Automatica, Vol. 18, No. 1, pp. 1–13, 1982
Printed in Great Britain

0005-1098/82/010001–13$03.00/0
Pergamon Press Ltd.
© 1982 International Federation of Automatic Control

Exponential Convergence of Adaptive Identification and Control Algorithms*

BRIAN D. O. ANDERSON† and C. RICHARD JOHNSON, JR‡

The exponential convergence rate of output and equation error formulated adaptive identifiers and adaptive inverse control is proven with sufficient excitation conditions only on the input and reference signal, respectively.

Key Words—Adaptive control; adaptive systems; convergence of numerical methods; filtering; identification; parameter estimation; stability.

Abstract—Output and equation error adaptive identification algorithms are shown to be exponentially convergent under a deterministic or stochastic persistently exciting (or spanning) condition on the system inputs together with several other standard conditions. An adaptive control algorithm is shown to be exponentially convergent under a deterministic or stochastic persistently exciting condition on the reference trajectory together with some standard conditions.

1. INTRODUCTION

THIS PAPER is concerned with establishing the exponential convergence of certain algorithms used in adaptive identification and control. At the outset, one can reasonably raise the question, why bother establishing exponential convergence, as opposed simply to convergence? The main reason is that systems which are exponentially convergent are generally systems which are tolerant of modelling errors, noise, and various inadequacies of realization. Furthermore, if the real reason for using adaptive identification or control is to cope with a slowly-varying plant, it is highly doubtful that a rate of convergence slower than exponential in the stationary plant case will translate to satisfactory behavior in the nonstationary case.

The earliest contributions to the study of exponential convergence of adaptive parameter estimators (Lion, 1967) required signals to be periodic, so that Lyapunov theorems on periodic linear systems could be appealed to. Through misunderstanding of these Lyapunov theorems, a number of authors later asserted that when the system signals comprised linear combinations of sinusoids, exponential stability could be claimed. The Lion result was correctly extended to this almost periodic case, for the first time it is believed, in Anderson (1974a, b). Anderson (1974a) applied a new and powerful tool to the stability analysis which had been developed in Anderson and Moore (1969)—a time-varying Lyapunov lemma which can be used to deduce an exponential stability result precisely when a uniform observability property holds. The proof in Anderson (1974a) contained a proof of exponential stability under even more general conditions than apply in the almost periodic case. The alternative derivation of Morgan and Narendra (1977a) of the same result offers many insightful examples and proves the necessity and sufficiency of the exponential stability condition rather than just sufficiency; on the other hand, it is lengthy, due to the fact that the general tools of Anderson and Moore (1969) are not appealed to, but rather sometimes rederived in highly specific form. Anderson (1977a) develops the results of Anderson (1974a) and includes the main results of Morgan and Narendra (1977a) and also Morgan and Narendra (1977b). The application of Anderson (1977a) to problems of adaptive identification is described in Anderson (1977b). Sondhi and Mitra (1976) and Weiss and Mitra (1979), written more from the viewpoint of the adaptive equalization than the control literature, establish exponential convergence results, provide valuable information on convergence rates, and explain the practical importance of convergence that is specifically exponential. Kreisselmeier (1977) obtains independently and in another way

*Received 30 June 1980; revised 6 February 1981. The original version of this paper was presented at the 8th IFAC World Congress on Control Science and Technology for the Progress of Society, which was held in Kyoto, Japan during August 1981. The published proceedings of this IFAC meeting may be ordered from: Pergamon Press Ltd, Headington Hill Hall, Oxford OX3 0BW, U.K. This paper was recommended for publication in revised form by associate editor I. D. Landau. This work was supported by the Australian Research Grants Committee and National Science Foundation Grant ECS-79-08695.

†Department of Electrical Engineering, University of Newcastle, New South Wales, 2308, Australia.

‡Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, U.S.A.

a number of the results of Anderson (1974a, b, 1977a, b) and Morgan and Narendra (1977a, b). Yuan and Wonham (1977) discuss the design of system input signals to ensure exponential convergence. Virtually all of the above references deal with adaptive situations where the plant is excited by deterministic signals. New techniques are required to handle the presence of random signals (e.g. Jones, 1973; Farden, Goding and Saygood, 1979; Bitmead and Anderson, 1978; 1980a, b).

The novelty of this paper arises on two distinct grounds. First, the paper presents results for the exponential stability of an *output* error identification algorithm and of an adaptive control algorithm. Second, the conditions are free of a deficiency in the 'persistent excitation' conditions in most but not all of the references cited above. Consider the discrete-time identification problem. Let $\{u_k\}$ and $\{y_k\}$ be the input and output sequence of the unknown plant and $\{z_k\}$ the output of the adaptive identifier. Then typical conditions for exponential stability are given (e.g. Bitmead and Anderson, 1980b), in terms of a vector $x_k = [u_{k-1} \quad u_{k-2} \cdots u_{k-m} \; z_{k-1} \cdots z_{k-n}]'$ for some m, n. Of course, a condition in terms of $\{u_k\}$ only is desired. (For the adaptive control problem, a desirable condition is one on the reference trajectory which the plant output is supposed to track.) This paper explains how such a condition can be obtained. Previous contributions have obtained such a condition on the input alone only when equation error identification has been considered, for which x_k is replaced by

$$x_k = [u_{k-1} u_{k-2} \cdots u_{k-m} y_{k-1} \cdots y_{k-n}]$$

and $\{u_k\}$ has been almost periodic. For adaptive identification (or control), the main requirement for exponential convergence is shown to be a persistently exciting or spanning condition on the input (or reference) trajectory.

2. REFORMULATING PERSISTENT EXCITATION CONDITIONS

This section provides tools allowing conversion of 'persistent excitation' conditions involving input and output quantities to ones involving, for the identification problem, input quantities alone, and for the adaptive control problem, output quanties alone.

Lemma 2.1. Consider the single-input, single-output, time-invariant system

$$y_k + \alpha_1 y_{k-1} + \cdots + \alpha_n y_{k-n} = \beta_1 u_{k-1} + \cdots + \beta_m u_{k-m} \tag{1}$$

and assume that $z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n$ and $\beta_1 z^{n-1} + \cdots + \beta_m z^{n-m}$ are coprime polynomials. Suppose further that with

$$x_k = [y_k \; y_{k-1} \cdots y_{k-n+1} \; u_k \; u_{k-1} \cdots u_{k-m+1}]' \tag{2}$$

the matrix $\sum_{k=j}^{j+S} x_k x_k'$ is singular for some j and S. Then

(a) there exists a nonzero $(n+m)$-vector θ such that

$$\theta'[u_{k+n} u_{k+n-1} \cdots u_{k-m+1}]' = 0 \tag{3a}$$

for $k \in [j, j+S-n+1]$; and

(b) there exists a nonzero $(n+m)$-vector $\bar{\theta}$ such that

$$\bar{\theta}'[y_{k+m} y_{k+m-1} \cdots y_{k-n+1}]' = 0 \tag{3b}$$

for $k \in [j, j+S-m+1]$. In case $\beta_1 = \beta_2 = \cdots \beta_{d-1} = 0$, $\beta_d \neq 0$, then (3b) holds for $k \in [j, j+S-m+d]$.

The proof is contained in Appendix A. □

Equation (3a) is relevant for identification problems; it is trivial if the number of equalities, viz. $S-n+2$, is smaller than the dimensions of the θ vector, viz. $(n+m)$. Thus for the result to be of interest, we require

$$S \geq 2n + m - 2. \tag{4}$$

Using Lemma 2.1, we obtain results appropriate for systems with $\{u_k\}$ deterministic or appropriate with $\{u_k\}$ stochastic.

Theorem 2.2. Consider the single-input, single-output system (1), assume that $\alpha(z) = z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n$ and $\beta(z) = \beta_1 z^{n-1} + \cdots + \beta_m z^{n-m}$ are coprime and that $\{y_k\}$ is a bounded sequence. Then with x_k as in (2) and S any integer with $S \geq 2n + m - 2$

$$\rho_2 I > \sum_j^{j+S} x_k x_k' > \rho_1 I > 0 \tag{5}$$

if

$$\rho_4 I > \sum_j^{j+S-n+1} \begin{bmatrix} u_{k+n} \\ \vdots \\ u_{k-m+1} \end{bmatrix} [u_{k+n} \cdots u_{k-m+1}] > \rho_3 I > 0 \tag{6}$$

for some ρ_3, ρ_4 and all j.

The proof is contained in Appendix A. □

Remarks. (a) The upper bound in (6) ensures $\{u_k\}$ is bounded, and if $\alpha(z)$ has all roots in $|z| < 1$, $\{y_k\}$ is bounded. However, it is conceivable that the system be unstable, but located in a stable closed-loop. (b) Suppose $\{u_k\}$ consists of a sum of ρ distinct sinusoidal terms and possibly a constant component; then $\{u_k\}$ also satisfies a difference equation, of order $q = 2p+1$ or $2p$, according to whether the constant term is present or not. If $q \geq n+m$, it is not hard to check that (6) must hold. So if there are an appropriate number of sinusoids present in $\{u_k\}$, (5) holds with $S = 2n+m-2$. This type of condition arises frequently in the literature (see e.g. Lion, 1967; Kreisselmeier, 1977; Yuan and Wonham, 1977). For stochastic systems, we have the following theorem.

Theorem 2.3. Consider the single-input, single-output system (1), assume that $\alpha(z) = z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n$ and $\beta(z) = \beta_1 z^{n-1} + \cdots + \beta_m z^{n-m}$ are coprime, and that $\{u_k\}$ and $\{y_k\}$ are stochastic with $E[y_k^2]$ bounded. Then with x_k as in (2), for all j, and some $S \geq n-1$ there exists $\rho_1 > 0$, $\rho_2 > 0$ such that

$$\rho_2 I > \sum_j^{j+S} E[x_k x_k'] > \rho_1 I \tag{7}$$

provided that

$$\rho_4 I > \sum_j^{j+S-n+1} E\left\{\begin{bmatrix} u_{k+m} \\ \vdots \\ u_k \\ \vdots \\ u_{k-m+1} \end{bmatrix} [u_{k+n} \cdots u_{k-m+1}]\right\} > \rho_3 I \tag{8}$$

for some positive ρ_3, ρ_4.

The proof is contained in Appendix A. □

Remark. A variant is that

$$\rho_2 I > \liminf_{q\to\infty} \frac{1}{q} \sum_0^q E[x_k x_k'] > \rho_1 I \tag{9}$$

provided that

$$\rho_4 I > \liminf_{q\to\infty} \frac{1}{q} \sum_0^q E\left\{\begin{bmatrix} u_{k+n} \\ \vdots \\ u_k \\ \vdots \\ u_{k-m+1} \end{bmatrix} [u_{k+n} \cdots u_{k-m+1}]\right\} > \rho_3 I. \tag{10}$$

Finally, we state a result relevant to adaptive control. The result comes from the second conclusion of Lemma 2.1 in the same manner as Theorems 2.2 and 2.3 come from the first conclusion and accordingly is stated without proof.

Theorem 2.4. Consider the single-input, single-output system (1) and assume that $\alpha(z) = z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n$ and $\beta(z) = \beta_1 z^{n-1} + \cdots + \beta_m z^{n-m}$ are coprime. Let x_k be as in (2). (a) Assume that $\{u_k\}$ is a bounded deterministic sequence and let S be any integer $\geq 2m+n-2$. Then

$$\rho_2 I > \sum_j^{j+S} x_k x_k' > \rho_1 I > 0 \tag{5}$$

for all j and some ρ_1, ρ_2 if

$$\rho_4 I > \sum_j^{j+S-m+1} \begin{bmatrix} y_{k+m} \\ \vdots \\ y_{k-n+1} \end{bmatrix} [y_{k+m} \cdots y_{k-n+1}] > \rho_3 I > 0 \tag{11}$$

for some ρ_3, ρ_4 and all j. In case it is known that $\beta_1 = \beta_2 = \cdots = \beta_{d-1} = 0$ and $\beta_d \neq 0$, the result holds for $S \geq 2m+n-1-d$. (b) Assume that $\{u_k\}$ is stochastic with $E[u_k^2]$ bounded and let S be any integer $\geq (m-1)$. Then

$$\rho_2 I > \sum_j^{j+S} E[x_k x_k'] > \rho_1 I > 0 \tag{7}$$

for all j and some ρ_1, ρ_2 if

$$\rho_4 I > \sum_j^{j+S-m+1} E\left\{\begin{bmatrix} y_{k+m} \\ \vdots \\ y_{k-n+1} \end{bmatrix} [y_{k+m} \cdots y_{k-n+1}]\right\} > \rho_3 I \tag{8}$$

for some ρ_3, ρ_4 and all j. □

3. OUTPUT AND EQUATION ERROR IDENTIFICATION

We shall begin by studying output error *identification* (Landau, 1976, 1978; Dugard and Landau, 1980). We state the basic exponential convergence problem and summarize the solution. Then we reformulate the problem using state variable equations and obtain exponential convergence conditions, initially involving input and output quantities. Finally, we use the ideas of Section 2 to obtain an exponential convergence condition involving input quantities only. We then look at equation error identification, which is far simpler.

The basic problem; plant and identifier structures

We follow the notation of Johnson (1979b) and postulate a model

$$y_k + \alpha_1 y_{k-1} + \cdots + \alpha_n y_{k-n} = \beta_1 u_{k-1} + \cdots + \beta_m u_{k-m} \tag{12}$$

(with n, m assumed known). We assume that n is minimal. We also assume that coefficients γ_i are known such that the transfer function

$$H(z) = \left(1 + \sum_{i=1}^n \gamma_i z^{-i}\right) \Big/ \left(1 + \sum_{i=1}^n \alpha_i z^{-i}\right) \tag{13}$$

is strictly positive real. [The conjunction of positive real transfer functions, such as (13), with adaptive problems goes back at least to Butchart and Shackcloth (1965) and Parks (1966), if not earlier.]

This means that the model is necessarily stable. An adaptive identifier is constructed which at time k is parameterized by $\hat{\alpha}_i(k)$, $\hat{\beta}_i(k)$. It is considered as having two outputs, an a priori output $\hat{y}_k$ and an *a posteriori* output z_k

$$\hat{y}_k = -\sum_{i=1}^{n} \hat{\alpha}_i(k) z_{k-i} + \sum_{j=1}^{m} \hat{\beta}_j(k) u_{k-j} \tag{14}$$

$$z_k = -\sum_{i=1}^{n} \hat{\alpha}_i(k+1) z_{k-i} + \sum_{j=1}^{m} \hat{\beta}_j(k+1) u_{k-j}. \tag{15}$$

The parameters are updated with the aid of a smoothed error defined as

$$v_k = y_k - z_k + \sum_{i=1}^{n} \gamma_i [y_{k-i} - z_{k-i}] \tag{16}$$

with the update equations

$$\hat{\alpha}_i(k+1) = \hat{\alpha}_i(k) - \mu_i z_{k-i} v_k, \qquad \mu_i > 0 \tag{17}$$

$$\hat{\beta}_j(k+1) = \hat{\beta}_j(k) + \rho_j u_{k-j} v_k, \qquad \rho_j > 0. \tag{18}$$

Calculation of $\hat{\alpha}_i(k+1)$, according to (17), needs v_k, which according to (16) needs z_k, which according to (15) needs $\hat{\alpha}_i(k+1)$; to avoid this circularity, one uses the following equivalent expression for v_k

$$v_k = \left(y_k - \hat{y}_k + \sum_{i=1}^{n} \gamma_i [y_{k-i} - z_{k-i}]\right) \Big/ \left(1 + \sum_{i=1}^{n} \mu_i z_{k-i}^2 + \sum_{j=1}^{m} \rho_j u_{k-j}^2\right). \tag{19}$$

With bounded inputs, it can be shown by a simple Lyapunov analysis (Lin and Narendra, 1980; Johnson, 1980a), or by appeal to hyperstability arguments (Landau, 1976, 1978; Johnson, 1979a) that $\hat{y}_k \to y_k$, i.e. that the identifier output converges to the plant output. Hyperstability methods and the Lyapunov analysis for this problem are closely linked concepts; nevertheless the latter offers more prospect of establishing qualitative properties like exponential convergence, and even rate of convergence in particular instances.

It is easy to see that the convergence of $\hat{y}_k$ to y_k need not be exponentially fast—this can be checked for the example with $y_k = \beta_1 u_{k-1}$, where $u_{k-1} = 0$ except when $k - 1 = 2^\alpha$, $\alpha = 1, 2, 3, \ldots$ and then $\rho u_{k-1}^2 = 1$; a simple calculation shows that $\hat{y}_k - y_k \to 0$ at a $1/k$ rate. Slower rates again are possible. It is also easy to see that even though u_k might be such as to apparently cause $\hat{y}_k$ to approach y_k exponentially fast, this convergence might not be robust, in that after convergence has apparently occurred, variation of u_k may mean that $\hat{y}_k$ no longer tracks y_k. Consider for example a plant with nonconstant transfer function, and suppose that u_k is constant over $[0, K]$ for some very large, arbitrary but fixed K. Then in $[0, K]$, $\hat{y}_k \to y_k$ exponentially fast. But if, at time $K+1$, u_k becomes sinusoidal, $\hat{y}_k$ no longer tracks y_k, at least until the identifier has learnt the value of the plant transfer function at the relevant frequency. To avoid both of these kinds of behavior, we want the identifier parameters to approach the plant parameters exponentially fast. Accordingly, we are led to the basic problem: find conditions on $\{u_k\}$ such that $\hat{y}_k \to y_k$, $\hat{\alpha}_i(k) \to \alpha_i$, and $\hat{\beta}_i(k) \to \beta_i$.

Solution to basic problem

Most of the remainder of this section will be devoted to establishing the following result.

Theorem 3.1. Consider the plant (12) where n, m are assumed known, and $\alpha(z) = z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n$ and $\beta(z) = \beta_1 z^{n-1} + \cdots + \beta_n z^{n-m}$ are coprime but otherwise unknown. Suppose also that γ_i are known such that $H(z)$ in (13) is strictly positive real, and that the adaptive identification scheme of (14)–(19) is used. Then $y_k - \hat{y}_k$, $\alpha_i - \hat{\alpha}_i(k)$ and $\beta_j - \hat{\beta}_j(k)$ converge exponentially to zero provided that (a) or (b) holds

(a) $\{u_k\}$ is deterministic, and for S some integer $\geq 2n + m - 2$, for all j and some ρ_3, ρ_4

$$\rho_4 I > \sum_{j}^{j+S-n+1} \begin{bmatrix} u_{k+n} \\ \vdots \\ u_{k-m+1} \end{bmatrix} [u_{k+n} \cdots u_{k-m+1}] > \rho_3 I > 0$$

(b) $\{u_k\}$ is stochastic, and for S some integer $\geq n - 1$, for all j and some ρ_3, ρ_4

$$\rho_4 I > \sum_{j}^{j+S-n+1} E\left\{ \begin{bmatrix} u_{k+n} \\ \vdots \\ u_{k-m+1} \end{bmatrix} [u_{k+n} \cdots u_{k-m+1}]\right\} > \rho_3 I > 0.$$

Reformulation using state-variable equations

Following Johnson (1980a), we set up state-variable equations. Make the definitions

$$e_k = \begin{bmatrix} y_{k-1} - z_{k-1} \\ \vdots \\ y_{k-n} - z_{k-n} \end{bmatrix}, \qquad \phi_k = \begin{bmatrix} -\alpha_1 + \hat{\alpha}_1(k) \\ \vdots \\ -\alpha_n + \hat{\alpha}_n(k) \\ \beta_1 - \hat{\beta}_1(k) \\ \vdots \\ \beta_m - \hat{\beta}_m(k) \end{bmatrix}, \tag{20}$$

$$x_k = [z_{k-1} z_{k-2} \cdots z_{k-n} u_{k-1} \cdots u_{k-m}],$$

$$w_k = \sum_{i=1}^{n} [-\alpha_i + \hat{\alpha}_i(k+1)] z_{k-1} + \sum_{j=1}^{m} [\beta_j - \hat{\beta}_j(k+1)] u_{k-j}$$

and

$$A = \begin{bmatrix} -\alpha_1 & -\alpha_2 & \cdots & -\alpha_n \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 \cdots & 0 & 1 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad h' = [\gamma_1 - \alpha_1 \cdots \gamma_n - \alpha_n] \tag{21}$$

$$\Gamma = \operatorname{diag}[\mu_1 \cdots \mu_n \, \rho_1 \cdots \rho_n].$$

Then one obtains

$$\begin{aligned} e_{k+1} &= Ae_k + bw_k \\ v_k &= h'e_k + w_k \\ w_k &= \phi_k' x_k - x_k' \Gamma x_k v_k \\ \phi_{k+1} &= \phi_k - \Gamma v_k x_k. \end{aligned} \tag{22}$$

Equations (22) can be thought of as defining a feedback system with forward part F_1 defined by the strictly positive real transfer function $1 + h'(zI - A)^{-1}b$ and feedback part F_2 defined by a time-varying passive system with input v_k and output w_k as in Fig. 1. This can be used as the basis of a general asymptotic stability (but not exponential stability) argument via hyperstability concepts. As it turns out, to prove exponential stability, some reformulation of (22) is necessary. We use a device of Landau (1978), Landau and Silveira (1979) and Bitmead and Anderson (1980). Define, for a fixed δ

$$\bar{w}_k = w_k - \delta v_k$$

and consider the arrangement of Fig. 2 where δ is a positive constant, chosen sufficiently small that F_1' still has a strictly positive real transfer function.* The scheme of Fig. 2 is of course equivalent to that of Fig. 1; suppose the defining equations are

$$\begin{aligned} \bar{e}_{k+1} &= \bar{A}\bar{e}_k + \bar{b}\bar{w}_k \\ v_k &= \bar{h}'\bar{e}_k + \bar{d}\bar{w}_k \\ \bar{w}_k &= \phi_k' x_k - (x_k'\Gamma x_k + \delta) w_k \\ \phi_{k+1} &= \phi_k - \Gamma v_k x_k. \end{aligned} \tag{23}$$

As simple calculations show, these equations may be reformulated as

$$\begin{bmatrix} e_{k+1} \\ \phi_{k+1} \end{bmatrix} = F_k \begin{bmatrix} e_k \\ \phi_k \end{bmatrix} \tag{24}$$

where

$$F_k = \begin{bmatrix} \bar{A} - \dfrac{\delta + x_k'\Gamma x_k}{1 + \bar{d}x_k'\Gamma x_k + \bar{d}\delta}\bar{b}\bar{h}' & \dfrac{1}{1 + \bar{d}x_k'\Gamma x_k + \bar{d}\delta}\bar{b}x_k' \\ -\dfrac{1}{1 + \bar{d}x_k'\Gamma x_k + \bar{d}\delta}\Gamma x_k h' & I - \dfrac{\bar{d}}{1 + \bar{d}x_k'\Gamma x_k + \bar{d}\delta}\Gamma x_k x_k' \end{bmatrix} \tag{25}$$

The basic problem now becomes one of establishing exponential stability of (24) under suitable conditions on $\{u_k\}$.

*Later, we shall note why $\delta = 0$ would cause difficulty.

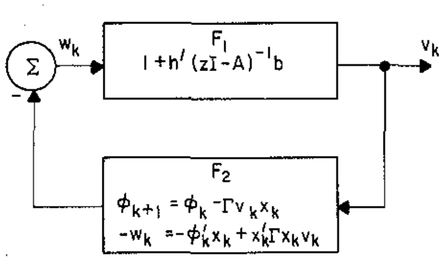

FIG. 1. Arrangement of adaptive scheme as two passive systems in feedback loop.

Exponential convergence in terms of the x_k *vector*

The following argument is based on arguments for the stochastic x_k case in Bitmead and Anderson (1980a, b). The deterministic results we obtain are discrete-time analogues of continuous time results in, e.g. Anderson (1977a).

We associate with the linear equation (24) the quadratic Lyapunov function $V \triangleq e_k' P e_k + \phi_k' \Gamma^{-1} \phi_k$, where P is a real positive definite matrix satisfying the discrete-time Kalman–Yakubovich lemma equations, (see e.g. Hitz and

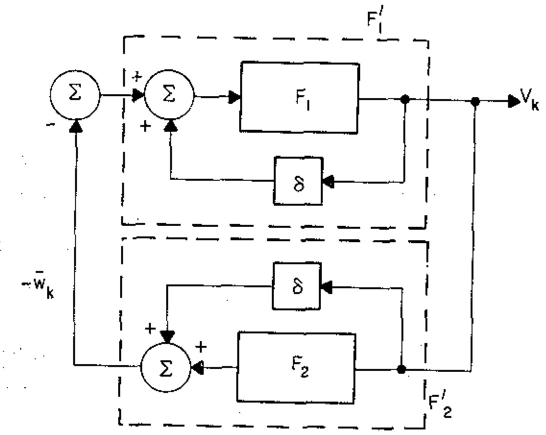

FIG. 2. Rearrangement of scheme of Fig. 1.

Anderson (1969), appropriate for a strict positive real transfer function

$$\begin{aligned} \bar{A}'P\bar{A} - P &= -\bar{L}\bar{L}' - \sigma^2 I \\ \bar{A}'P\bar{b} &= \bar{h} + \nu\bar{L} \\ 2\bar{d} - \bar{b}'P\bar{b} &= \nu^2 \end{aligned} \tag{26}$$

with $\bar{L}$ real, $\sigma \neq 0$, $\nu \neq 0$, and $\{\bar{A}, \bar{b}, \bar{h}, \bar{d}\}$ minimal, i.e. the numerator and denominator in (13) can have no common factors. Should common factors exist, there exists a neighboring set of γ_i giving rise to a neighboring $H(z)$, still strictly positive real, for which no common factor exists. This is one way around the difficulty. The other is to introduce a minimal realization to replace that defined $\{\bar{A}, \bar{b}, \bar{c}, \bar{d}\}$; the nonminimal modes are guaranteed decaying since they are defined by stable zeros of $1 + \Sigma\, \alpha_i z^{-i}$.

As direct calculation shows Bitmead and Anderson (1980b)

$$F_k' \begin{bmatrix} P & 0 \\ 0 & \Gamma^{-1} \end{bmatrix} F_k - \begin{bmatrix} P & 0 \\ 0 & \Gamma^{-1} \end{bmatrix} = -H_k H_k' \tag{27}$$

where

$$H_k = \begin{bmatrix} \sigma I & L + \nu(\delta + x_k'\Gamma x_k)\chi_k h & \xi_k \chi_k h \\ 0 & -\chi_k \nu x_k & \xi_k \chi_k \bar{d} x_k \end{bmatrix} \tag{28}$$

with scalars

$$\chi_k = (1 + \bar{d}\delta + \bar{d}x_k'\Gamma x_k)^{-1}, \quad \xi_k = [2\delta + x_k'\Gamma x_k]^{1/2}. \tag{29}$$

Note that equation (27) reflects the fact that

$$V(e_{k+1}, \phi_{k+1}) - V(e_k, \phi_k) = -[e_k'\phi_k']H_k H_k'[e_k'\phi_k']'.$$

The discrete-time, time-varying equation (27) is the prototype equation for applying the time-varying Lemma of Lyapunov, see Anderson and Moore (1969) and Anderson (1977a) for a continuous-time version and Anderson and Moore (1981) for a discrete-time version. This lemma allows us to conclude that (24) is exponentially stable provided the pair $[F_k, H_k]$ is uniformly observable. Now $[F_k, H_k]$ is uniformly observable if and only if $[F_k + K_k H_k', H_k]$ is uniformly observable (Anderson and Moore, 1981) for any bounded sequence K_k. Taking

$$K_k = \begin{bmatrix} 0 & \nu^{-1}b & 0 \\ 0 & 0 & \xi_k^{-1}\Gamma x_k \end{bmatrix} \tag{30}$$

we obtain

$$F_k + K_k H_k' = \begin{bmatrix} A + \nu^{-1}\bar{b}\bar{L} & 0 \\ 0 & I \end{bmatrix} \tag{31}$$

while we may write

$$H_k = \begin{bmatrix} \sigma I & h & 0 \\ 0 & 0 & x_k \end{bmatrix} \begin{bmatrix} I & \sigma^{-1}L & 0 \\ 0 & \nu(\delta + x_k'\Gamma x_k)\chi_k I & \xi_k \chi_k I \\ 0 & -\chi_k \nu I & \xi_k \chi_k \bar{d} I \end{bmatrix} \tag{32}$$

The second matrix on the right-hand side of (32), call it R_k, is bounded and has a bounded inverse.* It is easily checked that $[F_k + K_k H_k', H_k]$ is uniformly observable if and only if $[F_k + K_k H_k', H_k R_k^{-1}]$ is uniformly observable. The diagonal structure of $F_k + K_k H_k'$ and $H_k R_k^{-1}$ means that this pair is uniformly observable if and only if both $[\bar{A} + \nu^{-1}\bar{b}\bar{L}, \sigma I]$ is completely observable, and $[I, x_k]$ is uniformly observable. The first pair is trivially observable; the second is uniformly observable if and only if

$$\rho_2 I > \sum_{k=j}^{j+S} x_k x_k' > \rho_1 I > 0 \tag{33}$$

for some S and all j. The upper bound is guaranteed if $\{u_k\}$ is bounded, by virtue of the stability of the system (12).

In case x_k is a stationary random process, arguments as in Bitmead and Anderson (1980a, b) which are based on a random time-varying Lyapunov Lemma, allow us to conclude exponential stability of (24) in case

$$E[x_k x_k'] > 0 \tag{34}$$

and in case x_k is nonstationary but satisfies a mixing property (requiring past and future to be asymptotically independent) exponential stabil-

*A little calculation will show that if $\delta = 0$, R_k may not have a bounded inverse.

ity follows if

$$\rho_2 I > \liminf \frac{1}{n} \sum_1^n E[x_k x_k'] > \rho_1 I > 0. \quad (35)$$

Exponential convergence in terms of the input vector

Satisfaction of (33) for deterministic $\{u_k\}$ [and $\{z(k)\}$], (34) for ergodic $\{u_k\}$ [and $\{z(k)\}$], and (35) for a mixing $\{u_k\}$ [and $\{z(k)\}$] imply exponential stability. We seek conditions free of the $\{z_k\}$.

Now recall the definition (20) of x_k. If the entry z_{k-i} of x_k were replaced by y_{k-i}, the material of Section 2 would immediately allow us to replace the conditions (33), (34), or (35) by a condition involving the $\{u_k\}$ sequence only. We now argue that the fact that x_k contains z_{k-i} rather than y_{k-i} makes no difference to this condition. Let $\bar{x}_k$ denote x_k with $z_{k-1}, \ldots, z_{k-n}$ replaced by $y_{k-1}, \ldots, y_{k-n}$ and let $\bar{F}_k$ denote the corresponding change to F_k. The basic convergence result that a simple Lyapunov analysis yields is that $z_k \to y_k$ and so $F_k \to \bar{F}_k$. We now observe.

Lemma 3.2. Suppose that $\bar{\eta}_{k+1} = \bar{F}_k \bar{\eta}_k$ is exponentially stable, and $F_k \to \bar{F}_k$. Then $\eta_{k+1} = F_k \eta_k$ is exponentially stable.†

The proof is a direct analog of the continuous-time case (Willems, 1970). □

Now we construct the following chain of argument. Conditions of the type stated in Theorem 2.2 and 2.3 on $\{u_k\}$ imply conditions of the type in (32) through (34) on $\bar{x}_k$, by Theorems 2.2 and 2.3. Accordingly, $\bar{\eta}_{k+1} = \bar{F}_k \bar{\eta}_k$ is exponentially stable. But then by Lemma 3.2, $\eta_{k+1} = F_k \eta_k$ is exponentially stable; in particular, since $\eta_k = [e_k' \phi_k']'$, output and parameter errors both approach zero exponentially fast.

Equation error identification

Results for equation error identification (Mendel, 1973; Johnson, 1979a) are much easier to achieve. In the simplest approach to equation error identification, we consider the plant (12) as before, i.e. n, m known, n minimal. We also explicitly assume the plant is stable. We also consider the identifier

$$\hat{y}_k = -\sum_{i=1}^n \hat{\alpha}_i(k) y_{k-i} + \sum_{j=1}^m \hat{\beta}_j(k) u_{k-j} \quad (36)$$

where we set

$$\hat{\alpha}_i(k+1) = \hat{\alpha}_i(k) \frac{\mu_i}{1 + \sum_{i=1}^n \mu_i y_{k-i}^2 + \sum_{j=1}^m \rho_j u_{k-j}^2} \times y_{k-i}(y_k - \hat{y}_k)$$

$$\hat{\beta}_j(k+1) = \hat{\beta}_j(k) - \frac{\rho_j}{1 + \sum_{i=1}^n \mu_i y_{k-i}^2 + \sum_{j=1}^m \rho_j u_{k-j}^2} \times u_{k-j}(y_k - \hat{y}_k).$$

With ϕ_k the parameter error vector of (20) and Γ as in (21) and with

$$\bar{x}_k = [y_{k-1} \cdots y_{k-n}\ u_{k-1} \cdots u_{k-m}] \quad (37)$$

we have

$$\phi_{k+1} = \left[I - \frac{\Gamma \bar{x}_k \bar{x}_k'}{1 + \bar{x}_k' \Gamma \bar{x}_k} \right] \phi_k. \quad (38)$$

A simplified version of the argument applying for output error, where we studied (24), yields (33), (34) or (35) with x_k replaced by $\bar{x}_k$ in the deterministic, ergodic, or nonstationary mixing cases, respectively, as guaranteeing exponential convergence. [With $V(\phi_k) \triangleq \phi_k' \Gamma^{-1} \phi_k$, one can show that $V(\phi_{k+1}) - V(\phi_k) \leq -\|H_k' \phi_k\|^2$ where $H_k = \bar{x}_k(1 + \bar{x}_k' \Gamma \bar{x}_k)^{-(1/2)}$; then one uses the uniform observability trick.] Section 2 results then allow reformulation of these conditions in terms of $\{u_k\}$ only, and a theorem like Theorem 3.1 results. Of course, exponential convergence of ϕ_k to zero assures exponential convergence of $\hat{y}_k$ to y_k.

Summary

For the output error and equation error identifiers described in this section, applied to the plant (12) where n, m are known, n is *minimal*, $\{u_k\}$ *is bounded and the plant is stable*, the condition of (6), (8), (9) or (10), as appropriate, guarantees exponential convergence of parameter estimates and identifier outputs to the correct values. Of course, for the output error identifier, coefficients γ_i must be known causing $H(z)$ in (13) to be strictly positive real.

4. ADAPTIVE CONTROL

Without sacrificing the essential ideas, we shall consider a single-input, single-output system. We shall follow the first projection algorithm of Goodwin, Ramadge and Caines (1980) shown in Johnson (1980b) to underlie several approaches to adaptive control. We shall show that when the output of the unknown plant is required to follow a sufficiently rich reference trajectory, in the sense that conditions like the output-only conditions of Section 2—see especially Theorem 2.4—are satisfied by the reference trajectory, then the estimate of the plant parameters will converge exponentially fast to the true value, as will the plant output to

†If $\Phi_{k,l}$ and $\bar{\Phi}_{k,l}$ denote the two transition matrices, $\|\Phi_{k,l}\| \leq A\beta^{k-l}$ for some $\beta < 1$, then $\|\bar{\Phi}_{k,l}\| < \bar{A}\bar{\beta}^{k-l}$ for some A, with $\bar{A}$ possibly far larger than A and any $\bar{\beta} > \beta$.

the reference trajectory. The argument seems reasonably representative, i.e. with minor modification it should apply to other adaptive control algorithms.

We shall first set up the basic convergence problem (which involves specifying the controller algorithm) and summarize its solution; next we shall indicate how stability can be described with Lyapunov methods and conclude by obtaining the output-only condition for exponential stability.

The basic problem; plant and controller structures

We assume the plant is modelled by

$$y_k + \alpha_1 y_{k-1} + \cdots + y_{k-n} = \beta_d u_{k-d} + \cdots + \beta_m u_m \qquad \beta_d \neq 0 \quad (39)$$

where the values of d, m, n are assumed known, but not the values of the α_i, β_j. We assume that $\alpha(z) = z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n$ and $\beta(z) = \beta_d z^{n-d} + \cdots + \beta_m z^{n-m}$ are coprime,† and that $\beta(z)$ has all its zeros in $|z| < 1$. This means that if $\{y_k\}$ is bounded, so is $\{u_k\}$.

We now describe the adaptive controller. As is known, an alternative representation for the plant is

$$y_{k+d} = a_0 y_k + a_1 y_{k-1} + \cdots + a_{n-1} y_{k-n+1} + b_0 u_k + \cdots + b_{m-1} u_{k-m+1} \quad (40)$$

for some a_i, b_j related to the α_i, β_j. [See Appendix B for the simple derivation of (40) based on ideas of Åström (1970).] The control objective is to ensure that

$$\lim_{k\to\infty} (y_k - y_k^*) = 0 \quad (41)$$

where y_k^* is a bounded reference trajectory. Define

$$x_k = [y_k \, y_{k-1} \cdots y_{k-n+1} \, u_k \cdots u_{k-m+1}]' \quad (42)$$

and

$$\theta_0 = [a_0 \cdots a_{n-1} \, b_0 \cdots b_{m-1}]'. \quad (43)$$

Notice that

$$y_k = x'_{k-d}\theta_0. \quad (44)$$

Now suppose that $\hat{\theta}_{k-1}$ is an estimate of θ_0 available after measurement of y_{k-1}. We recursively set

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \nu_k x_{k-d}(1 + x'_{k-d}x_{k-d})^{-1}(y_k - x'_{k-d}\hat{\theta}_{k-1}) \quad (45)$$

with ν_k as defined below, and we choose u_k to ensure satisfaction of

$$x'_k\hat{\theta}_k = y^*_{k+d}. \quad (46)$$

That this equation can be satisfied is ensured by the rule for choosing ν_k. If the choice $\nu_k = 1$ causes the $(n+1)$th entry of $\hat{\theta}_k$ in (45) to be nonzero, we make this choice of ν_k. Otherwise, we set $\nu_k = \gamma$ for some arbitrary but fixed $\gamma \in (0, 2)$ such that the $(n+1)$th entry of $\hat{\theta}(k)$ is nonzero. Let us also define

$$\phi_k = \theta_k - \theta_0. \quad (47)$$

Then (44) and (45) yield

$$\phi_k = \left[I - \frac{\nu_k x_{k-d}x'_{k-d}}{1 + x'_{k-d}x_{k-d}}\right]\phi_{k-1}. \quad (48)$$

In Goodwin, Ramadge and Caines (1980) the following facts are established

$$x_k \quad \text{is bounded} \quad (49)$$

and

$$\lim_{k\to\infty} (y_k - y_k^*) = 0. \quad (50)$$

Convergence is however not guaranteed to be exponential. (It is easy to contrive a situation where it will not be exponential.) This leads us to the basic problem. Find conditions on y_k^* guaranteeing that $y_k - y_k^*$ converges exponentially fast. The conditions we shall find will also ensure that ϕ_k converges exponentially fast to zero; this means that we cannot have a situation where $y_k - y_k^*$ apparently converges to zero and then, when y_k^* changes it character, $y_k - y_k^*$ becomes large.

Solution to the basic problem

Theorem 4.1. Consider the plant (4.1) where n, m and d are assumed known, $\alpha(z) = z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n$ and $\beta(z) = \beta_d z^{n-d} + \cdots + \beta_m z^{n-m}$ are coprime but otherwise unknown and all zeros of $\beta(z)$ lie in $|z| < 1$. Let $\{y_k^*\}$ be a bounded reference trajectory, and suppose the adaptive control scheme of (42), (45) and (46) is used. Then $y_k - y_k^*$ and $\phi_k = \hat{\theta}_k - \theta_0$ converge exponentially fast to zero [θ_0 being defined by (40) and (43)] provided that (a)

†Coprimeness is not assumed in Goodwin, Ramadge and Caines (1980). It is clear that given a common factor, no convergence of the parameter estimation error to zero could be expected.

or (b) holds

(a) $\{y_k^*\}$ is deterministic and for S some integer $\geq 2m+n-2$, all $j \geq$ some j_0 and some ρ_3, ρ_4

$$\rho_4 I > \sum_j^{j+S-m+1} \begin{bmatrix} y_{k+m}^* \\ \vdots \\ y_{k-n+1}^* \end{bmatrix} [y_{k+m}^* \cdots y_{k-n+1}^*] > \rho_3 I > 0 \tag{51}$$

(b) $\{y_k^*\}$ is stochastic, and for S some integer $\geq m-1$, for all $j \geq$ some j_0 and some ρ_3, ρ_4

$$\rho_4 I > \sum_j^{j+S-m+1} E\{ \begin{bmatrix} y_{k+m}^* \\ \vdots \\ y_{k-n+1}^* \end{bmatrix} [y_{k+m}^* \cdots y_{k-n+1}^*]\} > \rho_3 I > 0. \tag{52}$$

Lyapunov function formulation of exponential stability

We shall study (48). If we can show $\phi_k = \hat{\theta}_k - \theta_0 \to 0$ exponentially fast, then $x_k' \phi_k = y_{k+d}^* - y_{k+d}$ has the same property, since x_k is bounded. Define

$$F_{k-1} = I - \frac{\nu_k x_{k-d} x_{k-d}'}{1 + x_{k-d}' x_{k-d}}. \tag{53}$$

Adopt as a Lyapunov function for (48) $V(\phi_k) = \phi_k' \phi_k$. Then

$$\begin{aligned} V(\phi_k) - V(\phi_{k-1}) &= \phi_{k-1}'[F_{k-1}' F_{k-1} - I]\phi_{k-1} \\ &= -\phi_{k-1}' H_{k-1}' H_{k-1} \phi_{k-1} \end{aligned} \tag{54}$$

where

$$H_{k-1} = \frac{\nu_k}{(1 + x_{k-d}' x_{k-d})^{1/2}} \left[2 - \nu_k \frac{x_{k-d}' x_{k-d}}{1 + x_{k-d}' x_{k-d}}\right]^{1/2} \tag{55}$$

Notice that

$$F_{k-1}' F_{k-1} - I = -H_{k-1} H_{k-1}'. \tag{56}$$

As we known, (48) will be exponentially stable provided that $[F_k, H_k]$ is uniformly observable, or equivalently $[F_k + K_k H_k', H_k R_k]$ is uniformly observable where K_k is a bounded sequence, and R_k is bounded, together with its inverse. Let us take

$$R_{k-1} = \frac{(1 + x_{k-d}' x_{k-d})^{1/2}}{\nu_k^{1/2}} \left[2 - \nu_k \frac{x_{k-d}' x_{k-d}}{1 + x_{k-d}' x_{k-d}}\right]^{-1/2} \tag{57}$$

$$K_{k-1} = \nu_k \frac{x_{k-d}}{1 + x_{k-d}' x_{k-d}} R_{k-1}$$

which ensures that

$$F_k + K_k H_k' = I \qquad H_k R_k = x_{k-d}. \tag{58}$$

Arguing as in the last section, we conclude that in the deterministic $\{y_k^*\}$ case, a sufficient condition for the desired exponential stability is

$$\rho_1 I > \sum_j^{j+S} x_{k-d} x_{k-d}' > \rho_2 I > 0 \tag{59}$$

for all j and some S and in the stochastic case

$$\rho_1 I > \sum_j^{j+S} E[x_{k-d} x_{k-d}'] > \rho_2 I > 0 \tag{60}$$

for all j and some S.

Exponential stability using reference trajectory conditions only

We shall argue the deterministic case only, the stochastic case being very similar. By Theorem 2.4, (59) is guaranteed to hold for some $S \geq 2m+n-2$ and all j provided that

$$\rho_4 I > \sum_j^{j+S-m+1} \begin{bmatrix} y_{k+m} \\ \vdots \\ y_{k-n+1} \end{bmatrix} [y_{k+m} \cdots y_{k-n+1}] > \rho_3 I > 0 \tag{61}$$

for ρ_3, ρ_4 and all j. Since $y_k \to y_k^*$ as $k \to \infty$ irrespective of the conditions on y_k^*, (61) is guaranteed to hold for all suitably large j, if for all suitably large j

$$\rho_4 I > \sum_j^{j+S-m+1} \begin{bmatrix} y_{k+m}^* \\ \vdots \\ y_{k-n+1} \end{bmatrix} [y_{k+m}^* \cdots y_{k-n+1}^*] > \rho_3 I > 0. \tag{62}$$

For exponential convergence of (48), it is enough that (59) hold for all $j \geq$ some j_0, and accordingly that (62) hold for all $j \geq$ some j_0.

5. CONCLUSIONS

We summarize the main results as follows. If we model the order of a system correctly, if it is stable, and if the input is persistently spanning in an appropriate deterministic or stochastic sense, then we shall secure exponential convergence of identifier output and parameters to unknown system output and parameters in both output error and equation error identification, provided that in the former we also have positive realness of a certain transfer function. For adaptive control, if we model the plant order correctly, if the plant has zeros inside $|z| = 1$, or we otherwise ensure the control input remains provably bounded, and if the reference trajectory is persistently spanning, we get exponential convergence.

What do these results lead on to? First, as noted in the introduction, exponential convergence implies robustness of adaptive algorithms, e.g. given noise or slowly-varying parameters. But if one assumes a plant order larger than is really the case, difficulty must be expected, perhaps more so for adaptive control. For then we cannot expect to identify the plant, and indeed one simulation study reported to us by E. J. Hannan in a private communication has resulted in the adaptive model predicting a pole-zero cancellation which migrates to +1. For similar reasons, we can be concerned about the absence of persistent excitation conditions. In equation error identification, one can show that, with absence of persistence excitation and with noise in the calculations, one or more linear functionals of the parameter estimation vector will diverge. As against these conclusions though, one could well argue of any real plant that it is impossible to overmodel its order, so that these difficulties are largely illusory.

More practically, proving exponential stability opens up the possibility of results on approximate identification and adaptive control, including results applicable when the unknown plant order is underestimated. The relevant equations are normally perturbations of those equations for which exponential stability has been established; certain perturbations of exponentially stable equations are always possible, in the sense that at worst bounded errors will result. Similar calculations to those justifying approximate identification should also justify the robustness of various algorithms when the plant is slowly time-varying. Some persistent excitation requirement can still be expected.

There are several directions in which the ideas of this paper could be extended in a straightforward manner. Multiple-input, multiple-output systems with appropriate (efficient) parameterizations should be easily treated; other adaptive control algorithms for trajectory following [which all seem to fall into one family (Johnson, 1980b)] should be equally amenable to analysis. Continuous-time problems raise a slightly more subtle issue. In passing from input-output conditions for exponential stability to input-only conditions, it is necessary to demand that the inputs not become faster and faster (else the contributions to the output become smaller and smaller). This can be accomplished by using a technical device due to Yuan and Wonham (1977) also used in Anderson (1977a, b).

Less straightforward extensions are required to address questions concerning the actual convergence rate. This paper simply proves the existence of an exponential convergence rate but does not provide a formula for that rate. Given such a formula, questions could be addressed concerning, e.g., (i) the relative convergence speeds of output and equation error parameter estimators, (ii) the effects of input $\{u(k)\}$ composition and designer-selected parameters, such as the step-size constants in Γ and the error smoothing coefficients γ_i, on the convergence rate, (iii) the possible benefits of various time-varying matrix gains as in Landau and Silveira (1979), and (iv) the possible benefits of more general parameter estimator algorithms such as in Kreisselmeier (1977, 1979) and Ljung (1981). The fact, proven in this paper, that such an exponential rate exists should spur efforts in these directions.

REFERENCES

Anderson, B. D. O. (1974a). Multivariable adaptive identification. Technical Report EE 7402, University of Newcastle, Australia.

Anderson, B. D. O. (1974b). Adaptive identification of multiple-input, multiple-output plants. *Proc.* 1974 *IEEE Conf. Decision and Control*, 273.

Anderson, B. D. O. (1977a). Exponential stability of linear equations arising in adaptive identification. *IEEE Trans Aut. Control*, **AC-22**, 83.

Anderson, B. D. O. (1977b). An approach to multivariable system identification. *Automatica*, **13**, 401.

Anderson, B. D. O. and J. B. Moore (1969). New results in linear system stability. *SIAM J. Control*, **7**, 398.

Anderson, B. D. O. and J. B. Moore (1981). Detectability and stabilizability of discrete-time linear systems. *SIAM J. Contr. and Opt.*, (to be published).

Åström, K. J. (1970). *Introduction to Stochastic Control Theory.* Academic Press, New York.

Bitmead, R. R. and B. D. O. Anderson (1978). Exponentially convergent behaviour of simple stochastic adaptive estimation algorithms. *Proc.* 1978 *IEEE Conf. Decision and Control*, 580.

Bitmead, R. R. and B. D. O. Anderson (1980a). Lyapunov techniques for the exponential stability of linear difference equations with random coefficients. *IEEE Trans Aut. Control*, **AC-25**, 782.

Bitmead, R. R. and B. D. O. Anderson (1980b). Performance of adaptive estimation algorithms in dependent random environments. *IEEE Trans Aut. Control*, **AC-25**, 788.

Butchart, R. L. and B. Shackloth (1965). Synthesis of model reference adaptive control systems by Lyapunov's second method. *Proc. IFAC Symp. on Adaptive Control*, Teddington, England.

Dugard, L. and I. D. Landau (1980). Recursive output error identification algorithms—theory and evaluation. *Automatica*, **16**, 443.

Farden, D. C., J. C. Goding, Jr. and K. Sayood (1979). On the 'desired behaviour' of adaptive signal processing algorithms. *Proc. IEEE Conf. on Acoustics, Speech and Signal Processing.*

Goodwin, G. C., P. J. Ramadge and P. E. Caines (1980). Discrete time multivariable adaptive control. *IEEE Trans Aut. Control*, **AC-25**, 449.

Hitz, K. L. and B. D. O. Anderson (1969). Discrete positive-real functions and their applications to system stability. *Proc. IEE*, **116**, 153.

Hodge, W. V. D. and D. Pedoe (1968). *Methods of Algebraic Geometry*, Vol. 1. Cambridge University Press, p. 151.

Johnson, C. R., Jr (1979a). Adaptive parameter matrix and output vector estimation via an equation error formulation. *IEEE Trans Syst., Man, Cybernet.*, **SMC-9**, 392.

Johnson, C. R., Jr (1979b). A convergence proof for a hyperstable adaptive recursive filter. *IEEE Trans Info. Theory*, **IT-25**, 745.

Johnson, C. R., Jr (1980a). Another use of the Lin–Narendra error model: HARF. *IEEE Trans Aut. Control*, **AC-25**, 985.

Johnson, C. R., Jr (1980b). Input matching, error augmentation, self-tuning, and output error identification: Algorithmic similarities in discrete adaptive model-following. *IEEE Trans Aut. Control*, **AC-25**, 697.

Jones, S. K. (1973). Adaptive filtering with correlated training samples: a mean square error analysis. Ph.D. Thesis, Southern Methodist University.

Kreisselmeier, G. (1977). Adaptive observers with exponential rate of convergence. *IEEE Trans Aut. Control*, **AC-22**, 2.

Kreisselmeier, G. (1979). The generator of adaptive law structures for globally convergent adaptive observers. *IEEE Trans Aut. Control*, **AC-24**, 510.

Landau, I. D. (1976). Unbiased recursive identification using model reference adaptive techniques. *IEEE Trans Aut. Control*, **AC-21**, 194.

Landau, I. D. (1978). An addendum to "Unbiased recursive identification using model reference adaptive techniques." *IEEE Trans Aut. Control*, **AC-23**, 97.

Landau, I. D. and H. M. Silveira (1978). A stability theorem with applications to adaptive control. *IEEE Trans Aut. Control*, **AC-24**, 305.

Lin, Y. H. and K. S. Narendra (1980). A new error model for adaptive systems. *IEEE Trans Auto. Control*, **AC-25**, 585.

Lion, P. M. (1967). Rapid identification of linear and non-linear systems. *AIAA J.*, **5**, 1835.

Ljung, L. (1981). Analysis of a general recursive prediction error algorithm. *Automatica*, **17**, 89.

Mendel, J. M. (1973). *Discrete Techniques of Parameter Estimation: The Equation Error Formulation.* Marcel Dekker, New York.

Morgan, A. P. and K. S. Narendra (1977a). On the uniform asymptotic stability of certain linear nonautonomous differential equations. *SIAM J. Control*, **15**, 5.

Morgan, A. P. and K. S. Narendra (1977b). On the stability of nonautonomous differential equations $\dot{x}=[A+B(t)]x$ with skew symmetric matrix $B(t)$. *SIAM J. Contr. and Opt.*, **15**, 163.

Parks, P. C. (1966). Lyapunov redesign of model reference adaptive control systems. *IEEE Trans Aut. Control*, **AC-11**, 362.

Sondhi, M. M. and D. Mitra (1976). New results on the performance of a well-known class of adaptive filters. *Proc. IEEE*, **64**, 1583.

Weiss, A. and D. Mitra (1979). Digital adaptive filters: conditions for convergence, rates of convergence, effects of noise and errors arising from the implementation. *IEEE Trans Info. Theory*, **IT-25**, 637.

Willems, J. L. (1970). *Stability Theory of Dynamical Systems*, Nelson, London, p. 113.

Yuan, J. S-C. and W. M. Wonham (1977). Probing signals for model reference identification. *IEEE Trans Aut. Control*, **AC-22**, 530.

APPENDIX A

Proof of Lemma 2.1.

We shall first establish (3a). Since

$$\sum_{k=j}^{j+S} x_k x_k'$$

is singular, there exist constants $\gamma_1,\ldots,\gamma_n$, $\delta_1,\ldots,\delta_m$, not all zero, such that

$$[\gamma_1\ldots\gamma_n \quad \delta_1\ldots\delta_m]\begin{bmatrix} y_k \\ \vdots \\ y_{k-n+1} \\ u_k \\ \vdots \\ u_{k-m+1}\end{bmatrix}=0 \quad k\in[j,j+S]. \tag{A1}$$

If all γ_i are zero, we are done. So suppose not all γ_i are zero, and let γ_p be the first in the sequence $\gamma_1,\gamma_2,\ldots$ which is nonzero, and suppose that the polynomials

$$\alpha(z)\triangleq z^n+\alpha_1 z^{n-1}+\cdots+\alpha_n \tag{A2a}$$

$$\gamma(z)\triangleq \gamma^p z^{n-p}+\cdots+\gamma_n \tag{A2b}$$

have a greatest common divisor of degree t (with $t=0$ corresponding to coprimeness).

Then it is a consequence of the theory of resultants (see Hodge and Pedoe, 1968) for almost the required result) that the matrix R displayed below has row nullity 1

$$R=\underbrace{\begin{bmatrix} 0\ldots\ldots 0 & 1 & \alpha_1\ldots\ldots\alpha_n \\ 0\ldots 0 & 1 & \alpha_1\ldots\ldots\alpha_n \quad 0 \\ & \ddots & \ddots \\ 1 & \alpha_1\ldots\ldots\alpha_n & 0\ldots\ldots 0 \\ 0\ldots\ldots\ldots 0 & \gamma_p\ldots\ldots\gamma_n \\ 0\ldots\ldots\ldots & \gamma_p\ldots\ldots\gamma_n & 0 \\ & \ddots & \ddots \\ \gamma_p\ldots\ldots\ldots\gamma_n & 0\ldots\ldots 0 \end{bmatrix}}_{2n-p-t+1}\begin{matrix} \left.\right\} n-p-t+1 \\ \\ \left.\right\} n-t+1 \end{matrix} \tag{A3}$$

Denote a (nonzero) left null-vector as

$$w'=[\phi_{n-p-t+1}\cdots\phi_1 \quad \psi_{n-t+1}\cdots\psi_1]. \tag{A4}$$

On defining

$$\phi(z)\triangleq\phi_1 z^{n-p-t}+\ldots+\phi_{n-p-t+1} \tag{A5a}$$

$$\psi(z)\triangleq\psi_1 z^{n-t}+\ldots+\psi_{n-t+1} \tag{A5b}$$

the equation $w'R=0$ yields easily

$$\phi(z)\alpha(z)+\psi(z)\gamma(z)=0. \tag{A6}$$

Since $w\neq 0$, we cannot have both $\phi(z)$ and $\psi(z)$ identically zero. Equation (A6) then shows that neither $\phi(z)$ nor $\psi(z)$ can be identically zero. In fact, $\phi(z)$ and $\psi(z)$ are coprime, for the following reason. Let $d(z)$ be the greatest common divisor of $\alpha(z)$ and $\gamma(z)$, and write $\alpha_1(z)=\alpha(z)/d(z)$, $\gamma_1(z)=\gamma(z)/d(z)$. Then $\phi(z)\alpha_1(z)+\psi(z)\gamma_1(z)=0$, or $\phi/\psi=-\gamma_1/\alpha_1$, and since γ_1,α_1 are coprime, $\deg\psi\geq\deg\alpha_1$ with $\deg\psi=\deg\alpha_1$ precisely when ϕ,ψ are coprime. However, from (A5), $\deg\psi\leq n-t=\deg\alpha_1$. Hence ϕ,ψ are coprime.

Now use (A1) and the system equation (1) connecting $\{y_k\}$ to $\{u_k\}$ to obtain, for $k=j,j+1,\ldots,j+S-(n-t)$

$$\begin{matrix} n-p-t+1 \\ \\ \\ \\ \\ \end{matrix}\underbrace{\begin{bmatrix} 0\ldots\ldots\ldots 1 & \alpha_1\ldots\ldots\alpha_n \\ 0\ldots\ldots 1 & \alpha_1\ldots\ldots\alpha_n \quad 0 \\ \vdots \;\; \ddots & \ddots \\ 1 \quad \alpha_1\ldots\alpha_n & \\ & \gamma_p\ldots\ldots\gamma_n \\ \ddots & \gamma_p\ldots\ldots\gamma_n \\ \gamma_p\ldots\ldots\gamma_n & \end{bmatrix}}_{2n-p-t+1}\begin{bmatrix} y_{k+n-p-t+1} \\ \vdots \\ \vdots \\ y_{k-n+1}\end{bmatrix}$$

$$
= \underbrace{\begin{bmatrix} 0 \cdots\cdots 0 & -\beta_1 \cdots\cdots -\beta_m & \\ 0 \cdots 0 & -\beta_1 \cdots\cdots -\beta_m & \\ & \delta_1 \cdots\cdots \delta_m & \\ & \delta_1 \cdots\cdots \delta_m & 0 \\ \delta_1 \cdots\cdots \delta_m & & \end{bmatrix}}_{m+n-t} \begin{bmatrix} -u_{k+n-t} \\ \vdots \\ -u_k \\ \vdots \\ -u_{k-m+1} \end{bmatrix} \tag{A7}
$$

The equations embodied in the first $n-p-t-1$ rows actually hold for all k, while those in the last $n-t+1$ hold for $k \in [j, j+S-(n-t)]$.

Premultiply by the row vector w'. There results

$$
w' \begin{bmatrix} 0 \cdots\cdots 0 & -\beta_1 \cdots -\beta_m \\ 0 \cdots 0 \;\; -\beta_1 \cdots\cdots -\beta_m & \\ & \delta_1 \cdots \delta_m \\ \delta_1 \cdots\cdots \delta_m & \end{bmatrix} \begin{bmatrix} u_{k+n-t} \\ \vdots \\ u_k \\ \vdots \\ u_{k-m+1} \end{bmatrix} = 0 \tag{A8}
$$

for $k \in [j, j+S-(n-t)]$. We shall now argue that

$$
v' \triangleq w' \begin{bmatrix} 0 \cdots\cdots 0 & -\beta_1 \cdots\cdots -\beta_m \\ 0 \cdots -\beta_1 \cdots\cdots -\beta_m & \\ & \delta_1 \cdots\cdots \delta_m \\ \delta_1 \cdots\cdots \delta_m & \end{bmatrix} \neq 0. \tag{A9}
$$

If v' were zero, we should have, as shown by an easy calculation

$$
\phi(z)\frac{\beta(z)}{z^{n-m}} + \psi(z)\delta(z) = 0 \tag{A10}
$$

where $\beta(z) \triangleq \beta_1 z^{n-1} + \ldots + \beta_m z^{n-m}$ and $\delta(z) \triangleq \delta_1 z^{m-1} + \ldots + \delta_m$. Now (A6) and the coprimeness of $\psi(z)$ and $\phi(z)$ imply $\psi(z)$ is a divisor of $\alpha(z)$, while if (A10) were true, $\psi(z)$ is a divisor of the polynomial $\beta(z)/z^{n-m}$. Hence $\alpha(z)$ and $\beta(z)$ would have a common divisor $\psi(z)$. Since $\alpha(z)$ and $\beta(z)$ are coprime by assumption, $\psi(z)$ is a constant, but then because $\deg \phi(z) < \deg \psi(z)$, we get $\phi(z) = 0$, a contradiction. Hence (A9) holds. Equivalently, (A8) is not trivial. From (A8), we conclude, encompassing all possibilities for $p \geq 1$ and $t \geq 0$ that there exists a nonzero $(n+m)$-vector θ for which

$$
\theta' \begin{bmatrix} u_{k+n} \\ \vdots \\ u_{k-m+1} \end{bmatrix} = 0 \quad k \in [j, j+S-n]. \tag{A11}
$$

We turn now to proving (3b). The proof is similar to that for (3a) and consequently will be given in abbreviated form. In (A1), if all δ_i are zero, we are done. Assume δ_p is the first in the sequence $\delta_1, \delta_2, \ldots$ which is nonzero, and suppose that β_d is the first in the sequence $\beta_1, \beta_2, \ldots$ which is nonzero. Suppose the polynomials $\beta(z) = \beta_d z^{m-d} + \ldots + \beta_m$ and $\delta(z) = \delta_p z^{m-p} + \ldots + \delta_m$ have a greatest common divisor of degree t, which implies that the following matrix $\bar{R}$ has row nullity 1

$$
\bar{R} = \underbrace{\begin{bmatrix} 0 \cdots\cdots 0 & \beta_d \cdots\cdots\cdots \beta_m \\ 0 \cdots\cdots \beta_d \cdots\cdots\cdots \beta_m & 0 \\ \beta_d \cdots\cdots \beta_m & \\ & \delta_p \cdots\cdots \delta_m \\ & \delta_p \cdots\cdots \delta_m \;\; 0 \\ \delta_p \cdots \delta_m & \end{bmatrix}}_{2m-d-p-t+1} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} m-p-t+1 \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} m-d-t+1 \end{matrix}
$$

Denote a left nullvector by

$$
\bar{w}' = [\bar{\phi}_{m-p-t+1} \cdots \bar{\phi}_1 \quad \bar{\psi}_{m-d-t+1} \cdots \bar{\psi}_1] \tag{A12}
$$

and define $\bar{\phi}(z) = \bar{\phi}_1 z^{m-p-t} + \cdots + \bar{\phi}_{m-p-t+1}$ and $\bar{\psi}(z)$ similarly. This ensures that

$$
\bar{\phi}(z)\beta(z) + \bar{\psi}(z)\delta(z) = 0 \tag{A13}
$$

and that $\bar{\phi}(z)$, $\bar{\psi}(z)$ are coprime.

Now use (A1) and the system equation connecting $\{y_k\}$ and $\{u_k\}$ to yield

$$
\begin{bmatrix} 0 \cdots\cdots 0 & \beta_d \cdots\cdots \beta_m \\ \beta_d \cdots\cdots \beta_m & 0 \cdots\cdots 0 \\ 0 \cdots\cdots 0 \cdots\cdots & \delta_p \cdots \delta_m \\ \delta_p \cdots \delta_m & 0 \cdots\cdots 0 \cdots 0 \end{bmatrix} \begin{bmatrix} u_l \\ \vdots \\ u_{k-m+1} \end{bmatrix}
$$

$$
= \begin{bmatrix} 0 \;\; 0 \;\; 1 & \alpha_1 \cdots \alpha_n \\ 1 \;\; \alpha_1 \cdots \alpha_n & \\ & -\gamma_1 \cdots -\gamma_n \\ 0 \cdots 0 & -\gamma_1 \cdots\cdots -\gamma_n \end{bmatrix} \begin{bmatrix} y_{d+l} \\ \vdots \\ y_{k-n+1} \end{bmatrix} \tag{A14a}
$$

where $l = (k+m-d-pt+1)$ in case $p < d+1$ and to yield

$$
\begin{bmatrix} 0 \cdots\cdots 0 & \beta_d \cdots\cdots \beta_m \\ \beta_d \cdots\cdots \beta_m & 0 \cdots\cdots 0 \\ 0 \cdots\cdots 0 \cdots\cdots & \delta_p \cdots \delta_m \\ \delta_p \cdots \delta_m & 0 \cdots\cdots 0 \cdots 0 \end{bmatrix} \begin{bmatrix} u_l \\ \vdots \\ u_{k-m+1} \end{bmatrix}
$$

$$
= \begin{bmatrix} 0 \cdots\cdots 0 \;\; 1 & \alpha_1 \cdots \alpha_n \\ 0 \cdots 1 \;\; \alpha_1 \cdots \alpha_n & \\ 0 \cdots\cdots\cdots 0 & -\gamma_1 \cdots -\gamma_n \\ -\gamma_1 \cdots -\gamma_n & 0 \cdots\cdots\cdots 0 \end{bmatrix} \begin{bmatrix} y_{p+l-1} \\ \vdots \\ y_{k-n+1} \end{bmatrix} \tag{A14b}
$$

in case $p \geq d+1$. Equation (A14a) holds for $k \in [j, j+S-m+d+t)$ and (A14b) for $k \in [j, j+S-m+(p+1)+t]$. Premultiply both equations by $\bar{w}'$. One obtains a nontrivial linear relation among the y_k so long as the following polynomial identity does not hold

$$
\bar{\phi}(z)\alpha(z) - \bar{\psi}(z)\gamma(z) = 0. \tag{A15}
$$

[Here, $\gamma(z)$ is defined as $\gamma_1 z^{n-1} + \ldots + \gamma_n$.] We now deduce a contradiction to (A15): from (A13), and coprimeness of $\bar{\phi}$ and $\bar{\psi}$, $\bar{\psi}$ divides $\beta(z)$. Use this, (A15) and coprimeness of α and β to conclude that $\deg \bar{\psi} = 0$. But in (A15), $\alpha(z)$ has degree n and $\gamma(z)$ degree $(n-1)$, which with $\deg \bar{\psi} = 0$ is a contradiction. Encompassing all possibilities [$d \geq 1$, $t \geq 0$, $p \geq 1$ with $p \geq d+1$ in case (A14b) holds], the nontrivial relation among the y_j can be written as

$$
\bar{\theta}'[y_{k+m} \; y_{k+m-1} \cdots y_{k-n+1}]' = 0
$$

for $k \in [j, j+S-m+1]$. In case d is known, the same result follows for $k \in [j, j+S-m+d]$. □

Proof of Theorem 2.2

Assume that (6) holds, while (5) does not, in order to obtain a contradiction. Because (6) ensures that $\{u_k\}$ is bounded and $\{y_k\}$ is bounded by hypothesis, the upper bound in (5) holds. So assume the lower bound fails, i.e. for arbitrary $\epsilon > 0$, there exists a j and a vector $[\gamma_1 \ldots \gamma_n \delta_1 \ldots \delta_m]$ of unit length such that

$$[\gamma_1 \dots \gamma_n\, \delta_1 \dots \delta_m]x_k = O(\epsilon) \quad k \in [j, j+2n+m-2].$$

The argument used for proving Lemma 2.1 carries through to yield the existence of an $(n+m)$-vector θ of unit norm such that

$$\theta'[u_{k+n} \dots u_{k-m+1}] = O(\epsilon)$$

for $k \in [j, j+n+m-1]$. Since ϵ is arbitrarily small, we violate (6). □

Proof of Theorem 2.3

We again proceed by contradiction. The upper bound in (7) is guaranteed by the upper bound in (8) and by hypothesis. Accordingly, suppose the lower bound fails, i.e. for arbitrary $\epsilon > 0$, there exists a j and a vector $[\gamma_1 \dots \gamma_n\, \delta_1 \dots \delta_n]$ of unit length such that

$$\sum_j^{j+S} E\{\|[\gamma_1 \dots \gamma_n\, \delta_1 \dots \delta_n]x_k\|^2\} < \epsilon.$$

Define p and t as in the proof of Lemma 2.1, and let R and S denote the matrices on the left-hand and right-hand sides of (A7), respectively. Then we can argue that

$$\sum_j^{j+S} E\{\|R\begin{bmatrix} y_{k+n-p-t} \\ \vdots \\ y_{k-n+1}\end{bmatrix} - S\begin{bmatrix} -u_{k+n-t} \\ \vdots \\ -u_{k-m+1}\end{bmatrix}\|^2\} < \epsilon.$$

Introducing w as before leads to

$$\sum_j^{j+S-n+1} E\{\|\frac{\theta'}{\|\theta\|}\begin{bmatrix} u_{k+n} \\ \vdots \\ u_{k-m+1}\end{bmatrix}\|^2\} < \epsilon$$

i.e. the lower bound in (8) fails. □

APPENDIX B

Derivation of alternative plant representation (40)

Since $\alpha(z^{-1})$ and z^{-d}, regarded as polynomials in z^{-1}, are coprime, there exist unique polynomials $F(z^{-1})$, $G(z^{-1})$, with $F(\cdot)$ of degree $< d$ such that

$$1 = F(z^{-1})\alpha(z^{-1}) + G(z^{-1})z^{-d}.$$

[Generically, $F(\cdot)$ has degree $d-1$ and $G(\cdot)$ has degree $n-1$.] Then

$$\frac{\beta(z^{-1})}{\alpha(z^{-1})} = \frac{F(z^{-1})\beta(z^{-1})}{1 - z^{-d}G(z^{-1})}.$$

The coefficients of $F(z^{-1})\beta(z^{-1})$ and $G(z^{-1})$ yield the b_i and a_i in (40).