Committee members:

prof.dr.ir. M. Steinbuch
dr. S. Weiland
prof.ir. O. H. Bosgra
dr.ir. G. Z. Angelis (coach at Philips)
ir. A. J. den Hamer (coach at TU/e)
Applications of Transfer Function Data

Abstract

For mechatronic systems, accurate frequency response data (FRD) can be obtained at low cost. Controller design techniques exist that can be applied directly to FRD, e.g. classical loop-shaping. However, loop-shaping cannot guarantee that the designed controller is optimal in the sense of closed-loop norm specifications, and the closed-loop pole locations cannot be computed directly. Besides analysis, applying controller synthesis methods on TFD such as pole-placement for example is even more challenging. A different approach is to construct a parametric model of the system. This enables the use of advanced controller synthesis methods such as pole-placement, optimal control, robust control, etc. However, in the modeling phase it is often not clear what aspects of the model are relevant for controller design. It would be advantageous to combine the best of both approaches; to use advanced synthesis methods on the commonly used and well accepted measured FRD.

In previous research a way to connect these two approaches is found [12]. In [12], a method has been developed to compute a data-based equivalent of the transfer function for lightly damped mechanical systems, by extrapolating FRD. The data-based version of the transfer function gives transfer function data (TFD) which gives information on the transfer function for the whole s-plane, while FRD only gives information on the value of the transfer function on the imaginary axis. Model-based results described in terms of a transfer function can be applied data-based using TFD. The potential of TFD has been demonstrated in previous research, therefore this project is started to explore further applications of TFD.

In this thesis, the computation of TFD and three applications of TFD are studied and simulations are performed to demonstrate the proposed theory. Each of these parts will be described shortly.

Regarding the computation of TFD, various methods can be used to compute the TFD for the RHP. Three methods that are discussed are Laplace transformation, convolution and Cauchy contour integral methods. The transfer function of lightly damped mechanical systems is symmetric with respect to the imaginary axis. This makes it possible to compute TFD in the LHP from the RHP.

The first application that is studied, is an extension of the Nyquist stability theorem. It is derived based on two known generalizations of the conventional Nyquist stability theorem; the generalization to MIMO systems and the generalization to alternative contours. It is shown that while the conventional MIMO Nyquist method can give no interpretation of stability margins, the proposed method can give such an interpretation; it shows that the poles in the system have a guaranteed amount of relative damping if it is determined from the generalized Nyquist plot that the system is relative stable. The proposed method can be used data-based, by using the TFD.

The second application of the TFD is the symmetric root locus (SRL). The SRL gives the optimal pole locations that minimize a quadratic performance criterion. The SRL can be computed data-based by searching for points at which the TFD is negative and real. The gain of the resulting root locus is equal to the parameter in the criterion that determines the relative importance of the output compared to the input. The data-based SRL lies close to the value of the SRL computed with a model of the system and is therefore a good approximation. Finally, it is shown that a generalization of the SRL exists, the optimal return difference equation, which is the frequency domain version of the Riccati equation. This equation is used in the next section to an optimal controller that can be computed data-based.

The data-based computation of an optimal controller is the third application of TFD that is studied in this thesis. An optimal regulator, that can be computed data-based because it is described completely in terms of $H(s)$, is found in literature. This optimal controller can be derived from the optimal return difference equation that is the frequency domain version of the Riccati equation. The optimal controller can be computed data-based, by computing a data-based spectral decomposition. A data-based method to choose the optimality parameter $\rho$ is proposed too. Closed-loop pole locations can be selected on the SRL and the corresponding value of $\rho$ can be computed. Two examples of application of this theory are
given.

The objective to find applications of TFD is met, which is demonstrated by the three applications described in this report. It can be expected that other useful applications can be found. Furthermore, it might be possible to compute TFD for other classes of systems, which would create new opportunities to apply this theory.
Symbols and Acronyms

\( j \) The complex number
\( s \) Laplace variable
\( s_i \) Sampled Laplace variable, i.e. points \( i \) in the complex plane
\( \omega_i \) Frequency points
\( H(s) \) System model
\( H \) \( H(-s) \)
\( H(s_i) \) Transfer function data, i.e. complex response at complex frequency point \( s_i \)
\( H(j\omega_i) \) Frequency response data
\( C(s) \) Controller
\( L(s) \) Loop gain \( CH \)
\( N(s) \) Numerator of a transfer function
\( D(s) \) Denominator of a transfer function
\( T(s) \) Return difference
\( z \) Zero
\( p \) Pole
\( Z, P \) Number of poles and zeros
\( m, k, d \) Mass, stiffness, damping
\( A, B, C, D \) State-space matrices
\( P \) Solution of the Riccati equation
\( R, Q \) Weighting of the input and output (optimal control)
\( J \) Cost function
\( \rho \) Optimality parameter
\( \gamma \) spectral decomposition
\( (\cdot)^+ \) RHP spectral factor
\( (\cdot)^- \) LHP spectral factor
\( (\cdot)^* \) Complex conjugate
\( \mathcal{F} \) Fourier transform
\( \mathcal{L} \) Laplace transform
\( C \) Contour in the complex plane
\( \mathbb{N}_1 \) Natural numbers greater than zero \( 1, 2, 3 \ldots \)

\( LHP \) Left half plane, i.e. \( Re(s) < 0 \)
\( RHP \) Right half plane, i.e. \( Re(s) > 0 \)
\( SISO \) Single input single output
\( MIMO \) Multi input multi output
\( SRL \) Symmetric root locus
\( LTR \) Loop transfer recovery
\( LQR \) Linear quadratic regulator
\( LQG \) Linear quaratic Gaussian
\( FRF \) Frequency response function
\( FRD \) Frequency response data
\( TFD \) Transfer function data
Applications of Transfer Function Data
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>i</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>iii</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Background</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Objective</td>
<td>2</td>
</tr>
<tr>
<td>1.4 Outline</td>
<td>2</td>
</tr>
<tr>
<td>2 Transfer Function Data</td>
<td>3</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>3</td>
</tr>
<tr>
<td>2.2 Computing TFD for the RHP</td>
<td>3</td>
</tr>
<tr>
<td>2.2.1 Laplace</td>
<td>3</td>
</tr>
<tr>
<td>2.2.2 Convolution</td>
<td>4</td>
</tr>
<tr>
<td>2.2.3 Contour integral methods</td>
<td>5</td>
</tr>
<tr>
<td>2.2.4 Data-based computation of $H(s_i)$</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Computing the LHP from the RHP</td>
<td>7</td>
</tr>
<tr>
<td>2.4 Relevance of lightly damped mechanical systems</td>
<td>8</td>
</tr>
<tr>
<td>2.5 Applications of contour integral methods</td>
<td>9</td>
</tr>
<tr>
<td>3 Generalized MIMO Nyquist</td>
<td>15</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>15</td>
</tr>
<tr>
<td>3.2 Methods</td>
<td>15</td>
</tr>
<tr>
<td>3.2.1 Properties of transfer functions</td>
<td>15</td>
</tr>
<tr>
<td>3.2.2 Argument principle</td>
<td>16</td>
</tr>
<tr>
<td>3.2.3 Conformal mapping theorem</td>
<td>18</td>
</tr>
<tr>
<td>3.3 Conventional Nyquist stability theorem</td>
<td>20</td>
</tr>
<tr>
<td>3.3.1 Application of the argument principle</td>
<td>20</td>
</tr>
<tr>
<td>3.3.2 Application of the conformal mapping theorem</td>
<td>20</td>
</tr>
<tr>
<td>3.3.3 Example</td>
<td>21</td>
</tr>
<tr>
<td>3.4 Generalized Nyquist stability theorem</td>
<td>21</td>
</tr>
<tr>
<td>3.4.1 Generalized contour</td>
<td>21</td>
</tr>
<tr>
<td>3.4.2 MIMO generalization</td>
<td>22</td>
</tr>
<tr>
<td>3.4.3 Combination: Generalized MIMO Nyquist</td>
<td>24</td>
</tr>
<tr>
<td>3.4.4 Example</td>
<td>25</td>
</tr>
<tr>
<td>3.5 More on the relative stability analysis</td>
<td>27</td>
</tr>
<tr>
<td>3.5.1 Evaluation of multiple contours</td>
<td>27</td>
</tr>
<tr>
<td>3.5.2 Loops in the Nyquist plot</td>
<td>27</td>
</tr>
<tr>
<td>3.6 Data-based application of the generalized MIMO Nyquist</td>
<td>30</td>
</tr>
<tr>
<td>3.6.1 Procedure</td>
<td>30</td>
</tr>
<tr>
<td>3.6.2 Example</td>
<td>30</td>
</tr>
</tbody>
</table>
4 Data-Based Symmetric Root Locus 35
4.1 Introduction 35
4.2 Optimal control 35
4.3 Symmetric Root Locus 36
4.4 Data-based symmetric root locus 38
4.4.1 Introduction 38
4.4.2 Example data-based SRL 38
4.4.3 Computation $H(s_i)$ from $H(j\omega_i)$ 39
4.4.4 Computing the SRL 39
4.4.5 SRL and optimal control 41
4.5 Energy in the system 43
4.6 Optimal return difference 45

5 Data-based optimal control 49
5.1 Introduction 49
5.2 Derivation of the optimal controller 49
5.3 Interpretation of the optimal controller 50
5.4 The optimal controller versus LQG 53
5.5 Data-based computation of the optimal controller 53
5.5.1 Data-based spectral factorization 54
5.5.2 Choosing the optimality parameter $\rho$ 55
5.6 Data-based application of the optimal controller 56
5.6.1 Introduction 56
5.6.2 Example 1: collocated control 56
5.6.3 Example 2: non-collocated control 59
5.7 Introducing roll-off in the controllers 60
5.8 Notes and suggestions for further work 63

6 Conclusions and Recommendations 65
6.1 Conclusions 65
6.2 Recommendations 65

A Proof of the argument principle 67
B Return Difference 69
C Phase reconstruction from gain behavior 71

Bibliography 74
Chapter 1

Introduction

1.1 Motivation

Since mechatronic systems are designed to possess linear, time-invariant behavior they can be characterized by their frequency response behavior very well. For mechatronic systems, large amounts of input-output data can be obtained at low cost. Therefore, the majority of the controllers for this class of systems is designed using data-based controller synthesis methods in industrial practice. In this approach, frequency response data (FRD) is used to tune PID controllers to realize the desired closed-loop behavior of a system. But the methods for controller design are limited to loop-shaping which gives only limited insight in properties as pole locations and damping of the system, and cannot give guarantees about the optimality of the resulting controller in the sense of closed-loop norm specifications.

Many model-based controller synthesis methods are available too. Pole-placement, optimal control, $H_\infty$ robust control and the root locus design method are a few examples of such techniques. These have in common that an accurate parametric model of the system must be available to be able to use them in practice. Obtaining a low-order accurate model can be a laborious task for complex systems. This is due to the fact that in the modeling phase of the design process it is not yet clear which aspects of the model are relevant for the controller design. This can unnecessarily complicate the controller design and implementation because a large-order model will have to be used.

Apparently both data-based and model-based controller design methods have their advantages and disadvantages. This is summarized in Figure 1.1. The objective of this thesis is to combine best of both; to use the accurate data for the system description and use the powerful model-based design techniques.

![Diagram showing tradeoffs of model-based and data-based control design]

Figure 1.1: Tradeoffs of model-based and data-based controller design approaches.
1.2 Background

This graduation project is part of a research program at Philips Applied Technologies that aims to extend manual loop shaping by bridging the gap between data-based and model-based controller design methods. Prior to this study, work has been done to extend manual loop-shaping for single-input-single-output (SISO) systems to multi-input-multi-output (MIMO) systems [5] [22]. Progress has also been made by enabling closed-loop pole locations to be a part of the controller design [12]. This can be used to tune for vibration reduction, settling-time or other objectives closely related to pole-placement. In this research a method is developed to extrapolate FRD to so-called transfer function data (TFD). TFD provides information of a system for the whole s-plane instead of only the imaginary axis, without the requirement of a parametric model. The extension of FRD to TFD is an essential step that is necessary to enable the use of model-based techniques.

1.3 Objective

TFD can be seen as a data-based version of the transfer function. Its potential is not limited to the applications described in [12]. The availability of TFD provides a way to evaluate model-based controller synthesis and analysis criteria in a data-based way. Therefore this graduation project was initiated to investigate applications of TFD that enable the use of model-based controller design techniques. The goal of this thesis is

Explore applications of transfer function data, that aim at the data-based use of model-based controller design and analysis methods.

The main focus of this thesis are lightly damped mechanical systems. The reason for this is that, besides their practical relevance, at the moment TFD can only be computed for lightly damped mechanical systems. However, it might be possible to develop new methods to compute TFD for other classes of systems.

1.4 Outline

The computation of TFD is first analyzed to get insight in its potential and limits. Next, a search for model-based design methods described in terms of transfer functions was performed in literature. The three most promising results that were found, are discussed in this thesis. In Chapter 2 the derivation and computation of TFD is discussed. TFD is computed from FRD by means of a Cauchy integral. TFD is the basis needed to apply the model-based controller design methods. Next, three applications of TFD are treated. Firstly, a generalization of the Nyquist criterion is discussed in Chapter 3. The data-based application of this generalization is only possible by using TFD. Secondly, the data-based application of the symmetric root locus is derived in Chapter 4. The symmetric root locus provides an analysis method which can be used to analyze the closed-loop pole locations of a system. These closed-loop pole locations are optimal for a quadratic optimality criterion. Thirdly, a data-based method to compute an optimal controller is discussed in Chapter 5. In literature a frequency domain solution for the optimal control problem was found, described in terms of the transfer function. The derivation and implementation of this synthesis method is discussed in this chapter. After these applications conclusions are drawn and recommendations for further research are given.
Chapter 2

Transfer Function Data

2.1 Introduction

The frequency response function (FRF) gives information about the response $H(j\omega)$ at the frequencies $\omega$, i.e. on the imaginary axis in the complex plane. However, many controller synthesis and analysis results found in literature are formulated in terms of $H(s)$ instead of $H(j\omega)$. An example of such a result is the symmetric root locus equation

$$1 + \rho H(-s)H(s) = 0 \quad s \in \mathbb{C}. \quad (2.1)$$

The solutions $s$ of this equation are the closed-loop pole locations that are optimal for a certain quadratic criterion with optimality parameter $\rho$. The symmetric root locus is the topic of Chapter 4, it is mentioned here to show the advantage of having information about $H(s)$ compared to $H(j\omega)$.

Without information of the transfer function at each point in the complex plane, this equation cannot be used data-based. The optimal pole locations will probably not lie on the imaginary axis, so FRD is not sufficient to compute the solutions. A parametric model could be used to solve this equation if the model is available. But a different approach is used in this thesis, namely to compute $H(s)$ data-based, which has the advantage that a parametric model of the system is not required. It has been shown in [12] that this is possible for lightly damped mechanical systems. To emphasize the difference between the model-based transfer function and TFD, the model-based transfer function will be denoted $H(s)$ and TFD is denoted $H(s_i)$ consistently. The subscript $i \in \mathbb{N}_1$ indicates that $s_i$ is a sampled complex frequency, which is different from the Laplace variable $s$, but satisfies $s_i \in s$.

This chapter treats the computation of TFD for lightly damped mechanical systems, which consists of two main steps. The first step is to compute $H(s_i)$ in the right half $s$-plane (RHP), and the second step is to compute $H(s_i)$ in the left half $s$-plane (LHP) from $H(s_i)$ in the RHP.

2.2 Computing TFD for the RHP

Lightly damped mechanical systems, which are the focus of this research, have their poles and zeros in the LHP. Therefore the transfer function of the system is analytical in the RHP, which enables the computation of $H(s)$ as will become clear in this section. Various methods are available to compute the transfer function $H(s)$ in the RHP [12]. Laplace transformation, convolution and contour integral methods are discussed in this section.

These methods can be used to compute $H(s_i)$ data-based by evaluating $H(s)$ for a grid of points $s_i$, which is discussed at the end of this section. Computing $H(s_i)$ in the RHP does not provide additional information on the system. It is only an extrapolation of the FRD of the system. However, this extrapolation is very useful in this case because it transforms the data to the format that is needed, as described in the previous section.

2.2.1 Laplace

The first method to compute $H(s)$ that is described is to use the Laplace transformation. This method uses time-domain data to compute $H(s)$. Similarly to the use of the Fourier transform to obtain the
FRF $H(j\omega)$ of the system, the Laplace transform can be used to obtain the transfer function $H(s)$. A time-domain signal $x(t)$ can be transformed to the s-domain using

$$X(s) = \int_0^\infty e^{-st}x(t)dt.$$ \hspace{1cm} (2.2)

This can also be seen as the decomposition of the signal $x(t)$ into Laplace series

$$x(t) = \sum_{i=0}^{N-1} C_i e^{-st},$$ \hspace{1cm} (2.3)

where $C_i$ are the basis functions

$$C_i = \frac{1}{e^{-s_i t}, e^{-s_i t}} \int_0^T x(t)e^{-s_i t}.$$ \hspace{1cm} (2.4)

So $x(t)$ is projected on a finite set of sinusoidal basis functions described by damped exponentials. This method is also known as Prony’s Method [10],[30].

Convergence of the Laplace integral is not guaranteed for all $s$ due to poles of the system under consideration. Suppose a dynamical system $H(s)$ has an impulse response $h(t)$, described in terms of its poles $p_i$ and weightings $\alpha_i$,

$$h(t) = \sum_{i=1}^N \alpha_i e^{p_i t}.$$ \hspace{1cm} (2.5)

Taking the Laplace transform of the impulse response yields

$$H(s) = \int_0^\infty \sum_{i=1}^N e^{(p_i - s)t} dt.$$ \hspace{1cm} (2.6)

which does not converge if $Re(s) < Re(p_i)$, thus the region of convergence is $Re(s) > Re(p_i)$. This shows that it is only possible to compute the Laplace transform for an analytic region [16]. For lightly damped mechanical systems, which are stable, the Laplace transform can be computed for the whole RHP, but not for the LHP because it contains the poles and zeros of the system.

The transfer function is obtained by transforming the input signal $u(t)$ and output signal $y(t)$ to $Y(s)$ and $U(s)$ and computing their quotient

$$H(s) = \frac{Y(s)}{U(s)}.$$ \hspace{1cm} (2.7)

Alternatively, if the impulse response $h(t)$ of a system is available, the transfer function can be computed from

$$H(s) = \int_0^\infty h(t)e^{-st}dt.$$ \hspace{1cm} (2.8)

### 2.2.2 Convolution

A second method to compute $H(s)$ uses FRD instead of time domain data to compute the Laplace transform. Instead of computing the Laplace transform of the time signals, the Laplace transform can be computed from a convolution of two frequency domain functions. Consider again the Laplace transform $H(s)$ of $h(t)$, but now $s = \sigma + j\omega$ is substituted, yielding

$$H(s) = \int_0^\infty h(t)e^{-\sigma t}e^{-j\omega t}dt.$$ \hspace{1cm} (2.9)

$$= \mathcal{F}[h(t)e^{-\sigma t}] \mathcal{F}[e^{-\omega t}],$$ \hspace{1cm} (2.10)

where $\mathcal{F}$ denotes the Fourier transform. Thus, the Laplace transform can also be seen as the convolution (in the frequency domain) of the Fourier transforms of the functions $h(t)$ and $e^{-\sigma t}$. The Fourier transform
of the time function \( h(t) \) is of course the FRF of the system. The Fourier transform of the second function is

\[
\mathcal{F}[e^{-\sigma t}] = \frac{1}{\sigma + s},
\]

(2.12)

which enables the computation of \( H(s) \) by evaluation of the following convolution in the frequency domain

\[
H(s) = \int_{-\infty}^{\infty} H(j\omega - \tau) \frac{1}{\sigma + j\tau} \, d\tau,
\]

(2.13)

where \( (\quad)^* \) denotes the complex conjugate. The convolution effectively averages \( H(j\omega) \) with its surrounding points to compute \( H(s) \). The higher the value of \( \sigma \), the more the FRF is averaged, so it becomes smoother. This can be understood intuitively. Imagine the magnitude of the transfer function as a landscape with mountain peaks, caused by the poles. Since there are no poles in the RHP for which \( H(s) \) is computed, the farther in the RHP, the smoother the landscape becomes. This is visualized beautifully by the 3D Bode diagram of Figure 2.1.

A 3D Bode diagram is a generalization of the conventional Bode diagram in which an extra axis is added to cope with complex frequencies \( s = \sigma + j\omega \). It can be used to visualize \( H(s) \) for all \( s \) in the complex plane, while the conventional Bode diagram can only visualize one cross-section of this diagram, i.e. \( s = j\omega \) [44].

### 2.2.3 Contour integral methods

The third and last method that is discussed uses contour integral methods to compute \( H(s) \). A transfer function is a complex function or mapping \( \mathbb{C} \to \mathbb{C} \) that maps the \( s \)-plane to the \( H(s) \)-plane. Thus complex function theory can be applied to the transfer function. An important theorem in complex function theory states the following.

**Theorem 2.1. (Cauchy integral theorem)** Let \( f(z) \) be an analytical function in a region \( G \) and \( C \) a Jordan curve in \( G \). Then

\[
\int_C f(z) \, dz = 0
\]

(2.14)

**Proof.** For the proof, see [21] or [32].

Another well known result from complex function theory can be derived from this theorem. The value of a function \( f(a) \) in the point \( z = a \) can be computed with a contour integral. This is known as the Cauchy integral formula, see the following theorem.
Applications of Transfer Function Data

Figure 2.2: Computation of $H(s)$ from $H(j\omega)$ by evaluation of a contour integral.

**Theorem 2.2.** (Cauchy integral formula) Let $f(z)$ be an analytical function in a region $G$, $C$ a simple closed (Jordan) curve in $G$ and let $a$ denote a point in the $C$. Then

$$
\frac{1}{2\pi j} \int_C \frac{f(z)}{z-a} \, dz = f(a)
$$

(2.15)

**Proof.** For the proof, see [21] or [32].

The Cauchy integral formula shows a remarkable property of analytical functions, the value of a function $f(a)$ at a point $z = a$ within the contour is fully determined by the value of the function on the contour that encloses this point, provided that the function is analytic in this region.

This can be applied to obtain the value of a transfer function $H(s)$ in an analytic region. As mentioned, no singularities may lie in the contour used to compute $H(s)$. Thus if the RHP is to be computed, the system must be stable since no poles may lie in the RHP. Applying the Cauchy integral formula yields

$$
H(s) = \frac{1}{2\pi j} \int_C \frac{H(s_c)}{(s-s_c)} \, ds_c,
$$

(2.16)

where $s_c$ are all points $s$ on the contour, thus

$$
s_c = \{ s \mid s \in C \}.
$$

(2.17)

Equation 2.16 states that $H(s)$ can be computed from the values of the function on a contour that encloses $s$. When FRD of the system is available, a suitable choice for the contour is the D-contour in the complex plane, see Figure 2.2. Assume that the response at infinite frequency is zero, which is a valid assumption for systems with a relative degree of two or higher. Then the semi-circle at infinity does not contribute to the contour integral, and thus the value of the transfer function is known for all points on the contour, enabling the computation of $H(s)$ for all $s$ in the RHP. In this case the contour integral changes into an integral over all frequency points.

$$
H(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega) \frac{Re(s)}{(j\omega - s)} \, d\omega.
$$

(2.18)

Thus $H(s)$ can be computed from the FRF $H(j\omega)$. There also exist alternative contour integral methods [32],[14] that can be derived from the Cauchy integral. Two examples of alternative contour integrals are the Schwarz

$$
H(s) = \frac{1}{\pi} \int_{-\infty}^{\infty} H(j\omega) \frac{Re(s)}{Re(s)^2 + (Im(s) - \omega)^2} \, d\omega
$$

(2.19)

and Poisson

$$
H(s) = \frac{1}{\pi} \int_{-\infty}^{\infty} Re(H(j\omega)) \frac{1}{s - j\omega} \, d\omega
$$

(2.20)

contour integrals.
2.2.4 Data-based computation of \( H(s_i) \)

Three methods to compute \( H(s) \) have been discussed. However, only the Cauchy contour integral is used during this research to compute \( H(s_i) \) from \( H(j\sigma) \). The reason for this is that good results of computing \( H(s_i) \) were obtained with the Cauchy methods during previous research. Furthermore, it will become clear in Section 2.5 that the Cauchy integral can also be used for other useful applications. In order to evaluate the Cauchy integral of Equation 2.18 numerically it will be approximated by the summation

\[
H(s_i) = \frac{1}{2\pi} \sum_k \frac{H(j\omega_k)}{(j\omega_k - s_i)} \Delta\omega,
\]

provided that the function \( H(j\omega) \) is sufficiently smooth. A system that has poles on the imaginary axis violates this constraint, because it is certainly not smooth. A sufficient condition on \( H(j\omega) \) is that the damping of the poles is much larger than the spacing of the frequency grid \( \Delta\omega \). More information on this smoothness condition can be found in [13].

2.3 Computing the LHP from the RHP

The previous section shows that it is possible to compute \( H(s) \) and therefore also \( H(s_i) \) in the RHP for stable systems. But the LHP cannot be computed with the methods discussed previously since it contains the poles and zeros of the system which makes the LHP a non-analytic region. Therefore another method is needed to compute the transfer function in the LHP. A method to compute \( H(s) \) for the LHP is given in this section, but this technique can also be applied to TFD \( H(s_i) \).

For undamped resonant systems, symmetry in the s-plane can be utilized to compute \( H(s) \) for the LHP. This is stated in the following theorem.

**Theorem 2.3.** For undamped resonant systems, it holds that

\[
H(s) = H(-s).
\]  

**Proof.** The transfer function of a resonant system in modal representation is [26], [29], [19]

\[
H(s) = \sum_n R_n s^2 + 2\zeta_n \omega_n s + \omega_n^2
\]

which is a summation of \( N \) modes \( n \), which have residue \( R_n \), damping \( \zeta_n \) and natural frequency \( \omega_n \). For zero damping \( \zeta_n = 0 \) Equation 2.23 becomes

\[
H(s) = \sum_n R_n s^2 + \omega_n^2
\]

which is a summation of symmetric functions since \( s^2 = (-s)^2 \).

Theorem 2.3 implies that \( H(s) \) in the LHP can be computed from \( H(s) \) in the RHP. A graphical interpretation of Equation 2.22 is also possible, see Figure 2.3. In the figure, the pole-zero diagram of an undamped resonant system \( H(s) \) is shown, which has its poles and zeros on the imaginary axis. Points \( A \) and \( B \) satisfy \( B = -A \) and therefore \( H(A) = H(B) \), satisfying Equation 2.22. This can be seen as follows. A transfer function can be written in terms of its poles \( p \) and zeros \( z \)

\[
H(s) = \frac{\prod_n(s - z_n)}{\prod_m(s - p_m)}.
\]  

The amplitude of \( H(s) \) can be computed from a product of the distances from \( s \) to the pole and zero locations

\[
|H(s)| = \frac{\prod_u|s - z_u|}{\prod_v|s - p_v|},
\]

while its phase is

\[
\angle H(s) = \sum_u \arg(s - z_u) - \sum_v \arg(s - p_v).
\]
Applications of Transfer Function Data

Figure 2.3: Graphical interpretation of the relation $H(s) = H(-s)$ [12].

which is a summation of all angles from the point $s$ to the pole and zero locations. From the figure it is clear that the distance from the points $A$ and $B$ to the poles and zeros is the same for $A$ as for $B$. Furthermore, the angles between $A$ and $B$ and the poles and zeros are also the same for $A$ as for $B$. This shows that $H(A)$ and $H(B)$ have the same amplitude and phase and are therefore equal.

If damping is present in the system, the poles in Figure 2.3 will shift into the LHP. If all poles shift with the same amount, say $\sigma$, a similar symmetry condition as Equation 2.22 can be found. Instead of symmetry with respect to the imaginary axis, the system is now symmetric in the line $Re(s) = -\sigma$, yielding the new symmetry condition

$$H(s) = H(-s + 2\sigma).$$

If the poles do not have equal amounts of damping, the symmetry condition does not hold anymore. But if there is a limited amount of damping, only small deviations will occur. Consider again Figure 2.3. Shifting the poles a small distance into the LHP will result in small differences in the amplitude and phase of $H(A)$ and $H(B)$. The amount of damping that is allowed such that Equation 2.22 still gives a good approximation depends on the value of $s$. $s$ should lie sufficiently far from the imaginary axis relative to the distance of the poles $p_n$ to imaginary axis, that is

$$|Re(s)| >> |Re(p_n)|.$$  \hspace{1cm} (2.29)

Then Equation 2.22 can be used to compute a good approximation of $H(s)$ in the LHP from information about $H(s)$ in the RHP.

The symmetry condition described in this section can also be used data-based to obtain TFD $H(s_i)$ for the LHP, from $H(s_i)$ computed for the RHP. This enables the computation of $H(s_i)$ for a grid of points $s_i$ in the whole s-plane.

2.4 Relevance of lightly damped mechanical systems

As mentioned before, the approach is only possible for undamped or lightly damped resonant systems. However, there are many mechanical systems that have this property. Modern design methods recommend the use of light and stiff constructions because they result in high resonance frequencies and therefore allow accurate motion and positioning [23]. But because of the use of undamped materials such as steel and ceramics these constructions typically have lightly damped resonances. Therefore the proposed method can potentially be applied in the (high-tech) industry in the future. In this early stadium of the research
on this topic, the theory is applied to an academic example system; a fourth order mass-spring-damper system, see Figure 2.4. The system can be described by a set of differential equations

\[
\begin{align*}
    m_1 \ddot{x}_1 &= F_1 - k_1 x_1 - k_2 (x_1 - x_2) - d_1 \dot{x}_1 - d_2 (\dot{x}_1 - \dot{x}_2) \\
    m_2 \ddot{x}_2 &= F_2 + k_2 (x_1 - x_2) + d_2 (\dot{x}_1 - \dot{x}_2)
\end{align*}
\]  

(2.30) (2.31)

which can be put into state-space format

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

(2.32) (2.33)

with

\[
A = \begin{bmatrix}
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1 \\
    -\frac{k_1}{m_1} & \frac{k_2}{m_2} & -\frac{d_1}{m_1} & \frac{d_2}{m_1} \\
    \frac{k_1}{m_2} & -\frac{k_2}{m_2} & -\frac{d_1}{m_2} & \frac{d_2}{m_2}
\end{bmatrix}
\]  

(2.34)

\[
B = \begin{bmatrix}
    0 & 0 \\
    0 & 0 \\
    1 & 0 \\
    0 & 1
\end{bmatrix}
\]  

(2.35)

\[
C = \begin{bmatrix}
    1 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0
\end{bmatrix}
\]  

(2.36)

If \( d \approx 0 \), the poles and zeros of this system are lightly damped, and will lie close to the imaginary axis. Then the symmetry condition is approximately satisfied and TFD can be computed for the example system. A 3D Bode diagram of the magnitude of \( H(s) \), computed with the proposed method, is shown in Figure 2.5. Notice the symmetry with respect to the imaginary axis. The two-mass system will be used in the sequel to explore the derived synthesis methods that are found for TFD.

### 2.5 Applications of contour integral methods

As described in the previous sections, the Cauchy contour integral method can be used to compute \( H(s) \) from FRD of a lightly damped mechanical system. Other interesting applications of the Cauchy contour integral are described in this section.

**Finding poles**

Contour integrals can be used to find poles of a system. The Cauchy integral in Theorem 2.1 states that the contour integral for a function \( f(z) \) that is analytical in the contour is zero. A contour of a function that contains one pole (at \( z = a \)) in the contour, will give the residue of this pole defined by

\[
\text{Res}_{z=a} f(z) := \frac{1}{2\pi j} \int_C f(z) dz.
\]  

(2.37)

This residue is of course related to the term residue in the modal description of a transfer function. Consider a transfer function

\[
H(s) = \frac{2s + 2}{s^2 + 2s + 2} = \frac{1}{s + 1 - j} + \frac{1}{s + 1 + j},
\]  

(2.38)
where the last part is the modal format of the transfer function. Thus there are two poles at \( s = -1 \pm j \) which each have residue 1. The contour integral of curve \( C_1 \) around the pole at \( s = -1 + j \) is computed numerically, see Figure 2.6. A second contour integral of another curve \( C_2 \) that does not encircle the pole is computed too. The results of the integrals are

\[
I_{C1} = 1.0000 - 0.0000i \quad I_{C2} = 1.0481e-016 + 7.4163e-017i
\]

This demonstrates that a contour integral that does not contain a pole gives zero, while a pole that does contain a pole gives the residue of the pole that is encircled. When there is more than one pole in the contour, the integral is equal to the summation of the residues [21]. Therefore caution is necessary when applying this in practice, because the theorem does not give information on the number of poles in the contour.

### Modal Identification

The circle fit method is a simple technique to derive a parametric model of a resonant system. It estimates pole locations and residues to compute the parametric model. But the estimation of the residues is often inaccurate due to interaction between various modes of the system. Using TFD and Cauchy integrals, the estimation of residues in this method can be improved as will become clear in this section.

The circle fit method computes a least squares fit of the resonance circles of the response of a system. Figure 2.7 shows a circle fit (-) of one resonance of the system (···). The modes of the system cannot have much interaction or much damping, else the resonance circles of the different modes cannot be distinguished. The modal parameters, pole location \( p_k = \mu_k + j\nu_k \) and residue \( R_k \), can then be computed for each mode \( k \) of the system. The imaginary part of the pole can be computed from the phase change as the response traverses the resonance circle. The phase change is maximal at the pole location, i.e.

\[
\nu_k = \max_\omega \frac{\partial \angle H}{\partial \omega}.
\]

(2.39)

The real part of the pole, which is related to the damping of the pole, can be computed from two surrounding points on the circle

\[
\mu_k = \frac{\omega_1 - \omega_2}{\tan(\phi_1/2) - \tan(\phi_2/2)}.
\]

(2.40)

Computing the residue can be done from the radius \( r_k \) of the circle

\[
R_k = 2r_k |\mu_k|.
\]

(2.41)

Interaction with other modes can influence the radius of this circle which results in poor estimation of the residues.
Figure 2.6: Finding poles of the system by computing a contour integral.

Figure 2.7: Computing modal parameters from the circle fit method.
Applications of Transfer Function Data

Figure 2.8: Residues can be computed from Cauchy integrals on circles (---) around each pole ×.

There are many alternatives for this method [29]. Here a method is proposed that uses a contour integral to compute the residue. Define a small contour in the complex plane around each pole. Evaluating the Cauchy integral for these contours gives the residue for the poles, see Equation 2.37. This requires only limited information of the pole location, it is only important that the contour that is defined contains (only) the pole of which the residue has to be computed. The advantage of this method is that the contour integral will give the correct residue because other modes do not influence the computation. The value of the transfer functions on these contours can be computed data-based, using Equation 2.21. An example of the application of this improvement of the circle fit method is shown in Figures 2.8 and 2.9 for a resonant system with a number of modes. Figure 2.8 shows the contours that are used to compute the residues. In Figure 2.9 the system (black) is fitted using the conventional circle fit method (red) and the circle fit method with improved residue computation (blue). The original system is hardly visible, because the fit made with the contour integral approach is very good. Both methods use the same algorithm to compute the pole locations, and therefore they are estimated equally well. But the zero locations are estimated much better with the contour integral approach, because the residues are estimated more accurately. Notice that the phase estimation is also almost perfect for the alternative method.

The circle fit method fails in this case because the fourth and fifth mode are close to each other. The alternative method is apparently less sensitive for closely spaced modes. When the contour integral around a pole can be chosen such that it only encircles the pole of which the residue has to be computed, the method will provide accurate results. This depends of course on the density of the frequency grid and the spacing of the modes.

This method can also be applied to MIMO systems. The pole locations will be the same for each transfer function in the MIMO system. The residues, however, will depend on the inputs and outputs. Using the proposed technique, the residues for each transfer function in the MIMO system can be computed. The result is an accurate model that has a structure that corresponds to the physics of the real system.

Noise filtering

The contour integral can also be used to filter a measured FRF to reduce noise. The conventional method to reduce noise is to perform more measurements and average the results. A different approach is proposed in this section. As explained in Section 2.2.2 the contour integral averages the points $H(j\omega)$ on the contour to compute $H(s)$. This property can also be used to filter an FRF by choosing $H(s_i) = H(j\omega_i)$. Caution must be taken when evaluating this contour, because now the point that is evaluated lies on the contour itself. Recall the application of the Cauchy integral formula, 2.18

$$H(s_i) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{H(j\omega)}{(j\omega - s_i)} d\omega,$$

(2.42)
Figure 2.9: Modal identification for a model (black) using the circle fit (red) and circle fit with Cauchy (blue) methods.

Because $s_i$ now lies on the contour, the denominator $(j\omega - s_i)$ will go to zero while traversing the contour. Approximating the contour integral with a summation will now be non-convergent because of this. A solution for this problem is to make small indentations in the curve at the point that is evaluated, see Figure 2.10. Using this contour, the integral of Equation 2.42 can be divided into three parts. The first part consists of the points on the imaginary axis, the second part is the small indentation and the third part is the semicircle at infinity. Since it is assumed that the response of the system is zero at infinity, the last part does not contribute to the integral. Therefore the integral becomes

$$H(s_i) = \frac{1}{2\pi} \int_{-\infty}^{s_i-\epsilon} H(j\omega) \frac{1}{(j\omega - s_i)} d\omega + \frac{1}{2\pi} \int_{s_i+\epsilon}^{\infty} H(j\omega) \frac{1}{(j\omega - s_i)} d\omega +$$

$$\frac{1}{2\pi j} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2\pi} H(j\omega) \frac{1}{(j\omega - s_i)} e^{j\phi} d\phi. \tag{2.43}$$

It can be derived that the last part converges to $\frac{1}{\epsilon} H(s_i)$ for $\epsilon \to 0$. Therefore $H(s_i)$ can be computed from

$$H(s_i) = \frac{1}{\pi} \left( \int_{-\infty}^{s_i-\epsilon} \frac{H(j\omega)}{(j\omega - s_i)} d\omega + \int_{s_i+\epsilon}^{\infty} \frac{H(j\omega)}{(j\omega - s_i)} d\omega \right). \tag{2.44}$$

Because this integral has no singular points, it can be approximated by a summation and therefore $H(s_i)$ can be computed from the measured frequency response. This operation is a local averaging which preserves the holomorphic function in a correct way, but averages out noise. This operation can be done multiple times on the same FRF in order to smooth it even further. The number of iterations is not unlimited though, because small numeric errors are made during each step. An example of the application of this method is shown in Figure 2.11. The figure shows the FRF of the model of the system (black, - -), FRF plus noise (red) and the filtered noisy FRF (blue), which is the result of filtering the red FRF ten times. It can be observed that the original amplitude and phase are reconstructed very well. The contour integral assures that the phase-gain relationship is preserved and even reconstructed, while filtering the noise away. Notice also that the $-2$ slope at high frequencies becomes visible in the filtered FRF.
Figure 2.10: Contour integral used for noise filtering.

Figure 2.11: Noise filtering with contour integrals. The FRF of the model of the system (black --), FRF plus noise (red) and the filtered noisy FRF (blue) are shown.
Chapter 3

Generalized MIMO Nyquist

3.1 Introduction

This chapter describes a method to analyze the relative stability of MIMO feedback controlled dynamical systems, in a data-based way. Relative stability means that the poles of the system do not only lie in the LHP, but they have a certain amount of damping, which gives a certain stability margin for the system. The analysis can be performed for lightly damped mechanical systems, using only FRD of the system. This method is based on generalizations of the conventional Nyquist stability theorem [36]. In most literature the term generalized refers to the generalization from Single-Input-Single-Output (SISO) to Multi-Input-Multi-Output (MIMO) systems. In earlier literature [37], however, the term generalized refers to generalization of the contour that is used. When an alternative contour is used, a so-called relative stability analysis is performed, to guarantee that the system has a certain amount of damping. To avoid confusion, the MIMO generalization will be addressed as MIMO Nyquist consistently.

The proposed method combines these two generalizations in order to analyze the so called ‘relative’ stability of (MIMO) dynamical systems. It accomplishes this by drawing the generalized MIMO Nyquist plot for a number of contours in the s-plane as will become clear in Section 3.4.1.

The outline of this chapter is as follows. First some important properties of transfer functions are discussed that are used later to derive the argument principle and conformal mapping. These techniques are necessary to analyze the Nyquist plots. Next, the conventional Nyquist stability theorem is discussed, after which the proposed generalizations of the Nyquist theorem and their application are described. Then, the opportunities and limits of the proposed theory are evaluated. Finally, the data-based application of this technique is discussed.

3.2 Methods

3.2.1 Properties of transfer functions

Transfer functions belong to the class of meromorphic functions that are defined as functions that are analytic everywhere on a certain domain $D$, except for the pole locations [23]. A function $f(z)$ is called analytic if it possesses a derivative at $z$ and at every point in some neighborhood of $z$. The notion of differentiability has to be specified more precisely, because complex functions are considered.

**Definition 3.1. (Complex derivative) The complex derivative of a complex function $f(z) : \mathbb{C} \rightarrow \mathbb{C}$ is defined as**

$$
\frac{df(z)}{dz} = f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}
$$

(3.1)

This form of this definition resembles the normal definition of derivative, but the function and its argument are complex. The complex derivative of $f(z)$ only exists when the value of the derivative is the same for a given $\Delta z$, regardless of its orientation. This is the case for analytic functions. A complex
Applications of Transfer Function Data

The function \( f(z) = u(x, y) + jv(x, y) \), where \( z = x + jy \), is analytic if it satisfies the Cauchy-Riemann conditions

\[
\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{(3.2)}
\]

\[
\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \quad \text{(3.3)}
\]

For more information on the Cauchy-Riemann equations see e.g. [32]. The concept of conformal mapping, which will be used to analyze Nyquist plots, can now be defined.

**Definition 3.2.** (Conformal mapping) A mapping \( f(z) : \mathbb{C} \to \mathbb{C} \) is called a conformal mapping if \( f(z) \) is analytic and \( f'(z) \neq 0 \) in a certain region \( R \) [32].

Since transfer functions are meromorphic functions, which are differentiable infinitely many times, the mapping from the s-plane to the open-loop plane is a conformal mapping, except for the pole locations. Figure 3.1 shows an example of a conformal mapping by an arbitrarily chosen transfer function. Note that although the lines are curved after mapping, the lines are locally still perpendicular. This demonstrates that a conformal mapping preserves the magnitude and sense of angles. This is proved in the following theorem.

**Theorem 3.3.** A conformal mapping preserves the the magnitude and sense of angles.

**Proof.** Consider the mapping \( w = f(z) \). Let \( \Delta z = |\Delta z|e^{j\theta} \) and \( \Delta w = |\Delta w|e^{j\phi} \) be small variations of the points \( z \) and \( w \). The derivative can be written as

\[
f'(z) = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} e^{j(\phi - \theta)} \quad \text{(3.4)}
\]

From this it is clear that the amplitude and phase of \( f'(z) \) satisfy

\[
|f'(z)| = \lim_{\Delta z \to 0} \left| \frac{\Delta w}{\Delta z} \right| \Rightarrow |\Delta w| = |f'(z)||\Delta z| \quad \text{(3.5)}
\]

\[
\arg(f'(z)) = \lim_{\Delta z \to 0} (\phi - \theta) \Rightarrow \arg(\Delta w) = \arg(\Delta z) + \arg(f'(z)) \quad \text{(3.6)}
\]

to any degree of approximation. If the mapping is conformal, \( f'(z) \) exists and is independent of the path along which \( \Delta z \) goes to zero. Therefore its magnitude \( |f'(z)| \) and phase \( \arg(f'(z)) \) are also independent of this path. Equation 3.5 implies that the length of an infinitesimal segment is scaled by a factor \( |f'(z)| \). Similarly Equation 3.6 shows that angles are rotated by \( \arg(f'(z)) \) independently of the initial direction of the segment. Hence, two segments forming an angle before transformation (at a certain intersection point of the two segments), will have the same angle after transformation [32].

Theorem 3.3 shows why the lines in Figure 3.1 are locally tangential both before and after the mapping.

### 3.2.2 Argument principle

The conventional Nyquist stability theorem relates the open-loop frequency response to the number of closed-loop poles in the RHP. It uses the argument principle to relate the number of poles in the RHP to the number of encirclements of \(-1\) by the Nyquist curve [19]. Therefore the argument principle is discussed next shortly. As will become apparent in Section 3.3, both the poles and zeros of the open-loop transfer function are relevant in this analysis.

**Theorem 3.4.** (Argument principle) A clockwise contour map of a complex function will encircle the origin \( Z - P \) times in clockwise direction, where \( Z \) is the number of zeros and \( P \) is the number of poles of the function inside the contour.

**Proof.** See Appendix A.
Figure 3.1: Conformal mapping of \( H(s) = \frac{1}{(s + 5)^2} \)

Figure 3.2: The mapping encircles the origin if the contour contains a zero.
Thus, the number of poles and zeros within a contour can be related to the number of encirclements of the origin. This is now illustrated by means of an example. The top part of Figure 3.2 shows the mapping of a dynamical system \( s \rightarrow H(s) \) with two poles and one zero, indicated by (\( \times \)) and (\( \circ \)) respectively. The points \( s \) indicated by the circular contour on the left are mapped to \( H(s) \) on the right. As \( s \) traverses the contour from the arbitrary starting point \( s_0 \), \( \arg(s - z_k) \) will increase and decrease, but will not undergo a change of 360°. Therefore \( \arg H(s) \) will not encircle the origin. The bottom part of Figure 3.2 shows a similar mapping, but now the zero lies inside the contour. Hence the argument of \( \arg(s - z_k) \) undergoes a net change of 360°, which causes the mapping to encircle the origin. The behavior would have been similar if the contour would have enclosed a pole instead of a zero, but then the direction of encirclement would be reversed.

### 3.2.3 Conformal mapping theorem

Although the argument principle is well-known and often used, it is not the only way of determining whether poles and zeros lie in a certain contour. The conformal mapping property can be used to determine this directly.

**Theorem 3.5.** Points that lie inside the contour before mapping, will also lie inside the contour after mapping.

**Proof.** This theorem follows directly from Theorem 3.3. \( \square \)

The question of course is, how to interpret what the inside and outside of the contour is. It is not trivial for mappings that have a complex shape. Graphically it can be done as follows. The contour in Figure 3.2 is shaded on the inside. The shading is created by calculating the mapping of a number of points that lie inside the contour and drawing lines from these points to the contour. The inside of the contour can also be identified from the direction of the contour.

The zeros \( z_k \) of the system are, by definition, mapped on the origin. Thus,

\[
s = z_k \rightarrow H(s) = 0. \tag{3.7}
\]

It can easily be observed from Figure 3.2 that the origin lies outside the mapping in the top part while it is inside the mapping in the bottom graph. Hence, only the bottom contour encloses a zero.

Poles \( p_l \) of the system are of course defined by

\[
s = p_l \rightarrow H(s) = \infty. \tag{3.8}
\]

Thus imagine the poles as lying far outside the visible part of the plot. Because the shading is on the outside of the circle, the poles that lie at infinity are outside the contour. If another system is considered that has the poles inside the contour, see Figure 3.3, it can be seen that the mapped contour is turned inside out. The shading is on the outside of the circle, indicating that the poles lie within the contour. Note that in this case the zero lies outside the contour. The mapping encircles the origin two times, because there are two poles inside the contour.

The advantage of the conformal mapping way of looking at the mapping becomes obvious when we consider a system that has a pole and a zero inside the contour. The argument principle tells us that the mapping will not encircle the origin, because \( Z - P = 0 \). But if only the mapping is known, which is the case when using the Nyquist stability criterion, it is impossible to tell the difference between this case and the case when no poles nor zeros lie inside the contour. The conformal mapping property, however, can be used to determine what the inside and what the outside of the contour is. In this way it can be determined whether the poles are inside or outside the contour. The conformal mapping theorem gives information \( P \) and \( Z \) separately, while the argument principle only gives information on \( Z - P \). Thus it is very easy to draw conclusions on stability using the conformal mapping theorem, and this does not require information about the open-loop poles.

This is illustrated by Figure 3.4 which shows the case when one pole and one zero lie inside the contour. Note that the origin is outside the mapped contour. The shading is also on the outside of the contour indicating that a zero lies outside the contour. Because the shading is also pointing towards infinity, a pole lies inside the contour too. Compare this figure to the top part of Figure 3.2 where the shading lies on the inside of the contour, indicating that no poles or zeros lie inside the contour.

Figure 3.4 also shows the disadvantage of the conformal mapping approach. The small clockwise loop (close to zero) is caused by the pole that lies outside the contour, but is hardly visible. The influence of poles can become invisible in the mapping if the poles lie far from the contour. But many cases the conformal mapping approach can give valuable direct insights.
Figure 3.3: Argument principle: The zero lies outside the contour while the poles (mapped at infinity) are inside the contour.

Figure 3.4: Argument principle: A pole and a zero lie inside the contour.
3.3 Conventional Nyquist stability theorem

In this section, it is discussed how the theory of Sections 3.2.2 and 3.2.3 can be applied to analyze Nyquist plots. The strength of the Nyquist stability theorem lies in the fact that open-loop information can be used to determine the stability of the closed-loop [19]. The following lemma explains why this is possible.

Lemma 3.6. Consider the open loop transfer function \( L(s) \) of a negative feedback system. It holds that

- The poles of \( 1 + L(s) \) are the poles of the open-loop transfer function \( L(s) \)
- The zeros of \( 1 + L(s) \) are the poles of the corresponding closed-loop system

Proof. The closed-loop transfer function is defined as

\[
\frac{Y(s)}{R(s)} = \frac{L(s)}{1 + L(s)},
\]

where \( Y(s) \) is the output and \( R(s) \) the reference of the system. From this equation it is clear that the closed-loop poles are given by

\[
1 + L(s) = 0 \quad (3.10)
\]

Let \( N(s) \) and \( D(s) \) denote the numerator and denominator of \( L(s) \). Then \( 1 + L(s) \) can be written as

\[
1 + L(s) = 1 + \frac{N(s)}{D(s)} = \frac{D(s) + N(s)}{D(s)}, \quad (3.11)
\]

from which it can be seen that the poles of \( 1 + L(s) \) are the poles of the open-loop transfer function and the zeros of \( 1 + L(s) \) are the closed-loop poles.

3.3.1 Application of the argument principle

Next, it is discussed how the argument principle can be used to determine the stability of a closed-loop system using the Nyquist plot of the open-loop.

Theorem 3.7. (Application argument principle) The net number of encircllements \( N \) of the point \(-1\) by the open-loop transfer function \( L(s) \), for \( s \) traversing the D-contour, is equal to the number of unstable closed-loop poles \( Z \) minus the number of open-loop unstable poles \( P \), i.e. \( N = Z - P \) [19].

Proof. \( 1 + L(s) \) is evaluated on a D-contour that encircles the whole Right Half Plane (RHP). Equation 3.11 shows that the zeros of \( 1 + L(s) \) are the closed-loop poles \( Z \) and the poles of \( 1 + L(s) \) are the open-loop poles \( P \). By Theorem 3.4, \( 1 + L(s) \) will encircle the origin \( Z - P \) times. Because the Nyquist plot of \( 1 + L(s) \) is the same as \( L(s) \) shifted 1 to the right, plotting \( L(s) \) and evaluating the encirclements of \(-1\) yields the same result.

If the number of unstable open-loop poles is known, the stability of the closed-loop system can be determined. The number of closed-loop unstable poles can then simply be computed from \( Z = N + P \).

3.3.2 Application of the conformal mapping theorem

We can also use the conformal mapping property very easily to determine the stability of closed-loop systems on the basis of the frequency response function of the open-loop transfer function.

Theorem 3.8. (Application conformal mapping) A system \( L(s) \) is closed-loop stable if the \(-1\) point lies outside the mapping of the D-contour by \( L(s) \).

Proof. Equation 3.10 shows that the closed-loop poles are mapped on the point \(-1\). Applying Theorem 3.5 shows that the closed-loop poles that do not lie inside the contour, are in the LHP. Hence, the closed-loop system is stable.

It has to be emphasized that while the conventional Nyquist stability theorem requires information about the number of open-loop unstable poles, this information is not necessary when using the conformal mapping property. However, all poles must be visible in the Nyquist plot in order to be able to prove that the system is stable. Instability can often be seen directly with this method. A thorough analysis of the applicability of the conformal mapping theorem will be given in Section 3.5.
3.3.3 Example

Figure 3.5 shows an example of the application of the Nyquist stability theorem. The system $H(s)$ in this example has two unstable open-loop poles ($P = 2$) which are stabilized by unity feedback ($Z = 0$). A simplification in the application can be made if the system has more poles than zeros. Theoretically, the mapping of the Nyquist D-contour has to be evaluated. Fortunately, the mapping of the semicircle at infinity is zero for such a system, so evaluating $H(s)$ for all frequencies $\omega$, i.e. $H(j\omega)$ gives the correct Nyquist plot. This is also the case for the system in this example.

The argument principle argues that there will be $N = Z - P = -2$ encirclements of the point $-1$, as seen in the figure. If the number of unstable open-loop poles is known, stability of the system is proved because the mapping encircles the $-1$ point $-2$ times. The conformal mapping theorem applies as well. The open-loop poles will map to infinity. Thus the open-loop poles lie inside the contour before and after the mapping. This can be seen in the figure by observing that the shading of the mapped contour points towards infinity. The closed-loop poles are mapped to $-1$. By looking at the shading of the mapped contour again, it can be observed that the closed-loop poles indeed lie outside the contour before and after the mapping. Stability can be determined easily by looking whether the point $-1$ lies inside the contour, without knowledge of the number of open-loop unstable poles, but assuming all poles are visible in the Nyquist plot (see Section 3.5).

3.4 Generalized Nyquist stability theorem

Two generalizations of the Nyquist stability theorem are possible. Firstly generalized contours can be defined to check the so-called relative stability. Furthermore, the theory can be extended to MIMO systems. The generalizations and their applications are the topic of this section.

3.4.1 Generalized contour

The Nyquist stability theorem shows stability by checking whether poles lie in the RHP. It accomplishes this by defining a contour that encloses the whole RHP and checking whether the poles lie in this contour. As discussed, this can be done either by using the argument principle or the conformal mapping theorem. This D-contour is shown in Figure 3.6a. By choosing a different contour, a certain amount of absolute or relative damping is guaranteed when the poles do not lie inside the contour [37],[34]. Figures 3.6b shows a contour that can be used to guarantee absolute damping $\sigma$ and the contour of Figure 3.6c assures the poles have a relative damping greater than $\theta$. When the Nyquist stability theorem is evaluated using these generalized contours, and it shows that the system is ‘stable’, this indicates that the poles do not lie in the generalized contour. Thus a certain amount of damping is guaranteed. This is called a relative stability analysis and leads to the following definition.
Applications of Transfer Function Data

Figure 3.6: Contours in the s-plane. a: conventional Nyquist contour, b: contour absolute damping $\sigma$, c: contour relative damping $\theta$.

**Definition 3.9.** *(Relative stability)* A system is called relative stable if the closed-loop poles of the system lie outside a predefined contour $C$.

### 3.4.2 MIMO generalization

The other generalization that can be made, is generalizing the theorem to MIMO systems. The closed-loop transfer function matrix of system $L(s)$ with factorization $L(s) = N(s)D(s)^{-1}$ is given by

\[
L_{cl}(s) = L(s)(I + L(s))^{-1} = N(s)(D(s) + N(s))^{-1}.
\]

(3.12)

(3.13)

$\det(D(s) + N(s))$ is the closed-loop characteristic polynomial. Zeros of the characteristic polynomial correspond to the closed-loop poles. Consequently, the closed-loop system is stable if no zeros of the characteristic polynomial lie in the RHP. An additional requirement is that $\det(I + L(s)) \neq 0$, else the closed-loop system is non-proper, see Equation 3.12. Substituting $L(s) = N(s)D(s)^{-1}$ gives

\[
\det(I + L(s)) = \frac{\det(D(s) + N(s))}{\det(D(s))}.
\]

(3.14)

Comparison of Equations 3.14 and 3.11 shows that there is a relation between the open-loop and closed-loop poles just like in the SISO case [15]. Therefore Equation 3.14 can be used to assess the stability of MIMO systems.

**Argument principle in a MIMO context**

The argument principle can be used to determine whether the closed-loop system is stable [15].

**Theorem 3.10.** *(Application argument principle, MIMO)* Let $P$ denote the number of open-loop unstable poles of a MIMO transfer function $L(s)$ and $C$ denote the contour on which $L(s)$ is evaluated. The closed-loop system $L_{cl}(s)$ as defined in Equation 3.12 is stable if and only if [15]:

1. $\det(I + L(s))_{C} \neq 0$
2. $\det(I + L(s))_{C}$ encircles the origin $P$ times (counterclockwise)

**Proof.** Similarly to the SISO case, zeros of Equation 3.14 are closed-loop poles and poles of Equation 3.14 are open-loop poles of the system. By Theorem 3.4 the number of encirclements is determined by these poles and zeros. The number of unstable closed-loop poles is zero when the mapping encircles the origin $P$ times counterclockwise. \(\square\)
Conformal mapping in a MIMO context

Just as in the SISO case, the conformal mapping theorem can be applied to analyze the stability of the closed-loop system. But first it has to be shown that the mapping

\[ s \to \det(I + L(s)) \]  

(3.15)

is indeed a conformal mapping. The following two results are necessary to show this.

**Lemma 3.11.** The determinant of a $2 \times 2$ transfer function matrix is a rational of polynomials in $s$.

**Proof.** Let $G(s)$ define a MIMO transfer function composed of four SISO transfer functions,

\[ G(s) = \begin{bmatrix} N_{11}(s) & N_{12}(s) \\ D_{11}(s) & D_{12}(s) \\ N_{21}(s) & N_{22}(s) \\ D_{21}(s) & D_{22}(s) \end{bmatrix}, \]

(3.16)

where $N_{kl}(s)$ and $D_{kl}(s)$ are the numerator and denominator polynomials of the SISO transfer functions. The determinant of this matrix is

\[
\det(G) = \frac{N_{11}N_{22}D_{12}D_{21} - N_{12}N_{21}D_{11}D_{22}}{D_{11}D_{21}D_{22}D_{12}}.
\]

(3.17)

(3.18)

where the dependence of $s$ is omitted for convenience of notation.

**Lemma 3.12.** The mapping $s \to \det(I + L(s))$ is conformal.

**Proof.** It will be shown that $\det(I + L(s))$ is a rational of two polynomials in $s$, just like SISO transfer functions. As already mentioned in Section 3.2.1, these functions belong to the class of meromorphic functions and therefore define a conformal mapping. In general $I + L(s)$ is a $n \times n$ matrix for which the determinant can be expressed as a function of lower order determinants by means of the Laplace expansion [38]. The determinant of an $n \times n$ matrix $A$ with elements $a_{kl}$ can be expressed as

\[
\det(A) = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}
\]

(3.19)

where $C_{kl}$ are the cofactors of the matrix. These are defined as

\[
C_{kl} = (-1)^{k+l} \det(M_{kl})
\]

(3.20)

with $M_{kl}$ the $(n-1) \times (n-1)$ minors of $A$ obtained by deleting row $k$ and column $l$ from $A$. Equation 3.19 defines the Laplace expansion along the first row of $A$. Any other row or column can also be chosen of course. By using the Laplace expansion repeatedly, the determinant can be expressed as a function of second order determinants. Lemma 3.11 shows that these second order determinants are rational functions. In this Laplace expansion the second order determinants are multiplied with elements of the matrix $I + L(s)$, which are also rational. Hence the determinant is a summation of products of rational functions and is therefore a rational of polynomials, which is meromorphic and therefore defines a conformal mapping.

**Theorem 3.13.** (Application conformal mapping, MIMO) Let $L(s)$ denote a MIMO transfer function and $\det(I + L(s))_C$ denote the evaluation of the characteristic equation along contour $C$. Then, the closed-loop system $L_C(s)$ as defined in Equation 3.12 is relative stable if and only if the origin does not lie inside the mapping $\det(I + L(s))_C$.

**Proof.** Equation 3.14 shows that the closed-loop poles are mapped on the origin by $\det(I + L(s))_C$. Lemma 3.12 shows that $\det(I + L(s))_C$ defines a conformal mapping. By Theorem 3.5, poles that lie outside the contour before the mapping, will also lie outside the contour after mapping. Therefore if the origin lies inside the contour the closed-loop system must contain unstable closed-loop poles.
Applications of Transfer Function Data

\[ m_1 \quad m_2 \]
\[ k_1 \quad k_2 \]
\[ d_1 \quad d_2 \]

Figure 3.7: Fourth order MIMO system consisting of two masses connected via springs and dampers.

\[ \begin{array}{ccc}
\sum (4) & K (4) & H(s) (4) \\
\end{array} \]

Figure 3.8: Feedback scheme for the mass-spring-damper system.

3.4.3 Combination: Generalized MIMO Nyquist

When both generalizations are combined, a relative stability analysis can be performed on a MIMO system, which gives a certain robustness margin. Using the generalized contours, the minimum amount of damping in the system can be computed. This is an improvement compared to the conventional MIMO Nyquist theorem that cannot give a clear interpretation of stability margins. In the conventional MIMO Nyquist theorem the gain and phase margins can at best be interpreted as the independent gain and phase margins that are allowed in each input channel [27].

To combine the two generalizations evaluate

\[
\text{det}(I + L(s))_C
\]

where \( C \) is now a generalized contour, recall Figures 3.6b and 3.6c. Determining the stability of the system requires some precaution. In the conventional Nyquist theorem, assumptions on the open-loop pole locations have to be made in order to be able to assess the stability of the system. Often, it can be assumed that the system is open-loop stable or that the number of unstable poles is known. Then the argument principle can readily be used to determine stability, see Theorem 3.10. However, when a generalized contour is used, the notion of stability and instability changes. Poles that lie in the LHP can be relative unstable, depending on the chosen contour. Therefore making assumptions about the number of relative unstable open-loop poles is not as easy as making assumption on the number of open-loop RHP poles. It would require knowledge on the pole locations, which is not available since the theory will be used on FRD, not a parametric model. The proposed technique is the following. Start with the conventional D-contour and gradually increase the amount of relative or absolute damping.

Each time an open-loop pole is passed, one of the loops in the Nyquist plot will go to infinity and its direction of rotation will change. By counting the number of times a loop ‘flips’ via infinity, the number of open-loop relative unstable poles can be counted. Then the argument principle can be used to determine whether the system is unstable.

When the contour passes a closed-loop pole, the contour will pass the \(-1\) point somewhere, which is the indication for a system that initially was stable to become closed-loop unstable. This can be determined by applying the argument principle, but also by using the conformal mapping theorem. When the Nyquist curve passes the \(-1\) point, this means that a closed-loop pole now lies in the contour, hence the system is unstable.

The argument principle can also be used to determine relative stability by investigating whether the \(-1\) point lies inside or outside the contour, by applying Theorem 3.13. Multiple contours have to be evaluated for this method too, as will become clear in Section 3.5.
3.4.4 Example

Consider a fourth order mass spring damper system, as shown in Figure 3.7. The system is stable because all coefficients are chosen positive. There are two inputs for the system, each mass is driven by a force \( F \). For both masses the position and velocity are measured, which implies that the number of outputs is four. The system is controlled by means of state feedback. The feedback gain is calculated using a standard pole placement algorithm, that places the poles more into the LHP. The feedback loop is shown in Figure 3.8, where \( K \) is the \( 4 \times 2 \) feedback matrix and \( H(s) \) is the \( 2 \times 4 \) system transfer function matrix. The numbers above the connections show the dimensions of the signals. Let \( L(s) = KH(s) \) denote the loop gain.

Figure 3.9 demonstrates the application of the proposed method. The generalized MIMO Nyquist plot is plotted for five different contours, indicated by a-e, which differ in the amount of relative damping. From a to e the damping increases. Of course, contours with absolute damping could have been used equally well.

For each contour the left plot shows the s-plane with the open-loop and closed-loop pole locations and the contour. The poles of the system seem to have a considerable amount of damping, but this is not the case. The scale of the real axis is very small to be able to determine whether a pole lies in the contour or not. The fourth order system under consideration has four poles, of course, which are symmetric with respect to the real axis. Although the left plot only shows the poles in the upper half of the plane, the other poles also contribute to the mapping. The full contour of Figure 3.6c is evaluated. In conventional Nyquist nomenclature this would be equivalent to considering both positive and negative frequencies.

The plot on the right shows the mapping of each contour via the function \( \det(I + H(s)) \). By plotting contours with different amounts of damping, the minimum amount of damping can be determined. The assumption is made that the open-loop has no unstable poles, which is the same assumption that is needed for a conventional Nyquist analysis. The system becomes relative unstable when the Nyquist plot passes the origin. The origin should be considered instead of the \( -1 \) point, because this is a MIMO case and therefore theorem 3.10 applies. The conformal mapping theorem can also be used to analyze the plots. As explained before, an open-loop pole is mapped to infinity and a closed-loop pole to zero. Therefore, if the origin lies within the contour, the system is closed-loop relative unstable. A onus which is pointing towards infinity (counterclockwise loop) indicates the existence of an open-loop relative unstable pole. This can be seen for each plot individually, without making assumptions on the number open-loop poles in the contour. So the number of relative unstable open-loop poles \( P \) and the number of relative unstable closed-loop poles \( Z \) can be determined directly. This will now be clarified by discussing cases a-e of Figure 3.9.

- **a** This case is equivalent to the conventional MIMO Nyquist. The contour is chosen to have zero relative damping and therefore it lies on the imaginary axis. The assumption is made that all open-loop poles lie in the LHP. Therefore the argument principle states that \( P = 0 \), from the figure it is observed that \( N = 0 \) and therefore \( Z = 0 \).
  Conformal mapping gives directly: \( P = 0, Z = 0 \).

- **b** For the second case, the contour has a certain amount of relative damping. This causes one of the open-loop poles to lie inside the contour. A part of the Nyquist plot has flipped via infinity, indicating an open-loop pole is passed. The system is still closed-loop relative stable, since the contour has not passed the origin yet.
  Conformal mapping gives directly: \( P = 1, Z = 0 \).

- **c** The contour has now passed the origin, indicating that the system is closed-loop relative unstable. The minimum amount of relative damping of the closed-loop system is found somewhere between the contours of subfigures b and c.
  Conformal mapping gives directly: \( P = 1, Z = 1 \).

- **d** Another open-loop pole is passed again, which can be seen because again one of the loops has flipped via infinity.
  Conformal mapping gives directly: \( P = 2, Z = 1 \).

- **e** Since the Nyquist plot has passed the origin again another closed-loop pole must lie in the contour now.
  Conformal mapping gives directly: \( P = 2, Z = 2 \).

This example shows that by evaluating different contours, a MIMO relative stability analysis can be performed. With this method it can be determined what the minimum amount of damping in the system is, which gives a certain robustness margin. Alternatively, this method can be used to design a controller such that the origin is not passed for a contour that has the desired amount of damping.
Figure 3.9: Generalized MIMO Nyquist applied to the fourth order dynamical system. Figures a to e contain evaluation of contours that have an increasing amount of relative damping.
3.5 More on the relative stability analysis

3.5.1 Evaluation of multiple contours

The proposed method can derive the relative stability of a system by considering a number of contours in the s-plane. This is necessary because the number of relative unstable open-loop poles is unknown for a certain generalized contour, because no parametric model is available. The conformal mapping theorem however, does not use information on the number of open-loop unstable poles. Therefore the question could arise whether it is possible to determine the stability of the system by considering only one contour with a certain amount of damping and assess stability using the conformal mapping theorem.

In general this is not possible because not all aspects of a system are visible in a generalized Nyquist plot. When a pole of the system lies far from the contour that is plotted, its influence on the mapping can be so small that it is not visible in the Nyquist plot.

Figure 3.10 gives an illustrative example where an unstable pole is not visible which could lead to wrong conclusions on stability. The generalized Nyquist plot for a SISO system with four open-loop poles \((\times ol)\) and four closed-loop poles \((\times cl)\) is plotted for two generalized contours which differ in the amount of absolute damping. Only positive frequencies are considered for convenience. Applying the conformal mapping theorem on the dotted contour leads to the conclusion that the system is stable, because the \(-1\) point lies outside the contour. But in reality there is an unstable pole which is not visible in this Nyquist plot. The solid Nyquist plot shows that there is a counterclockwise loop and therefore the system is unstable, since for this loop the \(-1\) point lies within the contour. Both contours should of course give the same result because they are both on the same side with respect to all poles. In order to be sure this effect does not lead to wrong conclusions, a number of contours should be considered, starting from a situation in which the number of unstable closed loop poles is known. If then a closed-loop pole is passed it will be clearly visible in the Nyquist plot, because then the mapping will cross the \(-1\) point.

3.5.2 Loops in the Nyquist plot

Although it is not always possible to prove that a system is stable, there are cases in which it is clear that the closed-loop system is unstable, even without information on the number of relative unstable open-loop poles. This can be done by analyzing resonance circles in the Nyquist plot. For mechanical systems, a transfer function \(H(s)\) can be written as a summation of \(n\) modal components

\[
H(s) = \sum_{m=1}^{n} \frac{R_m}{s - p_m} + \frac{R_m^*}{s - p_m^*}
\]  

(3.22)

where \(R_m\) and \(p_m\) are the residue and pole location of mode \(m\). Evaluating this expression on the imaginary axis for one of the modes gives

\[
H_m(j\omega) = \frac{R_m}{-a_m + j(\omega - b_m)} + \frac{R_m^*}{-a_m + j(\omega + b_m)}
\]

(3.23)
where $a_m$ and $b_m$ are the real and imaginary part of the pole, so $p_m = a_m + b_mj$. Around mode $m$, $\omega \approx b_m$, the second term is much smaller than the first term, provided that the damping of the pole, $a_m$, is small. Then $H_m(j\omega)$ can be rewritten

\begin{align}
H_m(j\omega) &= \frac{R_m}{-a_m + j(\omega - b_m)} \\
&= \frac{-R_m}{2a_m} + \frac{R_m}{2a_m} \left[ \frac{(a_m + j(\omega - b_m))^2}{a_m^2 + (\omega - b_m)^2} \right] \\
&= \frac{-R_m}{2a_m} \left[ 1 + e^{j\phi_m} \right]
\end{align}

(3.24)

(3.25)

(3.26)

where

$$\phi_m = 2 \arctan \left( \frac{\omega - b_m}{a_m} \right).$$

Equation 3.26 shows that the transfer function of a single mode forms a circle in the $s$-plane. The summation of $n$ modes as in Equation 3.22 will result in a combination of $n$ (parts of) circles. These circles can either have a clockwise or a counterclockwise direction, depending on the stability of the corresponding pole. In the following lemma it is derived that that counterclockwise loops in a Nyquist plot correspond to unstable open-loop poles.

**Lemma 3.14.** A Nyquist plot which contains a counterclockwise loop represents a transfer function that has at least one open-loop RHP pole.

**Proof.** A transfer function can be represented as

$$H(s) = \frac{N(s)}{D(s)}.$$ 

(3.28)

Let $P$ denote the number of unstable RHP poles of $H(s)$, $Z$ the number of RHP zeros of $H(s)$ and $N$ the number of (clockwise) encirclements of the origin by the Nyquist plot of $H(s)$. There are 2 cases:

**Case 1:** The counterclockwise loop encircles the origin.

The argument principle yields:

$$Z - P = N = -1$$ 

(3.29)

As $Z \geq 0$ it is necessary that $P \geq 1$, which means that $H(s)$ contains at least one unstable pole.

**Case 2:** The counterclockwise loop does not encircle the origin.

In this case, the transfer function must be shifted such that the origin is encircled. This does not change the pole locations of the system. To see this, add a complex number $c$ to the transfer function, yielding

$$H(s) = \frac{N(s)}{D(s)} + c = \frac{N(s) + cD(s)}{D(s)} = \frac{N'(s)}{D(s)}.$$ 

(3.30)

Let $Z'$ denote the number of RHP poles of $N(s) + cD(s)$. Then it holds that

$$Z' - P = N = -1$$ 

(3.31)

As $Z' \geq 0$ it is necessary that $P \geq 1$, which again means that $H(s)$ contains at least one unstable pole.

This lemma can be used to prove that a system is unstable in certain cases. Consider Figure 3.11, where four cases for a part of the Nyquist plot are shown. A resonance circle can be clockwise or counterclockwise and can encircle the -1 point or not. The four cases are

**a) Counterclockwise loop not encircling -1.** Lemma 3.14 shows that $P > 0$, hence $Z = 0$ only if $N < 0$. A counterclockwise encirclement of the -1 point is required for stability. The only way to achieve this is to make the counterclockwise loop encircle -1. One could argue that an other part of the Nyquist plot could create the necessary counterclockwise encirclement. But if there is another counterclockwise loop, there are at least 2 unstable open-loop poles and hence 2 counterclockwise encirclements are required. So a counterclockwise loop that does not encircle the -1 point indicates

28
that the system is unstable. This can be derived using the conformal mapping theorem, because the $-1$ point lies inside the contour when a counterclockwise loop does not encircle $-1$.

The converse is not true unfortunately. If there are no counterclockwise loops in the Nyquist plot the system can still be unstable, when a counterclockwise loop is not visible because it is related to a damped pole, recall the example corresponding to Figure 3.10.

b) **Counterclockwise loop encircling $-1$.** Lemma 3.14 shows that $P > 0$. Without information on the exact number of unstable open-loop poles no conclusion can be drawn from this plot. The conformal mapping theorem gives no indications for instability, since $-1$ lies outside the contour. But there can be counterclockwise loops that are not visible.

c) **Clockwise loop not encircling $-1$.** Same information as contour b; without information on the exact number of unstable open-loop poles, no conclusion can be drawn from this plot. The conformal mapping theorem gives no indications for instability, since $-1$ lies outside the contour. But again there can be counterclockwise loops that are not visible in the Nyquist plot.

d) **Clockwise loop encircling $-1$.** Another case in which the system can immediately be identified as closed-loop unstable is when there are clockwise loops around $-1$. The argument principle gives

$$Z - P = N > 0, \quad P \geq 0 \Rightarrow Z \geq 1 \quad (3.32)$$

Hence the system is closed-loop unstable. This also holds when there is a clockwise loop around $-1$, but the total number of encirclements is smaller than zero. This can only be achieved by adding counterclockwise loops around $-1$. But with each added loop not only the number of counterclockwise loops is changed ($N = N - 1$), but also at least one unstable pole is added ($P = P + 1$), so this does not change the number of unstable poles. The conformal mapping theorem would directly give that the $-1$ point lies in the contour, which can not be changed by adding loops in an other part of the contour.

In cases a) and d) the conclusion can be drawn that the system is unstable, without information on the number of relative unstable open-loop poles. The conformal mapping is an intuitive approach that immediately shows that in these cases a closed-loop pole lies in the contour.
3.6 Data-based application of the generalized MIMO Nyquist

3.6.1 Procedure

The generalized MIMO Nyquist theorem can also be applied data-based for a certain class of systems. As discussed in Chapter 2, for lightly damped mechanical systems a non-parametric transfer function $H(s_i)$ can be calculated using Frequency Response Function (FRF) measurements. This means $H(s_i)$ is available for the whole complex plane, not only the imaginary axis. When the controller $C(s)$ is known, the open-loop $L(s)$ can be computed for each complex frequency point $s_i$, i.e.

$$L(s_i) = H(s_i)C(s_i).$$ (3.33)

Because the response of the open-loop is known at every point in the complex plane, generalized contours that do not lie on the imaginary axis can be evaluated as discussed in this chapter. With this method, a relative stability analysis can be performed to estimate the minimum amount of damping in the closed-loop system. Thus the procedure is as follows

1. Compute $H(s_i)$ on generalized contours
2. Compute $\det(I + L(s_i))$
3. Plot the the imaginary versus the real part of this determinant
4. Analyze the plot with the argument principle and/or the conformal mapping theorem

This procedure will now be demonstrated by means of an example.

3.6.2 Example

The fourth order mass-spring damper system, described in Chapter 2, is used as a setup to test the proposed theory. Since the positions of both masses are measured and forces are applied to both masses, it is a MIMO system. The damping constants $d_i$ of the system, which is depicted in Figure 3.12, are chosen small such that the system is lightly damped. TFD can be computed for this system using the Cauchy methods also described in the previous chapter, which is the first step of the procedure. A controller is designed to add damping to the system. The controller

$$C(s) = \begin{bmatrix} 50(s + 1) & 0 \\ 0 & 5(s + 1) \end{bmatrix}$$ (3.34)

adds damping to the system because of the PD like structure. It will shift the poles of the system more into the LHP. Since the controller is known, it is now possible to compute

$$\det(I + L(s_i)) = \det(I + H(s_i)C(s_i))$$ (3.35)

data-based for every $s_i$. This completes step 2 of the procedure.

The generalized MIMO Nyquist plot can now be drawn, which is the third step. The mapping of four generalized contours that differ in the amount of absolute damping are computed, see Figure 3.13. The left part of the figure shows the four contours in the $s$-plane. The open-loop ($\times_{ol}$) and closed-loop ($\times_{cl}$) pole locations, computed with a parametric model of the system, are also shown. The right part shows the resulting generalized MIMO Nyquist plots, which are mappings of the contours on the left to the $\det(I+L)$ plane. The red lines correspond to the data-based approach described in this section. The results obtained from a parametric model of the system are shown in black to be able to check the data-based results. The shading of the line indicated the inside of the generalized contour. If the origin lies inside the contour, the system is relative unstable. The final step is to analyze the resulting plots a-d, which is done next, using the conformal mapping approach.
Applications of Transfer Function Data

Figure 3.13: Data-based application of the generalized MIMO Nyquist theorem.
Figure 3.14: Gain of the system and the data-based approximation as a function of \( \text{Re}(s) \).

\( a \) The contour contains no poles, and therefore the origin must lies outside the contour. The shading shows that the origin is indeed outside the contour. The data-based approach gives exactly the same mapping as the parametric model-based approach, since the FRD is not extrapolated.

\( b \) In this case, all open-loop poles lie inside the contour. Compared to \( a \), all loops in the mapping have flipped via infinity into the LHP which indicates that all the open-loop poles have been passed. Since the shading does not point towards the origin, the origin and therefore also the closed-loop poles lie outside the contour. Thus the system is relative stable for this contour. Although the data-based approach has an error compared to the parametric model-based approach, the conclusions are the same for both mappings.

\( c \) Now two closed-loop poles are inside the contour. This can be seen in the mapping since the origin has been crossed by two of the loops. The system is relative unstable for this contour, since the shading of the smaller loops points towards the origin. This shows that the origin lies inside the contour and therefore there must be relative unstable closed-loop poles.

\( d \) In this last case the relative damping of the contour is so large, that all poles lie inside the contour. The shading shows that the origin is indeed in the contour, since it points towards the origin. Again there is a difference between the model-based and data-based approach, but the conclusion is again the same for both approaches.

By evaluating different contours that have a relative damping close to 10, the minimum amount of damping is estimated with the data-based approach. This can be seen as shifting the contour in Figure 3.13b into the LHP until the red contour passes the origin. The result is the following:

<table>
<thead>
<tr>
<th>Damping</th>
<th>Model</th>
<th>Approximation</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.4</td>
<td></td>
<td>10.2</td>
</tr>
</tbody>
</table>

This shows that the amount of damping can be estimated accurately, the difference is only 2%. Note that the result is conservative; the actual performance in terms of damping is better than the estimation predicts. This is true for all controllers that move the poles sufficiently far into the LHP. Sufficiently far in this case is, when the closed-loop poles have more damping than the open-loop poles. Then all the resonance circles of the original Nyquist plot will flip into the LHP as in plot \( b \) of the example. So all resonance circles encircle the origin counterclockwise, which means that the origin lies outside the contour and is therefore relative stable. For contours with more damping, these resonance circles will become smaller. At some points there will be a resonance circle that crosses the origin. Then the closed-loop poles with the least damping have been reached.

It can be shown that TFD always gives a smaller gain than the real system, see Figure 3.14. This Figure shows the computation of \( H(s) \) for a fixed value of \( \text{Im}(s) \) as a function of \( \text{Re}(s) \). Because the computation of \( H(s) \) is done on basis of the value of the FRF which is measured at \( \text{Re}(s) = 0 \), the actual value of \( H(s) \) will be higher, since the resonances of the system do not lie on the imaginary axis. Because the TFD predicts a slightly lower gain, the resonance circles of the generalized MIMO Nyquist plot are smaller and will pass the origin sooner when evaluating contours with an increasing amount of damping. This leads to the conservatism in the estimation of the minimum amount of damping.

This example has demonstrated that it is possible to perform a relative stability analysis for a MIMO system by applying the generalized MIMO Nyquist theorem as proposed in this chapter. It provides a
way to estimate the minimum amount of damping of the closed-loop poles of a lightly damped mechanical system, given a certain controller.
Chapter 4

Data-Based Symmetric Root Locus

4.1 Introduction

This chapter is focused on the derivation of a symmetric root locus (SRL) of a system using only frequency response data (FRD). A SRL gives optimal closed-loop pole locations for a quadratic optimality criterion.

It is very special that the SRL can be derived from FRD, because it gives information on the closed-loop pole locations of the system. The computation of closed-loop pole locations is in general not possible without making a parametric model of the system from the FRD. But because transfer function data (TFD) $H(s)$ can be computed from the FRD $H(j\omega)$, it is possible to compute the SRL from the FRD.

The outline of this chapter is as follows. First, the theory of optimal control and of the SRL is introduced. Then, an example is given in which the SRL of a two-mass system is computed and relations to optimal control and the closed-loop pole locations are discussed. Finally, insights on energy in the system and a generalization of the SRL, the optimal return difference equation, are discussed.

4.2 Optimal control

The SRL can be used to compute the optimal closed-loop pole locations for a quadratic performance criterion. It will be derived in this section that an LQR controller can achieve these optimal pole locations. Therefore it is useful to briefly review the derivation of an LQR controller. An LQR controller minimizes the following integral criterion

$$J = \int \left[ y^T Q y + u^T R u \right] dt,$$

(4.1)

for all initial conditions and trajectories of the state-space system defined by the matrices $A$, $B$ and $C$ and the following equations

$$\dot{x} = Ax + Bu \quad \quad \quad (4.2)$$

$$y = Cx \quad \quad \quad (4.3)$$

where $x$ is the state, $u$ the input and $y$ the output.

The Euler-Lagrange equations can be used to find the optimal $u$. Define the Hamiltonian matrix $\mathcal{H}$ as

$$\mathcal{H} = \frac{1}{2}(y^T Q y + u^T R u) + \lambda^T (Ax + Bu),$$

(4.4)

where $\lambda$ is a Lagrange multiplier. The complete derivation of the Euler-Lagrange equations can be found for example in [11]. Here only the solution is given. In Equations 4.5 and 4.6 the Lagrange equations on the left result in the conditions on the right.

$$\dot{\lambda} = -\frac{\partial \mathcal{H}}{\partial x} \quad \Rightarrow \quad \dot{\lambda} = -A^T \lambda - C^T Q C x \quad \quad \quad (4.5)$$

$$\frac{\partial \mathcal{H}}{\partial u} = 0 \quad \Rightarrow \quad Ru + B^T \lambda = 0 \quad \quad \quad (4.6)$$

From Equation 4.6 the optimal control law is found to be

$$u = -R^{-1}B^T \lambda \quad \quad \quad (4.7)$$
Substituting this \( u \) in the system equations, together with Equation 4.5 yields the following set of \( 2n \) equations

\[
\begin{bmatrix}
\dot{x} \\
\lambda
\end{bmatrix} =
\begin{bmatrix}
A & -BR^{-1}B^T \\
-C^TQC & -A^T
\end{bmatrix}
\begin{bmatrix}
x \\
\lambda
\end{bmatrix} = S \begin{bmatrix}
x \\
\lambda
\end{bmatrix},
\]

(4.8)

which is called the Hamiltonian system [20],[27]. It can be proved that the matrix \( S \) has its poles symmetric with respect to the imaginary axis.

\[
det(sI - S) = \prod_i (s - p_i) \prod_i (s + p_i)
\]

(4.9)

This has a direct relation with the symmetric root locus which will be derived in Section 4.3; the stable poles of this matrix are the optimal closed-loop poles that will be achieved by the LQR. The solution of Equation 4.8 may be written as [42]

\[
\lambda(t) = P(t)x(t),
\]

(4.10)

Substituting Equation 4.10 in 4.8 yields, after elimination of \( \dot{x} \),

\[
(\dot{P} + PA + A^TP - PBR^{-1}B^TP + C^TQC)x = 0.
\]

(4.11)

Because this holds for all \( x \) the expression in the brackets, which is the Riccati equation, should be zero. From this equation \( P \) has to be solved. Fortunately, for controllable processes the solution of the Riccati equation converges quickly to the steady state solution \( P \), so \( \dot{P} \) can be taken zero, resulting in the algebraic Riccati equation (ARE),

\[
PA + A^TP - PBR^{-1}B^TP + C^TQC = 0.
\]

(4.12)

With the solution \( P \) of the ARE, the optimal control law is

\[
u = -R^{-1}B^TPx = Kx,
\]

(4.13)

which is a constant state feedback gain. Using this gain in the feedback scheme of Figure 4.1 defines the system

\[
\begin{align*}
\dot{x} &= (A - BK)x \\
y &= Cx.
\end{align*}
\]

(4.14)

(4.15)

The poles of this closed-loop system are equal to the stable poles of the Hamiltonian system \( S \). The unstable poles of \( S \) are the images of these poles in the \( \text{Im-axis} \).

### 4.3 Symmetric Root Locus

The LQR controller derived in the previous section can be used to compute a state feedback gain which is optimal for a certain criterion. In this section the Symmetric Root Locus (SRL) equation is derived which can be used to compute the optimal closed-loop pole locations for a given system for the simplified cost criterion

\[
J = \int [\rho y^2 + u^2]dt,
\]

(4.16)

which is parameterized by the single parameter \( \rho \). This criterion follows from Equation 4.1 by taking \( R = 1 \) and \( Q = \rho \). The SRL equation can now be derived.
Theorem 4.1. (Symmetric Root Locus) The state feedback gain $K$ that is optimal in the sense of Equation 4.16, places the closed-loop poles $p_i$ of a system $H$ at the stable roots of the SRL equation

$$p_i(s) = \{ s \mid 1 + \rho H^T(-s)H(s) = 0, \ Re(s) < 0 \} , \quad (4.17)$$

where $H(s) = C(sI - A)^{-1}B$ is the open-loop transfer function.

Proof. Transform the system dynamics and Euler-Lagrange Equations 4.5 and 4.6 to the s-domain,

$$sx = Ax + Bu$$
$$s\lambda + A^T\lambda = -Qx$$
$$u = -R^{-1}B^T\lambda = Kx.$$ (4.19)

Eliminating $x$ and $\lambda$ results in

$$u = B^T(-sI + A^T)^{-1}C^TC(sI - A)^{-1}Bu \quad (4.21)$$

$$\Rightarrow u = -\rho H^T(-s)H(s)u$$

$$\Rightarrow 0 = [1 + \rho H^T(-s)H(s)]u$$

$$\Rightarrow 0 = 1 + \rho H^T(-s)H(s),$$ (4.24)

where the last equation holds because Equation 4.23 should hold for all $u$. \hfill \Box

By investigating the SRL of the open-loop of a system, the closed-loop pole locations as function of $\rho$ are visualized. Notice that Equation 4.17 is similar to a conventional root-locus equation. Like the conventional root locus, the SRL is plotted for all values of $\rho$ to show the influence of $\rho$ on the pole locations. $\rho$ indicates the relative importance of control cost versus performance.

A large $\rho$ indicates the cheap control case, because performance is relatively more important than control cost. Suppose the system $H(s)$ can be partitioned in a numerator $N(s)$ and denominator $D(s)$, so $H(s) = \frac{N(s)}{D(s)}$. For the extreme case $\rho \rightarrow \infty$ Equation 4.17 yields

$$H(-s)H(s) = \frac{N(-s)N(s)}{D(-s)D(s)} = \frac{-1}{\rho} \rightarrow 0,$$ (4.25)

which means that the closed-loop poles go to the zeros of the system. If there are more poles than zeros in the system, some poles will go to infinity in the SRL. Physical systems never have more zeros than poles, since systems with more zeros than poles are non proper and therefore a-causal and cannot represent a physical system.

Conversely, a small $\rho$ will cause the control cost to contribute more to the criterion, resulting in lower control signals, which is desirable if the cost of control is high. Consider the extreme case, where $\rho \rightarrow 0$ i.e. the high control cost case, for which Equation 4.17 yields

$$H(-s)H(s) = \frac{N(-s)N(s)}{D(-s)D(s)} = \frac{-1}{\rho} \rightarrow \infty.$$ (4.26)

this for $\rho = 0$ the SRL points are the open-loop poles.

When plotting this root-locus, it has to be determined whether the positive or negative root-locus has to be drawn. Because $\rho$ is taken positive, the sign of $H^T(-s)H(s)$ determines this. The following lemma can be used to determine which locus to draw.

Lemma 4.2. Let $n$ and $m$ denote the degree of the numerator $N(s)$ and denominator $D(s)$ of a transfer function $H(s) = \frac{N(s)}{D(s)}$ respectively. To draw the correct SRL, draw the $0^\circ$ (180$^\circ$) degree root locus if $n - m$ is odd (even).

Proof. The SRL equation contains $H^T(-s)H(s) = \frac{NT(-s)N(s)}{D^T(-s)D(s)}$. Consider the following two examples:

$N(s) = (s + 1) \Rightarrow N(-s)N(s) = (-s + 1)(s + 1) = -s^2 + \ldots$
$N(s) = s^2 + s + 1 \Rightarrow N(-s)N(s) = ((-s)^2 - s + 1)(s^2 + s + 1) = s^4 + \ldots$

This shows that an even degree results in a positive numerator and an odd degree in a negative numerator. Same holds for the degree of the denominator, of course. When $n - m$ is even, the numerator and denominator are either both even, or both odd. Therefore either both numerator and denominator have a minus sign, which cancel each other, or both have no minus sign. Thus $H(s)$ is positive for large $s$ in this case. If $n - m$ is odd, either the numerator or denominator are odd and therefore negative, resulting in a negative $H(s)$ for large $s$. \hfill \Box

37
4.4 Data-based symmetric root locus

4.4.1 Introduction

Equation 4.17, used to describe the SRL, is a function of the open-loop transfer function of a system $H(s)$. As explained in Chapter 1, $H(s_i)$ can be computed from a Frequency Response Function (FRF) measurement of the system $H(j\omega)$, provided that the system is lightly damped. Thus the SRL of a system can be computed data-based by rewriting Equation 4.17:

$$1 + \rho H(-s_i)H(s_i) = 0$$

(4.27)

$$H(-s_i)H(s_i) = -\frac{1}{\rho}$$

(4.28)

$$H(s_i)H(s_i) = -\frac{1}{\rho}$$

(4.29)

where the property from Section 2.3, $H(s) = H(-s)$ is used for the last step. Thus, the SRL can be computed from $H(s_i)H(s_i)$ by searching for points with phase $-180^\circ$ ($-\rho$ is a negative real number). Note that this is equivalent to searching for points of $H(s_i)$ with phase $-90^\circ$. The corresponding value of $\rho$ of the SRL can also be reconstructed by rewriting Equation 4.17:

$$\rho = -\frac{1}{H(s_i)H(s_i)}$$

(4.30)

Recall that the gain $\rho$ is not a conventional root-locus gain, but corresponds to the weighing of the output relative to the input in the optimal control criterion 4.16.

In the following sections the SRL will be computed for the two mass system using only the FRD $H(j\omega_i)$ of the system. First $H(s_i)$ is computed from $H(j\omega_i)$ using Cauchy’s formula, see Equation 2.16. Next the SRL is computed and compared with a SRL that is calculated using the model of the system. Finally an optimal controller is designed using the LQR method and the closed-loop pole locations are compared with the SRL. The last part uses the model of the system, because it is not the scope of this chapter to derive the optimal controller directly from TFD $H(s_i)$.

4.4.2 Example data-based SRL

Consider a fourth order mass-spring-damper system, as shown in Figure 4.2. All coefficients are chosen positive such that the system is stable. The masses are $m_1 = m_2 = 1$ kg, the springs are $k_1 = k_2 = 1000$ N/m and the dampers have $d_1 = d_2 = 1$ Ns/m. The control objective is to keep the position $x_1$ of the first mass as small as possible using the input $F_1$ acting on the first mass. Again, the state feedback controller structure as shown in Figure 4.1 is used.

The frequency response of the transfer from $F_1$ to $x_1$ is shown in Figure 4.3. The SRL can be computed data-based from the frequency response by performing the following steps:

1. Compute the data-based transfer function $H(s_i)$ from the FRF data
2. Search for points $s$ where $\angle H(s)H(s) = -180^\circ$, these points form the SRL
3. Reconstruct $\rho$ using Equation 4.30
4.4.3 Computation $H(s_i)$ from $H(j\omega_i)$

In Chapter 2 it was explained that for lightly damped systems $H(s_i)$ can be computed from $H(j\omega_i)$ using Cauchy’s equation,

$$H(s_i) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{H(j\omega)}{(j\omega - s)} \, d\omega.$$  \hfill (4.31)

The Cauchy equation will now be used to compute $H(s_i)$ for a grid of points $s_i$ in the s-plane, see Figure 4.4. Figure 4.4a shows the grid in the s-plane. Each line $s_i = \sigma + j\omega$ is plotted for the same values for $\omega$ which range from 0 to 100 but a different value of $\sigma$ per line, which ranges from 0.5 until 10. Note that a blue line-color corresponds to lines close to the imaginary axis, i.e. small $\sigma$, while an orange line-color corresponds to a line farther from the imaginary axis, i.e. a large $\sigma$. Figure 4.4b shows the corresponding plots of $H(s_i)$. The colors of the lines in Figure 4.4b have the same meaning as in Figure 4.4a. This can also be derived from the properties of the system. Because the poles of the system lie close to the imaginary axis, lines evaluated farther from the poles show more damped behavior than the lines close to the poles. The 3D-Bode diagram visualizes this very clearly. A 3D-Bode diagram can be used to visualize the magnitude and phase behavior on the complex frequency $s_i = \sigma + j\omega$. The plot is three dimensional because the dependency on two variables, $\omega$ and $\sigma$, is visualized. The magnitude of $H(s_i)$ of Figure 4.4b is plotted as a function of $\text{Re}(s_i)$ and $\text{Im}(s_i)$ in Figure 4.5.

4.4.4 Computing the SRL

Now that $H(s_i)$ is computed, the points that lie on the SRL can be computed by searching for points where $H(s_i)H(s_i^*)$ has phase $-180^\circ$. Numerically this is done by searching for points $s_i$ that have a phase smaller than $-180^\circ$, and have at least one neighboring point $s_{i+1}$ or $s_{i-1}$ that has a phase greater than $-180^\circ$. Figure 4.6 shows the frequency response of $H(s_i)H(s_i^*)$. For small $\sigma$ values (blue line) there are three points where the phase is $-180^\circ$. For large values of $\sigma$, however, there is only one point where the phase is $-180^\circ$.

The points in the s-plane with phase $-180^\circ$ found in this way are plotted with a (*) in Figure 4.7. As expected from the discussion on Figure 4.6 there are three points at small $\sigma$ values and only one point at large $\sigma$ values. The corresponding symmetric root-locus gains $\rho$, computed from Equation 4.30, are shown above these points. As $\rho$ increases, the closed-loop poles travel from the open-loop poles to the zeros or to infinity, as expected. The SRL computed from the model of the system is plotted as a solid (blue) line.
Figure 4.4: Grid in the s-plane (a) and $H(s_j)$ computed with Cauchy (b).

Figure 4.5: 3D-Bode magnitude diagram of the two mass system, computed with Cauchy.
Figure 4.6: Frequency response of $H(s_i)H(s_i)$. Points where the phase is $-180^\circ$ are points on the SRL.

It can be observed that the closed-loop pole locations computed from the FRF $H(j\omega)$ lie on the SRL computed from the model.

The data-based SRL does not match the model when points are calculated that lie closer to the imaginary axis than the open-loop poles. Points computed at $\sigma = 0.5$ and $\sigma = 1$ should not be part of the SRL, but the computation does include them in the SRL. The explanation for this is that the assumption that was made, $H(s) = H(-s)$, does not hold for these locations.

The gain of the SRL computed with this data-based approach is compared to the theoretical SRL gain computed with the model of the system, using equation 4.17. Figures 4.8a and 4.8b show the gain the upper and lower branch of the first quadrant of the SRL of Figure 4.7. The gain is plotted as a function of the points that are computed on the SRL, starting at the pole and traveling towards the zero, or towards infinity. The lower branch has a better match with the model than the upper branch. This could be expected because the pole and zero of the lower branch have less damping and therefore correspond better to the assumption that is made, namely that the system is undamped. In Figure 4.8b it can be seen that the gain computed from the model has a minimum at point 3, indicating that a pole is passed. It is not visible in the plot due to the logarithmic scale, but the gains corresponding to the first two points are actually negative, and therefore do not belong to the SRL.

### 4.4.5 SRL and optimal control

In this section a state feedback controller will be derived using the model of the system to show the connection between the SRL and optimal control. The closed-loop poles of the controlled system will then be compared to the SRL. Suppose the objective of the system is to maximize the damping of the low frequent pole. By looking closer at Figure 4.7, it is found that the $\rho$ corresponding to the point that has the largest real part is ca. $\rho = 3.15 \times 10^6$. Computing now the optimal control using the criterion

$$J = \int [3.15E\dot{y}^2 + u^2]dt,$$

(4.32)

gives the state feedback law

$$u = -Kx$$

(4.33)

$$= - \begin{bmatrix} 822 & 215 & 39 & 17 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix}$$

(4.34)
Figure 4.7: Symmetric root locus from data-based approach (*) and model (-).

Figure 4.8: SRL gains for the upper and lower branch from model (-) and data-based (*) approach.
Computing the poles and zeros of the controlled system

\[
\begin{align*}
\dot{x} &= (A - BK)x \\
y &= Cx,
\end{align*}
\]

(4.35) (4.36)

gives

\[
\begin{align*}
p_1 &= -14.0 \pm 50.5j \\
p_2 &= -7.0 \pm 26.3j \\
z_1 &= -0.5 \pm 31.6j.
\end{align*}
\]

(4.37) (4.38) (4.39)

Figure 4.9 shows the data-based SRL and the computed optimal pole locations for the chosen value of \( \rho \). Of course, the poles lie in the LHP, because the optimal control always results in a stable system. Indeed the optimal damping for the low frequent pole is achieved. The improvement of the damping can be seen when comparing the impulse response of the original system with the impulse response of the closed-loop system, see Figure 4.10. This concludes the example.

4.5 Energy in the system

It is insightful to consider the energy in a system that is controlled with an LQR controller. Useful relations between the value of the criterion \( J \), Lyapunov functions and the Riccati equation are derived in this section. It will also be shown that the energy in the time domain is equal to the energy in the frequency domain, using Parseval's theorem.

The criterion that is used to derive the optimal state-feedback controller

\[
J = \int_0^\infty [x(t)^T C^T Q C x(t) + u(t)^T R u(t)] dt
\]

(4.40)
can be interpreted as an energy function. It represents the total energy of the system when it is released from an initial condition \( x_0 \). The remaining energy in the system from an arbitrary time \( t_1 \) until infinity
can be computed by changing the integral bounds

\[ J(t_1) = \int_{t_1}^{\infty} [x(t)^T C^T QCx(t) + u(t)^T Ru(t)] dt. \] (4.41)

The energy in the system can also be computed by considering the Lyapunov function

\[ V(t) = x(t)^T Px(t) \] (4.42)

where \( P \) is a positive definite matrix. For this function it holds that

\[ \dot{V}(t) > 0 \quad \forall x \neq 0 \] (4.43)
\[ \dot{V}(t) \leq 0 \] (4.44)

Equations 4.40 and 4.42 are related to each other. In the optimal controller synthesis, the derivative of the Lyapunov function is not only smaller than zero, but is less than or equal to a certain function of the state \( x \) and input \( u \) [6]. Suppose there are matrices \( P, Q \) and \( R \) such that the following inequality is satisfied

\[ \dot{V}(t) \leq -[x(t)^T C^T QCx(t) + u(t)^T Ru(t)] \leq 0. \] (4.45)

This means that the energy in the system decreases with at least a certain rate. Integrating this expression yields

\[ V(\infty) - V(t_1) \leq -\int_{t_1}^{\infty} [x(t)^T C^T QCx(t) + u(t)^T Ru(t)] dt \] (4.46)

which clarifies the energy interpretation of the cost function \( J \). In the case that the system goes to equilibrium it holds that \( V(\infty) = 0 \). Then Equation 4.46 is equal to Equation 4.40. When \( P \) is the solution of the algebraic Riccati equation

\[ PA + A^T P - PBR^{-1}B^T P + C^T QC = 0, \] (4.47)

the derivative of the Lyapunov function is equal to the integrand of the criterion which can be derived by substitution of the system equations and the optimal control law

\[ \dot{x}(t) = Ax(t) + Bu(t) \] (4.48)
\[ K = R^{-1}B^T P \] (4.49)

and the Riccati equation in Equation 4.45, yielding

\[ \dot{V} \leq -[x^T C^T QCx + u^T Ru] \] (4.50)
\[ \dot{x}^T Px + x^T P \dot{x} \leq -[x^T C^T QCx + x^T K^T RKx] \] (4.51)
\[ x^T (A^T - K^T B^T)Px + x^T P(A - BK)x \leq -[x^T (C^T QC + K^T RK)x] \] (4.52)
\[ x^T (A^T P + PA - 2PBR^{-1}B^T P)x = -[x^T (-A^T P - PA + 2PBR^{-1}B^T P)x]. \] (4.53)
The dependence on time is omitted for clarity of the notation. The last equation shows that the \( \leq \) sign indeed becomes an \( = \) sign.

There is a third way to compute the energy in the system. The integral expression for the cost criterion, Equation 4.40, is an integral in the time domain. This integral can also be computed in the frequency domain by using Parseval’s theorem,

\[
\int_0^\infty |f(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(j\omega)|^2 d\omega
\]

which states that the integral of the square of a function \( f(t) \) in the time domain is equal to the integral of its transform \( F(s) \) in the frequency domain [16]. This relation holds because the time domain signal and the frequency domain signal must contain the same amount of energy. Applying Parseval’s theorem to the cost function yields

\[
J(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [X(j\omega)^T C^T Q C X(j\omega) + U(j\omega)^T R U(j\omega)] d\omega,
\]

with

\[
X(j\omega) = \int_{-\infty}^{\infty} e^{-j\omega \tau} x(\tau) d\tau
\]

\[
U(j\omega) = \int_{-\infty}^{\infty} e^{-j\omega \tau} u(\tau) d\tau.
\]

Three different ways to compute the energy in the system have been discussed. They can now be applied to the two-mass system under consideration. The system is given an initial condition

\[
x_0 = \begin{bmatrix} 6 & -4 & 7 & 18 \end{bmatrix} \times 10^{-2}
\]

and the response of the system is monitored. The response of the system is shown in Figure 4.11 on the left. The energy in the system is computed using Equations 4.41, 4.42 and 4.55. In the right part of the figure, the energy in the system is plotted for the three methods. The computed energy is of course the same for each method.

### 4.6 Optimal return difference

The SRL is a relationship between the open-loop transfer function of a system and the closed-loop pole locations. There exists also a more general relationship that connects the open-loop transfer function to the return difference equation. This relationship is named the optimal return difference relationship but is also known as the Kalman Frequency Domain Equality (KFDE). First, some notation has to be defined.
Consider again the state feedback scheme as depicted in Figure 4.1. The transfer function of the state space system is

$$H(s) = C(sI - A)^{-1}B$$  \hspace{1cm} (4.59)

The return difference equation, for which a derivation is given in Appendix B yields

$$T(s) = [I + K(sI - A)^{-1}B]$$  \hspace{1cm} (4.60)

Also needed are the hermitian transpose of the transfer function and the return difference equation as will become clear later.

$$H^T(-s) = B^T(-sI - A)^{-1}C^T$$  \hspace{1cm} (4.61)

$$T^T(-s) = [I + B^T(-sI - A^T)^{-1}K^T]$$  \hspace{1cm} (4.62)

Remember that $K$ is the optimal gain for the state feedback controller defined by

$$K = R^{-1}B^TP$$  \hspace{1cm} (4.63)

**Theorem 4.3.** (Optimal Return Difference Relationship) For the LQR system with the optimality criterion described in Equation 4.1 the optimal return difference relationship is

$$T^T(-s)RT(s) = H^T(-s)QH(s) + R$$  \hspace{1cm} (4.64)

with the notation described in Equations 4.59-4.63. The system $H(s)$ should at least be stabilizable and detectable.

**Proof.** [28] Start from the Algebraic Ricatti Equation

$$PA + A^TP - PBR^{-1}B^TP + C^TQC = 0.$$  \hspace{1cm} (4.65)

Add and subtract $sP$ gives

$$sP - sP + PA + A^TP - PBR^{-1}B^TP + C^TQC = 0$$  \hspace{1cm} (4.66)

$$-(sI - A^T)P - P(sI - A) - PBR^{-1}B^TP = -C^TQC.$$  \hspace{1cm} (4.67)

Premultiplying with $-B^T(-sI - A^T)^{-1}$ and $(sI - A)^{-1}B$ and adding $R$ to both sides yields

$$R + B^T(sI - A)^{-1}B + B^T(-sI - A^T)^{-1}PB$$

$$+ B^T(-sI - A^T)^{-1}C^TQC(sI - A)^{-1}B$$

$$= R + B^T(-sI - A^T)^{-1}C^TQC(sI - A)^{-1}B.$$  \hspace{1cm} (4.68)

Factorizing the left hand side gives

$$[I + B^T(-sI - A^T)K^T][I + K(sI - A)^{-1}B]$$

$$= R + B^T(-sI - A^T)^{-1}C^TQC(sI - A)^{-1}B.$$  \hspace{1cm} (4.69)

Using the definitions described in Equations 4.59-4.63 gives

$$T^T(-s)RT(s) = H^T(-s)QH(s) + R$$  \hspace{1cm} (4.70)

The optimal return difference is the frequency domain equivalent of the Riccati equation. It is an important relation since it will be used in the next chapter to derive an optimal controller. Note that the solutions of the return difference $T(s)$, Equation 4.60, are the closed loop poles. Therefore, the optimal return difference equation, Equation 4.64, gives a relation between the open-loop and closed-loop poles for an optimal controller $K$. The optimal return difference equation is in fact a generalization of the SRL equation that was discussed earlier. Because it holds for all $s$, it also holds for the special case that $s$ is a closed-loop pole location. In this case the return difference is zero which results in

$$0 = pH^T(-s)H(s) + 1,$$  \hspace{1cm} (4.71)

where $R = 1$ and $Q = p$ are chosen as before. This equation is equal to the SRL Equation 4.17. This proves once more that the closed-loop pole locations can be found by solving this equation as demonstrated in the data-based example of Section 4.4.
There is also an application for the optimal return difference equation itself. As demonstrated by the famous article of John C. Doyle [17], there are no guaranteed stability margins when applying LQG control. There is however a technique called Loop Transfer Recovery (LTR) which can retrieve the stability margins of a LQR controller partly [19]. To achieve the recovery, some of the estimator poles are placed at (or near) the zeros of the plant and the remaining poles are moved into the left half plane. The idea is to redesign the estimator such that the loop gain approximates the loop gain corresponding to LQR control of the system under consideration. This is done in two steps:

1. Target loop $L$ design:
   Design the target loop $L$, given by
   \[
   L = K (sI - A)^{-1} B = H_{LQ}
   \]  
   (4.72)
   The optimal return difference equation aids in the shaping of the open-loop $H_{OL}$, because information on the resulting closed-loop can be derived from
   \[
   [I + H_{LQ}(j\omega)]H[I + H_{LQ}(j\omega)] = I + \left[ \frac{1}{\sqrt{\rho}} H_{OL}(j\omega) \right]^H \frac{1}{\sqrt{\rho}} H_{OL}(j\omega)
   \]  
   (4.73)

2. Target loop recovery:
   Design a model-based estimator that recovers the designed $L$.

So the optimal return difference gives insight in the closed-loop transfer function from the open-loop transfer function. Here only the link to the optimal return difference is clarified, further details on the LTR technique can be found in [33] or [19] for example.
Applications of Transfer Function Data
Chapter 5

Data-based optimal control

5.1 Introduction

The data-based SRL of the previous chapter can be used to compute the optimal closed-loop pole locations given a certain system and criterion. This chapter describes a method to compute an optimal controller data-based, that achieves the optimal closed-loop pole locations of the SRL. Because a method to compute TFD $H(s_i)$ is available, see Chapter 2, a search in literature for an optimal controller description in terms of $H(s)$ is performed.

In literature, some time domain results are known. In [1] a method to compute a data-based observer and optimal controller using Markov parameters of the system is described. Other time domain results for example are [35], which also uses Markov parameters and [31] which uses Cholesky factorizations. However, these results are less useful in our context, because they do not have a controller description in terms of $H(s)$.

There are also frequency domain approaches to the optimal control problem. These approaches use the frequency domain version of the Riccati equation, which is the optimal return difference equation discussed in the previous chapter. Spectral factorization is used in [3] to compute the optimal gains of a state feedback controller for a simply supported beam. But the most promising result that was encountered was found in “The frequency domain solution of regulator problems”[45]. The following part of this article (using our own notation in the equation for consistency) drew immediate attention;

\[
\ldots \text{then the optimal control law is}\ldots \\
C(s) = \left[ (I + \rho H(s) \tilde{H}(s))^{+} - I \right] H(s)^{-1}
\]

which is an optimal controller specified completely in terms of $H(s)$ and the optimality parameter $\rho$. Therefore it is possible to compute this controller using TFD $H(s_i)$. In this chapter the derivation, data-based computation and data-based application of this controller are discussed.

5.2 Derivation of the optimal controller

In [45], the optimal controller is derived from a comparison between a Kalman optimal control problem and a Wiener filtering problem. An alternative derivation is presented here, which uses the optimal return difference that was derived in Chapter 4. Using our choice $R = 1$ and $Q = \rho$, the optimal return difference equation yields

\[
\bar{T}T = \rho \bar{T}H + 1,
\]

where the return difference is now defined as

\[
T = 1 + C(s)H(s).
\]

This form of the return difference is more general than the return difference defined in Chapter 4.17. This is required since the controller that achieves this optimal return difference will be derived from $T(s)$. The optimal return difference minimizes the cost criterion

\[
J = \int_{0}^{\infty} \rho x^T C^T C x + u^T u \ dt.
\]
Applications of Transfer Function Data

for the system $H(s)$ described by

$$
\dot{x} = Ax + Bu \\
y = Cx
$$

(5.5)

It will become clear that by extracting $T(s)$ from Equation 5.2, it is possible to compute an optimal controller. To compute the optimal controller, a spectral factorization is necessary. The spectral factorization is defined as follows.

**Definition 5.1.** (Spectral factorization) Let $Z(s)$ denote a transfer function. Suppose $Z(s)$ can be written as

$$
Z(s) = Y(s)Y(-s).
$$

(5.6)

Let $Y(s)$ is a transfer function that has all its poles and zeros in the LHP and consequently, $Y(-s)$ is the same function but it has all its poles and zeros in the RHP, which are the mirror images of the poles and zeros of $Y(s)$. Then

$$
(Z(s))^+ = Y(s) \quad \text{and} \quad (Z(s))^− = Y(-s).
$$

(5.7)

The following theorem gives the derivation of the optimal controller.

**Theorem 5.2.** The controller that minimizes the cost criterion of Equation 5.4 is given by

$$
C(s) = ((ρHH + 1)^+ - 1)H^{−1}
$$

(5.8)

**Proof.** Since it is required that $T(s)$ leads to a stable, minimum-phase closed-loop system, it must have all its poles and zeros in the LHP. Therefore, $T(s)$ can be derived from a spectral factorization of the optimal return difference. Since

$$
TT = ρHH + 1 = (ρHH + 1)^−(ρHH + 1)^+
$$

(5.9)

it must hold that

$$
T = (ρHH + 1)^+
$$

(5.10)

Because the return difference is defined by $T = 1 + CH$, the controller can be computed from

$$
C(s) = (T − 1)H^{−1} = ((ρHH + 1)^+ − 1)H^{−1}
$$

(5.11)

\[\square\]

Notice that this controller is specified completely in terms of $H(s)$ and can therefore be evaluated using TFD $H(s)$. This will be done in Section 5.5, but first the interpretation of this controller and the comparison to conventional optimal controllers will be given in the next two sections.

### 5.3 Interpretation of the optimal controller

This section discusses the interpretation of the optimal controller. It will become clear that the optimal controller can be divided in two parts; an observer and a state-feedback controller. In order to make this interpretation, it is necessary to rewrite Equation 5.8 which requires the introduction of some notation. The transfer function of a system $H(s)$ can be written as

$$
H(s) = \frac{q(s)}{p(s)} = \frac{\sum_{i=0}^{m} q_i s^i}{\sum_{i=0}^{n} p_i s^i}, \quad m < n
$$

(5.12)

where $q(s)$ and $p(s)$ denote the numerator and denominator polynomial of $H(s)$. Then define the following spectral factor

$$
γ(s) = (p^o + ρq)^+.
$$

(5.13)
using $\gamma(s)$, $p(s)$ and $q(s)$, it is possible to rewrite the optimal controller as

$$C(s) = \left( (\rho HH + 1)^+ - 1 \right) H^{-1}$$

$$= \left( \left( \frac{\bar{p}p + \rho \bar{q}q}{\bar{p}p} \right)^+ - 1 \right) \frac{P}{q}$$

$$= \left( \frac{2}{p} - 1 \right) \frac{P}{q}$$

$$= \frac{\gamma(s) - p(s)}{q(s)}$$

(5.17)

where the dependency on $s$ is omitted in the intermediate steps. It will be proven in Proposition 5.3 that the $\frac{1}{q(s)}$ part of the optimal controller of Equation 5.17 is an observer. Proposition 5.5 will show that the $\gamma(s) - p(s)$ part is a state feedback controller that places the poles at locations that are optimal for the criterion of Equation 5.4. Figure 5.1 visualizes this division of the optimal controller in two parts.

**Proposition 5.3.** The $\frac{1}{q(s)}$ part of Equation 5.17 is an observer for the states of the system $H(s)$.

**Proof.** Assume that the system can be transformed using a similarity transformation to controllable canonical form

$$A = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
-p_0 & -p_{n-2} & -p_{n-1} & 0 & 1
\end{bmatrix}$$

$$B = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}^T$$

$$C = \begin{bmatrix} q_0 & \cdots & q_{n-1} \end{bmatrix}$$

The states of this special representation satisfy

$$x_2 = \frac{d}{dt} x_1$$

$$x_3 = \frac{d}{dt} x_2 = \frac{d^2}{dt^2} x_1$$

$$\vdots$$

$$x_n = \frac{d}{dt} x_{n-1} = \frac{d^{n-1}}{dt^{n-1}} x_1$$

Thus all states can be represented as derivatives of the first state $x_1$. The output $y(t)$ is computed from

$$y(t) = C x(t)$$

$$= \begin{bmatrix} q_0 x_1 & q_1 x_2 & \cdots & q_{n-1} x_n \end{bmatrix}$$

$$= \begin{bmatrix} q_0 & q_1 \frac{d}{dt} & \cdots & q_{n-1} \frac{d^{n-1}}{dt^{n-1}} \end{bmatrix} x_1,$$
which uses the special structure of the states of the canonical representation. Transforming to the Laplace domain yields

\[ Y(s) = [q_0 \ q_1s \ \cdots \ q_{n-1}s^{n-1}]X_1(s) \]

\[ = q(s)X_1(s) \]  

(5.21)

Therefore the state \( X_1(s) \) can be computed from

\[ X_1(s) = \frac{Y(s)}{q(s)} \]  

(5.22)

which shows that the \( \frac{1}{q(s)} \) part of Equation 5.17 is an observer for the state \( X_1(s) \) of the system \( H(s) \). Because of the canonical form, the information of all other states is contained in \( X_1 \), since they are equal to the derivatives of \( X_1 \).

The observer part of the optimal controller is clear now. Before the state feedback part \( \gamma(s) - p(s) \) can be described, first some important properties of \( \gamma \) are described in the following lemma.

**Lemma 5.4.** Let \( H(s) \) denote a transfer function which can be written in terms of its numerator and denominator polynomial, \( H(s) = q(s)/p(s) \). Consider the controller \( C = (\gamma - p)/q \), where \( \gamma = (pq + qpq)^+ \).

In this case it holds that:

1. \( \gamma(s) \) is the closed-loop characteristic polynomial
2. The roots of \( \gamma(s) \) are the optimal closed-loop pole locations given the cost function of Equation 5.4.

**Proof.**

1) Substitution of the controller and system in the denominator of closed loop transfer function gives

\[ \frac{1}{1 + CH} = \frac{1}{1 + \frac{\gamma - p}{p}} = \frac{p}{\gamma} \]  

(5.23)

which shows that \( \gamma \) is the closed-loop characteristic polynomial.

2) The zeros of \( \gamma \) are equal to the LHP solutions of the symmetric root locus equation. This can be seen by considering

\[ (1 + \rho H)\gamma = (pq + qpq)^+ = \frac{\gamma}{p} \]  

(5.24)

from which it is clear that the roots of \( \gamma \) are the solutions of the SRL. Theorem 4.1 proves that the solutions of the SRL are the optimal closed loop poles. Therefore the roots of \( \gamma \) correspond to the optimal closed-loop poles.

**Proposition 5.5.** The \( \gamma(s) - p(s) \) part of the optimal controller of Equation 5.17 places the closed-loop poles at the locations that are optimal for the cost criterion of Equation 5.4.

**Proof.** From Lemma 5.4 it is clear that \( \gamma \) is the desired closed-loop characteristic polynomial that is optimal for the cost criterion of Equation 5.4. This is exactly the closed-loop polynomial that the state feedback controller \( \gamma - p \) achieves, when it is applied to the state \( X_1(s) \). Utilizing the special structure of the states in canonical form of Equation 5.18 again gives

\[ u(s) = [\gamma(s) - p(s)]\hat{X}(s) \]  

(5.25)

\[ = [\gamma_0 + \gamma_1s + \gamma_2s^2 + \cdots + \gamma_{n-1}s^{n-1} - p_0 - p_1s - p_2s^2 - \cdots - p_{n-1}s^{n-1}]\hat{X}(s) \]

\[ = [(\gamma_0 - p_0) + (\gamma_1 - p_1)s + (\gamma_2 - p_2)s^2 + \cdots + (\gamma_{n-1} - p_{n-1})s^{n-1}]\hat{X}(s) \]

\[ = k_0x_1 + k_1x_2 + k_2x_3 + \cdots + k_{n-1}x_n \]

This is equivalent to a standard pole placement algorithm as found in [19] for example. In such a pole placement algorithm, the gains for each state are equal to the difference between the coefficients of the characteristic equation of the system \( p_i \) and the coefficients of the desired closed-loop characteristic equation \( \gamma_i \), thus in this case

\[ k_i = \gamma_i - p_i \]

(5.26)

which completes the proof.
Summarizing, the optimal controller can be interpreted as an observer that estimates $X_1(s)$ from which all states can be computed by differentiating, combined with a state feedback controller that places the poles at the optimal locations, given by the roots of $\gamma(s)$.

5.4 The optimal controller versus LQG

It was shown in the previous section that the $`\gamma` − p$ part of the controller achieves pole placing at optimal pole locations. Therefore it is equivalent to a standard linear quadratic regulator (LQR). The complete optimal regulator solution also contains a state estimation part. Therefore the question arises whether the optimal controller is equivalent to a linear quadratic Gaussian (LQG) controller, because it also consists of a state estimation and a state feedback part. The system description for the LQG problem is

\[
\dot{x} = Ax + Bu + w \\
y = Cx + v
\]

where $w$ denotes the process noise and $v$ the measurement noise. To estimate the state $x$ of this system, a Kalman filter is used in an LQG controller. A Kalman filter uses a simulation of a model of the real system that has the same inputs as the real system to compute the state vector of the system. The error between the simulation and the real system is minimized by feeding back the error with a certain estimator gain $L$. This gain is computed by solving the following (steady-state) Riccati equation,

\[
0 = PA^T + AP + R_v - PC^T R_w^{-1} CP,
\]

where $R_v$ is the covariance of the measurement noise and $R_w$ the covariance of the process noise. The optimal estimator gain is

\[
L = PC^T R_w^{-1}.
\]

It is known that a technique called loop transfer recovery (LTR) can be used to approximates the properties of an LQR controller for an LQG controller. It does this by redesigning the estimator, such that the poles of the estimator approach the zeros of the plant. The covariances of the measurement and process noise are used as design parameters to shape the estimator such that the resulting controller approximates an LQR controller. The motivation to do so is that LQR controllers have good margins, while the margins of an LQG controller can be very poor. The covariances are chosen

\[
R_v = 1 \\
R_w = \alpha BB^T
\]

where $\alpha$ is a scalar. For $\alpha \rightarrow \infty$, the poles of the estimator approach the zeros of the plant and the resulting LQG controller approximates an LQR controller [18]. The tradeoff is that the noise suppression abilities of the estimator decrease because $R_v$ is relatively small compared to $R_w$. Therefore the estimator will put more effort in the state estimation than in the noise filtering. The estimator will rely too much on the measurement which will deteriorate the state estimation.

Both the LQG regulator in combination with the LTR technique and the optimal controller have the estimator poles at the zeros of the plant. They also both use state feedback to achieve certain optimal pole locations. Thus they are expected to be equivalent when the problem description is the same. A simulation is performed to illustrate this observation.

The optimal controller and the LQG/LTR controller are computed for the two mass system. The LQG/LTR controller is computed for three different values of $\alpha$. The FRFs of the resulting controllers look similar at low frequencies, see Figure 5.2. The higher the value of $\alpha$, the more the LQG controller resembles the optimal controller at higher frequencies. But due to the higher gain at high frequencies, the noise filtering properties of these controllers will be less good. This was expected since the covariance $R_v$ is assumed to be low.

5.5 Data-based computation of the optimal controller

The derivation and the principles behind the optimal controller are clear now. The optimal controller derived in Theorem 5.2 is described in terms of $H(s)$ and can therefore be evaluated data-based by using TFD $H(s_i)$. The only difficulty is how to compute the spectral decomposition in a data-based way. This will be discussed in Section 5.5.1. The selection of a suitable value for the optimality parameter $\rho$ is the topic of Section 5.5.2.
5.5.1 Data-based spectral factorization

The problem of computing the controller derived in Theorem 5.2, is how to compute the spectral factor

$$(1 + \rho H \bar{H})^+$$  \hspace{1cm} (5.33)

using only the FRD of the system. Recall that the spectral decomposition separates the LHP and RHP terms of the transfer function. The function

$$(1 + \rho H \bar{H}) = (1 + \rho H \bar{H})^+(1 + \rho H \bar{H})^-$$  \hspace{1cm} (5.34)

is the product of both spectral factors, which can be computed from the FRD directly. The amplitude of Equation 5.33 can be computed from this function. Since the amplitude of the LHP factor is equal to the amplitude of the RHP factor, the amplitude of Equation 5.34 is the square of the amplitude of Equation 5.33, which is the function that has to be computed. Thus

$$| (1 + \rho H \bar{H})^+ | = \sqrt{| (1 + \rho H \bar{H}) |}$$  \hspace{1cm} (5.35)

gives the amplitude of the spectral factorization, but the phase is still not known. The main idea to obtain the phase of the stable spectral factor is the following. Since Equation 5.33 is the LHP factor it must be a stable minimum phase transfer function. For stable minimum phase systems, the Bode gain-phase relationship can be used to compute the phase for a given gain behavior [4]. Implementations of the Bode gain-phase relationship can be found in [8], for example. Appendix C provides more information on phase reconstruction and its relation to the Hilbert transformation. The formula to compute the phase in this case is

$$\angle [1 + \rho H(s) \bar{H}(-s)]^+ = \frac{2s}{\pi} \int_0^{\infty} \frac{\alpha(y) - \alpha(s)}{y^2 - s^2} dy,$$  \hspace{1cm} (5.36)

where

$$\alpha = \log \left( \sqrt{|(1 + \rho H \bar{H})|} \right)$$  \hspace{1cm} (5.37)

and $y$ defines a line in the complex plane along which the phase has to be reconstructed. Figure 5.3 shows an example of the computation of the LHP spectral factor for the two mass system, with $\rho = 1.4 \times 10^9$ and

Figure 5.2: The controller can be divided in two parts.
Figure 5.3: Example of computing the spectral factor. The spectral factor computed from a model of the system (-) is approximated very well by the reconstructed spectral factor (*). Also shown is the amplitude of \((1 + \rho H \bar{H})\), from which the amplitude of the spectral factor is computed by taking its square root.

\[ s = j\omega \] In the figure the spectral factor computed from a model (-) and the spectral factor computed from the FRF (*) are shown. The amplitude is computed by taking the square root of the amplitude of \((1 + \rho H \bar{H})\), which is also shown (···) in the magnitude plot of the figure. This result enables the complete data-based computation of the optimal controller of Equation 5.8 by substituting \(H(s_i)\) for \(H(s)\), giving

\[ C(s_i) = [(\rho H(-s_i)H(s_i) + 1)^+ - 1][H(s_i)]^{-1} \] (5.38)

### 5.5.2 Choosing the optimality parameter \(\rho\)

The optimality parameter \(\rho\) in the optimal controller defines the importance of the performance relative to the importance of the costs that are made to achieve the performance. There is a tradeoff between the size of the output \(y\) and input \(u\). Using the method described above, it is possible to compute an optimal controller for given \(\rho\). But how should this parameter be chosen? Various methods to choose \(\rho\) are known from model-based optimal control theory. One way to choose it is to assess the criterion that has to be minimized

\[ J = \int_0^\infty y^TQy + u^TRu \, dt. \] (5.39)

Recall that \(R = 1\) and \(Q = \rho\) are chosen in our case. The more general criterion is mentioned here in order to discuss weighting selection methods known in literature. The difficulty is to know what a reasonable values of \(R\) and \(Q\) are. The inverse square method by Bryson [7] is an often used method. In this method, the maximum allowable deviations in the input and output are used to compute the weighting functions \(Q\) and \(R\). In our case \(R = 1\) and \(Q = \rho\), such that in this case Bryson’s rule yields

\[ \rho = \frac{u_{\text{max}}}{y_{\text{max}}}. \] (5.40)

The disadvantage of this technique is that if the response of the system is not satisfactory with this value of \(\rho\), there is no pragmatic way to improve the response.

Therefore, a different method was proposed by Thompson [43]. In this method the weighting matrices are parameterized such that they depend on a scalar parameter \(\rho\). Our choice for \(R\) and \(Q\) satisfies exactly
Applications of Transfer Function Data

Figure 5.4: Fourth order system consisting of two masses connected via springs and dampers.

this condition! The advantage of this choice is that the closed-loop pole locations can be visualized in a root-locus. This provides a graphical way of adjusting the optimality parameter such that the system has certain closed-loop pole locations. In our case, the optimal root-locus is the symmetric root locus that was encountered in the previous section,

\[ 1 + \rho H(-s)H(s) = 0. \]  \hspace{1cm} (5.41)

It was shown in the previous chapter that the symmetric root locus can be computed data-based. This provides a data-based graphical way of determining the closed-loop pole locations that will be achieved with the controller. Alternatively, if certain conditions on the closed-loop poles are known, the value of \( \rho \) that achieves this can be computed. This is the method that will be used in the next section to compute \( \rho \).

A concluding remark on the selection of \( \rho \) has to be made, namely that there are also model-based methods to choose the weighting functions. One example of a method that can be applied when a model of the system is available is called optimal eigenstructure assignment [20]. This method enables the design of the closed-loop poles and eigenvectors. The design of eigenvectors can be used to adjust the eigenmodes of the system. But since our approach focuses on data-based methods of optimal controller design, these weight selection methods are less relevant for this research.

5.6 Data-based application of the optimal controller

5.6.1 Introduction

In this section two examples of the data-based application of the optimal controller will be given. The proposed theory will be applied to the two mass system that is shown in Figure 5.4, for which a model was derived in Chapter 2. The input \( F_1 \) is used in both examples, but the first example uses the position of the first mass \( x_1 \) as measurement, while the position of the second mass \( x_2 \) is used in the second example. The theory is thus applied on a colocated and non-colocated SISO system. Only the FRF \( H(j\omega) \) of the systems are used to compute the optimal controllers, the method is completely data-based.

The analysis of the resulting controller will be done in a data-based way as much as possible too. A model of the system will be used to check the results and to compute time domain responses. Throughout this section, model-based results will be plotted with solid lines (−), while data-based results are plotted with (∗) consistently.

5.6.2 Example 1: colocated control

In this example an optimal controller \( C(s) \) is computed for the system shown in Figure 5.5. FRD \( H(j\omega) \)
Applications of Transfer Function Data

Figure 5.6: (a) SRL showing three different choices for $\rho$ based on closed-loop pole locations. (b) Resulting optimal controllers for the three choices of $\rho$. model-based (-) data-based (*).

is computed from the model of this system, where the force on the first mass $F_1$ is the input and the position of the first mass $x_1$ is the output, thus

$$H(j\omega) = \frac{x_1}{F_1}.$$  \hspace{1cm} (5.42)

The TFD is computed from FRD using the Cauchy integral.

**Computation of the controller $C(s)$**

The controller can be computed by evaluating Equation 5.38

$$C(s_i) = [(1 + \rho H(s_i)H(-s_i))^+ - 1]H(s_i)^{-1}. \hspace{1cm} (5.43)$$

The controller depends on the optimality parameter $\rho$, which determines the relative importance of the output versus the required input. As proposed earlier, the value of $\rho$ can be chosen using the SRL of the system, which can be computed using TFD. Figure 5.6a shows the SRL of one quadrant of the system, computed with the data-based method (*) and with the model of the system (-). The SRL computed with the data-based method has only small deviations from the SRL computed from the model of the system. In the other three quadrants that are not shown, the SRL is the mirror image with respect to the real and imaginary axis of this part of the SRL. The poles (×) and zeros (◦) of the system are plotted as thick black markers. Recall that the SRL shows the closed-loop pole locations as a function of $\rho$. The SRL makes it possible to compute the value of $\rho$ that belongs to a certain desired closed-loop pole position. Three desired closed-loop pole locations are chosen, by selecting three points on the SRL, which are indicated by the colored blue, green and red square behind the SRL points. From these points on the SRL, the corresponding value of $\rho$ can be computed using

$$\rho = \frac{-1}{H(s_i)H(s_i)}, \hspace{1cm} (5.44)$$

see Chapter 4. These values are shown in the figure too. For each color, the theoretical closed-loop pole locations (×) that belong to the computed value of $\rho$ are plotted in the corresponding color. In theory, these locations should be equal to the chosen points on the SRL. Differences between these points are caused by errors in the computation of the SRL. The assumption in the computation of the SRL is that the poles of the system have zero damping and thus lie on the imaginary axis. This is clearly not the case and therefore the actual pole locations are different from the selected SRL points. Note that the upper
Applications of Transfer Function Data

Figure 5.7: (a) Open loop for the three controllers (b) Sensitivity of the three controllers, model-based (-) data-based (*).

branch of the SRL shows only the blue pole. A green and red closed-loop pole lie also on this branch, but lie outside the plotted part of the graph.

Using these values of $\rho$, the corresponding optimal controllers are computed using Equation 5.8. In this computation, the data-based spectral decomposition as described in Section 5.5.1 has to be performed. Figure 5.6b shows an FRF of the resulting controllers, both computed from the model of the system (-) and with the data-based method (*). It can be seen that the data and model-based methods yield the same result.

When comparing the three controllers it can be observed that the controller corresponding to the high value of $\rho$, plotted in red, has a high gain. This can be explained because a high value of $\rho$ means that the output is weighted strongly in the criterion. This results in a high gain controller that shifts the pole far from its original location, while following the SRL. The controller that corresponds to a low $\rho$ value, plotted in blue, on the other hand has a low gain. Because the inputs are now more important in the cost function, the gain of the controller is lower. Therefore the performance of the output becomes less, and the poles are shifted only a small distance from their original location. The green controller shows a special case, because in this case the low frequent pole is chosen to have maximum damping. All three controllers have a $+1$ slope at high frequencies and thereby give phase advance to the system. For the controller corresponding to high $\rho$, the phase advance is given at higher frequencies than in the case of the controller corresponding to the low $\rho$ value. Also, the higher the value of $\rho$, the less dynamics are visible in the controller, and the more the controller starts to resemble a pure PD controller.

Performance of the controller

The open loop $CH$ and sensitivity $(1 + CH)^{-1}$ have also been computed for the three controllers, See Figure 5.7. Again it can be seen that a large $\rho$ corresponds to a high gain controller, which has a low sensitivity and that a small $\rho$ corresponds to a low gain controller with a sensitivity close to 0 dB. None of the controllers have a sensitivity that exceeds the 0 dB line, and therefore the Bode sensitivity integral does not hold. This can be explained because the requirement for the Bode sensitivity integral to hold is that the system must have at least two more poles than zeros. These systems have only one more pole than zeros and therefore the Bode sensitivity integral does not hold. The observation that the sensitivity does not exceed 1 can be proven easily. Recall the optimal return difference equation, and substitute $s = j\omega$

$$\bar{TT} = \rho \bar{H} + 1 \quad \Rightarrow \quad |T(j\omega)|^2 = \rho |H(j\omega)|^2 + 1$$

From this equation it is clear that $T \geq 1$ and therefore $S = T^{-1} \leq 1$. 

58
The open-loop of the green controller that was chosen to optimize the damping of the low frequent pole, crosses the 0 dB line exactly between the resonance and anti-resonance. This corresponds to a design rule proposed by [25], stated to optimize the damping in the system.

The three open loops have margins that approximate the LQR margins: 60 degrees phase margin and an infinite upwards gain margin. As shown in Section 5.4, the observer achieves perfect state estimation, which results in an controller that approximates an LQR controller.

The impulse responses of the closed-loop systems are evaluated too, see Figure 5.8. The impulse response gives information on the ability of the controller to control the system back to its initial position. This analysis is performed model-based, since time domain simulations cannot be performed on FRD directly. The position of the first mass is shown in the upper plot, while the position of the second mass is shown in the lower plot. The expectation could be that a higher value of ρ will lead to a better response, because the output is penalized more. But since the closed-loop pole location corresponding to high ρ has low damping, its response will not decay the fastest. The medium ρ value (green) that was chosen to optimize damping in the system leads to the best response for the first mass. This shows that it is very useful to be able to select the desired closed loop pole locations using the SRL.

5.6.3 Example 2: non-collocated control

The system used in this example is shown in Figure 5.9. The FRD \( H(j\omega) \) is computed from the model of this system, where the force on the first mass \( F_1 \) is the input but now the position of the second mass \( x_2 \) is the output, thus

\[
H(j\omega) = \frac{x_2}{F_1}
\]  

(5.46)

The difference with the first example is that point where the position is measured does not coincide with the point at which the force is applied. Thus this setup is a case of non-collocated control.
Computation of the controller $C(s)$

The controller can be computed in exactly the same as in Example 1. Thus TFD is computed with Cauchy after which the SRL is computed, see Figure 5.10a. Three desired closed-loop pole locations are chosen, depicted with the blue, green and red square behind the SRL points. From these locations the $\rho$ value is derived using Equation 5.44. Substitution of these rho values in the SRL equation, but using now the model of the system, gives the actual closed-loop pole locations that belong to these rho values. These locations are plotted with (×) in the same color as the corresponding chosen SRL points (the squares).

Then the optimal controllers are computed using Equation 5.8. The FRF’s of the resulting controllers are shown in Figure 5.10b. Again it can be seen that a high $\rho$ value gives a high gain controller. All controllers have a +2 slope at high frequencies to add phase advance and thereby stabilize the system.

Performance of the controller

The open loop and sensitivity are plotted for the three controllers, see Figure 5.11. Again note that a high gain and low sensitivity is found for the red controller, corresponding to a high $\rho$ value. The open loop responses have 60 degrees phase margin and infinite gain margin, due to the +2 slope of the controller at high frequencies.

An interesting result of this example is shown by the time domain response of the closed-loop system, see Figure 5.12. It shows that the response decays faster for a higher $\rho$ value. Due to the higher gain of the controller, however, the amplitude of the response is higher. Compare this to the time domain response of the collocated case. In practice the experience is that non-collocated control gives more restrictions on the performance than non-collocated control. The controller computed in this example performs so well because of the +2 slope. In practice this controller cannot be implemented directly, because the high gain at high frequencies would amplify noise very much and therefore will not perform good. Therefore it is desirable to have roll-off in the controller. This will be discussed in the next section.

5.7 Introducing roll-off in the controllers

The controllers derived in this chapter have infinite gain at high frequencies. This means that in practice they will not work, because high frequent measurement noise, which is always present in practice, will be amplified. To suppress measurement noise, the amplification at high frequencies should be low. Therefore a method to design a controller that has roll-off at high frequencies is necessary. One method is to add
Figure 5.11: (a) Open loop for the three controllers (b) Sensitivity of the three controllers, model-based (-) data-based (★).

Figure 5.12: Impulse response of the controlled system for the three controllers.

Figure 5.13: Second order weighting filter.
Figure 5.14: (a) SRL (black) and closed-loop poles and zeros (blue) of the system and SRL (green) and closed-loop poles and zeros (red) of the system with roll-off. (b) Resulting controller for the original system (black) and for the system with roll-off (green).

A weighting function that emphasizes the high frequency behavior of the plant. The controller will be designed for a weighted plant, such that roll-off is induced in the resulting controller. The weighted plant can be computed from

\[ H_W(s) = H(s)W(s) \]  

where \( W(s) \) is a filter that has gain 1 at low frequencies and a positive slope at high frequencies. Then the controller synthesis method from the previous sections is performed on the weighted plant \( H_W(s) \).

An example of the application of adding roll-off is presented next. Roll-off will be added to a controller for the system of Example 1. The weighting function, shown in Figure 5.13, has a +2 slope at high frequencies. A controller is designed for the weighted plant and compared to the original controller. The SRL of the plant with weighting (green) and of the original system (black) are shown in Figure 5.14a. Also shown are the optimal pole locations for the weighted (blue) and original (red) system. It can be observed that both the SRL and the optimal pole locations are the same for the weighted and the original system, for the part of the SRL that is shown. This could be expected since the SRL equation is the same for the weighted system and the original system

\[ 1 + \rho H = 0 \quad \Leftrightarrow \quad 1 + \rho H W = 0 \]  

provided that \( W(s) \approx 1 \) for small \( s \) values in the region of the s-plane that contains the relevant closed-loop poles. The SRL equation can also be used to compute the value of \( \rho \) for the weighted and original system. When the values of \( \rho \) and the closed-loop pole locations are equal for the weighted and original system, it means that the controllers are optimal for the same cost criterion

\[ J = \int_0^\infty \rho y^2 + u^2 \, dt. \]  

This property holds at low \( s \) values. However, the SRL with the weighting function has a different behavior at large values of \( s \) due to the added weighting function. It has two extra poles at the location of the zeros of the weighting filter (not shown in the graph). These poles cause the roll-off behavior.

The controller designed for the original system and the controller designed for the weighted system are shown in Figure 5.14b. The controller designed for the weighted system is similar to the original controller at low frequencies, but it has roll-off at high frequencies.

The controller designed for the weighted plant can be used to control the original system and will have good noise suppression results. However, it is not clear whether this controller is still optimal for the original system, since it is designed for the weighted system. Interestingly, the controller with roll-off resembles the LQG controller of Figure 5.2. The connection between the optimal controller with roll-off and an LQG controller is an interesting topic for future research.
5.8 Notes and suggestions for further work

Introducing tracking performance in the controller

It would also be desirable to be able to design an optimal controller that has good tracking performance.
In order to have good tracking performance, the gain of the controller should be high at low frequencies.
Some first attempts to add good tracking performance to the controller have been made. Firstly, adding
another weighting function to the plant was tried, similar to the weighting function used to add roll-off.
Unfortunately, this did not lead to a high gain at low frequencies.

Secondly, it was tried to improve the tracking performance with the following standard procedure

1. Augment plant dynamics with dynamics of \( \frac{1}{s} \) (to ensure that the loop contains an integrator)
2. Design LQR controller for augmented plant
3. Move integrator from plant to controller when implemented

But this did not give the desired result either. A final option is to add the integrator parallel to the the
designed controller, but the consequences with respect to optimality should then be investigated, which is
also left for future research.

MIMO

The proposed design method could also be applied to a MIMO system. In fact, Equation 5.1, that triggered
the research on this controller synthesis method, is proposed in [45] as an extension of their theory to
MIMO systems. But in order to compute the controller for a MIMO system data-based, the data-based
spectral decomposition should then be extended to MIMO systems. Currently, a such a MIMO data-based
spectral decomposition is not available.

Using \( C(s) \ vs \ C(j\omega) \)

In this chapter, the controller is only evaluated on the imaginary axis, so \( C(j\omega) \). It will be beneficial
to compute \( C(s) \) for all values of \( s \). This enables the use of the generalized MIMO Nyquist, in order to
analyze the relative stability of the system in a data-based way. Furthermore, this enables the use of the
Cauchy integral to find closed-loop poles of the system as proposed in Chapter 2.

Other performance variables

It can be beneficial to consider other, more general performance variables [45] using a cost function of the
type

\[
J = \int_{0}^{\infty} \left[ b \left( \frac{d}{dt} y(t) \right)^2 + u(t)^2 \right] dt.
\]  

(5.50)

With this performance measure, it is possible to include derivatives of \( y \) in the criterion. For the examples
given in this section, it might be possible to include a dynamic part which enables it to take the position
of the second mass as a performance measure while the position of the first mass is measured.

The optimum compensator in this case is still given by Equation 5.17, but the spectral decomposition
is now computed from

\[
\gamma = (\bar{p} + \bar{h}\bar{q})^+.
\]  

(5.51)

Substituting this equation in the optimal return difference and taking squares on both sides gives

\[
|1 + CH|^2 = 1 + |hH|^2.
\]  

(5.52)

It has been shown by Kalman [24] that this is both a necessary and sufficient condition for the controller
to be optimal.
Solution Riccati equation

The (data-based) computation of the optimal controller in the frequency domain using the spectral decomposition is related to solving the Riccati equation in the time domain approach. Therefore it might be possible to compute the solution of the Riccati equation $P$ by comparing the two approaches. Let

\begin{align*}
\gamma_r &= [\gamma_0 \ldots \gamma_{n-1}]^T \\
p_r &= [p_0 \ldots p_{n-1}]^T
\end{align*}

(5.53) (5.54)

where $n$ is the order of the system. Then it can be proved that

$$\rho P B = \gamma_r - p_r.$$  

(5.55)

A large step has to be made to be able to compute this data-based, since the order of the system $n$ and the separate coefficients of the spectral decomposition $\gamma_i$ and $p_i$ cannot be determined directly from the TFD.
Chapter 6

Conclusions and Recommendations

6.1 Conclusions

The main conclusion of this thesis is that TFD, can be exploited to

- Analyze the relative stability of MIMO systems using the generalized MIMO Nyquist theorem
- Compute optimal closed-loop pole locations for a system using the SRL
- Synthesize an optimal controller for this system that achieves the closed-loop pole locations computed with the SRL

using only FRD of the system, without the use of a parametric model.

This is not possible by using conventional methods, which would require a parametric model of the system. A limitation to the proposed theory is that it can only be applied to lightly damped mechanical systems, because TFD can only be computed for this class of systems. Nevertheless, there are many mechanical systems that meet these requirements because of modern design methods.

The relative stability analysis performed with the generalized MIMO Nyquist theorem, derived in Chapter 3, gives an interpretation of stability margins for MIMO systems, which is not possible when using the conventional Nyquist theorem. It can be used to compute the minimum amount of damping of the closed-loop poles in a system.

The SRL, discussed in Chapter 4, combined with the optimal controller synthesis method, provides a new controller design method as discussed in Chapter 5. With this method it can be computed what the closed-loop pole locations are that can be achieved with an optimal state feedback controller. The proposed synthesis method can then be used to design an output feedback controller that achieves the desired optimal closed-loop pole locations.

Three applications of TFD are found and applied successfully.

6.2 Recommendations

This thesis has shown that the availability of TFD provides a way to apply several model-based controller synthesis techniques when only FRD of the system is available, without computing a parametric model. But not all possibilities for the application of this technique have been studied yet, and the class of systems for which this technique is available could be extended. Therefore it is highly recommended to continue research on this topic. A few suggestions for further research are given in this section.

Generalized MIMO Nyquist

The generalized MIMO Nyquist theorem provides a way to estimate the minimum amount of damping of the closed-loop poles of a system. Using the techniques used in this thesis, it is possible to compute all the closed-loop pole locations of the system. Because $H(s_i)$ is known, the Cauchy integral can be used to compute the closed-loop pole locations of $(1 + CH)^{-1}$ for a given controller $C(s)$, using the pole finding technique described in Section 2.5. This is an alternative for the data-based root locus method described in [12], that can also be used to compute the closed-loop pole locations for a given system and controller.
Data-based optimal control

The data-based optimal controller could be improved by adding integral action to improve tracking performance. Furthermore, it is interesting to study the relation between the data-based optimal controller with roll-off and conventional LQG controllers, since their FRFs have a similar shape. Investigating the possibility to use other performance criteria in the optimal controller synthesis, by adding derivatives of the output is recommended too.

The extension to MIMO systems is highly recommended since for MIMO systems the loop shaping methods are very limited, while it might be possible to compute an optimal MIMO controller data-based using the approach followed in Chapter 5. Studying the possibility to compute a MIMO spectral decomposition is essential to achieve this goal.

Other applications of TFD

The applications studied in this thesis are certainly not the only applications of TFD. Before the state-space approach became popular in the 1960s, a lot of results were derived in the frequency domain. It is recommended to look for results in terms of \( H(s) \) that can be computed data-based, using TFD.

The behavior approach is another field in which results in terms of \( H(s) \) might be found. For example, a system representation known as the kernel representation \( R \) in the behavior approach can be related to \( H(s) \),

\[
R \left( \frac{d}{dt} \right) w = 0 \quad \Leftrightarrow \quad [1 - H(s)] \begin{bmatrix} y \\ u \end{bmatrix} = 0
\]

for \( w = [u \ y]^T \). Some work is necessary to connect these two, since the kernel representation is in the time-domain, while \( H(s) \) is in the Laplace domain. Furthermore, the behavior theory for rational functions such as \( H(s) = \frac{N(s)}{D(s)} \) has not been fully developed yet.

Extension of the class of systems

The proposed theory is now limited to lightly damped mechanical systems, since TFD can only be computed because of the symmetry in the s-plane. Research could be performed to extend the class of systems for which TFD can be computed. A starting point for this research could be to try to compute TFD for first order systems. Inspiration can be found in [40], [41] and [39] in which identification tools for first order systems are developed.

Numerical improvements

The numerical computation of TFD could be improved, in order to decrease the computation time. As proposed in [12], it might be possible to derive a Laplace version of the fast Fourier transform (FFT).
Appendix A

Proof of the argument principle

Let \( G(s) \) be a complex function with a pole at \( s = s_0 \) with multiplicity \( l \). Then \( h(s) := (s - s_0)^l G(s) \) is analytic. The derivative of

\[
G(s) = \frac{h(s)}{(s - s_0)^l}
\]

is

\[
G'(s) = \frac{-lh(s)}{(s - s_0)^{l+1}} + \frac{h'(s)}{(s - s_0)^l}.
\]

The quotient of \( G(s) \) and its derivative is

\[
\frac{G'(s)}{G(s)} = \frac{-l}{(s - s_0)^{l+1}} + \frac{h'(s)}{h(s)}.
\]

Because \( \frac{h'(s)}{h(s)} \) is analytic, the contour integral of this part is zero. Therefore, the contour integral along a closed path yields

\[
\frac{1}{2\pi i} \oint_C \frac{G'(s)}{G(s)} \, ds = -l.
\]

The contour integral gives the number of poles that lie inside the contour. This can be generalized for the other poles and zeros of the complex function. Denote the number of poles and zeros inside the contour by \( P \) and \( Z \) respectively. Then,

\[
\frac{1}{2\pi i} \oint_C \frac{G'(s)}{G(s)} \, ds = Z - P.
\]

This can be formulated differently as follows:

\[
Z - P = \frac{1}{2\pi i} \oint_C d[\ln(G(s))] \, ds = \frac{1}{2\pi i} \ln|G(s)|_C \tag{A.6}
\]

\[
= \frac{1}{2\pi i} \ln|G(s)|_C + i \arg G(s)_C \tag{A.7}
\]

\[
= \frac{1}{2\pi} \arg G(s)_C \tag{A.8}
\]

\[
= \frac{1}{2\pi} \text{variation of } \arg G(s) \text{ around } C \tag{A.9}
\]

Note that equation A.8 follows from equation A.7 because for any complex number \( z \) it holds that \( \ln(z) = \ln(|z|e^{i\arg z}) = \ln|z| + i \arg z \). Furthermore equation A.9 follows from equation A.8 since \( \ln|G(s)|_C \) is zero because \( C \) is a closed contour.

Equation A.10 shows that the difference in number of poles and zeros of \( G(s) \) is equal to the number of encirclements of the origin by \( G(s) \).
Appendix B

Return Difference

The return difference matrix incorporates both open-loop and closed-loop stability parameters in its determinant. Therefore it is important in the analysis of stability in the frequency domain. The term return ‘difference’ originates from the reasoning depicted in Figure B.1. A state-space system $(A,B,C)$ is controlled by means of state feedback $K_\infty$. If the loop depicted in the figure is broken at ‘$X$’, the difference between the injected signal $\alpha$ and returned signal $\beta$ is

$$\alpha(s) - \beta(s) = T(s)\alpha(s). \quad (B.1)$$

Thus $T$ is defined as

$$T(s) \equiv [I + K_\infty(sI - A)^{-1}B] \quad (B.2)$$

For other control strategies, it is also possible to define the return difference in an equivalent way. For a standard negative feedback scheme with system $H$ and controller $C$, the return difference is given by

$$T(s) = 1 + C(s)H(s). \quad (B.3)$$

Figure B.1: Derivation of the return difference
Applications of Transfer Function Data
Appendix C

Phase reconstruction from gain behavior

Let $H(j\omega)$ denote a transfer function that can be decomposed in its real $R(\omega)$ and imaginary $I(\omega)$ parts

$$H(j\omega) = R(\omega) + jI(\omega) \quad (C.1)$$

Then $R(\omega)$ and $I(\omega)$ form a Hilbert pair, for which it holds that

$$R(\omega) = R(\infty) - \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{I(y)}{y-\omega} dy \quad (C.2)$$
$$I(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R(y)}{y-\omega} dy. \quad (C.3)$$

This implies that if the real part is known, the imaginary part can be computed and vice versa. If the system is minimum phase and stable, $H(j\omega)$ will be analytic in the RHP. Taking the logarithm of $H(j\omega)$ then gives

$$\ln[H(j\omega)] = \alpha(\omega) + j\beta(\omega) \quad (C.4)$$

where $\alpha(\omega)$ is the gain in nepers and $\beta(\omega)$ is the phase. The phase can then be determined uniquely from

$$\beta(\omega) = \frac{2\omega}{\pi} \int_{0}^{\infty} \frac{\alpha(y) - \alpha(\omega)}{y^2 - \omega^2} dy, \quad (C.5)$$

which gives a relation that enables the computation of the phase of the system, given a certain gain behavior. Because for $\omega = y$ the integrand goes to infinity, direct implementation of this integral is not possible. Various implementations of this relation can be found in [8].

One way to compute the integral is to consider it as a contour integral as depicted in Figure C.1. Because of Cauchy’s theorem this contour integral will give the value of $\beta(\omega)$. It can be proved that the integral of the small indent around the singular point $s_i$ converges to $\frac{1}{2} \beta(\omega)$. Therefore, the rest of the integral must also equal $\frac{1}{2} \beta(\omega)$. Thus

$$\beta(\omega) = \frac{\omega}{\pi} \int_{0}^{\omega-\epsilon} \frac{\alpha(y) - \alpha(\omega)}{y^2 - \omega^2} dy + \frac{\omega}{\pi} \int_{\omega+\epsilon}^{\infty} \frac{\alpha(y) - \alpha(\omega)}{y^2 - \omega^2} dy, \quad (C.6)$$

provides a way to compute the integral, which provides a method to compute the phase from the gain behavior of the system. This integral can be computed by simply converting Equation C.6 to a summation.
Figure C.1: Contour integral used for phase reconstruction.
Bibliography


Applications of Transfer Function Data