Topics in Dual Control

Sternby, Jan

1977

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

Jan Sternby

Topics in Dual Control

Lund 1977

Dokumentutgivare
Lund Institute of Technology
Dept of Automatic Control

Handläggare
Karl Johan Åström

Författare
Jan Sternby

Dokumentnamn
REPORT

Dokumentbeteckning
LUTFD2/(TFRT-1012)/1-135/(1977)

Utgivningsdatum
May 1977

Ärendebeteckning

Dokumenttitel och undertitel
Topics in Dual Control

Referat (sammandrag)
In the first part of this thesis, a particular dual control problem is formulated, based on a controlled discrete-state Markov chain. The optimal dual control law and several suboptimal schemes are derived analytically and compared.

A condition for strong consistency of the least squares identification method is given in part II. A Bayesian point of view is adopted in deriving the results from Martingale theory.

In part III a new suboptimal dual control algorithm is given. Computer simulations indicate that it can manage to control quite difficult systems with time-varying stochastic parameters. Using simulated examples, a comparison is made of the performance of some previously suggested suboptimal dual control schemes, and the new one.

Referat skrivet av
Author

Förslag till ytterligare nyckelord

Klassifikationssystem och -klass(er)

Indextermer (ange källa)
Adaptive systems, Automatic control, Markov chains, Control simulation, Least squares method, Consistency (Thesaurus of Engineering and Scientific Terms, Engineers Joint Council, N.Y., USA).

Omfång
135 pages

Övriga bibliografiska uppgifter

Språk
English

Sekretessuppgifter

ISSN

ISBN

Dokumentet kan erhållas från
Department of Automatic Control
Lund Institute of Technology
P O Box 725, S-220 07 Lund 7, Sweden

Mottagarens uppgifter

Pris

TOPICS IN DUAL CONTROL

av

Jan Sternby

Civ ing, Ld

Akademisk avhandling som för avläggande av teknisk doktorsexamen vid tekniska fakulteten av universitetet i Lund kommer att offentligen försvaras i sal M:B, Maskinhuset, Lunds Tekniska Högskola, fredagen den 27 maj 1977 kl 10.15.

TABLE OF CONTENTS

	Page
Dual Control	7
References	12
Part I - A Simple Dual Control Problem with an Analytical Solution	13
I. Introduction	13
II. Two Problems	13
III. Derivation of Functional Equations	13
IV. Solutions to the Two Problems	14
V. Comparison of Different Strategies	15
VI. Concluding Remarks	16
References	17
Part II - Consistency for the Method of Least Squares Using Martingale Theory	19
1. Introduction	21
2. The Least Squares Identification Method. Notations	22
3. General Results	25
4. Main Results. The Gaussian Case	27
5. Conclusions	37
6. References	38
Appendix	40
Part III - Regulators for Time-varying Stochastic Systems	43
1. Introduction	45
2. The System and the Criterion	47
3. Some Different Regulators	56
4. The Two-step Regulator	72
5. Simulations	85
6. Conclusions	117
Appendix A	119
Appendix B	122
Appendix C	125
References	131

Dual Control

The PI-regulator has been used to control industrial processes for much more than 100 years. Two reasons for its extensive use are its simplicity and its good performance in many situations. With this regulator there is no control error if the closed-loop system is stable. The regulator parameters are fairly easy to tune, and, in principle, very little knowledge of the controlled process is required for the tuning. These properties have resulted in the wide-spread use of the PI-regulator in industry.

There are, however, processes that need a more complex regulator. Some disturbances cannot be suppressed sufficiently well by a PI-regulator. The process characteristics may vary so much with time that no constant PI-regulator could control the system satisfactorily. The arrival of digital computers has now made it possible to implement control laws that require a substantial amount of computing for each new control action. Control algorithms that were previously too complex can now be implemented. Buth, with increasing regulator complexity, the number of regulator parameters to be chosen has also increased. It may then be difficult to find the best combination of parameter values. This tuning problem is one reason why the more complex regulators have not been used in practice.

To overcome the tuning problem the so called self-tuning regulator was derived some years ago; see Kalman (1958), Wieslander and Wittenmark (1971), Peterka (1970) or Åström and Wittenmark (1973). The basic idea is to estimate first the process dynamics on-line using a real-time identification algorithm, and then calculate the regulator parameters as if the true system were perfectly described by the estimated system model. This is an adaptive regulator, since it can change its parameters to follow changes in process dynamics. The self-tuning regulator has been tested on several industrial processes and has been found to be a valuable complement to conventional PI-controllers;

see Åström, Borisson, Ljung, Wittenmark (1977).

The self-tuning regulator was primarily designed to handle systems with constant but unknown parameters. Practical applications of the self-tuning regulator and simulated examples have indicated that it can also manage to control slowly time-varying processes. But, with larger or more rapid parameter variations in the system, it is expected that more complicated regulators will be required. Such problems are discussed in this thesis.

Parameter variations may be modelled by assuming that the system parameters are stochastic processes. Stochastic optimal control theory can then be applied to derive an optimal control law. In general however, the calculations become quite involved, and only for very simple cases will this procedure yield a practical regulator. But Feldbaum (1960) noted that the optimal control law has a particular structure. The input has two different tasks. It must, of course, govern the system as usual. But it must, at the same time, make the parameter estimation as accurate as possible in order to improve future control of the system. These tasks are often conflicting.

The regulator must compromise between good control and good identification. Such regulators are called dual. This thesis treats various aspects of dual control, and consists of this summary and the following papers:

I. A Simple Dual Control Problem with an Analytical Solution. IEEE Trans Automatic Control AC-21 (1976) 840-844.

II. On Consistency for the Method of Least Squares Using Martingale Theory. Submitted to IEEE Trans Automatic Control (1977).

III. Regulators for Time-varying Stochastic Systems.

There are very few cases for which the optimal dual control law has actually been calculated. In order to gain more insight into the nature of dual control it is desirable to have examples where the optimal control can be computed analytically. Part I of this thesis contains one such example. It is a discrete-state Markov chain. The transition probabilities are determined by the inputs in a particular way. It is then not only possible to calculate the optimal control law, but also to compare the performance of several suggested suboptimal schemes for dual control.

In a dual control problem unknown system parameters must be estimated. It is then important to know the behaviour of the estimation algorithm. It seems reasonable to demand that for constant system parameters the estimates should converge to the true values. This is called consistency.

Part II of the thesis considers consistency for the method of least squares. A modification of this identification scheme is employed in part III, where the processes, however, have time-varying parameters.

A Bayesian point of view is adopted in discussing consistency. Compared with previous results, the underlying assumptions about the distribution of the random variables involved are more restrictive. On the other hand, the assumptions concerning feedback and system stability are less restrictive. The least squares method is a basic estimation algorithm and it is therefore important to understand all aspects of it. The results of part II clearly show the relation between least squares estimates and martingales.

In most cases the optimal dual control problem can not be solved. It is therefore interesting to study the effects of making different approximations. This is done in part III of this thesis. A number of suggested algorithms are discussed, and a new, suboptimal, dual control law is derived. Some

algorithms require a large amount of computing effor. So far, this has been considered a serious drawback. With modern computer technology, however, prices are rapidly decreasing, and it is now possible to have small, low cost computers, dedicated to certain control tasks. Potential regulators need no longer be disqualified because they require lengthy calculations or large memory space. It is, however, still important to find out for what kind of systems more complicated algorithms will be advantageous, and when simpler regulators will suffice. This point is examined, by means of simulations, in part III, where also different suboptimal regulators are compared.

As in the simple PI-regulator, all the adaptive regulators employ feedback from the output; the only difference is that, in adaptive regulators, the computations required to deduce the control action from the output are more complex. Feedback is thus a very general and useful concept in automatic control, even for complicated time-varying stochastic systems.

Simulation is a valuable tool to get a feeling for the properties of suboptimal control laws. But a thourough analysis and a true understanding of dual control is still missing. New formulations of the problem must probably be sought in order to arrive at complete solutions. This is an interesting area for future research in dual control.

In preparing this thesis I was supported very much by a number of people around me. First of all I want to thank my supervisor Professor Karl Johan Åström. He is always open for discussions and has given numerous valuable comments on different versions of the manuscript. I also want to thank my colleagues at the Department of Automatic Control, who form a group with an open and stimulating atmosphere.

The manuscripts were excellently typed by Gudrun Christensen and Eva Dagnegård, and all the figures were skilfully prepared by Britt-Marie Carlsson.

This work was partly supported by the Swedish Board for Technical Development under contracts No. 74-3476 and 76-3804, which is gratefully acknowledged.

References

Åström, K J, Borisson, U, Ljung, L, Wittenmark, B (1977): Theory and Applications of Self-tuning Regulators. To appear in Automatica 13.

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

Feldbaum, A A (1960): Dual Control Theory I-IV. Automation and Remote Control 21, 874-880, 1033-1039, and 22, 1-12, 109-121.

Kalman, R E (1958): Design of a Self-optimizing Control System. ASME Trans 80, No 2, 468-478. Also in: Oldenburger, R (ed.): Optimal Self-optimizing Control, MIT Press (1966) 440-449.

Peterka, V (1970): Adaptive Digital Regulation of Noisy Systems. IFAC Symp on Identification and Process Parameter Estimation, Prague, paper No. 6.2.

Wieslander, J, Wittenmark, B (1971): An Approach to Adaptive Control Using Real-time Identification. Automatica 7, 211-217.

Part I

A Simple Dual Control Problem with an Analytical Solution

JAN STERNBY

***Abstract*—A stochastic control problem for which the optimal dual control law can be calculated analytically is given. The system is a four state Markov chain with transition probabilities that depend on the control variable. The performance of the optimal dual control law and of some suboptimal control laws are calculated and compared.**

I. Introduction

It is, in general, very difficult to solve control problems leading to dual control laws in the sense of Feldbaum [5]. A few examples of this type have been solved numerically by extensive computer calculations; see, e.g., [3], [6], [8], [9], and [15]. The motivation for solving these necessarily rather simple problems has been to obtain some insight into how dual control laws actually work and thus to understand how to make good suboptimal dual controllers for more difficult problems.

But numerical solutions do not, however, give as much and as detailed information as analytical solutions. For one thing, with just a numerical solution it is not known what happens between data points. For this reason an example is given in this short paper which is completely solvable over any time interval. With the analytical solution given, one can study in detail how the dual controller works. An analytical comparison is also made between the performance of different control strategies. However, this is a very special problem, and consequently nothing can be said in general about other and more realistic problems.

The example is based on governed Markov chains as in [14], but may also be looked upon as a simplification of the example in [8]. In Section II, two problems are formulated corresponding to open- and closed-loop control. Functional equations for the two cases are set up in Section III and their solutions appear in Section IV. Some different control laws are considered, including an open-loop feedback control, see, e.g., [1], [4], [10], and [13]. The performance of these controls is analyzed in Section V. The last section is a discussion of the results.

Manuscript received March 1, 1976; revised August 5, 1976. Paper recommended by Y. Bar-Shalom, Chairman of the IEEE S-CS Stochastic Control Committee. This work was supported in part by the Swedish Board for Technical Development under Contract 74-3476.

The author is with the Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

The calculations needed are easy but tedious, and are therefore omitted. They are, however, stated in [11].[1]

II. Two Problems

Consider a Markov chain with four states called x_1 to x_4. The transition probabilities depend on a control variable u $(0 \leqslant u \leqslant 1)$, and are shown in Table I. The functions p_i are given in Fig. 1.

For every u

$$p_1+p_3+p_5=p_2+p_4+p_6=1.$$

The p_i:s are chosen piecewisely linear in u to make the calculations simple and, for the same reason, some of the transition probabilities are identical and others are zero. The desire to achieve a dual effect in the resulting regulator also restricts the possible choices of p_i:s.

Introduce the loss function h which assigns a loss to the states in the following way:

$$h(x_1)=h(x_4)=1, \quad h(x_2)=h(x_3)=0.$$

The loss function thus puts the states together into two groups: x_1 and x_4 in one and x_2 and x_3 in another. For obvious reasons these groups will be called the one-states and the zero-states, respectively.

Let us first consider the open-loop problem. No measurements are made and thus the only available information about the state is its initial probability distribution which is known in advance. The value of the state at time t is denoted by x^t.

Problem 1 (open loop): Determine at $t=t_0$ a sequence $\bar{v}$ of values for the control variable to minimize the criterion

$$\bar{W}_n=E\sum_{t=t_0+1}^{t_0+n} h(x^t)$$

where n is chosen in advance.

The second problem is the corresponding closed-loop situation. At every time t, $t_0+1 \leqslant t \leqslant t_0+n$, the value of the loss function is measured. This means that it is possible to separate a zero-state from a one-state but impossible to separate the two states within the group.

Admissible control laws at time t may use all information available at that time, which is now the initial probability distribution for the state *and* the outcome of all measurements up to and including that at time t. Due to the Markov property this information is contained in the conditional probability distribution for x^t.

Now consider Problem 2.

Problem 2 (closed loop): Find a sequence v of admissible control laws to minimize the criterion

$$W_n=E\sum_{t=t_0+1}^{t_0+n} h(x^t)$$

where n is chosen in advance. This is the problem which leads to a dual control law.

III. Derivation of Functional Equations

Using dynamic programming, functional equations for the minimal loss in the two cases are easily established. The system is completely time-invariant, so that the functional equations can be set up in such a way that the minimization involved is a minimization with respect to the *initial* control variable at $t=t_0$.

Since all the transition probabilities from states x_1 and x_2 are equal the future development of the process will be the same if the current state is x_1 or x_2. The same property is also true for states x_3 and x_4. Therefore, introduce

q = the probability for the initial state to be x_1 or x_2.

[1]This report is available on request from the Department of Automatic Control, Lund Institute of Technology, P.O. Box 725, S-220 07 Lund 7, Sweden.

TABLE I
THE TRANSITION PROBABILITIES

next state \ current state	x_1	x_2	x_3	x_4
x_1	$p_1(u)$	$p_1(u)$	$p_2(u)$	$p_2(u)$
x_2	$p_3(u)$	$p_3(u)$	0	0
x_3	0	0	$p_4(u)$	$p_4(u)$
x_4	$p_5(u)$	$p_5(u)$	$p_6(u)$	$p_6(u)$

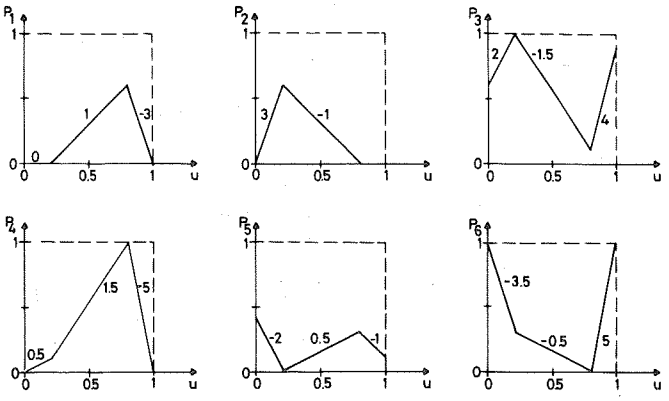


Fig. 1. The transition probabilities as functions of u. Each line is marked by its slope.

Then the corresponding probability for x_3 or x_4 will be $1-q$. Also define $P^0(q,u)$ and $P^1(q,u)$ as

$P^0(q,u)$ = the probability for x^{t_0+1} to be a zero-state if the control variable at $t=t_0$ was u.

$P^1(q,u)$ = the corresponding probability for a one-state.

Table I then gives

$$P^0(q,u)=q\cdot p_3(u)+(1-q)\cdot p_4(u) \tag{1}$$

$$P^1(q,u)=q\cdot[p_1(u)+p_5(u)]+(1-q)\cdot[p_2(u)+p_6(u)]. \tag{2}$$

The Open-Loop Case

In this case $\bar{q}(q,u)$ is introduced as

$\bar{q}(q,u)$ = the probability for x^{t_0+1} (the state at time $t=t_0+1$) to be x_1 or x_2, if the control variable at $t=t_0$ is u.

Then from Table I

$$\bar{q}(q,u)=q[p_1(u)+p_3(u)]+(1-q)\cdot p_2(u). \tag{3}$$

Put

$$\bar{F}_n=\min_{\bar{v}}\bar{W}_n$$

where n is the number of steps considered. $\bar{F}_n$ will be a function of q and for $n\geqslant 2$

$$\bar{F}_n(q)=\min_{\bar{v}_0}\left\{P^1(q,\bar{v}_0)+\bar{F}_{n-1}[\bar{q}(q,\bar{v}_0)]\right\} \tag{4}$$

with

$$\bar{F}_1(q)=\min_{\bar{v}_0}P^1(q,\bar{v}_0). \tag{5}$$

Looking just one step ahead the probability for the next state to be a one-state is minimized in (5) (and in (8) below).

The Closed-Loop Case

For this case set

$q^0(q,u)$ = the probability for x^{t_0+1} (the state at time $t=t_0+1$) to be x_2, given that it is a zero-state.

$q^1(q,u)$ = the probability for x^{t_0+1} to be x_1, given that it is a one-state.

By the multiplication rule of probabilities this means that

$$q^0(q,u)=\frac{q\cdot p_3(u)}{P^0(q,u)} \quad \text{and} \quad q^1(q,u)=\frac{q\cdot p_1(u)+(1-q)\cdot p_2(u)}{p^1(q,u)}. \tag{6}$$

As in the previous section put (without overbars)

$$F_n=\min_{v}W_n$$

and obtain for $n\geqslant 2$

$$F_n(q)=\min_{v_0}\left\{P^1(q,v_0)+F_{n-1}[q^1(q,v_0)]\cdot P^1(q,v_0)+F_{n-1}[q^0(q,v_0)]\cdot P^0(q,v_0)\right\} \tag{7}$$

with

$$F_1(q)=\min_{v_0}P^1(q,v_0). \tag{8}$$

The expression to be minimized in (7) will be denoted by $J_n(q,v_0)$. In it the first term is the immediate loss, while the last two terms add up to the expected loss for the next $n-1$ steps knowing that measurements are going to be made.

IV. SOLUTIONS TO THE TWO PROBLEMS

By repeated use of (4) or (7) for increasing values of n, it is now possible to calculate $\bar{F}_n(q)$ or $F_n(q)$ for any n. In [11] it is shown that the expression to be minimized is always a piecewisely concave function. As an example of this, $J_\infty(q,v_0)$ is shown in Fig. 2 for $q=0.4$, 0.5, 0.7, 0.8, and 0.9.

For every value of n, $J_n(q,v_0)$ is of the same type as $J_\infty(q,v_0)$. The only possible minimizing values of the control variable are $v_0=0$, 0.2, 0.8, and 1, independently of n and q. The minimization is then done by a direct comparison of the four corresponding losses.

It should be emphasized that for most choices of piecewisely linear transition probabilities the number of values for the control variable that has to be checked will *not* be four, or even constant, but will increase with n. For such a case it would then not be possible to get an analytical solution valid for all n. Using (2), (5) and (8) give

$$\bar{F}_1(q)=F_1(q)=\begin{cases}0.9q, & q\leqslant 1/2\\ 0.9(1-q), & q\geqslant 1/2\end{cases}$$

and

$$\bar{v}_0^{\text{opt}}(q)=v_0^{\text{opt}}(q)=\begin{cases}0.8, & q<1/2\\ 0.2, & q>1/2\end{cases} \quad \text{(either for } q=1/2).$$

Since a complete knowledge about the state means $q=0$ or $q=1$ it is clear that the best one-step regulator chooses v_0 as if the most probable state was the true one.

The Open-Loop Case

Starting with the $\bar{F}_1(q)$ above, it is shown by induction in [11] that

$$\bar{F}_n(q)=\min\left[\bar{K}_1^n q,\bar{K}_2^n(1-q)\right]$$

with

$$\bar{v}_0^{\text{opt}}=\begin{cases}0.8, & \text{if } \bar{K}_1^n\cdot q<\bar{K}_2^n(1-q) \quad \text{(small } q\text{'s)}\\ 0.8 \text{ or } 0.2, & \text{if } \bar{K}_1^n\cdot q=\bar{K}_2^n(1-q) \quad \text{(medium } q\text{'s)}\\ 0.2, & \text{if } \bar{K}_1^n\cdot q>\bar{K}_2^n(1-q) \quad \text{(big } q\text{'s)}\end{cases}$$

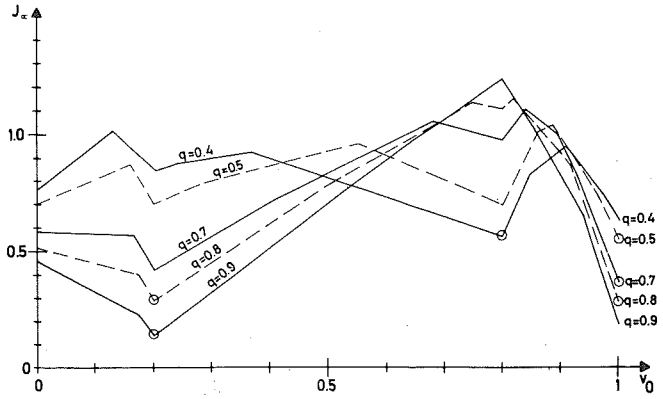


Fig. 2. The expected loss for infinitely many steps as a function of v_0 plotted for some values of q. The rings mark out the global minimum for each q.

where

$$\bar{K}_1^n = 3(1-0.7^n) \to 3 \quad \text{and} \quad \bar{K}_2^n = 1.5(1-0.4^n) \to 1.5.$$

The limiting optimal control for $n \to \infty$ is then

$$\bar{v}_0^{\text{opt}} = \begin{cases} 0.8, & q < 1/3 \\ 0.2, & q > 1/3. \end{cases} \quad \text{(either for } q = 1/3)$$

This control scheme can be used either in an open-loop mode (if no measurements are made) or as the suboptimal closed-loop control law called "open-loop feedback optimal control." In Section V analytical expressions for the performance of these two and some other controls are given. Note that all control laws in this short paper will be discontinuous as functions of q. In the following they will be chosen to be right-continuous.

The Closed-Loop Case

In [11] it is shown by induction that for $n \geq 2$

$$F_n(q) = \min[K^n q, 1-0.9q, K^n(1-q)]$$

with

$$v_0^{\text{opt}} = \begin{cases} 0.8, & \text{if } F_n(q) = K^n q \quad \text{(small } q\text{'s)} \\ 1, & \text{if } F_n(q) = 1-0.9q \quad \text{(medium } q\text{'s)} \\ 0.2, & \text{if } F_n(q) = K^n(1-q) \quad \text{(big } g\text{'s)} \end{cases}$$

where $K^n = 1.4(1-0.1^{n-1}) \to 1.4$ $n \to \infty$.

For q-values close to $1/2$, $v_0 = 1$ is chosen. This v_0-value will never be used by the one-step regulator (or if the state is completely known), and so the n-step regulator is essentially different for $n \geq 2$.

The value $v_0 = 1$ gives an identification step since $q^0(1) = 1$ and $q^1(1) = 0$, i.e., the exact state becomes known. Then the future loss will be zero.

Thus, the n-step regulator for $n \geq 2$ works as follows: For q:s close to zero or one, i.e., good knowledge about the current state, v_0 will be chosen as by the one-step regulator, whereas for q:s close to $1/2$, i.e., poor knowledge about the current state, an identification step will be taken.

The function F_n is shown in Fig. 3 for some values of n. Note that for all $n \geq 2$ the loss is $1-0.9q$ for q close to $1/2$. This is the expected loss when taking an identification step.

The number of identification steps will be either one or zero. It will be zero if the true initial state is probable enough. Then all the time control actions will be taken as if the most probable state was the true one, and the probability for this state will increase.

The identification step, if there is one, does not have to be the first step. If the true initial state has a very small probability, an "incorrect" control action is taken, but no identification is made until the mistake is discovered. This will happen when transition into a one-state occurs.

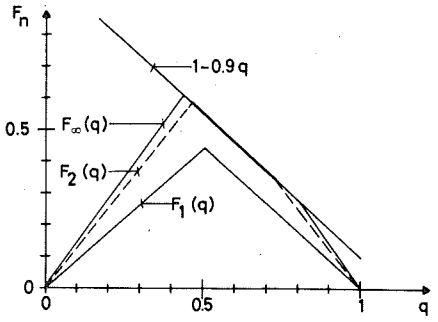


Fig. 3. Graphs of the function F_n for some values of n.

V. Comparison of Different Strategies

In this section the expected n-step loss is listed for six different control strategies. The loss will be denoted by $F_n^j(q)$, $j = a, b, c, d, e,$ or f according to

F^a— open-loop control (no measurements);
F^b— open-loop feedback optimal control (OLFO);
F^c— one-step regulator;
F^d— two-step regulator;
F^e— approximate multistep regulator;
F^f— optimal dual regulator.

In order to include certainty-equivalent control in the above list a state estimate is needed. A natural choice is to take the most probable state as the estimate. Table I and Fig. 1 then give the certainty-equivalent control and it turns out that it is identical to the one-step regulator.

The approximate multistep regulator is calculated in the following way. When (7) is minimized to give v_0 the optimal *open*-loop loss for an infinite number of steps, $\bar{F}_\infty$, is used instead of the corresponding closed-loop loss. This corresponds closely to what is done for more general systems in [12]. The approximate multistep regulator and the OLFO control both use the optimal open-loop loss. However, the OLFO control is computed as if no further measurements are made, but when computing the approximate multistep regulator it is assumed that a final measurement will be made at the next step. Another method using an approximate future loss in the minimization is the one given in [7] leading to "neutral" control.

To calculate the F^j:s an equation similar to (4) (case a) or (7) is used, where the minimization with respect to v_0 is removed. Instead the v_0-value inserted should be some function of q, depending on which regulator is used.

In this section v_k will denote the control used at time $t = t_0 + k$, a function of q_k, the *a posteriori* probability for the state at time $t = t_0 + k$ to be x_1 or x_2. The q previously used is thus equivalent to q_0.

5.1. Open-Loop Control

The control law is

$$v_k = \begin{cases} 0.8, & q_k < 1/3 \\ 0.2, & q_k \geq 1/3. \end{cases}$$

No measurements are made, and q_k can be computed from a formula similar to (3)

$$q_{k+1} = q_k[p_1(v_k) + p_3(v_k)] + (1-q_k)p_2(v_k).$$

Thus, all q_k:s and v_k:s can be computed in advance.

Now from [11]

$$F_n^a = \begin{cases} 3(1-0.7^n)q_0, & q_0 < 1/3 \\ 1.5(1-0.4^n)(1-q_0), & q_0 \geq 1/3 \end{cases} \to \begin{cases} 3q_0 \\ 1.5(1-q_0) \end{cases} \quad \text{as } n \to \infty.$$

Note that $F_\infty^a = \bar{F}_\infty$ of Section IV.

5.2. Open-Loop Feedback Control

Again the control law is

$$v_k = \begin{cases} 0.8, & q_k < 1/3 \\ 0.2, & q_k > 1/3 \end{cases}$$

but now the measurements are used to update q_k. Depending on the measurements, an equation similar to one of those in (6) will be used where $q^0(q,u)$ and $q^1(q,u)$ are replaced by q_{k+1}, u by v_k, and q by q_k. Then from [11]

$$F_n^b = \begin{cases} 1.5(1-0.4^n)q_0, & q_0 < 1/3 \\ 1.5(1-0.4^n)(1-q_0), & q_0 > 1/3. \end{cases}$$

5.3 One-Step Regulator (Certainty-Equivalent Control)

Now

$$v_k = \begin{cases} 0.8, & q_k < 1/2 \\ 0.2, & q_k > 1/2 \end{cases}$$

where q_k is calculated as in 5.2. From [11]

$$F_n^c = \begin{cases} 1.5(1-0.4^n)q_0, & q_0 < 1/2 \\ 1.5(1-0.4^n)(1-q_0), & q_0 > 1/2. \end{cases}$$

5.4. Two-Step Regulator

The control law is

$$v_k = \begin{cases} 0.8 & 0 \leq q_k < 25/54 \quad (\sim 0 \leq q_k < 0.46) \\ 1, & 25/54 \leq q_k < 39/54 \quad (\sim 0.46 \leq q_k < 0.72) \\ 0.2, & 39/54 \leq q_k \leq 1 \quad (\sim 0.72 \leq q_k \leq 1) \end{cases}$$

and q_k is again computed as in 5.2. From [11]

$$F_n^d = \begin{cases} (1.4-5\cdot 0.1^n)q_0, & 0 \leq q_0 < 25/54 \\ 1-0.9q_0, & 25/54 \leq q_0 < 39/54 \\ (1.4-5\cdot 0.1^n)(1-q_0), & 39/54 \leq q_0 \leq 1. \end{cases}$$

5.5. Approximate Multistep Regulator

To obtain the control law minimize

$$J^*(q_k, v_k) = P^1(q_k, v_k) + \bar{F}_\infty[q^1(q_k, v_k)] \cdot P^1(q_k, v_k) + \bar{F}_\infty[q^0(q_k, v_k)] \cdot P^0(q_k, v_k)$$

with respect to v_k. This gives

$$v_k = \begin{cases} 0.8, & 0 \leq q_k < 120/306 \quad (\sim 0 \leq q_k < 0.39) \\ 1, & 120/306 \leq q_k < 255/306 \quad (\sim 0.39 \leq q_k < 0.83) \\ 0.2, & 255/306 \leq q_k \leq 1 \quad (\sim 0.83 \leq q_k \leq 1). \end{cases}$$

As previously, q_k should be computed as in 5.2. From [11]

$$F_n^e = \begin{cases} (1.4-5\cdot 0.1^n)q_0, & 0 \leq q < 120/306 \\ 1-0.9q_0, & 120/306 \leq q_0 < 255/306 \\ (1.4-5\cdot 0.1^n)(1-q_0), & 255/306 \leq q_0 \leq 1. \end{cases}$$

5.6. Optimal Dual Regulator

This regulator is

$$v_k = \begin{cases} 0.8, & 0 \leq q_k < 50/115 \quad (\sim 0 \leq q_k < 0.43) \\ 1, & 50/115 \leq q_k < 92/115 \quad (\sim 0.43 \leq q_k < 0.80) \\ 0.2, & 92/115 \leq q_k \leq 1 \quad (\sim 0.80 \leq q_k \leq 1) \end{cases}$$

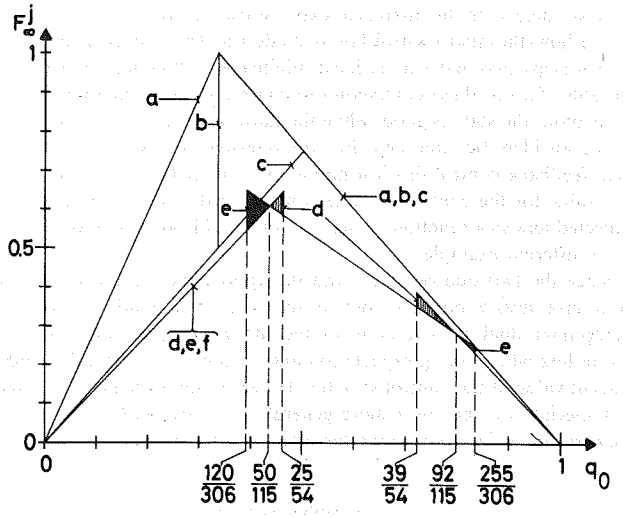


Fig. 4. The expected losses for infinitely many steps, F_∞^j, as a function of the probability for the initial state to be x_1 or x_2 when using different control strategies. a) open-loop control; b) open-loop feedback optimal control; c) one-step regulator; d) two-step regulator; e) approximate multistep regulator; f) optimal dual regulator. The bottom curve is f), from which d) and e) differ only by the two dashed and black areas, respectively.

with q_k as in 5.2–5.5. [11] gives

$$F_n^f = \begin{cases} (1.4-5\cdot 0.1^n)q_0, & 0 \leq q_0 < 50/115 \\ 1-0.9q_0, & 50/115 \leq q_0 < 92/115 \\ (1.4-5\cdot 0.1^n)(1-q_0), & 92/115 \leq q_0 \leq 1. \end{cases}$$

In Fig. 4 the expected losses for infinitely many steps when using these six regulators are compared. Note that most of the dual effect is present for the two suboptimal dual controllers. The difference between the three can be explained as follows.

The two-step regulator is designed to be used for only two steps, and the need for identification is therefore less than with the optimal dual controller. This means that q_0 must be closer to $1/2$ before an identification step is taken.

The approximate multistep regulator, however, must try to find a good estimate of the state immediately, since it is designed as if no measurements are made after the first one.

Fig. 4 classifies the six regulators into three groups. The first group is just the open-loop control, which, of course, gives the biggest loss. The second group consists of the open-loop feedback control and the one-step regulator. For these two the slope of F_∞^j is ± 1.5 depending on the q_0. These strategies do not use any identification steps and so this type of regulator is called *passively adaptive* by Bar-Shalom and Tse [2]. Note that the one-step regulator is better than open-loop feedback control for this example. The third group consists of the two-step regulator, the approximate multistep regulator and the optimal dual control. Here the slope of F_∞^j is ± 1.4 for small and big q_0's, but the loss is decreased by taking an identification step when q_k is close to $1/2$. These controls are called *actively adaptive* in the terms of [2].

VI. Concluding Remarks

The most interesting feature of the given example is that the problem is solvable analytically. This means that it has been possible to examine in detail how different strategies work.

It turns out that in this case the best one-step regulator is equivalent to certainty equivalence control. The best two-step regulator, however, is essentially different in that it gives the possibility for making identification steps. Having this feature built in it performs nearly as good as the optimal dual regulator. The same thing is also true for the approximate multistep regulator.

An interesting detail is that for this example the one-step regulator performs better than open-loop feedback control, but, of course, the problem is a very special one, and again nothing can be said in general.

Nevertheless, as the analytical expressions are given it is also possible to see how the dual control law is made up. The expected loss for the next n steps may have three local minima as a function of the control variable. Two of these correspond to control actions taken when knowledge about the state is good, while the third one corresponds to a control giving an identification step. For the one-step regulator and the open-loop feedback control this last minimum is never the lowest one, and so this value for the control variable is never used. Multiple minima in the expected loss as a function of the control variable are also reported in [3] for a different example.

Since the two-step regulator and the approximate multistep regulator both give nearly minimal loss, it seems as if a good way to derive suboptimal dual regulators is to include *some* approximation of the future loss when taking expectation and minimizing in order to find the current value of the control variable. However, the example in this paper is a special one and for a more general case it can only be said that it may be interesting to examine the effect of such approximations.

Acknowledgment

The author wishes to thank Prof. K. J. Åström for many valuable comments and discussions on this subject.

References

[1] Y. Bar-Shalom and R. Sivan, "On the optimal control of discrete-time linear systems with random parameters," *IEEE Trans. Automat. Contr.*, vol. AC-14, Feb. 1969.
[2] Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence and separation in stochastic control," in *Proc. Joint Automatic Control Conf.*, 1974.
[3] T. Bohlin, "Optimal dual control of a simple process with unknown gain," IBM Nordic Laboratory, Sweden, Tech. Paper TP 18.196, 1969.
[4] S. Dreyfus, "Some types of optimal control of stochastic systems," *SIAM J. Contr.*, vol. 2, no. 1, 1964.
[5] A. A. Feldbaum, *Optimal Control Systems*. New York: Academic, 1965.
[6] J. J. Florentin, "Optimal probing adaptive control of a simple Bayesian system," *J. Elec. Contr.*, vol. 13, no. 2, 1962.
[7] O. L. R. Jacobs and D. J. Hughes, "Simultaneous identification and control using a neutral control law," presented at the IFAC Symp., Boston, MA, 1975, paper 30-2.
[8] O. L. R. Jacobs and S. M. Langdon, "An optimal extremum control system," in *1969 IFAC Conf., Preprints*.
[9] O. L. R. Jacobs and J. W. Patchell, "Caution and probing in stochastic control," *Int. J. Contr.*, vol. 16, no. 1, 1972.
[10] R. Ku and M. Athans, "On the adaptive control of linear systems using the open-loop-feedback-optimal approach," *IEEE Trans. Automat. Contr.*, vol. AC-18, Oct. 1973.
[11] J. Sternby, "Analytical solution of a simple dual control problem," Dep. Automat. Contr., Lund Inst. Technol., Lund, Sweden, Rep. 7430, 1974.
[12] E. Tse, Y. Bar-Shalom, and L. Meier, "Wide-sense adaptive dual control for nonlinear stochastic systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, Apr. 1973.
[13] E. Tse and M. Athans, "Adaptive stochastic control for a class of linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-17, Feb. 1972.
[14] K. J. Åström, "Optimal control of Markov processes with incomplete state information—Part I," *J. Math. Anal. Appl.*, vol. 10, 1965. ——, "Optimal control of Markov processes with incomplete state information—Part II," *J. Math. Anal. Appl.*, vol. 2, 1969.
[15] K. J. Åström and B. Wittenmark, "Problems of identification and control," *J. Math. Anal. Appl.*, vol. 34, Apr. 1971.

Part II- On Consistency for the Method of Least Squares Using Martingale Theory

ABSTRACT

Least Squares Identification is considered from the Bayesian point of view. A necessary and sufficient condition for consistency almost everywhere is given under the assumption that the data are generated by a regression model with white and Gaussian noise.

1. INTRODUCTION

The method of Least Squares (LS) has been treated by many authors, starting with Gauss. Mann and Wald (1943) were the first ones to apply it to time-series modelling, and also to prove its consistency for this case. Åström (1968) extended the consistency result to systems with an input, and Ljung (1976) has shown convergence and consistency under very mild conditions that e.g. include general feedback situations.

By using a Bayesian approach to the identification procedure it is possible to show convergence of the LS estimate just by combining a couple of known theorems. In the present paper the true system parameters are thus regarded as random variables and not constants, which is usually the case. The same approach has been used previously by Åström and Wittenmark (1971).

The main result is a necessary and sufficient condition for consistency a.e. for the LS method in the Gaussian white noise case under a weak condition. This is proved in a series of theorems, where the basic idea is that the LS estimate is a conditional mean, and therefore converges a.e. according to martingale theory.

2. THE LEAST SQUARES IDENTIFICATION METHOD. NOTATIONS

Let a system with p outputs $y(\cdot)$ and r inputs $u(\cdot)$ for every integer t be governed by the vector difference equation

$$y(t) + A_1 y(t-1) + \ldots + A_n y(t-n) =$$

$$= B_1 u(t-1) + \ldots + B_m u(t-m) + v(t) \qquad (1)$$

where $v(t)$ is some (vector) disturbance. Introduce

$$\varphi(t)^T = [-y(t-1)^T \ \ldots \ -y(t-n)^T \ u(t-1)^T \ \ldots \ u(t-m)^T]$$

and

$$\theta^T = [A_1 \ \ldots \ A_n \ B_1 \ \ldots \ B_m]$$

To get a formal similarity with the filtering problem it is convenient to represent the unknown parameters as a vector. Therefore, introduce

$$x = \operatorname{col} \theta$$

which is obtained by writing the columns of θ under each other. This is, of course, not crucial, but it simplifies comparison with the well-known Kalman filter.

Then with

$$\Phi(t)^T = \begin{pmatrix} \varphi(t)^T & & 0 \\ & \ddots & \\ 0 & & \varphi(t)^T \end{pmatrix} \qquad (p \text{ rows})$$

(1) can be written

$$y(t) = \Phi(t)^T x + v(t) \qquad (2)$$

With the ordinary non-Bayesian approach a weighted LS estimate of x at time t, $\hat{\zeta}_t$, is obtained by minimizing with respect to ζ

$$V_t(\zeta) = \frac{1}{t - t_0} \sum_{s=t_0+1}^{t} \| y(s) - \Phi(s)^T \cdot \zeta \|_W^2 \qquad (3)$$

(where W is a positive definite weighting matrix).

Notice that for every diagonal W the estimates will be the same as if W is the identity matrix.

In this paper, however, the parameter vector x is considered as a random vector with known *a priori* mean and covariance. The LS estimate $\hat{\xi}_t$ of x is defined to minimize

$$E(x-\xi)(x-\xi)^T \qquad (3a)$$

with respect to ξ, and is allowed to be a function of the *a priori* information and all measurements up to and including time t.

Now let all random variables be defined on a sample space Ω. In treating consistency, the true system parameters are usually considered as constant but unknown. Then each point ω in the sample space gives a certain realization of the noise sequence $\{v(t)\}$ (and of the input sequence $\{u(t)\}$ in case of random input), whereas the true system parameters are the same for every ω. The concept "almost everywhere", a.e., then means "for almost every realization" for the particular system given. Ljung (1976) has given very general conditions for the LS estimates to be consistent a.e. in the above meaning.

With the true system parameters considered as random variables, the choice of system is regarded as part of the experiment, and each realization starts by picking a system. Then if the distribution function of the parameters is continuous the probability will be zero to have the same true system in two different experiments. Each point ω in the sample space will thus give

1) the true system parameters and 2) a noise sequence. The sample space may be regarded as a product space, so that $\omega = (\omega_1 \quad \omega_2)$, where ω_1 determines the true system and ω_2 determines the noise sequence.

With this point of view, the concept a.e. means "for almost every realization for almost every system". This must be remembered when comparing the results of this paper with other results. In particular nothing can be said about a certain given system when the probability for the system parameters to take any special values is zero.

The following additional notations will be used:

$\mathcal{F}_t$ - the σ-algebra generated by all measurements of y and u up to and including time t.

$\mathcal{F}_\infty$ - the smallest σ-algebra containing $\mathcal{F}_t$ for every t

$\hat{x}_t = E(x|\mathcal{F}_t)$ - the conditional expectation of x given $\mathcal{F}_t$

$P_t = E\left((x-\hat{x}_t)(x-\hat{x}_t)^T|\mathcal{F}_t\right)$ - the conditional covariance of x given $\mathcal{F}_t$

$1_B: \Omega \to R$ - the indicator function for the set B $\left(1_B(\omega) = 1 \text{ if } \omega \in B \text{ otherwise } 1_B(\omega) = 0\right)$

$P(B|\mathcal{F}_t)$ - the conditional probability for B given $\mathcal{F}_t$

3. GENERAL RESULTS

It is well-known (see e.g. Kalman (1960) or Jazwinski (1970)) that under very general circumstances the LS-estimate minimizing (3a) is the conditional mean. This fact makes the following theorem interesting.

Theorem 1: Suppose that the distribution of the true parameters x has finite second moments. Then $\hat{x}_t$ and P_t converge a.e. The limits are denoted by $\hat{x}_\infty$ and P_∞.

Proof: According to Theorem 9.4.5 in Chung (1968) the conditional mean of an integrable variable is a martingale that converges a.e. Now x has finite second moments and each component of the vector $\hat{x}_t$ is a conditional mean. Moreover,

$$(P_t)_{ij} = E\left((x-\hat{x}_t)_i \cdot (x-\hat{x}_t)_j \mid F_t\right) = E(x_i x_j \mid F_t) - (\hat{x}_t)_i \cdot (\hat{x}_t)_j$$

where the first term is a conditional mean and the second one has already been shown to converge. □

In fact, Chung (1968) also shows that the limit $\hat{x}_\infty = E(x \mid F_\infty)$ a.e. In the next theorem this limit is examined.

Theorem 2: With the assumptions of Theorem 1, if M is the set $\{\omega \mid P_\infty = 0\}$ then $1_M \cdot \hat{x}_\infty = 1_M \cdot x$ a.e.
If $P(M) = 1$ then also $\hat{x}_t \to x$ in L^2.

Proof: It is sufficient to consider the scalar case, since $P_t \to 0$ implies that all its diagonal elements tend to zero. Since $M \in F_\infty$, $E(1_M \mid F_t) \to 1_M$ a.e. according to Lévy's zero-or-one law, Chung (1968) p. 313. Then

$$E(1_M \mid F_t)^2 \cdot P_t \to 0 \text{ a.e.}$$

But

$$0 \le E(1_M \mid F_t)^2 \cdot P_t = E(1_M \mid F_t)^2 \cdot E(x^2 \mid F_t) - E(1_M \mid F_t)^2 \cdot \hat{x}_t^2$$

Now both terms of the right member are uniformly integrable since they are less than $E(x^2|F_t)$ (a.e.), which is a conditional mean and thus uniformly integrable by martingale theory [Chung (1968), Theorem 9.4.3]. Also both terms converge a.e. and so they must converge in L^1 [Chung (1968), Theorem 4.5.4]. Then the left member converges in L^1 and a.e., and the limits must be equal, i.e. zero. This means that

$$E\left\{E(1_M|F_t)^2 \cdot E\left((\hat{x}_t-x)^2|F_t\right)\right\} = E\left\{E\left(E(1_M|F_t)^2(\hat{x}_t-x)^2|F_t\right)\right\} =$$

$$= E\left(E(1_M|F_t)(\hat{x}_t-x)\right)^2 \to 0$$

so that

$$E(1_M|F_t)\cdot(\hat{x}_t-x) \to 0 \text{ in } L^2$$

and the last part of the theorem is proven. But Lévy's zero-or-one law and Theorem 1 together imply that

$$E(1_M|F_t)\cdot(\hat{x}_t-x)$$

converges a.e. The limit must be zero because it is in L^2. Then also

$$1_M\cdot(\hat{x}_t-x) \to 0 \text{ a.e.}$$

which proves the theorem. □

Remark: From the proof it is evident that the theorem can be applied component-wise.

These two theorems might also be used in connection with other identification schemes than the LS-method. Then it must be shown that the difference between the estimate and the conditional mean tends to zero. The conditional mean is unfortunately difficult to calculate in general. For the Gaussian case, however, it is equal to the linear LS-estimate, which is given by the Kalman filter equations.

4. MAIN RESULTS. THE GAUSSIAN CASE

The following theorem, given in Åström, Wittenmark (1971), couples the Kalman filter equations to the conditional mean. The LS-estimate $\hat{\xi}_t$ is then given recursively by the same equations since $\hat{\xi}_t = \hat{x}_t$.

Theorem 3 (Åström, Wittenmark): Suppose that the true parameter vector x is Gaussian with *a priori* mean m and *a priori* covariance P_0, $\{v(t)\}$ is a sequence of independent, equally distributed normal vectors with zero mean value and positive definite covariance R, and x and v(t) are independent for all t. Let the output vector of the system be generated by (1). Then the conditional distribution of x given $\mathcal{F}_t$ is normal with mean $\hat{x}_t$ and covariance P_t, where $\hat{x}_t$ and P_t satisfy the difference equations

$$\hat{x}_t = \hat{x}_{t-1} + K(t)[y(t) - \Phi(t)^T\hat{x}_{t-1}] \tag{4}$$

$$P_t = P_{t-1} - P_{t-1}\Phi(t)[R + \Phi(t)^T P_{t-1}\Phi(t)]^{-1}\Phi(t)^T P_{t-1} \tag{5}$$

where

$$K(t) = P_{t-1}\Phi(t)[R + \Phi(t)^T P_{t-1}\Phi(t)]^{-1} = P_t\Phi(t)R^{-1} \tag{6}$$

and the initial conditions are $\hat{x}_{t_0} = m$, $P_{t_0} = P_0$.

Proof: For the single-input single-output case the proof is indicated in Åström, Wittenmark (1971). The extension to the multivariable case is straightforward. The last equality in (6) is proved by multiplying (5) with $\Phi(t)R^{-1}$ from the right to get

$$P_t\Phi(t)R^{-1} = P_{t-1}\Phi(t)\{R + \Phi(t)^T P_{t-1}\Phi(t)\}^{-1} \cdot$$

$$\cdot\left(\{R + \Phi(t)^T P_{t-1}\Phi(t)\}R^{-1} - \Phi(t)^T P_{t-1}\Phi(t)R^{-1}\right) =$$

$$= K(t) \cdot I$$

□

Corollary 1: Under the assumptions of the theorem it follows from Theorem 2 that the estimate $\hat{\xi}_t = \hat{x}_t$ is consistent a.e. and in L^2 provided $P_t \to 0$ a.e.

Note that for Theorem 3 to hold it is necessary to use the correct initial conditions. Equation (4), however, gives a parameter estimate also for other initial values.

Corollary 2: If the parameter estimates are not used in the control law and $P_t \to 0$, then the estimate given recursively by (4) - (6) is consistent a.e. for any initial values $\hat{x}_{t_0}$ and $P_{t_0} > 0$.

Proof: Using the well-known matrix identity

$$(A+BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}$$

(5) can be written

$$P_t^{-1} = P_{t-1}^{-1} + \Phi(t)R^{-1}\Phi(t)^T \tag{7}$$

Rewrite (4) and use (7) to get

$$\hat{x}_t = \left(I - P_t\Phi(t)R^{-1}\Phi(t)^T\right)\cdot\hat{x}_{t-1} + P_t\Phi(t)R^{-1}y(t) =$$

$$= P_tP_{t-1}^{-1}\hat{x}_{t-1} + P_t\Phi(t)R^{-1}y(t)$$

Summing up from $t = t_0$ gives

$$\hat{x}_t = P_tP_{t_0}^{-1}\hat{x}_{t_0} + P_t\sum_{k=t_0+1}^{t}\Phi(k)R^{-1}y(k) \tag{8}$$

Now suppose that S_t satisfies (5) with $S_{t_0}^{-1} = \Delta + P_{t_0}^{-1}$. Then (7) shows that $S_t^{-1} = \Delta + P_t^{-1}$ for all $t \geq t_0$, so that the above matrix identity with $B = \Delta$ and $C = D = I$ gives

$$S_t = P_t - P_t\Delta\{I + P_t\Delta\}^{-1}P_t \tag{9}$$

Since the estimates are not supposed to influence the system, Φ and y will be the same for any initial $\hat{x}_{t_0}$ and P_{t_0}.

Now replace P_t in (8) by S_t and then insert (9). Suppose $P_t \to 0$. The first term of (8) will then tend to zero. Because of (9) the second term will be split up into two, of which the last one tends to zero. The remaining term is equal to the second term of (8), and thus converges to $\hat{x}_\infty$ for any initial values. □

The initial values may, however, be important for other reasons than consistency. Some problems in connection with the initialization of equations (4) - (6) have been studied by e.g. Hagander (1973) and Lainiotis (1976).

Remark: It is well-known (see e.g. Åström (1968)) that the weighted LS-estimate $\hat{\zeta}_t$ is also given by equations (4) - (6) if the weighting matrix $W = R^{-1}$. The initial condition then is $P_{t_0}^{-1} = 0$. Using Corollary 2 when P_t has become invertible a consistency result is obtained also for this case.

Moreover, if R is diagonal then $\hat{\zeta}_t$ is also the ordinary LS-estimate, i.e. it minimizes $V_t(\cdot)$ for $W = I$. Since R must be known it can also be made diagonal by a transformation of variables, and so it is no restriction to assume R diagonal.

Corollary 3: Under the assumptions of the theorem

$$\sum_{s=t_0+1}^{\infty} K(s)K(s)^T < \infty \text{ a.e.}$$

which gives a lower bound to the convergence rate of K(t).

Proof: By Theorem 1 P_t converges a.e. and (5) and (6) give

$$P_t = P_0 - \sum_{s=t_0+1}^{t} K(s)[R + \Phi(s)^T P_s \Phi(s)]K(s)^T$$

The sum is thus convergent. But R positive definite implies $R + \Phi(s)^T P_s \Phi(s) \geq \varepsilon \cdot I$ for some $\varepsilon > 0$.

Now a condition is needed to guarantee that $P_t \to 0$. This is given in the next theorem.

Theorem 4: With notations and assumptions as in Theorem 3 and $P_{t_0} > 0$

$$\{\omega \mid P_t \to 0\} =$$

$$= \left\{\omega \mid \sum_{s=t_0+1}^{\infty} [a^T \varphi(s)]^2 \text{ divergent for every constant column vector } a \neq 0\right\}$$

To prove this the following lemma is needed.

Lemma: Let $\{P_t\}$ be a sequence of positive definite matrices such that $P_t \to P_\infty$ and $P_t - P_\infty$ positive semidefinite for all t. Then

$$P_\infty = 0 \Leftrightarrow a^T P_t^{-1} a \to \infty \quad t \to \infty \quad \text{for every constant column vector } a \neq 0.$$

Proof: The proof is given in Appendix.

Proof of Theorem 4: Theorem 1 gives $P_t \to P_\infty$ a.e. for some $P_\infty \geq 0$. The formula in Theorem 3 for computing P_{t+1} shows that $P_{t+1} - P_t \leq 0$ and so $P_t - P_\infty \geq 0$ for all t. Then the lemma gives

$$P_t \to 0 \Leftrightarrow \tilde{a}^T P_t^{-1} \tilde{a} \to \infty \quad t \to \infty \text{ for every constant vector } \tilde{a} \neq 0$$

Summing up (7) gives

$$P_t^{-1} = P_0^{-1} + \sum_{s=t_0+1}^{t} \Phi(s) R^{-1} \Phi(s)^T$$

Now use that R is positive definite so that $\varepsilon_1 \cdot I \leq R^{-1} \leq \varepsilon_2 \cdot I$ for some $\varepsilon_1 > 0$ and ε_2 and get

$$\tilde{a}^T(P_t^{-1}-P_0^{-1})\tilde{a} = \sum_{s=t_0+1}^{t} \tilde{a}^T\Phi(s)R^{-1}\Phi(s)^T\tilde{a} \leq$$

$$\leq \sum_{s=t_0+1}^{t} [\tilde{a}_1^T \ldots \tilde{a}_p^T] \begin{pmatrix} \varphi(s) & & 0 \\ & \ddots & \\ 0 & & \varphi(s) \end{pmatrix} \cdot \varepsilon_2 I \cdot \begin{pmatrix} \varphi(s)^T & & 0 \\ & \ddots & \\ 0 & & \varphi(s)^T \end{pmatrix} \begin{pmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_p \end{pmatrix} =$$

$$= \varepsilon_2 \cdot \sum_{j=1}^{p} \sum_{s=t_0+1}^{t} (\tilde{a}_j^T\varphi(s)\varphi(s)^T\tilde{a}_j)$$

In the same way

$$\tilde{a}^T(P_t^{-1}-P_0^{-1})\tilde{a} \geq \varepsilon_1 \sum_{j=1}^{p} \sum_{s=t_0+1}^{t} (\tilde{a}_j^T\varphi(s)\varphi(s)^T\tilde{a}_j)$$

Thus

$\tilde{a}^T P_t^{-1}\tilde{a} \to \infty$ for every constant $\tilde{a} \neq 0$

$\Leftrightarrow$

$\sum_{s=t_0+1}^{t} [a^T\varphi(s)]^2$ divergent for every constant $a \neq 0$

This completes the proof. □

Theorem 4 shows that in the Gaussian case (with $v(\cdot)$ being white noise) the only condition needed for consistency a.e. is that

$$\sum_{s=t_0+1}^{\infty} [a^T\varphi(s)]^2$$

be divergent a.e. for every constant vector $a \neq 0$. This condition will be referred to as CC (Consistency Condition).

Now for simplicity consider single-input single-output systems. Then in the open-loop case CC is a condition on the input only, provided the input and the noise are independent. This is realized as follows. For CC not to be fulfilled $a^T\varphi(t)$ must tend to zero as $t \to \infty$. But if any of the a-components corresponding to y-components of $\varphi(t)$ are non-zero, $a^T\varphi(t)$ will contain a noise term from the y. This noise term cannot be cancelled by any other term when the input and the noise are independent, and $a^T\varphi(t)$ cannot tend to zero.

To prove consistency in the open-loop case a common assumption is that the input is persistently exciting. For a definition of this concept see e.g. Åström and Bohlin (1965). To show its relation to CC consider the case $\varphi(t) = u(t-1)$. Then u is persistently exciting of order one only if

$$\lim_{N\to\infty} \frac{1}{N} \sum^{N} u(t)^2 > 0$$

whereas CC only demands that

$$\sum^{\infty} u(t)^2 \text{ diverges}$$

so that u(t) may e.g. decrease to zero with increasing t.

In the closed-loop case CC gives a condition on the feedback. If it is linear and constant, it must be of such a high order that not all of its terms are components in the vector $\varphi(t)$. If it is time-varying it must not converge too fast to a linear and constant feedback of low order.

Example (from Ljung (1974)): Consider the system

$$y(t+1) + x_1 y(t) = x_2 u(t) + e(t+1)$$

where $\{e(t)\}$ is a sequence of independent Gaussian random variables with zero mean and unit variance. Let the input be given by the time-varying feedback

$$u(t) = f(t)y(t)$$

where $f(t) \to f$ as $t \to \infty$. Then with $a^T = [a_1 \quad a_2]$ CC is

$$\sum^{\infty}[a_1 + a_2 f(t)]^2 y(t)^2$$

diverges for every $a \neq 0$. Now $y(t) \not\to 0$ because of the noise, so there must be a subsequence for which $\{y(t)^2\}$ is bounded from below. For CC not to be satisfied it is then necessary that $[a_1 + a_2 f(t)]^2 \to 0$, i.e. $a_1 + a_2 f = 0$. Thus CC is satisfied if

$$\sum^{\infty}[f(t) - f]^2$$

diverges since

$$\sum^{\infty}[a_1 + a_2 f(t)]^2 = a_2^2 \cdot \sum^{\infty}\left[f(t) + \frac{a_1}{a_2}\right]^2 = a_2^2 \cdot \sum^{\infty}[f(t) - f]^2$$

For the case when the minimization of $V_t(\cdot)$ is restricted to a finite set of parameter values this result and the consistency condition (CC) was shown in Ljung (1974), cf. also Ljung (1976). □

Finally the question of non-consistency will be treated. In order to conclude non-consistency a.e. it is not sufficient that P_∞, the *a posteriori* covariance after all the measurements, is nonzero. But if the *a posteriori* distribution is continuous and x is a constant, then with $P_\infty > 0$ the probability will be zero for the estimate to take any particular value, especially the true one. When x is a stochastic variable, however, this is not sufficient. Then x and $\hat{x}_\infty$ might depend on the realization in similar ways, to give a nonzero probability

for $x = \hat{x}_\infty$. The next theorem treats non-consistency in the Gaussian case and couples it to P_∞ being non-zero.

Theorem 5: With the assumptions of Theorem 3

$$P(P_\infty \neq 0,\ \hat{x}_\infty = x) = 0$$

Proof: As in Theorem 2 it is no restriction to consider the scalar case only. According to Theorem 3 the conditional distribution of x given $\mathcal{F}_t$ is normal with mean $\hat{x}_t$ and covariance P_t. Introduce the sets $M_\gamma = \{\omega | P_\infty < \gamma\}$ and $M = \{\omega | P_\infty = 0\}$. Then

$$P(|x-\hat{x}_t| < \varepsilon | \mathcal{F}_t) = \frac{1}{\sqrt{2\pi\ P_t}} \int_{-\varepsilon}^{\varepsilon} e^{-s^2/2P_t} ds \leq$$

$$\leq \begin{cases} 1 \text{ if } \omega \in M_\gamma \\ \\ \dfrac{1}{\sqrt{2\pi\ P_\infty}} \cdot 2\varepsilon \leq \dfrac{1}{\sqrt{2\pi\gamma}} \cdot 2\varepsilon = k(\gamma)\cdot\varepsilon \text{ if } \omega \notin M_\gamma \end{cases}$$

Taking expectations on both sides gives

$$P(|x-\hat{x}_t| < \varepsilon) \leq k(\gamma)\cdot\varepsilon + P(M_\gamma)$$

for all $t > t_0$. Now $P(M_\gamma)$ can be made arbitrarily close to $P(M)$ by choosing γ small enough.

But $\hat{x}_t \to \hat{x}_\infty$ a.e. and so by Egorov's theorem (see e.g. Halmos (1950)) for any $\delta > 0$ there exists a set N with $P(N) > 1 - \delta$ such that $\hat{x}_t \to \hat{x}_\infty$ uniformly on N. Then there is a $T(\varepsilon)$ so that

$$\sup_{\omega\in N} \left| \hat{x}_t - \hat{x}_\infty \right| < \varepsilon$$

for all $t > T(\varepsilon)$.

This gives

$$P(\hat{x}_\infty=x) \leq P(|\hat{x}_\infty-x| < \varepsilon) \leq P(\{|\hat{x}_\infty-x| < \varepsilon\} \cap N) + \delta \leq$$

$$\leq P(|\hat{x}_t-x| < 2\varepsilon) + \delta$$

if $t > T(\varepsilon)$. Now the right member can be made less than $P(M) + 3\delta$ for any $\delta > 0$ by choosing first δ, then γ to make $P(M_\gamma) < P(M) + \delta$ then $\varepsilon\left(0 < \varepsilon < \delta/2k(\gamma)\right)$ and finally $t > T(\varepsilon)$. Thus

$$P(\hat{x}_\infty=x) \leq P(M) = P(P_\infty=0)$$

Then

$$P(\hat{x}_\infty=x, \quad P_\infty \neq 0) = P(\hat{x}_\infty=x) - P(\hat{x}_\infty=x, P_\infty=0) =$$

$$= P(\hat{x}_\infty=x) - P(P_\infty=0) \leq 0$$

where the last equality is implied by Theorem 2. This completes the proof. □

Remark: Theorems 2 and 5 together show that the sets $\{\omega|P_\infty=0\}$ and $\{\omega|\hat{x}_\infty=x\}$ can differ only by a null-set.

Theorems 4 and 5 should be combined to show different cases of non-consistency. A constant and linear feedback of sufficiently low order is, of course, one case, since then $a^T\varphi(t) \equiv 0$ everywhere for some $a \neq 0$, so that CC is satisfied nowhere.

The only difficult cases are when the feedback converges too fast to a linear and constant one. Then the exact limit in convergence rate separating consistency from non-consistency will depend on the stability of the limiting closed-loop system.

Example (continued): Consider again the first-order example given above. If the closed-loop system is stable and

$$\sum^{\infty}[f(t) - f]^2$$

converges then $P_t \neq 0$. But if the closed-loop system is unstable then f(t) must converge faster in order to make

$$\sum^{\infty}[a_1 + a_2 f(t)]^2 y(t)^2$$

convergent, and the required convergence rate depends on how unstable the closed-loop system is, which in turn depends on f, x_1 and x_2.

5. CONCLUSIONS

The two main ideas and results of this paper are 1) the way of looking at the true system parameters as random variables with some assumed *a priori* distribution and 2) the coupling of consistency and non-consistency for the LS-method to the divergence or convergence of a certain series (CC, the Consistency Condition). This condition is shown to be sufficient *and* necessary in the Gaussian white noise case. It may be interpreted as a condition that the input should "shake" the system long enough, in the open-loop as well as in the closed-loop case.

It is interesting to note that the theorems do not require any conditions on the stability of the systems, as do most results previously given. However, in showing consistency using CC, unstable systems seem to require a "less exciting" input than do stable systems.

As for extensions, the case with time-varying noise covariance could be treated. This would affect only Theorem 4 and CC would include the noise covariance. Theorems 1 and 2 are given in a general form, but their possible application to other cases has not been investigated. However, Theorem 4 may be used for any method containing a P-equation as in the LS-case.

ACKNOWLEDGEMENT

The author is glad to express his gratitude to professor L Ljung for many fruitful discussions and good suggestions.

6. REFERENCES

Åström, K J (1968):
Lectures on the Identification Problem - The Least Squares Method. Report 6806, Dept. of Automatic Control, Lund Institute of Technology.

Åström, K J, and Bohlin, T (1965):
Numerical Identification of Linear Dynamic Systems from Normal Operating Records.
IFAC Symp. on Theory of Self-Adaptive Control Systems, Teddington, England. Also in P H Hammond (Ed.), Theory of Self-Adaptive Control Systems, Plenum Press, 1966.

Åström, K J, and Wittenmark, B (1971):
Problems of Identification and Control.
J of Mathematical Analysis and Applications, Vol. 34, pp 90-113.

Chung, K L (1968):
A Course in Probability Theory. Harcourt, Brace & World Inc., 1st edition.

Hagander, P (1973):
Kalman Filters for Processes with Unknown Initial Values.
Report 7332, Dept. of Automatic Control, Lund Institute of Technology.

Halmos, P R (1950):
Measure Theory. D van Nostrand Company Inc.

Jazwinski, A H (1970):
Stochastic Processes and Filtering Theory.
Academic Press.

Lainiotis, D G (1976):
Partitioned Riccati Solutions and Integration-Free Doubling Algorithms.
IEEE Trans on Automatic Control, Vol. AC-21, No. 5, Oct. 1976.

Ljung, L (1974):
On Consistency for Prediction Error Identification Methods. Report 7405, Dept. of Automatic Control, Lund Institute of Technology.

Ljung, L (1976):
Consistency of the Least Squares Identification Method. IEEE Trans on Automatic Control, Vol. AC-21, No. 5, Oct. 1976.

Kalman, R E (1960):
A New Approach to Linear Filtering and Prediction Problems. ASME J Basic Engineering, Vol. 82, pp. 35-45.

Mann, H B, and Wald, A (1943):
On the Statistical Treatment of Linear Stochastic Difference Equations. Econometrica, Vol. 11, pp. 173-220.

APPENDIX

Lemma: Let $\{P_t\}$ be a sequence of positive definite matrices such that $P_t \to P_\infty$ and $P_t - P_\infty$ positive semidefinite for all t. Then

$$P_\infty = 0 \Leftrightarrow a^T P_t^{-1} a \to \infty \qquad t \to \infty$$

for every constant column vector $a \neq 0$.

Proof of the lemma: First suppose that $P_\infty = 0$. Let λ_t be the smallest eigenvalue of P_t^{-1}. Then

$$a^T P_t^{-1} a \geq \lambda_t a^T a \to \infty \qquad t \to \infty$$

for every constant $a \neq 0$, since all eigenvalues of P_t tend to zero.

Next suppose $P_\infty \neq 0$, and assume that it is diagonal. This is no restriction since it is symmetric and thus can be diagonalized. At least one of the elements of P_∞ must be non-zero, say the (1,1)-element. Then put

$$P_\infty = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & \end{pmatrix} \quad \text{and} \quad \bar{P} = \begin{pmatrix} \lambda_1 & & 0 \\ & 0 & \\ & & \ddots \\ 0 & & 0 \end{pmatrix}$$

so that $P_\infty \geq \bar{P}$. Also introduce $A_t = P_t - \bar{P}$ and $\tilde{A}_t$ with

$$A_t = \left(\begin{array}{c|ccc} a_{11}^t & a_{12}^t & a_{13}^t & \cdots \\ \hline a_{21}^t & a_{22}^t & a_{23}^t & \cdots \\ a_{31}^t & a_{32}^t & a_{33}^t & \\ \vdots & \vdots & & \ddots \end{array}\right) = \left(\begin{array}{c|cc} a_{11}^t & a_{12}^t & \cdots \\ \hline a_{21}^t & & \\ \vdots & & \tilde{A}_t \\ \vdots & & \end{array}\right)$$

Then $A_t \geq 0$.

Now

$$\det P_t = \left|\begin{array}{c|c} \lambda_1+a_{11}^t & a_{12}^t \cdots\cdots \\ \hline a_{21}^t & \\ \vdots & \tilde{A}_t \\ \vdots & \end{array}\right| =$$

$$= \left|\begin{array}{c|c} \lambda_1 & a_{12}^t \cdots\cdots \\ \hline 0 & \\ \vdots & \tilde{A}_t \\ 0 & \end{array}\right| + \left|\begin{array}{c|c} a_{11}^t & a_{12}^t \cdots\cdots \\ \hline a_{21}^t & \\ \vdots & \tilde{A}_t \\ \vdots & \end{array}\right| =$$

$$= \lambda_1 \det \tilde{A}_t + \det A_t \geq \lambda_1 \det \tilde{A}_t$$

so that for the (1,1)-element of P_t^{-1}

$$(P_t^{-1})_{11} = \frac{\det \tilde{A}_t}{\det P_t} \leq \frac{\det \tilde{A}_t}{\lambda_1 \det \tilde{A}_t} = \frac{1}{\lambda_1}$$

Thus with $a^T = [1 \quad 0 \quad 0 \ \ldots \ 0]$

$$a^T P_t^{-1} a \leq \frac{1}{\lambda_1} a^T a \not\to \infty$$

as $t \to \infty$.

Part III – Regulators for Time-varying Stochastic Systems

ABSTRACT

Suboptimal dual regulators for stochastic adaptive systems are considered. A new algorithm is proposed and a survey of previously suggested schemes is given. It is discussed for what kind of system the new algorithm will be advantageous, and when a simpler scheme will suffice. This discussion is illustrated by simulations where different regulators are compared.

1. INTRODUCTION

Every control system is subject to disturbances from its environment. A successful regulator design then requires that the disturbances are taken into account. This indicates the need to model the disturbances. One way to do this is to use stochastic processes. This leads to stochastic control theory.

There is one type of stochastic control problems whose optimal solution has been known for long, namely the linear-quadratic-gaussian problem. In this problem the system dynamics and the measurements are assumed to be linear with known parameters and additive gaussian noise. The object of control is to minimize a cost functional, which in the discrete time case is a sum of squares of deviations of the input and the output from their desired values. In an admissible control law the input is a function of all the present information, i.e. initial data and currently available measurements.

This problem admits an analytical solution through dynamic programming. A recursive equation for the loss, the Bellman equation, is then obtained, which will give the optimal solution. The linear-quadratic-gaussian model has been used in many practical problems, often with a very good result. One drawback with this model is that it requires a great deal of knowledge about the system to be controlled. In practice the system parameters are often not known and may even be time-varying. They might then be modelled as stochastic processes, just as the disturbances. This gives a nonlinear stochastic control problem. A survey of different regulators for stochastic adaptive systems is given in Wittenmark (1975b).

For such problems, the Bellman equation for the loss may still be derived. Even if it cannot be solved the Bellman equation gives useful information about the character of the optimal control law. Feldbaum (1960) realized this and introduced the concept of dual control. By this is meant that the purpose of control is twofold. It shall estimate unknown parameters and at the same time keep the output at a desired value. These objec-

tives are often conflicting, since estimation is usually improved by increasing the input. In some situations, however, this is not the case. It may happen in special cases that the quality of estimation will not at all be influenced by the inputs. An example of this is the above mentioned linear-quadratic-gaussian problem. Such systems are called neutral by Feldbaum (1960).

In general it is not possible to calculate optimal dual control laws. Much effort has therefore been spent in trying to find good suboptimal controls having dual properties. A straightforward method is to first estimate the system parameters using all available information, and then use the estimates in a control law as if they were the true values. This is called certainty equivalence. Cautious controls take also the estimate covariances into account when calculating the input.

In Jacobs and Patchell (1972) the effects of dual control laws are explained using three terms. These are: a certainty equivalence input, a cautious correction term because of uncertain parameters, and a probing term to improve the estimates. Suboptimal dual control laws are often derived as some combination of these 3 parts.

One approach in designing suboptimal dual control laws is to neglect probing. Whatever estimates that come out of the estimation algorithm are then used, and nothing is done to improve future estimates. Parameter learning is then completely accidental. Such control laws are called passively adaptive by Bar-Shalom and Tse (1974). They use the term actively adaptive for the opposite case, when the control law is designed also to improve the quality of future estimates.

This report considers different regulators having dual properties. A model for the system and the optimality criterion are given in chapter 2. The algorithms considered are discussed in chapter 3. A new control law is described in chapter 4. To compare regulators some simulations were performed. They are collected in chapter 5. The last chapter contains some concluding remarks.

2. THE SYSTEM AND THE CRITERION

The reason for setting up a certain mathematical model of a system can be either practical or theoretical. In the first case the model should describe a specific real system as well as possible in some sense. Such models are often complicated - nonlinear, time-varying and maybe infinite-dimensional or stochastic. Analysis is then difficult or even impossible.

If the model is set up for theoretical reasons, it is chosen to fit in with the analysis needed to show the desired point. In such a way a lot of results have come out that are valid for certain classes of systems. An example of this is the common use of linear, time-invariant, finite-dimensional models. One measure of the value of the theory is then how well it works on real systems.

For dual control problems, however, it is difficult to find theoretical models that give tractable calculations. Only in simple or special cases has it been possible to calculate an optimal dual control law and express it explicitly as a function of the available information, see e.g. Sternby (1976).

In dual control theory it is thus necessary to use simple models and to make approximations in order to get results. The models should then be chosen to give as few and as accurate approximations as possible. On the other hand the models should also capture some properties of real systems.

Models

In this report only single-input single-output systems will be considered. The model used is (cf Åström and Wittenmark (1971))

$$y(t) = \varphi(t)\theta(t) + e(t) \qquad (2.1)$$

$\varphi(t)$ is a row-vector consisting of functions of old outputs and

inputs (up to and including time t-1). The unknown parameters are collected in the column vector $\theta(t)$, and $e(t)$ is assumed to be (scalar) gaussian white noise with zero mean and variance σ^2. The evolution of $\theta(t)$ is given by

$$\theta(t+1) = \Phi\theta(t) + v(t+1) \tag{2.2}$$

where Φ is a known matrix and $\theta(0)$ is a gaussian random vector with mean m and covariance P_o. The noise $v(t)$ is also assumed gaussian and white with zero mean and covariance R. The sequences $\{e(t)\}$ and $\{v(t)\}$ and the vector $\theta(0)$ are independent.

The parameter model (2.2) could be made more sophisticated. The parameters may e.g. be taken as the outputs from a general discrete time Markov process. However, from a practical point of view even the model (2.2) is a bit too detailed. For example, it seems unrealistic to assume knowledge about the dynamics of unknown parameters. Thus in practice Φ will have to be a unity matrix.

With

$$\varphi(t) = [-y(t-1),\ldots,-y(t-n_a),\ u(t-1),\ldots,u(t-n_b)] \tag{2.3}$$

(2.1) becomes an ordinary linear difference equation model, but the formulation allows also nonlinear models, as long as they are linear in the parameters. In the following $\varphi(t)$ will be assumed to have the form (2.3), and the parameter vector is taken as

$$\theta(t)^T = [a_1(t),\ldots,a_{n_a}(t),\ b_1(t),\ldots,b_{n_b}(t)]$$

Note that the case with constant but unknown parameters is obtained by putting $\Phi = I$ (the identity matrix) and $R = 0$. Unfortunately, this does not simplify the calculations.

State-space models

In the literature it is common to consider linear state-space equations

$$x(t+1) = A\big(t,\theta(t)\big)x(t) + B\big(t,\theta(t)\big)u(t) + w(t) \qquad (2.4)$$

$$y(t) = C\big(t,\theta(t)\big)x(t) + e(t) \qquad (2.5)$$

The unknown parameters $\theta(t)$ are often assumed to enter linearly into the equations, and to vary as in (2.2). The model (2.4) - - (2.5) can then be transformed into a difference equation with dependent equation noise, where the unknown parameters enter linearly as in (2.1). Unlike (2.1), however, all parameters will not get the same time index. This fact shows up also in the estimation.

The state-space approach has been taken by e.g. Bar-Shalom and Sivan (1969), Tse and Athans (1972), Ku and Athans (1973), Saridis and Lobbia (1972), Frost (1970) and Jacobs and Hughes (1975). The reason for using this model is probably the successful and wide-spread use of it in the known parameter case. With unknown parameters, however, some difficulties are added. As will be shown later on, most dual controllers need an estimate of the parameters $\theta(t)$. The best estimate in the least squares sense is the conditional mean. With the model (2.4) - (2.5) it is not possible to calculate this exactly, unless the whole state is measured exactly, or A and C are completely known. Thus approximations must be made already at this stage, e.g. by using an extended Kalman filter as is usually done. This is somewhat unsatisfactory, since the convergence properties of extended Kalman filters are not yet fully understood. This problem is avoided by using the model (2.1).

Estimation

Most regulators need some estimate of the parameters. The conditional mean is a good choice. To calculate this the amount of available information must be specified. This is an important point, which is connected to the choice of admissible control laws. In this report F_t is defined as the σ-algebra generated by all outputs up to and including time t, i.e. $y(t)$, $y(t-1)$, ..., and all previously applied inputs, i.e. $u(t-1)$, $u(t-2)$, ...

For the model (2.1) - (2.2) the conditional distribution of future values of θ is gaussian and it can easily be computed using an ordinary Kalman filter, Åström, Wittenmark (1971). The estimate of $\theta(t)$ calculated at time t-1 will be denoted by $\hat{\theta}(t)$ and its covariance matrix by $P(t)$.

Then

$$\hat{\theta}(t) = E[\theta(t) \mid F_{t-1}]$$

$$P(t) = E[(\theta(t) - \hat{\theta}(t))(\theta(t) - \hat{\theta}(t))^T \mid F_{t-1}]$$

Recursive equations for $\hat{\theta}(t)$ and $P(t)$ are given by

$$\hat{\theta}(t+1) = \Phi\hat{\theta}(t) + K(t)[y(t) - \varphi(t)\hat{\theta}(t)] \tag{2.6}$$

$$P(t+1) = \Phi P(t)\Phi^T + R - \frac{\Phi P(t)\varphi(t)^T\varphi(t)P(t)\Phi^T}{\sigma^2 + \varphi(t)P(t)\varphi(t)^T} \tag{2.7}$$

where

$$K(t) = \Phi P(t)\varphi(t)^T(\sigma^2 + \varphi(t)P(t)\varphi(t)^T)^{-1} \tag{2.8}$$

The model (2.1) thus admits an exact solution of the estimation problem.

To handle time-varying parameters the matrix R in (2.7) should be different from zero. Another way to treat this case is to keep R = 0, but instead divide the right member of (2.7) by a so-called forgetting factor, which should be close to, but less than unity. Then $\hat{\theta}(t)$ as given by (2.6) will no longer be the conditional mean of $\theta(t)$, but may still be a good estimate.

The noise assumption

The equation noise e is in this report assumed to be white. The case with dependent noise can, however, be handled if the dependence is known. Let

$$e(t) = w(t) + c_1 w(t-1) + \ldots + c_{n_c} w(t-n_c) \qquad (2.9)$$

where w(t) is gaussian white noise. Then the c:s could be included in $\varphi(t)$ and the w:s in $\theta(t)$. However, if the c-parameters are not known, it is again not possible to calculate the conditional means exactly. This is, of course, the normal situation, but will not be considered in this report, since the white noise case is difficult enough.

One suboptimal way to handle unknown c-parameters may be first to estimate them using e.g. a recursive ML-method, and then take the estimates as being the correct values. This use of the certainty equivalence principle can be motivated. It has been noted that in many cases the accuracy of the c-parameter estimates does not depend as much on the input as that of the other estimates. This is illustrated in the following example.

Example 2.1: Consider the system

$$y = bu + cv + e$$

where b and c are unknown parameters, v is a measurable disturbance and e is random with zero mean and unit variance. All time indices are dropped, since the system will be studied for one step only. If v had been the old value of e this would have

been an ordinary system with dependent noise, but v is assumed measurable here to make computations easy. In the previous notations

$$\varphi = [\, u \quad v \,]$$

$$\theta^T = [\, b \quad c \,]$$

Suppose

$$P_0 = \begin{pmatrix} p_b^0 & 0 \\ 0 & p_c^0 \end{pmatrix}$$

Then (2.7) becomes

$$\begin{pmatrix} p_{bb} & p_{bc} \\ p_{bc} & p_{cc} \end{pmatrix} = P = P_0 - \frac{P_0 \varphi^T \varphi P_0}{1 + \varphi P_0 \varphi^T}$$

which gives for p_{bb} and p_{cc}

$$p_{bb} = \frac{p_b^0 (1 + p_c^0 v^2)}{p_b^0 u^2 + (1+p_c^0 v^2)}$$

$$p_{cc} = \frac{p_c^0 + p_b^0 p_c^0 u^2}{p_b^0 u^2 + (1+p_c^0 v^2)}$$

The input u may be chosen to reduce the uncertainties, but v is fixed. The variance of $\hat{b}$ will stay between 0 and p_b^0 depending on the value of u. To get a small variance u should be made big.

But the variance of $\hat{c}$ can vary only between $\left(1/(1+p_c^0 v^2)\right) p_c^0$ and p_c^0. Now put $v^2 = 1$ = the variance of e, which is the expected value of v^2 if v were the previous e. Then if p_c^0 is not too big

the variance of $\hat{c}$ is always very close to p_c^0 and little can be gained by choosing u. Note also that as expected the best choice of u to get a good $\hat{c}$ is always u = 0. In fig. 1 is shown how p_{bb}/p_b^0 and p_{cc}/p_c^0 vary with u for $p_c^0 = 0.25$.

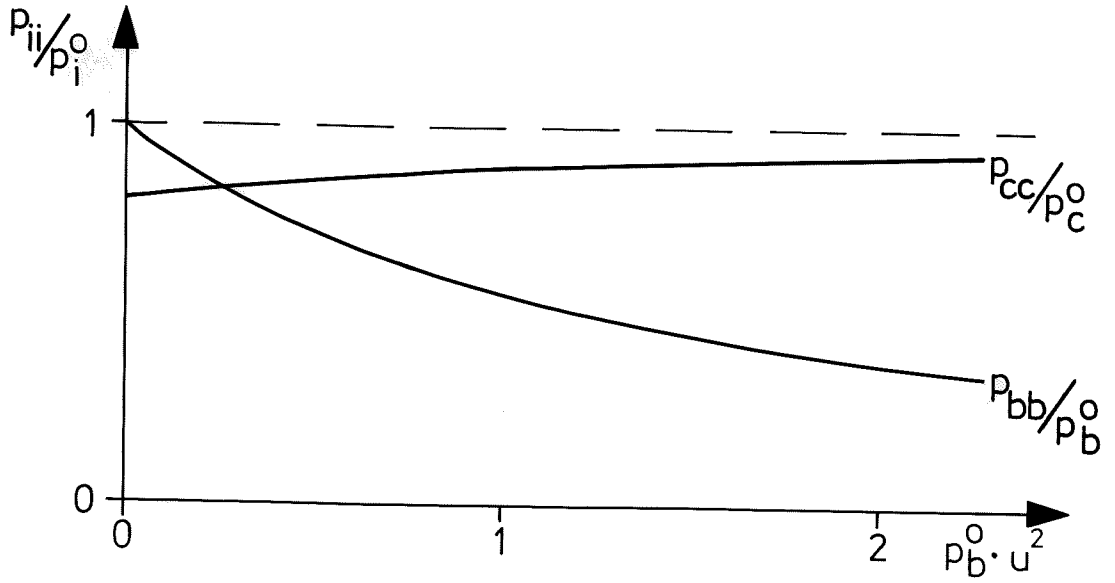


Fig. 1 - The covariances of $\hat{b}$ (p_{bb}) and $\hat{c}$ (p_{cc}) related to their previous values.

□

The example suggests that the estimates of c-parameters are not much influenced by the input signal. For general dual control problems, however, nothing is known about the effect of using certainty equivalence to take care of c-parameters.

Criteria

For the model (2.1) - (2.3) a cost functional is given as

$$J_N = \sum_{k=1}^{N} \left(y(k) - y_r(k)\right)^2 \qquad (2.10)$$

where $y_r(t)$ is the reference value for the output. The optimal input is determined by minimizing the expected value of (2.10).

A penalty on the control actions can be included by adding the term $q(u(k) - u_r(k))^2$ in (2.10). This is important in practice because it makes it possible to limit the size of the control signal. No penalty on the control will, however, be included in this study because it will require the determination of an extra parameter.

Some authors have used a time-varying weighting of the terms in (2.10). This is necessary to treat problems where e.g. only the final state is important and not the whole trajectory. Such problems are not studied here.

When the model (2.4) - (2.5) is used, the cost functional is usually quadratic in the states and in the input. This choice makes computations feasible in the known parameter case. With unknown parameters approximations must again be made.

When minimizing the expected value of J_N in (2.10) the value of N will influence the characteristics of the resulting control law. Two different situations will be considered in this report.

For N small the control law must try to estimate the parameters quickly in order to gain anything in the following steps from an accurate estimation. This situation is of interest when dealing with industrial batch processes, e.g. cement-making, pulp-making and some types of steel-making. In chapter 5 some simulations for small N are shown where the true parameters have been taken as constants ($\Phi = I$ and $R = 0$). These simulations thus show how fast the parameter-adaptation will be with different regulators.

A large value of N corresponds to a steady state situation. Suppose that the true parameters are constants and that the true system is described correctly by (2.1). With $\{e(t)\}$ white and gaussian the estimates will then converge to the true parameters with probability one under very general conditions, see e.g. Ljung (1976) or Sternby (1977). Every good regulator should then converge to its deterministic counterpart, and produce a

corresponding steady state loss. This situation can be analyzed through the ordinary differential equations given by Ljung (1977). For large values of N and constant parameters the problem is thus understood to a reasonable degree. Little is, however, known about the time-varying case for large N. This is investigated through simulations in chapter 5.

A third property to study could be the ability to track changes in the reference value. In Norman (1976) three regulators are compared in this respect by simulations of a first-order system. This question will, however, be left aside in this report.

Admissible Control Laws

Finally must be defined what is an admissible control law. It is natural to let the input at time t, u(t), be a function of all information available at this moment. This can be expressed in probabilistic terms as $u(t) \in \mathcal{F}_t$. The *a priori* known statistics of the noise and the parameters are then considered as constants.

For the gaussian case all this information are collected in the estimate and its covariance together with the vector φ of old inputs and outputs.

The optimal control problem may thus be summarized as follows. Find an admissible control law for the system (2.1) - (2.3) that minimizes the expected value of (2.10) for a certain value of N.

3. SOME DIFFERENT REGULATORS

A functional equation for the minimal expected value of the loss J_N in (2.10) is derived in Åström, Wittenmark (1971).

Introduce the vector $\tilde{\varphi}(t)$, which is $\varphi(t)$ with $u(t-1)$ replaced by a zero. The triplet $(\hat{\theta}(t), P(t), \tilde{\varphi}(t))$ will be denoted by ζ_t. Also introduce V_t as

$$V_t = V_t(\zeta_t) = \min E\left[\sum_{k=t}^{N} (y(k) - y_r(k))^2 \,\middle|\, F_{t-1}\right] \tag{3.1}$$

The minimization is with respect to the present and future admissible control laws. Then by dynamic programming

$$V_t(\zeta_t) = \min_{u(t-1)} E\left[(y(t) - y_r(t))^2 + V_{t+1}(\zeta_{t+1}) \,|\, F_{t-1}\right] \tag{3.2}$$

This equation, the Bellman equation, can be solved only in a few special cases. One example is the ordinary linear-quadratic problem, which is obtained by putting $P(t) \equiv R = 0$. Another one is a system where the parameters vary as white noise, i.e. $\Phi = 0$ in (2.2), see e.g. Tou (1963) or Gunckel and Franklin (1963). Unfortunately, these cases all give non-dual control laws, whereas the interesting cases of dual control remain unsolved.

The Bellman equation has also been solved numerically for a couple of problems. However, even with only a few unknown parameters this leads to extensive computer calculations, so this is no method to use in practice.

The nature of difficulties of solving (3.2) will now be discussed. Let $\ell = [0,\ldots,0,1,0,\ldots,0]$ be a vector with the same dimension as $\varphi(t)$ and with its only non-zero element at the position corresponding to $u(t-1)$. Then V_N is

$$V_N = \sigma^2 + \tilde{\varphi}(N)P(N)\tilde{\varphi}(N)^T + [y_r(N) - \tilde{\varphi}(N)\hat{\theta}(N)]^2 -$$

$$- \frac{[\ell\hat{\theta}(N)(y_r(N) - \tilde{\varphi}(N)\theta(N)) - \ell P(N)\tilde{\varphi}(N)^T]^2}{[\ell\hat{\theta}(N)]^2 + \ell P(N)\ell^T} \qquad (3.3)$$

and

$$u(N-1) = \frac{\ell\hat{\theta}(N)(y_r(N) - \tilde{\varphi}(N)\hat{\theta}(N)) - \ell P(N)\tilde{\varphi}(N)^T}{[\ell\hat{\theta}(N)]^2 + \ell P(N)\ell^T} \qquad (3.4)$$

The next step should now be to insert (3.3) into (3.2). Then two difficulties arise. Firstly $\hat{\theta}(N)$ and $\tilde{\varphi}(N)$ are stochastic variables that are not measurable w.r.t F_{N-2}. With the conditionally gaussian variable $\hat{\theta}(N)$ in the denominator of (3.3) the expectation cannot be given in a form suitable for further calculations.

Moreover, the right member of (3.3) depends in a very complicated way on u(N-2). Via (2.7) P(N) is itself a function of u(N-2). Also the distributions of $\hat{\theta}(N)$ and $\tilde{\varphi}(N)$ depend on u(N-2). Thus even with an analytical expression for the conditional expectation there would be very little hope for a possibility to minimize (3.2) analytically. In this report various approximations will first be discussed and then compared by simulations.

The resulting regulators can be classified into two main groups with some sub-groups each. These are

I REGULATORS WITH ACCIDENTAL LEARNING
- a) Self-tuning regulators (STURE)
- b) Cautious regulators
- c) Open-loop-optimal feedback type regulators (OLOF)

II REGULATORS WITH ACTIVE LEARNING
- a) Based on modifications of the one-step loss function
- b) Based on approximations of the future loss

This taxonomy is not perfect. Some of the regulators could in fact be placed in either of two groups. However, the classification may be of some help to look at different proposed methods in a systematic way.

Note that the regulators in the first group may also show active learning. This is because they are modified to be able to handle also time-varying systems. The classification is based on the original algorithms.

The different regulators will now be discussed.

3.1. Regulators with Accidental Learning

3.1.1. Self-tuning regulators

This regulator is treated to some detail in Åström, Wittenmark (1973) and in Åström, Borisson, Ljung, Wittenmark (1977). A similar regulator is treated in Clarke and Gawthrop (1975), where particular attention is given to command following by feed forward. The regulator is there called a Self-Tuning Controller.

The basic self-tuning regulator is obtained by calculating the minimum variance regulator with the true system parameters replaced by their estimates (certainty equivalence). With only one time-delay in the system the best one-step prediction of y should then be made equal to y_r. From (2.1)

$$\varphi(t+1)\hat{\theta}(t+1) = y_r(t+1)$$

which gives

$$u(t) = \frac{1}{\hat{b}_1(t+1)}[y_r(t+1) - \tilde{\varphi}(t+1)\hat{\theta}(t+1)] \qquad (3.5)$$

It is shown in Åström and Wittenmark (1971) that for known systems, the minimum variance regulator minimizes the expected value of (2.10) for any N. The self-tuning regulator is thus

the approximative solution obtained by considering the estimates as being completely correct. The estimate covariances P are not taken into account, so STURE is a non-cautious regulator. It is also clear that the inputs are not at all chosen to improve the estimates, i.e. learning is accidental.

In practice it has been found useful to limit the control signal. This may affect learning.

This regulator is primarily designed for unknown, but constant systems. For constant minimum-phase systems driven by white noise it has been shown always to converge to the true minimum variance controller, Ljung, Wittenmark (1974).

Taking $R \neq 0$ in the parameter model (2.2) the identification procedure can also track time-varying systems, and STURE can be made to work at least in some cases. It is shown in Appendix A that for a system with only the first a-parameter unknown and $y_r \equiv 0$ STURE is optimal if $N \leq 3$ in (2.10).

The self-tuning regulator has been applied successfully to several real life systems. A review is given in Åström, Borisson, Ljung, Wittenmark (1977).

Simulations have also shown that STURE performs well on many time-varying systems for a large N in (2.10), see Wittenmark (1973), Abramowicz, Stymne (1975), and chapter 5 of this report. There is, however, one problem. For constant parameters STURE will meet difficulties if the estimate of the leading b-parameter has the wrong sign. The closed-loop system is then unstable, and the outputs will grow. In Ljung, Wittenmark (1974) and (1976) it is shown that the large outputs will improve the identification to give good estimates more quickly. The closed-loop system will then be stable again after a while. For time-varying systems this may not be enough if the b_1-parameter changes sign often, or is frequently very close to zero.

As can be seen from (3.5) STURE will give very large inputs

when $\hat{b}_1(t)$ is close to zero. If the true b_1-parameter is not small this will cause the output to be large. It is therefore necessary to put a bound on the amplitude of the input. This may on the other hand be dangerous if the system is unstable, since it restricts the possibilities to drive the system back to normal if the output has once drifted away.

To summarize, it seems as if the self-tuning regulator can be used successfully on some time-varying systems, as long as the leading b-parameter does not change sign or comes very close to zero frequently. Some precautions must then be taken.

3.1.2. Cautious regulators

A cautious regulator is obtained by taking N = 1 in (2.10) and using the resulting control law (3.4) all the time (with N replaced by t+1). The input generated in this way is denoted $u_1(t)$. Again learning will be accidental, since the design is made as if the new estimates will never be used (N = 1). The difference compared to STURE is that the cautious regulator also takes the uncertainties of the estimates into account.

This regulator will not have the same problems as STURE for small $\hat{b}_1$, because the input given by (3.4) will be finite even if $\hat{b}_1 = 0$. There will, however, be other problems. The following may happen. Suppose that the input for some reason is small for a while. Then identification will be poor, so that the variance term in the denominator of the input expression will grow. This causes the input to remain small and so on. This effect is called turn-off, and has been discussed by several authors, e.g. Wittenmark (1971), Åström and Wittenmark (1971), Hughes and Jacobs (1974) and Alster and Bélanger (1974).

When turn-off does not occur, the cautious regulator seems to work very well. Thus it is natural to try preventing turn-off in some way. This can be done simply by adding an extra signal directly to the input to ensure good identification all the time. Then some questions arise. What is the most suitable size

and form of the extra signal? When should an extra signal be added, all the time or just now and then?

A simple method is to take the extra input as

$$u_e(t) = c(-1)^t \tag{3.6}$$

where the constant c has to be chosen in advance. Of course, (3.6) could be replaced by a Pseudo Random Binary Signal, which is often employed in system identification. The optimal choice of c will depend on the true system. This method is discussed in e.g. Wieslander, Wittenmark (1971). In Nakamura, Nakamura (1973) white noise is used as an extra signal.

Since the turn-off problem comes from $\ell P \ell^T$ (the variance of $\hat{b}_1$) growing too large, it seems better to assure that the sign of u_e is such that it helps decreasing $\ell P \ell^T$. A straightforward way is to give $u_e(t)$ the same sign as $u_1(t)$. This gives

$$u_e(t) = c \cdot \text{sign}[u_1(t)] \tag{3.7}$$

A closer examination of the equation for P, (2.7), shows that this may not be the best way. From (2.7)

$$\ell P(t+2)\ell^T = \ell\Phi P(t+1)\Phi^T\ell^T - \frac{\ell\Phi P(t+1)\varphi(t+1)^T\varphi(t+1)P(t+1)\Phi^T\ell^T}{\sigma^2 + \varphi(t+1)P(t+1)\varphi(t+1)^T} + \ell R\ell^T \tag{3.8}$$

Only the second term of the right member depends on u(t). With $\varphi(t) = \tilde{\varphi}(t) + \ell u(t)$ this term is

$$\frac{[\ell\Phi P(t+1)\tilde{\varphi}(t+1)^T + \ell\Phi P(t+1)\ell^T u(t)]^2}{\sigma^2 + \tilde{\varphi}(t+1)P(t+1)\tilde{\varphi}(t+1)^T + 2\ell P(t+1)\tilde{\varphi}(t+1)^T u(t) + \ell P(t+1)\ell^T u(t)^2} \tag{3.9}$$

The maximum of (3.8) is obtained as the minimum of (3.9), i.e. by making the numerator equal to zero. Now, in practice, Φ is usually chosen to be diagonal since the dynamics of the para-

meters is seldom known. Then (3.8) is symmetric around its maximum and decreasing with increasing distance from it. The sign of $u_e(t)$ should then be chosen to maximize the distance of the total input $u_1(t) + u_e(t)$ from the maximum of (3.8).

For long periods of time turn-off does not occur, and accidental learning is sufficient. It is then unnecessary to use any extra input, which just will increase the loss. One possibility is to apply the extra input only when the variance of $\hat{b}_1$ is greater than a certain limit. This gives

$$u_e(t) = \begin{cases} c \cdot \text{sign}[u_1(t)] & \text{if } \ell P(t+1)\ell^T \geq P_{lim} \\ 0 & \text{if } \ell P(t+1)\ell^T < P_{lim} \end{cases} \qquad (3.10)$$

This scheme has been tested by simulations in chapter 5.

A slight modification is to test $P_{b_1}/\hat{b}_1^2$ against a limit instead of just P_{b_1}. The motivation for this is that the difficult cases are when $\hat{b}_1$ is close to zero compared to the uncertainty.

Another possibility is to add the extra signal if the absolute value of the ordinary cautious input (3.4) is smaller than a preset value. Again this value has to be chosen in advance.

The master thesis by Mannerfelt (1977) shows a method to determine the size of the extra input. The idea is to allow the minimal one-step loss to increase by a certain fraction λ due to the extra input. With $u(t) = u_1(t) + u_e(t)$ the one-step loss is

$$E\left\{[y(t+1) - y_r(t+1)]^2 | F_t\right\} = [\tilde{\varphi}(t+1)\hat{\theta}(t+1) + \ell\hat{\theta}(t+1)u(t) - y_r(t+1)]^2 +$$

$$+ [\tilde{\varphi}(t+1) + \ell u(t)]P(t+1)[\tilde{\varphi}(t+1) + \ell u(t)]^T + \sigma^2 =$$

$$= V_1^* + u_e(t)^2\left[(\ell\hat{\theta}(t+1))^2 + \ell P(t+1)\ell^T\right] = (1+\lambda)V_1^* \qquad (3.11)$$

where V_1^* is the minimal one-step loss. Some terms have disappeared in the second last expression, because $u_1(t)$ minimizes (3.11). The last equality of (3.11) is the defining relation for $|u_e(t)|$.

An advantage with this method is that it makes the choice of magnitude of $u_e(t)$ more problem-independent, while the input itself depends very much on parameters, noise and reference value. It also makes $|u_e(t)|$ time-varying, hopefully according to the need for identification. A disadvantage is that with known parameters the total cost will be precisely the fraction λ greater than its achievable minimum. This could be taken care of by letting λ depend on the uncertainty with $\lambda = 0$ for $P = 0$.

Hughes and Jacobs (1974) combine the questions of the size of the extra input and when to apply it. They take the total input $u_1 + u_e$ as the function of u_1 shown in fig. 2.

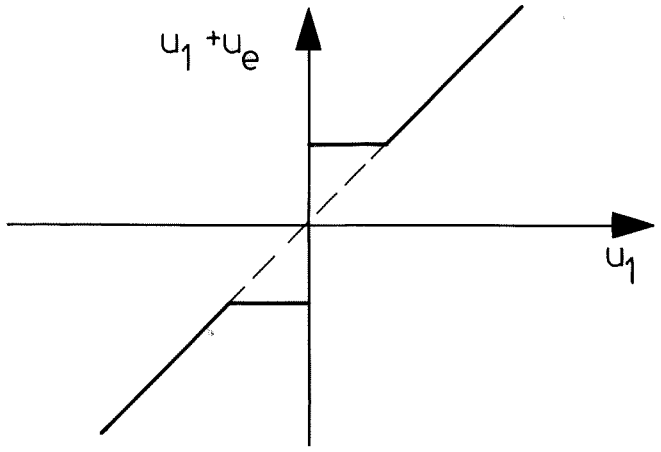


Fig. 2 - The total input of Hughes and Jacobs (1974).

Again the threshold value has to be chosen *a priori*.

It is worth noting that the difficult case for both the basic self-tuning regulator and for the basic cautious one is when $\hat{b}_1$ is close to zero. Both regulators then give extreme inputs, but in different ways. Consider a case where only one parameter is estimated, the first b-parameter, so that the last term

in the numerator of (3.4) disappears. Then for $\hat{b}_1 = 0$ STURE will give an infinite input, while the cautious regulator gives zero input. Also for all other values of $\hat{b}_1$ the self-tuning input is larger than the cautious one. These observations suggest the use of an input which lies between those of STURE and the cautious regulator. The arithmetic mean will give an infinite input in the above example and cannot be used. Instead the harmonic mean could be tried. For the case considered this gives

$$u_h(t) = \frac{\ell\hat{\theta}(t+1)[y_r(t+1) - \tilde{\varphi}(t+1)\hat{\theta}(t+1)]}{[\ell\hat{\theta}(t+1)]^2 + \frac{1}{2}\ell P(t+1)\ell^T} \tag{3.12}$$

The only difference to (3.4) is that in (3.12) P is divided by 2, which makes this regulator less cautious.

Florentin (1962) has calculated numerically the optimal control for the system

$$y(t+1) = y(t) + bu(t) + e(t+1)$$

where only b is unknown. The loss function was as (2.10) except that it also included the input. The reference value was $y_r(t) \equiv \equiv 0$. It turned out that the optimal control could be approximated by

$$u_f(t) = -\frac{\ell\hat{\theta}(t+1)y(t)}{[\ell\hat{\theta}(t+1)]^2 + \frac{1}{2}\ell P(t+1)\ell^T + 1} \tag{3.13}$$

It is remarkable that the only difference with (3.12) is the last term of the denominator, which naturally comes in when the input is included in the loss function. Because of the form of the inputs (3.12) and (3.13) it is, however, questionable whether either has any dual action.

3.1.3. Open-loop-optimal feedback (OLOF) type controllers

An early discussion of the OLOF strategy is given in Dreyfus (1964). A discrete state three-stage decision problem is discussed. It is shown that OLOF gives a performance which is very close to optimal. On the other hand, in the Markov chain example of Sternby (1976), certainty equivalence control is better than OLOF.

The basic approximation in controllers of OLOF type is that all future inputs are taken as functions of present information only. This is in contrast with the active learning case, where future inputs are allowed to be functions of future information. Using the OLOF approximation the expected value of the loss is minimized by dynamic programming. Unlike the previously discussed regulators the effects of the present input on future outputs are taken into account to some extent.

Further approximations must be made in the general case. They give different versions of the OLOF controllers. In Tse and Athans (1972) the model (2.4) - (2.5) is considered with θ entering only in the B-matrix. The loss function is quadratic in state and input.

The behaviour of the algorithm can be tested only by simulations. This was done to some extent in Tse and Athans (1972). In Ku and Athans (1973) the results are extended to include also unknown A-parameters. Further approximations must then be done.

Mehra (1974) has given a similar derivation of the optimal open loop loss for an impulse response model. The solution is obtained from a two point boundary value problem. Applying only the first input and recalculating at every step then yields an OLOF controller.

In Jacobs and Hughes (1975) another approximation of future loss is made. A corresponding deterministic problem is considered, where all random variables are replaced by their means.

The optimal closed loop loss for this problem is then used to approximate the future loss in the original problem. As usual the calculations have to be redone at every step. The resulting regulator, which could also be classified as a cautious regulator, is called neutral.

When using an OLOF controller the input is calculated as if no further measurements are going to be made. Another approach is to assume that measurements are made, but that they will not be used to improve the estimates or their covariances. The future parameter values are then assumed to be random variables that are independent of the inputs, the outputs and the noise. Their distribution is determined by the current estimate and its covariance. This gives a Riccati equation for the expected value of the loss. The resulting regulator will still only give accidental learning of parameters, since the input is calculated as if it could not improve the estimates. Such a regulator can be expected to handle changes in the reference value quite well since the output is measured and fed back. In Norman (1976) simulations indicate that this is actually the case. The regulator is there called open-loop-mean-variance (OLMV). The amount of computing necessary is a draw-back with this method.

Murphy (1968) makes still another approximation. When calculating the expected value of future loss the input is supposed to be a time-varying linear feedback from future estimates. The time-varying gains are then calculated using dynamic programming based on linearization around a predicted trajectory. The feedback thus consists of two factors, the gains and the future estimates. The gains are deterministic, but the future estimates are random variables. This regulator is something in between an OLOF regulator and one with active learning. The approximation made may be rather crude, since the optimal regulator is certainly not a linear feedback from the estimates.

3.2. Regulators with Active Learning

3.2.1. Modifications of the one-step loss function

A dual control law must compromise between good control and good parameter identification. Since the quality of identification can be expressed by the variance of the estimates, one possibility is to include the variance in the loss function.

Alster and Bélanger (1974) use as cost functional the expected value of (2.10) with N = 1. However, the minimization is done under the constraint

$$\text{tr } P(t+2)^{-1} \geq M \tag{3.14}$$

where M is a constant to be chosen. This gives a lower limit to the input when P is large since

$$P(t+2)^{-1} = P(t+1)^{-1} + \frac{1}{\sigma^2}\varphi(t+1)^T\varphi(t+1) \tag{3.15}$$

Thus

$$u(t)^2 \geq M - \text{tr } P(t+1)^{-1} - \frac{1}{\sigma^2}\tilde{\varphi}(t+1)\tilde{\varphi}(t+1)^T \tag{3.16}$$

When the right member of (3.16) is less than zero this regulator is identical to the basic cautious regulator. The effect of (3.16) is to decrease the periods of turn-off. They are unfortunately not completely eliminated.

Wittenmark (1975a) uses the cost functional

$$W_W = E\left\{[y(t+1) - y_r(t+1)]^2 | F_t\right\} + \lambda f(P(t+2)) \tag{3.17}$$

where f is a scalar function of the covariance matrix of the estimates in the next step. The weighting factor λ has to be chosen in advance. Both terms in W_W depend on u(t), and the compromise between control (= the first term) and identification (= second term) is evident.

The choice of the function f is also discussed in Wittenmark (1975a). It is pointed out that the leading b-parameter is the most important one. Therefore a good choice may be $f(P) = \ell P \ell^T$, i.e. to pick out the variance of the leading b-parameter. Another choice could be $f(P) = \mathrm{tr}(P)$.

In Wittenmark (1975a) simulations indicated that the value of λ is not very critical. This point will be further examined in the simulations of chapter 5. To minimize (3.17) its explicit dependence on u(t) must be shown. The model (2.1), (2.7) and the definitions of ℓ and $\tilde{\varphi}$ give

$$
\begin{aligned}
W_w = {} & \left[u(t)\ell\hat{\theta}(t+1) + \tilde{\varphi}(t+1)\hat{\theta}(t+1) - y_r(t+1)\right]^2 + \\
& + \left[\tilde{\varphi}(t+1) + \ell u(t)\right] P(t+1) \left[\tilde{\varphi}(t+1) + \ell u(t)\right]^T + \sigma^2 + \\
& + \lambda f \Bigg(\Phi P(t+1) \Phi^T + R - \\
& - \frac{\Phi P(t+1)\left[\tilde{\varphi}(t+1) + \ell u(t)\right]^T \left[\tilde{\varphi}(t+1) + \ell u(t)\right] P(t+1) \Phi^T}{\sigma^2 + \left[\tilde{\varphi}(t+1) + \ell u(t)\right] P(t+1) \left[\tilde{\varphi}(t+1) + \ell u(t)\right]^T} \Bigg)
\end{aligned}
\qquad (3.18)
$$

From this equation it seems likely that W_w has several local minima. As a possibility for minimizing (3.18) Wittenmark (1975a) suggested numerical optimization. Some precaution must then be taken to get convergence to the correct local minimum.

A systematic way of choosing the smallest minimum is obtained by the following argument. Suppose that ℓ^T is an eigenvector of Φ^T. This means that $b_1(t+1)$ may depend on $b_1(t)$ and noise, but not on the other parameters. In practice this is mostly fulfilled, since Φ usually has to be chosen diagonal. Also suppose that $f(P) = \ell P \ell^T$. The first term of (3.17) is minimized by the cautious input $u_1(t)$. With $u_2(t)$ denoting the maximum of $\ell P(t+2)\ell^T$ the dependence of u(t) in the two terms of (3.17) may look as in fig. 3.

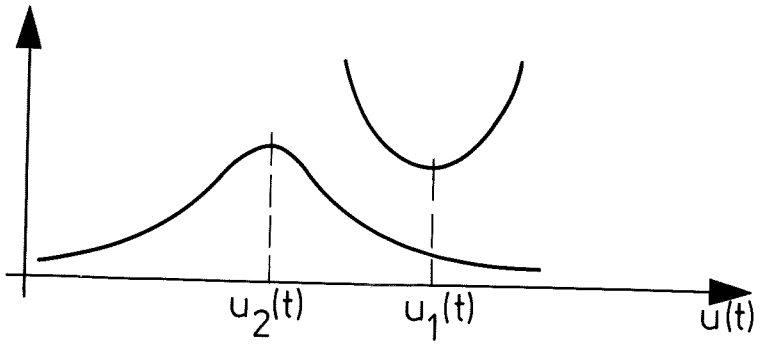


Fig. 3 - Typical behaviour of the two terms of (3.17).

It is proven in appendix B that there is exactly one local minimum of (3.17) to the right of $u_1(t)$, and that this is also the global minimum. Because of symmetry a corresponding result holds if $u_1(t) < u_2(t)$. If $u_1(t) = u_2(t)$ one of two things may happen. Either $u_1(t) = u_2(t)$ is the only local minimum, or it is a local maximum. In the latter case there is exactly one local minimum on each side with the same value of (3.17).

The starting value for the numerical minimization can be taken as the cautious input $u_1(t)$ with a first step away from $u_2(t)$.

With $f(P) = \mathrm{tr}(P)$ or $f(P) = \ell P \ell^T$ another method can be used to avoid the problems. Then W_w is a rational function of $u(t)$, and its derivative can be calculated analytically. The numerator of the derivative will be a fifth order polynomial. Using a root-finding algorithm the zeroes of the derivative can be found and the global minimum is determined by comparing the corresponding values of (3.17). The root-finding may be time consuming (of the order of tenths of seconds on a PDP-15). Numerical minimization has the advantage that it can be stopped after any number of iterations.

3.3.3. Approximations of future loss

The most complicated of the proposed approximative dual control laws is the wide-sense adaptive dual control algorithm suggested by Tse, Bar-Shalom and Meier (1973). They consider a general nonlinear system with additive noise in both states and measurements. The cost functional is also nonlinear. Tse and Bar-Shalom (1973) applied this algorithm to the model (2.4) - (2.5) with a corresponding cost functional, quadratic in the states and the input. The main steps are as follows.

For each value of u(t) considered, an approximate loss is calculated. To do so, first x(t+1) is predicted from u(t) and the estimates at time t. A nominal future trajectory is then chosen, e.g. as given by an OLOF controller or the solution to a corresponding deterministic linear-quadratic problem. The system is linearized around this nominal and the loss is minimized up to second order. Finally the immediate loss is added to give the total expected loss, which is minimized by selecting u(t). One way to look upon the algorithm is to assume that measurements are going to be made in the next step, but not thereafter. With an OLOF controller it is assumed that no measurements are made even in the next step.

The wide-sense adaptive algorithm thus gives a linear-quadratic problem to solve for each new input that is tried. The time interval considered in this problem should, of course, equal the time to go in the original problem, but in order to save computing time a shorter time interval may be used.

All these steps have to be gone through for several input values at each sample point, since the total loss has to be minimized numerically to obtain the optimal input. Therefore this algorithm may require substantial computations, especially when the number of steps to go is large. Another drawback is that the algorithm is hard to analyze, and little is known about the effects of the various approximations. The numerical minimization also requires some attention, since the loss function may well have several local minima.

Because of the complexity of this method it has not been included in the simulations of chapter 5. For a simple integrator example a comparison with the cautious regulator and the two-step regulator defined below can be found in Sternby and Pernebo (1977). It is shown that for the particular example the wide-sense adaptive and the two-step regulators give similar results. In particular, they seem to generate almost the same inputs.

The wide-sense adaptive algorithm is, however, very general, and can be used not only to keep a system in steady-state. In Tse and Bar-Shalom (1973) simulations are shown for cases when the objective has instead been to achieve a certain final state.

Another regulator based on approximation of future loss has been suggested by Chow (1975). His model is of the state equation type and the whole state is observed. The criterion is quadratic, but not additive, because it contains also products of outputs from different times. An approximate dynamic programming is performed corresponding to (3.2). In each step the minimal future loss is approximated with a quadratic function of all previous states. This is done by calculating the derivatives numerically. It is then possible to calculate the expected value of this quadratic function and continue the dynamic programming. This is also a rather laborious method. Some possible simplifications are also discussed by Chow (1975).

The last regulator to be discussed in this report is obtained by putting $N = 2$ in (2.10). This is believed to introduce some of the required dual effect. It is, however, not possible to solve even this problem exactly. Further approximations must be made. This regulator will be discussed to some detail in the next chapter.

4. THE TWO-STEP REGULATOR

4.1. Motivation

Taking N = 2 in (2.10) may give a control law that works satisfactorily also for longer times. To show this a few examples will be given.

Example 4.1.1. The first one is taken from Sternby (1976), where a controlled Markov chain is considered. For special choices of transition probabilities and loss function it is possible to calculate e.g. the optimal control law, certainty equivalence control and the two-step regulator analytically. For this example the difference in performance between the optimal control and the two-step regulator is negligible, while certainty equivalence control performs somewhat worse. □

Example 4.1.2. Another example is given in Sternby and Pernebo (1977). A wide-sense adaptive dual control and the two-step regulator are compared on a first order system, an integrator with unknown and time-varying gain. The comparison is made using Monte Carlo simulations. The wide-sense adaptive dual control is derived for 3 different values of the number of steps to go, N = 1, N = 2 and N = 20. With N = 20 the Riccati-equation for the corresponding linear quadratic problem will reach very close to its steady state. The results can therefore not be expected to be improved much by increasing N further.

As expected the control law obtained with N = 20 performs better than that with N = 2, but the difference is surprisingly small, and within the estimated standard deviation. The average loss per step is 1.51±0.04 with N = 20 and 1.54±0.04 with N = 2. A fairly good performance with a loss of 1.65±0.05 per step is also obtained for N = 1. One reason is that for the wide-sense adaptive regulator the estimate is not allowed to become close to zero, because that would make the corresponding linear-quadratic problem singular. □

Example 4.1.3. As a last example consider the scalar system

$$x(t+1) = ax(t) + bu(t) + v(t+1) \tag{4.1}$$

$$y(t) = cx(t) + e(t) \tag{4.2}$$

where a, b and c are known parameters and $\{v(t)\}$ and $\{e(t)\}$ are independent sequences of independent gaussian variables with zero mean and unit variance. With known parameters any value of N in the loss function (2.10) will give the same control law. To get a difference it is necessary to include also the input in the loss function. The input is thus determined to minimize the expected value of

$$\frac{1}{N} \sum_{s=t+1}^{t+N} x(s)^2 + ku(s-1)^2 \tag{4.3}$$

where k is some given constant. The aim with this example is to compare the control laws obtained by using N = 1, N = 2 or letting $N \to \infty$ in (4.3). This is a standard problem, and its solution can be found in standard texts such as Åström (1970).

All control laws have the form

$$u(t) = L_N \hat{x}(t) \tag{4.4}$$

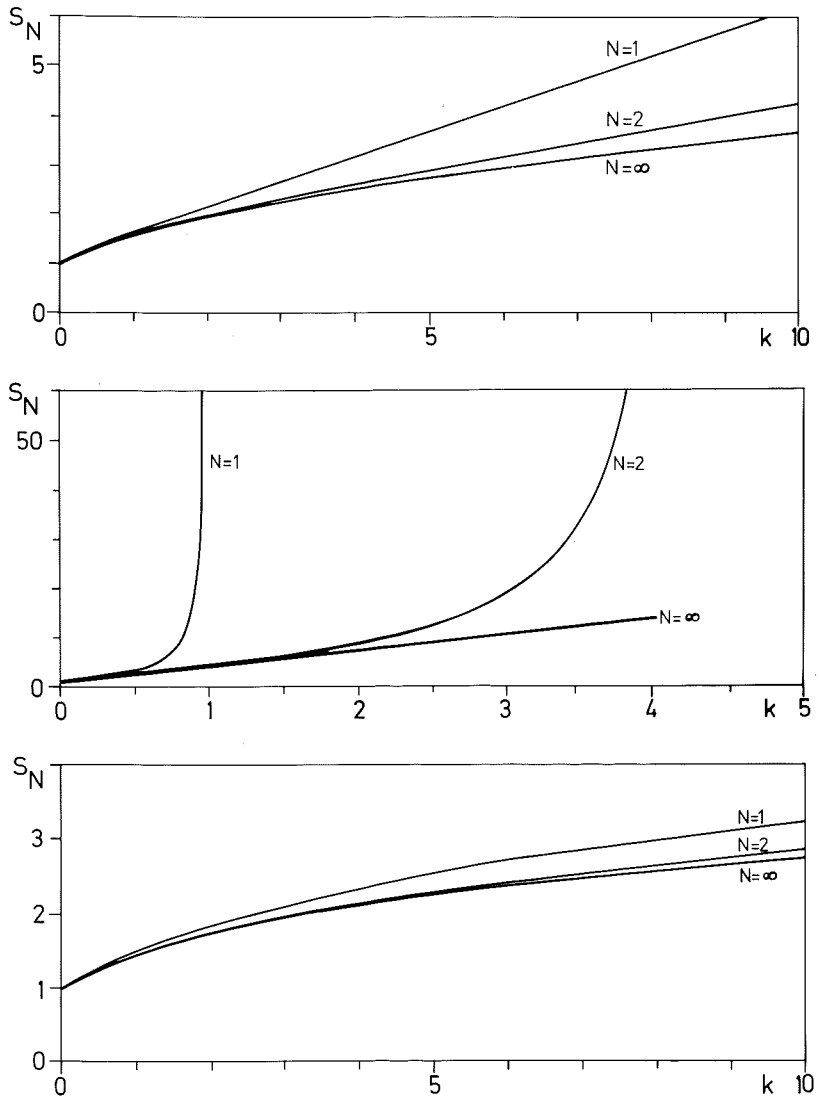
The average value of the loss per step is

$$P + S_N \frac{a^2 c^2 P^2}{c^2 P + 1} \tag{4.5}$$

where P is the stationary value of the estimate variance and S_N is the stationary solution to the associated Riccati equation. The only difference between the three control laws is the value of S, whereas P will be the same.

It turns out that S_N depends only on N, a and k/b^2. Therefore

choose b equal to one. Figs. 4 - 6 show S_N as a function of k for some values of a. Each figure contains one curve for each of the three N-values.



Figs. 4 - 6 - S_N for N = 1, 2 and ∞ with a = 1 (fig. 4), a = 2 (fig. 5) and a = 0.9 (fig. 6).

In fig. 4 a = 1. Then S_2 is rather close to S_∞. The two-step approximation is thus quite good in this case. It can be shown that as $k \to \infty$ $S_\infty \sim \sqrt{k}$, $S_1 \sim \frac{k}{2}$ and $S_2 \sim \frac{k}{4}$.

When the system is unstable, k must not be taken too large if a finite N-value is used. This would result in an unstable closed-loop system, because the input is kept too small to stabilize it. This is seen in fig. 5 where a = 2. Still N = 2

is much better than N = 1, but only on infinite N keeps S finite for all k.

For $|a| < 1$ S_N tends to the same limit with any N as $k \to \infty$. Thus N = 1 can be used without much extra loss if k is very large or very small. With a = 0.9 as in fig. 6 there is an interval for k where S_2 is much closer to S_∞ than to S_1. Then again the two-step approximation is good. When $|a|$ is further decreased the differences between the three control laws will rapidly vanish completely for all values of k. Then, of course, the one-step approximation is excellent, and there is no need to have N = 2. □

In conclusion, the examples of this section have shown that there are cases when the two-step approximation works very well. There are, however, other cases when the improvement over the one-step approximation is small, or when the approximations do not work at all.

Another problem is caused by non-minimum-phase systems. It is well-known (Åström (1970)) that even for systems with known parameters the criterion (2.10) will then give an optimal system that is extremely sensitive to parameter errors. For unknown parameters control laws obtained with this criterion can therefore not be expected to work well. To overcome this problem the input should be included in the loss function. But even then there will be problems if too few steps are considered in the input calculation.

4.2. Derivation and Approximations

In chapter 3 the Bellman equation (3.2) was stated. The optimal loss in the last step, V_N, was also given in equation (3.3). To compute the best two-step input u(N-2), it is necessary to calculate $E(V_N | F_{N-2})$. As was discussed in chapter 3 this cannot be done analytically, so some approximation must be made. It is then important to remember that the difficult situations are when the estimate of the first b-parameter is close to zero.

The approximation chosen should therefore be accurate at least in such cases. It would also be an advantage if the approximation could simplify minimization of the expected loss w.r.t. the input. It is then preferable to be able to do this analytically. But also with a numerical minimization it must be possible to keep track of every local minimum to assure that the global minimum is found. Some different approximations will now be discussed.

With F_{N-2} given all random parts of V_N in (3.3) are generated by $y(N-1)$. However, V_N can also be written with $\ell\hat{\theta}(N)/\sqrt{\ell P(N)\ell^T} = x$ as the basic random variable. Using this notation the problem is to find approximations to

$$E\,\frac{f(x)}{1+x^2} \tag{4.6}$$

where $x \in N(m,\sigma)$. They should be most accurate for small m/σ, and should not give too difficult a minimization to perform afterwards.

Gauss approximation

One method is to replace all random variables by their means. This is called Gauss approximation and gives for (4.6)

$$E\,\frac{f(x)}{x^2+1} \approx \frac{f(m)}{m^2+1} \tag{4.7}$$

The uncertainty σ does not enter into (4.7) at all, which is probably a disadvantage. For $m = 0$ this approximation is good only when $\sigma \approx 0$, i.e when the b-parameter estimate is very accurate.

Denominator expansion

Another possibility is to make a serial expansion of the denominator around the mean of x^2 before taking the expectation. This gives

$$E\frac{f(x)}{1+x^2} = E\frac{1}{1+m^2+\sigma^2}\cdot\frac{f(x)}{1+\frac{x^2-m^2-\sigma^2}{1+m^2+\sigma^2}} \approx$$

$$\approx \frac{1}{1+m^2+\sigma^2}\cdot\sum_{k=0}^{n} Ef(x)\left(\frac{m^2+\sigma^2-x^2}{1+m^2+\sigma^2}\right)^k \qquad (4.8)$$

With n = 0 (4.8) is

$$E\frac{f(x)}{1+x^2} \approx \frac{Ef(x)}{1+m^2+\sigma^2} = \frac{E(\text{numerator})}{E(\text{denominator})} \qquad (4.9)$$

This is the approximation used for the two-step regulator. It turns out that for small values of m/σ this is a much better approximation than (4.7). To illustrate this the correct value (4.7), (4.9) and (4.8) with n = 2 have been calculated for a simple case where f(x) = 1. Fig. 7 shows the result for m = 0 as functions of σ.

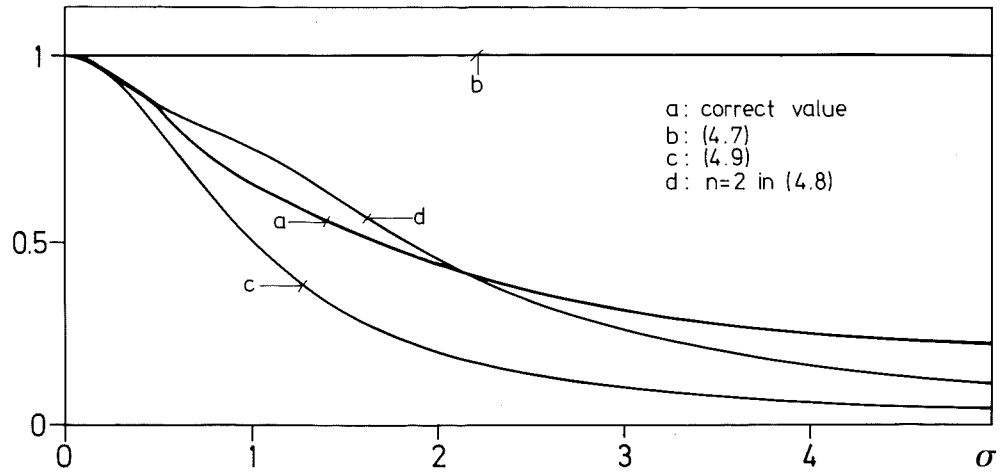


Fig. 7 - Approximations of (4.6) for m = 0.

The accuracy of (4.9) decreases with growing m. For m = 1 and σ close to zero the correct value increases with σ while (4.9) decreases. This is seen in fig. 8. This approximation is thus not very good for small σ-values when m/σ is not close to zero. Inclusion of one more term from (4.8) obviously takes care of these problems. However, this extra term would make the two-step regulator much more complicated.

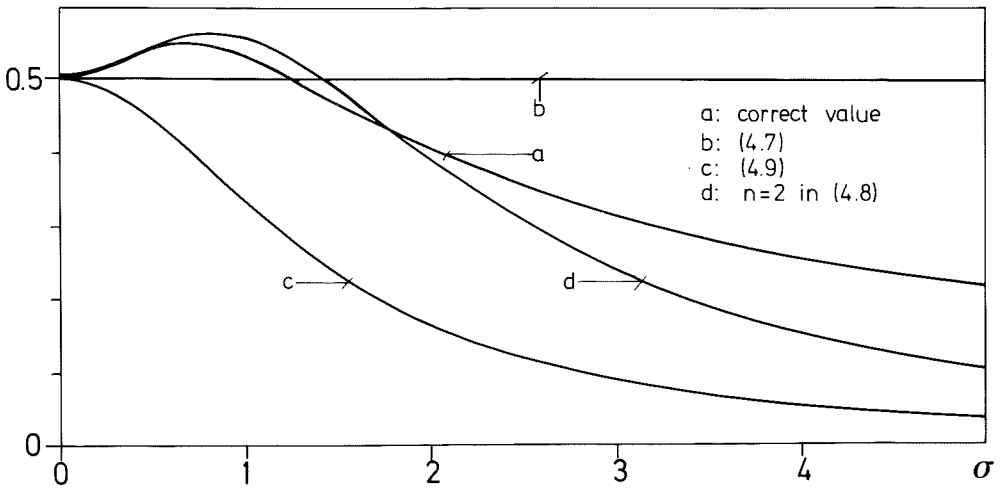


Fig. 8 - Approximations of (4.6) for m = 1.

With a minimization to follow, the derivatives are also of interest. Fig. 8 shows that for m = 1 (4.9) may be not so bad for $\sigma \gtrsim 0.8$. The dependence on m for a fixed σ (= $1/\sqrt{2}$) is shown in fig. 9. Again (4.9) is best for small m/σ-values, but the error is almost constant, and the error of the derivative will thus be small.

To summarize, the approximation (4.9) seems to be good when m/σ is small, which is a desired property. For larger m-values there is an area around σ = 0 where it is not so good. Comparing (4.7) and (4.9) the latter is better, at least in the important case of small m/σ-values. This is one reason for choosing (4.9) in this report. Another reason is that, as pointed out below, the minimization to follow will be easier than with (4.7).

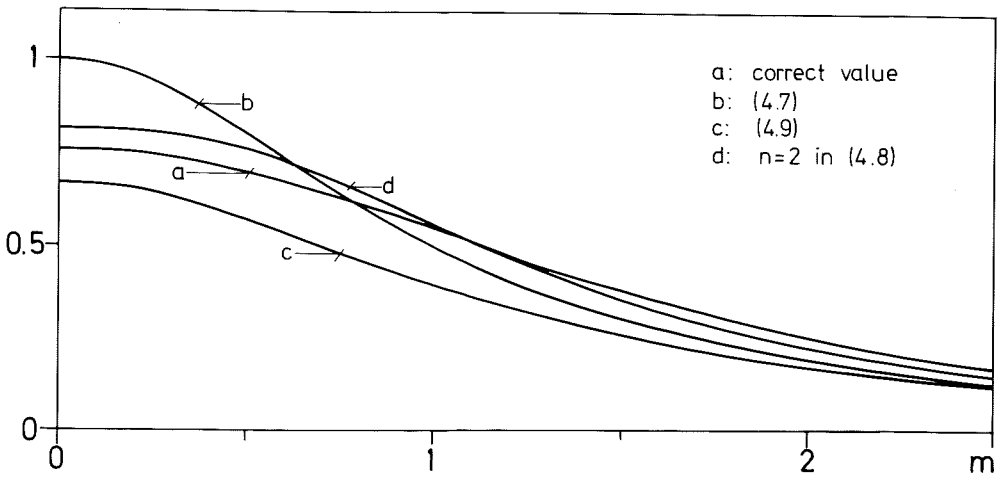


Fig. 9 - Approximations of (4.6) for $\sigma = 1/\sqrt{2}$.

The two-step regulator

When using (4.9) on two different expressions for the same quantity the results may be different as for

$$E(x^2) = m^2 + \sigma^2$$

and

$$E\left\{\frac{x^2(1+x^2)}{1+x^2}\right\} \approx \frac{E(x^2+x^4)}{E(1+x^2)} = m^2 + \sigma^2 + \frac{2\sigma^2[\sigma^2 + 2m^2]}{1 + m^2 + \sigma^2}$$

This is taken into account in the following way when $E(V_N | F_{N-2})$ is calculated from (3.3). The last two terms of (3.3) is a difference between similar terms if P is small. The difference is zero if P is zero. Therefore (4.9) is applied to the difference, and not just to the last term, even though the second last term could be treated without approximation.

The result from using (4.9) on $E(V_N | F_{N-2})$ is now inserted into (3.2). This gives a rational function of u(N-2), which is de-

rived in appendix C. To find the minimum the derivative is taken and equated to zero. This gives a 5th order polynomial to solve. The input, finally, is determined by a direct comparison of the 5 corresponding function values.

It is worth noting that the expected value of the denominator of (3.3) does not depend on u(N-2). This fact simplifies the minimization when using approximation (4.9). On the other hand it is possible that this simplification introduces a major difference to the optimal regulator. Anyway, if Gauss approximation were used, the denominator would depend on u(N-2), and the derivative would be a 9th order polynomial instead. This would more than triple the time required to solve for its zeroes, at least with the root-finding algorithm used in the simulations of this report.

Since the approximation (4.9) is not very accurate when m/σ is large, it might be better not to use it then. For such cases the basic cautious or self-tuning regulators are probably good enough. A modification of the two-step algorithm would then be to use (4.9) only when m/σ is small and (3.4) or (3.5) otherwise. The effect of this can only be tested by simulations.

4.3. Relations to Other Regulators

Wittenmark's regulator

There is one special case when the two-step regulator and the regulator by Wittenmark (1975a) minimizing (3.17) are very similar. This is a system with no a-parameters and only one b-parameter. Then (3.3) reduces to

$$\sigma^2 + \frac{y_r(N)^2 P_{bb}(N)}{\hat{b}(N)^2 + P_{bb}(N)}$$

on which the approximation (4.9) is applied. Using the Kalman filter equations (2.6) - (2.8) the expected value of the denominator is

$$E\left\{[\hat{b}(N)]^2 + P_{bb}(N) \mid F_{N-2}\right\} =$$

$$= E\left\{\left[\Phi_{bb}\hat{b}(N-1) + K(N-1)\left(y(N-1) - u(N-2)\hat{b}(N-1)\right)\right]^2 + \right.$$

$$\left. + \Phi_{bb}^2 P_{bb}(N-1) + R_{bb} - \frac{\Phi_{bb}^2 P_{bb}(N-1)^2 u(n-2)^2}{\sigma^2 + u(N-2)^2 P_{bb}(N-1)} \,\middle|\, F_{N-2}\right\} =$$

$$= [\Phi_{bb}\hat{b}(N-1)]^2 + \Phi_{bb}^2 P_{bb}(N-1) + R_{bb} \qquad (4.10)$$

which is independent of $u(N-2)$. With this approximation (3.2) becomes

$$V_{N-1} = \min_{u(N-2)}\left\{E\left[\left(y(N-1) - y_r(N-1)\right)^2 \mid F_{N-2}\right] + \right.$$

$$\left. + \sigma^2 + \frac{y_r(N)^2}{[\Phi_{bb}\hat{b}(N-1)]^2 + \Phi_{bb}^2 P_{bb}(N-1) + R_{bb}} P_{bb}(N)\right\} \qquad (4.11)$$

But minimizing (4.11) is exactly the same thing as minimizing (3.17) for $t = N-2$ if

$$f\left(P(N)\right) = P_{bb}(N) \qquad (4.12)$$

and

$$\lambda = \frac{y_r(N)^2}{[\Phi_{bb}\hat{b}(N-1)]^2 + \Phi_{bb}^2 P_{bb}(N-1) + R_{bb}} \qquad (4.13)$$

Thus in this case the two-step regulator is equivalent to the regulator of Wittenmark (1975a) with a time-varying λ. Intuitively it seems useful to let λ vary, since the need for identification varies, and is largest when $\hat{b}_1^2/P_{bb}$ is close to zero. It could therefore be interesting to try using (4.13) or some

modification of this type in Wittenmark's regulator also for more complicated systems.

An advantage with Wittenmark's regulator is that numerical minimization can be used to find the global minimum of the loss function. Because of the possibilities for several local minima a root-finding algorithm is employed for the two-step regulator, which is a bit time-consuming. A middle way could be to assume all parameters known except for the first b-parameter, and then use the (4.9) approximation. This might simplify the minimization, and if the leading b-parameter is much more important than the others the increase in loss might be small. This possibility has not yet been examined closer.

Cautious regulator

An interesting question to ask is: How big is the difference between the two-step regulator and the basic cautious one? This point will now be studied for the system treated above with only one b-parameter and no a-parameters. To simplify further put $\Phi_{bb} = y_r = 1$. Firstly (4.11) is expressed in terms of u(N-2) with λ as in (4.13). For easy notations the arguments are dropped. The first term is

$$L_1 = (\hat{b}u-1)^2 + u^2P + \sigma^2 = (P+\hat{b}^2)u^2 - 2\hat{b}u + 1 + \sigma^2 \qquad (4.14)$$

and the last term is

$$L_2 = \lambda\left(P + R - \frac{P^2u^2}{\sigma^2 + u^2P}\right) = \lambda\left(\frac{\sigma^2P}{\sigma^2 + u^2P} + R\right) \qquad (4.15)$$

The graphs of L_1 and L_2 may look as shown in fig. 10.

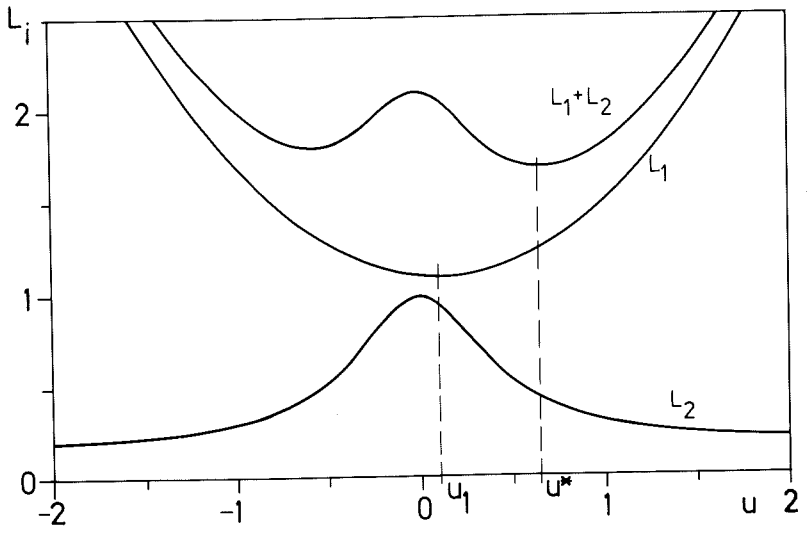


Fig. 10 - A typical behaviour of L_1, L_2 and L_1+L_2.

The minimum of L_1 is denoted u_1 and the total minimum u^*. The derivative of (4.14) is

$$\frac{dL_1}{du} = 2u(P+\hat{b}^2) - 2\hat{b} \tag{4.16}$$

which gives

$$u_1 = \frac{\hat{b}}{P + \hat{b}^2} \tag{4.17}$$

As shown in appendix B u^* and the maximum of L_2 are always on different sides of u_1. When u_1 is away from zero, u^* and u_1 are relatively close. But if u_1 is close to zero and the maximum of L_2 is narrower than the minimum of L_1, then u^* and u_1 are quite different. This is the case if $\hat{b}$ is small, and if estimation can be much influenced by choosing u, which it can if the second derivative of L_1+L_2 is negative at $u = 0$. Now

$$\left.\frac{d^2(L_1+L_2)}{du^2}\right|_{u=0} = 2(P+\hat{b}^2) - \frac{2}{\hat{b}^2+P+R}\cdot\frac{P^2}{\sigma^2} =$$

$$= 2\hat{b}^2 + 2P\,\frac{\hat{b}^2+R+P\left[1-\frac{1}{\sigma^2}\right]}{\hat{b}^2+P+R} \qquad (4.18)$$

To get (4.18) negative it is necessary that $\sigma^2 < 1$. This means that if the measurements are very bad, then L_2 will not change very much with u, and the cautious regulator u_1 may just as well be used. If $\sigma^2 < 1$ then (4.18) will be negative if $\hat{b}^2$ is small and P big enough. This means that when measurements are accurate enough, u_1 and u* may differ significantly when the estimate is close to zero compared to its uncertainty. To illustrate, the following numbers are given

$\hat{b} = 0.05$
$P = 0.5$
$R = 0.1$
$\sigma^2 = 0.1$

In fact these are the numbers used for fig. 10. They give

$u_1 \approx 0.1$ and $u^* \approx 0.65$

which is quite a difference.

Multiple minima

Numerical solutions to dual control problems in Bohlin (1969) and Åström and Wittenmark (1971) indicate that the control law may be a discontinuous function. The two-step regulator is indeed discontinuous. In the above example the minimizing u for L_1+L_2 will be discontinuous at $\hat{b} = 0$. Changing $\hat{b}$ will not influence L_2, but the minimum of L_1 will move. When decreasing $\hat{b}$ from a positive value the global minimum of L_1+L_2 will jump at $\hat{b} = 0$ from the right local minimum to the left one.

5. SIMULATIONS

In all the simulations the system output is generated by (2.1) with $\varphi(t)$ as in (2.3). The corresponding parameter vector $\theta(t)$ is generated as shown in (2.2) unless otherwise stated. The noise was generated by summing 12 sequential outputs from a mixed congruential generator with 131 072 states. The period then was 16 384. For each simulation the performance is measured as

$$V_s = \frac{1}{N - t_d + 1} \sum_{t=t_d}^{N} \left(y(t) - y_r(t)\right)^2 \tag{5.1}$$

When examining steady-state behaviour t_d was usually set to $t_d = 50$ in order to decrease the influence of initial values.

Most simulations were performed M times to get a measure of the spread of the performance function. The different runs were considered as independent, and the mean value and the standard deviation of the collection of M values of V_s were computed as

$$\overline{V_s} = \frac{1}{M} \sum_{i=1}^{M} V_s^i \tag{5.2}$$

$$\Sigma = \sqrt{\frac{1}{M-1} \sum_{i=1}^{M} (V_s^i - \overline{V_s})^2} \tag{5.3}$$

where V_s^i is the result of the i:th run. Then Σ tells how close to $\overline{V_s}$ the result of single runs will usually be. This value will, of course, depend on the length of the simulation.

To get a measure of the accuracy of $\overline{V_s}$ its standard deviation $\overline{\Sigma}$ is estimated as

$$\overline{\Sigma} = \frac{1}{\sqrt{M}} \Sigma \tag{5.4}$$

since the different runs are considered as independent.

Parameter estimation is an important part of the considered regulators. The same estimation algorithm is used for all regulators in order to make conditions as equal as possible. The estimation algorithm contains a model of the system, and the parameters of that model must be chosen in advance. In cases where the model structure is correct, all parameters were set to their true values. In some of the simulations, however, the true system was generated in another way. The parameters for estimation then had to be tried out by simulation. This tuning was done using the cautious regulator, and the same parameter values were then used for the other regulators.

For easy reference the regulators tested are listed and numbered in table 1.

Regulator	Parameters
1. Self-tuning regulator (3.5)	Input limit u_{lim}
2. Cautious regulator (3.4)	
3. Cautious, perturbed with extra input	
3a. according to (3.6)	amplitude u_e of extra input
3b. according to (3.7)	u_e
3c. according to (3.10)	u_e and uncertainty limit P_{lim}
3d. according to Mannerfelt (3.11)	λ in (3.11)
4. Wittenmark's regulator, $f(P) = \ell P \ell^T$	λ in (3.17)
5. Two-step regulator	

Table 1 - List of regulators

Note that regulators 2 and 5 have no parameters apart from those of the estimation algorithm. As shown in table 1 the other regulators have additional parameters that have to be chosen in advance. This may be a serious drawback, especially if no *a priori* knowledge about the optimal value is available or if the performance is sensitive to the parameters. In the

simulations the extra parameters were determined experimentally, and are well-tuned.

The minimization in Wittenmark's regulator was done numerically. A line search based on the fitting of a quadratic to 3 points was used. It was found that 5 iterations was normally sufficient.

Simulation 1: This example illustrates that for slow parameter variations, where the leading b-parameter keeps away from zero, the non-dual regulators will do quite well. The system is of first order with

$$a(t) = 0.9 + 0.5\sin 0.01t$$
$$b(t) = 1.5 - 0.9\sin 0.01t$$
$$\sigma^2 = 1$$
$$y_r(t) \equiv 0$$

Note that sometimes $a(t) > 1$, so that with no control the output would drift away. The optimal control for known parameters is $u(t) = \frac{a(t)}{b(t)} \cdot y(t)$. The closed-loop system then has a pole at the origin and disturbances will be eliminated in one step. The gain of the optimal regulator is shown in fig. 11. It varies between 0.17 and 2.33 with a period of $200\pi \approx 628$ steps. The gain variation is thus very slow compared to the dynamics of the optimal closed-loop system.

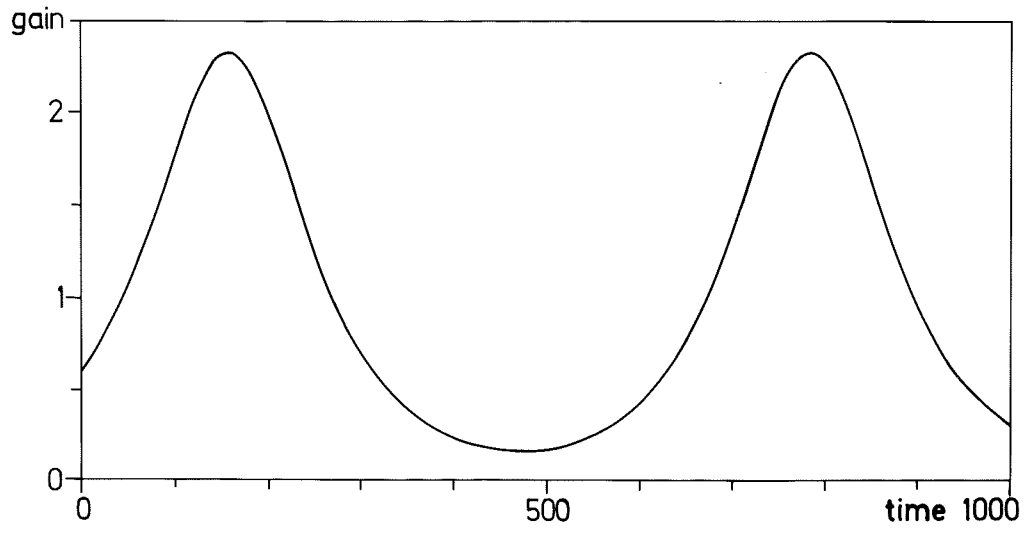


Fig. 11 - Gain of the optimal regulator for known parameters.

With sinusoidal a(t) and b(t) the parameters in the estimation algorithm have to be tried out by simulation. The covariance matrix for the parameter noise was first determined from the maximal derivative of the parameters. This is the maximal parameter change in one step, and could therefore be a reasonable value for the standard deviation. This would give

$R = \text{diag}[0.000\ 009 \quad 0.000\ 081]$

However, with sinusoidal parameters the changes are coordinated, and simulations showed that a larger R-value had to be used. The following values were finally chosen

$\Phi = I$ (unity matrix)
$R = 0.005 \cdot I$
$\sigma^2 = 1$
$\hat{\theta}(0)^T = [0.9 \quad 1]$
$P(0) = 0.1 \cdot I$

Regulators 1 (with no input limit), 2 and 5 were used, each in 1 run of 1000 steps. As expected all 3 regulators gave very similar results with the following loss per step

STURE	1.072
Cautious	1.067
Two-step	1.071

With known parameters the loss would be $V_s = 1.0$. To show the similarity the regulator gains in the interval [200,300] are given in fig. 12. The parameter estimates are also similar and are shown only for the two-step regulator in fig. 13 and 14. Notice that the gain of the regulators agree well with the optimal gain although the parameter estimates a and b vary significantly from the true values.

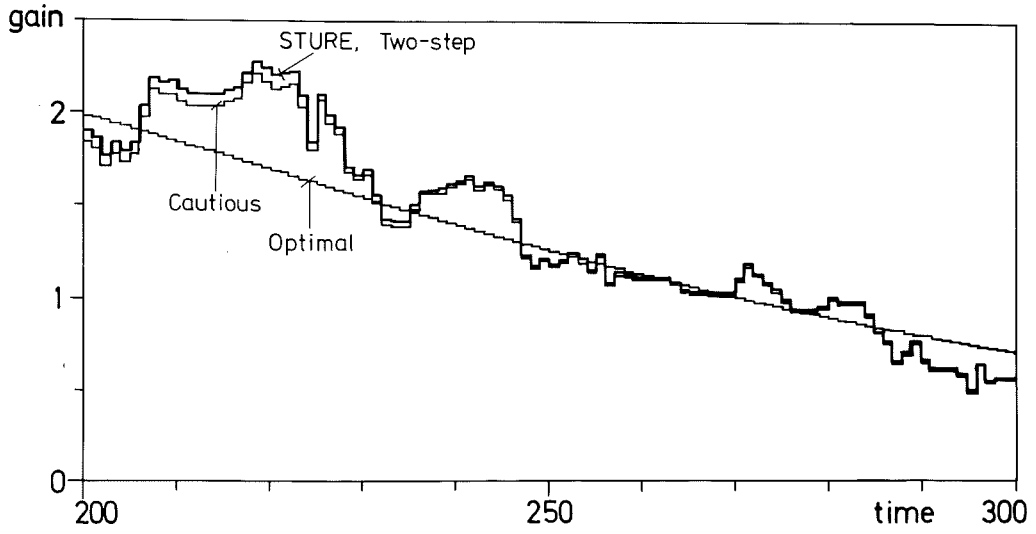


Fig. 12 - Regulator gains of STURE, the cautious and the two-step regulators and the optimal gain for known parameters.

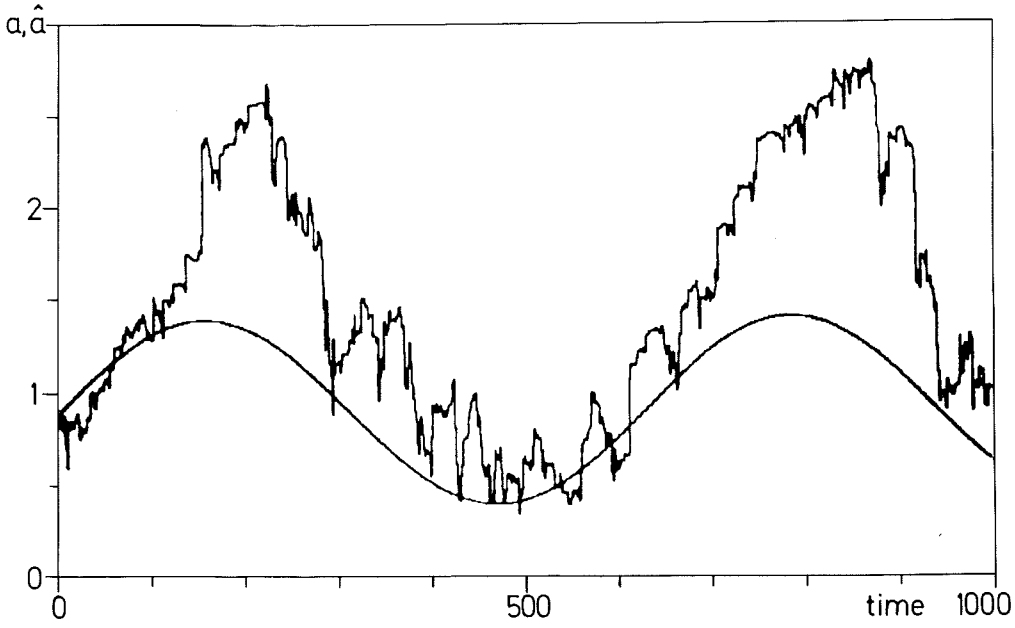


Fig. 13 - The a-parameter and its estimate.

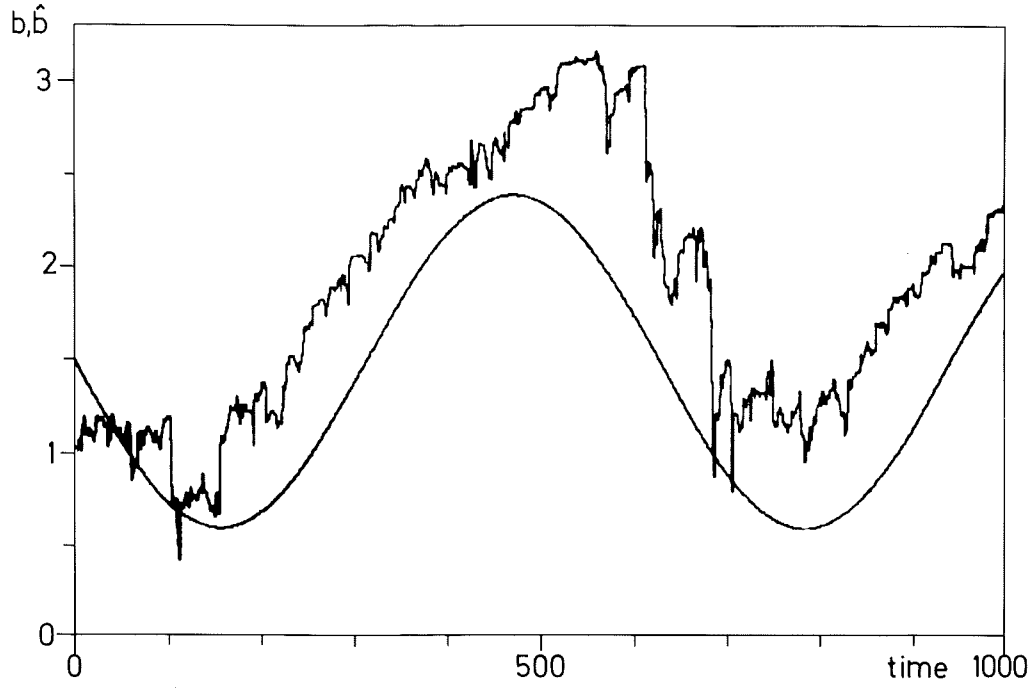


Fig. 14 - The b-parameter and its estimate.

It is obviously not necessary for the regulators to have accurate estimates of both parameters. In fact, a closer look at (3.4) and (3.5) reveals that in both cases with $y_r(t) \equiv 0$ the input depends on the quotient of the estimates, but not on the estimates themselves. Fig. 15 shows $\hat{a}(t)/\hat{b}(t)$ and $a(t)/b(t)$ for the two-step regulator.

It is possible to estimate the system parameters better by applying an extra input as is done with the cautious regulators 3. But this is not necessary for controlling the system, and will increase the loss. Figs. 16 and 17 show the estimates when using regulator 3a with $u_e = 0.5$. Estimation is better, at least in the second period, but the loss was $V_s = 1.731$.

With a non-zero reference value (3.4) and (3.5) show that both parameter estimates are needed and in simulations their accuracy was much higher than with $y_r(t) \equiv 0$.

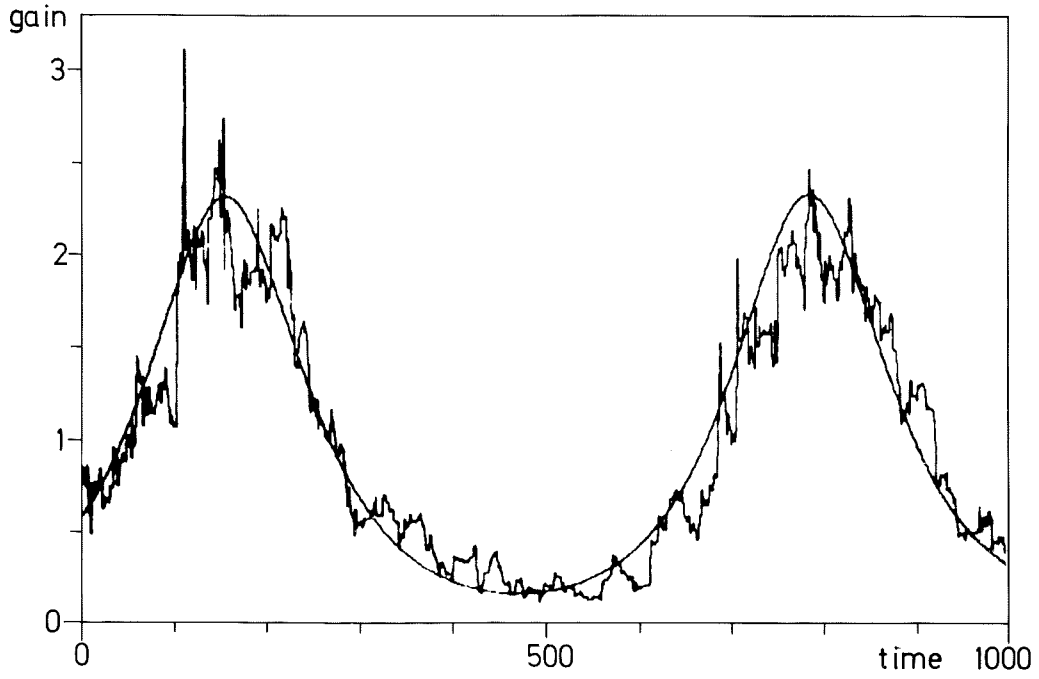


Fig. 15 - Gain of the two-step regulator and the optimal gain for known parameters.

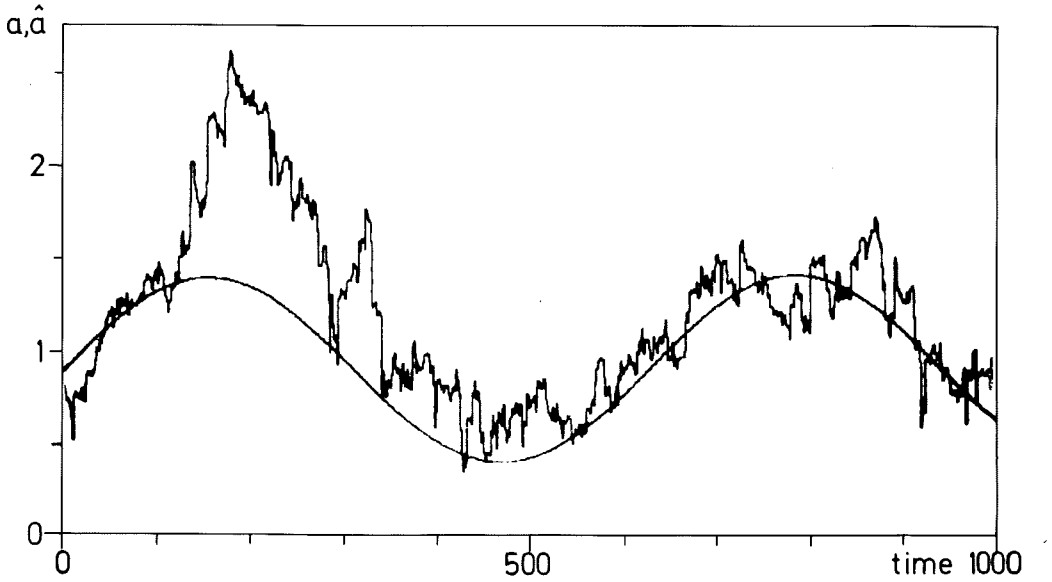


Fig. 16 - The a-parameter and its estimate when using an extra input.

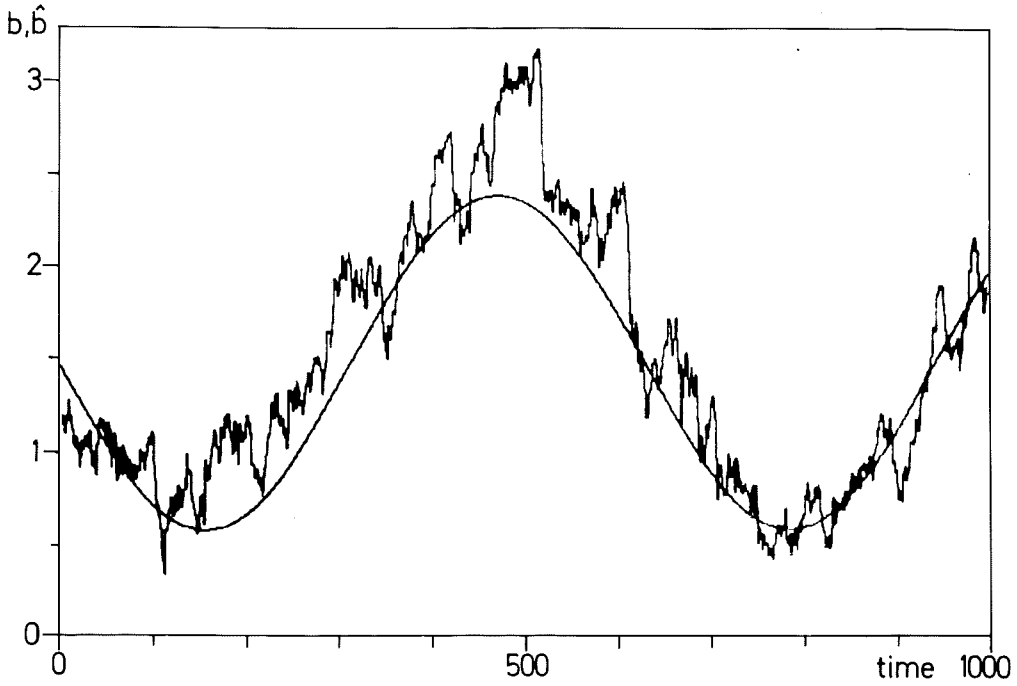


Fig. 17 - The b-parameter and its estimate when using an extra input.

In order to get a feeling for the limitations of the regulators the frequency of the parameter variation was increased. Two additional frequencies were tried, 0.1 and 0.3, which correspond to periods of T = 63 and T = 21 steps respectively. The result is given in table 2.

Regulator \ Parameter period (steps)	T=628	T=63	T=21
STURE	1.072	1.566	3.158
Cautious	1.067	2.193	2.911
Two-step	1.071	1.582	2.905

Table 2 - The loss for different speeds in parameter variations.

The large loss with the cautious regulator for T = 63 is due to short periods of turn-off. Apart from this all 3 regulators

still behave very similarly. The performance limit is set by the estimation algorithm, which cannot follow too rapid parameter variations.

Simulation 2: This example is more difficult than the previous one. The parameters are stochastic processes, and the variation in the gain is much larger. The system is of first order with the parameters generated as

$$a(t+1) = 0.98a(t) + v_1(t+1)$$

$$b(t) = 2 + x(t)^2$$

where

$$x(t+1) = x(t) + v_2(t+1)$$

and

$$\sigma^2 = 0.04$$
$$y_r(t) \equiv 0$$
$$R_v = 0.04 \cdot I$$

As in simulation 1 the optimal control with known parameters is $u(t) = \frac{a(t)}{b(t)} \cdot y(t)$. This gives a closed-loop system with a pole at the origin. The gain variations of the optimal regulator are much faster and larger than in the previous simulation. But the b-parameter does still not change sign, and is always greater than 2. One realization of the optimal gain is shown in fig. 18 and the corresponding parameters a(t) and b(t) are shown in fig. 19, where the two straight lines indicate ±1.

Note that the a-parameter is sometimes large with e.g. $a(275) \approx \approx -2$. With no control the output would then drift away very quickly. Towards the end of this realization the optimal gain is close to zero for a while. This is because b(t) increases drastically, with a maximum of 20.5. The gain must then be estimated accurately, since errors will be multiplied by a large b(t).

Fig. 18 - One realization of the optimal regulator gain for known parameters.

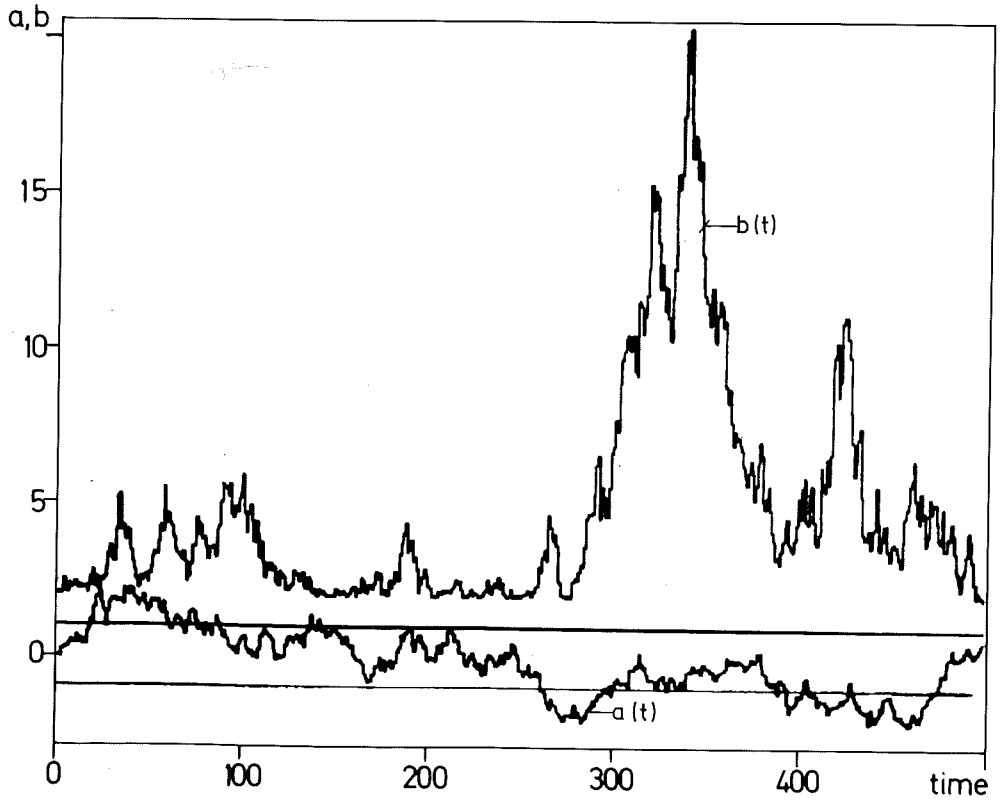


Fig. 19 - One realization of a(t) and b(t).

Most parameters in the estimation algorithm were given their correct values. The true b-parameter, however, does not fit in the model (2.2) which is used in the estimation. The model for the b-parameter was found empirically by simulations. The initial guess for the standard deviation of the b-parameter noise was based on the derivative of the b-parameter equation. Then $x(t) = 1$ was used as a common value for x. This gave

$$R_b = [2x(t)\sigma_x]^2 = 0.16$$

Simulations showed better results with a somewhat larger R_b, and the final choice was

$$\Phi = I$$
$$R = \text{diag}[0.04 \quad 0.4]$$
$$\sigma^2 = 0.04$$
$$\hat{\theta}(0)^T = [0.1 \quad 1]$$
$$P(0) = \text{diag}[10 \quad 1]$$

For each regulator 30 runs with $N = 500$ and $t_d = 50$ were performed. The result is shown in fig. 20, where $\bar{V}_s \pm \bar{\Sigma}$ is marked out for all cases.

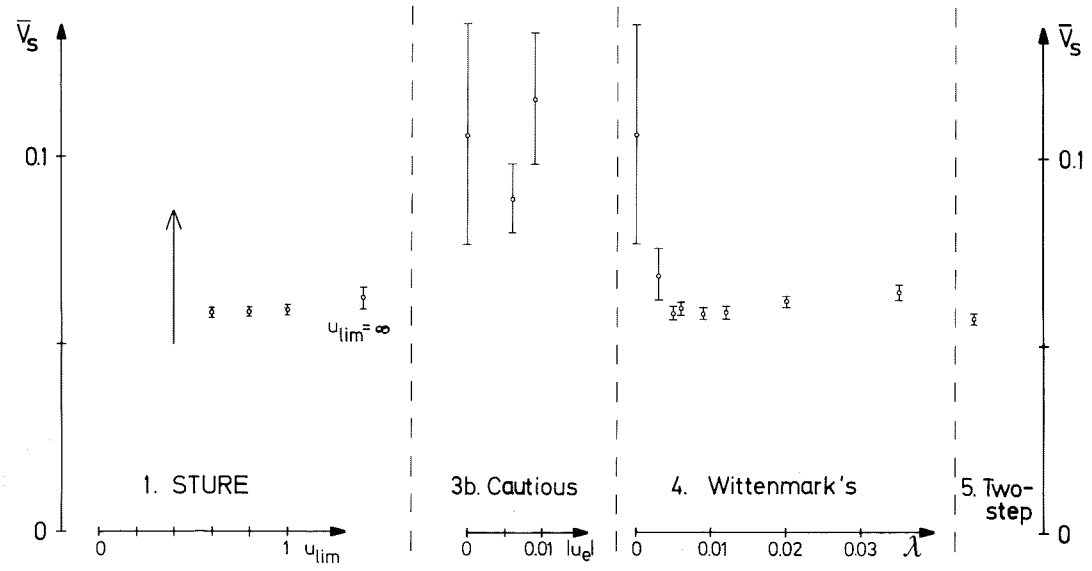


Fig. 20 - The loss with different regulator parameters.

The results for the best parameter choices are also listed in table 3.

Regulator	$\bar{V}_s \pm \bar{\Sigma}$	Parameter
1. STURE	0.0623±0.0028	$u_{lim} = \infty$
1. STURE	0.0582±0.0012	$u_{lim} = 0.6$
2. Cautious	0.106±0.030	- - -
3b. Perturbed cautious	0.089±0.009	$\|u_e\| = 0.006$
4. Wittenmark's	0.0586±0.0018	$\lambda = 0.005$
5. Two-step	0.0571±0.0012	- - -

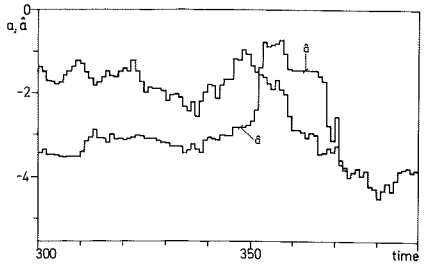
Table 3 - Best results in simulation 2.

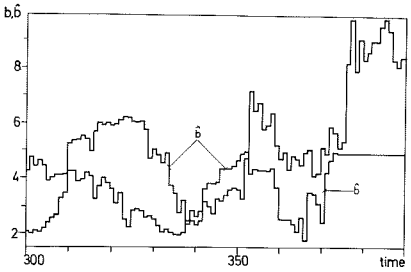
With the cautious regulator and no extra input a few cases of turn-off occurred. With $|u_e| = 0.003$ one of the runs gave V_s = = 6.449. The reason was a short period with negative $\hat{b}_1$, which by chance was fatal to that particular run. The mean loss for the perturbed cautious regulator with $|u_e| = 0.003$ is then very inaccurate with $\bar{V}_s = 0.28 \pm 0.21$.

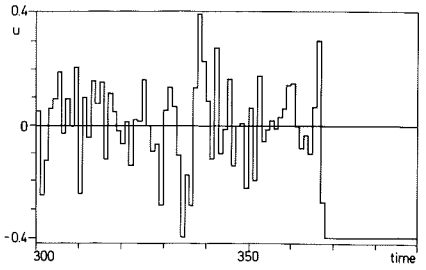
The arrow in fig. 20 at $u_{lim} = 0.4$ for STURE indicates that one simulation gave an unstable closed-loop system. This may happen when the input is limited, since a large input is required when the system is unstable. Figs. 21 - 24 show what happened.

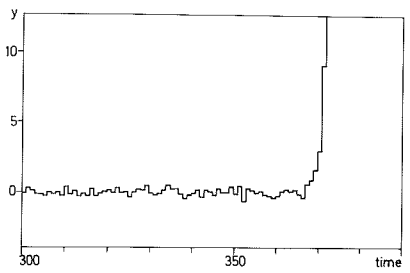
Just before t = 370 the a-parameter drifts down to $a \approx -4$. This demands a large negative input. With a limit of -0.4, this cannot be generated and the output drifts away in just a few steps. Note also that the large output makes $\hat{a}$ very accurate.

With the large a-parameter variations of this example it is thus difficult to find a suitable input limit. Moreover, the decrease in the loss will be fairly small. In a practical problem the a-parameter usually varies less. It is then often possible to find quite reasonable bounds for the control signal.









Figs. 21-24 - An unstable realization with STURE.

Fig. 21 - the a-parameter and its estimate
Fig. 22 - the b-parameter and its estimate
Fig. 23 - the input
Fig. 24 - the output

To summarize this simulation, the two-step regulator, STURE with no input limit and Wittenmark's regulator give approximately the same performance. But STURE is the simplest one, and could therefore be preferred.

Simulation 3: In order to find the limitations of the self-tuning regulator, the b-parameter is now moved closer to zero, and the variation is speeded up a little. The first order system is generated with the same a-parameter, equation noise and reference value as in the previous simulation. The b-parameter is now

$$b(t) = 0.5 + x(t)^2$$

where

$$x(t+1) = 0.999 \cdot x(t) + v_2(t+1)$$

with

$$E(v_2(t))^2 = 0.09$$

The gain of the optimal regulator for known parameters may now change considerably in just one step. One realization is shown in fig. 25 with the corresponding b-parameter in fig. 26. The gain changes are about as fast as the dynamics of the optimal closed-loop system, which has a pole at the origin.

The parameters in the estimation algorithm were chosen in the same way as in simulation 2. Intuitively, R_{bb} should be larger than before, since the b-parameter noise is larger. It turned out, however, that with 0.5 as the lower limit for b(t), it is more important to keep the estimates "calm" in order to avoid sign errors. The same value as before was therefore chosen.

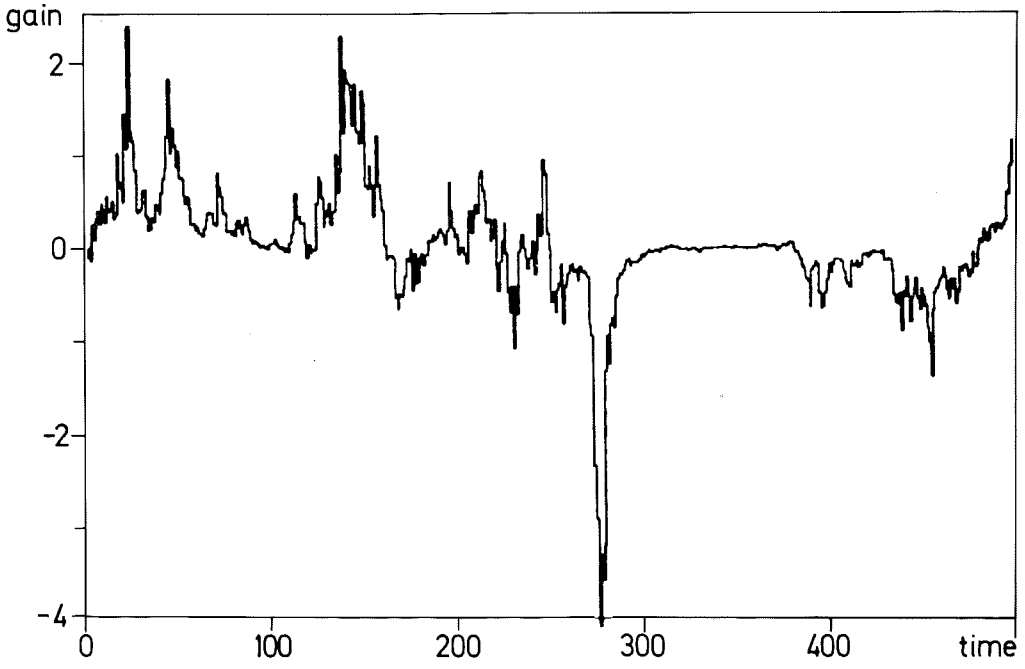


Fig. 25 - One realization of the optimal regulator gain for known parameters.

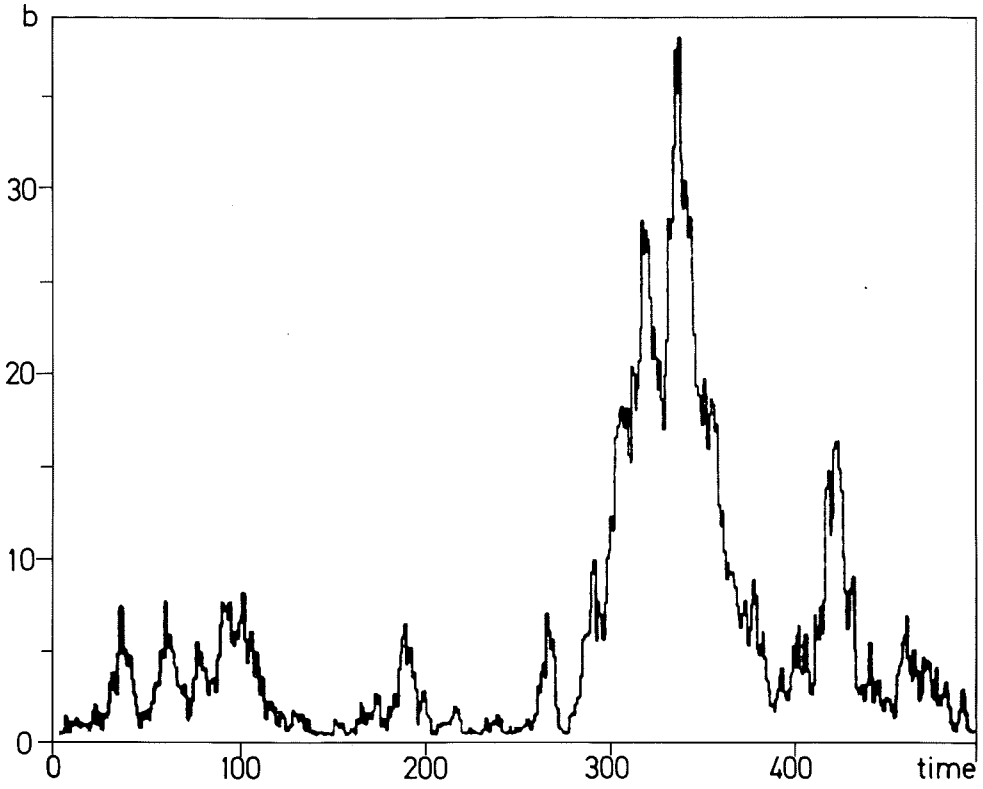


Fig. 26 - One realization of the b-parameter.

Again 30 runs with N = 500 and t_d = 50 were performed for each regulator. The gain changes are very large and the a-parameter is sometimes greater than 1. This makes control very difficult, and the result from just 500 steps will depend very much on the actual noise sequences. The mean loss is then not an appropriate measure. Instead the results from each run is marked out by a cross in fig. 27, which indicates the probability distribution of the loss function.

Fig. 27 clearly shows the difficulties in using mean values only for problems of this type. The two-step regulator gives the most concentrated distribution. It has only one run with $V_s > 1$, and this maximum value is only ~4.12.

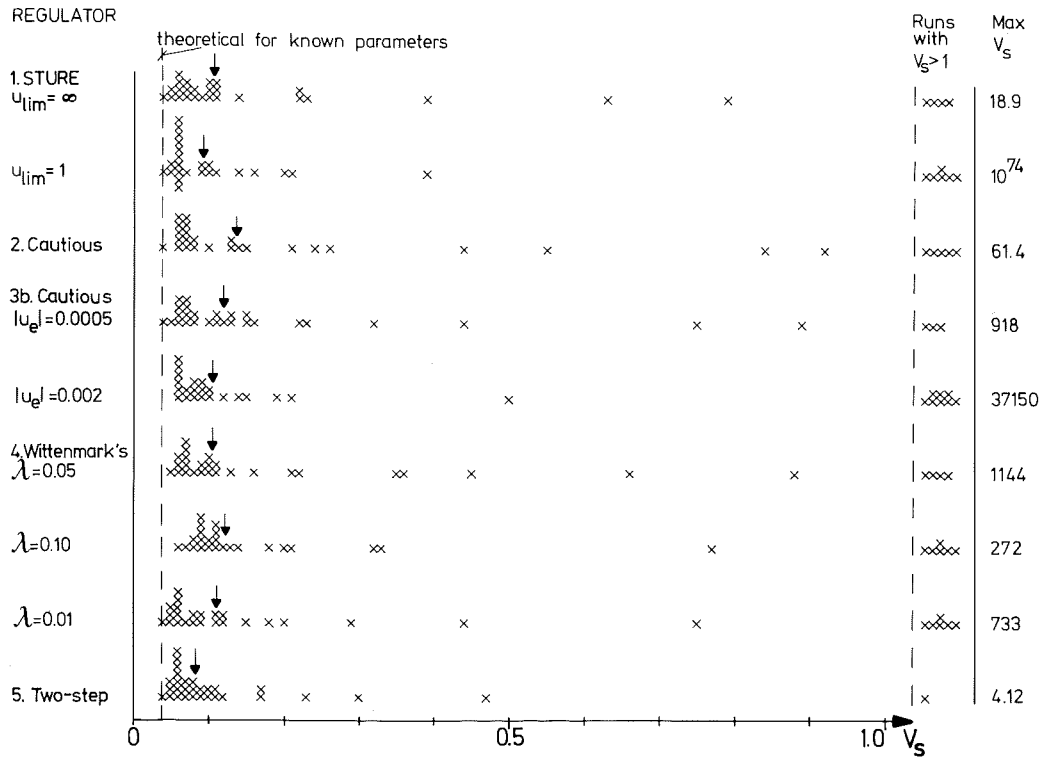


Fig. 27 - The distribution of the loss function. The arrows denote median values.

The minimal value of the loss for known parameters is 0.04. The self-tuning regulator can manage most of the realizations. This regulator has, however, difficulties when $\hat{b}(t)$ changes sign. The two-step regulator is safer in such cases, and also gives a lower median value, 0.08 compared to 0.10 for STURE.

The large loss (10^{74}) for STURE with $u_{lim} = 1$ again shows the danger of limiting the input. In 6 runs this caused instability. The loss then grew infinitely. Occasional big losses for all other regulators arose during comparatively short periods of bad control.

Large inputs can be generated by STURE in two different ways. One is when $\hat{b}$, the denominator of (3.5), is close to zero. If the true b-parameter is not small this effect is not wanted. The other case is when the open-loop system is unstable, so that big inputs are needed for stabilization. In this case a limit on the input may be fatal, as shown in the simulation above.

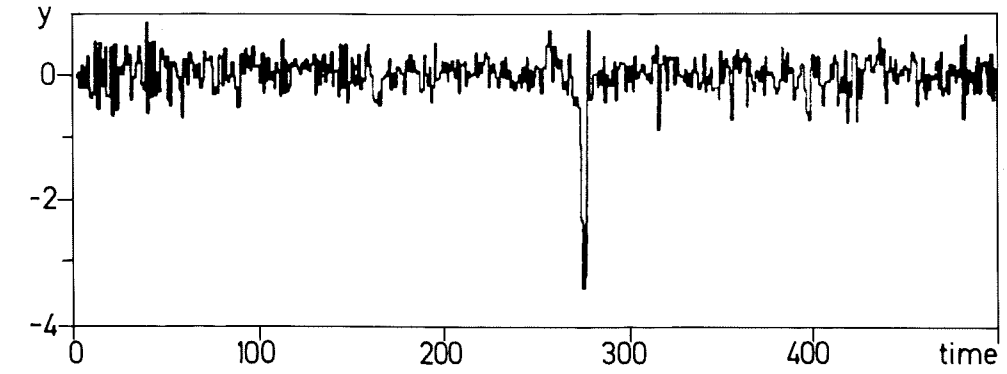
Instead of limiting the input, $|\hat{b}|$ could be kept away from zero by force. The size of the forbidden region could then be determined from the variance of the b-parameter estimate. Another alternative is to limit the regulator gain. This might be more natural in practical problems. These two methods have, however, not been tested in this report.

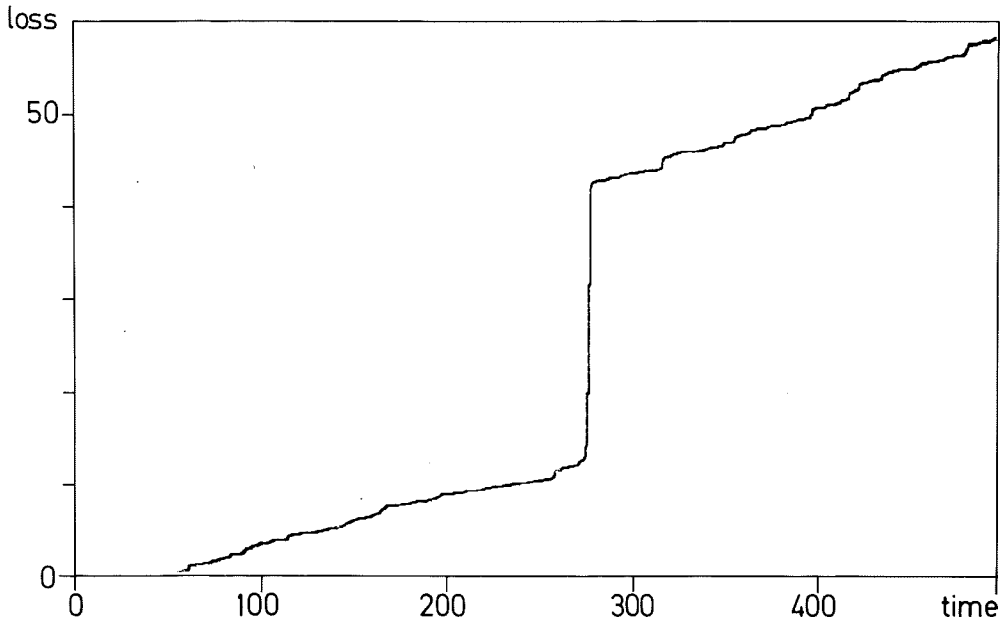
For this difficult system, the large losses in some of the runs are usually generated in just a few steps. An example is shown in figs. 28-29, which is the output and the loss in a run with Wittenmark's regulator and $\lambda = 0.01$. A major part of the loss is generated at $t \approx 275$.

This explains the wide-spread results from different runs. Difficult periods occur just now and then. If such periods are completely absent in one run, the loss will be very small and otherwise big. The only way to get less spread out results is to use longer runs, so that all runs contain some difficult periods.

On the other hand, the criterion is maybe not very well-chosen for this kind of problems. It will focus the efforts of the controller totally to the short difficult periods. This indicates that what is needed is a detailed study of what happens over a few sampling intervals when the output is large.

In this simulation the two-step regulator is uniformly better than all the others. The need for extra computing efforts is well compensated for by improved performance.





Figs. 28 - 29 - The output and the loss in one realization with Wittenmark's regulator.

Simulation 4: Sofar, the b-parameter has not been allowed to change sign. The next simulation is taken from the master thesis by Mannerfelt (1977), where an integrator with a zero mean stochastic process as gain is studied. The system is

$$y(t) - y(t-1) = b(t)u(t-1) + e(t)$$

The parameter variations are described by

$$b(t+1) = 0.9 \cdot b(t) + v(t+1)$$

and

$$E e(t)^2 = 0.25$$

$$E v(t)^2 = 1$$

$$y_r(t) \equiv 1$$

$$b(0) = 0$$

The b-parameter model implies frequent sign changes. A typical realization of b(t) is shown in fig. 30.

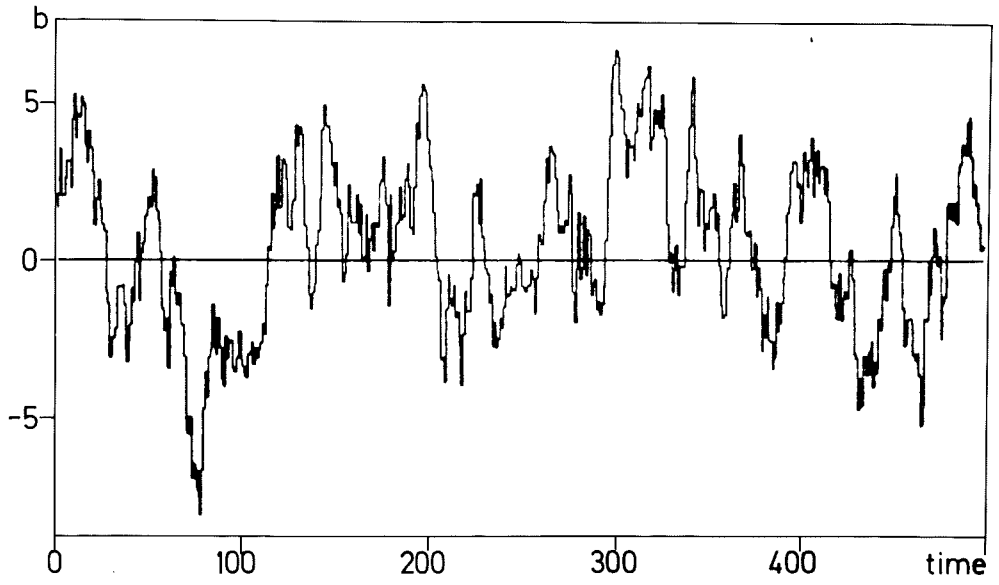


Fig. 30 - One realization of the b-parameter.

When b(t) is zero the system is not controllable, and the output variance will grow linearly. But also with b(t) small, but non-zero, estimation errors will make control difficult, and the input may easily get the wrong sign. This would give an unstable closed-loop system if b(t) were constant. Particularly difficult situations occur when the true b-parameter stays close to zero for a while as e.g. around t = 285 in fig. 30.

It can be expected that the self-tuning regulator will have difficulties to control the system. With a small b-parameter estimate in the denominator, this input is very sensitive to estimation errors.

In the estimation the a-coefficient -1 was assumed to be known. Only b(t) was estimated. At first all parameters in the estimation algorithm were set to their true values with

$\hat{b}(0) = 0.1$

$P(0) = 10$

Then also $\Phi_{bb} = 1$ was tried instead of the true value $\Phi_{bb} = 0.9$. For each regulator 50 runs of 500 steps were recorded with $t_d = 100$. The additional parameters of regulators 1, 3 and 4 were well-tuned using simulations. The best result for each regulator is shown in table 4.

Regulator	$\bar{V}_s \pm \bar{\Sigma}$ $\Phi_{bb} = 0.9$	Parameter	$\bar{V}_s \pm \bar{\Sigma}$ $\Phi_{bb} = 1$	Parameter
1. STURE	1.160±0.043	$u_{lim} = 0.7$	1.110±0.040	$u_{lim} = 0.65$
2. Cautious	1.680±0.040	none	1.605±0.043	none
3a. Perturbed cautious	1.095±0.023	$\lvert u_e \rvert = 0.15$	1.084±0.023	$\lvert u_e \rvert = 0.15$
3b. -"-	0.995±0.021	$\lvert u_e \rvert = 0.14$	0.965±0.019	$\lvert u_e \rvert = 0.16$
3c. -"-	1.025±0.022	$\lvert u_e \rvert = 0.16$, $P_{lim} = 2.0$	1.012±0.024	$\lvert u_e \rvert = 0.17$, $P_{lim} = 2.0$
3d. -"-	0.855±0.028	$\lambda = 0.1$	0.860±0.027	$\lambda = 0.15$
4. Wittenmark's	0.905±0.025	$\lambda = 0.3$	0.905±0.025	$\lambda = 0.2$
5. Two-step	0.915±0.036	none	0.950±0.030	none

Table 4: Results from integrator example.

The best performance is obtained with regulators 3d and 4. But to get these low losses the λ-value had to be well adjusted. Regulator 3d with $\Phi_{bb} = 0.9$ and $\lambda = 0.3$ gave e.g. $\bar{V}_s = 1.01$. Also for regulators 1, 3a and 3b $\bar{V}_s$ has a distinct minimum at the indicated parameter value. In a practical situation it would probably not be possible to tune the parameters optimally. The two-step regulator is then a good alternative without any extra parameters.

The self-tuning regulator performs surprisingly well compared to the other regulators of table 4, despite the fact that the b-parameter changes sign. It is, however, necessary to limit the input, and furthermore the loss increases considerably when u_{lim} is moved away from its optimal value. With $u_{lim} = 2.22$ and $\Phi_{bb} = 0.9$ the loss is e.g. $\bar{V}_s = 2.46$. Limiting the input is not too dangerous in this case, since the system is just on the stability limit and not exponentially unstable. This is why it is possible to make STURE work at all in this example.

Turn-off is probably a major reason for the big loss with the cautious regulator. With an extra input the loss is decreased considerably. Unexpectedly regulator 3c showed no improvement to regulator 3b. It seems to be more important to give the extra input a correct sign than to remove it when it is not needed (cf. regulators 3a, b and c).

It is noteworthy that the results are in most cases better with $\Phi = I$ than with the true value. This is good, because for real systems Φ would not be known, and $\Phi = I$ is then a natural choice.

Some runs were also made to test the sensitivity of the results to the R-value in the regulator. A factor of 10 up or down approximately doubled the loss for regulators 2, 3c, 4 and 5.

In this example the difference in performance between the regulators is thus small if they are all well-tuned. The two-step regulator, however, gives no additional parameters to tune.

Simulation 5: With two integrators instead of one the differences between the regulators ought to be more evident. This is shown in this simulation. The system is

$$y(t) - 2y(t-1) + y(t-2) = b(t)u(t-1) + e(t)$$

with

$$b(t+1) = 0.95 \cdot b(t) + v(t+1)$$

and

$$Ee(t)^2 = 0.0009$$

$$Ev(t)^2 = 0.09$$

$$y_r(t) \equiv 1$$

$$b(0) = 0.5$$

The b-parameter variations are similar to those of simulation 4, and there are the same problems due to sign changes. In this simulation, however, the effects are amplified by the second integrator. Fig. 31 below shows one realization of the b-parameter.

Again the a-coefficients are assumed known, and all parameters for the estimation algorithm were set to their true values with

$$\hat{b}(0) = 0.5$$

$$P(0) = 10$$

As in simulation 3 short periods of bad control may give dominating contributions to the total loss. Such periods will occur when the true b-parameter varies around zero. Then even an optimal control law would have trouble, since the estimate of $b(t+1)$ cannot be more accurate than the noise added to $b(t)$ allows. With $b(t) = 0$ the system is not controllable and the output will drift away rather quickly.

Figs. 31 - 36 show the b-parameter, the estimation error and covariance, the input, the output and the loss for one run with the two-step regulator.

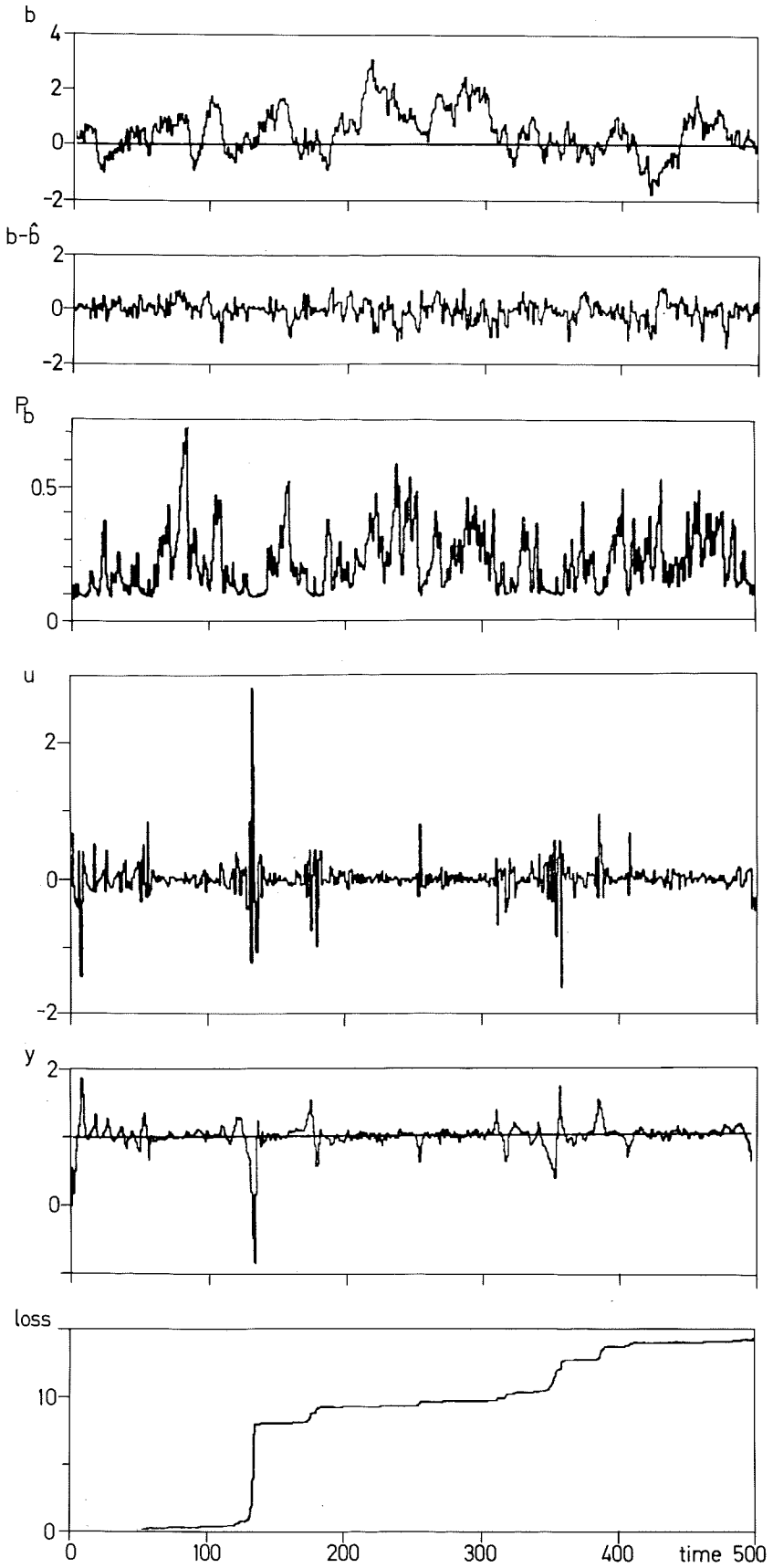
In order to be able to use the quadratic loss function to compare different regulators very long runs were made. The probability is then small for a run not to contain any difficult periods, and the results should be less wide-spread. Due to the limited period of the random number generator the number of runs was cut down to 25 with 5000 steps in each and $t_d = 50$. Fig. 37 shows $\bar{V}_s \pm \bar{\Sigma}$ for some choices of regulator parameters. Notice that the result for regulator 2 (cautious) is the point $\lambda = 0$ or $|u_e| = 0$ in regulators 3a, 3b, 3d and 4.

For this system the best modification of the cautious regulator is regulator 3b. It is less sensitive to the value of $|u_e|$, and also gives a smaller minimal loss.

The two regulators 4 and 5 give in this case a significantly smaller loss, which could compensate for their larger computing requirements. The advantage with the two-step regulator is, of course, the absence of an extra parameter to tune such as λ. No results for the self-tuning regulator have been given. This is because it was impossible to make it work satisfactorily. The best result in one run for any value of u_{lim} was $V_s = 0.4798$, and more than half of the runs gave $V_s > 1$.

The result for the best parameter value of each regulator is listed in table 5. For further illustration these results are also given in fig. 38, where 2 cases with N = 500 are also shown. Each run is marked out by a cross as in fig. 27. The arrows denote the median (upper arrow) and the mean (lower arrow).

With longer runs the relative distance between the median and the mean or between the maximum value and the mean is smaller. This is most evident for regulators 4 and 5. The results for the other regulators are still quite spread out.



Figs. 31 - 36 - One realization of the double integrator with the two-step regulator. From above: the b-parameter, the estimation error, the estimate variance, the input, the output and the loss.

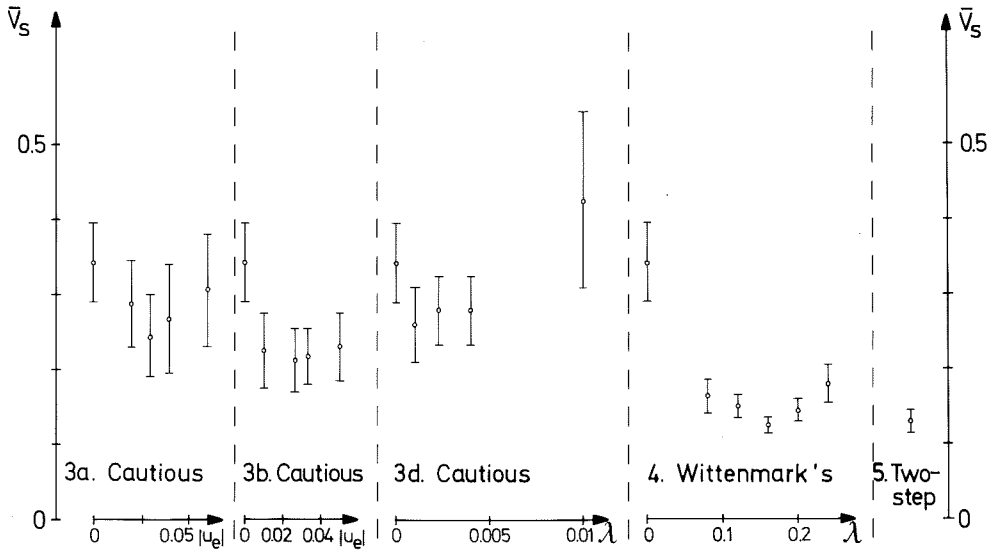


Fig. 37 - The loss with different regulator parameters.

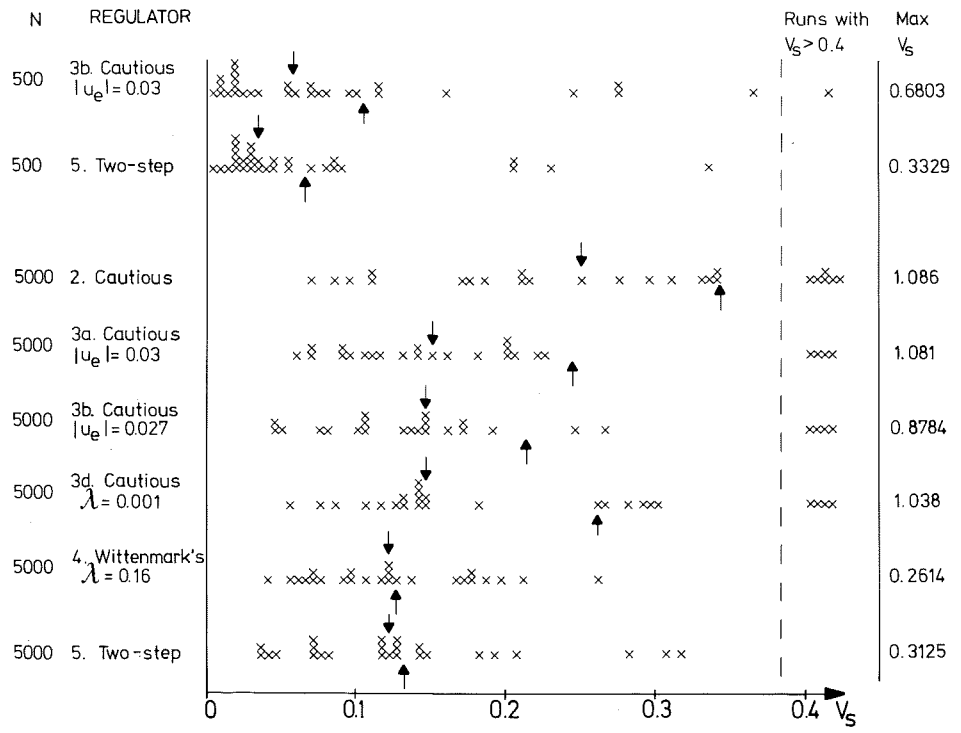


Fig. 38 - The distribution of the loss function.

Regulator	$\bar{V}_s \pm \bar{\Sigma}$	Parameter
1. STURE	11.5±7.2	$u_{lim} = 1$
2. Cautious	0.343±0.054	none
3a.Perturbed cautious	0.244±0.055	$\|u_e\| = 0.03$
3b. -"- -"-	0.213±0.043	$\|u_e\| = 0.027$
3d. -"- -"-	0.261±0.049	$\lambda = 0.001$
4. Wittenmark's	0.124±0.011	$\lambda = 0.16$
5. Two-step	0.130±0.016	none

Table 5 - Results for double integrator.

It appears in fig. 38 as if the mean loss per step is lower with N = 500 than with N = 5000. This is, however, a mere coincidence. The 30 runs with 500 steps should give the same amount of information as only 3 runs with N = 5000. By chance these "3 runs" may well give a much too low mean loss. This fact was demonstrated when 51 more simulations with the two-step regulator and N = 500 were performed. In the last but one the loss per step was $V_s = 3.03$. This run alone raised the mean loss from $\bar{V}_s = 0.084$ to $\bar{V}_s = 0.12$.

In problems of this kind it is thus dangerous to compare mean values only, not to mention single runs. It is necessary to get to know the whole probability distribution of the loss function.

In this simulation there are large differences in performance between the regulators. It is not possible to find any suitable input limit that will make STURE work satisfactorily. An extra input will improve the performance of the cautious regulator only to a certain extent. This clearly motivates the use of Wittenmark's regulator or the two-step regulator since they may improve the performance further by approximately a factor of 2. In this example the result is sensitive to variations in the extra regulator parameters. For this reason the two-step regulator is the best choice.

Simulation 6: All previous simulations have illustrated steady-state performance. The behaviour in the first 10 steps will now be studied. The system is a double integrator with unknown but constant gain

$$y(t) - 2y(t-1) + y(t-2) = bu(t-1) + e(t)$$

with

$$Ee(t)^2 = 0.01$$

$$y_r(t) \equiv 0$$

$$y(-1) = y(-2) = 0$$

The purpose of this simulation is to study the cost of learning the b-parameter. This is done in 50 runs for each regulator with $N = 10$ and $t_d = 1$. For a certain value of the initial estimate $\hat{b}(0)$ and covariance $P(0)$ the result will, of course, depend on the true b-parameter value. This effect is averaged out by taking the true b-parameter for each run from a probability distribution. Previous simulations have shown that sign errors in the estimate for b will give problems. To decrease the risk for sign errors the true b-parameter is always made positive by taking

$$b = \begin{cases} x & \text{if } x \geq 20 \\ 10[1 + \exp(0.1x-2)] & \text{if } x < 20 \end{cases}$$

where x is gaussian with mean and standard deviation both 50.

In the estimation algorithm the a-coefficients are assumed known. The b-coefficient is known to be constant but with an initial uncertainty. All parameters in the estimation algorithm were set to their correct values, disregarding the limit on b.

This means

$\Phi = I$

$R = 0$

$\sigma^2 = 0.01$

$\hat{b}(0) = 50$

$P_b(0) = 2500$

The results from this simulation are given in the first column of table 6. The probability distribution of the loss function has long tails in this example too, so that just a few of the runs may give dominating contributions to the mean loss. It turned out that in all such runs, the true b-parameter was close to 10, which is the least possible value. The explanation is that due to large initial uncertainty, there is a risk that the estimate of b will go too far when decreasing from $\hat{b}(0) = 50$ towards the true b-value. The estimate may even get the wrong sign.

This situation was tested separately by making 50 new runs with

$b = 5$

$\hat{b}(0) = 50$

and all other parameters as before. For comparison 50 runs with

$b = 50$

$\hat{b}(0) = 5$

were also made. These two cases are shown as cases b and c of table 6.

A comparison of cases b and c indicates that it is safer to underestimate the b-parameter than to overestimate it. The only exception is the cautious regulator with a well-tuned extra input, where the two minimal loss values are approximately equal. The true b-parameter is, however, different in the two cases, which may influence the result.

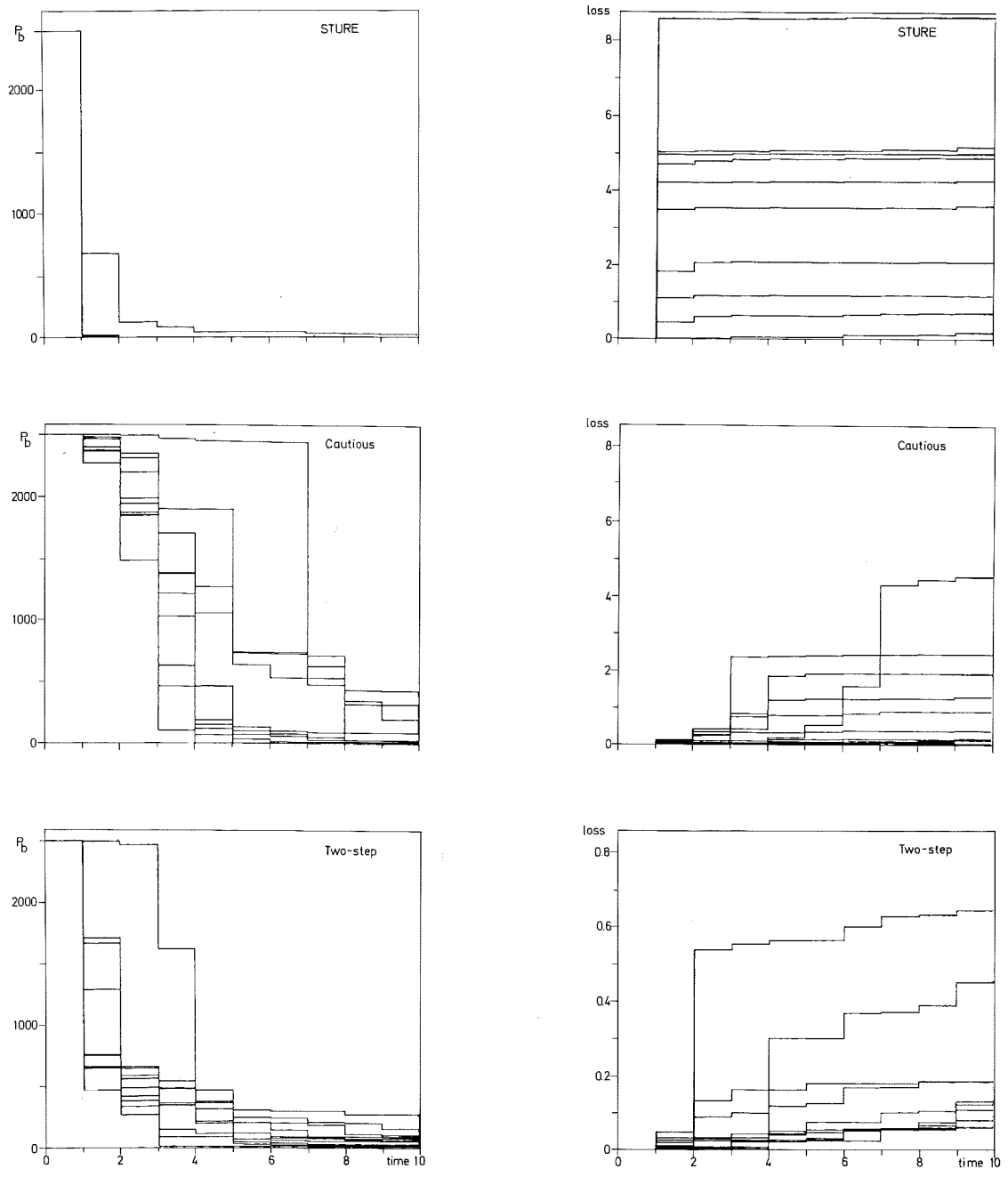
Regulator	Parameter	Case a $\hat{b}(0) = 50$ b varying	Case b $\hat{b}(0) = 50$ b = 5	Case c $\hat{b}(0) = 5$ b = 50
1. STURE	$u_{lim}=\infty$	3.6±3.5	2.8±2.4	0.331±0.060
	=0.005	0.166±0.071	1.45±0.29	0.025±0.004
	=0.01	0.073±0.029	1.15±0.26	0.037±0.005
	=0.02	0.032±0.009	0.47±0.13	0.112±0.012
	=0.05	0.031±0.009	0.111±0.034	0.344±0.052
	=0.08	0.038±0.013	0.090±0.025	0.419±0.090
	=0.15	0.062±0.031	0.091±0.027	0.335±0.062
2. Cautious	-	0.034±0.008	0.197±0.038	0.120±0.029
3b. Perturbed cautious	u_e=0.0001	0.034±0.008	0.175±0.028	0.068±0.011
	=0.0005	0.028±0.005	0.141±0.020	0.033±0.006
	=0.001	0.028±0.004	0.115±0.016	0.024±0.004
	=0.005	0.140±0.023	0.042±0.006	0.068±0.002
	=0.009	0.423±0.074	0.031±0.005	0.199±0.003
	=0.02	-	0.023±0.002	0.975±0.008
	=0.04	-	0.045±0.002	3.933±0.016
4. Wittenmark's	λ =0.00002	0.025±0.005	0.125±0.017	0.015±0.002
	=0.00005	0.022±0.004	0.109±0.017	0.014±0.001
	=0.0002	0.024±0.003	0.062±0.008	0.016±0.001
	=0.0005	0.028±0.003	0.047±0.007	0.018±0.001
	=0.001	0.032±0.004	0.039±0.006	0.021±0.001
	=0.002	0.040±0.005	0.038±0.007	0.027±0.002
	=0.008	0.071±0.011	0.025±0.005	0.045±0.002
	=0.02	0.111±0.018	0.024±0.004	0.069±0.003
	=0.10	0.235±0.039	0.022±0.002	0.152±0.005
	=0.25	0.366±0.064	0.026±0.002	0.237±0.006
5. Two-step	-	0.028±0.006	0.090±0.010	0.015±0.002

Table 6 - Results from 10-step simulations. The best values for each regulator are underlined.

It is obviously necessary to limit the input if the self-tuning regulator is used. The optimal limit-value will depend on the true b-parameter because the amplitude of the optimal input with known parameters depends on b. The problem associated with overestimating the gain (case b) was discussed above. When STURE is used underestimation (case c) may also cause difficulties. The inputs generated will then be too big in the first few steps, since the denominator of the input equation is $\hat{b}$. This will, however, speed up estimation, so that control is near optimal very soon. But the first few steps of large inputs cause the loss to be higher than with any other regulator. As shown in table 6 this effect is overcome by limiting the input ($u_{lim} = 0.005$). In practice it may be difficult to find an appropriate limit to use in just 10 steps.

In case a the cautious regulator works well even without any extra input. Using this facility the loss may be decreased by less than 20%. In cases b and c with large initial estimation errors more is gained. Without the extra input the regulator is overly cautious, and parameter estimation is too slow. But a well-tuned extra input makes this regulator far better than STURE in case b and about as good in case c. A comparison of the speed of parameter estimation is given for case c in figs. 39 - 44, where the loss and the variance of $\hat{b}$ in 10 runs are shown for STURE, the unperturbed cautious and the two-step regulators.

With STURE the variance is very small after just one step, but the price for this is a large initial loss. The cautious regulator decreases the variance too slowly, so that the estimate may drift around too much. It can even become negative. The output is then big, and produces a jump in the loss, but also a much decreased variance. With the two-step regulator the loss increase is about the same in each step with just two exceptions. Note also the different scales of fig. 44 from those of figs. 40 and 42. With the two-step regulator the speed of parameter estimation is thus well adjusted to give a small total loss.



Figs. 39 - 44 - The variance of $\hat{b}$ and the loss in 10 realizations with STURE (39-40)
unperturbed cautious regulator (41-42)
two-step regulator (43-44)

(Notice the scales!)

Wittenmark's regulator is the best one in all 3 cases and its performance is not very sensitive to the value of λ. The two-step regulator is almost as good in case c, but cannot handle case b very well. Thus if the gain is not overestimated the two-step algorithm is again a good choice, since it introduces no extra parameters. If possible to do, it may, however, be worth while to try out a good λ-value for Wittenmark's regulator.

Comparing the two-step and cautious regulators shows that in cases b and c of this example a great deal is gained by looking two steps ahead instead of one. But the large loss for the two-step regulator in case b indicates that two steps are also insufficient.

Summing up this simulation the cautious regulator is found to be better than STURE. The lowest loss is obtained with Wittenmark's regulator, but if the b-parameter is not overestimated the two-step regulator also works well. In practical problems with just a few steps to go it is hardly possible to tune any parameters. The two-step regulator is then better than all the others, since they are in this example sensitive to the parameter values.

6. CONCLUSIONS

There are three main contributions of this report. The first one is the attempt in chapter 3 to present current ideas on dual control in a systematic manner. Next a new algorithm was given, which has been shown to work in some simulated examples. In the simulations chapter finally, examples were given to show when a suboptimal dual control law is motivated and in which cases simpler adaptive schemes will do.

The main problem with the self-tuning regulator is that small values for the estimate of the leading b-parameter will produce too large inputs. This is most apparent when the true b-parameter may change sign or is frequently close to zero. The cure used in this report was to limit the control signal. This may be difficult if the system is unstable. It may also be difficult to find suitable limits in practice, e.g. if there are only a few steps to go.

In many cases the cautious regulator works very well as long as turn-off does not occur. When this happens control may be very bad. An extra input will then improve the performance. In some cases, however, it shows up that this regulator is myopic.

In the most difficult examples of this report the two-step and the Wittenmark regulators can manage much better than the others. This will then compensate for their larger computing requirements. It should be emphasized that the two-step regulator introduces no extra parameters to tune. This is of great importance in practice, since no *a priori* knowledge about the optimal value of these parameters is usually available. They may also be difficult to tune well enough in a practical application.

Throughout the simulations the estimation algorithm has been supplied with an accurate system model, in some cases even the true one. This is unrealistic in practical situations,

and it remains to be examined what are the effects of errors in the model. The noise characteristics are, of course, not known in general. In many cases the model order and the model structure may not be known. This makes it even more important to keep down the number of parameters to choose in advance. It has been suggested to estimate also the noise characteristics on line. This would reduce the amount of *a priori* information needed, but the effect when combining this with e.g. the two-step regulator has not yet been studied.

In this report only independent equation noise has been treated. It is not clear how the results should best be extended to systems with c-parameters. This is a subject for future research. Moreover, systems with a longer time delay than one sampling period has not been considered. This is also a problem to consider in the future. Nonlinear systems may often be rewritten as time-varying linear systems. It would therefore be interesting to test different suboptimal dual schemes on nonlinear systems.

As for improvements of the two-step regulator the approximation of future loss could be made more accurate. Two terms in the serial expansion of future loss could e.g. be used instead of one. It may, however, be more rewarding to find approximations that will make it possible to calculate multi-step regulators. Such approximations must then preserve their structure when taken through one more step of a dynamic programming procedure that includes taking expectation and minimizing w.r.t. the input.

In some simulations the distribution of the loss function was found to have long tails. This raised the question whether the quadratic criterion is suitable or not for this kind of problems. Future research in dual control could maybe produce better criteria and also better system models. This is probably what is needed also to get a deeper understanding of the nature of dual control. There is still a lot to be done in this area.

APPENDIX A

A 3-step Example

Let the system be described by (2.1) with only $a_1(t)$ as in (2.2) and the other parameters time-varying but known. The only non-zero covariance is then the covariance for $a_1(t)$ which will be denoted by $P(t)$ in this appendix. With zero reference value (3.3) is

$$V_N = \sigma^2 + y(N-1)^2 P(N) \tag{A.1}$$

The last two terms of (3.3) cancel since $b_1(t)$ is known so that $\ell P(N) = 0$. The minimizing $u(N-1)$ from (3.4) is

$$u(N-1) = \frac{1}{b_1(N)}[-\tilde{\varphi}(N)\hat{\theta}(N)] \tag{A.2}$$

which is the self-tuning regulator (3.5) with $y_r = 0$ and $b_1(t)$ known.

Inserting (A.1) into (3.2) gives

$$V_{N-1} = \min_{u(N-2)} E[y(N-1)^2 + \sigma^2 + y(N-1)^2 P(N) | F_{n-2}] =$$

$$= \min_{u(N-2)} \left\{\sigma^2 + \left[1 + P(N)\right]\left[\left(\varphi(N-1)\hat{\theta}(N-1)\right)^2 + \right.\right.$$

$$\left.\left. + y(N-2)^2 P(N-1) + \sigma^2\right]\right\} =$$

$$= \sigma^2 + [1 + P(N)][y(N-2)^2 P(N-1) + \sigma^2] \tag{A.3}$$

where the minimizing $u(N-2)$ makes

$$\varphi(N-1)\hat{\theta}(N-1) = 0 \tag{A.4}$$

i.e.

$$u(N-2) = \frac{1}{b_1(N-1)}\left[-\tilde{\varphi}(N-1)\hat{\theta}(N-1)\right] \tag{A.5}$$

This is still the self-tuning input.

Now from (2.7)

$$P(N) = \Phi^2 P(N-1) + R - \frac{\Phi^2 P(N-1)^2 y(N-2)^2}{\sigma^2 + y(N-2)^2 P(N-1)} \tag{A.6}$$

With $P(N)$ from (A.6) equation (A.3) becomes

$$V_{N-1} = \sigma^2 + [1 + \Phi^2 P(N-1) + R][y(N-2)^2 P(N-1) + \sigma^2] -$$

$$- \Phi^2 P(N-1)^2 y(N-2)^2 =$$

$$= \sigma^2(2+R) + \sigma^2\Phi^2 P(N-1) + [1+R]P(N-1)y(N-2)^2 \tag{A.7}$$

Again (3.2) is used to give

$$V_{N-2} = \min_{u(N-3)} E\left[y(N-2)^2 + V_{N-1} \mid F_{N-3}\right] =$$

$$= \min_{u(N-3)} \Big[\{1 + (1+R)P(N-1)\}\{[\varphi(N-2)\hat{\theta}(N-2)]^2 +$$

$$+ y(N-3)^2 P(N-2) + \sigma^2\} + \sigma^2(2+R) + \sigma^2\Phi^2 P(N-1)\Big] =$$

$$= \Big[\{1 + (1+R)P(N-1)\}\{y(N-3)^2 P(N-2) + \sigma^2\} +$$

$$+ \sigma^2(2+R) + \sigma^2\Phi^2 P(N-1)\Big] \tag{A.8}$$

Also $u(N-3)$ should be chosen to yield

$$\varphi(N-2)\hat{\theta}(N-2) = 0 \tag{A.9}$$

which gives the self-tuning input

$$u(N-3) = \frac{1}{b_1(N-2)}[-\tilde{\varphi}(N-2)\hat{\theta}(N-2)] \qquad (A.10)$$

In the next step, however, the last term of (A.8) contains $P(N-1)$, and its denominator contains $y(N-3)$. This makes the next expectation impossible to calculate. To summarize, the self-tuning regulator gives an optimal input for this system when looking just 3 steps ahead.

APPENDIX B

The Location of the Global Minimum of (3.17)

Introduce

$$a = \ell P(t+1)\ell^T$$

$$b = \ell P(t+1)\tilde{\varphi}(t+1)^T$$

$$c = a\left(\sigma^2 + \tilde{\varphi}(t+1)P(t+1)\tilde{\varphi}(t+1)^T\right) - b^2$$

$$x = au(t) + b$$

With $f(P) = \ell P \ell^T$ the last term from the argument of f in (3.18) is

$$-\lambda \frac{\ell\Phi P(t+1)[\tilde{\varphi}(t+1) + \ell u(t)]^T[\tilde{\varphi}(t+1) + \ell u(t)]P(t+1)\Phi^T\ell^T}{\sigma^2 + [\tilde{\varphi}(t+1) + \ell u(t)]P(t+1)[\tilde{\varphi}(t+1) + \ell u(t)]^T} \qquad (B.1)$$

But ℓ^T is an eigenvector of Φ^T. The eigenvalue is denoted Φ_b. Then with the definitions above (B.1) becomes

$$-\lambda\Phi_b^2 a \frac{x^2}{c + x^2} = \lambda\Phi_b^2 ac \frac{1}{x^2 + c} - \lambda\Phi_b^2 a \qquad (B.2)$$

All other terms of (3.18) are quadratic in u(t). Now W_w is split up into

$$W_w = f(x) + g(x) \qquad (B.3)$$

where g(x) is the first term in the right member of (B.2), and f(x) is quadratic in x. Then f(x) will have a minimum corresponding to the u(t)-value given by the cautious regulator (3.4). This point is denoted x_1. The functions f and g may typically look as shown in fig. 45.

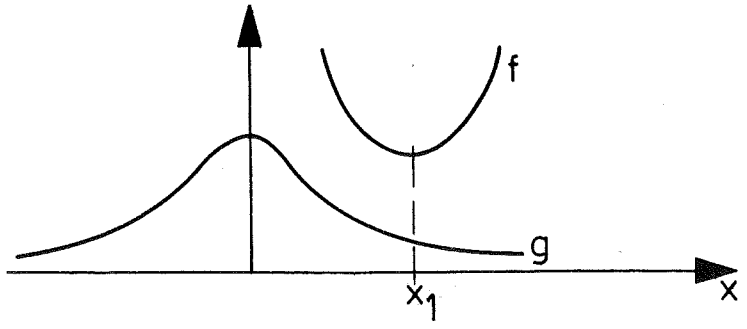


Fig. 45 - f and g of (B.3).

It is evident that g(x) is symmetric around its maximum at x = 0 and tends to zero as $x \to \infty$. It is then no restriction to assume $x_1 \geq 0$. The opposite case follows similarly. Now suppose that x_1 and x^*, the global minimum of W_w, have opposite signs. Then $g(x^*) = g(-x^*)$ but $f(x^*) > f(-x^*)$, at least if $x_1 \neq x^*$. Thus $W_w(-x^*) < W_w(x^*)$ which proves that x_1 and x^* have the same sign. The global minimum must then be outside x_1 since f + g is decreasing from 0 to x_1. To continue, the derivatives of g are needed.

$$g'(x) = -\lambda\Phi_b^2 ac \frac{2x}{[x^2 + c]^2} \tag{B.4}$$

$$g''(x) = \lambda\Phi_b^2 ac \frac{6x^2 - 2c}{[x^2 + c]^3} \tag{B.5}$$

$$g'''(x) = -\lambda\Phi_b^2 ac \frac{24x[x^2 - c]}{[x^2 + c]^4} \tag{B.6}$$

Now $g''(x) \geq 0$ if $x^2 \geq c/3$ and $f''(x) > 0$ give

$$f''(x) + g''(x) > 0 \quad \text{for} \quad x^2 \geq c/3$$

But $g'''(x) = f'''(x) + g'''(x) > 0$ when $0 < x^2 \leq c/3$ shows that $f''(x) + g''(x)$ is increasing so that one of two things will occur.

1) $f''(x) + g''(x) > 0 \qquad \forall\ x > 0$

2) $f''(x) + g''(x) \begin{cases} < 0 & \text{for } 0 \leq x < x_0 \\ > 0 & \text{for } x > x_0 \end{cases}$

Both cases are treated simultaneously by putting $x_0 = 0$ for the first case.

First assume that $x_1 \geq x_0$. Then $f'(x) + g'(x)$ is increasing for $x > x_1$ and if $x_1 \neq 0$ $f'(x_1) + g'(x_1) = g'(x_1) < 0$. There is then exactly one point $x^* > x_1$ with $f'(x^*) + g'(x^*) = 0$. This must be the global minimum.

If $x_1 = 0$ $f'(x) + g'(x) > 0$ $\forall\ x > 0$, and $x = x_1 = 0$ is the only stationary point (global minimum).

Next suppose $x_1 < x_0$. Then again $f'(x_1) + g'(x_1) \leq 0$ and $f' + g'$ is decreasing for $x < x_0$ so that $f'(x) + g'(x) < 0$ for $x_1 < x \leq x_0$. For $x > x_0$ $f' + g'$ is increasing and there is exactly one $x^* > x_1$ with $f'(x^*) + g'(x^*) = 0$. This x^* must then be the global minimum. If $f'(x_1) + g'(x_1) = 0$ then x_1 is a local maximum, since the second derivative is negative.

APPENDIX C

The Two-step Regulator

The optimal two-step input is determined by minimizing w.r.t. $u(N-2)$

$$E\left\{\left[y(N-1) - y_r(N-1)\right]^2 + V_N \middle| F_{N-2}\right\} =$$

$$= \left[\varphi(N-1)\hat{\theta}(N-1) - y_r(N-1)\right]^2 + \varphi(N-1)P(N-1)\varphi(N-1)^T + \sigma^2 +$$

$$+ \sigma^2 + E\left\{\tilde{\varphi}(N)P(N)\tilde{\varphi}(N)^T \middle| F_{N-2}\right\} +$$

$$+ E\left\{\left[y_r(N) - \tilde{\varphi}(N)\hat{\theta}(N)\right]^2 \middle| F_{N-2}\right\} -$$

$$- E\left\{\frac{\left[\ell\hat{\theta}(N)\left(y_r(N) - \tilde{\varphi}(N)\hat{\theta}(N)\right) - \ell P(N)\tilde{\varphi}(N)^T\right]^2}{\left[\ell\hat{\theta}(N)\right]^2 + \ell P(N)\ell^T} \middle| F_{N-2}\right\} \qquad (C.1)$$

In (C.1) eq. (3.3) has been inserted for V_N. The first expectation can be calculated exactly. As discussed in chapter 4 the approximation (4.9) is applied to the difference of the last two terms.

In the following $u(N-2)$ will be denoted by just u. Some additional notations are needed.

$$\varphi(N) = [-y(N-1), \ldots, u(N-1), \ldots]$$

$$\tilde{\varphi}(N) = [-y(N-1), \ldots, 0, u(N-2), \ldots]$$

$$\varphi(N)S = [0, -y(N-1), \ldots, 0, u(N-1), \ldots]$$

$$\ell = [0, \ldots, 0, 1, 0, \ldots]$$

$$\tilde{\ell} = [1, 0, \ldots, 0, 0, 0, \ldots]$$

S is thus a shift matrix. These definitions give

$$\varphi(N) = \tilde{\varphi}(N) + u(N-1)\ell \tag{C.2}$$

$$\tilde{\varphi}(N) = \varphi(N-1)S - y(N-1)\tilde{\ell} \tag{C.3}$$

Also introduce

$$\zeta = y(N-1) - \varphi(N-1)\hat{\theta}(N-1)$$

$$r = E[\zeta^2 | F_{N-2}] = \varphi(N-1)P(N-1)\varphi(N-1)^T + \sigma^2 =$$

$$= \ell P(N-1)\ell^T u^2 + 2\ell P(N-1)\tilde{\varphi}(N-1)^T u +$$

$$+ \tilde{\varphi}(N-1)P(N-1)\tilde{\varphi}(N-1)^T + \sigma^2$$

$$\alpha = \tilde{\varphi}(N-1)\hat{\theta}(N-1)$$

$$\beta = \ell\hat{\theta}(N-1) = \hat{b}_1(N-1)$$

$$\gamma = y_r(N)$$

$$a^T = \Phi\hat{\theta}(N-1)$$

$$b = \ell S - \tilde{\ell}\beta$$

$$d = \tilde{\varphi}(N-1)S - \tilde{\ell}\alpha$$

$$e^T = \Phi P(N-1)\tilde{\varphi}(N-1)^T$$

$$f^T = \Phi P(N-1)\ell^T$$

$$H = \Phi P(N-1)\Phi^T + R$$

The first row of (C.1) is then

$$[\alpha - y_r(N-1) + \beta u]^2 + r \tag{C.4}$$

All these quantities except ζ are known at time N-2, and u enters only in ζ and r. The next step is now to extract the dependence on ζ in (C.1) and calculate the expectations. The Kalman gain is

$$K(N-1) = \frac{1}{r}(e^T + f^T u)$$

In the sequel just K will denote K(N-1). Also

$$\tilde{\varphi}(N) = \varphi(N-1)S - \tilde{\ell}y(N-1) = d + bu - \tilde{\ell}\zeta$$

$$\hat{\theta}(N) = a^T + K\zeta$$

$$P(N) = H - rKK^T$$

Then for the first term of expectation in (C.1)

$$\tilde{\varphi}(N)P(N)\tilde{\varphi}(N)^T = (d+bu-\tilde{\ell}\zeta)(H-rKK^T)(d+bu-\tilde{\ell}\zeta)^T$$

Since $E[\zeta|F_{N-2}] = 0$ the expectation is

$$(d+bu)(H-rKK^T)(d+bu)^T + \tilde{\ell}(H-rKK^T)\tilde{\ell}^T r =$$

$$= (d+bu)H(d+bu)^T - \frac{1}{r}[(d+bu)(e^T+f^Tu)]^2 +$$

$$+ r\tilde{\ell}H\tilde{\ell}^T - (\tilde{\ell}e^T+\tilde{\ell}f^Tu)^2 \qquad (C.5)$$

Since u enters into r, (C.5) is no polynomial, but a rational function of u.

The common denominator of the last two terms of (C.1) gives

$$[\ell\hat{\theta}(N)]^2 + \ell P(N)\ell^T = [\ell a^T+\ell K\zeta]^2 + \ell(H-rKK^T)\ell^T$$

The expectation of this is

$$[\ell a^T]^2 + (\ell K)^2 r + \ell H\ell^T - r\ell KK^T\ell^T = [\ell a^T]^2 + \ell H\ell^T \qquad (C.6)$$

The expected value of the denominator does thus not depend on u.

The last two terms should be combined before taking expectations. The numerator is then

$$[y_r(N) - \tilde{\varphi}(N)\hat{\theta}(N)]^2\left\{[\ell\hat{\theta}(N)]^2 + \ell P(N)\ell^T\right\} -$$

$$- \left\{\ell\hat{\theta}(N)[y_r(N) - \tilde{\varphi}(N)\hat{\theta}(N)] - \ell P(N)\tilde{\varphi}(N)^T\right\}^2 =$$

$$= \ell P(N)\ell^T[y_r(N) - \tilde{\varphi}(N)\hat{\theta}(N)]^2 +$$

$$+ 2\ell\hat{\theta}(N)[y_r(N) - \tilde{\varphi}(N)\hat{\theta}(N)]\ell P(N)\tilde{\varphi}(N)^T -$$

$$- [\ell P(N)\tilde{\varphi}(N)^T]^2 =$$

$$= \ell(H-rKK^T)\ell^T[\gamma - (d+bu-\tilde{\ell}\zeta)(a^T+K\zeta)]^2 +$$

$$+ 2(\ell a^T+\ell K\zeta)[\gamma - (d+bu-\tilde{\ell}\zeta)(a^T+K\zeta)]\ell(H-rKK^T)(d+bu-\tilde{\ell}\zeta)^T -$$

$$- [\ell(H-rKK^T)(d+bu-\tilde{\ell}\zeta)^T]^2 =$$

$$= A + B\zeta + C\zeta^2 + D\zeta^3 + E\zeta^4 \qquad (C.7)$$

where

$$A = \ell(H-rKK^T)\ell^T[\gamma - (d+bu)a^T]^2 +$$

$$+ 2\ell a^T[\gamma - (d+bu)a^T]\ell(H-rKK^T)(d+bu)^T -$$

$$- [\ell(H-rKK^T)(d+bu)^T]^2 \qquad (C.8)$$

$$C = \ell(H-rKK^T)\ell^T\left\{[\tilde{\ell}a^T - (d+bu)K]^2 +\right.$$

$$\left. + 2[\gamma - (d+bu)a^T]\tilde{\ell}K\right\} +$$

$$+ 2\ell a^T\left\{\tilde{\ell}K\ell(H-rKK^T)(d+bu)^T -\right.$$

$$\left. - [\tilde{\ell}a^T - (d+bu)K]\ell(H-rKK^T)\tilde{\ell}^T\right\} +$$

$$+ 2\ell K\left\{[\tilde{\ell}a^T - (d+bu)K]\ell(H-rKK^T)(d+bu)^T -\right.$$

$$\left. - [\gamma - (d+bu)a^T]\ell(H-rKK^T)\tilde{\ell}^T\right\} -$$

$$- [\ell(H-rKK^T)\tilde{\ell}^T]^2 \qquad (C.9)$$

$$E = \ell(H-rKK^T)\ell^T[\tilde{\ell}K]^2 - 2(\ell K)(\tilde{\ell}K)\ell(H-rKK^T)\tilde{\ell}^T \qquad (C.10)$$

The expressions for B and D are of no relevance since they will disappear in the expectation. With

$$E[\zeta^4|F_{N-2}] = 3r^2 \qquad (C.11)$$

the expected value of the numerator of the last two terms is

$$A + Cr + 3Er^2 \qquad (C.12)$$

where A, C and E are given above.

In (C.8) - (C.10) only r, K and u itself depend on the input. It can then be shown that

$$Ar + Cr^2 + 3Er^3 \qquad (C.13)$$

is a 4th order polynomial in u. To see this, notice that rK is a linear function of u, and r is quadratic in u. It turns out that the non-polynomial parts of (C.13) will cancel. They are

$$-\frac{1}{r}[\ell rKrK^T(d+bu)^T]^2 \qquad (C.8a)$$

and

$$-\frac{1}{r}[\ell rKrK^T\ell^T\{(d+bu)K\}^2] +$$

$$+\frac{1}{r}2\ell rK(d+bu)rK\ell rKrK^T(d+bu)^T \qquad (C.9a)$$

The complete expression for the approximate two-step loss is then

$$\frac{1}{r}\left\{r[\alpha - y_r(N-1) + \beta u]^2 + r^2 + \sigma^2 r +\right.$$

$$+ r(d+bu)H(d+bu)^T - [(d+bu)(e^T+f^Tu)]^2 +$$

$$+ r^2\tilde{\ell}H\tilde{\ell}^T - r(\tilde{\ell}e^T+\tilde{\ell}f^Tu)^2 +$$

$$+ \frac{1}{[\ell a^T]^2 + \ell H\ell^T} \cdot (Ar + Cr^2 + 3Er^3)\Big\} \qquad (C.14)$$

with A, C and E given above.

To find the minimum of (C.14) its derivative is calculated. This will give a 5th order numerator, since (C.14) has a 4th order numerator and a 2nd order denominator. Using a root-finding algorithm the zeroes of the derivative can be found, and the corresponding values of (C.14) can be compared to find the global minimum.

REFERENCES

Abramowicz, H., Stymne, K.J. (1975):
Duala regulatorer (in Swedish). TFRT-5159, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Alster, J., Bélanger, P.R. (1974):
A Technique for Dual Adaptive Control. Automatica 10, pp. 627-634.

Åström, K.J. (1970):
Introduction to Stochastic Control. Academic Press.

Åström, K.J., Wittenmark, B. (1971):
Problems of Identification and Control. J. of Math. Anal. & Appl. 34, pp. 90-113.

Åström, K.J., Wittenmark, B. (1973):
On Self-Tuning Regulators. Automatica 9, pp. 185-199.

Åström, K.J., Borisson, U., Ljung, L., Wittenmark, B. (1977):
Theory and Applications of Self-Tuning Regulators. To appear in Automatica 13.

Bar-Shalom, Y., Sivan, R. (1972):
On the Optimal Control of Discrete-Time Linear Systems with Random Parameters. IEEE Trans. Automatic Control, AC-14, pp. 3-8.

Bar-Shalom, Y., Tse, E. (1974):
Dual Effect, Certainty Equivalence, and Separation in Stochastic Control. IEEE Trans. Automatic Control, AC-19, pp. 494-500.

Bohlin, T. (1969):
Optimal Dual Control of A Simple Process with Unknown Gain. Technical Paper TP 18.196, IBM Nordic Laboratory, Lidingö, Sweden.

Chow, G.C. (1975):
Analysis and Control of Dynamic Economic Systems, Chapter 11. John Wiley & Sons.

Clarke, D.W., Gawthrop, P.J. (1975):
Self-tuning Controller. Proc. IEE Control & Science 122, pp. 929-934.

Dreyfus, S. (1964):
Some Types of Optimal Control of Stochastic Systems. SIAM J. Control 2, 120-134.

Feldbaum, A.A. (1960):
Dual Control Theory I-IV. Automation & Remote Control 21, pp. 874-880, 1033-1039, and 22, 1-12, 109-121.

Florentin, J.J. (1962):
Optimal, Probing, Adaptive Control of a Simple Bayesian System. J. Electron Control 11, pp. 165-177.

Frost, V.A. (1970):
Identification and Control of Linear Discrete Time Systems with Noisy Output Measurements. Int. J. Control 11, pp. 571-580.

Gunckel, T.L., Franklin, G.F. (1963):
A General Solution for Linear Sampled-Data Control. Trans. ASME, J. Basic. Eng. 85, pp. 197-203.

Hughes, D.J., Jacobs, O.L.R. (1974):
Turn-off, Escape and Probing in Nonlinear Stochastic Control. Preprint IFAC Symp. on Stochastic Control, Budapest, Hungary.

Jacobs, O.L.R. (1974):
Structure of Stochastic Control Laws. IMA Conf. on Stochastic Programming, Oxford, England.

Jacobs, O.L.R., Hughes, D.J. (1975):
Simultaneous Identification and Control using a Neutral Control Law. Proc. IFAC 6th World Congress, paper 30.2, Boston, USA.

Jacobs, O.L.R., Patchell, J.W. (1972):
Caution and Probing in Stochastic Control. Int. J. Control 16, pp. 189-199.

Ku, R., Athans, M. (1973):
On the Adaptive Control of Linear Systems Using the Open-Loop-Feedback-Optimal Approach. IEEE Trans. Automatic Control AC-18, pp. 489-493.

Ljung, L. (1976):
Consistency of the Least-Squares Identification Method. IEEE Trans. Automatic Control AC-21, pp. 779-781.

Ljung, L. (1977):
Analysis of Recursive Stochastic Algorithms. To appear in IEEE Trans. Automatic Control.

Ljung, L., Wittenmark, B. (1974):
Asymptotic Properties of Self-Tuning Regulators. TFRT-3071, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Ljung, L., Wittenmark, B. (1976):
On a Stabilizing Property of Adaptive Regulators. Preprint IFAC Symp. on Identification and System Parameter Estimation, Tbilisi, USSR.

Mannerfelt, C.F. (1977):
Undersökning av några duala regulatorer (in Swedish). Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, CODEN: LUTFD2/(TFRT-5194)/1- 33 /(1977).

Mehra, R. (1974):
Dual Control of Linear Dynamic Systems with Unknown Parameters Using Bayesian Preposterior Analysis. Allerton Conf., October 1974.

Murphy, W.J. (1968):
Optimal Stochastic Control of Discrete Linear Systems with Unknown Gain. IEEE Trans. Automatic Control AC-13, pp. 338--344.

Nakamura, S., Nakamura, K. (1973):
On the Effects of Perturbation Signal in Dual Control System. Research Report of Auto. Cont. Lab., Vol. 20, Nagoya Univ.

Norman, A.L. (1976):
First Order Dual Control. Ann. of Economic and Social Measurement, 5/3 1976, pp. 311-321.

Saridis, G.N., Lobbia, R.N. (1972):
Parameter Identification and Control of Linear Discrete-Time Systems. IEEE Trans. Automatic Control AC-17, 52-60.

Sternby, J. (1976):
A Simple Dual Control Problem with an Analytical Solution. IEEE Trans. Automatic Control AC-21, pp. 840-844.

Sternby, J. (1977):
On Consistency for the Method of Least Squares Using Martingale Theory. Submitted to IEEE Trans. Automatic Control.

Sternby, J., Pernebo, L. (1977):
A Comparison of Two Suboptimal Dual Controllers on a First Order System. To be published.

Tou, J.T. (1963):
Optimal Design of Digital Control Systems. Academic Press.

Tse, E., Athans, M. (1972):
Adaptive Stochastic Control for a Class of Linear Systems. IEEE Trans. Automatic Control AC-17, pp. 38-51.

Tse, E., Bar-Shalom, Y. (1973):
An Actively Adaptive Control for Linear Systems with Random Parameters via the Dual Control Approach. IEEE Trans. Automatic Control AC-18, pp. 109-117.

Tse, E., Bar-Shalom, Y., Meier, L. (1973):
Wide-Sense Adaptive Dual Control for Nonlinear Stochastic Systems. IEEE Trans. Automatic Control AC-18, 98-108.

Wieslander, J., Wittenmark, B. (1971):
An Approach to Adaptive Control Using Real-Time Identification. Automatica 7, pp. 211-217.

Wittenmark, B. (1971):
On the Turn-off Phenomenon in Adaptive Control. TFRT-3030, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Wittenmark, B. (1973):
A Self-Tuning Regulator. TFRT-3054, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Wittenmark, B. (1975a):
An Active Suboptimal Dual Controller for Systems with Stochastic Parameters. Automatic Control Theory and Application 3, pp. 13-19.

Wittenmark, B. (1975b):
Stochastic Adaptive Control Methods: A Survey. Int. J. Control 21, pp. 705-730.

Sigma-Tryck
TLTH Lund