Robust Adaptive Control

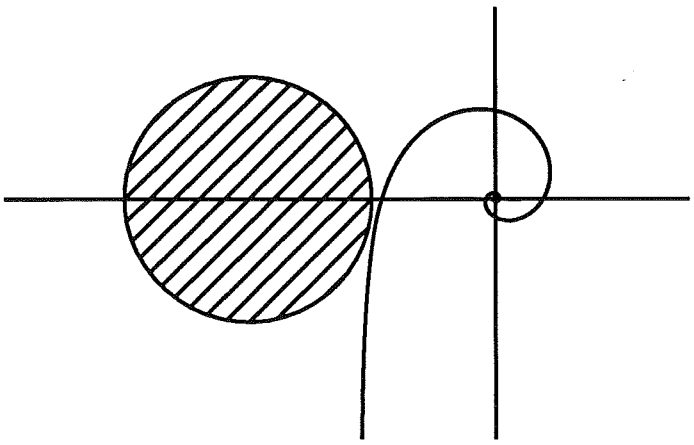
Lundh, Michael

1991

Document Version:
Publisher's PDF, also known as Version of record

Robust Adaptive Control

Michael Lundh



Lund 1991

Department of Automatic Control, Lund Institute of Technology

To Maria, Joachim and Henric

Department of Automatic Control
Lund Institute of Technology
Box 118
S-221 00 LUND
Sweden

Published 1991
Printed in Sweden by Studentlitteratur AB

Department of Automatic Control
Lund Institute of Technology
P.O. Box 118
S-221 00 Lund Sweden

Document name
DOCTORAL DISSERTATION

Date of issue
December 1991

Document Number
CODEN: LUTFD2/(TFRT–1035)/1-112/(1991)

Author(s)
Michael Lundh

Supervisor
Karl Johan Åström

Sponsoring organisation

Title and subtitle
Robust Adaptive Control

Abstract

This thesis deals with two problems in robust and adaptive control. First a method for robust control system design is presented. A constrained convex optimization problem is used to find a controller that guarantees robust performance for an uncertain process where the uncertainty is described as a combination of structured and unstructured uncertainties. Adaptive control is an alternative way to deal with uncertain processes. However, adaptive controllers may be difficult to commission. A method for automatic initialization of an adaptive controller is presented. The method uses information from an experiment with relay feedback to initialize a robust adaptive controller.

Key words
Robust Control, Optimization, Adaptive Control, Initialization, Pretuning

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title

ISBN

Language
English

Number of pages
112

Recipient's notes

Security classification

Contents

Preface

In industry there exist a huge number of controllers that through feedback provide processes to operate more reliable, more exact or more economically despite disturbances affecting the process. To obtain this it is required that the controllers are properly tuned. Two different approaches for control system design are taken in this thesis. One approach is a robust design method that requires substantial engineering effort but can guarantee a control system with certain properties. The other approach is suited when the amount of feedback loops makes it impossible to spend too much effort on each controller. Using a procedure for automatic initialization of an adaptive controller allows a control system to be tuned by simply pushing a button. The research behind this thesis began initially with a study of some adaptive controllers. Trying to understand their behavior led me into the area of robust control design.

Acknowledgements

First of all I would like to thank my supervisor Professor Karl Johan Åström who suggested the thesis subject. He has been an inexhaustible source of encouragement and enthusiasm. I would also like to thank all my colleagues at the department. Special thanks goes to Björn Wittenmark, Per Hagander, Tore Hägglund, and Mats Lilja who have given many valuable comments on the manuscript. The excellent computer facilities at the department are supported by Leif Andersson and Anders Blomdell. Eva Dagnegård has managed the library of BIB-TEX references and has assisted with type setting problems. Britt-Marie Mårtensson has skillfully prepared the figures. Finally, I would like to thank my wife and children for their love and support.

1

Introduction

The purpose of feedback control is to make a system behave in a predicted manner irrespective of its environment, or more safe, or more economically. Reduction of the effects of disturbances and process variations are the main reasons for using feedback on open loop stable processes. This is accomplished by a well designed feedback. If no disturbances were acting on the process and a perfect description of the process is available, then feedforward control would be sufficient. If on the other hand the open loop process is unstable these objectives are only secondary because the main purpose of feedback control is then to stabilize the process.

There are many different requirements on a control system that have to be considered in the design process. The demands are sometimes contradictory and it may be necessary to make trade-offs between conflicting design goals. Some important issues are model uncertainty, command signal following, rejection of load disturbances and sensor noise, and actuator saturation. Unfortunately there exists no design method that considers all these issues simultaneously. Different design methods focus on different issues. Requirements that are not explicitly considered in the design should be analyzed afterwards. Optimization is a general method that can deal with many requirements. This idea of formulating control system design as an optimization problem was pioneered in [Zakian and Al-Naib, 1973]. Fundamental contributions over a long period of time is summarized in [Polak *et al.*, 1984]. The design procedure leads to a semi-infinite programming problem because constraints are given by frequency and time domain functions. Consideration of a finite number of frequencies and times leads to significant simplifications, that has been pursued in [Boyd and Barratt, 1991]. Another aspect of con-

trol system design is the amount of engineering effort that is spent in the design. Will a simple, computationally cheap design method give sufficient performance of the system, or is it necessary to use an advanced, perhaps computationally intensive design method to obtain desired performance.

Two different approaches of control system design will be taken in this thesis. First, a new design method for single input single output processes is presented. It considers uncertainty in the process model explicitly. The controller is obtained by solving a large optimization problem with frequency domain constraints. A substantial engineering effort is required by this method. It can give the limit of some performance objective. If this performance is not satisfactory, as may happen for processes with large uncertainties, an adaptive controller is an alternative to improve performance.

Adaptive control has the potential to give desired performance for a control system with uncertainty. A drawback is that adaptive controllers may be difficult to commission. For this reason the most used adaptive technique is automatically tuned PID-controllers [Kraus and Myron, 1984] and [Åström and Hägglund, 1988a]. These controllers have been developed to the stage where tuning is performed simply by pushing a button. It would be desirable to make adaptive controllers as easy to use as the PID autotuner. A procedure for automatic initialization of an adaptive controller will be developed in this thesis. The key idea is to use information from an experiment with relay feedback to initialize the adaptive controller. The initialization procedure together with the adaptive controller gives a new auto-tuning approach that works for a larger class of processes than those which can be well controlled by PID-controllers. The use of an auto-tuner is a way to reduce the engineering effort for the design. This is particularly interesting in the process industry where there may be thousands of control loops.

The thesis is organized as follows. Chapter 2 defines the control system and presents pole placement design. A robust design problem using constrained convex functions is formulated in Chapter 3. These are used in the optimization problem discussed in Chapter 4, where some examples also are given. Chapter 5 presents an adaptive controller. Chapter 6 studies an experiment with relay feedback and suggest information that can be obtained from this experiment. A procedure for automatic initialization of an adaptive controller is proposed in Chapter 7, where some examples demonstrate properties of the procedure. Conclusions are given in Chapter 8. A test batch of processes, used to evaluate the initialization procedure for the adaptive controller, is given in Appendix A.

2

Preliminaries

The purpose of this chapter is to introduce the feedback system and the pole placement design that will be considered in this thesis.

2.1 The Feedback System

The process to be controlled has one output signal y, that is supposed to follow a desired reference, one input signal u that may be manipulated by a control law. Furthermore two disturbances l and d affect the process. These are assumed to act at the process input and at the process output respectively. The process itself is assumed to be described by a rational transfer function. This is a ratio of two polynomials B and A in the differential operator $\frac{d}{dt}$ for continuous-time systems, or in the shift operator q for discrete-time systems. The process output is then given by

$$y = \frac{B}{A}(u + l) + d \tag{2.1}$$

It is assumed that the measurement of y is corrupted by measurement noise n. The control actions thus have to be based on the signal z given by

$$z = y + n \tag{2.2}$$

The controller is linear with two degrees of freedom. A general finite dimensional control law is then given by

$$u = \frac{S}{R}\left(\frac{B_{\mathrm{ff}}}{A_{\mathrm{ff}}}r - z\right) \tag{2.3}$$

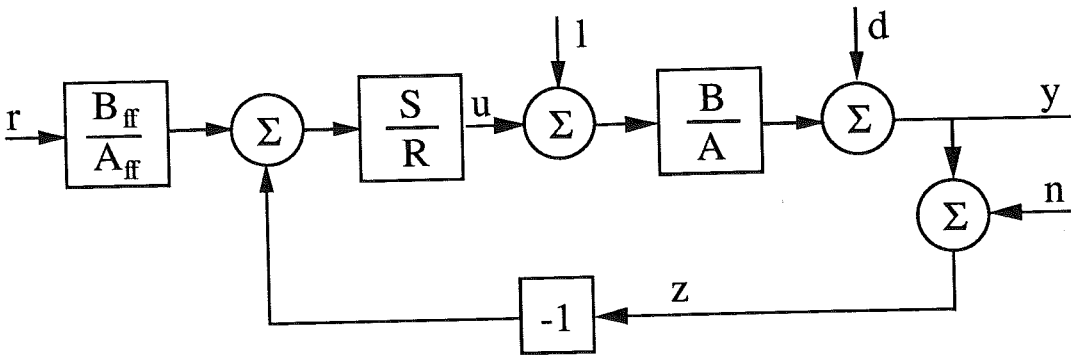


Figure 2.1 The closed loop system

where r is the reference signal. The controller is described by four polynomials. The feedback design determines the polynomials S and R to obtain desired disturbance rejection properties and desired measurement noise attenuation. Once the feedback loop is determined, the feedforward design provides the polynomials B_{ff} and A_{ff} for a desired reference signal response.

A block diagram of the feedback system is shown in Figure 2.1. The process output y and the process input u for the closed loop system are obtained from (2.1), (2.2) and (2.3).

$$\begin{aligned} y &= \frac{BS}{AR+BS}\left(\frac{B_{ff}}{A_{ff}}r - n\right) + \frac{AR}{AR+BS}d + \frac{BR}{AR+BS}l \\ u &= \frac{AS}{AR+BS}\left(\frac{B_{ff}}{A_{ff}}r - n - d\right) - \frac{BS}{AR+BS}l \end{aligned} \tag{2.4}$$

Introduce

$$\begin{aligned} \mathcal{T} &= \frac{BS}{AR+BS} \\ \mathcal{S}_o &= \frac{AR}{AR+BS} \\ \mathcal{S}_i &= \frac{BR}{AR+BS} \\ \mathcal{S}_n &= \frac{AS}{AR+BS} \end{aligned} \tag{2.5}$$

where $\mathcal{T}$ and $\mathcal{S}_o$ are recognized as the complementary sensitivity function and the sensitivity function for the closed loop system [Doyle and Stein, 1981]. Two other sensitivity functions are defined, the sensitivity for input disturbances $\mathcal{S}_i$ and the input sensitivity for measurement noise $\mathcal{S}_n$. Using

these transfer functions equation (2.4) may now be written as

$$\begin{aligned} y &= \mathcal{T}\frac{B_{\text{ff}}}{A_{\text{ff}}}r + \mathcal{S}_i l + \mathcal{S}_o d - \mathcal{T}n \\ u &= \mathcal{S}_n\frac{B_{\text{ff}}}{A_{\text{ff}}}r - \mathcal{T}l - \mathcal{S}_n d - \mathcal{S}_n n \end{aligned} \tag{2.6}$$

Design objectives may be given in terms of desired properties of the transfer functions in (2.5). From equation (2.6) it is obvious that control design is a trade-off between conflicting objectives. For some of the transfer functions high bandwidth would be demanded and for others a low bandwidth is desired. However, these transfer functions are strongly dependent, since they have the same characteristic polynomials.

The loop transfer function is another important quantity. It is composed of the feedback controller and the open loop process in series, i.e

$$L = \frac{BS}{AR}$$

Many qualities of the closed loop system may be demonstrated by the properties of L. In classical servo design the compensator S/R is modified to assign desired properties to the loop transfer function [Chestnut and Máyer, 1959]. The complementary sensitivity function and the sensitivity function may be expressed in the loop transfer function.

$$\begin{aligned} \mathcal{T} &= \frac{L}{1+L} \\ \mathcal{S}_o &= \frac{1}{1+L} \end{aligned}$$

Also notice that $\mathcal{T} + \mathcal{S}_o = 1$.

2.2 Pole Placement Design

Pole placement is a simple design method where a two degree of freedom control law may be achieved without iterative calculations. The design method focuses on command signal following and load disturbance rejection [Åström and Wittenmark, 1990]. A desired closed loop characteristic polynomial is chosen and the feedback controller is provided by the solution to a polynomial equation.

The open loop process is given by (2.1) and the controller is given by

$$Ru = Tr - Sz$$

which corresponds to (2.3) with $B_{\text{ff}} = T$ and $A_{\text{ff}} = S$. The closed loop system is given by (2.4).

The design is performed in two steps, first the feedback controller S/R is designed to give the desired response for the disturbances l or d or both and to give the desired measurement noise attenuation. Once the feedback controller is designed, the polynomial T may be chosen for the feedforward to give the desired response to reference signals.

Feedback Design

The feedback controller design may be formulated such that constant disturbances l and d are eliminated in stationarity. Consider a step disturbance l to be generated by $A_r l = e_l$ where e_l is a pulse and $A_r = s$ in the continuous time case or $A_r = 1 - q^{-1}$ in the discrete time case. The influence of l in the output signal is then

$$y_l = \frac{BR}{AR + BS} l = \frac{BR}{A_r(AR + BS)} e_l$$

Provided that $AR + BS$ is asymptotically stable and A_r is a factor of R, it follows that

$$\lim_{t \to \infty} y_l = 0$$

i.e. the effect from the disturbance is eliminated in stationarity. This corresponds to integral control. The controller will then have high gain at low frequencies. The controller properties required to eliminate output disturbances are derived similarly. However, it may be noted that constant output disturbances always are eliminated in stationarity if the open loop process contains an integrator.

To carry out the design the process model numerator is factorized as

$$B = B^+ B^-$$

where B^+ is a monic polynomial whose zeros are stable and well damped. These zeros are canceled by the controller. The remaining zeros B^- are not canceled and will appear in the closed loop transfer functions $\mathcal{T}$ and $\mathcal{S}_i$. Since B^+ is canceled it is a factor of the closed loop characteristic polynomial. The zeros of B^+ will correspond to modes in the control signal u that are unobservable in the process output y and B^+ should therefore only contain stable and well damped zeros.

The closed loop characteristic polynomial consists of B^+ and A_c with the desired closed loop poles

$$AR + BS = A_c B^+ \tag{2.7}$$

Since B^+ is a factor of B and also is present in the right hand side of (2.7) it must be a factor in R. Then (2.7) may be simplified to the Diophantine-Aryabhatta-Bezout equation

$$AA_rR' + B^-S = A_c \tag{2.8}$$

from which R' and S are achieved. To ensure a proper controller the characteristic polynomial A_c must be chosen such that

$$\deg A_c \geq \min \deg A_c = 2\deg A - \deg B^+ + \deg A_r - 1 \tag{2.9}$$

[Åström and Wittenmark, 1990]. If the degree of the polynomial A_c is higher than necessary the solution of the Diophantine-Aryabhatta-Bezout equation does not yield a unique proper controller. This freedom can be used for other purposes, e.g. improving robustness, as will be proposed in Chapters 3 and 4. The feedback compensator is given by S in (2.8) and

$$R = R'A_rB^+$$

which then is forced to contain the factor A_r.

Freedom of the Diophantine-Aryabhatta-Bezout Equation

If there exists one solution R'_0 and S_0 to the Diophantine-Aryabhatta-Bezout equation (2.8) there is an infinite number of solutions

$$\begin{aligned} R' &= R'_0 - QB^- \\ S &= S_0 + QAA_r \end{aligned}$$

parameterized by the the polynomial Q. Assume that R'_0 and S_0 yields a proper controller. Not all choices of Q lead to proper controllers. The degree of the closed loop characteristic polynomial A_c determines the admissible degree of Q. The degree of the controller is $\deg R = \deg A_c + \deg B^+ - \deg A$. To have a proper controller the degree of S should satisfy $\deg R \geq \deg S = \max(\deg S_0, \deg Q + \deg A + \deg A_r)$. If $\deg A_c = \min \deg A_c$ then only $Q = 0$ gives a proper controller. If $\deg A_c > \min \deg A_c$ then (2.9) gives

$$\begin{aligned} \deg Q &\leq \deg A_c - 2\deg A + \deg B^+ - \deg A_r \\ &= \deg A_c - \min \deg A_c - 1 \end{aligned} \tag{2.10}$$

Equation (2.10) shows that each increase in the degree of A_c allows the same increase in the degree of Q.

If Q is not restricted to be a polynomial but a stable rational transfer function, any such $Q = N/D$ satisfying

$$\deg N - \deg D \leq \deg A_c - \min \deg A_c - 1 \tag{2.11}$$

gives a proper controller

$$\frac{S}{R} = \frac{S_0 + QAA_r}{A_r(R_0' - QB)} = \frac{S_0 + QAA_r}{R_0 - QBA_r} \tag{2.12}$$

Here $B^+ = 1$, which is no restriction. The closed loop characteristic polynomial is the product A_cD. There exist polynomials N and D such that the controller may have common stable factors in S and R that can be used to cancel factors in A_c. Unstable common factors can not appear in S and R. This means that all stabilizing controllers with the factor A_r in the denominator can be parameterized as (2.12) where the casual controller S_0/R_0 stabilizes the process B/A and $Q = N/D$ satisfies (2.11). This resembles the the Youla parameterization or the Q parameterization [Vidyasagar, 1985].

Solution of the Diophantine-Aryabhatta-Bezout Equation

The controller polynomials are given by the solution of equation (2.8). This equation is equivalent to a system of linear equations. Define the process model and the controller as polynomials in z.

$$\begin{aligned}
A(z)A_r(z) &= \alpha_0 z^{\eta} + \alpha_1 z^{\eta-1} + \cdots + \alpha_\eta \\
B^-(z) &= \beta_0 z^{\mu} + \beta_1 z^{\mu-1} + \cdots + \beta_\mu \\
A_c(z) &= z^{\eta+d_r} + a_{c1} z^{\eta+d_r-1} + \cdots + a_{c(\eta+d_r)} \\
R'(z) &= r_0 z^{d_r} + r_1 z^{d_r-1} + \cdots + r_{d_r} \\
S(z) &= s_0 z^{d_s} + s_1 z^{d_s-1} + \cdots + s_{d_s}
\end{aligned}$$

The polynomial equation (2.8) can then be written as the linear equation

$$Mx = P \tag{2.13}$$

where P is a column vector with the coefficients of A_c, x is a vector with the controller coefficients

$$x = (\, r_0 \quad \cdots \quad r_{d_r} \quad s_0 \quad \cdots \quad s_{d_s} \,)^T \in \mathbb{R}^{d_r+d_s+2}$$

and M is a $(\eta + d_r + 1) \times (d_r + d_s + 2)$ Sylvester matrix

$$M = \begin{pmatrix} \alpha_0 & & & \beta_0 & & \\ \alpha_1 & \ddots & & \beta_1 & \ddots & \\ \vdots & \ddots & \alpha_0 & \vdots & \ddots & \beta_0 \\ \alpha_\eta & \ddots & \alpha_1 & \beta_\mu & \ddots & \beta_1 \\ & \ddots & \vdots & & \ddots & \vdots \\ & & \alpha_\eta & & & \beta_\mu \end{pmatrix} \tag{2.14}$$

The matrix has $d_r + 1$ columns with α-coefficients and $d_s + 1$ columns with β-coefficients. Leading β coefficients may well be zero here since it is required that

$$\eta + d_r = \mu + d_s$$

to make the block matrices with α and β coefficients have the same number of rows.

If AA_r and B^- are coprime, the linear equation (2.13) has a solution. It is unique if the degree of A_c is minimal. If on the other hand the polynomials AA_r and B^- have a common factor G, the Sylvester matrix M is singular and thus not invertible. If G also is a factor in A_c, there exists a solution to the Diophantine-Aryabhatta-Bezout equation, otherwise not. Numerical problems may occur if the polynomials AA_r and B^- have factors that are close.

Choice of Characteristic Polynomial

Mathematically the closed loop poles may be chosen arbitrarily. However, certain choices may lead to a closed loop system that is very sensitive to variations in the process model, and should therefore be avoided. A reasonable choice is to let the open loop properties, i.e poles, guide the choice of closed loop poles. For an open loop stable continuous time process it is often reasonable to let the closed loop poles have the same distance to the origin as the open loop process poles. Choices far away from this recommendation may lead to unstable controllers or controllers with high gain, both undesirable from a sensitivity point of view. See [Lilja, 1989] for further discussion.

Design of Feedforward

The choice of feedback compensator was motivated from the disturbances l or d. There is, however, still freedom in the control law to select the polynomial T in the feedforward to give an appropriate response to reference signals. For the pole placement controller, the transfer function from reference signal to output is

$$y = \frac{BT}{AR + BS} r = \frac{B^- T}{A_c} r \tag{2.15}$$

where the degree of T satisfies $\deg T \leq \deg R$. The closed loop characteristic polynomial may be factored as $A_c = A_m A_o$. Select $T = B'_m A_o$ and (2.15) gives

$$y = \frac{B^- B'_m}{A_m} r \tag{2.16}$$

The factor A_o may be interpreted as an observer polynomial that will be canceled in the transfer function from r to y [Åström and Wittenmark, 1990]. The role of the polynomial B'_m is to adjust the static gain and, if desired, introduce additional zeros in the transfer function (2.15).

The reference signal response may be further manipulated if the feedforward instead is chosen as the stable proper transfer function B_{ff}/A_{ff}.

Discussion

Any design method using a finite dimensional model described by a rational transfer function as (2.1) could be interpreted as a pole placement design. Therefore the question of where the poles should be located is highly interesting.

Pole placement design, as described in this chapter, does not consider model uncertainty. If the process model is uncertain, either the robustness properties of the control system must be checked or the design has to be modified to cope with an uncertain process. There exist some methods for incorporating model uncertainty into pole placement design [Soh *et al.*, 1987, Boyd *et al.*, 1988].

Due to the simplicity of the pole placement design method it is popular in adaptive controllers where the computational complexity must be limited.

3

Objectives for Robust Design

Chapter 3 and 4 deal with robust design of uncertain single input single output processes. Some of the ideas here were presented in [Lundh, 1990]. The purpose of this chapter is to describe uncertainty of process models and to use this description to define frequency domain objectives for robust design. The focus is on design of a feedback compensator, the feedforward design problem is briefly mentioned at the end of the chapter.

3.1 Introduction

In Chapter 2 the process was assumed to be described by a rational transfer function. This model is only an approximation. If the model deviates much from the real process, the actual behavior of the real closed loop system will not be as predicted from the model. Bad performance or even instability may be encountered. It is, therefore, of great significance to describe the accuracy of the model and to design the control law such that the closed loop system performance is acceptable for all processes within the uncertainty envelope for the model. The process model uncertainty influences the achievable performance. Small uncertainty implies that the process behavior is rather well defined. It is then possible to design with small robustness margins. Hence, the performance may be improved.

There are many sources for model uncertainty. Variations may occur due to imprecise manufacturing, a cheap component with large tolerances

may be used instead of an expensive one with small tolerances. This is actually one of the strong motivations for using feedback control. Other sources of uncertainty are wear and heating. Real processes are nonlinear and a linearized model is describing the process for some operating conditions. Different operating conditions may be described by an uncertain linear model. It is assumed that all these variations occur rarely or are so slow such that the process may be considered as time invariant.

Two types of model uncertainties are discussed here. Each with its own characteristics. A combination of the two types of uncertainty descriptions is proposed. It is shown that this combination of uncertainty descriptions has some nice properties.

Only continuous time processes are considered here since structured uncertainty is often coupled to physical parameter variations, and these couplings become very complex when the process is sampled [Åström and Wittenmark, 1989]. It may well be argued that the physical structure is completely destroyed by the sampling.

Feedback Performance Objectives

Feedback control design is a trade-off between different objectives, e.g. load disturbance rejection and measurement noise attenuation. Assume first that the real process is perfectly described by the model. Then, the only reason for feedback control of a stable process is to reduce the effects from the disturbances.

Let $\mathcal{S}$ be one of the closed loop transfer functions in (2.5). The feedback compensator design gives this transfer function desired properties. The specifications can be expressed as limits on the magnitude of the frequency response. Desired disturbance rejection is achieved by making $\mathcal{S}$ small in a certain frequency interval. A convenient way to limit $\mathcal{S}$ is to introduce a weight $W(\omega)$ and the inequality

$$W(\omega)|\mathcal{S}(i\omega)| \leq 1, \qquad \forall\omega$$

The weight $W(\omega)$ may be the frequency response magnitude function of some filter as in H_∞ control design.

The disturbance rejection properties are described by the three strictly stable transfer functions $\mathcal{S}_o$, $\mathcal{S}_i$, and $\mathcal{S}_n$ in (2.5). The disturbance rejection objectives are defined through the inequalities

$$\begin{aligned} W_o(\omega)|\mathcal{S}_o(i\omega)| &\leq 1 \\ W_i(\omega)|\mathcal{S}_i(i\omega)| &\leq 1 \\ W_n(\omega)|\mathcal{S}_n(i\omega)| &\leq 1 \end{aligned} \tag{3.1}$$

which should hold for all frequencies ω.

The first inequality impose a limit on the transfer function from output disturbance to process output, i.e. the sensitivity function. The second inequality limits the transfer function from input disturbance to process output. The third inequality limits the transfer function from measurement noise to control signal. The inequalities (3.1) are referred to as the *nominal performance criteria.*

Certain properties of the weights in (3.1) leads to certain closed loop behavior. Step disturbances affecting the system at the process input or at the process output are eliminated in the output y in stationarity provided that either $\mathcal{S}_i(0) = 0$ or $\mathcal{S}_o(0) = 0$. A sufficient condition for this is to require that $R(0) = 0$ in the feedback controller. The controller is then said to have integral action. This correspond to weights $W_i(0) = \infty$ and $W_o(0) = \infty$ respectively. Measurement noise is present in all systems. High noise energy in the control signal u is undesirable since it leads to frequent control actions that causes actuator wear. Therefore the transfer function $\mathcal{S}_n$ from noise n to control signal u should be limited. The weight function $W_n(\omega)$ is chosen with respect to spectral characteristics of the noise n acting on the system.

3.2 Unstructured Uncertainty

This section reviews and extends some well known results from robust design [Doyle and Stein, 1981, Morari and Doyle, 1986]. Relations between process model uncertainty and performance are discussed in the frequency domain.

Unstructured uncertainty is used to describe the accuracy of a model. Let the real process be defined by the transfer function G^0 and let G be the transfer function of the model of the process. The unstructured uncertainty is specified by the function $W_u(\omega)$, defined by

$$\left|\frac{G^0(i\omega) - G(i\omega)}{G(i\omega)}\right| \leq W_u(\omega), \qquad \forall\omega$$

The real process is then considered to be described by

$$G^0 = GG_u = \frac{B}{A}(1 + \Delta W_u) \tag{3.2}$$

where $G(s) = B(s)/A(s)$ is the model of the process and

$$G_u(i\omega) = (1 + \Delta(i\omega)W_u(\omega))$$

is a stable transfer function with $|\Delta(i\omega)| \leq 1$. The multiplicative unstructured uncertainty may be viewed as a disc in the Nyquist diagram, to which

the frequency response of the real process belongs for a certain frequency. The center of the disc is given by $G(i\omega)$ and it has the radius $|G(i\omega)W_u(\omega)|$.

A related type of unstructured uncertainty is additive unstructured uncertainty. Both additive and multiplicative unstructured uncertainties are limited to stable perturbations of the process, i.e. it is assumed that the number of unstable open loop poles is constant. This is a drawback with these types of uncertainty descriptions. It is overcome by using the coprime factorization unstructured uncertainty [Vidyasagar, 1985]. This considers unstructured perturbations of both the numerator and the denominator of the process. The process is described as a fraction of coprime transfer functions with all poles in a prescribed region.

The criteria in (3.1) are established for a process that is accurately described by the model. If this is not the case, the real closed loop system may be unstable or may perform badly. It must of course be required that the closed loop system is stable for all processes (3.2). This is referred to as *robust stability*. There exists a simple criterion for this [Doyle and Stein, 1981].

LEMMA 3.1—Robust Stability
The process (3.2) is controlled by the controller (2.3). The closed loop system is stable for unstructured multiplicative uncertainties $W_u(\omega)$ provided that the closed loop system is stable for $W_u = 0$ and that

$$W_u(\omega)|\mathcal{T}(i\omega)| < 1, \qquad \forall\omega \tag{3.3}$$

where $\mathcal{T}$ is defined in (2.5). □

Robust stability is not sufficient for a nice closed loop performance when the process is uncertain. A stronger requirement is *robust performance*, which is obtained when the nominal performance criteria (3.1) are satisfied for all processes G^0 in (3.2), described by B/A and W_u [Morari and Doyle, 1986]. The following three sub-sections modify the nominal performance criteria (3.1) to establish criteria for robust performance.

Output Disturbance Rejection

Consider first rejection of an output disturbance d. The output due to such a disturbance is given by

$$y_d = \mathcal{S}_o^0 d = \frac{AR}{AR + BG_u S}\, d$$

as was derived in Chapter 2. Assuming that the specifications require that the transfer function $\mathcal{S}_o^0$ based on G^0 should be smaller than $1/W_o$. Hence

$$W_o \left| \frac{AR}{AR + BSG_u} \right| = W_o \left| \frac{AR}{AR + BS(1 + \Delta W_u)} \right| \leq 1, \qquad \forall\omega, \forall|\Delta| \leq 1$$

This inequality should be fulfilled for all frequencies and all uncertainties $|\Delta(\omega)| \leq 1$ for the process G^0. It can be rewritten as

$$W_o \left| \frac{\frac{AR}{AR+BS}}{1 + \frac{BS}{AR+BS}\Delta W_u} \right| \leq 1, \qquad \forall\omega, \forall|\Delta| \leq 1$$

Using the transfer functions in (2.5) the robust performance criterion for output disturbance rejection is

$$W_o \left| \frac{\mathcal{S}_o}{1 + \Delta W_u \mathcal{T}} \right| \leq 1 \qquad \forall\omega, \forall|\Delta| \leq 1 \tag{3.4}$$

A necessary and sufficient condition for (3.4) is established by a lemma in [Francis, 1988].

LEMMA 3.2—Robust Performance
Assume that Lemma 3.1 holds. A necessary and sufficient condition for robust performance (3.4) is

$$W_o|\mathcal{S}_o| + W_u|\mathcal{T}| \leq 1 \qquad \forall\omega \tag{3.5}$$

Proof: Proof of sufficiency: Lemma 3.1 yields that $W_u|\mathcal{T}| < 1$. Then (3.5) implies

$$1 \geq \frac{W_o|\mathcal{S}_o|}{1 - W_u|\mathcal{T}|} \geq \frac{W_o|\mathcal{S}_o|}{|1 + \Delta W_u \mathcal{T}|} \qquad \forall\omega, \forall|\Delta| \leq 1$$

To prove the necessity assume that $W_u|\mathcal{T}| < 1$ and that (3.4) holds. Pick a frequency ω_0 for which

$$\frac{|W_o\mathcal{S}_o|}{1 - |W_u\mathcal{T}|}$$

obtains its maximum. Pick Δ_0 such that $|1 + \Delta_0 W_u\mathcal{T}| = 1 - |W_u\mathcal{T}|$. Then

$$\max_\omega \frac{W_o(\omega)|\mathcal{S}_o(i\omega)|}{1 - W_u(\omega)|\mathcal{T}(i\omega)|} = \frac{W_o(\omega_0)|\mathcal{S}_o(i\omega_0)|}{|1 + \Delta_0(i\omega_0)W_u(\omega_0)\mathcal{T}(i\omega_0)|} \leq 1$$

and (3.5) holds. □

The inequality (3.5) may be written

$$W_o(\omega) + W_u(\omega)|L(i\omega)| \leq |1 + L(i\omega)| \qquad \forall\omega \tag{3.6}$$

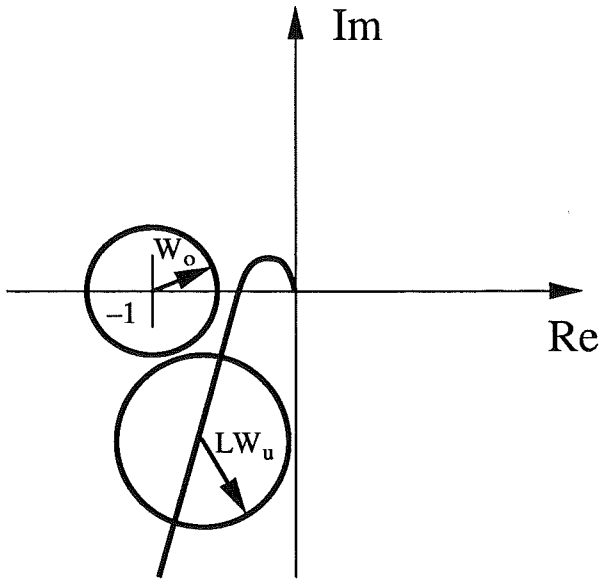


Figure 3.1 Nyquist diagram interpretation of robust performance for output disturbance rejection

where $L = BS/AR$ is the loop transfer function. This inequality and thus Lemma 3.2 has a nice interpretation in the Nyquist diagram. See Figure 3.1. The right-hand side of (3.6) is the distance in the complex plane between $L(i\omega)$ and -1. The unstructured uncertainty bound $W_u(\omega)$ limits the real loop transfer function for the frequency ω to be inside a circle with center in $L(i\omega)$ and with radius $W_u(\omega)L(i\omega)$. This is represented by the lower circle in Figure 3.1. The nominal performance criterion is represented by the circle with center in -1 and with radius $W_o(\omega)$. The robust performance criterion is satisfied for the frequency ω if these two circles do not intersect. Robust performance is achieved if the circles are disjoint for all frequencies.

Effective disturbance rejection is achieved by a large W_o over a wide frequency interval, i.e. a tight bound on $\mathcal{S}_o$. Large uncertainties correspond to large W_u implying that $\mathcal{T}$ must be kept small. Since $\mathcal{S}_o + \mathcal{T} = 1$, these two requirements can not both be fulfilled for same frequencies. Note that

$$\begin{aligned} 1 &\geq W_o(\omega)|\mathcal{S}_o(i\omega)| + W_u(\omega)|\mathcal{T}(i\omega)| \\ &\geq \min(W_o(\omega), W_u(\omega))(|\mathcal{S}_o| + |\mathcal{T}|) \\ &\geq \min(W_o(\omega), W_u(\omega))(|\mathcal{S}_o + \mathcal{T}|) \\ &\geq \min(W_o(\omega), W_u(\omega)) \qquad \forall\omega \end{aligned}$$

implying that if $\min(W_o(\omega), W_u(\omega)) > 1$ for any ω then no controller exists that fulfills the robust performance criterion.

If $W_u(\omega) < 1, \forall\omega$ the controller gain may be increased to infinity and still satisfy Lemma 3.2. This is never the case for real plants. True plants

always have roll-off for high frequencies which implies that $|B(i\omega)/A(i\omega)|$ is small for high frequencies. To have W_u less than 1 for high frequencies thus requires a very accurate plant description.

Still there exist no analytic method to obtain a controller that satisfies the condition of Lemma 3.2. A sufficient condition for SISO systems that allows the problem to be solved using H_∞ methods, is given in [Francis, 1988]. Alternatively the problem may be solved numerically using convex programming as suggested in [Boyd and Barratt, 1991].

Input Disturbance Rejection

Consider the case of reducing the effects in the output y of an input disturbance l. For a process without uncertainty, l may be expressed by an equivalent output disturbance $d = G^0 l$. Then, the desired input disturbance rejection is specified by the weight $W_o = W_i|G^0|$ in the condition for output disturbance rejection. In presence of model uncertainty, the uncertainty must be accounted for in the expression for an equivalent disturbance. This is considered in the derivation of the design criterion for robust input disturbance rejection here.

The contribution in the output caused by l is given by

$$y_l = S_i^0 l = \frac{BG_u R}{AR + BG_u S} l$$

Using the transfer functions in (2.5) the robust performance criterion for input disturbance rejection is

$$W_i \frac{|S_i(1 + \Delta W_u)|}{|1 + \Delta W_u \mathcal{T}|} \leq 1, \qquad \forall\omega, \forall|\Delta| \leq 1 \tag{3.7}$$

provided that Lemma 3.1 holds.

LEMMA 3.3
Assume that Lemma 3.1 is satisfied. A sufficient condition for robust performance (3.7) is that the inequality

$$W_i(1 + W_u)|S_i| + W_u|\mathcal{T}| \leq 1 \tag{3.8}$$

holds for all ω.

Proof: From Lemma 3.1 it follows that $W_u|\mathcal{T}| < 1$. If (3.8) holds, then

$$1 \geq \frac{W_i(1 + W_u)|S_i|}{1 - W_u|\mathcal{T}|} \geq \frac{W_i(1 + W_u)|S_i|}{|1 + \Delta W_u \mathcal{T}|} \geq \frac{W_i|1 + \Delta W_u||S_i|}{|1 + \Delta W_u \mathcal{T}|}$$

for all frequencies ω and for all $|\Delta| \leq 1$. □

It may be noted that (3.8) is not a necessary condition for (3.7). Some conservatism is incorporated since $|1 + \Delta W_u|$ not necessarily achieves its maximum for the same Δ as $|1 + \Delta W_u \mathcal{T}|$ achieves its minimum. However, this happens for frequencies where $\mathcal{T}(i\omega)$ is a negative real number. Thus, the condition in this lemma is tight for these frequencies.

Measurement Noise Attenuation

A criterion for robust measurement noise attenuation is derived in the same way as for output disturbance rejection.

LEMMA 3.4
Assume that Lemma 3.1 is satisfied. A necessary and sufficient condition for robust measurement noise attenuation is

$$W_n|\mathcal{S}_n| + W_u|\mathcal{T}| \leq 1 \quad \forall\omega \tag{3.9}$$

Proof: The proof is analogous to Lemma 3.2. □

Under mild assumptions on the process, Lemma 3.4 implies the robust stability condition in Lemma 3.1.

LEMMA 3.5
Assume that $W_n(\omega) > 0, \forall\omega$ and that $W_u(\omega_0) \neq 1, \forall\omega_0$ such that $A(i\omega_0) = 0$. Then (3.9) is a sufficient condition for $W_u|\mathcal{T}| < 1$.

Proof: For $\omega : |\mathcal{S}_n(i\omega)| > 0$ it follows that $W_u(\omega)|\mathcal{T}(i\omega)| < 1$. Consider then $\omega_0 : |\mathcal{S}_n(i\omega_0)| = |AS/(AR + BS)| = 0$. If $S(i\omega_0) = 0$ then $\mathcal{T}(i\omega) = 0$ and the inequality holds. Otherwise $A(i\omega_0) = 0$ and (3.9) is $W_u|BS/(0 + BS)| \leq 1$. Since $W_u(\omega_0) \neq 1$, then (3.9) implies $W_u(\omega_0)|\mathcal{T}(i\omega_0)| < 1$. □

Summary

Conditions for closed loop robust stability and robust performance have been established for the uncertain process (3.2). If the nominal performance criterion (3.1) holds, the assumptions in Lemma 3.5 are satisfied, and if

$$\begin{aligned} W_o|\mathcal{S}_o| + W_u|\mathcal{T}| &\leq 1 \\ W_i(1 + W_u)|\mathcal{S}_i| + W_u|\mathcal{T}| &\leq 1 \\ W_n|\mathcal{S}_n| + W_u|\mathcal{T}| &\leq 1 \end{aligned}$$

for all ω, then closed loop robust stability and robust performance are obtained.

3.3 Structured and Unstructured Uncertainties

This section presents process models with both structured and unstructured uncertainties. The feedback system was introduced in Chapter 2. Assume that the real process is described by the transfer function

$$G^0 = G_s G_u \tag{3.10}$$

where $G_s(s)$ is the ratio of the two polynomials B and A. These polynomials describe a parametric transfer function model. Uncertainty is represented as variations in the polynomial coefficients in known intervals. This type of uncertainty is referred to as *structured uncertainty*. Furthermore G_u is a stable transfer function of the same type as discussed in (3.2). This describes the unstructured uncertainty.

Model with Structured Uncertainty

The parametric part of the process model (3.10) is described by the strictly proper rational transfer function

$$G_s(s,p) = \frac{B(s,p)}{A(s,p)} = \frac{B_{oo}(s)B_p(s,p)}{A_{oo}(s)A_p(s,p)} \tag{3.11}$$

where B_{oo}, A_{oo}, B_p and A_p are polynomials of finite degrees in the Laplace variable s. The polynomials B_{oo} and A_{oo} are accurately known, i.e. their coefficients have no interval variation. The degrees of the polynomials B_p and A_p are known and

$$B_p(s,p) = b_0 s^{d_b} + b_1 s^{d_b-1} + \cdots + b_{d_b}$$
$$A_p(s,p) = a_0 s^{d_a} + a_1 s^{d_a-1} + \cdots + a_{d_a}$$

are described by a vector containing their coefficients

$$p_{ab} = (\, a_0 \quad \cdots \quad a_{d_a} \quad b_0 \quad \cdots \quad b_{d_b} \,)^T \in \mathbb{R}^{d_{ab}}$$

The elements of this vector belong to certain given intervals. Sometimes the variations in different elements are correlated. A better description of the structured uncertainty can be obtained if the variations in p_{ab} can be described as an affine mapping $\mathbb{R}^{d_p} \mapsto \mathbb{R}^{d_{ab}}$

$$p_{ab}(p) = p_{ab0} + M_0\, p \tag{3.12}$$

where the vector p is known from modeling or identification to belong to a hyperrectangle in $\mathbb{R}^{d_p}$. The simplest case of variation is when $p_{ab0} = 0$ and $M_0 = I$. Then the uncertainty is directly expressed in the polynomial coefficients.

Two sets are introduced to describe the transfer function $G_s(s,p)$.

DEFINITION 3.1—Set of Uncertain Parameters
Let p be as in (3.12). The transfer function family $G_s(s,p)$ with structured uncertainty is defined by the set

$$\mathcal{P} = \left\{p : p_i \in (p_i^-, p_i^+) \quad i = 1, \ldots, d_p\right\}$$

where p_i^- and p_i^+ are the lower and the upper limits of the i:th component of p. □

DEFINITION 3.2—Set of Vertices
Let the set $\mathcal{P}_c$ contain the $d_{pc} = 2^{d_p}$ vertices of the set $\mathcal{P}$. □

Any element in the set $\mathcal{P}$ may be formed as a convex combination of elements in the set $\mathcal{P}_c$. Design complexity may be reduced considerably if only the elements in $\mathcal{P}_c$ has to be considered for the design. This will be pursued later using properties of convex functions. It may be noted that variations in many parameters will make the set $\mathcal{P}_c$ very large. For computational reasons it is desirable to keep this set reasonably small.

Discussion

One reason for introducing structured uncertainty instead of unstructured uncertainty is that it is a more accurate way of modeling situations where the uncertainty can be related to parameter variations. Attempts to describe such variations as unstructured uncertainty results in an unnecessarily conservative design since phase information is discarded. The uncertainty displayed in the Nyquist diagram shows this. The region of $G(i\omega,p)$ in the Nyquist plane may have any geometry in the structured uncertainty case. If unstructured uncertainty is used to describe the same type of uncertainty, a circle is drawn, that encircles the region of uncertainty.

Another reason for using structured uncertainty is that unstructured uncertainty requires that perturbations are such that the number of unstable poles of the process remains the same under the perturbations. With unstructured uncertainty it is thus not possible to capture the case when a process changes from being stable to being unstable due to the perturbations.

Unmodeled dynamics is normally present at high frequencies and is preferably described using unstructured uncertainty. On the other hand, process variations at low frequencies are often adequately described by structured uncertainty. It may be very conservative to use unstructured uncertainty to describe this type of model uncertainty. A combination of the two types of uncertainties is proposed. This idea has also been proposed in [Wei and Yedavalli, 1989].

An example describes how a model with both structured and unstructured uncertainties may be constructed.

EXAMPLE 3.1—Ship Dynamics
A linearized model of a ship moving under constant velocity is given in [Åström, 1990, pp. 41–42]. The transfer function from rudder angle to yaw angle is

$$G^0(s) = \frac{(b_0 s + 1)b_1}{s(s + a_0)(s + a_1)}$$

The parameters of this model depend on operating conditions like speed, trim, and loading. Five different operating conditions k are considered with the parameters given in the table below.

k	b_0	b_1	a_0	a_1
1	1.07	0.75	1.96	−0.70
2	1.05	0.74	1.66	−0.59
3	0.93	0.85	1.86	−0.47
4	0.71	1.29	2.02	−0.21
5	0.89	1.83	2.35	0.05

Since parameter a_1 changes sign the number of open loop unstable poles is not constant. The behavior of the process, thus, can not be described using only unstructured uncertainty. The essential parametric variations occur in the process gain b_1 and in the pole a_1. These variations are related and can approximately be described by the linear relation $b_1 \approx 1.6 + 1.3a_1$. One way to model this process and its uncertainty is to use the model in (3.11) with

$$\frac{B(s,p)}{A(s,p)} = \frac{B_{oo}(s)B_p(s,p)}{A_{oo}(s)A_p(s,p)} = \frac{0.9s + 1}{s(s + 2.00)} \cdot \frac{b_1}{s + a_1} \tag{3.13}$$

The structured uncertainty is described by the coefficient variation

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1.6 \end{pmatrix} + \begin{pmatrix} 1 \\ 1.3 \end{pmatrix} p \qquad -0.70 \leq p \leq 0.05$$

Hence, the set of vertices has two elements

$$\mathcal{P}_c = \{ -0.70 \quad 0.05 \}$$

The accuracy of the model (3.13) is described by unstructured uncertainty. A bound on the unstructured uncertainty can be obtained in the following way. Consider one of the operating conditions k. For each k assign a value to a_1 from the table. The gain b_1 is then estimated from a_1 using the

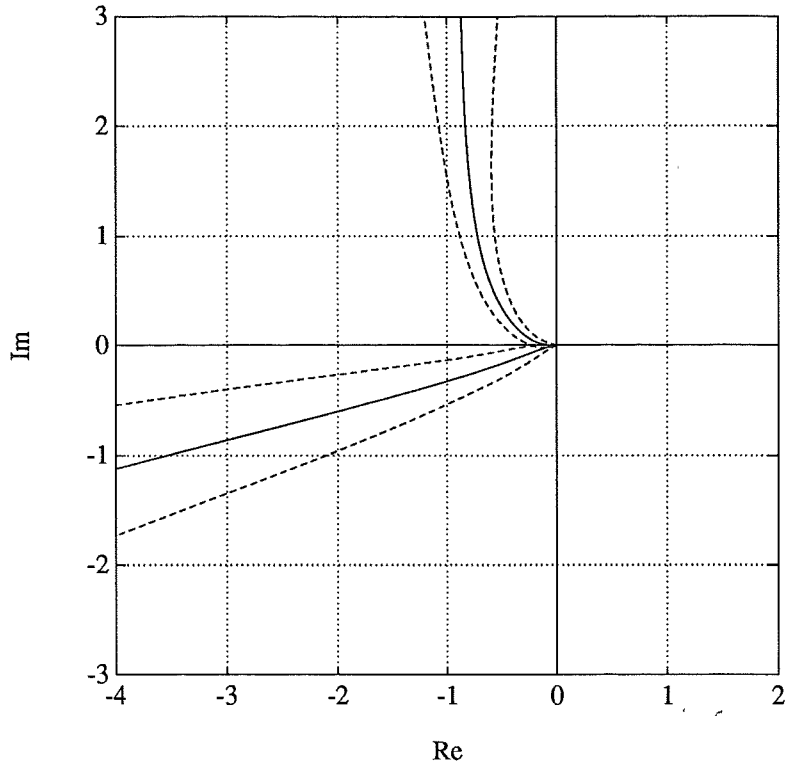


Figure 3.2 Frequency responses $B(i\omega,p)/A(i\omega,p)$ and corresponding regions describing unstructured uncertainties for the two models in $\mathcal{P}_c$ in Example 3.1.

linear approximation. The error in the gain and the variations in the other pole and in the zero gives an error in the model. These errors are covered by unstructured uncertainty which is bounded by

$$W_u(\omega) = \max_k \left| \frac{G^{0k}(i\omega) - \dfrac{B(i\omega,p_k)}{A(i\omega,p_k)}}{\dfrac{B(i\omega,p_k)}{A(i\omega,p_k)}} \right|$$

A simple curve fit gives

$$W_u(\omega) = \left| \frac{0.3i\omega + 0.12}{i\omega + 1.2} \right|$$

The solid lines in Figure 3.2 show the frequency responses for the two models $B(i\omega,p)/A(i\omega,p)$ in (3.13) for $p \in \mathcal{P}_c$. The bound $W_u(\omega)$ allows variations in each of these frequency responses between the dashed lines surrounding the curve for the respective frequency response. □

3.4 Design Criteria

In this section it is shown how the design criteria for processes with combined structured and unstructured uncertainties can be formulated in a manner analogous to Section 3.2. This requires a reformulation of the design criteria. The key idea is to express them as constrained convex functions of the process parameter p.

Controller Parameterization

A controller was presented in Chapter 2. The parameterization (2.12) is a convenient way to characterize all proper stabilizing controllers having the factor A_r in the denominator when the process is accurately described by the fraction B/A. Here the problem is extended to a family of transfer function models described by the set $\mathcal{P}$. One element in this set is regarded as the nominal model.

DEFINITION 3.3—Nominal Model
The nominal process model is chosen as one element $p_0 \in \mathcal{P}$ corresponding to $B_0(s) = B(s,p_0)$ and $A_0(s) = A(s,p_0)$, with $\deg A_0(s) = \max \deg A(s,p)$ for $p \in \mathcal{P}$. □

The nominal controller S_0/R_0, is designed to give desired closed loop properties for the nominal model. The controller used in the sequel is given by

$$\frac{S}{R} = \frac{S_0 + QA_rA_0}{A_r(R_0' - QB_0)} \tag{3.14}$$

where $Q = N/D$ is a stable rational transfer satisfying the degree condition (2.11). This form guarantees that S and R do not have common zeros in the closed right half plane. It should also be noted that this controller only guarantees stability for the closed loop system based on the nominal model. For other transfer functions in the family $\mathcal{P}$, the closed loop system may, however, be unstable. Criteria for stability are implied by the robust performance criteria given later.

Throughout the rest of this chapter it will be assumed that the function Q is fixed. Later, in Chapter 4, the function Q will be chosen to obtain desired closed loop performance objectives.

Output Disturbance Rejection

Incorporation of structured uncertainty in the robust performance criteria is pursued for the output disturbance rejection case. With straightforward modifications it is applicable for input disturbance rejection and measurement noise attenuation.

Introduce the functions

$$f(\omega,p) = W_o(\omega)|A(i\omega,p)R(i\omega)| + W_u(\omega)|B(i\omega,p)S(i\omega)|$$

and

$$A_c(s,p) = A(s,p)R(s) + B(s,p)S(s) \tag{3.15}$$

where $B(s,p)$ and $A(s,p)$ are defined in (3.11) and $R(s)$ and $S(s)$ are defined in (3.14). Then, the robust performance criterion (3.5) in Lemma 3.2 for output disturbance rejection may be written as

$$f(\omega,p) - |A_c(i\omega,p)| \leq 0, \qquad \forall\omega,\forall p \in \mathcal{P} \tag{3.16}$$

Notice that the functions $f(\omega,p)$ and $A_c(s,p)$ depend on the rational function Q. Furthermore, A_c is not a polynomial since it contains Q.

A drawback of using the inequality (3.16) as a design criterion is that the left hand side is not necessarily a convex function in p, since the second term $-|A_c(i\omega,p)|$ is concave in p. A modified criterion which is convex will therefore be formulated. To do this introduce a bounding function $\psi(\omega,p)$ satisfying

$$\psi(\omega,p) \leq |A_c(i\omega,p)|, \qquad \forall\omega,\forall p \in \mathcal{P} \tag{3.17}$$

The function $\psi(\omega,p)$ will be constructed below. It should be concave in p. The function

$$g(\omega,p) = f(\omega,p) - \psi(\omega,p) \tag{3.18}$$

will then have the desired convexity properties. Using (3.18) the design objective may be expressed by the inequality

$$g(\omega,p) \leq 0 \tag{3.19}$$

that should hold for all ω and for all $p \in \mathcal{P}$. The condition (3.19) is then a sufficient condition for the robust performance criterion (3.16).

If (3.19) holds for the vertices $p \in \mathcal{P}_c$, then it follows from the convexity of $g(\omega,p)$ that (3.19) will be satisfied for all convex combinations of the elements in $\mathcal{P}_c$, namely the set $\mathcal{P}$. The condition

$$g(\omega,p) \leq 0, \qquad \forall\omega,\forall p \in \mathcal{P}_c$$

is thus a sufficient condition for robust output disturbance rejection. The fact that $g(\omega,p)$ is convex in p thus reduces the complexity of the design problem considerably.

Construction of a Bounding Function

The use of a bounding function $\psi(\omega, p)$ was suggested above. This function is constructed here.

Introduce the structured uncertainty as perturbations of the nominal model

$$\delta_B(s,p) = B(s,p) - B_0(s)$$
$$\delta_A(s,p) = A(s,p) - A_0(s)$$

The function (3.15) is split into

$$A_c(s,p) = A_{c0}(s) + \delta_{Ac}(s,p) \tag{3.20}$$

where the first term

$$A_{c0}(s) = A_0(s)A_r(s)R_0'(s) + B_0(s)S_0(s) \tag{3.21}$$

is independent of p and Q. The second term

$$\delta_{Ac}(s,p) = \delta_A A_r R_0' + \delta_B S_0 + Q A_r \left(A_0 \delta_B - B_0 \delta_A\right) \tag{3.22}$$

is affine in p. Also notice that δ_{Ac} is affine in Q. This important fact will be useful in Chapter 4.

The magnitude of the function $A_c(i\omega, p)$ satisfies

$$\begin{aligned} |A_c| = |A_{c0} + \delta_{Ac}| &= |A_{c0}| \sqrt{\left(1 + \mathrm{Re}\,\frac{\delta_{Ac}}{A_{c0}}\right)^2 + \left(\mathrm{Im}\,\frac{\delta_{Ac}}{A_{c0}}\right)^2} \\ &\geq |A_{c0}| \left(1 + \mathrm{Re}\,\frac{\delta_{Ac}}{A_{c0}}\right) \end{aligned} \tag{3.23}$$

The bounding function ψ is now chosen as

$$\psi(\omega, p) \equiv |A_{c0}(i\omega)| \left(1 + \mathrm{Re}\,\frac{\delta_{Ac}(i\omega, p)}{A_{c0}(i\omega)}\right) \tag{3.24}$$

It follows from (3.23) that ψ satisfies (3.19). Further, $\psi(s,p)$ is affine in p, i.e. concave. This means that g is convex in p.

The use of ψ instead of $|A_c|$ implies that the constraints are unnecessarily conservative. The amount conservatism can be expressed by the quantity

$$\delta_\psi(\omega, p) = \frac{|A_c(i\omega, p)| - \psi(\omega, p)}{|A_{c0}(i\omega)|} \tag{3.25}$$

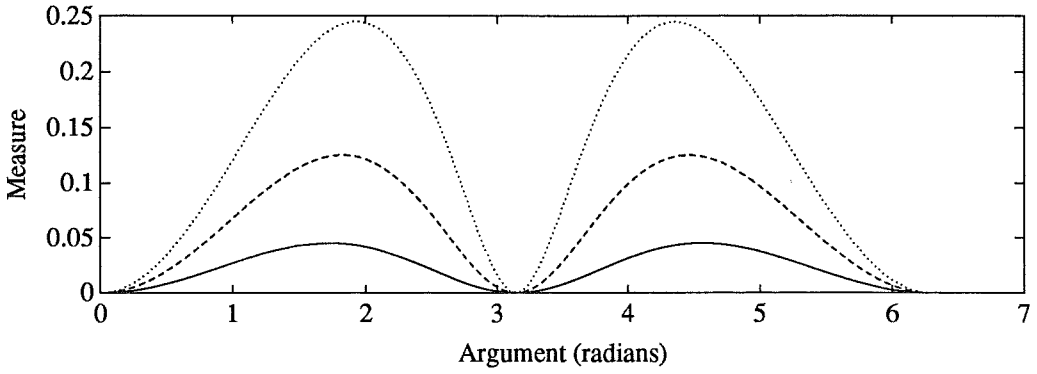


Figure 3.3 Measure of conservatism δ_ψ as function of the argument ϕ for $r = 0.3$ (solid line), $r = 0.5$ (dashed line), and $r = 0.7$ (dotted line) in Example 3.2.

The normalization intends to make the first term approximately one. It corresponds to the scaling that will be used in the optimization problems in Chapter 4. The function $\delta_\psi(s,p)$ is convex in p. The worst case is then given by

$$\max_{\omega,p} \delta_\psi(\omega,p), \quad \forall p \in \mathcal{P}_c \tag{3.26}$$

A δ_ψ which is considerably smaller than one implies that the inequality in (3.23) is close and the use of ψ is not particularly conservative. An example will explore this.

EXAMPLE 3.2
Consider the ratio

$$\frac{A_c}{A_{c0}} = 1 + re^{i\phi}$$

with the interpretations $r = |\delta_{Ac}/A_{c0}|$ and $\phi = \arg(\delta_{Ac}/A_{c0})$. Figure 3.3 shows the measure δ_ψ in (3.25) as function of the argument $\phi \in (0, 2\pi)$ for $r = 0.3$ (solid line), for $r = 0.5$ (dashed line), and for $r = 0.7$ (dotted line). □

Stability

A basic demand for robustness is that the closed loop system has all poles strictly in the left half plane for all $p \in \mathcal{P}$. The closed loop poles correspond to the zeros of the function $A_c(s,p)$ in (3.20). Closed loop stability for a family of processes with structured uncertainty is established by the following theorem.

THEOREM 3.1—Stability of a Family of Transfer Functions
Consider the functions in (3.20)–(3.22). If the polynomial $A_{c0}(s)$ has all roots strictly in the left half plane and if $\psi(\omega, p) > 0, \forall\omega$ and $\forall p \in \mathcal{P}_c$, then $A_c(s,p)$ has all zeros in the open left half plane for all $p \in \mathcal{P}$.

Proof: The polynomial $A_{c0}(s)$ has no roots in the closed right half plane, $\mathbb{C}^+$. Then it is possible to write $A_c = A_{c0}A_{c\delta}$ with

$$A_{c\delta}(s,p) = 1 + \frac{\delta_{Ac}(s,p)}{A_{c0}(s)}$$

for $s \in \mathbb{C}^+$ and it is sufficient to check if $A_{c\delta}(s,p)$ has zeros in $\mathbb{C}^+$.

The condition on ψ is equivalent to

$$1 + \operatorname{Re}\frac{\delta_{Ac}(i\omega,p)}{A_{c0}(i\omega)} > 0, \qquad \forall\omega, \forall p \in \mathcal{P}_c \tag{3.27}$$

which implies that the map $A_{c\delta}(s,p)$, for a contour $s \in \Gamma$ encircling the closed right half plane, always has a positive real part $\forall p \in \mathcal{P}_c$. The function $\delta_{Ac}(s,p)$ is affine in p and thus also $\operatorname{Re} A_{c\delta}(s,p)$. Then (3.27) is satisfied $\forall p \in \mathcal{P}$. The degree condition in Definition 3.3 yields that $\deg A_0 \geq \deg \delta_A$ and $\deg A_0 > \deg \delta_B$. Then δ_{Ac}/A_{c0} is proper and $A_{c\delta}(s,p)$ is bounded for $s \in \mathbb{C}^+$. Thus the map $A_{c\delta}(s,p)$ for the contour $s \in \Gamma$ will never encircle the origin for any $p \in \mathcal{P}$. Since $A_{c0}(s)$ and $D(s)$ have no roots in $\mathbb{C}^+$, $A_{c\delta}(s,p)$ will have no poles in $\mathbb{C}^+$. Thus $A_{c\delta}(s,p)$ has no zeros in $\mathbb{C}^+$ from the principle of argument variation. □

Robust Performance

The robust performance criteria for processes with both structured and unstructured uncertainties are summarized in three theorems. They extend the results on robust performance in Section 3.2.

THEOREM 3.2—Robust Output Disturbance Rejection
Let the process (3.10) be controlled by the fixed controller (3.14). The structured uncertainty is described by $p \in \mathcal{P}$ and the unstructured uncertainty is described by $W_u(\omega)$. The performance objective for output disturbance rejection is defined by $W_o(\omega)$. Assume that the closed loop system is stable for all uncertainties $p \in \mathcal{P}$ and W_u.

Define

$$\begin{aligned} f_o(\omega,p) &= \left|\frac{A(i\omega,p)R(i\omega)}{A_{c0}(i\omega)}\right| \\ f_{u\psi}(\omega,p) &= W_u(\omega)\left|\frac{B(i\omega,p)S(i\omega)}{A_{c0}(i\omega)}\right| - \left(1 + \operatorname{Re}\frac{\delta_{Ac}(i\omega,p)}{A_{c0}(i\omega)}\right) \end{aligned} \tag{3.28}$$

If the function

$$g_o(\omega,p) = W_o(\omega)f_o(\omega,p) + f_{u\psi}(\omega,p) \leq 0 \tag{3.29}$$

for all ω and for all $p \in \mathcal{P}_c$, then the condition (3.5) in Lemma 3.2 holds for all processes within the uncertainty region.

Proof: The condition (3.29) is equivalent to (3.19) for $p \in \mathcal{P}_c$. The function g_o is convex in p, and hence it is sufficient to satisfy $g_o(\omega,p) \leq 0$ for the vertices, $p \in \mathcal{P}_c$ to yield that $g_o(\omega,p) \leq 0, \forall p \in \mathcal{P}$. Then (3.19) holds $\forall p \in \mathcal{P}$. This is a sufficient condition for the inequality (3.5) in Lemma 3.2. □

The condition in Theorem 3.2 guarantees that the closed loop system satisfies the nominal performance condition for all processes within the region of uncertainty. Thus, (3.29) is a sufficient condition for robust performance. The use of convex functions in (3.29) reduces complexity, since not all processes with structured uncertainty have to be investigated.

In Theorem 3.2, it was assumed that the closed loop system is stable for all uncertain processes under consideration. This assumption will, however, be eliminated in a forthcoming theorem. First, analog conditions for robust input disturbance rejection and robust measurement noise attenuation are presented.

THEOREM 3.3—Robust Input Disturbance Rejection
Let the process (3.10) be controlled by the fixed controller (3.14). The structured uncertainty is described by $p \in \mathcal{P}$ and the unstructured uncertainty is described by $W_u(\omega)$. The performance objective for input disturbance rejection is defined by $W_i(\omega)$. Assume that the closed loop system is stable for all uncertainties $p \in \mathcal{P}$ and W_u.

Define

$$f_i(\omega,p) = (1 + W_u(\omega)) \left| \frac{B(i\omega,p)R(i\omega)}{A_{c0}(i\omega)} \right|$$

$$f_{u\psi}(\omega,p) = W_u(\omega) \left| \frac{B(i\omega,p)S(i\omega)}{A_{c0}(i\omega)} \right| - \left(1 + \mathrm{Re}\, \frac{\delta_{Ac}(i\omega,p)}{A_{c0}(i\omega)} \right)$$

If the function

$$g_i(\omega,p) = W_i(\omega)f_i(\omega,p) + f_{u\psi}(\omega,p) \leq 0 \tag{3.30}$$

for all ω and for all $p \in \mathcal{P}_c$, then the condition (3.8) in Lemma 3.3 holds for all processes within the uncertainty region.

Proof: Analogous to the proof of Theorem 3.2. □

THEOREM 3.4—Robust Measurement Noise Attenuation
Let the process (3.10) be controlled by the fixed controller (3.14). The structured uncertainty is described by $p \in \mathcal{P}$ and the unstructured uncertainty is described by $W_u(\omega)$. The performance objective for measurement noise attenuation is defined by $W_n(\omega)$. Assume that the closed loop system is stable for all uncertainties $p \in \mathcal{P}$ and W_u.

Define

$$f_n(\omega,p) = \left| \frac{A(i\omega,p)S(i\omega)}{A_{c0}(i\omega)} \right|$$

$$f_{u\psi}(\omega,p) = W_u(\omega) \left| \frac{B(i\omega,p)S(i\omega)}{A_{c0}(i\omega)} \right| - \left(1 + \mathrm{Re}\, \frac{\delta_{Ac}(i\omega,p)}{A_{c0}(i\omega)} \right)$$

If the function

$$g_n(\omega,p) = W_n(\omega) f_n(\omega,p) + f_{u\psi}(\omega,p) \leq 0 \tag{3.31}$$

for all ω and for all $p \in \mathcal{P}_c$, then the condition (3.9) in Lemma 3.4 holds for all processes within the uncertainty region.

Proof: Analogous to the proof of Theorem 3.2. □

Main Result

In the three Theorems 3.2, 3.3, and 3.4 it was assumed that the closed loop system is stable for all uncertainties described by $p \in \mathcal{P}$ and W_u. Theorem 3.1 guarantees closed loop stability for all $p \in \mathcal{P}$. Now it only remains to establish a condition that guarantees that the unstructured uncertainty will not destabilize the closed loop system. This condition will also be sufficient for the inequality $\psi > 0$ in Theorem 3.1 to hold.

THEOREM 3.5—Robust Performance and Robust Stability
Let the process (3.10) be controlled by the fixed controller (3.14). The structured uncertainty is described by $p \in \mathcal{P}$ and the unstructured uncertainty is described by $W_u(\omega)$. Assume that $B(s,p)$ and $A(s,p)$ have no common zeros on the imaginary axis and that all weights W_o, W_i, W_n, and W_u are positive $\forall\omega$. Assume further, for any $p \in \mathcal{P}$ that $\forall\omega_0$ such that $A(p,i\omega_0) = 0$ it follows that $W_u(\omega_0) \neq 1$. If the conditions

$$\begin{aligned} g_o(\omega,p) &\leq 0 \\ g_i(\omega,p) &\leq 0 \\ g_n(\omega,p) &\leq 0 \end{aligned} \tag{3.32}$$

in the theorems 3.2, 3.3, and 3.4 hold $\forall\omega$ and $\forall p \in \mathcal{P}_c$ then the closed loop system is stable for all models within the uncertainty region.

Proof: The inequalities above may be written

$$\begin{aligned} W_o \left|\frac{AR}{A_{c0}}\right| + W_u \left|\frac{BS}{A_{c0}}\right| &\leq \left(1 + \mathrm{Re}\,\frac{\delta_{Ac}}{A_{c0}}\right) \\ W_i(1+W_u) \left|\frac{BR}{A_{c0}}\right| + W_u \left|\frac{BS}{A_{c0}}\right| &\leq \left(1 + \mathrm{Re}\,\frac{\delta_{Ac}}{A_{c0}}\right) \qquad \forall\omega, \forall p \in \mathcal{P}_c \\ W_n \left|\frac{AS}{A_{c0}}\right| + W_u \left|\frac{BS}{A_{c0}}\right| &\leq \left(1 + \mathrm{Re}\,\frac{\delta_{Ac}}{A_{c0}}\right) \end{aligned}$$

Since neither B and A nor S and R have common zeros on the imaginary axis, at least one of the terms on the left-hand side is positive. Then

$$1 + \mathrm{Re}\,\frac{\delta_{Ac}(i\omega, p)}{A_{c0}(i\omega)} > 0, \quad \forall\omega, \forall p \in \mathcal{P}_c$$

and stability is obtained from Theorem 3.1.

It remains to show that the unstructured uncertainty not will destabilize the closed loop system, i.e.

$$W_u\,|BS| < |AR + BS| \qquad \forall\omega, \forall p \in \mathcal{P} \tag{3.33}$$

The condition for noise attenuation (3.31) yields

$$W_n\,|AS| + W_u\,|BS| \leq |A_{c0}| \left(1 + \mathrm{Re}\,\frac{\delta_{Ac}}{A_{c0}}\right) \leq |AR + BS| \qquad \forall\omega, \forall p \in \mathcal{P} \tag{3.34}$$

For $\omega : W_n|AS| > 0$ it directly follows that (3.34) implies (3.33). Consider $\omega_0 : |A(i\omega_0, p)S(i\omega_0)| = 0$. Theorem 3.1 yields that $A_c(i\omega, p) \neq 0$, then one of $A(i\omega_0, p) = 0$ and $S(i\omega_0) = 0$ must be nonzero. If $S(i\omega_0) = 0$, (3.33) holds. For $A(i\omega_0, p) = 0$ and if the bounding function ψ is conservative, the right inequality in (3.34) is strict, which implies (3.33). For $A(i\omega_0, p) = 0$ and a nonconservative ψ, (3.34) gives $W_u|BS| \leq |BS|, \forall p \in \mathcal{P}$ which holds for $W_u(\omega_0) \leq 1$. But $W_u(\omega_0) \neq 1$ and (3.33) holds. □

The inequalities (3.32) will be used as constraints in an optimization problem. This will be treated in Chapter 4.

3.5 Design of Feedforward

The feedback design above determines the complementary sensitivity function. The command signal response may be shaped further by a feedforward $B_{\rm ff}(s)/A_{\rm ff}(s)$, since the output signal is given by

$$y = \frac{B_{\rm ff}}{A_{\rm ff}} T^0 r$$

where $T^0 = T^0(s,p,\Delta)$ is the complementary sensitivity function based on the real process G^0 in (3.10). The feedforward ought to be chosen such that variations in the response to command signals are limited although $T^0(s,p,\Delta)$ may have different characteristics due to the process uncertainties $p \in \mathcal{P}$ and Δ.

One way to obtain well defined command signal responses is to limit the frequency content in the signal

$$y_r = \frac{B_{\rm ff}}{A_{\rm ff}} r$$

by chosing $B_{\rm ff}(s)/A_{\rm ff}(s)$ as a low pass filter. Introduce the fictive error signal

$$e_f = y_r - y = S_o^0 y_r = \frac{B_{\rm ff}}{A_{\rm ff}} S_o^0 r$$

where $S_o^0 = S_o^0(s,p,\Delta)$ is based on G^0 in (3.10). Then select the feedforward to limit the magnitude of the transfer function from r to e_f, i.e.

$$\left| \frac{B_{\rm ff}(i\omega)}{A_{\rm ff}(i\omega)} S_o^0(i\omega,p,\Delta) \right| \leq \epsilon_f(\omega), \quad \forall\omega, \forall p \in \mathcal{P}, \forall|\Delta| \leq 1 \tag{3.35}$$

for some weight $\epsilon_f(\omega) > 0$. This condition resembles the condition in the robust output feedback problem. An analogous reformulation is

$$\frac{1}{\epsilon_f}\left|\frac{B_{\rm ff}}{A_{\rm ff}}\right|\left|\frac{AR}{A_{c0}}\right| + W_u\left|\frac{BS}{A_{c0}}\right| - \left(\mathrm{Re}\left(\frac{\delta_{Ac}}{A_{c0}}\right)+1\right) \leq 0, \quad \forall\omega, \forall p \in \mathcal{P}_c \tag{3.36}$$

Fulfillment of (3.36) for $p \in \mathcal{P}_c$ is sufficient for (3.36) to hold for $p \in \mathcal{P}$ due to convexity. Given a certain feedforward $B_{\rm ff}/A_{\rm ff}$ (3.36) is a sufficient condition for (3.35). Hence, the responses to command signals will be close for all uncertainties under consideration.

The feedforward must satisfy $B_{ff}(0)/A_{ff}(0) = k$, where $1/k$ is the DC-gain of $\mathcal{T}$ to yield a unit DC-gain of the transfer function form r to y. For an integrating controller it follows that $k = 1$.

A simple way to select the feedforward is to find the smallest T_{ff} for the feedforward

$$\frac{B_{ff}}{A_{ff}} = \frac{k}{(sT_{ff}+1)^{n_{ff}}} \tag{3.37}$$

for which (3.36) holds. Given the process $B(s,p)/A(s,p)$ with unstructured uncertainty $W_u(\omega)$, the nominal model $B_0(s)/A_0(s)$, and the feedback controller $S(s)/R(s)$ it is possible to explicitly calculate $\min T_{ff}$ to make the response for the reference signal as fast as possible while (3.36) holds. This is exemplified in Chapter 4.

3.6 Conclusions

Since a model of a real process is never exact it is useful to characterize the uncertainty of the model. In this chapter it has been proposed to do this using structured and unstructured uncertainties. Design methods for processes with unstructured uncertainty have been developed by [Morari and Doyle, 1986], who showed that robust output disturbance rejection could be expressed by an inequality. This criterion has here been extended to criteria for robust input disturbance rejection and for robust measurement noise attenuation. These results have also been generalized for models with structured and unstructured uncertainties. These criteria for robust performance is the main contribution of this chapter.

A design method that considers these conditions for robust performance will be presented in Chapter 4.

4

Robust Design Using Optimization

Many different requirements have to be considered in control system design. Command signal following, rejection of load disturbances, rejection of measurement noise, and reduction of the effects from model uncertainty are examples of typical requirements. In the previous chapter it has been shown that many requirements can be expressed in terms of inequalities involving convex functions. In this chapter a design method will be developed that considers such inequalities.

There are some methods for control system design that considers model uncertainty explicitly. The H_∞-method [Zames, 1981] and the LQG/LTR approach [Stein and Athans, 1987] deal with unstructured uncertainty only. Quantitative feedback design [Horowitz, 1963] and robust pole placement design [Soh *et al.*, 1987] deal with structured uncertainty. The μ-design [Doyle, 1987], offers a way to deal with complex valued structured perturbations. It is, however, inordinately complex at the moment. A design method that provides closed loop robust stability for a process with a combination of structured and unstructured uncertainties is described in [Wei and Yedavalli, 1989]. Robust performance is, however, not guaranteed.

The method that will be presented in this chapter also considers a combination of structured and unstructured uncertainties. It provides robust performance for the closed loop system. It may be viewed as an extension of the design approach proposed in [Boyd and Barratt, 1991]. A certain formulation of the weight functions W in (3.29), (3.30), and (3.31) provides the

design method to search for a controller that yields, in some sense, the best disturbance rejection or the best noise attenuation.

4.1 The Optimization Problem

Design criteria for processes with structured and unstructured uncertainties were formulated in Chapter 3. It was shown in Theorems 3.2, 3.3, and 3.4 that output disturbance rejection, input disturbance rejection, and measurement noise attenuation could be expressed by inequalities involving the functions $g_o(w,p)$, $g_i(w,p)$, and $g_n(w,p)$ which are convex in p.

The controller was given by (3.14). It is parameterized in terms of the rational function Q. The choice of a good controller Q will now be investigated. First some approximations are made.

The transfer function Q is approximated by

$$Q(s,x) = \frac{N(s,x)}{D(s)} \tag{4.1}$$

where $D(s)$ is a fixed stable polynomial and

$$N(s,x) = x_1 s^{n-1} + x_2 s^{n-2} + \cdots + x_n \tag{4.2}$$

From (2.11) it follows that $\deg N \leq \deg D + \deg R_0' - \deg A_0$. The motivation for this particular choice is that Q is linear in $x \in \mathbb{R}^n$. The controller is now parameterized by x. A similar approach is taken in [Boyd *et al.*, 1988].

It will now be investigated how the functions g_o, g_i, and g_n depend on x. Consider e.g. the transfer function g_o in (3.29). It follows from (3.14) that

$$S = S_0 + QA_rA_0$$
$$R = A_r(R_0' - QB_0)$$

and from (3.22) that

$$\delta_{Ac} = \delta_A A_r R_0' + \delta_B S_0 + QA_r\left(A_0\delta_B - B_0\delta_A\right)$$

The dependence of R, S, and δ_{Ac} on x is thus affine, and it follows from (3.28) and (3.29) that g_o is convex in x.

The function $g_o(\omega,p)$ will now be redefined as $g_o(\omega,p,x)$ to show the dependence on x explicitly. Analog redefinitions give $g_i(\omega,p,x)$ and $g_n(\omega,p,x)$. These functions are all convex in p for fixed x and convex in x for fixed p. The conditions in Chapter 3 on a feasible controller x now becomes

$$\begin{aligned} g_o(\omega,p,x) &= W_o(\omega)f_o(\omega,p,x) + f_{u\psi}(\omega,p,x) \quad &\leq 0 \\ g_i(\omega,p,x) &= W_i(\omega)f_i(\omega,p,x) + f_{u\psi}(\omega,p,x) \quad &\leq 0 \\ g_n(\omega,p,x) &= W_n(\omega)f_n(\omega,p,x) + f_{u\psi}(\omega,p,x) \quad &\leq 0 \end{aligned} \tag{4.3}$$

for all ω and $p \in \mathcal{P}_c$. Since the functions are convex in x, control design can be done by convex optimization. There are many controllers x that satisfy these inequalities. To find a unique controller that gives, in some sense, best performance, introduce an additional constraint function

$$g_y(y,\omega,p,x) = yW_y(\omega)f_y(\omega,p,x) + f_{u\psi}(\omega,p,x) \leq 0 \tag{4.4}$$

where y is a performance measure that will be maximized, $W_y(\omega)$ is a non-negative weight, and f_y is one of the functions f_o, f_i, or f_n in (4.3) depending on the optimization objective. A maximization of y forces $W_y(\omega)f_y(\omega,p,x)$ to be minimized while other constraints are satisfied. The optimization problem is

$$\begin{array}{lll} \text{maximize} & y & \\ \text{subject to} & g_y(y,\omega,p,x) \leq 0 & \\ & g_o(\omega,p,x) \leq 0 & \\ & g_i(\omega,p,x) \leq 0 & \\ & g_n(\omega,p,x) \leq 0 & \forall\omega, \forall p \in \mathcal{P}_c \end{array} \tag{4.5}$$

with constraint functions defined in (4.3) and (4.4). It is important that g_o, g_i, and g_n all are present and have the properties assumed in Theorem 3.5, otherwise closed loop stability can not be assured.

Two examples are used to illustrate how certain design objectives can be formulated as the optimization problem (4.5).

EXAMPLE 4.1—Improved Load Disturbance Rejection
A controller for fast rejection of constant output load disturbances is obtained if $\mathcal{S}_o(i\omega,p,x)$ has small magnitude for low frequencies and is bounded for higher frequencies. Such design criteria are provided by replacing the weight $W_o(\omega)$ by $W_{yo}(\omega,\omega_b)$ in the first constraint in (4.3), i.e.

$$W_{yo}(\omega,\omega_b)f_o(\omega,p,x) + f_{u\psi}(\omega,p,x) \leq 0 \qquad \forall\omega, \forall p \in \mathcal{P}_c$$

The new weight is defined as

$$W_{yo}(\omega,\omega_b) = \max\left(\sqrt{2}(\frac{\omega_b}{\omega})^\sigma, \frac{1}{M_p}\right)$$

where M_p is the maximum peak of the frequency function from disturbance to output, σ is the low frequency slope of $\mathcal{S}_o(i\omega,p,x)$, and ω_b may be interpreted as the closed loop bandwidth. Figure 4.1 shows the inverse of the weight $W_{yo}(\omega,\omega_b)$. Let $\mathcal{S}_o^0(i\omega,p,\Delta,x)$ be the sensitivity function based on the uncertain process G^0 in (3.10). A feasible controller is achieved when

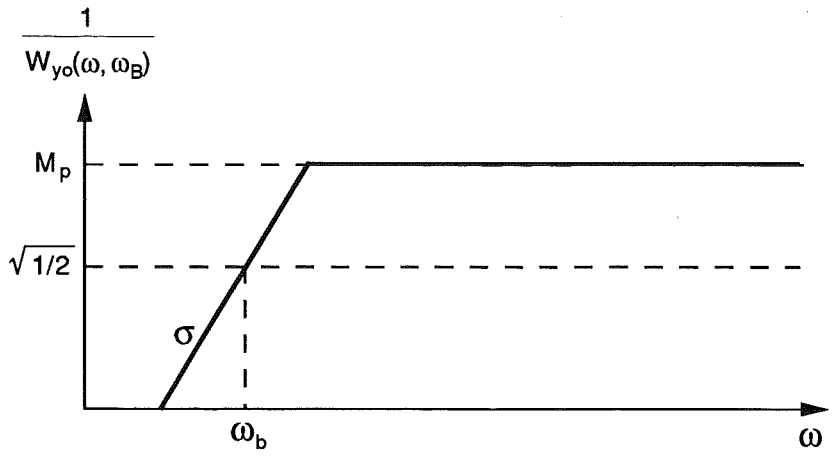


Figure 4.1 Performance bound on disturbance rejection.

$|S_o^0(i\omega, p, \Delta, x)|$ is lower than $1/W_{yo}(\omega, \omega_b)$ for all ω, for all $p \in \mathcal{P}$ and for all $|\Delta| \leq 1$. Increasing ω_b means that the output load disturbance rejection is improved through a lowered low-frequency slope. The fastest load disturbance rejection is achieved for the maximal attainable ω_b subject to the constraints (4.3).

An optimization problem of the form (4.5) for maximally fast load disturbance rejection is obtained by chosing the weights

$$W_y(\omega) = \frac{\sqrt{2}}{\omega^\sigma} \qquad \text{and} \qquad W_o(\omega) = \frac{1}{M_p} \tag{4.6}$$

the function $f_y = f_o$, and the performance measure $y = \omega_b^\sigma$. □

Remark. The weight $W_y(\omega)$ in (4.6) may also be used to optimize the input disturbance rejection. Then $f_y = f_i$, $W_i(\omega) = 1/M_p$, and σ is the low frequency slope of $S_i(i\omega, p, x)$. □

EXAMPLE 4.2—Improved Measurement Noise Attenuation
The measurement noise attenuation is bounded by the third constraint in (4.3). A normal design objective is to minimize the maximum magnitude of the transfer function from measurement noise to control signal. Such a controller is obtained by chosing $f_y = f_n$ and $W_y(\omega) = 1$ in (4.5). At the optimum the magnitude of $S_n^0(i\omega, p, \Delta, x)$ is bounded by the inverse of

$$W_{yn}(\omega, y) = \max(y, W_n(\omega))$$

for all ω, $p \in \mathcal{P}$, and $|\Delta| \leq 1$ according to Theorem 3.4. A large y corresponds to good measurement noise attenuation. □

The optimization problem (4.5) has constraints that are functions of the frequency ω. Such constraint functions lead to a semi-infinite optimization problem [Polak *et al.*, 1984]. To solve such an optimization problem here, it is a approximated by a finite dimensional problem. Based on knowledge of the interesting frequency range for the control system, a set of frequencies

$$\{\omega_j, j = 1, \ldots, m\} \tag{4.7}$$

is selected. It is assumed that if the constraints in (4.3) and (4.4) are satisfied for the frequencies in (4.7) they are satisfied for all frequencies.

The optimization problem (4.5) can now be reformulated using the finite set of frequencies in (4.7). It is

$$\begin{array}{lll} \text{maximize} & y & \\ \text{subject to} & g_y(y, \omega_j, p, x) \leq 0 & \\ & g_o(\omega_j, p, x) \leq 0 & \\ & g_i(\omega_j, p, x) \leq 0 & \\ & g_n(\omega_j, p, x) \leq 0 & j = 1, \ldots, m, \quad \forall p \in \mathcal{P}_c \end{array} \tag{4.8}$$

This optimization problem has $4m2^{d_{Pc}}$ constraint functions $\mathbb{R}^n \mapsto \mathbb{R}$, since each element $p \in \mathcal{P}_c$ produces $4m$ constraints and the set $\mathcal{P}_c$ consists of $2^{d_{Pc}}$ elements.

It is important that the number of frequencies m is considerably larger than the dimension n of x. Then x can not be chosen such that $R(i\omega_j, x) = 0$, $\forall \omega_j$, or $S(i\omega_j, x) = 0$, $\forall \omega_j$ in the controller (3.14). This means that $\forall x$ there exist $f_y(\omega_j, p, x) > 0$ in (4.4). Inspection of (3.28) and (4.4) gives with the same arguments that for some ω_j and $0 < y < \infty$ it follows that $g_y(y, \omega_j, p, x) > 0$ when $|x| \to \infty$. Hence the set of feasible x is bounded.

The controller that is obtained from the solution to (4.8) is often of high order. Hence, model reduction may be used to reduce the controller complexity.

Properties of the Optimization Problem

This section discusses properties of the optimization problem defined in (4.8). It will be shown that the optimization problem has only one extremum. The constraint functions in (4.8) all have a special structure which is given by

$$g_i(y, x) = y f_{1i}(x) + f_{2i}(x), \qquad i \in I = \{1, \ldots, 4m2^{d_{Pc}}\} \tag{4.9}$$

where $f_{1i}(x)$ is convex and non-negative and $f_{2i}(x) > -\infty$ and convex. Constraints in (4.8) without y correspond to $f_{1i} = 0$. For all x, at least one $f_{1i}(x) > 0$ and at least one $g_i(y, x) > 0$ for $0 < y < \infty$ and $|x| \to \infty$. The constraints define a convex set.

DEFINITION 4.1—Feasible Set
Let $g_i(y,x)$ be given by (4.9). Define for fixed $y > 0$ the convex set $x \in \mathbb{R}^n$ of feasible solutions

$$\Omega(y) = \{x : g_i(y,x) \leq 0, \forall i \in I\} \qquad \square$$

As noted above $\Omega(y)$ is also bounded. A lemma relates the sets $\Omega(y)$ for different values of y.

LEMMA 4.1
Assume that the convex set $\Omega(y_2)$ in Definition 4.1 is non-empty. If $y_2 > y_1$ then $\Omega(y_2) \subseteq \Omega(y_1)$.

Proof: Consider the convex set $\Omega_i(\cdot)$ defined by the function $g_i(\cdot,\cdot)$. It follows that

$$\begin{aligned} g_i(y_1,x) &= y_1 f_{1i}(x) + f_{2i}(x) \leq y_2 f_{1i}(x) + f_{2i}(x) \\ &= g_i(y_2,x) \leq 0 \end{aligned}$$

Thus, if $x \in \Omega_i(y_2)$ then $x \in \Omega_i(y_1)$, and $\Omega_i(y_2) \subseteq \Omega_i(y_1)$. The intersection of all $\Omega_i(y_2), i \in I$ form a convex set which is a subset of $\Omega(y_1)$. $\square$

The optimization problem (4.8) can now be reformulated as

$$\begin{array}{ll} \text{maximize} & y \\ \text{subject to} & x \in \Omega(y) \\ & y > 0 \end{array} \qquad (4.10)$$

If the set $\Omega(\cdot)$ is empty, no feasible solution exists. Otherwise the following theorem shows that the optimization problem has one global extremum.

THEOREM 4.1—Uniqueness
If a feasible solution exists for the optimization problem (4.10) then the set $X^* = \Omega(y^*) \subset \mathbb{R}^n$ corresponding to the maximum $y = y^*$ is convex and nonempty and any relative maximum is a global maximum.

Proof: Given an initial $y_0 > 0$, assume that $\Omega(y_0) \neq \emptyset$, i.e. a feasible solution exists. Increasing y generates ordered subsets of $\Omega(y)$ according to Lemma 4.1. Introduce $y^* = \sup \hat{y}$, $\Omega(\hat{y}) \neq \emptyset$, and a sequence $\Omega(y_i)$ with y_i monotonically increasing to y^*. Then $y^* < \infty$, since $\forall x$, at least one $f_{1i}(x) > 0$. Since the sets $\Omega(y_i)$ are closed and bounded it follows from Cantor's intersection theorem, e.g. [Simmons, 1963, pp. 73–74] that $\Omega(y^*) \neq \emptyset$. Optimum is thus achieved for the nonempty convex set $X^* = \Omega(y^*)$. $\square$

4.2 Practical Considerations

The proposed design method uses optimization to find a controller for certain closed loop behavior. By making several approximations an optimization problem with a unique optimum was obtained. The consequences of the approximations will now be discussed and it will also be shown how a feasible solution can be obtained.

Iteration over Q

The controller is parameterized by the numerator coefficients x of $Q(s,x)$ in (4.1). The choice of denominator of $Q(s,x)$ affects the solution to the problem. If $Q(s,x)$ is sufficiently complex it is a reasonable approximation of the infinite dimensional set of all $Q(s)$. After a solution x^* to the optimization problem (4.8) has been found, an investigation of $Q(i\omega,x^*)$ will show if a new denominator $D(s)$ further may improve the optimum of (4.8). It is not realistic to assume that the first choice of denominator has the requested properties, but an acceptable choice is obtained after a few iterations since the choice of $D(s)$ is not so crucial. This will be demonstrated in an example below.

Initial Feasible Controller

The optimization problem in (4.8) requires an initial controller parameterized by x in (4.1). For a reasonable guess of the initial controller the optimization routine itself, finds a feasible solution to the problem (4.8). For some problems the routine may however fail here. Searching for a feasible controller may in such cases be done in two steps. First, find a controller that yields a stable closed loop system for all processes $G(s,p), p \in \mathcal{P}$, then find a feasible solution to the optimization problem (4.8) for a given $y = y_0$.

Stabilizing Solution: A controller that stabilizes all processes $G(s,p), p \in \mathcal{P}$ yields a closed loop system that satisfies the condition in Theorem 3.1. The use of the lower bounding function (3.24) instead of the magnitude of the closed loop characteristic polynomial may cause some conservatism in the constraints (4.3). A measure of the conservatism is proposed in (3.26). The minimal conservatism may be found through the optimization problem

$$\begin{aligned} \text{minimize} \quad & \max_{\omega_j,p} \delta_\psi(\omega_j,p,x) \\ \text{subject to} \quad & 1 + \text{Re}\,\frac{\delta_{Ac}(\omega_j,p,x)}{A_{c0}(i\omega_j)} \geq \epsilon > 0 \qquad (4.11) \\ & j = 1,\ldots,m, \quad \forall p \in \mathcal{P}_c \end{aligned}$$

The objective function is convex in x since it is the maximum of some convex functions $\delta_\psi(\omega,p,x)$. The constraint function is affine in x. Hence, the

optimization problem (4.11) has one minimum that could be found using optimization. For a feasible solution to (4.11), Theorem 3.1 guarantees closed loop stability for all processes $G(s,p), p \in \mathcal{P}$.

Feasible Solution: A stabilizing solution to the optimization problem (4.11) may be improved so that the constraints in (4.8) are satisfied for $y = y_0$. An optimization problem of the type (4.10) is formulated using the constraints in (4.8). Hence, a feasible solution is found from

$$\begin{aligned}
&\text{maximize} && \gamma \\
&\text{subject to} && 0 \le \gamma \le 1 \\
& && \gamma\,(g_y(y_0,\omega_j,p,x)+1) \le 1 \\
& && \gamma\,(g_o(\omega_j,p,x)+1) \le 1 \\
& && \gamma\,(g_i(\omega_j,p,x)+1) \le 1 \\
& && \gamma\,(g_n(\omega_j,p,x)+1) \le 1 \qquad j=1,\ldots,m, \quad \forall p \in \mathcal{P}_c
\end{aligned} \tag{4.12}$$

A feasible solution to (4.8) is achieved if the optimal $\gamma^* = 1$. If $\gamma^* < 1$ no feasible solution is found.

If the weights in (4.12) give acceptable performance there is no need to search for an optimal solution. Arbitrary positive weights may then be used.

Implementation

The optimization problem is solved using a sequential quadratic programming algorithm implemented in Fortran [Gill *et al.*, 1986]. The Fortran subroutine is linked to PRO-MATLAB [MathWorks, 1990], where the problem is set up and the result is evaluated.

4.3 First Order Example

The first example is taken from [Masten and Cohen, 1989]. It will show that a controller satisfying the desired specifications is obtained after a few iterations. The reason for this iteration is that the initial choice of $D(s)$ in (4.1) does not yield a sufficiently good approximation of the infinite set $Q(s)$.

The process is considered to consist of two parts $G^0(s) = G_s G_u$. The first part is a first order transfer function describing parametric variations. It is given by

$$G_s(s,p) = \frac{b}{s+a}$$

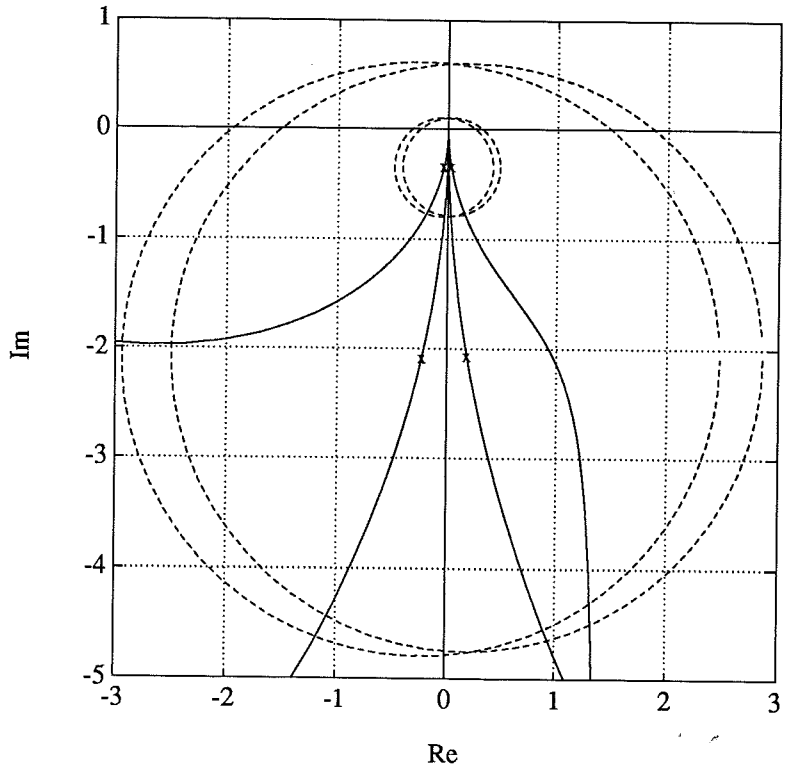


Figure 4.2 Nyquist curves for the loop transfer functions $L_0(i\omega, p)$, for the four $p \in \mathcal{P}_c$ with uncertainty circles drawn for the frequency $\omega = 20$ radians/s.

with $0.5 \leq b \leq 3.0$ and $-1.0 \leq a \leq 3.0$. Thus, the set $\mathcal{P}_c$ has four elements. Characteristics for these will be shown in the figures for this example. The second part of the transfer function represent unmodeled dynamics. It is assumed to be

$$G_u(s) = \frac{1}{1+sT}\, e^{-s\tau} \tag{4.13}$$

with $0 \leq T \leq 0.1$ and $0 \leq \tau \leq 0.05$. The unmodeled dynamics is bounded using unstructured uncertainty. This bound is given by

$$W_u(\omega) = \left| \frac{e^{-i0.05\omega}}{(1+i0.1\omega)} - 1 \right|$$

A feedback controller will be designed for effective input disturbance rejection. The initial controller is a PI controller

$$\frac{S_0(s)}{R_0(s)} = \frac{17s+14}{s}$$

The Nyquist curves for the loop transfer functions

$$L_0(i\omega,p) = \frac{S_0(i\omega)}{R_0(i\omega)} G_s(i\omega,p)$$

for $p \in \mathcal{P}_c$ are shown with solid lines in Figure 4.2. The controller stabilizes all four processes $p \in \mathcal{P}_c$ without unstructured uncertainty. Presence of unstructured uncertainty may, however, destabilize the closed loop system. For each frequency response in Figure 4.2, the frequency $\omega = 20$ radians/s is marked with a cross. Due to unstructured uncertainty, each of these points may vary inside respective dashed circle with center at the cross, i.e at $L_0(i20,p)$. The radius of these circles are given by $W_u(20)L_0(i20,p)$. Since some of these circles surround -1, the closed loop will be unstable for some processes G^0 within the uncertainty envelope.

A stabilizing controller which rejects input disturbances as fast as possible will now be determined. It is required that the sensitivity constraint $|\mathcal{S}_o^0| \leq 2.5$ should hold for all processes G^0 within specified region of uncertainty. This is captured by the optimization problem (4.8) with $f_y = f_i$, $W_y(\omega) = \sqrt{2}/\omega$, $y = \omega_b$, $W_i = 1/2$, and $W_o = 1/2.5$. The denominator of $Q(s,x)$ in (4.1) is first chosen as

$$D(s) = (s+1)(s+10)(s+100)$$

Optimization gives $\omega_b \approx 2.4$ radians/s and the measure of conservatism (3.26) is $\max \delta_\psi \approx 0.10$. The resulting controller $S_1(s)/R_1(s)$ is of fourth order. It will be used as the initial controller in the forthcoming optimization problems. A new optimization problem attempts to increase the bandwidth ω_b. For the initial controller S_1/R_1 and the denominator

$$D(s) = (s+1)(s+10)$$

optimization finds a controller that yields $\omega_b \approx 2.5$ radians/s. An inspection of the frequency response of Q reveals that this transfer function is a high pass filter. An appropriate choice would therefore be

$$D(s) = (s+10)^2(s+20)^3$$

where also the degree is increased for more freedom in $Q(s,x)$. This denominator is used in a new optimization problem. It follows that the maximal attainable bandwidth now is $\omega_b \approx 3.5$ radians/s. Model reduction of the ninth order controller from the optimization yields a third order controller

$$\frac{S_2(s)}{R_2(s)} = \frac{54.9(s+1.0)((s+17.0)^2+13.0^2)}{s((s+27.5)^2+52.3^2)}$$

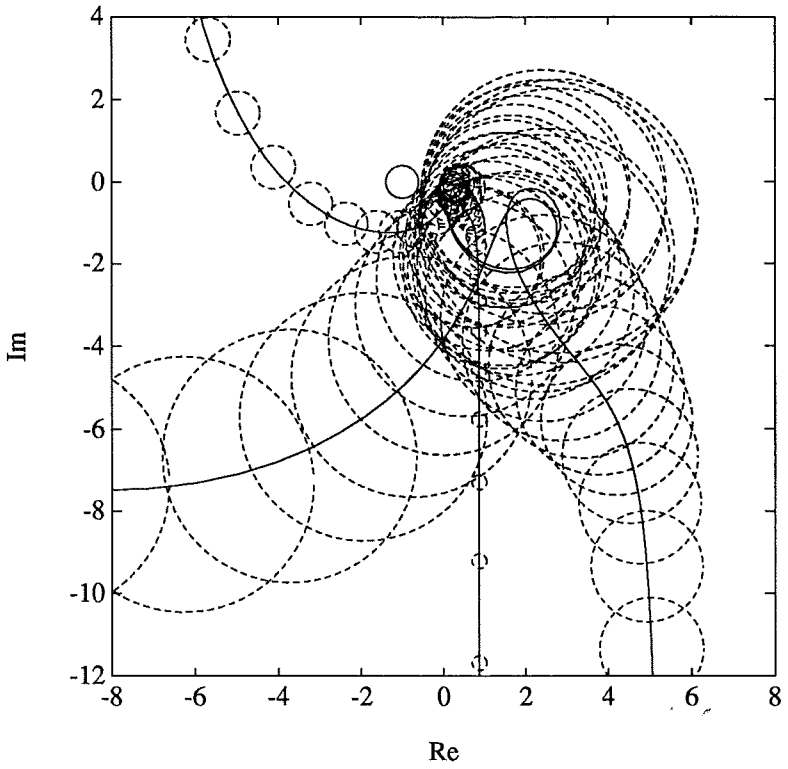


Figure 4.3 Nyquist curves for the loop transfer functions $L_2(i\omega, p)$, for $p \in \mathcal{P}_c$ with uncertainty circles.

Figure 4.3 shows the Nyquist curves for the loop transfer functions

$$L_2(i\omega, p) = \frac{S_2(i\omega)}{R_2(i\omega)} G_s(i\omega, p)$$

for $p \in \mathcal{P}_c$. Each of these curves has a set of circles describing the possible variation due to unstructured uncertainty for different frequencies. None of these uncertainty circles intersects the circle around -1 for the sensitivity function constraint. Hence, this constraint holds for all frequencies and for all processes G^0 under consideration.

Iterations with other initial controllers and with other denominators $D(s)$ does not improve the bandwidth further. The dashed lines in Figure 4.4 show $|\mathcal{S}_i(i\omega, p)|$ for the four processes with $p \in \mathcal{P}_c$. The solid lines show the maximal $|\mathcal{S}_i^0(i\omega, p, \Delta)|$, $|\Delta| \leq 1$ for the four $p \in \mathcal{P}_c$. The unstructured uncertainty increases the magnitude of the functions $|\mathcal{S}_i|$. The dotted line shows the inverse of the optimal performance function $1/W_{yi}(\omega, 3.5)$. As requested $1/W_{yi}$ is an upper bound on $|\mathcal{S}_i^0|$.

Figure 4.5 shows simulations for the four processes $p \in \mathcal{P}_c$ when the unmodeled dynamics $G_u(s)$ in (4.13) is defined by $T = 0.1$ and $\tau = 0.05$.

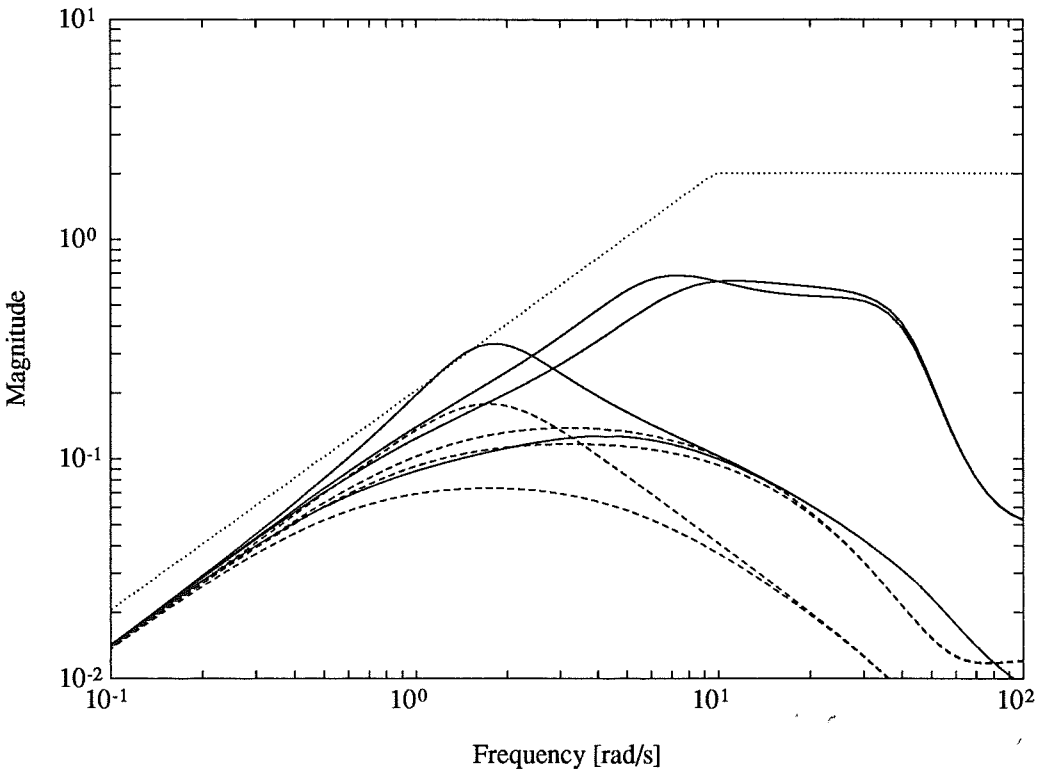


Figure 4.4 Magnitude of $\mathcal{S}_i$ for $p \in \mathcal{P}_c$ with unstructured uncertainty (solid) and without unstructured uncertainty (dashed). The inverse of the input disturbance performance function W_{yi}^{-1} (dotted).

The controller is S_2/R_2 and the feedforward is $B_{\text{ff}}/A_{\text{ff}} = 1/(0.88s + 1)$. The reference signal is one, at $t = 10$ s a constant input load ($l = -1$) affects the system and at $t = 18$ s a constant output load ($d = -0.25$) affects the system. The response to the input load disturbance is acceptable. The response to the output disturbance has a large overshoot, but this was not considered in the design.

4.4 Ship Steering Example

The second example considers the ship in Example 3.1. The process is unstable for certain operating conditions. A controller will be designed such that the sensitivity function is constrained by $|\mathcal{S}_o| \leq 1/W_o = 1.5$ and the noise sensitivity is constrained by $|\mathcal{S}_n| \leq 1/W_n = 100$ for all processes $p \in \mathcal{P}$ with unstructured uncertainty bounded by $W_u(\omega)$. Load disturbances, affecting the process at the input, should be eliminated in stationarity. This requires

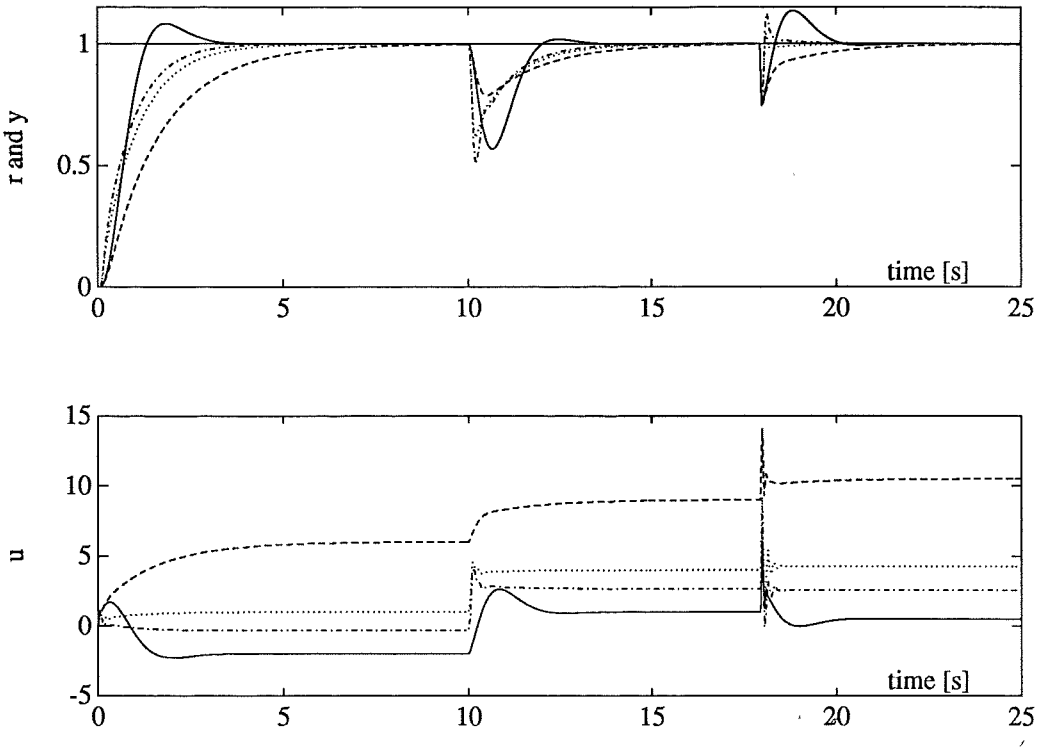


Figure 4.5 Simulation of closed loop system for four processes $p \in \mathcal{P}$ with unmodeled dynamics $G_u(s)$.

an integrating controller. The initial controller

$$\frac{S_0(s)}{R_0(s)} = \frac{3.75((s+0.24)^2+0.27^2)(s+1.71)}{s(s+1.11)(s+2.39)}$$

stabilizes the nominal process $p = p_0 \equiv (-0.70+0.05)/2 = -0.325$. It does not stabilize all processes within the prescribed range of uncertainty. This is seen in Figure 4.6, where the solid curves show the frequency responses of the loop transfer functions

$$L_0(i\omega, p) = \frac{B(i\omega, p)S_0(i\omega)}{A(i\omega, p)R_0(i\omega)}$$

for the two elements $p \in \mathcal{P}_c$. The dashed curves on each side of these frequency responses show the region of possible variation due to the unstructured uncertainty. The frequency responses do not encircle -1 for all uncertain processes under consideration, and hence robust stability is not achieved.

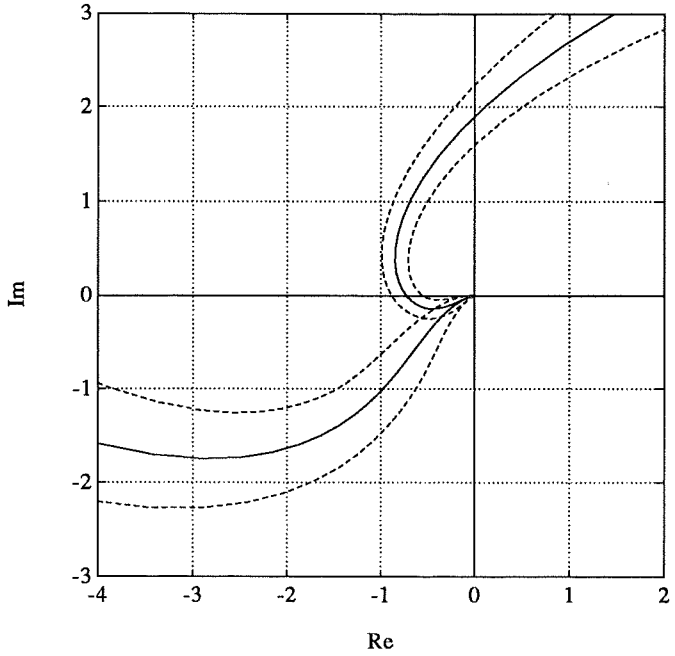


Figure 4.6 Nyquist curves $L_0(i\omega, p)$ with uncertainty regions of the loop transfer function for the system, with controller S_0/R_0 and for the two values of the process parameter $p \in \mathcal{P}_c$.

A controller satisfying the bounds on $|\mathcal{S}_o|$ and $|\mathcal{S}_n|$ is given by a feasible solution x^f to the optimization problem (4.8). The constraints are defined at 25 frequencies logarithmically spaced between 10^{-2} and 10^2. The controller (3.14) is parameterized by $Q(s, x)$ in (4.1) with

$$D(s) = (s+1)^4(s+5)^2(s+20)^2$$

A feasible solution to the optimization problem (4.8) yields an eleventh order controller. Its frequency response is adequately described by a fourth order controller, that is obtained using model reduction. It is given by

$$\frac{S_1(s)}{R_1(s)} = \frac{54.71((s+0.13)^2+0.17^2)(s+1.99)(s+46.21)}{s(s+1.27)(s+21.21)(s+28.95)}$$

Optimal Input Disturbance Rejection

A controller for optimized input disturbance rejection will now be designed. Define $W_y(\omega) = \sqrt{2}/\omega$ and $f_y = f_i$. Keep $W_o = 1/1.5$ and $W_n = 0.01$ as above. The controller is defined by the nominal controller S_1/R_1 above and it is parameterized by $Q(s, x)$ in (4.1) with

$$D(s) = (s+0.5)^2(s+1)^3(s+5)^2$$

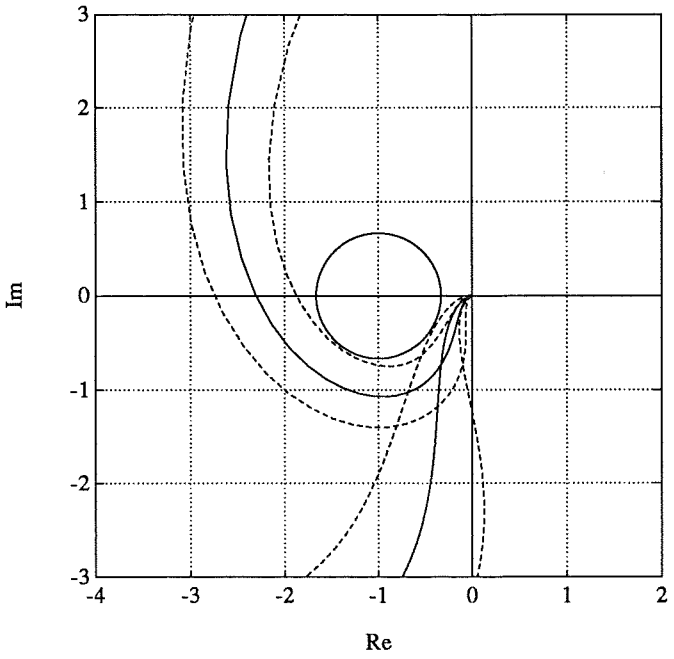


Figure 4.7 Nyquist curves $L_2(i\omega, p)$ with uncertainty regions of the loop transfer function for the system, with controller S_2/R_2 and for the two values of the process parameter $p \in \mathcal{P}_c$.

This $D(s)$ is obtained iteratively as demonstrated in the previous example. A controller for optimized disturbance rejection is given by the solution to the optimization problem (4.8). The controller is of eleventh order. Model reduction then gives a fourth order controller

$$\frac{S_2(s)}{R_2(s)} = \frac{93.72((s+0.42)^2+0.36^2)(s+2.00)(s+36.35)}{s(s+1.86)(s+22.85)(s+28.97)}$$

The solid curves in Figure 4.7 show the frequency responses of the loop transfer functions

$$L_2(i\omega, p) = \frac{B(i\omega, p)S_2(i\omega)}{A(i\omega, p)R_2(i\omega)}$$

for the two elements $p \in \mathcal{P}_c$. The circle around -1 has radius $1/1.5$. It corresponds to the sensitivity constraint $|S_o| \leq 1.5$. The uncertainty regions due to unstructured uncertainty are shown by dashed lines as in Figure 4.6. These regions do not intersect the circle around -1 as requested by the design constraint on the sensitivity function.

The conservatism of using ψ in (3.24) instead of the norm of the closed

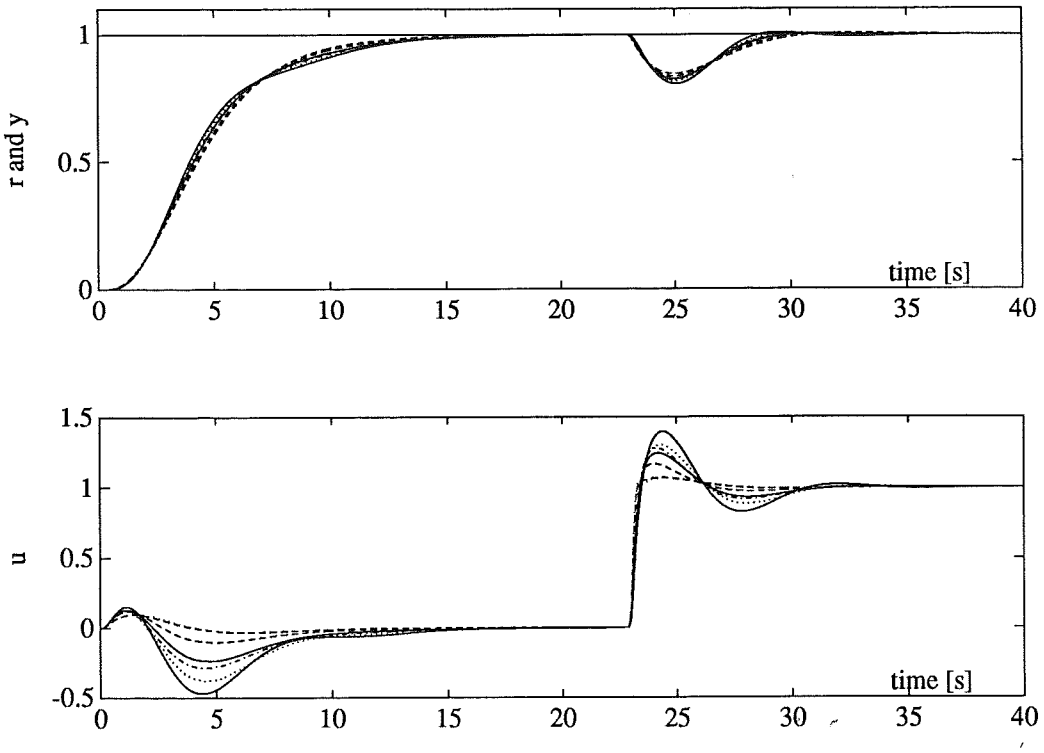


Figure 4.8 Simulation of seven closed loop systems in the ship example.

loop characteristic polynomial is measured by (3.26). For this problem it is

$$\max_{\omega,p\in\mathcal{P}_c} \delta_\psi(\omega,p,x^*) \approx 0.07$$

which is considerably smaller than one.

The bandwidth, as defined by Figure 4.1, is $\omega_b = 0.96$ radians/s. Simulations of the closed loop system for seven different open loop processes within the uncertainty bound are shown in Figure 4.8. The controller S_2/R_2 was used, and the feedforward

$$\frac{B_{ff}(s)}{A_{ff}(s)} = \frac{1}{(1+1.61s)^3}$$

was chosen as suggested in Chapter 3. The reference signal is a step and an input load disturbance affects the process at $t = 23$ s. The feedforward slows down the response time for reference inputs to avoid overshoots for some processes. The responses for the different processes are then close as seen in Figure 4.8.

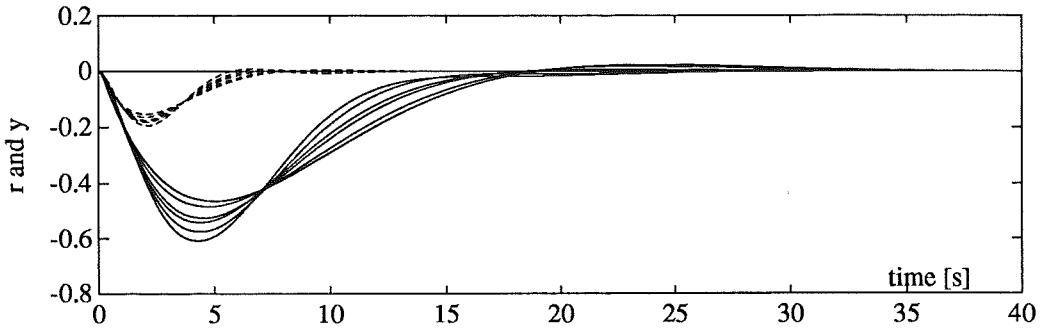


Figure 4.9 Simulation of seven closed loop systems for two different controllers.

Optimal Noise Attenuation

Assume now that a lower bandwidth than above is acceptable. It is considered as acceptable if it satisfies $\omega_b \geq 0.10$ radians/s, which is ten times lower than the optimal bandwidth above. Another controller will now be designed to minimize the measurement noise contribution in the control signal. To do this define $W_y = 1$, $f_y = f_n$, $W_i = \sqrt{2}\cdot 0.1/\omega$ and $W_o = 1/1.5$ for the optimization problem (4.8). An optimal solution is obtained for $y = 0.021$ which means that $|\mathcal{S}_n|$ is less than 48 for all uncertain processes under consideration. The low demand on bandwidth here, allows lower noise sensitivity compared to the controller for fast load disturbance rejection. After model reduction, the controller is

$$\frac{S_3(s)}{R_3(s)} = \frac{47.21((s+0.15)^2+0.16^2)(s+2.01)}{s(s+1.57)(s+12.78)}$$

Figure 4.9 show the responses to a constant input load disturbance for seven different processes within the uncertainty bound when the controllers S_2/R_2 (dashed lines) and S_3/R_3 (solid lines) are used. The fastest response is achieved for the controller S_2/R_2 that is optimized with respect to input disturbance rejection. The slowest response is achieved for the controller S_3/R_3 that is optimized with respect to measurement noise attenuation, with little emphasize on load disturbance rejection properties.

Finally Figure 4.10 show the frequency responses for the four controllers that have been designed. The initial controller $S_0(i\omega)/R_0(i\omega)$ (dash-dotted) does not have sufficient phase advance to stabilize all uncertain processes. The feasible controller $S_1(i\omega)/R_1(i\omega)$ (dotted), has more phase lead, and the price for this is high gain at high frequencies. This property is also found for the controller $S_2(i\omega)/R_2(i\omega)$ (dashed). However, the low frequency gain is increased for this controller. This means faster disturbance rejection. The

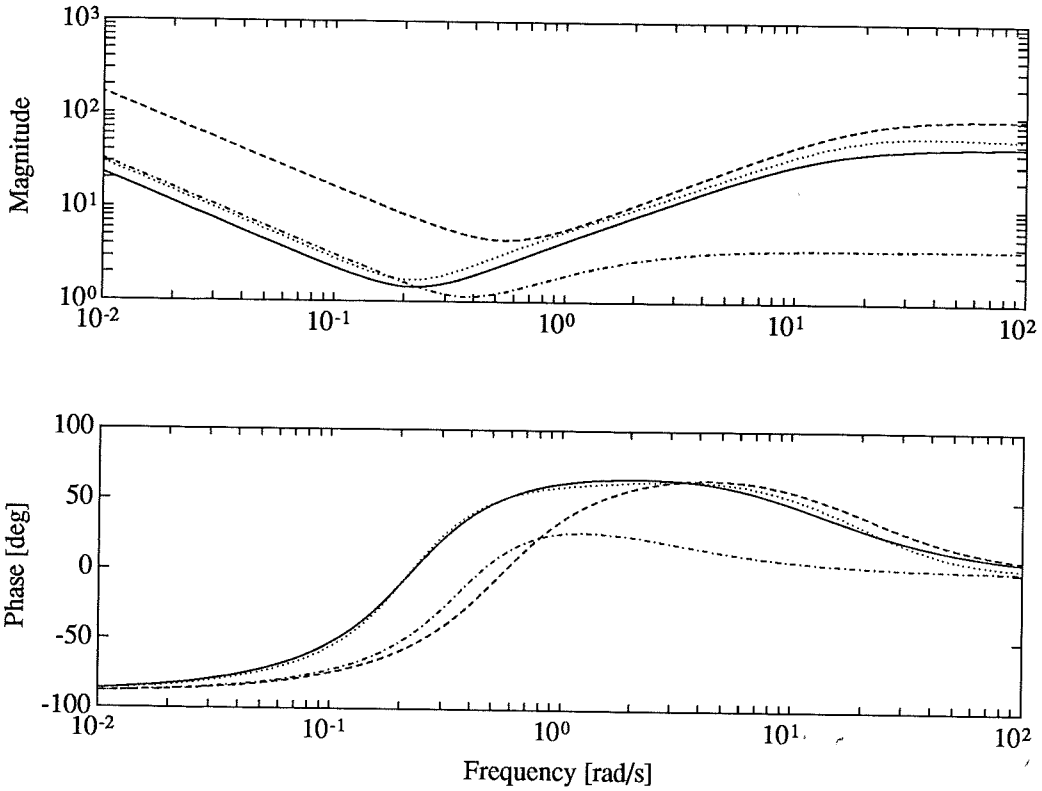


Figure 4.10 Frequency responses for the controllers in the ship example.

controller $S_3(i\omega)/R_3(i\omega)$ (solid) for best noise attenuation has lower gain than the other controllers that stabilize all processes under consideration. None of these controllers have high frequency roll-off. For a practical implementation such would be desired.

4.5 Conclusions

A new method for robust control system design has been presented here. The method considers processes with combined structured and unstructured uncertainties. The controller parameters are determined through a convex optimization problem which is formulated using the frequency domain conditions for robust performance in Chapter 3.

Two examples have demonstrated the use of the design method. The first example showed how iteration over different denominators D and initial controllers S_0/R_0 leads to a controller with desired properties. The second example showed how the method can be used to achieve a controller that is optimized in some sense, while regarding other design criteria and uncertainty.

5

Robust and Adaptive Control

Robust control of uncertain processes has been discussed in Chapter 3 and Chapter 4. Adaptive control is another way to deal with variations in the process dynamics. For very large uncertainties adaptive control may be the only way to obtain an acceptable control system. Adaptive controller are reasonable well understood. See e.g. the textbooks [Åström and Wittenmark, 1989] and [Goodwin and Sin, 1984]. Although adaptive controllers have the potential to give excellent performance, it has been observed that adaptive systems may also be difficult to commission. For this reason the most extensively used adaptive technique is automatic tuning of simple controllers of the PID type [Kraus and Myron, 1984, Åström and Hägglund, 1988a]. These controllers have been developed to the stage where tuning is performed simply by pushing a tuning button. It would be desirable to make adaptive controllers as easy to use as the simple autotuners. Such a procedure will be developed in Chapters 5, 6, and 7. The key idea is to use information from an experiment with relay feedback to initialize the adaptive controller.

Section 5.1 gives a brief background to adaptive control and the equations are summarized for the adaptive controller that is used here. This controller is a well known algorithm based on pole placement design and recursive least squares estimation. It is known that pole placement control design may lead to closed loop systems with poor robustness. The controller itself may e.g. be unstable. This is illustrated in Section 5.2. The design procedure developed in Chapter 4 is too complicated to perform on line. In

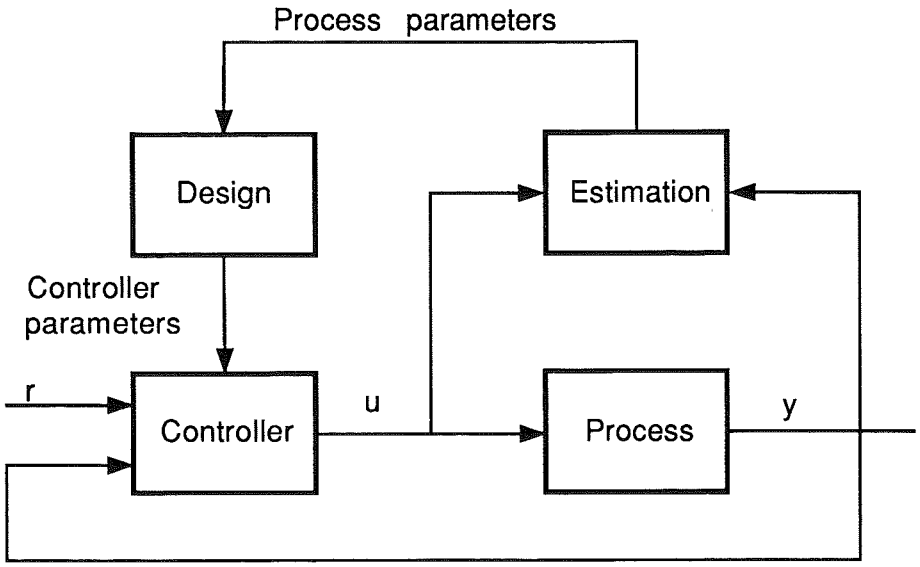


Figure 5.1 An indirect adaptive controller.

Section 5.2 a much simpler way to improve the robustness of the pole placement controller is therefore suggested. This procedure, which is believed to be novel, is introduced in the adaptive controller.

5.1 Adaptive Control

An adaptive control system is a special type of nonlinear feedback system. The controller has the ability to adjust itself from knowledge, successively gained about the process, such that the closed loop system behaves as desired.

In this thesis an indirect adaptive controller with the structure shown in Figure 5.1 is used. Only single input single output systems are considered. The control design is based on a robustified version of pole placement design. Parameter estimation is performed by recursive least squares.

The selection of desired closed loop poles is critical for the success of a pole placement design. This is done based on the characteristics of the process. The controller is a sampled data system. The selection of sampling period is critical for control design and especially for a successful parameter estimation. These choices are discussed in Chapter 7.

The process is described by the input-output relation

$$y(k) = \frac{q^{-d}B^*(q^{-1})}{A^*(q^{-1})}\,u(k) \tag{5.1}$$

where $B^*(q^{-1})$ and $A^*(q^{-1})$ are polynomials of degree m and n respectively in the delay operator q^{-1} and d is the delay in the process. It is required that

$d > 0$ to have a strictly proper model of the process. In an adaptive controller it is essential to know the delay. The polynomials in (5.1) are estimated by a recursive least squares algorithm

$$\begin{aligned} \theta(k) &= \theta(k-1) + K(k)\left(y_f(k) - \varphi_f^T(k)\theta(k-1)\right) \\ K(k) &= P(k-1)\varphi_f(k)/\left(\lambda + \varphi_f^T(k)P(k-1)\varphi_f(k)\right) \\ P(k) &= \left(I - K(k)\varphi_f^T(k)\right)P(k-1)/\lambda \end{aligned} \tag{5.2}$$

with the polynomial coefficients defined in

$$\theta(k) = \begin{pmatrix} a_1 & \dots & a_n & b_0 & \dots & b_m \end{pmatrix}^T$$

The estimator uses a forgetting factor λ to put more emphasize on recent measurements than on older. The input and output signals of the process are filtered before they are fed into the estimation algorithm.

$$y_f(k) = H_f^*(q^{-1})y(k) \qquad u_f(k) = H_f^*(q^{-1})u(k)$$

The filtered output $y_f(k)$ and the filtered regression vector

$$\varphi_f(k) = \begin{pmatrix} -y_f(k-1) \\ \vdots \\ -y_f(k-n) \\ u_f(k-d)) \\ \vdots \\ u_f(k-m-d)) \end{pmatrix}$$

are then used in the estimator. The filter $H_f^*(q^{-1})$ may be interpreted as a frequency domain weighting of the error in frequency response of the estimated process model. The filter is often chosen with band pass characteristics. The passband should be chosen such that an accurate process model is obtained for frequencies where the loop transfer function L is close to -1 in the Nyquist diagram. The regression filter would therefore be chosen from the closed loop characteristics.

The controller is defined in the delay operator as

$$R^*(q^{-1})u(k) = T^*(q^{-1})r(k) - S^*(q^{-1})y(k) \tag{5.3}$$

Its coefficients are determined in the design block, using pole placement design that was introduced in Chapter 2. The Diophantine-Aryabhatta-Bezout equation

$$A^*(q^{-1})A_r^*(q^{-1})R'^*(q^{-1}) + q^{-d}B^*(q^{-1})S^*(q^{-1}) = A_c^*(q^{-1}) \tag{5.4}$$

is defined by the estimated parameters θ and the desired closed loop specifications. It is solved at every sampling instant.

A Priori Knowledge

Any control system synthesis is guided by specifications on the desired behavior of the closed loop system. The specifications may be in time or frequency domain or in locations of the closed loop poles. An adaptive controller must also be given information of what it is intended to achieve. It is not a black box that may be connected to any real process and be expected to do a good job. The indirect adaptive controller above needs the following a priori information.

Sampling interval should be chosen with respect to the closed loop system. The essential dynamics of the closed loop system should be reflected in the discrete time model.

Process model structure is defined by the delay d and the degrees of the polynomials B^* and A^*. It is used to set up the recursive estimator (5.2).

Closed loop specifications are given by the characteristic polynomial $A_c^* = A_o^* A_m^*$ and the polynomial A_r^*. An integrating controller is introduced with $A_r^* = 1 - q^{-1}$. A such will be used in the sequel. The complexity of the process model determines the degrees of A_o^* and A_m^*. Appropriate closed loop poles may be selected from properties of the open loop process.

Variables for the recursive estimator are the initial estimate of the parameters $\theta(0)$ and an initial covariance matrix $P(0)$. The regression filter H_f^* and the forgetting factor λ must also be given.

The data above should be given before the controller is commissioned. They are essential for the robustness of the adaptive controller. Information gained by a simple experiment under relay feedback will here be used to obtain the data.

5.2 A Simple Robustification

The design method in Chapter 4 used the rational Q in the controller (3.14) to satisfy various robustness criteria. The controller was obtained using convex optimization. Such calculations are too complicated to do on line in an adaptive algorithm. There are, however, some simple choices that will lead to controllers with significantly improved performance.

The properties of a controller may be influenced by requiring that its transfer function has certain specified poles and zeros. A common case is to require that the controller has integral action, i.e. a pole at $q = 1$ in

Table 5.1 Lower bound p_{n_z} of pole p for $0 \leq n_z \leq 2$ controller zeros in $q = -1$ that gives a stable controller.

d	n_c	p_0	p_1	p_2
1	1	0.80	0.11	0.00
	2	0.40	0.06	0.00
2	1	0.80	0.62	0.51
	2	0.46	0.37	0.31
3	1	0.80	0.72	0.66
	2	0.52	0.47	0.45
4	1	0.80	0.76	0.72
	2	0.56	0.53	0.52
5	1	0.80	0.77	0.75
	2	0.58	0.57	0.56

the discrete time case. Similarly it can be required that a discrete time controller has a zero at $q = -1$. This implies that the controller has zero gain at the Nyquist frequency. This is a simple way to improve robustness and insensitivity to high frequency measurement noise. A simple example illustrates the benefits.

EXAMPLE 5.1
The process

$$\frac{B^*(q^{-1})}{A^*(q^{-1})} = \frac{q^{-(d+1)}}{1 - 0.8q^{-1}} \tag{5.5}$$

is controlled with an integrating controller. This means that the denominator of the controller has the structure $R^*(q^{-1}) = R'^*(q^{-1})(1-q^{-1})$. The numerator is chosen to have n_z zeros at $q = -1$, hence $S^*(q^{-1}) = S'^*(q^{-1})(1+q^{-1})^{n_z}$. The closed loop characteristic polynomial is specified as

$$A_c^*(q^{-1}) = (1 - pq^{-1})^{n_c} \tag{5.6}$$

with $0 \leq p < 1$. The controller is obtained from the solution to (2.8). The freedom $Q(q^{-1})$ in (2.12) is used to give $S(q^{-1})$ the desired zeros. The controller obtained will be unstable if p is chosen too small. Table 5.1 shows how fast the pole p can be chosen for different delays d and different degrees n_c to have a stable controller. A controller with n_z zeros at $q = -1$ is stable for $p \geq p_{n_z}$. The last three columns of Table 5.1 show p_{n_z} for $n_z = 1, 2, 3$. It is observed that for processes with small delays d, the pole p can be chosen faster when the controller has zero in $q = -1$ compared with a controller without such a zero. For large delays, however, the difference is small.

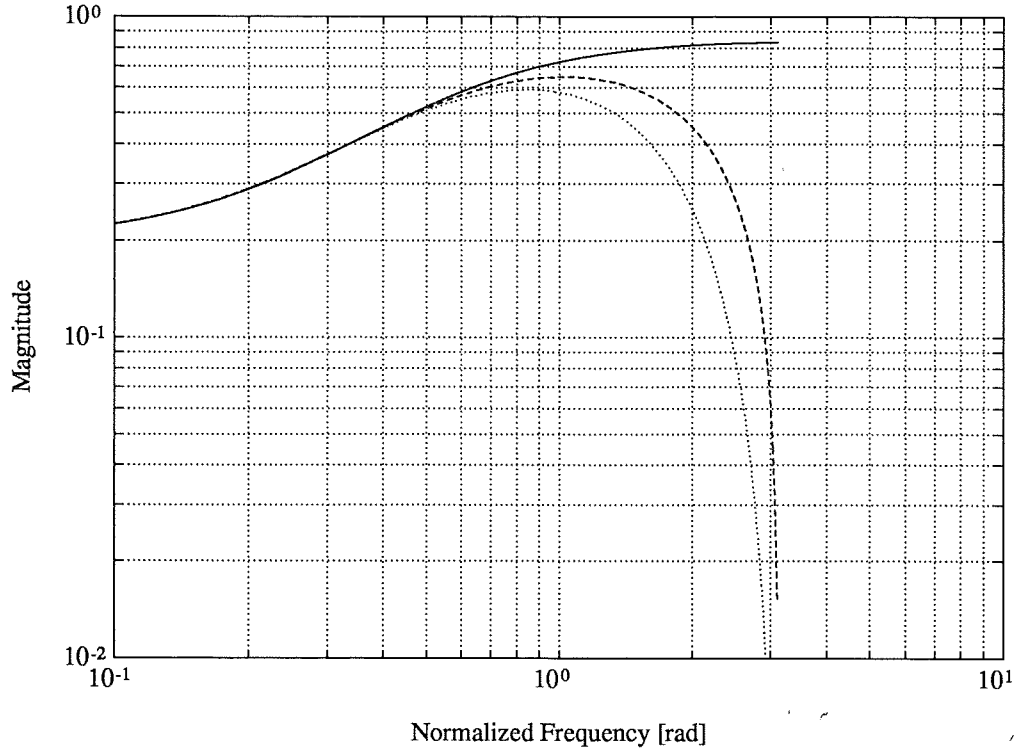


Figure 5.2 Noise sensitivity $\mathcal{S}_n(e^{-i\omega h})$ for controllers with $n_z = 0$ (solid line), $n_z = 1$ (dashed line), and $n_z = 2$ (dotted line).

Controller stability is not the only reason for introducing zeros at $q = -1$ in the controller. A controller with such a zero has low gain at high frequencies. This gives low sensitivity to high frequency measurement noise. Figure 5.2 shows the noise sensitivity for different n_z. The process (5.5) has delay $d = 2$. The closed loop characteristic polynomial (5.6) is defined by $p = 0.6$ and $n_c = 2$. The noise sensitivities $|\mathcal{S}_n(e^{-i\omega h})|$ are shown for $n_z = 0$ (solid line), $n_z = 1$ (dashed line), and $n_z = 2$ (dotted line). The magnitude of $\mathcal{S}_n(e^{-i\omega h})$ decreases with increasing n_z. From Table 5.1 it follows that the controllers are stable for these three n_z. It should also be noticed that the sensitivity functions $\mathcal{S}_o(e^{-i\omega h})$ are close for the three cases. This means that the disturbance rejection properties are similar. Analogous behavior of $\mathcal{S}_n(e^{-i\omega h})$ and of $\mathcal{S}_o(e^{-i\omega h})$ are observed for all processes in this example. □

The example has demonstrated some benefits of having a low controller gain at high frequencies. The interval with low gain may be enlarged by introducing more zeros in the controller. This will not be pursued here. Thus, the suggested robustification of the pole placement design that will be used in Chapter 7 is simply that the controller should have one zero at $q = -1$.

6

Relay Feedback

One successful method to tune a PID-controller is based on relay feedback [Åström and Hägglund, 1988a]. The idea is that many processes will exhibit limit cycle oscillations under relay feedback. For a large class of processes good controller parameters can be obtained from the amplitude and the period of the limit cycle. In [Åström and Hägglund, 1988b] and [Hägglund and Åström, 1991] it is shown that simple discrete time models can also be calculated from the waveform of the limit cycle.

This chapter discusses information about the process that can be obtained from an experiment with relay feedback. Both frequency domain information and time domain information may be useful to characterize the process. It will be shown in Chapter 7 that an experiment with a relay feedback provides sufficient a priori information to commission the adaptive controller described in Chapter 5.

6.1 The Relay Feedback System

Relay feedback is a classical topic in control theory. A key problem is to characterize the behavior of linear systems under relay feedback. Major advances were made in the fifties. Main results are found in the textbooks [Atherton, 1975] and [Tsypkin, 1984]. Systems with with relay feedback have a very complex behavior and there are still important problems that are not resolved. A complete characterization of first order processes with dead time has recently been published [Holmberg, 1991].

The block diagram of a linear system with relay feedback is shown in

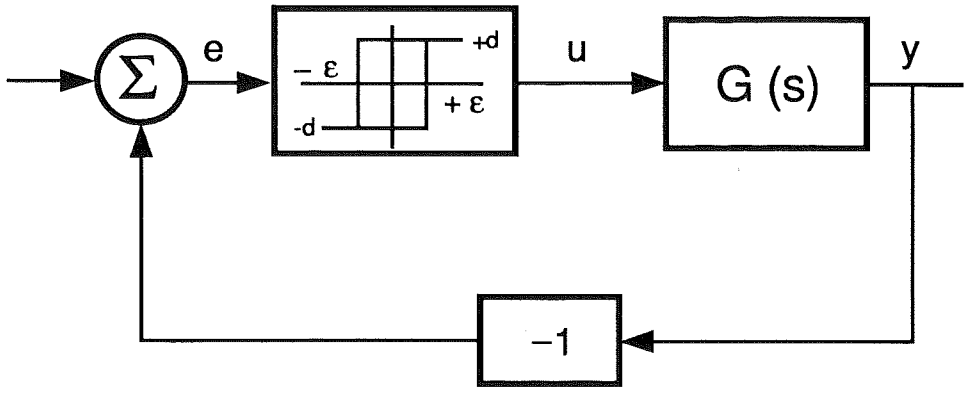


Figure 6.1 Relay feedback

Figure 6.1. The relay output signal is given by

$$u(t) = \begin{cases} d, & \text{if } e \geq \epsilon \text{ or } (e > -\epsilon \text{ and } u(t-) = d) \\ -d, & \text{if } e \leq -\epsilon \text{ or } (e < \epsilon \text{ and } u(t-) = -d) \end{cases}$$

where the relay amplitude is d and the hysteresis is ϵ.

Under relay feedback when a stable symmetric oscillation around zero is established, the relay output will be a square wave with amplitude d and period $T_{osc} = 2\pi/\omega_{osc}$. The Fourier series expansion of the process input signal is

$$u(t) = \frac{4d}{\pi} \sum_{m \in \mathcal{M}} \frac{1}{m} \sin(\omega_{osc} m t) \tag{6.1}$$

where $\mathcal{M} = \{1, 3, 5, \ldots\}$. Similarly the Fourier series expansion of the process output is

$$y(t) = \frac{4d}{\pi} \sum_{m \in \mathcal{M}} \frac{1}{m} |G(i\omega_{osc} m)| \sin\big(\omega_{osc} m t + \arg G(i\omega_{osc} m)\big) \tag{6.2}$$

Practical Considerations

Assume that the system initially is in stationarity at the desired operating point. The input signal is kept constant for a while to estimate the noise variance σ and the maximal amplitude of the noise v_{max}. The relay hysteresis level is then chosen from the maximal noise amplitude, so that erroneous relay switches due to noise are avoided. A significant change in the error signal will be required for the relay to switch. The relay feedback is now

introduced and the system will oscillate. A predefined value of the relay amplitude d is used initially. It is adjusted under the experiment so that the oscillation will have a desired amplitude y_{max}. There is a trade-off between the information content in the output signal, due to the signal to noise ratio, and the permissible perturbation.

Assume that the first harmonic dominates the process output. The ratio between the oscillation amplitude

$$y_{max} = \frac{4d}{\pi} |G(i\omega_{osc})|$$

and the relay hysteresis determines the phase shift of $G(i\omega_{osc})$. Describing function analysis gives

$$\text{Im } G(i\omega_{osc}) = -\frac{\pi\epsilon}{4d}$$

and it follows that

$$\arg G(i\omega_{osc}) = -\pi + \arcsin\left(\frac{\epsilon}{y_{max}}\right) \tag{6.3}$$

The ratio ϵ/y_{max} is given by the user. The choice $\epsilon/y_{max} = 0.5$ implies that $\arg G(i\omega_{osc}) \approx -5\pi/6$.

Asymmetric oscillations may occur if the relay is asymmetric or if there are small load disturbances. Large load disturbances may quench the oscillations. Constant disturbances may be compensated for by introducing a bias in the relay [Hang and Åström, 1988]

6.2 Frequency Domain Information

Under relay feedback the input signal is a square-wave when a periodic oscillation is established. The Fourier series expansion of the input signal (6.1) shows that the signal has its energy at the frequencies $\omega = \omega_{osc}m$ with $m \in \mathcal{M}$. It is therefore appropriate to estimate the frequency response $G(i\omega)$ at these frequencies.

The slope of the frequency response magnitude is locally given by

$$n(\omega) = -\frac{\partial \log |G(i\omega)|}{\partial \log \omega} \tag{6.4}$$

The maximal slope is a measure of the process complexity. It may be used to estimate the relative degree of the process. If the slope is steep in some frequency interval a high order process model will be required for proper modeling in this interval.

Frequency Response Estimation

The frequency response of the process is estimated using the discrete Fourier transforms of u and y. When the signals are sampled with interval h the discrete Fourier transform is given by

$$Z_N(\Omega) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} z(k) e^{-i\Omega k}$$

where $\Omega = 2\pi k/N$, $k = 1, 2, \ldots, N$. The frequency response G is then estimated by

$$\hat{H}(e^{i\Omega_m}) = \frac{Y_N(\Omega_m)}{U_N(\Omega_m)} \tag{6.5}$$

for $m \in \mathcal{M}$. Notice that an estimate is obtained only for the frequencies $\Omega_m = \omega_{osc} mh$, $m \in \mathcal{M}$ where the input spectrum is nonzero.

To illustrate the accuracy of estimates, assume that the process is defined by

$$y(k) = H_0(q)u(k) + v(k) \tag{6.6}$$

where $v(k)$ is measurement noise. Assume further that $v(k)$ is a white noise sequence with zero mean value and with variance σ^2. The m:th component of the sampled input signal

$$u_m(k) = \frac{4d}{\pi m} \sin(\Omega_m k) \tag{6.7}$$

has the spectrum

$$|U_N(\Omega)|^2 = \begin{cases} 4N \left(\frac{d}{\pi m}\right)^2, & \text{if } \Omega = \pm\Omega_m \\ 0, & \text{if } \Omega \neq \pm\Omega_m \end{cases} \tag{6.8}$$

If the summation time for the discrete Fourier transform is a multiple of the period, the frequency response estimate (6.5) of the process (6.6) is unbiased

$$E\left[\hat{H}(e^{i\Omega})\right] = H_0(e^{i\Omega})$$

and the covariance of the estimate is

$$\text{Var}\left[\hat{H}(e^{i\Omega_{m_1}}), \hat{H}(e^{i\Omega_{m_2}})\right] = \begin{cases} \frac{\sigma^2}{|U_N(\Omega)|^2}, & \text{if } m_1 = m_2 \\ 0, & \text{otherwise.} \end{cases}$$

for frequencies where the input spectrum is nonzero. From (6.8) it then follows that the variance of the frequency response estimate is

$$\text{Var}\left[H(e^{i\Omega_m})\right] = \frac{\pi^2\sigma^2 m^2}{4Nd^2} \tag{6.9}$$

for the normalized frequency Ω_m. Assuming that y_{max} is dominated by the first harmonic it is possible to rewrite (6.9) as

$$\frac{\text{Var}\left[H(e^{i\Omega_m})\right]}{|H(e^{i\Omega_1})|^2} = \frac{4}{N}\left(\frac{\sigma}{y_{max}}\right)^2 m^2 \tag{6.10}$$

The variance in the estimate for a certain harmonic is determined by the signal to noise ratio y_{max}/σ, the number of measurements and the number m of the harmonic. The variance of the frequency response estimate (6.5) increases quadratic with m. The estimates deteriorate fast with increasing frequency. Increasing the number of measurements will decrease the variance of the transfer function estimate. A decrease may either be accomplished by faster sampling or summation over more periods.

Estimation of Slope

The slope (6.4) of the transfer function at the crossover frequency is of interest to discriminate between different models. It is obtained from the frequency response estimates (6.5). The slope at the frequency $\Omega = \sqrt{\Omega_{m_1}\Omega_{m_2}}$ is estimated by

$$\hat{n}\left(\left|H(e^{i\Omega_{m_1}})\right|, \left|H(e^{i\Omega_{m_2}})\right|\right) = -\frac{\log\left|\hat{H}(e^{i\Omega_{m_1}})\right| - \log\left|\hat{H}(e^{i\Omega_{m_2}})\right|}{\log(m_1) - \log(m_2)} \tag{6.11}$$

To estimate the accuracy of the estimated slope, assume that the magnitude of the frequency response in the interesting frequency interval is described by

$$\left|\hat{H}(e^{i\Omega_m})\right| = \left|\hat{H}(e^{i\Omega_1})\right|\left(\frac{1}{m}\right)^{\alpha} \tag{6.12}$$

where the slope α is a real number. Gauss' approximation formula gives

$$\text{E}\left[\hat{n}\right] = -\frac{\log \text{E}\left[\left|\hat{H}(e^{i\Omega_{m_1}})\right|\right] - \log \text{E}\left[\left|\hat{H}(e^{i\Omega_{m_2}})\right|\right]}{\log(m_1) - \log(m_2)} = \alpha$$

Table 6.1 Standard deviation of slope in Example 6.1.

α	m_1	m_2	$\sigma_{\hat{n}}$
1	1	3	0.12
	1	5	0.22
	3	5	0.74
2	1	3	0.35
	1	5	1.10
	3	5	3.54
3	1	3	1.04
	1	5	5.49
	3	5	17.45

Since the estimates $h_m = |\hat{H}(e^{i\Omega_m})|$ are uncorrelated for different m it follows that

$$\text{Var}\,[\hat{n}(h_{m_1}, h_{m_2})] = \left(\frac{\partial \hat{n}}{\partial h_{m_1}}\right)^2 \text{Var}\,[h_{m_1}] + \left(\frac{\partial \hat{n}}{\partial h_{m_2}}\right)^2 \text{Var}\,[h_{m_2}] \quad (6.13)$$

Inserting (6.10) and (6.12) in (6.13) gives

$$\text{Var}\,[\hat{n}] = \frac{4}{N\left(\log(m_1/m_2)\right)^2}\left(\frac{\sigma}{y_{max}}\right)^2 \left(m_1^{2+2\alpha} + m_2^{2+2\alpha}\right) \quad (6.14)$$

Notice that this expression for the variance is approximative. Gauss' approximation formula may be inaccurate when the variance of $|\hat{H}(e^{i\Omega_m})|$ is large compared to the magnitude $|\hat{H}(e^{i\Omega_m})|$. This may happen for large α.

The relay amplitude is chosen so that the signal to noise ratio is sufficiently large. If not, the only way to achieve accurate estimates of n at the oscillation frequency is to increase the number of measurements. The sampling interval is limited by the available computer and it is not practical to have experiments over many periods of oscillation. Here a trade-off is necessary.

EXAMPLE 6.1
Consider a relay experiment over $N = 200$ samples. Assume that the signal to noise ratio is $y_{max}/\sigma = 10$. Table 6.1 shows the standard deviation of the estimates of the slope $\hat{n}$ using (6.14) for different values of the slope α. The standard deviation of the estimate of the slope increases fast with increasing harmonic and increasing order of the process. Both these effects are due to decreasing frequency response magnitude of the open loop process. □

Summary

The conclusion from Example 6.1 is that with a reasonable signal to noise ratio and not a too long experiment, only the two first harmonics of the signals should be used for the slope estimation. The estimated slope is hereafter considered to represent the local slope of the frequency response at $\omega = \sqrt{3}\omega_{osc}$, which is the geometric mean of the frequencies for the two first harmonics of the oscillation.

6.3 Time Domain Information

In the previous section it was shown how frequency domain information could be obtained from the relay experiment and how the slope of the frequency response magnitude could be estimated. Time domain information can also be extracted from the relay experiment. This can be used to determine the time constant T and the delay τ for the model

$$G(s) = \frac{b}{(sT+1)^n} e^{-s\tau} \tag{6.15}$$

where the degree n is obtained from the frequency response estimate. Processes with monotone step responses may be approximated by (6.15). Although this may be a crude approximation, the parameters of the model (6.15) can be used for initialization of an adaptive controller. This will be pursued in Chapter 7.

The phase shift of the model (6.15) is given by

$$\arg G(i\omega) = -\omega\tau - n \arctan \omega T \tag{6.16}$$

and the slope is locally given by

$$n_l(\omega) = -\frac{\partial \log |G(i\omega)|}{\partial \log \omega} = \frac{(\omega T)^2}{1+(\omega T)^2} n \tag{6.17}$$

If the argument $\arg G(i\omega)$ is known for some frequency ω, Equation (6.16) defines a relation between ωT and $\omega\tau$. If one of these quantities is known the other may then be obtained from (6.16). Consider especially the frequency ω_π, defined by $\arg G(i\omega_\pi) = -\pi$, which may be considered as an upper bound of the oscillation frequency ω_{osc}. Define the two quantities $\omega_\pi\tau$ and $\omega_\pi T$ as the *normalized delay* and the *normalized time constant.* They may be considered as measures of the phase shift due to non-minimum phase properties and to minimum phase properties of the process respectively.

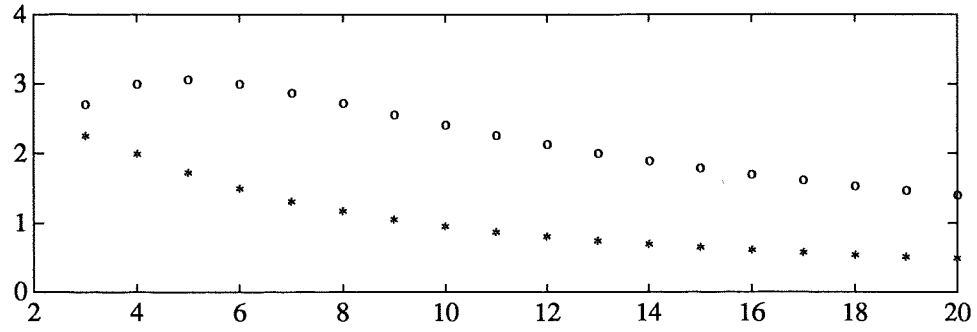


Figure 6.2 The slope n_l at the frequencies ω_π (stars) and $\sqrt{3}\omega_\pi$ (circles) as function of the degree n in Example 6.2.

A high order process with monotone step response may be approximated by a lower order process model with a time delay for frequencies below and in the neighborhood of ω_π. An example will suggest an upper limit on the process model degree n.

EXAMPLE 6.2
Consider the process (6.15) with $\tau = 0$ and $n \geq 3$. The slope of the frequency response magnitude is given by (6.17). The frequency ω_π satisfies $\omega_\pi T = \tan(\pi/n)$. Then the slope at the frequency $\omega = \omega_\pi m$ is given by

$$n_l(\omega_\pi m) = \frac{\left(m \tan \frac{\pi}{n}\right)^2}{1 + \left(m \tan \frac{\pi}{n}\right)^2} n$$

which is shown in Figure 6.2. The stars represent the slope at ω_π and the circles represent the slope at $\sqrt{3}\omega_\pi$ for different n. The maximal slope at ω_π is achieved for $n = 3$. It is $n_l(\omega_\pi) = 2.25$. The maximum slope at $\sqrt{3}\omega_\pi$ is $n_l(\sqrt{3}\omega_\pi) \approx 3.06$. □

This example shows that it is sufficient to consider process models (6.15) with $1 \leq n \leq 3$. The reason is that the degree n of the model is chosen from the estimated slope at the frequency $\sqrt{3}\omega_{osc}$, and the slope $n_l(\sqrt{3}\omega_{osc}) \leq n_l(\sqrt{3}\omega_\pi)$. If there also is a time delay in the process, the phase will decrease faster and the slope will be less steep at the frequency $\sqrt{3}\omega_{osc}$.

For a process (6.15) of degree n with considerable time delay, the local slope $n_l(\sqrt{3}\omega_{osc})$ is less than $n - 1$. To choose a model of the same degree n as the process it is required that

$$n_l(\sqrt{3}\omega_{osc}) > n - 1$$

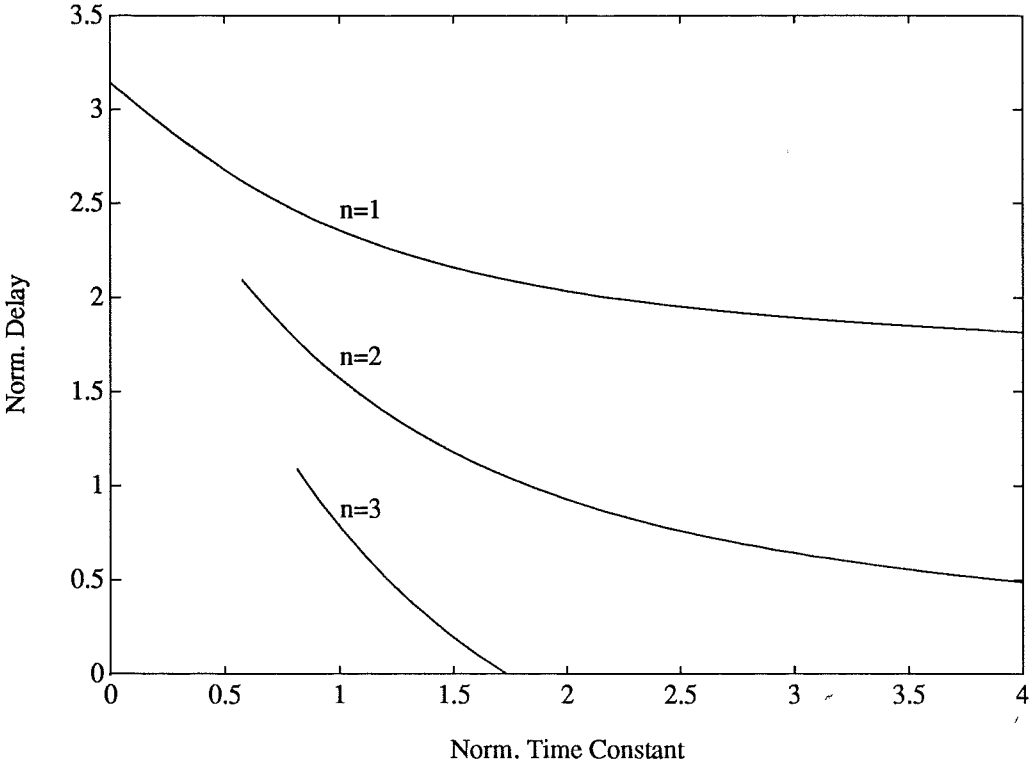


Figure 6.3 Relation between the normalized time constant $\omega_\pi T$ and the normalized delay $\omega_\pi \tau$ for processes (6.15) of degree one, two and three.

It then follows from Equation (6.17) that this is equivalent to

$$\omega_{osc} T > \sqrt{\frac{n-1}{3}}$$

Since $\omega_\pi T > \omega_{osc} T$, a lower limit on the normalized time constant may now be stated for models (6.15) of degree two and three.

$$\begin{cases} \omega_\pi T > \omega_{osc} T > \sqrt{\frac{1}{3}}, & \text{if } n = 2 \\ \omega_\pi T > \omega_{osc} T > \sqrt{\frac{2}{3}}, & \text{if } n = 3 \end{cases} \tag{6.18}$$

Figure 6.3 shows the relation between the normalized time constant $\omega_\pi T$ and the normalized delay $\omega_\pi \tau$ for processes (6.15) of degree $1 \leq n \leq 3$. Only curves for processes satisfying (6.18) are displayed. Third order processes have short normalized delays. Second order processes may have longer normalized delays, but still limited. Processes with dominating delays will have

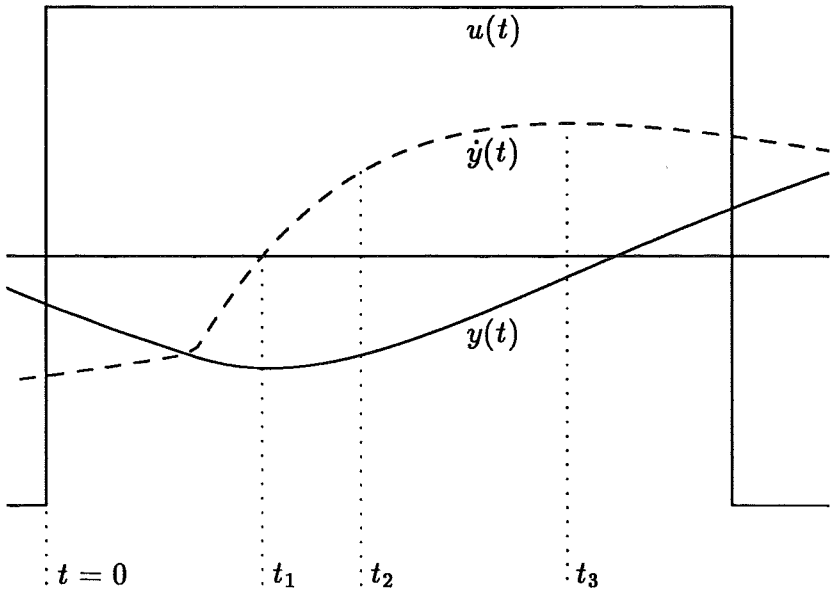


Figure 6.4 Half a period of oscillation for definition of times t_1, t_2, and t_3.

large normalized delays $\omega_\pi \tau$ and consequently small normalized time constants $\omega_\pi T$. Such processes will be considered as first order, and thus modeled by a first order model. Higher order processes are modeled by second or third order models for moderate delays.

Time Domain Estimates

There are several useful features that can be extracted from the the waveform. To see this, the waveform is analyzed for some simple processes. The features that will be considered are the three times t_1, t_2, and t_3 that are shown in Figure 6.4. This figure shows the input $u(t)$, the output $y(t)$, and its derivative $\dot{y}(t)$ for half a period of oscillation. Let $t = 0$ be chosen as the time for the relay switch.

- The output y achieves it extremum at time t_1.
- The first extremum of the derivative $\dot{y}$ for $t \geq t_1$ is at time t_3.
- The derivative $\dot{y}$ reach 63 % of it extreme value at t_2.

It then follows that $0 \leq t_1 \leq t_2 \leq t_3$. It is assumed that the derivative $\dot{y}$ reaches its extreme value before the delayed input due to the next relay switch affects the output, i.e

$$t_3 \leq \frac{T_{osc}}{2} + \tau$$

This inequality may be violated for processes without delay when the hysteresis level ϵ is very small.

It may sometimes be difficult to estimate t_1, t_2, and t_3. Consider e.g. a process that resembles an integrator, then $|\dot{y}|$ switches between two constant levels. For such a process t_2 and t_3 are undefined. For nonminimum-phase processes the output initially moves in the wrong direction which leads to erroneous estimation of the time t_1 for the amplitude peak. Measurement noise corrupts the process output y. This makes it difficult to differentiate y. However, the signal quality can be improved considerably through filtering. Since the signals are stored during the relay experiment, non-casual filtering may be employed, e.g. taking mean value over several periods.

The times t_1, t_2, and t_3 are used to estimate the parameters of the model (6.15). Depending on the degree n, the times are used to determine either the time constant T or the delay τ. Given ω_{osc} and $\arg G(i\omega_{osc})$ the other parameter τ or T can be calculated from (6.16). Processes of degree $n \leq 3$ are analyzed below.

First Order Process

The first order process is described by the transfer function (6.15) with $n = 1$. It gives the differential equation

$$T\dot{y}(t) + y(t) = bu(t - \tau) \tag{6.19}$$

Under relay feedback the input signal u is piecewise constant $\pm d$. The solution to (6.19) from the time when the relay switches from $u(t) = -d$ to $u(t) = +d$ for the initial condition $y(0) = -\epsilon$ is

$$y_1(t) = -bd\left(1 - e^{-t/T}\right) - \epsilon e^{-t/T}, \quad 0 \leq t \leq \tau$$

At $t = \tau$ the delayed input signal will affect the process output. The peak amplitude $\max|y|$ is reached here, and it is given by

$$\max|y| = bd\left(1 - e^{-\tau/T}\right) + \epsilon e^{-\tau/T}$$

It then follows that the time t_1 is given by $t_1 = \tau$ for this first order process. For t larger than τ the solution to (6.19) is given by

$$y_2(t) = bd\left(1 + e^{-t/T}\right) - 2bde^{-(t-\tau)/T} - \epsilon e^{-t/T}, \quad t \geq \tau \tag{6.20}$$

The peak of the output derivative $|\dot{y}|$ is achieved at $t_3 = \tau^+$. For this first order process it follows that $t_1 = t_2 = t_3$.

The relay switches again after half a period of the oscillation. It then follows that $y_2(T_{osc}/2) = \epsilon$. From (6.20) the period of oscillation is given by

$$T_{osc} = 2T \log\left(\frac{2e^{\tau/T} - 1 + \epsilon/bd}{1 - \epsilon/bd}\right) \tag{6.21}$$

If the delay $\tau = 0$ it follows from Taylor expansion of (6.21) that

$$T_{osc} \approx \frac{4T\epsilon}{bd}$$

The period of oscillation does then not reflect the process dynamics. For this type of processes $\max|y| = \epsilon$, that may be used to identify this type of processes. Another special case is when the process is an integrator with delay, i.e. $T \to \infty$. A similar analysis yields that

$$T_{osc} \approx 4\tau + \frac{4\epsilon}{bd}$$

The process then oscillates with a triangular wave form. The derivative of the process output switches between two constant levels, with the same magnitude but with different signs. If the phase shift of the process is dominated by the delay, i.e. $\tau \gg T$, the period of oscillation is approximately

$$T_{osc} \approx 2\tau$$

Second Order Process

Consider a process with the transfer function

$$G(s) = \frac{b}{(sT_1 + 1)(sT_2 + 1)} e^{-s\tau} \tag{6.22}$$

For this process it follows that

$$t_3 - t_1 = T_1 T_2 \frac{\log T_1/T_2}{T_1 - T_2} \tag{6.23}$$

with t_1 and t_3 as the time instants for extrema in $|y|$ and $|\dot{y}|$ respectively. To see this consider the situation when there is a limit cycle. Let $t = 0$ be the time when the delayed input affects the second order dynamics. The solution of the differential equation corresponding to (6.22) from $t = 0$ is

$$y(t) = bd - \frac{T_1(bd + y_0 + T_2\dot{y}_0)e^{-t/T_1}}{T_1 - T_2} + \frac{T_2(bd + y_0 + T_1\dot{y}_0)e^{-t/T_2}}{T_1 - T_2}$$

Calculation of $\dot{y}(t_1') = 0$ and $\ddot{y}(t_3') = 0$ gives after some simplification

$$\exp\left(\frac{t_1'}{T_2} - \frac{t_1'}{T_1}\right) = \frac{bd + y_0 + T_1\dot{y}_0}{bd + y_0 + T_2\dot{y}_0}$$

and

$$\exp\left(\frac{t_3'}{T_2} - \frac{t_3'}{T_1}\right) = \frac{T_1}{T_2} \cdot \frac{bd + y_0 + T_1\dot{y}_0}{bd + y_0 + T_2\dot{y}_0}$$

These equations give t_1' and t_3'. Then $t_3' - t_1' = t_3 - t_1$ establishes (6.23).

From the limit for (6.23)

$$\lim_{T_1 \to T_2} T_1 T_2 \frac{\log T_1/T_2}{T_1 - T_2} = T_1$$

it follows in the special case where $T = T_1 = T_2$ that

$$t_3 - t_1 = T \tag{6.24}$$

This important relation will be used to estimate the time constant of the model (6.15).

Another special case of (6.22) is when $T_2 \to \infty$. Then the process is given by

$$G(s) = \frac{b}{s(sT + 1)} e^{-\tau s}$$

The first derivative $\dot{y}$ of the process output will then increase until the next relay switch affects the dynamics. This means that T may be overestimated and consequently τ will be underestimated. Instead it is better to use the time instant t_2. For a second order process with multiple poles like (6.15) it follows that

$$t_2 - t_1 \approx 0.320T$$

which can be used for estimation of T.

Third Order Process

A third order process described by (6.15) with $n = 3$ can be analyzed in the same way as the second order process above. The analysis does not give simple explicit expressions, so it will be done numerically instead. The intervals $t_3 - t_1$ and $t_2 - t_1$ depend in a complicated way on the delay τ. The first interval satisfies $t_3 - t_1 \leq 2$ where equality is achieved for processes with

Table 6.2 Oscillation analysis of a third order process.

$\arg G(i\omega_{osc})$	$\omega_{osc}T$	$(t_3 - t_1)/T$	$(t_2 - t_1)/T$
$-17\pi/18$	0.82	1.58	0.58
	1.31*	1.52	0.55
$-16\pi/18$	0.82	1.58	0.57
	1.31*	1.52	0.55
$-15\pi/18$	0.82	1.57	0.57
	1.19	1.54	0.55
$-14\pi/18$	0.82	1.59	0.58
	1.06	1.55	0.56
$-13\pi/18$	0.82	1.58	0.58
	0.94	1.55	0.56

a dominating delay. However, these are not considered as third order since the slope is too flat at $\sqrt{3}\omega_{osc}$.

The properties of this third order process under relay feedback are analyzed for different arguments $\arg G(i\omega_{osc})$ and for different ratios τ/T. The period of oscillation T_{osc} and the instants t_1, t_2, and t_3 were determined. Table 6.2 displays properties for different $\arg G(i\omega_{osc})$. The argument depends on the hysteresis level and the oscillation amplitude as shown by (6.3). The second column shows the lower and upper values for $\omega_{osc}T$. The lower limit is given by (6.18). The upper limit is either given by $\omega_{osc}T = \tan(-\arg G(i\omega_{osc})/3)$ or it is the limit for which $t_3 \leq T_{osc}/2 + \tau$. This limit for too short period is obtained for $\arg G(i\omega_{osc})$ that are close to $-\pi$. These values for $\omega_{osc}T$ are marked with an asterisk. The two last columns show the corresponding values of the normalized intervals $(t_3 - t_1)/T$ and $(t_2 - t_1)/T$.

Table 6.2 cover third order process with $\omega_{osc}T$ in the interesting interval for relay oscillations. The quantities in the two last columns not are varying much. Therefore the following approximations are suggested.

$$t_3 - t_1 \approx 1.55T$$
$$t_2 - t_1 \approx 0.55T$$

They will be used for the estimation of the time constant T.

Summary

The Model (6.15) is used to characterize processes using information from an experiment with relay feedback. The slope $\hat{n}$ and the argument $\arg G(i\omega_{osc})$ are determined using discrete Fourier transforms. This gives a process model

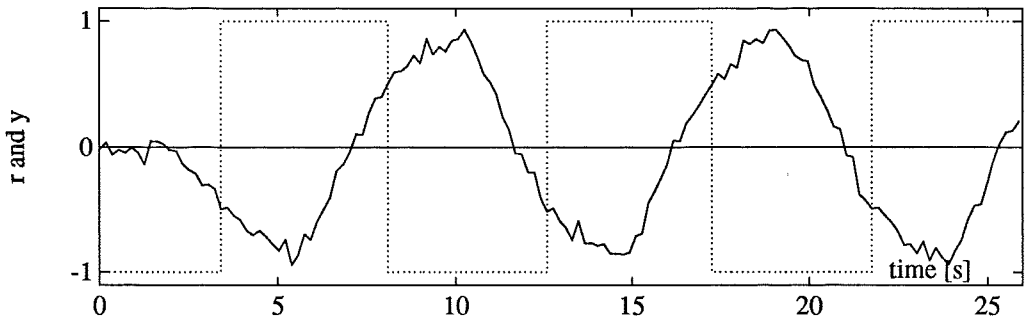


Figure 6.5 Process input and output during relay experiment.

of first, second, or third order. For first order models the delay τ is determined from the time instant for extreme y, i.e. $\tau = t_1$. The time constant is then given by (6.16), i.e.

$$T = \frac{1}{\omega_{osc}} \tan\bigl(-\arg G(i\omega_{osc}) - \omega_{osc}\tau\bigr)$$

For higher order models the time constant is determined from the numbers $t_3 - t_1$ or $t_2 - t_1$. The time constant is estimated by

$$T = \min\bigl(t_3 - t_1, (t_2 - t_1)/0.32\bigr)$$

for a second order model, and by

$$T = \min\bigl((t_3 - t_1)/1.55, (t_2 - t_1)/0.55\bigr)$$

for a third order model. Then (6.16) gives the delay

$$\tau = \frac{1}{\omega_{osc}}\bigl(-\arg G(i\omega_{osc}) - n \arctan \omega_{osc} T\bigr)$$

An example will now demonstrate the use of the model (6.15) for process characterization.

EXAMPLE 6.3
Consider the process

$$G_0(s) = \frac{1}{(s+1)^2} e^{-2s}$$

where white noise with standard deviation $\sigma_n = 0.05$ is added to the output. Figure 6.5 shows the process input (dotted line) and the process output (solid

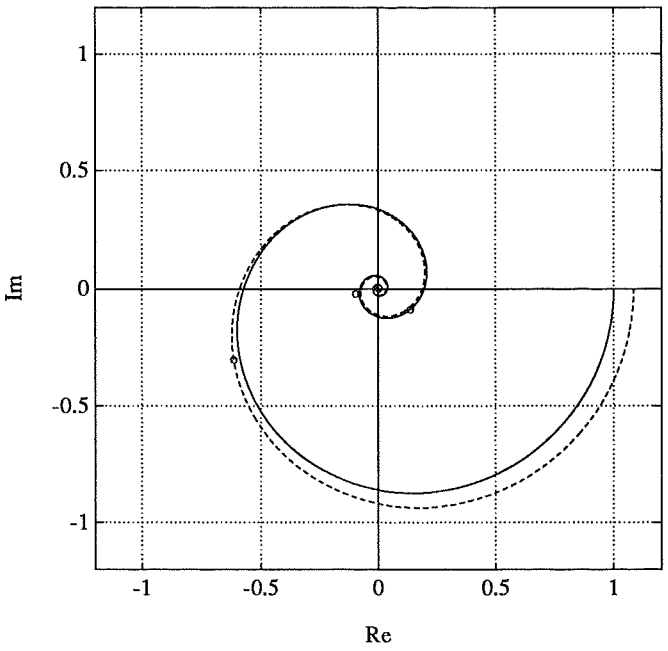


Figure 6.6 Frequency responses in Example 6.3 for the true process $G_0(i\omega)$ (solid line) and for the estimated model $G(i\omega)$ (dashed line). The circles show the frequency response estimate $\hat{H}(e^{i\Omega_m})$ for $m = \{1, 3, 5\}$.

line) during the relay experiment. At $t \approx 8$ s it is assumed that a stationary oscillation is established. The following two periods were then analyzed. The oscillation frequency is $\omega_{osc} = 0.71$ radians/s. The slope of the frequency response was estimated to $\hat{n} = 1.32$ using ω_{osc} and $3\omega_{osc}$. This gives $n = 2$ for the model $G(s)$ in (6.15). The mean value over three half periods of the process output gave a signal that was so smooth that it could be differentiated. Analysis of the signal gave $t_3 - t_1 = 1.10$ s and $(t_2 - t_1)/0.32 = 1.08$ s, which yields $T = 1.08$ s. The phase shift at the oscillation frequency is $\arg G(i\omega_{osc}) = -2.7$ radians. The delay τ is then calculated to $\tau = 1.97$ s. The solid line in Figure 6.6 shows the frequency response $G_0(i\omega)$ of the true process, the dashed line shows the frequency response of the model $G(i\omega)$ where the gain is adjusted such that $|G(i\omega_{osc})| = |G_0(i\omega_{osc})|$. The circles show the frequency response estimate $\hat{H}(e^{i\Omega_m})$ for $m = \{1, 3, 5\}$, which was obtained using discrete Fourier transforms. The curves in the figure show that the estimated model is a good approximation of the true process. □

6.4 Conclusions

A system with relay feedback system was analyzed in this chapter. It was assumed that a unique limit cycle is achieved when the process is controlled with a relay with hysteresis.

An analysis of the output from the process under relay feedback reveals interesting properties of the process. The information discussed here are

- The oscillation frequency $\omega_{osc} = 2\pi/T_{osc}$
- The frequency response $\hat{H}(e^{i\Omega_m}), m \in \mathcal{M}$
- The times t_1, t_2 and t_3

A simple model (6.15) is determined from this information. The model is used to characterize the process. It is described by the degree n, the time constant T, and the delay τ. This model will be used in Chapter 7 for initializing an adaptive controller.

7

Initialization of Adaptive Controllers

This chapter is concerned with initialization of an indirect adaptive controller. The controller and the information required to start the controller were presented in Chapter 5. The idea is to use an experiment with relay feedback, discussed in Chapter 6, to provide initial data. In this Chapter it will be shown how information from the experiment with relay feedback may be used to design and commission the adaptive controller. Properties of the desired closed loop system are discussed in Section 7.1. In Section 7.2 and 7.3 it is discussed how the model structure can be chosen and how the estimator is initialized. Section 7.4 presents an algorithm for initialization of an indirect adaptive controller from an experiment with relay feedback. Some examples, given in Section 7.5, demonstrate its properties for some different processes.

7.1 Properties of the Closed Loop System

The achievable properties of the closed loop system depend critically on properties of the open loop process. Reasonable specifications on the closed loop system is a trade-off between command signal following, load disturbance rejection, measurement noise attenuation, and robustness for process model uncertainty. The robust design approach in Chapter 3 and 4 provided a trade-off between these objectives. It is, however, too complicated to perform on line in an adaptive controller.

Heuristics

The response time of a system may be characterized by the frequency ω_π where the phase lag is π radians. The response time consists of two parts, the pure time delay and the time for the transient. These parts are characterized by the normalized delay $\omega_\pi\tau$ and the normalized time constant $\omega_\pi T$ that were defined in Chapter 6. They correspond to the non-minimum phase and to the minimum phase properties respectively.

Define $\omega_{\pi l}$ as the frequency where the loop transfer function L has π radians phase lag. The desired closed loop system is faster than the open loop system if $\omega_{\pi l} > \omega_\pi$. This increase in speed is obtained with phase lead in the feedback controller. To understand the achievable performance two cases are considered. If the slope of the phase curve is small, as for processes with small normalized delay $\omega_\pi\tau$, a controller with moderate lead may increase $\omega_{\pi l}$ considerably. This will decrease the response time of the closed loop system. On the other hand, if the phase curve is steep, the response time is dominated by the delay and may only be decreased marginally. Increasing $\omega_{\pi l}$ requires a controller with considerable phase lead.

The required phase lead in the controller increases with the decrease in response time. A controller with much phase lead must either be of high order or be unstable. A high order controller may be difficult to tune. On the other hand, an unstable controller in series with a stable process yields a loop transfer function with a Nyquist curve that must encircle -1. This may be bad for the robustness of the system [Lilja, 1989]. An unstable controller may also create problems if the actuator saturates. Another drawback is that unstable controllers may cause limit cycles in systems with static friction [Wallenborg, 1987]. A fourth reason to avoid unstable controllers are the precautions that have to be taken for safe changes between different operating modes, e.g. between manual and automatic control. It is, thus, important not to require too much phase lead in the controller.

Specifications for the Pole Placement Controller

The adaptive controller, described in Chapter 5, operates in discrete time. The closed loop discrete time transfer function from reference input to plant output is

$$H_m^*(q^{-1}) = \frac{q^{-d}A_m^*(1)B^*(q^{-1})}{B^*(1)A_m^*(q^{-1})} \tag{7.1}$$

where $B^*(q^{-1})$ and d are the delay and the numerator polynomial of the discrete time process model (5.1). This choice of $H_m^*(q^{-1})$ means that the process zeros are not canceled by the controller. The controller (5.3) has integral action and its numerator contains the factor $1 + q^{-1}$, as suggested

in Chapter 5. The controller is obtained from the Diophantine-Aryabhatta-Bezout equation (5.4) with

$$A_c(q^{-1}) = A_m^*(q^{-1})A_o^*(q^{-1})$$

Given the structure of the control system, the closed loop performance is mainly affected by the choice of the polynomials $A_m^*(q^{-1})$ and $A_o^*(q^{-1})$, and the sampling interval h. The degrees of A_m^* and A_o^* are both chosen equal to the degree of the process model A^*. The degree of A^* is given by the slope of the frequency response magnitude, which can be obtained from the experiment with relay feedback. This will be discussed below. Since the real process is continuous it is convenient to give closed loop specifications in continuous quantities. The characteristic polynomials $A_m^*(q^{-1})$ and $A_o^*(q^{-1})$ are discrete time counterparts to continuous time polynomials. These are parametrized by ω_0, which is the distance from the poles to the origin. The polynomials are chosen as

$$\begin{aligned} A_{mc}(s) &= s + \omega_0 \\ A_{oc}(s) &= s + \omega_0 \end{aligned} \tag{7.2}$$

for first order processes, as

$$\begin{aligned} A_{mc}(s) &= s^2 + \sqrt{2}\omega_0 s + \omega_0^2 \\ A_{oc}(s) &= s^2 + \sqrt{3}\omega_0 s + \omega_0^2 \end{aligned} \tag{7.3}$$

for second order processes and as

$$\begin{aligned} A_{mc}(s) &= (s + \omega_0)(s^2 + \sqrt{2}\omega_0 s + \omega_0^2) \\ A_{oc}(s) &= (s + \omega_0)(s^2 + \sqrt{3}\omega_0 s + \omega_0^2) \end{aligned} \tag{7.4}$$

for third order processes. These continuous time polynomials have their poles spread on a circle with radius ω_0 and with center at the origin.

Choice of Sampling Interval

In a discrete time control system, the choice of sampling interval is one of the most crucial issues. The sampling interval influences many properties of the closed loop system. Load disturbances and measurement noise that affects the system are not detected and compensated for until a new control signal is affecting the process. A long sampling interval also requires a more accurate model of the process since a sampled system operates in open loop between the samples. One reason for using long sampling intervals are cost, less powerful computers may be used.

The sampling interval limits the closed loop bandwidth that may be obtained. The sampling interval must not be longer than the essential process dynamics is covered in the discrete time model. Here the sampling interval h will be chosen from ω_0 that specifies the closed loop characteristic polynomial. For a first order process the sampling interval is chosen such that $\omega_0 h \approx 0.3$. For second and third order processes the choice of h is given by $\omega_0 h \approx 0.5$ and $\omega_0 h \approx 0.7$ respectively. These choices are within the recommendation in [Åström and Wittenmark, 1989], which says that it is reasonable to select the sampling interval such that there will be 5–20 samples in a closed loop step response.

The recommendations above are reasonable for processes with relatively small time delays. For processes with long time delays another aspect of the sampling interval selection appears. A short sampling interval gives high order process models and the pole placement design then yields a controller of high order. A reasonable choice is to select the sampling interval such that the model of the open loop process has not more than five additional poles at the origin. This correspond to a sampling interval that is a fifth of the time delay.

It will be shown below that there are some advantages to select the sampling interval h as a fraction of the period of oscillation.

$$h = \frac{T_{osc}}{2p} \tag{7.5}$$

The relay output is assumed to be a symmetric square wave. The sampled signal will then also be symmetric. Each half period of the relay output will then have p samples. In [Åström and Hägglund, 1988b], the choice $p = 3$ was motivated from a polynomial equation. Here a larger value on p will be used. The recommendations above on $\omega_0 h$ and the delay d give p between five and ten.

Performance Assignment

In an automatic design approach as in an adaptive controller, the trade-off between noise attenuation and load disturbance rejection may be difficult to consider. Reasonable specifications on the closed loop system will instead be chosen such that the sensitivity function magnitude $|S_o|$ not is too large. This is achieved if the loop transfer function has a frequency response sufficiently far away from -1 in the Nyquist diagram. The design objective is to make the closed loop system as fast as possible with sufficient robustness margins. The controller is also required to be stable. The robustness is considered as acceptable provided that the frequency response of the loop transfer function satisfies

$$\min_{\omega} \left|1 + L\left(e^{i\omega h}\right)\right| \geq 0.5 \tag{7.6}$$

This is equivalent to $|S_o| \leq 2$. The peak of the sensitivity function depends on the parameter ω_0 for the closed loop poles.

To find a rule for a reasonable choice of ω_0, pole placement design is applied on the process (6.15) for many different values of τ and T, and for $n = 1, 2, 3$. A discrete time integrating controller (5.3) is designed to give a closed loop characteristic polynomial $A_o^*(q^{-1})A_m^*(q^{-1})$ that corresponds to (7.2), (7.3), or (7.4) depending on the degree n. The numerator of the controller is required to contain the factor $1 + q^{-1}$, which was discussed in Chapter 5. The sampling interval h satisfies $\omega_0 h = 0.1 + 0.2n$. The parameter ω_0 is increased until the sensitivity reaches the value 2 or the controller becomes unstable. In this way the largest values of ω_o that can be achieved subject to the constraints are obtained. The results are illustrated in Figure 7.1. The figure shows that ω_0 is related to ω_π, and that the relation is conveniently expressed in terms of the normalized time constant $\omega_\pi T$. The empirical results for $n = 1$ are marked with circles ($\circ$). For $n = 2$ they are marked with plus signs ($+$) and for $n = 3$ they are marked with asterisks ($*$). The function

$$\frac{\omega_0}{\omega_\pi} = 0.5 + \frac{1}{\omega_\pi T} + 0.12\left(\frac{1}{\omega_\pi T}\right)^2 \tag{7.7}$$

is a reasonable approximation to the points in Figure 7.1. The solid line shows this function. Notice that $\omega_\pi T$ and $\omega_\pi \tau$ are related through Equation (6.16).

For processes with small relative delay, τ/T, the normalized time constant $\omega_\pi T$ is large. Consider, e.g. a first order process with $T = 1$ and $\tau = 0.1$. For this process Equation (7.7) gives $\omega_0 \approx 0.6\omega_\pi$. This means that the controller is similar to a PID-controller, tuned by Ziegler-Nichols method [Ziegler and Nichols, 1942], which has $\omega_0 \approx 0.7\omega_\pi$ for this process.

For processes with larger relative delay, τ/T, it follows from (7.7) that it is possible to have a significant larger ratio ω_0/ω_π. This is equivalent to classical controllers with dead time compensation. Such controllers can give significant faster response than PID-controllers, tuned by Ziegler-Nichols method, which in this case have $\omega_0 \approx 0.3\omega_\pi$.

For processes with large relative delays the response to command signals and load disturbances are limited by the time delay. There is some benefit in reducing the rise time. The effect is marginal when the rise time is an order of magnitude smaller than the dead time. For first order systems this correspond to $\omega_0\tau = 10$. This means that $\omega_0/\omega_\pi \approx 10/\pi$. It is therefore reasonable to limit the ratio ω_0/ω_π to 4.

The quantity $\omega_0\tau$ is introduced to facilitate comparisons with other design methods. A PID-controller tuned by Ziegler-Nichols method gives $\omega_0\tau \approx 1$ [Åström *et al.*, 1989]. In [Middleton, 1991] it is suggested that

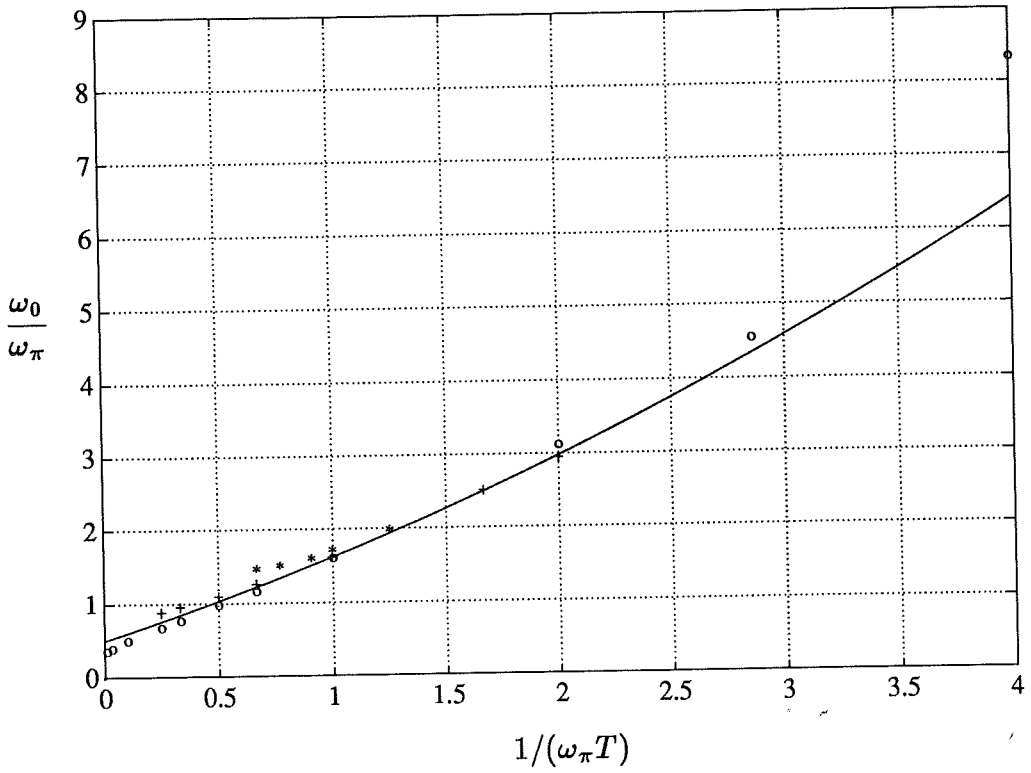


Figure 7.1 Relation (7.7) between ω_0 and ω_π. Empirical results for $n = 1$ (o), $n = 2$ (+), and $n = 3$ (*).

$\omega_0 \tau \leq 2$. The design rule (7.7) gives a value of $\omega_0 \tau$ that ranges from 1 to 12. The number increases with the ratio τ/T. Equation (7.7) gives ω_0 as a function of T and ω_π. Normally $\omega_{osc} < \omega_\pi$, but it has been found that ω_{osc} can be used as an approximation of ω_π.

7.2 Choice of Model Structure

The indirect adaptive controller identifies parameters in the discrete time process model (5.1). The model structure is determined by the degree of the polynomials, $\deg B^*$ and $\deg A^*$, and the delay d. The structure of this model should reflect the properties of the real process in a certain frequency interval.

It is neither of interest to have a much more complex model of the process than the process itself nor to have a too simple process model with few parameters. The model should cover the essential process dynamics. If there are too many parameters in the model of the process, higher demands

are required on the excitation in the system to make the estimator converge and to minimize the variance of the estimated parameters. If the model is too complex it may have almost common zeros and poles which may lead to numerical problems in the pole placement design. If the model class is too simple it will not cover the real process dynamics. The nature of the excitation will determine the identified model. This may differ dramatically from the real process [Rohrs *et al.*, 1985].

High order continuous processes may accurately be described by a lower order process model at low frequencies. In a sampled system the frequency content of the signals is limited by the Nyquist frequency $\omega_n = \pi/h$. With sufficiently slow sampling the steep high frequency slope of the frequency response for a high order process will not be represented by the discrete time process model. In most cases the process may be well described by a model with some poles that model the dominating dynamics and with a delay to achieve a correct phase response.

Model Order

The order of the process model is $d + \deg A^*$. The order may be reflected in the frequency response. For processes without zeros, the high frequency slope of the frequency response magnitude is equal to $\deg A^*$, but if the process model contains zeros this is not true. However, the degree of $A^*(q^{-1})$ will here be estimated as if no zeros were present. The degree is limited to $\deg A^* \leq 3$, as discussed in Chapter 6.

The degree of $A^*(q^{-1})$ will be estimated from the slope of the frequency response magnitude. The slope of the frequency response $\hat{n}$ is estimated of from the first two harmonics in the relay oscillation. The slope estimate is given by (6.11). The degree of $A^*(q^{-1})$ is chosen as

$$\deg A^*(q^{-1}) = \begin{cases} 1, & \text{if } \hat{n} \leq 1 + \Delta_{12} \\ 2, & \text{if } 1 + \Delta_{12} < \hat{n} \leq 2 + \Delta_{23} \\ 3, & \text{if } 2 + \Delta_{23} < \hat{n} \end{cases} \tag{7.8}$$

where Δ_{12} and Δ_{23} are thresholds, introduced not to select a higher order model than necessary due to the uncertainty in $\hat{n}$. The thresholds can be chosen from the uncertainty in the estimate of the slope.

Time Delay and Numerator

Once the degree of $A^*(q^{-1})$ is chosen it is possible to select the structure of the process model numerator. This may be viewed as a high order polynomial, whose d leading coefficients are zero and the remaining coefficients form the polynomial $B^*(q^{-1})$. The time delay will appear in the desired closed loop transfer function. An accurate estimate of the delay d is essential.

When sampling a continuous time process, the degree of $B^*(q^{-1})$ is generically equal to $\deg A^*$. It thus seems natural to select $\deg B^* = \deg A^*$ for the discrete time process model (5.1). Sampling of a process with a wide spread of poles gives a sampled system with poles close to the origin. Such systems are conveniently approximated by models with $\deg A^* < \deg B^*$. Models of this type which have the advantage that they use fewer parameters have been found useful in several applications [Åström, 1980, Bengtsson, 1989, Åström *et al.*, 1991].

The structure of the process model numerator is chosen as follows. Consider the process model

$$H_0^*(q^{-1}) = \frac{b_{01}q^{-1} + b_{02}q^{-2} + \cdots + b_{0m}q^{-m}}{1 + a_{01}q^{-1} + a_{02}q^{-2} + \cdots + a_{0n}q^{-n}} \tag{7.9}$$

where $n = \deg A^*$ and m is the smallest integer satisfying $m \geq \deg A^* + \tau/h$. Estimate the parameters in this model using least squares estimation. The discrete time signals $u(k)$ and $y(k)$ for the estimation are provided by interpolating the signals from the experiment with relay feedback. The sampling period is given by (7.5). The sampling is synchronized with the relay switches.

Significant coefficients in the numerator of (7.9) are considered as coefficients in $B^*(q^{-1})$ and the rest with small magnitude are neglected. Leading coefficients will have small magnitude for processes with delay. The following method is proposed to select the degree of $B^*(q^{-1})$ and the delay d: Initially let $\deg B^* = \deg A^*$. Then choose the delay in the process as the integer d that maximizes

$$J(d) = \sum_{i=d}^{\deg B^*+d} |b_{0i}| \tag{7.10}$$

where b_{0i} are the estimated coefficients in the numerator of (7.9). This objective means that the coefficients in $B^*(q^{-1})$ are the $\deg B^* + 1$ consecutive b_{0i}'s with the largest sum of coefficient magnitude.

If the neglected coefficients have magnitudes considerably lower than the largest $|b_{0i}|$, this is an indication that the model class is appropriate. If not, the degree of $B^*(q^{-1})$ is increased. A coefficient b_{0k} can be neglected if it satisfies

$$|b_{0k}| < \kappa \max_{i \in I_{B^*}} |b_{0i}|, \qquad I_{B^*} = \{i : d \leq i \leq d + \deg B^*\} \tag{7.11}$$

where the choice $\kappa = 0.5$ has turned out well in simulations. From experience it is recommended to consider coefficients b_{0k} only for $k < d$ here. If any of

these do not satisfy (7.11) the degree $\deg B^*$ is increased to let $B^*(q^{-1})$ contain these coefficients. The delay d is decreased so that

$$d + \deg B^* = \text{constant}$$

Trailing coefficients with large magnitude will always be neglected since these may take strange values in the estimation due to insufficient excitation from the experiment with relay feedback. The input spectrum is nonzero only for certain frequencies and the energy decays as $1/\omega$. Another rule for estimating the delay is proposed in [Isermann, 1980].

7.3 Initialization of the Estimator

The adaptive controller was described in Chapter 5. It is based on recursive identification of a process model. The structure of the model is chosen as suggested above.

Initial Parameters

The least squares estimator (5.2) updates the parameter estimates $\theta(k)$ and the covariance matrix $P(k)$ every sample. If the model structure is appropriate and the input in persistently exciting the estimates will converge to some point θ^*. A bad choice of the initial model $\theta(0)$ may give a large transient after startup. Such a transient may be avoided if information from the experiment with relay feedback is used to assign the initial $\theta(0)$ and $P(0)$.

First the model (6.15) is sampled with the sampling interval h for the adaptive controller. This gives $\theta(-k_r)$ where k_r is the length of the experiment with relay feedback. This estimate is then improved by recursive least squares estimation, using the stored input and output signals from the experiment with relay feedback. This estimation gives $\theta(0)$ and $P(0)$.

The process input is a square wave during the experiment with relay feedback. Its spectrum is non-zero for frequencies ω_{osc}, $3\omega_{osc}$, etc., and is inversely proportional to frequency. This means from a practical point of view that the excitation is sufficient for estimating the parameters in a model with four parameters. Thus, models of first and second order, that are estimated using the signals from the experiment with relay feedback, can be expected to be good estimates of the process. A third order model requires more excitation to provide an accurate estimate of the process.

Forgetting Factor

Discounting of old data is essential in all adaptive algorithms. It is also known that discounting may cause some problems [Åström and Wittenmark, 1989]. There are many ways to do the discounting. The particular choice depends on the nature of the parameter variations. Ultimately it would be desirable to let the algorithm explore this itself. Good ways to do this have not been explored. The forgetting factor would be chosen with respect to the sampling period. In a self-tuning controller, where tuning is switched off after some time when the controller parameters have converged, the forgetting factor λ can be chosen close to 1.

Regression Filters

Least squares estimation is based on minimization of a quadratic loss function that is formulated in the time domain. The loss function may also be interpreted in the frequency domain [Ljung, 1987, Wittenmark, 1989]. With the filter $H_f^*(q^{-1})$, weighting is introduced in the loss function. The passband of the filter determines the frequency range where the Nyquist curves of the identified model and the real process should be close. For a robust design it is important that the frequency response of the estimated model is accurate in the frequency range where the Nyquist curve for the loop transfer function L is close to -1.

The processes under consideration here are stable. The specifications are chosen to give a stable controller. Then the phase of the loop transfer function decreases for frequencies above ω_π. An accurate loop transfer function L is then required in an interval surrounding the frequency ω_f, which is defined by $\arg L(i\omega_f) = -3\pi/4$. For a process with dominating delay it follows that $\omega_\pi \tau \approx \omega_{\pi l} \tau \approx \pi$. Then a good choice is $\omega_f = 3\omega_\pi/4$. For a process without delay the important interval is in the neighborhood of ω_0. The ratio ω_0/ω_π in Figure 7.1 for second and third order processes yield that ω_f/ω_π is between one and two. The phase response is, however, flat for this type of process and the suggestion above for ω_f will also do very well here. A reasonable bandwidth of the filter is one decade. For large τ the upper limit of the passband of the regression filter is reached for $\arg L \approx -9\pi/4$. The regression filter can then be chosen as

$$H_f^*(q^{-1}) = \frac{k(1-q^{-1})}{A_f^*(q^{-1})} \tag{7.12}$$

where the characteristic polynomial $A_f^*(q^{-1})$ is the discrete time counterpart to the continuous time characteristic polynomial

$$A_f(s) = \left(s + \frac{\omega_f}{\sqrt{10}}\right)\left(s^2 + 1.4\sqrt{10}\omega_f s + 10\omega_f^2\right)$$

where $\omega_f = 0.75\omega_\pi$. The filter is sampled with the same interval as the adaptive controller.

7.4 An Initialization Procedure

The choice of design variables for an indirect adaptive controller based on pole placement design and recursive least squares estimation have been discussed in this chapter. The choices originated from the information available from the experiment with relay feedback in Chapter 6. The controller is designed to have a pole in $q = 1$, i.e. integral action, and a zero in $q = -1$ for improved robustness and measurement noise attenuation.

A procedure for automatic initialization of the adaptive controller will now be given. From an operational point of view the algorithm is similar to the procedure for tuning a PID-controller based on relay feedback, see [Åström and Hägglund, 1988a]. The procedure has the following steps.

Step 1: Bring the process output close to the desired set point by manual control and wait for stationarity. Estimate the noise level and determine the hysteresis level and relay amplitude.

Step 2: Introduce relay feedback. Measure the process input and output. Store these values with shortest possible sampling interval. Adjust the relay amplitude to obtain a limit cycle with desired amplitude. Stop the experiment after 2 or 3 periods with a stationary oscillation.

Step 3: Determine the period of oscillation T_{osc}. Estimate the frequency response for the frequencies ω_{osc} and $3\omega_{osc}$ using discrete Fourier transforms. Estimate the slope $\hat{n}$ of the amplitude curve from (6.11). Determine $\deg A^*$ from (7.8). Let $\Delta_{12} = 0.3$ and $\Delta_{23} = 0.1$. Use ω_{osc} as an approximation for ω_π in the sequel.

Step 4: Determine the times when the limit cycle and its derivative have their extreme values. Estimate T and τ of the model (6.15) as proposed in Chapter 6.

Step 5: Determine ω_0 from Equation (7.7).

Step 6: Select the sampling interval as $h = T_{osc}/2p$, where p is the smallest integer such that $\omega_0 h \leq 0.1 + 0.2 \deg A^*$. If $4h < \tau$ the process model will have too many delays. Then decrease p until $4h > \tau$.

Step 7: The closed loop characteristic polynomials are then chosen as the sampled counterpart to (7.2), (7.3), or (7.4). The regression filter is chosen as (7.12) with $\omega_f = 3\omega_\pi/4$.

Step 8: Determine the delay d and $\deg B^*$ for the process model (5.1) using (7.9), (7.10), and (7.11).

Step 9: Initialize the recursive estimator with parameters computed from n, τ, T, and h. Run the estimator with the data from step 2. The estimator will then have the same states as if it had been running since the experiment with relay feedback started.

Step 10: Select a reference input signal that excites the system so that the controller will tune. One possibility is to let a square wave with frequency $\omega_{osc}/3$ excite the system for some periods.

In a self-tuning controller the estimation is switched off after step 10. The parameters are then assumed to have converged during the excitation from the reference input. If the estimation continues, it is from a practical point of view recommended to monitor the excitation. Then it is possible to adapt only, when both the input u and the output y are varying considerably.

7.5 Examples

Part of the heuristics for determining the design parameters, like Figure 7.1 were based on simple models. To determine the validity of the approach the initialization procedure was tested on a large test batch. The results of the initialization procedure for the processes in this batch are given in Appendix A. This section demonstrates the initialization procedure on some examples.

A Second Order Process with Time Delay

An experiment with relay feedback was demonstrated in Example 6.3 for the process

$$G_A(s) = \frac{1}{(s+1)^2} e^{-2s}$$

The information from this experiment will here be used to initialize an adaptive controller. The estimate of the slope $\hat{n} = 1.32$ gives $\deg A^* = 2$. Equation (7.7) gives $\omega_0 = 1.42$ radians/s. Step 6 in the initialization procedure above gives the sampling interval h as follows. The condition $\omega_0 h \leq 0.5$ first gives $h = T_{osc}/26 = 0.34$ s. However, this gives $\tau/h = 5.7$ which is larger than 4. A longer sampling interval is thus required to limit the controller complexity. For $h = T_{osc}/18 = 0.49$ s, it follows that $\tau/h = 4$ which is acceptable. This choice gives $\omega_0 h = 0.71$. The delay of the discrete time process model is then estimated to $d = 5$ and $\deg B^* = 2$.

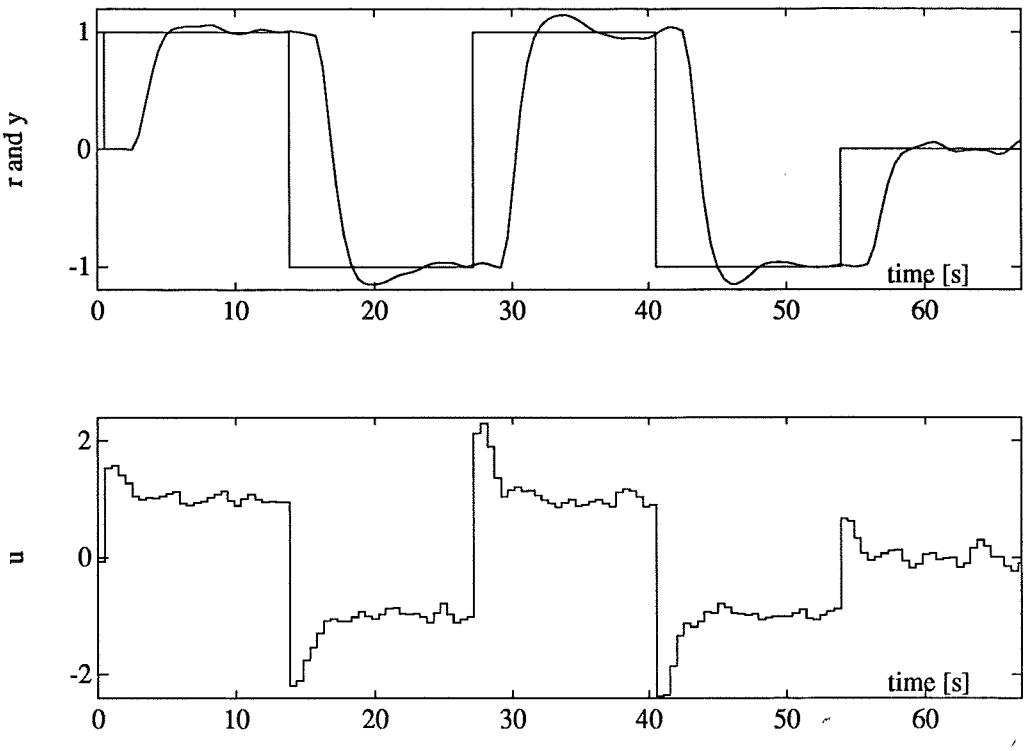


Figure 7.2 Adaptive control of the process G_A.

The adaptive controller is started after the initialization. Figure 7.2 shows a simulation of the system. The standard deviation of the measurement noise is $\sigma_n = 0.05$. Table 7.1 shows the parameters after the initialization procedure ($t = 0$) and after 67 seconds of adaptive control. For comparison, the true parameters of the process are also given. The estimated parameters at $t = 0$ and at $t = 67$ s do not differ much. The estimated parameters deviate somewhat from the true parameters θ_{true}. The two time constants in the model (6.22) can be calculated from the estimated θ. They are $T_1 \approx 0.75$ s and $T_2 \approx 1.8$ s. The main reason for this deviation is insufficient excitation during the relay feedback and the adaptive control. In spite of this, the loop

Table 7.1 True and estimated parameters of the model (5.1).

	a_1	a_2	b_0	b_1	b_2
$\theta(0)$	−1.0850	0.2736	0.1233	0.0317	0.0308
$\theta(67)$	−1.1073	0.2875	0.1057	0.0463	0.0282
θ_{true}	−1.2198	0.3720	0.0819	0.0702	0.0001

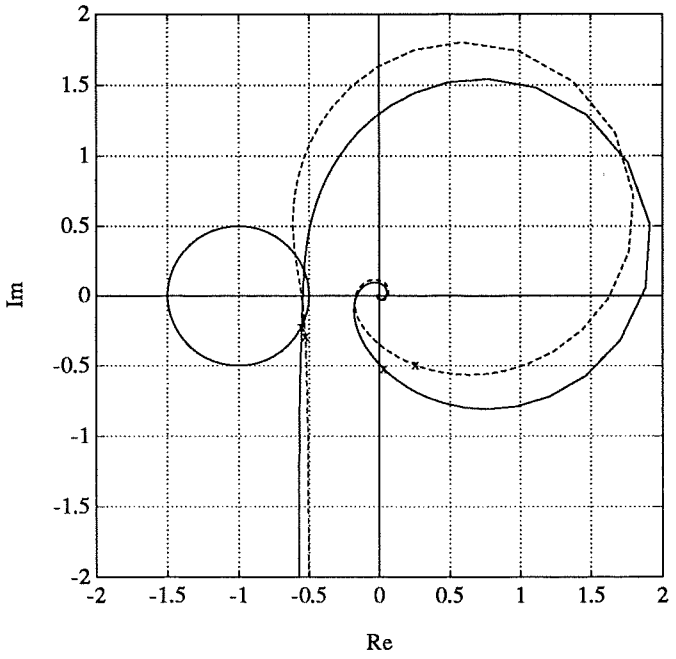


Figure 7.3 Nyquist plot of the loop transfer functions based on the estimated process model $\theta(67)$ (solid line) and on the true process θ_{true} (dashed line). The frequencies ω_{osc} and $3\omega_{osc}$ are marked.

transfer function based on the estimated model $\theta(67)$ and the loop transfer function based on the true process θ_{true} do not differ much. The Nyquist curves of these loop transfer functions are shown in Figure 7.3. The solid curve corresponds to the estimated model and the dashed curve corresponds to the true process. The same eighth order controller, $G_{R1}(q^{-1})$, is used for both curves. This controller is obtained when the adaptation stops at $t = 67$ s. Thus, it is calculated from the model $\theta(67)$. The curves are adjacent at ω_{osc}, which means that the estimate is good in the critical region. The robustness margin is satisfactory since the loop transfer function is sufficiently far away from -1. Notice also that the sensitivity of the closed loop system is slightly larger than 2 although the specifications for this system were chosen to give $|\mathcal{S}_o| \leq 2$. This is due to the fact that the excitation is not able to make $\|\theta(67) - \theta_{true}\| = 0$.

Comparison with PID-Control. The controller $G_{R1}(q^{-1})$ is compared with two other controllers to assess the quality of this adaptive design method:

- The PI-controller $G_{R2}(s)$ is tuned automatically using a method that is recommended for processes with long time delays [Hägglund and Åström, 1991]. Its parameters are $K = 0.25/|G_A(i\omega_{osc})| = 0.34$ and $T_I = 1.6/\omega_{osc} = 2.27$.

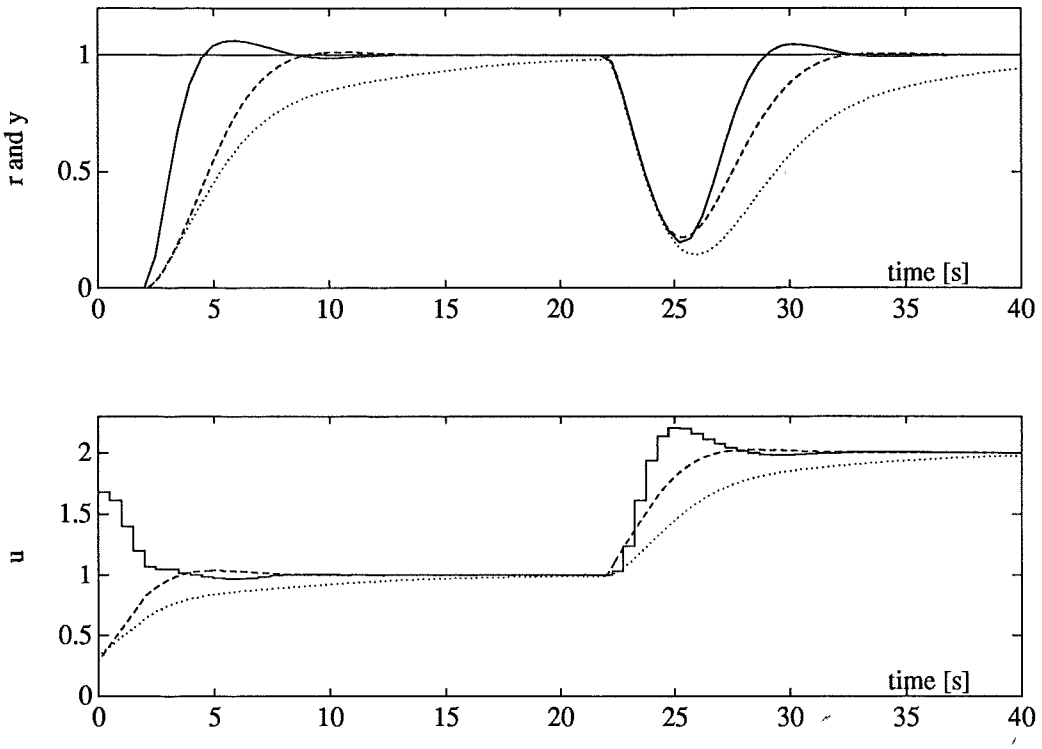


Figure 7.4 Comparison between different tuning methods. Simulation using the controllers G_{R1} (solid line), G_{R2} (dotted line), and G_{R3} (dashed line).

- The PID-controller $G_{R3}(s)$ is determined using dominant pole design [Persson, 1992]. This method gives a very well tuned PID-controller with $K = 0.58$, $T_I = 2.21$, and $T_D = 0.56$.

Figure 7.4 shows simulations where the process G_A is controlled with these fixed controllers. An input load disturbance affects the process at $t = 20$ s. The solid line corresponds to G_{R1}, the dotted line to G_{R2}, and the dashed line to G_{R3}. The behavior of the system for the controller G_{R1} is satisfactory. The response for the controller G_{R2} is very slow. This type of automatic tuning is not suited for this type of process. A well tuned PID-controller can, however, do a good job as the response for G_{R3} indicates.

Influence of Sampling Interval. The controller G_{R1} is compared with two other controllers to demonstrate different choices of sampling interval. The other controllers are obtained by similar initialization procedures followed by adaptive control for 67 s. The only difference from the way G_{R1} was obtained is the choice of sampling interval:

- The sampling interval is chosen as $h = T_{osc}/26 = 0.34$ s for the controller $G_{R4}(q^{-1})$. This h was initially proposed above to meet $\omega_0 h \leq 0.5$. It

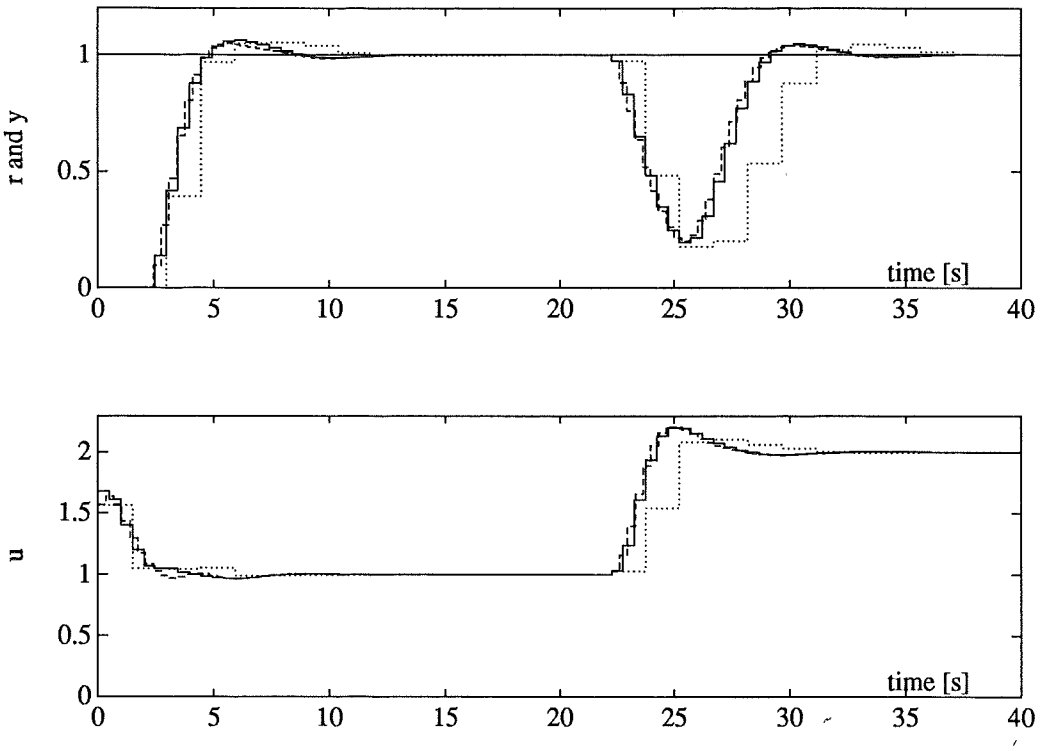


Figure 7.5 Comparison between different sampling intervals. Control of process G_A using the controllers G_{R1} (solid line), G_{R4} (dashed line), and G_{R5} (dotted line).

gives $\omega_0 h = 0.49$. The delay is $d = 7$ and the controller G_{R4} is of tenth order.

- The sampling interval is $h = T_{osc}/6 = 1.48$ s for the controller $G_{R5}(q^{-1})$. This is very slow sampling since $\omega_0 h = 2.12$. The control strategy resembles dead-beat since the closed loop poles are close to the origin. This controller is of fifth order.

Figure 7.5 shows simulations where the process G_A is controlled with the controllers G_{R1} (solid line), G_{R4} (dashed line), and G_{R5} (dotted line). In order to emphasize the effects of sampling the sampled signals are shown. The responses for the reference input are close for all three controllers. For the load disturbance, the responses for G_{R1} and G_{R4} are similar, but the response for G_{R5} is slower. The reason for this slow response is that the disturbance is not detected so fast with a long sampling interval.

Such a performance degradation can be avoided if the sampling interval is sufficiently short, as for G_{R1}. An even shorter sampling interval, as for G_{R4}, improve performance very little. This illustrates that it is sound practice to bound the number of samples per dead time to 4 or 5.

A Benchmark Problem

One of the processes in the test batch has been used as a benchmark problem for adaptive control strategies [M'Saad, 1991]. Here, it is used to show the benefit of the proposed method for chosing the structure of the process model numerator in Section 7.2.

The process is given by

$$G_B(s) = \frac{(1-s)\omega_p^2}{(1+s)(s^2+2\zeta\omega_p s+\omega_p^2)}\, e^{-s\tau}$$

where ω_p, ζ, and τ belong to certain intervals. Here, $\omega_p = 15$, $\zeta = 0.5$, and $\tau = 0.4$ are considered. The experiment with relay feedback gives $\omega_{osc} = 1.11$ radians/s. The slope of the transfer function is estimated to $\hat{n} \approx 0$. Hence, the non-minimum phase properties dominate the process at the oscillation frequency. A first order process model is chosen. Equation (7.7) gives $\omega_0 = 0.66$ radians/s. The sampling interval is $h = 0.40$ s.

The structure of the process model numerator is determined in step 8 of the initialization procedure. After normalization with the largest coefficient, the numerator of (7.9) is given by

$$B_0^*(q^{-1}) = 0.00q^{-1} - 0.62q^{-2} - 0.14q^{-3} + 1.00q^{-4} + 0.35q^{-5}$$

Initially, let $\deg B^* = \deg A^* = 1$, then (7.10) is maximized for $d = 3$. However, the coefficient b_{02} violates (7.11). This is an indication of that the suggested process model structure may not be suited for the actual process. Therefore, the structure of the process model numerator is changed to $\deg B^* = 3$ and $d = 1$, as proposed in Section 7.2. Figure 7.6 shows simulations of two adaptive controllers with different structure of the process model numerator. The solid lines correspond to the choices $\deg B^* = 3$ and $d = 1$, and the dashed lines correspond to $\deg B^* = 1$ and $d = 3$. The convergence is much faster for the more complex model, since this matches the real process much better. After many reference steps the two strategies give, however, almost the same behavior. This is observed for both the reference step at $t \approx 68$ s and the load disturbance at $t \approx 76$ s.

Different Choices of Model Order

The degree of the model is given by (7.8). It determines the complexity of the controller. When the slope $\hat{n}$ of the frequency response is close to a borderline for the model order in (7.8), a minor change in $\hat{n}$ leads to another model order. This example demonstrates what different choices of model order means for the resulting closed loop system.

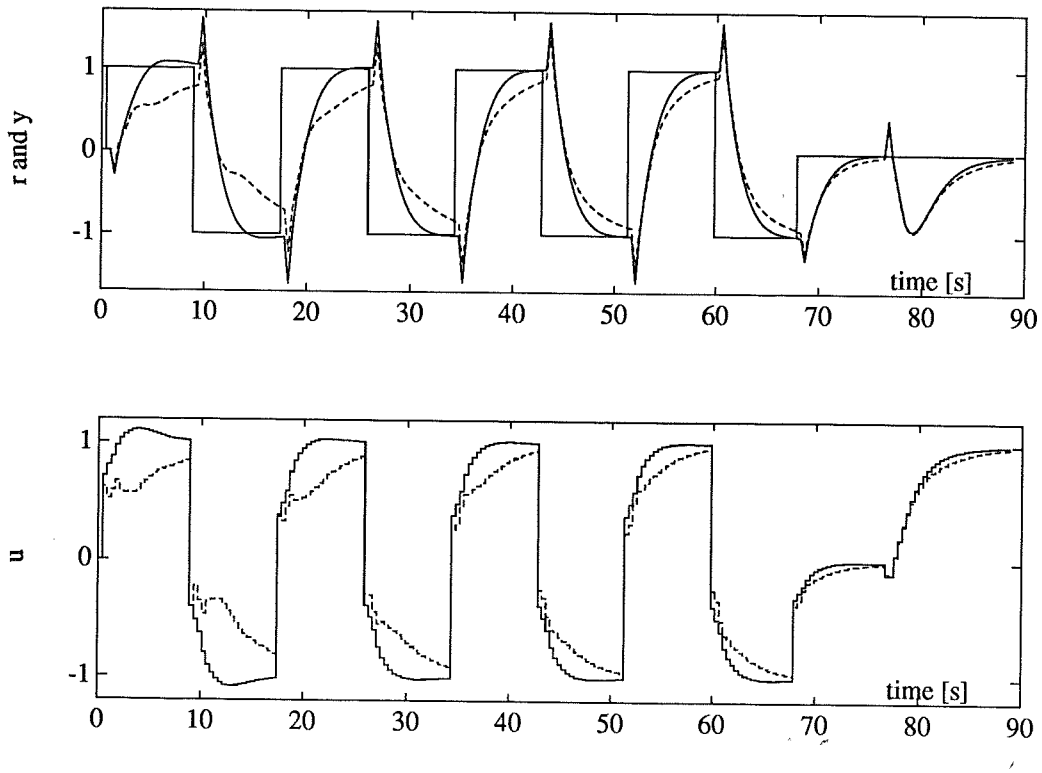


Figure 7.6 Comparison between different structures of the process model numerator. The solid lines correspond to $d = 1$ and $\deg B^* = 3$ and the dashed line to $d = 3$ and $\deg B^* = 1$.

Consider the process

$$G_C(s) = \frac{8}{(s+1)(s+2)(s+4)} e^{-0.09s}$$

The value of the time delay was judiciously chosen to obtain an estimate of the slope at the borderline. An experiment with relay feedback gives the oscillation frequency $\omega_{osc} = 2.05$ radians/s. Waveform analysis gives $t_2 - t_1 = 0.27$ s and $t_3 - t_1 = 0.89$ s. The slope of the frequency response is $\hat{n} = 2.1$ which is on the borderline between the choices $\deg A^* = 2$ and $\deg A^* = 3$. Table 7.2 shows what these choices of degree mean for some parameters in this adaptive design method. The delay in the model is in both cases $d = 1$, and $\deg B^* = \deg A^*$. The sampling interval is $h = 0.22$ s.

The process G_C is controlled by two different adaptive controllers. The controller G_{C2} is based on the second order process model. The solid curves in Figure 7.7 show the responses for this system. The other adaptive controller G_{C3} is based on the third order model. The behavior of this system is shown

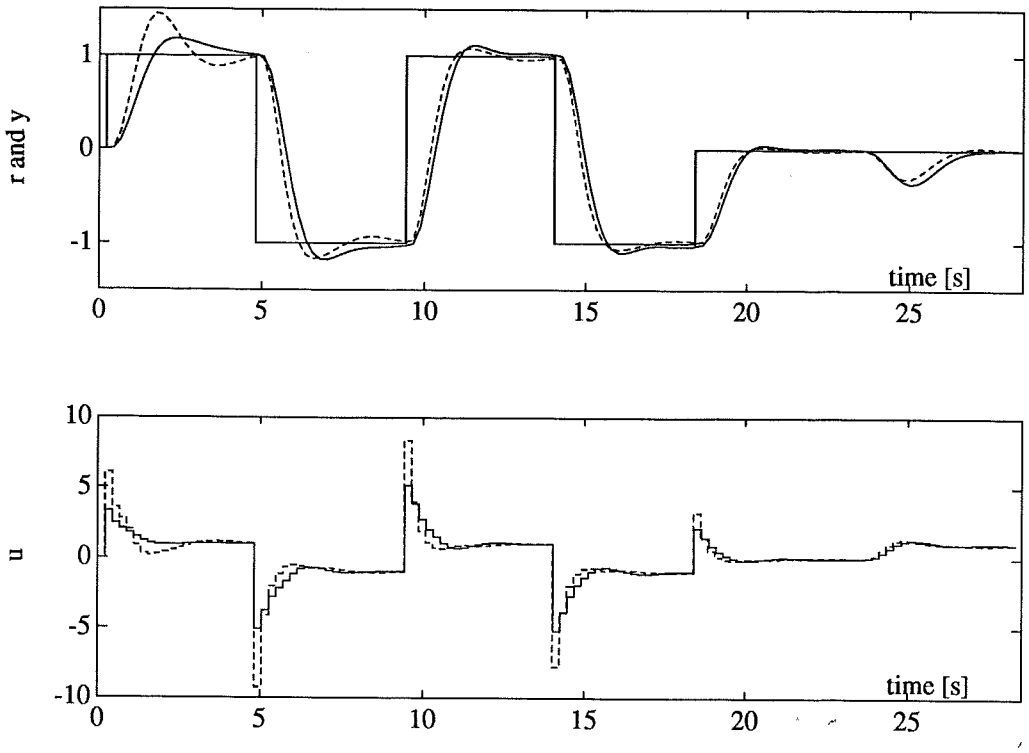


Figure 7.7 Adaptive control of process $G_C(s)$. The solid curves represent the system based on the second order model. The dashed curves represent the system for the third order model.

by the dashed curves. First, the reference input is two periods of a square wave with frequency $\omega_{osc}/3$, then it is zero for ten seconds. An input load disturbance affects the system at $t \approx 23$ s.

The system with the controller G_{C3} does not tune as fast as the one with G_{C2}. The reason is that a more complex process model requires more excitation to converge. After this initial transient the systems behave similarly as long as the reference signal excites the system. This is also verified by the Nyquist curves for the loop transfer functions at $t = 22$ s. These are close, as shown in Figure 7.8.

Table 7.2 Parameters in the adaptive design method.

deg A^*	2	3	
T	0.89	0.58	s
τ	0.24	0.01	s
ω_0	2.21	2.92	rad/s

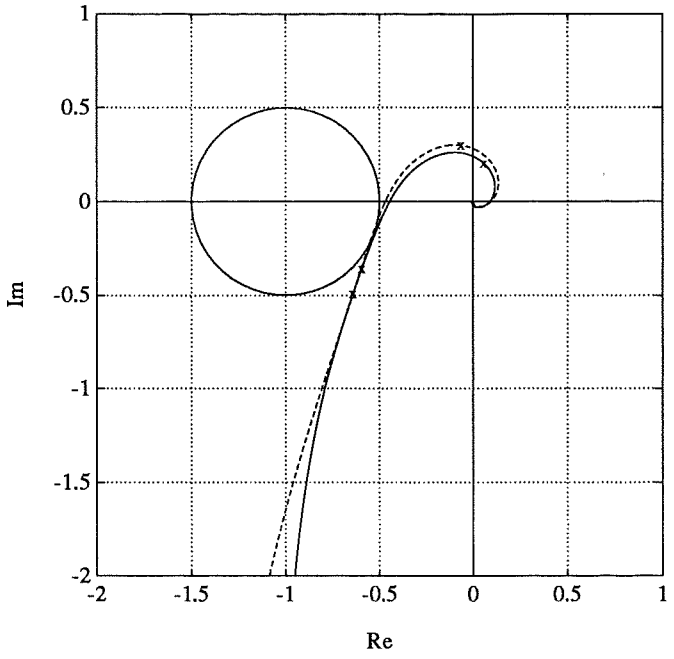


Figure 7.8 Nyquist curves for the loop transfer functions at $t = 22$ s. The solid curve represents the system based on the second order model and the dashed curve represents the third order model.

7.6 Conclusions

A procedure for automatic initialization of an adaptive controller based on pole placement design and recursive least squares identification has been presented. The procedure uses information from an experiment with relay feedback to chose sampling interval, model structure and closed loop specifications. Data from the experiment with relay feedback is also used to provide the recursive estimator with appropriate initial parameters. Properties of this initialization procedure have been illustrated by some examples. Since the procedure is based on relay feedback it will only work for systems where stable limit cycles are obtained under relay feedback. Extensive experience indicates that this occurs for a wide variety of processes that are encountered in industry, see [Hägglund and Åström, 1991].

8

Conclusions

This thesis has treated two problems in robust and adaptive control. The results are a robust design method and a procedure for automatic initialization of adaptive controllers.

A new method for robust control system design has been presented in Chapters 3 and 4. This method is based on ideas given in [Boyd and Barratt, 1991] which are generalized to processes where the uncertainty of the process model is described as a combination of structured and unstructured uncertainties. A main contribution is the formulation of design criteria as constrained convex functions. These criteria guarantee robust performance of the closed loop system for all processes within the specified region of uncertainty. This is developed in Chapter 3. The convexity of the constraints is used in two ways. First, to decrease the complexity of the design problem to a finite dimensional problem. The controller is obtained using optimization as shown in Chapter 4. The convexity implies that the optimization problem has a unique solution. This design method searches for a controller to meet several different design criteria. The design problem can also be formulated so that it gives the limit of some performance objective.

For processes with uncertainties ranging over large intervals, a linear controller that meets all design objectives may not exist. Adaptive control is an alternative in such cases. Such controllers have the potential to give excellent performance, but adaptive systems may be difficult to commission. The other main contribution of the thesis is a method for automatic initialization of an adaptive controller. This method uses information from an experiment with relay feedback. The robust design procedure developed in the first part of the thesis is too complicated to be performed on line in an adaptive con-

troller. It also requires manual interaction. A strongly simplified version that can be used on line is developed in Chapter 5. This procedure is used in the adaptive controller. This controller may give poor performance if the closed loop poles are not chosen properly. In Chapter 7 a simple formula is proposed for choosing the closed loop poles. This choice is intended to give a closed loop system with a sensitivity around 2.

The procedure for automatic initialization of an adaptive controller is, from an operational point of view, similar to the PID auto-tuners that are available today. The main difference is that the method in this thesis works well for a larger class of processes than the PID auto-tuner, e.g. processes with considerable delays can now be automatically tuned.

Natural future extensions are initialization procedures for other types of adaptive controllers and modifications of the experiment to cover also other types of processes.

9

References

ÅSTRÖM, K. J. (1980): "Self-tuning control of a fixed-bed chemical reactor system." *International Journal of Control*, **32:2**, pp. 221–256.

ÅSTRÖM, K. J. (1990): "Regulation of a ships heading." In DAVISON, Ed., *Benchmark Problems for Control System Design*, pp. 41–42. IFAC Theory Committee, Laxenburg, Austria.

ÅSTRÖM, K. J., B. BERNHARDSSON, and A. RINGDAHL (1991): "Solution using robust adaptive pole placement." In COMMAULT *et al.*, Eds., *Proceedings of the First European Control Conference, ECC 91,* Grenoble, France, volume 3, pp. 2340–2345, Paris. Hermes.

ÅSTRÖM, K. J. and T. HÄGGLUND (1988a): *Automatic Tunings of PID Controllers.* Instrument Society of America, Research Triangle Park, North Carolina.

ÅSTRÖM, K. J. and T. HÄGGLUND (1988b): "A new auto-tuning design." In *Preprints IFAC Int. Symposium on Adaptive Control of Chemical Processes, ADCHEM '88*, Lyngby, Denmark.

ÅSTRÖM, K. J., C. C. HANG, and P. PERSSON (1989): "Towards intelligent PID control." In *Preprints IFAC Workshop on Artificial Intelligence in Real Time Control, Shenyang, P. R. China*, pp. 38–43, Oxford, UK. Pergamon Press.

ÅSTRÖM, K. J. and B. WITTENMARK (1989): *Adaptive Control.* Addison-Wesley, Reading, Massachusetts.

ÅSTRÖM, K. J. and B. WITTENMARK (1990): *Computer Controlled Systems—Theory and Design.* Prentice-Hall, Englewood Cliffs, New Jersey, second edition.

ATHERTON, D. P. (1975): *Nonlinear Control Engineering—Describing Function Analysis and Design.* Van Nostrand Reinhold Co, London, UK.

BENGTSSON, G. (1989): "First microcontroller – A self-adaptive process control system." Technical report, First Control Systems AB, Västerås, Sweden.

BOYD, S. P., V. BALAKRISHNAN, C. H. BARRATT, N. M. KHRAISHI, X. LI, D. G. MEYER, and S. A. NORMAN (1988): "A new CAD method and associated architectures for linear controllers." *IEEE Transactions on Automatic Control*, **33**, pp. 268–282.

BOYD, S. P. and C. H. BARRATT (1991): *Linear Controller Design – Limits of Performance.* Prentice Hall Inc., Englewood Cliffs, New Jersey.

CHESTNUT, H. and R. W. MAYER (1959): *Servomechanisms and Regulating System Design.* Wiley, New York.

DOYLE, J. C. (1987): "A review of μ—For case studies in robust control." In *Preprints 10th IFAC World Congress*, pp. 395–402, Munich, Germany.

DOYLE, J. C. and G. STEIN (1981): "Multivariable feedback design: Concepts for a classical/modern synthesis." *IEEE Transactions on Automatic Control*, **AC-26**, pp. 4–16.

FRANCIS, B. A. (1988): "Snippets of H_∞ control theory." Technical Report 8802, Systems Control Group, Department of Electrical Engineering, University of Toronto, Toronto, Canada.

GILL, P. E., W. MURRAY, M. A. SAUNDERS, and M. H. WRIGHT (1986): *User's Guide for NPSOL.* Department of Operations Research, Stanford University, Stanford, California. Technical Report SOL 86-2.

GOODWIN, G. C. and K. S. SIN (1984): *Adaptive Filtering, Prediction and Control.* Prentice-Hall, Englewood Cliffs, New Jersey.

HÄGGLUND, T. and K. J. ÅSTRÖM (1991): "Industrial adaptive controllers based on frequency response techniques." *Automatica*, **27**, pp. 599–609.

HANG, C. C. and K. J. ÅSTRÖM (1988): "Practical aspects of PID autotuners based on relay feedback." In *Preprints IFAC Int. Symposium on Adaptive Control of Chemical Processes, ADCHEM '88*, Lyngby, Denmark.

HOLMBERG, U. (1991): *Relay Feedback of Simple Systems*, PhD thesis TFRT-1034. Dept. of Automatic Control, Lund Inst. of Technology, Lund, Sweden.

HOROWITZ, I. M. (1963): *Synthesis of Feedback Systems.* Academic Press, New York.

ISERMANN, R. (1980): "Practical aspects of process identification." *Automatica*, **16**, pp. 575–587.

KRAUS, T. W. and T. J. MYRON (1984): "Self-tuning PID controller uses pattern recognition approach." *Control Engineering*, June, pp. 106–111.

LILJA, M. (1989): *Controller Design by Frequency Domain Approximation*, PhD thesis TFRT-1031. Dept. of Automatic Control, Lund Inst. of Technology, Lund, Sweden.

LJUNG, L. (1987): *System Identification—Theory for the User*. Prentice Hall, Englewood Cliffs, New Jersey.

LUNDH, M. (1990): "Optimization based robust design of uncertain SISO systems." In *Preprints 11th IFAC World Congress*, volume 5, pp. 260–265, Tallinn, Estonia.

MASTEN, M. K. and H. E. COHEN (1989): "An advanced showcase of adaptive controller designs." In *Proceedings American Control Conference*, Pittsburg, Pennsylvania.

MATHWORKS (1990): *Pro-Matlab User's Guide*. The MathWorks Inc, South Natick, Massachusetts.

MIDDLETON, R. H. (1991): "Trade-offs in linear control system design." *Automatica*, **27**, pp. 281–292.

MORARI, M. and J. C. DOYLE (1986): "A unifying framework for control system design under uncertainty and its implications for chemical process control." In MORARI and MCAVOY, Eds., *Chemical Process Control—CPCIII*, Proceedings 3rd International Conference on Chemical Process Control, pp. 5–51. Elsevier, New York.

M'SAAD, M. (1991): "A showcase of adaptive control designs." In COMMAULT *et al.*, Eds., *Proceedings of the First European Control Conference, ECC 91*, Grenoble, France, volume 3, pp. 2374–2375, Paris. Hermes.

PERSSON, P. (1992): *PID-Controller Design*, PhD thesis to appear. Dept. of Automatic Control, Lund Inst. of Technology, Lund, Sweden.

POLAK, E., D. Q. MAYNE, and D. M. STIMLER (1984): "Control systems design via semi-infinite optimization – A review." *IEEE Proceedings*, **72**, pp. 1777–1794.

ROHRS, C., L. S. VALAVANI, M. ATHANS, and G. STEIN (1985): "Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics." *IEEE Transactions on Automatic Control*, **AC-30**, pp. 881–889.

SIMMONS, G. F. (1963): *Introduction to Topology and Modern Analysis*. McGraw-Hill, Singapore.

SOH, Y. C., R. J. EVANS, I. R. PETERSEN, and R. E. BETZ (1987): "Robust pole assignment." *Automatica*, **23**, pp. 601–610.

STEIN, G. and M. ATHANS (1987): "The LQG/LTR procedure for multivariable feedback control design." *IEEE Transactions on Automatic Control*, **AC-32**, pp. 105–114.

TSYPKIN, Y. Z. (1984): *Relay Control Systems.* Cambridge University Press, Cambridge, UK.

VIDYASAGAR, M. (1985): *Control System Synthesis: A Factorization Approach.* MIT Press, Cambridge, Massachusetts.

WALLENBORG, A. (1987): "Control of flexible servo systems." Lic Tech thesis TFRT-3188, Dept. of Automatic Control, Lund Inst. of Technology, Lund, Sweden.

WEI, K. and R. K. YEDAVALLI (1989): "Robust stabilizability for linear systems with both parameter variation and unstructured uncertainty." *IEEE Transactions on Automatic Control*, **34**, pp. 149–156.

WITTENMARK, B. (1989): "Adaptive control: Implementation and application issues." In SHAH and DUMONT, Eds., *Adaptive Control Strategies for Industrial Use*, Proceedings of a Workshop, Kananaskis, Canada, 1988, volume 137 of *Lecture Notes in Control and Information Sciences*, pp. 103–120. Springer-Verlag.

ZAKIAN, V. and L. AL-NAIB (1973): "Design of dynamical and control systems by the method of inequalities." *Proc. Inst. Elec. Eng.*, **120:11**.

ZAMES, G. (1981): "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses." *IEEE Transactions on Automatic Control*, **AC-26**, pp. 301–320.

ZIEGLER, J. G. and N. B. NICHOLS (1942): "Optimum settings for automatic controllers." *Trans. ASME*, **64**, pp. 759–768.

A

The Batch of Processes

The procedure for initialization of adaptive controllers in Chapter 7 involves choices of many variables. These choices were based on heuristics rules. To determine the validity of the initialization procedure, it was tested on a batch of processes. These processes are listed here. The outcome of the initialization procedure for the processes were measured by the peaks of the sensitivity functions.

Processes

Processes with the following transfer functions were considered.

$$G_1(s) = \frac{1}{(s+1)^2} e^{-s\tau} \qquad 0.01 \leq \tau \leq 10$$

$$G_2(s) = \frac{1}{(s+1)^n} \qquad 3 \leq n \leq 20$$

$$G_3(s) = \frac{1-\alpha s}{(s+1)^3} \qquad 0.1 \leq \alpha \leq 2$$

$$G_4(s) = \frac{1}{s} e^{-s}$$

$$G_5(s) = \frac{1}{s(s+1)} e^{-s\tau} \qquad 0.05 \leq \tau \leq 2$$

$$G_6(s) = \frac{1-\alpha s}{s(s+1)} \qquad 0.05 \leq \alpha \leq 0.75$$

$$G_7(s) = \frac{1}{s+1} e^{-s\tau} \qquad 0.01 \leq \tau \leq 10$$

$$G_8(s) = \frac{\omega_p^2}{s(s^2 + 2\zeta\omega_p s + \omega_p^2)} e^{-s} \qquad \zeta = 0.2, \omega_p = 4.3$$

$$G_9(s) = \left(\frac{1}{(s+1)^2} + \frac{0.5}{s+0.5}\right) e^{-s\tau} \qquad 0.01 \leq \tau \leq 5$$

$$G_{10}(s) = \frac{25}{(s+1)(s+3)(s^2+2s+25)} e^{-3s}$$

$$G_{11}(s) = \frac{(1-s)\omega_p^2}{(1+s)(s^2 + 2\zeta\omega_p s + \omega_p^2)} e^{-s\tau} \qquad 10 \leq \omega_p \leq 15, 0.2 \leq \tau \leq 0.4$$

$$G_{12}(s) = \frac{1}{(s+1)^3} e^{-s\tau} \qquad 0.01 \leq \tau \leq 10$$

$$G_{13}(s) = \frac{1}{(s+1)(s+p)(s+p^2)(s+p^3)} \qquad 2 \leq p \leq 10$$

The processes $G_1(s)$, $G_7(s)$, and $G_{12}(s)$ were used to heuristically determine the rule (7.7) for choosing the parameter ω_0. Processes $G_1(s)$ to $G_6(s)$ have been used to assess performance of PID-controllers [Åström *et al.*, 1989]. The process $G_{11}(s)$ is a benchmark problem for adaptive control [M'Saad, 1991]. The process $G_8(s)$ is constructed to have a resonance at $3\omega_{osc}$, which leads to an erroneous estimate of the slope. The processes $G_9(s)$ and $G_{10}(s)$ are inspired by processes where PID auto-tuning has failed. The process G_{13} is a modification of $G_2(s)$.

Validation

The initialization procedure for adaptive controllers in Section 7.4 have been tested on the processes above. First, the initialization procedure was applied to the process, then the adaptive controller was commissioned. Under adaptive control, the reference input was first two periods of a square wave with the frequency $\omega_{osc}/3$ then it was zero for 30 samples. The frequency response of the sensitivity function $\mathcal{S}_o$ was calculated using the controller that was obtained at the time when the adaptation stopped. The design objectives in Section 7.2 were that the sensitivity function should satisfy $|\mathcal{S}_o| \leq 2$ and that the controller should be stable. The controllers were stable for all the investigated processes. The peaks of the sensitivity functions are shown in Figure A.1. Processes modeled with $\deg A^* = 1$ are marked with a circle ($\circ$). The plus marks ($+$) correspond to $\deg A^* = 2$ and the stars ($*$) correspond to $\deg A^* = 3$. Some of the sensitivity functions have peaks higher than 2. The highest peak is obtained for the process $G_9(s)$ with $\tau = 1$. This process would preferably be modeled by a first order model.

Although the formula (7.7) for choosing the design parameter ω_0 was derived using simple processes, it works rather well for all these processes.

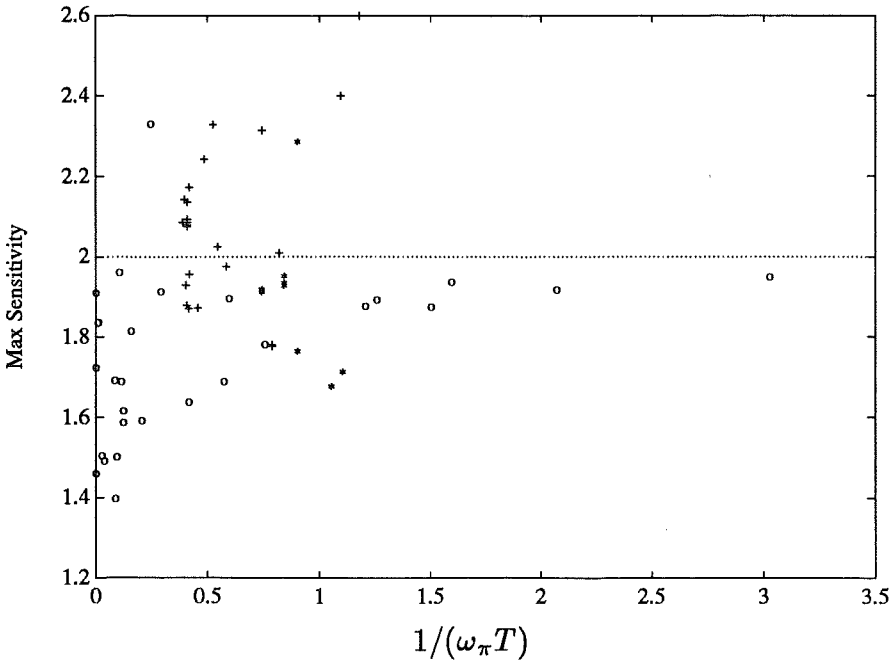


Figure A.1 Peaks of sensitivity functions. The circles (o), the plus marks (+), and the stars (*) represent processes modeled with $\deg A^* = 1$, $\deg A^* = 2$, and $\deg A^* = 3$ respectively.

Table A.1 shows the result of the validation procedure for the process $G_k(s)$. The second column shows the value of the parameter that varies for the different processes. An exception is the process G_{11} where $p_k = a$ means $\omega_p = 15, \tau = 0.2$, $p_k = b$ means $\omega_p = 15, \tau = 0.4$, $p_k = c$ means $\omega_p = 10, \tau = 0.2$, and $p_k = d$ means $\omega_p = 10, \tau = 0.4$. Other columns in the table show the estimate of the normalized time constant $\omega_{osc}T$, the peak of $|S_o|$, the oscillation frequency ω_{osc}, the design parameter ω_0, and the estimated slope $\hat{n}$. The chosen model structure is defined by $\deg A^*$, $\deg B^*$, and d. These are also given. Last column shows the estimated delay τ in (6.15).

Table A.1 Valuation of processes in test batch.

k	p_k	$\omega_{osc}T$	$\lvert\mathcal{S}_o\rvert$	ω_{osc}	ω_0	$\hat{n}$	$\deg A^*$	$\deg B^*$	d	τ
1	0.01	2.40	1.87	2.66	2.50	1.79	2	2	1	0.06
1	0.03	2.45	1.88	2.53	2.34	1.80	2	2	1	0.06
1	0.10	2.47	1.93	2.31	2.13	1.79	2	2	1	0.11
1	0.32	1.83	2.02	1.65	1.78	1.77	2	2	2	0.30
1	1.00	0.91	2.40	0.99	1.73	1.55	2	2	4	1.15
1	3.20	1.33	1.78	0.52	0.69	0.93	1	1	5	3.39
1	10.00	0.48	1.92	0.24	0.74	0.31	1	1	4	10.34
2	3	1.35	1.92	1.16	1.51	2.24	3	3	1	0.00
2	4	1.11	1.76	0.74	1.12	2.52	3	3	1	0.15
2	6	0.95	1.68	0.45	0.76	2.39	3	3	2	0.73
2	8	0.91	1.71	0.34	0.60	2.39	3	3	3	1.43
2	10	1.27	1.78	0.27	0.37	2.04	2	2	4	3.30
2	15	1.27	1.78	0.18	0.24	1.49	2	2	5	5.02
2	20	2.42	1.64	0.13	0.12	1.16	1	1	5	11.45
3	0.10	1.19	1.95	1.07	1.53	2.21	3	3	1	0.02
3	0.25	1.71	1.97	0.98	1.10	2.03	2	2	1	0.56
3	0.50	1.34	2.31	0.87	1.14	1.72	2	2	1	0.87
3	1.00	12.13	1.69	0.72	0.42	1.22	1	2	1	1.47
3	1.50	8.89	1.69	0.62	0.38	0.91	1	2	1	1.64
3	2.00	37.62	1.50	0.56	0.29	0.69	1	2	1	1.57
4	–	83.64	1.84	0.94	0.48	1.02	1	1	2	1.08
5	0.05	2.44	2.07	1.43	1.33	1.70	2	2	1	0.22
5	0.10	2.44	2.08	1.27	1.18	1.70	2	2	1	0.24
5	0.20	2.56	2.08	1.06	0.96	1.67	2	2	1	0.26
5	0.40	2.44	2.14	0.83	0.77	1.61	2	2	1	0.36
5	0.60	2.39	2.17	0.69	0.65	1.56	2	2	1	0.45
5	0.80	2.06	2.24	0.56	0.56	1.46	2	2	2	0.63
5	1.00	1.90	2.33	0.49	0.52	1.42	2	2	2	0.84
5	2.00	> 100	1.84	0.32	0.16	1.26	1	1	2	3.20
6	0.05	2.44	2.08	1.43	1.33	1.68	2	2	1	0.22
6	0.10	2.44	2.09	1.27	1.18	1.65	2	2	1	0.24
6	0.25	2.51	2.14	0.93	0.85	1.54	2	2	1	0.24
6	0.50	9.54	1.96	0.71	0.43	1.27	1	1	2	1.59
6	0.75	> 100	1.91	0.60	0.30	1.11	1	1	2	1.52

Table A.1 Cont'd

k	p_k	$\omega_{osc}T$	$\lvert S_o\rvert$	ω_{osc}	ω_0	$\hat{n}$	deg A^*	deg B^*	d	τ
7	0.01	11.55	1.40	14.30	8.40	1.00	1	1	1	0.01
7	0.03	8.34	1.59	12.06	7.50	0.99	1	1	1	0.04
7	0.10	6.36	1.82	8.16	5.39	0.99	1	1	2	0.11
7	0.32	3.43	1.91	3.77	3.02	1.02	1	1	4	0.34
7	1.00	1.68	1.90	1.56	1.77	0.90	1	1	4	1.02
7	3.20	0.63	1.94	0.67	1.61	0.57	1	1	4	3.34
7	10.00	0.33	1.95	0.26	1.05	0.17	1	1	4	10.02
8	–	4.09	2.33	0.86	0.65	0.70	1	1	4	1.45
9	0.01	>100	1.46	7.70	3.85	1.06	1	1	1	0.02
9	0.10	>100	1.72	5.00	2.50	1.10	1	1	1	0.14
9	1.00	0.85	2.60	1.13	2.09	1.39	2	2	3	1.08
9	5.00	0.83	1.88	0.40	0.76	0.70	1	1	4	5.23
10	–	0.80	1.89	0.61	1.19	0.55	1	1	4	3.31
11	*a*	27.21	1.49	1.20	0.64	−0.02	1	1	1	0.49
11	*b*	10.68	1.50	1.11	0.66	−0.01	1	3	1	0.69
11	*c*	8.24	1.62	1.18	0.73	−0.04	1	1	1	0.58
11	*d*	4.92	1.59	1.08	0.76	−0.03	1	1	2	0.80
12	0.00	1.35	1.92	1.16	1.51	2.24	3	3	1	0.00
12	0.01	1.35	1.92	1.15	1.50	2.24	3	3	1	0.00
12	0.03	1.35	1.91	1.13	1.48	2.24	3	3	1	0.00
12	0.10	1.19	1.93	1.07	1.53	2.25	3	3	1	0.02
12	0.32	1.11	2.29	0.94	1.41	2.18	3	3	2	0.12
12	1.00	1.22	2.01	0.71	0.99	1.91	2	2	3	1.19
12	3.20	1.74	1.69	0.44	0.49	1.15	1	2	4	3.76
12	10.00	0.67	1.88	0.22	0.50	0.41	1	1	4	10.58
13	2.00	1.19	1.93	1.96	2.80	2.23	3	3	1	0.01
13	4.00	2.38	1.95	4.37	4.11	1.86	2	2	1	0.05
13	10.00	2.19	1.87	7.82	7.67	1.61	2	2	1	0.00