# TRANSIENT ANALYSIS OF ADAPTIVE CONTROL

by

*R.L. Kosut,*[1,4,5] *I.M.Y. Mareels,*[2,6] *B.D.O. Anderson,*[2]
*R.R. Bitmead,*[2] *and C.R. Johnson, Jr.*[3,4]

**Abstract** *The transient properties of adaptive control systems are examined. A simple example system is presented which exhibits both slow and fast transients. Methods for analyzing the transient behavior are discussed, including averaging, Floquet analysis, and fixed point theory.*

## 1. INTRODUCTION

In this paper we consider an analysis of the transient response of adaptive control systems. An understanding of the transient is required in order to satisfy practical requirements such as those arising from constraints on tracking response and disturbance attenuation. For example, consider an adaptive system subject to abrupt set-point changes, e.g., step inputs. Typical system requirements are stated in terms of rise-time, overshoot, undershoot, and settling-time. Unlike a non-adaptive system, two sets of such requirements are needed; one set determined by the goal of the adaptive system, i.e., when the adaptive parameters are near convergence, and another set of requirements dealing with the transient, i.e., when the adaptive system is learning. The latter requirements include reasonable length of time for learning as well as bounds on responses during learning imposed by hardware limitations.

Analysis of the adaptive systems transient will require sharper estimates of signal bounds and rate of convergence than currently exist. Consider the ideal case of perfect model matching, i.e., when there exists a *constant* unique setting of the adaptive parameters which produce zero error for all inputs. In this situation although it is possible to prove global stability and exponential parameter convergence, the system states can be arbitrarily large and the theory does not offer guidelines for adjustment, e.g., Goodwin and Sin (1984). Local stability analysis based on the method of averaging -- which is valid also in the non-ideal case -- provides some transient information but the theoretical result is restricted to parameter trajectories which vary slowly in a subset of the constant-parameter stability set, see e.g., Astrom (1983, 1984), Bodson et al. (1985), Riedle and Kokotovic (1986), and Anderson et al. (1986). Certainly one can argue that the latter is not restrictive in practical system tuning when the plant is slowly varying and initial parametrizations are close to a tuned setting. The drawback is that although the convergence rate is exponential, it is also very slow, whereas simulations show that onset of instability may produce very rapid learning, see e.g., Anderson (1983). Moreover, estimates of the convergence rate and the region of attraction obtained from the method of averaging can be quite conservative, and hence, do not provide a complete representation of the achievable transients leading to good performance, e.g., Mareels et al. (1986).

In this paper we present a simple example which demonstrates some of the transient behavior which is typical of an adaptive system. We show how an averaging analysis predicts the long term (slow) transient behavior, and does surprisingly well for some fast transients. Other methods for analyzing the transient are also discussed, including short-time interval averaging, Floquet analysis, and the use of the contraction mapping principal. Further detail on these methods can be found in Kosut and Bitmead (1986).

## 2. ADAPTIVE ERROR SYSTEM

A general structure for an adaptive system is shown in Fig. 1 and is described by the operator equations

$$\begin{bmatrix} e \\ \phi \end{bmatrix} = P \begin{bmatrix} w \\ \mu \end{bmatrix} = \begin{bmatrix} P_{ew} & P_{e\mu} \\ P_{\phi w} & P_{\phi\mu} \end{bmatrix} \begin{bmatrix} w \\ \mu \end{bmatrix} \tag{2.1}$$

$$\mu = -F(\theta)\phi \tag{2.2}$$

$$\theta = \Omega(\theta_0, e, \phi) \quad , \quad \theta(t_0) = \theta_0 \tag{2.3}$$

The adaptive system consists of three subsystems:

(1) the *plant subsystem* $P$ which takes exogenous inputs $w$ -- consisting of references and disturbances -- and the adaptive control inputs $\mu$ into the *error* $e$ and *regressor* $\phi$;

(2) the *control subsystem* which transforms $\phi$ into $u$ via the control matrix $F(\theta)$, which is parametrized by $\theta$ the adaptive parameter vector; and

(3) the *adaptation subsystem* $\Omega$ which uses the error signal $e$, the regressor signal $\phi$, and the initial parameter value $\theta_0$ to generate the parameter $\theta$.

The structure of (2.1)-(2.3) can describe most of the standard forms of either continuous-time, discrete-time, or hybrid adaptive controller. Details on this structure can be found elsewhere, e.g., Kosut and Johnson (1983), Anderson et al. (1986). For example, a typical form for (2.2) is the bilinear structure

$$\mu = -\theta^T\phi \tag{2.4}$$

Typical forms for (2.3) include the simple continuous time gradient algorithm

$$\dot{\theta} = \varepsilon\phi e \tag{2.5}$$

---

[1] Integrated Systems, Inc. and Dept. of EE, Stanford University, Palo Alto and Stanford, CA, USA.

[2] Dept. of Systems Engineering, Australian National University, Canberra, Australia

[3] School of Electrical Engineering, Cornell University, Ithaca, NY, USA.

[4] Research supported by NSF Grant INT-85-13400.

[5] Research supported by AFOSR Contract F49620-85-C-0094.

[6] Automatic Control Lab, University of Gent, Belgium.

or the discrete-time normalized gradient algorithm

$$\delta\theta = \varepsilon \frac{\phi e}{1+\rho|\phi|^2} \quad (2.6)$$

where $\delta$ is the difference operator, i.e., $(\delta\theta)(t) = \theta(t) - \theta(t-1)$. In (2.5), (2.6), $\varepsilon$ and $\rho$ are positive constants and $|\cdot|$ is the Euclidean norm, i.e., $|\phi| = (\phi^T\phi)^{1/2}$. The error signal $e$ depends on the adaptive system structure, for example, in model reference adaptive control $e = y - y_m$, where $y_m$ is the reference model output, whereas in equation error identification $e = y - \theta^T\phi$.

For illustrative purposes we will concentrate first on the continuous-time gradient algorithm (2.5). A convenient form for analysis is the adaptive *error system* which is formed by introducing the parameter error

$$\tilde{\theta}(t) = \theta(t) - \theta_* \quad (2.7)$$

We refer to $\theta_*$ as a *tuned* parameter, which is a constant vector of parameters producing desirable performance properties of (2.1), (2.4). When $\theta(t)$ is held fixed at $\theta_*$ the resulting system is referred to as the tuned system and is described by

$$\begin{bmatrix} e_* \\ \phi_* \end{bmatrix} = P \begin{bmatrix} w \\ \mu_* \end{bmatrix}$$

$$\mu_* = -F(\theta_*)\phi_* = -\theta_*^T\phi_* \quad (2.8)$$

The signals $e_*$, $\phi_*$, and $u_*$ are referred to as the tuned error, regressor, and control, respectively. If $P$ is a linear-time-invariant (LTI) operator with transfer matrix $P(s)$, then it can be shown (Kosut and Friedlander, 1985), that the adaptive system (2.1), (2.4), and (2.5) can be described in *error form* by

$$\begin{aligned} \dot{\tilde{\theta}} &= \varepsilon\phi e \quad , \quad \tilde{\theta}(t_0) = \tilde{\theta}_0 \\ e &= e_* - H_{ev}(\phi^T\tilde{\theta}) \\ \phi &= \phi_* - G(\phi^T\tilde{\theta}) \end{aligned} \quad (2.9)$$

where $G, H$ are stable LTI operators, dependent on $\theta_*$, with stable, proper transfer functions $G(s)$, $H(s)$, respectively. Stability of $G$, $H$ follows from the definition of the tuned parameter setting.

The system (2.9) can be shown to be globally stable, i.e., stable for all $\tilde{\theta}_0 \in \mathbb{R}^p$, provided that $H(s)$ is SPR (strictly positive real), and the tuned error $e_*(t)$ is bounded and decays exponentially fast to zero. Moreover, if $\phi_*(t)$ is persistently exciting then (2.9) is globally exponentially stable, i.e., $\tilde{\theta}(t) \to 0$ exponentially for all $\tilde{\theta}_0 \in \mathbb{R}^p$. These results are typical, but not particularly useful for a transient analysis because the signal bounds are crude. The following example will illustrate this.

### An Example

Consider the adaptive system of Figure 2 with:

$$\begin{aligned} &\text{Plant:} && \dot{y} = -a_p y + u \quad , \; y(0) = y_0 \\ &\text{Control:} && u = \theta y + a_m(r - y) \\ &\text{Reference:} && \dot{y}_m = a_m(r - y_m) \quad , \; y_m(0) = y_{mo} \\ &\text{Estimator:} && \dot{\theta} = -\varepsilon y e \quad , \; \theta(0) = \theta_0 \\ & && e = y - y_m \end{aligned} \quad (2.10)$$

with $a_m$ a positive constant. The error, regressor, and adaptive control signals are then, respectively

$$\begin{aligned} e &= y - y_m \\ \phi &= -y \\ \mu &= -\theta\phi \end{aligned} \quad (2.11)$$

By choosing the tuned parameter setting

$$\theta_* = a_p \quad (2.12)$$

the transfer function with $\theta(t) = \theta_*$ from $r \to y$ is identical with that of the reference model $r \to y_m$. Hence, the error system of (2.9) becomes

$$\begin{bmatrix} \dot{\tilde{\theta}} \\ \dot{e} \end{bmatrix} = \begin{bmatrix} 0 & \varepsilon\phi \\ -\phi & -a_m \end{bmatrix} \begin{bmatrix} \tilde{\theta} \\ e \end{bmatrix} \quad (2.13a)$$

$$\phi = -y_m - e \quad (2.13b)$$

with initial conditions

$$\begin{aligned} \tilde{\theta}_0 &= \theta_0 - \theta_* \\ e_0 &= y_0 - y_{mo} \end{aligned} \quad (2.13c)$$

In terms of (2.9) we then have

$$\begin{aligned} \phi_*(t) &= -y_m(t) - e_0\exp(-a_m t) \\ e_*(t) &= e_0\exp(-a_m t) \\ H(s) &= \frac{1}{s + a_m} \\ G(s) &= -\frac{1}{s + a_m} \end{aligned} \quad (2.14)$$

The stability properties of (2.13) can be inferred from the Lyapunov function

$$V = \frac{1}{\varepsilon}\tilde{\theta}^2 - e^2 \quad (2.15)$$

Differentiating along (2.13) gives

$$\dot{V} = -2a_m e^2 \quad (2.16)$$

Hence, we immediately establish that

$$e = -(\phi + y_m) \in L_2 \cap L_\infty \quad (2.17)$$

and from LaSalle's Theorem [LaSalle and Lefchetz (1961)] it follows that $(e, \tilde{\theta}) = (0,0)$ is globally asymptotically stable (in the sense of Lyapunov). Since there is no further information to be extracted from the Lyapunov analysis, the transient properties of (2.13) can now be obtained by considering (2.13a) above, as if $\phi(t)$ is some exogenous function satisfying (2.17) and

$$\phi(t) + y_m(t) \to 0 \quad \text{as} \quad t \to \infty \quad (2.18)$$

The system of (2.13a), however, is a time-varying, linear system of a special form. Using the results in Anderson (1977) or Chapter II of Anderson et al. (1986), if $\phi(t)$ is *persistently exciting* (PE), i.e., there are positive constants $T$ and $\alpha$ such that

$$\frac{1}{T}\int_s^{s+T} \phi^2(t)dt \geq \alpha \quad , \quad \forall s \geq 0 \quad (2.19)$$

then $(e, \tilde{\theta}) \to (0,0)$ exponentially fast. Given the properties of $\phi(t)$ in (2.17), (2.18), if $r(t) = r_0$, a non-zero constant, then it can be established, with a little work, that $\phi(t) \in$ PE.

The difficulty with the above results is that they only provide *qualitative* information, but no fine detail. In general, the signal bounds and rate of convergence information is too coarse to provide any transient information. For example, if in our example system (2.10) we have the initial conditions

$$\begin{aligned} y(0) &= y_m(0) = 0 \\ \theta(0) &= \theta_0 \end{aligned} \quad (2.20)$$

then (2.15), (2.16) gives

$$\begin{aligned} \|e\|_2 &\le (2\varepsilon a_m)^{1/2}|\theta_0 - \theta_*| \\ \|e\|_\infty &\le (\varepsilon)^{1/2}|\theta_0 - \theta_*| \qquad (2.21) \\ \|\theta - \theta_*\|_\infty &\le |\theta_0 - \theta_*| \end{aligned}$$

Recall that $\theta_* = a_p$ in this case (2.12).

Rate of convergence estimates are also coarse. If $r(t) = r_0$, a constant, then it follows from (2.17) and (2.18) that for sufficiently large time $t$, $(e, \tilde\theta)$ are approximated by solutions of

$$\begin{bmatrix} \dot{\tilde\theta} \\ \dot e \end{bmatrix} = \begin{bmatrix} 0 & -\varepsilon r_0 \\ r_0 & -a_m \end{bmatrix} \begin{bmatrix} \tilde\theta \\ e \end{bmatrix} \qquad (2.22)$$

which in this case is an LTI system. Hence, the *asymptotic* rate of exponential convergence of $(e, \tilde\theta) \to (0,0)$ is[*]

$$\lambda = \sup\{ -\mathrm{Re}\, s : \sigma^2 + a_m s + a_m s + \varepsilon r_0^2 = 0 \} \qquad (2.23)$$

For small $\varepsilon > 0, \lambda \approx \frac{1}{2} r_0^2 \varepsilon$ whereas for large $\varepsilon > 0$, $\lambda \approx a_m/2$. Because the bounds in (2.23) are inversely proportional to $\varepsilon^{1/2}$, one suspects that (2.25) is better for large $\varepsilon$.

### The Multi-Parameter Case

When there is more than one adaptive parameter, i.e., when $\theta(t) \in \mathbb{R}^p$, $p > 1$, similar results are obtained provided, as mentioned before, $H(s)$ is strictly-positive-real (SPR), e.g., $H(s)$ is stable and there is a constant $\rho > 0$ such that[†]

$$\mathrm{Re} H(j\omega) \ge \rho |H(j\omega)|^2 \quad , \quad \forall\, \omega \in \mathbb{R} \qquad (2.24)$$

This is certainly the case in our example where $H(s) = 1/(s + a_m)$. The PE condition (2.19) in the vector case becomes,

$$\frac{1}{T} \sum_{s}^{s+T} \phi(t)\phi^T(t)\theta t \ge \alpha I_p \quad , \quad \forall\, s \ge 0 \qquad (2.25)$$

where now $\phi(t) \in \mathbb{R}^p$. The SPR condition provides a Lyapunov function, see, e.g., Chapter II of Anderson et al. (1986). Hence, The Lyapunov analysis leading to (2.17), (2.18) now yields

$$\begin{aligned} \phi - \phi_* &\in L_2 \cap L_\infty \\ \phi(t) - \phi_*(t) &\to 0 \quad \text{as} \quad t \to \infty \end{aligned} \qquad (2.27)$$

Under these conditions, it can be shown that $\phi_*(t) \in \mathrm{PE} \to \phi(t) \in \mathrm{PE}$. Since $\phi_*(t)$ is the output of a linear time invariant system, vis-a-vis (2.9), it follows that $\phi_*(t) \in \mathrm{PE}$ if $r(t)$ is *sufficiently rich*, meaning roughly that it contains enough sinusoids, see, e.g., Boyd and Sastry (1986) or Anderson et al. (1986).

### Simulation Results

In Figures 3 and 4 we show simulations of (2.10) with parameter values

$$a_p = 1 \quad , \quad a_m = 2 \qquad (2.28a)$$

reference signal

$$r(t) = 1 \qquad (2.28b)$$

initial conditions

$$\begin{aligned} y(0) &= y_m(0) = 0 \\ \theta(0) &= 2 \end{aligned} \qquad (2.28c)$$

for varying values of adaptation gain

$$\varepsilon = .01 \,,\, .1 \,,\, 1 \,,\, 10 \qquad (2.28d)$$

In Table 1 we compare the upper bound on $\|e\|_\infty$ from (2.23) with the simulation results. With the values above we have $\|e\|_\infty \le 1/\sqrt{\varepsilon}$. Observe that the estimated bound is practically useless for small $\varepsilon$, although much improved for large $\varepsilon$. This tendency is the opposite of what is hoped for, since small $\varepsilon$ is the prudent choice, usually favoring some amount of caution. In fact, for small $\varepsilon$, we can produce more accurate more accurate estimates of the transient behavior. The analysis is based on the classical method of averaging which is discussed in the next section.

In Table 2 we compare the asymptotic rate of convergence from (2.24) with the simulation results.

## 3. AVERAGING ANALYSIS

When the adaptation gain is small, we can consider an averaging analysis of our example system. Following Anderson et al. (1986), we compare the error solution $e(t)$ of (2.13) with the "frozen" system state $\bar e(t, \tilde\theta_0)$, that is

$$\begin{aligned} \tilde\theta(t) &= \tilde\theta_0 \\ \bar e(t, \tilde\theta_0) &= \psi(t, \tilde\theta_0)e_0 + \left(\int_0^t \psi(t-\tau, \tilde\theta_0) y_m(\tau)d\tau\right)\tilde\theta_0 \qquad (3.1) \\ \psi(t, \tilde\theta_0) &= \exp\{ -(a_m - \tilde\theta_0)t \} \end{aligned}$$

For $\tilde\theta_0 < a$ the solution is bounded on $t > 0$ and as $t \to \infty$,

$$\bar e(t, \tilde\theta_0) \to \frac{r_0}{a_m - \tilde\theta_0} \qquad (3.2)$$

exponentially fast. Obviously, (3.1) can only be a good approximation on short intervals. The long term parameter behavior, however, is governed by the "averaged system:"

$$\dot{\tilde\theta}_a = -\varepsilon f(\tilde\theta_a) \qquad (3.3a)$$

where

$$\begin{aligned} f(\tilde\theta_a) &= \lim_{T \to \infty} \frac{1}{T} \int_0^T [y_m(t) + \bar e(t, \tilde\theta_a)]\bar e(t, \tilde\theta_a)dt \\ &= a_m r_0^2 \tilde\theta_a/(a_m - \tilde\theta_a)^2 \qquad (3.3b) \end{aligned}$$

It is easily verified that $\tilde\theta_a = 0$ is a globally asymptotically stable solution of (3.3), and hence, for large $t$ (and small $\varepsilon$), the error is approximately

$$e_a = r_0 \tilde\theta_a/(a_m - \tilde\theta_a) \qquad (3.4)$$

The system (3.3)-(3.4) describes the trajectories in the "slow-manifold" as explained by Riedle and Kokotovic (1986).

From (3.3) the asymptotic rate of convergence is

$$\lambda_a = \varepsilon \frac{\partial f}{\partial \tilde\theta_a}\Big|_{\tilde\theta_a = 0} = \varepsilon r_0^2/a_m \qquad (3.5)$$

For the values as given in (2.26), $\lambda_a/\varepsilon = .5$, which compares favorably with the results in Table 2. In Figure 5 we show $\theta(t)$ and $\theta_{avg}(t) = \tilde\theta_a(t) + \theta_*$ for $\varepsilon = .01, .1$, and 1, starting with $\theta(0) = \theta_{avg}(0) = 2$. These are the same initial conditions used in the simulation results shown in Figures 3 and 4. Observe that $\theta_{avg}(t)$ compares

[*] The exponential rate of convergence to zero of a function $x(t)$ is defined as the number

$$\lambda = \sup\{\sigma > 0 : \exp(\sigma t)x(t) \to 0 \,,\, \text{as}\ t \to \infty\}$$

[†] The inequality is sufficient for SPR but not necessary.

favorably with $\theta(t)$ even for $\varepsilon = 1$. However, in Figure 6, the comparison is not good for $\varepsilon = 10$. One can expect this for large $\varepsilon$, but the averaging *theory* produces conservative estimates, e.g., $\varepsilon_{max} \cong .01$ [see Chapter III, Anderson et al. (1986)]. Clearly, in this example, the averaged system parameter trajectories provide a valid approximation for a reasonable range of adaptation gains. Many other example systems show the same validity, and apparently breaks down for excessive values of adaptation gains, that is, values which would never be chosen in practical situations. Other limitations on gain size, which we are not addressing here, have to do with disturbances, noise, and round-off error.

To fill out the character of our example system, we simulated (2.10) with $\varepsilon = .1$ for a variety of initial conditions, in particular, all combinations of $e_0 = -2, 0, 2$ and $\tilde{\theta}_0 = -2, -1, 0, 1, 2$. The results are shown in the phase-plane diagram of Figure 7. We can now gather the following "picture" of the behavior of (2.10):

**A. "Stable" Transient:** If $\tilde{\theta}_0 - a_m \ll \varepsilon$ then there is a fast, well behaved transient in the "boundary layer" (3.1), followed by a slow transient governed by the "averaged" systems (3.3)-(3.4). The latter is characterized by an asymptotically exponentially slow convergence rate $\varepsilon/a_m$.

**B. "Unstable" Transient:** If $\tilde{\theta}_0 - a_m \gg \varepsilon$ then there is a fast explosive transient in the boundary layer (3.1), during which the error grows excessively (this is predicted by (3.1) as long as $\theta(t)$ changes slowly). Following this, there is a "learning period" during which $\tilde{\theta}(t) - a_m \to O(\varepsilon)$ and $\tilde{\theta}(t) - a_m$ changes sign, thus arriving back into situation A.

In this example, because the Lyapunov analysis provides for *global* asymptotic stability, there are, of course, no initial conditions which produce unbounded trajectories. This cannot be expected, though, when the SPR condition is violated, as would be the case in practical situations where high frequency dynamics are neglected. Moreover, non-vanishing disturbances can also produce undesirable behavior. In both of these non-ideal cases -- unmodeled dynamics and disturbances -- although the SPR based (passivity) theory no longer applies to insure boundedness, the averaging theory still applies, and produces stability in the sense of Lyapunov, i.e., a local theory. In our example we avoided discussing this problem, which although of great importance, nonetheless clouds the issue insofar as determining the transient characteristics. These are predominantly local phenomena and are *not* predicted by passivity theory.

In closing this section, we call attention one more time to the simulations which show (see Figure 5) that the parameter trajectories of the "averaged systems" are quite good approximations to the actual parameter trajectories for a wide range of adaptation gain, much more extensive then demanded by the theory, see, e.g., Anderson et al. (1986). This calls into question the *theoretical* restrictions imposed by the averaging analysis.

## 4. CONCLUDING DISCUSSION

In this investigation of the transient properties of adaptive control systems, we have demonstrated some of the difficulties associated with predicting transient behavior, even for an ideal adaptive system. The long-term behavior, particularly under slow adaptation, can be predicted with good accuracy from the method of averaging. Although the averaging theory produces rather small estimates of allowable adaptation gains, our simulations here, as well as elsewhere, seem to indicate that averaging analysis is justified for much larger gains. Where averaging fails is for very large values of the initial parameter error, particularly at settings which are not in the constant parameter stability set. Thus, there are two limitations with the averaging analysis in regard to the transient: (1) conservative adaptation gain estimates, and (2) conservative rate of convergence estimates far away from convergence.

In order to transcend these limitations, it is probably necessary to go beyond averaging or else use averaging in a different way. To provide a focus for our speculations, consider the error system of (2.9), i.e.,

$$\begin{aligned} \dot{\tilde{\theta}} &= \varepsilon\phi e \quad , \quad \tilde{\theta}(0) = \tilde{\theta}_0 \\ e &= e_* - H(\phi^T\tilde{\theta}) \\ \phi &= \phi_* - G(\phi^T\tilde{\theta}) \end{aligned} \tag{4.1}$$

This is not the most general error system from which we can extract Figure 1, but we use it for illustrative purposes. Assume temporarily that $\phi(t)$ is a *given* exogenous functions. Then,

$$\dot{\tilde{\theta}} = \varepsilon\phi\left[e_* - H(\phi^T\tilde{\theta})\right] \tag{4.2}$$

is a *linear* time-varying system. There is no known general method of analysis except when $\phi(t)$ is periodic. Then one can utilize Floquet Theory. For example, if $\phi_*(t)$ is $T$-periodic, condition (2.27) holds, and $e_*(t) \to 0$ exponentially, then $\tilde{\theta}(t) \to 0$ exponentially *if and only if*

$$|\lambda\{\psi(T, 0)\}| < 1 \tag{4.3}$$

where $\psi(t, \tau)$ is the transition matrix associated with (4.7). Obviously, (4.8) can be significantly less conservative then (4.4), both in regard to the spectral constant of $\phi_*(t)$ and the limitation on $\varepsilon$, the adaptation gain. The proviso, of course, is whether (2.27) or a similar condition, holds. Observe that if $H$ is SPR and $\phi_*$ is PE, then (4.3) will always hold. If $H$ is not SPR then $\phi_*$ being PE is only necessary. Results of Floquet analysis of (4.2) can be found in Mareels et al. (1986).

Another way to approach the transient analysis is to utilize the method of averaging over a finite time interval. Since the transient occurs in the neighborhood of $\theta_0$, it is more sensible to generate the error system using $\theta_0$ rather than $\theta_*$, particularly if $\theta_0 - \theta_*$ is large. Thus, for sufficiently small time intervals -- not necessary small $\varepsilon$ -- the parameter transient behaves like the solution to the linear system

$$\frac{d}{dt}(\theta - \theta_0) = \varepsilon[\bar{\phi}(t, \theta_0)\bar{e}(t, \theta_0) - \bar{R}(t, \theta_0)(\theta - \theta_0)] \tag{4.4}$$

where $\bar{\phi}(t, \theta_0)$ and $\bar{e}(t, \theta_0)$ are the system regressor and error, respectively, with $\theta(t)$ "frozen" at $\theta_0$, and $\bar{R}(t, \theta_0)$ is given by

$$\bar{R}(t, \theta_0) = \bar{\phi}(H\,\bar{\phi})^T + \bar{e}(G\,\bar{\phi})^T \tag{4.5}$$

In (4.5), $H$ and $G$ have their usual definitions, vis a vis (2.9), except now are dependent on $\theta_0$ rather than $\theta_*$. Averaging analysis proceeds from here in the usual way.

A more general approach to the analysis of (4.1) can be formulated by appealing to the Contraction Mapping Principal(CMP) of Banach and Cacciopoli. We need the following definitions and statement of the CMP from Hale(1969).

If $\mathcal{M}$ is a subset of a Banach space $B$ with norm $\|\cdot\|$, and $\Gamma$ is an operator mapping $\mathcal{M} \to B$, then $\Gamma$ is a *contraction on* $\mathcal{M}$ if there is a constant $\sigma \in [0,1)$ such that

$$\|\Gamma x - \Gamma y\| \leq \sigma \|x - y\|, \forall x, y \in \mathcal{M} \qquad (4.6)$$

The constant $\sigma$ is referred to as the *contraction constant for* $\Gamma$ *on* $\mathcal{M}$. A *fixed point* of $\Gamma$ in $\mathcal{M}$ is a point ( function) $x \in \mathcal{M}$ such that $x = \Gamma x$. We now state the following result.

**Contraction Mapping Principal:** *If $\mathcal{M}$ is a closed subset of a Banach space $B$, and $\Gamma : \mathcal{M} \to \mathcal{M}$ is a contraction on $\mathcal{M}$, then $\Gamma$ has a unique fixed point in $\mathcal{M}$.*

In order to apply the CMP to the adaptive system (4.1) we need to identify the operator $\Gamma$ and the space $\mathcal{M}$. For example, let $\Gamma$ be the operator mapping $\bar{\theta} \mapsto \tilde{\theta}$ defined implicitly by

$$\phi = \phi_* - G(\bar{\theta}^T \phi) \qquad (4.7a)$$

$$\dot{\tilde{\theta}} = \epsilon \phi [e_* - H(\phi^T \tilde{\theta})] \qquad (4.7b)$$

It is clear that fixed points of $\Gamma$ are solutions to the adaptive system, i.e., $\tilde{\theta} = \Gamma \tilde{\theta}$ is equivalent to (4.1) with $\phi$ implicitly defined. A convenient choice of the space $\mathcal{M}$ is

$$\mathcal{M} = \{\tilde{\theta} \in C[0,T] : \|\tilde{\theta}(t)\| \leq \delta \exp(-\lambda t)\} \qquad (4.8)$$

where $C[0,T]$ is the Banach space of continuous, bounded functions on $[0,T]$, and the positive constants $\delta, \lambda$ are chosen so as to satisfy the conditions on $\Gamma$ imposed by the CMP. This means, in the first place, that $\Gamma$ must map $\mathcal{M}$ into itself. The proceedure is to determine the properties of $\phi$ from (4.7a) for all $\bar{\theta} \in \mathcal{M}$. Then, with $\phi$ so classified, determine the characteristics of $\tilde{\theta}$ from (4.7b). Observe that (4.7b) is of the same form as (4.2), but now $\phi$ behaves truely as an exogenous input. Thus, the stability properties of (4.7b) can be established from either averaging or Floquet analysis, as appropriate. The next step is to establish that $\Gamma$ is contractive on $\mathcal{M}$, which imposes further restrictions. Some of these can be relaxed by appealing to less demanding fixed point theory, e.g., the Schauder fixed point theorem gives up uniqueness for some additional requirements on $\mathcal{M}$. A further discussion of fixed point based approaches is presented in more detail in Kosut and Bitmead(1986).

As a final remark we point out that although our analysis tools, such as averaging, involve straightforward calculations, it is clear that the level of complexity of a realistic adaptive system is well beyond hand calculation. Hence, an area for further work is in the development of software tools which can eliminate some of the tedious parts of the analysis.

## REFERENCES


Anderson, B.D.O., 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*, MIT Press, 1986.

Anderson, B.D.O., R.R. Bitmead, C.R. Johnson, Jr., and R.L. Kosut, (1984), "Stability theorems for the relaxation of strictly positive real condition in hyperstable adaptive schemes," *Proc. 23rd IEEE CDC*, Las Vegas, NV, 1286-1291.

Anderson, B.D.O., (1985), "Adaptive systems, lack of persistency of excitation and bursting phenomena," *Automatica*, 21(3):247-258.

Astrom, K.J., (1983), "Analysis of Rohrs counterexamples to adaptive control," *Proc. 22nd IEEE Conf. on Dec. and Contr.*, San Antonio, TX, 982.

Astrom, K.J., (1984), "Interactions between excitation and unmodeled dynamics in adaptive control," *Proc. 23rd IEEE CDC*, Las Vegas, NV, 1276-1281.

Bodson, M., S. Sastry, B.D.O. Anderson, I. Mareels, and R.R. Bitmead, (1986), "Nonlinear averaging theorems, and the determination of parameter convergence rates in adaptive control," *Systems and Contr. Letters*, to appear.

Boyd, S. and S. Sastry (1986), "Necessary and sufficient conditions for parameter convergence in adaptive control," *Automatica*, 22(6):629-640.

Goodwin, G.C., and K.S. Sin, (1984), *Adaptive Filtering, Prediction and Control*, Prentice Hall, New Jersey.

Hale, J.K., (1980), *Ordinary Differential Equations*, Kreiger, Molaban, FL, originally published (1969), Wiley (Interscience), New York.

Kosut, R.L., and C.R. Johnson, Jr., (1984), "An input-output view of robustness in adaptive control," *Automatica*, 20(5):569-581.

Kosut, R.L., and B. Friedlander, (1985), "Robust adaptive control: conditions for global stability," *IEEE Trans. on Aut. Contr.*, AC-30(7):610-624.

Kosut, R.L. and R.R. Bitmead, "Fixed-point theorems for the analysis of adaptive systems", *Proc. 2nd IFAC Workshop on Adaptive Systems*, Lund, Sweden, July 1986.

Mareels, I.M.Y., B.D.O. Anderson, R.R. Bitmead, M. Bodson, and S.S. Sastry, (1986), "Revisiting the MIT-Rule for adaptive systems," *Proc. 2nd IFAC Workshop on Adaptive Systems*, Lund, Sweden, pp. 67-71.

Riedle, B.D., and P.V. Kokotovic, (1986), "Integral manifold approach to slow adaptation," *IEEE Trans. on Aut. Contr.*, AC-31(4):316-325.


*Table 1*

Peak error: $\|e\|_\infty$

| $\epsilon$ | simulation | from(2.21) |
|---|---|---|
| .01 | .90 | 10 |
| .1 | .72 | 3.16 |
| 1 | .43 | 1 |
| 10 | .2 | .316 |

*Table 2*

Asymptotic rate of parameter convergence

| $\epsilon$ | simulation | from(2.23) |
|---|---|---|
| .01 | .013 | .005 |
| .1 | .069 | .051 |
| 1 | 1.17 | 1 |
| 10 | 1.53 | 1 |

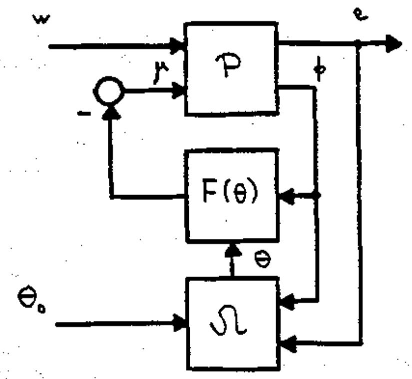


Fig. 1. - Adaptive system

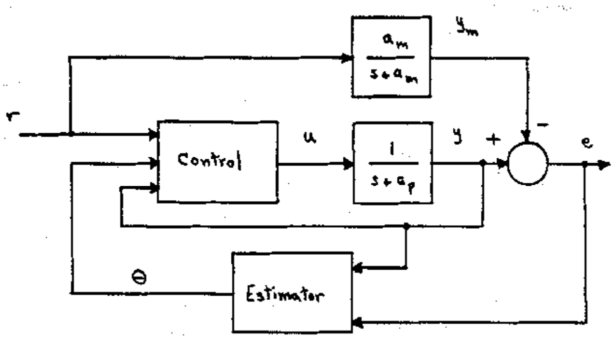


Fig. 2 - Example system

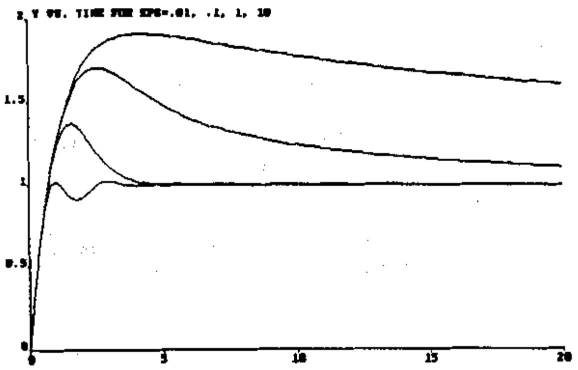


Fig. 3 $y(t)$ for $\varepsilon = .01, .1, 1, 10$

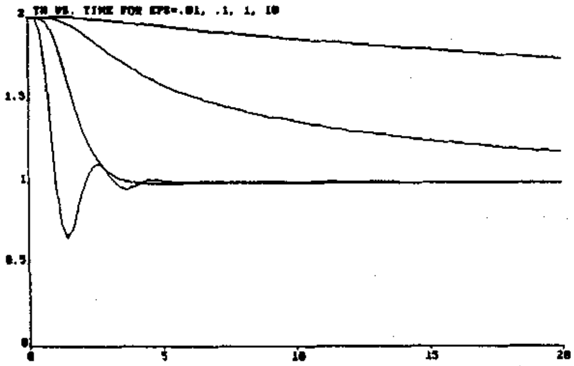


Fig. 4 - $\theta(t)$ for $\varepsilon = .01, .1, 1, 10$

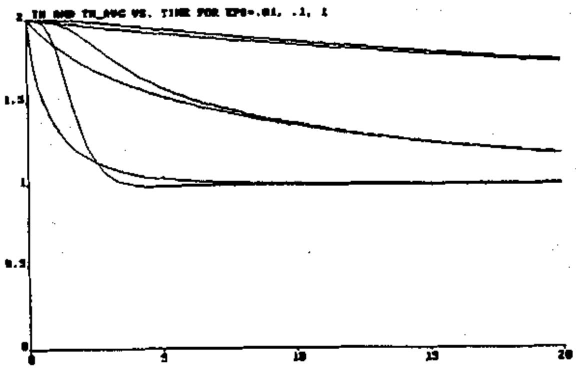


Fig 5. - $\theta(t)$ and $\bar{\theta}_a(t)$ for $\varepsilon = .01, .1, 1$

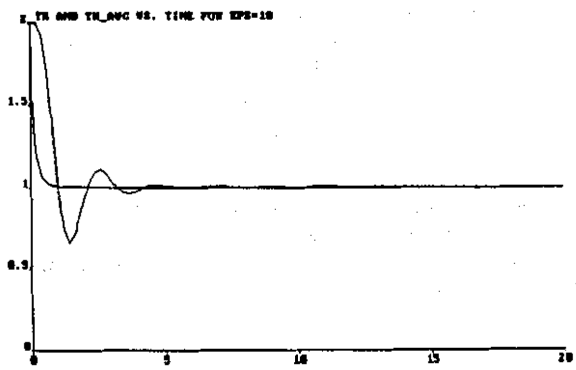


Fig. 6 - $\theta(t)$ and $\theta_a(t)$ for $\varepsilon = 10$

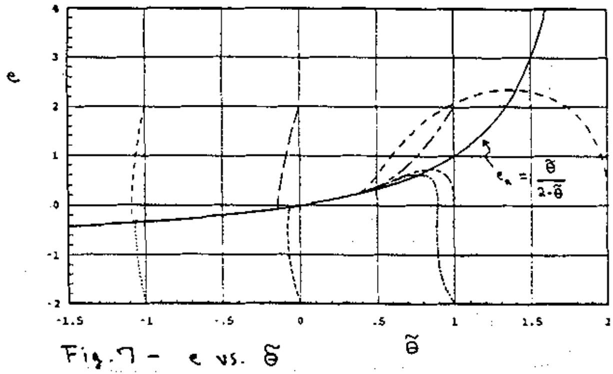


Fig. 7 - e vs. $\tilde{\theta}$