Stability Theory for Adaptive Systems: Method of Averaging and Persistency of Excitation

ROBERT L. KOSUT, MEMBER, IEEE, BRIAN D. O. ANDERSON, FELLOW, IEEE, AND IVEN M. Y. MAREELS

Abstract—**A method of averaging is developed for the stability analysis of linear differential equations with small time-varying coefficients which do not necessarily possess an average value. The technique is then applied to determine the stability of a linear equation which arises in the study of adaptive systems where the adaptive parameters are slowly varying. The stability conditions are stated in the frequency domain, which shows the relation between persistent excitation and unmodeled dynamics.**

I. Introduction

For a large class of adaptive feedback systems, as well as for some output error identification schemes, a stability analysis in the neighborhood of the desired behavior leads to investigating the stability of the following linear system of differential-operator equations (see, e.g., [1]-[3], [20])

$$\dot{\theta} = \epsilon[f - \phi H(\phi'\theta)] \tag{1.1a}$$

where $\theta(0) = \theta_0 \in \mathbb{R}^p$, ϵ is a positive constant, $f(\cdot)$, $\phi(\cdot): \mathbb{R}_+ \to \mathbb{R}^p$ are regulated and bounded, and H is a linear-time-invariant convolution operator with kernel $h(t)$ and transfer function $H(s)$, i.e.,

$$(Hu)(t) = \int_0^t h(t-\tau)u(\tau)\ d\tau. \tag{1.1b}$$

We consider the case when $H(s)$ is strictly proper and exponentially stable; thus, $h(t)$ is bounded by a decaying exponential. The strictly proper assumption is not necessary for analysis, but it is more often the case when (1.1) arises from dynamical systems. The same can be said for considering the general convolution (1.1b) and not just the case of rational $H(s)$.

The specific problem we consider is slow adaptation (small $\epsilon > 0$), and to determine sufficient conditions for which the map $(f, \theta_0) \to \theta$, defined implicitly by (1.1), is exponentially stable, i.e., there are positive constants K, α such that

$$|\theta(t)| \leq \int_0^t Ke^{-\alpha(t-\tau)}|f(\tau)|\ d\tau + Ke^{-\alpha t}|\theta_0|. \tag{1.2}$$

When such a condition exists, it then follows that the adaptive systems from which (1.1) arose is locally stable.

Manuscript received March 8, 1985; revised April 10, 1986 and August 7, 1986. Paper recommended by Past Associate Editor, J. K. Tugnait. This paper was presented at the 24th IEEE Conference on Decision and Control, December 1985. The work of R. L. Kosut was supported by the Air Force Office of Scientific Research under Contract F49620-C-84. The work of I. M. Y. Mareels was supported by the National Fund for Scientific Research, Belgium.

R. L. Kosut is with Integrated Systems, Inc., Palo Alto, CA 94301 and with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305.

B. D. O. Anderson and I. M. Y. Mareels are with the Department of Systems Engineering, Australian National University, Canberra, A.C.T. 2600, Australia.

IEEE Log Number 8611299.

A. Linearization and Local Stability

In [2], for example, system (1.1) is obtained as a result of linearization of the adaptive system in the neighborhood of a "tuned" system, i.e., a system where the adaptive parameters are set to a constant value $\theta_* \in \mathbb{R}^p$ and whose behavior is deemed acceptable. Hence, in (1.1), $\theta(t)$ is the vector of parameter errors between the parameter estimate at time t and the tuned value θ_*, $\phi(t)$ is the regressor vector from the tuned system (e.g., filtered versions of measured signals), and the scalar ϵ is the magnitude of the adaptation gain which essentially controls the rate of adaptation. The operator H depends on the actual system being controlled or identified and also on the tuned parameter setting θ_*.

It is shown in [2], [3] that if system (1.1) is exponentially stable, then the adaptive system is locally stable, i.e., the adaptive system behavior will remain in a neighborhood of the desired behavior provided the initial parameter error $\theta(0)$ and the effect of external disturbances are sufficiently small. Although the results in [2], [3] were arrived at using input-output properties [16], the local stability property also follows from the results on "total stability" [4], [20].

B. Unmodeled Dynamics and Slow Adaptation

In the ideal case there are a sufficient number of adaptive parameters (the number p) such that the tuned parameter setting results in $H(s)$ being strictly positive real (SPR), i.e., Re $H(j\omega) > 0$, $\forall\omega \in \mathbb{R}_+$. Under these conditions, we have the following results (see e.g., [5]-[8], [1]): 1) system (1.1) is stable, i.e., $\theta(t)$ is bounded but not necessarily constant; 2) if, in addition, $\phi(t)$ is persistently exciting, then system (1.1) is exponentially stable. The trouble starts when there are an insufficient number of parameters to obtain $H(s) \in$ SPR, as is the case in adaptive control when the plant has unmodeled dynamics (see, e.g., [2], [7], [12]).

In this paper we will examine the stability of (1.1) when ϵ is small, $\phi(t)$ is persistently exciting, and $H(s)$ is not necessarily SPR but only exponentially stable. Riedle and Kokotovic [9] refer to this case as "slow adaptation" and by using the method of averaging described by Hale [10], they show that the stability of (1.1) is critically dependent on the spectrum of the excitation in relation to the frequency response $H(j\omega)$. With the same assumptions, Astrom [11] uses averaging techniques to analyze the interaction between unmodeled dynamics and external inputs in the counterexample posed by Rohrs *et al.* [12]. Both of these analyses require the assumption that $\phi(t)$ is almost periodic and that $H(s)$ is rational. In this case Riedle and Kokotovic [9] show that system (1.1) is exponentially stable if

$$\lambda\left(\sum_{\omega\in\Omega}[\alpha(\omega)\alpha(\omega)^*]\ \mathrm{Re}\ H(j\omega)\right) > 0 \tag{1.3}$$

where Ω and $\{\alpha(\omega),\ \omega \in \Omega\}$ are, respectively, the Fourier exponents and coefficients of $\phi(t)$. Condition (1.2) can be considered as a *signal dependent positivity condition*, but unlike

0018-9286/87/0100-0026$01.00

the SPR condition, Re $H(j\omega)$ is not required to be positive at *all* frequencies.

The main contribution of this paper is to extend the theory of averaging to include the case when $\phi(t)$ does not have a (generalized) Fourier series representation, but is only known to be regulated and bounded. Thus, $\phi(t)$ need not be almost periodic nor even possess an average value. We also state stability conditions in the frequency domain in a form similar to (1.2). Moreover, $H(s)$ need not be rational. Analogous results can be stated for discrete-time systems; see, e.g., [13].

C. Averaging: Uses and Limitations

The averaging theory developed here, as well as averaging theory in general, has its uses and limitations for adaptive systems. In the first place, the theory requires slow adaptation, which can be counterproductive because performance can be below par for the long period of time it takes for the parameters to adjust. Second, the averaging results developed in the sequel concern linear time-varying systems only, so that application of these results to the nonlinear adaptive system requires a linearization. In this sense we can obtain information, including frequency domain information, about the dynamical behavior of the adaptive system in the neighborhood of the tuned system. Both stability and instability conditions are discussed. The results arising from a combination of small gain theory and perturbation methods, e.g., [2], [3], [14], [15], are restricted to stability results, and are far less quantitative.

D. Organization of Paper

The paper is organized as follows. Section II develops a method of averaging for linear systems with *sample averages*. In Section III we apply the general results of Section II to (1.1) and obtain conditions for stability and instability. In Section IV these are interpreted in terms of frequency domain stability conditions. In Section V we provide a general discussion.

E. Notation

The symbol $|\cdot|$ denotes both the vector norm as well as its induced matrix norm. Similarly, $\|\cdot\|_p$, $p \in [1, \infty]$, denotes the L_p-norm of a vector or matrix function, i.e., for $p \in [1, \infty)$, $\|F\|_p = (\int_0^\infty |F(t)|^p\, dt)^{1/p}$, and $\|F\|_\infty = \text{ess sup}\, \{|F(t)|: t \geq 0\}$. $\lambda_i(A)$ denotes the ith eigenvalue of matrix A and $\sigma_i(A)$ denotes the ith singular value of A, i.e., $\sigma_i(A) = [\lambda_i(A^*A)]^{1/2}$. An operator H is *L_p-stable* if $\exists$ constants k, b such that $\|Hu\|_p \leq k\|u\|_p + b$, $\forall u \in L_p$. The smallest k is referred to as the *L_p-gain*, and is denoted by $\gamma_p(H)$.

II. Method of Averaging for Linear Homogeneous Systems

In this section we will consider the homogeneous linear time-varying system

$$\dot{x} = \epsilon A(t)x. \tag{2.1}$$

Lemma 2.1: Suppose in (2.1) that ϵ is a real constant and $A(\cdot): \mathbb{R}_+ \to \mathbb{R}^{n\times n}$ is regulated and bounded. Then $\forall s, \tau \in \mathbb{R}_+$, the transition matrix $F(s+\tau, s)$ of (2.1) is given by

$$F(s+\tau, s) = \exp\,[\epsilon\tau \bar{A}_\tau(s)] + R(s, \epsilon\tau) \tag{2.2}$$

where

$$\bar{A}_\tau(s) = \frac{1}{\tau}\int_s^{s+\tau} A(t)\, dt \tag{2.3}$$

is referred to as the *sample average value* of $A(t)$ on the interval $s \leq t \leq s+\tau$, and

$$\|R(\cdot, \epsilon\tau)\|_\infty \leq (\epsilon\tau\|A\|_\infty)^2 \exp\,(\epsilon\tau\|A\|_\infty) := r(\epsilon\tau\|A\|_\infty). \tag{2.4}$$

Proof: Using the Peano–Baker series representation for the transition matrix of (2.1) gives

$$F(s+\tau, s) = I + \epsilon\int_s^{s+\tau} A(t)\, dt + \sum_{k=2}^\infty \epsilon^k \int_s^{s+\tau} A(t_1)\int_s^{t_1} A(t_2)\cdots\int_s^{t_{k-1}} A(t_k)\, dt_1\cdots dt_k.$$

Using definitions (2.2) and (2.3) for $R(s, \epsilon\tau)$ and $\bar{A}_\tau(s)$, respectively, together with the series expansion for exp $(\epsilon\tau\bar{A}_\tau(s))$ results in

$$\begin{aligned} R(s, \epsilon\tau) &= \sum_{k=2}^\infty \Big[-(\epsilon\tau\bar{A}_\tau(s))^k/k! + \epsilon^k\int_s^{s+\tau} A(t_1) \cdot \int_s^{t_1} A(t_2)\cdots\int_s^{t_{k-1}} A(t_k)\, dt_1\cdots dt_k\Big] \\ &\leq 2\sum_{k=2}^\infty (\epsilon\tau\|A\|_\infty)^k/k!, \qquad \forall s \in \mathbb{R}_+ \\ &= (\epsilon\tau\|A\|_\infty)^2 \exp\,(\epsilon\tau\|A\|_\infty) \end{aligned}$$

since $\|\bar{A}_\tau(\cdot)\|_\infty \leq \|A(\cdot)\|_\infty$. This proves (2.4). □

Remarks:

1) Assuming that $A(t)$ is regulated and bounded is sufficient for the existence and uniqueness of solutions [17].

2) Observe that Lemma 2.2 is valid $\forall s, \tau \in \mathbb{R}_+$ and $\forall \epsilon \in \mathbb{R}$. In the sequel we use Lemma 2.2 only for the case when $\epsilon > 0$ and $\epsilon\tau$ is small.

The stability properties of (2.1) can be established by application of Lemma 2.2 as stated in Theorem 2.1 below. We first require the following.

Definition: The function $\mu(\cdot): \mathbb{R}^{n\times n} \to \mathbb{R}$, defined by

$$\mu(M) = \lim_{\alpha\downarrow 0}\, (|I+\alpha M| - 1)/\alpha \tag{2.5}$$

is called the *measure of the matrix M*, where $|\cdot|$ is an induced matrix norm on $\mathbb{R}^{n\times n}$.

For any induced matrix norm and its corresponding measure, the following properties hold (see, e.g., [16]):

P1)

$$-|M| \leq -\mu(-M) \leq \text{Re}\ \lambda(M) \leq \mu(M) \leq |M|, \qquad \forall M \in \mathbb{R}^{n\times n}. \tag{2.6a}$$

P2)

$$\mu(M_1+M_2) \leq \mu(M_1) + \mu(M_2), \qquad \forall M_1, M_2 \in \mathbb{R}^{n\times n}. \tag{2.6b}$$

P3) The transition matrix $F(t, \tau)$ of $\dot{x} = M(t)x$ satisfies

$$\exp\left(-\int_\tau^t \mu[-M(s)]\, ds\right) \leq |F(t, \tau)| \leq \exp\left(\int_\tau^t \mu[M(s)]\, ds\right). \tag{2.6c}$$

P4) If the vector norm on $\mathbb{R}^n$ is $|x| = (x'Px)^{1/2}$, $P = P' > 0$, then

$$|M| = \max_i \sigma_i(P^{1/2}MP^{-1/2}) \tag{2.6d}$$

$$\mu(M) = \frac{1}{2}\max_i \lambda_i(P^{1/2}MP^{-1/2} + P^{-1/2}M'P^{1/2}). \tag{2.6e}$$

These properties, together with Lemma 2.2, yield the following stability result for system (2.1).

Theorem 2.1: Suppose $A(t)$ in (2.1) is regulated and bounded with the sequence of sample averages $\{\bar{A}_T(kT), \forall k \in \mathbb{Z}_+\}$. Then:

i) if $\exists T > 0$ and $\alpha > 0$ such that

$$\mu[\bar{A}_T(kT)] \leq -\alpha, \qquad \forall k \in \mathbb{Z}_+ \tag{2.7}$$

then $\exists \eta > 0$ such that $\forall \epsilon T \in (0, \eta)$, the zero solution of (2.1) is u.a.s.;

ii) if $\exists T > 0$ and $\alpha > 0$ such that

$$\mu[-\bar{A}_T(kT)] \leq -\alpha, \qquad \forall k \in \mathbb{Z}_+ \tag{2.8}$$

then $\exists \eta > 0$ such that $\forall \epsilon T \in (0, \eta)$, the zero solution of (2.1) is completely unstable.

Proof: Combining (2.6c) with (2.2) gives

$$\begin{aligned}|\exp[-\epsilon\tau\mu(-\bar{A}_\tau(s))]| &\leq |F(s+\tau, s) - R(s, \epsilon\tau)| \\ &= |\exp[\epsilon\tau\bar{A}_\tau(s)]| \\ &\leq \exp[\epsilon\tau\mu(\bar{A}_\tau(s))], \qquad \forall s, \tau \in \mathbb{R}_+\end{aligned}$$

which implies

$$|F(s+\tau, s)| \leq \exp[\epsilon\tau\mu(\bar{A}_\tau(s))] + r(\epsilon\tau m) \tag{2.9}$$

$$|F(s+\tau, s)| \geq \exp[-\epsilon\tau\mu(\bar{A}_\tau(s))] - r(\epsilon\tau m) \tag{2.10}$$

where we have used (2.4) with $\|A\|_\infty = m$.

We first prove part i) by using condition (2.7) and inequality (2.9) with $\tau = T$ and $s = kT$. This gives

$$|F((k+1)T, kT)| \leq \exp(-\epsilon T\alpha) + r(\epsilon Tm), \qquad \forall k \in \mathbb{Z}_+.$$

We now establish that for all small $\epsilon T > 0$, $|F((k+1)T, kT)| < 1$, i.e., the map $\theta(kT) \to \theta((k+1)T)$ is a contraction. From the definition of $r(\cdot)$ in (2.4), it follows that for any $\alpha > 0$ there is an $\eta > 0$ such that

$$\exp(-\eta\alpha) + r(\eta m) = 1. \tag{2.11}$$

Hence, for all $\epsilon T \in (0, \eta)$, there is a $\beta > 0$ such that

$$\exp(-\epsilon T\alpha) + r(\epsilon Tm) = \exp(-\epsilon T\beta) < 1 \tag{2.12}$$

which shows the contraction property.

Now, for any $t, s \in \mathbb{R}_+$ with $t \geq s$, there exists an integer $k \geq 0$ such that $s + kT \leq t \leq s + (k+1)T$. Thus

$$\begin{aligned}|F(t, s)| &= |F(t, s+kT)F(s+kT, s+(k-1)T) \cdots F(s+T, s)| \\ &\leq |F(t, s+kT)| \exp(-\epsilon kT\beta) \\ &\leq |F(t, s+kT)| \exp(-\epsilon(t-s-T)\beta), \\ &\qquad \text{by } kT \geq t-s-T \\ &\leq \exp(\epsilon T(m+\beta) \exp(-\epsilon(t-s)\beta).\end{aligned}$$

The last line follows from Property (2.6c), i.e.,

$$\begin{aligned}|F(t, s+kT)| &\leq \exp\left(\int_{s+kT}^{t} \mu[\epsilon A(\tau)]\, d\tau\right) \\ &\leq \exp(\epsilon m(t-s-kT)), \qquad \text{by } \mu[A(\tau)] \leq |A(\tau)| \leq m \\ &\leq \exp(\epsilon mT), \qquad \text{by } t-s-kT \in (0, T).\end{aligned}$$

This proves part i) of Theorem 2.1. The proof of part ii) follows from the above analysis, but starting with inequality (2.10). □

Using the same technique, but allowing $A(\cdot)$ [equivalently $\bar{A}_T(\cdot)$] to possess a uniform average, we obtain the following sharper result.

Theorem 2.2: Suppose $A(t)$ in (2.1) is regulated, bounded, and has a *uniform average* $\bar{A} \in \mathbb{R}^{n\times n}$, i.e.,

$$\lim_{T\to\infty} \bar{A}_T(s) = \bar{A} \tag{2.13}$$

uniformly $\forall s \in \mathbb{R}$. Under these conditions:

i) if $\exists \alpha > 0$ such that

$$\text{Re } \lambda(\bar{A}) \leq -\alpha \tag{2.14}$$

then $\exists \epsilon_0 > 0$ such that $\forall \epsilon \in (0, \epsilon_0)$, the zero solution of (2.1) is u.a.s.;

ii) if $\exists \alpha > 0$ such that $\text{Re } \lambda(\bar{A}) \neq 0$ and

$$\max \text{ Re } \lambda(\bar{A}) \geq \alpha \tag{2.15}$$

then $\exists \epsilon_0 > 0$ such that $\forall \epsilon \in (0, \epsilon_0)$, the zero solution of (2.1) is unstable.

Proof: We first prove part i). Assumption (2.13) means that $\forall \delta > 0$, $\exists T(\delta) > 0$ such that

$$|\bar{A}_T(s) - \bar{A}| \leq \delta, \qquad \forall s \in \mathbb{R}_+. \tag{2.16}$$

From (2.9) with $\|A(\cdot)\|_\infty = m$, we have

$$\begin{aligned}|F(s+T, s)| &\leq \exp[\epsilon T\mu(\bar{A} + \bar{A}_T(s) - \bar{A}] + (\epsilon Tm)^2 \exp(\epsilon Tm) \\ &\leq \exp[\epsilon T(\mu(\bar{A}) + \delta)] + (\epsilon Tm)^2 \exp(\epsilon Tm).\end{aligned}$$

Since $\text{Re } \lambda(\bar{A}) < 0$, there is a constant matrix $P = P' > 0$ which satisfies the Lyapunov equation,

$$\bar{A}'P + P\bar{A} + 2I = 0. \tag{2.18}$$

Now, choose as a norm on $\mathbb{R}^n$

$$|x| = (x'Px)^{1/2}. \tag{2.19}$$

From (2.6d) and using (2.18), we then have

$$\begin{aligned}\mu(\bar{A}) &= -\frac{1}{2}\max_i \lambda_i\{P^{-1/2}(\bar{A}'P + P\bar{A})P^{1/2}\} \\ &= -\min_i \lambda_i\{P^{-1}\} = -\alpha.\end{aligned} \tag{2.20}$$

Hence, (2.17) becomes

$$|F(s+\tau, s)| \leq \exp[-\epsilon T(\alpha - \delta)] + (\epsilon TM)^2 \exp(\epsilon TM). \tag{2.21}$$

By assumption (2.13) it is always possible to select $T(\delta)$ in (2.16) such that $\delta < \alpha$. By inspection of (2.21), there then exists $\epsilon_0 > 0$ such that $\forall \epsilon \in (0, \epsilon_0)$, $|F(s + T, s)| < 1$, $\forall s \in \mathbb{R}_+$, which completes the proof of part i). Part ii) can be proven in an analogous manner, starting with (2.10) and using (2.18) with $\bar{A}$ replaced by $-\bar{A}$. Note that here $\text{Re } \lambda(\bar{A}) \neq 0$ is required explicitly. □

Discussion: The results in Theorems 2.1 and 2.2 generalize some results obtained by averaging methods such as those described by Hale [10], or as obtained by Coppel [18] using the notion of integral smallness. Theorem 2.2 is a classical result of averaging theory, except that, as stated, it allows for functions which are not necessarily almost periodic. The class of functions allowed in Theorem 2.2—regulated, bounded, with a uniform average—is not precisely characterized. Obviously it includes the class of asymptotically almost periodic functions of the form

$$A(t)=A_0(t)+A_1(t) \tag{2.22}$$

where $A_0(t)$ is almost periodic and $A_1(\cdot) \in L_1^{n\times n}$.

Theorem 2.1 considers a larger class of functions—those without an average—at the expense of a weaker result: the stability–instability boundary is not as sharp as in Theorem 2.2.

An example of a function which satisfies the conditions of Theorem 2.1, but not of Theorem 2.2, is

$$A(t)=A_0+(1/\sqrt{2})A_1\ (\sin\ \log\ t+\cos\ \log\ t) \tag{2.23}$$

where $A_i = A_i' > 0$, $i = 0, 1$, such that $A_0 - A_1 > 0$. This function does not have a uniform average, as can be seen from

$$\frac{1}{T}\int_s^{s+T} A(t)\ dt=A_0 + \frac{A_1}{\sqrt{2}}\left(\frac{s+T}{T}\sin\ \log\ (s+T)-\frac{s}{T}\sin\ \log\ s\right).$$

However, it satisfies the conditions of Theorem 2.1 because from (2.23)

$$\frac{1}{T}\int_s^{s+T} A(t)\ dt \geq A_0-A_1>0, \qquad \forall s \in \mathbb{R}_+,\ \ \forall T>0. \tag{2.25}$$

Condition (2.7), which is the basis for the u.a.s. property, has some interesting interpretations. In the first place, since $\alpha > 0$ is a constant, condition (2.10) provides a uniform bound on the sequence of sample-average measures $\{\mu[\bar{A}_T(kT)],\ k \in \mathbb{Z}_+\}$. From the definition (2.5), the measure is dependent on the underlying vector norm. Suppose we choose as the vector norm $|x| = (x'Px)^{1/2}$ with $P = P' > 0$ a constant matrix. This was done in the proof of Theorem 2.2 where P was given as the solution to (2.18). In general, however, we have from (2.6e) that

$$\mu[\bar{A}_T(kT)]=\frac{1}{2}\max_i \lambda_i\{P^{-1/2}[\bar{A}_T'(kT)P+P\bar{A}_T(kT)]P^{-1/2}\}. \tag{2.26}$$

If there is a constant matrix $P = P' > 0$ such that

$$\frac{1}{2}\max_i \lambda_i\{\bar{A}_T'(kT)P+P\bar{A}_T(kT)\} \leq -1, \qquad \forall k \in \mathbb{Z}_+ \tag{2.27}$$

then (2.7) holds with the choice

$$\alpha=\min_i \lambda_i(P^{-1}). \tag{2.28}$$

Observe that (2.27) is *not* equivalent to

$$\mathrm{Re}\ \lambda\{\bar{A}_T(kT)\}<0, \qquad \forall k \in \mathbb{Z}_+. \tag{2.29}$$

This latter condition means there is a *sequence* of matrices $\{\bar{P}(k) = \bar{P}(k)' > 0,\ k \in \mathbb{Z}_+\}$ which satisfy

$$\max_i \lambda_i\{\bar{A}_T'(k)\bar{P}(k)+\bar{P}(k)\bar{A}_T(kT)\} = -2, \qquad \forall k \in \mathbb{Z}_+. \tag{2.30}$$

Unfortunately, it makes no sense to choose a time-varying norm, e.g., $|x| = (x'\bar{P}(k)x)^{1/2}$. Hence, condition (2.27) provides a means to satisfy (2.7), provided that *constant* P can be found.

A simple sufficient condition for (2.20) is that

$$\max_i \lambda_i\{\bar{A}(k)+\bar{A}(k)'\} \leq -2\alpha_0, \qquad \forall k \in \mathbb{Z}_+ \tag{2.31}$$

where α_0 is a positive constant. Hence, we can take $P = (1/\alpha_0)I$ in (2.27), and thus (2.7) holds with $\alpha = \alpha_0$. We will discuss condition (2.31) further when we specialize Theorem 2.1 for adaptive systems in Section IV.

Theorem 2.1 also requires that $\epsilon T > 0$ be sufficiently small, i.e., that $\epsilon T \in (0,\eta)$. From the proof of Theorem 2.1 we can extract a value for η and also state bounds on the exponential rates of growth or decay of the transition matrix $F(t,\tau)$ for all $t \geq \tau$. Specifically, we have the following.

Corollary 2.1: If $A(t)$ is regulated and bounded with $\|A(\cdot)\|_\infty \leq m$, then the following holds.

i) Whenever (2.7) holds for some $T > 0$, the zero solution of (2.1) is u.a.s. $\forall \epsilon T \in (0, \eta)$, i.e.,

$$|F(t,\ \tau)| \leq M\ \exp\ (-\epsilon(t-\tau)\beta) \tag{2.32}$$

where η, M, and β satisfy:

$$\exp\ (-\eta\alpha)+r(\eta m)=1$$
$$M=\exp\ (\epsilon T(m+\beta))>1$$
$$\exp\ (-\epsilon T\beta)=\exp\ (-\epsilon T\alpha)+r(\epsilon Tm)<1. \tag{2.33}$$

ii) Whenever (2.8) holds for some $T > 0$, the zero solution of (2.1) is unstable $\forall \epsilon T \in (0, \eta)$, i.e.,

$$|F(t,\ \tau)| \geq M\ \exp\ (\epsilon(t-\tau)\beta) \tag{2.34}$$

where η, M, and β satisfy

$$\exp\ (\eta\alpha)-r(nm)=1$$
$$M=\exp\ (-\epsilon T(m+\beta))<1$$
$$\exp\ (\epsilon T\beta)=\exp\ (\epsilon T\alpha)-r(\epsilon Tm)<1. \tag{2.35}$$

□

III. Stability of Linearized Adaptive System

In this section we apply the results of Section II to the linearized adaptive system (1.1) under slow adaptation, i.e., small $\epsilon > 0$. The first step is to transform (1.1) into a form suitable for application of Theorem 2.1. This is accomplished by a *time-scale decomposition*. That is, under slow adaptation the parameters $\theta(t)$ change much more slowly than the internal states of the dynamical system H. This suggests approximating (1.1) by the system.

$$\dot{\theta}=\epsilon[f-(\phi H\phi')\theta] \tag{3.1}$$

for which Theorem 2.1 would apply, i.e., replace $A(t)$ in (2.1) with $-\epsilon(\phi H\phi')(t)$. We start with the following intermediate result, developed in [20] and based on the discrete-time formulation in [21].

Lemma 3.1: System (1.1) is equivalent to

$$\dot{\theta}=\epsilon[f-R\theta+\epsilon W(f,\ \theta)] \tag{3.2}$$

where R is the time-varying matrix

$$R(t)=(\phi H\phi')(t) \tag{3.3}$$

and $W(f,\theta)$ is the linear integral operator

$$W(f,\ \theta)=\phi G_{\phi}[f-\phi H(\phi'\theta)] \tag{3.3}$$

with G_{ϕ}, the linear integral operator whose kernel is

$$g_{\phi}(t,\ \tau)=\int_0^{\tau} h(t-s)\phi'(s)\ ds, \qquad 0\le\tau\le t. \tag{3.5}$$

Proof: Integrating by parts gives

$$H(\phi'\theta)=(H\phi')\theta-G_{\phi}\dot{\theta}.$$

Thus,

$$W(f,\ \theta)=\frac{1}{\epsilon}\ \phi[(H\phi')\theta-H(\phi'\theta)]$$

$$=\frac{1}{\epsilon}\ \phi G_{\phi}\dot{\theta} \qquad \text{by (3.5)}$$

$$=\phi G_{\phi}[f-\phi H(\phi'\theta)]$$

by (1.1). □

Discussion: The decomposition of (1.1) into (3.2) is illustrated by the feedback system shown in Fig. 1. Hence, for small ϵ, the operator $\epsilon W(f,\ \theta)$ has little effect on system stability, and $\dot{\theta}=\epsilon(f-R\theta)$ provides the dominating stabilizing force. We will prove this assertion in Theorem 3.1 below.

If $H(s)$ is rational, i.e., $\exists A\in\mathbb{R}^{n\times n}$ and $b,\ c\in\mathbb{R}^{n}$ such that

$$H(s)=c'(sI-A)^{-1}b$$

then the decomposition (3.2) is essentially equivalent to the L-transformation in [9], which is also a Lyapunov transformation, i.e., both original and transformed systems have identical (Lyapunov) stability properties; see [24, p. 117]. In this case system (3.1) is equivalently represented in state form as

$$\dot{\theta}=\epsilon(f(t)-\phi(t)c'z), \qquad \theta(0)=\theta_0 \tag{3.7a}$$

$$\dot{z}=Az+b\phi'(t)\theta \quad z(0)=0. \tag{3.7b}$$

Using the "L-transformation" [9]

$$\xi=z-L(t)\theta \tag{3.8a}$$

where $L(t)$ satisfies

$$\dot{L}=AL+b\phi'(t) \tag{3.8b}$$

gives

$$\dot{\theta}=\epsilon[f(t)-\phi(t)c'(L(t)\theta+\xi)] \tag{3.9a}$$

$$\dot{\xi}=A\xi-\epsilon L(t)[f(t)-\phi(t)c'(L(t)\theta+\xi)]. \tag{3.9b}$$

Since $\theta(0)=\theta_0$ and $z(0)=0$ by definition (1.1), it follows that by assigning $L(0)=0$ we have[1]

$$R(t)=\phi(t)c'L(t) \tag{3.10}$$

and hence, (3.9a) becomes

$$\dot{\theta}=\epsilon[f(t)-R(t)\theta-\phi(t)c'\xi]. \tag{3.11}$$

Since $\xi(0)=0$ from (3.8) and Re $\lambda(A)<0$ because $H(s)$ is stable, it follows that $\xi(t)=O(\epsilon)$. Thus, (3.11) is dominated for small ϵ by $\dot{\theta}=\epsilon(f-R\theta)$. Consequently, both the "L-transformation" and the operator decomposition (3.2) are qualitatively equivalent for small ϵ.

[1] Note that the choice of initial condition for $L(0)$ is immaterial when discussing asymptotic stability properties, i.e., since A is stable, different initial conditions give rise to different exponentially fast decaying transients.

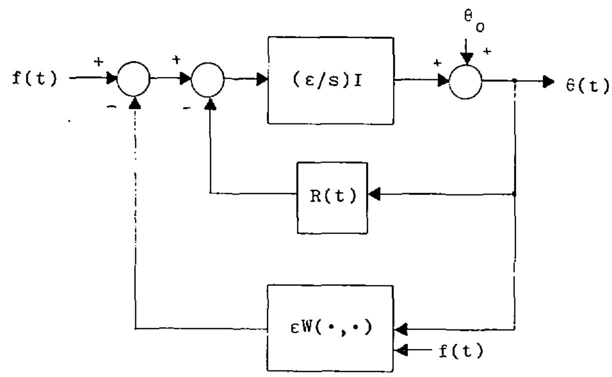

Fig. 1.

Using Lemma 3.1, we now state conditions for exponential stability of system (1.1), i.e., the map $(\theta_0, f)\to\theta$.

Theorem 3.1: Assume that

A1) $\dot{\theta}=-\epsilon R(t)\theta,\ R(t)=(\phi H\phi')(t)$, is u.a.s. with transition matrix $F(t,\tau)$ overbounded by

$$|F(t,\ \tau)|\le Me^{-\epsilon\beta(t-\tau)}, \qquad \forall t\ge\tau\ge 0 \tag{3.12}$$

A2) The impulse response $h(t)$ of H satisfies

$$|h(t)|\le Ke^{-\alpha t}, \qquad \forall t\ge 0. \tag{3.13}$$

Under these conditions, $\exists\epsilon_0>0$ such that $\forall\epsilon\in(0,\epsilon_0)$, system (1.1) is exponentially stable. Specifically, if

$$\epsilon_0=\min\ \{\alpha/\beta,\ \epsilon_1\}, \qquad \rho(\epsilon_1)=0 \tag{3.14}$$

then

$$|\theta(t)|\le M|\theta_0|e^{-\epsilon\rho(\epsilon)t}+\int_0^t \epsilon m(\epsilon)e^{-\epsilon\rho(\epsilon)(t-\tau)}|f(\tau)|\ d\tau \tag{3.15}$$

where

$$m(\epsilon)=M[1+\epsilon\|\phi\|_{\infty}^{2}K/(\alpha-\epsilon\beta)^2] \tag{3.16}$$

$$\rho(\epsilon)=\beta-\epsilon M\|\phi\|_{\infty}^{4}K^2/(\alpha-\epsilon\beta)^2. \tag{3.17}$$

Proof: Using the decomposition from Lemma 3.1 gives the following expression for (1.1):

$$\theta(t)=F(t,\ 0)\theta_0+\epsilon(W_1 f)(t)-\epsilon^2(W_2\theta)(t) \tag{3.18}$$

where $W_1,\ W_2$ are linear integral operators given by

$$W_1=F(I+\epsilon\phi G_{\phi}) \tag{3.19a}$$

$$W_2=F\phi G_{\phi}\phi H\phi' \tag{3.19b}$$

and where F has kernel $F(t,\tau)$. We first show that W_1 and W_2 are exp. stable integral operators.

For *any* integral operator W with kernel $w(t,\tau)$ we have

$$(Wu)(t)=\int_0^t w(t,\ \tau)u(\tau)\ d\tau$$

$$=\int_0^t e^{-\sigma(t-\tau)}[e^{\sigma(t-\tau)}w(t,\ \tau)]u(\tau)\ d\tau.$$

Therefore, W is exponentially stable iff $\exists\sigma>0$ such that

$$\sup_{t\ge\tau}\ |e^{\sigma(t-\tau)}w(t,\ \tau)|<\infty.$$

Let the superscript notation $(\cdot)^\sigma$ denote *exponential weighting*, i.e., $(x^\sigma)(t) = e^{\sigma t}x(t)$. Hence, $(Wu)^\sigma = W^\sigma u^\sigma$ where W^σ is the linear integral operator with kernel $e^{\sigma(t-\tau)}w(t, \tau)$. Now, following [16, p. 119], let $\|W\|_b$ be defined by

$$\|W\|_b = \sup_{t\geq\tau} |w(t, \tau)|. \quad (3.20)$$

Hence, W is exponentially stable iff $\exists\ \sigma > 0$ such that

$$\|W^\sigma\|_b < \infty.$$

Observe also that if G_1 and G_2 are linear integral operators, then

$$\|G_1 G_2\|_b \leq \|G_1\|_b \gamma_1(G_2). \quad (3.21)$$

Applying these relations to (3.19) for some $\sigma > 0$ gives

$$\|W_1^\sigma\|_b \leq \|F^\sigma\|_b[1+\epsilon\|\phi\|_\infty \gamma_1(G_\phi^\sigma)]$$

$$\|W_2^\sigma\|_b \leq \|F^\sigma\|_b \|\phi\|_\infty^3 \gamma_1(G_\phi^\sigma)\gamma_1(H^\sigma).$$

Choose $\sigma = \epsilon\beta$ with $\epsilon \in (0, \epsilon_0)$, ϵ_0 given by (3.14). Using (3.19)–(3.21) gives

$$\|F^\sigma\|_b \leq M$$

$$\gamma_1(H^\sigma) \leq K/(\alpha - \epsilon\beta)$$

$$\gamma_1(G_\phi^\sigma) \leq \|\phi\|_\infty K/(\alpha-\epsilon\beta)^2.$$

Hence,

$$\|W_1^\sigma\|_b \leq M[1+\epsilon\|\phi\|_\infty^2 K/(\alpha-\epsilon\beta)^2] = m(\epsilon)$$

$$\|W_2^\sigma\|_b \leq M\|\phi\|_\infty^4 K^2/(\alpha-\epsilon\beta)^3 = [\beta - \rho(\epsilon)]/\epsilon$$

where $m(\epsilon)$ and $\rho(\epsilon)$ are defined in (3.16) and (3.17). Going back to (3.18), we now have

$$|\theta(t)| \leq Me^{-\epsilon\beta t}|\theta_0| + \int_0^t e^{-\epsilon\beta(t-\tau)}\{\epsilon m(\epsilon)|f(\tau)| + \epsilon[\beta-\rho(\epsilon)]|\theta(\tau)|\}\ d\tau. \quad (3.22)$$

The result (3.15)–(3.17) follows by directly applying the Bellman–Gronwall lemma to (3.22). □

Discussion: Under slow adaptation, Theorem 3.1 shows that (1.1) is exponentially stable if $\dot{\theta} = -\epsilon R(t)\theta$ is u.a.s. Hence, we can apply Theorem 2.1, with $A(t)$ replaced by $-\epsilon R(t)$, and arrive at stability condition (2.7), that is:

System (1.1) is exponentially stable for all small $\epsilon > 0$ if $\exists T > 0$ such that

$$\mu[-\bar{R}(k)] < 0, \qquad \forall k \in \mathbb{Z}_+ \quad (3.23)$$

where $\bar{R}(k)$ is the kth sample average

$$\bar{R}(k) = \frac{1}{T}\int_{kT}^{(k+1)T} R(t)\ dt. \quad (3.24)$$

Using (2.27), condition (3.23) holds if there is a constant matrix $P = P' > 0$ such that

$$\frac{1}{2}\min_i \lambda_i\{\bar{R}(k)P + P\bar{R}(k)'\} \geq 1, \qquad \forall k \in \mathbb{Z}_+. \quad (3.25)$$

Moreover, a sufficient condition for (3.23) is that

$$\min_i \lambda_i\{\bar{R}(k) + \bar{R}(k)'\} > \alpha_0, \qquad \forall k \in \mathbb{Z}_+ \quad (3.26)$$

where α_0 is a positive constant. Comparing (3.26) to (3.25) reveals that $P = (1/\alpha_0)I$, which means the interval contraction of $\theta(kT) \to \theta((k+1)T)$ is scaled uniformly, i.e., $\theta'(kT+T)\theta(kT+T) < \theta'(kT)\theta(kT)$. The scaling implications are discussed further in Section IV to follow.

IV. Frequency-Domain Stability Conditions

In this section we reformulate condition (3.12) in the frequency domain. This involves the Fourier transform $H(j\omega)$ and an appropriately defined expression for the spectrum of $\phi(t)$. We show that (3.12) requires that $\phi(t)$ have a persistent excitation property, and that the dominant excitation be at those frequencies for which Re $H(j\omega) > 0$.

The first requirement is that $\phi(t)$ be restricted to those functions which have a Fourier series representation on any finite interval. A known class of such functions is defined as follows (see, e.g., [19]).

Definition: A function $f(\cdot): \mathbb{R}_+ \to \mathbb{R}^n$ is a C_δ^n function if it is regulated, bounded and $\exists$ a constant $\delta > 0$ such that any two points $t_1, t_2 \in \mathbb{R}_+$ where $f(\cdot)$ is discontinuous are separated by at least an interval δ, i.e., $|t_1 - t_2| \geq \delta$.

Frequency-domain stability conditions for the stability of (3.1) can now be stated.

Theorem 4.1: Assume in (1.1) that

A1)

$$|h(t)| \leq Ke^{-\alpha t}, \qquad \forall t \geq 0. \quad (4.1)$$

A2) $\phi \in C_\delta^p$ with piecewise Fourier series representation $\forall k \in \mathbb{Z}_+$:

$$\phi(t) \sim \sum_{\omega\in\Omega_k} \alpha_k(\omega)e^{j\omega t}, \qquad \forall t \in [kT, (k+1)T), \quad T \geq \delta \quad (4.2)$$

where Ω_k is the set of distinct Fourier exponents and $\alpha_k(\cdot)$ the corresponding Fourier coefficients. Let $B(k) \in \mathbb{R}^{p\times p}$ be defined by[2]

$$B(k) = \sum_{\omega\in\Omega_k} \alpha_k(\omega)\bar{\alpha}_k(\omega)'H(-j\omega), \qquad \forall k \in \mathbb{Z}_+. \quad (4.3)$$

Under these conditions, the following holds.

i) If $\exists T \geq \delta$ such that

$$\mu[-B(k)] < -2\|\phi\|_\infty^2(K/\alpha^2)/T, \qquad \forall k \in \mathbb{Z}_+ \quad (4.4)$$

then $\exists\epsilon_0 > 0$ such that $\forall\epsilon \in (0, \epsilon_0)$, system (1.1) is exponentially stable.

ii) If $\exists T \geq \delta$ such that

$$\mu[B(k)] < -2\|\phi\|_\infty^2(K/\alpha^2)/T, \qquad \forall k \in \mathbb{Z}_+ \quad (4.5)$$

then $\exists\epsilon_0 > 0$ such that $\forall\epsilon \in (0,\epsilon_0)$, system (1.1) is unstable.

Remarks:

1) The representation (4.2) for $\phi(t)$ specifies the local frequency content over $t \in [kT, (k+1)T]$. Such a representation—if not given—can always be found if $\phi(t) \in C_\delta^p$ [17]; then (4.2) can be obtained via the Fourier series of the T-periodic function

$$\phi_k(t) = \phi(t+mT) \qquad t \in [(k-m)T, (k-m+1)T]$$
$$\forall k \in \mathbb{N}, \forall m \in \mathbb{Z}.$$

Notice $\phi_k(t)$ is well defined on $\mathbb{R}$ and has a Fourier series representation:

$$\phi_k(t) \sim \sum_{m\in Z} \alpha_k(j\omega_m)e^{j\omega_m t}; \qquad \forall t \in \mathbb{R}$$

[2] Overbar $(\bar{\cdot})$ denotes complex conjugation.

where

$$\alpha_k = \alpha^*_{-k}, \quad \omega_m = 2\pi m/T$$

and hence

$$\phi_k(t) \sim \sum_{m \in Z} \alpha_k(j\omega_m) e^{j\omega_m t}; \qquad \forall t \in [kT, (k+1)T] \quad (4.6)$$

which is of the form (4.2).

2) The matrix $B(k)$ can be equivalently expressed as the sample average value of the T-periodic part of $(\phi_k H \phi_k')(t)$, i.e.,

$$B(k) = \frac{1}{T} \int_{kT}^{(k+1)T} \phi_k(t) \Psi_k(t)' \, dt \quad (4.7)$$

where $\Psi_k(t)$ is the T-periodic part of $H(\phi_k)(t)$, i.e.,

$$\Psi_k(t) = \int_{-\infty}^{t} h(t-\tau)\phi_k(\tau) \, d\tau = \sum_{\omega \in \Omega_k} H(j\omega)\alpha_k(\omega)e^{j\omega t}, \qquad \forall t \in \mathbb{R}. \quad (4.8)$$

Proof: To prove part i), it follows from Lemma 3.1 and Theorem 3.1 that it is only necessary to show that (4.4) implies

$$\mu[-\bar{R}(k)] < 0, \qquad \forall k \in \mathbb{Z}_+$$

where $\bar{R}(k)$ is given by (3.24). We start by defining

$$\tilde{B}(k) = \bar{R}(k) - B(k)$$

with $B(k)$ from (4.3). Using (4.8) gives

$$\tilde{B}(k) = \frac{1}{T} \int_{kT}^{(k+1)T} \phi_k(t) \tilde{\Psi}_k'(t) \, dt$$

where

$$\tilde{\Psi}_k(t) = (H\phi)(t) - \Psi_k(t) = \int_0^t h(t-\tau)\phi(\tau) \, d\tau - \int_{-\infty}^{t} h(t-\tau)\phi_k(\tau) \, d\tau = \int_0^{kT} h(t-\tau)\phi(\tau) \, d\tau - \int_{-\infty}^{kT} h(t-\tau)\phi_k(\tau) \, d\tau.$$

The last line follows from (4.6), i.e., $\phi_k(t) = \phi(t)$ for $t \in [kT, (k+1)T)$. Using (4.1) gives

$$|\tilde{\Psi}_k(t)| \le (2\|\phi\|_\infty K/\alpha) e^{-\alpha t}$$

from which it follows that

$$|\tilde{B}(k)| \le 2\|\phi\|_\infty^2 (K/\alpha^2)/T, \qquad \forall k \in \mathbb{Z}_+.$$

This together with inequality (2.6a) proves part i). Part ii) follows analogously by replacing $\bar{R}(k)$ with $-\bar{R}(k)$. □

If $\phi(t)$ is further restricted so that it has a uniform average, then we can sharpen the stability-instability boundary. For example, if $\phi(t)$ is almost periodic, then a Fourier series representation exists $\forall t \in \mathbb{R}_+$, and thus, it has an average [10]. The stability conditions for this case are stated as follows.

Theorem 4.2: Suppose in (1.1) that $\phi(t)$ is almost periodic with generalized Fourier series

$$\phi(t) \sim \sum_{\omega \in \Omega} \alpha(\omega) e^{j\omega t}, \qquad \forall t \in \mathbb{R}_+ \quad (4.9)$$

where $\Omega \in \mathbb{R}$ are the distinct Fourier exponents and $\{\alpha(\omega), \omega \in \Omega\}$ are the Fourier coefficients. Define the matrix B by

$$B = \sum_{\omega \in \Omega} \alpha(\omega)\bar{\alpha}(\omega)' H(-j\omega). \quad (4.10)$$

If Re $\lambda(B) \neq 0$, then $\exists \epsilon_0 > 0$ such that $\forall \epsilon \in (0, \epsilon_0)$, system (1.1) is

i) exponentially stable if

$$\text{Re } \lambda(B) < 0 \quad (4.11)$$

ii) unstable if

$$\max_i \text{Re } \lambda_i(B) > 0. \quad (4.12)$$

Remark: The proof of Theorem 4.2 is entirely analogous to that of Theorem 4.1. Theorem 4.2 is the result obtained in [9] when $\phi(t)$ is almost periodic. Theorem 4.1 is a generalization to $\phi(\cdot) \in C_\delta^p$.

V. Discussion of Results

A. Effect of Transients on Sample Average

An informative interpretation of stability condition (4.4) is that the average energy in the T-periodic part of $(\phi_k H \phi_k')(t)$ must dominate (or overcome) the possibly negative efforts of the transient terms. In other words, the period T must be sufficiently larger than the dominant time constant of H, i.e., $T \gg 1/\alpha$. Note that the term $2\|\phi\|_\infty^2(K/\alpha^2)$ essentially arises from initial conditions or stored energy in H at $t = kT$. Obviously, when $\phi(t)$ has a uniform average, it is always possible to select T to be sufficiently large, e.g., as shown in the proof of Theorem 2.2.

Using (2.27), condition (4.4) holds if there is a constant matrix $P = P > 0$ such that $\forall k \in \mathbb{Z}_+$,

$$\min_i \lambda_i[Q(k)] \le 1, \qquad \forall k \in \mathbb{Z}_+ \quad (5.1a)$$

where

$$Q(k) = \frac{1}{2} \sum_{\omega \in \Omega_k} H(-j\omega)[PX_k(\omega) + \bar{X}_k(\omega)P] \quad (5.1b)$$

$$X_k(\omega) = \alpha_k(\omega)\bar{\alpha}_k(\omega)' = [\bar{X}_k(\omega)]'. \quad (5.1c)$$

B. Relation to Persistent Excitation

A necessary condition for the existence of P which satisfies (5.1) is that, for some finite integer $q \ge (p-1)/2$ and $\forall k \in \mathbb{Z}_+$,

$$\text{rank } [\alpha_k(0), \alpha_k(\omega_1), \cdots, \alpha_k(\omega_q), \bar{\alpha}_k(\omega_1), \cdots, \bar{\alpha}_k(\omega_q)] = p. \quad (5.2)$$

If this were not the case, then $\min_i \lambda_i[Q_k(P)] = 0$, $\forall k \in \mathbb{Z}_+$, and $\forall P = P' > 0$. Hence, Theorem 4.1 implicitly restricts $\phi(\cdot) \in C_\delta^p$ to those functions whose (time-varying) Fourier coefficients satisfy the rank condition above. This class of functions, however, are precisely those which can be categorized as persistently exciting [1].

Definition: A function $f(\cdot): \mathbb{R}_+ \to \mathbb{R}^n$ is *persisently exciting* (PE) over an interval h if it is regulated, bounded, and $\exists$ constants $h > 0$ and $\beta > 0$ such that

$$\min_i \lambda_i \left(\frac{1}{h} \int_s^{s+h} f(t) f(t)' \, dt \right) \ge \beta, \qquad \forall s \in \mathbb{R}_+. \quad (5.3)$$

Denote such functions by $f(\cdot) \in \text{PE}^n(h, \beta)$.

It follows from the definition that if $\phi(\cdot) \in \mathrm{PE}^p(h, \beta) \cap C_\delta^p$, then the rank condition (5.2) will hold for any $T \geq h \geq \delta$, and thus, (5.1) *may* be satisfied for some matrix P. The point to emphasize is that persistent excitation is not sufficient for stability, except in the case when $H(s)$ is SPR [11]. Thus, we can view (5.1) as a *signal dependent positivity condition.* In general, the PE condition is necessary for stability, but as seen from (4.5) in Theorem 4.1, even if it holds, the system can still be unstable.

C. Parameter Scaling

The matrix P in (5.1) can be viewed as a *scaling* of the parameter vector. That is, if (5.1) holds for some P, then for all small $\epsilon > 0$, $\dot{\theta} = -\epsilon R(t)\theta$ is u.a.s. in the sense that $\theta(t)'P\theta(t) \to 0$ exponentially fast as $t \to \infty$. Thus, parameters will tend to converge with different scalings. If for some *given* signal $\phi(t)$, the determined scaling matrix P gives unwanted responses, then the signal can be reshaped so as to produce a more desirable scaling. The difficulty is in finding the matrix P. If $\phi(t)$ is almost periodic, then Theorem 4.3 holds, and we can take P as the solution to $P\bar{A} + \bar{A}'P = -2I$. When $\phi(t)$ has a sample-average, there is no simple means to find P.

If there is sufficient *a priori* knowledge about the effect of parameters on the system, then this information will provide the *desired* scaling in the following sense. It is always possible to prescale θ and then select $P = (1/\alpha_0)I$ where α_0 is some positive constant. With this choice, condition (5.1) becomes

$$\min_i \lambda_i \left\{ \sum_{\omega \in \Omega_k} H(-j\omega) \operatorname{Re}\,[X_k(\omega)] \right\} \geq \alpha_0, \qquad \forall k \in \mathbb{Z}_+. \tag{5.4}$$

This is equivalently expressed as

$$\min_i \lambda_i \left\{ \sum_{\omega \in \Omega_k} \operatorname{Re}\,[H(j\omega)] \operatorname{Re}\,[X_k(\omega)] \right\} \geq \alpha_0/2, \qquad \forall k \in \mathbb{Z}_+ \tag{5.5}$$

which has a more informative interpretation in terms of the usual positivity conditions on H. For example, a strictly proper transfer function $H(s)$ is strictly positive real (SPR) if it is exponentially stable and $\exists$ constant $\rho > 0$ such that [16]

$$\operatorname{Re}\,[\hat{H}(j\omega)] \geq \rho |\hat{H}(j\omega)|^2, \qquad \forall \omega \in \mathbb{R}_+.$$

This condition must hold at *every* frequency, whereas (5.5) requires $\operatorname{Re}[\hat{H}(j\omega)] > 0$ at those *discrete* frequencies in $\mathbb{R}_+$ where the magnitude of the input spectrum is large. Conversely, at those frequencies in $\mathbb{R}_+$ where $\operatorname{Re}\,[\hat{H}(j\omega)] < 0$, the magnitude of the input spectrum should be small. Since (5.5) will fail if $\operatorname{Re} \hat{H}(j\omega) < 0$, $\forall \omega \in \mathbb{R}_+$, it follows that $\operatorname{Re} \hat{H}(j\omega) > 0$ at some frequencies; hence the motivation to refer to (5.1) as a positivity condition.

D. Bounds on ϵ

The upper bound ϵ_0 on the size of $\epsilon > 0$ to ensure stability can be extracted from the proof of Theorem 4.1. Looking back over Theorem 3.1, Theorem 2.1, and subsequent discussions, we have

$$\epsilon_0 = \min\ \{\alpha/\beta,\ \epsilon_1,\ \epsilon_2\} \tag{5.6}$$

where ϵ_1 and ϵ_2 satisfy

$$\beta - \epsilon_1 M \|\phi\|_\infty^4 K^2/(\alpha - \epsilon_1\beta)^2 = 0 \tag{5.7}$$

$$\exp\ (-\epsilon_2 T\alpha_0) + (\epsilon_2 T\alpha_0)^2 \exp\ (\epsilon_2 T\alpha_0) = 1. \tag{5.8}$$

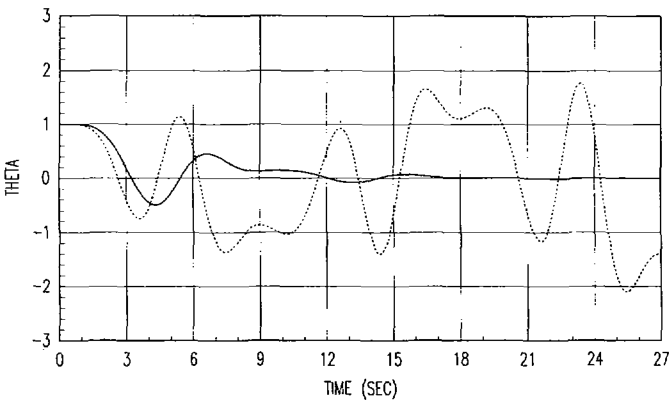

Fig. 2.

Recall from the proof of Theorem 3.1 that α_0, α, β, M, and K are defined from

$$\mu \left[-\frac{1}{T} \int_{kT}^{(k+1)T} (\phi H \phi')(t)\ dt \right] \leq -\alpha_0, \qquad \forall k \in \mathbb{Z}_+ \tag{5.9}$$

$$M = \exp\ (\epsilon T(m + \beta)) \tag{5.10}$$

$$|h(t)| \leq K \exp\ (-\alpha t) \tag{5.11}$$

$$\exp\ (-\epsilon T\beta) = \exp\ (-\epsilon T\alpha_0) + (\epsilon T\alpha_0)^2 \exp\ (\epsilon T\alpha_0). \tag{5.12}$$

E. A Limitation Arising from Averaging

Suppose $H(s)$ is SPR and (4.4) holds. Hence, system (1.1) is exponentially stable for all *small* $\epsilon > 0$. Since (4.4) holds, it follows that $\phi(t)$ is persistently exciting. However, from other arguments (see, e.g., [1]) we know that under these same conditions the zero solution of (3.1) is u.a.s. for *all* $\epsilon > 0$. Thus, Theorem 4.1 is conservative in this case in regard to the limitations on ϵ. However, when $H(s)$ is not SPR, Theorem 4.1 is now applicable, whereas the results in [1] do not apply. In fact, in this latter case, when ϵ gets too large then system (1.1) can be unstable, even if (4.4) holds. For example, if in (1.1) $\phi(t) = \sin(0.35t)$ and $H(s) = 1/(s^2 + 2s + 2)$, then condition (4.4) is satisfied. The simulations in Fig. 2 with $\theta(0) = 1$ show that the zero solution is u.a.s. for $\epsilon = 4$ but is completely unstable for $\epsilon = 8$.

F. Comparison with Averaging Analysis of Stochastic Recursive Algorithms

Comparing our results with the ordinary differential equation approach (ODE), used in the analysis of stochastic recursive algorithms [22], [23], we notice the following differences.

1) The ODE approach can deal with nonlinear recursions, while our analysis is restricted to the linear case. It is possible to extend our results to the nonlinear case, but this would introduce more technicalities (see, e.g., [20]), perhaps obscuring the main idea of "sample-averaging."

2) In the ODE approach it is assumed that the adaptable gain (our ϵ) converges to zero (and is not summable), while in the present contribution ϵ is a small positive constant.

3) The main difference lies in the condition imposed on the regressor vector sequence. Typically, the ODE approach relies on an ergodicity or mixing assumption to infer the existence of cesaro-mean along the sample paths ($\equiv$ average). Our conditions only involve finite sample path properties of the regressor vector and related quantities; this makes the concept of cesaro-mean or global average meaningless. In this sense, the ODE approach is closer to [9] where periodicity or almost periodicity are invoked to guarantee the existence of averages. In a way, our conditions

allow for a second slow time scale, the slower time scale on which the nature of the regressor vector is allowed to change.

4) The present approach yields instability results as well, a point not touched upon in the ODE approach.

VI. Conclusion

We have presented a method of averaging for linear time varying systems, allowing one to deal with general time motions, thus removing the classical restriction of almost periodicity.

This method can be applied to the nonlinear adaptive control problem after linearizing the system in the neighborhood of the tuned solutions. Both (local) stability and instability have been discussed. The conditions obtained to guarantee local stability can be expressed in frequency domain terms.

Acknowledgment

The authors are grateful to two reviewers for their careful reading and helpful comments. The authors have also benefited from "persistently exciting" discussions with R. Bitmead, C. R. Johnson, Jr., P. Kokotovic, L. Praly, and B. Riedle.

References

[1] B. D. O. Anderson, "Exponential stability of linear equations arising in adaptive identification," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 83-88, Feb. 1977.
[2] R. L. Kosut and C. R. Johnson, Jr., "An input-output view of robustness in adaptive control," *Automatica*, vol. 20, pp. 569-581, Sept. 1984.
[3] R. L. Kosut and B. D. O. Anderson, "A local stability analysis for a class of adaptive systems," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 86-89, Jan. 1986.
[4] J. LaSalle and S. Lefshetz, *Stability by Liapunov's Direct Method*. New York: Academic, 1961.
[5] Y. D. Landau, *Adaptive Control: The Model Reference Approach*. New York: Dekker, 1979.
[6] K. S. Narendra, Y. H. Lin, and L. Valavani, "Stable adaptive control design," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 440-448, June 1980.
[7] R. L. Kosut and B. Friedlander, "Robust adaptive control: Conditions for global stability," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 610-624, July 1985.
[8] S. Boyd and S. S. Sastry, "Necessary and sufficient conditions for parameter convergence in adaptive control," ERL Memo UCB/ERL M84/25, Berkeley, CA, Mar. 1984.
[9] B. Riedle and P. Kokotovic, "A stability-instability boundary for disturbance-free slow adaptation and unmodeled dynamics," in *Proc. 23rd IEEE Conf. Decision Contr.*, Las Vegas, NV, Dec. 1984, pp. 998-1001.
[10] J. Hale, *Ordinary Differential Equations*. New York: Wiley, 1969; Melbourne, FL: Krieger, 1980 (reprint).
[11] K. J. Astrom, "Interactions between excitation and unmodeled dynamics in adaptive control," in *Proc. 23rd IEEE Conf. Decision Contr.*, Las Vegas, NV, Dec. 1984, pp. 1276-1281.
[12] C. E. Rohrs, L. S. Valavani, M. Athans, and G. Stein, "Robustness of adaptive control algorithms in the presence of unmodeled dynamics," in *Proc. 21st IEEE Conf. Decision Contr.*, Orlando, FL, Dec. 1982.
[13] R. R. Bitmead, B. D. O. Anderson, and T. Ng, "Convergence rate estimates for LMS filters," in *Proc. IFAC*, Budapest, Hungary, July 1984.
[14] B. D. O. Anderson, R. R. Bitmead, C. R. Johnson, Jr., and R. L. Kosut, "Stability theorems for the relaxation of the strictly positive real condition in hyperstable adaptive systems," in *Proc. 23rd IEEE Conf. Decision Contr.*, Las Vegas, NV, Dec. 1984, pp. 1286-1291.
[15] P. A. Ioannou, and P. V. Kokotovic, *Adaptive Systems with Reduced Models*. New York: Springer-Verlag, 1983.
[16] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. New York: Academic, 1975.
[17] J. Dieudonne, *Foundations of Modern Analysis*. New York: Academic, 1960.
[18] W. A. Coppel, "Dichotomies and reducibility (II)," *J. Diff. Equations*, vol. 4, pp. 386-398, July 1968.
[19] J. S.-C. Yuan and W. M. Wohham, "Probing signals for model reference identification," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 530-538, Aug. 1977.
[20] B. D. O. Anderson, R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, I. M. Y. Mareels, L. Praly, and B. D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis*. Cambridge, MA: M.I.T. Press, 1986.
[21] R. R. Bitmead and C. R. Johnson, Jr., "Discrete averaging principles and robust adaptive identification," *Contr. Dynamic Syst.*, vol. 24, to be published, 1986.
[22] M. Metivier and P. Priouret, "Applications of a Kushner-Clark lemma to general class of stochastic algorithms," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 140-150, Mar. 1984.
[23] L. Ljung, "Analysis of stochastic gradient algorithms for linear regression problems," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 151-160, Mar. 1984.
[24] F. R. Gantmacher, *The Theory of Matrices*, vol. 2. New York: Chelsea, 1959.

Robert L. Kosut (S'67-M'75) was born in the Bronx, NY, on May 28, 1942. He received the B.S. degree in electrical engineering from the City College of New York in 1965, and the M.S. and Ph.D. degrees from the University of Pennsylvania, Philadelphia, in 1967 and 1969, respectively.

He is currently with Integrated Systems, Inc., Palo Alto, CA, and is also a Consulting Professor in the Department of Electrical Engineering at Stanford University, Stanford, CA. He was a Visiting Fellow on two occasions in the Department of Systems Engineering at the Australian National University, Canberra. He has previously held positions at Grumman Aerospace Corp., Bethpage, NY, the Singer Company (both at the Link Division, Binghamton, NY, and then at the Kearfott Division, Little Falls, NJ), and at Systems Control Technology, Inc., Palo Alto. He has also managed a small business designing electronic musical products, as well as managing a television repair service. He is a coauthor of the text *Stability of Adaptive Systems: Passivity and Averaging Analysis* (Cambridge, MA: M.I.T. Press, 1986). His research interests are in adaptive control, system identification, and robust control.

Dr. Kosut is an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Brian D. O. Anderson (S'62-M'66-SM'74-F'75) was born in Sydney, Australia, in 1941. He received the B.S. degrees in pure mathematics and electrical engineering from the University of Sydney, Sydney, Australia, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1966.

He is currently Professor and Head of Department of Systems Engineering at the Australian National University, Canberra, Australia. From 1967 through 1981, he was Professor of Electrical Engineering at the University of Newcastle, New South Wales, Australia. He has also held appointments as Visiting Professor at a number of universities in the United States, Australia, and Europe. He is coauthor of several books and his research interests are in control and signal processing.

Dr. Anderson is a Fellow of the Australian Academy of Science, the Australian Academy of Technological Sciences, and an Honorary Fellow of the Institute of Engineers, Australia. Previously, he was Editor of *Automatica* and is currently a Vice-President of the IFAC.

Iven M. Y. Mareels was born in Aalst, Belgium, in 1959. He received the B.D. degree in electromechanical engineering from the State University of Ghent, Belgium, in 1982, and the Ph.D. degree from the Australian National University, Canberra, in July 1986.

Since October 1982, he has been a Research Assistant with the National Fund for Scientific Research, Belgium.