm of $G(s)$
$\Gamma$ Hankel operator of $G(s)$
$H(t, \tau)$ impulse response to the states
$P(t_0, t)$ controllability grammian of $(A,B)$
$Q(t_0, t)$ observability grammian of $(A,C)$
$P_{\text{bal}}, Q_{\text{bal}}$ balanced controllability, respectively observability matrix
$\Sigma$ balanced controllability and observability matrix
$t$ time
$x(t)$ states, $x(t) \in \mathbb{R}^n$
$u(t)$ input signal, $u(t) \in \mathbb{R}^m$
$v(t)$ modified input; $v(-t) = u(t) \forall t < 0$, $v(t) = 0 \forall t \geq 0$
$y(t)$ output signal, $y(t) \in \mathbb{R}^p$
$w(t)$ modified output; $w(t) = y(t) \forall t > 0$, $w(t) = 0 \forall t \leq 0$
$max$ maximum
$\text{supr}$ supremum
Appendix A

A.1 The Forbinius-norm

Definition A.1.1:
Let $\Sigma \in \mathbb{C}^n$ with $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ then the Forbinius-norm is defined as

$$\| \Sigma \|_F = \left( \sum_{i=1}^{n} \sigma_i^2 \right)^{1/2}$$

A.2 The Lebesgue space

Definition A.2.1:
Let $X(s): \mathbb{C} \to \mathbb{C}^n$ then $X(j\omega) \in L^2(-\infty, \infty)$ if and only if

$$\int_{-\infty}^{\infty} X^*(j\omega) X(j\omega) d\omega < \infty$$

and denote

$$\| X(j\omega) \|_{L^2}^2 = \int_{-\infty}^{\infty} X^*(j\omega) X(j\omega) d\omega$$

$$\langle Y(j\omega), X(j\omega) \rangle = \int_{-\infty}^{\infty} Y^*(j\omega) X(j\omega) d\omega$$

Definition A.2.2:
Let $G(s): \mathbb{C} \to \mathbb{C}^{m \times n}$ then $G(j\omega) \in L^\infty(-\infty, \infty)$ if and only if $\exists M < \infty$ such that $\sigma_{\text{max}}(G(j\omega)) \leq M$ $\forall \omega \in \mathbb{R}$ and denote

$$\| G(j\omega) \|_{L^\infty} = \sup \{ \sigma_{\text{max}}(G(j\omega)) \}$$

A.3 Hardy spaces

Definition A.3.1:
Let $X(s): \mathbb{C} \to \mathbb{C}^n$ then $X \in H^\infty_+ (\text{respectively } H^\infty_-)$ if and only if $X$ is analytic in the open right (respectively left) half-plane and $\| X(\sigma + j\omega) \|_{L^\infty} \leq M < \infty \ \forall \sigma \geq 0$ (respectively $\sigma \leq 0$).

An analytic function can be written as a converge Taylor-sequence. The Hardy-space is thus defined for real-rational functions, which have an limited number of poles. Notice that $L^\infty(-\infty, \infty)$ corresponds to the transfer functions bounded on $s=j\omega$. Whereas $H^\infty_+$ corresponds to the causal bounded transfer functions (i.e.
poles in the right half-plane) and $H^\infty_-$ corresponds to the anticausal bounded transfer functions (i.e. poles in the right half-plane).

Definition A.3.2:
Let $G(s): \mathbb{C} \rightarrow \mathbb{C}^\infty$ then $G \in H^\infty_+$ (respectively $H^\infty_-$) if and only if $G$ is analytic in the open right (respectively left) half-plane and bounded in the closed right (respectively left) half-plane.
This definition implies that $H^\infty_+$ and $H^\infty_-$ are subspaces of $L^\infty$.

A.4 Theorem of Parseval

Theorem A.4.1:
Let $w(t) \in L^2(-\infty, \infty)$ and $W(j\omega) \in L^2(-\infty, \infty)$, then by Parseval applies

$$\|w(t)\|_{L^2}^2 = \int_{-\infty}^{\infty} w^*(t) w(t) \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} W^*(j\omega) W(j\omega) \, d\omega = \|W(j\omega)\|_{L^2}^2,$$

where

$$W(j\omega) = \int_{-\infty}^{\infty} w(t) e^{-j\omega t} \, dt$$
Appendix B

B.1 All-pass transferfunctions

A square matrix with all-pass transfer functions $G(s)$ satisfies $G(s)G^*(-s) = I$. Then all Hankel singular values are equal to one. The phase change of an $n$th-order all-pass transfer function is necessary $\pi n$ with constant amplitude. Therefore it is impossible to approximate an $n$th-order all-pass system adequately. A lower order model will have a smaller phase change, so that the error will become large by some frequencies.

Assume that realisation $(A, B, C)$ with $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{n \times m}$, $C \in \mathbb{C}^{m \times n}$ is completely controllable and observable.

If there exist a matrix $D$ such that $G(s)G^*(-s) = \sigma^2 I$ \forall s with $G(s) = D + C(sI-A)^{-1}B$, then there exist $P, Q \in \mathbb{C}^{n \times n}$ such that

\[
\begin{align*}
P &= P^*, & Q &= Q^* \\
AP + PA^* + BB^* &= 0 \\
A^*Q + QA + C^*C &= 0 \\
PQ &= \sigma^2 I
\end{align*}
\]

and the other way around. This implies that when the second statement is satisfied then there exist a matrix $D$ which makes $G(s)$ all-pass and

\[
\begin{align*}
D^*D &= \sigma^2 I \\
D^*C + B^*Q &= 0 \\
DB^* + CP &= 0
\end{align*}
\]

* The proof is given by Glover [1984].
Appendix C

C.1 Proof of theorem 1.4.2

Let \((A, B, C, D)\) with \(A \in \mathbb{C}^{nxn}, B \in \mathbb{C}^{nxm}, C \in \mathbb{C}^{nxn}, D \in \mathbb{C}^{nxm}\) be a system that satisfies

\[AP + PA^T + BB^T = 0 \tag{C.1.1}\]

\[A^TQ + QA + C^TC = 0 \tag{C.1.2}\]

for

\[P = P^T = \text{diag}(\Sigma_1, \sigma)\]
\[Q = Q^T = \text{diag}(\Sigma_2, \sigma)\]

with \(\Sigma_1\) and \(\Sigma_2\) diagonal, \(\sigma \neq 0\) and \((\Sigma_1\Sigma_2 - \sigma^2 I)\) without zeros at the imaginary axis. It is seen that \(P = Q\) is not inevitably for the further derivation. However for balanced truncation \(P\) is made equal to \(Q\) and diagonal, which applies when \(\Sigma_1 = \Sigma_2\) is assumed. Notice that at this point the choice of the Hankel singular value \(\sigma\) is not definite.

Partition \((A, B, C, D)\) conform \(P\) as

\[A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \quad D = D\]

and define

\[
\dot{\hat{A}} = \Lambda^{-1}(\sigma^2 A_{11}^T + \Sigma_2 A_{11} \Sigma_1 - \sigma C_1^T U B_1^T) \\
\dot{\hat{B}} = \Lambda^{-1}(\Sigma_2 B_1 + \sigma C_1^T U) \\
\dot{\hat{C}} = C_1 \Sigma_1 + \sigma U B_1^T \\
\dot{\hat{0}} = D - \sigma U
\]

with \(\Lambda \propto \Sigma_1 \Sigma_2 - \sigma^2 I\) and \(U\) an unitary matrix satisfying \(B_2 = -C_2^T U\) (see section C.2).

Define system \(E(s) \Delta D + C(sI - \hat{A})^{-1} B\) by

\[A_s = \begin{bmatrix} \hat{A} & 0 \\ 0 & \hat{\Lambda} \end{bmatrix}, \quad B_s = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \quad C_s = \begin{bmatrix} C & -\hat{C} \end{bmatrix}, \quad D_s = \begin{bmatrix} D & -\hat{0} \end{bmatrix}\]

for the reduced system \((\hat{A}, \hat{B}, \hat{C}, \hat{D})\) of order \(k\) where \(\hat{A}(s) = \hat{D} + \hat{C}(sI - \hat{A})^{-1} \hat{B}\).

Beside (C.1.1) we assume that applies (see section (C.3))

\[A_{11} + \hat{A}_{11}^* + B_1 \hat{B}^* = 0 \]
\[A_{21} + B_2 \hat{B}^* = 0 \]
\[\hat{A} \Sigma_2 \Sigma_2^{-1} + \Sigma_2 \Sigma_2^{-1} \hat{A} = \hat{B} \hat{B}^* + 0\]

if \(\sigma = 1\) is assumed (wich can be made true by rescaling \(B, C\) and \(\Sigma\)). Such that \((A_s, B_s, C_s)\) satisfies

\*

* The system may not have any zeros at the imaginary axis while these will be lost by this type of model reduction.
Similarly we assume together with (C.1.2) that (see appendix (C.3))
\[ A_1^\dagger (-\Gamma) + (-\Gamma)\hat{A}^* + C_1\hat{C}^* = 0 \]
\[ A_2^\dagger (-\Gamma) + C_2\hat{C}^* = 0 \]
\[ \hat{A}\Sigma_1\Gamma + \Sigma_1\Gamma\hat{A} + \hat{C}^*\hat{C} = 0 \]
and thus
\[ A_0^*Q_e + Q_eA_e + C_e^*C_e = 0 \]

\[ \text{met } \quad Q_e = \begin{bmatrix} \Sigma_2 & 0 \\ -\Gamma & \Sigma_1\Gamma \end{bmatrix} \]  \quad (C.1.5)

It is seen that \( P_eQ_e = \sigma^2 I \). This implies that there exists an \( D_e \) such that \( E(s) \) is all-pass\(^*\). Therefore we determine \( D_e \) such that it satisfies:
\[ D_e^*D_e = \sigma^2 I \]
\[ D_e^*C_e + B_e^*Q_e = 0 \]
\[ D_eB_e^* + C_eP_e = 0 \]

The second equation results from the definition of \( \hat{A}, \hat{C}, D_e \) and \( Q_e \). The third equation is calculated by multiplying for with \( D_e \) and after with \( P_e \).

### C.2 Proof \( U \) is an unitary matrix

By block (2,2) of equation (2.3.1) and (2.3.2) is given
\[ A_{22} + A_{22}^T + B_2B_2^T = 0 \]

and thus \( B_2B_2^T = C_2^TC_2 \).

\( B_2 \) has a singular value decomposition \( B_2 = U_2\Sigma V_1^T \) with \( U_2U_2^T = I, \ V_1V_1^T = I, \ \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \), \( \Sigma_1 > 0 \) and diagonal.

Hence \( C_2^TC_2B_2B_2^T = U_2\Sigma^T U_1^T \), so that \( C_2^T \) has a singular value decomposition
\[ C_2^T = U_1(\Sigma, 0) \begin{bmatrix} V_2^T \\ V_3^T \end{bmatrix} \]

This implies that \( C_2^T = U_1\Sigma V_2^T = XV_1V_2^T \). Let \( U = V_1V_2^T \), then \( UU^T = I \), while \( V_2^TV_2 = I \). Thus \( U \) is an unitary matrix.

\(^*\) See appendix (B) for all-pass systems.
C.3 Derivation Lyapunov equations for an error system

Together with (C.1.2) is assumed
\[ B_2 \dot{\Phi}^T = B_2 (B_1^T \Sigma_2 + U^T C_i) \Lambda^{-1} \]

and while \( B_2 = C_2^T U \) this give
\[ B_2 \dot{\Phi}^T = (B_2 B_1^T \Sigma_2 - C_2^T C_i) \Lambda^{-1} \]
\[ = ((-A_{12}^T \Sigma_2) - A_{12}^T \Sigma_2 + A_{12}^T \Sigma_2 + A_{21}) \Lambda^{-1} \]
\[ = -A_{21} \]

where \( B_2 B_1^T \) and \( C_2^T C_i \) are substituted and block (2,1) of (C.1.1) is used and (C.1.2). This applies
\[ A_{12} + B_2 \dot{\Phi}^T = 0. \]

Furthermore
\[ B_i \dot{Z}^T = (B_i B_i^T \Sigma_2 + B_i U^T C_i) \Lambda^{-1} \]
\[ = ((-A_{11}^T \Sigma_1 - A_{11}^T \Sigma_2 + A_{11}^T \Sigma_2 + B_i U^T C_i) \Lambda^{-1} \]
\[ = -A_{11} - \dot{\Lambda}^T \]

where \( B_i \dot{Z}^T \) is substituted and block (1,1) of (C.1.1) and (C.1.2). Resulting \( A_{11} + \dot{\Lambda}^T + B_i \dot{\Phi}^T = 0. \)

Consider
\[ \Lambda \ddot{\Sigma} + \Sigma \dot{\Lambda} \Lambda = (A_{11}^T + \Sigma_2 A_{11}^T - C_i^T U B_i) \Sigma_2 + (A_{11}^T + \Sigma_1 A_{11}^T - B_i U^T C_i) \]
\[ = -(\Sigma_2 B_i + C_i^T (\dot{B}_i \Sigma_2 + U^T C_i) + (A_{11}^T \Sigma_2 + \Sigma_2 A_{11} + C_i^T C_i) \Sigma_2 \]
\[ + \Sigma_2 (A_{11}^T + \Sigma_1 A_{11}^T + B_i B_i^T) \Sigma_2 \]
\[ = -\Gamma \dot{\Phi}^T \Lambda \]

where block (1,1) of (C.1.1) and (C.1.2) is used. Then applies \( \dot{A} \Sigma_2 \Lambda^{-1} + \Sigma_2 \Lambda^{-1} \dot{\Lambda}^T + \dot{\Phi}^T = 0 \) and is (C.1.4) proved.

Similarly equation (C.1.5) can be proofed. Together with (C.1.1) is given
\[ C_2^T \dot{\Lambda} = C_2^T (A_{11} + UB_i^T) \]
\[ = (-A_{12}^T \Sigma_2 - A_{11}^T + B_2 \dot{B}_1^T) \]
\[ = -A_{12}^T \Lambda \]

This implies \( A_{12}^T (-\Lambda) - C_2^T \dot{\Lambda} = 0. \) Also is given
\[ C_i^T \dot{\Lambda} = C_i^T (C_i \Sigma_1 + C_i^T U B_i) \]
\[ = -A_{11}^T \Sigma_1 + \Sigma_2 A_{11}^T + C_i^T U B_i \]
\[ = -A_{11}^T \Lambda - \dot{\Lambda} \dot{\Lambda} \]

and thus \( A_{11}^T (-\Lambda) + (-\Lambda) \dot{\Lambda} + C_i^T \dot{\Lambda} = 0. \) Finally consider
\[ \dot{A}^T \Sigma_1 \Lambda + \Sigma_1 \Lambda \dot{A} + \dot{C}^T \dot{C} = 0. \] which result in \( \dot{A}^T \Sigma_1 \Lambda + \Sigma_1 \Lambda \dot{A} + \dot{C}^T \dot{C} = 0. \) Now (C.1.5) is proven too.
Appendix D

D.1 A reduction algorithm for Hankel-norm approaches

A specific approximation algorithm for model-reduction by Hankel-norm approach is given. The algorithm is borrowed from Glover [1984]. Assume is that the full system is real, stable and minimal.

1. Let a state-space realisation of the transfer function $G(s)$, $(A,B,C)$ be given for a completely observable and controllable system. The McMillan-degree of the approximation is choose to be $k$.

2. Determine the balanced realisation $G(s)$. Let the Hankel singular values be $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > \sigma_{k+1} \geq \sigma_{k+2} \geq \ldots \geq \sigma_r > 0$ and let the realisation be reordered so that

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

$$ \Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k, \sigma_{k+2}, \ldots, \sigma_n, \sigma_{k+1}) $$

where $(A,B,C)$ is partitioned conformably with $\Sigma$ (i.e. $A_{ij} \in \mathbb{C}^{n \times n}$)

3. Form $\hat{G}(s) = F(s)$ by

$$ U = - (C_2 B_2) / (B_2 B_2^T) $$

$$ \Lambda = (\Sigma^2 - \sigma_{k+1}^2 I) $$

$$ \hat{A} = \Lambda^{-1}(\sigma_{k+1}^2 A_{11} + \Sigma_1 A \Sigma - \sigma_{k+1} C_1^T U B_1^T) $$

$$ \hat{B} = \Lambda^{-1}(\Sigma_1 B_1 + \sigma_{k+1} C_1^T U) $$

$$ \hat{C} = C_1 + \sigma_{k+1} U B_1^T $$

$$ \hat{D} = - \sigma_{k+1} U $$

4. Block diagonalise $\hat{A}$

a. Reduce $\hat{A}$ to a real upper Schur form: Find $V_1$ such that $V_1^T V_1 = I$ and $V_1^T \hat{A} V_1$ are in upper Schur form.

b. Find an orthogonal matrix $V_2$ such that

$$ V_2^T V_1^T \hat{A} V_1 V_2 = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix} $$

where $\text{Re}(\lambda_i(\hat{A}_{11})) < 0$ and $\text{Re}(\lambda_i(\hat{A}_{22})) > 0$.

c. Find $X \in \mathbb{R}^{k \times (n-k)}$ such that

$$ \hat{A}_{11} X - X \hat{A}_{22} + \hat{A}_{12} = 0 $$

by the Bartels-Stewart algorithm.

d. Let

$$ T = V_1 V_2 \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = [T_1 \ T_2] $$
APPENDIX D

\[ S = \begin{bmatrix} 1 - \mathbf{X} & \mathbf{V}_2^T \mathbf{V}_1^T \end{bmatrix} \]

\[ \mathbf{S}_1 \]

\[ \mathbf{S}_2 \]

e. And let

\[ \mathbf{\hat{A}}_1 = \mathbf{S}_1 \mathbf{\hat{B}} \]

\[ \mathbf{\hat{A}}_2 = \mathbf{S}_2 \mathbf{\hat{B}} \]

\[ \mathbf{\hat{C}}_1 = \mathbf{\hat{C}} \mathbf{T}_1 \]

\[ \mathbf{\hat{C}}_2 = \mathbf{\hat{C}} \mathbf{T}_2 \]

5. Determine the \( \mathbf{D} \)-matrix

a. Form a balanced realisation of the system \((-\mathbf{\hat{A}}_2, \mathbf{\hat{B}}_2, \mathbf{\hat{C}}_2, \mathbf{\hat{D}})\) with Hankel singular values \(\mu_1 > \mu_2 > ... > \mu_{n-k+1} > 0\).

b. Let \( \mathbf{Z}, \mathbf{Y} \in \mathbb{R}^{mq(n-k-1)} \) be given by

\[ \mathbf{Z} = \begin{bmatrix} \mathbf{B}_3^T \\
\mathbf{0} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} \mathbf{C}_3 \\
\mathbf{0} \end{bmatrix} \]

and let be \( z_i, y_i \) the \( i \)-th columns of \( \mathbf{Z} \) and \( \mathbf{Y} \) respectively. \((q = p + m)\)

c. For \( k = 1 \) to \( n-k-1 \)

(i) Find Householder transformations (Steward [1973, p.233]) such that

\[ (I - \pi_1^{-1} \mathbf{W}_1 \mathbf{W}_1^T) \mathbf{y}_j = -(\alpha, 0, 0, \ldots, 0)^T \]

\[ (I - \pi_2^{-1} \mathbf{W}_2 \mathbf{W}_2^T) \mathbf{z}_j = -(\beta, 0, 0, \ldots, 0)^T \]

(ii) and let

\[ \mathbf{U} := (I - \pi_1^{-1} \mathbf{W}_1 \mathbf{W}_1^T) \begin{bmatrix} \alpha & 0 & 0 & 0 \\
\beta & 0 & 0 & 0 \\
0 & 0 & I_{p-1} & 0 \\
0 & 0 & 0 & I_{m-1} \\
0 & 0 & 0 & I_{p+m-1} \end{bmatrix} (I - \pi_2^{-1} \mathbf{W}_2 \mathbf{W}_2^T) \]

(iii) if \( i < n-k-1 \) then for \( j = (i + 1) \) to \((n-k-1)\)

\[ \mathbf{y}_j := -(y_j \mu_j + U \mathbf{z}_j \mu_j) (\mu_i^2 - \mu_j^2)^{\frac{1}{2}} \]

\[ \mathbf{z}_j := (z_j \mu_j + U^T \mathbf{y}_j \mu_j) (\mu_i^2 - \mu_j^2)^{\frac{1}{2}} \]

(iv) let

\[ \mathbf{\hat{D}} := \mathbf{\hat{D}} + (-1)^j \mu_j \begin{bmatrix} I_p \\
0 \end{bmatrix} \]

6. \[ \mathbf{\hat{G}}(s) = \mathbf{\hat{D}} + \mathbf{\hat{C}}_1 (sI - \mathbf{\hat{A}}_1)^{-1} \mathbf{\hat{B}}_1 \]

and

\[ \| \mathbf{G}(s) - \mathbf{\hat{G}}(s) \|_\infty \leq \sigma_{k+1} + \mu_1 + \mu_2 + ... + \mu_{n-k-1} \]