On Positive Real Transfer Functions and the Convergence of Some Recursive Schemes

Ljung, Lennart

1976

TFRT-3138

ON POSITIVE REAL TRANSFER FUNCTIONS AND THE CONVERGENCE OF SOME RECURSIVE SCHEMES

L. LJUNG

Report 7630(C) June 1976
Department of Automatic Control
Lund Institute of Technology

On Positive Real Transfer Functions and the Convergence of Some Recursive Schemes

Lennart Ljung

Dept of Automatic Control
Lund Institute of Technology
Box 725
S-220 07 LUND 7 SWEDEN

Abstract

The convergence with probability one of a recently suggested recursive identification method by Landau is investigated. The positive realness of a certain transfer function is shown to play a crucial role, both for the proof of convergence and for convergence itself. A completely analogous analysis can be performed also for the Extended Least squares method and for the Self-tuning regulator of Åström and Wittenmark. Explicit conditions for convergence of all these schemes are given.

1. INTRODUCTION

In a recent paper by Landau [1], a recursive algorithm for the identification of dynamical systems is suggested. The literature on various such schemes is already quite extensive, but Landau's scheme and analysis exhibit some particularly interesting features. From a theoretical viewpoint, the use and importance of positive real transfer functions for the convergence analysis in [1] is interesting, since no such properties seem to have been used previously in the literature on recursive identification (except in Landau's own previous work).

In [1] convergence of the algorithm is proved for the noiseless-case utilizing hyper-stability theory. It is also shown that in the case when noise affects the system, the obtained estimates are asymptotically unbiased. Here a stronger result will be shown, namely convergence w.p. 1 of the estimates to the true values. For this proof the method of analysis of stochastic, recursive algorithms discussed in [2] will be used. An interesting feature of the proof is that the positive realness property here enters in a completely different way compared to the analysis in [1].

A perhaps more important property is that other recursive schemes like the extended least squares method (ELS), and the self-tuning regulator exhibit a very similar structure, and hence also the convergence of these schemes depend on the positive realness of certain transfer functions. The ELS method (also known as "Panuska's method", "the approximate maximum likelihood method", or "the extended matrix method") has been widely discussed, e.g. in [3], [4], [5], [6]. In [7] it was shown that ELS does not converge for certain systems, but with the results of the present paper and of [13] a more complete picture of its convergence properties will be obtained. The self-tuning regulator was suggested in [8] and its convergence properties were discussed in [9], [10], but the present results imply a considerable extension of the set of systems for which this adaptive regulator is known to converge to the optimal one.

The paper is organized as follows. First, in Section 2, the three algorithms are described. Section 3 contains a brief summary of the results of [2], and shows how this analysis method can be applied to the algorithms under discussion. The role that the positive real-property plays in this context is illustrated in Section 4, while possible convergence points of the algorithms are determined in Section 5. The global convergence of the algorithms is analysed in Section 6, where also explicit conditions for convergence are given.

All of this development is made in parallel for the three algorithms. Most of the notation, like $\theta(t)$, $\varphi(t)$, $\tilde{\varphi}(t,\theta)$, $\tilde{G}(\theta)$ etc, to be used below will be common to the different algorithms and the current interpretation of the symbols will depend on the context. This should however not lead to any ambiguities.

Finally, in Section 7 the results are summarized and discussed.

2. THE ALGORITHMS

Landau's Scheme

First the scheme of [1] will be described briefly. For all details it is referred to [1]. The notation will differ from [1], in order to make it consistent with that of [2] and [7] and Table 1 gives a cross-reference between the notation of [1] and the present one.

The true system is assumed to be given by

$$A(q^{-1})y(t) = B(q^{-1})u(t) + w(t) \tag{1}$$

where q^{-1} is the delay operator

$$q^{-1}\, y(t) = y(t-1)$$

and

$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_n q^{-n}$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + \ldots + b_m q^{-m}$$

The term $w(t)$ represents some (unmeasurable) disturbance acting on the system. With

$$\theta_0 = (a_1, \ldots, a_n, b_0, \ldots, b_m)^T$$

and

$$\varphi_0(t) = \big(-y(t-1), \ldots, -y(t-n), u(t), \ldots, u(t-m)\big)^T$$

(1) can be written

$$y(t) = \theta_0^T \varphi_0(t) + w(t) \tag{2}$$

In addition, a model based on current estimates of the parameters $\hat{a}_i(t)$, $\hat{b}_i(t)$ is used:

$$y_M(t) + \hat{a}_1(t)\, y_M(t-1) + \ldots + \hat{a}_n(t)\, y_M(t-n) =$$

$$= \hat{b}_0(t)\, u(t) + \ldots + \hat{b}_m(t)\, u(t-m) \qquad (3)$$

which can be written

$$y_M(t) = \hat{\theta}(t)^T \varphi(t) \qquad (4)$$

with

$$\hat{\theta}(t) = \left(\hat{a}_1(t), \ldots, \hat{a}_n(t), \hat{b}_0(t), \ldots, \hat{b}_m(t)\right)^T$$

$$\varphi(t) = \left(-y_M(t-1), \ldots, -y_M(t-n), u(t), \ldots, u(t-m)\right)^T$$

The sequence of estimates $\{\hat{\theta}(t)\}$ will be defined recursively by (7) below. Introduce

$$\varepsilon_0(t) = y(t) - y_M(t) \qquad (5)$$

and

$$\varepsilon(t) = y(t) - \hat{\theta}(t-1)^T \varphi(t) + d_1\, \varepsilon_0(t-1) + \ldots + d_n \varepsilon_0(t-n) \qquad (6)$$

where d_i are some suitably chosen numbers, discussed below. Then the estimates $\hat{\theta}(t)$ are recursively defined by:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{1}{1 + \frac{1}{t}\left[\varphi(t)^T R^{-1}(t-1)\varphi(t) - 1\right]} R^{-1}(t-1) \cdot$$

$$\cdot\ \varphi(t)\ \varepsilon(t) \qquad (7a)$$

$$R(t) = R(t-1) + \frac{1}{t}\left(\varphi(t)\varphi(t)^T - R(t-1)\right) \qquad (7b)$$

(In [1] the updating formula (7b) is given for $F_{t-1} = \frac{1}{t} R^{-1}(t-1)$ directly.)

We can also define a Stochastic approximation variant of (7) by replacing the matrix $R(t)$ by the scalar $r(t) = tr\ R(t)$:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{1}{t}\left(\frac{1}{r(t-1)} \cdot \varphi(t)\ \varepsilon(t)\right) \tag{8a}$$

$$r(t) = r(t-1) + \frac{1}{t}\left(\varphi(t)^T \varphi(t) - r(t-1)\right) \tag{8b}$$

Table 1

Present Notation Compared to that of [1].

Here	In [1]
t	k
y	θ_p
u	ρ
θ_0	p
$\hat{\theta}(t)$	$\hat{p}(t)$
$\varphi_0(t)$	x_{t-1}
y_M	θ_s
ε	ν^o
$\varphi(t)$	y_{t-1}
$\varepsilon_0(t)$	ε_t
$R(t)$	$(t \cdot F_t)^{-1}$
a_i	$-a_i$

Extended Least Squares

Let us now proceed by describing the ELS-method. For more details, see e.g. [7].

The true system is supposed to be

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})e(t) \tag{9}$$

where A and B are as before and

$$C(q^{-1}) = 1 + c_1 q^{-1} + \ldots + c_n q^{-1} \tag{10}$$

With

$$\varphi_0(t) = \left(-y(t-1),\ldots,-y(t-n), u(t),\ldots,u(t-m), e(t-1),\ldots,e(t-n)\right)^T$$

$$\theta_0 = (a_1,\ldots,a_n, b_0,\ldots,b_m, c_1,\ldots,c_n)^T$$

(9) can be written

$$y(t) = \theta_0^T \varphi_0(t) + e(t) \tag{11}$$

The disturbance $\{e(t)\}$ is supposed to be white noise. Now $\{e(t)\}$ is not measurable so $\varphi_0(t)$ is not available, but we may recursively define $\varepsilon(t)$ and $\varphi(t)$ from

$$\varphi(t) = \left(-y(t-1),\ldots,-y(t-n), u(t),\ldots,u(t-m), \varepsilon(t-1),\ldots,\varepsilon(t-n)\right)^T \tag{12a}$$

$$\varepsilon(t) = y(t) - \hat{\theta}^T(t-1)\varphi(t) \tag{12b}$$

where $\hat{\theta}(t)$ is the current estimate of θ_0 defined by (7) (or (8) for a stochastic approximation variant) with the new interpretations of $\varepsilon(t)$, $\varphi(t)$ given by (12).

Self-tuning Regulator

The true system is supposed to be (9) with the additional assumption that $b_0 = 0$ (i.e. there is a time delay in the system) and that b_1 is known.

The model then is

$$y(t) = b_1 u(t-1) + \hat{\theta}^T \varphi(t) + w(t) \tag{13}$$

where

$$\hat{\theta} = (\hat{a}_1, \ldots, \hat{a}_n, \hat{b}_2, \ldots, \hat{b}_m)^T$$

and

$$\varphi(t) = \Big(-y(t-1), \ldots, -y(t-n),\ u(t-2), \ldots, u(t-m)\Big)^T \tag{14a}$$

The current estimate of θ, $\hat{\theta}(t)$ is given by (7) (or (8) for a stochastic approximation variant) with $\varphi(t)$ as in (14a) and $\varepsilon(t)$ given by

$$\varepsilon(t) = y(t) - b_1 u(t-1) - \hat{\theta}(t-1)^T \varphi(t) \tag{14b}$$

The input to the process u(t) is chosen, based on the estimate $\hat{\theta}(t)$ as

$$u(t) = -\frac{1}{b_1} \hat{\theta}(t)^T \varphi(t+1) \tag{15}$$

The rationale for this choice of regulator is explained in [8]. Basically, the reason is that if $\hat{\theta}(t)$ were equal to θ_0 and w(t) was white noise; (i.e. $C(z) \equiv 1$) then this control would be the minimum variance control, [11]. It is interesting to note, that even if C(z) is different from unity, minimum variance control can still be obtained using (15) with

$$u(t) = -\frac{1}{b_1} \theta_{MV}^T \varphi(t+1) \tag{16}$$

where

$$\theta_{MV} = (a_1-c_1,\dots,a_n-c_n,\ b_2,\dots,b_m)^T \tag{17}$$

see [8]. The most important feature of the self-tuning regulator [8] is that experience both in simulations and industrial applications shows that indeed $\theta(t)$ tends to θ_{MV} "usually", even though it has been proved, [12], that such convergence does not take place for all systems (9).

3. METHOD OF ANALYSIS

The idea of the convergence analysis is to associate (7) (or (8)) with a differential equation that contains all necessary information about the asymptotic behaviour. This approach is explained in detail in [2]. To define the differential equation associated with (7) (or (8)) stationary processes $\{\bar{\varepsilon}(t,\theta)\}$ and $\{\bar{\varphi}(t,\theta)\}$ will be introduced. Precise definitions are given below, and let it suffice here to say that these processes would be obtained as $\{\varepsilon(t)\}$ and $\{\varphi(t)\}$ respectively if the estimate sequence was fixed to a constant value θ in the relations defining ε and φ.

Then define

$$f(\theta) = E\,\bar{\varphi}(t,\theta)\,\bar{\varepsilon}(t,\theta) \tag{18}$$

$$G(\theta) = E\,\bar{\varphi}(t,\theta)\,\bar{\varphi}(t,\theta)^T \tag{19}$$

$$g(\theta) = E\,\bar{\varphi}(t,\theta)^T\,\bar{\varphi}(t,\theta) \tag{20}$$

where the expectation is over the stochastic processes involved in the respective problem ($\{w(t)\}$ and $\{u(t)\}$ for Landau's scheme; $\{e(t)\}$ and $\{u(t)\}$ for ELS and $\{e(t)\}$ for the self-tuning regulator).

The differential equation associated with (7) then is, according to [2],

$$\frac{d}{d\tau}\,\theta(\tau) = R^{-1}(\tau)\ f\big(\theta(\tau)\big) \tag{21a}$$

$$\frac{d}{d\tau}\,R(\tau) = G\big(\theta(\tau)\big) - R(\tau) \tag{21b}$$

and the stochastic approximation variant (8) is associated with

$$\frac{d}{d\tau}\,\theta(\tau) = \frac{1}{r(\tau)}\ f\big(\theta(\tau)\big) \tag{22a}$$

$$\frac{d}{d\tau}\,r(\tau) = g\big(\theta(\tau)\big) - r(\tau) \tag{22b}$$

Essentially, global stability of (21) (or (22)) implies convergence w.p.1 of (7) ((8)), and furthermore, the possible convergence points of the algorithms are only the stable stationary points of the corresponding differential equation. More details of this will be given in sections 5 and 6. The results are valid under certain regularity conditions of the algorithm and of the stochastic processes. That (7) (and (8)) satisfies these regularity conditions is verified in [2], example 3. Further conditions on the stochastic processes will be given below.

The remainder of this section will deal with more specific assumptions upon the respective algorithms, the precise definitions of $\bar{\varepsilon}$ and $\bar{\varphi}$ and with the structure of the function $f(\theta)$ defined by (18). This analysis will be given separately for the three schemes, but with the notation in common as before.

Landau's Scheme

Assume that $\{w(t)\}$ and $\{u(t)\}$ are stationary stochastic processes with rational spectral densities such that all moments exist. Assume also that the system operates in open loop (i.e. that $\{u(t)\}$ and $\{w(t)\}$ are independent) and that it is stable.

Let $\bar{y}_m(t,\theta)$ and $\bar{\varphi}(t,\theta)$ be the stationary processes defined by, cf (4),

$$\bar{y}_m(t,\theta) = \theta^T \bar{\varphi}(t,\theta) \tag{23a}$$

$$\bar{\varphi}(t,\theta) = \left(-\bar{y}_m(t-1,\theta),\ldots,-\bar{y}_m(t-n,\theta),u(t),\ldots,u(t-m)\right)^T \tag{23b}$$

Let $\bar{\varepsilon}_0(t,\theta)$ and $\bar{\varepsilon}(t,\theta)$ be defined by, cf (5),(6),

$$\overline{\varepsilon}_0(t,\theta) = y(t) - \overline{y}_m(t,\theta) \tag{24a}$$

$$\overline{\varepsilon}(t,\theta) = D(q^{-1})\ \overline{\varepsilon}_0(t,\theta) \tag{24b}$$

where

$$D(q^{-1}) = 1 + d_1 q^{-1} + \ldots + d_n q^{-n} \tag{25}$$

Notice that these variables are defined only for such θ for which the corresponding $\hat{A}$-part is stable, i.e. for θ, such that the recursion (23) is stable so that a stationary process really can be defined. We shall denote this area by D_s:

$$D_s = \left\{ \theta \mid \theta = (\hat{a}_1, \ldots, \hat{a}_n, \hat{b}_0, \ldots, \hat{b}_m)^T; \quad z^n + \hat{a}_1 z^{n-1} + \ldots + \hat{a}_n = 0 \Rightarrow |z| < 1 \right\} \tag{26}$$

Let us now proceed to analyse the quantities $\overline{\varepsilon}(t,\theta)$ and $\overline{\varphi}(t,\theta)$ in order to obtain a more explicit expression for $f(\theta)$.

We have from (2)

$$y(t) = \theta_0^T\ \varphi_0(t) + w(t);$$

and using (23) and (24)

$$\begin{aligned}\overline{\varepsilon}_0(t,\theta) &= \theta_0^T\ \varphi_0(t) - \theta^T\ \overline{\varphi}(t,\theta) + w(t) = \\ &= \varphi_0(t)^T\ \theta_0 - \overline{\varphi}(t,\theta)^T\theta_0 + \overline{\varphi}(t,\theta)^T\theta_0 - \overline{\varphi}(t,\theta)^T\theta + w(t) = \\ &= \left(\varphi_0(t) - \overline{\varphi}(t,\theta)\right)^T\theta_0 + \overline{\varphi}(t,\theta)^T(\theta_0-\theta) + w(t)\end{aligned}$$

But

$$\begin{aligned}&\left(\varphi_0(t) - \overline{\varphi}(t,\theta)\right)^T = \\ &= \left(\overline{y}_m(t-1,\theta)-y(t-1), \ldots, \overline{y}_m(t-n,\theta)-y(t-n),\ 0,\ldots,0\right) = \\ &= \left(-\overline{\varepsilon}_0(t-1,\theta), \ldots, -\overline{\varepsilon}_0(t-n,\theta),\ 0,\ldots,0\right)\end{aligned}$$

and

$$\left(\varphi_0(t) - \bar{\varphi}(t,\theta)\right)^T \theta_0 = \left(1-A(q^{-1})\right) \bar{\varepsilon}_0(t,\theta)$$

Hence

$$A(q^{-1}) \bar{\varepsilon}_0(t,\theta) = \bar{\varphi}(t,\theta)^T (\theta_0-\theta) + w(t)$$

and

$$\bar{\varepsilon}(t,\theta) = \frac{D(q^{-1})}{A(q^{-1})} \bar{\varphi}(t,\theta)^T (\theta_0-\theta) + \frac{D(q^{-1})}{A(q^{-1})} w(t)$$

Introduce $\tilde{\varphi}(t,\theta)$ as the stationary process

$$\tilde{\varphi}(t,\theta) = H(q^{-1}) \bar{\varphi}(t,\theta) \tag{27}$$

where

$$H(q^{-1}) = \frac{D(q^{-1})}{A(q^{-1})} \tag{28}$$

Then

$$f(\theta) = E \bar{\varphi}(t,\theta) \bar{\varepsilon}(t,\theta) = E \bar{\varphi}(t,\theta) \tilde{\varphi}(t,\theta)^T (\theta_0-\theta)$$

since $w(t)$ is independent of $\{u(t)\}$ and $\{\bar{y}_m(t,\theta)\}$.
Introduce the notation

$$\tilde{G}(\theta) = E \bar{\varphi}(t,\theta) \tilde{\varphi}(t,\theta)^T \tag{29}$$

$$\theta^* = \theta_0 \tag{30}$$

Then

$$f(\theta) = \tilde{G}(\theta) (\theta^*-\theta) \tag{31}$$

Remark. Notice that the assumption of open loop operation is vital in these calculations!

The ELS Scheme

Assume that $\{e(t)\}$ is a stationary sequence of independent random variables and that $\{u(t)\}$ is a stationary stochastic process with rational spectral density. Assume further that all moments of $e(t)$ and $u(t)$ exist. The input may partly be determined as linear output feedback. Assume that the (closed loop) system is stable.

Let $\bar{\varepsilon}(t,\theta)$ and $\bar{\varphi}(t,\theta)$ be the stationary process defined by, cf (12),

$$\bar{\varepsilon}(t,\theta) = y(t) - \theta^T \bar{\varphi}(t,\theta)$$

$$\bar{\varphi}(t,\theta) = \left(-y(t-1),\ldots,-y(t-n),u(t),\ldots,u(t-m),\bar{\varepsilon}(t-1,\theta),\ldots,\bar{\varepsilon}(t-n,\theta)\right)^T \tag{32}$$

Notice that these processes are defined only if the recursion in (32) is stable, i.e. for θ such that the corresponding $\hat{C}$-polynomial is stable. Let this set be denoted by D_s:

$$D_s = \Big\{ \theta \,|\, \theta = (\hat{a}_1,\ldots,\hat{a}_n,\ \hat{b}_0,\ldots,\hat{b}_m,\ \hat{c}_1,\ldots,\ \hat{c}_n)^T; \quad z^n + \hat{c}_1 z^{n-1} + \ldots + \hat{c}_n = 0 \Rightarrow |z| < 1 \Big\} \tag{33}$$

It is an important observation, that $f(\theta)$ also in this case can be written analogously to (31). To show this, we have from (32),

$$\bar{\varepsilon}(t,\theta) = y(t) - \theta^T \bar{\varphi}(t,\theta) = \theta_0^T \varphi_0(t) - \theta^T \bar{\varphi}(t,\theta) + e(t) =$$

$$= \left(\varphi_0(t) - \bar{\varphi}(t,\theta)\right)^T \theta_0 + \bar{\varphi}(t,\theta)^T (\theta_0 - \theta) + e(t)$$

Here

$$\left(\varphi_0(t) - \bar{\varphi}(t,\theta)\right)^T = \left(0,\ldots,0,\ldots,0,\ldots,0,e(t-1) - \bar{\varepsilon}(t-1,\theta),\ldots \right.$$

$$\left. \ldots,e(t-n) - \bar{\varepsilon}(t-n,\theta)\right)$$

and

$$\left(\varphi_0(t) - \bar{\varphi}(t,\theta)\right)^T \theta_0 = \left(C(q^{-1})-1\right)\left(e(t) - \bar{\varepsilon}(t,\theta)\right)$$

Hence

$$C(q^{-1})\bar{\varepsilon}(t,\theta) = C(q^{-1})e(t) + \bar{\varphi}(t,\theta)^T(\theta_0-\theta)$$

and

$$\bar{\varepsilon}(t,\theta) = \frac{1}{C(q^{-1})} \bar{\varphi}(t,\theta)^T (\theta_0-\theta) + e(t)$$

Introduce

$$\tilde{\varphi}(t,\theta) = H(q^{-1}) \bar{\varphi}(t,\theta) \tag{34}$$

where

$$H(q^{-1}) = \frac{1}{C(q^{-1})} \tag{35}$$

Then, since $e(t)$ is independent of $y(s)$, $u(s)$, $\bar{\varepsilon}(s,\theta)$ for $s < t$ (also for closed loop operation), we have

$$f(\theta) = E\, \bar{\varphi}(t,\theta)\, \bar{\varepsilon}(t,\theta) = \tilde{G}(\theta)\, (\theta^*-\theta) \tag{36}$$

where

$$\tilde{G}(\theta) = E\, \bar{\varphi}(t,\theta)\, \tilde{\varphi}(t,\theta)^T \tag{37}$$

$$\theta^* = \theta_0 \tag{38}$$

The Self-tuning Regulator

Assume that $\{e(t)\}$ is a stationary sequence of independent random variables, such that all moments exist. Assume further that the system (9) is minimum phase.

Both y(t) and u(t) are influenced by the sequence of estimates $\{\hat{\theta}(t)\}$ through the control law (15). Define $\bar{y}(t,\theta)$, $\bar{u}(t,\theta)$, $\bar{\varphi}(t,\theta)$, $\bar{\varepsilon}(t,\theta)$ as the stationary process which would be obtained with a constant control law corresponding to θ.

$$A(q^{-1})\ \bar{y}(t,\theta) = B(q^{-1})\ \bar{u}(t,\theta) + C(q^{-1})\ e(t) \tag{39a}$$

$$\bar{\varphi}(t,\theta) = \left(-\bar{y}(t-1,\theta),\dots,-\bar{y}(t-n,\theta),\bar{u}(t-2,\theta),\dots,\bar{u}(t-m,\theta)\right)^T \tag{39b}$$

$$\bar{u}(t,\theta) = -\frac{1}{b_1}\ \theta^T\ \bar{\varphi}(t+1,\theta) \tag{39c}$$

$$\bar{\varepsilon}(t,\theta) = \bar{y}(t,\theta) - b_1\ \bar{u}(t-1,\theta) - \theta^T\ \bar{\varphi}(t,\theta) = \bar{y}(t,\theta) \tag{39d}$$

These processes are defined only for those θ that give a stable closed loop system. Let this set be denoted by D_S:

$$\begin{aligned} D_s = \Big\{ \theta \,|\, \theta = (\hat{a}_1,\dots,\hat{a}_n,\ \hat{b}_2,\dots,\hat{b}_m);\\ (b_1 z^m + \hat{b}_2 z^{m-1} + \dots + \hat{b}_m)(z^n + a_1 z^{n-1} + \dots + a_n) - \\ - (b_1 z^m + b_2 z^{m-1} + \dots + b_m)(\hat{a}_1 z^{n-1} + \dots + \hat{a}_n) = 0 \\ \Rightarrow |z| < 1 \Big\} \end{aligned} \tag{40}$$

Furthermore, we have

$$\begin{aligned} \bar{\varepsilon}(t,\theta) = \bar{y}(t,\theta) &= \theta_0^T\ \bar{\varphi}(t,\theta) + b_1\bar{u}(t-1,\theta) + C(q^{-1})\ e(t) = \\ &= (\theta_0-\theta)^T\ \bar{\varphi}(t,\theta) + C(q^{-1})\ e(t) = \\ &= (\theta_0-\theta_{MV})^T\ \bar{\varphi}(t,\theta) + \bar{\varphi}(t,\theta)^T(\theta_{MV}-\theta) + C(q^{-1})\ e(t) \end{aligned}$$

where θ_{MV} is given by (17).

Hence

$$(\theta_0-\theta_{MV})^T \bar{\varphi}(t,\theta) = \left(1-C(q^{-1})\right)\bar{y}(t,\theta) = \left(1-C(q^{-1})\right)\bar{\varepsilon}(t,\theta)$$

and

$$C(q^{-1}) \bar{\varepsilon}(t,\theta) = \bar{\varphi}(t,\theta)^T (\theta_{MV}-\theta) + C(q^{-1}) e(t)$$

and

$$\bar{\varepsilon}(t,\theta) = \tilde{\varphi}(t,\theta)^T (\theta_{MV}-\theta) + e(t)$$

where

$$\tilde{\varphi}(t,\theta) = H(q^{-1}) \bar{\varphi}(t,\theta) \tag{41}$$

$$H(q^{-1}) = \frac{1}{C(q^{-1})} \tag{42}$$

Introduce

$$\tilde{G}(\theta) = E \bar{\varphi}(t,\theta) \tilde{\varphi}^T(t,\theta) \tag{43}$$

and

$$\theta^* = \theta_{MV} \tag{44}$$

Then, since $e(t)$ is independent of $y(s)$, $u(s)$; $s < t$, we have

$$f(\theta) = E \bar{\varphi}(t,\theta) \bar{\varepsilon}(t,\theta) = \tilde{G}(\theta)(\theta^*-\theta) \tag{45}$$

as before.

In summary then, we find that the algorithm (7) is associated with the differential equation

$$\frac{d}{d\tau} \theta(\tau) = R^{-1}(\tau) \tilde{G}\left(\theta(\tau)\right)\left(\theta^*-\theta(\tau)\right) \tag{46a}$$

$$\frac{d}{d\tau} R(\tau) = G\left(\theta(\tau)\right) - R(\tau) \tag{46b}$$

which is defined for $\theta(\tau) \in D_s$, $R(\tau) > 0$.

In (46)

$$G(\theta) = E\,\bar{\varphi}(t,\theta)\,\bar{\varphi}^T(t,\theta) \tag{47}$$

$$\tilde{G}(\theta) = E\,\bar{\varphi}(t,\theta)\,\tilde{\varphi}^T(t,\theta) \tag{48}$$

where

$$\tilde{\varphi}(t,\theta) = H(q^{-1})\,\bar{\varphi}(t,\theta). \tag{49}$$

This description is common to all three considered schemes, but the interpretations of θ^*, D_s, $\bar{\varphi}(t,\theta)$ and $H(q^{-1})$ depend on the actual scheme as specified above.

Similarly, the stochastic approximation variant (8) is associated with

$$\frac{d}{d\tau}\,\theta(\tau) = \frac{1}{r(\tau)}\,\tilde{G}\big(\theta(\tau)\big)\big(\theta^*-\theta(\tau)\big) \tag{50a}$$

$$\frac{d}{d\tau}\,r(\tau) = g\big(\theta(\tau)\big) - r(\tau) \tag{50b}$$

where

$$g(\theta) = E\,\bar{\varphi}^T(t,\theta)\,\bar{\varphi}(t,\theta). \tag{51}$$

4. THE IMPORTANCE OF POSITIVE REALNESS

In Landau's analysis [1], an assumption that the transfer function

$$\frac{D(z)}{A(z)}$$

is positive real plays a key role. The reason is that hyperstability theory is applied to the system. It is however not quite clear from the analysis if this condition is necessary for convergence.

In this section we shall show the implications of positive realness of the transfer function upon the differential equations discussed in the previous section.

In all the schemes the matrix $\tilde{G}(\theta)$ is given by (48), (49). In (49) $H(q^{-1})$ is a linear filter, which for Landau's method equals $D(q^{-1})/A(q^{-1})$ and for ELS and the self-tuning regulator equals $1/C(q^{-1})$.

We shall now show that the matrix $\tilde{G}(\theta) + \tilde{G}^T(\theta)$ is positive semidefinite if $H(q^{-1})$ is positive real; i.e. if it is stable and

$$\text{Re } H(e^{i\omega}) > 0 \qquad -\pi < \omega \leq \pi \tag{52}$$

To do this consider for a column vector L

$$L^T\left(\tilde{G}(\theta) + \tilde{G}(\theta)^T\right)L = 2\ L^T\ \tilde{G}(\theta)L = 2\ E\ \bar{z}(t,\theta)\ \tilde{z}(t,\theta)$$

where

$$\bar{z}(t,\theta) = L^T\ \bar{\varphi}(t,\theta)$$

and

$$\tilde{z}(t,\theta) = H(q^{-1})\ \bar{z}(t,\theta)$$

are scalars. Now the covariance between $\bar{z}(t,\theta)$ and $\tilde{z}(t,\theta)$

can be given by integration over the unit circle, see Åström [11],

$$E\ \bar{z}(t,\theta)\ \tilde{z}(t,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{i\omega})\ \Phi_{\bar{z}}(\omega)\ d\omega \tag{53}$$

where $\Phi_{\bar{z}}(\omega)$ is the spectral density of the stationary process $\bar{z}(t,\theta)$.

Since $\Phi_{\bar{z}}(\omega)$ and $E\ \bar{z}(t,\theta)\ \tilde{z}(t,\theta)$ are real we have

$$E\ \bar{z}(t,\theta)\ \tilde{z}(t,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Re}\left(H(e^{i\omega})\right)\ \Phi_{\bar{z}}(\omega)\ d\omega \geq 0 \tag{54}$$

if (52) holds. Notice that equality in (54) holds only if $\Phi_{\bar{z}}(\omega) \equiv 0$, i.e. if $\bar{z}(t,\theta) \equiv 0$.

Therefore the matrix $\tilde{G}(\theta) + \tilde{G}^T(\theta)$ is positive semidefinite if $H(q^{-1})$ is positive real.

In the same way it follows that the matrix

$$\tilde{G}(\theta) + \tilde{G}^T(\theta) - G(\theta) \tag{55}$$

is positive semidefinite if the transfer function

$$H(q^{-1}) - \frac{1}{2} \tag{56}$$

is positive real.

This is seen as follows:

$$L^T\left(\tilde{G}(\theta) + \tilde{G}^T(\theta) - G(\theta)\right)L = 2E\ \bar{z}(t,\theta)\left(\tilde{z}(t,\theta) - \frac{1}{2}\ \bar{z}(t,\theta)\right) =$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{\bar{z}}(\omega)\left(H(e^{i\omega}) - \frac{1}{2}\right) d\omega$$

In Section 5 the eigenvalues of matrices of the type

$$S\ \tilde{G}(\theta)$$

will be considered. Then the following result is useful

- o If S is positive definite symmetric and $\tilde{G} + \tilde{G}^T$ is positive semidefinite, then $S\tilde{G}$ has all eigenvalues in the right half plane (including the imaginary axis). (57)

A simple way to realize this is to consider the linear differential equation $\dot{X} = -S\tilde{G}X$. Since for $P = S^{-1}$

$$P(-S\tilde{G}) + (-S\tilde{G})^T P = -(\tilde{G} + \tilde{G}^T) \leq 0$$

it follows from the well known stability criterion for linear systems that it is stable, which implies that $S\tilde{G}$ has all its eigenvalues in the right half plane.

In this context it may be of interest to characterize positive real, stable polynomials $C(q^{-1})$. It is easy to show by straightforward calculations that

- o For all stable first order polynomials

  $$C(q^{-1}) = 1 + cq^{-1}, \qquad |c| < 1$$

  it is true that

  $$\frac{1}{C(q^{-1})} \quad \text{and} \quad \frac{1}{C(q^{-1})} - \frac{1}{2}$$

  are positive real.

- o For stable second order polynomials

  $$C(q^{-1}) = 1 + c_1 q^{-1} + c_2 q^{-2}$$

  $\frac{1}{C(q^{-1})}$ is positive real for c_1, c_2 as shown in Figure 1 and $\frac{1}{C(q^{-1})} - \frac{1}{2}$ is positive real for c_1, c_2 as shown in Figure 2.

This means that for low order systems, the assumption of positive realness is not as restrictive as it may seem.

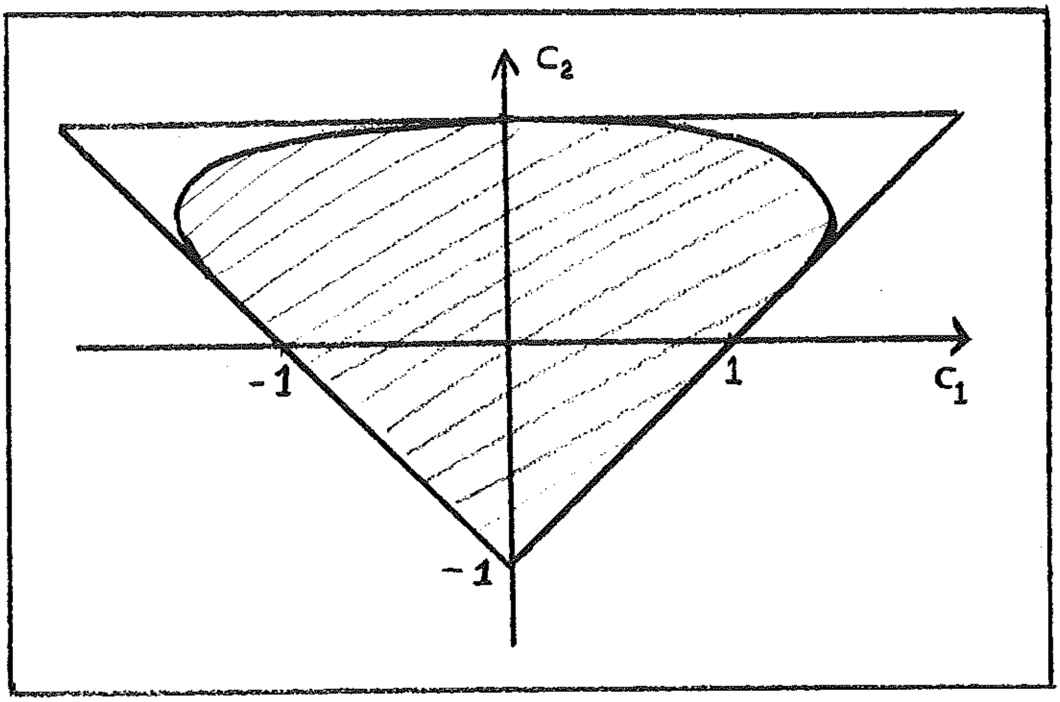

Figure 1. Triangle: stability region for second order filters $C(q^{-1})$. Shaded area: Region such that $\frac{1}{C(q^{-1})}$ is positive real.

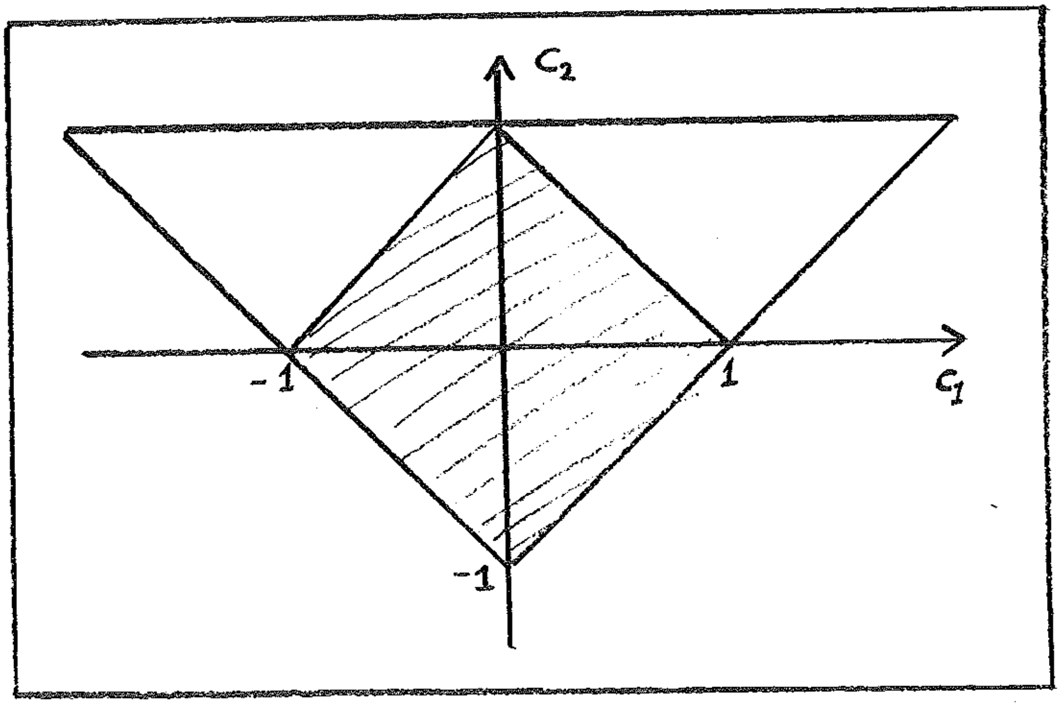

Figure 2. As Figure 1, but shaded area is region such that $\frac{1}{C(q^{-1})} - \frac{1}{2}$ is positive real.

5. STATIONARY POINTS AND LOCAL CONVERGENCE

From Theorem 2 of [2] it follows that under certain regularity conditions, only stable, stationary points of the differential equation (46) ((50)) are possible convergence points of the algorithm (7) ((8)).

Let us first consider the stationary points of (46) and (50). They are given by

$$\frac{d}{d\tau}\,\theta = R^{-1}\,f(\theta) = 0$$

$$\frac{d}{d\tau}\,R = G(\theta) - R = 0$$

i.e.

$$f(\theta) = \tilde{G}(\theta)(\theta^{*}-\theta) = 0 \tag{58}$$

and

$$R = G(\theta) \tag{59}$$

where $\tilde{G}(\theta)$ and θ^{*} are defined as before. The stationary points of R (r) therefore follow trivially from those of θ and we may consequently concentrate on eq. (58).

Clearly $\theta = \theta^{*}$ is always a solution, "the desired convergence point". Assume now that the filter $H(q^{-1})$ in (49) is positive real, i.e. (52) holds. In the previous section it was shown that

$$\tilde{G}(\theta)L = 0$$

implies that

$$\bar{z}(t,\theta) = L^{T}\,\bar{\varphi}(t,\theta) \equiv 0$$

Therefore the stationary points $\bar{\theta}$ of (58) are characterized by

$$\theta^{*T}\,\bar{\varphi}(t,\bar{\theta}) \equiv \bar{\theta}^{T}\,\bar{\varphi}(t,\bar{\theta}) \tag{60}$$

We shall now show that (60) implies that all stationary points corresponds to "good" estimates; i.e. parameters that give a correct input-output description of the system in the Landau and ELS cases and a minimum variance controller in the self-tuning regulator case.

Landau's Scheme

We have

$$\bar{y}_M(t,\bar{\theta}) = \bar{\theta}^T \bar{\varphi}(t,\bar{\theta}) = \theta_0^T \bar{\varphi}(t,\bar{\theta})$$

$$\bar{y}_M(t,\theta_0) = \theta_0^T \bar{\varphi}(t,\theta_0)$$

Now, with

$$\Delta\bar{y}_M(t) \triangleq \bar{y}_M(t,\bar{\theta}) - \bar{y}_M(t,\theta_0)$$

we have

$$\bar{\varphi}(t,\bar{\theta}) - \bar{\varphi}(t,\theta_0) = \left(-\Delta\bar{y}_M(t-1),\ldots,-\Delta\bar{y}_M(t-n), 0,\ldots,0\right)^T$$

and

$$\Delta\bar{y}_M(t) = \theta_0^T\left(\bar{\varphi}(t,\bar{\theta}) - \bar{\varphi}(t,\theta_0)\right) = (1-A(q^{-1}))\ \Delta\bar{y}_M(t)$$

or

$$A(q^{-1})\ \Delta\bar{y}_M(t) = 0$$

Since A(·) is a stable filter this implies that

$$\bar{y}_M(t,\bar{\theta}) \equiv \bar{y}_M(t,\theta_0) \qquad (61)$$

i.e. the estimate $\bar{\theta}$ gives a description of the system that is equivalent to true system from an input-output point of view. Moreover, if the order of the model does not exceed that of the true system, (61) implies that

$$\bar{\theta} = \theta_0$$

since the input is persistently exciting.

Remark. We have in our notation implicitly assumed the same "order", i.e. the same number of parameters in the true system and in the model. With this convention, the requirement on the system is that it be observable and controllable, i.e. that the A- and B-polynomials do not have common factors.

Therefore, in this case the true parameter value θ_0 is the only stationary point of (46) or (50).

The ELS-Scheme

We have from (32) and (60)

$$\bar{\varepsilon}(t,\bar{\theta}) = y(t) - \bar{\theta}^T \varphi(t,\bar{\theta}) = y(t) - \theta_0^T \bar{\varphi}(t,\bar{\theta})$$

$$\bar{\varepsilon}(t,\theta_0) = e(t) = y(t) - \theta_0^T \bar{\varphi}(t,\theta_0) = y(t) - \theta_0^T \varphi_0(t)$$

and analogously to the previous analysis

$$\varepsilon(t,\bar{\theta}) - e(t) = \left(1-C(q^{-1})\right)\left(\varepsilon(t,\bar{\theta}) - e(t)\right)$$

or

$$C(q^{-1}) \left(\varepsilon(t,\bar{\theta}) - e(t)\right) = 0$$

and

$$\varepsilon(t,\bar{\theta}) \equiv e(t) \tag{62}$$

This implies that the model corresponding to $\bar{\theta}$ has equivalent input-output behaviour to the true system. Moreover, if the order of the model does not exceed that of the true system, i.e. there is no factor common to A(z), B(z) and C(z) in (9), then (62) implies that

$$\bar{\theta} = \theta_0$$

as above.

The Self-tuning Regulator

We have from (39) and (60)

$$\bar{y}(t,\bar{\theta}) = -\bar{\theta}^T \bar{\varphi}(t,\bar{\theta}) + \theta_0^T \bar{\varphi}(t,\bar{\theta}) + C(q^{-1})\ e(t) =$$

$$= (-\theta^*+\theta_0)^T \bar{\varphi}(t,\bar{\theta}) + C(q^{-1})\ e(t)$$

and, since

$$(\theta_0-\theta^*)^T = (c_1,\ldots,c_n,\ 0,\ldots,0)$$

we have

$$C(q^{-1})\ \bar{y}(t,\bar{\theta}) = C(q^{-1})\ e(t)$$

or

$$\bar{y}(t,\bar{\theta}) = e(t) \tag{63}$$

so that $\bar{\theta}$ gives minimum variance control of the process. Again, if the model order is not overestimated, (63) implies

$$\bar{\theta} = \theta^*$$

so a unique stationary point is obtained. This latter result is true regardless of the properties of $C(q^{-1})$, see [8].

Local convergence

We shall now proceed to investigate the stability properties of the linearized equation around θ^*. Let us assume that the respective filter $H(q^{-1})$ obeys (52) and that the model orders are not overestimated so that θ^* is a unique stationary point. This implies in particular that $G(\theta^*)$ is invertible and strictly positive definite.

It is immediate to linearize (46) around $\theta=\theta^*$, $R=G(\theta^*)$:

$$\frac{d}{d\tau}(\theta-\theta^*) = -\ G(\theta^*)^{-1}\ \tilde{G}(\theta^*)(\theta-\theta^*) \tag{64a}$$

$$\frac{d}{d\tau}\big(R-G(\theta^*)\big) = \frac{d}{d\theta}G(\theta)\bigg|_{\theta=\theta^*}(\theta-\theta^*) - \big(R-G(\theta^*)\big) \tag{64b}$$

Clearly (64) is stable if the eigenvalues of

$$G(\theta^*)^{-1}\,\tilde{G}(\theta^*)$$

are in the right half plane. This is however the case if (52) holds, due to (57). Similarly, linearizing (50) gives

$$\frac{d}{d\tau}\,(\theta-\theta^*) = -\,\frac{1}{g(\theta^*)}\,\tilde{G}(\theta^*)(\theta-\theta^*) \tag{65a}$$

$$\frac{d}{d\tau}\,\big(r-g(\theta^*)\big) = \frac{d}{d\theta}\,g(\theta)\Big|_{\theta=\theta^*}(\theta-\theta^*) - \big(r-g(\theta^*)\big) \tag{65b}$$

which also is stable under the assumption (52). However, if $H(q^{-1})$ is not positive real, that is $\mathrm{Re}\,H(e^{i\omega_0}) < 0$ for some ω_0, then certain choices of system and input signal characteristics can always make the signal $\bar{z}(t,\theta^*)$ have a sharp resonance at the frequency ω_0. Then $\mathrm{tr}\,\tilde{G}(\theta^*)$ can be made negative and (65) becomes unstable. Consequently the estimates will then not, w.p.1, converge to the desired limit. This method was demonstrated in [7] for the ELS-case, where the A-polynomial is given a resonance at ω_0, and in [12] for the self-tuning regulator, where the B-polynomial is given a resonance at ω_0. For Landau's scheme, e.g. the input can be chosen so that it has a frequence peak at ω_0 to yield nonconvergence of the scheme.

A more thourough analysis of $G(\theta^*)^{-1}\,\tilde{G}(\theta^*)$ is given by Holst [13], where explicit expression for its eigenvalues are given. In particular, systems and choices of $D(q^{-1})$ for which Landau's scheme does not converge can easily be generated in that way.

6. GLOBAL CONVERGENCE

From Theorem 1 of [2] it follows that, under certain regularity conditions, if

1) $\theta(t)$ belongs to a compact subset of D_s infinitely often w.p.1 and $|\varphi(t)|$ for those t is bounded

2) and the differential equation (46) $\left((50)\right)$ has a stationary point (θ^*, R^*), $\left((\theta^*, r^*)\right)$ with global domain of attraction

then for $\theta(t)$ defined by (7) $\left((8)\right)$ $\theta(t) \to \theta^*$ w.p.1 (and $R(t) \to R^*$ w.p.1) as $t \to \infty$.

Moreover, if the trajectories of (46) tend into a set $\bar{D}$, then it follows that the estimates of the corresponding algorithm converge w.p.1 into $\bar{D}$.

The regularity conditions in question are verified for the algorithm-structure (7) in Example 3 of [2].

Let us first comment somewhat on condition 1). As remarked before, when $\theta(t)$ does not belong to D_s the filter that defines $\varphi(t)$ is unstable. Therefore some mechanizm is required that keeps $\theta(t)$ inside D_s. For Landau's scheme and ELS this is easy to do since D_s is a known area, and it can be checked that the estimated A-polynomial (C-polynomial) is stable, using, e.g., a Routh's scheme. If not, the stepsize in the algorithm is reduced until a stable polynomial is obtained. Such a modification is usually introduced in the algorithm also in practice to avoid "explosion". Therefore (see also Thm 4 of [2]) $\theta(t)$ is assured to belong to D_s. However it could happen that this procedure introduces a cluster point of $\{\theta(t)\}$ on the boundary of D_s. This is the case if the trajectories of (46) $\left((50)\right)$ point out of D_s at the boundary of D_s. However, a cluster point on the boundary of D_s is easy to detect and undesirable, and we shall assume that some measure is taken in the algorithm to avoid such an effect.

For the self-tuning regulator the area D_s is not known to the user, and the described projection procedure cannot be used. However, it follows from [14] that the regulator possesses a stabilizing feature which assures that 1) holds. Also in this case we shall assume that the trajectories of (46) ((50)) do not point out of D_s at the boundary.

It could be remarked that a similar stabilizing feature has been shown also for the ELS-scheme [16]. This means that the estimated C-polynomial will be forced back into D_s. It also, as for the self-tuning regulator, gives a strong indication that indeed the trajectories at the boundary of D_s point into D_s.

Consider first the differential equation (50):

$$\begin{aligned} \frac{d}{d\tau}\,\theta(\tau) &= \frac{1}{r(\tau)}\,\tilde{G}\big(\theta(\tau)\big)\big(\theta^{*}-\theta(\tau)\big) \\ \frac{d}{d\tau}\,r(\tau) &= g\big(\theta(\tau)\big) - r(\tau) \end{aligned} \tag{66}$$

where the definition of θ^{*} and $\tilde{G}$ are given in Section 3 for the different methods.

We shall prove stability of (66) using the Lyapunov function

$$V(\theta,r) = |\theta-\theta^{*}|^{2}$$

We have

$$\frac{d}{d\tau}\,V\big(\theta(\tau),r(\tau)\big) = -\frac{1}{r(\tau)}\,\big(\theta(\tau)-\theta^{*}\big)^{T}\Big(\tilde{G}\big(\theta(\tau)\big)+\tilde{G}^{T}\big(\theta(\tau)\big)\Big)\cdot$$

$$\cdot\,\big(\theta(\tau)-\theta^{*}\big) \le 0 \tag{67}$$

if $H(q^{-1})$ is positive real, according to Section 4. Equality in (67) holds only for θ such that $f(\theta)=0$, which proves that all solutions to (66) converge into this set. If the order of the model is such that only θ^{*} is a solution of $f(\theta)=0$, then consequently θ^{*} is a globally asymptotically stable solution of (66).

This implies that the Stochastic Approximation variant of Landau's scheme, the ELS-method and the self-tuning regulator converge w.p.1 to desired limits if $D(z)/A(z)$, $C(z)$ and $C(z)$ respectively are positive real transfer functions. If the model order does not exceed that of the true system, then strong convergence to θ_0, θ_0 and θ_{MV}, respectively, follows.

Let us now turn to the variant (7). The corresponding differential equation is

$$\frac{d}{d\tau}\,\theta(\tau) = R(\tau)^{-1}\,\tilde{G}\big(\theta(\tau)\big)\,\big(\theta^*-\theta(\tau)\big)$$
$$\frac{d}{d\tau}\,R(\tau) = G\big(\theta(\tau)\big) - R(\tau) \tag{68}$$

For this differential equation we try the Lyapunov function

$$V(\theta,R) = (\theta-\theta^*)^T\,R(\theta-\theta^*)$$

and we have

$$\frac{d}{d\tau}\,V\big(\theta(\tau),R(\tau)\big) = -\big(\theta(\tau)-\theta^*\big)^T\Big(\tilde{G}\big(\theta(\tau)\big) + \tilde{G}^T\big(\theta(\tau)\big) - G\big(\theta(\tau)\big)\Big)\cdot$$
$$\cdot\big(\theta(\tau)-\theta^*\big)$$

This expression is non-positive if the transfer function

$$H(q^{-1}) - \frac{1}{2}$$

is positive real as shown in Section 4.

This is a slightly stronger condition than for the stochastic approximation variant. However, there are several reasons to believe that by a more clever choice of Lyapunov function it will be possible to reduce this condition to the condition (52). For example, if $R(t)$ in (7a) is replaced by any positive definite constant matrix it would be sufficient to require (52). Moreover, if the stepsize is increased so that $R(t)$ is replaced by $\alpha R(t)$ in (7a) then it is sufficient that $H(q^{-1})-\frac{1}{2\alpha}$ be positive real.

In summary then we have shown convergence w.p.1 of the schemes under discussion, variant (7), to the desired limit if

$$D(z)/A(z) - 1/2 \quad \text{(Landau)}$$

$$1/C(z) - 1/2 \quad \text{(ELS and self-tuning regulator)}$$

are positive real.

7. SUMMARY AND CONCLUDING REMARKS

Let us first summarize the convergence results obtained in Sections 5 and 6.

Landau's scheme

Consider the system given by (1).
Assume that

(i) $\{u(t)\}$ and $\{w(t)\}$ are stationary stochastic processes with rational spectral densities and such that all moments exist.

(ii) The system operates in open loop, i.e. that $\{u(t)\}$ and $\{w(t)\}$ are independent.

(iii) The identification scheme defined by (3)-(7) (or the stochastic approximation variant (8)) is used together with a projection feature to keep $\hat{\theta}(t)$ in a compact subset of D_s, such that $\{\hat{\theta}(t)\}$ does not have a cluster-point on the boundary of this subset. (Such a feature is usually included in practical implementations.)

(iv) The filter $A(q^{-1})$ is stable.

(v) The filter $D(q^{-1})$ is chosen so that the transfer function $D(q^{-1})/A(q^{-1})$ is positive real.

Then,

(a) The stochastic approximation variant (8) gives estimates that converge w.p.1 to a true input-output description of the system (1).

(b) If the model order is not overestimated, i.e. the polynomials $A(z)$ and $B(z)$ do not have common factors, then it follows that $\hat{\theta}(t) \to \theta_0$ w.p.1 as $t \to \infty$.

(c) The convergence results a) and b) have been proven also for the variant (7) under the strengthend condition that $D(q^{-1})/A(q^{-1}) - 1/2$ be positive real.

(d) The true value Θ_0 is always a possible convergence point of the algorithm (both (7) and (8)) in the sense that the corresponding linearized differential equation is stable.

(e) If (v) does not hold there exists an input process {u(t)}, such that the procedure (8) (and (7)) will not, w.p.1 converge to Θ_0, nor to any other value.

The Extended Least Squares Scheme

Consider the system given by (9).
Assume that

(i) {e(t)} is stationary sequence of independent random variables and {u(t)} is a stationary stochastic process with rational spectral density. All moments of e(t) and u(t) are assumed to exist.

(ii) u(t) is independent of e(s); $s > t$, but may be (partly) obtained as linear output feedback by a regulator, independent of the parameter estimates (no adaptive control).

(iii) The identification scheme defined by (12), (7) (or (8)) is used together with a feature to keep the estimated C-polynomial stable as explained under (iii) above.

(iv) The closed loop system is stable.

(v) The filter $C(q^{-1})$ is positive real.

Then, the conclusions (a)(b)(c)(d) and (e) hold exactly as for Landau's scheme. The requirement under (b) is that there is no factor common to all the polynomials A(z), B(z) and C(z). The strengthening under (c) is that $1/C(q^{-1}) - 1/2$ be positive real, and the non-convergence described under (e) can also be obtained for suitable chosen A-polynomials even for the case with no input, see [7].

The Self-tuning Regulator

Consider the system given by (9) controlled by the adaptive regulator defined by (14), (15) and (7) (or (8)).
Assume that

(i) $\{e(t)\}$ is a stationary sequence of independent random variables.

(ii) The system is minimum phase, i.e. $B(z) = 0 \Rightarrow |z| > 1$.

(iii) The trajectories of the associated differential equation do not leave the area D_s given by (40).

(iv) The filter $C(q^{-1})$ is positive real.

Then,

(a) The stochastic approximation variant (8) gives a controller that converges w.p.1 to the minimum variance control law.

(b) If the model order is not overestimated, i.e. if there is no factor common to $A(z)$,$B(z)$ and $C(z)$, then $\hat{\theta}(t) \rightarrow \theta_{MV}$ w.p.1 as $t \rightarrow \infty$.

(c) The convergence results (a) and (b) hold also for the variant (7) if the transfer function $1/C(q^{-1}) - 1/2$ is positive real.

(d) The value θ_{MV} corresponding to minimum variance control is always a possible convergence point (both for (8) and (7)).

(e) If (iv) does not hold, then there exists A- and B--polynomials such that $\hat{\theta}(t)$ produced by (7) or (8) cannot, w.p.1, converge to θ_{MV} nor to any other value.

All these convergence results are new, and they give, together with the analysis in [13], a fairly complete picture of the convergence problems.

In addition, there should be some interest in the manner in which the property of positive realness enters here compared to Landau's analysis.

It is interesting to note that this property is not only a formal trick to make the proof go through. While convergence certainly may take place even if the corresponding filter is not positive real, it has been demonstrated under point (e) above, that as a condition on D/A or C alone it is necessary for convergence.

Therefore the positive-real property is something inherently tied to the convergence behaviour of the algorithms. This is an important conclusion drawn from the analysis here, but the observation of this (maybe surprising) fact is of course due to Landau.

Another point worthy of attention is that hyperstability theory and hence positive realness has been widely used in connection with model reference adaptive systems, as a complement to Lyapunov design, see e.g. [15]. A perhaps unexpected connection has thus here been established between model-reference adaptive regulators and the self-tuning regulator, which has been designed from a completely different viewpoint.

The analysis gives a fairly good understanding of the convergence "mechanism". It was demonstrated in [7] how the analysis gave direct ideas to a scheme related to ELS with improved convergence. Here we may note that Landau's scheme can be changed so that in (6), the d-parameters are taken as the current a-estimates: $d_i = \hat{a}_i(t-1)$. Then the $H(q^{-1})$-filter in (28) will of course depend on Θ: $H(q^{-1},\Theta)$, and $H(q^{-1},\Theta_0) = 1$. This in turn implies that $\tilde{G}(\Theta_0) = G(\Theta_0)$ and that Θ_0 is a stable stationary point both to (46) and (50). Therefore this variant of Landau's scheme always yields local convergence without any a priori knowledge necessary to allow for a good choice of D.

We may finally remark that, according to [2], all what has been said here also holds for algorithms like (7) and (8) where $1/t$ is replaced by more general sequences $\{\gamma(t)\}$.

REFERENCES

[1] I D Landau: "Unbiased Recursive Identification Using Model Reference Adaptive Techniques", IEEE Trans Automat Contr Vol AC-21, pp 194-202, April 1976.

[2] L Ljung: "Analysis of Recursive Stochastic Algorithms", submitted to IEEE Trans Automat Contr, 1976.

[3] V Panuska: "An Adaptive Recursive Least Squares Identification Algorithm", Proc 8th IEEE Symp Adaptive Processes, 1969.

[4] P C Young, S H Shellswell, and C G Neethling: "A Recursive Approach to Time-Series Analysis", Dept Eng, Univ Cambridge, Cambridge, England, Tech Note CN/70/1, 1970.

[5] J L Talmon and A J W van den Boom: "On the Estimation of Transfer Function Parameters of Process- and Noise Dynamics Using a Single-stage Estimator", in Proc 3rd Int Fed Automat Contr Symp Identification and System Parameter Estimation, The Hague, The Netherlands, 1973.

[6] R L Kashyap: "Estimation of Parameters in a Partially Whitened Representation of a Stochastic Process", IEEE Trans Automat Contr Vol AC-19, pp 13-21, Feb 1974.

[7] L Ljung, T Söderström, and I Gustavsson: "Counterexamples to General Convergence of a Commonly Used Recursive Identification Method", IEEE Trans Automat Contr Vol AC-20, pp 643-652, Oct 1975.

[8] K J Åström and B Wittenmark: "On Self-tuning Regulators", Automatica Vol 9, pp 185-199, 1973.

[9] L Ljung and B Wittenmark: "Analysis of a Class of Adaptive Regulators", in Proc Int Fed Automat Contr Symp. Stochastic Control, Budapest, Hungary, Sept 1974.

[10] K J Åström, U Borisson, L Ljung, and B Wittenmark: "Theory and Applications of Adaptive Regulators Based on Recursive Identification", in Proc 6th Int Fed Automat Contr World Congr, Boston, Mass, 1975.

[11] K J Åström: "Introduction to Stochastic Control Theory", New York: Academic Press, 1970.

[12] L Ljung and B Wittenmark: "Asymptotic Properties of Self-tuning Regulators", Report 7407, Dept Automat Contr, Lund Inst Tech, Lund, Sweden, Feb 1974.

[13] J Holst: "Local Convergence for some Recursive Estimators", in preparation.

[14] L Ljung and B Wittenmark: "On a Stabilizing Property of Adaptive Regulators", in Preprints Int Fed Automat Contr Symp Identification, Tbilisi, USSR, Sept 1976.

[15] I D Landau: "A Survey of Model Reference Adaptive Techniques - Theory and Applications", Automatica Vol 10, July 1974.

[16] C Doncarli: "A Real Time Extended Least Squares Method", in preparation.