*Automatica*, Vol. 13, pp. 401–408. Pergamon Press. 1977. Printed in Great Britain

Brief Paper

# An Approach to Multivariable System Identification*

BRIAN D. O. ANDERSON†

**Key Word Index**—Identification; linear systems; adaptive systems; multivariable systems; multivariable control systems; model reference.

**Summary**—Two prototype identifiable structures are presented which make possible the identification via an equation-error model reference adaptive system of linear plants with rational transfer function matrices. The structures include as specialisations many of the particular structures presented hitherto in the literature. Convergence properties are also discussed, and several modes of convergence are distinguished: model output to plant output, model transfer function matrix to plant transfer function matrix, and model parameters to plant parameters. Conditions are presented for exponentially fast convergence in the absence of noise.

## 1. *Introduction*

IN THIS paper, we study schemes for model reference adaptive identification. The schemes are of the series-parallel, or output-error type[1]. At the outset, it must be acknowledged that such schemes have a serious deficiency in that in the presence of noise, they generally provide biased estimates of quantities of interest. Nonetheless, they have found application, perhaps because in general convergence properties are often better than those schemes which, provided they converge, give unbiased estimates.

The first point the paper emphasises is that virtually all equation error schemes in the literature are examples of the identification of one of two prototype structures, each scheme being obtained by a specialisation of certain parameters. The prototype structures therefore unify; at the same time, they help make much clearer the procedures which may be used for tackling multivariable plants. They may also stimulate the development of further equation error schemes with appealing properties.

Having discovered what structures are identifiable, one can then consider the problem of casting an identification problem into a form consistent with the structures. The second main point of the paper is that the problem of identification of stable transfer function matrices with entries having unknown denominator and numerator can be cast into the form of identifying either of the two prototypes. This involves showing that a parameterisation of transfer function matrices can be found consistent with the identifiable structures.

The preceding two paragraphs suggest it is helpful to view *any* identification problem as a pair of problems: finding identifiable structures, and finding a parameterisation consistent with the structure.

The third point of the paper is that it is important to distinguish three modes of convergence of model reference adaptive schemes. At the most primitive level, one can seek convergence of the model's prediction of the plant output to the plant output. At the next level, one can seek convergence of the model's prediction of the plant transfer function matrix to the plant transfer function matrix. Finally, one can seek convergence of an adjustable gain matrix in the model to a fixed gain matrix in a certain representation of the plant. The first type of convergence is almost always achieved; the second requires suitably complex inputs and the third requires a parameterisation process to have a certain uniqueness property. Obviously, the first and second convergence aspects are far more important than the third.

The fourth main point of the paper is that with suitable conditions on the plant input, the convergence rate in all three modes of convergence is exponentially fast. As far as we are aware, the first proof of this result was in [2], a precursor and fuller account of some of the issues explored in this paper.

The earlier work which has most influenced this paper is that of Lion[3], Narendra and coworkers as summarised in, for example, [4], Carroll and Lindorff[5] and Parks[6]. The results of Lion, many of which were obtained earlier by Young[7], can be recovered by specializing the first prototype structure, and those of [4, 5] by specializing the second structure. The second structure involves positive real matrices: it was Parks[6] who alerted workers to the implications of involving positive real functions in the identification process, an idea subsequently developed by many others, including the authors of [4], and [5]. Another important aspect of the problem of obtaining identifiable structures can be found in for example, [5]; by exploring the notion of state-variable filters perhaps due originally to Rucker[8], reference [5] suggests the possibility of whole classes of identification schemes of which schemes in the literature are representative examples.

Nonspecialists would also find very helpful the surveys [9], placing model reference adaptive identification schemes along side other identification schemes, and [1], with a very extensive and thorough discussion of model reference adaptive principles, theories and applications. Another recent survey can be found in [10].

The paper is structured as follows. In Section 2, the identifiable prototype structures are presented, and it is shown how they may accommodate the problem of transfer function identification. In Section 3, the adaptive law and its convergence properties are explored for the two structures. Section 4 contains a number of miscellaneous comments. The Appendix contains the bulk of the convergence proofs.

## 2. *Prototype situations for identification*

In this section, we argue the following points.

1. There are two prototype structures which include the great majority of structures used for adaptive identification which are of the equation error type (or series-parallel MRAS). For specializations of the structures, see e.g. [3–5, 7, 10–16].

2. These structures are appropriate for tackling the problem of transfer function matrix identification for multivariable systems. Further, results on unique parameterization of transfer function matrices can be obtained.

Convergence aspects of the identifiers are discussed in the next section.

*The first prototype structure.* Let us first make contact with known ideas. Numerous workers have considered the problem of identifying the numerator and denominator coefficients of a stable transfer function $W^p(s)$ of known order, linking the Laplace transforms of quantities $v^p(\cdot)$ and $y^p(\cdot)$, assumed measurable, according to $Y^p(s) = W^p(s)V^p(s)$. Various schemes have been proposed which involve driving a dynamic system (termed the model) with $v^p(\cdot)$ and $y^p(\cdot)$ and adjusting coefficients within the model using an error signal derived from $y^p(\cdot)$ and a further signal $y^m(\cdot)$ generated by the model. The adjustment process ceases when $y^m = y^p$, and at this time one aims to have the adjustable model coefficients equal to the coefficients of the plant transfer function.

*A great many of these schemes are of the general arrangement of Fig. 1.* Both $M(s)$ and $N(s)$ are transfer function matrices (having

*Received 16 July 1974; revised 26 August 1976; revised 11 February 1977. The original version of this paper was not presented at any IFAC meeting. It was recommended for publication in revised form by associate editor I. Landau.

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

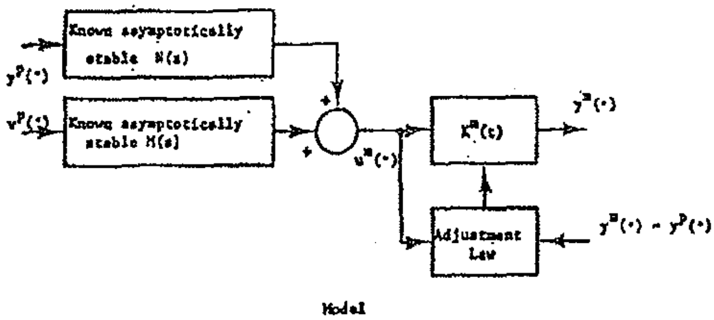


FIG. 1. The first prototype structure.

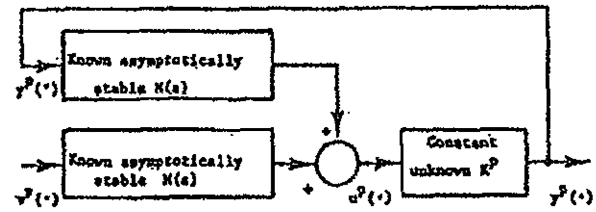


FIG. 2. System with same transfer function matrix as plant.

vector outputs) and in particular schemes, often take particular values e.g.

$$N(s)=\left[\frac{a}{s+a}\ \frac{a^2}{(s+a)^2}\cdots\frac{a^n}{(s+a)^n}\right],$$

see Young[7]. The blocks labelled $M(\cdot)$ and $N(\cdot)$ are often termed state variable filters.

To explain why this model might do as is claimed, consider Fig. 2. The transfer function of the arrangement shown is

$$K^p[I-N(s)K^p]^{-1}M(s)=[I-K^pN(s)]^{-1}K^pM(s). \tag{2.1}$$

It turns out that provided $M(\cdot)$ and $N(\cdot)$ are taken to have suitable dimensions and complexity, one may so arrange that for every $W^p(s)$ there is a $K^p$ such that

$$W^p(s)=[I-K^pN(s)]^{-1}K^pM(s). \tag{2.2}$$

In other words, any plant is input–output equivalent to a scheme of the form of Fig. 2 and all the unknownness in the plant transfer function matrix $W^p(\cdot)$ is located in $K^p$. Further, if one takes care, one can ensure that to any one $W^p(s)$, there corresponds only one $K^p$. Identifying $W^p(s)$ is then equivalent to identifying $K^p$. The model reconstructs the input to the $K^p$ block. [Asymptotically, $u^m(t)=u^p(t)$ because $M(s)$ and $N(s)$ are asymptotically stable and driven by the same signals in the model and the system with transfer function $W^p(s)$]. Using the difference between the output of the $K^p$ block and the output of the model's estimate $K^m(t)$ of the $K^p$ block, the model adjusts its estimate $K^m(t)$ until there is zero error. We regard identification as having occurred once there is zero error.

There are evidently two crucial aspects to the identification problem, the first being to locate the unknownness of $W^p(\cdot)$ in $K^p$ and the second to devise the adjustment law and describe its convergence properties. We shall discuss the first aspect further in this section, and the second aspect in the next section.

*Scalar plants.* The following result describes an extremely wide choice available for $M(\cdot)$ and $N(\cdot)$. *Earlier devices in the literature appear to correspond to specialisations of the theorem.*

*Theorem* 1. Let $F$ be a fixed $n\times n$ matrix and $g$ a fixed $n$-vector, both arbitrary save that $[F, g]$ is completely controllable and $\mathrm{Re}\,\lambda_i[F]<0$. Set

$$N(s)=\begin{bmatrix}(sI-F)^{-1}g\\ 0\end{bmatrix}\quad M(s)=\begin{bmatrix}0\\ (sI-F)^{-1}g\end{bmatrix}. \tag{2.3}$$

Let $K^p=[k_1'\ k_2']$ where $k_1$, $k_2$ are $n$-vectors. Then to every transfer function $W^p(s)$ with denominator of degree $n$ and numerator of degree $(n-1)$ or less, there corresponds a $K^p$ such that (2.2) holds.

*Proof.* By direct calculation, we have for (2.2) and (2.3)

$$\begin{aligned}W^p(s)&=\frac{k_2'(sI-F)^{-1}g}{1-k_1'(sI-F)^{-1}g}\\ &=k_2'[sI-(F+gk_1')]^{-1}g.\end{aligned} \tag{2.4}$$

It is then easily seen that any denominator of $W^p(\cdot)$ is obtainable by adjustment of $k_1$; because of the complete controllability of $[F, g]$; then any numerator is obtained by adjustment of $k_2$. ∇∇

*Remark.* Suppose $W^p(s)$ is of order $n'<n$ and one acts as if it has dimension $n$ by seeking $k_1$, $k_2$ to satisfy (2.4). Then vectors $k_1$, $k_2$ certainly exist, but they are not unique. One can write $W^p(s)$ with a denominator polynomial of degree $n$ and with an arbitrary common factor of degree $(n-n')$ between denominator and numerator. For each choice of common factor, there will be a different $k_1$, $k_2$ pair.

*Multivariable plants—unique parameterization.* There is no intrinsic property of the arrangements of Fig. 1 or 2 which requires $W^p(\cdot)$ to be scalar. We indicate now the extension to matrix $W^p(\cdot)$, dealing first with the case when $W^p(\cdot)$ gives rise to a unique $K^p$. Suppose $W^p(\cdot)$ is $q\times r$ and the degree $n$ of the least common denominator of all entries of $W^p(s)$ is known. Then

$$W^p(s)=\frac{\sum_{i=1}^{n}B_is^{i-1}}{s^n+\sum_{i=1}^{n}a_is^{i-1}} \tag{2.5}$$

for some scalar $a_i$ and matrices $B_i$ which are unknown but fixed uniquely by $W^p(s)$.

*Theorem* 2. Let $F$ be a fixed $n\times n$ matrix and $g$ a fixed $n\times n$ vector, both arbitrary save that $[F, g]$ is completely controllable and Re $\lambda_i[F]<0$. Set*

$$N(s)=\begin{bmatrix}(sI-F)^{-1}g\otimes I_q\\ 0\end{bmatrix}$$

$$M(s)=\begin{bmatrix}0\\ (sI-F)^{-1}g\otimes I_r\end{bmatrix}.$$

Let

$$K^P=[k_{11}I_q\ldots k_{1n}I_q \vdots K_{21}\ldots K_{2n}] \qquad (2.6)$$

where the $k_{1j}$ are scalar and $K_{2j}$ are $q\times r$ matrices. Then to every transfer function matrix $W^P(s)$ of the form of (2.5) there corresponds a $K^P$ such that (2.2) holds, with $K^P$ unique in case $n$ is the degree of the least common denominator of entries of $W^P(s)$; in case $W^P(s)$ has least common denominator of degree $n'<n$, there corresponds an infinity of $K^P$ such that (2.2) holds.

*Proof.* We first consider $[I-K^PN(s)]^{-1}$. Set $k_1'=[k_{11}\ldots k_{1n}]$. Then $K^PN(s)=[k_1'\otimes I_q][(sI-F)^{-1}g\otimes I_q]=k_1'(sI-F)^{-1}g\otimes I_q=[k_1'(sI-F)^{-1}g]I_q$. This yields

$$[I-K^PN(s)]^{-1}K^PM(s)$$
$$=[K_{21}\ldots K_{2n}][(sI-\overline{F+gk_1'})^{-1}g\otimes I_r].$$

Choose $k_1$ so that

$$\det[sI-(F+gk_1')]=s^n+\sum_{i=1}^{n}a_is^{i-1},$$

with $k_1$ existing because of the controllability of $[F,g]$. Let $T$ be such that $(TFT^{-1},Tg)$ is in controllable canonical form, i.e. $TFT^{-1}$ is a companion matrix and $g=[0\ldots0\ \ 1]'$. Then

$$(sI-\overline{F+gk_1'})^{-1}g\otimes I_r$$
$$=\frac{1}{s^n+\sum_{i=1}^{n}a_is^{i-1}}T^{-1}\begin{bmatrix}1\\ s\\ \vdots\\ s^{n-1}\end{bmatrix}\otimes I_r$$
$$=\frac{1}{s^n+\sum_{i=1}^{n}a_is^{i-1}}(T^{-1}\otimes I_r)\begin{bmatrix}1\\ s\\ \vdots\\ s^{n-1}\end{bmatrix}\otimes I_r$$

Now choose the $K_{2i}$ as

$$[K_{21}\ K_{22}\ldots K_{2n}]=[B_n\ B_{n-1}\ldots B_1](T\otimes I_r).$$

It follows that

$$[K_{21},K_{22}\ldots K_{2n}][(sI-\overline{F+gk_1'})^{-1}g\otimes I_r]$$
$$=\frac{1}{s^n+\sum_{i=1}^{n}a_is^{i-1}}[B_n,B_{n-1}\ldots B_1]\begin{bmatrix}1\\ s\\ \vdots\\ s^{n-1}\end{bmatrix}\otimes I_r$$
$$=\frac{\sum_{i=1}^{n}B_is^{i-1}}{s^n+\sum_{i=1}^{n}a_is^{i-1}}$$

*The Kronecker product $A\otimes B$ is the matrix with $i-j$ block entry $a_{ij}B$, where $A=(a_{ij})$. Given compatible dimensions, $(A\otimes B)(C\otimes D)=AC\otimes BD$.

and so (2.2) holds. That the quantities $k_{1i}$ and $K_{2i}$ are unique unless $W^P(s)$ has least common denominator of degree $n'<n$ is easily seen. ∇∇

To obtain a unique parameterisation above, one assumed knowledge of only one quantity, the degree of the least common denominator of $W(s)$. If one assumes knowledge of more quantities, one can get other unique parameterisations. For example, suppose that the degree $n_j$ of the least common denominator of the elements of the $j$th row of $W(s)$ is known for $j=1,2,\ldots,q$. One then regards the problem of identifying $W(s)$ as $q$ separate tasks of identifying the $j$-th row for $j=1,2,\ldots q$

$$W_j(s)=\frac{\sum_{i=1}^{n_j}B_i^js^{i-1}}{s^{n_j}+\sum_{i=1}^{n_j}a_i^js^{i-1}}.$$

Here $n_j$ is known, the scalars $a_i^j$ are unknown, and the row $r$-vectors $B_i^j$ are unknown. The procedure of Theorem 2 will provide a unique parameterisation. What is interesting is that every entry in the gain matrix $K^{Pj}$ corresponding to $W_j(s)$ is a free parameter, whereas in (2.6) when $q>1$ certain entries are constrained to be zero, or equal.

The structure suggested in the theorem, as opposed to the remarks following the theorem, is not generally the best structure for the following reasons:

(a) In general, but not always, the number of parameters will exceed that of the scheme following the theorem, or that of Guidorzi[25]. The scheme of the theorem requires $n(qr+1)$ parameters, that of the scheme mentioned just above requires

$$\sum_{j=1}^{q}n_j(r+1).$$

(b) The state-space representation identified will not be readily useable for control, since it is in general nonminimal.

(c) There appear to be sensitivity difficulties in case pole-zero near-cancellations occur. While this will be the case with almost any method attempting parameteric identification, including that of the scheme following the theorem or [25], the problem will be heightened by the scheme of the theorem, because more near-pole-zero cancellations can be expected due to the nonmimality of the scheme.

The identification of a multi-input, multi-output system *in state-variable form* is, *in part*, a problem of choosing a good representation. If one is prepared to do structural identification first, as per [25], choosing a good representation is easier.

If however, state-variable representations are not needed, but rather the value of the transfer function at certain frequencies, these sensitivity issues are irrelevant. As argued in the next section, use of sinusoidal inputs will cause model and plant transfer functions evaluated at the input frequencies to approach one another. The data from such identification will be adequate if one is seeking to design controllers according to the British school of multi-variable frequency domain methods.

The key point of this subsection is that almost any model reference adaptive structure will have to take the form of Fig. 2.

*The second prototype structure.* The second prototype structure, see Fig. 3, has its genesis in work of Parks[6] on adaptive control, and appears to have first been used by Lüders and Narendra[11] for adaptive identification. The reason for its use rests in a saving of complexity (initially not obvious) as compared with the first structure. Description of the adjustment law and its convergence properties will be postponed to the next section. Here, our aim is to explain the connection with the task of locating the unknownness of a plant transfer function matrix in $K^P$. As before, $\bar{M}(s)$ and $\bar{N}(s)$ are known and asymptotically stable, while $Z(s)$ is also known and has the property that it is nonsingular for almost all $s$, $Z(\infty)$ is finite and symmetric and $Z(s-\sigma)$ is positive real for some $\sigma>0$,[17]. These properties are enough to ensure the existence of a satisfactory adjustment law, but for the purposes of getting nice parameterizations, we shall further restrict $\bar{M}(s)$, $\bar{N}(s)$ and $Z(s)$. In the process, we extend the

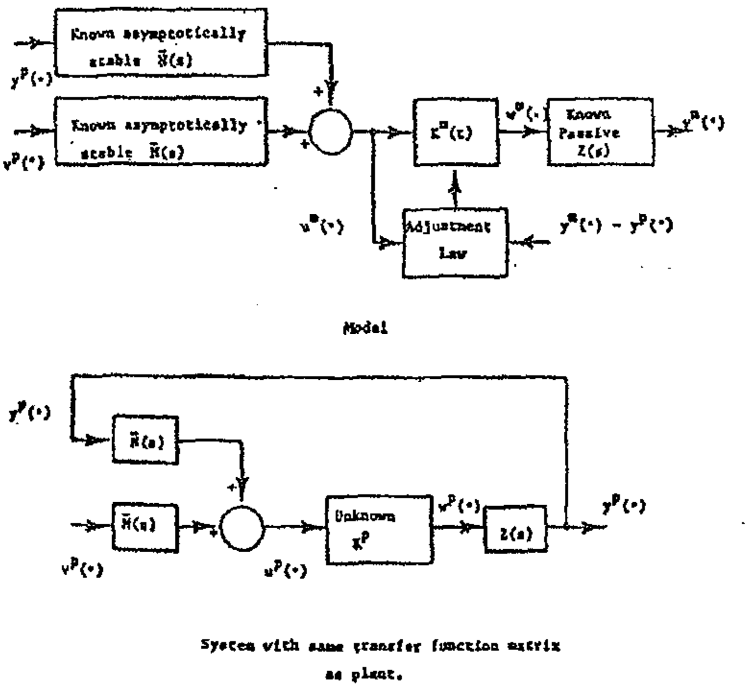


FIG. 3. The second prototype structure

ideas of [11–13] to the multivariable situation. Some multivariable ideas can also be found in [4].

*Lemma* 1. With $W^p(s)$ a real rational $q\times r$ transfer function matrix, let $M(s)$, $N(s)$ and $K^p$ be such that

$$W^p(s)=[I-K^pN(s)]^{-1}K^pM(s).$$

Define for arbitrary $\alpha>0$

$$\bar{M}(s)=(s+\alpha)M(s)$$
$$\bar{N}(s)=(s+\alpha)N(s)$$
$$Z(s)=(s+\alpha)^{-1}I_q.$$

Then

$$W^p(s)=[I-Z(s)K^p\bar{N}(s)]^{-1}Z(s)K^p\bar{M}(s).$$

The proof is trivial. The main use of the idea comes when $\alpha$ is chosen so that $(s+\alpha)$ cancels a pole of $M(s)$ and $N(s)$; thus if $M(s)$ and $N(s)$ are as in Theorem 2, $-\alpha$ is an eigenvalue of $F$. Then $\bar{M}$ and $\bar{N}$ have lower McMillan degree and are easier to implement physically in the model. The penalty paid is the need to implement $Z(s)$, but there is still a saving. Note that $(s+\alpha)^{-1}I_q$ is such that $Z(s-\sigma)$ is positive real for some $\sigma>0$.

*Remark.* If the denominator of a transfer function matrix is known one can dispense with the $\bar{N}(s)$ block and part of $K^p$; if the transfer function matrix is of the form $(sI-A)^{-1}B$, i.e. $y^p(t)$ is a state vector, significant simplification is possible here too. (This case has always been recognized as easier.)

*Remark.* There are many identification problems in which part knowledge of the system equations is available. It would be useful to have a catalog of those situations which could be cast into the form of one of the prototype situations without throwing away the *a priori* knowledge.

*Remark.* Reference [5] introduced the notion of adaptive observers. Remarks concerning adaptive observers applicable to multivariable systems can be found in [2] and [8].

## 3. *Adaptive identification laws*

In this section, our aim is to describe the adjustment laws for the prototype structures described in the last section, and to indicate the convergence properties, including convergence rates. There are actually several different convergence questions

1. Under what conditions does the model input-output behaviour become coincident with the plant input-output behaviour for the particular plant input used, i.e. when does $y^p(\cdot)\to y^m(\cdot)$ for the particular $v^p(\cdot)$?
2. Under what conditions does the 'frozen' transfer function matrix predicted by the model approach that of the plant, i.e. for the first situation, when does $[I-K^m(t)N(s)]^{-1}K^m(t)M(s)\to[I-K^pN(s)]^{-1}K^pM(s)$ for fixed $s$ as $t\to\infty$?
3. Under what conditions does the model gain matrix $K^m(t)$ approach the plant gain matrix $K^p$ as $t\to\infty$?

Until further notice in this section, we shall assume that in the first prototype situation, $M(s)$ and $N(s)$ are asymptotically stable and in the second prototype situation that $\bar{M}(s)$ and $\bar{N}(s)$ are known and asymptotically stable, and that $Z(s)$ is known and nonsingular for almost all $s$, $Z(\infty)$ is finite and symmetric and $Z(s-\sigma)$ is positive real for some $\sigma>0$. We do not need to assume that $M(s)$, $N(s)$ etc. have any of the special forms described in the last section—the convergence results are general.

*Adjustment law.* For both prototype structures, the adjustment law is obtained as follows. Let $k_i^{m'}$ denote the $i$th row of $K^m$ and let $\Lambda_i$ be an arbitrary positive definite symmetric matrix. Then, with $y_i^m$, $y_i^p$ the $i$th entries of $y^m$ and $y^p$,

$$\dot{k}_i^m=-\Lambda_i u^m(t)[y_i^m(t)-y_i^p(t)]. \tag{3.1}$$

It is interesting to note that in the second prototype situation the error signal $y_i^m(t)-y_i^p(t)$ depends on the difference $K^m(\tau)-K^p$ for time $\tau$ earlier than $t$. Because of the integrating effect in $Z(s)$, (3.1) might then be thought of as defining a proportional plus integral adaptation scheme.

*Reformulation of adjustment law—first prototype situation.* Set

$$x_i=(\Lambda_i^{1/2})^{-1}(k_i^m-k_i^p).$$

Here, $\Lambda^{1/2}$ denotes the positive definite symmetric square root of $\Lambda$. Then, since $\dot{k}_i^p=0$, (3.1) yields

$$\dot{x}_i=-\Lambda_i^{1/2}u^m u^{m'}\Lambda_i^{1/2}x_i-\Lambda_i^{1/2}u^m(u^m-u^p)'k_i^p. \tag{3.2}$$

If $v^p$ is bounded, so that $u^m$, $u^p$ and $y^p$ are also bounded, the second term on the right side of (3.2) decays exponentially fast.

*Reformulation of adjustment law—second prototype situation.* For convenience, let us assume all $\Lambda_i$ are the same. Suppose that

$$Z(s) = D + C'(sI - A)^{-1}B$$

with

$$\begin{bmatrix} 2D & C - B \\ (C-B)' & -(A+A') \end{bmatrix} = \begin{bmatrix} 2D & \sqrt{2}LD^{1/2} \\ \sqrt{2}D^{1/2}L' & -LL' - \sigma I \end{bmatrix}$$

for some $L$ and scalar $\sigma > 0$. The existence of such $A$, $B$, $C$, $D$ is guaranteed by the Kalman–Yakubovic lemma[19]. The equations of the adaptive scheme become

$$\dot{\lambda} = A\lambda + B(w^m - w^p)$$

$$y^m - y^p = C'\lambda + D(w^m - w^p)$$

(which describes the $Z(s)$ block) together with the following equation, obtained from (3.1) with

$$x_i = (\Lambda^{1/2})^{-1}(k_i^m - k_i^p)$$

and describing the gain adjustment block:

$$\dot{x}_i = -\Lambda^{1/2}u^m[y_i^m - y_i^p].$$

Rearrangement leads, with $x = [x_1' \ldots x_q' \ \lambda']'$, to

$$\dot{x} = \begin{bmatrix} -D \otimes \bar{u}^m\bar{u}^{m'} & -C' \otimes \bar{u}^m \\ B \otimes \bar{u}^{m'} & A \end{bmatrix} x + \eta(t). \qquad (3.3)$$

Here, $\bar{u}^m$ stands for $(\Lambda^{1/2})^{-1}u^m$ and $\eta(t)$ for a term involving $u^m(t) - u^p(t)$ which is exponentially decaying in case $v^p(\cdot)$ is bounded.

*Convergence of outputs.* Below, we easily show that

$$\int_{t}^{\infty} [y^p - y^m]'[y^p - y^m]\,dt \to 0$$

as $t \to \infty$; with more work we shall conclude that $y^p \to y^m$. To obtain pointwise convergence, we place a limitation on $u^p(\cdot)$ [which can be traced back to a limitation on $v^p(\cdot)$].

Define a class $\mathscr{V}$ of functions $R \to R^a$ for arbitrary $a$ in the following way. For arbitrary $\Delta > 0$, let $C$ be a set of discrete points $\{\ldots t_i, t_{i+1}, \ldots\}$ in $(-\infty, \infty)$ such that $t_{i+1} - t_i > \Delta$ for all $i$. Then $u(\cdot) \in \mathscr{V}$ if $u(\cdot)$ and $\dot{u}(\cdot)$ are continuous and bounded on $R - C_\Delta$ for some $\Delta$ and $C_\Delta$, and $\lim_{t \uparrow t_i}$ exist and are finite for $u$, $\dot{u}$, and all $t_i$. Obviously if $v^p$ and $\dot{v}^p$ are continuous and bounded, $u^p \in \mathscr{V}$. But looser conditions on $v^p$ suffice. Thus $v^p(\cdot)$ could, for example, be a square wave. (In our earlier work on convergence, [2], $u^p(\cdot)$ had to be smooth.)

The main result is then as follows:

*Theorem* 3. (Model Following Result) Suppose that the above conditions on $M(\cdot)$ etc. are in force, and that the adjustment law (3.1) is used. Then if $u^p$ is bounded,

$$\int_{t}^{\infty} [y^p - y^m]'[y^p - y^m]\,dt \to 0$$

and if $u^p \in \mathscr{V}$,

$$y^p - y^m \to 0$$

in both prototype situations.

For a proof, see the appendix.

Notice that the above theorem makes no assumptions of a persistently or completely exciting nature, and draws no conclusions about the rate of convergence or transfer function identification.

*Convergence rate.* To obtain a result on the convergence of $y^m(t)$ to $y^p(t)$ we shall assume that $v^p(\cdot)$ is a (possibly infinite) linear combination of periodic functions:

$$v^p(t) = \sum_i \mathrm{Re}[\exp(j\omega_i t)V_i] \quad \omega_i \neq \omega_j \text{ for } i \neq j. \qquad (3.4)$$

Moreover, we shall suppose that $u^p \in \mathscr{V}$ in both prototype situations. This is an inessential restriction. Notice that $v^p(\cdot)$ is not necessarily periodic or almost periodic*—take $v^p(\cdot)$ as a sum of two square waves of incommensurate frequencies.

*Theorem* 4. Suppose the conditions of Theorem 3 hold and that $v^p(\cdot)$ is as in (3.4). Let

$$u^m(t) = \tilde{u}^m(t) + 0(e^{-\alpha t}), \quad \alpha > 0,$$

where $\tilde{u}^m(t)$ is a linear combination of sinusoids. Then as $t \to \infty$, $K^m(t)$ approach a limit exponentially fast, $[K^m(t) - K^p]\mathscr{A}[\tilde{u}^m(\cdot)] \to 0$ exponentially fast, and $y^m(t) - y^p(t) \to 0$ exponentially fast. Here $\mathscr{A}[\tilde{u}^m(\cdot)]$ denote the space spanned by $\tilde{u}^m(\tau)$ for all $\tau \in [0, \infty)$.

The theorem is proved in the appendix. The key is to use some recent results[20] on the exponential convergence of equations like (3.2) and (3.3).

*Remark.* As an examination of the proof of the Theorem will show, the crucial property required of $v^p(\cdot)$ is that the resulting $u^p(\cdot)$ be expressible as

$$u^p(t) = -H\begin{bmatrix} v(t) \\ 0 \end{bmatrix} + 0(e^{-\alpha t}) \qquad (3.5)$$

where $H$ is a constant matrix, the entries of $v(t)$ are linearly independent and for some suitably large $T$ and positive $\alpha_1, \alpha_2$

$$\alpha_1 I \leq \frac{1}{T}\int_s^{s+T} v(t)v'(t)\,dt \leq \alpha_2 I \quad \text{for all } s. \qquad (3.6)$$

The condition (3.4) ensures that $u^p(\cdot)$ has the property, and it is of course a natural condition. Obviously, it is not necessary.

*Convergence of transfer function matrices.* Since the plant transfer matrix is, for the first prototype situation,

$$W^p(s) = K^p[I - N(s)K^p]^{-1}M(s)$$

$$= [I - K^p N(s)]^{-1}K^p M(s) \qquad (3.7)$$

it is natural to take as the transfer function matrix predicted by the model at time $t$ the above expression with $K^p$ replaced by $K^m(t)$:

$$W^m(s;t) = K^m(t)[I - N(s)K^m(t)]^{-1}M(s)$$

$$= [I - K^m(t)N(s)]^{-1}K^m(t)M(s). \qquad (3.8)$$

The second prototype situation is obtained by an obvious variation. The question arises as to whether $\lim_{t\to\infty} W^m(s;t) = W^p(s)$ for each fixed $s$. The key to obtaining this convergence is to take $v^p(t)$ to be sufficiently rich; then the convergence of $y^m(t)$ to $y^p(t)$ ensures the transfer function convergence.

As a preliminary result, we have the following:

*Lemma* 2. Under the hypothesis of Theorem 4,

$$\lim_{t\to\infty} W^m(j\omega_i;t)V_i = W^p(j\omega_i)V_i \qquad (3.9)$$

with the convergence rate exponential.

For a proof, see the Appendix.

It then follows that with sufficient variety among the $\omega_i$ and the $V_i$ and with (3.9) holding for all $i$, $W^m(j\omega_i;t) \to W^p(j\omega_i)$. A *persistently exciting* condition suffices, see e.g. [21] for such a condition for multivariable systems. An interesting recent

*Almost periodic functions are (possibly infinite) linear combinations of sinusoids which, *inter alia*, are normally taken as continuous.

treatment of the type of complexity condition needed can be found in [22].

A specialisation of this condition can be found in the following way. Let $n$ be the degree of the least common denominator of all entries of $W^p(s)$ and let $e_l$ be the unit vector with 1 in the $l$th position. Take $V_1, V_2, \ldots = V_n = e_1$, $V_{n+1}, \ldots V_{2n} = e_2$, etc., there being $nr$ nonzero frequencies $\omega_i$ involved. Then (3.9) implies

$$j\text{th column of } W^m(j\omega_i; t) = j\text{th column of } W^p(j\omega_i)$$

for $n$ frequencies $\omega_i$. Thus for each entry of $W^p$ and $W^m$ we have equality at $n$ frequencies. This and the upper bound of $n$ on the degree of the denominator ensures that equality occurs for all frequencies. Essentially the same argument holds if $e_1, \ldots e_r$ are replaced by an arbitrary set of $r$ linearly independent $r$-vectors.

*Convergence of gain matrix.* In one sense, the question of whether or not $K^m(t) \to K^p$ is irrelevant. The goal of identification should after all be the determination of an input–output description of the object being identified. What one can say however is that convergence of the gain matrix $K^m(t)$ to $K^p$ is guaranteed to occur if, first, the mapping

$$K^p \to [I - K^p N(s)]^{-1} K^p M(s) = W^p(s)$$

is one-to-one i.e. is $W^p(s)$ is uniquely parameterisable by one $K^p$, and second, $W^m(s; t) \to W^p(s)$.

In the last section, a parameterisation of a matrix $W^p(s)$ was described in which certain entries of $K^p$ were equal and others were zero. The convergence theory of this section needs minor adjustment to cope with the incorporation of a similar constraint in $K^m(t)$ for all $t$ if one is to guarantee that when $W^m(s; t) \to W^p(s)$, one also has $K^m(t) \to K^p$.

*Remark.* Exponentially fast convergence of the transfer function matrix and even of the gain matrix can occur with inputs other than linear combinations of sinusoids. The input $v^p(\cdot)$ must be sufficiently complex as to ensure that if $y^p(t) - y^m(t) \to 0$ then $W^m(s; t) \to W^p(s)$ and it must have a persistently exciting property carrying over to a property on $u^p(\cdot)$ embodied in (3.5) and (3.6) in order to guarantee convergence is exponential.

*Remark.* Lemma 2 is consistent with one's intuition that input signals in the passband of the system being identified will be better for identification purposes than those outside. The best choice of input signal is however a major study in itself.

*Remark.* Arbitrarily fast convergence can theoretically be obtained by using more complicated $M(s)$ and $N(s)$. Lion's work[4] explains the core of the idea, and Kreisselmeier[16] develops it further. One would however expect in practice noise to be a more severe limitation the faster one tried to identify.

*Remark.* It is highly likely that if the $Z(s)$ block were replaced by a possibly nonlinear, strictly passive block (perhaps even involving relays), convergence could still be proved.

## 4. *Conclusions*

In this paper we have taken some mildly unconventional approaches to adaptive identification. Sometimes explicitly and frequently implicitly we have been arguing several points:

1. Finding identifiable structures is a problem in itself.
2. Organising an identification problem into the form of an identifiable structure is a problem in itself.
3. There are several types of convergence to think about, the most important normally being transfer function matrix convergence, *not* gain convergence.
4. Convergence rates can be assured to be exponential by appropriate constraints on the plant input.

One would expect the results carry over to discrete time. Some of them can be found in [23]. There are several questions however of great interest still remaining.

1. How robust are the algorithms in their tolerance of nonlinearities, noise, and parasitic modes of the plant? [To an extent, simulations can answer this question, but are unlikely to give general results.]
2. Can one obtain results for identifying unstable plants?
3. How should one choose input signals?
4. To what extent can one replace the $Z(s)$ block in the second prototype scheme by another possibly nonlinear passive block (perhaps allowing switchings)?

*Acknowledgement*—Work supported by the Australian Research Grants Committee.

*References*

[1] I. D. Landau: A survey of model reference adaptive techniques—theory and applications. *Automatica* **10**, (4) 353–380 (1974).
[2] B. D. O. Anderson: Multivariable adaptive identification. Tech. Report EE7402, University of Newcastle (1974).
[3] P. M. Lion: Rapid identification of linear and nonlinear systems. *AIAA Jl* **5**, (10) 1835–1842 (1967).
[4] K. S. Narendra and P. Kudva: Stable adaptive schemes for system identification and control—part II. *IEEE Trans. Systems, Man and Cybernetics*. SMC-4, (6) 552–560 (1974).
[5] R. L. Carroll and D. P. Lindorff: An adaptive observer for single-input, single-output linear systems. *IEEE Trans. Aut. Control* AC-18, (5) 428–435 (1973).
[6] P. C. Parks: Lyapunov redesign of model reference adaptive control systems. *IEEE Trans. Aut. Control* AC-11, (3) 362–367 (1966).
[7] P. C. Young: Discussion on 'In-flight dynamic checkout' by J. E. Walker. *IEEE Trans. Aerospace* AS-2, (3) 1106–1111 (1964).
[8] R. A. Rucker: Real time system identification in the presence of noise. *Proc. Western Electronics Conv.* Paper 2–3, August (1963).
[9] K. J. Åström and P. Eykhoff: System identification, a survey. *Automatica* **7**, (2) 123–162 (1971).
[10] R. L. Carroll and R. V. Monopoli: Model reference adaptive control, estimation and identification using only input and output signals. *Proc. 6th IFAC World Congress*, Boston, Part I, Paper 58.3 (1975).
[11] G. Lüders and K. S. Narendra: An adaptive observer and identifier for a linear system. *IEEE Trans. Aut. Control* AC-18, (5) 496–499 (1973).
[12] G. Lüders and K. S. Narendra: A new canonical form for an adaptive observer. *IEEE Trans. Automatic Control* AC-19, (2) 117–119 (1974).
[13] P. Kudva and K. S. Narendra: Synthesis of an adaptive observer using Lyapunov's direct method. *Int. J. Control* **18**, (6) December 1201–1210 (1973).
[14] K. S. Narendra and P. Kudva: Stable adaptive schemes for system identification and control—part I. *IEEE Trans. on Systems, Man and Cybernetics* AMC-6, (6) 542–551 (1974).
[15] D. O. Molnar: New form for an adaptive observer. *AIAA J.* **13**, (2) 243–244 (1975).
[16] G. Kreisselmeier: Adaptive observers with arbitrary exponential rate of convergence. *IEEE Trans. Aut. Control*, to appear.
[17] R. W. Newcomb: *Linear Multiport Systhesis*. McGraw-Hill, New York. (1966).
[18] B. D. O. Anderson: Adaptive identification of multiple-input multiple-output plants. *Proc. 1974, IEEE Conf. on Decision and Control* pp. 273–281.
[19] B. D. O. Anderson and S. Vongpanitlerd: *Network Analysis and Synthesis—A Modern System Theory Approach*, Prentice-Hall, Englewood Cliffs, NJ (1973).
[20] B. D. O. Anderson: Exponential stability of linear equations arising in adaptive identification. *IEEE Trans. Aut. Control* AC-22, (1) 83–88 (1977).
[21] D. Q. Mayne: A canonical model for identification of linear multivariable systems. *IEEE Trans. Aut. Control* AC-17, (5) 728–729 (1972).
[22] A. V. Mathew and F. W. Fairman: Tranfer function matrix identification. *IEEE Trans. on Circuits and Systems* CAS-21, (5) 584–588 (1974).
[23] P. Kudva and K. S. Narendra: The discrete adaptive observer. *Proc. 1974 IEEE Conf. on Decision and Control*, Phoenix, pp. 307–312.
[24] J. S. C. Yuan and W. M. Wonham: Asymptotic identification using stability criteria, part I: proving signal

description. University of Toronto Control Systems, Report 7422, November (1974).
[25] R. Guidorzi: Canonical structures in the identification of multivariable systems. *Automatica* 11, (4) 361–374 (1975).

## *Appendix*

### *Convergence results*

*Convergence of outputs.* To prove Theorem 3, we need a lemma. It is well known that if $g(\cdot)$ is a function for which $\int_0^\infty g(t)\,dt$ is finite and for which $g, \dot{g}$ are continuous and bounded, then $g(t) \to 0$ as $t \to \infty$. The lemma, due to Yuan and Wonham[24] generalizes this.

*Lemma* A1. If $\int_0^\infty g(t)\,dt$ is finite and $g(\cdot) \in \mathscr{U}$, then $g(t) \to 0$ as $t \to \infty$.

*Proof of Theorem* 3. *First prototype situation.* Since $y^m - y^p = K^m u^m - K^p u^p$ is a decaying exponential, it is enough to prove that for each $i$ $u^{m'}(k_i^m - k_i^p) = x_i' \Lambda_i^{1/2} u^m$ has the properties demanded of $y^m - y^p$. Set $V = x_i' x_i$. Then (3.2) implies for a certain function $\eta(t)$ approaching zero exponentially fast that

$$\dot{V} = -\sum_i (x_i' \Lambda_i^{1/2} u^m)^2 - 2(x_i' \Lambda_i^{1/2} u^m)\eta(t)$$

$$= -\sum_i [x_i' \Lambda_i^{1/2} u^m + \eta(t)]^2 + \eta^2(t).$$

From this and the positive definite character of $V$ we see that

$$\int_t^\infty [x_i' \Lambda_i^{1/2} u^m + \eta(t)]^2\,dt$$

and then

$$\int_t^\infty [x_i' \Lambda_i^{1/2} u^m]^2\,dt$$

exist and are finite.

It easily follows that if $u^p \in \mathscr{U}$, then $u^m \in \mathscr{U}$ and then from (3.2) that $x_i \in \mathscr{U}$. In turn $[x_i' \Lambda_i^{1/2} u^m]^2 \in \mathscr{U}$ and finally $y^m \to y^p$ by the lemma.

*Second prototype situation.* Let $V = x'x$ and compute $\dot{V}$ using (3.3). One obtains for some $\alpha > 0$,

$$\dot{V} = -x' \begin{bmatrix} -2D \otimes \bar{u}^m \bar{u}^{m'} & (B' - C') \otimes \bar{u}^m \\ (B - C) \otimes \bar{u}^{m'} & -LL' - \sigma I \end{bmatrix} x + 0(e^{-\alpha t})$$

$$-x' \begin{bmatrix} \sqrt{2} D^{1/2} \otimes \bar{u}^m & 0 \\ L & -\sigma^{1/2} I \end{bmatrix} \begin{bmatrix} \sqrt{2} D^{1/2} \otimes \bar{u}^m & 0 \\ L & -\sigma^{1/2} I \end{bmatrix}' x + 0(e^{-\alpha t})$$

$$= [\sqrt{2}(w^m - w^p)' D^{1/2} + \lambda' L \quad -\sigma^{1/2} \lambda'] \begin{bmatrix} \sqrt{2} D^{1/2}(w^m - w^p) + L'\lambda \\ -\sigma^{1/2} \lambda \end{bmatrix} + 0(e^{-\alpha t}).$$

Using arguments as for the first prototype situation, one concludes that

$$\int_t^\infty \lambda' \lambda\,dt < \infty$$

and

$$\int_t^\infty (w^m - w^p)' D (w^m - w^p)\,dt < \infty.$$

Since

$$y^m - y^p = D(w^m - w^p) + C'\lambda,$$

it is clear that

$$\int_t^\infty (y^m - y^p)'(y^m - y^p)\,dt < \infty.$$

That $y^m - y^p \to 0$ if $u^p \in \mathscr{U}$ is proved essentially as for the first prototype structure.

An easily proved corollary is

*Corollary.* Under the hypothesis that $u^p$ is bounded, $x_i$ is bounded in the first prototype situation and $x$ is bounded in the second prototype situation.

*Exponential convergence.* The following is a standard result

*Lemma* A.2. Suppose that $\dot{x} = A(t)x$ is exponentially asymptotically stable. Let $C(t) \to 0$ exponentially fast. Then the solution of $\dot{x} = A(t)x + C(t) \to 0$ exponentially fast.

*Proof of Theorem* 4. *First prototype situation.* With the condition on $u^p(\cdot)$, there exists $\bar{u}^m(t)$ which is a linear combination of sinusoids and such that

$$u^m(t) - \bar{u}^m(t) = 0(e^{-\alpha t})$$

and

$$u^p(t) - \bar{u}^m(t) = 0(e^{-\alpha t})$$

for some $\alpha > 0$.

Let $\Omega$ be a constant orthogonal matrix such that

$$\Lambda_i^{1/2} \bar{u}^m(t) = \Omega \begin{bmatrix} v(t) \\ \hline 0 \end{bmatrix}$$

where the entries of $v(\cdot)$ are linearly independent functions and linear combinations of sinusoids. Equation (3.2) then yields

$$\dot{x}_i = -\Omega \begin{bmatrix} v(t)v'(t) & 0 \\ 0 & 0 \end{bmatrix} \Omega' x_i + 0(e^{-\alpha t}) x_i + 0(e^{-\alpha t}).$$

Set

$$\Omega' x_i = \begin{bmatrix} \bar{x}_{i1} \\ \bar{x}_{i2} \end{bmatrix}.$$

Then

$$\dot{\bar{x}}_{i1} = -v(t)v'(t)\bar{x}_{i1} + 0(e^{-\alpha t})\bar{x}_{i1} + 0(e^{-\alpha t})\bar{x}_{i2} + 0(e^{-\alpha t}).$$

Since $\bar{x}_i$ is bounded,

$$\dot{\bar{x}}_{i1} = -v(t)v'(t)\bar{x}_{i1} + 0(e^{-\alpha t}).$$

Similarly, we obtain

$$\dot{\bar{x}}_{i2} = 0(e^{-\alpha t}).$$

The properties of $v(\cdot)$ ensure that for suitably large $T$, there exist positive $\alpha_1, \alpha_2$ such that

$$\alpha_1 I \leq \frac{1}{T} \int_s^{s+T} v(t)v'(t)\,dt \leq \alpha_2 I \quad \text{for all} \quad s.$$

By a result of [20], this and Lemma A2 ensures that $\bar{x}_{i1}(t) \to 0$ exponentially fast.

The exponential bound on $\dot{\bar{x}}_{i2}$ also guarantees that $\lim_{t\to\infty} \bar{x}_{i2}(t)$ exists and is approached exponentially fast.

Since $k_i^m(t) - k_i^p = \Lambda_i^{1/2} x_i$, $\lim_{t\to\infty} k_i^m(t)$ exists.

Since

$$\bar{x}_{i1} = x_i' \Omega \begin{bmatrix} I \\ 0 \end{bmatrix} = [k_i^m(t) - k_i^p]' \Lambda_i^{-1/2} \Omega \begin{bmatrix} I \\ 0 \end{bmatrix}$$

$$= [k_i^m(t) - k_i^p]' \mathscr{B}[\bar{u}^m(\cdot)]$$

the second result of the theorem is immediate. Also

$$y_i^m(t) - y^p(t) = [k_i^m(t) - k_i^p]' \bar{u}^m(t) + 0(e^{-\alpha t})$$

and the third result follows.

*Second prototype situation.* The proof is analogous to that above, relying on a result of [20] for the exponential convergence of equation of the form (3.3).

*Convergence of transfer function matrices*

*Proof of Lemma 2. First Prototype situation.* It is easily checked that with $\tilde{u}^n(t)$ as above,

$$\tilde{u}^n(t) = \mathrm{Re}\left\{\sum_i \exp(j\omega_i t)[I - N(j\omega_i)K^p]^{-1} M(j\omega_i)V_i\right\}.$$

Since the functions $\exp(j\omega_1 t)$ and $\exp(j\omega_2 t)$ are linearly independent whenever $\omega_1 \neq \omega_2$, it follows that

$$\mathscr{R}[\tilde{u}^n(\cdot)] = \mathscr{R}\{[I - N(j\omega_i)K^p]^{-1} M(j\omega_i)V_i \quad \text{for all} \quad i\}.$$

By Theorem 4, for each $i$ one has

$$\lim_{t\to\infty} K^n(t)[I - N(j\omega_i)K^p]^{-1} M(j\omega_i)V_i = K^p[I - N(j\omega_i)K^p]^{-1} M(j\omega_i)V_i \qquad \text{(A1)}$$

and so

$$\lim_{t\to\infty} [I - N(j\omega_i)K^n(t)][I - N(j\omega_i)K^p]^{-1} M(j\omega_i)V_i$$
$$= [I - N(j\omega_i)K^p][I - N(j\omega_i)K^p]^{-1} M(j\omega_i)V_i$$
$$= M(j\omega_i)V_i$$

whence

$$[I - N(j\omega_i)K^p]^{-1} M(j\omega_i)V_i = \lim_{t\to\infty} [I - N(j\omega_i)K^n(t)]^{-1} M(j\omega_i)V_i.$$

Using (A1) again, it follows that

$$\lim_{t\to\infty} K^n(t)[I - N(j\omega_i)K^n(t)]^{-1} M(j\omega_i)V_i = K^p[I - N(j\omega_i)K^p]^{-1} M(j\omega_i)V_i$$

which is (3.9). The proof for the second prototype situation proceeds analogously. Convergence is exponential as a result of Theorem 4.