Identification of Continuous Time Dynamic Systems

Johansson, Rolf

1986

CODEN: LUTFD2/(TFRT-7318)/1-29/(1986)

Identification of Continuous Time Dynamic Systems

Rolf Johansson

Department of Automatic Control
Lund Institute of Technology
March 1986

LUND INSTITUTE OF TECHNOLOGY
DEPARTMENT OF AUTOMATIC CONTROL
Box 118
S 221 00 Lund Sweden

Document name
Technical report

Date of issue
March 1986

Document number
CODEN:LUTFD2/TFRT-7318/1-29/1986

Author(s)

Rolf Johansson

Supervisor

Sponsoring organization

NFR-CNRS

Title and subtitle

IDENTIFICATION OF CONTINUOUS TIME DYNAMIC SYSTEMS

Abstract

It is sometimes of interest to monitor on-line a certain physical parameter represented as some coefficient of a differential equation

$$p^n y(t)+a_1 p^{n-1} y(t)+..+a_n y(t) = b_1 p^{n-1} u(t)+..+b_n u(t); \quad p=\frac{d}{dt}$$

when only the system input u and the output y can be measured. Traditional discrete time identification is applicable but is less attractive because it results in sampling time dependent nonlinear transformation of the coefficients. Earlier state variable filter based approaches to identification seem superior. This paper shows that the operator transformation $z=1/(1+p\tau)$ gives an exact reciprocal model

$$y(t)+\alpha_1 \left[zy(t)\right]+..+ \alpha_n \left[z^n y(t)\right] = \beta_1 \left[zu(t)\right]+..+ \beta_n \left[z^n u(t)\right]$$

where α_i, β_i; i=1,..,n are known linear combinations of a_i, b_i. Sampling of this linear model allows e.g. estimation of α_i, β_i and consequently also a_i, b_i. A state space analysis is made where transformation matrices for parameters and state vectors are given.

Key words

Adaptive control, convergence, estimation, identification

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title

ISBN

Language
English

Number of pages
29

Security classification

Recipient's notes

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.

INTRODUCTION

Identification and parameter estimation of dynamic systems are fundamental for all model based design of control systems. The methodology of identification is often very different in cases of on-line and off-line estimation.

Many excellent methods of off-line estimation like frequency analysis, transient analysis, and some off-line methods like maximum-likelihood identification are not directly applicable for on-line estimation. The needs in real-time applications for on-line estimation have stimulated the development of recursive versions of many identification methods, see (Ljung and Söderström, 1983). The methods of recursive estimation for dynamic systems are usually based on discretized linear models formulated as ARMAX-models.

There are however systematic problems associated with the approach of ARMAX-model based parameter estimation in some applications.

A simple observation is the following. Assume that a parameter of an ARMAX-type model changes abruptly to a new value. Detection of a such a change and convergence of the parameter estimate to the new value would require a time proportional to the sampling period and to the number of estimated parameters. This delay may be intolerably long. It is sometimes impossible to improve the response time. A shorter sampling period may be incompatible with good parameter identifiability.

Adaptive control applications are often associated with recursive estimation methods. One obvious constraint of many discrete-time adaptive control schemes is the feature that the sampling period must be chosen to provide good identifiability rather than good control action.

Another problem is the application of discrete-time parameter estimation to the identification of continuous-time models. Firstly, the model representation formulated in the differential operator must be translated to a shift operator formulation. There are several ways to do this. An exact parameter translation does however typically require some matrix exponentiations. This means that a certain continuous-time parameter will have a nonlinearly distributed influence on all the discrete-time parameters. A consequence is that it becomes very difficult to focus the interest on a certain continuous-time parameter. n order to monitor one continuous-time parameter it is in general needed to estimate the full order discrete-time parameter vector. In other words - it is difficult to separate known parameters from unknown ones and partitioning is not easy. Moreover, the discrete-time parameters become dependent on the sampling interval. This is an unattractive property and the discrete-time parameters have often no physical meaning. There exist some linear discrete-time approximations of the differential operator e.g. the simple backward difference or the bilinear Tustin's approximation with better properties of parameter translation. These methods are however approximations with limited applicability.

Secondly, the discrete-time identification requires anti-aliasing filters of the input-output data before sampling. High frequency dynamics will otherwise corrupt the estimation. The assumption of elimination of high frequency influence may be acceptable in off-line identification where some time series signal processing may be exercised. Efficient elimination of the high frequency components is however difficult in the case of on-line identification. A good frequency cut-off property of a sampling filter would require noncausal operations which cannot be implemented. A causal filter with sufficiently good damping of high frequencies may on the other hand cut away too much of the useful low frequency contents or introduce a delay. Moreover, the sampling filter will be incorporated as a part of discrete-time process model and it is difficult to separate filter parameters from process parameters of physical significance. This situation is sometimes a dilemma and the shortcomings of discrete-time ARMAX-type estimation become obvious.

There are other cases of application where the approach of traditional time series analysis is unsatisfactory. Consider a case where a system can be separated into two parts where one part is well known and where the other part is unknown or time varying. It is in this case certainly not good common sense to estimate the well known parameters along with the unknown ones.

MOTIVATION

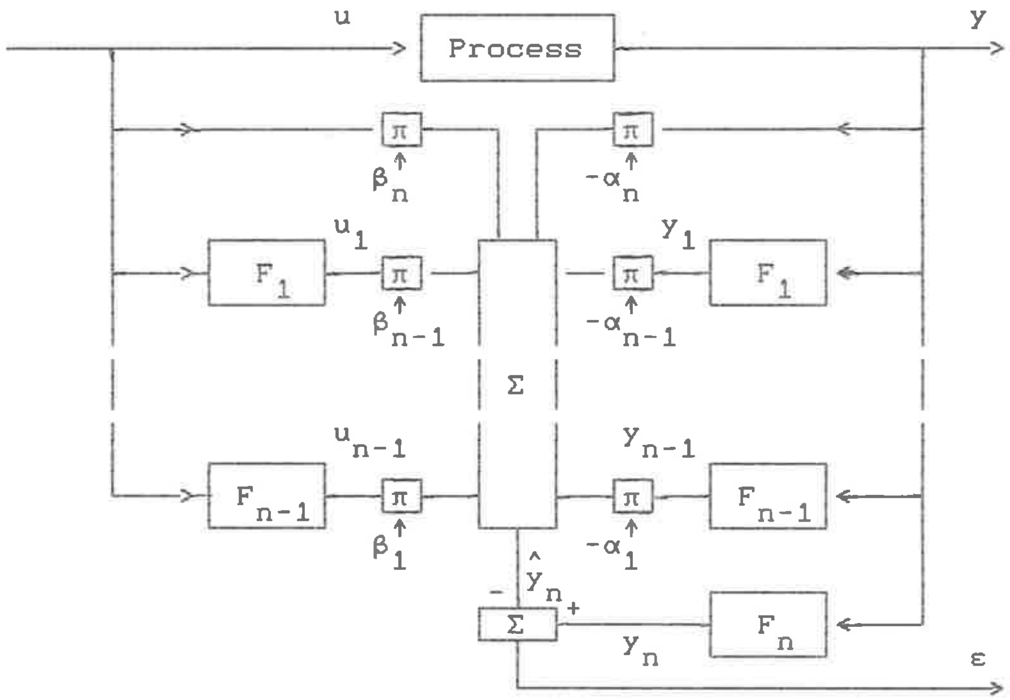
The arguments in the introduction have shown that there is an interest for methods of recursive identification with good properties for estimation of continuous-time systems. In applications of hybrid adaptive control it should be possible to choose sampling frequencies of control and identification independently. It is desirable to make identification of partitioned, coupled and partially known systems.

From a mathematical point of view there are some problems with traditional approaches to recursive estimation for continuous-time systems. Operator transformations are e.g. usually approximative or nonlinear. The discrete-time parameters are nonlinear in the original parameters and the estimates are difficult to translate in real time. All sets of discrete-time parameters are sampling time dependent. A certain continuous-time parameter becomes distributed over all discrete parameters. Requirements on anti-aliasing may be too restrictive or even non-realistic. It is also unsatisfactory that the anti-aliasing filters become part of the process model.

It could therefore be asked why there is no analogue to ARMAX-models for continuous-time systems. The succesful ARMAX-models correspond to polynomials in the forward or the backward shift operators with advantages for modelling and signal processing, respectively.

Polynomials in the differential operator can not be used for identification immediately due to the implementation problems associated with differentiation. There is no commonly used operator in parameter estimation that corresponds to the backward shift operator of discrete-time systems.

There is however one approach to identification of continuous-time systems that developed in the days of analog computers. It was - as far as the author knows - pioneered by Clymer (1959) and Young (1965, 1969). An account of these ideas is given in chapter 9 of (Eykhoff, 1974).



The idea is to have "state variable filters" $F_1,\dots,F_n$ acting on the inputs and outputs of the continuous-time process. One possible alternative is to choose

$$F_1 = \frac{s}{1+s\tau}, \dots, \quad F_n = \left(\frac{s}{1+s\tau}\right)^n \tag{1}$$

Let all y_i, u_i ; $0 \le i \le n$ be outputs from the filter assembly. These filter outputs will then give approximative derivatives of the inputs and outputs. If the original input-output model is

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_n y = b_1 \frac{d^{n-1} u}{dt^{n-1}} + \dots + b_n u \tag{2}$$

then we can fit the parameters $\alpha_1,\dots,\alpha_n$, $\beta_1,\dots,\beta_n$ of the model

$$y_n + \alpha_1 y_{n-1} + \dots + \alpha_n y = \beta_1 u_{n-1} + \dots + \beta_n u \tag{3}$$

by parameter adjustment until y_n and $\hat{y}_n$ coincide and $\varepsilon \to 0$. For good choices of the state variable filters it holds that

$$\alpha_i \approx a_i \ , \quad \beta_i \approx b_i ; \quad 1 \le i \le n \tag{4}$$

although the filtered signals are only approximatively equal to the true state variables.

The problems to measure and reconstruct derivatives have been reflected in the difficulties to design adaptive control, see (Parks, 1966). The concept of state variables filter has also had a certain influence in the theory of adaptive control, see (Monopoli, 1974) and (Elliott, 1982).

Another source of inspiration is found in algebraic system theory. Pernebo (1981) has developed an algebraic theory for linear systems using polynomials in causal and stable operators. He obtains a representation where stability and causality are treated by the same algebraic criteria. One possibility to use these ideas is to introduce an operator translation so that the continuous time linear modelling is made in terms of some low pass filter operator

$$z = \frac{a}{p+a} = \frac{1}{1+p\tau} \; ; \; \tau = 1/a \tag{5}$$

with a time constant τ.

PROBLEM FORMULATION

The rest of this paper tries to systematize and develop some results on estimation of parameters of continuous-time models. There is a potential for adaptive control of continuous-time systems control although these aspects are not elaborated here.

The idea is to find a causal, stable, realizable linear operator that may replace the differential operator while keeping an exact transfer function. This shall be done in such a way that we obtain a linear model for estimation of the original transfer function parameters a_i, b_i. We will consider cases where we obtain a linear model in low pass filter operators. It will be shown that there is always a linear one-to-one transformation which relates the continuous-time parameters and the convergence points for each choice of filter.

This approach has been used for continuous-time adaptive control by Johansson (1983, 1985). Canudas de Wit (1985) has performed a number of case studies with investigations of convergence rates and other properties.

The paper starts with a model transformation. It is then shown that there always exists a parameter transformation back to the original continuous-time model parameters. Then follows investigations on the state space properties of the introduced filters and the original model. The convergence rate of the parameter estimates is then considered. Finally, there are two examples with applications to time invariant and time varying systems, respectively.

A MODEL TRANSFORMATION

Consider a linear n-th order transfer operator formulated with a differential operator p=d/dt and unknown coefficients a_i, b_i.

$$G_0(p) = \frac{b_1 p^{n-1} + \ldots + b_n}{p^n + a_1 p^{n-1} + \ldots + a_n} = \frac{B(p)}{A(p)} \tag{6}$$

It is assumed that A and B are coprime. Introduce now the operator

$$z = \frac{a}{p+a} = \frac{1}{1 + p\tau} \; ; \quad \tau = 1/a \tag{7}$$

This allows us to make the following transformation

$$G_0(p) = \frac{B(p)}{A(p)} = \frac{B^*(z)}{A^*(z)} = G_0^*(z) \tag{8}$$

with

$$A^*(z) = 1 + \alpha_1 z + \alpha_2 z^2 + \dots + \alpha_n z^n$$

$$B^*(z) = \beta_1 z + \beta_2 z^2 + \dots + \beta_n z^n \tag{9}$$

An input -output model is now easily formulated as

$$A^*(z)y(t) = B^*(z)u(t) \tag{10}$$

$$y(t) = -\alpha_1[zy](t) - \dots - \alpha_n[z^n y](t) + \beta_1[zu](t) + \dots + \beta_n[z^n u](t)$$

This is now a linear model of a dynamical system at all points of time. Notice that [zu], [zy] etc. mean filtered inputs and outputs. The parameters α_i, β_i may now be estimated by any suitable method for estimation of parameters of a linear model. A reformulation of the model (10) is

$$y(t) = \theta_\tau^T \varphi_\tau(t) \qquad t \in R \tag{11}$$

$$\theta_\tau = \left[-\alpha_1 \; -\alpha_2 \; \dots -\alpha_n \; \beta_1 \; \dots \beta_n\right]^T \tag{12}$$

$$\varphi_\tau(t) = \left[[zy](t) \; [z^2 y](t) \; \dots \; [zu](t) \dots [z^n u](t)\right]^T \tag{13}$$

We may now have the following continuous-time input-output relations.

$$y(t) = G_0(p)u(t) = G_0^*(z)u(t) \tag{14}$$

$$y(t) = \theta_\tau^T \varphi_\tau(t) \tag{15}$$

$$Y(s) = \theta_\tau^T \Phi_\tau(s) \quad \text{with} \quad \Phi_\tau(s) = L\{\varphi_\tau(t)\}(s) \tag{16}$$

where L means a Laplace-transform. Finally, a Laplace transformation of (14) gives

$$Y(s) = G_0^*(z(s))U(s) \tag{17}$$

A particularly rich and attractive feature is the fact that the same linear relation holds in both the time domain and the frequency domain. Notice that this property hold without any approximations or any selection of data.

PARAMETER TRANSFORMATIONS

Before we proceed to signal processing aspects we should make clear the relation between the parameters α_i, β_i of (9) and the original parameters a_i, b_i of the transfer function (6). Let the vector of original parameters be denoted by

$$\theta = [-a_1 \ -a_2 \ \ldots -a_n \ b_1 \ \ldots \ b_n]^T \tag{18}$$

The relation between (12) and (18) is then

$$\theta_\tau = F_\tau \theta + G_\tau \tag{19}$$

Using the definition of z (7) and (8) it can be shown that the 2nx2n matrix F_τ and the 2nx1 vector G_τ are given by

$$F_\tau = \left[\begin{array}{c|c} M_\tau & 0 \\ \hline 0 & M_\tau \end{array}\right] \tag{20}$$

with

$$M_\tau = \begin{bmatrix} m_{11} & 0 \cdots & 0 \\ \vdots & \ddots & 0 \\ m_{n1} & \cdots & m_{nn} \end{bmatrix} \quad \text{with} \quad m_{ij} = (-1)^{i-j}\binom{n-j}{i-j}\tau^j \tag{21}$$

Further

$$G_\tau = [g_1 \ldots g_n \ 0 \ldots 0]^T \quad \text{with} \quad g_i = \binom{n}{i}(-1)^i \tag{22}$$

The matrix F_τ is invertible when M_τ is invertible i.e. for all $\tau>0$. The parameter transformation is then one-to-one and

$$\theta = F_\tau^{-1}\left[\theta_\tau - G_\tau\right] \tag{23}$$

We may then conclude that the parameters a_i, b_i of the continuous-time transfer function G_0 may be reconstructed from the parameters α_i, β_i of θ_τ. As an alternative we may estimate the original parameters a_i, b_i of θ from the linear relation

$$y(t) = \theta_\tau^T\varphi_\tau(t) = \left[F_\tau\theta + G_\tau\right]^T\varphi_\tau(t) \tag{24}$$

$$y(t) = \theta^T\left[F_\tau^T\varphi(t)\right] + G_\tau^T\varphi_\tau(t) \tag{25}$$

where F_τ and G_τ are known matrices for each τ.

STATE SPACE TRANSFORMATION

It is of major concern that no information is lost when doing the operator transformation. This is not obvious from the original approach with state space filters where a filtered state variable could only approximate the true one due to the low pass filter properties. In this section we will show that there is a one-to-one mapping between the state space associated with the original system description and that of the transformed description.

Consider therefore the transfer function

$$G_0(s) = \frac{b_1 s^{n-1} + \ldots + b_n}{s^n + a_1 s^{n-1} + \ldots + a_n} \tag{26}$$

The controllable canonical form of (26) with a state vector x and the differential operator p may be written as

$$p\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} -a_1 & \cdots & & -a_n \\ 1 & & & \\ 0 & 1 & & \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}u(t) \tag{27}$$

$$y(t) = \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}x(t)$$

This may be associated with the fractional form

$$A(p)\xi(t) = u(t)$$

$$y(t) = B(p)\xi(t) \tag{28}$$

with ξ as a scalar 'partial state', see (Kailath, 1980). The components x_i of the state vector x may now be related to ξ via the the correspondence

$$x_i(t) = p^{n-i}\xi(t) \qquad 1 \le i \le n \tag{29}$$

The representation (27) is sufficient to describe the dynamics of the identification object. The system order is however increased by the introduced state variable filters. The filters will increase the minimal order of the system. It is possible to find a state space of the order 2n to describe both the process and the filters although the realisation often is non-minimal.

$$A'(p)\xi'(t) = u(t)$$

$$y(t) = B'(p)\xi'(t) \tag{30}$$

$$A'(p) = A(p)\left[p+a\right]^n = p^{2n} + \dots + a'_{2n}$$

$$B'(p) = B(p)\left[p+a\right]^n = b'_1 p^{2n-1} + \dots + b'_{2n} \tag{31}$$

A state space realization is given by

$$p\begin{bmatrix} x'_1(t) \\ \vdots \\ x'_{2n}(t) \end{bmatrix} = \begin{bmatrix} -a'_1 & \cdots & \cdots & -a'_{2n} \\ 1 & & & \\ 0 & 1 & & \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x'_1(t) \\ \vdots \\ x'_{2n}(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u(t) \tag{32}$$

with

$$x'_i(t) = p^{2n-i}\xi'(t) \tag{33}$$

Each of the components of φ_τ may now be expressed as a linear combination of the state vector components. We have with the arguments of (7), (31), (33).

$$[zu](t) = \frac{a}{p+a}A(p)\left[p+a\right]^n\xi'(t) = \tag{34}$$

$$= a\left[p^{2n-1}\xi'(t)\right]+\ldots+ a^n a_n\left[p^0\xi'(t)\right] = ax_1'(t) +\ldots+ a^n a_n x_{2n}'(t)$$

$\vdots$

$$[z^n y](t) = a^n\left[b_1 \ldots b_n \; 0 \ldots 0\right]x'(t) = \frac{B(p)}{A(p)}\cdot\left(\frac{a}{p+a}\right)^n u(t) \qquad (35)$$

The original state vector x of (27) is related to x' by the following relations.

$$\left[p+a\right]^n\xi'(t) = \xi(t) \qquad (36)$$

$$x_i(t) = p^{n-i}\xi(t) = p^{n-i}\left[p+a\right]^n\xi'(t) \qquad (37)$$

From (33) and (37) we find

$$x_i(t) = \sum_{j=1}^{n}\binom{n}{j}a^{n-j}x'_{n+i-j} \quad 1\le i\le n \qquad (38)$$

Consider now the full regression vector

$$\varphi_\tau(t) = \left\{[zy](t)\ldots[z^n y](t) \; [zu](t)\ldots[z^n u](t)\right\} \qquad (39)$$

The φ_τ-vector is therefore related to the state vector x' by a linear transformation matrix M' containing coefficients obtained from (34), (35) and (38)

$$\varphi_\tau(t) = M'x'(t) \qquad (40)$$

Notice that all components of the state space are observable from $[z^n y]$ provided there are no common factors of A and B. This means that the states of x'(t) and x(t) are observable from φ_τ. From the construction of (32) we also see that the state x' is controllable from u provided there are no common factors of A and B. Neither should there be any factor of (p+a) in B. This means that it is in principle possible to determine an input u such that x obtains any direction in the 2n-dimensional space. This means in a sense that no information is lost in the filtering process. The following theorem can be shown.

Theorem:

Let G be a rational function such that

$$G(p) = \frac{B(p)}{A(p)}\cdot\left(\frac{a}{p+a}\right)^n = \frac{B'(p)}{A'(p)} \; ; \; deg(A)=n \; ; \; deg(B)=m\le n-1$$

where the polynomial factorization is such that B has no common factor with A or (p+a). Let the following strictly proper transfer operator relation hold between input u and output y

$$y(t) = \frac{B(p)}{A(p)}u(t)$$

Let furthermore z be the operator

$$z=\frac{a}{p+a}$$

Let φ_τ be the vector of filtered inputs and outputs

$$\varphi_\tau(t) = \left[[zu](t)\ldots[z^n u](t)\ [zy](t)\ldots[z^n y](t)\right] \quad (39)$$

and let x' be the state vector of the controllable canonical form of G.

Then there exists a linear transformation such that

$$x'(t) = T_\tau \varphi_\tau(t) \quad (41)$$

for an invertible matrix T_τ.

<u>Proof</u>: See appendix.

REMARK

The theorem above has shown that φ_τ is a sufficient state vector for the system to be identified and the filter state. The controllability of x' and φ_τ means that any direction in the 2n-dimensional space can be reached. Active improvement of identifiability by choice of the input u is also in principle possible.

SIGNAL PROCESSING FILTERS

It has been shown in the previous sections that the transfer operator may be exactly transformed to the linear model

$$y(t) = \theta_\tau^T \varphi_\tau(t) \quad (11)$$

with

$$\varphi_\tau(t) = \left[(zy)\ (z^2 y)\ \ldots\ (zu)\ (z^2 u)\ldots\right]^T \quad (13)$$

It is straightforward to implement the filtered inputs and outputs of φ and we may estimate θ.

Sometimes it may be desirable to make some data selection. Let us thus denote such a data selection by the filter f in the time domain or F in the frequency domain. Let subscript f denote a signal filtered by f or F.

The estimation algorithm will then fit parameters to data from the relation

$$y_f(t) = \theta_\tau^T \varphi_f(t)$$

or

$$Y_f(s) = \theta_\tau^T \Phi_f(s) \qquad (42)$$

This means that we have the the possiblity to make filtering operations in the time domain or in the frequency domain or both. Filtering operations in the frequency domain will lead to weightings and selections in certain frequency ranges. Time domain filtering will mean choices of recording times, averaging or sampling.

An interesting possibility of time domain filtering is to pick y(t) and all components of $\varphi_\tau(t)$ at certain points of time $t=t_1, t_2,\ldots$ The linear relation (41) then of course still holds between y and φ_τ. These sampled data may now be used to fit parameters to the continuous time model (41), (42) by using ordinary discrete time recursive estimation methods. Notice that there is no discrete time model involved although we use sampled data and discrete time estimation.

This type of data sampling does not need to be equidistant and "slow sampling" may be used if a lower convergence rate can be accepted. Notice also that the sampling for constant parameters θ may be performed without any anti-aliasing filter. This is due to the fact that θ rather than y is the reconstructed entity. It would be necessary, however, to choose the sampling frequency properly when it is desirable to track a time varying θ.

CHOICE OF THE LOW PASS FILTER z

It is of practical interest to consider the choice of the time constant $\tau=1/a$ of the low pass filter z used in the modelling. Notice from (8), (11) that the input-output relation is

$$Y(s) = G_0(s)U(s) = G_0^*(z(s))U(s) = \theta_\tau^T \Phi_\tau(s) \qquad (43)$$

The accuracy of a parameter estimate $\hat{\theta}_\tau$ is often evaluated with a a quadratic criterion formulated

$$J_t\left[\hat{\theta}_\tau(t)\right] = \int_0^t \left[y_f(r)-\hat{\theta}_\tau^T(t)\varphi_f(r)\right]^2 dr \tag{44}$$

It is also possible to give criteria in the frequency domain. We will make statements for "long" but finite time intervals [0,t] and assume that Parseval's relation holds between the time domain and the frequency domain. A counterpart to (44) in the the frequency domain is then the following

$$J_\omega\left[\hat{\theta}_\tau(t)\right] = \int_{-\infty}^{+\infty} \left|Y_f(i\omega)-\hat{\theta}_\tau^T(t)\Phi_f(i\omega)\right|^2 d\omega \tag{45}$$

Introduce the parameter error vector as

$$\tilde{\theta}_\tau(t) = \hat{\theta}_\tau(t) - \theta_\tau \tag{46}$$

and the weighting matrix

$$P^{-1}(t) = \int_{-\infty}^{+\infty}\Phi_\tau(-i\omega)\Phi_\tau^T(i\omega)d\omega \tag{47}$$

with Φ_τ defined in (16). The criterion (45) may now be written as

$$J_\omega = \tilde{\theta}_\tau^T(t)\left[\int_{-\infty}^{+\infty}\Phi_\tau(-i\omega)\Phi_\tau^T(i\omega)d\omega\right]\tilde{\theta}_\tau(t) = \tilde{\theta}_\tau^T(t)P^{-1}(t)\tilde{\theta}(t) \tag{48}$$

All components of Φ_τ are dependent on the input U(s). From (13-17) it is found that the vector Φ_τ may be decomposed into

$$\Phi_\tau(s) = \Gamma_\tau(s)U(s) \tag{49}$$

with

$$\Gamma_\tau(s) = \left[z(s)G_0(s)\ldots z^n(s)G_0(s)\ \ z(s)\ldots z^n(s)\right]^T \tag{50}$$

In (50) we see that P depends on the spectrum of the input signal u. There is also a dependence of the unknown transfer function G_0. It is therefore difficult to derive any result on how to choose τ optimally on the basis of this type of pure quadratic criterion.

Another approach is to request a certain convergence rate of the parameter estimates. This may be reflected in a weighted least squares criterion

$$J'_t\left[\hat{\theta}(t)\right] = \int_0^t e^{2\alpha r}\left[y_f(r)-\hat{\theta}_\tau^T(t)\varphi_f(r)\right]^2 dr \tag{52}$$

where $\alpha>0$ is some constant rate of desired exponential convergence. The weighting matrix modifies to

$$P_\tau^{-1}(t) = \int_0^t e^{2\alpha r}\varphi_f(r)\varphi_f^T(r)dr \tag{53}$$

when evaluated in the time domain. The frequency domain counterpart of (52) is

$$J'_\omega\left[\hat{\theta}_\tau(t)\right] = \frac{1}{2\pi i}\int_{0-i\infty}^{0+i\infty}\left|Y_f(s-\alpha)-\hat{\theta}_\tau(t)\Phi_f(s-\alpha)\right|^2 ds \tag{54}$$

By examination the integrand of (54) we find that the convergence properties of (54) are related to the properties of $U(s)$ and $\Gamma(s)$ which depends on $z(s)$ and $G_0(s)$. One finds that

$$P_\tau^{-1}(t) = \int_{-\infty}^{+\infty}\Gamma_\tau(-i\omega-\alpha)\Gamma_\tau^T(i\omega-\alpha)U(-i\omega-\alpha)U(i\omega-\alpha)d\omega \tag{55}$$

For a non-zero input $U(s)$ we have the following condition for convergence of (55).

1: $G_0(s-\alpha)$ stable

2: $z(s-\alpha)$ stable $\Rightarrow$ $\tau<1/\alpha$

This determines the limits of convergence rates for different parameter estimations. It means that we have to require that G_0 is stable and responds rapidly enough to the input u. It is also necessary that the filter time constant τ is smaller than the desired time constant of convergence $1/\alpha$.

IMPLEMENTATION OF LEAST SQUARES ESTIMATION

We will look at recursive least squares estimation of parameters of the linear model. A minimization of (44) in the continuous time domain gives an algorithm of the type

$$\dot{\hat{\theta}}_\tau(t) = P_c(t)\varphi_\tau(t)\left[y(t)-\hat{\theta}_\tau(t)\varphi_\tau(t)\right] \tag{56}$$

$$\dot{P}_c(t) = -P_c(t)\varphi(t)\varphi^T(t)P_c(t) \tag{57}$$

and a convergence rate given by

$$\frac{d}{dt}\left[\tilde{\theta}_\tau^T(t)P_c^{-1}(t)\tilde{\theta}_\tau(t)\right] = -2\left[\tilde{\theta}_\tau^T(t)\varphi_\tau(t)\right] \tag{58}$$

We normally prefer a discrete-time estimation for reasons of implementation although it is suboptimal. A discretization of (1) and (2) at time instants t=0,h,..,kh and a Riemann sum type approximation of integration gives the familiar recursive least squares identification, see e.g. (Ljung and Söderström, 1983).

$$\hat{\theta}_\tau(t) = \hat{\theta}_\tau(t-h) + P_s(t)\varphi_f(t)\left[y_f(t)-\hat{\theta}_\tau^T(t-h)\varphi_\tau(t)\right] \tag{59}$$

$$P_s^{-1}(t) = P_s^{-1}(t-h) + \varphi_f(t)\varphi_f^T(t) \quad ; \quad t=0,h,\ldots,kh \tag{60}$$

Some manipulations of (60) give a formula to update P_s, see e.g. (Ljung and Söderström, 1983).

More sophisticated numerical integration routines may of course also be utilized. With trapezoidal interpolation we may replace (60) by

$$P_s^{-1}(kh) = P_s^{-1}(kh-h) + \frac{h}{2}\left[\varphi_f(kh)\varphi_f^T(kh)+\varphi_f(kh-h)\varphi_f^T(kh-h)\right] \tag{61}$$

to obtain a better approximation of (44).

The sampling rate will certainly influence the parameter accuracy both with respect to convergence rate and with respect to the accuracy at different "frequency points". The former aspect has been treated in the previous section. The accuracy of the transfer function at different frequency points may be investigated by the following arguments. The accuracy is determined by the value of the cost criterion J and its matrix

$$P_s^{-1}(\ell h) = \sum_{k=0}^{\ell} \varphi_f(kh)\varphi_f^T(kh) \tag{61}$$

Introduce the pulse train function of ℓ pulses

$$W_\ell(t) = \sum_{k=0}^{\ell} \delta(t-kh) \tag{62}$$

Then we may rewrite (61) as

$$P_s^{-1}(t) = \int_{-\infty}^{+\infty} W_\ell(t-r)\varphi_f(r)\varphi_f^T(r)dr = \varphi_f(t)\varphi_f^T(t)*W_\ell(t) \tag{63}$$

Plancherel's theorem for Laplace transform of a convolution gives

$$L\{P_s^{-1}(\ell h)\}(s) =$$

$$= L\{\varphi_f(t)\varphi_f^T(t)*W_\ell(t)\} = L\{\varphi_f(t)\varphi_f^T(t)\}\cdot L\{W_\ell(t)\} \qquad (64)$$

Poisson's formula now gives for large values of ℓ

$$L\{W_\ell(t)\} \to L\{\sum_{k=0}^{\infty} \delta(t-kh)\} = \frac{1}{h}\{\sum_{k=-\infty}^{\infty} \delta(s-i\frac{2\pi k}{h})\} \quad ; \ \ell\to\infty \qquad (65)$$

The result of (64) and (65) is that there is a weighting in the frequency domain such that when $t\to\infty$ there will be better accuracy of the estimated transfer function at the frequencies

$$\omega_k = \frac{2\pi k}{h} \quad ; \ k=0, \pm 1, \pm 2, \pm 3, \ldots \qquad (66)$$

This means that the accuracy at multiples of the sampling frequency ω_s will be favoured.

EXAMPLE 1 - Estimation of two constant parameters

Consider the system with input u, output y, and the transfer operator G_0

$$y(t) = G_0(p)u(t) = \frac{b_1}{p + a_1}u(t) \tag{67}$$

Use the operator transformation z of (5)

$$z = \frac{1}{1 + p\tau} \tag{68}$$

This gives the transformed model

$$G_0^*(z) = \frac{b_1 \tau z}{1 + \left(a_1\tau - 1\right)z} = \frac{\beta_1 z}{1 + \alpha_1 z} \tag{69}$$

A linear estimation model of the type (11) is given by

$$y(t) = -\alpha_1 [zy](t) + \beta_1 [zu](t) = \theta_\tau^T(t)\varphi_\tau(t) \tag{70}$$

with

$$\varphi_\tau(t) = \left([zy](t) \;\; [zu](t)\right)^T \tag{71}$$

and the parameter vector

$$\theta_\tau = \left(-\alpha_1 \;\; \beta_1\right)^T \tag{72}$$

The original parameters are found via the relations

$$a_1 = \left(\alpha_1 + 1\right)/\tau \qquad b_1 = \beta_1/\tau \tag{73}$$

Sampling and application of the recursive least squares estimation algorithm (59-60) give the following simulation results for different choices of the filter time constant τ and the sampling interval h. All simulations have started with initial values at zero for the parameter estimates and the filters. The P-matrix of the recursive least squares estimation has been initialized to the same value in all simulations.

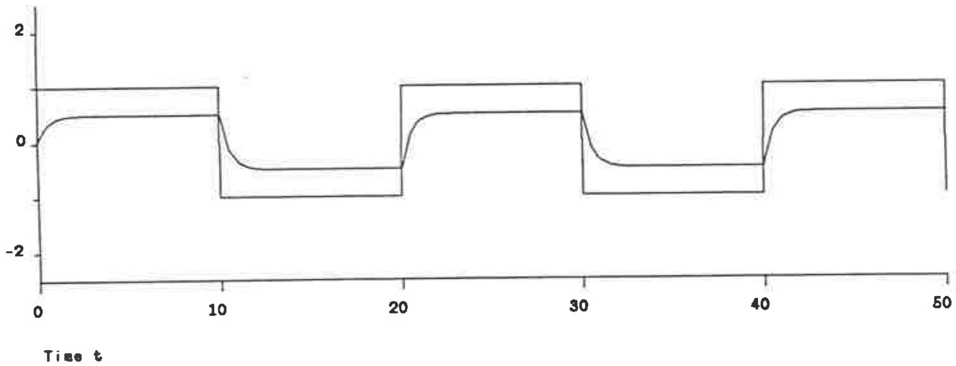


Figure 1: Input u and output y of the process

The simulations have been performed with $a_1=2$ and $b_1=1$ and a moderate excitation.

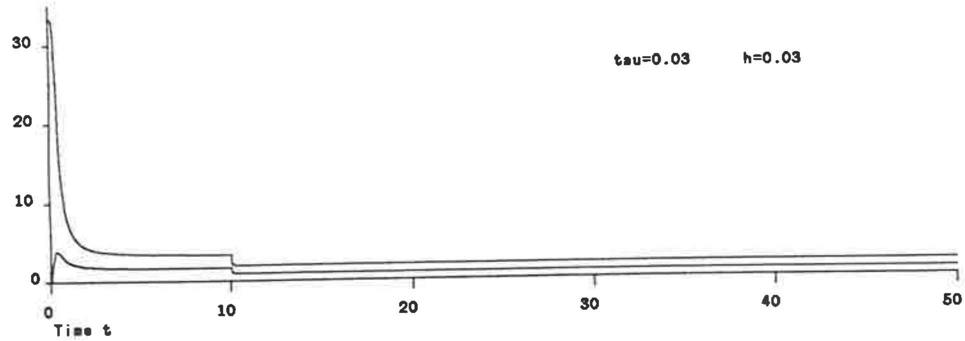


Figure 2: Estimates $\hat{a}_1$ and $\hat{b}_1$ for h=0.03 and τ=0.03

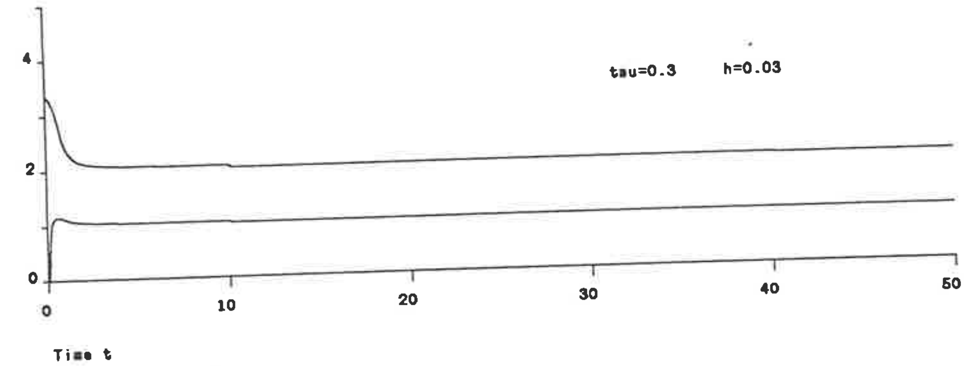


Figure 3: Estimates $\hat{a}_1$ and $\hat{b}_1$ for h=0.03 and τ=0.3

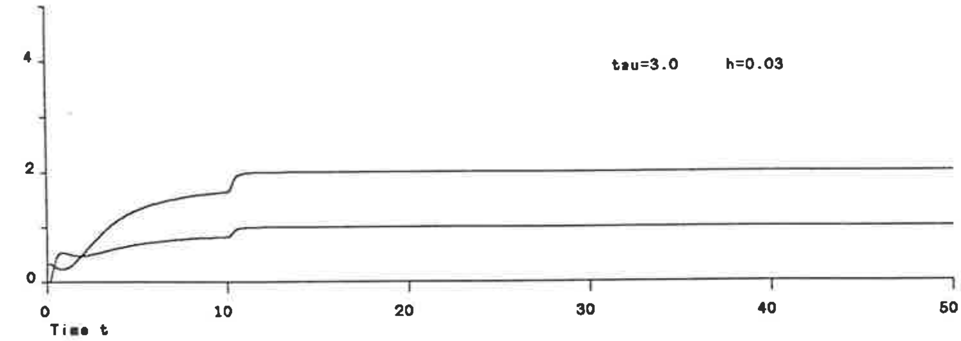


Figure 4: Estimates $\hat{a}_1$ and $\hat{b}_1$ for h=0.03 and τ=3.0

The simulations above show that the convergence works properly over at least two decades of values of τ. The convergence rate is faster for a shorter τ but the convergence transient may be violent for "too" short time constants τ. The estimates are accurate for all the cases of simulation above. The recursive estimation has been performed every 0.03 s and is no limiting factor for the convergence rate here.

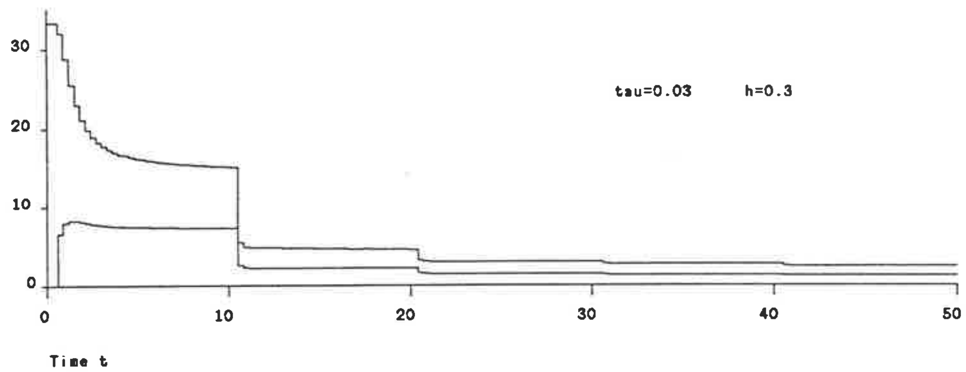


Figure 5: Estimates $\hat{a}_1$ and $\hat{b}_1$ for h=0.3 and τ=0.03

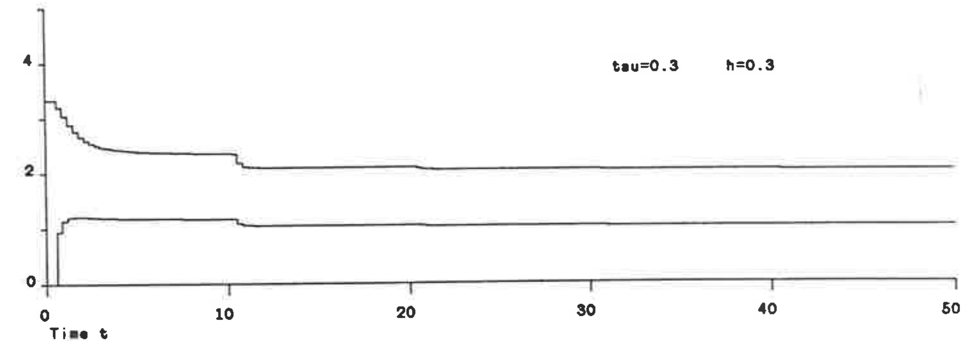


Figure 6: Estimates $\hat{a}_1$ and $\hat{b}_1$ for h=0.3 and τ=0.3

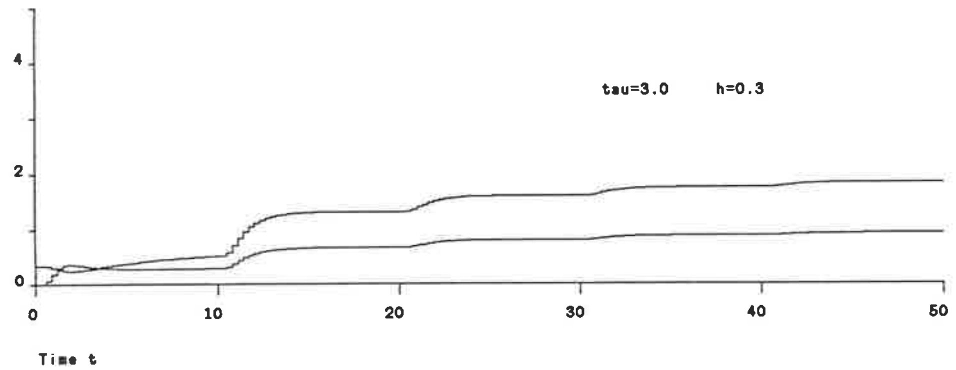


Figure 7: Estimates $\hat{a}_1$ and $\hat{b}_1$ for h=0.3 and τ=3.0

The last three simulations show the convergence rate when doing slower sampling than above. The convergence rate is still good in the figure 6 where $h=\tau=0.3$ i.e. of the same order of magnitude as the process time constant $1/a_1$. The cases of figures 5 and 7 exhibit slower convergence but still good accuracy at the end of the time scales.

It can be seen from the figures that there is acceptable convergence rates over a large range of values of the time constant τ. Notice that the convergence rate is higher for small values of τ but the parameter transient tends to be more violent.

Finally, this paper does not treat the properties of identification in the presence of noise. The following simulation shows however the convergence when "white noise" with variance σ^2=0.1 is corrupting the output y. The simulations have been run with $h=\tau=0.3$. The figure shows the parameter estimates both with and without noise. As expected there is a certain bias of estimates when noise is present.

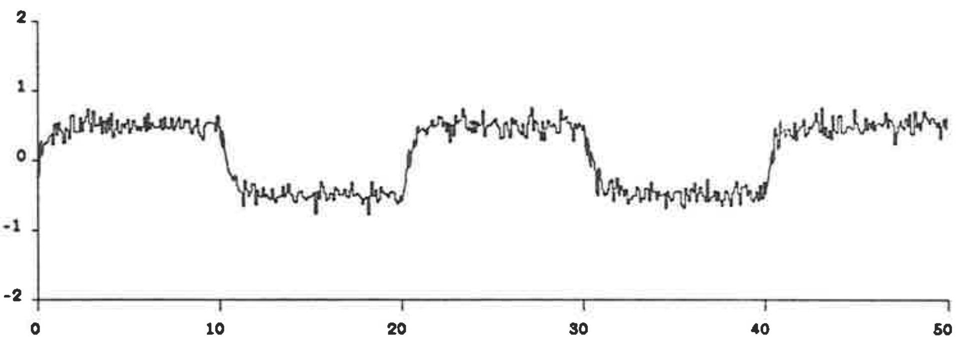


Figure 8: Noise influence on output y

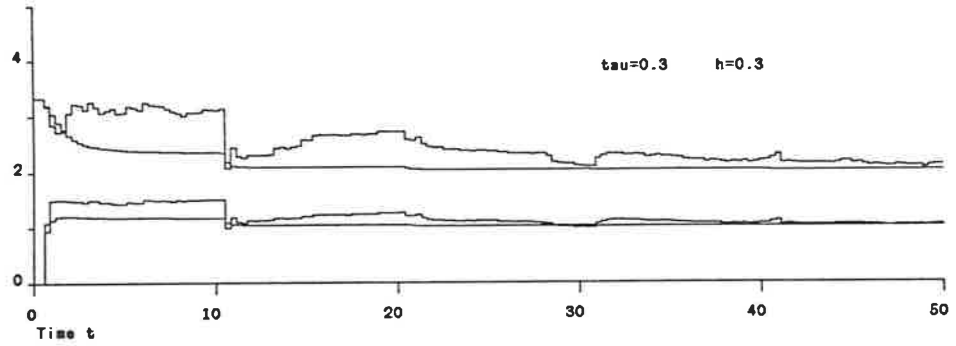
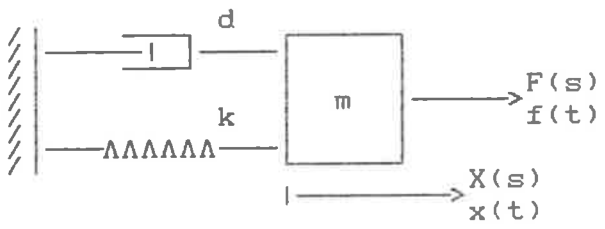


Figure 9: Noise influence on $\hat{a}_1$ and $\hat{b}_1$ for h=0.3 and τ=0.3.

EXAMPLE 2 - Estimation of a time varying parameter



Assume that the spring coefficient k and the mass m are well known and constant. The damping coefficient d is however unknown and time varying. Assume that the force f and the position x are measurable. The force f is assumed to be the control input variable. The transfer function from input f to output x is given by

$$\frac{X(s)}{F(s)} = \frac{\frac{1}{m}}{s^2 + \frac{d}{m}s + \frac{k}{m}} = \frac{b_2}{s^2 + a_1 s + a_2} \qquad (74)$$

The operator translation (5) gives the transformed transfer operator from force f to position x

$$\frac{b_2\tau^2 z^2}{\left[1-z\right]^2 + a_1\tau\left[1-z\right]z + a_2\tau^2 z^2} \qquad (75)$$

The unknown coefficient is a_1 for which we find the relation

$$a_1 \cdot \varphi_\tau(t) = y(t) \qquad (76)$$

with

$$\varphi_\tau(t) = \tau z\left[1-z\right]x(t) \qquad (77)$$

$$y(t) = -\left[1-z\right]^2 x(t) - a_2\tau^2 z^2 x(t) + \tau^2 b_2 z^2 f(t) \qquad (78)$$

A simple heuristical tracking algorithm for a_1 is the following threshold algorithm

$$\hat{a}_1(kh) = \begin{cases} \hat{a}_1(kh-h) & \text{if } |\varphi_\tau(kh)|<0.1 \\ y(kh)/\varphi_\tau(kh) & \text{if } |\varphi_\tau(kh)|\geq 0.1 \end{cases} \qquad (79)$$

The potential for adaptive control is obvious. Assume that we want a continuous-time adaptive controller matches the input-output behaviour to the model

$$\frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \tag{80}$$

with ζ=0.8 and ω_0=0.7. It is possible to utilize the estimated parameter $\hat{a}_1$ to modify the controller on-line. Some simulation studies are presented in the figures below.

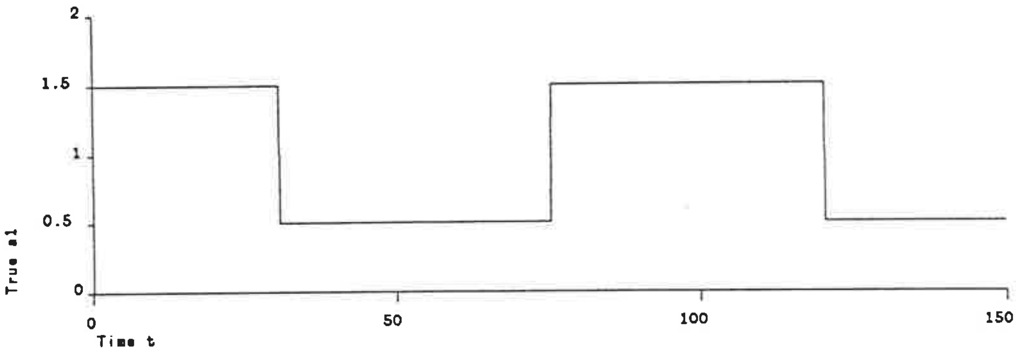


Figure 10: The true parameter a_1

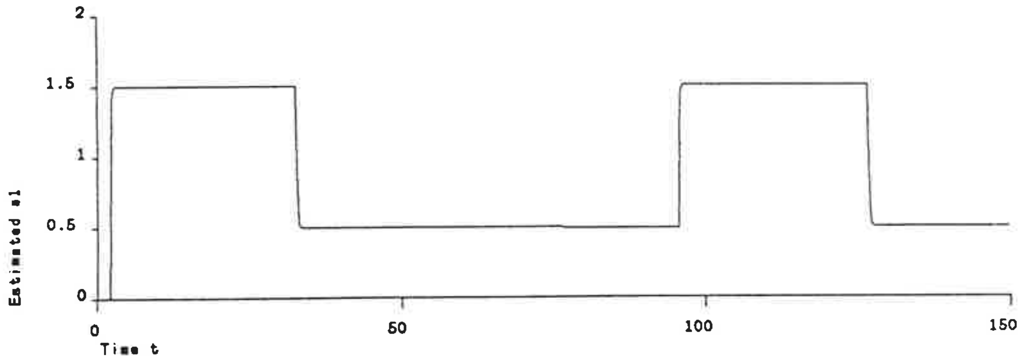


Figure 11: Estimate of a_1

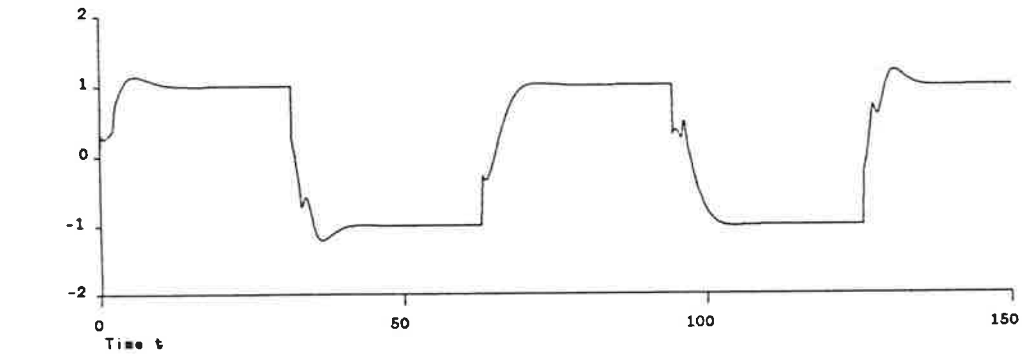


Figure 12: Force F

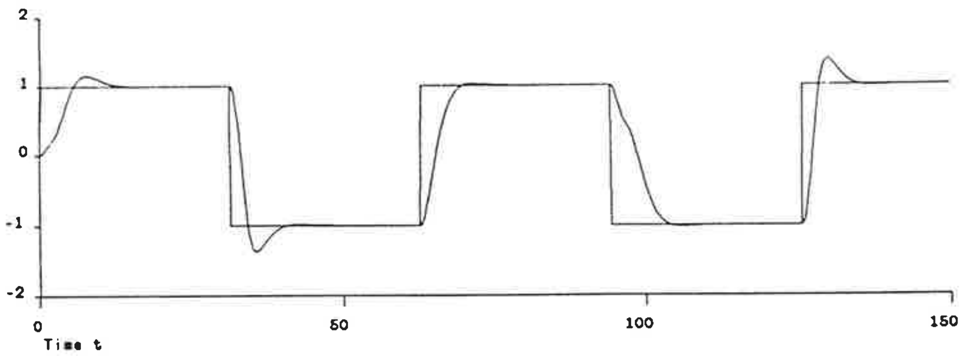


Figure 13: Reference value for position x_2 and actual position x_2

CONCLUSIONS

The old method of state variable filters for estimation of parameters in continuous time dynamic systems has been revised. It has been shown that this method can be made rigourous by reformulation of the model in terms of a realizable operator. The problem formulation has also been made different with respect to the state variables. It is not claimed that the derivatives of the inputs and outputs are reconstructed by the filters. In fact, the filter outputs tend to be bad approximates of the desired derivatives when the filter time constants become longer. Instead we claim that the filter outputs are linear combinations of certain state vector components. We have also given an invertible linear transformation to find the original parameters from the new parameter set. For each value of the filter constant τ there is a certain transformation.

It has been demonstrated that the convergence results do not depend critically on a certain choice of the low pass filter constant.

There are certain advantages from the model formulation point of view. It is possible to maintain a connection between modelling and parameter estimation. Parameters do not need to become abstract and anonymous due to discretization transformations. It is easier to partition the object of identification into known and unknown parts. This means that the estimation may give faster results with a moderate burden of computation.

The potential for adaptive control is very interesting. A sampling period for the regulator may be chosen independently from that of the identification. Unlike the ARMAX-model based discrete time adaptive regulators there is no obligation to choose the sampling period to satisfy the needs of both control and identication.

ACKNOWLEDGEMENTS

This work was supported by Centre National de la Recherche Scientifique under the agreement between C.N.R.S. and the Science Research Council NFR of Sweden.

This work was also supported by Skandinaviska Enskilda Bankens Skånska Stipendiefond, Malmö, Sweden.

REFERENCES

Canudas de Wit, C. (1985). Recursive estimation of continuous-time parameters, Report, Dept. of Autom. Control, Lund Inst. of Technology, Lund, Sweden.

Clymer, A.B. (1959). Direct system synthesis by means of computers, Trans. AIEE, 77, part I, 798-806.

Elliott, H. (1982). Hybrid adaptive control of continuous time systems, Trans. Autom. Contr., AC-27, 419-426.

Eykhoff, P. (1974). System identification, parameter and state estimation, John Wiley & Sons, London.

Johansson, R. (1983). Multivariable adaptive control. Lund Inst. of Technology, Lund, Sweden.

Johansson, R. (1985). Estimation and direct adaptive control of delay-differential systems, Proc VII IFAC Conf. on Ident, York, England.

Kailath, T. (1980). Linear Systems. Prentice-Hall, Englewood Cliffs, USA.

Ljung, L., Söderström, T. (1983). Theory and practice of recursive identification. MIT Press, Cambridge, USA.

Monopoli, R.V. (1974). Model reference adaptive control with an augmented error signal, IEEE Trans. Autom. Contr., AC-19, 474-484.

Parks, P.C. (1966). Liapunov redesign of model reference adaptive control systems, IEEE Trans. Autom. Contr., AC-11, 362-367.

Pernebo, L. (1981). Algebraic control theory for linear multivariable systems, Part I-II, IEEE Trans. Autom. Contr., AC-26, 171-193.

Young, P.C. (1965). Process parameter estimation and self-adaptive adaptive control, Proc. 2nd IFAC Symp. - The theory of self-adaptive control systems, Plenum Press, New York.

Young, P.C. (1969): Applying parameter estimation to dynamic systems, Control engineering, 16, Oct 119-125, Nov 118-124.

APPENDIX

<u>Theorem:</u>

Let G be a rational function such that

$$G(p) = \frac{B(p)}{A(p)} \cdot \left(\frac{a}{p+a}\right)^n = \frac{B'(p)}{A'(p)} \quad ; \quad \deg(A)=n \quad ; \quad \deg(B)=m \leq n-1$$

where the polynomial factorization is such that B has no common factor with A or (p+a). Let the following strictly proper transfer operator relation hold between input u and output y

$$y(t) = \frac{B(p)}{A(p)} u(t)$$

Let furthermore z be the operator

$$z = \frac{a}{p+a}$$

Let φ_τ be the vector of filtered inputs and outputs

$$\varphi_\tau(t) = \left\{[zu](t) \ldots [z^n u](t) \quad [zy](t) \ldots [z^n y](t)\right\} \qquad (39)$$

and let x' be the state vector of the controllable canonical form of G.

Then there exists a linear transformation such that

$$x'(t) = T_\tau \varphi_\tau(t) \qquad (41)$$

for an invertible matrix T_τ.

<u>Proof</u>

Let y_i be the output of 'i' operators z operating on y

$$y_i(t) = [z^i y](t) \qquad (A1)$$

The transfer operator from the input u to the output y_n is then

$$y_n(t) = \frac{B'(p)}{A'(p)} u(t) \qquad (A2)$$

with

$$A'(p) = A(p)\left[p+a\right]^{n} = p^{2n} + a_1' p^{2n-1} + .. + a_{2n}'$$

$$B'(p) = B(p)\left[p+a\right]^{n} = b_1' p^{2n-1} + .. + b_{2n}' \qquad (31)$$

A fractional form for (A2) is

$$A'(p)\xi'(t) = u(t)$$

$$y(t) = B'(p)\xi'(t) \qquad (30)$$

The state space realization on the controllable canonical form is given by

$$p\begin{bmatrix} x_1'(t) \\ \vdots \\ \vdots \\ x_{2n}'(t) \end{bmatrix} = \begin{bmatrix} -a_1' & \cdots & \cdots & -a_{2n}' \\ 1 & & & \\ 0 & 1 & & \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 \quad 1 & 0 \end{bmatrix} \begin{bmatrix} x_1'(t) \\ \vdots \\ \vdots \\ x_{2n}'(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u(t) \qquad (32)$$

$$y_n(t) = \left[b_1' \cdots\cdots b_{2n}'\right] x'(t)$$

where the state vector components are given by

$$x_i'(t) = p^{2n-i}\xi'(t) \qquad (33)$$

Consider now the fractional form (28) relating u and y

$$A(p)\xi(t) = u(t)$$

$$y(t) = B(p)\xi(t) \qquad (28)$$

with

$$\left[p+a\right]^{n}\xi'(t) = \xi(t) \qquad (36)$$

With this state representation it holds that

$$[zu](t) = \frac{a}{p+a} A(p)\left[p+a\right]^{n}\xi'(t) =$$

$$= a\left[p^{2n-1}\xi'(t)\right] + .. + a^{n}a_n\left[p^{0}\xi'(t)\right] = a x_1'(t) + ... + a^{n}a_n x_{2n}'(t)$$

:

$$[z^i u](t) = A(p)a^i(p+a)^{n-i}\xi'(t)$$

$$[z^i y](t) = B(p)a^i(p+a)^{n-i}\xi'(t)$$

$$[z^n y](t) = a^n\left[b_1 \ldots b_n \ 0 \ldots 0\right]x'(t) = \frac{B(p)}{A(p)} \cdot \left[\frac{a}{p+a}\right]^n u(t) \tag{35}$$

This means that all components of φ_τ may be expressed as linear combinations of the components of x'. Hence

$$\varphi_\tau(t) = M'x'(t) \qquad (40),\ (A3)$$

The next step is to show that x' may be expressed as a linear transformation of φ_τ.

Another form of (28) is the fractional form expressed in the operator z, see (8-10) and (Pernebo, 1981).

$$A^*(z)\xi_z(t) = u(t)$$

$$y(t) = B^*(z)\xi_z(t) \tag{A4}$$

with coprime A^* and B^* and with

$$\xi_z(t) = \left[p+a\right]^n \xi(t) \tag{A5}$$

Recall that

$$\left[p+a\right]^n \xi'(t) = \xi(t) \tag{36}$$

This gives that

$$\xi'(t) = \left[\frac{1}{p+a}\right]^{2n} \xi_z(t) \tag{A6}$$

From (33) and (A6) it is found that

$$x_i'(t) = p^{2n-i}\left[\frac{1}{p+a}\right]^{2n} \xi_z(t) = P_i(z)\xi_z(t) \tag{A7}$$

where P_i for $1 \le i \le 2n$ are polynomials in the operator z.

$$P_i(z) = \frac{\left[(p+a)-a\right]^{2n-i}}{\left[p+a\right]^{2n}} = a^{-i}\left(\frac{a}{p+a}\right)^i\left(1 - \frac{a}{p+a}\right)^{2n-i}$$

It can be seen from the following relation that all P_i contain powers of z from 1 to 2n.

$$P_i(z) = a^{-i}z^i\left[1-z\right]^{2n-i} = \sum_{j=0}^{2n-i}\binom{2n-i}{j}a^{-i}(-z)^{2n-j} \tag{A8}$$

The factorization polynomials $A^*(z)$ and $B^*(z)$ are coprime. The ring of polynomials is an integral domain and the Diophantine equations

$$A^*(z)R_i^*(z) + B^*(z)S_i^*(z) = P_i^*(z) \quad ; \; 1 \le i \le 2n \tag{A9}$$

have therefore solutions for all i in the given interval. The solutions are such that there are solutions with

$$R_i^*(z) = r_{i1}z + \ldots + r_{in}z^n$$

$$S_i^*(z) = s_{i1}z + \ldots + s_{in}z^n \tag{A10}$$

From (A4) and (A9) it is found for $1 \le i \le 2n$

$$R_i^*(z)u(t) + S_i^*(z)y(t) = P_i^*(z)\xi_z(t) = x_i'(t) \tag{A11}$$

The constraints on the polynomial degrees gives a possibility to express (A11) on the form

$$x_i'(t) = \left[r_{i1} \ldots r_{in} \; s_{i1} \ldots s_{in}\right]\varphi_\tau(t) \tag{A12}$$

with

$$\varphi_\tau(t) = \left[[zy](t) \cdots [z^n y](t) \; [zu](t) \ldots [z^n u](t)\right] \tag{39}$$

Let the matrix T_τ be

$$T_\tau = \begin{bmatrix} r_{11} & \cdots & r_{1n} & s_{11} & \cdots & s_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ r_{(2n)1} & \cdots & r_{(2n)n} & s_{(2n)1} & \cdots & s_{(2n)n} \end{bmatrix} \tag{A13}$$

Then it holds that

$$x'(t) = T_{\tau}\varphi_{\tau}(t) \tag{A14}$$

A conclusion from (A3) and (A14) gives that

$$T_{\tau}^{-1} = M'$$

Hence, T_{τ} is an invertible matrix relating x' and φ_{τ}.