On the Role of Filters in Adaptive Control

Wittenmark, Björn

1986

ON THE ROLE OF FILTERS IN ADAPTIVE CONTROL

by

Bjorn Wittenmark*

Technical Report EE8662

December 1986

* Department of Electrical and Computer Engineering, University of Newcastle, New South Wales 2308, Australia. On leave from Lund Institute of Technology, Sweden.

ABSTRACT

This note gives a summary and some illustrative examples of the role of different filters in adaptive control that are needed to make a robust implementation. In adaptive controllers it is important to filter the signals in order to avoid erroneous results in the estimation part of the algorithm. In discrete time adaptive controllers as well as in sampled data controllers it is important to filter the signals to avoid the aliasing problem. An attempt is made to quantify and illustrate some of the "folklore" in this area.

TABLE OF CONTENTS

1. INTRODUCTION

Adaptive controllers are now more and more used in industrial applications, see the surveys Åström (1983b) and Seborg et al (1986). The convergence and stability issues for some classes of adaptive controllers were solved in the late 1970's for ideal cases such as known orders and delays of the process. The robustness properties of adaptive controllers have been debated intensively during the last years. The concencus of these discussions seems to be that adaptive control can be useful, but that extreme care must be taken when implementing the controllers. Different fixes have been suggested based partly on experience and partly on analysis.

The implementation and robustness issues are therefore of great importance. This report discusses some of these aspects. The signal processing, i.e. the choice of different filters, in adaptive controllers is treated. An attempt is made to quantify some of the "folklore" around the implementation issues. We will concentrate on discrete time adaptive controllers. Some of the signal processing issues are then the same as for implementation of fixed parameter sampled data controllers. The problems are for instance the choices of the sampling period, the antialiasing filters and the specifications. These choices are, however, also influenced by the process, which now is unknown. The estimator part of the adaptive controller is also of great importance since it must produce accurate process models in appropriate frequency bands. Finally the interaction between the estimator and the controller may cause problems. Practical issues have been discussed for instance in Astrom (1983a,b), Rohrs et al (1984) and Wittenmark and Åström (1984). The following quotations give some general advices and statements about the implementation.

Quote 1.1 (Ljung and Söderström (1983), p 139-140)

"... even if the true system is more complex than our models, the identification procedure will pick the best approximation of the system. ... The recursive algoirthm converges to that approximation that is best under the input signal used during the experiment."

Quote 1.2 (Ljung and Söderström (1983), p 267)

"It is thus good practice to choose an input for the identification experiment that as far as possible is similar to inputs to be used for the system at later occasions."

Quote 1.3 (Aström (1983a), p 986)

"Do not estimate unless the data is good."

Quote 1.4 (Aström (1983a), p 986)

"... beneficial to use a design method which gives a high gain at low frequencies and use adaptation only to find the characteristics around the cross-over frequency."

Quote 1.5 (Rohrs et al (1984), p 653)

"Keep the adaptation gain of the system small and let the adaptation proceed slowly."

Quote 1.6 (Rohrs et al (1984), p 653)

"Design the nominal control loop so that it is robust and that approximate model matching can be achieved even in the presence of unmodelled dynamics."

These quotations and references pinpoints some of the fundamental issues when implementing adaptive controllers. In this paper we will try to quantify these advices and arrive at an overall view of the implementation problem. The adaptive control problem is first defined in Section 2. The true process is in general more complex than the estimated model. To get a good approximation it is important to adapt the estimator to the selected class of models. This problem is discussed in Section 3. The starting point is a generalization (or rather adaptation) of results in Wahlberg and Ljung (1986) and Gevers and Ljung (1986). Based on these results some general rules of thumb can be given. Section 4 gives a discussion of the implementation of the recursive estimator. One main question is how to judge when the received data contains useful information. Sampled data control aspects as the choices of antialiasing filters, sampling interval and the bandwidth compromise is discussed in Section 5. The interaction between the estimator and the controller is finally treated in Section 6. This section also contains an example, which illustrates the points discussed in the report. Sections 7 and 8 contain conclusions and references.

The contributions of this paper relative to previous work are:

a) An overall treatment of the implementation problem.

b) Quantitative rules of thumb are given to help the implementor.

c) Both analysis and examples are used to illustrate the different points.

ACKNOWLEDGEMENTS

The author's visit in Australia has been possible through support by Australian Research Grant Scheme, Swedish Board of Technical Development (STU), Troedssons Forskningsfond and Rydins Stiftelse.

2. THE ADAPTIVE CONTROL PROBLEM

An adaptive controller can often be structured as in Figure 2.1. The system consists of two loops: an inner loop with a conventional regulator and an outer loop which changes the parameters in the regulator. The regulator parameters are calculated based on the specifications on the closed loop system and the estimated parameters. Even if the system is adaptive the user has to specify for instance what the specifications should be and which design method that should be used. These choices depend on the physical insight into the process and the disturbances acting on the process. This implies that even an adaptive controller can not be implemented without knowledge about the process. This knowledge is used as for the case with a known system to determine suitable specifications and the structure of the controller. The process is also more complex than the estimated model. This implies that we have to decide over which frequencies we want to have a good approximation, compare Quotes 1.4 and 1.6.

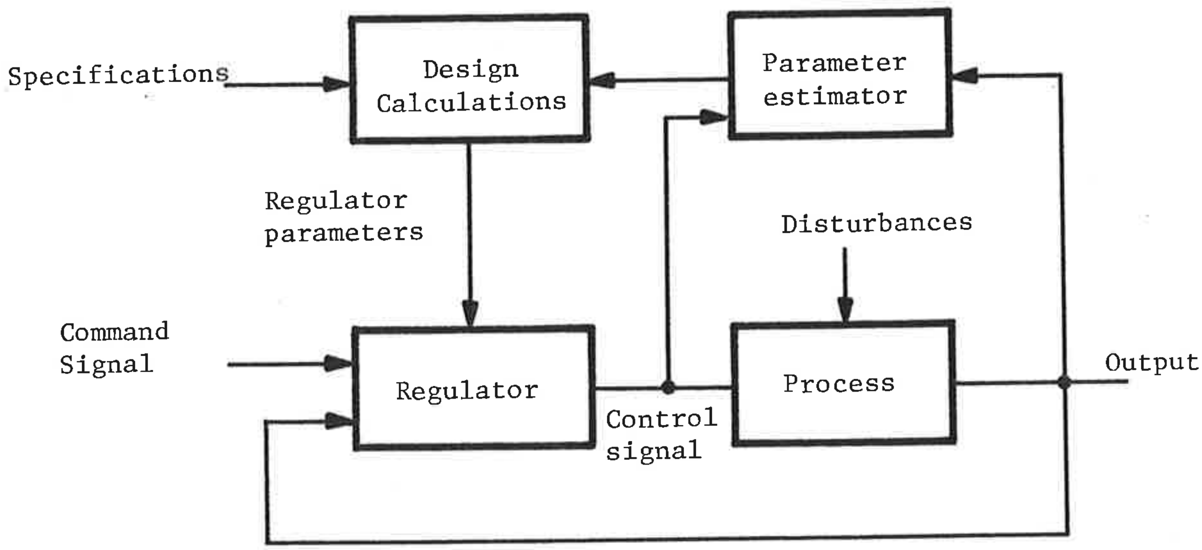


Figure 2.1 Structure of an Adaptive Control System

To be more specific we will assume that the process is described by the single input single output discrete time model

$$y(t) = G_o(q)u(t) + H_o(q)e(t) + d(t) \tag{2.1}$$

where

$$G_o(q) = \frac{B_o(q)}{A_o(q)} \tag{2.2}$$

$$H_o(q) = \frac{C_o(q)}{D_o(q)} \tag{2.3}$$

and y and u are the output and input of the process. Notice that we allow possible antialiasing filters to be a part of the process $G_o(q)$. The pulse transfer operator $G_o(q)$ is also a function of the sampling period h. This dependence is, however, surpressed in the formulas. The distubance $e(t)$ is assumed to be a sequence of independent random variables with zero mean values and variances σ^2. The second disturbance term $d(t)$ is a purely deterministic signal of known form but unknown magnitude. It may be a level, a ramp or a sinusoidal signal.

It is assumed that the desired specifications of the closed loop system is given as the bandwidth or in terms of a desired closed loop model

$$G_m(q) = \frac{B_m(q)}{A_m(q)} \tag{2.4}$$

Adaptive controllers can be divided into direct and indirect methods. In the direct methods the controller parameters are estimated directly. This implies that the design block in Figure 2.1 is very simple. The specifications are in this case used to make a reparameterization of the model, see for instance Åström and Wittenmark (1980). In the indirect methods we first estimate the parameters in a model such as (2.1). The estimated parameters are then used in the design calculations as if they are the true ones.

Based on the specifications and the knowledge of the process we will in the following sections discuss the implementation problem and how to make an appropriate filtering of the signals.

3. ADJUSTMENT TO THE PREJUDICE MODEL

When determining the structure and complexity of a controller we always have some more or less well defined "feeling" of how the process can be described. For instance its complexity, amount of time delay, types of disturbances. The purpose of the estimator part of the adaptive controller is to find the parameters of this prejudice model. In this section we will discuss how to adjust the measured data (Sic!) to the prejudice model.

Elimination of known types of disturbances

The known types of disturbances was represented by $d(t)$ in (2.1). $d(t)$ can be thought of as generated by pulses into known dynamic systems. It is assumed that it is not known a priori when the pulses occur and the amplitude of the pulses. Such signals are called piecewise deterministic signals, see Åström and Wittenmark (1984, Ch 6). This type of disturbance can be generated by

$$d(t) = H_d(q)\delta_s(t) = \frac{D_n(q)}{D_d(q)}\delta_s(t)$$

where $\delta_s(t)$ is a sequence of pulses and $H_d(q)$ a filter. For instance if $d(t)$ is an unknown level then

$$H_d(q) = \frac{q}{q-1}$$

The signal $d(t)$ can be eliminated (except for occational pulses or finite length disturbances) by filtering the input and the output by $D_d(q)$.

The disturbance annihilation filter $D_d(q)$ is of high pass or notch type. It can then be advisable to use a filter of the form

$$H_f(q) = \frac{D_d(q)}{D_d'(q)} \tag{3.1}$$

where $D_d'(q)$ is a stable polynomial and H_f has a high attenuation at high frequencies. Filtering by H_f will then remove low frequencies in the signal and eventually also some other frequencies within a narrow band. The filter $H_f(q)$ should thus have an amplitude curve as in Figure 3.1. The lower break frequency $\omega_{f\ell}$ depends on the desired cross over frequency of the closed loop system. As a rule of thumb $\omega_{f\ell}$ should be at least one decade below the cross over frequency.

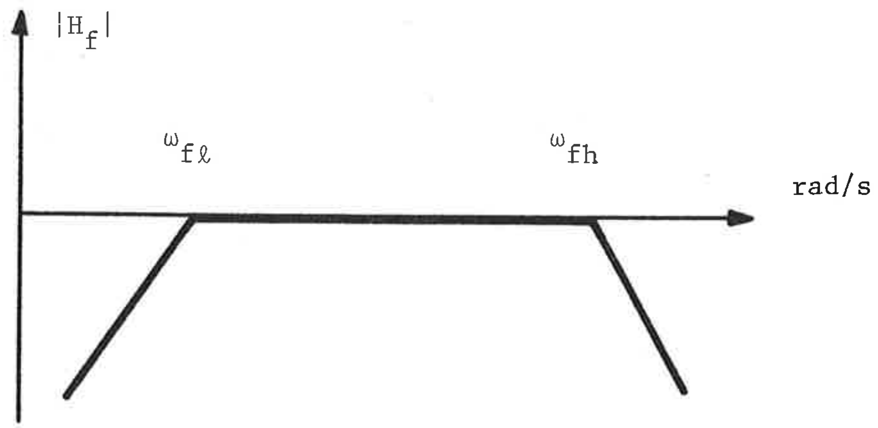


Figure 3.1 Amplitude curve for the disturbance annihilation filter $H_f(q)$.

The high frequency break frequency will be discussed later in this section. In several references it is pointed out that it is advantageous for the estimation to eliminate levels in the signals. This is the case when $d(t)$ is constant. See for instance Peterka (1984). These recommendations are thus followed if the filter H_f is used.

With the filter H_f we will remove $d(t)$ from the data. The estimator will thus not be confused by low frequency drift or disturbances at specific frequencies. These disturbances are instead taken into consideration when the controller is designed by using the internal model principle, see for instance Garcia and Morari (1982). This implies that a level is removed by having an integrator in the controller and so on. See Section 5. Notice that the regulator is still using the unfiltered sampled signal.

<u>The bias distribution problem</u>

In the following it is assumed that $d(t) \equiv 0$ in (2.1) or that this disturbance has been properly eliminated by the filter H_f in (3.1).

In the remaining part of this section we will analyze how the asymptotic estimates of the parameters in the process are influenced by design variables such as

- input signal and reference signal spectrum
- model orders and noise structure
- prefilters
- desired closed loop performance

We will follow the lines of thought for open loop estimation given in Wahlberg and Ljung (1986). Their results are modified to suit the adaptive control problem. One result in this section is a limiting expression for the loss function that is minimized by the estimator. This result can then be used for a more quantitative discussion <u>in the frequency domain</u> of the influence of the design variables. Similar aspects are also discussed in Gevers and Ljung (1986).

The purpose of the estimator is to provide an estimate of the process and eventually also the noise characteristics. Let the prejudice model be described by

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{3.2}$$

where the pulse transfer operators G and H are parameterized by the unknown parameters θ. The structure such as model orders and time-delays are parts of the parameterization. The model (3.2) includes for instance ARMAX models of different orders.

The estimation is done by comparing the model with observed data. The one step ahead prediction of $y(t + h)$ can be written as

$$\hat{y}(t+h|t, \theta) = H(q,\theta)^{-1} G(q,\theta) u(t+h) + (1 - H(q,\theta)^{-1})y(t+h) \tag{3.3}$$

Notice that the right hand side is a function of data observed up to and including time t. Define the prediction error

$$\begin{aligned} \epsilon(t + h, t, \theta) &= y(t + h) - \hat{y}(t + h|t, \theta) \\ &= (G_o(q) - G(q,\theta))H(q,\theta)^{-1} u(t + 1) \\ &\quad + (H_o(q) - H(q,\theta))H(q,\theta)^{-1}e(t+1) + e(t+1) \\ &= \Delta G(q,\theta)H(q,\theta)^{-1}u(t+1) \\ &\quad + \Delta H(q,\theta)H(q,\theta)^{-1}u(t+1) + e(t+1) \end{aligned} \tag{3.4}$$

and the filtered prediction error

$$\epsilon_r(t,t-h,\theta) = H_p(q)\epsilon(t,t-h,\theta) \tag{3.5}$$

where $H_p(q)$ is a stable causal prefilter. The estimate is now obtained by using the criterion

$$V_N(\theta) = \frac{1}{N}\sum_{i=1}^{N} \epsilon_f(ih,(i-1)h,\theta)^2 \tag{3.6}$$

and the estimate is given by

$$\hat{\theta}_N = \arg\min_{\theta} V_N(\theta) \tag{3.7}$$

The estimate $\hat{\theta}_N$ in (3.7) is influenced in a complex way by many factors, design variables, such as sampling period, model structure prefilter and number of observations. To be able to draw any conclusions we will first derive the limiting pulse transfer operator estimate as the number of data tends to infinity.

Introduce the weighting function

$$W(\omega,\theta) = \left| \frac{H_p(e^{i\omega h})}{H(e^{i\omega h},\theta)} \right|^2 \tag{3.8}$$

In Gevers and Ljung (1986) it is shown that the limiting estimate as $N \to \infty$ is

$$\theta^{*} = \arg\min_{\theta} \overline{V}(\theta)$$

where

$$\overline{V}(\theta) = \int_{-\pi/h}^{\pi/h} B^T(e^{i\omega h},\theta)\phi(\omega)B(e^{-i\omega h},\theta)W(\omega,\theta)d\omega \tag{3.9}$$

where

$$B(e^{i\omega h},\theta) = \begin{bmatrix} G(e^{i\omega h},\theta) - G_o(e^{i\omega h}) \\ H(e^{i\omega h},\theta) - H_o(e^{i\omega h}) \end{bmatrix} \tag{3.10}$$

and

$$\phi(\omega) = \begin{bmatrix} \phi_u(\omega) & \phi_{ue}(\omega) \\ \phi_{eu}(\omega) & \phi_e(\omega) \end{bmatrix}$$

where $\phi_u(\omega)$ is the input spectrum, $\phi_{ue}(\omega)$ is the cross-spectrum between the input and the noise e and $\phi_e(\omega) = \sigma^2$. The conditions for the excistence of ϕ_u and ϕ_{ue} are given in Gevers and Ljung (1986).

In Wahlberg and Ljung (1986) it is assumed that $\phi_{uv}(\omega) \equiv 0$, since the estimation is done in open loop. In the adaptive case we must take into consideration that the input signal is generated through feedback.

The simplicity of (3.9) is misleading since it is in fact a complicated function of θ and the design variables. Qualitatively we can say that the estimate is weighted in frequency. This implies that the estimates will depend on the experiment conditions. To be able to draw further conclusions it is necessary to be more specific with respect to model structure and feedback.

If the parameterization is such that the true process model $\{G_o(q), H_o(q)\}$ is among the model set $\{G(q,\theta), H(q,\theta)\}$ then the estimate will under weak conditions converge to the true parameters independent of the filtering and the frequency dependence of the weighting function $W(\omega,\theta)$. However if the model set does not include the true system then the estimate will be biased. The bias distribution will be a function of the design variables. We refer to Wahlberg and Ljung (1986) for a thorough discussion of this important property for open loop estimation.

Mannerfelt (1981) discusses estimation of simple parametric models for high order systems. Equation (3.9) is derived for open loop least squares estimation. It is for instance shown that if the input spectrum ϕ_u has a discrete spectrum with n delta functions where n is the number of parameters in $G(q,\theta)$ then there exists a unique solution θ^* to the estimation problem such that

$$G(e^{i\omega_j h},\theta^*) = G_o(e^{i\omega_j h})$$

where ω_j are the frequencies of the spectrum of ϕ_u. The Nyquist curves of the true system and the model will thus coincide at the frequencies of the input spectrum. This gives an indication of the importance of the frequency content in the input signal. Too high frequencies in the input signal will give good models in the "wrong" frequency band. The important frequency band is determined by the design method for the controller and the specifications of the closed loop system.

<u>Discussion of the choice of prefilter</u>

To get further insight into the bias distribution problem and guidance to choose the prefilter we make the following assumptions:

- The true process is described by

$$A_o(q)y(t) = B_o(q)u(t) + C_o(q)\ e(t) \tag{3.10}$$

where $e(t)$ is white Gaussian noise with variance σ^2.

- The least squares (LS) or maximum likelihood (ML) methods are used for the estimation. This implies that the model is

$$\begin{aligned} A(q)y(t) &= B(q)u(t) + e(t) \qquad &(LS) \\ A(q)y(t) &= B(q)u(t) + C(q)\ e(t) \qquad &(ML) \end{aligned} \tag{3.11}$$

and

$$H = 1/A \qquad (LS)$$

$$H = C/A \qquad (ML)$$

- The input signal to the process is generated by the <u>fixed</u> control law

$$R(q)u(t) = - S(q)y(t) + T(q)y_r(t) \tag{3.12}$$

Using (3.12) in (3.10) gives the closed loop system

$$y(t) = \frac{B_oT}{A_oR + B_oS} y_r(t) + \frac{C_oR}{A_oR + B_oS} e(t)$$

$$= \frac{B_m}{A_m} y_r(t) + \frac{C_oR}{A_m} e(t)$$

where we have assumed that the controller is chosen such that $G_m(q) = B_m(q)/A_m(q)$ is the desired pulse transfer operator from y_r to y. We are thus assuming that the control signal has the optimal (= desired) properties, compare Quote 1.2. Simple calculations now give

$$u(t) = \frac{A_o}{B_o}\frac{B_m}{A_m} y_r(t) - \frac{C_o S}{A_m} e(t)$$

If we assume that y_r is independent of e then the input spectrum is

$$\phi_u(\omega) = | G_o^{-1}(e^{i\omega h}) G_m(e^{i\omega h})|^2 \phi_{y_r}(\omega)$$

$$+ \left| \frac{C_o(e^{i\omega h}) S(e^{i\omega h})}{A_m(e^{i\omega h})} \right|^2 \sigma^2$$

Further

$$\phi_{ue}(\omega) = - \frac{C_o(e^{i\omega h}) S(e^{i\omega h})}{A_m(e^{i\omega h})} \sigma^2$$

For the least squares case we now have

$$B(e^{i\omega h},\theta) = \begin{bmatrix} \frac{B}{A} - G_o \\ \frac{1}{A} - H_o \end{bmatrix}$$

$$W(e^{i\omega h},\theta) = |H_p(e^{i\omega h}) A(e^{i\omega h})|^2$$

The unknown parameters are the coefficients in the A and B polynomials. The expression for (3.9) is still quite complex and we will first consider the noise free case ($\sigma = 0$). Then

$$\bar{V}(\theta) = \int_{-\pi/h}^{\pi/h} |AG_o - B|^2 \; |G_o^{-1} \; G_m|^2 \; |H_p|^2 \phi_{yr}(\omega)d\omega \qquad (3.13)$$

The part $|G_o^{-1} \; G_m H_p|^2 \phi_{yr}$ is a frequency dependent weighting, which determines the frequency band where the estimate should be good.

The desired closed loop transfer operator G_m is in general constant for low frequencies and the gain decreases for frequencies above the bandwidth. Further G_o is in general of low pass character. This implies that G_o^{-1} will be of high pass character.

For the design of controllers it is important to have a good model around the cross over frequency. This implies that the prefilter H_p should be chosen such that the weighting is largest around the desired cross over frequency and low at frequencies above the bandwidth. H_p should thus in general be of low pass type or band pass type with the pass band around the desired cross over frequency. If $H_p = 1$ then the estimator will concentrate on too high frequencies, which is a well known property of least squares estimation, see Ljung and Söderström (1983).

From (3.9) it is seen that the noise will introduce a bias and more weight on higher frequencies. This makes it even more important that a proper weighting is introduced through prefiltering of the data before it is used in the estimator. It is also seen that the desired bandwidth of the closed loop system will influence the bias through G_m and A_m. A high desired bandwidth implies that the model must be accurate over a large frequency band. It may then be necessary to increase the order of the model. Finally increased input power or excitation will decrease the relative influence of the noise terms. This again in accordance with rules of thumb in identification.

Direct or indirect adaptive algorithms

In the literature it is sometimes, without any quantitative results, argued that direct adaptive controllers are more robust than indirect methods. One reason why this may be correct is given in Quote 1.1. Direct methods are essentially based on prediction of the output and direct estimation of the parameters in the predictor. Prediction error methods are robust even for low order models. Also numerical difficulties in the design calculations are avoided in direct methods.

Some direct methods are, however, based on cancellation of process zeros, for instant the basic self-tuning regulator in Åström and Wittenmark (1973). These controllers are not suitable for nonminimum phase systems. A modification of the basic self-tuner is given in Åström and Wittenmark (1985) where moving average controllers are obtained by increasing the prediction horizon. There will not be any zero cancellations if the prediction horizon is increased sufficiently.

Examples

A couple of examples will illustrate the analysis given above.

Example 3.1 - Influence of desired closed loop bandwidth.

The example is taken from Wittenmark and Åström (1980). The process is a fourth order system.

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

The process is identified as a second order sampled data system using least squares estimation without any prefiltering of the signals. The controller is of the form (3.12) and has the same structure as a digital PID-controller. The desired closed loop response is given as the model

$$G_m(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

with $\zeta = 0.7$. The system and the controller was simulated for different values of ω. Figure 3.2 shows the output and the control signal when the parameters have converged. For $\omega \leq 0.3$ the control is good. The behaviour starts to deteriorate when ω is increased to 0.4 and 0.45. The reason is the difficulty for the estimator to find a good second order model of the process in the desired frequency band. The example shows that too simple models can restrict the performance of the system, due to modelling errors.

∇∇∇

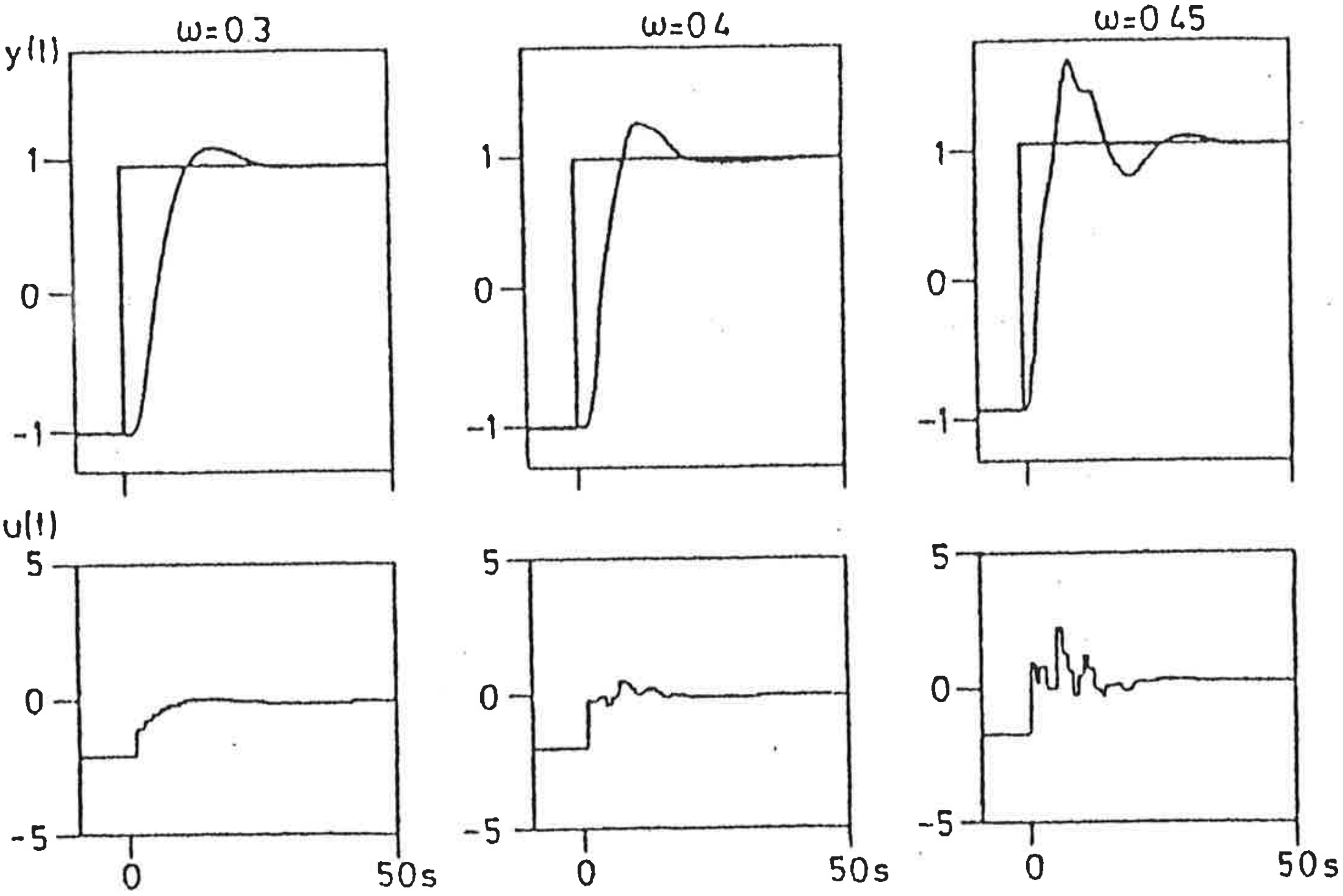


Figure 3.2: The output and the control signals for a fourth order system when ω = 0.3, 0.4 and 0.45. For ω = 0.3 the control is good but the performance starts to deteriorate when ω is increased.

Summary

The aim of the signal processing is to guarantee that the true process can be well approximated by the prejudice model such as (3.11) within the desired bandwidth of the closed loop system.

To summarize we can draw the following conclusions from the discussion above.

- Eliminate known types of disturbances such as $d(t)$ in (2.1) using high pass or notch filters. The effect of the disturbance $d(t)$ is then eliminated through the choice of the regulator structure. Compare Middleton et al (1985).
- Use a prefilter H_p such that the process can be well described by the model within the desired frequency band. H_p should be of low pass or band pass type.
- Increase in the desired closed loop performance in general implies that higher order models should be used.
- Estimate only when there is useful information in the system i.e. when the system is excited, compare Åström (1983a). This point is further discussed in Section 4.

4. THE ESTIMATOR ALGORITHM

The previous section discussed the shaping of the data to adjust it to the prejudice model. In this section we will discuss implementation of recursive estimation algorithms. We will primarily discuss prediction error methods for the least squares or maximum likelihood models (3.11). The process model is now written as

$$y(t) = \varphi^T(t)\theta + \epsilon(t) \tag{4.1}$$

where

$$\varphi^T(t) = [-y(t-1), \ldots, -y(t-n_y)\ u(t-d), \ldots, u(t-d-n_b)\ \hat{\epsilon}(t-1), \ldots, \hat{\epsilon}(t-n_c)]$$

$$\theta^T = [a_1 \ldots a_{n_a}\ b_o \ldots b_{n_b}\ c_1 \ldots c_{n_c}]$$

The variables a_i, b_i and c_i are the coefficients of the A, B and C polynomials respectively. Further

$$\hat{\epsilon}(t) = y(t) - \varphi^T(t)\hat{\theta}(t)$$

If c_i and $\hat{\epsilon}$ are omitted from θ and $\varphi(t)$ respectively we get the least squares model.

The most widespread way to estimate θ is defined by the least squares estimate

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)(y(t) - \varphi^T(t)\hat{\theta}(t-1)) \tag{4.2}$$

$$K(t) = P(t-1)\varphi(t)(\lambda + \varphi^T(t)P(t-1)\varphi(t))^{-1} \tag{4.3}$$

$$P(t) = [P(t-1) - P(t-1)\varphi(t)(\lambda + \varphi^T(t)P(t-1)\varphi(t))^{-1}\varphi^T(t)P(t-1)]/\lambda$$

$$= (I - K(t)\varphi^T(t))P(t-1)/\lambda \tag{4.4}$$

We have also introduced the exponential forgetting factor λ. It is assumed that $0 < \lambda \leq 1$. If $\lambda < 1$ then there will be an exponential weighting of old data and the estimator will minimize the loss function

$$\sum_{i=1}^{N} \lambda^{N-i} \epsilon(i)$$

If $\lambda = 1$ then $P(t) \to 0$ as the number of data points is increased. This is not a desirable feature in an adaptive regulator since the estimator will then not be able to follow changes in the process.

The drawback with a constant forgetting factor as in (4.3) and (4.4) is that the estimator has a fading memory. As a rule of thumb we can say that the memory window is

$$N_{mem} = \frac{2}{1-\lambda}$$

Datapoints further away are weighted by less than 0.1.

Exponential forgetting works well only if the process is properly excited all the time. There are several problems with exponential forgetting when the excitation of the process changes. A typical situation is when the main source of excitation is changes in the set point. There may then be long periods with no excitation and the P matrix will grow. This can be called <u>estimator windup</u> (compare with integrator windup in conventional integral control). The problem can be understood from (4.4). The negative term on the right hand side represents the reduction in uncertainty due to the last measurement. If there is no information in the last measurement then $P(t-1)\varphi(t)$ will be zero and (4.4) reduces to

$$P(t+1) = P(t)/\lambda$$

$P(t)$ will thus grow exponentially until φ changes if $\lambda < 1$. If there is no excitation for a long period of time then $P(t)$ may be very large. Since $P(t)$ also influences the gain in (4.3) then there may be large changes in the estimated parameters when new information is coming into the system, for instance when the reference value changes. The estimator windup may then cause a burst in the output of the process.

Other ways have therefore been suggested to tackle this problem:

- Covariance resetting, see Goodwin and Sin (1984)
- Time variable forgetting factors, see Fortescue et al (1981), Wellstead and Sanoff (1981)
- Constant trace algorithm, see Irving (1979)
- Directional forgetting, see Hägglund (1985)

The main idea with these modifications is to ensure that P stays bounded.

Following the advice in Quote 1.3 it has in practice and theory been shown that it is advantageous to introduce a deadzone in the estimator and turn it off when the residual ϵ is sufficiently small. See Egardt (1979) Kreisselmeier and Narendra (1982) and Middleton et al (1985).

One version of the regularized constant trace algorithm that has shown good properties in experiments is

$$\hat{\theta}(t) = \hat{\theta}(t-1) + a(t)K(t)(y(t) - \varphi^T(t)\hat{\theta}(t-1)) \tag{4.5}$$

$$K(t) = P(t-1)\varphi(t)(1 + \varphi(t)^T P(t-1)\varphi(t) + \bar{c}\,\varphi(t)^T\varphi(t))^{-1} \tag{4.6}$$

$$\bar{P}(t) = \bar{P}(t-1) - a(t)\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1 + \varphi(t)^T P(t-1)\varphi(t) + \bar{c}\,\varphi(t)^T\varphi(t)} \tag{4.7}$$

$$P(t) = c_1\frac{\bar{P}(t)}{tr(\bar{P}(t))} + c_2 I \tag{4.8}$$

$$a(t) = \begin{cases} \bar{a} & |y(t) - \varphi^T(t)\hat{\theta}(t-1)| > 2\Delta \\ 0 & \text{otherwise} \end{cases} \tag{4.9}$$

where $c_1 > 0$, $c_2 \geq 0$, $\bar{c} \geq 0$ and Δ is an estimate of the magnitude of the noise. Typical values for the parameters can be, Middleton (1986).

$$\bar{a} \in [0.1, 0.5]$$
$$c_1/c_2 \approx 10^4$$
$$\varphi^T\varphi \cdot c_1 >> 1$$
$$\varphi^T\varphi\bar{c} \approx 1$$

for typical values of $\varphi^T\varphi$.

If c_2 is used in (4.8) then it is not necessary to use $\bar{c}$.

A new approach to the problem of estimator windup is given in Hägglund (1985). The main idea is to forget information only in the directions in which new information is gathered. The following estimator is then obtained.

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)(y(t) - \varphi(t)^T\hat{\theta}(t-1)) \tag{4.10}$$

$$K(t) = \frac{P(t-1)\varphi(t)}{v(t) + \varphi(t)^T P(t-1) \varphi(t)(1-\alpha(t)v(t))} \tag{4.11}$$

$$P(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi(t)^T P(t-1)}{[v(t)^{-1} - \alpha(t)]^{-1} + \varphi(t)^T P(t-1)\varphi(t)} \tag{4.12}$$

$$\alpha(t) = \begin{cases} 0 & \alpha_d \leq 0 \\ \alpha_d & 0 < \alpha_d \leq \frac{1}{\varphi^T P\varphi} \\ \frac{1}{\varphi^T P\varphi} & \frac{1}{\varphi^T P\varphi} < \alpha_d \leq v^{-1} + \frac{1}{\varphi^T P\varphi} \\ 0 & \alpha_d > v^{-1} + \frac{1}{\varphi^T P\varphi} \end{cases} \tag{4.13}$$

$$\delta_d(t) = \frac{\dfrac{\varphi(t)^T P(t-1)P(t-1)P(t-1)\varphi(t)}{\varphi^T(t)P(t-1)P(t-1)\varphi(t)} - a}{\varphi(t)^T P(t-1)^2\varphi(t)} \tag{4.14}$$

$$\alpha_d(t) = v(t)^{-1} + \frac{\delta_d(t)}{\delta_d(t)\ \varphi(t)^T P(t-1)\varphi(t) - 1} \tag{4.15}$$

The estimates will then converge to values such that $P(t) = a \cdot I$ where a should be chosen as a small value. Finally $v(t)$ should be an estimate of the variance of ϵ in (4.1).

The covariance matrix P in (4.4), (4.7) or (4.12) should be a symmetric positive definite matrix. A straight forward implementation of the equations may lead to numerical problems in the same way as when implementing Kalman filters. The algorithm will become more numerical robust if U-D factorization or square root algorithms are used for the updating of the P-matrix, see Bieman (1977).

Example

The following example from Wittenmark and Åström (1984) illustrates the problem with estimator windup.

Example 4.1 - Estimator windup

Let the process to be controlled be described by

$$y(t) - 0.9y(t - 1) = u(t - 1)$$

It is desired that the pulse transfer operator from the reference signal to the output has a pole in 0.7 and that the gain is unity. This is achieved with the controller

$$u(t) = 0.3y(t) - 0.35y(t-1) + u_c(t) - 0.5u_c(t-1)$$

The process is controlled using an direct pole placement algorithm where the parameters in the controller are estimated. Figure 4.1 shows the diagonal elements of the P-matrix (4.4) or (4.12) when different estimation schemes are used. The reference signal is a square wave with unit magnitude and period 100 up to time 300. After that the reference signal is constant. In Figure 4.1a the estimation algorithm described by (4.2) - (4.4) is used with $\lambda = 0.99$. When the reference signal is constant and the output has settled there is no information in the measurements. The variance will then start to increase. In Figure 4.1b the estimation routine described by (4.10) - (4.15) is used. The variances of the estimates now settle on constant values and there is no estimator windup.

∇∇∇

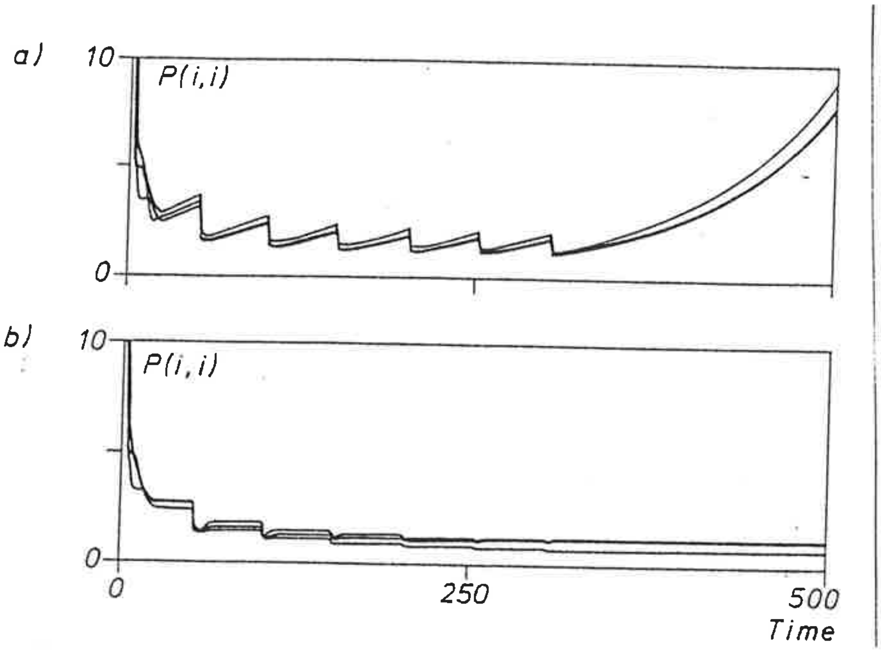


Figure 4.1: The diagonal elements of the P-matrix when controlling the process in Example 4.1 a)Constant exponential forgetting factor λ = 0.99. b)Forgetting according to Hägglund (1984).

5. SAMPLED DATA CONTROL ASPECTS

Design and implementation of digital controllers are treated in books on digital control, see for instance Chapter 15 of Åström and Wittenmark (1984). The following points are discussed in this section

- Antialiasing filters
- Choice of sampling period
- Antireset windup

Antialiasing filters

In all digital control applications it is important to have a proper filtering of the signals before they are sampled. Due to the aliasing problem connected with the sampling procedure it is necessary to eliminate all frequencies above the Nyquist frequency before sampling the signals. High frequencies may otherwise be interpreted as low frequencies and introduce disturbances in the controller. This implies that the signal conditioning is related to the choice of the sampling interval. Suitable choices of antialiasing filters are second or fourth order Butterworth, ITAE (Integral Time Absolute Error) or Bessel filters. They consist of one or two cascaded filters of the form

$$G_f(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

Let ω_B be the desired bandwidth of the filter. The damping ζ and the frequency ω should then be chosen according to Table 5.1.

Table 5.1 Damping and natural frequency of second and fourth order Butterworth, ITAE, and Bessel filters. The filters have the bandwidth ω_B

Order	Butterworth ω/ω_B	ζ	ITAE ω/ω_B	ζ	Bessel ω/ω_B	ζ
2	1	0.71	1	0.71	1.27	0.87
4	1	0.38	1.48	0.32	1.59	0.62
	1	0.92	0.83	0.83	1.42	0.96

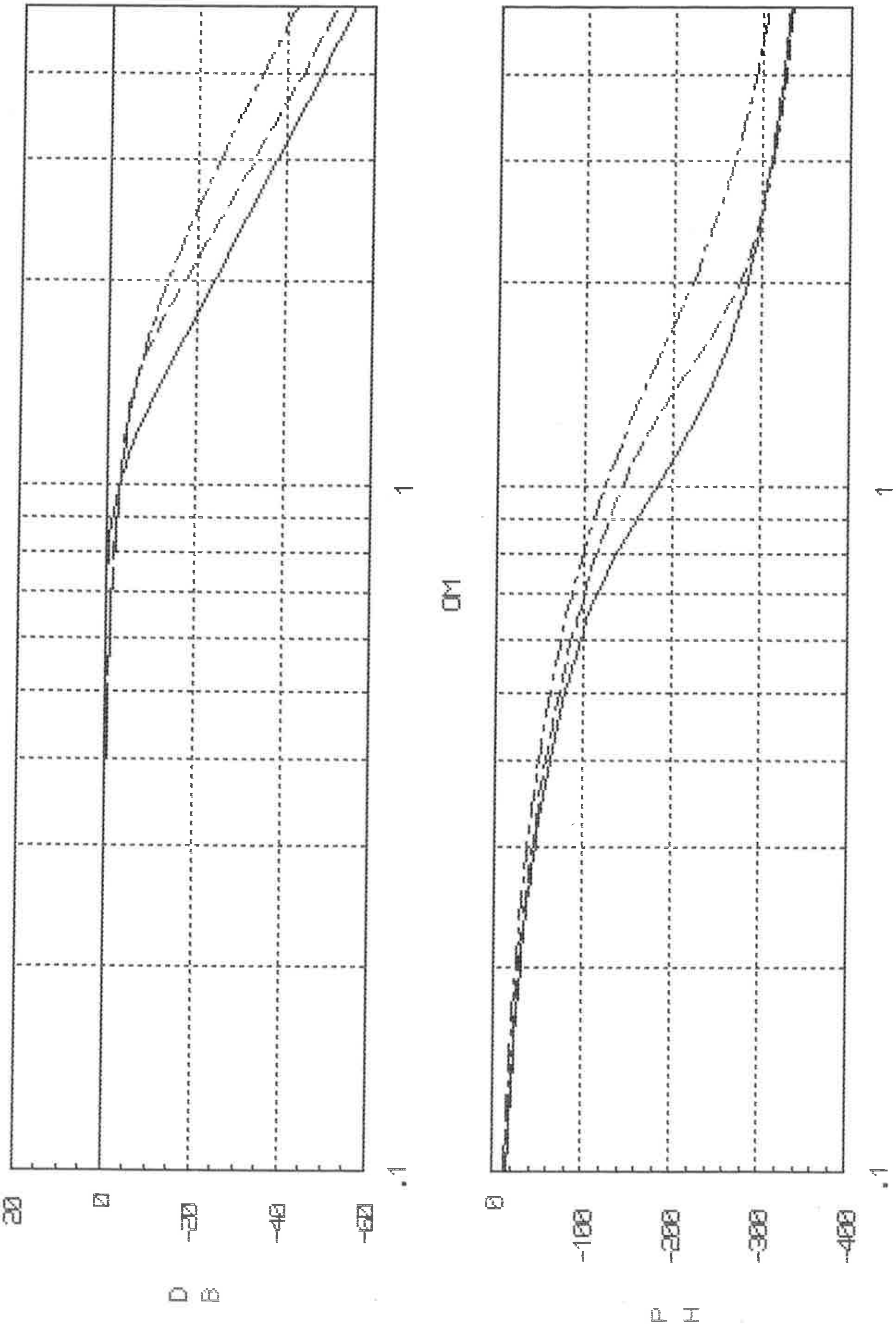
Figure 5.1 shows the Bode diagrams for fourth order Butterworth, ITAE and Bessel filters and Figure 5.2 shows the step responses for the same filters.

The antialiasing filter with then naturally low pass filter the outputs of the process. The filter will also make if possible for the estimator to get good models in correct frequency bounds. Compare the discussion of H_p in Section 3. This feature of discrete time systems may be one reason why the type of difficulties discussed by Rohrs et al (1982) were first noticed in continuous time adaptive controllers. The natural filtering obtained through the sampling helps the estimator to get better models.

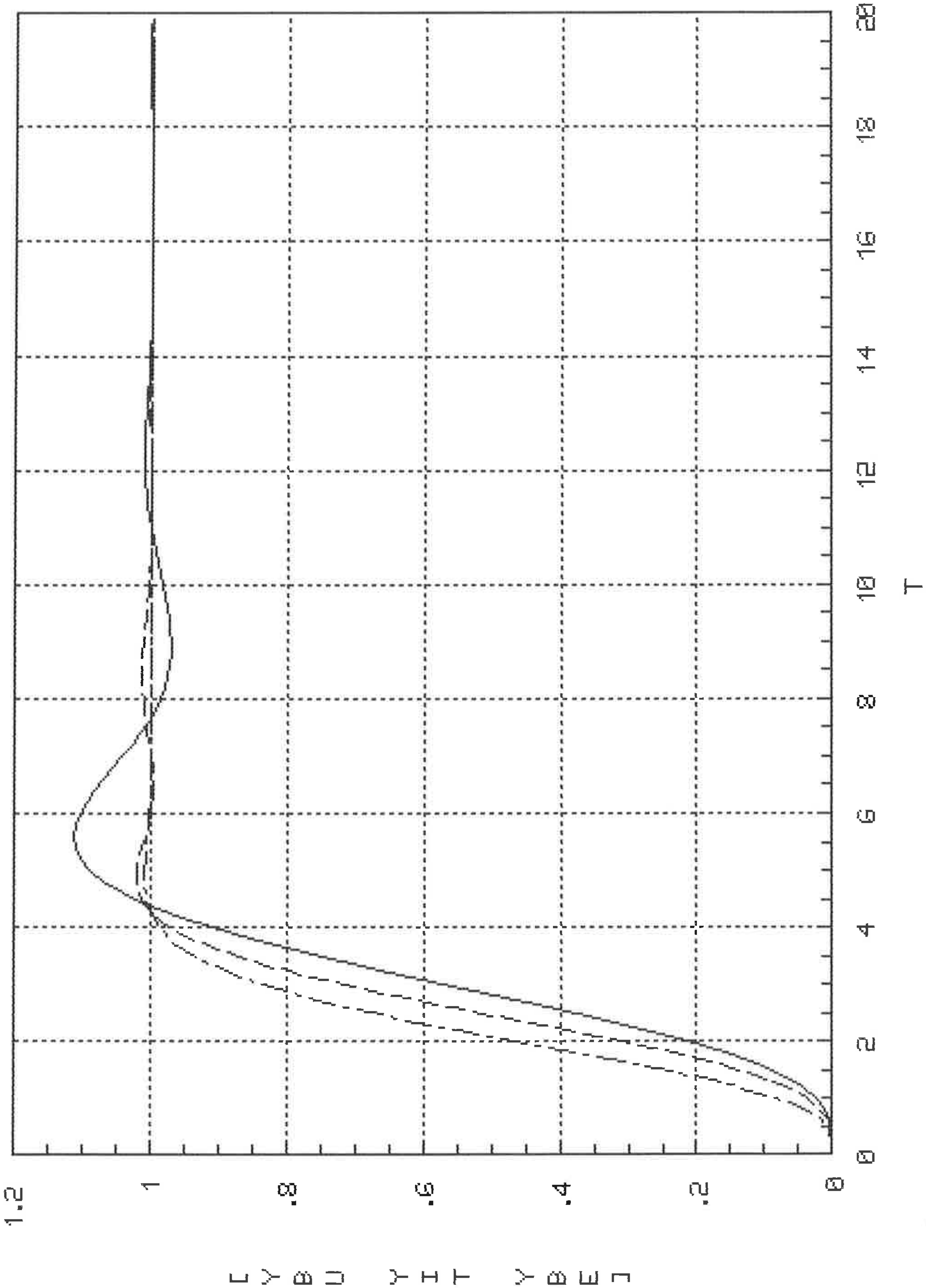
How will the antialiasing filter influence the sampled data model? The Besel filter is a good approximation of a time delay to frequencies up to around ω_B. For instance the fourth order filter can be well approximated by a time delay of

$$T_d = \frac{2\pi}{3\omega_B} \text{ [s]}$$

Figure 5.1 Bode diagrams for fourth order Butterworth, ITAE, and Bessel filters. All filters have the bandwidth $\omega_B = 1$ rad/s.



<u>Figure 5.2</u> Step responses for fourth order Butterworth, ITAE, and Bessel filters. All filters have the bandwidth ω_B = 1 rad/s.



This implies that the sampled data model can be assumed to contain additional time delay compared to the process. Assume that the bandwidth of the filter is chosen as

$$|G_{aa}(i\omega_N)| = \beta$$

where $\omega_N = \pi/h$ is the Nyquist frequency and $G_{aa}(s)$ is the transfer function of the filter. The attenuation β may be in the range 0.1 - 0.5.

Table 5.2 gives some values of T_d as a function of β.

Table 5.2: The time delay T_d as a function of the desired attenuation β at the Nyquist frequency for a fourth order antialiason Bessel filter.

β	ω_N/ω_B	T_d
0.05	3.1	2.1 h
0.1	2.5	1.7 h
0.2	2.0	1.3 h
0.5	1.4	0.9 h
0.7	1.0	0.7 h

The table shows that, for a fixed sampling period or Nyquist frequency, the delay will increase if a higher attenuation is desired. This will also imply that the antialiasing filter must be taken into account when estimating the process model and in the design. For reasonable values of β we can approximate the filter with a delay of one or two sampling periods. Sampled data models of systems with time delays are discussed in Åström and Wittenmark (1984). It is shown that the sampled data model then is such that the A and B polynomials are of the same order. (Opposed to systems without time delays where the order of B is one less than the order of A.) The price for having the antialiasing filter is that one extra parameter has to be estimated in the model. However, if the desired closed loop bandwidth is less than 0.05 - 0.1 times the band-width of the antialiasing filter than the filter will have very little influence on the process model around the desired cross over frequency.

In many process industry applications it can be sufficient to have long sampling periods (1-30 minutes). A corresponding antialising filter will then have quite large component values. This problem can be avoided by making a faster sampling, with an appropriate antialising filter, and then filter the sampled signals digitally to remove frequencies above the Nyquist frequency for the control signal.

Example 5.1 - The effect of the antialiasing filter.

Let the process to be controlled by given by

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

which can be a model for a normalized motor. The system is sampled with $h = 0.5$ and the desired closed loop system is assumed to be described by $\omega_m = 1$ rad/s and $\zeta_m = 0.7$. A controller of the form (3.12) is then determined. The output of the system is disturbed by a sinusoidal signal i.e. that the measured signal is

$$y_m(t) = y(t) + a_d \sin(\omega_d t)$$

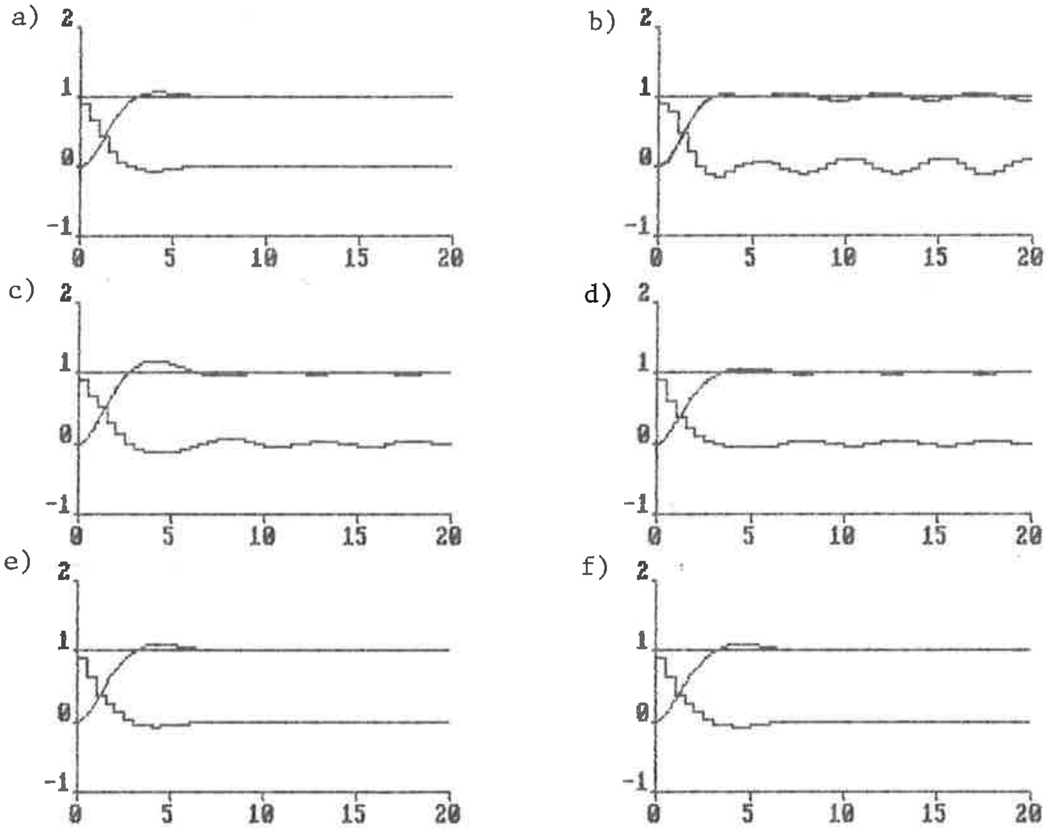


Figure 5.3 Output, reference value and control signal for the system in Example 5.1.

a) $a_d = 0$, $\omega_B = 25$ rad/s b) $a_d = 0.1$, $\omega_B = 25$ rad/s

c) $a_d = 0.1$, $\omega_B = 6.28$ rad/s d) $a_d = 0.1$, $\omega_B = 6.28$ rad/s and the regulator compensated for a delay of 0.7 h

e) $a_d = 0$, $\omega_B = 2.51$ rad/s and the regulator compensated for a delay of 1.7 h.

f) same as e) but with $a_d = 0.1$.

This signal is filtered through a fourth order Bessel filter with the bandwidth ω_B. Figure 5.3 shows the influence of the disturbance and the antialising filter. Figure 5.3 a and b show the influence of the disturbance when $\omega_B = 25$ rad/s, which is far too high for the disturbance frequency. The Nyquist frequency is 6.28 rad/s. The disturbance frequency, which is 11.3 rad/s, is folded and gives rise to a low frequency disturbance in the sampled measurement. The controller then tries to compensate this "imaginary" frequency. Figure 5.3 c shows the effect of the antialising filter when $\omega_B = \omega_N$. The distubance is smaller but still seen in the input and the output. The filter will, however, also influence the closed loop performance. Figure 5.3 d shows the same experiment as in Figure 5.3 c but when the controller is modified. The controller is now determined for the true process and when the antialiasing filter is approximated by a delay of 0.7 h, compare Table 5.2. The closed loop response is improved compared to Figure 5.3 c. Finally Figure 5.3 e and f shows the case when $\omega_B = \omega_N/2.5$ and when the filter is approximated by a delay of 1.7 h in the design of the controller.

The experiments show that it is important to use an antialiasing filter and that the filter has to influence the design. It is, however, sufficient to approximate the filter by a time delay. This will in the adaptive case simplify the estimation and the design.

∇∇∇

Choice of Sampling Period

The choice of the sampling period depends on the design algorithm as well as of the desired closed loop performance. One rule of thumb that is useful for deterministic design methods is to let the sampling interval h be chosen such that

$$\omega_o h \approx 0.25 - 1$$

where ω_o is the natural frequency of the dominating poles of the desired closed loop system. If the dominating pole is real one can use

$$h/T_o \approx 0.25 - 1$$

where T_o is the time constant of the dominating pole. These rules imply that there will be 5-20 samples on a step response of the closed loop system.

The choice of the sampling interval must also be influenced by the disturbances acting on the system. If the process is disturbed by occationally large disturbances it may be necessary to shorten the sampling interval to ensure that the disturbance is detected and compensated for as soon as possible.

For stochastic design methods it is of importance how the performance of the closed loop system is measured. One way is to consider the output and input signal variances at the sampling instances. Another and from a practical point of view more relevant measure is to consider the average variance of the continuous time output of the process. The reason is that the variance in "steady state" is periodic over the sampling period. The periodicity is a fundamental property of sampled data systems due to the periodicity of the changes in the control signal. The variance can be large at times between the sampling instances. See de Souza and Goodwin (1984) and Lennartsson (1986). This periodical behaviour can be considered as the stochastic correspondence to hidden oscillations in deterministic sampled data control. Lennartsson (1986) discusses the choices of the design method and the sampling period. One conclusion is that it is advisable to choose a design method that has a continuous time counterpart for instance linear quadratic control. The sampling period can be chosen as the rules of thumb above for deterministic controller design with a preference for the higher limit of the sampling frequence.

The choice of the sampling interval also influences if the antialiasing filter should be taken into account in the design. First assume that the sampling frequency is high (20 - 100 times) compared with the desired closed loop bandwidth.

It is then not necessary to take the filter into account when making the design. The amplitude of the filter is about unity and the phase lag is neglectable at the cross over frequency. The influence of the filter can, however, not be neglected in the design if the Nyguist frequency is only a little bit larger than the desired cross over frequency. This in general implies that the estimated model must be of higher order than if a higher sampling rate is chosen. Compare Example 5.1. This is therefore one motivation why the sampling period should be quite short. Other aspects on the choice of the sampling interval is discussed in Goodwin (1985).

<u>Antireset windup</u>

In the discussion in Section 3 it was pointed out that unknown levels, ramps etc are taken care off by introducing integrators in the controller. An integrator is an unstable system and it may happen that the integral term in the controller can assume very large values if the control signal is limited when there still is an error. This is called reset windup or integrator windup. Special precautions must be taken in order to avoid this. Different ways are discussed in Åström and Wittenmark (1984) for different controller structures.

Consider a regulator described by

$$R^*(q^{-1})u(t) \;=\; - S^*(q^{-1})y(t) + T^*(q^{-1})y_r(t)$$

The regulator may contain unstable modes. One way to solve the reset windup problem is to rewrite the regulator by adding $A_o^*(q^{-1})u(t)$ to both sides. This gives

$$A_o^*(q^{-1})u(t) \;=\; (A_o^*(q^{-1}) - R^*(q^{-1}))\, u(t) - S^*(q^{-1})y(t) + T^*(q^{-1})y_r(t) \qquad (5.1)$$

Antiset compensation is then obtained by using

$$\begin{aligned} A_o^*(q^{-1})v(t) \;&=\; (A_o^*(q^{-1}) - R^*(q^{-1}))u(t) - S^*(q^{-1})y(t) + T^*(q^{-1})y_r(t) \\ u(t) \;&=\; \text{sat}\,(v(t)) \end{aligned} \qquad (5.2)$$

where sat(•) is the saturation function. This regulator is equivalent to (5.1) when the control signal does not saturate. A_o^* is chosen as a stable polynomial and can be interpreted as the observer dynamics of the controller. A block diagram of (5.2) is shown in Figure 5.4.

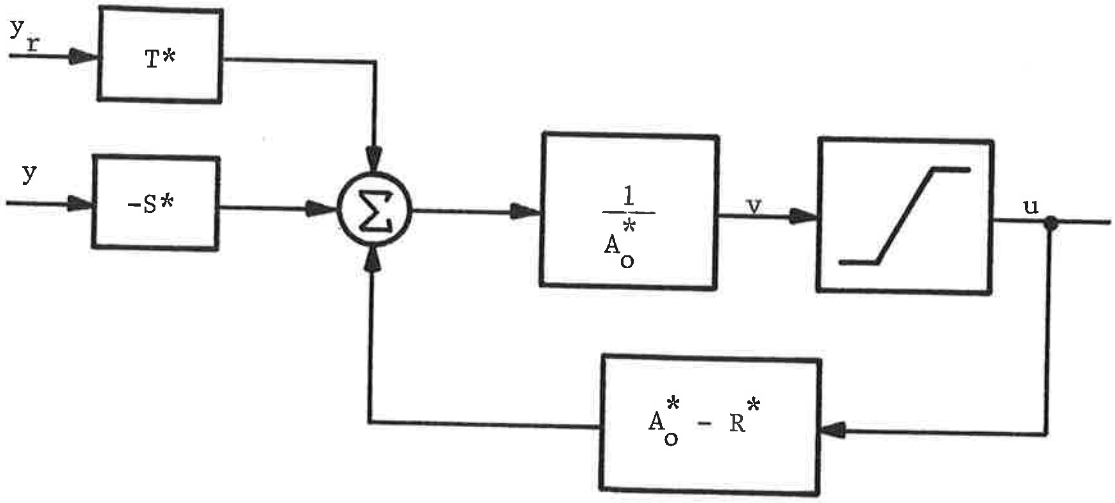


Figure 5.4 Block diagram of (5.2) which avoids integrator windup.

6. THE INTERACTION BETWEEN ESTIMATION AND CONTROL

The estimation, design and controller implementation have been discussed as seperate subjects in the previous sections. We have, however, also seen that there is a strong interaction between these parts. For instance the desired closed loop performance will influence the frequency content of the input and output signals. This will then influence the properties of the estimator. This interaction is complex as discussed above. One way to decrease the interaction is to follow the advice in Quote 1.5 and use a low gain in the adaptation loop. This will exaggerate the two time scale property indicated in Figure 2.1. A low gain in the adaptation loop makes it, however, more difficult to follow fast variations in the parameters of the process.

The estimator gain is determined by the exponential forgetting factor λ in (4.2) - (4.4), c_1 and c_2 in (4.5) - (4.8) or a in (4.10) - (4.15).

Startup procedures

There are several ways in which an adaptive algorithm can be initialized depending on the a priori information about the process. One case is if nothing is known about the process initially. The initial values of the parameters in the estimator can then be chosen to zero or such that the initial controller is a proportional or integral controller with low gain. The inputs and outputs should be scaled such that they are of the same magnitude. This will improve the numerical properties of the estimator and controller parts. The initial value of the covariance matrix P can be 1-100 times a unit matrix. These values are usually not critical since the estimator will obtain reasonable values in a short period of time. The experience is that 10-50 samples are sufficient to get a good controller. During the initial phase it can be advantageous to add a perturbation signal to speed up the convergence of the estimator.

One way to initialise an adaptive algorithm is to use the auto-tuner discussed in Åström and Hägglund (1984). The auto-tuner generates a suitable input signal and will give safe initial values for the parameters in the controller. Another situation occurs if the process has been controlled before with a conventional or an adaptive controller. The initial values should then be such that they correspond to the controller used before.

Sometimes it is important to have as small disturbance as possible at the start up of the adaptive controller. There are then two precautions that can be taken. First the estimator can be used for some sampling periods before the adaptive contoller is allowed to apply any control actions. During that period a safe simple controller should be used. It is also possible and desirable to limit the control signal. The allowable magnitude can be very small during the first period of time and can then be increased when better parameter estimaters are obtained. This kind of "soft" start-up can for instance be used in Asea's Novatune, see Bengtson and Egardt (1984). A drawback of having small input signals is that the excitation of the process will be poor and that it will take longer time to get good parameter estimates.

Influence of the design variables

In the previous sections we have tried to isolate the influences of the different design variables by either looking at the estimation or the controller design. We will now illustrate the interdependence through a simulated example.

Example 6.1

Consider again the process and output disturbance in Example 5.1. The antialiasing filter is a fourth order Bessel filter. The parameters in the process is estimated from the model

$$y(t)+a_1y(t-1)+a_2y(t-2) = b_1u(t-1)+b_2u(t-2)+b_3u(t-3)+e(t) \quad (6.1)$$

i.e. it is assumed that the antialiasing filter can be approximated by a delay that is less than one sampling period.

The controller is designed based on the estimated parameters such that the corresponding continuous closed loop system is described by the natural frequency ω_m and the damping ζ_m. Figure 6.1a shows the output and reference value when ω_B = 25 rad/s, $\omega_B = 1$, $\zeta_m = 0.7$ and $a_d = 0$. Figure 6.1b shows the influence of the disturbance. Compare Figure 5.3b. Figure. Figure 6.2 is the same as Figure 6.1 but when $\omega_B = \Pi/h = \omega_N$. Compare Figure 5.3d. These two figures show that the antialiasing filter is necessary and that it is automatically compensated for by using the model structure (6.1). Figure 6.3a shows a simulation of the adaptive controller when $\omega_B = 2.512$. The closed loop response is not satisfactory since (6.1) will not give an adequate model. To be able to use a lower ω_B it would be necessary to change (6.1) to

$$y(t)+a_1y(t-1)+a_2b(t-2)=b_2u(t-2)+b_3u(t-3)+b_4u(t-4)+e(t)$$

i.e. to introduce an extra time delay in the estimated model. Compare Table 5.2.

Figure 6.3b shows the same as Figure 6.3a but when ω_m=0.5. By decreasing the desired closed frequency it is possible to again get a good model using (6.1). This implies that there is an intricate relationship between the desired response, the model structure, and the frequency content of the input original.

∇∇∇

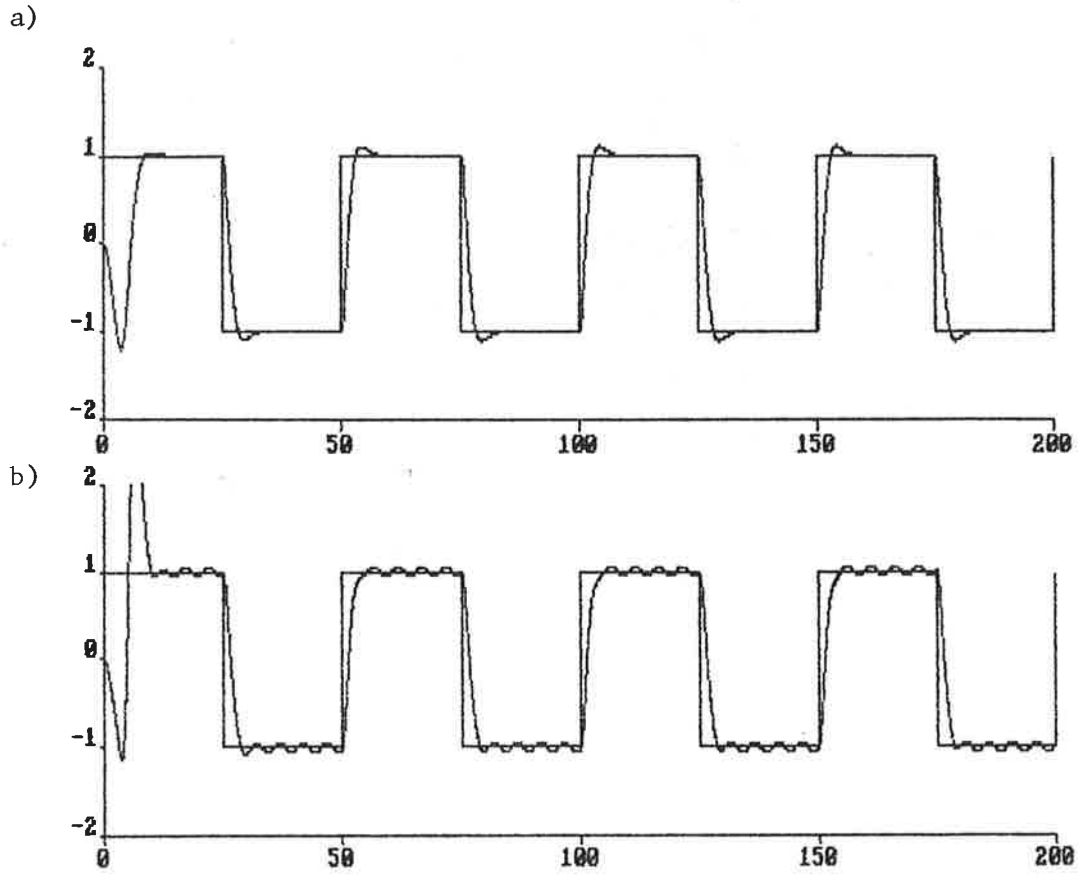


Figure 6.1 Output and reference value when using an adaptive controller on the process in Example 6.1.
a) h=0.5, ω_m=1, ω_B=25, a_d=0
b) same as (a) but with a_d=0.1

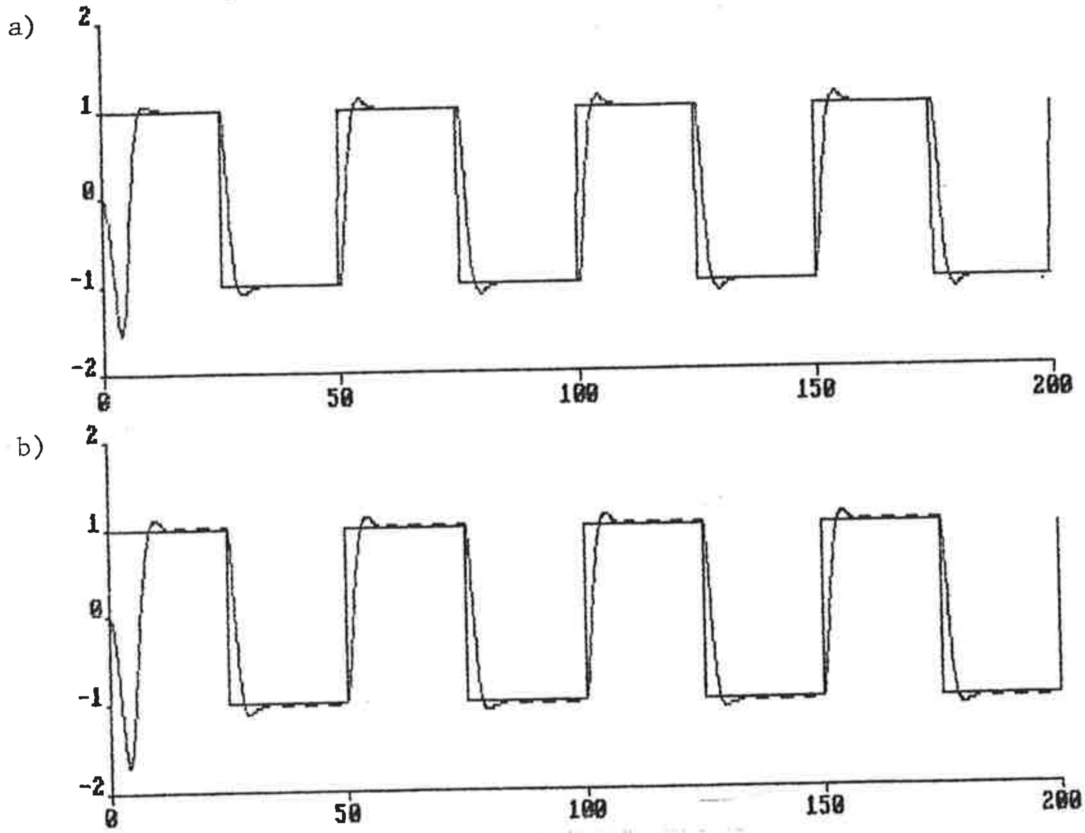


Figure 6.2 Same as Figure 6.1 but with ω_B=6.28

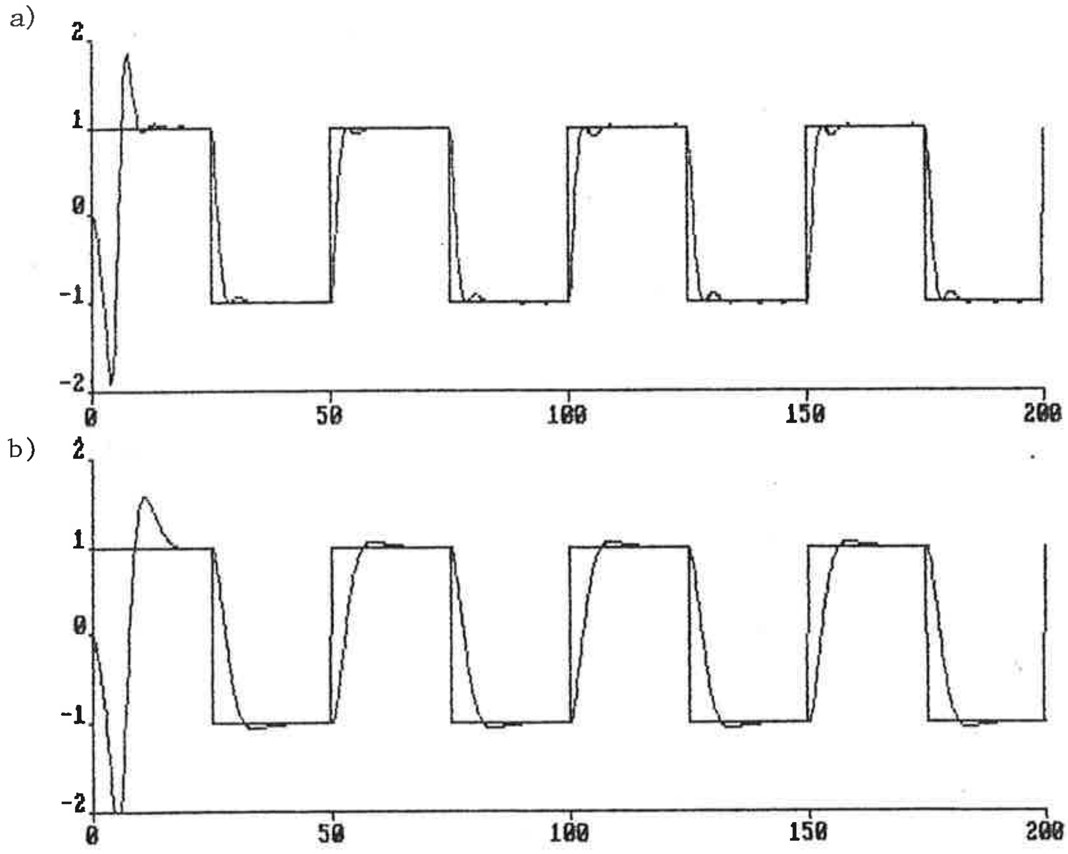


Figure 6.3 Output and reference value when using an adaptive controller on the process in Example 6.1.

a) $h=0.5$, $\omega_m=1$, $\omega_B=2.512$, $a_d=0.1$

b) Same as (a) but with $\omega_m=0.5$.

7. CONCLUSIONS

In this paper we have discussed different filters that are used in the implementation of adaptive controllers. The filters and signal processing that are used are quite different in the different parts of the controller. The filters play an important role in order to obtain a robust implementation of an adaptive controller. The algorithm will contain the following steps:

- Analog antialiasing filter
  Second or fourth order filter with bandwidth below or around the Nyguist frequency π/h, where h is the sampling period.

- High pass filtering of the sampled signal to remove low frequency disturbances such as levels and ramps. Known sinusoidals can also be removed using notch filters. The lower limit of the passband should be at least one decade below the desired cross over frequency.

- Low pass filter with H_p to get a weighting in the estimator in an appropriate frequency band.

- Estimate a low order model using an algorithm with time variable exponential forgetting, regularized constant trace or directional forgetting. The estimator should also contain a dead zone. Finally the estimator may contain a "switch", which detects if the system is sufficiently excited. The "switch" can measure the power in different frequency bands and controls if the estimator should be active or not.

- The design method for the controller should be robust against unmodelled dynamics. Level and ramp disturbances are eliminated by introducing integrators in the controller. The control signal should be limited and the controller should include antireset windup.

A more detailed block diagram with added filters is given in Figure 7.1.

All these points contain choices of parameters. The discussions in the previous sections contain guidelines that can be used for the selection of the parameters.

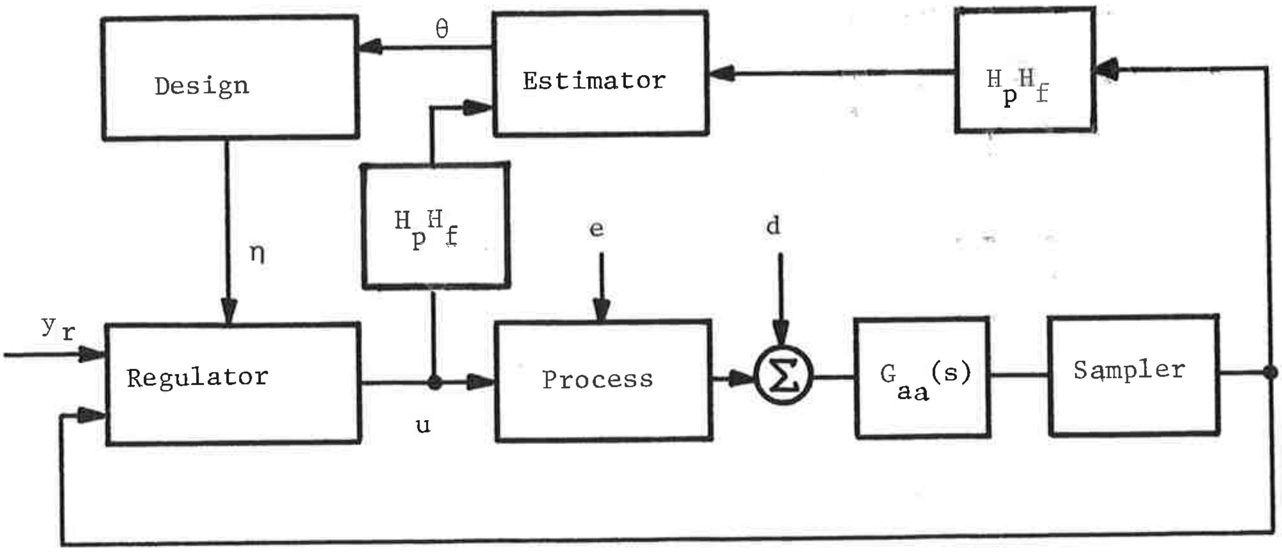


Figure 7.1 An adaptive control system with added filters. θ are the estimated parameters and η the controller parameters.

REFERENCES

Åström, K.J. (1983a): Analysis of Rohrs counter examples to adaptive control. Reprints 22nd IEEE Conference on Decision and Control, San Antonia TX, 982-987.

Åström, K.J. (1983b): Theory and applications of adaptive control - A survey, Automatica, 19, 471.

Åström, K.J. and T. Hägglund (1984): Automatic tuning of simple regulators with specifications on phase and amplitude margins. Automatica, 20, 645 - 651.

Åström, K.J. and B. Wittenmark (1973): On Self-tuning regulators. Automatica, 9, 185-199.

Åström, K.J. and B. Wittenmark (1980): Self-tuning controllers based on pole-zero placement. Proc. IEEE, Part D, 128, 120-130.

Åström, K.J. and B. Wittenmark (1984): Computer Controlled Systems, Prentice Hall Inc., Englewood Cliffs N.J.

Åström, K.J. and B. Wittenmark (1985): The self-turning regulator re-visited. Preprints 7th IFAC Symposium on Identification and System Parameter estimation, York, UK.

Bengtson G. and B. Egardt (1984): Experience with self-tuning control in the process industry. Preprints 9th World IFAC Congress, Budapest, Paper CS-7.

Bierman, G.J. (1977): Factorization Methods for Discrete Sequential Estimation, Academic Press, New York.

de Souza, C.E., G.C. Goodwin, D.Q. Mayne and M. Palaniswami (1986): An adaptive control algorithm for linear systems having unknown time delay. Technical Report EE88626. Department of Electrical and Computer Engineering, University of Newcastle, New South Wales.

de Souza C.E., and G.C. Goodwin (1984): Intersample variances in discrete minimum variance control. IEEE Trans Automatic Control, AC-29, 759 - 761.

Egardt B., (179): Stability of Adaptive Controllers, Springer-Verlag, New York.

Fortescue T.R., L.S. Kershenbaum and B.F. Ydstie (1981): Implementation of self-tuning regulators with variable forgetting factors, Automatica, 17, 831-835.

Garcia C.E., and M. Morari (1982): Internal model control I - A unifying review and some new results. Ind. Engineering Chem. Proc. Des. Dev. 1982, 21, 308.

Gevers M. and L. Ljung (1986): Optimal experiment design with respect to the intended model application. Automatica, 22, 543-555.

Goodwin G. C. (1985): Some observations on robust estimation and control. Preprints 7th IFAC Symposium on Identification and System Parameter Estimation. York.

Goodwin G.C., and K.S. Sin (1984): Adaptive Filtering, Prediction and Control, Prentice-Hall Inc., Englewood Cliffs, N.J.

Hägglund T. (1985): Recursive estimation of slowly time varying parameters. Preprints 7th IFAC Symposium on Identification and System Parameter Estimation, York, 1137-1142. Accepted for publication in Automatica.

Irving E., (1979): New developments in improving power network stability with adaptive control. Yale Workshop on Applications of Adaptive Systems.

Kreisselmeier G., and K.S. Narendra (1982): Stable model reference adaptive control in the presence of bounded disturbances. IEEE Trans. Aut. Control, AC-27, 1169-1175.

Ljung L and T. Soderstrom (1983): Theory and Practice of Recursive Identification, MIT Press, Cambridge, MA.

Lennartsson B. (1986): On the design of stochastic control systems with multirate sampling. Ph.D. Thesis, Control Engineering Laboratory, Chalmers University of Technology, Gothenburg, Sweden.

Mannerfelt, C.F., (1981): Robust Control Design with Simplified Models, TFRT-1021. Ph.D. Thesis, Lund Institute of Technology.

Middleton, R.H. (1986): Private Communication.

Middleton, R.H., G.C. Goodwin, D.J. Hill and D.Q. Mayne (1985): Robust adaptive control (Convergence, stability and performance). Technical report EE8544. Department of Electrical and Computer Engineering, University of Newcastle, New South Wales.

Peterka V. (1984): Predictor-based self-tuning control, Automatica, 20, 39-50.

Rohrs, C.E., M. Athans, L. Valavani and G. Stein (1984): Some design guide lines for discrete-time adaptive controllers. Automatica, 20, 653-660.

Rohrs, C.E., L. Valavani, M. Athans and G. Stein (1982): Robustness of adaptive control algorithms in the presence of unmodelled dynamics. Preprints 21st IEEE Conference on Decision and Control, Orlando, FL, 3-11.

Seborg, D.E., T.F. Edgar and S.L. Shah (1986): Adaptive control strategies for process control : A survey . AIChE Journal, 32, 881-913.

Wahlberg B., and L. Ljung (1986): Design variables for bias distribtuion in transfer function estimation. IEEE Trans. on Automatic Control, AC-31, 134-144.

Wellstead, P.E., and S.P. Sanoff (1981): Extended self-tuning algorithm. Int. J. Control, 345, 433-455.

Wittenmark, B., and K.J. Åström (1980): Simple self-tuning controllers. International Symposium on Adaptive Control, Bochum.

Wittenmark, B., and K.J. Åström (1984): Practical issues in the implementation of adaptive control. Automatica, 20, 595-605.