Automatica, Vol. 22, No. 6, pp. 629–639, 1986
Printed in Great Britain.

0005-1098/86 $3.00+0.00
Pergamon Journals Ltd.

Necessary and Sufficient Conditions for Parameter Convergence in Adaptive Control*

STEPHEN BOYD† and S. S. SASTRY‡

A complete description of parameter convergence in model reference adaptive control may be given in terms of the spectrum of the exogenous reference input signal.

Key Words—Adaptive control; adaptive systems; identifiability; parameter estimation.

Abstract—Using Generalized Harmonic Analysis, a complete description of parameter convergence in Model Reference Adaptive Control (MRAC) is given in terms of the *spectrum* of the exogenous reference input signal. Roughly speaking, if the reference signal "contains enough frequencies" then the parameter vector converges to its correct value. If not, it converges to an easily characterizable subspace in parameter space.

1. INTRODUCTION AND PROBLEM STATEMENT

IN RECENT work (Narendra and Valavani, 1978; Nardendra *et al.*, 1980; Morse, 1980) on continuous time model reference adaptive control, it has been shown that under a suitable adaptive control law the output y_P of the plant asymptotically tracks the output y_M of a stable reference model, despite the fact that the parameter error vector may not converge to zero (indeed, it may not converge at all). Results that have appeared in the literature on parameter error convergence (Morgan, 1977; Anderson, 1977; Kriesselmeier, 1977; Yuan and Wonham, 1977) have established the exponential stability of adaptive schemes under a certain *persistent excitation* (PE) condition. As is widely recognized (e.g. in Anderson and Johnson, 1982) the drawback to this condition is that it applies to a certain vector of signals $w(t)$ appearing inside the non-linear feedback loop around the unknown plant.

In earlier work (Boyd, 1983) this shortfall was remedied by showing that the persistent excitation condition can be moved from w to w_M, a vector of signals analogous to w but appearing in the linear, time invariant (LTI) model loop. Unlike w, w_M is simply *the output of a LTI system driven by the reference signal* r, and it is thus much easier to determine whether or not it is persistently exciting.

In Boyd (1983) one simple condition was given which ensures that w_M is PE:

If the reference input $r(t)$ contains as many spectral lines as there are unknown parameters, then w_M is PE and consequently the model-plant output error and the parameter error converge exponentially to 0.

Note that a *real* reference signal with a spectral line at frequency ν also has a spectral line at $-\nu$. Thus, for example, a reference signal with a (non-zero) average (d.c.) value and at least one other spectral line will guarantee exponential convergence of the parameter error vector to zero in a three parameter MRAC system. Related results for the scheme of Morse (1980) have appeared in Dasgupta *et al.* (1983).

These results made precise the following intuitive argument: assuming the parameter vector *does* converge (but perhaps to the wrong value) the plant loop is "asymptotically time invariant". If the reference input r has spectral lines at frequencies $\nu_1, \ldots, \nu_k$, one expects y_P will also; since $y_P \to y_M$, one "concludes" that the asymptotic closed loop plant transfer function matches the model transfer function at $s = j\nu_1, \ldots, j\nu_k$. If k is large enough, this implies that the asymptotic closed loop transfer function is *precisely* the model transfer function so that the parameter error converges to *zero*.

In this paper, this idea that the reference signal must be "rich enough", i.e. "contain enough frequencies" for the parameter error to converge to zero is pursued further. Simple *necessary and sufficient* conditions on the reference input r for the parameter error to converge to zero are derived. Roughly speaking, the condition is:

* Received 20 February 1985; revised 15 November 1985; revised 25 March 1986; revised 10 June 1986. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor G. Kreisselmeier under the direction of Editor P. C. Parks. The research was supported in part by NASA under grant NAG2-243.

† EE Department, Standford University, Stanford, CA 94035, U.S.A.

‡ EECS Department and the Electronics Research Laboratory, University of California, Berkeley, CA 94720, U.S.A.

A reference input $r(t)$ results in parameter error convergence to zero *unless* its spectrum is concentrated on $k < N$ lines, where N is the number of unknown paramters in the adaptive scheme.

Precisely what is meant by spectrum is detailed in the sequel. The results have been announced without proof in Boyd and Sastry (1984). Related work appears in Narendra and Annaswamy (1983a, b).

In Section 2 the MRAC system is briefly described when the plant has relative degree one. In Section 3 the basic notions of Generalized Harmonic Analysis: autocovariance and spectral measure, are reviewed. In Section 4 the main result, on necessary and sufficient conditions for parameter convergence, is stated and proved.

In Section 5, partial convergence, i.e. behaviour of the parameter error vector when w is *not* PE is discussed. This will be the case when the reference signal has its spectrum concentrated on $k < N$ lines, where N is the number of unknown parameters: then the parameter vector can be shown to converge to an affine subspace of dimension $N - k$. The Partial Convergence Theorem of Section 5 also implies the results of Morgan and Narendra (1977) and Anderson (1977), but gives a greatly simplified proof.

In Section 6 higher relative degree cases are considered and the results of the previous sections are shown to hold, despite the more complicated control strategies. It is shown that one can never guarantee convergence of the gain parameter associated with the augmented error signal. This clarifies the misleading statement in Boyd and Sastry (1983) that "the extension of the results presented (there) to higher relative degree cases is straightforward".

The appendix contains proofs of the theorems of Generalized Harmonic Analysis used in the paper. Although some of these theorems are analogous to results from the theory of wide-sense stationary stochastic processes, these proofs are not, to the authors' knowledge, in the literature.

2. THE MODEL REFERENCE ADAPTIVE SYSTEM

To fix notation, the model reference adaptive system of Narendra and Valavani (1978) and Narendra *et al.* (1980) is reviewed. The single input single output plant is assumed to be represented by a transfer function

$$\hat{W}_P(s) = k_P \frac{\hat{n}_P(s)}{\hat{d}_P(s)}, \tag{2.1}$$

where $\hat{n}_P, \hat{d}_P$ are relatively prime monic polynomials of degree $n - 1, n$ respectively and k_P is a scalar. The following are assumed known about the plant transfer function:

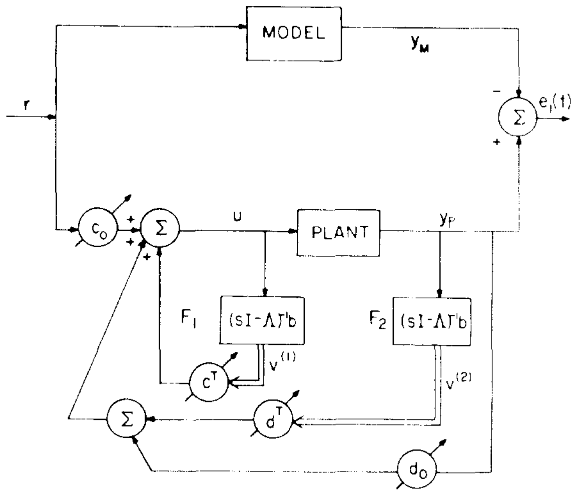

FIG. 1. The adaptive system for the relative degree one case.

(A1) the degree of the polynomial $\hat{d}_P$, i.e. n, is known;

(A2) the sign of k_P is known (say, + without loss of generality);

(A3) the transfer function $\hat{W}_P$ is assumed to be minimum phase, i.e. $\hat{n}_P$ is Hurwitz.

The objective is to build a compensator so that the plant output asymptotically matches that of a stable reference model $\hat{W}_M(s)$ with input $r(t)$ and output $y_M(t)$ and transfer function

$$\hat{W}_M(s) = k_M \frac{\hat{n}_M(s)}{\hat{d}_M(s)}, \tag{2.2}$$

where $k_M > 0$ and $\hat{n}_M$, $\hat{d}_M$ are monic polynomials of degree $n - 1$ and n, respectively ($\hat{n}_M$ and $\hat{d}_M$ need not be relatively prime). If the input and output of the plant are denoted $u(t)$ and $y_P(t)$, respectively, the objective may be stated: find $u(t)$ so that $y_P(t) - y_M(t) \to 0$ as $t \to \infty$. By suitable prefiltering, if necessary, one may assume that the model $W_M(s)$ is strictly positive real.

The scheme proposed by Narendra *et al.* is shown in Fig. 1. The dynamic compensator blocks F_1 and F_2 are identical one input, $(n - 1)$ output systems, each with transfer function

$$(sI - \Lambda)^{-1}b; \quad \Lambda \in R^{n-1 \times n-1},\ b \in R^{n-1},$$

where Λ is chosen so that its eigenvalues are the zeros of $\hat{n}_M$. Assume that the pair (Λ, b) is in controllable canonical form so that

$$(sI - \Lambda)^{-1}b = \frac{1}{\hat{n}_M(s)} \begin{bmatrix} 1 \\ s \\ \vdots \\ s^{n-2} \end{bmatrix}. \tag{2.3}$$

The parameters $c \in R^{n-1}$ in the precompensator block serve to tune the closed loop plant zeros, $d \in R^{n-1}$, $d_0 \in R$ in the feedback compensator assign the closed loop plant poles. The parameter c_0 adjusts the overall gain of the closed loop plant. Thus, the vector of $2n$ adjustable parameters denoted θ is

$$\theta^T = [c_0, c^T, d_0, d^T].$$

If the signal vector $w \in R^{2n}$ is defined by

$$w^T = [r, v^{(1)T}, y_P, v^{(2)T}]. \tag{2.4}$$

we see that the input to the plant is given by

$$u = \theta^T w. \tag{2.5}$$

It may be verified that there is a unique constant $\theta^* \in R^{2n}$ such that when $\theta = \theta^*$, the transfer function of the plant plus controller equals $\hat{W}_M(s)$.† If $r(t)$ is bounded (an assumption henceforth made) it can be shown that under the parameter update law

$$\dot{\theta} = -e_1 w = -(y_P - y_M)w \tag{2.6}$$

all signals in the loop, i.e. u, $v^{(1)}$, $v^{(2)}$, y_P, y_M are bounded, and in addition $\lim_{t\to\infty} e_1(t) = 0$, i.e. the plant output matches the model output and thus the overall objective has been achieved. However the convergence need not be exponential.

Despite the fact that $e_1(t) \to 0$, the parameter vector θ does not necessarily converge to θ^* (it may not converge at all). Various authors (Morgan and Narendra, 1977; Anderson, 1977; Kriesselmeier, 1977) have established that $e_1(t) \to 0$ and $\theta(t) \to \theta^*$ (i.e. the parameter error converges to 0) exponentially iff the signal vector $w(t)$ is persistently exciting (PE). If $\dot{r}$ is bounded (an assumption henceforth made) then PE can be simply stated: there are $\delta, \alpha > 0$ such that for all $s \geq 0$

$$\int_s^{s+\delta} ww^T \,dt \geq \alpha I. \tag{2.7}$$

Several comments are in order here. First, if $\dot{r}$ is not bounded then the PE condition is similar to but not exactly equivalent to (2.7). A complete discussion of this can be found in Morgan and Narendra (1977). Second, since r is bounded, there a β such that

$$\beta I \geq \int_s^{s+\delta} ww^T \,dt \geq \alpha I,$$

which is the form in which the PE condition often appears in the literature.

Since $w(t)$ contains the signals $v^{(1)}(t)$, $v^{(2)}(t)$, $y_P(t)$ generated inside the non-linear plant loop, translating the PE condition (2.7) on w into an equivalent condition on the exogenous reference input $r(t)$ would seem difficult if even possible. This is precisely what will now be done. Amazingly enough, the condition is very simple when expressed in the frequency domain.

3. REVIEW OF GENERALIZED HARMONIC ANALYSIS

The integral (2.7) appearing in the definition of PE reminds one of an autocovariance.

Definition 3.1 (*Autocovariance*). A function $u: R_+ \to R^n$ is said to have *autocovariance* $R_u(\tau) \in R^{n\times n}$ iff

$$\lim_{T\to\infty} \frac{1}{T} \int_s^{s+T} u(t)u(t+\tau)^T \,dt = R_u(\tau) \tag{3.1}$$

with the limit uniform in s.

This concept is well known in the theory of time series analysis. There is a strong analogy between (3.1) and $R_u^{\text{stoch}}(\tau) = Eu(t)u(t+\tau)$ for u a wide sense stationary stochastic process. Indeed, for a wide sense stationary *ergodic* process $u(t,\omega)$, $R_u(\tau,\omega)$ exists and is $R_u^{\text{stoch}}(\tau)$ for almost all ω. An autocovariance is a completely deterministic notion. Its relation to the notion of PE is simple.

Lemma 3.2 (*PE lemma*). Suppose w has autocovariance $R_w(\tau)$. Then w is PE iff $R_w(0) > 0$.

Proof. The "if" part is clear. Suppose now that w has an autocovariance R_w and is PE. Let $c \in R^n$, $c \neq 0$. From (3.1), for all n

$$\frac{1}{n\delta} \int_s^{s+n\delta} (w^T c)^2 \,dt \geq \frac{\alpha}{\delta} \|c\|^2.$$

Hence

$$\overline{\lim_{T\to\infty}} \frac{1}{T} \int_s^{s+T} (w^T c)^2 \,dt \geq \frac{\alpha}{\delta} \|c\|^2. \tag{3.2}$$

Because w has an autocovariance,

† Indeed θ^* consists of k_M/k_P and the coefficients of the polynomials $\hat{n}_P - \hat{n}_M$ and $\hat{d}_P - \hat{d}_M$.

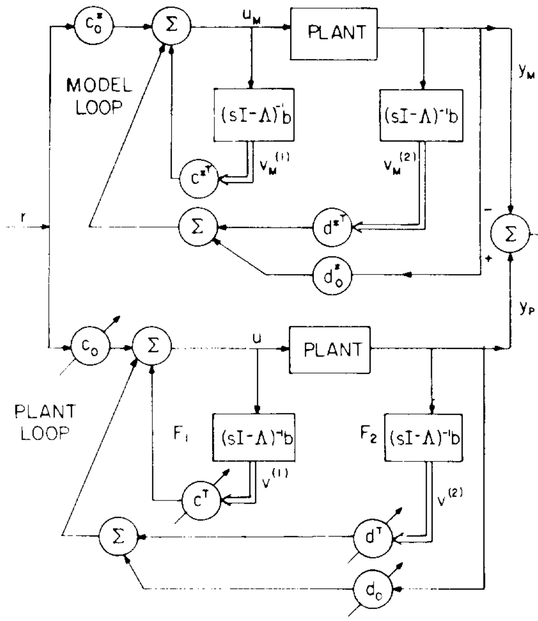

FIG. 2. The adaptive system of Fig. 1 with a new representation for the model.

$$\lim_{T\to\infty}\frac{1}{T}\int_s^{s+T}(w^Tc)^2\,dt = c^TR_w(0)c. \qquad (3.3)$$

From (3.2) and (3.3), $c^TR_w(0)c \geq \alpha/\delta\|c\|^2$, thus $R_w(0) \geq \alpha/\delta > 0$.

A few more simple lemmas concerning autocovariances are required. The proofs and a more detailed discussion of Generalized Harmonic Analysis appear in the Appendix.

Lemma 3.3. $R_u(\tau)$ is a positive semi-definite function.

Thus, provided $R_u(\tau)$ is continuous at $\tau = 0$ (an assumption henceforth made†) it has a Bochner representation:

$$R_u(\tau) = \int e^{i\nu\tau}S_u(d\nu), \qquad (3.4)$$

where S_u is a positive semi-definite matrix of bounded measures, which is called the *spectral measure* of u. If u is scalar valued, then S_u is just a positive bounded measure; $2S_u([\omega_0, \omega_1])$ can then be interpreted as the *average energy* contained in u in the frequency band $[\omega_0, \omega_1]$. Thus, for example, if a scalar valued u has a spectral line of amplitude a_ν at ν, then S_u has a point mass at ν of size $|a_\nu\|^2$.

Lemma 3.4 (*Linear filter lemma*). Suppose $u: R_+ \to R^n$ has autocovariance $R_u(\tau)$, its spectral measure S_u, and h is an $m \times n$ matrix of bounded measures. Then $y = h{*}u$ has an autocovariance R_y. Its spectral measure is given by:

$$S_y(d\nu) = H(j\nu)S_u(d\nu)H(j\nu)^*. \qquad (3.5)$$

In particular,

$$R_y(0) = \int H(j\nu)S_u(d\nu)H(j\nu)^*, \qquad (3.6)$$

where $H(j\nu)$ is the Fourier transform of h.

The reader should note that these formulas are identical to those from the theory of stochastic processes.

Lemma 3.5. If $u - v \in L^2$ and u has an autocovariance R_u, then v has autocovariance R_u.

Thus transients of finite energy do not affect the autocovariance of a signal.

The main result can now be proved.

4. NECESSARY AND SUFFICIENT CONDITIONS FOR PARAMETER CONVERGENCE

As in Boyd and Sastry (1983), redraw Fig. 1 as Fig. 2 with the model represented in non-minimal form as the plant with compensator and $\theta = \theta^*$. The signal $w_M \in R^{2n}$ in the model loop is given by

$$w_M^T = [r, v_M^{(1)T}, y_M, v_m^{(2)T}].$$

It is shown in Narendra and Valavani (1978); Narendra *et al.* (1980) that $w - w_M \in L_2$.

Note that w_M is the output of a stable LTI system driven by $r(t)$ and its transfer function is

$$\hat{Q}(s) = \begin{bmatrix} 1 \\ \hat{W}_M\hat{W}_P^{-1}(sI-\Lambda)^{-1}b \\ \hat{W}_M \\ \hat{W}_M(sI-\Lambda)^{-1}b \end{bmatrix}.$$

The only property of $\hat{Q}$ which will be needed is that there is a *constant* invertible matrix M such that

$$\hat{Q}^T(s)M = \frac{1}{\hat{n}_P(s)\hat{d}_M(s)}[\hat{d}_P(s), \ldots, \hat{d}_P(s)s^{n-2}, \hat{n}_P(s), \ldots, \hat{n}_P(s)s^n]. \qquad (4.1)$$

(This is shown in Boyd and Sastry (1983).)

The following is assumed: *r has an autocovariance.*‡

† A discontinuity in R_u at $\tau = 0$ means, roughly speaking, that u contains energy at "infinite frequency", which does not happen in practice. An example: $u(t) = e^{it^2}$.

‡ Not all rs have autocovariances (e.g. $r(t) = \cos\log(1 + t)$) but reasonable ones, whose general characteristics do not change drastically over time, do.

Let the spectral measure of r be denoted S_r. An explicit formula will now be derived for $R_w(0)$.

By Lemmas 3.4 and 3.5 and the discussion above, w_M has an autocovariance, with spectral measure

$$S_{w_M}(\mathrm{d}\nu) = \hat{Q}(j\nu)S_r(\mathrm{d}\nu)\hat{Q}(j\nu)^*.$$

Since $w - w_M \in L^2$, another application of Lemma 3.5 shows that w has an autocovariance, its spectral measure also given by

$$S_w(\mathrm{d}\nu) = \hat{Q}(j\nu)S_r(\mathrm{d}\nu)\hat{Q}(j\nu)^*$$

and autocovariance at 0 given by

$$R_w(0) = \int \hat{Q}(j\nu)S_r(\mathrm{d}\nu)\hat{Q}(j\nu)^*. \tag{4.2}$$

By the PE lemma, then:

$$w \text{ is PE} \quad \text{iff } R_w(0) = \int \hat{Q}(j\nu)S_r(\mathrm{d}\nu)\hat{Q}(j\nu)^* > 0. \tag{4.3}$$

Main Theorem. w is PE iff the spectral measure of r is *not* concentrated on $k < 2n$ points.

Proof. Suppose first that S_r is concentrated at $\nu_1, \ldots, \nu_k$, where $k < 2n$. Then

$$R_w(0) = \int \hat{Q}(j\nu)S_r(\mathrm{d}\nu)\hat{Q}(j\nu)^* = \sum_{m=1}^{k} \hat{Q}(j\nu_m)S_r(\{\nu_m\})\hat{Q}(j\nu_m)^*.$$

Being the sum of $k < 2n$ dyads, $R_w(0)$ is singular so by (4.3) w is not PE. •

Suppose now that w is not PE. Then by the PE lemma there is a non-zero $c \in R^{2n}$ such that

$$0 = c^T R_w(0)c = \int |\hat{Q}(j\nu)^*c|^2 S_r(\mathrm{d}\nu). \tag{4.4}$$

Since $|\hat{Q}(j\nu)^*c|^2$ is continuous in ν, (4.4) implies that $\hat{Q}(j\nu)^*c$ vanishes for all ν in $\mathrm{Supt}(S_r)$, the support of S_r. Thus for all $\nu \in \mathrm{Supt}(S_r)$,

$$0 = \hat{Q}(j\nu)^*c = \hat{Q}(j\nu)^*M(M^{-1}c), \tag{4.5}$$

where M is the constant non-singular matrix referred to in (4.1). If $\tilde{c} = M^{-1}c$, noting that $\hat{d}_P(j\nu)\hat{n}_M(j\nu) \neq 0$, (4.5) says for all $\nu \in \mathrm{Supt}(S_r)$,

$$0 = \hat{Q}(j\nu)^*M\tilde{c} = a(j\nu)\hat{d}_P(j\nu) + b(j\nu)\hat{n}_P(j\nu), \tag{4.6}$$

where the polynomials $a(s)$ and $b(s)$ are defined by

$$a(s) = \sum_{m=1}^{n-1} \tilde{c}_m s^{m-1}, \quad b(s) = \sum_{m=n}^{2n} \tilde{c}_m s^{m-n}. \tag{4.7}$$

Now if $\mathrm{Supt}(S_r)$ contains $2n$ or more points, (4.6) vanishes identically since its right hand side is a polynomial of degree $< 2n$, that is

$$a\hat{d}_P + b\hat{n}_P = 0. \tag{4.8}$$

But this contradicts coprimeness of $\hat{d}_P$ and $\hat{n}_P$, since (4.8) implies $\hat{n}_P/\hat{d}_P = -a/b$ and $\partial a \leq n - 2 < \partial \hat{n}_P$. So $\mathrm{Supt}(S_r)$ must contain $k < 2n$ points, and the Main Theorem is proved.

4.1. *Discussion*

The following has been proved:

Suppose the reference input $r(t)$ to the MRAC system of Section 2 has an autocovariance. Then the model-plant mismatch error $y_P - y_M$ and the parameter error $\theta - \theta^*$ tend to 0 exponentially iff the spectral measure of r is not supported on $k < 2n$ points.

Thus in general, one has parameter convergence: only for very special reference signals (which unfortunately sometimes include analytical favourites such as $1(t)$, $\cos(\omega t)$) does one not have $\theta \to \theta^*$.

It is instructive to see how previous (Boyd and Sastry, 1983) *sufficient* conditions on $r(t)$ fit into the theory above. If r has an autocovariance and has $2n$ spectral lines, then its spectral measure S_r has point masses at the $2n$ frequencies. Thus

$$R_w(0) = \int \hat{Q}(j\nu)S_r(\mathrm{d}\nu)\hat{Q}(j\nu)^* \geq \sum_{l=1}^{2n} \hat{Q}(j\nu_l)S_r(\{\nu_l\})\hat{Q}(j\nu_l)^* > 0$$

since the vectors $\hat{Q}(j\nu_l)$ are linearly independent by the argument above.

The terms *sufficiently rich* (SR) and *persistently* *exciting* (PE) have been used somewhat interchangeably in the literature. It is proposed that PE refers to property (2.7) for a vector of signals, and that *sufficient richness* be a property of the *reference signal* (scalar valued). A vector of signals is thus PE or not, but whether or not a reference signal is SR depends on the MRAC being studied. More specifically it depends only on the number of unknown parameters in the system, so it is proposed that a reference signal which results in a PE w in an N-parameter MRAC be referred to as *sufficiently rich of order* N. Then the following characterization results.

If r has an autocovariance, then it is SR of order N iff the support of its spectral measure S_r contains at least N points.

Thus, for example, if r has any *continuous spectrum* (see Wiener, 1930 for examples of such rs) then r is SR of all orders.

5. PARTIAL CONVERGENCE

If w is *not* PE, then the parameter error need not converge to zero (it may not converge at all). In this case S_r is concentrated on $k < 2n$ frequencies $\nu_1, \ldots, \nu_k$. Intuition suggests that although θ need not converge to θ^*, it should converge to the set of θs for which the *closed loop plant matches the model* at the frequencies $s = j\nu_1, \ldots, j\nu_k$. This is indeed the case.

Before starting the theorem, this idea is discussed more formally. Suppose that the parameter vector θ is *constant*. Then the plant loop of the MRAC system is LTI: w is in this case Qr. Since the input to the plant is $u = \theta^T w$, the overall closed loop plant transfer function is $\hat{W}_P(s)\theta^T\hat{Q}(s)$. This transfer function matches $\hat{W}_M$ at $s = j\nu_1, \ldots, j\nu_k$ iff

$$\begin{bmatrix} \hat{W}_P(j\nu_1)\hat{Q}(j\nu_1)^T \\ \vdots \\ \hat{W}_P(j\nu_k)\hat{Q}(j\nu_k)^T \end{bmatrix} \theta = \begin{bmatrix} \hat{W}_M(j\nu_1) \\ \vdots \\ \hat{W}_M(j\nu_k) \end{bmatrix}. \qquad (5.1)$$

Call the set of θs for which (5.1) holds Θ. Since $\theta^* \in \Theta$,

$$\Theta = \theta^* + \text{Nullspace} \begin{matrix} \hat{W}_P(j\nu_1)\hat{Q}(j\nu_1)^T \\ \vdots \\ \hat{W}_P(j\nu_k)\hat{Q}(j\nu_k)^T \end{matrix}. \qquad (5.2)$$

Thus Θ has dimension $2n - k$. In terms of the parameter error vector $\phi = \theta - \theta^*$, Θ has the simple description

$$\theta \in \Theta \quad \text{iff } R_w(0)\phi = 0. \qquad (5.3)$$

The verification of this is left to the reader; recall that here

$$R_w(0) = \sum_{m=1}^{k} S_r(\{\nu_m\})\hat{Q}(j\nu_m)\hat{Q}(j\nu_m)^*.$$

Partial Convergence Theorem. Bearing the above discussion in mind, suppose that $\dot{r}$ is bounded, then

$$\lim_{t\to\infty} R_w(0)\phi(t) = 0. \qquad (5.4)$$

Remark. If $R_w(0) > 0$, then this theorem tells us nothing more than Theorem 1: $\phi \to 0$. But if w is not PE, the conclusion (5.4) can be interpreted as:

$$\theta(t) \to \Theta \quad \text{as} \quad t \to \infty,$$

which means $\mathrm{dist}(\theta(t), \Theta) \to 0$, *not* $\theta(t) \to \theta(\infty)$ for some $\theta(\infty) \in \Theta$. In particular, θ need not converge to any point as $t \to \infty$.

Proof. Since ϕ and w are bounded, find K such that $\|\phi(t)\|$, $\|w(t)\| \le K$.

Let $\epsilon > 0$ be given. Find T_0 such that for $t > T_0$, $\|R_w(0)\phi(t)\| \le \epsilon$.

First choose T_1 large enough that for all s,

$$\left\| R_w(0) - \frac{1}{T_1}\int_s^{s+T_1} w(t)w(t)^T\,\mathrm{d}t \right\| \le \frac{\varepsilon}{3K^2}. \qquad (5.5)$$

Thus for all t

$$\left| \phi^T(t)R_w(0)\phi(t) - \phi(t)^T\frac{1}{T_1} \times \int_t^{t+T_1} w(\tau)w(\tau)^T\,\mathrm{d}\tau\,\phi(t) \right| \le \frac{\varepsilon}{3}. \qquad (5.6)$$

From our update law $\dot{\phi} = \dot{\theta} = -we_1$; since $e_1 \to 0$, $\dot{\phi}(t) \to 0$ as $t \to \infty$. The hypothesis $\dot{r}$ bounded implies that $\phi(t)^T w(t) \to 0$ (Narendra and Valavani, 1978). Now find T_0 so that for $t \geq T_0$

$$(\phi(t)^T w(t))^2 \leq \frac{\varepsilon}{3} \tag{5.7a}$$

and

$$\|\dot{\phi}(t)\| \leq \frac{\varepsilon}{3K^3 T_1}. \tag{5.7b}$$

Then for $t \geq T_0$,

$$\left|\phi(t)^T \frac{1}{T_1}\int_t^{t+T_1} w(\tau)w(\tau)^T \mathrm{d}\tau \phi(t) - \frac{1}{T_1}\int_t^{t+T_1} \phi(\tau)^T w(\tau)^T \phi(\tau)\mathrm{d}\tau\right| \tag{5.8a}$$

$$= \left|\frac{1}{T_1}\int_t^{t+T_1} w(\tau)^T(\phi(t) - \phi(\tau))w(\tau)^T \times (\phi(t) + \phi(\tau))\,\mathrm{d}\tau\right| \leq \frac{\varepsilon}{3} \tag{5.8b}$$

using (5.7b). From (5.7a), for $t \geq T_0$,

$$\left|\frac{1}{T_1}\int_t^{t+T_1} \phi(\tau)^T w(\tau)w(\tau)^T \phi(\tau)\,\mathrm{d}\tau\right| \leq \frac{\varepsilon}{3}. \tag{5.9}$$

From (5.6), (5.8) and (5.9), for $t \geq T_0$

$$|\phi(t)^T R_w(0)\phi(t)| \leq \varepsilon,$$

which completes the proof of the Partial Convergence Theorem.

Remark 1. The proof relies only on the assumptions (5.7), which state, roughly speaking, that *the parameter error eventually becomes orthogonal to w* and that *the updating slows down.* These are nearly universal properties of adaptive systems, so this theorem actually applies quite generally, not just to Narendra's scheme. For example, it applies to all of the schemes described in Goodwin *et al.* (1980).

Remark 2. While the $2n - k$ dimensional set Θ to which $\theta(t)$ converges depends only on the frequencies $\nu_1, \ldots, \nu_k$ and not on the average powers $S_r(\{\nu_1\})$, ..., $S_r(\{\nu_k\})$ contained in the reference signal at those frequencies, the *rate of convergence* of θ to Θ depends on both.

Remark 3. As mentioned above, if w is PE then $R_w(0) > 0$ and consequently this theorem yields the original parameter convergence results of Morgan and Narendra (1977) and Anderson (1977): uniform, asymptotic convergence of ϕ to zero (and consequently exponential convergence). This proof, however, is considerably simpler than the original proofs.

6. PLANT RELATIVE DEGREE ≥ 2

The scheme of Section 2 needs to be modified (Narendra and Valavani, 1978) when the relative degree of the plant to be controlled is ≥ 2, i.e. the plant has the transfer function (2.1) with $\hat{n}_P$, $\hat{d}_P$ relatively prime monic polynomials of degree m, n respectively. In addition to the assumptions (A1)–(A3) the new assumption (A4) is added:†

(A4) The relative degree of the plant, i.e. $(n - m)$, is known.

The model has the form (2.2) with the difference that $\hat{n}_M$ has degree m. The objective of the adaptive control is as before: to get $e_1 = y_P - y_M$ to converge to zero as $t \to \infty$.

Although the control scheme in this case is considerably more complicated, it will be shown that the necessary and sufficient conditions for exponential parameter error convergence to zero are *identical* to those given in Section 4 for the relative degree one case: namely, that Supt(S_r) contain at least $2n$ points.

6.1. *The relative degree* 2 *case*

Consider first the scheme of Fig. 1 with the difference that Λ is chosen exponentially stable so that its eigenvalues (there are $n - 1$ of them) include the zeros of $\hat{n}_M$ (there are m of them). It may again be verified that there is a unique constant $\theta^* \in R^{2n}$ such that when $\theta = \theta^*$ the transfer function of the plant plus controller equals $\hat{W}_M(s)$. The relationship between θ^* and the coefficients of $\hat{n}_P$ and $\hat{d}_P$ is more complex in this case than in Section 2. In this case since $\hat{W}_M$ has relative degree 2 it cannot be chosen positive real; however, it may be assumed (using suitable prefiltering, if necessary) that there is $\hat{L}(s) = (s + \delta)$ with $\delta > 0$ such that $\hat{W}_M\hat{L}$ is strictly positive real.

Now, modify the scheme of Fig. 1 by replacing each of the gains θ_i, i.e. c_0, d_o, c, d, with the gains

† Of course, (A4) appears implicitly in the relative degree one case.

$\hat{L}\theta_i\hat{L}^{-1}$ which in turn are given by

$$\hat{L}\theta_i\hat{L}^{-1} = \theta_i + \dot{\theta}_i\hat{L}^{-1} \quad i = 1,\ldots,2n. \qquad (6.1)$$

Now define the signal vector

$$\zeta^T(t) \triangleq [\hat{L}^{-1}r, \hat{L}^{-1}v^{(1)}, \hat{L}^{-1}y_P, \hat{L}^{-1}v^{(2)}]. \qquad (6.2)$$

Then

$$\dot{\theta} = -e_1\zeta$$

yields that $e_1(t) \to 0$ as $t \to \infty$ provided $r(t)$ is bounded. The persistent excitation condition for exponential parameter and error convergence is on the signal vector $\zeta(t)$ of (6.2): there are $\alpha, \delta > 0$ such that for all $s \geq 0$,

$$\int_s^{s+\delta} \zeta(t)\zeta(t)^T \, \mathrm{d}t \geq \alpha I. \qquad (6.3)$$

Now, define the analogous signal vector for the model

$$\zeta_M^T = [\hat{L}^{-1}r, \hat{L}^{-1}v_M^{(1)T}, \hat{L}^{-1}y_M, \hat{L}^{-1}v_M^{(2)T}],$$

i.e. ζ_M is obtained by filtering each component of w_M through the stable system with transfer function $\hat{L}^{-1}$.

Suppose now that r has an autocovariance. The output of a LTI filter driven by r is ζ_M, so it has an autocovariance; since $\zeta - \zeta_M \in L^2$ (see [Narendra and Valavani, 1978]), ζ has an autocovariance identical to that of ζ_M. In fact

$$R_\zeta(0) = \int |\hat{L}^{-1}(j\nu)|^2 \hat{Q}(j\nu) S_r(\mathrm{d}\nu) \hat{Q}(j\nu)^*.$$

Thus $R_\zeta(0) > 0$ iff $R_w(0) > 0$ and hence the necessary and sufficient conditions on r for exponential parameter convergence are exactly the same as in the relative degree one case.

6.2. *Relative degree* ≥ 3

As in Section 6.1, pick a Hurwitz polynomial $\hat{L}$ so that $\hat{L}\hat{W}_M$ is strictly positive real. The trick used, namely, to replace each θ_i by $\hat{L}\theta_i\hat{L}^{-1}$, is no longer possible since $\hat{L}\theta_i\hat{L}^{-1}$ depends on second (and possibly higher) derivatives of θ_i. To obtain a positive real error equation, retain the configuration of Fig. 1 and attempt to augment the model output by

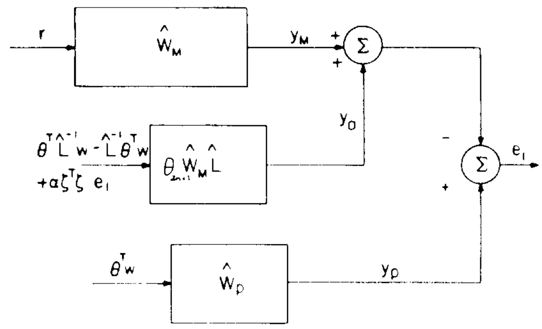

FIG. 3. Modification of the adaptive scheme when the relative degree ≥ 3.

$$\frac{k_P}{k_M}\hat{W}_M\hat{L}[\theta^T\hat{L}^{-1} - \hat{L}^{-1}\theta^T]w. \qquad (6.4)$$

The difficulty in implementing (6.4) arises from the fact that k_P is unknown. Consequently the model output is augmented, not by (6.4) but by

$$\hat{W}_M\hat{L}\theta_{2n+1}(t)[\theta^T\hat{L}^{-1} - \hat{L}^{-1}\theta^T]w \qquad (6.5)$$

with θ_{2n+1} being a new adaptive parameter expected to converge to k_P/k_M. To obtain $\phi \in L^2$ and prove stability of the augmented scheme an additional quadratic term is also added as shown in Fig. 3 to (6.5) to get

$$\hat{W}_M\hat{L}\theta_{2n+1}(t)\{(\theta^T\hat{L}^{-1} - \hat{L}^{-1}\theta^T)w + \alpha\zeta^T\zeta e_1\} \qquad (6.6)$$

where $\alpha > 0$ and ζ is as defined in (6.2). If ξ is defined to be $(\theta^T\hat{L}^{-1} - \hat{L}^{-1}\theta T)w$ then the update law

$$\dot{\theta} = -e_1\zeta$$
$$\dot{\theta}_{2n+1} = e_1\xi$$

yields that as $t \to \infty$, $e_1(t) \to 0$, that is $y_M \to y_P$.

For the scheme of Fig. 3, there are $2n + 1$ parameters to be considered and the sufficient richness condition for parameter convergence reads: there are $\alpha, \delta > 0$ such that for all $s \geq 0$,

$$\int_s^{s+\delta} \begin{bmatrix}\zeta \\ \xi\end{bmatrix}[\zeta^T \xi] \, \mathrm{d}t \geq \alpha I, \qquad (6.7)$$

where $\xi \triangleq (\theta^T\hat{L}^{-1} - \hat{L}^{-1}\theta^T)w$. However, condition (6.7) can never be satisfied since $\zeta \to 0$ as $t \to \infty$ as pointed out by Anderson and Johnson (1982). From the preceding discussion, it follows that the addition of the new parameter θ_{2n+1} in the augmented output signal is what causes this difficulty. If k_p is known, of course, θ_{2n+1}, ξ are unnecessary and the parameter convergence condition (6.7) reduces to (6.3), which is satisfied if $r(t)$ is sufficiently rich of order $\geq 2n$.

When k_p is unknown, and when $r(t)$ is SR of order $\geq 2n$ if follows that the autocovariance at zero of the signal vector $[\zeta^T, \xi]^T$ is given by

$$\begin{bmatrix} R_\zeta(0) & 0 \\ 0 & 0 \end{bmatrix} \in R^{2n+1 \times 2n+1} \tag{6.8}$$

with $R_\zeta(0) > 0$. By the Partial Convergence Theorem of Section 5, it follows that the parameter error converges to the null space of the matrix in (6.8). Thus *all but the* $(2n + 1)th$ parameter errors converge to zero. But the $(2n + 1)th$ parameter is inconsequential since it is the gain parameter associated with the augmented model output y_a.

7. A SIMPLE SIMULATION

In this section the simplest simulation which will illustrate the results above is presented: a two parameter MRAC system, with plant $2/(s + 1)$ and reference model $3/(s + 3)$, as shown in Fig. 4. The correct values of the adjustable parameters are $c_0^* = 1.5$ and $d_0^* = -1.0$. The parameter update law (2.6),

$$\dot{c}_0 = -er, \quad \dot{d}_0 = -ey_p$$

was used.

In the first simulation the constant reference input $r(t) = 2$ was used, and all initial conditions were zero. This r has spectral measure $4\delta(v)$, that is, one spectral line at $v = 0$. Since there are two parameters, here $w^T = [r, y_p]$ is *not* PE, and hence the parameters need not converge to their correct values. In fact the parameters do converge, to 0.85 and -0.35, respectively, which yields an asymptotic closed loop response $1.7(s + 1.7)^{-1}$. This is not the model transfer function $3(s + 3)^{-1}$, but it does match the model transfer at $s = 0$, as required by the Partial Convergence Theorem. Figure 5 shows c_0 and d_0 for $0 \leq t \leq 10$.

In the second simulation (Fig. 6) the reference input $r(t) = 2$ was kept, but the parameter initial conditions were changed: $c_0(0) = -0.75$ and $d_0(0) = 0.75$. Once again c_0 and d_0 converge, but this time to 1.25 and -0.75, respectively, yielding an asymptotic closed loop response of $2.5(s + 2.5)^{-1}$. As in the first simulation, this matches the model transfer function at $s = 0$.

In the third simulation (Fig. 7) reference input $r(t) = 4 \sin 1.5t$ was used, which has spectral measure $4\delta(v - 1.5) + 4\delta(v + 1.5)$ and thus is SR of order two. Of course $c_0(t) \to 1.5$ and $d_0(t) \to -1$,

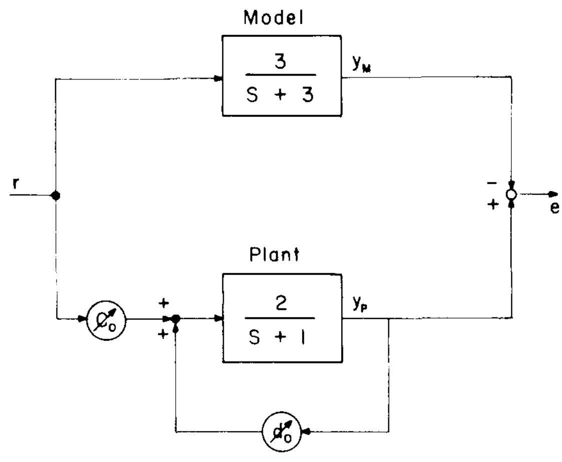

FIG. 4. Simple two parameter MRAC system simulated.

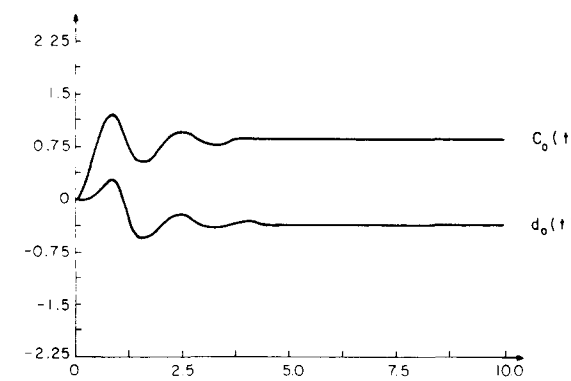

FIG. 5. Plot of c_0 and d_0 for $0 \leq t \leq 10$, with reference input $r(t) = 2$, all initial conditions zero.

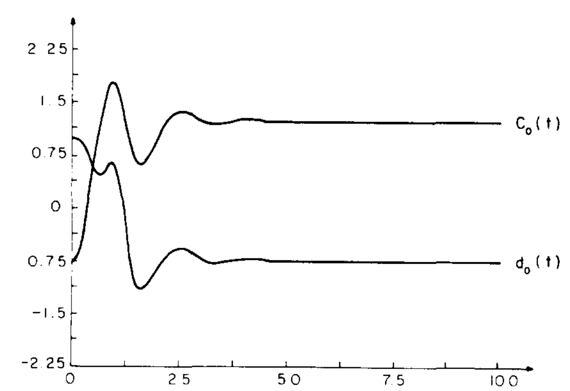

FIG. 6. Plot of c_0 and d_0 for $0 \leq t \leq 10$, with reference input $r(t) = 2$, with non-zero parameter initial conditions.

yielding asymptotic closed loop response $3(s + 3)^{-1}$.

Even this simplest example imparts something useful: when MRAC systems are used in regulator applications, and thus have constant reference inputs (as in the first two simulations) only parameter convergence to the set of parameters which yield unity closed loop gain (an affine space of dimension $2n - 1$, if there are $2n$ parameters) can be expected.

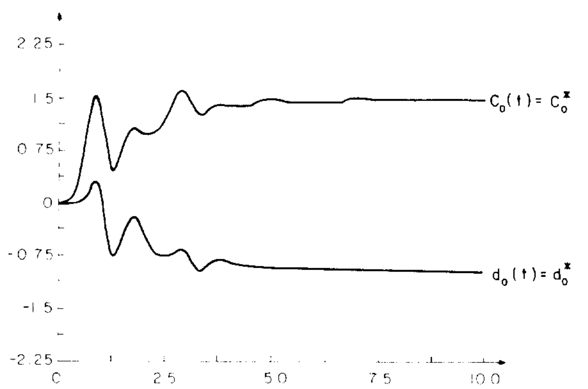

FIG. 7. Plot of c_0 and d_0 for $0 \leq t \leq 10$, with SR reference input $r(t) = 4 \sin 1.5t$.

8. CONCLUDING REMARKS

We have shown that a *complete description of parameter convergence* can be given in terms of the *spectrum of the reference input signal.*

Specifically, regardless of the relative degree:

(1) the parameter error ϕ converges exponentially to zero iff $\mathrm{Supt}(S_r)$ contains at least $2n$ points;

(2) if $\mathrm{Supt}(S_r)$ contains only $k < 2n$ points, then ϕ need not converge to zero. Instead it converges to a subspace of dimension $2n - k$, which corresponds *precisely to the set of parameter values for which the closed loop plant matches the model at the frequencies contained in* $\mathrm{Supt}(S_r)$.

Acknowledgements—The authors thank Prof. K. J. Åström for making his simulation package SIMNON available to them. They also thank the reviewers for several good suggestions and clarifications. This research was supported in part by NASA under grant NAG 2-243. Mr Boyd gratefully acknowledges the support of the Fannie and John Hertz Foundation.

REFERENCES

Anderson, B. D. O. (1977). Exponential stability of linear equations arising in adaptive identification. *IEEE Trans. Aut. Control*, **22**, 83–88.
Anderson, B. D. O. and C. R. Johnson (1982). Exponential convergence of adaptive identification and control algorithms. *Automatica*, **18**, 1–13.
Boyd, S. and S. Sastry (1983). On parameter convergence in adaptive control. *Syst. Control Lett.*, **3**, 311–319.
Boyd, S. and S. Sastry (1984). Necessary and sufficient conditions for parameter convergence in adaptive control. *Proc.* 1984 *Am. Control Conf.*, San Diego, June.
Dasgupta, S., B. D. O. Anderson and A. C. Tsoi (1983). Input conditions for continuous time adaptive system problems. *Proc. IEEE Conf. Decis. Control*, San Antonio, Texas, December, pp. 211–216.
Goodwin, G. C., P. J. Ramadge and P. E. Caines (1980). Discrete time multivariable adaptive control. *IEEE Trans. Aut. Control*, **25**, 449–456.
Kriesselmeier, G. (1977). Adaptive observers with exponential rate of convergence. *IEEE Trans. Aut. Control*, **22**, 2–9.
Koopmans, L. H. (1974). *The Spectral Analysis of Time Series*. Academic Press, New York.
Morgan, A. P. and K. S. Narendra (1977). On the uniform asymptotic stability of certain linear non-autonomous differential equations. *SIAM J. Control Opt.*, **15**, 5–24.
Morse, A. S. (1980). Global stability of parameter-adaptive control systems. *IEEE Trans. Aut. Control*, **25**, 433–440.
Narendra, K. S. and L. S. Valavani (1978). Stable adaptive controller design—direct control. *IEEE Trans. Aut. Control*, **23**, 570–583.
Narendra, K. S., Y.-M. Lin and L. S. Valavani (1980). Stable adaptive controller design, Part II: Proof of stability. *IEEE Trans. Aut. Control*, **25**, 440–448.
Narendra, K. S. and Y.-H. Lin (1980). Stable discrete adaptive control. *IEEE Trans. Aut. Control*, **25**, 456–461.
Narendra, K. S. and A. M. Annaswamy (1983a). Persistent excitation and robust adaptive algorithms. *Proc. 3rd Yale Workshop Applic. Adaptive Syst. Theory*, Yale University, New Haven, Connecticut, June.
Narendra, K. S. and A. M. Annaswamy (1983b). Persistent excitation and robust adaptive systems. Tech. Report #8302, Center for Systems Science, Yale University, New Haven, Connecticut, October.
Wiener, N. (1930). Generalized harmonic analysis. *Acta Mathematica*, **55**, 117–258.
Yuan, J. S.-C. and W. M. Wonham (1977). Probing signals for model reference identification. *IEEE Trans. Aut. Control*, **22**, 530–538.

APPENDIX: GENERALIZED HARMONIC ANALYSIS

The first careful treatment of the notion of autocovariance was Wiener (1930). The idea is well known in the theory of time series analysis (see e.g. Koopmans, 1974), and is usually presented in the context of stochastic processes. A clear modern discussion of autocovariances which does not make use of the connection with wide sense stationary stochastic processes could not be found. Since the proofs of the various lemmas used here are neither difficult nor long, they are given below.

The analogy between *autocovariance* and *stochastic autocovariance* mentioned in Section 3 is not complete—for example the limit in the definition of R_u makes the proof of the linear filter lemma trickier than the proof of its stochastic analogue (which is little more than interchanging integrals and expectation via the Fubini theorem) and there is no stochastic analogue of Lemma 3.5.

For the remainder of this section it is assumed that $u: \mathbf{R}_+ \to \mathbf{R}^n$ has autocovariance R_u. Note that the integral (3.1) in the definition of autocovariance makes sense iff u is locally square integrable, i.e. $u \in L^2_{\mathrm{loc}}$.

Lemma 3.3. R_u is a positive semi-definite function.

Proof. Suppose $\tau_1, \ldots, \tau_K \in R$, $c_1, \ldots, c_K \in C^n$. It must be shown that

$$\sum_{i,j} c_i^* R_u(\tau_j - \tau_i) c_j \geq 0. \tag{A1}$$

Define the scalar valued function v by:

$$v(t) \triangleq \sum_{k=1}^{K} c_k^* u(t + \tau_k).$$

Then for all $T > 0$

$$0 \leq \frac{1}{T} \int_0^T |v(t)|^2 \, dt \tag{A2}$$

$$= \sum_{i,j} c_i^* \left[\frac{1}{T} \int_0^T u(t + \tau_i) u(t + \tau_j)^* \, dt \right] c_j$$

$$= \sum_{i,j} c_i^* \left[\frac{1}{T} \int_{\tau_i}^{\tau_i + T} u(t) u(t + \tau_j - \tau_i)^* \, dt \right] c_j. \tag{A3}$$

Since u has an autocovariance, as $T \to \infty$ (A3) converges to

$$\sum_{i,j} c_i^* R_u(\tau_j - \tau_i) c_j.$$

From (A2), (A3) is non-negative, so (A1) follows.

Proposition (A1) implies that R_u is the transform of a positive semi-definite matrix S_r of bounded measures, that is

$$R_u(\tau) = \int e^{i\nu\tau} S_r(\mathrm{d}\nu). \tag{A4}$$

(This is the matrix analogue of Bochner's theorem.) S_r is symmetric, both in ν and as a matrix, since $R_u(\tau)$ is a real symmetric matrix.

Lemma 3.4 (*Linear filter lemma*). Suppose that $y = h*u$, where h is an $m \times n$ matrix of bounded measures. Then y has an autocovariance R_y given by

$$R_y(\tau) = \iint h(\mathrm{d}\tau_1) R_u(\tau + \tau_1 - \tau_2) h(\mathrm{d}\tau_2)^T \tag{A5}$$

and spectral measure S_y given by

$$S_y(\mathrm{d}\nu) = H(j\nu) S_u(\mathrm{d}\nu) H(j\nu)^*. \tag{A6}$$

Proof. First, establish that y has an autocovariance:

$$\frac{1}{T}\int_s^{s+T} y(t)y(t+\tau)^T \,\mathrm{d}t \tag{A7}$$

$$= \frac{1}{T}\int_s^{s+T} [h(\mathrm{d}\tau_1)u(t-\tau_1)][u(t+\tau-\tau_2)^T h(\mathrm{d}\tau_2)^T]\,\mathrm{d}t. \tag{A8}$$

For each T, the integrals in (A8) exist absolutely so the order of integration may be changed:

$$= \iint h(\mathrm{d}\tau_1)\left[\frac{1}{T}\int_{s-\tau_1}^{s-\tau_1+T} u(t)u(t+\tau+\tau_1-\tau_2)^T\,\mathrm{d}t\right] h(\mathrm{d}\tau_2)^T. \tag{A9}$$

The bracketed expression in (A9) converges to $R_u(\tau + \tau_1 - \tau_2)$ as $T \to \infty$, uniformly in s. Furthermore the bracketed expression in (A9) is *bounded* as a function of T, s, τ_1 and τ_2, for $T \geq 1$, since by Cauchy-Schwartz†

$$\left|\frac{1}{T}\int_{s-\tau_1}^{s-\tau_1+T} u(t)u(t+\tau+\tau_1-\tau_2)^T\,\mathrm{d}t\right|$$

$$\leq \sup_{s,T\geq 1} \frac{1}{T}\int_s^{s+T} \|u(t)\|^2\,\mathrm{d}t < \infty. \tag{A10}$$

So by dominated convergence (A9) converges, uniformly in s, as $T \to \infty$, to

$$\iint h(\mathrm{d}\tau_1) R_u(\tau+\tau_1-\tau_2) h(\mathrm{d}\tau_2)^T. \tag{A11}$$

† The restriction $T \geq 1$ is required if u is not bounded but only in L^2_{loc}.

Thus y has an autocovariance, given by (A11). This establishes (A5); to finish the proof, substitute the Bochner integral for R_u in (A11):

$$R_y(\tau) = \iint h(\mathrm{d}\tau_1) \int e^{i(\tau+\tau_1-\tau_2)\nu} S_u(\mathrm{d}\nu) h(\mathrm{d}\tau_2)^T \tag{A12}$$

$$= \int e^{i\nu\tau}\left[\int e^{-i\tau_1\nu} h(\mathrm{d}\tau_1)\right] S_u(\mathrm{d}\nu) \left[\int e^{-i\tau_2\nu} h(\mathrm{d}\tau_2)\right]^* \tag{A13}$$

(since all the measures are finite)

$$= \int e^{i\nu\tau} H(j\nu) S_u(\mathrm{d}\nu) H(j\nu)^*. \tag{A14}$$

This is the Bochner representation of R_y, so

$$S_y(\mathrm{d}\nu) = H(j\nu) S_u(\mathrm{d}\nu) H(j\nu)^* \tag{A15}$$

establishing the linear filter lemma.

Lemma 3.5 (*Transient lemma*). Suppose $e(t) = u(t) - v(t) \in L^2$ (and u has autocovariance R_m). Then v also has autocovariance R_u.

Proof.

$$\left|\frac{1}{T}\int_s^{s+T} u(t)u(t+\tau)^T\,\mathrm{d}t - \frac{1}{T}\int_s^{s+T} v(t)v(t+\tau)^T\,\mathrm{d}t\right| \tag{A16}$$

$$= \left|\frac{1}{T}\int_s^{s+T} e(t)u(t+\tau)^T\,\mathrm{d}t + \frac{1}{T}\int_s^{s+T} u(t)e(t+\tau)^T\,\mathrm{d}t + \frac{1}{T}\int_s^{s+T} e(t)e(t+\tau)^T\,\mathrm{d}t\right|$$

$$\leq \frac{1}{\sqrt{T}}\|e\|_2\left[\frac{1}{T}\int_s^{s+T} \|u(t+\tau)\|^2\right]^{1/2} + \frac{1}{\sqrt{T}}\|e\|_2\left[\frac{1}{T}\int_s^{s+T} \|u(t)\|^2\right]^{1/2} + \frac{1}{T}\|e\|_2^2 \tag{A17}$$

using the Cauchy-Schwartz inequality. The two bracketed expressions in (A17) converge uniformly in s as $T \to \infty$ to $\operatorname{Trace} R_u(0)$, so the entire expression (A17), and thus (A16), converges to zero, uniformly in s, as $T \to \infty$. Thus

$$\frac{1}{T}\int_s^{s+T} v(t)v(t+\tau)^T\,\mathrm{d}t \to R_u(\tau) \quad \text{as } T \to \infty$$

uniformly in s, and Lemma 3.5 is proved.

Remark. Actually the hypothesis can be weakened to $R_e = 0$, that is, e has *zero average energy*.