Robust Control Design with Simplified Models

Mannerfelt, Carl Fredrik

1981

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

Link to publication

Citation for published version (APA):
Mannerfelt, C. F. (1981). *Robust Control Design with Simplified Models*. [Doctoral Thesis (monograph), Department of Automatic Control]. Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

CODEN: LUTFD2/(TFRT-1021)/1-153/(1981)

Robust Control Design with Simplified Models

Carl Fredrik Mannerfelt

Department of Automatic Control | Lund Institute of Technology

Carl Fredrik Mannerfelt

Robust Control Design with Simplified Models

80 ⅙

Robust Control Design with Simplified Models

Carl Fredrik Mannerfelt

Lund 1981

(1)

ROBUST CONTROL DESIGN

WITH SIMPLIFIED MODELS

Carl Fredrik Mannerfelt

LUND 1981

LUND INSTITUTE OF TECHNOLOGY
DEPARTMENT OF AUTOMATIC CONTROL
Box 725
S 220 07 Lund 7 Sweden

Document name: DOCTORAL DISSERTATION

Date of issue: August 1981

Document number: CODEN: LUTFD2/(TFRT-1021)/1-153/(1981)

Author(s): Carl Fredrik Mannerfelt

Supervisor: Karl Johan Åström

Sponsoring organization:

Title and subtitle

Robust Control Design with Simplified Models.

Abstract

Simplified system models are in widespread use in Automatic Control. Controllers are usually designed to stabilize and control a model of a real system. Stability does however not necessarily follow when the controller is used on the real process. This problem is refered to as the robustness problem. Several new robustness results are given in this thesis. The results are given in terms of the frequency responses of the system and the simplified model. The parameters of linear system models can be estimated by system identification methods. We examine the properties of low order estimated models. The real process is implicitly assumed to be more complex than the simplified model. Finally, an algorithm for a simple self-tuning controller is proposed. Because of the simplicity of the controller, it is possible to derive L^{∞}-stability results when the regulated process is more complex then the simple model.

Key words:

Classification system and/or index terms (if any):

Supplementary bibliographical information:

ISSN and key title:

ISBN:

Language: English

Number of pages: 153

Recipient's notes:

Security classification:

Distribution: The report may be ordered from the Department of Automatic Control or borrowed through the University Library 2, Box 1010, S-221 03 Lund, Sweden, Telex: 33248 lubbis lund.

To Anna
who never saw me doing it

CONTENTS

ROBUST CONTROL DESIGN WITH SIMPLIFIED MODELS

In many branches of modern science, the use of mathematical models to describe dynamical systems, has been found an indispensable tool. The models are needed for a multitude of purposes. In the analysis and simulation of dynamical processes, it is always a model of a process that is examined. When designing controllers for dynamical systems, the controllers are usually designed to control a model of the system. Another important application is in the diagnosis of certain states of operation of a process. This particular application will probably have practical consequences in future medical treatment. In socio-economics as well as in telecommunications, there are many important applications of process models in the prediction of future developements of system variables.

When deriving a system model from physical principles, the result is a process description often containing non-linear, infinite dimensional differential equations. Such models give good global descriptions, but are very difficult, if not impossible to work with. If it is assumed that the system is operating with small deviations around some fix level of operation, it is in general acceptable to regard the system model as being linear locally. Further simplification of the linearized model is to approximate it with a linear model of finite (low) order. This is the type of process model that is used in most practical applications. The theory for analysis and control of such models is well developed.

In this thesis we are dealing with some practical consequences of the model simplification. The thesis is organized into three parts:

PART I - Robustness and Stability with Uncertain Models.

PART II - Estimation of Simple Parametric Models for High Order Systems.

PART III - A Simple Self-Tuning Controller.

In the first part, the problem of robustness in control systems is treated. The controller in a control system is usually designed to control a simple linear model of the real process. It is then not evident that the real process shows good behaviour together with the controller. Discrepancies between the process and the simple model may be such that the real control system becomes unstable. It is thus important to obtain design principles that take into account the uncertainties and errors in the model, giving a robust control system. In Part I we give conditions for stability of the real control system. From the results it is also possible to draw conclusions about means how to achieve robust design methods.

During the completion of the thesis, there has been a publication of new robustness results, Doyle and Stein (1981). These results are very similar to one of the present results (Theorem 4.3 of Part I).

A common way to obtain the values of parameters in finite dimensional system models, is by System Identification methods. Much theoretical work has been carried out under the assumption that the real process has the same structure as the system model. In Part II, we examine the properties of estimated low order models, when it is implicitly assumed that the real system is more complex than the model.

Part III is concerned with a simple self-tuning regulator. The controller is intended to control stable processes that can be accurately described by a pure time delay and a constant gain. The controller is shown to have appealing properties when the model error is not too large. Because of the simple structure of the self-tuning controller, it is possible to derive stability results when the controlled process is more complex than the simple model.

ACKNOWLEDGEMENTS

I want to thank my supervisor Professor Karl Johan Åström for his excellent guidance and for many enlightening discussions. His enthusiasm for control theory and related subjects, by which he influences all coworkers at the Department of Automatic Control, has been a power-source needed to complete this work.

I also want to thank my colleagues at the department for many and helpful discussions. In particular Per Hagander should be mentioned, who read the manuscript and pointed out ways to improve some of the results. The figures and drawings were masterfully arranged by Britt-Marie Carlsson, Eslöv.

REFERENCE

Doyle, J.C. and Stein, G.(1981): Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis. IEEE Trans. Aut. Control 26, 4-16.

PART I - ROBUSTNESS AND STABILITY WITH UNCERTAIN MODELS

ABSTRACT

In the control of linear systems, feedback regulators are usually designed for simple process models. It is therefore important to obtain design principles that take into account model uncertainties and modelling errors, in order to get robust control systems.

New robustness results are presented, that can be applied to both stable and unstable systems. The results are formulated in terms of the difference between the frequency responses of the loop transfer functions of the real system and the model. These robustness results have implications on the practical design of feedback controllers.

1. INTRODUCTION

In the control of linear, time-invariant dynamical systems, feedback has been used for a long time. Reasons for introducing feedback in a control system are e.g.

* to stabilize the system if it has insufficient stability properties.
* to reduce the effects of noise disturbances acting on the system.
* to achieve a specified input-output performance.
* to reduce the effects of variations in the plant, such as parametric variations and neglected or unmodelled high-order process dynamics.

These items can all be covered in "reducing the effects of disturbances", where now "disturbances" has a wide meaning - it includes

* unwanted effects of initial conditions, such as unstable, poorly damped or slow transients.
* external noise disturbances entering the process in an additive fashion.
* variations in the dynamics of the process.

Black (1934) found that high feedback gain in amplifiers reduced the sensitivity to parameter variations. Bode (1945) generalized this result to single-input - single-output (SISO) dynamical systems. He found that if the magnitude of the loop transfer function was much larger than unity, the closed-loop system was insensitive to open-loop system variations. This could also be expressed in terms of the magnitude of the return difference, i.e. one minus the loop transfer function.

Practical considerations of the magnitude, or gain, of feedback controllers were elaborated by Horowitz (1963). He found that process and measurement noise, as well as uncertainties in the plant (model), restrict the frequency range over which the feedback gain can be kept high.

In connection with the development of the state-space theory for dynamical systems, several efficient design methods for SISO and multivariable (MIMO) systems have been presented. These modern design methods are based on computations in "closed" form on parametric system models, and are hence well suited for computer-aided design.

System models in state-space (internal) representation are necessarily of finite dimension, and must therefore always be regarded as approximations in the real, infinite-dimensional world. This is of course also the case for finite-dimensional models on input-output (external) representation, i.e. rational transfer function models. The process models used in practical design of control systems, are almost always simple and of rather low order. In many situations, the models can be regarded as "uncertain", e.g. it may be difficult to obtain a good model of the process under certain operating conditions. It is therefore important and interesting to know, when and why (why not) feedback controllers obtained from modern and classical design methods work well with real plants.

In the following, the term robustness will be used quite frequently. A control system is said to be <u>robust</u>, if variations in the open-loop process dynamics do not destroy the stability of the closed-loop system. From a practical point of view, it is not only closed-loop stability that is important. There are usually constraints on the behaviour of the control system that have to be satisfied. The preservation of such properties under process variations, will be called <u>robustness of performance</u>.

Since real processes in general can be considered as infinite-dimensional, it is natural to work with input-output descriptions of systems when analyzing robustness properties. The relevant system models are transfer functions and impulse responses, or even better - the parameter-free representation of the frequency response. This has some immediate consequences:

* only the controllable and observable part of a system is considered.

* continuous and discrete time systems can be treated with the same formalism.

The difference between the frequency responses (transfer functions) of the real process G and the model G_m, will be called the <u>modelling error</u>. It is often assumed that the modelling error is additive. This leads to the "absolute" error

$$\Delta G_a = G - G_m \tag{1.1}$$

This structure is natural when considering the frequency response of the real process in a Nyquist diagram. The distance between corresponding points on the Nyquist curves of the real process and the model, is then equal to the absolute value of the error $|\Delta G_a|$.

We can also consider the relative error between the transfer functions of the real system and the model

$$\Delta G_r = [G - G_m]/G_m \tag{1.2}$$

The absolute error can then be written in a multiplicative form

$$\Delta G_a = G_m \cdot \Delta G_r \tag{1.3}$$

This structure has a nice consequence when plotting the absolute error in a Bode diagram. The logarithm of the error becomes additive, according to

$$\log|\Delta G_a| = \log|G_m| + \log|\Delta G_r| \tag{1.4}$$

We illustrate these error concepts with an example:

EXAMPLE 1.1

Consider a stable process with a small time-delay, having the transfer function

$$G(s) = \frac{2}{s+1} e^{-0.1s} \tag{1.5}$$

A simplified, first order model of the system has the transfer function

$$G_m(s) = \frac{2}{s+1} \tag{1.6}$$

The Nyquist plots of the two transfer functions are shown in

Fig. 1.1. It is seen that the model is a good approximation of the process at low frequencies.

A plot of the absolute error $|\Delta G_a(i\omega)| = |G(i\omega) - G_m(i\omega)|$ as a function of the frequency ω, is shown in Fig. 1.2. As a comparison, the frequency response of the relative error $|\Delta G_r| = |[G - G_m]/G_m|$ is shown in a Bode plot in Fig. 1.3. Although the absolute error is small and decreasing for high frequencies, the relative error is large. This is usually the case for real systems and models. □

The robustness problem has been analyzed in several

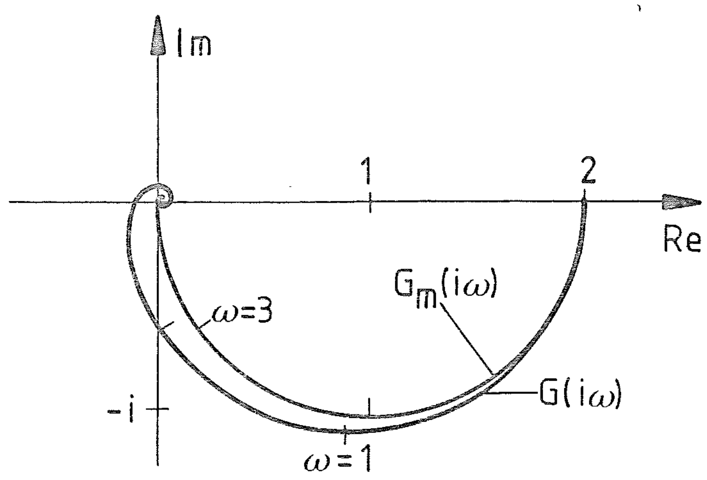


Fig. 1.1 - The Nyquist curves for the system and the model.

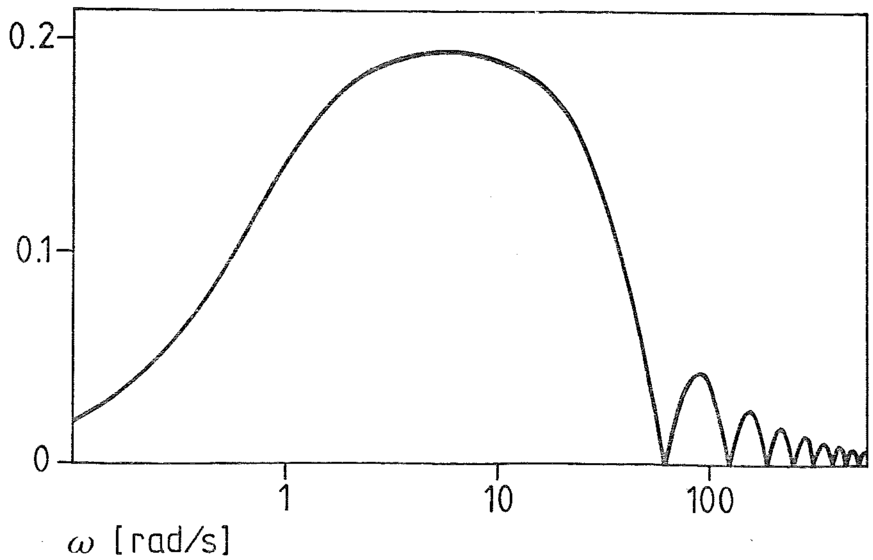


Fig. 1.2 - A plot of the absolute error $|\Delta G_a(i\omega)|$.

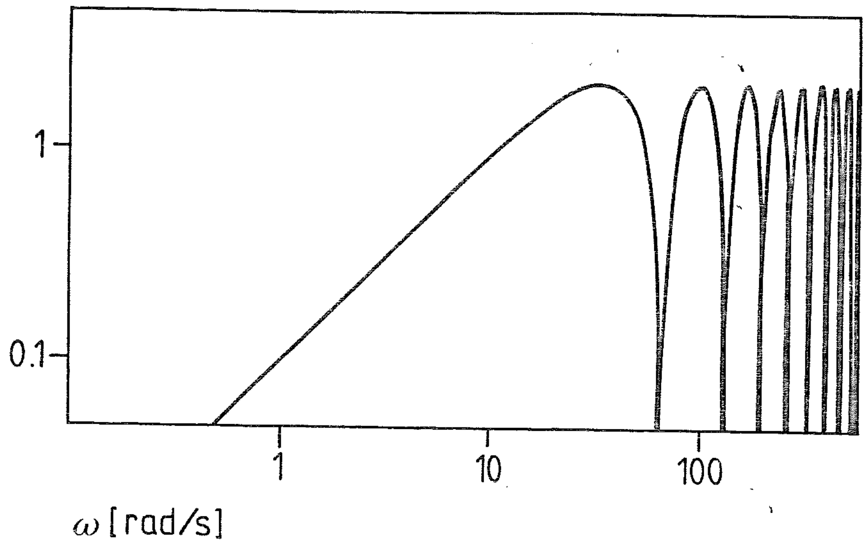


Fig. 1.3 - A Bode plot of the relative error $|\Delta G_r(i\omega)|$.

different contexts.

Kalman (1964) showed the celebrated robustness property of control systems designed with the tools of linear-quadratic optimal control. The closed-loop system has infinite amplitude margin, and a guaranteed phase margin of at least 60 degrees.

Further results in the area of robustness have been published by Safonov (1977), Doyle (1978) and Sandell Jr. (1979). These results state that if the Nyquist curves of the loop transfer functions of the real system and model are close, then the closed-loop stability of the model system implies closed-loop stability for the real control system. Both the real plant and the model must be stable, or have exactly the same set of unstable poles, in order to apply the results. Since exact modelling of the unstable open-loop poles of a process is highly unrealistic, the results can only be used for stable systems in practice.

These results are reviewed in Chapter 3 of the present work, in the case of SISO systems.

We now come to the contributions of Part I of the present thesis. The robustness results of Doyle (1978) are extended to be valid for open-loop unstable systems as well. It is shown that an important consideration when modelling a system, is the number of unstable poles. The exact locations of the unstable poles are not crucial. There is no further assumption on the dimensionality of the systems, as was the case in Doyle (1978). Both SISO and MIMO versions of the robustness result are given. Continuous and discrete time systems are treated within the same formalism.

A new type of robustness result is also presented. It involves the "unstable" zeros of the systems and the transfer functions of the inverse systems. In order to apply this result, it is no longer necessary that the open-loop transfer functions of the real process and the model have the same number of unstable poles.

The robustness results have some immediate consequences in the design of feedback controllers. Here it is theoretically motivated that the gain of the loop transfer function should be high for low frequencies where usually the modelling error is small. For high frequencies, where it is difficult to obtain good models, the controller gain should be small to ensure robustness. At the bandwidth frequency of the closed-loop system, where the magnitude of the loop transfer function is around unity, the robustness is poor. Therefore it is important to use a good model in this frequency interval.

This part of the thesis is organized as follows. In Chapter 2, we formulate assumptions on the systems to be considered. Important and well-known stability criteria in suitable forms, are also summarized. As was previously mentioned, the robustness results of Doyle (1978) are reviewed in Chapter 3. The proofs are given in a fashion that indicates a way how to derive more general robustness results. Chapter 4 contains the main contributions of this part of the thesis, robustness results for general perturbations between the real system and the model. In Chapter 5 some different approaches to analysis of robustness of the performance of control systems are considered. The famous formula for closed-loop sensitivity by Bode, is generalized to arbitrary variations in the open-loop system transfer function. To illustrate the results of the previous chapters, some examples are given in Chapter 6. Chapter 7 contains conclusions, and the references are listed in Chapter 8.

2. PRELIMINARIES

In this chapter, some basic assumptions are made on the systems to be analyzed. A standard feedback configuration, well suited for stability analysis, is motivated. Finally, two wellknown stability criteria are presented. These criteria will in the following be used in the derivation of the robustness results.

2.1 THE SYSTEMS CONSIDERED

The systems considered in this part of the thesis, are assumed to be linear and time-invariant. They have one input and one output (SISO), unless otherwise stated. Both continuous and discrete time systems are considered. There is no assumption on the dimensionality of the systems. This is important, since real processes can always be considered as infinite dimensional and any reasonable model is of finite dimension.

Associated with a system G is a transfer function G(z), where z is a complex variable.

We will use the following stability concept:
A system is said to be stable if its transfer function G(z) has all its poles in the region Z. For continuous time systems $Z = \{z: \mathrm{Re}(z) < -\sigma \ , \ \sigma > 0\}$, and for discrete time systems $Z = \{z: |z| < 1-\sigma \ , \ \sigma > 0\}$. A system is said to be unstable if it is not stable. We assume that the systems considered have at most a finite number of unstable poles.

This concept of stability is equivalent to exponential stability.

A convenient way to incorporate practical stability constraints is to change the definition of the stability region Z. This is discussed in Section 5.2.

It is assumed that the systems considered do not contain any unstable modes that are uncontrollable or unobservable. This is a natural assumption since we work with input-output descriptions of systems.

2.2 THE STANDARD CONFIGURATION

A general configuration for linear feedforward - feedback control of a single-input system (process) G_p is shown in Fig. 2.1. Here the scalar u is the system input, and y is a vector of measured system variables used for feedback generation. The vector r contains external signals, such as command inputs and measured disturbances forcing the system. The controller consists of two parts, a feedforward compensator G_{ff}, and a feedback regulator G_{fb}.

From a stability point of view it is only the closed-loop system, formed by G_p and G_{fb}, that is relevant. For this reason introduce a SISO system G with transfer function

$$G(z) = G_{fb}(z)G_p(z) \tag{2.1}$$

Stability of the control system in Fig. 2.1 is then equivalent to the stability of the system G with unit feedback applied around it. The equivalent feedback system is shown in Fig. 2.2. It is this basic feedback configuration that will be analyzed with respect to robustness in the following.

The control system shown in Fig. 2.1 consists of a SIMO process and a MISO feedback regulator. We can of course also analyze control systems with MISO processes and SIMO controllers, the only important thing is that somewhere in the feedback loop, the loop transfer function can be regarded as SISO, like (2.1).

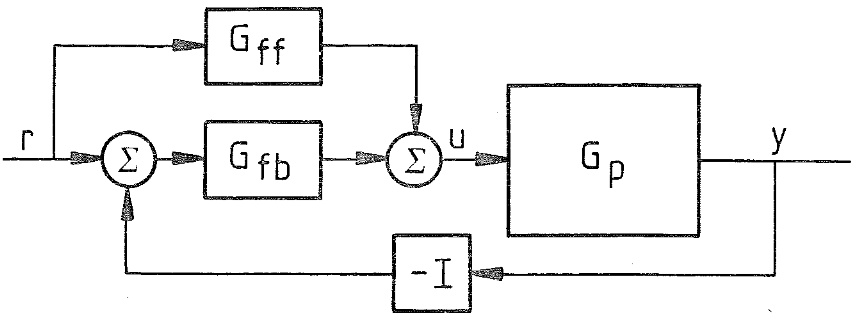


Fig. 2.1 - A general control configuration.

In the light of what has been said above, an open-loop SISO system will in the sequel be regarded as consisting of a plant and a regulator connected in cascade.

2.3 FREQUENCY DOMAIN STABILITY CRITERIA

In the stability analysis of a closed-loop system in the standard configuration of Fig. 2.2, there are some powerful frequency domain criteria available. Before formulating these stability results, we present two important concepts.

First, introduce the Nyquist Contour Γ as a curve that simply encloses the instability region C\Z (the complement of the stability region). When a transfer function is evaluated along the Nyquist Contour and the transfer function has poles on the boundary of the instability region, the Nyquist Contour is modified to make small semi-circles (indentations) around the singularities. In this way, small neighborhoods of the poles are added to the encircled region, and there are no singularities on the Nyquist Contour itself. A transfer function $G(z)$ is thus analytic in a region of the complex plane, containing the Nyquist Contour Γ. In Figs. 2.3 and 2.4 are shown the Nyquist Contours for the continuous and discrete time cases respectively. The contours are modified to include a small neighborhood of the point where integrator poles appear, to the encircled instability region.

Next, introduce the Nyquist Curve $G(\Gamma)$ as the directed closed curve

$$G(\Gamma) = \{G(z): z \in \Gamma\} \tag{2.2}$$

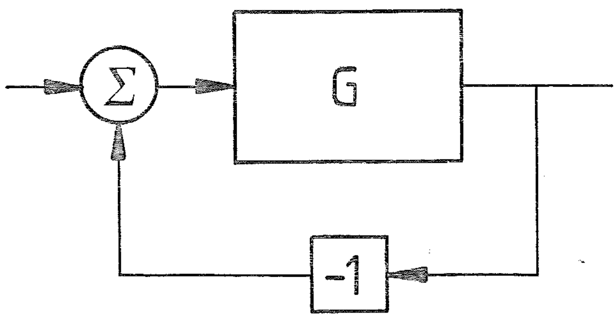


Fig. 2.2 - The standard feedback configuration.

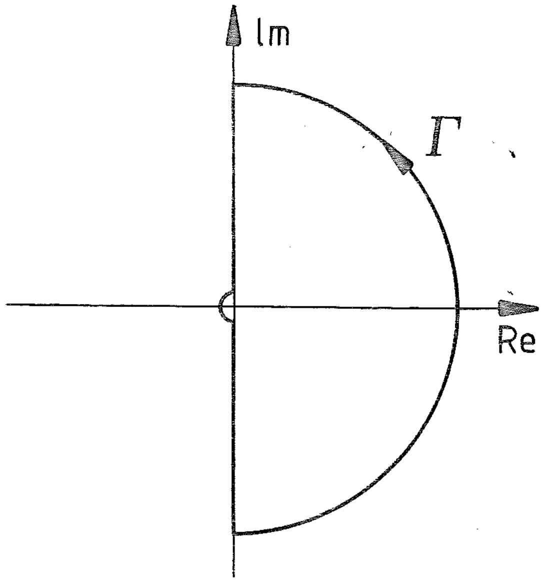


Fig. 2.3 - The contour Γ in the continuous time case.

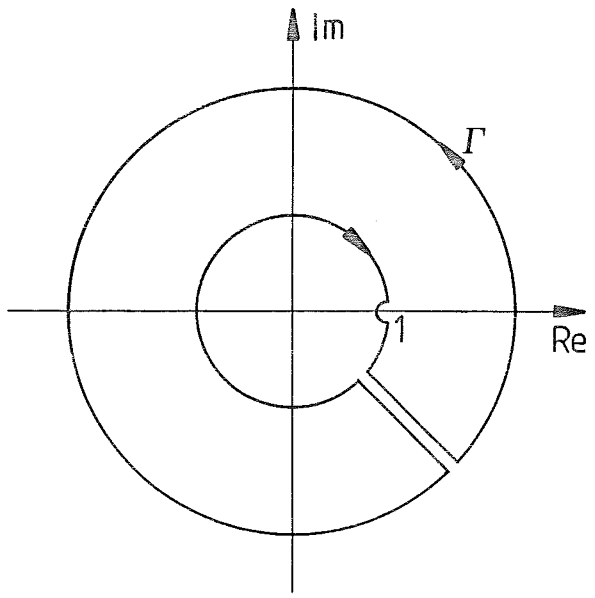


Fig. 2.4 - The contour Γ in the discrete time case.

obtained when the Nyquist Contour Γ is traversed in a positive direction (anti-clockwise), as viewed from an interior point of the instability region. The major part of the Nyquist Curve G(Γ), as it is defined here, can be obtained from ordinary frequency analysis. Now, we have the following two wellknown stability theorems:

The Nyquist Criterion. Consider a closed-loop system in the standard configuration (Fig. 2.2). Require that G(z) is a stable transfer function.

Then the closed-loop system is stable if and only if the Nyquist Curve $G(\Gamma)$ does not encircle the point $z = -1 + i \cdot 0$ □

The Nyquist Criterion is a special case of the more general:

Principle of Variation of the Argument. Consider a closed-loop system in the standard configuration (Fig. 2.2).

Then the closed-loop system has

$$n = p + \frac{1}{2\pi} \Delta_{\Gamma} \arg [1 + G(z)] \qquad (2.3)$$

unstable poles (the zeros of $1 + G(z)$ in the instability region). Here p is the number of open-loop unstable poles of $G(z)$, and $1/2\pi \cdot \Delta_{\Gamma} \arg[1+G(z)]$ is the number of times the Nyquist Curve $G(\Gamma)$ encircles the point $z = -1 + i \cdot 0$ in a positive direction (anti-clockwise). □

The Principle of Variation of the Argument is also called the Generalized Nyquist Criterion. The probably most general formulation is found in Desoer and Wang (1980), which also treats the multivariable case.

3. ROBUSTNESS

In this chapter we present the robustness problem in a form that is suited for our analysis. Some already published robustness results are reviewed. The proofs given are very simple and provide a good understanding of the mechanism that may cause instability. It will also be pointed out the reason why the results only apply to stable systems.

3.1 THE PROBLEM

The background of the robustness problem was given in the introduction. Since it is difficult to analyze every special controller-process configuration, we recall the discussion in Section 2.2 and focus our attention on the closed-loop system in the standard configuration, shown in Fig. 2.2. The element in the forward path is considered as a process and a feedback controller in series.

We will deal with two different systems:

<u>The model system</u> G_m, consisting of a model of the process and a controller designed to control this model. The closed-loop system formed from G_m is stable, and will meet the design specifications.

<u>The real system</u> G, consisting of the real process and the controller designed for the model. Since the model system and the real system in general will be different, it is not at all evident that the real system is closed-loop stable just because the model system has this property.

The problem is to derive conditions under which closed-loop stability of the model system G_m implies closed-loop stability of the real system G. The results will involve bounds on the magnitude of otherwise arbitrary errors between G and G_m.

3.2 ADDITIVE PERTURBATIONS

Consider a model system G_m in the standard feedback configuration shown in Fig. 2.2. Assume that this closed-loop system with the transfer function

$$H_m(z) = \frac{G_m(z)}{1 + G_m(z)} \tag{3.1}$$

is stable. We will now assume that the real system can be described as the model system G_m subjected to an additive perturbation, such that the transfer function is

$$G(z) = G_m(z) + \Delta G_a(z) \tag{3.2}$$

where $\Delta G_a(z)$ is small in some sense. It then seems plausible that the closed-loop system obtained should be stable. This is formally stated in the following theorem.

THEOREM 3.1
Consider the closed-loop system obtained by applying unit feedback around a system with the transfer function

$$G(z) = G_m(z) + \Delta G_a(z) \tag{3.3}$$

Assume that the system with the transfer function

$$H_m(z) = \frac{G_m(z)}{1 + G_m(z)} \tag{3.4}$$

is stable.

Then the closed-loop system considered is stable if

(i) $\Delta G_a(z)$ is stable.

(ii) $|\Delta G_a(z)| < |1 + G_m(z)|$ on the Nyquist Contour Γ.

Proof. Assume that the conditions (i) and (ii) are satisfied. The input-output relation of the real closed-loop system is

$$(1 + G(z))y(z) = G(z)y_c(z) \tag{3.5}$$

where y is the output and y_c is a command input. Using (3.3), this relation can be written as

$$(1 + G_m(z))y(z) = G_m(z)y_c(z) + \Delta G_a(z)(y_c(z) - y(z)) \tag{3.6}$$

A block diagram of this feedback system is shown in Fig. 3.1.

From a stability point of view, it is only the feedback loop that is interesting. The corresponding open-loop system with the loop transfer function

$$\frac{\Delta G_a(z)}{1 + G_m(z)} \tag{3.7}$$

is stable according to condition (i), and the condition (ii) implies that

$$\sup_{z\in\Gamma} \left| \frac{\Delta G_a(z)}{1 + G_m(z)} \right| < 1 \tag{3.8}$$

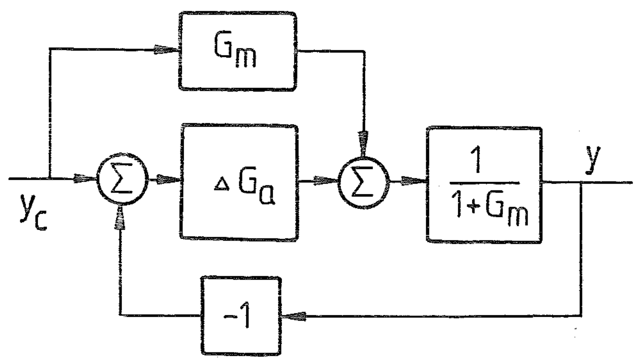


Fig. 3.1 - A representation of the closed-loop system.

Hence, application of the Nyquist Criterion yields closed-loop stability. □

The block diagram in Fig. 3.1 shows that, in the terminology of the "Small Gain Theorem", the closed-loop system is stable if the loop gain in the "error"-loop is less than unity. It is important to get a feeling for how sharp this robustness result is. If the inequality in the theorem is violated, there are ΔG_a's that will cause the closed-loop system to be unstable. Hence, for some ΔG_a the theorem gives both necessary and sufficient conditions for stability. On the other hand, there are ΔG_a that violate the inequality in the theorem but the closed-loop system still is stable. Therefore for a <u>specific</u> ΔG_a, Theorem 3.1 gives sufficient conditions for stability. For <u>arbitrary</u> ΔG_a, only characterized by the magnitude, the conditions are both necessary and sufficient. If the error has a partially or completely known structure, it may be better to use some other method for stability analysis, dedicated to systems with the known structure.

In a real situation $G(z)$ and $\Delta G_a(z)$ usually are unknown. Since the model transfer function $G_m(z)$ is an approximation of $G(z)$, it is reasonable that there is an upper bound on the error $|\Delta G_a|$ available. If the model is obtained from system identification, such a bound may be a function of the frequency variable and therefore known on the Nyquist Contour Γ, e.g.

$$|\Delta G_a(i\omega)| \le f(i\omega) \tag{3.9}$$

where $f(i\omega)$ is an upper bound on the error. Since $f(i\omega)$ is usually the best available information about $|\Delta G_a|$, stability criteria like

$$f(i\omega) < |1 + G_m(i\omega)| \tag{3.10}$$

are as sharp as possible. This observation should be kept in mind when reading the rest of this, and the next chapter.

3.3 MULTIPLICATIVE PERTURBATIONS

A result similar to Theorem 3.1, holds if the model system transfer function is perturbed in a multiplicative way instead

$$G(z) = G_m(z)[1 + \Delta G_r(z)] \qquad (3.11)$$

We then have:

THEOREM 3.2
Consider the closed-loop system obtained by applying unit feedback around a system with the transfer function

$$G(z) = G_m(z)[1 + \Delta G_r(z)] \qquad (3.12)$$

Assume that the system with the transfer function

$$H_m(z) = \frac{G_m(z)}{1 + G_m(z)} \qquad (3.13)$$

is stable.

Then the closed-loop system considered is stable if

(i) $\Delta G_r(z)$ is stable.

(ii) $|\Delta G_r(z)| < \left| \frac{1 + G_m(z)}{G_m(z)} \right|$

on the Nyquist Contour Γ

<u>Proof</u>. Simple algebraic manipulations applied to the equations for the closed-loop system, give that it can be represented as is shown in Fig. 3.2. The rest of the proof is analogous to the proof of Theorem 3.1. □

<u>Remark 1</u>. Notice that in both Theorem 3.1 and 3.2, the stability of the perturbation $\Delta G(z)$ is crucial for the application of the Nyquist Criterion. □

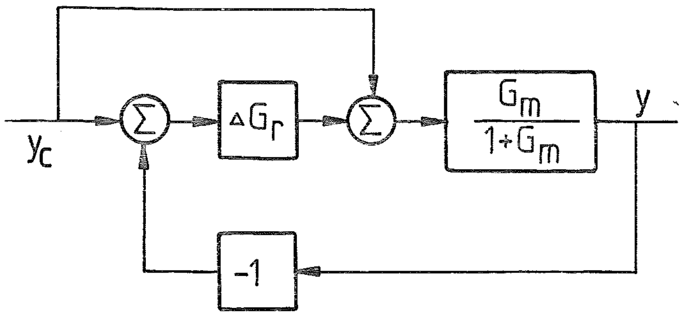


Fig. 3.2 - A representation of the closed-loop system.

Remark 2. Notice that in both Theorem 3.1 and 3.2, an important measure of the stability robustness is the magnitude of the Return Difference

$$R(z) = 1 + G_m(z) \tag{3.14}$$

□

The results in Theorems 3.1 and 3.2 are SISO versions of robustness results for multivariable systems due to Doyle (1978) and Sandell Jr. (1979). In the multivariable case the sufficient conditions for stability unfortunately become rather difficult to grasp. This will become apparent in the next chapter.

Since the proofs of Theorem 3.1 and 3.2 are based on the Nyquist Criterion, it can not be expected that the same method of proof will work when the perturbation $\Delta G(z)$ is unstable. This is due to the inherent property of the Nyquist Criterion, that the open-loop system should be stable. An obvious remedy is to utilize the Principle of Variation of the Argument, which also applies to unstable open-loop systems.

4. NEW ROBUSTNESS RESULTS

The robustness results presented in the previous chapter were derived with the assumption that the modelling error transfer function ΔG was stable. This means that in order to apply Theorems 3.1 and 3.2 to open-loop unstable systems, the unstable poles of the real process must also be the unstable poles of the model. In a practical situation, such good modelling is never possible. Consequently the use of the theorems is restricted to stable systems.

In the following, robustness results which can be applied to unstable systems are derived with the aid of the Principle of the Variation of the Argument.

4.1 GENERAL PERTURBATIONS

With the background of the previous chapter, we are now able to state a robustness result for systems with general perturbations. The results in Theorems 3.1 and 3.2 both turn out to be special cases of this theorem.

THEOREM 4.1
Consider the closed-loop system obtained by applying unit feedback around a system with the transfer function $G(z)$. Assume that the system with the transfer function

$$H_m(z) = \frac{G_m(z)}{1 + G_m(z)} \tag{4.1}$$

is stable. Assume further that the inequality

$$|G(z) - G_m(z)| < |1 + G_m(z)| \tag{4.2}$$

is satisfied on the Nyquist Contour Γ.

Then the closed-loop system considered is stable if and only if G and G_m have the same number of unstable poles.

<u>Proof</u>. Introduce the quotient between the return differences

of the real system and the model

$$f(z) = \frac{1 + G(z)}{1 + G_m(z)} \tag{4.3}$$

The function $f(z)$ has $p + p(G_m)$ zeros in the instability region and $p(G)$ unstable poles. Here p is the number of unstable poles of the closed-loop system (zeros of $1+G(z)$), and $p(G)$ and $p(G_m)$ are the number of unstable poles of G and G_m respectively.

Simple calculations give that

$$f(z) = 1 + \frac{G(z) - G_m(z)}{1 + G_m(z)} \tag{4.4}$$

The Principle of the Variation of the Argument applied to $f(z)$ gives

$$p + p(G_m) - p(G) = \frac{1}{2\pi}\Delta_\Gamma \arg\left[1 + \frac{G(z) - G_m(z)}{1 + G_m(z)}\right] \tag{4.5}$$

Use of assumption (4.2) in this equation renders the result

$$p = p(G) - p(G_m) \tag{4.6}$$

The stability result in the theorem follows from the relation (4.6). □

Remark 1. Notice that the inequality (4.2) is satisfied whenever

$$|G(z)| < 1/3 \quad \text{and} \quad |G_m(z)| < 1/3 \tag{4.7}$$

For systems with strictly proper transfer functions, the inequality is thus always satisfied at $|z| = \infty$. □

Remark 2. Theorems 3.1 and 3.2 are special cases of Theorem 4.1. □

Theorem 4.1 provides the interesting result, that it is important to know the number of unstable poles of a process to ensure robustness. The exact locations of the unstable poles are not crucial. Notice also, that it is <u>not</u> sufficient to know that the frequency curves corresponding to G and G_m are close.

This is illustrated by the following example.

EXAMPLE 4.1

Consider a servo system with the transfer function

$$G(s) = \frac{100}{s(s+1)(100-s)} \qquad (4.8)$$

A model of the process has the transfer function

$$G_m(s) = \frac{1}{s(s+1)} \qquad (4.9)$$

The Nyquist curves of these two transfer functions are close for all frequencies ω, $s = i\omega$

$$|G(i\omega) - G_m(i\omega)| \le 0.01 \qquad (4.10)$$

It follows from Theorem 4.1 that, although the closed-loop system formed from the model is stable, the real closed-loop system is unstable.

The unstable pole will only influence the frequency response at high frequencies $\omega > 100$. For these frequencies the transfer function will have a very small magnitude $< 10^{-4}$, because of the poles at $s = 0$ and $s = -1$ (high frequency roll-off). □

As was previously mentioned, the model transfer function $G_m(z)$ is composed of the transfer functions of the process model and the regulator

$$G_m(z) = G_{fb}(z)G_p(z) \qquad (4.11)$$

The error between the real system and the model loop transfer functions, G and G_m, then becomes

$$\Delta G(z) = G_{fb}(z)\Delta G_p(z) \tag{4.12}$$

and the robustness inequality (4.2) can be formulated as

$$|G_{fb}(z)\Delta G_p(z)| < |1 + G_{fb}(z)G_p(z)| \tag{4.13}$$

This result gives a largest allowable bound on the absolute error $|\Delta G_p|$ between the process and the process model transfer functions.

It is clear from this relation, that the controller G_{fb} can be designed to improve the robustness properties of the system. The high frequency parts of the process are often difficult to model correctly. The effects of time-delays, neglected high order dynamics, and electro-mechanical resonances can be quite pronounced. To satisfy the robustness inequality, it is then necessary to make sure that the magnitude of the regulator transfer function $G_{fb}(z)$ is small for high frequencies, as is seen from (4.13).

To obtain a robust control system, the regulator should thus be designed so that the loop gain drops rapidly for frequencies above the desired bandwidth. It is also important not to choose the closed-loop bandwidth too high, since the model will then be poor in the working range of the system. At the bandwidth frequency the loop transfer function has a magnitude of about unity. The return difference may consequently be small, less than one in magnitude, giving poor robustness and large sensitivity to modelling errors. Wellknown concepts for stability, amplitude and phase margins, essentially tell how small the return difference becomes in this frequency interval.

4.2 A ROBUSTNESS RESULT INVOLVING SYSTEM ZEROS

In the previous section, a rather general robustness result was given. The essence was that it is possible to draw conclusions about closed-loop stability properties for two systems with close Nyquist curves, if they have the same number of unstable open-loop poles. It is in fact possible to circumvent this, by instead looking at the inverse transfer functions. We then obtain the following result:

THEOREM 4.2
Consider the closed-loop system obtained by applying unit feedback around a system with the transfer function G(z). Assume that the system with the transfer function

$$H_m(z) = \frac{G_m(z)}{1 + G_m(z)} \tag{4.14}$$

is stable. Assume further that the inequality

$$|[G(z) - G_m(z)]/G(z)| < |1 + G_m(z)| \tag{4.15}$$

is satisfied on the Nyquist Contour Γ.

Then the closed-loop system considered is stable if and only if G and G_m have the same (finite) number of zeros in the instability region.

<u>Proof</u>. Form the quotient between the closed-loop transfer functions of the model and the real system

$$f(z) = \frac{G_m(z)}{1 + G_m(z)} \cdot \frac{1 + G(z)}{G(z)} \tag{4.16}$$

The function $f(z)$ has $p + n(G_m)$ zeros in the instability region and $n(G)$ unstable poles. Here p is the number of unstable poles of the closed-loop system, and $n(G)$ and $n(G_m)$ are the number of zeros in the instability region for G and G_m respectively.

Now straightforward calculations give

$$f(z) = 1 + \frac{[G_m(z) - G(z)]/G(z)}{1 + G_m(z)} \tag{4.17}$$

The assumption (4.15), equation (4.17) and the Principle of Variation of the Argument applied to $f(z)$ give

$$p + n(G_m) - n(G) =$$

$$= \frac{1}{2\pi}\Delta_{\Gamma} \arg\left[1 + \frac{[G_m(z) - G(z)]/G(z)}{1 + G_m(z)}\right] = 0 \qquad (4.18)$$

The result in the theorem follows from equation (4.18). □

Remark 1. The robustness inequality (4.15) can also be written as

$$|1/G(z) - 1/G_m(z)| < |1 + 1/G_m(z)| \qquad (4.19)$$

which illustrates the close connection of this result with that of Theorem 4.1. □

Remark 2. The robustness inequality (4.15) is always satisfied when

$$|G(z)| > 3 \quad \text{and} \quad |G_m(z)| > 3 \qquad (4.20)$$

This simplifies the verification of the robustness inequality (4.15) on the Nyquist Contour Γ. □

The robustness result Theorem 4.2 is obtained in terms of the number of non-minimum phase zeros of the transfer functions of the process and the model. It is thus important to use a model with the same number of "unstable" zeros as the real system. Just closeness of the Nyquist curves corresponding to G and G_m is not sufficient for stability.

The robustness inequality (4.15) gives insight into the problem of regulator design. For low frequencies the model of the process often is good. Even if there are large modelling errors at low frequencies, e.g. uncertain steady-state gain, the robustness inequality will be satisfied if the regulator has sufficiently high gain, as is seen from (4.15). This can easily be achieved by including integral action in the feedback controller.

At high frequencies, where there are often large relative errors between the real process and the model, the inequality (4.15) may be difficult to satisfy. At infinity $|z| = \infty$, a necessary condition is that the model G_m has larger or equal pole-zero excess (number of time-delays) than the the real continuous time (discrete time) system G. It is thus "safe" to over-estimate the number of time-delays, when modelling discrete time systems.

For discrete time systems we can draw special conclusions:

The discrete time processes that are encountered in real life, are with few exceptions sampled continuous time systems. The continuous time signals are then filtered in a pre-sampling filter, before sampled. This action removes the high-frequency components in the signals, with frequencies higher than the Nyquist frequency (half the sampling frequency). Any remaining high frequency component is after the sampling represented as a low frequency signal, according to the Sampling theorem (the Aliasing effect). The problems normally encountered at high frequencies for continuous time systems, are thus reduced or even removed for sampled systems. The longer the sampling period is chosen, the easier it is to find a simple model that is a good approximation of the real system. Long sampling periods thus faciliate the design of robust control systems, although the achievable design specifications (e.g. speed) may be limited.

To summarize desirable properties of a robust controller; it should have high gain at low frequencies and low gain at high frequencies. Consequently, somewhere between, the gain must be reduced. This occurs around the desired system bandwidth, where

$$|G_m(z)| = |G_{fb}(z)G_p(z)| \approx 1 \qquad (4.21)$$

It is very important to have a good model of the system in this frequency interval. Or conversely, the closed-loop system bandwidth should be chosen at a reasonably high frequency, where it is still possible to obtain a good model.

Notice that in both Theorems 4.1 and 4.2, the magnitude of the Return Difference

$$R(z) = 1 + G_m(z) \qquad (4.22)$$

is an important measure for how robust a system is to modelling errors. If the return difference is large in some frequency interval, considerable modelling errors between system and model transfer functions can be tolerated. The importance of the magnitude of the return difference to sensitivity and closed-loop stability, has been known for a long time, see e.g. Black (1934), Bode (1945) and Horowitz (1963).

4.3 MULTIVARIABLE RESULTS

In Doyle (1978), the results of Theorems 3.1 and 3.2 have been generalized to square multivariable systems. It is tempting to believe that there is a multivariable extension of Theorem 4.1 as well. Following the approach of Doyle, we proceed to derive this result and a multivariable result corresponding to Theorem 4.2.

First we introduce some concepts related to square complex matrices:

Let $G(z)$ be a $n\times n$ - matrix, with elements $g_{ij}(z)$ that are complex-valued functions of a complex variable z. It is assumed that $G(z)$ is nonsingular for almost all values of z.

The eigenvalues of the matrix $G(z)$ are refered to as the characteristic loci, and are denoted

$$\lambda_i[G(z)] \quad ; \ i = 1,\ldots,n \tag{4.23}$$

The singular values of $G(z)$

$$\sigma_i[G(z)] \quad ; \ i = 1,\ldots,n \tag{4.24}$$

are defined as the square-roots of the real, non-negative eigenvalues of the matrix

$$G(z)^* G(z) \tag{4.25}$$

where * means complex conjugated transpose.

The largest and smallest singular values can alternatively be defined as

$$\bar{\sigma}[G(z)] = \max_{\|x\|=1} \|G(z)x\| \tag{4.26}$$

$$\underline{\sigma}[G(z)] = \min_{\|x\|=1} \|G(z)x\| \tag{4.27}$$

where $\|\cdot\|$ is the usual Euclidean norm. According to (4.26) and (4.27), they can be interpreted as the extreme values of the gain that the matrix G(z) represents.

The following relations will be useful in the sequel

$$\max_i |\lambda_i [G(z)]| \le \bar{\sigma}[G(z)] \qquad (4.28)$$

$$\bar{\sigma}[A \cdot B^{-1}] \le \bar{\sigma}[A]/\underline{\sigma}[B] \qquad (4.29)$$

We have the following result for multivariable systems:

THEOREM 4.3
Consider the closed-loop system obtained by applying unit feedback around a square multivariable system with the transfer matrix G(z). Assume that the system with the transfer matrix

$$H_m(z) = G_m(z)[I + G_m(z)]^{-1} \qquad (4.30)$$

is stable. Assume further that the inequality

$$\bar{\sigma}[G(z) - G_m(z)] < \underline{\sigma}[I + G_m(z)] \qquad (4.31)$$

is satisfied on the Nyquist Contour Γ.

Then the closed-loop system considered is stable if and only if G and G_m have the same number of unstable poles.

Proof. As in Desoer and Vidyasagar (1975) we have a relation between open-loop (OLCP) and closed-loop (CLCP) characteristic polynomials

$$CLCP(z)/OLCP(z) = \det[I + G(z)] \qquad (4.32)$$

$$CLCP_m(z)/OLCP_m(z) = \det[I + G_m(z)] \qquad (4.33)$$

These two determinants will now be related through a simple calculation (the argument z will be omitted for convenience)

$$\det[I + G] = \det[I + (G-G_m)(I+G_m)^{-1}]\cdot\det[I + G_m] \qquad (4.34)$$

and hence

$$CLCP/OLCP\cdot OLCP_m/CLCP_m = \det[I + (G-G_m)(I+G_m)^{-1}] \qquad (4.35)$$

Let p be the number of unstable poles of the closed-loop system, and $p(G)$ and $p(G_m)$ the number of open-loop unstable poles of G and G_m respectively. We will now examine the variation of the argument of this expression as z traverses the Nyquist Contour Γ. From the Generalized Nyquist Criterion of Desoer and Wang (1980) it follows that

$$p + p(G_m) - p(G) = \frac{1}{2\pi}\Delta_\Gamma \arg \det[I + (G-G_m)(I+G_m)^{-1}] =$$

$$= \frac{1}{2\pi}\Delta_\Gamma \arg \prod_{i=1}^{n} [1 + \lambda_i[(G-G_m)(I+G_m)^{-1}]] =$$

$$= \frac{1}{2\pi}\sum_{j=1}^{m} \Delta_\Gamma \arg [1 + \gamma_j] \qquad (4.36)$$

Here $\{\gamma_j ; j=1,\ldots,m\}$ is an indexed family of closed circuits (graphs), formed from the Nyquist curves of the characteristic loci $\{\lambda_i(\Gamma); i=1,\ldots,n\}$.

Now assumption (4.31), and the relations (4.28) and (4.29) give that

$$\bar{\sigma}[(G(z) - G_m(z))(I + G_m(z))^{-1}] < 1 \qquad (4.37)$$

and

$$\max_i |\lambda_i[(G(z) - G_m(z))(I + G_m(z))^{-1}]| < 1 \qquad (4.38)$$

Hence

$$\max_{j} \sup_{z \in \gamma_j} |z| < 1 \qquad (4.39)$$

Using this last inequality in equation (4.36) gives the result in the theorem:

$$p + p(G_m) - p(G) = 0 \qquad (4.40)$$

□

Remark. The robustness results for multivariable systems by Doyle (1978), are special cases of Theorem 4.3. □

The multivariable robustness result corresponding to Theorem 4.2 is derived in a similar way:

THEOREM 4.4
Consider the closed-loop system obtained by applying unit feedback around a square multivariable system with the transfer matrix $G(z)$.

Assume that the system with the transfer matrix

$$H_m(z) = G_m(z)[I + G_m(z)]^{-1} \qquad (4.41)$$

is stable. Assume further that the inequality

$$\bar{\sigma}[G(z)^{-1} - G_m(z)^{-1}] < \underline{\sigma}[I + G_m(z)^{-1}] \qquad (4.42)$$

is satisfied on the Nyquist Contour Γ.

Then the closed-loop system considered is stable if and only if G and G_m have the same (finite) number of zeros in the instability region.

Proof. The closed-loop characteristic polynomial and zero polynomial ZP(z) are related by

$$CLCP(z)/ZP(z) = \det[I + G(z)^{-1}] \qquad (4.43)$$

$$CLCP_m(z)/ZP_m(z) = \det[I + G_m(z)^{-1}] \qquad (4.44)$$

By simply substituting the equations (4.43) and (4.44) for the equations (4.32) and (4.33), the proof of Theorem 4.3 gives the result. □

The robustness results for multivariable systems, Theorems 4.3 and 4.4, involve stability conditions given in terms of the singular values of transfer matrices. These concepts are not trivial to visualize. However, using (4.26) and (4.27) as interpretations of the singular values as matrix gains (norms), the robustness inequalities (4.31) and (4.42) can be seen as gain-constraints on the loop transfer functions. It is then obvious that the implications on feedback controller design, drawn for SISO systems, are valid also for MIMO systems. The controller should thus have high gain for low frequencies and low gain for high frequencies. In the frequency interval where the singular values of the loop transfer function of the model are close to unity, it is important to have a good model of the process. This frequency region can be interpreted as a generalization of the bandwidth concept for SISO systems.

5. ROBUSTNESS OF PERFORMANCE

In a practical situation, closed-loop stability is the first, and basic requirement on a control system. For most applications, however, it is not enough. The closed-loop system must also meet some design specifications. In this chapter we will briefly discuss some different approaches to the analysis of performance robustness.

5.1 DESIGN SPECIFICATIONS

There are mainly three distinct ways to give specifications of a control system. One is to directly specify how the system should behave under working conditions in the time domain. Another is to give specifications in the frequency domain, and hereby indirectly give the system desirable time domain properties. The last alternative is to specify that the system should optimize a criterion related to system performance in either time or frequency domain.

Time Domain Specifications. The time domain specifications for control systems are mainly formulated for servo systems. It is common to give characteristics of the closed-loop system step response. Such features may be delay, rise-time and over-shoot. Unless the system is of low order, there is no simple way to translate the specified time-domain properties to characteristics of the transfer function. This has as a consequence that analysis of sensitivity and performance robustness often is carried out in the time-domain if the specifications are given there. Such analysis may incorporate the tools of Geometric Control Theory, see e.g. Molander (1979).

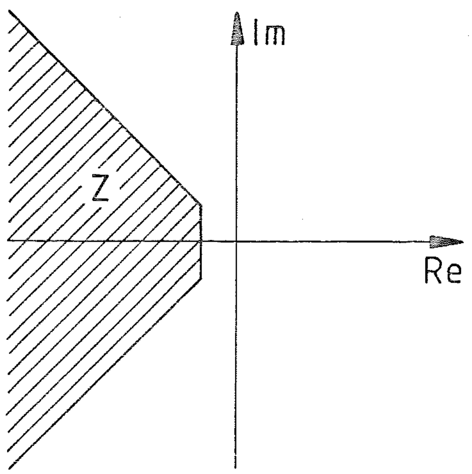
Frequency Domain Specifications. The specifications in the frequency domain can be given in two different ways. First there are "qualitative" characterizations of the frequency response of the closed-loop transfer function. These may be bandwidth, resonance peak and high-frequency roll-off. Corresponding properties of the open-loop transfer function can also be given, e.g. phase- and amplitude margins. As was seen in Theorem 4.1 and 4.2, these concepts are important to the robustness of a system.

The other way to give specifications in the frequency domain, is to directly specify locations for some or all of the closed-loop system poles. E.g. in a continuous time servo system, a pair of control-poles are specified and the rest of the poles are restricted to lie far to the left in the left complex half-plane. In this way, the system behaviour will approximate that of a second order system.

<u>Optimization</u>. The formulation of this type of specification is usually that the control system should optimize (minimize) some criterion related to performance. The most common and frequently used method is the Linear Quadratic (Gaussian) technique. It can be used both for systems in time domain and frequency domain description.

5.2 PRACTICAL STABILITY

When system specification is made by explicitly prescribing a set of closed-loop poles, one way to analyze the robustness is to examine the sensitivity of the closed-loop pole locations to small modelling errors in the transfer functions. This approach was used in Åström (1979a). Another approach is to use practical stability - a system has desirable properties if and only if all the poles lie in a subset, or restriction, of the stability region (see Chapter 2). Examples of restricted stability regions are shown in Figs. 5.1 and 5.2. Analysis of performance robustness is then to examine whether the system is stable in the restricted sense or not.



<u>Fig. 5.1</u> - A restricted stability region for continuous time systems.

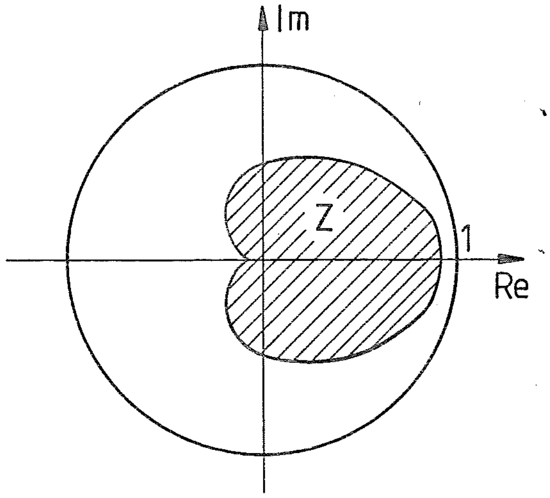


Fig. 5.2 - A restricted stability region for discrete time systems.

The analysis is carried out by redefining Z as the restricted stability region. The Nyquist Contour Γ is now a curve that encircles $C \backslash Z$, the complement of the restricted stability region. With this definition of Z, all the results in Chapters 3 and 4 still are valid, except that "stability" should be replaced by "restricted stability".

Unfortunately the bound on the model error $|\Delta G(z)|$ along the Nyquist Contour Γ now loses its physical interpretation as a function of the frequency (Compare the discussion in Section 3.2).

The restricted stability introduced above, is equivalent to the concept of λ-stability, introduced by Pernebo (1981).

5.3 SENSITIVITY OF THE TRANSFER FUNCTION

Often the specifications of a system are given in terms of the frequency response of the closed-loop system transfer function. The performance analysis can then be to examine how an error in the open-loop transfer function is related to an error in the closed-loop transfer function.

In the following, we assume that both the real and the model systems are closed-loop stable. Unless this assumption is satisfied, the analysis of robustness of performance is quite meaningless.

Introduce the transfer functions of the closed-loop systems for the real and the model systems

$$H(z) = \frac{G(z)}{1 + G(z)} \tag{5.1}$$

$$H_m(z) = \frac{G_m(z)}{1 + G_m(z)} \tag{5.2}$$

We now have a result on robustness of performance:

THEOREM 5.1
The relative error in the closed-loop transfer function is given by

$$\frac{H(z) - H_m(z)}{H(z)} = \frac{1}{1 + G_m(z)} \cdot \frac{G(z) - G_m(z)}{G(z)} \tag{5.3}$$

Proof. Insertion of equations (5.1) and (5.2) in relation (5.3) verifies the result. □

Corollary. In case of small differences

$$dG(z) = G(z) - G_m(z) \tag{5.4}$$

the relation (5.3) reduces to Bode's wellknown sensitivity formula

$$\frac{dH(z)}{H(z)} = \frac{1}{1 + G(z)} \cdot \frac{dG(z)}{G(z)} \tag{5.5}$$

□

Remark. The multivariable version of result (5.3) is

$$[H - H_m]H^{-1} = [I + G_m]^{-1}[G - G_m]G^{-1} \tag{5.6}$$

where the argument z has been omitted for simplicity. □

The relations (5.3) and (5.6) are true independently of the magnitude of the difference between G and G_m, and hence provide a very useful tool when examining sensitivity with respect to large variations in the plant transfer function.

Theorem 5.1 gives a nice characterization of how an error in the open-loop transfer function affects the closed-loop transfer function. The return difference

$$R(z) = 1 + G_m(z) \tag{5.7}$$

once more is noted to be important to robustness and sensitivity. Accuracy of the open-loop system model is only crucial in those frequency intervals where the return difference is small.

In a normal control system, the feedback controller is designed to give the closed-loop system certain properties. The next chapter provides some examples that illustrate how controllers can be designed to give acceptable robustness.

6. EXAMPLES

In this chapter a number of examples are given. It is believed that the theory of the previous chapters can be more appreciated with the supplement of some illustrations.

EXAMPLE 6.1

Consider the design of a continuous time position servo. A model of the open-loop system typically has the transfer function

$$G(s) = \frac{1}{s(sT + 1)} \qquad (6.1)$$

A simple proportional controller

$$u = K \cdot [y_c - y] \qquad (6.2)$$

with $K = 1/(4T\zeta^2)$ gives a closed-loop system with the transfer function

$$H(s) = \frac{1}{(s/\omega_0)^2 + 2\zeta(s/\omega_0) + 1} \qquad (6.3)$$

where $\omega_0 = 1/(2\zeta T)$. The damping ζ, is a design parameter which is commonly chosen between 0.5 and 1.0. Here y_c is a command input.

As has been pointed out in the previous chapters, the return difference R(s) is an important measure of the robustness properties of a system. In this example, the return difference is

$$R(s) = 1 + KG(s) = \frac{(s/\omega_0)^2 + 2\zeta(s/\omega_0) + 1}{(s/\omega_0)^2 + 2\zeta(s/\omega_0)} \qquad (6.4)$$

In Fig. 6.1, the logarithm of $|R(i\omega)|$ is shown as a function of the normalized frequency ω/ω_0 for some different values of ζ. It is seen that in this case, it is possible to divide the frequency axis into three parts where the return difference has principally different behaviour.

Low frequencies. ($\omega \ll \omega_0$)

In this region, the return difference is large. This will always be the case if the open-loop transfer function is much larger than unity. According to the results in the previous chapters, rather large modelling errors therefore can be permitted without affecting stability or performance. This is in fact one of the reasons why it is desirable to have integral action in the loop transfer function.

High frequencies. ($\omega \gg \omega_0$)

For high frequencies the return difference is close to unity. Since the open-loop system transfer function is small in this region, moderate modelling errors can be tolerable. One wellknown exception is when there are poorly damped (mechanical) resonances in the system. In this case it is important to use a feedback controller with drastically decreasing gain for high frequencies. This is illustrated in Example 6.3. The need for decreasing loop

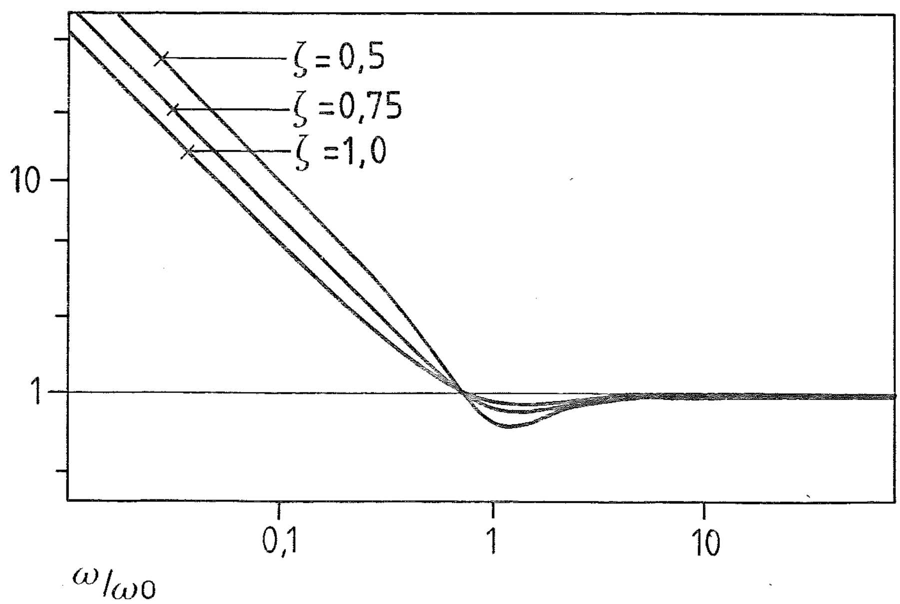


Fig. 6.1 - The return difference for a typical servo.

gain at high frequencies can be included in the control system specifications as desired high-frequency roll-off.

Intermediate frequencies. ($\omega \approx \omega_0$)

From Fig. 6.1, it is clear that in this region the return difference may attain small values, and consequently the robustness can be poor. The frequencies where such interesting properties as bandwidth, phase- and amplitude margins are defined, are in this interval. These facts emphasize the importance of correct modelling and suitably designed controllers for intermediate frequencies.

For further discussion on the return difference of servo systems, see e.g. Kwakernaak and Sivan (1972).

EXAMPLE 6.2

The behaviour from rudder to heading angle of a 255 000 tdw oil tanker can be modelled rather accurately by a 2:nd order model, the Nomoto model, with transfer function

$$G_1(s) = \frac{K}{s(s + a)} \qquad (6.5)$$

The coefficient a is very sensitive to variations in trim and loading of the ship, it can even change sign which indicates that the ship is sometimes unstable. Under normal operations the value of the gain is $K = 2 \cdot 10^{-4}$, and the pole $s = -a$ lies in the interval

$$-0.01 \leq a \leq 0.01 \qquad (6.6)$$

The model and the numerical values of the parameters are taken from Åström (1980).

Now, introduce a simplification of the Nomoto model (6.5)

$$G_2(s) = K/s^2 \qquad (6.7)$$

This model is reasonable for most ω-values, and will be used for controller design purposes.

The simplest controller that stabilizes this model is a PD-controller with the transfer function

$$G_{PD}(s) = K_P + K_D s \qquad (6.8)$$

Introduce the loop transfer functions of the real (Nomoto) and model (simplified Nomoto) systems

$$G(s) = G_1(s)G_{PD}(s) \qquad (6.9)$$

$$G_m(s) = G_2(s)G_{PD}(s) \qquad (6.10)$$

Design. The closed-loop transfer function of the model system, formed by G_2 and the PD-controller is

$$H_m(s) = \frac{KK_D s + KK_P}{s^2 + KK_D s + KK_P} \qquad (6.11)$$

We let the parameters of the controller be related through

$$KK_D/2 = (KK_P)^{1/2} = \omega_0 \qquad (6.12)$$

which corresponds to placing both closed-loop poles at $s = -\omega_0$.

Robustness. We will now examine if the closed-loop system obtained from the Nomoto model G_1 and the PD-controller is stable. For certain values of the a-parameter the open-loop systems G and G_m have different number of unstable poles, and since the models have no zeros in the instability region, the robustness result we should use is Theorem 4.2.

To verify the robustness inequality (4.15), we need the following functions

$$|[G(s) - G_m(s)]/G(s)| = |a/s| \qquad (6.13)$$

and

$$|1 + G_m(s)| = |(s^2 + 2\omega_0 s + \omega_0^2)/s^2| \qquad (6.14)$$

The inequality (4.15) now can be expressed as

$$|Q(s)| \equiv \left|\frac{1}{1 + G_m(s)} \cdot (G(s) - G_m(s))/G(s)\right| < 1 \qquad (6.15)$$

A Bode plot of the function $|Q(i\omega)|$ is shown in Fig. 6.2. It is clear that the stability condition (6.15) is satisfied if

$$|Q(i\omega)| \leq \left|\frac{a}{2\omega_0}\right| \leq \frac{0.01}{2\omega_0} < 1 \qquad (6.16)$$

Combining this result with (6.12) and the value of K, yields the stability bound

$$K_D > 50 \qquad (6.17)$$

and K_P given by (6.12).

Robustness of Performance. The real closed-loop system has the transfer function

$$H(s) = \frac{G(s)}{1 + G(s)} \qquad (6.18)$$

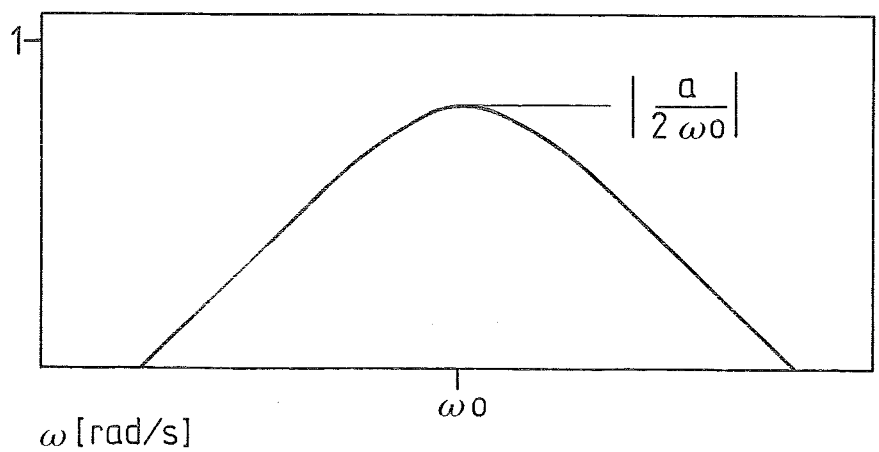


Fig. 6.2 - A plot of the function $|Q(i\omega)|$ in a logarithmic diagram.

It follows from Theorem 5.1, that the transfer functions for the closed-loop systems are related through

$$[H(s) - H_m(s)]/H(s) =$$

$$= \frac{1}{1 + G_m(s)} \cdot [G(s) - G_m(s)]/G(s) = Q(s) \qquad (6.19)$$

Suppose now that we want the relative error (6.19) to be less than 50 % . Inspection of Fig. 6.2 gives that

$$|Q(i\omega)| \le \left|\frac{a}{2\omega_0}\right| \le \frac{0.01}{2\omega_0} < 0.5 \qquad (6.20)$$

or equivalently, that

$$K_D > 100 \qquad (6.21)$$

and K_P given by (6.12).

Obviously the performance is improved when the derivative gain K_D is increased. This is however, not the whole truth.

The models (6.5) and (6.7) do not take noise disturbances into account. If the feedback is corrupted by high-frequency measurement noise, the high gain of the controller will cause the input to be too large. The rudder actuator might saturate and control will be poor. In a more complex situation like this, it may give pay-off to use a more complex controller.

EXAMPLE 6.3

A rigid body model of the pitch channel dynamics of an aircraft is

$$\dot{x} = \begin{bmatrix} -1.3 & 332 & 0.118 \\ -0.141 & -1.21 & -0.0389 \\ 0 & 1 & 0 \end{bmatrix} x + \begin{bmatrix} 81.9 \\ 39.5 \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0.092 & 0.09 & -0.023 \\ 0 & 1 & 0 \end{bmatrix} x \qquad (6.22)$$

The input u is the rudder deflection and the output y_1 is the vertical component of the aircraft velocity. The other output y_2, is a measured gyro signal. The poles of this model are

$-1.4 \cdot 10^{-3}$
$-1.25 \pm i \cdot 6.84$

The effects of the slow time-constant (≈ 10 minutes) is manually compensated for by the pilot, but the fast oscillatory modes need compensation to give a suitable aircraft performance. System specifications prescribe a desired step response with rise-time of about 1 (sec.) and a small over-shoot, for the channel between rudder and vertical velocity.

In Fig. 6.3, simulations of the step response is shown for some different choices of proportional control from y_2. It is seen that the controller

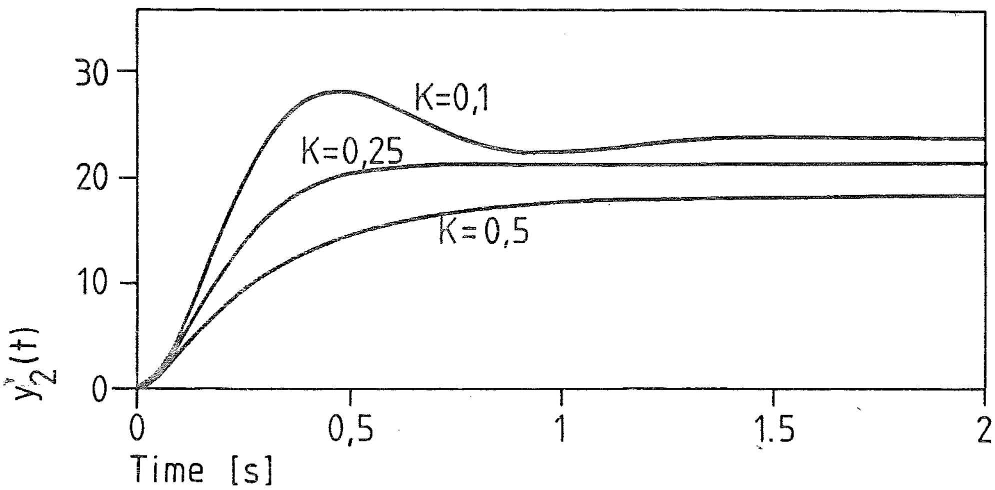


Fig. 6.3 - Step responses for the controlled model (6.22).

$$u = K(y_c - y_2) \tag{6.23}$$

with K = 0.25 gives a satisfactory result. Here y_c is an external command input.

It is known that there exists some poorly damped bending modes in the aircraft. These modes have characteristic frequencies in the interval between 10 and 100 (rad/sec.). A seventh order model of the pitch channel dynamics that includes the bending modes is

$$\dot{x} = \begin{bmatrix} -1.3 & 334 & 0.118 & 17.86 & 0.119 & 67.5 & -0.044 \\ -0.134 & -1.21 & -0.0389 & 5.72 & 0.068 & 10.5 & 0.042 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 5.22 & 14.4 & 0.143 & -3300 & 11.37 & -312 & 0.723 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1.74 & 3.86 & 0.121 & -10.4 & 0.074 & -5385 & -5.18 \end{bmatrix} x + \begin{bmatrix} 86.6 \\ 38.2 \\ 0 \\ 0 \\ -995 \\ 0 \\ -124 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0.11 & -0.3 & -0.026 & 21.0 & 0.03 & 84.3 & 0.054 \\ 0 & 1 & 0 & 0 & 0.045 & 0 & 0.09 \end{bmatrix} x \tag{6.24}$$

The poles of this model are

$-1.4 \cdot 10^{-3}$

$-1.22 \pm i \cdot 6.65$

$-1.28 \pm i \cdot 73.4$

$-1.94 \pm i \cdot 57.4$

To examine if the complicated model (6.24) is stable together with the controller (6.23) derived for the simple model, we utilize Theorem 4.1.

Introduce $G_m(s)$ and $G(s)$ as the transfer functions from u to y_2 for the models (6.22) and (6.24), respectively. The inequality (4.2) can now be written as

$$|G(s) - G_m(s)| < |1/K + G_m(s)| \qquad (6.25)$$

A Bode plot of these two functions is shown in Fig. 6.4. The nominal value K = 0.25 has been used.

Since both models are stable, and the inequality (6.25) apparently is satisfied on the Nyquist Contour Γ (the frequency axis), Theorem 4.1 gives that the complex model (6.24) is closed-loop stable.

In Fig. 6.4 it is seen that the two curves are close at the frequencies where G(s) has resonance peaks corresponding to the bending modes. This suggests that the closed-loop system might be oscillatory. Simulation of the step response for the model (6.24) together with some different proportional controllers are shown in Figs. 6.5 to 6.7. The simulations confirm the suspicion about poor control.

From Fig. 6.4 it is seen that if the curve corresponding to

$$\frac{1}{K} + G_m(s) \qquad (6.26)$$

is "raised" in the frequency interval where the resonance peaks are, the robustness can be improved. A practical way to achieve this can be to insert a "notch" filter in the feedback loop. Here we let it suffice to show a simulation of a step response for the complex model (6.24) when the controller uses a low-pass filtered measurement of the output y_2

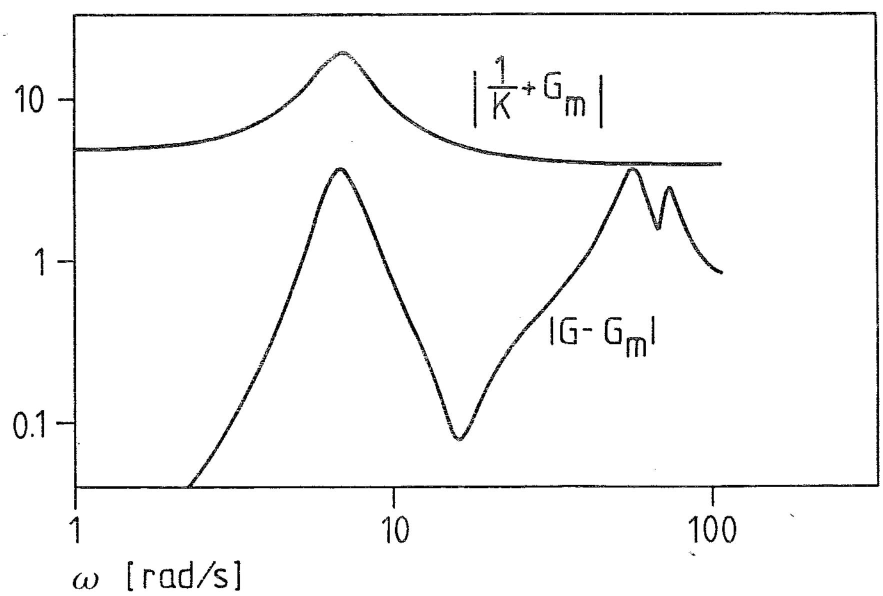
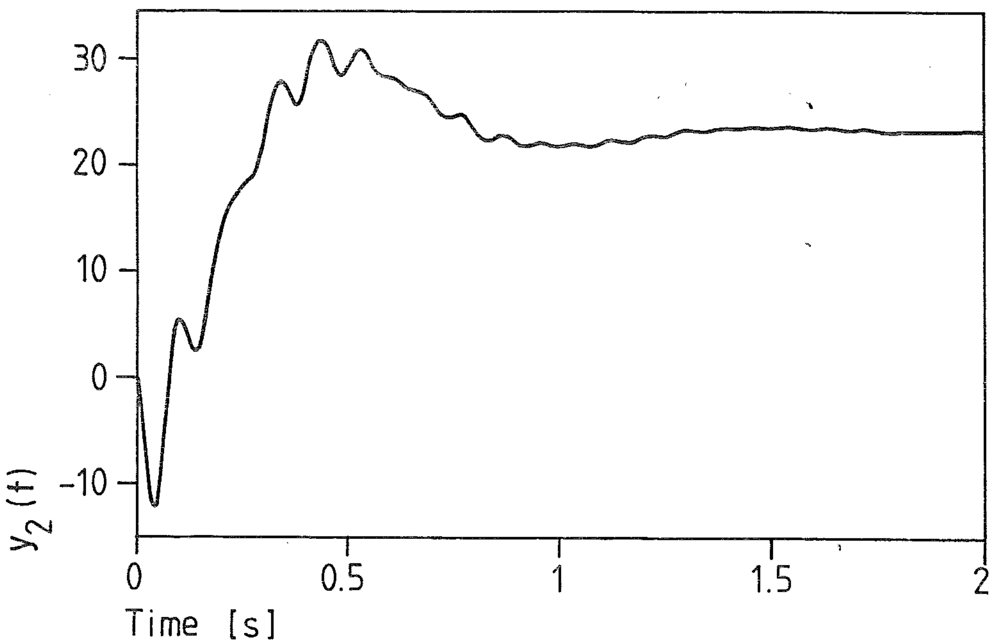
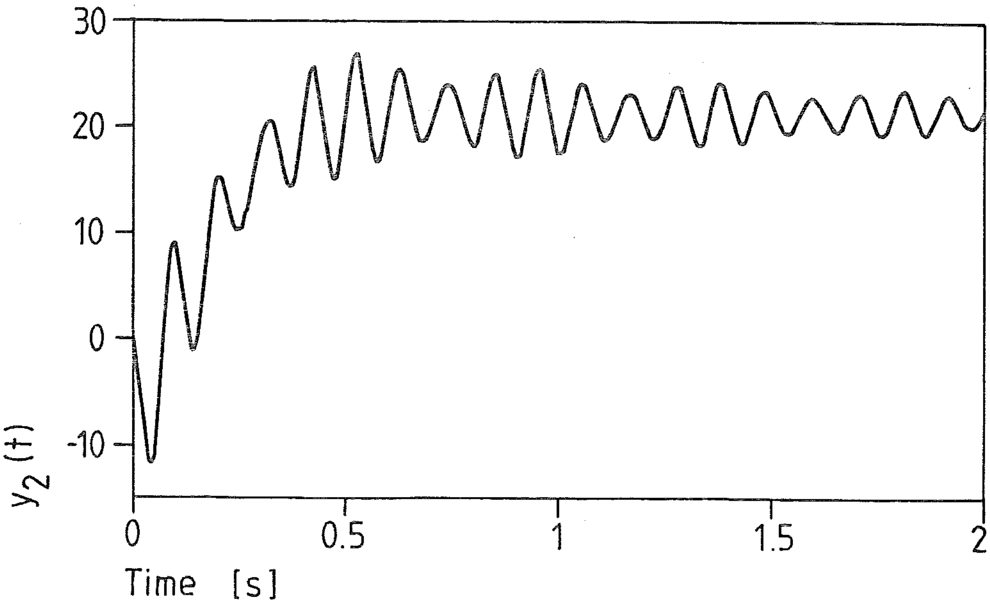


Fig. 6.4 - Bode plot of the two functions in (6.25).



<u>Fig. 6.5</u> - Step response when K = 0.1.



<u>Fig. 6.6</u> - Step response with nominal gain K = 0.25.

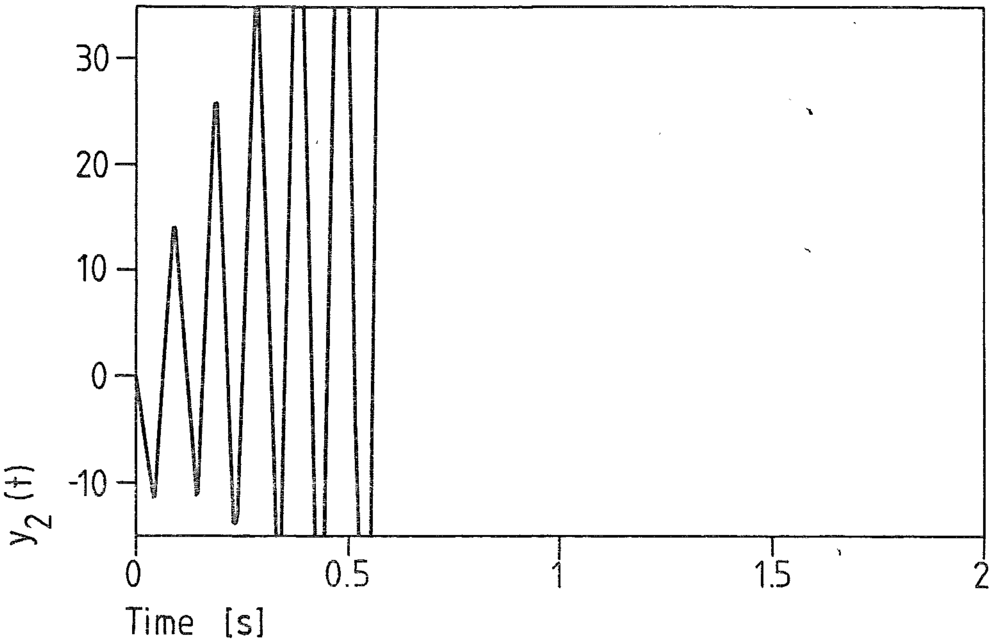


Fig. 6.7 - Unstable step response when K = 0.5.

$$u(s) = 0.25\left[y_c(s) - \frac{700}{s^2 + 50s + 700}\, y_2(s)\right] \qquad (6.27)$$

This simulation is shown in Fig. 6.8. An intuitive explanation of the acceptable result is that the low-pass filter in the feedback loop reduces the possibility of excitation of the resonant bending modes.

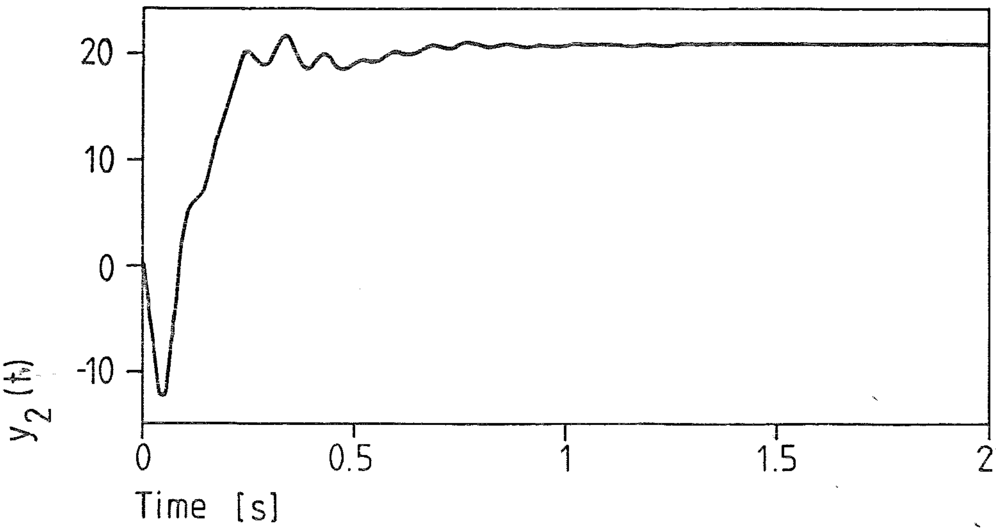


Fig. 6.8 - Step response with low-pass filtered feedback.

EXAMPLE 6.4

In this final example, we will present a design method for stable, multivariable systems. First some background on the subject is presented.

<u>Background 1</u>. In Richalet et al. (1978), IDCOM, a control package for computer control of stable multivariable systems is presented. The method can be viewed as multivariable <u>set-point control</u>, and is intended to be used in connection with ordinary PID-controllers.

The method starts with the off-line identification of a process model. This model is a truncated (finite length) discrete time impulse response, used as predictor for future system outputs. The control signals are computed by an algorithm that takes into account constraints on the control signal and possibly other auxiliary system variables.

The control package IDCOM is reported to have been successfully implemented on several industrial processes.

<u>Background 2</u>. Davison et al. (1980) introduced the concept of Tuning Regulator. It is a combination of feedforward and integrating feedback control, intended for stable multivariable continuous time processes.

The design method only requires a static model of the steady state gain-matrix of the process. The integration part of the controller is tuned manually, on line.

In the mentioned reference, the method is illustrated with a design example on a multivariable heat exchanger.

<u>A Design Method</u>. Consider a stable multivariable process in continuous time. It is assumed that the process exhibits desirable stability behaviour, and that the purpose of control is to keep the outputs at constant reference values. This design method is thus <u>not</u> intended as a "super"-design for systems that are hard to control. Slow disturbances like drifts and offsets are to be regulated to zero. The stability assumption is usually satisfied on industrial plants. There the different subsystems of the process are controlled by single-loop PID-controllers, giving a reasonable stability.

When sampling the process, it can be represented by the relation

$$y(t) = P(q)u(t) \qquad (6.28)$$

where u and y of dimension n, are the system input and output respectively, and q is the forward shift operator. The pulse transfer operator P(q) can be written in terms of the impulse response

$$P(q) = \sum_{i=1}^{\infty} P_i q^{-i} \tag{6.29}$$

Here P_i are constant nxn-matrices and q^{-1} is the backward shift operator.

We now follow the idea of Richalet et al. (1978) and use a truncation of the impulse response as a process model

$$P_m(q) = \sum_{i=1}^{N} P_{mi} q^{-i} \tag{6.30}$$

where P_{mi} is an estimate of P_i. The number of terms N to be retained in the model, depends on the sampling interval and the dominating time constants of the process. It is assumed that the steady state gain matrices P(1) and $P_m(1)$ of the process and model respectively, are nonsingular.

We now proceed to the controller. The controller structure is a generalization of the feedforward, integrating feedback controller by Davison et al. (1980). The control input u(t) is given by

$$u = P_m(1)^{-1} y_r + P_m(1)^{-1} [P_m(q)u - y] \tag{6.31}$$

or equivalently

$$[P_m(1) - P_m(q)]u = y_r - y \tag{6.32}$$

Here y_r is a n-dimensional reference value for the output y. The integral action of the controller makes it possible to achieve zero steady state error between y_r and y. Constant disturbances are also regulated to zero.

The model closed-loop system has the transfer matrix

$$H_m(z) = P_m(z)P_m(1)^{-1} \quad , \tag{6.33}$$

and is stable.

Robustness Analysis. We will now analyze the robustness of the real closed-loop system. Introduce the loop transfer matrices for the real and the model systems in series with the controller (6.32)

$$G(z) = P(z)[P_m(1) - P_m(z)]^{-1} \tag{6.34}$$

$$G_m(z) = P_m(z)[P_m(1) - P_m(z)]^{-1} \quad , \tag{6.35}$$

The return difference of the model system is given by

$$I + G_m(z) = I + P_m(z)[P_m(1) - P_m(z)]^{-1} =$$

$$= P_m(1)[P_m(1) - P_m(z)]^{-1} \tag{6.36}$$

According to Theorem 4.3, the real closed-loop system is stable if the inequality

$$\bar{\sigma}\{[P(z) - P_m(z)][P_m(1) - P_m(z)]^{-1}\} <$$

$$< \underline{\sigma}\{P_m(1)[P_m(1) - P_m(z)]^{-1}\} \tag{6.37}$$

is satisfied on the unit circle and at infinity (the Nyquist Contour Γ).

Since both $P(z)$ and $P_m(z)$ are assumed to be strictly proper (no direct feed-through), the inequality is always satisfied at infinity.

At the unit circle the analysis becomes more involved. In order to get expressions that are easier to verify, we will derive a stronger inequality that implies (6.37). Application of relation (4.29) on the left-hand side of (6.37) gives

$$\bar{\sigma}\{[P(z) - P_m(z)][P_m(1) - P_m(z)]^{-1}\} \le$$

$$\le \bar{\sigma}[P(z) - P_m(z)]/\underline{\sigma}[P_m(1) - P_m(z)] \qquad (6.38)$$

The inequality (6.37) is satisfied if the stronger condition

$$\bar{\sigma}[P(z) - P_m(z)] <$$

$$< \underline{\sigma}[P_m(1) - P_m(z)] \cdot \sigma\{\underline{P}_m(1)[P_m(1) - P_m(z)]^{-1}\} \qquad (6.39)$$

is satisfied on the unit circle.

For SISO systems this inequality (or (6.37)) reduces to

$$|P(z) - P_m(z)| < |P_m(1)| \qquad (6.40)$$

A special case of this (SISO) robustness result is reported in Åström (1979b).

We now illustrate the design method and the robustness of the controller on a multivariable example. Consider "Rosenbrock's system", Rosenbrock (1970), with transfer matrix

$$G(s) = \begin{bmatrix} \dfrac{1}{s+1} & \dfrac{2}{s+3} \\ \dfrac{1}{s+1} & \dfrac{1}{s+1} \end{bmatrix} \qquad (6.41)$$

Suppose this system is sampled with a sampling interval of h = 4 (secs.). Although the continuous time system (6.41) is famous for being difficult to control - it has a right half-plane zero - the long sampling interval has as a consequence that the sampled system is minimum-phase and rather easy to control. For simplicity, the impulse response model that is going to be used, contains only one term:

$$P_m(q) = \begin{bmatrix} 0.98 & 0.67 \\ 0.98 & 0.98 \end{bmatrix} \cdot q^{-1} \qquad (6.42)$$

In this case it is simple to calculate the right-hand side of the expression (6.39)

$$\underline{\sigma}[P_m(1) - P_m(z)] \cdot \underline{\sigma}\{P_m(1)[P_m(1) - P_m(z)]^{-1}\} =$$

$$= \underline{\sigma}[P_m(1)] = 0.167 \qquad (6.43)$$

The largest singular value for the model error is more difficult to calculate in a closed form, and we settle for an upper bound from the standard formula

$$(\bar{\sigma}[A])^2 \leq \|A\|_1 \cdot \|A\|_\infty \qquad (6.44)$$

which applied to (6.39) gives

$$\bar{\sigma}[P(z) - P_m(z)] = \bar{\sigma}\begin{bmatrix} \frac{0.98 \cdot 0.02}{z(z-0.02)} & \frac{4 \cdot 10^{-6}}{z \cdot z} \\ \frac{0.98 \cdot 0.02}{z(z-0.02)} & \frac{0.98 \cdot 0.02}{z(z-0.02)} \end{bmatrix} \leq$$

$$\leq \frac{2 \cdot 0.98 \cdot 0.02}{|z - 0.02|} \leq 0.04 \quad \text{for } |z| = 1 \qquad (6.45)$$

Consequently, the real closed-loop system is stable. It should be mentioned that the robustness properties grow poorer fast as the sampling interval is decreased. For shorter sampling intervals, the number of terms in the impulse response model P_m must be increased in order to maintain good control.

A simulation of the closed-loop system step response is shown in Fig. 6.9. During the transient behaviour there are rather large coupling effects from y_2 to y_1. The non-minimum phase characteristics of the continuous time system is

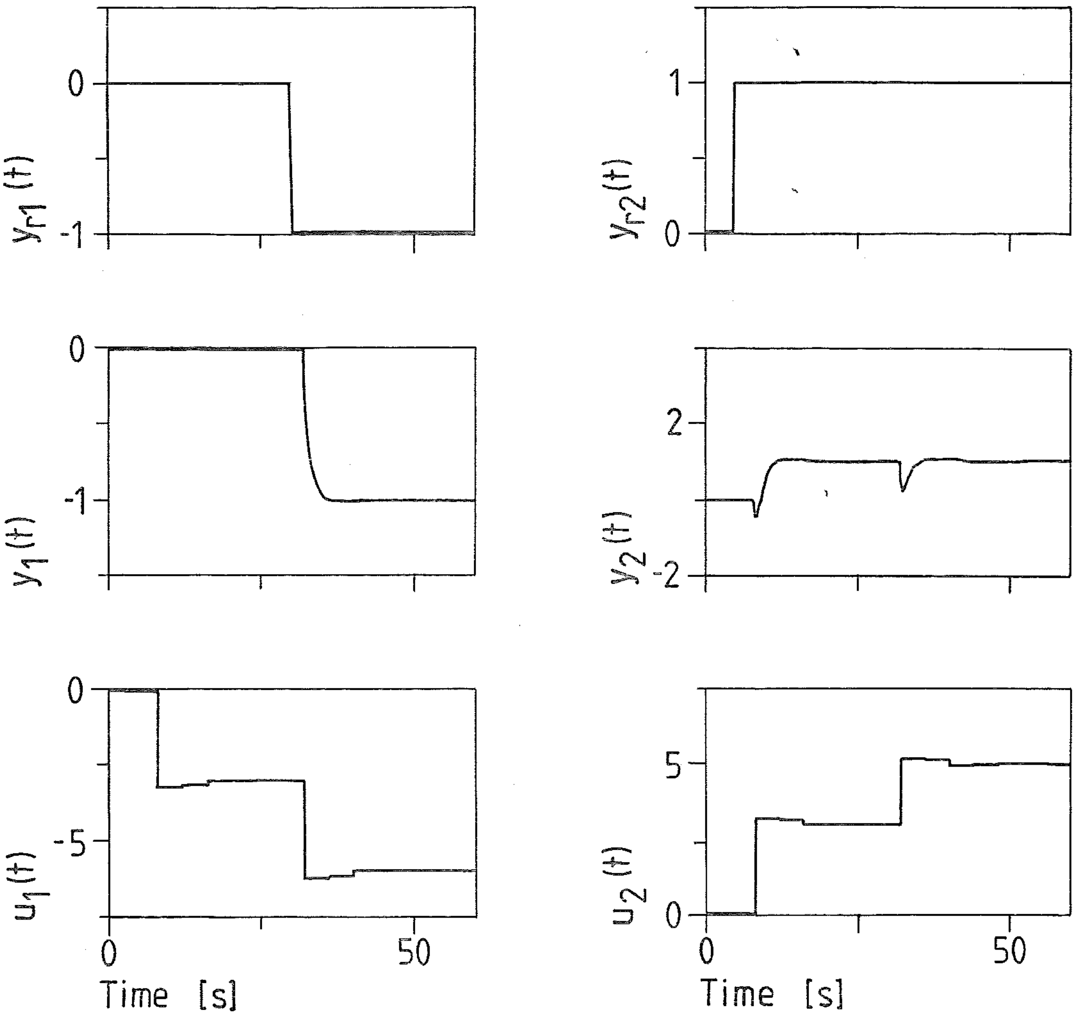


Fig. 6.9 - The closed-loop system step response.

clearly seen in y_1. The asymptotic behaviour of the system is according to the specifications, and quite acceptable.

7. CONCLUSIONS

In this part of the thesis, the robustness problem is treated for linear, time-invariant control systems, designed with uncertain process models. The main contributions are the new robustness results for systems with general perturbations. Both continuous time and discrete time systems are treated with the same formalism. The results are formulated in terms of closeness between the frequency responses of the loop transfer functions of the real process and the process model. Provided the Nyquist curves are close, it is possible to conclude closed-loop stability for the real process, if the model and the real system have either the same number of unstable poles, or the same number of non-minimum phase zeros. These two numbers are thus important aspects when modelling a process.

The closeness of the Nyquist curves are expressed with inequalities in the frequency domain. These robustness inequalities depend on the choice of the feedback controller, and imply certain design principles. The loop gain should be high for low frequencies, and low for high frequencies over the desired closed-loop bandwidth. The need for a good process model is greatest around the bandwidth frequency, where the robustness is poor. Thus there is an analytical confirmation of practical regulator design intuition.

A goal for further research, is to derive robustness results for systems controlled by adaptive regulators. Such results would provide insight into the construction of adaptive controller algorithms that guarantee closed-loop stability, even if the real processes are more complex than the model structure that the controller is designed for. The analysis can not be expected to be kept entirely in the frequency domain, since such systems contain nonlinearities with memory.

8. REFERENCES

Åström, K.J.(1979a): Robustness of a Design Method Based on Assignment of Poles and Zeros. Lund Inst. of Technology, Lund, Sweden. CODEN: LUTFD2/(TFRT-3153)/1-14/(1979).

Åström, K.J.(1979b): A robust sampled regulator for stable processes with monotone step responses. Lund Inst. of Technology, Lund, Sweden.
CODEN: LUTFD2/(TFRT-7173)/1-13/(1979).

Åström, K.J.(1980): Why use adaptive techniques for steering large tankers? Int. J. Control 32, 689-708.

Black, H.S.(1934): Stabilized Feedback Amplifiers. Bell System Techn. Jour. 13 (1934), 1-18.

Bode, H.W.(1945): Network Analysis and Feedback Amplifier Design. New York: van Nostrand 1945.

Davison, E.J., Taylor, P.A. and Wright, J.D.(1980): On the Application of Tuning Regulators to Control a Commercial Heat Exchanger. IEEE Trans. Aut. Control 25, 361-375.

Desoer, C.A. and Vidyasagar, M.(1975): Feedback Systems: Input-Output Properties. Academic Press, New York - San Francisco - London.

Desoer, C.A. and Wang, Y.-T.(1980): On the Generalized Nyquist Stability Criterion. IEEE Trans. Aut. Control 25, 187-196.

Doyle, J.C.(1978): Robustness of Multiloop Feedback Systems. 1978 IEEE CDC, 12-18.

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

Kalman, R.E.(1964): When is a linear control system optimal? Trans. ASME, Ser.D:J. Basic Eng., Vol.86, March 1964, 1-10.

Kwakernaak, H. and Sivan, R.(1972): Linear Optimal Control Systems. J. Wiley, New York.

Molander, P.(1979): Stabilisation of Uncertain Systems, Lund Inst. of Technology, Lund, Sweden. CODEN: LUTFD2/(TFRF-1020)/1-111/(1979).

Pernebo, L.(1981): An Algebraic Theory for the Design of Controllers for Linear Multivariable Systems, Part I and II. IEEE Trans. Aut. Control 26, To be published.

Richalet, J., Rault, A., Testud, J.L. and Papon, J.(1978): Model Predictive Heuristic Control: Application to industrial processes. Automatica 14, 413-429.

Rosenbrock, H.H.(1970): State-space and Multivariable Theory. Nelson, London.

Safonov, M.G.(1977): Robustness and Stability Aspects of Stochasic Multivariable Feedback System Design, PhD-thesis. Massachusetts Institute of Technology.

Sandell, N.R.Jr.(1979): Robust Stability of Systems with Application to Singular Perturbations. Automatica 15, 467- 470.

PART II - ESTIMATION OF SIMPLE PARAMETRIC MODELS FOR HIGH ORDER SYSTEMS

ABSTRACT

Linear parametric models, obtained by system identification methods, have found many applications in the design of control systems as well as in signal prediction and filtering. The estimated models practically always have lower order than any real system. It is therefore important to characterize the properties of estimated low order models, in order to obtain design principles for identification experiments.

We analyze the properties of low order parametric models obtained by the Least Squares method and the Stochastic Approximation method, when the identification input is a periodic signal. The analysis is carried out mainly in the frequency domain. It is shown that the estimated model and the system have the same transfer function at the frequencies of the input signal. We also give a characterization of the Least Squares estimate of the model parameters, when there are disturbances acting on the system. This characterization is given in the frequency domain, and involves the spectral densities of the input signal and the disturbance.

1. INTRODUCTION

In Automatic Control there are many situations where it is necessary to have a model of a system:

* Most design methods for controllers are based on a model of the system.

* Analysis of the properties of a system is usually carried out for a system model.

* Simulation of a process always demands a system model.

Systems are unfortunately generally described by nonlinear timevarying infinite-dimensional differential equations. Since such descriptions with few exceptions are impossible to work with, simplified models are used in practice.

There are many different ways to obtain simplified models. Depending on how, and under what circumstances the model is derived, it gives a more or less accurate description of the real system. It is therefore interesting and valuable to have some characterization of relations between the model and the real system.

For practical use, linear parametric models of rather low order have the greatest importance. They often allow simple analysis and simple design methods for controllers. A common way to obtain linear parametric models, is by system identification with a (recursive) algorithm of the Least Squares type. The parameters of the model then correspond to the minimum of a quadratic loss-function of prediction errors. The shape of the loss-function depends on the input-output data, and hence on the choice of input sequence. The parameters of the identified model therefore can be highly dependent on the properties of the input signal, as the following example illustrates.

EXAMPLE 1.1

Consider a continuous time system described by the transfer function

$$G(s) = \frac{(s+1)}{(s+0.1)(s+10)} \tag{1.1}$$

The system is sampled with a sample period of $T_s = 0.1$ s.

Introduce a first order discrete time model for the sampled system

$$G_m(q) = \frac{b}{q + a} \qquad (1.2)$$

The parameters a and b are estimated with the Least Squares method. During the identification experiment, the input signal to the system is a sinusoid with angular frequency ω rad/s.

In Table 1.1 the estimated models and their continuous time counterparts are listed for some different values of ω.

ω [rad/s]	Discrete Model	Cont. Model
0.1	$\frac{0.011}{q - 0.988}$	$\frac{0.11}{s + 0.12}$
1.0	$\frac{0.019}{q - 0.91}$	$\frac{0.20}{s + 0.94}$
10.0	$\frac{0.063}{q - 0.44}$	$\frac{0.93}{s + 8.3}$

Table 1.1 - The identified models and their continuous time counterparts.

To give a feeling for how the real system and the identified models are related, the Nyquist curves are shown in Figs. 1.1 to 1.3. It is seen that the frequency responses of the real transfer function and the continuous time models coincide at the respective frequency of the identification input. The simplified models thus are good approximations of the true system at the input frequency. Despite of this, the identified models obviously have quite different properties.

Details of this identification experiment are given in Appendix A. □

Here it is emphasized that before a (simple) model is derived, it should be considered for what purpose it is going to be used. The model should give a good description of the system in some sense, depending on the attempted use.

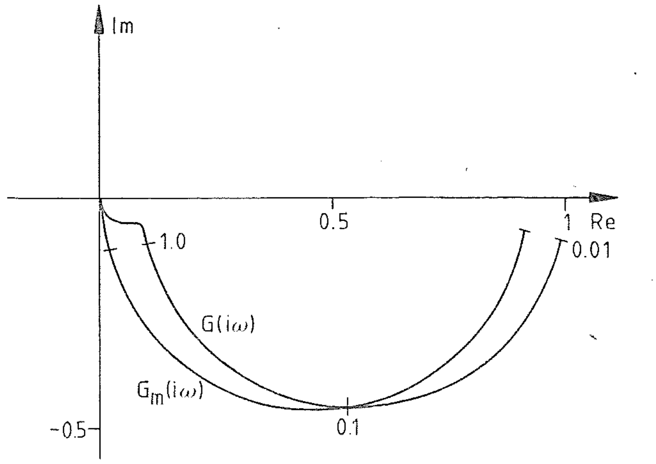


Fig. 1.1 - Nyquist curves for G and G_m when ω = 0.1.

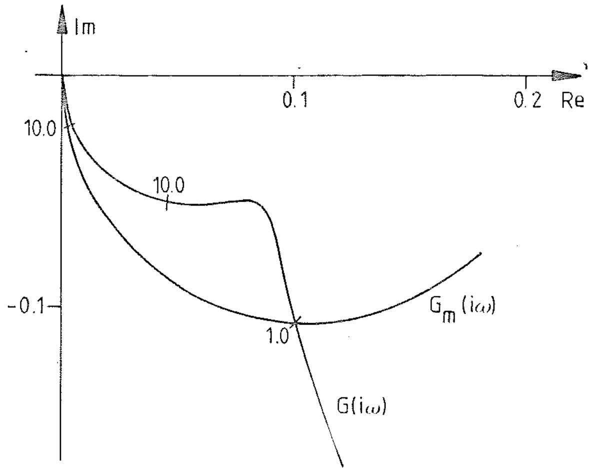


Fig. 1.2 - Nyquist curves for G and G_m when ω = 1.0.

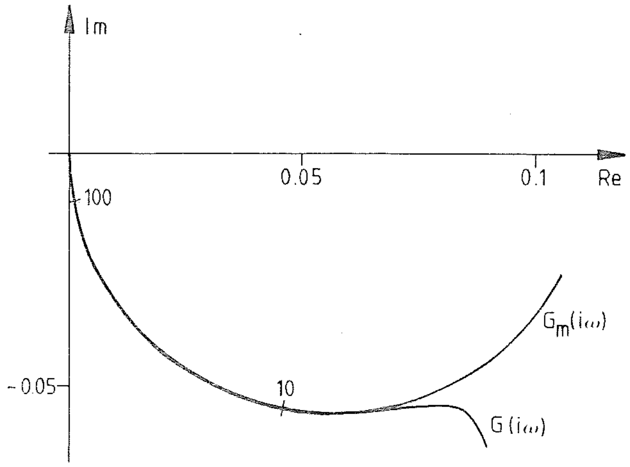


Fig. 1.3 - Nyquist curves for G and G_m when ω = 10.

It is interesting to see that the same philosophy could be used in the area of Model Reduction. Here the original high-order model is simplified to a new model of reduced order. It is desirable that the reduced order model should approximate the original model well in some sense.

In this part of the thesis, we investigate relations between the real system and an estimated model obtained from system identification. We primarily consider an open-loop identification experiment where the input signal consists of a sum of sinusoids. The number of sinusoids is related to the number of parameters in the model. It is shown that, provided a certain matrix related to the input-output data has full rank, there exists a unique member of the model class with a very interesting characterization. The unique model has the same value of the transfer function as the real system, at the frequencies of the input signal. When estimating the parameters of a model with the specified structure by some identification method, e.g. the Least Squares method or the Recursive Stochastic Approximation method, the estimated parameters converge to the parameters of the unique model. It is emphasized that this result is true even if the system is much more complex than the model structure.

This is an appealing characterization of the relation between the real system and an estimated model. The result immediately suggests a convenient way to obtain the frequency response of a system: Estimate the parameters of a model with some identification method, when the system input is a sum of sinusoids. The frequency response is computed by evaluating the transfer function of the estimated model at the respective frequencies of the input signal.

Another area of application for this result is in the design of robust control systems. In Part I of this thesis, it was shown that an important measure of robustness for a feedback system is that the Nyquist curves of the real system to be controlled and a system model used for design, are close. The previous characterization of estimated models give that these Nyquist curves are equal at the frequencies of the identification input. If the input contains sufficiently many sinusoids, the Nyquist curves will be close for <u>all</u> frequencies.

Still another application of the previously mentioned results is in the area of Model Reduction. We propose a method to obtain a reduced order model from a given high order system. The reduced order model has the same value of the transfer function as the given high order system at a certain number of frequencies, free to choose.

We also investigate the asymptotic properties of the Least Squares estimate of the model parameters, when the system identification input may be more complex than just a sum of sinusoids, and in addition the system may be subjected to unmeasurable disturbances. The result in this case is that the LS estimate minimizes a quadratic integral criterion, involving the spectral densities of the input signal and the disturbance.

Part II of the thesis is organized as follows. Chapter 2 is a short presentation of the general identification problem. In Chapter 3, we make assumptions on the system to be identified. The model class is specified, and we give a characterization of the input signal. An analysis of the output of the system leads to the characterization of a unique member of the model class, in Section 3.4. The results from various parameter estimation methods are analyzed in Chapter 4. In Chapter 5, we examine the properties of the asymptotic LS estimate, in case of general input signal and disturbances acting on the system. Some applications of the main results from Chapter 3 and 4, are presented in Chapter 6. Chapter 7 contains the conclusions, and the references are listed in Chapter 8. Finally, four appendices provide the details of an identification example, and the proofs of some theorems and lemmas.

2. THE PROBLEM

An identification problem is usually formulated using the following concepts, see Ljung et al.(1974).

$\mathcal{M}$ a class of models

$\mathcal{D}$ input-output data

$\mathcal{C}$ a criterion

The data $\mathcal{D}$ is obtained from experiments χ on a system S. The identification problem is then to find a model M in the class $\mathcal{M}$ so that the fit to the data $\mathcal{D}$ is the best possible according to the criterion $\mathcal{C}$.

A lot of investigations have been carried out under the assumption that

$$S \in \mathcal{M} \tag{2.1}$$

i.e. the system can be exactly described within the model class. This is of course a highly unrealistic assumption, since most processes are quite complex and the models in $\mathcal{M}$ are often based on strong simplifications.

One purpose of this study is to find out what happens when the model is much simpler than the system which generated the data.

3. ANALYSIS FOR PERIODIC INPUTS

In this chapter assumptions on the system, model and input signal are made. It will be shown that although the system can be much more complex than the model, for certain input signals the data can be described as generated by a member of the model class.

3.1 THE SYSTEM S

The system to be identified is assumed to be SISO, exponentially stable, time-invariant and linear. When sampled it can be described by the pulse transfer function G(q) in the shift operator q

$$y(t) = G(q)u(t) \tag{3.1}$$

where u(t) is the input and y(t) is the output. Equivalently, in the frequency domain with z-transformed variables

$$Y(z) = G(z)U(z) \tag{3.2}$$

The effect of disturbances acting on the system is assumed to be negligeable. With no loss of generality, time is normalized so that the sampling interval is unity. This simplifies notations in the sequel.

It is throughout this work implicitly assumed that the system (plant) is more complex than the low order models usually used in practical situations.

3.2 THE MODEL CLASS M

The model that we will use, is an ARMA prediction model with general structure

$$y(t) + a_1 y(t-1) + \ldots + a_k y(t-k) =$$

$$= b_1 u(t-d-1) + \ldots + b_\ell (t-d-\ell) \qquad (3.3)$$

The number of a and b parameters, k and ℓ respectively, as well as the number of time delays d, are chosen a priori to suit any particular application. The order of the model is $\max(k, d+\ell)$, and the number of parameters is $n = k + \ell$.

Associated with the model (3.3) is a model transfer function $G_m(z)$

$$G_m(z) = \frac{z^{-d}(b_1 z^{-1} + \ldots + b_\ell z^{-\ell})}{1 + a_1 z^{-1} + \ldots + a_k z^{-k}} \qquad (3.4)$$

Here the transfer function $G_m(z)$ is expressed as the quotient between two polynomials in the variable z^{-1}. This facilitates the analysis in the following, and it further reveals the close connection between the ARMA model (3.3) and the transfer function $G_m(z)$.

Models of this type are well suited for modern methods for system analysis and control design.

3.3 THE INPUT SIGNAL

The results from a system identification experiment, depend critically on the experimental conditions χ. We will here primarily consider open-loop experiments. Such an experiment is uniquely characterized by the input signal.

It is a well known fact that a periodic input signal with only one frequency, provides enough "excitation" on a two-parameter system to make possible estimation of the two parameters. This will be used in the following, in that the input will be just so rich in frequencies as to permit estimation of the number of parameters in the model.

For simplicity it will be assumed that the input signal is the sum of a number of sinusoids and possibly a constant. Such a signal can be generated as the output of linear, time-invariant, discrete time and autonomous system. Assume that the generating system is of order

$$n = k + \ell$$

equal to the number of parameters in the model. The system has distinct poles $\{z_i : i = 1,\ldots,n\}$ on the unit circle $\{z : |z| = 1\}$. Complex poles appear in conjugate pairs. The generated signal is composed of modes $u_i(t)$ corresponding to the poles of the system. The weighting of each mode $u_i(t)$ depends on the initial conditions of the system. The input signal thus has the frequency components $\{z_i\}$.

The signal can be written

$$u(t) = \sum_{i=1}^{n} u_i(t) \tag{3.5}$$

$$u_i(t) = c_i z_i^t \quad , \; i = 1,\ldots,n \tag{3.6}$$

or with z-transformed variables

$$U(z) = \sum_{i=1}^{n} U_i(z) \tag{3.7}$$

$$U_i(z) = \frac{c_i}{z - z_i} \tag{3.8}$$

The signal u(t) has the property that

$$\frac{1}{N} \sum_{t=t_0}^{t_0+N-1} \mu(t)\mu(t)^T > 0 \tag{3.9}$$

for all t_0 and $N \geq n$. Here

$$\mu(t)^T = [\ u(t-1)\ \dots\ u(t-n)\ ] \qquad (3.10)$$

The limit of (3.9) as time goes to infinity exists, and is independent of t_0. The property (3.9) implies that u(t) is <u>persistently exciting</u> (p.e.) of order n (but not of order higher than n); see e.g. Ljung (1971).

3.4 ANALYSIS OF THE SYSTEM OUTPUT

The system output can be written as

$$y(t) = G(q)u(t) \qquad (3.11)$$

if it assumed that the system is in equilibrium when the input signal is applied. This can also be expressed in the frequency domain

$$Y(z) = G(z)U(z) =$$

$$= G(z) \sum_{i=1}^{n} U_i(z) \qquad (3.12)$$

This is an analytic function which can be partitioned into two parts, having poles <u>on</u> the unit circle and <u>inside</u> the unit circle, respectively:

$$Y(z) = \sum_{i=1}^{n} G(z_i)U_i(z) + Y_{tr}(z) \qquad (3.13)$$

Here $Y_{tr}(z)$ is a z-transformed exponentially decaying transient.

This means that asymptotically, when the transient has vanished, the system output can be described by

$$y(t) = \sum_{i=1}^{n} G(z_i)u_i(t) \tag{3.14}$$

We therefore, for this special choice of input signal, have an input-output representation of the system in n complex numbers $\{G(z_i)\}$.

To get further, we introduce a matrix of second moments, associated with the model structure.

Define the matrix $R_{\varphi\varphi}$ of order $n = k + \ell$, by

$$R_{\varphi\varphi} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=t_0}^{N+t_0-1} \varphi(t)\varphi(t)^T \tag{3.15}$$

where

$$\varphi(t)^T = [-y(t-1) \; \ldots \; -y(t-k) \; u(t-d-1) \; \ldots \; u(t-d-\ell)] \tag{3.16}$$

Due to the periodicity of the input signal, the limit (3.15) is independent of the starting point $t = t_0$ of the summation. The choice $t_0 = 1$ is assumed in the sequel, unless otherwise is stated.

Remark 1. A necessary condition for $R_{\varphi\varphi}$ to have full rank n, is that the input is p.e. at least of order n. □

Remark 2. The matrix $R_{\varphi\varphi}$ is always of full rank when k = 0 and ℓ = n (i.e. there are no y-values in the φ-vector). This result follows directly from equation (3.9). In this case the model is a MA-process or equivalently, an impulse response model. □

Now, we state the following result:

THEOREM 3.1
There exists a unique model with transfer function

$$G_m(z) = \frac{z^{-d}(b_1 z^{-1} + \dots + b_\ell z^{-\ell})}{1 + a_1 z^{-1} + \dots + a_k z^{-k}} \qquad (3.17)$$

such that

$$G_m(z_i) = G(z_i) \qquad (3.18)$$

at the frequencies of the input

$$\{z_i \; ; \; i = 1,\dots,n\} \qquad (3.19)$$

if and only if the matrix $R_{\varphi\varphi}$ has full rank.

To prove this, we use the following lemma:

LEMMA 3.1
A necessary and sufficient condition for the set of equations

$$\begin{bmatrix} -G(z_1)z_1^{-1} \dots -G(z_1)z_1^{-k} & z_1^{-d-1} \dots z_1^{-d-\ell} \\ \vdots & \vdots \\ -G(z_n)z_n^{-1} \dots -G(z_n)z_n^{-k} & z_n^{-d-1} \dots z_n^{-d-\ell} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_k \\ b_1 \\ \vdots \\ b_\ell \end{bmatrix} = \begin{bmatrix} G(z_1) \\ \vdots \\ G(z_n) \end{bmatrix} \qquad (3.20)$$

to have a unique solution, is that the matrix $R_{\varphi\varphi}$ has full rank $n = k + \ell$.

Proof. See Appendix B.

Proof of Theorem 3.1. According to the Lemma, the system of equations (3.20) has a unique solution if and only if the matrix $R_{\varphi\varphi}$ has full rank n. Call this solution

$$\theta_0^T = [a_1 \ldots a_k\, b_1 \ldots b_\ell] \tag{3.21}$$

The equations in (3.20) can then be written as

$$(1 + a_1 z_i^{-1} + \ldots + a_k z_i^{-k}) G(z_i) =$$

$$= z_i^{-d} (b_1 z_i^{-1} + \ldots + b_\ell z_i^{-\ell}) \tag{3.22}$$

for i = 1,...,n and (3.18) follows. □

Remark 1. Notice that although the system G is stable, the unique model G_m characterized in Theorem 3.1 might be unstable! □

Remark 2. If the matrix $R_{\varphi\varphi}$ is singular, there exist infinitely many solutions to the system of equations (3.20). Consequently, in this case there are infinitely many models G_m satisfying $G_m(z_i) = G(z_i)$. In the following, we will avoid this situation by assuming that $R_{\varphi\varphi}$ is of full rank, an assumption that is further discussed in Section 4.4. □

Equation (3.13) together with Theorem 3.1 imply that, disregarded from an exponentially decaying transient that arises when the input initially is applied to the system, it is possible to describe the system input-output data as generated by a unique member of the model class (3.17), provided $R_{\varphi\varphi}$ has full rank:

$$y(t) = \sum_{i=1}^{n} G_m(z_i) u_i(t) + y_{tr}(t) \tag{3.23}$$

With suitable initial conditions in the unique model G_m this can also be written as

$$y(t) = G_m(q)u(t) + y_{tr}(t) \tag{3.24}$$

This fact is due to the special choice of input sequence, and is not true for more general input signals. In the next chapter we will utilize the result in Theorem 3.1 to characterize the estimates from various parameter identification algorithms.

4. IDENTIFICATION RESULTS

In the previous chapter we showed that for a special choice of input signal, the system could be described by a unique member of the model class, provided the matrix $R_{\varphi\varphi}$ had full rank. It will now be shown that parameter identification by minimization of a quadratic loss-function (criterion) of output prediction errors, yields an estimate that converges to the parameters of this unique model. It will also be shown that certain recursive gradient algorithms converge to the same parameters. The experimental conditions are the same as were described in the previous chapter.

4.1 PRELIMINARY

We will start by making an important assumption. This assumption is in fact necessary for the following results to be valid.

Assumption 4.1. It will in the sequel be assumed that the matrix $R_{\varphi\varphi}$ has full rank. This requirement is further discussed in Section 4.4.

Using Theorem 3.1 and Assumption 4.1, the system output can be described as generated by a unique model G_m

$$y(t) = \sum_{i=1}^{n} G_m(z_i)u_i(t) + y_{tr}(t) \qquad (4.1)$$

or according to equation (3.24)

$$y(t) = G_m(q)u(t) + y_{tr}(t) \qquad (4.2)$$

This relation can also be written in ARMA form

$$y(t) = - a_1 y(t-1) - \ldots - a_k y(t-k) +$$
$$+ b_1 u(t-d-1) + \ldots + b_\ell u(t-d-\ell) +$$
$$+ y_{tr}(t) + a_1 y_{tr}(t-1) + \ldots + a_k y_{tr}(t-k) \qquad (4.3)$$

Introduce

$$w(t) = y_{tr}(t) + a_1 y_{tr}(t-1) + \ldots + a_k y_{tr}(t-k) \qquad (4.4)$$

and the parameters of G_m in vector representation

$$\theta_0^T = [a_1 \ \ldots \ a_k b_1 \ \ldots \ b_\ell] \qquad (4.5)$$

Equation (4.3) can then be written in the compact form

$$y(t) = \varphi(t)^T \theta_0 + w(t) \qquad (4.6)$$

4.2 LEAST SQUARES IDENTIFICATION

Consider the problem of parameter identification by minimizing the criterion

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} [y(t) - \varphi(t)^T \theta]^2 \qquad (4.7)$$

with respect to the parameters θ.

We have the following result:

THEOREM 4.1
Assume that the matrix

$$\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi(t)^T \qquad (4.8)$$

has full rank.

Then the value of θ that minimizes $V_N(\theta)$ is given by the Least Squares (LS) estimate

$$\theta_{LS}(N) = \left[\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi(t)^T\right]^{-1}\left[\frac{1}{N}\sum_{t=1}^{N}\varphi(t)y(t)\right] \qquad (4.9)$$

which asymptotically converges to the parameters θ_O of the unique model G_m characterized in Theorem 3.1

$$\lim_{N\to\infty}\theta_{LS}(N) = \theta_O \qquad (4.10)$$

<u>Proof</u>. The first part of the theorem is a standard result, proven e.g. in Goodwin and Payne (1977). To prove result (4.10), recall that $\{y_{tr}(t)\}$ is an exponentially decaying sequence, and the same therefore also holds for $\{w(t)\}$.

Now, using (4.6) we obtain the following limit

$$\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)y(t) =$$

$$= \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi(t)^T\theta_O + \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)w(t) \qquad (4.11)$$

The asymptotic LS estimate therefore satisfies

$$\theta_{LS}(\infty) = \theta_0 + R_{\varphi\varphi}^{-1} \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \varphi(t)w(t) = \theta_0 \qquad (4.12)$$

since

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} |w(t)| = 0 \qquad (4.13)$$

and the vectors $\varphi(t)$ are bounded. This concludes the proof. □

EXAMPLE 4.1

Consider the system and models in Example 1.1. With the background of Theorems 3.1 and 4.1, it is straightforward to show that the system and model transfer functions are equal at the respective frequencies of the input signal. This was also seen from the Nyquist curves in Figs. 1.1 to 1.3. □

4.3 RECURSIVE ESTIMATION ALGORITHMS

In Theorem 4.1 we stated that the LS estimate minimized the quadratic loss-function $V_N(\theta)$. It is well known, that instead of computing the LS estimate off-line in a "batch mode", it can be computed recursively on line. The equations for the recursive LS estimate are

$$\begin{cases} \theta(t+1) = \theta(t) + P(t+1)\varphi(t)[y(t) - \varphi^T(t)\theta(t)] & (4.14a) \\ P(t+1) = P(t) - \dfrac{P(t)\varphi(t)\varphi^T(t)P(t)}{1 + \varphi^T(t)P(t)\varphi(t)} & (4.14b) \end{cases}$$

In the area of System Identification and Parameter Estimation, there are many recursive algorithms that resembles the recursive Least Squares method. Some of these methods can be described as gradient algorithms to minimize $V_N(\theta)$. A common method of this type is the Stochastic

Approximation Algorithm, for which we have the following convergence result:

THEOREM 4.2
Assume that the $R_{\varphi\varphi}$-matrix is of full rank. An identification algorithm of the form

$$\Theta(t+1) = \Theta(t) + \alpha(t)\varphi(t)[y(t) - \varphi(t)^T\Theta(t)] \qquad (4.15)$$

converges exponentially to the parameters of the unique model G_m, provided the sequence $\{\alpha(t)\}$ is non-negative and bounded. It is further assumed that there exists an $\varepsilon > 0$, such that

$$(1 - \alpha(t)\varphi(t)^T\varphi(t))^2 \leq 1 - \varepsilon\cdot\varphi(t)^T\varphi(t) \leq 1 \qquad (4.16)$$

<u>Proof</u>. According to equation (4.6), the system output is given by

$$y(t) = \varphi(t)^T\Theta_0 + w(t) \qquad (4.17)$$

where Θ_0 is the parameter vector of the unique model G_m. Introduce the parameter estimation error

$$\tilde{\Theta}(t) = \Theta(t) - \Theta_0 \qquad (4.18)$$

Subtracting Θ_0 from both sides of (4.15), and using (4.17), we obtain a recursive equation for the error:

$$\tilde{\Theta}(t+1) = [I - \alpha(t)\varphi(t)\varphi(t)^T]\tilde{\Theta}(t) + \alpha(t)\varphi(t)w(t) \qquad (4.19)$$

This is a linear, time-varying difference equation. For the homogenous part we have the following result:

LEMMA 4.1
The solution $\tilde{\Theta}(t)$ to the homogenous difference equation

$$\tilde{\theta}(t+1) = [I - \alpha(t)\varphi(t)\varphi(t)^T]\tilde{\theta}(t) \qquad (4.20)$$

is exponentially stable, provided the $R_{\varphi\varphi}$ -matrix is of full rank and $\alpha(t)$ satisfies relation (4.16).

Proof. See Appendix C. □

The solution to (4.19) can be written

$$\tilde{\theta}(t) = \Phi(t,t_0)\tilde{\theta}(t_0) + \sum_{k=t_0}^{t-1} \alpha(k)\Phi(t,k)\varphi(k)w(k) \qquad (4.21)$$

where $\Phi(t,t_0)$ is the transition matrix associated with the homogenous system (4.20). According to the lemma, the homogenous equation is exponentially stable. The norm of the transition matrix then satisfies

$$\|\Phi(t,t_0)\| \le C\cdot\lambda^{(t-t_0)} \qquad (4.22)$$

for some $C > 0$ and $\lambda < 1$.

From the assumptions in the theorem, the sequence $\{\alpha(t)\}$ is bounded. The sequence $\{\varphi(t)\}$ is bounded, since it consists of bounded inputs $\{u(t)\}$, and outputs $\{y(t)\}$ from an exponentially stable system G driven by $u(t)$.

The error $\tilde{\theta}(t)$, given by (4.21), is thus the state of an exponentially stable system, driven by an exponentially decaying input $\{w(t)\}$, and hence is an exponentially decaying sequence. □

Corollary. The result in Theorem 4.2 is valid for algorithms of the form

$$\theta(t+1) = \theta(t) + \alpha(t)K\varphi(t)[y(t) - \varphi(t)^T\theta(t)] \qquad (4.23)$$

where K is a positive definite and constant symmetric matrix, provided condition (4.16) is changed to

$$(1 - \alpha(t)\varphi(t)^TK\varphi(t))^2 \le 1 - \varepsilon\cdot\varphi(t)^TK\varphi(t) \le 1 \qquad (4.24)$$

Proof. The recursive equation for the parameter estimation error is

$$\tilde{\theta}(t+1) = [I - \alpha(t)K\varphi(t)\varphi(t)^T]\tilde{\theta}(t) + \alpha(t)K\varphi(t)w(t) \qquad (4.25)$$

Introduce the transformed variables

$$z(t) = K^{-1/2}\,\tilde{\theta}(t) \qquad (4.26)$$

$$\Psi(t) = K^{1/2}\,\varphi(t) \qquad (4.27)$$

Equation (4.25) can now be written as

$$z(t+1) = [I - \alpha(t)\Psi(t)\Psi(t)^T]z(t) + \alpha(t)\Psi(t)w(t) \qquad (4.28)$$

which is of the same form as equation (4.19). The rest of the proof is completely analogous to the proof of Theorem 4.2. □

Remark. A common choice of the gain $\alpha(t)$ is

$$\alpha(t) = \frac{1}{\beta + \varphi(t)^T\varphi(t)} \qquad (4.29)$$

where β is a small and positive constant. □

Notice that there is no assumption that the gain $\{\alpha(t)\}$ should tend to zero as time goes to infinity. The choice of a non-decreasing gain has importance for the convergence properties of the algorithm. The transient $\{y_{tr}(t)\}$ gives contribution to the data, that in general is inconsistent with the unique model θ_m, and hence the parameter estimate will initially be forced away from θ_0. In the Least Squares method, this "bad" information will always be present, and its weight will only decrease as 1/t, giving a rather slow convergence. In the Stochastic Approximation method with non-decreasing gain, the estimates will only be poor during the transient behaviour of the system, and the convergence rate is exponential.

The previous discussion is not relevant when the system is subjected to non-decreasing additive noise disturbances. In this case, the estimates will never converge unless the gain in the algorithm tends to zero.

4.4 THE RANK CONDITION ON $R_{\varphi\varphi}$

To derive the results of Theorems 4.1 and 4.2, we had to assume that the matrix $R_{\varphi\varphi}$ was of full rank. This is an assumption often made in the literature on System Identification. It for example ensures that the Least Squares estimate (4.9) exists.

In Söderström (1973), there is a short discussion and some results on the rank of $R_{\varphi\varphi}$ for different combinations of system and model order, and degree of excitation of the input signal. Suppose that the input signal is p.e. of order n (i.e. the number of parameters in the model). We have then three different cases:

i) The system is of lower order than the model. The matrix $R_{\varphi\varphi}$ is then singular, and there are infinitely many models G_m satisfying $G_m(z_i) = G(z_i)$, at the frequencies of the input signal.

2) The system has the same order as the model. The matrix $R_{\varphi\varphi}$ is then of full rank. There exists a unique model G_m such that $G_m(z_i) = G(z_i)$ for all choices of frequencies $\{z_i\}$ in the input signal.

3) The system is of higher order than the model. This is the situation implied in this work, and the most realistic one. In Söderström (1973), it is commented as "Nothing general can be stated", which is highly unsatisfactory. To see what a rank defect of the matrix $R_{\varphi\varphi}$ means in this case, recall Lemma 3.1. The $R_{\varphi\varphi}$-matrix has full rank if and only if the matrix

$$\begin{bmatrix} -G(z_1)z_1^{-1} & \cdots & -G(z_1)z_1^{-k} & z_1^{-d-1} & \cdots & z_1^{-d-\ell} \\ \vdots & & \vdots & \vdots & & \vdots \\ -G(z_n)z_n^{-1} & \cdots & -G(z_n)z_n^{-k} & z_n^{-d-1} & \cdots & z_n^{-d-\ell} \end{bmatrix} \qquad (4.30)$$

is nonsingular. Singularity of this matrix can be expressed as an algebraic equation in the matrix elements, by setting the determinant equal to zero. The sets of points $\{z_1,\ldots,z_n\}$ for which the matrix looses rank, forms an algebraic manifold, or hypersurfaces, in C^n.

Now, let n-2 of the z_i's be fixed on the unit circle, and the remaining two z_i's be chosen as z and 1/z. The system G is assumed to be exponentially stable, and therefore both G(z) and G(1/z) are analytical in regions of the complex plane, containing the unit circle. The determinant of (4.30), viewed as a function of the variable z, is thus analytic in a region of the complex plane containing the unit circle. The zeros of an analytic function are isolated, and consequently, the determinant of the matrix (4.30) can only have isolated zeros <u>on</u> the unit circle (some of these are the other fixed n-2 frequencies z_i). Moreover, if G(z) is rational, there can at most be a finite number of zeros.

We therefore have a nice characterization of frequencies where (4.30) is singular, or equivalently the $R_{\varphi\varphi}$-matrix looses rank: If the $R_{\varphi\varphi}$-matrix is singular for a given set of frequencies $\{z_i\}$ in the input signal, it suffices to change <u>one</u> of the frequencies (z_i and its complex conjugate) slightly, to regain the full rank property. In this sense, the $R_{\varphi\varphi}$-matrix is nonsingular for almost all choices of frequencies in the input, provided the system is not of lower order than the model.

We illustrate this with a simple example:

EXAMPLE 4.2

Consider an identification experiment on a system with the transfer function G(z), of order higher than one. The model has the transfer function

$$G_m(z) = \frac{bz^{-1}}{1 + az^{-1}} = \frac{b}{z + a} \qquad (4.31)$$

and the $\varphi(t)$-vector is

$$\varphi(t)^T = [-y(t-1)\ \ u(t-1)] \qquad (4.32)$$

In this example, the matrix (4.30) is

$$\begin{bmatrix} -G(z_1)z_1^{-1} & z_1^{-1} \\ -G(z_2)z_2^{-1} & z_2^{-1} \end{bmatrix} \qquad (4.33)$$

The matrix (4.33) is singular, or equivalently the $R_{\varphi\varphi}$-matrix is singular, if and only if

$$G(z_1) = G(z_2) \qquad (4.34)$$

We can separate two cases:

1) If $z_1 = 1$ and $z_2 = -1$, the $R_{\varphi\varphi}$-matrix is singular if and only if

$$G(1) = G(-1) \qquad (4.35)$$

2) If $z_1 = z$ and $z_2 = \bar{z}$, the $R_{\varphi\varphi}$-matrix is singular if and only if

$$\mathrm{Im}[G(z)] = 0 \qquad (4.36)$$

Equation (4.36) is satisfied when the Nyquist curve of G(z) crosses the real axis. This only happens a finite number of times for any reasonable system. □

In a practical identification experiment, it is possible to incorporate a rank test on the matrix $R_{\varphi\varphi}$. Such a test may also compute the conditioning number, thus indicating possible convergence troubles for the parameter estimates.

In previous works on System Identification, the characterization of the asymptotic LS estimate has mainly been in terms of its definition

$$\Theta_{LS} = R_{\varphi\varphi}^{-1} \cdot R_{\varphi y} \tag{4.37}$$

and in a stochastic framework. Since the result (4.37) is expressed in cross-products of time-functions, it is a time domain result. It is wellknown from many branches of applied mathematics, that it may be valuable to work in both time and frequency domains. Results tend to be more easily understood in one or the other. The results in this work, Theorems 4.1 and 4.2 together with Theorem 3.1, give a nice characterization of the asymptotic parameter estimates in the frequency domain: The identified model has the same transfer function as the real system at the frequencies of the input signal.

In the next chapter, we consider the situation when the input signal may be more general than just a sum of sinusoids. In addition, the system can be subjected to unmeasurable disturbances.

5. GENERAL INPUT SIGNALS

Estimation of low order parametric models with a more general class of input signals, will now be considered.

5.1 SYSTEM DESCRIPTION

The system to be identified is assumed to be exponentially stable, linear and time-invariant. It can be represented by the input-output relation

$$y(t) = G(q)u(t) + e(t) \tag{5.1}$$

where $\{u(t)\}$ is the input signal and $\{e(t)\}$ is a sequence of disturbances acting on the system.

The input is assumed to have existing second moments

$$r_u(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u(t)u(t+\tau) \quad ; \forall \tau \tag{5.2}$$

This choice of input class is convenient - it includes e.g. both ergodic stationary stochastic processes and purely deterministic signals, such as sinusoids and constants.

The disturbance $\{e(t)\}$ is assumed to have existing second moments

$$r_e(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} e(t)e(t+\tau) \quad ; \forall \tau \tag{5.3}$$

and is independent of the input u(t) in the sense that

$$r_{ue}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u(t)e(t+\tau) = 0 \quad ; \forall \tau \tag{5.4}$$

The input and disturbance signals have spectral densities given by

$$\Phi_u(\omega) = \sum_{\tau=-\infty}^{\infty} r_u(\tau)e^{-i\omega\tau} \tag{5.5}$$

$$\Phi_e(\omega) = \sum_{\tau=-\infty}^{\infty} r_e(\tau)e^{-i\omega\tau} \tag{5.6}$$

The input is further assumed to be persistently exciting of at least order n, i.e. the number of parameters in the model

$$G_m(z) = \frac{z^{-d}(b_1 z^{-1} + \dots + b_\ell z^{-\ell})}{1 + a_1 z^{-1} + \dots + a_k z^{-k}} \tag{5.7}$$

5.2 LEAST SQUARES ESTIMATION

According to Chapter 4, the LS estimate of the parameter vector of the model, is given by

$$\Theta_{LS}(N) = \left[\frac{1}{N}\sum_{t=1}^{N} \varphi(t)\varphi(t)^T\right]^{-1} \left[\frac{1}{N}\sum_{t=1}^{N} \varphi(t)y(t)\right] \tag{5.8}$$

Relation (5.2) for the input signal and the corresponding property for the disturbance, together with the stability of the system, give that the sequences $\{y(t)\}$ and $\{\varphi(t)\}$ also have existing second moments. The LS estimate then asymptotically satisfies

$$\lim_{N\to\infty} \Theta_{LS}(N) = R_{\varphi\varphi}^{-1} \cdot R_{\varphi y} \tag{5.9}$$

where $R_{\varphi\varphi}$ and $R_{\varphi y}$ are constant matrices. Here it has been assumed that the matrix $R_{\varphi\varphi}$ has full rank n.

Asymtotically, the prediction error $\{\varepsilon(t)\}$

$$\varepsilon(t) = y(t) - \varphi(t)^T\Theta_{LS}(\infty) \tag{5.10}$$

therefore is a sequence with existing second moment. The LS estimate $\Theta_{LS}(N)$ minimizes the criterion

$$V_N(\Theta) = \frac{1}{N}\sum_{t=1}^{N}[y(t) - \varphi(t)^T\Theta]^2 \tag{5.11}$$

As time goes to infinity, the minimum of this function becomes

$$\lim_{N\to\infty} V_N(\Theta_{LS}) = \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\varepsilon(t)^2 = r_\varepsilon(0) \tag{5.12}$$

Using Parcevals relation, (5.12) can be formulated in the frequency domain:

THEOREM 5.1
Asymptotically, the LS estimate minimizes the function

$$V_\infty(\Theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi}|G(e^{i\omega})A^*(e^{-i\omega}) - B^*(e^{-i\omega})|^2\,\Phi_u(\omega)d\omega +$$

$$+ \frac{1}{2\pi}\int_{-\pi}^{\pi}|A^*(e^{-i\omega})|^2\,\Phi_e(\omega)d\omega \tag{5.13}$$

with respect to the parameters Θ in the polynomials

$$A^*(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_k z^{-k} \tag{5.14}$$

$$B^*(z^{-1}) = z^{-d}[b_1 z^{-1} + \dots + b_\ell z^{-\ell}] \tag{5.15}$$

Proof. See Appendix D.

Remark 1. In the special situation when $\Phi_e(\omega) = 0$ and $\Phi_u(\omega)$ has a discrete spectrum with $n = k + \ell$ delta funtions

$$\Phi_u(\omega) = \sum_{j=1}^{n} \varrho_j \delta(e^{i\omega} - z_j) \tag{5.16}$$

i.e. the case treated in Chapters 3 and 4, the minimum of the criterion (5.13) becomes

$$r_\varepsilon(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{i\omega})A^*(e^{-i\omega}) - B^*(e^{-i\omega})|^2 \Phi_u(\omega)d\omega =$$

$$= \frac{1}{2\pi} \sum_{j=1}^{n} |G(z_j)A^*(z_j^{-1}) - B^*(z_j^{-1})|^2 \varrho_j = 0 \tag{5.17}$$

where it has been used that there exists a unique solution Θ_0 of coefficients in the polynomials $A^*(z^{-1})$ and $B^*(z^{-1})$, such that all n terms in (5.17) simultaneously are zero (Compare Lemma 3.1). This of course is the unique solution characterized in Theorem 3.1. □

Remark 2. Consider the asymptotic LS estimate of the model parameters, when there are no disturbances acting on the system:

$$\bar{\Theta}_{LS} = \left[R_{\bar{\varphi}\bar{\varphi}}\right]^{-1} \cdot R_{\bar{\varphi}\bar{y}} \tag{5.18}$$

The bar superscript ("−") indicates that there are no disturbances associated with the variables.

Suppose now that the system is subjected to a small disturbance signal $\sigma \cdot e(t)$, where $\{e(t)\}$ is a disturbance sequence with normalized power, and σ^2 is a factor that represents the inverted signal-to-noise ratio. The system output now becomes

$$y(t) = \bar{y}(t) + \sigma \cdot e(t) \tag{5.19}$$

and the regression vector

$$\varphi(t) = \bar{\varphi}(t) + \sigma \cdot \psi(t) \tag{5.20}$$

where

$$\psi(t)^T = [-e(t-1) \ldots -e(t-k)\ 0 \ldots 0] \tag{5.21}$$

The asymptotic LS estimate of the model parameter vector in this latter case, is given by

$$\theta_{LS} = R_{\varphi\varphi}^{-1} \cdot R_{\varphi y} = [R_{\bar{\varphi}\bar{\varphi}} + \sigma^2 \cdot R_{\psi\psi}]^{-1} [R_{\bar{\varphi}y} + \sigma^2 \cdot R_{\psi e}] \tag{5.22}$$

where it has been used that the input signal $\{u(t)\}$ and the unperturbed output $\{\bar{y}(t)\}$ are independent of the disturbance. For small values of σ^2 (large signal-to-noise ratios) the parameter estimate (5.22) approximately satisfies

$$\theta_{LS} \approx \bar{\theta}_{LS} + \sigma^2 \left[R_{\bar{\varphi}\bar{\varphi}} \right]^{-1} [R_{\psi e} - R_{\psi\psi} \bar{\theta}_{LS}] \tag{5.23}$$

This means that for large signal-to-noise ratios, the asymptotic LS parameter estimate will converge to approximately the same value as if the disturbance would have been zero, provided the matrix

$$R_{\bar{\varphi}\bar{\varphi}} = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{\infty} \bar{\varphi}(t) \bar{\varphi}(t)^T \tag{5.24}$$

is well conditioned. Such insensitivity to small disturbances is appealing, since in a practical identification experiment there is always measurement noise etc. corrupting the data. □

Here it should be mentioned that the asymptotic LS estimate (5.9) is only depending on the matrices of second moments, $R_{\varphi\varphi}$ and $R_{\varphi y}$. Totally different kinds of signals can give rise to the same matrices in the time limit. For finite times, however, the matrices in (5.8) will be highly dependent on the choice of input signal. The parameter estimate will have fast convergence at least if the matrices in (5.8) rapidly approaches their stationary values. This indicates that it may be advantageous to use deterministic signals such as sinusoids, with completely known behavior, rather than stochastic inputs which can have a rather dramatic behaviour during periods of time.

The way the asymptotic estimate is influenced by the input signal is clearly seen from (5.13). The reduced order model can be interpreted as an approximation of the true transfer function. The approximation is a weighted least squares approximation, and the weighting depends on the spectral density of the input signal.

Results similar to that of Theorem 5.1, have previously appeared in the literature, see e.g. Söderström (1973) and Solo (1978).

6. CONSEQUENCES AND APPLICATIONS

In the previous chapters we have characterized the estimated parameters of low order models, obtained by Least Squares identification. We will now proceed to various applications of these results.

6.1 CONTROLLER DESIGN

One of the most important and interesting aims in the area of Automatic Control, is to derive design methods for controllers. There are numerous (classical) design methods available, and each has its dedicated supporters. No matter which design method is used, it is based on the a priori knowledge of a model of the system to be controlled. The model can be either a transfer function or a state space model, depending on the design method.

If a controller is derived with some design method based on a model, and a closed-loop system is formed from the model and the controller, this system will behave as expected and system specifications will be fullfilled. However, in the real control situation, when the controller is controlling the real process, it is not at all clear what will happen. Relevant questions are: Is the closed-loop system stable? How does the system perform? The first basic requirement on a control system, naturally is that it should be stable. In the present context the question of stability is refered to as the "Robustness Problem". In Part I of this thesis, there are a number of results on robustness of stability. These results and the characterization of properties of estimated low order parametric models in Chapters 3 and 4, provide a possible tool for robust controller design.

A ROBUST DESIGN PROCEDURE

Consider the discrete time feedback system shown in Fig. 6.1, where $G_R(z)$ and $G(z)$ are the transfer functions of the controller and the real system, respectively.

Assume that the real system is exponentially stable, and that $G_m(z)$ is the transfer function of an exponentially stable model of the system. The parameters of the (low order) model $G_m(z)$ are estimated with the Least Squares method in an open-loop identification experiment on the real

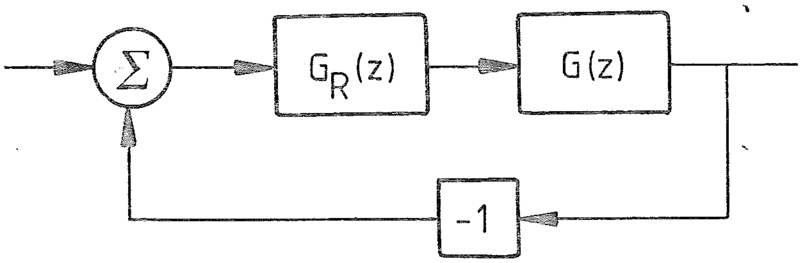


Fig. 6.1 - The feedback system.

system. The input sequence in the identification is chosen as a sum of sinusoids, according to Chapter 3.

From Theorems 3.1 and 4.1, we know that

$$G_m(z_i) = G(z_i) \quad ; \; i = 1,\dots,n \tag{6.1}$$

at the frequencies of the input $\{z_i\}$.

Suppose now that the controller is designed by some method, to appropriately control the model. The model closed-loop transfer function

$$H_m(z) = \frac{G_R(z)G_m(z)}{1 + G_R(z)G_m(z)} \tag{6.2}$$

then corresponds to a well-behaved and stable control system.

For the real closed-loop system, shown in Fig. 6.1, we have the following stability result from Theorem 4.1 of Part I of this thesis:

The real closed-loop system is asymptotically stable if

$$|G(z) - G_m(z)| < |1/G_R(z) + G_m(z)| \tag{6.3}$$

for all $|z| = 1$

Now, consulting the result (6.1) we see the interesting fact that the left-hand side of the inequality (6.3) is zero at the n frequencies $\{z_i\}$ of the identification input.

Since these frequencies are free to choose, inequality (6.3) is guaranteed to be satisfied around any n points $\{z_i\}$ on the unit circle.

If the frequencies are chosen in the range where the right-hand side of inequality (6.3) normally is small, e.g. around the desired closed-loop bandwidth frequency, it is reasonable that (6.3) is satisfied on the whole of the unit circle. This of course requires that the frequency response of the real system is "sufficiently smooth", and have no large resonance peaks. This in turn would imply that the real closed-loop system is asymptotically stable.

To illustrate the robustness of such a design, we give an example.

EXAMPLE 6.1

Consider the problem of controlling a continuous time system with the transfer function

$$G(s) = \frac{1}{(s + 1)^4} \tag{6.4}$$

A simulation of a step response for this system is shown in Fig. 6.2. The system will be controlled by a discrete time pole-placement controller, see Åström (1979).

Design specifications are of the servo system type, the closed-loop system should have a bandwidth of about $\omega = 0.8$ rad./sec., and a relative damping of $\zeta = 0.7$ giving a small over-shoot in a step response. The step response corresponding to these specifications has a rise-time of around 6 seconds, which gives at hand that a reasonable sample interval is $h = 0.6$ secs (10 samples on a step response transient).

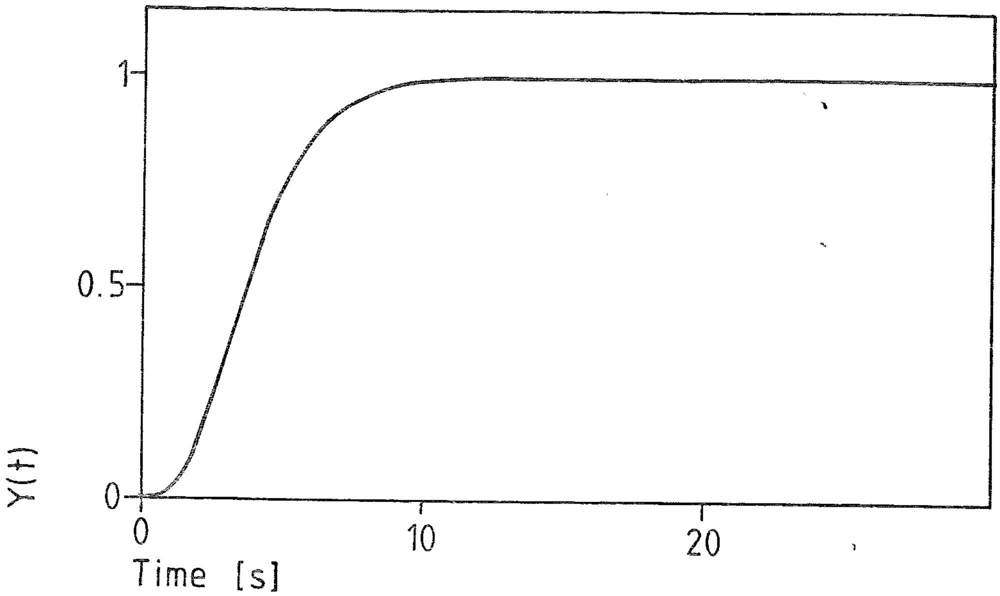


Fig. 6.2 - The open-loop system step response.

We settle for a second order model for the sampled system, having the transfer function

$$G_m(z) = \frac{B_m(z)}{A_m(z)} = \frac{b_1 z + b_2}{z^2 + a_1 z + a_2} \tag{6.5}$$

The closed-loop system, corresponding to the model (6.5) and the design specifications, has the transfer function

$$H_m(z) = \frac{b_1 z + b_2}{z^2 + p_1 z + p_2} \cdot \frac{1 + p_1 + p_2}{b_1 + b_2} \tag{6.6}$$

where the coefficients in the characteristic polynomial P(z) are given by

$$p_1 = -2\cos(\omega h\sqrt{1-\zeta^2})\exp(-\zeta\omega h) \tag{6.7}$$

$$p_2 = \exp(-2\zeta\omega h) \tag{6.8}$$

A controller, that together with the model (6.5) gives this closed-loop system is

$$R(z)u = -S(z)y + T(z)u_c \tag{6.9}$$

Here u_c is the command input. The polynomials R(z) and S(z) are derived from the polynomial equation

$$zP(z) = A_m(z)R(z) + B_m(z)S(z) \tag{6.10}$$

and T(z) is chosen so that the steady state gain from u_c to y, is equal to one.

We now proceed to the LS identification of the parameters of the reduced order model (6.5). According to the previous discussion, the input signal in the identification experiment is chosen as the sum of two sinusoids (corresponding to $\{z_1, \bar{z}_1, z_2, \bar{z}_2\}$). The angular frequencies are chosen as

$$\omega_1 = 0.2 \quad \text{rad/sec.}$$

$$\omega_2 = 0.8 \quad \text{rad/sec.}$$

This choise is motivated by:

ω_1 lies in the frequency interval of typical operation for the system. According to (6.1) this gives a good model for low frequencies.

ω_2 is equal to the desired closed-loop system bandwidth, where the return difference is usually small, indicating that the sensitivity may be poor. The inequality (6.3) is then guaranteed to be satisfied at this important point.

A LS identification experiment on the real sampled system, with the input signal as above, gives the parameters of the model

$a_1 = -1.567$

$a_2 = 0.636$

$b_1 = -0.0455$

$b_2 = 0.112$

Computation of the coefficients in the characteristic polynomial, and solving the polynomial equation (6.10) give the polynomials in the controller

$R(z) = z + 0.366$

$S(z) = 3.129z - 2.084$

$T(z) = 2.41z$

A simulation of a step response for the real process (6.4) together with the controller (6.9) is shown in Fig. 6.3. The control input u is shown in Fig. 6.4. The simulation shows that the control is satisfactory and that the design specifications are met.

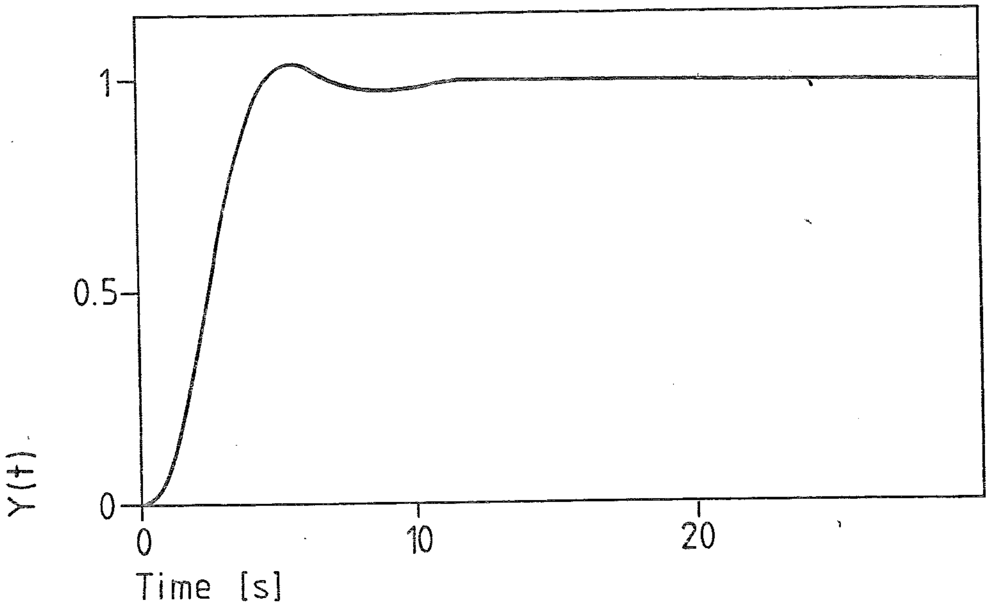


Fig. 6.3 - Step response for the real controlled system.

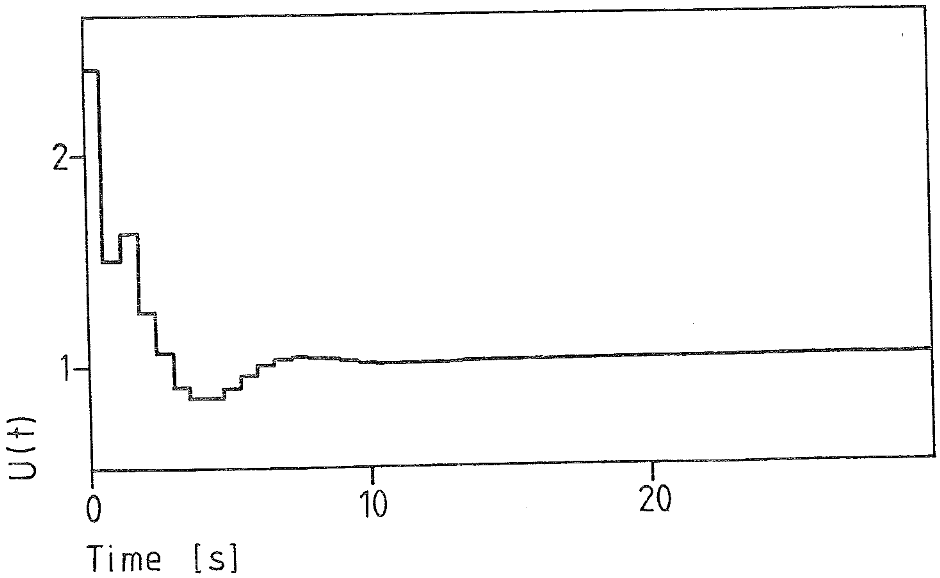


Fig. 6.4 - The control input for the controlled system.

6.2 MODEL REDUCTION

The Model Reduction Problem has attracted interest in the literature on Automatic Control for a long time. The aim is to find a simple model in order to facilitate system analysis and controller design.

The starting point is a completely known (linear) model of a system. By some means, a new reduced order model is derived from the original one. The problem is to choose a reduction method, such that the reduced order model approximates the original model as well as possible in some sense.

It is obvious that this problem is closely related to the problem of identification of low order models, discussed in Chapters 3 and 4. Use of Theorem 3.1 and Lemma 3.1 lead to the following model reduction procedure:

A MODEL REDUCTION PROCEDURE

Assume that a high order linear model with transfer function G(z) is given, and that the transfer function of a reduced order model is restricted to be of the form

$$G_m(z) = \frac{z^{-d}(b_1 z^{-1} + \dots + b_\ell z^{-\ell})}{1 + a_1 z^{-1} + \dots + a_k z^{-k}} \tag{6.11}$$

The parameters of this reduced order model are given by the unique solution (if it exists) to the set of linear equations

$$\begin{bmatrix} -G(z_1)z_1^{-1} \dots -G(z_1)z_1^{-k} & z_1^{-d-1} \dots z_1^{-d-\ell} \\ \vdots & \vdots \\ -G(z_n)z_n^{-1} \dots -G(z_n)z_n^{-k} & z_n^{-d-1} \dots z_n^{-d-\ell} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_k \\ b_1 \\ \vdots \\ b_\ell \end{bmatrix} = \begin{bmatrix} G(z_1) \\ \vdots \\ G(z_n) \end{bmatrix} \tag{6.12}$$

The reduced order model has the property, that

$$G_m(z_i) = G(z_i) \quad ; \ i = 1,\ldots,n \tag{6.13}$$

where

$$\{z_i : i = 1,\ldots,n\} \tag{6.14}$$

is a set of distinct numbers on the unit circle. Complex numbers appear in conjugated pairs.

<u>Remark</u>. This method works for continuous time systems as well. The numbers $\{z_i\}$ then lie on the imaginary axis. □

The method thus consists of equalizing the Nyquist curve of the reduced order model, at n points, to the Nyquist curve of the original model. If the points $\{z_i\}$ are chosen in the frequency range where the model is going to be used, the reduced order model will no doubt be a good approximant of the original model.

One way to practically solve the set of equations (6.12), is to perform an identification experiment on the original model G, designed according to Chapters 3 and 4.

According to Remark 1 of Theorem 3.1, the reduced order model G_m may be unstable even if the original model G is stable. This feature of a model reduction method is in general unacceptable, and hence there have to be some stability test incorporated to guarantee the stability of G_m.

7. CONCLUSIONS

In this part of the thesis, we analyze the properties of estimated low order models obtained from open-loop identification experiments. It is throughout the work implicitly assumed that all real systems are more complex than any estimated model. The main contribution is a characterization in the frequency domain of the properties of the low order models. When the identification input consists of a sum of sinusoids, the real system and the estimated model have the same value of the transfer functions at the frequencies of the sinusoids. This type of time series analysis thus turns out to be equivalent to a frequency response method. The fact that the transfer functions of the system and the estimated model are equal at a number of frequencies, suggests that the Nyquist curves may be close for all frequencies. According to Part I of this thesis, the model may therefore be used in the design of robust control systems.

We also consider the situation when the identification input is a general signal and the system is subjected to noise disturbances. The asymptotic LS estimate is characterized in the frequency domain by an integral criterion, involving the spectral densities of the input signal and the disturbance.

Further research in this area, is to derive new useful relations between estimated low order models and the identified systems. It would also be interesting to extend the present results to identification experiments in closed-loop. This may provide tools for the design of on-line tuning devices for simple regulators, and possibly new methods for analysis of self-tuning controllers.

8. REFERENCES

Åström, K.J.(1979): Simple Self-Tuners I. Lund Inst. of Technology, Lund, Sweden.
CODEN: LUTFD2/(TFRT-7184)/1-063/(1979).

Goodwin, G.C. and Payne, R.L.(1977): Dynamic System Identification: Experiment Design and Data Analysis. Academic Press, New York - San Francisco - London.

Ljung, L.(1971): Characterization of the Concept of "Persistently Exciting" in the Frequency Domain. Lund Inst. of Technology, Lund, Sweden.
CODEN: Report 7119. 1971.

Ljung, L., Gustavsson, I. and Söderström, T.(1974): Identification of Linear, Multivariable Systems Operating under Linear Feedback Control.
IEEE Trans. Aut. Control 19, 836-840.

Söderström, T.(1973): Uniqueness and On-line Algorithms in Identification of Linear Dynamic Systems, PhD-thesis. Lund Inst. of Technology, Lund, Sweden.
CODEN: LUTFD2/(TFRT-1004)/1-292/(1973).

Solo, V.(1978): Time Series Recursion and Stochastic Approximation, PhD-thesis. Australian National University, Australia 1978.

Wieslander, J.(1980): Idpac Commands - User's Guide. Lund Inst. of Technology, Lund, Sweden.
CODEN: LUTFD2/(TFRT-3157)/1-108/(1980).

Wittenmark, B. and Bar-Shalom, Y.(1979): Model Validation from Estimated Closed Loop Performance. Preprints 5th IFAC Symposium, Darmstadt, Federal Republic of Germany, 1103-1109.

APPENDIX A

At the Department of Automatic Control at Lund Institute of Technology, Lund, there is a programming package for computer-aided data analysis, Wieslander (1980). This package, IDPAC, includes among many other routines the possibilities to generate sinusoids, to estimate model parameters with the Least Squares method and to simulate the output signal from dynamical systems.

The identification experiment in Example 1.1 was thus performed by first generating the sinusoidal input signal sequences. These sequences were of the length 200 seconds, and consisted of 2000 samples. The system output sequences were generated by simulation of a discrete time version of the "real" system (1.1)

$$G(z) = \frac{10^{-2}(6.655z - 6.023)}{z^2 - 1.358z + 0.364} \qquad (A.1)$$

The estimates of the model parameter vector were then computed with the Least Squares method.

The continuous time counterparts of the estimated discrete time models have been derived with standard formulas.

APPENDIX B

Proof of Lemma 3.1. The matrix $R_{\varphi\varphi}$ is given by

$$R_{\varphi\varphi} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \varphi(t)\varphi(t)^T \tag{B.1}$$

where

$$\varphi(t)^T = [-y(t-1) \;\ldots\; -y(t-k) \;\; u(t-d-1) \;\ldots\; u(t-d-\ell)] \tag{B.2}$$

The system is described by the relation

$$y(t) = \sum_{i=1}^{n} G(z_i)u_i(t) + y_{tr}(t) \tag{B.3}$$

Using this, the $\varphi(t)$-vector can be written as

$$\varphi(t) = \begin{bmatrix} -G(z_1)z_1^{-1} & \ldots & -G(z_n)z_n^{-1} \\ \vdots & & \vdots \\ -G(z_1)z_1^{-k} & \ldots & -G(z_n)z_n^{-k} \\ z_1^{-d-1} & \ldots & z_n^{-d-1} \\ \vdots & & \vdots \\ z_1^{-d-\ell} & \ldots & z_n^{-d-\ell} \end{bmatrix} \begin{bmatrix} u_1(t) \\ \vdots \\ \vdots \\ u_n(t) \end{bmatrix} + \begin{bmatrix} -y_{tr}(t-1) \\ \vdots \\ -y_{tr}(t-k) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{B.4}$$

or in a more compact form

$$\varphi(t) = \Phi U(t) + \Psi(t) \tag{B.5}$$

Now, the $R_{\varphi\varphi}$-matrix can be expressed as

$$R_{\varphi\varphi} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} [\Phi U(t) + \Psi(t)][U(t)^{*} \Phi^{*} + \Psi(t)^{T}] \qquad (B.6)$$

where * means complex conjugated transpose. Since the elements of $\{\Psi(t)\}$ are exponentially decaying functions, the terms in (B.6) involving $\Psi(t)$ give no contribution to $R_{\varphi\varphi}$.

We obtain

$$R_{\varphi\varphi} = \lim_{N\to\infty} \frac{1}{N} \Phi \sum_{t=1}^{N} U(t)U(t)^{*} \Phi^{*} \qquad (B.7)$$

The matrix

$$\frac{1}{N} \sum_{t=1}^{N} U(t)U(t)^{*} \qquad (B.8)$$

has full rank n for all $N \geq n$, because of the special choice of input signal.

Hence, from (B.7) we see that there is equivalence between the full rank of $R_{\varphi\varphi}$ and the invertibility of the matrix Φ. □

APPENDIX C

Proof of Lemma 4.1. The norm of the error $\tilde{\theta}$ satisfies

$$\|\tilde{\theta}(t+1)\|^2 = \|\tilde{\theta}(t)\|^2 -$$

$$- \alpha(t)[2 - \alpha(t)\varphi(t)^T\varphi(t)]\ \tilde{\theta}(t)^T\varphi(t)\varphi(t)^T\tilde{\theta}(t) \tag{C.1}$$

By subtracting 1 from both sides of (4.16), we obtain that

$$- \alpha(t)[2 - \alpha(t)\varphi(t)^T\varphi(t)] \le -\varepsilon \tag{C.2}$$

Use of relation (C.2) in equation (C.1) gives the inequality

$$\|\tilde{\theta}(t+1)\|^2 \le \|\tilde{\theta}(t)\|^2 - \varepsilon\ \tilde{\theta}(t)^T\varphi(t)\varphi(t)^T\tilde{\theta}(t) \tag{C.3}$$

The norm of the estimation error $\tilde{\theta}(t)$ is thus monotone and non-increasing. Summing up N steps of this equation from an arbitrary time t_0, gives

$$\|\tilde{\theta}(t_0+N)\|^2 \le \|\tilde{\theta}(t_0)\|^2 -$$

$$- \varepsilon\ \tilde{\theta}(t_0)^T \sum_{t=t_0}^{t_0+N-1} \Phi(t,t_0)^T\varphi(t)\varphi(t)^T\Phi(t,t_0)\ \tilde{\theta}(t_0) \tag{C.4}$$

where $\Phi(t,t_0)$ is the transition matrix associated with the homogenous system (4.20)

$$\Phi(t,t_0) = \begin{cases} I & \text{if } t = t_0 \\ \prod_{k=t_0}^{t-1} [I - \alpha(k)\varphi(k)\varphi(k)^T] & \text{if } t > t_0 \end{cases} \tag{C.5}$$

According to Assumption 4.1, the $R_{\varphi\varphi}$-matrix is of full rank, and therefore positive definite. Due to the periodicity of the input signal $\{u(t)\}$, the following holds for all t_0 and some $\sigma > 0$

$$R_{\varphi\varphi} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=t_0}^{t_0+N-1} \varphi(t)\varphi(t)^T \geq \sigma \cdot I \qquad (C.6)$$

Therefore, to all t_0 there exists an $N \geq n$ independent of t_0, such that

$$\frac{1}{N} \sum_{t=t_0}^{t_0+N-1} \varphi(t)\varphi(t)^T \geq \frac{1}{2} \cdot \sigma \cdot I \qquad (C.7)$$

and consequently

$$\sum_{t=t_0}^{t_0+N-1} \varphi(t)\varphi(t)^T = P(t_0)P(t_0)^T \geq \frac{N}{2} \cdot \sigma \cdot I \qquad (C.8)$$

The matrix $P(t_0)$ is given by

$$P(t_0) = [\varphi(t_0) \; . \; . \; . \; \varphi(t_0+N-1)] \qquad (C.9)$$

Introduce the vectors $v(t)$

$$v(t) = \Phi(t,t_0)^T \varphi(t) \qquad ; \; t \geq t_0 \qquad (C.10)$$

and the matrix $V(t_0)$

$$V(t_0) = [v(t_0) \; . \; . \; . \; v(t_0+N-1)] \qquad (C.11)$$

This matrix satisfies

$$V(t_0)V(t_0)^T = \sum_{t=t_0}^{t_0+N-1} \Phi(t,t_0)^T\varphi(t)\varphi(t)^T\Phi(t,t_0) \tag{C.12}$$

The matrices $V(t_0)$ and $P(t_0)$ are related by

$$V(t_0) = P(t_0)\cdot T(t_0) \tag{C.13}$$

where

$$T(t_0) = \begin{bmatrix} 1 & & & \\ & \ddots & t_{ij}(t_0) & \\ & & \ddots & \\ 0 & & & \ddots \\ & & & 1 \end{bmatrix} \tag{C.14}$$

The elements $t_{ij}(t_0)$ have no simple structure, but they can easily be computed from (C.9) and (C.10). For example

$$t_{12}(t_0) = -\alpha(t_0)\varphi(t_0)^T\varphi(t_0+1) \tag{C.15}$$

Both $\{\alpha(t)\}$ and $\{\varphi(t)\}$ are (uniformly) bounded sequences, and hence the elements $t_{ij}(t_0)$ are uniformly bounded with respect to t_0 for fixed N. The matrix $T(t_0)$ is invertible and the inverse has the structure

$$T(t_0)^{-1} = \begin{bmatrix} 1 & & & \\ & \ddots & s_{ij}(t_0) & \\ & & \ddots & \\ 0 & & & \ddots \\ & & & 1 \end{bmatrix} \tag{C.16}$$

where the $s_{ij}(t_0)$ are uniformly bounded with respect to t_0 for fixed N. We obtain a bound on the minimum singular value of $T(t_0)$:

$$\sigma_{min}[T(t_0)] = \frac{1}{\sigma_{max}[T(t_0)^{-1}]} \geq \frac{1}{\|T(t_0)^{-1}\|_F} \geq \zeta > 0 \qquad (C.17)$$

for some ζ only depending on N. Here $\|A\|_F = [\Sigma|a_{ij}|^2]^{1/2}$ is the Frobenius norm.

Consequently,

$$\sum_{t=t_0}^{t_0+N-1} \Phi(t,t_0)^T\varphi(t)\varphi(t)^T\Phi(t,t_0) = V(t_0)V(t_0)^T =$$

$$= P(t_0)T(t_0)T(t_0)^T P(t_0)^T \geq \zeta^2 P(t_0)P(t_0)^T =$$

$$= \zeta^2 \sum_{t=t_0}^{t_0+N-1} \varphi(t)\varphi(t)^T \geq \frac{\zeta^2 \cdot \sigma \cdot N}{2} \cdot I \qquad (C.18)$$

The inequality relations in (C.18) refer to the smallest eigenvalue of the respective matrices.

Using relation (C.18) in inequality (C.4), we obtain

$$\|\tilde{\theta}(t_0+N)\|^2 \leq [1 - \varepsilon\zeta^2\sigma N/2]\|\tilde{\theta}(t_0)\|^2 \qquad (C.19)$$

The parameter estimation error $\{\tilde{\theta}(t)\}$ is thus exponentially decreasing to zero, and the transition matrix satisfies

$$\|\Phi(t,t_0)\| \leq C\cdot\lambda^{(t-t_0)} \qquad (C.20)$$

for some $C > 0$ and $\lambda < 1$. □

APPENDIX D

Proof of Theorem 5.1. Consider the loss-function (5.11) in the time limit

$$V_{\infty}(\theta) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} [y(t) - \varphi(t)^T\theta]^2 \tag{D.1}$$

Let the parameter vector θ have the structure

$$\theta^T = [a_1 \ \dots \ a_k \ b_1 \ \dots \ b_\ell] \tag{D.2}$$

Using the definition of $\varphi(t)$, we obtain

$$V_{\infty}(\theta) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} [A^*(q^{-1})y(t) - B^*(q^{-1})u(t)]^2 \tag{D.3}$$

where

$$A^*(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_k q^{-k} \tag{D.4}$$

$$B^*(q^{-1}) = q^{-d}[b_1 q^{-1} + \dots + b_\ell q^{-\ell}] \tag{D.5}$$

The output $\{y(t)\}$ is generated as

$$y(t) = G(q)u(t) + e(t) \tag{D.6}$$

where $\{u(t)\}$ and $\{e(t)\}$ are independent in the sense of (5.4). Hence

$$V_{\infty}(\Theta) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \{[G(q)A^{*}(q^{-1}) - B^{*}(q^{-1})]u(t)\}^{2} +$$

$$+ \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} [A^{*}(q^{-1})e(t)]^{2} \qquad (D.7)$$

Using equations (5.2) and (5.3), this limit can be written as a linear combination of $r_u(\tau)$'s and $r_e(\tau)$'s. These functions can be expressed in their corresponding spectral densities:

$$r_u(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) e^{i\omega\tau} d\omega \qquad (D.8)$$

and

$$r_e(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_e(\omega) e^{i\omega\tau} d\omega \qquad (D.9)$$

Using these relations, the limit (D.7) can also be expressed in the frequency domain (Parceval's formula)

$$V_{\infty}(\Theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{i\omega})A^{*}(e^{-i\omega}) - B^{*}(e^{-i\omega})|^{2} \Phi_u(\omega) d\omega +$$

$$+ \frac{1}{2\pi} \int_{-\pi}^{\pi} |A^{*}(e^{-i\omega})|^{2} \Phi_e(\omega) d\omega \qquad (D.10)$$

The result that the Least Squares estimate minimizes the criterion $V_{\infty}(\Theta)$ with respect to Θ, has already been shown in Theorem 4.1. This concludes the proof. □

PART III - A SIMPLE SELF-TUNING CONTROLLER

ABSTRACT

The theory for self-tuning regulators has received numerous contributions in the recent past. Many technical journals contain reports of successful applications of self-tuning control to industrial processes. However, some important questions regarding stability remain unanswered. It is for instance highly desirable to obtain conditions for practical stability of simple self-tuning regulators, when controlling complex plants.

Here we present a new algorithm for a simple self-tuning controller. Due to the simplicity of the regulator, it is possible to derive stability results when the controlled system is more complex than the implied model structure.

1. INTRODUCTION

The theory for self-tuning controllers has developed considerably during the 1970's. Asymptotic properties of the parameter estimates can in some schemes be analyzed with the method of Ljung (1977), provided the closed-loop system is ℓ^{∞}-stable. Global asymptotic stability for some algorithms has been proven when there are no disturbances acting on the system, Egardt (1979) and Goodwin et al (1978). These proofs, however, are derived with the assumption that the "real" system can be exactly described within the model class. In Egardt (1979), there is also a result on ℓ^{∞}-stability when the system is subjected to bounded disturbances.

It is desirable to gain insight into the stability properties of self-tuning controllers for the more realistic situation, when the process is more complex than the model.

For the continuous time case, there are some simple results in this direction based on Lyapunov theory, Parks (1966). These results are derived under the assumption that the system to be controlled has a strictly positive real (s.p.r.) transfer function. Sampled continuous time systems never have s.p.r. transfer functions, and consequently the robustness results of Parks can not in practice be applied to discrete time systems.

In this part of the thesis, a simple self-tuning controller for stable discrete time systems is proposed. The underlying model of the process consists of an unknown gain and a pure time-delay. Such simplified models have recently attracted attention in the literature on robust controllers, see Åström (1980) and Davison et al (1980). The present controller algorithm can be regarded as an adaptive version of a simple IDCOM algorithm, see Richalet et al (1978). The purpose of control of such simplified models is not to achieve an outstanding performance, but rather to keep the process at a specified working point, irrespectively of slow disturbances. As a consequence of the moderate design specifications it is feasible to use long sampling intervals. The simple model structure is thus justified. The controller is shown to have good performance when controlling systems that are well approximated by the simple model.

Due to the structural simplicity of the proposed self-tuning controller, it is possible to derive stability results when the controlled system is more complex than the simple model that the controller is designed for. Sufficient conditions for ℓ^{∞}-stability, and bounds on the system output variable are given. It is believed, that this is the first time that such results are presented.

Part III is organized as follows: In Chapter 2 the controller algorithm is presented. Global asymptotic stability for the closed-loop system is shown to be obtained when the model is correct. A simulation example is included to illustrate the properties of the controller. Chapter 3 is devoted to the stability analysis when the process is more complex than the simple model. First, ℓ^{∞}-stability for the system subjected to bounded disturbances is proven. Then the degree of complexity is further increased by letting the system dynamics be more general than the simple model. Finally, we show that the common class of systems with monotone step responses apply to the ℓ^{∞}-stability results obtained. In Chapter 4, two examples are given. The first is concerned with the presented stability results, and the second example illustrates an application of the self-tuning controller to a laboratory process. A discussion of the results and their possible extensions to general self-tuning controllers is held in Chapter 5. This also serves as the conclusions. References are listed in Chapter 6. An appendix is included with the programming code of the controller in SIMNON, a simulation language.

2. A SIMPLE SELF-TUNING CONTROLLER

The algorithm for a simple self-tuning controller for stable systems will now be presented. The resulting closed-loop system is shown to be globally asymptotically stable if the model of the plant is correct. A simulation example is included, to illustrate the appealing behaviour of the controller.

2.1 THE ALGORITHM

Consider a stable discrete time system, described by the following simple input-output model

$$y(k) = \frac{1}{\beta} u(k-d) \tag{2.1}$$

where d is the known number (integer) of time delays in the system, and β is a constant inverse gain satisfying

$$0 < \beta < \infty \tag{2.2}$$

The aim of control is to follow a command input $y_c(k)$ satisfying

$$0 < |y_c(k)| \leq C \tag{2.3}$$

The exclusion of the value $y_c = 0$ is included simply to avoid division by zero in the following estimation algorithm, and in practice puts no hard limitations on the use of the algorithm.

Let the system be controlled by the control law

$$u(k) = \hat{\beta}(k) y_c(k) \tag{2.4}$$

where $\hat{\beta}(k)$ is an estimate of β, recursively given by the gradient type parameter estimation algorithm

$$\begin{cases} \hat{\beta}(k) = \hat{\beta}(k-d) - \beta_0 [y(k) - y_c(k-d)]/y_c(k-d) & (2.5) \\ \hat{\beta}(0) = \ldots = \hat{\beta}(d-1) = \beta_0 & (2.6) \end{cases}$$

Here β_0 is an initial estimate, or guess, of β.

Although the controller (2.4) looks like a proportional feedforward regulator, the identification algorithm (2.5) introduces integrating feedback from the output y. The self-tuning controller (2.4) to (2.6), is closely related to the "quotient-regulator", proposed by Jensen (1978). However, it turns out that for the present algorithm it is possible to show stability in rather general control situations, while no such proof is yet available for Jensen's controller.

In some situations, it may be unnecessarily restrictive to use the simple model (2.1). We can also consider more general system models of the type

$$y(k) = \frac{1}{\beta} F(q)u(k-d) \qquad (2.7)$$

where F(q) is a completely known, possibly non-linear, stable operator with stable and causal inverse. The corresponding control law will then be

$$u(k) = F^{-1}(q)\hat{\beta}(k)y_c(k)) \qquad (2.8)$$

In this way, known actuator non-linearities and "nice" process dynamics may be included in the model.

An interesting fact is that the simple controller (2.4), (2.5) and (2.6) can be used with non-uniform sampling as well. The sampling time index k is then replaced by the actual sampling instant t_k and the index k-d is replaced by the largest sampling time t_s, such that $t_k - t_s \geq$ (the time delay in the system). This feature may be valuable in the control of many industrial processes, where the measurements of system variables are often obtained at irregular times. Such measurements can e.g. be the results from laboratory tests.

2.2 GLOBAL ASYMPTOTIC STABILITY

If the system to be controlled is governed by the relation (2.1), the following stability result holds:

THEOREM 2.1
Consider a system with input-output relation (2.1) controlled by the control algorithm (2.4),(2.5) and (2.6).

Then the closed-loop system is globally asymptotically stable if and only if

$$0 < \beta_0/\beta < 2 \tag{2.9}$$

Further, the tracking error is bounded from above by an exponentially decreasing function

$$|y(k) - y_c(k-d)| \leq C|1 - \beta_0/\beta|^{\frac{k+1-d}{d}} \tag{2.10}$$

where C is the bound on the command input.

<u>Proof</u>. Equations (2.1) and (2.4) give the tracking error

$$y(k) - y_c(k-d) = \left[\frac{\hat{\beta}(k-d) - \beta}{\beta}\right] y_c(k-d) \tag{2.11}$$

which together with (2.5) yields

$$\hat{\beta}(k) - \beta = [1 - \beta_0/\beta][\hat{\beta}(k-d) - \beta] \tag{2.12}$$

This is a linear, time-invariant difference equation, which is globally asymptotically stable if and only if

$$0 < \beta_0/\beta < 2 \tag{2.13}$$

The initial conditions (2.6) give d subsequencies of the solution to (2.12)

$$\hat{\beta}(nd+\tau) - \beta = [1 - \beta_0/\beta]^n [\beta_0 - \beta] \tag{2.14}$$

$\tau = 0 , 1 , \ldots , d-1$ and $n = 0 , 1 , \ldots$

where $k = nd + \tau$.

It is then evident that the estimation error satisfies

$$|\hat{\beta}(k) - \beta| \leq |1 - \beta_0/\beta|^{\frac{k+1-d}{d}} |\beta_0 - \beta| \tag{2.15}$$

with equality at the time instants

$k = d-1 , 2d-1 , \ldots$

Equations (2.3) and (2.11) give

$$|y(k) - y_c(k-d)| \leq C|1 - \beta_0/\beta|^{\frac{k+1-d}{d}} \tag{2.16}$$

which concludes the proof. □

The stability condition (2.9) is often met in the literature on self-tuning controllers, see e.g. Åström and Wittenmark (1973) and Egardt (1979). From the proof of Theorem 2.1, it is clear that the estimation error $\hat{\beta}(k)-\beta$ is uncontrollable from the command input $y_c(k)$, and that it is observable in the system output $y(k)$. In addition, the characteristic polynomial for these uncontrollable modes in the estimator is

$$z^d - (1 - \beta_0/\beta) = 0 \tag{2.17}$$

With the initial conditions given by (2.6), the estimation error will be monotone and exponentially decreasing if

$$0 < \beta_0/\beta \leq 1 \tag{2.18}$$

and oscillating but exponentially decreasing if

$$1 < \beta_0/\beta < 2 \qquad (2.19)$$

This feature is illustrated by the simulation example in the following section.

2.3 A SIMULATION EXAMPLE

In the previous sections, a simple self-tuning controller was presented. It was shown in Theorem 2.1, that the controller together with a process of the type

$$y(k) = \frac{1}{\beta} u(k-d) \qquad (2.20)$$

gives a stable control system, provided the parameter β_0 satisfies

$$0 < \beta_0/\beta < 2 \qquad (2.21)$$

We will now illustrate these properties with a simulation example.

Consider the "trivial" process described by the relation

$$y(k) = u(k-1) \qquad (2.22)$$

Hence, with the notations used earlier, d = 1 and β = 1. The parameter estimation algorithm becomes

$$\begin{cases} \hat{\beta}(k) = \hat{\beta}(k-1) - \beta_0[y(k) - y_c(k-1)]/y_c(k-1) \\ \hat{\beta}(0) = \beta_0 \end{cases} \qquad (2.23)$$

and the control law is

$$u(k) = \hat{\beta}(k)y_c(k) \qquad (2.24)$$

The results from simulations are shown in Figs. 2.1 to 2.3, for some different values of β_0. The command input $y_c(k)$ is a square wave function, changing between the levels +1 and -1 at every 10-th sample interval. It is seen that the parameter estimate $\hat{\beta}(k)$ has monotone convergence when $\beta_0 < 1$, and oscillatory convergence when $\beta_0 > 1$. When $\beta_0 = 1$, i.e. β_0 is equal to the true parameter β, the estimate remains fixed. The nature of the convergence of the estimate, is clearly reflected in the system output.

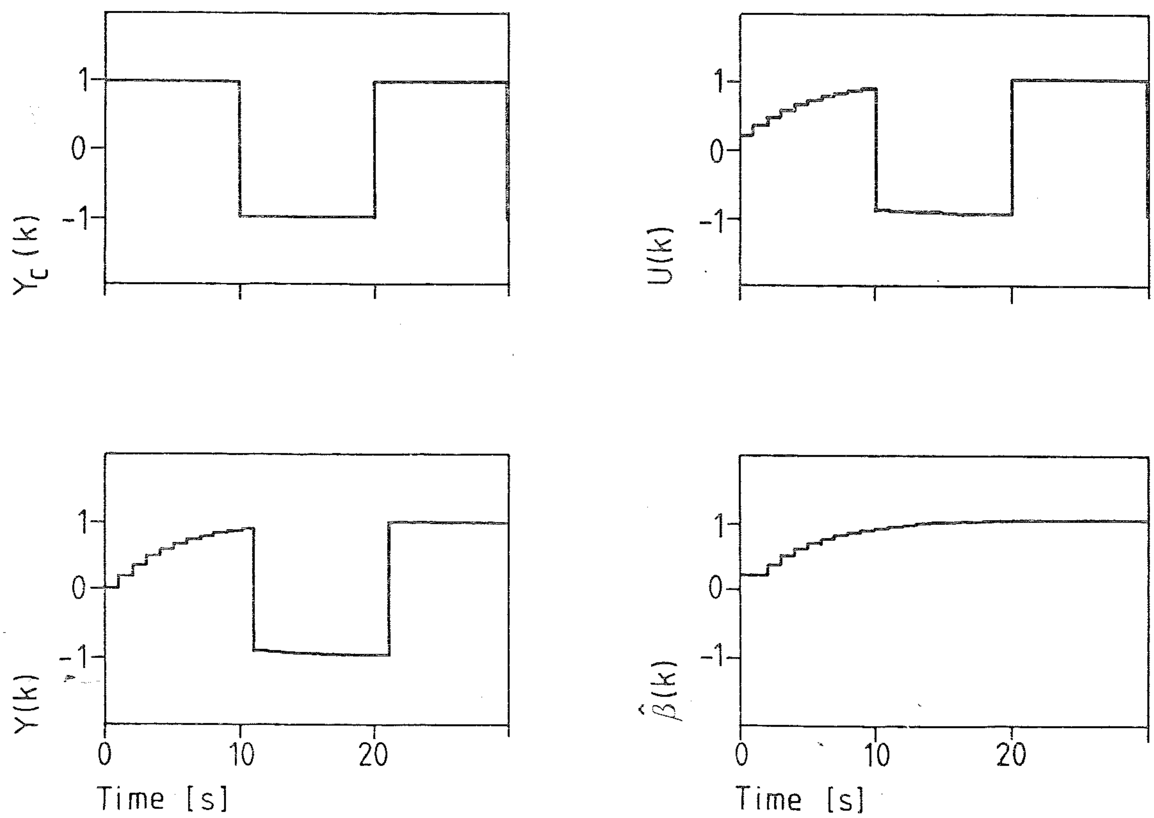


Fig. 2.1 - A plot of the system variables when $\beta_0 = 0.2$.

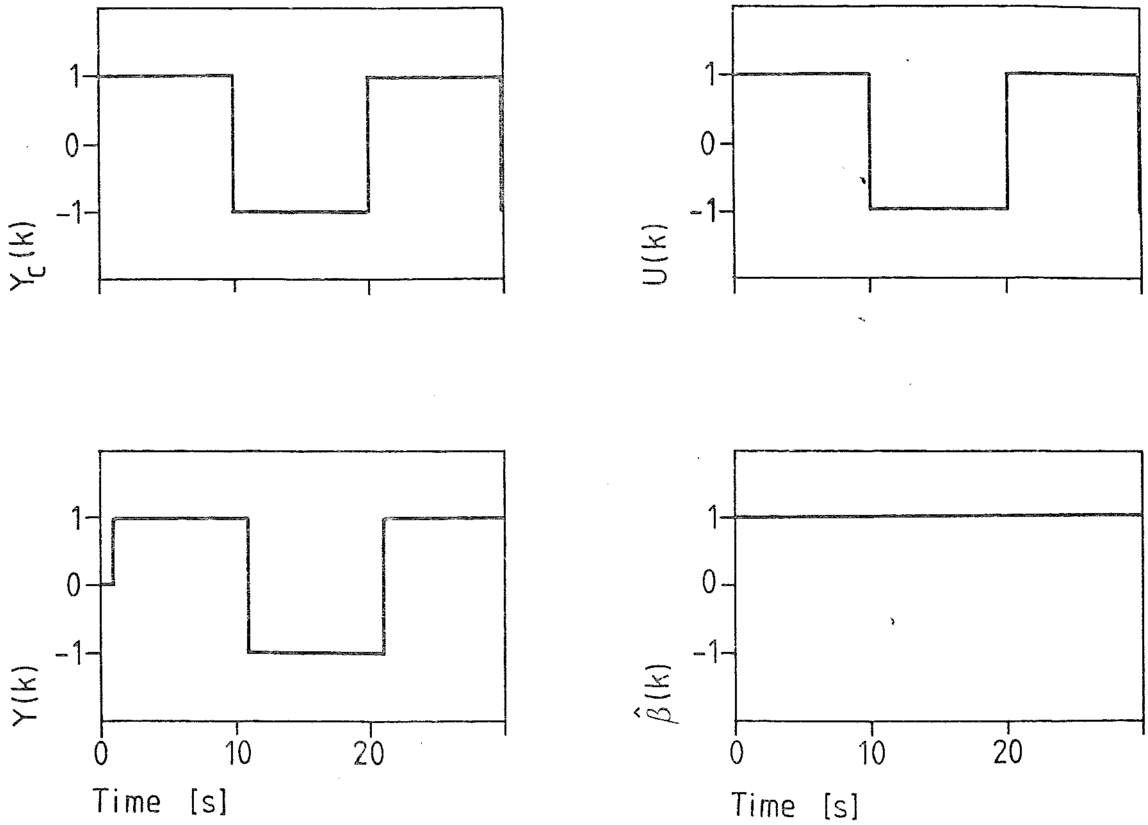


Fig. 2.2 - A plot of the system variables when $\beta_o = 1$.

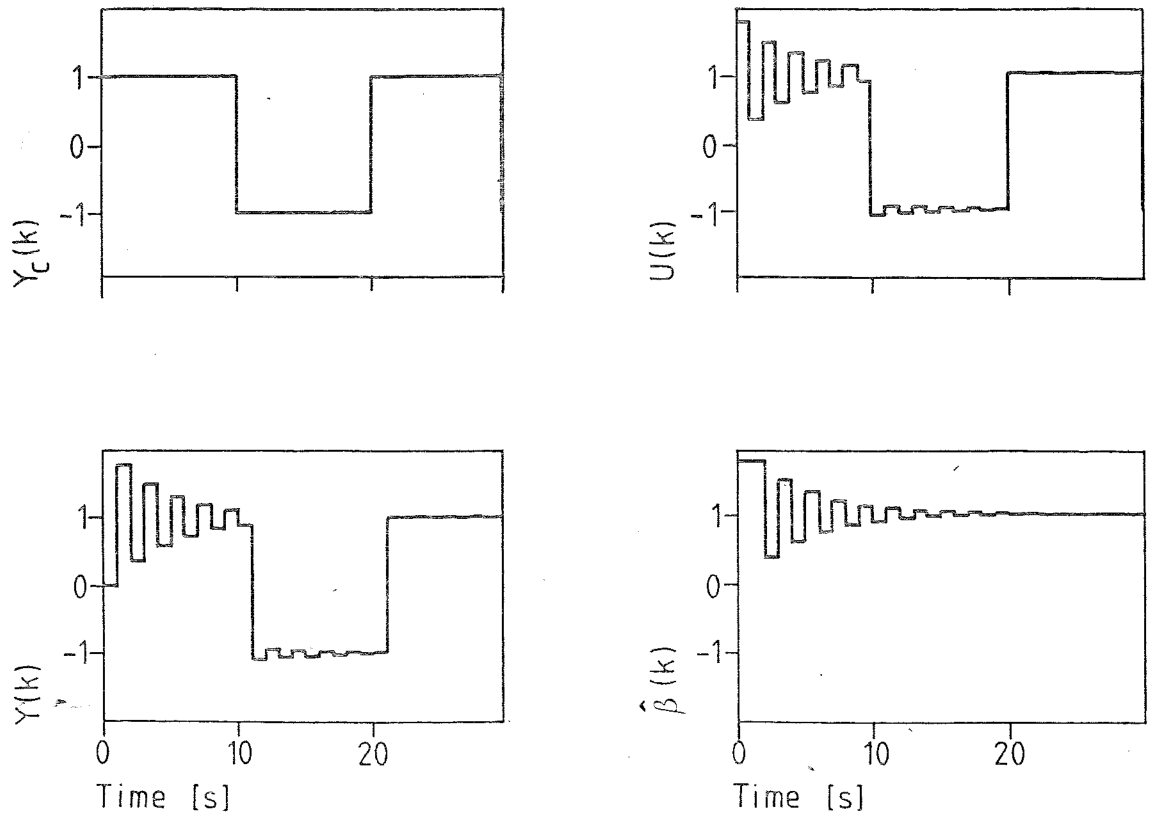


Fig. 2.3 - A plot of the system variables when $\beta_o = 1.8$.

3. STABILITY ANALYSIS

In this chapter, the controlled system is assumed to be more complex than the simple model (2.1). First, the system is augmented with additive disturbances, and later also with additional dynamics. Sufficient conditions for ℓ^{∞}-stability are given, and bounds on the output signal are computed. Finally, we show that the common class of stable systems with monotone step responses apply to the stability results obtained.

3.1 ADDITIVE DISTURBANCES

In the previous chapter the stability of the closed-loop system was established in case of correct model. What will happen when the system is subjected to additive bounded disturbances is stated in the following theorem.

THEOREM 3.1
Consider the system described by the relation

$$y(k) = \frac{1}{\beta} u(k-d) + w(k) \tag{3.1}$$

controlled by the regulator (2.4), (2.5) and (2.6). Assume that the disturbance w(k) and the command input satisfy

$$|w(k)| \le C_w \tag{3.2}$$

$$0 < C_1 \le |y_c(k)| \le C_2 \tag{3.3}$$

Then the closed-loop system is ℓ^{∞}-stable

if and only if

$$0 < \beta_0/\beta < 2 \tag{3.4}$$

and an upper bound on the tracking error is

$$|y(k) - y_c(k-d)| \le \left[1 + \frac{\beta_0/\beta}{1 - |1 - \beta_0/\beta|} C_2/C_1\right] C_w \qquad (3.5)$$

Proof. Assume that the conditions (3.2) and (3.3) are fullfilled. Equations (3.1), (2.4) and (2.5) yield

$$[\hat{\beta}(k) - \beta] = [1 - \beta_0/\beta][\hat{\beta}(k-d) - \beta] - \beta_0 \frac{w(k)}{y_c(k-d)} \qquad (3.6)$$

This is an asymptotically stable difference equation if and only if

$$0 < \beta_0/\beta < 2 \qquad (3.7)$$

The Schwartz inequality gives an upper bound on the parameter estimation error

$$|\hat{\beta}(k) - \beta| \le |\hat{\beta} - \beta| =$$

$$= |1 - \beta_0/\beta||\hat{\beta} - \beta| + \beta_0 C_w/C_1 \qquad (3.8)$$

or equivalently

$$|\hat{\beta}(k) - \beta| \le \frac{1}{1 - |1 - \beta_0/\beta|} \beta_0 C_w/C_1 \qquad (3.9)$$

The tracking error then satisfies

$$|y(k) - y_c(k-d)| \le \frac{\beta_0/\beta}{1 - |1 - \beta_0/\beta|} C_2/C_1 \cdot C_w + C_w \qquad (3.10)$$

and the theorem is proven. □

It is believed that this is the first time that an explicit bound on the tracking error is given for a self-tuning controller, when the controlled system is subjected to arbitrary disturbances.

The bound on the tracking error (3.5) can be simplified further

$$|y(k) - y_c(k-d)| \leq$$

$$\leq \begin{cases} [1 + C_2/C_1]C_w & \text{if } 0 < \beta_0/\beta \leq 1 \\ \left[1 + \dfrac{\beta_0/\beta}{2 - \beta_0/\beta} C_2/C_1\right]C_w & \text{if } 1 \leq \beta_0/\beta < 2 \end{cases} \quad (3.11)$$

This indicates that it might advantageous to underestimate β_0 if there are disturbances acting on the system. The part on the right-hand side of (3.11) that may give rise to large tracking errors, is proportional to C_w/C_1. This term can be interpreted as an inverted signal-to-noise ratio. The tracking error is reduced if the signal-to-noise ratio is increased.

The bound (3.11) is not unnecessarily pessimistic, in the sense that there are sequencies $\{w(k)\}$ and $\{y_c(k)\}$ such that the relation is satisfied with equality for some time instants k.

For the special case when the disturbance is a constant bias w and the command input is constant, the tracking error tends to zero. This is due to the integral action of the controller, introduced by the identification algorithm. This feature of the controller is further commented in the next section.

3.2 THE GENERAL CASE

It is now assumed that the dynamics of the exponentially stable system is more complex than the simple model (2.1). In addition, bounded disturbances may be acting on the system. An appropriate system description is then

$$y(k) = G(q)u(k) + w(k) \quad (3.12)$$

We will now rewrite the equations (2.4) and (2.5) for the controller, into one single equation for the control variable u(k). Multiplying the parameter estimation equation (2.5) with $y_c(k)$ yields a recursive equation for the control input

$$u(k) = y_c(k)/y_c(k-d)\cdot\left[u(k-d) + \beta_0[y_c(k-d) - y(k)]\right] \quad (3.13)$$

The equation (3.13) for the control law is very interesting. It immediately reveals that the controller has integral action. It also shows how the controller gain is adapted whenever there is a change in the setpoint $y_c(k)$. Finally, it represents a convenient way to implement the self-tuning controller on a computer.

A block diagram of the closed-loop system formed from (3.12) and (3.13), is shown in Fig. 3.1. By substituting the equation for y(k) into (3.13), we obtain a time-varying equation for u(k):

$$u(k) = y_c(k)/y_c(k-d)\cdot[\beta_0(y_c(k-d)-w(k)) + (q^{-d} - \beta_0 G(q))u(k)]$$

(3.14)

The closed-loop system, where u(k) is generated by (3.14), is shown in a block diagram form in Fig. 3.2. Since G(q) is a stable operator, the stability of this system is determined only by the closed loop. The following results hold:

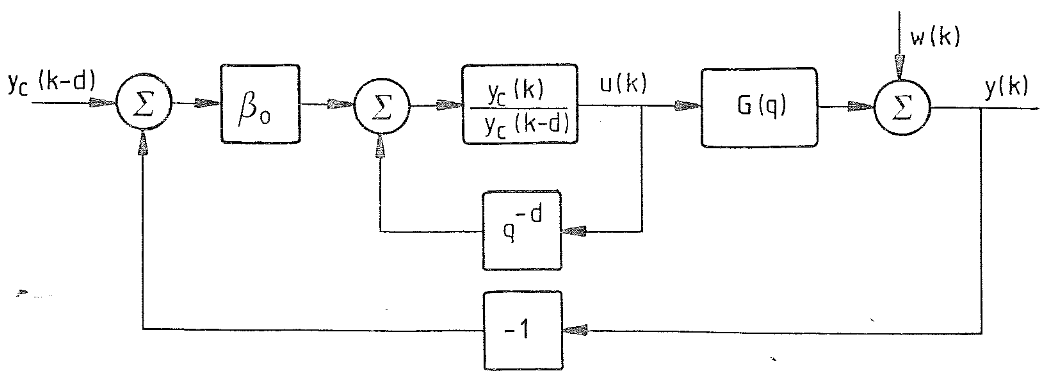


Fig. 3.1 - A block diagram of the closed-loop system.

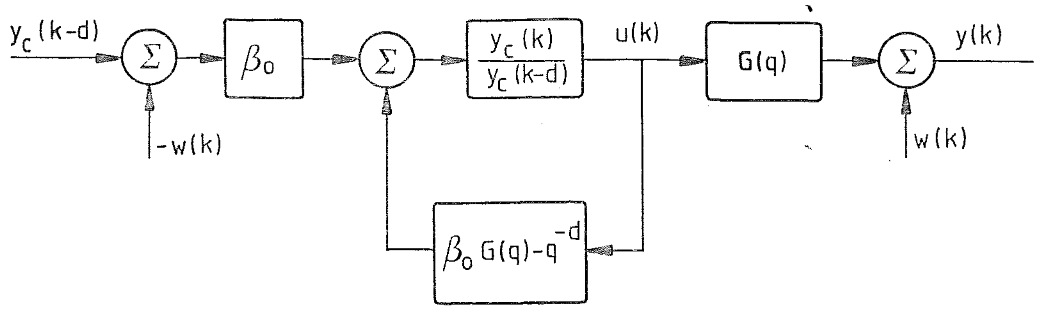


Fig. 3.2 - The transformed system.

THEOREM 3.2
Consider an exponentially stable system with input-output relation

$$y(k) = G(q)u(k) + w(k) \tag{3.15}$$

controlled by the control law (3.13). Assume that the disturbance $w(k)$ is bounded

$$|w(k)| < C_w \tag{3.16}$$

and that the command input satisfies

$$0 < |y_c(k)| < C \tag{3.17}$$

Introduce the following norms

$$g_1 = \| y_c(k)/y_c(k-d) \|_\infty \tag{3.18}$$

$$g_2 = \| \beta_0 G(q) - q^{-d} \|_1 \tag{3.19}$$

$$g_3 = \| G(q) \|_1 \tag{3.20}$$

Then the closed-loop system is ℓ^∞-stable if

$$g_1 \cdot g_2 < 1 \tag{3.21}$$

Further, a bound on the output is given by

$$\|y\|_\infty \le C_w + \frac{\beta_0 g_1 g_3}{1 - g_1 g_2} [C_w + C] \tag{3.22}$$

<u>Proof</u>. The system (3.15) is assumed to be exponentially stable, and hence the norms (3.19) and (3.20) exist. Application of the Small Gain Theorem, Desoer and Vidyasagar (1975), to the transformed system shown in Fig. 3.2, yields the ℓ^∞-stability result.

Now, assuming that the system is stable, a bound on the output signal is

$$\|y\|_\infty \le g_3 \|u\|_\infty + C_w \tag{3.23}$$

The input sequence is also bounded, and the bound can be computed from (3.14)

$$\|u\|_\infty \le g_1 [\beta_0 (C_w + C) + g_2 \|u\|_\infty] \tag{3.24}$$

which inserted in (3.23) yields

$$\|y\|_\infty \le C_w + \frac{\beta_0 g_1 g_3}{1 - g_1 g_2} [C_w + C] \tag{3.25}$$

□

<u>Remark</u>. If the command input is constant, the closed-loop system is linear and time-invariant. A sufficient condition for stability is then that

$$|\beta_0 G(z) - z^{-d}| < 1 \quad ; \quad \forall \ |z| = 1 \tag{3.26}$$

or equivalently, that the transfer function

$$\frac{z^{-d}}{\beta_0 G(z)} - \frac{1}{2} \qquad (3.27)$$

is strictly positive real. This is in general a weaker condition than (3.21). The stability result follows from the Nyquist criterion.

In the case of constant command input, the output tracking error satisfies

$$y(k) - y_c(k-d) = \frac{1 - q^{-d}}{1 + \beta_0 G(q) - q^{-d}} w(k) \qquad (3.28)$$

and an upper bound is given by

$$\|y - y_c\|_\infty = \frac{2}{1 - g_2} C_w \qquad (3.29)$$

If the disturbance is a constant bias w, it follows from (3.28) that the tracking error asymptotically tends to zero. □

The norm (3.19) is in general hard to establish. For a certain class of systems, however, it can be computed in terms of the system step response. The idea to use the step response of the system in the derivation of stability results, has previously been used by Åström (1980).

3.3 SYSTEMS WITH MONOTONE STEP RESPONSES

Consider the class of linear, time-invariant and stable systems with monotone, positive step responses. Let a system in this class be represented by the input-output relation

$$y(k) = G(q)u(k-d) \qquad (3.30)$$

where $G(\infty) \neq 0$, i.e. the system has the same number of time delays as the simple model (2.1). Introduce the step response $H(k)$ associated with the transfer function $G(z)$:

$$H(k) = G(q)s(k) \tag{3.31}$$

where the step input $s(k)$ is given by

$$s(k) = \begin{cases} 1 & \text{if} \quad k \geq 0 \\ 0 & \text{if} \quad k < 0 \end{cases} \tag{3.32}$$

It is now possible to express the transfer function $G(z)$ in terms of the step response $H(k)$

$$G(z) = H(0) + \sum_{i=1}^{\infty} (H(i) - H(i-1))z^{-i} \tag{3.33}$$

The following stability result holds:

THEOREM 3.3
Consider the system (3.30) controlled by the simple self-tuning controller. Assume that there exists a $\delta \geq 0$ such that

$$2H(0) \geq H(\infty) + \delta \tag{3.34}$$

and

$$2/\beta_0 \geq H(\infty) + \delta \tag{3.35}$$

Then the following is true

$$\max_{|z|=1} |\beta_0 z^{-d} G(z) - z^{-d}| \leq \|\beta_0 q^{-d} G(q) - q^{-d}\|_1 \leq 1 - \beta_0 \delta \tag{3.36}$$

and the closed-loop system is ℓ^{∞}-stable if the command input satisfies

$$\|y_c(k)/y_c(k-d)\|_{\infty} < 1/(1 - \beta_0 \delta) \tag{3.37}$$

Proof. In equation (3.33) all the differences $H(i) - H(i-1)$

are non-negative since the step response is monotone and positive. Thus, the following holds

$$\max_{|z|=1} |\beta_0 z^{-d} G(z) - z^{-d}| =$$

$$= \max_{|z|=1} \left| \beta_0 H(0) + \beta_0 \sum_{i=1}^{\infty} (H(i) - H(i-1)) z^{-i} - 1 \right| \leq$$

$$\leq |\beta_0 H(0) - 1| + \beta_0 [H(\infty) - H(0)] = \|\beta_0 q^{-d} G(q) - q^{-d}\|_1 \quad (3.38)$$

In the last equality, consider two separate cases:

I) $0 < \beta_0 H(0) \leq 1$

$$\|\beta_0 q^{-d} G(q) - q^{-d}\|_1 = 1 + \beta_0 [H(\infty) - 2H(0)] \leq 1 - \beta_0 \delta \quad (3.39)$$

from (3.34).

II) $1 \leq \beta_0 H(0)$

$$\|\beta_0 q^{-d} G(q) - q^{-d}\|_1 = \beta_0 H(\infty) - 1 \leq 1 - \beta_0 \delta \quad (3.40)$$

from (3.35). The stability result (3.37) then follows from Theorem 3.2. □

By straightforward computations, it can be shown that the value of β_0 that minimizes the norm

$$g_2 = \|\beta_0 q^{-d} G(q) - q^{-d}\|_1 = \|\beta_0 G(q) - 1\|_1 \quad (3.41)$$

is $\beta_0 = 1/H(0)$, and the corresponding minimum is

$$\min_{\beta_0} \|\beta_0 G(q) - 1\|_1 = \frac{H(\infty) - H(0)}{H(0)} \quad (3.42)$$

This gives a useful rule of thumb how to choose the parameter β_0 in order to guarantee a certain margin of stability, when controlling processes with monotone step responses. If the system has a smaller number of time delays than the model (2.1), it is still sometimes possible to conclude stability from the condition (3.21). The expression for the norm g_2 is in this case similar to the one presented in Theorem 3.3.

It was shown in Theorem 3.2 that the control system is ℓ^∞-stable if the command input $y_c(k)$ satisfies

$$|y_c(k)/y_c(k-d)| < 1/g_2 \qquad ; \forall k \tag{3.43}$$

where g_2 is given by (3.19). It may therefore be of interest to incorporate some action, so that the inequality (3.43) is satisfied. A nonlinear device, that operates on an arbitrary signal $\{u_c(k)\}$ and has an output $\{y_c(k)\}$ that satisfies (3.43) is given by

$$y_c(k) = \min[1/g_2 \cdot |y_c(k-d)|, \ |u_c(k)|] \cdot \mathrm{sign}[u_c(k)] \tag{3.44}$$

This device may easily be implemented in a computer, in connection with the simple self-tuning controller. As was pointed out in Chapter 2, the command input must for practical reasons be bounded away from zero. It is therefore suitable to implement a test in the computer to make sure that y_c satisfies

$$|y_c(k)| \geq \varepsilon \tag{3.45}$$

for some small positive ε. This has also as a consequence that a bound on the tracking error (3.5) will be obtained. The error becomes less sensitive to disturbances as ε is increased. It may also be of advantage to use some limitations on the magnitude and increments of the command input, as is usually done in industrial controllers.

4. EXAMPLES

In this chapter, two examples are given. The first is concerned with the stability results derived in Chapter 3. The second example illustrates the application of the self-tuning controller to a realistic laboratory process.

EXAMPLE 1

Consider a stable first order continuous time system with the transfer function

$$P(s) = \frac{2}{s + 2} \tag{4.1}$$

When this system is sampled with a sample period of $h = 1$ second, the resulting sampled system becomes

$$y(k) = G(q)u(k-1) \tag{4.2}$$

where the transfer operator $G(q)$ is given by

$$G(q) = \frac{0.865}{1 - 0.135q^{-1}} \tag{4.3}$$

It is easily verified that the step response $H(k)$ associated with the transfer function $G(z)$, is monotone and positive. With the notations of Chapter 3, we have that

$$H(0) = 0.865$$

$$H(\infty) = 1.0$$

Now, let the sampled process be controlled by the simple self-tuning controller. The parameter β_0 is chosen so that the norm $\|\beta_0 G - 1\|_1$ is minimized. According to Chapter 3, this choice is given by

$$\beta_0 = \frac{1}{H(0)} = 1.16 \tag{4.4}$$

and the corresponding minimum is

$$g_2 = \frac{H(\infty) - H(0)}{H(0)} = 0.16 \tag{4.5}$$

Application of Theorem 3.2 to the controlled system, gives that the system is ℓ^∞-stable if the command input $y_c(k)$ satisfies

$$|y_c(k)/y_c(k-1)|\cdot g_2 < 1 \tag{4.6}$$

or equivalently

$$|y_c(k)/y_c(k-1)| < 6.4 \tag{4.7}$$

We will now examine the practical stability properties of the system with some simulations. The command input $y_c(k)$ first is chosen as a square wave function, switching between the levels 2.0 and 1.0 at every 15-th sample interval. The function $y_c(k)/y_c(k-1)$ is then also a square wave function, switching between the levels 2.0 and 0.5. The stability condition (4.7) is thus satisfied. In Fig. 4.1 we show a simulation of the system output. The control of the system is quite acceptable.

When the command input instead is switching between the levels 8.0 and 1.0 at every sample interval, the relation (4.7) is violated at every second sample instant. A simulation of the system output is shown in Fig. 4.2. The output is growing rapidly and the system is clearly unstable. The latter simulation shows that violation of the (sufficient) stability bound (4.7), may cause an unstable system.

Here it is appropriate to point out, that the command input $y_c(k)$ used in this example to generate instability, is quite pathological. In practice the command input does not change between two levels at every sample instant!

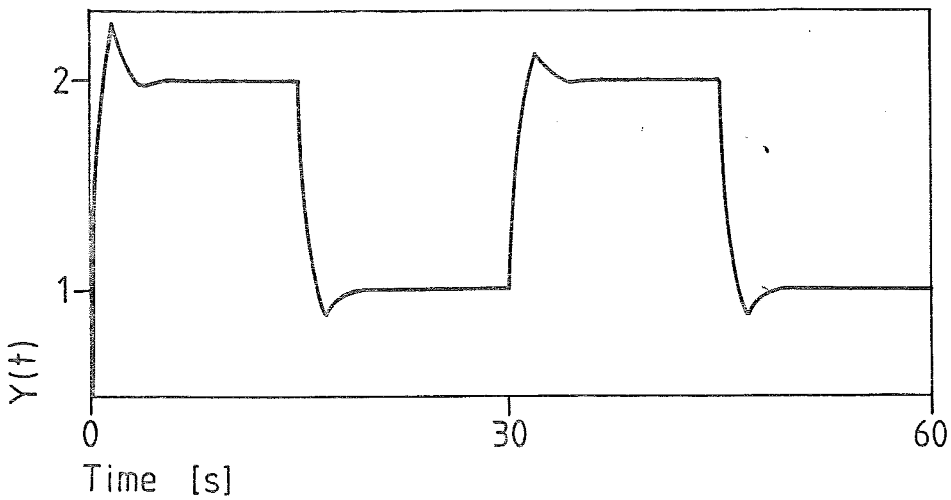


Fig. 4.1 - The system output when the command input changes level at every 15-th sample interval.

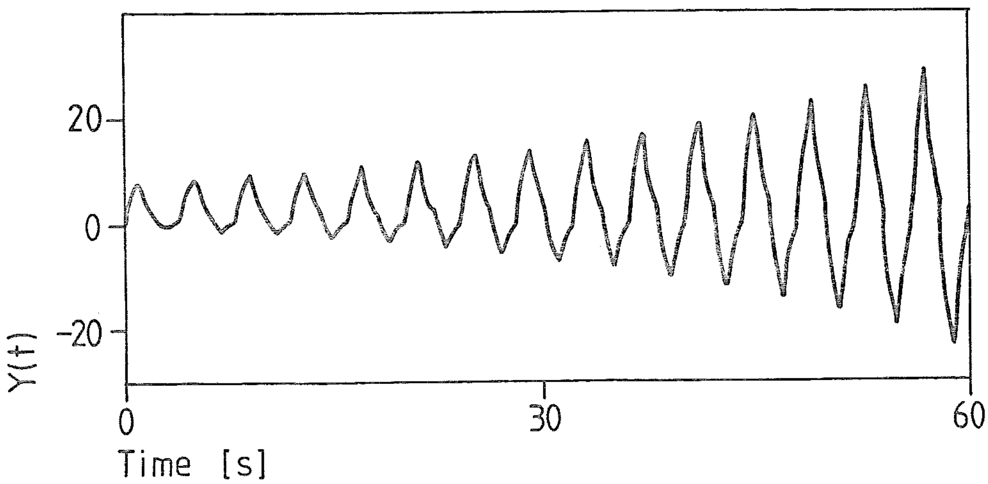


Fig. 4.2 - The system output when the command input changes level at every sample interval.

EXAMPLE 4.2

This final example illustrates the application of the simple self-tuning controller to a realistic laboratory process.

The Plant

The process to be controlled is a small-scale plant for continuous mixing of salt and water. The mixing is to be made to a prescribed concentration. A schematic picture of the plant is shown in Fig. 4.3. At the point A, concentrated salt solution is injected into a flow of pure water. The resulting salt-water solution is homogenized in the mixing tank B. A measurement of the salt concentration in the outflow, is made by measuring the conductivity at the point C.

A reasonable dynamical model for the mixing tank is a linear, first order system. The tubing from A to B and from B to C, introduce a small time delay before a change in the amount of injected salt becomes measurable at C.

The control variable u to the process, is the input signal to an electric pump motor controlling the injection of salt. The characteristics of the pump and the injection nozzle are slightly nonlinear, giving different steady state gain for different operating points. As was previously mentioned, the output variable y is a conductivity measurement of the outcoming fluid. A simple transfer function model for the input-output relation of the system is

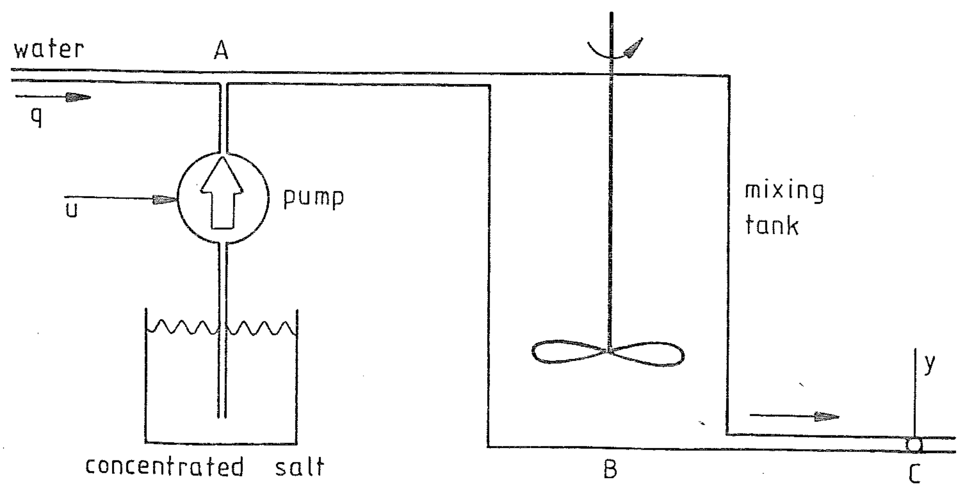


Fig.4.3 - A schematic diagram of the mixing process.

$$y(s) = \frac{K \cdot e^{-s\tau}}{sT + 1} u(s) \qquad (4.8)$$

The parameters K, τ and T all depend on the inflow of pure water q.

A recording of a step response experiment on the system is shown in Fig. 4.4. During the experiment the water flow was at the nominal level, i.e. 100 % . Approximate values of the parameters of the model (4.8) are at the nominal water flow

$K \approx 0.6$

$\tau \approx 13$ seconds

$T \approx 20$ seconds

During operation of the process, there are some obvious sources for disturbances. First, the water pressure at the inlet is varying somewhat, causing variations in the flow q. This in turn influences the parameters K, τ and T of the system. The conductivity measurement y is sometimes unreliable. When occational air-bubbles passes the measurement device, the output shows bursts of noise. Different batches of the concentrated salt solution to be injected may have different concentrations, having as an effect that the process gain K varies. Finally, the mixing in the tank is not perfect. This causes small "pockets" of fluid with high salt concentration to appear in the tank. When the water flow changes marginally, these "pockets" suddenly are stirred up by the turbulence. The symptoms of this, are unexpected indications in the output measurement.

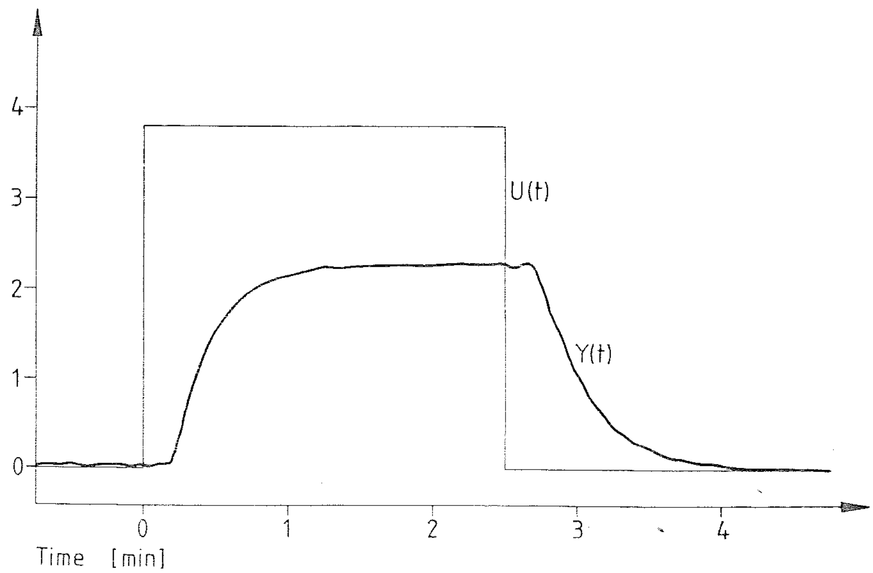


Fig. 4.4 - A step response from the process.

Control

The salt process will now be controlled by the simple self-tuning regulator described in Chapter 2. The aim of the control is to maintain the concentration of salt in the outflow at a prescribed reference value. Moderate variations in the system characteristics must not damage the control. Since the controller is able to adapt itself to slow variations in process gain and to slow disturbances, it appears to be suitable for the control purpose.

Studying the step response of the system in Fig. 4.4, gives at hand that a suitable sampling interval is h = 60 seconds. For this sampling period the simple model (2.1) gives a reasonably good approximation of the input-output behaviour of the system. We choose the gain parameter of the controller to $\beta_0 = 1.5$.

The simple self-tuning controller is implemented in the programming language Process Basic, on a micro-computer LSI-11. A recording of the process output y(t) is shown in Fig. 4.5. The output signal follows the reference level $y_c(t)$ in an acceptable manner. The demands on the speed of the control system can not be too high when this simple regulator is used. It is seen that the settling time of the regulated process is about the same as that for the unregulated plant. Within the span of the different working levels used in the recorded experiment, the steady state gain of the process varies between 0.5 and 0.85 depending on the nonlinearities in the actuator. This does not visibly affect the control.

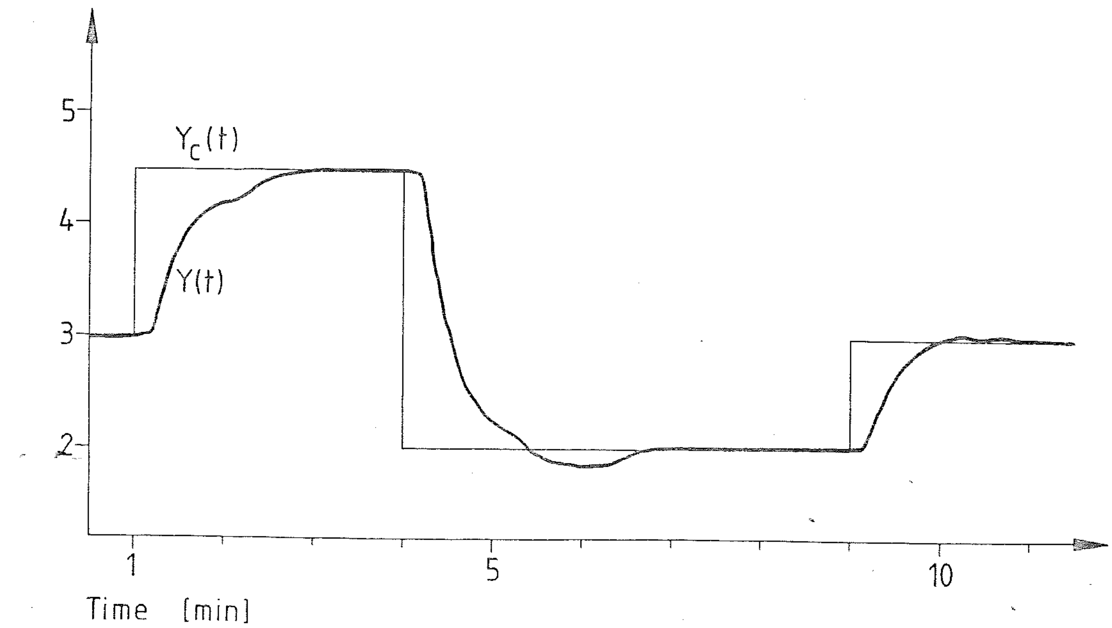


Fig. 4.5 - The output from the regulated process.

In Fig. 4.6, a recording of the process output is shown when the reference level of the concentration is kept constant. The inflow of the pure water is varied to, illustrate how well the controller can compensate for disturbances. For the shown variations in water flow, the parameters of the system K, τ and T change approximately ± 50 % . Normally, the changes in flow are slower than the illustrated step variations, and consequently the tracking tò the reference signal is usually much better.

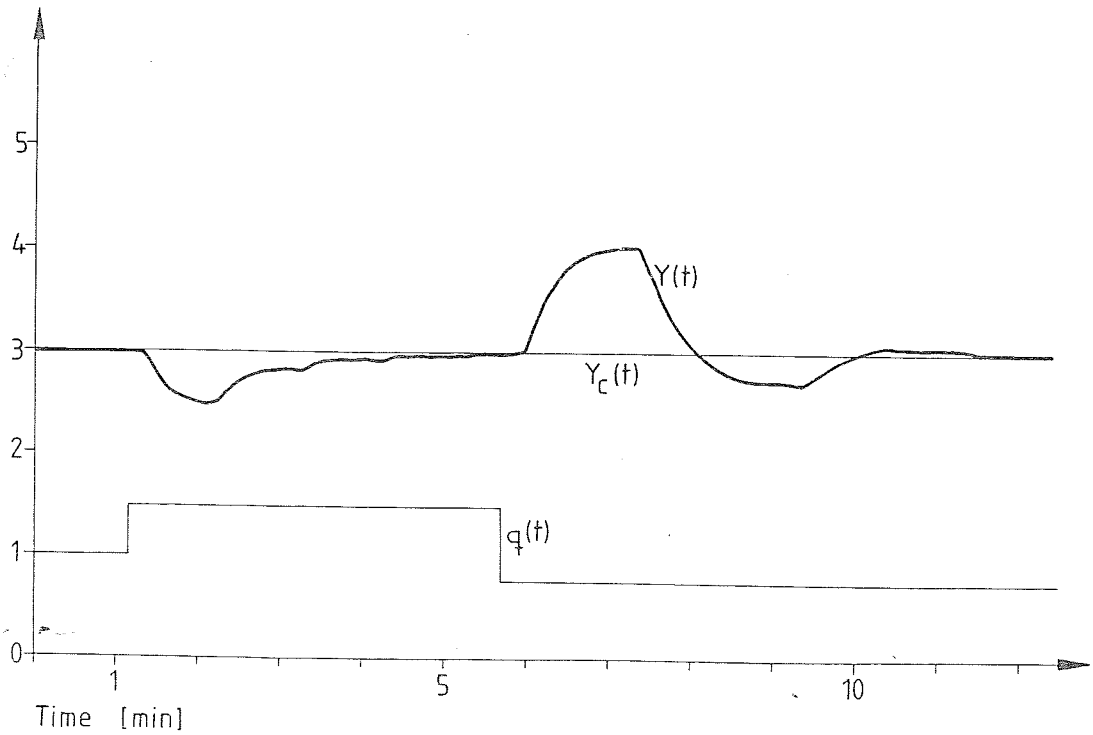


Fig. 4.6 - The process output when the water flow changes.

5. DISCUSSION

In this chapter we will give some brief comments and discussions around the previously derived results. This will also serve as the conclusions.

First of all, it is suitable to mention something about the basic controller algorithm (2.4) to (2.6). This is in principle an ordinary self-tuning controller of the type first described in Åström and Wittenmark (1973). The fact that the only regression variable in the parameter estimation algorithm is the command input $y_c(k)$, facilitates the analysis considerably. The controller is shown to have appealing properties when controlling systems with the same simple structure as the model. The application example in Chapter 4, illustrates that the control is acceptable also when controlling more complex processes, provided the demands for speed are moderate.

The necessary and sufficient condition for stability

$$0 < \beta_0/\beta < 2 \qquad (5.1)$$

is encountered in practically all literature on self-tuning control. For the present algorithm the parameter β_0 determines the locations of d uncontrollable and observable poles in the parameter estimator (observer). In the general self-tuning controller this is not strictly correct, but still a way to intuitively understand the significance of β_0.

The result in Theorem 3.1 on ℓ^∞-stability with disturbances, resembles the result in Egardt (1979) for more general self-tuning controllers. This could be expected since, as was mentioned above, the controllers essentially have the same structure. In the present case, however, it is possible to derive an explicit bound on the tracking error. It is believed to be the first time that such a result is obtained for a self-tuning controller.

The stability result Theorem 3.2 and it's applicability to practical processes through Theorem 3.3, is also believed to the first of it's kind. The basic idea is that a rapid change in the system input excitates unmodelled modes in the system. These modes become visible in the system output and thus influences the parameter estimates, who in turn influence the system input, etc. There is clearly a possible "loop of events" that may cause instability. Unfortunately the method of proof used in Theorem 3.2, the Small Gain Theorem, has not yet been possible to apply in the stability analysis of more general self-tuning controllers.

For more complex self-tuning controllers there exist instability phenomena similar to that encountered in Example 4.1, for the simple regulator. This is illustrated in the following example.

EXAMPLE 5.1

Consider a stable discrete time system with input-output relation

$$y(k) - 1.1y(k-1) + 0.25y(k-2) = u(k-1) \tag{5.2}$$

controlled by a standard self-tuning controller, designed for a first order model. The control signal is

$$u(k) = \hat{a}(k)y(k) + y_c(k) \tag{5.3}$$

and the parameter estimate $\hat{a}(k)$ is recursively given by

$$\hat{a}(k) = \hat{a}(k-1) - \frac{[y(k) - y_c(k-1)]}{1 + y(k-1)^2} y(k-1) \tag{5.4}$$

The controller (5.3) and (5.4) is discussed e.g. in Goodwin et al. (1978). The command input in this example is chosen as a square wave function, switching between the levels 10.0 and 30.0. Simulations of the system output are shown in Figs. 5.1 and 5.2, when the period of the square wave function is 60.0 and 4.0, respectively. In Fig. 5.1 it is seen that the performance of the system is rather poor. The output y(k) has a transient of oscillations after each step in the command input, but nevertheless the system remains stable.

Fig. 5.2 shows that when the period of the command input is decreased to 4.0, the system output behaves very strangely. There are occational bursts of large magnitude, totally destroying the performance. Clearly the system exhibits instability phenomena for some types of command inputs. □

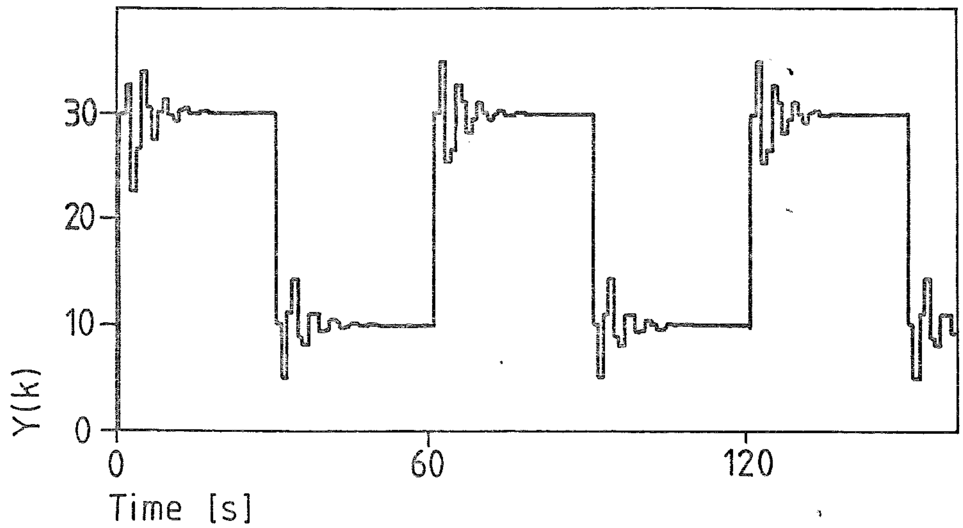


Fig. 5.1 - The system output when the command input has the period 60.0.

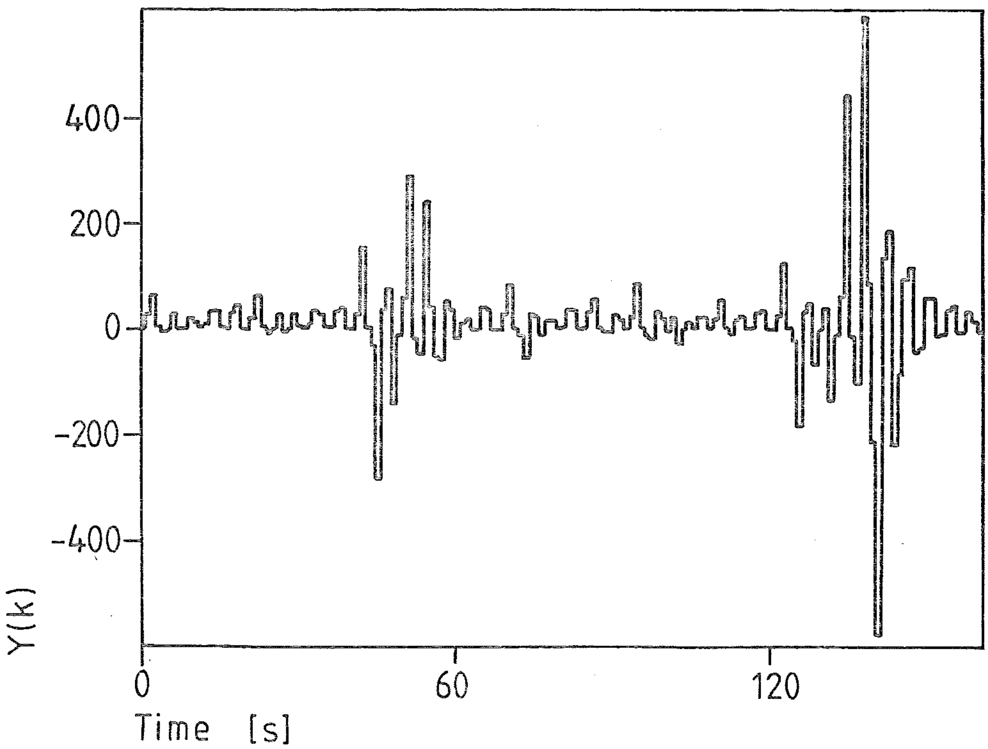


Fig. 5.2 - The system output when the command input has the period 4.0.

It is fairly simple to generate simulation examples that show instability for self-tuning controllers, as was illustrated above. Since there is an increasing number of practical applications of self-tuning controllers in the industry, it is highly desirable to obtain design principles that guarantee stability for such control systems. This can be done both by the trial-and-error method of simulation, and by theoretical stability analysis. An obvious goal for future research in this field, is to extend the presented stability results to general self-tuning and adaptive controllers. This may e.g. provide rules of thumb how to choose the complexity of the controller and how to select the sampling period for a given process. Another important area of research is to analyze if linear and non-linear filtering of the signals in self-tuning controllers can improve their stability properties. This was implied by the present work, and has earlier been suggested by simulations of model reference adaptive regulators, see Egardt (1979).

At the present we lack much understanding of the way that self-tuning controllers work. Analysis of very simple controller algorithms, like the present, may be the right way to attack this problem.

6. REFERENCES

Aström, K.J. and Wittenmark, B.(1973): On Self-Tuning Regulators. Automatica 9, 185-199.

Aström, K.J.(1980): A Robust Sampled Regulator for Stable Processes with Monotone Step Responses. Automatica 16, 313-315.

Desoer, C.A. and Vidyasagar, M.(1975): Feedback Systems - Input-Output Properties. AP, New York.

Egardt, B.(1979): Stability of Adaptive Controllers. Springer Verlag, Berlin.

Elmqvist, H.(1977): SIMNON - An Interactive Simulation Program for Nonlinear Systems. Proc. Simulation '77, Montreux.

Goodwin, G.C., Ramadge, P.J. and Caines, P.E.(1978): Discrete Time Multivariable Adaptive Control. Div..of Applied Science. Harvard University.

Jensen, L.(1978): Digital Reglering av Klimatprocesser. Lund Inst. of Technology, Lund, Sweden. (In Swedish). CODEN: LUTFD2/(TFRT-1014)/1-263/(1978).

Ljung, L.(1977): Analysis of Recursive Stochastic Algorithms. IEEE Trans. AC-22, 551-575.

Parks, C.P.(1966): Liapunov Redesign of Model Reference Adaptive Control Systems. IEEE Trans. AC-11, 362-367.

Richalet, J., Rault, A., Testud, J.L. and Papon, J.(1978): Model Predictive Heuristic Control: Application to Industrial Processes. Automatica 14, 413-429.

APPENDIX

In this appendix we give the programming code for the simple self-tuning controller in SIMNON, a simulation language. SIMNON is available at the Department of Automatic Control at Lund Institute of Technology, Lund, see Elmqvist (1977).

```
Discrete System SSTC
" A Simple Self-Tuning Controller designed for system
" models consisting of a pure time delay and a constant
" gain:
"
"     y(k) = 1/b * u(k-1)
"
" Author C F Mannerfelt 810724

" The names of variables and parameters are as follows:
" ************************************************************
" y: the process output
" yc: the command input
" ycold: the one step delayed value of yc
" u: the control input
" bhat: the actual estimate of b
" bhato: the old estimate of b
" b0: the initial estimate of b
" dt: the sampling interval

INPUT y yc
OUTPUT u
STATE bhato ycold
NEW bhat nycold
TSAMP ts
TIME t

INITIAL

bhato=b0

SORT

" Computation of the control signal
" **********************************

u=bhat*yc


" Updating of the parameter estimate
" **********************************
```

```
bhat=if t<dt then b0 else bhato-b0*(y-ycold)/ycold

" Updating of the regression variable
" ***********************************

nycold=yc

ts=t+dt

b0:1     " Setting of initial estimate of b
dt:1     " Setting of sampling interval

END
```