CHAPTER 4

PARAMETER CONVERGENCE USING AVERAGING TECHNIQUES

4.0 INTRODUCTION

Averaging is a method of analysis of differential equations of the form

$$\dot{x} = \epsilon f(t, x) \tag{4.0.1}$$

and relates properties of the solutions of system (4.0.1) to properties of the solutions of the so-called *averaged* system

$$\dot{x}_{av} = \epsilon f_{av}(x_{av}) \tag{4.0.2}$$

where

$$f_{av}(x) := \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} f(\tau, x)\, d\tau \tag{4.0.3}$$

assuming that the limit exists and that the parameter ϵ is sufficiently small. The method was proposed originally by Bogoliuboff & Mitropolskii [1961], developed subsequently by Volosov [1962], Sethna [1973], Balachandra & Sethna [1975] and Hale [1980]; and stated in a geometric form in Arnold [1982] and Guckenheimer & Holmes [1983].

Averaging methods were introduced for the stability analysis of deterministic adaptive systems in the work of Astrom [1983], Astrom [1984], Riedle & Kokotovic [1985] and [1986], Mareels *et al* [1986], and Anderson *et al* [1986]. We also find early informal use of averaging in Astrom & Wittenmark [1973], and, in a stochastic context, in Ljung &

Soderstrom [1983] (the ODE approach).

Averaging is very valuable to assess the stability of adaptive systems in the presence of unmodeled dynamics and to understand mechanisms of instability. However, it is not only useful in stability problems, but in general as an *approximation* method, allowing one to replace a system of *nonautonomous* (time varying) differential equations by an *autonomous* (time invariant) system. This aspect was emphasized in Fu, Bodson, & Sastry [1986], Bodson *et al* [1986], and theorems were derived for one-time scale and two-time scale systems such as those arising in identification and control. These results are reviewed here, together with their application to the adaptive systems described in previous chapters. Our recommendation to the reader not familiar with these results is to derive the simpler versions of the theorems for linear periodic systems. In the following section, we present examples of averaging analysis which will help to understand the motivation of the methods discussed in this chapter.

4.1 EXAMPLES OF AVERAGING ANALYSIS

One-Time Scale Averaging

Consider the *linear* nonautonomous differential equation

$$\dot{x} = -\epsilon \sin^2(t)\, x \qquad x(0) = x_0 \tag{4.1.1}$$

where x is a *scalar*. This equation is a special case of the parameter error equation encountered in Chapter 2

$$\dot{\phi} = -g\, w(t)\, w^T(t)\, \phi \qquad \phi(0) = \phi_0 \tag{4.1.2}$$

and corresponds to the identification of a single constant θ^* from measurements of

$$y_p(t) = \theta^* \sin(t) \tag{4.1.3}$$

using a gradient update law. The general solution of a first order linear differential equation of the form

$$\dot{x} = a(t)\, x \qquad x(0) = x_0 \tag{4.1.4}$$

is known analytically, and is given by

$$x(t) = e^{\int_0^t a(\tau)\, d\tau}\, x_0 \tag{4.1.5}$$

In particular, the solution of (4.1.1) is simply

$$x(t) = e^{-\epsilon\int_0^t \sin^2(\tau)\,d\tau} x_0 = e^{-\epsilon\int_0^t (\frac{1}{2} - \frac{1}{2}\cos(2\tau))d\tau} x_0$$

$$= e^{-\epsilon\frac{t}{2} + \frac{\epsilon}{4}\sin(2t)} x_0 \tag{4.1.6}$$

Note that when we replaced $\sin^2(\tau)$ by $\frac{1}{2} - \frac{1}{2}\cos(2\tau)$ in (4.1.6), we separated the integrand into its average and periodic part. Indeed, for all t_0, x

$$\lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} \sin^2(\tau)x\,d\tau = \frac{1}{2}x \tag{4.1.7}$$

Therefore, the averaged system defined by (4.0.2)–(4.0.3) is now given by

$$\dot{x}_{av} = -\frac{\epsilon}{2}x_{av} \qquad x_{av}(0) = x_0 \tag{4.1.8}$$

The solution of the averaged system is

$$x_{av}(t) = e^{-\epsilon\frac{t}{2}} x_0 \tag{4.1.9}$$

Let us now compare the solutions of the original system (4.1.6) and of the averaged system (4.1.9). The difference between the solutions, at a fixed t

$$|x(t) - x_{av}(t)| = e^{-\epsilon\frac{t}{2}}\left|e^{\frac{\epsilon}{4}\sin(2t)} - 1\right| \tag{4.1.10}$$

$$\to \left|\frac{\epsilon}{4}\sin(2t)\right| \qquad \text{as } \epsilon \to 0 \tag{4.1.11}$$

In other words, the solutions are arbitrarily close as $\epsilon \to 0$, so that we may *approximate* the original system by the averaged system. Also, both systems are *exponentially stable* (and if we were to change the sign in the differential equation, both would be unstable). As is now shown, the convergence rates are also identical.

Recall that the convergence rate of an exponentially stable system is the constant α such that the solutions satisfy

$$|x(t)| \le m\,e^{-\alpha(t-t_0)}\,|x(t_0)| \tag{4.1.12}$$

for all $x(t_0)$, $t_0 \ge 0$. A graphical representation may be obtained by plotting $\ln(|x(t)|^2)$, noting that

$$\ln(|x(t)|^2) \le \ln(m^2|x(t_0)|^2) - 2\alpha(t - t_0) \tag{4.1.13}$$

Therefore, the graph of $\ln(|x(t)|^2)$ is bounded by a straight line of slope -2α. In the above example, the original and the averaged system have identical convergence rate $\alpha = \frac{\epsilon}{2}$.

In this chapter, we will prove theorems stating similar results for more general systems. Then, the analytic solution of the original system is not available, and averaging becomes useful. The method of proof is completely different, but the results are essentially the same: closeness of the solutions, and closeness of the convergence rates as $\epsilon \to 0$. We devote the rest of this section to show how the averaged system may be calculated in more complex cases, using frequency-domain expressions.

One-Time Sale Averaging—Multiple Frequencies

Consider the system

$$\dot{x} = -\epsilon\,w^2(t)\,x \tag{4.1.14}$$

where $x \in \mathbb{R}$, and w contains multiple frequencies

$$w(t) = \sum_{k=1}^{n} a_k \sin(\omega_k t + \phi_k) \tag{4.1.15}$$

To define the averaged system, we need to calculate the average

$$\text{AVG}(w^2(t)) := \lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T} w^2(\tau)\,d\tau \tag{4.1.16}$$

Expanding w^2 will give us a sum of product of sin's at the frequencies ω_k. However, a product of two sinusoids at different frequencies has zero average, so that

$$\text{AVG}(w^2(t)) = \sum_{k=1}^{n}\frac{a_k^2}{2} \tag{4.1.17}$$

and the averaged system is

$$\dot{x}_{av} = -\epsilon\left[\sum_{k=1}^{n}\frac{a_k^2}{2}\right]x_{av} \tag{4.1.18}$$

The averaged system is exponentially stable as soon as w contains at least one sinusoid. Note also that the expression (4.1.18) is independent of the phases ϕ_k.

Two-Time Scale Averaging

Averaging may also be applied to systems of the form

$$\dot{x}(t) = -\epsilon\, w(t)\, y(t) \tag{4.1.19}$$

$$\dot{y}(t) = -a\, y(t) + b\, w(t)\, x(t) \tag{4.1.20}$$

where $x, y \in \mathbb{R}$, $a > 0$. In short, (4.1.20) is denoted

$$y = \frac{b}{s+a}(w x) = \hat{M}(w x) \tag{4.1.21}$$

where $\hat{M}$ is a stable transfer function. (4.1.19) becomes

$$\dot{x} = -\epsilon\, w\, \hat{M}(w x) \tag{4.1.22}$$

Equations (4.1.19)–(4.1.20) were encountered in model reference identification with x replaced by the parameter error ϕ, ϵ by the adaptation gain g, and y by the identifier error e_i.

When $\epsilon \to 0$, $x(t)$ varies slowly when compared to $y(t)$, and the time scales of their variations become separated. $x(t)$ is called the *slow state*, $y(t)$ the *fast state* and the system (4.1.19)–(4.1.20) a *two-time* scale system. In the limit as $\epsilon \to 0$, $x(t)$ may be considered frozen in (4.1.20) or (4.1.21), so that

$$\hat{M}(w x) = \hat{M}(w)\, x \tag{4.1.23}$$

The result of the averaging theory is that (4.1.22) may indeed be approximated by

$$\dot{x}_{av} = -\epsilon\, \mathrm{AVG}(w\, \hat{M}(w))\, x_{av} \tag{4.1.24}$$

Again, a frequency domain expression brings more interesting insight. Let w contain multiple sinusoids

$$w(t) = \sum_{k=1}^{n} a_k \sin(\omega_k t) \tag{4.1.25}$$

so that

$$\hat{M}(w) = \sum_{k=1}^{n} a_k\, |\hat{M}(j\omega_k)|\, \sin(\omega_k t + \phi_k) \tag{4.1.26}$$

where

$$\phi_k = \arg(\hat{M}(j\omega_k)) \tag{4.1.27}$$

The product $w\hat{M}(w)$ may be expanded as the sum of products of sinusoids. Further, $\sin(\omega_k t + \phi_k) = \sin(\omega_k t)\cos(\phi_k) + \cos(\omega_k t)$

$\sin(\phi_k)$. Now, products of sinusoids at different frequencies have zero average, as do products of sin's with cos's of any frequency. Therefore

$$\mathrm{AVG}\left[w\,\hat{M}(w)\right] = \sum_{k=1}^{n} \frac{a_k^2}{2}\, |\hat{M}(j\omega_k)|\, \cos(\phi_k)$$

$$= \sum_{k=1}^{n} \frac{a_k^2}{2}\, \mathrm{Re}\left[\hat{M}(j\omega_k)\right] \tag{4.1.28}$$

Using (4.1.28), a *sufficient* condition for the stability of the averaged system (4.1.24) is that

$$\mathrm{Re}\,\hat{M}(j\omega) > 0 \qquad \text{for all } \omega > 0 \tag{4.1.29}$$

The condition is the familiar SPR condition obtained for the stability of the original system in the context of model reference identification. The averaging analysis brings this condition in evidence *directly in the frequency domain*. It is also evident that this condition is *necessary*, if one does not restrict the frequency content of the signal $w(t)$. Otherwise, it is sufficient that the ω_k's be concentrated in frequencies where $\mathrm{Re}\,\hat{M}(j\omega) > 0$, so that the sum in (4.1.28) is positive.

Vector Case

In identification, we encountered (4.1.2), where ϕ was a vector. The solution (4.1.5) does not extend to the vector case, but the frequency domain analysis does, as will be shown in Section 4.3. We illustrate the procedure with the simple example of the identification of a first order system (cf. Section 2.0).

The regressor vector is given by

$$w = \begin{bmatrix} r \\ y_p \end{bmatrix} = \begin{bmatrix} r \\ \hat{P}(r) \end{bmatrix} \tag{4.1.30}$$

where $\hat{P} = k_p/(s + a_p)$. As before, we let the input r be periodic

$$r(t) = \sum_{k=1}^{n} r_k \sin(\omega_k t) \tag{4.1.31}$$

so that the averaged system is given by (the gain g plays the role of ϵ)

$$\dot{\phi}_{av} = -g\, \mathrm{AVG}(w w^T)\, \phi_{av}$$

$$= -g \begin{bmatrix} \mathrm{AVG}(r\, r) & \mathrm{AVG}(r\, \hat{P}(r)) \\ \mathrm{AVG}(r\, \hat{P}(r)) & \mathrm{AVG}(\hat{P}(r)\, \hat{P}(r)) \end{bmatrix} \phi_{av}$$

$$= -g \sum_{k=1}^{n} \frac{r_k^2}{2} \begin{bmatrix} 1 & \mathrm{Re}\,\hat{P}(j\omega_k) \\ \mathrm{Re}\,\hat{P}(j\omega_k) & |\hat{P}(j\omega_k)|^2 \end{bmatrix} \phi_{av} \tag{4.1.32}$$

$$= -g \sum_{k=1}^{n} \frac{r_k^2}{2} \begin{bmatrix} 1 & \dfrac{a_p k_p}{\omega_k^2 + a_p^2} \\ \dfrac{a_p k_p}{\omega_k^2 + a_p^2} & \dfrac{k_p^2}{\omega_k^2 + a_p^2} \end{bmatrix} \phi_{av} \tag{4.1.33}$$

The matrix above is symmetric and it may be checked to be positive semi-definite. Further, it is positive definite for all $\omega_k \neq 0$. Taking a Lyapunov function $v = \phi_{av}^T \phi_{av}$ shows that the averaged system is exponentially stable *as long as the input contains at least one sinusoid of frequency* $\omega \neq 0$. Thus, we directly recover a frequency-domain result obtained earlier for the original system through a much longer and laborious path.

Nonlinear Averaging

Analyzing *adaptive control* schemes using averaging is trickier because the schemes are usually nonlinear. This is the motivation for the derivation of nonlinear averaging theorems in this chapter. Note that it is possible to linearize the system around some nominal trajectory, or around the equilibrium. However, averaging allows us to approximate a nonautonomous system by an autonomous system, independently of the linearity or nonlinearity of the equations. Indeed, we will show that it is possible to keep the nonlinearity of the adaptive systems, and even obtain frequency domain results. The analysis is therefore not restricted to a neighborhood of some trajectory or equilibrium.

As an example, we consider the output error model reference adaptive control scheme for a first order system (cf. Section 3.0, with $k_p = k_m = c_0 = 1$)

$$\begin{aligned} \dot{y}_p &= -a_p y_p + u \\ \dot{y}_m &= -a_m y_m + r \end{aligned} \tag{4.1.34}$$

where $a_m > 0$, and a_p is unknown. The adaptive controller is defined by

$$\begin{aligned} u &= r + d_0 y_p \\ \dot{d}_0 &= -g\, e_0 y_p \end{aligned} \tag{4.1.35}$$

where $g > 0$ is the adaptation gain. The output error and the parameter error are given by

$$\begin{aligned} e_0 &= y_p - y_m \\ \phi &= d_0 - d_0^* = d_0 - (a_m - a_p) \end{aligned} \tag{4.1.36}$$

The adaptive system is completely described by

$$\begin{aligned} \dot{e}_0 &= -(a_m - \phi)\, e_0 + y_m \phi \\ \dot{\phi} &= -g\, e_0 (e_0 + y_m) \end{aligned} \tag{4.1.37}$$

where

$$y_m = \frac{1}{s + a_m}(r) \tag{4.1.38}$$

When g is small (g takes the place of ϵ in the averaging analysis), ϕ varies slowly compared to r, y_m and e_0. The averaged system is defined by calculating $\mathrm{AVG}(e_0(e_0 + y_m))$, assuming that ϕ is fixed. In that case

$$\begin{aligned} e_0 &= \frac{1}{s + a_m - \phi}(y_m)\ \phi \\ &= \frac{1}{s + a_m - \phi}\left[\frac{1}{s + a_m}(r)\right] \phi \end{aligned} \tag{4.1.39}$$

and

$$\begin{aligned} e_0 + y_m &= \frac{1}{s + a_m - \phi}(y_m)\ \phi + y_m \\ &= \frac{s + a_m}{s + a_m - \phi}(y_m) = \frac{1}{s + a_m - \phi}(r) \end{aligned} \tag{4.1.40}$$

Note that $s + a_m - \phi$ is the *closed-loop polynomial*, that is the polynomial giving the closed-loop pole for ϕ fixed. Assume again that r is of the form

$$r = \sum_{k=1}^{n} r_k \sin(\omega_k t) \tag{4.1.41}$$

and it follows that

$$\begin{aligned} &\mathrm{AVG}\left[e_0(e_0 + y_m)\right]_{\phi\ \mathrm{fixed}} \\ &= \sum_{k=1}^{n} \frac{r_k^2}{2} \frac{1}{\omega_k^2 + (a_m - \phi)^2} \frac{a_m}{\omega_k^2 + a_m^2}\ \phi \end{aligned} \tag{4.1.42}$$

so that the averaged system is given by

$$\dot{\phi}_{av} = -g \sum_{k=1}^{n} \frac{r_k^2}{2} \frac{1}{\omega_k^2 + (a_m - \phi_{av})^2} \frac{a_m}{\omega_k^2 + a_m^2} \phi_{av} \tag{4.1.43}$$

The averaged system is a scalar *nonlinear* system. Indeed, averaging did not alter the nonlinearity of the original system, only its time variation. Note that the averaged system is of the form

$$\dot{\phi}_{av} = -a(\phi_{av})\, \phi_{av} \tag{4.1.44}$$

where $a(\phi_{av})$ is a nonlinear function of ϕ_{av}. However, for all $h > 0$, there exists $\alpha > 0$ such that

$$a(\phi_{av}) \geq \alpha > 0 \qquad \text{for all } |\phi_{av}| \leq h \tag{4.1.45}$$

as long as r contains at least one sinusoid (including at $\omega = 0$). By taking a Lyapunov function $v = \phi_{av}^2$, it is easy to see that (4.1.43) is exponentially stable in B_h, with rate of convergence α. Since h is arbitrary, the system is not only locally exponentially stable, but also exponentially stable in any closed-ball. However, it is *not* globally exponentially stable, because α is not bounded below as $h \to \infty$.

Again, we recovered a result and a frequency domain analysis, obtained for the original system through a very different path. An advantage of the averaging analysis is to give us an expression (4.1.43) which may be used to predict parameter convergence quantitatively from frequency domain conditions.

The analysis of this section may be extended to the general identification and adaptive control schemes discussed in Chapter 2 and Chapter 3. We first present the averaging theory that supports the frequency-domain analysis.

4.2 AVERAGING THEORY—ONE-TIME SCALE

In this section, we consider differential equations of the form

$$\dot{x} = \epsilon f(t, x, \epsilon) \qquad x(0) = x_0 \tag{4.2.1}$$

where $x \in \mathbb{R}^n$, $t \geq 0$, $0 < \epsilon \leq \epsilon_0$, and f is piecewise continuous with respect to t. We will concentrate our attention on the behavior of the solutions in some closed ball B_h of radius h, centered at the origin.

For small ϵ, the variation of x with time is slow, as compared to the rate of time variation of f. The *method of averaging* relies on the assumption of the existence of the mean value of $f(t, x, 0)$ defined by the limit

$$f_{av}(x) = \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} f(\tau, x, 0)\, d\tau \tag{4.2.2}$$

assuming that the limit exists uniformly in t_0 and x. This is formulated more precisely in the following definition.

Definition Mean Value of a Function, Convergence Function

The function $f(t, x, 0)$ is said to have mean value $f_{av}(x)$ if there exists a continuous function $\gamma(T)$: $\mathbb{R}_+ \to \mathbb{R}_+$, strictly decreasing, such that $\gamma(T) \to 0$ as $T \to \infty$ and

$$\left| \frac{1}{T} \int_{t_0}^{t_0+T} f(\tau, x, 0)\, d\tau - f_{av}(x) \right| \leq \gamma(T) \tag{4.2.3}$$

for all $t_0 \geq 0$, $T \geq 0$, $x \in B_h$.

The function $\gamma(T)$ is called the *convergence function.*

Note that the function $f(t, x, 0)$ has mean value $f_{av}(x)$ if and only if the function

$$d(t, x) = f(t, x, 0) - f_{av}(x) \tag{4.2.4}$$

has zero mean value.

It is common, in the literature on averaging, to assume that the function $f(t, x, \epsilon)$ is periodic in t, or almost periodic in t. Then, the existence of the mean value is guaranteed, without further assumption (Hale [1980], theorem 6, p. 344). Here, we do not make the assumption of (almost) periodicity, but consider instead the assumption of the existence of the mean value as the starting point of our analysis.

Note that if the function $d(t, x)$ is periodic in t and is bounded, then the integral of the function $d(t, x)$ is also a bounded function of time. This is equivalent to saying that there exists a convergence function $\gamma(T) = a/T$ (i.e., of the order of $1/T$) such that (4.2.3) is satisfied. On the other hand, if the function $d(t, x)$ is bounded, and is not required to be periodic but almost periodic, then the integral of the function $d(t, x)$ need not be a bounded function of time, even if its mean value is zero (Hale [1980], p. 346). The function $\gamma(T)$ is bounded (by the same bound as $d(t, x)$) and converges to zero as $T \to \infty$, but the convergence function need not be bounded by a/T as $T \to \infty$ (it may be of order $1/\sqrt{T}$ for example). In general, a zero mean function need not have a bounded integral, although the converse is true. In this book, we do not make the distinction between the periodic and the almost periodic case, but we do distinguish the bounded integral case from the general case and indicate the importance of the function $\gamma(T)$ in the subsequent developments.

System (4.2.1) will be called the *original system* and, assuming the existence of the mean value for the original system, the *averaged system* is defined to be

$$\dot{x}_{av} = \epsilon f_{av}(x_{av}) \qquad x_{av}(0) = x_0 \tag{4.2.5}$$

Note that the averaged system is autonomous and, for T fixed and ϵ varying, the solutions over intervals $[0, T/\epsilon]$ are identical, modulo a simple time scaling by ϵ.

We address the following two questions:

(a) the closeness of the response of the original and averaged systems on intervals $[0, T/\epsilon]$,

(b) the relationships between the stability properties of the two systems.

To compare the solutions of the original and of the averaged system, it is convenient to transform the original system in such a way that it becomes a *perturbed* version of the averaged system. An important lemma that leads to this result is attributed to Bogoliuboff & Mitropolskii [1961], p. 450 and Hale [1980], lemma 4, p. 346. We state a generalized version of this lemma.

Lemma 4.2.1 Approximate Integral of a Zero Mean Function

If $d(t,x): \mathbb{R}_+ \times B_h \to \mathbb{R}^n$ is a bounded function, piecewise continuous with respect to t, and has zero mean value with convergence function $\gamma(T)$

Then there exists $\xi(\epsilon) \in K$ and a function $w_\epsilon(t,x): \mathbb{R}_+ \times B_h \to \mathbb{R}^n$ such that

$$|\epsilon w_\epsilon(t,x)| \le \xi(\epsilon) \tag{4.2.6}$$

$$\left| \frac{\partial w_\epsilon(t,x)}{\partial t} - d(t,x) \right| \le \xi(\epsilon) \tag{4.2.7}$$

for all $t \ge 0$, $x \in B_h$. Moreover, $w_\epsilon(0,x) = 0$, for all $x \in B_h$.

If, moreover, $\gamma(T) = a/T^r$ for some $a \ge 0$, $r \in (0,1]$

Then the function $\xi(\epsilon)$ can be chosen to be $2a\epsilon^r$.

Proof of Lemma 4.2.1 in Appendix.

Comments

The construction of the function $w_\epsilon(t,x)$ in the proof is identical to that in Bogoliuboff & Mitropolskii [1961], but the proof of (4.2.6), (4.2.7) is different and leads to the relationship between the convergence function $\gamma(T)$ and the function $\xi(\epsilon)$.

The main point of lemma 4.2.1 is that, although the exact integral of $d(t,x)$ may be an unbounded function of time, there exists a bounded function $w_\epsilon(t,x)$, whose first partial derivative with respect to t is arbitrarily close to $d(t,x)$. Although the bound on $w_\epsilon(t,x)$ may increase as $\epsilon \to 0$, it increases slower than $\xi(\epsilon)/\epsilon$, as indicated by (4.2.6).

It is necessary to obtain a function $w_\epsilon(t,x)$, as in lemma 4.2.1, that has some additional smoothness properties. A useful lemma is given by Hale ([1980], lemma 5, p. 349). At the price of additional assumptions on the function $d(t,x)$, the following lemma leads to stronger conclusions that will be useful in the sequel.

Lemma 4.2.2 Smooth Approximate Integral of a Zero Mean Function

If $d(t,x): \mathbb{R}_+ \times B_h \to \mathbb{R}^n$ is piecewise continuous with respect to t, has bounded and continuous first partial derivatives with respect to x and $d(t,0) = 0$, for all $t \ge 0$. Moreover, $d(t,x)$ has zero mean value, with convergence function $\gamma(T)|x|$ and $\dfrac{\partial d(t,x)}{\partial x}$ has zero mean value, with convergence function $\gamma(T)$

Then there exists $\xi(\epsilon) \in K$ and a function $w_\epsilon(t,x): \mathbb{R}_+ \times B_h \to \mathbb{R}^n$, such that

$$|\epsilon w_\epsilon(t,x)| \le \xi(\epsilon)|x| \tag{4.2.8}$$

$$\left| \frac{\partial w_\epsilon(t,x)}{\partial t} - d(t,x) \right| \le \xi(\epsilon)|x| \tag{4.2.9}$$

$$\left| \epsilon \frac{\partial w_\epsilon(t,x)}{\partial x} \right| \le \xi(\epsilon) \tag{4.2.10}$$

for all $t \ge 0$, $x \in B_h$. Moreover, $w_\epsilon(0,x) = 0$, for all $x \in B_h$.

If, moreover, $\gamma(T) = a/T^r$ for some $a \ge 0$, $r \in (0,1]$,

Then the function $\xi(\epsilon)$ can be chosen to be $2a\epsilon^r$.

Proof of Lemma 4.2.2 in Appendix.

Comments

The difference between this lemma and lemma 4.2.1 is in the condition on the partial derivative of $w_\epsilon(t, x)$ with respect to x in (4.2.10) and the dependence on $|x|$ in (4.2.8), (4.2.9).

Note that if the original system is linear, i.e.

$$\dot{x} = \epsilon A(t)x \qquad x(0) = x_0 \tag{4.2.11}$$

for some $A(t): \mathbb{R}_+ \to \mathbb{R}^{n \times n}$, then the main assumption of lemma 4.2.2 is that there exists A_{av} such that $A(t) - A_{av}$ has zero mean value.

The following assumptions will now be in effect.

Assumptions

For some $h > 0$, $\epsilon_0 > 0$

(A1) $x = 0$ is an equilibrium point of system (4.2.1), that is, $f(t, 0, 0) = 0$ for all $t \geq 0$. $f(t, x, \epsilon)$ is Lipschitz in x, that is, for some $l_1 \geq 0$

$$|f(t, x_1, \epsilon) - f(t, x_2, \epsilon)| \leq l_1 |x_1 - x_2| \tag{4.2.12}$$

for all $t \geq 0$, $x_1, x_2 \in B_h$, $\epsilon \leq \epsilon_0$.

(A2) $f(t, x, \epsilon)$ is Lipschitz in ϵ, linearly in x, that is, for some $l_2 \geq 0$

$$|f(t, x, \epsilon_1) - f(t, x, \epsilon_2)| \leq l_2 |x| \, |\epsilon_1 - \epsilon_2| \tag{4.2.13}$$

for all $t \geq 0$, $x \in B_h$, $\epsilon_1, \epsilon_2 \leq \epsilon_0$.

(A3) $f_{av}(0) = 0$ and $f_{av}(x)$ is Lipschitz in x, that is, for some $l_{av} \geq 0$

$$|f_{av}(x_1) - f_{av}(x_2)| \leq l_{av} |x_1 - x_2| \tag{4.2.14}$$

for all $x_1, x_2 \in B_h$.

(A4) the function $d(t, x) = f(t, x, 0) - f_{av}(x)$ satisfies the conditions of lemma 4.2.2.

Lemma 4.2.3 Perturbation Formulation of Averaging

If the original system (4.2.1) and the averaged system (4.2.5) satisfy assumptions (A1)–(A4)

Then there exist functions $w_\epsilon(t, x)$, $\xi(\epsilon)$ as in lemma 4.2.2 and $\epsilon_1 > 0$ such that the transformation

$$x = z + \epsilon w_\epsilon(t, z) \tag{4.2.15}$$

is a homeomorphism in B_h for all $\epsilon \leq \epsilon_1$ and

$$|x - z| \leq \xi(\epsilon) |z| \tag{4.2.16}$$

Under the transformation, system (4.2.1) becomes

$$\dot{z} = \epsilon f_{av}(z) + \epsilon p(t, z, \epsilon) \qquad z(0) = x_0 \tag{4.2.17}$$

where $p(t, z, \epsilon)$ satisfies

$$|p(t, z, \epsilon)| \leq \psi(\epsilon) |z| \tag{4.2.18}$$

for some $\psi(\epsilon) \in K$. Further, $\psi(\epsilon)$ is of the order of $\epsilon + \xi(\epsilon)$.

Proof of Lemma 4.2.3 in Appendix.

Comments

a) A similar lemma can be found in Hale [1980] (lemma 3.2, p. 192). Inequality (4.2.18) is a Lipschitz type of condition on $p(t, z, \epsilon)$, which is not found in Hale [1980] and results from the stronger conditions and conclusions of lemma 4.2.2.

b) Lemma 4.2.3 is fundamental to the theory of averaging presented hereafter. It separates the error in the approximation of the original system by the averaged system $(x - x_{av})$ into two components: $x - z$ and $z - x_{av}$. The first component results from a pointwise (in time) transformation of variable. This component is guaranteed to be small by inequality (4.2.16). For ϵ sufficiently small ($\epsilon \leq \epsilon_1$), the transformation $z \to x$ is invertible and as $\epsilon \to 0$, it tends to the identity transformation. The second component is due to the perturbation term $p(t, z, \epsilon)$. Inequality (4.2.18) guarantees that this perturbation is small as $\epsilon \to 0$.

c) At this point, we can relate the convergence of the function $\gamma(T)$ to the order of the two components of the error $x - x_{av}$ in the approximation of the original system by the averaged system. The relationship between the functions $\gamma(T)$ and $\xi(\epsilon)$ was indicated in lemma 4.2.1. Lemma 4.2.3 relates the function $\xi(\epsilon)$ to the error due to the averaging. If $d(t, x)$ has a bounded integral (i.e., $\gamma(T) \sim 1/T$), then both $x - z$ and $p(t, z, \epsilon)$ are of the order of ϵ with respect to the main term $f_{av}(z)$. It may indeed be useful to the reader to check the lemma in the linear periodic case. Then, the transformation (4.2.15) may be replaced by

$$x(t) = z(t) + \epsilon \left[\int_0^t (A(\tau) - A_{av})\, d\tau \right] z(t)$$

and $\psi(\epsilon)$, $\xi(\epsilon)$ are of the order of ϵ. If $d(t, x)$ has zero mean but unbounded integral, the perturbation terms go to zero as $\epsilon \to 0$, but possibly more slowly than linearly (as $\sqrt{\epsilon}$ for example). The proof of lemma 4.2.1 provides a direct relationship between the order of the convergence to the mean value and the order of the error terms.

We now focus attention on the approximation of the original system by the averaged system. Consider first the following assumption.

(A5) x_0 is sufficiently small so that, for fixed T and some $h'<h$, $x_{av}(t) \in B_{h'}$ for all $t \in [0, T/\epsilon]$ (this is possible, using the Lipschitz assumption (A3) and proposition 1.4.1).

Theorem 4.2.4 Basic Averaging Theorem

If the original system (4.2.1) and the averaged system (4.2.5) satisfy assumptions (A1)–(A5)

Then there exists $\psi(\epsilon)$ as in lemma 4.2.3 such that, given $T \geq 0$

$$|x(t) - x_{av}(t)| \leq \psi(\epsilon) b_T \tag{4.2.19}$$

for some $b_T \geq 0$, $\epsilon_T > 0$, and for all $t \in [0, T/\epsilon]$ and $\epsilon \leq \epsilon_T$.

Proof of Theorem 4.2.4

We apply the transformation of lemma 4.2.3, so that

$$|x - z| \leq \xi(\epsilon)|z| \leq \psi(\epsilon)|z| \tag{4.2.20}$$

for $\epsilon \leq \epsilon_1$. On the other hand, we have that

$$\frac{d}{dt}(z - x_{av}) = \epsilon\left[f_{av}(z) - f_{av}(x_{av})\right] + \epsilon p(t, z, \epsilon) \tag{4.2.21}$$

$$z(0) - x_{av}(0) = 0$$

for all $t \in [0, T/\epsilon]$, $x_{av} \in B_{h'}$, $h'<h$.

We will now show that, on this time interval, and for as long as $x, z \in B_h$, the errors $(z - x_{av})$ and $(x - x_{av})$ can be made arbitrarily small by reducing ϵ. Integrating (4.2.21)

$$|z(t) - x_{av}(t)| \leq \epsilon l_{av}\int_0^t |z(\tau) - x_{av}(\tau)|\, d\tau + \epsilon\psi(\epsilon)\int_0^t |z(\tau)|\, d\tau \tag{4.2.22}$$

Using the Bellman-Gronwall lemma (lemma 1.4.2)

$$\begin{aligned}|z(t) - x_{av}(t)| &\leq \epsilon\psi(\epsilon)\int_0^t |z(\tau)|\, e^{\epsilon l_{av}(t-\tau)} d\tau \\ &\leq \psi(\epsilon) h\left[\frac{e^{\epsilon l_{av} T} - 1}{l_{av}}\right] \\ &:= \psi(\epsilon) a_T \end{aligned} \tag{4.2.23}$$

Combining (4.2.20), (4.2.23)

$$\begin{aligned}|x(t) - x_{av}(t)| &\leq |x(t) - z(t)| + |z(t) - x_{av}(t)| \\ &\leq \psi(\epsilon)|x_{av}(t)| + (1 + \psi(\epsilon))|z(t) - x_{av}(t)| \\ &\leq \psi(\epsilon)(h + (1 + \psi(\epsilon_1)) a_T) \\ &:= \psi(\epsilon) b_T \end{aligned} \tag{4.2.24}$$

By assumption, $|x_{av}(t)| \leq h' < h$. Let ϵ_T (with $0 < \epsilon_T \leq \epsilon_1$) such that $\psi(\epsilon_T) b_T < h - h'$. It follows, from a simple contradiction argument, that $x(t) \in B_h$, and that the estimate in (4.2.24) is valid for all $t \in [0, T/\epsilon]$, whenever $\epsilon \leq \epsilon_T$. □

Comments

Theorem 4.2.4 establishes that the trajectories of the original system and of the averaged system are arbitrarily close on intervals $[0, T/\epsilon]$, when ϵ is sufficiently small. The error is of the order of $\psi(\epsilon)$, and the order is related to the order of convergence of $\gamma(T)$. If $d(t, x)$ has a bounded integral (i.e., $\gamma(T) \sim 1/T$), then the error is of the order of ϵ.

It is important to remember that, although the intervals $[0, T/\epsilon]$ are unbounded, theorem 4.2.4 does not state that

$$|x(t) - x_{av}(t)| \leq \psi(\epsilon) b \tag{4.2.25}$$

for all $t \geq 0$ and some b. Consequently, theorem 4.2.4 does not allow us to relate the stability of the original and of the averaged system. This relationship is investigated in theorem 4.2.5.

Theorem 4.2.5 Exponential Stability Theorem

If the original system (4.2.1) and the averaged system (4.2.5) satisfy assumptions (A1)–(A5), the function $f_{av}(x)$ has continuous and bounded first partial derivatives in x, and $x = 0$ is an exponentially stable equilibrium point of the averaged system

Then the equilibrium point $x = 0$ of the original system is exponentially stable for ϵ sufficiently small.

Proof of Theorem 4.2.5

The proof relies on the converse theorem of Lyapunov for exponentially stable systems (theorem 1.4.3). Under the hypotheses, there exists a function $v(x_{av}) : \mathbb{R}^n \to \mathbb{R}_+$ and strictly positive constants $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ such that, for all $x_{av} \in B_h$,

$$\alpha_1 |x_{av}|^2 \leq v(x_{av}) \leq \alpha_2 |x_{av}|^2 \tag{4.2.26}$$

$$\dot{v}(x_{av})\Big|_{(4.2.5)} \leq -\epsilon\alpha_3 |x_{av}|^2 \qquad (4.2.27)$$

$$\left|\frac{\partial v}{\partial x_{av}}\right| \leq \alpha_4 |x_{av}| \qquad (4.2.28)$$

The derivative in (4.2.27) is to be taken along the trajectories of the averaged system (4.2.5).

The function v is now used to study the stability of the perturbed system (4.2.17), where $z(x)$ is defined by (4.2.15). Considering $v(z)$, inequalities (4.2.26) and (4.2.28) are still verified, with z replacing x_{av}. The derivative of $v(z)$ along the trajectories of (4.2.17) is given by

$$\dot{v}(z)\Big|_{(4.2.17)} = \dot{v}(z)\Big|_{(4.2.5)} + \left[\frac{\partial v}{\partial z}\right](\epsilon p(t, z, \epsilon)) \qquad (4.2.29)$$

and, using previous inequalities (including those from lemma 4.2.3)

$$\dot{v}(z)\Big|_{(4.2.17)} \leq -\epsilon\alpha_3 |z|^2 + \epsilon\alpha_4\psi(\epsilon)|z|^2$$

$$\leq -\epsilon\left[\frac{\alpha_3 - \psi(\epsilon)\alpha_4}{\alpha_2}\right] v(z) \qquad (4.2.30)$$

for all $\epsilon \leq \epsilon_1$. Let ϵ_2' be such that $\alpha_3 - \psi(\epsilon_2')\alpha_4 > 0$, and define $\epsilon_2 = \min(\epsilon_1, \epsilon_2')$. Denote

$$\alpha(\epsilon) := \frac{\alpha_3 - \psi(\epsilon)\alpha_4}{2\alpha_2} \qquad (4.2.31)$$

Consequently, (4.2.30) implies that

$$v(z) \leq v(z(t_0))\, e^{-2\epsilon\alpha(\epsilon)(t-t_0)} \qquad (4.2.32)$$

and

$$|z(t)| \leq \left[\frac{\alpha_2}{\alpha_1}\right]^{\frac{1}{2}} |z(t_0)|\, e^{-\epsilon\alpha(\epsilon)(t-t_0)} \qquad (4.2.33)$$

Since $\alpha(\epsilon) > 0$ for all $\epsilon \leq \epsilon_2$, system (4.2.17) is exponentially stable. Using (4.2.16), it follows that

$$|x(t)| \leq \frac{1+\xi(\epsilon)}{1-\xi(\epsilon)}\left[\frac{\alpha_2}{\alpha_1}\right]^{\frac{1}{2}} |x(t_0)|\, e^{-\epsilon\alpha(\epsilon)(t-t_0)} \qquad (4.2.34)$$

for all $t \geq t_0 \geq 0$, $\epsilon \leq \epsilon_2$, and $x(t_0)$ sufficiently small that all signals remain in B_h. In other words, the original system is exponentially stable, with rate of convergence (at least) $\epsilon\alpha(\epsilon)$. □

Comments

a) Theorem 4.2.5 is a *local* exponential stability result. The original system will be *globally* exponentially stable if the averaged system is globally exponentially stable, and provided that *all* assumptions are valid globally.

b) The proof of theorem 4.2.5 gives a useful bound on the rate of convergence of the original system. As ϵ tends to zero, $\epsilon\alpha(\epsilon)$ tends to $\epsilon/2\,\alpha_3/\alpha_2$, which is the bound on the rate of convergence of the averaged system that one would obtain using (4.2.26)–(4.2.27). In other words, the proof provides a bound on the rate of convergence, and this bound gets arbitrarily close to the corresponding bound for the averaged system, provided that ϵ is sufficiently small. This is a useful conclusion because it is in general very difficult to obtain a guaranteed rate of convergence for the original, nonautonomous system. The proof assumes the existence of a Lyapunov function satisfying (4.2.26)–(4.2.28), but does not depend on the specific function chosen. Since the averaged system is autonomous, it is usually easier to find such a function for it than for the original system, and any such function will provide a bound on the rate of convergence of the original system for ϵ sufficiently small.

c) The conclusion of theorem 4.2.5 is quite different from the conclusion of theorem 4.2.4. Since both x and x_{av} go to zero exponentially with t, the error $x - x_{av}$ also goes to zero exponentially with t. Yet theorem 4.2.5 does not relate the bound on the error to ϵ. It is possible, however, to combine theorem 4.2.4 and theorem 4.2.5 to obtain a uniform approximation result, with an estimate similar to (4.2.25).

4.3 APPLICATION TO IDENTIFICATION

To apply the averaging theory to the identifier described in Chapter 2, we will study the case when $g = \epsilon > 0$ and the update law is given by (cf. (2.4.1))

$$\dot{\phi}(t) = -g\, e_1(t)\, w(t) \qquad \phi(0) = \phi_0 \qquad (4.3.1)$$

The evolution of the parameter error is described by

$$\dot{\phi}(t) = -g\, w(t)\, w^T(t)\, \phi(t) \qquad \phi(0) = \phi_0 \qquad (4.3.2)$$

In theorem 2.5.1, we found that system (4.3.2) is exponentially stable, provided that w is *persistently exciting*, i.e., there exist constants $\alpha_1, \alpha_2, \delta > 0$, such that

$$\alpha_2 I \geq \int_{t_0}^{t_0+\delta} w(\tau)\, w^T(\tau)\, d\tau \geq \alpha_1 I \qquad \text{for all } t_0 \geq 0 \qquad (4.3.3)$$

On the other hand, the averaging theory presented above leads us to the limit

$$R_w(0) := \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} w(\tau)\, w^T(t+\tau)\, d\tau \quad \in \mathbb{R}^{n \times n} \qquad (4.3.4)$$

where we used the notation of Section 1.6 for the *autocovariance* of w evaluated at 0. Recall that $R_w(t)$ may be expressed as the inverse Fourier transform of the positive *spectral measure* $S_w(d\omega)$

$$R_w(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} S_w(d\omega) \qquad (4.3.5)$$

Further, w is the output of a proper stable transfer function $\hat{H}_{wr}$ given by (cf. (2.2.16)–(2.2.17))

$$\hat{H}_{wr}(s) = \begin{bmatrix} (sI - \Lambda)^{-1} b_\lambda \\ (sI - \Lambda)^{-1} b_\lambda \hat{P}(s) \end{bmatrix} \quad \in \mathbb{R}^{2n}(s) \qquad (4.3.6)$$

Therefore, if the input r is stationary, then w is also stationary. Its spectrum is related to the spectrum of r through

$$S_w(d\omega) = \hat{H}_{wr}^{*}(j\omega)\, \hat{H}_{wr}^{T}(j\omega)\, S_r(d\omega) \qquad (4.3.7)$$

and, using (4.3.5) and (4.3.7), we have that

$$R_w(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{H}_{wr}^{*}(j\omega)\, \hat{H}_{wr}^{T}(j\omega)\, S_r(d\omega) \qquad (4.3.8)$$

Since $S_r(d\omega)$ is an even function of ω, $R_w(0)$ is also given by

$$R_w(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{Re}\left[\hat{H}_{wr}^{*}(j\omega)\, \hat{H}_{wr}^{T}(j\omega) \right] S_r(d\omega)$$

It was shown in Section 2.7 (proposition 2.7.1) that when w is stationary, w is persistently exciting (PE) if and only if $R_w(0)$ is positive definite. It followed (proposition 2.7.2) that this is true if the support of $S_r(d\omega)$ is greater than or equal to $2n$ points (the dimension of w = the number of unknown parameters = $2n$). Note that a DC component in $r(t)$ contributes one point to the support of $S_r(d\omega)$, while a sinusoidal component contributes two points (at $+\omega$ and $-\omega$).

With these definitions, the averaged system corresponding to (4.3.2) is simply

$$\dot{\phi}_{av} = -g R_w(0)\, \phi_{av} \qquad \phi_{av}(0) = \phi_0 \qquad (4.3.9)$$

This system is particularly easy to study, since it is linear.

Convergence Analysis

When w is persistently exciting, $R_w(0)$ is a positive definite matrix. A natural Lyapunov function for (4.3.9) is

$$v(\phi_{av}) = \frac{1}{2} |\phi_{av}|^2 = \frac{1}{2} \phi_{av}^T \phi_{av} \qquad (4.3.10)$$

and

$$-g\, \lambda_{\min}(R_w(0))\, |\phi_{av}|^2 \le -\dot{v}(\phi_{av}) \le -g\, \lambda_{\max}(R_w(0))\, |\phi_{av}|^2 \qquad (4.3.11)$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are, respectively, the minimum and maximum eigenvalues of $R_w(0)$. Thus, the rate of exponential convergence of the averaged system is at least $g\, \lambda_{\min}(R_w(0))$ and at most $g\, \lambda_{\max}(R_w(0))$. We can conclude that the rate of convergence of the original system for g small enough is close to the interval $[g\, \lambda_{\min}(R_w(0)),\, g\, \lambda_{\max}(R_w(0))]$.

Equation (4.3.8) gives an interpretation of $R_w(0)$ in the frequency domain, and also a mean of computing an estimate of the rate of convergence of the adaptive algorithm, given the spectral content of the reference input. If the input r is periodic or almost periodic

$$r(t) = \sum_k r_k \sin(\omega_k t) \qquad (4.3.12)$$

then the integral in (4.3.8) may be replaced by a summation

$$R_w(0) = \sum_k \frac{r_k^2}{2} \mathrm{Re}\left[\hat{H}_{wr}^{*}(j\omega_k)\, \hat{H}_{wr}^{T}(j\omega_k) \right] \qquad (4.3.13)$$

Since the transfer function $\hat{H}_{wr}$ depends on the unknown plant being identified, the use of (4.3.11) to determine the rate of convergence is limited. With knowledge of the plant, it could be used to determine the spectral content of the reference input that will optimize the rate of convergence of the identifier, given the physical constraints on r. Such a procedure is very reminiscent of the procedure indicated in Goodwin & Payne [1977] (Chapter 6) for the design of input signals in identification. The autocovariance matrix defined here is similar to the *average information matrix* defined in Goodwin & Payne [1977] (p. 134). Our interpretation is, however, in terms of rates of parameter convergence of the averaged system rather than in terms of parameter error covariance in a stochastic framework.

Note that the proof of exponential stability of theorem 2.5.1 was based on the Lyapunov function of theorem 1.4.1 that was an average of the norm along the trajectories of the system. In this chapter, we averaged the *differential equation* itself and found that the norm becomes a Lyapunov function to prove exponential stability.

It is also interesting to compare the convergence rate obtained through averaging with the convergence rate obtained in Chapter 2. We found, in the proof of exponential convergence of theorem 2.5.1, that the estimate of the convergence rate tends to $g\alpha_1/\delta$ when the adaptation gain g tends to zero. The constants α_1, δ resulted from the PE condition (2.5.3), i.e., (4.3.3). By comparing (4.3.3) and (4.3.4), we find that the estimates provided by direct proof and by averaging are essentially identical for $g = \epsilon$ small.

Example

To illustrate the conclusions of this section, we consider the following example

$$\hat{P}(s) = \frac{k_p}{s + a_p} \tag{4.3.14}$$

The filter is chosen to be $\hat{\lambda}(s) = (s + l_2)/l_1$ (where $l_1 = 10.05$, $l_2 = 10$ are arbitrarily chosen such that $|\hat{\lambda}(j1)| = 1$). Although $\hat{\lambda}$ is not monic, the gain l_1 can easily be taken into account.

Since the number of unknown parameters is 2, parameter convergence will occur when the support of $S_r(d\omega)$ is greater than or equal to 2 points. We consider an input of the form $r = r_0 \sin(\omega_0 t)$, so that the support consists of exactly 2 points.

The averaged system can be found by using (4.3.9), (4.3.13)

$$\dot{\phi}_{av} = -g \frac{r_0^2}{2} \frac{l_1^2}{l_2^2 + \omega_0^2} \begin{bmatrix} 1 & \dfrac{a_p k_p}{\omega_0^2 + a_p^2} \\ \dfrac{a_p k_p}{\omega_0^2 + a_p^2} & \dfrac{k_p^2}{\omega_0^2 + a_p^2} \end{bmatrix} \cdot \phi_{av} \tag{4.3.15}$$

with $\phi_{av}(0) = \phi_0$. When $r_0 = 1$, $\omega_0 = 1$, $a_p = 1$, $k_p = 2$, the eigenvalues of the averaged system (4.3.15) are computed to be $-\frac{3+\sqrt{5}}{4}g = -1.309g$, and $-\frac{3-\sqrt{5}}{4}g = -0.191g$. The nominal parameter $\theta^{*T} = (k_p/l_1, \ (l_2 - a_p)/l_1)$. We let $\theta(0) = 0$, so that

$\phi^T(0) = (-0.199, -0.9)$.

Figures 4.1 to 4.4 show the plots of the parameter errors ϕ_1 and ϕ_2, for both the original and averaged systems, and with two different adaptation gains $g = 1$, and $g = 0.1$.

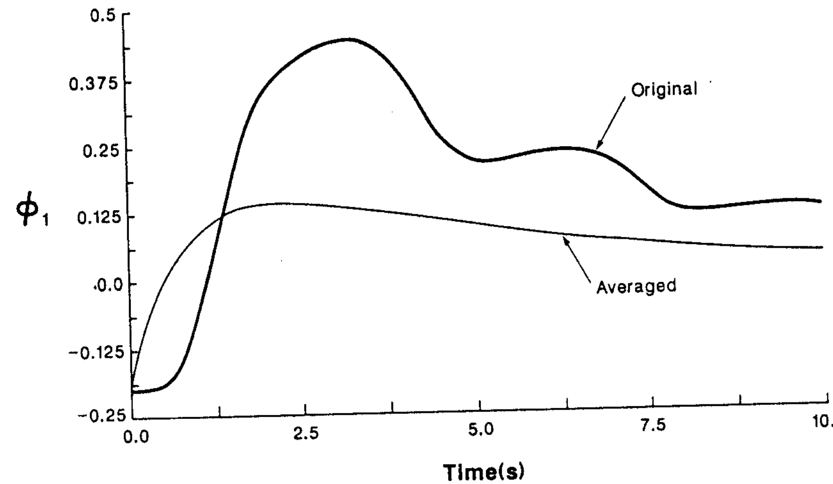

Figure 4.1: Parameter Error ϕ_1 ($g = 1$)

We notice the closeness of the approximation for $g = 0.1$.

Figures 4.5 and 4.6 are plots of the Lyapunov function (4.3.10) for $g = 1$ and $g = 0.1$, using a logarithmic scale. We observe the two slopes, corresponding to the two eigenvalues. The closeness of the estimate of the convergence rate by the averaged system can also be appreciated from these figures.

Figure 4.7 represents the two components of ϕ, one as a function of the other when $g = 0.1$. It shows the two subspaces corresponding to the small and large eigenvalues: the parameter error first moves fast along the direction of the eigenvector corresponding to the large eigenvalue. Then, it slowly moves along the direction corresponding to the small eigenvalue.

4.4 AVERAGING THEORY—TWO-TIME SCALES

We now consider a more general class of differential equations arising in the adaptive control schemes presented in Chapter 3.

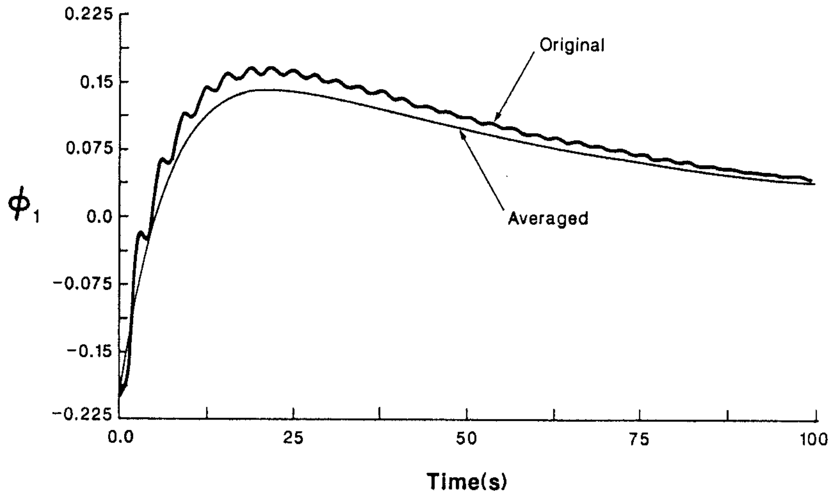

Figure 4.2: Parameter Error ϕ_1 (g = 0.1)

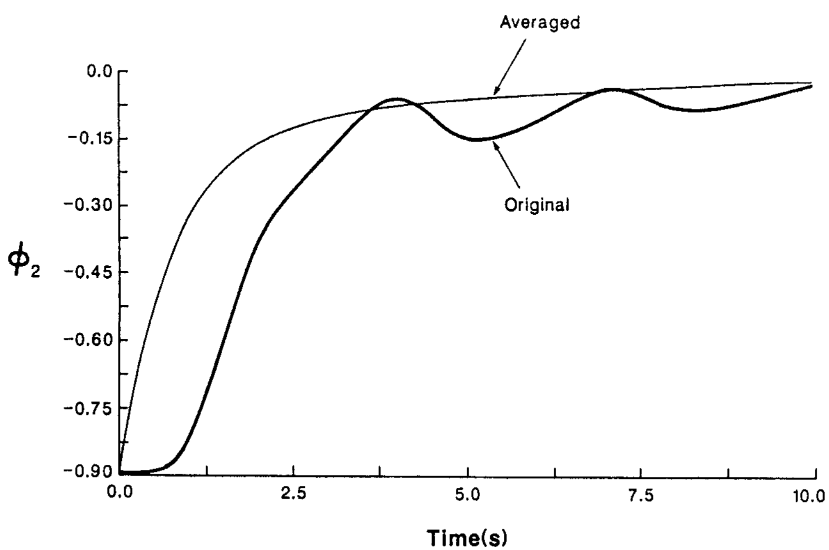

Figure 4.3: Parameter Error ϕ_2 (g = 1)

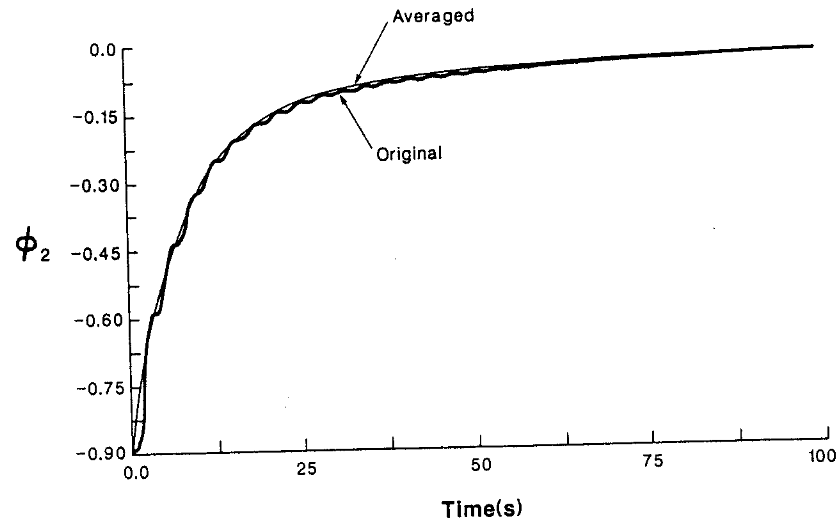

Figure 4.4: Parameter Error ϕ_2 (g = 0.1)

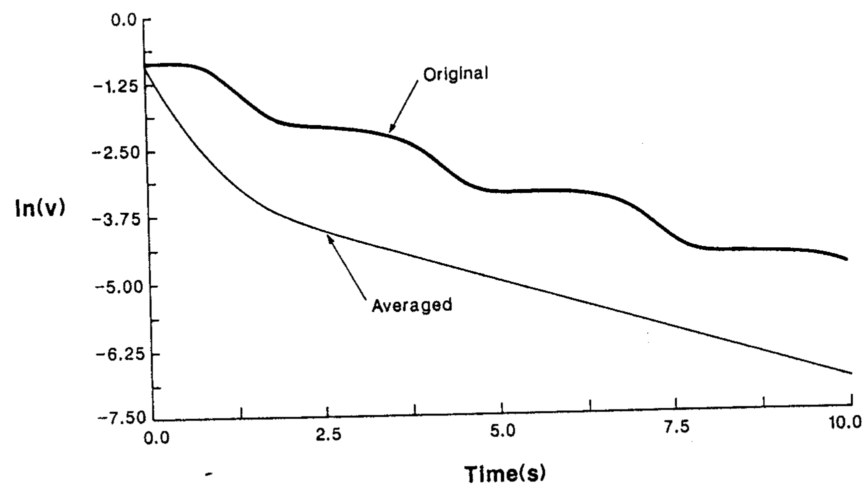

Figure 4.5: Logarithm of the Lyapunov Function (g = 1)

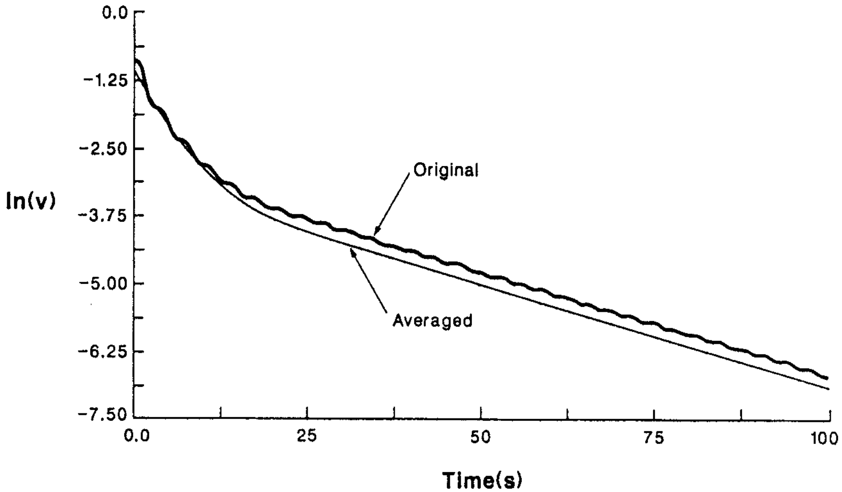

Figure 4.6: Logarithm of the Lyapunov Function ($g = 0.1$)

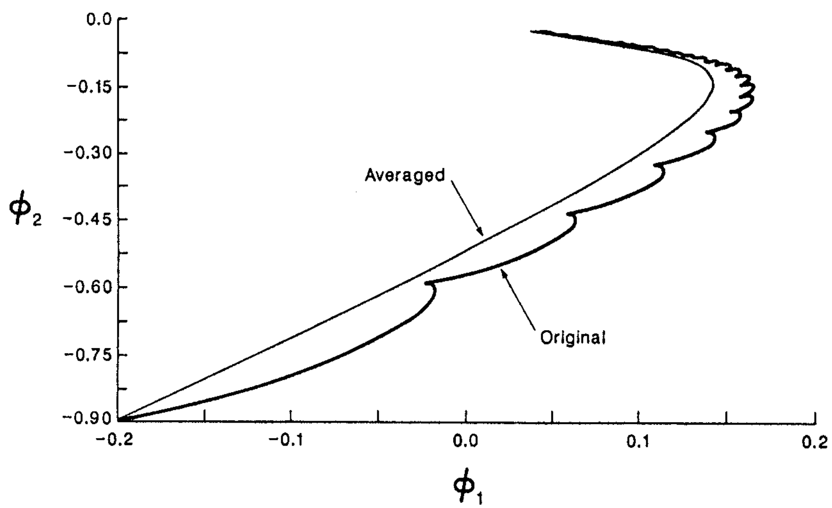

Figure 4.7: Parameter Error $\phi_2(\phi_1)$ ($g = 0.1$)

4.4.1 Separated Time Scales

We first consider the system of differential equations

$$\dot{x} = \epsilon f(t, x, y) \tag{4.4.1}$$

$$\dot{y} = A(x)y + \epsilon g(t, x, y) \tag{4.4.2}$$

where $x(0) = x_0$, $y(0) = y_0$, $x \in \mathbb{R}^n$, and $y \in \mathbb{R}^m$.

The state vector is divided into a fast state vector y and a slow state vector x, whose dynamics are of the order of ϵ with respect to the fast dynamics. The dominant term in (4.4.2) is linear in y, but is itself allowed to vary as a function of the slow state vector.

As previously, we define

$$f_{av}(x) = \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} f(\tau, x, 0)\, d\tau \tag{4.4.3}$$

and the system

$$\dot{x}_{av} = f_{av}(x_{av}) \qquad x_{av}(0) = x_0 \tag{4.4.4}$$

is the *averaged system* corresponding to (4.4.1)–(4.4.2). We make the following additional assumption.

Definition Uniform Exponential Stability of a Family of Square Matrices

The family of matrices $A(x) \in \mathbb{R}^{m \times m}$ is *uniformly exponentially stable for all* $x \in B_h$, if there exist m, λ, m', $\lambda' > 0$, such that for all $x \in B_h$ and $t \geq 0$

$$m' e^{-\lambda' t} \leq \| e^{A(x)t} \| \leq m\, e^{-\lambda t} \tag{4.4.5}$$

Comments

This definition is equivalent to require that the solutions of the system $\dot{y} = A(x)y$ are bounded above and below by decaying exponentials, independently of the parameter x.

It is also possible to show that the definition is equivalent to requiring that there exist p_1, p_2, q_1, $q_2 > 0$, such that for all $x \in B_h$, there exists $P(x)$ satisfying $p_1 I \leq P(x) \leq p_2 I$, and $-q_2 I \leq A^T(x)P(x) + P(x)A(x) \leq -q_1 I$.

We will make the following assumptions.

Assumptions

For some $h > 0$

(B1) The functions f and g are piecewise continuous functions of time, and continuous functions of x and y. Moreover, $f(t, 0, 0) = 0$, $g(t, 0, 0) = 0$ for all $t \geq 0$, and for some $l_1, l_2, l_3, l_4 \geq 0$

$$|f(t, x_1, y_1) - f(t, x_2, y_2)| \leq l_1 |x_1 - x_2| + l_2 |y_1 - y_2|$$

$$|g(t, x_1, y_1) - g(t, x_2, y_2)| \leq l_3 |x_1 - x_2| + l_4 |y_1 - y_2| \quad (4.4.6)$$

for all $t \geq 0$, $x_1, x_2 \in B_h$, $y_1, y_2 \in B_h$. Also assume that $f(t, x, 0)$ has continuous and bounded first partial derivatives with respect to x, for all $t \geq 0$, and $x \in B_h$.

(B2) The function $f(t, x, 0)$ has average value $f_{av}(x)$. Moreover, $f_{av}(0) = 0$, and $f_{av}(x)$ has continuous and bounded first partial derivatives with respect to x, for all $x \in B_h$, so that for some $l_{av} \geq 0$

$$|f_{av}(x_1) - f_{av}(x_2)| \leq l_{av} |x_1 - x_2| \quad (4.4.7)$$

for all $x_1, x_2 \in B_h$.

(B3) Let $d(t, x) = f(t, x, 0) - f_{av}(x)$, so that $d(t, x)$ has zero average value. Assume that the convergence function can be written as $\gamma(T)|x|$, and that $\frac{\partial d(t, x)}{\partial x}$ has zero average value, with convergence function $\gamma(T)$.

(B4) A(x) is uniformly exponentially stable for all $x \in B_h$ and, for some $k_a \geq 0$

$$\left\| \frac{\partial A(x)}{\partial x} \right\| \leq k_a \quad \text{for all } x \in B_h \quad (4.4.8)$$

(B5) For some $h' < h$, $|x_{av}(t)| \in B_{h'}$ on the time intervals considered, and for some $h_0, y_0 \in B_{h_0}$ (where h', h_0 are constants to be defined later). This assumption is technical, and will allow us to guarantee that all signals remain in B_h.

As for one-time scale systems, we first obtain the following preliminary lemma, similar to lemma 4.2.3.

Lemma 4.4.1 Perturbation Formulation of Averaging—Two-Time Scales

If the original system (4.4.1)–(4.4.2) and the averaged system (4.4.4) satisfy assumptions (B1)–(B3)

Then there exist functions $w_\epsilon(t, x)$, $\xi(\epsilon)$ as in lemma 4.2.2 and $\epsilon_1 > 0$, such that the transformation

$$x = z + \epsilon w_\epsilon(t, z) \quad (4.4.9)$$

is a homeomorphism in B_h for all $\epsilon \leq \epsilon_1$, and

$$|x - z| \leq \xi(\epsilon)|z| \quad (4.4.10)$$

Under the transformation, system (4.4.1) becomes

$$\dot{z} = \epsilon f_{av}(z) + \epsilon p_1(t, z, \epsilon) + \epsilon p_2(t, z, y, \epsilon) \quad (4.4.11)$$

$$z(0) = x_0$$

where

$$|p_1(t, z, \epsilon)| \leq \xi(\epsilon) k_1 |z|$$

$$|p_2(t, z, y, \epsilon)| \leq k_2 |y| \quad (4.4.12)$$

for some k_1, k_2 depending on l_1, l_2, l_{av}.

Proof of Lemma 4.4.1 in Appendix.

We are now ready to state the averaging theorems concerning the differential system (4.4.1)–(4.4.2). Theorem 4.4.2 is an approximation theorem similar to theorem 4.2.4 and guarantees that the trajectories of the original and averaged system are arbitrarily close on compact intervals, when ϵ tends to zero. Theorem 4.4.3 is an exponential stability theorem, similar to theorem 4.2.5.

Theorem 4.4.2 Basic Averaging Theorem

If the original system (4.4.1)–(4.4.2) and the averaged system (4.4.4) satisfy assumptions (B1)–(B5)

Then there exists $\psi(\epsilon)$ as in lemma 4.2.3 such that, given $T \geq 0$

$$|x(t) - x_{av}(t)| \leq \psi(\epsilon) b_T \quad (4.4.13)$$

for some $b_T \geq 0$, $\epsilon_T > 0$ and for all $t \in [0, T/\epsilon]$, and $\epsilon \leq \epsilon_T$.

Theorem 4.4.3 Exponential Stability Theorem

If the original system (4.4.1)–(4.4.2) and the averaged system (4.4.4) satisfy assumptions (B1)–(B5), the function $f_{av}(x)$ has continuous and bounded first partial derivatives in x, and $x = 0$ is an exponentially stable equilibrium point of the averaged system

Then the equilibrium point $x = 0, y = 0$ of the original system is exponentially stable for ϵ sufficiently small.

Comments

As for theorem 4.2.5, the proof of theorem 4.4.3 gives a useful bound on the rate of convergence of the nonautonomous system. As $\epsilon \to 0$, the rate tends to the bound on the rate of convergence of the averaged system that one would obtain using the Lyapunov function for the averaged system. Since the averaged system is autonomous, it is usually easier to obtain such a Lyapunov function for the averaged system than for the original nonautonomous system, and conclusions about its exponential convergence can be applied to the nonautonomous system for ϵ sufficiently small.

4.4.2 Mixed Time Scales

We now discuss a more general class of two-time scale systems, arising in adaptive control

$$\dot{x} = \epsilon f'(t, x, y') \tag{4.4.14}$$

$$\dot{y}' = A(x)y' + h(t, x) + \epsilon g'(t, x, y') \tag{4.4.15}$$

We will show that system (4.4.14)–(4.4.15) can be transformed into the system (4.4.1)–(4.4.2). In this case, x is a slow variable, but y' has both a fast and a slow component.

The averaged system corresponding to (4.4.14), (4.4.15) is obtained as follows. Define the function

$$v(t, x) := \int_0^t e^{A(x)(t-\tau)} h(\tau, x)\, d\tau \tag{4.4.16}$$

and assume that the following limit exists uniformly in t and x

$$f_{av}(x) = \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} f'(\tau, x, v(\tau, x))\, d\tau \tag{4.4.17}$$

Intuitively, $v(t, x)$ represents the steady-state value of the variable y' with x frozen and $\epsilon = 0$ in (4.4.15). Then, f is averaged with $v(t, x)$ replacing y' in (4.4.14).

Consider now the transformation

$$y = y' - v(t, x) \tag{4.4.18}$$

Since $v(t, x)$ satisfies

$$\frac{\partial}{\partial t} v(t, x) = A(x) v(t, x) + h(t, x) \qquad v(t, 0) = 0 \tag{4.4.19}$$

we have that

$$\dot{y} = A(x)y + \epsilon \Big[-\frac{\partial v(t, x)}{\partial x} f'(t, x, y + v(t, x)) + g'(t, x, y + v(t, x)) \Big] \tag{4.4.20}$$

so that (4.4.14), (4.4.20) is of the form of (4.4.1), (4.4.2) when

$$f(t, x, y) = f'(t, x, y + v(t, x)) \tag{4.4.21}$$

$$g(t, x, y) = -\frac{\partial v(t, x)}{\partial x} f'(t, x, y + v(t, x)) + g'(t, x, y + v(t, x)) \tag{4.4.22}$$

The averaged system is obtained by averaging the right-hand side of (4.4.21) with $y = 0$, so that the definitions (4.4.17), and (4.4.3) (with f given by (4.4.21)) agree.

To apply theorems 4.4.2 and 4.4.3, we require Assumptions (B1)–(B5) to be satisfied. In particular, we assume similar Lipschitz conditions on f', g', and the following assumption on $h(t, x)$

(B6) $h(t, 0) = 0$ for all $t \geq 0$, and $\left\| \frac{\partial h(t, x)}{\partial x} \right\|$ is bounded for all $t \geq 0$, $x \in B_h$.

This new assumption implies that $v(t, 0) = 0$. It also implies that $\left\| \frac{\partial v(t, x)}{\partial x} \right\|$ is bounded for all $t \geq 0$, $x \in B_h$, since

$$\frac{\partial v(t, x)}{\partial x_i} = \int_0^t \Big[e^{A(x)(t-\tau)} \frac{\partial h(\tau, x)}{\partial x_i} + \frac{\partial}{\partial x_i} \Big[e^{A(x)(t-\tau)} \Big] h(\tau, x) \Big] d\tau \tag{4.4.23}$$

and using the fact that $e^{A(x)(t-\tau)}$ and $\frac{\partial}{\partial x} e^{A(x)(t-\tau)}$ are bounded by exponentials ((4.4.5), and (A4.4.30) in the proof of theorem 4.4.3).

4.5 APPLICATIONS TO ADAPTIVE CONTROL

For illustration, we apply the previous results to the output error direct adaptive control-algorithm for the relative degree 1 case.

We established the complete description of the adaptive system in Section 3.5 with (3.5.28), i.e.,

$$\dot{e}(t) = A_m e(t) + b_m \phi^T(t) w_m(t) + b_m \phi^T(t) Q\, e(t)$$

$$\dot{\phi}(t) = -g\, c_m^T e(t)\, w_m(t) - g\, c_m^T e(t)\, Q\, e(t) \qquad (4.5.1)$$

where g is the adaptation gain. With the exception of the last terms (quadratic in e and ϕ), (4.5.1) is a set of linear time varying differential equations. They describe the adaptive control system, linearized around the equilibrium $e = 0$, $\phi = 0$. We first study these equations, then turn to the nonlinear equations.

4.5.1 Output Error Scheme—Linearized Equations

The linearized equations, describing the adaptive system for small values of e and ϕ, are

$$\dot{e}(t) = A_m e(t) + b_m w_m^T(t)\, \phi(t)$$
$$\dot{\phi}(t) = -g\, w_m(t)\, c_m{}^T e(t) \qquad (4.5.2)$$

Since w_m is bounded, it is easy to see that (4.5.2) is of the form of (4.4.14), (4.4.15) with the functions f' and h satisfying the conditions of Section 4.4. Recall that A_m is a stable matrix.

The function $v(t, \phi)$ defined in (4.4.16) is now

$$v(t, \phi) = \left[\int_0^t e^{A_m(t-\tau)} b_m w_m^T(\tau)\, d\tau \right] \phi \qquad (4.5.3)$$

and f_{av} is given by

$$f_{av}(\phi) = -\lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} w_m(t)\, c_m^T \cdot \left[\int_0^t e^{A_m(t-\tau)} b_m w_m^T(\tau)\, d\tau \right] dt\ \phi \qquad (4.5.4)$$

Frequency Domain Analysis

To derive frequency domain expressions, we assume that r is stationary. Since the transfer function from $r \to w_m$ is stable, this implies that w_m is stationary. The spectral measure of w_m is related to that of r by

$$S_{w_m}(d\omega) = \hat{H}^*_{w_m r}(j\omega)\, \hat{H}^T_{w_m r}(j\omega)\, S_r(d\omega) \qquad (4.5.5)$$

where the transfer function from $r \to w_m$ is given by (using (3.5.11))

$$\hat{H}_{w_m r} = \begin{bmatrix} 1 \\ (sI - \Lambda)^{-1} b_\lambda \hat{P}^{-1} \hat{M} \\ \hat{M} \\ (sI - \Lambda)^{-1} b_\lambda \hat{M} \end{bmatrix} \qquad (4.5.6)$$

which is a stable transfer function.

Define a filtered version of w_m to be

$$w_{mf}(t) = \int_0^t c_m^T e^{A_m(t-\tau)} b_m w_m(\tau)\, d\tau = \frac{1}{c_0^*} \hat{M}(w_m) \qquad (4.5.7)$$

where the last equality follows from (3.5.22). Note that the signal w_f was also used in the direct proof of exponential convergence in Chapter 2 (cf. (2.6.34)).

Since $c_m{}^T (sI - A_m)^{-1} b_m = \frac{1}{c_0^*} \hat{M}(s)$ is stable, $w_{mf}(t)$ is stationary. We let

$$R_{w_m w_{mf}}(0) = \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} w_m(t)\, w_{mf}^T(t)\, dt \qquad (4.5.8)$$

which was called the *cross correlation* between w_m and w_{mf} (evaluated at 0) in Section 1.6. Consequently, we may use (4.5.7) and (4.5.8) to obtain a frequency domain expression for $R_{w_m w_{mf}}(0)$ as

$$R_{w_m w_{mf}}(0) = \frac{1}{2\pi c_0^*} \int_{-\infty}^{\infty} \hat{H}^*_{w_m r}(j\omega)\, \hat{H}^T_{w_m r}(j\omega)\, \hat{M}(j\omega)\, S_r(d\omega) \qquad (4.5.9)$$

Since r is a scalar, $S_r(d\omega)$ is *real* and consequently an even function of ω (cf. Section 1.6). This may be used to show that

$$R_{w_m w_{mf}}(0) = \frac{1}{2\pi c_0^*} \int_{-\infty}^{\infty} \mathrm{Re}\left[\hat{H}^*_{w_m r}(j\omega)\, \hat{H}^T_{w_m r}(j\omega) \right] \mathrm{Re}\, \hat{M}(j\omega)\, S_r(d\omega)$$
$$+ \frac{1}{2\pi c_0^*} \int_{-\infty}^{\infty} \mathrm{Im}\left[\hat{H}^*_{w_m r}(j\omega)\, \hat{H}^T_{w_m r}(j\omega) \right] \mathrm{Im}\, \hat{M}(j\omega)\, S_r(d\omega)$$

and that the first matrix in the right-hand side is symmetric, while the second in antisymmetric.

With (4.5.7) and (4.5.8), (4.5.4) shows that the averaged system is a LTI system

$$\dot{\phi}_{av} = -g\,R_{w_m w_{mf}}(0)\,\phi_{av} \qquad \phi_{av}(0) = \phi_0 \tag{4.5.10}$$

Convergence Analysis

Since $\hat{M}(s)$ is strictly positive real, the matrix $R_{w_m w_{mf}}(0)$ is a positive semidefinite matrix. Unlike the matrix $R_w(0)$ of Section 4.3, $R_{w_m w_{mf}}(0)$ need not be symmetric, so that its eigenvalues need not be real. However, the real parts are guaranteed to be positive, and a natural Lyapunov function is again

$$v(\phi_{av}) = |\phi_{av}|^2 = \phi_{av}^T \phi_{av} \tag{4.5.11}$$

and

$$-\dot{v}(\phi_{av}) = g\,\phi_{av}^T \left[R_{w_m w_{mf}}(0) + R_{w_m w_{mf}}^T(0) \right] \phi_{av} \tag{4.5.12}$$

The matrix in parentheses is symmetric positive semidefinite. As previously, it is positive definite if w_m is PE.

When the reference input r is periodic or almost periodic, i.e.,

$$r(t) = \sum_k r_k \sin(\omega_k t) \tag{4.5.13}$$

an expression for $R_{w_m w_{mf}}(0)$ is

$$\begin{aligned} R_{w_m w_{mf}}(0) = {} & \frac{1}{c_0^*} \sum_k \left[\frac{r_k^2}{2} \operatorname{Re}\left[\hat{H}^*_{w_m r}(j\omega_k) \hat{H}^T_{w_m r}(j\omega_k) \right] \cdot \operatorname{Re} \hat{M}(j\omega_k) \right] \\ & + \frac{1}{c_0^*} \sum_k \left[\frac{r_k^2}{2} \operatorname{Im}\left[\hat{H}^*_{w_m r}(j\omega_k) \hat{H}^T_{w_m r}(j\omega_k) \right] \cdot \operatorname{Im} \hat{M}(j\omega_k) \right] \end{aligned} \tag{4.5.14}$$

Example

As an illustration of the preceding results, we consider the following example of a first order plant with an unknown pole and an unknown gain

$$\hat{P}(s) = \frac{k_p}{s + a_p} \tag{4.5.15}$$

We will choose values of the parameters corresponding to the Rohrs examples (Rohrs *et al* [1982], see also Section 5.2), when no unmodeled dynamics are present.

The adaptive process is to adjust the feedforward gain c_0 and the feedback gain d_0 so as to make the closed-loop transfer function match the model transfer function

$$\hat{M}(s) = \frac{k_m}{s + a_m} \tag{4.5.16}$$

To guarantee persistency of excitation, we use a sinusoidal input signal of the form

$$r(t) = r_0 \sin(\omega_0 t) \tag{4.5.17}$$

Thus, (4.5.2) becomes

$$\begin{aligned} \dot{e}_0(t) &= -a_m e_0(t) + k_p(\phi_r(t) r(t) + \phi_y(t) y_m(t)) \\ \dot{\phi}_r(t) &= -g\,e_0(t) r(t) \\ \dot{\phi}_y(t) &= -g\,e_0(t) y_m(t) \end{aligned} \tag{4.5.18}$$

where

$$\begin{aligned} \phi_r(t) &= c_0(t) - c_0^* \\ \phi_y(t) &= d_0(t) - d_0^* \end{aligned} \tag{4.5.19}$$

It can be checked, using (4.5.14), that the averaged system defined in (4.5.10) is now

$$\dot{\phi}_{av} = -g\,\frac{r_0^2}{2}\,\frac{k_p}{k_m} \begin{bmatrix} \dfrac{a_m k_m}{(a_m^2 + \omega_0^2)} & \dfrac{k_m^2 (a_m^2 - \omega_0^2)}{(a_m^2 + \omega_0^2)^2} \\ \dfrac{k_m^2}{(a_m^2 + \omega_0^2)} & \dfrac{a_m k_m^3}{(a_m^2 + \omega_0^2)^2} \end{bmatrix} \phi_{av} \tag{4.5.20}$$

With $a_m = 3$, $k_m = 3$, $a_p = 1$, $k_p = 2$, $r_0 = 1$, $\omega_0 = 1$, $g = 1$, the two eigenvalues of the averaged system are computed to be $-0.0163g$ and $-0.5537g$, and are both real negative. The nominal parameter $\theta^{*T} = (k_m/k_p, (a_p - a_m)/k_p)$. We let $\theta(0) = 0$, so that $\phi^T(0) = (-1.5, 1)$.

Figures 4.8, 4.9 and 4.10 show the plots of the parameter errors $\phi_y(\phi_r)$ for the original and averaged system, with three different frequencies ($\omega_0 = 1, 3, 5$).

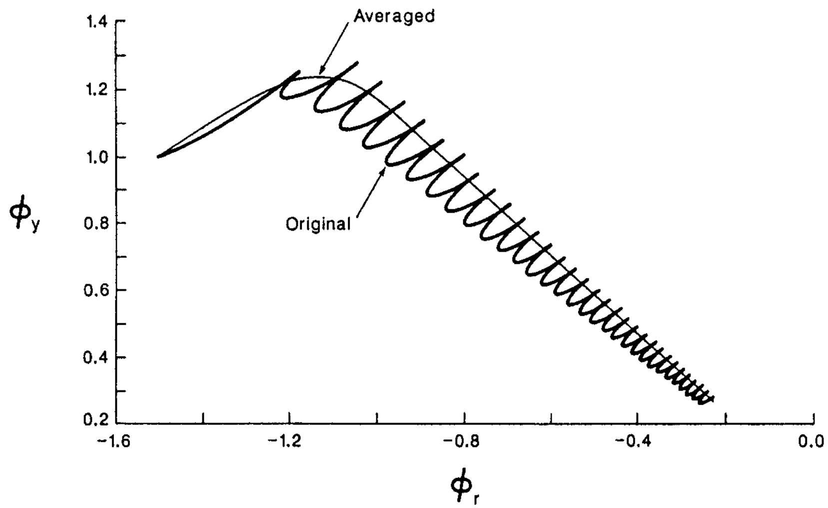

Figure 4.8: Parameter Error $\phi_y(\phi_r)$ ($r = \sin t$)

Figure 4.10 corresponds to a frequency of the input signal $\omega_0 = 5$, such that the eigenvalues of the matrix $R_{w_m w_{mf}}(0)$ are complex: $(-0.0553 \pm j\,0.05076)g$. This explains the oscillatory behavior of the original and averaged systems observed in the figure, which did not exist in the previous examples of Section 4.3.

4.5.2 Output Error Scheme—Nonlinear Equations

We now return to the complete, nonlinear differential equations

$$\dot{e}(t) = A_m e(t) + b_m \phi^T(t) w_m(t) + b_m \phi^T(t) Q\, e(t)$$

$$\dot{\phi}(t) = -g\, w_m(t) c_m^T e(t) - g\, Q\, e(t) c_m^T e(t) \qquad (4.5.21)$$

From (4.4.45)

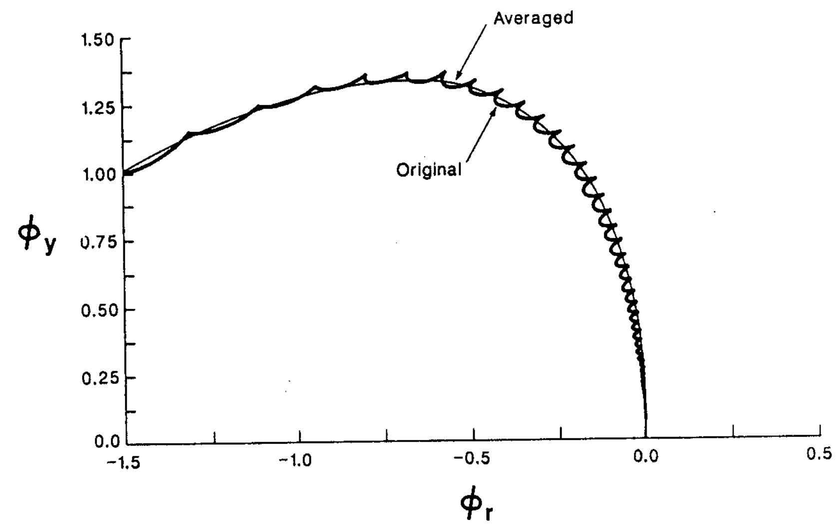

Figure 4.9: Parameter Error $\phi_y(\phi_r)$ ($r = \sin 3t$)

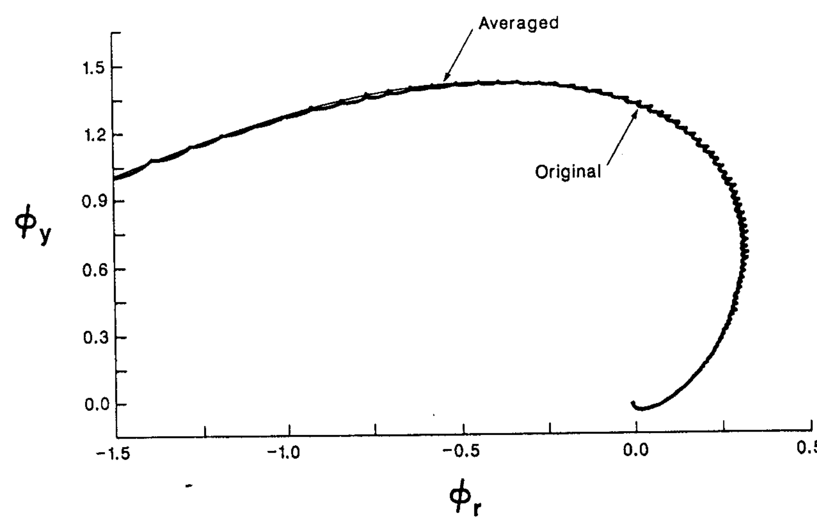

Figure 4.10: Parameter Error $\phi_y(\phi_r)$ ($r = \sin 5t$)

$$v(t, \phi) = \int_0^t e^{(A_m + b_m \phi^T Q)(t-\tau)} b_m \phi^T w_m(\tau) d\tau \tag{4.5.22}$$

so that the averaged system is

$$\dot{\phi}_{av} = g f_{av}(\phi_{av}) \qquad \phi_{av}(0) = \phi(0) \tag{4.5.23}$$

where f_{av} is defined by the limit

$$f_{av}(\phi) = -\lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} \Big[w_m(t) c_m^T v(t, \phi) + Q\, v(t, \phi) c_m^T v(t, \phi) \Big] dt \tag{4.5.24}$$

The assumptions of the theorems will be satisfied if the limit in (4.5.24) is uniform in the sense of (B3) and provided that the matrix $A_m + b_m \phi^T Q$ is uniformly exponentially stable for $\phi \in B_h$. This means that if the controller parameters are frozen at any point of the trajectory the resulting time invariant system must be closed-loop stable.

Frequency Domain Analysis

The expression of f_{av} in (4.5.24) can be translated into the frequency domain, noting that w_m is related to r through the vector transfer function $\hat{H}_{w_m r}$

$$f_{av}(\phi) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \Big[\hat{H}_{w_m r}(-j\omega) + Q(-j\omega I - A_m - b_m \phi^T Q)^{-1} b_m \phi^T \hat{H}_{w_m r}(-j\omega) \Big] \Big[c_m^T (j\omega I - A_m - b_m \phi^T Q)^{-1} b_m \phi^T \hat{H}_{w_m r}(j\omega) \Big] S_r(d\omega) \tag{4.5.25}$$

where $S_r(d\omega)$ is the spectral measure of r. Note that f_{av} can be factored as

$$f_{av}(\phi) = -A_{av}(\phi) \cdot \phi \tag{4.5.26}$$

where $A_{av}: \mathbb{R}^{2n} \to \mathbb{R}^{2n \times 2n}$ is similar to $R_{w_m w_{mf}}(0)$ in Section 4.5.1, but now depends nonlinearly on ϕ. The expression in (4.5.25) is more complex than in the linear case, but some manipulations will allow us to obtain a more interesting result.

Recall that (4.5.21) was obtained from the differential equation

$$\dot{e}(t) = A_m e(t) + b_m \phi^T(t) w(t)$$
$$\dot{\phi}(t) = -g\, w(t) c_m^T e(t) \tag{4.5.27}$$

by noting that $w(t) = w_m(t) + Q\, e(t)$. In general, (4.5.27) is of limited use, precisely because w depends on e. The signal w is not an external signal, but depends on internal variables. On the other hand, w_m is an exogeneous signal, related to r through a stable transfer function.

In the context of averaging, the differential equation describing the fast variable (i.e., e) is averaged, assuming that the slow variable (i.e., ϕ) is constant. However, when ϕ is constant, w *is* related to r through a linear time invariant system, with a transfer function depending on ϕ. If $\det(sI - A_m - b_m \phi^T Q)$ is Hurwitz (as we assume to apply averaging), this transfer function is stable. Therefore, *assuming that ϕ is fixed*, we can write

$$\hat{w} = \hat{H}_{wr}(s, \phi)\, \hat{r} \tag{4.5.28}$$

so that using (4.5.27), (4.5.25) can be replaced by (4.5.26), with an expression similar to the expression of $R_{w_m w_{mf}}(0)$ in (4.5.9), i.e.

$$A_{av}(\phi) = \frac{1}{2\pi c_0^*} \int_{-\infty}^{\infty} \hat{H}_{wr}^*(j\omega, \phi) \hat{H}_{wr}^T(j\omega, \phi) \hat{M}(j\omega) S_r(d\omega) \tag{4.5.29}$$

Explicit Expression for $\hat{H}_{wr}(s, \phi)$

Recall that $\overline{w}_m$ is related to r through the transfer function $\hat{H}_{\overline{w}_m r}$, whose poles are the zeros of $\det(sI - A_m)$. Let

$$\hat{\chi}_m(s) = \det(sI - A_m) \tag{4.5.30}$$

and write the transfer function $\hat{H}_{\overline{w}_m r}$ as the ratio of a vector polynomial $\hat{n}(s)$, and a characteristic polynomial $\hat{\chi}_m(s)$, i.e.,

$$\hat{H}_{\overline{w}_m r}(s) = \frac{\hat{n}(s)}{\hat{\chi}_m(s)} \tag{4.5.31}$$

We found in Section 3.5 (cf. (3.5.8), (3.5.11)) that

$$\overline{w} = \frac{\hat{n}(s)}{\hat{\chi}_m(s)} \left[r + \frac{1}{c_0^*} \phi^T w \right] \tag{4.5.32}$$

Denote $\phi_r = c_0 - c_0^*$, so that $\phi^T w = \phi_r r + \overline{\phi}^T \overline{w}$. Assuming that ϕ is constant, (4.5.32) becomes

$$\overline{w} = \left[\hat{\chi}_m(s) \cdot I - \frac{1}{c_0^*} \hat{n}(s) \overline{\phi}^T \right]^{-1} \hat{n}(s) \left[\left[1 + \frac{\phi_r}{c_0^*} \right] r \right]$$

$$= \frac{\hat{n}(s)}{\hat{\chi}_m(s) - \frac{1}{c_0^*} \overline{\phi}^T \hat{n}(s)} \left[\left[1 + \frac{\phi_r}{c_0^*} \right] r \right] \tag{4.5.33}$$

Denote

$$\hat{\chi}_\phi(s) := \hat{\chi}_m(s) - \frac{1}{c_0^*} \overline{\phi}^T \hat{n}(s) \tag{4.5.34}$$

$\hat{\chi}_\phi(s)$ is closed-loop characteristic polynomial, giving the poles of the adaptive system with feedback θ, that is, the poles of the model transfer function with feedback ϕ. Therefore, $\hat{\chi}_\phi(s)$ is also given by

$$\hat{\chi}_\phi(s) = \det(sI - A_m - b_m \phi^T Q) \tag{4.5.35}$$

With this notation, (4.5.33) can be written

$$\overline{w} = \frac{\hat{\chi}_m}{\hat{\chi}_\phi} \cdot \hat{H}_{\overline{w}_m r} \left[r + \frac{\phi_r}{c_0^*} r \right]$$

$$= \frac{\hat{\chi}_m}{\hat{\chi}_\phi} (\overline{w}_m) + \frac{\hat{\chi}_m}{\hat{\chi}_\phi} \left[\frac{\phi_r}{c_0^*} \overline{w}_m \right] \tag{4.5.36}$$

On the other hand

$$r = \frac{\hat{\chi}_m}{\hat{\chi}_\phi} \left[1 - \frac{\overline{\phi}^T \hat{n}}{c_0^* \hat{\chi}_m} \right] \cdot (r)$$

$$= \frac{\hat{\chi}_m}{\hat{\chi}_\phi} (r) - \frac{\hat{\chi}_m}{\hat{\chi}_\phi} \left[\frac{\overline{\phi}^T}{c_0^*} \cdot \overline{w}_m \right] \tag{4.5.37}$$

Define

$$B(\phi) := \begin{bmatrix} 0 & -\frac{1}{c_0^*} \overline{\phi}^T \\ 0 & \frac{\phi_r}{c_0^*} \cdot I \end{bmatrix} \in \begin{bmatrix} \mathbb{R}^{1 \times 1} & \mathbb{R}^{1 \times 2n-1} \\ \mathbb{R}^{2n-1 \times 1} & \mathbb{R}^{2n-1 \times 2n-1} \end{bmatrix} \tag{4.5.38}$$

i.e.,

$$B(\phi) \in \mathbb{R}^{2n \times 2n}$$

so that (4.5.36)–(4.5.37) can be written

$$w = \begin{bmatrix} r \\ \overline{w} \end{bmatrix} = \frac{\hat{\chi}_m}{\hat{\chi}_\phi} \begin{bmatrix} r \\ \overline{w}_m \end{bmatrix} + \frac{\hat{\chi}_m}{\hat{\chi}_\phi} B(\phi) \cdot \begin{bmatrix} r \\ \overline{w}_m \end{bmatrix} \tag{4.5.39}$$

The vector transfer function $\hat{H}_{wr}$ can therefore be expressed in terms of the vector transfer function $\hat{H}_{w_m r}$ by

$$\hat{H}_{wr}(s, \phi) = \frac{\hat{\chi}_m(s)}{\hat{\chi}_\phi(s)} (I + B(\phi)) \hat{H}_{w_m r}(s) \tag{4.5.40}$$

and, as expected

$$\hat{H}_{wr}(s, 0) = \hat{H}_{w_m r}(s) \tag{4.5.41}$$

Convergence Analysis

With (4.5.40), A_{av} can be written

$$A_{av}(\phi) = \frac{1}{2\pi c_0^*} \int_{-\infty}^{\infty} \left| \frac{\hat{\chi}_m(j\omega)}{\hat{\chi}_\phi(j\omega)} \right|^2 (I + B(\phi)) \hat{H}^*_{w_m r}(j\omega) \cdot \hat{H}^T_{w_m r}(j\omega)(I + B^T(\phi)) \hat{M}(j\omega)\, S_r(d\omega) \tag{4.5.42}$$

Consider now the trajectories of the averaged system and let $v(\phi_{av}) = |\phi_{av}|^2 = \phi_{av}^T \phi_{av}$. Note that by (4.5.38), it follows that

$$\phi^T \cdot B(\phi) = 0 \qquad \text{for all } \phi \tag{4.5.43}$$

Denote

$$R(\phi_{av}) := \frac{1}{2\pi c_0^*} \int_{-\infty}^{\infty} \left| \frac{\hat{\chi}_m(j\omega)}{\hat{\chi}_{\phi_{av}}(j\omega)} \right|^2 \hat{H}^*_{w_m r}(j\omega)$$

$$\cdot\ \hat{H}^T_{w_m r}(j\omega)\hat{M}(j\omega)S_r(d\omega) \qquad (4.5.44)$$

It follows that the derivative of v is given by

$$-\dot{v}(\phi_{av}) = g\,\phi_{av}^T(R(\phi_{av})+R^T(\phi_{av}))\,\phi_{av} \qquad (4.5.45)$$

which is identical to the expression for the linear case (4.5.12), provided that $R(\phi_{av})$ given in (4.5.44) replaces $R_{w_m w_{mf}}(0)$ given in (4.5.9). It is remarkable that this result differs from the expression obtained by linearization followed by averaging in Section 4.5.1 only by the *scalar* weighting factor $|\hat{\chi}_m/\hat{\chi}_\phi|^2$. Recall that $\hat{\chi}_m(s)$ defines the nominal closed-loop poles (i.e. when $\phi = 0$, while $\hat{\chi}_\phi$ defines the closed-loop poles with feedback gains $\theta = \phi + \theta$. The term $|\hat{\chi}_m/\hat{\chi}_\phi|^2$ is strictly positive, given any ϕ bounded, and it approaches unity continuously as ϕ approaches zero.

Since $\hat{M}(s)$ is strictly positive real, $R(\phi_{av})$ is at least positive semidefinite. As in the linearized case, it is positive definite if w_m is persistently exciting. Using the Lyapunov function $v(\phi_{av})$, this argument itself constitutes a proof of exponential stability of the averaged system, using (4.5.45). By theorem 4.4.3, the exponential stability of the original system is also guaranteed for g sufficiently small.

Rates of convergence can also be determined, using the Lyapunov function $v(\phi_{av})$, so that

$$\begin{aligned} -\dot{v} &= g\,\phi_{av}^T(R(\phi_{av})+R^T(\phi_{av})))\,\phi_{av} \\ &\geq g \inf_{\phi_{av}\in B_h}(\lambda_{\min}(R(\phi_{av})+R^T(\phi_{av})))\,v \;:=\; 2g\alpha v \end{aligned} \qquad (4.5.46)$$

and the guaranteed rate of parameter convergence of the averaged adaptive system is $g\alpha$. The rate of convergence of the original system can be estimated by the same value, for g sufficiently small.

It is interesting to note that, as $|\phi_{av}|$ increases, $\lambda_{\min}(R(\phi_{av})+R^T(\phi_{av})))$ tends to zero in some directions. This indicates that the adaptive control system may *not* be globally exponentially stable.

Example

We consider the previous two parameter example. The adaptive system is described by

$$\dot{e}_0(t) = -a_m e_0(t) + k_p(\phi_r(t)r(t) + \phi_y(t)e_0(t) + \phi_y(t)y_m(t))$$

$$\begin{aligned} \dot{\phi}_r(t) &= -g\,e_0(t)\,r(t) \\ \dot{\phi}_y(t) &= -g\,e_0^2(t) - g\,e_0(t)\,y_m(t) \end{aligned} \qquad (4.5.47)$$

Consider the case when $r = r_0\sin(\omega_0 t)$. The averaged system can be computed using (4.5.42). We can also verify the expression using (4.5.47) and the definition of the averaged system (4.5.22). After some manipulations, we obtain, for the averaged system (dropping the "av" subscripts for simplicity)

$$\dot{\phi}_r = -g\,k_p\frac{r_0^2}{2}\,\frac{1}{\omega_0^2+(a_m-k_p\phi_y)^2}\left[(a_m-k_p\phi_y)\phi_r + \left[\frac{a_m^2-\omega_0^2}{\omega_0^2+a_m^2}\,k_m\right]\phi_y - \frac{k_p a_m k_m}{\omega_0^2+a_m^2}\,\phi_y^2\right] \qquad (4.5.48)$$

$$\dot{\phi}_y = -g\,k_p\frac{r_0^2}{2}\,\frac{1}{\omega_0^2+(a_m-k_p\phi_y)^2}\left[k_m\phi_r + \frac{a_m k_m^2}{\omega_0^2+a_m^2}\,\phi_y + k_p\phi_r^2 + \frac{k_p a_m k_m}{\omega_0^2+a_m^2}\,\phi_r\phi_y\right] \qquad (4.5.49)$$

Using this result, or using (4.5.42)–(4.5.43), we find that for $v = \phi^T\phi$

$$-\dot{v} = 2g\left[\frac{\omega_0^2+a_m^2}{\omega_0^2+(a_m-k_p\phi_y)^2}\right]\frac{r_0^2}{2}\,\frac{k_p}{k_m}\cdot\phi^T\begin{bmatrix}\dfrac{a_m k_m}{\omega_0^2+a_m^2} & \dfrac{k_m^2(a_m^2-\omega_0^2)}{(\omega_0^2+a_m^2)^2} \\[2ex] \dfrac{k_m^2}{\omega_0^2+a_m^2} & \dfrac{a_m k_m^3}{(\omega_0^2+a_m^2)^2}\end{bmatrix}\phi \qquad (4.5.50)$$

It can easily be checked that when the first term in brackets is equal to 1 (i.e. with ϕ_y replaced by zero), the result is the same as the result obtained by first linearizing the system, then averaging it (cf. (4.5.20)). In fact, it can be seen, from the expressions of the averaged systems

((4.5.10) with (4.5.9), and (4.5.23) with (4.5.26), (4.5.38) and (4.5.42)), that the system obtained by linearization followed by averaging is *identical* to the system obtained by averaging followed by linearization. Also, given any prescribed B_h (but such that det $(sI - A_m - b_m \phi^T Q)$ is Hurwitz), (4.5.50) can be used to obtain estimates of the rates of convergence of the *nonlinear* system.

We reproduce here simulations for the following values of the parameters: $a_m = 3$, $k_m = 3$, $a_p = 1$, $k_p = 2$, $r_0 = 1$, $\omega_0 = 1$, $g = 1$. The first set of figures is a simulation for initial conditions $\phi_r(0) = -0.5$ and $\phi_y(0) = 0.5$. Figure 4.11 represents the time variation of the function $\ln(v = \phi^T \phi)$ for the original, averaged, and linearized-averaged systems (the minimum slope of the curve gives the rate of convergence).

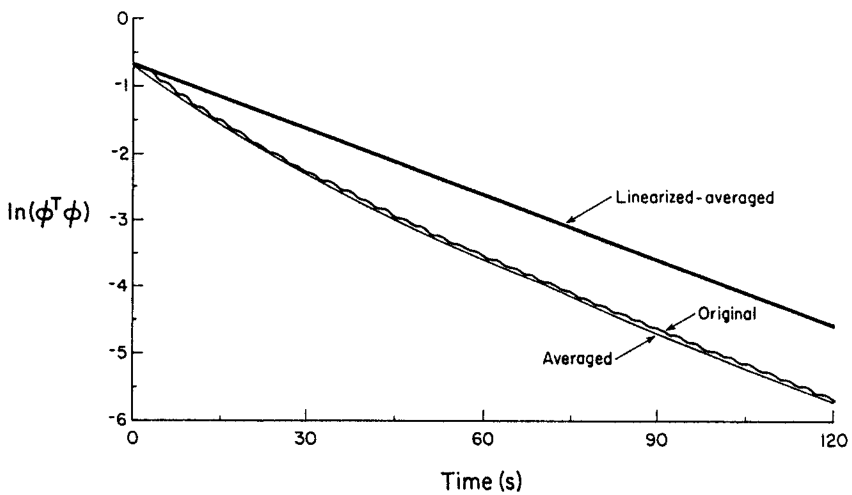

Figure 4.11: Logarithm of the Lyapunov Function

It shows the close approximation of the original system by the averaged system. The slope for the linearized-averaged system is asymptotically identical to that of the averaged system, since parameters eventually get arbitrarily close to their nominal values. Figures 4.12 and 4.13 show the approximation of the trajectories of ϕ_r and ϕ_y.

Figure 4.14 represents the logarithm of the Lyapunov function for a simulation with identical parameters, but initial conditions $\phi_r(0) = 0.5$, $\phi_y(0) = -0.5$. Due to the change of sign in $\phi_y(0)$, the rate of convergence of the nonlinear system is less now than the rate of the linearized

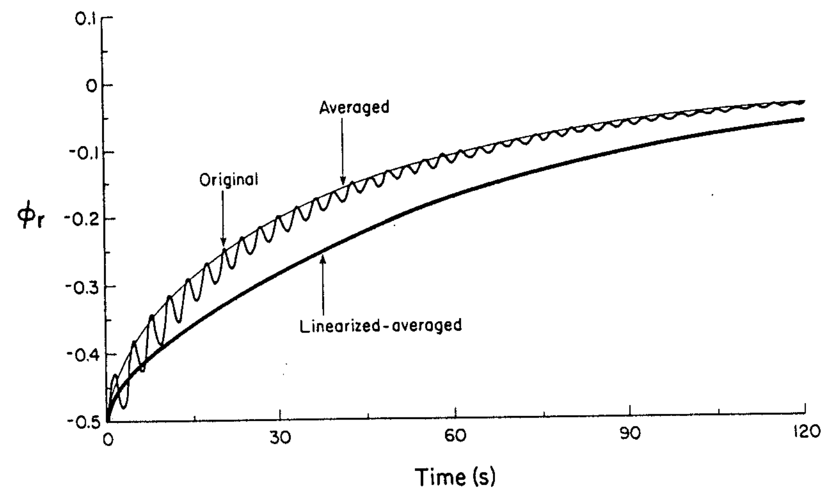

Figure 4.12: Parameter Error ϕ_r

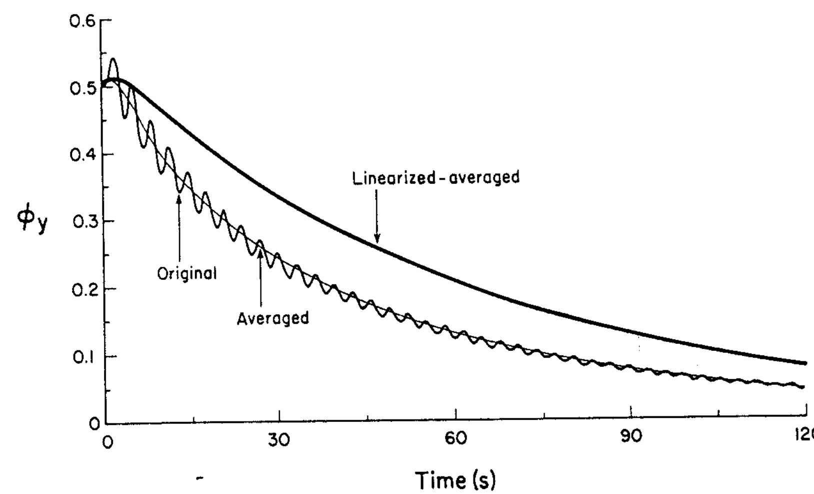

Figure 4.13: Parameter Error ϕ_y

system, while it was larger in the previous case.

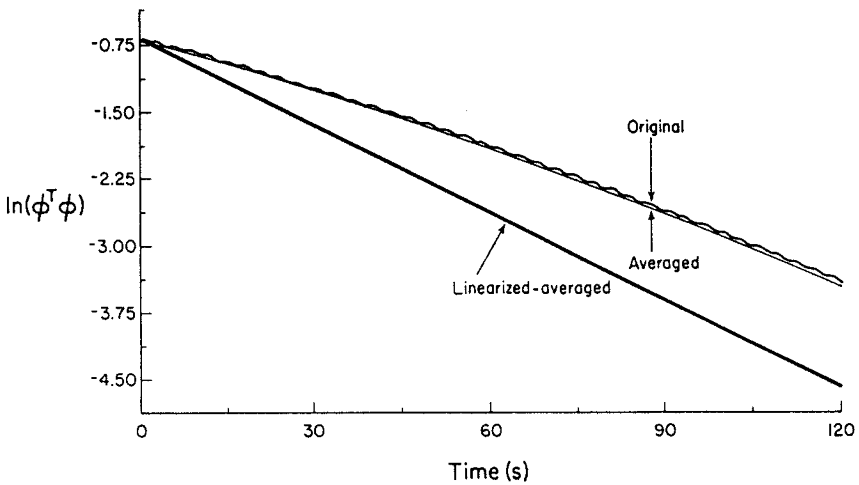

Figure 4.14: Logarithm of the Lyapunov Function

These simulations demonstrate the close approximation by the averaged system, and it should be noted that this is achieved despite an adaptation gain g equal to 1. This shows that the averaging method is useful for values of g which are not necessarily infinitesimal (i.e. not necessarily for very slow adaptation), but for values which are often practical ones.

Figure 4.15 shows the state-space trajectory $\phi_y(\phi_r)$, corresponding to Figure 4.10, that is with initial conditions $\phi_r(0) = -1.5, \phi_y(0) = 1$, and parameters as above except $\omega_0 = 5$. Figure 4.15 shows the distortion of the trajectories in the state-space, due to the nonlinearity of the differential system.

4.5.3 Input Error Scheme

An expression for the averaged system corresponding to the input error scheme may also be obtained. We consider the scheme for arbitrary relative degree. For simplicity, however, we neglect the normalization factor in the gradient algorithm and the projection for the feedforward gain c_0.

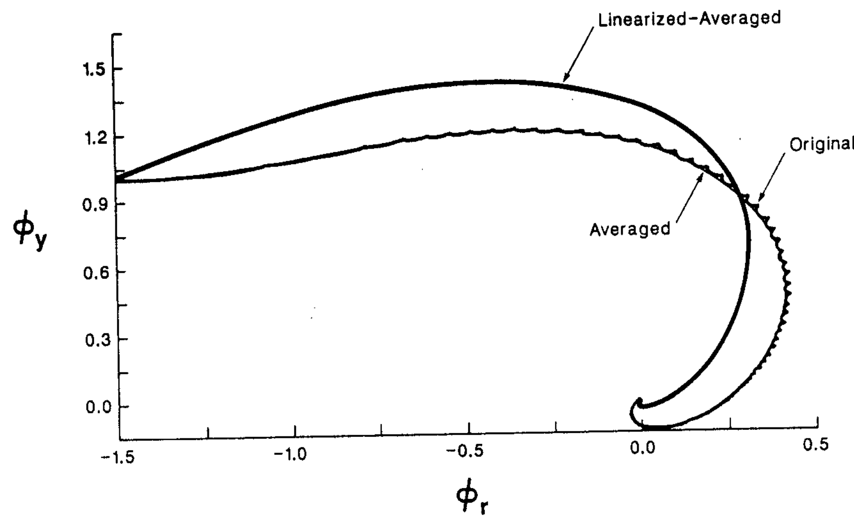

Figure 4.15: Parameter Error $\phi_y(\phi_r)(r = \sin 5t)$

The equation describing the parameter update is then simply

$$\dot{\phi}(t) = -g\, v(t) v^T(t) \phi(t) \tag{4.5.51}$$

so that the averaged system is again of the form

$$\dot{\phi}_{av} = -g\, A_{av}(\phi_{av}) \phi_{av} \tag{4.5.52}$$

where $A_{av}(\phi_{av})$ is the autocovariance of the vector v at $t = 0$. It depends on ϕ_{av} because v is obtained from a closed-loop system with feedback depending on the parameter error ϕ_{av}. Within the framework of averaging, the system is a linear time invariant system, so that we may write (as in (4.5.28))

$$\hat{v} = \hat{H}_{vr}(s, \phi) \cdot \hat{r} \tag{4.5.53}$$

Explicit Expression of $\hat{H}_{vr}(s, \phi)$

Recall that (cf. (3.5.10))

$$v = \hat{L}^{-1}(z) = \hat{L}^{-1}\begin{bmatrix} r_p \\ \overline{w} \end{bmatrix} = \hat{L}^{-1}\begin{bmatrix} r + \frac{1}{c_0^*}\phi^T w \\ \overline{w} \end{bmatrix} \qquad (4.5.54)$$

Since the controller is the same as for the output error scheme, we may use (4.5.36)–(4.5.37). First, rewrite (4.5.36) using $\phi_r = c_0 - c_0^*$ so that

$$\overline{w} = \frac{c_0}{c_0^*}\frac{\hat{\chi}_m}{\hat{\chi}_\phi}(\overline{w}_m) \qquad (4.5.55)$$

and

$$r + \frac{1}{c_0^*}\phi^T w = \frac{c_0}{c_0^*} r + \frac{1}{c_0^*}\overline{\phi}^T \overline{w}$$

$$= \frac{c_0}{c_0^*} r + \frac{c_0}{c_0^*}\frac{\hat{\chi}_m}{\hat{\chi}_\phi}\left[\frac{\overline{\phi}^T}{c_0^*}\overline{w}_m\right] \qquad (4.5.56)$$

With (4.5.37), (4.5.56) simply becomes

$$r + \frac{1}{c_0^*}\phi^T w = \frac{c_0}{c_0^*}\frac{\hat{\chi}_m}{\hat{\chi}_\phi}(r) \qquad (4.5.57)$$

Therefore

$$v = \frac{c_0}{c_0^*}\hat{L}^{-1}\frac{\hat{\chi}_m}{\hat{\chi}_\phi}\begin{bmatrix} r \\ \overline{w}_m \end{bmatrix} = \frac{c_0}{c_0^*}\hat{L}^{-1}\frac{\hat{\chi}_m}{\hat{\chi}_\phi}(w_m) \qquad (4.5.58)$$

and

$$\hat{H}_{vr}(s,\phi) = \frac{c_0}{c_0^*}\hat{L}^{-1}(s)\frac{\hat{\chi}_m(s)}{\hat{\chi}_\phi(s)}\hat{H}_{w_m r}(s) \qquad (4.5.59)$$

which is the equivalent of (4.5.40) for the input error scheme.

Convergence Analysis

Using the foregoing result, we may express $A_{av}(\phi)$ in the frequency domain as

$$A_{av}(\phi) = \frac{1}{2\pi}\left[\frac{c_0}{c_0^*}\right]^2 \int_{-\infty}^{\infty} |\hat{L}^{-1}(j\omega)|^2 \left|\frac{\hat{\chi}_m(j\omega)}{\hat{\chi}_\phi(j\omega)}\right|^2 \hat{H}^*_{w_m r}(j\omega)$$

$$\cdot\ \hat{H}^T_{w_m r}(j\omega)\ S_r(d\omega) \qquad (4.5.60)$$

Note that the matrix $A_{av}(\phi)$ is now symmetric and is a positive semidefinite matrix for all ϕ. It is positive definite if the input r is sufficiently rich. Again, parameter convergence rates may be estimated from the preceding expression. Although the convergence properties are quite similar, the symmetry of $A_{av}(\phi)$ guarantees that around the equilibrium, the linearized system is described by a linear time invariant system with only real eigenvalues. Therefore, the oscillatory behavior of the output error scheme is not observed for the input error scheme.

Example

We consider once again the example of Section 4.5.2, but for the input error scheme. The model transfer function is $\hat{M} = k_m/(s+a_m)$, and we choose $\hat{L} = (s+l_2)/l_1$. Note that $(\hat{M}\hat{L})^{-1}$ may be expressed as

$$(\hat{M}\hat{L})^{-1} = \frac{(s+a_m)l_1}{k_m(s+l_2)} = \frac{l_1}{k_m} + \frac{a_m - l_2}{k_m}\frac{l_1}{s+l_2} \qquad (4.5.61)$$

The equations describing the overall adaptive system with the input error scheme are

$$\dot{y}_p = -a_p y_p + k_p u$$

$$\dot{y}_m = -a_m y_m + k_m r$$

$$u = c_0 r + d_0 y_p$$

$$\dot{x}_1 = -l_2 x_1 + l_1 u \qquad \text{i.e.} \quad x_1 = \hat{L}^{-1}(u)$$

$$\dot{x}_2 = -l_2 x_2 + l_1 y_p \qquad \text{i.e.} \quad x_2 = \hat{L}^{-1}(y_p)$$

$$x_3 = \frac{l_1}{k_m}(y_p) + \frac{a_m - l_2}{k_m}(x_2) \qquad \text{i.e } x_3 = (\hat{M}\hat{L})^{-1}(y_p)$$

$$e_2 = c_0 x_3 + d_0 x_2 - x_1$$

$$\dot{c}_0 = -g e_2 x_3 \qquad \phi_r = c_0 - c_0^* = c_0 - k_m/k_p$$

$$\dot{d}_0 = -g e_2 x_2 \qquad \phi_y = d_0 - d_0^* = d_0 - \frac{(a_p - a_m)}{k_p}$$

Again, we neglected the normalization factor and the projection in the update law for simplicity.

When $r = r_0 \sin(\omega_0 t)$, the averaged system is

$$\begin{bmatrix}\dot{\phi}_r \\ \dot{\phi}_y\end{bmatrix} = -g\,\frac{r_0^2}{2}\,\frac{l_1^2}{\omega_0^2 + l_2^2}\left[\frac{\phi_r}{c_0^*} + 1\right]^2 \frac{\omega_0^2 + a_m^2}{\omega_0^2 + (a_m - k_p\phi_y)^2}$$

$$\begin{bmatrix} 1 & \dfrac{k_m a_r}{a_m^2 + \omega_0^2} \\ \dfrac{k_m a_m}{a_m^2 + \omega_0^2} & \dfrac{k_m^2}{a_m^2 + \omega_0^2}\end{bmatrix}\begin{bmatrix}\phi_r \\ \phi_y\end{bmatrix} \qquad (4.5.62)$$

When $\omega_0 = 5$, Figure 4.16 shows that trajectories of the output scheme exhibit an oscillatory type of response. Figure 4.17 shows the response for the input error scheme under comparable conditions.

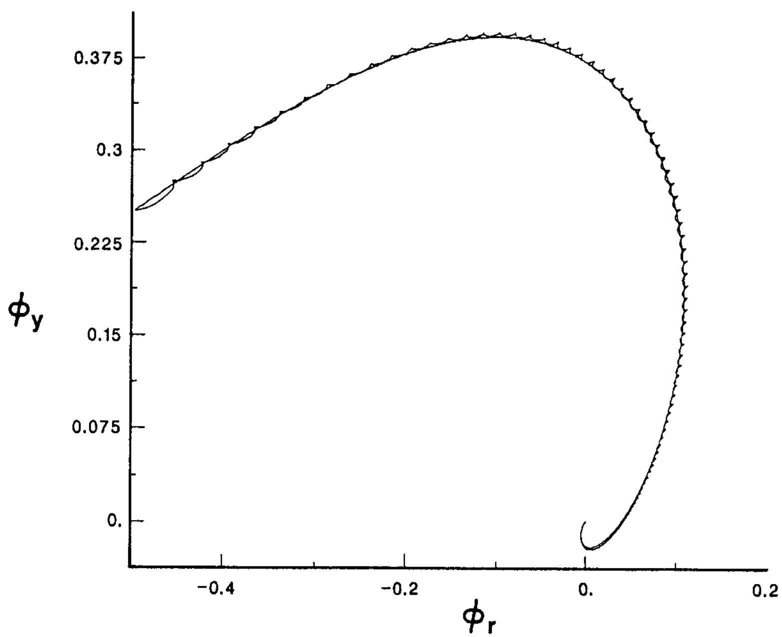

Figure 4.16: Parameter Error $\phi_y(\phi_r)$–Output Error Scheme

The parameters are $k_m = 3$, $a_m = 3$, $a_p = 1$, $k_p = 2$, $r_0 = 1$, $\omega_0 = 5$, $g = 1$, $l_1 = 10.05$, $l_2 = 10$, $\phi_r(0) = -0.5$, $\phi_y(0) = 0.25$. As may be observed, the

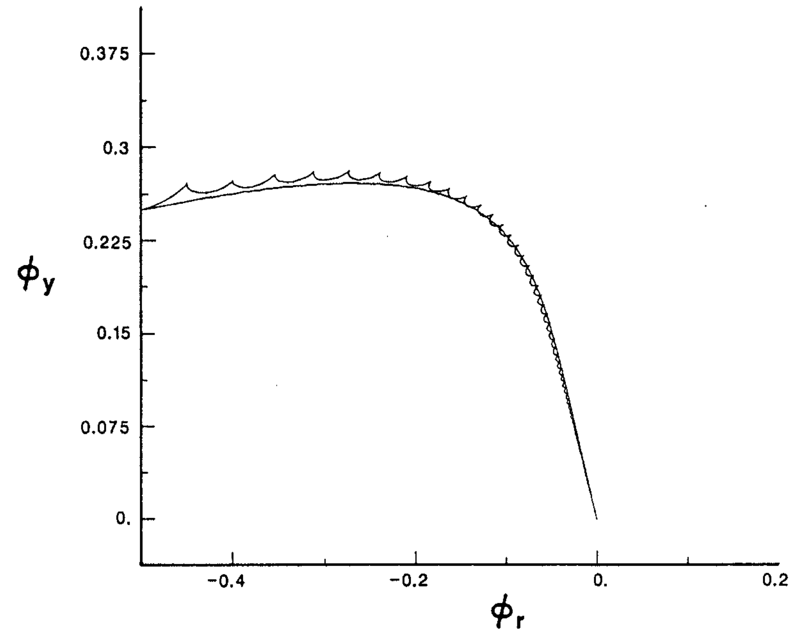

Figure 4.17: Parameter Error $\phi_y(\phi_r)$–Input Error Scheme

trajectories do not exhibit oscillatory behavior, reflecting the fact that the matrix above is symmetric and, therefore, has only real eigenvalues.

4.6 CONCLUSIONS

Averaging is a powerful tool to approximate nonautonomous differential equations by autonomous differential equations. In this chapter, we introduced averaging as a method of analysis of adaptive systems. The approximation of parameter convergence rates using averaging was justified by general results concerning a class of systems including the adaptive systems described in Chapters 2 and 3. The analysis had the interesting feature of considering nonlinear differential equations as well as linear ones. Therefore, the application was not restricted to linear or linearized systems, but extended to all adaptive systems considered in this work, including adaptive control systems.

The application to adaptive systems included useful parameter convergence rates estimates for identification and adaptive control systems. The rates depended strongly on the reference input and a frequency domain analysis related the frequency content of the reference input to the convergence rates, even in the nonlinear adaptive control case. These results are useful for the optimum design of reference input. They have the limitation of depending on unknown plant parameters, but an approximation of the complete parameter trajectory is obtained and the understanding of the dynamical behavior of the parameter error is considerably increased using averaging. For example, it was found that the trajectory of the parameter error corresponding to the linear error equation could be approximated by an LTI system with real negative eigenvalues, while for the strictly positive real (SPR) error equation it had possibly complex eigenvalues.

Besides requiring stationarity of input signals, averaging also required slow parameter adaptation. We showed however, through simulations, that the approximation by the averaged system was good for values of the adaptation gain that were close to 1 (that is, not necessarily infinitesimal) and for acceptable time constants in the parameter variations. In fact, it appeared that a basic condition is simply that parameters vary more slowly than do other states and signals of the adaptive system.

CHAPTER 5
ROBUSTNESS

5.1 STRUCTURED AND UNSTRUCTURED UNCERTAINTY

In a large number of control system design problems, the designer does not have a detailed state-space model of the plant to be controlled, either because it is too complex, or because its dynamics are not completely understood. Even if a detailed high-order model of the plant is available, it is usually desirable to obtain a reduced order controller, so that part of the plant dynamics must be neglected. We begin discussing the representation of such uncertainties in plant models, in a framework similar to Doyle & Stein [1981].

Consider the kind of prior information available to control a *stable* plant, and obtained for example by performing input-output experiments, such as sinusoidal inputs. Typically, Bode diagrams of the form shown in Figures 5.1 and 5.2 are obtained. An inspection of the diagrams shows that the data obtained beyond a certain frequency ω_H is unreliable because the measurements are poor, corrupted by noise, and so on. They may also correspond to the high-order dynamics that one wishes to neglect. What is available, then, is essentially no phase information, and only an "envelope" of the magnitude response beyond ω_H. The dashed lines in the magnitude and phase response correspond to the approximation of the plant by a finite order model, assuming that there are no dynamics at frequencies beyond ω_H. For frequencies below ω_H, it is easy to guess the presence of a zero near ω_1, poles in the neighborhood of ω_2, ω_3, and complex pole pairs in the neighborhood of ω_4, ω_5.